



## Εθνικό Μετσόβιο Πολυτεχνείο

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

ΕΡΓΑΣΤΗΡΙΟ ΕΙΚΟΝΩΝ, BINTEO ΚΑΙ ΠΟΛΥΜΕΣΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

Προσαρμοζόμενη Πρόσβαση σε Πολυμεσικό  
Περιεχόμενο με βάση το Εννοιολογικό Πλαίσιο και  
το Προφίλ του Χρήστη

Διδακτορική Διατριβή

του

**ΦΟΙΒΟΥ-ΑΠΟΣΤΟΛΟΥ Ι. ΜΥΛΩΝΑ**

Διπλωματούχου Ηλεκτρολόγου Μηχανικού &  
Μηχανικού Υπολογιστών Ε.Μ.Π. (2001)

Αθήνα, Απρίλιος 2008





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ & ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ & ΥΠΟΛΟΓΙΣΤΩΝ  
ΕΡΓΑΣΤΗΡΙΟ ΕΙΚΟΝΩΝ, ΒΙΝΤΕΟ & ΠΟΛΥΜΕΣΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

Προσαρμοζόμενη Πρόσβαση σε Πολυμεσικό Περιεχόμενο με βάση το Εννοιολογικό Πλαίσιο και το Προφίλ του Χρήστη

## Adaptive Multimedia Content Access based on Context and User Profiles

# Διδακτορική Διατριβή

# ΦΟΙΒΟΥ-ΑΠΟΣΤΟΛΟΥ Ι. ΜΥΛΩΝΑ

## Διπλωματούχου Ηλεκτρολόγου Μηχανικού & Μηχανικού Υπολογιστών Ε.Μ.Π. (2001)

Συμβουλευτική Επιτροπή: Στέφανος Κόλλιας  
Ανδρέας-Γεώργιος Σταφυλοπάτης  
Παναγιώτης Τσανάκας

Εγκρίθηκε από την επαυξελή εξεταστική επιτροπή την 22η Απριλίου 2008.

Σ. Κόλλιας                  Α.-Γ. Σταφυλοπάτης          Π. Τσανάκας  
Καθηγητής Ε.Μ.Π.          Καθηγητής Ε.Μ.Π.          Καθηγητής Ε.Μ.Π.

Τ. Σελλής Ι. Βασιλείου Ν. Παπασπύρου  
Καθηγητής Ε.Μ.Π. Καθηγητής Ε.Μ.Π. Επ. Καθηγητής Ε.Μ.Π.

Α. Ντελόπουλος  
Επ. Καθηγητής Α.Π.Θ.

Αθήνα, Απρίλιος 2008

.....

**ΦΟΙΒΟΣ-ΑΠΟΣΤΟΛΟΣ Ι. ΜΥΛΩΝΑΣ**  
Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Φοίβος-Απόστολος Ι. Μυλωνάς, 2008.  
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

στην οικογένειά μου



# Περιεχόμενα

Περιεχόμενα	iii
Κατάλογος Σχημάτων	ix
Κατάλογος Πινάκων	xiii
Πρόλογος	xvii
Ευχαριστίες	xxi
Περίληψη	xxiii
Περίληψη στα Αγγλικά	xxvii
Συντμήσεις και Αρκτικόλεξα	xxix
Κατάλογος Απόδοσης 'Ορων - Ελληνικά	xxxi
Κατάλογος Απόδοσης 'Ορων - Αγγλικά	xxxiii
<b>1 Εισαγωγή</b>	<b>1</b>
1.1 Γενικό Ερευνητικό Αντικείμενο .....	2
1.2 Γενικά Ερευνητικά Προβλήματα .....	2
1.3 Επιμέρους Αντικείμενο .....	4
1.4 Ερευνητική Ενασχόληση .....	5
1.4.1 Θεματική κατηγοριοποίηση εγγράφων .....	5
1.4.1.1 Επισκόπηση ερευνητικής περιοχής .....	5
1.4.1.2 Ερευνητική συνεισφορά .....	6
1.4.2 Θέματα προσωποποίησης .....	7
1.4.2.1 Επισκόπηση ερευνητικής περιοχής .....	7
1.4.2.2 Ερευνητική συνεισφορά .....	8
1.4.3 Θέματα οπτικού εννοιολογικού πλαισίου .....	9
1.4.3.1 Επισκόπηση ερευνητικής περιοχής .....	9
1.4.3.2 Ερευνητική συνεισφορά .....	10
1.4.4 Θέματα εννοιολογικού πλαισίου “ενδιάμεσου επιπέδου” .....	11
1.4.4.1 Επισκόπηση ερευνητικής περιοχής .....	11
1.4.4.2 Ερευνητική συνεισφορά .....	12

<b>2 Ερευνητικό πλαίσιο</b>	<b>15</b>
2.1 Εισαγωγή .....	15
2.2 Μηχανική Μάθηση / Ανάλυση και Εξόρυξη Δεδομένων .....	16
2.2.1 Μηχανική μάθηση .....	16
2.2.2 Εξόρυξη δεδομένων .....	17
2.2.3 Μη επιβλεπόμενη μάθηση .....	19
2.2.4 Μπεϋζιανή εξαγωγή συμπερασμάτων .....	19
2.2.5 Ταξινομητές .....	20
2.2.6 Μπεϋζιανός ταξινομητής .....	20
2.2.7 Κατηγοριοποίηση εγγράφων κειμένου .....	21
2.2.8 Μπεϋζιανά δίκτυα .....	23
2.3 Ανάλυση Πολυμεσικού Περιεχομένου Υποστηριζόμενη από τη Γνώση και το Εννοιολογικό Πλαίσιο .....	23
2.3.1 Ανάλυση πολυμεσικού περιεχομένου - μια περιληπτική σύνοψη .....	24
2.3.2 Αρχιτεκτονική ενός συστήματος ανάλυσης πολυμεσικού περιεχομένου .....	25
2.3.3 Ο ρόλος του εννοιολογικού πλαισίου .....	27
2.4 Πρόσβαση και Ανάκτηση Πολυμεσικής Πληροφορίας .....	29
2.4.1 Οι διαδικασίες ενός IR συστήματος .....	31
2.4.2 Μοντέλα IR συστημάτων .....	32
2.4.3 Μετρικές IR συστημάτων .....	35
2.4.4 Προβλήματα των IR συστημάτων .....	36
2.4.5 Αναζήτηση σε πολυμεσική πληροφορία .....	37
2.5 Προσωποποίηση Πολυμεσικού Περιεχομένου .....	38
2.5.1 Εξαγωγή προφίλ χρηστών .....	39
2.5.2 Συσχετιστική ανατροφοδότηση .....	39
2.5.3 Εξαγωγή προφίλ χρηστών με βάση τα πρότυπα MPEG .....	41
2.5.3.1 Εξαγωγή προφίλ χρηστών κατά MPEG-7 .....	41
2.5.3.2 Εξαγωγή προφίλ χρηστών κατά MPEG-21 .....	42
2.5.4 Κατάταξη προσωποποιημένων αποτελεσμάτων .....	43
<b>3 Μαθηματική σημειολογία, όροι και έννοιες</b>	<b>45</b>
3.1 Εισαγωγή .....	45
3.2 Ασαφή Σύνολα και Σχέσεις .....	46
3.3 Μεταβατικό Κλείσιμο .....	49
3.4 Οντολογίες .....	52
3.4.1 Εισαγωγή .....	52
3.4.2 Οντολογίες και αναπαράσταση γνώσης .....	55
3.4.3 Γλώσσες αναπαράστασης γνώσης: RDF/RDFS/OWL .....	56
3.4.4 Υποδομή οντολογιών - εργαλεία και υποστήριξη .....	58
3.5 Σημασιολογικές Σχέσεις .....	59
3.6 Σημασιολογική Περιγραφή Περιεχομένου .....	60
3.6.1 Περιγραφικά σχήματα .....	61
3.6.2 Αναπαράσταση περιγραφικών σχημάτων .....	62

<b>4 Εξόρυξη δεδομένων και κατηγοριοποίηση</b>	<b>65</b>
4.1 Εισαγωγή .....	65
4.1.1 Υπάρχουσα κατάσταση - Επισκόπηση ερευνητικής περιοχής .....	65
4.1.2 Δομή κεφαλαίου .....	69
4.2 Γενικές Αρχές .....	70
4.2.1 Συσταδοποίηση δεδομένων .....	71
4.2.1.1 Συσσωρευτική ιεραρχική συσταδοποίηση .....	73
4.2.2 Μείωση των διαστάσεων των χαρακτηριστικών γνωρισμάτων ...	76
4.2.3 Βελτιστοποίηση τεχνικών συσταδοποίησης μέσω κατηγοριοποίησης .....	77
4.3 Ασαφής Ταξινομική Συσταδοποίηση .....	80
4.3.1 Ασαφής ιεραρχική συσταδοποίηση .....	80
4.3.2 Ανάλυση συστάδων .....	81
4.3.3 Δημιουργία γνώσης .....	82
4.3.4 Ταξινομία εννοιολογικού πλαισίου .....	83
4.3.5 Ασαφής ιεραρχική συσταδοποίηση .....	84
4.4 Θεματική Κατηγοριοποίηση Εγγράφων .....	85
4.4.1 Ανάλυση εγγράφων .....	87
4.4.2 Επέκταση γνώσης .....	88
4.4.3 Ταξινομικό εννοιολογικό πλαισίο .....	89
4.4.4 Ερμηνεία του χαρακτηρισμού ενός πολυμεσικού εγγράφου .....	91
4.4.5 Σημασιολογικό ευρετήριο .....	93
4.4.6 Ασαφής ιεραρχική συσταδοποίηση εγγράφων .....	94
4.4.7 Αξιοποίηση υπηρεσιών ασαφούς εξαγωγής συμπερασμάτων .....	98
4.4.7.1 Αναπαράσταση γνώσης και συμβολισμοί .....	98
4.4.7.2 Δημιουργία σημασιολογικού ευρετηρίου και κατηγοριοποίησης κειμένου .....	101
4.4.7.3 Η ασαφής βάση γνώσης .....	102
4.5 Πειραματικά Αποτελέσματα .....	104
4.5.1 Θεματική κατηγοριοποίηση συνθετικών εγγράφων .....	104
4.5.2 Θεματική κατηγοριοποίηση πραγματικών εγγράφων .....	109
4.5.3 Θεματική κατηγοριοποίηση εγγράφων με ταυτόχρονη εξαγωγή συμπερασμάτων .....	115
<b>5 Προσωποποιημένη πρόσβαση</b>	<b>119</b>
5.1 Εισαγωγή .....	119
5.1.1 Υπάρχουσα κατάσταση - σχετικές εργασίες .....	120
5.1.2 Δομή κεφαλαίου .....	122
5.2 Εννοιολογικό Πλαίσιο και Προσωποποίηση .....	123
5.3 Αναπαράσταση Γνώσης .....	127
5.4 Αναπαράσταση Ασαφούς Εννοιολογικού Πλαισίου .....	129
5.4.1 Ασαφείς σημασιολογικές σχέσεις κατά την προσωποποίηση .....	131
5.4.2 Ταξινομικό εννοιολογικό πλαίσιο .....	132
5.4.3 Εννοιολογικό πλαίσιο κατά το χρόνο εκτέλεσης .....	134
5.5 Εξαγωγή Προφίλ Χρηστών .....	137
5.5.1 Ο ρόλος των προφίλ .....	137
5.5.2 Ενέργειες χρηστών .....	138
5.5.3 Ιστορικό χρήσης .....	139

5.5.4	Από τα έγγραφα στις προτυμήσεις χρηστών .....	140
5.5.4.1	Μοντελοποίηση προσέγγισης .....	140
5.5.4.2	Κλασική συσταδοποίηση .....	142
5.5.4.3	Ασαφοποίηση συστάδων .....	142
5.5.4.4	Εξαγωγή ασαφών προτυμήσεων .....	144
5.5.5	Εξαγωγή προφίλ χρηστών και ηλεκτρονική μάθηση .....	145
5.5.5.1	Στατική εξαγωγή προφίλ χρηστών .....	146
5.5.5.2	Δυναμική εξαγωγή προφίλ χρηστών .....	150
5.6	Ανάκτηση και Κατάταξη Προσωποποιημένης Πληροφορίας .....	156
5.6.1	Επίδραση εννοιολογικού πλαισίου .....	156
5.6.2	Κατάταξη προσωποποιημένων αποτελεσμάτων .....	159
5.7	Πειραματικές Μετρήσεις .....	161
5.8	Σενάρια Χρήσης .....	169
5.8.1	Πρώτο σενάριο χρήσης .....	169
5.8.2	Δεύτερο σενάριο χρήσης .....	174
<b>6</b>	<b>Εννοιολογικό πλαίσιο και οπτική πληροφορία</b>	<b>179</b>
6.1	Εισαγωγή .....	179
6.1.1	Υπάρχουσα κατάσταση - Ορισμός προβλήματος .....	179
6.1.2	Δομή κεφαλαίου .....	186
6.2	Εννοιολογική Αναπαράσταση Γνώσης .....	187
6.3	Ανάλυση Εικόνων με Χρήση του Εννοιολογικού Πλαισίου .....	191
6.3.1	Σχετικότητα εννοιολογικού πλαισίου .....	192
6.3.2	Μεθοδολογία και αλγόριθμος ανάλυσης .....	193
6.3.3	Μεθοδολογία και αλγόριθμος στην περίπτωση πολλαπλών θεματικών περιοχών .....	194
6.3.3.1	Ασαφής εννοιολογική αναπαράσταση γνώσης .....	195
6.3.3.2	Επεκτεταμένο μοντέλο γνώσης .....	196
6.3.3.3	Τροποποίηση αλγορίθμου ανάλυσης .....	199
6.3.3.4	Πειραματικά αποτελέσματα .....	200
6.4	Οπτικό Εννοιολογικό Πλαίσιο και Σημασιολογική Κατάτμηση Εικόνων	202
6.4.1	Αρχικοποίηση σημασιολογικής κατάτμησης .....	202
6.4.2	Επίδραση εννοιολογικού πλαισίου .....	205
6.4.3	Διαδικασία εκτέλεσης και αποτίμησης πειραματικών μετρήσεων	206
6.4.4	Πειραματικές μετρήσεις .....	208
6.4.5	Συνολικά αποτελέσματα αξιολόγησης .....	210
6.5	Οπτικό Εννοιολογικό Πλαίσιο και Γενετικός Αλγόριθμος Σημασιολογικού Χαρακτηρισμού Περιοχών .....	211
6.5.1	Βοηθούμενη από τη γνώση ανάλυση εικόνων .....	212
6.5.2	Κατάτμηση, εξαγωγή χαρακτηριστικών γνωρισμάτων και αρχική παραγωγή υποθέσεων .....	213
6.5.3	Υποδομή γνώσης .....	213
6.5.4	Βελτίωση εννοιολογικού πλαισίου .....	214
6.5.5	Πειραματικά αποτελέσματα .....	215
<b>7</b>	<b>Εννοιολογικό πλαίσιο ταξινόμησης</b>	<b>221</b>
7.1	Εισαγωγή .....	221
7.1.1	Υπάρχουσα κατάσταση - Ορισμός προβλήματος .....	223
7.1.2	Δομή κεφαλαίου .....	228

7.2	Ο Ρόλος του Εννοιολογικού Πλαισίου στην Ανίχνευση Εννοιών και στην Ταξινόμηση Εικόνων .....	229
7.3	Ανάλυση Τύπων Περιοχής .....	231
7.3.1	Ανίχνευση εννοιών με χρήση οπτικού θησαυρού .....	234
7.3.2	Κατασκευή οπτικού θησαυρού .....	236
7.3.3	Διανύσματα αναπαράστασης .....	239
7.3.4	Ταξινόμηση σκηνών .....	240
7.4	Εννοιολογικό Πλαίσιο Ταξινόμησης .....	242
7.4.1	Εννοιολογικό πλαίσιο ταξινόμησης εννοιών .....	243
7.4.1.1	Επίδραση βελτιστοποίησης εννοιολογικού πλαισίου ..	245
7.4.2	Εννοιολογικό πλαίσιο τύπων περιοχής .....	246
7.4.2.1	Επίδραση οπτικού εννοιολογικού πλαισίου .....	253
7.4.3	Μεικτό εννοιολογικό πλαίσιο .....	256
7.4.3.1	Μεικτή επίδραση εννοιολογικού πλαισίου .....	261
7.5	Πειραματικά Αποτελέσματα .....	264
7.5.1	Εννοιολογική προσέγγιση .....	265
7.5.2	Προσέγγιση τύπων περιοχής .....	267
7.5.3	Μεικτή προσέγγιση τύπων περιοχής και εννοιών .....	269
7.5.4	Ταξινόμηση σκηνών .....	274
<b>8</b>	<b>Συμπεράσματα και προτεινόμενη περαιτέρω έρευνα</b>	<b>277</b>
8.1	Προτεινόμενη Περαιτέρω Έρευνα .....	277
8.1.1	Μελλοντικές ερευνητικές επεκτάσεις .....	277
8.1.2	Μελλοντικές εφαρμογές .....	278
8.2	Συμπεράσματα .....	280
<b>9</b>	<b>Κατάλογος δημοσιεύσεων</b>	<b>281</b>
9.1	Περιοδικά .....	281
9.1.1	Δημοσιευμένα .....	281
9.1.2	Υπό δημοσίευση .....	281
9.1.3	Υποβεβλημένα προς χρήση .....	282
9.2	Κεφάλαια Βιβλίων .....	282
9.2.1	Δημοσιευμένα .....	282
9.2.2	Υπό δημοσίευση .....	283
9.2.3	Σε ετοιμασία .....	283
9.3	Συνέδρια .....	283
9.3.1	Δημοσιευμένα .....	283
9.3.2	Υπό δημοσίευση .....	286
	<b>Βιβλιογραφία</b>	<b>289</b>
	<b>Βιογραφικό Σημείωμα</b>	<b>313</b>
	<b>Παράρτημα Α</b>	<b>319</b>
	<b>Παράρτημα Β</b>	<b>321</b>
	<b>Ευρετήριο</b>	<b>323</b>



# Κατάλογος Σχημάτων

2.1 Δομικό διάγραμμα ανάλυσης πολυμεσικού περιεχομένου, υποβοηθούμενης από κάποιας μορφής γνώση.	25
2.2 Δομή ενός συστήματος IR	30
2.3 Κατηγοριοποίηση των μοντέλων IR	32
2.4 Ομαδοποίηση χρηστών	40
3.1 Παράδειγμα επαυξητικού (incremental) μεταβατικού κλεισίματος.	50
3.2 Μια οντολογική ταξινομία.	53
3.3 Ασαφής αναπαράσταση σχέσης σε OWL: OWL reification.	58
3.4 Εργαλεία του προτύπου MPEG-7 για την περιγραφή των εννοιολογικών πτυχών του πολυμεσικού περιεχομένου.	60
4.1 Δεδομένα προς συσταδοποίηση	75
4.2 Παράδειγμα ιεραρχικής συσσωρευτικής συσταδοποίησης δεδομένων - απεικόνιση όλων των βημάτων	75
4.3 Οι οντότητες που ευρετηριάζουν ένα έγγραφο (1,2,..,6) και οι σχέσεις μεταξύ τους.	87
4.4 Αλφάριθμος εννοιών της ασαφούς βάσης γνώσης.	103
4.5 Εργαλείο Κατηγοριοποίησης Εγγράφων	105
4.6 Παράδειγμα κατασκευής της σχέσης $T$ .	107
5.1 Σημασιολογική σύνδεση μεταξύ χρηστών και πολυμεσικού περιεχομένου μέσω των εννοιών μιας βάσης γνώσης.	124
5.2 Παράδειγμα ταξινομικού εννοιολογικού πλαισίου.	133
5.3 Παράδειγμα ασαφούς ταξινομικού πλαισίου.	133
5.4 Χρονικά παράθυρα ταιριάσματος εννοιολογικού πλαισίου και προτιμήσεων χρηστών.	135
5.5 Παράδειγμα εξαγωγής ασαφών προτιμήσεων χρηστών.	145
5.6 Η δομή ενός Προφίλ Χρήστη.	147
5.7 Δομή αντιστοίχισης των Προφίλ Χρηστών.	148
5.8 Παράδειγμα αντιστοίχισης στατικού προφίλ (1ο τμήμα του ηλεκτρονικού ερωτηματολογίου) - ανάλογα με την επιλογή του χρήστη στο αριστερό τμήμα προκύπτει η σχετική αντιστοίχιση σε προφίλ του δεξιού τμήματος.	148
5.9 Παράδειγμα αντιστοίχισης στατικού προφίλ (2ο τμήμα του ηλεκτρονικού ερωτηματολογίου) - 8 ειδικά βάρη αποδίδονται στις 8 διαφορετικές πιθανές επιλογές/απαντήσεις, σύμφωνα με τον χανόνα: $Eπιλογή(\alpha \dots \zeta) > 4 \Rightarrow \text{"Έμπειρος"}, 2 < Eπιλογή(\alpha \dots \zeta) < 4 \Rightarrow \text{"Προχωρημένος"}, 2 < Eπιλογή(\alpha \dots \zeta) \wedge Eπιλογή(\eta) \Rightarrow \text{"Αρχάριος"}.$	149

5.10 Παράδειγμα αντιστοίχισης στατικού προφίλ (3ο τμήμα του ηλεκτρονικού ερωτηματολογίου) .....	149
5.11 Αρχική αντιστοίχιση στατικών προφίλ - χαρακτηρισμός χρηστών για τρία ανεξάρτητα μέρη ενός ηλ. ερωτηματολογίου.....	150
5.12 Υποστριζόμενες γλώσσες δεδομένων. ....	151
5.13 Παράδειγμα στατιστικής ανάλυσης μιας ερώτησης ηλεκτρονικού ερωτηματολογίου, συνδεδεμένης με τις κατηγορίες των προφίλ. ....	155
5.14 Συνδυασμένη πληροφορία προφίλ με βάση τα αποτελέσματα της συσταδοποίησης και για τα τρία μέρη ενός ηλεκτρονικού ερωτηματολογίου.....	155
5.15 Σημασιολογική επέκταση των προτιμήσεων χρηστών και του εννοιολογικού πλαισίου κατά το χρόνο εκτέλεσης.....	157
5.16 Συγκριτική απόδοση της προσωποποιημένης αναζήτησης με και χωρίς επίδραση του εννοιολογικού πλαισίου για το ερώτημα: “επιχειρήσεις που είναι εισηγμένες στο Χρηματιστήριο Αξιών της Νέας Υόρκης και έχουν ένα εμπορικό σήμα στις Η.Π.Α.”. Το γράφημα παρουσιάζει i) την καμπύλη ακρίβειας-ανάκλησης και ii) την καμπύλη μέσης σχετικότητας-ανάκλησης. ....	164
5.17 Συγκριτική απόδοση της προσωποποιημένης αναζήτησης με και χωρίς την επίδραση του εννοιολογικού πλαισίου· μέσος όρος δέκα περιπτώσεων χρήσης.....	165
5.18 Συγκριτική επίδοση της προσωποποιημένης αναζήτησης με και χωρίς την επίδραση του εννοιολογικού πλαισίου· καταδεικνύεται η καμπύλη ακρίβειας-ανάκλησης για i) ένα από τα σενάρια και ii) το μέσο όρο των 10 σεναρίων. Τα αποτελέσματα στο γράφημα i) αντιστοιχούν στο ερώτημα “Εταιρίες που έχουν τη βάση τους σε οποιαδήποτε Ιαπωνική περιοχή”, σύμφωνα με το σενάριο χρήσης που περιγράφεται στην ενότητα 5.8. ....	166
5.19 Συγκριτικό ιστόγραμμα ακρίβειας της εξατομικευμένης αναζήτησης με και χωρίς εννοιολογικό πλαίσιο των δέκα περιπτώσεων χρήσης. Οι ελαφρώς χρωματισμένες μπάρες συγχρίνουν την εξατομικευμένη ανάκτηση με εννοιολογικό πλαίσιο και την απλή εξατομικευμένη ανάκτηση χωρίς εννοιολογικό πλαίσιο. Οι σκούρες μπάρες συγχρίνουν την εξατομικευμένη ανάκτηση με εννοιολογικό πλαίσιο και την ανάκτηση χωρίς καθόλου προσωποποίηση. ....	167
5.20 Συγκριτική επίδοση της προσωποποιημένης αναζήτησης με και χωρίς τη χρήση εννοιολογικού πλαισίου από 18 άτομα σε τρεις προτεινόμενους στόχους. Η γραφική παράσταση παρουσιάζει α) την καμπύλη ακρίβειας-ανάκλησης και β) την ακρίβεια στα σημεία διακοπών (cut-off points). Λαμβάνεται η μέση τιμή των αποτελεσμάτων από το σύνολο όλων των χρηστών και όλων των στόχων. ....	168
5.21 Παράδειγμα κατασκευής της σχέσης $T$ . ....	171
5.22 Ένα υποσύνολο της σχέσης $\widehat{T}$ , το οποίο συνδέει τις έννοιες που εμπλέκονται στην επέκταση του κατά το χρόνο εκτέλεσης εννοιολογικού πλαισίου της 'Έλλης.....	172
5.23 Ένα υποσύνολο της οντολογίας θεματικής περιοχής, το οποίο περιέχει τις έννοιες που περιλαμβάνονται στο δεύτερο σενάριο χρήσης..	175

5.24 Η εικόνα και το βίντεο που επιλέχθηκαν από το χρήστη.	176
6.1 Επάνω σειρά: χιόνι, σύννεφα και κύματα εκτός εννοιολογικού πλαισίου· κάτω σειρά: Οι ίδιες έννοιες εντός του εννοιολογικού τους πλαισίου.	180
6.2 Αναγνώριση παρόμοιων υλικών χωρίς τη βοήθεια του εννοιολογικού πλαισίου· τα επιλεγμένα τμήματα είναι αυτά που απεικονίζουν χιόνι στην πραγματικότητα.	180
6.3 Απομονωμένα αντικείμενα και αντικείμενα συνυφασμένα με το εννοιολογικό τους πλαίσιο.	181
6.4 Παράδειγμα χωρικών σχέσεων περιοχών μιας εικόνας παραλίας - (α) - (γ) Παραδείγματα εικόνων που αναπαριστούν μια παραλία. (δ) Υπερτοπικές χωρο-εννοιολογικές σχέσεις.	185
6.5 Αναπαράσταση μιας ασαφούς σχέσης σε RDF - χρήση της τεχνικής reification	190
6.6 Παραδείγματα οντολογιών θεματικών περιοχών.	192
6.7 Παράδειγμα αναπαράστασης γράφου - εκτίμηση δείκτη συμβατότητας.	193
6.8 Παράδειγμα αναπαράστασης της τεχνικής OWL reification.	197
6.9 Δείγμα γράφου οντολογίας - θεματικές υπο-περιοχές: παραλία, βουνό, πόλη	198
6.10 Ενδεικτικό παράδειγμα από μια εικόνα παραλίας.	200
6.11 Αρχική ανάθεση εικετών σε περιοχές με βάση τον ARG και την αντιστοίχηση των MPEG-7 Περιγραφέων.	203
6.12 Επίδραση του εννοιολογικού πλαισίου σε μία εικόνα παραλίας.	209
6.13 Αρχιτεκτονική υποβοηθούμενου από τη γνώση πλαισίου ανάλυσης.	212
6.14 Υποδομή Γνώσης.	214
6.15 Πειραματικά αποτελέσματα για τη θεματική περιοχή παραλία.	218
6.16 Πειραματικά αποτελέσματα για τη θεματική περιοχή βουνό.	219
6.17 Υπόδειγμα αποτελέσμάτων για τη θεματική περιοχή των καλοκαιρινών διακοπών. Παρουσιάζονται η αρχική εικόνα και τα αποτελέσματα της προσέγγισης πριν (β' στήλη) και μετά (γ' στήλη) την επίδραση του εννοιολογικού πλαισίου.	220
7.1 Μια εικόνα και η χονδροειδής κατάτμησή της.	222
7.2 Παράδειγμα συλλογής φωτογραφιών με έντονο το χρονικό εννοιολογικό πλαίσιο: η πρώτη σειρά εικόνων αναφέρεται στο ίδιο γεγονός, ενώ η δεύτερη σειρά περιέχει ξεκάθαρα δύο ξεχωριστές ομάδες γεγονότων, τα οποία απέχουν μεταξύ τους χρονικά.	227
7.3 Δεύτερο παράδειγμα αρχικής εικόνας και της χονδροειδούς κατάτμησής της.	232
7.4 Επιλογή τύπων περιοχής μέσω ιεραρχικής συσταδοποίησης. Οι έξι επιλεγμένοι τύποι περιοχής καταδεικνύονται με το κόκκινο περίγραμμα.	234
7.5 Μεθοδολογία ανίχνευσης εννοιών υψηλού επιπέδου	236
7.6 Μια αρχική εικόνα εισαγωγής και η χονδροειδής κατάτμησή της.	236
7.7 Αποστάσεις ανάμεσα σε περιοχές εικόνας και τύπους περιοχής: αριστερά παρουσιάζονται αποστάσεις ανάμεσα σε μία περιοχή της εικόνας και σε όλους τους τύπους περιοχής, ενώ δεξιά αποστάσεις ανάμεσα σε όλες τις περιοχές μίας εικόνας και έναν τύπο περιοχής	240
7.8 Παραδείγματα διαγυμνάστων αναπαράστασης	241

7.9 Τμήμα της οντολογίας εννοιολογικού πλαισίου από τη θεματική περιοχή παραλία για την περίπτωση χρήσης υψηλού επιπέδου εννοιών με βάση ενδιάμεσου επιπέδου χαρακτηριστικά. Η έννοια παραλία είναι το στοιχείο ρίζα. ....	245
7.10 Μεθοδολογία επίδρασης εννοιολογικού πλαισίου ταξινόμησης εννοιών .	245
7.11 Γωνία μεταξύ δύο σημείων και του X άξονα .....	250
7.12 Ασαφή σύνολα για τη χωρική σχέση τεσσάρων κατευθύνσεων: “αριστερά του”, “δεξιά του”, “πάνω” και “κάτω”. ....	250
7.13 Ιστόγραμμα που συσχετίζεται με το Θ. ....	251
7.14 Ένα τμήμα της οντολογίας τύπων περιοχής για τη θεματική περιοχή παραλία. ....	253
7.15 Επιλογή τύπων περιοχής μέσω ιεραρχικής συσταδοποίησης. ....	254
7.16 Τμήμα RDF οντολογίας τύπων περιοχής. ....	254
7.17 Μεθοδολογία επίδρασης εννοιολογικού πλαισίου ταξινόμησης τύπων περιοχής .....	255
7.18 Ένα τμήμα της οντολογίας του επεκτεταμένου οπτικού εννοιολογικού πλαισίου. Εμφανίζονται μόνο οι σχέσεις μεταξύ της υψηλού επιπέδου έννοιας θάλασσα και όλων των υπολοίπων εννοιών και τύπων περιοχής. Οι αριθμοί καταδεικνύουν τον ασαφή βαθμό εμπιστοσύνης για κάθε σχέση. ....	259
7.19 Ένα άλλο τμήμα της οντολογίας του επεκτεταμένου οπτικού εννοιολογικού πλαισίου. Εμφανίζονται μόνο οι σχέσεις μεταξύ του τέταρτου τύπου περιοχής (Τ4) και όλων των υπολοίπων εννοιών και τύπων περιοχής. Οι αριθμοί καταδεικνύουν τον ασαφή βαθμό εμπιστοσύνης για κάθε σχέση. ....	260
7.20 Διάγραμμα ροής της επίδρασης του μεικτού εννοιολογικού πλαισίου ταξινόμησης .....	262
7.21 Μεθοδολογία επίδρασης μεικτού εννοιολογικού πλαισίου ταξινόμησης .	264
7.22 Ένα απλό διάγραμμα ροής για ένα σενάριο χρήσης εννοιολογικού πλαισίου εννοιών. ....	265
7.23 Ενδεικτικό παράδειγμα εικόνας παραλίας. ....	266
7.24 Ενδεικτικό παράδειγμα οπτικής ομοιότητας περιοχών. ....	267
7.25 Ένα απλό διάγραμμα ροής για ένα σενάριο χρήσης εννοιολογικού πλαισίου τύπων περιοχής.....	268
7.26 Ενδεικτικό δείγμα αρχικής εικόνας. ....	268
7.27 1 <sup>η</sup> εικόνα παραλίας. ....	270
7.28 2 <sup>η</sup> εικόνα παραλίας. ....	270
7.29 3 <sup>η</sup> εικόνα παραλίας. ....	270
7.30 Ενδεικτικές εικόνες από τη συλλογή Corel. ....	272
7.31 Ενδεικτικές εικόνες από τη συλλογή TRECVID. ....	276
7.32 Ενδεικτικά αποτελέσματα καθολικής ταξινόμησης· εικόνες από τη θεματική περιοχή παραλία εμφανίζονται στην επάνω σειρά, ενώ εικόνες από την περιοχή βουνό παρουσιάζονται στην κάτω σειρά. ....	276

# Κατάλογος Πινάκων

3.1	Ασαφής αναπαράσταση σχέσης σε RDF: RDF reification.....	57
4.1	Ασαφείς ταξινομικές σχέσεις που χρησιμοποιούνται για τη δημιουργία του πρώτου μέρους του κατάλληλου μοντέλου γνώσης. ....	100
4.2	Βάση Γνώσης ( <i>TBox</i> ) .....	104
4.3	Ονόματα εννοιών και τα μνημονικά τους. Οι θεματικές κατηγορίες παριστάνονται με έντονα γράμματα. ....	106
4.4	Η ταξινομική σχέση $T$ . Μηδενικά και υπονοούμενα από τη μεταβατικότητα στοιχεία έχουν παραληφθεί. ....	106
4.5	Αποτελέσματα κατηγοριοποίησης εγγράφων. Τιμές κάτω από 0.1 έχουν αγνοηθεί. ....	108
4.6	Ασαφής μήτρα σύγχυσης $D$ για τη σημασιολογική κατηγοριοποίηση των πολυμεσικών προγραμμάτων. ....	112
4.7	Ασαφής μήτρα σύγχυσης $D$ για τη σημασιολογική κατηγοριοποίηση των ειδησεογραφικών αντικειμένων. ....	112
4.8	Αποτελέσματα αξιολόγησης σημασιολογικής κατηγοριοποίησης - 653 πολυμεσικά προγράμματα. ....	113
4.9	Αποτελέσματα αξιολόγησης σημασιολογικής κατηγοριοποίησης - 2733 ειδησεογραφικά αντικείμενα. ....	113
4.10	Αποτελέσματα κατηγοριοποίησης εγγράφων. Τιμές < 0.10 αγνοούνται. ....	115
4.11	Ασαφείς ισχυρισμοί εννοιών για το έγγραφο $d_2$ .....	116
4.12	Κατηγοριοποίηση Εγγράφου $d_2$ .....	116
5.1	Ασαφείς ταξινομικές σχέσεις που χρησιμοποιήθηκαν για την κατασκευή της σχέσης $T$ . ....	129
5.2	Επιλογή χαρακτηριστικών γνωρισμάτων. ....	152
5.3	Αποτελέσματα συσταδοποίησης για 100 χρήστες - 3 συστάδες. ....	153
5.4	Αποτελέσματα συσταδοποίησης για 100 χρήστες - 5 συστάδες. ....	153
5.5	Αποτελέσματα συσταδοποίησης για 100 χρήστες - 9 συστάδες. ....	153
5.6	Μέση ακρίβεια (MAP) για κάθε ένα από τα τρία είδη αναζήτησης .....	168
5.7	Ονόματα εννοιών και τα αντίστοιχα μνημονικά τους. Οι θεματικές κατηγορίες εμφανίζονται με έντονα μαύρα γράμματα .....	170
5.8	Μέρος της ταξινομικής σχέσης $T$ . ....	170
5.9	Μέρος της σχέσης $\widehat{T}$ .....	173
5.10	Προτιμήσεις χρήστη. ....	175
5.11	Βάρη διάδοσης των σημασιολογικών σχέσεων. ....	175
5.12	Σύνολο εννοιολογικού πλαισίου. ....	176
5.13	Επεκτεταμένο εννοιολογικό πλαίσιο. ....	177
5.14	Επεκτεταμένες προτιμήσεις χρήστη. ....	177

5.15 Προτιμήσεις χρήστη με βάση το εννοιολογικό πλαίσιο.....	177
6.1 Τελικοί βαθμοί συμμετοχής πριν και μετά την εφαρμογή του οπτικού εννοιολογικού πλαισίου για το παράδειγμα της εικόνας από τη θεματική περιοχή παραλία .....	201
6.2 Συνολικά σκορ ακρίβειας ανά έννοια για 573 εικόνες .....	202
6.3 Τελικοί βαθμοί ιδιότητας μέλους πριν και μετά από την ενίσχυση του οπτικού εννοιολογικού πλαισίου για ένα παράδειγμα εικόνας παραλίας - οι βαθμοί με έντονα μαύρα γράμματα καταδεικνύουν την πληροφορία απόλυτης αλήθειας της εικόνας και η πρώτη στήλη δείχνει το συγκεκριμένο αριθμό της περιοχής υπό εξέταση. ....	210
6.4 Συνολικά και ανά έννοια αποτελέσματα ανίχνευσης για τη θεματική περιοχή της παραλίας. Η στήλη α περιέχει τα αποτελέσματα του αλγορίθμου σημασιολογικής κατάτυμησης watershed, η 2η στήλη τα αποτελέσματα του ίδιου αλγορίθμου με χρήση του εννοιολογικού πλαισίου, η στήλη γ τα αποτελέσματα του παραδοσιακού αλγορίθμου κατάτυμησης watershed, η στήλη δ τα αποτελέσματα του σημασιολογικού RSST, η ε τα αποτελέσματα του σημασιολογικού RSST που χρησιμοποιεί το εννοιολογικό πλαίσιο και η ζ του παραδοσιακού RSST.....	211
6.5 Αριθμητική αποτίμηση για τη θεματική περιοχή παραλία. ....	217
6.6 Αριθμητική αποτίμηση για τη θεματική περιοχή βουνό. ....	217
7.1 Ασαφείς σημασιολογικές σχέσεις μεταξύ εννοιών. ....	243
7.2 Προτεινόμενες ασαφείς τοπολογικές σχέσεις. ....	249
7.3 Σχέσεις εννοιολογικού πλαισίου στην περίπτωση των τύπων περιοχής. 252	
7.4 Σχέσεις εννοιολογικού πλαισίου μεταξύ οντοτήτων. ....	257
7.5 Ασαφείς σημασιολογικές σχέσεις μεταξύ της υψηλού επιπέδου έννοιας $C_1/\text{ουρανος}$ και κάποιων ενδεικτικών υπόλοιπων οντοτήτων. Οι αριθμοί καταδεικνύουν τον ασαφή βαθμό εμπιστοσύνης για κάθε σχέση. 260	
7.6 Οι βαθμοί της σημασιολογικής σχέσης Συνοδός και της τοπολογικής σχέσης Συνύπαρξη για κάποια ενδεικτικά ζεύγη οντοτήτων που αυτές ορίζονται. Οι αριθμοί καταδεικνύουν τον ασαφή βαθμό εμπιστοσύνης για κάθε σχέση. ....	261
7.7 Συνολικά σκορ ακρίβειας και ανάκλησης πριν και μετά την επίδραση του εννοιολογικού πλαισίου ανά υψηλού επιπέδου έννοια από τη θεματική περιοχή παραλία .....	267
7.8 Συνολικά σκορ ακρίβειας και ανάκλησης για τις 1148 εικόνες και τις 6 υψηλού επιπέδου έννοιες θάλασσα, ουρανός, άμμος, χύμα, βλάστηση και βράχος. ....	269
7.9 Ζεύγη βαθμών εμπιστοσύνης ανά εικόνα παραλίας. ....	271
7.10 Συνολικά σκορ ακρίβειας ανά έννοια - 1148 εικόνες. ....	271
7.11 Ακρίβεια/ανάκληση ανά έννοια πριν και μετά την εφαρμογή του αλγορίθμου επίδρασης του μεικτού εννοιολογικού πλαισίου - 750 εικόνες Corel. ....	272
7.12 Σκορ ακρίβειας και ανάκλησης ανά έννοια πριν και μετά την εφαρμογή του μεικτού εννοιολογικού πλαισίου - 1250 εικόνες TRECVID. ....	272

7.13 Συγχριτικά αποτελέσματα ακρίβειας $P$ και ανάκλησης $R$ ανά έννοια για 6 διαφορετικές μεθοδολογίες ανίχνευσης εννοιών υφηλού επιπέδου πάνω στις 750 εικόνες του Corel. ....	274
7.14 Συγχριτικά αποτελέσματα ακρίβειας $P$ και ανάκλησης $R$ ανά έννοια για 6 διαφορετικές μεθοδολογίες ανίχνευσης εννοιών υφηλού επιπέδου πάνω στις 1147 εικόνες του TRECVID. ....	274
7.15 Σύνολα εκπαίδευσης και δοκιμών για την καθολική ταξινόμηση σε εικόνες παραλίας/βουνού. ....	275
7.16 Αποτελέσματα καθολικής ταξινόμησης εικόνων παραλίας και βουνού πριν και μετά την επίδραση του εννοιολογικού πλαισίου για όλες τις εικόνες (Συλλογές, TRECVID και Corel). ....	276



## ΠΡΟΛΟΓΟΣ

Αναμφίβολα, τόσο η φύση, όσο και η δομή της εποχής μας χαρακτηρίζονται από τη γρήγορη ανταλλαγή πληροφοριών και τις δυνατότητες άμεσης και παγκόσμιας επικοινωνίας. Το γεγονός αυτό επιδρά σημαντικά τους σύγχρονους ανθρώπους και επηρεάζει μια σειρά από ποικίλες καθημερινές δραστηριότητές τους, οι οποίες εκτείνονται από την επαγγελματική έως την προσωπική τους ζωή. Ειδικότερα στον τομέα της διασκέδασης και ψυχαγωγίας, τα πολυμέσα παίζουν σήμερα πρωταγωνιστικό ρόλο, μιας και συναντώνται με πάρα πολλές μορφές. Η σύγχρονη ψυχαγωγία του ανθρώπου κινείται γύρω από συσκευές τελευταίας τεχνολογίας, όπως είναι τα κινητά τηλέφωνα, οι τηλεοράσεις υψηλής ευκρίνειας (HDTV), οι συσκευές οικιακού κινηματογράφου (home cinema), η δορυφορική και ψηφιακή τηλεόραση, οι οποίες στηρίζουν κατά ένα μεγάλο ποσοστό την δημοτικότητά τους στο πολυμεσικό περιεχόμενο που διαχειρίζονται, παρουσιάζουν, αποθηκεύουν, αποστέλλουν, λαμβάνουν, κ.ο.κ.. Κατά συνέπεια, γίνεται κατανοητή η ιδιαίτερη βαρύτητα που αποκτά η αποδοτική διαχείριση πολυμεσικού περιεχομένου στη σημερινή εποχή. Τα χαρακτηριστικά της τελευταίας επηρεάζουν τον τρόπο αναζήτησης και ανάκτησης του πολυμεσικού περιεχομένου από τους εκάστοτε τελικούς χρήστες, καθώς και τον ίδιο τον ρυθμό που αυτός εξελίσσεται. Όμως, σε μια εποχή που όλοι και όλα αλλάζουν συνεχώς, η κοινότητα των πολυμέσων δε θα μπορούσε να μείνει απαθής και αμέτοχη, τη στιγμή μάλιστα, που το πολυμεσικό περιεχόμενο βρίσκεται στο επίκεντρο του παγκόσμιου ενδιαφέροντος με ολοένα και περισσότερες αυξητικές τάσεις. Δεδομένου ότι οι ανάγκες για ψυχαγωγία, η οποία θα βασίζεται στο πολυμεσικό περιεχόμενο, παρακολουθούν την ανάπτυξη και τις αλλαγές της γενικότερης κοινωνίας, διανύουμε μια περίοδο που όλες οι σχετικές παραδοσιακές τεχνικές αναθεωρούνται και επαναξιολογούνται, ενώ ταυτόχρονα αναπτύσσονται καινοτόμες ή ακόμα και ριζοσπαστικές, μερικές φορές, καινούριες μεθοδολογίες, προκειμένου να αντεπεξέλθουν στην αλματώδη ανάπτυξη του χώρου.

Όπως είναι φυσικό, πάνω απ' όλα βρίσκεται στο προσκήνιο το Διαδίκτυο (Internet) και διαδραματίζει έναν ιδιαίτερα σημαντικό ρόλο σε όλους τους παραπάνω τομείς, προωθώντας και αυτό με τον τρόπο του τις διαδικασίες της εξέλιξης. Αναμφίβολα, η ίδια η αλληλεπιδραστική υποδομή του Διαδικτύου έχει συμβάλει στον πλούτο της διαθέσιμης πολυμεσικής πληροφορίας. Στην πραγματικότητα, η φύση του πολυμεσικού περιεχομένου που είναι διαθέσιμο, το κάνει ελκυστικό στα διαφορετικά και ποικίλα σύνολα των χρηστών του, ανάλογα με το επίπεδο της εμπειρίας τους, τα ενδιαφέροντά τους, τις διανοητικές και τεχνολογικές τους δυνατότητες. Η τελευταία διαπίστωση καταδεικνύει τη νέα μορφή συστημάτων που απαιτούνται, καθώς επίσης και τη νέα μορφή χρηστών που αναπτύσσεται μέσα στο πλαίσιο της κοινωνίας μας, η οποία συνεχώς εμπλουτίζεται με διάφορα είδη τεχνολογικών “βελτιώσεων”. Σήμερα, οι χρήστες των πολυμεσικών συστημάτων διαφέρουν ριζικά από τους αντίστοιχους χρήστες της προηγούμενης δεκαετίας, όντας πρακτικότεροι και πιο εξοικειωμένοι με

την τεχνολογία από κάθε άλλη εποχή της ανθρωπότητας.

Όσον αφορά στο διαθέσιμο πολυμεσικό περιεχόμενο, η μεγάλη εξέλιξη των ηλεκτρονικών αποθηκευτικών μέσων αλλά και των δυνατοτήτων ψηφιακής απεικόνισης και αποθήκευσης έχει πλέον επιτρέψει τη δημιουργία μεγάλων ηλεκτρονικών βάσεων δεδομένων, τα περιεχόμενα των οποίων ποικίλουν από προσωπικές συλλογές φωτογραφιών ή/και ακολουθιών βίντεο σε ψηφιακά βιβλία έως οποιοδήποτε υλικό ηλεκτρονικής μορφής είναι σε θέση να αποθηκευθεί. Οι συλλογές αυτές, ως τεράστιοι αποθηκευτικοί χώροι, έχουν απωλέσει πλέον την έννοια της κλασικής βιβλιοθήκης. Εμφανίζουν, δε, μια σειρά από συγκριτικά πλεονεκτήματα έναντι των τελευταίων, όπως είναι η οικονομία χώρου, το χαμηλό κόστος λειτουργίας, η δυνατότητα επαναχρησιμοποίησης του διαθέσιμου υλικού, η δυνατότητα ταυτόχρονης διάθεσης του ίδιου υλικού σε πολλαπλούς χρήστες και η εύκολη πρόσβαση από όλους μέσω του Διαδικτύου.

Τα ανωτέρω πλεονεκτήματα, σε συνδυασμό με το μέγεθος των συστημάτων αυτών, καθιστούν ανεπαρκείς τις όποιες κλασικές αρχές βιβλιοθηκονομίας θα μπορούσαν τυχόν να εφαρμοστούν και απαιτούν την ανάπτυξη ευφυών τεχνικών και μεθοδολογιών αναζήτησης και ανάκτησης της πολυμεσικής πληροφορίας. Τα νέα αυτά συστήματα, τα συστήματα ανάκτησης πληροφορίας (Information Retrieval Systems - IRS), όπως ονομάζονται, θα πρέπει να είναι σε θέση να αναλάβουν την έως ένα βαθμό αυτοματοποιημένη οργάνωση των ηλεκτρονικών αποθηκών και τη διεξαγωγή γρήγορων και πολύπλοκων αναζητήσεων σε αυτές. Εν γένει, τα συστήματα ανάκτησης πληροφορίας επιχειρούν να εντοπίσουν τα επικέρυκα στοιχεία από το διαθέσιμο πολυμεσικό υλικό που ικανοποιούν τις ανάγκες και τις προτιμήσεις του χρήστη. Η διαδικασία αυτή αντιμετωπίζει διάφορα εγγενή προβλήματα που εστιάζονται κυρίως στην αδυναμία πλήρους περιγραφής των αναγκών και προτιμήσεων του χρήστη και του περιεχομένου των διαθεσίμων στοιχείων. Η έρευνα που επικεντρώνεται στα συστήματα ανάκτησης πληροφορίας επιχειρεί να λύσει τα παραπάνω προβλήματα με τη χρήση αλληλεπιδραστικών μεθόδων, τόσο για την οργάνωση όσο και για την αναζήτηση του υλικού.

Σκοπός της παρούσας εργασίας είναι να συνεισφέρει προς αυτή την κατεύθυνση, αξιοποιώντας μια βασισμένη στο εννοιολογικό πλαίσιο (context) θεώρηση της ανωτέρω διαδικασίας. Κατά κύριο λόγο, επιχειρεί να αντιμετωπίσει τα θέματα της επεξεργασίας, αναζήτησης και ανάκτησης πολυμεσικής πληροφορίας μέσα από το πρόσμα του εννοιολογικού πλαισίου, όπως αυτό εκφράζεται στο πεδίο της προσαρμογής της γνώσης και της πρόσβασης στην πληροφορία. Ειδικότερα, δύο είναι οι κεντρικοί άξονες που αποτέλεσαν και αποτελούν το κίνητρό μας για την έρευνα της συγκεκριμένης θεματικής περιοχής: (i) η προσωποποίηση του πολυμεσικού περιεχομένου και (ii) η ανάλυση του πολυμεσικού περιεχομένου με βάση το - ποικίλως εννοούμενο - εννοιολογικό πλαίσιο.

Αξίζει να τονιστεί ότι τόσο η προσωποποίηση πολυμεσικού περιεχομένου όσο και η ανάλυση πολυμεσικού περιεχομένου με βάση το εννοιολογικό πλαίσιο έχουν άμεση σχέση με το χρήστη, προσδίδοντας μία ανθρωποεντρική ματιά στην έρευνά μας. Η πρόσβαση στην πολυμεσική πληροφορία, σε συνδυασμό με την κατάλληλη προσαρμογή της γνώσης αποτελούν τον πυρήνα της διδακτορικής διατριβής και το αντικείμενο της επιχειρούμενης εμβάθυνσης. Πρόκειται, ουσιαστικά, για τις δύο όψεις του ίδιου νομίσματος: της αλληλεπίδρασης (interaction) με το εκάστοτε σύστημα αναζήτησης ή ανάκτησης πληροφορίας· θεωρώντας το πρόβλημα από την σκοπιά της προσωποποίησης πολυμεσικού περιεχομένου βρίσκεται στο επίκεντρο ο χρήστης, ενώ από την άλλη πλευρά της μοντελοποίησης, ανάλυσης και διαχείρισης του εννοιολογικού πλαισίου

βρίσκεται η βούληση του χρήστη με την μορφή των αντικειμένων της αναζήτησής του.

Το παρόν κείμενο συγκεντρώνει τους ερευνητικούς κόπους και τις προσπάθειες των τελευταίων 6 ετών και θα αποτελούσε σοβαρή παράλειψη αν ο υποφαινόμενος δεν εξέφραζε στο σημείο αυτό τις ειλικρινείς και θερμές ευχαριστίες του, τόσο προς επιβλέποντα καθηγητή Ε.Μ.Π. Δρ. Στέφανο Κόλλια, όσο και προς τον Δρ. Ιωάννη Αβρίθη, για την λαμπρή καθοδήγηση και αμέριστη υποστήριξή τους σε όλα τα στάδια της έρευνάς του. Το κείμενο της διατριβής συνιστά μια αναφορά της βιβλιογραφικής ενημέρωσής του σε θέματα που βρίσκονται στο γενικότερο πλαίσιο της ευφυούς αλληλεπίδρασης του χρήστη με το σύστημα, καθώς και μια περίληψη της πρωτότυπης ερευνητικής εργασίας που έχει προκύψει ως αποτέλεσμα της ενασχόλησής του με το χώρο. Περιέχει μια αναφορά της ερευνητικής διαδρομής του, ενώ στο τέλος του περιλαμβάνεται ένας πλήρης κατάλογος δημοσιεύσεων και λοιπών ερευνητικών δραστηριοτήτων. Ελπίζω αυτή η διδακτορική διατριβή να αποβεί χρήσιμη για έρευνα στις ραγδαία εξελισσόμενες περιοχές της πολυμεσικής ανάλυσης και ανάκτησης περιεχομένου από τους νεότερους.

*Φοίβος-Απόστολος Ι. Μυλωνάς  
Αθήνα, Απρίλιος 2008*



## ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες σε όλους όσους μου συμπαραστάθηκαν και με βοήθησαν σε όλη τη διάρκεια της εκπόνησης της παρούσας διδακτορικής διατριβής τα τελευταία έξι χρόνια. Κατ' αρχήν θα ήθελα να ευχαριστήσω ιδιαίτερα τον επιβλέποντα Καθηγητή Ε.Μ.Π., κ. **Στέφανο Κόλλια**, για την πολύτιμη βοήθεια, το ειλικρινές ενδιαφέρον και την αμέριστη υποστήριξή του, τόσο σε επιστημονικό, όσο και σε προσωπικό επίπεδο. Η ουσιαστική συνεισφορά του στη διατριβή μου υπήρξε αποφασιστική, ενώ τα σχόλια και η εποικοδομητική κριτική του στο κείμενο της διατριβής βοήθησαν καθοριστικά στην ποιότητά της. Πάνω απ' όλα όμως τον ευχαριστώ για την ελευθερία που μου άφησε στην ερευνητική κατεύθυνση της διατριβής και στην ανάπτυξη και αξιοποίηση νέων ιδεών και προτάσεων. Θερμές ευχαριστίες οφείλω και στα άλλα μέλη της Τριμελούς Επιτροπής, Καθηγητές Ε.Μ.Π. κ. **Ανδρέα Σταφυλόπατη** και κ. **Παναγιώτη Τσανάκα**.

Ξεχωριστές και ιδιαίτερες ευχαριστίες ανήκουν στο Δρ. **Ιωάννη Αβρίθη**, χωρίς την πολύτιμη βοήθεια, καθοδήγηση και αμέριστη υποστήριξη του οποίου, η ολοκλήρωση της διδακτορικής διατριβής μου δε θα ήταν δυνατή. Η ουσιαστική συνεισφορά του σε όλη τη διάρκεια της ερευνητικής μου πορείας υπήρξε καθοριστική, τόσο σε επιστημονικό, όσο και σε προσωπικό επίπεδο, ενώ τα σχόλια και η εποικοδομητική κριτική του στο κείμενο της διατριβής θεωρώ ότι έθεσαν τις βάσεις για την ποιότητά του. Τον ευχαριστώ, δε, ιδιαίτερα για την άπλετη παροχή ερευνητικών συμβουλών και καθοδήγησης, καθώς και για το αυθεντικό ενδιαφέρον που επέδειξε για την ανάπτυξη και αξιοποίηση νέων ιδεών και προτάσεων στην ερευνητική κατεύθυνση της διατριβής. Οι γνώσεις, η εμπειρία και η διάθεσή του οδήγησαν τη συγκεκριμένη ερευνητική δουλειά σε αυτό το σημείο, ενώ οποιαδήποτε στιγμή ήταν έτοιμος να προσφέρει ό,τι ήταν δυνατό ώστε να ξεπεραστεί κάθε δυσκολία.

Θα ήθελα επίσης να ευχαριστήσω όλους τους συναδέλφους μου στο Εργαστήριο Ψηφιακής Επεξεργασίας Εικόνας, Βίντεο και Πολυμέσων (ΨΕΕΒΠ) για τη βοήθειά τους στη δημιουργία ενός ευχάριστου περιβάλλοντος εργασίας όλα αυτά τα χρόνια, αλλά ιδιαίτερα εκείνους με τους οποίους συνεργάστηκα στο πλαίσιο της ερευνητικής μου εργασίας. Ειδικότερα είμαι υπόχρεος στον συνάδελφο Δρ. **Μανόλη Wallace** για την πολύτιμη προσφορά του στα πρώμα στάδια της έρευνάς μου και τη σημαντική συμβολή του στις ερευνητικές διεργασίες που περιγράφονται στο κεφάλαιο 4, ενώ ιδιαίτερη αναφορά θα πρέπει να γίνει στον Δρ. **Γιώργο Ακρίβα** για την σημαντική συνεισφορά του όσον αφορά στο θεωρητικό υπόβαθρο της παρούσης εργασίας. Επίσης, ξεχωριστή μνεία θα πρέπει να γίνει στους συναδέλφους **Θάνο Αθανασιάδη**, **Νίκο Σίμου**, **Βαγγέλη Σπύρου** και **Γιώργο Τόλια** για την άριστη συνεργασία που είχαμε στο πλαίσιο των ερευνητικών προσπαθειών που παρουσιάζονται στα κεφάλαια 6, 4 και 7, αντίστοιχα. Τέλος, θα ήθελα να εκφράσω την εκτίμησή μου και τις ευχαριστίες μου προς τους συναδέλφους και φίλους Γιώργο Ανδρέου, Κώστα Καρπούζη, Πέτρο

Καφάλα, Βασίλη Τζουβάρα και Παρασκευή Τζούβελη για τις ερευνητικές συνεργασίες που είχαμε καθ' όλο το χρονικό διάστημα εκπόνησης της διατριβής.

Οφείλω, τέλος, να εκφράσω τις βαθύτατες ευχαριστίες μου στη γυναίκα μου, **Κλειώ Στάμου**, για την ηθική υποστήριξη και τη συνεχή ενθάρρυνση που μου παρείχε στη διεξαγωγή της ερευνητικής μου δραστηριότητας, καθώς και για την υπομονή που επέδειξε κατά τη διάρκεια συγγραφής του κειμένου της διατριβής. Τέλος, είμαι ευγνώμων προς την οικογένειά μου για την ανεκτίμητη συμπαράστασή της σε όλη τη διάρκεια των σπουδών μου. Χωρίς την πολύπλευρη βοήθειά τους η ολοκλήρωση της διδακτορικής μου διατριβής δε θα ήταν δυνατή.

*Φοίβος-Απόστολος Ι. Μυλωνάς  
Αθήνα, Απρίλιος 2008*

## ΠΕΡΙΛΗΨΗ

Αντικείμενο της παρούσης διδακτορικής διατριβής είναι η αντιμετώπιση θεμάτων ανάλυσης, αναζήτησης και ανάκτησης πολυμεσικού περιεχομένου με προσαρμογή στο εκάστοτε εννοιολογικό πλαίσιο της πρόσβασης σε αυτό. Δύο είναι οι κεντρικοί άξονες που αποτέλεσαν το χίνητρο αλλά και τον προσανατολισμό της έρευνας στη συγκεκριμένη θεματική περιοχή: (i) η προσωποποιημένη πρόσβαση στο πολυμεσικό περιεχόμενο, προσαρμοζόμενη στις προτιμήσεις των χρηστών και στο διαθέσιμο ιστορικό πρόσβασης σε αυτό και (ii) η αποδοτική ανάλυση και αναζήτηση του πολυμεσικού περιεχομένου με βάση το εννοιολογικό πλαίσιο. Προκειμένου να πραγματοποιηθεί μια εμπεριστατωμένη και όσο το δυνατόν πλήρης παρουσίαση των ερευνητικών πεπραγμένων, παρουσιάζονται τόσο θεωρητικά αποτελέσματα, όσο και εφαρμογές. Συγκεκριμένα, τα περιεχόμενα της διατριβής χωρίζονται στις ακόλουθες τέσσερις βασικές ενότητες:

- *Εξόρυξη δεδομένων και κατηγοριοποίηση.* Η ενότητα αυτή επιχειρεί να παρουσιάσει ορισμένες βασικές μεθοδολογίες αντιμετώπισης ερευνητικών θεμάτων ιδιαίτερου ενδιαφέροντος που άπονται άμεσα των περιοχών της εξόρυξης δεδομένων και της κατηγοριοποίησης πολυμεσικών εγγράφων. Συγκεκριμένα, προτείνεται αρχικά μια επέκταση της κλασικής μεθόδου της ιεραρχικής συσταδοποίησης, η οποία χρησιμοποιεί μια ασαφή επιλογή χαρακτηριστικών γνωρισμάτων για τον καθορισμό της μετρικής που καθορίζει τις αποστάσεις των συστάδων μεταξύ τους. Η τεχνική αυτή οδηγεί σε μία μείωση των διαστάσεων των χαρακτηριστικών γνωρισμάτων και βρίσκει εφαρμογή σε περιπτώσεις συνόλων δεδομένων που χαρακτηρίζονται από υψηλές διαστάσεις, όπως ακριβώς είναι αυτά που απαντώνται κατά την ανάλυση πολυμεσικής πληροφορίας.

Στη συνέχεια, προτείνοντας μια νέα μεθοδολογία βελτιστοποίησης της ανωτέρω τεχνικής συσταδοποίησης που βασίζεται στη χρήση μίας τεχνικής κατηγοριοποίησης, αντιμετωπίζονται κάποια από τα θεμελιώδη προβλήματα των ιεραρχικών αλγορίθμων συσταδοποίησης, όπως είναι η ευαισθησία τους σε λάθη στα αρχικά βήματα. Τα αποτελέσματα της αρχικής συσταδοποίησης βρίσκουν εφαρμογή στην ασαφή θεματική κατηγοριοποίηση εγγράφων πολυμεσικής πληροφορίας. Στη διαδικασία ορίζεται, μοντελοποιείται και λαμβάνεται υπόψη το ταξινομικό εννοιολογικό πλαίσιο με ιδιαίτερα ελπιδοφόρα αποτελέσματα. Τέλος, παρουσιάζεται μια βελτιωμένη εκδοχή της παραπάνω πρότασης κατηγοριοποίησης, η οποία λαμβάνει υπόψη της μια σειρά από υπηρεσίες εξαγωγής συμπερασμάτων, προκειμένου να προσδιοριστεί αποτελεσματικά η σημασιολογία των εγγράφων.

- *Προσωποποιημένη πρόσβαση.* Η προσωποποίηση του πολυμεσικού περιεχομένου εξετάζεται από την σκοπιά των προτιμήσεων των χρηστών και του ιστορικού

χρήσης που δημιουργείται. Δίνεται ιδιαίτερη έμφαση στη χρήση και δημιουργία κατάλληλων οντολογιών για την αναπαράσταση της απαιτούμενης γνώσης, στη χρήση στοιχείων που πηγάζουν από τη θεωρία της ασαφούς σχεσιακής άλγεβρας, στην τροποποίηση και χρήση κατάλληλων ασαφών σημασιολογικών σχέσεων, καθώς και στην επιρροή που ασκεί η έννοια του εννοιολογικού πλαισίου στη διαδικασία ανάκτησης πληροφοριών. Ειδικότερα, προτείνεται η χρήση ασαφών σχέσεων ως επέκταση των σχέσεων που λαμβάνουν λογικές τιμές 0 ή 1 και εμφανίζονται στις κλασικές οντολογίες, καθώς είναι πολύ πιο εκφραστικές κατά την αναπαράσταση της γνώσης και πολύ πιο κοντά στην αναπαράσταση των πραγματικών συνθηκών ζωής. Οι σχέσεις αυτές υλοποιούνται ακολουθώντας τα τελευταία πρότυπα του Σημασιολογικού Ιστού, όπως είναι οι γλώσσες RDF και OWL, καθώς και η τεχνική reification.

Στη συνέχεια, ορίζεται και παρουσιάζεται η έννοια του εννοιολογικού πλαισίου, η οποία στην παρούσα ερευνητική εργασία είναι διττή και λαμβάνει δύο από τις πολλαπλές πιθανές εκφάνσεις της: αυτή του ταξινομικού εννοιολογικού πλαισίου και αυτή του εννοιολογικού πλαισίου κατά το χρόνο εκτέλεσης. Και στις δύο περιπτώσεις, όμως, ορίζεται (με τη χρήση και τον κατάλληλο συνδυασμό ασαφών σημασιολογικών σχέσεων) και περιγράφεται η μεθοδολογία εκτίμησης του εννοιολογικού πλαισίου, ξεκινώντας από ένα ασαφές σύνολο σημασιολογικών οντοτήτων μιας οντολογίας. Καθορίζεται η απαιτούμενη αναπαράσταση γνώσης και εξετάζονται ζητήματα που αφορούν στον ορισμό, στην εξαγωγή και στη χρήση προτυπίσεων και κατατομών (προφίλ) χρηστών.

Τέλος, παρουσιάζονται τα αποτελέσματα της ερευνητικής δραστηριότητάς μας πάνω σε θέματα σχετικά με την ανάκτηση προσωποποιημένης πολυμεσικής πληροφορίας. Συζητείται η αυτόματη θεματική κατηγοριοποίηση πολυμεσικών εγγράφων μέσω της ιεραρχικής συσταδοποίησης των οντοτήτων που περιέχουν, η αναπαράσταση, η ανίχνευση και ο χειρισμός των προφίλ των χρηστών στο σημασιολογικό επίπεδο με τη χρήση του εννοιολογικού πλαισίου και η ανίχνευση προφίλ χρηστών από μεταδεδομένα, με χρήση της συσταδοποίησης υψηλών διαστάσεων και της ασαφούς επιλογής χαρακτηριστικών γνωρισμάτων.

- **Εννοιολογικό πλαίσιο και οπτική πληροφορία.** Ανάμεσα στους κύριους στόχους της ερευνητικής εργασίας μας βρίσκεται και η πραγματοποίηση υψηλού επιπέδου ανάλυσης και μοντελοποίησης πολυμεσικών εγγράφων<sup>1</sup>, με απώτερο στόχο την αυτόματη κατηγοριοποίηση και σημασιολογική αναζήτησή τους. Παρουσιάζεται, λοιπόν, στην τρίτη ενότητα, μια πρωτότυπη μεθοδολογία ανάλυσης της πληροφορίας κειμένου αλλά και της οπτικής πληροφορίας, η οποία βασίζεται σε γνώση αποθηκευμένη σε οντολογίες (ontologies), η ανάπτυξη των οποίων βασίζεται σε συγκεκριμένες δομές και προηγείται χρονικά των αλγορίθμων ανάλυσης. Καθώς οι οντολογίες αποτελούν ένα μέσο διαχείρισης και οργάνωσης της γνώσης με σκοπό την εξαγωγή συμπερασμάτων, η χρήση τους έχει σημαντικά αποτελέσματα όσον αφορά στις επιδόσεις αρχειοθέτησης, κατηγοριοποίησης και αναζήτησης της πολυμεσικής πληροφορίας.

Στη συνέχεια παρουσιάζεται ο ρόλος του εννοιολογικού πλαισίου κατά την ανάλυση πολυμεσικής πληροφορίας, ενώ προτείνονται τεχνικές μοντελοποίησής του,

<sup>1</sup>Με τον όρο “πολυμεσικό έγγραφο” θα αναφερόμαστε από τούδε και στο εξής σε οποιοδήποτε διαθέσιμο πολυμεσικό υλικό, όπως π.χ. μία φημιακή φωτογραφία, μια ταινία, ένα αρχείο κειμένου, κ.ο.κ.

καθώς και μεθοδολογίες υπολογισμού του. Στο πεδίο της ανάλυσης, το εννοιολογικό πλαίσιο ορίζεται (όπως και στην περίπτωση της προσωποποίησης περιεχομένου) και πάλι με τη βοήθεια οντολογιών και πληροφορίας κατηγοριοποίησης. Τέλος, γίνεται μια εκτενής αναφορά στις εφαρμογές της χρήσης του εννοιολογικού πλαισίου σε τεχνικές και υλοποίησεις κατά τη διαδικασία ανάλυσης πολυμεσικής πληροφορίας και εξετάζονται τα αποτελέσματα της έρευνάς μας σε τομείς όπως η βασισμένη σε γνώση ανάλυση εικόνων, η κατάτμηση εικόνων και η εξαγωγή σημασιολογικών συμπερασμάτων με χρήση ενός γενετικού αλγορίθμου.

- *Εννοιολογικό πλαίσιο ταξινόμησης.* Στο τελευταίο μέρος της διατριβής επεκτείνεται η ερευνητική εργασία μας πάνω στα ενδιάμεσου επιπέδου χαρακτηριστικά γνωρίσματα των εικόνων και στην εκμετάλλευση της αντίστοιχης πληροφορίας. Ειδικότερα, εστιάζουμε στην αναγνώριση υψηλού επιπέδου εννοιών μέσα σε πολυμεσικά έγγραφα με την εισαγωγή καινοτόμων παραστάσεων, όπως είναι οι οντότητες ενδιάμεσου επιπέδου και οι αντίστοιχες οντολογίες τους. Εισάγονται τρία σχετικά είδη, τα οποία επεκτείνουν την ανεπτυγμένη στην προηγούμενη ενότητα κλασική θεώρηση της επίδρασης του εννοιολογικού πλαισίου, τόσο προς την κατεύθυνση των σημασιολογικών εννοιών, όσο και των τύπων περιοχής (region types). Μοντελοποιείται και υλοποιείται η απαιτούμενη γνώση, καθορίζονται οι σχέσεις μεταξύ των νέων ειδών και εξηγείται πώς οι προτεινόμενοι αλγόριθμοι και μεθοδολογίες διαχείρισης του οπτικού εννοιολογικού πλαισίου μπορούν να συνδυαστούν αποδοτικά για να επιτευχθεί ο στόχος της σημασιολογικής πολυμεσικής ανάλυσης. Τέλος, παρουσιάζονται ποικίλα πειραματικά αποτελέσματα από την εν λόγω ενσωμάτωση του εννοιολογικού πλαισίου στη διαδικασία της ανάλυσης, μέσω της εξαγωγής τύπων περιοχών και ομάδων οντοτήτων.



## **ABSTRACT**

The main research objective of this Thesis is to tackle issues related to multimedia content processing, search and retrieval, under the prism of context, as the latter is expressed in the fields of knowledge adaptation and information access. More specifically, the main research motivation was caused by two major research fields: (i) multimedia content personalization and (ii) multimedia content analysis based on visual context. It tackles issues such as data mining, thematic categorization of multimedia documents, multimedia personalization, retrieval and ranking of personalized multimedia documents, knowledge-assisted analysis optimization through visual context exploitation, mid-level visual analysis and context utilization, contextual image classification problems, etc. Towards this direction, it presents research results and indicative implementations/applications, in order to facilitate the proposed interpretation.



# Συντμήσεις και Αρκτικόλεξα

<b>ACM</b>	:	Association for Computing Machinery
<b>CS</b>	:	Classification Scheme
<b>DAML+OIL</b>	:	DARPA Agent Markup Language + Ontology Inference Layer
<b>DDL</b>	:	Data Definition Language
<b>DI</b>	:	Digital Item
<b>DIA</b>	:	Digital Item Adaptation
<b>DL</b>	:	Description Logic
<b>DS</b>	:	Description Scheme
<b>GT</b>	:	Ground Truth
<b>ICT</b>	:	Information and Communications Technology
<b>IEEE</b>	:	Institute of Electrical and Electronics Engineers
<b>IR</b>	:	Information Retrieval
<b>IRS</b>	:	Information Retrieval Systems
<b>KAA</b>	:	Knowledge Assisted Analysis
<b>KB</b>	:	Knowledge Base
<b>kNN</b>	:	k-Nearest Neighbour
<b>LSA</b>	:	Latent Semantic Analysis
<b>LSI</b>	:	Latent Semantic Indexing
<b>MAP</b>	:	Mean Average Precision
<b>MDS</b>	:	Multimedia Description Scheme
<b>MPEG</b>	:	Motion Picture Experts Group
<b>NLP</b>	:	Natural Language Processing
<b>OWL</b>	:	Web Ontology Language
<b>RDF</b>	:	Resource Description Framework
<b>RDFS</b>	:	Resource Description Framework Schema
<b>RF</b>	:	Relevance Feedback
<b>RSST</b>	:	Recursive Shortest Spanning Tree
<b>SEN</b>	:	Special Educational Needs
<b>SVD</b>	:	Singular Value Decomposition
<b>SVM</b>	:	Support Vector Machine
<b>TC</b>	:	Texture Classifier
<b>VCD</b>	:	Visual Content Detectors
<b>W3C</b>	:	World Wide Web Consortium
<b>WWW</b>	:	World Wide Web
<b>XML</b>	:	Extensible Markup Language
<b>ΨΕΕΒΠ</b>	:	Εργαστήριο Ψηφιακής Επεξεργασίας Εικόνας, Βίντεο και Πολυμέσων



# Κατάλογος Ελληνικών 'Ορων

ακρίβεια	:	precision
αναγνώριση προτύπων	:	pattern recognition
ανάκληση	:	recall
ανάκτηση πληροφορίας	:	Information Retrieval
ανάλυση εικόνας	:	image analysis
ανάλυση ιδιαίτουρων τιμών	:	Singular Value Decomposition
ανάλυση κύριων συνιστώσων	:	principal component analysis
αναρτηματοθέτηση	:	tagging
ανεπίβλεπτος/χωρίς επίβλεψη	:	unsupervised
απογραμμικός	:	offline
ασαφές σύνολο	:	fuzzy set
ασαφής σχέση	:	fuzzy relation
ασαφής τροποποιητής	:	modifier / linguistic hedge
ασθενής	:	patient
αφανής	:	implicit
βαθμωτή πληθυκότητα	:	scalar cardinality
βάση γνώσης	:	knowledge base
δεδομένα εκπαίδευσης	:	training data
διάνυσμα αναπαράστασης	:	model vector
διάνυσμα χαρακτηριστικών	:	feature vector
ελεγχόμενο λεξιλόγιο	:	controlled vocabulary
έλλειψη διακριτικότητας	:	obtrusiveness
εμφανής	:	explicit
έννοια	:	concept
εννοιολογικό πλαίσιο	:	context
εντατική οντολογία	:	intensional ontology
εξαγωγή συμπερασμάτων	:	reasoning
εξαγωγή χαρακτηριστικών γνωρισμάτων	:	feature extraction
εξόρυξη δεδομένων	:	data mining
επεκτατική οντολογία	:	extensional ontology
επιβλεπόμενος/με επίβλεψη	:	supervised
επιγραμμικός	:	online
επιλογή χαρακτηριστικών γνωρισμάτων	:	feature selection
ερώτημα αναζήτησης	:	query
ερώτημα με χρήση παραδείγματος	:	query by example
ετικέτα	:	label
ευαισθησία	:	sensitivity
ευρετήριο	:	index

θεματική κατηγοριοποίηση	:	thematic categorization
θεματική περιοχή	:	domain
θησαυρός/λεξικό συνωνύμων	:	thesaurus
κεντρικό στοιχείο	:	centroid
κοινή ερμηνεία	:	common meaning
λανθάνουσα σημασιολογική ανάλυση	:	Latent Semantic Analysis
λανθάνουσα σημασιολογική ευρετηρίαση	:	Latent Semantic Indexing
μεταδεδομένα	:	metadata
μηχανή διανυσμάτων υποστήριξης	:	support vector machine
μηχανική μάθηση	:	machine learning
μονοσύνολο	:	singleton
οντολογία	:	ontology
οπτική λέξη	:	visual word
παρακείμενος όρος	:	surrounding term
πλάνο	:	shot
πλοιήγηση	:	browsing
προσαρμοζόμενα συστήματα υπερμέσων	:	adaptive hypermedia systems
πρότυπο	:	pattern
σάκος λέξεων	:	bag of words
σημασιολογική οντότητα	:	semantic entity
σημασιολογική περιγραφή	:	semantic description
σημασιολογική σχέση	:	semantic relation
συμβολισμός αθροίσματος	:	sum notation
συμπερίληψη	:	inclusion
συστάδα	:	cluster
συσταδοποίηση	:	clustering
σύσταση	:	recommendation
συσχετιστική ανάδραση	:	relevance feedback
Σχέδια Περιγραφής MPEG-7	:	MPEG-7 Description Schemes
σχέδιο περιγραφής πολυμέσων	:	multimedia description scheme
ταξινόμηση	:	classification
ταξινομία	:	taxonomy
τμήμα καμπύλης	:	curve fragment
υπαγωγή	:	subsumption
υπερφόρτωση πληροφοριών	:	information overload
υπόθεση αλήθειας	:	ground truth
υποταυτοδυναμία	:	subidempotency
χαρακτηρισμός	:	annotation
χαρακτηριστικό γνώρισμα	:	feature
χαρακτηριστικό καρέ	:	keyframe
χώρος εννοιών	:	concept space

Σημ.: μετάφραση όρων σύμφωνα με τη Βάση Τηλεπικοινωνιακών 'Ορων του ΕΛΟΤ, Ιανουάριος 2008, <http://www.moto-teleterm.gr/>, 120.397 λήμματα.

# Κατάλογος Αγγλικών 'Ορων

<b>adaptive hypermedia systems</b>	:	προσαρμοζόμενα συστήματα υπερμέσων
<b>annotation</b>	:	χαρακτηρισμός
<b>bag of words</b>	:	σάκος λέξεων
<b>browsing</b>	:	πλοήγηση
<b>centroid</b>	:	κεντρικό στοιχείο
<b>classification</b>	:	ταξινόμηση
<b>cluster</b>	:	συστάδα
<b>clustering</b>	:	συσταδοποίηση
<b>common meaning</b>	:	χοινή ερμηνεία
<b>concept</b>	:	έννοια
<b>concept space</b>	:	χώρος εννοιών
<b>context</b>	:	εννοιολογικό πλαίσιο, πλαίσιο γνώσης
<b>controlled vocabulary</b>	:	ελεγχόμενο λεξιλόγιο
<b>curve fragment</b>	:	τμήμα καμπύλης
<b>data mining</b>	:	εξόρυξη δεδομένων
<b>domain</b>	:	θεματική περιοχή
<b>explicit</b>	:	εμφανής
<b>extensional ontology</b>	:	επεκτατική οντολογία
<b>feature</b>	:	μετρήσιμο χαρακτηριστικό γνώρισμα
<b>feature extraction</b>	:	εξαγωγή χαρακτηριστικών γνωρισμάτων
<b>feature selection</b>	:	επιλογή χαρακτηριστικών γνωρισμάτων
<b>feature vector</b>	:	διάνυσμα χαρακτηριστικών
<b>ground truth</b>	:	υπόθεση αληθείας
<b>image analysis</b>	:	ανάλυση εικόνας
<b>implicit</b>	:	αφανής
<b>inclusion</b>	:	συμπερίληψη
<b>index</b>	:	ευρετήριο
<b>information overload</b>	:	υπερφόρτωση πληροφοριών
<b>Information Retrieval</b>	:	ανάκτηση πληροφορίας
<b>intensional ontology</b>	:	εντατική οντολογία
<b>keyframe</b>	:	χαρακτηριστικό καρέ
<b>knowledge base</b>	:	βάση γνώσης
<b>label</b>	:	ετικέτα
<b>Latent Semantic Indexing</b>	:	λανθάνουσα σημασιολογική ευρετηρίαση
<b>machine learning</b>	:	μηχανική μάθηση
<b>metadata</b>	:	μεταδεδομένα
<b>model vector</b>	:	διάνυσμα αναπαράστασης
<b>modifier/linguistic hedge</b>	:	ασαφής τροποποιητής

<b>MPEG-7 Description Schemes</b>	:	MPEG-7 Σχέδια Περιγραφής
<b>multimedia description scheme</b>	:	σχέδιο περιγραφής πολυμέσων
<b>obtrusiveness</b>	:	έλλειψη διαχριτικότητας
<b>offline</b>	:	απογραμμικός
<b>online</b>	:	επιγραμμικός
<b>ontology</b>	:	οντολογία
<b>patient</b>	:	ασθενής
<b>pattern</b>	:	πρότυπο
<b>pattern recognition</b>	:	αναγνώριση προτύπων
<b>precision</b>	:	ακρίβεια
<b>principal component analysis</b>	:	ανάλυση κύριων συνιστωσών
<b>query</b>	:	ερώτημα αναζήτησης
<b>query by example</b>	:	ερώτημα με χρήση παραδείγματος
<b>reasoning</b>	:	εξαγωγή συμπερασμάτων
<b>recall</b>	:	ανάκληση
<b>recommendation</b>	:	σύσταση
<b>relevance feedback</b>	:	συσχετιστική ανάδραση
<b>scalar cardinality</b>	:	βαθμωτή πληθικότητα
<b>semantic description</b>	:	σημασιολογική περιγραφή
<b>semantic entity</b>	:	σημασιολογική οντότητα
<b>semantic relation</b>	:	σημασιολογική σχέση
<b>sensitivity</b>	:	ευαισθησία
<b>shot</b>	:	πλάνο
<b>singleton</b>	:	μονοσύνολο
<b>singular value decomposition</b>	:	ανάλυση ιδιαζουσών τιμών
<b>subidempotency</b>	:	υποταυτοδυναμία
<b>subsumption</b>	:	υπαγωγή
<b>sum notation</b>	:	συμβολισμός αθροίσματος
<b>supervised</b>	:	εποπτευμένος / με επίβλεψη
<b>support vector machine</b>	:	μηχανή διανυσμάτων υποστήριξης
<b>surrounding term</b>	:	παρακείμενος όρος
<b>tagging</b>	:	αναρτηματοθέτηση
<b>taxonomy</b>	:	ταξινομία
<b>thematic categorization</b>	:	θεματική κατηγοριοποίηση
<b>thesaurus</b>	:	θησαυρός / λεξικό συνανύμων
<b>training data</b>	:	δεδομένα εκπαίδευσης
<b>unsupervised</b>	:	ανεπίβλεπτος / χωρίς επίβλεψη
<b>visual word</b>	:	οπτική λέξη

Σημ.: μετάφραση όρων σύμφωνα με τη Βάση Τηλεπικοινωνιακών 'Ορων του ΕΛΟΤ, Ιανουάριος 2008, <http://www.moto-teleterm.gr/>, 120.397 λήμματα.





# Κεφάλαιο 1

## Εισαγωγή

Η παρούσα εργασία επιχειρεί να αντιμετωπίσει τα θέματα της επεξεργασίας, αναζήτησης και ανάκτησης πολυμεσικής πληροφορίας μέσα από το πρόσμα του εννοιολογικού πλαισίου, όπως αυτό εκφράζεται στο πεδίο της προσαρμογής της γνώσης και της πρόσβασης στην πληροφορία. Ειδικότερα, δύο είναι οι κεντρικοί άξονες που αποτέλεσαν το κίνητρο για την έρευνα της συγκεκριμένης θεματικής περιοχής: (i) η προσωποποίηση πολυμεσικού περιεχομένου και (ii) η ανάλυση πολυμεσικού περιεχομένου με βάση το - ποικίλως εννοούμενο - εννοιολογικό πλαίσιο.

Αναμφίβολα, τόσο η προσωποποίηση πολυμεσικού περιεχομένου όσο και η ανάλυση πολυμεσικού περιεχομένου με βάση το εννοιολογικό πλαίσιο έχουν άμεση σχέση με το χρήστη, προσδίδοντας μία ανθρωποκεντρική ματιά στην έρευνά μας. Η πρόσβαση στην πολυμεσική πληροφορία και η κατάλληλη προσαρμογή της γνώσης αποτελούν το κύριο αντικείμενο της διδακτορικής διατριβής και το αντικείμενο της επιχειρούμενης εμβάθυνσης. Πρόκειται, ουσιαστικά, για τις δύο όψεις του ίδιου νομίσματος, της αλληλεπίδρασης (interaction), δηλαδή, του χρήστη με το εκάστοτε σύστημα αναζήτησης ή ανάκτησης πληροφορίας. Εκείνο που διαφοροποιείται σε κάθε περίπτωση είναι η έμφαση η οποία δίνεται: από τη σκοπιά της προσωποποίησης πολυμεσικού περιεχομένου, στο επίκεντρο βρίσκεται ο χρήστης, ενώ από την σκοπιά της μοντελοποίησης, ανάλυσης και διαχείρισης του εννοιολογικού πλαισίου, βρίσκεται η βούληση του χρήστη, με την μορφή των αντικειμένων της αναζήτησής του.

Κατά συνέπεια και χωρίς βλάβη της γενικότητας, μπορούμε να ισχυριστούμε ότι ο συνδετικός χρίκος των παραπάνω εννοιών είναι το εννοιολογικό πλαίσιο. Κι αυτό γιατί η προσωποποίηση του πολυμεσικού περιεχομένου μπορεί να θεωρηθεί ως το ίδιο το εννοιολογικό πλαίσιο της ανάκτησής του, καθώς οι τελικοί χρήστες ενός συστήματος ανάκτησης πολυμεσικής πληροφορίας ικανοποιούν τις ανάγκες πληροφόρησής τους. Σύμφωνα με αυτή τη θεώρηση, η προσωποποίηση πολυμεσικού περιεχομένου αποτελεί μία από τις ποικιλόμορφες εκφάνσεις του εννοιολογικού πλαισίου. Όμως, προς μεγάλη χαρά του δραστήριου και ανήσυχου ερευνητή, αλλά και όπως καταδεικνύεται στο πλαίσιο της παρούσης διατριβής, το τελευταίο εκδηλώνεται με μια τέτοια ποικιλία πιθανών μορφών, που η έρευνα του χώρου δεν είναι σε θέση να αντιμετωπίσει με έναν καθολικό τρόπο, παρά μόνο αποσπασματικά, εμβαθύνοντας ταυτόχρονα στα κατά τόπους ανακύπτοντα προβλήματα. Συνακόλουθα, όπως ίσως είναι αναμενόμενο στον αναγνώστη από τον ίδιο τον τίτλο της, η παρούσα θεώρηση άπτεται πολλών ερευνητικών περιοχών και προβλημάτων, όπως λ.χ. είναι η εξόρυξη δεδομένων, η αναζήτηση πληροφορίας, η αναπαράσταση γνώσης, ο ορισμός, η μοντελοποίηση και η ανάλυση του εννοιολογικού

πλαισίου, η υποβοηθούμενη από τη γνώση ανάλυση εικόνας, η προσωποποίηση πολυμεσικού περιεχομένου, η ανάκτηση και κατάταξη προσωποποιημένων αποτελεσμάτων αναζήτησης, κ.α..

Συνοψίζοντας, το κεφάλαιο 2 που ακολουθεί είναι αφιερωμένο στο γενικότερο ερευνητικό πλαίσιο της εργασίας, ενώ το κεφάλαιο 3 παρουσιάζει συνοπτικά το μαθηματικό υπόβαθρο, τους όρους και τις βασικές έννοιες πάνω στις οποίες κινείται η ερευνητική δραστηριότητά της. Το κεφάλαιο 4 περιγράφει τις βασικές αρχές και προτεινόμενες επεκτάσεις των ερευνητικών θεμάτων εξόρυξης δεδομένων και τεχνικών μηχανικής μάθησης που άπτονται άμεσα των κεφαλαίων που έπονται και την προτεινόμενη θεώρηση της κατηγοριοποίησης πολυμεσικών εγγράφων. Το κεφάλαιο 5 παρουσιάζει πρωτότυπη εργασία, έρευνα και αποτελέσματα σε θέματα προσωποποίησης πολυμεσικού περιεχομένου, ενώ στο κεφάλαιο 6 παρουσιάζεται η συνεισφορά του οπτικού εννοιολογικού πλαισίου στην αντιμετώπιση προβλημάτων ανάλυσης πολυμεσικού περιεχομένου (γενικά) και επεξεργασίας εικόνας (ειδικότερα). Στο κεφάλαιο 7 επεκτείνεται η παραπάνω θεώρηση, εισάγοντας μια καινοτόμο θεώρηση πολυμεσικής ανάλυσης και επεξεργασίας περιεχομένου, η οποία βασίζεται σε ενδιάμεσου επιπέδου χαρακτηριστικά γνωρίσματα και το αντίστοιχο εννοιολογικό πλαισιο. Το κεφάλαιο 8 αναφέρει με συντομία τις βασικές τρέχουσες τάσεις τις έρευνας στα πεδία που σχετίζονται με την παρούσα διδακτορική διατριβή και περιγράφει κάποιες πιθανές μελλοντικές επεκτάσεις της, ενώ στο κεφάλαιο 9 περιλαμβάνεται ο κατάλογος των σχετικών ερευνητικών δημοσιεύσεων.

## 1.1 Γενικό Ερευνητικό Αντικείμενο

Το γενικό ερευνητικό αντικείμενο της ερευνητικής εργασίας μας είναι η αποτελέσματική και ευφυής αντιμετώπιση των προβλημάτων ανάλυσης, αναζήτησης και ανάκτησης πολυμεσικού περιεχομένου. Μεταξύ διαφόρων πιθανών ερευνητικών διαδρομών, ακολουθήσαμε τους πιο ανθρωποκεντρικούς από αυτούς. Όπως συμβαίνει σε μία πληθώρα επιστημών και ερευνητικών προσεγγίσεων, ο χρήστης-άνθρωπος και η αλληλεπίδρασή του με το πολυμεσικό περιεχόμενο αποτελεί το κεντρικό στοιχείο της έρευνάς μας. Θεωρώντας την πρόσβαση στην πολυμεσική πληροφορία και την κατάλληλη προσαρμογή της γνώσης ως τις δύο όψεις του ίδιου νομίσματος, από τη μία βρίσκεται η βούληση του χρήστη με την μορφή των προτιμήσεων του επί των αντικειμένων της αναζήτησής του και από την άλλη η (προ-)επεξεργασία του περιεχομένου για την αποδοτικότερη ως προς τον χρήστη του αναζήτησή του. Έτσι, αντιμετωπίσαμε θέματα που άπτονται των ερευνητικών περιοχών της προσωποποίησης πολυμεσικού περιεχομένου, καθώς και της μοντελοποίησης, ανάλυσης και διαχείρισης του εννοιολογικού πλαισίου στο ίδιο το περιεχόμενο, ως το μέσο για την αποτελεσματικότερη διαχείρισή του.

## 1.2 Γενικά Ερευνητικά Προβλήματα

Τα ερευνητικά προβλήματα που καλείται να αντιμετωπίσει η παρούσα διατριβή ποικίλουν τόσο σε είδος όσο και σε πλήθος. Αρχικά, το πρόβλημα της θεματικής κατηγοριοποίησης πολυμεσικών εγγράφων απαντάται ως επί το πλείστον στο σύνολο των σύγχρονων συστημάτων ανάκτησης πληροφορίας (IRS). Ο ερευνητής καλείται να προτείνει ικανοποιητικές λύσεις στο πεδίο της (ημι-)αυτόματης εξαγωγής πληροφορίας

από τα πολυμεσικά έγγραφα. Η διαδικασία αποτελεί βασικό εργαλείο για την αντιμετώπιση του σημαντικού προβλήματος της υπερφόρτωσης πληροφοριών (information overload), ενώ διαμορφώνει ένα ιδιαίτερα σημαντικό και αποτελεσματικό από πλευράς κόστους βήμα προς την κατεύθυνση της αποδοτικής διαχείρισης και κυρίως κατανόησης του πολυμεσικού περιεχομένου. Παραμένει, όμως, γεγονός, ότι η εξαγωγή σημασιολογικής πληροφορίας μέσω της απόκτησης γνώσης και της αναγνώρισης υψηλού επιπέδου σημασιολογικών χαρακτηριστικών παραμένει ένα ανοιχτό και ενδιαφέρον ερευνητικό πρόβλημα. Τα χαρακτηριστικά προβλήματα που απαντώνται συνήθως στα συστήματα ανάκτησης κειμενικής πληροφορίας, όπως είναι η υπερφόρτωση και η διαφορετικότητα πληροφορίας [214], εμφανίζονται και στην περίπτωση της πολυμεσικής ανάκτησης [34]. Οι περισσότερες σύγχρονες προσεγγίσεις τείνουν να συνδυάζουν την πολυμεσική εξαγωγή και την ανάπτυξη της γνώσης σε ένα κοινό πλαίσιο. Η ανάπτυξη της κατάλληλης γνώσης είναι ιδιαίτερα σημαντική και εντείνει την ευρωστία της διαδικασίας εξαγωγής, ενώ την ίδια στιγμή η διαρκής εξαγωγή σημασιολογικής πληροφορίας από το περιεχόμενο, αξιοποιείται για τον περαιτέρω εμπλουτισμό της γνώσης. Κατά συνέπεια, υβριδικές προσεγγίσεις που αποτελούνται τόσο από τεχνικές κατηγοριοποίησης κειμένων, όσο και από κατάλληλες μηχανές εξαγωγής συμπερασμάτων, είναι ιδιαίτερα ευπρόσδεκτες.

Το γενικότερο ερευνητικό πρόβλημα της προσωποποίησης ή αλλιώς της εξατομίκευσης πολυμεσικού περιεχομένου αποτελεί τα τελευταία χρόνια ένα δημοφιλές πεδίο έρευνας και ανάπτυξης. Ο απώτερος στόχος της προσωποποίησης είναι να καταστεί δυνατή η προσφορά στον τελικό χρήστη πολυμεσικού περιεχομένου, το οποίο να βρίσκεται πολύ κοντά στις επιθυμίες του. Εν γένει, η προσωποποίηση των πολυμέσων αποσκοπεί στο να προσαρμόσει το φηφιακό οπτικοακουστικό περιεχόμενο στους εκάστοτε χρήστες του, με βάση προσωπικές λεπτομέρειες ή χαρακτηριστικά που παρέχουν οι τελευταίοι. Επιτρέπει σε έναν προμηθευτή περιεχομένου να προσαρμόσει συγκεκριμένο πολυμεσικό περιεχόμενο σύμφωνα με τα ατομικά πρότυπα και τις προτιμήσεις των χρηστών του. Θα λέγαμε ότι αποτελεί μια τεχνολογία που επιτρέπει σε ένα σύστημα να ταιριάζει το διαθέσιμο πολυμεσικό περιεχόμενο, τις εφαρμογές και τις διαθέσιμες μορφές ανθρώπινης αλληλεπίδρασης με αυτό, τόσο με στατικές και ρητώς δεδηλωμένες, όσο και με δυναμικές, προτιμήσεις ενός χρήστη.

Οι προτιμήσεις και κατ' επέκταση τα προφίλ των χρηστών είναι δυνατόν να μελετηθούν και να αναγνωρισθούν με την ανάπτυξη κατάλληλων αλγορίθμων εξαγωγής και διαχείρισης. Ειδικότερα, κάτι τέτοιο είναι ερευνητικά εφικτό, είτε μέσω της άμεσης παροχής πληροφοριών από τους χρήστες για τα ενδιαφέροντά τους, είτε μέσω της εκπαίδευσης ενός συστήματος με την πάροδο του χρόνου και καθώς ο χρήστης αλληλεπιδρά με αυτό. Αναμφίβολα, η συλλογή και αναπαράσταση των προτιμήσεων των χρηστών αποτελεί από μόνη της ένα μεγάλο πεδίο έρευνας [97], γνωστό και ως μοντελοποίηση χρηστών (user modeling). Και πάλι, ιδιαίτερη σημασία πρέπει να δωθεί στην χρησιμοποιούμενη γνώση. Καθώς τα προφίλ και τα ενδιαφέροντα των χρηστών αποτελούν ένα από τα χαρακτηριστικότερα παραδείγματα μεγεθών που μετά βίᾳς τοποθετούνται κάτω από μία σαφή και ευδιάκριτη έννοια, γίνεται κατανοητό ότι ο προσεκτικός σχεδιασμός και εκμετάλλευση της γνώσης είναι εκ των ουκ άνευ για την αποτελεσματικότητα της οποιασδήποτε ερευνητικής μεθόδου. Η χρονική εξάρτηση, οι δυναμικοί βαθμοί συσχέτισης, η αβεβαιότητα και το ποικιλόμορφο εννοιολογικό πλαίσιο που συνοδεύει τις εκάστοτε προτιμήσεις, τις καθιστούν μια ιδιαίτερα σύνθετη έννοια. Η σημασιολογία του περιεχομένου παίζει σημαντικό ρόλο κατά τη διαδικασία της προσωποποίησης, ιδιαίτερα, δε, στην περίπτωση που ερευνώνται αυτόματες μέθοδοι

εξαγωγής και χρήσης. Προκειμένου να περιγραφεί η διαδικασία με έναν ολοκληρωμένο τρόπο, απαιτείται η εκμετάλλευση του εννοιολογικού πλαισίου, ούτως ώστε αυτή να λάβει την καλύτερη δυνατή ερμηνεία.

Όσον αφορά στην ανάλυση του πολυμεσικού περιεχομένου, η αξιοποίηση της έννοιας του οπτικού εννοιολογικού πλαισίου ως της πληροφορίας εκείνης που είναι σχετική με το οπτικό περιεχόμενο της σκηνής μιας εικόνας, αποτελεί από μόνη της ένα εκτεταμένο ερευνητικό πεδίο. Αυτή η οπτική πληροφορία που ενυπάρχει εγγενώς στο πολυμεσικό περιεχόμενο μπορεί να αποβεί ιδιαίτερα χρήσιμη για την αντιμετώπιση κλασικών ερευνητικών προβλημάτων της επεξεργασίας εικόνας, όπως είναι η υποβοηθούμενη από τη γνώση χαμηλού επιπέδου ανάλυση εικόνων, η σημασιολογική ή κλασική κατάταξη εικόνων και η εξαγωγή σημασιολογικών συμπερασμάτων για το πολυμεσικό περιεχόμενο. Αυτή η μορφή του εννοιολογικού πλαισίου μπορεί να λάβει ποικίλες διαφορετικές υποστάσεις. Σε κάθε περίπτωση, πάντως, θα πρέπει να βασισθεί σε κάποιας μορφής γνώση, η οποία θα πρέπει να μοντελοποιηθεί και να αποθηκευθεί κατάλληλα. Ως εκ τούτου, είναι γεγονός ότι πλέον οι περισσότεροι σημερινοί ερευνητές μελετούν και υψηλοτέρου επιπέδου χαρακτηριστικά γνώσης, εκτός από αυτά του χαμηλού επιπέδου (π.χ. χρώμα, υφή, σχήμα, κ.α.) αναζητώντας αποτελεσματικότερες μεθόδους έρευνας. Το εννοιολογικό πλαίσιο αποτελεί έναν από τους πιο σημαντικούς παράγοντες για την ξεκάθαρη διαδικασία ανάλυσης και επεξεργασίας του πολυμεσικού περιεχομένου, καθώς συνήθως αναφέρεται στις σημασιολογικές, τοπολογικές ή και χρονικές σχέσεις που διέπουν τα διαφορετικά αντικείμενα στη σκηνή μιας εικόνας και εκμεταλλεύεται την σημασιολογία τους. Σε αυτό το πλαίσιο, ο προσδιορισμός υψηλού επιπέδου εννοιών μέσα στα πολυμεσικά έγγραφα μέσω της εισαγωγής ενδιάμεσου επιπέδου χαρακτηριστικών γνωρισμάτων και της εκμετάλλευσης της αντίστοιχης πληροφορίας εννοιολογικού πλαισίου συνθέτει μία ζεχωριστή, αυτοδύναμη προσέγγιση αναπαράστασης ενδιάμεσου επιπέδου. Η τελευταία απαιτεί τη μοντελοποίηση, σχεδίαση και κατασκευή ζεχωριστής γνώσης, ενώ η εισαγωγή, αξιοποίηση και επέκταση των σχετικών αλγορίθμων διαχείρισης του οπτικού εννοιολογικού πλαισίου αποτελεί αυτοδύναμο ερευνητικό πρόβλημα και ζεχωριστό πεδίο έρευνας στο χώρο.

### 1.3 Επιμέρους Αντικείμενο

Το επιμέρους ερευνητικό αντικείμενο της παρούσης διδακτορικής διατριβής είναι η προσαρμογή των παραπάνω διαδικασιών ανάλυσης, αναζήτησης και ανάκτησης πολυμεσικού περιεχομένου στο εκάστοτε εννοιολογικό πλαίσιο της πρόσβασης στο περιεχόμενο. Ιδιαίτερη έμφαση δίνεται στο εννοιολογικό πλαίσιο της ανάκτησης πολυμεσική πληροφορίας με τη μορφή της προσωποποιημένης πρόσβασης σε αυτή. Προτείνονται ευφυές και καινοτόμες τεχνικές για την κατάλληλη προσαρμογή του περιεχομένου ανάλογα με τις προτιμήσεις και τα προφίλ των χρηστών του, τα οποία εξάγονται από το εκάστοτε διαθέσιμο ιστορικό πρόσβασης σε αυτό. Επίσης, σημαντικό ρόλο παίζει η αποδοτική ανάλυση του πολυμεσικού περιεχομένου με βάση το οπτικό εννοιολογικό πλαίσιο. Εδώ η προσέγγιση που ακολουθείται περιλαμβάνει ποικίλες προσεγγίσεις, καθεμία από τις οποίες είναι κατάλληλη για την αντιμετώπιση συγκεκριμένου είδους προβλημάτων ανάλυσης, όπως η υποβοηθούμενη από τη γνώση ανάλυση εικόνων (π.χ. η σημασιολογική κατάταξη τους), η εξαγωγή σημασιολογικών συμπερασμάτων με χρήση γενετικών αλγορίθμων και η αξιοποίηση ενδιάμεσου επιπέδου χαρακτηριστικών γνωρισμάτων για την ανίχνευση υψηλού επιπέδου εννοιών και ταξινόμηση σκηνών. Προκειμένου, δε, να πραγματοποιηθεί μια εμπεριστατωμένη

και όσο το δυνατόν πλήρης παρουσίαση των ερευνητικών πεπραγμένων, παρουσιάζονται τόσο θεωρητικά αποτελέσματα, όσο και εφαρμογές.

## 1.4 Ερευνητική Ενασχόληση

Όπως καταδεικνύει και ο τίτλος της παρούσης διατριβής, η ερευνητική ενασχόλησή μας κινείται γύρω από την προσαρμοζόμενη πρόσβαση του χρήστη στο πολυμεσικό περιεχόμενο, με βάση προσωπικά χαρακτηριστικά του, όπως το ποικίλως εννοούμενο εννοιολογικό πλαίσιο και οι προτιμήσεις/προφίλ του. Τα παραπάνω συνοψίζονται στις ακόλουθες τέσσερις βασικές θεματικές περιοχές: (i) τη θεματική κατηγοριοποίηση πολυμεσικών εγγράφων, (ii) την προσωποποίηση του περιεχομένου από τη σκοπιά των προτιμήσεων του χρήστη, (iii) την επίδραση του οπτικού εννοιολογικού πλαισίου στα πρόβληματα ανάλυσης εικόνας και (iv) την θεώρηση των προηγουμένων κάτω από ένα νέο πρίσμα “ενδιάμεσου επιπέδου” επεξεργασίας και ανάλυσης, το οποίο συνδυάζει χαμηλού και υψηλού επιπέδου στοιχεία.

### 1.4.1 Θεματική κατηγοριοποίηση εγγράφων

Η θεματική κατηγοριοποίηση πολυμεσικών εγγράφων περιλαμβάνει τα βήματα της επιλογής των αντιπροσωπευτικών εγγράφων κάθε θέματος, τον προσδιορισμό μέσα στα επιλεγμένα έγγραφα, στοιχείων που είναι χαρακτηριστικά για κάθε θέμα, τον καθορισμό ενός ή περισσοτέρων συντελεστών για κάθε έγγραφο ενός αντιπροσωπευτικού συντελεστή σχετικότητας του εκάστοτε στοιχείου με το αντίστοιχο θέμα και την τελική εξαγωγή χαρακτηριστικών θεμάτων για κάθε υπό εξέταση έγγραφο. Τα θέματα αυτά χαρακτηρίζουν σημασιολογικά το εκάστοτε έγγραφο και το κατατάσσουν στην κατάλληλη κατηγορία.

#### 1.4.1.1 Επισκόπηση ερευνητικής περιοχής

Είναι γεγονός, ότι η θεματική κατηγοριοποίηση των πολυμεσικών εγγράφων αποτελεί ένα ιδιαίτερα σημαντικό και αποτελεσματικό από πλευράς κόστους βήμα για την αποδοτική διαχείριση και κατανόηση του πολυμεσικού περιεχομένου. Τα τελευταία χρόνια πολλές ερευνητικές προσπάθειες προσπάθησαν να αντικετωπίσουν το γενικότερο πρόβλημα της αναγνώρισης του πολυμεσικού περιεχομένου που επιθυμεί ο χρήστης. Σε αυτό το πλαίσιο, η εξαγωγή σημασιολογικής πληροφορίας μέσω της απόκτησης γνώσης και της αναγνώρισης υψηλού επιπέδου σημασιολογικών χαρακτηριστικών παραμένει ένα άκρως ενδιαφέρον ερευνητικό πρόβλημα, το οποίο εξακολουθεί να είναι διάχυτο στα σύγχρονα συστήματα ανάκτησης πληροφορίας.

Τυπικά προβλήματα των συστημάτων αυτών είναι η υπερφόρτωση και η διαφορετικότητα της πληροφορίας [214], [34]. Οι σύγχρονες προσεγγίσεις τείνουν να συνδυάζουν την πολυμεσική με την κειμενική πληροφορία. Μια επισκόπηση των ερευνητικών προσπαθειών που εξετάζουν την ταξινόμηση εγγράφων κειμένου παρουσιάζεται στη δημοσίευση [222]. Καινούρια στατιστικά μοντέλα για την ταξινόμηση δομημένων εγγράφων πολυμέσων παρουσιάζονται στην [77], ενώ ποικίλοι αλγόριθμοι συσταδοποίησης έχουν προταθεί για την συσταδοποίηση κειμένων στη βιβλιογραφία [251].

Σε αντιδιαστολή με το απλό κείμενο οι προσεγγίσεις ταξινόμησης εγγράφων πλούσιου πολυμεσικού περιεχομένου έρχονται αντιμέτωπες με επιπρόσθετες ερευνητικές

προκλήσεις, κυρίως λόγω της φύσης του ίδιου του περιεχομένου που εξετάζουν. Διάφοροι αλγόριθμοι έχουν προταθεί κατά καιρούς, τόσο για την περίπτωση επιβλεπόμενης, όσο και ανεπίβλεπτης κατηγοριοποίησης εγγράφων. Π.χ., ο αλγόριθμος ασαφών c-μέσων (fuzzy c-means) [25] αποτελεί σημείο αναφοράς για την περίπτωση της επιβλεπόμενης κατηγοριοποίησης πολυμεσικών εγγράφων, ενώ τεχνικές προβολής [220], αλλά και η συσταδοποίηση k-μέσων [211] προτείνονται για να επιταχυνθεί ο υπολογισμός των αποστάσεων της απαιτούμενης συσταδοποίησης. Ένας αρκετά διαδεδομένος τρόπος για να επιτευχθεί η επιθυμητή αντιστοίχηση εγγράφων σε όρους είναι και αυτός που χρησιμοποιεί διάφορες παραλλαγές της Λανθάνουσας Σημασιολογικής Ανάλυσης (LSA) [147], [74], τον οποίο και θα εξετάσουμε σε αντιδιαστολή με την δική μας προσέγγιση στο κεφάλαιο 4.

#### 1.4.1.2 Ερευνητική συνεισφορά

Η ερευνητική συνεισφορά της ενότητας αυτής εστιάζει στο πεδίο της επεξεργασίας και ανάκτησης πολυμεσικού περιεχομένου από τη σκοπιά της κατηγοριοποίησης πολυμεσικών εγγράφων σε σημασιολογικές κατηγορίες, λαμβάνοντας υπόψη το σχετικό εννοιολογικό πλαίσιο και ως εκ τούτου, λαμβάνει περισσότερες της μίας μορφές. Αρχικά, επεκτείναμε κάποιες από τις βασικές αρχές και αλγορίθμους συσταδοποίησης και κατηγοριοποίησης δεδομένων και προτείναμε καινοτόμες προσθήκες και αλλαγές σε αυτούς. Ειδικότερα παρουσιάσαμε μια τεχνική μείωσης των διαστάσεων των χαρακτηριστικών γνωρισμάτων κατά τη διαδικασία της συσταδοποίησης, η οποία οδηγεί στη βελτιστοποίηση της απόδοσης και της επίδοσή της. Η πρότασή μας αποφαίνεται ιδιαίτερα χρήσιμη όταν οι χωρικές διαστάσεις των δεδομένων εισαγωγής (input space dimensions) είναι μεγάλες και οι ακλίμακες των χαρακτηριστικών γνωρισμάτων είναι ποικιλόμορφες. Στη συνέχεια, εισαγάγαμε την έννοια της ασαφούς ταξινομικής συσταδοποίησης, παρουσιάσαμε αναλυτικά την μεθοδολογία της και προτείναμε μία επιπρόσθετη μεθοδολογία ελέγχου και αποτίμησης του σχετικού αλγορίθμου, συνδυάζοντας την με την ευρέως διαδεδομένη τεχνική κατηγοριοποίησης του k-κοντινότερου γείτονα (kNN). Ειδικότερα, αναλύσαμε μία χωρίς επιβλεψη (unsupervised) επέκταση της ιεραρχικής συσταδοποίησης, με την έννοια της επιλογής χαρακτηριστικών γνωρισμάτων, με την επιπρόσθετη χρήση μιας δεύτερης τεχνικής κατηγοριοποίησης ύστερα από την πρώτη φάση. Η κύρια συνεισφορά της προτεινόμενης επέκτασης είναι η αντιμετώπιση των κλασικών προβλημάτων της ιεραρχικής συσταδοποίησης, όπως είναι η κατάρα της διαστατικότητας (curse of dimensionality) και η αρχική διάδοση λάθους (initial error propagation), αυξάνοντας με τον τρόπο αυτό την ευρωστία της τεχνικής. Στο πλαίσιο της προτεινόμενης ασαφούς ταξινομικής συσταδοποίησης, ορίσαμε, μοντελοποιήσαμε και αξιοποιήσαμε την έννοια του ταξινομικού εννοιολογικού πλαισίου (taxonomic context), αξιοποιώντας σε αλγορίθμικό επίπεδο την κοινή έννοια των στοιχείων, προκειμένου να καθοριστούν καλύτερα οι συστάδες που προκύπτουν από το σύνολο των δεδομένων. Τελικά, εφαρμόσαμε τις παραπάνω τεχνικές και μεθοδολογίες για την αποτελεσματική και αυτόματη, ασαφή θεματική κατηγοριοποίηση πολυμεσικών εγγράφων. Στη διαδικασία, εισαγάγαμε και εξελίξαμε μεθοδολογίες ανάλυσης εγγράφων, δημιουργίας σημασιολογικού ευρετηρίου και αντιστοίχησης ενός πολυμεσικού εγγράφου σε μία ή περισσότερες κατηγορίες, ανάλογα με τα περιεχόμενά του. Ως επιστέγασμα της σχετικής έρευνάς μας, παρουσιάζουμε μια περαιτέρω βελτιωμένη εκδοχή της αρχικής πρότασής μας, η οποία αναλύει και εκμεταλλεύεται τον χαρακτηρισμό (με τη μορφή κειμένου) που συνοδεύει ένα πολυμεσικό έγγραφο, προκειμένου να εξάγει την σημασιολογία του, να κατασκευάσει ένα σημασιολογικό

ευρετήριο και τελικά να κατηγοριοποιήσει τα έγγραφα σε θεματικές κατηγορίες. Το ευρετήριο αυτό βασίζεται σε ένα πρωτότυπο και ενοποιημένο μοντέλο αναπαράστασης γνώσης και σημασιολογίας που προτείνεται στο πλαίσιο της διατριβής αυτής, το οποίο περιλαμβάνει έναν συνδυασμό χαρακτηριστικών από την ασαφή επέκταση της εκφραστικής γλώσσας περιγραφής λογικής *SHIN*, την f-*SHIN*, και ορισμένων βασικών υπηρεσιών εξαγωγής συμπερασμάτων από το πολυμεσικό περιεχόμενο (με τη μορφή κανόνων συμπερασμού - inference rules). Η τελευταία αυτή προσέγγιση αξιοποιείται για την περαιτέρω εκλέπτυνση και βελτιστοποίηση των αρχικών αποτελεσμάτων της θεματικής κατηγοριοποίησης.

## 1.4.2 Θέματα προσωποποίησης

Η προσωποποίηση του πολυμεσικού περιεχομένου και κατ' επέκταση της ίδιας της πρόσβασης σε αυτό, περιλαμβάνει ευφυείς και καινοτόμες τεχνικές προσαρμογής του περιεχομένου, ανάλογα με τις προτιμήσεις και τα προφίλ των χρηστών του. Τα τελευταία εξάγονται με βάση το εκάστοτε διαθέσιμο ιστορικό πρόσβασης στο περιεχόμενο, τις διαθέσιμες πράξεις των χρηστών και τις κατάλληλες μορφές του εννοιολογικού πλαισίου που ενυπάρχουν στη διαδικασία.

### 1.4.2.1 Επισκόπηση ερευνητικής περιοχής

Όντας ένα δημοφιλές πεδίο έρευνας και ανάπτυξης, η προσωποποίηση πολυμεσικού περιεχομένου έχει λάβει ποικίλες ερευνητικές εκφράσεις τα τελευταία χρόνια. Ο απώτερος στόχος της προσωποποίησης είναι να καταστεί δυνατή η προσφορά στον τελικό χρήστη πολυμεσικού περιεχομένου, το οποίο να βρίσκεται πολύ κοντά στις ανάγκες και τις επιθυμίες του. Η μορφή και ο ρόλος του εννοιολογικού πλαισίου στη διαδικασία είναι καθοριστικός προκειμένου αυτό να ενταχθεί ομαλά στην προσωποποίηση της ανάκτησης πληροφορίας πολυμεσικού περιεχομένου. Η προσωποποίηση επιδιώκει να βελτιώσει την υποκειμενική απόδοση της ανάκτησης αυτής, και διάφορες ερευνητικές προτάσεις έχουν γίνει κατά καιρούς ([55], [96], [131], [159], [172]).

Ο στόχος της βελτίωσης των μοντέλων ανάκτησης πληροφορίας προς την κατεύθυνση μοντέλων που λαμβάνουν υπόψη το εννοιολογικό πλαίσιο (context-aware models) έχει προκαλέσει τελευταίο το ολοένα και αυξανόμενο ενδιαφέρον της ερευνητικής κοινότητας. Προσδιορίζεται, δε, ως ένα βασικό βήμα προκειμένου να αντιμετωπιστεί η συνεχής αύξηση της πολυμεσικής πληροφορίας παγκοσμίως, κάτι το οποίο θέτει σοβαρές προκλήσεις όσον αφορά στις τρέχουσες τεχνολογίες αναζήτησης και ανάκτησης πολυμεσικού περιεχομένου. Στη βιβλιογραφία υπάρχουν αρκετοί αλγόριθμοι ανάκτησης πληροφορίας οι οποίοι λαμβάνουν κάποιο είδος εννοιολογικού πλαισίου υπόψη τους. Όμως, οι περισσότεροι εξ αυτών βασίζονται σε στατιστικά γλωσσικά μοντέλα για να συνδυάσουν τα προγενέστερα ερωτήματα του χρήστη με τις τρέχουσες επιλογές του, έτσι ώστε να επιτύχουν μια καλύτερη ταξινόμηση των πολυμεσικών εγγράφων ([31], [90], [111], [149]). Κάποιοι αλγόριθμοι συσχετιστικής, αλλά και αφανούς (im-plicit) ανατροφοδότησης εκμεταλλεύονται το εννοιολογικό πλαίσιο του χρήστη ως πηγή πληροφορίας που έρχεται να συμπληρώσει τα σαφή ερωτήματα χρηστών και να καθοδηγήσει τη διαδικασία της ανάκτησης ([207], [51], [136], [224], [276]).

Στη γενική περίπτωση, το δυναμικό παραγόμενο πολυμεσικό περιεχόμενο μπορεί να συμπεριλάβει διάφορες μορφές. Προσποτοιημένο πολυμεσικό περιεχόμενο με βάση το κείμενο, μπορεί να υπάρξει εύκολα στον Ιστό (Web) και στα προσαρμοζόμενα συστήματα υπερμέσων (adaptive hypermedia systems) [46]. Η κατόπιν παραγγελίας

παραγωγή και δημοσίευση προσωποποιημένου πολυμεσικού περιεχομένου εισάγεται με ερευνητικές προσπάθειες, όπως το προσωποιημένο λεύκωμα MyPhotos [248] και υπηρεσίες, όπως το Flickr [292] ή το Youtube [320]. Στον τομέα των δυναμικά παραγόμενων προσωποποιημένων πολυμεσικών παρουσιάσεων, οι [99] και [29] αποτελούν ενδεικτικά ερευνητικά παραδείγματα, ενώ το SlideShare [309] εισήγαγε όρους όπως τα “webinars” και το “συνεργατικό λογισμικό” (collaborative software) στο ευρύ κοινό. Πρόσφατα, υπηρεσίες κοινωνικών δικτύων και “mashups” σηματοδότησαν την απαρχή μιας νέας εποχής για την προσωποποίηση, τόσο σε προσωπικό (π.χ. Facebook [291], Myspace [306], Flixster [293]), όσο και σε επαγγελματικό (π.χ. LinkedIn [303], Plaxo [307]) επίπεδο χρηστών, αξιοποιώντας πληροφορίες αναρτηματοθέτησης (tagging) και συστάσεων (recommendations). Σε όλες τις περιπτώσεις, η προσωποποίηση βελτιώνει την υποκειμενική απόδοση της ανάκτησης, όπως αυτή γίνεται αντιληπτή από τους μεμονωμένους χρήστες ([131], [96], [159], [172], [55]), ενώ την ενισχύει με το να συμπληρώνει τα εμφανή (explicit) αιτήματα των χρηστών με τις σχετικές αφανείς (implicit) προτιμήσεις τους. Με τον τρόπο αυτό, η προσωποποίηση έρχεται να ικανοποιήσει καλύτερα τις ιδιαίτερες ανάγκες των χρηστών τέτοιων συστημάτων [141], αλλά και να ανακουφίσει από το γνωστό πρόβλημα της υπερφόρτωσης πληροφοριών [63].

#### 1.4.2.2 Ερευνητική συνεισφορά

Η ερευνητική συνεισφορά της ενότητας αυτής δίνει έμφαση στην προσωποποίηση του πολυμεσικού περιεχομένου από την σκοπιά της πρόσβασης σε αυτό και ειδικότερα, από τη σκοπιά του ορισμού, μοντελοποίησης και εκμετάλλευσης των κατάλληλων μορφών εννοιολογικού πλαισίου στη διαδικασία. Σε αυτό το πλαίσιο, όπου η ανοχή στις ανακριβείς περιγραφές αποτελεί υπόθεση εργασίας, η μοντελοποίηση του σημασιολογικού εννοιολογικού πλαισίου παίρνει έναν βασικό ρόλο στην εκμετάλλευση του βαθμού ασάφειας που περιλαμβάνεται στη γενικότερη θεώρηση εξατομίκευσης και κατά συνέπεια, εισάγονται κατάλληλα μοντέλα για να διαχειριστεί αυτή η αβεβαιότητα (uncertainty). Η συμβολή στο χώρο αποτελείται από την εκμετάλλευση της οντολογικής πληροφορίας ως της πηγής του σημασιολογικού εννοιολογικού πλαισίου και ως μιας βοήθειας για την συσχέτιση διαφορετικών μερών του βασισμένου στο εννοιολογικό πλαίσιο γνώσης πεδίου κατά τη διαδικασία της ανάκτησης πολυμεσικού περιεχομένου. Η πρόσθετη σημασιολογία (ακριβής ταξινόμηση, ρητές σχέσεις μεταξύ των εννοιών) που εισάγεται, παρέχει μια πλούσια πηγή πρόσθετης γνώσης, επιτρέποντας σημαντικές βελτιώσεις σε σύγκριση με τα αποτελέσματα που επιτυγχάνονται π.χ. με τη χρήση ανεξάρτητων, απλών λέξεων-κλειδιών.

Η συνεισφορά μας στο χώρο βελτιώνει την αποτελεσματικότητα της προσωποποίησης πολυμεσικού περιεχομένου, όπως η τελευταία γίνεται αντιληπτή σε ένα συγκεκριμένο εννοιολογικό πλαίσιο, μειώνοντας τελικά κάποια από τα περιστασιακά μειονεκτήματά της, όπως π.χ. είναι η έλλειψη διακριτικότητας (obtrusiveness), η ανακριβεία (inaccuracy), η ασυνέπεια (inconsistency) και η απόσπαση της προσοχής (distraction). Αυτό επιτυγχάνεται κυρίως με το να καθιστά περισσότερο σχετική προς το εκάστοτε εννοιολογικό πλαίσιο (context-relevant), καθώς και εννοιολογικά συνεπή, τη διαδικασία της προσωποποίησης. Τα μοντέλα και οι τεχνικές που προτείνονται στο πλαίσιο της διατριβής εξετάζουν την αυτόματη εξαγωγή των εμμένοντων (persistent), βασισμένων στο περιεχόμενο (content-based) προτιμήσεων των χρηστών, καθώς επίσης και των “ζωντανών”, επί τούτω ενδιαφερόντων τους, κατά τέτοιο τρόπο έτσι ώστε ο συνδυασμός και των δύο να δημιουργεί κατάλληλα μοντέλα χρηστών. Τα τε-

λευταία βασίζονται στο εννοιολογικό πλαίσιο και εφαρμόζονται στη συνέχεια για να αυξήσουν την ακρίβεια κατά την προσωποποίηση των αποτελεσμάτων της ανάκτησης, με άκρως ικανοποιητικά αποτελέσματα (βλ. λ.χ. ενότητα 5.7). Παρουσιάζουμε, δε, αποτελέσματα και πειραματικές μετρήσεις από τη διαδικασία της ανάκτησης προσωποποιημένου πολυμεσικού περιεχομένου με βάση το εννοιολογικό πλαίσιο, καθώς και επιλεγμένες περιπτώσεις χρήσης, προς επίρρωση των παραπάνω.

Λαμβάνοντας υπόψη τα παραπάνω, η προσέγγισή μας συμπληρώνει την βασισμένη σε οντολογίες προοπτική με ασαφείς έννοιες για την αναπαράσταση των προτιμήσεων των χρηστών, του εννοιολογικού πλαισίου των χρηστών, της σημασιολογίας του περιεχομένου και των σχέσεων μεταξύ των σημασιολογικών εννοιών. ειδικά για το εννοιολογικό πλαίσιο, εισάγει καινοτόμες μορφές αναπαράστασης και αξιοποίησης του εννοιολογικού πλαισίου κατά τη διαδικασία της προσωποποίησης πολυμεσικού περιεχομένου, καθορίζοντάς το μέσω ασαφών σημασιολογικών σχέσεων και οντολογίων. Οι προτεινόμενες μέθοδοι για την εξαγωγή προφίλ χρηστών και την εξατομικευμένη ανάκτηση περιεχομένου με βάση το εννοιολογικό πλαίσιο βασίζονται στις αρχές των ασαφών συνόλων και της ασαφούς σχεσιακής άλγεβρας και εκμεταλλεύονται κανούριες και ήδη διαθέσιμες τεχνικές της ερευνητικής περιοχής, οι οποίες θεωρούνται κατάληγες για να εξετάσουν τα προβλήματα που περιλαμβάνουν ασαφή μεγέθη [139], [286].

### 1.4.3 Θέματα οπτικού εννοιολογικού πλαισίου

Η παρούσα διατριβή κινείται επίσης και γύρω από την ερευνητική περιοχή της ανάλυσης πολυμεσικού περιεχομένου με βάση το εννοιολογικό πλαίσιο. Η οπτική πληροφορία που ενυπάρχει στο πολυμεσικό περιεχόμενο μπορεί να αποβεί ιδιαίτερα χρήσιμη σε περιπτώσεις αναγνώρισης αντικειμένων και ταξινόμησης σκηνών. Η αξιοποίηση αυτής της πληροφορίας είναι άρρηκτα συνδεδεμένη με την έννοια του εννοιολογικού πλαισίου, η οποία μπορεί να λάβει ποικίλες διαφορετικές υποστάσεις.

#### 1.4.3.1 Επισκόπηση ερευνητικής περιοχής

Η έννοια και η ιστορία της μοντελοποίησης του οπτικού εννοιολογικού πλαισίου περιγράφεται συνοπτικά από τον Torralba στη δημοσίευση [257]. Περαιτέρω μελέτες των Biederman [33] και Palmer [199] δίνουν έμφαση στην επίδραση του εννοιολογικού πλαισίου στο χρόνο επεξεργασίας κατά την αναγνώριση αντικειμένων σε μία σκηνή. Σε παρόμοιο πνεύμα, διάφορες ερευνητικές μελέτες υποστηρίζουν την ιδέα ότι η σημασιολογία μιας σκηνής είναι δυνατόν να είναι διαθέσιμη από τα αρχικά στάδια της επεξεργασίας των πληροφοριών ([221], [253], [196], [221]).

Στη σημερινή εποχή, είναι πλέον γεγονός, ότι ποικίλες μεθοδολογίες ανάλυσης εικόνας είναι δυνατόν να χρησιμοποιηθούν για την επίτευξη μιας αποτελεσματικής κατανόησης του περιεχομένου της και ότι οι σημερινοί ερευνητές κοιτάζουν πέρα από τα χαμηλού επιπέδου χαρακτηριστικά γνωρίσματα (όπως το χρώμα, η υφή και το σχήμα), αναζητώντας αποτελεσματικότερες μεθόδους έρευνας. Οι περισσότερες από αυτές τις εργασίες, όμως, βασίζονται αποκλειστικά σε χαμηλού επιπέδου χαρακτηριστικά και/ή δεν λαμβάνουν υπόψη τους το οπτικό εννοιολογικό πλαίσιο στη διαδικασία ([13], [92], [152], [52]). Κατά συνέπεια, εξακολουθεί να εμφανίζεται ένας σημαντικός αριθμός λανθασμένων ταξινομήσεων στα αποτελέσματά τους, μιας και όσο καλή και να είναι μία τέτοια μέθοδος, το γεγονός ότι δεν λαμβάνει υπόψη της την ενυπάρχουσα

περιβαλλοντική γνώση (όποια μορφή κι αν έχει η τελευταία), δρα όχρως αυτοπεριοριστικά. Σήμερα, υπάρχουν κάποια λίγα τέτοια συστήματα που συνδυάζουν τις οπτικές πληροφορίες εννοιολογικού πλαισίου με τη χαμηλού επιπέδου επεξεργασία. Επίσης, υπάρχουν και αρκετές σχετικά παρόμοιες μέθοδοι στην πρόσφατη βιβλιογραφία της υπολογιστικής όρασης (computer vision), οι οποίες συνδυάζουν προγενέστερη γνώση (δηλ. προσδοκίες) με τρέχουσες εκτιμήσεις (δηλ. την παραγωγή της χαμηλού επιπέδου επεξεργασίας) ([105], [106], [107]). Η χρήση της γνώσης και κατ' επέκταση των κατάλληλων μορφών του εννοιολογικού πλαισίου στη διαδικασία αναγνώρισης και επεξεργασίας, βελτιστοποιεί τα αποτελέσματά τους και εκμεταλλεύεται τη σημασιολογία των οπτικών χαρακτηριστικών. Κάποιοι λίγοι ερευνητές προσπαθούν να αντισταθμίσουν αυτήν την έλλειψη γνώσης με την εστίαση σε τεχνικές μάθησης χωρίς επίβλεψη ([271], [209]). Όμως, η ξεκάθαρη θεώρηση για το οπτικό εννοιολογικό πλαισίο που ακολουθείται στην παρούσα διατριβή αναφέρεται τόσο στις σχέσεις που διέπουν τη θέση διαφορετικών αντικειμένων στη σκηνή μιας εικόνας [182], όσο και στις σημασιολογικές σχέσεις που συσχετίζουν τα τελευταία και αποτελεί, εν γένει, μια πλήρη και βασισμένη στην ελλοχεύουσα γνώση, ερευνητική προσέγγιση.

#### 1.4.3.2 Ερευνητική συνεισφορά

Καθώς η προτεινόμενη προσέγγιση μοντελοποίησης του οπτικού εννοιολογικού πλαισίου το εμπλέκει σε όλες τις επιμέρους διεργασίες, εκ πρώτης όψεως το εννοιολογικό πλαισίο φαίνεται να αποτελεί μια από τις αιτίες για το φαινόμενο της υπερφόρτωσης πληροφορίας. Όμως, το προτεινόμενο εννοιολογικό πλαισίο μπορεί ταυτόχρονα να υπαγορεύσει και το ποιά αντικείμενα πρέπει να παρατηρηθούν και τι μπορεί να αγνοηθεί, κάτι το οποίο αποτελεί ένα ερευνητικό επίτευγμα, όταν εστιάζουμε στα τρέχοντα μοντέλα οπτικής προσοχής και ανάλυσης εικόνας. Η χρήση της έννοια του οπτικού εννοιολογικού πλαισίου οδηγεί σε επιτυχείς αναγνωρίσεις τόσο των αντικειμένων όσο και των γεγονότων, αλλά και συμβάλλει στο ευρύτερο ερευνητικό πεδίο της ανάλυσης και επεξεργασίας του πολυμεσικού περιεχομένου.

Η συμβολή στο χώρο αποτελείται από την εκμετάλλευση της οντολογικής πληροφορίας ως πηγής του οπτικού εννοιολογικού πλαισίου και ως μιας βοήθειας για την συσχέτιση διαφορετικών μερών του βασισμένου στο εννοιολογικό πλαισίο γνώσης πεδίου κατά τη διαδικασία της ανάλυσης πολυμεσικού περιεχομένου. Και σε αυτή την περίπτωση, η πρόσθετη σημασιολογία που εισάγεται με τη μορφή ρητών σχέσεων μεταξύ των εννοιών/αντικειμένων, παρέχει μια ιδιαίτερα πλούσια πηγή πρόσθετης γνώσης, επιτρέποντας σημαντικές βελτιώσεις σε σύγχριση με τα αποτελέσματα που επιτυγχάνονται χωρίς την αξιοποίηση αυτού του είδους γνώσης (βλ. λ.χ. υποενότητες 6.4.4 και 6.5.5).

Η προτεινόμενη προσέγγιση ορίζει, μοντελοποιεί και κατασκευάζει την απαιτούμενη, βασισμένη σε οντολογίες, γνώση, χρησιμοποιώντας βασικές αρχές της ασαφούς σχεσιακής άλγεβρας (π.χ. ασαφή σύνολα, ασαφείς δυαδικές σχέσεις, κ.λ.π.) για την αναπαράστασή της. Προτείνουμε πρωτότυπες δομές αναπαράστασης “ασαφοποιημένων” οντολογιών εννοιολογικού πλαισίου, οι οποίες βασίζονται στη γλώσσα RDF [316], αλλά και επεκτείνουμε και ενημερώνουμε τις δομές αυτές προς την κατεύθυνση της γλώσσας OWL [319], υιοθετώντας την ευρέως διαδεδομένη άποφη ότι η τελευταία αποτελεί την πιο σύγχρονη και με περισσότερες δυνατότητες (λ.χ. εξαγωγής συμπερασμάτων) έκφανση των τεχνολογιών του Σημασιολογικού Ιστού. Με τον τρόπο αυτό, εισάγονται καινοτόμες μορφές αναπαράστασης και διευρύνονται οι ορίζοντες της αξιοποίησης του οπτικού εννοιολογικού πλαισίου. Η σχετική έρευνα μας,

προχωράει, λοιπόν, ένα βήμα παραπέρα, και προτείνει συγκεκριμένους αλγορίθμους εκμετάλλευσης της επιπλέον γνώσης, κατάλληλους για να μελετήσουν προβλήματα που περιλαμβάνουν ασαφή μεγέθη [139], [286], εκμεταλλευόμενη το μοντελοποιημένο εννοιολογικό πλαίσιο. Οι αλγόριθμοι αυτοί βρίσκουν εφαρμογή τόσο σε περιπτώσεις συγκεκριμένης θεματικής περιοχής, όσο και σε περιπτώσεις πολλαπλών θεματικών περιοχών. Με άλλα λόγια, εμβαθύνουμε, σε θέματα επέκτασης και βελτίωσης της απόδοσης του απλού αλγορίθμου οπτικού εννοιολογικού πλαισίου προς την κατεύθυνση της ταυτόχρονης υποστήριξης πολλαπλών θεματικών περιοχών και μεγάλου αριθμού εννοιών [185]. Οι προσεγγίσεις που προτείνουμε αποδεικνύονται, όπως διαφαίνεται και από τις πειραματικές μετρήσεις μας, ιδιαίτερα αποδοτικές στα πεδία της σημασιολογικής κατάτμησης εικόνων και της εξαγωγής συμπερασμάτων από το πολυμεσικό περιεχόμενο, ενώ η επιτυχής υιοθέτησή των προτεινόμενων αναπαραστάσεων γνώσης πιστεύουμε ότι θα ανοίξει το δρόμο για την προτυποποίηση των προτεινόμενων μορφών του εννοιολογικού πλαισίου στον ευρύτερο ερευνητικό χώρο.

#### 1.4.4 Θέματα εννοιολογικού πλαισίου “ενδιάμεσου επιπέδου”

Ως τελευταίο μέρος της παρούσας διατριβής έρχεται η περαιτέρω επέκταση της ερευνητικής ενασχόλησής μας γύρω από ευφυείς τεχνικές ανάλυσης του πολυμεσικού περιεχομένου, πάντα με βάση το εκάστοτε εννοιολογικό πλαίσιο. Η ανίχνευση και αναγνώριση υψηλού επιπέδου εννοιών μέσα στα πολυμεσικά έγγραφα με την εισαγωγή οντοτήτων και γνώσης εννοιολογικού πλαισίου ενδιάμεσου επιπέδου αποτελεί το αντικείμενο της ενότητας αυτής.

##### 1.4.4.1 Επισκόπηση ερευνητικής περιοχής

Είναι γεγονός, ότι το εννοιολογικό πλαίσιο αποτελεί κρίσιμο παράγοντα κατά την διαδικασία αναγνώρισης αντικειμένων από τους ανθρώπους, δεδομένου ότι το ανθρώπινο οπτικό σύστημα αντίληφης εκμεταλλεύεται εκτενώς το περιβάλλον για να διευκολύνει την οποιαδήποτε ανίχνευση αντικειμένου [33].

Ο συνδυασμός του οπτικού εννοιολογικού πλαισίου ενδιάμεσου επιπέδου με τεχνικές (ημι-αυτόματης κατηγοριοποίησης εικόνων ή/και σκηνών, προκειμένου να επιτύχουμε καλύτερα σημασιολογικά αποτελέσματα κατά την ανάλυση του πολυμεσικού περιεχομένου αποτελεί από μόνο του ένα ευρύ ερευνητικό πεδίο-πρόκληση για τον ερευνητή. Εν γένει, οι ερευνητικές προσεγγίσεις της ανάλυσης πολυμεσικού περιεχομένου εξακολουθούν να διαιρούνται σε δύο μάλλον διαχριτές κατηγορίες: αφενός, στις χαμηλού επιπέδου μεθόδους και εργαλεία ανάλυσης πολυμέσων (όπως π.χ. η [205]) και αφετέρου, στις υψηλού επιπέδου μεθόδους και εργαλεία σημασιολογικού σχολιασμού (όπως π.χ. οι [259], [24]). Μόνο πρόσφατα άρχισε να αξιοποιείται και κάποιας μορφής σημασιολογία στη διαδικασία. Η ιδέα του συνδυασμού της τυποποιημένης γνώσης και ενός συνόλου χαρακτηριστικών γνωρισμάτων για να περιγραφεί το οπτικό περιεχόμενο μιας εικόνας παρουσιάζεται λ.χ. στη δημοσίευση [266], ενώ στις δημοσιεύσεις [285] και [158] πραγματοποιείται μια προσπάθεια για την εκμετάλλευση περιορισμών χωρικού εννοιολογικού πλαισίου. Επιπλέον, στις [57], [234] εισάγεται μια παραπλήσια, καθοδηγούμενη από λεξικό, προσέγγιση.

Στο ιδιαίτερα ενδιαφέρον ερευνητικό πεδίο του χαρακτηρισμού εικόνων με βάση τις κατατμημένες περιοχές τους, υπάρχουν διάφορες ερευνητικές προσεγγίσεις ([100], [135], [155], [18], [132], [134], [85], [283]), αλλά οι περισσότερες δεν ασχολούνται με τις συσχετίσεις μεταξύ των μεμονωμένων μερών και κατά συνέπεια την ελλοχεύουσα

γνώση. Η χρήση πληροφοριών ενδιάμεσου επιπέδου, όπως είναι οι πληροφορίες που λαμβάνονται από την εφαρμογή επιβλεπόμενων ή ανεπίβλεπτων μεθοδολογιών μηχανικής μάθησης για τα χαμηλού επιπέδου χαρακτηριστικά, μπορεί να χρησιμοποιηθεί για να βελτιώσει τα αποτελέσματα της παραδοσιακής - βοηθούμενης από τη γνώση - ανάλυσης, καθώς συνδιάζει τόσο χαμηλού επιπέδου οπτική, όσο και υψηλού επιπέδου "εννοιολογική" πληροφορία.

#### 1.4.4.2 Ερευνητική συνεισφορά

Στο τμήμα αυτό της διατριβής παρουσιάζονται τρεις νέες θεωρήσεις του (οπτικού) εννοιολογικού πλαισίου. Ειδικότερα, εισάγεται η θεμελιώδης έννοια του εννοιολογικού πλαισίου "ενδιάμεσου επιπέδου", η οποία περιλαμβάνει στοιχεία τόσο προς την κατεύθυνση των σημασιολογικών εννοιών (περιλαμβάνοντας καινούριες και διαφορετικής, σε σχέση με το απλό οπτικό εννοιολογικό πλαίσιο, μορφής σύνθετες έννοιες), όσο και χαρακτηριστικών "μεσαίου επιπέδου", με τη μορφή μιας καινούριας έννοιας, αυτής των "τύπων περιοχής" (region types). Οι τελευταίοι ορίζονται ως μια ενδιάμεση περιγραφή, η οποία μπορεί να περιγραφεί σημασιολογικά, αλλά δεν εκφράζει ούτε μία υψηλού επιπέδου έννοια, ούτε μία χαμηλού επιπέδου περιγραφή που εξάγεται απ' ευθείας από το πολυμεσικό έγγραφο. Σχεδιάζονται, ορίζονται και μοντελοποιούνται μια σειρά από κατάλληλες οντολογίες εννοιολογικού πλαισίου, καθορίζονται οι σχέσεις τους και ενσωματώνονται στο γενικότερο μοντέλο αναπαράστασης γνώσης που προτείνουμε με αυτή τη διατριβή. Και στην περίπτωση αυτή, σχεδιάζονται και υλοποιούνται αλγόριθμοι για την υπολογιστικά αποδοτική διαχείριση του "μεσαίου επιπέδου" οπτικού εννοιολογικού πλαισίου, ενώ εκμεταλλευόμενοι τα αποτελέσματα της έρευνας στον τομέα της εξαγωγής ενδιάμεσου επιπέδου χαρακτηριστικών, εξηγείται ο τρόπος με τον οποίο οι σχετικοί αλγόριθμοι και μεθοδολογίες μπορούν να συνδυαστούν αποδοτικά, προκειμένου να αναλύσουμε σημασιολογικά το πολυμεσικό περιεχόμενο και να το "φέρουμε" έτσι πιο κοντά στις επιθυμίες και τις ανάγκες του τελικού χρήστη του.

Εν ολίγοις, η ερευνητική προσπάθειά μας στο κεφάλαιο αυτό εστιάζει αρχικά σε μια ολοκληρωμένη προσέγγιση που προσφέρει έναν ενοποιημένο και ανεπίβλεπτο χειρισμό του πολυμεσικού περιεχομένου. 'Όπως και προηγουμένως, η προτεινόμενη προσέγγιση ορίζει, μοντελοποιεί και κατασκευάζει την απαιτούμενη, βασισμένη σε οντολογίες, γνώση, χρησιμοποιώντας ασαφή σύνολα και διαφόρων ειδών ασαφείς δυαδικές σχέσεις για την αναπαράστασή της. Με τον τρόπο αυτό, εισάγονται πρωτότυπες μορφές αναπαράστασης και αξιοποίησης του εννοιολογικού πλαισίου ενδιάμεσου επιπέδου, οι οποίες συμμορφώνονται πλήρως με τα πρότυπα του Σημασιολογικού Ιστού (όπως π.χ. συμβαίνει με την τεχνική αναπαράστασης RDF reification που προτείνουμε για τα fuzzy RDF triplets). Και στην περίπτωση αυτή, σχεδιάζουμε και υιοθετούμε συγκεκριμένους αλγορίθμους εκμετάλλευσης της επιπλέον οντολογικής γνώσης, κατάλληλους για να μελετήσουν τα σχετικά προβλήματα σημασιολογικής ανάλυσης του πολυμεσικού περιεχομένου.

Ενεργούμε συμπληρωματικά στην τρέχουσα υπάρχουσα κατάσταση, μιας και, εστιάζοντας στη σημασιολογική ανάλυση των πολυμέσων, συμβάλλουμε στο γεφύρωμα του χάσματος [232] μεταξύ της σημασιολογικής και της ακατέργαστης φύσης του πολυμεσικού περιεχομένου και αντιμετωπίζουμε ένα από τα πιο ενδιαφέροντα προβλήματα της ανάλυσης πολυμεσικού περιεχομένου, αυτό της ανίχνευσης υψηλού επιπέδου εννοιών μέσα σε αυτό. Αποδεικνύουμε, δε, με σχετικές πειραματικές μετρήσεις, ότι η χρήση των μεσαίου επιπέδου πληροφοριών που εισάγονται με την παρούσα, βελτιώνει

τα αποτελέσματα της κλασικής - βοηθούμενης από τη γνώση - πολυμεσικής ανάλυσης και εκφράζουμε με πληρότητα το γεγονός ότι ο συνδυασμός της έννοιας του εννοιολογικού πλαισίου με τεχνικές (ημι-)αυτόματης κατηγοριοποίησης εικόνων ή/και σκηνών, επιτυγχάνει καλύτερα σημασιολογικά αποτελέσματα κατά την ανάλυση του πολυμεσικού περιεχομένου.

□



# Κεφάλαιο 2

## Ερευνητικό πλαίσιο

### 2.1 Εισαγωγή

Το παρόν κεφάλαιο επιχειρεί να παρουσιάσει το ερευνητικό υπόβαθρο πάνω στο οποίο αποπειράται να κινηθεί η παρούσα ερευνητική εργασία. Συγκεκριμένα, το κεφάλαιο χωρίζεται στους ακόλουθους τέσσερις βασικούς άξονες:

- *Μηχανική Μάθηση / Ανάλυση και Εξόρυξη Δεδομένων.* Στην ενότητα 2.2 γίνεται μία εισαγωγή στις κυριότερες έννοιες του ερευνητικού πεδίου της μηχανικής μάθησης (γενικότερα) και της εξόρυξης γνώσης (ειδικότερα), καθώς αποτελούν μέρος των βασικών δομών ολόκληρης της παρούσας εργασίας.
- *Ανάλυση Πολυμεσικού Περιεχομένου Υποστηριζόμενη από τη Γνώση και το Εννοιολογικό Πλαίσιο.* Στην ενότητα 2.3 επιχειρείται η εισαγωγή στις βασικές έννοιες της υποστηριζόμενης από τη γνώση ανάλυσης του πολυμεσικού περιεχομένου, η οποία αποτελεί και τη βάση της μετέπειτα παρουσιαζόμενης συνεισφοράς μας με την έννοια του εννοιολογικού πλαισίου γνώσης (κεφάλαιο 6 της παρούσης).
- *Πρόσβαση και Ανάκτηση Πολυμεσικής Πληροφορίας.* Στην ενότητα 2.4 παραθέτονται βασικές αρχές και εισαγωγικά στοιχεία της θεματικής περιοχής της πρόσβασης και ανάκτησης πολυμεσικής πληροφορίας. Παρουσιάζονται οι διαδικασίες που διέπουν τα σύγχρονα συστήματα ανάκτησης πληροφορίας και πραγματοποιείται μια πρώτη απόπειρα μοντελοποίησής τους. Συζητούνται, επίσης, διαφόρων ειδών κατάλληλες μετρικές συστημάτων ανάκτησης πληροφορίας, ενώ εστιάζουμε και στα πιθανά προβλήματά τους.
- *Προσωποποίηση Πολυμεσικού Περιεχομένου.* Στην ενότητα αυτή (2.5) παρουσιάζονται συνοπτικά οι βασικές αρχές της ερευνητικής περιοχής της προσωποποίησης περιεχομένου, όπως είναι η εξαγωγή προφίλ χρηστών τόσο υπό την ευρεία έννοια, όσο και σύμφωνα με τα πρότυπα MPEG-7 και MPEG-21 και η συσχετιστική ανατροφοδότηση. Συζητείται, επίσης, και η ιδιαίτερα διαδεδομένη εφαρμογή της προσωποποίησης σε συστήματα IR μέσω της κατάταξης προσωποποιημένων αποτελεσμάτων αναζήτησης.

## 2.2 Μηχανική Μάθηση / Ανάλυση και Εξόρυξη Δεδομένων

Η μάθηση, όπως και η νοημοσύνη, καλύπτει εν γένει μια τόσο ευρεία σειρά διαδικασιών που είναι ιδιαίτερα δύσκολο να καθοριστεί με ακρίβεια και σαφήνεια. Σε αυτή την ενότητα θα ασχοληθούμε αποκλειστικά με τη μηχανική μάθηση (machine learning), η οποία, όπως προδίδει και ο τίτλος της, αφορά τις μηχανές ή καλύτερα τους υπολογιστές. Οι τελευταίοι μαθαίνουν, εν γένει, όποτε αλλάζει η δομή, το πρόγραμμα, ή τα στοιχεία τους, σύμφωνα με μια λογική αλληλουχία στις εισόδους τους ή κάποιες άλλες εξωτερικές πληροφορίες με έναν τέτοιο τρόπο, έτσι ώστε η αναμενόμενη μελλοντική απόδοσή τους να βελτιώνεται. Μερικές από αυτές τις αλλαγές, όπως π.χ. η προσθήκη ενός αρχείου σε μια βάση δεδομένων, εμπίπτουν αναμφίβολα στη θεματική περιοχή άλλων επιστημών και δε γίνονται απαραίτητα καλύτερα κατανοητές με τη χρήση του όρου μάθηση. Από την άλλη μεριά, όμως, όταν βελτιώνεται π.χ. η απόδοση ενός συστήματος αναγνώρισης ομιλίας μετά από την ακοή διαφόρων δειγμάτων της ομιλίας ενός προσώπου, αισθανόμαστε αρκετά σίγουροι να ισχυριστούμε ότι το σύστημα αυτό έχει όντως μάθει κάτι νέο!

Στην παρούσα ενότητα θα παρουσιάσουμε εν συντομίᾳ τις βασικές αρχές της μηχανικής μάθησης (2.2.1) και της εξόρυξης δεδομένων (2.2.2). Θα αναλύσουμε τις πιο διαδεδομένες τεχνικές και μεθοδολογίες, όπως η μη επιβλεπόμενη μάθηση, η Μπεϋζιανή εξαγωγή συμπερασμάτων, οι ταξινομητές και τα Μπεϋζιανά δίκτυα, έχοντας πάντοτε ως στόχο την καλύτερη κατανόηση των επόμενων κεφαλαίων της εργασίας μας, μιας και στην πορεία θα εκμεταλλευτούμε και θα επεκτείνουμε τις περισσότερες από αυτές τις τεχνικές. Σε αυτό το πλαίσιο θα δώσουμε ιδιαίτερη έμφαση στον Μπεϋζιανό ταξινομητή και στις τεχνικές κατηγοριοποίησης κειμένου.

### 2.2.1 Μηχανική μάθηση

Υπάρχουν διάφοροι λόγοι για τους οποίους είναι σημαντική η μηχανική μάθηση. Μερικοί από αυτούς είναι και οι ακόλουθοι:

- Ο ορισμός ορισμένων υπολογιστικών προβλημάτων και στόχων πολλές φορές δεν μπορεί να πραγματοποιηθεί καλά και με σαφήνεια παρά μόνο μέσω παραδειγμάτων. Με άλλα λόγια μπορεί να είμαστε σε θέση να διευκρινίσουμε τις εισόδους και εξόδους ενός προβλήματος, αλλά όχι μία συνοπτική σχέση μεταξύ των εισόδων και των τελικών επιθυμητών αποτελεσμάτων. Θα επιθυμούσαμε οι μηχανές/υπολογιστές να είναι σε θέση να (αυτο-)ρυθμίζουν την εσωτερική δομή τους, έτσι ώστε να παράγουν σωστές εξόδους για ένα μεγάλο αριθμό δειγμάτων εισόδου και με τον τρόπο αυτό να θέτουν περιορισμούς στη συνάρτηση εισόδου-εξόδου τους, ούτως ώστε να είναι σε θέση να προσεγγίσουν την υπονοούμενη σχέση των παραδειγμάτων. Στο σημείο αυτό θα πρέπει να αναφερθεί, βέβαια, ότι μπορεί να κρύβονται ιδιαίτερα σημαντικές σχέσεις και συσχετισμοί μεταξύ μεγάλων όγκων δεδομένων και ότι οι μέθοδοι μηχανικής μάθησης είναι σε θέση να εξάγουν αυτές τις δεδομένες σχέσεις.
- Οι άνθρωποι σχεδιάζουν συχνά υπολογιστές/μηχανές, που δε λειτουργούν στο επιθυμητό επίπεδο στα περιβάλλοντα στα οποία χρησιμοποιούνται. Στην πραγματικότητα, είναι αρκετά συχνό το φαινόμενο να μην είναι γνωστά ορισμένα

χαρακτηριστικά του εργασιακού περιβάλλοντος κατά τον αρχικό χρόνο σχεδίασης. Επομένως, οι μέθοδοι μηχανικής μάθησης μπορούν να χρησιμοποιηθούν για τη βελτίωση τέτοιων σχεδίων υπολογιστών/μηχανών, ακόμα και μετά από μια τέτοια λανθασμένη αρχική εκτίμηση.

- Το μέγεθος της απαιτούμενης γνώσης για την εκπλήρωση ορισμένων υπολογιστικών στόχων μπορεί να φτάσει σε υψηλά μεγέθη. Ως εκ τούτου, είναι πιθανό να καταστεί αδύνατη η σαφής κωδικοποίησή του από ανθρώπους. Μηχανές ή υπολογιστές, που είναι σε θέση να μάθουν αυτήν τη γνώση σταδιακά, μπορούν να συλλέξουν τεράστιες ποσότητες γνώσης με σημαντικά λιγότερη προσπάθεια απ' ότι οι ανθρώποι. Επίσης, οι μηχανές, που είναι σε θέση να προσαρμοστούν σε ένα περιβάλλον με την πάροδο του χρόνου, θα μείωναν αποτελεσματικά την ανάγκη ύπαρξης ενός σταθερού πλάνου επανασχεδιασμού τους, καθώς δε χωράει αμφιβολία ότι νέα γνώση ανακαλύπτεται ή δημιουργείται διαρκώς από τους ανθρώπους. Ένας συνεχιζόμενος επανασχεδιασμός των συστημάτων τεχνητής νοημοσύνης, προκειμένου αυτά να προσαρμοστούν ιδιαίτερα στην εκάστοτε νέα γνώση και πραγματικότητα, δεν αποτελεί μια αποδοτική και αποτελεσματική τακτική. Απεναντίας, οι μέθοδοι μηχανικής μάθησης είναι σε θέση να παρακολουθήσουν και να υλοποιήσουν ένα μεγάλο μέρος από τις παραπάνω αλλαγές και να σχεδιάσουν αποδοτικά και σταθερά υπολογιστικά συστήματα.

Συνοψίζοντας, στο κεφάλαιο αυτό πραγματοποιείται μια εισαγωγή στις κυριότερες έννοιες του ερευνητικού πεδίου της μηχανικής μάθησης (γενικότερα) και της εξόρυξης γνώσης (ειδικότερα), καθώς θεωρούμε ότι αποτελούν μέρος των θεμελιωδών δομών ολόκληρης της παρούσας εργασίας. Αναμφίβολα, η πρόσφατη πρόοδος των σχετικών αλγορίθμων και θεωριών, η ολοένα αυξανόμενη προσφορά διαθέσιμων online δεδομένων, αλλά και η ολοένα αυξανόμενη υπολογιστική ισχύς οριοθετούν ένα ζωτικό χώρο για έρευνα με όχημα τη μηχανική μάθηση. Κάποιος είναι σε θέση να αναγνωρίσει ποικίλες θεματικές περιοχές που λειτουργούν ως υποδοχείς της μηχανικής μάθησης, μεταξύ των οποίων συγκαταλέγονται:

- Η εξόρυξη δεδομένων, δηλαδή, με άλλα λόγια η χρησιμοποίηση προγενέστερων ιστορικών δεδομένων προς όφελος και βελτίωση της λήψης αποφάσεων.
- Η ανάπτυξη εφαρμογών λογισμικού, σε περιπτώσεις όπου ο προγραμματισμός δεν μπορεί να πραγματοποιηθεί με το χέρι είτε για λόγους υπολογιστικής πολυπλοκότητας, είτε για λόγος χρόνου και δυνατοτήτων του ανθρωπίνου είδους (π.χ. η αναγνώριση φωνής).
- Η περίπτωση των αυτο-διοργανούμενων προγραμμάτων (όπως π.χ. ένας αναγνώστης ειδήσεων, ο οποίος μαθαίνει (από) τα ενδιαφέροντα των χρηστών του).

## 2.2.2 Εξόρυξη δεδομένων

Εάν εστιάζαμε στην πρώτη από τις παραπάνω κατηγορίες και αν θέλαμε να εξετάσουμε μια κλασική περίπτωση εξόρυξης δεδομένων (data mining), θα αναφερόμασταν στο κλασικό παράδειγμα των ιατρικών φακέλων γεννήσεων, όπου ο κάθε φάκελος περιέχει πληροφορίες κυήσεων και γεννήσεων με πολλαπλά χαρακτηριστικά γνωρίσματα (όπως π.χ. για το αν πρόκειται για την πρώτη γέννα της γυναίκας ή όχι, κ.ο.κ.) και όπου οι μελλοντικές ασθενείς κατατάσσονται ανάλογα με το προϋπάρχον ιστορικό.

Η εξόρυξη δεδομένων είναι μια διαδικασία αυτόματης αναζήτησης προτύπων μέσα σε μεγάλους όγκους δεδομένων, αφού χρησιμοποιήσουμε εργαλεία, όπως η ταξινόμηση/κατηγοριοποίηση, η εξόρυξη κανόνων συσχέτισης, η συσταδοποίηση, κ.λ.π.. Η εξόρυξη δεδομένων αποτελεί από μόνη της ένα σύνθετο ερευνητικό θέμα και προσδίδει ιδιαίτερη αξία σε αν μη τι άλλο δημιουργικές υπολογιστικές τεχνικές στις θεματικές περιοχές της στατιστικής, της ανάκτησης πληροφορίας, της μηχανικής μάθησης και της αναγνώρισης προτύπων.

Ένα μάλλον απλό αλλά συνάμα κατανοητό παράδειγμα εξόρυξης δεδομένων, το οποίο απαντάται συχνά στη βιβλιογραφία ως “η ανάλυση του καλαθιού αγορών”, είναι η χρήση της στις λιανικές πωλήσεις. Σε μια υπεραγορά (super market), μέσω της ανάλυσης των συναλλαγών κατά τη διάρκεια μιας μακριάς χρονικής περιόδου, διαπιστώνεται λ.χ. ότι δύο φαινομενικά ανόμοια προϊόντα, όπως η μπύρα και οι πάνες, αγοράστηκαν από κοινού με ιδιαίτερη συχνότητα. Αν και η εξήγηση αυτής της σχέσης μπορεί να είναι δύσκολη (αν βασιστούμε μονάχα στην κοινή λογική), η εκμετάλλευσή της όμως είναι ευκολότερη. Παραδείγματος χάριν, το κατάστημα θα μπορούσε να αξιοποιήσει την παραπάνω πληροφορία και να τοποθετήσει τις πάνες που του αποφέρουν υψηλό κέρδος κοντά σε εκείνες τις μπύρες που του αποφέρουν επίσης υψηλό κέρδος. Το παράδειγμα αυτό είναι μεν απλοϊκό, καταδεικνύει όμως περίτρανα την ύπαρξη κανόνων συσχέτισης μέσα στα βασισμένα στις συναλλαγές δεδομένα.

Η εξόρυξη δεδομένων έχει στο παρελθόν οριστεί ως “η μη τετριμένη εξαγωγή υπονοούμενων, προηγουμένως άγνωστων και ενδεχομένως χρήσιμων πληροφοριών από τα δεδομένα” [93] αλλά και ως “η επιστήμη της εξαγωγής των χρήσιμων πληροφοριών από μεγάλα σύνολα δεδομένων” [108]. Περιλαμβάνει, δε, την ταξινόμηση μεγάλων όγκων δεδομένων, καθώς και την επιλογή μόνο των σχετικών πληροφοριών. Χρησιμοποιείται όλο και περισσότερο στις επιστήμες για την εξαγωγή πληροφοριών από τεράστια σύνολα δεδομένων, φαινόμενο που δημιουργείται όλο και πιο συχνά εξαιτίας των σύγχρονων πειραματισμών και της πληθώρας πληροφοριών που τους συνοδεύει. Τα μεταδεδομένα (metadata), ή αλλιώς τα δεδομένα για ένα δοσμένο σύνολο δεδομένων εκφράζονται συχνά είτε με μία συμπυκνωμένη διάταξη (format) που επιτρέπει την εξόρυξη δεδομένων, είτε με μία διάταξη που διευκολύνει την πρακτική της εξόρυξης δεδομένων. Συνήθη τέτοια παραδείγματα θεωρούνται οι διαφόρων ειδών περιλήψεις περιεχομένου, πολυμεσικού ή μη.

Αν και η εξόρυξη δεδομένων είναι ένας σχετικά νέος όρος, η τεχνολογία που τη συνοδεύει δεν είναι. Εδώ και αρκετό καιρό είναι γνωστό ότι οι επιχειρήσεις χρησιμοποιούν ισχυρούς υπολογιστές, που “κοσκινίζουν” τον όγκο των δεδομένων που συλλέγονται, όπως π.χ. τα δεδομένα από τους ανιχνευτές γραμμωτού κώδικα (bar code) που ενέσκηψαν προσφάτως στα super markets: Σχετικές εκθέσεις ερευνών αγοράς κάνονται όλο και πιο συχνά την εμφάνισή τους στις στήλες των οικονομικών εφημερίδων. Οι συνεχείς καινοτομίες και βελτιώσεις στην επεξεργαστική ισχύ των υπολογιστών, οι ολοένα και φθηνότεροι αποθηκευτικοί χώροι με τη μορφή σκληρών δίσκων και επανεγγράψιμων μνημών (flash memory), καθώς και το ολοένα και πιο εξελιγμένο λογισμικό στατιστικής αυξάνουν εντυπωσιακά την ακρίβεια και τη χρησιμότητα μιας τέτοιας ανάλυσης και η εξόρυξη δεδομένων είναι εδώ, για να προσδιορίσει τις τάσεις μέσα στα δεδομένα που υπερβαίνουν την απλή, επιφανειακή ανάλυση.

Τέλος, θα πρέπει να σημειώσουμε ότι ο όρος της εξόρυξης δεδομένων χρησιμοποιείται συχνά, για να περιγράψει και τις δύο, ανεξάρτητες διαδικασίες της ανακάλυψης και της πρόβλεψης γνώσης (knowledge discovery και knowledge acquisition, αντίστοιχα). Η ανακάλυψη γνώσης παρέχει τις ρητές πληροφορίες που έχουν μια ανα-

γνώσιμη μορφή και μπορούν να γίνουν κατανοητές από έναν χρήστη. Η πρόβλεψη, ή αλλιώς το προβλεπτικό μοντέλο παρέχει τις προβλέψεις των μελλοντικών γεγονότων και μπορεί να είναι διαφανείς και αναγνώσιμες σε μερικές προσεγγίσεις (όπως π.χ. στα βασισμένα σε κανόνες συστήματα) ή αδιαφανείς σε άλλες (όπως π.χ. τα νευρωνικά δίκτυα). Επιπλέον, μερικά συστήματα εξόρυξης δεδομένων, όπως τα νευρωνικά δίκτυα, τείνουν εγγενώς προς την πρόβλεψη παρά προς την ανακάλυψη της γνώσης.

### **2.2.3 Μη επιβλεπόμενη μάθηση**

Η μη επιβλεπόμενη μάθηση είναι μια μέθοδος μηχανικής μάθησης, κατά την οποία ένα μοντέλο προσαρμόζεται σε κάποιες παρατηρήσεις. Ξεχωρίζει από την επιβλεπόμενη μάθηση από το γεγονός ότι δεν υπάρχει καμία εκ των προτέρων γνωστή έξοδος. Στην μη επιβλεπόμενη μάθηση συγκεντρώνεται ένα σύνολο δεδομένων που αποτελείται από αντικείμενα εισόδου, τα οποία στη συνέχεια αντικετωπίζονται ως ένα σύνολο τυχαίων μεταβλητών. Κατόπιν, κατασκευάζεται ένα μοντέλο συνδυασμένης πυκνότητας για το σύνολο των δεδομένων.

Η μη επιβλεπόμενη μάθηση μπορεί να χρησιμοποιηθεί από κοινού με την εξαγωγή των Μπεϋζιανών συμπερασμάτων, προκειμένου να παραχθούν οι δεσμευμένες πιθανότητες (δηλαδή η επιβλεπόμενη μάθηση) για οποιαδήποτε από τις παραπάνω τυχαίες μεταβλητές με δεδομένες τις υπόλοιπες. Η μη επιβλεπόμενη μάθηση είναι επίσης χρήσιμη κατά τη συμπίεση δεδομένων, καθώς όλοι οι αλγόριθμοι συμπίεσης δεδομένων, είτε εμφανώς είτε σιωπηρά, στηρίζονται σε μια κατανομή πιθανότητας πάνω σε ένα σύνολο εισόδων.

Μια άλλη μορφή μη επιβλεπόμενης μάθησης είναι η συσταδοποίηση, για την οποία θα μιλήσουμε εκτενέστερα σε επόμενο κεφάλαιο της παρούσης και η οποία πολλές φορές δεν είναι καν πιθανοτική.

### **2.2.4 Μπεϋζιανή εξαγωγή συμπερασμάτων**

Η Μπεϋζιανή εξαγωγή συμπερασμάτων είναι ουσιαστικά μία στατιστική εξαγωγή συμπερασμάτων, κατά την οποία τα στοιχεία ή οι παρατηρήσεις χρησιμοποιούνται, για να ενημερώσουν ή για να συμπεράνουν εκ νέου την πιθανότητα ότι μια υπόθεση μπορεί να είναι αληθινή. Ο όρος “Μπεϋζιανή” προέρχεται από τη συχνή χρήση του θεωρήματος του Μπέυζ κατά τη διαδικασία εξαγωγής των συμπερασμάτων. Το θεώρημα του Bayes (που είναι επίσης γνωστό και ως κανόνας του Μπέυζ ή νόμος του Μπέυζ) είναι ένα αποτέλεσμα στη θεωρία πιθανοτήτων, το οποίο αφορά τις δεσμευμένες και οριακές κατανομές πιθανότητας τυχαίων μεταβλητών. Σε μερικές πιθανοτικές ερμηνείες το θεώρημα του Μπέυζ προβλέπει την ενημέρωση ή αναθεώρηση των πεποιθήσεων, λαμβάνοντας υπόψη τα νέα στοιχεία, ήτοι μία εκ των υστέρων ενημέρωσή τους.

Ειδικότερα, η δεσμευμένη πιθανότητα ενός γεγονότος  $A$  με δεδομένο ένα άλλο γεγονός  $B$  είναι εν γένει διαφορετική από την πιθανότητα του  $B$  με δεδομένο το  $A$ . Εντούτοις, υπάρχει μια καθορισμένη σχέση μεταξύ των δύο και το θεώρημα του Μπέυζ είναι η δήλωση αυτής ακριβώς της σχέσης. Το θεώρημα του Μπέυζ ισχύει σε όλες τις πιθανοτικές ερμηνείες. Εντούτοις, σε διάφορες προσεγγίσεις καταγράφονται διαφωνίες σχετικά με το σε ποια είδη πραγμάτων είναι δυνατόν να ανατεθούν πιθανότητες. Στη γενική περίπτωση, η Μπεϋζιανή προσέγγιση μπορεί να ορίσει τις πιθανότητες σε προτάσεις που περιέχουν αβεβαιότητα (uncertainty). Η επέκταση της σχετικής συζήτησης στο πλαίσιο αυτής της διατριβής θεωρούμε ότι θα ξέφευγε σαφώς

από το στόχο της και για το λόγο αυτό παραπέμπουμε τον αναγνώστη στη σχετική βιβλιογραφία [175].

### 2.2.5 Ταξινομητές

Στα μαθηματικά ένας ταξινομητής είναι μια αντιστοίχιση από ένα (διακριτό ή συνεχές) διάστημα χαρακτηριστικών γνωρισμάτων  $X$  σε ένα διακριτό σύνολο ετικετών  $Y$ . Οι ταξινομητές έχουν πρακτική εφαρμογή σε πολλούς κλάδους, χυρίως των επιστημών, αλλά και της ίδιας της κοινωνίας. Αν θεωρήσουμε μια αφαιρετική προσέγγιση των ταξινομητών, οι τελευταίοι μπορεί να θεωρηθούν ως συστήματα απόφασης, τα οποία δέχονται ως είσοδο τις τιμές μερικών χαρακτηριστικών γνωρισμάτων ή χαρακτηριστικών μιας κατάστασης και παράγουν ως έξοδο μια διακριτή ετικέτα που σχετίζεται με τις τιμές εισαγωγής. Ένα τέτοιο παράδειγμα αποτελεί ένας ταξινομητής ο οποίος λόγου χάριν δέχεται τις λεπτομέρειες του μισθού ενός προσώπου, την ηλικία του, την οικογενειακή του κατάσταση, τη διεύθυνση κατοικίας του και το πιστοληπτικό ιστορικό του και ταξινομεί το πρόσωπο αυτό ως ικανό ή μη ικανό να λάβει ένα νέο δάνειο ή μια νέα πιστωτική κάρτα.

Οι εφαρμογές των ταξινομητών είναι ποικίλες και πολλές φορές δε γίνονται καν αντιληπτές. Χρησιμοποιούνται ευρέως στην ιατρική (ανάλυση φαρμάκων, ανάλυση δεδομένων MRI), στην οικονομία (ανάλυση μετοχών, πρόβλεψη χρηματιστηριακών δεικτών), στην κινητή τηλεφωνία (αποκωδικοποίηση σήματος, διόρθωση λαθών), στην όραση υπολογιστών (ανάλυση προσώπων, παρακολούθηση στόχων), στην αναγνώριση φωνής και στην εξόρυξη δεδομένων (ανάλυση αγορών super markets) κ.ο.κ.. Οι ταξινομητές μπορεί να είναι είτε σταθεροί ταξινομητές, είτε ταξινομητές οι οποίοι μαθαίνουν με την πάροδο του χρόνου, οι οποίοι και διακρίνονται σε επιβλεπόμενους και μη επιβλεπόμενους ταξινομητές μάθησης. Έναν τέτοιο ταξινομητή θα κατασκευάσουμε στα πλαίσια της ενότητας 4.4.6 της παρούσης με άμεση εφαρμογή στη θεματική κατηγοριοποίηση εγγράφων.

### 2.2.6 Μπεϋζιανός ταξινομητής

Ένας αφελής ταξινομητής Μπέυζ είναι ένας απλός πιθανοτικός ταξινομητής, βασισμένος στην εφαρμογή του θεωρήματος Μπέυζ με τη χρήση ισχυρών (αφελών) υποθέσεων ανεξάρτησίας. Ένας όρος που περιγράφει ίσως καλύτερα το ελλοχεύον μοντέλο πιθανότητας θα ήταν το “ανεξάρτητο μοντέλο χαρακτηριστικών γνωρισμάτων”.

Ανάλογα με την ακριβή φύση του μοντέλου πιθανότητας οι αφελείς ταξινομητές Μπέυζ μπορούν να εκπαιδευθούν πολύ αποτελεσματικά σε μια επιβλεπόμενη ρύθμιση μάθησης. Σε πολλές πρακτικές εφαρμογές η εκτίμηση παραμέτρων για τα αφελή πρότυπα Μπέυζ χρησιμοποιεί τη μέθοδο της μέγιστης πιθανότητας. Με άλλα λόγια, κάποιος μπορεί να λειτουργήσει με το αφελές πρότυπο Μπέυζ, χωρίς να βασίζεται στην Μπεϋζιανή πιθανότητα ή χωρίς τη χρησιμοποίηση οποιωνδήποτε Μπεϋζιανών μεθόδων.

Παρά τον αφελή σχεδιασμό τους και τις προφανώς υπεραπλουστευμένες υποθέσεις τους οι αφελείς ταξινομητές Μπέυζ λειτουργούν συχνά πολύ καλύτερα απ' ότι αναμένεται κάτω από σύνθετες, πραγματικές καταστάσεις. Πρόσφατα, μάλιστα, η προσεκτική ανάλυση του Μπεϋζιανού προβλήματος ταξινόμησης έδειξε ότι υπάρχουν ισχυροί θεωρητικοί λόγοι για την εκ πρώτης όψεως αδικαιολόγητη αποτελεσματικότητα αυτών των αφελών ταξινομητών.

### 2.2.7 Κατηγοριοποίηση εγγράφων κειμένου

Η ταξινόμηση/κατηγοριοποίηση εγγράφων κειμένου είναι ένα από τα κύρια προβλήματα της επιστήμης των υπολογιστών γενικά και της ανάκτησης πληροφορίας ειδικότερα. Ο απότερος στόχος είναι να αποδοθεί ένα ηλεκτρονικό έγγραφο σε μια ή περισσότερες κατηγορίες με βάση το περιεχόμενό του. Οι στόχοι της ταξινόμησης εγγράφων μπορούν να διαιρεθούν σε δύο είδη: (i) την επιβλεπόμενη ταξινόμηση εγγράφων, όπου κάποιος εξωτερικός μηχανισμός (όπως π.χ. η ανθρώπινη ανατροφοδότηση) παρέχει τις πληροφορίες για τη σωστή ταξινόμηση των εγγράφων, και (ii) την μη επιβλεπόμενη ταξινόμηση εγγράφων κειμένου, όπου η ταξινόμηση γίνεται εξ ολοκλήρου χωρίς αναφορά στις εξωτερικές πληροφορίες.

Η ταξινόμηση εγγράφων κειμένου περιλαμβάνει μεθόδους, όπως:

- ο αφελής ταξινομητής Μπένζ
- ο αλγόριθμος tf-idf
- η λανθάνουσα σημασιολογική ευρετηρίαση
- οι διανυσματικές μηχανές υποστήριξης
- τα τεχνητά νευρωνικά δίκτυα
- ο αλγόριθμος kNN
- τα δένδρα απόφασης (όπως π.χ. τα ID3)
- η εξόρυξη εννοιών
- καθώς και προσεγγίσεις βασισμένες στην επεξεργασία φυσικής γλώσσας.

Μια εφαρμογή της ταξινόμησης εγγράφων κειμένου μπορεί να αναπαρασταθεί με ένα παράδειγμα ενός αφελούς ταξινομητή Μπένζ. Ας θεωρήσουμε ένα πρόβλημα ταξινόμησης εγγράφων κειμένου ανάλογα με το περιεχόμενό τους, όπως π.χ. τον διαχωρισμό σε ανεπιθύμητη (spam) και επιθυμητή (όχι-spam) ηλεκτρονική αλληλογραφία. Ας υποθέσουμε, τώρα, ότι τα έγγραφα κειμένου επιλέγονται από έναν αριθμό από κατηγορίες εγγράφων κειμένου, τα οποία μπορούν να μοντελοποιηθούν ως σύνολα από λέξεις, όπου η (ανεξάρτητη) πιθανότητα ότι η  $i$ -οστη λέξη ενός δοσμένου εγγράφου εμφανίζεται σε ένα έγγραφο της κατηγορίας  $C$  μπορεί να γραφεί ως:<sup>1</sup>

$$p(w_i|C) \quad (2.1)$$

Τότε, η πιθανότητα ενός δοσμένου εγγράφου  $D$ , με δεδομένη μία κατηγορία  $C$ , είναι

$$p(D|C) = \prod_i p(w_i|C) \quad (2.2)$$

Η ερώτηση που θέλουμε να απαντήσουμε είναι: “ποια είναι η πιθανότητα ότι ένα δοσμένο έγγραφο  $D$  ανήκει σε μία δεδομένη κατηγορία  $C$ ”; Με άλλα λόγια, ποια είναι η  $p(C|D)$ ; Σύμφωνα με τον ορισμό τους,

$$p(D|C) = \frac{p(D \cap C)}{p(C)} \quad (2.3)$$

<sup>1</sup>Σημ.: ακολουθούμε μία απλοποιημένη θεώρηση, υποθέτοντας ότι η πιθανότητα μιας λέξης σε ένα έγγραφο κειμένου είναι ανεξάρτητη από το μήκος του εγγράφου ή αλλιώς ότι όλα τα έγγραφα κειμένου έχουν το ίδιο μήκος

και

$$p(C|D) = \frac{p(D \cap C)}{p(D)} \quad (2.4)$$

Το θεώρημα του Μπέυζ χειραγωγεί τα παραπάνω σε μία δήλωση πιθανότητας:

$$p(C|D) = \frac{p(C)}{p(D)} p(D|C) \quad (2.5)$$

Θεωρώντας για λίγο ότι υπάρχουν μόνο δύο κατηγορίες  $S$  και  $\neg S$  (π.χ. *spam* και *όχι spam*), θα έχουμε:

$$p(D|S) = \prod_i p(w_i|S) \quad (2.6)$$

και

$$p(D|\neg S) = \prod_i p(w_i|\neg S) \quad (2.7)$$

Χρησιμοποιώντας το παραπάνω αποτέλεσμα, μπορούμε να γράψουμε:

$$p(S|D) = \frac{p(S)}{p(D)} \prod_i p(w_i|S) \quad (2.8)$$

$$p(\neg S|D) = \frac{p(\neg S)}{p(D)} \prod_i p(w_i|\neg S) \quad (2.9)$$

Διαχωρίζοντας το ένα από το άλλο, έχουμε:

$$\frac{p(S|D)}{p(\neg S|D)} = \frac{p(S)}{p(\neg S)} \frac{\prod_i p(w_i|S)}{\prod_i p(w_i|\neg S)} \quad (2.10)$$

το οποίο μπορεί να ξαναγραφτεί ως:

$$\frac{p(S|D)}{p(\neg S|D)} = \frac{p(S)}{p(\neg S)} \prod_i \frac{p(w_i|S)}{p(w_i|\neg S)} \quad (2.11)$$

Συνεπώς, ο πιθανοτικός λόγος  $p(S|D)/p(\neg S|D)$  μπορεί να εκφραστεί με τη μορφή μιας σειράς από πιθανοτικούς λόγους. Η πραγματική πιθανότητα  $p(S|D)$  μπορεί να υπολογιστεί με ευκολία από το  $\log(p(S|D)/p(\neg S|D))$ , με βάση την παρατήρηση ότι  $p(S|D) + p(\neg S|D) = 1$ .

Παίρνοντας τον λογάριθμο όλων των λόγων, έχουμε:

$$\ln \frac{p(S|D)}{p(\neg S|D)} = \ln \frac{p(S)}{p(\neg S)} + \sum_i \ln \frac{p(w_i|S)}{p(w_i|\neg S)} \quad (2.12)$$

Τελικά, το έγγραφο μπορεί να ταξινομηθεί ως ακολούθως: είναι *spam* αν  $\ln \frac{p(S|D)}{p(\neg S|D)} > 0$ , διαφορετικά δεν είναι *spam*.

Βέβαια, στο σημείο αυτό θα πρέπει να τονίσουμε ότι παρά τις διάφορες ιδιότητες, που καθιστούν έναν αφελή ταξινομητή Μπέυζ ιδιαίτερα δημοφιλή για πρακτικά προβλήματα, πολλές φορές οι υπερβολικές υποθέσεις/απαιτήσεις ανεξαρτησίας του τον καθιστούν ανακριβή για προβλήματα που σχετίζονται με την πραγματική ζωή. Ανάλογα, όμως, με την ακριβή φύση του πιθανοτικού μοντέλου, που καλούμαστε να εφαρμόσουμε σε κάθε πρόβλημα, ο αφελής ταξινομητής Μπέυζ μπορεί να εκπαιδευθεί ιδιαίτερα αποδοτικά στο πλαίσιο ενός μοντέλου επιβλεπόμενης μάθησης, το οποίο εφαρμόζεται για την αντιμετώπιση του προβλήματος κατηγοριοποίησης εγγράφων κειμένου.

### 2.2.8 Μπεϋζιανά δίκτυα

Ένα Μπεϋζιανό δίκτυο είναι μια μορφή πιθανοτικού γράφου επίσης γνωστή και ως Μπεϋζιανό δίκτυο πεποίθησης ή απλώς ως δίκτυο πεποίθησης. Ένα Μπεϋζιανό δίκτυο μπορεί να αναπαρασταθεί με ένα γράφο (όπως στη θεωρία γράφων) με συνημμένες τις πιθανότητές του. Με τον τρόπο αυτό ένα Μπεϋζιανό δίκτυο αντιπροσωπεύει ένα σύνολο μεταβλητών μαζί με μια συνδυασμένη κατανομή πιθανότητας με ρητές υποθέσεις ανεξαρτησίας. Ειδικότερα, ένα Μπεϋζιανό δίκτυο είναι ένας κατευθυνόμενος ακυκλικός γράφος, του οποίου οι κόμβοι αντιπροσωπεύουν τις μεταβλητές και τα τόξα μεταξύ των κόμβων, τις στατιστικές σχέσεις εξάρτησης μεταξύ των μεταβλητών και τις τοπικές κατανομές πιθανότητας για κάθε μεταβλητή. Οι κόμβοι μπορούν να αντιπροσωπεύουν οποιοδήποτε είδος μεταβλητής, είτε πρόκειται για μια μετρήσιμη παράμετρο, είτε για μια λανθάνουσα μεταβλητή ή μια υπόθεση.

Εάν υπάρχει ένα τόξο από τον κόμβο  $A$  σε έναν άλλο κόμβο  $B$ , τότε η μεταβλητή  $B$  εξαρτάται άμεσα από τη μεταβλητή  $A$  και το  $A$  καλείται γονέας του  $B$ . Εάν για κάθε μεταβλητή  $X_i, i \in \{1, \dots, n\}$  το σύνολο των μεταβλητών γονέων δίνεται από τους  $\text{parents}(X_i)$ , τότε η συνδυασμένη κατανομή των μεταβλητών είναι προϊόν των τοπικών κατανομών

$$Pr(X_1, \dots, X_n) = \prod_{i=1}^n Pr(X_i | \text{parents}(X_i)) \quad (2.13)$$

Αν το  $X_i$  δεν έχει κανέναν γονέα, η τοπική κατανομή πιθανότητάς του λέγεται ότι είναι απεριόριστη, διαφορετικά είναι υπό όρους. Εάν η μεταβλητή που αντιπροσωπεύεται από έναν κόμβο παρατηρείται, τότε ο κόμβος αυτός λέγεται ότι είναι ένας κόμβος στοιχείων. Τέλος, το κύριο πλεονέκτημα των Μπεϋζιανών δικτύων είναι ότι είναι διαισθητικά ευκολότερο για έναν άνθρωπο να καταλάβει τις άμεσες εξαρτήσεις και τις τοπικές κατανομές από την πλήρη συνδυασμένη κατανομή και είναι αυτό που τα κάνει ιδιαίτερα δημοφιλή στις εφαρμογές.

## 2.3 Ανάλυση Πολυμεσικού Περιεχομένου Υποστηριζόμενη από τη Γνώση και το Εννοιολογικό Πλαίσιο

Η ανάλυση πολυμεσικού περιεχομένου είναι ένας ιδιαίτερα απαιτητικός χώρος έρευνας στο πεδίο της επεξεργασίας εικόνας και βίντεο και της όρασης υπολογιστών. Πράγματι, αποτελεί ένα από τα κρισιμότερα βήματα προς την κατεύθυνση της κατανόησης της εικόνας ή του βίντεο, που είναι εδώ και καιρό ένας από τους βασικούς απώτερους στόχους των σχετικών ερευνητικών προσπαθειών [143]. Ποικίλες εφαρμογές, όπως η αναγνώριση αντικειμένων, ο χαρακτηρισμός εικόνων, η κωδικοποίηση και η ευρετηρίαση εικόνων, χρησιμοποιούν σε κάποιο σημείο έναν αλγόριθμο ανάλυσης και η απόδοσή τους εξαρτάται ιδιαίτερα από την ποιότητα των τελευταίων. Στη συνέχεια θα ορίσουμε τον ερευνητικό χώρο, με τον οποίο καταπιάνεται η ανάλυση πολυμεσικού περιεχομένου, και θα παρουσιάσουμε εν συντομίᾳ τις βασικές αρχές και τα προβλήματά της. Επίσης, θα δούμε τον τρόπο με τον οποίο αντιμετωπίσαμε στο πλαίσιο των ερευνών μας το γενικότερο πρόβλημα μέσω της χρήσης της γνώσης και του εννοιολογικού πλαισίου στη διαδικασία. Το τελευταίο αποτελεί βέβαια και ένα εκ των βασικών αντικειμένων της παρούσης διατριβής και κατά συνέπεια θα αναλυθεί επισταμένως σε επόμενα κεφάλαιά της.

### 2.3.1 Ανάλυση πολυμεσικού περιεχομένου - μια περιληπτική σύνοψη

Είναι γεγονός ότι εδώ και αρκετές δεκαετίες οι σχετικές έρευνες έχουν προτείνει κάποιους αποδοτικούς αλγορίθμους για την αυτόματη κατάτμηση εικόνας και βίντεο [94], [35], [200], καθώς επίσης και τη δόμηση του πολυμεσικού περιεχομένου [61], [22]. Σύντομα, όμως, έγινε κατανοητό ότι η απλή ανάλυση εικόνας και βίντεο, η οποία δεν εκμεταλλεύεται κάποια μορφή γνώσης στην πορεία, δεν είναι σε θέση να υπερινησει προβλήματα, όπως μεταξύ άλλων είναι η εύκαμπτη κίνηση αντικειμένων και η υπερκατάτμηση και να φθάσει το απαραίτητο επίπεδο αποδοτικότητας, ακρίβειας και ευρωστίας που απαιτείται για τις εφαρμογές.

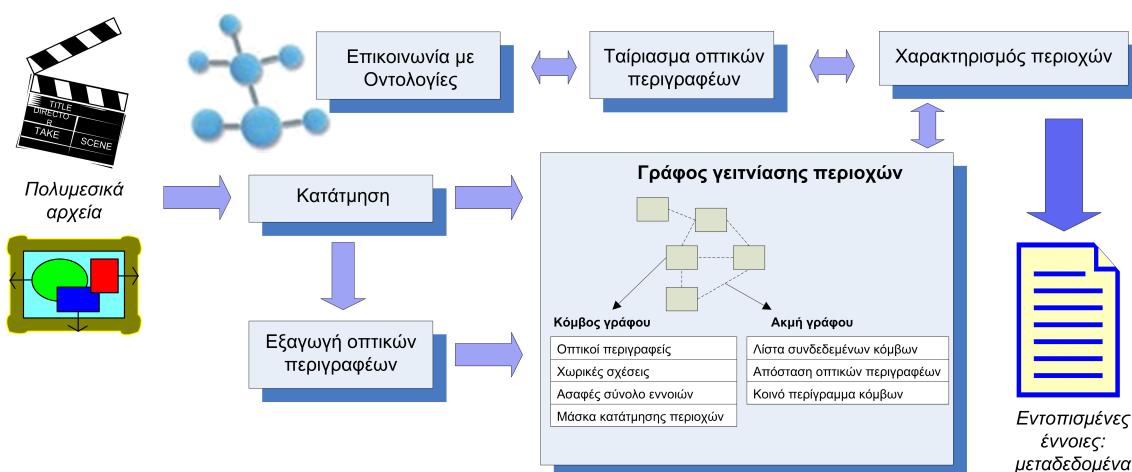
Στην περιοχή της ανάλυσης πολυμέσων η εκ των προτέρων γνώση για τις θεματικές περιοχές [125] αποτελεί μια ελπιδοφόρα προσέγγιση, σύμφωνα με την οποία η σημασιολογία υψηλότερου επιπέδου μπορεί να ενσωματωθεί στις τεχνικές αυτόματης ανάλυσης του πολυμεσικού περιεχομένου [232]. Οι τελευταίες ολοένα και περισσότερο απευθύνονται στις προσεγγίσεις διαχείρισης γνώσης, συμπεριλαμβανομένων των τεχνολογιών του Σημασιολογικού Ιστού, προκειμένου να επιλύσουν το πρόβλημα αυτό [284]. Εκ των προτέρων καθορισμένα μοντέλα γνώσης χρησιμοποιούνται ως βάσεις γνώσης, οι οποίες συνεισφέρουν στη βασισμένη στη σημασιολογία ταξινόμηση και συσταδοποίηση [170]. Στη δημοσίευση [250] χρησιμοποιούνται οι σημασιολογικές οντότητες, όπως αυτές καθορίζονται στο πρότυπο MPEG-7, για την υποβοηθούμενη από τη γνώση ανάλυση βίντεο και την ανίχνευση αντικειμένων, επιτρέποντας κατά συνέπεια τη σημασιολογική ευρετηρίαση. Στο άρθρο [171] αντιστοιχίζονται αυτόματα συμβατοί με το MPEG-7 χαμηλού επιπέδου περιγραφές σε κατάλληλους ενδιάμεσου επιπέδου περιγραφές, διαμορφώνοντας μία οντολογία απλών λεξικογραφικών όρων. Επιπρόσθετα, στο [144] παρουσιάζεται μια οντολογία αντικειμένων, προκειμένου να διευκολύνει την αντιστοίχιση των χαμηλού επιπέδου στα υψηλού επιπέδου χαρακτηριστικά γνωρίσματα και να επιτρέψει τον καθορισμό των σχέσεων μεταξύ των τυμάτων της πολυμεσικής πληροφορίας. Αυτό το τελευταίο παράδειγμα οντολογίας συνδέεται και με έναν μηχανισμό συσχετιστικής ανατροφοδότησης (relevance feedback), ώστε να επιτρέψει την επίτευξη υψηλής ακρίβειας κατά την ανάκτηση του επιθυμητού περιεχομένου. Όσον αφορά στον σημασιολογικό χαρακτηρισμό [109], οι μέχρι τώρα έρευνες έχουν κυρίως εστιάσει στα κείμενα [110] ή τον απλό χαρακτηρισμό φωτογραφιών [277], [120]. Η αυτόματη υποστήριξη των χρηστών στην επέκταση του χαρακτηρισμού εικόνων με πρόσθετες χωρικές πληροφορίες συζητείται στο [121]. Τέλος, μια μάλλον απλή προσέγγιση για την ανάλυση και το χαρακτηρισμό ακολουθιών βίντεο με χρήση των τεχνολογιών του Σημασιολογικού Ιστού παρουσιάζεται στο [71].

Τελευταία, η έρευνα έχει στραφεί στο συνδυασμό γενικών χαρακτηριστικών γνωρισμάτων της εικόνας με τοπικά οπτικά χαρακτηριστικά γνωρίσματα, έτσι ώστε να μπορέσει να επιλύσει τις ασάφειες μεταξύ των περιοχών που εμφανίζονται κατά την κατάτμηση μιας εικόνας. Στο [160] χρησιμοποιείται ένα Μπεϋζιανό δίκτυο για την ενσωμάτωση της γνώσης από το χαμηλό επίπεδο σε μέσου επιπέδου χαρακτηριστικά γνωρίσματα με απώτερο στόχο την ταξινόμηση των εικόνων σε δύο κατηγορίες: εσωτερικού και εξωτερικού χώρου. Στο [41] προτείνεται μια ταξινόμηση πολλαπλών ετικετών, για να περιγράψει τις φυσικές σκηνές που ανήκουν συνήθως σε πολλαπλές σημασιολογικές κατηγορίες, γεγονός που προσομοιάζει πολύ μια προσέγγιση βασισμένη στην ασαφή λογική. Η ασαφής ταξινόμηση υιοθετήθηκε στο [216] για την ταξινόμηση εικόνας με βάση τις περιοχές της. Στο [192] η εκμάθηση ενός Μπεϋζιανού

δικτύου με χρήση ταξινομητών στους κόμβους του αντιστοιχίζει τη χαμηλού επιπέδου αναπαράσταση ακολουθιών βίντεο σε σημασιολογία υψηλού επιπέδου, προκειμένου να επιτευχθεί αποδοτική ευρετηρίαση, φιλτράρισμα και ανάκτηση του βίντεο. Τέλος, η γνώση των σχετικών χωρικών θέσεων μεταξύ των περιοχών της εικόνας χρησιμοποιείται στο [174] ως επιπρόσθετος περιορισμός, βελτιώνοντας έτσι τη μεμονωμένη αναγνώριση των περιοχών.

### 2.3.2 Αρχιτεκτονική ενός συστήματος ανάλυσης πολυμεσικού περιεχομένου

Η πρόθεσή μας σε αυτήν την εργασία είναι να λειτουργήσουμε σε ένα σημασιολογικό επίπεδο, όπου οι περιοχές μιας εικόνας συνδέονται με έναν αριθμό από πιθανές υψηλού επιπέδου ετικέτες εννοιών και όχι μόνο με τα χαμηλού επιπέδου οπτικά χαρακτηριστικά γνωρίσματά τους. Κατά συνέπεια, πρέπει να χρησιμοποιηθεί μια αναπαράσταση γνώσης, για να αποθηκεύσει τόσο την χαμηλού επιπέδου, όσο και την υψηλού επιπέδου σημασιολογική πληροφορία. Στο πλαίσιο προηγούμενων εργασιών έχει αναγνωρισθεί ένα σύνολο από κατάλληλους οπτικούς περιγραφείς MPEG-7, όπως π.χ. το κυριαρχού χρώμα (Dominant Color - DC) και η ομοιογενής υφή (Homogeneous Texture - HT) [164], οι οποίοι χρησιμοποιούνται, για να αντιπροσωπεύσουν κάθε περιοχή της εικόνας με ένα διάνυσμα χαρακτηριστικών γνωρισμάτων, ενώ ασαφή σύνολα υποψηφίων εννοιών χρησιμοποιούνται, για να μοντελοποιήσουν την υψηλού επιπέδου πληροφορία. Για αυτό το λόγο, στο πλαίσιο της παρούσας χρησιμοποιήθηκε ένας ευρέως αποδεκτός αλγόριθμος ανάλυσης πολυμεσικού περιεχομένου, ο οποίος συζητείται σε βάθος στο [9] και έχει εφαρμοσθεί με επιτυχία σε ποικίλες ερευνητικές και μη εφαρμογές. Η βασική αρχιτεκτονική του συνολικού διαγράμματος ροής της διαδικασίας ανάλυσης του πολυμεσικού περιεχομένου παρουσιάζεται στο Σχήμα 2.1.



Σχήμα 2.1: Δομικό διάγραμμα ανάλυσης πολυμεσικού περιεχομένου, υποβοηθούμενης από κάποιας μορφής γνώση.

Στο κέντρο του παραπάνω διαγράμματος υπάρχει ένας γράφος γειτνίασης περιοχών, ο οποίος χρησιμοποιείται ως αναπαράσταση της εικόνας κατά τη διάρκεια της διαδικασίας της ανάλυσης. Κάθε χορυφή του γράφου αντιστοιχεί σε μία συνδεδεμένη περιοχή της υπό εξέταση εικόνας, ενώ μια ακμή του γράφου αναπαριστά το σύνδεσμο μεταξύ δύο περιοχών και είναι αυτή που περιέχει όλη την πληροφορία γειτνίασης [9].

Έχοντας καθορίσει τη δομή αυτή για την αναπαράσταση μιας εικόνας, ακολουθούν οι αλγόριθμοι που κάνουν την πραγματική επεξεργασία, αλληλεπιδρώντας δυναμικά (δηλ. αρχικοποιώντας, ενημερώνοντας και διαβάζοντας) το γράφο.

Σε κάθε περίπτωση, το πρώτο βήμα της διαδικασίας ανάλυσης θα πρέπει να είναι ένας αλγόριθμος κατάτμησης της εικόνας, ο οποίος θα παράσχει μερικές δεκάδες από συνδεδεμένες περιοχές και θα αρχικοποιήσει τον γράφο αναπαράστασης. Η κατάτμηση μπορεί να χρησιμοποιήσει εν γένει οποιονδήποτε αλγόριθμο κατάτμησης, στο πλαίσιο όμως των δικών μας ερευνών και, όπως θα παρουσιασθεί αναλυτικότερα στο κεφάλαιο 6, χρησιμοποιήσαμε μια επέκταση του ευρέως γνωστού και εφαρμοσμένου αλγόριθμου RSST, η οποία βασίζεται σε ένα νέο μοντέλο χρωμάτων και συντακτικών χαρακτηριστικών γνωρισμάτων [1]. Ο ίδιος ο γράφος παρέχει τις πληροφορίες για το κατά πόσο δύο περιοχές είναι γειτονικές ή όχι, αλλά επιπρόσθετα θα πρέπει να γνωρίζουμε και τη χωρική τους σχέση· για παράδειγμα, η περιοχή  $X$  βρίσκεται πάνω από την περιοχή  $Y$ , ή μια απόλυτη σχέση, όπως π.χ. η περιοχή  $Z$  βρίσκεται κάτω από όλες τις περιοχές. Κατόπιν, εξάγουμε για κάθε περιοχή τους κατά MPEG-7 οπτικούς περιγραφείς χυρίαρχου χρώματος και μορφής περιοχής και τους αποθηκεύουμε στην αντίστοιχη κορυφή του γράφου. Το επόμενο βήμα είναι να υπολογιστεί για κάθε περιοχή η απόσταση μεταξύ των δύο οπτικών περιγραφέων της, καθώς και των αντίστοιχων περιγραφέων όλων των πρωτοτύπων (instances) των εννοιών στην αντίστοιχη οντολογία. Το αποτέλεσμα αυτού του ταιριάσματος είναι μια απόσταση για κάθε περιγραφέα, η οποία όμως από μόνη της δεν είναι πολύ χρήσιμη, καθώς απαιτείται μια εν γένει μοναδική, συνδυασμένη απόσταση.

Προς αυτήν την κατεύθυνση, εξετάσαμε τόσο την απλή προσέγγιση ενός σταθμισμένου αθροίσματος των δύο αποστάσεων, όσο και την περίπτωση της εκπαίδευσης ενός νευρωνικού δικτύου [237] και μπορούμε να πούμε ότι και οι δύο προσεγγίσεις λειτουργούν απροβλημάτιστα, όταν χρησιμοποιούνται μόνο δύο οπτικοί περιγραφείς. Χρησιμοποιώντας μια χαρακτηριστική συνάρτηση κανονικοποίησης, μετατρέπουμε, εν συνεχείᾳ, την απόσταση σε ένα βαθμό σχετικότητας, ο οποίος και αποτελεί από εδώ και πέρα το κριτήριο ομοιότητας για όλες τις διαδικασίες ταιριάσματος και συγχωνεύσεων. Τελικά, από την παραπάνω διαδικασία προκύπτει ένας κατάλογος πιθανών εννοιών μαζί με έναν βαθμό σχετικότητας για όλες τις περιοχές της εικόνας και αποθηκεύεται κατάλληλα στο γράφο.

Στην περίπτωση που δύο ή περισσότερες γειτονικές περιοχές έχουν αποδοθεί σε μία μόνο έννοια, ή άλλες πιθανές έννοιες έχουν έναν βαθμό μικρότερο από ένα προκαθορισμένο κατώτατο όριο, υποθέτουμε ότι αυτές οι περιοχές είναι μέρος μιας μεγαλύτερης περιοχής, που δεν κατατμήθηκε σωστά, χυρίως λόγω των γνωστών περιορισμών που διέπουν την διαδικασία της κατάτμησης. Η προτεινόμενη λύση για το πρόβλημα αυτό είναι η συγχώνευση όλων εκείνων των περιοχών που εμφανίζουν τα παραπάνω συμπτώματα. Με άλλα λόγια, διορθώνουμε το παραπάνω φαινόμενο με τη συγχώνευση κορυφών του γράφου και την ενημέρωση όλων των σχετικών πεδίων του γράφου (δηλ. εξάγοντας και πάλι τους οπτικούς περιγραφείς, ενημερώνοντας το περίγραμμα της περιοχής, αναπροσαρμόζοντας τις ακμές του γράφου, κ.ο.κ..). Σε κάθε περίπτωση, το τελικό αποτέλεσμα για κάθε περιοχή της εικόνας είναι ο κατάλογος των πιθανών εννοιών της μαζί με τους εκάστοτε βαθμούς σχετικότητάς/εμπιστοσύνης τους.

### 2.3.3 Ο ρόλος του εννοιολογικού πλαισίου

Όποτε παρέχεται καινούριο πολυμεσικό περιεχόμενο ως είσοδος της διαδικασίας ανάλυσης, η εκ των προτέρων υπάρχουσα βάση γνώσης χρησιμοποιείται, για να συγχρίνει (με τη βοήθεια του παραπάνω ταιριάσματος των οπτικών περιγραφέων MPEG-7) κάθε περιοχή του γράφου με τα πρωτότυπα στιγμιότυπα των πολυμεσικών θεματικών περιοχών από την οντολογία. Για αυτόν τον λόγο, το εκάστοτε σύστημα ανάλυσης πρέπει να έχει πλήρη πρόσβαση στη γενική βάση γνώσης, που αποτελείται από όλες τις περιπτώσεις πρωτοτύπων εννοιών ανά θεματική περιοχή. Αυτές οι περιπτώσεις εφαρμόζονται ως αναφορές στους αλγορίθμους ανάλυσης και με τη βοήθεια των κατάλληλων κανόνων που σχετίζονται με τις υποστηριζόμενες θεματικές περιοχές, η υποβοηθούμενη από τη γνώση ανάλυσης εξάγει τις σημασιολογικές έννοιες, που σχετίζονται με τις συγκεκριμένες περιοχές της εικόνας ή της ακολουθίας βίντεο.

Η ανωτέρω διαδικασία καταλήγει σε ένα αρχικό ασαφές μαρκάρισμα όλων των περιοχών της εικόνας με έννοιες από τη βάση γνώσεων ή με άλλα λόγια καταλήγει σε ένα σύνολο, του οποίου τα στοιχεία είναι τα ασαφή σύνολα όλων των περιοχών στην εικόνα. Προφανώς, κάτι τέτοιο δεν αποτελεί έναν απλό και εύκολο στόχο και η αποδοτικότητά της μεθόδου εξαρτάται ιδιαίτερα από τη θεματική περιοχή, στην οποία εφαρμόζεται, καθώς επίσης και από την ποιότητα της βάσης γνώσης. Οι κύριοι περιορισμοί αυτής της προσέγγισης είναι η εξάρτηση από την αρχική κατάτμηση και τη δημιουργία των αντιπροσωπευτικών περιπτώσεων πρωτοτύπων των εννοιών. Η τελευταία διαπίστωση είναι εν γένει ευκολότερο να αντιμετωπιστεί, ενώ, όπως θα δούμε στο πλαίσιο της παρούσης, η πρώτη διαπίστωση θα αντιμετωπιστεί με μια επέκταση κάποιων παραδοσιακών αλγορίθμων κατάτμησης, η οποία στηρίζεται στην κατάτμηση των περιοχών με βάση ένα σημασιολογικό επίπεδο και στη χρήση της ασαφούς λογικής.

Στη συνέχεια με τον όρο (οπτικό) εννοιολογικό πλαίσιο θα αναφερόμαστε σε όλες εκείνες τις πληροφορίες που είναι σχετικές με το οπτικό περιεχόμενο της σκηνής μιας ακίνητης εικόνας ή μιας ακολουθίας βίντεο και που μπορούν να φανούν χρήσιμες για την ανάλυσή της. Το εννοιολογικό πλαίσιο σχετίζεται άμεσα με δύο προβλήματα της ανάλυσης εικόνας, την ταξινόμηση σκηνής (scene classification) και την ανίχνευση αντικειμένων (object detection). Από τη μία, η ταξινόμηση σκηνής είναι μια προσέγγιση από επάνω προς τα κάτω, όπου τα χαμηλού επιπέδου οπτικά χαρακτηριστικά γνωρίσματα υιοθετούνται, για να αναλύσουν συνολικά το περιεχόμενο της σκηνής και να την ταξινομήσουν σε μια εκ των προκαθορισμένων κατηγοριών (π.χ. εσωτερικός/υπαίθριος χώρος, πόλη/φύση, και λοιπά). Από την άλλη, η ανίχνευση (ή αναγνώριση) αντικειμένων είναι μια από κάτω προς τα επάνω προσέγγιση, η οποία εστιάζει στην τοπική ανάλυση, για να ανιχνεύσει και να αναγνωρίσει τα συγκεκριμένα αντικείμενα σε περιορισμένες περιοχές μιας εικόνας, χωρίς να διαθέτει κάποια ρητή γνώση του περιβάλλοντος πλαισίου (π.χ. να αναγνωρίσει ένα κτίριο ή ένα δέντρο). Οι δύο αυτοί σημαντικοί τομείς της ανάλυσης εικόνας περιλαμβάνουν ένα πρόβλημα που μοιάζει με το γνωστό ερώτημα “η κότα έκανε το αυγό ή το αυγό την κότα;”. Κι αυτό γιατί, για παράδειγμα, η ανίχνευση ενός κτιρίου στη μέση μιας εικόνας μπορεί να υπονοήσει ότι η εικόνα αυτή είναι μια εικόνα πόλης με μια αρκετά υψηλή πιθανότητα, ενώ η προ-ταξινόμηση της εικόνας αυτής ως “πόλη” θα ευνοούσε την αναγνώριση ενός κτιρίου σε αντιδιαστολή με ένα δέντρο.

Το θέμα της αυτόματης ανίχνευσης των σημαντικών ή ενδιαφερουσών περιοχών σε μια εικόνα έχει αντιμετωπιστεί από διάφορους ερευνητές κατά τη διάρκεια της τελευταίας δεκαετίας [288], [198], [247]. Για παράδειγμα, μια τέτοια μεθοδολογία

παρατίθεται στο [173], όπου αναπτύσσεται ένα υπολογιστικό μοντέλο οπτικής προσοχής, με το να συνδυάζεται η γνώση για το ανθρώπινο οπτικό σύστημα με τις τεχνικές όρασης των υπολογιστών. Η έννοια της απλής κατάτμησης περιοχών στοχεύει, όπως είδαμε, στην παραγωγή περιοχών της εικόνας, οι οποίες θα διέπονται από ομοιογενείς ιδιότητες, όπως π.χ. το χρώμα ή/και η υφή, με στόχο την υψηλότερου επιπέδου εξαγωγή συμπερασμάτων και ερμηνεία τους. Προφανώς, δε χωράει αμφιβολία ότι το εν λόγω θέμα θεωρείται γενικά ως κάτι που ένας ανθρώπινος παρατηρητής εκτελεί με σχετική ευκολία, αλλά συγχρόνως είναι δύσκολο να γίνει κατανοητό και κατά συνέπεια αντιμετωπίσουμε για ένα αυτοματοποιημένο σύστημα ή υπολογιστή.

Είναι γεγονός ότι στο πλαίσιο της αναζήτησης και ανάκτησης εικόνας, που είναι βασισμένη στο περιεχόμενο, όλο και περισσότεροι ερευνητές κοιτάζουν πέρα από τα χαμηλού επιπέδου χαρακτηριστικά, όπως είναι το χρώμα, η υφή και τα χαρακτηριστικά γνωρίσματα μορφής, καθώς αναζητούν αποτελεσματικότερες μεθόδους έρευνας. Η ανίχνευση φυσικών αντικειμένων στις εσωτερικές ή υπαίθριες σκηνές, δηλ. ο προσδιορισμός βασικών τύπων αντικειμένων, όπως π.χ. ο ουρανός, η χλόη, το φύλλωμα, το νερό ή το χιόνι, μπορεί να διευκολύνει τις εφαρμογές που βασίζονται στο περιεχόμενο και που ποικίλουν από τη βελτίωση εικόνων ως την κωδικοποίηση ή και άλλες χρήσιμες εφαρμογές πολυμέσων. Εντούτοις, σε όλα τα παραπάνω εμφανίζεται ένας σημαντικός αριθμός λανθασμένων ταξινομήσεων. Αυτό συμβαίνει συνήθως εξαιτίας των ομοιοτήτων, που απαντώνται στα χαρακτηριστικά χρώματος και υφής των διάφορων τύπων αντικειμένων, καθώς και εξαιτίας της παντελούς έλλειψης πληροφοριών εννοιολογικού πλαισίου, η οποία και αποτελεί έναν σημαντικό περιορισμό όλων των μεμονωμένων ανιχνευτών αντικειμένων. Μια ενδιαφέρουσα προσέγγιση για την επίλυση του τελευταίου προβλήματος είναι αυτή που παρουσιάζεται στο [161]. Ειδικότερα, προτείνεται η χρήση ενός χωρικού συστήματος ανίχνευσης αντικειμένων, το οποίο μπορεί να υποστηριχθεί ότι λαμβάνει υπόψη του μια μορφή εννοιολογικού πλαισίου, συνδυάζοντας αρχικά την παραγωγή των μεμονωμένων ανιχνευτών αντικειμένου, προκειμένου να παραχθεί ένα σύνθετο διάνυσμα πεποιθησης για τα αντικείμενα που είναι ενδεχομένως παρόντα σε μια εικόνα. Στη συνέχεια, χρησιμοποιούνται χωρικοί περιορισμοί εννοιολογικού πλαισίου, για να μειώσουν τη λανθασμένη ταξινόμηση, με το να περιορίζουν τις αντίστοιχες πεποιθήσεις. Οι χωρικοί αυτοί περιορισμοί έχουν τη μορφή συναρτήσεων πυκνότητας πιθανότητας, οι οποίες και δημιουργούνται με χρήση τεχνικών εκμάθησης.

Άλλες ερευνητικές προσπάθειες στην περιοχή περιλαμβάνουν αυτή που προτείνεται στο [191], όπου ένας κατάλογος σημασιολογικών αντικειμένων, συμπεριλαμβανομένων των: ουρανός, χιόνι, βράχος, νερό και δάσος, χρησιμοποιείται για τη σημασιολογική ευρετηρίαση και ανάκτηση ακολουθιών βίντεο. Επιπλέον, το χρώμα αποτελεί ένα από τα κεντρικά χαρακτηριστικά γνωρίσματα των υπαρχόντων εργασιών για την ανίχνευση φυσικών αντικειμένων. Παραδείγματος χάριν, στο [210] χρησιμοποιείται η ταξινόμηση με βάση το χρώμα, προκειμένου να ανιχνευθεί ο ουρανός. Στο πλαίσιο της ανάκτησης εικόνων με βάση το περιεχόμενο ο Smith και ο Li [233] υπέθεσαν ότι μία εκτεταμένη μπλε περιοχή στην κορυφή μιας εικόνας είναι πιθανό να αντιπροσωπεύει τον καθαρό ουρανό. Επίσης, πιο πρόσφατα παρουσιάστηκε μια προσέγγιση βασισμένη σε υποδείγματα, η οποία χρησιμοποιεί έναν συνδυασμό χαρακτηριστικών γνωρισμάτων χρώματος και υφής, για να ταξινομήσει μικρά μπλοκ μιας υπαίθριας σκηνής ως ουρανό ή βλάστηση, υποθέτοντας ότι η εικόνα διαθέτει το σωστό προσανατολισμό [262]. Κάτι τέτοιο θέτει σαφώς το ζήτημα της χρησιμοποίησης των πληροφοριών προσανατολισμού από το εννοιολογικό πλαίσιο μέσα στους αλγορίθμους ανίχνευσης της

κατηγορίας ενός αντικειμένου, στόχος, όμως, που αποφεύγεται γενικά εξαιτίας του γεγονότος ότι τέτοιες πληροφορίες που βασίζονται στο εννοιολογικό πλαίσιο δεν είναι πάντα διαθέσιμες.

Μέχρι τώρα, καμία από τις ανωτέρω μεθόδους και τεχνικές δε χρησιμοποιεί ρητά το εννοιολογικό πλαίσιο. Αυτό τείνει να είναι το κύριο μειονέκτημα αυτών των μεμονωμένων ανιχνευτών αντικειμένων, δεδομένου ότι εξετάζουν μόνο τις απομονωμένες λωρίδες των καθαρών υλικών του αντικειμένου, χωρίς να λαμβάνουν υπόψη τους το εννοιολογικό πλαίσιο της σκηνής ή των ίδιων των μεμονωμένων αντικειμένων. Αυτό είναι πολύ σημαντικό και εξαιρετικά προκλητικό, ακόμη και για τους ανθρώπινους παρατηρητές. Η έννοια του οπτικού εννοιολογικού πλαισίου είναι σε θέση να βοηθήσει στην κατεύθυνση των μεθοδολογιών ανίχνευσης φυσικών αντικειμένων, μιμούμενη την ανθρώπινη προσέγγιση σε παρόμοια προβλήματα. Πολλά υλικά αντικειμένων έχουν την ίδια εμφάνιση, όταν τα εξετάζει κανείς αποκλειστικά από την άποψη του χρώματος και της υφής τους, αλλά το ίδιο αντικείμενο μπορεί να εμφανίζεται εντελώς διαφορετικό κάτω από διαφορετικές συνθήκες απεικόνισης (όπως π.χ. είναι ο φωτισμός, η μεγέθυνση, η απόσταση κ.α.). Αν μη τι άλλο, ένα από τα πιο σημαντικά γνωρίσματα των ανθρώπων είναι ότι εξετάζουν όλα τα αντικείμενα σε μία σκηνή, πριν λάβουν μια τελική απόφαση σχετικά με την ταυτότητα των μεμονωμένων αντικειμένων. Όπως θα δούμε παρακάτω (κεφ. 6), η χρήση του οπτικού εννοιολογικού πλαισίου αποτελεί το κλειδί αυτής της σαφέστατης διαδικασίας αναγνώρισης.

## 2.4 Πρόσβαση και Ανάκτηση Πολυμεσικής Πληροφορίας

Η ανάκτηση πληροφορίας (Information Retrieval) είναι εν γένει η επιστήμη, η οποία ασχολείται με την αναζήτηση πληροφορίας σε πολυμεσικά ή μη έγγραφα, με την αναζήτηση των ίδιων των πολυμεσικών ή μη εγγράφων, με την αναζήτηση μεταδεδομένων, που περιγράφουν αυτά τα έγγραφα, ή ακόμα και με αναζήτηση κειμένου, ήχου, εικόνων και δεδομένων μέσα σε διαφόρων ειδών βάσεις δεδομένων, όπως π.χ. οι σχεσιακές βάσεις δεδομένων ή/και οι δικτυωμένες βάσεις δεδομένων υπερκειμένων, όπως το Διαδίκτυο ή ένα μεγάλων διαστάσεων εσωτερικό δίκτυο (intranet). Συνακόλουθα, υπάρχει και ένα είδος σύγχυσης αναφορικά με τις θεματικές περιοχές της ανάκτησης δεδομένων, της ανάκτησης εγγράφων, της ανάκτησης πληροφοριών και της ανάκτησης κειμένων, καθώς κάθε μία από αυτές διαθέτει τη δική της βιβλιογραφία, θεωρία, πράξη και τεχνολογία.

Τα αυτοματοποιημένα συστήματα IR χρησιμοποιούνται ως επί το πλείστον για τη μείωση του φαινομένου της υπερφόρτωσης πληροφοριών (information overload). Πολλά πανεπιστήμια και ερευνητικά ιδρύματα χρησιμοποιούν τέτοια συστήματα IR, για να παρέχουν πρόσβαση σε βιβλία, περιοδικά και άλλα έγγραφα. Ανάμεσα στις πιο γνωστές εφαρμογές IR συστημάτων συγκαταλέγονται οι μηχανές αναζήτησης στο Διαδίκτυο, όπως η Google<sup>2</sup>, η Live.com<sup>3</sup>, ή το Yahoo<sup>4</sup>. Στα πλαίσια της παρούσης, θα περιοριστούμε στα συστήματα ανάκτησης πληροφορίας, τα οποία επιχειρούν να εντοπίζουν εκείνα τα στοιχεία από το διαθέσιμο πολυμεσικό υλικό που ικανοποιούν τις επιθυμίες του χρήστη. Η διαδικασία αυτή παρουσιάζει μια πληθώρα προβλημάτων

<sup>2</sup><http://www.google.com>

<sup>3</sup><http://www.live.com>

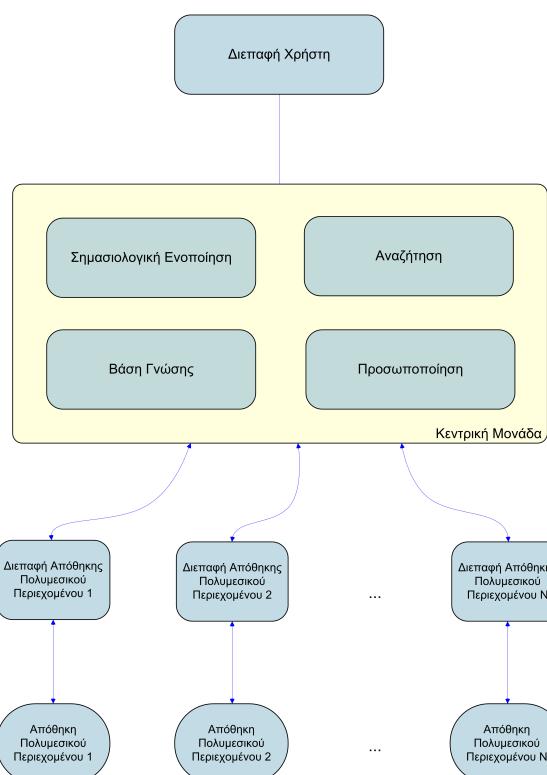
<sup>4</sup><http://www.yahoo.com>

που χρήζουν ερευνητικής αντικειμενικής, τα οποία εστιάζονται κυρίως στην αδυναμία πλήρους περιγραφής των επιθυμιών του χρήστη, αλλά και του περιεχομένου των διαθεσίμων δεδομένων. Οι ερευνητικές προσπάθειες των τελευταίων ετών προς την κατεύθυνση της εξμετάλλευσης προϋπάρχουσας ή κατάλληλα δημιουργούμενης γνώσης, καθώς και η χρήση προσαρμοζόμενων μεθόδων οργάνωσης του περιεχομένου, αποτέλεσαν και το δικό μας κίνητρο για την ενασχόληση με αυτή τη θεματική περιοχή.

Όπως θα δούμε, τα συστήματα IR σχετίζονται ιδιαίτερα με τα αντικείμενα της αναζήτησης και το ίδιο το ερώτημα αναζήτησης (query) που θέτει ένας χρήστης. Τα ερωτήματα αποτελούν επίσημες δηλώσεις των αναγκών πληροφόρησης των χρηστών. Ένα αντικείμενο είναι μια οντότητα που κρατά ή αποθηκεύει πληροφορίες σε μια βάση δεδομένων. Στη γενική περίπτωση, τα ερωτήματα των χρηστών αντιστοιχίζονται κατάλληλα με αντικείμενα που βρίσκονται αποθηκευμένα στη βάση δεδομένων. Ένα έγγραφο είναι, επομένως, ένα αντικείμενο δεδομένων, το οποίο συνήθως δεν αποθηκεύεται καν στο ίδιο το IR σύστημα, αλλά αντιπροσωπεύεται σε αυτό.

Θεωρώντας το πολυμεσικό περιεχόμενο ως ένα σύνολο από διαθέσιμα δεδομένα, ένα σύστημα IR αναλαμβάνει τόσο την αποθήκευση, όσο και την πρόσβαση σε αυτά. Η πρώτη διαδικασία ασχολείται με την αναπαράσταση των δεδομένων και τον τρόπο αποθήκευσης, ενώ η δεύτερη περιλαμβάνει τους τρόπους πρόσβασης σε αυτά. Επιπρόσθετα, αντικείμενο της πρόσβασης ενός συστήματος IR δεν πάνε να αποτελεί και η διεπαφή με τους τελικούς χρήστες με την έννοια της ικανοποίησης των επιθυμιών πληροφόρησής τους και την ανάκτηση των δεδομένων εκείνων που αντιστοιχούν στα ερωτήματα που τίθενται.

Η αφηρημένη δομή ενός τυπικού συστήματος ανάκτησης πληροφορίας παρουσιάζεται στο Σχήμα 2.2. Μια ενιαία διεπαφή με το χρήστη παρέχει μια ενοποιημένη



Σχήμα 2.2: Δομή ενός συστήματος IR

## Κεφάλαιο 2. Ερευνητικό πλαίσιο

πρόσβαση σε κάθε μεμονωμένη αποθήκη δεδομένων, ενώ οι διεπαφές των αποθηκών του πολυμεσικού περιεχομένου είναι αρμόδιες για την επικοινωνία μεταξύ της κεντρικής μονάδας του συστήματος και κάθε αποθήκης πολυμέσων. Η κεντρική μονάδα αποτελείται από τέσσερις επιμέρους μονάδες:

- τη βάση γνώσης,
- τη μονάδα σημασιολογικής ενοποίησης,
- τη μονάδα αναζήτησης και
- τη μονάδα προσωποποίησης.

Η βάση γνώσης αποτελείται από το μοντέλο γνώσης, το σημασιολογικό ευρετήριο και τα προφίλ των χρηστών. Η μονάδα σημασιολογικής ενοποίησης εξετάζει τη δημιουργία και την ενημέρωση του σημασιολογικού ευρετηρίου, ενώ η μονάδα προσωποποίησης χειρίζεται την ενημέρωση των προφίλ των χρηστών. Η μονάδα αναζήτησης αναλύει τα ερωτήματα των χρηστών, πραγματοποιεί το ταίριασμα με το ευρετήριο και επιστρέφει τα ανακτημένα έγγραφα στο χρήστη. Εν γένει, η περιγραφή της λειτουργίας κάθε μονάδας ακολουθεί έναν βασικό διαχωρισμό σε δύο κύριες περιόδους λειτουργίας: Κατά την περίοδο ερωτημάτων τα ερωτήματα των χρηστών υποβάλλονται σε επεξεργασία, συγκεντρώνονται και παρουσιάζονται οι αντίστοιχες απαντήσεις. Κατά την περίοδο αναπροσαρμογής ενημερώνονται το σημασιολογικό ευρετήριο και τα προφίλ των χρηστών, έτσι ώστε να απεικονίσουν τις όποιες αλλαγές και ενημερώσεις περιεχομένου και χρήσης έχουν πραγματοποιηθεί στο μεσοδιάστημα.

### 2.4.1 Οι διαδικασίες ενός IR συστήματος

Εμβαθύνοντας στις επιμέρους διαδικασίες ενός συστήματος ανάκτησης πληροφορίας, παρατηρούμε ότι αυτό διέπεται από μια σειρά διαδικασιών, οι οποίες σε κάθε περίπτωση οφείλουν να περιλαμβάνουν τις εξής τέσσερις:

- την ευρετηρίαση των δεδομένων
- το ταίριασμα των δεδομένων
- τη μετάφραση των ερωτημάτων των χρηστών
- την ανατροφοδότηση του συστήματος με τις γνώμες των χρηστών

Ειδικότερα, μέσω της διαδικασίας ευρετηρίασης κάθε δεδομένο  $\tilde{d}$  του συνόλου  $\tilde{D} = \{\tilde{d}_1, \tilde{d}_2, \tilde{d}_3, \dots\}$  των διαθεσίμων δεδομένων αντιστοιχίζεται σε κάποιο στοιχείο  $d$  του συνόλου  $D = \{d_1, d_2, d_3, \dots\}$  των πιθανών αναπαραστάσεων. Η αντιστοιχίση αυτή μπορεί να θεωρηθεί ως μία συνάρτηση  $f_1$  ορισμένη στο σύνολο  $\tilde{D}$ , η οποία δεν είναι απαραίτητα 1-1, ή αλλιώς:  $f_1 : \tilde{D} \rightarrow D$ . Εν γένει, τα διαθέσιμα δεδομένα μπορεί να έχουν οποιαδήποτε μορφή, ηλεκτρονική ή μη, ενώ η διαδικασία ευρετηρίασης μπορεί να γίνεται είτε αυτόματα από το σύστημα, είτε με τη συμβολή των ίδιων των χρηστών του συστήματος.

Αντίστοιχα, η διαδικασία της μετάφρασης των ερωτημάτων των χρηστών αναπαριστά την επιθυμία του εκάστοτε χρήστη σε μία μορφή κατανοητή από το σύστημα ανάκτησης πληροφορίας, μια διαδικασία επίσης όχι απαραίτητα 1-1. Έτσι, το ερώτημα  $q$  του χρήστη αντιστοιχίζεται μέσω της συνάρτησης  $f_2$  στο στοιχείο  $q$  του συνόλου

$Q = \{q_1, q_2, q_3, \dots\}$  των δυνατών αναπαραστάσεων του ερωτήματος του χρήστη, ή αλλιώς:  $f_2 : Q \rightarrow Q$ .

Ο μηχανισμός ταιριάσματος επιλέγει από τα διαθέσιμα δεδομένα ένα υποσύνολο  $\tilde{D}_{\alpha\pi} \subseteq D$ , που περιέχει τα δεδομένα που ταιριάζουν περισσότερο με το ερώτημα  $\tilde{q}$  του χρήστη. Ανάλογα με την περίπτωση η επιλογή μπορεί να είναι διαδική ή να γίνεται με τη βοήθεια βαθμών ταιριάσματος των αναπαραστάσεων των διαθεσίμων δεδομένων με την αναπαράσταση του εκάστοτε ερωτήματος. Συνακόλουθα, ο μηχανισμός ταιριάσματος επεξεργάζεται την αναπαράσταση  $q$  του ερωτήματος και τις αναπαραστάσεις  $D$  των διαθεσίμων δεδομένων και όχι τα ίδια τα  $\tilde{q}$  και  $\tilde{D}$ . Έτσι, ο μηχανισμός ταιριάσματος μπορεί να οριστεί ως η διαδικασία που επιλέγει το υποσύνολο  $D_{\alpha\pi} \subseteq D$  των διαθεσίμων δεδομένων, τα οποία σχετίζονται με το ερώτημα  $q$ , ή αλλιώς:  $D_{\alpha\pi} = \{d \in D : d \text{ σχετικό με } q\}$ .

Μαθηματικά η παραπάνω διαδικασία μπορεί να περιγραφεί από τη συνάρτηση  $f_3$ , που αντιστοιχεί κάθε ερώτημα σε ένα υποσύνολο του συνόλου των διαθέσιμων δεδομένων, ή αλλιώς:  $f_3 : Q \rightarrow \mathcal{P}(D)$ . Επομένως, όταν ένας χρήστης θέτει το ερώτημα  $\tilde{q}$ , η απάντηση του συστήματος δίδεται από τη σχέση  $\tilde{D}_{\alpha\pi} = (f_2 \circ f_3 \circ f_1^{-1})(\tilde{q})$ .

Από τα παραπάνω είναι μάλλον εμφανές για τον αναγνώστη ότι ο μηχανισμός ευρετησίας είναι υπεύθυνος για την αναπαράσταση της πληροφορίας, ο μηχανισμός μετάφρασης του ερωτήματος για την κατανόηση της επιθυμίας του χρήστη και ο μηχανισμός ταιριάσματος για την επιλογή των δεδομένων εκείνων που ικανοποιούν αυτή την επιθυμία. Αυτοί οι τρεις μηχανισμοί μπορούν να θεωρούνται ως βασικοί για ένα σύστημα ανάκτησης πληροφορίας.

#### 2.4.2 Μοντέλα IR συστημάτων

Για να είναι επιτυχής η ανάκτηση των πληροφοριών, είναι απαραίτητο να αναπαρασταθούν τα (πολυμεσικά) έγγραφα με κάποιο αποδοτικό τρόπο. Υπάρχουν διάφορα μοντέλα στη βιβλιογραφία για το σκοπό αυτό. Τα τελευταία μπορούν να κατηγοριοποιηθούν σε δύο διαστάσεις, όπως φαίνεται στο Σχήμα 2.3, ανάλογα με τη μαθηματική τους βάση και τις ιδιότητες του εκάστοτε μοντέλου. Στην πρώτη διάσταση ανήκουν

Ιδιότητες μοντέλου Μαθηματικό υπόβαθρο	χωρίς αλληλεξαρτήσεις όρων	με αλληλεξαρτήσεις όρων	
		δεδομένες αλληλεξαρτήσεις όρων	μεγάλες αλληλεξαρτήσεις όρων
συνολο-θεωρητικά			
αλγεβρικά			
πιθανοτικά			

Σχήμα 2.3: Κατηγοριοποίηση των μοντέλων IR

τα βασισμένα στα μαθηματικά μοντέλα. Αυτά περιλαμβάνουν:

## Κεφάλαιο 2. Ερευνητικό πλαίσιο

- Τα κλασικά σύνολο-θεωρητικά μοντέλα, τα οποία αναπαριστούν τα έγγραφα με σύνολα. Οι ομοιότητες προέρχονται συνήθως από τις σύνολο-θεωρητικές διαδικασίες σε αυτά τα σύνολα. Συνήθη τέτοια μοντέλα είναι:
  - Το στάνταρ Boolean μοντέλο
  - Το επεκτεταμένο Boolean μοντέλο
  - Η ασαφής αναζήτηση
- Τα αλγεβρικά μοντέλα, τα οποία αναπαριστούν τα έγγραφα και τα ερωτήματα ως διανύσματα ή μήτρες. Τα διανύσματα αυτά μετασχηματίζονται με την χρήση ενός πεπερασμένου αριθμού αλγεβρικών διαδικασιών σε μια μονοδιάστατη μετρική ομοιότητας. Διαχρίνονται συνήθως στα εξής:
  - Το μοντέλο διανυσματικού χώρου
  - Το γενικευμένο μοντέλο διανυσματικού χώρου
  - Το βασισμένο σε κατηγορίες μοντέλο διανυσματικού χώρου
  - Το επεκτεταμένο Boolean μοντέλο
  - Το ενισχυμένο, βασισμένο σε κατηγορίες, μοντέλο διανυσματικού χώρου
  - Η λανθάνουσα σημασιολογική ευρετηρίαση (LSI) (επίσης γνωστή και ως “λανθάνουσα σημασιολογική ανάλυση”-(LSA))
- Τα πιθανοτικά μοντέλα που μεταχειρίζονται τη διαδικασία της ανάκτησης εγγράφων σαν ένα πολυβάθμιο τυχαίο πείραμα. Οι ομοιότητες αναπαριστώνται με τον τρόπο αυτό ως πιθανότητες. Τα πιθανοτικά θεωρήματα, όπως το θεώρημα του Μπέυζ, χρησιμοποιούνται συχνά σε αυτά τα μοντέλα. Διαχρίνονται και αυτά σε:
  - Δυαδική ανάκτηση ανεξαρτησίας
  - Αβέβαιη εξαγωγή συμπερασμάτων
  - Μοντέλα γλώσσας
  - Μοντέλα απόκλισης από την τυχαιότητα

Στη δεύτερη διάσταση του Σχήματος 2.3 απαντούμε τα μοντέλα κατηγοριοποιημένα ανάλογα με τις ιδιότητές τους. Έτσι, παρατηρούμε ότι:

- τα μοντέλα χωρίς αλληλεξαρτήσεις όρων μεταχειρίζονται τους διαφορετικούς όρους/λέξεις ως μη αλληλοεξαρτώμενους. Αυτό το γεγονός αντιπροσωπεύεται συνήθως στα μοντέλα διανυσματικού χώρου από την υπόθεση ορθογωνιότητας των διανυσμάτων των όρων ή στα πιθανολογικά μοντέλα από μια υπόθεση ανεξαρτησίας για τις μεταβλητές των όρων.
- τα μοντέλα με δεδομένες αλληλεξαρτήσεις όρων επιτρέπουν μια κάποια αναπαράσταση των αλληλεξαρτήσεων μεταξύ των όρων. Εντούτοις, ο βαθμός της αλληλεξάρτησης μεταξύ δύο όρων καθορίζεται από το ίδιο το μοντέλο. Συνήθως προέρχεται άμεσα ή έμμεσα (π.χ. από τη μείωση των διαστάσεων) από την επανάληψη των όρων αυτών σε ολόκληρο το σύνολο των εγγράφων.

- τα μοντέλα με μεγάλες αλληλεξαρτήσεις όρων επιτρέπουν μια κάποια αναπαράσταση των αλληλεξαρτήσεων μεταξύ των όρων, αλλά δεν καθορίζουν την αλληλεξάρτηση μεταξύ δύο όρων. Με άλλα λόγια, απλά αναμεταδίδουν μια εξωτερική πηγή για το βαθμό αλληλεξάρτησης μεταξύ δύο όρων. Τέτοιες πηγές μπορούν π.χ. να αποτελέσουν περίπλοκοι αλγόριθμοι.

Τα Boolean μοντέλα αναζήτησης βασίζονται στην κλασική θεωρία συνόλων και αναπαριστούν κάθε διαθέσιμο δεδομένο με ένα σύνολο από όρους, π.χ.:  $d = \{t_1, t_2, \dots\}$ . Ένα ερώτημα  $q$  ενός χρήστη μεταφράζεται σε μια δυαδική αναπαράσταση όρων  $q = (t_1 \wedge t_2 \wedge t_3 \wedge \dots) \vee (t_4 \wedge t_5 \wedge t_6 \wedge \dots) \vee \dots$  και το ταίριασμα γίνεται με την απαίτηση της ικανοποίησης της αναπαράστασης αυτής από την ευρετηρίαση των διαθέσιμων δεδομένων. Από τα σημαντικά πλεονεκτήματα ενός Boolean μοντέλου αναζήτησης είναι οι σημαντικά μικρότερες απαιτήσεις του σε χώρο αποθήκευσης των αναπαραστάσεων των δεδομένων και είναι αυτό ακριβώς που το κάνει ιδιαίτερα δημοφιλές στις εμπορικές εφαρμογές. Επίσης, είναι σε θέση να χειριστεί εύκολα μεγάλες συλλογές διαθεσίμων δεδομένων, κάτι ιδιαίτερα σημαντικό σήμερα, μιας και η εποχή μας διέπεται από το φαινόμενο της υπερφόρτωσης πληροφοριών. Από την ερευνητική σκοπιά, όμως, η αδυναμία ενός τέτοιου μοντέλου να περιγράψει την αναχρίβεια και την υποκειμενικότητα, που χαρακτηρίζει τη διαδικασία αναζήτησης της πληροφορίας, το καθιστά μάλλον αδιάφορο.

Το αμέσως επόμενο πιο δημοφιλές μοντέλο στη βιβλιογραφία είναι το μοντέλο διανυσματικού χώρου ή πιο απλά διανυσματικό μοντέλο [213]. Το διανυσματικό μοντέλο είναι ένα αλγεβρικό μοντέλο που χρησιμοποιείται για το φιλτράρισμα πληροφοριών, την ανάκτηση πληροφοριών, την ευρετηρίαση και τις ταξινομήσεις σχετικότητας. Αντιπροσωπεύει έγγραφα φυσικής γλώσσας (ή οποιαδήποτε αντικείμενα, εν γένει) κατά τρόπο επίσημο μέσω της χρήσης διανυσμάτων (των προσδιοριστικών όρων, όπως π.χ. των δεικτών ευρετηρίασης) σε ένα πολυδιάστατο γραμμικό διάστημα. Η πρώτη χρήση του μοντέλου αυτού έγινε στο σύστημα ανάκτησης πληροφοριών SMART [47]. Τα έγγραφα αναπαριστώνται ως διανύσματα των όρων ευρετηρίασης (λέξεις κλειδιά). Η ευρετηρίαση των διανυσματικών μοντέλων τους επιτρέπει να περιγράφουν πολύ καλύτερα τους διάφορους βαθμούς συσχέτισης των όρων με τα διαθέσιμα δεδομένα. Έχουν, όμως, το σημαντικό μειονέκτημα της απαίτησης  $N$  θέσεων αποθήκευσης δεδομένων για κάθε διαθέσιμο δεδομένο. Όταν τα διαθέσιμα δεδομένα είναι πολλά και ποικίλου περιεχομένου, ο αριθμός  $|N|$  των όρων που χρησιμοποιεί το σύστημα γίνεται μεγάλος και αποτρέπει τη χρησιμοποίηση του μοντέλου αυτού.

Τέλος, τα πιθανοτικά μοντέλα [265] βασίζονται στην παραδοχή ότι σε ένα σύστημα ανάκτησης πληροφορίας η σχετικότητα ενός δεδομένου με το ερώτημα ενός χρήστη μπορεί μόνο να υποτεθεί. Ειδικότερα, στο [265] προτείνεται ένας αλγόριθμος που επιχειρεί μέσω δοκιμαστικών αναζητήσεων να βελτιώσει την αξιοπιστία της υπόθεσης, που γίνεται γύρω από τη σχετικότητα ερωτήματος και δεδομένου. Η αναπαράσταση των δεδομένων γίνεται μέσω δυαδικών διανυσμάτων και το ταίριασμα είναι μια στοχαστική διαδικασία που βασίζεται στο θεώρημα του Μπέυζ. Συνήθως το ταίριασμα γίνεται για κάθε όρο χωριστά, οπότε χρειάζεται να υποθέσουμε στατιστική ανεξαρτησία των όρων, για να συνδυάσουμε τα επιμέρους αποτελέσματα σε μια συνολική εκτίμηση για την πιθανότητα να είναι ένα δεδομένο σχετικό με το ερώτημα του χρήστη.

### 2.4.3 Μετρικές IR συστημάτων

Υπάρχουν διάφοροι τρόποι μέτρησης της απόδοσης ενός συστήματος ανάκτησης πληροφοριών. Οι σχετικές μετρικές στηρίζονται στην ύπαρξη μιας συλλογής εγγράφων και ενός ερωτήματος, για τα οποία είναι γνωστή η σχετικότητα των εγγράφων. Όλες οι συνήθεις μετρικές που περιγράφονται παρακάτω υποθέτουν την ύπαρξη δυαδικής σχετικότητας, δηλαδή ένα έγγραφο είναι είτε σχετικό, είτε απολύτως άσχετο. Στην πράξη τα ερωτήματα μπορούν να είναι λανθασμένα και μπορεί να υπάρχουν διαφορετικές εκφάνσεις της σχετικότητας. Οι σχετικές εξισώσεις για την ακρίβεια (precision), την ανάκληση (recall) και κάποιες άλλες αξιοσημείωτες μετρικές ακολουθούν.

Έστω  $R(\tilde{q})$  το σύνολο των δεδομένων που ικανοποιούν την επιθυμία ενός χρήστη που θέτει το ερώτημα  $\tilde{q}$ . Οι δύο κλασικές μετρικές αποτίμησης της ανάκτησης, η ακρίβεια και η ανάκληση, ορίζονται ως εξής:

$$p(q) = \frac{|\tilde{D}_{\alpha\pi}(q) \cap R(\tilde{q})|}{|\tilde{D}_{\alpha\pi}(q)|} \quad (2.14)$$

$$r(q) = \frac{|\tilde{D}_{\alpha\pi}(q) \cap R(\tilde{q})|}{|R(\tilde{q})|} \quad (2.15)$$

Η ακρίβεια ορίζεται ως το ποσοστό των ανακτημένων (retrieved) και σχετικών (relevant) εγγράφων προς το σύνολο των εγγράφων που ανακτήσαμε, ήτοι:

$$\text{precision} = \frac{|\{\text{relevant documents}\} \cap \{\text{retrieved documents}\}|}{|\{\text{retrieved documents}\}|} \quad (2.16)$$

Στη δυαδική ταξινόμηση η ακρίβεια είναι ανάλογη με τη θετική προβλεπόμενη τιμή. Η ακρίβεια λαμβάνει υπόψη το σύνολο των ανακτημένων εγγράφων. Να σημειωθεί, επίσης, ότι η παραπάνω έννοια και χρήση του όρου “ακρίβεια” στο χώρο της ανάκτησης πληροφοριών διαφέρει κατά πολύ από τον ορισμό της ακρίβειας σε άλλους επιστημονικούς κλάδους.

Από την άλλη, η ανάκληση ορίζεται ως το ποσοστό των σχετικών εγγράφων που ανακτώνται από το σύνολο των διαθέσιμων σχετικών εγγράφων, ήτοι:

$$\text{recall} = \frac{|\{\text{relevant documents}\} \cap \{\text{retrieved documents}\}|}{|\{\text{relevant documents}\}|} \quad (2.17)$$

Η επίτευξη ποσοστού ανάκλησης 100% είναι προφανώς μια εύκολη αποστολή, αν επιστρέφουμε όλα τα έγγραφα ως απάντηση σε οποιαδήποτε ερώτημα. Κατά συνέπεια, η ανάκληση από μόνη της δεν είναι αρκετή, αλλά πρέπει να λάβουμε υπόψη μας και τον αριθμό των άσχετων εγγράφων, π.χ. με τον ταυτόχρονο υπολογισμό της ακρίβειας.

Στην περίπτωση της δυαδικής ταξινόμησης εισάγονται τα μεγέθη της ιδιομορφίας (specificity -  $sp$ ), της ευαισθησίας (sensitivity -  $sn$ ) και της αποτελεσματικότητας (effectiveness -  $e$ ). Στην περίπτωση αυτή η ανάκληση ταυτίζεται με την ευαισθησία. Τα παραπάνω μεγέθη ορίζονται ως εξής: έστω ότι ο αριθμός των σχετικών εγγράφων που αναγνωρίζονται σωστά ή με άλλα λόγια αυτών που ταξινομούνται με σωστό τρόπο αναπαρίστανται με τον όρο “αληθινά και θετικά -  $TP$ ” και ότι ο αριθμός των εγγράφων που αναγνωρίζεται εσφαλμένα, δηλ. αυτά που ταξινομούνται εσφαλμένα, αναπαριστώνται με τον όρο “ψεύτικα και αρνητικά -  $FN$ ”. Ομοίως, έστω ότι ο αριθμός των εγγράφων που δεν είναι σχετικά και ταξινομούνται είτε σωστά, είτε λανθασμένα

παρέχεται από τους όρους “αληθινά και αρνητικά -  $TN$ ” και “λανθασμένα και θετικά -  $FP$ ”, αντίστοιχα. Τότε:

$$sp = specificity = \frac{TN}{TN + FP} \quad (2.18)$$

$$sn = sensitivity = \frac{TP}{TP + FN} \quad (2.19)$$

$$e = effectiveness = \frac{1}{a(1/sp) + (1-a)(1/sn)} \quad (2.20)$$

όπου η παράμετρος  $a$  επηρεάζει την εκτίμηση της αποτελεσματικότητας ε επιτρέποντας τη διαφορετική στάθμιση της ιδιομορφίας και της ευαισθησίας, δηλ. μια χαμηλή τιμή του  $a$  ευνοεί την ευαισθησία, ενώ μια υψηλή τιμή του ευνοεί την ιδιομορφία.

To *fall-out* ορίζεται ως το ποσοστό των άσχετων εγγράφων που ανακτώνται από το σύνολο όλων των άσχετων εγγράφων που είναι διαθέσιμα, ήτοι:

$$\text{fall-out} = \frac{|\{\text{irrelevant documents}\} \cap \{\text{retrieved documents}\}|}{|\{\text{irrelevant documents}\}|} \quad (2.21)$$

Ο σταθμισμένος αρμονικός μέσος όρος ακρίβειας και ανάκλησης, γνωστός και ως φ-μετρική ή ως *ισορροπημένο φ-σκορ* είναι:

$$F = 2 \cdot \text{precision} \cdot \text{recall}/(\text{precision} + \text{recall}) \quad (2.22)$$

όπου η ακρίβεια και η ανάκληση είναι ομοιόμορφα σταθμισμένες. Ο γενικός τύπος για έναν μη αρνητικό πραγματικό αριθμό α είναι:

$$F_\alpha = (1 + \alpha) \cdot \text{precision} \cdot \text{recall}/(\alpha \cdot \text{precision} + \text{recall}) \quad (2.23)$$

Τέλος, είναι γεγονός ότι η ακρίβεια και η ανάκληση βασίζονται σε ολόκληρο τον κατάλογο των εγγράφων που επιστρέφονται ως απάντηση σε ένα ερώτημα του χρήστη. Η μέση ακρίβεια δίνει έμφαση στο να επιστρέψει νωρίτερα τα πιο σχετικά έγγραφα. Αποτελεί το μέσο μεταξύ ενός συνόλου από ακρίβειες, που υπολογίζονται μετά από την διαδοχική περικοπή του καταλόγου μετά από κάθε ένα από τα σχετικά έγγραφα, ως εξής:

$$\text{Ave } P = \frac{\sum_{r=1}^N (P(r) \times \text{rel}(r))}{\text{number of relevant documents}} \quad (2.24)$$

όπου το  $r$  είναι η κατάταξη,  $N$  ο αριθμός ανάκτησης,  $\text{rel}(r)$  μία διαδική συνάρτηση της σχετικότητας μιας δοσμένης κατάταξης και  $P(r)$  η ακρίβεια σε μία αποκομμένη κατάταξη. Στην περίπτωση που υπάρχουν διάφορα ερωτήματα με γνωστές σχετικότητες η μέση ακρίβεια καταδεικνύει τη μέση τιμή των μέσων ακρίβειών, που υπολογίζονται για κάθε ένα από τα ερωτήματα ξεχωριστά.

#### 2.4.4 Προβλήματα των IR συστημάτων

Είναι γνωστό ότι τα συστήματα ανάκτησης πληροφορίας δεν αποδίδουν σύμφωνα με το αναμενόμενο, ακόμη και όταν οι αναζητήσεις γίνονται σε απλά αρχεία κειμένου και όχι σε πολυμεσικά έγγραφα. Έως ένα βαθμό η παραπάνω παρατήρηση είναι λογική, καθώς ο ρόλος ενός τέτοιου συστήματος έγκειται αποκλειστικά στον εντοπισμό δεδομένων, που είναι σχετικά με αυτό που ο χρήστης (δείχνει να) επιθυμεί. Προς τούτο, κανένα πλήρως αυτοματοποιημένο σύστημα δεν μπορεί να αντεπεξέλθει τέλεια, κυρίως λόγω των παρακάτω θεμάτων:

## Κεφάλαιο 2. Ερευνητικό πλαίσιο

1. Συνήθως τα ερωτήματα των χρηστών προς τις μηχανές αναζήτησης αποτελούνται από ελάχιστες λέξεις (συνήθως, δε, μόνο από δύο!) [67], [60], οι οποίες και δεν αρκούν, για να χαρακτηριστούν πλήρως οι περίπλοκες έννοιες που σχετίζονται με την επιθυμία του χρήστη.
2. Οι τελικοί χρήστες των συστημάτων ανάκτησης πληροφορίας χρησιμοποιούν πολύ συχνά διαφορετικούς όρους από εκείνους που έχουν χρησιμοποιήσει οι εμπειρογνόμονες ή το σύστημα κατά το χαρακτηρισμό του αναζητήσιμου υλικού [154].
3. Οι χρήστες δεν ξέρουν πάντα από την αρχή τι ακριβώς ψάχνουν [21]. Αντίθετα, τα αποτελέσματα της αναζήτησής τους συχνά επηρεάζουν τις επιθυμίες τους.
4. Η αβεβαιότητα είναι εγγενής στον ορισμό της σχετικότητας των εγγράφων με την επιθυμία του χρήστη [145].

Λαμβάνοντας υπόψη τα παραπάνω, είμαστε σε θέση να ισχυριστούμε ότι πρέπει να ληφθούν μία σειρά από μέτρα για την αντιμετώπιση της αστοχίας των υπαρχόντων συστημάτων IR. Προς την κατεύθυνση αυτή και στα πλαίσια της παρούσης εργασίας θα επιχειρήσουμε να συνοφίσουμε τα βασικά συμπεράσματα που προκύπτουν από τις παραπάνω αδιαμφισβήτητες διαπιστώσεις:

1. Η ύπαρξη πολλαπλών όρων που περιγράφουν κοινές έννοιες, όπως και η ύπαρξη όρων που περιγράφουν πολλαπλές έννοιες, οδηγεί στη χρήση λεξικών συνανύμων κατά τη διαδικασία της αναζήτησης.
2. Τα αποτελέσματα της αναζήτησης επηρεάζουν την επιθυμία του χρήστη, οπότε είναι λογικό να αναζητηθούν τρόποι, ώστε η διαδικασία να αποκτήσει περισσότερο αλληλεπιδραστικό χαρακτήρα. Η ανάγκη αξιοποίησης της ανατροφοδότησης από το χρήστη καθίσταται απολύτως απαραίτητη, καθώς μπορεί να επιλύσει τα προβλήματα που σχετίζονται με τις πολλαπλές έννοιες των όρων.
3. Η αδυναμία των όρων να περιγράφουν επαρκώς ένα ερώτημα χρήστη (ή ακόμα και τα διαθέσιμα δεδομένα, όποια μορφή και να έχουν αυτά) μάς οδηγεί στη διαπίστωση ότι τα συστήματα IR θα πρέπει να χειρίζονται έννοιες (concepts) και όχι όρους (terms).
4. Η στενή διασύνδεση της αβεβαιότητας με τις διαδικασίες αναζήτησης οδηγεί στη μοντελοποίηση των συστημάτων ανάκτησης πληροφορίας με βάση καινοτόμες θεωρήσεις, όπως τα ασαφή σύνολα και οι ασαφείς σχέσεις.
5. Τέλος, αφού το μέσο μήκος ενός ερωτήματος χρήστη κρίνεται ανεπαρκές, θα πρέπει να αναζητηθούν μέθοδοι επέκτασής του, ώστε να γίνει πληρέστερο. Η επέκταση του ερωτήματος, δηλαδή ο εμπλουτισμός του με όρους που δεν έδωσε ο χρήστης αλλά είναι σχετικοί με το ερώτημα, αποτελεί στην ουσία μια τεχνική βελτίωσης της ανάκλησης (του πλήθους των σωστών απαντήσεων).

### 2.4.5 Αναζήτηση σε πολυμεσική πληροφορία

Αναμφίβολα, τα τελευταία χρόνια η αφθονία ποικιλών μορφών πολυμεσικού περιεχομένου έχει συμβάλει τα δέοντα, έτσι ώστε το κόστος δημιουργίας, μετάδοσης και

αποθήκευσής του να είναι πλέον συγκριτικά πολύ μικρό. Όμως, όσο εύκολη και χωρίς κόστος τείνει να γίνει η συλλογή πολυμεσικού υλικού, τόσο πιο περίπλοκη, δύσκολη και με αυξανόμενο κόστος είναι η αποδοτική αναζήτηση και ανάκτηση τμημάτων του πολυμεσικού υλικού που μας ενδιαφέρουν. Οι σύγχρονες ερευνητικές προσπάθειες επικεντρώνονται στην ανάπτυξη μεθοδολογιών και τρόπων γρήγορου και εύκολου εντοπισμού των τμημάτων μέσα στο πολυμεσικό περιεχόμενο, που σε κάθε περίπτωση άπτονται του ενδιαφέροντος των εκάστοτε χρηστών.

Με δεδομένο ότι προς το παρόν η μορφή του πολυμεσικού υλικού είναι διαφορετική από εκείνη στην οποία τίθενται τα ερωτήματα των χρηστών, ποικίλα και γνωστά από τις αντίστοιχες διαδικασίες εγγράφων κειμένου ([213], [43]) προβλήματα πηγάζουν τόσο από το πλήθος των διαθέσιμων πολυμεσικών εγγράφων, όσο και από τις έμφυτες δυσκολίες της απαραίτητης διαδικασίας ταιριάσματος [75], [217]. Είναι πλέον κοινό μυστικό ότι η ανάκτηση και αναζήτηση πληροφορίας σε πολυμεσικό περιεχόμενο είναι δυσκολότερο να αντιμετωπιστεί σε σύγχριση με την απλή ανάκτηση κειμένων, καθώς σε αυτήν την περίπτωση είναι δυσκολότερο να ταιριάζουν τα ερωτήματα των χρηστών με τα διαθέσιμα πολυμεσικά έγγραφα. Η τελευταία διαπίστωση καθιστά ιδιαίτερα σημαντικό τον ρόλο των προφίλ χρηστών στη διαδικασία [7].

Το πρόβλημα της ανάλυσης και αναζήτησης του περιεχομένου ενός πολυμεσικού εγγράφου είναι αρκετά διαφορετικό από αυτή της ανάλυσης και αναζήτησης σε ένα έγγραφο κειμένου και αρκετά πιο περίπλοκο. Από τη μία, οι οντότητες που πρόκειται να ευρετηριασθούν δεν απαντώνται άμεσα μέσα στο έγγραφο, αλλά τα αναγνωρίσιμα χαρακτηριστικά γνωρίσματα πρέπει πρώτα να εξαχθούν και εν συνεχείᾳ να αντιστοιχθούν με εκείνα που βρίσκονται στη βάση γνώσης. Από την άλλη, ένα πολυμεσικό έγγραφο περιέχει τα αντικείμενα και τα γεγονότα, των οποίων οι σχέσεις είναι κυρίως χωροχρονικές, παρά καθαρά γραμματικές. Τέλος, αφηρημένες έννοιες, όπως “αθλητισμός” και “τέχνες”, δεν απαντώνται σε καμία περίπτωση άμεσα μέσα στα πολυμεσικά έγγραφα και συνεπώς πρέπει να προκύψουν από τα συγκεκριμένα αντικείμενα και γεγονότα, καθώς επίσης και τα χαρακτηριστικά γνωρίσματα (όπως π.χ. οι “συνθήκες φωτισμού”) που δεν αποδίδονται σε ένα ιδιαίτερο αντικείμενο ή ένα γεγονός.

Οι βασικές προσεγγίσεις που έχουν επιχειρηθεί έως τώρα προς την κατεύθυνση της αναζήτησης σε πολυμεσική πληροφορία είναι το ευρέως διαδεδομένο ερώτημα με χρήση παραδείγματος (query by example) και η σημασιολογική δεικτοδότηση (semantic indexing) [287]. Στην πρώτη περίπτωση, οπτικοακουστικές και λοιπές πληροφορίες εξάγονται από το παράδειγμα και ταιριάζονται με αυτές των διαθεσίμων εγγράφων για την επιλογή του συνόλου των διαδεδομένων που σχετίζονται με το ερώτημα [64]. Αν και αποφεύγεται το ταίριασμα μεταξύ λεξικών και πολυμεσικών αναπαραστάσεων, άλλου είδους προβλήματα είναι πιθανόν να δημιουργηθούν, όπως π.χ. είναι η εύρεση ακατάλληλου παραδείγματος. Η δεύτερη προσέγγιση διατηρεί τη λεξική μορφή του ερωτήματος του χρήστη και προσπαθεί να εξάγει σημασιολογικές οντότητες από την πολυμεσική πληροφορία, όπως π.χ. αντικείμενα και γεγονότα, με τα οποία μπορούμε εν συνεχείᾳ να ταιριάζουμε το ερώτημα.

## 2.5 Προσωποποίηση Πολυμεσικού Περιεχομένου

Η δυναμική προσαρμογή περιεχομένου στο χρήστη ή αλλιώς η προσωποποίησή του είναι ένα από τα ζητούμενα για κάθε σύστημα IR το οποίο σέβεται τον εαυτό του, αλληλεπιδρά με πολλούς διαφορετικούς χρήστες και έχει τη δυνατότητα να τους προσφέρει πολυάριθμες επιλογές. Αναμφίβολα, η αλληλεπιδραση του χρήστη με το σύστημα

απλοποιείται, αν το σύστημα μπορεί να “μαντέψει” εκ των προτέρων κάποιες από τις επιλογές του χρήστη, μειώνοντας έτσι τις εντολές που ο χρήστης θα πρέπει να δώσει, για να πάρει ένα ικανοποιητικό αποτέλεσμα κατά την αναζήτηση μιας πληροφορίας.

Στα πλαίσια της ερευνητικής μας ενασχόλησης αναπτύζαμε μεθοδολογίες, που περιλαμβάνουν την καταγραφή και ανάλυση του ιστορικού χρήστης των χρηστών ενός τέτοιου συστήματος (2.5.1). Στην ενότητα αυτή θα πραγματοποιηθεί μια πρώτη εισαγωγική παρουσίαση, ενώ η αναλυτική παρουσίασή τους θα λάβει χώρα στο κεφάλαιο 5 της παρούσης. Συνάμα, για λόγους πληροτητας, θα παρουσιάσουμε εδώ και την περίπτωση της συσχετιστικής ανατροφοδότησης ως την πλέον δημοφιλή μέθοδο προσαρμογής (2.5.2). Τέλος, γίνεται μια μικρή αναφορά στις δυνατότητες προσωποποίησης που περιγράφονται στο πλαίσιο δύο εκ των τελευταίων προτύπων του χώρου, του MPEG-7 (2.5.3.1) και του MPEG-21 (2.5.3.2), αντίστοιχα, ενώ παρουσιάζονται και οι βασικές αρχές που διέπουν την αποτελεσματική κατάταξη των αποτελεσμάτων της προσωποποίησης περιεχομένου σε ένα IR σύστημα (2.5.4).

### 2.5.1 Εξαγωγή προφίλ χρηστών

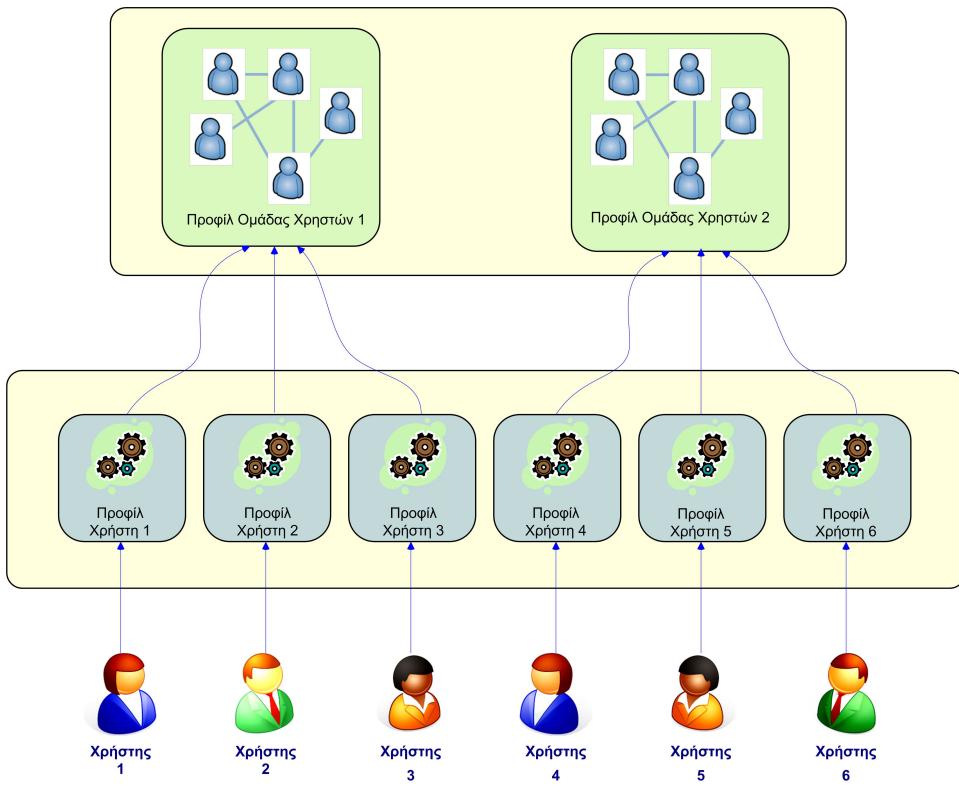
Εν γένει, η εξαγωγή προφίλ μπορεί να αναφέρεται σε ποικίλες δραστηριότητες και ερευνητικά πεδία, όπως είναι τα στερεότυπα, η εξαγωγή ψυχολογικών προφίλ, η εξαγωγή κοινωνικών προφίλ, η ανάλυση της απόδοσης ενός λογισμικού, κ.ο.κ.. Τα στερεότυπα π.χ., μπορούν να οριστούν ως απόψεις για κάποια από τα μέλη συγκεκριμένων ομάδων, βασισμένες απλώς και μόνο στην ιδιότητα μέλους σε εκείνη την ομάδα. Σύμφωνα με αυτή την προσέγγιση, οι διαφορές ανάμεσα στους χρήστες οφείλονται αποκλειστικά σε διαφορές στο επίπεδο της γνώσης τους. Έτσι, μετά από μια σειρά ερωτήσεων το σύστημα κατατάσσει το χρήστη κάπου και στο μέλλον ενεργεί ανάλογα.

Σε αντιδιαστολή με τη συσχετιστική ανάδραση, η οποία δεν επιτρέπει την μόνιμη προσαρμογή του συστήματος στις προτιμήσεις ενός συγκεκριμένου χρήστη, η καταγραφή των προτιμήσεων και του ιστορικού του τελευταίου μπορεί να οδηγήσει σε αποδοτική προσαρμογή του συστήματος στα εκάστοτε ενδιαφέροντά του. Συνάμα, το σύστημα είναι σε θέση να εκπαιδευτεί, ώστε στο μέλλον να προσαρμόζεται πιο γρήγορα και πιο σωστά σε αυτά.

Στα σύγχρονα συστήματα, λοιπόν, όλες οι ενέργειες του χρήστη καταγράφονται, ώστε να δημιουργηθεί ένα ιστορικό χρήστης. Στη συνέχεια το ιστορικό αυτό αναλύεται, για να παραχθεί το προφίλ του χρήστη. Η μεθοδολογία αυτή βασίζεται στην υπόθεση ότι οι ενέργειες του χρήστη στο παρελθόν χαρακτηρίζουν και αυτές που θα κάνει στο μέλλον. Μια ελαφρώς διαφορετική αντιμετώπιση επιτρέπει τη δημιουργία ομάδων ή και κοινωνιών χρηστών, των οποίων τα ιστορικά αναλύονται από κοινού. Βασική υπόθεση σε αυτή την περίπτωση είναι πως όμοιοι χρήστες αντιδρούν όμοια μεταξύ τους, κάτω από όμοιες καταστάσεις. Αυτή η προσέγγιση είναι χρήσιμη, όταν το ιστορικό κάθε χρήστη δεν είναι αρκετά μεγάλο, ώστε από μόνο του να αρκεί για την εξαγωγή αξιοποιήσιμων συμπερασμάτων (Σχήμα 2.4).

### 2.5.2 Συσχετιστική ανατροφοδότηση

Η πληροφορία που λαμβάνεται σχετικά με την εξέλιξη της αναζήτησης από την αλληλεπίδραση με το χρήστη λέγεται ανατροφοδότηση (feedback). Συνηθέστερα, τα συστή-



Σχήμα 2.4: Ομαδοποίηση χρηστών

ματα αναζήτησης αξιοποιούν τη συσχετιστική ανατροφοδότηση (relevance feedback), δηλαδή την αξιολόγηση που προσφέρει ο χρήστης γύρω από το βαθύ, στον οποίο τα έγγραφα που το σύστημα έχει επιλέξει σχετίζονται με την αρχική του επιθυμία. Η συσχετιστική ανατροφοδότηση αποτελεί πλέον ένα αναπόσπαστο γνώρισμα των περισσότερων συστημάτων ανάκτησης πληροφοριών. Ένα από τα πρώτα συστήματα IR που επιχείρησε να αξιοποιήσει τη συσχετιστική ανατροφοδότηση είναι το σύστημα SMART [47], που υλοποιεί το διανυσματικό μοντέλο που εξετάσαμε προηγουμένως. Η ιδέα πίσω από τη συσχετιστική ανατροφοδότηση είναι να ληφθούν τα αποτελέσματα που επιστρέφονται αρχικά από κάποια ερώτηση και να χρησιμοποιηθούν πληροφορίες για το κατά πόσο αυτά τα αποτελέσματα είναι ή όχι σχετικά, για να πραγματοποιήσουν μια νέα ερώτηση.

Μπορούμε να διακρίνουμε μεταξύ τριών ειδών ανατροφοδότησης: τη ρητή ανατροφοδότηση (explicit relevance feedback), την υπονοούμενη ανατροφοδότηση (implicit relevance feedback) και την “ψεύδο-” ή τυφλή ανατροφοδότηση. Η ρητή ανατροφοδότηση λαμβάνεται με το να αναγκάσουμε τους χρήστες να μαρκάρουν συγκεκριμένα έγγραφα ως σχετικά ή άσχετα. Η υπονοούμενη ανατροφοδότηση προκύπτει από τη συμπεριφορά των χρηστών, όπως π.χ. η καταγραφή από το σύστημα του ποια έγγραφα επιλέγουν ή δεν επιλέγουν για απεικόνιση ή/και πόσο χρόνο βλέπουν αυτά τα έγγραφα. Τέλος, η τυφλή ή “ψεύδο-” ανατροφοδότηση λαμβάνεται με το να υποθέσει κανείς ότι τα *n* κορυφαία έγγραφα στην κατάταξη των αποτελεσμάτων είναι αυτά που θεωρούν οι χρήστες πραγματικά σχετικά και ενδιαφέροντα. Είναι εμφανές ότι στην περίπτωση της ρητής συσχετιστικής ανατροφοδότησης ο χρήστης καλείται ο ίδιος να παρέμβει και να συνεισφέρει πληροφορίες σχετικές με την εκάστοτε αναζήτηση, ενώ, αντίθετα, κατά την υπονοούμενη συσχετιστική ανατροφοδότηση το σύστημα παρεμβαίνει και συλλέγει πληροφορίες εκ μέρους του χρήστη, χωρίς ο ίδιος να το αντιλαμβάνεται ή

να το προκαλεί.

Αν θέλαμε να εντρυφήσουμε περαιτέρω στη διαδικασία της ευρέως χρησιμοποιούμενης ρητής ανατροφοδότησης, θα παρατηρούσαμε ότι το σύστημα επιστρέφει ένα σύνολο  $D_{\alpha\pi}$  από δεδομένα ως το αποτέλεσμα της επεξεργασίας του αρχικού ερωτήματος του χρήστη. Από αυτά ο χρήστης επιλέγει όσα θεωρεί σχετικά με την αναζήτησή του, έστω το σύνολο  $D_{\sigma\chi} \subseteq D_{\alpha\pi}$ . Το σύνολο  $D_{\alpha\sigma} = D_{\alpha\pi} - D_{\sigma\chi}$  θεωρείται ότι έχει χαρακτηριστεί ως σύνολο μη σχετικών στοιχείων. Για τη συνέχιση της αναζήτησης το διάνυσμα  $q_\pi$ , που αναπαριστούσε το αρχικό ερώτημα, αντικαθίσταται πλέον από το διάνυσμα

$$q_\nu = \alpha \cdot q_\pi + \beta \sum_{d \in D_{\sigma\chi}} \frac{d}{|D_{\sigma\chi}|} - \gamma \sum_{D_{\alpha\sigma}} \frac{d}{|D_{\alpha\sigma}|} \quad (2.25)$$

όπου  $\alpha, \beta, \gamma$  σταθερές. Συνοπτικά, η σχέση (2.25) σημαίνει πως η αναζήτηση θα πρέπει (σε σχέση πάντα με το προηγούμενο ερώτημα) να κινηθεί λίγο πιο κοντά στα δεδομένα που χαρακτηρίστηκαν σχετικά και λιγότερο κοντά στα δεδομένα που χαρακτηρίστηκαν μη σχετικά. Οι παραπάνω σταθερές ορίζονται αυθαίρετα και ρυθμίζουν τη βαρύτητα που έχει το αρχικό ερώτημα, η θετική συσχετιστική ανάδραση και η αρνητική συσχετιστική ανάδραση.

Φυσικά, η παραπάνω εξίσωση αποτελεί μόνο μία από τις διαθέσιμες σχέσεις για την τροποποίηση του ερωτήματος που έχουν προταθεί κατά καιρούς από τα διάφορα μοντέλα αναζήτησης. Οφείλουμε να αναφέρουμε ότι η πιο διαδεδομένη υλοποίηση της συσχετιστικής ανάδρασης είναι εκείνη που βασίζεται στον αλγόριθμο Rocchio [207]. Επίσης, αξιζεί να σημειωθεί στο σημείο αυτό πως, αν και οι τύποι αυτοί είναι από τους παλαιότερους που προτάθηκαν, εφαρμόζονται ακόμη και σήμερα, πολλές φορές χωρίς καμία τροποποίηση [48].

### 2.5.3 Εξαγωγή προφίλ χρηστών με βάση τα πρότυπα MPEG

Δεδομένου ότι το πολυμεσικό περιεχόμενο έχει πολλαπλασιαστεί κατά τη διάρκεια των τελευταίων ετών, οι χρήστες έχουν αρχίσει να αναμένουν ότι το περιεχόμενο προσεγγίζεται σχετικά εύκολα σύμφωνα με τις δικές τους προτιμήσεις. Ένας από τους αποτελεσματικότερους τρόπους, για να επιτευχθεί αυτή η διαπίστωση, είναι μέσω της χρησιμοποίησης των προτύπων MPEG-7 και MPEG-21, τα οποία μπορούν να βοηθήσουν στο να αντιμετωπιστούν αποτελεσματικά τα ζητήματα που συνδέονται με το σχεδιασμό ενός συστήματος προσωποποίησης πολυμεσικού περιεχομένου, ακόμα και κάτω από ετερογενή περιβάλλοντα χρήστης [260].

#### 2.5.3.1 Εξαγωγή προφίλ χρηστών κατά MPEG-7

Στα πλαίσια του MPEG-7, τα προφίλ έχουν βασικά προταθεί ως ένα μέσο για τη μείωση της πολυπλοκότητας των MPEG-7 περιγραφών [296]. Όπως συμβαίνει και σε άλλα πρότυπα του MPEG, τα προφίλ αποτελούν υποσύνολα των προτύπων που καλύπτουν ορισμένες λειτουργικότητες, ενώ υπάρχουν και παραλλαγές των προφίλ με διαφορετική πολυπλοκότητα. Στο MPEG-7 τα προφίλ είναι υποσύνολα των εργαλείων περιγραφής για συγκεκριμένους τομείς εφαρμογών. Η προτεινόμενη διαδικασία του καθορισμού ενός προφίλ αποτελείται, εν γένει, από τα εξής τρία βήματα:

1. Επιλογή των εργαλείων που υποστηρίζονται από το προφίλ, δηλ. το υποσύνολο των περιγραφέων (descriptor) και των σχεδίων περιγραφής (description schemes) που χρησιμοποιούνται στην περιγραφή.

2. Καθορισμός των περιορισμών που διέπουν αυτά τα εργαλεία, όπως π.χ. οι περιορισμοί στην πληθικότητα των στοιχείων και στη χρήση των ιδιοτήτων.
3. Καθορισμός των περιορισμών επάνω στη σημασιολογία των εργαλείων, τα οποία περιγράφουν ακριβέστερα τη χρήση τους στο προφίλ.

Το αποτέλεσμα της επιλογής των εργαλείων και του καθορισμού των περιορισμών γι' αυτά προτυποποιείται, χρησιμοποιώντας την MPEG-7 Γλώσσα Περιγραφής (MPEG-7 DDL), και καταλήγει σε ένα XML σχήμα (XML Schema), όπως ακριβώς και το πλήρες πρότυπο.

Διάφορα είδη και τύποι προφίλ βρίσκονται αρχικά υπό εξέταση. Ξεχωρίζουμε το UserInteraction DS, το οποίο περιγράφει τις προτιμήσεις των χρηστών σχετικά με την κατανάλωση του οπτικοακουστικού περιεχομένου, καθώς επίσης και το ίδιο το ιστορικό χρήσης. Οι περιγραφές οπτικοακουστικού περιεχομένου του MPEG-7 μπορούν να αντιστοιχηθούν στις περιγραφές προτίμησης, έτσι ώστε να είναι δυνατή η επιλογή και προσωποποίηση του οπτικοακουστικού περιεχομένου για την αποδοτικότερη και αποτελεσματικότερη πρόσβαση σε αυτό, την παρουσίασή αλλά και την κατανάλωσή του. Το UserPreference DS περιγράφει τις προτιμήσεις για τους διαφορετικούς τύπους περιεχομένου και τους τρόπους πλοήγησης, συμπεριλαμβανομένης της εξάρτησης του εννοιολογικού πλαισίου, τόσο από την άποψη του χρόνου, όσο και του τόπου. Το UserPreference DS περιγράφει επίσης τη στάθμιση σπουδαιότητας των διαφορετικών προτιμήσεων, τα χαρακτηριστικά μυστικότητας των προτιμήσεων και το εάν οι προτιμήσεις υπόκεινται σε πιθανή αναπροσαρμογή, όπως π.χ. από έναν πράκτορα (agent) που μαθαίνει αυτόματα μέσω της αλληλεπίδρασης με το χρήστη. Το UsageHistory DS περιγράφει επίσης την ιστορία των ενεργειών, που πραγματοποιούνται από έναν χρήστη ενός συστήματος πολυμέσων. Οι περιγραφές του ιστορικού χρήσης μπορούν να ανταλλαχθούν μεταξύ των καταναλωτών, των πρακτόρων, των προμηθευτών περιεχομένου και των συσκευών τους και μπορούν στη συνέχεια να χρησιμοποιηθούν, για να καθορίσουν τις προτιμήσεις του χρήστη, όσον αφορά το πολυμεσικό περιεχόμενο [212].

### 2.5.3.2 Εξαγωγή προφίλ χρηστών κατά MPEG-21

Το όραμα του προτύπου MPEG-21 είναι να καθορίσει ένα ανοικτό πλαίσιο πολυμέσων, το οποίο θα επιτρέψει τη διαφανή και αυξημένη χρήση των πολυμεσικών πόρων πάνω από ένα ευρύ φάσμα δικτύων και συσκευών που χρησιμοποιούνται από διαφορετικές κοινότητες. Ο στόχος είναι να καλυφθεί ολόκληρη η αλυσίδα παράδοσης πολυμεσικού περιεχομένου, ζεκινώντας από την δημιουργία, την παραγωγή, την παράδοση και φτάνοντας έως το εμπόριο και την κατανάλωσή του. Στο πλαίσιο αυτό ο χρήστης είναι σε θέση να δημιουργήσει, να τροποποιήσει, να προστατεύσει, να προσαρμόσει και να καταναλώσει Ψηφιακά Αντικείμενα (Digital Items). Ο σημαντικότερος στόχος της Προσαρμογής Ψηφιακών Αντικειμένων (Digital Item Adaptation), που καθορίζεται στο πρότυπο MPEG-21 [295], είναι να υποστηριχθεί, στο τρέχον περιβάλλον χρήσης, η προσαρμογή των MPEG-21 ψηφιακών αντικειμένων, τα οποία αποτελούν τις θεμελιώδεις μονάδες της διανομής και της συναλλαγής μέσα στο πλαίσιο του προτύπου. Τα εργαλεία της DIA είναι βασισμένα στα εργαλεία σχημάτων, τα οποία παρέχουν ομοιόμορφα στοιχεία ρίζας για όλες τις DIA περιγραφές, και στους χαμηλού επιπέδου τύπους αντικειμένων, οι οποίοι παρέχουν μερικούς χαμηλού επιπέδου και βασικούς τύπους δεδομένων που μπορούν να χρησιμοποιηθούν ανεξάρτητα από ποικίλα εργαλεία DIA. Τα εργαλεία DIA είναι εν γένει ταξινομημένα σε οκτώ κατηγορίες:

- Τα εργαλεία περιγραφής περιβάλλοντος χρήσης,
- το BSDLink,
- την περιγραφή σύνταξης Bitstream (BSD),
- τα εργαλεία τερματικών και ποιότητας υπηρεσιών (QoS),
- τα καθολικά εργαλεία περιγραφής περιορισμών,
- τα εργαλεία προσαρμοστικότητας μεταδεδομένων,
- τα εργαλεία κινητικότητας συνόδου και
- τα εργαλεία διαμόρφωσης DIA.

Τα σημαντικότερα και πιο ενδιαφέροντα για την παρούσα ερευνητική δουλειά είναι τα εργαλεία περιγραφής περιβάλλοντος χρήσης, όπου καθορίζονται οι ικανότητες των τερματικών, τα φυσικά χαρακτηριστικά περιβάλλοντος, τα χαρακτηριστικά των δικτύων και τα χαρακτηριστικά των χρηστών. Στο πλαίσιο των χαρακτηριστικών των χρηστών συλλέγονται τα χαρακτηριστικά γνωρίσματα των χρηστών, συμπεριλαμβανομένων των πληροφοριών για τις περιγραφές χρηστών, τις προτιμήσεις χρηστών, τις προτιμήσεις του ιστορικού χρήσης, παρουσίασης και μετατροπής μέσων, τα χαρακτηριστικά δυνατότητας πρόσβασης και τις πληροφορίες που επιτρέπουν την καλύτερη υποστήριξη των βασισμένων στο εννοιολογικό πλαίσιο υπηρεσιών (όπως τα χαρακτηριστικά κινητικότητας, τον προορισμό, και την εστίαση της προσοχής). Και εδώ οι περιγραφές των προτιμήσεων των χρηστών και του ιστορικού χρήσης διαμορφώνονται σύμφωνα με τους ομώνυμους τύπους των MPEG-7 MDS. Μια περιγραφή προτίμησης χρηστών αποτελείται από (i) ένα σύνολο στοιχείων προτιμήσεων φιλτραρίσματος και αναζήτησης (FASP), τα οποία περιγράφουν τις προτιμήσεις των χρηστών σχετικά με το φιλτράρισμα και την αναζήτηση του πολυμεσικού περιεχομένου και (ii) ένα σύνολο στοιχείων προτιμήσεων πλοήγησης, τα οποία περιγράφουν τις προτιμήσεις των χρηστών σχετικά με τις περιλήψεις του πολυμεσικού περιεχόμενου.

#### 2.5.4 Κατάταξη προσωποποιημένων αποτελεσμάτων

Όπως είδαμε στην ενότητα 2.4.2, ένα από τα πιο δημοφιλή μοντέλα IR είναι το διανυσματικό μοντέλο, το οποίο απαντάται ευρέως και στην περίπτωση της προσωποποίησης της διαδικασίας ανάκτησης. Ειδικότερα, η κατάταξη των αποτελεσμάτων που επιστρέφει ένα σύστημα IR σε μια αναζήτηση λέξης κλειδιού μπορεί να υπολογιστεί χρησιμοποιώντας τις υποθέσεις της θεωρίας ομοιότητας εγγράφων, συγχρίνοντας την απόκλιση των γωνιών μεταξύ κάθε διανύσματος εγγράφων και του αρχικού διανύσματος της ερώτησης. Αντί του υπολογισμού των πραγματικών γωνιών μεταξύ των διανυσμάτων υπολογίζεται και συγχρίνεται το συνημίτονο της γωνίας μεταξύ των διανυσμάτων:

$$\cos \theta = \frac{\mathbf{v1} \cdot \mathbf{v2}}{\|\mathbf{v1}\| \|\mathbf{v2}\|} \quad (2.26)$$

Μια τιμή συνημιτόνου ίση με μηδέν σημαίνει ότι το ερώτημα και το διάνυσμα εγγράφων ήταν ορθογώνια και δεν είχαν καμία αντιστοιχία (δηλ. ο όρος του ερωτήματος δεν υπήρξε στο έγγραφο που εξετάζεται). Το κλασικό διανυσματικό μοντέλο, όπως

αυτό προτάθηκε από τον Salton [213], ενσωμάτωσε και τις τοπικές και σφαιρικές παραμέτρους στην εξίσωση των βαρών των όρων (γνωστή και ως tf-idf):

$$W = \text{tf} \cdot \log \frac{D}{d} \quad (2.27)$$

όπου:  $\text{tf}$  είναι η συχνότητα ενός όρου, μία παράμετρος καθοριζόμενη ανά έγγραφο ή τοπική και  $\log \frac{D}{d}$  είναι η αντίστροφη συχνότητα εγγράφων, μία παράμετρος που καθορίζεται ανά σύνολο εγγράφων. Το  $D$  είναι ο συνολικός αριθμός εγγράφων που υπάρχει στο σύνολο των εγγράφων, ενώ το  $d$  είναι ο αριθμός των εγγράφων που περιέχουν έναν όρο.

Βασική παραδοχή του μοντέλου είναι πως για την ευρετηρίαση όλων των διαθέσιμων δεδομένων και τη μετάφραση όλων των δυνατών ερωτημάτων του χρήστη αρκεί ένα πεπερασμένο πλήθος όρων. Αν  $N$  είναι το σύνολο των όρων, τότε η ευρετηρίαση αντιστοιχίζει σε κάθε διαθέσιμο δεδομένο  $\tilde{d}$  ένα διάνυσμα  $d$  του χώρου  $R^{|N|}$ . Η τιμή στη θέση  $i$  του διανύσματος δείχνει το βαθμό στον οποίο το συγκεκριμένο δεδομένο σχετίζεται με τον όρο  $t_i$ . Όμοια, το ερώτημα  $\tilde{q}$  του χρήστη αντιστοιχίζεται σε ένα διάνυσμα  $q$  του χώρου  $R^{|N|}$ . Στο διανυσματικό μοντέλο, δηλαδή, η αναπαράσταση του ερωτήματος συμπίπτει με την αναπαράσταση του ιδανικού για το χρήστη δεδομένου. Η διαδικασία του ταϊριάσματος υπολογίζει τις αποστάσεις ανάμεσα στις αναπαραστάσεις των διαθεσίμων δεδομένων και την αναπαράσταση του ερωτήματος του χρήστη. Οποιοδήποτε μετρική ορισμένη στο χώρο  $R^{|N|}$  μπορεί να χρησιμοποιηθεί για το ταϊριασμα αυτό, αλλά η πλέον συνηθισμένη είναι το εσωτερικό γινόμενο των αναπαραστάσεων, ήτοι:

$$\text{sim}(q, d) = \sum_{i \in N_{|N|}} q_i \cdot d_i \quad (2.28)$$

□

# Κεφάλαιο 3

## Μαθηματική σημειολογία, όροι και έννοιες

### 3.1 Εισαγωγή

Το παρόν κεφάλαιο επιχειρεί να παρουσιάσει το μαθηματικό υπόβαθρο και το γενικότερο πλαίσιο πάνω στα οποία κινείται η ερευνητική δραστηριότητά μας. Θα πρέπει να σημειωθεί εδώ ότι το παρόν κεφάλαιο έχει ως αποκλειστικό σκοπό την εξοικείωση του αναγνώστη με τα στοιχειώδη μαθηματικά και τις βασικές έννοιες και όρους που θα απαντηθούν σε όλο το εύρος του κειμένου από εδώ και στο εξής. Συγκεκριμένα, το κεφάλαιο χωρίζεται στις ακόλουθες πέντε βασικές ενότητες:

- **Ασαφή σύνολα και σχέσεις.** Αναμφίβολα, βασικό στοιχείο της αληθευτικής με το χρήστη, καθώς και της αυτοματοποιημένης ανάλυσης δεδομένων και πολυμεσικών κειμένων είναι η αβεβαιότητα, με όποια μορφή κι αν εμφανίζεται αυτή. Η ασαφής άλγεβρα έχει κεντρική θέση ανάμεσα στα μαθηματικά εργαλεία που χρησιμοποιούνται στο χώρο αυτό, μιας και αποτελεί μία από τις καλύτερες μαθηματικές εκφράσεις της αβεβαιότητας. Στη συνέχεια παρουσιάζονται οι βασικοί ορισμοί και οι βασικές αρχές που διέπουν το ερευνητικό πεδίο και θα αποβούν ιδιαίτερα χρήσιμοι για την κατανόηση των υπολογίων κεφαλαίων της παρούσης διατριβής.
- **Μεταβατικό κλείσιμο.** Στην ενότητα αυτή παρουσιάζεται συνοπτικά ο μαθηματικός ορισμός του μεταβατικού κλεισίματος μιας δυαδικής σχέσης, καθώς και παραδείγματα χρήσης του. Το σύνολο των προσεγγίσεων των επόμενων κεφαλαίων στηρίζονται στο μεταβατικό κλείσιμο και τις ιδιότητες των σχέσεων που αυτό επιβάλλει.
- **Οντολογίες.** Ανάμεσα στους στόχους της ερευνητικής μας ενασχόλησης βρίσκεται και η πραγματοποίηση υψηλού επιπέδου ανάλυσης και μοντελοποίησης πολυμεσικών εγγράφων με κύριο στόχο την αυτόματη κατηγοριοποίηση και σημασιολογική αναζήτησή τους. Στο πλαίσιο της εργασίας αυτής, η ανάλυση της πληροφορίας κειμένου αλλά και της οπτικής πληροφορίας θα δομηθεί γύρω από μία γνώση, αποθηκευμένης με τη μορφή οντολογιών (ontologies), η ανάπτυξη των οποίων βασίζεται σε συγκεκριμένες δομές και τεχνικές. Ειδικότερα, μια οντολογία είναι εν γένει ένα μοντέλο δεδομένων που αντιπροσωπεύει μια θεματική περιοχή (domain) και χρησιμοποιείται συνήθως για την εξαγωγή συμπερασμάτων (reasoning) σχετικά με τα αντικείμενα της περιοχής αυτής και τις

σχέσεις που τα διέπουν. Οι οντολογίες χρησιμοποιούνται ευρέως στην τεχνητή νοημοσύνη, το Σημασιολογικό Ιστό (Semantic Web) και την τεχνολογία λογισμικού ως μία μορφή αναπαράστασης γνώσης για έναν κόσμο ή κάποιο μέρος του και στην παρούσα ενότητα θα γίνει μια απόπειρα να παρουσιασθούν με τρόπο συνεκτικό και περιεκτικό.

- **Σημασιολογικές σχέσεις.** Το πρότυπο MPEG-7 έχει προτυποποιήσει έναν αριθμό κοινών και ευρέως χρησιμοποιούμενων σήμερα κανονιστικών (σαφών) σημασιολογικών σχέσεων. Οι σημασιολογικές αυτές σχέσεις μπορούν να περιγράψουν τον τρόπο με τον οποίο συνδέονται διάφορες σημασιολογικές οντότητες μεταξύ τους, καθώς και τις σχέσεις μεταξύ αντικειμένων ή/και γεγονότων. Προκειμένου να αντιμετωπίσουμε τα ερευνητικά προβλήματα με τα οποία ασχολείται η παρούσα διδακτορική διατριβή, διαμορφώνουμε κατάλληλα τις παραπάνω ευρέως διαδεδομένες κλασικές (σαφείς) σχέσεις ως ασαφείς σχέσεις ταξινόμησης και τις συνδυάζουμε για την παραγωγή ενός μεστού νοήματος συνδυασμού ασαφών ταξινομικών σχέσεων. Συνεπώς, για να αντιμετωπίσουμε τέτοιους σύνθετους τύπους σχέσεων, προτείνουμε μία μέθοδο “ασαφοποίησης” (“fuzzification”) των σημασιολογικών σχέσεων του προτύπου MPEG-7, η οποία χρησιμοποιεί ένα ασαφές σχεσιακό σύστημα, για να εμπλουτίσει μια υπάρχουσα ενοιολογική περιγραφή, και λειτουργεί με την εφαρμογή ενός συνόλου ασαφών διαδικασιών σχέσης στη σημασιολογική περιγραφή. Αυτές οδηγούν στον εμπλουτισμό της σημασιολογικής περιγραφής και στην κατασκευή των συσχετίσεων, που αποτελούν την περαιτέρω αξιοποιούμενη (ασαφή) γνώση.
- **Σχέδια Περιγραφής.** Τα Σχέδια Περιγραφής (Description Schemes - DSs) είναι μέρος του προτύπου MPEG-7 [189], [190], [299] και αποτελούν δομές μεταδεδομένων για την περιγραφή και το σχολιασμό του οπτικοακουστικού περιεχομένου. Τα σχέδια περιγραφής παρέχουν έναν τυποποιημένο τρόπο (σε γλώσσα XML) για την περιγραφή των σημαντικών εννοιών, που σχετίζονται με την οπτικοακουστική περιγραφή και διαχείριση του πολυμεσικού περιεχομένου, έτσι ώστε να διευκολυνθούν η έρευνα, η ευρετηρίαση, το φιλτράρισμα και η ίδια η πρόσβαση σε αυτό. Τα DSs καθορίζονται χρησιμοποιώντας τη γλώσσα καθορισμού περιγραφής, η οποία είναι βασισμένη στη γλώσσα σχημάτων XML [313]. Οι προκύπτουσες περιγραφές μπορούν να εκφραστούν με τη μορφή κειμένου ή και με συμπλεκτική διαδική μορφή. Σε αυτή την υποενότητα παρέχουμε μια γενική επισκόπηση των MPEG-7 MDSs και περιγράφουμε συνοπτικά τη λειτουργία και τη χρήση τους στις εφαρμογές πολυμέσων.

## 3.2 Ασαφή Σύνολα και Σχέσεις

Στο κεφάλαιο αυτό γίνεται μια απόπειρα για την εισαγωγή του αναγνώστη στις κυριότερες έννοιες της ασαφούς σχεσιακής άλγεβρας και της θεωρίας ασαφών συνόλων. Τις βασικότερες πηγές για τη θεωρία αποτέλεσαν τα συγγράμματα [139], [177] και [286]. Η μαθηματική σημειολογία που χρησιμοποιείται για το σκοπό αυτό συνοψίζεται στα ακόλουθα.

Με δεδομένο ένα σύμπαν  $U$ , ένα κλασικό (ή σαφές, σε αντιδιαστολή προς το ασαφές) σύνολο  $S$  περιγράφεται από μια συνάρτηση ιδιότητας μέλους (ή συνάρτηση συμμετοχής)

$$\mu_S : U \rightarrow \{0, 1\} \quad (3.1)$$

### Κεφάλαιο 3. Μαθηματική σημειολογία, όροι και έννοιες

Ένα ασαφές σύνολο  $F$  στο  $S$  περιγράφεται από μια συνάρτηση ιδιότητας μέλους

$$\mu_F : S \rightarrow [0, 1] \quad (3.2)$$

Συχνά το σύμβολο της συνάρτησης συμμετοχής είναι το ίδιο το σύμβολο  $F$ . Για  $s \in S$ , το  $F(s)$  συμβολίζει το βαθμό, στον οποίο το  $s$  ανήκει στο  $F$ , ή ισοδύναμα την τιμή  $\mu_F(s)$ , στην οποία η συνάρτηση συμμετοχής αντιστοιχεί το  $s$ . Μπορούμε να περιγράψουμε το ασαφές σύνολο χρησιμοποιώντας το συμβολισμό αθροίσματος (sum notation)[177]:

$$F = \sum_i s_i / w_i = \{s_1/w_1, s_2/w_2, \dots, s_n/w_n\} \quad (3.3)$$

όπου  $i \in N_n$ ,  $n = |S|$ ,  $w_i = M_F(s_i)$  ή πιο απλά  $w_i = F(s_i)$  και  $s_i \in S$ . Το  $|S|$  αναπαριστά την πληθυκότητα του  $S$ , δηλ. τον αριθμό των στοιχείων συνόλου του  $S$ , μια έννοια γνωστή από την κλασική θεωρία συνόλων, η οποία και γενικεύεται από τη θεωρία ασαφών συνόλων. Η μέγιστη τιμή της συνάρτησης συμμετοχής για ένα ασαφές σύνολο  $F$  ονομάζεται ύψος (height) του ασαφούς συνόλου και συμβολίζεται ως:

$$h(F) = \max_i(F(s_i)) \quad (3.4)$$

Ένα ασαφές σύνολο  $F$  καλείται κανονικό, αν το ύψος του  $h(F)$  είναι ίσο με 1. Μια άλλη έννοια γνωστή από τα κλασικά σύνολα είναι αυτή του υποσυνόλου, όπου με δεδομένα δύο ασαφή σύνολα  $A$  και  $B$  ορίζεται ως  $A \subseteq B$ :

$$A \subseteq B \Leftrightarrow A(s) \leq B(s), \forall s \in S \quad (3.5)$$

Αντίθετα με την κλασική θεωρία συνόλων, κατά τη θεώρηση των ασαφών συνόλων υπάρχουν πολλές επιλογές για τις συναρτήσεις της τομής, της ένωσης και του συμπληρώματος. Ειδικότερα, οι έννοιες της τομής  $t$

$$t(x, 1) = x \quad (3.6)$$

$$x_1 \leq x_2 \Leftrightarrow t(x, x_1) \leq t(x, x_2) \quad (3.7)$$

$$t(x_1, x_2) = t(x_2, x_1) \quad (3.8)$$

$$t(x_1, t(x_2, x_3)) = t(t(x_1, x_2), x_3) \quad (3.9)$$

της ένωσης  $u$

$$u(x, 0) = x \quad (3.10)$$

$$x_1 \leq x_2 \Leftrightarrow u(x, x_1) \leq u(x, x_2) \quad (3.11)$$

$$u(x_1, x_2) = u(x_2, x_1) \quad (3.12)$$

$$u(x_1, u(x_2, x_3)) = u(u(x_1, x_2), x_3) \quad (3.13)$$

και του συμπληρώματος  $c$

$$c(0) = 1 \quad (3.14)$$

$$c(1) = 0 \quad (3.15)$$

$$x_1 \leq x_2 \Leftrightarrow c(x_1) \geq c(x_2) \quad (3.16)$$

εξακολουθούν να ορίζονται ως είθισται, οι συνηθέστερες όμως επιλογές στην περίπτωση των ασαφών συνόλων είναι για την τομή  $\eta$

$$t(x_1, x_2) = \min(x_1, x_2) \quad (3.17)$$

για την ένωση η

$$u(x_1, x_2) = \max(x_1, x_2) \quad (3.18)$$

και για το συμπλήρωμα η

$$c(x) = 1 - x \quad (3.19)$$

Αν μια νόρμα τομής  $t$  ( $t$ -νόρμα) είναι συνεχής και ισχύει

$$t(x, x) \leq x \quad \forall x \in [0, 1] \quad (3.20)$$

τότε η νόρμα λέγεται Αρχιμήδεια.

Τέλος, μία ασαφής δυαδική σχέση στο  $S$  είναι μία συνάρτηση  $R : S^2 \rightarrow [0, 1]$ . Η αντίστροφη σχέση του ορίζεται ως

$$R^{-1}(x, y) = R(y, x) \quad (3.21)$$

Η διατομή, η ένωση και η sup-t σύνθεση δύο ασαφών σχέσεων  $P$  και  $Q$  που καθορίζονται στο ίδιο σύνολο  $S$  ορίζονται ως:

$$(P \cap Q)(x, y) = t(P(x, y), Q(x, y)) \quad (3.22)$$

$$(P \cup Q)(x, y) = u(P(x, y), Q(x, y)) \quad (3.23)$$

$$(P \circ Q)(x, y) = \sup_z t(P(x, z), Q(z, y)) \quad (3.24)$$

όπου  $t$  και  $u$  είναι μία  $t$ -νόρμα και μία  $t$  co-νόρμα, αντίστοιχα. Η στάνταρ  $t$ -νόρμα και  $t$ -coνόρμα είναι οι συναρτήσεις  $\min$  και  $\max$ , αν και είναι δυνατόν να χρησιμοποιηθεί μια πληθώρα σχέσεων, όπως καταδεικνύεται στο Παράρτημα A. Μία Αρχιμήδεια  $t$ -νόρμα ικανοποιεί επίσης τις ιδιότητες της συνέχειας και της υποταυτόδυναμιας (subidempotency), δηλ.  $t(a, a) < a, \forall a \in (0, 1)$ .

Η σχέση ταυτότητας  $R_I$  είναι το στοιχείο ταυτότητας της sup-t σύνθεσης:

$$R \circ R_I = R_I \circ R = R, \forall R \quad (3.25)$$

Οι ιδιότητες της ανακλαστικότητας, της συμμετρικότητας και της sup-t μεταβατικότητας καθορίζονται ως εξής:

- η  $R$  ονομάζεται **ανακλαστική** ανν  $R_I \subseteq R$  ή με άλλα λόγια αν για όλα τα  $x \in X$  ισχύει ότι  $R(x, x)$ . Παράδειγμα ανακλαστικής σχέσης είναι η δυαδική σχέση “μεγαλύτερο ή ίσο”, ενώ η δυαδική σχέση “μεγαλύτερο” δεν είναι.
- η  $R$  ονομάζεται **συμμετρική** ανν  $R = R^{-1}$  ή με άλλα λόγια αν για όλα τα  $x, y \in X$  ισχύει ότι αν  $R(x, y)$ , τότε  $R(y, x)$ . Π.χ. η σχέση “είναι συγγενής εξ αίματος” είναι μία συμμετρική σχέση, γιατί ο  $x$  είναι συγγενής εξ αίματος με τον  $y$  αν και μόνο αν ο  $y$  είναι συγγενής εξ αίματος με τον  $x$ .<sup>1</sup>
- η  $R$  ονομάζεται **αντισυμμετρική** ανν  $R \cap R^{-1} \subseteq R_I$  ή με άλλα λόγια αν για όλα τα  $x, y \in X$  ισχύει ότι αν  $R(x, y)$  και  $R(y, x)$ , τότε  $x = y$ . Π.χ. η σχέση “μεγαλύτερο ή ίσο” είναι μια αντισυμμετρική σχέση, μιας και αν  $x \geq y$  και  $y \geq x$ , τότε  $x = y$ .

<sup>1</sup>Προφανώς και σύμφωνα με τα παραπάνω, μία δυαδική σχέση είναι ίση με την αντίστροφή της, αν και μόνο αν είναι συμμετρική.

- η  $R$  ονομάζεται **sup-t μεταβατική** (ή απλά **μεταβατική**) ανν  $R \circ R \subseteq R$  ή με άλλα λόγια αν για όλα τα  $x, y, z \in X$  ισχύει ότι αν  $R(x, y)$  και  $R(y, z)$ , τότε  $R(x, z)$ . Π.χ. η σχέση “πρόγονος” είναι μία μεταβατική σχέση, μιας και αν ο  $x$  είναι πρόγονος του  $y$  και ο  $y$  είναι πρόγονος του  $z$ , τότε ο  $x$  είναι πρόγονος του  $z$ .

Θα πρέπει να σημειωθεί ότι η τομή δύο μεταβατικών σχέσεων είναι πάντα μεταβατική· γνωρίζοντας ότι οι σχέσεις “γεννήθηκε πριν” και “έχει το ίδιο μικρό όνομα” είναι μεταβατικές, μπορούμε με ασφάλεια να συμπεράνουμε ότι η σχέση “γεννήθηκε πριν και έχει το ίδιο μικρό όνομα” είναι επίσης μεταβατική. Η ένωση, όμως, δύο μεταβατικών σχέσεων δεν είναι πάντοτε μεταβατική· π.χ. η σχέση “γεννήθηκε πριν ή έχει το ίδιο μικρό όνομα” δεν είναι μία μεταβατική σχέση στη γενική περίπτωση.

Το **μεταβατικό κλείσιμο** μιας σχέσης είναι η μικρότερη μεταβατική σχέση που περιέχει την αρχική σχέση. Όπως αναλύεται και στην επόμενη υποενότητα, το μεταβατικό κλείσιμο  $Tr(R)$  μιας σχέσης  $R$  δίδεται από την

$$Tr(R) = \bigcup_{i=1}^{\infty} R^{(i)} \quad (3.26)$$

όπου  $R^{(i)} = R \circ R^{(i-1)}$  και  $R^{(1)} = R$ . Αποδεικνύεται ότι, αν η  $R$  είναι ανακλαστική, τότε το μεταβατικό κλείσιμό της δίνεται από την  $Tr(R) = R^{(n-1)}$ , όπου  $n = |S|$ .

Μία **ασαφής σχέση διάταξης** είναι μία ασαφής δυαδική σχέση, η οποία είναι αντισυμμετρική και μεταβατική. Μία **σχέση μερικής διάταξης** είναι επιπρόσθετα και ανακλαστική. Τέλος, μία **ασαφής σχέση μερικής διάταξης  $R$**  ορίζει για κάθε στοιχείο  $s \in S$  το ασαφές σύνολο των προγόνων του (ή αλλιώς της κλάσης που κυριαρχεί)  $R_{\geq[s]}(x) = R(s, x)$  και των απογόνων του (ή αλλιώς της κλάσης που κυριαρχείται)  $R_{\leq[s]}(x) = R(x, s)$ .<sup>2</sup>

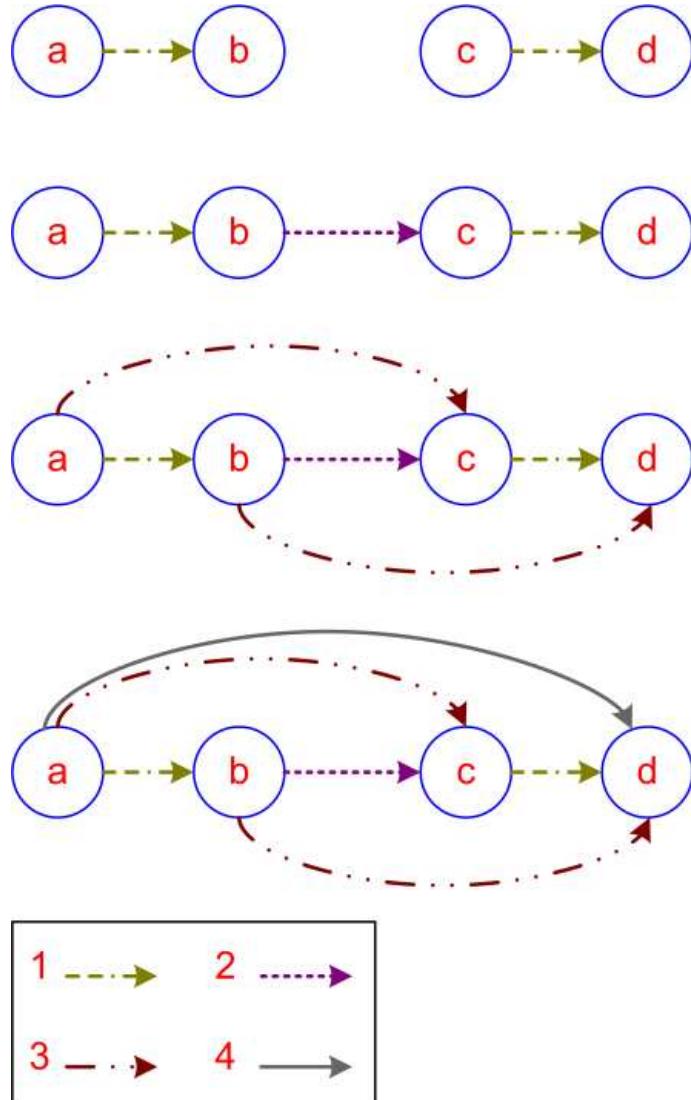
### 3.3 Μεταβατικό Κλείσιμο

Η ιδιότητα της μεταβατικότητας είναι μια από τις πιο σημαντικές ιδιότητες, όσον αφορά στις ασαφείς δυαδικές σχέσεις, ειδικά στην περίπτωση που οι τελευταίες χρησιμοποιούνται για την αναπαράσταση χαρακτηριστικών ομοιότητας ή και διάταξης της πληροφορίας. Ο αυστηρός μαθηματικός ορισμός του μεταβατικού κλεισίματος μιας δυαδικής σχέσης  $R$  σε ένα σύνολο  $X$  το ορίζει ως την μικρότερη μεταβατική σχέση στο  $X$  που περιέχει την  $R$ . Παραδείγματος χάριν, εάν το  $X$  είναι το σύνολο των ανθρώπων (ζωντανών ή νεκρών) και η  $R$  είναι η σχέση “γονέας”, τότε το μεταβατικό κλείσιμο της  $R$  είναι η σχέση “ $x$  είναι πρόγονος του  $y$ ,  $x, y \in X$ ”. Ή με άλλα λόγια, αν το  $X$  είναι ένα σύνολο αεροδρομίων,  $x, y \in X$  και  $R(x, y)$  σημαίνει “υπάρχει μια απ’ ευθείας πτήση από το αεροδρόμιο  $x$  στο αεροδρόμιο  $y$ ”, τότε το μεταβατικό κλείσιμο της  $R$  είναι η σχέση “είναι δυνατόν να πετάξει κανείς από το  $x$  στο  $y$  με μια ή περισσότερες πτήσεις”.

Σχηματικά, τα παραπάνω απεικονίζονται στο Σχήμα 3.1, όπου και παραθέτουμε τα διαδοχικά βήματα του μεταβατικού κλεισίματος μιας σχέσης  $R$  μεταξύ 4 στοιχείων  $a, b, c, d \in X$ . Ειδικότερα, έστω ότι η  $R$  είναι μία μεταβατική σχέση, η οποία απεικονίζεται στο Σχήμα με γραμμές τύπου 1. Αν υποθέσουμε ότι εισάγεται η  $R(b, c)$  (γραμμή τύπου 2), τότε δεν μπορούμε να ισχυριζόμαστε ότι η σχέση  $R$  εξακολουθεί

<sup>2</sup>Για λόγους απλότητας θα χρησιμοποιούμε στο εξής τον εναλλακτικό συμβολισμό  $R(s)$  αντί του  $R_{\leq[s]}(x)$ .

να είναι μεταβατική. 'Υστερα από μία αυτο-σύνθεση (self-composition), ο πρόγονος  $a$  του  $b$  συνδέεται με το  $c$  και το  $b$  συνδέεται με τον απόγονο  $d$  του  $c$  (γραμμές τύπου 3). Τελικά, μετά από ακόμα ένα βήμα αυτό-σύνθεσης, το  $a$  συνδέεται με το  $d$  (γραμμή τύπου 4). Με τον επαυξητικό αυτό τρόπο επιτυγχάνεται η μεταβατικότητα της σχέσης ακόμα και μετά την μεταβολή ενός στοιχείου της.



**Σχήμα 3.1:** Παράδειγμα επαυξητικού (incremental) μεταβατικού κλεισίματος.

Στο σημείο αυτό θα πρέπει να αναφέρουμε ότι το μεταβατικό κλεισίμου ορίζεται πάντοτε και για οποιαδήποτε σχέση  $R$ . Για να γίνει κατανοητό το τελευταίο, θα πρέπει να σημειωθεί ότι η τομή οποιαδήποτε οικογένειας μεταβατικών σχέσεων είναι και πάλι μεταβατική, ενώ υπάρχει τουλάχιστον μια μεταβατική σχέση που περιέχει την  $R$ . Το μεταβατικό κλεισίμου της  $R$  δίνεται από την τομή όλων των μεταβατικών σχέσεων που περιέχουν την  $R$ . Επιπρόσθετα, μπορούμε να περιγράψουμε το μεταβατικό κλεισίμου της  $R$  ως εξής: 'Εστω ότι καθορίζουμε μια σχέση  $T$  στο  $X$  λέγοντας ότι  $T(x, y)$ , αν και μόνο αν υπάρχει μια πεπερασμένη ακολουθία στοιχείων  $(x_i)$  τέτοια, ώστε  $x = x_0$  και  $x_0Rx_1, x_1Rx_2, \dots, x_{n-1}Rx_n$  και  $x_nRy$ . Ο επίσημος συμβολισμός συνοψίζεται στο:

$$R^+ = \bigcup_{i \in \mathbb{N}} R^i \quad (3.27)$$

ενώ ο αναγνώστης καλείται να αναζητήσει περαιτέρω πληροφορίες στον ορισμό της συνάρτησης σύνθεσης. Είναι εύχολο να δει κανείς ότι η σχέση  $T$  είναι μεταβατική και περιέχει την  $R$ . Επιπλέον, οποιαδήποτε μεταβατική σχέση που περιέχει την  $R$  πρέπει επίσης να περιέχει την  $T$ , έτσι η  $T$  είναι το μεταβατικό κλείσιμο της  $R$ .

Για να επιδείξουμε τώρα ότι η  $T$  είναι όντως η μικρότερη μεταβατική σχέση που περιέχει την  $R$ , θεωρούμε τα ακόλουθα: Έστω ότι το  $A$  είναι ένα οποιοδήποτε σύνολο στοιχείων. Υποθέτουμε ότι  $\exists G_A$  μεταβατική σχέση  $R_A \subseteq G_A \wedge T_A \not\subseteq G_A$ . Συνεπώς,  $\exists (a, b) \notin G_A$ . Έτσι αυτό το συγκεκριμένο  $(a, b) \notin R_A$ . Τώρα, εξ ορισμού της  $T$  ξέρουμε ότι  $\exists n \in \mathbb{N} \quad (a, b) \in R_A^n$ . Κατόπιν,  $\forall i \in \mathbb{N}, i < n \Rightarrow e_i \in A$ . Έτσι, υπάρχει ένα μονοπάτι από το  $a$  στο  $b$  όπως αυτό:  $aR_Ae_1R_A\dots R_Ae_{n-1}R_Ab$ . Αλλά, από τη μεταβατικότητα του  $G_A$  στο  $R_A$ ,  $\forall i \in \mathbb{N}, i < n \Rightarrow (a, e_i) \in G_A$ , έτσι  $(a, e_{(n-1)}) \in G_A \wedge (e_{(n-1)}, b) \in G_A$ , έτσι από τη μεταβατικότητα του  $G_A$ , παίρνουμε  $(a, b) \in G_A$ , το οποίο βρίσκεται σε αντίφαση με το  $(a, b) \notin G_A$ . Επομένως,  $\forall (a, b) \in A \times A, (a, b) \in T_A \Rightarrow (a, b) \in G_A$ . Αυτό σημαίνει ότι  $T \subseteq G$ , για οποιοδήποτε μεταβατικό  $G$  που περιέχει την  $R$ . Έτσι, η  $T$  είναι η μικρότερη μεταβατική σχέση που περιέχει την  $R$ . Κατά συνέπεια: Αν η  $R$  είναι μεταβατική, τότε  $R = T$ .

Επιστρέφοντας στο Σχήμα 3.1, η εισαγωγή μιας μεταβατικής σχέσης στο γράφο δεν εξασφαλίζει την μεταβατικότητα της νέας δομής. Ειδικότερα, έστω ότι στην αρχική κατάσταση υποθέτουμε τις (μη μηδενικές) σχέσεις  $R(a, b)$  και  $R(c, d)$  (πράσινες γραμμές). Ακολούθως, εισάγεται η σχέση  $R(b, c)$ , όπως καταδεικνύεται στο δεύτερο βήμα του Σχήματος (μωβ γραμμή), οπότε δεν μπορούμε να εξακολουθήσουμε να θεωρούμε ότι η σχέση  $R$  είναι μεταβατική. Μετά από την πρώτη σύνθεση ο πρόγονος  $a$  του  $\beta$  σχετίζεται με το  $c$  και το  $b$  σχετίζεται με τον απόγονο  $d$  του  $c$  (καφέ γραμμές). Υπερα από ακόμη μία σύνθεση το  $a$  συνδέεται με το  $d$  (γκρι γραμμή). Καθώς η σχέση  $R$  είχε υποτεθεί αρχικά ότι είναι μεταβατική, οι δύο συνθέσεις είναι αρκετές για να επιβάλλουν εκ νέου την μεταβατικότητα.

Η χρήση της έννοιας του μεταβατικού κλεισίματος είναι ποικίλη. Εν προκειμένω, ιδιαίτερα σημαντικό είναι το γεγονός ότι η ένωση δύο μεταβατικών σχέσεων δεν είναι απαραίτητα μεταβατική και, προκειμένου να συντηρηθεί η μεταβατικότητα, πρέπει κάποια εκ των δύο να πάρει το μεταβατικό κλείσιμο. Παράδειγμα τέτοιας περίπτωσης αποτελεί η ένωση δύο σχέσεων ισοδυναμίας. Προκειμένου να ληφθεί μια νέα σχέση ισοδυναμίας, μία εκ των δύο αρχικών πρέπει να πάρει το μεταβατικό κλείσιμο. (Επίσης, σημειώστε ότι στην περίπτωση των σχέσεων ισοδυναμίας η ανακλαστικότητα και η συμμετρία είναι αυτόματες).

Εξαιτίας της φυσικής τους σημασίας η μεταβατικότητα των δυαδικών σχέσεων είναι μια ιδιότητα που έχει συνδεθεί στενά με τη μελέτη των γράφων. Σε αυτό το πλαίσιο, το μεταβατικό κλείσιμο ισοδυναμεί με τον εντοπισμό των ζευγαριών των κόμβων που συνδέονται είτε άμεσα, είτε έμμεσα μέσω κάποιου μονοπατιού. Ο αναγνώστης μπορεί να αναζητήσει στη βιβλιογραφία εκτενέστερες αναφορές στη χρήση του μεταβατικού κλεισίματος ενός γράφου [274], [273], [252], [14], [223], αλλά και αποδοτικούς αλγορίθμους για τον υπολογισμό του [203], [194]. Ιδιαίτερη μνεία θα πρέπει να αποδοθεί στον Warshall [274], μιας και οι περισσότερες εργασίες βασίζονται στις δικές του ερευνητικές προσπάθειες. Όσον αφορά το μεταβατικό κλείσιμο ασαφών δυαδικών σχέσεων, αυτό έχει μελετηθεί στα πλαίσια των [73], [76] και [151].

Συνοψίζοντας, στα πλαίσια της παρούσης διατριβής το μεταβατικό κλείσιμο  $Tr^t(r)$  μια πλήρους δυαδικής σχέσης  $r$ , με δεδομένη την  $t$ -νόρμα, υπολογίζεται με την ακόλουθη μέθοδο [139]:

$$Tr^t(r) = \bigcup_{f=1}^{\infty} r^f \quad (3.28)$$

$$r^{f+1} = r^f \circ^t r \quad (3.29)$$

$$r^1 = r \quad (3.30)$$

Στη σχέση 3.28 χρησιμοποιείται η κλασική ένωση *max*. Υποθέτοντας πως το καθολικό σύνολο  $S$  είναι φραγμένο ( $|S| = n$ ), η σχέση 3.28 γράφεται τελικά ως:

$$Tr^t(r) = \bigcup_{f=1}^{n-1} r^f \quad (3.31)$$

## 3.4 Οντολογίες

Αναγνωρίζοντας στον άνθρωπο το μεγάλο πλεονέκτημα της εκμετάλλευσης της τρέχουσας, αλλά και της προηγούμενης συμβολικής γνώσης και του συνδυασμού της με μη συμβολικά δεδομένα, δίνεται πλέον ολοένα και πιο μεγάλη έμφαση στην αυστηρή κωδικοποίηση της κοινής γνώσης που μοιράζονται οι άνθρωποι, ώστε η τελευταία να μπορεί να αξιοποιηθεί από αυτοματοποιημένες διεργασίες. Αν θέλαμε να ορίσουμε αυστηρά τον όρο “γνώση”, θα λέγαμε ότι αυτός αναφέρεται σε αποθηκευμένη πληροφορία ή μοντέλα που χρησιμοποιούνται από έναν άνθρωπο ή μια μηχανή, για να εξηγήσουν, να προβλέψουν και να αντιδράσουν κατάλληλα στον εξωτερικό κόσμο [91]. Αναμφίβολα, ο πιο σχετικός ως προς τη γνώση ερευνητικός χώρος είναι ο χώρος των οντολογιών και αυτόν ακριβώς θα μελετήσουμε και θα αξιοποιήσουμε στο πλαίσιο της παρούσης διατριβής.

### 3.4.1 Εισαγωγή

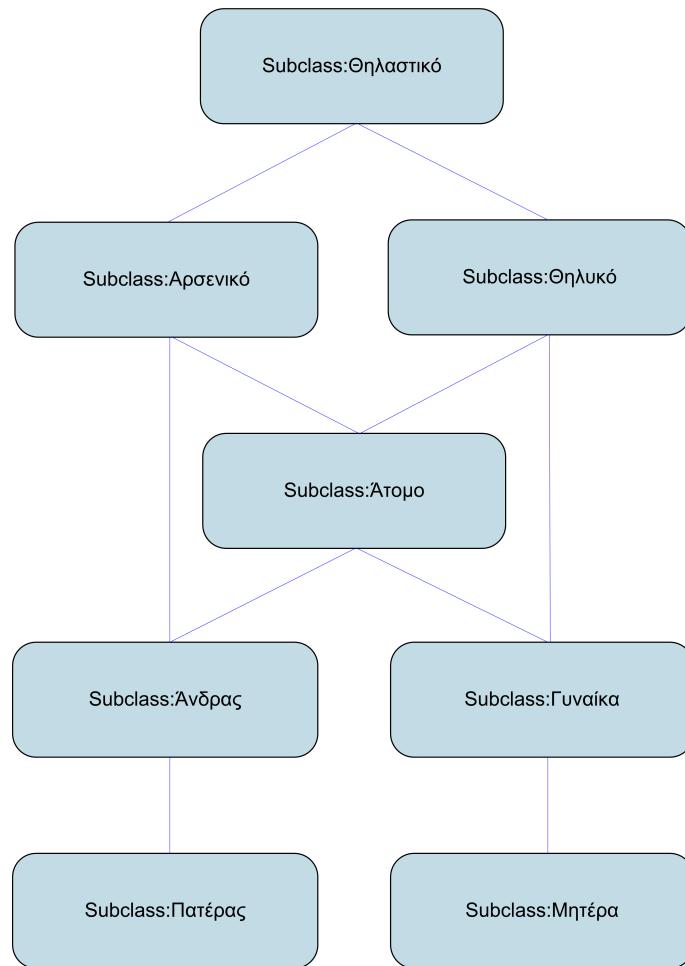
Ανάμεσα στους στόχους της προτεινόμενης ερευνητικής ενασχόλησής μας βρίσκεται μεταξύ άλλων και η πραγματοποίηση υψηλού επιπέδου ανάλυσης και μοντελοποίησης πολυμεσικών εγγράφων με κύριο στόχο την αυτόματη κατηγοριοποίηση και τη σημασιολογική αναζήτησή τους. Η ανάλυση της πληροφορίας κειμένου αλλά και της οπτικής πληροφορίας βασίζεται κατά κόρον σε γνώση αποθηκευμένη σε οντολογίες (ontologies), η ανάπτυξη των οποίων βασίζεται σε συγκεκριμένες δομές και προηγείται χρονικά των αλγορίθμων ανάλυσης. Οι οντολογίες αποτελούν ένα μέσο διαχείρισης και οργάνωσης της γνώσης με σκοπό την εξαγωγή συμπερασμάτων και η χρήση τους στα πλαίσια των ερευνητικών δραστηριοτήτων που περιγράφονται στο παρόν κείμενο έχει σημαντικά αποτελέσματα, όσον αφορά τις επιδόσεις αρχειοθέτησης, κατηγοριοποίησης και αναζήτησης πληροφορίας.

Όπως έχει ήδη αναφερθεί, μια οντολογία (βλ. Σχήμα 3.2) είναι εν γένει ένα μοντέλο δεδομένων που αντιροστούει μια θεματική περιοχή, τα αντικείμενά του και τις σχέσεις που τα διέπουν. Αποτελεί, δε, μια αυστηρή περιγραφή ενός πεδίου. Εν γένει αποτελείται από τρία βασικά στοιχεία [163]:

$$O = \{S, R, I\} \quad (3.32)$$

όπου  $S$  είναι το σύνολο των σημασιολογικών οντοτήτων,  $R$  το σύνολο των σχέσεων στο  $S$  και  $I$  ένα σύνολο κανόνων που χρησιμοποιούνται για την αναπαράσταση πληροφοριών, που δεν μπορούν να κωδικοποιηθούν με τη χρήση απλών δυαδικών σχέσεων. Γενικά, αν και η εξέλιξη στο χώρο της αναπαράστασης απλών δεδομένων, δηλαδή στη δημιουργία των  $S$  και  $R$ , έχει προχωρήσει αρκετά, δεν έχουν γίνει ακόμη πολλά βήματα προς την αυτόματη εξαγωγή πιο σύνθετων συμπερασμάτων με χρήση του  $I$ .

### Κεφάλαιο 3. Μαθηματική σημειολογία, όροι και έννοιες



**Σχήμα 3.2:** Μια οντολογική ταξινομία.

Οι σύγχρονες οντολογίες μοιράζονται πολλές δομικές ομοιότητες, ανεξάρτητα από τη γλώσσα στην οποία εκφράζονται. Οι περισσότερες από αυτές περιγράφουν στιγμιότυπα (instances), έννοιες (concepts), ιδιότητες (properties) και σχέσεις (relations):

- Στιγμιότυπα: τα βασικά αντικείμενα (ή “βασικού επιπέδου” αντικείμενα)
- Έννοιες: σύνολα, συλλογές ή τύποι αντικειμένων
- Ιδιότητες: ιδιότητες, χαρακτηριστικά γνωρίσματα, γενικά χαρακτηριστικά ή παράμετροι, που μπορούν να έχουν ή και να μοιραστούν τα αντικείμενα
- Σχέσεις: οι τρόποι με τους οποίους μπορούν να συσχετισθούν τα αντικείμενα μεταξύ τους

Στη συνέχεια θα ασχοληθούμε σύντομα με καθένα από αυτά τα στοιχεία ζεχωριστά.

- Τα στιγμιότυπα είναι τα βασικά τμήματα μιας οντολογίας και μπορούν να συμπεριλαβούν συγκεκριμένα αντικείμενα, όπως π.χ. άνθρωπος, ζώο, πίνακας, αυτοκίνητο, κ.λ.π., καθώς επίσης και αφηρημένα στιγμιότυπα, όπως αριθμούς και λέξεις. Αν ακολουθήσουμε μια αυστηρή θεώρηση μιας οντολογίας, αυτή δε χρειάζεται να περιλάβει κανένα στιγμιότυπο. Παρ' όλα αυτά ένας από τους γενικούς σκοπούς μιας οντολογίας είναι η παροχή ενός τρόπου κατηγοριοποίησης των στιγμιότυπων, ακόμα κι αν αυτά δεν αποτελούν ρητά μέρος της οντολογίας.

- Οι έννοιες είναι αφηρημένες ομάδες, σύνολα ή συλλογές αντικειμένων. Μπορούν να περιέχουν στιγμιότυπα, άλλες έννοιες ή και έναν συνδυασμό και των δύο. Μερικά (αυθαίρετα) παραδείγματα εννοιών περιλαμβάνουν:

- Πρόσωπο, η έννοια όλων των ανθρώπων
- Μόριο, η έννοια όλων των μορίων
- Αριθμός, η έννοια όλων των αριθμών
- Όχημα, η έννοια όλων των οχημάτων
- Αυτοκίνητο, η έννοια όλων των αυτοκινήτων
- Άτομο, που αντιπροσωπεύει την έννοια όλων των ατόμων
- Κατηγορία, που αντιπροσωπεύει την έννοια όλων των κατηγοριών
- Πράγμα, που αντιπροσωπεύει την έννοια όλων των πραγμάτων
- κ.ο.κ.

Οι οντολογίες ποικίλλουν ανάλογα με το εάν οι έννοιες που τις απαρτίζουν μπορούν να περιέχουν άλλες έννοιες, εάν μια έννοια μπορεί να ανήκει στον εαυτό της, εάν υπάρχει μια καθολική έννοια (δηλαδή μια έννοια που να περιέχει τα πάντα), κ.λ.π.. Μερικές φορές, βέβαια, επιβάλλονται στα παραπάνω κάποιοι αυθαίρετοι περιορισμοί, προκειμένου να αποφευχθούν ορισμένα γνωστά παράδοξα.

Οι έννοιες μιας οντολογίας μπορούν να είναι επεκτατικές (extensional) ή εντατικές (intensional). Μια έννοια είναι επεκτατική, εάν και μόνο εάν χαρακτηρίζεται απλώς από την ιδιότητα μέλους της. Ακριβέστερα, μια έννοια  $C$  είναι επεκτατική, εάν και μόνο εάν για οποιαδήποτε έννοια  $C'$  ισχύει ότι: εάν η  $C'$  έχει ακριβώς τα ίδια μέλη με την  $C$ , τότε οι  $C$  και  $C'$  είναι ίδιες. Εάν μια έννοια δεν ικανοποιεί αυτόν τον όρο, τότε είναι εντατική. Ενώ οι επεκτατικές έννοιες συμπεριφέρονται καλύτερα και είναι πιο κατανοητές από μαθηματική άποψη, δεν επιτρέπουν τις λεπτεπίλεπτες διακρίσεις που καλούνται να κάνουν συχνά οι οντολογίες. Παραδείγματος χάριν, μια οντολογία μπορεί να πρέπει να διακρίνει μεταξύ της έννοιας όλων των πλασμάτων με ένα νεφρό και της έννοιας όλων των πλασμάτων με μια καρδιά, ακόμα κι αν αυτές οι έννοιες συμβαίνει να έχουν ακριβώς τα ίδια μέλη. Επιπρόσθετα, μια έννοια μπορεί να εντάξει ή να ενταχθεί σε άλλες έννοιες. Για παράδειγμα, το όχημα εντάσσει το αυτοκίνητο, καθώς οτιδήποτε είναι μέλος της τελευταίας έννοιας είναι απαραίτητα και μέλος της προηγούμενης. Η σχέση υπαγωγής (subsumption) χρησιμοποιείται ευρέως, για να δημιουργήσει μια ιεραρχία εννοιών, στην οποία υπάρχει συνήθως μια ανώτατη γενική έννοια, όπως π.χ. το “Πράγμα” στην κορυφή, και πολύ συγκεκριμένες έννοιες, όπως π.χ. το “Rover 214” στα κατώτατα σημεία.

- Τα αντικείμενα μιας οντολογίας μπορούν να περιγραφούν με την ανάθεση ιδιότητων σε αυτά. Κάθε ιδιότητα έχει τουλάχιστον ένα όνομα και μια τιμή και χρησιμοποιείται, για να αποθηκεύσει τις πληροφορίες που είναι συγκεκριμένες για το αντικείμενο στο οποίο είναι συνημμένη. Παραδείγματος χάριν, το αντικείμενο “Rover 214” μπορεί να έχει ιδιότητες όπως:

- Όνομα: Rover 214
- Χρώμα: μπλε

- Πόρτες: 3
- Κυβισμός: {1400cc, 1600cc, 1800cc}
- Μετάδοση: 5-ταχύτητες

Η τιμή μιας ιδιότητας μπορεί να είναι ένας σύνθετος τύπος δεδομένων. Παρατηρούμε ότι στο παραπάνω παράδειγμα η τιμή της ιδιότητας “Κυβισμός” αποτελείται από μία λίστα τριών τιμών, αντί για μία και μόνο τιμή. Όταν δεν καθορίζονται ιδιότητες για τις έννοιες, τότε είτε έχουμε την περίπτωση της ταξινομίας (taxonomy) (εάν περιγράφονται και οι σχέσεις που διέπουν τις έννοιες μεταξύ τους), είτε έχουμε ένα αποκαλούμενο ελεγχόμενο λεξιλόγιο (controlled vocabulary), παρόλο που το τελευταίο δε θεωρείται αυθεντική οντολογία με τη στενή έννοια του όρου.

- Μια σημαντική χρήση των ιδιοτήτων είναι να περιγράφουν τις σχέσεις που υπάρχουν μεταξύ των αντικειμένων σε μια οντολογία. Συνήθως, μια σχέση είναι μια ιδιότητα, της οποίας η τιμή είναι είναι ένα άλλο αντικείμενο στην οντολογία. Παραδείγματος χάριν, στην οντολογία που περιέχει το “Rover 214” και το “Rover 200”, το αντικείμενο “Rover 200” μπορεί να έχει την ιδιότητα “Διάδοχος”: “Rover 214”. Κάτι τέτοιο μας καθιστά σαφές ότι το “Rover 200” είναι ο διάδοχος που αντικατέστησε το “Rover 214”. Ένα μεγάλο μέρος της δύναμης των οντολογιών προέρχεται από αυτή ακριβώς τη δυνατότητα περιγραφής αυτών των σχέσεων, το σύνολο των οποίων περιγράφει τη σημασιολογία μιας θεματικής περιοχής, κάτι ιδιαίτερα σημαντικό και χρήσιμο, όπως θα δούμε στο πλαίσιο της έρευνάς μας.

### 3.4.2 Οντολογίες και αναπαράσταση γνώσης

Προκειμένου να είναι δυνατή η χρησιμοποίηση μιας οποιασδήποτε οντολογίας για την αναπαράσταση της γνώσης, αυτή θα πρέπει να καθοριστεί, δηλαδή να αναπαρασταθεί και η ίδια μέσω μίας τυποποιημένης και κατανοητής από υπολογιστές αναπαράστασης. Για το σκοπό αυτό υπάρχει μία μεγάλη ποικιλία γλωσσών που είναι σε θέση να αναπαραστήσουν εννοιολογικά μοντέλα, καθεμία με τα δικά της χαρακτηριστικά όσον αφορά την εκφραστικότητά τους, την ευκολία χρήσης και την υπολογιστική πολυπλοκότητα που τις διέπει. Αναμφίβολα, η περιοχή της τεχνητής νοημοσύνης έχει μία μακρά παράδοση στην ανάπτυξη μεθόδων, εργαλείων και γλωσσών για τη δόμηση γνώσης και πληροφορίας, κάτι το οποίο αναπτύχθηκε και βελτιώθηκε τελευταία στα πλαίσια των δραστηριοτήτων που αναπτύσσονται γύρω από τον Σημασιολογικό Ιστό (Semantic Web)[314].

Ιδιαίτερο ρόλο κατά την επιλογή της αναπαράστασης παίζουν η εκφραστικότητα της γλώσσας κωδικοποίησης, η σημασιολογία της γλώσσας αναπαράστασης και η υπολογιστική πολυπλοκότητα που απαιτείται για την εξαγωγή συμπερασμάτων με βάση τη γνώση που μπορεί να αναπαρασταθεί. Ειδικότερα:

- Η εκφραστική δύναμη μιας γλώσσας αναπαράστασης γνώσης είναι ένα μέτρο του εύρους των δομών που μπορούν να χρησιμοποιηθούν, προκειμένου να περιγραφούν με τυπικό και ελαστικό τρόπο, ρητά και με ακρίβεια τα συστατικά μιας οντολογίας.
- Η σημασιολογία μιας γλώσσας αναφέρεται στην ερώτηση εάν είναι σαφές το τι σημαίνουν οι επιμέρους γλωσσικές δομές. Οι σαφείς ορισμοί της σημασιολογίας

και μία καλή κατανόηση αυτών είναι ουσιαστικοί, εάν η οντολογία πρόκειται να χρησιμοποιηθεί για την ανταλλαγή πληροφοριών.

- Η πολυπλοκότητα της εξαγωγής συμπερασμάτων είναι ένα μέτρο το οποίο περιγράφει το μέγεθος των υπολογιστικών πόρων που απαιτούνται, για να υπολογίσουν τις εκφράσεις στη δεδομένη γλώσσα αναπαράστασης γνώσης, και αναπόφευκτα συσχετίζεται με την εκφραστικότητα της χρησιμοποιούμενης γλώσσας. Ενώ η υψηλότερη εκφραστικότητα επιτρέπει συνήθως όλο και πιο σύνθετες δηλώσεις σχετικά με τη μοντελοποιούμενη περιοχή, απαιτεί περισσότερους υπολογιστικούς πόρους, για να υπολογίσει όλες τις πιθανές εκφράσεις.

### 3.4.3 Γλώσσες αναπαράστασης γνώσης: RDF/RDFS/OWL

Στα πλαίσια του Σημασιολογικού Ιστού έχουν αναπτυχθεί διάφορες γλώσσες αναπαράστασης γνώσης κατά τη διάρκεια των τελευταίων ετών, κυρίως ως γλώσσες οντολογιών, που είναι βασισμένες στο συντακτικό της γλώσσας XML [312]. Στη συνέχεια, θα εστιάσουμε σε τρεις από τις τρέχουσες γλώσσες αναπαράστασης γνώσης που αναπτύσσονται από τις ομάδες εργασίας της κοινοπραξίας του Παγκόσμιου Ιστού (World Wide Web) (W3C) [314]:

- την RDF [316]
- την RDFS [317]
- και την OWL [319]

οι οποίες και χρησιμοποιήθηκαν εκτενώς στα πλαίσια της παρούσης διδακτορικής διατριβής. Η RDFS αποτελεί ουσιαστικά μια απλή γλώσσα μοντελοποίησης πάνω από την RDF. Η ίδια η RDF δεν αποτελεί ένα σύστημα αναπαράστασης γνώσης, αλλά προσπαθεί να βελτιώσει τη διαλειτουργικότητα των δεδομένων στον Ιστό. Το τελευταίο επιτυγχάνεται με την εξειδίκευση του μοντέλου δεδομένων της XML μέσω ενός μοντέλου δεδομένων, το οποίο βασίζεται σε έναν γράφο παρόμοιο με το φορμαλισμό των σημασιολογικών δικτύων. Με τον τρόπο αυτό η RDF παρέχει ένα ελαφρύ πρότυπο αντικειμένων, προκειμένου να είναι δυνατός ο καθορισμός άλλων γλωσσών πάνω από την ίδια την RDF.

Αν προσπαθούσαμε να δώσουμε μια σύντομη περιγραφή της RDF, θα λέγαμε ότι η RDF αποτελείται από τρεις τύπους αντικειμένων: τους πόρους (resources), οι οποίοι περιγράφονται από εκφράσεις RDF και ακολουθούν την ονοματολογία των Uniform Resource Identifiers (URIs) και την κωδικοποίηση από την W3C του συντακτικού ονομάτων και διευθύνσεων, τόσο σημερινών όσο και μελλοντικών αντικειμένων στο Διαδίκτυο. Ακολούθως, κάποιες ιδιότητες της γλώσσας καθορίζουν συγκεκριμένες πτυχές, χαρακτηριστικά, ιδιότητες ή σχέσεις που χρησιμοποιούνται, για να περιγράψουν έναν πόρο και κάποιες δηλώσεις της ορίζουν μια τιμή για μια ιδιότητα σε έναν συγκεκριμένο πόρο (όπου αυτή η τιμή μπορεί να είναι μια άλλη δήλωση RDF).

Η RDFS, ως μία γλώσσα περιγραφής του λεξιλογίου της RDF, καθορίζει μία απλή γλώσσα μοντελοποίησης πάνω από την RDF. Η RDFS προσφέρει τα πρωτογενή στοιχεία για τον ορισμό των μοντέλων γνώσης, που είναι πιο κοντά στις βασισμένες σε πλαίσια προσεγγίσεις. Εισάγει ως βασικά της χαρακτηριστικά τις κατηγορίες, τις σχέσεις ένταξης (subsumption relationships) πάνω στις κατηγορίες και τις ιδιότητες

### Κεφάλαιο 3. Μαθηματική σημειολογία, όροι και έννοιες

και θέτει γενικούς περιορισμούς στα πεδία και στο εύρος των ιδιοτήτων. Ο προκύπτων φορμαλισμός επιτρέπει τη μοντελοποίηση οντολογιών ως μία ταξινομική δομή εννοιών και ιδιοτήτων. Επίσης, επιτρέπει τη μοντελοποίηση σχέσεων μεταξύ εννοιών ως ιδιότητες οι οποίες συνδέονται με κάποια έννοια.

Εν προκειμένω, στην παρούσα εργασία θα χρησιμοποιηθεί ευρέως και μια ζεχωριστή ιδιότητα της RDF/RDFS, το reification [44]. Η τελευταία χρησιμοποιείται, για να επιτύχουμε την επιθυμητή εκφραστικότητα και να αξιοποιήσουμε την αυξημένη λειτουργικότητα που εισάγει η ασάφεια. Όπως θα δούμε αναλυτικά σε επόμενα κεφάλαια της παρούσης, η αναπαράσταση ενός βαθμού εμπιστοσύνης που σχετίζεται με κάποια δυαδική σχέση πραγματοποιείται μέσω μιας επιπρόσθετης δήλωσης (statement) για την αρχική δήλωση, η οποία επιπρόσθετη δήλωση περιέχει την πληροφορία του βαθμού. Η αναπαράσταση της ασάφειας με τέτοιου είδους δηλώσεις αποτελεί ένα καινοτόμο αλλά συνάμα απολύτως αποδεκτό τρόπο, μιας και τελικά η εν λόγω δήλωση δε θα πρέπει να βεβαιώνεται αυτόματα. Για παράδειγμα, μια δήλωση της μορφής: “Σκηνή Μηχανοκίνητου Αθλητισμού μέρος Αυτοκίνητο”, η οποία σημαίνει ότι ένα αυτοκίνητο είναι μέρος μιας σκηνής μηχανοκίνητου αθλητισμού και η οποία συνοδεύεται από ένα βαθμό εμπιστοσύνης ίσο με 0.75, δε σημαίνει κατ’ ανάγκη ότι ένα αυτοκίνητο είναι πάντοτε μέρος μιας σκηνής μηχανοκίνητου αθλητισμού, αλλά, όπως είναι φυσικό, μπορεί να βρεθεί και κάτω από άλλο εννοιολογικό πλαίσιο.

#### Πίνακας 3.1: Ασαφής αναπαράσταση σχέσης σε RDF: RDF reification.

```
<rdf:Description rdf:about="#s1">
<rdf:subject rdf:resource="#&dom;wrc"/>
<rdf:predicate rdf:resource="#&dom;specializationOf"/>
<rdf:object> rdf:resource="#&dom;rally" </rdf:object>
<rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Statement"/>
<context:specializationOf rdf:datatype="http://www.w3.org/2001/XMLSchema#float">0.90</context:specializationOf>
</rdf:Description>
```

Ένα μικρό αντιπροσωπευτικό παράδειγμα αυτής της τεχνικής δίδεται στον Πίνακα 3.1, επιλέγοντας ένα στιγμιότυπο της MPEG-7 σχέσης εξειδίκευσης *Sp*. Όπως ορίζεται στο πρότυπο:  $Sp(x, y) > 0$  σημαίνει ότι η ερμηνεία του  $x$  “περιλαμβάνει” την ερμηνεία του  $y$ . οι συνήθεις μορφές εξειδίκευσης είναι η υπο-κατηγοριοποίηση (sub-classing), δηλαδή το  $x$  είναι μια γενίκευση του  $y$ , και η θεματική κατηγοριοποίηση, δηλαδή το  $x$  είναι η θεματική κατηγορία του  $y$ . Στο παράδειγμά μας, το υποκείμενο της RDF *wrc* (Παγκόσμιο Πρωτάθλημα Ράλλυ) έχει τη σχέση *specializationOf* ως RDF κατηγόρημα και το *rally* αποτελεί το RDF αντικείμενο. Επιπρόσθετα, η προτεινόμενη τεχνική του reification εισάγει μια δήλωση για την προηγούμενη δήλωση/τριπλέτα πάνω στον πόρο *specializationOf*, δηλώνοντας ότι το 0.90 είναι ο βαθμός ιδιότητας μέλους αυτής της σχέσης. Η αναπαράσταση, τώρα, του παραδείγματος αυτού στη γλώσσα OWL δεν είναι μια απλή διαδικασία λόγω των πολλών περιορισμών και κανόνων επικύρωσής που τη διέπουν. Εντούτοις, το μοντέλο που προτείνουμε είναι πλήρως επικυρωμένο και σύμφωνο με όλους τους συντακτικούς κανόνες της OWL. Επομένως, το παραπάνω παράδειγμα μεταφράζεται τελικά στη γλώσσα OWL στο Σχήμα 3.3.

```

<rdf:RDF
  xmlns:rdf = "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#" ... >
<owl:Class rdf:ID="beach"> <rdfs:subClassOf
rdf:resource="#holidays" />
...
</owl:Class> <owl:Class rdf:ID="motorsports"/> ...
<owl:Class rdf:ID="sand"/> <rdf:Description rdf:about="s1">
<rdf:subject rdf:resource="#wrc"/>
<rdf:predicate rdf:resource="#specializationOf"/>
<rdf:object rdf:resource="#rally"/>
<rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Statement"/>
<context:isRelated rdf:datatype="http://www.w3.org/2001/XMLSchema
#float">0.90</context:isRelated>
</rdf:Description> ... </rdf:RDF>

```

Σχήμα 3.3: Ασαφής αναπαράσταση σχέσης σε OWL: OWL reification.

### 3.4.4 Υποδομή οντολογιών - εργαλεία και υποστήριξη

Η επιτυχής δημιουργία και εκμετάλλευση μιας υποδομής γνώσης απαιτούν αναμφίβολα την υποστήριξη εργαλείων και συστημάτων σε όλα τα στάδια της γνώσης κατά τη διάρκεια του κύκλου ζωής της υποδομής. Ο τελευταίος περιλαμβάνει αφενός τη δημιουργία, την αναπροσαρμογή, την ενημέρωση και την εξέλιξη των οντολογιών και των βάσεων γνώσεων γενικότερα και αφετέρου ορισμένα προηγμένα συστήματα βάσεων δεδομένων, καθώς και μηχανές εζαγωγής συμπερασμάτων, προκειμένου να εκμεταλλευτούν τις πηγές της γνώσης. Στο πλαίσιο της παρούσης θα ασχοληθούμε μόνο με το πρώτο σκέλος.

Γενικά, τα εργαλεία και τα συστήματα αυτής της περιοχής είναι αρκετά διασκορπισμένα και ασύνδετα μεταξύ τους. Πολλά εργαλεία είναι διαθέσιμα μόνο ως ερευνητικά πρωτότυπα, ενώ άλλα έχουν αρχίσει να εμπορευματοποιούνται τα τελευταία χρόνια. Η συμβατότητα με τις γλώσσες RDF/RDFS ολοένα και αυξάνεται, αλλά εξαιτίας των παραλλαγών του συντακτικού των επιλεγμένων γλωσσών αναπαράστασης και των συγκεκριμένων σχεδιαστικών επιλογών των εργαλείων τα τελευταία δεν είναι ακόμα αρκετά αποδοτικά κατά την χρήση. Εντούτοις, τα περισσότερα σύγχρονα συστήματα υποστηρίζουν ικανοποιητικά τους ώριμους φορμαλισμούς, όπως η RDF και η OWL, κάτι το οποίο καθιστά τις γλώσσες αυτές μια ελπιδοφόρο επιλογή.

Όσον αφορά στα εργαλεία υποστήριξης μιας οντολογίας, όπως π.χ. είναι τα εργαλεία διόρθωσης και αλλαγών, αυτά παρέχουν στους χρήστες τη δυνατότητα να δημιουργήσουν και να διατηρήσουν οντολογίες, ενίστε και με τη βοήθεια ενός γραφικού περιβάλλοντος. Ευρέως διαδεδομένα παραδείγματα τέτοιων εργαλείων αποτελούν το KAON Ontology Infrastructure Modeler [302] και το OntoEdit[315]: το OntoEdit έχει αναπτυχθεί από την OnToPrise GmbH, ενώ το KAON OI-Modeler αποτελεί το βασικό εργαλείο υποστήριξης του Πανεπιστημίου της Καρλσρούης. Το τελευταίο χρησιμοποιήθηκε και για τη δημιουργία της τελικής έκδοσης των οντολογιών που παρουσιάζονται στο παρόν κείμενο, καθώς παρέχει ένα πλήρες πλαίσιο αποδοτικής και αξιόπιστης μοντελοποίησης, ανάπτυξης, διαχείρισης και επαλήθευσης της ορθότητας μιας οντολογίας.

### 3.5 Σημασιολογικές Σχέσεις

Τα τελευταία χρόνια είναι γεγονός ότι έχει υπάρξει μια ιδιαίτερα σημαντική αύξηση στο διαθέσιμο πολυμεσικό περιεχόμενο και στις τεχνολογίες πρόσβασης σε αυτό. Εντούτοις, η εξαγωγή χρήσιμης πληροφορίας από το πολυμεσικό περιεχόμενο και η αποδοτική εφαρμογή αυτής σε πρακτικά συστήματα, όπως π.χ. είναι οι μηχανές αναζήτησης πολυμεσικού περιεχομένου, παραμένουν ακόμα και σήμερα ανοικτά ερευνητικά πεδία. Στο πλαίσιο αυτό το πρότυπο MPEG-7 έχει προτυποποιήσει μια σειρά εργαλείων, που είναι κατάλληλα για τις διάφορες πτυχές και τα επίπεδα αφαίρεσης που απαντώνται στα πολυμέσα. Τα εργαλεία αυτά είναι σε θέση να περιγράψουν τη σημασιολογία συγκεκριμένων στιγμιοτύπων πολυμέσων, όπως π.χ. είναι μια εικόνα ή ένα κομμάτι βίντεο, αλλά μπορούν επίσης και να γενικεύσουν αυτές τις περιγραφές είτε για πολλαπλά τέτοια στιγμιότυπα, είτε για ένα σύνολο σημασιολογικών περιγραφών. Τα βασικά συστατικά των σημασιολογικών περιγραφών είναι οι σημασιολογικές οντότητες (όπως π.χ. τα αντικείμενα και τα γεγονότα), οι ιδιότητες αυτών των οντοτήτων (όπως π.χ. οι ετικέτες) και, τελικά, οι σχέσεις αυτών των οντοτήτων (όπως π.χ. ένα αντικείμενο που είναι ο ασθενής (patient) ενός γεγονότος). Η περιγραφική δύναμη και οι δυνατότητες επαναχρησιμοποίησης αυτών των εργαλείων έχουν παρουσιαστεί ευρέως σε πολυάριθμα πειράματα και εφαρμογές.

Το πρότυπο MPEG-7 έχει προτυποποιήσει έναν αριθμό κοινών και ευρέως χρησιμοποιούμενων σήμερα κανονιστικών (σαφών) σημασιολογικών σχέσεων (βλ. Παράρτημα B), όπως π.χ. η σχέση πράκτορας, αλλά επιτρέπει την περιγραφή και μη-κανονιστικών σχέσεων. Τα εργαλεία σημασιολογικών σχέσεων του προτύπου περιλαμβάνουν το SemanticRelation CS, το οποίο καθορίζει τις σημασιολογικές σχέσεις που διέπουν τις οντότητες εκείνες που περιέχουν σημασιολογική πληροφορία. Αυτές οι κανονιστικές σημασιολογικές σχέσεις μπορούν να ισχύσουν, μεταξύ άλλων, για οντότητες, όπως είναι οι σημασιολογικές οντότητες, οι συλλογές, τα πρότυπα και τα τμήματα.

Οι κανονιστικές σημασιολογικές σχέσεις μπορούν να περιγράψουν τον τρόπο με τον οποίο συνδέονται διάφορες σημασιολογικές οντότητες μεταξύ τους στο πλαίσιο ενός αφηγήματος ή μιας ιστορίας (π.χ. σχέσεις πράκτορας, ασθενής, συνοδός, κ.α.). Οι σημασιολογικές σχέσεις μπορούν επίσης να περιγράψουν τις σχέσεις μεταξύ των αντικειμένων ή/και των γεγονότων που βασίζονται σε χαρακτηριστικές γλωσσικές σχέσεις, όπως ένα αντικείμενο που μπορεί να είναι μεταξύ άλλων ο πράκτορας, ο ασθενής, ο συνοδός, το ερέθισμα, η αιτία, ο στόχος, ή το όργανο ενός γεγονότος, ή ένα γεγονός το οποίο είναι το αποτέλεσμα ή μια περίληψη ενός άλλου γεγονότος.

Άλλες κανονιστικές σημασιολογικές σχέσεις μπορούν να περιγράψουν πώς συνδέονται οι ορισμοί διάφορων σημασιολογικών οντοτήτων μεταξύ τους (π.χ. ο συνδυασμός, ειδικεύεται και εξηγεί). Ειδικότερα, οι σημασιολογικές σχέσεις μπορούν να περιγράψουν μια σημασιολογική οντότητα, που είναι ο συνδυασμός των εννοιών δύο ή περισσότερων σημασιολογικών οντοτήτων, ή μια σημασιολογική οντότητα που είναι μεταξύ άλλων μια ειδίκευση, παρόμοια κατά έννοια με, αντίθετη κατά έννοια με, ή ένα παράδειγμα μιας άλλης σημασιολογικής οντότητας.

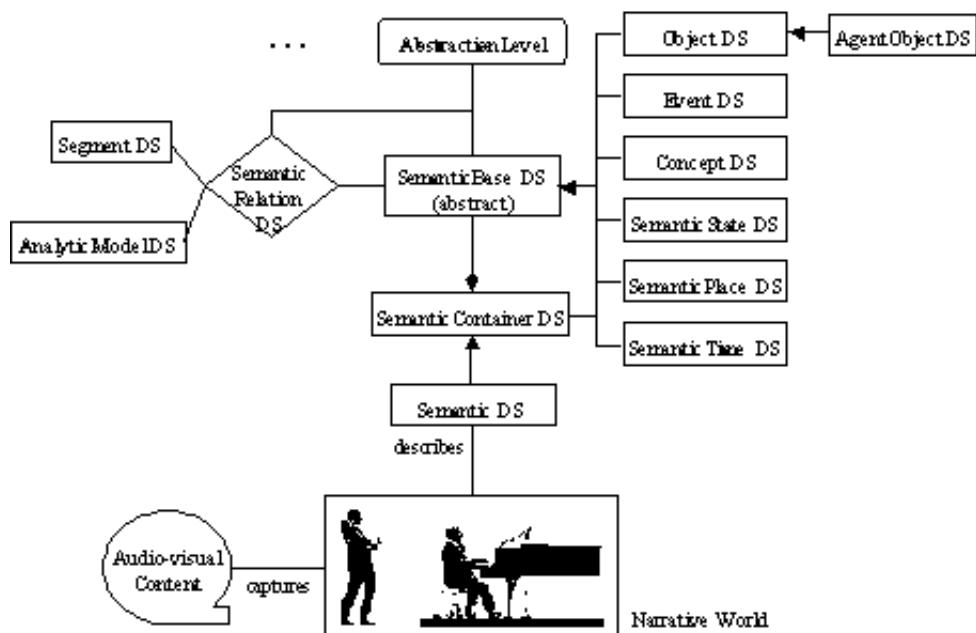
Συνάμα, υπάρχουν και σημασιολογικές οντότητες που μπορούν να περιγράψουν το χωρικό ή και χρονικό προσδιορισμό των σημασιολογικών οντοτήτων σε τμήματα, πρότυπα, συλλογές ή άλλες σημασιολογικές οντότητες (π.χ. οι οντότητες: απεικονίζει, συμβολίζει και πλαίσιο). Παραδείγματος χάριν, οι σημασιολογικές σχέσεις μπορούν να περιγράψουν: ένα γεγονός που πραγματοποιείται σε έναν συγκεκριμένο

τόπο και χρόνο, ένα αντικείμενο που γίνεται αντιληπτό ή που συμβολίζεται σε ένα τμήμα ή μια συλλογή, έναν αφηγηματικό κόσμο που είναι το πλαίσιο ή η ερμηνεία ενός άλλου κ.ο.χ..

Τέλος, υπάρχουν εργαλεία MPEG-7 για την περιγραφή γραφικών (π.χ., ισοδύναμο), βασικών (π.χ., μέλος), χωρικών (π.χ., αριστερό), και χρονικών (π.χ., πριν) σχέσεων, οι οποίες μπορούν επίσης να περιγράψουν σχέσεις μεταξύ των σημασιολογικών οντοτήτων. Ο αναγνώστης καλείται να ανατρέξει στο Παράρτημα Β για τον πλήρη κατάλογο των κανονιστικών (σαφών) σχέσεων που καθορίζει το πρότυπο.

### 3.6 Σημασιολογική Περιγραφή Περιεχομένου

Τα Σχέδια Περιγραφής Πολυμέσων (MDSs) είναι εν γένει δομές μεταδεδομένων για την περιγραφή και το σχολιασμό του πολυμεσικού περιεχομένου. Καθώς η εννοιολογική πτυχή μιας πολυμεσικής περιγραφής μπορεί να οργανωθεί σε ένα δέντρο ή σε ένα γράφο, η δομή μιας τέτοιας αναπαράστασης γράφου καθορίζεται από ένα σύνολο κόμβων, που αντιπροσωπεύουν τις σημασιολογικές έννοιες, και ένα σύνολο ακμών, που διευκρινίζουν τη σχέση μεταξύ των κόμβων. Οι ακμές περιγράφονται αναλυτικά μέσω του Περιγραφικού Σχήματος της Σημασιολογική Σχέσης (Semantic Relation DSs), το οποίο αποτελεί μέρος του προτύπου MPEG-7, διαφαίνεται και από το παράδειγμα του Σχήματος 3.4 και περιγράφει με σαφήνεια τις σημασιολογικές σχέσεις που ορίζει το πρότυπο.



**Σχήμα 3.4:** Εργαλεία του προτύπου MPEG-7 για την περιγραφή των εννοιολογικών πτυχών του πολυμεσικού περιεχομένου.

Εκτός από τη σημασιολογική περιγραφή μεμονωμένων περιπτώσεων μέσα στο οπτικοακουστικό πολυμεσικό περιεχόμενο, τα σημασιολογικά Περιγραφικά Σχήματα του προτύπου MPEG-7 παρέχουν επίσης έναν τυποποιημένο τρόπο περιγραφής (σε XML) των σημαντικών εννοιών που είναι σχετικές με την οπτικοακουστική περιγραφή και τη διαχείριση περιεχομένου, προκειμένου να διευκολυνθούν η αναζήτηση, η ευρετηρίαση, το φιλτράρισμα και η πρόσβαση στην πληροφορία. Τα DSs καθορίζονται

χρησιμοποιώντας την MPEG-7 γλώσσα καθορισμού περιγραφής (DDL), η οποία βασίζεται στη γλώσσα XML Schema[313], και αρχικοποιούνται ως έγγραφα ή ρεύματα. Οι προκύπτουσες περιγραφές μπορούν να εκφραστούν σε μορφή κειμένου (δηλαδή σε ανθρωπίνως κατανοητή XML, προκειμένου να είναι δυνατή η διόρθωση, η αναζήτηση και το φιλτράρισμα) ή σε συμπιεσμένη δυαδική μορφή (έτσι ώστε να είναι δυνατή η αποθήκευση και η μετάδοση). Στην επόμενη υποενότητα, παρέχουμε μια σύντομη επισκόπηση των Σχεδίων Περιγραφής Πολυμέσων του MPEG-7 και περιγράφουμε τη λειτουργία και τη χρήση τους στις εφαρμογές πολυμέσων. Ο στόχος της παρούσης είναι η επισκόπηση εκείνων των Σχεδίων Περιγραφής Πολυμέσων, που είναι σχετικά με το αντικείμενο της προτεινόμενης ερευνητικής εργασίας και για το λόγο αυτό θα περιοριστούμε μόνο στις δομές μεταδεδομένων του MPEG-7, που θα μας είναι χρήσιμες περαιτέρω και όχι σε όλα τα Περιγραφικά Σχήματα που αυτό περιέχει.

### 3.6.1 Περιγραφικά σχήματα

Ο στόχος του προτύπου MPEG-7 είναι να επιτραπεί η αναζήτηση, η ευρετηρίαση, το φιλτράρισμα και η πρόσβαση του πολυμεσικού περιεχομένου με τη διευκόλυνση της διαλειτουργικότητας μεταξύ των συσκευών και των εφαρμογών που ασχολούνται με την περιγραφή του. Το MPEG-7 ασχολείται με συγκεκριμένα χαρακτηριστικά γνωρίσματα του πολυμεσικού περιεχομένου, καθώς επίσης και με την πληροφορία που σχετίζεται με τη διαχείριση του τελευταίου. Όπως ειπώθηκε προηγουμένως, οι περιγραφές του MPEG-7 λαμβάνουν δύο πιθανές μορφές: (i) μια απλή μορφή κειμένου σε XML, κατάλληλη για διορθώσεις, αναζητήσεις και φιλτράρισμα, και (ii) μια δυαδική μορφή, κατάλληλη για αποθήκευση, μετάδοση και ομαλή ροή του πολυμεσικού περιεχομένου. Συνολικά, το πρότυπο καθορίζει τέσσερις τύπους κανονιστικών στοιχείων: Περιγραφείς, Σχέδια Περιγραφής (DSs), μια Γλώσσα Καθορισμού Περιγραφής (DDL) και Σχέδια Κωδικοποίησης.

Οι Περιγραφείς του MPEG-7 έχουν ως κύριο σκοπό να περιγράψουν τα χαμηλού επιπέδου ακουστικά ή οπτικά χαρακτηριστικά γνωρίσματα, όπως το χρώμα, η υφή, η κίνηση, η ακουστική ενέργεια και ούτω καθ' εξής, καθώς επίσης και ιδιότητες του οπτικοακουστικού περιεχομένου, όπως η θέση, ο χρόνος, η ποιότητα, κ.ο.κ.. Όπως αναμένεται οι περισσότεροι Περιγραφείς για τα χαμηλού επιπέδου χαρακτηριστικά γνωρίσματα εξάγονται αυτόματα κατά τη χρήση τους στις εφαρμογές.

Από την άλλη πλευρά, τα MPEG-7 DSs έχουν ως κύριο σκοπό να περιγράψουν τα υψηλότερου επιπέδου οπτικοακουστικά χαρακτηριστικά γνωρίσματα, όπως οι περιοχές, τα τμήματα, τα αντικείμενα, τα γεγονότα και άλλα αμετάβλητα μεταδεδομένα που αφορούν τη δημιουργία, την παραγωγή και τη χρήση πολυμεσικού περιεχομένου. Τα DSs παράγουν πιο σύνθετες περιγραφές με το να ενσωματώνουν μαζί πολλαπλούς περιγραφείς και DSs, καθώς και με τη δήλωση των σχέσεων που διέπουν τα τμήματα περιγραφής. Στο πλαίσιο του προτύπου MPEG-7, τα DSs είναι ταξινομημένα ανάλογα με το αν αναφέρονται σε πολυμέσα, σε ήχο ή αποκλειστικά στην οπτική περιοχή. Στην τυπική περίπτωση τα πολυμεσικά DSs περιγράφουν το περιεχόμενο που αποτελείται από έναν συνδυασμό ακουστικών και οπτικών στοιχείων, ενδεχομένως δε και κάποιων στοιχείων κειμένου, ενώ τα ακουστικά ή τα οπτικά DSs αναφέρονται συγκεκριμένα στα χαρακτηριστικά γνωρίσματα που είναι μοναδικά στην ακουστική ή την οπτική περιοχή αντίστοιχα. Σε μερικές περιπτώσεις μπορούν να χρησιμοποιηθούν αυτόματα εργαλεία για την αρχικοποίηση των DSs, αλλά σε πολλές περιπτώσεις η αρχικοποίηση αυτή απαιτεί την ανθρώπινη παρέμβαση και τη χρήση υποβοηθούμενων

εργαλείων εξαγωγής ή δημιουργίας.

Εν γένει, τα Σχέδια Περιγραφής Πολυμέσων του MPEG-7 παρέχουν τα δομικά στοιχεία για τον καθορισμό των μεταδεδομένων για την περιγραφή του πολυμεσικού περιεχομένου και των υπηρεσιών του. Τα MPEG-7 MDS έχουν ορισθεί χρησιμοποιώντας την MPEG-7 Γλώσσα Καθορισμού Περιγραφής, που είναι ουσιαστικά η γλώσσα σχημάτων XML [84], [313], εμπλουτισμένη με τους βασικούς τύπους δεδομένων που είναι απαραίτητοι για τον καθορισμό των σχεδίων περιγραφής (DSs) των MPEG-7 MDS. Τα DSs είναι ουσιαστικά σύνθετοι τύποι δεδομένων, που χρησιμοποιούνται για την περιγραφή των εννοιών στο πεδίο τους. Τα MPEG-7 MDS αποτελούνται από τα ακόλουθα σημαντικά συστατικά [296]:

1. Τα βασικά στοιχεία που περιλαμβάνουν τα εργαλεία σχημάτων, τους βασικούς τύπους δεδομένων, τις μαθηματικές δομές, τα εργαλεία σύνδεσης και εντοπισμού μέσων, καθώς επίσης και βασικά DSs, που χρησιμοποιούνται ως στοιχειώδη συστατικά πιο σύνθετων DSs.
2. Τα στοιχεία περιγραφής και διαχείρισης περιεχομένου, που παρέχουν (i) εργαλεία δημιουργίας, παραγωγής και χρήσης, προκειμένου να συλλεχθεί πληροφορία διαχείρισης περιεχομένου και (ii) εργαλεία δομικών και σημασιολογικών πτυχών, προκειμένου να συλλεχθεί πληροφορία περιγραφής περιεχομένου.
3. Τα στοιχεία πλοήγησης και πρόσβασης, όπου η πλοήγηση (browsing) υποστηρίζεται μέσω περιλήψεων του πολυμεσικού περιεχομένου και διαφόρων παραλλαγών τους.
4. Τα στοιχεία αλληλεπίδρασης χρηστών που χρησιμοποιούνται, για να περιγράφουν τις προτιμήσεις των χρηστών σχετικά με το πολυμεσικό περιεχόμενο και τα οποία θα μας απασχολήσουν σε επόμενα κεφάλαια της παρούσης.
5. Τα στοιχεία οργάνωσης περιεχομένου, όπου η οργάνωση του πολυμεσικού περιεχομένου εξετάζεται υπό το πρίσμα της ταξινόμησης, της διαμόρφωσης και των συλλογών πολυμεσικών εγγράφων.

Τέλος, εκτός από τα DSs, τα MPEG-7 MDS παρέχουν τα λεγόμενα Σχέδια Ταξινόμησης (CSs), τα οποία είναι στην ουσία ένα είδος θησαυρών (thesauri) που αποτελούνται από ιεραρχίες όρων. Τα 40 τυποποιημένα CSs που παρέχονται στα MPEG-7 MDS περιέχουν τις τιμές των συγκεκριμένων δομικών στοιχείων (ιδιότητες ή απλά στοιχεία τύπων) των DSs των MPEG-7 MDS.

### 3.6.2 Αναπαράσταση περιγραφικών σχημάτων

Τελευταία, εξαιρετικά ενδιαφέρουσες ερευνητικές προσεγγίσεις έχουν καταβάλει μια σημαντική προσπάθεια για την αναπαράσταση όλου ή μέρους του προτύπου MPEG-7 και των Περιγραφικών Σχημάτων του σε μια οντολογία βασισμένη στις γλώσσες RDF/OWL. Μπορούμε να διακρίνουμε τέσσερις αξιόλογες προσπάθειες:

- Την Αγώτερη MPEG-7 MDS Οντολογία από την Hunter [124]. Η οντολογία αυτή είναι χρονολογικά η πρώτη απόπειρα αναπαράστασης και αρχικά είχε αναπτυχθεί στη γλώσσα RDFS [317], ενώ κατόπιν μετατράπηκε σε DAML+OIL [289] και πλέον είναι διαθέσιμη σε OWL-Full [319]. Καλύπτει το πάνω μέρος των MDS του προτύπου MPEG-7 και αποτελείται από περίπου 60 κατηγορίες και 40 ιδιότητες.

### *Κεφάλαιο 3. Μαθηματική σημειολογία, όροι και έννοιες*

- Την MPEG-7 MDS Οντολογία από την Τσιναράκη [261]. Ξεκινώντας από την προηγούμενη οντολογία, αυτή η οντολογία καλύπτει το πλήρες μέρος των Σχεδίων Περιγραφής Πολυμέσων του προτύπου MPEG-7. Περιέχει 420 κατηγορίες και 175 ιδιότητες και είναι γραμμένη σε OWL DL.
- Την MPEG-7 Οντολογία από το Rhizomik [95]. Αυτή η MPEG-7 οντολογία έχει παραχθεί αυτόματα από το πρότυπο MPEG-7, προκειμένου να του δοθεί μια επίσημη σημασιολογία. Για το σκοπό αυτό υλοποιήθηκε μια γενική χαρτογράφηση XSD2OWL και οι ορισμοί των τύπων της XML Schema και των στοιχείων του προτύπου του ISO έχουν μετατραπεί σε ορισμούς της OWL, σύμφωνα με τον πίνακα που δίνεται στο [95]. Αυτή η οντολογία αποσκοπεί να καλύψει ολόκληρο το πρότυπο και αποτελεί έτσι την πληρέστερη, σε σχέση με τις προαναφερθείσες, προσέγγιση. Περιέχει 2372 κατηγορίες και 975 ιδιότητες και είναι μια OWL-Full οντολογία.
- Την Οντολογία INA [127]. Αυτή η οντολογία δεν είναι πραγματικά μια MPEG-7 οντολογία, δεδομένου ότι δεν καλύπτει ολόκληρο το πρότυπο. Είναι μάλλον μια οπτικοακουστική οντολογία που εμπνέεται από διάφορες ορολογίες, είτε τυποποιημένες (όπως π.χ. το MPEG-7 και το TV-Anytime), είτε ακόμα υπό ανάπτυξη (όπως το ProgramGuideML). Αυτή η οντολογία περιέχει αυτήν την περίοδο 1100 κατηγορίες και 220 ιδιότητες και είναι και αυτή μια OWL-Full οντολογία.

□



# Κεφάλαιο 4

## Εξόρυξη δεδομένων και κατηγοριοποίηση

### 4.1 Εισαγωγή

Στο κεφάλαιο αυτό παρουσιάζουμε την ερευνητική συνεισφορά μας στο πεδίο της επεξεργασίας και ανάκτησης πολυμεσικού περιεχομένου από τη σκοπιά της κατηγοριοποίησης πολυμεσικών εγγράφων σε σημασιολογικές κατηγορίες, λαμβάνοντας, φυσικά, υπόψη το εννοιολογικό πλαίσιο, το οποίο και αποτελεί τον πυρήνα της παρούσας διατριβής. Αρχικά, θα παρουσιάσουμε ορισμένες θεμελιώδεις μεθοδολογίες που είναι σχετικές με το ερευνητικό πεδίο της διαχείρισης γνώσης και εξόρυξης δεδομένων. Η προτεινόμενη διαδικασία θεματικής κατηγοριοποίησης παρουσιάζεται κατόπιν μέσα από το πρίσμα τόσο των κλασικών προσεγγίσεων από το χώρο της μηχανικής μάθησης και της τεχνητής νοημοσύνης, όσο και της αξιοποίησης του εννοιολογικού πλαισίου στη διαδικασία. Οι τεχνικές και αλγόριθμοι που παρουσιάζονται, είτε χρησιμοποιήθηκαν με μικρές αλλαγές και τροποποιήσεις, είτε επεκτάθηκαν και εμπλουτίστηκαν ριζικά για την αντιμετώπιση των προβλημάτων που θέτει η ερευνητική περιοχή. Ειδικότερα, αξιοποιήσαμε και επεκτείναμε βασικές αρχές και αλγορίθμους εξόρυξης δεδομένων και γνώσης, όπως είναι π.χ. οι τεχνικές συσταδοποίησης και κατηγοριοποίησης δεδομένων, ενώ σε περιπτώσεις, όπου η έρευνά μας το απαίτησε, προτείνουμε καινοτόμες προσθήκες και αλλαγές, όπως είναι η μείωση των διαστάσεων κατά τη συσταδοποίηση, προκειμένου να βελτιστοποιήσουμε την απόδοση και την επίδοσή τους.

#### 4.1.1 Υπάρχουσα κατάσταση - Επισκόπηση ερευνητικής περιοχής

Η πιο πρόσφατη τάση που καταγράφεται στο χώρο της μηχανικής μάθησης είναι η απορρόφηση κάποιων εκ των κλασικών τεχνικών της από ευρέως διαδεδομένες (εμπορικές) εφαρμογές. Είναι αλήθεια ότι η μηχανική μάθηση σχετιζόταν ανέκαθεν με εφαρμογές, αλλά συνήθως οι τελευταίες υλοποιούνταν αποκλειστικά από ερευνητές που ήθελαν να καταδείξουν τη βιωσιμότητα της προσέγγισής τους πάνω σε ένα πραγματικό πρόβλημα. Σήμερα, οι ερευνητές καινοτόμων περιοχών, όπως π.χ. η βιοπληροφορική, αναγνωρίζουν όλο και περισσότερο τη χρησιμότητα των τεχνικών της μηχανικής μάθησης στην έρευνά τους, παράλληλα φυσικά με τις υπόλοιπες παραδοσιακές τεχνικές της στατιστικής. Ιδανικά, η μηχανική μάθηση μπορεί να αποτελέσει ακόμα ένα εργαλείο στην εργαλειοθήκη ενός αναλυτή δεδομένων, κάτι το οποίο θα την καθιστούσε ίσως μία από τις πιο επιτυχημένες εκφάνσεις της τεχνητής νοημο-

σύνης και η αλήθεια είναι ότι σημαντική πρόοδος έχει σημειωθεί τελευταία προς τον στόχο αυτόν. Ένα βασικό σημείο που θα κρίνει την επιτυχία ή όχι των εφαρμογών μηχανικής μάθησης είναι ο βαθμός, με τον οποίο είναι σε θέση να προσαρμόζονται με ευκολία στη θεματική περιοχή της εκάστοτε έρευνας. Κι αυτό, γιατί η ίδια η δυναμική των “έτοιμων” και τελείως ανεξάρτητων μεθόδων μπορεί να αποτελέσει συγχρόνως και αδυναμία, και όσο πιο περίπλοκη είναι η θεματική περιοχή της εφαρμογής, τόσο πιο αισθητή γίνεται η αδυναμία αυτή. Εν γένει, οι ερευνητές δε χρειαζόμαστε μια γενικής χρήσης μετρική απόστασης, αλλά αντίθετα χρειαζόμαστε μία που να λειτουργεί καλά, όταν εφαρμοστεί π.χ. σε ακολουθίες DNA και μία άλλη που να λειτουργεί εξίσου καλά, όταν εφαρμοστεί π.χ. σε έγγραφα κειμένου.

Μερικές από τις πιο εμπειριστατωμένες προσπάθειες περιγραφής της περιοχής της μηχανικής μάθησης και της εξόρυξης δεδομένων περιλαμβάνονται στη βιβλιογραφία. Ξεχωρίζουμε τρία από τα πιο πρόσφατα εγχειρίδια για τη μηχανική μάθηση. Ειδικότερα, το βιβλίο του Mitchell [175] αποτελεί ένα λεπτομερές εγχειρίδιο και θεωρείται το πιο καθιερωμένο στο χώρο. Ο Langley [148] από την άλλη, ενώ παρέχει μια κάπως πιο προσωπική άποψη για τον τομέα της μηχανικής μάθησης, διέπεται από ένα άριστο μαθηματικό υπόβαθρο. Τέλος, το βιβλίο των Witten και Frank [280] συνοδεύεται από μια σουίτα εργαλείων και εφαρμογών υλοποιημένων στη γλώσσα προγραμματισμού Java [301] και αποτελεί μια πρώτης τάξης αναφορά για τη μετάβαση από τη θεωρία στην πράξη. Παρόλο που το ύφος της παρουσίασης είναι λιγότερο ακαδημαϊκό σε σχέση με αυτό που ακολουθεί ο Mitchell, η εργαλειοθήκη που παρέχεται είναι αξιομνημόνευτη. Επιπλέον, το βιβλίο καλύπτει μερικές πτυχές της μηχανικής μάθησης, όπως η εκμάθηση κανόνων συσχέτισης (association rule learning) και οι μηχανές υποστήριξης διανυσμάτων (support vector machines), που δεν αντιμετωπίζονται στα προηγούμενα συγγράμματα. Τέλος, θα πρέπει να αναφερθεί ότι και ο Weiss στο [275] στοχεύει στη δημιουργία μιας πεπατημένης για την ανάπτυξη εφαρμογών εξόρυξης δεδομένων.

Από μια άλλη σκοπιά, τώρα, μπορεί να ισχυριστεί κανείς ότι η εξόρυξη δεδομένων “έγινε της μόδας” και, παρόλο που η εμπορική εκμετάλλευσή της μπορεί να είναι εν γένει κάτι καλό, καθώς δημιουργεί ενδιαφέροντας γι' αυτή και ίσως χρηματοδοτεί και σχετικές έρευνες, η καλλιέργεια ουτοπικών προσδοκιών και οι συναχόλουθες απογοητεύσεις μπορεί να οδηγήσουν τελικά στην ολική απώλεια ενδιαφέροντος για το χώρο (σε πλήρη αναλογία με το τι συνέβη π.χ. στην περιοχή των έμπειρων συστημάτων τις δεκαετίες του 1970 και του 1980). Σε αυτό το πλαίσιο αξίζει να αναφερθεί η εργασία των Barry και Linoff [28]. Μια πιο απλοϊκή προσέγγιση στο θέμα ακολουθείται από τον Adriaans [2], όπου δίδεται ιδιαίτερη προσοχή στη δημιουργία ενός περιβάλλοντος εξόρυξης δεδομένων και στη χρήση εφαρμογών βγαλμένων από την πραγματική ζωή. Τέλος, δε θα πρέπει να αμελήσουμε και τη συμβολή της μηχανικής μάθησης στα πεδία εκ πρώτης όψεως άσχετων επιστημών, όπως π.χ. η βιοπληροφορική, η βιοϊατρική και η βιολογία. Ειδικά για τον ερευνητή της μηχανικής μάθησης, τέτοιες αναφορές είναι σημαντικές τουλάχιστον για δύο λόγους. Πρώτον, γιατί καταδεικνύουν ότι η μηχανική μάθηση έχει κερδίσει το σεβασμό ως μία ιδιαίτερα χρήσιμη μορφή ανάλυσης δεδομένων, και δεύτερον, γιατί παρέχουν νέες προοπτικές εξέλιξης για την ίδια την έρευνα της μηχανικής μάθησης [15], [81], [89].

Στον τομέα του άμεσου ενδιαφέροντός μας, τώρα, αποτελεί μάλλον κοινό τόπο ότι το (πολυμεσικό ή μη) περιεχόμενο είναι απαραίτητο σε μία πληθώρα σύγχρονων εφαρμογών και ότι οι αποκαλούμενες αποθήκες περιεχομένου έχουν αποκτήσει πλέον τεράστιες διαστάσεις και αυξάνονται με γοργούς ρυθμούς. Συνεπώς, καλούμαστε ως

#### Κεφάλαιο 4. Εξόρυξη δεδομένων και κατηγοριοποίηση

ερευνητές, που εν γένει το αντικείμενο της έρευνάς μας περιλαμβάνει το πολυμεσικό περιεχόμενο, να εμβαθύνουμε σε παλαιότερες και νέες, καινοτόμες τεχνικές και εργαλεία για την εξαγωγή της ενίστε χρυμμένης, αλλά πάντοτε χρήσιμης, γνώσης που βρίσκεται ενσωματωμένη μέσα στις συλλογές αυτές και με τον τρόπο αυτό να αποκαλύψουμε τις σχέσεις που συνδέουν τα διάφορα στοιχεία και τη γνώση μεταξύ τους. Τα τελευταία χρόνια έχουν γίνει σημαντικές προσπάθειες στον τομέα τόσο της εξόρυξης πολυμεσικών δεδομένων [201], [249], [80], όσο και στον τομέα της εξόρυξης κειμένου [119], [50], [147]. Το αυξανόμενο ενδιαφέρον για την εξόρυξη δεδομένων παρακινείται από ένα κοινό ερευνητικό πρόβλημα, το οποίο συνοφίζεται στο εξής ερώτημα [108]: πώς μπορεί κανείς να αποθηκεύει, να έχει πρόσβαση, να διαμορφώνει, να περιγράψει και τελικά να καταλάβει ένα πολύ μεγάλο σύνολο δεδομένων;

Ιστορικά οι διαφορετικές πτυχές της εξόρυξης δεδομένων έχουν εξεταστεί ανεξάρτητα από τις διάφορες επιστήμες. Η εξόρυξη κειμένου, για παράδειγμα, η οποία είναι επίσης γνωστή και ως ανάσυρση δεδομένων κειμένου [113] ή ανακάλυψη γνώσης [88], αναφέρεται γενικά στη διαδικασία εξαγωγής ενδιαφερόντων και ουχί τετρικμένων προτύπων ή γνώσης από μη δομημένα έγγραφα κειμένων. Μπορεί να θεωρηθεί επέκταση της κλασικής εξόρυξης δεδομένων ή ακόμα και συγγενής της διαδικασίας ανακάλυψης γνώσης από (δομημένες) βάσεις δεδομένων [86], [227]. Καθώς η έως σήμερα πιο διαδεδομένη μορφή αποθήκευσης πληροφοριών είναι το απλό κείμενο, η εξόρυξη κειμένου θεωρείται ότι διέπεται από μεγάλες δυνατότητες ανάπτυξης. Μια πρόσφατη μελέτη έδειξε ότι πάνω από το 80% των πληροφοριών των μεγάλων Ευρωπαϊκών επιχειρήσεων περιλαμβάνεται σε έγγραφα κειμένων. Η εξόρυξη κειμένων, εντούτοις, αποτελεί έναν πολύ πιο σύνθετο στόχο (σε σχέση με την εξόρυξη δεδομένων), καθώς περιλαμβάνει την εξέταση κειμένων, τα οποία είναι εγγενώς μη δομημένα και ασαφή. Η εξόρυξη κειμένων τείνει να γίνει ένας τεράστιος διεπιστημονικός τομέας, ο οποίος “ακουμπάει” έως ένα βαθμό την ανάκτηση πληροφορίας, την ανάλυση κειμένων, την εξαγωγή πληροφοριών, τη συγκέντρωση, την κατηγοριοποίηση, την απεικόνιση, τις τεχνολογίες των βάσεων δεδομένων, τη μηχανική μάθηση, την εξόρυξη δεδομένων, κ.ο.κ..

Τις τελευταίες δεκαετίες, πολλοί ερευνητές προσπάθησαν να αντιμετωπίσουν το πρόβλημα της αναγνώρισης του πολυμεσικού περιεχομένου που επιθυμεί ο χρήστης. Σε αυτό το πλαίσιο η θεματική κατηγοριοποίηση των πολυμεσικών εγγράφων αποτελεί ένα ιδιαίτερα σημαντικό και αποτελεσματικό από πλευράς κόστους βήμα προς την κατεύθυνση της αποδοτικής διαχείρισης και κατανόησης του πολυμεσικού περιεχομένου. Είναι, όμως, γεγονός ότι η εξαγωγή σημασιολογικής πληροφορίας μέσω της απόκτησης γνώσης και της αναγνώρισης υψηλού επιπέδου σημασιολογικών χαρακτηριστικών παραμένει ένα ανοιχτό και ενδιαφέρον ερευνητικό πρόβλημα. Τα χαρακτηριστικά προβλήματα που απαντώνται συνήθως στα συστήματα ανάκτησης κειμενικής πληροφορίας, όπως είναι η υπερφόρτωση και η διαφορετικότητα πληροφορίας [214], εμφανίζονται και στην περίπτωση της πολυμεσικής ανάκτησης [34]. Οι τρέχουσες προσεγγίσεις τείνουν να συνδυάζουν την πολυμεσική εξαγωγή και την ανάπτυξη της γνώσης σε ένα κοινό πλαίσιο. Η ανάπτυξη της γνώσης εντείνει την ευρωστία της διαδικασίας εξαγωγής, ενώ την ίδια στιγμή η συνεχόμενη εξαγωγή σημασιολογικής πληροφορίας αξιοποιείται για τον εμπλουτισμό της γνώσης. Προκειμένου να μειώσουμε περαιτέρω την πολυπλοκότητα της σημασιολογικής πολυμεσικής ανάλυσης, εισάγονται οι έννοιες της ενοποιημένης μοντελοποίησης και της οντολογικής αναπαράστασης της πολυμεσικής γνώσης.

Οι ερευνητικές προσπάθειες που εξετάζουν την ταξινόμηση εγγράφων κειμένου

έχουν πλέον ωριμάσει και έχουν παραγάγει ενδιαφέροντα αποτελέσματα κατά τη διάρκεια της τελευταίας δεκαετίας. Μια ολοκληρωμένη επισκόπηση αρκετών μοντέλων κατηγοριοποίησης παρουσιάζεται στη [222]. Νέα στατιστικά μοντέλα για την ταξινόμηση δομημένων εγγράφων πολυμέσων παρουσιάζονται στην [77]. Στο ίδιο πλαίσιο αρκετό ενδιαφέρον παρουσιάζει και ο τομέας της ανεπίβλεπτης συσταδοποίησης εγγράφων, όπου τα κειμενικά έγγραφα είναι συσταδοποιημένα σε ομάδες παρόμοιου περιεχομένου σύμφωνα με ένα προκαθορισμένο κριτήριο ομοιότητας, το οποίο στις περισσότερες περιπτώσεις απεικονίζεται από τον ευρέως αποδεκτό συντελεστή συνημιτόνου (cosine coefficient) του μοντέλου διανυσματικού χώρου. Αλγόριθμοι συσταδοποίησης, όπως ο αλγόριθμος πλήρους συνδέσμου, ο αλγόριθμος ενιαίου συνδέσμου ή ακόμα και οι ιεραρχικοί αλγόριθμοι συσταδοποίησης μέσου όρου, χρησιμοποιούνται κατά κόρον για τη συσταδοποίηση κειμένων στη σχετική βιβλιογραφία [251]. Η πιο αποδοτική υπολογιστικά μέθοδος είναι αυτή του μοναδικού συνδέσμου και γι' αυτό έχει χρησιμοποιηθεί εκτενώς στη βιβλιογραφία. Παρ' όλα αυτά η μέθοδος πλήρους συνδέσμου θεωρείται αποτελεσματικότερη, αν και απαιτεί ένα υψηλότερο υπολογιστικό κόστος [278]. Τέλος, θα πρέπει να αναφέρουμε ότι ο MacLeod [162] πρότεινε μέχρι και τη χρησιμοποίηση κατάλληλων μοντέλων νευρωνικών δικτύων για τη συσταδοποίηση ως μία εναλλακτική λύση στα διανυσματικά μοντέλα ομοιότητας των εγγράφων.

'Όταν δώσουμε έμφαση στο θέμα του προσδιορισμού των θεματικών κατηγοριών και, σε αντιδιαστολή με το απλό κείμενο, εστιάσουμε αποκλειστικά σε πλούσιο πολυμεσικό περιεχόμενο, παρατηρούμε ότι έχουν προταθεί και χρησιμοποιηθεί κατά καιρούς γρήγοροι αλγόριθμοι για την πλοήγηση μεγάλων ποσοτήτων πολυμεσικού περιεχομένου [69]. Η λανθάνουσα σημασιολογική ανάλυση (LSA) [147], [74], [27] χρησιμοποιεί την τεχνική της ανάλυσης ιδιαίτερων τιμών (singular value decomposition - SVD), για να αντιστοιχήσει έγγραφα και όρους από τον διανυσματικό χώρο αναπαράστασης σε ένα λανθάνοντα χώρο χαμηλότερων διαστάσεων, αποκαλύπτοντας τις σημασιολογικές σχέσεις μεταξύ των οντοτήτων που μας ενδιαφέρουν. Μια χωρίς επιβλεψη γενίκευση της LSA, η πιθανοτική-LSA (pLSA) [119], η οποία στηρίζεται σε στατιστικά θεμέλια, αναπαριστά τα έγγραφα σε ένα χώρο σημασιολογικών εννοιών και εξάγει τις έννοιες με αυτόματο τρόπο. Επιπρόσθετα, τεχνικές αναγνώρισης προτύπων και μηχανικής μάθησης έχουν εφαρμοστεί στην κατηγοριοποίηση εγγράφων, όπως π.χ. είναι ο αλγόριθμος ασαφών c-μέσων (fuzzy c-means) [25] για την περίπτωση της επιβλεπόμενης κατηγοριοποίησης πολυμεσικών εγγράφων. Τέλος, τεχνικές προβολής [220] και η συσταδοποίηση k-μέσων [211] προτείνονται για να επιταχυνθεί ο υπολογισμός των αποστάσεων της συσταδοποίησης· η επίδοσή τους εξετάζεται στο [49], αλλά θα πρέπει να σημειώσουμε ότι τα αποτελέσματα της συσταδοποίησης εγγράφων εξαρτώνται από το αρχικό σύνολο δεδομένων του πολυμεσικού περιεχομένου.

Στον τομέα, τώρα, της εξαγωγής χρήσιμων συμπερασμάτων από το πολυμεσικό περιεχόμενο, μια ιδιαίτερα δημοφιλής οικογένεια γλωσσών αναπαράστασης γνώσης, στην οποία δόθηκε ιδιαίτερη προσοχή κατά τη διάρκεια της τελευταίας δεκαετίας, είναι οι Περιγραφικές Λογικές (Description Logics - DLs) [12]. Οι DLs αποτελούν μια λογική αναδημιουργία των αποκαλούμενων βασισμένων σε πλαίσιο γλωσσών αναπαράστασης γνώσης. Ως κύριο στόχο έχουν την παροχή μιας απλής, καθιερωμένης και δηλωτικής σημασιολογίας, για να συλλάβουν την έννοια των δημοφιλέστερων χαρακτηριστικών γνωρισμάτων της δομημένης αναπαράστασης της γνώσης. Η καλά ορισμένη σημασιολογία τους, μαζί με τις προσδιορίσιμες μεθόδους εξαγωγής συμπερασμάτων, τις καθιστά εφαρμόσιμες σε ποικίλες εφαρμογές [12]. Από την άλλη, παρά την πλούσια εκφραστικότητα των κλασικών DLs, είναι μάλλον ανεπαρκείς να εξετάσουν

ασαφείς και αβέβαιες πληροφορίες που απαντώνται συνήθως σε πολλές πραγματικές εφαρμογές, όπως είναι το πολυμεσικό περιεχόμενο. Για το σκοπό αυτό έχει προταθεί μια ποικιλία DLs, οι οποίες είναι σε θέση να διαχειριστούν διάφορα είδη ανακρίβειας στην πληροφορία, όπως είναι π.χ. οι πιθανοτικές [116], [129] ή οι ασαφείς [244], [246].

Ως εκ τούτου, στο τέλος του κεφαλαίου αυτού θα παρουσιάσουμε μια υβριδική προσέγγιση που αποτελείται από δύο διακριτά μέρη, δηλαδή μια υποενότητα κατηγοριοποίησης κειμένων και μια ασαφή μηχανή εξαγωγής συμπερασμάτων. Το κομμάτι της κατηγοριοποίησης βασίζεται στη γενική ιδέα του εννοιολογικού πλαισίου, όπως αυτή ορίζεται και παρουσιάζεται στο πλαίσιο της παρούσης διατριβής, αλλά και κοινών, ευρέως χρησιμοποιούμενων τεχνικών συσταδοποίησης. Χρησιμοποιεί τις αρχές της ασαφούς άλγεβρας, προκειμένου να επιτύχει την επιθυμητή θεματική κατηγοριοποίηση των πολυμεσικών εγγράφων. Το δεύτερο τμήμα, της ασαφούς μηχανής εξαγωγής συμπερασμάτων έχει υλοποιηθεί με βάση τη γλώσσα περιγραφής *f-SHIN* [244], η οποία, όπως θα δούμε, αποτελεί την ασαφή επέκταση της εκφραστικής γλώσσας περιγραφής *SHIN* [123]. Στο πλαίσιο της εργασίας μας η έννοια της εξαγωγής συμπερασμάτων αναφέρεται στην αυτόματη παραγωγή και βελτιστοποίηση των αποτελεσμάτων της κατηγοριοποίησης κειμένων, από την άποψη της εκλέπτυνσης των υψηλού επιπέδου σημασιολογικών αποτελεσμάτων θεματικής κατηγοριοποίησης, μέσω της χρησιμοποίησης της παρεχόμενης γνώσης (η οποία μπορεί, εν γένει, να είναι γενική, δομική ή κάποια γνώση θεματικής περιοχής, κ.ο.κ..).

#### 4.1.2 Δομή κεφαλαίου

Το γενικότερο πλαίσιο, στο οποίο κινείται η δομή του κεφαλαίου, χωρίζεται σε τέσσερις βασικές θεματικές ενότητες. Η ενότητα 4.2 αναφέρεται στις γενικές αρχές, που διέπουν την ερευνητική περιοχή και εστιάζει στις τεχνικές συσταδοποίησης και κατηγοριοποίησης δεδομένων που θα χρησιμοποιηθούν στο πλαίσιο της παρούσης. Παρουσιάζεται η ερευνητική συνεισφορά στο χώρο μέσω της επέκτασης χλασικών προσεγγίσεων προς την κατεύθυνση της μείωσης των διαστάσεων των χαρακτηριστικών γνωρισμάτων και της γενικότερης βελτιστοποίησης της επίδοσης των τεχνικών συσταδοποίησης μέσω της επιπρόσθετης εκμετάλλευσης συγκεκριμένων τεχνικών κατηγοριοποίησης. Στη συνέχεια, στην ενότητα 4.3 πραγματοποιείται μια εκτενής αναφορά στην προτεινόμενη μεθοδολογία ασαφούς ταξινομικής συσταδοποίησης με βάση το εννοιολογικό πλαίσιο, η οποία βασίζεται στην ασαφή σχεσιακή άλγεβρα και αποτελεί τον κορμό της θεματικής κατηγοριοποίησης εγγράφων.

Οι βασικές αρχές της θεματικής κατηγοριοποίησης εγγράφων παρουσιάζονται στην ενότητα 4.4, η οποία περιλαμβάνει την προαναφερθείσα υβριδική προσέγγιση χρήσης DLs στη διαδικασία. Αναλυτικά πειραματικά αποτελέσματα από την εφαρμογή των προτεινόμενων σε συνθετικά και πραγματικά σύνολα δεδομένων παρατίθενται στην ενότητα 4.5. Το κεφάλαιο αυτό της διατριβής παρουσιάζει ιδιαίτερο ενδιαφέρον, μια και η θεματική κατηγοριοποίηση εγγράφων αποτελεί μία από τις πολλά υποσχόμενες περιοχές εφαρμογής τεχνικών ιεραρχικής συσταδοποίησης για τη διαχείριση πολυμεσικού περιεχομένου. Θα πρέπει, βέβαια, να σημειώσουμε ότι η εξαγωγή θεματικών κατηγοριών βρίσκει εφαρμογή τόσο σε πολυμεσικά, όσο και σε απλά έγγραφα κειμένου. Συνακόλουθα, η προσέγγιση που ακολουθείται είναι η ίδια και για τις δύο περιπτώσεις, καθώς η δεύτερη θεωρείται απλή υποπερίπτωση της πρώτης. Ο γενικότερος στόχος είναι η αυτόματη ανίχνευση των κατηγοριών - σε εννοιολογικό/σημασιολογικό επίπεδο - που σχετίζονται με ένα συγκεκριμένο έγγραφο και η ανάλογη κατηγοριοποίηση του

τελευταίου.

## 4.2 Γενικές Αρχές

Καθώς οι τεχνολογίες αποθήκευσης, καταγραφής, ψηφιοποίησης, αλλά και μετάδοσης δεδομένων έχουν προχωρήσει σημαντικά τα τελευταία χρόνια, όλο και περισσότερα δεδομένα είναι διαθέσιμα σε ψηφιακή μορφή. Πλέον, το βασικό πρόβλημα δεν είναι τόσο η απόκτηση των δεδομένων, όσο η εξαγωγή χρήσιμων συμπερασμάτων από αυτά και η μετέπειτα εκμετάλλευση αυτών των συμπερασμάτων. Αυτά είναι τα βασικά αντικείμενα της σύγχρονης μηχανικής μάθησης. Βασική θέση στη μηχανική μάθηση κατέχουν η συσταδοποίηση (clustering) δεδομένων, από την πλευρά της ανάλυσης των δεδομένων, και η κατηγοριοποίηση (classification) δεδομένων, από την πλευρά τόσο της ανάλυσης, όσο και της αξιοποίησης της εξαχθείσας πληροφορίας.

Η συσταδοποίηση των δεδομένων είναι ένα πρόβλημα που συσχετίζεται με πολυάριθμους ερευνητικούς και εφαρμοσμένους τομείς [118]. Αν και οι ερευνητές στον τομέα της εξόρυξης δεδομένων έχουν εργαστεί σε αυτήν την κατεύθυνση για μεγάλο χρονικό διάστημα και είναι αλήθεια ότι υπάρχουν πολυάριθμα σχετικά κείμενα στη βιβλιογραφία, το πρόβλημα αυτό θεωρείται ακόμα ένα ανοικτό ερευνητικό ζήτημα. Ιδιαίτερη δυσκολία εμφανίζεται στις περιπτώσεις που τα δεδομένα χαρακτηρίζονται από πολυάριθμα μετρήσιμα χαρακτηριστικά, κάτι που αναφέρεται συχνά ως το πρόβλημα της “κατάρας των υψηλών διαστάσεων”.

Από την άλλη πλευρά, η έρευνα στον τομέα της κατηγοριοποίησης δεδομένων εστιάζει στη χρήση ήδη χαρακτηρισμένων δεδομένων, τα οποία ονομάζονται συνήθως δεδομένα εκπαίδευσης (training data), για την αυτόματη παραγωγή συστημάτων που θα είναι σε θέση να κατηγοριοποιήσουν (ή με απλά λόγια να χαρακτηρίσουν) μελλοντικά δεδομένα. Αυτή η κατηγοριοποίηση στηρίζεται στην ομοιότητα των εισερχόμενων δεδομένων σε σύγκριση με τα δεδομένα εκπαίδευσης. Ο κύριος στόχος είναι να παραχθούν αυτόματα συστήματα που είναι σε θέση να κατηγοριοποιήσουν σωστά τα μελλοντικά εισερχόμενα δεδομένα [118].

Αν και μια πρώτη ματιά οι στόχοι της κατηγοριοποίησης και της συσταδοποίησης μοιάζουν κατά πολύ μεταξύ τους, εντούτοις υπάρχει μια σημαντική διαφορά: ενώ στην κατηγοριοποίηση το σημαντικότερο σημείο είναι η διάκριση μεταξύ των κατηγοριών, δηλαδή η ανίχνευση των ορίων κάθε κατηγορίας, κατά τη συσταδοποίηση το κύριο σημείο είναι ο προσδιορισμός των χαρακτηριστικών συστάδων. Το τελευταίο αντιμετωπίζεται συνήθως μέσω της επιλογής των αντιπροσώπων συστάδων ή αλλιώς των κεντρικών στοιχείων *centroids* των συστάδων.

Στην τυπική περίπτωση, προκειμένου να επιτευχθεί η αυτόματη δημιουργία τέτοιων συστημάτων κατηγοριοποίησης, θα πρέπει πρώτα να ανιχνευθούν τα πρότυπα που υποβόσκουν στα δεδομένα, σε αντίθεση με το απλά να διαμεριστούν τα δείγματα των δεδομένων με βάση τις διαθέσιμες ετικέτες [258], και κατόπιν να μελετηθεί ο τρόπος με τον οποίο αυτά τα πρότυπα σχετίζονται με κατηγορίες που έχουν κάποιο νόημα. Αποδοτικές λύσεις έχουν προταθεί στη βιβλιογραφία και για τους δύο στόχους, συνήθως όμως για την περίπτωση κατά την οποία καθορίζεται ένα μοναδικό μέτρο ομοιότητας ή ανομοιότητας μεταξύ των δεδομένων εισόδου [251]. Ακόμα και αν χρησιμοποιηθούν αυτοεκπαίδευσημένα συστήματα, όπως π.χ. είναι τα νευρωνικά δίκτυα ανακατανομής πόρων, τα οποία είναι σε θέση να προσαρμοστούν δυναμικά στα δεδομένα εκπαίδευσης, καλά αποτελέσματα μπορούν μόνο να επιτευχθούν, όταν τα πρότυπα είναι γνωστά από πριν, έτσι ώστε να μπορούν να χρησιμοποιηθούν για

κατάλληλη αρχικοποίηση [112].

Στην περίπτωση που τα δεδομένα χαρακτηρίζονται από πολλαπλά, ανεξάρτητα μεταξύ τους χαρακτηριστικά γνωρίσματα, και με τον τρόπο αυτόν μπορούν να ορισθούν και να έχουν νόημα περισσότερα του ενός μέτρα ομοιότητας ή ανομοιότητας, τότε και οι δύο παραπάνω στόχοι γίνονται δυσκολότερο να επιτευχθούν. Μια διαδεδομένη προσέγγιση στο πρόβλημα είναι η μείωση των διαστάσεων εισαγωγής (dimensionality reduction), κάτι το οποίο μπορεί να πραγματοποιηθεί αγνοώντας μερικά από τα διαθέσιμα χαρακτηριστικά γνωρίσματα (feature selection) [142] ή με την εφαρμογή κάποιου χωρικού μετασχηματισμού. Στην περίπτωση, όμως, που τα χαρακτηριστικά γνωρίσματα εισαγωγής είναι ανεξάρτητα, ή όταν δεν είναι γνωστή εκ των προτέρων η σχέση μεταξύ τους, όπως συχνά συμβαίνει με τα δεδομένα από την πραγματική ζωή, δεν μπορεί να συντελεσθεί μια μείωση των χωρικών διαστάσεων χωρίς την απώλεια κάποιας πληροφορίας. Επομένως, εάν η σχέση μεταξύ των χαρακτηριστικών δεν είναι γνωστή από πριν και ο στόχος είναι να ανιχνευθούν τα πρότυπα που υπάρχουν στα δεδομένα, η μείωση των διαστάσεων δεν είναι δυνατή. Επιπλέον, οι διαφορές στην κλίμακα μέτρησης μεταξύ των διαφορετικών χαρακτηριστικών δυσχεραίνει συνήθως την όλη διαδικασία.

#### 4.2.1 Συσταδοποίηση δεδομένων

Μία από τις βασικότερες προσεγγίσεις στο χώρο της μηχανικής μάθησης και της εξόρυξης δεδομένων είναι, λοιπόν, η συσταδοποίηση, η οποία μπορεί να θεωρηθεί ως ένα από τα σημαντικότερα προβλήματα ανεπίβλεπτης μάθησης. Βέβαια, ο όρος “συσταδοποίηση δεδομένων” αναφέρεται, εν γένει, σε μία ευρέως διαδεδομένη τεχνική που χρησιμοποιείται για τη στατιστική ανάλυση δεδομένων σε πολλά πεδία των σύγχρονων επιστημών, όπως η μηχανική μάθηση (machine learning), η εξόρυξη δεδομένων (data mining), η αναγνώριση προτύπων (pattern recognition) και η επεξεργασία εικόνας (image processing). Όπως και κάθε άλλο πρόβλημα αυτού του είδους, η συσταδοποίηση εξετάζει την εύρεση μιας δομής σε μια συλλογή από μη χαρακτηρισμένα (unlabeled) στοιχεία. Ένας μάλλον “χαλαρός” ορισμός της συσταδοποίησης θα μπορούσε να την ορίζει ως “τη διαδικασία της οργάνωσης των αντικειμένων σε ομάδες, των οποίων τα μέλη είναι παρόμοια κατά κάποιο τρόπο”. Μια συστάδα είναι επομένως μια συλλογή αντικειμένων που είναι “παρόμοια” μεταξύ τους και είναι “ανόμοια” ως προς τα αντικείμενα που ανήκουν σε άλλες συστάδες. Με άλλα λόγια, η συσταδοποίηση δεδομένων μπορεί να θεωρηθεί ως η κατηγοριοποίηση παρόμοιων αντικειμένων σε διαφορετικές ομάδες ή η διαμέριση ενός συνόλου δεδομένων σε υποσύνολα (συστάδες - clusters), έτσι ώστε τα δεδομένα κάθε συστάδας να χαρακτηρίζονται ιδανικά από κάποια κοινά γνωρίσματα. Είναι δε ιδιαίτερα σύνηθες να χρησιμοποιείται η εγγύτητα των δεδομένων, σύμφωνα με κάποια εκ των προτέρων ορισμένη μετρική απόστασης, ως το χαρακτηριστικό γνώρισμά τους.

Στην ενότητα αυτή καταβάλλεται μια προσπάθεια εισαγωγής του αναγνώστη στις κυριότερες τεχνικές συσταδοποίησης, με έμφαση στις τεχνικές iεραρχικής συσταδοποίησης (hierarchical clustering), μια και οι τελευταίες είναι κατάλληλες για ταυτόχρονη χρήση σε ένα περιβάλλον πολυμεσικής πληροφορίας αλλά και διαχείρισης γνώσης. Οι βασικές πηγές μας από τη βιβλιογραφία είναι οι [139] και [177]. Η ουσία της συσταδοποίησης δεδομένων συνοψίζεται στον προσδιορισμό ομοιογενών ομάδων αντικειμένων με βάση τις τιμές των ιδιοτήτων τους. Πρόκειται για ένα πρόβλημα που συσχετίζεται με διάφορους επιστημονικούς και εφαρμοσμένους τομείς και έχει κατά

χόρον χρησιμοποιηθεί τόσο στις επιστήμες γενικά, όσο και στον τομέα της εξόρυξης δεδομένων ειδικότερα. Οι εφαρμογές της τεχνικής κυμαίνονται από την τεχνητή νοημοσύνη και την αναγνώριση προτύπων ως τις βάσεις δεδομένων και τη στατιστική [118]. Είναι, δε, κοινός τόπος ότι υπάρχουν διαφορετικά είδη αλγορίθμων συσταδοποίησης δεδομένων ανάλογα με τους διαφορετικούς τύπους εφαρμογών και μια ευρέως διαδεδομένη διάκριση που χρησιμοποιείται είναι αυτή μεταξύ αλγορίθμων *ιεραρχικής συσταδοποίησης* (hierarchical clustering) και *καταμερίσμου* (partitioning). Παρ' όλα, δε, τα πολυάριθμα άρθρα που υπάρχουν στη βιβλιογραφία η συσταδοποίηση δεδομένων θεωρείται ακόμα ένα ανοικτό ερευνητικό ζήτημα/πρόβλημα, κυρίως επειδή είναι δύσκολο να αντιμετωπιστεί αποδοτικά στις περιπτώσεις, όπου τα δεδομένα χαρακτηρίζονται από πολυάριθμα μετρήσιμα χαρακτηριστικά γνωρίσματα. Το τελευταίο πρόβλημα αναφέρεται συχνά ως “η κατάρα των διαστάσεων” (curse of dimensionality) και θα παρουσιαστεί αναλυτικά στην επόμενη υποενότητα 4.2.2.

Οι τεχνικές συσταδοποίησης έχουν μελετηθεί ευρέως τα τελευταία χρόνια σε πολλούς ενεργούς ερευνητικούς τομείς. Έρευνες όπως οι [26], [115], [130] και [19] συμφωνούν στην διάκρισή τους σε τέσσερις κύριες κατηγορίες: (i) *ιεραρχική συσταδοποίηση*, (ii) *συσταδοποίηση καταμερίσεων*, (iii) *συσταδοποίηση με βάση την πυκνότητα* (density-based clustering) και (iv) *συσταδοποίηση με βάση το πλέγμα* (grid-based clustering). Η πρώτη στηρίζεται στην αναπαράσταση των δεδομένων σε μία δενδρική δομή, η οποία στη γενική περίπτωση κατασκευάζεται με βάση την ομοιότητα μεταξύ κάθε ζεύγους δεδομένων. Στη συνέχεια, αποκόπτοντας το δένδρο σε διάφορα επίπεδα, είμαστε σε θέση να λάβουμε το ζητούμενο αριθμό και το ζητούμενο μέγεθος των συστάδων Η δεύτερη κατηγορία προσπαθεί να δημιουργήσει μια διαμέριση των δεδομένων με τη μεγιστοποίηση της απόστασης μεταξύ των διαφορετικών συστάδων και την ταυτόχρονη ελαχιστοποίηση της απόστασης των στοιχείων στο εσωτερικό κάθε συστάδας. Οι πιο δημοφιλείς αλγόριθμοι αυτής της κατηγορίας είναι ο αλγόριθμος *k*-μέσων και ο BSAS (Basic Sequential Algorithmic Scheme). Οι βασισμένες στην πυκνότητα προσεγγίσεις συγκεντρώνουν τα δεδομένα σύμφωνα με κάποια χρήσιμα κριτήρια πυκνότητας. Κατά συνέπεια, οι συστάδες που δημιουργούνται είναι περιοχές υψηλής πυκνότητας που χωρίζονται από χαμηλής πυκνότητας περιοχές. Σε αντίθεση με τις δύο προηγούμενες περιπτώσεις, το σύνολο των συστάδων εδώ δεν αποτελεί συνήθως μια διαμέριση των δεδομένων, μιας και οι περιοχές χαμηλής πυκνότητας δεν συμπεριλαμβάνονται συνήθως στο τελικό σύνολο συστάδων. Τέλος, η βασισμένη στο πλέγμα προσέγγιση ποσοτικοποιεί το χώρο των δεδομένων έτσι ώστε αυτός να αναπαριστά τα δεδομένα σε ένα πλέγμα. Κατόπιν, η συσταδοποίηση υλοποιείται χρησιμοποιώντας όχι απευθείας τα ίδια τα δεδομένα, αλλά τα συνδεδεμένα κύτταρα (cells) του πλέγματος. Με τον τρόπο αυτό μπορούν να επιτευχθούν διαφορετικά μεγέθη και αριθμοί συστάδων, αλλάζοντας κάθε φορά την κοκκοποίηση (granularity) του πλέγματος.

Στην πραγματικότητα, οι παραπάνω τεχνικές είναι εντελώς διαφορετικές μεταξύ τους και έχουν προταθεί για να ικανοποιήσουν διαφορετικούς ερευνητικούς στόχους. Π.χ., ενώ οι καταμερίσεις λαμβάνουν υπόψη τους όλα τα δεδομένα (συμπεριλαμβανομένου και του πιθανού θορύβου σε αυτά), οι βασισμένοι στην πυκνότητα αλγόριθμοι αγνοούν τις χαμηλής πυκνότητας συστάδες/περιοχές. Όσον αφορά στην ανάκτηση πληροφορίας, οι αλγόριθμοι συσταδοποίησης δεδομένων που χρησιμοποιούνται κατά κόρον ανήκουν, εν γένει, είτε στους ιεραρχικούς αλγόριθμους, είτε στις καταμερίσεις. Βασική διαφορά των δύο προσεγγίσεων είναι ότι οι ιεραρχικοί αλγόριθμοι βρίσκουν τις διαδοχικές ομάδες χρησιμοποιώντας σε κάθε βήμα τις συστάδες που έχουν

#### Κεφάλαιο 4. Εξόρυξη δεδομένων και κατηγοριοποίηση

δημιουργηθεί στο αμέσως προηγούμενο βήμα, ενώ οι αλγόριθμοι καταμερισμού καθορίζουν όλες τις συστάδες σε ένα βήμα. Οι iεραρχικοί αλγόριθμοι διαχωρίζονται περαιτέρω σε συσσωρευτικούς (agglomerative) (προσέγγιση: “από κάτω προς τα επάνω”) και διαχωριστικούς (divisive) (προσέγγιση: “από επάνω προς τα κάτω”). Οι συσσωρευτικοί αλγόριθμοι αρχίζουν θεωρώντας κάθε στοιχείο ως μία ξεχωριστή συστάδα. Τις συστάδες που δημιουργούνται τις συγχωνεύουν σε διαδοχικά μεγαλύτερες συστάδες σε κάθε βήμα. Αντίθετα, οι διαχωριστικοί αλγόριθμοι αρχίζουν θεωρώντας ολόκληρο το σύνολο των δεδομένων ως μία συστάδα και προχωρούν διαιρώντας το σε κάθε βήμα σε διαδοχικά μικρότερες συστάδες.

Αν και οι iεραρχικές μέθοδοι συσταδοποίησης είναι πιο ευέλικτες από τις αντίστοιχες καταμερίσεις δεδομένων, μια και δε χρειάζεται να γνωρίζουν εκ των προτέρων τον αριθμό των συστάδων ως είσοδο, είναι λιγότερο εύρωστες, όταν ληφθούν υπόψη κάποιες επιπλέον παράμετροι της διαδικασίας. Πιο συγκεκριμένα, τα λάθη που τυχόν πραγματοποιούνται κατά τα αρχικά βήματα ενός iεραρχικού αλγορίθμου τείνουν να διαδοθούν σε όλο το εύρος της διαδικασίας ώστε και την τελική έξοδό του. Αυτό είναι ένα σημαντικό πρόβλημα, καθώς μπορεί να οδηγήσει σε παραπλανητικά και ακατάλληλα συμπεράσματα για τα αντίστοιχα σύνολα δεδομένων. Επιπλέον, η αρκετά υψηλότερη υπολογιστική πολυπλοκότητα που, στην τυπική περίπτωση, διέπει τους iεραρχικούς αλγορίθμους, τους καθιστά μη εφαρμόσιμους στις περισσότερες καταστάσεις της πραγματικής ζωής λόγω του μεγάλου μεγέθους των αντίστοιχων συνόλων δεδομένων.

Ο αλγόριθμος που προτείνεται στο πλαίσιο αυτής της διδακτορικής διατριβής αποτελεί μια επέκταση της συσσωρευτικής συσταδοποίησης προς την παραπάνω κατεύθυνση και είναι βασισμένος σε μια ευέλικτη επιλογή των χαρακτηριστικών γνωρισμάτων που λαμβάνονται υπόψη κατά τη σύγκριση των δεδομένων. Τα αποτελέσματα της αρχικής συσταδοποίησης, που εκτελέσται σε ένα μικρό μόνο μέρος του αρχικού συνόλου δεδομένων, εκλεπτύνονται στη συνέχεια μέσω ενός βήματος κατηγοριοποίησης· αυτό το βήμα, αν και χωρίς επίβλεψη (unsupervised), είναι βασισμένο στις αρχές της κατηγοριοποίησης με βάση τον k-κοντινότερο γείτονα (kNN) και εφαρμόζεται σε ολόκληρο το σύνολο των δεδομένων. Κατ' αυτό τον τρόπο υπερνικάμε δύο σημαντικά μειονεκτήματα που χαρακτηρίζουν εν γένει τη συσσωρευτική συσταδοποίηση: (i) αυτό της αρχικής διάδοσης λάθους και (ii) αυτό που σχετίζεται με ζητήματα υπολογιστικής πολυπλοκότητας. Τέλος, αυτό το σημαντικό βήμα *a posteriori* κατηγοριοποίησης συμβάλλει επίσης στην πειραματική αξιολόγηση της αποδοτικότητας της προτεινόμενης μεθόδου συσταδοποίησης, γεγονός το οπόιο από μόνο του αποτέλεσε ισχυρό κίνητρο για την έρευνά μας.

##### 4.2.1.1 Συσσωρευτική iεραρχική συσταδοποίηση

Όπως έχει ήδη αναφερθεί, οι πειρισσότερες μέθοδοι συσταδοποίησης ανήκουν σε μία εκ των δύο γενικών κατηγοριών: καταμερισμού και iεραρχικές. Οι μέθοδοι καταμερισμού δημιουργούν μία σαφή ή ασαφή συσταδοποίηση ενός δεδομένου συνόλου στοιχείων, αλλά απαιτούν τον αριθμό των συστάδων ως είσοδο. Όταν το πλήθος των προτύπων που υπάρχουν σε ένα σύνολο δεδομένων δεν είναι γνωστό εκ των προτέρων, οι μέθοδοι αυτές δεν είναι εφαρμόσιμες και πρέπει να χρησιμοποιηθεί ένας iεραρχικός αλγόριθμος συσταδοποίησης. Οι iεραρχικές μέθοδοι διαιρούνται περαιτέρω στις συσσωρευτικές και τις διαχωριστικές. Από αυτές, οι πρώτες έχουν μελετηθεί εκτενώς και έχουν εφαρμοστεί σε ευρύτερα επιστημονικά πεδία, ενώ συνάμα θεωρούνται πιο αποδοτικές και εύρωστες. Κατά συνέπεια, μια τέτοια συσσωρευτική μέθοδος υιοθε-

τείται και στην παρούσα εργασία, ενώ η απόλυτα γενική δομή της είναι η ακόλουθη [177]:

1. Μετατρέπουμε κάθε στοιχείο των δεδομένων εισόδου σε μονοσύνολο (singleton), δηλ. σε μία δική του συστάδα.
2. Για κάθε ζεύγος συστάδων  $c_1, c_2$  υπολογίζουμε ένα δείκτη συμβατότητας  $CI(c_1, c_2)$ . Ο  $CI$  αναφέρεται επίσης και ως μετρική ομοιότητας ή ανομοιότητας των συστάδων ή μετρική της απόστασης των συστάδων.
3. Συγχωνεύουμε το ζευγάρι των συστάδων που έχουν τον καλύτερο δείκτη  $CI$ . Ανάλογα με το αν ο  $CI$  είναι μια μετρική ομοιότητας ή ανομοιότητας, ο καλύτερος δείκτης θα μπορούσε να επιλεγεί χρησιμοποιώντας το  $\max$  ή το  $\min$ , αντίστοιχα.
4. Συνεχίζουμε στο βήμα 2, έως ότου ικανοποιηθεί ένα κριτήριο τερματισμού. Το κριτήριο τερματισμού που χρησιμοποιείται συνήθως είναι ο καθορισμός ενός κατωφλίου για την τιμή του καλύτερου δείκτη συμβατότητας.

Τα δύο βασικά σημεία στην ιεραρχική συσταδοποίηση είναι ο προσδιορισμός σε κάθε βήμα των συστάδων προς συγχώνευση, δηλαδή ο καθορισμός μίας μετρικής με νόημα για το  $CI$  και ο προσδιορισμός του βέλτιστου βήματος τερματισμού, δηλαδή ο καθορισμός ενός κατάλληλου κριτηρίου τερματισμού. Σημαντικά μειονεκτήματα των συσσωρευτικών μεθόδων είναι η υψηλή πολυπλοκότητά τους και η ευαισθησία τους στα λάθη κατά τα αρχικά βήματα, τα οποία διαδίδονται μέχρι την τελική έξοδο του αλγορίθμου.

Στη συνέχεια, θα δώσουμε ένα παράδειγμα εφαρμογής των παραπάνω. Έστω ότι πρόκειται να συσταδοποιηθούν τα δεδομένα του Σχήματος 4.1 και ότι χρησιμοποιείται η Ευκλείδια μετρική απόστασης [281], η οποία στη γενική περίπτωση μεταξύ σημείων  $P = (p_1, p_2, \dots, p_n)$  και  $Q = (q_1, q_2, \dots, q_n)$  στον Ευκλείδιο  $n$ -χώρο, ορίζεται ως εξής:

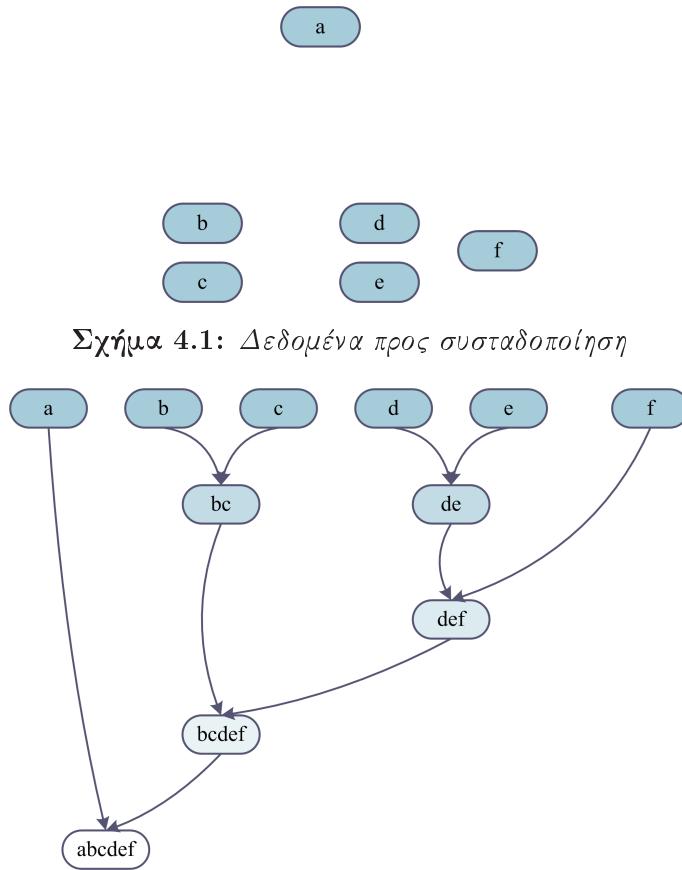
$$\sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_n - q_n)^2} = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (4.1)$$

και εν προκειμένω, ως:

$$d : \mathbb{R}^2 \longrightarrow \mathbb{R} \quad \forall x = (x_1, x_2), \quad y = (y_1, y_2) \quad d(x, y) = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2} \quad (4.2)$$

Τότε το δενδροδιάγραμμα της ιεραρχικής συσταδοποίησης θα είχε τη μορφή που παρουσιάζεται στο Σχήμα 4.2.

Αυτή η μέθοδος χτίζει την ιεραρχία από τα μεμονωμένα στοιχεία, συγχωνεύοντας σε κάθε βήμα τις κατάλληλες συστάδες. Θεωρώντας ότι έχουμε τα έξι στοιχεία που απεικονίζονται στο Σχήμα 4.1:  $\{a\}, \{b\}, \{c\}, \{d\}, \{e\}$  και  $\{f\}$ , το πρώτο βήμα είναι να αποφασιστούν ποια στοιχεία θα συγχωνευτούν σε μια συστάδα. Συνήθως, θέλουμε να πάρουμε τα δύο πιο κοντινά στοιχεία, επομένως πρέπει να καθορίσουμε μια απόσταση  $d(element1, element2)$  μεταξύ των στοιχείων. Υποθέτοντας ότι έχουμε συγχωνεύσει τα δύο πιο κοντινά στοιχεία  $b$  και  $c$ , θα έχουμε τώρα τις ακόλουθες ομάδες:  $\{a\}, \{b, c\}, \{d\}, \{e\}$  και  $\{f\}$ , τις οποίες θέλουμε να τις συγχωνεύσουμε περαιτέρω. Για να πραγματοποιηθεί το τελευταίο, θα πρέπει να πάρουμε την απόσταση



Σχήμα 4.1: Δεδομένα προς συσταδοποίηση

Σχήμα 4.2: Παράδειγμα ιεραρχικής συσσωρευτικής συσταδοποίησης δεδομένων - απεικόνιση όλων των βημάτων

μεταξύ  $\{a\}$  και  $\{b, c\}$ , και επομένως να καθορίσουμε την απόσταση μεταξύ δύο συστάδων. Συνήθως, η απόσταση μεταξύ δύο συστάδων, έστω  $\mathcal{A}$  και  $\mathcal{B}$ , είναι μια από τις ακόλουθες:

- Η μέγιστη απόσταση μεταξύ των στοιχείων κάθε συστάδας, η οποία ονομάζεται επίσης και συσταδοποίηση πλήρους συνδέσμου:

$$\max\{d(x, y) : x \in \mathcal{A}, y \in \mathcal{B}\} \quad (4.3)$$

- Η ελάχιστη απόσταση μεταξύ των στοιχείων κάθε συστάδας, η οποία ονομάζεται επίσης και συσταδοποίηση μοναδιαίου συνδέσμου:

$$\min\{d(x, y) : x \in \mathcal{A}, y \in \mathcal{B}\} \quad (4.4)$$

- Η μέση απόσταση μεταξύ των στοιχείων κάθε συστάδας, η οποία ονομάζεται επίσης και συσταδοποίηση μέσου συνδέσμου:

$$\frac{1}{\text{card}(\mathcal{A})\text{card}(\mathcal{B})} \sum_{x \in \mathcal{A}} \sum_{y \in \mathcal{B}} d(x, y) \quad (4.5)$$

- Το άθροισμα όλων των ενδο-ομαδιαίων διαφορών (*intra-cluster variance*).
- Η αύξηση της διαφοράς για τη συστάδα που συγχωνεύεται, η οποία ονομάζεται και κριτήριο του *Ward* (*Ward's criterion*).

- Η πιθανότητα οι υποψήφιες συστάδες να προέρχονται από την ίδια συνάρτηση κατανομής (*V-linkage*).

Κάθε συσσώρευση συμβαίνει σε μια μεγαλύτερη απόσταση μεταξύ των συστάδων συγκριτικά με την προηγούμενη συσσώρευση, και η διαδικασία μπορεί να λάβει τέλος (χριτήριο τερματισμού), είτε όταν οι συστάδες βρίσκονται εξαιρετικά μακριά, ώστε να συγχωνευθούν (χριτήριο απόστασης), είτε όταν υπάρχει ένας αρκετά μικρός αριθμός συστάδων (χριτήριο πλήθους).

#### 4.2.2 Μείωση των διαστάσεων των χαρακτηριστικών γνωρισμάτων

Στη γενική περίπτωση, η μείωση των διαστάσεων των μετρήσιμων χαρακτηριστικών γνωρισμάτων μπορεί να χωρισθεί σε δύο κατηγορίες: (i) στην επιλογή των χαρακτηριστικών γνωρισμάτων (feature selection) και (ii) στην εξαγωγή των χαρακτηριστικών γνωρισμάτων (feature extraction). Το φαινόμενο της μείωσης των διαστάσεων των χαρακτηριστικών γνωρισμάτων συναντάται κυρίως στο πεδίο της στατιστικής ανάλυσης, αλλά και σε άλλες θεματικές περιοχές, όπως π.χ. η φυσική.

Η επιλογή χαρακτηριστικών γνωρισμάτων βασίζεται σε προσεγγίσεις που προσπαθούν να βρουν ένα υποσύνολο των αρχικών χαρακτηριστικών γνωρισμάτων. Δύο διαδεδομένες στρατηγικές είναι αυτές του φιλτραρίσματος (π.χ. κέρδος πληροφορίας - information gain) και της προσέγγισης περιτυλίγματος (π.χ. ένας γενετικός αλγόριθμος). Από την άλλη, η εξαγωγή χαρακτηριστικών γνωρισμάτων εφαρμόζει μια αντιστοίχιση του πολυδιάστατου χώρου σε ένα χώρο λιγότερων διαστάσεων. Αυτό σημαίνει ότι ο αρχικός χώρος των χαρακτηριστικών γνωρισμάτων μετασχηματίζεται με την εφαρμογή π.χ. ενός γραμμικού μετασχηματισμού μέσω μιας ανάλυσης κύριων συνιστωσών (principal components analysis).

Σε αυτή την ενότητα παρουσιάζεται μία χωρίς επίβλεψη (unsupervised) επέκταση της ιεραρχικής συσταδοποίησης, σύμφωνα με την έννοια της επιλογής χαρακτηριστικών γνωρισμάτων. Το φαινόμενο της μείωσης της διαστατικότητας των χαρακτηριστικών γνωρισμάτων συναντάται εύλογα και στο πεδίο των τεχνικών συσταδοποίησης, μια και η τελευταία βασίζεται ακριβώς σε μετρήσιμα χαρακτηριστικά γνωρισμάτων. Όπως ίσως έχει ήδη διαφανεί από τα προηγούμενα, το σημείο αναφοράς του παραπάνω γενικού αλγορίθμου συσταδοποίησης (βλ. ενότητα 4.2.1.1) είναι η δυνατότητα να καθοριστεί μια μοναδική απόσταση μεταξύ οποιουδήποτε ζευγαριού συστάδων. Επομένως, όταν ο χώρος εισόδου διέπεται από περισσότερες της μίας διαστάσεις, χρησιμοποιείται μια αθροιστική συνάρτηση απόστασης, όπως η Ευκλείδια απόσταση [281]. Κάτι τέτοιο, βέβαια, δεν είναι πάντοτε κάτι το οποίο έχει νόημα και υπάρχουν περιπτώσεις, όπου πρέπει να εκτελεσθεί μια επιλογή των σημαντικών χαρακτηριστικών γνωρισμάτων πριν από τον υπολογισμό μιας τέτοιας απόστασης [267]. Με άλλα λόγια, για ένα δεδομένο σύνολο στοιχείων μπορεί να μην είναι δυνατό να επιλεγεί μια ενιαία μετρική απόστασης, η οποία θα ισχύει σε όλες τις περιπτώσεις.

Επιπλέον, ένα μετρήσιμο χαρακτηριστικό γνώρισμα μπορεί να είναι σημαντικότερο από άλλα, ενώ όλα τα χαρακτηριστικά γνωρίσματα είναι χρήσιμα, καθένα στο δικό του βαθμό. Στο πλαίσιο της έρευνάς μας αντιμετωπίζουμε τη στάθμιση των χαρακτηριστικών γνωρισμάτων με βάση την ακόλουθη αρχή: ενώ αναμένουμε τα στοιχεία ενός συνόλου να παρουσιάζουν τυχαίες αποστάσεις μεταξύ τους, σύμφωνα με τα περισσότερα χαρακτηριστικά τους γνωρίσματα, αναμένουμε

επίσης να έχουμε μικρές αποστάσεις, σύμφωνα με τα χαρακτηριστικά γνωρίσματα που τα συσχετίζουν. Στηριζόμαστε σε αυτήν τη διαφορά στην κατανομή των τιμών απόστασης, προκειμένου να προσδιοριστεί το εννοιολογικό πλαίσιο ενός συνόλου στοιχείων, δηλ. ο υποχώρος στον οποίο καθορίζεται καλύτερα το ίδιο το σύνολο.

Θέλοντας να ακολουθήσουμε έναν πιο αυστηρό φορμαλισμό, έστω ότι  $c_1$  και  $c_2$  είναι δύο συστάδες στοιχείων. Έστω, επίσης, ότι  $d_i$ , όπου  $i \in \mathbb{N}_F$ , είναι η μετρική που συγχρίνει το  $i$ -οστό χαρακτηριστικό γνώρισμα και το  $F$  είναι η γενική αρίθμηση των χαρακτηριστικών γνωρίσματων (δηλ. με άλλα λόγια η διάσταση του χώρου εισόδου). Ένα μέτρο της απόστασης μεταξύ των δύο συστάδων, κατά την εξέταση ακριβώς του  $i$ -οστού χαρακτηριστικού γνωρίσματος, δίνεται από την:

$$f_i(c_1, c_2) = \sqrt[k]{\frac{\sum_{a \in c_1, b \in c_2} [r_i(a_i, b_i)]^k}{|c_1| |c_2|}} \quad (4.6)$$

όπου ο δείκτης  $i$  δείχνει το  $i$ -οστό χαρακτηριστικό γνώρισμα ενός στοιχείου,  $|c|$  είναι η πληθυκότητα των στοιχείων συνόλου της συστάδας  $c$  και  $k \in \mathbb{R}$  είναι μια σταθερά (με χαρακτηριστική τιμή ίση με 2). Η γενική απόσταση μεταξύ  $c_1$  και  $c_2$  υπολογίζεται ως:

$$d(c_1, c_2) = \sum_{i=1}^F [x_i(c_1, c_2)]^\lambda f_i(c_1, c_2) \quad (4.7)$$

όπου το  $x_i$  είναι ο βαθμός στον οποίο το  $i$  και επομένως και το  $f_i$ , συμπεριλαμβάνεται στην ευέλικτη επιλογή των χαρακτηριστικών γνωρίσματων,  $i \in N_F$  και  $\lambda \in \mathbb{R}$  είναι μια σταθερά. Η χαρακτηριστική τιμή που χρησιμοποιείται για το  $\lambda$  είναι το 2. Με βάση την αρχή που παρουσιάστηκε παραπάνω, οι τιμές του διανύσματος  $x_i$  επιλέγονται μέσω της ελαχιστοποίησης της απόστασης  $d$  [270]. Τα χαρακτηριστικά γνωρίσματα που σχετίζουν τα  $c_1$  και  $c_2$  είναι “πιθανότατα” αυτά που παράγουν τελικά τις μικρότερες αποστάσεις  $f_i$ .

#### 4.2.3 Βελτιστοποίηση τεχνικών συσταδοποίησης μέσω κατηγοριοποίησης

Η συσταδοποίηση δεδομένων αποτελεί από μόνη της ένα δύσκολο πρόβλημα, το οποίο σχετίζεται με ποικίλα επιστημονικά πεδία και εφαρμογές. Η πρόκληση για τον ερευνητή γίνεται μεγαλύτερη, καθώς αυξάνουν οι χωρικές διαστάσεις εισαγωγής (input space dimensions) των χαρακτηριστικών γνωρίσματων και οι κλίμακες των τελευταίων αρχίζουν και ποικίλουν. Στο πλαίσιο της ενότητας αυτής παρουσιάζεται, όπως και προηγουμένως, μία χωρίς επίβλεψη (unsupervised) επέκταση της ιεραρχικής συσταδοποίησης, με την έννοια της επιλογής χαρακτηριστικών γνωρίσματων, με την επιπρόσθετη χρήση μιας δεύτερης τεχνικής κατηγοριοποίησης ύστερα από την πρώτη φάση. Ο κύριος στόχος της προτεινόμενης επέκτασης είναι η αντιμετώπιση κλασικών προβλημάτων της ιεραρχικής συσταδοποίησης, όπως είναι η κατάρα της διαστατικότητας (curse of dimensionality) και η αρχική διάδοση λάθους (initial error propagation), αυξάνοντας με τον τρόπο αυτόν την ευρωστία της τεχνικής.

Ο αρχικός στόχος των αλγορίθμων συσταδοποίησης δεν έγκειται στο να ταξινομήσουν/κατηγοριοποιήσουν σωστά τα δεδομένα, αλλά στο να προσδιοριστούν τα πρότυπα που υποβόσκουν κάτω από τα δεδομένα και να παραγάγουν συστάδες παρόμοιων δειγμάτων δεδομένων. Επομένως, τα “λανθασμένα” στοιχεία των συστάδων

μπορούν να είναι εν γένει αποδεκτά, εφόσον όμως η γενική συστάδα περιγράφει σωστά ένα υπάρχον και με νόημα πρότυπο. Στο σημείο αυτό θα πρέπει να αναφερθεί ότι στην πραγματικότητα και, όπως έχουν δείξει σχετικές έρευνες [258], οι συστάδες που περιέχουν κακώς ορισμένα δείγματα δεδομένων μπορούν να είναι ακόμα και καλύτερες από τους θεωρητικά τέλειους διαχωρισμούς του συνόλου των δεδομένων κατά την περιγραφή των ελλοχευόντων προτύπων και με τον τρόπο αυτό μπορούν να οδηγήσουν στην καλύτερη αρχική συνθήκη για την εφαρμογή των ταξινομητών. Κάτι τέτοιο υπονοεί φυσικά ότι, εάν “ταΐσουμε” στον αλγόριθμο τα επονομαζόμενα στοιχεία και μετρήσουμε το ποσοστό ταξινόμησης, τότε αυτό μπορεί να μην είναι αρκετό να αξιολογήσει την πραγματική αποδοτικότητα του αλγορίθμου.

Για να μπορέσει ένας αλγόριθμος συσταδοποίησης να αξιολογηθεί κατάλληλα, θα πρέπει να αξιολογηθούν τα πρότυπα που περιγράφονται από τις συστάδες στην έξοδό του. Η εφαρμογή τους κατά τη διαδικασία παραγωγής ενός ταξινομητή και η αξιολόγηση του προκύπτοντος ταξινομητή είναι ένας αποδοτικός τρόπος προς αυτήν την κατεύθυνση. Στο πλαίσιο αυτής της υποενότητας εξετάζουμε εάν τα συγκεκριμένα αποτελέσματα ενός τέτοιου αλγορίθμου, που εφαρμόζονται σε διάφορα (ευρέως διαδεδομένα στο χώρο της μηχανικής μάθησης) σύνολα δεδομένων, έχουν κάποιο νόημα, κυρίως με την αξιολόγηση των αποτελεσμάτων από έναν ταξινομητή του κκοντινότερου γείτονα, ο οποίος δημιουργείται με τη χρησιμοποίησή τους.

Χωρίς αμφιβολία, στη βιβλιογραφία υπάρχουν ποικίλα σχέδια ταξινόμησης [157]. Επιλέξαμε να εργαστούμε με τον ταξινομητή του k-κοντινότερου γείτονα (kNN), αν και διάφοροι άλλοι ταξινομητές θα μπορούσαν να έχουν επιλεγεί στη θέση του, κυρίως λόγω της φύσης της ίδιας της βασισμένης σε στιγμιότυπα μεθόδου μάθησης που ακολουθεί και της απλής προσέγγισής του [175]. Ο αλγόριθμος kNN είναι πράγματι σχετικά σχεδιαστικά απλός<sup>1</sup>, ταυτόχρονα όμως είναι ιδιαίτερα εύρωστος, έχει χρησιμοποιηθεί σε πολλές εφαρμογές και μπορεί να εφαρμοστεί ακίνδυνα σε όλα τα είδη των συνόλων δεδομένων, είτε πρόκειται για δεδομένα από την πραγματική ζωή, είτε συνθετικά, ανεξάρτητα από τους όποιους συμβιβασμούς μεγέθους ή χρόνου, οδηγώντας έτσι σε υψηλής ποιότητας επιστημονικές παρατηρήσεις. Ο kNN είναι επίσης εξαιρετικά κατάλληλος, για να χρησιμοποιηθεί σε περιπτώσεις, όπου τα στιγμιότυπα αντιστοιχούν σε σημεία του  $\mathbb{R}^n$ , σε περιπτώσεις όπου υπάρχουν ιδιαιτέρως πολλά δεδομένα εκπαίδευσης και σε περιπτώσεις όπου μετά από την εφαρμογή μιας ευέλικτης επιλογής μετρήσιμων χαρακτηριστικών γνωρισμάτων έχουμε λιγότερες από 20 ιδιότητες ανά στιγμιότυπο.

Αναγνωρίζοντας το γεγονός ότι ο kNN συνήθως εξετάζει όλες τις ιδιότητες από όλα τα στοιχεία, τα πιθανά μειονεκτήματά του υπερνικώνται εύκολα με την εφαρμογή της αρχικής διαδικασίας συσταδοποίησης μόνο σε ένα μικρό υποσύνολο των διαθέσιμων δεδομένων, μειώνοντας κατά συνέπεια τον αριθμό στοιχείων που θα πρέπει να λάβει υπόψη του το σχέδιο ταξινόμησης, προκειμένου να ταξινομηθεί κάθε εισερχόμενο δείγμα. Η προσέγγιση αυτή είναι εξαιρετικά κατάλληλη για χρήση σε περιπτώσεις online ταξινόμησης. Συγκεκριμένα, ο αλγόριθμος kNN υποθέτει ότι όλα τα στοιχεία αντιστοιχούν σε σημεία στον n-διάστατο χώρο  $\mathbb{R}^n$ . Οι γείτονες ενός στοιχείου καθορίζονται με τη μορφή κάποιας μετρικής απόστασης. 'Οπως έχει προαναφερθεί, ποικίλες μετρικές [146], [169], [139] μπορούν να χρησιμοποιηθούν ως αποστάσεις στον

<sup>1</sup>Πρόκειται, άλλωστε, για τον αλγόριθμο που χρησιμοποιούν πολλά γνωστά διαδικτυακά καταστήματα, όπως π.χ. το ευρέως γνωστό Amazon.com, όταν προσφέρουν παραπλήσια, ως προς αυτό που έχει ήδη επιλέξει ο πελάτης, προϊόντα προς αγορά.

#### Κεφάλαιο 4. Εξόρυξη δεδομένων και κατηγοριοποίηση

αλγόριθμο, όπως η Ευκλείδια απόσταση:

$$d(a, b) = \sqrt{\sum_{i \in N_F} (a_i - b_i)^2} \quad (4.8)$$

η απόσταση Minkowsky:

$$d(a, b) = \sum_{i \in N_F} |a_i - b_i| \quad (4.9)$$

η απόσταση minimax

$$d(a, b) = \max \sum_{i \in N_F} |a_i - b_i| \quad (4.10)$$

η απόσταση Mahalanobis

$$d(a, b) = \sqrt{(a_i - b_i)^T \Sigma^{-1} (a_i - b_i)} \quad (4.11)$$

και άλλες. Στη συνέχεια, θεωρούμε κάθε αρχικό στοιχείο και υπολογίζουμε την απόστασή του από κάθε άλλο στοιχείο στο σύνολο των δεδομένων. Καθορίζουμε εκ των προτέρων τον αριθμό των κοντινότερων, στο υπό εξέταση στοιχείο, γειτόνων  $k$ , οι οποίοι πρόκειται να διαδραματίσουν έναν σημαντικό ρόλο στο χαρακτηρισμό της συστάδας του στοιχείου σε επόμενο στάδιο. Για να το πετύχουμε αυτό, χρησιμοποιούμε ένα κατάλληλο κατώτατο όριο σχετικά με την ακρίβεια της ταξινόμησης. Σαφώς, εάν το  $k$  γίνει πολύ μεγάλο, τότε οι ταξινομήσεις θα γίνουν όλες ίδιες. Γενικά, υπάρχει κάποια λογική στο να θεωρηθεί το  $k > 1$ , αλλά από την άλλη δε θα πρέπει να καταστήσουμε το  $k$  ίσο με τον αριθμό των στοιχείων εκπαίδευσης.

Στην τυπική περίπτωση, έστω ότι  $e_q$  είναι το εκάστοτε στοιχείο προς ταξινόμηση και ότι τα  $e_1, e_2, \dots, e_k$  δείχνουν τα στοιχεία  $k$  που είναι τα κοντινότερα στο  $e_q$ . Έστω επίσης ότι το  $c(a)$  ορίζεται ως:

$$c(a, j) = \begin{cases} 1 & , \text{ αν } \text{ το } a \text{ ανηκει στην κατηγορια } j \\ 0 & \text{αλλιως} \end{cases} \quad (4.12)$$

Τότε το  $e_q$  κατηγοριοποιείται ως ακολούθως:

$$c(e_q) = z : \sum (e_i, z) = \max \left( \sum_{j=1}^{\text{Πληθος} \text{Κατηγοριων}} c(e_i, j) \right) \quad (4.13)$$

όπου το  $e_i$  είναι το κοντινότερο στο  $e_q$  στιγμιότυπο εκπαίδευσης. Με άλλα λόγια, το  $e_q$  κατηγοριοποιείται στην κατηγορία, στην οποία ανήκουν οι περισσότεροι από τους  $k$  κοντινότερους γείτονές του.

Προφανώς, προκειμένου να εφαρμοστεί το σχέδιο ταξινόμησης του kNN, απαιτείται ένα μικρό σύνολο από ήδη χαρακτηρισμένα δείγματα δεδομένων. Στο πλαίσιο της έρευνας που διεξαγάγαμε περιγράφουμε την χωρίς επίβλεψη κατηγοριοποίηση των δεδομένων και έτσι υποθέτουμε ότι τέτοια πληροφορία δεν είναι διαθέσιμη. Χρησιμοποιούμε τις ετικέτες δεδομένων στα πειράματά μας μόνο και μόνο προκειμένου να μετρηθεί το ποσοστό ταξινόμησης και με τον τρόπο αυτό η απόδοση του αλγορίθμου. Επομένως, υποθέτουμε ότι καθεμία από τις συστάδες που ανιχνεύθηκε κατά τη διάρκεια του βήματος της ιεραρχικής συσταδοποίησης αντιστοιχεί σε μια πλήρως διακριτή κατηγορία.

Χρησιμοποιώντας το σχέδιο ταξινόμησης που περιγράφεται παραπάνω, καθώς και τις αναθέσεις συστάδων των συσταδοποιημένων δειγμάτων δεδομένων ως ετικέτες κατηγορίας, μπορούμε να προχωρήσουμε και να ταξινομήσουμε όλα τα διαθέσιμα στοιχεία δεδομένων. Εάν η αρχική συσταδοποίηση ήταν επιτυχής στην αποκάλυψη των προτύπων που υποβόσκουν κάτω από τα δεδομένα, τότε αυτή η διαδικασία θα εκλεπτύνει την έξοδο και θα βελτιώσει το ποσοστό ταξινόμησης με την απομάκρυνση μερικών μελών συστάδων που προήλθαν ως αποτέλεσμα λαθών κατά τα αρχικά βήματα. Κατά συνέπεια, αυτή η διαδικασία προσφέρει μια ένδειξη της αληθινής απόδοσης της ιεραρχικής συσταδοποίησης. Επιπλέον, καθιστά το συνολικό αλγόριθμο πιο γερό και εύρωστο, σε αντιδιαστολή με την απλή ιεραρχική συσταδοποίηση, δεδομένου ότι είναι πιο ανθεκτικός στα λάθη κατά τα αρχικά βήματα.

Τέλος, πρέπει να σημειωθεί ότι το βήμα της ταξινόμησης είναι αυτό που επεκτείνει τα συμπεράσματα της αρχικής συσταδοποίησης σε ολόκληρο το σύνολο των δεδομένων, επιτρέποντας κατά συνέπεια στα πρώτα να εφαρμόζονται μόνο σε μια μερίδα του συνόλου των δεδομένων. Αυτό είναι κάτι ιδιαίτερα σημαντικό, δεδομένου ότι χωρίς αυτό δε θα ήταν δυνατό να υπάρξουν τα οφέλη της ιεραρχικής συσταδοποίησης κατά την εξέταση μεγαλύτερων συνόλων δεδομένων. Επιπλέον, η επαναληπτική φύση του αλγορίθμου παίζει ένα σημαντικό ρόλο κατά τη διαδικασία της ταξινόμησης. Ειδικότερα, η είσοδος του αλγορίθμου είναι η ίδια με την έξοδό του, γεγονός που επιτρέπει διάφορες επαναληπτικές εφαρμογές του αλγορίθμου, έως ότου οι αναθέσεις συστάδων των στοιχείων παραμείνουν αμετάβλητες.

### 4.3 Ασαφής Ταξινομική Συσταδοποίηση

Προκειμένου να καταλήξουμε σε μια επιτυχή ανίχνευση κατηγοριών που σχετίζονται με ένα πολυμεσικό έγγραφο, μπορούμε να χρησιμοποιήσουμε τις κλασικές αρχές και μεθοδολογίες που παρουσιάστηκαν στην προηγούμενη ενότητα. Κάτι τέτοιο, όμως, δε θα αποτελούσε καινοτόμο συνεισφορά στο χώρο και, χυρίως, δε θα είχε τα επιθυμητά αποτελέσματα, όσον αφορά στην απόδοση και επίδοση κατηγοριοποίησης της προσέγγισής μας. Το κύριο έναυσμα πίσω από την ερευνητική ενασχόλησή μας ήταν η χρήση της κοινής σημασίας/ερμηνείας των εννοιών που απαντώνται στο πολυμεσικό περιεχόμενο ή με άλλα λόγια η εκμετάλλευση του εννοιολογικού πλαισίου στη διαδικασία. Προκειμένου να επιτύχουμε τον παραπάνω στόχο, πρέπει να επεκτείνουμε τους κλασικούς αλγόριθμους συσταδοποίησης προς την κατεύθυνση της ασαφοποίησης, έτσι ώστε να είναι σε θέση να αντιμετωπίσουν προβλήματα της πραγματικής ζωής, τα οποία, εν γένει, διέπονται από βαθμούς. Πρέπει, επίσης, να αναγνωρίσουμε κατάλληλες μετρικές απόστασης για τη διαδικασία της συσταδοποίησης, καθώς και να ορίσουμε τη δομή και τη μορφή της χρησιμοποιούμενης γνώσης, δύο παράμετροι που είναι εξαιρετικής σημασίας για την ευρωστία της μετέπειτα προσέγγισής μας στην περίπτωση των εγγράφων.

#### 4.3.1 Ασαφής ιεραρχική συσταδοποίηση

Προκειμένου να επεκτείνουμε τους κλασικούς αλγόριθμους συσταδοποίησης προς την κατεύθυνση της ευφυούς ασαφοποίησης, θα πρέπει να συσταδοποιήσουμε το σύνολο των στοιχείων σύμφωνα με μια κοινή τους έννοια. Στη γενική περίπτωση, έστω ότι έχουμε ένα σύνολο  $X$  στοιχείων προς συσταδοποίηση. Θεωρώντας το  $Y$  ως το σύνολο

#### Κεφάλαιο 4. Εξόρυξη δεδομένων και κατηγοριοποίηση

των συστάδων που προκύπτουν εφαρμόζοντας τον κλασικό αλγόριθμο συσσωρευτικής ιεραρχικής συσταδοποίησης (βλ. ενότητα 4.2.1.1), τότε κάθε συστάδα  $c \in Y$  αποτελεί ένα κλασικό σύνολο από στοιχεία. Παρ' όλα αυτά κάτι τέτοιο από μόνο του δεν είναι αρκετό. Κατά συνέπεια, χωρίς απώλεια της λειτουργικότητας ή επιπλέον υπολογιστικό κόστος αντικαθιστούμε τις κλασικές συστάδες  $c$  με ασαφείς κανονικοποιημένες συστάδες  $c^n$ . Κατόπιν, αθροίζοντας το εννοιολογικό πλαίσιο κάθε συστάδας  $c^n$ , προσδιορίζουμε το τελικό ασαφές σύνολο των στοιχείων. Στη συνέχεια αυτής της υποενότητας θα προχωρήσουμε στην αναλυτική παρουσίαση των αρχικών βημάτων της συσταδοποίησης εννοιών και της διαδικασίας “ασαφοποίησης” των συστάδων που προκύπτουν.

#### 4.3.2 Ανάλυση συστάδων

Κατά το σχεδιασμό ενός τέτοιου αλγορίθμου πρέπει να αντιμετωπιστεί μια σειρά από προκύπτοντα ζητήματα, όπως π.χ.:

- το στοιχείο μιας συστάδας μπορεί να σχετίζεται με πολλαπλές, ανεξάρτητες συστάδες, καθεμία με ένα συγκεκριμένο βαθμό συμμετοχής.
- οι συστάδες μπορεί να περιέχουν λανθασμένα αναγνωρισμένα στοιχεία.

Σύμφωνα με το πρώτο ζήτημα, ένα στοιχείο που ανήκει σε μια συστάδα μπορεί ταυτόχρονα να σχετίζεται και με άλλες, ανεξάρτητες μεταξύ τους συστάδες. Επομένως, είναι απαραίτητο για τον αλγόριθμο να είναι σε θέση να καθορίσει ποιες από αυτές τις συστάδες σχετίζονται πράγματι με ένα συγκεκριμένο στοιχείο. Προκειμένου να επιτευχθεί αυτός ο στόχος με έναν τρόπο που να έχει νόημα, πρέπει να εξεταστεί η κοινή έννοια των υπόλοιπων στοιχείων που συντάσσουν το σύνολο των δεδομένων υπό εξέταση. Από την άλλη, όταν ένα στοιχείο σχετίζεται με περισσότερες από μία, ανεξάρτητες συστάδες, δεν πρέπει να αναμένουμε ότι όλα τα στοιχεία που απαρτίζουν το σύνολο των δεδομένων θα σχετίζονται με κάθε μία από τις εν λόγω συστάδες. Αντιθέτως, πρέπει να αναμένουμε ότι τα περισσότερα στοιχεία θα σχετίζονται με ακριβώς μία από αυτές. Επομένως, είναι απαραίτητο να εφαρμοστεί μια εμπλουτισμένη συσταδοποίηση των στοιχείων, η οποία θα βασίζεται στην κοινή έννοια τους. Σε αυτήν την διαδικασία, τα στοιχεία που είναι παραπλανητικά (π.χ. στοιχεία που συμπεριελήφθησαν στο σύνολο δεδομένων εκ παραδρομής) δε θα βρεθεί ότι σχετίζονται με άλλα στοιχεία του συνόλου δεδομένων. Επομένως, η πληθικότητα των συστάδων που θα προκύψουν μπορεί να χρησιμοποιηθεί για να αντιμετωπίσει και το δεύτερο ζήτημα.

Η προτεινόμενη προσέγγιση είναι δυνατόν να αποσυντεθεί στα παρακάτω βήματα:

- Δημιουργούμε ένα μοντέλο γνώσης, όπως είναι λ.χ. μια οντολογία, το οποίο να είναι κατάλληλο προς χρήση από το μοντέλο της θεματικής κατηγοριοποίησης, για το οποίο το προορίζουμε.
- Καθορίζουμε τον αριθμό των διαφορετικών συστάδων, με τις οποίες σχετίζεται ένα στοιχείο, εκτελώντας μια διαμέριση των στοιχείων και χρησιμοποιώντας μια έκφανση του εννοιολογικού πλαισίου ή αλλιώς την κοινή σημασία τους (common meaning) ως κριτήριο της συσταδοποίησης.
- Ασαφοποιούμε τη διαμέριση, προκειμένου να επιτρέψει πιθανές επικαλύψεις των συστάδων και των ασαφών βαθμών ιδιότητας μέλους.

- Λαμβάνουμε υπόψη την πληθυκότητα των συστάδων και αγνοούμε τις παραπλανητικές συστάδες.

### 4.3.3 Δημιουργία γνώσης

Για τον προσδιορισμό του μοντέλου γνώσης που απαιτείται μπορεί να αναζητηθεί ένα κατάλληλο πλαίσιο για την αναπαράσταση της κοινής σημασίας και του εννοιολογικού πλαισίου στις οντολογίες. Ακολουθώντας μια υπεραπλουστευμένη ερμηνεία, μια οντολογία είναι, εν γένει, ένα σύνολο στοιχείων που καθορίζονται με κάποιο τρόπο, προκειμένου να δημιουργήσουν ένα συμφωνημένο λεξιλόγιο για την ανταλλαγή πληροφοριών και τη μοντελοποίηση εννοιών του πραγματικού κόσμου. Επιπλέον, το εννοιολογικό πλαίσιο καθορίζει την προοριζόμενη σημασία κάθε στοιχείου, δηλ. ένα στοιχείο που χρησιμοποιείται υπό το πρίσμα διαφορετικών εννοιολογικών πλαισίων είναι δυνατόν να λάβει διαφορετικές ερμηνείες. Συνεπώς, ένας πιθανός τρόπος να εξαχθεί και να χρησιμοποιηθεί το εννοιολογικό πλαίσιο για την επέκταση της συσταδοποίησης είναι να καθοριστεί μέσω της αξιοποίησης των ασαφών σχέσεων που υπάρχουν μεταξύ των στοιχείων. Για να δοθεί μια περαιτέρω επεξήγηση, μια απλή οντολογία είναι δυνατόν να περιγραφεί γενικά ως:

$$O = \{C, \{R_{c_i, c_j}\}\}, i, j = 1..n, i \neq j \quad (4.14)$$

$$R_{c_i, c_j} : C \times C \rightarrow \{0, 1\} \quad (4.15)$$

όπου το  $O$  αναπαριστά μία οντολογία, το  $C$  είναι το σύνολο όλων των πιθανών στοιχείων που περιγράφει και  $R_{c_i, c_j}$  είναι μια οποιαδήποτε σχέση μεταξύ δύο στοιχείων  $c_i, c_j$ .

Μια οντολογία μπορεί να περιέχει οποιοδήποτε τύπο σχέσεων, στο πλαίσιο, όμως, της προτεινόμενης διαδικασίας αξιοποίησης του εννοιολογικού πλαισίου κατά τη συσταδοποίηση εστιάζουμε μόνο σε έναν τύπο σχέσεων, τις ταξινομικές (taxonomic ή ordering) σχέσεις. Όπως θα δούμε σε επόμενα κεφάλαια, η χρήση ταξινομικών σχέσεων είναι απαραίτητη για τον προσδιορισμό του εννοιολογικού πλαισίου ενός πολυμεσικού εγγράφου. Κατά συνέπεια, το επόμενο βήμα είναι η ουσιαστική εκμετάλλευση των πληροφοριών που περιλαμβάνονται σε αυτή την ταξινομική σχέση μέσα στην οντολογία. Η δήλωση αυτή περιγράφει σαφώς μια χαρακτηριστική μορφή ιεραρχιών ταξινόμησης (taxonomy hierarchies), εντούτοις οι τελευταίες είναι στη γενική περίπτωση σαφείς (χλασικές) και, δεδομένου ότι οι έννοιες της πραγματικής ζωής διέπονται σχεδόν πάντα από ασάφεια, εύκολα γίνεται κατανοητό ότι μοντελοποιούνται καλύτερα χρησιμοποιώντας ασαφείς σχέσεις, όπως άλλωστε περιγράφεται και στο [4].

Κατά συνέπεια, οφείλουμε να εισαγάγουμε ένα νέο ορισμό του παραπάνω οντολογικού μοντέλου γνώσης. Το νέο μοντέλο είναι βασισμένο στις σχέσεις που μπορούν να μοντελοποιηθούν ως ασαφείς ταξινομικές σχέσεις διαταγής και συνοψίζονται στον ακόλουθο ορισμό οντολογίας:

$$O_F = \{C, \{r_{c_i, c_j}\}\}, i, j = 1..n, i \neq j \quad (4.16)$$

$$F(R_{c_i, c_j}) = r_{c_i, c_j} : C \times C \rightarrow [0, 1] \quad (4.17)$$

όπου το  $O_F$  εκφράζει την νέα οντολογία και η  $F(R_{c_i, c_j}) = r_{c_i, c_j}$  την ασαφή σχέση μεταξύ δύο στοιχείων  $c_i, c_j$ . Η τελευταία θα αποτελέσει τη βάση για την ταξινομία του εννοιολογικού πλαισίου.

#### 4.3.4 Ταξινομία εννοιολογικού πλαισίου

Ο απώτερος στόχος της ενότητας αυτής είναι να χρησιμοποιήσουμε την κοινή έννοια των στοιχείων, προκειμένου να καθοριστούν καλύτερα οι συστάδες που προκύπτουν από το σύνολο των δεδομένων. Θα αναφερόμαστε σε αυτό τον όρο ως **ταξινομικό εννοιολογικό πλαίσιο** (taxonomic context). Στη γενική περίπτωση, ο όρος ταξινομικό εννοιολογικό πλαίσιο αναφέρεται σε οτιδήποτε είναι κοινό μεταξύ ενός συνόλου στοιχείων. Για την ανίχνευση του ταξινομικού εννοιολογικού πλαισίου ενός συνόλου στοιχείων χρησιμοποιείται, εν γένει, μια κατάλληλα διαμορφωμένη σχέση  $r_{c_i, c_j}$ , η οποία σχετίζει όλα τα ζεύγη στοιχείων  $c_i, c_j$  μεταξύ τους. Επομένως, το ταξινομικό εννοιολογικό πλαίσιο ενός συνόλου δεδομένων καθορίζεται πάλι μέσω των στοιχείων που σχετίζονται με αυτό και καθορίζονται, εξάγονται και χρησιμοποιούνται μέσω της παραπάνω σχέσης.

Ορίζουμε το ταξινομικό εννοιολογικό πλαίσιο  $K(s)$  ενός στοιχείου  $x \in X$  ως το σύνολο των προγόνων του στη σχέση  $r$ , όπου το  $X$  είναι το σύνολο όλων των στοιχείων που περιλαμβάνονται στην οντολογία. Ειδικότερα, ακολουθώντας τον προτυποποιημένο συμβολισμό συνόλων-υποσυνόλων από την ασαφή σχεσιακή άλγεβρα [139], θα έχουμε στη γενική περίπτωση:  $K(x) = r_{\leq}(x)$ . Υποθέτοντας ότι ένα σύνολο στοιχείων  $Z \subseteq X$  είναι σαφές (κλασικό), το ταξινομικό εννοιολογικό πλαίσιο της συστάδας, καθορίζεται ως το σύνολο των κοινών προγόνων τους:

$$K(Z) = K(x_i), \quad x_i \in Z \quad (4.18)$$

Προφανώς, όσο περισσότερα στοιχεία λαμβάνονται υπόψη, το ταξινομικό εννοιολογικό πλαίσιο γίνεται “στενότερο”, δηλ. περιέχει λιγότερα στοιχεία και σε μικρότερους βαθμούς:

$$Z \supset W \rightarrow K(X) \subseteq K(W) \quad (4.19)$$

Όταν ο καθορισμός του ταξινομικού εννοιολογικού πλαισίου επεκταθεί στην περίπτωση των ασαφών συνόλων, η παραπάνω ιδιότητα πρέπει να ισχύει ακόμα. Απαιτούμε, λοιπόν, ότι, όταν το  $X$  είναι ένα κανονικό ασαφές σύνολο, το ταξινομικό εννοιολογικό πλαίσιο που λαμβάνεται υπόψη  $K(x)$  του στοιχείου  $x$ , δηλ. το ταξινομικό εννοιολογικό πλαίσιο του στοιχείου, όταν λαμβάνουμε υπόψη το βαθμό συμμετοχής του στο σύνολο, είναι χαμηλό, όταν ο βαθμός συμμετοχής  $Z(x)$  είναι υψηλός, ή όταν το ταξινομικό εννοιολογικό πλαίσιο του κλασικού στοιχείου  $K(x)$  είναι χαμηλό. Επομένως:

$$cp(K(x)) = cp(K(x)) \cap (X \cdot Z(x)) \quad (4.20)$$

όπου  $cp$  είναι ένα περιπλεκτικό (involutive) ασαφές συμπλήρωμα. Εφαρμόζοντας το νόμο de Morgan, έχουμε:

$$K(x) = K(x) \cup cp(X \cdot Z(x)) \quad (4.21)$$

Τότε, το συνολικό ταξινομικό εννοιολογικό πλαίσιο του συνόλου υπολογίζεται εύκολα ως:

$$K(Z) = K(x_i), \quad x_i \in Z \quad (4.22)$$

Το ύψος (height) του ταξινομικού εννοιολογικού πλαισίου  $h(K(Z))$ , δηλ. ο μέγιστος βαθμός ιδιότητας μέλους που εμφανίζεται σε αυτό, μπορεί να χρησιμοποιηθεί ως μέτρο του σημασιολογικού συσχετισμού των στοιχείων στο σύνολο  $Z$ . Θα αναφερόμαστε σε αυτή τη μετρική ως **ένταση** (intensity) του ταξινομικού εννοιολογικού πλαισίου, η οποία και θα φανεί ιδιαίτερα χρήσιμη μετέπειτα, κατά την εφαρμογή της αυτόματης κατηγοριοποίησης εγγράφων.

### 4.3.5 Ασαφής ιεραρχική συσταδοποίηση

Το πρώτο βήμα της δομής του αλγορίθμου ασαφούς ιεραρχικής συσταδοποίησης είναι όμοιο με αυτό που περιγράφεται στην ενότητα 4.2.1.1 για τον κλασικό αλγόριθμο:

1. Μετατρέπουμε κάθε στοιχείο των δεδομένων εισόδου σε μονοσύνολο (singleton), δηλ. σε μία δική του συστάδα.
2. Για κάθε ζεύγος συστάδων  $c_1, c_2$  υπολογίζουμε ένα δείκτη συμβατότητας  $CI(c_1, c_2)$ . Ο  $CI$  αναφέρεται επίσης και ως μετρική ομοιότητας ή ανομοιότητας των συστάδων ή μετρική της απόστασης των συστάδων.
3. Συγχωνεύουμε το ζευγάρι των συστάδων που έχουν τον καλύτερο δείκτη  $CI$ . Ανάλογα με το αν ο  $CI$  είναι μια μετρική ομοιότητας ή ανομοιότητας, ο καλύτερος δείκτης θα μπορούσε να επιλεγεί χρησιμοποιώντας το  $\max$  ή το  $\min$ , αντίστοιχα.
4. Συνεχίζουμε στο βήμα 2, έως ότου ικανοποιηθεί ένα κριτήριο τερματισμού. Το κριτήριο τερματισμού που χρησιμοποιείται συνήθως είναι ο καθορισμός ενός κατωφλίου για την τιμή του καλύτερου δείκτη συμβατότητας.

Τα δύο βασικά σημεία στον παραπάνω αλγόριθμο είναι ο προσδιορισμός των συστάδων προς συγχώνευση σε κάθε βήμα, δηλ. ο καθορισμός μίας μετρικής με νόημα για το  $CI$  και ο προσδιορισμός του βέλτιστου βήματος τερματισμού, δηλ. ο καθορισμός ενός κριτηρίου τερματισμού που να έχει κάποιο νόημα. Η ιδανική μετρική απόστασης για δύο συστάδες  $c_1, c_2$  είναι αυτή που ποσοτικοποιεί το σημασιολογικό συσχετισμό τους και μια τέτοια μετρική είναι η ένταση των κοινών εννοιολογικών πλαισίων γνώσης τους. Η διαδικασία της συσταδοποίησης ολοκληρώνεται, όταν συσταδοποιηθούν τα στοιχεία στα σύνολα που αντιστοιχούν στις διαφορετικές συστάδες, τα οποία και προσδιορίζονται από το γεγονός ότι τα κοινά εννοιολογικά πλαίσια γνώσης τους θα έχουν χαμηλή ένταση. Επομένως, το κριτήριο τερματισμού είναι ένα κατώτατο όριο στην επιλεγμένη μετρική συμβατότητας. Η έξοδος του αλγορίθμου αποτελείται από ένα σύνολο συστάδων  $Y$ , όπου κάθε συστάδα  $c \in Y$  αποτελεί ένα κλασικό σύνολο από στοιχεία.

Εώς εδώ η μέθοδος συσταδοποίησης, ώντας ιεραρχική, καθορίζει επιτυχώς το πλήθος των διαφορετικών συστάδων. Παρ' όλα αυτά δημιουργεί μόνο σαφείς (κλασικές) συστάδες και μόνο καταμερισμούς, δηλ. δεν επιτρέπει επικαλύψεις μεταξύ των προκυπτουσών συστάδων. Οι παρατηρήσεις αυτές αποτελούν δύο μεγάλα μειονεκτήματα για το συγκεκριμένο πρόβλημα και, προκειμένου να υπερνικηθούν, περιγράφουμε στη συνέχεια της υποενότητας μια μέθοδο για την ασαφοποίηση της διαμέρισης. Με τον τρόπο αυτό, οι κλιμακωτές πληθικότητες των συστάδων θα διορθωθούν, έτσι ώστε να μπορούν να χρησιμοποιηθούν αργότερα για το φιλτράρισμα των παραπλανητικών στοιχείων.

Κάθε συστάδα περιγράφεται από το σαφές σύνολο των στοιχείων που ανήκουν σε αυτή. Χρησιμοποιώντας τα στοιχεία αυτά, κατασκευάζουμε έναν ασαφή ταξινομητή, δηλ. μια συνάρτηση  $C_c : X \rightarrow [0, 1]$  που μετρά το βαθμό συσχέτισης μιας έννοιας  $x$  με την συστάδα  $c$ . Προφανώς, μια έννοια  $x$  πρέπει να θεωρηθεί συσχετισμένη με την συστάδα  $c$ , εάν σχετίζεται με την κοινή ερμηνεία (common meaning) των στοιχείων στην  $c$ . Για το λόγο αυτό, η ποσότητα  $C_1(c, x) = h(K(c \cup \{x\}))$  αποτελεί ένα κατάλληλο μέτρο της συσχέτισης. Φυσικά, δεν είναι όλες οι συστάδες εξίσου συμπαγείς. Εν γένει, μπορούμε να μετρήσουμε την πυκνότητα μιας συστάδας χρησιμοποιώντας

#### Κεφάλαιο 4. Εξόρυξη δεδομένων και κατηγοριοποίηση

την ομοιότητα μεταξύ των στοιχείων που αυτή περιέχει, δηλαδή χρησιμοποιώντας την ένταση του εννοιολογικού πλαισίου της συστάδας. Επομένως, το ανωτέρω μέτρο της συσχέτισης θα πρέπει να αναπροσαρμοστεί στα χαρακτηριστικά της εκάστοτε υπό εξέτασης συστάδας:

$$C_c(x) = \frac{C_1(c, x)}{h(K(c))} = \frac{h(K(c \cup \{x\}))}{h(K(c))} \quad (4.23)$$

Χρησιμοποιώντας αυτόν τον ταξινομητή, μπορούμε να επεκτείνουμε τις ανιχνευμένες κλασικές (σαφέις) συστάδες, έτσι ώστε να συμπεριλάβουν περισσότερα στοιχεία. Για το σκοπό αυτό η συστάδα  $c$  αντικαθίσταται από την ασαφή συστάδα:  $c^f \supseteq c : c^f = \sum_{x \in X} x / C_c(x)$ , χρησιμοποιώντας τον ευρύτατα διαδεδομένο συμβολισμό αθροίσματος (sum notation) [177] για την αναπαράσταση των ασαφών συνόλων.

Προκειμένου να ολοκληρωθεί η προσέγγισή μας, πρέπει να ληφθεί υπόψη το εννοιολογικό πλαίσιο και η βαθμωτή πληθικότητα κάθε συστάδας  $|c^f|$ . Δεδομένου ότι το εννοιολογικό πλαίσιο έχει καθοριστεί μόνο ως προς κανονικά ασαφή σύνολα (βλ. ενότητα 3.2), θα πρέπει να ομαλοποιήσουμε την συστάδα ως εξής:

$$c^n(x) = \frac{c^f(x)}{h(c^f(x))}, \forall x \in X \quad (4.24)$$

Στην περίπτωση, τέλος, που ανιχνεύονται περισσότερες από μια συστάδες, τότε είναι επιτακτικό να εξεταστούν και οι πληθικότητες των συστάδων. Οι συστάδες με εξαιρετικά χαμηλό αριθμό στοιχείων συνόλου περιέχουν πιθανότατα μόνο παραπλανητικά στοιχεία και επομένως πρέπει να αγνοηθούν. Αντίθετα, οι συστάδες με υψηλό αριθμό στοιχείων συνόλου αποτελούν την έξοδο του αλγορίθμου και είναι εκείνες ακριβώς που θα ληφθούν σοβαρά υπόψη στα επόμενα βήματα αξιοποίησής του. Θα πρέπει να σημειωθεί στο σημείο αυτό ότι η έννοια του “υψηλού αριθμού στοιχείων συνόλου” ή της μεγάλης πληθικότητας μοντελοποιείται με τη χρήση ενός “μεγάλου” ασαφούς αριθμού  $L(\cdot)$ , ο οποίος αποτελεί μια συνάρτηση από το σύνολο των πραγματικών θετικών αριθμών στο διάστημα  $[0, 1]$ , ποσοτικοποιώντας την έννοια του “μεγάλου” ή “υψηλού”.

## 4.4 Θεματική Κατηγοριοποίηση Εγγράφων

Μία από τις πολλές περιοχές όπου βρίσκουν εφαρμογή οι προαναφερθείσες τεχνικές και μεθοδολογίες της μηχανικής μάθησης (γενικότερα) και της ιεραρχικής συσταδοποίησης (ειδικότερα) είναι και η θεματική κατηγοριοποίηση εγγράφων. Η τελευταία, αποτελεί έναν από τους θεμέλιους λίθους της επίτευξης αποτελεσματικής διαχείρισης του πολυμεσικού περιεχομένου και γι' αυτό ακριβώς το λόγο έχει εξέχουσα βαρύτητα στη βιβλιογραφία και θέση μεταξύ της έρευνάς μας. Είναι γεγονός ότι η ταξινόμηση θεματικών κατηγοριών μπορεί να εφαρμοστεί τόσο σε πολυμεσικά, όσο και σε απλά έγγραφα κειμένου. Η προσέγγιση που θα ακολουθηθεί είναι η ίδια και για τις δύο περιπτώσεις, καθώς η δεύτερη θεωρείται υποπερίπτωση της πρώτης και συνακόλουθα στο υπόλοιπο του κεφαλαίου θα χρησιμοποιούμε τον όρο “έγγραφο” και δεν θα κάνουμε καμία σχετική διάκριση. Ο γενικότερος στόχος μας είναι η αυτόματη ανίχνευση των κατηγοριών - σε εννοιολογικό/σημασιολογικό επίπεδο - που σχετίζονται με ένα συγκεκριμένο έγγραφο και η ανάλογη κατηγοριοποίηση του τελευταίου.

Στη γενική περίπτωση, η αυτόματη θεματική κατηγοριοποίηση εγγράφων εκτελεί ένα ταίριασμα μεταξύ των μεταδεδομένων κειμένων που συνδέονται με το πολυμεσικό

υλικό και των οντοτήτων που περιέχονται στη βάση γνώσης. Παίρνει ως είσοδο μια σειρά θεματικών κατηγοριών από τη βάση γνώσης, δηλ. τις επιλεγμένες σημασιολογικές οντότητες, τις ενοποιεί με τις θεματικές κατηγορίες των πολυμεσικών αρχείων και παράγει ως έξοδο ένα σύνδεσμο (link) ο οποίος καταδεικνύει την ακριβή θέση των θεματικών κατηγοριών που σχετίζονται με κάθε πολυμεσικό ή μη έγγραφο. Η έξοδος της αυτόματης θεματικής κατηγοριοποίησης εγγράφων αποθηκεύεται στο σημασιολογικό ευρετήριο, το οποίο περιέχει ένα σύνολο δεικτών εγγράφων (document locators) για κάθε θεματική κατηγορία και σημασιολογική οντότητα της βάσης γνώσης (καθώς επίσης και για ένα μεγάλο σύνολο σύνθετων σημασιολογικών οντοτήτων) και χρησιμοποιείται για τη γρήγορη ανάκτηση των πολυμεσικών εγγράφων.

Μια σημαντική πτυχή του θέματος είναι ότι οι θεματικές κατηγορίες που αποθηκεύονται στη βάση γνώσης και συνεπώς και η ίδια η διαδικασία της αυτόματης θεματικής κατηγοριοποίησης εγγράφων, έχουν μια δυναμική φύση, με την έννοια ότι οι θεματικές κατηγορίες της βάσης γνώσης δεν ακολουθούν μία ένα-προς-ένα (one-to-one) αντιστοίχιση με τις θεματικές κατηγορίες των μεμονωμένων αρχείων. Ως εκ τούτου, ο ρόλος της αυτόματης θεματικής κατηγοριοποίησης εγγράφων είναι όχι μόνο να αναζητεί και να αποθηκεύει (ως συνδέσεις) τα πολυμεσικά έγγραφα που ανήκουν σε κάθε θεματική κατηγορία, αλλά και να υπολογίζει ένα βάρος συσχέτισης, το οποίο αντιπροσωπεύει το βαθμό εμπιστοσύνης με τον οποίο τα έγγραφά αυτά χαρακτηρίζονται από αυτή τη θεματική κατηγορία. Εδώ ο όρος βάρος συσχέτισης αναφέρεται στην ασαφή κατηγοριοποίηση κάθε πολυμεσικού εγγράφου σε μία ή περισσότερες θεματικές κατηγορίες αποθηκευμένες στη βάση γνώσης.

Η αυτόματη θεματική κατηγοριοποίηση εγγράφων λαμβάνει χώρα κατά τη διάρκεια της κατασκευής της βάσης γνώσης και των πολυμεσικών εγγράφων, προκειμένου να καταστεί δυνατή η ενημέρωση του τμήματος του σημασιολογικού ευρετηρίου που αναφέρεται στις θεματικές κατηγορίες. Κάθε φορά που εμπλουτίζονται τα πολυμεσικά αρχεία με νέα έγγραφα, η αυτόματη θεματική κατηγοριοποίηση εκτελείται για τα νέα έγγραφα, προκειμένου να ενημερώσει το σημασιολογικό ευρετήριο. Επίσης, κάθε φορά που ενημερώνεται η βάση γνώσης, η αυτόματη θεματική κατηγοριοποίηση εκτελείται για το σύνολο των πολυμεσικών εγγράφων, έτσι ώστε να δημιουργήσει ένα νέο σημασιολογικό ευρετήριο. Η λύση μιας βηματικής ανανέωσης δεν είναι κατάλληλη στην περίπτωση αυτή, δεδομένου ότι οι νέες οντότητες θα υποβληθούν σε διαφορετική ανάλυση, όσον αφορά τις περιγραφές των εγγράφων.

Όπως αναφέρθηκε ήδη, η αυτόματη θεματική κατηγοριοποίηση εγγράφων δέχεται ως είσοδο το σημασιολογικό ευρετήριο και τη γνώση του συστήματος. Η εκτίμηση του συνόλου των σημασιολογικών σχέσεων είναι περιττή διαδικασία και κοστίζει ιδιαίτερα σε υπολογιστικούς πάρους. Επομένως, χωρίς βλάβη της γενικότητας θα εξετάσουμε ένα μικρό αλλά αντιπροσωπευτικό υποσύνολο των διαθέσιμων σχέσεων. Στα πλαίσια της προτεινόμενης εργασίας θα περιοριστούμε σε σχέσεις που θα παρουσιαστούν αναλυτικά σε επόμενα κεφάλαια, ούτως ώστε να διατηρηθεί η συνολική συνοχή της.

Ειδικότερα, η αυτόματη θεματική κατηγοριοποίηση εγγράφων πρέπει να καθορίσει για κάθε πολυμεσικό έγγραφο το ασαφές σύνολο θεματικών κατηγοριών με το οποίο σχετίζεται, δηλ. εν προκειμένω ζευγάρια θεματικών κατηγοριών και βαθμών. Αυτοί οι βαθμοί μπορούν να ερμηνευθούν είτε ως βαθμοί εμπιστοσύνης στην έξοδο, είτε ως βαθμοί συσχέτισης του εκάστοτε εξεταζόμενου εγγράφου με κάθε θεματική κατηγορία. Δεδομένου ότι οι δύο βαθμοί δεν είναι ανεξάρτητοι (οι υψηλοί βαθμοί συσχέτισης ανιχνεύονται συνήθως με υψηλούς βαθμούς εμπιστοσύνης και αντίστροφα), ένας ενιαίος βαθμός είναι τελικά επαρκής ως έξοδος. Στη συνέχεια, θα παρουσιάσουμε τον

#### Κεφάλαιο 4. Εξόρυξη δεδομένων και κατηγοριοποίηση

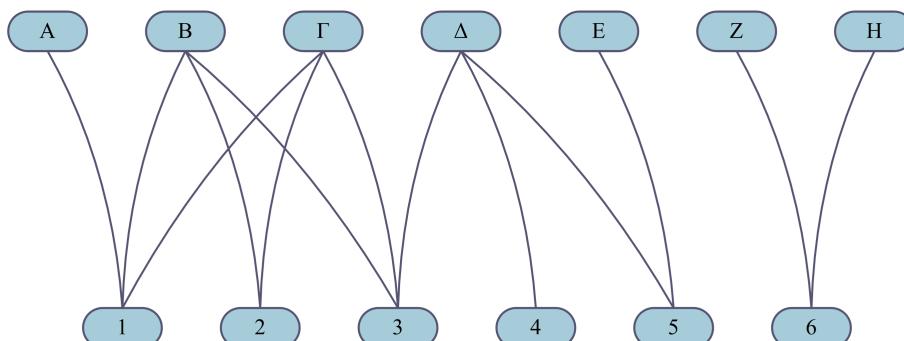
αλγόριθμο αυτόματης θεματικής κατηγοριοποίησης εγγράφων που προτείνουμε και ο οποίος χρησιμοποιεί πληροφορίες από το σημασιολογικό ευρετήριο.

##### 4.4.1 Ανάλυση εγγράφων

Η προσέγγιση που ακολουθείται για την αναπαράσταση, την εξαγωγή και τη χρησιμοποίηση των θεματικών κατηγοριών είναι βασισμένη σε μια μεθοδολογία που βασίζεται στην ασαφή σχεσιακή άλγεβρα, όπως αυτή παρουσιάστηκε στην ενότητα 3.2. Η υποενότητα αυτή αναφέρεται στην ανάλυση του σημασιολογικού ευρετηρίου με στόχο την εξαγωγή της σημασιολογίας ενός εγγράφου. Με άλλα λόγια, στοχεύουμε να ανιχνεύσουμε ποιες σημασιολογικές οντότητες σχετίζονται πράγματι με ένα έγγραφο και μέχρι ποιο σημείο. Ειδικότερα, δεχόμαστε ως είσοδο τη σημασιολογική ευρετηρίαση των διαθέσιμων εγγράφων, δηλ. το ίδιο το σημασιολογικό ευρετήριο. Στο τελευταίο, κάθε έγγραφο αντιπροσωπεύεται ως ένα κανονικό ασαφές σύνολο ορισμένο στο σύνολο των σημασιολογικών οντοτήτων. Με βάση αυτό το σύνολο και τη γνώση που περιλαμβάνεται στις διαθέσιμες σημασιολογικές σχέσεις στοχεύουμε να ανιχνεύσουμε το βαθμό, στον οποίο ένα δεδομένο έγγραφο  $d$  είναι πράγματι σχετικό με μια σημασιολογική οντότητα  $s \in S$ . Θα αναφερθούμε σε αυτόν τον βαθμό ως  $R_T(s, d)$ . Με άλλα λόγια, προσπαθούμε να υπολογίσουμε μια σχέση:

$$R_T : S \times D \rightarrow [0, 1] \quad (4.25)$$

όπου το  $D$  είναι το σύνολο των διαθέσιμων εγγράφων. Κατά το σχεδιασμό ενός τέτοιου αλγορίθμου, ο οποίος θα είναι σε θέση να υπολογίσει αυτήν την σχέση με ένα τέτοιο τρόπο που να έχει νόημα, πρέπει να αντιμετωπιστεί μια σειρά από ανακύπτοντα ζητήματα, όπως φαίνεται και στο ακόλουθο Σχήμα 4.3. Τα ζητήματα που πρέπει να εξεταστούν κατά την ανίχνευση των σχετικών κατηγοριών είναι τα εξής: (i) δεν υπάρχει καμία κατηγορία που να αφορά σε όλες τις οντότητες (ii) η οντότητα 1 σχετίζεται με τις κατηγορίες A, B και Γ, αλλά μόνο οι τελευταίες δύο σχετίζονται με ολόκληρο το έγγραφο εξαιτίας των πολλαπλών ( $> 1$ ) σχέσεων τους με διαφορετικές κατηγορίες μέσα στο έγγραφο (iii) η οντότητα 3 σχετίζεται με δύο διαφορετικές κατηγορίες· οι κατηγορίες Β και Γ θεωρούνται ως μία ενιαία ομάδα, καθώς σχετίζονται με ακριβώς τις ίδιες οντότητες (iv) οι κατηγορίες Z και H σχετίζονται με μόνο μια από τις οντότητες του εγγράφου, γεγονός που θα μπορούσε να είναι συμπτωματικό. Ειδικότερα και όσον αφορά τη σχέση οντοτήτων και κατηγοριών:



**Σχήμα 4.3:** Οι οντότητες που ευρετηριάζουν ένα έγγραφο  $(1, 2, \dots, 6)$  και οι σχέσεις μεταξύ τους.

- Μια σημασιολογική οντότητα μπορεί να σχετίζεται με πολλαπλές, ανεξάρτητες κατηγορίες.
- Ένα έγγραφο μπορεί να σχετίζεται με πολλαπλές, ανεξάρτητες κατηγορίες.
- Το σημασιολογικό ευρετήριο μπορεί να περιέχει λανθασμένα αναγνωρισμένες οντότητες.

Σύμφωνα με το πρώτο ζήτημα, μια σημασιολογική οντότητα μπορεί να αντιστοιχεί σε πολλές, ανεξάρτητες μεταξύ τους κατηγορίες. Επομένως, είναι απαραίτητο για τον αλγόριθμο να είναι σε θέση να καθορίσει ποιες από αυτές τις κατηγορίες σχετίζονται πράγματι με ένα συγκεκριμένο έγγραφο. Προκειμένου να επιτευχθεί αυτός ο στόχος με έναν τρόπο που να έχει νόημα, πρέπει να εξεταστεί η κοινή έννοια των υπόλοιπων οντοτήτων που συντάσσουν το ευρετήριο για το δεδομένο έγγραφο. Από την άλλη, όταν ένα έγγραφο σχετίζεται με περισσότερες από μία, ανεξάρτητες κατηγορίες, όπως καταδεικνύει το δεύτερο ζήτημα, δεν πρέπει να αναμένουμε ότι όλες οι οντότητες που συντάσσουν το ευρετήριό του θα σχετίζονται με κάθε μία από τις εν λόγω κατηγορίες. Αυτιθέτως, πρέπει να αναμένουμε ότι οι περισσότερες οντότητες θα σχετίζονται με ακριβώς μία από αυτές. Επομένως, είναι απαραίτητο να εφαρμοστεί μια συσταδοποίηση των σημασιολογικών οντοτήτων, βασισμένη στην κοινή έννοιά τους. Σε αυτήν τη διαδικασία οι οντότητες που είναι παραπλανητικές (π.χ. οντότητες που προέκυψαν από την ανακριβή ανίχνευση των οντοτήτων στο έγγραφο) δε θα βρεθεί ότι σχετίζονται με άλλες οντότητες που συντάσσουν το ευρετήριο του δεδομένου εγγράφου. Επομένως, η πληθυκότητα των ομάδων που θα προκύψουν μπορεί να χρησιμοποιηθεί, για να αντιμετωπίσει το τρίτο από τα παραπάνω ζητήματα.

Η προτεινόμενη προσέγγιση είναι δυνατόν να αποδομηθεί στα παρακάτω βήματα:

- Δημιουργούμε ένα μοντέλο γνώσης, όπως είναι μια συνδυασμένη σημασιολογική σχέση με τη μορφή μιας οντολογίας, κατάλληλο προς χρήση από το μοντέλο της θεματικής κατηγοριοποίησης.
- Καθορίζουμε τον αριθμό των διαφορετικών κατηγοριών με τις οποίες σχετίζεται ένα έγγραφο, εκτελώντας μια διαμέριση των σημασιολογικών οντοτήτων, χρησιμοποιώντας μια έκφανση του εννοιολογικού πλαισίου ή αλλιώς την κοινή σημασία τους (common meaning) ως χριτήριο της συσταδοποίησης.
- Ασφοποιούμε τη διαμέριση, προκειμένου να επιτρέψει πιθανές επικαλύψεις των συστάδων και των ασαφών βαθμών ιδιότητας μέλους.
- Προσδιορίζουμε την θεματική κατηγορία που σχετίζεται με κάθε συστάδα.
- Αθροίζουμε τις θεματικές κατηγορίες για τις διαφορετικές συστάδες, προκειμένου να εξαχθεί ένα συνολικό αποτέλεσμα για το έγγραφο.

#### 4.4.2 Επέκταση γνώσης

Η ύπαρξη μιας πληθώρας σημασιολογικών σχέσεων στη βάση γνώσης οδηγεί αναπόφευκτα στη διαίρεση της διαθέσιμης γνώσης μεταξύ τους, το οποίο με τη σειρά του οδηγεί στην ανάγκη για χρησιμοποίηση περισσότερων σχέσεων για την παραγωγή μιας επαρκούς ταξινομικής σχέσης, όπως θα δούμε και σε επόμενα κεφάλαια. Ειδικότερα, χρησιμοποιούμε μια ταξινομική σχέση που έχει παραχθεί με τη χρήση των ακόλουθων, βασισμένων στο πρότυπο MPEG-7, σημασιολογικών σχέσεων (βλ. Παράρτημα B):

- Μέρος  $P$ , ανεστραμμένη.
- Ειδίκευση  $Sp$ .
- Παράδειγμα  $Ex$ . Το  $Ex(a, b) > 0$  δείχνει ότι το  $b$  είναι ένα παράδειγμα του  $a$ . Για παράδειγμα, το  $a$  μπορεί να είναι ένας “παίκτης” και το  $b$  μπορεί να είναι ο “Jordan”.
- Όργανο  $Ins$ . Το  $Ins(a, b) > 0$  δείχνει ότι το  $b$  είναι ένα όργανο του  $a$ . Για παράδειγμα, το  $a$  μπορεί να είναι η “μουσική” και το  $b$  μπορεί να είναι τα “υτραύλια”.
- Θέση  $Loc$ , ανεστραμμένη. Το  $Loc(a, b) > 0$  δείχνει ότι το  $b$  είναι μία θέση του  $a$ . Για παράδειγμα, το  $a$  μπορεί να είναι ένα “κονσέρτο” και το  $b$  η “σκηνή”.
- Ιδιότητα  $Pr$ , ανεστραμμένη. Το  $Pr(a, b) > 0$  δείχνει ότι το  $b$  είναι μια ιδιότητα του  $a$ . Για παράδειγμα, το  $a$  μπορεί να είναι ο “Jordan” και το  $b$  μπορεί να είναι “αστέρι”.

$$T = (P^{-1} \cup Sp \cup Ex \cup Ins \cup Loc^{-1} \cup Pr^{-1})^{n-1} \quad (4.26)$$

Με βάση τη σημασιολογία των παραπάνω σχέσεων είναι εύκολο να δει κανείς ότι η  $T$  είναι ιδανική για τον προσδιορισμό των κατηγοριών, με τις οποίες μπορεί να σχετίζεται μία οντότητα. Η ταξινομική αυτή σχέση αποτελεί μια προσαρμογή αυτής που καθορίζεται στην ενότητα 3.2, όπου μπορεί κανείς να ανατρέξει για περισσότερες θεωρητικές πληροφορίες.

#### 4.4.3 Ταξινομικό εννοιολογικό πλαίσιο

Κατά τη χρησιμοποίηση μιας οντολογίας, όπως αυτή που χρησιμοποιήσαμε στα πλαίσια της παρούσης, ιδιαίτερη σημασία έχει το ταξινομικό εννοιολογικό πλαίσιο ενός όρου. Με άλλα λόγια, η αληθινή πηγή πληροφοριών είναι η επανεμφάνιση ορισμένων οντοτήτων και όχι καθεμία ανεξάρτητα. Κατά συνέπεια, στο στάδιο της ανάλυσης των εγγράφων θα πρέπει, όπως είδαμε, να χρησιμοποιήσουμε την κοινή έννοια των σημασιολογικών οντοτήτων, προκειμένου να καθοριστούν καλύτερα οι κατηγορίες που σχετίζονται με κάθε υπό εξέταση έγγραφο. Από εδώ και πέρα θα αναφερόμαστε σε αυτή την έννοια με τον όρο ταξινομικό εννοιολογικό πλαίσιο (taxonomic context). Στη γενική περίπτωση, ο όρος ταξινομικό εννοιολογικό πλαίσιο αναφέρεται σε οτιδήποτε είναι κοινό μεταξύ ενός συνόλου στοιχείων. Για την ανίχνευση του ταξινομικού εννοιολογικού πλαισίου ενός συνόλου σημασιολογικών οντοτήτων θα χρησιμοποιηθεί μία κατάλληλα διαμορφωμένη σχέση. Γενικά, θεωρούμε ότι ένα έγγραφο  $d$  αντιπροσωπεύεται αποκλειστικά από την αντιστοίχισή του στις σημασιολογικές οντότητες μέσω ενός σημασιολογικού ευρετηρίου. Επομένως, το ταξινομικό εννοιολογικό πλαίσιο ενός εγγράφου καθορίζεται πάλι μέσω των σημασιολογικών οντοτήτων που συσχετίζονται με αυτό και καθορίζονται, εξάγονται και χρησιμοποιούνται μέσω της παραπάνω σχέσης.

Στη συνέχεια, βασιζόμενοι στη σημασιολογία της σχέσης  $T$ , καθορίζουμε το ταξινομικό εννοιολογικό πλαίσιο  $K(s)$  μιας σημασιολογικής οντότητας  $s \in S$  ως το σύνολο των προγόνων της στη σχέση  $T$ , όπου το  $S$  είναι το σύνολο όλων των σημασιολογικών οντοτήτων που περιλαμβάνονται στην οντολογία. Τυπικότερα:  $K(s) =$

$T_{\leq}(s)$ , ακολουθώντας τον προτυποποιημένο συμβολισμό συνόλων-υποσυνόλων από την ασαφή σχεσιακή άλγεβρα [139]. Υποθέτοντας ότι ένα σύνολο οντοτήτων  $A \subseteq S$  είναι ασαφές, δηλ. όλες οι εξεταζόμενες οντότητες ανήκουν στο σύνολο με βαθμό 1, το ταξινομικό εννοιολογικό πλαίσιο της ομάδας, που είναι πάλι ένα σύνολο σημασιολογικών οντοτήτων, μπορεί να καθοριστεί απλά ως το σύνολο των κοινών προγόνων τους:

$$K(A) = \bigcap_i K(s_i), \quad s_i \in A \quad (4.27)$$

Προφανώς, όσο περισσότερες οντότητες λαμβάνονται υπόψη, το ταξινομικό εννοιολογικό πλαίσιο γίνεται στενότερο, δηλ. περιέχει λιγότερες οντότητες και σε μικρότερους βαθμούς:

$$A \supset B \rightarrow K(A) \subseteq K(B) \quad (4.28)$$

Όταν ο καθορισμός του ταξινομικού εννοιολογικού πλαισίου επεκτείνεται στην περίπτωση των ασαφών συνόλων σημασιολογικών οντοτήτων, η παραπάνω ιδιότητα πρέπει να ισχύει ακόμα. Απαιτούμε, λοιπόν, ότι, όταν το  $A$  είναι ένα κανονικό ασαφές σύνολο, το ταξινομικό εννοιολογικό πλαίσιο που λαμβάνεται υπόψη  $K(s)$  της  $s$ , δηλ. το ταξινομικό εννοιολογικό πλαίσιο της οντότητας, όταν λαμβάνουμε υπόψη το βαθμό συμμετοχής της στο σύνολο, είναι χαμηλό, όταν ο βαθμός συμμετοχής  $A(s)$  είναι υψηλός, ή όταν το ταξινομικό εννοιολογικό πλαίσιο της σαφούς οντότητας  $K(s)$  είναι χαμηλό. Επομένως:

$$cp(K(s)) \doteq cp(K(s)) \cap (S \cdot A(s)) \quad (4.29)$$

όπου  $cp$  είναι ένα περιπλεκτικό (involutive) ασαφές συμπλήρωμα. Με την εφαρμογή του νόμου de Morgan, λαμβάνουμε:

$$K(s) = K(s) \cup cp(S \cdot A(s)) \quad (4.30)$$

Τότε, το συνολικό ταξινομικό εννοιολογικό πλαίσιο του συνόλου υπολογίζεται εύκολα ως:

$$K(A) = \bigcap_i K(s_i), \quad s_i \in A \quad (4.31)$$

Προκειμένου να επεξηγηθούν περαιτέρω τα παραπάνω, χρησιμοποιούμε το ακόλουθο αριθμητικό παράδειγμα. Έστω  $A$  ένα κανονικό ασαφές σύνολο  $A = \mu\alpha\lambda\alpha/1.0 + \delta\iota\alpha\iota\tau\eta\tau\eta\varsigma/0.6 + \kappa\alpha\lambda\alpha\theta\iota/0.8$  ορισμένο στο κλασικό σύνολο  $S = \{\mu\alpha\lambda\alpha, \delta\iota\alpha\iota\tau\eta\tau\eta\varsigma, \kappa\alpha\lambda\alpha\theta\iota\}$ . Οι ποσότητες  $cp(S \cdot A(\mu\alpha\lambda\alpha))$ ,  $cp(S \cdot A(\delta\iota\alpha\iota\tau\eta\tau\eta\varsigma))$  και  $cp(S \cdot A(\kappa\alpha\lambda\alpha\theta\iota))$  υπολογίζονται ως εξής:

$$\begin{aligned} cp(S \cdot A(\mu\alpha\lambda\alpha)) &= cp(\mu\alpha\lambda\alpha/1.0 + \delta\iota\alpha\iota\tau\eta\tau\eta\varsigma/0.6 + \kappa\alpha\lambda\alpha\theta\iota/0.8) \\ &= \mu\alpha\lambda\alpha/0 + \delta\iota\alpha\iota\tau\eta\tau\eta\varsigma/0 + \kappa\alpha\lambda\alpha\theta\iota/0 \\ cp(S \cdot A(\kappa\alpha\lambda\alpha\theta\iota)) &= cp(\mu\alpha\lambda\alpha/0.6 + \delta\iota\alpha\iota\tau\eta\tau\eta\varsigma/0.4 + \kappa\alpha\lambda\alpha\theta\iota/0.4) \\ &= \mu\alpha\lambda\alpha/0.4 + \delta\iota\alpha\iota\tau\eta\tau\eta\varsigma/0.4 + \kappa\alpha\lambda\alpha\theta\iota/0.4 \quad (4.32) \\ cp(S \cdot A(\delta\iota\alpha\iota\tau\eta\tau\eta\varsigma)) &= cp(\mu\alpha\lambda\alpha/0.8 + \delta\iota\alpha\iota\tau\eta\tau\eta\varsigma/0.2 + \kappa\alpha\lambda\alpha\theta\iota/0.2) \\ &= \mu\alpha\lambda\alpha/0.2 + \delta\iota\alpha\iota\tau\eta\tau\eta\varsigma/0.2 + \kappa\alpha\lambda\alpha\theta\iota/0.2 \end{aligned}$$

Σύμφωνα με την εξίσωση (4.30) έχουμε:

$$\begin{aligned}\mathcal{K}(\mu\pi\alpha\alpha) &= \text{ποδοσφαιρο}/0.8 + \text{καλαθοσφαιριση}/0.8 \\ \mathcal{K}(\kappa\alpha\lambda\theta\iota) &= \text{ποδοσφαιρο}/0.75 + \text{καλαθοσφαιριση}/0.75 + \\ &\quad \text{μπαλα}/0.4 + \text{διαιτητης}/0.4 + \text{καλαθι}/0.4 \\ \mathcal{K}(\delta\iota\alpha\iota\tau\eta\tau\eta\varsigma) &= \text{καλαθοσφαιριση}/0.75 + \text{μπαλα}/0.2 + \\ &\quad \text{διαιτητης}/0.2 + \text{καλαθι}/0.2\end{aligned}\quad (4.33)$$

Κατά συνέπεια, το  $K(A)$  υπολογίζεται από την εξίσωση (4.31) ως εξής:

$$\begin{aligned}K(A) &= \mathcal{K}(\mu\pi\alpha\alpha) \cap \mathcal{K}(\delta\iota\alpha\iota\tau\eta\tau\eta\varsigma) \cap \mathcal{K}(\kappa\alpha\lambda\theta\iota) \\ &= \text{ποδοσφαιρο}/0 + \text{καλαθοσφαιριση}/0.75 + \text{μπαλα}/0 + \\ &\quad \text{διαιτητης}/0 + \text{καλαθι}/0 \\ &= \text{καλαθοσφαιριση}/0.75\end{aligned}\quad (4.34)$$

Εξετάζοντας την σημασιολογία της σχέσης  $T$  και την διαδικασία του προσδιορισμού του ταξινομικού εννοιολογικού πλαισίου, είναι εύκολο να συνειδητοποιήσουμε ότι, όταν οι οντότητες σε ένα σύνολο σχετίζονται ιδιαίτερα με μια κοινή έννοια, το ταξινομικό εννοιολογικό πλαίσιο θα έχει υψηλούς βαθμούς ιδιότητας μέλους για τις οντότητες που αντιπροσωπεύουν αυτήν την κοινή έννοια. Επομένως, το ύψος (height) του ταξινομικού εννοιολογικού πλαισίου  $h(K(A))$ , δηλ. ο μέγιστος βαθμός ιδιότητας μέλους που εμφανίζεται σε αυτό, μπορεί να χρησιμοποιηθεί ως μέτρο του σημασιολογικού συσχετισμού των οντοτήτων στο σύνολο  $A$ . Θα αναφερόμαστε σε αυτή τη μετρική ως ένταση (intensity) του ταξινομικού εννοιολογικού πλαισίου, η οποία και θα φανεί ιδιαίτερα χρήσιμη παρακάτω.

#### 4.4.4 Ερμηνεία του χαρακτηρισμού ενός πολυμεσικού εγγράφου

Όπως έχει ήδη αναφερθεί, οι ορισμοί των σημασιολογικών οντοτήτων στην οντολογία περιέχουν ακολουθίες ετικετών, που καθεμία παρέχει μια διαφορετική κειμενική μορφή της σημασιολογικής οντότητας, ενδεχομένως, δε, και σε περισσότερες από μία γλώσσες. Ταιριάζοντας τα παραπάνω με τους όρους που περιέχονται στο χαρακτηρισμό (annotation) ενός πολυμεσικού εγγράφου, μπορούμε να αποκτήσουμε τη σημασιολογική αναπαράσταση του ίδιου του χαρακτηρισμού. Φυσικά, στις περισσότερες περιπτώσεις αυτό δεν είναι καθόλου απλό ή τετριμένο: η αντιστοιχιση μεταξύ των όρων και των σημασιολογικών οντοτήτων είναι μια σχέση πολλών-μεπολλά (many-to-many relation), το οποίο σημαίνει ότι υπάρχουν πολλαπλές πιθανές σημασιολογικές ερμηνείες για έναν και μοναδικό χαρακτηρισμό κειμένου. Π.χ. ας εξετάσουμε την περίπτωση του όρου “στοιχείο”. Τουλάχιστον δύο διαφορετικές σημασιολογικές οντότητες αντιστοιχούν σε αυτόν: “στοιχείο1”, που σχετίζεται με τη χημεία και “στοιχείο2”, το οποίο σχετίζεται με την γλώσσα XML. Εάν υποτεθεί ότι ένας χαρακτηρισμός κειμένου περιλαμβάνει τον όρο “στοιχείο”, το σύστημα πρέπει να είναι σε θέση να καθορίσει αυτόματα σε ποια σημασιολογική οντότητα στην οντολογία θα πρέπει να αντιστοιχηθεί ο όρος, προκειμένου να ανακτηθούν τα αντίστοιχα πολυμεσικά ή μη έγγραφα από το ευρετήριο. Στο ίδιο παράδειγμα, εάν οι υπόλοιποι όροι του χαρακτηρισμού σχετίζονται με τη χημεία, τότε είναι αρκετά ασφαλές να υποτεθεί ότι το κείμενο αναφέρεται στη σημασιολογική οντότητα “στοιχείο1”, παρά στην σημασιολογική οντότητα “στοιχείο2”. Τα παραπάνω υπονοούν ότι το ταξινομικό

εννοιολογικό πλαίσιο του χαρακτηρισμού μπορεί να χρησιμοποιηθεί, για να διευκολυνθεί η διαδικασία του σημασιολογικού προσδιορισμού οντοτήτων στην περίπτωση αμφισημειών.

Εντούτοις, η εκτίμηση του ταξινομικού εννοιολογικού πλαισίου του χαρακτηρισμού απαιτεί ως είσοδο την αναπαράσταση του χαρακτηρισμού ως ένα ασαφές σύνολο από οντότητες και έτσι δεν μπορεί να εκτελεσθεί, προτού να ολοκληρωθεί η ερμηνεία του χαρακτηρισμού. Συνακόλουθα, η τελευταία πρέπει να λάβει χώρα ταυτόχρονα με την εκτίμηση του ταξινομικού εννοιολογικού πλαισίου. Ακολουθούμε την εξής μέθοδο: έστω ότι ο χαρακτηρισμός του κειμένου περιέχει τους όρους  $\{t_i\}$ , όπου  $i = 1, \dots, T$ . Έστω, ακόμη, ότι το  $t_i$  είναι η βασισμένη στο κείμενο περιγραφή των σημασιολογικών οντοτήτων  $\{s_{ij}\}$  με  $j = 1, \dots, T_i$ . Τότε, υπάρχουν  $N_Q = \prod_i T_i$  ξεχωριστοί συνδυασμοί από σημασιολογικές οντότητες που μπορούν να χρησιμοποιηθούν για την αναπαράσταση του χαρακτηρισμού του κειμένου. Για κάθε μία από αυτές υπολογίζουμε το αντίστοιχο ταξινομικό εννοιολογικό πλαίσιο. Όπως έχει ήδη εξηγηθεί, η ένταση του ταξινομικού εννοιολογικού πλαισίου είναι ένα σημασιολογικό μέτρο της συσχέτισης των οντοτήτων σε ένα σύνολο. Κατά συνέπεια, από τους υποψήφιους χαρακτηρισμούς  $\{q_k\}$ , όπου  $k = 1, 2, \dots, N_Q$ , αυτός που παράγει το εντονότερο ταξινομικό εννοιολογικό πλαίσιο είναι αυτός που περιέχει τις σημασιολογικές οντότητες που συσχετίζονται περισσότερο μεταξύ τους. Αυτός είναι και ο συνδυασμός που επιλέγεται ως έξοδος για την διαδικασία ερμηνείας του χαρακτηρισμού:

$$q = q_i \in \{q_1, \dots, q_{N_Q}\} : h(q_i) \geq h(q_j) \quad \forall q_j \in \{q_1, \dots, q_{N_Q}\} \quad (4.35)$$

Αυτή η ερμηνεία του σημασιολογικού χαρακτηρισμού είναι εξαντλητική, υπό την έννοια ότι πρέπει να εξετάσει όλες τις πιθανές ερμηνείες ενός δεδομένου χαρακτηρισμού. Το τελευταίο δεν αποτελεί πρόβλημα, εφόσον:

- οι χαρακτηρισμοί δεν περιέχουν μεγάλο αριθμό από όρους
- ο αριθμός των ξεχωριστών σημασιολογικών οντοτήτων που μπορούν να έχουν μια κοινή κειμενική περιγραφή δεν είναι μεγάλος
- το κέρδος στην ποιότητα του σημασιολογικού περιεχομένου του ερμηνευμένου χαρακτηρισμού, όπως υποδεικνύεται από τη διαφορά στην ακρίβεια της απάντησης του εκάστοτε συστήματος, είναι κατά ένα μεγάλο μέρος σημαντικότερο από τον προστιθέμενο υπολογιστικό φόρτο

Πηγαίνοντας ένα βήμα παρακάτω, μια πιθανή διαδικασία επέκτασης του χαρακτηρισμού των όρων του κειμένου εμπλουτίζει τον σημασιολογικό χαρακτηρισμό, προκειμένου να αυξηθεί η πιθανότητα της αντιστοιχίας μεταξύ του χαρακτηρισμού και του ευρετηρίου εγγράφων. Η παρουσία διαφόρων σημασιολογικών οντοτήτων στο χαρακτηρισμό καθορίζει κάποιο ταξινομικό εννοιολογικό πλαίσιο, το οποίο μπορούμε να χρησιμοποιήσουμε, προκειμένου να κατευθυνθεί η διαδικασία επέκτασης, έτσι ώστε να παράγει επεκτεταμένους χαρακτηρισμούς που να παρέχουν ενισχυμένη ανάκληση (recall) στο αποτέλεσμα, χωρίς να υποστούμε την παρενέργεια της φτωχής ακρίβειας (precision). Βέβαια, σε μία τέτοια περίπτωση τα βέλτιστα αποτελέσματα προκύπτουν μόνο, εάν η προέλευση των νέων οντοτήτων στον επεκταθέντα χαρακτηρισμό είναι γνωστή. Με άλλα λόγια, θα πρέπει να ξέρουμε σε ποια οντότητα αντιστοιχεί κάθε καινούρια οντότητα στον αρχικό χαρακτηρισμό.

#### 4.4.5 Σημασιολογικό ευρετήριο

Σε ένα απλό έγγραφο κειμένου το θέμα του περιεχομένου και η σημασιολογία που σχετίζεται με αυτό είναι δυνατόν να καθορίζονται ρητά ή τουλάχιστον να περιέχονται μέσα στο ίδιο το κείμενο ως λέξεις. Το πρόβλημα της ανάλυσης περιεχομένου ενός πολυμεσικού ή μη εγγράφου γενικότερα είναι ελαφρώς διαφορετικό και γι' αυτό απαιτεί και διαφορετικού είδους μεταχείριση, καθώς είναι αρκετά πιο περίπλοκο. Αφενός, οι οντότητες που θα μπουν στο ευρετήριο δεν απαντώνται κατευθείαν μέσα στο έγγραφο, αλλά τα αναγνωρίσιμα χαρακτηριστικά πρέπει να εξαχθούν και να ταιριάζουν σε αυτά που βρίσκονται μέσα στην εκάστοτε βάση γνώσης. Αφετέρου, ένα πολυμεσικό έγγραφο περιέχει εν γένει αντικείμενα και γεγονότα, των οποίων η αλληλοσυσχέτιση είναι κυρίως χωροχρονική, αντί για αυστηρά γραμματική. Τέλος, αφηρημένες έννοιες, όπως “τα σπορ” ή “η τέχνη” δεν εμφανίζονται κατευθείαν μέσα στα πολυμεσικά έγγραφα, αλλά πρέπει να τις συμπεράνουμε από τα συγκεκριμένα αντικείμενα και γεγονότα, καθώς και από χαρακτηριστικά (όπως π.χ. ο φωτισμός) που δεν αποδίδονται αποκλειστικά σε ένα ιδιαίτερο αντικείμενο ή γεγονός.

Μέχρι τώρα υποθέσαμε ότι η περιγραφή ενός εγγράφου (είτε αυτή εκφράζεται μέσω όρων, στην απλή περίπτωση των εγγράφων κειμένου, είτε μέσω σημασιολογικών εννοιών, στην πιο σύνθετη περίπτωση των πολυμεσικών εγγράφων) περιέχει οντότητες ίσης σημασίας. Μία πιο ελαστική, ωστόσο και πιο ακριβής προσέγγιση είναι να εισάγουμε βάρη στην σημαντικότητα των όρων περιγραφής, δηλ. να τοποθετήσουμε έναν βαθμό σημαντικότητας σε κάθε έναν από αυτούς. Στην περίπτωση ενός εγγράφου κειμένου ένας αυστηρά σημασιολογικός βαθμός σημαντικότητας απαιτεί την κατανόηση της φυσικής γλώσσας, πράγμα όχι πάντοτε εφικτό. Πιο απλές τεχνικές για την εύρεση των βαρών των όρων του ευρετηρίου περιλαμβάνουν συνήθως την στατιστική ανάλυση των όρων, όπως μπορεί να παρατηρήσει κανείς ανατρέχοντας στη σχετική βιβλιογραφία. Από την άλλη πλευρά, ένα πολυμεσικό έγγραφο παρέχει κάποια στοιχεία τα οποία είναι χρήσιμα για την εξαγωγή των αντικειμένων που μας ενδιαφέρουν. Τέτοια στοιχεία είναι π.χ. η κίνηση (π.χ. ένα άτομο που κινείται είναι πιο σημαντικό για ένα πλάνο απ' ότι ένα ακίνητο τραπέζι), το βάθος (ένα αντικείμενο στο προσκήνιο είναι στη γενική περίπτωση πιο σημαντικό από ένα αντικείμενο στο παρασκήνιο) και, εφόσον υπάρχει, το εννοιολογικό πλαίσιο (για παράδειγμα, είναι σημαντικό να γνωρίζουμε αν πρόκειται για ένα πλάνο εσωτερικού ή εξωτερικού χώρου, είτε αυτό είναι στη θάλασσα, είτε σε ένα λιβάδι).

Επίσης, μια διαφορετική πηγή βαρών είναι και η αβεβαιότητα. Στη γενική περίπτωση, η αβεβαιότητα παίζει έναν μικρό ρόλο κατά την δημιουργία ενός ευρετηρίου κειμένου και μετράται με τεχνικές όπως το stemming<sup>2</sup>. Κατά την ανάκτηση κειμένου, όμως, η αβεβαιότητα υπεισέρχεται μαζί με το εννοιολογικό πλαίσιο. Για παράδειγμα, η ίδια ακριβώς λέξη μπορεί να έχει πολύ διαφορετική σημασία και επεξήγηση, ανάλογα με το εννοιολογικό πλαίσιο και το περιβάλλον μέσα στα οποία χρησιμοποιείται. Παρόλο που η σημαντικότητα και η αβεβαιότητα είναι εν γένει δύο πολύ διαφορετικά είδη πληροφορίας, αλληλοσυσχέτιζονται, καθώς ένα αντικείμενο που θεωρείται σημαντικό για μία σκηνή (όπως π.χ. ένα άτομο στο προσκήνιο), δεν αναγνωρίζεται πάντοτε με μεγαλύτερη βεβαιότητα από ένα που θεωρείται ασήμαντο. Κατά συνέπεια και λαμβάνοντας υπόψη όλα τα παραπάνω, στα πλαίσια της παρούσης θα χρησιμοποιούμε μόνο ένα είδος βαθμού για τη στάθμιση των οντοτήτων. Ειδικότερα, έστω ότι ένα έγγραφο περιγρά-

<sup>2</sup>To stemming είναι η διαδικασία με την οποία μια κλινόμενη ή και παραγόμενη λέξη ελαχιστοποιείται στο κυρίως μόρφημά της.

φεται μέσω του σημασιολογικού ευρετήριου  $I$  ως ένα ασαφές σύνολο σημασιολογικών οντοτήτων.

$$I : S \times D \rightarrow [0, 1] \quad (4.36)$$

Το σημασιολογικό ευρετήριο πρέπει να είναι κανονικό για κάθε έγγραφο, δηλ.:

$$\forall d \in D \quad \exists s \in S : I(s, d) = 1 \quad (4.37)$$

όπου το  $D$  είναι το σύνολο των εγγράφων και  $d \in D$  είναι ένα έγγραφο.

Το σημασιολογικό ευρετήριο μπορεί να παραχθεί και σε ένα πολυμεσικό έγγραφο μέσω της αυτόματης αναγνώρισης αντικειμένων και γεγονότων και αντιστοίχησής τους σε σημασιολογικές οντότητες. Το τελευταίο, όμως, παραμένει σαφώς ένα ανοικτό και πολύπλοκο πρόβλημα. Εν τούτοις, παρόμοια αποτελέσματα μπορούν να προκύψουν από την ανάλυση των δομημένων πληροφοριών του κειμένου που περιλαμβάνονται στα μεταδεδομένα, τα οποία συνοδεύουν τα χαρακτηρισμένα πολυμεσικά έγγραφα (και σε συμφωνία με τη διαδικασία που περιγράφηκε στην υποενότητα 4.4.4).

#### 4.4.6 Ασαφής ιεραρχική συσταδοποίηση εγγράφων

Προκειμένου να πραγματοποιήσουμε επιτυχώς την ανίχνευση των κατηγοριών που σχετίζονται με ένα έγγραφο  $d$ , πρέπει να συσταδοποιήσουμε το σύνολο των σημασιολογικών οντοτήτων που σχετίζονται με το έγγραφο, σύμφωνα με την κοινή έννοιά τους. Πιο συγκεκριμένα, το σύνολο προς συσταδοποίηση αποτελείται από την υποστήριξη του εγγράφου:

$$S_d = {}^0d = \{s \in S : I(s, d) > 0\} \quad (4.38)$$

όπου το  $I$  αναπαριστά, όπως είδαμε, το σημασιολογικό ευρετήριο και αποτελεί μία ασαφή σχέση μεταξύ των εγγράφων  $d$  και των εννοιών  $s$ , δηλαδή το  $I(s, d)$  αναπαριστά το βαθμό ιδιότητας μέλους της έννοιας  $s$  στο έγγραφο  $d$ . Θεωρώντας το  $G_d$  ως το σύνολο των συστάδων που αναγνωρίζονται στο  $d$ , τότε κάθε συστάδα  $c \in G_d$  είναι ένα κλασικό (σαφές) σύνολο από έννοιες,  $c \subseteq S_d$ . Παρ' όλα αυτά, κάτι τέτοιο από μόνο του δεν είναι αρκετό για την προσέγγισή μας, καθώς εν γένει καλούμαστε να υποστηρίξουμε και περιπτώσεις εγγράφων, που μπορεί να ανήκουν σε πολλαπλές και διαφορετικές μεταξύ τους θεματικές περιοχές και με ποικίλους βαθμούς, αλλά ταυτόχρονα να διατηρήσουμε την αποδοτικότητα και ευρωστία που χαρακτηρίζει την ιεραρχική συσταδοποίηση. Κατά συνέπεια, χωρίς απώλεια της λειτουργικότητας ή επιπλέον υπολογιστικό κόστος αντικαθιστούμε τις κλασικές συστάδες  $c$  με ασαφείς κανονικοποιημένες συστάδες  $c^n$ . Κατόπιν, αθροίζοντας το εννοιολογικό πλαίσιο κάθε συστάδας  $c^n$ , προσδιορίζουμε το ασαφές σύνολο  $W_d$  των θεματικών περιοχών που σχετίζονται με το έγγραφο  $d$ . Στη συνέχεια αυτής της υποενότητας θα προχωρήσουμε σε αναλυτική παρουσίαση των αρχικών βημάτων της συσταδοποίησης εννοιών, της διαδικασίας “ασαφοποίησης” των συστάδων που προκύπτουν, καθώς επίσης και της τελικής ταξινόμησης ενός εγγράφου με βάση τη θεματική περιοχή με την οποία αυτό σχετίζεται.

Όπως έχουμε εξηγήσει ήδη, οι περισσότερες μέθοδοι συσταδοποίησης ανήκουν σε μία εκ των δύο γενικών κατηγοριών: καταμερισμού και ιεραρχικές [251]. Οι μέθοδοι καταμερισμού δημιουργούν μία σαφή ή ασαφή συσταδοποίηση ενός δεδομένου συνόλου στοιχείων, αλλά απαιτούν τον αριθμό των ομάδων ως είσοδο. Δεδομένου ότι ο αριθμός των κατηγοριών που υπάρχουν σε ένα έγγραφο δεν είναι γνωστός εκ των

#### Κεφάλαιο 4. Εξόρυξη δεδομένων και κατηγοριοποίηση

προτέρων, οι μέθοδοι αυτές είναι μη εφαρμόσιμες για το συγκεκριμένο πρόβλημα [177]. Κατά συνέπεια, πρέπει να εφαρμοστεί ένας ιεραρχικός αλγόριθμος συσταδοποίησης.

Η γενική δομή του αλγορίθμου, κατάλληλα προσαρμοσμένη στις ανάγκες του προβλήματος που θέλουμε να επιλυθεί στα πλαίσια της ενότητας αυτής, ακολουθεί τα εξής βήματα:

1. Κατά την εξέταση του εγγράφου  $d$ , μετατρέπουμε κάθε σημασιολογική οντότητα  $s \in {}^0+d$  σε μονοσύνολο (singleton), δηλ. σε μία αυτόνομη συστάδα  $c$ .
2. Για κάθε ζεύγος συστάδων  $c_1, c_2$  υπολογίζουμε ένα δείκτη συμβατότητας  $CI(c_1, c_2)$ . Ο  $CI$  αναφέρεται επίσης και ως μετρική ομοιότητας ή ανομοιότητας συστάδων.
3. Συγχωνεύουμε το ζευγάρι των συστάδων που έχουν τον καλύτερο  $CI$ . Ανάλογα με το εάν αυτός είναι μία μετρική ομοιότητας ή ανομοιότητας, ο καλύτερος δείκτης θα μπορούσε να επιλεγεί χρησιμοποιώντας το  $\max$  ή το  $\min$ , αντίστοιχα.
4. Συνεχίζουμε στο βήμα 2, έως ότου ικανοποιηθεί το κριτήριο τερματισμού. Το κριτήριο τερματισμού που χρησιμοποιείται συνήθως είναι ο καθορισμός ενός κατωφλίου για την τιμή του καλύτερου δείκτη συμβατότητας.

Τα δύο βασικά σημεία στον παραπάνω αλγόριθμο είναι ο προσδιορισμός των συστάδων προς συγχώνευση σε κάθε βήμα, δηλ. ο καθορισμός μίας μετρικής με νόημα για το  $CI$  και ο προσδιορισμός του βέλτιστου βήματος τερματισμού, δηλ. ο καθορισμός ενός κριτηρίου τερματισμού που να έχει κάποιο νόημα. Κατά την συσταδοποίηση σημασιολογικών οντοτήτων η ιδανική μετρική απόσταση για δύο συστάδες  $c_1, c_2$  είναι αυτή που ποσοτικοποιεί το σημασιολογικό συσχετισμό τους. Μία τέτοια μετρική μπορεί να ορισθεί ως η ένταση των κοινών εννοιολογικών πλαισίων γνώσης τους (βλ. ενότητα 4.4.3). Η διαδικασία της συγχώνευσης πρέπει να ολοκληρωθεί, όταν συσταδοποιούνται οι οντότητες στα σύνολα που αντιστοιχούν στις διαφορετικές κατηγορίες. Μπορούμε να προσδιορίσουμε τέτοια σύνολα από το γεγονός ότι τα κοινά εννοιολογικά πλαισία γνώσης τους θα έχουν χαμηλή ένταση. Επομένως, το κριτήριο τερματισμού θα είναι ένα κατώτατο όριο στην επιλεγμένη μετρική συμβατότητας. Η έξοδος αποτελείται από ένα σύνολο συστάδων  $G_d$ , όπου κάθε συστάδα  $c \in G_d$  αποτελεί ένα κλασικό σύνολο από έννοιες:  $c \subseteq S_d$ .

Αυτή η μέθοδος συσταδοποίησης, όντας ιεραρχική, θα καθορίσει επιτυχώς το πλήθος των διαφορετικών ομάδων που υπάρχουν στο  $S_d$ . Παρ' όλα αυτά θεωρείται κατώτερη από τις αντίστοιχες προσεγγίσεις καταμερισμού ως προς τις ακόλουθες παρατηρήσεις:

- Δημιουργεί μόνο κλασικές συστάδες, δηλ. δεν επιτρέπει την εξαγωγή βαθμών ιδιότητας μέλους στην έξοδο.
- Δημιουργεί μόνο καταμερισμούς, δηλ. δεν επιτρέπει επικαλύψεις μεταξύ των συστάδων που προκύπτουν.

Τα παραπάνω είναι δύο μεγάλα μειονεκτήματα για το συγκεκριμένο πρόβλημα, δεδομένου ότι δεν είναι συμβατά με τη σημασιολογία του στόχου: στην πραγματική ζωή μια σημασιολογική οντότητα μπορεί να σχετίζεται με μία κατηγορία με έναν βαθμό διαφορετικό από 0 ή 1 και μπορεί επίσης να σχετίζεται με περισσότερες της μιας διαφορετικές κατηγορίες. Προκειμένου να υπερνικηθούν τέτοιους είδους προβλήματα, περιγράφουμε στη συνέχεια μια μέθοδο για την ασαφοποίηση της διαμέρισης. Με τον

τρόπο αυτό, οι κλιμακωτές πληθικότητες των συστάδων θα διορθωθούν, έτσι ώστε να μπορούν να χρησιμοποιηθούν αργότερα για το φιλτράρισμα των παραπλανητικών οντοτήτων.

Όπως είδαμε παραπάνω, κάθε συστάδα περιγράφεται από το σαφές σύνολο των σημασιολογικών οντοτήτων  $\{s\} \subseteq S_d$  που ανήκουν σε αυτή. Χρησιμοποιώντας τις έννοιες αυτές, μπορούμε να κατασκευάσουμε έναν ασαφή ταξινομητή, δηλ. μια συνάρτηση  $C_c : S \rightarrow [0, 1]$  που μετρά το βαθμό συσχέτισης μιας έννοιας  $s$  με την συστάδα  $c$ . Προφανώς, μια έννοια  $s$  πρέπει να θεωρηθεί συσχετισμένη με τη συστάδα  $c$ , εάν σχετίζεται με την κοινή ερμηνεία (common meaning) των έννοιων στην  $c$ . Για το λόγο αυτό η ποσότητα  $C_1(c, s) = h(K(c \cup \{s\}))$  αποτελεί ένα κατάλληλο μέτρο της συσχέτισης. Φυσικά, δεν είναι όλες οι συστάδες εξίσου συμπαγείς. Εν γένει, μπορούμε να μετρήσουμε την πυκνότητα μιας συστάδας, χρησιμοποιώντας την ομοιότητα μεταξύ των έννοιων που αυτή περιέχει, δηλαδή χρησιμοποιώντας την ένταση του εννοιολογικού πλαισίου της συστάδας. Επομένως, το ανωτέρω μέτρο της συσχέτισης θα πρέπει να αναπροσαρμοστεί στα χαρακτηριστικά της εκάστοτε υπό εξέτασης συστάδας:

$$C_c(s) = \frac{C_1(c, s)}{h(K(c))} = \frac{h(K(c \cup \{s\}))}{h(K(c))} \quad (4.39)$$

Είναι εύκολο να παρατηρήσει κανείς ότι το παραπάνω μέτρο συσχέτισης διέπεται από τις εξής ιδιότητες:

- $C_c(s) = 1$ , αν η σημασιολογία της  $s$  υπαινίσσεται ότι αυτή ανήκει στην  $c$ . Για παράδειγμα:  $C_c(s) = 1, \forall s \in c$
- $C_c(s) = 0$ , αν η σημασιολογία της  $s$  υπαινίσσεται ότι αυτή δεν ανήκει στην  $c$ .
- $C_c(s) \in (0, 1)$ , αν η  $s$  σχετίζεται, αλλά και δε σχετίζεται με την  $c$ !

Χρησιμοποιώντας αυτόν τον ταξινομητή, μπορούμε να επεκτείνουμε τις ανιχνευμένες κλασικές συστάδες, έτσι ώστε να συμπεριλάβουν περισσότερες έννοιες. Για το σκοπό αυτό η συστάδα  $c$  αντικαθίσταται από την ασαφή συστάδα:  $c^f \supseteq c : c^f = \sum_{s \in S_d} s / C_c(s)$ , χρησιμοποιώντας τον ευρύτατα διαδεδομένο συμβολισμό αθροίσματος (sum notation) [177] για την αναπαράσταση των ασαφών συνόλων.

Θα πρέπει να έχει γίνει σαφές ότι η έως τώρα διαδικασία της ασαφούς ιεραρχικής συσταδοποίησης βασίστηκε στα κλασικά σύνολα, αγνοώντας την ύπαρξη της ασάφειας στο σημασιολογικό ευρετήριο. Προκειμένου να ενσωματωθεί η τελευταία κατά τον υπολογισμό των συστάδων που περιγράφουν το περιεχόμενο ενός εγγράφου, προχωράμε στη ρύθμιση των βαθμών ιδιότητας μέλους των συστάδων ως εξής:

$$c^i(s) = t(c^f(s), I(s, d)), \forall s \in S_d \quad (4.40)$$

όπου  $t$  είναι μία  $t$ -νόρμα. Η σημασιολογική φύση αυτής της λειτουργίας απαιτεί η  $t$  να είναι μία Αρχιμήδεια νόρμα<sup>3</sup>. Καθεμία από τις προκύπτουσες συστάδες αντιστοιχεί σε ένα από τα ευδιάκριτα θέματα του εγγράφου. Προκειμένου να καθοριστούν τα θέματα που σχετίζονται με μια συστάδα, δύο πράγματα πρέπει να ληφθούν υπόψη: η

<sup>3</sup>Μία  $t$ -νόρμα καλείται Αρχιμήδεια, αν το 0 και το 1 είναι οι μόνοι “ανεπηρέαστοι όροι”. Ένας “ανεπηρέαστος όρος” είναι κάτι το οποίο - δεδομένου ενός δυαδικού τελεστή όπως αυτός που χρησιμοποιείται εδώ - όταν πολλαπλασιαστεί με (ή στην περίπτωση μιας συνάρτησης όταν συντεθεί με) τον εαυτό του, παράγει τον ίδιο του τον εαυτό ως αποτέλεσμα.

#### Κεφάλαιο 4. Εξόρυξη δεδομένων και κατηγοριοποίηση

Βαθμωτή πληθικότητα της συστάδας  $|c^i|$  και το εννοιολογικό πλαίσιο της. Δεδομένου ότι το εννοιολογικό πλαίσιο έχει καθοριστεί μόνο ως προς κανονικά ασαφή σύνολα (βλ. ενότητα 3.2), θα πρέπει να ομαλοποιήσουμε τη συστάδα ως εξής:

$$c^n(s) = \frac{c^i(s)}{h(c^i(s))}, \forall s \in S_d \quad (4.41)$$

Το πρώτο βήμα, προκειμένου να προσδιοριστεί το ασαφές σύνολο  $W_d$  των θεματικών κατηγοριών που σχετίζονται με το έγγραφο  $d$ , είναι ο υπολογισμός του  $W(c)$ , δηλαδή το σύνολο των θεματικών κατηγοριών που σχετίζονται με κάθε συστάδα  $c$ . Υπολογίζουμε αρχικά το  $W^*$ , το οποίο προέρχεται από την ομαλοποιημένη συστάδα  $c^n$  και αντιστοιχεί στην παραγόμενη έξοδο της διαδικασίας στην περίπτωση που δεν ληφθούν υπόψη οι πληθικότητες των συστάδων. Γενικά, οι έννοιες που δεν περιλαμβάνονται στο εννοιολογικό πλαίσιο του  $c^n$  δεν μπορούν να θεωρηθούν ότι σχετίζονται με τη θεματική κατηγορία της συστάδας. Επομένως:  $W(c) \subseteq W^*(c^n) = w(K(c^n))$ , όπου  $w$  είναι ένας ασαφής τροποποιητής (weak modifier). Οι τροποποιητές, οι οποίοι απαντώνται στη βιβλιογραφία και ως γλωσσικοί φράκτες (linguistic hedges) [139], χρησιμοποιούνται, εν γένει, για να ρυθμίσουν τις από μαθηματική άποψη υπολογισμένες τιμές, ώστε να ταιριάζουν με τα σημασιολογικά προσδοκώμενα αντίστοιχά τους.

Εάν οι έννοιες που δεικτοδοτούν το έγγραφο  $d$  είναι όλες συσταδοποιημένες σε μία και μοναδική συστάδα  $c^i$ , τότε η  $W_d = W^*(c^n)$  είναι μία κατάλληλη προσέγγιση. Από την άλλη, στην περίπτωση που ανιχνεύονται περισσότερες από μια συστάδες, τότε είναι επιτακτικό να εξεταστούν και οι πληθικότητες των συστάδων. Οι συστάδες με εξαιρετικά χαμηλό αριθμό στοιχείων συνόλου περιέχουν πιθανότατα μόνο παραπλανητικές έννοιες και επομένως πρέπει να αγνοηθούν κατά την εκτίμηση του  $W_d$ . Αντίθετα, οι συστάδες με υψηλό αριθμό στοιχείων συνόλου αντιστοιχούν σχεδόν με βεβαιότητα στα ευδιάκριτα θέματα που σχετίζονται με το  $d$  και πρέπει να εξεταστούν κατά την εκτίμηση του  $W_d$ . Η έννοια του “υψηλού αριθμού στοιχείων συνόλου” ή της μεγάλης πληθικότητας μοντελοποιείται με τη χρήση ενός “μεγάλου” ασαφούς αριθμού  $L(\cdot)$ , ο οποίος αποτελεί μια συνάρτηση από το σύνολο των πραγματικών θετικών αριθμών στο διάστημα  $[0, 1]$ , ποσοτικοποιώντας την έννοια του “μεγάλου” ή “υψηλού”. Οι θεματικές κατηγορίες που σχετίζονται με κάθε συστάδα υπολογίζονται, αφού ρυθμιστούν οι βαθμοί ιδιότητας μέλους σύμφωνα με τις βαθμωτές πληθικότητες ως εξής:  $W(c) = W^*(c^n) \cdot L(|c^i|)$ .

Το σύνολο των θεματικών κατηγοριών που αντιστοιχούν σε ένα έγγραφο είναι το σύνολο των θεματικών κατηγοριών που ανήκουν σε κάποια από τις ανιχνευμένες συστάδες εννοιών, οι οποίες ευρετηριάζουν το δοσμένο έγγραφο:  $W_d = \bigcup_{c \in G_d} W(c)$ , όπου  $\bigcup$  είναι μία ασαφής co-νόρμα και  $G_d$  είναι το σύνολο των συστάδων που ανιχνεύθηκαν στο  $d$ . Είναι σχετικά εύκολο να παρατηρήσει κανείς ότι το  $W_d(s)$  θα είναι υψηλό, αν ανιχνευθεί μία συστάδα  $c^i$  στο  $d$ , της οποίας το εννοιολογικό πλαίσιο περιέχει την  $s$  και επιπρόσθετα, αν η πληθικότητα της  $c^i$  είναι μεγάλη και ο βαθμός ιδιότητας μέλους της  $s$  στο εννοιολογικό πλαίσιο της συστάδας είναι επίσης υψηλός (δηλαδή εάν η θεματική κατηγορία σχετίζεται με την συστάδα και η συστάδα δεν αποτελείται από παραπλανητικές έννοιες). Τέλος, προκειμένου να επικυρωθούν τα αποτελέσματα της ασαφούς ταξινόμησης, δηλ. προκειμένου να βεβαιώσουμε ότι το σύνολο των θεματικών κατηγοριών  $W_d$  που αντιστοιχούν σε ένα έγγραφο  $d$  προέρχεται από ένα γνωστό σύνολο όλων των πιθανών θεματικών κατηγοριών  $Y$ , υπολογίζουμε την ποσότητα  $Z_d = W_d \cap Y$ . Το σύνολο  $Z_d$  αποτελεί τελικά το ζητούμενο σύνολο των θεματικών

κατηγοριών του εγγράφου *d*.

#### 4.4.7 Αξιοποίηση υπηρεσιών ασαφούς εξαγωγής συμπερασμάτων

Όπως διαπιστώσαμε, η κατηγοριοποίηση πολυμεσικών εγγράφων αποτελεί ένα γνωστό, ενδιαφέρον και ιδιαίτερο πρόβλημα στο ερευνητικό πεδίο της ανάκτησης πληροφορίας. Εξάλλου, όπως παρουσιάσαμε θεωρητικά στις προηγούμενες ενότητες και όπως θα παρουσιάσουμε και αναλυτικά με παραδείγματα στην επόμενη υποενότητα 4.5.1, η γενική αρχή που ακολουθείται είναι η αντιστοίχηση ενός πολυμεσικού εγγράφου σε μία ή περισσότερες κατηγορίες σύμφωνα με τα περιεχόμενά του. Προκειμένου, δε, να πετύχουμε μια αποδοτική διαχείριση και θεματική κατηγοριοποίηση των εγγράφων, απαιτείται συνήθως η εξαγωγή της υποβόσκουσας σημασιολογίας από αυτά. Στην παρούσα ενότητα παρουσιάζουμε μια βελτιωμένη εκδοχή της πρότασής μας, η οποία αξιοποιεί ως δεδομένο, αναλύει και εν τέλει εκμεταλλεύεται τον χαρακτηρισμό (με τη μορφή κειμένου) που συνοδεύει ένα πολυμεσικό έγγραφο, προκειμένου να εξάγει τη σημασιολογία του, να κατασκευάσει κατά τα γνωστά ένα σημασιολογικό ευρετήριο και τελικά να κατηγοριοποιήσει τα έγγραφα σε θεματικές κατηγορίες. Η διαδικασία αυτή βασίζεται στο πρωτότυπο ενοποιημένο μοντέλο αναπαράστασης γνώσης και σημασιολογίας, που προτείνεται στα πλαίσια της διατριβής αυτής, καθώς και σε ορισμένες από τις βασικές αρχές της ασαφούς σχεσιακής άλγεβρας. Επιπρόσθετα, μια ασαφής επέκταση της εκφραστικής γλώσσας περιγραφής λογικής *SHIN*, η *f-SHIN* και οι υπηρεσίες εξαγωγής συμπερασμάτων που τη συνοδεύουν χρησιμοποιούνται για την περαιτέρω εκλέπτυνση και βελτιστοποίηση των αρχικών αποτελεσμάτων της κατηγοριοποίησης.

##### 4.4.7.1 Αναπαράσταση γνώσης και συμβολισμοί

Το μοντέλο γνώσης που χρησιμοποιήθηκε είναι διττό και βασίζεται αφ' ενός σε ένα σύνολο από έννοιες και σχέσεις μεταξύ των εννοιών και αφ' ετέρου σε μια οικογένεια από βασισμένους στη λογική φορμαλισμούς αναπαράστασης γνώσης. Οι τελευταίοι έχουν σχεδιαστεί έτσι, ώστε να αναπαριστούν και να εξάγουν συμπεράσματα σχετικά με τη γνώση μιας θεματικής περιοχής με έναν δομημένο και πλήρως κατανοητό τρόπο. Αυτή η αναπαράσταση γνώσης επιτρέπει τη δημιουργία μιας λεπτομερούς περιγραφής όλων των ειδών εγγράφων και των περιεχομένων τους με έναν μοναδικό και ενοποιημένο τρόπο. Στην περίπτωση των ακίνητων ή και κινούμενων εικόνων η γνώση αυτή κατασκευάζεται έτσι, ώστε να ταιριάζει με τους συγκεκριμένους κειμενικούς χαρακτηρισμούς των αντίστοιχων πολυμεσικών εγγράφων.

Το πρώτο μέρος του προτεινόμενου μοντέλου γνώσης αποτελείται από ένα σύνολο από έννοιες και τις σχέσεις που τις διέπουν. Όπως έχουμε παρατηρήσει πολλάκις στα πλαίσια αυτής της διατριβής, οι περισσότερες σχέσεις μεταξύ πραγματικών εννοιών εμπεριέχουν συνήθως κάποιο βαθμό εμπιστοσύνης. Κατά συνέπεια, αποτελεί πεποίθησή μας ότι τέτοιου είδους σχέσεις μοντελοποιούνται καλύτερα με τη χρήση ασαφών σχέσεων. Για το λόγο αυτό ακολουθούμε και σε αυτή την ερευνητική περιπλάνηση την αυστηρή μεθοδολογία και τους μαθηματικούς συμβολισμούς που εισήχθησαν από τον Klir [139] και τον [177] και περιγράφονται αναλυτικά στην υποενότητα 3.2 της παρούσης.

Όπως περιγράφει ο Salembier [212], οι προσεγγίσεις που στο παρελθόν βασίστηκαν αποκλειστικά σε λεξικογραφικούς όρους πάσχουν από την προβληματική αντιστοίχηση των όρων σε έννοιες. Ειδικότερα, καθώς παραπάνω του ενός όροι είναι δυνατόν

#### Κεφάλαιο 4. Εξόρυξη δεδομένων και κατηγοριοποίηση

να σχετίζονται με την ίδια έννοια και περισσότερες από μία έννοιες μπορεί να σχετίζονται με τον ίδιο λεξικογραφικό όρο, η επεξεργασία του εκάστοτε ερωτήματος και της πληροφορίας ευρετηρίου δεν είναι εύκολη. Προκειμένου να αντιμετωπίσουμε τέτοιους είδους προβλήματα, εμείς δουλεύουμε αποκλειστικά με έννοιες και όχι με όρους. Συνεπώς, στη συνέχεια, ορίζουμε το κλασικό (σαφές) σύνολο των γνωστών εννοιών ως:  $S = \{s_1, s_2, \dots, s_n\}$ . Εν γένει, ένα μοντέλο αναπαράστασης γνώσης μπορεί να αποτελείται από τους ορισμούς των παραπάνω εννοιών μαζί με τις κειμενικές περιγραφές τους (δηλαδή τους αντίστοιχους λεξικογραφικούς όρους), καθώς και ένα σύνολο από σχέσεις μεταξύ των εννοιών αυτών. Ο απώτερος στόχος είναι η κατασκευή ενός μοντέλου, στο οποίο το εννοιολογικό πλαίσιο θα καθορίζει την εννοούμενη σημασία κάθε λεξικογραφικού όρου και όπου ένας λεξικογραφικός όρος που χρησιμοποιείται κάτω από διαφορετικό εννοιολογικό πλαίσιο μπορεί να έχει διαφορετικές ερμηνείες. Ο ορισμός ενός μοντέλου που ικανοποιεί τις απαιτήσεις αυτές δίδεται από την εξίσωση:

$$O = \{S, \{\mathcal{R}_i\}\}, \quad i = 1 \dots n \quad (4.42)$$

$$\mathcal{R}_i : S \times S \rightarrow \{0, 1\}, \quad i = 1 \dots n \quad (4.43)$$

όπου το  $O$  είναι το μοντέλο γνώσης και η  $\mathcal{R}_i$  αναπαριστά την  $i$ -στη δυαδική σχέση μεταξύ των εννοιών.

Στη γενική περίπτωση, μπορούμε να χρησιμοποιήσουμε οποιοδήποτε είδος σχέσεων για την κατασκευή του παραπάνω μοντέλου, όμως εδώ, για λόγους πληρότητας των περιγραφών, αλλά και λαμβάνοντας υπόψη την ασάφεια των σχέσεων, καταλήγουμε στον ακόλουθο μαθηματικό φορμαλισμό/ορισμό του μοντέλου μας, ως:

$$O_F = \{S, \{r_i\}\}, \quad i = 1 \dots n \quad (4.44)$$

$$r_i = F(\mathcal{R}_i) : S \times S \rightarrow [0, 1], \quad i = 1 \dots n \quad (4.45)$$

όπου το  $F$  καταδεικνύει την “ασαφοποίηση” των αρχικών σχέσεων  $\mathcal{R}_i$ . Σύμφωνα με τις σχέσεις  $r_i$ , κατασκευάζουμε κατά τα γνωστά το συνδυασμό:

$$T = \bigcup_i r_i^{p_i}, \quad p_i \in \{-1, 1\}, \quad i = 1 \dots n \quad (4.46)$$

όπου η τιμή του  $p_i$  καθορίζεται από τη σημασιολογία κάθε επιμέρους σχέσης που εμπλέκεται στην κατασκευή της  $T$  και του κατά πόσο αυτή έρχεται ή όχι σε αντιδιαστολή με την κοινή ερμηνεία της κάθε σχέσης για το συγκεκριμένο ζεύγος εννοιών υπό εξέταση. (Γι' αυτό και η σειρά των ορισμάτων  $x, y$  στον Πίνακα 4.1 παίζει ιδιαίτερο ρόλο). Όπως και προηγουμένως (βλ. υποενότητα 4.4.2), κάποιες σχέσεις θα πρέπει να αντιστραφούν, πριν χρησιμοποιηθούν για την κατασκευή της  $T$ . Η τελευταία κατασκευάζεται σύμφωνα με όσα περιγράφονται στα προηγούμενα, ως:

$$T = Tr^t(Sp \cup P^{-1} \cup Ins \cup Pr^{-1} \cup Pat \cup L \cup Ex) \quad (4.47)$$

Όπως παρατηρούμε στην παραπάνω εξίσωση, χρησιμοποιήθηκαν 7 ασαφείς σχέσεις για την κατασκευή της  $T$ , των οποίων η σημασιολογία καθορίζεται στο πρότυπο MPEG-7 και συνοψίζεται στον Πίνακα 4.1. Σύμφωνα με τη σημασιολογία των επιλεγμένων σχέσεων η  $T$  καθίσταται ιδανική για τον καθορισμό των θεμάτων, με τα οποία μπορεί να σχετίζεται μία έννοια, καθώς και για τον καθορισμό των θεματικών

Όνομα	Αντίστροφο	Σύμβολο	Ερμηνεία
Μέρος	Μέρος Του	$P(x,y)$	το $y$ είναι μέρος του $x$
Εξειδίκευση Γενίκευση		$Sp(x,y)$	το $x$ είναι μια γενίκευση του $y$
Παράδειγμα Παράδειγμα Του		$Ex(x,y)$	το $y$ είναι ένα παράδειγμα του $x$
'Οργανο	'Οργανο Του	$Ins(x,y)$	το $y$ είναι ένα όργανο του $x$
Θέση	Θέση Του	$L(x,y)$	το $y$ είναι η θέση του $x$
Ασθενής	Ασθενής Του	$Pat(x,y)$	το $y$ είναι ένας ασθενής του $x$
Ιδιότητα	Ιδιότητα Του	$Pr(x,y)$	το $y$ είναι μια ιδιότητα του $x$

**Πίνακας 4.1:** Ασαφείς ταξινομικές σχέσεις που χρησιμοποιούνται για τη δημιουργία του πρώτου μέρους του κατάλληλου μοντέλου γνώσης.

κατηγοριών που μπορούν να εντοπιστούν σε ένα δεδομένο σύνολο από έννοιες σε ένα έγγραφο.

Το δεύτερο μέρος του προτεινόμενου μοντέλου γνώσης βασίζεται στην περιγραφική λογική. Οι περιγραφικές λογικές (Description Logics - DLs) αποτελούν μια οικογένεια βασισμένων στη λογική γλωσσών αναπαράστασης γνώσης, οι οποίες μπορούν να χρησιμοποιηθούν, προκειμένου να αναπαρασταθεί η γνώση μιας θεματικής περιοχής με έναν δομημένο και ανθρωπίνως κατανοητό φορμαλισμό. Οι DLs κληρονομούν τις περιγραφές των εννοιών από τους προγόνους τους, τις γλώσσες περιγραφής (*description languages*), δηλαδή περιγράφουν σημαντικές έννοιες της θεματικής περιοχής χρησιμοποιώντας εκφράσεις που κατασκευάζονται από ατομικές έννοιες και ατομικούς ρόλους. Στη συνέχεια της ενότητας αυτής, θα παρουσιάσουμε εν συντομίᾳ τον συμβολισμό της  $f\text{-SHIN}$ , η οποία είναι μια ασαφής επέκταση της DL  $SHIN$  [123]. Μια γλώσσα περιγραφής αποτελείται από ένα αλφάριθμο διακεκριμένων ονομάτων εννοιών (**C**), ονομάτων ρόλων (**R**) και ονομάτων ατόμων (**I**), μαζί με ένα σύνολο από κατασκευαστές που κατασκευάζουν τις περιγραφές των εννοιών και των ρόλων. Π.χ., αν  $R$  είναι ένας ρόλος, τότε  $R^-$  είναι επίσης ένας ρόλος, ο οποίος εκφράζει το αντίστροφο του  $R$ . Οι  $f\text{-SHIN}$ -έννοιες καθορίζονται ματευτικά ως ακολούθως:

1. Αν  $C \in \mathbf{C}$ , τότε  $C$  είναι μια  $f\text{-SHIN}$ -έννοια.
2. Αν  $C$  και  $D$  είναι έννοιες, το  $R$  είναι ένας ρόλος και  $n \in \mathbb{N}$ , τότε  $(\neg C)$ ,  $(C \sqcup D)$ ,  $(C \sqcap D)$ ,  $(\forall R.C)$ ,  $(\exists R.C)$ ,  $(\geq nR)$  και  $(\leq nR)$  είναι επίσης  $f\text{-SHIN}$ -έννοιες.

Σε αντίθεση με τις κλασικές (σαφείς) DLs, η σημασιολογία των ασαφών DLs παρέχεται σύμφωνα με τους Straccia[246] και Stoilo[245] από μια ασαφή ερμηνεία. Μια ασαφής ερμηνεία είναι ένα ζευγάρι  $\mathcal{I} = \langle \Delta^{\mathcal{I}}, \cdot^{\mathcal{I}} \rangle$ , όπου η θεματική περιοχή  $\Delta^{\mathcal{I}}$  είναι ένα μη-κενό σύνολο από αντικείμενα και  $\cdot^{\mathcal{I}}$  είναι μία συνάρτηση ασαφούς ερμηνείας, η οποία αντιστοιχεί ένα ατομικό όνομα σε στοιχεία του  $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$  και ένα όνομα έννοιας  $A$  (όνομα ρόλου  $R$ ) σε μια συνάρτηση συμμετοχής  $A^{\mathcal{I}} : \Delta^{\mathcal{I}} \rightarrow [0, 1]$ .

Μια  $f\text{-SHIN}$  βάση γνώσης (KB) ορίζεται ως μία τριπλέτα  $\langle T, R, A \rangle$ , όπου το  $T$  είναι ένα ασαφές  $TBox$ , το  $R$  είναι ένα ασαφές  $RBox$  και το  $A$  είναι ένα ασαφές  $ABox$ . Το  $TBox$  είναι ένα πεπερασμένο σύνολο από ασαφή αξιώματα εννοιών, τα οποία έχουν είτε τη μορφή  $C \equiv D$ , οπότε ονομάζονται αξιώματα ισοδυναμίας ασαφών εννοιών (fuzzy concept equivalence axioms), είτε  $C \sqsubseteq D$ , οπότε ονομάζονται αξιώματα συμπερίληψης ασαφών εννοιών (fuzzy concept inclusion axioms), όπου

#### Κεφάλαιο 4. Εξόρυξη δεδομένων και κατηγοριοποίηση

$C, D$  είναι έννοιες που υποδηλώνουν ότι η  $C$  είναι ισοδύναμη ή μια υπο-έννοια της  $D$ , αντίστοιχα. Κατ' αναλογία, το  $RBox$  είναι ένα πεπερασμένο σύνολο από αξιώματα ασαφών ρόλων (fuzzy role axioms) της μορφής  $\text{Trans}(R)$ , τα οποία ονομάζονται αξιώματα ασαφών μεταβατικών ρόλων (fuzzy transitive role axioms), είτε  $R \sqsubseteq S$ , οπότε ονομάζονται αξιώματα συμπερίληψης ασαφών ρόλων (fuzzy role inclusion axioms), όπου ο  $R$  είναι μεταβατικός και ο  $R$  είναι ενας υπο-ρόλος του  $S$ , αντίστοιχα. Τέλος, το  $ABox$  είναι ένα πεπερασμένο σύνολο από ασαφείς ισχυρισμούς της μορφής  $\langle a : C \bowtie n \rangle, \langle (a, b) : R \bowtie m \rangle$ , όπου το  $\bowtie$  συμβολίζει το  $\geq, >, \leq, <$  ή  $a \neq b$ , για  $a, b \in \mathbf{I}$ . Διαισθητικά, μπορούμε να ισχυριστούμε ότι ένας ασαφής ισχυρισμός της μορφής  $\langle a : C \geq n \rangle$  σημαίνει ότι ο βαθμός εμπιστοσύνης του  $a$  στην έννοια  $C$  είναι τουλάχιστον  $n$  σημαίνει ότι ο βαθμός εμπιστοσύνης του  $a$  στην έννοια  $C$  είναι τουλάχιστον  $n$  με  $n$ . Τους ισχυρισμούς που ορίζονται με  $\geq, >$  τους καλούμε θετικούς ισχυρισμούς, ενώ αυτούς που ορίζονται με  $\leq, <$ , τους καλούμε αρνητικούς ισχυρισμούς [244].

Η κύρια υπηρεσία εξαγωγής συμπερασμάτων (reasoning service) για μία κλασική DL είναι η συνοχή της βάσης γνώσης (KB). Επιπρόσθετες υπηρεσίες εξαγωγής συμπερασμάτων είναι οι ακόλουθες: *ικανοποίηση έννοιας* (concept satisfiability), υπαγωγή έννοιας (concept subsumption) και *συνεπαγωγή* (entailment), οι οποίες όμως περιορίζονται στην συνοχή της βάσης γνώσης (KB). Προκειμένου να αποφασίσουμε για τη συνοχή της τελευταίας, υιοθετούμε έναν αλγόριθμο ο οποίος προσπαθεί να κατασκευάσει ένα μοντέλο της με το να αποδομεί τα βασικά δομικά στοιχεία της βάσης γνώσης και εξάγοντας με τον τρόπο αυτό νέους περιορισμούς για τα στοιχεία του μοντέλου. Οι παραπάνω υπηρεσίες εξαγωγής συμπερασμάτων είναι διαθέσιμες και από την  $f\text{-SHIN}$  μαζί με ερωτήματα μεγαλύτερου μικρότερου ορίου (greatest lower bound queries), τα οποία είναι αυτά που εκμεταλλεύονται την ασάφεια. Η μηχανή ασαφούς εξαγωγής συμπερασμάτων, που χρησιμοποιήθηκε για τα πειράματά μας, είναι το λογισμικό FiRE<sup>4</sup> που βασίζεται στην  $f\text{-SHIN}$ . Με άλλα λόγια, όλες οι υπηρεσίες εξαγωγής συμπερασμάτων μπορούν να αναπαρασταθούν ως ερωτήματα. Ως εκ τούτου, τα ερωτήματα ασαφούς συνεπαγωγής ρωτούν κατά πόσο ένα άτομο συμμετέχει σε μία έννοια και σε ποιό βαθμό, ενώ τα ερωτήματα υπαγωγής θέτουν ερωτήματα σχετικά με το κατά πόσο μία έννοια είναι υπο-έννοια μιας άλλης έννοιας (π.χ. *FootBallCategory*  $\sqsubseteq$  *Sports*). Τέλος, μιας και ένα ασαφές  $ABox$  μπορεί να περιέχει πολλούς θετικούς ισχυρισμούς για το ίδιο άτομο (ή για το ίδιο ζεύγος ατόμων) χωρίς να οδηγείται σε αντίφαση, θα μας ήταν ωφέλιμο να υπολογίσουμε ποια είναι τα καλύτερα όρια ενός ασαφούς ισχυρισμού και με δεδομένο ότι η έννοια του μεγαλύτερου μικρότερου ορίου ενός ασαφούς ισχυρισμού σύμφωνα με την βάση γνώσης (KB) ορίστηκε από τον Straccia [246] ως ο βαθμός συμμετοχής ενός ατόμου σε μία έννοια.

##### 4.4.7.2 Δημιουργία σημασιολογικού ευρετηρίου και κατηγοριοποίηση κειμένου

'Οπως παρουσιάστηκε αναλυτικά στην ενότητα 4.4.5, η ολοκληρωμένη διαχείριση των έννοιών και των σχέσεων που προέρχονται από τα κειμενικά μεταδεδομένα που συνοδεύουν κάθε πολυμεσικό έγγραφο είναι δυνατή μόνο με την αξιοποίηση και την κατασκευή ενός σημασιολογικού ευρετηρίου. Και στην περίπτωση του υπό εξέταση συνδυασμού της ασαφούς σχεσιακής άλγεβρας με τις παραπάνω υπηρεσίες εξαγω-

<sup>4</sup>Το λογισμικό FiRE είναι διαθέσιμο στην ιστοσελίδα <http://www.image.ece.ntua.gr/~nsimou/FiRE/>, μαζί με οδηγίες εγκατάστασης και παραδείγματα χρήσης.

γής συμπερασμάτων η διαδικασία αυτή βρίσκεται στο κέντρο της προσέγγισής μας και αποκτά ιδιαίτερη βαρύτητα. Βασίζεται, δε, στο διττό ασαφές μοντέλο γνώσης, που περιγράφηκε στην προηγούμενη ενότητα. Μετά τη δημιουργία του σημασιολογικού ευρετηρίου, τα πολυμεσικά έγγραφα κατηγοριοποιούνται σε θεματικές κατηγορίες (π.χ. αθλητικά, πολιτική, κ.λ.π..) μέσω της μεθοδολογίας ασαφούς συσταδοποίησης των εννοιών που σχετίζονται με κάθε έγγραφο. Το σημασιολογικό ευρετήριο που συνδέει τις εννοιες με τα πολυμεσικά έγγραφα κατασκευάζεται σε δύο βήματα. Αρχικά, όλα τα διαθέσιμα πολυμεσικά έγγραφα αντιστοιχίζονται σε εννοιες μαζί με ένα σχετικό βαθμό εμπιστοσύνης. Η αντιστοίχιση αυτή περιλαμβάνει τη σημασιολογική ερμηνεία των λεξικογραφικών όρων τους, που προέρχεται από τον συνοδευτικό κειμενικό χαρακτηρισμό. Σε ένα δεύτερο βήμα, το σημασιολογικό ευρετήριο (δηλαδή οι συσχετισμένες εννοιες) κάθε εγγράφου αναλύεται περαιτέρω, προκειμένου να εκτιμήσουμε το βαθμό, στον οποίο το δεδομένο έγγραφο σχετίζεται με κάθε ένα από τα προκαθορισμένα σύνολα από θεματικές κατηγορίες. Το τελευταίο επιτυγχάνεται με την εφαρμογή μιας κατάλληλης προσέγγισης συσταδοποίησης των εννοιών που σχετίζονται με ένα έγγραφο σύμφωνα με την κοινή τους ερμηνεία και το εννοιολογικό πλαίσιο, όπως ακριβώς περιγράφηκε στις υποενότητες 4.2.1-4.4.6.

#### 4.4.7.3 Η ασαφής βάση γνώσης

Σαν δεύτερο βήμα της προσέγγισής μας, μια ασαφής βάση γνώσεων (KB) χρησιμοποιείται για την κατηγοριοποίηση των πολυμεσικών εγγράφων. Η τελευταία αποτελείται από ένα ασαφές *TBox*, ένα ασαφές *RBox* και ένα ασαφές *ABox*. Τα *TBox* και *RBox* εισάγουν την ορολογία, δηλ. το λεξιλόγιο της θεματικής περιοχής, ενώ το *ABox* περιέχει τους ισχυρισμούς για τα ονομασμένα άτομα (named individuals) από την άποψη αυτού του λεξιλογίου. Όπως θα διευκρινιστεί περαιτέρω στην υποενότητα αυτή, η τρέχουσα θεματική περιοχή ενδιαφέροντός μας είναι οι ειδήσεις και τα πολυμεσικά της έγραφα μπορούν να κατηγοριοποιηθούν σε ποικίλες κατηγορίες σύμφωνα με το περιεχόμενό τους.

Η βασισμένη στη σημασιολογία επέκταση της απλής προσέγγισης κατηγοριοποίησης κειμένου που παρουσιάστηκε στην υποενότητα 4.4.7.2, αρχικοποιεί το *ABox* της ασαφούς βάσης γνώσης. Ο αλγόριθμος σημασιολογικής ευρετηρίασης αξιολογεί τα πολυμεσικά έγγραφα παράγοντας ένα σύνολο εννοιών, οι οποίες υπάρχουν στο εκάστοτε έγγραφο, μαζί με έναν βαθμό  $w \in [0, 1]$  που όσο πιο υψηλός είναι, τόσο περισσότερα συνώνυμα της συγκεκριμένης εννοιας απαντώνται στο έγγραφο. Αυτό το είδος πληροφορίας αναπαριστάται ως θετικοί ισχυρισμοί στην κατασκευασμένη ασαφή βάση γνώσης, χρησιμοποιώντας το έγγραφο ως το άτομο που συμμετέχει στην εννοια με βαθμό  $w$ . Μιας και ο αριθμός των χρησιμοποιούμενων εννοιών κατά τη διάρκεια του προηγούμενου βήματος της κατηγοριοποίησης κειμένου διαμορφώνει ένα εξαιρετικά μεγάλο λεξιλόγιο (περίπου 150.000 εννοιες που προέρχονται από το σύνολο δεδομένων Wordnet<sup>5</sup>), ήταν απαραίτητη μιας μορφής μείωσή του για τη δημιουργία του ασαφούς αλφαριθμητού γνώσης. Η μείωση αυτή υλοποιήθηκε επιλέγοντας ένα αντιπροσωπευτικό σύνολο εννοιών για κάθε κατηγορία. Το σύνολο των εννοιών που περιλαμβάνουν το αλφαριθμητο της ασαφούς βάσης γνώσης μας είναι το ακόλουθο:

Η αποτελεσματική εξαγωγή της υπονοούμενης γνώσης απαιτεί μια εκφραστική ορολογία, η οποία είναι σε θέση να καθορίσει υψηλότερες εννοιες. Οι κατηγορίες, στις οποίες κατηγοριοποιούνται τα έγγραφα, καθορίζονται από την ορολογία που πα-

<sup>5</sup><http://wordnet.princeton.edu/>

#### Κεφάλαιο 4. Εξόρυξη δεδομένων και κατηγοριοποίηση

Εννοιες = {γεγονός, γκολ, τερματοφύλακας, ποδόσφαιρο, χέρια, πέναλτι, ανταγωνισμός, πρόσθετος, γενικός, παρεμπόδιση, ομάδα, ακροατήριο, τύρφη, σύγχρονος, παίκτης, σήμερα, αέρας, αθλητικά, σκορ, επαφή, λάκτισμα, φάουλ, δοκάρι, λίγκα, νικητής, περιλαμβάνει, ποικίλλει, αριθμοί, ρόλος, ζημία, χρόνοι, αρχικοί, φανέλα, όπλα, άμυνα, συστολή, τοπικός, αντικειμενικός, αγαπημένος, ομιλία, ερασιτεχνικός, μεμονωμένος, επαγγελματικός, απλός, φθορά, συμπεριφορά, θηλυκό, αρσενικό, υποδήματα, σύνολο, πρωτάθλημα, χώρα, ομάδα, έθνος, έτος, δεκάλεπτο, σουτ, ριμπάουντ, πόντοι, ταιμ-αυτ, αόψιμο, άρρωστη, σκάψιμο, καρφί, σερβίς, σετ, άσσος, κυβέρνηση, νόμος, υπουργός, εκλογή, κοινοβούλιο, τιμή, κέρδος, τράπεζα, εμπορικός, αόστος, πτώση, αγορά, απώλεια, ποσοστό, άνοδος, φοιτητής, όρος, χρηματοδότηση, πανεπιστήμιο, ακαδημία, σχολείο, διοίκηση, συν, βασικός, χρήματα, προϋπολογισμός, δισεκατομμύριο, εβδομάδα, διακοπές, μήνας, επιλογή, δειγματοληψία, διορισμός, γενικός, γένος, αναλυτής, προϊστάμενος, μέρος, ταυτότητα, γεγονός, ανταγωνιστής, αποστασία, διαγωνισμός, αιχμηρός, ισχυρισμός, κοινό, πολιτική, οδήγηση, διοργανωτής}

Σχήμα 4.4: Αλφαριθμητικό εννοιών της ασαφούς βάσης γνώσης.

ρουσιάζεται στον Πίνακα 4.2. Όπως παρατηρούμε, οι κατηγορίες: *Εργασία, Εκπαίδευση και Πολιτική* είναι υποκατηγορίες της κατηγορίας *Ειδήσεις* και οι παρόμοιες κατηγορίες: *Ποδόσφαιρο Μπάσκετ* και *Βόλλεϋ* είναι υποκατηγορίες της κατηγορίας *Αθλητικά*.

Οι αθλητικές υποκατηγορίες καθορίστηκαν με παρόμοιο τρόπο. Τα αξιώματα καθορισμού τους απαιτούν ένα έγγραφο να αποτελείται από μερικές απαραίτητες έννοιες και μία από τις προαιρετικές έννοιες. Οι απαραίτητες έννοιες για κάθε ένα από αυτά είναι η έννοια *Ομάδα* μαζί με το χριτήριο νίκης (δηλ. την έννοια *Πόντοι* στην περίπτωση του μπάσκετ και του βόλλεϋ και *Γκολ* στην περίπτωση του ποδοσφαίρου). Οι προαιρετικές έννοιες αποτελούνται από τους ειδικούς όρους που χρησιμοποιούνται για κάθε αθλημα. Κάποιος μπορεί να αναρωτηθεί γιατί η έννοια *Ποδόσφαιρο* δεν είναι μια από τις απαραίτητες λέξεις-κλειδιά για την κατηγοριοποίηση ενός εγγράφου στην κατηγορία *Ποδόσφαιρο*. Η απάντηση είναι ότι πολλά έγγραφα μπορεί να περιγράφουν ένα ποδοσφαιρικό παιχνίδι, αλλά χωρίς απαραιτήτως να χρησιμοποιήσουν αυτήν την έννοια.

Από την άλλη μεριά, οι υποκατηγορίες ειδήσεων καθορίστηκαν με έναν πιο γενικό τρόπο λόγω κυρίως του αφηρημένου περιεχομένου τους. Ως εκ τούτου, οι κατηγορίες *Πολιτική* και *Εργασία* καθορίζονται από αξιώματα που απαιτούν την ύπαρξη δύο έννοιών. Κάθε μια από αυτές τις έννοιες ανήκει σε ένα ευρύτερο σύνολο λέξεων-κλειδιών που είναι σχετικές με τη θεματική κατηγορία, παρόλο που από μόνες τους είναι ανεπαρκείς. Ας θεωρήσουμε λ.χ. ένα έγγραφο που περιλαμβάνει την έννοια *Εκλογή*. Αυτή η έννοια δε συμπεριλαμβάνεται απαραιτήτως σε ένα έγγραφο που περιγράφει ένα πολιτικό γεγονός. Αυτίθετα, εάν αυτή η έννοια συμπεριλαμβάνεται σε ένα έγγραφο μαζί με μια από τις έννοιες *Κυβέρνηση, Νόμος, Υπουργός* και *Κοινοβούλιο*, τότε το έγγραφο αυτό περιγράφει ενδεχομένως ένα πολιτικό γεγονός. Αναμφίβολα, η κατηγορία *Εκπαίδευση* είναι λιγότερο περιγραφική από τις υπόλοιπες, δεδομένου ότι κάθε εμφάνιση κάποιων σχετικών με τη θεματική περιοχή έννοιών μέσα στο έγγραφο, ταξινομεί το έγγραφο αυτό σε αυτήν την κατηγορία.

---

$\mathcal{T} = \{\text{Μπασκετ} \equiv \text{Ομαδα} \sqcap \text{Ποντοι}$
$\sqcap (\Delta \text{εκαλεπτο} \sqcup \Sigma \text{ουτ} \sqcup \text{Ριμπαουντ}$
$\sqcup \text{Ταιμ} - \alpha \text{ουτ} \sqcup \text{Κοφιμο} \sqcup \text{Καρφωμα}),$
$\text{Βολλευ} \equiv \text{Ομαδα} \sqcap \text{Ποντοι}$
$\sqcap (\Sigma \text{καφιμο} \sqcup \text{Καρφι} \sqcup \Sigma \text{ερβις} \sqcup \Sigma \text{ετ} \sqcup \text{Κοφιμο} \sqcup \text{Ασσος}),$
$\text{Ποδοσφαιρο} \equiv \text{Ομαδα} \sqcap \text{Γκολ}$
$\sqcap (\text{Ποδοσφαιρο} \sqcup \text{Πεναλτι} \sqcup \text{Τερματοφυλακας} \sqcup \text{Σκορ}),$
$\text{Αθλητικα} \equiv \text{Ποδοσφαιρο}$
$\sqcup \text{Μπασκετ} \sqcup \text{Βολλευ},$
$\text{Πολιτικη} \equiv (\text{Κυβερνηση} \sqcup \text{Νομος} \sqcup \text{Υπουργος} \sqcup \text{Κοινοβουλιο})$
$\sqcap (\text{Κυβερνηση} \sqcup \text{Εκλογη} \sqcup \text{Κοινο}),$
$\text{Εργασια} \equiv (\text{Τιμη} \sqcup \text{Κερδος} \sqcup \text{Τραπεζα} \sqcup \text{Εμπορικος} \sqcup \text{Κοστος})$
$\sqcap (\text{Πτωση} \sqcup \text{Αγορα} \sqcup \text{Απωλεια} \sqcup \text{Ποσοστο} \sqcup \text{Ανοδος}),$
$\text{Εκπαιδευση} \equiv \text{Φοιτητης} \sqcup \text{Σχολειο} \sqcup \text{Ορος}$
$\sqcup \text{Χρηματοδοτηση} \sqcup \text{Ακαδημια} \sqcup \text{Πανεπιστημιο},$
$\text{Ειδησεις} \equiv \text{Πολιτικη} \sqcup \text{Εργασια} \sqcup \text{Εκπαιδευση},$

---

Πίνακας 4.2: Βάση Γνώσης (TBox)

## 4.5 Πειραματικά Αποτελέσματα

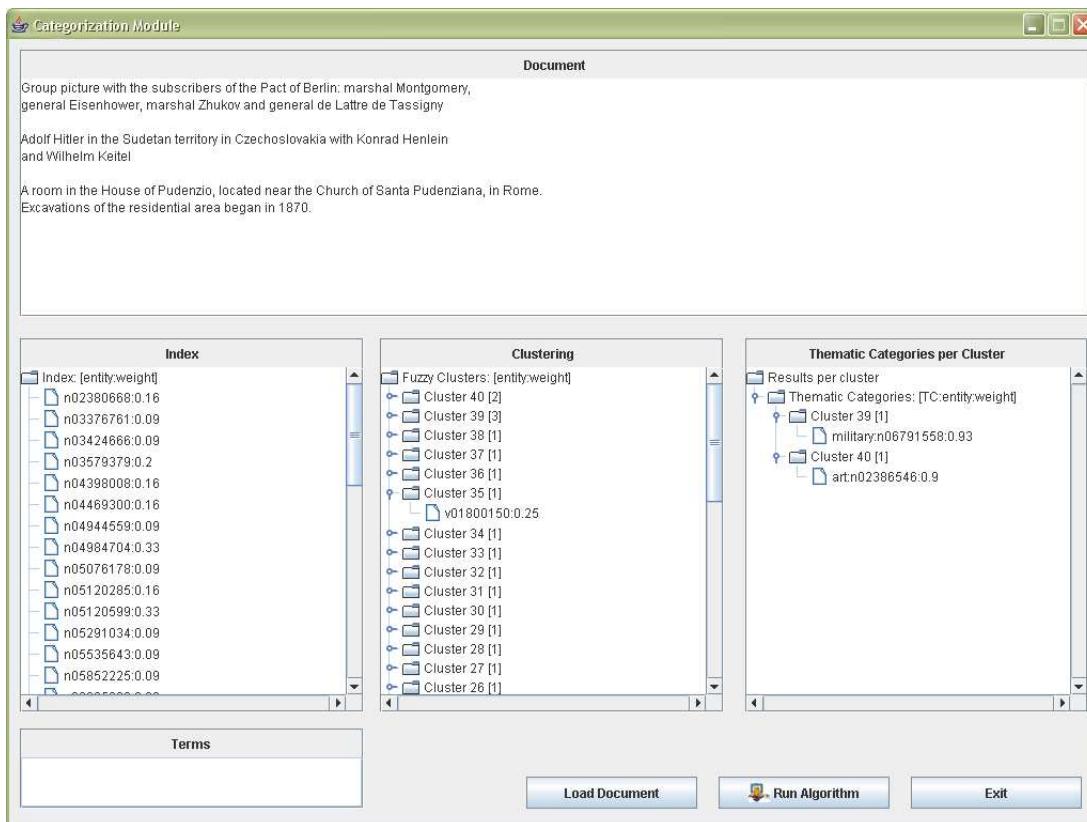
Στη συνέχεια, παρατίθενται μια ενδεικτική επιλογή από τις πειραματικές μετρήσεις που πραγματοποιήσαμε κατά την ενασχόλησή μας με την ερευνητική περιοχή της θεματικής κατηγοριοποίησης πολυμεσικών εγγράφων. Παρουσιάζονται παραδείγματα από την εφαρμογή της προτεινόμενης μεθοδολογίας τόσο σε συνθετικά πολυμεσικά έγγραφα (βλ. ενότητα 4.5.1), όσο και σε πραγματικά πολυμεσικά έγγραφα (βλ. ενότητα 4.5.2).

### 4.5.1 Θεματική κατηγοριοποίηση συνθετικών εγγράφων

Αρχικά, παρουσιάζουμε ένα πρώτο πείραμα για την επικύρωση της προτεινόμενης μεθοδολογίας, το οποίο περιλαμβάνει τη θεματική κατηγοριοποίηση πέντε πολυμεσικών εγγράφων με τη χρήση του εργαλείου που παρουσιάζεται στο Σχήμα 4.5 και που αναπτύχθηκε στα πλαίσια της διδακτορικής διατριβής ειδικά για το σκοπό αυτό. Στο παράδειγμα αυτό τα έγγραφα έχουν υποβληθεί σε επεξεργασία και έχει καθοριστεί το σημασιολογικό ευρετήριό τους. Ένα περιορισμένο σύνολο σημασιολογικών οντοτήτων έχει εξαχθεί αυτόματα από το χαρακτηρισμό (ή αλλιώς την τεκμηρίωση) του κειμένου. Οι σημασιολογικές οντότητες που περιλαμβάνονται στη βάση γνώσης παρουσιάζονται στον ακόλουθο Πίνακα 4.3, με τις θεματικές κατηγορίες να εμφανίζονται με έντονα μαύρα γράμματα.

Η υπόθεση εργασίας περιλαμβάνει πέντε (συνθετικά) έγγραφα ( $d_1, \dots, d_5$ ), ένα σύνολο από έννοιες  $S$  και μία μικρή ταξινομική σχέση  $T$  ορισμένη πάνω στο σύνολο  $S$ . Το σύνολο των εννοιών μαζί με τα μνημονικά ονόματά τους παρουσιάζεται στον Πίνακα 4.3, η σχέση  $T$  στον Πίνακα 4.4 και στο Σχήμα 4.6 και τα αποτελέσματα της κατηγοριοποίησης στον Πίνακα 4.5. Τα στοιχεία της σχέσης που υπονοούνται εξαι-

#### Κεφάλαιο 4. Εξόρυξη δεδομένων και κατηγοριοποίηση



**Σχήμα 4.5:** Εργαλείο Κατηγοριοποίησης Εγγράφων

τίας της μεταβατικότητας έχουν αγνοηθεί για λόγους σαφήνειας των Πινάκων, ενώ έχει υποτεθεί sup-γινόμενο για την υλοποίηση της μεταβατικότητας και η  $t$ -νόρμα που χρησιμοποιήθηκε για το μεταβατικό κλείσιμο της σχέσης  $T$  είναι η νόρμα Yager με παράμετρο 3. Κατά την υλοποίηση των βημάτων που περιγράφονται στην ενότητα 4.4.6 χρησιμοποιήθηκε ο ασαφής τροποποιητής  $w(a) = \sqrt{a}$  και η στάνταρ co-νόρμα  $max$  (για την τελική εξαγωγή των θεματικών κατηγοριών). Τέλος, το κατώφλι που χρησιμοποιήθηκε για το κριτήριο τερματισμού του αλγορίθμου της συσταδοποίησης ήταν το 0.3 και ως “μεγάλος” ασαφής αριθμός  $L(\cdot)$  χρησιμοποιείται ο τριγωνικός ασαφής αριθμός  $(1.3, 3, \infty)$ <sup>6</sup> [139].

Το έγγραφο  $d_1$  έχει κατασκευαστεί υποθέτοντας ότι περιέχει ένα πλάνο από μια θεατρική αίθουσα. Το έργο που παίζεται είναι σχετικό με τον πόλεμο. Υποθέτουμε ότι τα αντικείμενα και τα γεγονότα έχουν ανιχνευθεί με έναν περιορισμένο βαθμό βεβαιότητας, όπως αναμένεται συνήθως να συμβεί κατά τη διαδικασία εξαγωγής των εννοιών κατευθείαν από το ακατέργαστο περιεχόμενο. Επιπλέον, οι ανιχνευμένες έννοιες δεν σχετίζονται πάντοτε με τη γενική θεματική κατηγορία του εγγράφου. Παραδείγματος χάριν, ένα άρμα μάχης ( $tnk$ ) είναι δυνατόν να εμφανιστεί σε ένα πλάνο από ένα θέατρο ως μέρος της θεατρικής παράστασης, αλλά αυτές οι πληροφορίες δεν είναι δυνατόν να βοηθήσουν στο στάδιο της σημασιολογικής ταξινόμησης και συνεπώς αγνοούνται από τον αλγόριθμο εξαγωγής θεματικών κατηγοριών. Το ίδιο πράγμα ισχύει και για την έννοια ομιλία ( $spk$ ). Εν προκειμένω, η σημασιολογική

<sup>6</sup> Εστω ότι  $a, b, c \in R$ ,  $a < b < c$ . Ο ασαφής αριθμός  $u : R \rightarrow [0, 1]$  που αναπαριστάται με  $(a, b, c)$  και ορίζεται ως  $u(x) = 0$ , αν  $x \leq a$  ή  $x \geq c$ ,  $u(x) = \frac{x-a}{b-a}$ , αν  $x \in [a, b]$  και  $u(x) = \frac{c-x}{c-b}$ , αν  $x \in [b, c]$  ονομάζεται τριγωνικός ασαφής αριθμός.

**Πίνακας 4.3:** Ονόματα εννοιών και τα μνημονικά τους. Οι θεματικές κατηγορίες παριστάνονται με έντονα γράμματα.

Έννοια	Μνημονικό	Έννοια	Μνημονικό
πόλεμος	war	ηθοποιός	prf
τανκς	tnk	ομιλία	spk
πύραυλος	msl	θέατρο	thr
έκρηξη	exp	καθήμενο άτομο	spr
εκτόξευση πυράυλου	lms	οθόνη	scr
μαχητικό αεροπλάνο	far	κουρτίνα	crn
στρατιωτική στολή	unf	θέση	sit
F16	f16	κερκίδα	tir
λάκτισμα	sht	ποδόσφαιρο	fbl
ποτάμι	riv	γρασίδι	lwn
τέχνες	art	γκολ	gol
σινεμά	cnm	ποδοσφαιριστής	fpl
σκηνή	scn	τερματοφύλακας	glk

ευρετηρίαση του εγγράφου  $d_1$  είναι η ακόλουθη:

$$d_1 = \text{prf}/0.9 + \text{spr}/0.9 + \text{spk}/0.6 + \text{sit}/0.7 + \text{crn}/0.8 + \text{scn}/0.9 + \text{tnk}/0.7 \quad (4.48)$$

Το έγγραφο  $d_2$  περιέχει ένα πλάνο από έναν κινηματογράφο. Προκειμένου να δυσκολέψουμε τον αλγόριθμο, θεωρούμε ότι το έργο είναι και πάλι σχετικό με τον πόλεμο. Η σημασιολογική ευρετηρίαση του  $d_2$  αναπαρίσταται ως εξής:

$$d_2 = \text{spr}/0.9 + \text{spk}/0.8 + \text{sit}/0.9 + \text{scr}/1 + \text{tnk}/0.4 \quad (4.49)$$

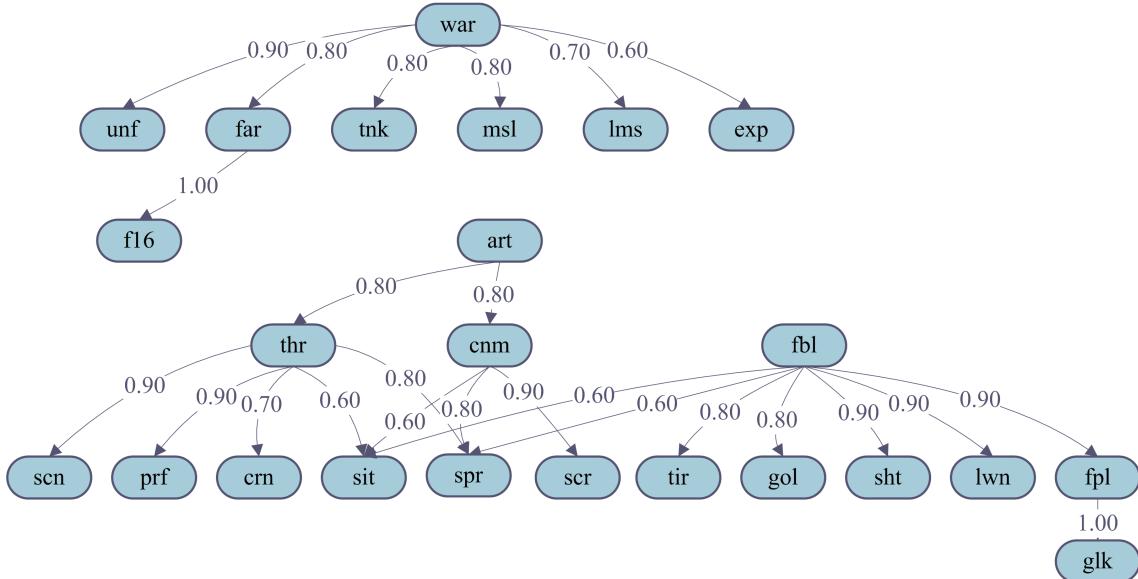
Η διαδικασία συσταδοποίησης των εννοιών καταλήγει τελικά σε 3 κλασικές συστάδες:

$$G_{d_2} = \{c_1, c_2, c_3\} = \{(spr, scr, sit), spk, tnk\} \quad (4.50)$$

Εξαιτίας της απλότητας του περιεχομένου του εγγράφου  $d_2$  και του μικρού αριθμού ανιχνευμένων εννοιών η χρήση του βασισμένου στο εννοιολογικό πλαίσιο ταξινομητή που παρουσιάστηκε στην ενότητα 4.4.6 δεν επεκτείνει τις ανιχνευμένες κλασικές συστάδες, έτσι ώστε να συμπεριλάβει και άλλες έννοιες. Αυτό είναι κάτι το αναμενόμενο, ιδίως αν παρατηρήσουμε τη δομή της σχέσης  $T$  στο Σχήμα 4.6, καθώς η σημασιολογία

**Πίνακας 4.4:** Η ταξινομική σχέση  $T$ . Μηδενικά και υπονοούμενα από τη μεταβατικότητα στοιχεία έχουν παραληφθεί.

<b><math>s_1</math></b>	<b><math>s_2</math></b>	<b><math>T</math></b>	<b><math>s_1</math></b>	<b><math>s_2</math></b>	<b><math>T</math></b>	<b><math>s_1</math></b>	<b><math>s_2</math></b>	<b><math>T</math></b>
war	unf	0.90	war	lms	0.70	war	exp	0.60
war	far	0.80	tbl	lwn	0.90	tbl	gol	0.80
war	tnk	0.80	cnm	scr	0.90	tbl	sit	0.60
war	msl	0.80	cnm	spr	0.80	cnm	sit	0.60
thr	scn	0.90	tbl	spr	0.60	tbl	sht	0.90
thr	prf	0.90	thr	sit	0.60	tbl	tir	0.80
thr	spr	0.80	thr	crn	0.70	tbl	fpl	0.90
far	f16	1.00	art	thr	0.80	art	cnm	0.80
fpl	glk	1.00						



Σχήμα 4.6: Παράδειγμα κατασκευής της σχέσης  $T$ .

όλων των εννοιών στο έγγραφο  $d_2$  υπονοεί είτε μία συνολική, είτε μία εντελώς απούσα σχέση με την  $c$ . Κατόπιν, ρυθμίζουμε περαιτέρω τους βαθμούς ιδιότητας μέλους για τις συστάδες και χρησιμοποιώντας το γινόμενο ως  $t$ -νόρμα έχουμε:

$$\begin{aligned} c_1^i(s) &= spr/0.9 + scr/1.0 + sit/0.9 \\ c_2^i(s) &= spk/0.8 \\ c_3^i(s) &= tnk/0.4 \end{aligned} \quad (4.51)$$

Καθεμία από τις ανωτέρω συστάδες αντιστοιχεί σε μία από τις διακριτές θεματικές κατηγορίες που συνδέονται με το  $d_2$  και προκειμένου να τις αναγνωρίσουμε θα πρέπει σύμφωνα με τα παραπάνω να εξετάσουμε και τον βαθματό αριθμό στοιχείων του συνόλου κάθε συστάδας, καθώς επίσης και το εννοιολογικό πλαίσιο της. Για κάθε συστάδα του εγγράφου  $d_2$  έχουμε:

$$\begin{aligned} h(c_1^i(s)) &= 1.0 \quad και \quad |c_1^i(s)| = 3 \\ h(c_2^i(s)) &= 0.8 \quad και \quad |c_2^i(s)| = 1 \\ h(c_3^i(s)) &= 0.4 \quad και \quad |c_3^i(s)| = 1 \end{aligned} \quad (4.52)$$

Τελικά, η κανονικοποίηση των παραπάνω συστάδων καταλήγει στα:

$$\begin{aligned} c_1^n(s) &= spr/0.9 + scr/1.0 + sit/0.9 \\ c_2^n(s) &= spk/1.0 \\ c_3^n(s) &= tnk/1.0 \end{aligned} \quad (4.53)$$

Το εννοιολογικό τους πλαίσιο υπολογίζεται ως εξής:

$$\begin{aligned} K(c_1^n(s)) &= cnm/0.6 + art/0.58 \\ K(c_2^n(s)) &= \emptyset \\ K(c_3^n(s)) &= war/0.8 \end{aligned} \quad (4.54)$$

Εφαρμόζοντας τον ασαφή τροποποιητή  $w(a) = \sqrt{a}$ , έχουμε:

$$\begin{aligned} W^*(c_1^n) &= w(K(c_1^n)) = \sqrt{K(c_1^n)} = cnm/0.77 + art/0.76 \\ W^*(c_2^n) &= w(K(c_2^n)) = \emptyset \\ W^*(c_3^n) &= w(K(c_3^n)) = \sqrt{K(c_3^n)} = war/0.89 \end{aligned} \quad (4.55)$$

Οι συστάδες  $c_2^n$  και  $c_3^n$  χαρακτηρίζονται από εξαιρετικά χαμηλή πληθυκότητα και κατά συνέπεια περιέχουν παραπλανητικές έννοιες, όσον αφορά τις θεματικές κατηγορίες του εγγράφου  $d_2$ . Αφού ρυθμίσουμε τους βαθμούς ιδιότητας μέλους των συστάδων σύμφωνα με τις βαθμωτές πληθυκότητές τους χρησιμοποιώντας τον τριγωνικό ασαφή αριθμό  $(1.3, 3, \infty)$ , και οι δύο συστάδες αγνοούνται κατά την εκτίμηση του ζητούμενου  $W_{d_2}$ . Τέλος, το σύνολο των θεματικών κατηγοριών του εγγράφου  $d_2$  δίνεται από την:

$$W_{d_2} = \bigcup_{c \in G_d} W(c) = W(c_1) = W^*(c_1^n) = cnm/0.77 + art/0.76 \quad (4.56)$$

Αν και μερικές έννοιες είναι κοινές μεταξύ  $d_1$  και  $d_2$  και σχετίζονται και με τα δύο θέματα (θέατρο και κινηματογράφος), ο αλγόριθμος ανιχνεύει σωστά ότι η γενική θεματική κατηγορία τους είναι διαφορετική. Αυτό πραγματοποιείται με το να θεωρήσουμε ότι η έννοια οθόνη αλλάζει το εννοιολογικό πλαίσιο και κατά συνέπεια τη γενικότερη ερμηνεία.

**Πίνακας 4.5:** Αποτελέσματα κατηγοριοποίησης εγγράφων. Τιμές κάτω από 0.1 έχουν αγνοηθεί.

Θέμα	Έγγραφα				
	$d_1$	$d_2$	$d_3$	$d_4$	$d_5$
πόλεμος					0.77
τέχνες	0.84	0.77			0.85
κινηματογράφος			0.76		0.86
θέατρο		0.89			0.33
ποδόσφαιρο				0.84 0.37	0.76

Κατόπιν, εξετάζουμε τα έγγραφα  $d_3$  και  $d_4$ , τα οποία σχετίζονται με το ποδόσφαιρο. Η διαφορά τους έγκειται στη βεβαιότητα με την οποία έχουν ανιχνευθεί οι έννοιες σε αυτά. Όπως μπορεί κανείς να παρατηρήσει, ο αλγόριθμός μας ενσωματώνει με επιτυχία την αβεβαιότητα της εισόδου στο αποτέλεσμά του:

$$\begin{aligned} d_3 = & spr/0.8 + unf/0.9 + lwn/0.6 + gol/0.9 + \\ & tir/0.7 + spk/0.9 + glk/0.6 + sht/0.5 \end{aligned} \quad (4.57)$$

$$\begin{aligned} d_4 = & spr/0.2 + unf/0.3 + lwn/0.4 + gol/0.3 + \\ & tir/0.4 + spk/0.2 + glk/0.3 + sht/0.4 \end{aligned} \quad (4.58)$$

Τέλος, το έγγραφο  $d_5$  είναι μια ακολουθία πλάνων από ένα δελτίο ειδήσεων. Λόγω της ποικιλομορφίας των ιστοριών που παρουσιάζονται στο δελτίο οι έννοιες που ανιχνεύονται και περιλαμβάνονται στο σημασιολογικό ευρετήριο είναι αρκετά ανεξάρτητες η μια από την άλλη. Προκειμένου να διευκρινιστεί πλήρως η διαδικασία της ασαφοποίησης των συστάδων, θα αναλύσουμε εν συντομίᾳ τα συγκεκριμένα βήματα για το

#### Κεφάλαιο 4. Εξόρυξη δεδομένων και κατηγοριοποίηση

έγγραφο  $d_5$ . Η σημασιολογική ευρετηρίαση του εγγράφου είναι:

$$\begin{aligned} d_5 = & \text{spr}/0.9 + \text{unf}/0.8 + \text{lwn}/0.5 + \text{gol}/0.9 + \text{tir}/0.7 + \\ & \text{spk}/0.9 + \text{glk}/0.8 + \text{sht}/0.5 + \text{prf}/0.7 + \text{sit}/0.9 + \text{crn}/0.7 + \\ & \text{scn}/0.8 + \text{tnk}/0.9 + \text{msl}/0.8 + \text{exp}/0.9 + \text{riv}/1.0 \end{aligned} \quad (4.59)$$

Θεωρώντας την ασάφεια του ευρετηρίου, όπως αυτή περιγράφεται σε προηγούμενες ενότητες, υπολογίζουμε τις ακόλουθες πέντε ασαφείς συστάδες εννοιών για το έγγραφο  $d_5$ :

$$\begin{aligned} c_1^i &= \text{spk}/0.9 \\ c_2^i &= \text{riv}/1.0 \\ c_3^i &= \text{spr}/0.9 + \text{prf}/0.7 + \text{sit}/0.77 + \text{crn}/0.7 + \text{scn}/0.8 \\ c_4^i &= \text{spr}/0.9 + \text{lwn}/0.5 + \text{gol}/0.9 + \text{tir}/0.7 + \text{glk}/0.8 + \text{sht}/0.5 + \text{sit}/0.9 \\ c_5^i &= \text{unf}/0.8 + \text{tnk}/0.9 + \text{msl}/0.8 + \text{exp}/0.9 \end{aligned} \quad (4.60)$$

Οι πρώτες δύο συστάδες  $c_1^i$  και  $c_2^i$  αγνοούνται κατά το στάδιο προσδιορισμού του ασαφούς συνόλου θεματικών κατηγοριών  $W_{d_5}$  που σχετίζονται με το έγγραφο λόγω της χαμηλής πληθικότητάς τους. Με βάση τη μεθοδολογία που έχει ήδη παρουσιαστεί και η οποία εξετάζει το εννοιολογικό πλαίσιο καθεμιάς από τις υπόλοιπες συστάδες  $c_3^i$ ,  $c_4^i$  και  $c_5^i$  και αναγνωρίζοντας το γεγονός ότι το τελευταίο έχει καθοριστεί μόνο για τα κανονικά ασαφή σύνολα, προσδιορίζουμε τις σχετικές με το  $d_5$  θεματικές κατηγορίες, όπως περιγράφεται στην τελευταία στήλη του Πίνακα 4.5. Όπως παρατηρεί κανείς, ο αλγόριθμος προσδιορίζει επιτυχώς την ύπαρξη περισσότερων του ενός ευδιάκριτων θεματικών κατηγοριών στο έγγραφο. Επιπλέον, έννοιες, όπως το κάθισμα και το καθήμενο πρόσωπο, εμπεριέχονται σε περισσότερες από μια συστάδες, δεδομένου ότι σχετίζονται με περισσότερα του ενός εννοιολογικά πλαίσια που ανιχνεύονται στο έγγραφο.

#### 4.5.2 Θεματική κατηγοριοποίηση πραγματικών εγγράφων

Για να επεξηγήσουμε περαιτέρω την απόδοση της μεθοδολογίας μας, πραγματοποιήσαμε και μια σειρά πειραμάτων αξιολόγησης πάνω σε ένα πραγματικό σύνολο  $Q$  από πολυμεσικά έγγραφα, που προήλθε από τα αρχεία πολυμέσων της Ελληνικής Ραδιοφωνίας Τηλεόρασης (EPT), του Αυστριακού Κινηματογραφικού Αρχείου (FAA), της Ταινιοθήκης της Ελλάδας (FAG) και της Αυστριακής Εταιρίας Ραδιοτηλεοπτικής Μετάδοσης (ORF). Όλα τα έγγραφα διέπονται από μεγάλη ποικιλομορφία από την άποψη του περιεχομένου τους, γεγονός το οποίο μετουσιώνει την προσπάθεια ενοποίησης σε μία πραγματική πρόκληση. Το πολυμεσικό υλικό συνοψίζεται σε περίπου 80 ώρες οπτικοακουστικών προγραμμάτων, ντοκιμαντέρ και δελτίων ειδήσεων, τα οποία και κατανέμονται σε ένα σύνολο 484 ξεχωριστών πολυμεσικών εγγράφων, με τη διάρκειά τους να κυμαίνεται από 55'' ως 35'.28''. Κάθε πολυμεσικό έγγραφο έχει χαρακτηρισθεί εκ των προτέρων με ανθρωπίνως κατανοητές λέξεις-κλειδιά, περιλαμβάνει διάφορα πολυμεσικά προγράμματα και αντικείμενα ειδήσεων και κάθε αντικείμενο περιέχει από μία έως μερικές δεκάδες έννοιες, γεγονός που καταλήγει στην ύπαρξη περίπου 30000 εννοιών μέσα στους χαρακτηρισμούς των εγγράφων.

Εξαιτίας των διαφορών στο μέγεθος και το περιεχόμενο των διαθέσιμων τηλεοπτικών προγραμμάτων αποφασίσαμε να ακολουθήσουμε στη συνέχεια μια διττή προσέγγιση, όσον αφορά τα διαθέσιμα πολυμεσικά έγγραφα: αφ' ενός πραγματοποιήσαμε τα πειράματά μας με το αρχικό σύνολο των πολυμεσικών προγραμμάτων και αφ' ετέρου πραγματοποιήσαμε εκ νέου τα ίδια πειράματα με ένα επεκτεταμένο σύνολο από αντικείμενα ειδήσεων (news items). Καθώς ένας αριθμός από αντικείμενα ειδήσεων αποτελεί ένα πολυμεσικό πρόγραμμα, αυτά είναι αφ' ενός σημαντικά μικρότερα σε διάρκεια (η οποία κυμαίνεται από 1'.28" έως 5'.12"), αλλά αφετέρου θεωρούνται έγγραφα πολυμέσων και συγχρόνως ο αριθμός τους είναι μεγαλύτερος (συνολικά είναι διαθέσιμα 1976 διακριτά αντικείμενα ειδήσεων). Στα πλαίσια αυτής της προσέγγισης χειρίζόμαστε τόσο τα προγράμματα πολυμέσων όσο και τα αντικείμενα ειδήσεων με έναν ενοποιημένο τρόπο ως πολυμεσικά έγγραφα. Το γεγονός αυτό, μας οδήγησε σε δύο σύνολα πειραμάτων, όπως αυτά παρουσιάζονται ακολούθως. Για λόγους απλότητας και εξελιξιμότητας (scalability) επιλέχτηκαν  $|Y| = N = 13$  ενδεικτικές θεματικές κατηγορίες μεταξύ των οποίων:

- αθλητισμός
- πολιτική
- θρησκεία
- ειδήσεις
- ηγέτες
- στρατιωτικά
- τέχνες
- υγεία
- ταξίδια
- happenings
- εκπαίδευση
- διαδηλώσεις
- ιστορία

'Όλα τα έγγραφα είχαν ταξινομηθεί χειρωνακτικά εκ των προτέρων, προκειμένου να κατασκευασθεί η απόλυτη αλήθεια (ground truth) και να είναι εφικτή η μετέπειτα αξιολόγηση του αλγορίθμου ταξινόμησής μας. Λόγω της υποκειμενικότητας που εισάγεται από αυτή τη χειρωνακτική διαδικασία η ταξινόμηση της απόλυτης αλήθειας ακολούθησε τις αρχές της κλασικής θεωρίας συνόλων, σε αντιδιαστολή με την ασαφή προσέγγιση της μεθόδου μας. Εντούτοις ένα έγγραφο - είτε αυτό είναι πολυμεσικό πρόγραμμα, είτε αντικείμενο ειδήσεων - θα μπορούσε να ανήκει σε πολλαπλές θεματικές κατηγορίες εξαιτίας των πιθανών θεματικών μερών που μπορεί αυτό να περιέχει,

#### Κεφάλαιο 4. Εξόρυξη δεδομένων και κατηγοριοποίηση

με αποτέλεσμα την τεχνητή διεύρυνση του αρχικού συνόλου των πολυμεσικών προγραμμάτων και αντικειμένων ειδήσεων από 484 σε  $Q = 653$  προγράμματα πολυμέσων και από 1976 σε  $Q = 2733$  αντικείμενα ειδήσεων:

$$Q = \bigcup_{i=1}^N Q_i, Q_i \cap Q_j \neq \emptyset, i, j \in \{1, \dots, N\} \quad (4.61)$$

όπου το  $Q_i$  είναι το σύνολο των πολυμεσικών εγγράφων που σχετίζεται πραγματικά με τη θεματική κατηγορία  $i$ . Θα πρέπει να σημειώσουμε ότι η παραπάνω προσέγγιση οδηγεί σε μια μάλλον απαισιόδοξη αξιολόγηση της μεθοδολογίας μας, όπως αυτή παρουσιάζεται στους Πίνακες 4.6 και 4.7, εντούτοις η εφαρμογή της δεν απαιτεί περίπλοκες ικανότητες από τους εμπειρογνώμονες που χαρακτηρίζουν τα πολυμεσικά έγγραφα και αντιπροσωπεύει πλήρως μια κατάσταση της πραγματικής ζωής.

Πραγματοποιήθηκε μια σειρά πειραμάτων, για να μετρήσουμε την αποτελεσματικότητα και την απόδοση της προτεινόμενης ταξινόμησης, χρησιμοποιώντας τα μεγέθη της ιδιομορφίας (specificity -  $sp$ ), της ευαισθησίας (sensitivity -  $sn$ ) και της αποτελεσματικότητας (effectiveness -  $e$ ), όπως αυτά ορίστηκαν στην υποενότητα 2.4.3. Έστω ότι ο αριθμός των εγγράφων που σχετίζονται με ένα θέμα και που αναγνωρίζονται σωστά ότι ανήκουν στο συγκεκριμένο θέμα, δηλ. με άλλα λόγια τα πολυμεσικά έγγραφα που ταξινομούνται με σωστό τρόπο, αναπαρίστανται με τον όρο “αληθινά και θετικά -  $TP$ ” και ότι ο αριθμός των εγγράφων που αναγνωρίζεται εσφαλμένα να μην ανήκει σε αυτό το θέμα, δηλ. τα πολυμεσικά έγγραφα που ταξινομούνται εσφαλμένα, αναπαριστώνται με τον όρο “ψεύτικα και αρνητικά -  $FN$ ”. Ομοίως, έστω ότι ο αριθμός εγγράφων που δε σχετίζονται πραγματικά με ένα συγκεκριμένο θέμα υπό εξέταση και ταξινομούνται είτε σωστά, είτε λανθασμένα, δείχνεται από τους όρους “αληθινά και αρνητικά -  $TN$ ” και “λανθασμένα και θετικά -  $FP$ ”, αντίστοιχα. Τότε:

$$sp = \text{ιδιομορφια} = \frac{TN}{TN + FP} \quad (4.62)$$

$$sn = \text{ευαισθησια} = \frac{TP}{TP + FN} \quad (4.63)$$

$$e = \text{αποτελεσματικοτητα} = \frac{1}{a(sp) + (1-a)(1/sn)} \quad (4.64)$$

όπου η παράμετρος  $a$  επηρεάζει την εκτίμηση της αποτελεσματικότητας  $e$  επιτρέποντας τη διαφορετική στάθμιση της ιδιομορφίας και της ευαισθησίας, δηλ. μια χαμηλή τιμή του  $a$  ευνοεί την ευαισθησία, ενώ μια υψηλή τιμή του ευνοεί την ιδιομορφία. Ο στόχος οποιουδήποτε πειράματος (πρέπει να) είναι η μεγιστοποίηση και των δύο μεγεθών,  $sp$  και  $sn$ . Εντούτοις, πειραματικά και θεωρητικά έχει αποδειχθεί ότι μια αύξηση σε μια εκ των δύο τιμών οδηγεί στη μείωση της άλλης. Ο Πίνακας 4.6 και ο Πίνακας 4.7 παρουσιάζουν δύο ασαφείς μήτρες σύγχυσης (fuzzy confusion matrices) που περιέχουν τις πληροφορίες για τα πραγματικά και ανιχνευμένα θέματα των πολυμεσικών εγγράφων (δηλ. των προγραμμάτων και των αντικειμένων ειδήσεων, αντίστοιχα), όπου:

$$D_{ij} = \sum_{d \in Q_i} W_d(j), i, j \in \{1, \dots, N\} \quad (4.65)$$

και  $W_d(j)$  είναι ο βαθμός στον οποίο το το έγγραφο  $d$  κατηγοριοποιείται στη θεματική κατηγορία  $j$ .

Στην περίπτωση αυτή, οι τιμές  $TP_t, FP_t, TN_t, FN_t$  για κάθε θεματική κατηγορία  $t \in \{1, \dots, N\}$  ορίζονται ως:

$$TP_t = D_{tt}, FP_t = \sum_{\substack{k=1 \\ k \neq t}}^N D_{tk}, TN_t = \sum_{\substack{k=1 \\ k \neq t}}^N D_{kk}, FN_t = \sum_{\substack{k=1 \\ k \neq t}}^N D_{kt} \quad (4.66)$$

**Πίνακας 4.6:** Ασαφής μήτρα σύγχυσης  $D$  για τη σημασιολογική κατηγοριοποίηση των πολυμεσικών προγραμμάτων.

Πραγματικά θέματα	Ανιχνευθέντα θέματα												
	αθλητικά πολιτική θρησκεία ειδήσεις ηγέτες στρατιωτικά τέχνες υγεία ταξίδια happening εκπαίδευση διαδηλώσεις ιστορία												
αθλητικά	45.00	1.04	0.00	1.44	2.10	0.00	0.13	0.66	0.60	1.26	0.32	0.00	0.14
πολιτική	0.72	35.10	1.05	1.17	1.35	0.90	1.92	0.48	1.16	2.59	2.97	1.15	2.75
θρησκεία	0.00	0.72	22.36	0.12	2.31	0.90	2.17	0.51	0.00	0.00	1.76	0.44	1.45
ειδήσεις	0.76	0.80	0.00	49.28	0.60	2.45	1.32	1.35	0.58	0.57	1.44	0.88	4.80
ηγέτες	0.00	0.21	1.95	0.26	33.44	1.86	0.00	0.48	0.45	0.11	0.24	0.60	2.97
στρατιωτικά	0.00	1.14	0.64	0.60	0.28	29.88	0.00	0.00	0.00	0.00	0.00	1.68	2.45
τέχνες	0.00	0.00	0.15	0.00	1.40	0.00	26.40	0.00	1.53	2.80	0.90	0.00	1.02
υγεία	0.00	0.00	0.00	0.35	0.00	0.00	0.00	22.14	1.26	1.28	1.10	0.00	1.45
ταξίδια	0.68	0.00	0.16	0.84	0.00	0.00	0.00	0.00	25.52	0.00	0.00	0.00	0.84
happening	0.72	0.70	0.27	0.16	1.35	0.19	2.80	1.20	1.92	33.93	1.12	1.04	0.42
εκπαίδευση	0.00	0.44	0.52	0.58	0.19	0.00	1.25	1.08	0.60	0.32	25.20	0.35	0.64
διαδηλώσεις	0.00	0.33	0.72	1.50	1.20	0.54	0.00	0.42	0.60	0.54	1.30	26.66	0.87
ιστορία	0.54	0.78	0.70	1.20	0.62	0.46	0.00	0.32	1.60	0.96	1.20	1.08	43.00
Σύνολο:	48.42	41.26	28.52	57.50	44.84	37.18	35.99	28.64	35.82	44.36	37.55	33.88	62.80

Για παράδειγμα:

$$TP_{\alpha\theta\lambda\eta\tau\iota\kappa} = 45$$

$$FP_{\alpha\theta\lambda\eta\tau\iota\kappa} = 1.04 + 0 + 1.44 + 2.10 + 0 + 0.13 + 0.66 + 0.60 + 1.26 + 0.32 + 0 + 0.14 = 7.69$$

$$TN_{\alpha\theta\lambda\eta\tau\iota\kappa} = 35.10 + 22.36 + \dots + 43.00 = 372.91$$

και

$$FN_{\alpha\theta\lambda\eta\tau\iota\kappa} = 0.72 + 0 + 0.76 + 0 + 0 + 0 + 0.68 + 0.72 + 0 + 0 + 0.54 = 3.42$$

**Πίνακας 4.7:** Ασαφής μήτρα σύγχυσης  $D$  για τη σημασιολογική κατηγοριοποίηση των ειδήσεογραφικών αντικειμένων.

Πραγματικά θέματα	Ανιχνευθέντα θέματα												
	αθλητικά πολιτική θρησκεία ειδήσεις ηγέτες στρατιωτικά τέχνες υγεία ταξίδια happening εκπαίδευση διαδηλώσεις ιστορία												
αθλητικά	189.00	2.45	0.00	7.91	8.79	0.00	0.41	4.14	0.75	3.96	0.50	0.00	2.20
πολιτική	0.75	141.30	10.55	26.94	45.22	16.96	1.51	0.75	2.73	3.49	9.33	12.28	19.63
θρησκεία	0.00	3.39	90.30	1.51	5.18	0.00	0.97	0.00	0.00	1.38	1.38	1.38	10.93
ειδήσεις	17.90	47.73	0.00	197.12	7.54	7.69	7.60	2.83	9.11	4.77	16.96	4.84	6.28
ηγέτες	0.00	4.62	3.67	0.82	140.80	7.79	0.00	0.00	0.94	0.35	1.88	5.02	15.54
στρατιωτικά	0.00	7.16	0.50	9.42	2.64	119.52	0.00	0.00	0.00	0.00	0.00	5.93	24.18
τέχνες	0.00	0.00	2.36	1.26	0.63	0.00	100.00	0.00	0.53	21.35	5.65	0.00	1.07
υγεία	0.00	0.00	0.00	2.20	0.00	0.00	0.00	91.84	3.96	1.51	4.84	0.00	1.82
ταξίδια	0.00	0.00	0.00	4.62	0.00	0.00	0.00	0.00	110.88	0.00	0.00	0.00	0.66
happening	8.29	1.76	0.85	1.51	0.85	0.00	15.07	0.00	0.00	137.46	1.00	8.16	0.88
εκπαίδευση	0.00	3.45	0.00	5.46	1.19	0.00	2.36	5.65	1.88	0.50	109.20	15.39	3.01
διαδηλώσεις	0.00	19.69	1.51	14.13	3.77	1.70	0.00	0.00	0.00	2.83	8.16	103.20	5.46
ιστορία	0.00	11.02	8.79	5.65	10.71	8.67	0.00	0.50	1.00	0.00	5.02	6.78	176.30
Σύνολο:	215.94	242.57	118.53	278.54	227.31	162.32	127.91	105.72	131.79	176.21	163.93	162.99	267.96

Σύμφωνα με τους ανωτέρω ορισμούς, χρησιμοποιώντας μια τιμή για το  $a$  ίση με 0.5, δηλ.  $e(a = 0.5) = \frac{2(sp)(sn)}{sp+sn}$ , και λαμβάνοντας υπόψη τις ασαφείς μήτρες σύγχυσης που παρουσιάστηκαν στους Πίνακες 4.6 και 4.7, μετρήσαμε την ιδιομορφία, την ευαισθησία και την αποτελεσματικότητα για καθένα από τα 13 θέματα κατ' επίκληση της απόλυτης αλήθειας, όπως απεικονίζεται στον Πίνακα 4.8 και τον Πίνακα 4.9, για τα προγράμματα πολυμέσων και τα ειδήσεογραφικά αντικείμενα, αντίστοιχα.

Παρατηρώντας τους δύο Πίνακες, δεν υπάρχει καμία σημαντική διαφορά των δεικτών ιδιομορφίας, ευαισθησίας ή αποτελεσματικότητας μεταξύ των διαφορετικών θεμάτων. Παραδείγματος χάριν, εξετάζοντας και τα 13 θέματα, παρατηρούμε ότι η

#### Κεφάλαιο 4. Εξόρυξη δεδομένων και κατηγοριοποίηση

αποτελεσματικότητα κυμαίνεται μεταξύ 0.80 και 0.95 (βλ. π.χ. Πίνακα 4.8), ενώ παρατηρούμε επίσης ότι, εν γένει, η ιδιομορφία είναι υψηλότερη από την ευαισθησία. Επιπλέον, όλες οι τιμές ευαισθησίας των θεμάτων βρίσκονται κοντά στην αντίστοιχη ιδιομορφία, γεγονός που δείχνει μια γενική ικανοποιητική απόδοση του προτεινόμενου πλαισίου πάνω σε ολόκληρο το σύνολο των διαθέσιμων δεδομένων, λαμβάνοντας υπόψη και την ποικιλία των θεματικών κατηγοριών. Συγκρίνοντας τις μέσες τιμές ιδιομορφίας και ευαισθησίας για τα πολυμεσικά προγράμματα στον Πίνακα 4.8 και τα ειδησεογραφικά αντικείμενα στον Πίνακα 4.9, παρατηρούμε ότι και στις δύο περιπτώσεις η ιδιομορφία παραμένει μάλλον υψηλή, κάτι το οποίο σημαίνει ότι στα πολυμεσικά έγγραφα δεν αποδίδονται ανακριβή πρόσθετα θέματα, ενώ στην περίπτωση των 653 προγραμμάτων πολυμέσων η τιμή της ευαισθησίας κατατροπώνει την αντίστοιχη τιμή ευαισθησίας των 2773 ειδησεογραφικών αντικειμένων. Το παραπάνω δικαιολογείται από το γεγονός ότι τα πολυμεσικά προγράμματα έχουν μεγαλύτερη χρονική διάρκεια με τον τρόπο αυτό ευρετηριάζονται με περισσότερες έννοιες κατά τη δημιουργία του σημασιολογικού ευρετηρίου και σχετίζονται με περισσότερες θεματικές κατηγορίες, κάτι το οποίο δικαιώνει πλήρως την επιλογή μας για ένα ασαφές σημασιολογικό ευρετήριο και για μία ασαφή προσέγγιση συσταδοποίησης. Χωρίς αυτά η κατηγοριοποίηση ενός εγγράφου σε περισσότερα του ενός θέματα δε θα ήταν εφικτή.

**Πίνακας 4.8:** Αποτελέσματα αξιολόγησης σημασιολογικής κατηγοριοποίησης - 653 πολυμεσικά προγράμματα.

Θέματα	GT	TP	FP	TN	FN	εννοιολογικό πλαίσιο			pLSA		
						sp	sn	e	sp	sn	e
αθλητικά	60	45.00	7.69	372.91	3.42	0.98	0.93	0.95	0.95	0.91	0.93
πολιτική	59	35.10	18.21	382.81	6.16	0.95	0.85	0.90	0.93	0.83	0.87
θρησκεία	37	22.36	10.38	395.55	6.16	0.97	0.78	0.87	0.89	0.72	0.80
ειδήσεις	82	49.28	15.55	368.63	8.22	0.96	0.86	0.91	0.94	0.85	0.89
ηγέτες	51	33.44	9.13	384.47	11.40	0.98	0.75	0.85	0.88	0.69	0.77
στρατιωτικά	46	29.88	6.79	388.03	7.30	0.98	0.80	0.88	0.89	0.74	0.81
τέχνες	41	26.40	7.80	391.51	9.59	0.98	0.73	0.84	0.91	0.67	0.77
υγεία	32	22.14	5.44	395.77	6.50	0.99	0.77	0.87	0.94	0.71	0.81
ταξίδια	35	25.52	2.52	392.39	10.30	0.99	0.71	0.83	0.90	0.66	0.76
happening	54	33.93	11.89	383.98	10.43	0.97	0.76	0.86	0.94	0.70	0.80
εκπαίδευση	41	25.20	5.97	392.71	12.35	0.99	0.67	0.80	0.96	0.62	0.75
διαδηλώσεις	44	26.66	8.02	391.25	7.22	0.98	0.79	0.87	0.88	0.72	0.79
ιστορία	71	43.00	9.46	374.91	19.80	0.98	0.68	0.80	0.95	0.68	0.79
macro-average						0.97	0.78	0.87	0.92	0.74	0.82
micro-average						0.99	0.77		0.91	0.71	

**Πίνακας 4.9:** Αποτελέσματα αξιολόγησης σημασιολογικής κατηγοριοποίησης - 2733 ειδησεογραφικά αντικείμενα.

Θέματα	GT	TP	FP	TN	FN	εννοιολογικό πλαίσιο			pLSA		
						sp	sn	e	sp	sn	e
αθλητικά	247	189.00	31.12	1517.92	26.94	0.98	0.88	0.92	0.96	0.86	0.91
πολιτική	301	141.30	150.12	1565.62	101.27	0.91	0.58	0.71	0.90	0.57	0.70
θρησκεία	135	90.30	24.74	1616.62	28.23	0.98	0.76	0.86	0.90	0.70	0.79
ειδήσεις	355	197.12	133.23	1509.80	81.42	0.92	0.71	0.80	0.92	0.71	0.80
ηγέτες	212	140.80	40.63	1566.12	86.51	0.97	0.62	0.76	0.88	0.57	0.69
στρατιωτικά	196	119.52	49.83	1587.40	42.80	0.97	0.74	0.84	0.85	0.68	0.76
τέχνες	160	100.00	32.84	1606.92	27.91	0.98	0.78	0.87	0.93	0.72	0.81
υγεία	136	91.84	14.32	1615.08	13.88	0.99	0.87	0.93	0.94	0.80	0.86
ταξίδια	146	110.88	5.28	1596.04	20.91	0.99	0.84	0.91	0.90	0.77	0.83
happening	204	137.46	38.37	1569.46	38.75	0.98	0.78	0.87	0.94	0.74	0.83
εκπαίδευση	181	109.20	38.90	1597.72	54.73	0.98	0.67	0.79	0.94	0.61	0.74
διαδηλώσεις	182	103.20	57.24	1603.72	59.79	0.97	0.63	0.76	0.91	0.58	0.71
ιστορία	278	176.30	58.15	1530.62	91.66	0.96	0.66	0.78	0.94	0.64	0.77
macro-average						0.96	0.72	0.82	0.92	0.68	0.78
micro-average						0.99	0.72		0.92	0.66	

Η μεθοδολογία ευρετηρίασης και ταξινόμησης πολυμεσικών εγγράφων που προτείνουμε χρησιμοποιεί ένα εννοιολογικό πλαίσιο και ένα ταξινομικό πρότυπο γνώσης βασισμένα στη σχέση  $T$ . Σε αυτό το σημείο θα δείξουμε και θα συγχρίνουμε τα αποτελέσματά της ενάντια στη χρήση μιας τεχνικής ευρετηρίασης και ταξινόμησης εγγράφων, η οποία δε χρησιμοποιεί τη γνώση ή το εννοιολογικό πλαίσιο στη διαδικασία. Μια γνωστή, κατάλληλη τεχνική που προσπαθεί να εξάγει την υπονοούμενη σημασιολογία χωρίς τη χρήση γνώσης είναι αυτή που βασίζεται στη λανθάνουσα σημασιολογική ανάλυση (LSA). Η τελευταία δε χρησιμοποιεί κάποιο ανθρώπινα κατασκευασμένο λεξικό, βάση γνώσεων, σημασιολογικό δίκτυο, γραμματική, συντακτικό αναλυτή ή μορφολογία και παίρνει ως είσοδο μόνο το ακατέργαστο κείμενο, όπως σαφώς δηλώνεται στη βιβλιογραφία [147]. Σε αυτήν την περίπτωση, θεωρούμε τα έγγραφα και τις έννοιες που αποκτήθηκαν από την ανάλυση κειμένων ως την είσοδο του αλγορίθμου και, χωρίς να χρησιμοποιήσουμε κάποια πληροφορία γνώσης, προσπαθούμε να προσδιορίσουμε τις θεματικές κατηγορίες με το να τις αντιστοιχίσουμε σε απαρατήρητες κατηγορίες. Το πρόβλημα αυτό αντιμετωπίζεται πολύ καλά από μια πιθανοτική θεώρηση της LSA, την πιθανοτική λανθάνουσα σημασιολογική ανάλυση (pLSA) [119], η οποία έχει μια δομημένη στατιστική υποδομή και χρησιμοποιεί ένα λανθάνον μεταβλητό πρότυπο για τις απαρατήρητες κατηγορίες. Ξεκινώντας από έγγραφα και έννοιες, η pLSA μπορεί να προσδιορίσει τις υπονοούμενες απαρατήρητες μεταβλητές κατηγορίας και έχει χρησιμοποιηθεί μερικώς για την κατηγοριοποίηση κειμένων [50]. Στα πλαίσια αυτής της εργασίας εφαρμόσαμε το πρότυπο pLSA χρησιμοποιώντας έννοιες (προερχόμενες είτε από τον κειμενικό χαρακτηρισμό, είτε από τα αποτελέσματα της ανάλυσης του βίντεο), πολυμεσικά έγγραφα και θεματικές κατηγορίες αντί για όρους, έγγραφα και λανθάνουσες κατηγορίες, αντίστοιχα, προκειμένου να πραγματοποιηθεί η ταξινόμηση των πολυμεσικών εγγράφων στα 13 θέματα. Τα αποτελέσματα της pLSA παρουσιάζονται στις τελευταίες τρεις στήλες των Πινάκων 4.8 και 4.9.

Ο μακρο-υπολογισμός μέσου όρου (macro-average) και ο μικρο-υπολογισμός μέσου όρου (micro-average) που παρατηρεί κανείς στους δύο παραπάνω Πίνακες αποτελούν δύο συμβατικές μεθόδους για τον υπολογισμό μέσου όρου της απόδοσης στα θέματα. Τα μακρο-υπολογισμένα αποτελέσματα είναι υπολογισμένα κατά μέσο όρο πάνω στον αριθμό των θεμάτων, ενώ τα μικρο-υπολογισμένα αποτελέσματα είναι υπολογισμένα κατά μέσο όρο πάνω στον αριθμό όλων των μετρημένων εγγράφων. Σαν πρώτη συγχριτική παρατήρηση, η προτεινόμενη προσέγγιση παρέχει καλύτερα αποτελέσματα κατά την εξέταση των ακριβέστερων θεμάτων, όπως π.χ. η υγεία, το ταξίδι ή η θρησκεία, σε σχέση με γενικότερα, όπως ο αθλητισμός, η πολιτική ή οι ειδήσεις και κάτι τέτοιο είναι αναμενόμενο, μιας και η έννοια των πρώτων είναι ευκολότερο να προσδιορίσει στην ταξινομική γνώση. Επιπλέον, τόσο στην περίπτωση των προγραμμάτων όσο και των ειδησεογραφικών αντικειμένων (Πίνακες 4.8 και 4.9), τα αποτελέσματα που επιτυγχάνονται από την προτεινόμενη μέθοδο σημασιολογικής ταξινόμησης ξεπερνούν αυτά που λαμβάνονται από την εφαρμογή του αλγορίθμου pLSA και επιτυγχάνουν μια βελτίωση έως και 14%. Κάτι τέτοιο είναι αναμενόμενο, μιας και στην προσέγγισή μας χρησιμοποιούμε το εννοιολογικό πλαίσιο και ένα ταξινομικό μοντέλο γνώσης, που καθορίζει ρητά τις σχέσεις μεταξύ των εννοιών και των θεμάτων. Το τελευταίο είναι αδύνατο να επιτευχθεί στα πλαίσια της προσέγγισης που ακολουθεί η pLSA, η οποία, όμως, ακόμη και χωρίς τη χρήση της γνώσης, παρέχει εν γένει ελπιδοφόρα αποτελέσματα. Συνεπώς, η προσεκτική παρατήρηση των αποτελεσμάτων καταδεικνύει ότι μια στατιστική προσέγγιση, όπως αυτή που ακολουθείται

από την pLSA, δεν είναι σε θέση να ξεπεράσει μια καθοδηγούμενη από τη γνώση προσέγγιση, όπως είναι αυτή που προτείνουμε στην παρούσα εργασία, γεγονός που καθιστά ιδιαίτερα σημαντική την ερευνητική προσφορά της τελευταίας.

### 4.5.3 Θεματική κατηγοριοποίηση εγγράφων με ταυτόχρονη εξαγωγή συμπερασμάτων

Τέλος, η προτεινόμενη τεχνική αξιοποίησης υπηρεσιών εξαγωγής ασαφούς συμπερασμού εφαρμόστηκε και δοκιμάστηκε σε ένα σύνολο από πραγματικά πολυμεσικά έγγραφα, τα οποία συλλέχθηκαν από το Διαδίκτυο και τον Παγκόσμιο Ιστό<sup>7</sup>, καθώς και από προσωπικές συλλογές και τα αποτελέσματα της εφαρμογής της είναι άκρως ενθαρρυντικά για την έρευνά μας. Εξελίξαμε, για το σκοπό αυτό, ένα σενάριο χρήσης της υβριδικής μεθοδολογίας πάνω σε ένα σύνολο από πολυμεσικά έγγραφα από τις θεματικές κατηγορίες αθλητικά και πολιτική. Το σενάριο αποτελείται από 9 έγγραφα ( $d_1, \dots, d_9$ ), ένα σύνολο από έννοιες  $S$  που περικλείει το αλφάριθμο της ασαφούς βάσης γνώσης που κατασκευάστηκε (βλ. ενότητα 4.4.7.3) και μία σχέση  $T$ , ορισμένη πάνω στο σύνολο  $S$ , όπως ακριβώς περιγράφεται στην υποενότητα 4.4.7.1. Τα συνολικά αποτελέσματα της κατηγοριοποίησης των 9 εγγράφων παρουσιάζονται στον Πίνακα 4.10. Το αριστερό τμήμα κάθε στήλης εγγράφου του Πίνακα υποδεικνύει τα αποτελέσματα της κατηγοριοποίησης απλού κειμένου, ενώ το δεξιό τμήμα κάθε στήλης παρουσιάζει (με έντονα μαύρα γράμματα) τα αποτελέσματα που προέκυψαν από την εφαρμογή της προτεινόμενης προσέγγισης εξαγωγής σημασιολογικών συμπερασμάτων.

Θέμα	Έγγραφα									
	$d_1$	$d_2$	$d_3$	$d_4$	$d_5$	$d_6$	$d_7$	$d_8$	$d_9$	
αθλητικά	0.88	<b>0.69</b>	0.86	<b>0.67</b>	0.94	<b>0.56</b>	0.88	<b>0.52</b>	0.91	<b>0.71</b>
ποδόσφαιρο		<b>0.68</b>		<b>0.67</b>		<b>0.51</b>		<b>0.54</b>		<b>0.71</b>
πολιτική	0.85			0.91		0.92		0.91		0.88
εργασία										<b>0.62</b>
εκπαίδευση	0.89		0.89			0.89		0.89		
ειδήσεις						0.97				

Πίνακας 4.10: Αποτελέσματα κατηγοριοποίησης εγγράφων. Τιμές < 0.10 αγνοούνται.

Προκειμένου να αποσαφηνίσουμε τη διαδικασία της θεματικής κατηγοριοποίησης, θα αναλύσουμε περισσότερο τα επιμέρους βήματα για το έγγραφο  $d_2$ . Το έγγραφο  $d_2$  αποτελείται από μια ακολουθία πλάνων από ένα δελτίο ειδήσεων. Λόγω της ανομοιότητας των ιστοριών που υπάρχουν στην ακολουθία των πλάνων οι έννοιες που ανιχνεύονται και περιλαμβάνονται στο σημασιολογικό ευρετήριο είναι ανεξάρτητες και ασυσχέτιστες μεταξύ τους. Το σημασιολογικό ευρετήριο του εγγράφου  $d_2$  δίνεται από τον Πίνακα 4.11 τα περιεχόμενα του οποίου λειτουργούν ταυτόχρονα και ως οι ασαφείς ισχυρισμοί εννοιών για την επερχόμενη εφαρμογή της προσέγγισης εξαγωγής ασαφών συμπερασμάτων. Με βάση τη διαδικασία ιεραρχικής συσταδοποίησης και ασαφοποίησης υπολογίζουμε τις ακόλουθες ασαφείς συστάδες εννοιών για το έγγραφο (οι συστάδες που εμφανίζουν εξαιρετικά χαμηλή πληθικότητα παραλείπονται):

$$\begin{aligned} c_{475} &= \text{game}/0.11 + \text{play}/0.01 \\ c_{474} &= \text{fair}/0.20 + \text{union}/0.09 \\ c_{473} &= \text{score}/0.05 \end{aligned} \tag{4.67}$$

<sup>7</sup>Ειδικότερα από τα web-sites: <http://www.cnn.com> και <http://en.wikipedia.org>.

Σύμφωνα με τη μεθοδολογία που παρουσιάστηκε στην υποενότητα 4.4.7.2, ανιχνεύουμε τα θέματα αθλητικά και εκπαίδευση ως σχετικά με το έγγραφο  $d_2$  και όπως παρουσιάζεται στη σχετική στήλη του Πίνακα 4.10. Και σε αυτή την περίπτωση, παρατηρούμε ότι η προτεινόμενη μεθοδολογία ανιχνεύει επιτυχώς την ύπαρξη περισσοτέρων του ενός διακριτών θεμάτων στο έγγραφο.

Το σύνολο των εννοιών που περικλείονται στο έγγραφο μαζί με ένα βαθμό σημαντικότητας ανήκουν στο σύνολο των *Εννοιών* που ορίστηκε στην υποενότητα 4.4.7.3 ως το αλφάριθμο της ασαφούς βάσης γνώσης, ενώ οι ισχυρισμοί που δηλώνονται για το έγγραφο αυτό παρουσιάζονται στον Πίνακα 4.11.

Έννοια	Βαθμός
Επαφη	0.5
Λαχτισμα	0.5
Φαουλ	0.5
Γκολ	0.75
Πεναλτι	0.67
Δοκαρι	0.67
Ομαδα	0.67
Λιγκα	0.75
Νικητης	0.67
Περιλαμβανει	0.75
Ποικιλλει	0.75
Σκορ	0.5

Πίνακας 4.11: Ασαφείς ισχυρισμοί εννοιών για το έγγραφο  $d_2$

Προκειμένου να κατηγοριοποιήσουμε το έγγραφο σε μια θεματική κατηγορία, χρησιμοποιείται η υπηρεσία εξαγωγής συμπερασμάτων (reasoning service) του υψηλότερου χαμηλότερου ορίου (greatest lower bound). Διεξάγονται ερωτήματα υψηλότερου χαμηλότερου ορίου για όλες τις ορισμένες θεματικές κατηγορίες. Οι θεματικές κατηγορίες, στις οποίες κατηγοριοποιείται το έγγραφο  $d_2$ , παρουσιάζονται στον Πίνακα 4.12. Παρατηρούμε ότι, μιας και το  $d_2$  συμμετέχει στις έννοιες Ομάδα με βαθμό 0.67, Γκολ με βαθμό 0.75 και Πέναλτι με βαθμό 0.67, κατηγοριοποιείται στην κατηγορία Ποδόσφαιρο με βαθμό 0.67, που είναι το χαμηλότερο επιτρεπτό όριο για τη συμμετοχή σε αυτό. Είναι επιπλέον ταξινομημένο στην κατηγορία Αθλητικά με τον ίδιο βαθμό, μιας και η κατηγορία Ποδόσφαιρο είναι υποκατηγορία της κατηγορίας Αθλητικά.

Έννοια	Βαθμός
Ποδοσφαιρο	0.67
Αθλητικα	0.67

Πίνακας 4.12: Κατηγοριοποίηση Εγγράφου  $d_2$

Η ανωτέρω βάση γνώσης μπορεί να χρησιμοποιηθεί, για να ταξινομήσει οποιοδήποτε πολυμεσικό έγγραφο σε μια κατηγορία με έναν βαθμό συμμετοχής, με τη

#### *Κεφάλαιο 4. Εξόρυξη δεδομένων και κατηγοριοποίηση*

χρησιμοποίηση της υπηρεσίας εξαγωγής συμπερασμάτων (reasoning service) του υψηλότερου χαμηλότερου ορίου (greatest lower bound), η οποία αξιοποιεί την ασάφεια. Η προσέγγισή μας αξιολογήθηκε σε μια ποικιλία από διάφορα πολυμεσικά έγγραφα, παράγοντας, όπως είδαμε, αξιόλογα αποτελέσματα και καταδεικνύοντας με τον τρόπο αυτό την ισχυρή δυναμική της.

□



# Κεφάλαιο 5

## Προσωποποιημένη πρόσβαση

### 5.1 Εισαγωγή

Η προσωποποίηση του πολυμεσικού περιεχομένου στοχεύει κυρίως στη βοήθεια του τελικού χρήστη σε περιβάλλοντα τεράστιων όγκων δεδομένων. Έρχεται, δηλαδή, να συμβάλλει στις δραστηριότητες του εκάστοτε χρήστη κατά τις περιπτώσεις εκείνες όπου ο τελευταίος στηρίζεται σε μεγάλο ποσοστό σε ισχυρά εργαλεία ανάκτησης και δυνατότητες αυτόματης βοήθειας για να ικανοποιήσει τις ανάγκες του. Ο γενικός στόχος της είναι να αναπτυχθεί μια κατά το δυνατόν αυτόνομη μονάδα, η οποία θα προσφέρει στους χρήστες του πολυμεσικού περιεχομένου αυτόματη εξαγωγή των σημασιολογικών προτιμήσεών τους, καθώς και ενισχυμένες, εξατομικευμένες ικανότητες αναζήτησης και ανάκτησης του περιεχομένου. Ένας από τους στόχους αυτής της προσπάθειας είναι, εν γένει, η προσφορά σημασιολογικών απόψεων (views) για το υπάρχον πολυμεσικό περιεχόμενο, μέσω της χρήσης κατάλληλων βάσεων γνώσεων και θεματικών περιοχών, αλλά και η προσωποποίηση των απόψεων εκείνων σύμφωνα με το εννοιολογικό πλαίσιο που καθορίζεται από το προφίλ των μεμονωμένων χρηστών. Από το τελευταίο γίνεται σαφές, ότι η απαιτούμενη σημασιολογική ερμηνεία των παραπάνω στηρίζεται σε μεγάλο ποσοστό στο εννοιολογικό πλαίσιο, το οποίο εξαρτάται με τη σειρά του από το συγκεκριμένο προφίλ που εξάγεται.

Όπως ακριβώς η σημασιολογική γνώση μιας θεματικής περιοχής χρησιμοποιείται κατά την ανάλυση πολυμεσικού περιεχομένου (βλ. ενότητα 2.3.1) για να εκλεπτύνει και να συμπεράνει καινούρια μεταδεδομένα από τα χαρακτηριστικά γνωρίσματα του πολυμεσικού περιεχομένου σε διάφορα επίπεδα, έτσι και εδώ, οι οντολογίες [243] μπορούν να χρησιμοποιηθούν για να εκλεπτύνουν και να προβλέψουν καινούρια ενδιαφέροντα χρηστών, μέσω της εξαγωγής συμπερασμάτων από τις σχέσεις και τα γεγονότα που περιλαμβάνουν τις έννοιες που θεωρούνται ότι είναι προτιμητέες από τους χρήστες. Η βασισμένη στο εννοιολογικό πλαίσιο προσωποποίηση αντιπροσωπεύει το γεγονός ότι δεν είναι πάντοτε όλες οι προτιμήσεις χρηστών σχετικές με την εκάστοτε αναζήτηση [59]. Το ποιο υποσύνολο προτιμήσεων θα πρέπει να είναι ενεργό κάθε φορά εξαρτάται κατά ένα μεγάλο βαθμό από το εννοιολογικό πλαίσιο των τρέχουσων δραστηριοτήτων και στόχων του χρήστη. Σε αυτό το πλαίσιο, το εννοιολογικό πλαίσιο παρουσιάζεται εδώ ως το σύνολο των αντικειμένων ή των θεμάτων που είναι σχετικά με τις τρέχουσες ενέργειες των χρηστών, ενώ η προσωποποίηση που βασίζεται σε αυτό αντιπροσωπεύει την απόρριψη εκείνων των προτιμήσεων που δεν συσχετίζονται τελικά με το τρέχον εννοιολογικό πλαίσιο του εκάστοτε χρήστη.

Έτσι, στο κεφάλαιο αυτό επιχειρούμε να παρουσιάσουμε τις βασικές αρχές των

ερευνητικών προσπαθειών μας αναφορικά με την παραπάνω διαδικασία της προσωποποίησης του πολυμεσικού περιεχομένου. Η τελευταία είναι, όπως θα δούμε, άρρηκτα συνδεδεμένη με την έννοια του εννοιολογικού πλαισίου, η οποία σε αυτό το ερευνητικό πεδίο λαμβάνει δύο διαφορετικές υποστάσεις, αυτή του ταξινομικού εννοιολογικού πλαισίου και αυτή του εννοιολογικού πλαισίου κατά το χρόνο εκτέλεσης. Στο πλαίσιο αυτό αρχικά θα καθορίσουμε την απαιτούμενη αναπαράσταση γνώσης και θα εξετάσουμε ζητήματα που αφορούν τον ορισμό, την εξαγωγή και τη χρήση προτιμήσεων και κατατομών (προφίλ) χρηστών. Θα παρουσιάσουμε την επίδραση του εννοιολογικού πλαισίου σε θέματα ανάκτησης προσωποποιημένης πληροφορίας, αλλά και κατάταξης των αποτελεσμάτων της ανάκτησης. Τέλος, θα εμβαθύνουμε σε μια πρακτική εφαρμογή προσωποποιημένης ηλεκτρονικής εκπαίδευσης, συνοψίζοντας τα όσα θεωρητικά θα παρουσιασθούν στο κεφάλαιο αυτό.

### 5.1.1 Υπάρχουσα κατάσταση - σχετικές εργασίες

Αν επιχειρούσαμε να ορίσουμε με απλό τρόπο την έννοια της προσωποποίησης, θα λέγαμε ότι αποτελεί μια τεχνολογία που επιτρέπει σε ένα σύστημα να ταιριάζει το διαθέσιμο πολυμεσικό περιεχόμενο, τις εφαρμογές και τις διαθέσιμες μορφές ανθρώπινης αλληλεπίδρασης με αυτό, τόσο με στατικές και ρητώς δεδηλωμένες, όσο και με δυναμικές, προτιμήσεις ενός χρήστη. Με άλλα λόγια, η προσωποποίηση έρχεται να κάνει τη ζωή πιο εύκολη και πιο απλή για τον τελικό χρήστη ενός τέτοιου συστήματος, φροντίζοντας για την κάλυψη των πραγματικών αναγκών και ενδιαφερόντων του. Με τον τρόπο αυτό αποφεύγονται περιπτώσεις, όπου ο χρήστης “πελαγώνει” από την επιστροφή μακροσκελών λιστών αποτελεσμάτων κατά την αναζήτηση εικόνων ή/και videos ή του επιστρέφονται άσχετες πληροφορίες, όταν προσπαθεί π.χ. να προσπελάσει τις τελευταίες ειδήσεις της ημέρας.

Ο απώτερος στόχος της προσωποποίησης είναι να καταστεί δυνατή η προσφορά στον τελικό χρήστη πολυμεσικού περιεχομένου, το οποίο να βρίσκεται πολύ κοντά στις επιθυμίες του. Ποικίλες μέθοδοι έχουν προταθεί τα τελευταία χρόνια, για να εκπληρώσουν αυτό τον στόχο. Π.χ., μια διαδεδομένη μέθοδος είναι η εφαρμογή ενός είδους φιλτραρίσματος του περιεχομένου, το οποίο φιλτράρισμα επιλέγει το εκάστοτε κατάλληλο περιεχόμενο από ένα σύνολο διαθέσιμου περιεχομένου, ανάλογα με τις προτιμήσεις ενός χρήστη [63]. Οι προτιμήσεις και κατ' επέκταση τα προφίλ των χρηστών είναι δυνατόν να μελετηθούν και να αναγνωρισθούν, είτε μέσω της άμεσης παροχής πληροφοριών από τους χρήστες για τα ενδιαφέροντά τους, είτε μέσω της εκπαίδευσης του συστήματος με την πάροδο του χρόνου και καθώς ο χρήστης αλληλεπιδρά με αυτό. Αναμφίβολα, η συλλογή και αναπαράσταση των προτιμήσεων των χρηστών αποτελεί από μόνη της ένα μεγάλο πεδίο έρευνας [97], γνωστό και ως μοντελοποίηση χρηστών (user modeling).

Αν και σύμφωνα με τις τελευταίες ερευνητικές τάσεις η θεώρηση των πραγμάτων μέσω μιας ενοποιημένης και βασισμένης σε οντολογίες άποψης τείνει προς την θεωρητικά ιδανική άποψη του κόσμου, τα προφίλ και τα ενδιαφέροντα των χρηστών αποτελούν ένα από τα χαρακτηριστικότερα παραδείγματα μεγεθών, που μετά βίας τοποθετούνται κάτω από την ομπρέλα μιας σαφούς και ευδιάχριτης έννοιας. Είναι κοινός τόπος, πλέον, στο χώρο της προσωποποίησης ότι οι προτιμήσεις των χρηστών αλληλοσχετίζονται μεταξύ τους και είναι, εν γένει, πολυδιάστατες, χρονικά εξαρτημένες, αλλά και εξαρτώμενες από τον εκάστοτε στόχο· συνοδεύονται, δε, από διαφορετικούς βαθμούς, οι οποίοι είναι δυναμικοί και σχετίζονται με μια ευρεία ποικιλία παραγόντων

## Κεφάλαιο 5. Προσωποποιημένη πρόσθαση

που άπτονται άμεσα ή έμμεσα του εννοιολογικού πλαισίου. Εκτός από τα παραπάνω, η αβεβαιότητα (uncertainty), που απαντάται εγγενώς στην αναπαράσταση ή/και την πρόβλεψη των τάσεων των χρηστών μέσα σε ένα τέτοιο σύστημα λογισμικού, αποτελεί έναν σημαντικό επιπρόσθετο παράγοντα, που πρέπει να ληφθεί υπόψη. Κατά συνέπεια, οι βασισμένοι στο εννοιολογικό πλαίσιο όροι είναι εξίσου δύσκολο να καθοριστούν και να γίνουν κατανοητοί με τρόπο που να είναι απαλλαγμένος από την αβεβαιότητα και την ανακρίβεια. Είναι γεγονός ότι η σημασιολογία του περιεχομένου παίζει σημαντικό ρόλο κατά τη διαδικασία της προσωποποίησης και αποτελεί μια ιδιαίτερα σύνθετη έννοια. Προκειμένου η τελευταία να περιγραφεί με φορμαλισμό και με ένα ολοκληρωμένο και αδιαμφισβήτητο τρόπο, χρειάζεται να δανειστεί περαιτέρω πληροφορίες από το εννοιολογικό πλαίσιο, ούτως ώστε να λάβει μια ακριβή ερμηνεία. Η προαναφερθείσα αβεβαιότητα αυξάνει σημαντικά, όταν οι σημασιολογικές περιγραφές εξάγονται με αυτόματα μέσα, π.χ. μέσω τεχνικών ανάλυσης πολυμεσικού ή μη περιεχομένου. Τέλος, ακόμα και όταν η σημασία είναι σαφής, οι σχέσεις μεταξύ των εννοιών της πραγματικής ζωής είναι πολύ συχνά ένα ζήτημα βαθμού και επομένως μοντελοποιούνται καλύτερα χρησιμοποιώντας ασαφείς σχέσεις.

Λαμβάνοντας υπόψη όλα τα παραπάνω, η προσέγγισή μας συμπληρώνει τη βασισμένη σε οντολογίες προοπτική με ασαφείς έννοιες για την αναπαράσταση των προτιμήσεων των χρηστών, του εννοιολογικού πλαισίου των χρηστών, της σημασιολογίας του περιεχομένου και των σχέσεων μεταξύ των εννοιών. Οι προτεινόμενες μέθοδοι για την εξαγωγή προφίλ χρηστών και την εξατομικευμένη ανάκτηση περιεχομένου με βάση το εννοιολογικό πλαίσιο βασίζονται στις αρχές των ασαφών συνόλων και της ασαφούς σχεσιακής άλγεβρας και εκμεταλλεύονται καινούριες και ήδη διαθέσιμες τεχνικές της ερευνητικής περιοχής, οι οποίες θεωρούνται κατάλληλες, για να εξετάσουν τα προβλήματα που περιλαμβάνουν ασαφή μεγέθη [139], [286]. Η προτεινόμενη μοντελοποίηση με βάση τις αρχές της ασάφειας και τις ασαφείς σημασιολογικές δυαδικές σχέσεις που προέρχονται από το πρότυπο MPEG-7 και η συνακόλουθη υλοποίησή της, αποτελεί πρωτότυπη ερευνητική εργασία και συνεισφορά στο ερευνητικό πεδίο.

Σε αυτό το πλαίσιο, όπου η ανοχή στις ανακριβείς περιγραφές αποτελεί μια αναγνωρισμένη υπόθεση εργασίας, η μοντελοποίηση του εννοιολογικού πλαισίου παίζει έναν βασικό ρόλο στην εκμετάλλευση του βαθμού ασάφειας που περιλαμβάνεται στη γενικότερη θεώρηση. Τα σχετικά μοντέλα που έχουν αναπτυχθεί σε παλαιότερες εργασίες (με βάση π.χ. την ανατροφοδότηση που στηρίζεται στους χρήστες [51], [136], [137], [208], [224], [276], τις προτιμήσεις χρηστών [54], [96], [131], [159], [172], το περικλείον περιβάλλον [45], [114], [122] και τις καταστάσεις στόχου [66], τις χωρικές σχέσεις μεταξύ των αντικειμένων σε μια εικόνα [72], [182], τις γλωσσικές σχέσεις μεταξύ των λέξεων σε ένα κείμενο [90], [264], τα θέματα παρασκηνίου [111], κ.α.) θα μπορούσαν να μας φανούν χρήσιμα και να μας βοηθήσουν να χειριστούμε την αβεβαιότητα. Η δική μας συμβολή, όμως, στο χώρο αποτελείται από την πρότασή μας για διεύρυνση των προγενέστερων εργασιών του χώρου με βάση την εκμετάλλευση της οντολογικής πληροφορίας ως πηγής του εννοιολογικού πλαισίου και/ή ως βοήθειας για την συσχέτιση διαφορετικών τμημάτων περιεχομένου κατά τη διαδικασία της ανάκτησης. Η πρόσθετη σημασιολογία (ακριβής ταξινόμηση, ρητές σχέσεις μεταξύ των εννοιών) παρέχει μια πλούσια πηγή πρόσθετης γνώσης, επιτρέποντας σημαντικές βελτιώσεις σε σύγκριση με τα αποτελέσματα που μπορούν να επιτευχθούν με την χρήση ανεξάρτητων και απλών λέξεων-κλειδιών.

Ο στόχος του εμπλουτισμού των μοντέλων και των μεθόδων ανάκτησης πληρο-

φορίας προς την κατεύθυνση των μοντέλων, που είναι ενήμερα για το εννοιολογικό πλαίσιο (context-aware), έχει αποσπάσει τελευταία την προσοχή και το ενδιαφέρον της ερευνητικής κοινότητας. Επίσης, θεωρείται ως το βασικό εργαλείο, προκειμένου να αντιμετωπιστεί η συνεχής αύξηση της πληροφορίας παγκοσμίως (όποια μορφή κι αν παίρνει η αύξηση σε σχέση με αποθήκες πληροφοριών, δίκτυα, χρήστες, κ.α.), κάτι το οποίο μπορεί να θέσει στο μέλλον σοβαρές προκλήσεις, όσον αφορά στις τρέχουσες τεχνολογίες αναζήτησης. Σε ένα ιδιαίτερα απαιτητικό και ολοένα αυξανόμενο και ανταγωνιστικό περιβάλλον τα ερωτήματα των χρηστών συχνά δεν επαρκούν σε μια σύγχρονη μηχανή αναζήτησης, για να καλύψουν τις ανάγκες πληροφόρησης με έναν αποτελεσματικό τρόπο που να ικανοποιεί πλήρως τις προσδοκίες των χρηστών. Σε μια δύσκολη και σύνθετη περίπτωση ο χρήστης μπορεί να πρέπει να τροποποιήσει το ερώτημά του και να προσπελάσει πολλαπλά επίπεδα ταξινομημένων αποτελεσμάτων/εγγράφων, προτού μπορέσει να ικανοποιήσει την πρωταρχική ανάγκη πληροφόρησής του. Σε ένα τέτοιο σενάριο διαλογικής ανάκτησης οι πληροφορίες που είναι διαθέσιμες στο σύστημα ανάκτησης είναι συνήθως περισσότερες από απλά και μόνο την τρέχουσα ερώτηση του χρήστη και τη συλλογή των διαθέσιμων εγγράφων. Εν γένει, το αυθαίρετο ιστορικό της αλληλεπίδρασης πρέπει να καταστεί διαθέσιμο στο σύστημα ανάκτησης, συμπεριλαμβανομένων και των παλαιότερων ερωτήσεων των χρηστών, των εγγράφων που ο κάθε χρήστης έχει επιλέξει να δει, ακόμη και του τρόπου που έχει επιλέξει ένας χρήστης να έχει πρόσβαση σε ένα έγγραφο.

Τέλος, θα πρέπει να σημειωθεί ότι σήμερα υπάρχουν στη βιβλιογραφία αλγόριθμοι ανάκτησης που εξαρτώνται σε μικρό ή μεγάλο βαθμό από το εννοιολογικό πλαίσιο. Σε αντίθεση με τη δική μας προσέγγιση, όμως, οι περισσότεροι βασίζονται σε στατιστικά γλωσσικά πρότυπα, για να συνδυάσουν τις προηγούμενες ερωτήσεις των χρηστών και τις επιλεγμένες περιλήψεις των εγγράφων με την τρέχουσα ερώτηση, προκειμένου να επιτύχουν καλύτερη ταξινόμηση των εγγράφων [31], [90], [111], [149]. Η συσχετιστική ανατροφοδότηση [208] και η πιο πρόσφατη, υπονοούμενη ανατροφοδότηση [51], [136], [224], [276], εκμεταλλεύονται τη βασισμένη στο εννοιολογικό πλαίσιο εισαγωγή του χρήστη (user input) ως μία πηγή πληροφοριών, για να συμπληρώσει τις ρητές ερωτήσεις των χρηστών και να καθοδηγήσει τη διαδικασία της ανάκτησης.

### 5.1.2 Δομή κεφαλαίου

Η προσέγγισή μας στο κεφάλαιο αυτό χωρίζεται στους εξής βασικούς άξονες:

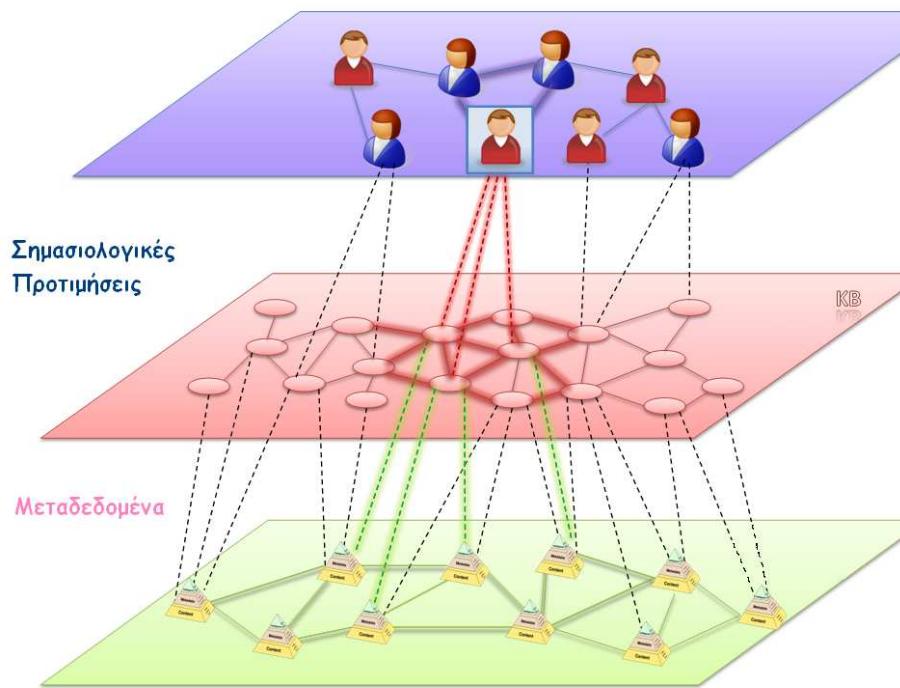
- **Υπάρχουσα κατάσταση - σχετικές εργασίες:** Στην ενότητα αυτή παρουσιάστηκε η συντομία μια επισκόπηση της υπάρχουσας κατάστασης στο χώρο της προσωποποιημένης πρόσβασης σε πολυμεσικό περιεχόμενο. Επίσης, μελετώνται και αναφέρονται οι τελευταίες ερευνητικές εξελίξεις και προσεγγίσεις για την αντιμετώπιση των σχετικών ερευνητικών προβλημάτων, καταδεικνύοντας ταυτόχρονα τις βασικές διαφορές με την προτεινόμενη προσέγγισή μας.
- **Εννοιολογικό πλαίσιο και προσωποποίηση:** Παρουσιάζεται μια αναλυτική επισκόπηση της χρήσης και αξιοποίησης των διαφόρων μορφών του εννοιολογικού πλαισίου κατά την προσωποποίηση πολυμεσικού περιεχομένου και δίνεται έμφαση στις βασικές αρχές της δικής μας προσέγγισης. Ειδικότερα, αναφερόμαστε στο διττό ρόλο που λαμβάνει το εννοιολογικό πλαίσιο στην προσωποποίηση, δηλ. το ταξινομικό εννοιολογικό πλαίσιο και το εννοιολογικό πλαίσιο που σχηματίζεται κατά το χρόνο εκτέλεσης.

## Κεφάλαιο 5. Προσωποποιημένη πρόσβαση

- **Αναπαράσταση γνώσης:** Στην ενότητα αυτή περιγράφονται οι μαθηματικοί φορμαλισμοί της ερευνητικής προσέγγισής μας χρησιμοποιώντας βασικά στοιχεία από τη θεωρία της σχεσιακής ασαφούς άλγεβρας. Παρουσιάζονται, επίσης, οι θεμελιώδεις αρχές της οντολογίας και των σημασιολογικών σχέσεων που θα χρησιμοποιηθούν στη διαδικασία.
- **Αναπαράσταση ασαφούς εννοιολογικού πλαισίου:** Γίνεται μια αναλυτική παρουσίαση των 2 μορφών του εννοιολογικού πλαισίου που είναι κατάλληλες και χρήσιμες για την προσωποποίηση, ενώ παρουσιάζονται κάποιες βασικές διαφορές και ομοιότητες μεταξύ τους.
- **Εξαγωγή προφίλ χρηστών:** Η ενότητα αυτή περιγράφει τη μεθοδολογία που ακολουθείται προκειμένου να εξαχθούν τα ενδιαφέροντα των χρηστών από το ιστορικό χρήσης του συστήματος ανάκτησης πληροφορίας μέσω της αλληλεπίδρασης των τελευταίων με αυτό. Περιγράφεται αναλυτικά η διαδικασία με την οποία είναι σε θέση κανείς να καταλήξει στην εξαγωγή των προτιμήσεων των χρηστών, ξεκινώντας από τα διαθέσιμα πολυμεσικά έγγραφα. Στο τέλος της ενότητας δίδεται ένα εκτενές παράδειγμα εφαρμογής των προηγουμένων στο πεδίο της ηλεκτρονικής μάθησης.
- **Ανάκτηση και κατάταξη προσωποποιημένης πληροφορίας:** Η ενότητα αυτή εξηγεί την επίδραση του εννοιολογικού πλαισίου πάνω στις προτιμήσεις των χρηστών κατά το χρόνο ανάκτησης, προκειμένου να επιτευχθεί η με βάση το εννοιολογικό πλαίσιο προσωποποίηση των αποτελεσμάτων της ανάκτησης, καθώς και η αποδοτικότερη κατάταξή τους.
- **Πειραματικά αποτελέσματα:** Στην ενότητα αυτή παρουσιάζονται αποτελέσματα από τρία διαφορετικά σύνολα πειραμάτων.
- **Ένα σενάριο χρήσης:** Εδώ παρουσιάζονται δύο στοιχειώδη σενάρια χρήσης της προτεινόμενης μεθοδολογίας, τα οποία και καταδεικνύουν τον τρόπο με τον οποίο εφαρμόζονται οι τεχνικές αυτές στην πράξη καθώς και την υψηλή αποδοτικότητά τους.

## 5.2 Εννοιολογικό Πλαίσιο και Προσωποποίηση

Ανάμεσα στους σχετικούς με την προσωποποίηση στόχους της παρούσης διατριβής βρίσκεται η δημιουργία κατάλληλων δομών γνώσης που θα βασίζονται σε οντολογίες, καθώς και εργαλείων αυτόματης ανάλυσης πολυμεσικού περιεχομένου. Η τριπλή αυτή συνέργια και η σημασιολογική σύνδεση του περιεχομένου, της βάσης γνώσης και των χρηστών μπορεί να περιγραφεί σχηματικά με το ακόλουθο Σχήμα 5.1. Τα προβλήματα που θα εξεταστούν στη συνέχεια περιλαμβάνουν την αναπαράσταση των προφίλ των χρηστών, χρησιμοποιώντας την εκφραστική δύναμη των οντολογιών και των μεταδεδομένων που χρησιμοποιούνται για να περιγράψουν το πολυμεσικό περιεχόμενο, την ανάπτυξη τεχνικών αυτόματης μάθησης, προκειμένου να ενημερωθούν οι προτιμήσεις των χρηστών και σύμφωνα με το εννοιολογικό τους πλαίσιο και την εκάστοτε συμπεριφορά τους, και τον ορισμό μιας σειράς βασικών αλγορίθμων και μέτρων για την υποστήριξη της προσωποποίησης περιεχομένου σε συστήματα ανάκτησης πληροφορίας.



**Σχήμα 5.1:** Σημασιολογική σύνδεση μεταξύ χρηστών και πολυμεσικού περιεχομένου μέσω των εννοιών μιας βάσης γνώσης.

Αν ανατρέξει, λοιπόν, κανείς σε ένα δημοφιλές λεξικό, όπως π.χ. το Merriam-Webster [304], μπορεί να δει ότι το εννοιολογικό πλαίσιο ορίζεται ως “οι αλληλένδετοι όροι στους οποίους κάτι υπάρχει ή εμφανίζεται”. Στο πλαίσιο της έρευνάς μας οι ερωτήσεις που διατυπώνονται από έναν χρήστη ενός συστήματος Ανάκτησης Πληροφορίας (IR) κατά τη διάρκεια μιας διαλογικής συνόδου ανάκτησης πληροφορίας (information retrieval session) προς την κατεύθυνση της ικανοποίησης μιας ανάγκης πληροφόρησής του αποτελούν τα υπό εξέταση εμφανιζόμενα γεγονότα. Οι παρακείμενοι όροι περιλαμβάνουν:

1. Μακροπρόθεσμες προτιμήσεις, που βρίσκονται στο παρασκήνιο και που είτε φανερώνονται ρητά, είτε παρέχονται σιωπηρά από το χρήστη κατά τις προγενέστερες συνόδους του με το σύστημα ανάκτησης πληροφορίας.
2. Την χρονικά σύντομη εστίαση όρων-χρηστών, που παρέχονται σιωπηρά από την αλληλεπίδραση (π.χ. τα διάφορα χλικ) και τις επερωτήσεις του χρήστη κατά τη διάρκεια μιας τρέχουσας συνόδου.
3. Το σημασιολογικό πεδίο (π.χ. τη θεματική περιοχή) των πληροφοριών που αναζητούνται από το χρήστη ή στις οποίες έχει πρόσβαση ο χρήστης, πάλι κατά τη διάρκεια μιας τρέχουσας συνόδου ανάκτησης.

Η έμφαση στην αμοιβαία σχέση στον ανωτέρω ορισμό σχετίζεται άμεσα με τη δική μας θεώρηση των πραγμάτων και αντιμετωπίζεται επισταμένως στη συνέχεια. Πριν προχωρήσουμε σε περαιτέρω λεπτομέρειες της προσέγγισής μας, αξίζει να σημειωθεί ότι το κίνητρο και η ανάπτυξη της έννοιας του εννοιολογικού πλαισίου στο περιβάλλον της προσωποποίησης πολυμεσικού περιεχομένου καθώς και οι σχετικές αλγορίθμικές

## Κεφάλαιο 5. Προσωποποιημένη πρόσβαση

μέθοδοι στηρίζονται σε ένα σύνολο προβλημάτων, υποθέσεων, απόψεων και αποφάσεων σχεδίασης, τα οποία και παρουσιάζονται συνοπτικά στις επόμενες παραγράφους της παρούσας ενότητας του κεφαλαίου 5.

Η έρευνά μας στο πεδίο της προσωποποίησης θα βασισθεί στο ακόλουθο υποθετικό, αλλά συνάμα άκρως ρεαλιστικό, σενάριο ανάκτησης πληροφοριών από ένα σύστημα IR: ένα σύνολο από χρήστες  $U$  αλληλεπιδρά με ένα χώρο ανάκτησης  $D$  μέσω μιας διεπαφής ανάκτησης, η οποία περιλαμβάνει λειτουργίες αναζήτησης και πλοήγησης του πολυμεσικού περιεχομένου. Οι τελευταίες επιτρέπουν την επιθεώρηση των αποτελεσμάτων αναζήτησης ή την άμεση πλοήγηση στο χώρο ανάκτησης καθώς και την επιλογή και επίδειξη αντικειμένων πληροφορίας. Ο χώρος ανάκτησης  $D$  αποτελείται από αντικείμενα πληροφοριών, τα οποία συνήθως (αν και όχι υποχρεωτικά) περιέχουν μία σεβαστή ποσότητα μη δομημένου ή ημι-δομημένου περιεχομένου (π.χ. κείμενο και αντικείμενα ή/και έγγραφα πολυμέσων). Στα αντικείμενα πληροφοριών προστίθενται σχόλια μεταδεδομένων, τα οποία αποτελούνται από έννοιες, ιδιότητες και τιμές που καθορίζονται σύμφωνα με την οντολογία  $\mathcal{O}$  μιας θεματικής περιοχής, και που αποθηκεύονται κατάλληλα σε μια Βάση Γνώσης (knowledge base - KB).

Σύμφωνα με αυτή τη θεώρηση οι χρήστες ενδιαφέρονται για εν γένει διαφορετικά θέματα και “πράγματα”. Πολλές από αυτές τις προτιμήσεις μπορούν να είναι ανεξάρτητες από το συγκεκριμένο σώμα της ανάκτησης, παρ' όλα αυτά στο πλαίσιο της έρευνάς μας υποθέτουμε την ύπαρξη ενός υποσυνόλου ενδιαφερόντων χρηστών (το οποίο από εδώ και πέρα θα ονομάσουμε προτιμήσεις  $P$ , από τον αγγλικό όρο “preferences”)), το οποίο συνδέεται με κάποιο τρόπο με το σώμα αυτό. Το είδος και η μορφή μιας τέτοιας σύνδεσης ποικίλλουν· πιο συγκεκριμένα, μπορούν να σχετίζονται με εξωτερικές ιδιότητες των αντικειμένων πληροφοριών, όπως είναι η φύση τους (π.χ. εγκυκλοπαιδικές ιδιότητες, δημοσιογραφικές, επιστημονικές, κ.α.), ο σκοπός τους (αποκτώντας τη γνώση, βρίσκοντας τις κατευθύνσεις, διασκεδάζοντας), η ποιότητά τους, η εμπορική τους αξία, η προηγούμενη εμπειρία χρηστών με τα αντικείμενα και διάφορα άλλα. Εκτός από αυτές τις ιδιότητες είναι γνωστό ότι μια ιδιαιτέρως ενδιαφέρουσα (από την άποψη της αξίας και της γενικότητάς της) πλευρά των προτιμήσεων των χρηστών σχετίζεται με την εσωτερική σημασιολογία που μεταβιβάζεται από τα αντικείμενα πληροφοριών, κάτι το οποίο αναμφίβολα αποτελεί αντικείμενο της επικρατούσας έρευνας στον τομέα της Ανάκτησης Πληροφορίας (IR) [213].

Ακολουθώντας αυτή την πεπατημένη, ορίζουμε το  $P$  ως ένα σύνολο εννοιών (concepts) που μπορεί να βρεθούν ή να αναφερθούν μέσα στα στοιχεία του χώρου ανάκτησης  $D$ . Πέρα από τις ακατέργαστες λέξεις-κλειδιά και τους περιγραφές πολυμέσων, που χρησιμοποιούνται συνήθως κατά την συμβατική ανάκτηση πληροφοριών ως τα δομικά στοιχεία της σημασιολογικής αναπαράστασης για τις ανάγκες των χρηστών, εμείς ερευνούμε τη χρήση οντολογιών (ontologies) στο χώρο αυτό, καθώς θεωρούνται ο πυρήνας μιας ποιοτικά υψηλότερης εκφραστικότητας και της ακρίβειας τέτοιων περιγραφών [54], [96], [138], [202]. Στην προσέγγιση που ακολουθείται στα πλαίσια της διδακτορικής διατριβής οι προτιμήσεις των χρηστών  $P$  περιγράφονται ως ένα σύνολο σημασιολογικών οντοτήτων, για τις οποίες ενδιαφέρεται ο χρήστης σε διαφορετικό βαθμό για κάθεμία και όπου ακριβώς εξαιτίας της ύπαρξης βαθμών έχει χρησιμοποιηθεί η ίδια οντολογία, όπως και για το σχολιασμό του σώματος ανάκτησης. Το γεγονός αυτό παρέχει ένα αρκετά ακριβές, εκφραστικό και ενοποιημένο μοντέλο αναπαράστασης, στο οποίο, τόσο τα ενδιαφέροντα των χρηστών, όσο και η ίδια η έννοια του περιεχομένου αντιπροσωπεύονται στον ίδιο χώρο, στον οποίο και μπορούν να συγχριθούν με άνεση [55].

Η έννοια του πλαισίου είναι διτή στην προσέγγισή μας και μπορεί να λάβει δύο όψεις· η πρώτη εφαρμόζεται κατά τον χρόνο εξαγωγής των προφίλ και η δεύτερη κατά τον χρόνο ανάκτησης των πληροφοριών. Και στις δύο φάσεις το εννοιολογικό πλαίσιο αποτελείται από μια ασαφή περιοχή μιας οντολογίας θεματικών περιοχών και χρησιμοποιείται για να βοηθήσει, να στρέψει ή να επεκτείνει την ερμηνεία των ενδιαφερόντων των χρηστών από το σύστημα προς μια συγκεκριμένη σημασιολογική περιοχή. Κατά την φάση εξαγωγής των προφίλ (η οποία λαμβάνει χώρα απογραμμικά (offline) και πάντως όχι ταυτόχρονα με την αλληλεπίδραση με τους τελικούς χρήστες) το σύστημα ανιχνεύει τα πρότυπα των προτιμήσεων των χρηστών με την ανάλυση ενός μεγάλου συνόλου καταγεγραμμένων ενεργειών και αιτημάτων. Το σύστημα αναλύει τις σημασιολογικές σχέσεις, για να βρει κοινές θεματικές βάσεις για τα διαφορετικά υποσύνολα του ιστορικού χρήστης, ακολουθώντας μια προσέγγιση βασισμένη στις τεχνικές συσταδοποίησης που περιγράφηκαν αναλυτικά στα κεφάλαια 4.2.1-4.2.3. Η έννοια του εννοιολογικού πλαισίου που εφαρμόζεται εδώ είναι ταξινομική και περιοριστικής φύσης και χρησιμοποιείται, για να μειώσει το θόρυβο και την αβεβαιότητα, με το να αγνοεί τις άσχετες ενέργειες χρηστών και με το να εστιάζει στις πιο συνεκτικές, από τις οποίες είναι τελικά σοφότερο και ασφαλέστερο να προβλεφθούν τα ενδιαφέροντα των χρηστών. Το ταξινομικό εννοιολογικό πλαίσιο αναφέρεται σε οτιδήποτε είναι σημασιολογικά κοινό μεταξύ ενός συνόλου στοιχείων, κάτι το οποίο μπορεί να αναφέρεται στην κοινή σημασία ενός συνόλου εννοιών ή στο γενικό θέμα ενός εγγράφου, αντίστοιχα. Το ταξινομικό εννοιολογικό πλαίσιο μιας έννοιας είναι αυτό που παρέχει την αληθινά προοριζόμενη σημασία της, όταν χρησιμοποιείται μια οντολογική αναπαράσταση γνώσης (όπως αυτή που προτείνεται εν τω παρόντι) για να ερμηνευθεί η έννοια ενός αντικειμένου πληροφοριών. Με άλλα λόγια, η αληθινή πηγή πληροφόρησης είναι οι σημασιολογικές κοινωνίες των εννοιών και όχι η καθεμία έννοια ανεξάρτητα. Με τον τρόπο αυτό χρησιμοποιείται η κοινή σημασία των εννοιών για να καθορίσει καλύτερα είτε τα θέματά τους, είτε τις σχετικές προτιμήσεις των χρηστών, στις οποίες πρέπει να αντιστοιχηθούν.

Κατά το χρόνο εκτέλεσης, οι παραπάνω αρχές εφαρμόζονται με έναν ελαφρώς διαφορετικό τρόπο. Ακόμα κι όταν ο χρήστης θεωρείται ότι διαθέτει ένα συνεχές σύνολο ενδιαφερόντων το οποίο έχει προκύψει είτε εκπαιδεύοντας το σύστημα κατά τη φάση εξαγωγής των προφίλ, είτε παρέχεται από τον ίδιο το χρήστη με το χέρι, είναι γεγονός ότι τέτοιου είδους ενδιαφέροντα δεν παραμένουν στατικά, αλλά μεταβάλλονται με την πάροδο του χρόνου και εξαρτώνται από την εκάστοτε κατάσταση. Κατά συνέπεια, το προτεινόμενο μοντέλο διακρίνει μεταξύ ενός συνεχούς τμήματος προτιμήσεων χρηστών, το οποίο καθορίζεται εκ των προτέρων και το οποίο εξελίσσεται με έναν πιο αργό ρυθμό, και ενός προσωρινού, ειδικού τμήματος, το οποίο εξαρτάται από το “ζωντανό” εννοιολογικό πλαίσιο μέσα στο οποίο συμμετέχει και κινείται ο χρήστης κατά τη διατύπωση των στόχων της ανάκτησης περιεχομένου. Στην προσέγγισή μας, το τελευταίο τμήμα λαμβάνει τη μορφή μιας ρητής, δυναμικής αναπαράστασης του “ζωντανού”, σημασιολογικού εννοιολογικού πλαισίου ως ένα ασαφές σύνολο εννοιών μιας θεματικής περιοχής, το οποίο και κατασκευάζεται από τα οντολογικά στοιχεία που περιλαμβάνονται στις διάφορες ενέργειες των χρηστών. Αυτή η αναπαράσταση του εννοιολογικού πλαισίου κατά το χρόνο εκτέλεσης χρησιμοποιείται σε συνδυασμό με τις συνεχείς προτιμήσεις των χρηστών, προκειμένου να υπολογιστεί ένα συγκεκριμένο και βασισμένο στο εννοιολογικό πλαίσιο σύνολο, το οποίο θα περιγράφει τα συνολικά ενδιαφέροντα των χρηστών. Ο υπολογισμός αυτού του συνόλου επιτυγχάνεται σε δύο βήματα, αποτελούμενο από μια βασισμένη στο εννοιολογικό πλαίσιο

επέκταση (contextual expansion), η οποία ακολουθείται στη συνέχεια από μία συστολή (contextual contraction). Στο πρώτο βήμα ολοκληρώνονται τα αρχικά σύνολα προτιμήσεων και εννοιολογικού πλαισίου, έτσι ώστε να διαμορφώσουν σημασιολογικά συνεπή υπερσύνολα (με βάση τη θεωρία της ασαφούς σύνθεσης και ένωσης συνόλων). Κατά τη συστολή καθορίζεται μια ασαφής τομή των υπερσυνόλων αυτών, όπως άλλωστε θα περιγραφεί και παρακάτω στο παρόν κείμενο. Ο τελικός στόχος είναι να χρησιμοποιηθούν τα βασισμένα στο εννοιολογικό πλαίσιο ενδιαφέροντα των χρηστών για να επιτύχουμε μια καλύτερη, ακριβέστερη και πιο αξιόπιστη εξατομίκευση των αποτελεσμάτων αναζήτησης, που ανακτήθηκαν από το σύστημα ως απάντηση στις ερωτήσεις που τέθηκαν από τους χρήστες του.

### 5.3 Αναπαράσταση Γνώσης

Γενικά, η έννοια του εννοιολογικού πλαισίου σχετίζεται άμεσα και έντονα με την έννοια των οντολογιών, αφού μια οντολογία θεωρείται ότι αποτελεί ένα πλαίσιο για αναπαράσταση γνώσης [226] ή με άλλα λόγια το πρότυπο λειτουργίας οντοτήτων και αλληλεπιδράσεων σε κάποια ιδιαίτερη σημασιολογική περιοχή γνώσης. Σύμφωνα με αυτό, μια **οντολογία** είναι ένα σύνολο εννοιών (όπως τα αντικείμενα, τα γεγονότα, και οι σχέσεις) που καθορίζονται με κάποιο τρόπο (όπως η φυσική γλώσσα), προκειμένου να δημιουργήσουν ένα συμφωνημένο λεξιλόγιο για την ανταλλαγή πληροφοριών και να διαμορφώσουν μια προσπάθεια για τη μοντελοποίηση εννοιών του πραγματικού κόσμου. Επιπλέον, το **εννοιολογικό πλαίσιο** καθορίζει την προοριζόμενη έννοια κάθε έννοιας, δηλ. μια έννοια που χρησιμοποιείται υπό το πρόσμα διαφορετικών εννοιολογικών πλαισίων είναι δυνατόν να λάβει διαφορετικές ερμηνείες. Συνεπώς, ένας πιθανός τρόπος να εξαχθεί και να χρησιμοποιηθεί το εννοιολογικό πλαίσιο κατά τη διαδικασία της προσωποποίησης περιεχομένου είναι να καθοριστεί μέσω ασαφών σημασιολογικών σχέσεων. Για να δοθεί μια περαιτέρω επεξήγηση, μια απλή οντολογία είναι δυνατόν να περιγραφεί γενικά ως:

$$O = \{C, \{R_{c_i, c_j}\}\}, i, j = 1..n, i \neq j \quad (5.1)$$

$$R_{c_i, c_j} : C \times C \rightarrow \{0, 1\} \quad (5.2)$$

όπου το  $O$  αναπαριστά μία συσχετισμένη με μία συγκεκριμένη σημασιολογική περιοχή οντολογία, το  $C$  είναι το σύνολο όλων των πιθανών εννοιών που περιγράφει και  $R_{c_i, c_j}$  είναι η σχέση μεταξύ δύο εννοιών  $c_i, c_j$ .

Αν και εν γένει οι οντολογίες μπορούν να περιέχουν οποιοδήποτε τύπο σχέσεων, στο πλαίσιο της προτεινόμενης διαδικασίας αξιοποίησης του εννοιολογικού πλαισίου κατά την προσωποποίηση εστιάζουμε μόνο σε έναν τύπο σχέσεων, τις ταξινομικές (taxonomic ή ordering) σχέσεις, καθότι θεωρείται ότι αυτές καλύπτουν πλήρως τις ανάγκες διαχείρισης του πολυμεσικού περιεχομένου. Όπως συζητείται και στο [3], η χρήση ταξινομικών σχέσεων είναι απαραίτητη για τον προσδιορισμό του εννοιολογικού πλαισίου ενός πολυμεσικού εγγράφου. Κατά συνέπεια, το επόμενο βήμα είναι η ουσιαστική εκμετάλλευση των πληροφοριών που περιλαμβάνονται σε αυτές τις ταξινομικές σχέσεις μέσα στην οντολογία. Η παραπάνω δήλωση περιγράφει σαφώς μια χαρακτηριστική μορφή ιεραρχιών ταξινόμησης (taxonomy hierarchies), εντούτοις οι τελευταίες είναι στη γενική περίπτωση σαφείς και, δεδομένου ότι οι έννοιες της πραγματικής ζωής διέπονται σχεδόν πάντα από ασάφεια, εύκολα γίνεται κατανοητό ότι

μοντελοποιούνται καλύτερα χρησιμοποιώντας ασαφείς σχέσεις, όπως άλλωστε περιγράφεται και στο [4].

Κατά συνέπεια, οφείλουμε να εισάγουμε ένα νέο ορισμό του παραπάνω οντολογικού μοντέλου γνώσης. Το νέο μοντέλο είναι βασισμένο στις σχέσεις που μπορούν να μοντελοποιηθούν ως ασαφείς ταξινομικές σχέσεις διαταγής και μπορούν να συνοψιστούν στον ακόλουθο ορισμό οντολογίας:

$$O_F = \{C, \{r_{c_i, c_j}\}\}, i, j = 1..n, i \neq j \quad (5.3)$$

$$F(R_{c_i, c_j}) = r_{c_i, c_j} : C \times C \rightarrow [0, 1] \quad (5.4)$$

όπου το  $O_F$  εκφράζει την νέα οντολογία και η  $F(R_{c_i, c_j}) = r_{c_i, c_j}$  την ασαφή σχέση μεταξύ δύο εννοιών  $c_i, c_j$ .

Όπως προαναφέρθηκε, στο [3] προτάθηκε ένα σύνολο βασικών σχέσεων βασισμένων στο εννοιολογικό πλαίσιο, το οποίο μπορεί να χρησιμοποιηθεί, για να μοντελοποιήσει ταξινομικές ιεραρχίες εννοιολογικών πλαισίων, ενώ οι σχέσεις αυτές καθαυτές αναπαριστούν βαθύτερα σημασιολογικά νοήματα από μια απλή ταξινομική σχέση. Οι σχέσεις που αποφαίνονται κατάλληλες για την προσωποποίηση συνοψίζονται στις ακόλουθες:

- *Ειδίκευση*: Αντιστοιχεί στην ευρέως διαδεδομένη σχέση υποκατηγοριοποίησης.
- *Πλαίσιο*: Μοντελοποιεί το γεγονός ότι μία έννοια παρέχει το πλαίσιο μιας άλλης έννοιας.
- *Όργανο*: Μία έννοια αποτελεί το όργανο για μία άλλη έννοια.
- *Μέρος*: Περιγράφει ότι μια έννοια είναι μέρος μιας άλλης έννοιας.
- *Ασθενής*: Μία έννοια είναι το υποχείριο μιας άλλης έννοιας.
- *Θέση*: Μία έννοια είναι η θέση μιας άλλης έννοιας.
- *Πράκτορας*: Κάποια έννοια είναι ο πράκτορας μιας άλλης έννοιας.

Δεν υποστηρίζουμε, φυσικά, ότι οι παραπάνω σχέσεις είναι επαρκείς, για να μοντελοποιήσουν κάθε τύπο εννοιολογικού πλαισίου, αλλά αντιθέτως θεωρούμε ότι είναι αρκετά γενικές, για να αποτελέσουν μια χρήσιμη βάση για το μοντέλο γνώσης με το οποίο ασχολούμαστε στην έρευνά μας. Όπως θα δούμε παρακάτω, μια σύνθετη ταξινομική σχέση μπορεί να κατασκευαστεί από τις παραπάνω βασικές σχέσεις. Οι τελευταίες θα αποτελέσουν τη βάση για την ταξινομία του εννοιολογικού πλαισίου, καθώς όλες οι ανωτέρω σχέσεις που υπάρχουν στην οντολογία ορίζονται στο σύνολο των εννοιών της. Δεδομένου δε ότι οι σχέσεις αυτές περιλαμβάνουν πληροφορίες σημασιολογικής φύσης μεταξύ των εννοιών της οντολογίας, αποτελούν τη βασική μορφή αποθηκευμένης γνώσης. Η ασάφεια των σχέσεων παίζει επίσης έναν σημαντικό ρόλο. Παραδείγματος χάριν, η αυξανόμενη ή μειούμενη τιμή της σχέσης ειδίκευσης, όπως αυτή καθορίζεται μεταξύ δύο εννοιών, υποδηλώνει το βαθμό στον οποίο η αποθηκευμένη γνώση δείχνει ότι ένα στιγμιότυπο μιας έννοιας *b* σε ένα πολυμεσικό έγγραφο ικανοποιεί ένα αίτημα για μια άλλη έννοια *a* σε μια ερώτηση αναζήτησης. Παρομοίως, όλες οι άλλες σχέσεις μπορούν να συσχετισθούν άμεσα με έναν βαθμό ικανοποίησης του τελικού χρήστη.

## Κεφάλαιο 5. Προσωποποιημένη πρόσβαση

Βασιζόμενοι στις παραπάνω σχέσεις κατασκευάζουμε την ακόλουθη σχέση  $T$ :

$$T = Tr^t(\bigcup_i r_i^{p_i}), \quad p_i \in \{-1, 1\}, \quad i = 1 \dots n \quad (5.5)$$

όπου η τιμή του  $p_i$  καθορίζεται από τη σημασιολογία της κάθε σχέσης που χρησιμοποιείται για την κατασκευή της  $T$  (π.χ. από τη σειρά των ορισμάτων  $a, b$  στον Πίνακα 7.1), καθώς μερικές από τις σχέσεις ίσως χρειαστεί να αντιστραφούν, πριν χρησιμοποιηθούν στη δημιουργία της σχέσης  $T$ . Το μεταβατικό κλείσιμο στην εξισωση (5.5) είναι απαραίτητο, προκειμένου αυτή να είναι ταξινομική, καθώς η ένωση μεταβατικών σχέσεων δεν είναι απαραιτήτως μεταβατική. Για το σκοπό της ανάλυσης πολυμεσικών εγγράφων η σχέση  $T$  δημιουργήθηκε με τη χρήση ενός συνόλου από ασαφείς ταξινομικές σχέσεις, των οποίων η σημασιολογία καθορίζεται στο πρότυπο MPEG-7 (βλ. [296] και Παράρτημα B) και οι οποίες συνοψίζονται στον Πίνακα 7.1.

**Πίνακας 5.1:** Ασαφείς ταξινομικές σχέσεις που χρησιμοποιήθηκαν για την κατασκευή της σχέσης  $T$ .

'Όνομα	Αντίστροφο	Σύμβολο	Ερμηνεία	Παράδειγμα	
				<i>a</i>	<i>b</i>
Έξειδικευση	Γενίκευση	$Sp(a, b)$	το <i>b</i> είναι μια εξειδίκευση της ερμηνείας του <i>a</i>	Θηλαστικό	σκύλος
Μέρος	Μέρος Του	$P(a, b)$	το <i>b</i> είναι ένα μέρος του <i>a</i>	Νέα Υόρκη	Μανχάταν
Παράδειγμα	Παράδειγματου	$Ex(a, b)$	το <i>b</i> είναι ένα παράδειγμα του <i>a</i>	πρόεδρος	Bill Clinton
'Οργανο	'Οργανο Του	$Ins(a, b)$	το <i>b</i> είναι ένα όργανο του <i>a</i> πραγματοποιείται από το <i>a</i>	κοπή	μαχαίρι
Θέση	Θέση Του	$Loc(a, b)$	το <i>b</i> είναι η θέση του <i>a</i>	συναυλία	σκηνή
Ασθενής	Ασθενής Του	$Pat(a, b)$	το <i>b</i> επηρεάζεται από ή υπομένει την δράση του <i>a</i>	δίνω	βιβλίο
Ιδιότητα	Ιδιότητα Του	$Pr(a, b)$	το <i>b</i> είναι μια ιδιότητα του <i>a</i>	μπανάνα	ωριμότητα

Στην περίπτωση αυτή η  $T$  λαμβάνει τη μορφή [268]:

$$T = Tr^t(Sp \cup P^{-1} \cup Ins \cup Pr^{-1} \cup Pat \cup Loc \cup Ex) \quad (5.6)$$

Με βάση τη σημασιολογία των παραπάνω σχέσεων είναι εύκολο να παρατηρήσει κανείς ότι η  $T$  είναι ιδιαίτερη για τον προσδιορισμό των θεματικών κατηγοριών με τις οποίες είναι δυνατόν να σχετίζεται μία έννοια, καθώς επίσης και για την εκτίμηση της κοινής ερμηνείας (common meaning), δηλ. μιας μορφής εννοιολογικού πλαισίου, ενός συνόλου από έννοιες. Όλες οι σχέσεις που χρησιμοποιούνται για την κατασκευή της  $T$  είναι μερικώς διατακτικές (partial ordering). Παρ' όλ' αυτά, δεν υπάρχει καμία ένδειξη ότι η ένωσή τους εξακολουθεί να παραμένει αντισυμμετρική, μια ιδιότητα που απαιτείται προκειμένου αυτή η σχέση να είναι και ταξινομική. Απεναντίας, η  $T$  μπορεί, εν γένει, να ποικίλει από μια μερικώς διατεταγμένη σχέση μέχρι μια σχέση ισοδυναμίας. Αυτή η τελευταία παρατήρηση είναι ιδιαίτερα σημαντική, καθώς η απόλυτη συμμετρία και η απόλυτη αντισυμμετρία θα πρέπει συχνά να εγκαταλείπονται κατά τη μοντελοποίηση της πραγματικής ζωής. Παρ' όλ' αυτά, η σημασιολογία των χρησιμοποιημένων σχέσεων, καθώς επίσης και τα αποτελέσματα των πειραμάτων μας, καταδεικνύουν ότι η  $T$  βρίσκεται πολύ κοντά σε μία αντισυμμετρική σχέση. Επομένως, δικαιούμαστε να τη θεωρούμε και να τη χρησιμοποιούμε από εδώ και εμπρός ως μία σχεδόν-διατακτική (quasi-ordering) ή σχεδόν-ταξινομική (quasi-taxonomic) σχέση.

## 5.4 Αναπαράσταση Ασαφούς Εννοιολογικού Πλαισίου

Το προτεινόμενο μοντέλο προσωποποίησης που βασίζεται στο εννοιολογικό πλαίσιο μπορεί να εκφραστεί κατά τρόπο επίσημο με τη χρήση βασικών στοιχείων σημασιολογικής ερμηνείας, όπως είναι οι έννοιες (concepts), οι σχέσεις (relations) μεταξύ

των εννοιών και οι θεματικές κατηγορίες (topics), οι οποίες και χτίζουν μια δομή οντολογίας. Δεδομένου ότι οι σχέσεις μεταξύ των εννοιών της πραγματικής ζωής διέπονται σχεδόν πάντα από βαθμούς και, όπως είδαμε στην προηγούμενη ενότητα 5.3, μοντελοποιούνται καλύτερα χρησιμοποιώντας ασαφείς σχέσεις, η εν τω παρόντι ακολουθούμενη προσέγγιση βασίζεται στην επίσημη μεθοδολογία και το μαθηματικό συμβολισμό της ασαφούς σχεσιακής άλγεβρας [286], [139]. Οι βασικές αρχές της τελευταίας συνοψίζονται στην ενότητα 3.2, αλλά και στην έννοια μιας οντολογίας, που παρουσιάζεται στην ενότητα 3.4 της διατριβής.

Η τελική μορφή της προτεινόμενης ασαφούς οντολογίας περιέχει έννοιες, σχέσεις και θεματικές κατηγορίες και μπορεί να τυποποιηθεί με επίσημο τρόπο ως εξής:

$$\mathcal{O} = \{S, R, Z\} \quad (5.7)$$

Στην παραπάνω εξίσωση (5.7) το  $\mathcal{O}$  αναπαριστά μία ασαφή οντολογία, το  $S$  είναι το κλασικό σύνολο εννοιών που περιγράφεται από την οντολογία, το  $R$  είναι το κλασικό σύνολο των ασαφών σημασιολογικών σχέσεων μεταξύ των εννοιών και το  $Z$  είναι το κλασικό σύνολο από θεματικές κατηγορίες που είναι διαθέσιμες στην  $\mathcal{O}$ , όπου  $Z \subset S$ .

Με δεδομένο το σύνολο όλων των ασαφών συνόλων στο  $S$ ,  $\mathcal{F}_S$ , τότε  $F \in \mathcal{F}_S$ . 'Εστω ότι  $\mathcal{U}$  είναι το σύνολο όλων των χρηστών  $\widehat{\mathcal{U}}$  στο προτεινόμενο πλαίσιο, δηλαδή ένας χρήστης  $\widehat{u} \in \mathcal{U}$ . 'Εστω, τώρα, ότι  $\mathcal{P}$  είναι το σύνολο όλων των προτιμήσεων χρηστών και  $\mathcal{P}_{\mathcal{O}}$  είναι το σύνολο όλων των προτιμήσεων χρηστών στην  $\mathcal{O}$ . Τότε  $\mathcal{P}_{\mathcal{O}} \subset \mathcal{F}_S$  και  $\mathcal{P}_{\mathcal{O}} = \mathcal{F}_Z \subset \mathcal{F}_S$ , ενώ το  $P_{\widehat{u}} \in \mathcal{P}_{\mathcal{O}}$  αναπαριστά μία συγκεκριμένη προτίμηση χρήστη και περιγράφεται από ένα ασαφές σύνολο στο  $Z$ . Καθώς το γεγονός ότι η προτίμηση ενός χρήστη σχετίζεται άμεσα με τον χρήστη είναι ολοφάνερο, θα παραλείψουμε από τούδε και στο εξής το  $\widehat{u}$  ως δείκτη και θα χρησιμοποιούμε απλώς το  $P$ , εφόσον η ερμηνεία του είναι ξεκάθαρη.

'Εστω ότι  $\mathcal{C}_{\mathcal{O}}$  καταδεικνύει το σύνολο όλων των εννοιολογικών πλαισίων στην  $\mathcal{O}$ ,  $\mathcal{C}_{\mathcal{O}} \subseteq \mathcal{F}_S$ . 'Οπως και στην περίπτωση των προτιμήσεων χρηστών, έτσι και εδώ  $\widehat{C} \in \mathcal{C}_{\mathcal{O}}$  είναι ένα ασαφές σύνολο ορισμένο πάνω στο κλασικό σύνολο των εννοιών  $S$  και συμβολίζει το εννοιολογικό πλαίσιο κατά το χρόνο εκτέλεσης. Κατά συνέπεια, το  $C$  καταδεικνύει το σύνολο όλων των εννοιολογικών πλαισίων κατά το χρόνο εκτέλεσης. 'Εστω, ακόμη, ότι το κλασικό σύνολο εννοιών που χαρακτηρίζουν το κλασικό (ταξινομικό) εννοιολογικό πλαίσιο συμβολίζεται ως  $C'$ , ενώ το ασαφές αντίστοιχό του  $C$  παρέχει το ταξινομικό εννοιολογικό πλαίσιο με την έννοια ενός ασαφούς συνόλου από έννοιες στο  $S$ ,  $C \in \mathcal{C}_{\mathcal{O}}$ . Τέλος, έστω ότι το  $\mathcal{D}$  είναι το κλασικό σύνολο όλων των διαθέσιμων αντικειμένων πληροφορίας (π.χ. κείμενα ή πολυμεσικά έγγραφα), το  $S_d$  είναι το ασαφές σύνολο εννοιών συσχετισμένων με το  $d \in \mathcal{D}$ , ενώ  $S_d \in \mathcal{F}_S$ , και  $I(s, d)$  είναι το κατά τα γνωστά κατασκευασμένο σημασιολογικό ευρετήριο μεταξύ εγγράφων και εννοιών [6] (βλ. ενότητα 4.4.5).

Στο τελευταίο βήμα ορίζουμε την επίδραση του εννοιολογικού πλαισίου στις προτιμήσεις των χρηστών ως μία αντιστοίχιση  $\Phi : \mathcal{P} \times \widehat{C} \rightarrow \mathcal{P}$ , έτσι ώστε για όλα τα  $p \in \mathcal{P}$  και  $c \in \widehat{C}$ ,  $p \models \Phi(p, c)$ . Σε αυτό το πλαίσιο η συνεπαγωγή (entailment)  $p \models q$  σημαίνει ότι όποιο συμπέρασμα μπορεί να εξαχθεί από το  $q$  μπορεί να εξαχθεί και από το  $p$ . Για παράδειγμα, δεδομένου ενός χρήστη  $\widehat{u} \in \mathcal{U}$ , αν  $P_{\widehat{u}} = q$  υπονοεί ότι ο  $\widehat{u}$  “ενδιαφέρεται για το  $x$ ” (ό,τι κι αν σημαίνει αυτό), τότε ο  $\widehat{u}$  θα ενδιαφέρεται και για το  $x$ , αν η προτίμησή του ήταν το  $p$ .

#### 5.4.1 Ασαφείς σημασιολογικές σχέσεις κατά την προσωποποίηση

Προκειμένου να ορίσουμε, να εξαγάγουμε και να χρησιμοποιήσουμε τόσο το ταξινομικό εννοιολογικό πλαίσιο, όσο και το εννοιολογικό πλαίσιο κατά το χρόνο εκτέλεσης από ένα σύνολο εννοιών, θα βασιστούμε στη σημασιολογία των ασαφών σημασιολογικών τους σχέσεων. Όπως συζητήθηκε ήδη στην ενότητα 3.2, μία ασαφής δυαδική σχέση στο  $S$  ορίζεται ως μία συνάρτηση  $R_i : S \times S \rightarrow [0, 1]$ ,  $i = 1, \dots, M$ . Η αντίστροφη σχέση μιας σχέσης  $R_i(x, y)$ ,  $x, y \in S$  ορίζεται ως  $R_i^{-1}(x, y) = R_i(y, x)$ . Θα χρησιμοποιούμε τον συμβολισμό προθέματος (prefix notation)  $R_i(x, y)$  για τις ασαφείς σχέσεις παρά τον ένθετο συμβολισμό (infix notation)  $xR_iy$ , καθώς ο πρώτος θεωρείται καταλληλότερος για τον απλό αναγνώστη [177].

Η λειτουργία της ένωσης των ασαφών σχέσεων μπορεί να γενικευτεί για  $M$  σχέσεις. Αν  $R_1, R_2, \dots, R_M$  είναι ασαφείς σχέσεις στο  $S \times S$ , τότε η ένωσή τους  $R^u$  είναι μία σχέση ορισμένη στο  $S \times S$ , τέτοια ώστε για όλα τα  $(x, y) \in S \times S$ ,  $R^u(x, y) = u(R_i(x, y))$ . Το μεταβατικό κλείσιμο μιας σχέσης  $R_i$  είναι η μικρότερη μεταβατική σχέση που περιέχει την αρχική σχέση και έχει τα λιγότερα δυνατά μέλη. Στη γενική περίπτωση το κλείσιμο μιας σχέσης είναι η μικρότερη επέκταση της σχέσης, η οποία έχει μια συγκεκριμένη ιδιότητα, όπως είναι η ανακλαστικότητα, η συμμετρία ή η μεταβατικότητα, σύμφωνα με τον Klir [139]. Το sup- $t$  μεταβατικό κλείσιμο  $Tr^t(R_i)$  μιας ασαφούς σχέσης  $R_i$  δίδεται με επίσημο τρόπο ως εξής:

$$Tr^t(R_i) = \bigcup_{j=1}^{\infty} R_i^{(j)} \quad (5.8)$$

όπου  $R_i^{(j)} = R_i \circ R_i^{(j-1)}$  και  $R_i^{(1)} = R_i$ . Κατά τα γνωστά, αν η  $R_i$  είναι ανακλαστική, τότε το μεταβατικό κλείσιμο της δίνεται από την  $Tr^t(R_i) = R_i^{(n-1)}$ , όπου  $n = |S|$  [139].

Με βάση τις σχέσεις  $R_i$  κατασκευάζουμε την ακόλουθη συνδυασμένη σχέση  $T$ , η οποία και χρησιμοποιείται στον ορισμό του ταξινομικού εννοιολογικού πλαισίου,  $C$ :

$$T = Tr^t(\bigcup_i R_i^{p_i}), \quad p_i \in \{-1, 0, 1\}, \quad i = 1 \dots M \quad (5.9)$$

όπου η τιμή του  $p_i$  καθορίζεται από τη σημασιολογία κάθε επιμέρους σχέσης  $R_i$  που χρησιμοποιείται στην κατασκευή της  $T$ . Ειδικότερα::

- $p_i = 1$ , αν η σημασιολογία της  $R_i$  υπονοεί ότι πρέπει να χρησιμοποιηθεί ως έχει
- $p_i = -1$ , αν η σημασιολογία της  $R_i$  υπονοεί ότι θα πρέπει να θεωρηθεί το αντίστροφό της
- $p_i = 0$ , αν η σημασιολογία της  $R_i$  δεν επιτρέπει την συμμετοχή της στην κατασκευή της συνδυασμένης σχέσης  $T$ .

Το μεταβατικό κλείσιμο στην εξίσωση (5.9) απαιτείται, προκειμένου η  $T$  να είναι ταξινομική, καθώς, όπως έχει ήδη ειπωθεί, η ένωση μεταβατικών σχέσεων δεν είναι απαραίτητα μεταβατική, ανεξάρτητα από την ασαφή  $t$ -conorm που χρησιμοποιείται. Σε αυτό το πλαίσιο, μία ασαφής σημασιολογική σχέση καθορίζει, για κάθε στοιχείο  $s \in S$ , το ασαφές σύνολο των προγόνων και των απογόνων του. Για παράδειγμα, αν η γνώση μας δηλώνει ότι ο “Αμερικανικός εμφύλιος” έγινε πριν τον “1ο Παγκόσμιο Πόλεμο” και ο “1ος Παγκόσμιος Πόλεμος” έγινε πριν τον “2ο Παγκόσμιο Πόλεμο”, τότε δεν είναι σίγουρο ότι η γνώμη μας δηλώνει επίσης ότι ο “Αμερικανικός εμφύλιος” έγινε πριν τον “2ο Παγκόσμιο Πόλεμο”. Το μεταβατικό κλείσιμο έρχεται να

διορθώσει αυτή την ανακολουθία. Με παρόμοιο τρόπο, με το να πραγματοποιούνται τα σχετικά κλεισίματα σε σχέσεις που συσχετίζουν ζεύγη εννοιών του ιδίου συνόλου, εξασφαλίζουμε την συνέπειά τους.

Στη συνέχεια, με βάση ένα διαφορετικό σύνολο από σχέσεις  $R_i$ , κατασκευάζουμε τη συνδυασμένη σχέση  $\widehat{T}$ , προκειμένου να την χρησιμοποιήσουμε για τον καθορισμό του εννοιολογικού πλαισίου κατά το χρόνο εκτέλεσης  $\widehat{C}$ :

$$\widehat{T} = \bigcup_i (R_i^{\widehat{p}_i}), \quad \widehat{p}_i \in \{0, 1\}, \quad i = 1 \dots \widehat{M} \quad (5.10)$$

Έχοντας στο μιαλό την ανάλυση κειμένου, αλλά και εγγράφων πολυμέσων, η σχέση  $T$  κατασκευάστηκε με τη χρήση ενός μικρού συνόλου ασαφών ταξινομικών σχέσεων, των οποίων η σημασιολογία προέρχεται από το πρότυπο MPEG-7 [296] και συνοψίζονται στον Πίνακα 7.1. Από την άλλη πλευρά, η σχέση  $\widehat{T}$  κατασκευάστηκε με τη χρήση ολόκληρου του διαθέσιμου συνόλου σχέσεων από τη βάση γνώσης. Η προσέγγιση αυτή είναι ιδιαίτερη για την εφικτή επέκταση των δύο διαφορετικών ειδών του εννοιολογικού πλαισίου και των προτιμήσεων χρήστη που ακολουθούνται εν τω παρόντι. αρχικά, κατά τη θεώρηση του γενικού προφίλ χρήστη δίνεται έμφαση στη σημασιολογία των αφηρημένων εννοιών υψηλού επιπέδου, ενώ κατά τη διάρκεια της φάσης ανάκτησης απαιτείται πρόσθετη ακρίβεια και μια πιο συγκεκριμένη άποψη, καθώς λαμβάνει χώρα η επέκταση των προτιμήσεων χρηστών κατά το χρόνο εκτέλεσης. Το τελευταίο απαιτεί τη χρησιμοποίηση όλων των διαθέσιμων πληροφοριών από τη βάση γνώσης. Βέβαια, όπως υπονοεί και η κατασκευή της σχέσης  $\widehat{T}$ , πριν από την εφαρμογή της βασισμένης στην ταξινομία τεχνικής επέκτασης, απαιτείται ένα ενδιάμεσο βήμα για την αφάρεση πιθανών κύκλων, που οφείλονται στη χρησιμοποίηση όλων των σχέσεων και των αντιστρόφων τους.

Τέλος, η ασάφεια των παραπάνω σχέσεων έγκειται στο εξής: υψηλές τιμές του  $Sp(a, b)$ , για παράδειγμα, υπονοούν ότι η ερμηνεία του  $b$  πλησιάζει την ερμηνεία του  $a$ , ενώ καθώς το  $Sp(a, b)$  μειώνεται, η ερμηνεία του  $b$  γίνεται ολοένα και πιο στενή απ' ότι η ερμηνεία του  $a$ . Μια παρόμοια ερμηνεία δίνεται και στην ασάφεια των υπολοίπων σημασιολογικών σχέσεων. Με βάση τους ασαφείς ρόλους και τις σημασιολογικές ερμηνείες των  $R_i$  είναι εύκολο να δει κανείς ότι και οι δύο παραπάνω συνδυασμένες σχέσεις (5.9) και (5.10), τις συνδυάζουν με έναν ευθύνη και με νόημα τρόπο, χρησιμοποιώντας την αντίστροφη λειτουργικότητα των επιμέρους σχέσεων, όπου αυτό απαιτείται σημασιολογικά. Ειδικότερα, η σχέση  $T$  χρησιμοποιεί το ακόλουθο υποσύνολο σχέσεων (βλ. ενότητα 5.3):

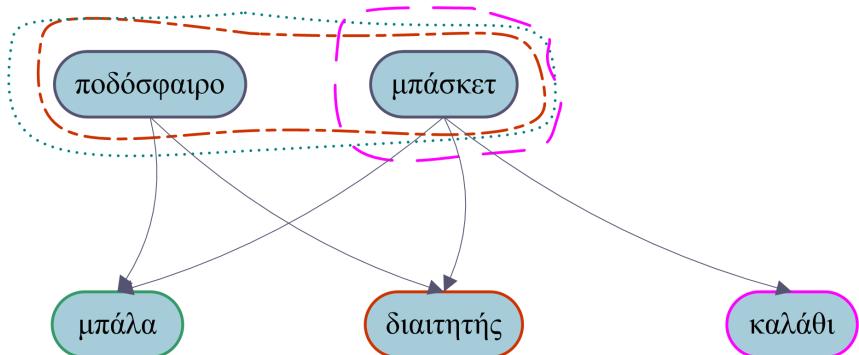
$$T = Tr^t(Sp \cup P^{-1} \cup Ex \cup Ins \cup Loc^{-1} \cup Pat \cup Pr^{-1}) \quad (5.11)$$

ενώ η σχέση  $\widehat{T}$  χρησιμοποιεί τελικά τις σχέσεις (βλ. Παράρτημα B):

$$\widehat{T} = Tr^t(Sp \cup P^{-1} \cup Sim \cup Ex \cup Sub^{-1}) \quad (5.12)$$

#### 5.4.2 Ταξινομικό εννοιολογικό πλαίσιο

Η παραπάνω σχέση  $T$  είναι εξαιρετικής σημασίας, καθώς επιτρέπει τον ορισμό, την εξαγωγή και τη χρήση του ταξινομικού εννοιολογικού πλαισίου ενός συνόλου από έννοιες. Στηριζόμενοι στη σημασιολογία της σχέσης  $T$ , ορίζουμε το κλασικό ταξινομικό εννοιολογικό πλαίσιο  $C'$  μιας μοναδικής έννοιας  $s \in S$  ως το σύνολο των προγόνων της που παρέχεται από την σχέση  $T$  στην οντολογία.



Σχήμα 5.2: Παράδειγμα ταξινομικού εννοιολογικού πλαισίου.

Όπως παρατηρούμε στο Σχήμα 5.2, οι έννοιες ποδόσφαιρο και μπάσκετ είναι οι πρόγονοι των εννοιών μπάλα και διαιτητής στη σχέση  $T$ , ενώ η έννοια μπάσκετ είναι ο μόνος πρόγονος της έννοιας καλάθι. Πιο αυστηρά, ακολουθώντας το γνωστό συμβολισμό υπερ/υπό-συνόλων από την ασαφή σχεσιακή άλγεβρα, το κλασικό εννοιολογικό πλαίσιο  $C'(s)$  μιας μοναδικής έννοιας  $s \in S$  δίνεται από την:

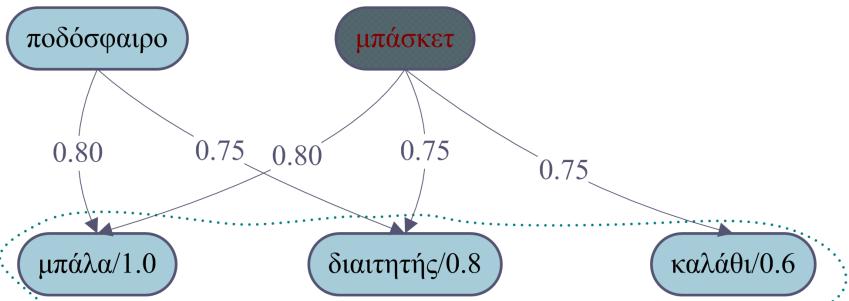
$$C'(s) = T_{\leq}(s), \quad (5.13)$$

Υποθέτοντας ξανά ότι ένα σύνολο από έννοιες  $S$  είναι κλασικό (σαφές), δηλαδή ότι όλες οι έννοιες ανήκουν στο σύνολο με βαθμό ίσο με 1, το εννοιολογικό πλαίσιο ολόκληρου του συνόλου, το οποίο είναι και πάλι ένα σύνολο από έννοιες, μπορεί να ορισθεί απλά ως το σύνολο των κοινών προγόνων τους:

$$C'(S) = \bigcap C'(s_i), \quad s_i \in S \quad (5.14)$$

Όπως φαίνεται στο Σχήμα 5.2, η έννοια μπάσκετ είναι ο μόνος κοινός πρόγονος και των τριών εννοιών μπάλα, διαιτητής και καλάθι στη σχέση  $T$ , δηλαδή το μπάσκετ είναι το εννοιολογικό πλαίσιο των εννοιών μπάλα, διαιτητής και καλάθι.

Καθώς λαμβάνουμε υπόψη περισσότερες έννοιες, το εννοιολογικό πλαίσιο γίνεται στενότερο, δηλαδή περιλαμβάνει λιγότερες έννοιες και σε μικρότερους βαθμούς. Όταν ο παραπάνω ορισμός του εννοιολογικού πλαισίου επεκταθεί στην περίπτωση των ασαφών συνόλων εννοιών (Σχήμα 5.3), το κλασικό ταξινομικό εννοιολογικό πλαίσιο  $C'$  αντικαθίσταται από το ασαφές αντίστοιχό του, δηλαδή το ασαφές ταξινομικό εννοιολογικό πλαίσιο  $C$ . Προφανώς, η σημασιολογική ερμηνεία του ασαφούς εννοιολογικού πλαισίου παραμένει η ίδια, όπως και στην κλασική θεώρηση, δηλαδή η παραπάνω ιδιότητα πρέπει και πάλι να ισχύει.



Σχήμα 5.3: Παράδειγμα ασαφούς ταξινομικού πλαισίου.

Το εννοιολογικό πλαίσιο  $C$  του κανονικού ασαφούς συνόλου  $F$  στο  $S$  υπολογίζεται ως:

$$C(F) = \bigcap_i \mathcal{K}(s_i), s_i \in F \quad (5.15)$$

όπου  $\mathcal{K}(s_i)$  είναι το “δεδομένο” εννοιολογικό πλαίσιο του  $s_i$ , δηλαδή το εννοιολογικό πλαίσιο της έννοιας όταν λαμβάνεται υπόψη ο βαθμός συμμετοχής της στο σύνολο. Το  $\mathcal{K}(s_i)$  ορίζεται ως:

$$\mathcal{K}(s_i) \doteq C'(s_i) \cup cp(S \cdot F(s_i)) \quad (5.16)$$

όπου  $S \cdot F(s_i)$  είναι το γινόμενο του συνόλου  $S$  με τον βαθμό συμμετοχής  $F(s_i)$ , όπως καθορίστηκε στην ενότητα 3.2, το σύμβολο “ $\doteq$ ” καταδεικνύει την εξ ορισμού ισότητα, το  $cp$  είναι ένα ασαφές συστροφικό (involutive) συμπλήρωμα και  $C'(s_i)$  είναι το κλασικό εννοιολογικό πλαίσιο μιας μοναδικής έννοιας  $s_i$ .

Παρατηρούμε επίσης ότι εξαιτίας της φύσης των ασαφών συνόλων οι ακόλουθες ιδιότητες ισχύουν:

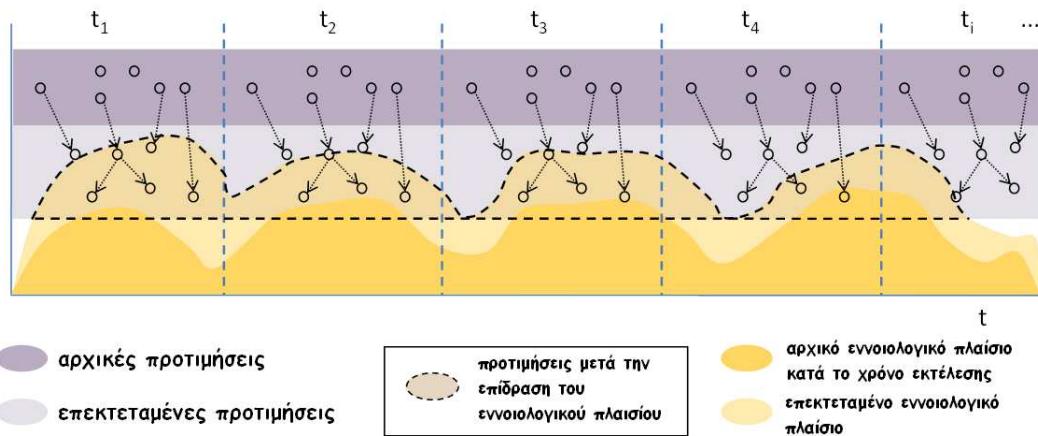
- $F(s_i) = 0 \Rightarrow C(F) = C(F - \{s_i\})$ , δηλαδή κανένας περιορισμός του εννοιολογικού πλαισίου
- $F(s_i) = 1 \Rightarrow C(F) \subseteq C(s_i)$ , δηλαδή ολικός περιορισμός του εννοιολογικού πλαισίου
- το  $C(F)$  μειώνεται μονοτονικά και αναλογικά με το  $F(s_i)$

Θεωρώντας τη σημασιολογία της σχέσης  $T$  και την παραπάνω διαδικασία καθορισμού του εννοιολογικού πλαισίου, είναι εύκολα κατανοητό ότι, όταν οι έννοιες ενός συνόλου σχετίζονται με μεγάλο βαθμό με μια κοινή ερμηνεία, το εννοιολογικό πλαίσιο θα έχει υψηλούς βαθμούς συμμετοχής για τις έννοιες που αναπαριστούν αυτή την κοινή ερμηνεία. Για το λόγο αυτό το ύψος του εννοιολογικού πλαισίου  $h(C(F))$  θα χρησιμοποιηθεί στη συνέχεια ως μία μετρική της σημασιολογικής συσχέτισης των εννοιών στο σύνολο  $F$ . Αυτή η μετρική αναπαριστά επίσης και το βαθμό ομοιότητας των εννοιών στο σύνολο.

#### 5.4.3 Εννοιολογικό πλαίσιο κατά το χρόνο εκτέλεσης

Η σχέση  $\widehat{T}$  είναι επίσης εξαιρετικού ενδιαφέροντος, καθώς επιτρέπει τον ορισμό, την εξαγωγή και τη χρήση του εννοιολογικού πλαισίου κατά το χρόνο εκτέλεσης. Ο σκοπός του τελευταίου είναι κυρίως να βρίσκει τα μονοπάτια που έχουν νόημα και συσχετίζουν τις προτιμήσεις με το εννοιολογικό πλαίσιο. Τα μονοπάτια που λαμβάνονται τελικά υπόψη αποτελούνται από υπάρχουσες σημασιολογικές σχέσεις μεταξύ των εννοιών στην οντολογία της βάσης γνώσης. Όσο πιο μικρά, πιο ισχυρά και πιο πολυάριθμα είναι αυτά τα μονοπάτια σύνδεσης, τόσο πιο συσχετισμένη με το εννοιολογικό πλαίσιο θα πρέπει να θεωρείται μία προτίμηση. Χρησιμοποιώντας αυτά τα σημασιολογικά μονοπάτια, επιλέγεται ένα σύνολο από προτιμήσεις, οι οποίες είναι βαθμονομημένες σύμφωνα με το επίπεδο συσχέτισης με το τρέχον εννοιολογικό πλαίσιο. Το σύνολο αυτό ενημερώνεται δυναμικά και περιέχει πάντοτε τις πιο συσχετισμένες προτιμήσεις που είναι αντιπροσωπευτικές για κάθε χρήστη σύμφωνα με τη σύνοδο του τελευταίου. Με άλλα λόγια, θα πρέπει να θεωρείται ένα προκαθορισμένο χρονικό παράθυρο προτιμήσεων προκειμένου να ληφθεί το εκάστοτε καλύτερο σύνολο προτιμήσεων από αυτές (βλ. Σχήμα 5.4).

## Κεφάλαιο 5. Προσωποποιημένη πρόσβαση



**Σχήμα 5.4:** Χρονικά παράθυρα ταιριάσματος εννοιολογικού πλαισίου και προτιμήσεων χρηστών.

Στην προσέγγιση που ακολουθούμε το εννοιολογικό πλαίσιο κατά το χρόνο εκτέλεσης παίρνει τη μορφή μιας σαφούς, δυναμικής αναπαράστασης του “ζωντανού” σημασιολογικού εννοιολογικού πλαισίου ως ένα ασαφές σύνολο από έννοιες, το οποίο κατασκευάζεται με τη συλλογή στοιχείων της οντολογίας που απαντώνται στις πράξεις των χρηστών. Αυτή η κατά το χρόνο εκτέλεσης αναπαράσταση του εν λόγω εννοιολογικού πλαισίου χρησιμοποιείται σε συνδυασμό με τις στατικές προτιμήσεις των χρηστών, προκειμένου να υπολογιστεί ένα συγκεκριμένο και υπό την επίδραση του εννοιολογικού πλαισίου σύνολο από ενδιαφέροντα χρηστών. Ο υπολογισμός αυτού του συνόλου επιτυγχάνεται σε δύο βήματα, που αποτελούνται αφενός από μια βασισμένη στο εννοιολογικό πλαίσιο επέκταση και αφετέρου από μια συστολή. Στην πρώτη φάση τα αρχικά σύνολα προτίμησης και πλαισίου ολοκληρώνονται, για να διαμορφώσουν σημασιολογικά συνεπή υπερσύνολα, ενώ κατά τη φάση της συστολής καθορίζεται ένα είδος τομής των υπερσυνόλων αυτών. Με τον τρόπο αυτό το σημασιολογικό εννοιολογικό πλαίσιο κατά το χρόνο εκτέλεσης χρησιμοποιείται, για να ενεργοποιήσει διαφορετικά υποσύνολα ενδιαφερόντων χρηστών, έτσι ώστε να επιτευχθεί μια συνοχή με το θεματικό πεδίο των ενεργειών των χρηστών, κατά τέτοιο τρόπο ώστε οι εκτός εννοιολογικού πλαισίου προτιμήσεις να απορρίπτονται. Τέλος, τα ενδιαφέροντα των χρηστών που βρίσκονται υπό την επίδραση του εννοιολογικού πλαισίου χρησιμοποιούνται, για να επιτύχουν μια καλύτερη, ακριβέστερη και πιο αξιοπιστη προσωποποίηση των αποτελεσμάτων της ανάκτησης που επέστρεψε το σύστημα ως απάντηση στις ερωτήσεις των χρηστών του.

Το εννοιολογικό πλαίσιο κατά το χρόνο εκτέλεσης αναπαριστάται ή προσεγγίζεται από ένα ασαφές σύνολο από έννοιες που υπάρχουν μέσα σε μια οντολογία μιας θεματικής περιοχής. Για παράδειγμα, αν ένας χρήστης θέτει ερωτήματα, αναζητεί και διαβάζει για την οικολογική καταστροφή σε μια συγκεκριμένη γεωγραφική περιοχή, το εννοιολογικό πλαίσιο κατά το χρόνο εκτέλεσης μπορεί να αποτελείται από έννοιες της θεματικής περιοχής, όπως η φωτιά, τα τοξικά απόβλητα, ο αέρας, το ποτάμι, η βλάστηση, κ.ο.κ.. Αποτελεί αναμφίβολα μέρος του γενικού πλαισίου, κατάλληλο για την ανάλυση στην προσωποποίηση και μπορεί να οριστεί ως το υπόβαθρο πάνω από το οποίο συμβαίνουν οι δραστηριότητες των χρηστών κατά τη διάρκεια μιας δεδομένης μονάδας του χρόνου. Από αυτή την άποψη τα προβλήματα που πρέπει να αντιμετωπίστούν περιλαμβάνουν τον τρόπο αναπαράστασης αυτού του είδους του εννοιολογικού πλαισίου, τον τρόπο καθορισμού του κατά το χρόνο εκτέλεσης καθώς και τον τρόπο

χρησιμοποίησής του, προκειμένου να επηρεάσει την ενεργοποίηση των προτιμήσεων των χρηστών, να επιβάλει την επίδρασή του σε αυτές και να προβλέψει ή να λάβει υπόψη τη μεταβολή των προτιμήσεων κατά τη διάρκεια του χρόνου (τόσο βραχυπρόθεσμα, όσο και μακροπρόθεσμα).

Στο πλαίσιο ενός συστήματος ανάκτησης περιεχομένου, όπως αυτό που εξετάζουμε στην παρούσα διατριβή, ορίζουμε το σημασιολογικό εννοιολογικό πλαίσιο του χρόνου εκτέλεσης ως το σύνολο των εννοιών που έχουν εμπλακεί άμεσα ή έμμεσα στην αλληλεπίδραση ενός χρήστη  $u$  με το σύστημα κατά τη διάρκεια μιας συνόδου ανάκτησης. Επομένως, σε κάθε χρονικό σημείο  $t$  αναπαριστούμε το πλαίσιο της ανάκτησης  $C(u, t)$  ως ένα ασαφές σύνολο εννοιών με βάρη, όπου για κάθε  $x \in \mathcal{O}$  ορίζεται ένα βάρος  $C_x(u, t) \in [0, 1]$ . Ο χρόνος μετριέται από τον αριθμό των αιτημάτων των χρηστών μέσα σε μια σύνοδο ανάκτησης. Δεδομένου, όμως, ότι το πλαίσιο σχετίζεται πάντοτε και σαφέστατα με κάποιον χρήστη, στη συνέχεια θα παραλείψουμε την μεταβλητή  $u$  και θα αναφερόμαστε σε αυτό ως  $C(t)$  ή απλά  $C$  για συντομία.

Στην προσέγγισή μας το  $C(t)$  κατασκευάζεται ως ένας συσσωρευτικός συνδυασμός των εννοιών που περιλαμβάνονται σε διαδοχικά αιτήματα χρηστών, κατά τέτοιο τρόπο ώστε η σημασία των εννοιών να εξασθενεί με την πάροδο του χρόνου. Η διαδικασία αυτή μιμείται τη φυσική μετακίνηση του επικέντρου του ενδιαφέροντος των χρηστών κατά τη διάρκεια του χρόνου. Αμέσως μετά το αίτημα ενός χρήστη καθορίζεται ένα σύνολο ζήτησης (request) εννοιών, το οποίο περιέχει το σύνολο των εννοιών που εμπεριέχονται στην ερώτηση που τέθηκε (εάν, φυσικά, το αίτημα αποτελείται από μια ερώτηση - βλ. ενότητα 5.5.2). Στην περίπτωση αυτή οι έννοιες μπορούν να εξαχθούν από μία ερώτηση φυσικής γλώσσας ή από μια βασισμένη στις λέξεις-κλειδιά ερώτηση, χρησιμοποιώντας τις τελευταίες τεχνικές εξαγωγής πληροφοριών [202]. Εάν το αίτημα είναι του τύπου “επισκόπηση εγγράφου”, το παραπάνω σύνολο μπορεί να καθοριστεί από τις σχετικές έννοιες που χαρακτηρίζουν το έγγραφο. Εάν το αίτημα είναι ένα βήμα επαναληπτικής συσχετιστικής ανατροφοδότησης, το παραπάνω σύνολο μπορεί να είναι το μέσο σύνολο που αντιστοιχεί στο σύνολο των εγγράφων που χαρακτηρίζονται ως σχετικά από το χρήστη. Τέλος, παρόμοιες στρατηγικές μπορούν να καθοριστούν για να χτίσουν τα παραπάνω σύνολα εννοιών από αιτήματα πλοήγησης, άλλα και από άλλους τύπους ενεργειών χρηστών κατά τη διαδικασία της ανάκτησης, όμως η περαιτέρω ενασχόληση με αυτές δεν αποτελεί μέρος της παρούσας διατριβής και ο αναγνώστης καλείται να ανατρέξει στη σχετική βιβλιογραφία.

Στη συνέχεια, καθορίζουμε ένα αρχικό ασαφές σύνολο εννοιολογικού πλαισίου  $C(t)$ , συνδυάζοντας το πρόσφατα κατασκευασμένο αίτημα με το εννοιολογικό πλαίσιο  $C(t-1)$  που υπολογίζεται στο αμέσως προηγούμενο βήμα, όπου τα βάρη του εννοιολογικού πλαισίου στο βήμα  $t-1$  μειώνονται αυτόματα κατά έναν παράγοντα αποσύνθεσης  $\xi \in [0, 1]$ . Συνεπώς, σε μία δεδομένη στιγμή  $t$  ενημερώνουμε το  $C_x(t)$  ως:

$$C_x(t) = \xi \cdot C_x(t-1) + (1 - \xi) \cdot R_x(t) \quad (5.17)$$

όπου  $R_x(t)$  είναι το αίτημα του χρήστη. Το παραπάνω μπορεί να φαίνεται παρόμοιο με ένα βήμα σχετικής ανατροφοδότησης [51], [136], εντούτοις, στην περίπτωση αυτή το εννοιολογικό πλαίσιο  $C(t)$  δε χρησιμοποιείται, για να αναδιατυπώσει την ερώτηση, αλλά για να επικεντρώσει το σύνολο των προτιμήσεων των χρηστών, όπως, άλλωστε, θα δειχθεί και παρακάτω.

## 5.5 Εξαγωγή Προφίλ Χρηστών

Μέχρι τώρα επεξηγήσαμε τη μοντελοποίηση της βασισμένης στο εννοιολογικό πλαίσιο εξάρτησης μεταξύ των εννοιών και των σχέσεών τους, χρησιμοποιώντας μια αναπαράσταση ασαφούς άλγεβρας και δύο συνδυασμένων σημασιολογικών σχέσεων. Συνεχίζουμε με την παρουσίαση του ρόλου που έχουν τα προφίλ των χρηστών στην προσέγγιση της προσωποποίησης πολυμεσικού περιεχομένου, την έννοια των προτιμήσεων χρηστών καθώς επίσης και την παρουσίαση, την εξαγωγή και τη χρήση αυτών των προτιμήσεων στη διαδικασία.

### 5.5.1 Ο ρόλος των προφίλ

Είναι γεγονός ότι η αβεβαιότητα εξουσιάζει τη διαδικασία της ανάκτησης πληροφοριών [60], δεδομένου ότι ένα περιορισμένο σύνολο όρων δεν μπορεί να περιγράψει πλήρως την επιθυμία του χρήστη. Ο ρόλος της προσωποποίησης έγκειται στο να μειώσει αυτή την αβεβαιότητα με τη χρησιμοποίηση περισσότερων πληροφοριών για τις επιθυμίες του χρήστη από μόνο το τοπικό ενδιαφέρον. Η συμβολή των προφίλ χρηστών (user profiles) στην κατανόηση της έμφυτης αυτής επίδρασης στην ανάκτηση πληροφοριών είναι κρίσιμη, ιδιαίτερα όταν π.χ. δύο ανεξάρτητοι χρήστες οι οποίοι θέτουν παρόμοια ερωτήματα λαμβάνουν ως απάντηση διαφορετικά υποσύνολα ανακτημένων εγγράφων και σε διαφορετικούς βαθμούς. Ένα προφίλ χρήστη παράγεται μέσω του συνεχούς και σταθερού ελέγχου της αλληλεπίδρασης του χρήστη, η οποία περιέχει λιγότερη αβεβαιότητα λόγω της φύσης των ενεργειών του, εφόσον, βέβαια, η σχετική περίοδος ελέγχου είναι ικανοποιητική και αντιπροσωπευτική των προτιμήσεων του χρήστη. Επομένως, το προφίλ ενός χρήστη, το οποίο περιέχει πολύτιμες πληροφορίες σχετικά με το γενικότερο ενδιαφέρον του χρήστη (δηλ. πληροφορίες σχετικά με τις προτιμήσεις του χρήστη κατά τη διάρκεια μιας μακριάς χρονικής περιόδου), μπορεί να χρησιμοποιηθεί, όποτε το ερώτημα (δηλαδή η τοπική προτίμηση του χρήστη ή με άλλα λόγια το πεδίο της τρέχουσας αλληλεπίδρασής του) παρέχει ανεπαρκείς πληροφορίες για το χρήστη και το τοπικό ενδιαφέρον του.

Προκειμένου να επεξεργαστούμε το προφίλ ενός χρήστη χρησιμοποιώντας την εκάστοτε αποθηκευμένη γνώση, απαιτείται πρώτα η ύπαρξη μιας σωστής, πλήρους και συμβατής με το οντολογικό μοντέλο αναπαράστασης της γνώσης. Καθώς όλες οι σχέσεις  $R_i$  που υπάρχουν στην οντολογία  $\mathcal{O}$  ορίζονται πάνω στο κλασικό σύνολο  $S$  των εννοιών, ορίζουμε και τις προτιμήσεις των χρηστών πάνω στο ίδιο σύνολο: Η αναπαράστασή τους, που επιτρέπει επίσης τη χρήση βαθμών προτίμησης, αποτελείται από ένα μοναδικό ασαφές σύνολο που ορίζεται στο σύνολο των εννοιών, όπως κατά τα γνωστά περιγράφεται στην ενότητα 3.2.

Στο σημείο αυτό θα πρέπει να τονισθεί ότι, όταν ο χρήστης θέτει ένα ερώτημα που στην πραγματικότητα σχετίζεται με μια από τις προτιμήσεις του, αυτή η προτίμηση μπορεί να χρησιμοποιηθεί, για να διευκολύνει την ερμηνεία της ερώτησης, καθώς επίσης και την ταξινόμηση των επιλεγμένων προς ανάκτηση εγγράφων. Η χρήση, όμως, προτιμήσεων που είναι άσχετες προς το αρχικό ερώτημα μπορεί να θεωρηθεί μόνο ως προσθήκη θορύβου, δεδομένου ότι οποιαδήποτε εγγύτητα μεταξύ των επιλεγμένων εγγράφων και αυτών των προτιμήσεων είναι τελείως συμπτωματική. Κατά συνέπεια, εκτός από τις θετικές προτιμήσεις  $P^+$ , ειδική προσοχή πρέπει να δοθεί και για την αναπαράσταση και ξεχωριστή αποθήκευση των αρνητικών προτιμήσεων  $P^-$ , έτσι ώστε να μπορούμε να τις επεξεργαστούμε ξεχωριστά. Συνεπώς, μια προτίμηση  $P$  αποτελεί-

ται στην πραγματικότητα από δύο ζεχωριστά ασαφή σύνολα θετικών και αρνητικών προτιμήσεων:

$$P = \{P^+, P^-\}, \quad P \in \mathcal{F}_Z \quad (5.18)$$

### 5.5.2 Ενέργειες χρηστών

Στο στάδιο του προσδιορισμού και των δύο ειδών προτιμήσεων χρηστών αρχίζουμε πρώτα από το σύνολο των εγγράφων που είναι διαθέσιμα στο ιστορικό χρήστη του κάθε χρήστη. Η προτεινόμενη υλοποίηση των προφίλ χρήστη λαμβάνει αυτό το ιστορικό χρήστης ως είσοδο και παράγει τις αντίστοιχες προτιμήσεις χρηστών ως έξοδο. Προκειμένου να επιτευχθεί αυτό, η διαδικασία έχει επίσης πρόσβαση στο σημασιολογικό ευρετήριο και την οντολογική γνώση. Το σύνολο των διαθέσιμων εγγράφων στο ιστορικό χρήστης κατασκευάζεται ως αποτέλεσμα της εφαρμογής τεσσάρων τύπων ενεργειών χρηστών κατά τη διάρκεια της αλληλεπίδρασης του χρήστη με το σύστημα ανάκτησης πληροφορίας. Αυτές οι ενέργειες χαρακτηρίζουν τον εκάστοτε χρήστη και εκφράζουν την προσωπική άποψή του για το χώρο της αναζήτησης περιεχομένου. Αυτές οι ενέργειες συνδέονται άμεσα με τα αιτήματα ή τις ερωτήσεις των χρηστών και επομένως θα χρησιμοποιήσουμε τον όρο ερώτημα (query) στη συνέχεια. Οι τέσσερις πιθανοί τύποι μιας ενέργειας που μπορεί να θέσει ένας χρήστης ενός τέτοιου συστήματος ανάκτησης πληροφοριών είναι: βασισμένες στις λέξεις-κλειδιά ερωτήσεις, ερωτήσεις επισκόπησης εγγράφων, ερωτήσεις συσχετιστικής ανατροφοδότησης και ερωτήσεις πλοιήγησης. Ορίζουμε το σύνολο κάθε τέτοιου τύπου ερώτησης ως ένα ασαφές σύνολο εννοιών στο  $S$ , του οποίου οι βαθμοί ιδιότητας μέλους λαμβάνονται με την παρακολούθηση και τον έλεγχο των συγκεκριμένων πιθανοτήτων εμφάνισης των τύπων ερώτησης ως εξής:

1. Ερωτήματα με βάση λέξεις-κλειδιά  $Q^k$ , επίσημα ορισμένα ως ένα ασαφές σύνολο από έννοιες στο  $S$ , δηλαδή  $Q^k \in \mathcal{F}_S$ . Οι λέξεις-κλειδιά μπορούν να εξαχθούν από ένα ερώτημα φυσικής γλώσσας ή ένα ερώτημα βασισμένο στις λέξεις-κλειδιά και αντιστοιχίζονται σε έννοιες στο χαρακτηρισμό των εγγράφων, χρησιμοποιώντας τις πιο πρόσφατες τεχνικές εξαγωγής πληροφορίας [202].
2. Ερωτήματα επισκόπησης εγγράφων  $Q^v$ , επίσημα ορισμένα ως ένα ασαφές σύνολο από έννοιες στο  $S$ , δηλαδή  $Q^v \in \mathcal{F}_S$ . Στην περίπτωση αυτή οι έννοιες εμφανίζονται κατευθείαν μέσα στο σχολιασμό ενός εγγράφου και ανακτώνται με τη βοήθεια του σημασιολογικού ευρετηρίου [263], συνεπώς  $Q^v = S_d$  σε αυτή την περίπτωση.
3. Ερωτήματα συσχετιστικής ανατροφοδότησης  $Q^{rf}$ , τα οποία ικανοποιούν τα αιτήματα συσχετιστικής ανατροφοδότησης των χρηστών και αποτελούνται από δύο μέρη, τα θετικά και τα αρνητικά αιτήματα συσχετιστικής ανατροφοδότησης:

$$Q^{rf} = \{Q^+, Q^-\} \quad (5.19)$$

Και τα δύο αιτήματα ορίζονται και πάλι ως ασαφή σύνολα πάνω στο σύνολο των εννοιών  $S$ , δηλαδή  $Q^+, Q^- \in \mathcal{F}_S$ . Το  $Q^+$  αντιστοιχεί στο χαρακτηρισμό του συνόλου των εγγράφων που μαρκάρονται ως σχετικά από το χρήστη και για το λόγο αυτό ορίζεται ως:

$$Q^+ = \bigcup_{d \in D^+} S_d \quad (5.20)$$

## Κεφάλαιο 5. Προσωποποιημένη πρόσβαση

ενώ το  $Q^-$  αντιστοιχεί στο χαρακτηρισμό του συνόλου των εγγράφων που μαρκάρονται ως μη σχετικά από το χρήστη και για το λόγο αυτό ορίζεται ως:

$$Q^- = \bigcup_{d \in D^-} S_d \quad (5.21)$$

Το  $D^+$  υποδηλώνει το σύνολο των εγγράφων που υποδείχτηκαν ως θετικά από το χρήστη κατά τη διάρκεια των επαναλήψεων της συσχετιστικής ανατροφοδότησης, ενώ το  $D^-$  υποδηλώνει το σύνολο των εγγράφων που υποδείχτηκαν ως αρνητικά.

4. Ερωτήματα πλοήγησης  $Q^b$ , σύμφωνα με μία συγκεκριμένη θεματική κατηγορία πλοήγησης ή κατηγορία εγγράφων ή εννοιών. Το  $Q^b$  ορίζεται ως ένα ασαφές σύνολο από θεματικές κατηγορίες που ζητήθηκαν για πλοήγηση από το χρήστη, δηλαδή αναγνωρίζεται επίσημα ως  $Q^b \in \mathcal{F}_Z$ .

Στο σημείο αυτό θα πρέπει να τονίσουμε ότι ανάλογα με τον τύπο του ερωτήματος χρήστη, το εννοιολογικό πλαίσιο κατά το χρόνο εκτέλεσης διαμορφώνεται ως ακολούθως. Έστω ότι ορίζουμε το σύνολο όλων των διαθέσιμων χρονοθυρίδων (time slots) ως:  $\mathcal{T} = \{1, \dots, M\}$ , δηλ.  $t \in \mathcal{T}$ . Ορίζουμε και το  $Q_t$  ως το ασαφές σύνολο εννοιών που δημιουργείται αμέσως μετά από κάθε ερώτημα ενός χρήστη σε μια δεδομένη χρονική στιγμή  $t$ , δηλ.  $Q_t \in \mathcal{F}_S$ . Προφανώς, σύμφωνα με όσα παρατίθενται παραπάνω, θα έχουμε:

$$Q_t = Q_t^k \cup Q_t^v \cup Q_t^+ \cup Q_t^b \quad (5.22)$$

όπου:

$$Q^k = \bigcup_{t \in \mathcal{T}} Q_t^k, \quad Q^v = \bigcup_{t \in \mathcal{T}} Q_t^v, \quad Q^+ = \bigcup_{t \in \mathcal{T}} Q_t^+, \quad Q^b = \bigcup_{t \in \mathcal{T}} Q_t^b \quad (5.23)$$

### 5.5.3 Ιστορικό χρήσης

Το ιστορικό χρήσης ενός χρήστη περιλαμβάνει έναν συνδυασμό όλων των παραπάνω τύπων ενεργειών, υπό τον όρο ότι σε μία δεδομένη στιγμή ένας χρήστης είναι σε θέση να εκτελέσει οποιοδήποτε τύπο ενέργειας. Υπάρχει μια συσχέτιση μεταξύ των σχετικών εγγράφων του ιστορικού και των εννοιών μέσω της χρησιμοποίησης του σημασιολογικού ευρετηρίου, το οποίο κατασκευάζεται εκ των προτέρων κατά τη διάρκεια της ανάλυσης είτε του ακατέργαστου πολυμεσικού περιεχομένου, είτε του σχετικού κειμενικού σχολιασμού. Υποδηλώνουμε με επίσημο τρόπο ολόκληρο το ιστορικό χρήσης του κάθε χρήστη, δηλαδή τις έννοιες που συνδέονται με τα έγγραφα του ιστορικού χρήσης του, ως:

$$H = \{H^+, H^-\} \quad (5.24)$$

το οποίο αποτελείται τόσο από θετικά  $H^+$ , όσο και αρνητικά  $H^-$  μέρη. Θα πρέπει να έχει καταστεί σαφές έως αυτό το σημείο ότι ο όρος θετικά αντιστοιχεί σε πράγματα που αρέσουν στο χρήστη, ενώ ο όρος αρνητικά αντιστοιχεί σε πράγματα που δεν του αρέσουν. Σε αυτό το πλαίσιο, το  $H^+$  ορίζεται ως το ασαφές σύνολο εννοιών που λαμβάνεται από την ένωση όλων των εννοιών που σχετίζονται με όλα τα ερωτήματα του χρήστη, οπότε:

$$H^+ = Q^k \cup Q^v \cup Q^+ \cup Q^b, \quad H^+ \in \mathcal{F}_S \quad (5.25)$$

και

$$H^- = Q^-, \quad H^- \in \mathcal{F}_S \quad (5.26)$$

## 5.5.4 Από τα έγγραφα στις προτιμήσεις χρηστών

### 5.5.4.1 Μοντελοποίηση προσέγγισης

Ο επίσημος ορισμός των προτιμήσεων χρηστών ως ένα ασαφές σύνολο εννοιών (όπως αυτό περιγράφεται στις ενότητες 5.4 και 5.5.1), επιτρέπει τη συμμετοχή μιας έννοιας σε πολλαπλές προτιμήσεις και με διαφορετικούς βαθμούς. 'Όπως έχει ήδη ειπωθεί, το ιστορικό  $H$  του χρήστη αντιπροσωπεύεται ως ένα ασαφές σύνολο στο σύνολο των εννοιών που σχετίζονται με αυτό και αποτελείται τόσο από θετικά ( $H^+$ ), όσο και αρνητικά ( $H^-$ ) μέρη. Οι προτιμήσεις εξάγονται με τη χρησιμοποίηση και των δύο παραπάνω μερών ως είσοδο και εφαρμόζοντας κατάλληλους αλγορίθμους συσταδοποίησης. Χρησιμοποιώντας την έννοια του εννοιολογικού πλαισίου στη διαδικασία, εξάγουμε τελικά δύο ανεξάρτητα σύνολα θετικών και αρνητικών προτιμήσεων χρηστών ως έξοδο και τις συνδυάζουμε με έναν τρόπο που να έχει νόημα για να λάβουμε το  $P$ .

'Όπως έχουμε ήδη αναφέρει (βλ. ενότητα 4.2.1.1), οι περισσότερες μέθοδοι συσταδοποίησης που απαντώνται στη βιβλιογραφία ανήκουν σε μία εκ των δύο γενικών κατηγοριών, τις διαμεριστικές και τις ιεραρχικές [251]. Οι ιεραρχικές μέθοδοι δεν απαιτούν τον αριθμό των συστάδων ως είσοδο, σε αντίθεση με τις αντίστοιχες διαμεριστικές. Δεδομένου ότι ο αριθμός των προτιμήσεων που μπορεί να απαντηθούν σε ένα έγγραφο δεν είναι γνωστός εκ των προτέρων, οι τελευταίες δεν είναι εφαρμόσιμες [177]. Το ίδιο πράγμα ισχύει και για τη χρήση μιας εποπτευμένης μεθόδου συσταδοποίησης, η οποία επιτρέπει σε μια έννοια να ανήκει σε δύο ή περισσότερες συστάδες, όπως π.χ. οι ασαφείς c-μέσοι όροι [25]. Αυτός ο αλγόριθμος απαιτεί τον αριθμό των συστάδων εννοιών ως είσοδο, δηλαδή χρησιμοποιεί ένα "σκληρό" κριτήριο τερματισμού πάνω στον αριθμό των συστάδων. Κατά συνέπεια, δεν μπορεί να υιοθετηθεί για την επίλυση του παρόντος προβλήματος. Αντ' αυτού χρησιμοποιούμε μια υβριδική προσέγγιση, η οποία βασίζεται στην ασαφοποίηση ενός συσσωρευτικού ιεραρχικού αλγόριθμου συσταδοποίησης.

'Εστω ότι  $K' = \{k'_i\}$  είναι το σύνολο των κλασικών συστάδων που αναγνωρίζονται στο  $H^+$ , όπου κάθε συστάδα  $k'_i$  είναι ένα κλασικό σύνολο από έννοιες. Εντούτοις, κάτι τέτοιο δεν είναι αρκετό για την προσέγγισή μας, δεδομένου ότι πρέπει να υποστηρίξουμε έγγραφα που ανήκουν σε πολλαπλές ανεξάρτητες προτιμήσεις και με διαφορετικούς βαθμούς, αλλά και να διατηρήσουμε συγχρόνως την ευρωστία και την αποδοτικότητα της ιεραρχικής προσέγγισης συσταδοποίησης. Κατά συνέπεια και χωρίς οποιαδήποτε απώλεια της λειτουργικότητας ή αύξηση του υπολογιστικού κόστους αντικαθιστούμε τις κλασικές συστάδες  $k'_i$  με ασαφείς ομαλοποιημένες συστάδες  $k_i$  με την κατασκευή ενός ασαφούς ταξινομητή από  $K' = \{k'_i\} \rightarrow K = \{k_i\}$ , όπου  $K = \{k_i\}$  είναι το σύνολο των λαμβανομένων ασαφών συστάδων από έννοιες. 'Όπως περιγράφεται στη συνέχεια, για κάθε ασαφή συστάδα  $k_i$  λαμβάνουμε το ασαφές σύνολο των προτιμήσεων που σχετίζεται με αυτή με το να εκμεταλλεύμαστε τις πληροφορίες εννοιολογικού πλαισίου και πληθικότητας που τη συνοδεύουν. Κατόπιν, με τη γενίκευση της διαδικασίας σε ολόκληρο το σύνολο των ασαφών συστάδων, προσδιορίζουμε το ασαφές σύνολο των προτιμήσεων που σχετίζονται με το αρχικό σύνολο των εγγράφων που βρίσκονται στο ιστορικό χρήστη του χρήστη μετά από την εκλέπτυνσή του σύμφωνα με το υπάρχον προκαθορισμένο σύνολο όλων των πιθανών προτιμήσεων χρηστών.

Η παραπάνω διαδικασία ασαφοποίησης/συσταδοποίησης είναι παρεμφερής με τη μεθοδολογία που ακολουθείται στην υποενότητα 4.4.6 της παρούσης διδακτορικής

## Κεφάλαιο 5. Προσωποποιημένη πρόσβαση

διατριβής. Εν γένει, είναι δυνατόν να αναπαραστήσουμε το τριπλό αυτό μοντέλο (i) της αρχικής συσταδοποίησης εννοιών, (ii) της ασαφοποίησης των συστάδων και (iii) της τελικής εξαγωγής προτιμήσεων χρηστών με έναν αφηρημένο τρόπο ως μία συνάρτηση:

$$\mathcal{Y} = G(\mathcal{X}) \quad (5.27)$$

χωρίς να προχωρούμε σε καμία υπόθεση για τον τρόπο με τον οποίο μπορεί να αναπαρασταθεί και να αρχικοποιηθεί τόσο η είσοδος όσο και η έξοδος της συνάρτησης. Η συνάρτηση λαμβάνει ως είσοδο ένα ασαφές σύνολο  $\mathcal{X}$  και παράγει ένα διαφορετικό ασαφές σύνολο  $\mathcal{Y}$  ως έξοδο. Σε αυτό το πλαίσιο μπορούμε να συγκεκριμένοποιησουμε την παραπάνω εξίσωση για τη συγκεκριμένη περίπτωση των θετικών προτιμήσεων χρηστών και ιστορικού χρήσης. Η συνάρτηση  $G$  μπορεί να χρησιμοποιηθεί, για να υπολογίσουμε το σύνολο  $P^+$  από το σύνολο  $H^+$  ως εξής:

$$P^+ = G(H^+) \quad (5.28)$$

Η προτεινόμενη προσέγγιση μπορεί να αναλυθεί περαιτέρω στα ακόλουθα τέσσερα βήματα:

1. Εκτελέστε μια σαφή συσταδοποίηση των εννοιών  $H^+$ , έτσι ώστε να καθοριστεί το πλήθος των ξεχωριστών θετικών προτιμήσεων  $P^+$  με τις οποίες μπορεί να σχετίζεται ένα έγγραφο από το ιστορικό
2. Κατασκευάστε έναν ασαφή ταξινομητή, ο οποίος θα μετρά το βαθμό συσχέτισης μιας έννοιας  $s_j$  με την συστάδα  $k_i$ .
3. Θεωρήστε το εννοιολογικό πλαίσιο και την πληθικότητα των ασαφών συστάδων  $k_i$  που προκύπτουν και προσαρμόστε από μαθηματικής απόψεως τις υπολογισμένες τιμές τους, έτσι ώστε να ταιριάζουν με τα σημασιολογικά προσδοκώμενα αντίστοιχά τους.
4. Προσδιορίστε τις θετικές προτιμήσεις χρηστών  $P^+$  που σχετίζονται με κάθε συστάδα σύμφωνα με το εκ των προτέρων γνωστό σύνολο όλων των πιθανών προτιμήσεων χρηστών, προκειμένου να εξαχθεί ένα συνολικό γενικό αποτέλεσμα.

Ακριβώς η ίδια διαδικασία ακολουθείται και στην περίπτωση της εφαρμογής της συνάρτησης  $G$  στο σύνολο  $H^-$ , προκειμένου να εξαχθεί το σύνολο  $P^-$  ως:

$$P^- = G(H^-) \quad (5.29)$$

Όπως έχει ήδη αναφερθεί, το τελικό σύνολο προτιμήσεων  $P$  που αντιστοιχεί στο ιστορικό χρήσης του χρήστη είναι ο συνδυασμός τόσο του συνόλου των θετικών προτιμήσεων  $P^+$ , όσο και του συνόλου των αρνητικών προτιμήσεων  $P^-$ , οπότε χρησιμοποιώντας και πάλι το συμβολισμό αθροίσματος για τα ασαφή σύνολα θα έχουμε τελικά:

$$P = \sum_{s \in S} s / \max(0, P^+(s) - P^-(s)) \quad (5.30)$$

όπου το  $P(s) = \max(0, P^+(s) - P^-(s))$  αναπαριστά τον τελικό βαθμό συσχέτισης προτιμήσεων για κάθε έννοια  $s$ .

#### 5.5.4.2 Κλασική συσταδοποίηση

Το πρώτο βήμα, λοιπόν, προς τον προσδιορισμό των προτιμήσεων των χρηστών είναι η υλοποίηση μιας κλασικής μορφής σαφούς συσταδοποίησης του συνόλου των εννοιών που απαντώνται στο ιστορικό του χρήστη. Η γενική δομή μιας ιεραρχικής προσέγγισης συσταδοποίησης κατάλληλα προσαρμοσμένη για τις ανάγκες του προβλήματος που θα αντιμετωπίσουμε ακολουθεί. Χωρίς απώλεια της γενικότητας συγκεκριμενοποιούμε την προσέγγισή μας μόνο για τις θετικές προτιμήσεις  $P^+$ , λαμβάνοντας υπόψη ότι ακριβώς η ίδια προσέγγιση ισχύει και για τις αρνητικές προτιμήσεις  $P^-$ . Έτσι τα βήματά της συνοψίζονται στα ακόλουθα:

1. Θεωρήστε το σύνολο των διαθέσιμων εννοιών προς συσταδοποίηση  $H^+$  και ότι κάθε μία από αυτές αποτελεί ένα μονοσύνολο (singleton), δηλ. μια συστάδα  $k'_i$  από μόνη της.
2. Για κάθε ζεύγος συστάδων  $k'_1, k'_2$  υπολογίστε το δείκτη συμβατότητας τους  $d(k'_1, k'_2)$ . Ο δείκτης αυτός αποτελεί και δείγμα ομοιότητας των συστάδων ή μετρική απόστασής τους.
3. Συγχωνεύεται το ζεύγος συστάδων που διαθέτει τον καλύτερο δείκτη συμβατότητας  $d(k'_1, k'_2)$ . Ανάλογα με τον εάν αυτός ο δείκτης αποτελεί μέτρο ομοιότητας ή μετρική απόστασης, ο καλύτερος δείκτης μπορεί να επιλεχθεί χρησιμοποιώντας τον τελεστή  $\max$  ή  $\min$ , αντίστοιχα.
4. Επιστροφή στο βήμα 2, εκτός εάν εκπληρώνονται τα κριτήρια τερματισμού. Το πιο διαδεδομένο κριτήριο τερματισμού αποτελεί ένα ουσιώδες κατώφλι για την τιμή του καλύτερου δείκτη συμβατότητας  $d(k'_1, k'_2)$ .

Όπως σε όλες τις χαρακτηριστικές προσέγγισεις ιεραρχικής συσταδοποίησης, έτσι και εδώ τα δύο καίρια σημεία της ανωτέρω διαδικασίας είναι ο προσδιορισμός των συστάδων προς συγχώνευση σε κάθε βήμα και ο προσδιορισμός του βέλτιστου βήματος ολοκλήρωσης του αλγορίθμου. Στην παρούσα φάση επιλέχθηκε μια κατάλληλη μετρική απόστασης: το ύψος του εννοιολογικού πλαισίου  $h(C(k'_1 \cup k'_2))$ . Το ύψος του εννοιολογικού πλαισίου χρησιμοποιείται ως μετρική απόστασης για δύο συστάδες  $k'_1, k'_2$ , ποσοτικοποιώντας με τον τρόπο αυτό το σημασιολογικό συσχετισμό τους. Η διαδικασία ολοκληρώνεται, όταν συσταδοποιηθούν οι έννοιες σε επιμέρους σύνολα, τα οποία αντιστοιχούν στις ευδιάκριτες προτιμήσεις, οι οποίες προσδιορίζονται από το γεγονός ότι το κοινό εννοιολογικό πλαίσιο τους έχει χαμηλό ύψος. Επομένως, το κριτήριο τερματισμού αποτελεί ένα κατώφλι ή κατώτατο όριο για την επιλεγμένη μετρική συμβατότητας. Η έξοδος αυτού του βήματος είναι ένα κλασικό σαφές σύνολο συστάδων  $K'$ , όπου κάθε συστάδα  $k'_i \in K'$  είναι ένα σαφές σύνολο από έννοιες,  $k'_i \in S_d$ .

#### 5.5.4.3 Ασαφοποίηση συστάδων

Η ανωτέρω μέθοδος συσταδοποίησης καθορίζει αφενός με επιτυχία το πληθος των ξεχωριστών συστάδων που υπάρχουν, αλλά αφετέρου δημιουργεί μόνο σαφείς συστάδες, δηλ. δεν επιτρέπει την ύπαρξη βαθμών ιδιότητας μέλους στην έξοδο και δεν επιτρέπει την ύπαρξη επικάλυψης μεταξύ των ανιχνευμένων συστάδων. Εντούτοις, μια έννοια μπορεί στην πραγματικότητα να αφορά μια προτίμηση χρηστών κατά ένα βαθμό διάφορο του 1 ή 0. Μπορεί επίσης να αφορά περισσότερες της μίας ξεχωριστές

## Κεφάλαιο 5. Προσωποποιημένη πρόσβαση

προτιμήσεις. Προκειμένου να υπερινικηθούν τέτοιους είδους προβλήματα, απαιτείται η επιβολή μιας κατάλληλης διαδικασίας ασαφοποίησης (*fuzzification*) των παραγόμενων συστάδων. Ειδικότερα, κατασκευάζουμε έναν ασαφή ταξινομητή, δηλ. μια συνάρτηση:

$$G_c : S \rightarrow [0, 1] \quad (5.31)$$

η οποία μετράει τον βαθμό συσχέτισης μιάς έννοιας  $s_j \in S$  με την συστάδα  $k'_i$ . Προφανώς, μία έννοια  $s_j$  θα πρέπει να θεωρείται συσχετισμένη με τη συστάδα  $k'_i$ , εφόσον σχετίζεται με το κοινό νόημα των εννοιών στο  $k'_i$ . Για το λόγο αυτό η ποσότητα

$$G_c(s_j) = \frac{h(C(k'_i \cup \{s_j\}))}{h(C(k'_i))} \quad (5.32)$$

αποτελεί μια κατάλληλη μετρική συσχέτισης. Η μετρική αυτή παρουσιάζει τις εξής ιδιότητες:

- $G_c(s_j) = 1$ , αν η σημασιολογία της  $s_j$  υπαινίσσεται ότι πρέπει να ανήκει στο  $k'_i$ . Για παράδειγμα:  $G_c(s_j) = 1, \forall s_j \in k'_i$ .
- $G_c(s_j) = 0$ , αν η σημασιολογία της  $s_j$  υπαινίσσεται ότι δεν πρέπει να ανήκει στο  $k'_i$ .
- $G_c(s_j) \in (0, 1)$ , αν η  $s_j$  δε σχετίζεται καθολικά με το  $k'_i$ .

Χρησιμοποιώντας αυτόν τον ταξινομητή, επεκτείνουμε τις ανιχνευμένες κλασικές συστάδες, για να συμπεριλάβουμε περισσότερες έννοιες. Κατά συνέπεια, η συστάδα  $k'_i$  αντικαθίσταται από την ασαφή συστάδα  $k_i \supseteq k'_i$  και  $k_i = \sum_{s_j \in S_d} s_j / G_c(s_j)$ , χρησιμοποιώντας και πάλι το συμβολισμό αθροίσματος για τα ασαφή σύνολα.

Το τελευταίο σημείο που πρέπει να εξεταστεί κατά τη διάρκεια του βήματος της ασαφοποίησης είναι το γεγονός ότι μέχρι τώρα η διαδικασία της ασαφούς εραρχικής συσταδοποίησης έχει βασιστεί στο κλασικό σύνολο  $S_d$ , αγνοώντας κατά συνέπεια την ασάφεια που υπάρχει στο σημασιολογικό ευρετήριο. Προκειμένου να ενσωματωθεί αυτή η πληροφορία κατά τον υπολογισμό των συστάδων, θα πρέπει να ρυθμίσουμε τους βαθμούς ιδιότητας μέλους τους  $k_i(s_j)$  σύμφωνα με την πληροφορία που ενυπάρχει στο σημασιολογικό ευρετήριο  $I(s_j, d)$ . Κατόπιν, καθεμία από τις παραγόμενες συστάδες αντιστοιχεί σε μια από τις ξεχωριστές προτιμήσεις χρηστών του εγγράφου. Προκειμένου να καθοριστούν οι προτιμήσεις που σχετίζονται με μια συστάδα  $k_i$ , πρέπει να θεωρήσουμε τόσο τον αριθμό των στοιχείων του συνόλου της (ή αλλιώς την πληθικότητα της συστάδας)  $|k_i|$ , όσο και το εννοιολογικό πλαίσιο της. Δεδομένου ότι το ταξινομικό εννοιολογικό πλαίσιο ορίστηκε προηγουμένως μόνο για κανονικά ασαφή σύνολα, κάθε βαθμός συσχέτισης θα προκύψει τελικά ως:

$$k_i(s_j) = \frac{t(k_i(s_j), I(s_j, d))}{h(t(k_i(s_j), I(s_j, d)))}, \forall s_j \in H^+ \quad (5.33)$$

όπου εξαιτίας της σημασιολογικής φύσης της παραπάνω λειτουργίας η  $t$  είναι - κατά τα γνωστά - μια Αρχικήδεια  $t$ -νόρμα.

#### 5.5.4.4 Εξαγωγή ασαφών προτιμήσεων

Προκειμένου να προσδιοριστεί το ασαφές σύνολο  $\{W(k_i)\}$  των προτιμήσεων που σχετίζονται με το σύνολο των εννοιών υπό εξέταση, πρέπει να υπολογίσουμε κάθε  $W(k_i)$ , δηλ. το σύνολο των προτιμήσεων που σχετίζονται με κάθε συστάδα  $k_i$ . Τα τελευταία υπολογίζονται ως εξής:

$$W(k_i) = \tilde{w}(C(k_i)) \cdot L(|k_i|) \quad (5.34)$$

όπου  $\tilde{w}$  είναι ένας αδύναμος τροποποιητής (weak modifier) και  $L(\cdot)$  είναι ένας “μεγάλος” ασαφής αριθμός. Ο αδύναμος τροποποιητής χρησιμοποιείται, για να ρυθμίσει τις από μαθηματική άποψη υπολογισμένες τιμές, έτσι ώστε αυτές να ταιριάζουν με τα σημασιολογικά προσδοκώμενα αντίστοιχά τους. Ο  $\tilde{w}(a) = \sqrt{a}$  είναι ένας συνήθης αδύναμος τροποποιητής [139]. Ο “μεγάλος” ασαφής αριθμός μοντελοποιεί την έννοια της “υψηλής πληθικότητας” των συστάδων και διαμορφώνει μια συνάρτηση από το σύνολο των πραγματικών θετικών αριθμών στο διάστημα  $[0, 1]$ , ποσοτικοποιώντας την έννοια του “μεγάλου” ή “υψηλού”. Στην παρούσα διατριβή ο “μεγάλος” ασαφής αριθμός ορίζεται με τη βοήθεια του τριγωνικού ασαφή αριθμού  $(1.3, 3, \infty)$  [139]<sup>1</sup>.

Προφανώς, αν υπάρχει μόνο μία συστάδα  $k_i$ , τότε η  $\{W(k_i)\} = \tilde{w}(C(k_i))$  είναι μία προσέγγιση που έχει νόημα, η οποία και παρέχει την έξοδο της διαδικασίας στην περίπτωση που αγνοηθούν οι πληθικότητες των συστάδων. Από την άλλη πλευρά, όταν αναγνωρίζονται περισσότερες της μίας συστάδες, τότε είναι επιτακτικό να εξετασθούν και οι τελευταίες. Συστάδες με εξαιρετικά χαμηλό αριθμό στοιχείων συνόλου περιέχουν πιθανότατα αποκλειστικά παραπλανητικές έννοιες και επομένως πρέπει να αγνοηθούν κατά την εκτίμηση του  $\{W(k_i)\}$ . Αυτίθετα, οι συστάδες με υψηλό αριθμό στοιχείων συνόλου αντιστοιχούν σχεδόν με βεβαιότητα στις ξεχωριστές προτιμήσεις και πρέπει να εξεταστούν κατά την εκτίμησή του σύμφωνα με την εξίσωση (5.34).

Τελικά, το σύνολο των προτιμήσεων που αντιστοιχούν στο σύνολο των εγγράφων του ιστορικού χρήσης που συνδέεται με τα ερωτήματα του χρήστη είναι το σύνολο των προτιμήσεων που ανήκουν σε οποιεσδήποτε από τις ανιχνευμένες συστάδες των εννοιών που ευρετηριάζουν τα δεδομένα έγγραφα. Παραδείγματος χάριν για το σύνολο των θετικών προτιμήσεων έχουμε:

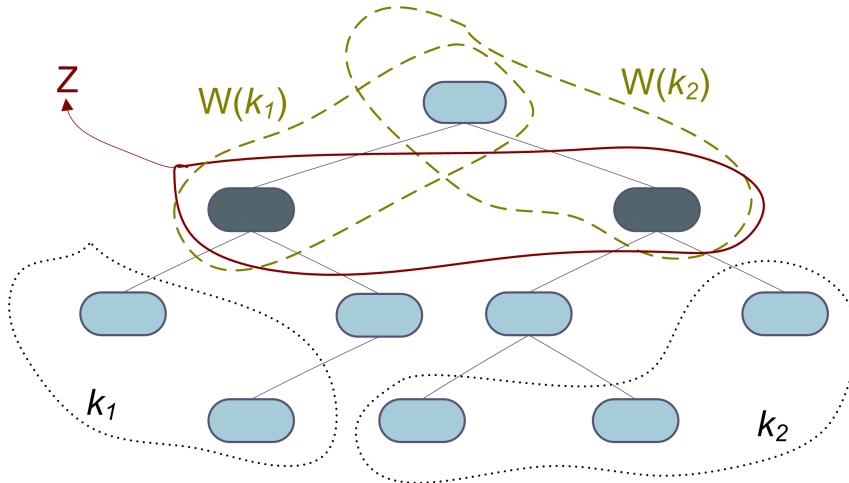
$$\{W(k_i)\} = \bigcup_{k \in K} W(k) \quad (5.35)$$

όπου  $\bigcup$  είναι μία ασαφής co-νόρμα και το  $K$  περιέχει το σύνολο των συστάδων που έχουν ανιχνευθεί στο  $H^+$ . Παρατηρούμε ότι το  $\{W(k_i)\}(s_j)$  θα είναι υψηλό, αν μια συστάδα  $k_i$ , της οποίας το εννοιολογικό πλαίσιο περιέχει την  $s_j$ , ανιχνευθεί στο  $H^+$ , και επιπρόσθετα, αν η πληθικότητα του  $k_i$  είναι υψηλή και ο βαθμός συσχέτισης της  $s_j$  στο πλαίσιο της συστάδας είναι επίσης υψηλός (ή με άλλα λόγια αν η θεματική κατηγορία σχετίζεται με τη συστάδα και η συστάδα δεν αποτελείται από παραπλανητικές έννοιες). Τέλος, προκειμένου να επικυρώσουμε τα αποτελέσματα της ασαφούς ταξινόμησης, δηλ. να βεβαιωθούμε ότι το σύνολο των θεμάτων  $\{W(k_i)\}$  που αντιστοιχούν στο σύνολο των εγγράφων  $H^+$  προέρχονται από το σύνολο όλων των πιθανών προτιμήσεων χρηστών  $Z$ , υπολογίζουμε την ποσότητα:

$$P^+ = \{W(k_i)\} \cap Z \quad (5.36)$$

<sup>1</sup> Εστω  $a, b, c \in R$ ,  $a < b < c$ . Ο ασαφής αριθμός  $u : R \rightarrow [0, 1]$  που αναπαριστάται με  $(a, b, c)$  και ορίζεται από το  $u(x) = 0$  αν  $x \leq a$  ή  $x \geq c$ ,  $u(x) = \frac{x-a}{b-a}$ , αν  $x \in [a, b]$  και  $u(x) = \frac{c-x}{c-b}$  αν  $x \in [b, c]$  ονομάζεται τριγωνικός ασαφής αριθμός.

Ακολουθώντας ακριβώς τα ίδια βήματα και για τις αρνητικές προτιμήσεις  $P^-$ , προχύπτουν οι συνολικές προτιμήσεις των χρηστών  $P$ . Ένα επεξηγηματικό παράδειγμα δίνεται στο Σχήμα 5.5. Όπως παρατηρεί κανείς στο Σχήμα, το  $W(k_1)$  αντιστοι-



Σχήμα 5.5: Παράδειγμα εξαγωγής ασαφών προτιμήσεων χρηστών.

χεί στο σύνολο των προτιμήσεων που σχετίζονται με την συστάδα  $k_1$  και το  $W(k_2)$  είναι το σύνολο των προτιμήσεων που σχετίζεται με την συστάδα  $k_2$ . Το σύνολο των προτιμήσεων που ανήκουν σε οποιαδήποτε από τις δύο συστάδες δίδεται από την  $W(k_1) \cup W(k_2)$ , δηλαδή εν προκειμένω το σύνολο και των τριών εννοιών. Η εφαρμογή της αρχής ύπαρξης ενός προκαθορισμένου συνόλου από πιθανές προτιμήσεις  $Z$  περιορίζει, όπως αναμένεται, το σύνολο των προτιμήσεων χρηστών στις δύο σκιασμένες θεματικές κατηγορίες που φαίνονται στο Σχήμα.

### 5.5.5 Εξαγωγή προφίλ χρηστών και ηλεκτρονική μάθηση

Στον τομέα της “εξ αποστάσεως εκπαίδευσης” ή ηλεκτρονικής μάθησης (e-learning) αποτελεί κοινό τόπο ότι τα σύγχρονα πρότυπα ηλεκτρονικής μάθησης χαρακτηρίζονται από ελλιπείς αρχές αναπαράστασης τόσο του ίδιου του εκπαιδευτικού περιεχομένου, όσο και της προσωποποίησης των χρηστών. Σε αυτό το πλαίσιο, η αυτοματοποιημένη εξαγωγή προφίλ χρηστών (τα οποία χρησιμοποιούνται σε ένα σύστημα ηλεκτρονικής μάθησης και παροχής πολυμεσικού εκπαιδευτικού περιεχομένου) διαμορφώνει ένα ενδιαφέρον και συνάμα σημαντικό ερευνητικό πρόβλημα. Στην υποενότητα αυτή παρουσιάζουμε μια προσέγγιση σχεδίασης, καθώς και την υλοποίηση ενός τέτοιου βασισμένου στα προφίλ συστήματος. Στα πλαίσια του συστήματος αυτού αντιστοιχείται το περιεχόμενό του στο περιβαλλοντικό του πλαίσιο, έτσι ώστε να μπορεί να προσαρμοστεί στις ανάγκες και τις ικανότητες των χρηστών του. Η προτεινόμενη προσέγγιση επεκτείνει τις ερευνητικές προσπάθειές μας πάνω στην εξαγωγή προφίλ χρηστών μέσω τεχνικών συσταδοποίησης στην περίπτωση ενός ολοκληρωμένου συστήματος ηλεκτρονικής μάθησης. Το τελευταίο βασίζεται στο θεμελιώδες μοντέλο ηλεκτρονικής μάθησης που προτείνει η IEEE [294], το οποίο προσαρμόζεται κατάλληλα, προκειμένου αφενός να απεικονίσει και αφετέρου να εστιάσει στα προφίλ των χρηστών του συστήματος.

Σε αυτό το σημείο είναι απαραίτητη μια συνοπτική παρουσίαση του εν λόγω συστήματος προσωποποίησης εκπαιδευτικού περιεχομένου<sup>1</sup>. Ειδικότερα, στο πλαίσιο

<sup>1</sup>Το σύστημα ηλεκτρονικής μάθησης και προσωποποίησης πολυμεσικού εκπαιδευτικού περιεχο-

του τελευταίου παρουσιάζεται μια νέα μέθοδος για τη συλλογή πληροφοριών και την εκτίμηση του επιπέδου Τεχνολογίας Πληροφόρησης και Επικοινωνίας (Information and Communications Technology - ICT) των μαθητών σε όλους τους τομείς της εκπαίδευσης. Αυτό επιτυγχάνεται μέσω βασισμένων στον Ιστό και φιλικών προς το χρήστη διεπαφών (e-questionnaires), οι οποίες λαμβάνουν υπόψη τους προσωποποιημένα και βασισμένα στα προφίλ σχέδια. Η μεθοδολογία αυτή έχει ως σκοπό να επιτρέψει στους αρχαρίους χρήστες να αυξήσουν βαθμίατα την ICT γνώση τους καθώς και να προσφέρει σε όλους αξιόπιστες πληροφορίες και υλικό, όπως είναι κατάλληλα επιλεγμένο εκπαιδευτικό περιεχόμενο, ηλεκτρονικά μαθήματα (e-courses) και πολυμεσικό εκπαιδευτικό περιεχόμενο (φωτογραφίες, βίντεο, παρουσιάσεις, κ.α.).

Η επίτευξη των παραπάνω στόχων υπαγορεύει τη χρησιμοποίηση δύο εκ πρώτης όψεως διαφορετικών προσεγγίσεων εξαγωγής προφίλ χρηστών· αν και οι δύο τρόποι φαίνονται αρχικά πολύ διαφορετικοί μεταξύ τους, τελικά συνδυάζονται σε μια μεταγενέστερη φάση. Η πρώτη προσέγγιση ακολουθείται στο αρχικό στάδιο της κατασκευής της απολύτης αλήθειας των προφίλ και χαρακτηρίζεται από έναν στατικό μηχανισμό εξαγωγής προφίλ, ενώ η δεύτερη εκμεταλλεύεται τα αποτελέσματα του προηγούμενου βήματος προς την κατεύθυνση της δυναμικής εξαγωγής των προφίλ τρέχοντων αλλά και μελλοντικών χρηστών του συστήματος. Το πρώτο βήμα υλοποιείται μέσω μιας σταθερής αντιστοίχισης των ηλεκτρονικών ερωτηματολογίων (e-questionnaires), που καλούνται να συμπληρώσουν οι χρήστες κατά την εισαγωγή τους στο σύστημα σε συγκεκριμένους τύπους προφίλ χρήστη, όπως θα εξηγηθεί στη συνέχεια. Το δεύτερο βήμα ενεργεί πάνω από το πρώτο και βασίζεται στην εφαρμογή της τεχνικής συσταδοποίησης που αναλύθηκε στην ενότητα 4.2.1 στα δεδομένα εισαγωγής. Συνεπώς, το σύνολο των διαθέσιμων δεδομένων διαιρέθηκε σε δύο μέρη: ένα αρχικό, το οποίο χρησιμοποιήθηκε από τους εμπειρογνόμονες του συστήματος κατά τη διάρκεια του πρώτου βήματος και ένα άλλο που χρησιμοποιήθηκε κατά την υλοποίηση της αυτοματοποιημένης μεθοδολογίας της συσταδοποίησης.

#### 5.5.5.1 Στατική εξαγωγή προφίλ χρηστών

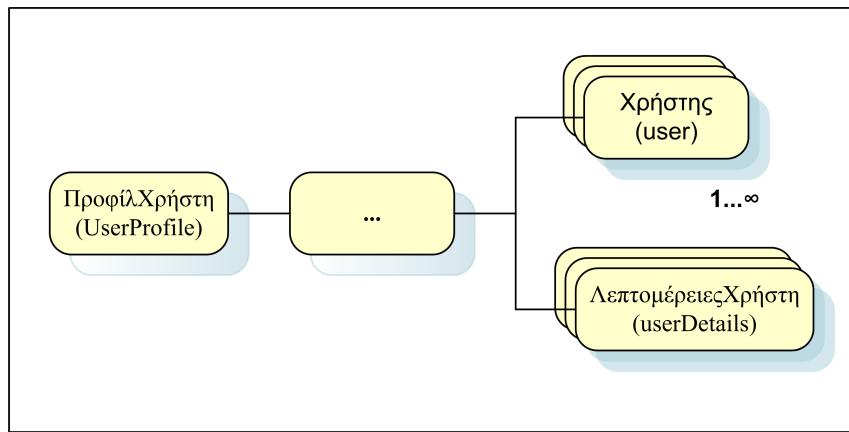
Σε πρώτη φάση θα επεξηγήσουμε τη λειτουργία της αρχικής αναπαράστασης προφίλ, η οποία βασίζεται σε ένα στατικό μηχανισμό προφίλ. Πιο συγκεκριμένα, οι εμπειρογνόμονες του συστήματος, βασισμένοι στην εμπειρία και τη διαίσθησή τους, καθορίζουν ένα σύνολο χαρακτηρισμών χρηστών (στην προκειμένη περίπτωση τριών), που διαμορφώνουν μια στατική αντιπροσώπευση των προφίλ τους. Αυτό το βήμα αποτελεί ένα βήμα προεπεξεργασίας, το οποίο εκτελείται με το χέρι και συνδυάζεται με τη μεταγενέστερη στατιστική ανάλυση και την ακόλουθη επιλογή συσχέτισης των χαρακτηριστικών γνωρισμάτων, η οποία και καταδεικνύει τα σημαντικότερα εξ αυτών. Θεωρείται πράγματι εξαιρετικά αξιόπιστο και ακριβές, δεδομένου ότι αποτέλεσε το αποτέλεσμα τριών ετών εντατικής συνεργασίας μεταξύ των εμπειρογνωμόνων από 8 ευρωπαϊκές χώρες, οι οποίοι εξέτασαν και ανέλυσαν τα σχετικά ερωτηματολόγια. Αυτοί οι αρχικοί χαρακτηρισμοί χρηστών χρησιμοποιούνται επιπρόσθετα και σε μεταγενέστερο στάδιο, κατά τη διάρκεια της ευφυούς διαδικασίας εξαγωγής των προφίλ και παρέχουν ένα στέρεο σημείο αναφοράς, αφού, αν και είναι στατικοί, παράγονται στην πραγματικότητα αυτόματα από το σύστημα, σύμφωνα με τους αρχικούς ορισμούς τους. Η βάση της εξαγωγής τους στηρίζεται στις πληροφορίες που παρέχονται από τα δεδομένα εισόδου των χρηστών, τα οποία λαμβάνονται από τα ηλεκτρονικά

μένου αναπτύχθηκε στο πλαίσιο του προγράμματος Leonardo SPERO [310]

## Κεφάλαιο 5. Προσωποποιημένη πρόσβαση

ερωτηματολόγια. Σε αυτήν την περίπτωση η διαδικασία της προσωποποίησης είναι απαραίτητη, προκειμένου να βοηθήσει την επιλογή του τελικού εκπαιδευτικού πολυμεσικού περιεχομένου που προσφέρεται στους χρήστες από το σύστημα.

Ο προτεινόμενος μηχανισμός εξαγωγής προφίλ δημιουργεί, αναπροσαρμόζει και χρησιμοποιεί τα προφίλ των χρηστών του συστήματος, ταιριάζοντας συγκεκριμένες ερωτήσεις από τα ηλεκτρονικά ερωτηματολόγια με αναγνωρισθέντα πρότυπα. Η σχεδίαση του μοντέλου των προφίλ διευκολύνει τόσο τη διαδικασία χρήσης των προτιμήσεων των χρηστών κατά τη δημιουργία των προφίλ, όσο και τη διαδικασία της ανίχνευσης/παρακολούθησης των προτιμήσεων αυτών. Επιπλέον, έχει σχεδιασθεί με έναν τρόπο που επιτρέπει την αυτοματοποιημένη εξαγωγή των προφίλ του χρήστη με βάση αυτές τις προτιμήσεις και το ιστορικό εισαγωγής δεδομένων των χρηστών. Αυτό το μοντέλο διαμορφώνει μία αρχική στατική εκδοχή του προφίλ ενός χρήστη, η οποία καταδεικνύεται από την οντότητα *UserProfile* που παρουσιάζεται στο Σχήμα 5.6. Όπως παρατηρεί κανείς στο Σχήμα, η κύρια αφηρημένη δομή του σύνθετου τύ-

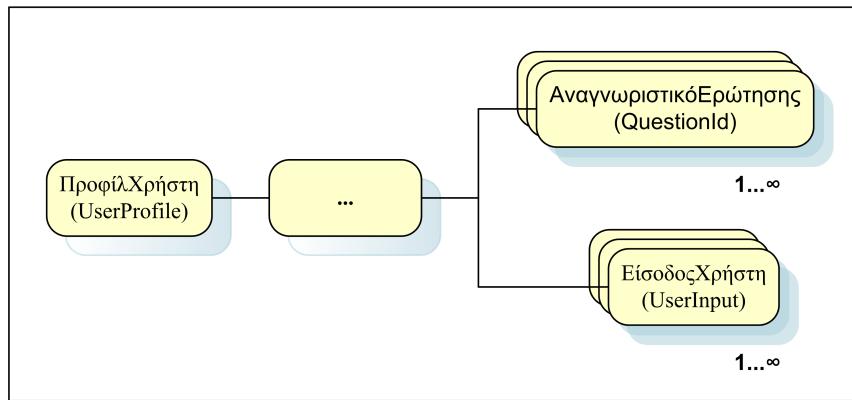


Σχήμα 5.6: Η δομή ενός Προφίλ Χρήστη.

που *Προφίλ Χρήστη* (*UserProfile*) περιλαμβάνει δύο στοιχεία. Το πρώτο (*user*) αποθηκεύει τις πληροφορίες για το ιστορικό του χρήστη, ενώ το δεύτερο (*userDetails*) αποθηκεύει τις προτιμήσεις των χρηστών. Καθώς αρχικοποιείται η διαδικασία εξαγωγής των προφίλ, όλα τα σχετικά προφίλ αποθηκεύονται μέσα σε μια ενιαία, κεντρική δομή αντιστοίχισης, της οποίας το αφηρημένο μοντέλο παρουσιάζεται στο Σχήμα 5.7. Το *UserProfile* αντιστοιχείται με τις πληροφορίες που ανακτώνται είτε από το ίδιο το εκάστοτε ηλεκτρονικό ερωτηματολόγιο, είτε άμεσα από την εισαγωγή δεδομένων που πραγματοποιούν οι χρήστες. Εδώ, το πρώτο στοιχείο (*QuestionId*) φυλάσσει όλες τις πληροφορίες που απαιτούνται για τον προσδιορισμό της ελλοχεύουσας ερώτησης του ηλεκτρονικού ερωτηματολογίου, καθώς επίσης και του τύπου της<sup>2</sup>, στοχεύοντας στην καλύτερη κατανόηση και ταίριασμα του παραγόμενου προφίλ. Το δεύτερο στοιχείο (*UserInput*) περιέχει τις σχετικές με τα στοιχεία πληροφορίες χρηστών, όπως είναι οι απαντήσεις των χρηστών. Και οι δύο ακόλουθες στοιχείων (*QuestionId* και *UserInput*) δείχνουν την ύπαρξη μιας σημαντικής ποικιλίας ηλεκτρονικών ερωτηματολογίων, ερωτήσεων και δεδομένων εισόδου των χρηστών.

Η ουσία της περιγραφόμενης μεθοδολογίας συνοψίζεται στο ακόλουθο βήμα της χρήσης ειδικών βαρών, το οποίο εκτελείται σύμφωνα με τις παρακάτω βασικές αρχές. Όταν ένας χρήστης απαντά σε μια ερώτηση του ηλεκτρονικού ερωτηματολογίου

<sup>2</sup> Πιθανοί τύποι ερωτήσεων ενός ηλεκτρονικού ερωτηματολογίου είναι μεταξύ άλλων: radio button, multiple choice, drop-down box, free text box, κ.α.



**Σχήμα 5.7:** Δομή αντιστοίχισης των Προφίλ Χρηστών.

εισαγωγής, ένας βαθμός σχετικότητας συνδέεται με αυτή και προσαρμόζεται στο συγκεκριμένο στοιχείο *QuestionId* του. Με τον τρόπο αυτό ο βαθμός αυτός διαδίδεται και στο στοιχείο *UserProfile*. Καθώς όλο και περισσότερες απαντήσεις από τον τελικό χρήστη εισάγονται στη δομή *UserProfile*, καταχωρούνται και οι πρόσθετοι βαθμοί σχετικότητας στα αντίστοιχα στοιχεία *QuestionID*. Ανάλογα με την εκάστοτε ερώτηση, καθώς επίσης και με το τμήμα του ηλεκτρονικού ερωτηματολογίου στο οποίο ανήκει η τελευταία<sup>3</sup>, διαδίδονται και διαφορετικοί βαθμοί. Το τελευταίο βήμα βασίζεται στη σύγκριση των παρεχόμενων αριθμητικών τιμών με τη σειρά των τιμών που συνδέονται εκ των προτέρων με τα προφίλ. Προκειμένου να γίνουν καλύτερα κατανοητές οι ελλοχεύουσες δομές αντιστοίχισης, παρουσιάζονται τα παραδείγματα των Σχημάτων 5.8, 5.9 και 5.10, τα οποία προκύπτουν με άμεσο τρόπο από τα ηλεκτρονικά ερωτηματολόγια.

B.1.4. Διδακτική Εμπειρία	
Κάτω από 5 χρόνια	Αρχάριος (Beginner)
Από 5 έως 10 χρόνια	Προχωρημένος (Advanced)
Από 11 έως 20 χρόνια	Προχωρημένος/Εμπειρος (Advanced/Expert)
Πάνω από 20 χρόνια	Εμπειρος (Expert)

**Σχήμα 5.8:** Παράδειγμα αντιστοίχισης στατικού προφίλ (1ο τμήμα του ηλεκτρονικού ερωτηματολογίου) - ανάλογα με την επιλογή του χρήστη στο αριστερό τμήμα προκύπτει η σχετική αντιστοίχιση σε προφίλ του δεξιού τμήματος.

Το ηλεκτρονικό ερωτηματολόγιο ενεργεί ως μεσάζων κατά τη διαδικασία συλλογής πληροφοριών και, καθώς αυξάνει ο αριθμός των απαντημένων ερωτήσεων, οι περισσότερες καταχωρήσεις συνοψίζονται στη δομή *UserProfile*. Κατά συνέπεια, η συνολική διαδικασία οδηγεί σε μία σταθμισμένη αθροιστική αντιστοίχιση του τελικού χρήστη σε ένα συγκεκριμένο προφίλ, το οποίο είναι διαφορετικό ανάλογα με τις απαντήσεις του κάθε χρήστη και εξαρτάται από την ιδιαίτερη φύση τους. Αυτή η αντιστοίχιση διατηρείται προσωρινά και, καθώς πραγματοποιείται η συμπλήρωση του ηλεκτρονικού ερωτηματολογίου, οι προαναφερθείσες σταθμισμένες αντιστοιχήσεις ενημερώνονται διαρκώς. Με αυτό τον τρόπο αλλάζουν συνεχώς και δυναμικά την αντιστοίχιση κάθε χρήστη στο προφίλ, έως ότου επιτυγχάνεται ένα τελικό σημείο

<sup>3</sup>Ένα ηλεκτρονικό ερωτηματολόγιο αποτελείται, εν γένει, από τα εξής τρία μέρη: (i) Επαγγελματική κατάρτιση, (ii) Προσωπικό ICT υπόβαθρο και (iii) Χρήση ICT στη διδασκαλία.

## Κεφάλαιο 5. Προσωποποιημένη πρόσβαση

B.2.5. Ποιες από τις παρακάτω εργασίες έχετε πραγματοποιήσει τουλάχιστον 1 φορά?	
Εγκατάσταση λογισμικού	α
Εγκατάσταση εκτυπωτή	β
Δημιουργία αντιγράφων ασφαλείας	γ
Εγκατάσταση εφαρμογών (π.χ. MS® Office)	δ
Επίλυση προβλήματος λογισμικού	ε
Επίλυση προβλήματος εκτυπωτή ή κάρτας δικτύου	ζ
Τίποτα από τα παραπάνω	η
Άλλο (παρακαλώ προσδιορίστε)	κείμενο / text

**Σχήμα 5.9:** Παράδειγμα αντιστοίχισης στατικού προφίλ (3ο τμήμα του ηλεκτρονικού ερωτηματολογίου) - 8 ειδικά βάρη αποδίδονται στις 8 διαφορετικές πιθανές επιλογές/απαντήσεις, σύμφωνα με τον χανόνα: Επιλογή ( $\alpha \dots \zeta$ )  $> 4 \Rightarrow$  "Έμπειρος",  $2 < \text{Επιλογή} (\alpha \dots \zeta) < 4 \Rightarrow$  "Προχωρημένος",  $2 < \text{Επιλογή} (\alpha \dots \zeta) \text{ ή } \text{Επιλογή} (\eta) \Rightarrow$  "Αρχάριος".

B.3.4. Πόσες ώρες την εβδομάδα (κατά μέσο όρο) χρησιμοποιείτε το Internet ή εκπαιδευτικό λογισμικό με τους μαθητές σας?	
Γενικοί	αν (ώρες Γενικών ή/και SEN) $> 5 \rightarrow$ «Έμπειρος»
SEN	αλλιώς αν $2 < \text{ώρες} < 5 \rightarrow$ «Προχωρημένος» αν ώρες $< 2 \rightarrow$ «Αρχάριος»

**Σχήμα 5.10:** Παράδειγμα αντιστοίχισης στατικού προφίλ (3ο τμήμα του ηλεκτρονικού ερωτηματολογίου).

ισορροπίας του προφίλ. Τα πειραματικά αποτελέσματα καταδεικνύουν ότι, αφότου έχει απαντηθεί ένα δείγμα περίπου 50% των ερωτήσεων του ερωτηματολογίου, είναι δυνατό να επιτευχθεί μία κατάσταση ισορροπίας που να διέπεται από στέρεα, στατικά, αρχικά προφίλ χρήστη με αρκετά μεγάλη εμπιστοσύνη.

Κατά συνέπεια, ολόκληρο το βήμα της προεπεξεργασίας διέπεται από μεγάλη ακρίβεια και αξιοπιστία, αν και γενικά είναι δύσκολο να το διαχειριστούμε σε τέτοιες περιπτώσεις, όπου τα δεδομένα χαρακτηρίζονται από πολυάριθμα, μετρήσιμα χαρακτηριστικά γνωρίσματα, όπως είναι οι απαντήσεις στα ηλεκτρονικά ερωτηματολόγια. Αυτό συμβαίνει, όταν πολλαπλά ανεξάρτητα χαρακτηριστικά γνωρίσματα χαρακτηρίζουν τα δεδομένα και κατά συνέπεια μπορούν να καθοριστούν περισσότερες από μία μετρικές ομοιότητας ή ανομοιότητας που να έχουν νόημα. Μια αποδοτική αντιμετώπιση του προβλήματος είναι η ελάττωση των διαστάσεων εισαγωγής [142], που μπορεί να ολοκληρωθεί με το να αγνοηθούν μερικά από τα διαθέσιμα χαρακτηριστικά γνωρίσματα/απαντήσεις και είναι και ο τρόπος που ακολουθείται εν τω παρόντι, σύμφωνα με όσα παρουσιάστηκαν στην ενότητα 4.2.2 της διατριβής. Επιπρόσθετα, εκτελέσαμε και μια στατιστική ανάλυση συσχέτισης της σπουδαιότητας των χαρακτηριστικών γνωρίσματων. Η βασική αρχή που ακολουθήθηκε είναι ότι, ενώ αναμένουμε τα στοιχεία ενός δεδομένου συνόλου ηλεκτρονικών ερωτηματολογίων να παρουσιάζουν τυχαίες αποστάσεις μεταξύ τους, σύμφωνα με τα περισσότερα χαρακτηριστικά γνωρίσματα, αναμένουμε, επίσης, να εμφανίζουν μικρές αποστάσεις σύμφωνα με τα χαρακτηριστικά γνωρίσματα που τα συσχετίζουν. Στηριζόμαστε σε αυτήν τη διαφορά κατά την κατανομή των τιμών απόστασης, προκειμένου να προσδιορίσουμε το εννοιολογικό πλαίσιο ενός συνόλου στοιχείων, δηλαδή τον υποχώρο στον οποίο ορίζεται καλύτερα το σύ-

νολο και παρέχει τα σημαντικότερα αποτελέσματα από την άποψη της σημασιολογικής σαφήνειας.

Το αποτέλεσμα της παραπάνω διαδικασίας είναι ένα σύνολο 44 σημαντικών ερωτήσεων ηλεκτρονικών ερωτηματολογίων. Η τελική έξοδος μετά και από την εφαρμογή των βαρών είναι η εξαγωγή μιας “1-1” σχέσης μεταξύ των προφίλ και των χρηστών. Με αυτό τον τρόπο κάθε τελικός χρήστης ταξινομείται σε ένα αρχικό, στατικό προφίλ, που χαρακτηρίζει τη συμπεριφορά του, τα ενδιαφέροντά του και την περαιτέρω επεξεργασία/διαχείρισή του από το σύστημα. Αυτός ο ίδιαίτερος χαρακτηρισμός προφίλ αποτελεί τη βάση της επόμενης ευφυούς διαδικασίας συσταδοποίησης, η οποία περιλαμβάνει την έννοια της εξαγωγής και της ολοκλήρωσης ενός προφίλ. Στο Σχήμα 5.11 παρουσιάζουμε ένα ενδεικτικό δείγμα της στατικής αντιστοίχισης των τελικών χρηστών, που εξάγεται από την προηγουμένως αναλυθείσα διαδικασία και σύμφωνα με τις απαντήσεις των χρηστών που συλλέγονται από τα τρία ανεξάρτητα μέρη των ηλεκτρονικών ερωτηματολογίων του συστήματος.

A/A	Επαγγελματική Κατάταση	Προσωπικό ICT Υπόβαθρο	Χρήση ICT στη Διδασκαλία
1. 509	αρχάριος/beginner	ακαθόριστο/unspecified	αρχάριος/beginner
2. 659	αρχάριος/beginner	ακαθόριστο/unspecified	αρχάριος/beginner
3. 807	έμπειρος/expert	έμπειρος/expert	αρχάριος/beginner
4. 808	ακαθόριστο/unspecified	αρχάριος/beginner	αρχάριος/beginner
5. 809	έμπειρος/expert	έμπειρος/expert	προχωρημένος/advanced
6. 811	έμπειρος/expert	έμπειρος/expert	προχωρημένος/advanced
7. 813	έμπειρος/expert	έμπειρος/expert	προχωρημένος/advanced
8. 814	έμπειρος/expert	έμπειρος/expert	προχωρημένος/advanced

**Σχήμα 5.11:** Αρχική αντιστοίχιση στατικών προφίλ - χαρακτηρισμός χρηστών για τρία ανεξάρτητα μέρη ενός ηλ. ερωτηματολογίου.

### 5.5.5.2 Δυναμική εξαγωγή προφίλ χρηστών

Σε αυτό το σημείο, μετά και από την ανάθεση των στατικών προφίλ σε κάθε χρήστη σύμφωνα με τις απαντήσεις που δίνονται στα ηλεκτρονικά ερωτηματολόγια, θεωρούμε ότι οι εμπειρογνώμονες έχουν επιτυχώς αποδώσει σημασιολογικές ετικέτες στο πρώτο μέρος του συνόλου των υπό εξέταση δεδομένων. Το επόμενο βήμα αποτελείται από την εφαρμογή των ευφυών τεχνικών συσταδοποίησης, προκειμένου να συσταδοποιηθούν παρόμοια προφίλ στο σύνολο των δεδομένων. Τα παραδοσιακά μοντέλα ταξινόμησης δεν βρίσκουν εφαρμογή στην περίπτωση αυτή, καθώς η ίδια η φύση των δεδομένων υπαγορεύει μια ευφυή μεθοδολογία ταξινόμησης. Αυτό το βήμα είναι απαραίτητο για την απρόσκοπη και ανεπίβλεπτη λειτουργία του συστήματος, όπου δεν είναι διαθέσιμο κανενός είδους χειρωνακτικό σημασιολογικό μαρκάρισμα. Φυσικά, στο σημείο αυτό είναι δυνατόν να προκύψουν διάφορα προφίλ και κατά συνέπεια τα οποία αποτελέσματα δε θεωρούνται άμεσα συνδεδεμένα με τον μέχρι τώρα χειρωνακτικό χαρακτηρισμό που παρέχεται από τους εμπειρογνώμονες του συστήματος. Η τελευταία προσέγγιση μπορεί να οδηγήσει σε περισσότερα από τρία προφίλ χρηστών και απαιτεί την εφαρμογή τεχνικών κατηγοριοποίησης, προκειμένου να αναγνωρισθούν τα κυριαρχα προφίλ. Το τελευταίο είναι απαραίτητο, προκειμένου να είμαστε σε θέση να συγχρίνουμε τα αποτελέσματα μεταξύ των δύο φάσεων εξαγωγής των προφίλ.

Η λεπτομερής θεωρητική ανάλυση της χρησιμοποιούμενης τεχνικής συσταδοποίησης παρουσιάστηκε στην ενότητα 4.2.1. Στην προσέγγιση αυτή και ειδικά στην περίπτωση της εξαγωγής προφίλ χρηστών από ένα σύστημα ηλεκτρονικών ερωτηματολο-

## Κεφάλαιο 5. Προσωποποιημένη πρόσβαση

γίων, τα αποτελέσματα της εφαρμογής της σε ένα μέρος του συνόλου των δεδομένων επεκτάθηκαν στη συνέχεια σε ολόκληρο το σύνολο των δεδομένων. Η απόδοση της προτεινόμενης μεθοδολογίας συγκρίνεται τελικά με το προηγούμενο βήμα της σταθερής εξαγωγής προφίλ, χρησιμοποιώντας τους προκαθορισμένους χαρακτηρισμούς προφίλ ως μία εκ των προτέρων γνωστή σημασιολογική πληροφορία ετικετών.

Στη συνέχεια, παρουσιάζουμε την προτεινόμενη εφαρμογή του αλγορίθμου συσταδοποίησης στο σύνολο των στοιχείων του συστήματός μας, χρησιμοποιώντας την Ευκλείδια απόσταση ως μετρική απόστασης. Ο αλγόριθμος συσταδοποίησης εφαρμόστηκε αρχικά σε ένα μικρό μέρος του συνόλου των δεδομένων (ειδικότερα σε ένα 10% των χρηστών του συστήματος), που περιλάμβανε 100 στοιχεία/χρήστες και 44 σημαντικά χαρακτηριστικά γνωρίσματα/ερωτήσεις. Ολόκληρο το σύνολο των δεδομένων αποτελείτο από απαντήσεις 1000 χρηστών στις ερωτήσεις των ηλεκτρονικών ερωτηματολογίων που συμπληρώθηκαν σε 8 ευρωπαϊκές χώρες και σε 7 ευρωπαϊκές γλώσσες, όπως παρουσιάζεται στο Σχήμα 5.12, πάνω στο ίδιο σύνολο των 44 επιλεγμένων χαρακτηριστικών γνωρισμάτων. Προηγουμένως, η στατιστική ανάλυση και η επιλογή της



Σχήμα 5.12: Υποστηριζόμενες γλώσσες δεδομένων.

συσχέτισης των χαρακτηριστικών γνωρισμάτων κατέδειξαν τα 44 σημαντικότερα από αυτά, τα οποία και εξετάστηκαν στη συνέχεια. Αν και τα παραπάνω σημαντικά χαρακτηριστικά γνωρίσματα αποτέλεσαν μόνο ένα 25% (44 από συνολικά 176 χαρακτηριστικά γνωρίσματα), αυτό αποδείχθηκε επαρκές και αποδοτικό κατά τη διαδικασία. Τα γνωρίσματα αυτά αντιστοιχούν σε ένα σύνολο ερωτήσεων ερωτηματολογίων μαζί με τις πιθανές επιλογές απάντησής τους, τα οποία συνοψίζονται στον ακόλουθο Πίνακα 5.2 και έχουν θεωρηθεί ότι είναι ενδεικτικά της διαδικασίας εξαγωγής των προφίλ. Τα χαρακτηριστικά γνωρίσματα (πρώτη στήλη) έχουν συσταδοποιηθεί σύμφωνα με το αντίστοιχο αναγνωριστικό της ερώτησης (δεύτερη στήλη) του ηλεκτρονικού ερωτηματολογίου. Για κάθε λεκτική περιγραφή ερώτησης που παρουσιάζεται στην τρίτη στήλη του Πίνακα 5.2 παρουσιάζεται ο αριθμός/αναγνωριστικό της σχετικής ερώτησης και τα χαρακτηριστικά γνωρίσματα που χρησιμοποιούνται κατά τη διαδικασία της συσταδοποίησης.

Σύμφωνα με την προηγούμενη ανάλυση τα ανωτέρω στοιχεία ανήκουν στην πραγματικότητα και στις τρεις σταθερές κατηγορίες των στατικών προφίλ, αλλά αυτή η πληροφορία μαρκαρίσματος δεν χρησιμοποιήθηκε κατά τη διάρκεια της αυτόματης συσταδοποίησης. Εν τούτοις, οι ετικέτες αυτές χρησιμοποιήθηκαν για την αξιολόγηση της ποιότητας της διαδικασίας συσταδοποίησης πριν από την προβολή/επέκταση των αποτελεσμάτων σε ολόκληρο το σύνολο των δεδομένων, όπως αυτή περιγράφεται με λεπτομέρειες στη δημοσίευση [181]. Πιο συγκεκριμένα, κάθε ανιχνευμένη συστάδα κατατάχτηκε στην εκάστοτε στατική κατηγορία που την εξουσίασε. Στη γενική περίπτωση οι προσδιορισμένες συστάδες καθορίζουν συγκεκριμένα ενδιαφέροντα και προφίλ, τα οποία δεν αντιστοιχούν απαραίτητα στις εκ των προτέρων γνωστές κατηγορίες που χρησιμοποιούνται κατά τη διάρκεια της πρώτης φάσης. Αυτές οι συστάδες

**Πίνακας 5.2:** Επιλογή χαρακτηριστικών γνωρισμάτων.

Χαρακτηριστικό γνωρισμάτων	Αναγνωριστικό ερώτησης	Περιγραφή
1	125	Are you a teacher dedicated to or working in Special Education Needs?
2, 3, 4	126, 127, 128	Qualification/training in Special Education Needs?
5	130	Teaching Experience
6	133	Do you have a computer at home?
7	134	Do you have access to the Internet from your home?
8	139	How often do you personally use your Internet connection at home?
9, 10, 11, 12, 13	141, 142, 143, 144, 145	For which of the following did you use the computer at least once in the past month?
14, 15, 16, 17, 18, 19, 20	147, 148, 149, 150, 151, 152, 153	Which of the following tasks have you performed at least once, without any help?
21	155	Are there any computers in your work environment?
22, 23	300, 301	How often did you use the computer last week at the school?
24, 25	168, 169	Do you have access to the Internet or educational software in your work environment?
26, 27	174, 175	In your teaching, how many hours a week, on average, do you use the Internet or educational software with your students?
28, 29	176, 177	Do you use the Internet for search and retrieval of information relating to the needs and problems faced by SEN <sup>4</sup> students?
30, 31	182, 183	Do you use the Internet from the school in order to find additional sources of educational material?
32, 33	400, 401	Do you use the Internet to connect with other schools?
34, 35, 36, 37	261, 262, 264, 265	Is your post permanent - temporary?
38, 39, 40, 41, 42, 43	267, 268, 270, 271, 273, 274	Age of your students
44	275	Area served by your school

είναι χρήσιμες για την παραγωγή συνεργατικών συστάσεων (collaborative recommendations) εκπαιδευτικού περιεχομένου στους τελικούς χρήστες σε ένα μεταγενέστερο στάδιο, όπως περιγράφεται και στο [269]. Τα αποτελέσματα της αυτόματης συσταδοποίησης παρουσιάζονται στους Πίνακες 5.3, 5.4 και 5.5, ενώ οι αριθμοί που χωρίζονται από κόμματα μέσα στην παρένθεση καταδεικνύουν τα στοιχεία που ανήκουν σε καθεμία από τις τρεις κατηγορίες προφίλ σε κάθε βήμα. Η εκτέλεση της αρχικής συσταδοποίησης σε ένα μόνο υποσύνολο 10% του συνόλου των δεδομένων είναι όχι μόνο υπολογιστικά αποδοτικότερη και σοφή, αλλά είναι επίσης και καλύτερη από την άποψη της ποιότητας και της απόδοσης, όταν συγχρίνεται με την προσέγγιση της εφαρμογής της ιεραρχικής διαδικασίας συσταδοποίησης σε ολόκληρο το σύνολο των δεδομένων. Παρόλο που η συσταδοποίηση αυτού του 10% του συνόλου δεδομένων οδήγησε σε διαφορετικούς αριθμούς πιθανών συστάδων, τα βέλτιστα αποτελέσματα επιτεύχθηκαν για έναν αριθμό από εγγέα συστάδες, όπως φαίνεται και στους ακόλουθους Πίνακες

## Κεφάλαιο 5. Προσωποποιημένη πρόσβαση

**5.3, 5.4 και 5.5**, όπου παρουσιάζονται τα αποτελέσματα της συσταδοποίησης για τρεις παραλλαγές των συστάδων (3, 5 και 9 συστάδες):

**Πίνακας 5.3:** Αποτελέσματα συσταδοποίησης για 100 χρήστες - 3 συστάδες.

Συστάδες	Στοιχεία	%
1η	(2, 6, <b>9</b> )	(11.77%, 35.29%, <b>52.94%</b> )
2η	(11, 2, <b>25</b> )	(28.95%, 5.26%, <b>65.79%</b> )
3η	(14, 1, <b>30</b> )	(31.11%, 2.22%, <b>66.67%</b> )

**Πίνακας 5.4:** Αποτελέσματα συσταδοποίησης για 100 χρήστες - 5 συστάδες.

Συστάδες	Στοιχεία	%
1η	(3, 1, <b>7</b> )	(27.27%, 9.09%, <b>63.64%</b> )
2η	(5, 1, <b>8</b> )	(35.72%, 7.14%, <b>57.14%</b> )
3η	(5, 1, <b>13</b> )	(26.32%, 5.26%, <b>68.42%</b> )
4η	(5, 9, <b>11</b> )	(20.00%, 36.00%, <b>44.00%</b> )
5η	(11, 1, <b>19</b> )	(35.48%, 3.23%, <b>61.29%</b> )

**Πίνακας 5.5:** Αποτελέσματα συσταδοποίησης για 100 χρήστες - 9 συστάδες.

Συστάδες	Στοιχεία	%
1η	(1, 1, <b>4</b> )	(16.66%, 16.66%, <b>66.66%</b> )
2η	(0, 1, <b>6</b> )	(0.00%, 14.28%, <b>85.71%</b> )
3η	(4, 2, <b>5</b> )	(36.36%, 18.18%, <b>45.45%</b> )
4η	(3, 3, <b>6</b> )	(25.00%, 25.00%, <b>50.00%</b> )
5η	(4, 2, <b>5</b> )	(36.36%, 18.18%, <b>45.45%</b> )
6η	( <b>8</b> , 4, 5)	( <b>47.05%</b> , 23.52%, 29.41%)
7η	( <b>4</b> , 1, 4)	( <b>44.44%</b> , 11.11%, <b>44.44%</b> )
8η	(3, <b>10</b> , 6)	(15.78%, <b>52.63%</b> , 31.57%)
9η	(1, 4, 3)	(12.50%, <b>50.00%</b> , 37.50%)

Πιο συγκεκριμένα, ο Πίνακας 5.3 παρουσιάζει τα αποτελέσματα συσταδοποίησης 100 χρηστών, όπου ο ιεραρχικός αλγόριθμος συσταδοποίησης ολοκληρώθηκε, όταν έφθασε σε ένα κατώτατο όριο 3 συστάδων. Η πρώτη συστάδα αποτελείται από 17 χρήστες, δηλαδή 2 χρήστες που ανήκουν στην κατηγορία των Ειδημόνων/Εμπειρογνωμόνων, 6 στην κατηγορία των Αρχαρίων και 9 στην ενδιάμεση κατηγορία των Προηγμένων χρηστών. Η αντίστοιχη κατανομή των ποσοστών δείχνει σαφώς ότι οι 9 προηγμένοι χρήστες εξουσιάζουν την πρώτη συστάδα. Η δεύτερη συστάδα αποτελείται από 38 χρήστες: 11 Εμπειρογνώμονες, 2 Αρχάριους και 25 Προηγμένους. Οι τελευταίοι είναι κυριαρχούντες και σε αυτή την συστάδα, ενώ από την άποψη των ποσοστών η κυριαρχία τους είναι δεδομένη (με 65,79%). Τέλος, η τρίτη συστάδα περιέχει 45 χρήστες, 14 εκ των οποίων είναι Εμπειρογνώμονες, 1 είναι Αρχάριος και 30 είναι Προηγμένοι. Η κυριαρχία των Προηγμένων είναι και εδώ εμφανής, δεδομένου ότι το 66,67% ως ποσοστό τους δίνει ένα σαφές προβάδισμα και χαρακτηρίζει τη συστάδα.

Ο Πίνακας 5.4 παρουσιάζει τα αποτελέσματα συσταδοποίησης των ίδιων 100 χρηστών, εντούτοις ένα νέο κατώτατο όριο 5 συστάδων ολοκληρώνει τον αλγόριθμο συσταδοποίησης 2 βήματα νωρίτερα. Σε αυτήν την περίπτωση όλες οι συστάδες περιέχουν λιγότερους χρήστες σε σύγκριση με την προηγούμενη περίπτωση. Η πρώτη

συστάδα αποτελείται από 11 χρήστες: 3 Εμπειρογνώμονες, 1 Αρχάριο και 7 Προηγμένους. Η αντίστοιχη κατανομή των ποσοστών δείχνει σαφώς ότι οι 7 Προηγμένοι χρήστες εξουσιάζουν την πρώτη συστάδα. Το ίδιο ισχύει επίσης για τη δεύτερη συστάδα με 14 χρήστες, όπου 5 χρήστες ανήκουν στους Εμπειρογνώμονες, 1 χρήστης είναι Αρχάριος και 8 χρήστες βρίσκονται στο ενδιάμεσο στάδιο, δηλ. Προηγμένοι. Η τρίτη, τέταρτη και πέμπτη συστάδα καταδεικνύουν όλες την υπεροχή των Προηγμένων χρηστών: η τρίτη συστάδα περιέχει 19 χρήστες (που διανέμονται αναλόγως σε 5 Εμπειρογνώμονες, 1 Αρχάριο και 13 Προηγμένους), η τέταρτη συστάδα περιέχει 25 χρήστες (5 Εμπειρογνώμονες, 9 Αρχάριοι και 11 Προηγμένοι) και η πέμπτη συστάδα περιέχει 31 χρήστες (11 Εμπειρογνώμονες, 1 Αρχάριος και 19 Προηγμένοι).

Στον Πίνακα 5.5, τώρα, τα αποτελέσματα του βήματος της συσταδοποίησης καταδεικνύουν μία σαφή ελλοχεύουσα τάση στα δεδομένα εισόδου του συστήματος: οι χρήστες του χαρακτηρίζονται από ενδιάμεσες ICT δεξιότητες και ενδιάμεσου επιπέδου σχετική εμπειρία. Αυτή η παρατήρηση είναι εξαιρετικά εμφανής στην τρίτη στήλη του Πίνακα 5.5, η οποία δείχνει σαφώς ότι οι περισσότεροι χρήστες του συστήματος ανήκουν στο στατικό, ενδιάμεσο προηγμένο προφίλ. Οι πρώτες δύο συστάδες που προσδιορίζονται από τον αλγόριθμό μας κυριαρχούνται σαφέστατα από την τρίτη κατηγορία προφίλ, δηλ. τους Προηγμένους χρήστες. Επιπλέον, οι συστάδες 3, 4 και 5 δείχνουν και αυτές μια ζεκάθαρη πλειοφηφία της τρίτης κατηγορίας στα στοιχεία τους. Συνεπώς, 5 από τις 9 συστάδες (ποσοστό 55.55%) δείχνουν ένα σαφές πλεονέκτημα των Προηγμένων χρηστών. Επιπλέον, η συστάδα 7 ενεργεί ως μεσάζων μεταξύ των Προηγμένων και των Εμπειρογνώμονων, καθώς εμφανίζει μια ισορροπία στοιχείων μεταξύ των δύο κατηγοριών προφίλ. Οι συστάδες 8 και 9 κυριαρχούνται από την κατηγορία προφίλ Αρχαρίων, ενώ η συστάδα 6 αποτελεί έναν ζεκάθαρο αντιπρόσωπο των Εμπειρογνωμόνων.

Η παραπάνω προσέγγιση συσταδοποίησης διαμορφώνει τη βασική διαδικασία, με τη βοήθεια της οποίας κάθε τελικός χρήστης του συστήματος ταξινομείται αυτόματα σε μια συγκεκριμένη κατηγορία προφίλ, που χαρακτηρίζει τη συμπεριφορά του, τα μελλοντικά ενδιαφέροντά του και τις επιλογές του μέσα στο σύστημα. Επιπλέον, κάθε ανιχνευμένη συστάδα σχετίζεται άμεσα με τις απαντήσεις και τα συγκεκριμένα σχόλια των ηλεκτρονικών ερωτηματολογίων. Εν ολίγοις, χρησιμοποιήσαμε τη συσταδοποίηση των αποτελεσμάτων, για να ταξινομήσουμε τις απαντήσεις των χρηστών στα συγκεκριμένα μέρη των ηλεκτρονικών ερωτηματολογίων, εξάγοντας τις σχετικές πληροφορίες από τα εκάστοτε προφίλ χρήστη. Η αυτόματη εξαγωγή των συστάδων παρείχε στους εμπειρογνώμονες του συστήματος τη δυνατότητα να διαχρίνουν και να ξεχωρίσουν τις απαντήσεις των συσταδοποιημένων χρηστών σύμφωνα με πιο συγκεκριμένα και αξιόπιστα προφίλ χρηστών. Ειδικότερα, συνδυάσαμε πληροφορίες για τα προφίλ των χρηστών προερχόμενες από διαφορετικά μέρη των ηλεκτρονικών ερωτηματολογίων, σεβόμενοι την εισαγωγή στοιχείων των χρηστών. Παράδειγμα στατιστικής ανάλυσης της τελευταίας παρουσιάζεται στο Σχήμα 5.13, ενώ αποτελέσματα κατηγοριοποίησης ενός μέρους των χρηστών του συστήματος με βάση τα ανιχνευμένα προφίλ για τα 3 επιμέρους τμήματα των ηλεκτρονικών ερωτηματολογίων παρατίθενται στο Σχήμα 5.14.

Θα πρέπει, επίσης, να αναφερθεί ότι η προτεινόμενη μεθοδολογία μπορεί να συνδυαστεί και με τεχνικές μάθησης νευρωνικών δικτύων, χρησιμοποιώντας τη συσταδοποίηση για την αρχικοποίηση ενός ταξινομητή νευρωνικών δικτύων, ο οποίος θα περιέχει πληροφορίες προσωποποίησης για την εξαγωγή του τοπικού ενδιαφέροντος των χρηστών. Εντούτοις, κάτι τέτοιο δε θεωρείται ότι αποτελεί αντικείμενο της πα-

## Κεφάλαιο 5. Προσωποποιημένη πρόσβαση

### Ερωτήσεις για τον εκπαιδευτικό

A. Παιδαγωγική χρήση των Τεχνολογιών Πληροφορικής

#	Profile	Color
1.	Beginner	
2.	Advanced	
3.	Expert	
4.	Unspecified	

A.1. Αλλαγές στις καθημερινές δραστηριότητες του διδακτικού προσωπικού που οφείλονται στη χρήση των Τεχνολογιών της Πληροφορικής και Επικοινωνίας		
A.1.1. Η ευρεία χρήση των ΤΠΕ ως ?εργαλείων? στο σχολείο θα συντελέσει σε αλλαγές απόψεων σχετικά με την πρακτική στη διδασκαλία. Παρακαλούμε δηλώστε πού επικεντρώνεται αυτή η αλλαγή.		
Αλλαγές στη σάση του διδακτικού προσωπικού είναι απαραίτητες	Πολύ συχνά	Σπάνια
73	16.11%	8   1.77%
180	39.74%	31   6.84%
135	29.8%	25   5.52%
0	0%	0   0%

**Σχήμα 5.13:** Παράδειγμα στατιστικής ανάλυσης μιας ερώτησης ηλεκτρονικού ερωτηματολογίου, συνδεδεμένης με τις κατηγορίες των προφίλ.



### Users' Profiles - Προφίλ χρηστών B questionnaire

	Αρχάριοι/Beginners	Προχωρημένοι/Advanced	Έμπειροι/Experts
Επαγγελματική Κατάρτιση	84	223	168
Προσωπικό ICT Υπόβαθρο	197	215	63
Χρήση ICT στη Διδασκαλία	395	11	69

**Σχήμα 5.14:** Συνδυασμένη πληροφορία προφίλ με βάση τα αποτελέσματα της συσταδοποίησης και για τα τρία μέρη ενός ηλεκτρονικού ερωτηματολογίου.

ρούσης διατριβής, η οποία εστιάζει στη χρησιμοποίηση της συσταδοποίησης προς την παροχή ικανοποιητικών συνεργατικών συστάσεων (collaborative recommendations) περιεχομένου προς τους τελικούς χρήστες και για το λόγο αυτό δε θα αναπτυχθεί περαιτέρω. Σε αυτό το πλαίσιο, σύμφωνα με το προφίλ στο οποίο ανήκει τελικά κάθε χρήστης, του προσφέρεται και το κατάλληλο εκπαιδευτικό περιεχόμενο. Το τελευταίο επιλέγεται από τους εμπειρογνώμονες του συστήματος ανάλογα με τις ικανότητες/δυνατότητες που χαρακτηρίζουν το εκάστοτε προφίλ. Τέλος, θα πρέπει να σημειωθεί στο σημείο αυτό ότι εξαιτίας της δεδομένης ευελιξίας, αλλά και για λόγους προστασίας κρίσιμων προσωπικών δεδομένων το βήμα του χαρακτηρισμού των χρηστών που περιγράφηκε σε αυτή την υποενότητα παρέχεται μόνο ως προστιθέμενη αξία στους χρήστες του συστήματος που είναι πρόθυμοι να το χρησιμοποιήσουν. Κατάλληλες διαδικασίες επαλήθευσης εξασφαλίζουν ότι το φιλτράρισμα των χαρακτηριστικών γνωρισμάτων και του εκπαιδευτικού πολυμεσικού περιεχομένου επιτρέπονται μόνο με τη συγχατάθεση του εκάστοτε τελικού χρήστη.

## 5.6 Ανάκτηση και Κατάταξη Προσωποποιημένης Πληροφορίας

Επιστρέφοντας στο πρόβλημα που σχετίζεται με την εξαγωγή των προτιμήσεων ή και των προφίλ χρηστών με βάση το εννοιολογικό τους πλαίσιο, μια άκρως ενδιαφέρουσα εφαρμογή της μεθοδολογίας που περιγράφηκε απαντάται κατά τη διαδικασία της ανάκτησης περιεχομένου, εν συνεχείᾳ, δε, και κατά την κατάταξη των αποτελεσμάτων αυτής.

### 5.6.1 Επίδραση εννοιολογικού πλαισίου

Ειδικότερα, στο πλαίσιο ενός συστήματος ανάκτησης πληροφορίας και σύμφωνα με όσα έχουν παρουσιαστεί στα προγενέστερα κεφάλαια της διατριβής ορίζουμε το εννοιολογικό πλαίσιο της σημασιολογικής ανάκτησης των χρηστών κατά το χρόνο εκτέλεσης ως το σύνολο των εννοιών, που, άμεσα ή έμμεσα, διαδραματίζουν κάποιο ρόλο κατά την αλληλεπίδραση ενός χρήστη  $\hat{u}$  με το σύστημα κατά τη διάρκεια μιας συνόδου ανάκτησης. Επομένως, σε κάθε χρονικό σημείο  $t$  αναπαριστούμε το εννοιολογικό πλαίσιο ανάκτησης  $\hat{C}_t(\hat{u})$  ως ένα ασαφές σύνολο εννοιών. Ο χρόνος υπολογίζεται με τον αριθμό των αιτημάτων χρήστη κατά τη διάρκεια μιας τέτοιας συνόδου. Μιας και η συσχέτιση του εννοιολογικού πλαισίου με κάποιον χρήστη είναι, εν γένει, ζεκάθαρη, στο εξής θα παραλείψουμε τη μεταβλητή  $\hat{u}$  και θα χρησιμοποιούμε το  $\hat{C}_t$ , εφόσον φυσικά παραμένει ίδιο και προφανές το νόημα.

Στην προσέγγισή μας το σημασιολογικό εννοιολογικό πλαίσιο κατά το χρόνο εκτέλεσης  $\hat{C}_t$  δημιουργείται ως ένας συνδυασμός των εννοιών που εμπλέκονται σε διαδοχικές αιτήσεις ή/και ερωτήσεις χρηστών με έναν τέτοιο τρόπο, έτσι ώστε η σοβαρότητα των εννοιών να ελαττώνεται με το πέρασμα του χρόνου. Η διαδικασία αυτή προσομοιώνει τη φυσική αλλαγή του ενδιαφέροντος των χρηστών με την πάροδο του χρόνου. Έστω ότι ορίζουμε το σύνολο όλων των διαθέσιμων χρονικών τμημάτων ως  $\mathcal{T} = \{1, \dots, M\}$ , i.e.  $t \in \mathcal{T}$ . Ορίζουμε, επίσης, το  $Q_t$  ως το ασαφές σύνολο των εννοιών που δημιουργείται αμέσως μετά από κάθε επερώτηση χρήστη σε μια δεδομένη χρονική στιγμή  $t$ , δηλ.  $Q_t \in \mathcal{F}_S$ . Σύμφωνα με την ανάλυση της ενότητας 5.5.2 θα έχουμε:

$$Q_t = Q_t^k \cup Q_t^v \cup Q_t^+ \cup Q_t^b \quad (5.37)$$

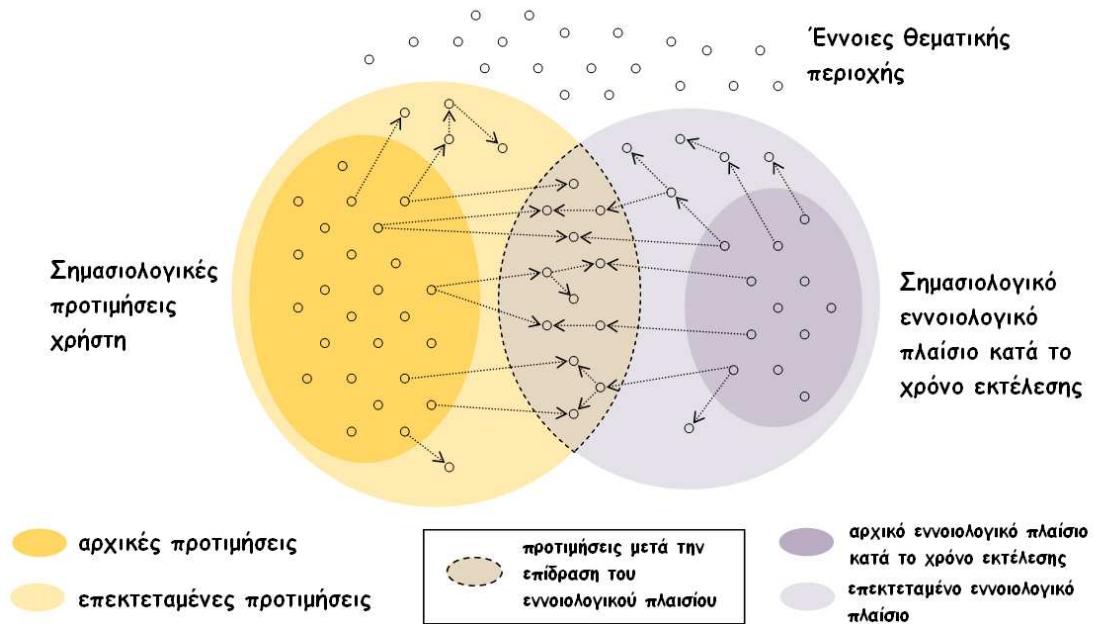
όπου:

$$Q^k = \bigcup_{t \in \mathcal{T}} Q_t^k, \quad Q^v = \bigcup_{t \in \mathcal{T}} Q_t^v, \quad Q^+ = \bigcup_{t \in \mathcal{T}} Q_t^+, \quad Q^b = \bigcup_{t \in \mathcal{T}} Q_t^b \quad (5.38)$$

Στη συνέχεια, ορίζεται το εννοιολογικό πλαίσιο κατά το χρόνο εκτέλεσης  $\hat{C}_t$  στο χρόνο ερώτησης  $t$  ως ο συνδυασμός του καινούριου ασαφούς συνόλου  $Q_t$  και του εννοιολογικού πλαίσιο κατά το χρόνο εκτέλεσης  $\hat{C}_{t-1}$ , το οποίο υπολογίστηκε στο αμέσως προηγούμενο βήμα, όπου τα βάρη του εννοιολογικού πλαισίου που υπολογίστηκαν στο βήμα  $t-1$  μειώνονται αυτόματα κατά ένα παράγοντα μείωσης (decay factor)  $\beta \in [0, 1]$ . Συνακόλουθα, σε οποιαδήποτε χρονική στιγμή  $t > 1$  ενημερώνουμε το  $\hat{C}_t$  ως:

$$\hat{C}_t = \beta \hat{C}_{t-1} + (1 - \beta) Q_t \quad (5.39)$$

Στην παραπάνω εξίσωση χρησιμοποιούνται το αλγεβρικό άθροισμα και το αλγεβρικό γινόμενο για την υλοποίηση της πρόσθεσης και του πολλαπλασιασμού μεταξύ δύο οποιονδήποτε ασαφών συνόλων [177]. Είναι εύκολο να παρατηρήσει κανείς ότι  $\hat{C}_t \in$



Σχήμα 5.15: Σημασιολογική επέκταση των προτιμήσεων χρηστών και του εννοιολογικού πλαισίου χατά το χρόνο εκτέλεσης.

Ο  $\mathcal{C}_0$  και ότι η εξίσωση (5.39) ισχύει για  $\widehat{C}_0 = \emptyset$  και  $\widehat{C}_1 = Q_1$ . Το  $Q_t$  αποτελείται από μια ποικιλία αιτήσεων χρηστών και το ασαφές σύνολο του εννοιολογικού πλαισίου χατά το χρόνο εκτέλεσης δε χρησιμοποιείται, για να επαναδιατυπώσουμε την ερώτηση, αλλά για να εστιάσουμε στο σύνολο των προτιμήσεων. Το γεγονός αυτό διαφοροποιεί την προσέγγισή μας σε σχέση με κλασικές τεχνικές διαδραστικής ανατροφοδότησης, όπως οι [51] και [136].

Στο σημείο αυτό θεωρούμε ότι έχουμε ολοκληρώσει την παρουσίαση τόσο της offline αναπαράστασης των προτιμήσεων χρηστών  $P$ , που σχετίζεται με το σύνολο του ιστορικού χρήστη  $H$  κάθε χρήστη, όσο και του εννοιολογικού πλαισίου χατά το χρόνο εκτέλεσης  $\widehat{C}_t$ . Η επιλεκτική ενεργοποίηση των προτιμήσεων χρηστών βασίζεται στην εύρεση των σημασιολογικών μονοπατιών μεταξύ των εννοιών προτίμησης και εννοιολογικού πλαισίου (βλ. και ενότητα 5.4.3). Τα μονοπάτια αυτά χρησιμοποιούν την κατασκευασμένη σημασιολογική σχέση  $\widehat{T}$  μεταξύ του συνόλου των εννοιών  $S$  που είναι διαθέσιμες στην οντολογία της περιοχής  $\mathcal{O}$ . Κεντρικό σημείο της προτεινόμενης στρατηγικής μας αποτελεί η σημασιολογική επέκταση μέσω μιας ασαφούς σημασιολογικής διατομής μεταξύ των προτιμήσεων χρηστών  $P$  και του σημασιολογικού εννοιολογικού πλαισίου χατά το χρόνο εκτέλεσης  $\widehat{C}_t$ . Κατά τη διάρκεια της σημασιολογικής επέκτασης ορίζονται στις περιληφθείσες έννοιες τα βάρη προτίμησης και τα βασισμένα στο εννοιολογικό πλαισίο βάρη, τα οποία και μειώνονται βαθμιαία, καθώς η επέκταση απομακρύνεται από τα αρχικά σύνολα. Αυτή η διαδικασία μπορεί να ερμηνευθεί επίσης ως μια ασαφής σημασιολογική διατομή μεταξύ των προτιμήσεων χρηστών και του σημασιολογικού εννοιολογικού πλαισίου χατά το χρόνο εκτέλεσης, όπου το τελικό υπολογισμένο βάρος κάθε έννοιας αντιπροσωπεύει το βαθμό στον οποίο ανήκει σε κάθε σύνολο (βλ. και Σχήμα 5.15).

Αρχικά θα επιχειρήσουμε να ορίσουμε την έννοια της σημασιολογικής επέκτασης (semantic expansion) μιας αφηρημένης έννοιας  $X$  με μία συνάρτηση  $E$ . Η έννοια  $X$  μπορεί να είναι είτε οι προτιμήσεις χρήστη, είτε το ίδιο το εννοιολογικό πλαίσιο χατά

το χρόνο εκτέλεσης, μιας και η προτεινόμενη μεθοδολογία είναι κατάλληλη και για τα δύο. Έστω ότι το  $X$  είναι ένα ασαφές σύνολο από έννοιες στο  $S$ , δηλαδή  $X \in \mathcal{F}_S$ . Τότε (και σε αναλογία με το φορμαλισμό που εισάγαμε στην εξίσωση (5.8)):

$$X_0 = X \quad \text{και} \quad X_{i+1} = X_i \circ \widehat{T}, i > 0 \quad (5.40)$$

Συνακόλουθα:

$$E(X) = X_{\mathcal{L}}, \quad (5.41)$$

όπου το σημείο που σταματά η αναδρομή και συγχλίνει η εξίσωση (5.40) συμβολίζεται με  $\mathcal{L}$ . Η αναδρομή σταματά, όταν το αποτέλεσμα του αμέσως προηγούμενου βήματος είναι ίσο με το αποτέλεσμα του βήματος υπό εξέταση, δηλαδή:  $X_{\mathcal{L}} = X_{\mathcal{L}-1}$ , ή με άλλα λόγια όταν  $X_{\mathcal{L}-1} \circ \widehat{T} = X_{\mathcal{L}-1}$ . Στο σημείο αυτό θα πρέπει να σημειώσουμε ότι στη γενική περίπτωση ο γράφος που ορίζεται από την  $\widehat{T}$  δεν είναι DAG<sup>5</sup> και η επανάληψη μπορεί να μη συγχλίνει σε ένα πεπερασμένο αριθμό βημάτων. Προκειμένου να αποφύγουμε μια τέτοια εξέλιξη, καθώς και μια ανεπιθύμητη retro-feeding επίδραση της επέκτασης (π.χ. οι αρχικές μη-μηδενικές προτιμήσεις χρήστη δε θα πρέπει να αυξηθούν από την επέκταση), μετατρέπουμε την  $\widehat{T}$  σε ακυκλική, πριν ξεκινήσουμε την αναδρομική διαδικασία. Το τελευταίο επιτυγχάνεται με το να απομακρύνουμε τις κατάλληλες ακμές πριν την επέκταση και ειδικότερα αυτές που εξαφανίζουν όλους τους κύκλους και ταυτόχρονα μεγιστοποιούν το παραγόμενο  $E(X)$ . Το γεγονός ότι η παραπάνω εξίσωση συγχλίνει εξασφαλίζεται από αυτό το μετασχηματισμό που επιβάλλεται πριν από το βήμα της ασαφοποίησης και περιγράφεται στον ακόλουθο αλγόριθμο.

Στις παραπάνω σχέσεις το σύμβολο “ο” συμβολίζει την ασαφή σύνθεση (fuzzy composition) μεταξύ μίας ασαφούς σχέσης και ενός ασαφούς συνόλου. Στη γενική περίπτωση και με δεδομένη μια ασαφή σχέση  $\mathcal{R} : \mathcal{X} \times \mathcal{Y} \rightarrow [0, 1]$  και ένα ασαφές σύνολο  $\mathcal{A}' : \mathcal{X} \rightarrow [0, 1]$ , η ασαφής σύνθεση ορίζεται ως:

$$\mathcal{B}' = \mathcal{A}' \circ \mathcal{R} : \mathcal{Y} \rightarrow [0, 1] \quad (5.42)$$

και:

$$\mathcal{B}' = \bigcup_{x \in \mathcal{X}} (\mathcal{A}' \cap \mathcal{R}) \quad \text{or} \quad \mu_{\mathcal{B}'}(y) = u_{x \in \mathcal{X}}(t(\mu_{\mathcal{A}'}(x), \mu_{\mathcal{R}}(x, y))) \quad (5.43)$$

όπου η  $t$  και η  $u$  είναι μια ασαφής  $t$ -νόρμα και μια ασαφής  $t$ -συνορμα, αντίστοιχα.

Στο πλαίσιο της προσέγγισής μας υλοποιούμε την παραπάνω λειτουργία της σύνθεσης με την ακόλουθη διαδικασία:

```

expand_set( $X, E(X)$ )
  for  $x \in S$  do
     $E(X)(x) \leftarrow X(x)$ 
    in_path[x]  $\leftarrow false$ 
  for  $x \in supp(X)$  do
    expand_concept( $x, 0$ )

expand_concept( $x, prev\_x$ )
  in_path[x]  $\leftarrow true$ 
  for  $y \in \{z \in S | \widehat{R}(x, z) > 0\}$  do

```

<sup>5</sup>Κατευθυνόμενος Ακυκλικός Γράφος - Directed Acyclic Graph

```

if not in_path[y] and  $X(y) = 0$  and  $E(X)(y) < 1$  then
     $prev\_y \leftarrow E(X)(y)$ 
     $E(X)(y) \leftarrow (E(X)(y) - \widehat{R}(x, y) * prev\_x) / (1 - \widehat{R}(x, y) * prev\_x)$ 
     $E(X)(y) \leftarrow E(X)(y) + (1 - E(X)(y)) * \widehat{R}(x, y) * E(X)(x)$ 
    if  $E(X)(y) > \varepsilon$  then expand_concept(y, prev_y)
    in_path[x]  $\leftarrow false$ 

```

Ο αλγόριθμος (που για λόγους ευκολότερης αναγνωσιμότητας παρουσιάζεται στο σημείο αυτό με ψευδοκώδικα και στην αναδρομική μορφή του) υλοποιήθηκε στην πραγματικότητα ως μία επανάληψη χρησιμοποιώντας μια στοίβα. Το σύνολο  $supp(X)$  καταδεικνύει την κλασική (σαφή) υποστήριξη του  $X$ , δηλαδή το σύνολο  $\{x \in S | X(x) > 0\}$ . Η ιδιότητα  $in\_path[x]$  κατά τη διαδικασία επέκτασης εννοιών είναι αυτό που εξαλείφει τις σωστές ακμές στην σχέση  $\widehat{R}$ . Με άλλα λόγια, οι ακμές, μέσω των οποίων μία έννοια  $x$  θα συνεισφέρει στην ίδια της την επέκταση, είναι αυτές που απενεργοποιούνται προσωρινά κατά την επανάληψη. Η τιμή  $\varepsilon$  αποτελεί ένα ελάχιστο κατώφλι, κάτω από την οποία η τιμή του  $E(X)(x)$  δεν επεκτείνεται προς τη σημασιολογική γειτονιά της  $x$ .

Μπορεί να αποδειχθεί ότι ο παραπάνω αλγόριθμος επιτυγχάνει την επέκταση με μια πολυπλοκότητα της τάξης του  $O(|supp(X)| \cdot |S| \cdot |supp(\widehat{R})|)$ , όπου το  $supp(\widehat{R})$  απεικονίζει την κλασική (σαφή) υποστήριξη του  $\widehat{R}$ , δηλαδή τα σύνολα ζευγών  $(x, y) \in S \times S$  με  $\widehat{R}(x, y) > 0$ . Παρ' όλ' αυτά στην πράξη το κόστος είναι ακόμα μικρότερο, καθώς το  $E(X)(y)$  μειώνεται γρήγορα κάτω από το  $\varepsilon$  καθώς το  $y$  απομακρύνεται από τις αρχικές έννοιες που έχουν  $X(x) > 0$  (όπου ο όρος “μακριά” (ή κοντά) καθορίζεται από τον αριθμό των ακμών  $\widehat{R}$  που χρειάζονται, για να προσεγγίσουμε το  $y$  από αυτό το σύνολο). Οι πειραματικές μετρήσεις που πραγματοποιήσαμε δείχνουν ότι ο χρόνος που αφιερώνεται στην ίδια την λειτουργία της επέκτασης είναι ασήμαντος σε σύγκριση με το κόστος άλλων λειτουργιών, όπως π.χ. είναι η πρόσβαση στη Βάση Γνώσης (KB).

Το επεκταθέν εννοιολογικό πλαίσιο κατά το χρόνο εκτέλεσης  $E(\widehat{C}_t)$ , καθώς και το επεκταθέν σύνολο από προτιμήσεις χρηστών  $E(P)$  υπολογίζονται σύμφωνα με την εξίσωση (5.41) και κατά συνέπεια η ακριβής έκφραση των σύμφωνα με το εννοιολογικό πλαίσιο υπολογισμένων προτιμήσεων χρηστών  $CP_t$  δίνεται από το αλγεβρικό γινόμενο των δύο ασαφών συνόλων:

$$CP_t = E(P)E(\widehat{C}_t) \quad (5.44)$$

Το  $CP_t$  μπορεί τώρα να θεωρηθεί ως μία συνδυασμένη μετρική της πιθανότητας να είναι προτιμητέα μία έννοια και του κατά πόσο σχετική είναι η έννοια αυτή με το υπό εξέταση εννοιολογικό πλαίσιο. Θα πρέπει να σημειώσουμε εδώ ότι αυτό το ασαφές σύνολο εξαρτάται στην πραγματικότητα τόσο από τους χρήστες όσο και από το χρόνο, δηλαδή  $CP_t(\widehat{u})$ . Στο σημείο αυτό έχουμε επιτύχει την με βάση το εννοιολογικό πλαίσιο αντιστοίχιση των προτιμήσεων. ειδικότερα:  $\Phi(P, \widehat{C}_t) = CP_t$ , όπου  $P \models \Phi(P, \widehat{C}_t)$ , αφού  $CP_t > P$  μόνο όταν το  $E(P)$  έχει προκύψει από το σύνολο  $P$  και  $CP_t < E(P)$ .

### 5.6.2 Κατάταξη προσωποποιημένων αποτελεσμάτων

Με δεδομένο, τώρα, ένα έγγραφο  $d \in \mathcal{D}$  (όπου, όπως έχει ήδη εξηγηθεί, το  $\mathcal{D}$  αντιπροσωπεύει το σύνολο όλων των εγγράφων στο χώρο ανάκτησης), το προλεγόμενο

(predicted) ενδιαφέρον (στο οποίο θα αναφερόμαστε από εδώ και στο εξής ως: προσωπική μετρική κατάταξης (personal ranking measure)  $r_P(d, t)$ ), του χρήστη  $\hat{u}$  για το έγγραφο  $d$  σε ένα δεδομένο στιγμιότυπο  $t$  κατά τη διάρκεια μίας συνόδου λαμβάνει τιμές στο διάστημα  $[0, 1]$ , σύμφωνα με τις προτιμήσεις του  $P$  και υπολογίζεται από την εξίσωση:

$$r_P(d, t) = \cos(S_d, CP_{t-1}) \quad (5.45)$$

όπου  $S_d$  είναι το ασαφές σύνολο των εννοιών που σχετίζονται με το έγγραφο  $d \in \mathcal{D}$  και  $CP_{t-1}$  είναι το σύνολο των με βάση το εννοιολογικό πλαίσιο υπολογισμένων προτιμήσεων, όπως αυτές προέκυψαν κατά τα περιγραφόμενα στην προηγούμενη ενότητα 5.6.1. Η εξίσωση (5.45) ισχύει, μιας και σύμφωνα με τον Egghe [83], θεωρώντας δεδομένα δύο ασαφή σύνολα  $\mathcal{X}, \mathcal{Y} \in \mathcal{F}_S$ , η μετρική ομοιότητάς τους με βάση το συνημίτονο (cosine similarity measure) ορίζεται ως εξής:

$$\cos(\mathcal{X}, \mathcal{Y}) = \frac{|\mathcal{X} \cap \mathcal{Y}|}{\sqrt{|\mathcal{X}| |\mathcal{Y}|}} \quad (5.46)$$

Προκειμένου να μελετήσουμε περαιτέρω την εξίσωση (5.46), υιοθετούμε την επέκταση της έννοιας της πληθυκότητας ενός κλασικού (σαφούς) συνόλου και στην περίπτωση ενός ασαφούς συνόλου. Η τελευταία ορίζεται ως ακολούθως:

$$|\mathcal{X}| = \sum_{x \in \mathcal{X}} P_{\mathcal{X}}(x) \quad (5.47)$$

χρησιμοποιώντας την νόρμα  $\min$  για την ασαφή τομή των ασαφών συνόλων  $\mathcal{X}$  και  $\mathcal{Y}$ . Με τον τρόπο αυτό:

$$r_P(d, t) = \frac{|S_d \cap CP_{t-1}|}{\sqrt{|S_d| |CP_{t-1}|}} \quad (5.48)$$

Στο πλαίσιο ενός συστήματος ανάκτησης περιεχομένου, όπου οι χρήστες ανακτούν το πολυμεσικό περιεχόμενο που τους ενδιαφέρει με το να δημιουργούν και να εισάγουν συγκεκριμένα αιτήματα και ερωτήματα, η μετρική  $r_P(d, t)$  συνδυάζεται με κάποιες κατάλληλες τιμές κατάταξης, προκειμένου να οδηγηθούμε σε ένα τελικό και προσωποποιημένο - σύμφωνα με το εννοιολογικό πλαίσιο - σκορ κατάταξης για το εκάστοτε έγγραφο. Οι τιμές αυτές είναι ανεξάρτητες προς το εκάστοτε ερώτημα και αντιστοιχούνται στα εκάστοτε αποτελέσματα αναζήτησης.

Το τελικό, προσωποποιημένο - σύμφωνα με το εννοιολογικό πλαίσιο - σκορ κατάταξης  $r(d, t)$  για το έγγραφο  $d$  δίνεται από την εξίσωση:

$$r(d, t) = f(r_P(d, t), r_S(d, t)) \quad (5.49)$$

Η μετρική ομοιότητας  $r_S(d, t)$  είναι δυνατόν να χρησιμοποιηθεί για οποιαδήποτε τεχνική κατάταξης, προκειμένου να κατατάξει ένα έγγραφο  $d$  σύμφωνα με ένα ερώτημα ή αίτημα χρήστη τη δεδομένη χρονική στιγμή  $t$ . Η  $r_S(d, t)$  υπολογίζεται σύμφωνα με τα δεδομένα πιθανά ερωτήματα χρήστη, όπως αυτά περιγράφηκαν στην ενότητα 5.5.2. Για παράδειγμα, στην περίπτωση ερωτημάτων με λέξεις-κλειδιά θα έχουμε:

$$r_S(d, t) = \cos(S_d, Q_t^k) \quad (5.50)$$

Στην περίπτωση της θεματικής πλοήγησης, ο βαθμός με τον οποίο ένα έγγραφο  $d$  κατηγοριοποιείται σε ένα θέμα  $Q_t^b$ , δίνεται από την εξίσωση:

$$r_S(d, t) = G_d(z) \quad (5.51)$$

όπου το  $z = Q_t^b$  είναι το συγκεκριμένο θέμα και  $G_d = G(S_d)$  είναι το αποτέλεσμα της διαδικασίας θεματικής κατηγοριοποίησης. Η τελευταία ακολουθεί τις ίδιες γενικές κατευθυντήριες γραμμές που περιγράφηκαν εκτενώς στο κεφάλαιο 4. Με άλλα λόγια, τόσο η διαδικασία της εξαγωγής προτιμήσεων χρηστών, όσο και η διαδικασία θεματικής κατηγοριοποίησης ακολουθούν και υλοποιούν τον ίδιο αλγόριθμο και μπορούν να οριστούν στο πλαίσιο μίας και μοναδικής συνάρτησης  $G$ :

- Το  $G(S_d) = G_d$  παρέχει το ασαφές σύνολο όλων των θεμάτων που σχετίζονται με το συγκεκριμένο έγγραφο  $d$ ,
- Το  $z$  αντιστοιχεί σε ένα συγκεκριμένο θέμα/στοιχείο του ασαφούς συνόλου, και
- Το  $G_d(z)$  αντιπροσωπεύει το βαθμό στον οποίο το  $z$  ανήκει στο  $G_d$  ή με άλλα λόγια το βαθμό ιδιότητας μέλους του.

Τα έγγραφα κατατάσσονται σύμφωνα με την ομοιότητά τους προς το προκαθορισμένο θέμα της αναζήτησης, ενώ στην περίπτωση ενός μοναδικού ερωτήματος επισκόπησης παρουσιάζεται απλά το σχετικό έγγραφο  $d$  στον χρήστη.

Στη γενική περίπτωση, μπορούμε να χρησιμοποιήσουμε το παραπάνω σκορ (5.49), προκειμένου να εισάγουμε ένα προσωποποιημένο bias σε οποιαδήποτε τεχνική κατάταξης, η οποία υπολογίζει το  $r_S(d, t)$  και η οποία θα μπορούσε να βασίζεται σε εικόνες, οντολογίες, συσχετιστική ανάδραση, κ.α.. Η συνδυαστική συνάρτηση  $f$  μπορεί να οριστεί για παράδειγμα ως ένας γραμμικός συνδυασμός  $f(x, y) = \lambda \cdot \bar{x} + (1 - \lambda) \cdot \bar{y}$ . Ο όρος  $\lambda$  αποτελεί τον παράγοντα προσωποποίησης (personalization factor), ο οποίος θα καθορίσει τον βαθμό της προσωποποίησης που θα εφαρμοστεί κατά την κατάταξη των αποτελεσμάτων της αναζήτησης και ο οποίος κυμαίνεται μεταξύ  $\lambda = 0$  (οπότε δεν εφαρμόζεται καμία προσωποποίηση αποτελεσμάτων) και  $\lambda = 1$  (οπότε αγνοείται το ερώτημα και τα αποτελέσματα της αναζήτησης κατατάσσονται μόνο με βάση τα γενικά ενδιαφέροντα των χρηστών). Ο γενικός κανόνας που ακολουθείται είναι ότι το  $\lambda$  θα πρέπει να μειώνεται σύμφωνα με τον βαθμό της αβεβαιότητας για τις προτιμήσεις των χρηστών και να αυξάνει σύμφωνα με τον βαθμό της αβεβαιότητας που υπάρχει στο ερώτημα. Το πρόβλημα του δυναμικού ορισμού της τιμής του  $\lambda$  δεν αποτελεί μέρος αυτής της ερευνητικής εργασίας και ο αναγνώστης καλείται να συμβουλευτεί τη σχετική βιβλιογραφία [55]. Τέλος, τα  $\bar{x}$  και  $\bar{y}$  καταδεικνύουν τις κανονικοποιημένες τιμές του σκορ  $x$  και  $y$ , οι οποίες χρειάζονται, πριν πραγματοποιηθεί ο συνδυασμός τους, έτσι ώστε να εξασφαλίσουμε ότι αυτές θα κυμαίνονται στην ίδια κλίμακα. Η τελική τιμή του  $r(d, t)$  καθορίζει τη θέση κάθε εγγράφου  $d$  στην τελική κατάταξη, καθώς παρουσιάζονται τα προσωποποιημένα αποτελέσματα της αναζήτησης στον εκάστοτε τελικό χρήστη.

## 5.7 Πειραματικές Μετρήσεις

Με στόχο να αποκτήσουμε μια πιο σφαιρική εικόνα για τις παραπάνω τεχνικές εφαρμογής του εννοιολογικού πλαισίου στη διαδικασία της προσωποποίησης του πολυμεσικού περιεχομένου, υλοποιήσαμε ένα πειραματικό πρωτότυπο με ένα μεσαίου μεγέθους σώμα δεδομένων. Βέβαια, η αξιολόγηση της προσωποποίησης είναι, εν γένει, ένα δύσκολο και υπολογιστικά ακριβό έργο [204], [279]. Προκειμένου να μετρήσουμε την απόδοση ενός συστήματος ανάκτησης πληροφορίας, τόσο χωρίς, όσο και με τη χρήση

των προτεινόμενων τεχνικών προσωποποίησης είναι απαραίτητο να συγχρίνουμε την επίδοση της ανάκτησης i) χωρίς προσωποποίηση, ii) με απλή προσωποποίηση και iii) με χρήση προσωποποίησης με βάση το εννοιολογικό πλαίσιο. Οι τυπικές μετρικές αξιολόγησης από το ερευνητικό πεδίο της Ανάκτησης Πληροφορίας (IR) απαιτούν την ύπαρξη χειρωνακτικών διαβαθμίσεων περιεχομένου σύμφωνα με i) τη σχετικότητα του ερωτήματος, ii) τη σχετικότητα του ερωτήματος και τις γενικές προτιμήσεις χρήστη και iii) τη σχετικότητα ερωτήματος και τις ειδικές προτιμήσεις χρήστη (δηλαδή αυτές που είναι περιοριμένες από το εννοιολογικό πλαίσιο της εκάστοτε εργασίας του).

Για το σκοπό αυτό πραγματοποιήσαμε δύο σύνολα πειραμάτων που βασίζονται στον ίδιο χώρο αναζήτησης, ο οποίος αποτελείται από 145,316 έγγραφα (445MB) από το δικτυακό τόπο του CNN<sup>6</sup> και την οντολογία θεματικής περιοχής και βάση γνώσης KIM [138]. Η τελευταία είναι διαθέσιμη ως μέρος της πλατφόρμας KIM που έχει αναπτυχθεί από το εργαστήριο Ontotext Lab<sup>7</sup>. Η εν λόγω βάση γνώσης περιέχει ένα σύνολο από 281 RDF [140] κατηγορίες, 138 ιδιότητες, 35,689 στιγμιότυπα και 465,848 προτάσεις. Τα έγγραφα του CNN συνοδεύονται από πολυμεσικούς σχολιασμούς με έννοιες από τη βάση γνώσης, το μέγεθος των οποίων ξεπερνάει τα τρία εκατομμύρια. Οι τιμές των ασαφών σχέσεων καθορίστηκαν αρχικά με βάση μια διαισθητική προσέγγιση και στη συνέχεια ρυθμίστηκαν εκ νέου εμπειρικά μετά και από την εκτέλεση των πρώτων δοκιμών. Για τα πειράματα αυτά αξιοποιήθηκε ένα σύστημα ανάκτησης πληροφορίας, γνωστό και ως σημασιολογική μηχανή αναζήτησης [54]. Η μηχανή αυτή επιλέχθηκε με κύριο γνώμονα την καλύτερη απόδοσή της σε σχέση με παραδοσιακά - βασισμένα στις λέξεις-κλειδιά - συστήματα, όπως π.χ. είναι το Jakarta Lucene library<sup>8</sup>, όταν είναι διαθέσιμη μιας κάποιας μορφής οντολογική γνώση. Με τον τρόπο αυτό παρέχεται ένα καλύτερο μέτρο σύγκρισης του συστήματος και συνάμα ανυψώνεται σημαντικά και ο πήχυς της αξιολόγησής του. Στο σημείο αυτό θα πρέπει να αναφέρουμε ότι τα πειραματικά αποτελέσματα που παρουσιάζονται στην παρούσα διατριβή μετρούν μόνο την επίδοση της φάσης ανάκτησης, λαμβάνοντας υπόψη κάποιες εκ των προτέρων υπολογισμένες αρχικές προτιμήσεις χρήστη. Όπως θα δούμε παρακάτω, στο πρώτο σετ πειραμάτων οι προτιμήσεις χρήστη προσομοιώνονται κατάλληλα, ενώ στο δεύτερο σετ παρέχονται χειρωνακτικά από πραγματικούς χρήστες.

Δεδομένου ότι οι τεχνικές χρήσης και αξιοποίησης του εννοιολογικού πλαισίου εφαρμόζονται κατά τη διάρκεια μιας συνόδου ενός χρήστη, ένας τρόπος να αξιολογηθούν είναι να καθοριστεί μια σειρά από βήματα, όπου οι τεχνικές αυτές παρεμβαίνουν στη διαδικασία. Αυτή είναι και η προσέγγιση που ακολουθείται στο πρώτο σύνολο πειραμάτων, για το οποίο έχουμε κατασκευάσει ένα δοκιμαστικό σενάριο (testbed) που αποτελείται από ένα λεπτομερές, σταθερό σύνολο υποθετικών καταστάσεων χρήσης εννοιολογικού πλαισίου. Το δοκιμαστικό σενάριο περιλαμβάνει δέκα μικρής έκτασης περιπτώσεις χρήσης, συμπεριλαμβανομένων αυτών που παρουσιάζονται αναλυτικά στις επόμενες ενότητες 5.8 και 5.8.2. Κάθε σενάριο αποτελείται από μια ακολουθία ενεργειών χρηστών που καθορίζονται εκ των προτέρων, συμπεριλαμβανομένων των ερωτημάτων και των επικροτήσεων (clicks) στα αποτελέσματα αναζήτησης. Κατά τον υπολογισμό μεγεθών, όπως η ακρίβεια (precision) και η ανάκληση (recall), αυτή η προσέγγιση δε λαμβάνει με εύκολο τρόπο λεπτομερείς αξιολογήσεις χρηστών (αλήθεια εδάφους - ground truth), λόγω της προσπάθειας και της δυσκολίας που περιλαμβάνεται στην αξιολόγηση των αποτελέσμάτων κάτω από ένα μεγάλο σύνολο τεχνητών,

<sup>6</sup>[http://dmoz.org/News/Online\\_Archives/CNN.com](http://dmoz.org/News/Online_Archives/CNN.com)

<sup>7</sup><http://www.ontotext.com/kim/index.html>

<sup>8</sup><http://lucene.apache.org>

## Κεφάλαιο 5. Προσωποποιημένη πρόσβαση

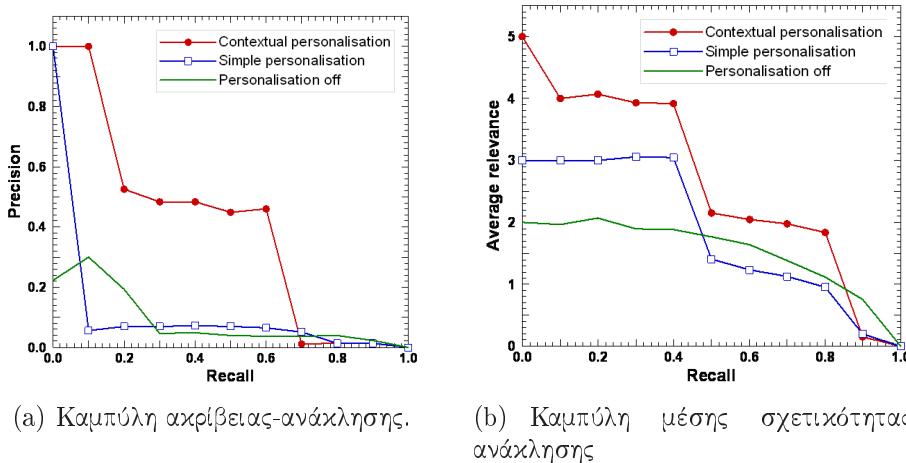
σύνθετων και απαιτητικών υποθέσεων. Επομένως, έχουμε ακολουθήσει μια χειρωνακτική κατάταξη των εγγράφων/ερωτημάτων/προτιμήσεων και του εννοιολογικού πλαισίου με βάση τους υποθετικούς χρήστες, για τους οποίους προσομοιώνονται τα προφίλ. Αν και υποκειμενική, αυτή η προσέγγιση επιτρέπει την εξαγωγή σημαντικών παρατηρήσεων, καθώς και την αποτελεσματική δοκιμή των δυνατοτήτων υλοποίησης και αξιοπιστίας των καθορισμένων προτύπων και αλγορίθμων. Τα αποτελέσματα αυτά συμπληρώνονται κατόπιν με μια πιο αντικειμενική αξιολόγηση με πραγματικούς χρήστες, η οποία και ακολουθεί.

Ένα πρώτο σενάριο δειγμάτων ακολουθεί την εξής πορεία: 'Εστω ότι στον Αλέξανδρο αρέσουν όλα τα άρθρα σχετικά με πολυτελή και μοντέρνα αντικείμενα. Οι προτιμήσεις του περιλαμβάνουν, μεταξύ άλλων, φανταχτερά εμπορικά σήματα, όπως Rolex, Maybach, Lexus, Hilton, Aston Martin, Bentley, Louis Vuitton, Sony, Apple, Rolls-Royce, Mercedes, Ferrari, Prada, και BMW. Ο Αλέξανδρος αρχίζει μια σύνοδο αναζήτησης με μια ερώτηση για τις ειδήσεις σχετικά με την Daimler-Chrysler και τα διαφορετικά εμπορικά σήματα που ανήκουν στην εταιρία. Στην Daimler-Chrysler ανήκουν πολυτελή εμπορικά σήματα, όπως η Mercedes ή η Maybach, καθώς και άλλα πιο συνηθισμένα, όπως η Dodge ή η Setra, τα οποία και δεν ενδιαφέρουν τον Αλέξανδρο.

Ενώ το σύστημα ανάκτησης δεν ταξινομεί κάποια εμπορικά σήματα πολυτέλειας υψηλότερα από άλλα, η προσωποποίηση αναδιατάσσει τα αποτελέσματα σύμφωνα με τις προτιμήσεις του Αλεξάνδρου. Συνακόλουθα, παρουσιάζει πρώτα τα έγγραφα σχετικά με την Daimler-Chrysler και τα πολυτελή εμπορικά σήματα των Mercedes ή Maybach και ωθεί προς τα κάτω άλλα έγγραφα σχετικά με τα υπόλοιπα μη-πολυτελή εμπορικά σήματα της επιχείρησης. Κατά συνέπεια, η εξατομικευμένη αναζήτηση αποδίδει καλύτερα από την άποψη του χρήστη. Δεδομένου ότι αυτό ήταν το πρώτο ερώτημα της συνόδου, δεν υπάρχει ακόμα κανένα εννοιολογικό πλαίσιο, οπότε οι προτιμήσεις των χρηστών δε φιλτράρονται και δεν υπάρχει καμία μετρήσιμη απόδοση της επίδρασης του εννοιολογικού πλαισίου.

Στη συνέχεια, ο Αλέξανδρος ανοίγει μερικά έγγραφα από τα αποτελέσματα της αναζήτησης, τα οποία ασχολούνται με το εμπορικό σήμα της Mercedes και το πώς η Daimler-Chrysler πρόκειται να "ρίξει" στην αγορά ένα νέο μοντέλο. Ανοίγει επίσης μια πολυμεσική παρουσίαση σχετικά με το νέο μοντέλο "Maybach 62". Το όργανο ελέγχου του εννοιολογικού πλαισίου (context monitor) εξάγει την έννοια της Mercedes από τα σχολιασμένα έγγραφα και τις εικόνες που προσπέλασε ο χρήστης μαζί με την έννοια Maybach, δεδομένου ότι το επιλεγμένο περιεχόμενο ήταν ως επί το πλείστον σχετικό με αυτά τα δύο εμπορικά σήματα. Το εννοιολογικό πλαίσιο επεκτείνεται σε νέες έννοιες, όπως Daimler-Chrysler, ιδιοκτήτης Mercedes και Maybach μαζί με όλα τα εμπορικά σήματά του. Έπειτα, ο Αλέξανδρος κάνει μια νέα ερώτηση: "Επιχειρήσεις που δραστηριοποιούνται στο Χρηματιστήριο Αξιών της Νέας Υόρκης και έχουν καταθέσει εμπορικά τους σήματα στις Η.Π.Α.". Τα αποτελέσματα της ερώτησης επαναταξινομούνται σύμφωνα με τις βασισμένες στο εννοιολογικό πλαίσιο προτιμήσεις του Αλεξάνδρου. Τα έγγραφα που αναφέρουν την Daimler-Chrysler και τη Mercedes προωθούνται προς τα επάνω στο σύνολο των αποτελεσμάτων. Η προσωπική σχετικότητα αυξάνει και στα έγγραφα που σχολιάζονται με την έννοια Maybach, το έτερο αγαπημένο από το χρήστη εμπορικό σήμα της Daimler-Chrysler. Ο Αλέξανδρος βρίσκει κι άλλες επιχειρήσεις και εμπορικά σήματα που συναλλάσσονται στο Χρηματιστήριο Αξιών της Νέας Υόρκης και ταιριάζουν με τις προτιμήσεις του, όπως είναι π.χ. η εταιρία Sony, αλλά αυτές δε βρίσκονται σημασιολογικά κοντά στα εμπορικά σήματα στο εννοιολογικό πλαίσιο και επομένως λαμβάνουν μια χαμηλότερη

ταξινόμηση σε σχέση με το υπόλοιπο πολυμεσικό περιεχόμενο που ταφιάζει περισσότερο με το πλαίσιο των προηγούμενων ενεργειών του χρήστη. Το τελευταίο ταφιάζει με τα πραγματικά τρέχοντα (υπονοούμενα) ενδιαφέροντα του χρήστη, κάτι το οποίο εξηγεί τη βελτίωση που παρουσιάζεται στο Σχήμα 5.16.

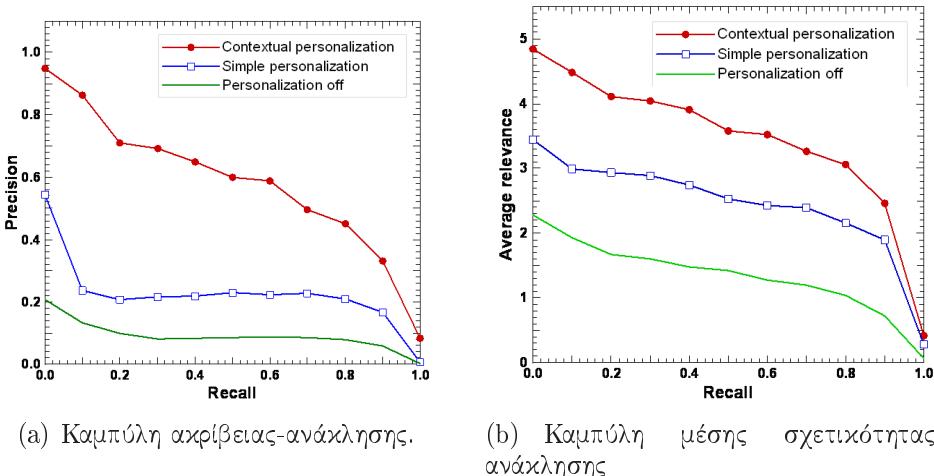


**Σχήμα 5.16:** Συγχριτική απόδοση της προσωποποιημένης αναζήτησης με και χωρίς επίδραση του εννοιολογικού πλαισίου για το ερώτημα: “επιχειρήσεις που είναι εισηγμένες στο Χρηματιστήριο Αξιών της Νέας Υόρκης και έχουν ένα εμπορικό σήμα στις Η.Π.Α.”. Το γράφημα παρουσιάζει i) την καμπύλη ακρίβειας-ανάκλησης και ii) την καμπύλη μέσης σχετικότητας-ανάκλησης.

Το πείραμα αυτό αποτελεί ένα ξεκάθαρο παράδειγμα, όπου η προσωποποίηση από μόνη της δε θα έδινε τα καλύτερα αποτελέσματα, ή θα απέδιδε ακόμη και χειρότερα από την μη-προσαρμοστική ανάκτηση (βλ. την πτώση της ακρίβειας για ανάκληση μεταξύ 0.1 και 0.3 στο Σχήμα 5.16(a)), επειδή οι άσχετες, μακροπρόθεσμες προτιμήσεις (όπως στο παράδειγμα τα αγαπημένα εμπορικά σήματα πολυτελείας και οι επιχειρήσεις του χρήστη που δε σχετίζονται με την τρέχουσα εστίασή του στο εννοιολογικό πλαίσιο της αυτοκινητοβιομηχανίας) θα εμπόδιζαν τον χρήστη. Το πείραμα επιδεικνύει πώς η προσέγγισή μας μπορεί να αποφύγει αυτήν την επίδραση και να ενισχύσει σημαντικά την προσωποποίηση, αφενός με την αφαίρεση τέτοιων ενδιαφερόντων χρηστών που βρίσκονται εκτός του εννοιολογικού πλαισίου και αφετέρου με το να αφήσει αυτά που είναι πράγματι σχετικά με το τρέχον σχέδιο δράσης.

Η τεχνική της επίδρασης του εννοιολογικού πλαισίου οδηγεί με συνέπεια σε καλύτερη απόδοση, όσον αφορά την απλή προσωποποίηση, όπως μπορεί να παρατηρηθεί στο Σχήμα 5.17, το οποίο παρουσιάζει το μέσο όρο των αποτελεσμάτων για δέκα σενάρια χρήσης. Οι ελάχιστες περιπτώσεις, όπου η τεχνική μας φαίνεται να αποδίδει χειρότερα οφέλονταν στην έλλειψη κατάλληλων πληροφοριών στη Βάση Γνώσης (KB), συνέπεια του οποίου ήταν το γεγονός ότι το σύστημα δεν μπόρεσε να διαπιστώσει ότι ορισμένες προτιμήσεις χρηστών αφορούσαν πράγματι το συγκεκριμένο εννοιολογικό πλαίσιο. Ένας άλλος περιορισμός της προσέγγισής μας είναι ότι υποθέτει πως οι διαδοχικές ερωτήσεις χρηστών τείνουν να συσχετίζονται, κάτι το οποίο δε φαίνεται να ισχύει κατά τη διάρκεια αιφνίδιων αλλαγών του ενδιαφέροντος του χρήστη. Εντούτοις, κατά μέσο όρο όχι μόνο οι γενικές βελτιώσεις της προσέγγισης είναι πολλαπλές, αλλά και η πιθανή ελάττωση της απόδοσης σε τέτοιες περιπτώσεις εξαφανίζεται μετά από δύο ή τρία ερωτήματα, δεδομένου ότι το βάρος των βασισμένων στο εννοιολογικό πλαίσιο εννοιών μειώνεται εκθετικά, καθώς ο χρήστης συνεχίζει να αλληλεπιδρά με το σύστημα.

## Κεφάλαιο 5. Προσωποποιημένη πρόσβαση

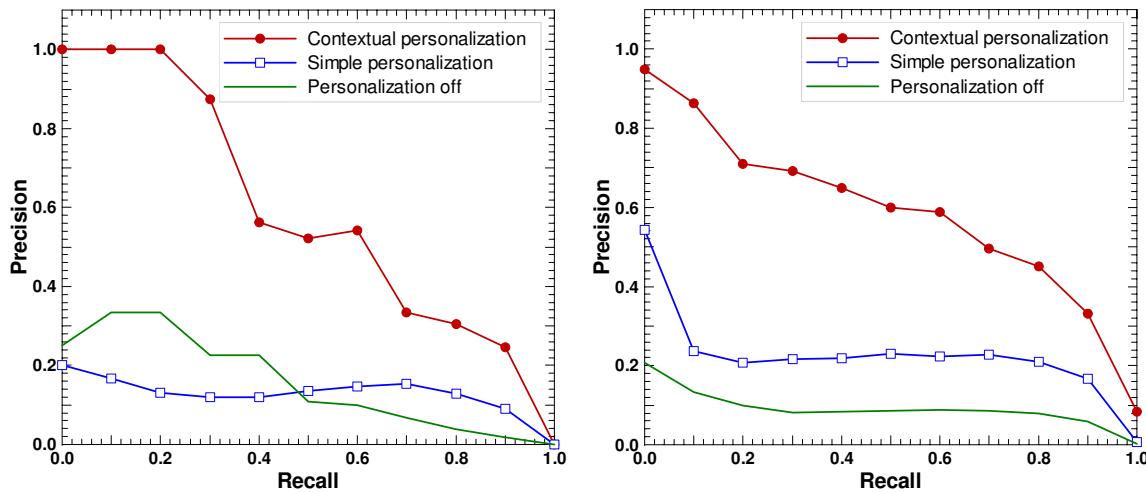


**Σχήμα 5.17:** Συγκριτική απόδοση της προσωποποιημένης αναζήτησης με και χωρίς την επίδραση του εννοιολογικού πλαισίου μέσος όρος δέκα περιπτώσεων χρήσης.

Κατά κάποιο τρόπο το μοντέλο των βασισμένων στο εννοιολογικό πλαίσιο προτιμήσεων χρηστών μπορεί να αντιμετωπισθεί ως μια προσέγγιση στα βραχυπρόθεσμα, “ζωντανά” ενδιαφέροντα των χρηστών, σε αντιδιαστολή με ολόκληρο το σύνολο των προτιμήσεων, το οποίο θα αντιπροσώπευε τις μακροπρόθεσμες προτιμήσεις. Εντούτοις, το μοντέλο μας δεν αποτυπώνει ρητά τα περιστασιακά βραχυπρόθεσμα ενδιαφέροντα υπό αυτήν τη μορφή, εκτός κι αν αποθηκεύονται διαρκώς στο προφίλ του χρήστη. Δεδομένου ότι είναι αρκετά δύσκολο να διακριθεί ένα περιστασιακό, “ζωντανό” ενδιαφέρον χρήστη από μια απλώς και μόνο βασισμένη στο εννοιολογικό πλαίσιο έννοια, περιλαμβάνουμε τα πρώτα μέσα στα δεύτερα, με έναν τρόπο, έτσι ώστε στην πράξη οι βραχυπρόθεσμες προτιμήσεις χρηστών μπορούν να επηρεάσουν τις απαντήσεις του συστήματος. Επιπλέον, αν το υπονοούμενο “ζωντανό” ενδιαφέρον είναι τελείως ανεξάρτητο από τις σταθερές (επίμονες) προτιμήσεις, τότε ο αντίκτυπός του θα είναι ελάχιστος ή ακόμα και μηδενικός. Αυτό αποτελεί σαφώς ένα ανοικτό ερευνητικό πρόβλημα που ζεψεύγει από τα όρια της παρούσης διαδικτορικής διατριβής.

Στη συνέχεια, στο Σχήμα 5.18a) παρουσιάζονται τα αποτελέσματα αυτής της παρούσης πειραματικής προσέγγισης για την περίπτωση του σεναρίου χρήσης που παρουσιάζεται στην ενότητα 5.8. Πρόκειται και πάλι για ένα ξεκάθαρο παράδειγμα, όπου η προσωποποίηση από μόνη της δε θα μας έδινε καλύτερα αποτελέσματα ή θα μας έδινε ακόμα και χειρότερα αποτελέσματα σε σχέση με την μη-προσαρμοζόμενη ανάκτηση (βλ. την πτώση της ακρίβειας για τιμές ανάκλησης μεταξύ 0.1 και 0.4 στο Σχήμα 5.18a). Κάτι τέτοιο δικαιολογείται από το γεγονός ότι οι άσχετες μακροπρόθεσμες προτιμήσεις (όπως οι εταιρίες πληροφορικής, στο παράδειγμά μας, που δε σχετίζονται με την τρέχουσα εστίαση του χρήστη στις Ιαπωνικές εταιρίες) θα εμπόδιζαν αισθητά τον χρήστη. Το δεύτερο αυτό πείραμα επιδεικνύει επίσης τον τρόπο με τον οποίο η προτεινόμενη διαδικασία της προσωποποίησης μπορεί να αποφύγει την επίδραση αυτή και να ενισχύσει σημαντικά την προσωποποίηση των αποτελεσμάτων.

Μπορούμε, επίσης, να παρατηρήσουμε στο Σχήμα 5.18b ότι η τεχνική της προσωποποίησης με βάση το εννοιολογικό πλαίσιο καταλήγει σταθερά σε καλύτερη επίδοση σε σχέση με την απλή προσωποποίηση. Το τελευταίο καταδεικνύει το μέσο όρο των αποτελεσμάτων για τα 10 σενάρια και το Σχήμα 5.19 δείχνει το μέσο ιστόγραμμα ακρίβειας συγχρίνοντας την χωρίς και με εννοιολογικό πλαίσιο προσωποποίηση κατά το χρόνο εκτέλεσης.



**Σχήμα 5.18:** Συγκριτική επίδοση της προσωποποιημένης αναζήτησης με και χωρίς την επίδραση του εννοιολογικού πλαισίου. καταδεικνύεται η καμπύλη ακρίβειας-ανάκλησης για i) ένα από τα σενάρια και ii) το μέσο όρο των 10 σεναρίων. Τα αποτελέσματα στο γράφημα i) αντιστοιχούν στο ερώτημα “Εταιρίες που έχουν τη βάση τους σε οποιαδήποτε Ιαπωνική περιοχή”, σύμφωνα με το σενάριο χρήσης που περιγράφεται στην ενότητα 5.8.

Στην επόμενη πειραματική προσέγγιση ανατίθενται σε πραγματικούς ανθρώπους τρεις διαφορετικοί στόχοι ανάκτησης. Κάθε ένας από τους στόχους εκφράζει μια συγκεκριμένη ανάγκη πληροφόρησης, έτσι ώστε οι χρήστες καλούνται να αναζητήσουν όσο το δυνατόν περισσότερα έγγραφα που να εκπληρώνουν όλες τις δεδομένες ανάγκες. Σε αυτό το πείραμα η ακολουθία των ενεργειών δεν είναι προκαθορισμένη, όπως στην προηγούμενη περίπτωση, αλλά καθορίζεται με πλήρη ελευθερία από τους χρήστες, καθώς οι τελευταίοι επιδιώκουν να επιτύχουν τους προτεινόμενους στόχους.

Ένα σύνολο από 18 άτομα επιλέχθηκαν για την εκτέλεση του πειράματος, ενώ ορίστηκαν και τρία είδη σχετικών αναζητήσεων<sup>9</sup> ως ακολούθως:

1. Ειδήσεις σχετικά με εταιρικές συμφωνίες.
2. Παρουσιάσεις νέων ηλεκτρονικών προϊόντων.
3. Πληροφορίες σχετικά με πόλεις που διοργανώνουν κάποιο γεγονός μηχανοκίνητου αθλητισμού.

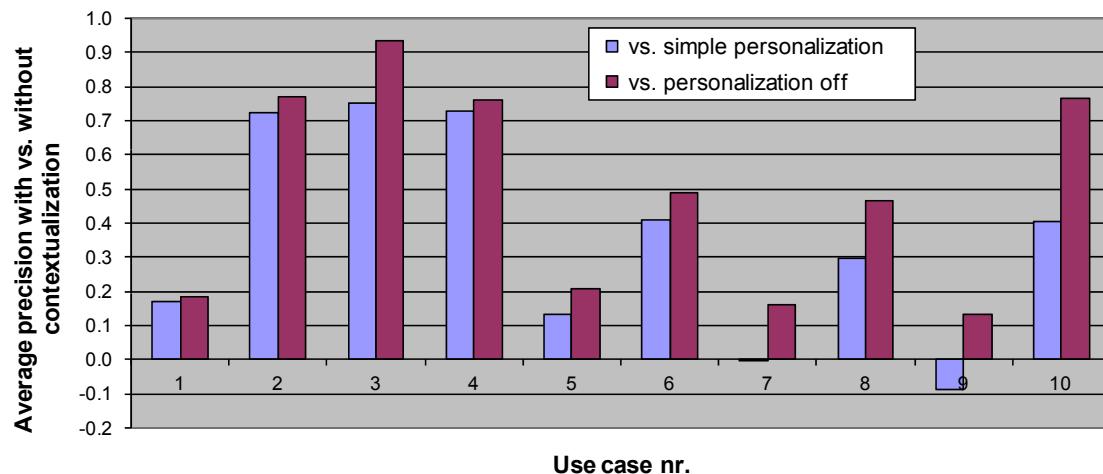
Κάθε είδος αναζήτησης εκτελέσθηκε i) με προσωποποίηση και με βάση το εννοιολογικό πλαίσιο, ii) με απλή προσωποποίηση και iii) χωρίς προσωποποίηση. Προκειμένου να μην επαναληφθούν αναζητήσεις από τους χρήστες, κάθε ένα από τα τρία είδη χρησιμοποιήθηκε από 6 χρήστες ( $3 \text{ είδη} \times 6 \text{ χρήστες} = 18 \text{ τεστ}$  για κάθε είδος), με ένα τέτοιο τρόπο έτσι ώστε κάθε χρήστης να δοκιμάσει και τα τρία είδη (i), (ii) και (iii), ακριβώς μία φορά, ακολουθώντας μια πειραματική σχεδίαση λατινικού τετραγώνου<sup>10</sup>. Με τον τρόπο αυτό κάθε είδος τεστάρεται ακριβώς 18 φορές: μία φορά για

<sup>9</sup> Στην πράξη, στους χρήστες δόθηκε μια αναλυτικότερη και εκτενέστερη περιγραφή του θέματος, έτσι ώστε αυτό να οριστεί με όσο το δυνατόν μεγαλύτερη ακρίβεια και να αποφευχθούν αμφισημίες.

<sup>10</sup> Ένα λατινικό τετράγωνο είναι ένας πίνακας  $n \times n$  ο οποίος αποτελείται από  $n$  διαφορετικά σύμβολα κατά τέτοιο τρόπο, ώστε κάθε σύμβολο να εμφανίζεται ακριβώς μία φορά σε κάθε σειρά

και ακριβώς μιά φορά σε κάθε στήλη. Π.χ.:  $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}$

## Κεφάλαιο 5. Προσωποποιημένη πρόσβαση



**Σχήμα 5.19:** Συγκριτικό ιστόγραμμα ακρίβειας της εξατομικευμένης αναζήτησης με και χωρίς εννοιολογικό πλαίσιο των δέκα περιπτώσεων χρήσης. Οι ελαφρώς χρωματισμένες μπάρες συγχρίνουν την εξατομικευμένη ανάκτηση με εννοιολογικό πλαίσιο και την απλή εξατομικευμένη ανάκτηση χωρίς εννοιολογικό πλαίσιο. Οι σκούρες μπάρες συγχρίνουν την εξατομικευμένη ανάκτηση με εννοιολογικό πλαίσιο και την ανάκτηση χωρίς καθόλου προσωποποίηση.

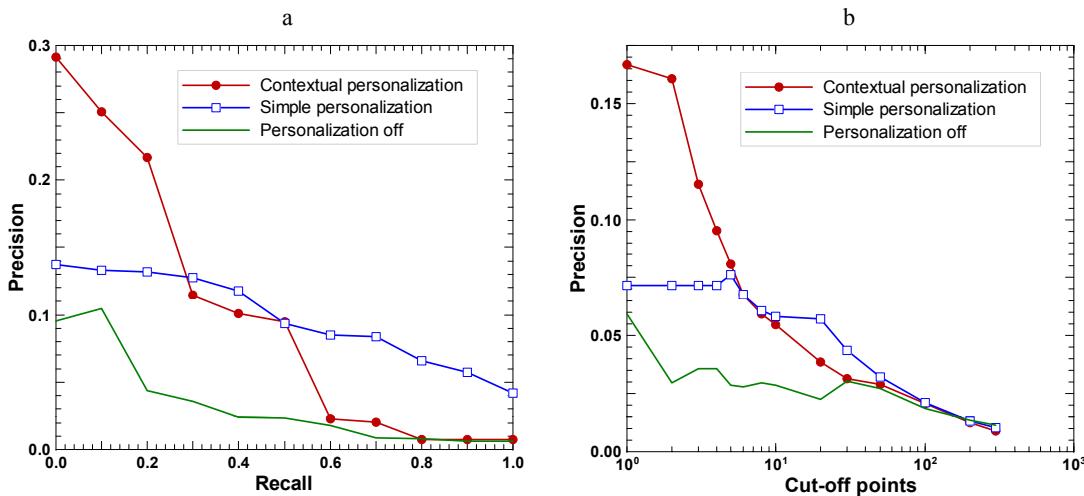
κάθε χρήστη και 6 φορές για κάθε έργο, έτσι ώστε κανένα είδος να μην αδικείται ή να μην επωφελείται από τη δυσκολία της κάθε αναζήτησης ή τις ικανότητες του εκάστοτε χρήστη. Οι προτιμήσεις των χρηστών λαμβάνονται χειρωνακτικά από τους χρήστες, ζητώντας τους να κατατάξουν/καθορίσουν ρητά μια προκαθορισμένη λίστα από έννοιες της θεματικής περιοχής στην αρχή κάθε συνόδου.

Τα σχετικά έγγραφα για κάθε αναζήτηση μαρκάρονται εκ των προτέρων από έναν ειδήμονα (εν προκειμένω από τον υποφαινόμενο), έτσι ώστε οι χρήστες να απαλλάσσονται από το να προσδίδουν εκτενείς κρίσεις σχετικότητας. Παρ' όλα αυτά οι χρήστες ενθαρρύνονται να προσπελαύνουν τα έγγραφα που εμφανίζονται να είναι πιο σχετικά σύμφωνα με τα υποκειμενικά τους κριτήρια, έτσι ώστε να προσφέρουν στο σύστημα περισσότερους πόλους εννοιολογικού πλαισίου. Η πληροφορία εννοιολογικού πλαισίου συλλέγεται με βάση τις έννοιες που βρίσκονται στο χρακτηρισμό (annotation), τα επιλεγμένα αποτελέσματα καθώς και τις έννοιες που σχετίζονται με τις λέξεις-κλειδιά στα ερωτήματα των χρηστών (χρησιμοποιώντας την αντιστοίχηση λέξεων-κλειδιών και εννοιών που παρέχεται στη Βάση Γνώσης KIM).

Στο τέλος κάθε δοκιμασίας αναζήτησης το σύστημα ζητά από τον χρήστη να μαρκάρει τα έγγραφα μεταξύ των τελικών αποτελεσμάτων ως **σχετικά** (related) ή **άσχετα** (unrelated), τόσο σε σχέση με τα προσωπικά του ενδιαφέροντα, όσο και με τη δοκιμασία της αναζήτησης. Για τον υπολογισμό της ακρίβειας και της ανάκλησης πραγματοποιούνται δύο απλοποιήσεις στη θεώρηση κάθε διαδραστικής ακολουθίας (δηλαδή για κάθε δοκιμασία αναζήτησης και χρήστη):

- Ο χώρος αναζήτησης απλοποιείται, έτσι ώστε να αποτελεί το σύνολο όλων των εγγράφων, που έχουν επιστραφεί από το σύστημα σε κάποιο σημείο στην επαναληπτική διαδικασία ανάκτησης για τη δοκιμασία που κατευθύνεται από τον συγκεκριμένο χρήστη.
- Το σύνολο των σχετικών εγγράφων θεωρείται ότι είναι η διατομή (intersection) των εγγράφων του χώρου αναζήτησης που χαρακτηρίζονται ως σχετικά κατά

την κρίση του ειδήμονα και αυτών που χαρακτηρίζονται ως σχετικά από το χρήστη σύμφωνα με τα ιδιαίτερα ενδιαφέροντά του.



**Σχήμα 5.20:** Συγκριτική επίδοση της προσωποποιημένης αναζήτησης με και χωρίς τη χρήση εννοιολογικού πλαισίου από 18 άτομα σε τρεις προτεινόμενους στόχους. Η γραφική παράσταση παρουσιάζει α) την καμπύλη ακρίβειας-ανάκλησης και β) την ακρίβεια στα σημεία διακοπών (cut-off points). Λαμβάνεται η μέση τιμή των αποτελεσμάτων από το σύνολο όλων των χρηστών και όλων των στόχων.

Το Σχήμα 5.20 καταδεικνύει τα αποτελέσματα που λαμβάνονται από την παραπάνω μεθοδολογία. Η καμπύλη στα αριστερά του Σχήματος δείχνει μια εμφανή βελτίωση της προσέγγισης με βάση το εννοιολογικό πλαίσιο στα υψηλά επίπεδα ακρίβειας, τόσο σε σχέση με την απλή προσωποποίηση, όσο και σε σχέση με την περίπτωση της καθόλου προσωποποίησης. Ωστόσο, παρατηρούμε ότι η βελτίωση αυτή μειώνεται στα υψηλά επίπεδα ανάκλησης. Παρόμοια βελτίωση παρατηρείται και στην περίπτωση της καμπύλης ακρίβειας στα σημεία διακοπών, ιδιαίτερα, όμως, στα κορυφαία 10 αποτελέσματα. Η απλή προσωποποίηση επιτυγχάνει σημαντικότερες τιμές ακρίβειας στα κορυφαία έγγραφα, καταδεικνύοντας το γεγονός ότι η προτεινόμενη τεχνική αποφεύγει ένα σημαντικό αριθμό από λανθασμένα θετικά (false positives), τα οποία μπορεί να εμφανιστούν, όταν οι προτιμήσεις χρηστών δε λαμβάνουν υπόψη τους το εννοιολογικό πλαίσιο. Οι μέσες τιμές ακρίβειας (MAP), που παρουσιάζονται στον Πίνακα 5.7 για την περίπτωση της με βάση το εννοιολογικό πλαίσιο της απλής και της καθόλου προσωποποίησης, επιβεβαιώνουν ότι η τεχνική που προτείνουμε τα καταφέρνει αρκετά καλύτερα.

**Πίνακας 5.6:** Μέση ακρίβεια (MAP) για κάθε ένα από τα τρία είδη αναζήτησης

Είδος αναζήτησης	MAP
Προσωποποίηση με βάση το Εννοιολογικό Πλαίσιο	0.1353
Απλή Προσωποποίηση	0.1061
Χωρίς Προσωποποίηση	0.0463

Θα πρέπει να σημειώσουμε και πάλι ότι αποτελεί εν γένει περιορισμό της προσέγγισής μας το γεγονός ότι υποθέτει πως τα διαδοχικά ερωτήματα χρηστών τείνουν να είναι συσχετισμένα μεταξύ τους, κάτι το οποίο δεν ισχύει σε περιπτώσεις ξαφνικής

αλλαγής του ενδιαφέροντος του χρήστη. 'Όμως, η προτεινόμενη τεχνική εμφανίζει κατά μέσο όρο μια ιδιαίτερα αξιόλογη βελτίωση και μια πιθανή μείωση της απόδοσής της σε τέτοιες ακραίες περιπτώσεις εξαφανίζεται μετά από δύο ή τρία ερωτήματα, δεδομένου ότι ο βαθμός των βασισμένων στο εννοιολογικό πλαίσιο εννοιών μειώνεται εκθετικά, καθώς ο χρήστης συνεχίζει να αλληλεπιδρά με το σύστημα, όπως άλλωστε εξηγείται αναλυτικά στην υποενότητα 5.6.1.

## 5.8 Σενάρια Χρήσης

### 5.8.1 Πρώτο σενάριο χρήσης

Ας θεωρήσουμε, τώρα, το ακόλουθο σενάριο χρήσης της παραπάνω μεθοδολογίας ως μια πρώτη απόπειρα ολοκληρωμένης παρουσίασης της εφαρμογής της στην πράξη. 'Εστω, λοιπόν, ότι η 'Ελλη έχει εγγραφεί σε ένα ειδησεογραφικό πρακτορείο οικονομικών ειδήσεων. Η 'Ελλη εργάζεται σε μία μεγάλη εταιρία τροφίμων, οπότε, όπως είναι φυσικό, προτιμάει τις ειδήσεις που σχετίζονται με αυτόν τον τομέα. Από την άλλη μεριά, προσπαθεί να είναι επίσης ενημερωμένη για όλες τις τελευταίες τεχνολογικές εξελίξεις του χώρου, καθώς τόσο η εταιρία της, όσο και η ίδια ενδιαφέρονται να χρησιμοποιήσουν την τελευταία λέξη της τεχνολογίας, για να βελτιώσουν την αλυσίδα παραγωγής τροφίμων της εταιρίας και να παραμείνουν ανταγωνιστικοί στην αγορά.

Η 'Ελλη σχεδιάζει να ταξιδέψει στο Τόκιο και στο Κιότο της Ιαπωνίας. Ο στόχος της είναι να συλλέξει ιδέες και τεχνογνωσία από αντίστοιχες προς τη δική της εταιρίες της 'Απω Ανατολής. Ψάχνοντας πληροφορίες για διάφορες Ιαπωνικές εταιρίες, στρέφεται στο ειδησεογραφικό πρακτορείο που είναι μέλος και ξεκινά μια σύνοδο αναζήτησης (search session). Το σχετικό σενάριο περιλαμβάνει τα έγγραφα, ένα σύνολο από έννοιες  $S$  και τις σχέσεις  $T$  και  $\hat{T}$  πάνω σε αυτό, όπως αυτές ορίστηκαν στην υποενότητα 5.4.1. Το πρώτο βήμα της μεθοδολογίας μας αποτελείται από τον υπολογισμό του συνόλου  $P$  των απογραμμικών (offline) προτιμήσεων χρήστη για την 'Ελλη. Με άλλα λόγια, υπολογίζονται τα βαθμονομημένα σημασιολογικά ενδιαφέροντα της οντολογίας της θεματικής περιοχής που περιλαμβάνονται στο προφίλ της 'Ελλης. Το σύνολο των παραπάνω εννοιών μαζί με τα μνημονικά που χρησιμοποιήσαμε για ευκολία στους υπολογισμούς παρουσιάζονται στον Πίνακα 5.7 και στην απεικόνιση της σχέσης  $T$  στο Σχήμα 5.21, ενώ οι τιμές τους περιλαμβάνονται στον Πίνακα 5.8. 'Οπως παρατηρεί κανείς, το σύνολο των εννοιών περιλαμβάνει ποικίλες εταιρίες τροφίμων, ποτών και τσιγάρων, καθώς και αρκετές εταιρίες πληροφορικής. Για λόγους απλότητας στην απεικόνιση παρουσιάζονται μόνο οι σχετικές έννοιες της οντολογίας μαζί με τους σχετικούς βαθμούς συμμετοχής. Τα παραπάνω θα οδηγήσουν τελικά στον ορισμό του ασαφούς συνόλου  $P$  για την 'Ελλη, σύμφωνα με όσα περιγράφονται αναλυτικά στη συνέχεια.

Θα πρέπει να σημειωθεί εδώ ότι τα στοιχεία της σχέσης που εννοούνται λόγω μεταβατικότητας έχουν παραληφθεί από την απεικόνιση για λόγους πολυπλοκότητας του Σχήματος. Επίσης, θεωρούμε το *sup*-γινόμενο για τη μεταβατικότητα και η  $t$ -νόρμα που χρησιμοποιούμε για το μεταβατικό κλείσιμο της σχέσης  $T$  είναι η  $t$ -νόρμα του Yager<sup>11</sup> (βλ. και Παράρτημα B) με παράμετρο  $p = 3$ . Επιπρόσθετα, η *co*-νόρμα που χρησιμοποιείται στην εξίσωση (5.16) είναι το δεσμευμένο άθροισμα, ενώ στην εξίσωση (5.33) χρησιμοποιείται ως  $t$ -νόρμα το γινόμενο. Για την τελική εξαγωγή

<sup>11</sup>  $T_p^Y(x, y) = \max(0, 1 - ((1-x)^p + (1-y)^p)^{1/p})$ , για  $0 < p < +\infty$

**Πίνακας 5.7:** Ονόματα εννοιών και τα αντίστοιχα μνημονικά τους. Οι θεματικές κατηγορίες εμφανίζονται με έντονα μαύρα γράμματα

Έννοια	Μνημονικό	Έννοια	Μνημονικό
<b>Food companies</b>	<b>fcp</b>	Japan Tobacco Inc.	jti
Mc' Donalds	med	Big Mac	bgm
Yamazaki Baking Co.	yam	<b>Technology companies</b>	<b>tcp</b>
<b>Microsoft Corp.</b>	<b>msc</b>	Apple Computers Inc.	apl
Microsoft Office Suite	ofc	Personal Computer	pcm
Macintosh	mac	<b>Linux Community</b>	<b>lnx</b>
Tux	tux	X Windows System	xws
Programming shell	shl	Windows media player	wmp
Microsoft Visio	vso	Windows Mathtype	mtp
Dunkin Donuts	dnk	Coca Cola	cok
Food, Beverage & Tobacco Sector	fbt	Makoto Tajima	mkt
Microsoft	mis	Apple	ape
McDonald's Corp.	mdc	Macintosh G3	mcg

**Πίνακας 5.8:** Μέρος της ταξινομικής σχέσης  $T$ .

$s_1$	$s_2$	$T$	$s_1$	$s_2$	$T$	$s_1$	$s_2$	$T$
msc	wmp	0.70	tcp	apl	0.80	tcp	msc	0.80
fcp	med	0.80	lnx	xws	0.80	med	bgm	1.00
fcp	jti	0.80	apl	mac	0.90	lnx	ofc	0.60
fcp	yam	0.80	apl	pcm	0.80	apl	ofc	0.60
msc	mtp	0.90	lnx	pcm	0.60	lnx	tux	0.90
msc	vso	0.90	msc	ofc	0.60	lnx	shl	0.90
msc	pcm	0.80						

των προτιμήσεων αξιοποιείται η standard co-νόρμα  $max$ . Τέλος, το κατώφλι που χρησιμοποιείται για το κριτήριο τερματισμού του αλγορίθμου συσταδοποίησης είναι το 0.3.

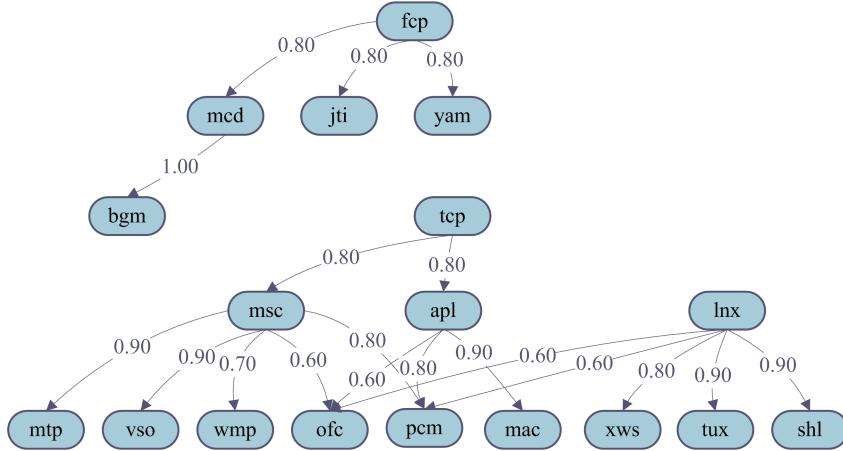
Το σημασιολογικό ευρετήριο αναπαρίσταται ως εξής:

$$I(s_j, d) = \text{pcm}/0.9 + \text{dnk}/0.8 + \text{ofc}/0.9 + \text{mac}/1 + \text{jti}/0.4 \quad (5.52)$$

Η διαδικασία συσταδοποίησης εννοιών καταλήγει σε 3 κλασικές (σαφείς) συστάδες:

$$K' = \{k'_1, k'_2, k'_3\} = \{\text{(pcm, mac, ofc)}, \text{dnk}, \text{jti}\} \quad (5.53)$$

Λόγω της απλότητας του περιεχομένου αυτής της πρώτης συνόδου της Έλλης και του μικρού αριθμού από τις αναγνωρισμένες έννοιες η χρήση του βασισμένου στο εννοιολογικό πλαίσιο ταξινομητή, που παρουσιάσθηκε στην υποενότητα 5.5.4.3, δεν οδηγεί σε μια επέκταση των ανιχνευθέντων κλασικών (σαφών) συστάδων, έτσι ώστε να συμπεριλάβουν και άλλες έννοιες. Κάτι τέτοιο είναι αναμενόμενο σύμφωνα με τη δομή της σχέσης  $T$  στο Σχήμα 5.21, αφού η σημασιολογία όλων των εννοιών υπαινίσσεται είτε μια συνολική, είτε μια τελείως απούσα σχέση. Στη συνέχεια, παρεμβαίνουμε περαιτέρω στους βαθμούς συμμετοχής των παραπάνω συστάδων, χρησιμοποιώντας την  $t$ -νόρμα γινόμενο και σύμφωνα με την εξίσωση (5.33) θα έχουμε:



Σχήμα 5.21: Παράδειγμα κατασκευής της σχέσης  $T$ .

$$\begin{aligned}
 k_1 &= pcm/0.9 + mac/1.0 + ofc/0.9 \\
 k_2 &= dnk/1.0 \\
 k_3 &= jti/1.0
 \end{aligned} \tag{5.54}$$

Κάθε μία από τις παραπάνω συστάδες αντιστοιχεί σε μία από τις διακεχριμένες προτιμήσεις χρήστη που σχετίζονται με την Έλλη και, προκειμένου να εκτιμήσουμε τις τελευταίες, λάβαμε υπόψη μας τόσο την γραμμική πληθυκότητα (scalar cardinality) κάθε συστάδας, όσο και το εννοιολογικό της πλαίσιο. Ειδικότερα, για κάθε συστάδα έχουμε:

$$\begin{aligned}
 h(k_1) &= 1.0 \quad και \quad |k_1| = 3 \\
 h(k_2) &= 0.8 \quad και \quad |k_2| = 1 \\
 h(k_3) &= 0.4 \quad και \quad |k_3| = 1
 \end{aligned} \tag{5.55}$$

Το εννοιολογικό τους πλαίσιο υπολογίζεται ως ακολούθως:

$$\begin{aligned}
 C(k_1) &= apl/0.6 + tcp/0.58 \\
 C(k_2) &= \emptyset \\
 C(k_3) &= fcp/0.8
 \end{aligned} \tag{5.56}$$

Εφαρμόζοντας τον αδύναμο τροποποιητή  $w(a) = \sqrt{a}$ , παίρνουμε:

$$\begin{aligned}
 w(C(k_1)) &= \sqrt{C(k_1)} = apl/0.77 + tcp/0.76 \\
 w(C(k_2)) &= \emptyset \\
 w(C(k_3)) &= \sqrt{C(k_3)} = fcp/0.89
 \end{aligned} \tag{5.57}$$

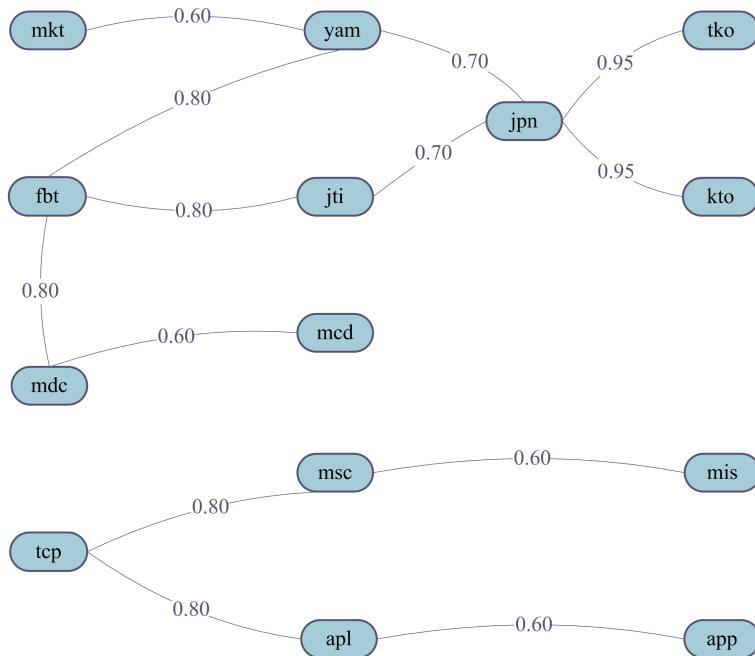
Σύμφωνα με τη θεωρία οι συστάδες  $k_2$  και  $k_3$  εμφανίζουν εξαιρετικά χαμηλή πληθυκότητα και για το λόγο αυτό περιλαμβάνουν παραπλανητικές οντότητες. Αφού ρυθμίσουμε κατάλληλα τους βαθμούς ιδιότητας μέλους των συστάδων σύμφωνα με τις γραμμικές πληθυκότητές τους και χρησιμοποιώντας τον τριγωνικό ασαφή αριθμό  $(1.3, 3, \infty)$ , οι δύο παραπάνω συστάδες αγνοούνται κατά τον υπολογισμό του  $\{W(k_i)\}$ .

Τελικά, το σύνολο των προτιμήσεων χρήστη που σχετίζεται με την Έλλη στο σημείο αυτό δίνεται από την:

$$\{W(k_i)\} = \bigcup_{k \in K} W(k) = W(k_1) = apl/0.77 + tcp/0.76 \quad (5.58)$$

Με την πάροδο του χρόνου και χωρίς βλάβη της γενικότητας μπορούμε να υποθέσουμε ότι το προταθέν πλαισιο έμαθε κάποιες από τις προτιμήσεις της Έλλης ακολουθώντας την προτεινόμενη μεθοδολογία και τα σχετικά ερωτήματά της προς το σύστημα. Με άλλα λόγια, θεωρούμε ότι το προφίλ της Έλλης εμπλουτίσθηκε και περιλαμβάνει πλέον τα βαθμονομημένα σημασιολογικά ενδιαφέροντά της για τις έννοιες της θεματικής περιοχής της οντολογίας. Οι τελευταίες περιλαμβάνουν διάφορες εταιρίες από τον τομέα τροφίμων, ποτών και τσιγάρων, καθώς και αρκετές εταιρίες πληροφορικής. Και πάλι, παρουσιάζουμε στη συνέχεια μόνο τις σχετικές έννοιες μαζί με τους βαθμούς συμμετοχής τους. Όπως είδαμε, η παραπάνω προσέγγιση μάς οδηγεί στην εκτίμηση του ασαφούς συνόλου  $P$  ως:

$$P = \{yam/0.85 + jti/0.92 + mcd/0.74 + apl/0.77 + msc/0.66 + tcp/0.76\} \quad (5.59)$$



**Σχήμα 5.22:** Ένα υποσύνολο της σχέσης  $\widehat{T}$ , το οποίο συνδέει τις έννοιες που εμπλέκονται στην επέκταση του κατά το χρόνο εκτέλεσης εννοιολογικού πλαισίου της Έλλης.

Προκειμένου να προχωρήσουμε στο επόμενο βήμα της ανίχνευσης του εννοιολογικού πλαισίου κατά το χρόνο εκτέλεσης, εκμεταλλευόμαστε τις σχέσεις μεταξύ των εννοιών του συνόλου  $P$ , όπως αυτές έχουν οριστεί σύμφωνα με τη σχέση  $\widehat{T}$  και παρουσιάζονται στο Σχήμα 5.22. Οι τιμές της σχέσης  $\widehat{T}$  ορίστηκαν αρχικά χρησιμοποιώντας όλη τη διαθέσιμη πληροφορία που βρίσκεται μέσα στη Βάση Γνώσης (KB), αναλύοντας χειρωνακτικά και ελέγχοντας την επίδραση της επέκτασης σε μια λίστα από σενάρια χρήσης για τη συνδυασμένη σχέση, και παραμετροιήθηκαν εμπειρικά μετέπειτα. Η ανάπτυξη μεθόδων για την αυτόματη εκμάθηση των παραπάνω τιμών αποτελεί μέρος των μελλοντικών επεκτάσεων της παρούσας διατριβής.

## Κεφάλαιο 5. Προσωποποιημένη πρόσβαση

Όταν η 'Ελλη εισάγει ένα ερώτημα (σημ.: θεωρούμε ότι η μηχανή αναζήτησης με βάση τα ερωτήματα είναι ένα "μαύρο κουτί" για την προτεινόμενη τεχνική), το σύστημα προσωποποίησης προσαρμόζει το αποτέλεσμα της κατάταξης στις προτιμήσεις της συνδυάζοντας την μετρική ομοιότητας με βάση το ερώτημα  $r_S(d, t)$  και τα με βάση τις προτιμήσεις σκορ  $r_P(d, t)$  για κάθε έγγραφο  $d$  που ταιριάζει με το ερώτημα, όπως περιγράφεται στην υποενότητα 5.6.2. Στο σημείο αυτό η προσαρμογή δεν έχει ολοκληρωθεί και δε λαμβάνει υπόψη της το εννοιολογικό πλαίσιο, μιας και η 'Ελλη έχει μόλις αρχικοποιήσει τη σύνοδο αναζήτησης και το εννοιολογικό πλαίσιο κατά το χρόνο εκτέλεσης παραμένει κενό (δηλαδή, για  $t = 0$ ,  $\widehat{C}_0 = \emptyset$ ).

Ας υποθέσουμε τώρα ότι η ανάγκη πληροφόρησης που εκφράζεται με το πρώτο ερώτημα σχετίζεται κατά κάποιο τρόπο με τις έννοιες Τόκιο (*tko*) και Κιότο (*kto*), μιας και η 'Ελλη θέλει να βρει πληροφορίες για τις πόλεις που θα επισκεφθεί. Η 'Ελλη προσπελαύνει και αποθηκεύει κάποια έγγραφα γενικών πληροφοριών σχετικά με τη ζωή και τις οικονομικές δραστηριότητες στις δύο αυτές πόλεις. Κατά συνέπεια, το σύστημα κατασκευάζει ένα εννοιολογικό πλαίσιο κατά το χρόνο εκτέλεσης από τα μεταδεδομένα των επιλεγμένων εγγράφων και το ερώτημα που έχει υποβληθεί, το οποίο σχηματίζει το ασαφές σύνολο  $\widehat{C}$ :

$$\widehat{C}_1 = \{tko/1.0 + kto/1.0\} \quad (5.60)$$

Τώρα η 'Ελλη θέλει να δει κάποιες γενικές πληροφορίες για Ιαπωνικές εταιρίες. Ενεργοποιείται, λοιπόν, ο μηχανισμός του εννοιολογικού πλαισίου ως ακολούθως:

1. Αρχικά, το σύνολο του εννοιολογικού πλαισίου επεκτείνεται μέσω των σημασιολογικών σχέσεων με την προσθήκη περισσότερων βαθμονομημένων εννοιών (που καταδειχνύονται με έντονα γράμματα) στο ασαφές σύνολο  $E(\widehat{C}_t)$  για την 'Ελλη, ακολουθώντας το μαθηματικό συμβολισμό που χρησιμοποιήθηκε στην υποενότητα 5.6.1 και το μέρος της σχέσης  $\widehat{T}$  που παρουσιάζεται στον Πίνακα 5.9.

Πίνακας 5.9: Μέρος της σχέσης  $\widehat{T}$

$s_1$	$s_2$	$\widehat{T}$	$s_1$	$s_2$	$\widehat{T}$	$s_1$	$s_2$	$\widehat{T}$
tko	jpn	0.95	kto	jpn	0.95	jpn	yam	0.70
jpn	jti	0.70	jti	fbt	0.80	fbt	yam	0.80
mkt	yam	0.60	fbt	mdc	0.80	mdc	mcd	0.60
tcp	msc	0.80	msc	mis	0.60	tcp	apl	0.80
apl	app	0.60						

Εφαρμόζοντας τη μεθοδολογία σημασιολογικής επέκτασης που περιγράφεται στην υποενότητα 5.6.1 και χρησιμοποιώντας στην περίπτωση αυτή το αλγεβρικό άθροισμα και το αλγεβρικό γινόμενο ως την ασαφή  $t$ -conorm και ασαφή  $t$ -νόρμα στην εξίσωση (5.43), αντίστοιχα, έχουμε:

$$E(\widehat{C}_1) = \{tko/1.00 + kto/1.00 + jpn/1.00 + jti/0.89 + \\ yam/0.64 + mkt/0.64 + fbt/0.78 + \\ mdc/0.67 + mcd/0.45\} \quad (5.61)$$

2. Σε αντίστοιχα με την παραπάνω διαδικασία οι αρχικές προτιμήσεις της 'Ελλης επεκτείνονται μέσω των σημασιολογικών σχέσεων. Οι προτιμήσεις που

επεκτάθηκαν και είναι αποθηκευμένες στο ασαφές σύνολο  $E(P)$  είναι οι ακόλουθες, όπου οι καινούριες/ενημερωμένες έννοιες καταδεικνύονται με έντονα μαύρα γράμματα:

$$\begin{aligned} E(P) = & \{yam/0.85 + jti/0.92 + tcp/0.76 + msc/0.66 + \\ & mcd/0.74 + apl/0.77 + \mathbf{jpn/0.89} + \mathbf{tko/0.86} + \\ & \mathbf{kto/0.86} + \mathbf{fbt/0.95} + \mathbf{mkt/0.83} + \mathbf{mdc/0.90} + \\ & \mathbf{mis/0.73} + \mathbf{app/0.75}\} \end{aligned} \quad (5.62)$$

3. Οι προτιμήσεις με βάση το εννοιολογικό πλαίσιο υπολογίζονται, όπως περιγράφεται στην υποενότητα 5.6.1, και καταλήγουν στο ακόλουθο ασαφές σύνολο  $CP$  (οι έννοιες με βαθμό συμμετοχής ίσο με 0 αγνοούνται):

$$CP_1 = \{yam/0.54 + jti/0.82 + mcd/0.33 + jpn/0.89 + tko/0.86 + \\ kto/0.86 + fbt/0.74 + mkt/0.53 + mdc/0.60\} \quad (5.63)$$

Αν συγχρίνουμε το παραπάνω με τις αρχικές προτιμήσεις της 'Ελλης στο προφίλ της, παρατηρούμε ότι η Microsoft Corp., η Apple Computers Inc. και οι εταιρίες Πληροφορικής εξαλείφονται ως προτιμήσεις εκτός εννοιολογικού πλαισίου, ενώ η Japan Tobacco Inc., η McDonald's και η Yamazaki Baking Co. παρέμειναν, επειδή σχετίζονται σημασιολογικά τόσο με τις αρχικές προτιμήσεις της 'Ελλης (τομέας τροφίμων), όσο και με το τρέχον εννοιολογικό πλαίσιο (Ιαπωνία). Επίσης, η Ιαπωνία, το Τόκιο και το Κιότο προστέθηκαν στις προτιμήσεις σύμφωνα με το αρχικό εννοιολογικό πλαίσιο. Αξίζει να σημειωθεί ότι η Ιαπωνία δε συμπεριλαμβανόταν στο αρχικό εννοιολογικό πλαίσιο κατά το χρόνο εκτέλεσης και προστίθεται εξαιτίας της σχέσης  $\widehat{T}$ . Τελικά, η Makoto Tajima, η McDonald's Corp. και ο Τομέας Τροφίμων, Ποτών & Τσιγάρων λαμβάνονται επίσης υπόψη με μικρότερους βαθμούς συμμετοχής ως προτιμήσεις χρήστη που βρίσκονται μέσα στο εννοιολογικό πλαίσιο.

4. Χρησιμοποιώντας τις παραπάνω προτιμήσεις σύμφωνα με το εννοιολογικό πλαίσιο, υπολογίζεται μια διαφορετική προσωποποιημένη κατάταξη ως απάντηση στο τρέχον ερώτημα χρήστη με βάση το ασαφές σύνολο  $E(\widehat{C}_1)$ , αντί για το βασικό ασαφές σύνολο προτιμήσεων  $P$ .

### 5.8.2 Δεύτερο σενάριο χρήσης

Για την καλύτερη απεικόνιση των προτεινόμενων τεχνικών προσωποποίησης με βάση το εννοιολογικό πλαίσιο εξετάζουμε, στη συνέχεια, ένα ακόμα σενάριο χρήσης: 'Εστω ότι η οικογένεια της Κλειούς και οι φίλοι της έχουν δημιουργήσει έναν κοινό αποθηκευτικό χώρο στο Διαδίκτυο, όπου φορτώνουν και μοιράζονται τις φωτογραφίες και τα βίντεο τους. Η Κλειώ δεν έχει ελέγξει τη συλλογή αυτή για αρκετό καιρό και θέλει να ρίξει μια ματιά στις φωτογραφίες που έχουν "ανεβάσει" οι συγγενείς της από τις διακοπές του περασμένου καλοκαιριού. Υποθέτουμε ότι το προτεινόμενο πλαίσιο έχει μάθει μερικές από τις προτιμήσεις της Κλειούς κατά τη διάρκεια του χρόνου, δηλ. έστω ότι το προφίλ της Κλειούς περιλαμβάνει τα βαθμονομημένα σημασιολογικά ενδιαφέροντα για τις έννοιες της θεματικής περιοχής της οντολογίας που παρουσιάζονται στον Πίνακα 5.10, όπου ο *Rex* είναι το κατοικίδιο ζώο του αδελφού της και όλα

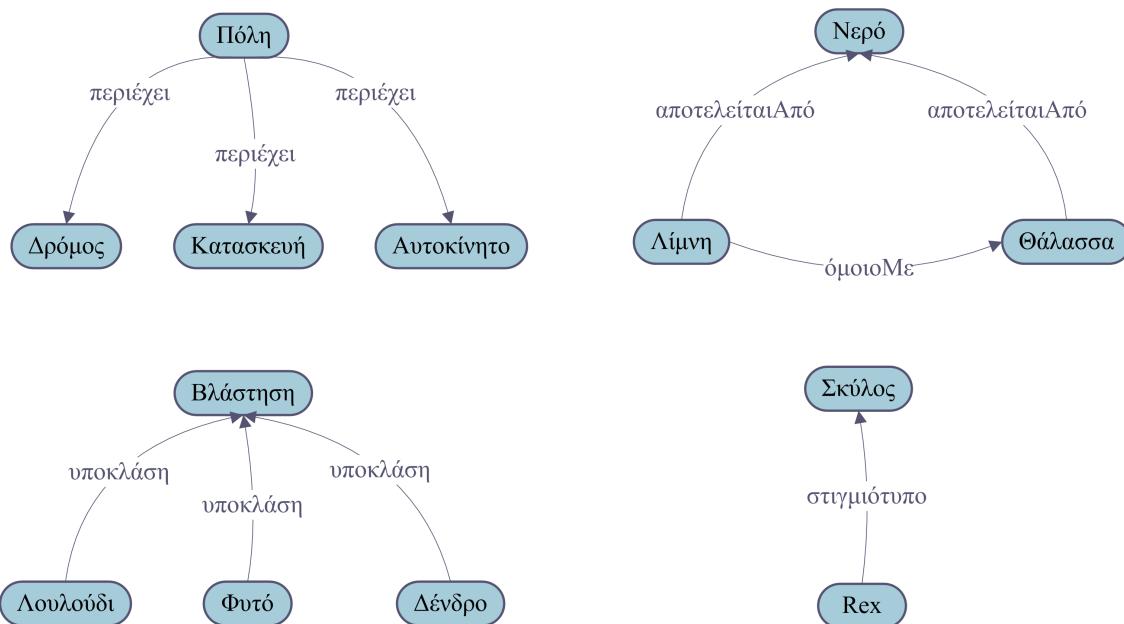
## Κεφάλαιο 5. Προσωποποιημένη πρόσσβαση

**Πίνακας 5.10:** Προτιμήσεις χρήστη.

	P(Κλειώ)
Αυτοκίνητο	1.0
Πόλη	1.0
Θάλασσα	1.0
Rex	1.0
Βλάστηση	1.0

τα βάρη έχουν ορισθεί ως 1.0, προκειμένου να απλοποιήσουμε το παράδειγμα. Με τον τρόπο αυτό καθορίζεται αρχικά το σύνολο  $P$ .

Στην καινούρια θεώρησή μας οι παραπάνω έννοιες ορίζονται σε μια οντολογία θεματικής περιοχής, η οποία περιέχει άλλες έννοιες και τις σχέσεις μεταξύ τους, ένα υποσύνολο των οποίων παρουσιάζεται στο Σχήμα 5.23.



**Σχήμα 5.23:** Ένα υποσύνολο της οντολογίας θεματικής περιοχής, το οποίο περιέχει τις έννοιες που περιλαμβάνονται στο δεύτερο σενάριο χρήσης.

Το βάρος διάδοσης που συσχετίζεται χειρωνακτικά με κάθε σημασιολογική σχέση παρουσιάζεται στον Πίνακα 5.11. Τα βάρη ετέθησαν μέσω μιας αρχικής ανάλυσης και ενός αρχικού ελέγχου της επίδρασης της διάδοσής τους σε έναν κατάλογο σεναρίων χρήσης για κάθε σχέση και συντονίστηκαν με εμπειρικό τρόπο στη συνέχεια.

**Πίνακας 5.11:** Βάρη διάδοσης των σημασιολογικών σχέσεων.

Σχέση r	w(r)	w(r <sup>-1</sup> )
περιέχει	0.6	0.5
στιγμιότυπο	1.0	0.3
αποτελείται Από	0.7	0.6
όμοιο Με	0.8	0.8
υποκλάση	1.0	0.3

Όταν η Κλειώ εισάγει ένα ερώτημα  $q_1$  (η βασισμένη στα ερωτήματα μηχανή αναζήτησης μπορεί και πάλι να θεωρηθεί ουσιαστικά ως ένα κλειστό “μαύρο κουτί” για

την τεχνική μας), το σύστημα προσωποποίησης προσαρμόζει το αποτέλεσμα της ταξινόμησης στις προτιμήσεις της Κλειούς, συνδυάζοντας τα σκορ του βασισμένου στο ερώτημα  $sim(d, q)$  και του βασισμένου στις προτιμήσεις  $prm(d, Kλειω)$  για κάθε πολυμεσικό έγγραφο  $d$  που ταιριάζει με το ερώτημα. Σε αυτό το σημείο, η προσαρμογή δεν έχει λάβει υπόψη της το εννοιολογικό πλαίσιο, δεδομένου ότι η Κλειώ μόλις άρχισε τη σύνοδο αναζήτησης, και το εννοιολογικό πλαίσιο του χρόνου εκτέλεσης είναι ακόμα κενό (δηλ.  $t = 0, C(0) = \emptyset$ ). Ας υποθέσουμε, στη συνέχεια, ότι η Κλειώ επιλέγει και “κατεβάζει” μια φωτογραφία και ένα βίντεο, τα οποία παρουσιάζονται στο Σχήμα 5.24, από τα αποτελέσματα της αναζήτησης.



**Σχήμα 5.24:** Η εικόνα και το βίντεο που επιλέχθηκαν από το χρήστη.

Κατά συνέπεια, το σύστημα χτίζει ένα εννοιολογικό πλαίσιο κατά το χρόνο εκτέλεσης από τα μεταδεδομένα των επιλεγμένων πολυμεσικών εγγράφων, συμπεριλαμβανομένων των στοιχείων που παρουσιάζονται στον Πίνακα 5.12. Αυτό αντιστοιχεί στο σύνολο  $C$  (που για  $t = 1$  είναι ίσο με το  $R(t)$ ).

**Πίνακας 5.12:** Σύνολο εννοιολογικού πλαισίου.

$C(Kλειώ, 1)$
Κατασκευή 1.0
Φυτό 1.0

Τώρα, η Κλειώ θέλει να δει κάποια φωτογραφία των μελών της οικογένειάς της και θέτει ένα νέο ερώτημα  $q_2$ . Ο μηχανισμός επίδρασης του εννοιολογικού πλαισίου ενεργοποιείται, ως εξής.

1. Αρχικά, επεκτείνεται το σύνολο του εννοιολογικού πλαισίου μέσω των σημασιολογικών σχέσεων προσθέτοντας δύο επιπλέον σταθμισμένες έννοιες, που παρουσιάζονται με έντονους μαύρους χαρακτήρες στον Πίνακα 5.13. Αυτό διαμορφώνει το σύνολο  $EC$ .
2. Ομοίως, επεκτείνονται και οι προτιμήσεις της Κλειούς μέσω των σημασιολογικών σχέσεων. Οι επεκταθείσεις προτιμήσεις που αποθηκεύονται στο σύνολο  $EP$  παρουσιάζονται στον Πίνακα 5.14, όπου οι νέες έννοιες διαχρίνονται με έντονους μαύρους χαρακτήρες.

*Κεφάλαιο 5. Προσωποποιημένη πρόσβαση*

**Πίνακας 5.13:** *Επεκτεταμένο εννοιολογικό πλαίσιο.*

$EC(Kλειώ, 1)$	
Κατασκευή	1.0
Πόλη	0.6
Φυτό	1.0
<b>Βλάστηση</b>	0.4

**Πίνακας 5.14:** *Επεκτεταμένες προτιμήσεις χρήστη.*

$EP(Kλειώ)$	
Αυτοκίνητο	1.0
Πόλη	1.0
Κατασκευή	0.7
Σκύλος	0.3
Λίμνη	0.8
Λουλούδι	1.0
Φυτό	1.0
<b>Δένδρο</b>	1.0
<b>Δρόμος</b>	0.5
Θάλασσα	1.0
Rex	1.0
Βλάστηση	1.0
<b>Νερό</b>	0.7

3. Οι προτιμήσεις με βάση το εννοιολογικό πλαίσιο υπολογίζονται από το αλγερικό γινόμενο των συνόλων  $EC$  και  $EP$ , τα οποία παράγουν το σύνολο  $CP$  που παρουσιάζεται στον Πίνακα 5.15 (οι έννοιες με βάρος 0 παραλείπονται).

**Πίνακας 5.15:** *Προτιμήσεις χρήστη με βάση το εννοιολογικό πλαίσιο.*

$CP(Kλειώ, 1)$	
Κατασκευή	0.7
Πόλη	0.6
Φυτό	1.0
<b>Βλάστηση</b>	0.4

Συγκρίνοντας το τελευταίο με τις αρχικές προτιμήσεις στο προφίλ της Κλειούς, είμαστε σε θέση να δούμε ότι οι έννοιες *Αυτοκίνητο*, *Θάλασσα* και *Rex* αγνοούνται ως εκτός εννοιολογικού πλαισίου προτιμήσεις, ενώ προστέθηκαν οι έννοιες *Κατασκευή* και *Λουλούδι*, επειδή αυτές σχετίζονται έντονα σημασιολογικά τόσο με τις αρχικές προτιμήσεις της Κλειούς, όσο και με το τρέχον εννοιολογικό πλαίσιο.

4. Χρησιμοποιώντας τις παραπάνω προτιμήσεις, υπολογίζεται μια διαφορετική, προσωποποιημένη ταξινόμηση σε απάντηση του τρέχοντος ερωτήματος του χρήστη  $q_2$ , σύμφωνα με το σύνολο  $EC(Kλειώ, 1)$ , αντί του βασικού συνόλου προτιμήσεων  $P(Kλειώ)$ .

Το παραδειγμα αυτό επεξηγεί πως η μέθοδός μας μπορεί να χρησιμοποιηθεί για την εκμετάλλευση του εννοιολογικού πλαισίου κατά την προσωποποίηση σε ένα βασισμένο στα ερωτήματα σύστημα αναζήτησης, όπου τα ερωτήματα μπορούν να είναι οποιουδήποτε είδους: οπτικά, ερωτήματα λέξεων-κλειδιών, ερωτήματα φυσικής γλώσσας, κ.λ.π.. Η προσέγγιση αυτή θα μπορούσε να εφαρμοστεί με παρόμοιο τρόπο και σε άλλους τύπους υπηρεσιών πρόσβασης περιεχομένου, όπως είναι οι εξατομικευμένες ικανότητες ζεψυλλίσματος για αποθήκες πολυμέσων, η αυτόματη παραγωγή εξατομικευμένων διαφανειών επίδειξης, η παραγωγή εξατομικευμένων τηλεοπτικών περιλήψεων, κ.λ.π..



# Κεφάλαιο 6

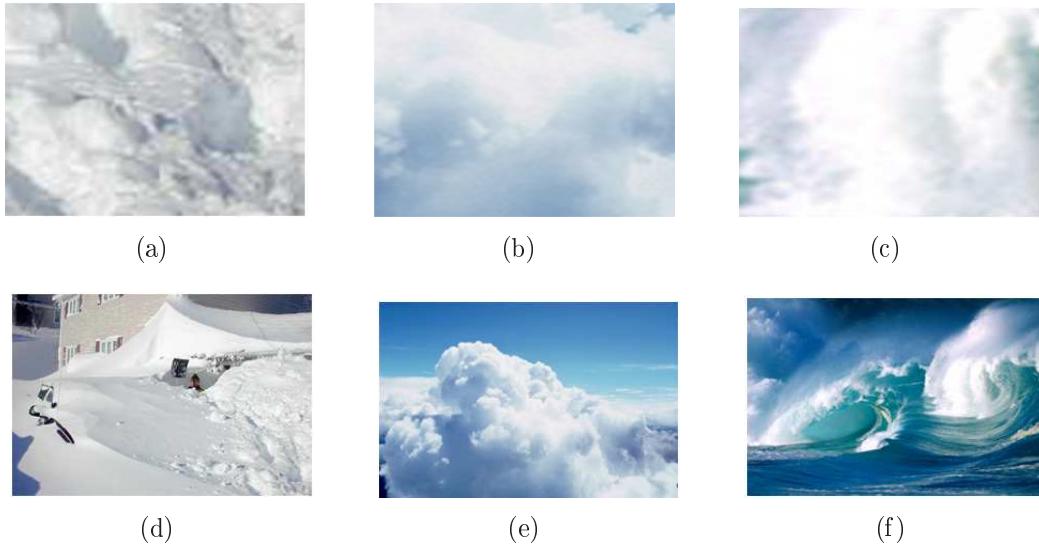
## Εννοιολογικό πλαίσιο και οπτική πληροφορία

### 6.1 Εισαγωγή

Στο κεφάλαιο αυτό παρουσιάζουμε τις βασικές αρχές της έρευνάς μας πάνω στην επίδραση του οπτικού εννοιολογικού πλαισίου κατά την ανάλυση των πολυμεσικών εγγράφων. Η οπτική πληροφορία που ενυπάρχει στο πολυμεσικό περιεχόμενο μπορεί να αποβεί ιδιαίτερα χρήσιμη σε περιπτώσεις αναγνώρισης αντικειμένων και ταξινόμησης σκηνών. Η αξιοποίηση αυτής της πληροφορίας είναι, όπως θα δούμε, συνδεδεμένη με την έννοια του εννοιολογικού πλαισίου, η οποία σε αυτό το ερευνητικό πεδίο μπορεί να λάβει ποικίλες διαφορετικές υποστάσεις. Η τελευταία βασίζεται σε μίας μορφής γνώση που ορίζεται, μοντελοποιείται και αποθηκεύεται σε οντολογίες (ontologies), η ανάπτυξη των οποίων βασίζεται σε συγκεκριμένες δομές και προηγείται χρονικά των αλγορίθμων ανάλυσης. Τέλος, πραγματοποιείται μια εκτενής αναφορά στις εφαρμογές της χρήσης του εννοιολογικού πλαισίου σε τεχνικές και υλοποιήσεις κατά τη διαδικασία της ανάλυσης της πολυμεσικής πληροφορίας και εξετάζονται τα αποτελέσματα της έρευνάς μας σε τομείς, όπως η υποβοηθούμενη από τη γνώση χαμηλού επιπέδου ανάλυση εικόνων, η κατάτμηση εικόνων και η εξαγωγή σημασιολογικών συμπερασμάτων με χρήση ενός γενετικού αλγορίθμου.

#### 6.1.1 Υπάρχουσα κατάσταση - Ορισμός προβλήματος

Είναι, εν γένει, αποδεκτό ότι όλα τα αντικείμενα του πραγματικού κόσμου είναι συνυφασμένα με κάποιο εννοιολογικό πλαίσιο. Επίσης, είναι μάλλον προφανές ότι ένας ακριβής ορισμός ενός οποιοδήποτε τέτοιου “αντικειμένου”, χωρίς να λάβει κανείς υπόψη του το σκοπό και το εννοιολογικό του πλαίσιο, είναι ιδιαίτερα δύσκολος έως αδύνατος. ’Όταν στηριζόμαστε αποκλειστικά και μόνο σε χαμηλού επιπέδου χαρακτηριστικά, όπως είναι το χρώμα και η υφή μιας περιοχής μιας εικόνας, είναι ιδιαίτερα εύκολο να υποπέσουμε σε λανθασμένες εκτιμήσεις. Για παράδειγμα, βασιζόμενοι στην πρώτη σειρά του Σχήματος 6.1, είναι μάλλον αρκετά δύσκολο να αποφασίσουμε χωρίς αμφιβολία αν πρόκειται για εικόνες με σύννεφα, χιόνι ή κύματα. Αντιθέτως, στη δεύτερη σειρά η παραπάνω απόφαση είναι ξεκάθαρη, μιας και υπεισέρχεται το εννοιολογικό πλαίσιο της εκάστοτε εικόνας. Στην πραγματικότητα, η σωστή αναγνώριση απομονωμένων τμημάτων καθαρών υλικών χωρίς τη βοήθεια του εννοιολογικού τους πλαισίου αποτελεί μια διαδικασία-πρόκληση ακόμα και για τους ανθρώπους. ’Οπως



**Σχήμα 6.1:** Επάνω σειρά: χιόνι, σύννεφα και κύματα εκτός εννοιολογικού πλαισίου. κάτω σειρά: Οι ίδιες έννοιες εντός του εννοιολογικού τους πλαισίου.

παρατηρούμε στο Σχήμα 6.2, απομονωμένα τμήματα χιονιού εμφανίζονται ανάμεικτα με απομονωμένα τμήματα σύννεφου και υφάσματος, πράγμα που καθιστά εξαιρετικά δύσκολη την αναγνώρισή του καθενός με απόλυτη σιγουριά. Αρκετοί ερευνητές έχουν προτείνει διάφορες μεθόδους για την αναγνώριση συγκεκριμένων υλικών, όπως είναι π.χ. ο ουρανός ή το γρασίδι ([191], [210], [233], [262], [229]), αλλά οι περισσότερες από αυτές δεν αξιοποιούν την πληροφορία του εννοιολογικού πλαισίου, η οποία μπορεί να βοηθήσει αποτελεσματικά τη διαδικασία αναγνώρισης συγκεχυμένων υλικών που εμφανίζουν παρόμοια χαμηλού επιπέδου χαρακτηριστικά χρώματος και υφής.



**Σχήμα 6.2:** Αναγνώριση παρόμοιων υλικών χωρίς τη βοήθεια του εννοιολογικού πλαισίου. τα επιλεγμένα τμήματα είναι αυτά που απεικονίζουν χιόνι στην πραγματικότητα.

Επίσης, διάφορες ενδείξεις που βασίζονται σε στατιστικά στοιχεία από τον καθημερινό οπτικό κόσμο που μας περιβάλλει μπορούν να φανούν χρήσιμες και να καθοδηγήσουν μια τέτοια απόφαση [254]. Αυτό το εννοιολογικό πλαίσιο μπορεί να λάβει τη μορφή καθολικών στατιστικών από εικόνες, τα οποία χαρακτηρίζουν έναν συγκεκριμένο τύπο περιβάλλοντος, όπως π.χ. μια σκηνή εσωτερικού χώρου γραφείων ή μια υπαίθρια σκηνή ενός κήπου ([20], [85], [193], [282]). Για παράδειγμα, η εικόνα

## Κεφάλαιο 6. Εννοιολογικό πλαίσιο και οπτική πληροφορία

ενός αλόγου είναι πιθανότερο να είναι παρούσα σε ένα περιβάλλον εξωτερικών τοπίων, όπως π.χ. ένα πράσινο λιβάδι, παρά σε κάποιο εσωτερικό χώρο, ενώ από την άλλη συνήθως αναμένουμε ότι ένα έπιπλο, όπως π.χ. ένας καναπές, θα βρίσκεται στο εσωτερικό ενός κτιρίου.



(a) απομονωμένο αντικείμενο  
(b) αντικείμενο ενταγμένο στο πλαίσιό του  
(c) απομονωμένο αντικείμενο  
(d) αντικείμενο ενταγμένο στο πλαίσιό του

**Σχήμα 6.3:** Απομονωμένα αντικείμενα και αντικείμενα συνυφασμένα με το εννοιολογικό τους πλαίσιο.

Και σε επίπεδο αντικειμένων μια εικόνα από μόνη της δεν επαρκεί, για να επιτρέψει σε έναν αλγόριθμο ή ακόμα και σε έναν άνθρωπο να αποφασίσει ξεκάθαρα πού αρχίζει το ένα αντικείμενό της και πού τελειώνει το άλλο. Ο προσδιορισμός ενός αντικειμένου σε μια εικόνα, ή ακόμα και μιας άλλης εικόνας του ίδιου αντικειμένου, όταν αυτό βρίσκεται σε πρώτο πλάνο, μπορεί να είναι εξαιρετικά δύσκολος χωρίς τη συνοδεία χρήσιμης και βασισμένης στο εννοιολογικό πλαίσιο πληροφορίας. Επεκτείνοντας τον συλλογισμό αυτό και, όπως απεικονίζεται στο Σχήμα 6.3, οι εικόνες μιας χαρτιέρας ή μιας τοστιέρας σε πρώτο πλάνο είναι δυσκολότερο να αναγνωρισθούν, όταν παρατηρούνται χωρίς τις περιβαλλοντικές πληροφορίες του υπολοίπου της εικόνας (δηλαδή του δωματίου της κουζίνας εν προκειμένω).

Η έννοια του εννοιολογικού πλαισίου από την άποψη των κινητών ή πανταχού παρόντων υπολογιστικών συστημάτων εμφανίστηκε αρχικά στα μέσα της δεκαετίας του '90 και συνδέθηκε με την εργασία του Schilit [219]. Στην περίπτωση αυτή, το εννοιολογικό πλαίσιο συνδέεται με τις σταθερές αλλαγές που εμφανίζονται σε τρία επίπεδα περιβάλλοντος: το υπολογιστικό περιβάλλον, το περιβάλλον του χρήστη και το φυσικό περιβάλλον. Στην παρούσα ενότητα της εργασίας μας περιορίζουμε την έννοια του εννοιολογικού πλαισίου στο οπτικό περιβάλλον. Σε μια πρόσφατη έρευνα ο Torralba [257] περιγράφει συνοπτικά την ιστορία της μοντελοποίησης του οπτικού εννοιολογικού πλαισίου. Περαιτέρω μελέτες των Biederman [33] και Palmer [199] δίνουν έμφαση στην επίδραση του εννοιολογικού πλαισίου στο χρόνο επεξεργασίας κατά την αναγνώριση αντικειμένων σε μία σκηνή. Ο Rensink [206] έχει δείξει ότι οι αλλαγές στον πραγματικό κόσμο παρατηρούνται γρηγορότερα στα αντικείμενα ή στις περιοχές ενδιαφέροντος, προτείνοντας κατά συνέπεια μια προνομιακή επέκταση της προσοχής σε αυτά τα μέρη μιας σκηνής. Οι Henderson και Hollingworth [117] έχουν εκθέσει αποτελέσματα που προτείνουν ότι η επιλογή των παραπάνω περιοχών κυβερνάται όχι μόνο από τη χαμηλού επιπέδου περίοπτη θέση τους, αλλά και από τη σημασιολογία της σκηνής. Σε παρόμοιο πνεύμα διάφορες ερευνητικές μελέτες υποστηρίζουν την ιδέα ότι η σημασιολογία μιας σκηνής είναι δυνατόν να είναι διαθέσιμη από τα αρχικά στάδια της αλυσίδας επεξεργασίας των πληροφοριών [221], [253], και τολμούν να ισχυριστούν ότι η αναγνώριση της σκηνής μπορεί και να μην απαιτεί την αναγνώριση των αντικειμένων ως το πρώτο βήμα [196], [221].

Με βάση τα παραπάνω, θα αναφερόμαστε από τούδε και στο εξής στο

(οπτικό) εννοιολογικό πλαίσιο ως την πληροφορία εκείνη που είναι σχετική με το οπτικό περιεχόμενο της σκηνής μιας ακίνητης εικόνας ή μιας ακολουθίας βίντεο. Το τελευταίο αποφαίνεται εξαιρετικά χρήσιμο στις περιπτώσεις:

1. *Ταξινόμησης σκηνής*, όπου τα χαμηλού επιπέδου οπτικά χαρακτηριστικά γνωρίσματα υιοθετούνται, για να αναλύσουν συνολικά το περιεχόμενο της σκηνής και να το ταξινομήσουν σε μια σειρά διαφόρων προκαθορισμένων κατηγοριών (όπως π.χ. εσωτερικός/υπαίθριος χώρος, πόλη/τοπίο, κ.ο.κ.).
2. *Ανίχνευσης/αναγνώρισης αντικειμένων*, η οποία εστιάζει στην τοπική (local) ανάλυση (μιας περιοχής) χρησιμοποιώντας προγενέστερη γνώση, για να ανιχνεύσει και να αναγνωρίσει γνωστά αντικείμενα στις περιορισμένες περιοχές μιας εικόνας.

Θα μπορούσαμε να ορίσουμε το εννοιολογικό πλαίσιο ενός συγκεκριμένου αντικειμένου από τη σκοπιά άλλων - προηγουμένως αναγνωρισμένων - αντικειμένων μέσα στη σκηνή. Στην περίπτωση αυτή η αναπαράσταση του εννοιολογικού πλαισίου είναι αντικειμενοστραφής και απαιτεί την αναγνώριση του αντικειμένου ως το πρώτο βήμα. Το θέμα της αυτόματης ανίχνευσης των σημαντικών περιοχών σε μια εικόνα και της ανάθεσης των κατάλληλων ετικετών σε αντικείμενα έχει αντικειμενική στο παρελθόν από διάφορους ερευνητές [288], [198], [247]. Μια τέτοια μεθοδολογία παρουσιάζεται στο [173], όπου αναπτύσσεται ένα υπολογιστικό μοντέλο οπτικής προσοχής (visual attention), το οποίο βασίζεται στο συνδυασμό της γνώσης για το ανθρώπινο οπτικό σύστημα με τεχνικές όρασης υπολογιστών. Η έννοια της κατάτμησης περιοχών στοχεύει στην παραγωγή περιοχών με ομοιογενείς - χαμηλού επιπέδου - ιδιότητες, όπως το χρώμα και η υφή, οι οποίες αποτελούν τις βασικές μονάδες μιας σκηνής, καθώς και μια αποδοτική ενδιάμεση αναπαράστασή της με απότερο στόχο την υψηλού επιπέδου εξαγωγή συμπερασμάτων. Το ερευνητικό αντικείμενο αυτό θεωρείται γενικά ως κάτι που ένας ανθρώπινος παρατηρητής εκτελεί με σχετική ευκολία, αλλά συγχρόνως είναι ιδιαίτερα δύσκολο να γίνει κατανοητό από ένα αυτοματοποιημένο σύστημα.

Εντούτοις, η αναπαράσταση του εννοιολογικού πλαισίου δεν απαιτεί τη γραμματική ανάλυση μιας εικόνας προκειμένου να κατασκευασθεί η αναπαράσταση μιας σκηνής. Όπως προτείνεται στο [196], είναι δυνατό να κατασκευασθεί μια περιγραφή της σκηνής, που παρακάμπτει τις ταυτότητες των αντικειμένων και στην οποία η σκηνή αναπαριστάται από μία ενιαία οντότητα. Μια τέτοια αναπαράσταση βασίζεται στον προσδιορισμό διάφορων ιδιοτήτων, που σχετίζονται με τη σκηνή και που δεν αναφέρονται σε μεμονωμένα αντικείμενα. Η πιο πρόσφατη εργασία του Torralba εισάγει ένα πρώτο μοντέλο για την ενσωμάτωση των βασισμένων στο εννοιολογικό πλαίσιο πληροφοριών στη διαδικασία ανίχνευσης αντικειμένων και ταξινόμησης σκηνής [257], [255], [256]. Η προσέγγιση αυτή είναι βασισμένη στη βελτίωση των στατιστικών των εννοιολογικών χαρακτηριστικών γνωρισμάτων της σκηνής σύμφωνα με την παρουσία ή την απουσία κάποιων κατηγοριών αντικειμένων.

Το ανθρώπινο οπτικό σύστημα, από την άλλη, είναι σε θέση να ελαττώσει το μέγεθος των εισερχόμενων οπτικών πληροφοριών σε μία σχετικά μικρή αλλά απόλυτα ουσιώδη ποσότητα πληροφοριών, προκειμένου να επιτευχθεί μια υψηλότερου επιπέδου γνωστική επεξεργασία. Στη βιβλιογραφία έχουν προταθεί δύο συμπληρωματικοί μηχανισμοί για την επιλογή των μεμονωμένων αντικειμένων: η “από κάτω προς τα επάνω” επιλεκτική προσοχή και η συσταδοποίηση που βασίζεται στην κατάτμηση. Ενώ η προσοχή επικεντρώνεται στις αντιθέσεις των χαρακτηριστικών γνωρισμάτων

[128], η συσταδοποίηση και η κατάτμηση προσπαθούν να βρουν ομοιογενείς περιοχές με βάση ορισμένα χαρακτηριστικά γνωρίσματα [225], [165], [17]. Οι βασισμένες στην προσοχή τεχνικές στηρίζονται σε μηχανισμούς, που δρούν “από κάτω προς τα επάνω” και που είναι με τέτοιον τρόπο σχεδιασμένοι, ώστε να εστιάζουν σε περιοχές υψηλής αντίθεσης, οι οποίες αντιστοιχούν συχνά στα αντικείμενα, ενώ οι τεχνικές συσταδοποίησης συνδέουν με έναν αποδοτικό τρόπο τα αποσπασματικά χαρακτηριστικά γνωρίσματα μιας εικόνας (όπως π.χ. τις άκρες). Οι μέθοδοι αυτές (σκόπιμα έως ένα βαθμό) δεν περιέχουν κάποιου είδους γνώση για το πώς μπορεί μια εικόνα ή μια ακολουθία βίντεο να χρησιμοποιηθεί σε ένα σύστημα για την επίτευξη ενός στόχου υψηλοτέρου επιπέδου, όπως π.χ. είναι η ακριβής ανίχνευση ενός αντικειμένου (μέσω του καθορισμού των ορίων του) ή η αναγνώρισή του. Η ανάλυση και χρήση του εννοιολογικού πλαισίου είναι και σε αυτή την περίπτωση το κλειδί που θα επιτρέψει την αποδοτική επίτευξη των στόχων αυτών.

Είναι γεγονός ότι ποικίλες μεθοδολογίες ανάλυσης εικόνας και ακολουθιών βίντεο είναι δυνατόν να χρησιμοποιηθούν για την επίτευξη μιας αποτελεσματικής κατανόησης του περιεχομένου και ότι οι σημερινοί ερευνητές κοιτάζουν πέρα από τα χαμηλού επιπέδου χαρακτηριστικά γνωρίσματα, όπως το χρώμα, η υφή και το σχήμα, αναζητώντας αποτελεσματικότερες μεθόδους έρευνας. Για παράδειγμα, οι Naphade και Huang [191] χρησιμοποιούν έναν κατάλογο σημασιολογικών αντικειμένων, συμπεριλαμβανομένων των ουρανός, χιόνι, βράχος, νερό και δάσος, σε ένα παραγοντικό πλαίσιο γράφου (factor graph-based framework) για τη σημασιολογική ευρετηρίαση και ανάκτηση ακολουθιών βίντεο. Οι περισσότερες από αυτές τις εργασίες ανήκουν στη λεγόμενη αναπαράσταση μέσου επιπέδου, συμπεριλαμβανομένων και των πολύ συχνά χρησιμοποιούμενων ιστογραμμάτων και των τοπικών μεθόδων στατιστικών υφής [13], [92], [152], [52]. Εντούτοις, ένας σημαντικός αριθμός λανθασμένων ταξινομήσεων εξακολουθεί να εμφανίζεται, συνήθως λόγω των ομοιοτήτων που υφίστανται στα χαρακτηριστικά γνωρίσματα χρώματος και υφής των διάφορων τύπων αντικειμένων. Όσον αφορά στην ταξινόμηση σκηνής, οι περισσότερες από τις παραπάνω μέθοδοι δε χρησιμοποιούν καμιά πληροφορία εννοιολογικού πλαισίου. Κατά συνέπεια, παράγουν συνήθως φτωχά αποτελέσματα, ειδικά όταν συγχρίνονται με την ανθρώπινη αντίληψη και την ανθρώπινη δυνατότητα να ταξινομεί περιοχές και να ανιχνεύει/αναγνωρίζει αντικείμενα. Η ενσωμάτωση των πληροφοριών οπτικού εννοιολογικού πλαισίου, που προτείνεται στη διατριβή αυτή, θα ενισχύσει την απόδοση των μεθοδολογιών ανίχνευσης φυσικών αντικειμένων, μικρούμενη ως ένα βαθμό την ανθρώπινη προσέγγιση σε παρόμοια προβλήματα. Η μετάβαση από την αντιπροσώπευση εικονοστοιχείων σε μία αντιπροσώπευση μέσου επιπέδου, στην οποία η εικόνα/το βίντεο αντιπροσωπεύεται ως σύνολο συνεπών περιοχών σε ένα διάστημα χαρακτηριστικών γνωρίσματων, παραμένει εξαιρετικά σημαντική. Π.χ. ο Greenspan και κάποιοι άλλοι [103], [101], [102] προτείνουν μια μέθοδο για την χωρίς επίβλεψη συσταδοποίηση της εικόνας, η οποία αφενός χρησιμοποιεί συνεχή πιθανοτικά γκαουσσιανά μοντέλα ανάμιξης (GMMs) για την αναπαράσταση της εικόνας και αφετέρου θεωρητικές αρχές πληροφοριών για την ταξινόμηση της εικόνας και την οργάνωση των σχετικών βάσεων δεδομένων. Η μέθοδος του παρέχει ένα συνεπή και μάλλον απλό στην υλοποίηση αλγόριθμο για την αναπαράσταση του οπτικού περιεχομένου, ο οποίος μπορεί να ρυθμιστεί, για να λάβει ως έσοδο τα αποτελέσματα ευρέως διαδεδομένων προσεγγίσεων κατάτμησης περιοχών.

Η χρήση του εννοιολογικού πλαισίου αποτελεί το κλειδί για μία ζεκάθαρη διαδικασία αναγνώρισης, καθώς αφ' ενός αναφέρεται στις σχέσεις που διέπουν τη θέση

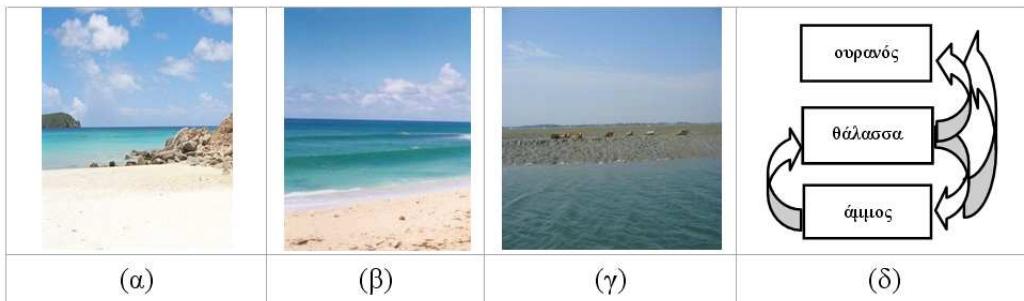
διαφορετικών αντικειμένων στη σκηνή της εικόνας και αφ' ετέρου εκμεταλλεύεται την ίδια τη σημασιολογία τους. Ως εκ τούτου, το εννοιολογικό πλαίσιο μπορεί, εν γένει, να λάβει ποικίλες μορφές [182]. Δύο από τις πιο συνηθισμένες από αυτές είναι το χωρικό και το χρονικό εννοιολογικό πλαίσιο. Το χωρικό εννοιολογικό πλαίσιο ([228], [285], [53], [156], [228]) σχετίζεται με τις χωρικές σχέσεις μεταξύ αντικειμένων ή περιοχών σε μία ακίνητη εικόνα ή σε μια ακολουθία βίντεο, ενώ το χρονικό εννοιολογικό πλαίσιο [38] έχει να κάνει με τις χρονικές σχέσεις που διέπουν τα αντικείμενα, τις περιοχές ή τις σκηνές στις ακολουθίες βίντεο. Στις απλές εικόνες μπορούν να αναγνωριστούν τουλάχιστον δύο τύποι χωρικών σχέσεων εννοιολογικού πλαισίου που να έχουν νόημα. Πρώτον, υφίστανται σχέσεις μεταξύ της επαναλαμβανόμενης εμφάνισης και της συνύπαρξης ορισμένων αντικειμένων στις εικόνες. Για παράδειγμα, η αναγνώριση χιονιού σε μια εικόνα υπαινίσσεται μία μικρή πιθανότητα για την ύπαρξη γρασίδιου στην ίδια εικόνα. Δεύτερον, υπάρχουν σχέσεις μεταξύ των χωρικών τοποθετήσεων συγκεκριμένων αντικειμένων σε μία εικόνα: το γρασίδι τείνει να βρίσκεται κάτω από τον ουρανό, ο ουρανός πάνω από το χιόνι, κ.ο.χ.. Φυσικά, το σύνολο των χωρικών σχέσεων μπορεί να είναι εξαιρετικά πλούσιο (π.χ. πολλές χωρικές σχέσεις με απειροελάχιστες διαφορές μεταξύ τους) ή αραιό (λιγότερες διαχριτές σχέσεις). Σε κάθε περίπτωση, όμως, οι χωρικές σχέσεις ορίζουν την απόλυτη ή σχετική χωρική πληροφορία μεταξύ των αντικειμένων. Ο απώτερος στόχος είναι η ανάπτυξη μιάς μεθόδου, η οποία δε θα εξαρτάται από τη συγκεκριμένη σκηνή για τη δημιουργία μοντέλων εννοιολογικού πλαισίου, που θα αποβούν χρήσιμα για την γενικότερη κατανόηση των σκηνών μιας εικόνας.

Επιπρόσθετα, υπάρχουν ακόμα δύο παράγοντες που πρέπει να ληφθούν υπόψη κατά την αντικειμενική θεμάτων κατηγοριοποίησης σκηνών με βάση την πληροφορία που μας παρέχει κάποια από τις μορφές του εννοιολογικού πλαισίου. Αφ' ενός, είναι εξαιρετικά σημαντικό να συνειδητοποιήσουμε ότι η συνολική οργάνωση των μερών μιας σκηνής επηρεάζει κατά πολύ την ερμηνεία της. Αρκετές ερευνητικές μελέτες έχουν δείξει ότι ένα ερεθισμα τοποθετημένο στη σωστή χωρική σειρά επιτρέπει την ακριβέστερη και γρηγορότερη ανίχνευση του ιδίου ή ακόμα και των μερών του, παρά το ίδιο ερεθισμα με ανακριβείς χωρικές σχέσεις (π.χ. μια εικόνα παραλίας και το ασαφές αντίστοιχό της) [16], [32], [56]. Αφ' ετέρου, οι σχετικές σχέσεις φωτεινότητας ή χρώματος μεταξύ των περιοχών της εικόνας είναι σημαντικές για τις ταξινομήσεις των σκηνών με βάση την κοινή αντίληψη, ακριβώς όπως οι σχετικές χωρικές σχέσεις επηρεάζουν την ίδια διαδικασία.

Στην περίπτωση της κατηγοριοποίησης σκηνής, για παράδειγμα, όπου η πληροφορία δεν είναι διαθέσιμη με τη μορφή αντικειμένων, αλλά με τη μορφή περιοχών, μια “από επάνω προς τα κάτω” τεχνική είναι απαραίτητη. Κατά την υλοποίηση της ταξινόμησης των εικόνων ή των ακολουθιών βίντεο με βάση το περιεχόμενό τους η πληροφορία εννοιολογικού πλαισίου με τη μορφή του χωρικού σχεδιαγράμματος των περιοχών σε μια εικόνα μπορεί να αποτελέσει ένα συγκριτικό πλεονέκτημα. Παραδείγματος χάριν, μια κατηγορία εικόνων που αναπαριστούν μια ηλιόλουστη παραλία μπορεί να περιγραφεί ως μία εικόνα που διαθέτει κατά κύριο λόγο τρεις εμφανείς περιοχές: (α) μια “υπλε” περιοχή που αντιπροσωπεύει τη θάλασσα, (β) μια “υπεζ” περιοχή που αντιπροσωπεύει την άμμο και (γ) μια “γαλάζια” περιοχή που αντιπροσωπεύει τον ασυννέφιαστο ουρανό. Σε όλες τις περιπτώσεις και οι δύο περιοχές (α) και (β) θα είναι πάντα κάτω από την περιοχή (γ), ενώ σε μερικές περιπτώσεις η περιοχή (β) θα μπορούσε να βρεθεί κάτω από την περιοχή (α) ή η περιοχή (α) θα μπορούσε να βρεθεί κάτω από την περιοχή (β), δηλ. οι περιοχές (α) και (β) μπορούν να ανταλλάξουν

## Κεφάλαιο 6. Εννοιολογικό πλαίσιο και οπτική πληροφορία

τις χωρικές θέσεις τους σε ορισμένα πλάνα σύμφωνα με την εκάστοτε χρησιμοποιούμενη προοπτική (βλ. Σχήμα 6.4). Το παραπάνω παράδειγμα καταδεικνύει ότι η επιθυμητή ταξινόμηση μιας σκηνής μπορεί να παραμείνει έγκυρη εφόσον οι σχετικές και βασισμένες στο εννοιολογικό πλαίσιο χωρικές διαπεριφερειακές σχέσεις μεταξύ των περιοχών της εικόνας παραμένουν οι ίδιες. Το τελευταίο ισχύει, ακόμα κι αν οι απόλυτες τιμές των περιοχών υπόκεινται σε αλλαγές. Εν γένει, μπορούν να καθοριστούν πολυάριθμες σχέσεις μεταξύ δύο περιοχών σε μία σκηνή με τον ίδιο τρόπο, όπως μεταξύ δύο αντικειμένων.



**Σχήμα 6.4:** Παράδειγμα χωρικών σχέσεων περιοχών μιας εικόνας παραλίας - (α) - (γ) Παραδείγματα εικόνων που αναπαριστούν μια παραλία. (δ) Υπερτοπικές χωρο-εννοιολογικές σχέσεις.

Η επιπλέον πληροφορία που εισάγει και αξιοποιεί το οπτικό εννοιολογικό πλαίσιο περιορίζει το τι περιμένει κανείς και το πού πρέπει να κοιτάξει (δηλ. την οπτική προσοχή), διευκολύνοντας έτσι την αναζήτηση και την αναγνώριση των αντικειμένων που βρίσκονται ενσωματωμένα σε σύνθετες εικόνες ή ακολουθίες βίντεο. Σε ένα πραγματικό σενάριο η διαδικασία της ενοποίησης της πληροφορίας οπτικού εννοιολογικού πλαισίου σε έναν αλγόριθμο αναγνώρισης μπορεί να υλοποιηθεί με έναν τέτοιο τρόπο, ώστε να παρέχει αλληλεπίδραση μεταξύ της υψηλού επιπέδου γνώσης (εννοιολογικό πλαίσιο) και των χαμηλοτέρου επιπέδου πληροφοριών (όπως π.χ. το χρώμα, η υφή, η κίνηση κ.λ.π.). Η χρήση του εννοιολογικού πλαισίου μπορεί να βοηθήσει σημαντικά την επέκταση της οπτικής προσοχής, ενώ και η αντίστροφη διαδικασία μπορεί να παρέχει μια πληρέστερη περιγραφή της σκηνής. Για παράδειγμα, αν αναλύουμε ακολουθίες βίντεο αθλημάτων, όπου συμμετέχουν πάντα δύο παίκτες, όπως η αντισφαίριση, η επιτραπέζια αντισφαίριση κ.λ.π., αναμένουμε ότι η χαμηλού επιπέδου επεξεργασία θα παρέχει τελικά δύο περιοχές παικτών. Εάν επιστρέφεται μόνο μια περιοχή παικτών, τότε ο αλγόριθμος πρέπει να αναγκαστεί να παρέχει και την περιοχή ενός άλλου παικτη που βρίσκεται πιθανώς στην αντίθετη πλευρά από την προηγούμενη. Εντούτοις, μετά από την ανίχνευση των παικτών το σύστημα θα πρέπει να επέλθει σε μια καταλληλότερη περιγραφή εννοιολογικού πλαισίου, η οποία θα ταιριάζει καλύτερα στο συγκεκριμένο άθλημα, έτσι ώστε να παραχθεί μια ολοκληρωμένη αναπαράσταση της σκηνής που του δόθηκε ως είσοδος. Ως εκ τούτου, εάν η χαμηλού επιπέδου υπομονάδα παρέχει μια πρόσθετη περιοχή που αντιστοιχεί π.χ. σε ένα καφέ αγωνιστικό χώρο, τότε το σύστημα θα πρέπει να αλλάξει προς την κατεύθυνση μιας θεματικής περιοχής σχετικής με την αντισφαίριση και όχι προς την κατεύθυνση μιας θεματικής περιοχής σχετικής με την επιτραπέζια αντισφαίριση.

Σήμερα, υπάρχουν κάποια λίγα τέτοια συστήματα που συνδυάζουν τις οπτικές πληροφορίες εννοιολογικού πλαισίου με τη χαμηλού επιπέδου επεξεργασία. Εντούτοις, υπάρχουν και αρκετές σχετικά παρόμοιες μέθοδοι στην πρόσφατη βιβλιογραφία

της γνωστικής όρασης (computer vision), οι οποίες συνδυάζουν προγενέστερη γνώση (δηλ. προσδοκίες) με τρέχουσες εκτιμήσεις (δηλ. την παραγωγή της χαμηλού επιπέδου επεξεργασίας). Ο Hamker [105], [106], [107] για παράδειγμα, συνδυάζει την προγενέστερη γνώση με μια παρατήρηση που λαμβάνεται από την εικόνα από ένα συμπέρασμα βασισμένο στον πληθυσμό, προκειμένου να ενημερώσει δυναμικά τα ευδιάκριτα χαρακτηριστικά γνωρίσματα. Αυτά τα χαρακτηριστικά γνωρίσματα περιλαμβάνουν το χρώμα, τον προσανατολισμό, την ένταση χ.λ.π.. Οποιαδήποτε απόφαση, όπως π.χ. η ανίχνευση ενός αντικειμένου, βασίζεται σε αυτά τα κατανευμημένα ευδιάκριτα γνωρίσματα. Συχνά, στα πραγματικά σενάρια δεν υπάρχει καμία προγενέστερη πληροφορία για τη σκηνή και η μοντελοποίηση ενός σχετικού εννοιολογικού πλαισίου μπορεί να φαίνεται αδύνατη. Εντούτοις, οι ερευνητές προσπαθούν να αντισταθμίσουν αυτήν την έλλειψη γνώσης με την εστίαση στις τεχνικές μάθησης χωρίς επίβλεψη που παρουσιάσαμε συνοπτικά στο κεφάλαιο 4. Για παράδειγμα, ο Walther [271], [209] προτείνει μια τεχνική βασισμένη στην οπτική προσοχή για την εκμάθηση αναπαραστάσεων πολλαπλών αντικειμένων από εικόνες χωρίς πληροφορία ετικετών. Διαλέγει τις εμφανείς περιοχές, που είναι πιθανό να περιέχουν τα αντικείμενα, βασισμένα στην “από κάτω προς τα επάνω” οπτική προσοχή και δείχνει ότι η απόδοση της εκμάθησης και της αναγνώρισης μπορεί να βελτιωθεί ακόμη και παρουσία μεγάλων ποσοτήτων θορύβου.

Συνοψίζοντας, το οπτικό εννοιολογικό πλαίσιο είναι παρόν σχεδόν σε όλες τις πράξεις της καθημερινής μας ζωής και αντίληψης. Τα αντικείμενα και τα γεγονότα σπανίως εμφανίζονται απομονωμένα και συνήθως εμφανίζονται μέσα σε ένα πλούσιο, δομημένο εννοιολογικό πλαίσιο που πηγάζει από άλλες οπτικές πληροφορίες. Αν και εκ πρώτης όψεως το εννοιολογικό πλαίσιο είναι μια από τις αιτίες για το φαινόμενο της υπερφόρτωσης πληροφορίας, μπορεί ταυτόχρονα να υπαγορεύσει και το ποια αντικείμενα πρέπει να παρατηρηθούν και τι μπορεί ακίνδυνα να αγνοηθεί, κάτι το οποίο δεν παύει από το να αποτελεί έναν από τους κύριους στόχους των υπολογιστικών συστημάτων και των μοντέλων οπτικής προσοχής και ανάλυσης εικόνας. Προφανώς, μια κατάλληλη αλληλεπίδραση μεταξύ των τελευταίων και του εννοιολογικού πλαισίου μπορεί να οδηγήσει σε επιτυχείς αναγνωρίσεις τόσο των αντικειμένων όσο και των γεγονότων, αλλά και να συμβάλλει τα μάλα στο ευρύτερο ερευνητικό πεδίο της ανάλυσης και επεξεργασίας πολυμέσων.

### 6.1.2 Δομή κεφαλαίου

Το παρόν κεφάλαιο επιχειρεί να παρουσιάσει το γενικότερο πλαίσιο του εννοιολογικού πλαισίου, το οποίο βασίζεται στην οπτική πληροφορία μιας εικόνας, και να εστιάσει σε συγκεκριμένες υλοποιήσεις αναφορικά με το ευρέως διαδεδομένο πρόβλημα της ανίχνευσης αντικειμένων και υψηλού επιπέδου εννοιών. Χωρίζεται σε ορισμένες βασικές δομικές ενότητες:

- **Εννοιολογική αναπαράσταση γνώσης.** Η ενότητα 6.2 παρουσιάζει την απαραίτητη υποδομή γνώσης και τη μοντελοποίηση του εννοιολογικού πλαισίου, προκειμένου να πραγματοποιηθεί μία αποδοτική και σωστή χρήση και εκμετάλλευσή του κατά τη διαδικασία της πολυμεσικής ανάλυσης. Χρησιμοποιούνται βασικές αρχές των οντολογιών κατάλληλα προσαρμοσμένες στο πεδίο της ανάλυσης, ενώ λαμβάνονται υπόψη και οι πιο πρόσφατες εξελίξεις για την αναπαράσταση γνώσης από το χώρο του Σημασιολογικού Ιστού.

- Ανάλυση εικόνων με χρήση του εννοιολογικού πλαισίου. Στην ενότητα 6.3 παρουσιάζεται η έννοια της “σχετικότητας πλαισίου” (context relevance), ένας αναδρομικός αλγόριθμος για τον επαναπροσδιορισμό του βαθμού ιδιότητας μέλους (ή εμπιστοσύνης) των αποτελεσμάτων της υποβοηθούμενης από τη γνώση πολυμεσικής ανάλυσης, καθώς και το επεκτεταμένο μοντέλο του εννοιολογικού πλαισίου στην περίπτωση πολλαπλών θεματικών περιοχών. Ο αλγόριθμος στηρίζεται στην προτεινόμενη θεώρηση του εννοιολογικού πλαισίου και της σχετικής εννοιολογικής γνώσης, καθώς και στην έννοια της σχετικότητας του εννοιολογικού πλαισίου.
- Οπτικό εννοιολογικό πλαίσιο και σημασιολογική κατάτμηση εικόνων. Στην ενότητα 6.4 του κεφαλαίου πραγματοποιείται μία επισκόπηση/παρουσίαση μιας εκ των δύο κυριότερων εφαρμογών της προτεινόμενης προσέγγισης του οπτικού εννοιολογικού πλαισίου σε επίπεδο επεξεργασίας των περιοχών μιας εικόνας. Ειδικότερα, γίνεται αναφορά στην επίδραση που ασκεί το εννοιολογικό πλαίσιο, για να βελτιώσει τη σημασιολογική κατάτμηση εικόνων. Παρουσιάζονται συνοπτικά τα βασικά χαρακτηριστικά των επιμέρους πρακτικών εφαρμογών, καθώς και κάποια ενδεικτικά ερευνητικά αποτελέσματα που καταδεικνύουν τη χρησιμότητα της προτεινόμενης μεθόδου.
- Οπτικό εννοιολογικό πλαίσιο και γενετικός αλγόριθμος. Στην ενότητα 6.5 συνδυάζουμε την προτεινόμενη θεώρηση που εισάγει το οπτικό εννοιολογικό πλαίσιο με αυτή ενός γενετικού αλγορίθμου προς την επίτευξη αρτιότερων αποτελεσμάτων σημασιολογικής ανάλυσης, ήτοι αποδοτικότερης ανίχνευσης υψηλού επιπέδου εννοιών στην εικόνα. Και σε αυτή την περίπτωση παρουσιάζονται συνοπτικά τα βασικά χαρακτηριστικά της εφαρμογής, καθώς και ορισμένα σχετικά πειραματικά αποτελέσματα που καταδεικνύουν το ρόλο και τη χρησιμότητα του εννοιολογικού πλαισίου στη διαδικασία.

## 6.2 Εννοιολογική Αναπαράσταση Γνώσης

Ανάμεσα σε όλους τους πιθανούς φορμαλισμούς, που μπορούν να χρησιμοποιηθούν για την αναπαράσταση της γνώσης και κατ’ επέκταση του ίδιου του (οπτικού) εννοιολογικού πλαισίου, οι οντολογίες [104], [163], [243] ξεχωρίζουν, μιας και εμφανίζουν ένα σημαντικό αριθμό από πλεονεκτήματα. Στο πλαίσιο αυτής της εργασίας, αλλά και γενικότερα οι οντολογίες είναι κατάλληλες, για να εκφράσουν με μία αυστηρή, επίσημη και συνάμα επεξεργάσιμη από υπολογιστές αναπαράσταση τη σημασιολογία του πολυμεσικού περιεχομένου, η οποία επιτρέπει τη χειρωνακτική ή αυτόματη ανάλυση και περαιτέρω επεξεργασία των εξαγομένων σημασιολογικών περιγραφών. Καθώς μια οντολογία μπορεί να αποτελέσει “μια επίσημη προδιαγραφή της κοινής κατανόησης μιας επιλεγμένης θεματικής περιοχής” [104], αυτή η επίσημη προδιαγραφή υλοποιείται συνήθως χρησιμοποιώντας μια ιεραρχία υποκατηγοριών, η οποία διέπεται από συγκεκριμένες σημασιολογικές σχέσεις μεταξύ των κατηγοριών, καθώς και σύνθετες περιγραφές των τελευταίων (π.χ. σε DL [12] ή σε OWL [319]). Ανάμεσα, λοιπόν, σε όλους τους πιθανούς τρόπους περιγραφής των οντολογιών επιλέγουμε τον επόμενο, προκειμένου να μοντελοποιήσουμε αποδοτικά και συνάμα απλά μια οντολογία:

$$O = \{C, \{R_{pq}\}\}, \quad \text{που} \quad R_{pq} : C \times C \rightarrow \{0, 1\} \quad (6.1)$$

Στην εξίσωση 6.1, το  $O$  αναπαριστά μια οντολογία, το  $C$  είναι το σύνολο των εννοιών που περιγράφονται από αυτή, το  $p$  και το  $q$  είναι δύο οποιεσδήποτε έννοιες  $p, q \in C$  και  $R_{pq}$  είναι η σημασιολογική σχέση μεταξύ αυτών των δύο εννοιών. Το προτεινόμενο μοντέλο γνώσης είναι βασισμένο σε ένα σύνολο εννοιών και σχέσεων μεταξύ τους, οι οποίες και διαμορφώνουν τα βασικά δομικά στοιχεία προς την κατεύθυνση της σημασιολογικής ερμηνείας της τελευταίας στο πλαίσιο της παρούσας ερευνητικής προσπάθειας. Γενικά, οι σημασιολογικές σχέσεις περιγράφουν συγκεκριμένα είδη συνδέσεων ή σχέσεων μεταξύ οποιωνδήποτε δύο εννοιών. Στην περίπτωση των κλασικών (σαφών) συνόλων, μια σημασιολογική σχέση είτε συσχετίζει ( $R_{pq} = 1$ ), είτε δε συσχετίζει ( $R_{pq} = 0$ ) ένα ζευγάρι εννοιών  $p, q$  μεταξύ τους. Αν και σχεδόν οποιοδήποτε είδος σχέσης μπορεί να ληφθεί υπόψη για την κατασκευή μιας τέτοιας αναπαράστασης γνώσης, οι δύο κατηγορίες που χρησιμοποιούνται συνήθως είναι οι ταξινομικές (ή αλλιώς διατακτικές) σχέσεις και οι σχέσεις συμβατότητας (ή αλλιώς συμμετρικές). Κατά συνέπεια, μια πρώτη ερευνητική πρόκληση αποτελεί η κατάλληλη χρησιμοποίηση των πληροφοριών που περιλαμβάνονται στις σχέσεις αυτές με στόχο την εκμετάλλευση του εννοιολογικού πλαισίου κατά την ανάλυση εικόνων και/ή βίντεο.

Επιπλέον, για να είναι ένα μοντέλο γνώσης ιδιαίτερα περιγραφικό, πρέπει να περιέχει έναν μεγάλο αριθμό ευδιάκριτων και διαφορετικών σχέσεων μεταξύ των εννοιών. Μια σημαντική παρενέργεια αυτής της προσέγγισης είναι το γεγονός ότι σε αυτή την περίπτωση θα διασκορπιστούν οι διαθέσιμες πληροφορίες, γεγονός που καθιστά κάθεμία από αυτές τις σχέσεις ανεπαρκή να περιγράψει ένα εννοιολογικό πλαίσιο με έναν ουσιαστικό τρόπο. Συνεπώς, οι σχέσεις αυτές πρέπει να συνδυαστούν, για να παρέχουν τη μορφή της γνώσης που απαιτείται για τον καθορισμό και την εκτίμηση του οπτικού εννοιολογικού πλαισίου. Σε αυτό το τμήμα της παρούσας εργασίας χρησιμοποιούμε κυρίως τρεις τύπους σχέσεων, των οποίων η σημασιολογία καθορίζεται στο ίδιο το πρότυπο του MPEG-7 ([296], [23], Παράρτημα B): τη σχέση *Ειδίκευσης Sp*, τη σχέση *Μέρους P* και τη σχέση *Ιδιοκτησίας Pr*.

Ένα τελευταίο σημείο προς εξέταση κατά το σχεδιασμό ενός τέτοιου μοντέλου γνώσης είναι η διάχυτη στο πλαίσιο της διατριβής διαπίστωση ότι τα δεδομένα της πραγματικής ζωής διαφέρουν συχνά από τα ερευνητικά δεδομένα. Οι πραγματικές οπτικές πληροφορίες κυριαρχούνται, εν γένει, από αβεβαιότητα και ασάφεια, κατά συνέπεια η μοντελοποίησή τους πρέπει να είναι βασισμένη στις ασαφείς σχέσεις. Για τα προβλήματα που καλούμαστε να αντιμετωπίσουμε στο πεδίο της επεξεργασίας και ανάλυσης εικόνας, οι ανωτέρω, ευρέως απαντούμενες, σαφείς σχέσεις μπορούν να μοντελοποιηθούν ως ασαφείς διατακτικές σχέσεις και μπορούν να συνδυαστούν για την παραγωγή μιας ασαφούς ταξινομικής σχέσης, που να έχει κάποιο ουσιαστικό νόημα. Συνεπώς, προκειμένου να αντιμετωπίσουμε τέτοιους σύνθετους τύπους σχέσεων, προτείνουμε μια “ασαφοποίηση” του παραπάνω γενικού ορισμού μιας οντολογίας (6.1) ως εξής:

$$O_F = \{C, \{r_{pq}\}\}, \quad \text{οπου} \quad r_{pq} = F(R_{pq}) : C \times C \rightarrow [0, 1] \quad (6.2)$$

Στην εξίσωση (6.2) το  $O_F$  καθορίζει μια “ασαφοποιημένη” οντολογία, το  $C$  είναι πάλι το σύνολο όλων των πιθανών εννοιών που αυτή περιγράφει και το  $r_{pq}$  υποδεικνύει μια ασαφή σχέση μεταξύ των δύο εννοιών  $p, q \in C$ . Στην ασαφή περίπτωση, μια ασαφής σημασιολογική σχέση συσχετίζει ένα ζευγάρι εννοιών  $p, q$  μεταξύ τους με έναν δεδομένο βαθμό ιδιότητας μέλους, δηλ. η τιμή του  $r_{pq}$  βρίσκεται μέσα στο διάστημα  $[0, 1]$ . Πιο συγκεκριμένα και όπως είδαμε και στην ενότητα 3.2, με δεδομένο έναν κόσμο  $U$ , ένα κλασικό (σαφές) σύνολο  $C$  περιγράφεται από μια συνάρτηση ιδιότητας μέλους  $\mu_C : U \rightarrow \{0, 1\}$  (όπως παρατηρείται ήδη στη σαφή περίπτωση για την  $R_{pq}$ ),

ενώ σύμφωνα με τον Klir [139], ένα ασαφές σύνολο  $F$  στο  $C$  περιγράφεται από μια συνάρτηση ιδιότητας μέλους  $\mu_F : C \rightarrow [0, 1]$ . Μπορούμε, ωστόσο, να περιγράψουμε το ασαφές σύνολο, χρησιμοποιώντας τον ευρέως εφαρμοσμένο μαθηματικό συμβολισμό αθροίσματος [177]:

$$F = \sum_{i=1}^n c_i / w_i = \{c_1/w_1, c_2/w_2, \dots, c_n/w_n\} \quad (6.3)$$

όπου  $n = |C|$  είναι η πληθυκότητα (δηλ. ο αριθμός των στοιχείων) του συνόλου  $C$  και η έννοια  $c_i \in C$ . Ο βαθμός ιδιότητας μέλους  $w_i$  περιγράφει τη συνάρτηση ιδιότητας μέλους  $\mu_F(c_i)$ , δηλ.  $w_i = \mu_F(c_i)$ , ή χάριν απλότητας  $w_i = F(c_i)$ . Όπως ακριβώς στο [139], μια ασαφής σχέση στο  $C$  είναι μια συνάρτηση  $r_{pq} : C \times C \rightarrow [0, 1]$  και η αντίστροφη σχέση της ορίζεται ως  $r_{pq}^{-1} = r_{qp}$ . Με βάση τις σχέσεις  $r_{pq}$  και σε πλήρη αναλογία με τα όσα περιγράψαμε αναλυτικά στο κεφάλαιο 5 (αυτή τη φορά, όμως, με απώτερο σκοπό την ανάλυση εικόνων), κατασκευάζουμε την ακόλουθη σχέση  $T$  με τη χρήση ενός συνόλου ασαφών ταξινομικών σχέσεων, το οποίο σε πρώτη φάση περιορίζεται στις σχέσεις  $Sp$ ,  $P$  και  $Pr$ :

$$T = Tr^t(Sp \cup P^{-1} \cup Pr^{-1}) \quad (6.4)$$

Στις προαναφερθείσες σχέσεις η ασάφεια έχει την ακόλουθη έννοια: υψηλές τιμές της  $Sp(p, q)$  υπονοούν ότι η έννοια του  $q$  προσεγγίζει την έννοια του  $p$ , από την άποψη ότι, όταν μια εικόνα σχετίζεται σημασιολογικά με το  $q$ , τότε πιθανότατα συσχετίζεται και με το  $p$ . Από την άλλη, καθώς μειώνεται το  $Sp(p, q)$ , η έννοια του  $q$  γίνεται “στενότερη” από αυτή του  $p$ , υπό την έννοια ότι η συσχέτιση μιας εικόνας με το  $q$  δεν υπονοεί μια ταυτόχρονη συσχέτιση και με το  $p$  με μια επίσης μεγάλη πιθανότητα ή με έναν υψηλό βαθμό. Με ανάλογο τρόπο οι βαθμοί των άλλων δύο σχέσεων  $P$  και  $Pr$  μπορούν να ερμηνευθούν ως δεσμευμένες πιθανότητες ή ως βαθμοί υπονοούμενης σχετικότητας. Τα MPEG-7 MDS [296], που παρουσιάσαμε αναλυτικά στην ενότητα 3.6, καθορίζουν όλους τους τύπους των σημασιολογικών σχέσεων, ενώ παρέχουν και τους ορισμούς όλων των αντίστροφών τους, όπου, φυσικά, ορίζονται τα τελευταία. Μερικές φορές, η σημασιολογική ερμηνεία μιας σχέσης δε βγάζει κάποιο νόημα για ένα ζευγάρι εννοιών, σε αντίθεση με το αντίστροφό του, το οποίο μπορεί να στέκει νοηματικά. Στην περίπτωσή μας, η σχέση Μέρους  $P(p, q)$  ορίζεται ως:  $p$  μέρος  $q$  εάν και μόνο εάν το  $q$  είναι μέρος του  $p$ . Παραδείγματος χάριν, έστω ότι το  $p$  είναι η Νέα Υόρκη και το  $q$  το Μανχάταν. Είναι προφανές ότι μόνο η αντίστροφη σχέση μέρος του,  $P^{-1}$ , έχει νόημα σημασιολογικά, δεδομένου ότι το Μανχάταν είναι μέρος της Νέας Υόρκης και ουχί το ανάποδο. Ομοίως και για τη σχέση ιδιοκτησίας  $Pr$  επιλέγεται το αντίστροφό της. Από την άλλη, μετά από τον παραπάνω ορισμό της σχέσης Ειδίκευσης  $Sp(p, q)$ , το  $p$  είναι μια ειδίκευση του  $q$ , εάν και μόνο εάν το  $q$  είναι μια ειδίκευση της έννοιας του  $p$ . Για παράδειγμα, έστω ότι το  $p$  είναι ένα θηλαστικό και το  $q$  είναι ένας σκύλος: το  $Sp(p, q)$  σημαίνει ότι ο σκύλος είναι μία ειδίκευση ενός θηλαστικού, κάτι το οποίο εκφράζει επακριβώς τη σημασιολογική ερμηνεία που επιθυμούμε να χρησιμοποιήσουμε (και προφανώς όχι το αντίστροφό του). Με βάση τους ρόλους και τις σημασιολογικές ερμηνείες των  $Sp$ ,  $P$  και  $Pr$  είναι εύκολο να δει κανείς ότι η εξίσωση (6.4) συνδυάζει κατάλληλα όλα τα παραπάνω και χρησιμοποιεί την αντίστροφη λειτουργία των σχέσεων, όπου αυτό είναι σημασιολογικά αρμόζον, δηλ. όπου η έννοια μιας σχέσης είναι σημασιολογικά αντιφατική προς την έννοια των υπολοίπων σχέσεων πάνω στο ίδιο σύνολο εννοιών. Στο σημείο αυτό θα πρέπει να σημειωθεί ότι

το μεταβατικό κλείσιμο  $Tr^t$  (βλ. ενότητα 3.3) απαιτείται για να παραμείνει η σχέση  $T$  ταξινομική, καθώς η ένωση διαφόρων μεταβατικών σχέσεων δεν παραμένει απαραίτητως μεταβατική, όπως άλλωστε καταδεικνύεται και σε προγενέστερες ερευνητικές εργασίες [3].

Η αναπαράσταση, τώρα, του παραπάνω εννοιο-κεντρικού μοντέλου γνώσης για το οπτικό εννοιολογικό πλαίσιο ακολουθεί το πρότυπο της γλώσσας RDF [316], το οποίο και έχει προταθεί στο πλαίσιο του Σημασιολογικού Ιστού. Η RDF αποτελεί το πλαίσιο στο οποίο εκφράζονται οι δηλώσεις μεταδεδομένων του Σημασιολογικού Ιστού και οι οποίες αντιπροσωπεύονται συνήθως ως γράφοι. Το μοντέλο της RDF είναι βασισμένο στην ιδέα της παραγωγής δηλώσεων για τους πόρους (resources) με τη μορφή της ακόλουθης έκφρασης: **υποκείμενο-κατηγόρημα-αντικείμενο**. τα κατηγορήματα είναι γνωρίσματα ή πτυχές ενός πόρου και εκφράζουν μια σχέση μεταξύ του υποκειμένου και του αντικειμένου. Η σχέση  $T$  μπορεί να αναπαρασταθεί ως γράφος, στον οποίο κάθε κόμβος αντιπροσωπεύει μια έννοια και κάθε ακμή μεταξύ δύο κόμβων αποτελεί μια βασισμένη στο εννοιολογικό πλαίσιο σχέση μεταξύ των αντίστοιχων εννοιών. Επιπρόσθετα, κάθε ακμή φέρει και έναν σχετικό βαθμό ιδιότητας μέλους, ο οποίος αντιπροσωπεύει την ασάφεια του μοντέλου του εννοιολογικού πλαισίου. Η αναπαράσταση του γράφου αυτού σε RDF αποτελεί έναν εφικτό στόχο, δεδομένου ότι η ίδια η δομή της RDF βασίζεται σε ένα παρόμοιο μοντέλο γράφου, όπως παρουσιάστηκε στην υποενότητα 3.4.3.

Στην παρούσα προσέγγιση χρησιμοποιήθηκε ευρέως η τεχνική του reification [318], προκειμένου να επιτευχθεί η επιθυμητή εκφραστικότητα και να εκμεταλλευτούμε την ενισχυμένη λειτουργικότητα που εισήχθη με τη χρήση της ασάφειας. Η αναπαράσταση του βαθμού ιδιότητας μέλους που συνδέεται με κάθε σχέση πραγματοποιείται χρησιμοποιώντας την τεχνική αυτή, δηλ. πραγματοποιώντας μια δήλωση για την εκάστοτε δήλωση, η οποία και περιέχει την επιθυμητή πληροφορία του βαθμού. Η αναπαράσταση της ασάφειας με τέτοιους είδους δηλώσεις αποτελεί έναν νέο, καινοτόμο αλλά ταυτόχρονα πλήρως αποδεκτό τρόπο, δεδομένου ότι μια τέτοια δήλωση δεν ισχύει ούτε βεβαιώνεται αυτόματα. Παραδείγματος χάριν, θεωρώντας μια δήλωση: “Ουρανός Μέρος Της Σκηνής Παραλίας” και έναν βαθμό ιδιότητας μέλους 0.75 για αυτή την δήλωση, προφανώς δε συνεπάγεται ότι ο ουρανός είναι πάντα μέρος μιας σκηνής παραλίας. Ένα μικρό επεξηγηματικό παράδειγμα για ένα στιγμιότυπο της σχέσης *Eidίκευσης Sp* παρουσιάζεται στο Σχήμα 6.5:

```
<rdf:Description rdf:about="#s1">
  <rdf:subject rdf:resource="&dom;wrc"/>
  <rdf:predicate rdf:resource="&dom;specializationOf"/>
  <rdf:object><rdf:resource="&dom;rally"></rdf:object>
  <rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Statement"/>
  <context:specializationOf rdf:datatype="http://www.w3.org/2001/XMLSchema#float">
    0.90</context:specializationOf>
</rdf:Description>
```

Σχήμα 6.5: Αναπαράσταση μιας ασαφούς σχέσης σε RDF - χρήση της τεχνικής reification

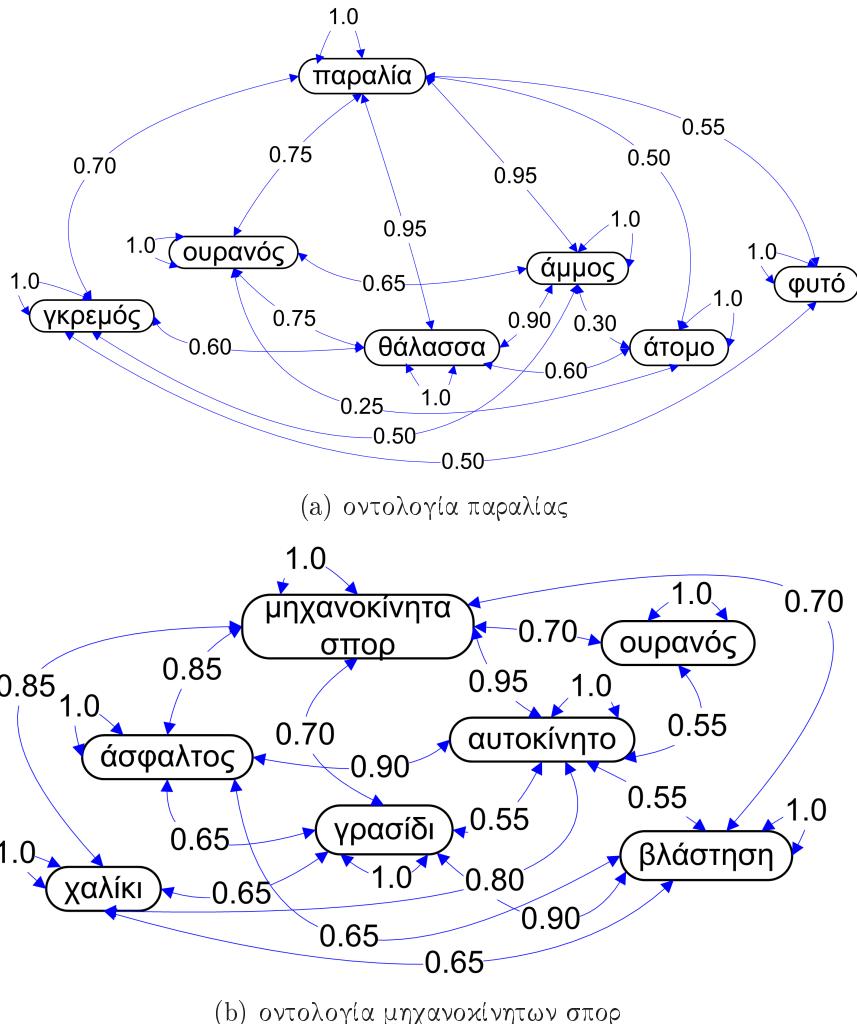
Όπως έχει ήδη συζητηθεί, το  $Sp(x, y) > 0$  σημαίνει ότι η έννοια  $x$  “περιλαμβάνει” την έννοια  $y$ . Οι πιο συνήθεις μορφές ειδίκευσης είναι η υποκατηγοριοποίηση, δηλ. το  $x$  είναι μια γενίκευση του  $y$  και η θεματική κατηγοριοποίηση, δηλ. το  $x$  είναι η θεματική κατηγορία του  $y$ . Στο παραπάνω παράδειγμα, το RDF υποκείμενο *wrc* (Πα-

γκόσμιο Πρωτάθλημα Ράλλυ - World Rally Championship) έχει την specializationOf ως RDF κατηγόρημα και το ράλλυ αποτελεί το RDF αντικείμενο. Επιπλέον, η προτεινόμενη διαδικασία του reification εισάγει μια δήλωση για την προηγούμενη δήλωση στον πόρο specializationOf με το να δηλώνει ότι το 0.90 είναι ο βαθμός ιδιότητας μέλους αυτής της σχέσης.

Η βασισμένη στο εννοιολογικό πλαίσιο γνώση που χρησιμοποιείται στις σειρές πειραμάτων που θα παρουσιαστούν στις επόμενες ενότητες της παρούσης βασίζεται εξ ολοκλήρου στην βασισμένη στις οντολογίες αναπαράσταση γνώσης του εννοιολογικού πλαισίου που περιγράφηκε σε αυτή την ενότητα. Ειδικότερα και σύμφωνα με την προτεινόμενη προσέγγιση αναπτύχθηκαν συγκεκριμένες οντολογίες θεματικών περιοχών για την αναπαράσταση των τυμάτων εκείνων της γνώσης, που απαιτούν τον ρητό καθορισμό τους. Παραδείγματα τέτοιων οντολογιών παρουσιάζονται στο Σχήμα 6.6. Αυτά περιλαμβάνουν τις σημασιολογικές έννοιες που θεωρούνται ενδιαφέρουσες για την κάθε μία υπό εξέταση θεματική περιοχή (π.χ. για τη θεματική περιοχή της παραλίας: θάλασσα, άμμος, ουρανός, πρόσωπο, φυτό, βράχος ή για τη θεματική περιοχή μηχανοκίνητα σπορ: αυτοκίνητο, άσφαλτος, αυμοχάλικο, χλόη, βλάστηση, ουρανός), καθώς επίσης και τις σημασιολογικές σχέσεις που τις διασυνδέουν από την άποψη των βαθμών ιδιότητας μέλους, που αντιστοιχούν σε καθεμία εκ των τριών ασαφών σημασιολογικών σχέσεων που χρησιμοποιούνται, δηλ. Sp, P και Pr. Σε αντιδιαστολή με τις ίδιες τις έννοιες, που καθορίζονται χειροκίνητα από εμπειρογνώμονες των θεματικών περιοχών, οι βαθμοί ιδιότητας μέλους για κάθε ζευγάρι των υπό εξέταση εννοιών (οι οποίοι απαιτούνται για την κατασκευή των σχέσεων κατά τη διάρκεια της διαδικασίας κατασκευής της οντολογίας), εξάγονται στο πλαίσιο της παρούσης χρησιμοποιώντας ένα σύνολο από εικόνες, που λειτουργεί ως “σύνολο εκπαίδευσης”, σε συνδυασμό με κατάλληλα πιθανοτικά στατιστικά μοντέλα πληροφοριών που εφαρμόζονται στο τελευταίο. Τέλος, θα πρέπει να σημειωθεί ότι οι εκάστοτε θεματικές περιοχές εφαρμογής (παραλία, μηχανοκίνητα σπορ, βουνό, πόλη, τέννις, ποδόσφαιρο, κ.α.), επιλέχθηκαν με βάση τη δημοτικότητά τους και το μέγεθος του πολυμεσικού περιεχομένου που ήταν διαθέσιμο για καθεμία από αυτές κατά την εκτέλεση των πειραμάτων μας.

### 6.3 Ανάλυση Εικόνων με Χρήση του Εννοιολογικού Πλαισίου

Η ιδέα πίσω από τη χρήση των οπτικών πληροφοριών του εννοιολογικού πλαισίου ανταποκρίνεται στο γεγονός ότι δεν είναι όλες οι ανθρώπινες πράξεις σχετικές με όλες τις καταστάσεις και αυτό ισχύει επίσης και κατά την εξέταση των προβλημάτων που διέπουν την ανάλυση εικόνας. Δεδομένου ότι το οπτικό εννοιολογικό πλαίσιο είναι μια μάλλον δυσνόητη και ίσως ακόμα και απροσδιόριστη έννοια [182], την περιορίζουμε εν τω παρόντι στην έννοια του οντολογικού εννοιολογικού πλαισίου. Το τελευταίο ορίζεται ως τμήμα της “ασαφοποιημένης” έκδοσης των παραδοσιακών οντολογιών, όπως αυτές παρουσιάζονται στην ενότητα 6.2. Σε αυτή την ενότητα τα προβλήματα που εξετάζονται περιλαμβάνουν το πώς θα επαναπροσδιορίσουμε τους βαθμούς ιδιότητας μέλους (με τέτοιο τρόπο ώστε να έχουν νόημα) των ασαφών περιοχών μιας εικόνας και το πώς θα χρησιμοποιήσουμε το οπτικό εννοιολογικό πλαίσιο, για να επηρεάσουμε τα συνολικά αποτελέσματα της ανάλυσης, η οποία επωφελείται από τη γνώση προς μια καλύτερη απόδοση και αποδοτικότητα.



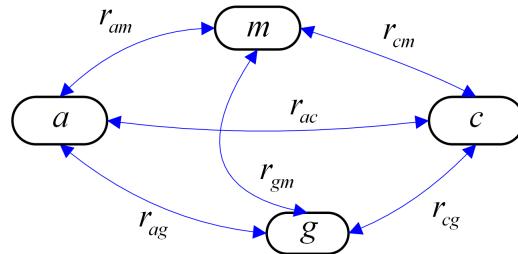
Σχήμα 6.6: Παραδείγματα οντολογιών θεματικών περιοχών.

### 6.3.1 Σχετικότητα εννοιολογικού πλαισίου

Με βάση το μαθηματικό υπόβαθρο που περιγράφεται λεπτομερώς στις προηγούμενες υποενότητες (βλ. π.χ. ενότητα 3.2), παρουσιάζουμε στη συνέχεια τον αλγόριθμο που χρησιμοποιείται, για να επαναπροσδιορίσει το βαθμό ιδιότητας μέλους  $\mu_a(c)$  κάθε έννοιας  $c$  στο ασαφές σύνολο  $L_a$  που συνδέεται με μια περιοχή μιας εικόνας. Κάθε συγκεκριμένη έννοια  $k \in C$ , που είναι παρούσα στην οντολογία της θεματικής περιοχής, αποθηκεύεται μαζί με τους βαθμούς σχέσης της  $r_{kl}$  με οποιαδήποτε άλλη έννοια  $l \in C$  με την οποία σχετίζεται. Για να αντιμετωπίσουμε τις περιπτώσεις, όπου περισσότερες από μια έννοιες συσχετίζονται με πολλαπλές έννοιες, εισάγουμε τον όρο της “σχετικότητας πλαισίου”  $cr_{dm}(k)$ , ο οποίος αναφέρεται στη γενική σχετικότητα της έννοιας  $k$  με το στοιχείο ρίζας που χαρακτηρίζει κάθε θεματική περιοχή  $dm$ . Παραδείγματος χάριν, το στοιχείο ρίζας της παραλίας και των μηχανοκίνητων σπορ είναι οι έννοιες παραλία και μηχανοκίνητα σπορ, αντίστοιχα. Η προτεινόμενη προσέγγιση αξιοποιεί ως ένα βαθμό την έννοια του μεταβατικού κλεισίματος, όπως αυτή παρουσιάστηκε στην ενότητα 3.3, μιας και λαμβάνονται υπόψη όλες οι πιθανές διαδρομές όλων των διαθέσιμων επιπέδων στο γράφο, κάτι που διαμορφώνει μια εξαντλητική προσέγγιση για τη θεματική περιοχή, ενώ όλες οι διαδρομές/διασυνδέσεις μεταξύ των έννοιών θεωρούνται και είναι αμοιβαίες.

Η εκτίμηση της τιμής κάθε έννοιας προέρχεται από τις άμεσες και τις έμμεσες σχέσεις της έννοιας με άλλες έννοιες, χρησιμοποιώντας έναν κατάλληλο δείκτη συμβατότητας ή με άλλα λόγια μια κατάλληλη μετρική απόστασης. Ανάλογα με τη φύση των θεματικών περιοχών υπό εξέταση ο καλύτερος δείκτης θα μπορούσε να επιλεχθεί χρησιμοποιώντας π.χ. τον τελεστή  $tax$ . Φυσικά, η ιδανική μετρική απόστασης για δύο έννοιες είναι αυτή που ποσοτικοποιεί το σημασιολογικό συσχετισμό τους. Για το υπό εξέταση πρόβλημα και λαμβάνοντας υπόψη τις θεματικές περιοχές παραλία και μηχανοκίνητα σπορ, το  $tax$  είναι όντως ένα κατάλληλο μέτρο του συσχετισμού και για τις δύο. Ένα απλουστευμένο παράδειγμα που υποθέτει ότι οι μόνες διαθέσιμες έννοιες είναι: μηχανοκίνητα σπορ (το στοιχείο  $r_1$ , που αναφέρεται ως  $m$ ), άσφαλτος ( $a$ ), χλόη ( $g$ ) και αυτοκίνητο( $c$ ), παρουσιάζεται στο ακόλουθο Σχήμα 6.7 και συνοψίζεται στα εξής: έστω ότι η έννοια  $a$  σχετίζεται άμεσα με τις έννοιες  $m$ ,  $g$  και  $c$  με τις:  $r_{am}$ ,  $r_{ag}$  και  $r_{ac}$ , ενώ η έννοια  $g$  σχετίζεται με την έννοια  $m$  με  $r_{gm}$  και η έννοια  $c$  σχετίζεται με την έννοια  $m$  με  $r_{cm}$ . Επιπρόσθετα, η  $c$  σχετίζεται με την  $g$  με  $r_{cg}$ . Τότε, υπολογίζουμε την τιμή της σχετικότητας πλαισίου για την έννοια άσφαλτος  $cr_{dm}(a)$  ως εξής:

$$cr_{dm}(a) = \max\{r_{am}, r_{ag}r_{gm}, r_{ac}r_{cm}, r_{ag}r_{cg}r_{cm}, r_{ac}r_{cg}r_{gm}\} \quad (6.5)$$



Σχήμα 6.7: Παράδειγμα αναπαράστασης γράφου - εκτίμηση δείκτη συμβατότητας.

### 6.3.2 Μεθοδολογία και αλγόριθμος ανάλυσης

Η γενική δομή του αλγορίθμου επαναπροσδιορισμού των βαθμών ιδιότητας μέλους συνοψίζεται στα ακόλουθα βήματα:

- Προσδιορίστε μια βέλτιστη παράμετρο κανονικοποίησης  $pr$ , που θα χρησιμοποιηθεί κατά τα βήματα εκτέλεσης του αλγορίθμου, κατάλληλη για την εκάστοτε εξεταζόμενη θεματική περιοχή. Η παράμετρος  $pr$  αναφέρεται επίσης και ως μέτρο ομοιότητας ή ανομοιότητας των θεματικών περιοχών και  $pr \in [0, 1]$ .
- Για κάθε έννοια  $k$  του ασαφούς συνόλου  $L_a$  μιας σκηνής της εικόνας, η οποία συνοδεύεται από έναν βαθμό ιδιότητας μέλους  $\mu_a(k)$ , εξάγετε την ιδιάτερη και βασισμένη στο εννοιολογικό πλαίσιο πληροφορία υπό την μορφή των σχέσεων της με το σύνολο οποιωνδήποτε άλλων εννοιών:  $\{r_{kl} : l \in C, l \neq k\}$ .
- Υπολογίστε το νέο βαθμό ιδιότητας μέλους  $\mu_a(k)$  της περιοχής  $a$  με βάση την  $pr$  και την σχετικότητα  $cr_k$  του εννοιολογικού πλαισίου. Στην περίπτωση της ύπαρξης πολλαπλών σχέσεων μιας έννοιας στην οντολογία, που συσχετίζουν την έννοια  $k$  με περισσότερες από μια έννοιες (αντί να την συσχετίζουν αποκλειστικά

με το “στοιχείο ρίζας”  $r^e$ ), θα πρέπει να εφαρμοστεί ένα ενδιάμεσο βήμα για την  $k$ , ήτοι:  $cr_k = \max\{r_{k^{re}}, \dots, r_{km}\}$ . Εκφράζουμε τον υπολογισμό του  $\mu_a(k)$  με τον αναδρομικό τύπο:

$$\mu_a^n(k) = \mu_a^{n-1}(k) - np(\mu_a^{n-1}(k) - cr_k) \quad (6.6)$$

όπου το  $n$  καταδεικνύει την εκάστοτε επανάληψη. Ισοδύναμα, για μία αυθαίρετη επανάληψη:

$$\mu_a^n(k) = (1 - np)^n \cdot \mu_a^0(k) + (1 - (1 - np)^n) \cdot cr_k \quad (6.7)$$

όπου το  $\mu_a^0(k)$  αντιπροσωπεύει τον αρχικό βαθμό ιδιότητας μέλους.

Στην πράξη οι χαρακτηριστικές τιμές του  $n$  βρίσκονται μεταξύ του 3 και του 5.

Η ερμηνεία των παραπάνω εξισώσεων (6.6) και (6.7) υπονοεί ότι η προτεινόμενη και βασισμένη στο εννοιολογικό πλαίσιο προσέγγιση θα ευνοήσει τους σχετικά βέβαιους βαθμούς ιδιότητας μέλους για την έννοια μιας περιοχής σε αντιδιαστολή με τους σχετικά αβέβαιους ή παραπλανητικούς βαθμούς. Θα ενισχύσει τις διαφορές τους, ενώ από την άλλη θα ελαττώσει την εμπιστοσύνη μας στις σαφώς παραπλανητικές έννοιες για μια συγκεκριμένη περιοχή. Περαιτέρω, με βάση την παρεχόμενη οντολογική γνώση θα διευκρινίσει και θα λύσει τις ασάφειες σε περιπτώσεις παρόμοιων εννοιών ή εννοιών που είναι αντικειμενικά δύσκολο να προκύψουν από τη χαμηλού επιπέδου ανάλυση.

Ένα από τα βασικά σημεία σε αυτήν την προσέγγιση παραμένει ο καθορισμός μιας κατάλληλης παραμέτρου κανονικοποίησης  $pr$ . Κατά την επαναξιολόγηση αυτών των τιμών η ιδανική  $pr$  καθορίζεται πάντα, λαμβάνοντας υπόψη τις ιδιαιτερότητες της εκάστοτε θεματικής περιοχής της γνώσης και είναι αυτή που ποσοτικοποιεί το σημασιολογικό συσχετισμό των εννοιών στη θεματική περιοχή. Στο πλαίσιο της παρούσης πραγματοποίησαμε μια σειρά πειραμάτων σε ένα σύνολο εκπαίδευσης εικόνων κατάληλα επιλεγμένων από τις εκάστοτε θεματικές περιοχές εφαρμογής (π.χ. παραλία ή/και μηχανοκίνητα σπορ) και επιλέξαμε την  $pr$ , που οδήγησε στις καλύτερες τιμές αξιολόγησης των αποτελεσμάτων για κάθε θεματική περιοχή.

### 6.3.3 Μεθοδολογία και αλγόριθμος στην περίπτωση πολλαπλών θεματικών περιοχών

Στη συνέχεια, προτείνουμε την επέκταση της αρχικής ερευνητικής προσέγγισής μας προς δύο ενδιαφέρουσες ερευνητικές κατευθύνσεις: αυτή της ταυτόχρονης ύπαρξης πολλαπλών θεματικών περιοχών και αυτή της αναπαράστασης του εννοιολογικού πλαισίου με βάση την νεότερη γλώσσα του Σημασιολογικού Ιστού, της OWL [319]. Η προτεινόμενη αναπαράσταση είναι σε θέση να χρησιμοποιηθεί με οποιαδήποτε μεθοδολογία υποβοηθούμενης από τη γνώση ανάλυσης και ο απώτερος στόχος αυτής της επέκτασης είναι η εφαρμογή της ανεξάρτητα από την επιλεγμένη θεματική περιοχή. Βέβαια, στο πλαίσιο της έρευνάς μας αυτοπεριοριστήκαμε, όπως θα δούμε και στην υποενότητα των σχετικών αποτελεσμάτων, για λόγους απλότητας και σαφήνειας σε συγκεκριμένες θεματικές περιοχές, όμως κάτι τέτοιο δεν αποτελεί περιορισμό της προτεινόμενης εργασίας. Συνακόλουθα, ο μόνος περιορισμός που τίθεται για το εκάστοτε υπό εξέταση πολυμεσικό περιεχόμενο είναι το γεγονός ότι αυτό θα πρέπει να εμπίπτει στο σύνολο των θεματικών περιοχών που υποστηρίζονται από την κατασκευασμένη οντολογία του εννοιολογικού πλαισίου.

Όπως προκύπτει από τα προηγούμενα, τόσο η υπάρχουσα έννοια του οπτικού εννοιολογικού πλαισίου, όσο και η έως τώρα προτεινόμενη υλοποίηση του αλγορίθμου βασίζονται στην ύπαρξη μιας εκ των προτέρων γνωστής θεματικής περιοχής εφαρμογής. Όμως για την αποδοτική αντικειώπιση των προβλημάτων του πραγματικού κόσμου των πολυμέσων, κάτι τέτοιο αποτελεί μια μάλλον ισχυρή και προαπαιτούμενη υπόθεση εργασίας, η οποία περιορίζει σε κάποιο βαθμό την ευελιξία και την προσαρμοστικότητα της προσέγγισής μας. Για το λόγο αυτό θεωρούμε ότι ο αλγόριθμος που χρησιμοποιείται, για να επαναπροσδιορίσει το βαθμό ιδιότητας μέλους κάθε έννοιας, που συνδέεται με μια περιοχή μιας εικόνας ή και με ολόκληρη την εικόνα, θα λειτουργεί αποδοτικότερα, εφόσον δεν εξαρτάται από τη θεματική περιοχή εφαρμογής. Μια πρώτη προσέγγιση της προτεινόμενης επέκτασης έχει ήδη υλοποιηθεί και δημοσιευθεί στη [185], ενώ ολοκληρωμένα αποτελέσματα παρουσιάζονται στη συνέχεια στην υποενότητα 6.3.3.4.

Και σε αυτή την περίπτωση, κατά τη διάρκεια της διαδικασίας επίδρασης του εννοιολογικού πλαισίου, εισάγουμε μια μεθοδολογία για τη βελτίωση των αποτελεσμάτων της κατάτμησης εικόνας με την έννοια των εννοιών κάθε περιοχής/τμήματος, σύμφωνα με την επιπρόσθετη πληροφορία που εισάγει το εννοιολογικό πλαίσιο. Σε σύγκριση με την ερευνητική συνεισφορά της προηγούμενης ενότητας παρουσιάζεται εδώ μια καινοτόμα οντολογική αναπαράσταση πολλαπλών θεματικών περιοχών για το εννοιολογικό πλαίσιο, η οποία συνδυάζει την θεωρία της ασάφειας και της ασαφούς άλγεβρας [177] με τη γλώσσα OWL [319] και την τεχνική του reification [318]. Ακολουθώντας τη διαδικασία αυτή, οι βαθμοί συμμετοχής των εννοιών της κάθε περιοχής επανεκτιμώνται κατάλληλα σύμφωνα με μια τροποποιημένη εκδοχή του βασισμένου στο εννοιολογικό πλαίσιο αλγορίθμου επαναπροσδιορισμού τους.

### 6.3.3.1 Ασαφής εννοιολογική αναπαράσταση γνώσης

Η προτεινόμενη επέκταση του μοντέλου γνώσης μπορεί να εκφραστεί και πάλι με τη χρήση των βασικών δομικών στοιχείων, όπως είναι οι έννοιες, οι σχέσεις μεταξύ τους και οι θεματικές περιοχές που δημιουργούν μια οντολογική δομή [104]. Καθώς, δε, οι σχέσεις μεταξύ εννοιών που απαντώνται στην πραγματική ζωή μοντελοποιούνται καλύτερα, χρησιμοποιώντας βασικές αρχές της ασαφούς σχεσιακής άλγεβρας, όπως οι τελευταίες παρουσιάστηκαν στην ενότητα 3.2, παραπλήσιοι κανόνες οφείλουν να διέπουν και τις εδώ χρησιμοποιούμενες σχέσεις.

Εν συντομίᾳ, ορίζουμε και πάλι ένα κλασικό σύνολο εννοιών  $S = \{s_i\}$ ,  $i = 1, \dots, N$ ,  $s_i \in S$ , το οποίο περιγράφεται από τη συνάρτηση συμμετοχής  $\mu_S : U \rightarrow \{0, 1\}$ . Το αντίστοιχο ασαφές σύνολο  $F$  στο  $S$  περιγράφεται από τη συνάρτηση συμμετοχής  $\mu_F : S \rightarrow [0, 1]$ . Ορίζουμε επίσης το κλασικό σύνολο σχέσεων μεταξύ εννοιών ως  $\{r_i\}$ , όπου  $r_i : S \times S \rightarrow \{0, 1\}$ ,  $i = 1, \dots, M$  και  $M \leq \frac{N!}{(N-2)!}$ , ενώ το κλασικό σύνολο ασαφών σχέσεων μεταξύ εννοιών δίνεται από την  $R = \{R_i\}$ , όπου  $R_i : S \times S \rightarrow [0, 1]$ ,  $i = 1, \dots, M$ . Έστω ότι  $Z$  είναι το κλασικό σύνολο όλων των θεματικών υπο-περιοχών (οι οποίες την ίδια στιγμή είναι και έννοιες και οι ίδιες). τότε  $Z \subseteq S$  και  $Z = \{z_i\}$ ,  $i = 1, \dots, L$ ,  $L \leq N$ ,  $z_i \in S$ . Τέλος, η (αν-)οιμοιότητα της τρέχουσας θεματικής υπο-περιοχής καταδεικνύεται από το  $p_i \in [0, 1]$ , το οποίο ορίζεται εκ των προτέρων από ειδήμονες των θεματικών περιοχών και χαρακτηρίζει την εκάστοτε θεματική υπο-περιοχή. Ο δείκτης  $i$  δείχνει τη συγκεκριμένη θεματική υπο-περιοχή και θεωρούμε ότι  $\eta |Z|$  προσδιορίζει τον αριθμό των διακεχριμένων θεματικών υπο-περιοχών  $z_i$ .

Με βάση τα παραπάνω, κατασκευάζεται μια ασαφής οντολογία, η οποία περιέχει

θεματικές υπο-περιοχές, έννοιες και ασαφείς σχέσεις μαζί με βαθμούς συμμετοχής για κάθε σχέση που συνδέει δύο έννοιες μεταξύ τους. Η τελευταία δήλωση μπορεί να διατυπωθεί ως:  $\mathcal{O} = \{S, R, Z\}$ , όπου η  $\mathcal{O}$  καταδεικνύει την ασαφή οντολογία, το  $S$  είναι το κλασικό σύνολο των εννοιών που περιγράφονται από την οντολογία, το  $R$  είναι το κλασικό σύνολο όλων των ασαφών σχέσεων μεταξύ των εννοιών αυτών και το  $Z$  είναι το κλασικό σύνολο των θεματικών υπο-περιοχών που περικλείονται στην  $\mathcal{O}$ .

Έστω, επίσης, ότι το  $C$  είναι το ασαφές σύνολο των εννοιολογικών πλαισίων στο  $S$  και το  $C$  είναι το σύνολο όλων των εννοιολογικών πλαισίων στην  $\mathcal{O}$ ,  $C \subseteq \mathcal{F}_S$ . Θεωρούμε ένα σαφές σύνολο εννοιολογικών πλαισίων  $C$ , το οποίο σχετίζεται με κάθε θεματική υπο-περιοχή  $z_i \in Z$  της οντολογίας, ή με άλλα λόγια ορίζουμε την έννοια του εννοιολογικού πλαισίου για μία θεματική υπο-περιοχή  $z_i \in Z$  ως το ασαφές σύνολο  $C_{z_i} = \sum_j c_j / w_j = \{c_1/w_1, c_2/w_2, \dots, c_n/w_n\}$ ,  $n \leq N$  όλων των εννοιών  $c_j \in S$  που σχετίζονται άμεσα με τη θεματική υπο-περιοχή  $z_i$ , μαζί με τους βαθμούς συμμετοχής  $w_j$  των σχέσεων τους. Ο βαθμός συμμετοχής  $w_j$  περιγράφει την συνάρτηση συμμετοχής  $\mu_F(c_j)$ , δηλ.  $w_j = \mu_F(c_j)$ , ή για λόγους απλότητας  $w_j = F(c_j)$ . Ο όρος άμεσα χρησιμοποιείται λαμβάνοντας υπόψη την σχετική μετρική απόστασης (π.χ. έννοιες που βρίσκονται λιγότερο από δύο χορυφές μακριά από την έννοια της θεματικής υπο-περιοχής).

Προκειμένου να ορίσουμε, να εξάγουμε και να εκμεταλλευτούμε κάθε εννοιολογικό πλαίσιο  $C$  ενός αυθαίρετου συνόλου εννοιών  $S$ , βασιζόμαστε στη σημασιολογία των ασαφών σημασιολογικών σχέσεων τους. Με βάση τις σχέσεις  $R_i$  κατασκευάζουμε μια συνδυασμένη σχέση  $I$ , η οποία εντείνει και εμπλουτίζει τον ορισμό της ασάφειας μεταξύ εννοιών στην οντολογία  $\mathcal{O}$ :  $I = \bigcup_i R_i^{q_i}$ ,  $q_i \in \{-1, 0, 1\}$ ,  $i = 1 \dots M$ . Η τιμή  $q_i$  καθορίζεται σύμφωνα με τη σημασιολογία<sup>1</sup> κάθε σχέσης  $R_i$ , όπως αυτή ορίζεται στο πρότυπο MPEG-7 [297], όπως π.χ. είναι η εξειδίκευση, το μέρος, η ιδιότητα, κ.α.. Όπως υπονοείται από την κατασκευή της σχέσης  $I$ , θα πρέπει να προνοήσουμε για την αποφυγή πιθανών κύκλων, οι οποίοι μπορούν να εμφανιστούν εξαιτίας των συνδέσεων μεταξύ θεματικών υπο-περιοχών που ορίζουν δύο έννοιες. Οι κύκλοι αποφεύγονται, αν δε λάβουμε υπόψη έννοιες που την ίδια στιγμή αποτελούν οι ίδιες και θεματικές υπο-περιοχές. Έτσι, υπολογίζουμε το ασαφές σύνολο των εννοιολογικών πλαισίων για όλες τις έννοιες εκτός από αυτές που είναι οι ίδιες και θεματικές υπο-περιοχές. Η σχέση  $I$  είναι εξαιρετικής σημασίας, καθώς μας επιτρέπει τον ορισμό, την εξαγωγή και χρήση των εννοιολογικών πλαισίων  $C$  για κάθε δεδομένο σύνολο από έννοιες  $S$ .

### 6.3.3.2 Επεκτεταμένο μοντέλο γνώσης

Η προτεινόμενη εννοιολογική αναπαράσταση γνώσης παρέχει και αυτή τα μέσα για την εκμετάλλευση της πληροφορίας εννοιολογικού πλαισίου στο πλαίσιο που θέτει η ανάλυση πολυμεσικού περιεχομένου. Η βασική διαφορά σε σχέση με τις υπόλοιπες προσεγγίσεις της παρούσας διδακτορικής έγκειται στη χρήση της γλώσσας OWL στη διαδικασία. Θεωρούμε ότι, για να αντιμετωπίσουμε την περίπτωση χρήσης και εκμετάλλευσης του εννοιολογικού πλαισίου πολλαπλών θεματικών περιοχών, η γλώσσα OWL περιγράφει και αντιμετωπίζει καλύτερα τις απαιτήσεις του προβλήμα-

<sup>1</sup>Η άμεση (δηλ.  $q_i = 1$ ) σημασιολογική ερμηνεία μιας σχέσης μπορεί να μην έχει νόημα, ενώ η αντίθετη μπορεί να έχει (δηλ.  $q_i = -1$ ), ενώ το  $q_i = 0$  εξαλείφει τη σχέση.

## Κεφάλαιο 6. Εννοιολογικό πλαίσιο και οπτική πληροφορία

τος και αναμένεται ότι θα επιτρέψει περαιτέρω εξαγωγή χρήσιμων συμπερασμάτων από τη γνώση. Επιπρόσθετα, προτείνουμε και την επέκταση της τεχνικής του reification στη γλώσσα OWL για την αναπαράσταση των ασαφών βαθμών συμμετοχής της οντολογίας.

Το προτεινόμενο μοντέλο γνώσης μπορεί να θεωρηθεί ως ένας γράφος, στον οποίο κάθε κορυφή αντιπροσωπεύει μία έννοια και κάθε ακμή μεταξύ δύο κορυφών αντιπροσωπεύει μια σχέση μεταξύ των δύο συγκεκριμένων έννοιών. Κάθε ακμή συνοδεύεται από έναν σχετικό βαθμό εμπιστοσύνης, ο οποίος εκφράζει την επιθυμητή ασάφεια του μοντέλου γνώσης. Μία από τις βασικές καινοτομίες και ο βασικός λόγος για τη χρήση της γλώσσας OWL αντί για τη γλώσσα RDF είναι το γεγονός ότι στην περίπτωση αυτή υπάρχει ένα σύνολο από “ειδικές” έννοιες, το κλασικό σύνολο των θεματικών υπο-περιοχών  $Z \subseteq S$ . Αυτές οι έννοιες υπο-περιοχής εισάγονται χειρωνακτικά

```
<rdf:RDF
  xmlns:rdf = "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#" ... >
<owl:Class rdf:ID="beach"> <rdfs:subClassOf
rdf:resource="#holidays" />
...
</owl:Class> <owl:Class rdf:ID="vacation"/> ...
<owl:Class rdf:ID="sand"/> <rdf:Description rdf:about="s1">
<rdf:subject rdf:resource="#sand"/>
<rdf:predicate rdf:resource="#isRelated"/>
<rdf:object rdf:resource="#beach"/>
<rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Statement"/>
<context:isRelated rdf:datatype="http://www.w3.org/2001/XMLSchema
#float">0,85</context:isRelated>
</rdf:Description> ... </rdf:RDF>
```

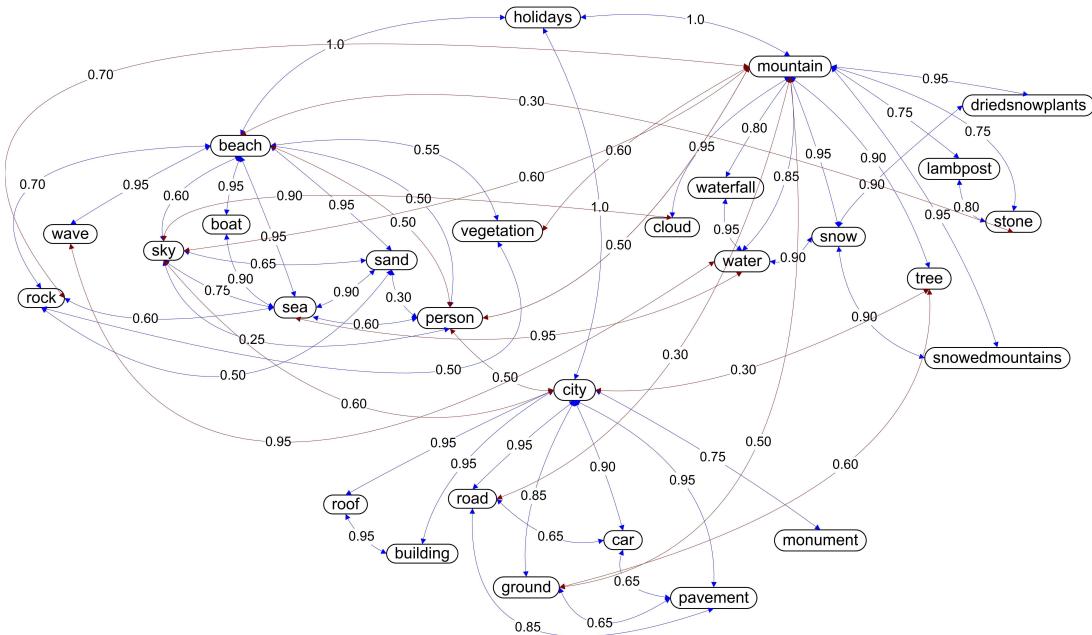
Σχήμα 6.8: Παράδειγμα αναπαράστασης της τεχνικής OWL reification.

στο μοντέλο του γράφου από ειδήμονες του ερευνητικού πεδίου και λειτουργούν ως μεσάζοντες μεταξύ του στοιχείου ρίζας και των υπόλοιπων έννοιών της οντολογίας. Οι θεματικές υπο-περιοχές δε θεωρούνται μέρος του συνόλου των έννοιών, που μπορεί να προήλθαν από κάποια κλασική μονάδα ανάλυσης πολυμεσικού περιεχομένου, αλλά είναι δηλωτικές των θεματικών περιοχών που υποστηρίζονται από τη γνώση του έννοιολογικού πλαισίου.

Η αναπαράσταση του παραπάνω γράφου στη γλώσσα OWL δεν είναι μια τετριμένη ούτε μια ξεκάθαρη εργασία κυρίως λόγω της πολυπλοκότητας και της ποικιλίας των περιορισμών και των κανόνων επικύρωσής της. Παρ' όλα αυτά, το μοντέλο που προτείνουμε είναι πλήρως επικυρωμένο και σύμφωνο με τους κανόνες της γλώσσας, καθώς κατασκευάστηκε, λαμβάνοντας υπόψη ολόκληρο το σύνολο των συντακτικών κανόνων της OWL. Ο ασαφής βαθμός εμπιστοσύνης που συσχετίζεται με κάθε σχέση μεταξύ δύο οποιονδήποτε έννοιών περιγράφεται με τη βοήθεια της τεχνικής του reification, αυτή τη φορά υλοποιημένης στη γλώσσα OWL. Ως γνωστόν, η τεχνική αυτή εισάγει μία επιπλέον δήλωση για κάθε κλασική δήλωση, η οποία επιπλέον δήλωση είναι αυτή που περιέχει την πληροφορία του βαθμού της σχέσης. Για το λόγο αυτό είναι ιδανική για την κατάλληλη αναπαράσταση της ασάφειας, μιας και η συνολική δήλωση δε θα πρέπει να βεβαιώνεται ούτε αυτόματα, ούτε και για καθεμία περίπτωση. Ένα σχετικό παράδειγμα παρουσιάζεται στο Σχήμα 6.8, όπου χρησιμοποιείται η αφηρημένη

σχέση *isRelated* για την αναπαράσταση της συνδυασμένης σχέσης *I* στις έννοιες παραλία και άμμος.

Θα πρέπει να σημειώσουμε εδώ ότι το στοιχείο *rίζας* παίζει έναν ιδιαίτερο ρόλο, καθώς η συνολική πληροφορία του εννοιολογικού πλαισίου υπολογίζεται σε σχέση με αυτό. Παρ’ όλα αυτά, δεν είναι αυτό που χαρακτηρίζει το πεδίο της εφαρμογής. Αντίθετα, το πλήθος και το είδος των θεματικών υπο-περιοχών είναι εκείνο που το καθορίζει και το γεγονός αυτό είναι που καθιστά την προτεινόμενη επέκταση της τεχνικής μας ιδιαίτερα σημαντική, μιας και μπορεί να εφαρμοστεί πλέον σε πολλαπλές και πλήρως διαχριτές/ανεξάρτητες μεταξύ τους θεματικές περιοχές.



**Σχήμα 6.9:** Δείγμα γράφου οντολογίας - θεματικές υπο-περιοχές: παραλία, βουνό, πόλη

Δεδομένου του δείγματος γράφου που παρουσιάζεται στο Σχήμα 6.9, λαμβάνονται υπόψη όλες οι πιθανές διαδρομές μεταξύ των διαθέσιμων κορυφών. Με τον τρόπο αυτό ακολουθείται μια εξαντλητική προσέγγιση η οποία, όμως, ακολουθεί τον κανόνα ότι οι διαδρομές μεταξύ εννοιών είναι αμοιβαίες και ότι οι διαδρομές μεταξύ δύο εννοιών δεν επιτρέπεται να περιλάβουν μία έννοια θεματικής υπο-περιοχής. Έτσι, η αρχική εκτίμηση της τιμής του εννοιολογικού πλαισίου για κάθε έννοια είναι συνδεδεμένη με τη συγκεκριμένη θεματική υπο-περιοχή, στην οποία ανήκει αρχικά η έννοια και απορρέει από τις σχέσεις της έννοιας με τις υπόλοιπες έννοιες της ίδιας υπο-περιοχής ή ακόμα και με τις έννοιες μιας διαφορετικής θεματικής υπο-περιοχής, λαμβάνοντας πάντοτε υπόψη τους βαθμονομημένους συνδέσμους τους. Πρόκειται, δηλαδή, για ακόμα μία περίπτωση υλοποίησης της αρχής του μεταβατικού κλεισίματος, όπως αυτή παρουσιάστηκε στην ενότητα 3.3 της παρούσης διατριβής. Φυσικά, στην παραπάνω διαδικασία χρησιμοποιείται και ένας κατάλληλος δείκτης συμβατότητας ή με άλλα λόγια μια κατάλληλη μετρική της απόστασης, άρα της ίδιας της σχετικότητας, των κορυφών. Ανάλογα και με τη φύση των θεματικών περιοχών που εξετάζονται κάθε φορά ο καλύτερος δείκτης επιλέγεται από μια ποικιλία κατάλληλων τελεστών της βιβλιογραφίας, όπως είναι π.χ. η *t*-νόρμα ή *t*-co-νόρμα. Στη γενική περίπτωση, η ιδανική μετρική απόστασης για δύο έννοιες είναι εκείνη που ποσοτικοποιεί τη σημασιολογική συσχέτισή τους. Για παράδειγμα, για τις τρεις υπο-περιοχές παραλία, βουνό και πόλη του παραδείγματος ένα κατάλληλο μέτρο της συσχέτισής τους είναι

η χρήση της μέγιστης τιμής:  $\max$ .

### 6.3.3.3 Τροποποίηση αλγορίθμου ανάλυσης

Η κύρια λειτουργικότητα της εκδοχής αυτής του αλγορίθμου του εννοιολογικού πλαισίου είναι η αποδοτική επανεκτίμηση των βαθμών συμμετοχής κάθε έννοιας που σχετίζεται με μία συγκεκριμένη περιοχή ή τιμή μιας εικόνας. Η πληροφορία αυτή μπορεί να προέρχεται από οποιαδήποτε μονάδα επεξεργασίας ή/και χαμηλού επιπέδου ανάλυσης πολυμεσικού περιεχομένου, γεγονός που καθιστά την προσέγγισή μας ιδιαίτερα χρήσιμη από την ερευνητική σκοπιά. Η καινοτομία που εισάγεται εστιάζει στην τιμή του εννοιολογικού πλαισίου, η οποία αξιοποιείται, προκειμένου να αντιμετωπιστούν περιπτώσεις, όπου μία έννοια σχετίζεται με περισσότερες από μία έννοιες και που οι έννοιες αυτές βρίσκονται στο πλαίσιο είτε της ίδιας θεματικής υπο-περιοχής, είτε διαφορετικών θεματικών υπο-περιοχών. Σε κάθε περίπτωση, η τιμή του εννοιολογικού πλαισίου μιας έννοιας αναφέρεται στη συνολική συνάφεια της έννοιας ως προς τη σχετική θεματική υπο-περιοχή και με τον τρόπο αυτό ως προς το στοιχείο ρίζας ολόκληρης της οντολογίας. Για παράδειγμα, στο Σχήμα 6.9, όπου το στοιχείο ρίζας είναι: διακοπές και οι σχετικές υπο-περιοχές είναι: παραλία, βουνό και πόλη, η έννοια νερό σχετίζεται με δύο από αυτές: παραλία και βουνό. Ο ρόλος των θεματικών υπο-περιοχών συνοψίζεται στο να λειτουργούν ως φράγμα, όταν λαμβάνονται υπόψη όλες οι πιθανές διαδρομές στο γράφο κατά την εκτίμηση της τιμής του εννοιολογικού πλαισίου κάθε έννοιας: καμία διαδρομή δεν επιτρέπεται να περάσει μέσω μιας έννοιας θεματικής υπο-περιοχής, μιας και κάτι τέτοιο θα αναμείγνυε με λάθος τρόπο τη σημασιολογία κάθε υπο-περιοχής. Χρησιμοποιώντας την τυπική  $t$ -co-νόρμα και το αλγεβρικό γινόμενο ως την  $t$ -νόρμα, καταλήγουμε στην ακόλουθη τροποποιημένη μορφή του αλγορίθμου του εννοιολογικού πλαισίου:

- Για κάθε υπο-περιοχή  $z_i \in Z$  προσδιορίστε το σύνολο των εννοιολογικών πλαισίων της:  $C_{z_i}$ .
- Για κάθε έννοια  $s_k \in S$ , που σχετίζεται με μία περιοχή μιας εικόνας με ένα βαθμό συμμετοχής  $\mu(s_k)$ , προσδιορίστε τις υπάρχουσες σχέσεις με το σύνολο των εννοιολογικών πλαισίων:  $\{R_{s_k, c_j} : c_j \in C_{z_i}\}$ .
- Αν και μόνο αν  $R_{s_k, c_j} \neq 0$ , τότε αποκτήστε τη συγκεκριμένη πληροφορία εννοιολογικού πλαισίου με τη μορφή ενός νέου βαθμού συμμετοχής  $\mu(s_k)$ , ο οποίος σχετίζεται με την περιοχή της εικόνας σύμφωνα με την αναδρομικό τύπο:

$$\mu^t(s_k) = \mu^{t-1}(s_k) - p_i \cdot (\mu^{t-1}(s_k) - w_j) \quad (6.8)$$

όπου το  $t$  δηλώνει την εκάστοτε επανάληψη.

- Αφαιρέστε από το σύνολο των εννοιών εκείνες που δε σχετίζονται άμεσα με τη θεματική υπο-περιοχή  $z_i$ , στην οποία ανήκει η υπό εξέταση εικόνα, σύμφωνα με τα αποτελέσματα της αρχικής κατηγοριοποίησής της<sup>2</sup>.

<sup>2</sup>Προκειμένου να μειώσουμε την επίδραση του εννοιολογικού πλαισίου στη διαδικασία μπορούμε να παρακάμψουμε αυτό το βήμα και με τον τρόπο αυτό να “ομαλοποιήσουμε” τη θεματική περιοχή εφαρμογής.

Η φυσική σημασία της εξίσωσης (6.8) είναι ότι το  $p_i$  ελέγχει το κατά πόσο θα εφαρμοστεί ή όχι η επίδραση του εννοιολογικού πλαισίου. Με άλλα λόγια,  $p_i = 0$  σημαίνει καμία επιρροή εννοιολογικού πλαισίου, ενώ  $p_i = 1$  καταδεικνύει μια ισχυρή επίδρασή του, η οποία τελικά αγνοεί παντελώς τους αρχικούς βαθμούς  $\mu(s_k)$ . Ισοδύναμα με την εξίσωση (6.8), για μια αυθαίρετη επανάληψη  $i$ , θα έχουμε:

$$\mu^t(s_k) = (1 - p_i)^t \cdot \mu^0(s_k) + (1 - (1 - p_i)^t) w_j \quad (6.9)$$

όπου το  $\mu^0(s_k)$  αναπαριστά τον αρχικό βαθμό συμμετοχής. Συνήθως, η επανάληψη σταματά, όταν το  $t$  γίνει ίσο με 3 ή 5.

#### 6.3.3.4 Πειραματικά αποτελέσματα

Προκειμένου να υλοποιήσουμε και να δοκιμάσουμε την αποτελεσματικότητα του παραπάνω αλγορίθμου, πραγματοποιήσαμε πειραματικές μετρήσεις χρησιμοποιώντας ένα σύνολο από 573 εικόνες. Το περιεχόμενο των εικόνων ήταν ισομερώς κατανεμημένο σε τρεις θεματικές περιοχές: 193 εικόνες παραλία, 184 εικόνες βουνό και 196 εικόνες πόλη και προήλθε από προσωπικές συλλογές και το Διαδίκτυο. Στη συνέχεια, παρουσιάζουμε ένα αντιπροσωπευτικό δείγμα τόσο από ενδεικτικά, όσο και από συνολικά ποσοτικά αποτελέσματα. Στο σημείο αυτό θα πρέπει να αναφέρουμε ότι η σχετική απόλυτη αλήθεια (ground truth) κατασκευάστηκε χειρωνακτικά και αποτελείται από ένα πλήθος συνδεδεμένων, μη επικαλυπτόμενων περιοχών συσχετισμένων με μία μοναδική έννοια.

Αρχικά, παρουσιάζονται δύο εκδοχές μιας ενδεικτικής εικόνας από τη θεματική περιοχή παραλία (Σχήμα 6.10): (i) η αρχική εικόνα και (ii) το σύνολο των 17 κατατυημένων περιοχών της. Προκειμένου να λάβουμε το αποτέλεσμα της κατάτυησης, υλοποιήσαμε τον αλγόριθμο κατάτυησης RSST (ο οποίος περιγράφεται στη δημοσίευση [179]), χρησιμοποιώντας ένα κατώφλι απόστασης για τον τερματισμό της σε έναν κατάλληλο αριθμό από περιοχές. Αξιοποιώντας 100 εικόνες (περίπου το 20% του συνόλου δεδομένων) ως το σύνολο εκπαίδευσης, επιλέξαμε τις τιμές του  $p_i$  που οδήγησαν στην καλύτερη αξιολόγηση.



(a) Αρχική εικόνα



(b) Περιοχές κατάτυησης

**Σχήμα 6.10:** Ενδεικτικό παράδειγμα από μια εικόνα παραλίας.

Ο Πίνακας 6.1 συνοφίζει τους τελικούς βαθμούς συμμετοχής πριν και μετά την εφαρμογή του οπτικού εννοιολογικού πλαισίου για το παράδειγμα της εικόνας από τη θεματική περιοχή παραλία. Οι βαθμοί που συμβολίζονται με έντονα γράμματα καταδεικνύουν την πληροφορία απόλυτης αλήθειας για τις περιοχές της εικόνας, ενώ

## Κεφάλαιο 6. Εννοιολογικό πλαίσιο και οπτική πληροφορία

η πρώτη στήλη του Πίνακα καταδεικνύει τον αύξοντα αριθμό της εκάστοτε περιοχής υπό εξέταση. Για λόγους απλότητας και εξοικονόμησης χώρου, οι έννοιες που δεν σχετίζονται καθόλου (όπως προκύπτει από το Σχήμα 6.9) με τη θεματική υπο-περιοχή παραλία, αγνοούνται.

Στη συνέχεια, παρουσιάζονται στον Πίνακα 6.2 συνολικά σκορ ακρίβειας για κάθε μία έννοια ζεχωριστά. Τα σκορ αυτά προέκυψαν από την εφαρμογή της μεθοδολογίας στο συνολικό σύνολο δεδομένων των 573 εικόνων και για τη συνολική θεματική περιοχή των διακοπών. Η γραμμή κάθε έννοιας δείχνει τα σκορ πριν και μετά την αξιοποίηση του εννοιολογικού πλαισίου.

**Πίνακας 6.1:** Τελικοί βαθμοί συμμετοχής πριν και μετά την εφαρμογή του οπτικού εννοιολογικού πλαισίου για το παράδειγμα της εικόνας από τη θεματική περιοχή παραλία

Α/A	Έννοιες											
	Βράχος		κύμα		ουρανός		βάρκα					
	πριν μετά	άμμος										
1	0.26	0.09	0.63	0.54	0.70	0.46	0.34	0.11	<b>0.71 0.79</b>	0.47	0.14	
2	0.63	0.42	0.13	0.07	0.67	0.24	0.70	0.28	0.17	0.03	0.77	0.46
3	0.73	0.48	0.82	0.41	0.69	0.25	0.68	0.27	0.37	0.07	0.98	0.59
4	0.33	0.11	0.70	0.59	0.67	0.45	0.31	0.10	<b>0.77 0.85</b>	0.27	0.08	
5	0.69	0.59	0.24	0.12	0.11	0.10	0.48	0.32	0.18	0.16	<b>0.85 0.94</b>	
6	0.67	0.58	0.25	0.13	0.13	0.12	0.37	0.24	0.29	0.26	<b>0.79 0.87</b>	
7	0.62	0.41	0.41	0.21	0.31	0.10	0.69	0.27	0.17	0.04	0.78	0.47
8	0.53	0.35	0.44	0.22	0.37	0.12	0.70	0.28	0.24	0.11	0.77	0.46
9	0.49	0.32	0.48	0.24	0.41	0.14	0.94	0.38	0.09	0.07	0.27	0.16
10	0.23	0.09	0.65	0.55	0.73	0.48	0.39	0.13	<b>0.79 0.87</b>	0.18	0.05	
11	0.40	0.24	0.25	0.13	0.22	0.05	0.78	0.34	0.44	0.33	0.92	0.55
12	0.37	0.22	0.36	0.18	0.47	0.11	0.70	0.30	0.67	0.51	0.77	0.46
13	0.53	0.32	0.42	0.21	0.39	0.09	0.68	0.29	0.35	0.26	0.98	0.59
14	0.73	0.62	0.30	0.15	0.27	0.18	0.31	0.20	0.25	0.23	<b>0.69 0.75</b>	
15	0.69	0.58	0.34	0.29	0.11	0.07	<b>0.81 0.87</b>	0.18	0.10	0.85	0.73	
16	0.67	0.57	0.35	0.30	0.13	0.09	<b>0.67 0.73</b>	0.07	0.04	0.77	0.66	
17	0.65	0.55	0.45	0.38	0.21	0.14	<b>0.69 0.74</b>	0.27	0.17	0.61	0.52	

Α/A	πρόσωπο	βλάστηση	σύννεφο	καταρράκτης	νερό	πέτρα						
	πριν μετά	πριν μετά	πριν μετά	πριν μετά	πριν μετά	πριν μετά						
1	0.21	0.05	0.15	0.02	0.70	0.56	0.36	0.18	0.71	0.68	0.19	0.06
2	<b>0.77 0.92</b>	0.76	0.27	0.37	0.30	0.30	0.15	0.27	0.14	0.76	0.42	
3	<b>0.73 0.87</b>	0.82	0.29	0.42	0.34	0.18	0.09	0.36	0.18	0.62	0.34	
4	0.14	0.04	0.24	0.04	0.67	0.54	0.50	0.25	0.70	0.63	0.12	0.04
5	0.59	0.53	0.74	0.67	0.31	0.22	0.48	0.24	0.18	0.14	0.74	0.67
6	0.67	0.61	0.75	0.67	0.33	0.23	0.22	0.11	0.14	0.11	0.75	0.67
7	<b>0.95 0.96</b>	0.81	0.28	0.31	0.25	0.29	0.15	0.67	0.22	0.81	0.45	
8	<b>0.83 0.87</b>	0.64	0.22	0.67	0.54	0.33	0.17	0.19	0.10	0.45	0.25	
9	<b>0.89 0.89</b>	0.88	0.31	0.51	0.41	0.44	0.22	0.09	0.07	0.61	0.34	
10	0.13	0.03	0.23	0.03	0.63	0.50	0.43	0.22	0.59	0.56	0.05	0.01
11	<b>0.80 0.88</b>	0.65	0.23	0.72	0.58	0.28	0.14	0.44	0.33	0.65	0.39	
12	<b>0.77 0.92</b>	0.76	0.27	0.17	0.14	0.30	0.15	0.26	0.20	0.76	0.46	
13	<b>0.73 0.87</b>	0.82	0.29	0.25	0.20	0.48	0.24	0.13	0.10	0.82	0.49	
14	0.73	0.62	0.70	0.63	0.67	0.47	0.71	0.35	0.26	0.19	0.70	0.63
15	0.69	0.23	0.74	0.35	0.51	0.41	0.18	0.09	0.38	0.28	0.74	0.56
16	0.67	0.22	0.75	0.36	0.43	0.34	0.24	0.12	0.47	0.34	0.75	0.56
17	0.95	0.31	0.81	0.39	0.36	0.29	0.30	0.15	0.50	0.37	0.81	0.61

**Πίνακας 6.2:** Συνολικά σκορ ακρίβειας ανά έννοια για 573 εικόνες

Έννοιες	πριν μετά	Έννοιες πριν μετά	Έννοιες	πριν μετά	
κτίριο	0.78 0.87	κύμα	0.54 0.60	ουρανός	0.75 0.86
στάνη	0.49 0.51	θάλασσα	0.74 0.82	άμμος	0.68 0.74
βλάστηση	0.78 0.83	πρόσωπο	0.73 0.82	σύννεφο	0.66 0.68
καταρράκτης	0.32 0.33	νερό	0.37 0.40	χιόνι	0.75 0.79
χιονισμένα-βουνά	0.46 0.49	πέτρα	0.64 0.70	βάρκα	0.67 0.75
χιονισμένα-φυτά	0.52 0.54	δένδρο	0.56 0.58	στέγη	0.74 0.83
μνημείο	0.57 0.65	έδαφος	0.69 0.76	αυτοκίνητο	0.84 0.89
πεζοδρόμιο	0.70 0.74	δρόμος	0.76 0.82	βράχος	0.67 0.73
<b>Σύνολο</b>	<b>0.64 0.70</b>				

## 6.4 Οπτικό Εννοιολογικό Πλαίσιο και Σημασιολογική Κατάτμηση Εικόνων

Ως αυτό το σημείο παρουσιάσαμε τις βασικές θεωρητικές αρχές της προτεινόμενης μεθοδολογίας, ενώ έγινε κατανοητή και η επέκτασή της στην περίπτωση πολλαπλών θεματικών περιοχών. Στην ενότητα αυτή θα εμβαθύνουμε στην πρακτική εφαρμογή της αξιοποίησης του οπτικού εννοιολογικού πλαισίου στο πεδίο της σημασιολογικής κατάτμησης.

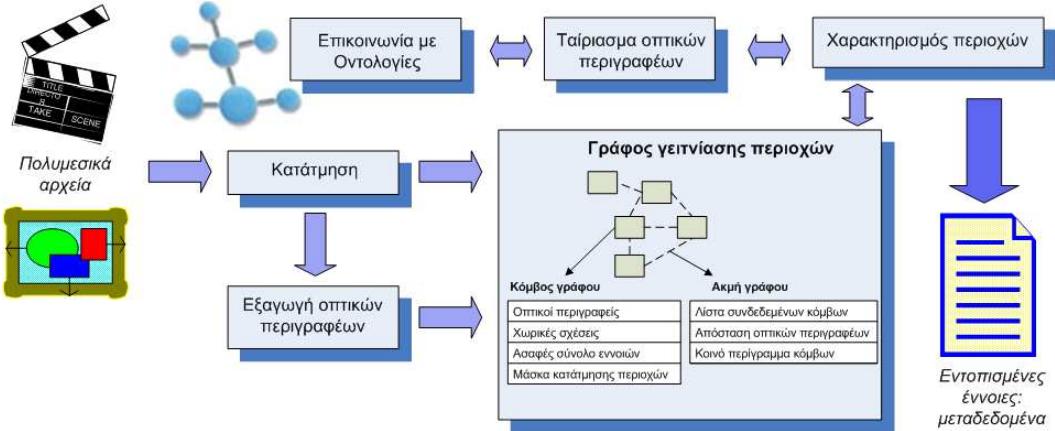
### 6.4.1 Αρχικοποίηση σημασιολογικής κατάτμησης

Ο προφανής στόχος της προσέγγισής μας είναι η αντιμετώπιση των προβλημάτων της πολυμεσικής ανάλυσης από τη σημασιολογική σκοπιά τους, όπου οι περιοχές μιας εικόνας συνδέονται εκτός από τα χαμηλού επιπέδου οπτικά χαρακτηριστικά τους και με μια σειρά πιθανών σημασιολογικών ετικετών. Συνακόλουθα, χρησιμοποιώντας την - πλήρως συμβατή με την έως τώρα οντολογική αναπαράσταση του εννοιολογικού πλαισίου - αναπαράσταση γράφου *ARG* για μια εικόνα (η οποία περιγράφεται αναλυτικά στη δημοσίευση [10]), είμαστε σε θέση να αποθηκεύσουμε τόσο τα χαμηλού επιπέδου χαρακτηριστικά, όσο και τη σημασιολογική πληροφορία με τη μορφή ετικετών. Αν και η προσέγγιση αυτή (η χρήση, δηλαδή, δύο διαφορετικών γράφων αναπαράστασης) μπορεί να φαίνεται ασυνήθιστη εκ πρώτης όψεως, θα πρέπει να σημειωθεί ότι η χρήση της γλώσσας *RDF* για την αναπαράσταση του μοντέλου της γνώσης δε συνεπάγεται αυτόματα τη χρήση γράφων βασισμένων σε *RDF* και για την αναπαράσταση μια εικόνας στο πεδίο της ανάλυσης. Αντίθετα, η χρήση του *ARG* υποστηρίζεται ξεκάθαρα για την περίπτωση της αναπαράστασης και ανάλυσης πολυμεσικού περιεχομένου, ενώ το βασισμένο στην *RDF* μοντέλο γνώσης θεωρείται ιδανικό για την αποθήκευση και ανάκτηση πληροφορίας από μια βάση γνώσης. Το κοινό στοιχείο και των δύο αναπαραστάσεων, που είναι και αυτό που ενοποιεί και ενδυναμώνει την παρούσα προσέγγιση, είναι η χρησιμοποίηση ενός κοινού συμβολισμού ασαφών συνόλων, ο οποίος και ισχυροποιεί τις συνδέσεις μεταξύ των δύο μοντέλων γνώσης.

Για την αναπαράσταση κάθε περιοχής της εικόνας στο χώρο των χαμηλού επιπέδου χαρακτηριστικών επιλέχθηκε η χρήση δύο *MPEG-7* Περιγραφέων, του Κυρίαρχου Χρώματος και της Ομοιογενούς Υφής [164], ενώ ασαφή σύνολα υποψήφιων εννοιών χρησιμοποιήθηκαν για τη μοντελοποίηση της υψηλού επιπέδου πληροφορίας. Για το σκοπό αυτό, στα πλαίσια παλαιότερης ερευνητικής εργασίας, σχεδιάστηκε και υλοποιήθηκε ένας αλγόριθμος - υποβοηθούμενης από τη γνώση - ανάλυσης [9], η γενική

## Κεφάλαιο 6. Εννοιολογικό πλαίσιο και οπτική πληροφορία

φιλοσοφία του οποίου ενσωματώθηκε στην προσέγγισή μας. Η γενική δομή της συνολικής αρχιτεκτονικής παρουσιάζεται στο Σχήμα 6.11.



**Σχήμα 6.11:** Αρχική ανάθεση ετικετών σε περιοχές με βάση τον ARG και την αντιστοιχηση των MPEG-7 Περιγραφέων.

Στο κέντρο του Σχήματος βρίσκεται ο γράφος *ARG*, ο οποίος αλληλεπιδρά με τις υπόλοιπες διεργασίες. Ο *ARG* κατασκευάζεται με βάση μια αρχική RSST κατάτμηση [1], η οποία παράγει μερικές δεκάδες περιοχές (όπως καταδεικνύουν τα πειράματά μας, περίπου 30-40 περιοχές). Για κάθε περιοχή εξάγονται οι Περιγραφείς Κυρίαρχου Χρώματος και Ομοιογενούς Υφής (π.χ. για μια περιοχή  $a$ :  $D_a = [DC_a HT_a]$ ) και αποθηκεύονται στην αντίστοιχη κορυφή του γράφου. Όπως περιγράφεται στο [164], ο τυπικός ορισμός των δύο Περιγραφέων Χρώματος και Υφής είναι αντίστοιχα:

$$DC \equiv [\{c_i, v_i, p_i\}, s], i = 1..N \quad (6.10)$$

όπου  $c_i$  είναι το  $i$ -οστό κυρίαρχο χρώμα,  $v_i$  η διακύμανση του χρώματος,  $p_i$  η ποσοστιαία τιμή του χρώματος,  $s$  η χωρική συνοχή και το  $N$  μπορεί να φτάσει έως οκτώ. Εν γένει, η συνάρτηση απόστασης για δύο Περιγραφείς Χρώματος  $DC_1, DC_2$  είναι:

$$d_{DC}(DC_1, DC_2) = \sqrt{\sum_{i=1}^{N_1} p_{1i}^2 + \sum_{j=1}^{N_2} p_{2j}^2 - \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} 2a_{1i,2j}p_{1i}p_{2j}} \quad (6.11)$$

όπου  $a_{1i,2j}$  είναι ένας συντελεστής ομοιότητας μεταξύ δύο χρωμάτων.

Ανάλογα, για τον Περιγραφέα Υφής  $HT$ , έχουμε:

$$HT \equiv [avg, std, e_1, .., e_{30}, d_1, .., d_{30}] \quad (6.12)$$

όπου  $avg$  είναι η μέση ένταση της περιοχής,  $std$  είναι η τυπική απόκλιση της έντασης της περιοχής, ενώ  $e_1$  και  $d_1$  είναι η ενέργεια και η απόκλιση για τριάντα  $i \in [1, \dots, 30]$  κανάλια συχνότητας. Μια συνάρτηση απόστασης δύο τέτοιων Περιγραφέων ορίζεται με παρόμοιο τρόπο ως:

$$d_{HT}(HT_1, HT_2) = \sum_{i=1}^{N_{HT}=62} \left| \frac{HT_1(i) - HT_2(i)}{\sigma_i} \right| \quad (6.13)$$

όπου το  $\sigma_i$  είναι μια παράμετρος κανονικοποίησης για κάθε κανάλι συχνότητας. Για λόγους απλότητας και αναγνωστικότητας θα χρησιμοποιήσουμε από τούδε και στο

εξής το συμβολισμό  $d_{DC}(DC_a, DC_b) \equiv d_{DC}(a, b)$  (και τον αντίστοιχο για το  $d_{HT}$ ), μιας και τελικά η ερευνητική ενασχόλησή μας περιορίζεται στις περιοχές  $a$  και  $b$  μιας εικόνας, που αναπαρίστανται πλήρως από τους Οπτικούς Περιγραφείς τους.

Η απόδοση σημασιολογικών ετικετών βασίζεται σε μία διαδικασία αντιστοίχησης μεταξύ των Οπτικών Περιγραφέων, που έχουν αποθηκευθεί σε κάθε κορυφή του  $ARG$ , και των αντίστοιχων Οπτικών Περιγραφέων όλων των διαθέσιμων εννοιών  $c \in C$ , που έχουν αποθηκευτεί με τη μορφή προτότυπων στιγμιοτύπων (prototype instances)  $P(c)$  στην οντολογική βάση γνώσης. Η αντιστοίχηση μιας περιοχής  $a \in G$  (όπου  $G$  είναι το σύνολο όλων των συνδεδεμένων, μη-επικαλυπτόμενων περιοχών μιας εικόνας) με ένα πρωτότυπο στιγμιοτύπο  $p \in P(c)$  μιας έννοιας  $c \in C$  εκτελείται με το συνδυασμό των επιμέρους αποστάσεων των δύο Περιγραφέων:

$$\begin{aligned} d(a, p) &= d([DC_a HT_a], [DC_p HT_p]) \\ &= w_{DC}(c) \cdot n_{DC}(d_{DC}(a, p)) + w_{HT}(c) \cdot n_{HT}(d_{HT}(a, p)) \end{aligned} \quad (6.14)$$

όπου  $d_{DC}$  και  $d_{HT}$  δίνονται από τις εξισώσεις (6.11) και (6.13),  $w_{DC}$  και  $w_{HT}$  είναι βάρη που εξαρτώνται από την εκάστοτε έννοια  $c$  και  $w_{DC}(c) + w_{HT}(c) = 1, \forall c \in C$ . Επιπλέον, οι  $n_{DC}$  και  $n_{HT}$  είναι συναρτήσεις κανονικοποίησης και ειδικότερα έχουν επιλεχθεί έτσι, ώστε να είναι γραμμικές:

$$n(x) = \frac{x - d_{\min}}{d_{\max} - d_{\min}}, n : [d_{\min}, d_{\max}] \rightarrow [0, 1] \quad (6.15)$$

όπου  $d_{\min}$  και  $d_{\max}$  είναι το ελάχιστο και το μέγιστο των δύο συναρτήσεων  $d_{DC}$  και  $d_{HT}$ , αντίστοιχα.

Έχετε από την εξαντλητική αντιστοίχηση μεταξύ των περιοχών και όλων των προτότυπων στιγμιοτύπων, το τελευταίο βήμα είναι η δημιουργία του ασαφούς συνόλου  $L_a = \sum_{i=1}^{|C|} c_i / \mu_a(c_i)$  για όλες τις κορυφές του γράφου. Το ασαφές σύνολο  $L_a$  είναι το σύνολο όλων των υποψήφιων σημασιολογικών ετικετών για την εκάστοτε περιοχή υπό εξέταση. Ο βαθμός συμμετοχής κάθε έννοιας  $c$  στο ασαφές σύνολο  $L_a$  υπολογίζεται ως εξής:

$$\mu_a(c) = 1 - \min_{p \in P(c)} d(a, p) \quad (6.16)$$

όπου το  $d(a, p)$  δίνεται από την εξίσωση (6.14). Η διαδικασία αυτή καταλήγει στη δημιουργία ενός αρχικού ασαφούς συνόλου από σημασιολογικές ετικέτες για όλες τις περιοχές της εικόνας. Οι ετικέτες αυτές προέρχονται από τις έννοιες που βρίσκονται στη βάση γνώσης. Πρόκειται, προφανώς, για ένα μη τετριγμένο έργο, του οποίου η αποδοτικότητα εξαρτάται κατά πολύ από τη θεματική περιοχή στην οποία εφαρμόζεται, καθώς και από την ποιότητα της βάσης γνώσης. Μερικοί από τους περιορισμούς που εισάγει η προσέγγιση αυτή έχουν να κάνουν με την εξάρτησή της από τα αποτελέσματα της αρχικής κατάτμησης και τη δημιουργία των αντιπροσωπευτικών προτότυπων στιγμιοτύπων των εννοιών, οι οποίοι ούμως αφενός ξεφεύγουν από τα όρια της παρούσης διδακτορικής διατριβής και αφετέρου αντιμετωπίζονται επαρκώς στα πλαίσια της δημοσίευσης [10]. Στη συνέχεια, χρησιμοποιώντας την κλασική εκδοχή του αλγόριθμου εννοιολογικού πλαισίου και το μαθηματικό του υπόβαθρο, θα εστιάσουμε στην αναπροσαρμογή του βαθμού συμμετοχής  $\mu_a(c)$  κάθε έννοιας  $c$  στο ασαφές σύνολο  $L_a$ , το οποίο σχετίζεται με μια περιοχή  $a \in G$  της σκηνής.

#### 6.4.2 Επίδραση εννοιολογικού πλαισίου

Με βάση τη μαθηματική σημειολογία που περιγράφεται λεπτομερώς στις προηγούμενες υποενότητες (βλ. π.χ. ενότητα 3.2) και για λόγους διευκόλυνσης της αναγνωσιμότητας της ενότητας αυτής, παρουσιάζουμε συνοπτικά τα βασικά βήματα της προσαρμογής στη σημασιολογική κατάταξη του αλγορίθμου που χρησιμοποιείται, για να επαναπροσδιορίσει το βαθμό ιδιότητας μέλους  $\mu_a(c)$  κάθε έννοιας  $c$  στο ασαφές σύνολο  $L_a$ , που συνδέεται με μια περιοχή  $a \in G$  μιας εικόνας  $\eta$  και με ολόκληρη την εικόνα. Κάθε συγκεκριμένη έννοια  $k \in C$  που είναι παρούσα στην οντολογία της θεματικής περιοχής αποθηκεύεται μαζί με τους βαθμούς σχέσης της  $r_{kl}$  με οποιαδήποτε άλλη έννοια  $l \in C$  με την οποία σχετίζεται. Για να αντιμετωπίσουμε τις περιπτώσεις, όπου περισσότερες από μια έννοιες συσχετίζονται με πολλαπλές έννοιες, χρησιμοποιούμε την έννοια της “σχετικότητας πλαισίου”, δύος αυτής ορίστηκε προηγουμένως. Ακολουθείται και πάλι μια εξαντλητική προσέγγιση για τη θεματική περιοχή, δηλαδή λαμβάνονται υπόψη όλες οι πιθανές διαδρομές στο γράφο. Η εκτίμηση της τιμής κάθε έννοιας προέρχεται από τις άμεσες και τις έμμεσες σχέσεις της έννοιας με άλλες έννοιες, χρησιμοποιώντας τον τελεστή  $\max$  ως τον πλέον κατάλληλο δείκτη συμβατότητας και ως μετρική της απόστασης που ποσοτικοποιεί το σημασιολογικό συσχετισμό δύο εννοιών.

Η γενική δομή του αλγορίθμου επαναπροσδιορισμού των βαθμών ιδιότητας μέλους συνοψίζεται στα ακόλουθα βήματα:

- Προσδιορίστε μια βέλτιστη παράμετρο κανονικοποίησης  $np$ , που θα χρησιμοποιηθεί κατά τα βήματα εκτέλεσης του αλγορίθμου, κατάλληλη για την εκάστοτε εξεταζόμενη θεματική περιοχή. Η παράμετρος  $np$  αναφέρεται επίσης και ως μέτρο ομοιότητας ή ανομοιότητας των θεματικών περιοχών και  $np \in [0, 1]$ .
- Για κάθε έννοια  $k$  του ασαφούς συνόλου  $L_a$ , που συνδέεται με μια περιοχή  $a \in G$  μιας σκηνής της εικόνας με έναν βαθμό ιδιότητας μέλους  $\mu_a(k)$ , εξάγετε την πληροφορία του εννοιολογικού πλαισίου με τη μορφή των σχέσεων της έννοιας με το σύνολο οποιωνδήποτε άλλων εννοιών:  $\{r_{kl} : l \in C, l \neq k\}$ .
- Υπολογίστε το νέο βαθμό ιδιότητας μέλους  $\mu_a(k)$  της περιοχής  $a$  με βάση την παράμετρο κανονικοποίησης  $np$  και τη σχετικότητα του εννοιολογικού πλαισίου. Όταν υπάρχουν πολλαπλές σχέσεις μιας έννοιας στην οντολογία, οι οποίες συσχετίζουν την έννοια  $k$  με περισσότερες από μια έννοιες (αντί να τη συσχετίζουν αποκλειστικά και μόνο με το “στοιχείο ρίζας”  $r^e$ ), θα πρέπει να εφαρμοστεί ένα ενδιάμεσο βήμα για την  $k$ , ήτοι:  $cr_k = \max \{r_{kr^e}, \dots, r_{km}\}$ . Κατά τα γνωστά, εκφράζουμε τον υπολογισμό του  $\mu_a(k)$  με τον αναδρομικό τύπο:

$$\mu_a^n(k) = \mu_a^{n-1}(k) - np(\mu_a^{n-1}(k) - cr_k) \quad (6.17)$$

όπου το  $n$  καταδεικνύει την εκάστοτε επανάληψη. Ισοδύναμα, για μία αυθαίρετη επανάληψη:

$$\mu_a^n(k) = (1 - np)^n \cdot \mu_a^0(k) + (1 - (1 - np)^n) \cdot cr_k \quad (6.18)$$

όπου το  $\mu_a^0(k)$  αντιπροσωπεύει τον αρχικό βαθμό ιδιότητας μέλους.

Στην πράξη οι χαρακτηριστικές τιμές του  $n$  βρίσκονται μεταξύ του 3 και του 5.

Η ερμηνεία των παραπάνω εξισώσεων (6.17) και (6.18) υποδηλώνει ξεκάθαρα ότι η προτεινόμενη προσέγγιση θα ευνοήσει τους σχετικά βέβαιους βαθμούς ιδιότητας

μέλους για την έννοια μιας περιοχής σε αντιδιαστολή με τους σχετικά αβέβαιους ή παραπλανητικούς βαθμούς και έτσι θα ενισχύσει τις διαφορές τους. Από την άλλη, θα ελαττώσει την εμπιστοσύνη μας στις σαφώς παραπλανητικές έννοιες για μια συγκεκριμένη περιοχή. Για την κατάλληλη επιλογή της παραμέτρου κανονικοποίησης *pr* πραγματοποιήσαμε μια σειρά πειραμάτων σε ένα σύνολο εκπαίδευσης εικόνων για δύο θεματικές περιοχές εφαρμογής (*παραλία* και *μηχανοκίνητα σπορ*) και, όπως θα δούμε κατά την επεξήγηση των πειραματικών αποτελεσμάτων, καταλήξαμε σε *pr* = 0.15 και *pr* = 0.20 αντίστοιχα.

Συνοψίζοντας, ο προτεινόμενος αλγόριθμος επαναπροσδιορίζει κατά τρόπο σημαντικό τους αρχικούς βαθμούς ιδιότητας μέλους, χρησιμοποιώντας τη σημασιολογία με τη μορφή βασισμένων στο εννοιολογικό πλαίσιο πληροφοριών που βρίσκονται στην - για το σκοπό αυτό - κατασκευασμένη “ασαφοποιημένη” οντολογία. Στην επόμενη ενότητα συζητάμε την πειραματική οργάνωση αυτών των προσπαθειών στα πλαίσια σχετικών εφαρμογών που αναπτύξαμε και παρουσιάζουμε τόσο περιγραφικά, όσο και συνολικά αποτελέσματα.

#### 6.4.3 Διαδικασία εκτέλεσης και αποτίμησης πειραματικών μετρήσεων

Προκειμένου να παρέχουμε μία ολοκληρωμένη αξιολόγηση της εργασίας μας στο πεδίο της σημασιολογικής κατάτμησης εικόνων, πραγματοποιήσαμε πειράματα στις θεματικές περιοχές παραλία και μηχανοκίνητα σπορ, χρησιμοποιώντας ένα σύνολο δεδομένων από συνολικά 602 εικόνες (443 εικόνες παραλίας και 159 εικόνες μηχανοκίνητων σπορ), οι οποίες αντλήθηκαν είτε από το Διαδίκτυο είτε από προσωπικές συλλογές εικόνων. Κατά το στάδιο της αξιολόγησης και των δοκιμών της προσέγγισής μας, όσον αφορά στην ανοχή της στην ανακρίβεια των αρχικών σημασιολογικών ετικετών των συγκεκριμένων περιοχών της εικόνας, που προέκυψαν από τα προηγούμενα στάδια της ανάλυσης (βλ. ενότητες [6.4.1](#) και [2.3](#)), πραγματοποιήσαμε μια σειρά πειραμάτων με ένα υποσύνολο 482 εικόνων από το ανωτέρω σύνολο δεδομένων, χρησιμοποιώντας 120 εικόνες (ένα υποσύνολο της τάξεως του 20%) ως το σύνολο εκπαίδευσης για τη βέλτιστη εκτίμηση των διαφόρων παραμέτρων και κατωφλίων (όπως π.χ. το *pr*). Αποτελεί, πράγματι, κοινό τόπο [\[58\]](#), [\[65\]](#) ότι η αντικειμενικότερη αξιολόγηση της κατάτμησης μιας εικόνας περιλαμβάνει μια μέθοδο αξιολόγησης που υιοθετεί μια αντίστοιχη υπόθεση απόλυτης αληθείας (ground truth). Για αυτόν το λόγο χρησιμοποιήσαμε ένα εργαλείο δημιουργίας χαρακτηρισμών για τη χειρωνακτική κατασκευή της υπόθεσης αληθείας. Εμπειρογνώμονες του χώρου κατέβαλαν προσπάθεια, για να επιλέξουν και να σχολιάσουν το υποσύνολο των εικόνων που χρησιμοποιήθηκαν κατά τη διάρκεια των βημάτων αξιολόγησης. Προκειμένου να παρουσιασθούν οι προτεινόμενες μεθοδολογίες και να παρακολουθηθούν τα αποτελέσματα των εκάστοτε μεμονωμένων αλγορίθμων βημάτων, ενσωματώσαμε τις περιγραφόμενες τεχνικές σε μια ενιαία εφαρμογή, η οποία ενισχύθηκε περαιτέρω και με ένα γραφικό περιβάλλον αλληλεπίδρασης με τον χρήστη (βλ. Σχήμα [6.12](#)).

Στη γενική περίπτωση, μια διαδικασία αξιολόγησης αποτελεί πάντα μια ιδιαίτερα κρίσιμη ερευνητική διεργασία, καθώς ποσοτικοποιεί την αποδοτικότητα ενός αλγορίθμου, βοηθώντας έτσι στην εξαγωγή συνεπών επιστημονικών συμπερασμάτων. Με δεδομένο ότι έως τώρα στο πλαίσιο της διδακτορικής διατριβής χρησιμοποιήσαμε ευρέως και με επιτυχία τα ασαφή σύνολα, υιοθετούμε και στην περίπτωση αυτή τις ασαφείς διαδικασίες συνόλων, για να αξιολογήσουμε τα αποτελέσματα. Το τελικό προϊόν των

## Κεφάλαιο 6. Εννοιολογικό πλαίσιο και οπτική πληροφορία

δύο παραλλαγών της σημασιολογικής επέκτασης των κατατυμένων περιοχών μιας εικόνας (που παρουσιάζονται στην υπενότητα αυτή), καθώς επίσης και του αλγόριθμου του εννοιολογικού πλαισίου, ο οποίος βελτιώνει τα αρχικά αποτελέσματά της, είναι μια μάσκα κατάτμησης μαζί με ένα ασαφές σύνολο  $L_a$  για οποιαδήποτε περιοχή α που περιέχει όλες τις υποψήφιες ετικέτες αντικειμένων/περιοχών με τους βαθμούς ιδιότητας μέλους τους. Η υπόθεση αληθείας μιας οποιασδήποτε εικόνας αποτελείται από διάφορα συνδεδεμένα, μη-επικαλυπτόμενα τμήματα  $seg_i$  που συνδέονται (χειρωνακτικά) με μια μοναδική σημασιολογική ετικέτα.

Αρχικά, υπολογίζουμε την επικάλυψη κάθε περιοχής της εικόνας με κάθε τμήμα που προκύπτει από την υπόθεση αληθείας, όπως περιγράφεται στην ακόλουθη εξίσωση, όπου η ποσότητα  $A(a)$  αποτελεί ένα μέτρο του μεγέθους μιας περιοχής  $a$  και είναι ο αριθμός των στοιχείων της εικόνας (pixels) που ανήκουν σε αυτή την περιοχή:

$$overlap(a, seg_i) = \frac{A(a \cap seg_i)}{A(a)} \quad (6.19)$$

Κατόπιν, υπολογίζουμε την Dombi t-norm [78] (βλ. και Παράρτημα A):

$$T_{Dombi}(x, y) = \begin{cases} 0 & \text{αν } x = 0 \text{ ή } y = 0 \\ T_{Dombi}(x, y) & \text{αν } p = 0 \\ T_{\min}(x, y) & \text{αν } p = +\infty \\ \frac{1}{1+((\frac{1-x}{x})^p+(\frac{1-y}{y})^p)^{1/p}} & \text{αλλιώς.} \end{cases} \quad (6.20)$$

με παράμετρο  $p = 3$  της επικάλυψης και του βαθμού ιδιότητας μέλους της αντίστοιχης (ως προς την υπόθεση αληθείας) ετικέτας ως εξής:

$$score_a(c) = T_{Dombi}(overlap(a, seg_i), \mu_a(c)) \quad (6.21)$$

όπου  $c$  είναι η έννοια που χαρακτηρίζει το τμήμα  $seg_i$ . Ακολουθώντας αυτή τη διαδικασία για το  $L_a$ , υπολογίζουμε το συνολικό σκορ της περιοχής χρησιμοποιώντας την Dombi t-conorm [78] (βλ. και Παράρτημα A) για όλες τις έννοιες:

$$score_a = \perp_{Dombi}(\{score_a(c) : c \in L_a\}) \quad (6.22)$$

Εξαιτίας του αξιώματος συσχέτισης ασαφών συνόλων (associativity axiom) των t-conorms η σειρά των παραμέτρων στην εξίσωση (6.22) είναι τελείως αδιάφορη. Η εξίσωση μας δίνει ένα σκορ αξιολόγησης μιας συγκεκριμένης περιοχής, κάτι το οποίο δεν είναι τελείως άχρηστο, αλλά δεν αποτελεί και χαρακτηριστική μετρική για ολόκληρη την εικόνα. Μια τέτοια συνολική μετρική μπορεί να αποκτηθεί με την εφαρμογή μιας λειτουργίας συνάθροισης όλων των μεμονωμένων  $score_a$ , λαμβάνοντας φυσικά υπόψη το μέγεθος κάθε περιοχής:

$$score = \frac{\sum_{a \in G} (A(a) \cdot score_a)}{\sum_{a \in G} A(a)} \quad (6.23)$$

Η εξίσωση (6.23) παρέχει ένα συνολικό σκορ αξιολόγησης κατάλληλο για τους αλγορίθμους, που παρουσιάζονται στο πλαίσιο της εφαρμογής και χρήσης του εννοιολογικού πλαισίου στην σημασιολογική κατάτμηση εικόνων. Στην παρούσα υλοποίηση

(βλ. και δημοσίευση [10]) χρησιμοποιήσαμε και επεκτείναμε δύο ιδιαίτερα δημοφιλείς αλγορίθμους κατάτμησης, τους RSST [1] και watershed [30]. Η παραπάνω στρατηγική αξιολόγησης ακολουθήθηκε, όπως θα δούμε, και κατά την αποτίμηση των αποτελεσμάτων της κλασικής κατάτμησης, κάτι το οποίο θεωρείται απαραίτητο για λόγους εποικοδομητικής σύγκρισης και σωστής ερευνητικής εργασίας.

#### 6.4.4 Πειραματικές μετρήσεις

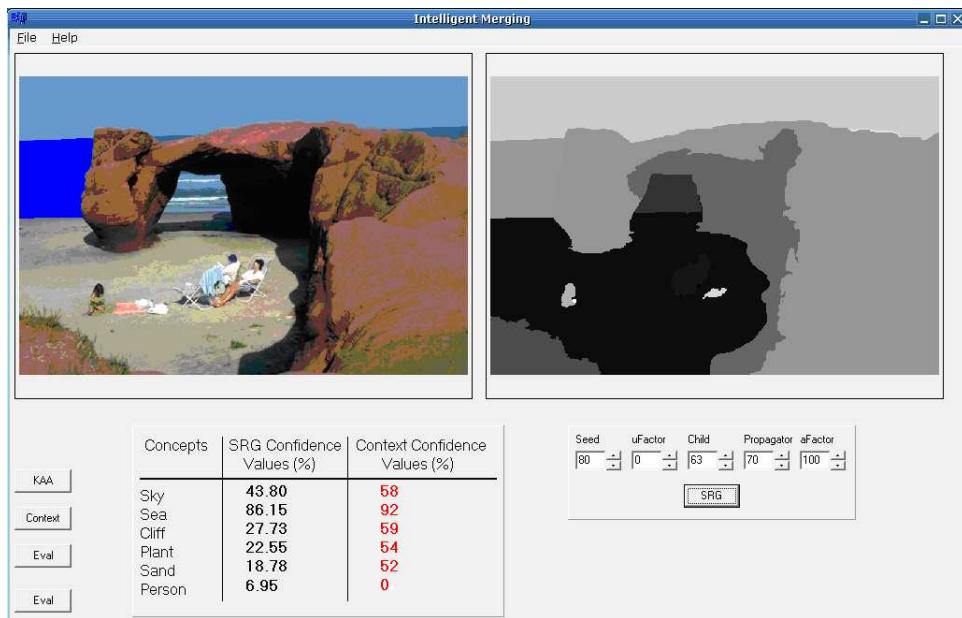
Τα αποτελέσματα που προέκυψαν από τις πειραματικές μετρήσεις που πραγματοποιήθηκαν σε ολόκληρο το σύνολο δεδομένων των εικόνων είναι ιδιαίτερα ελπιδοφόρα, ακόμη και σε περιπτώσεις όπου η ανίχνευση των συγκεκριμένων σημασιολογικών ετικετών των αντικειμένων κρίνεται ως ιδιαίτερα δύσκολη. Ως εκ τούτου, οι τελικές περιοχές της εικόνας συνδέονται με βαθμούς ιδιότητας μέλους και παρέχουν ένα ιδιαίτερα σημαντικό μέτρο της εμπιστοσύνης για την ακρίβεια της κατάτμησης. Στην υποενότητα αυτή θα παρουσιάσουμε συνοπτικά ένα χαρακτηριστικό παράδειγμα της επίδρασης του εννοιολογικού πλαισίου σε δύο αυθαίρετες εικόνες, για να τονίσουμε τόσο την προτεινόμενη προσέγγιση υπολογισμού ομοιότητας περιοχών<sup>3</sup>, όσο κυρίως τη βελτίωση που επιφέρει η χρήση του εννοιολογικού πλαισίου στη διαδικασία. Για να παρέχουμε, δε, μια επισκόπηση αξιολόγησης και για τις δύο θεματικές περιοχές εφαρμογής (παραλία και μηχανοκίνητα σπορ), παρουσιάζουμε αποτελέσματα αξιολόγησης για ολόκληρο το σύνολο δεδομένων που χρησιμοποιείται, εφαρμόζοντας τους προτεινόμενους αλγορίθμους κατάτμησης εικόνας με και χωρίς τη βελτιστοποίηση που επιφέρει το εννοιολογικό πλαίσιο.

Αρχικά, παρουσιάζουμε μια λεπτομερή απεικόνιση του βήματος της επίδρασης του εννοιολογικού πλαισίου που ακολουθείται στο πλαίσιο αυτής της εφαρμογής. Γενικά, ο αλγόριθμος του εννοιολογικού πλαισίου συμβάλλει στον επιτυχή προσδιορισμό των περιοχών στην εικόνα και διορθώνει τις παραπλανητικές συμπεριφορές, που προέρχονται από υπερβολική ή ελλιπή κατάτμηση, με το να ρυθμίζει κατάλληλα τους βαθμούς ιδιότητας μέλους τους. Όσον αφορά το *pr*, υιοθετήθηκαν ξεχωριστές τιμές για κάθε θεματική περιοχή. Με άλλα λόγια, μια τιμή του *pr* χρησιμοποιείται για τις εικόνες που ανήκουν στη θεματική περιοχή της παραλίας ( $pr_{\text{παραλίας}} = 0.15$ ) και μια διαφορετική τιμή χρησιμοποιείται κατά την εξέταση των εικόνων από τη θεματική περιοχή των μηχανοκίνητων σπορ ( $pr_{\text{μηχ.σπορ}} = 0.20$ ).

Στο Σχήμα 6.12 παραθέτουμε το βήμα της επίδρασης του οπτικού εννοιολογικού πλαισίου για μία εικόνα από τη θεματική περιοχή παραλία, έτσι όπως παρουσιάζεται μέσα από το εργαλείο ανάλυσης του εννοιολογικού πλαισίου που αναπτύχθηκε για το σκοπό αυτό. Η επίδραση του εννοιολογικού πλαισίου, η οποία κατά βάση λειτουργεί ανά περιοχή, εφαρμόζεται έπειτα από τη σημασιολογική επέκταση των περιοχών, έτσι ώστε να αποφέρει αποδοτικότερα αποτελέσματα. Στο Σχήμα 6.12 έχουμε προεπιλέξει την ενοποιημένη περιοχή της θάλασσας στο άνω αριστερά τμήμα της εικόνας, γεγονός το οποίο παριστάνεται με τεχνητό μπλε χρώμα. Τα αποτελέσματα της επίδρασης του εννοιολογικού πλαισίου παρουσιάζονται με κόκκινο χρώμα στη δεξιά στήλη στο κάτω μέρος του εργαλείου. Παρατηρούμε ότι το εννοιολογικό πλαίσιο ευνοεί έντονα το γεγονός πως η συγχωνευμένη περιοχή ανήκει στην έννοια θάλασσα, καθώς αυξάνει το βαθμό ιδιότητας μέλους της από 86.15% σε 92.00%. Ο βαθμός ιδιότητας μέλους του ατόμου, ο οποίος δε σχετίζεται με τη συγκεκριμένη περιοχή, εξαλείφεται, ενώ

<sup>3</sup>Ο αναγνώστης που επιθυμεί να εμβαθύνει στα πειραματικά αποτελέσματα της σημασιολογικής κατάτμησης μπορεί να βρει περαιτέρω πληροφορίες στη δημοσίευση [10].

## Κεφάλαιο 6. Εννοιολογικό πλαίσιο και οπτική πληροφορία



**Σχήμα 6.12:** Επίδραση του εννοιολογικού πλαισίου σε μία εικόνα παραλίας.

οι βαθμοί ιδιότητας μέλους για τις υπόλοιπες πιθανές έννοιες παραλίας αυξάνονται ελαφρώς εξαιτίας των οντολογικών σχέσεων γνώσης που υπάρχουν στο αντίστοιχο μοντέλο της γνώσης. Σε όλες, όμως, τις περιπτώσεις το εννοιολογικό πλαίσιο ομαλοποιεί τα αποτελέσματα κατά τρόπο σημαντικό, δηλαδή η κυριαρχη έννοια της εκάστοτε περιοχής ανιχνεύεται με αυξημένο βαθμό ιδιότητας μέλους.

Για να επεξηγήσουμε περαιτέρω την ενίσχυση του εννοιολογικού πλαισίου στην προσέγγισή μας, παραθέτουμε στη συνέχεια τον Πίνακα 6.3, ο οποίος παρουσιάζει τις συγχωνευμένες έννοιες περιοχών μαζί με τους βαθμούς ιδιότητας μέλους τους πριν και μετά από την ενίσχυση του οπτικού εννοιολογικού πλαισίου. Ο Πίνακας 6.3 παρέχεται, προκειμένου να συνοψιστεί η επιρροή του εννοιολογικού πλαισίου στις συγχωνευμένες περιοχές, δείχνοντας τις σημαντικές βελτιώσεις που εμφανίζονται στις περισσότερες περιπτώσεις. Η πρώτη στήλη του Πίνακα αντιπροσωπεύει την τελική συγχωνευμένη περιοχή μετά και από την εφαρμογή της σημασιολογικής προσέγγισης κατάτμησης της εικόνας<sup>4</sup>. Κάθε μια από τις επόμενες έξι στήλες εννοιών περιλαμβάνει μια διπλή τιμή: Τον βαθμό ιδιότητας μέλους χωρίς και με την ενίσχυση του εννοιολογικού πλαισίου. Τα ζευγάρια των τιμών με έντονα μαύρα γράμματα αναπαριστούν την απόλυτη αλήθεια για τη συγκεκριμένη περιοχή ή με άλλα λόγια τι απεικονίζει στην πραγματικότητα η περιοχή.

Εύκολα παρατηρεί κανείς ότι στην πλειοψηφία των περιπτώσεων το εννοιολογικό πλαίσιο βελτιώνει τα τελικά αποτελέσματα του σημασιολογικού μαρκαρίσματος από την σκοπιά της βελτίωσης του βαθμού ιδιότητας μέλους της κάθε έννοιας. Οι τιμές της απόλυτης αλήθειας για κάθε περιοχή εμφανίζονται στον παραπάνω Πίνακα, προκειμένου να παρασχεθούν άμεσα συγχριτικά αποτελέσματα στον αναγνώστη και με δεδομένο ότι αυτές είναι οι πιο ενδιαφέρουσες τιμές κατά τη διάρκεια της διαδικασίας αξιολόγησης. Παραδείγματος χάριν, κατά την εξέταση της περιοχής 0, που στην πραγματικότητα είναι μια περιοχή θάλασσας σύμφωνα με την απόλυτη αλήθεια, το εννοιολογικό πλαίσιο βελτιώνει το βαθμό ιδιότητας μέλους της έννοιας θάλασσα αυ-

<sup>4</sup>Για περισσότερες πληροφορίες υλοποίησης σχετικά με την τεχνική αυτή ο αναγνώστης καλείται να καταφύγει στη δημοσιευμένη εργασία [10])

**Πίνακας 6.3:** Τελικοί βαθμοί ιδιότητας μέλους πριν και μετά από την ενίσχυση του οπτικού εννοιολογικού πλαισίου για ένα παράδειγμα εικόνας παραλίας - οι βαθμοί με έντονα μαύρα γράμματα καταδεικνύουν την πληροφορία απόλυτης αλήθειας της εικόνας και η πρώτη στήλη δείχνει το συγκεκριμένο αριθμό της περιοχής υπό εξέταση.

Περιοχή	Έννοιες											
	ουρανός		θάλασσα		γκρεμός		φυτό		άμμος		άτομο	
	πριν	μετά	πριν	μετά	πριν	μετά	πριν	μετά	πριν	μετά	πριν	μετά
0	0.664	0.720	<b>0.818</b>	<b>0.900</b>	0.667	0.730	0.648	0.690	0.678	0.700	0.753	0.740
3	0.730	0.750	0.720	0.850	0.640	0.720	0.660	0.690	0.640	0.690	<b>0.730</b>	<b>0.750</b>
7	0.690	0.730	<b>0.781</b>	<b>0.880</b>	0.657	0.720	0.649	0.690	0.715	0.720	0.925	0.810
9	0.690	0.730	<b>0.660</b>	<b>0.830</b>	0.640	0.720	0.670	0.700	0.660	0.690	0.650	0.690
12	0.655	0.710	0.704	0.850	0.671	0.730	0.648	0.690	0.739	0.730	<b>0.810</b>	<b>0.910</b>
13	0.640	0.710	0.710	0.850	0.690	0.740	0.640	0.690	<b>0.730</b>	<b>0.750</b>	0.750	0.730
15	0.900	0.820	<b>0.770</b>	<b>0.880</b>	0.670	0.730	0.650	0.690	0.640	0.690	0.980	0.840
16	0.790	0.740	0.800	0.890	0.640	0.720	0.660	0.690	0.650	0.690	<b>0.730</b>	<b>0.750</b>
18	<b>0.840</b>	<b>0.950</b>	0.830	0.900	0.670	0.730	0.890	0.800	0.650	0.690	0.000	0.400
26	0.693	0.730	<b>0.898</b>	<b>0.930</b>	0.688	0.740	0.658	0.690	0.654	0.690	0.838	0.770
29	0.759	0.760	0.613	0.710	0.683	0.740	0.739	0.730	<b>0.703</b>	<b>0.750</b>	0.868	0.790

ξάνοντάς τον κατά 11.11%, από 0.81 σε 0.90. Ομοίως, εξετάζοντας την περιοχή 18, το εννοιολογικό πλαίσιο σημειώνει μια αύξηση 13.10% σχετικά με την πραγματική έννοια ουρανός, ενώ η περιοχή 29 παρουσιάζει μια αύξηση 6.69%, όσον αφορά το βαθμό ιδιότητας μέλους της έννοιας άμμος. Οι ανωτέρω βελτιώσεις των εννοιών της απόλυτης αλήθειας (ενδεικτικά παρατηρούμε μια συνολική βελτίωση της τάξεως του 12.55% για την έννοια θάλασσα) είναι ιδιαίτερα σημαντικές, τόσο όσον αφορά το ποσοστό της ίδιας της αύξησής, όσο και από το γεγονός ότι παρέχουν ένα βασικό εργαλείο αξιολόγησης της προτεινόμενης προσέγγισης. Σαν γενικό συμπέρασμα, είναι εμφανές ότι υπάρχει μια σαφής και υπολογίσιμη τάση βελτίωσης στις περισσότερες περιπτώσεις περιοχών, δηλαδή με άλλα λόγια η εφαρμογή του οπτικού εννοιολογικού πλαισίου έχει θετικές επιπτώσεις στη διαδικασία της σημασιολογικής κατάτυπησης εικόνων.

#### 6.4.5 Συνολικά αποτελέσματα αξιολόγησης

Στο πλαίσιο της περαιτέρω αξιολόγησης της εργασίας μας και της δοκιμής της ανοχής της στην ανακρίβεια, όσον αφορά στις αρχικές σημασιολογικές επικέτες των περιοχών, παρουσιάζουμε στη συνέχεια τα συνολικά αποτελέσματα από την εφαρμογή της προτεινόμενης μεθοδολογίας στο σύνολο των δεδομένων της θεματικής περιοχής της παραλίας. Στον Πίνακα 6.4 παρουσιάζονται συνολικά αποτελέσματα ανίχνευσης για κάθε έννοια, καθώς επίσης και το γενικό αποτέλεσμα που προέρχεται από τις έξι έννοιες της περιοχής παραλία. Παρατηρούμε ότι η έννοια ουρανός έχει το καλύτερο αποτέλεσμα μεταξύ των υπολοίπων, δεδομένου ότι το χρώμα και η σύστασή του παραμένουν σχετικά αμετάβλητα. Το οπτικό εννοιολογικό πλαίσιο βοηθά τη διαδικασία σημασιολογικού μαρκαρίσματος, δίνοντας μια συνολική οριακή βελτίωση της τάξεως του 2%, ένα γεγονός που δικαιολογείται κυρίως από την ποικιλομορφία και την ποιότητα του χρησιμοποιούμενου συνόλου εικόνων. Εκτός από τους παραπάνω παράγοντες η αποδοτικότητα του οπτικού εννοιολογικού πλαισίου εξαρτάται επίσης και από την ιδιαιτερότητα κάθε συγκεκριμένης έννοιας. Παραδείγματος χάριν, στον Πίνακα 6.4 παρατηρούμε ότι στην περίπτωση του σημασιολογικού αλγορίθμου watershed και για τις έννοιες θάλασσα και άτομο η βελτίωση που υπολογίζεται πάνω στο σύνολο των δεδομένων είναι της τάξεως του 5% και 7.2%, αντίστοιχα. Ομοίως, στην περίπτωση του σημασιολογικού RSST και για τις έννοιες θάλασσα, άμμο και άτομο βλέπουμε μια γενική αύξηση της βελτίωσης, η οποία βρίσκεται σημαντικά πάνω από το μέσο όρο του 2%, ειδικότερα: 7.2%, 7.3% και 14.6%, αντίστοιχα.

**Πίνακας 6.4:** Συνολικά και ανά έννοια αποτελέσματα ανίχνευσης για τη θεματική περιοχή της παραλίας. Η στήλη  $\alpha$  περιέχει τα αποτελέσματα του αλγορίθμου σημασιολογικής κατάτμησης watershed, η 2η στήλη τα αποτελέσματα του ίδιου αλγορίθμου με χρήση του εννοιολογικού πλαισίου, η στήλη γ τα αποτελέσματα του παραδοσιακού αλγορίθμου κατάτμησης watershed, η στήλη δ τα αποτελέσματα του σημασιολογικού RSST, η ε τα αποτελέσματα του σημασιολογικού RSST που χρησιμοποιεί το εννοιολογικό πλαίσιο και η ζ του παραδοσιακού RSST.

Έννοιες	$\alpha$	$\beta$	$\gamma$	$\delta$	$\varepsilon$	$\zeta$
ουρανός	0.95	0.94	0.93	0.93	0.93	0.90
θάλασσα	0.80	0.84	0.79	0.83	0.89	0.81
άμμος	0.79	0.81	0.72	0.82	0.88	0.72
άτομο	0.56	0.60	0.44	0.48	0.55	0.48
γκρεμός	0.75	0.77	0.61	0.76	0.78	0.66
φυτό	0.70	0.71	0.67	0.68	0.69	0.63
Σύνολο	<b>0.77</b>	<b>0.79</b>	<b>0.69</b>	<b>0.77</b>	<b>0.78</b>	<b>0.70</b>

Από την άποψη, τώρα, της υπολογιστικής πολυπλοκότητας η προσθήκη του οπτικού εννοιολογικού πλαισίου στους αλγορίθμους κατάτμησης δεν είναι ούτε υπολογιστικά ακριβή, ούτε χρονοβόρα διαδικασία. Οι μέσες χρονικές μετρήσεις για τη διαδικασία της επίδρασης του εννοιολογικού πλαισίου στο σύνολο των 355 εικόνων παραλίας καθιστούν σαφές ότι το οπτικό εννοιολογικό πλαίσιο είναι μια μάλλον γρήγορη διαδικασία, η οποία έχει ως άμεση συνέπεια τη γενική βέλτιστοποίηση των αποτελεσμάτων. Με βάση την υλοποίησή μας η αρχική κατάτμηση μιας εικόνας σε 30-40 περιοχές με βάση το χρώμα της απαιτεί περίπου 10 δευτερόλεπτα, ενώ η εξαγωγή οπτικών περιγραφέων και το αρχικό σημασιολογικό μαρκάρισμα των περιοχών χρειάζονται τον περισσότερο χρόνο, καθώς απαιτούν 60 και 30 δευτερόλεπτα, αντίστοιχα. Συγκρινόμενος με τους παραπάνω χρόνους, ο προτεινόμενος αλγόριθμος του οπτικού εννοιολογικού πλαισίου απαιτεί σημαντικά χαμηλότερο υπολογιστικό χρόνο της τάξεως του ενός δευτερολέπτου. Αξίζει επίσης να σημειωθεί ότι η προκύπτουσα επίδραση του εννοιολογικού πλαισίου επιτυγχάνει μια βέλτιστη αύξηση 13.10%, κάτι που σαφώς δικαιολογεί την αποτελεσματικότητά της στη διαδικασία της σημασιολογικής κατάτμησης εικόνων.

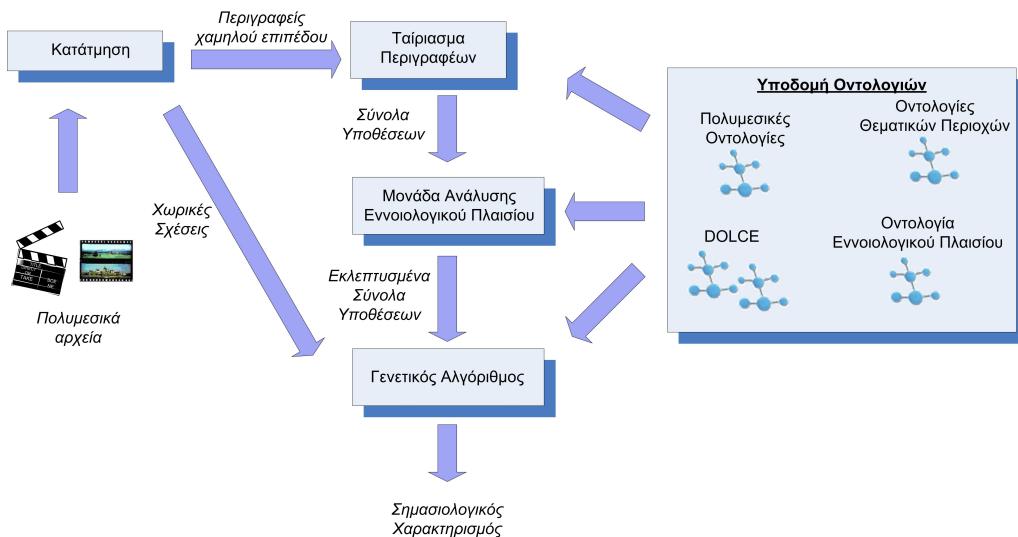
## 6.5 Οπτικό Εννοιολογικό Πλαίσιο και Γενετικός Αλγόριθμος Σημασιολογικού Χαρακτηρισμού Περιοχών

Στην τελευταία ενότητα του κεφαλαίου 6 παρουσιάζεται μια άλλη εκδοχή σημασιολογικής ανάλυσης πολυμεσικού περιεχομένου, η οποία αξιοποιεί το εννοιολογικό πλαίσιο ως μέρος της ίδιας της διαδικασίας. Σε αντίθεση με τη σημασιολογική κατάτμηση, όπου το εννοιολογικό πλαίσιο επιδρούσε στο τέλος της διαδικασίας, εδώ το εννοιολογικό πλαίσιο εισάγεται και αξιοποιείται στο μέσο της σημασιολογικής επεξεργασίας. Ειδικότερα, στην περίπτωση αυτή χρησιμοποιούνται οντολογίες, για να συλλάβουν τη γενική, χωρική και βασισμένη στο εννοιολογικό πλαίσιο γνώση μιας περιοχής, ενώ ένας γενετικός αλγόριθμος (genetic algorithm) εφαρμόζεται για να πραγματοποιήσει τον τελικό χαρακτηρισμό των περιοχών της εικόνας. Η θεωρούμενη γνώση των θεματικών περιοχών περιλαμβάνει υψηλού επιπέδου πληροφορίες με τη μορφή εννοιών ενδιαφέροντος (της υπό εξέτασης θεματικής περιοχής), πληροφορία εννοιολογικού

πλαισίου με τη μορφή ασαφών οντολογικών σχέσεων, καθώς επίσης και χαμηλού επιπέδου πληροφορίες με τη μορφή πρωτότυπων οπτικών περιγραφέων χαμηλού επιπέδου. Για να αντιμετωπίσουμε την έμφυτη ασάφεια των οπτικών πληροφοριών, εισάγουμε και εδώ την έννοια της αβεβαιότητας στο χωρικό καθορισμό των σχέσεων. Αρχικά, παράγεται ένα πρώτο σύνολο υποθέσεων, που αποτελείται από βαθμολογημένους χαρακτηρισμούς για κάθε περιοχή της εικόνας, και έπειτα αξιοποιείται το εννοιολογικό πλαίσιο για να ενημερώσει κατάλληλα τους κατ' εκτίμηση βαθμούς εμπιστοσύνης. Τέλος, ο γενετικός αλγόριθμος εφαρμόζεται, για να αποφασίσει τον πιο εύλογο χαρακτηρισμό, με τη χρησιμοποίηση των οπτικών και χωρικών ορισμών των εννοιών που περιλαμβάνονται στην οντολογία των θεματικών περιοχών. Οι πειραματικές μετρήσεις με μια συλλογή από εικόνες που ανήκουν σε δύο διαφορετικές θεματικές περιοχές (παραλία και βουνό) καταδεικνύουν την πολλά υποσχόμενη απόδοση της προτεινόμενης προσέγγισης.

### 6.5.1 Βοηθούμενη από τη γνώση ανάλυση εικόνων

Η συνολική αρχιτεκτονική του υποβοηθούμενου από τη γνώση πλαισίου ανάλυσης που ακολουθήθηκε παρουσιάζεται στο Σχήμα 6.13. Σε πρώτο στάδιο εφαρμόζεται μια



Σχήμα 6.13: Αρχιτεκτονική υποβοηθούμενου από τη γνώση πλαισίου ανάλυσης.

τεχνική κατάτμησης και κατόπιν εξάγονται χαμηλού επιπέδου περιγραφείς και χωρικές σχέσεις για τα παραγόμενα τμήματα της εικόνας. Με το που είναι διαθέσιμοι οι χαμηλού επιπέδου περιγραφείς παράγεται ένα αρχικό σύνολο υποθέσεων για κάθε τμήμα της εικόνας με βάση την απόσταση μεταξύ των εξαγομένων περιγραφέων του τμήματος και των πρωτότυπων περιγραφέων των εννοιών της θεματικής περιοχής που περιλαμβάνονται στη βάση γνώσης. Κατά συνέπεια, παράγεται ένα σύνολο από εύλογους σχολιασμούς (δηλ. έννοιες της θεματικής περιοχής) μαζί με τους αντίστοιχους βαθμούς εμπιστοσύνης τους για κάθε τμήμα της εικόνας. Αυτές οι βαθμολογημένες υποθέσεις παραδίδονται στη μονάδα ανάλυσης, η οποία, λαμβάνοντας υπόψη το εννοιολογικό πλαίσιο, τις εκλεπτύνει και τις βελτιώνει, χρησιμοποιώντας επί τούτου κατασκευασμένη και βασισμένη στο εννοιολογικό πλαίσιο γνώση με τη μορφή οντολογιών. Τα σύνολα των εκλεπτυσμένων υποθέσεων μαζί με τις χωρικές σχέσεις των τμημάτων της εικόνας μεταφέρονται στο γενετικό αλγόριθμο, ο οποίος με βάση τους

δεδομένους ορισμούς των εννοιών των θεματικών περιοχών αποφασίζει για το ποια είναι η βέλτιστη σημασιολογική τους ερμηνεία.

### 6.5.2 Κατάτμηση, εξαγωγή χαρακτηριστικών γνωρισμάτων και αρχική παραγωγή υποθέσεων

Προκειμένου να υλοποιηθεί η διαδικασία παραγωγής των αρχικών υποθέσεων, πρέπει να κατατμηθεί η υπό εξέταση εικόνα σε περιοχές και πρέπει να παραχθούν κατάλληλες χαμηλού επιπέδου περιγραφές για κάθε παραγόμενο τμήμα. Στην παρούσα υλοποίηση χρησιμοποιήθηκε μια επέκταση του αλγορίθμου RSST [179] για την κατάτμηση της εικόνας [1], ενώ επιλέχθηκαν συγκεκριμένοι χαμηλού επιπέδου περιγραφείς από το πρότυπο MPEG-7: Ομοιογενής Υφή, Σχήμα Περιοχής και Κυρίαρχο Χρώμα. Η εξαγωγή τους για κάθε μία από τις παραγόμενες περιοχές της εικόνας πραγματοποιείται σύμφωνα με τις κατευθυντήριες οδηγίες που παρέχονται από το MPEG-7 eXperimentation Model (XM) [305].

Για να εξάγουμε τα σύνολα υποθέσεων, χρειάστηκε να ορισθούν και μία σειρά από κατάλληλες μετρικές για την ποιοτική αποτίμηση της οπτικής ομοιότητας μεταξύ των υπό εξέταση τμημάτων της εικόνας και των ορισμένων πρωτότυπων εννοιών της εκάστοτε θεματικής περιοχής. Καθώς το πρότυπο MPEG-7 δεν παρέχει μια προτυποποιημένη μέθοδο για το συνδυασμό διαφορετικών αποστάσεων περιγραφέων ή για την εκτίμηση μίας μεμονωμένης απόστασης με βάση περισσότερους από έναν περιγραφείς, ακολουθήθηκε μία προσέγγιση με βάση το σταθμισμένο άθροισμα (weighted sum), η οποία και κατέληξε στον υπολογισμό μίας και μοναδικής βαθμωτής απόστασης  $D$  για κάθε υπόθεση. Με αυτό τον τρόπο παράγεται ένας βαθμός ομοιότητας  $DOC$  σε σχέση με τις ορισμένες έννοιες της θεματικής περιοχής για κάθε τμήμα ως εξής:

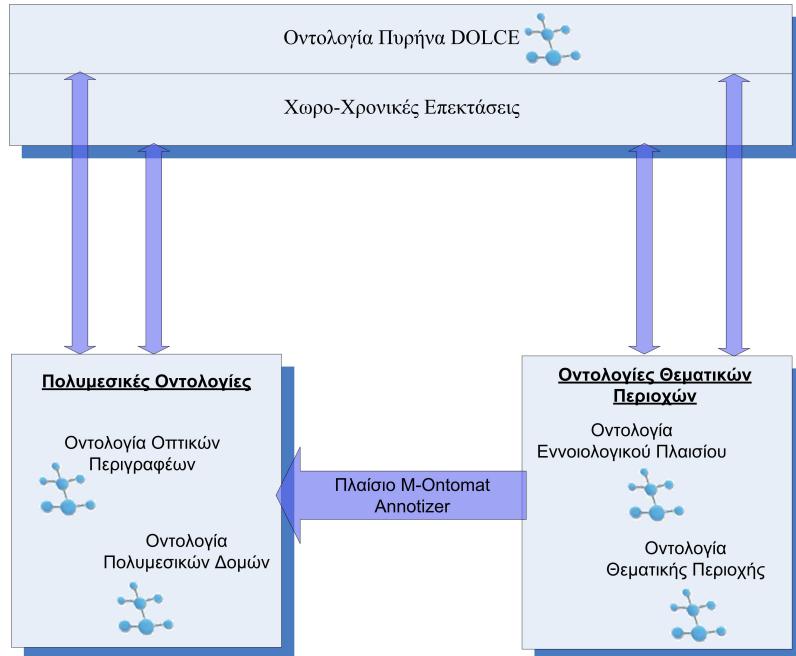
$$DOC = \frac{1}{e^{mD}} \quad (6.24)$$

όπου η παράμετρος  $m$  καθορίζεται από τις πειραματικές μετρήσεις. Τα ζεύγη των εννοιών της θεματικής περιοχής και των αντίστοιχων βαθμών εμπιστοσύνης, που προκύπτουν για κάθε τμήμα, αποτελούν το σύνολό των υποθέσεων του.

### 6.5.3 Υποδομή γνώσης

Όπως έχει ειπωθεί επανειλημμένα στα πλαίσια αυτής της διατριβής, ανάμεσα στους πολλούς δυνατούς φορμαλισμούς για την αναπαράσταση της γνώσης οι οντολογίες παρουσιάζουν έναν σημαντικό αριθμό από πλεονεκτήματα [243]. Ειδικότερα, για το συγκεκριμένο πρόβλημα παρέχουν ένα επίσημο πλαίσιο για την υποστήριξη ρητών, κατανοητών από υπολογιστές σημασιολογικών ορισμών και διευκολύνουν την εξαγωγή συμπερασμάτων και νέας γνώσης με βάση κανόνες, αλλά και προϋπάρχουσα γνώση. Με τον τρόπο αυτό οι οντολογίες είναι κατάλληλες για την έκφραση της σημασιολογίας πολυμεσικού περιεχομένου με μία κατανοητή από υπολογιστές αναπαράσταση, η οποία θα επιτρέψει την αυτόματη ανάλυση και περαιτέρω διαχείριση των εξαγομένων σημασιολογικών περιγραφών. Η χρησιμοποιούμενη οντολογία περιλαμβάνει (εκτός από την Οντολογία Εννοιολογικού Πλαισίου (Context Ontology), η δομή της οποίας αναλύθηκε στην ενότητα 6.2 της παρούσης) μία Οντολογία Πυρήνα (Core Ontology), της οποίας ο ρόλος είναι να λειτουργεί ως σημείο εκκίνησης για την κατασκευή νέων οντολογιών, μία Οντολογία Οπτικών Περιγραφέων (Visual

Descriptor Ontology), η οποία περιλαμβάνει τις αναπαραστάσεις των βασισμένων στο πρότυπο MPEG-7 περιγραφέων, μία Οντολογία Πολυμεσικών Δομών (Multimedia Structure Ontology), η οποία μοντελοποιεί μια σειρά από βασικές πολυμεσικές δομές των MPEG-7 MDS [296] καθώς και ποικίλες, εν γένει, Οντολογίες Θεματικών Περιοχών (Domain Ontologies), οι οποίες μοντελοποιούν το επίπεδο περιεχομένου του πολυμεσικού περιεχομένου, λαμβάνοντας υπόψη συγκεκριμένες θεματικές περιοχές του πραγματικού κόσμου (βλ. Σχήμα 6.14).



Σχήμα 6.14: Υποδομή Γνώσης.

#### 6.5.4 Βελτίωση εννοιολογικού πλαισίου

Εφόσον, όπως προαναφέραμε, έχει καθοριστεί η βασισμένη στο εννοιολογικό πλαίσιο δομή της γνώσης και εφαρμόζεται η αντίστοιχη αναπαράσταση που περιγράφεται στην ενότητα 6.2, μπορούμε να χρησιμοποιήσουμε, κατά τα γνωστά, μία υλοποίηση του αλγορίθμου αναπροσαρμογής των βαθμών εμπιστοσύνης. Η μορφή του τελευταίου ακολουθεί τις βασικές αρχές, που έχουν ήδη παρουσιασθεί επανειλημμένα στο κεφάλαιο αυτό (και για λόγους απλότητας δε θα επαναληφθούν στο σημείο αυτό), βασίζεται στο εννοιολογικό πλαίσιο και βοηθάει εμφανώς την ερευνητική περιοχή της ανάλυσης πολυμέσων. Η προσέγγισή μας στηρίζεται στην επίδραση του εννοιολογικού πλαισίου, ενεργεί ως ένα επιπρόσθετο βήμα επεξεργασίας πάνω από το αρχικό σύνολο υποθέσεων και επανεκτιμά τον αρχικό βαθμό εμπιστοσύνης των σημασιολογικών ετικετών για κάθε τμήμα της υπό εξέταση εικόνας. Κατά την παραπάνω διαδικασία χρησιμοποιεί πληροφορία εννοιολογικού πλαισίου, όπως αυτή έχει αποθηκευτεί στην κατασκευασμένη οντολογία του εννοιολογικού πλαισίου και μεταβιβάζει τα βελτιστοποιημένα αποτελέσματα στο γενετικό αλγόριθμο. Με τον τρόπο αυτό εκμεταλλευόμαστε στο έπακρο το εννοιολογικό πλαίσιο, το οποίο κατασκευάζεται από έναν σημασιολογικά κατάλληλο συνδυασμό επιλεγμένων ασαφών σχέσεων.

Πιο συγκεκριμένα, κάθε σημασιολογική ετικέτα σχετίζεται με μια συγκεκριμένη έννοια  $c_k$  της οντολογίας των θεματικών περιοχών εφαρμογής και είναι αποθηκευμένη

μαζί με τους βαθμούς συσχέτισής της με οποιαδήποτε άλλη σχετική έννοια. Για να αντιμετωπίσουμε τις περιπτώσεις, όπου περισσότερες από μια έννοιες σχετίζονται με πολλαπλές διαφορετικές έννοιες, εκμεταλλευόμαστε την σχετικότητα του εννοιολογικού πλαισίου  $cr_{dm}(c_k)$ , όπως αυτή ορίστηκε στην ενότητα 6.3.1 και η οποία αναφέρεται στη γενική σχετικότητα της έννοιας  $c_k$  ως προς το στοιχείο ρίζας της θεματικής περιοχής  $dm$ . Και στην περίπτωση αυτή ακολουθείται μια εξαντλητική προσέγγιση, εξετάζοντας όλες τις πιθανές διαδρομές στο σχετικό γράφο, λαμβάνοντας υπόψη το γεγονός ότι όλες οι διαδρομές μεταξύ των εννοιών είναι αμοιβαίες (reciprocal).

Αφού υπολογιστεί η τιμή της σχετικότητας του εννοιολογικού πλαισίου για κάθε έννοια και σύμφωνα με τον αλγόριθμο που περιγράφηκε στην υποενότητα 6.3.2, προσδιορίζουμε τη βέλτιστη παράμετρο κανονικοποίησης που είναι κατάλληλη για τη συγκεκριμένη θεματική περιοχή υπό εξέταση. Συνάμα καθορίζουμε την ελάχιστη τιμή που είναι σε θέση να λάβει ένας αρχικός βαθμός εμπιστοσύνης. Για κάθε σημασιολογική ετικέτα, που συνοδεύεται από έναν βαθμό εμπιστοσύνης υψηλότερο από αυτήν την ελάχιστη τιμή, εξετάζουμε την οντολογία της θεματικής περιοχής και προσδιορίζουμε την έννοια της θεματικής περιοχής που σχετίζεται με αυτή. Κατόπιν για κάθε τέτοια έννοια λαμβάνουμε τις σχετικές πληροφορίες εννοιολογικού πλαισίου (με τη μορφή των σχέσεων της έννοιας αυτής με το σύνολο οποιωνδήποτε άλλων εννοιών) και υπολογίζουμε το νέο βαθμό εμπιστοσύνης για τη σημασιολογική ετικέτα που συνδέεται με την περιοχή με βάση την παράμετρο κανονικοποίησης και την τιμή της σχετικότητας του εννοιολογικού πλαισίου. Στην περίπτωση κατά την οποία μια έννοια σχετίζεται με επιπρόσθετες έννοιες εκτός από το στοιχείο ρίζας της οντολογίας θα πρέπει να εφαρμοσθεί και ένα ενδιάμεσο βήμα συνάθροισης, για να υπολογίσει την τιμή της σχετικότητας του εννοιολογικού πλαισίου της έννοιας, όπως έχει ήδη εξηγηθεί στην υποενότητα 6.3.1.

Τα κύρια σημεία αυτής της προσέγγισης παραμένουν (i) ο προσδιορισμός των σχέσεων μεταξύ όλων των εννοιών, (ii) ο καθορισμός μιας παραμέτρου κανονικοποίησης που να έχει νόημα και (iii) ο προσδιορισμός της βέλτιστης τιμής έναρξης για τις αρχικές τιμές των βαθμών εμπιστοσύνης. Κατά την αξιολόγηση αυτών των τιμών η ιδανική παράμετρος κανονικοποίησης καθορίζεται πάντοτε, λαμβάνοντας υπόψη τη συγκεκριμένη θεματική περιοχή της γνώσης, και είναι εκείνη που ποσοτικοποιεί το σημασιολογικό συσχετισμό τους με την θεματική περιοχή. Η διαδικασία ολοκληρώνεται, όταν η εμπιστοσύνη στις αρχικές τιμές των ετικετών πάψει να είναι αρκετά υψηλή, δηλαδή δεν υπάρχουν πλέον σημασιολογικές ετικέτες με μια αποδεκτή αρχική τιμή εμπιστοσύνης που να βρίσκεται πάνω από την καθορισμένη τιμή έναρξης. Το τελικό αποτέλεσμα αυτού του βήματος επίδρασης του εννοιολογικού πλαισίου είναι η σημαντική αναπροσαρμογή των αρχικών βαθμών εμπιστοσύνης που συνοδεύουν κάθε τιμή μιας εικόνας, κάτι το οποίο αυξάνει την αποδοτικότητα και την ευρωστία της προτεινόμενης υβριδικής σημασιολογικής μεθοδολογίας ανάλυσης, καθώς και η παροχή βελτιστοποιημένων τιμών ως είσοδο στο γενετικό αλγόριθμο.

### 6.5.5 Πειραματικά αποτελέσματα

Στη συνέχεια, παρουσιάζουμε πειραματικά αποτελέσματα από τη δοκιμή της προτεινόμενης προσέγγισης πάνω σε πραγματικά δεδομένα (φωτογραφίες διακοπών) που προέρχονται από τις θεματικές περιοχές παραλία και βουνό. Κατ' αρχάς, αναπτύχθηκαν δύο ανεξάρτητες οντολογίες θεματικών περιοχών, για να αντιπροσωπεύσουν τις έννοιες των θεματικών περιοχών που μας ενδιαφέρουν. Για την περίπτωση της θε-

ματικής περιοχής παραλία χρησιμοποιήθηκαν, όπως και προηγουμένως, έξι έννοιες: ουρανός, θάλασσα, άμμος, φυτό, γκρεμός και άτομο. Εν συνεχείᾳ, επτά έννοιες: βράχος, χιόνι, έδαφος, βλάστηση, ουρανός, άτομο και νερό, καθορίστηκαν για την περίπτωση της θεματικής περιοχής βουνό.

Έχοντας κατασκευάσει τη γνώση των θεματικών περιοχών, είμαστε σε θέση να εκτελέσουμε το σημασιολογικό σχολιασμό των εικόνων. Για κάθε μια από τις υπό εξέταση εικόνες ακολουθούνται ορισμένα τυπικά βήματα χαμηλού επιπέδου ανάλυσης (τα οποία δεν εμπίπτουν στο αντικείμενο της παρούσης διδακτορικής διατριβής), όπως π.χ. η κατάτμηση, η εξαγωγή οπτικών περιγραφέων και η χωρική εξαγωγή σχέσεων. Για τα συγκεκριμένα πειράματα χρησιμοποιήθηκε ένα σύνολο κατάρτισης 200 εικόνων (100 για κάθε θεματική περιοχή), το οποίο αποτελείτο από μια πληθώρα εικόνων διακοπών που απεικόνιζαν τόσο παραλίες, όσο και βουνά. Οι εικόνες αυτές ήταν χαρακτηρισμένες “με το χέρι” και σύμφωνα με τις έννοιες που περιλαμβάνει η σχετική οντολογία των δύο θεματικών περιοχών. Εκτελέσθηκε, δε, μια μορφή κλασικής κατάτμησης (RSST) και εξήχθησαν τελικά τρεις οπτικοί MPEG-7 Περιγραφείς, δηλαδή το *Kυρίαρχο Χρώμα*, η *Ομοιογενής Υφή* και το *Σχήμα Περιοχής*. Κατόπιν, βασιζόμενοι στα πρωτότυπα στιγμιότυπα περιγραφέων (prototype descriptor instances), παρήχθησαν οι αρχικές υποθέσεις για τα υπό εξέταση τμήματα της εικόνας, οι οποίες είναι και αυτές που βελτιστοποιούνται στη συνέχεια μέσω της εφαρμογής της ανάλυσης εννοιολογικού πλαισίου που παρουσιάστηκε στην υποενότητα 6.3. Τελικά, οι ενημερωμένες βαθμολογημένες υποθέσεις περνούν στο γενετικό αλγόριθμο, που αποφαίνεται για την τελική ερμηνεία της υπό εξέταση εικόνας.

Στους Πίνακες 6.5 και 6.6 παρουσιάζονται ποσοτικά μέτρα της απόδοσης της προτεινόμενης τεχνικής για τις δύο παραπάνω θεματικές περιοχές, από την άποψη της ακρίβειας (precision) και της ανάκλησης (recall). Θα πρέπει να σημειωθεί ότι, προκειμένου να είναι ειφικτή η αριθμητική αξιολόγηση, οποιοδήποτε αντικείμενο, που ήταν παρόν στις υπό εξέταση εικόνες, αλλά που δεν είχε συμπεριληφθεί στις οντολογίες των θεματικών περιοχών, δεν λήφθηκε τελικά υπόψη. Στα Σχήματα 6.15 και 6.16 παρουσιάζονται ενδεικτικά αποτελέσματα των αρχικών εικόνων και των χαρακτηρισμών που προκύπτουν ως αποτέλεσμα της εφαρμογής, τόσο του γενετικού αλγορίθμου στις αρχικές υποθέσεις, όσο και του εννοιολογικού πλαισίου. Όπως είναι φανερό, η προτεινόμενη υβριδική προσέγγιση επιτυγχάνει ικανοποιητικά αποτελέσματα, τα οποία βελτιώνονται περαιτέρω μέσω της αξιοποίησης της γνώσης του εννοιολογικού πλαισίου. Με αυτόν τον τρόπο δικαιολογείται πλήρως η επιλογή της χρήσης ενός γενετικού αλγορίθμου, για να μεταχειριστεί τη σημασιολογική ερμηνεία μιας εικόνας ως πρόβλημα βελτιστοποίησης, καθώς επίσης και η προστιθέμενη αξία που συνεπάγεται η εισαγωγή και η χρησιμοποίηση του εννοιολογικού πλαισίου στην αλυσίδα της σημασιολογικής ανάλυσης και ερμηνείας.

Τέλος, το προτεινόμενο πλαίσιο δοκιμάστηκε αυτόνομα και στη θεματική περιοχή των καλοκαιρινών διακοπών και ειδικότερα λαμβάνοντας υπόψη τις έννοιες: θάλασσα, ουρανός, άμμος, φυτό, γκρεμός και πρόσωπο. Χρησιμοποιήθηκαν οι ακόλουθοι Περιγραφείς: *Κλιμακωτό Χρώμα*, *Ομοιογενής Υφή*, *Ιστόγραμμα Αχμών* και *Σχήμα Περιοχής*, ενώ οι υλοποιημένες ασφείς χωρικές σχέσεις περιλαμβάνουν τις κλασικές 8: αριστερά, επάνω, επάνω-αριστερά, δεξιά, κάτω, κάτω-δεξιά, κάτω-αριστερά, επάνω-δεξιά. Ενδεικτικά αποτελέσματα παρουσιάζονται στο Σχήμα 6.17, τα οποία και καταδεικνύουν την προστιθέμενη αξία της επίδρασης του εννοιολογικού πλαισίου και την απόδοση του γενετικού αλγορίθμου.

□

**Πίνακας 6.5:** Αριθμητική αποτίμηση για τη θεματική περιοχή παραλία.

Αντικείμενο	Αρχικές υποθέσεις		Εκλέπτυνση υποθέσεων		Τελική ερμηνεία	
	ακρίβεια	ανάκληση	ακρίβεια	ανάκληση	ακρίβεια	ανάκληση
ουρανός	83.33%	94.74%	92.78%	94.74%	95.79%	92.86%
θάλασσα	93.55%	87.00%	90.95%	95.50%	94.50%	90.00%
γκρεμός	51.92%	65.85%	59.02%	87.81%	82.93%	69.39%
φυτό	17.24%	50.00%	23.53%	40.00%	60.00%	33.33%
άμμος	82.69%	94.51%	89.58%	94.51%	96.70%	95.65%
άτομο	97.03%	71.02%	98.99%	71.02%	81.16%	99.12%
<b>Ορθότητα</b>	<b>82.76%</b>		<b>87.07%</b>		<b>89.66%</b>	

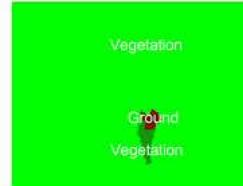
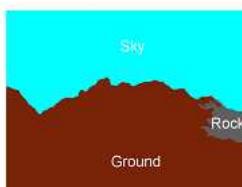
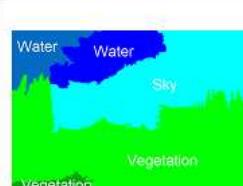
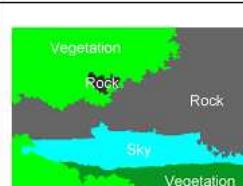
**Πίνακας 6.6:** Αριθμητική αποτίμηση για τη θεματική περιοχή βουνό.

Αντικείμενο	Αρχικές υποθέσεις		Εκλέπτυνση υποθέσεων		Τελική ερμηνεία	
	ακρίβεια	ανάκληση	ακρίβεια	ανάκληση	ακρίβεια	ανάκληση
βράχος	26.67%	28.57%	40.00%	28.57%	53.33%	57.14%
χιόνι	75.00%	60.00%	75.00%	60.00%	60.00%	60.00%
έδαφος	12.50%	50.00%	14.29%	50.00%	98.20%	99.10%
βλάστηση	87.00%	88.78	85.32%	94.90%	90.00%	91.84%
ουρανός	93.85%	85.92%	95.31%	85.92%	95.71%	94.37%
άτομο	37.50%	33.33%	33.33%	22.22%	50.00%	55.56%
νερό	60.00%	60.00%	60.00%	60.00%	100.00%	60.00%
<b>Ορθότητα</b>	<b>79.02%</b>		<b>81.46%</b>		<b>86.83%</b>	

<b>Αρχική Εικόνα</b>	<b>Αρχικές Υποθέσεις</b>	<b>Εκλέπτυνση Υποθέσεων</b>	<b>Τελικός Χαρακτηρισμός</b>

Σχήμα 6.15: Πειραματικά αποτελέσματα για τη θεματική περιοχή παραλία.

*Κεφάλαιο 6. Εννοιολογικό πλαίσιο και οπτική πληροφορία*

			
			
			
			
<b>Αρχική Ευκόνα</b>	<b>Αρχικές Υποθέσεις</b>	<b>Εκλέπτυνση Υποθέσεων</b>	<b>Τελικός Χαρακτηρισμός</b>

**Σχήμα 6.16:** Πειραματικά αποτελέσματα για τη θεματική περιοχή βουνό.



**Σχήμα 6.17:** Υπόδειγμα αποτελεσμάτων για τη θεματική περιοχή των καλοκαιρινών διαχοπών. Παρουσιάζονται η αρχική εικόνα και τα αποτελέσματα της προσέγγισης πριν ( $\beta'$  στήλη) και μετά ( $\gamma'$  στήλη) την επίδραση του εννοιολογικού πλαισίου.

# Κεφάλαιο 7

## Εννοιολογικό πλαίσιο ταξινόμησης

### 7.1 Εισαγωγή

Στο κεφάλαιο αυτό δίνουμε έμφαση στη θεώρηση και αξιοποίηση της έννοιας του εννοιολογικού πλαισίου από τη σκοπιά των συνολικών χαρακτηριστικών της εικόνας. Σε αντίθεση με το προηγούμενο κεφάλαιο, όπου εκμεταλλευθήκαμε το εννοιολογικό πλαίσιο στο επίπεδο των περιοχών μιας εικόνας, εδώ ορίζουμε αντίστοιχους αλγορίθμους, που είναι κατάλληλοι για την καθολική ανίχνευση εννοιών με βάση τα συνολικά χαρακτηριστικά της εικόνας. Ειδικότερα, επεκτείνουμε την έρευνά μας πάνω στην αναγνώριση υψηλού επιπέδου εννοιών μέσα στα πολυμεσικά έγγραφα με την εισαγωγή οντοτήτων ενδιάμεσου επιπέδου και των αντίστοιχων οντολογιών. Οι οντολογίες αυτές βασίζονται, βέβαια, στην προσέγγιση του οπτικού εννοιολογικού πλαισίου που προτείνεται στο κεφάλαιο 6, όμως η τελευταία επεκτείνεται εδώ, τόσο προς την κατεύθυνση των εννοιών (περιλαμβάνοντας καινούριες, σύνθετες και διαφορετικής μορφής καθολικές έννοιες), όσο και προς την κατεύθυνση μιας νέας μορφής οντοτήτων, των τύπων περιοχής (region types). Προτείνεται, επίσης, μία διευρυμένη προσέγγιση εννοιολογικού πλαισίου, η οποία στηρίζεται στον συνδυασμό τοπολογικών και σημασιολογικών σχέσεων και είναι κατάλληλη για την αποδοτική ανίχνευση υψηλού επιπέδου εννοιών σε εικόνες και την ταξινόμηση των τελευταίων σε κατάληξης κατηγορίες. 'Οσον αφορά στη γενική δομή της προσέγγισής μας, και στην περίπτωση αυτή κατασκευάζουμε μια σειρά από κατάλληλες οντολογίες, καθορίζουμε τις σχέσεις τους και τις ενσωματώνουμε στο μοντέλο αναπαράστασης γνώσης που προτείνουμε. 'Οπως στο προηγούμενο κεφάλαιο επιδοθήκαμε στον σχεδιασμό και την υλοποίηση αλγορίθμων για την υπολογιστικά αποδοτική διαχείριση του οπτικού εννοιολογικού πλαισίου, εδώ, εκμεταλλευόμενοι τα αποτελέσματα της βασικής έρευνας [241] στον τομέα της εξαγωγής ενδιάμεσου επιπέδου χαρακτηριστικών, εξηγούμε πώς αυτοί οι αλγόριθμοι και οι μεθοδολογίες μπορούν να συνδυαστούν αποδοτικά, για να έρθουμε λίγο πιο κοντά στην επίτευξη του υψηλότερου στόχου, αυτού δηλαδή της σημασιολογικής πολυμεσικής ανάλυσης.

Σε αυτό το πλαίσιο επεκτείνουμε την έρευνά μας για τον προσδιορισμό υψηλού επιπέδου εννοιών μέσα στα πολυμεσικά έγγραφα μέσω της εισαγωγής ενδιάμεσου επιπέδου χαρακτηριστικών γνωρισμάτων και της εκμετάλλευσης της αντίστοιχης πληροφορίας εννοιολογικού πλαισίου. 'Οπως θα δούμε αναλυτικά, η τελευταία βασίζεται στη γενική έννοια μιας οντολογίας εννοιολογικού πλαισίου και επεκτείνεται εν προκειμένω προς δύο κατευθύνσεις: αυτή των σημασιολογικών εννοιών (συμπεριλαμβανομένων καινούριων και σύνθετων σημασιολογικών εννοιών) και αυτή των τύπων περιοχής

(ενός νέου χαρακτηριστικού που εισάγεται σε αυτό το κεφάλαιο). Για να κατανοήσουμε καλύτερα την έννοια αυτών των χαρακτηριστικών ενδιάμεσου επιπέδου, δίνουμε ένα πρώτο σχετικό παράδειγμα στο Σχήμα 7.1: Θα μπορούσαμε να περιγράψουμε το οπτικό περιεχόμενο της εικόνας είτε με έναν υψηλού επιπέδου τρόπο (δηλ. η εικόνα περιέχει ουρανό, θάλασσα, άμμο και βλάστηση) ή σε ένα χαμηλότερο επίπεδο - αλλά ταυτόχρονα υψηλότερο ως προς μια εντελώς χαμηλού επιπέδου περιγραφή - (δηλ. η εικόνα περιέχει μια γαλάζια περιοχή, μια μπλε περιοχή, μια πράσινη περιοχή και μια γκρι περιοχή). Ονομάζουμε αυτά τα μεσαίου επιπέδου χαρακτηριστικά γνωρίσματα τύπους περιοχής, δεδομένου ότι κάθε εικόνα μπορεί να περιγραφεί τόσο διαισθητικά, όσο και αποτελεσματικά από ένα υποσύνολο ενός κατάλληλου συνόλου τους (λεξικό). Στη συνέχεια, είναι εξαιρετικά χρήσιμο να καθορίσουμε αυτό το σύνολο των τύπων περιοχής κατά τρόπο αποτελεσματικό, ώστε αυτό να μπορεί να περιγράψει κάθε εικόνα της θεματικής περιοχής του ενδιαφέροντός μας, έτσι ώστε να αξιοποιήσουμε αποδοτικά το σχετικό εννοιολογικό τους πλαίσιο.



Σχήμα 7.1: Μια εικόνα και η χονδροειδής κατάτμησή της.

Κατά συνέπεια, στο παρόν κεφάλαιο εισάγεται μια “ενδιάμεσου επιπέδου” γνώση ως το μέσο εκμετάλλευσης του οπτικού εννοιολογικού πλαισίου των εικόνων, από την άποψη των εννοιών υψηλού επιπέδου και των τύπων περιοχής ενδιάμεσου επιπέδου από τα οποία αποτελούνται. Ειδικότερα, σχεδιάζουμε, μοντελοποιούμε και κατασκευάζουμε τρεις τύπους ενισχυμένων οντολογιών εννοιολογικού πλαισίου, καθορίζουμε τις σχέσεις τους και τις ενσωματώνουμε σε ένα ενιαίο μοντέλο αναπαράστασης γνώσης. Κάθε ένας από τους τύπους αυτούς προτείνεται για την αντιμετώπιση και ενός παραπλήσιου ερευνητικού προβλήματος ταξινόμησης. Ειδικότερα, και όπως θα δούμε στη συνέχεια, η προσέγγιση τύπων περιοχής προτείνεται για την αποδοτική βελτίωση της βασικής έρευνας εξαγωγής ενδιάμεσου επιπέδου χαρακτηριστικών και η προσέγγιση καθολικών εννοιών έρχεται να συνδράμει απευθείας το πρόβλημα της ανίχνευσης εννοιών σε ολόκληρη την εικόνα, ενώ η μεικτή προσέγγιση υιοθετεί μια διττή μέθοδο, της οποίας το τελικό αποτέλεσμα είναι η ταξινόμηση των εικόνων στις αντίστοιχες κατηγορίες. Αξιοποιώντας και επεκτείνοντας τους αλγορίθμους αποδοτικής διαχείρισης του οπτικού εννοιολογικού πλαισίου που παρουσιάστηκαν έως τώρα, καθώς και την εξαγωγή ενδιάμεσου επιπέδου χαρακτηριστικών, είμαστε σε θέση να εξηγήσουμε πώς αυτοί οι διαφορετικοί αλγόριθμοι και μεθοδολογίες μπορούν να συνδυαστούν αποδοτικά, έτσι ώστε να επιτευχθεί ο απώτερος στόχος και των τριών μεθόδων, αυτός δηλαδή της σημασιολογικά πλούσιας και αποτελεσματικής ανάλυσης των πολυμέσων. Στο τέλος του κεφαλαίου παρουσιάζονται κάποια πειραματικά αποτελέσματα από την εφαρμογή των προτεινομένων μεθοδολογιών σε διάφορες θεματικές περιοχές προς επίρρωση των περιγραφούμενων.

### 7.1.1 Υπάρχουσα κατάσταση - Ορισμός προβλήματος

Είναι γεγονός ότι τα περισσότερα από τα σύγχρονα συστήματα ανάλυσης και ανάκτησης εικόνων με βάση το περιεχόμενο περιορίζονται από την υπάρχουσα ερευνητική κατάσταση (state-of-the-art) στο πεδίο της κατανόησης της εικόνας, με την έννοια ότι τα τελευταία υιοθετούν, συνήθως, μία σχετικά χαμηλού επιπέδου προσέγγιση και υπολείπονται, όσον αφορά στην υψηλότερου επιπέδου ερμηνεία και γνώση. Τεχνικές βασισμένες στη γνώση, στην ανθρώπινη αντίληψη και στην κατανόηση του περιεχομένου μιας σκηνής άρχισαν μόλις πρόσφατα να κάνουν δειλά-δειλά την εμφάνισή τους και να δίνουν έμφαση στη γεφύρωση του σημασιολογικού και εννοιολογικού χάσματος που υπάρχει μεταξύ των ανθρώπων και των υπολογιστών. Όπως είναι προφανές έως τώρα, η εκμετάλλευση των βασισμένων στο εννοιολογικό πλαίσιο πληροφοριών προς τη σημασιολογική κατανόηση του περιεχομένου μιας σκηνής αποτελεί κύριο στόχο της παρούσας διδακτορικής διατριβής. Το εννοιολογικό πλαίσιο αποτελεί κρίσιμο παράγοντα κατά την ανθρώπινη διαδικασία αναγνώρισης, δεδομένου ότι το ανθρώπινο οπτικό σύστημα αντίληψης εκμεταλλεύεται εκτενώς το περιβάλλον, για να διευκολύνει την οποιαδήποτε ανίχνευση αντικειμένου [33]. Κατά συνέπεια, το εννοιολογικό πλαίσιο μπορεί να χρησιμοποιηθεί, για να βελτιώσει την απόδοση τέτοιων αυτοματοποιημένων συστημάτων - υποβοηθούμενης από τη γνώση - ανάλυσης, σημασιολογικής ευρετηρίασης και ανάκτησης πολυμεσικού περιεχομένου. Σε αυτό το κεφάλαιο σκοπεύουμε να παρουσιάσουμε μια νέα προοπτική στην εκμετάλλευση του οπτικού εννοιολογικού πλαισίου, προκειμένου να επιτύχουμε καλύτερα αποτελέσματα, όσον αφορά στη σημασιολογική κατανόηση του περιεχομένου μιας σκηνής και των εννοιών που αυτή περιέχει.

Ο συνδυασμός του οπτικού εννοιολογικού πλαισίου με τεχνικές (ημι-)αυτόματης κατηγοριοποίησης εικόνων ή/και σκηνών, προκειμένου να επιτύχουμε καλύτερα σημασιολογικά αποτελέσματα κατά την ανάλυση του πολυμεσικού περιεχομένου, αποτελεί από μόνο του ένα ευρύ ερευνητικό πεδίο-πρόκληση για τον ερευνητή. Αν και το πρόβλημα του “σημασιολογικού χάσματος” (semantic gap) [232] είναι γνωστό εδώ και αρκετό καιρό, οι ερευνητικές προσεγγίσεις της ανάλυσης πολυμεσικού περιεχομένου εξακολουθούν να διαιρούνται σε δύο μάλλον διακριτές κατηγορίες: αφενός στις χαμηλού επιπέδου μεθόδους και εργαλεία ανάλυσης πολυμέσων (όπως π.χ. η [205]) και αφετέρου στις υψηλού επιπέδου μεθόδους και εργαλεία σημασιολογικού σχολιασμού (όπως π.χ. οι [259], [24]). Μόνο εξαιρετικά πρόσφατα άρχισαν κάποια σύγχρονα συστήματα ανάλυσης πολυμέσων να αξιοποιούν τις σημασιολογικές τεχνολογίες γνώσης, όπως οι τελευταίες ορίζονται από έννοιες, όπως οι οντολογίες [243], οι λαονομίες (folksonomies) [166] ή τα πρότυπα του Σημασιολογικού Ιστού και των οποίων τα πλεονεκτήματα (όταν αυτές χρησιμοποιούνται για τη δημιουργία, τη διαχείριση και τη μετα-επεξεργασία των πολυμεσικών μεταδεδομένων) είναι εμφανή σε πολυάριθμες σχετικές ερευνητικές ή μη δραστηριότητες.

Ένα από τα πιο ενδιαφέροντα προβλήματα στην ανάλυση πολυμεσικού περιεχομένου παραμένει η ανίχνευση υψηλού επιπέδου εννοιών μέσα στα πολυμεσικά έγγραφα. Αναγνωρίζοντας την ανάγκη για μια τέτοιου είδους προσέγγιση στην ανάλυση, πολλές ερευνητικές προσπάθειες εστιάζουν στην εξαγωγή χαμηλού επιπέδου χαρακτηριστικών γνωρισμάτων, έτσι ώστε να περιγράψουν αποτελεσματικά τα διάφορα οπτικοακουστικά χαρακτηριστικά ενός πολυμεσικού εγγράφου. Εντούτοις, το “σημασιολογικό χάσμα” χαρακτηρίζει συνήθως τις διαφορές μεταξύ των περιγραφών ενός αντικειμένου πολυμέσων, οι οποίες πηγάζουν από τις διαφορετικές αναπαραστάσεις του, και της σύνδεσης από τα χαμηλού- στα υψηλού επιπέδου χαρακτηριστικά γνωρίσματα.

Επιπλέον, είναι πλέον αδιαμφισβήτητο ότι η σημασιολογία κάθε αντικειμένου εξαρτάται από το εννοιολογικό πλαίσιο μέσα στο οποίο αυτό θεωρείται. Για τις εφαρμογές πολυμέσων αυτό σημαίνει ότι οποιαδήποτε επίσημη αναπαράσταση των πραγματικών στόχων ανάλυσης και επεξεργασίας απαιτεί τη μετάφραση των υψηλού επιπέδου εννοιών και σχέσεων (με τη μορφή της πολύτιμης γνώσης) στα στοιχειώδη και εκτενώς αξιολογημένα χαρακτηριστικά της χαμηλού επιπέδου ανάλυσης (όπως είναι οι οπτικές περιγραφές και τα χαμηλού επιπέδου οπτικά χαρακτηριστικά γνωρίσματα).

Ένα σημαντικό βήμα για τη γεφύρωση αυτού του χάσματος είναι η αυτοματοποίηση της διαδικασίας εξαγωγής σημασιολογικών χαρακτηριστικών γνωρισμάτων και του σχολιασμού των πολυμεσικών αντικειμένων με την ενίσχυση της διαδικασίας ταξινόμησης εικόνας και ακολουθιών βίντεο με σημασιολογικά χαρακτηριστικά. Η κύρια ιδέα που εισάγεται εν τω παρόντι στηρίζεται στην ενοποιημένη διαχείριση των εννοιών, που είναι εμφανείς μέσα στο πολυμεσικό περιεχόμενο. Η πρόσφατη πρόοδος στον ερευνητικό τομέα της βοηθούμενης από τη γνώση πολυμεσικής ανάλυσης μαζί με την εμφάνιση νέων αναπαραστάσεων περιεχομένου και μεταδεδομένων έχουν οδηγήσει όλο και περισσότερους ερευνητές στο να “κοιτάζουν” πέρα από τα απλά, χαμηλού επιπέδου χαρακτηριστικά γνωρίσματα (όπως π.χ. είναι το χρώμα, η υφή και το σχήμα), προκειμένου να καταλήξουν σε αποτελεσματικότερες υψηλού επιπέδου μεθόδους αναπαράστασης και ανάλυσης του πολυμεσικού περιεχομένου. Οι τρέχουσες, αλλά και κάποιες παλαιότερες ερευνητικές προσπάθειες έχουν στραφεί κυρίως στο συνδυασμό των χαμηλού επιπέδου περιγραφέων (οι οποίοι υπολογίζονται αυτόματα από το ακατέργαστο πολυμεσικό περιεχόμενο) και της σημασιολογίας (με τη μορφή της ανίχνευσης και εξαγωγής υψηλού επιπέδου χαρακτηριστικών γνωρισμάτων).

Η ιδέα αυτή του συνδυασμού της τυποποιημένης γνώσης και ενός συνόλου χαρακτηριστικών γνωρισμάτων, για να περιγραφεί το οπτικό περιεχόμενο μιας εικόνας, έχει παρουσιαστεί τελευταία από διάφορους ερευνητές. Παραδείγματος χάριν, στη δημοσίευση [266] παρουσιάζεται μια βασισμένη στις περιοχές προσέγγιση, η οποία χρησιμοποιεί MPEG-7 οπτικά χαρακτηριστικά γνωρίσματα και ουτολογική γνώση. Στη δημοσίευση [285] πραγματοποιείται μια προσπάθεια για την εκμετάλλευση περιορισμών χωρικού εννοιολογικού πλαισίου προς την κατεύθυνση του αυτόματου χαρακτηρισμού περιοχών. Παλαιότερα ο Lipson [158] είχε παρουσιάσει μια ανάλογη προσέγγιση μοντελοποίησης του χωρικού εννοιολογικού πλαισίου, αποκαλούμενη “μοντελοποίηση σκηνής με βάση τη διαμόρφωση” (“configuration-based scene modelling”), κατάλληλη για ευρετηρίαση και ανάκτηση με βάση το περιεχόμενο. Επιπλέον, στις [57], [234] εισάγεται μια αντίστοιχη προσέγγιση, η οποία καθοδηγείται από ένα κατάλληλο λεξικό.

Μερικές από τις εν λόγω ερευνητικές εργασίες εμπίπτουν στην κατηγορία της προσέγγισης “σάκος λέξεων” (bag-of-words). Στην περίπτωση αυτή μια εικόνα αποσυντίθεται σε ένα σύνολο “οπτικών λέξεων”, που παράγονται έπειτα από ένα βήμα συσταδοποίησης ή κατάτμησης της αρχικής εικόνας. Η μετάθεση αυτής της ιδιαίτερα δημοφιλούς μεθόδου από τη γλωσσική επεξεργασία στο “σάκο οπτικών λέξεων” ή στο “σάκο χαρακτηριστικών γνωρισμάτων” για τις εικόνες προσέδωσε στην ερευνητική κοινότητα νέες και ισχυρές μεθόδους για την περιγραφή των εικόνων. Οι δημοσιεύσεις [68], [150] και [230] προτείνουν εξαιρετικά ενδιαφέρουσες προσεγγίσεις για την ανάλυση και χρήση αυτής της μεθόδου τόσο για την ανάκτηση εικόνων με βάση το περιεχόμενο, όσο και για την ταξινόμηση σκηνών ή αντικειμένων.

Ειδικότερα, μερικές από τις πιο πρόσφατες ερευνητικές προσπάθειες εισήγαγαν ορισμένες ενδιάμεσες αναπαραστάσεις που εκμεταλλεύονται τοπικά στατιστικά στοι-

χεία στις εικόνες. Αυτές οι ενδιάμεσες αναπαραστάσεις κάνουν συχνά χρήση ανιχνευτών σημείων ενδιαφέροντος και μοντελοποιούν τις σκηνές ως μία συλλογή σημείων με ετικέτες. Οι τελευταίες προχύπτουν από ένα κατάλληλο λεξικό που κατασκευάζεται με το να κβαντοποιούνται αυτά τα σημεία ενδιαφέροντος, χρησιμοποιώντας τοπικά αλμετάβλητα χαρακτηριστικά γνωρίσματα. Για πράδειγμα, οι Fei-Fei και Perona [87] αναπαρέστησαν μια σκηνή ως μία συλλογή τοπικών τμημάτων, που τα ονόμασαν “λέξεις-κωδικες” και που συλλέχθηκαν με μία μέθοδο ανεπίβλεπτης μηχανικής μάθησης. Κάθε τμήμα αναπαραστήθηκε ως μέρος ενός “θέματος” και οι κατανομές των θεμάτων και των “λέξεων-κωδικών” πάνω στα θέματα εκπαιδεύτηκαν χωρίς κάποιου είδους επίβλεψη.

Καθώς οι παραπάνω μέθοδοι “σάκων λέξεων” δε λαμβάνουν υπόψη το χωρικό σχεδιάγραμμα των τμημάτων σε μια εικόνα, ο Lazebnik [150] χώρισε μια εικόνα σε λεπτά πλέγματα και υπολόγισε τα ιστογράμματα των τμημάτων που βρέθηκαν μέσα σε κάθε κύτταρο (cell) του πλέγματος. Η χωρική πυραμίδα που προέκυψε παρουσίασε σημαντικές βελτιώσεις σε σχέση με άλλες τέτοιοις τεχνικές, αλλά οι χωρικές πληροφορίες που κωδικοποιούνται από αυτό το μοντέλο εξακολουθούν να είναι αρκετά περιορισμένες, επειδή το τελευταίο υποθέτει ότι παρόμοια μέρη των σκηνών εμφανίζονται πολύ συχνά και σε παρόμοια κύτταρα του πλέγματος. Προκειμένου να εισάγουν εννοιολογικές πληροφορίες, ο Monay [178] χρησιμοποίησε την αναπαράσταση “σάκου λέξεων” μαζί με πιθανοτικά μοντέλα. Ένα εναλλακτικό μοντέλο για τις χωρικές πληροφορίες προτάθηκε από τον Boutil [42], ο οποίος ανέπτυξε μια συγκεκριμένη διαμόρφωση σκηνής, χρησιμοποιώντας τις περιοχές μιας εικόνας και τις ανά ζεύγη σχέσεις τους. Τέλος, ο Gemert [98] αναπαρέστησε τις σκηνές, χρησιμοποιώντας ιστογράμματα εμφάνισης (occurrence histograms).

Μια ακόμα προσέγγιση που βασίζεται στις περιοχές είναι και αυτή που παρουσιάζεται στη δημοσίευση [235] και η οποία χρησιμοποιεί τη λανθάνουσα σημασιολογική ανάλυση (LSA), ενώ στη [218] χρησιμοποιείται ένας αλγόριθμος μέσης-μετατόπισης (mean-shift algorithm) για τη συσταδοποίηση της εικόνας, προκειμένου να εξαχθούν κατόπιν οι χαμηλού επιπέδου έννοιες. Στην [126] χρησιμοποιείται μια πολύμορφη τεχνική μηχανικής μάθησης, προκειμένου να διαμορφωθούν οι σημασιολογικές έννοιες μέσα σε τηλεοπτικές ακολουθίες. Στην πιο πρόσφατη δημοσίευση [5] οι εικόνες χωρίζονται σε περιοχές, οι περιοχές συσταδοποιούνται, για να ληφθεί ένα λεξικό από τύπους περιοχής, και ακολούθως εφαρμόζεται μια προσέγγιση “σάκου λέξεων” για την αναπαράσταση της σκηνής. Στην [70] η οπική κατηγοριοποίηση επιτυγχάνεται χρησιμοποιώντας την ίδια κλασική προσέγγιση. Τέλος, στην [197] οι ερευνητές εκπαιδεύουν ξεχωριστούς ανιχνευτές σχημάτων χρησιμοποιώντας ένα σχετικό αλφαριθμητικό σχημάτων, το οποίο στην πραγματικότητα αποτελεί ένα λεξικό από τμήματα καμπυλών (curve fragments).

Στο πεδίο, τώρα, του χαρακτηρισμού εικόνων με βάση τις περιοχές μια εικόνας (region-based annotation), διάφορες ερευνητικές προσεγγίσεις υιοθετούν οπτικά χαρακτηριστικά γνωρίσματα που βασίζονται σε τμήματα είτε αυτά είναι πλέγματα (grids) [100], [135], [155], είτε κατατμημένες περιοχές [18], [132], [134], αλλά δε στοχεύουν να δημιουργήσουν τις συσχετίσεις μεταξύ των μεμονωμένων μερών και των σημασιολογικών ετικετών. Ο Fan [85] και ο Yang [283] έχουν προτείνει διάφορες προσεγγίσεις για τον χαρακτηρισμό εικόνων με βάση τις περιοχές, αρχικά με την καθιέρωση των συσχετίσεων μεταξύ των εμφανών (salient) (ή των αντιπροσωπευτικών) περιοχών και των σημασιολογικών ετικετών. Ωστόσο, τόσο οι τρέχουσες, όσο και οι προγενέστερες ερευνητικές δραστηριότητες εστιάζουν είτε σε χαμηλού-, είτε σε υψηλού επιπέδου

ερμηνείες με έναν εντελώς διακριτό τρόπο. Θεωρούμε ότι αυτού του είδους η προσέγγιση από μόνη της δεν είναι αρκετή για την αποδοτική επεξεργασία του πολυμεσικού περιεχομένου. Η πληροφορία εννοιολογικού πλαισίου με τη μορφή των συγκεκριμένων εννοιών, των αντικειμένων και των γεγονότων που απαντώνται συνήθως σε μια θεματική περιοχή (π.χ. σε μια παραλία, ένα βουνό ή μια πόλη), θα μπορούσε να αποτελέσει μια ιδιαίτερη πηγή χρήσιμων πληροφοριών [257]. Κι αυτό, διότι συνήθως εμφανίζεται ένας σημαντικός αριθμός λανθασμένων ταξινομήσεων εξαιτίας των ομοιοτήτων στα χαμηλού επιπέδου χαρακτηριστικά των διάφορων τύπων αντικειμένων και την έλλειψη μιας τέτοιας, υψηλού επιπέδου και βασισμένης στο εννοιολογικό πλαίσιο, πληροφορίας. Η διαπίστωση αυτή αποτελεί τον σημαντικότερο περιορισμό των σύγχρονων μεμονωμένων ανιχνευτών αντικειμένων. Οι γενικής φύσης αλγόριθμοι που έχουν προταθεί για την αυτόματη αναγνώριση αντικειμένων (π.χ. [18], [53], [82], [228]), τον χαρακτηρισμό σε επίπεδο περιοχών ([85], [283]) και την ταξινόμηση σκηνής [255], δεν παράγουν, δυστυχώς, ιδιαίτερα αξιόπιστα αποτελέσματα, ενώ ο περιορισμός του προβλήματος σε μια συγκεκριμένη θεματική περιοχή δεν παρέχει επίσης μια σφαιρική (και για το λόγο αυτό) ικανοποιητική λύση.

Η έννοια του οπτικού εννοιολογικού πλαισίου εισάγεται στις δημοσιεύσεις [257] και [182] ως μια πρόσθετη πηγή πληροφοριών, κατάλληλη τόσο για την ανίχνευση αντικειμένων, όσο και για την ταξινόμηση σκηνών. Η αλήθεια είναι ότι η ιδέα πίσω από τη χρήση τέτοιων πρόσθετων πληροφοριών στη διαδικασία της οπτικής ανάλυσης πολυμεσικού περιεχομένου ([182]) αναφέρεται στο αδιαμφισβήτητο ανθρώπινο γεγονός ότι δεν είναι όλα τα γεγονότα σχετικά με όλες τις καταστάσεις και κάτι τέτοιο ισχύει επίσης και κατά την εξέταση των προβλημάτων της ανάλυσης. Πρόσφατες έρευνες ([228], [40], [39]) ασχολήθηκαν με ορισμένους τύπους εννοιολογικού πλαισίου, όπως είναι το χωρικό εννοιολογικό πλαίσιο [158], [285] (δηλ. οι τοπολογικές σχέσεις μεταξύ των περιοχών στην ίδια σκηνή), το χρονικό εννοιολογικό πλαίσιο [38] (μέσα στις τηλεοπτικές ακολουθίες ή μεταξύ εικόνων που ανήκουν σε μια ιδιαίτερη συλλογή εικόνων (όπως π.χ. οι φωτογραφίες από τις διακοπές μιας οικογένειας στο Σχήμα 7.2)), ή το εννοιολογικό πλαίσιο απεικόνισης [37] (δηλ. οι ετικέτες μεταδεδομένων που συνοδεύουν τις φωτιακές φωτογραφίες και παράγονται αυτόμata από τις φωτογραφικές μηχανές (π.χ. ο χρόνος έκθεσης EXIF [290])). Οι παραπάνω τύποι εννοιολογικού πλαισίου συνδυάστηκαν από κοινού με χαρακτηριστικά γνωρίσματα του περιεχομένου της εικόνας με τη μορφή είτε χαμηλού επιπέδου χαρακτηριστικών γνωρισμάτων (όπως είναι π.χ. το χρώμα, η σύσταση και το σχήμα), είτε σημασιολογικών εννοιών (όπως είναι π.χ. οι έννοιες ουρανός, βλάστηση, άμμος, και θάλασσα). Καθίσταται εμφανές, λοιπόν, ότι η έννοια του (οπτικού) εννοιολογικού πλαισίου είναι, εν γένει, μια δύσκολη και πολύπλευρη έννοια, για να την αντιληφθεί και να την αξιοποιήσει κάποιος, και έτσι την (αυτο-)περιορίζουμε και εδώ στην έννοια του ασαφούς οντολογικού εννοιολογικού πλαισίου, όπως αυτό ορίζεται στην ενότητα 6.2. Η επιλογή μας ευθυγραμμίζεται πλήρως με τη σαφή ερευνητική τάση που υπάρχει στη βιβλιογραφία [215] προς την “ασαφοποίηση” (fuzzification) των γλωσσών περιγραφής μιας οντολογίας, όπως είναι π.χ. οι ασαφείς DL και η ασαφής OWL [244], συνεπικουρούμενη από το γεγονός ότι οι ικανότητες αναπαράστασης και εξαγωγής συμπερασμάτων που βασίζονται στην ασάφεια υπερβαίνουν σαφώς τις αντίστοιχες κλασικές προσεγγίσεις.

Η χρήση πληροφοριών ενδιάμεσου επιπέδου, όπως είναι οι πληροφορίες που λαμβάνονται από την εφαρμογή επιβλεπόμενων ή ανεπιβλεπτων μεθοδολογιών μηχανικής μάθησης για τα χαμηλού επιπέδου χαρακτηριστικά, μπορεί να χρησιμοποιηθεί, για να

## Κεφάλαιο 7. Εννοιολογικό πλαίσιο ταξινόμησης



**Σχήμα 7.2:** Παράδειγμα συλλογής φωτογραφιών με έντονο το χρονικό εννοιολογικό πλαίσιο: η πρώτη σειρά εικόνων αναφέρεται στο ίδιο γεγονός, ενώ η δεύτερη σειρά περιέχει ξεκάθαρα δύο ξεχωριστές ομάδες γεγονότων, τα οποία απέχουν μεταξύ τους χρονικά.

βελτιώσει τα αποτελέσματα της παραδοσιακής - βοηθούμενης από τη γνώση - ανάλυσης, με βάση τόσο την χαμηλού επιπέδου οπτική, όσο και την υψηλού επιπέδου "εννοιολογική" πληροφορία. Εν προκειμένω, τα αρχικά αποτελέσματα της ανάλυσης εικόνας ενισχύονται χρησιμοποιώντας σημασιολογική γνώση με τη μορφή, εν γένει, ανεξάρτητων από περιοχές, τύπων περιοχής (region types) και των σχέσεων μεταξύ τους. Σε γενικές γραμμές οι πληροφορίες ενδιάμεσου επιπέδου λαμβάνουν τη μορφή μιας ενδιάμεσης περιγραφής, η οποία μπορεί να περιγραφεί σημασιολογικά, αλλά ταυτόχρονα δεν εκφράζει τις υψηλού επιπέδου έννοιες. Έτσι, κατασκευάζεται μια σχετική οντολογία εννοιών ενδιάμεσου επιπέδου, η οποία περιέχει την παραπάνω πληροφορία. Αυτές οι πληροφορίες μπορούν να περιγραφούν ως:

- μια χαμηλού επιπέδου περιγραφή, αλλά και πάλι ένα επίπεδο πάνω από αυτή που εξάγεται απ' ευθείας από το πολυμεσικό έγγραφο
- μια υψηλού επιπέδου εννοιολογική περιγραφή, αλλά και πάλι ένα επίπεδο κάτω από τον θεωρητικό στόχο
- μια ενδιάμεση περιγραφή, η οποία μπορεί να περιγραφεί σημασιολογικά, αλλά δεν εκφράζει μια υψηλού επιπέδου έννοια

Η επέκταση της ερευνητικής προσπάθειάς μας στο κεφάλαιο αυτό εστιάζει αρχικά σε μια ολοκληρωμένη προσέγγιση που προσφέρει έναν ενοποιημένο και ανεπίβλεπτο χειρισμό του πολυμεσικού περιεχομένου. Ενεργεί συμπληρωματικά στην τρέχουσα υπάρχουσα κατάσταση και τις ερευνητικές εργασίες (state-of-the-art), δεδομένου ότι αντιμετωπίζει και τις δύο προαναφερθείσες ερευνητικές προκλήσεις. Εστιάζοντας στη σημασιολογική ανάλυση των πολυμέσων, συμβάλλει στο γεφύρωμα του χάσματος μεταξύ της σημασιολογικής και της ακατέργαστης φύσης του πολυμεσικού περιεχομένου και αντιμετωπίζει ένα από τα πιο ενδιαφέροντα προβλήματα της ανάλυσης πολυμεσικού περιεχομένου, αυτό της ανίχνευσης υψηλού επιπέδου εννοιών μέσα στα πολυμεσικά έγγραφα. Το τελευταίο επιτυγχάνεται με βάση τη σημασιολογία κάθε αντικειμένου από την άποψη της πληροφορίας του εκάστοτε οπτικού εννοιολογικού πλαισίου του. Αποδεικνύει, δε, ότι η χρήση πληροφοριών ενδιάμεσου επιπέδου βελτιώνει τα αποτελέσματα της κλασικής - βοηθούμενης από τη γνώση - ανάλυσης της εικόνας, λαμβά-

νοντας υπόψη τόσο την οπτική πληροφορία, όσο και την πληροφορία εννοιολογικού πλαισίου.

### 7.1.2 Δομή κεφαλαίου

Το κεφάλαιο αυτό δίνει έμφαση στην ενοποιημένη προσέγγιση αναπαράστασης και επεξεργασίας του πολυμεσικού περιεχομένου, συνδυάζοντας χαμηλού και υψηλού επιπέδου πληροφορίες κατά έναν πρωτότυπο “ενδιάμεσου επιπέδου” τρόπο. Σε αυτό το πλαισίο, ερευνούμε την αλληλεπίδραση μεταξύ ευφυών τεχνικών τοπικής και συνολικής ταξινόμησης και εκμεταλλευόμαστε και τους δύο τύπους της οπτικής και της βασισμένης στο εννοιολογικό πλαίσιο γνώσης, όπως αυτές περιγράφονται και λαμβάνουν χώρα στο πλαισίο μιας λογικής αλληλουχίας κατά την επεξεργασία του πολυμεσικού περιεχομένου. Τέλος, ερευνούμε περαιτέρω πιθανές προσεγγίσεις ενοποίησης, διεύρυνσης και προσαρμογής του εννοιολογικού πλαισίου τόσο προς την κατεύθυνση των υψηλού επιπέδου εννοιών, όσο και προς την κατεύθυνση των τύπων περιοχής. Η δομή του κεφαλαίου χαρακτηρίζεται από τις εξής τέσσερις διακριτές ενότητες:

- **Ο ρόλος του εννοιολογικού πλαισίου στην ανίχνευση εννοιών και στην ταξινόμηση σκηνής.** Στην ενότητα αυτή (7.2) πραγματοποιείται μια εισαγωγή του αναγνώστη στη μορφή και το ρόλο του εννοιολογικού πλαισίου κατά την αντιμετώπιση δύο συγκεκριμένων προβλημάτων της πολυμεσικής ανάλυσης, αυτών της ανίχνευσης εννοιών υψηλού επιπέδου (concept/object detection) και της ταξινόμησης σκηνής (scene classification). Τα παραπάνω δύο προβλήματα χρήζουν ιδιαίτερης ανάλυσης και μνείας κατά την επεξεργασία εικόνων, καθώς αφορούν μια καθολικής φύσεως προσέγγιση, η οποία είναι πιο προσιτή στην ανθρώπινη κατανόηση και παράγεται με εύκολο και γρήγορο τρόπο.
- **Ανάλυση τύπων περιοχής.** Στην ενότητα 7.3 παρουσιάζεται η διαδικασία ορισμού, μοντελοποίησης και υπολογισμού των μεσαίου επιπέδου χαρακτηριστικών γνωρισμάτων (τύποι περιοχής), που αξιοποιούνται στο κεφάλαιο αυτό μαζί με το εννοιολογικό τους πλαισίο προς την κατεύθυνση της συνολικά αποδοτικότερης πολυμεσικής ανάλυσης. Μελετούνται και αναλύονται οι διεργασίες κατασκευής και χρήσης του απαραίτητου οπτικού θησαυρού καθώς και των διανυσμάτων αναπαράστασης, προκειμένου αυτά να αξιοποιηθούν μετέπειτα για τον εντοπισμό υψηλού επιπέδου εννοιών. Στο σημείο αυτό αξίζει να σημειώσουμε ότι η εν λόγω διαδικασία είναι δυαδική της προσέγγισης που περιγράφεται στην ενότητα 2.3.1, μιας και προηγείται χρονικά της εφαρμογής και επίδρασης του εκάστοτε κατάλληλου εννοιολογικού πλαισίου.
- **Εννοιολογικό πλαίσιο ταξινόμησης.** Η ενότητα 7.4 μελετά και παρουσιάζει την επέκταση του προτεινόμενου εννοιολογικού πλαισίου στο “μεσαίο επίπεδο” με τη μορφή του αποκαλούμενου εννοιολογικού πλαισίου ταξινόμησης. Εγείρονται θέματα ορισμού, μοντελοποίησης και υπολογισμού του εννοιολογικού πλαισίου και προτείνονται νέες, κατάλληλες αναπαραστάσεις γνώσης και αλγόριθμοι υπολογισμού. Ειδικότερα, προτείνεται η χρήση τοπολογικών και σημασιολογικών σχέσεων και η εφαρμογή του εννοιολογικού πλαισίου τόσο κατά τη διάρκεια της ανάλυσης, όσο και ως βήμα μετα-επεξεργασίας. Ακολουθούνται, δε, τρεις διαφορετικές προσεγγίσεις ανάλογα με το είδος και τον τύπο της εκάστοτε εφαρμογής σε έννοιες υψηλού επιπέδου, τύπους περιοχής ή στη μεικτή κατάσταση.

- **Πειραματικά Αποτελέσματα.** Τα αποτελέσματα των τριών διακριτών εννοιολογικών προσεγγίσεων μαζί με τα σχετικά ερευνητικά αποτελέσματα της εκάστοτε εκδοχής του αλγορίθμου εννοιολογικού πλαισίου, που προτείνεται ειδικά για την περίπτωση ανίχνευσης εννοιών, παρουσιάζονται στο πλαίσιο της τελευταίας ενότητας 7.5 του κεφαλαίου. Για το σκοπό αυτό χρησιμοποιούνται 3 διαφορετικά σύνολα δεδομένων και χαρακτηριστικές μετρικές απόδοσης, όπως είναι η ακρίβεια (precision) και η ανάκληση (recall) (βλ. και ενότητα 2.4.3).

## 7.2 Ο Ρόλος του Εννοιολογικού Πλαισίου στην Ανίχνευση Εννοιών και στην Ταξινόμηση Εικόνων

Το οπτικό εννοιολογικό πλαίσιο διαμορφώνει μια μάλλον κλασική προσέγγιση απέναντι στο εννοιολογικό πλαίσιο, αντικειτωπίζοντάς το από τη σκοπιά των περιβαλλοντικών ή των φυσικών παραμέτρων που είναι εμφανείς στις πολυμεσικές εφαρμογές. Η προτεινόμενη αναπαράσταση εννοιολογικού πλαισίου υποστηρίζει την οπτικοακουστική πληροφορία (π.χ. τις συνθήκες φωτισμού, τις περιβαλλοντικές πληροφορίες, κ.λ.π.) και, όπως θα δούμε στη συνέχεια, αντικειτωπίζεται ξεχωριστά από κατάληλα μοντέλα οπτικού εννοιολογικού πλαισίου. Οι ερευνητικοί στόχοι του πεδίου περιλαμβάνουν την ανάλυση περιεχομένου με βάση το εννοιολογικό πλαίσιο, δηλ. την ανάλυση, λαμβάνοντας υπόψη τις εξαχθείσες/αναγνωρισμένες έννοιες κατά τη διάρκεια της ανάλυσης περιεχομένου, προκειμένου να ανιχνευθεί το συγκεκριμένο εννοιολογικό πλαίσιο, να εκφραστεί με μια δομημένη μορφή περιγραφής και να χρησιμοποιηθεί για τη βελτίωση ή τη συνέχιση των διεργασιών ανάλυσης, ευρετηρίασης και αναζήτησης, καθώς επίσης και προσωποποίησης περιεχομένου.

Στη συνέχεια αναφερόμενοι στο εννοιολογικό πλαίσιο που βασίζεται στην οπτική πληροφορία μιας εικόνας θα εννοούμε όλες εκείνες τις πληροφορίες, που είναι σχετικές με το οπτικό περιεχόμενο της σκηνής μιας ακίνητης εικόνας ή μιας ακολουθίας βίντεο και που μπορούν να φανούν χρήσιμες για την ανάλυσή της. Ανατρέχοντας στη βιβλιογραφία, ο όρος “εννοιολογικό πλαίσιο” [167] μπορεί να ερμηνευθεί και ακόμα και να ορισθεί με πολλούς διαφορετικούς τρόπους, οι οποίοι ποικίλλουν από την καθαρά φιλοσοφική διάσταση μέχρι την τελείως πρακτική [153]. Εντούτοις, μιας και, όπως είδαμε, ούτε στην αλληλουχία της πολυμεσικής ανάλυσης δεν υπάρχει μια καθολική άποψη για το εννοιολογικό πλαίσιο, είναι ιδιαίτερα σημαντικό να εδραιώσουμε μια αναπαράσταση εργασίας για αυτό, έτσι ώστε να μπορέσουμε να το εκμεταλλευτούμε και να συνεισφέρουμε στην προτεινόμενη στο κεφάλαιο αυτό ενδιάμεσου επιπέδου πολυμεσική ανάλυση. Από τα βασικά προς επίλυση προβλήματα είναι και πάλι ο τρόπος αναπαράστασης και προσδιορισμού του εννοιολογικού πλαισίου καθώς και ο τρόπος αξιοποίησής του για την βελτίωση των αποτελεσμάτων της ανάλυσης. Προφανώς, και όπως έχει ήδη εξηγηθεί αναλυτικά στο κεφάλαιο 6, τα τελευταία εξαρτώνται κατά πολύ από τη θεματική περιοχή στην οποία ανήκει μια εικόνα και κατά συνέπεια συνήθως δεν είναι επαρκή για την κατανόηση του πολυμεσικού περιεχομένου. Στη γενική περίπτωση, η έλλειψη εννοιολογικής πληροφορίας εμποδίζει σημαντικά την απόδοση της ανάλυσης [183] και μαζί με τις ομοιότητες των χαμηλού επιπέδου χαρακτηριστικών των διαφόρων τύπων αντικειμένων καταλήγουμε σε ένα σημαντικό αριθμό από λανθασμένες ταξινομήσεις.

Σύμφωνα με τα προηγούμενα, το εννοιολογικό πλαίσιο σχετίζεται άμεσα με το πρόβλημα της ταξινόμηση σκηνής και την ανίχνευση αντικειμένων. Από τη μία, η τα-

Ξινόμηση σκηνής είναι μια προσέγγιση “από επάνω προς τα κάτω” (top-down), όπου τα χαμηλού επιπέδου οπτικά χαρακτηριστικά γνωρίσματα υιοθετούνται, για να αναλύσουν συνολικά το περιεχόμενο της σκηνής και να την ταξινομήσουν σε μια εκ των προκαθορισμένων κατηγοριών (π.χ. εσωτερικός/υπαίθριος χώρος, πόλη/φύση, και λοιπά). Από την άλλη, η ανίχνευση (ή αναγνώριση) αντικειμένων είναι μια “από κάτω προς τα επάνω” (bottom-up) προσέγγιση, η οποία εστιάζει στην τοπική ανάλυση, για να ανιχνεύσει και να αναγνωρίσει τα συγκεκριμένα αντικείμενα σε περιορισμένες περιοχές μιας εικόνας, χωρίς να διαθέτει κάποια ρητή γνώση του περιβάλλοντος πλαισίου (π.χ. να αναγνωρίσει ένα κτίριο ή ένα δέντρο). Οι δύο αυτοί σημαντικοί τομείς της ανάλυσης εικόνας περιλαμβάνουν ένα πρόβλημα που μοιάζει με το γνωστό ερώτημα “η κότα έχανε το αυγό ή το αυγό την κότα;”. Κι αυτό γιατί, για παράδειγμα, η ανίχνευση ενός κτιρίου στη μέση μιας εικόνας μπορεί να υπονοήσει ότι η εικόνα αυτή είναι μια εικόνα πόλης με μια αρκετά υψηλή πιθανότητα, ενώ η προ-ταξινόμηση της εικόνας αυτής ως πόλη θα ευνοούσε την αναγνώριση ενός κτιρίου σε αντιδιαστολή με ένα δένδρο.

Εντούτοις, όπως ήδη ειπώθηκε κατ’ επανάληψη, σε όλα τα παραπάνω εμφανίζεται ένας σημαντικός αριθμός λανθασμένων ταξινομήσεων. Αυτό συμβαίνει συνήθως εξαιτίας των ομοιοτήτων, που απαντώνται στα χαρακτηριστικά χρώματος και υφής των διάφορων τύπων αντικειμένων, καθώς και εξαιτίας της παντελούς έλλειψης πληροφοριών εννοιολογικού πλαισίου, η οποία αποτελεί έναν σημαντικό περιορισμό όλων των μεμονωμένων ανιχνευτών αντικειμένων. Μια ενδιαφέρουσα προσέγγιση για την επίλυση του τελευταίου προβλήματος είναι αυτή που παρουσιάζεται στο [161]. Ειδικότερα, προτείνεται η χρήση ενός χωρικού συστήματος ανίχνευσης αντικειμένων, το οποίο μπορεί να θεωρηθεί ότι λαμβάνει υπόψη του μια μορφή εννοιολογικού πλαισίου, συνδυάζοντας αρχικά τα αποτελέσματα (output) των μεμονωμένων ανιχνευτών αντικειμένου, προκειμένου να παραχθεί ένα σύνθετο διάνυσμα πεποίθησης για τα αντικείμενα που είναι ενδεχομένως παρόντα σε μια εικόνα. Και πάλι όμως, μια τέτοια προσέγγιση από μόνη της δεν είναι αρκετή, καθώς, όπως θα δούμε παρακάτω, δεν αξιοποιεί την ιδιαίτερης σημασίας γνώση που βρίσκεται διαθέσιμη.

Αλλες ερευνητικές προσπάθειες στην περιοχή περιλαμβάνουν αυτή που προτάθηκε στη [191], όπου ένας κατάλογος σημασιολογικών αντικειμένων (μεταξύ άλλων: ουρανός, χιόνι, βράχος, νερό και δάσος) χρησιμοποιείται για τη σημασιολογική ευρετηρίαση και ανάκτηση ακολουθιών βίντεο. Επιπλέον, είναι γεγονός ότι το (χαμηλού επιπέδου) χρώμα είναι ένα από τα κεντρικά χαρακτηριστικά γνωρίσματα της υπάρχουσας έρευνας πάνω στην ανίχνευση φυσικών αντικειμένων. Παραδείγματος χάριν, στη [210] χρησιμοποιείται η ταξινόμηση χρώματος, προκειμένου να ανιχνευθεί ο ουρανός. Στο πλαίσιο της ανάκτησης εικόνων με βάση το περιεχόμενο ο Smith [233] υπέθεσε ότι ένα εκτεταμένο μπλε μπάλωμα στην κορυφή μιας εικόνας είναι πιθανό να αναπαριστά τον καθαρό ουρανό. Μια προσέγγιση που βασίζεται σε υποδείγματα (exemplars) παρουσιάστηκε πιο πρόσφατα, η οποία χρησιμοποιεί έναν συνδυασμό χαρακτηριστικών γνωρισμάτων χρώματος και υφής, για να ταξινομήσει υπό-μπλοκ σε μια υπαίθρια σκηνή ως ουρανό ή βλάστηση, υποθέτοντας έναν αρχικά σωστό προσανατολισμό της εικόνας [262]. Το τελευταίο θέτει το ζήτημα της χρησιμοποίησης των πληροφοριών προσανατολισμού ως μίας μορφής εννοιολογικό πλαίσιο στους αλγορίθμους ανίχνευσης της κατηγορίας ενός αντικειμένου, κάτι το οποίο αποφεύγεται γενικά εξαιτίας του γεγονότος ότι τέτοιες βασισμένες στο εννοιολογικό πλαίσιο πληροφορίες δεν είναι πάντα διαθέσιμες και η απόδοση των σχετικών αλγορίθμων χωρίς αυτές είναι μάλλον παραπάνω από επαρκής.

Μέχρι τώρα, καμία από τις παραπάνω μεθόδους και τεχνικές δε χρησιμοποιεί το προτεινόμενο οπτικό εννοιολογικό πλαίσιο με οποιαδήποτε μορφή. Αυτό τείνει να είναι και το κύριο μειονέκτημα των μεμονωμένων ανιχνευτών αντικειμένου, δεδομένου ότι εξετάζουν μόνο τις απομονωμένες λωρίδες των καθαρών υλικών του εκάστοτε αντικειμένου, χωρίς να λάβουν υπόψη τους το εννοιολογικό πλαίσιο της σκηνής ή των ίδιων των μεμονωμένων αντικειμένων. Αυτό είναι πολύ σημαντικό και επίσης εξαιρετικά προκλητικό ακόμη και για τους ανθρώπινους παρατηρητές. Η έννοια του οπτικού εννοιολογικού πλαισίου μπορεί σα φέρεται να βοηθήσει στην κατεύθυνση των μεθοδολογιών ανίχνευσης φυσικών αντικειμένων, μικρούμενη την ανθρώπινη προσέγγιση σε παρόμοια προβλήματα. Για παράδειγμα, πολλά υλικά αντικειμένων μπορούν να έχουν την ίδια εμφάνιση από την άποψη του χρώματος ή και της σύστασής τους, ενώ το ίδιο αντικείμενο μπορεί να έχει διαφορετικές εμφανίσεις κάτω από διαφορετικές συνθήκες απεικόνισης (π.χ. φωτισμός, μεγέθυνση). Εντούτοις, ένα σημαντικό ανθρώπινο γνώρισμα είναι ότι οι άνθρωποι εξετάζουν όλα τα αντικείμενα στη σκηνή, πριν λάβουν μια τελική απόφαση σχετικά με την ταυτότητα των μεμονωμένων αντικειμένων. Συνεπώς, η χρήση του εννοιολογικού πλαισίου στη διαδικασία της οπτικής ανάλυσης είναι εκείνη που διαμορφώνει την απαραίτητη προστιθέμενη αξία και παρέχει το κλειδί για μια τέτοια σταθερή και εύρωστη διαδικασία αναγνώρισης.

### 7.3 Ανάλυση Τύπων Περιοχής

Αρχικά, εστιάζοντας αποκλειστικά και μόνο στο οπτικό μέρος της διαδικασίας ανάλυσης εικόνων, είναι γεγονός ότι η καθολική ανίχνευση υψηλού επιπέδου εννοιών (τόσο σε εικόνες, όσο και σε ακολουθίες βίντεο) παραμένει ακόμα ένα άλυτο ερευνητικό πρόβλημα. Εντούτοις, λόγω του συνεχώς αυξανόμενου όγκου του πολυμεσικού περιεχομένου το πρόβλημα αυτό προσελκύει ολοένα και μεγαλύτερο ερευνητικό ενδιαφέρον. Οι δύο κύριες και πιο ενδιαφέρουσες πτυχές του είναι η επιλογή των χαμηλού επιπέδου χαρακτηριστικών γνωρισμάτων που θα εξαχθούν και η προσέγγιση που θα χρησιμοποιηθεί για την ανάθεση αυτών των χαμηλού επιπέδου περιγραφών στις υψηλού επιπέδου έννοιες<sup>1</sup>. 'Οπως είναι αναμενόμενο, διάφορες ερευνητικές εργασίες κινούνται προς αυτή την κατεύθυνση και τη λύση του προβλήματος. Συνοπτικά, στη δημοσίευση [226] παρουσιάζεται ένα σύστημα ανάλυσης και ανάκτησης πολυμεσικού περιεχομένου, το οποίο χρησιμοποιεί πολυτροπικές τεχνικές μηχανικής μάθησης, προκειμένου να μοντελοποιήσει τις σημασιολογικές έννοιες σε ακολουθίες βίντεο. Στη [235] παρουσιάζεται μια βασισμένη σε περιοχές προσέγγιση, η οποία χρησιμοποιεί τη λανθάνουσα σημασιολογική ευρετηρίαση (LSA). Στη [218] τα χαρακτηριστικά γνωρίσματα εξάγονται από τις κατατμημένες περιοχές της εικόνας. Επίσης, στη [266] εισάγεται μια βασισμένη στις περιοχές προσέγγιση, η οποία χρησιμοποιεί μια γνώση που κωδικοποιείται με τη μορφή μιας οντολογίας. Μια υβριδική προσέγγιση θησαυρού παρουσιάζεται στη [36], όπου πραγματοποιείται σημασιολογική αναγνώριση αντικειμένων με πεδίο εφαρμογής ακολουθίες βίντεο από δελτία ειδήσεων. Τέλος, ένα λεξικό, που χρησιμοποιείται σε μια προσέγγιση για ένα διαλογικό σύστημα ανάκτησης τηλεοπτικής πληροφορίας, παρουσιάζεται στη δημοσίευση [57].

Στις ακόλουθες υποενότητες παρουσιάζουμε αναλυτικά την προσέγγιση και υλοποίηση που πρωτοπαρουσιάστηκε στη δημοσίευση [239], προκειμένου να αντιμετωπίστε το παραπάνω πρόβλημα από μια διαφορετική και συνάμα πρωτοποριακή άποψη.

<sup>1</sup>'Η με άλλα λόγια η μελέτη του δημοφιλούς "σημασιολογικού χάσματος" [232]!

την άποψη, δηλαδή, που βασίζεται στην πληροφορία και τα χαρακτηριστικά γνωρίσματα ενδιάμεσου επιπέδου. Η ερευνητική αυτή προσπάθεια επεκτάθηκε και ενισχύθηκε περαιτέρω στις δημοσιεύσεις [186], [187], [240], [241] και [242] του υποφαινόμενου, αξιοποιώντας το εννοιολογικό πλαίσιο στη διαδικασία και επιτυγχάνοντας συγκριτικά καλύτερα ερευνητικά αποτελέσματα. Ειδικότερα, επιπρόσθετα της εισαγωγής και χρήσης των μεσαίου επιπέδου εννοιών και της δημιουργίας ενός θησαυρού από αυτές, επικεντρωθήκαμε, και θα παρουσιάσουμε στις επόμενες ενότητες του κεφαλαίου, στη δημιουργία μιας ειδικής οντολογικής αναπαράστασης γνώσης, η οποία περιέχει τόσο υψηλού επιπέδου χαρακτηριστικά γνωρίσματα (δηλ. υψηλού επιπέδου έννοιες), όσο και μεσαίου επιπέδου χαρακτηριστικά γνωρίσματα, καθώς και στην περαιτέρω εκμετάλλευσή της για τη βελτιστοποίηση της διαδικασίας της ανάλυσης.

Στη γενική περίπτωση, τα οπτικά χαρακτηριστικά γνωρίσματα που μπορούν να εξαχθούν από μια εικόνα ή ένα πολυμεσικό έγγραφο μπορεί να διαιρεθούν σε δύο σημαντικές κατηγορίες. Η πρώτη περιέχει τα χαμηλού επιπέδου οπτικά χαρακτηριστικά γνωρίσματα, τα οποία μπορούν να παρέχουν μια ποιοτική ή ποσοτική περιγραφή των οπτικών ιδιοτήτων. Συχνά αυτά τα χαρακτηριστικά γνωρίσματα προτυποποιούνται με τη μορφή ενός Οπτικού Περιγραφέα (ή απλά Περιγραφέα). Η δεύτερη κατηγορία περιέχει τα υψηλού επιπέδου χαρακτηριστικά γνωρίσματα, τα οποία περιγράφουν το οπτικό περιεχόμενο μιας εικόνας από την άποψη της σημασιολογίας του με τη μορφή σημασιολογικών εννοιών (ή απλά εννοιών). Μια θεμελιώδης διαφορά μεταξύ των κατηγοριών αυτών είναι ότι τα χαμηλού επιπέδου χαρακτηριστικά γνωρίσματα μπορούν να υπολογιστούν άμεσα από μια εικόνα ή ένα βίντεο, ενώ τα υψηλού επιπέδου χαρακτηριστικά γνωρίσματα δεν μπορούν να εξαχθούν άμεσα, αλλά καθορίζονται συνήθως με την κατάλληλη αξιοποίηση των χαμηλού επιπέδου χαρακτηριστικών γνωρισμάτων. Κατά συνέπεια, το πρόβλημα αυτό, το οποίο, όπως προείπαμε πλειστάκις, αναφέρεται συχνά ως το “σημασιολογικό χάσμα” [232], προσελκύει μεγάλο ενδιαφέρον εντός της ερευνητικής κοινότητας.



Σχήμα 7.3: Δεύτερο παράδειγμα αρχικής εικόνας και της χονδροειδούς κατάτμησής της.

Από αυτή την άποψη, στο κεφάλαιο αυτό προσπαθούμε να ενισχύσουμε την έννοια της οντολογίας του οπτικού εννοιολογικού πλαισίου με ενδιάμεσου επιπέδου έννοιες. Οι έννοιες αυτές μπορούν να παρέχουν μια ενδιάμεση περιγραφή, η οποία μπορεί να περιγραφεί σημασιολογικά, αλλά δεν εκφράζουν ούτε μια υψηλού, ούτε μια χαμηλού επιπέδου έννοια. Κατά συνέπεια, δίνουμε έμφαση σε μια ενοποιημένη πολυμεσική αναπαράσταση, συνδυάζοντας χαμηλού και υψηλού επιπέδου πληροφορίες κατά ένα αποδοτικό “μεσαίου επιπέδου” τρόπο και τη συνδέουμε με την οντολογία εννοιολογικού πλαισίου με τον καθορισμό ορισμένων κατάλληλων σχέσεων. Για την καλύτερη κατανόηση αυτών των εννοιών ενδιάμεσου επιπέδου παρουσιάζουμε ακόμα ένα σχετικό παράδειγμα στο Σχήμα 7.3. Αν θέλαμε να περιγράψουμε το οπτικό περιεχόμενο της εικόνας με έναν υψηλού επιπέδου τρόπο, θα λέγαμε απευθείας ότι η

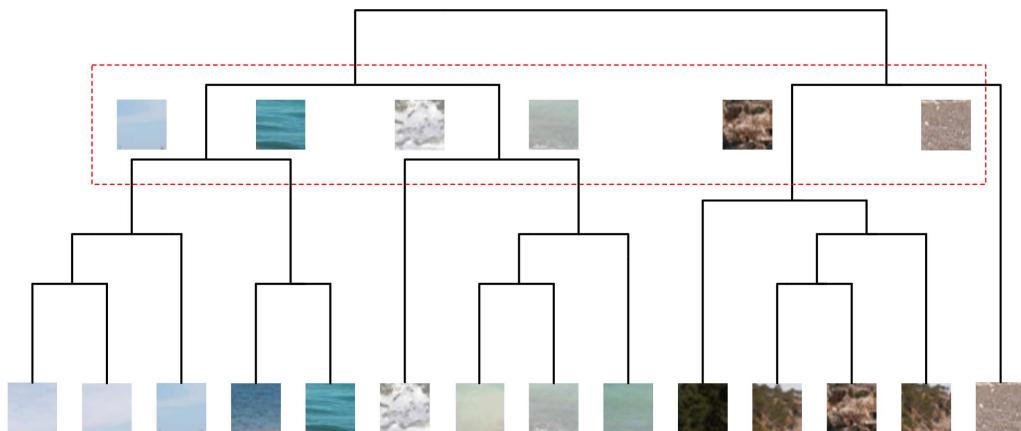
εικόνα περιέχει ουρανό, θάλασσα, άμμο και βλάστηση. Η περιγραφή της ίδιας εικόνας σε ένα χαμηλότερο επίπεδο, αλλά συγχρόνως υψηλότερο από μια εντελώς χαμηλού επιπέδου περιγραφή θα κατέληγε ότι η εικόνα αποτελείται από τέσσερις περιοχές: μια “γαλάζια”, μια “μπλε”, μια “πράσινη” και μια “γκρι” περιοχή. Ονομάσαμε αυτά τα ενδιάμεσου επιπέδου χαρακτηριστικά γνωρίσματα τύπους περιοχής (region types), δεδομένου ότι πιστεύουμε πως κάθε εικόνα μπορεί διαισθητικά ή και αποτελεσματικά να περιγραφεί από ένα σύνολο τους. Κατά συνέπεια, είναι ιδιαίτερα σημαντικό να καθορίσουμε αυτό το σύνολο τύπων περιοχής κατά ένα αποτελεσματικό τρόπο, έτσι ώστε αυτός να μπορεί να περιγράψει αποτελεσματικά οποιαδήποτε εικόνα μιας θεματικής περιοχής.

Το πρώτο βήμα για την κατασκευή της ενδιάμεσου επιπέδου οντολογίας, που θα αξιοποιεί το εννοιολογικό πλαίσιο, είναι η επιλογή των τύπων περιοχής που θα συμπεριλάβει. Κατά συνέπεια, απαιτείται ένας αυθαίρετα μεγάλος αριθμός υποψηφίων τύπων περιοχής. Αρχικά, ως ένα βήμα προ-επεξεργασίας, εφαρμόζεται ένας αλγόριθμος κατάτμησης χρώματος σε όλες τις εικόνες του διαθέσιμου συνόλου εκπαίδευσης, προκειμένου να συλλέξουμε τους κατάλληλους τύπους περιοχής. Στην παρούσα φάση, υιοθετήσαμε μια επέκταση της γνωστής μεθόδου κατάτμησης RSST [11], έτσι ώστε η τελευταία να παράγει μια αρχική χονδροειδή κατάτμηση των εικόνων σε περιοχές. Επίσης, επιλέξαμε αυτήν την προσέγγιση της χονδροειδούς κατάτμησης, δεδομένου ότι θέλουμε την παραχθείσα κατάτμηση, για να παρέχουμε διαισθητικά και μόνο μια ποιοτική περιγραφή της εκάστοτε εικόνας (βλ. Σχήμα 7.3). Στη συνέχεια, τα αποτελέσματα της κατάτμησης χρησιμοποιούνται, για να καθορίσουν τους υποψηφίους τύπους περιοχής από κάθε εικόνα.

Οι παραπάνω περιοχές αναπαριστώνται από έναν συνδυασμό των χαμηλού επιπέδου οπτικών χαρακτηριστικών γνωρισμάτων τους. Με αυτόν τον τρόπο, οι οπτικοί περιγραφείς από το πρότυπο MPEG-7 [62] επιλέγονται, για να συλλάβουν μια τυποποιημένη περιγραφή του οπτικού περιεχομένου των περιοχών. Για την αναπαράσταση των χαρακτηριστικών γνωρισμάτων χρώματος των περιοχών της εικόνας χρησιμοποιούνται συνήθως MPEG-7 Περιγραφείς Χρώματος, ενώ για την αναπαράσταση των χαρακτηριστικών υφής χρησιμοποιείται ένας κατάλληλος Περιγραφέας Υφής. Για την εξαγωγή των προαναφερθέντων περιγραφών ακολουθείται συνήθως η υλοποίηση eXperimentation Model (XM)[180] του προτύπου MPEG-7.

Λαμβάνοντας υπόψη ολόκληρο το σύνολο των περιοχών που προέρχεται από την προαναφερθείσα διαδικασία κατάτμησης σε όλες τις εικόνες του συνόλου εκπαίδευσης, καθώς και τα εξαχθέντα χαμηλού επιπέδου χαρακτηριστικά γνωρίσματά τους, μπορούμε εύκολα να παρατηρήσουμε δύο στοιχεία: (α) οι περιοχές εκείνες που ανήκουν σε παρόμοιες σημασιολογικές έννοιες, έχουν επίσης και παρόμοιες χαμηλού επιπέδου περιγραφές και (β) οι εικόνες που περιέχουν τις ίδιες υψηλού επιπέδου έννοιες αποτελούνται από παρόμοιες περιοχές. Σαν φυσική ακολουθία αυτής της παρατήρησης εφαρμόζουμε τον αλγόριθμο ιεραρχικής συσταδοποίησης, που περιγράφηκε αναλυτικά στο κεφάλαιο 4 στις περιγραφές μιας μικρής μερίδας του γενικού συνόλου δεδομένων, η οποία χρησιμοποιείται ως σύνολο εκπαίδευσης. Επιλέξαμε τη γνωστή πλέον προσέγγιση της ιεραρχικής συσταδοποίησης, δεδομένου ότι με αυτό τον τρόπο μπορούμε εύκολα να τροποποιήσουμε το μέγεθος του παραγόμενου θησαυρού (βλ. υποενότητα 7.3.2). Θα πρέπει, επίσης, να σημειώσουμε εδώ ότι μετά από την εφαρμογή της παραπάνω διαδικασίας κάθε συστάδα μπορεί ή δεν μπορεί να αναπαριστά ένα υψηλού επιπέδου χαρακτηριστικό γνώρισμα και ότι κάθε υψηλού επιπέδου χαρακτηριστικό γνώρισμα μπορεί να αντιπροσωπεύεται από μία ή και περισσότερες συστάδες. Με

άλλα λόγια, η έννοια άμμος μπορεί να έχει πολλά διαφορετικά στιγμιότυπα, όσον αφορά π.χ. στο χρώμα της. Επιπλέον, σε μια συστάδα που μπορεί να περιέχει διάφορες περιπτώσεις από μια σημασιολογική έννοια (π.χ. θάλασσα), ενδέχεται αυτές οι περιπτώσεις να μπορούσαν να αναμειχθούν με μέρη από μια άλλη οπτικά παρόμοια έννοια (όπως π.χ. είναι η έννοια ουρανός). Ένα κατάλληλο δενδροδιάγραμμα που περιγράφει την ιεραρχική συσταδοποίηση και την επιλογή των τύπων περιοχής απεικονίζεται στο Σχήμα 7.4. Σε αυτό το απλοϊκό παράδειγμα συσταδοποιείται ένα αρχικό σύνολο 14 υποψηφίων τύπων περιοχής και επιλέγουμε να κρατήσουμε 6 τύπους περιοχής, για να αναπαραστήσουμε το οπτικό περιεχόμενό τους από την άποψη των ενδιάμεσου επιπέδου εννοιών.



**Σχήμα 7.4:** Επιλογή τύπων περιοχής μέσω ιεραρχικής συσταδοποίησης. Οι έξι επιλεγμένοι τύποι περιοχής καταδεικνύονται με το κόκκινο περίγραμμα.

### 7.3.1 Ανίχνευση εννοιών με χρήση οπτικού θησαυρού

Στη συνέχεια, διαμορφώνουμε έναν κατάλληλο θησαυρό περιοχών (region thesaurus), προκειμένου να συνδυάσουμε και να χειριστούμε αποτελεσματικά το σχετικό κατάλογο κάθε τύπου περιοχής σε μια δεδομένη θεματική περιοχή της γνώσης (όπως π.χ. είναι η παραλία) και ενός συνόλου σχετικών περιοχών (συνώνυμα) για κάθε τύπο περιοχής. Όπως είδαμε, αυτοί οι τύποι περιοχής μπορούν να χαρακτηριστούν ως “ενδιάμεσου επιπέδου” έννοιες, οι οποίες ενσωματώνουν τόσο χαμηλού, όσο και υψηλού επιπέδου πληροφορίες. Κατόπιν, χρησιμοποιούμε το θησαυρό αυτό, για να διευκολύνουμε την ένωση των χαμηλού επιπέδου χαρακτηριστικών γνωρισμάτων της εικόνας με τις αντίστοιχες υψηλού επιπέδου έννοιες, διαμορφώνοντας ένα διάνυσμα αναπαράστασης (model vector) για κάθε εικόνα. Η διάσταση του διανύσματος αυτού είναι ίση με τον αριθμό των εννοιών που αποτελούν τον ίδιο το θησαυρό. Η απόσταση μιας περιοχής από έναν τύπο περιοχής υπολογίζεται ως ένας γραμμικός συνδυασμός των μέσων αποστάσεων των περιγραφέων, όπως αυτός ορίζεται στη δημοσίευση [237]. Υπολογίζοντας την απόσταση κάθε περιοχής της εικόνας από όλους τους τύπους περιοχής του κατασκευασμένου θησαυρού, το διάνυσμα αναπαράστασης  $D_m$  που περιγράφει σημασιολογικά το οπτικό περιεχόμενο της εικόνας διαμορφώνεται διατηρώντας τον υψηλότερο βαθμό εμπιστοσύνης για κάθε μεσαίου επιπέδου έννοια. Κάτι τέτοιο απεικονίζεται στην εξίσωση (7.1):

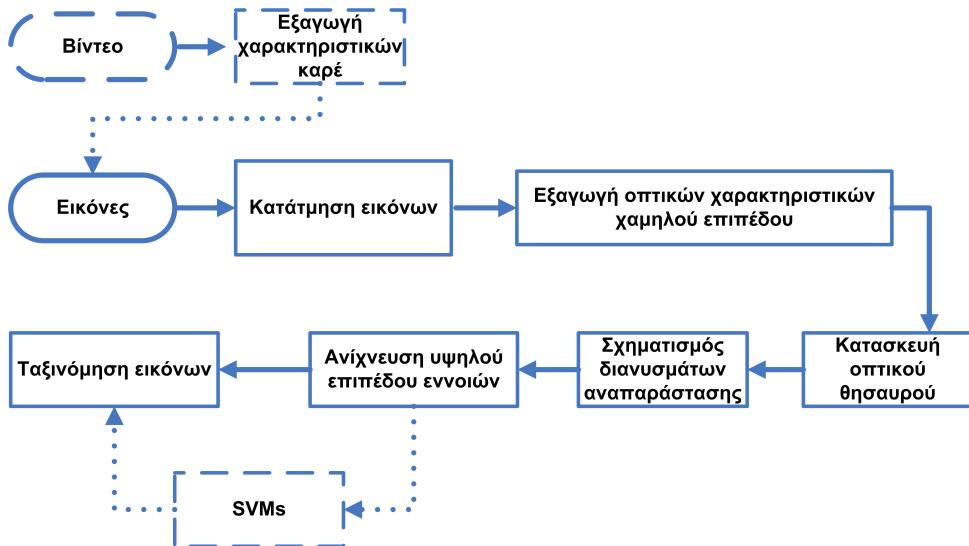
$$D_m = [\min\{d_i^1\}, \min\{d_i^2\}, \dots, \min\{d_i^{Nc}\}], i = 1, 2, \dots, NR \quad (7.1)$$

όπου  $d_i^j$  είναι ο βαθμός εμπιστοσύνης ότι η  $i$ -οστή περιοχή της εικόνας αντιστοιχεί στον  $j$ -οστό τύπο περιοχών,  $NR$  είναι ο αριθμός των κατατμημένων περιοχών της εικόνας και  $N_C$  το μέγεθος του θησαυρού περιοχών.

Για την εξαγωγή των χαμηλού επιπέδου χαρακτηριστικών γνωρισμάτων μιας εικόνας υπάρχουν γενικά δύο κατηγορίες προσεγγίσεων: η εξαγωγή των επιθυμητών Περιγραφέων συνολικά (δηλ. από ολόκληρη την εικόνα) ή τοπικά (δηλ. από τις περιοχές ενδιαφέροντος). Όπως προείπαμε, στην προσέγγισή μας εφαρμόζεται αρχικά ένας αλγόριθμος κατάτμησης χρώματος σε μια δεδομένη εικόνα, ο οποίος αποτελεί μια ειδική εφαρμογή της γνωστής μεθόδου RSST [11] κατάλληλα επεξεργασμένος, για να παράγει μια πρώτη χονδροειδή κατάτμηση. Με τον τρόπο αυτό η παραχθείσα κατάτμηση μπορεί διαισθητικά να παρέχει μια ποιοτική περιγραφή της εικόνας. Για να συλλάβουμε τα οπικά χαρακτηριστικά γνωρίσματα κάθε περιοχής της κατατμημένης εικόνας με έναν τυποποιημένο τρόπο, επιλέγουμε να εξάγουμε τα χαρακτηριστικά γνωρίσματα χρώματος και υφής. Δεδομένου ότι ένα σύνολο κυρίαρχων χρωμάτων σε μια εικόνα ή μια περιοχή ενδιαφέροντος έχει τη δυνατότητα να συλλάβει αποτελεσματικά τις ιδιότητες του χρώματός της, επιλέγουμε να εξαγάγουμε τον Περιγραφέα Κύριων Χρωμάτων MPEG-7 [164] από κάθε περιοχή της κατατμημένης εικόνας. Ακολουθώντας μια παρόμοια λογική, χρησιμοποιήθηκε ο Περιγραφέας Ομοιογενούς Υφής MPEG-7, για να συλλάβει τις ιδιότητες της υφής κάθε περιοχής. Οι ενεργειακές αποκλίσεις των Περιγραφέων απορρίφθηκαν, προκειμένου να απλοποιηθεί η συνολική περιγραφή. Όλες οι χαμηλού επιπέδου οπικές περιγραφές μιας εικόνας είναι ομαλοποιημένες και συγχωνευμένες σε ένα μοναδικό διάνυσμα. Αυτό το διάνυσμα θα αναφέρεται στη συνέχεια ως διάνυσμα χαρακτηριστικών. Για λόγους πληρότητας πρέπει να σημειώσουμε εδώ ότι για τις έννοιες που στοχεύουμε να ανιχνεύσουμε μόνο οι Περιγραφές χρώματος και υφής έχουν νόημα ανάμεσα στους Περιγραφές MPEG-7, οπότε οι σχετικοί Περιγραφές Σχήματος MPEG-7 έχουν παραλειφθεί.

Εν γένει, τα οπικά χαρακτηριστικά χαμηλού επιπέδου εξάγονται από όλες τις περιοχές κάθε εικόνας και οι οποίες έχουν προκύψει με κατάτμηση με βάση το χρώμα. Τα χαρακτηριστικά αυτά είναι στην ουσία οι MPEG-7 περιγραφές χρώματος και υφής. Συνδυάζοντας τους περιγραφές αυτούς, σχηματίζεται ένα διάνυσμα χαρακτηριστικών (feature vector) για κάθε περιοχή. Χρησιμοποιώντας έπειτα ένα μεγάλο αριθμό από εικόνες ή χαρακτηριστικά καρέ και εφαρμόζοντας τη μέθοδο ιεραρχικής συσταδοποίησης (βλ. ενότητα 4.2.1), κατασκευάζεται ένας θησαυρός περιοχών (region thesaurus) ή αλλιώς οπικός θησαυρός (visual thesaurus), ο οποίος περιέχει όλους τους τύπους περιοχής. Ο θησαυρός αυτός λειτουργεί σαν μια “βάση γνώσης” (KB) και διευκολύνει τη συσχέτιση των χαρακτηριστικών του χαμηλού με αυτά του υψηλού επιπέδου.

Σε κάθε περίπτωση επιλέγονται για τους τύπους περιοχής εκείνες οι περιοχές οι οποίες βρίσκονται πιο κοντά στα κεντρικά στοιχεία (centroids) της κάθε συστάδας και κάθε τύπος αναπαρίσταται από ένα συγχωνευμένο διάνυσμα (διάνυσμα χαρακτηριστικών). Μετρώντας τις αποστάσεις των περιοχών της εικόνας από τους τύπους περιοχής του θησαυρού, σχηματίζεται ένα διάνυσμα αναπαράστασης, το οποίο αποτυπώνει την σημασιολογία της εικόνας. Το διάνυσμα αυτό μπορεί να ληφθεί στη συνέχεια από έναν κατάλληλα εκπαιδευμένο ταξινομητή (π.χ. ένα Support Vector Machine - SVM) και να ταξινομήσει κατάλληλα την εικόνα με βάση το καθολικό της περιεχόμενο. Στις υπόλοιπες υποενότητες περιγράφονται συνοπτικά τα βήματα της κατασκευής οπικού θησαυρού και της δημιουργίας των διανυσμάτων αναπαράστασης, ενώ στο Σχήμα 7.5 παρουσιάζεται συνοπτικά η γενική μορφή της σχετικής μεθοδολογίας.



### 7.3.2 Κατασκευή οπτικού θησαυρού

Χρησιμοποιώντας τον αλγόριθμο κατάτμησης χρώματος, που αποτελεί μια εφαρμογή multiresolution της γνωστής RSST μεθόδου [11], λαμβάνουμε αρχικά μια χονδροειδή κατάτμηση των εικόνων, η οποία μας παρέχει διαισθητικά μια πρώτη ποιοτική περιγραφή τους. Στο Σχήμα 7.6 απεικονίζουμε μια τέτοια αρχική εικόνα μαζί με τη χονδροειδή κατάτμησή της. Η εικόνα αυτή θα μπορούσε π.χ. να περιγραφεί εύκολα από ένα χρήστη ως ένα σύνολο περιοχών: “μια γαλάζια περιοχή” (ουρανός), “δύο πράσινες περιοχές” (βλάστηση), “μια πορτοκαλί περιοχή” (φωτιά) κ.λ.π.. Κάθε περιοχή χαρακτηρίζεται από το διάνυσμα χαρακτηριστικών το οποίο εξάγεται με τη βοήθεια των χαμηλού επιπέδου οπτικών Περιγραφέων.



Σχήμα 7.6: Μια αρχική εικόνα εισαγωγής και η χονδροειδής κατάτμησή της.

Λαμβάνοντας, στη συνέχεια, υπόψη ολόκληρο το σύνολο εκπαίδευσης εικόνων και τα σχετικά χαμηλού επιπέδου χαρακτηριστικά γνωρίσματά τους, όπως αυτά εξάγονται χρησιμοποιώντας τους Περιγραφές που ορίζει το πρότυπο MPEG-7 [62]<sup>2</sup>, μπορούμε να διακρίνουμε ότι

1. εκείνες οι περιοχές, που ανήκουν σε παρόμοιες σημασιολογικές έννοιες, έχουν επίσης και παρόμοιες χαμηλού επιπέδου περιγραφές και επιπλέον,

<sup>2</sup>Για περισσότερες πληροφορίες σχετικά με την εξαγωγή των χαρακτηριστικών χαμηλού επιπέδου ο αναγνώστης καλείται να ανατρέξει στη δημοσίευση [188]

2. εκείνες οι εικόνες, που περιέχουν ίδιες υψηλού επιπέδου έννοιες, αποτελούνται από παρόμιοις περιοχές.

Αυτή η παρατήρηση για την ομοιότητα των περιοχών μπορεί και πρέπει να αξιοποιηθεί περαιτέρω, δεδομένου ότι η συνύπαρξη περιοχών χαρακτηρίζει συχνά τις έννοιες που υπάρχουν μέσα σε μια εικόνα.

Εκτελούμε, λοιπόν, μια συσταδοποίηση σε όλες τις διαφορετικές περιοχές που απαντώνται μέσα στις εικόνες του δεδομένου συνόλου εκπαίδευσης. Ο στόχος της συσταδοποίησης είναι να προσδιοριστούν οι πιο συχνές περιοχές, που απαντώνται μέσα στο σύνολο εκπαίδευσης, και να χρησιμοποιηθούν, για να κατασκευάσουν μια σχετική βάση γνώσης. Αυτές οι περιοχές θα αναφέρονται πλέον ως “τύποι περιοχής”. Όπως είναι μάλλον προφανές, δεν μπορούμε να γνωρίζουμε εκ των προτέρων τον ακριβή αριθμό των απαραίτητων κατηγοριών, προκειμένου να συσταδοποιηθούν αποτελεσματικά όλοι οι ετερογενείς τύποι περιοχής των εικόνων. Σε μία προσπάθεια να υπερνικήσουμε το πρόβλημα αυτό εφαρμόζουμε στο σημείο αυτό μια απλοποιημένη μορφή του αλγορίθμου ιεραρχικής συσταδοποίησης, που προτείνεται στο πλαίσιο της παρούσας διατριβής. Η μέθοδος αυτή επιλέγεται, για να εφαρμοστεί στο χαμηλού επιπέδου σύνολο περιγραφής, δεδομένου ότι μετά από τη συσταδοποίηση μπορούμε εύκολα να επιλέξουμε τον αριθμό των συστάδων που θα κρατήσουμε, χωρίς να χρειαστεί να επαναπροσαρμόσουμε τον ίδιο τον αλγόριθμο. Το τελευταίο επιτυγχάνεται εξαιτίας της ιεραρχικής φύσης του προτεινόμενου αλγορίθμου, επιλέγοντας τις τελικές συστάδες από διαφορετικά βηματικά επίπεδα κάθε φορά.

Μετά από το χωρισμό των περιοχών του συνόλου εκπαίδευσης στις κατάλληλες συστάδες μια αρχική παρατήρηση είναι ότι κάθε συστάδα μπορεί ή δεν μπορεί να περιέχει περιοχές από το ίδιο υψηλού επιπέδου χαρακτηριστικό γνώρισμα και ότι οι περιοχές από το ίδιο υψηλού επιπέδου χαρακτηριστικό γνώρισμα είναι δυνατόν να εμφανίζονται σε περισσότερες της μίας συστάδες. Για παράδειγμα, η υψηλού επιπέδου έννοια βλάστηση μπορεί να έχει περισσότερα από ένα στιγμιότυπα, τα οποία να διαφέρουν στο χρώμα των φύλλων των δέντρων. Κάθε στιγμιότυπο θα αναπαρασταθεί από το κεντρικό στοιχείο (centroid) μιας συστάδας και θα αναφερθεί ως τύπος περιοχής. Συγκεκριμένα θα αναπαρίσταται από το διάνυσμα χαρακτηριστικών της περιοχής αυτής, καθώς και ο αλγόριθμος συσταδοποίησης εφαρμόζεται πάνω στα διανύσματα χαρακτηριστικών του συνόλου των περιοχών. Επιπλέον, σε μία συστάδα που περιλαμβάνει στιγμιότυπα της έννοιας θάλασσα είναι δυνατόν αυτά τα στιγμιότυπα να έχουν αναμειχθεί με εκείνες της έννοιας ουρανός, εφόσον μία τυπική περιοχή ουρανού είναι σχεδόν πανομοιότυπη με συγκεκριμένες περιοχές θάλασσας. (βλ. Σχήμα 7.4).

Τελικά, ο θησαυρός  $T$  δίνεται από την εξίσωση (7.2).

$$T = \left\{ w_i, \quad i = 1 \dots N_T \right\}, \quad w_i \subset R \quad (7.2)$$

όπου  $\bigcup_i w = R$ ,  $i = 1 \dots N_T$  και  $\bigcap_{i,j} w = \emptyset$ ,  $i \neq j$ .  $N_T$  είναι ο αριθμός των τύπων περιοχής του θησαυρού (οπότε κατ' αναλογία και των συστάδων) και  $w_i$  είναι η  $i$ -οστή συστάδα, η οποία είναι ένα σύνολο από περιοχές εικόνων που ανήκουν στο  $R$ , όπως φαίνεται και στην εξίσωση (7.2). Ως τύπος περιοχής ή αλλιώς η “οπτική λέξη” (visual word) του οπτικού θησαυρού ονομάζεται το κεντρικό στοιχείο (centroid) της κάθε συστάδας  $w_i$ . Επίσης, από τα παραπάνω φαίνεται ότι η ένωση όλων των συστάδων δίνει όλο το σύνολο  $R$ , με την προϋπόθεση ότι όλες οι περιοχές έχουν χρησιμοποιηθεί για τη συσταδοποίηση και πως διαφορετικές συστάδες δεν έχουν κοινές περιοχές<sup>3</sup>.

<sup>3</sup>Προφανώς, σε μία μετέπειτα φάση βελτιστοποίησης της όλης διαδικασίας κάποιος θα μπορούσε

Γενικά, ένας θησαυρός συνδυάζει σε πρώτη φάση μία λίστα με τους όρους ενός δεδομένου πεδίου γνώσης κι ένα σύνολο από σχετικούς όρους για κάθε όρο στη λίστα, τα οποία είναι τα συνώνυμα του τρέχοντα όρου. Στην προσέγγιση μας ο κατασκευασμένος Θησαυρός Περιοχών περιέχει όλους τους τύπους περιοχής, οι οποίοι υπάρχουν μέσα στο σύνολο εκπαίδευσης. Επιλέγεται ο κάθε τύπος περιοχής να αναπαρίσταται από το διάνυσμα χαρακτηριστικών που είναι πιο κοντά στο διάνυσμα χαρακτηριστικών του κέντρου της κάθε συστάδας  $w_i$ . Ως κέντρο (ή “κεντρικό στοιχείο”) κάθε συστάδας  $w_i$  θεωρείται ο μέσος όρος όλων των σημείων της συστάδας. Με άλλα λόγια, στην περίπτωσή μας οι συντεταγμένες του κέντρου είναι οι αριθμητικοί μέσοι για κάθε διάσταση ξεχωριστά και για όλα τα σημεία της εκάστοτε συστάδας. Στην περίπτωση ενός  $N$ -διάστατου χώρου με  $M$  σημεία  $x_i = (x_{i1}, \dots, x_{iN})$ ,  $i = 1, \dots, M$ , το κεντρικό στοιχείο  $z$  δίνεται από τη σχέση:

$$z = \frac{1}{N} \left( \sum_{i=1, \dots, M} x_{i1}, \dots, \sum_{i=1, \dots, M} x_{iN} \right), \quad i = 1, \dots, N \quad (7.3)$$

Για παράδειγμα, αν υποθέσουμε ότι το σύνολο δεδομένων μας έχει τρεις διαστάσεις και η συστάδα δύο σημεία  $x = (x_1, x_2, x_3)$  και  $y = (y_1, y_2, y_3)$ , τότε το κεντρικό στοιχείο  $z$  θα είναι το  $z = (z_1, z_2, z_3)$ , όπου  $z_1 = \frac{(x_1+y_1)}{2}$ ,  $z_2 = \frac{(x_2+y_2)}{2}$  και  $z_3 = \frac{(x_3+y_3)}{2}$ . Υπολογίζοντας με τον τρόπο αυτό το κεντρικό στοιχείο  $z(w_i)$  όλων των διανυσμάτων χαρακτηριστικών που ανήκουν στην ίδια συστάδα, το διάνυσμα χαρακτηριστικών που επιλέγεται να αντιπροσωπεύει τον κάθε τύπο περιοχής δίνεται τελικά από την εξίσωση (7.4). Η απόσταση  $d$  ανάμεσα σε δύο διανύσματα χαρακτηριστικών ορίζεται στην επόμενη υποενότητα (εξίσωση (7.5)). Αυτοί οι τύποι περιοχής είναι οι κοντινότερες περιοχές στα κεντρικά στοιχεία των συστάδων και όλα τα υπόλοιπα διανύσματα χαρακτηριστικών της συστάδας είναι τα “συνώνυμά” τους. Σκοπός του θησαυρού είναι να αποτελέσει ένα ενδιάμεσο βήμα στην αντιστοιχία χαρακτηριστικών χαμηλού και υψηλού επιπέδου (έννοιες) και να διευκολύνει τη συσχέτισή τους.

$$f(w_i) = f\left(\arg \min_{r \in w_i} \left\{ d(f(r), z(w_i)) \right\}\right) \quad (7.4)$$

Ο κάθε τύπος περιοχής, λοιπόν, αναπαρίσταται από ένα διάνυσμα χαρακτηριστικών, το οποίο περιέχει όλη την εξαχθείσα πληροφορία χαμηλού επιπέδου για την περιοχή. Όπως είναι προφανές, μία περιγραφή χαμηλού επιπέδου δεν περιέχει καμία σημασιολογική πληροφορία, αποτελεί απλώς μία τυποποιημένη αναπαράσταση των οπτικών χαρακτηριστικών. Σημασιολογική πληροφορία και μόνο περιέχει μία έννοια υψηλού επιπέδου. Ένας τύπος περιοχής, λοιπόν, περιγράφεται ως κάτι ανάμεσα σε χαρακτηριστικά υψηλού και χαμηλού επιπέδου. Περιέχει τις απαραίτητες πληροφορίες, για να περιγράψει τυποποιημένα τα χαρακτηριστικά υφής και χρώματος, αλλά μπορεί επίσης να περιγράψει με μία περιγραφή χαμηλότερου επιπέδου από τις σημασιολογικές έννοιες υψηλού επιπέδου. Όπως θα δούμε παρακάτω, η επίδραση του εννοιολογικού πλαισίου πάνω σε αυτά τα ενδιάμεσου επιπέδου χαρακτηριστικά προσδίδει επιπρόσθετη σημασιολογική γνώση και βελτιώνει αισθητά τα αποτελέσματα της ανίχνευσης.

Για τον σχηματισμό του θησαυρού είναι απαραίτητος ένας αριθμός από διανύσματα χαρακτηριστικών από περιοχές εικόνων. Αν το σύνολο όλων των περιοχών δεν είναι πολύ μεγάλο, τότε όλες οι περιοχές θα χρησιμοποιηθούν για τον σχηματισμό του

να εφαρμόσει την πλήρη έκδοση της “ασαφοποιημένης” διαδικασίας συσταδοποίησης (βλ. κεφάλαιο 4), γεγονός, όμως, που ξεφεύγει από το πλαίσιο της παρούσης διατριβής

θησαυρού, ενώ στις ακολουθίες βίντεο ο αλγόριθμος συσταδοποίησης θα εφαρμοστεί στα διανύσματα χαρακτηριστικών όλων των χαρακτηριστικών καρέ. 'Όταν το σύνολο των εικόνων είναι πολύ μεγάλο και δεν υπάρχει δυνατότητα να χρησιμοποιηθεί ολόκληρο (για λόγους χρονοβόρων υπολογισμών αλλά και υπολογιστικής μνήμης), επιλέγεται ένα υποσύνολο αυτών. 'Ενας τρόπος με τον οποίο μπορεί να γίνει αυτό είναι να συλλέγονται περιοχές από εικόνες, οι οποίες απεικονίζουν τις επιλεγμένες προς εντοπισμό έννοιες και στη συνέχεια με τυχαίο τρόπο κάποιες ακόμα περιοχές. Αυτό, λοιπόν, το υποσύνολο θα χρησιμοποιηθεί, για να κατασκευασθεί ο θησαυρός στην περίπτωση ενός μεγάλου αριθμού συνολικών περιοχών. Στην περίπτωση αυτή δεν ισχύει  $\bigcup_i w = R$ ,  $i = 1 \dots N_T$ , αλλά ισχύει ότι  $\bigcup_i w \subset R$ ,  $i = 1 \dots N_T$ .

### 7.3.3 Διανύσματα αναπαράστασης

Σε αυτή την υποενότητα παρουσιάζεται ο αλγόριθμος που χρησιμοποιήθηκε, για να περιγραφεί η κάθε εικόνα, που είναι κατατμημένη σε περιοχές, με τη βοήθεια του κατασκευασμένου οπτικού θησαυρού. Ειδικότερα, εφαρμόζουμε την Ευκλείδια απόσταση, για να μετρήσουμε την απόσταση μεταξύ δύο διαφορετικών περιοχών, όσον αφορά τα διανύσματα χαρακτηριστικών τους. 'Εστω ότι θεωρούμε δύο οποιεσδήποτε περιοχές  $r_1, r_2 \in \mathcal{R}$ ,  $R \subset \mathcal{R}$ , οι οποίες έχουν διανύσματα χαρακτηριστικών  $f_1, f_2 \in \mathcal{F}$ ,  $F \subset \mathcal{F}$  και Ευκλείδια απόσταση  $d(f_1, f_2)$ , όπως αυτή δίνεται από την ακόλουθη εξίσωση (7.5).  $F$  είναι το σύνολο των διανυσμάτων χαρακτηριστικών για το συγκεκριμένο σύνολο περιοχών, ενώ  $\mathcal{F}$  είναι όλος ο χώρος των διανυσμάτων χαρακτηριστικών. Το ίδιο ισχύει και για τα  $R$  και  $\mathcal{R}$ . Πρέπει να σημειωθεί ότι το MPEG-7 δεν ορίζει αυστηρά μέτρα απόστασης για τους περιγραφές, παρά μόνο προτείνει κάποια ενδεικτικά, έτσι ώστε να αφήνει την ευελιξία να χρησιμοποιούνται και άλλοι τύποι αποστάσεων (βλ. σελίδα 66) και να ελέγχεται η αποτελεσματικότητα τους.

$$d(f_1, f_2) = \sqrt{\sum_{i=1}^n (f_1^i - f_2^i)^2} \quad (7.5)$$

'Έχοντας υπολογίσει την απόσταση της κάθε περιοχής της εικόνας από όλους τους τύπους περιοχής του κατασκευασμένου θησαυρού, το διάνυσμα αναπαράστασης που περιγράφει σημασιολογικά το οπτικό περιεχόμενο της εικόνας σχηματίζεται κρατώντας τη μικρότερη απόσταση από κάθε τύπο περιοχής του θησαυρού. Πιο συγκεκριμένα, το διάνυσμα αναπαράστασης που περιγράφει την εικόνα  $k_i$ , θα είναι:

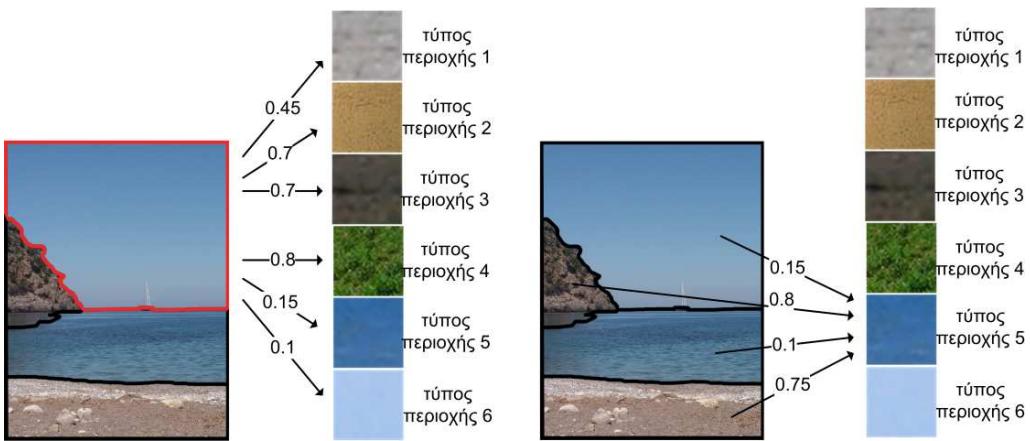
$$m_i = [m_i(1), m_i(2), \dots, m_i(j), \dots, m_i(N_T)], \quad i = 1 \dots N_K \quad (7.6)$$

όπου:

$$m_i(j) = \min_{r \in R(k_i)} \left\{ d(f(w_j), f(r)) \right\}, \quad i = 1 \dots N_K, \quad j = 1 \dots N_T \quad (7.7)$$

Κάθε διάνυσμα αναπαράστασης συμβολίζεται με  $m_i \in M$ ,  $i = 1 \dots N_K$ , όπου  $M$  το σύνολο όλων των διανυσμάτων αναπαράστασης και  $m_i$  το διάνυσμα αναπαράστασης της εικόνας  $k_i$ .

Για την καλύτερη κατανόηση των παραπάνω παρουσιάζουμε στο Σχήμα 7.7 ένα παράδειγμα, όπου έχουμε μία εικόνα χωρισμένη σε 4 περιοχές και 6 τύπους περιοχής ενός οπτικού θησαυρού. Στο αριστερό τμήμα του Σχήματος παρουσιάζονται οι αποστάσεις για την περιοχή του ουρανού από κάθε τύπο περιοχής και στο δεξιό τμήμα οι



**Σχήμα 7.7:** Αποστάσεις ανάμεσα σε περιοχές εικόνας και τύπους περιοχής: αριστερά παρουσιάζονται αποστάσεις ανάμεσα σε μία περιοχή της εικόνας και σε όλους τους τύπους περιοχής, ενώ δεξιά αποστάσεις ανάμεσα σε όλες τις περιοχές μίας εικόνας και έναν τύπο περιοχής

αποστάσεις κάθε περιοχής της εικόνας από τον τύπο περιοχής 5. Το διάνυσμα αναπαράστασης σχηματίζεται από τις μικρότερες αποστάσεις για κάθε τύπο περιοχής. Στην προκειμένη περίπτωση για τον τύπο περιοχής #5 η ελάχιστη απόσταση είναι 0.1. Το διάνυσμα αναπαράστασης για αυτή την εικόνα και για τον συγκεκριμένο θησαυρό θα είναι, όπως φαίνεται στην εξήσωση (7.8).

$$m = [m(1), m(2), m(3), m(4), m(5), m(6)] \quad (7.8)$$

όπου το  $m(5)$  θα είναι 0.1. Υπολογίζοντας όλες τις αποστάσεις ανάμεσα σε όλες τις περιοχές της εικόνας (4) και σε όλους τους τύπους περιοχής (6), συνολικά δηλαδή  $4 \times 6 = 24$  αποστάσεις, σχηματίζεται και το διάνυσμα αναπαράστασης. Τέλος, στο Σχήμα 7.8 φαίνονται τα διανύσματα αναπαράστασης για 5 εικόνες και για τον οπτικό θησαυρό 6 περιοχών, του οποίου οι περιοχές αναπαριστώνται στο Σχήμα.

### 7.3.4 Ταξινόμηση σκηνών

Το επόμενο βήμα της διαδικασίας που περιγράφηκε έως τώρα είναι μάλλον προφανές για τον αναγνώστη, ήτοι η εκπαίδευση κατάλληλων επιβλεπόμενων ταξινομητών SVM, οι οποίοι θα χρησιμοποιούν τα προαναφερθέντα διανύσματα αναπαράστασης ως το μέσο απεικόνισης για τα εξαχθέντα χαρακτηριστικά γνωρίσματα της δεδομένης εικόνας και οι οποίοι θα ταξινομούν τις εικόνες σε κατάλληλες κατηγορίες. Το εννοιολογικό πλαίσιο έχει πεδίο εφαρμογής τόσο κατά το μετασχηματισμό των διανυσμάτων αναπαράστασης, που χρησιμοποιούνται για την εκπαίδευση και δοκιμή των ανιχνευτών υψηλού επιπέδου εννοιών, όσο και μετά την εφαρμογή των παραπάνω ταξινομητών για την βελτίωση των αποτελεσμάτων της ανιχνευσης. Στην υποενότητα, λοιπόν, αυτή παρουσιάζεται συνοπτικά το υπο-πρόβλημα της αυτόματης ταξινόμησης εικόνων σε κατηγορίες με βάση τα χαρακτηριστικά ενδιάμεσου επιπέδου και πιο συγκεκριμένα της ταξινόμησης ανάμεσα σε δύο κατηγορίες, οπότε ανάγεται σε μία δυαδική απόφαση για την μία ή την άλλη κατηγορία. Το πρόβλημα αυτό στη συγκεκριμένη προσέγγιση αντιμετωπίζεται με ολικές περιγραφές της εικόνας και σε αυτό βασίζεται η επίδραση του εννοιολογικού πλαισίου ταξινόμησης.

Έστω ότι έχουμε  $N$  εικόνες οι οποίες αντιστοιχούν όλες σε μία από δύο κατηγορίες και κάθε εικόνα συμβολίζεται με  $k_i$ , ενώ το σύνολο όλων των εικόνων με



<i>Tύπος Περιοχής 1</i>	<i>Tύπος Περιοχής 2</i>	<i>Tύπος Περιοχής 3</i>	<i>Tύπος Περιοχής 4</i>	<i>Tύπος Περιοχής 5</i>	<i>Tύπος Περιοχής 6</i>
-------------------------	-------------------------	-------------------------	-------------------------	-------------------------	-------------------------

<i>Eικόνα 1</i>	[ 0.4 ]	[ 0.3 ]	[ 0.6 ]	[ 0.7 ]	[ 0.2 ]	[ 0.1 ]
<i>Eικόνα 2</i>	[ 0.5 ]	[ 0.7 ]	[ 0.1 ]	[ 0.8 ]	[ 0.3 ]	[ 0.4 ]
<i>Eικόνα 3</i>	[ 0.8 ]	[ 0.1 ]	[ 0.9 ]	[ 0.8 ]	[ 0.2 ]	[ 0.5 ]
<i>Eικόνα 4</i>	[ 0.1 ]	[ 0.7 ]	[ 0.3 ]	[ 0.6 ]	[ 0.2 ]	[ 0.4 ]
<i>Eικόνα 5</i>	[ 0.3 ]	[ 0.7 ]	[ 0.4 ]	[ 0.1 ]	[ 0.9 ]	[ 0.7 ]



Eικόνα 1



Eικόνα 2



Eικόνα 3



Eικόνα 4



Eικόνα 5

**Σχήμα 7.8:** Παραδείγματα διανυσμάτων αναπαράστασης

*K.*

$$k_i \in K, \quad i = 1 \dots N \quad (7.9)$$

Δημιουργούμε τα σχετικά διανύσματα χαρακτηριστικών, εξάγοντας Περιγραφείς MPEG-7 χρώματος και υψής συνολικά από την κάθε εικόνα. Ειδικότερα, οι Περιγραφείς που εξάγονται είναι: Περιγραφέας Κύριων Χρωμάτων -  $DCD$ , Κλιμακωτός Περιγραφέας Χρώματος -  $CLD$ , Περιγραφέας Διάταξης Χρώματος -  $SCD$ , Περιγραφέας Δομής Χρώματος -  $CSTD$ , Περιγραφέας Ομοιογενούς Υφής -  $EHD$  και Περιγραφέας Ιστογράμματος Ακμών -  $HT$ . Έτσι, σε κάθε εικόνα αντιστοιχεί και ένα διάνυσμα για κάθε Περιγραφέα. Συνολικά, έχουμε 6 διανύσματα για κάθε εικόνα  $k_i$ , τα:  $DCD_i$ ,  $CLD_i$ ,  $SCD_i$ ,  $CSTD_i$ ,  $EHD_i$  και  $HT_i$ . Δημιουργούμε, επίσης, κι ένα συγχωνευμένο διάνυσμα το οποίο περιέχει όλους τους Περιγραφείς για κάθε εικόνα, όπως φαίνεται στην εξίσωση (7.10). Στη συνέχεια, εφαρμόζεται μία ανάλυση πρωτογενών συνιστωσών (principal component analysis) σε όλα τα διανύσματα ανά Περιγραφέα, προκειμένου να αποσυσχετισθούν, και με τον τρόπο αυτό προκύπτουν διανύσματα μικρότερης διάστασης.

$$f_i = [DCD_i, CSTD_i, CLD_i, SCD_i, EHD_i, HT_i], \quad i = 1 \dots N \quad (7.10)$$

Οι κατηγορίες που θα εξετάσουμε περαιτέρω στην υποενότητα 7.5.4 είναι δύο: η πρώτη κατηγορία αφορά εικόνες από τη θεματική περιοχή παραλία και συμβολίζεται με  $c_1$ , ενώ η δεύτερη κατηγορία εικόνες από τη θεματική περιοχή βουνό και συμβολίζεται με  $c_2$ . Συμβολίζουμε με  $b(k_i)$  την έξοδο του ταξινομητή για την εικόνα  $k_i$ , η οποία είναι 1 για την πρώτη κατηγορία και 0 για τη δεύτερη κατηγορία. Σχηματίζεται έτσι ένα σύνολο  $D_1$ , το οποίο περιέχει τις εικόνες οι οποίες ταξινομήθηκαν στην πρώτη κατηγορία, και ένα σύνολο  $D_2$ , που περιέχει τις εικόνες της δεύτερης κατηγορίας.

$$b(k) \in \{0, 1\}, \quad k \in K \quad (7.11)$$

$$D_1 = \{k \in K : b(k) = 1\} \quad (7.12)$$

$$D_2 = \{k \in K : b(k) = 0\} \quad (7.13)$$

Με τον τρόπο αυτό καταλήγουμε στην επιθυμητή ταξινόμηση των εικόνων με βάση τα χαρακτηριστικά ενδιάμεσου επιπέδου τους. Στη συνέχεια, θα εμβαθύνουμε στην επίδραση του εννοιολογικού πλαισίου στη διαδικασία τόσο σε επίπεδο εννοιών, όσο και τύπων περιοχής, ενώ συνολικά αποτελέσματα παρέχονται στην ενότητα 7.5 του κεφαλαίου.

## 7.4 Εννοιολογικό Πλαίσιο Ταξινόμησης

Προκειμένου να αξιοποιήσουμε πλήρως την έννοια του εννοιολογικού πλαισίου και να τη συνδυάσουμε με την παραπάνω θεώρηση της ανάλυσης ενδιάμεσου επιπέδου, προτείνουμε στη συνέχεια μια τριπλή προσέγγιση. Η μεθοδολογία αυτή μπορεί να διαιρεθεί στις ακόλουθες ενότητες, ανάλογα με την επίδραση που ασκεί το εννοιολογικό πλαίσιο στις έννοιες και στους τύπους περιοχής. Η ερευνητική προσπάθειά μας εστιάζει στην ολοκληρωμένη προσέγγιση του θέματος και προσφέρει μια ενοποιημένη και ανεπίβλεπτη διαχείριση του πολυμεσικού περιεχομένου. Αποδεικνύεται ότι η χρήση μεσαίου επιπέδου πληροφοριών βελτιώνει τα αποτελέσματα της παραδοσιακής - βιοηθούμενης από τη γνώση - ανάλυσης της εικόνας με βάση τόσο οπτική πληροφορία, όσο και πληροφορία εννοιολογικού πλαισίου.

### 7.4.1 Εννοιολογικό πλαίσιο ταξινόμησης εννοιών

Στη συνέχεια, προτείνουμε μια παρεμφερή ως προς το κεφάλαιο 6 μέθοδο βελτιστοποίησης της ανίχνευσης εννοιών υψηλού επιπέδου με τη βασική διαφορά, όμως, ότι η εν λόγω προσέγγιση αφορά τώρα σε μία καθολική θεώρηση της εικόνας. Με άλλα λόγια, οι σχετικές υψηλού επιπέδου έννοιες είτε υπάρχουν, είτε δεν υπάρχουν στο σύνολο της εικόνας και όχι σε κάποια επιμέρους περιοχή της: η υπό εξέταση εικόνα περιέχει π.χ. την έννοια νερό, αλλά δεν παρέχεται κάποια πληροφορία για τον χωρικό προσδιορισμό της. Αν και εκ πρώτης όψεως μια τέτοια προσέγγιση δε φαίνεται εποικοδομητική, πρόκειται για μια ιδιαίτερα χρήσιμη και σημαντική μεθοδολογία, η οποία μας παρέχει άμεσα, γρήγορα και αυτόματα αρκετές πληροφορίες για το περιεχόμενο της εικόνας και κατ' επέκταση για την κατηγορία στην οποία αυτή ανήκει. Πρόκειται, άλλωστε, για την εγνωσμένης αξίας και ευρέως διαδεδομένη θεώρηση καθολικής κατηγοριοποίησης (global classification), που ακολουθείται και στο πλαίσιο του TRECVID [231].

Το εννοιολογικό πλαίσιο ταξινόμησης έρχεται εδώ να βελτιώσει τα αποτελέσματα της ανίχνευσης εννοιών (οι οποίες έννοιες προκύπτουν από την έξοδο κατάλληλα εκπαιδευμένων ανιχνευτών SVM), τα οποία προέρχονται από τη διαδικασία εξαγωγής και αξιοποίησης των μεσαίου επιπέδου χαρακτηριστικών που περιγράφηκε στις προηγούμενες υποενότητες. Προτείνουμε, λοιπόν, την περαιτέρω προσαρμογή των αποτελεσμάτων της χαμηλού επιπέδου και βασισμένης στους Περιγραφείς, πολυμεσικής ανάλυσης, αξιοποιώντας την έννοια των ενδιάμεσου επιπέδου τύπων περιοχής, αποκλειστικά, όμως, και μόνο σύμφωνα με μια υψηλού επιπέδου οντολογία εννοιών. Η τελευταία αποτελείται από ένα σύνολο εννοιών και σημασιολογικών σχέσεων μεταξύ των εννοιών μέσα σε ένα δεδομένο περιβάλλον εννοιολογικού πλαισίου. Μπορούμε να αποδομήσουμε μια τέτοια οντολογία  $\mathcal{O}_C$  σε δύο μέρη, δηλ.:

1. το σύνολο  $C$  όλων των σημασιολογικών εννοιών  $c_i \in C, i = 1 \dots n$  και
2. το σύνολο  $R_{c_i, c_j}$  όλων των σημασιολογικών σχέσεων μεταξύ δύο οποιονδήποτε εννοιών  $c_i, c_j, j = 1 \dots n$

Ακολουθώντας μια πιο επίσημη διατύπωση:

$$\mathcal{O}_C = \{C, R_{c_i, c_j}\}, \quad R_{c_i, c_j} : C \times C \rightarrow \{0, 1\} \quad (7.14)$$

**Πίνακας 7.1: Ασαφείς σημασιολογικές σχέσεις μεταξύ εννοιών.**

Όνομα	Αντίστροφο	Σύμβολο Ερμηνεία	Παραδειγματα	
			a	b
Ειδίκευση	Γενίκευση	$Sp(a, b)$ το $b$ είναι μια γενίκευση της ερμηνείας του $a$	ζώο	αγριογούρουνο
Μέρος	Μέρος Του	$P(a, b)$ το $b$ είναι μέρος του $a$	Αυστραλία	Σύδνεϊ
Παράδειγμα	Παράδειγμα Του	$Ex(a, b)$ το $b$ είναι ένα παράδειγμα του $a$	ηγέτης	Βεν
'Οργανο	'Οργανο Του	$Ins(a, b)$ το $b$ είναι ένα όργανο του ή εμπλέκεται από το $a$	τομή	μαχαίρι
Θέση	Θέση Του	$Loc(a, b)$ το $b$ είναι η θέση του $a$	σκηνή	παραλία
Ασθενής	Ασθενής Του	$Pat(a, b)$ το $b$ επηρεάζεται από ή υπομένει τη δράση του $a$	δίνω	φυσικοβούτυρο
Ιδιότητα	Ιδιότητα Του	$Pr(a, b)$ το $b$ είναι μια ιδιότητα του $a$	μπανάνα	άγουρη

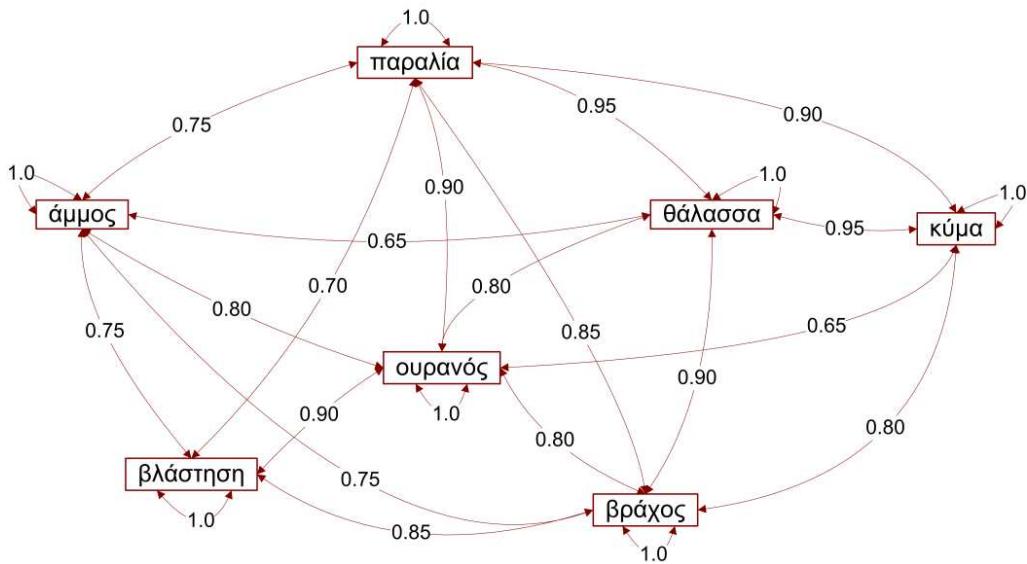
Όπως έχει αναφερθεί επανειλημμένα στο πλαίσιο της παρούσης, αλλά και στη βιβλιογραφία [185], με μια οντολογία αυτού του είδους είμαστε σε θέση να αναπαραστήσουμε οποιοδήποτε είδος σημασιολογικής σχέσης μπορεί να προκύψει. Εντούτοις, σε αυτή τη φάση περιοριζόμαστε σε μία “ασαφοποιημένη” εκδοχή μιας οντολογίας εννοιολογικού πλαισίου, ήτοι μια οντολογία που περιέχει έναν αντιπροσωπευτικό αριθμό

από διαχριτές και συνάμα ανόμοιες σχέσεις εννοιών. Η τελευταία εισάγεται, προκειμένου να εκφραστούν με έναν βέλτιστο τρόπο οι πραγματικές σχέσεις που υπάρχουν μεταξύ των συμμετεχουσών εννοιών κάθε περιοχής. Για να είναι η οντολογία μας αξιόπιστη και περιγραφική, θα πρέπει να περιέχει έναν αντιπροσωπευτικό αριθμό διαχριτών και διαφορετικών μεταξύ τους σχέσεων, έτσι ώστε να διασκορπίζονται οι πληροφορίες μεταξύ τους και να περιγράφεται έτσι το εννοιολογικό πλαίσιο τους με έναν ικανοποιητικό τρόπο. Επιπλέον, οι σχέσεις αυτές θα πρέπει να συνδυαστούν κατάλληλα, έτσι ώστε να προκύψει μια άποψη της γνώσης, που να αρκεί για τον καθορισμό και την εκτίμηση του εννοιολογικού πλαισίου. Κατά την ενσωμάτωση, δε, της αβεβαιότητας της πραγματικής ζωής στη μοντελοποίησή τους χρησιμοποιούμε ένα κατάλληλο υποσύνολο από ορισμένες στο πρότυπο MPEG-7 σημασιολογικές σχέσεις (βλ. Πίνακα 7.1) [23], οι οποίες είναι κατάλληλες για την ανάλυση πολυμέσων και καθορίζονται εκ των προτέρων από έναν εμπειρογνώμονα. Θα πρέπει να τονίσουμε ότι οι αρχικές σχέσεις επαναπροσδιορίσθηκαν, έτσι ώστε να συμπεριλάβουν και την ασάφεια στον ορισμό τους, δηλ. καθεμία από αυτές συσχετίστηκε με έναν βαθμό εμπιστοσύνης. Η οντολογία αυτή μπορεί να περιγραφεί από την ασαφοποιημένη εκδοχή της οντολογίας εννοιών (εξ. 7.15), όπου το  $C$  αντιπροσωπεύει και πάλι το σύνολο όλων των πιθανών εννοιών, η  $r_{c_i, c_j} : C \times C \rightarrow [0, 1]$  δείχνει μια ασαφή οντολογική σχέση μεταξύ δύο εννοιών  $c_i, c_j$  και η  $R_{c_i, c_j}$  δείχνει την κλασική, μη-ασαφή σημασιολογική σχέση μεταξύ των δύο εννοιών. Ο τελικός συνδυασμός των σχέσεων διαμορφώνει και στην περίπτωση αυτή έναν RDF γράφο και αποτελεί το αφηρημένο μοντέλο γνώσης που χρησιμοποιείται στη συνέχεια (βλ. Σχήμα 7.9).

$$\mathcal{O}_C^f = \{C, r_{c_i, c_j}\}, \quad i, j = 1 \dots n \quad (7.15)$$

Ο γράφος του προτεινόμενου μοντέλου περιέχει κόμβους (δηλ. έννοιες από την εκάστοτε θεματική περιοχή) και ακμές (δηλ. ένα κατάλληλο συνδυασμό<sup>4</sup> ασαφών σχέσεων εννοιολογικού πλαισίου). Ο βαθμός εμπιστοσύνης κάθε ακμής αντιπροσωπεύει την ασάφεια στο μοντέλο. Οι ανύπαρκτες ακμές υπονοούν ανύπαρκτες σχέσεις (δηλ. με άλλα λόγια, οι σχέσεις με μηδενικές τιμές εμπιστοσύνης παραλείπονται). Μια υφιστάμενη άκρη μεταξύ ενός δεδομένου ζευγαριού εννοιών παράγεται με βάση το σύνολο των ασαφών σχέσεων εννοιολογικού πλαισίου που είναι σημαντικές για το συγκεκριμένο ζευγάρι. Παραδείγματος χάριν, η ακμή του γράφου μεταξύ των εννοιών βράχος και άμμος παράγεται από το συνδυασμό των σχέσεων Τοποθεσία και Ασθενής, ενώ η ακμή μεταξύ των κόμβων νερό και θάλασσα χρησιμοποιεί τις σχέσεις Ειδίκευση, Μέρος, Παράδειγμα, Όργανο, Τοποθεσία και Ασθενής, προκειμένου να κατασκευαστεί και να καταλήξουμε στο συγκεκριμένο βαθμό εμπιστοσύνης που τη συνοδεύει. Φυσικά, κάθε έννοια έχει μια διαφορετική πιθανότητα, για να εμφανιστεί στη σκηνή της εικόνας και κατά συνέπεια και εδώ ένα επίπεδο μοντέλο εννοιολογικού πλαισίου, το οποίο δε θα αξιοποιούσε την εννοιολογική γνώση, δε θα ήταν επαρκές. Αντίθετα, το γεγονός ότι οι έννοιες σχετίζονται η μια με την άλλη υπονοεί ότι οι σχέσεις του γράφου είναι στην πραγματικότητα μεταβατικές. Η περιγραφή του βαθμού εμπιστοσύνης ακολουθεί τα όσα περιγράφηκαν σε προγενέστερα κεφάλαια της παρούσης και υλοποιείται με την τεχνική του RDF reification [318], όπου εισάγεται εν γένει μια επιπλέον RDF δήλωση για κάθε τριπλέτα που περιγράφει μια σχέση εννοιών.

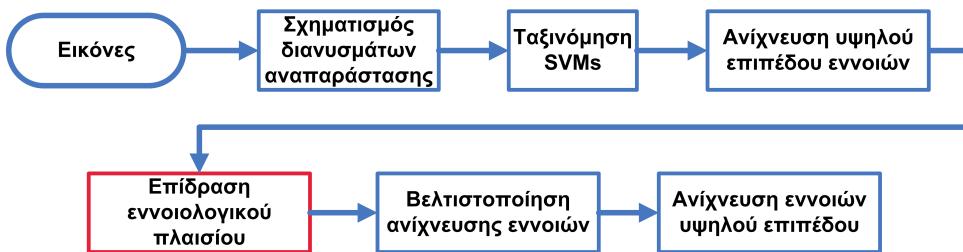
<sup>4</sup>Ο αποδοτικός συνδυασμός διαφορετικών ασαφών σχέσεων εννοιολογικού πλαισίου προς τη δημιουργία μιας στην πράξη εκμεταλλεύσιμης άποψης γνώσης εκτελείται με τη χρησιμοποίηση βασικών τελεστών της ασαφούς άλγεβρας και ειδικότερα της κλασικής  $t$ -νόρμας.



**Σχήμα 7.9:** Τμήμα της οντολογίας εννοιολογικού πλαισίου από τη θεματική περιοχή παραλία για την περίπτωση χρήσης υψηλού επιπέδου εννοιών με βάση ενδιάμεσου επιπέδου χαρακτηριστικά. Η έννοια παραλία είναι το στοιχείο ρίζα.

#### 7.4.1.1 Επίδραση βελτιστοποίησης εννοιολογικού πλαισίου

Μόλις οριστικοποιηθεί η παραπάνω δομή της γνώσης και υλοποιηθεί η αντίστοιχη αναπαράσταση, εφαρμόζεται μια παραλλαγή του βασισμένου στο εννοιολογικό πλαίσιο αλγορίθμου αναπροσαρμογής των βαθμών εμπιστοσύνης [183]. Η προτεινόμενη πρόσεγγιση επίδρασης του εννοιολογικού πλαισίου δημιουργεί ένα βήμα μετά-επεξεργασίας πάνω από το αρχικό σύνολο των ενδιάμεσου επιπέδου τύπων περιοχής, που εξάγονται κατά το βήμα της ανάλυσης (βλ. ενότητα 7.3). Επιδρά πάνω στα αποτελέσματα των ταξινομητών SVM και με τον τρόπο αυτό παρέχει μια βελτιωμένη επανεκτίμηση των βαθμών εμπιστοσύνης των αρχικών εννοιών που προήλθαν από την επεξεργασία των διανυσμάτων αναπαράστασης από αυτούς. Στη διαδικασία χρησιμοποιεί την υψηλού επιπέδου γνώση του εννοιολογικού πλαισίου από την κατασκευασμένη οντολογία εννοιολογικού πλαισίου. Η μεθοδολογία ενδιάμεσου επιπέδου λαμβάνει τη μορφή του Σχήματος 7.10. Με άλλα λόγια, ο παρακάτω αλγόριθμος ανάλυσης του εννοιολογικού πλαισίου ταξινόμησης αναπροσαρμόζει με έναν σημαντικό τρόπο τις αρχικές τιμές εμπιστοσύνης των καθολικών εννοιών που είναι παρούσες σε μία εικόνα και οι οποίες παράγονται από την ανάλυση των τύπων περιοχής.



**Σχήμα 7.10:** Μεθοδολογία επίδρασης εννοιολογικού πλαισίου ταξινόμησης εννοιών

Μια πρώτη εκτίμηση του βαθμού εμπιστοσύνης κάθε έννοιας προέρχεται από τις άμεσες και έμμεσες σχέσεις της έννοιας με άλλες έννοιες, χρησιμοποιώντας έναν κατάλληλο δείκτη συμβατότητας ή μια μετρική απόστασης. Κατά τα γνωστά, ανάλογα

με τη φύση των θεματικών περιοχών, που δημιούργησαν την εκάστοτε οντολογία θεματικών περιοχών, επιλέγομε και τον καλύτερο δείκτη συμβατότητας. Η γενική δομή του αλγορίθμου επανεκτίμησης των βαθμών ιδιότητας μέλους των καθολικών εννοιών είναι, στην περίπτωση αυτή, η ακόλουθη:

1. Η επιλεγμένη θεματική περιοχή επιβάλλει τη χρήση ενός μέτρου ομοιότητας (ή ανομοιότητας) μεταξύ των περιοχών:  $\mu \in [0, 1]$ .
2. Για κάθε τύπο περιοχής  $t$  εξετάστε ένα ασαφές σύνολο  $L_t$  με έναν βαθμό ιδιότητας μέλους  $\mu_t(c)$ , που περιέχει τους βαθμούς εμπιστοσύνης των πιθανών εννοιών.
3. Για κάθε έννοια  $c_i$  στο ασαφές σύνολο  $L_t$  με έναν βαθμό ιδιότητας μέλους  $\mu_t(c_i)$  λάβετε τις ιδιαίτερες βασισμένες στο εννοιολογικό πλαίσιο πληροφορίες με τη μορφή των σχέσεών του με το σύνολο οποιωνδήποτε άλλων εννοιών:  $\{r_{c_i, c_j} : c_i, c_j \in C, i \neq j\}$ .
4. Υπολογίστε το νέο βαθμό ιδιότητας μέλους  $\mu_t(c)$ , λαμβάνοντας υπόψη το μέτρο ομοιότητας κάθε θεματικής περιοχής. Στην περίπτωση σχέσεων πολλαπλών εννοιών στην οντολογία, κατά το συσχετισμό της έννοιας  $c$  με περισσότερες από την έννοια ρίζα έννοιες (Σχήμα 7.9), πρέπει να εφαρμοστεί ένα ενδιάμεσο βήμα συνάθροισης για την εκτίμηση του  $\mu_t(c)$  με τη θεώρηση της έννοιας  $c$  σχετικότητας  $cr_c$ :  $cr_c = \max \{r_{c, c_1}, \dots, r_{c, c_k}\}$ ,  $c_1 \dots c_k \in C$ . Εκφράζουμε τον υπολογισμό  $\mu_t(c)$  με τον αναδρομικό τύπο:

$$\mu_t^n(c) = \mu_t^{n-1}(c) - \mu(\mu_t^{n-1}(c) - cr_c) \quad (7.16)$$

όπου το  $n$  καταδεικνύει την εκάστοτε επανάληψη. Ισοδύναμα για μία αυθαίρετη επανάληψη  $n$  θα έχουμε:

$$\mu_t^n(c) = (1 - \mu)^n \cdot \mu_t^0(c) + (1 - (1 - \mu)^n) \cdot cr_c \quad (7.17)$$

όπου το  $\mu_t^0(c)$  αναπαριστά τον αρχικό βαθμό εμπιστοσύνης για την καθολική έννοια  $c$ .

#### 7.4.2 Εννοιολογικό πλαίσιο τύπων περιοχής

Στην ενότητα αυτή παρουσιάζεται μία προσέγγιση εννοιολογικού πλαισίου, η οποία στηρίζεται στις τοπολογικές σχέσεις μεταξύ των ενδιάμεσου επιπέδου τύπων περιοχής και είναι κατάλληλη για την αποδοτική ανίχνευση υψηλού επιπέδου εννοιών σε εικόνες. Οι κύριες διαφορές με τα όσα αναπτύχθηκαν έως τώρα έγκειται στο γεγονός ότι

- σε αυτή τη θεώρηση δε χρησιμοποιούνται σημασιολογικές σχέσεις για τη δημιουργία της γνώσης, παρά μόνο τοπολογικές,
- χρησιμοποιούνται ανιχνευτές υψηλού επιπέδου εννοιών για το σύνολο της εξεταζόμενης εικόνας και όχι για μία συγκεκριμένη περιοχή της και
- η εν λόγω διαδικασία αποτελεί μέρος της διαδικασίας της ανάλυσης και όχι εκ των υστέρων βήμα μετα-επεξεργασίας.

## Κεφάλαιο 7. Εννοιολογικό πλαίσιο ταξινόμησης

Με τον τρόπο αυτό ανιχνεύονται οι έννοιες οι οποίες είναι παρούσες στην υπό εξέταση εικόνα. Και εδώ ο χαρακτήρας της ανίχνευσης είναι καθολικός σε αντιδιαστολή με τον τοπικό χαρακτήρα των τεχνικών του προηγούμενου κεφαλαίου, οι οποίες έδιναν έμφαση στην εκάστοτε περιοχή μιας εικόνας. Στη συνέχεια, προτείνουμε μια σχετική αναπαράσταση γνώσης και αναλύουμε συνοπτικά τον προτεινόμενο αλγόριθμο αξιοποίησης αυτού του είδους εννοιολογικού πλαισίου κατά την ανίχνευση εννοιών στη διάρκεια της ανάλυσης πολυμεσικού περιεχομένου.

Τα αρχικά αποτελέσματα της ανάλυσης ενισχύονται μέσω της χρησιμοποίησης σημασιολογικής γνώσης με τη μορφή ανεξάρτητων από περιοχές τύπων περιοχής και των σχέσεων μεταξύ τους. Και εδώ, όπως και προηγουμένως, η θεώρηση είναι καθολική για το σύνολο της εικόνας, όμως η βοήθεια της σημασιολογικής γνώσης υπεισέρχεται κατά τη διάρκεια της ίδιας της ανάλυσης πάνω στον υπολογισμό των διανυσμάτων αναπαράστασης. Το κίνητρο για αυτού του είδους τη θεώρηση είναι το γεγονός ότι οι ενδιάμεσου επιπέδου πληροφορίες λαμβάνουν τη μορφή μιας ενδιάμεσης περιγραφής, η οποία μπορεί μεν να περιγραφεί σημασιολογικά, αλλά δεν εκφράζει τις υψηλού επιπέδου έννοιες. Έτσι, στη συνέχεια κατασκευάζεται μια σχετική οντολογία εννοιών ενδιάμεσου επιπέδου, η οποία περιέχει την παραπάνω πληροφορία.

Ειδικότερα, προτείνουμε μια βασισμένη στο εννοιολογικό πλαίσιο μέθοδο για τη βελτιστοποίηση των αποτελεσμάτων της χαμηλού επιπέδου πολυμεσικής ανάλυσης μέσω της χρησιμοποίησης της έννοιας των ενδιάμεσου επιπέδου τύπων περιοχής. Οι τελευταίοι κατασκευάζουν μια πρωτότυπη οντολογία εννοιολογικού πλαισίου ενδιάμεσου επιπέδου, η οποία περιγράφεται αποκλειστικά από το σύνολο των τύπων περιοχής και των χωρικών σχέσεων που τους διέπουν. Γενικά, μπορούμε να αποσυνθέσουμε μια τέτοια οντολογία  $\mathcal{O}_T$  σε δύο μέρη, στο σύνολο  $T$  όλων των τύπων περιοχής και στο σύνολο  $R_{t_i, t_j}$  όλων των σχέσεων μεταξύ δύο οποιονδήποτε τύπων περιοχής  $t_i, t_j$ . Τυπικότερα:

$$\mathcal{O}_T = \{T, R_{t_i, t_j}\}, \quad R_{t_i, t_j} : T \times T \rightarrow \{0, 1\}, \quad i, j = 1 \dots n \quad (7.18)$$

Όπως έχει γίνει πλέον σαφές σε αυτή τη διατριβή, είμαστε σε θέση να μοντελοποιήσουμε οποιαδήποτε είδος δυαδικής σχέσης (λ.χ. σημασιολογικές σχέσεις, τοπολογικές, χρονικές, χωροχρονικές, κ.ο.κ.) με μία οντολογία [243]. Έως τώρα στηρίζαμε την προσέγγισή μας κατά κόρον στις σημασιολογικές σχέσεις και λιγότερο στις χωρικές σχέσεις. Σε αυτή την ενότητα εστιάζουμε τη σχετική μεθοδολογία προς την κατεύθυνση των τοπολογικών και χωρικών σχέσεων, καθώς οι τελευταίες είναι εξίσου κατάλληλες, για να περιγράψουν με σαφήνεια το πολυμεσικό περιεχόμενο, και θεωρούνται ότι περιγράφουν καλύτερα το οπτικό περιεχόμενο μιας εικόνας. Θα πρέπει να σημειώσουμε στο σημείο αυτό ότι η εξαγωγή των σχέσεων αυτών πραγματοποιείται, όπως θα δούμε, είτε με πλήρως αυτόματο και ανεπίβλεπτο τρόπο, είτε με βάση στατιστικά δεδομένα από το εκάστοτε σύνολο εκπαίδευσης και για το λόγο αυτό θεωρούνται πιο εύρωστες σε σχέση με τις σημασιολογικές σχέσεις. Με βάση αυτά τα είδη σχέσεων, λοιπόν, εισάγουμε μια νέα αναπαράσταση γνώσης, την οποία τη μοντελοποιούμε με τη μορφή μιας βασισμένης στο εννοιολογικό πλαίσιο οντολογίας και τη χρησιμοποιούμε, προκειμένου να ενισχυθεί σημασιολογικά η διαδικασία της πολυμεσικής ανάλυσης. Οι χρησιμοποιούμενες σχέσεις θα πρέπει και πάλι να έχουν νόημα και να συνδυαστούν με τρόπο αποτελεσματικό, ούτως ώστε να προσφέρουν μία ολοκληρωμένη και περιεκτική άποψη της γνώσης. Προκειμένου αυτός ο τύπος οντολογίας να είναι και ιδιαίτερα περιγραφικός, θα πρέπει να περιέχει έναν αντιροσωπευτικό αριθμό ευδιάκριτων, αλλά συνάμα διαφορετικών σχέσεων μεταξύ

των τύπων περιοχής, ώστε να χρησιμοποιηθούν κατά τρόπο βέλτιστο οι βασισμένες στο εννοιολογικό πλαίσιο πληροφορίες που τους περιβάλλουν. Θέλοντας να λάβουμε υπόψη μας την αβεβαιότητα και την ασάφεια των πραγματικών συνθηκών, οι σχέσεις αυτές θα πρέπει να ενσωματώσουν ασαφή χαρακτηριστικά στον ορισμό τους.

Στη συνέχεια, ενσωματώνουμε ένα μάλλον κλασικό υποσύνολο κλασικών (σαφών) τοπολογικών και χωρικών σχέσεων μεταξύ τύπων περιοχής: {γειτονία, πάνω, κάτω, αριστερά, δεξιά}. Η ερμηνεία του συνόλου των τοπολογικών σχέσεων είναι σχετικά απλή και προφανής (βλ. και Πίνακα 7.3). Κατά τα γνωστά, θα προσαρτήσουμε στη συνέχεια έναν βαθμό εμπιστοσύνης σε κάθε μία από τις παραπάνω σχέσεις, έτσι ώστε να συμπεριλάβουμε την ασάφεια των πραγματικών συνθηκών στους υπολογισμούς μας. Προηγουμένως, όμως, θα ορίσουμε μια σειρά από στοιχειώδη σύνολα, τα οποία είναι απαραίτητα για τον καθορισμό και την εξαγωγή των μετέπειτα τοπολογικών σχέσεων. Ειδικότερα, έστω ότι στο πλαίσιο αυτής της ενότητας  $\mathcal{T}$  είναι το σύνολο όλων των τύπων περιοχής,  $\mathcal{P}$  είναι το σύνολο όλων των εικόνων του συνόλου εκπαίδευσης και  $\mathcal{S}$  είναι το σύνολο όλων των περιοχών όλων των εικόνων. Σε κάθε εικόνα  $p$  ορίζουμε:

- $\mathcal{T} = \{t_p\}, p = 1, 2, \dots, N_t$  το σύνολο όλων των τύπων περιοχής ενός δεδομένου οπτικού λεξικού. Όπως έχουμε αναφέρει, το  $\mathcal{T}$  μπορεί να προκύψει από την εφαρμογή ιεραρχικής συσταδοποίησης σε όλα τα στοιχεία του συνόλου  $\mathcal{S}$  και την επιλογή των περιοχών που βρίσκονται πιο κοντά στα κεντρικά στοιχεία (centroids) των συστάδων. Με βάση αυτούς τους τύπους περιοχής αναπτύσσεται και η σχετική οντολογία.
- $\mathcal{S} = \{s_p\}, p = 1, 2, \dots, N_s$  το σύνολο όλων των περιοχών (τμημάτων), όλων των εικόνων, όπως αυτές εξάγονται από κάποιο συγκεκριμένο εργαλείο κατάτμησης.
- $T_p = \{t_k^p\}, k = 1, 2, \dots, N_t^p, p \in \mathcal{P}$ , το σύνολο όλων των τύπων περιοχής (συστάδων) που είναι παρόντες στην εικόνα  $p$ . Όπως είναι προφανές,  $T_p \subset \mathcal{T}$ .
- $D_p = \{d_k^p\}, k = 1, 2, \dots, N_d^p, p \in \mathcal{P}$ , το σύνολο όλων των αρχικών τιμών των ανιχνευτών της εικόνας  $p$ . Οι αρχικές τιμές ανιχνευτών μιας εικόνας προέρχονται από την εφαρμογή κατάλληλων ανιχνευτών υψηλού επιπέδου εννοιών.

Θεωρώντας την  $R_1(t_1, t_2)$  ως μία δυαδική σχέση μεταξύ των τύπων περιοχής  $t_1$  και  $t_2$  και ότι η  $R_2(t_1, t_2)$  είναι η αντίθετη (opposite) σχέση της (π.χ. η “επάνω” είναι η αντίθετη σχέση της “κάτω”), ορίζουμε τελικά την αντίστροφη (inverse) σχέση ως:  $\mathbf{R}^{-1}: R_1^{-1}(t_1, t_2) = R_1(t_2, t_1)$  και την αντίθετη σχέση ως:  $\neg\mathbf{R}: \neg R_1(t_1, t_2) = R_2(t_1, t_2)$ . Η πληθικότητα ενός συνόλου ορίζεται ως  $| \cdot |$ .

Προκειμένου να καταλήξουμε σε ένα ουσιαστικό σύνολο από σχέσεις κατάλληλες για χρήση στα προβλήματα που θέτει η ανάλυση πολυμεσικού περιεχομένου από τη σκοπιά της καθολικής θεώρησης, επεκτείνουμε τις παραπάνω σχέσεις και ορίζουμε ένα σύνολο από ασαφείς τοπολογικές σχέσεις. Με τον τρόπο αυτό, έστω ότι η  $top$  καταδεικνύει οποιαδήποτε τοπολογική σχέση μεταξύ δύο δεδομένων τύπων περιοχής  $t_1$  και  $t_2$ :

$$R_{tt}^{top} = \{r_{t_1, t_2}^{top}\} = \{top(t_1, t_2)\}, t_1, t_2 \in \mathcal{T} \quad (7.19)$$

όπου η  $top \in \{adj, ab, bel, left, rgt\}$ . Καθεμία από τις τοπολογικές σχέσεις συνοψίζεται στον Πίνακα 7.2:

**Πίνακας 7.2:** Προτεινόμενες ασαφείς τοπολογικές σχέσεις.

Όνομα Σύμβολο Ιδιότητες		
γειτονία	$adj$	$adj(t_1, t_2) = adj^{-1}(t_1, t_2)$
πάνω	$ab$	$ab(t_1, t_2) \neq ab^{-1}(t_1, t_2)$
		$\neg ab(t_1, t_2) = bel(t_1, t_2)$
κάτω	$bel$	$bel(t_1, t_2) \neq bel^{-1}(t_1, t_2)$
		$\neg bel(t_1, t_2) = ab(t_1, t_2)$
αριστερά	$left$	$left(t_1, t_2) \neq left^{-1}(t_1, t_2)$
		$\neg left(t_1, t_2) = rgt(t_1, t_2)$
δεξιά	$rgt$	$rgt(t_1, t_2) \neq rgt^{-1}(t_1, t_2)$
		$\neg rgt(t_1, t_2) = left(t_1, t_2)$

Για να ορίσουμε την πρώτη από τις παραπάνω τοπολογικές σχέσεις γειτονία, θα πρέπει πρώτα να ορίσουμε τα ακόλουθα σύνολα:

- $B_{t_1, t_2} = \{(s_1, s_2) \in \mathcal{S}^2 : s_1 \in S_{t_1}, s_2 \in S_{t_2}\}$  είναι το σύνολο όλων των ζευγών περιοχών που συσχετίζονται η πρώτη με τον τύπο περιοχής  $t_1$  και η δεύτερη με τον τύπο περιοχής  $t_2$ .
- $B_{t_1, t_2}^{co} = \{(s_1, s_2) \in B_{t_1, t_2} : co(s_1, s_2)\}$  είναι το υποσύνολο των  $B_{t_1, t_2}$  που περικλείει εκείνα τα ζεύγη που συνυπάρχουν σε μία εικόνα.
- $B_{t_1, t_2}^{adj} = \{(s_1, s_2) \in B_{t_1, t_2} : adj(s_1, s_2)\}$  είναι το υποσύνολο των  $B_{t_1, t_2}$  που περικλείει εκείνα τα ζεύγη που είναι γειτονικά σε μία εικόνα.

Τότε η σχέση γειτονία ορίζεται και υπολογίζεται ως:

$$R_{tt}^{adj} = \{r_{t_1, t_2}^{adj}\}, \quad t_1, t_2 \in \mathcal{T} \quad (7.20)$$

Ο βαθμός στον οποίο ισχύει αυτή η σχέση υπολογίζεται από την εξίσωση 7.21:

$$r_{t_1, t_2}^{adj} = \frac{|B_{t_1, t_2}^{adj}|}{|B_{t_1, t_2}^{co}|} \quad (7.21)$$

Οι υπόλοιπες 4 τοπολογικές σχέσεις  $\{ab, bel, left, rgt\}$  καθορίζονται υιοθετώντας την υπόθεση εργασίας που εισάγει ο αλγόριθμος [176], ότι δηλαδή οι χωρικές σχέσεις μεταξύ δύο σημείων καθορίζονται από τη γωνία που σχηματίζει η γραμμή που περνάει μεταξύ των δύο σημείων και του X-άξονα (βλ. Σχήμα 7.11). Ειδικότερα, για δύο τύπους περιοχής  $t_1$  και  $t_2$  και μια εικόνα  $p$ , αν  $t_1 \in T_p \wedge t_2 \in T_p$ , τότε έστω ότι τα  $Q_{t_1} = \{q_{k_1}^{t_1}\}$ ,  $k_1 = 1, 2, \dots, n_1$  και  $Q_{t_2} = \{q_{k_2}^{t_2}\}$ ,  $k_2 = 1, 2, \dots, n_2$ , καταδεικνύουν τα αντίστοιχα σύνολα από εικονοστοιχεία (pixels). Τα σύνολα αυτά ορίζουν  $n_1 \times n_2$  ζεύγη από εικονοστοιχεία, τα οποία συμβολίζονται με  $(q_{k_1}^{t_1}, q_{k_2}^{t_2}) \in Q_{t_1} \times Q_{t_2}$ . Έστω ότι το  $\Theta$  συμβολίζει τη συλλογή των γωνιών  $\theta_{k_1, k_2} = \angle(q_{k_1}^{t_1}, q_{k_2}^{t_2})$ ,  $q_{k_1}^{t_1} \in Q_{t_1}, q_{k_2}^{t_2} \in Q_{t_2}$ .

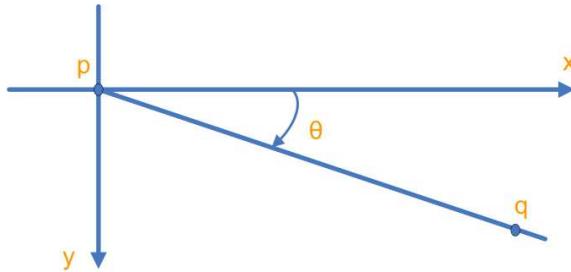
Για κάθε  $\theta \in \Theta$ , έστω ότι  $f_\theta$  είναι η συχνότητα της  $\theta$  στο  $\Theta$ , δηλαδή:

$$f_\theta = |\{(q_{k_1}^{t_1}, q_{k_2}^{t_2}) \in Q_{t_1} \times Q_{t_2} : \angle q_{k_1}^{t_1}, q_{k_2}^{t_2} = \theta\}|, \quad 1 \leq f_\theta \leq n_1 \times n_2 \quad (7.22)$$

Έστω ότι το  $H_\Theta(Q_{t_1}, Q_{t_2}) = \{(\theta, f_\theta)\}$  συμβολίζει το ιστόγραμμα που συσχετίζεται με το  $\Theta$  (βλ. Σχήμα 7.13). Χρησιμοποιώντας τις σχετικές συχνότητες

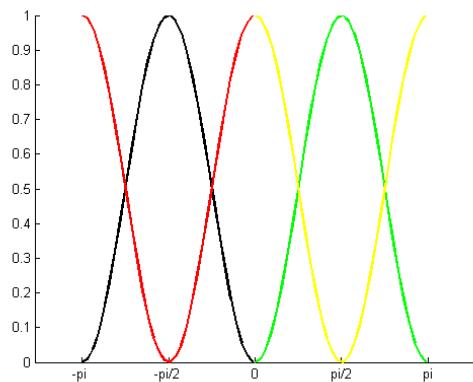
$\tilde{f}_\theta = f_\theta / (n_1 \times n_2)$  και διαιρώντας με τη μεγαλύτερη τιμή του  $H_\Theta$ , η οποία συμβολίζεται με  $\max(H_\Theta)$ , λαμβάνουμε το κανονικοποιημένο ιστόγραμμα  $\tilde{H}_\Theta$ :

$$\tilde{H}_\Theta(\tilde{f}_\theta) = \frac{H_\Theta(\tilde{f}_\theta)}{\max(H_\Theta(\tilde{f}_\theta))} \quad (7.23)$$



Σχήμα 7.11: Γωνία μεταξύ δύο σημείων και του  $X$  άξονα

Υιοθετούμε μία ακόμα υπόθεση εργασίας, ότι δηλαδή οι χωρικές σχέσεις μεταξύ των σημείων  $q_{k_1}^{t_1}$  και  $q_{k_2}^{t_2}$  είναι ασαφή σύνολα, των οποίων οι συναρτήσεις ιδιότητας μέλους δίνονται από τριγωνομετρικές συναρτήσεις της  $\theta$  [176] (βλ. και Σχήμα 7.12), χρησιμοποιώντας το  $\cos^2 \theta$  ή το  $\sin^2 \theta$  σε κατάλληλες περιοχές της γωνίας. Έστω ότι οι  $\mu_{left}(\theta)$ ,  $\mu_{rgt}(\theta)$ ,  $\mu_{ab}(\theta)$  και  $\mu_{bel}(\theta)$  συμβολίζουν τις συναρτήσεις ιδιότητας μέλους των ασαφών συνόλων “αριστερά του”, “δεξιά του”, “πάνω” και “κάτω”, αντίστοιχα.

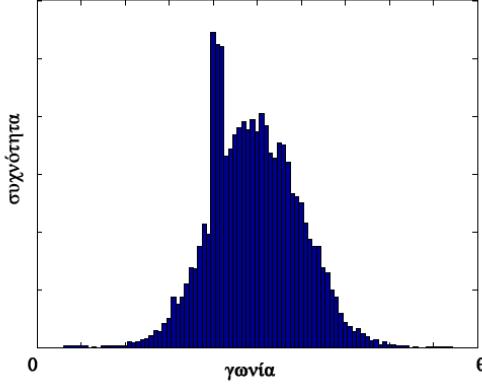


Σχήμα 7.12: Ασαφή σύνολα για τη χωρική σχέση τεσσάρων κατευθύνσεων: “αριστερά του”, “δεξιά του”, “πάνω” και “κάτω”.

Το ιστόγραμμα  $\tilde{H}_\Theta$  θεωρείται ως ένα αμαρκάριστο ασαφές σύνολο, το οποίο “συλλαμβάνει τη χωρική συσχέτιση μεταξύ των  $Q_{t_1}, Q_{t_2}$ ”. Με δεδομένα τα προηγούμενα ασαφή σύνολα ο βαθμός του  $\tilde{H}_\Theta$  λαμβάνεται χρησιμοποιώντας τη λειτουργία συμβατότητας μιας κατανομής σε ένα ασαφές σύνολο. Η συμβατότητα μιας κατανομής  $F$  σε ένα ασαφές σύνολο  $G$  είναι η επέκταση της αξιολόγησης ενός ασαφούς συνόλου σε ένα σημείο προς ένα ασαφές σύνολο. Το αποτέλεσμα είναι ένα ασαφές σύνολο  $CP(F; G)$ , του οποίου η συνάρτηση ιδιότητας μέλους  $\mu_{CP(F; G)}$  λαμβάνεται από την αρχή της επέκτασης. Στη συγκεκριμένη περίπτωση το ιστόγραμμα  $H_\Theta$  είναι η κατανομή  $F$  και το ασαφές σύνολο που συνδέεται με μια από τις τέσσερις χωρικές σχέσεις είναι το ασαφές σύνολο  $G$ . Κατόπιν, για κάθε μια από τις τέσσερις χωρικές σχέσεις  $X$ ,

$X \in \{left, rgt, ab, bel\}$ , το ασαφές σύνολο συμβατότητας  $C_X(\tilde{H}_\Theta; \mu_X)$  περιγράφεται από την συνάρτηση ιδιότητας μέλους  $\mu_{C_X}(F; \tilde{H}_\Theta; \mu_X) : [0, 1] \rightarrow [0, 1]$ :

$$\mu_{C_X(\tilde{H}_\Theta; \mu_X)}(v) = \begin{cases} sup_{\theta: v=\mu_X(\theta)} H_\Theta(\tilde{f}_\theta), & if \mu_G^{-1}(v) \neq \emptyset \\ 0, & if \mu_G^{-1}(v) = \emptyset \end{cases} \quad (7.24)$$



Σχήμα 7.13: Ιστόγραμμα που συσχετίζεται με το  $\Theta$ .

Τελικά, ο βαθμός στον οποίο ισχύει μία χωρική σχέση λαμβάνεται από το κέντρο της βαρύτητας του ασαφούς συνόλου συμβατότητας  $C_X(\tilde{H}_\Theta; \mu_X)$ . Σε αυτή τη μέθοδο [139] η απασαφοποιημένη (defuzzified) τιμή  $d_X(v)$  ορίζεται ως η τιμή μέσα στο πεδίο τιμών της μεταβλητής  $v$ , για την οποία η περιοχή κάτω από το γράφημα της συνάρτησης ιδιότητας μέλους  $\mu_{C_X(\tilde{H}_\Theta; \mu_X)}$  χωρίζεται σε δύο ίσες περιοχές. Η τιμή αυτή υπολογίζεται από την εξίσωση:

$$d_X(v) = \frac{\int_0^1 C_X(v) v dv}{\int_0^1 C_X(v) dv} \quad (7.25)$$

Θα πρέπει να σημειώσουμε εδώ ότι, εφόσον η  $\tilde{H}_\Theta$  είναι μία συνάρτηση του  $Q_{t_1}$  και  $Q_{t_2}$ ,  $d_X(v) = d_X(v; t_1, t_2) = d_X(t_1, t_2)$ . Προκειμένου, τώρα, να υπολογίσουμε τις παραπάνω τοπολογικές σχέσεις “πάνω”, “κάτω”, “αριστερά” και “δεξιά”, θα πρέπει να ορίσουμε επιπρόσθετα τα ακόλουθα σύνολα:

- $B_{t_1, t_2}^{ab} = \{(s_1, s_2) \in B_{t_1, t_2} : max\{d_X(t_1, t_2), X \in \{ab, bel, left, rgt\}\} = d_{ab}(t_1, t_2)\}$ . Είναι το υποσύνολο των  $B_{t_1, t_2}$  που περικλείει τα ζεύγη, όπου ο  $t_1$  είναι πάνω από τον  $t_2$ .
- $B_{t_1, t_2}^{bel} = \{(s_1, s_2) \in B_{t_1, t_2} : max\{d_X(t_1, t_2), X \in \{ab, bel, left, rgt\}\} = d_{bel}(t_1, t_2)\}$ . Είναι το υποσύνολο των  $B_{t_1, t_2}$  που περικλείει τα ζεύγη, όπου ο  $t_1$  είναι κάτω από τον  $t_2$ .
- $B_{t_1, t_2}^{left} = \{(s_1, s_2) \in B_{t_1, t_2} : max\{d_X(t_1, t_2), X \in \{ab, bel, left, rgt\}\} = d_{left}(t_1, t_2)\}$ . Είναι το υποσύνολο των  $B_{t_1, t_2}$  που περικλείει τα ζεύγη, όπου ο  $t_1$  είναι αριστερά του  $t_2$ .
- $B_{t_1, t_2}^{rgt} = \{(s_1, s_2) \in B_{t_1, t_2} : max\{d_X(t_1, t_2), X \in \{ab, bel, left, rgt\}\} = d_{rgt}(t_1, t_2)\}$ . Είναι το υποσύνολο των  $B_{t_1, t_2}$  που περικλείει τα ζεύγη, όπου ο  $t_1$  είναι δεξιά του  $t_2$ .

Όπως είναι πλέον εμφανές, οι αντίστοιχες σχέσεις μπορούν τώρα να υπολογιστούν ως εξής:

- $r_{t_1, t_2}^{ab} = \frac{|B_{t_1, t_2}^{ab}|}{|B_{t_1, t_2}^{co}|}$ ,  $t_1, t_2 \in \mathcal{T}$ .
- $r_{t_1, t_2}^{bel} = \frac{|B_{t_1, t_2}^{bel}|}{|B_{t_1, t_2}^{co}|}$ ,  $t_1, t_2 \in \mathcal{T}$ .
- $r_{t_1, t_2}^{left} = \frac{|B_{t_1, t_2}^{left}|}{|B_{t_1, t_2}^{co}|}$ ,  $t_1, t_2 \in \mathcal{T}$ .
- $r_{t_1, t_2}^{rgt} = \frac{|B_{t_1, t_2}^{rgt}|}{|B_{t_1, t_2}^{co}|}$ ,  $t_1, t_2 \in \mathcal{T}$ .

Τέλος, εισάγουμε και μια ιδιαίτερα σημαντική σχέση, τη σχέση συνύπαρξης (co-occurrence), η οποία ορίζεται με στατιστικό τρόπο πάνω στο σύνολο εκπαίδευσης των δεδομένων. Ορίζουμε:

$$R_{tt}^{co} = \{r_{t_1, t_2}^{co}\} = \{co(t_1, t_2)\}, \quad t_1, t_2 \in \mathcal{T} \quad (7.26)$$

όπου:

$$co(t_1, t_2) = \frac{|\{p \in \mathcal{P} : t_1 \in T_p \wedge t_2 \in T_p\}|}{|\{p \in \mathcal{P} : t_1 \in T_p \vee t_2 \in T_p\}|} \quad (7.27)$$

Κατά συνέπεια, καταλήγουμε σε ένα σύνολο 6 ασαφών σχέσεων (βλ. Πίνακα 7.3), που προέρχονται από τις αντίστοιχες κλασικές τοπολογικές και χωρικές σχέσεις. Όπως ακριβώς ορίστηκε στο κεφάλαιο 3, η εκάστοτε “ασαφοποιημένη” εκδοχή μιας

**Πίνακας 7.3:** Σχέσεις εννοιολογικού πλαισίου στην περίπτωση των τύπων περιοχής.

Όνομα	Αντίστροφο Σύμβολο	Ερμηνεία
Γειτονία	-	$adj(a, b)$ ένας τύπος περιοχής είναι παρακείμενος με έναν άλλο τύπο περιοχής
Συνύπαρξη	-	$co(a, b)$ ένας τύπος περιοχής συνυπάρχει με έναν άλλο τύπο περιοχής
Αριστερά	Δεξιά	$left(a, b)$ ένας τύπος περιοχής είναι αριστερά ενός άλλου τύπου περιοχής
Δεξιά	Αριστερά	$rgt(a, b)$ ένας τύπος περιοχής είναι δεξιά ενός άλλου τύπου περιοχής
Πάνω	Κάτω	$ab(a, b)$ ένας τύπος περιοχής είναι πάνω από έναν άλλο τύπο περιοχής
Κάτω	Πάνω	$bel(a, b)$ ένας τύπος περιοχής είναι κάτω από έναν άλλο τύπο περιοχής

οντολογίας τύπων περιοχής  $\mathcal{O}_T^f$ , η οποία εξαρτάται πάντοτε από την επιλεγμένη θεματική περιοχή του ενδιαφέροντός μας, θα έχει τη μορφή:

$$\mathcal{O}_T^f = \{T, r_{t_i, t_j}\}, \quad i, j = 1 \dots n, \quad i \neq j \quad (7.28)$$

όπου το  $T$  αναπαριστά το σύνολο όλων των πιθανών τύπων περιοχής, η

$$r_{t_i, t_j} : T \times T \rightarrow [0, 1] \quad (7.29)$$

καταδεικνύει μια ασαφή οντολογική σχέση μεταξύ δύο τύπων περιοχής  $t_i, t_j$  και η

$$R_{t_i, t_j} = \{R_{tt}^{top}, R_{tt}^{co}\} = \{adj, ab, bel, rgt, left, co\} \quad (7.30)$$

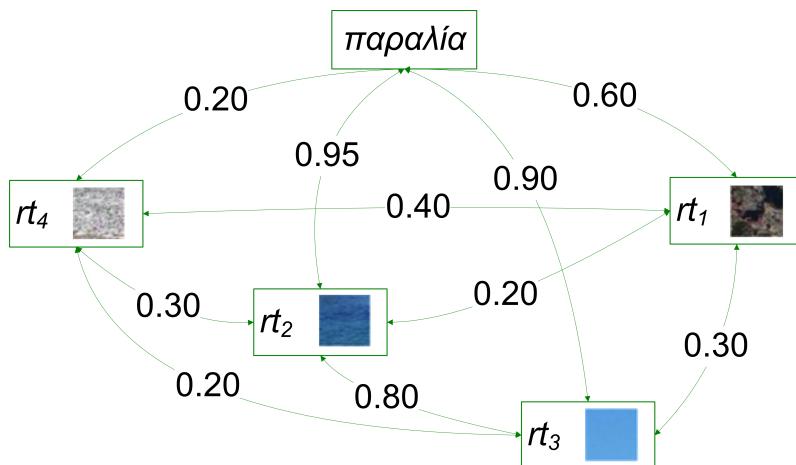
χαρακτηρίζει την εκάστοτε κλασική (σαφή) σχέση μεταξύ δύο τύπων περιοχής. Ο τελικός συνδυασμός των σχέσεων

$$\mathcal{C}_T = (\bigcup_i r_{t_i, t_j}^{p_i}), \quad p_i \in \{-1, 0, 1\}, \quad i = 1 \dots n \quad (7.31)$$

είναι στην ουσία ένας γράφος RDF [316] και αποτελεί το αφηρημένο μοντέλο γνώσης του εννοιολογικού πλαισίου, που θα χρησιμοποιηθεί κατά τη φάση της ανάλυσης (βλ. Σχήμα 7.14). Η τιμή του  $p_i$  καθορίζεται και πάλι από τη σημασιολογία κάθε σχέσης  $R_{t_i, t_j}$  που χρησιμοποιείται στην κατασκευή της συνδυασμένης σχέσης  $C_T$ . Ειδικότερα:

- $p_i = 1$ , αν η σημασιολογία της  $R_{t_i, t_j}$  υπονοεί ότι θα πρέπει να θεωρηθεί ως έχει
- $p_i = -1$ , αν η σημασιολογία της  $R_{t_i, t_j}$  υπονοεί ότι θα πρέπει να θεωρηθεί το αντίστροφό της
- $p_i = 0$ , αν η σημασιολογία της  $R_{t_i, t_j}$  δεν επιτρέπει τη συμμετοχή της στην κατασκευή της συνδυασμένης σχέσης  $C_T$ .

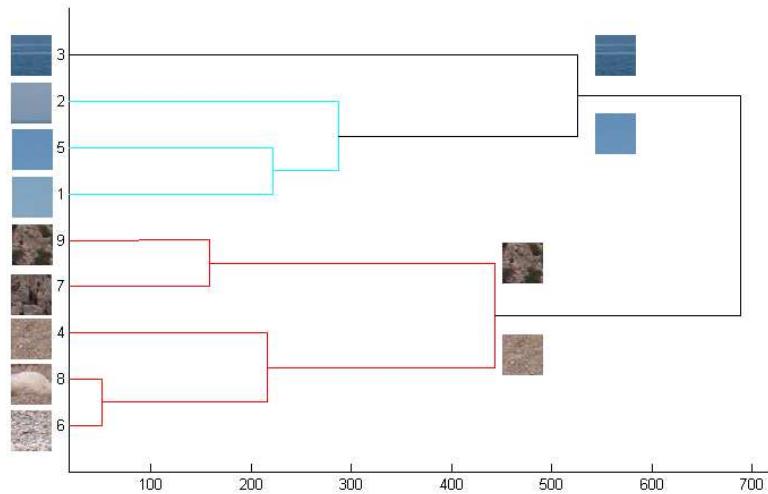
Ο γράφος του προτεινόμενου μοντέλου περιέχει κορυφές (δηλ. τύπους περιοχής) και ακμές (δηλ. βασισμένες στο εννοιολογικό πλαίσιο ασαφείς σχέσεις μεταξύ τύπων περιοχής). Ο βαθμός εμπιστοσύνης για κάθε ακμή αντιπροσωπεύει την ασάφεια του μοντέλου και περιγράφεται με την τεχνική του RDF reification [318]. Θα πρέπει να σημειώσουμε εδώ ότι, καθώς κάθε τύπος περιοχής έχει και διαφορετική πιθανότητα εμφάνισης στη σκηνή, ένα επίπεδο μοντέλο εννοιολογικού πλαισίου δε θα ήταν αρκετό σε αυτή την περίπτωση.



Σχήμα 7.14: Ένα τμήμα της οντολογίας τύπων περιοχής για τη θεματική περιοχή παραλία.

#### 7.4.2.1 Επίδραση οπτικού εννοιολογικού πλαισίου

Αφους υπολογίσουμε ένα κατάλληλο διάνυσμα αναπαράστασης για μια εικόνα (βλ. υποενότητα 7.3.3), εφαρμόζουμε μια τροποποιημένη έκδοση του βασισμένου στο εννοιολογικό πλαίσιο αλγορίθμου αναπροσαρμογής του βαθμού εμπιστοσύνης, έτσι ώστε να ικανοποιηθούν οι ανάγκες του συγκεκριμένου ερευνητικού προβλήματος. Η μεθοδολογία που ακολουθείται παρουσιάζεται στο Σχήμα 7.17. Η διαδικασία αυτή αποτελεί αναπόσπαστο κομμάτι της ανάλυσης, μιας και διαμορφώνει απευθείας το κατασκευαζόμενο διάνυσμα αναπαράστασης. Παρέχει, δε, μία βελτιστοποιημένη σε σχέση με την απλή περίπτωση επανεκτίμηση των βαθμών εμπιστοσύνης των τελικών ανιχνευμένων εννοιών υψηλού επιπέδου για τους επιλεγμένους τύπους περιοχής. Με άλλα λόγια, η προτεινόμενη τεχνική ενημερώνει τις τιμές κάθε διανύσματος αναπαράστασης, επιτρέποντας μια βελτιστοποιημένη διαδικασία εκπαίδευσης του ταξινομητή και επιτυγχάνοντας έτσι σαφώς καλύτερα αποτελέσματα αξιολόγησης.



Σχήμα 7.15: Επιλογή τύπων περιοχής μέσω ιεραρχικής συσταδοποίησης.

```

<rdf:Description rdf:about="#Relation1">
  <rdf:subject rdf:resource="#dom;rt1"/>
  <rdf:predicate rdf:resource="#dom;Part"/>
  <rdf:object>rdf:resource="#dom;rt2"</rdf:object>
  <rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Statement"/>
  <context:Part rdf:datatype="http://www.w3.org/2001/XMLSchema#float">0.85</context:Part>
</rdf:Description>

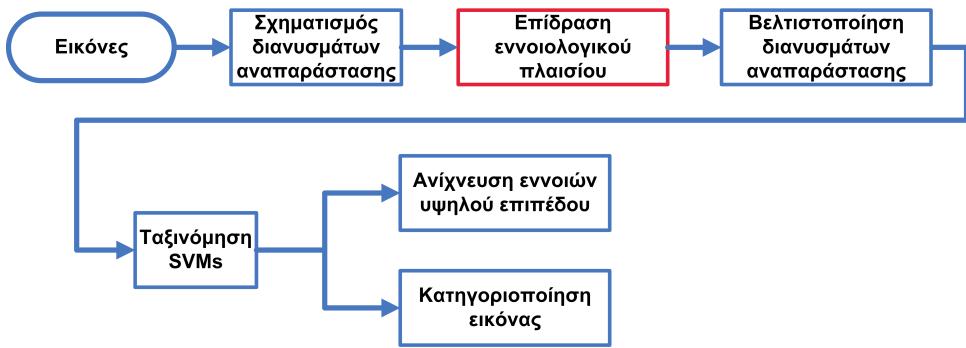
```

Σχήμα 7.16: Τμήμα RDF οντολογίας τύπων περιοχής.

Αν θέλαμε να ορίσουμε με έναν πιο επίσημο τρόπο το πρόβλημα που καλείται να εξετάσει αυτή η μεθοδολογία, αυτό συνοψίζεται στην ακόλουθη δήλωση: ο αλγόριθμος ενισχυμένης ανάλυσης του εννοιολογικού πλαισίου ταξινόμησης τύπων περιοχής αναπροσαρμόζει με έναν κατάλληλο τρόπο τις αρχικές τιμές του κατασκευαζόμενου διανύσματος αναπαράστασης. Το τελευταίο βασίζεται στις εκάστοτε τιμές των τύπων περιοχής, οι οποίοι παράγονται από το προηγούμενο βήμα της ανάλυσης των τύπων περιοχής (βλ. ενότητα 7.3). Σε αυτή την ενότητα τα προβλήματα που πρόκειται να εξεταστούν περιλαμβάνουν το πώς αναπροσαρμόζονται οι αρχικοί βαθμοί ιδιότητας μέλους και το πώς χρησιμοποιούν το εννοιολογικό πλαίσιο, για να επηρεάσουν τα συνολικά αποτελέσματα της - βοηθούμενης από τη γνώση - ανάλυσης προς μία συνολικά υψηλότερη απόδοση. Μια εκτίμηση του βαθμού ιδιότητας μέλους κάθε ενδιάμεσου επιπέδου τύπου περιοχής προέρχεται από τις άμεσες και έμμεσες σχέσεις των τελευταίων με άλλους τύπους περιοχής στο γράφο, χρησιμοποιώντας έναν κατάλληλο δείκτη συμβατότητας ή αλλιώς μια μετρική απόστασης. Ανάλογα με τη φύση και την πυκνότητα τύπων των θεματικών περιοχών, που περικλείονται στην οντολογία, κατάλληλους δείκτες αποτελούν οι τελεστές *max* ή *min*, αντίστοιχα. Φυσικά η ιδανική μετρική απόστασης για δύο τύπους περιοχής είναι, όπως και στην περίπτωση των υψηλού επιπέδου εννοιών, αυτή που ποσοτικοποιεί το σημασιολογικό συσχετισμό τους. Για το εν λόγω πρόβλημα πιστεύουμε ότι ο τελεστής *max* αποτελεί το καταλληλότερο μέτρο του συσχετισμού και για τους δύο.

Η γενική δομή του αλγορίθμου αναπροσαρμογής των βαθμών ιδιότητας μέλους μετασχηματίζεται στην περίπτωση αυτή ως ακολούθως. Θα πρέπει να σημειωθεί στο σημείο αυτό ότι ανάλογα με τα χαρακτηριστικά της εκάστοτε θεματικής περιοχής εφαρμογής επιβάλλεται η χρήση και μιας διαφορετικής μετρικής ομοιότητας (ή

## Κεφάλαιο 7. Εννοιολογικό πλαίσιο ταξινόμησης



Σχήμα 7.17: Μεθοδολογία επίδρασης εννοιολογικού πλαισίου ταξινόμησης τύπων περιοχής

ανομοιότητας) θεματικών περιοχών με τη μορφή μιας παραμέτρου κανονικοποίησης  $\mu \in [0, 1]$ .

- Για κάθε τύπο περιοχής  $t$  περιγράφουμε το ασαφές σύνολο  $L_t = F(T)$ , χρησιμοποιώντας τον ευρέως διαδεδομένο συμβολισμό αθροίσματος [177]  $L_t = \sum_{i=1}^{|T|} t_i/w_i = \{t_1/w_1, t_2/w_2, \dots, t_n/w_n\}$ , όπου το  $w_i$  περιγράφει τη συνάρτηση συμμετοχής:  $w_i = \mu_{L_t}(t_i)$ .
- Για κάθε τύπο περιοχής  $t_i$  στο ασαφές σύνολο  $L_t$  με βαθμό συμμετοχής  $w_i$ , λαμβάνουμε την ιδιαίτερη πληροφορία εννοιολογικού πλαισίου με τη μορφή των σχέσεών του με το σύνολο οποιωνδήποτε άλλων τύπων περιοχής:  $\{r_{t_i, t_j} : t_i, t_j \in T, i \neq j\}$ .
- Στην περίπτωση πολλαπλών σχέσεων ενδιάμεσου επιπέδου τύπων περιοχής, εφαρμόζουμε ένα ενδιάμεσο βήμα συνάθροισης για την εκτίμηση του βαθμού  $w_i$ , εξετάζοντας επιπλέον και την έννοια της σχετικότητας πλαισίου  $cr_{t_i}$  (βλ. ενότητα 6.3.1 και δημοσίευση [183]).
- Υπολογίζουμε το νέο βαθμό ιδιότητας μέλους  $w_i$ , λαμβάνοντας υπόψη το μέτρο ομοιότητας κάθε περιοχής σύμφωνα με τον αναδρομικό τύπο:

$$w_i^n = w_i^{n-1} - \mu(w_i^{n-1} - cr_{t_i}) \quad (7.32)$$

όπου το  $n$  δείχνει την επανάληψη που χρησιμοποιείται<sup>5</sup>. Ισοδύναμα για μία αυθαίρετη επανάληψη  $n$ :

$$w_i^n = (1 - \mu)^n \cdot w_i^0 + (1 - (1 - \mu)^n) \cdot cr_{t_i} \quad (7.33)$$

όπου το  $w_i^0$  αντιπροσωπεύει τον αρχικό βαθμό ιδιότητας μέλους για τον τύπο περιοχής  $t_i$ .

Με τον τρόπο αυτό επιτυγχάνουμε την αντικατάσταση των αρχικών τιμών/περιεχομένων του διανύσματος αναπαράστασης με βελτιστοποιημένες ως προς τη σημασιολογία τους νέες τιμές εννοιολογικού πλαισίου ταξινόμησης. Η συμβολή της μεθοδολογίας αυτής είναι ιδιαίτερα σημαντική, όπως θα δούμε και στην ενότητα 7.5.2 του κεφαλαίου.

<sup>5</sup> Σύμφωνα με τα πειράματά μας χαρακτηριστικές τιμές για το  $n$  αποτελούν οι 3, 4 και 5.

### 7.4.3 Μεικτό εννοιολογικό πλαίσιο

Σε αυτή την ενότητα επεκτείνουμε περαιτέρω την προτεινόμενη αντίληψη του εννοιολογικού πλαισίου ενδιάμεσου επιπέδου και προτείνουμε μια καινοτόμο προσέγγιση για την αναπαράσταση γνώσης με τη μορφή μιας επεκτεταμένης, μεικτής οντολογίας πλαισίου. Εισάγουμε μία πρωτότυπη αναπαράσταση για τη γνώση, η οποία έχει τη μορφή μιας επεκτεταμένης οντολογίας εννοιολογικού πλαισίου και η οποία πρωτοπαρουσιάστηκε στη δημοσίευση [186]. Από αυτή την άποψη εμπλουτίζουμε την έννοια της κλασικής οντολογίας του οπτικού εννοιολογικού πλαισίου με ενδιάμεση περιγραφή, η οποία μπορεί να περιγραφεί σημασιολογικά, αλλά δεν εκφράζουν ούτε μια υψηλού, ούτε μία χαμηλού επιπέδου έννοια. Κατά συνέπεια, εστιάζουμε σε μια ενοποιημένη πολυμεσική αναπαράσταση, συνδυάζοντας χαμηλού και υψηλού επιπέδου πληροφορίες με έναν “ενδιάμεσου επιπέδου” αποδοτικό τρόπο και τη συνδέουμε με την κλασική οντολογία του εννοιολογικού πλαισίου, καθορίζοντας νέες (τοπολογικές, χωρικές) και επεκτείνοντας παλαιότερες (σημασιολογικές) σχέσεις τους.

Ως εκ τούτου, η προτεινόμενη οντολογία περιγράφεται από ένα σύνολο υψηλού επιπέδου εννοιών, ένα σύνολο τύπων περιοχής και ένα σύνολο από σχέσεις μεταξύ τους. Το σύνολο των εννοιών ορίζεται εκ των προτέρων από έναν εμπειρογνώμονα της θεματικής περιοχής. Στη γενική περίπτωση αυτό το είδος οντολογίας  $\mathcal{O}_M$  μπορεί να αποδομηθεί σε τρία μέρη, το σύνολο  $C$  όλων των υψηλού επιπέδου εννοιών, το σύνολο  $T$  όλων των τύπων περιοχής και το συνόλο  $R_{x_i, y_j}$  όλων των δυαδικών σχέσεων μεταξύ οποιουδήποτε συνδυασμού εννοιών και τύπων περιοχής<sup>6</sup>. Πιο συγκεκριμένα, μπορεί να υπάρξουν μία ή και περισσότερες σχέσεις μεταξύ δύο υψηλού επιπέδου εννοιών, δύο δεδομένων τύπων περιοχής ή μιας υψηλού επιπέδου έννοιας και ενός τύπου περιοχών. Ειδικότερα:

$$\mathcal{O}_M = \{C, T, R_{x_i, y_j}\}, \quad R_{x_i, y_j} : X \times Y \rightarrow \{0, 1\}, \quad i, j = 1 \dots m, \quad i \neq j \quad (7.34)$$

όπου  $X, Y \in C \cup T$ . Με άλλα λόγια, δεδομένου ότι η προτεινόμενη οντολογία δεν περιορίζει τις σχέσεις μόνο μεταξύ των μελών του συνόλου  $C$  ή  $T$ , είναι δυνατό να είναι  $X \in C$  και  $Y \in T$  ή και ανάποδα. Επίσης, δεδομένου ότι για κάθε εφαρμοσμένη σημασιολογική σχέση υπάρχει το αντίστροφό της,  $m = |X| = |Y|$ , όπου το  $|\bullet|$  καταδεικνύει την πληθυκότητα ενός συνόλου.

Τελικά, χρησιμοποιούμε ένα σύνολο σχέσεων (βλ. Πίνακα 7.4), που προέρχονται αφενός από το σύνολο των MPEG-7 σημασιολογικών σχέσεων, αφετέρου δε από ένα σύνολο τοπολογικών και χωρικών σχέσεων μεταξύ στοιχείων της εικόνας. Ειδικότερα, επαναχρησιμοποιούνται οι 5 τοπολογικές σχέσεις και η σχέση συνύπαρξης που εισήχθησαν στην υποενότητα 7.4.2, καθώς και 6 επιλεγμένες σημασιολογικές σχέσεις  $\{Sim, Acc, P, Comp, Sp, Pr\}$ . Θεωρούμε ότι για τη συγκεκριμένη μεθοδολογία οι σχέσεις αυτές είναι οι πλέον κατάλληλες και τις επαναπροσδιορίζουμε με έναν τρόπο, έτσι ώστε να ενσωματωθεί σε αυτές η απαιτούμενη ασάφεια, δηλ., ένας βαθμός εμπιστοσύνης συνδέεται σε κάθε σχέση και βοηθά στη διάκριση μεταξύ των αντικειμένων, εκθέτοντας τα παρόμοια οπτικά χαρακτηριστικά. Το σύνολο των χρησιμοποιημένων σχέσεων περιέχει κατά συνέπεια τόσο χωρικές, όσο και σημασιολογικές σχέσεις, οι οποίες λαμβάνονται με τη χρησιμοποίηση είτε μιας αυτόματης στατιστικής προσέγγισης στο σύνολο των στοιχείων εκπαίδευσης (τεχνική που, όπως είδαμε στην

<sup>6</sup>Στη συνέχεια, θα χρησιμοποιήσουμε τον όρο οντότητες, όταν αναφερόμαστε είτε σε έννοιες, είτε σε τύπους περιοχής.

## Κεφάλαιο 7. Εννοιολογικό πλαίσιο ταξινόμησης

προηγούμενη υποενότητα, χρησιμοποιείται κυρίως για τον καθορισμό των χωρικών σχέσεων), είτε της γνώμης ενός ειδήμουνα (τεχνική που χρησιμοποιείται κυρίως για τον καθορισμό των σημασιολογικών σχέσεων).

**Πίνακας 7.4: Σχέσεις εννοιολογικού πλαισίου μεταξύ οντοτήτων.**

Όνομα	Αντίστροφο	Σύμβολο	Ερμηνεία	$C \times C$	$T \times T$	$C \times T$	Ορισμός
Παρόμοιος	Παρόμοιος	$Sim(a, b)$	ομοιότητα μεταξύ $a$ και $b$		•		στατιστικά
Συνοδός	ΣυνοδόςΤου	$Acc(a, b)$	συνύπαρξη των $a$ και $b$	•		•	ειδήμων/στατιστικά
Μέρος	ΜέροςΤου	$P(a, b)$	$\eta$ $a$ είναι μέρος της $b$	•	•	•	ειδήμων/στατιστικά
Συστατικό	ΣυστατικόΤου	$Comp(a, b)$	συνδυάζει την $a$ με τη $b$	•	•	•	ειδήμων/στατιστικά
Ειδίκευση	Γενίκευση	$Sp(a, b)$	$\eta$ $b$ εξειδικεύει τη σημασία της $a$	•			ειδήμων
Ιδιότητα	ΙδιότηταΤου	$Pr(a, b)$	$\eta$ $b$ είναι μια ιδιότητα της $a$		•	•	ειδήμων/στατιστικά
Γειτονία	-	$adj(a, b)$	$\eta$ $a$ είναι παρακείμενη με τη $b$		•		στατιστικά
Συνύπαρξη	-	$co(a, b)$	$\eta$ $a$ συνυπάρχει με τη $b$		•		στατιστικά
Αριστερά	Δεξιά	$left(a, b)$	$\eta$ $a$ είναι αριστερά της $b$		•		στατιστικά
Δεξιά	Αριστερά	$rgt(a, b)$	$\eta$ $a$ είναι δεξιά της $b$		•		στατιστικά
Πάνω	Κάτω	$ab(a, b)$	$\eta$ $a$ είναι πάνω από τη $b$		•		στατιστικά
Κάτω	Πάνω	$bel(a, b)$	$\eta$ $a$ είναι κάτω από τη $b$		•		στατιστικά

Κάθε οντότητα μπορεί να σχετίζεται με κάποια άλλη, χρησιμοποιώντας μια ή περισσότερες από τις προαναφερθείσες και βασισμένες στο εννοιολογικό πλαίσιο ασαφείς σχέσεις. Εντούτοις, θα πρέπει να σημειώσουμε εδώ ότι δεν είναι όλες οι σχέσεις κατάλληλες για οποιουδήποτε τύπου ζεύγη οντοτήτων. Παραδείγματος χάριν, η σχέση Παρόμοιος δεν ορίζεται μεταξύ δύο υψηλού επιπέδου εννοιών ή μεταξύ μιας υψηλού επιπέδου έννοιας και ενός τύπου περιοχής, δηλ. η θάλασσα δεν μπορεί να σχετίζεται με την άμμο χρησιμοποιώντας αυτήν την σχέση. Εντούτοις, η ομοιότητα είναι μια σημαντική μετρική για τη συσχέτιση δύο τύπων περιοχής και μπορεί να υπολογιστεί συγκρίνοντας τα χαρηλού επιπέδου χαρακτηριστικά γνωρίσματά τους. Οι πιθανές σχέσεις για κάθε ζευγάρι οντοτήτων απεικονίζονται στις τελευταίες τρεις στήλες του Πίνακα 7.4.

Ενώ η ερμηνεία και χρήση των τοπολογικών σχέσεων του παραπάνω Πίνακα δεν περιέχει αμφισημείες, στην περίπτωση των σημασιολογικών σχέσεων απαιτείται μια περαιτέρω εξήγησή τους. Η σχέση Παρόμοιος μπορεί να βρει εφαρμογή μόνο μεταξύ τύπων περιοχής, όπου π.χ. ένας “μπλε” τύπος περιοχής είναι Παρόμοιος με έναν “γαλάζιο” τύπο περιοχής. Η σχέση Συνοδός δείχνει τη συνύπαρξη δύο εννοιών ή μιας έννοιας και ενός τύπου περιοχής μέσα σε μια εικόνα/ένα βίντεο: η έννοια νερό αποτελεί Συνοδό της έννοιας θάλασσα. Η σχέση Μέρος δείχνει ότι μια οντότητα είναι μέρος μιας άλλης. Για παράδειγμα, η έννοια ουρανός μπορεί να είναι ένα Μέρος της έννοιας ύπαιθρος. Η σχέση Συστατικό δείχνει το συνδυασμό μιας οντότητας με μια άλλη και χρησιμοποιείται, όταν οι οντότητες αυτές συνδυάζονται και δημιουργούν μια άλλη. Δύο ή περισσότερες υψηλού επιπέδου έννοιες μπορούν να συνδυαστούν, για να διαμορφώσουν μια άλλη υψηλού επιπέδου έννοια, ενώ δύο ή περισσότεροι τύποι περιοχής μπορούν να συνδυαστούν όχι μόνο για να διαμορφώσουν έναν άλλο τύπο περιοχής, αλλά και μια υψηλού επιπέδου έννοια. Παραδείγματος χάριν, όταν συνδυαστούν ο ουρανός, η θάλασσα και η άμμος, διαμορφώνουν την υψηλού επιπέδου έννοια παραλία. Από την άλλη πλευρά, είναι προφανές ότι ένας συνδυασμός δύο τύπων περιοχής είναι ένας άλλος τύπος περιοχής, εντούτοις, όταν συνδυαστούν ένας “καφέ” και ένας “πράσινος” τύπος περιοχής, διαμορφώνουν την υψηλού επιπέδου έννοια δένδρο. Η σχέση Ειδίκευση ορίζεται μόνο για υψηλού επιπέδου έννοιες και επιτρέπει σε μια έννοια να εξειδικεύσει την ερμηνεία μιας άλλης έννοιας. Παραδείγματος χάριν, η κορομηλιά εξειδικεύει το δένδρο, το οποίο εξειδικεύεται επίσης από τη βλάστηση. Τέλος, μια πράσινη περιοχή μπορεί να είναι μια Ιδιότητα της βλάστησης.

Όλες οι ανωτέρω σχέσεις διαμορφώνουν ένα επεκτεταμένο μοντέλο ενδιάμεσου επιπέδου εννοιολογικού πλαισίου, το οποίο όμως και πάλι μπορεί να αναπαρασταθεί ως ένας και μόνο γράφος: κάθε κόμβος του γράφου αντιπροσωπεύει μία έννοια ή έναν τύπο περιοχής και κάθε ακμή μεταξύ δύο κόμβων, μια βασισμένη στο εννοιολογικό πλαίσιο σχέση μεταξύ των αντίστοιχων οντοτήτων. Επιπλέον, ένας σχετικός βαθμός εμπιστοσύνης συνδέεται με κάθε ακμή, εκφράζοντας την επιθυμητή ασάφεια μέσα στο μοντέλο του πλαισίου. Μια ακμή μεταξύ ενός δεδομένου ζευγαριού εννοιών παράγεται με βάση το σύνολο των ασαφών σχέσεων του εννοιολογικού πλαισίου που είναι σημαντικές για το συγκεκριμένο ζευγάρι. Παραδείγματος χάριν, η ακμή μεταξύ των εννοιών *βράχος* και *άμμος* παράγεται μόνο από τη σχέση *Συνοδός*, ενώ η ακμή μεταξύ *νερού* και *θάλασσας* χρησιμοποιεί τις σχέσεις *Ειδίκευση* και *Μέρος*, προκεμένου να κατασκευαστεί. Υπό την ίδια έννοια δύο τύποι περιοχής, δηλ. ένας “πράσινος” και ένας “μπλε” τύπος περιοχής, μπορεί να χρησιμοποιήσουν τις σχέσεις *Παρόμοιος*, *Συνοδός* και *Συστατικό*.

Όπως ορίζεται στον Klir [139], μια ασαφής σχέση στο  $T$  είναι μια συνάρτηση  $\mathcal{R}_{x_i, x_j} : X \times Y \rightarrow [0, 1]$  και η αντίστροφη σχέση της ορίζεται ως  $\mathcal{R}_{x_i, y_j}^{-1} = \mathcal{R}_{y_j, x_i}$ . Με βάση τις παραπάνω σχέσεις μια εξαρτώμενη από την εκάστοτε θεματική περιοχή, “ασαφοποιημένη” εκδοχή της προτεινόμενης οντολογίας μπορεί να περιγραφεί από την  $\mathcal{O}_M^f$ , ως εξής:

$$\mathcal{O}_M^f = \{C, T, \mathcal{R}_{x_i, x_j}\}, \quad i, j = 1, \dots, m, \quad i \neq j \quad (7.35)$$

όπου το  $C$  αναπαριστά και πάλι το σύνολο όλων των υψηλού επιπέδου εννοιών, το  $T$  το σύνολο όλων των πιθανών τύπων περιοχής, η

$$\mathcal{R}_{x_i, x_j} : X \times Y \rightarrow [0, 1] \quad (7.36)$$

καταδεικνύει μια ασαφή οντολογική σχέση μεταξύ δύο οντοτήτων  $x_i, y_j$  και

$$\mathcal{R}_{x_i, y_j} = \{Sim, Acc, P, Comp, Sp, Pr, adj, bel, ab, left, rgt, co\} \quad (7.37)$$

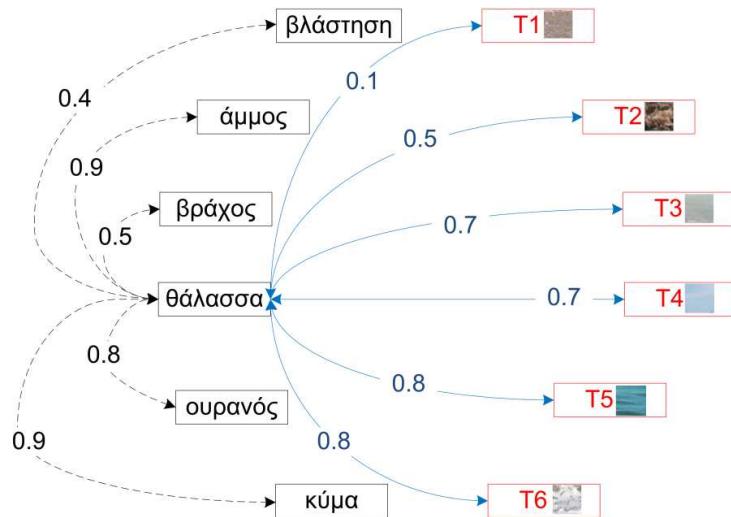
Ένας πιθανός συνδυασμός σχέσεων

$$\mathcal{C}_M = (\bigcup_{i,j} \mathcal{R}_{x_i, y_j}^{p_{ij}}), \quad p_{ij} \in \{-1, 0, 1\}, \quad i, j = 1 \dots m, \quad i \neq j \quad (7.38)$$

μπορεί να χρησιμοποιηθεί, για να διαμορφώσει έναν μοναδικό RDF γράφο [316], ο οποίος συνιστά το αφηρημένο μοντέλο του εννοιολογικού πλαισίου που εισάγεται στην παρούσα ενότητα και το οποίο θα αξιοποιηθεί κατά τη φάση της ανάλυσης. Η τιμή του  $p_{ij}$  καθορίζεται από τη σημασιολογία της εκάστοτε σχέσης  $\mathcal{R}_{x_i, y_j}$  που χρησιμοποιείται στην κατασκευή της  $\mathcal{C}_M$ . Ειδικότερα:

- $p_{ij} = 1$ , αν η σημασιολογία της  $\mathcal{R}_{x_i, y_j}$  υποδηλώνει ότι θα πρέπει να ληφθεί υπόψη, όπως είναι
- $p_{ij} = -1$ , αν η σημασιολογία της  $\mathcal{R}_{x_i, y_j}$  υποδηλώνει ότι θα πρέπει να ληφθεί υπόψη η αντίστροφή της
- $p_{ij} = 0$ , αν η σημασιολογία της  $\mathcal{R}_{x_i, y_j}$  δεν επιτρέπει τη συμμετοχή της στην κατασκευή της συνδυασμένης σχέσης  $\mathcal{C}_M$ .

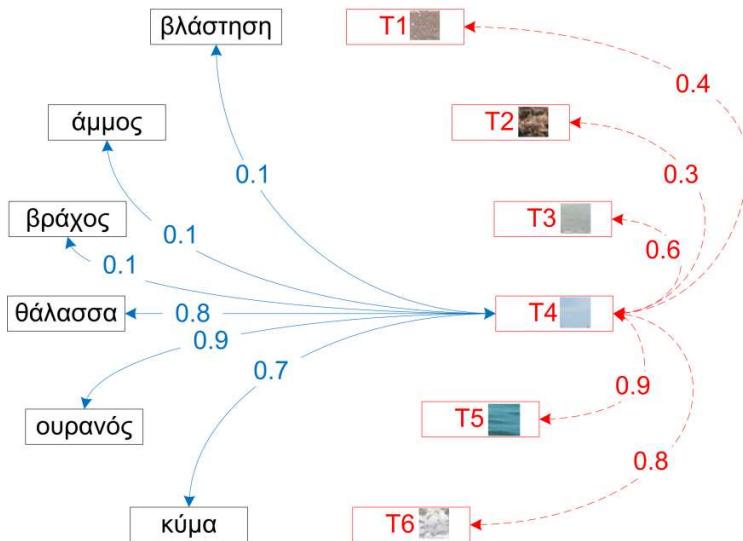
Στα Σχήματα 7.18 και 7.19 παρουσιάζουμε την απεικόνιση δύο τμημάτων του προαναφερθέντος γράφου RDF. Στο πρώτο εξ αυτών περιγράφονται μόνο οι σχέσεις



**Σχήμα 7.18:** Ένα τμήμα της οντολογίας του επεκτεταμένου οπτικού εννοιολογικού πλαισίου. Εμφανίζονται μόνο οι σχέσεις μεταξύ της υψηλού επιπέδου έννοιας θάλασσα και όλων των υπόλοιπων εννοιών και τύπων περιοχής. Οι αριθμοί καταδεικνύουν τον ασαφή βαθμό εμπιστοσύνης για κάθε σχέση.

μεταξύ της έννοιας υψηλού επιπέδου θάλασσα και όλων των υπόλοιπων εννοιών και τύπων περιοχής, ενώ στο δεύτερο εμφανίζονται μόνο οι σχέσεις μεταξύ του τέταρτου τύπου περιοχής (T4) και όλων των υπόλοιπων τύπων περιοχής και υψηλού επιπέδου εννοιών. Επίσης, θα πρέπει να αναφερθεί ότι οι παρακάτω γράφοι αντιπροσωπεύουν μόνο μικρά τμήματα ολόκληρης της εμπλουτισμένης οντολογίας του οπτικού εννοιολογικού πλαισίου. Αυτό συμβαίνει, γιατί μια οντολογία με 1 θεματική περιοχή, 10 έννοιες και 25 τύπους περιοχής (δηλ. 36 οντότητες στο σύνολο) απαιτεί ένα μέγιστο 630 σχέσεων και είναι ιδιαίτερα δύσκολο να απεικονιστεί με σαφήνεια. Ακόμα και αν λάβουμε υπόψη το γεγονός ότι οι εκάστοτε σημασιολογικές σχέσεις δεν είναι όλες εφαρμόσιμες παντού, όπως εξηγήθηκε παραπάνω, ένας τέτοιος περίπλοκος γράφος είναι δύσκολο να παρουσιαστεί σε ένα ευανάγνωστο Σχήμα. Ο γράφος του προτεινόμενου μοντέλου περιέχει κόμβους (δηλ. υψηλού επιπέδου έννοιες και τύπους περιοχής) και ακμές (δηλ. βασισμένες στο εννοιολογικό πλαίσιο ασαφείς σχέσεις μεταξύ των υψηλού επιπέδου εννοιών ή/και των τύπων περιοχής). Ο βαθμός εμπιστοσύνης κάθε ακμής αντιπροσωπεύει την ασάφεια στο μοντέλο. Προφανώς και εδώ οι ανύπαρκτες ακμές υπονοούν και ανύπαρκτες σχέσεις, ενώ όλες οι χρησιμοποιούμενες σχέσεις του γράφου είναι στην πραγματικότητα μεταβατικές.

Δεδομένου ότι συνήθως για ένα ζευγάρι οντοτήτων εφαρμόζονται περισσότερες της μίας ασαφείς σχέσεις, είναι δύσκολο να τις απεικονίσουμε όλες ταυτόχρονα σε ένα γράφο. Τα Σχήματα 7.18 και 7.19 παρουσιάζουν τμήματα της οντολογίας εννοιολογικού πλαισίου της θεματικής περιοχής παραλία, όπου κάθε σχέση μεταξύ δύο οντοτήτων μπορεί να είναι είτε μοναδική, είτε ένας συνδυασμός περισσότερων από μιας σχέσεων, όπως απεικονίζεται στην εξίσωση (7.38). Για να διευκολύνουμε την απεικόνιση της οντολογίας, παρουσιάζουμε στον Πίνακα 7.5 ένα μέρος από όλες τις πιθανές σημασιολογικές σχέσεις μεταξύ μιας υψηλού επιπέδου έννοιας  $C_{1/\text{ουρανός}}$  και όλων των άλλων οντοτήτων της οντολογίας. Οι μηδενικές τιμές καταδεικνύουν την απουσία της αντίστοιχης σχέσης. Επιπλέον, στον Πίνακα 7.6 παρουσιάζουμε τις αντίστοιχες ασαφείς τιμές για κάθε ζευγάρι οντοτήτων για την ασαφή σημασιολογική σχέση Συνοδός και το τοπολογικό ανάλογό της, τη σχέση Συνύπαρξη, οι οποίες



**Σχήμα 7.19:** Ένα άλλο τιμήμα της οντολογίας του επεκτεταμένου οπτικού εννοιολογικού πλαισίου. Εμφανίζονται μόνο οι σχέσεις μεταξύ του τέταρτου τύπου περιοχής ( $T_4$ ) και όλων των υπόλοιπων εννοιών και τύπων περιοχής. Οι αριθμοί καταδεικνύουν τον ασαφή βαθμό εμπιστοσύνης για κάθε σχέση.

εφαρμόζονται μεταξύ οποιονδήποτε δύο δεδομένων οντοτήτων αυτές ορίζονται.

**Πίνακας 7.5:** Ασαφείς σημασιολογικές σχέσεις μεταξύ της υψηλού επιπέδου έννοιας  $C_1/\text{ουρανός}$  και κάποιων ενδεικτικών υπόλοιπων οντοτήτων. Οι αριθμοί καταδεικνύουν τον ασαφή βαθμό εμπιστοσύνης για κάθε σχέση.

	$C_1$	$C_2$	...	$C_6$	$T_1$	$T_2$	...	$T_{25}$
$Sim$	0	0	...	0	0	0	...	0
$Acc$	1	0.5	...	0.9	0.7	0.8	...	0
$P$	1	0	...	0.3	0.7	0	...	0
$Comp$	1	0.2	...	0.9	0	0.5	...	0
$Sp$	0	0.8	...	0	0	0	...	0
$Pr$	0	0	...	0	0.5	0	...	0.7

Η περιγραφή του συνοδευτικού βαθμού εμπιστοσύνης κάθε ακμής μπορεί να πραγματοποιηθεί χρησιμοποιώντας ποικίλες μεθόδους, όμως στο πλαίσιο της παρούσας διατριβής προτείνουμε και επιλέξαμε να χρησιμοποιήσουμε τη μεθοδολογία του reification, όπως αυτή παρουσιάστηκε στην ενότητα 3.4.3. Η τεχνική χρησιμοποιείται ευρέως στο πεδίο της αναπαράστασης γνώσης, για να αντιπροσωπεύσει τα γεγονότα που πρέπει έπειτα να επεξεργαστούν με κάποιο τρόπο. Παραδείγματος χάριν, η τεχνική χρησιμοποιείται, για να συγχρίνει λογικούς ισχυρισμούς από διαφορετικούς μάρτυρες, έτσι ώστε να καθορίσει την αξιοπιστία τους. Π.χ. το μήνυμα “ο Ben είναι ο αρχηγός της ομάδας” είναι ένας ισχυρισμός αλήθειας που δεσμεύει τον αποστολέα στο γεγονός, ενώ η δήλωση “η Juliet ισχυρίζεται ότι ο Ben είναι ο αρχηγός της ομάδας” προσδίδει αυτή την δέσμευση στην Juliet. Κατ’ αυτό τον τρόπο οι δηλώσεις μπορούν να περιλάβουν ασαφείς πληροφορίες (δηλ. “ο Ben είναι ο αρχηγός της ομάδας με έναν βαθμό εμπιστοσύνης ίσο με 0.85”), χωρίς αυτό να δημιουργεί αντιφάσεις στην εξαγωγή συμπερασμάτων, δεδομένου ότι μια νέα ουσιαστικά δήλωση γίνεται για την αρχική δήλωση, η οποία νέα δήλωση είναι αυτή που περιέχει την πληροφορία του βαθμού. Φυσικά, η δήλωση αυτή δε βεβαιώνεται αυτόματα, ένα γεγονός που καθιστά τη χρήση της ανωτέρω τεχνικής για το επεκτεταμένο μοντέλο του εννοιολογικού πλαισίου πλήρως αποδεκτή. Παραδείγματος χάριν, έχοντας μια RDF τριπλέτα όπως η

**Πίνακας 7.6:** Οι βαθμοί της σημασιολογικής σχέσης Συνοδός και της τοπολογικής σχέσης Συνύπαρξη για κάποια ενδεικτικά ζεύγη οντοτήτων που αυτές ορίζονται. Οι αριθμοί καταδεικνύουν τον ασαφή βαθμό εμπιστοσύνης για κάθε σχέση.

	$C_1$	$C_2$	...	$C_6$	$T_1$	$T_2$	...	$T_{25}$
$C_1$	1	0.7	...	0	0.7	0.2	...	0.4
$C_2$	0.7	1	...	0.8	0.6	0.7	...	0.5
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$
$C_6$	0	0.8	...	1	0.6	0.7	...	0.8
$T_1$	0.7	0.2	...	0.4	1	0.3	...	0.5
$T_2$	0.6	0.7	...	0.5	0.3	1	...	0.1
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$
$T_{25}$	0.6	0.7	...	0.8	0.5	0.1	...	1

επόμενη: “μπλε Μέρος Του πράσινο” και έναν βαθμό εμπιστοσύνης ίσο με “0.85” για αυτήν τη δήλωση, προφανώς δε συνεπάγεται ότι ένας μπλε τύπος περιοχής θα είναι πάντα μέρος ενός πράσινου τύπου περιοχής σε μια σκηνή.

#### 7.4.3.1 Μεικτή επίδραση εννοιολογικού πλαισίου

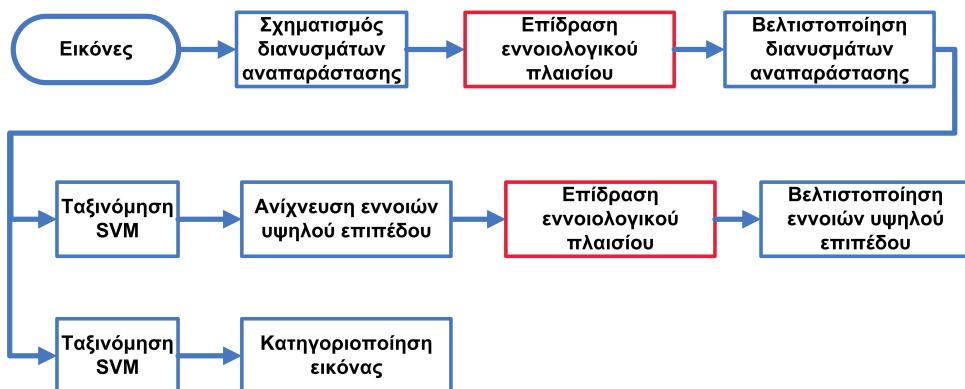
Με βάση τις αρχές και τα μαθηματικά θεμέλια της ασαφούς άλγεβρας [177] και της νέας προτεινόμενης αντίληψης για τη γνώση, παρουσιάζουμε στην υποενότητα αυτή ένα επιπρόσθετο βήμα και έναν αλγόριθμο βελτιστοποίησης εννοιολογικού πλαισίου ταξινόμησης ενδιάμεσου επιπέδου. Μια αφηρημένη εκδοχή της μεθοδολογίας οπτικοποιείται στο Σχήμα 7.21, ακολουθώντας μια τυπικότερη προσέγγιση, η προτεινόμενη διαδικασία επίδρασης του εννοιολογικού πλαισίου περιλαμβάνει συνοπτικά τα ακόλουθα βήματα:

1. Θεωρούμε ένα σύνολο  $T$  από αρχικούς τύπους περιοχής και ένα σύνολο  $C$  από αρχικές έννοιες.
2. Εισάγουμε το αρχικό σύνολο των τύπων περιοχής σε κατάλληλα εκπαιδευμένους ταξινομητές ως είσοδο και παράγουμε ένα σύνολο από εννοιολογικά ενημερωμένες έννοιες  $C'$ .
3. Εφαρμόζουμε το εννοιολογικό πλαίσιο πάνω στο αρχικό σύνολο  $T$ , προκειμένου να λάβουμε ένα σύνολο από εννοιολογικά ενημερωμένους τύπους περιοχής  $T'$ .
4. Χρησιμοποιούμε το ενημερωμένο σύνολο  $T'$  για να παράγουμε ένα καινούριο ενημερωμένο σύνολο τύπων περιοχής και επιστρέφουμε στο βήμα 2.
5. Εφαρμόζουμε το εννοιολογικό πλαίσιο πάνω στο σύνολο  $C'$  για να λάβουμε ένα τελικό σύνολο από εννοιολογικά ενημερωμένες έννοιες  $C''$ .

Θα πρέπει να σημειώσουμε εδώ ότι κατά τη διάρκεια των παραπάνω βημάτων το σύνολο  $C'$  θα αξιοποιηθεί μόνο κατά την τελευταία φάση, οπότε και επιλέγονται οι τελικές τιμές των εννοιών υψηλού επιπέδου  $C''$  και συνεπώς μπορεί να αγνοηθεί κατά τις αρχικές επαναλήψεις. Με άλλα λόγια, το μεικτό εννοιολογικό πλαίσιο λαμβάνει ως είσοδο τους αρχικούς τύπους περιοχής  $T$  και ένα σύνολο αρχικών τιμών για τις υψηλού επιπέδου έννοιες  $C$ . Με βάση τις εκάστοτε κατάλληλες σημασιολογικές, τοπολογικές και χωρικές σχέσεις, που ορίστηκαν στην προηγούμενη υποενότητα, είναι σε θέση να παράγει δύο διακριτά σύνολα οντοτήτων: ένα σύνολο από εννοιολογικά

ενημερωμένες τιμές εννοιών  $C'$  και ένα σύνολο από εννοιολογικά ενημερωμένους τύπους περιοχής  $T'$ . Το σύνολο  $C'$ , λοιπόν, θα χρησιμοποιηθεί μόνο κατά το τελευταίο στάδιο της επιλογής των τελικών τιμών για τις υψηλού επιπέδου έννοιες και προς το παρόν αγνοείται. Το σύνολο  $T'$  λαμβάνεται από τον ταξινομητή SVM και παράγει ένα άλλο σύνολο υψηλού επιπέδου εννοιών  $C''$ . Καθώς, όπως θα δούμε, ο αλγόριθμος επίδρασης του εννοιολογικού πλαισίου ταξινόμησης είναι αναδρομική φύσης, η παραπάνω διαδικασία επαναλαμβάνεται μέχρι να αποκομίσουμε το βέλτιστο σύνολο υψηλού επιπέδου εννοιών  $C'$  και τύπων περιοχής  $T'$ . Στο επίπεδο της προσέγγισης μας η εξαγωγή ενός ενημερωμένου συνόλου  $C'$  ολοκληρώνει την επίδραση του εννοιολογικού πλαισίου, ενώ το ενημερωμένο σύνολο  $T'$  μπορεί να χρησιμοποιηθεί από έναν άλλο ταξινομητή SVM, προκειμένου να ταξινομηθεί η υπό εξέταση εικόνα σε κάποια γενική κατηγορία.

Η παραπάνω διαδικασία αποτελεί αναπόσπαστο κομμάτι της ανάλυσης, μιας και διαμορφώνει απευθείας και αναδρομικά το κατασκευαζόμενο διάνυσμα αναπαράστασης. Παρέχει, δε, τη βέλτιστη σε σχέση με την απλή περίπτωση επανεκτίμηση των βαθμών εμπιστοσύνης των τελικών ανιχνευμένων εννοιών υψηλού επιπέδου, αλλά και των επιλεγμένων τύπων περιοχής. Με άλλα λόγια, η προτεινόμενη τεχνική ενημερώνει τις τιμές κάθε διανύσματος αναπαράστασης, επιτρέποντας μια βελτιστοποιημένη διαδικασία εκπαίδευσης του ταξινομητή που παράγει τελικά τα σύνολα των υποψήφιων υψηλού επιπέδου εννοιών, αλλά και του ταξινομητή που θα ταξινομήσει τελικά την εικόνα σε γενική κατηγορία (π.χ. ίππαιθρο/εσωτερικό, παραλία/βουνό), επιτυγχάνοντας έτσι τα βέλτιστα αποτελέσματα κατηγοριοποίησης σε σχέση με όλες τις προηγούμενες εκδοχές. Τελικά, το πρόβλημα που επιχειρούμε να εξετάσουμε συνοψίζεται ως εξής: προσπαθούμε να ενημερώσουμε κατά τρόπο σημαντικό τις αρχικές τιμές εμπιστοσύνης των εννοιών και των τύπων περιοχής που παράγονται από ένα αρχικό βήμα χαμηλού επιπέδου πολυμεσικής ανάλυσης. Το σχετικό διάγραμμα ροής της διαδικασίας παρουσιάζεται στο Σχήμα 7.20.



Σχήμα 7.20: Διάγραμμα ροής της επίδρασης του μεικτού εννοιολογικού πλαισίου ταξινόμησης

Μια εκτίμηση του βαθμού ιδιότητας μέλους κάθε ενδιάμεσου επιπέδου οντότητας προέρχεται από τις άμεσες και τις έμμεσες σχέσεις των τελευταίων με άλλες οντότητες στον κατασκευασμένο μεικτό γράφο, χρησιμοποιώντας έναν κατάλληλο δείκτη συμβατότητας ή μια κατάλληλη μετρική απόστασης. Ανάλογα με τη φύση των θεματικών περιοχών που περιέχονται στην οντολογία ο καλύτερος δείκτης θα μπορούσε να επιλεχθεί χρησιμοποιώντας τον τελεστή  $max$  ή  $min$ , αντίστοιχα. Φυσικά, η ιδιαίτερη μετρική απόστασης για δύο έννοιες ή τύπους περιοχής είναι και πάλι μια που

ποσοτικοποιεί το σημασιολογικό συσχετισμό τους. Για το συγκεκριμένο πρόβλημα επιλέγεται ο τελεστής  $\max$  ως ένα σημαντικό μέτρο της σημασιολογικής συσχέτισης και για τους δύο.

Η σύνθεση του προτεινόμενου αλγορίθμου επαναξιολόγησης των βαθμών ιδιότητας μέλους πραγματοποιείται προσαρμόζοντας και συνδυάζοντας κατάλληλα τις βασικές αρχές που διέπουν τους αντίστοιχους αλγόριθμους της απλής εννοιολογικής και οπτικής περίπτωσης. Εισάγονται, βέβαια, κάποια επιπλέον ιδιαίτερα συμπληρωματικά χαρακτηριστικά στη δομή του αλγορίθμου, τονίζοντας κυρίως την εκμετάλλευση των χαρακτηριστικών γνωρισμάτων της ασαφούς άλγεβρας τόσο για τις έννοιες υψηλού επιπέδου όσο και για τους τύπους περιοχής. Η γενική δομή του προτεινόμενου αλγορίθμου, για την οποία επιλέγουμε να χρησιμοποιήσουμε την τυποποιημένη (standard)  $t$ -co-νόρμα και το αλγεβρικό γινόμενο ως  $t$ -νόρμα, κατάλληλα προσαρμοσμένη για την αξιοποίηση του μεικτού εννοιολογικού πλαισίου, είναι η ακόλουθη:

1. Προσδιορίστε το μέτρο ομοιότητας (ή ανομοιότητας) περιοχών, που επιβάλλεται από τη φύση της εξεταζόμενης θεματικής περιοχής:  $\mu \in [0, 1]$ .
2. Για κάθε έννοια  $c \in C$  περιγράφουμε το ασαφές σύνολο  $L_c$ , χρησιμοποιώντας τον ευρέως διαδεδομένο συμβολισμό αθροίσματος [177]:  $L_c = \sum_{i=1}^{|C|} c_i / w_i = \{c_1/w_1, c_2/w_2, \dots, c_n/w_n\}$ , όπου το  $w_i$  περιγράφει τη συνάρτηση συμμετοχής:  $w_i = \mu_{L_c}(c_i)$
3. Για κάθε τύπο περιοχής  $t \in T$  περιγράφουμε το ασαφές σύνολο  $L_t$ :  $L_t = \sum_{i=1}^{|T|} t_i / z_i = \{t_1/z_1, t_2/z_2, \dots, t_n/z_n\}$ , όπου το  $z_i$  περιγράφει τη συνάρτηση συμμετοχής:  $z_i = \mu_{L_t}(t_i)$
4. Για κάθε έννοια  $c_i$  στο ασαφές σύνολο  $L_c$  με βαθμό συμμετοχής  $w_i$  λαμβάνουμε τη συγκεκριμένη πληροφορία εννοιολογικού πλαισίου με τη μορφή των σχέσεων του με το σύνολο όλων των υπόλοιπων οντοτήτων (εννοιών ή τύπων περιοχής):  $\{R_{c_i, x_j} : c_i \in C, x_j \in C \cup T, i \neq j\}$
5. Για κάθε τύπο περιοχής  $t_i$  στο ασαφές σύνολο  $L_t$  με βαθμό συμμετοχής  $z_i$  λαμβάνουμε τη συγκεκριμένη πληροφορία εννοιολογικού πλαισίου με τη μορφή των σχέσεων του με το σύνολο όλων των υπόλοιπων οντοτήτων:  $\{R_{t_i, x_j} : t_i \in T, x_j \in C \cup T, i \neq j\}$ .
6. Υπολογίζουμε τους νέους βαθμούς συμμετοχής  $w_i$  και  $z_i$ , λαμβάνοντας υπόψη τη μετρική ομοιότητας κάθε θεματικής περιοχής. Στην περίπτωση πολλαπλών σχέσεων, οι οποίες σχετίζουν την έννοια  $c_i$  ή τον τύπο περιοχής  $t_i$  και με άλλες οντότητες εκτός της έννοιας ρίζας, πρέπει να εφαρμοστεί ένα ενδιάμεσο βήμα συνάθροισης για την εκτίμηση των  $w_i$  και  $z_i$ , λαμβάνοντας υπόψη τη σχετικότητα πλαισίου που συζητήθηκε στην ενότητα 6.3.1:  $cr_{c_i}$  ή  $cr_{t_i}$ .

Εκφράζουμε τον υπολογισμό των  $w_i$  και  $z_i$  και για τις δύο περιπτώσεις με τον αναδρομικό τύπο:

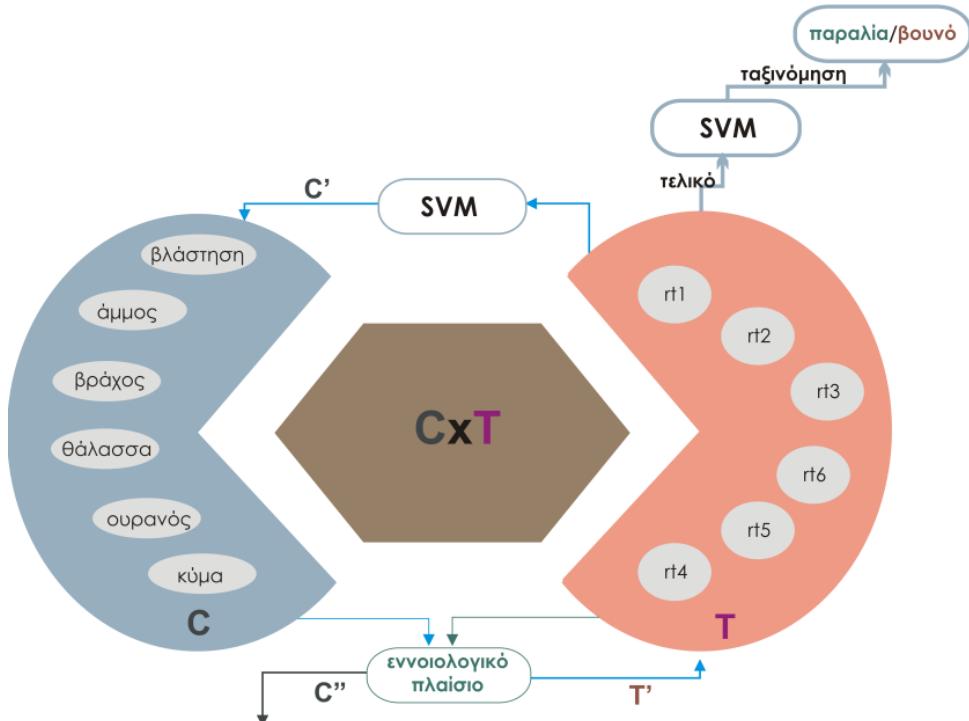
$$q_i^n = q_i^{n-1} - \mu(q_i^{n-1} - cr_{x_i}), \quad q_i = \{w_i, z_i\} \quad (7.39)$$

όπου το  $n$  καταδεικνύει την επανάληψη που χρησιμοποιείται και το  $x_i$  συμβολίζει είτε μία έννοια  $c_i$  ή ένα τύπο περιοχής  $t_i$ . Ισοδύναμα, για μία αυθαίρετη

επανάληψη  $n$ :

$$q_i^n = (1 - \mu)^n \cdot q_i^0 + (1 - (1 - \mu)^n) \cdot cr_{x_i} \quad (7.40)$$

όπου το  $q_i^0$  αναπαριστά τον αρχικό βαθμό συμμετοχής για την οντότητα  $x_i, x_i \in C \cup T$ .



Σχήμα 7.21: Μεθοδολογία επίδρασης μεικτού εννοιολογικού πλαισίου ταξινόμησης

## 7.5 Πειραματικά Αποτελέσματα

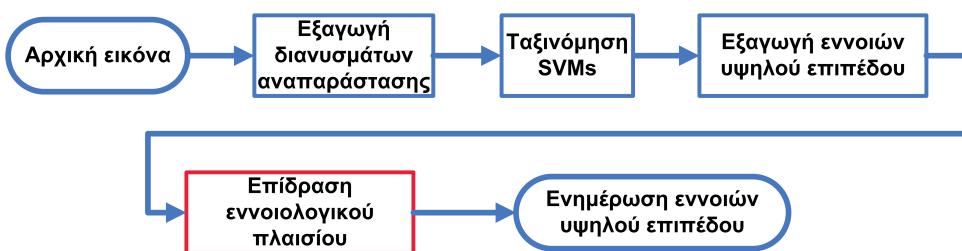
Στη συνέχεια, παρατίθενται μια ενδεικτική επιλογή πειραματικών μετρήσεων από τις τρεις αυτοτελείς περιοχές επίδρασης του εννοιολογικού πλαισίου ταξινόμησης ενδιαμεσου επιπέδου, καθώς και μια σειρά αποτελεσμάτων από την τελική ταξινόμηση εικόνων σε δύο κατηγορίες, με βάση τη θεώρηση που ενδιάμεσου επιπέδου που αναπτύχθηκε στο κεφάλαιο 7. Ειδικότερα, στην ενότητα αυτή περιλαμβάνονται αποτελέσματα από την εφαρμογή της προτεινόμενης μεθοδολογίας αξιοποίησης του εννοιολογικού πλαισίου από τη σκοπιά της καθαρά εννοιολογικής προσέγγισης (βλ. ενότητα 7.4.1), των τύπων περιοχής (βλ. ενότητα 7.4.2) και τη σκοπιά της κοινής θεώρησής τους (βλ. ενότητα 7.4.3). Αξιολογούμε επίσης την τελευταία προσέγγιση παραθέτοντας συγκριτικά αποτελέσματα από την εφαρμογή της πάνω στα ευρέως γνωστά σύνολα δεδομένων Corel και TRECVID, μαζί με μια σειρά από σύγχρονες τεχνικές ανίχνευσης εννοιών. Τέλος, τα αποτελέσματα από τη συνολική ταξινόμηση εικόνων σε κατηγορίες παρατίθενται στην ενότητα 7.5.4.

Θα πρέπει να σημειώσουμε εδώ ότι η δημιουργία των απαραίτητων οπτικών θησαυρών για το σύνολο των πειραμάτων μας έγινε, όπως αυτή περιγράφεται στην υποενότητα 7.3.2 και με βάση την κλασική μορφή του προτεινόμενου αλγόριθμου συσταδοποίησης. Καθώς όμως το σύνολο εκπαίδευσης ήταν πολύ μεγάλο, αποτελούμενο από εκατοντάδες χιλιάδες περιοχές, επιλέχθηκε ένα υποσύνολο αυτών των περιοχών, για

το οποίο η μεθοδολογία είναι ικανή να εκτελεστεί στο πλαίσιο ικανοποιητικών χρονικών ορίων και με επάρκεια μνήμης. Το υποσύνολο αυτό δημιουργήθηκε επιλέγοντας όλες τις περιοχές, οι οποίες ανήκαν σε εικόνες που περιείχαν τις σημασιολογικές έννοιες οι οποίες είχαν επιλεχθεί να εντοπιστούν, προσθέτοντας επίσης κι έναν μεγάλο αριθμό άλλων περιοχών που επιλέχθηκαν τυχαία.

### 7.5.1 Εννοιολογική προσέγγιση

Σε αυτή την ενότητα παρουσιάζουμε πειραματικά αποτελέσματα, που αναφέρονται στο πρώτο μέρος της προτεινόμενης προσέγγισης μεσαίου επιπέδου ανάλυσης, δηλ. σε αυτή που ασχολείται απευθείας με τις υψηλού επιπέδου έννοιες. Θα καταδείξουμε τη χρησιμότητα του αλγορίθμου επίδρασης του εννοιολογικού πλαισίου ταξινόμησης, όταν αυτός επιδρά πάνω σε πραγματικά πολυμεσικά προβλήματα και θα αποδείξουμε ότι η γνώση του εννοιολογικού πλαισίου με τη μορφή μιας οντολογίας εννοιολογικού πλαισίου είναι ικανή να συμπληρώσει αποτελεσματικά τις παραδοσιακές τεχνικές αναγνώρισης υψηλού επιπέδου εννοιών, να προάγει τις διαδικασίες τους και να εκλεπτύνει τα αποτελέσματά τους. Στο Σχήμα 7.22 απεικονίζεται ένα διάγραμμα ροής, το οποίο περιγράφει την αλληλεπίδραση μεταξύ του αλγορίθμου επίδρασης του εννοιολογικού πλαισίου και της διαδικασίας αναγνώρισης υψηλού επιπέδου εννοιών.



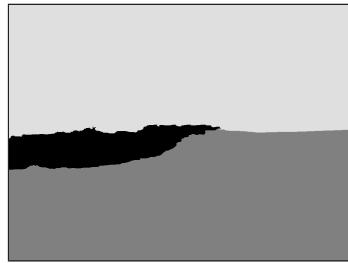
**Σχήμα 7.22:** Ένα απλό διάγραμμα ροής για ένα σενάριο χρήσης εννοιολογικού πλαισίου εννοιών.

Ειδικότερα, πραγματοποιήσαμε μια σειρά πειραμάτων χρησιμοποιώντας ένα σύνολο από 1148 εικόνες, 25 τύπους περιοχής και 6 έννοιες υψηλού επιπέδου {θάλασσα, κύμα, ουρανός, άμμος, βράχος, βλάστηση} από τη θεματική περιοχή παραλία. Οι εικόνες συλλέχθηκαν από προσωπικές συλλογές φωτογραφιών και το Διαδίκτυο. Κατασκευάστηκε, επίσης, χειρωνακτικά ένα σύνολο απόλυτης αλήθειας, αποτελούμενο από διάφορους τύπους περιοχής που συνδέθηκαν με μια μοναδική έννοια. Χρησιμοποιήσαμε 228 εικόνες (περίπου το 20% του συνόλου δεδομένων) ως το σύνολο εκπαίδευσης για τη συσταδοποίηση και μετά από μια εκτενή διαδικασία δοκιμών επιλέξαμε  $\mu = 0.12$  ως τη βέλτιστη παράμετρο χανονικοποίησης για τη δεδομένη θεματική περιοχή. Ένα ενδεικτικό παράδειγμα περίπτωσης χρήσης εικόνας από το παραπάνω σύνολο δεδομένων παρατίθεται στο Σχήμα 7.23, όπου: (α) είναι η αρχική εικόνα εισόδου και (β) η κατατμημένη εικόνα μετά την εφαρμογή της αρχικής χονδροειδούς κατάτμησης. Στην πρώτη αυτή περίπτωση θεωρούμε έναν απλό θησαυρό περιοχών, ο οποίος αποτελείται από 4 τύπους περιοχής. Προκειμένου να ληφθεί η σχετική κατάτμηση, υλοποιήσαμε τον αλγόριθμο RSST [11] και θέσαμε ένα κατάλληλο κατώφλι, προκειμένου ο αλγόριθμος να τερματίσει σε έναν ικανοποιητικό αριθμό από περιοχές.

Το πρόβλημα που καλούμαστε να αντιμετωπίσουμε σε αυτή την περίπτωση, βέβαια, είναι η καθολική αναγνώριση υψηλού επιπέδου εννοιών, δεδομένου ότι για τα περισσότερα προβλήματα αναγνώρισης υψηλού επιπέδου χαρακτηριστικών είναι διαθέσιμος



(a) Αρχική εικόνα



(b) Κατάτμηση

**Σχήμα 7.23:** Ενδεικτικό παράδειγμα εικόνας παραλίας.

μόνο ένας γενικός σχολιασμός για ολόκληρη την εκάστοτε εικόνα και όχι για κάθε περιοχή ξεχωριστά. Όπως είδαμε στην ενότητα 7.3, το πρόβλημα αυτό αντιμετωπίζεται με μια αναπαράσταση εικόνας που βασίζεται στον παραπάνω κατάλληλο θησαυρό περιοχών (δηλ. ένα σύνολο από τύπους περιοχής). Στο πλαίσιο των πειραματικών μας μετρήσεων επεκτείναμε αυτή την προσέγγιση εκμεταλλευόμενοι την οντολογία του εννοιολογικού πλαισίου και έτσι βελτιώσαμε τις σχετικές τιμές εμπιστοσύνης με έναν επαναληπτικό τρόπο, λαμβάνοντας υπόψη τις σχέσεις του εννοιολογικού πλαισίου μεταξύ των εννοιών που απαντώνται στην εικόνα.

Εν προκειμένω, για το δείγμα του Σχήματος 7.23 οι τιμές εμπιστοσύνης για τις υψηλού επιπέδου έννοιες θάλασσα, κύμα, ουρανός, άμμος, βράχος και βλάστηση, που παράγονται από τους εξειδικευμένους ταξινομητές SVM, παρουσιάζονται με τη μορφή διανύσματος στην εξίσωση (7.41).

$$\mathbf{C} = \{c_i\} = [ \begin{array}{ccccccc} 0.62 & 0.89 & 0.78 & 0.67 & 0.35 & 0.22 \end{array} ] \quad (7.41)$$

Με δεδομένο ότι μπορούμε να παρατηρήσουμε ότι η συγκεκριμένη εικόνα αποτελείται κατά βάση από περιοχές ουρανού και θάλασσας (βλ. Σχήμα 7.23(a) και 7.23(b)), πρέπει να αναμένουμε ότι οι υψηλού επιπέδου έννοιες που αντιστοιχούν σε αυτές τις σημασιολογικές έννοιες θα πρέπει να έχουν τις υψηλότερες τιμές. Στη συγκεκριμένη περίπτωση η θάλασσα έχει ένα αρκετά διαφορετικό χρώμα από τον τύπο περιοχής του θησαυρού που προέκυψε από μια περιοχή θάλασσα. Παρατηρούμε ότι το συγκεκριμένο χρώμα εμφανίζεται παρόμοιο ακόμα και με την περιοχή βράχος (βλ. Σχήμα 7.24). Κατά συνέπεια, αναμένουμε ότι η σχετική κατά SVM ταξινόμηση θα παρέχει έναν μάλλον μετριοπαθή βαθμό εμπιστοσύνης για την έννοια θάλασσα, πράγμα που επιβεβαιώνεται από την παραπάνω εξίσωση (7.41). Θα επιθυμούσαμε να αυξήσουμε αυτήν την εμπιστοσύνη και να μειώσουμε επίσης την εμπιστοσύνη που αντιστοιχεί στην έννοια βράχος (βλ. 1ο και 5ο στοιχείο της εξίσωσης). Αφού εφαρμόσουμε την κατάλληλη μορφή του αλγορίθμου του εννοιολογικού πλαισίου, οι τελικές τιμές των εννοιών γίνονται:

$$\mathbf{C}' = \{c'_i\} = [ \begin{array}{ccccccc} 0.83 & 0.91 & 0.85 & 0.70 & 0.23 & 0.19 \end{array} ] \quad (7.42)$$

Στη συνέχεια, προκειμένου να αξιολογήσουμε την προτεινόμενη τεχνική σε ολόκληρο το σύνολο δεδομένων μας, παρουσιάζουμε σκορ ακρίβειας και ανάκλησης από την εφαρμογή της σε ολόκληρο το σύνολο δεδομένων και ανά υψηλού επιπέδου έννοια (δηλ. το εννοιολογικό επιδρά, αφού εφαρμοστεί και το τελικό βήμα ταξινόμησης SVM κατά τη διαδικασία της ανάλυσης). Συνολικά αποτελέσματα αξιολόγησης για 6 υψηλού επιπέδου έννοιες από τη θεματική περιοχή παραλία παρουσιάζονται στον Πίνακα



(a) Αρχική εικόνα



(b) Κατάτμηση

**Σχήμα 7.24:** Ενδεικτικό παράδειγμα οπτικής ομοιότητας περιοχών.

**7.7.** Κάθε γραμμή παρουσιάζει την τιμή της ακρίβειας και της ανάκλησης πριν και μετά την επίδραση του εννοιολογικού πλαισίου.

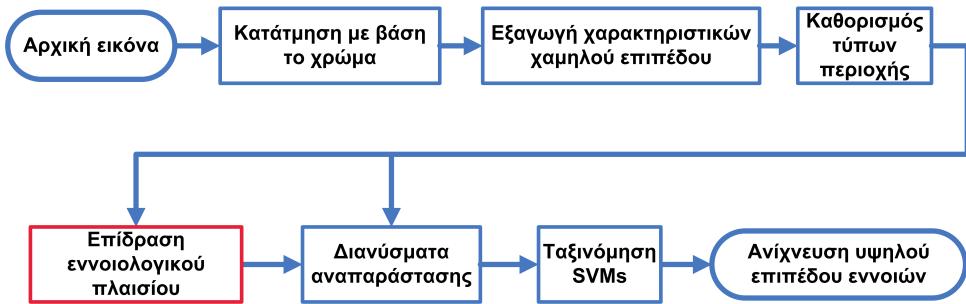
**Πίνακας 7.7:** Συνολικά σκορ ακρίβειας και ανάκλησης πριν και μετά την επίδραση του εννοιολογικού πλαισίου ανά υψηλού επιπέδου έννοια από τη θεματική περιοχή παραλία

Έννοιες	Ακρίβεια			Ανάκληση		
	πριν	μετά	%	πριν	μετά	%
θάλασσα	0.66	0.76	15.15%	0.76	0.72	-5.26 %
κύμα	0.59	0.69	16.95%	0.68	0.61	-10.29%
ουρανός	0.39	0.51	30.77%	0.55	0.48	-12.73%
άμμος	0.32	0.51	59.38%	0.45	0.39	-13.33%
βράχος	0.30	0.40	33.33%	0.38	0.33	-13.16%
βλάστηση	0.34	0.38	11.76%	0.39	0.34	-12.82%
<b>Σύνολο</b>	<b>0.43</b>	<b>0.54</b>	<b>25.58%</b>	<b>0.54</b>	<b>0.48</b>	<b>-11.11%</b>

### 7.5.2 Προσέγγιση τύπων περιοχής

Σε αυτή την ενότητα παρέχουμε πειραματικά αποτελέσματα που αναφέρονται στο δεύτερο μέρος της προτεινόμενης προσέγγισης, δηλ. σε αυτή που χρησιμοποιεί και ασχολείται αποκλειστικά με την αποδοτική διαχείριση των τύπων περιοχής. Για αυτό το δεύτερο σύνολο πειραμάτων χρησιμοποιήσαμε το ίδιο σύνολο δεδομένων, δηλαδή 1148 εικόνες, 6 έννοιες και 25 τύπους περιοχής από τη θεματική περιοχή παραλία, όπως και στην προηγούμενη περίπτωση. Στο Σχήμα 7.25 απεικονίζεται ένα διάγραμμα ροής, το οποίο περιγράφει την αλληλεπίδραση μεταξύ του αλγορίθμου επίδρασης του εννοιολογικού πλαισίου και της διαδικασίας της ανάλυσης. Χρησιμοποιήσαμε, επίσης, το ίδιο σύνολο εκπαίδευσης εικόνων για τη συσταδοποίηση και την ίδια παράμετρο κανονικοποίησης  $\mu = 0.12$ , μιας και η θεματική περιοχή εφαρμογής δε μεταβλήθηκε.

Προφανώς, και σε αυτή την περίπτωση το επιπρόσθετο πρόβλημα που καλούμαστε να αντιμετωπίσουμε και ο τρόπος, για να αξιολογήσουμε την προτεινόμενη μεθοδολογία εννοιολογικού πλαισίου, είναι η τελική ανίχνευση των υψηλού επιπέδου εννοιών και η εκτίμηση του κατά πόσο βελτιώθηκαν οι σχετικοί βαθμοί εμπιστοσύνης σε αυτές



**Σχήμα 7.25:** Ένα απλό διάγραμμα ροής για ένα σενάριο χρήσης εννοιολογικού πλαισίου τύπων περιοχής.

πριν και μετά τη χρήση της. Υλοποιώντας τη μεθοδολογία που περιγράφηκε στην ενότητα 7.3, κατασκευάζουμε το σχετικό θησαυρό περιοχών, ο οποίος αυτή τη φορά είναι πιο σύνθετος και περιλαμβάνει 6 τύπους περιοχής. Μια τυχαία αρχική εικόνα από το σύνολο των δεδομένων μας απεικονίζεται στο Σχήμα 7.26. Οι τιμές εμπιστοσύνης για όλους τους 6 τύπους περιοχής  $t_i$  αυτής της εικόνας απεικονίζονται με τη μορφή διανύσματος στην εξ. (7.43).

$$\mathbf{T} = \{t_i\} = [ 0.89 \ 0.62 \ 0.21 \ 0.68 \ 0.67 \ 0.31 ] \quad (7.43)$$



**Σχήμα 7.26:** Ενδεικτικό δείγμα αρχικής εικόνας.

Οι τιμές εμπιστοσύνης για τις υψηλού επιπέδου έννοιες θάλασσα, ουρανός, άμμος, κύμα, βλάστηση και βράχος, που παράγονται από τους εξειδικευμένους ταξινομητές SVM κατά το τελικό στάδιο της ανάλυσης χωρίς την αξιοποίηση του εννοιολογικού πλαισίου, παρουσιάζονται στην εξίσωση (7.44).

$$\mathbf{C} = \{c_i\} = [ 0.32 \ 0.91 \ 0.12 \ 0.87 \ 0.35 \ 0.09 ] \quad (7.44)$$

Όπως είναι απολύτως προφανές για έναν ανθρώπινο παρατηρητή, η αρχική εικόνα περιέχει τις υψηλού επιπέδου έννοιες θάλασσα, ουρανός και κύμα. Παρ' όλα αυτά, από αυτά που υπαγορεύει η εξίσωση (7.44) παρατηρούμε ότι για τη θάλασσα ο ταξινομητής SVM απέτυχε να παράγει μια ικανοποιητικά υψηλή τιμή εμπιστοσύνης. Μια πιθανή εξήγηση είναι ότι ο ταξινομητής SVM είχε εκπαιδευθεί, χρησιμοποιώντας διαφορετικά θετικά παραδείγματα για τη θάλασσα απ' ότι φαίνεται στο Σχήμα 7.26. Εντούτοις, μετά και από την αξιοποίηση του βήματος βελτιστοποίησης του εννοιολογικού πλαισίου το διάνυσμα που περιέχει τις τιμές εμπιστοσύνης για όλες τις υψηλού επιπέδου έννοιες μετατρέπεται σε αυτό της εξίσωσης (7.45).

$$\mathbf{C}' = \{c'_i\} = [ 0.62 \ 0.95 \ 0.18 \ 0.90 \ 0.29 \ 0.09 ] \quad (7.45)$$

Μέσω αυτού του απλοϊκού παραδείγματος είμαστε πλέον σε θέση να καταλάβουμε τη δύναμη του αλγορίθμου επίδρασης του εννοιολογικού πλαισίου και τη σημασία

## Κεφάλαιο 7. Εννοιολογικό πλαίσιο ταξινόμησης

του για το γενικό πρόβλημα της ανάλυσης πολυμεσικού περιεχομένου. Καθώς ο αλγορίθμος του εννοιολογικού πλαισίου είχε τις ακόλουθες πληροφορίες:

- επρόκειτο για μια εικόνα από τη θεματική περιοχή παραλία
- ο ουρανός είχε μια υψηλή τιμή εμπιστοσύνης
- το κύμα είχε μια υψηλή τιμή εμπιστοσύνης
- υπάρχει ένας “μπλε” τύπος περιοχής
- υπάρχει ένας “άσπρος” τύπος περιοχής
- ο ουρανός και το κύμα έχουν μια υψηλή συσχέτιση με τη θάλασσα

μπορούμε να παρατηρήσουμε ότι είχε ως αποτέλεσμα την αύξηση της τιμής εμπιστοσύνης της θάλασσας, ενώ για παράδειγμα οι τιμές της βλάστησης και της άμμου που ήταν χαμηλές παρέμειναν εξίσου χαμηλές. Επίσης, οι τιμές εμπιστοσύνης του ουρανού και του κύματος, οι οποίες ήταν εκ των προτέρων υψηλές, παρέμειναν πρακτικά αμετάβλητες.

Τέλος, συνολικά αποτελέσματα αξιολόγησης της προτεινόμενης τεχνικής για το σύνολο των 1148 εικόνων παρατίθενται στον Πίνακα 7.8, όπου παρουσιάζονται σκορ ακρίβειας και ανάκλησης τόσο πριν, όσο και μετά την επίδραση του εννοιολογικού πλαισίου (οι υψηλού επιπέδου έννοιες που ανιχνεύθηκαν στο σύνολο των δεδομένων παρέμειναν αμετάβλητες, ήτοι: θάλασσα, ουρανός, άμμος, κύμα, βλάστηση και βράχος).

**Πίνακας 7.8:** Συνολικά σκορ ακρίβειας και ανάκλησης για τις 1148 εικόνες και τις 6 υψηλού επιπέδου έννοιες θάλασσα, ουρανός, άμμος, κύμα, βλάστηση και βράχος.

Έννοιες	Ακρίβεια			Ανάκληση		
	πριν	μετά	%	πριν	μετά	%
θάλασσα	0.48	0.62	28.13%	0.60	0.55	-7.36 %
ουρανός	0.52	0.64	22.69%	0.65	0.57	-12.31%
άμμος	0.39	0.47	19.54%	0.48	0.40	-16.67%
κύμα	0.39	0.41	4.87 %	0.43	0.37	-13.95%
βλάστηση	0.24	0.32	35.17%	0.29	0.25	-13.79%
βράχος	0.25	0.35	38.25%	0.34	0.29	-14.71%
<b>Σύνολο</b>	<b>0.38</b>	<b>0.46</b>	<b>23.68%</b>	<b>0.47</b>	<b>0.41</b>	<b>-10.87%</b>

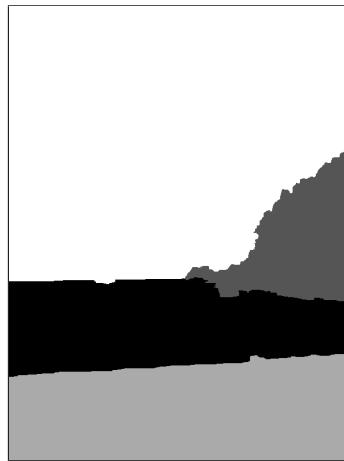
### 7.5.3 Μεικτή προσέγγιση τύπων περιοχής και εννοιών

Τελος, στην ενότητα αυτή παρουσιάζουμε ένα σύνολο πειραματικών αποτελεσμάτων από την εφαρμογή της τελευταίας εκδοχής της προτεινόμενης μεθοδολογίας, δηλ. της εκδοχής με την ταυτόχρονη αξιοποίηση τόσο των τύπων περιοχής, όσο και των εννοιών, καθώς και των σχέσεων που τους/τις διέπουν. Αρχικά πραγματοποιήσαμε μετρήσεις με το ίδιο σύνολο δεδομένων αποτελούμενο από τις 1148 εικόνες και τις 6 έννοιες, χρησιμοποιώντας την ίδια χειρωνακτικά κατασκευασμένη απόλυτη αλήθεια. Κάποια ενδεικτικά παραδείγματα εικόνων από το σύνολο δεδομένων παρουσιάζονται στα Σχήματα 7.27, 7.28 και 7.29, όπου καταδεικνύονται: (i) η αρχική εικόνα και (ii) το αποτέλεσμα της κατάτμησης και εξετάζεται η αποδοτικότητα του αλγορίθμου, αλλά

και συνολικά σκορ ακρίβειας και ανάκλησης για το σύνολο των δεδομένων (δηλ. τις 6 έννοιες και τις 1148 εικόνες).



(a) Αρχική εικόνα



(b) Κατάτμηση

**Σχήμα 7.27:** 1<sup>η</sup> εικόνα παραλίας.



(a) Αρχική εικόνα



(b) Κατάτμηση

**Σχήμα 7.28:** 2<sup>η</sup> εικόνα παραλίας.



(a) Αρχική εικόνα



(b) Κατάτμηση

**Σχήμα 7.29:** 3<sup>η</sup> εικόνα παραλίας.

Τα τελικά ζεύγη εννοιών/βαθμών εμπιστοσύνης για τις ανιχνευμένες σημασιολογικές έννοιες πριν και μετά την επίδραση του εννοιολογικού πλαισίου για τα παραπάνω τρία ενδεικτικά παραδείγματα συνοψίζονται στον Πίνακα 7.9. Οι έννοιες που δεν ανιχνεύθηκαν στο πλαίσιο των συγκεκριμένων εικόνων εμφανίζονται να έχουν μηδενικές τιμές.

Το πρόβλημα που εξετάζουμε στην περίπτωση αυτή είναι η ανίχνευση υψηλού επιπέδου εννοιών, αξιοποιώντας τη σύνθετη, μεικτή οντολογία του εννοιολογικού πλαισίου που προτείναμε στην ενότητα 7.4.3. Επομένως, στοχεύοντας να βελτιώσουμε τις

**Πίνακας 7.9:** Ζεύγη βαθμών εμπιστοσύνης ανά εικόνα παραλίας.

Έννοιες	Εικόνα 1		Εικόνα 2		Εικόνα 3	
	πριν	μετά	πριν	μετά	πριν	μετά
θάλασσα	0.77	0.85	0.65	0.75	0.62	0.72
βλάστηση	0.35	0.43	0.35	0.40	0.62	0.72
ουρανός	0.45	0.57	0.55	0.60	0.53	0.61
άμμος	0.69	0.75	0.45	0.56	0.52	0.60
βράχος	0.25	0.35	0.63	0.68	0.65	0.75
χύμα	0.00	0.00	0.25	0.34	0.20	0.27

τιμές εμπιστοσύνης με έναν επαναληπτικό τρόπο και λαμβάνοντας υπόψη τις βασισμένες στο εννοιολογικό πλαίσιο σχέσεις μεταξύ των τύπων περιοχής και των εννοιών που διαμορφώνουν την εικόνα, παρουσιάζουμε στον Πίνακα 7.10 συνολικά σκορ ακρίβειας και ανάκλησης ανά έννοια από την εφαρμογή της προτεινόμενης μεθοδολογίας σε ολόκληρο το σύνολο δεδομένων. Κάθε γραμμή καταδεικνύει τα σκορ πριν και μετά την εφαρμογή του προτεινόμενου αλγορίθμου εννοιολογικής επίδρασης.

**Πίνακας 7.10:** Συνολικά σκορ ακρίβειας ανά έννοια - 1148 εικόνες.

Έννοιες	Ακρίβεια			Ανάκληση		
	πριν	μετά	%	πριν	μετά	%
θάλασσα	0.49	0.55	12.35%	0.56	0.49	-12.66%
βλάστηση	0.27	0.34	27.04%	0.35	0.30	-14.29%
ουρανός	0.58	0.62	6.90 %	0.66	0.53	-19.63%
άμμος	0.53	0.57	6.57 %	0.61	0.53	-13.34%
βράχος	0.26	0.38	46.15%	0.41	0.28	-31.14%
χύμα	0.45	0.54	19.82%	0.58	0.38	-35.17%
<b>Σύνολο</b>	<b>0.43</b>	<b>0.53</b>	<b>16.28%</b>	<b>0.50</b>	<b>0.42</b>	<b>-20.75%</b>

Στη συνέχεια των πειραμάτων μας, παρουσιάζουμε ένα σετ από πειραματικά αποτελέσματα από την εφαρμογή του σχετικού αλγόριθμου μεικτού εννοιολογικού πλαισίου ταξινόμησης πάνω σε ένα άλλο σύνολο από 750 εικόνες και 6 υψηλού επιπέδου έννοιες. Το σύνολο των εικόνων που χρησιμοποιήθηκε ήταν ένα υποσύνολο των εικόνων της ευρέως διαδεδομένης και χρησιμοποιούμενης από τους ερευνητές συλλογής Corel [272] και για το λόγο αυτό χρήζει ιδιαίτερης μνείας. Ένα χαρακτηριστικό δείγμα αυτού του συνόλου εικόνων απεικονίζεται στο Σχήμα 7.30. Για την αρχική (δηλ. την χωρίς εννοιολογικό πλαίσιο) ανίχνευση των υψηλού επιπέδου εννοιών εφαρμόστηκε πιστά η προσέγγιση ανίχνευσης, που περιγράψαμε νωρίτερα στο κεφάλαιο αυτό [239]. Χρησιμοποιήσαμε 525 εικόνες, για να εκπαιδεύσουμε 6 ζεχωριστούς ταξινομητές SVM εννοιών και 225 εικόνες ως σύνολο δοκιμών.

Στον Πίνακα 7.11 παρουσιάζουμε τα αρχικά αποτελέσματα ανίχνευσης και εκείνα που προκύπτουν μετά από την επίδραση της μεικτής προσέγγισης εννοιολογικού πλαισίου. Μπορούμε να παρατηρήσουμε ότι η ακρίβεια βελτιώθηκε και για τις 6 έννοιες. Θα πρέπει να σημειώσουμε εδώ ότι οι έννοιες δρόμος και άμμος παρουσίαζαν αρχικά έναν μικρό αριθμό θετικών παραδειγμάτων σε σχέση με τις υπόλοιπες· κατά συνέπεια, ήταν αρκετά δύσκολο να εκπαιδευθούν αξιόπιστοι ταξινομητές SVM γι' αυτές. Εντούτοις, όπως παρατηρούμε, η βασισμένη στο εννοιολογικό πλαίσιο γνώση βελτίωσε και αυτά τα “αδύνατα” αποτελέσματα.

**Πίνακας 7.11:** Ακρίβεια/ανάκληση ανά έννοια πριν και μετά την εφαρμογή του αλγορίθμου επίδρασης του μεικτού εννοιολογικού πλαισίου - 750 εικόνες Corel.

'Εννοιες	Ακρίβεια			Ανάκληση		
	πριν	μετά	%	πριν	μετά	%
δρόμος	0.22	0.43	95.45%	0.40	0.35	-12.15%
άμμος	0.38	0.55	44.74%	0.50	0.44	-12.40%
θάλασσα	0.72	0.89	23.78%	0.85	0.80	-5.21 %
ουρανός	0.81	0.88	8.64 %	0.88	0.82	-6.37 %
χιόνι	0.48	0.72	50.00%	0.68	0.57	-15.88%
βλάστηση	0.67	0.81	21.41%	0.81	0.74	-8.78 %
<b>Συνολικά</b>	<b>0.55</b>	<b>0.71</b>	<b>29.09%</b>	<b>0.68</b>	<b>0.62</b>	<b>-8.82%</b>



**Σχήμα 7.30:** Ενδεικτικές εικόνες από τη συλλογή Corel.

Τέλος, παρουσιάζουμε ένα ακόμα σύνολο από πειραματικές μετρήσεις από την εφαρμογή του μεικτού εννοιολογικού πλαισίου ταξινόμησης πάνω σε 1247 εικόνες από τη συλλογή του TREC [231] και 7 υψηλού επιπέδου έννοιες {βλάστηση, δρόμος, φωτιά, ουρανός, χιόνι, άμμος, νερό}. Ένα χαρακτηριστικό δείγμα αυτού του συνόλου εικόνων απεικονίζεται στο Σχήμα 7.31. Χρησιμοποιήσαμε και πάλι 250 εικόνες, για να εκπαιδεύσουμε 7 ξεχωριστούς ταξινομητές SVM εννοιών και τις υπόλοιπες 997 εικόνες ως σύνολο δοκιμών. Στον Πίνακα 7.12 παρουσιάζουμε τα αρχικά αποτελέσματα ανίχνευσης και εκείνα που προκύπτουν μετά από την επίδραση της μεικτής προσέγγισης εννοιολογικού πλαισίου.

**Πίνακας 7.12:** Σκορ ακρίβειας και ανάκλησης ανά έννοια πριν και μετά την εφαρμογή του μεικτού εννοιολογικού πλαισίου - 1250 εικόνες TRECVID.

'Εννοιες	Ακρίβεια			Ανάκληση		
	πριν	μετά	%	πριν	μετά	%
βλάστηση	0.47	0.70	50.43%	0.69	0.49	-28.47%
δρόμος	0.21	0.39	85.71%	0.34	0.26	-22.12%
άμμος	0.79	0.90	13.92%	0.89	0.83	-6.97 %
νερό	0.57	0.73	28.30%	0.73	0.62	-15.18%
ουρανός	0.57	0.81	41.36%	0.86	0.62	-28.16%
χιόνι	0.41	0.51	25.74%	0.54	0.40	-25.93%
φωτιά	0.28	0.49	75.00%	0.51	0.39	-24.66%
<b>Συνολικά</b>	<b>0.47</b>	<b>0.65</b>	<b>38.30%</b>	<b>0.65</b>	<b>0.52</b>	<b>-20.00%</b>

Με βάση τα παραπάνω αποτελέσματα αποδειχνύεται ξεκάθαρα ότι οι υπάρχουσες σχέσεις μεταξύ των εννοιών βελτιώνουν την ακρίβεια των αποτελεσμάτων όχι μόνο

για τους ήδη καλά εκπαιδευμένους και αποτελεσματικούς ταξινομητές SVM, αλλά και για τους αδύνατους και μη-αποτελεσματικούς. Ο προτεινόμενος αλγόριθμος μεικτού εννοιολογικού πλαισίου χρησιμοποιεί και εκμεταλλεύεται αυτές τις σχέσεις και παρέχει μια διευρυμένη θεώρηση του ερευνητικού προβλήματος, η οποία βασίζεται σε έναν υβριδικό συνδυασμό πρωτότυπων τοπολογικών και σημασιολογικών σχέσεων.

Η ερμηνεία όλων των παραπάνω αποτελεσμάτων υποδηλώνει ότι η προτεινόμενη και βασισμένη στο εννοιολογικό πλαίσιο προσέγγιση θα ευνοήσει τους σχετικά βέβαιους βαθμούς εμπιστοσύνης για την ανίχνευση της έννοιας, που είναι παρούσα σε μία εικόνα, σε αντιδιαστολή με τους σχετικά αβέβαιους ή παραπλανητικούς βαθμούς. Θα ενισχύσει τις διαφορές τους, ενώ από την άλλη η επίδρασή του σε σχετικά βέβαιες έννοιες είναι σημαντική, αλλά μικρότερη. Το τελευταίο διαφοροποιεί ακόμα μία φορά την προσέγγιση αυτή από την περίπτωση της χρήσης του οπτικού εννοιολογικού πλαισίου στις φάσεις της ανάλυσης που εξετάστηκαν στο κεφάλαιο 6, καθώς με τον τρόπο αυτό δεν είναι δυνατόν να εντοπιστούν πιθανώς παραπλανητικές έννοιες για μια συγκεκριμένη εικόνα. Φυσικά, κάτι τέτοιο δεν αποδύναμωνει την προσέγγισή μας, δεδομένου ότι στην περίπτωση της ύπαρξης γενικού σχολιασμού για μια εικόνα η πιθανότητα ύπαρξης παραπλανητικών εννοιών πρακτικά εκμηδενίζεται<sup>7</sup>. Τέλος, με βάση την παρεχόμενη οντολογική γνώση ο προτεινόμενος αλγόριθμος θα διευκρινίσει και θα λύσει τις ασάφειες σε περιπτώσεις παρόμοιων εννοιών ή εννοιών, που είναι αντικειμενικά δύσκολο να προκύψουν από τη χαμηλού επιπέδου ανάλυση.

Προκειμένου να αξιολογήσουμε την ενοποιημένη προσέγγιση του εννοιολογικού πλαισίου, τη συγχρίνουμε με συναφείς σύγχρονες τεχνικές. Τα αποτελέσματα όλων των συγκρίσεων πάνω στα σύνολα δεδομένων του Corel και του TRECVID συνοψίζονται στους Πίνακες 7.13 και 7.14. Μια πρώτη κάπως αφελής προσέγγιση χρησιμοποιεί Περιγραφείς MPEG-7 από όλη την εικόνα κατά τρόπο παρεμφερή με τη δημοσίευση [237], μόνο που αυτή ασχολείται κυρίως με υλικά παρά με έννοιες σκηνής. Εκπαιδεύεται ένας ανεξάρτητος ανιχνευτής (ο οποίος βασίζεται σε ένα κατάλληλο νευρωνικό δίκτυο), για κάθε υψηλού επιπέδου έννοια σύμφωνα με καθολικά χαμηλού επιπέδου χαρακτηριστικά. Η τεχνική αυτή αναφέρεται ως “Καθολική”. Στη συνέχεια αναφερόμαστε στους Τύπους Περιοχής (RT) και παραθέτουμε αποτελέσματα σύμφωνα με την προσέγγιση της ενότητας 7.3, χωρίς τη χρήση του εννοιολογικού πλαισίου. Συγκρίνουμε επίσης και με μια επέκταση της παραπάνω τεχνικής, η οποία χρησιμοποιεί τη Λανθάνουσα Σημασιολογική Ανάλυση (LSA) [242]. Η τεχνική αυτή προσπαθεί να εκμεταλλεύεται τις “χρυφές” σχέσεις μεταξύ των τύπων περιοχής των εικόνων του συνόλου εκπαίδευσης και τροποποιεί κατάλληλα το σχετικό διάνυσμα αναπαράστασης. Τα αποτελέσματά της παρατίθενται ως “RT+LSA” στους παρακάτω Πίνακες.

Στη συνέχεια, υλοποιήσαμε ακόμα δύο τεχνικές για την συνολικότερη και πληρέστερη αξιολόγηση της προτεινόμενης τεχνικής. Η πρώτη (“Rel. LSA” [236]) εισάγει απευθείας δομικούς περιορισμούς στις οπτικές λέξεις του θησαυρού. Η κύρια διαφορά μεταξύ της παραδοσιακής LSA και αυτής της τεχνικής έγκειται στο γεγονός ότι κάθε πιθανό ζευγάρι συστάδων εκτός σειράς (unordered) θεωρείται τώρα ως μία οπτική λέξη. Με τον τρόπο αυτό δημιουργείται ένας οπτικός θησαυρός με πάρα πολλές λέξεις (δηλ. ζευγάρια συστάδων). Εντούτοις, τα χαμηλού επιπέδου χαρακτηριστικά γνωρίσματα που εξάγονται από κάθε περιοχή είναι απλούστερα σε σχέση με τα MPEG-7 χαρακτηριστικά γνωρίσματα που χρησιμοποιούνται στην ενότητα 7.3. Η δεύτερη προσέγγιση [133] εξάγει αρχικά μια σειρά από Τοπικά Σημεία Ενδιαφέροντος (Local

<sup>7</sup>Εν γένει, θεωρείται ότι ένας γενικός σχολιασμός είναι πιο εύκολος και πιο ακριβής από έναν σχολιασμό περιοχών, κυρίως για λόγους απλότητας και πολυπλοκότητας του ανθρώπινου παράγοντα.

**Πίνακας 7.13:** Συγχριτικά αποτελέσματα ακρίβειας  $P$  και ανάκλησης  $R$  ανά έννοια για 6 διαφορετικές μεθοδολογίες ανίχνευσης εννοιών υψηλού επιπέδου πάνω στις 750 εικόνες του Corel.

Έννοιες	Καθολική		Τύπ. Περ. (RT)		RT+LSA		RT+Ενν. Πλ.		LIPs [133]		Rel. LSA [236]	
	P	R	P	R	P	R	P	R	P	R	P	R
δρόμος	0.18	0.48	0.22	0.40	0.27	0.38	0.43	0.35	0.34	0.37	0.42	0.35
άμμος	0.30	0.60	0.38	0.50	0.42	0.48	0.55	0.44	0.47	0.46	0.52	0.45
θάλασσα	0.58	0.95	0.72	0.85	0.75	0.84	0.89	0.80	0.77	0.83	0.80	0.82
ουρανός	0.65	0.94	0.81	0.88	0.83	0.86	0.88	0.82	0.86	0.85	0.88	0.83
χιόνι	0.38	0.82	0.48	0.68	0.53	0.64	0.72	0.57	0.58	0.61	0.64	0.57
βλάστηση	0.54	0.90	0.67	0.81	0.70	0.78	0.81	0.74	0.73	0.76	0.76	0.73
Συνολικά:	<b>0.44</b>	<b>0.78</b>	<b>0.55</b>	<b>0.69</b>	<b>0.58</b>	<b>0.67</b>	<b>0.71</b>	<b>0.62</b>	<b>0.62</b>	<b>0.65</b>	<b>0.67</b>	<b>0.63</b>

Interest Points - LIPs) και την αναπαριστούμε ως "LIPs". Τα τοπικά αυτά σημεία ενδιαφέροντος τείνουν να έχουν σημαντικά διαφορετικές ιδιότητες έναντι όλων των άλλων pixels της γειτονιάς τους. Τελικά, ο σχετικός οπτικός θησαυρός κατασκευάζεται με μια κατάλληλη κβαντοποίηση των σημείων αυτών και κάθε εικόνα περιγράφεται ως ένα διάνυσμα οπτικών λέξεων, ενώ εκπαιδεύεται και ένας κατάλληλος ταξινομητής για κάθε υψηλού επιπέδου έννοια. Στο σημείο αυτό θα πρέπει να σημειώσουμε ότι επιλέξαμε να συγχρίνουμε την προσέγγισή μας με τις ανωτέρω τεχνικές κυρίως γιατί οι τελευταίες τείνουν να αντιμετωπίζουν τελικά το ίδιο πρόβλημα. Επιπρόσθετα, οι δύο τελευταίες προσεγγίσεις έχουν ήδη εφαρμοστεί με επιτυχία στο σύνολο δεδομένων του TRECVID. Όπως είναι μάλλον προφανές από την ερμηνεία των Πινάκων 7.13 και 7.14, η προτεινόμενη βασισμένη στο μεικτό εννοιολογικό πλαίσιο προσέγγιση (η οποία καταδεικνύεται ως "RT+Context") ξεπερνά σε γενικές γραμμές όλες τις συγχρινόμενες προσεγγίσεις από την άποψη της ακρίβειας, ενώ σε μερικές περιπτώσεις υστερεί από την άποψη της ανάκλησης.

**Πίνακας 7.14:** Συγχριτικά αποτελέσματα ακρίβειας  $P$  και ανάκλησης  $R$  ανά έννοια για 6 διαφορετικές μεθοδολογίες ανίχνευσης εννοιών υψηλού επιπέδου πάνω στις 1147 εικόνες του TRECVID.

Έννοιες	Καθολικά		Τύπ. Περ. (RT)		RT+LSA		RT+Ενν. Πλ.		LIPs [133]		Rel. LSA [236]	
	P	R	P	R	P	R	P	R	P	R	P	R
βλάστηση	0.38	0.83	0.47	0.69	0.49	0.66	0.70	0.49	0.51	0.64	0.53	0.61
δρόμος	0.17	0.41	0.21	0.34	0.26	0.33	0.39	0.26	0.31	0.32	0.38	0.30
άμμος	0.63	0.96	0.79	0.89	0.87	0.85	0.90	0.83	0.95	0.82	0.96	0.79
νερό	0.46	0.88	0.57	0.73	0.59	0.72	0.73	0.62	0.61	0.71	0.62	0.71
ουρανός	0.46	0.94	0.57	0.86	0.58	0.85	0.81	0.62	0.60	0.85	0.61	0.84
χιόνι	0.33	0.65	0.41	0.54	0.46	0.51	0.51	0.40	0.51	0.47	0.57	0.44
φωτιά	0.22	0.61	0.28	0.51	0.33	0.50	0.49	0.39	0.39	0.48	0.46	0.47
Συνολικά:	<b>0.38</b>	<b>0.75</b>	<b>0.47</b>	<b>0.65</b>	<b>0.51</b>	<b>0.63</b>	<b>0.65</b>	<b>0.52</b>	<b>0.55</b>	<b>0.61</b>	<b>0.59</b>	<b>0.59</b>

#### 7.5.4 Ταξινόμηση σκηνών

Προκειμένου να αξιολογήσουμε τη δυναμική του εννοιολογικού πλαισίου σε όλες τις πιθανές εφαρμογές του κατά την διαδικασία της ανάλυσης πολυμεσικού περιεχομένου, πραγματοποιήσαμε πειράματα και μετρήσεις και αναφορικά με το τελευταίο βήμα της ταξινόμησης των εικόνων σε κατηγορίες. Στα αποτελέσματα που ακολουθούν εστιάσαμε στη δυαδική κατηγοριοποίηση εικόνων, που απεικονίζουν είτε τοπία παραλίας, είτε τοπία βουνού. Για τα εν λόγω πειράματα χρησιμοποιήθηκαν και οι δύο τεχνικές επίδρασης του εννοιολογικού πλαισίου που επιδρούν στη συνολική ταξινόμηση σκηνών, δηλαδή τόσο το απλό εννοιολογικό πλαίσιο ταξινόμησης τύπων περιοχής

(βλ. υποενότητα 7.4.2), όσο και η τρίτη προσέγγιση του μεικτού εννοιολογικού πλαισίου ταξινόμησης που περιγράφηκε αναλυτικά στην προηγούμενη υποενότητα 7.4.3. Αξιοποιήσαμε και τα τρία σύνολα δεδομένων που παρουσιάσαμε έως τώρα (συλλόγες εικόνων, Corel και TRECVID), δηλ. ολοκληρώσαμε την κατηγοριοποίηση εικόνων μέσω SVMs στα σύνολα των 1148, των 750 και των 1247 εικόνων που είχαμε διαθέσιμα. Ειδικότερα, στο πρώτο σύνολο είχαμε 700 εικόνες παραλίας και 448 εικόνες βουνού, στο σύνολο εικόνων Corel είχαμε 460 εικόνες που απεικόνιζαν παραλία και 290 βουνό και στο σύνολο των 1247 εικόνων από τη συλλογή του TRECVID είχαμε 1033 που απεικόνιζαν παραλία και 214 βουνό. Τα σύνολα αυτά διαχωρίστηκαν με τυχαίο τρόπο σε σύνολα εκπαίδευσης και σύνολα δοκιμών, εκ των οποίων κάθε ένα περιλάμβανε 689 και 459 εικόνες για τις συλλογές, 830 και 417 εικόνες για το TRECVID και 600 και 150 εικόνες για το Corel, αντίστοιχα. Ειδικότερα, η κατανομή του συνόλου εκπαίδευσης για τις συλλογές ήταν 394 εικόνες παραλίας και 295 βουνού, για το TRECVID ήταν 690 εικόνες που απεικονίζουν παραλία και 140 που απεικονίζουν βουνό, ενώ για το Corel 370 εικόνες απεικόνιζαν παραλία και 230 βουνό. Το σύνολο δοκιμών αποτελείτο από 343 εικόνες που απεικονίζουν παραλία και 74 βουνό στο TRECVID, από 306 και 153 στις συλλογές και από 90 και 60 στο Corel, αντίστοιχα. Όλα τα παραπάνω συνοψίζονται στον Πίνακα 7.15.

	Σύνολο Εκπαίδευσης			Σύνολο Δοκιμών		
	Συλλογές	TRECVID	Corel	Συλλογές	TRECVID	Corel
παραλία	394	690	370	306	343	90
βουνό	295	140	230	153	74	60

Πίνακας 7.15: Σύνολα εκπαίδευσης και δοκιμών για την καθολική ταξινόμηση σε εικόνες παραλίας/βουνού.

Το σύνολο των εικόνων, οι οποίες πραγματικά ανήκουν στην πρώτη κατηγορία, ορίζεται ως  $G_1$ , ενώ όσες ανήκουν στη δεύτερη κατηγορία ορίζονται ως  $G_2$ . Το  $G_i$ , δηλαδή, είναι το σύνολο απόλυτης αλήθειας (ground-truth) και δίνεται από τη σχέση (7.46), όπου  $C(k)$  είναι η κατηγορία στην οποία ανήκει η εικόνα  $k$ .

$$G_i : \{k \in K : C(k) = c_i\} \quad (7.46)$$

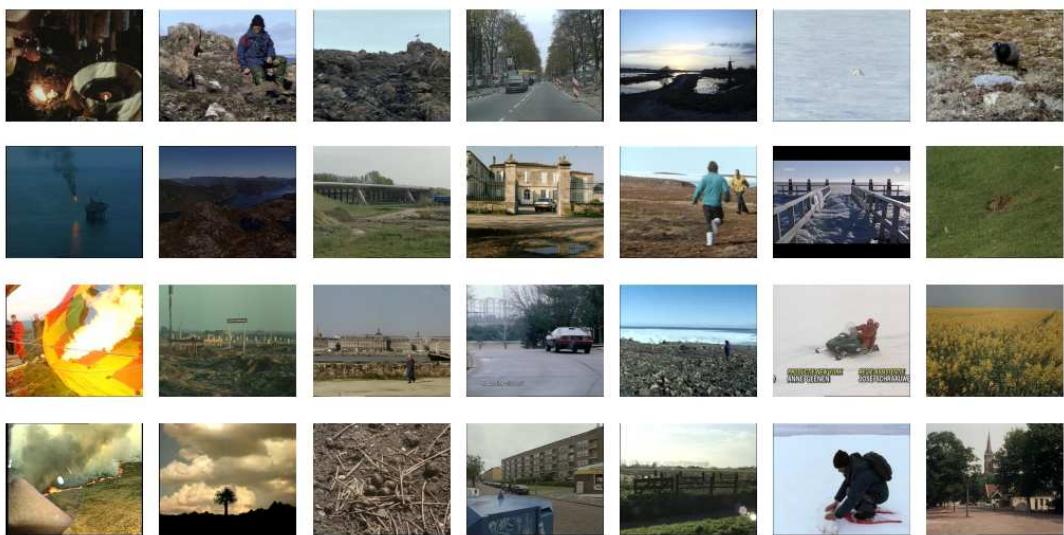
Οι εικόνες που ταξινομήθηκαν στην πρώτη κατηγορία σχηματίζουν το σύνολο  $D_1$  και οι εικόνες που ταξινομήθηκαν στη δεύτερη κατηγορία σχηματίζουν το σύνολο  $D_2$ . Για το συγκεκριμένο πρόβλημα υπολογίζουμε τα σκορ ακρίβειας και ανάκλησης κατά τα γνωστά από την ενότητα 2.4.3 ως  $P_i = \frac{|D_i \cap G_i|}{|D_i|}$  και  $R_i = \frac{|D_i \cap G_i|}{|G_i|}$  αντίστοιχα. Άρα, τελικά, χρησιμοποιώντας την απόλυτη αλήθεια (ground-truth) και την επίδραση του εννοιολογικού πλαισίου, υπολογίζονται τα σκορ ακρίβειας και ανάκλησης πριν και μετά την αξιοποίηση της επιπλέον εννοιολογικής γνώσης. Τα συνολικά αποτελέσματα για τις 2 κατηγορίες εικόνων παρουσιάζονται στον Πίνακα 7.16, ενώ κάποια ενδεικτικά αποτελέσματα ταξινόμησης παρατίθενται στο Σχήμα 7.32.

Γενικά, παρατηρούμε ικανοποιητικά αποτελέσματα βελτίωσης εξαιτίας της επίδρασης του εννοιολογικού πλαισίου και για τις δύο κατηγορίες εικόνων παραλίας και βουνού. Μικρότερη παρατηρείται η επίδραση στις εικόνες της κατηγορίας βουνό, κάτι το οποίο, όμως, είναι αναμενόμενο, δεδομένων των λιγότερων δειγμάτων τους στο αρχικό σύνολο εκπαίδευσης.

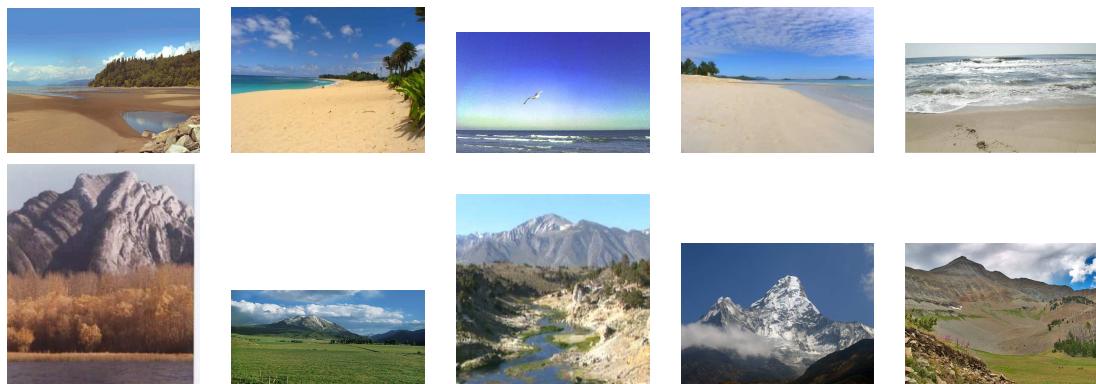
□

	παραλία						βουνό					
	ακρίβεια			ανάκληση			ακρίβεια			ανάκληση		
	πριν	μετά	%									
Συλλογές	0.75	0.82	9.33%	0.81	0.75	-7.41%	0.49	0.63	28.57%	0.37	0.35	-5.41%
TRECVID	0.88	0.98	11.36%	0.92	0.87	-5.43%	0.54	0.66	22.22%	0.42	0.40	-4.76%
Corel	0.83	0.95	14.46%	0.87	0.83	-4.60%	0.48	0.62	29.17%	0.38	0.36	-5.26%

**Πίνακας 7.16:** Αποτελέσματα καθολικής ταξινόμησης εικόνων παραλίας και βουνού πριν και μετά την επίδραση του εννοιολογικού πλαισίου για όλες τις εικόνες (Συλλογές, TRECVID και Corel).



**Σχήμα 7.31:** Ενδεικτικές εικόνες από τη συλλογή TRECVID.



**Σχήμα 7.32:** Ενδεικτικά αποτελέσματα καθολικής ταξινόμησης εικόνες από τη θεματική περιοχή παραλία εμφανίζονται στην επάνω σειρά, ενώ εικόνες από την περιοχή βουνό παρουσιάζονται στην κάτω σειρά.

# Κεφάλαιο 8

## Συμπεράσματα και προτεινόμενη περαιτέρω έρευνα

Ολοκληρώνοντας τις περιοχές που έχουν αποτελέσει το κύριο αντικείμενο μελέτης και έρευνάς μας, στο παρόν κεφάλαιο επιχειρείται μία επισκόπηση της προτεινόμενης περαιτέρω έρευνας στο πεδίο. Η αντιμετώπιση θεμάτων ανάλυσης, αναζήτησης και ανάκτησης πολυμεσικού περιεχομένου με προσαρμογή στο εκάστοτε εννοιολογικό πλαίσιο της πρόσβασης σε αυτό αποτέλεσε έως τώρα τον κύριο άξονα της ενασχόλησής μας. Είναι γεγονός ότι η προσωποποιημένη πρόσβαση και ανάκτηση πολυμεσικού περιεχομένου, καθώς και η ανάλυση και αναζήτηση του τελευταίου με βάση το εννοιολογικό πλαίσιο διαμορφώνουν ένα σαφές πεδίο έρευνας, το οποίο κάθε άλλο παρά κορεσμένο μπορεί να θεωρηθεί. Καινούρια ερευνητικά προβλήματα τίθενται συνέχεια στο χώρο, ενώ ακόμα και η λύση παλαιοτέρων προβλημάτων οδηγεί συχνά σε νέα ερευνητικά μονοπάτια. Κατά συνέπεια, το παρόν κεφάλαιο σκοπεύει να συνοψίσει την σχετική με τη διατριβή ερευνητική δραστηριότητά, καθώς και κάποιες εμφανείς μελλοντικές επεκτάσεις της.

### 8.1 Προτεινόμενη Περαιτέρω Έρευνα

Αποτελεί πεποίθησή μας ότι με βάση την παρούσα εργασία ανοίγονται περαιτέρω δρόμοι έρευνας, εφαρμογών και επεκτάσεων με όχημα το εννοιολογικό πλαίσιο. Είναι πλέον εμφανές ότι μπορεί να χρησιμοποιηθεί μια ευρεία ποικιλία προσεγγίσεων εννοιολογικού πλαισίου για να γεφυρώσει το σημασιολογικό χάσμα που υφίσταται κατά την ευρετηρίαση και ανάκτηση πολυμεσικού περιεχομένου. Πιστεύουμε, δε, ότι το εννοιολογικό πλαίσιο αντιπροσωπεύει μια σημαντική νέα ευκαιρία για την ερευνητική κοινότητα στο πεδίο της αναζήτησης για εύρωστα και ευφυή συστήματα ανάκτησης πολυμεσικού περιεχομένου. Θέλοντας για λόγους πληρότητας να ανατρέξουμε και σε πιθανές μελλοντικές και εκτός των αυστηρών ορίων αυτής της διαδικτορικής διατριβής εφαρμογές, παρουσιάζουμε στη συνέχεια κάποιες σχετικές ιδέες και προτάσεις μας, τόσο σε επίπεδο έρευνας, όσο και εφαρμογών.

#### 8.1.1 Μελλοντικές ερευνητικές επεκτάσεις

Από όσα συζητήθηκαν στο πλαίσιο της διατριβής αυτής, είναι φανερό, ότι η προσωποποίηση περιεχομένου έχει βρει πρόσφορο έδαφος σε μια σειρά από πραγματικές εφαρμογές, όπως είναι οι ειδήσεις, η εκπαίδευση, η διαφήμιση, ο τουρισμός ή και το

ηλεκτρονικό εμπόριο. Στο πλαίσιο αυτό μπορεί να περιγράψει μια σειρά προσωπικών χαρακτηριστικών των χρηστών, όπως το ενδιαφέρον τους για θέματα ή έννοιες (που εξάγεται άμεσα ή έμμεσα, μέσω του ελέγχου της συμπεριφοράς των χρηστών). Εντούτοις, ενώ η δημιουργία και η εκμετάλλευση μεμονωμένων μοντέλων προτιμήσεων και ενδιαφερόντων χρηστών έχουν εξερευνηθεί κατά ένα σημαντικό ποσοστό, η μοντελοποίηση των ίδιων χαρακτηριστικών στο πλαίσιο μιας ομάδας ή κατηγορίας χρηστών, η οποία συνδυάζει εν γένει τα μεμονωμένα μοντέλα των χρηστών, δεν έχει τύχει της ίδιας προσοχής [8], [168], [195].

Είναι, δε, πολύ συχνή η περίπτωση όπου οι χρήστες τέτοιων συστημάτων δεν δρουν μεμονωμένα. Ο πολλαπλασιασμός των ιδεατών κοινοτήτων και της συλλογικής αλληλεπίδρασης (π.χ. διάφοροι χρήστες μπροστά από ένα τηλεοπτικό δέκτη/αποκωδικοποιητή), επιτάσσουν την ανάγκη περαιτέρω έρευνας για τη μοντελοποίηση ομάδων χρηστών, ανοίγοντας νέες προοπτικές, αλλά και θέτοντας νέα ερευνητικά προβλήματα. Φυσικά, η ερώτηση που προκύπτει και δύναται να αντιμετωπιστεί στο πλαίσιο της μελλοντικής έρευνας, είναι πώς μπορεί ένα τέτοιο σύστημα να προσαρμοστεί σε μια ομάδα χρηστών, κατά τέτοιο τρόπο ώστε κάθε άτομο να απολαμβάνει ή και να ωφελείται από τα αποτελέσματα.

Όσον αφορά στο οπτικό εννοιολογικό πλαίσιο, άμεση επέκταση της παρούσης εργασίας είναι η υλοποίηση συνδυαστικής εφαρμογής κατάλληλων μορφών του εννοιολογικού πλαισίου σε διαφορετικές ερευνητικές περιοχές, όπως π.χ. είναι η σύνδεσή του με χαμηλού επιπέδου χαρακτηριστικά γνωρίσματα, με ανιχνευτές οπτικού περιεχομένου, καθώς και με τεχνικές εξαγωγής συμπερασμάτων για τη δημιουργία νέας γνώσης από ήδη υπάρχουσα.

Μία άλλη επέκταση της χρήσης του εννοιολογικού πλαισίου αφορά σε ευφυή χαρακτηρισμό πολυμεσικού περιεχομένου. Η δημιουργία αναγνωριστικών ενδείξεων για πολυμεσικό ή μη περιεχόμενο συνιστά μια νέα μέθοδο αναζήτησης πληροφοριών. Καθώς όλο και περισσότεροι άνθρωποι ανεβάζουν προσωπικό υλικό στο Διαδίκτυο, σε προσωπικές ή σε κοινόχρηστες βάσεις δεδομένων (όπως π.χ. το Flickr, το MySpace, το YouTube, κ.ά.), καλούνται ταυτόχρονα να δώσουν λέξεις-κλειδιά που προσδιορίζουν το υλικό αυτό, με απώτερο σκοπό την καλύτερη οργάνωση των πληροφοριών που παρουσιάζονται. Σύμφωνα με την τελευταία έρευνα του Ινστιτούτου Pew<sup>1</sup>, υπάρχει μια ολοένα αυξανόμενη τέτοια τάση (ιδιαίτερα ανάμεσα στους χρήστες του Διαδικτύου) παγκοσμίως.

### 8.1.2 Μελλοντικές εφαρμογές

Κατά τη γνώμη μας, η έννοια του εννοιολογικού πλαισίου είναι δυνατόν να προσαρμοστεί κατάλληλα, να επεκταθεί και να χρησιμοποιηθεί σε μια πληθώρα εφαρμογών. Θέλοντας για λόγους πληρότητας να ανατρέξουμε και σε πιθανές μελλοντικές και εκτός των αυστηρών ορίων αυτής της διδακτορικής διατριβής εφαρμογές, παρουσιάζουμε στη συνέχεια κάποιες σχετικές ιδέες και προτάσεις.

Όσον αφορά στις ίδιες τις βασικές δομές του εννοιολογικού πλαισίου, αλλά και στην ίδια τη χρήση του, θα ήταν ενδιαφέροντα η επέκταση της αναπαράστασης του εννοιολογικού πλαισίου για την υποστήριξη νέων δομών αναπαράστασης γνώσης, όπως είναι οι “λαονομίες” (folksonomies) και η επέκταση χρήσης του στους τομείς του έξυπνου χαρακτηρισμού (tagging) του πολυμεσικού περιεχομένου. Μια “λαονομία” είναι μια ταξινομία που παράγεται από τον ίδιο το χρήστη και χρησιμοποιείται για να ταξι-

<sup>1</sup> <http://pewresearch.org>

## Κεφάλαιο 8. Συμπεράσματα και προτεινόμενη περαιτέρω έρευνα

νομήσει και να ανακτήσει περιεχόμενο, όπως ιστοσελίδες, φωτογραφίες, video, χρησιμοποιώντας συγκεκριμένης μορφής σημασιολογικές ετικέτες που καλούνται tags. Συνήθως, οι “λαονομίες” βασίζονται στο Διαδίκτυο, αλλά η χρήση τους μπορεί να εμφανιστεί και σε άλλα πλαίσια/περιβάλλοντα. Η διαδικασία της επικόλλησης τέτοιων ετικετών έχει ως στόχο να καταστήσει ένα σημαντικό σώμα πληροφοριών εύκολο να αναζητηθεί, να ανακαλυφθεί και να πλοηγηθεί. Μια καλά ανεπτυγμένη “λαονομία” είναι στην ιδανική περίπτωση προσιτή ως ένα κοινό, διαμοιραζόμενο λεξιλόγιο που προέρχεται και απευθύνεται στους τελικούς χρήστες. Δύο ευρέως διαδεδομένα παραδείγματα ιστοσελίδων που χρησιμοποιούν τέτοιου είδους ετικέτες είναι το Flickr και το del.icio.us.

Η διαδικασία της τοποθέτησης ετικετών, γνωστή και ως “tagging”, αποκτά ιδιαίτερη σημασία για πολλούς δικτυακούς τόπους που φιλοξενούν φωτογραφίες ή video που ανεβάζουν οι χρήστες. Για παράδειγμα, περισσότερο από το ένα τέταρτο (28%) των Αμερικανών έχουν βάλει τέτοιου είδους “ετικέτες” στο περιεχόμενο που έχουν ανεβάσει στο Διαδίκτυο, είτε πρόκειται για κάποια φωτογραφία, κείμενο ή ακόμα και blog. Με τον τρόπο αυτό, γίνεται πιο εύκολη η αναζήτηση για τους επισκέπτες αυτών των ιστοσελίδων, ενώ ταυτόχρονα κατηγοριοποιούνται (ήμιτ-)αυτόματα και οι πληροφορίες. Το tagging μπορεί να πάρει πολλές μορφές. Η Google π.χ. χρησιμοποιεί τους “σελιδοδείκτες” για να κατηγοριοποιήσει τις σελίδες που φιλοξενεί στη μηχανή αναζήτησης της, ενώ άλλες ιστοσελίδες προσφέρουν τη δυνατότητα στους χρήστες τους να βάλουν ετικέτες στο περιεχόμενο τους χωρίς πολλές φορές να το αντιλαμβάνονται οι ίδιοι. Σύμφωνα με μελέτες του πανεπιστημίου Harvard των ΗΠΑ, αυτός ο νέος τρόπος οργάνωσης των πληροφοριών επιτρέπει στους ίδιους τους χρήστες να ορίσουν την ταξινόμηση των πληροφοριών, στοιχείο που μελλοντικά θα επιτρέψει την αμεσότερη συμμετοχή τους στην εξέλιξη τόσο του πολυμεσικού περιεχομένου όσο και του ίδιου του μέσου (δηλαδή του Διαδικτύου).

Σε ένα άλλο επίπεδο, η χρήση του εννοιολογικού πλαισίου σε καινοτόμες εφαρμογές που σχετίζονται με την προσωποποίηση περιεχομένου και την εξαγωγή των προφίλ χρηστών, όπως είναι οι εφαρμογές προσωποποίησης κατά τη χρήση μικρών συσκευών και κινητών τηλεφώνων, θα μπορούσαν να αποτελέσει άλλη μία άκρως ενδιαφέρουσα εφαρμογή υλοποίησης. Τέλος, η επέκταση του ρόλου του εννοιολογικού πλαισίου και η συνδρομή του κατά την εξαγωγή κοινωνικών προφίλ (social profiling) και χαρακτηριστικών κοινωνικής δικτύωσης (social networking), αποτελεί ακόμα μία περαιτέρω εφαρμογή του. Ο κύριος στόχος είναι η εξέλιξη της έννοιας των σημερινών δικτύων διανομής, συνδυάζοντας νέες δομές, όπως τα κοινωνικά δίκτυα, τις κοινωνικές υπηρεσίες λογισμικού και τα συστήματα διανομής πολυμεσικού περιεχομένου σε ένα συγκλίνον περιβάλλον, όπου οι χρήστες θα εξελιχθούν από απλοί παθητικούς δεκτες σε ενεργοί “λαοί” που συνεργάζονται και συμμετέχουν στη διαδικασία. Άμεση παραμένει και η πιθανή χρήση του εννοιολογικού πλαισίου σε καινοτόμες εφαρμογές που σχετίζονται με την προσωποποίηση περιεχομένου και την εξαγωγή των προφίλ χρηστών εκτός των στενών ορίων που θέτουν οι σημερινοί υπολογιστές, όπως π.χ. είναι οι εφαρμογές προσωποποίησης κατά τη χρήση μικρών συσκευών και κινητών τηλεφώνων.

Τέλος, στο πλαίσιο των έως τώρα εργασιών μας, επεκτείναμε την έννοια του σημασιολογικού ευρετήριου και εξηγήσαμε πώς το ευρετήριο αυτό μπορεί να αναλυθεί για την ανίχνευση κατηγοριών που σχετίζονται με οποιοδήποτε έγγραφο πολυμέσων. Εξετάστηκε επίσης, η ύπαρξη σφαλμάτων και αβεβαιότητας στο σημασιολογικό ευρετήριο. Η προσέγγισή μας είναι βασισμένη στην έννοια του ασαφούς εννοιολογικού

πλαισίου γνώσης και τη χρησιμοποίηση ασαφών ταξινομικών σχέσεων. Δεδομένων των απαιτήσεων και των χαρακτηριστικών της έρευνάς μας, η εφαρμογή αυτής της μεθοδολογίας έχει ιδιαίτερη σημασία και ο τομέας μελλοντικής έρευνας και επεκτάσεών της είναι η επιλογή των βέλτιστων ασαφών χειριστών για την εξαγωγή σημασιολογικής πληροφορίας με σαφές νόημα. Τα συμπεράσματά μας μέχρι τώρα δείχνουν ότι η επιλογή αυτή δεν είναι ανεξάρτητη από την ίδια τη γνώση που χρησιμοποιείται. Τέλος, μια ακόμη μελλοντική κατεύθυνση για την έρευνα στο πεδίο αυτό μπορεί να αποτελέσει και η χρησιμοποίηση των υπαρχουσών σαφών ταξινομιών για την παραγωγή της γνώσης που απαιτείται για την ανάλυση των πολυμεσικών εγγράφων.

## 8.2 Συμπεράσματα

Το γενικότερο πλαίσιο στο οποίο συνυπάρχουν και ολοκληρώνονται οι επιμέρους ερευνητικές εργασίες της διατριβής μας είναι η ευφυής εννοιολογική πρόσβαση στην πολυμεσική πληροφορία. Συνεπώς, η συνέχιση της ερευνητικής μας εργασίας δεν μπορεί παρά να είναι προς την ίδια γενική κατεύθυνση. Οι δύο επιμέρους τομείς που παρουσιάζουν το μεγαλύτερο ερευνητικό ενδιαφέρον τα τελευταία χρόνια, είναι αυτοί των οντολογιών και της σύνδεσης των χαρακτηριστικών χαμηλού επιπέδου των πολυμεσικών εγγράφων με σημασιολογικές οντότητες υψηλού επιπέδου.

Όσο αφορά στον τομέα των οντολογιών, τα πρόσφατα ερευνητικά ευρήματα αφορούν κυρίως στην αυστηρά ορισμένη αναπαράσταση της σημασιολογικής γνώσης με τη χρήση προκαθορισμένων σχέσεων. Η βασική υπόθεση είναι πως η γνώση αυτή θα είναι άμεσα αξιοποιήσιμη από συστήματα αναζήτησης και ανάκτησης πληροφορίας. Ωστόσο, δεν έχει δοθεί, ακόμη, προσοχή στο σχεδιασμό και την υλοποίηση αυτών των συστημάτων. Από την άλλη πλευρά, στην παρούσα εργασία έχουν ήδη παρουσιαστεί τεχνικές αξιοποίησης προϋπάρχουσας γνώσης για ευφυή αναζήτηση και ανάκτηση πληροφορίας. Μια συνέχεια της ερευνητικής εργασίας θα οδηγούσε σε συνδυασμό των δύο, μέσω είτε τροποποίησης των τεχνικών της διατριβής ώστε να δέχονται στην είσοδό τους οντολογική πληροφορία, είτε με αυτόματη εξαγωγή της επιθυμητής πληροφορίας από οντολογικές αναπαραστάσεις.

Όσο αφορά στον τομέα της σύνδεσης των χαρακτηριστικών χαμηλού επιπέδου με σημασιολογικές οντότητες, αυτό παραμένει βασικό πρόβλημα της ανάλυσης εικόνας, βίντεο και γενικότερα πολυμεσικής πληροφορίας για πολλά χρόνια. Η άποψη που αρχίζει να επικρατεί είναι πως η αναγνώριση οντοτήτων σε πολυμεσική πληροφορία με βάση τα χαρακτηριστικά τους θα επιτευχθεί μόνο με τον συνυπολογισμό του σχετικού πλαισίου γνώσης. Καθώς, όμως, σύμφωνα με τις τεχνικές που έχουμε έως τώρα παρουσιάσει, το πλαίσιο γνώσης θα μπορεί να είναι απόλυτα γνωστό μόνο μετά την πλήρη ανάλυση του περιεχομένου του πολυμεσικού εγγράφου, χρειάζεται επέκταση της θεωρίας ώστε να είναι δυνατή η εκτίμηση του πιθανού πλαισίου με βάση ατελή πληροφορία. Με αυτό τον τρόπο η ανάλυση του εγγράφου και η εκτίμηση του πλαισίου θα μπορούν να ολοκληρωθούν μέσω μιας συνδυασμένης επαναληπτικής διαδικασίας.

□

# Κεφάλαιο 9

## Κατάλογος δημοσιεύσεων

### 9.1 Περιοδικά

#### 9.1.1 Δημοσιευμένα

1. Ph. Mylonas, D. Vallet, P. Castells, M. Fernandez and Y. Avrithis, *Personalized information retrieval based on context and ontological knowledge*, Knowledge Engineering Review, Volume 23, Issue 1, March 2008.
2. Ph. Mylonas, P. Tzouveli and S. Kollias, *E-learning and intelligent content adaptation: an integrated approach*, International Journal of Continuing Engineering Education and Life-Long Learning (IJCEELL), Special Issue on Integrating Intelligent and Adaptive Hypermedia Techniques in Web-Based Education Systems, Inderscience, Vol. 17, No. 4/5, pp. 273 - 293, Inderscience, September 2007.
3. Th. Athanasiadis, Ph. Mylonas, Y. Avrithis and S. Kollias, *Semantic Image Segmentation and Object Labeling*, IEEE Transactions on Circuits and Systems for Video Technology, Volume 17, Issue 3, Page(s):298 - 312, March 2007.
4. D. Vallet, P. Castells, M. Fernandez, Ph. Mylonas and Y. Avrithis, *Personalized Content Retrieval in Context Using Ontological Knowledge*, IEEE Transactions on Circuits and Systems for Video Technology, Volume 17, Issue 3, Page(s):336-346, March 2007.
5. G. Th. Papadopoulos, Ph. Mylonas, V. Mezaris, Y. Avrithis, I. Kompatsiaris, *Knowledge-Assisted Image Analysis Based on Context and Spatial Optimization*, International Journal on Semantic Web and Information Systems, Vol. 2, No. 3, pp. 17-36, July-September 2006.

#### 9.1.2 Υπό δημοσίευση

6. E. Spyrou, G. Tolias, Ph. Mylonas and Y. Avrithis, *Concept Detection and Keyframe Extraction Using a Visual Thesaurus*, Multimedia Tools and Applications, Springer, 2008, accepted for publication.

7. Ph. Mylonas, Th. Athanasiadis, M. Wallace, Y. Avrithis and S. Kollias, *Semantic Representation of Multimedia Content - Part I: Knowledge Representation and Semantic Indexing*, Multimedia Tools and Applications, Springer, 2007, in press.
8. P. Tzouveli, Ph. Mylonas and S. Kollias, *An intelligent e-learning system based on learner profiling and learning resources adaptation*, Computers & Education, 2007, in press.

### 9.1.3 Υποβεβλημένα προς χρίση

9. Ph. Mylonas, E. Spyrou, Y. Avrithis and S. Kollias, *Using Visual Context and Region Semantics for High-Level Concept Detection*, IEEE Transactions on Multimedia, Special Issue on Integration of Context and Content for Multimedia Management, IEEE, 2009, submitted.

## 9.2 Κεφάλαια Βιβλίων

### 9.2.1 Δημοσιευμένα

1. S. Dasiopoulou, C. Saathoff, Ph. Mylonas, Y. Avrithis, Y. Kompatsiaris, S. Staab, *Semantic Multimedia and Ontologies: Theory and Applications*, Springer-Verlag, Heidelberg, March 2008.
2. E. Spyrou, Ph. Mylonas and Y. Avrithis, *Semantic Multimedia Analysis based on Region Types and Visual Context*, International Federation for Information Processing (IFIP), Artificial Intelligence and Innovations 2007: from Theory to Applications, Springer Boston, Volume 247/2007.
3. G. Th. Papadopoulos, Ph. Mylonas, V. Mezaris, Y. Avrithis and I. Kompat-siaris, *Semantic Image Analysis Optimization based on Context and Spatial Relations*, A. Sheth, M. Lytras (eds.), Advances in Semantic Web and Infor-mation Systems, Vol. 2., Innovations in Semantic Web Based Information Systems, October 2007.
4. Ph. Mylonas, G. Andreou, K. Karpouzis, *A Collaborative Filtering Approach to Personalized Interactive Entertainment using MPEG-21*, Emerging Artificial Intelligence Applications in Computer Engineering, Frontiers in Artificial Intelligence and Applications, vol. 160, IOS Press, Amsterdam, October 2007.
5. Ph. Mylonas, P. Tzouveli, S. Kollias, *An integrated approach towards intelligent educational content adaptation*, Emerging Artificial Intelligence Applications in Computer Engineering, Frontiers in Artificial Intelligence and Applications, vol. 160, IOS Press, Amsterdam, October 2007.
6. M. Wallace, Ph. Mylonas, G. Akrivas, Y. Avrithis and S. Kollias, *Automatic thematic categorization of multimedia documents using ontological information and fuzzy algebra Studies in Fuzziness and Soft Computing*, Soft Computing in Ontologies and Semantic Web, Springer, Ma, Z. (Ed.), Vol. 204, 2006.

7. Th. Athanasiadis, Ph. Mylonas and Y. Avrithis, *A Context-based Region Labeling Approach for Semantic Image Segmentation*, Lecture Notes in Computer Science - Semantic Multimedia, Springer, Vol. 4306/2006, pp. 212-225, 2006.
8. M. Wallace, Ph. Mylonas and S. Kollias, *Detecting and Verifying Dissimilar Patterns in Unlabelled Data*, Advances in Soft Computing, Soft Computing: Methodologies and Applications, Springer Berlin/Heidelberg, pp. 247-258, 2005.
9. P. Castells, M. Fernandez, D. Vallet, Ph. Mylonas and Y. Avrithis, *Self-Tuning Personalized Information Retrieval in an Ontology-Based Framework*, Lecture Notes in Computer Science - On the Move to Meaningful Internet Systems, Springer, Vol. 3762/2005, p. 977, 2005.
10. G. Andreou, Ph. Mylonas, M. Wallace, S. Kollias, *Offering Access to Personalized Interactive Video*, WSEAS Transactions on Systems, Issue 10, Volume 3, pp. 3120-3124, December 2004.
11. Ph. Mylonas, M. Wallace, S. Kollias, *Using k-nearest neighbor and feature selection as an improvement to hierarchical clustering*, Methods and Applications of Artificial Intelligence, Vouros G.A., Panayiotopoulos T. (Eds.), Lecture Notes in Computer Science 3025, Springer, 2004.

### **9.2.2 Υπό δημοσίευση**

12. Ph. Mylonas and Y. Avrithis, *Multimedia personalization*, Encyclopedia of Multimedia, 2nd Edition, Springer, 2008.

### **9.2.3 Σε εποιμασία**

13. Ph. Mylonas and Y. Avrithis, *Visual context and optimization for analysis*, K-Space Book, 2008.

## **9.3 Συνέδρια**

### **9.3.1 Δημοσιευμένα**

1. E. Spyrou, G. Tolias, Ph. Mylonas and Y. Avrithis A Semantic Multimedia Analysis Approach Utilizing a Region Thesaurus and LSA 9th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2008), Klagenfurt, Austria, 7-9 May 2008.
2. Ph. Mylonas, E. Spyrou and Y. Avrithis, *High-Level Concept Detection based on Mid-level Semantic Information and Contextual Adaptation*, 2nd International Workshop on Semantic Media Adaptation and Personalization (SMAP 2007), London, United Kingdom, 17-18 December 2007.

3. Ph. Mylonas, *A fuzzy contextual approach towards intelligent educational content adaptation*, 2nd International Workshop on Semantic Media Adaptation and Personalization (SMAP 2007), London, United Kingdom, 17-18 December 2007.
4. Ph. Mylonas, E. Spyrou, and Y. Avrithis, *Enriching a context ontology with mid-level features for semantic multimedia analysis*, 1st Workshop on Multimedia Annotation and Retrieval enabled by Shared Ontologies, co-located with SAMT 2007, Genova, Italy, 5-7 December 2007.
5. Ph. Mylonas, N. Simou, V. Tzouvaras and Y. Avrithis, *Towards Semantic Multimedia Indexing by Classification & Reasoning on Textual Metadata*, Knowledge Acquisition from Multimedia Content Workshop, co-located with the 2nd International Conference on Semantics and Digital Media Technologies (SAMT 2007), Genova, Italy, 5-7 December 2007.
6. E. Spyrou, P. Kapsalas, G. Tolias, Ph. Mylonas Y. Avrithis et al., *The COST292 experimental framework for TRECVID 2007*, 5th TRECVID Workshop, Gaithersburg, USA, November 2007.
7. E. Spyrou, Ph. Mylonas and Y. Avrithis, *Semantic Multimedia Analysis based on Region Types and Visual Context*, 4th IFIP Conference on Artificial Intelligence Applications & Innovations (AIAI), Athens, Greece, 19-21 September 2007.
8. Ph. Mylonas and Y. Avrithis, *Using Multiple Domain Visual Context in Image Analysis*, 8th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2007), Santorini, Greece, 6-8 June 2007.
9. G. Andreou, Ph. Mylonas and K. Karpouzis, *A collaborative filtering approach to personalized interactive entertainment using MPEG-21*, 3rd International Conference on Web Information Systems and Technologies (WEBIST 2007), Barcelona, Spain, 3-7 March 2007.
10. D. Vallet, M. Fernandez, P. Castells, Ph. Mylonas and Y. Avrithis, *Personalized Information Retrieval in Context*, 21st National Conference on Artificial Intelligence - 3rd International Workshop on Modeling and Retrieval of Context, Boston, USA, 16-17 July 2006.
11. D. Vallet, M. Fernandez, P. Castells, Ph. Mylonas and Y. Avrithis, *A contextual personalization approach based on ontological knowledge*, 17th European Conference on Artificial Intelligence (ECAI 2006) - Contexts and Ontologies: Theory, Practice and Applications Workshop, Riva del Garda, Italy, 28 August 2006.
12. Ph. Mylonas and M. Wallace, *Using ontologies and fuzzy relations in the multimedia personalization process*, 1st International Workshop on Semantic Media Adaptation and Personalization (SMAP 2006), Athens, Greece, 4-5 December 2006.

13. Ph. Mylonas, D. Vallet, M. Fernández, P. Castells and Y. Avrithis, *Ontology-based Personalization for Multimedia Content*, 3rd European Semantic Web Conference - Semantic Web Personalization Workshop, Budva, Montenegro, 11-14 June 2006.
14. Ph. Mylonas, Th. Athanasiadis and Y. Avrithis, *Image Analysis Using Domain Knowledge and Visual Context*, 13th International Conference on Systems, Signals and Image Processing (IWSSIP 2006), Budapest, Hungary, 21-23 September 2006.
15. Ph. Mylonas, Th. Athanasiadis and Y. Avrithis, *Improving image analysis using a contextual approach*, 7th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2006), Seoul, Korea, 19-21 April 2006.
16. S. Dasiopoulou, G. Th. Papadopoulos, Ph. Mylonas, Y. Avrithis, I. Kompatiari, *Using Context and a Genetic Algorithm for Knowledge-Assisted Image Analysis*, 1st International Conference on Semantics And Digital Media Technology (SAMT 2006), Athens, Greece, 6-8 December, 2006.
17. Th. Athanasiadis, Ph. Mylonas and Y. Avrithis, *A Context-based Region Labeling Approach for Semantic Image Segmentation*, 1st international conference on Semantics And digital Media Technology (SAMT), Athens, Greece, 6-8 December 2006.
18. D. Vallet, Ph. Mylonas, M. A. Corella, J. M. Fuentes, P. Castells and Y. Avrithis, *A Semantically-Enhanced Personalization Framework for Knowledge-Driven Media Services*, Proc. of IADIS International Conference on WWW / Internet (ICWI '05), Lisbon, Portugal, 19-22 October 2005.
19. P. Castells, M. Fernández, D. Vallet, Ph. Mylonas and Y. Avrithis, *Self-Tuning Personalized Information Retrieval in an Ontology-Based Framework*, 1st IFIP WG 2.12 & WG 12.4 International Workshop on Web Semantics (SWWS'05), Agia Napa, Cyprus, 31 Oct - 4 Nov 2005.
20. Ph. Mylonas and Y. Avrithis, *Context modelling for multimedia analysis*, Proc. of 5th International and Interdisciplinary Conference on Modeling and Using Context (CONTEXT '05), Paris, France, 5-8 July 2005.
21. Ph. Mylonas, P. Tzouveli and S. Kollias, *Intelligent content adaptation in the framework of an integrated e-learning system*, 16th ACM Conf. on Hypertext & Hypermedia - Workshop on Combining Intelligent & Adaptive Hypermedia Methods/Techniques in Web-Based Education Systems, Salzburg, Austria, 6-9 September 2005.
22. G. Andreou, Ph. Mylonas, M. Wallace, S. Kollias, *Offering Access to Personalized Interactive Video*, Proceedings of the International Conference on Mathematical Methods and Computational Techniques in Electrical Engineering, Vouliagmeni, Athens, December 2004.

23. M. Wallace, Ph. Mylonas, S. Kollias, *Automatic Extraction of Semantic Preferences from Multimedia Documents*, Proceedings of 5th International Workshop on Image Analysis for Multimedia Interactive Services, Lisboa, Portugal, April 2004.
24. Ph. Mylonas, G. Andreou and S. Kollias, *Personalised Interactive Video User Experience*, Proceedings of the International Conference on Multi-platform e-Publishing, Athens, Greece, November 2004.
25. Ph. Mylonas, K. Karpouzis, G. Andreou and S. Kollias, *Towards an integrated personalized interactive video environment*, Proceedings of the 6th IEEE International Symposium on Multimedia Software Engineering, Miami, Florida, USA, December 2004.
26. Ph. Mylonas, M. Wallace, S. Kollias, *Using k-nearest neighbor and feature selection as an improvement to hierarchical clustering*, Proceedings of 3rd Hellenic Conference on Artificial Intelligence, Samos, Greece, May 2004.
27. Ph. Mylonas, P. Tzouveli and S. Kollias, *Towards a personalized e-learning scheme for teachers*, Proceedings of the 4th IEEE International Conference on Advanced Learning Technologies, Joensuu, Finland, August 2004.
28. P. Tzouveli, Ph. Mylonas and S. Kollias, *SPERO - A personalized integrated e-learning system*, Proceedings of the International Association for Development of the Information Society (IADIS) WWW/Internet 2004 International Conference, Madrid, Spain, October 2004.
29. M. Wallace, Ph. Mylonas, S. Kollias, *Detecting and Verifying Dissimilar Patterns in Unlabelled Data*, 8th Online World Conference on Soft Computing in Industrial Applications (WSC8), September - October 2003.
30. M. Wallace, G. Akrivas, Ph. Mylonas, Y. Avrithis, S. Kollias, *Using context and fuzzy relations to interpret multimedia content*, Proceedings of the Third International Workshop on Content-Based Multimedia Indexing (CBMI), IRISA, Rennes, France, September 2003.

### 9.3.2 Υπό δημοσίευση

31. Ph. Mylonas, V. Solachidis, A. Geyer-Schulz, B. Hoser, S. Chapman, F. Ciravegna, S. Staab, P. Smrz, Y. Kompatsiaris and Y. Avrithis, *Efficient Media Exploitation towards Collective Intelligence*, 32nd Annual Conference of the German Classification Society (GfKl), Hamburg, Germany, 16-18 July 2008.
32. V. Solachidis, Ph. Mylonas, A. Geyer-Schulz, B. Hoser, S. Chapman, F. Ciravegna, S. Staab, C. Contopoulos, I. Gkika, P. Smrz, Y. Kompatsiaris and Y. Avrithis, *Generating Collective Intelligence*, 32nd Annual Conference of the German Classification Society (GfKl), Hamburg, Germany, 16-18 July 2008.

33. E. Spyrou, Ph. Mylonas and Y. Avrithis, *Using Region Semantics and Visual Context for Scene Classification*, 1st ICIP Workshop on Multimedia Information Retrieval: New Trends and Challenges, International Conference on Image Processing (ICIP), San Diego, California, U.S.A., 12 October 2008.

□



# Βιβλιογραφία

- [1] T. Adamek, N. O'Connor, N. Murphy, *Region-based Segmentation of Images Using Syntactic Visual Features*, In Proc. of Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), Montreux, Switzerland, April 13-15, 2005.
- [2] P. Adriaans and D. Zantinge, *Data mining*, Addison-Wesley, 1996.
- [3] G. Akrivas, M. Wallace, G. Andreou, G. Stamou and S. Kollias, *Context - Sensitive Semantic Query Expansion*, In Proc. of the IEEE International Conference on Artificial Intelligence Systems (ICAIS), Divnomorskoe, Russia, September 2002.
- [4] G. Akrivas, G. B. Stamou and S. Kollias, *Semantic Association of Multimedia Document Descriptions through Fuzzy Relational Algebra and Fuzzy Reasoning*, IEEE Transactions on Systems, Man, and Cybernetics, part A, Volume 34 (2), March 2004.
- [5] D. Gokalp and S. Aksoy, *Scene Classification Using Bag-of-Regions Representations*, In Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR '07), Minneapolis, Minnesota, USA, June 18-23, 2007.
- [6] W. Al-Khatib, Y.F. Day, A. Ghafoor, P.B. Berra, *Semantic modeling and knowledge representation in multimedia databases*, Knowledge and Data Engineering, IEEE Transactions on Publication vol. 11, issue 1, pp. 64-80, Jan/Feb 1999.
- [7] M. C. Angelides, *Special issue on Multimedia content modeling and personalization*, IEEE Multimedia 10(4), Oct-Dec 2003.
- [8] L. Ardissono, A. Goy, G. Petrone, M., Segnan and P. Torasso, *INTRIGUE: Personalised Recommendation of Tourist Attractions for Desktop and Handset Devices*, Applied Artificial Intelligence 17(8-9), 2003.
- [9] Th. Athanasiadis, V. Tzouvaras, K. Petridis, F. Precioso, Y. Avrithis and Y. Kompatsiaris, *Using a Multimedia Ontology Infrastructure for Semantic Annotation of Multimedia Content*, In Proc. of 5th International Workshop on Knowledge Markup and Semantic Annotation (SemAnnot '05), Galway, Ireland, November 2005.
- [10] Th. Athanasiadis, Ph. Mylonas, Y. Avrithis and S. Kollias, *Semantic Image Segmentation and Object Labeling*, IEEE Transactions on Circuits and Systems for Video Technology, Volume 17, Issue 3, Page(s):298 - 312, March 2007.

- [11] Y. Avrithis, A. Doulamis, N. Doulamis and S. Kollias, *A stochastic framework for optimal key frame extraction from mpeg video databases*, 1999.
- [12] F. Baader, D. Calvanese, D. L. McGuinness, D. Nardi, and P. F. Patel-Schneider, *The Description Logic Hand-book: Theory, Implementation and Application*, Cambridge University Press, 2002.
- [13] J.R. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Jain, C.F. Shu, *Virage image search engine: an open framework for image management*, In Jain R. (ed) Symposium on Electronic Imaging: Science and Technology - Storage and Retrieval for Image and Video databases IV, vol. IV, pp. 76-87, 1996.
- [14] J.J. Baker, *A note on multiplying Boolean matrices*, Comm. ACM 5(2):102, 1962.
- [15] P. Baldi and S. Brunak, *Bioinformatics: the machine learning approach*, MIT Press, 1998.
- [16] M. Bar and S. Ullman, *Spatial context in recognition*, Perception, vol. 25, pp. 342-352, 1996.
- [17] K. Barnard, P. Duygulu, R. Guru, P. Gabbur, D. Forsyth, *The effects of segmentation and feature choice in a translation model of object recognition*, Intern. Conf. On Computer Vision and Pattern Recognition, 2003.
- [18] K. Barnard, P. Duygulu, D. Forsyth, N. de Freitas, D. M. Blei, and M. I. Jordan, *Matching words and pictures*, J. Mach. Learn. Res., 3:1107-1135, 2003.
- [19] Y. Batistakis, M. Halkidi, M. Vaziriannis, *On clustering validation techniques*, Journal of Intelligent Information Systems, 107-120, 2001.
- [20] J. Batlle, A. Casals, J. Freixenet, and J. Martí, *A review on strategies for recognizing natural objects in colour images of outdoor scenes*, Image Vis. Comput., vol. 18, nos. 6-7, pp. 515-530, 2000.
- [21] N. J. Belkin, C. Cool, A. Stein, U. Thiel, *Cases, Scripts, and Information-Seeking Strategies: On the Design of Interactive Information Retrieval Systems*, Expert Systems with Applications 9(3):379-395, 1995.
- [22] A.B. Benitez et al., *Object-based multimedia content description schemes and applications for MPEG-7*, Image Communication Journal (ICJ), Invited Paper on a Special Issue on MPEG-7, Vol. 16, pp. 235-269, 2000.
- [23] A.B. Benitez, D. Zhong, S. Chang and J. Smith, *MPEG-7 MDS content description tools and applications*, In Proc. of International Conference on Computer Analysis of Images and Patterns (CAIP), Warsaw, Poland, 2001.
- [24] A.B. Benitez and S.-F. Chang, *Image Classification Using Multimedia Knowledge Networks*, Proceedings of the IEEE Int. Conf. on Image Processing (ICIP'03), Barcelona, Spain, 2003.

- [25] M. Benkhalifa, A. Bensaid and A. Mouradi, *Text categorization using the semi-supervised fuzzy c-means algorithm*, Proceedings of the 18<sup>th</sup> International Conference of the North American Fuzzy Information Processing Society - NAFIPS, pp. 561-5, 1999.
- [26] P. Berkhin, *Survey of clustering data mining techniques*, Technical report, Accrue Software, San Jose, CA, 2002.
- [27] M. W. Berry, S. T. Dumais and G. W. O'Brien, *Using linear algebra for intelligent information retrieval*, SIAM Review, volume 37(4), pp. 177-196, 1995.
- [28] M. Berry and G. Linoff, *Data mining techniques for marketing, sales, and customer support*, John Wiley, 1997.
- [29] M. J. F. Bes and F. Khantache, *A Generic Architecture for Automated Construction of Multimedia Presentations*, Intl. Conf. on Multimedia Modeling (MMM 2001), Amsterdam, The Netherlands, Nov. 5-7 2001.
- [30] S. Beucher and F. Meyer, *The Morphological Approach to Segmentation: The Watershed Transformation*, in Mathematical Morphology in Image Processing, E.R.Dougherty (Ed.), Marcel Dekker, NY, 1993.
- [31] K. Bharat, *SearchPad: Explicit capture of search context to support web search*, in Proc. of the 9<sup>th</sup> International World Wide Web Conference (WWW9). Amsterdam, The Netherlands, 2000.
- [32] I. Biederman, *Perceiving real world scenes*, Science, Vol. 177, pp. 77-80, 1972.
- [33] L. Biederman, R.J. Mezzanotte, J.C. Rabinowitz, *Scene perception: Detecting and judging objects undergoing relational violations*, Cogn. Psychol. Vol. 14, pp. 143-177, 1982.
- [34] A. Del Bimbo, *Visual Information Retrieval*, Morgan Kaufmann Publishers, 2001.
- [35] E. Borenstein, E. Sharon and S. Ullman, *Combining Top-Down and Bottom-Up Segmentation*, Computer Vision and Pattern Recognition Workshop, 27-02 June, 2004.
- [36] N. Boujemaa, F. Fleuret, V.G., Sahbi, H., *Visual content extraction for automatic semantic annotation of video news*, In: IS&T/SPIE Conference on Storage and Retrieval Methods and Applications for Multimedia, part of Electronic Imaging symposium, 2004.
- [37] M. Boutell and J. Luo, *Bayesian fusion of camera metadata cues in semantic scene classification*, in Proc. IEEE Conf. Computer Vision Pattern Recognition (CVPR), Washington, DC, vol. 2, pp. 623-630, 2004.
- [38] M. Boutell, J. Luo, and C.M. Brown, *A generalized temporal context model for classifying image collections*, ACM Multimedia Syst. J., vol. 11, no. 1, pp. 82-92, Nov. 2005.

- [39] M. Boutell, J. Luo, and C. Brown, *Improved semantic region labeling based on scene context*, in Proc. IEEE Int. Conf. Multimedia Expo (ICME), Amsterdam, The Netherlands, pp. 980-993, 2005.
- [40] M. Boutell, J. Luo, and C. Brown, *Learning spatial configuration models using modified Dirichlet priors*, in Proc. IEEE Workshop Statistical Relational Learning (in conjunction with ICML), Banff, Alberta, 2004.
- [41] M. Boutell, J. Luo, X. Shena and C. Brown, *Learning multi-label scene classification*, Pattern Recognition, 37(9), pp. 1757-1771, September 2004.
- [42] M. Boutell, J. Luo, and C. Brown, *Factor graphs for region-based whole-scene classification*, In CVPR, Semantic Learning Workshop, New York, NY, June 17-22, 2006.
- [43] C.M. Bowman, P.B. Danzing, U. Manber, F. Schwartz, *Scalable Internet resources discovery: research problems and approaches*, Communications of the ACM 37:98–107, 1994.
- [44] D. Brickley and R.V. Guha, *RDF Schema Specification 1.0, W3C Recommendation 10 February 2004*, <http://www.w3.org/TR/rdf-schema/>, 2004.
- [45] P. J. Brown, J. Bovey and X. Chen, *Context-Aware Applications: From the Laboratory to the Marketplace*, IEEE Personal Communications vol. 4, no. 5, pp. 58-64, 1997.
- [46] P. Brusilovsky, A. Kobsa, and J. Vassileva, *Adaptive Hypertext and Hypermedia*, Kluwer Acad. Publ., Dordrecht, 1998.
- [47] C. Buckley, *Implementation of the SMART Information Retrieval System*, Technical Report: TR85-686, Cornell University, Ithaca, NY, USA, 1985.
- [48] C. Buckley, J. Allan, G. Salton, *Automatic Routing and Retrieval Using Smart: TREC-2*, Information Processing and Management 31(3):315–326, 1995.
- [49] R. Burgin, *The retrieval effectiveness of five clustering algorithms as a function of indexing exhaustivity*, Journal of the American Society for Information Science, volume 46(8), pp. 562-572, 1995.
- [50] L. Cai and T. Hofmann, *Text Categorization by Boosting Automatically Extracted Concepts*, Proceedings of the 26th annual international ACM SIGIR Conference on Research and Development in Information Retrieval, Toronto, Canada, pp. 182-189, July/August 2003.
- [51] I. Campbell and C. J. Van Rijsbergen, *The ostensive model of developing information needs*, In Proc. of the 2<sup>nd</sup> International Conference on Conceptions of Library and Information Science, Copenhagen, Denmark: Royal School of Librarianship, pp. 251-268, 1996.
- [52] P. Campisi, A. Neri, G. Panci, G. Scarano, *Robust rotation-invariant texture classification using a model based approach*, JOSA-A, vol. 21, no. 6, pp. 913-925, June 2004.

- [53] P. Carbonetto, N. de Freitas, and K. Barnard, *A statistical model for general contextual object recognition*, In Proc. of ECCV 2004, pages 350-362, 2004.
- [54] P. Castells, M. Fernandez, D. Vallet, *An Adaptation of the Vector-Space Model for Ontology-Based Information Retrieval*, IEEE Transactions on Knowledge and Data Engineering vol. 19, no. 2, special issue on Knowledge and Data Engineering in the Semantic Web Era, pp. 261-272, February 2007.
- [55] P. Castells, M. Fernandez, D. Vallet, Ph. Mylonas and Y. Avrithis, *Self-Tuning Personalized Information Retrieval in an Ontology-Based Framework*, in Proc. of the 1<sup>st</sup> Int. Workshop on Web Semantics, Springer Verlag LNCS vol. 3762, pp. 977-986, 2005.
- [56] C. Cave and S. Kosslyn, *The role of parts and spatial relations in object identification*, Perception, Vol. 22, pp. 229-248, 1993.
- [57] D. C. K. Cees, G. M. Snoek, M. Worring and A. W. Smeulders, *Learned lexicon-driven interactive video retrieval*, 2006.
- [58] S. Chabrier, B. Emile, C. Rosenberger and H. Laurent, *Unsupervised Performance Evaluation of Image Segmentation*, EURASIP Journal on Applied Signal Processing, vol. 2006, Article ID 96306, 12 pages, 2006.
- [59] M. Chalmers, *A historical view of context*, Computer Supported Cooperative Work (CSCW), 13(3), pp. 223-247, 2004.
- [60] C. H. Chang, C. C. Hsu, *Integrating query expansion and conceptual relevance feedback for personalized Web information retrieval*, Computer Networks and ISDN Systems 30(1-7):621-623, 1998.
- [61] S.-F. Chang and H. Sundaram, *Structural and Semantic Analysis of Video*, IEEE International Conference on Multimedia and Expo (II), 2000.
- [62] S.-F. Chang, T. Sikora, A. Puri, *Overview of the mpeg-7 standard*, IEEE trans. on Circuits and Systems for Video Technology, 11(6) (2001), pp. 688–695, 2001.
- [63] P. M. Chen and F. C. Kuo, *An information retrieval system based on a user profile*, Journal of Systems and Software, 54, pp. 3-8, 2000.
- [64] G. Ciocca and R. Schettini, *A relevance feedback mechanism for content-based image retrieval*, Information Processing and Management 35(5):605–632, 1999.
- [65] P. L. Correia and F. Pereira, *Objective Evaluation of Video Segmentation Quality*, IEEE Transactions on Image Processing, Vol. 12, No. 2, February 2003.
- [66] J. Coutaz, J. Crowley, S. Dobson, D. Garlan, *Context is key*, Communications of the ACM vol. 48, no. 3, pp. 49-53, 2005.
- [67] B. Croft, R. Cook and D. Wilder, *Government Information on the Internet: Experiences with TOMAS*, Proceedings of Digital Libraries (DL'95), 1995.
- [68] G. Csurka, C.R. Dance, L. Fan, J. Willamowski, C. Bray, *Visual categorization with bags of keypoints*, In Proc. of ECCV Workshop on Statistical Learning in Computer Vision, pages 1-22, 2004.

- [69] D. Cutting, D. R. Karger, J. O. Pedersen and J. W. Tukey, *Scatter/Gather: A Cluster-based Approach to Browsing Large Document Collections*, Proceedings of the ACM/SIGIR, ACM, pages 318-329, 1992.
- [70] C. Dance, J. Willamowski, L. Fan, C. Bray and G. Csurka, *Visual categorization with bags of keypoints*, In Proc. of ECCV - International Workshop on Statistical Learning in Computer Vision, Prague, 2004.
- [71] S. Dasiopoulou and V. K. Papastathis, V. Mezaris, I. Kompatsiaris, and M.G. Strintzis, *An Ontology Framework For Knowledge-Assisted Semantic Video Analysis and Annotation*, In Proceedings of the 4th International Workshop on Knowledge Markup and Semantic Annotation at 3rd International Semantic Web Conference, November 2004.
- [72] S. Dasiopoulou, V. Mezaris, I. Kompatsiaris, V. K. Papastathis and M. G. Strintzis, *Knowledge-assisted semantic video object detection*, IEEE Transactions on Circuits and Systems for Video Technology vol. 15, no. 10, pp. 1210-1224, 2005.
- [73] P. Dawyndt, H. De Meyer, B. De Baets, *The complete linkage clustering algorithm revisited*, Soft Computing - A Fusion of Foundations, Methodologies and Applications, 9(5), pp. 385-392, May 2005.
- [74] S. C. Deerwester, S. T. Dumais, T. K. Landauer, G. W. Furnas and R. A. Harshman, *Indexing by latent semantic analysis*, Journal of the American Society of Information Science, volume 41(6), pp. 391-407, 1990.
- [75] A. Del Bimbo, *Visual Image Retrieval*, Morgan Kaufmann, San Francisco, 1999.
- [76] H. De Meyer, H. Naessens, B. De Baets, *Algorithms for computing the min-transitive closure and associated partition tree of a symmetric fuzzy relation*, European J. Oper. Res. vol. 155, pp. 226-238, 2004.
- [77] L. Denoyer, P. Gallinari, J.-N. Vitta, S. Bruneseaux, *Structured multimedia document classification*, Proceedings of the ACM DOCENG conference, Grenoble, France, 2003.
- [78] J. Dombi, *A general class of fuzzy operators, the De Morgan class of fuzzy operators and fuzziness measures induced by fuzzy operators*, Fuzzy Sets and Systems, 8(2), pp. 149-163, 1982.
- [79] R.O. Duda, P.E. Hart, D.G. Stork, *Pattern Classification*, 2 edn., Wiley Interscience, 2000.
- [80] M. H. Dunham, *Data Mining: Introductory and Advanced Topics*, Prentice Hall, 2003.
- [81] R. Durbin, S. Eddy, A. Krogh and G. Mitchinson, *Biological sequence analysis: probabilistic models of proteins and nucleic acids*, Cambridge University Press, 1998.

- [82] P. Duygulu, K. Barnard, N. de Freitas, and D. Forsyth, *Object recognition as machine translation: Learning a lexicon for a fixed image vocabulary*, In Proc. of ECCV 2002, pages 97-112, 2002.
- [83] L. Egghe and C. Michel, *Construction of weak and strong similarity measures for ordered sets of documents using fuzzy set techniques*, Information Processing and Management vol. 39, no. 5, pp. 771-807, September 2003.
- [84] D. Fallside, *XML Schema Part 0: Primer*, W3C Recommendation, <http://www.w3.org/TR/xmlschema-0/>, 2001.
- [85] J. Fan, Y. Gao, and H. Luo, *Multi-level annotation of natural scenes using dominant image components and semantic concepts*, In Proc. of ACM Multimedia 2004, pages 540-547, 2004.
- [86] U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, *From data mining to knowledge discovery: An Overview*, In Advances in Knowledge Discovery and Data Mining, U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, eds., MIT Press, Cambridge, Mass., 1-36, 1996.
- [87] L. Fei-Fei and P. Perona, *A Bayesian hierarchical model for learning natural scene categories*, In CVPR, volume 2, pages 524-531, San Diego, CA, June 20-25, 2005.
- [88] R. Feldman and I. Dagan, *Knowledge discovery in textual databases (KDT)*, In proceedings of the First International Conference on Knowledge Discovery and Data Mining (KDD-95), Montreal, Canada, August 20-21, AAAI Press, 112-117, 1995.
- [89] A. Fielding, Machine learning methods for ecological applications, Kluwer Academic Publishers, 1999.
- [90] L. Finkelstein, E. Gabrilovich, Y. Matias, E. Rivlin, Z. Solan, G. Wolfman, E. Ruppin, *Placing Search in Context: The Concept Revisited*. ACM Transaction on Information Systems vol. 20, no. 1, pp. 116-131, 2002.
- [91] R.A. Fischler, O. Firchein, *Intelligence: The eye, the brain and the computer*, Reading, MA, Addison-Wesley, 1997.
- [92] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, *Query by image and video content: the QBIC system*, IEEE Computer, vol. 28, no. 9, pp. 23-32, 1995.
- [93] W. Frawley and G. Piatetsky-Shapiro and C. Matheus (Fall 1992). *Knowledge Discovery in Databases: An Overview*. AI Magazine: pp. 213-228, ISSN 0738-4602, 1992.
- [94] H. Gao, W.-C. Siu and C.-H. Hou, *Improved techniques for automatic image segmentation*, IEEE Trans. on Circuits and Systems for Video Technology, vol.11, no. 12, pp. 1273-1280, December 2001.
- [95] R. Garcí'a, O. Celma, *Semantic Integration and Retrieval of Multimedia Metadata*, SemAnnot 2005.

- [96] S. Gauch, J. Chaffee and A. Pretschner, *Ontology-Based Personalized Search and Browsing*, Web Intelligence and Agent Systems vol. 1, no. 3-4, pp. 219-234, April 2004.
- [97] B. Geisler and V. Ha, *Modeling User Preferences via Theory Refinement*, Proceedings of International Conference on Intelligent User Interfaces, Santa Fe, New Mexico, United States pp. 87-90, 2001.
- [98] J. C. van Gemert, J. Geusebroek, C. J. Veenman, C. G. M. Snoek, and A. W. M. Smeulders, *Robust scene categorization by learning image statistics in context*, In CVPR, Semantic Learning Workshop, New York, NY, June 17-22, 2006.
- [99] J. Geurts, S. Bocconi, J. van Ossenbruggen and L. Hardman, *Towards Ontology-driven Discourse: From Semantic Graphs to Multimedia Presentations*, Second International Semantic Web Conference (ISWC2003), Sanibel Island, Florida, USA, <http://homepages.cwi.nl/~media/cuypers/>, October 20-23, 2003.
- [100] A. Ghoshal, P. Ircing, and S. Khudanpur, *Hidden Markov models for automatic annotation and content-based retrieval of images and video*, In Proc. of SIGIR 2003, pages 544-551, 2003.
- [101] J. Goldberger, S. Gordon and H. Greenspan, *An efficient image similarity measure based on approximations of KL-divergence between two Gaussian Mixtures*, In Proc. of International Conference on Computer Vision, pp. 487-493, Nice, France, 2003.
- [102] H. Greenspan, J. Goldberger and L. Ridel, *A continuous probabilistic framework for image matching*, Journal of Computer Vision and Image Understanding. pp. 384-406, 2001.
- [103] H. Greenspan, S. Gordon and J. Goldberger, *Probabilistic models for generating, modeling and matching image categories*, In Proc. of the International Conference on Pattern Recognition, Quebec, August 2002.
- [104] T.R. Gruber, *A Translation Approach to Portable Ontology Specification*, Knowledge Acquisition 5: 199-220, 1993.
- [105] F.H. Hamker, *The reentry hypothesis: linking eye movements to visual perception*, J. Vision, vol. 11, pp. 808-816, 2003.
- [106] F.H. Hamker, *A dynamic model of how feature cues guide spatial attention*, Vision Res., vol. 44, pp. 501-521, 2004.
- [107] F.H. Hamker, *The emergence of attention by population-based inference and its role in distributed processing and cognitive control of vision*, Eds. L. Itti, G. Rees, J. Tsotsos, Academic Press Inc. (London) Ltd, 2005.
- [108] D. Hand, H. Mannila, P. Smyth, *Principles of Data Mining*, MIT Press, Cambridge, MA., ISBN 0-262-08290-X, 2001.

- [109] S. Handschuh and S. Staab, editors. *Annotation for the Semantic Web*, IOS Press, 2003.
- [110] S. Handschuh and S. Staab, *Cream - creating metadata for the semantic web*, Computer Networks, Elsevier, 42:579–598, August 2003.
- [111] T. H. Haveliwala, *Topic-Sensitive PageRank*, in Proc. of the 11<sup>th</sup> International World Wide Web Conference (WWW 2002). Honolulu, Hawaii, USA, May 7–11, pp. 517–526, 2002.
- [112] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd edition, Prentice Hall, 1999.
- [113] M. A. Hearst, *Text data mining: Issues, techniques, and the relationship to information access*, Presentation notes for UW/MS workshop on data mining, July 1997.
- [114] J. Heer, A. Newberger, C. Beckmann and J. Hong, *liquid: Context-aware distributed queries*, in Proceedings of the 5<sup>th</sup> International Conference on Ubiquitous Computing, pp. 140–148, 2003.
- [115] M. Hegland, *Data mining, challenges, models, methods and algorithms*, 2003.
- [116] J. Heinsohn, *Probabilistic description logics*, Proceedings of UAI-94, pp. 311–318, 1994.
- [117] J.M. Henderson, A. Hollingworth., *High level scene perception*, Annu. Rev. Psychol., vol. 50, pp. 243–271, 1999.
- [118] K. Hirota, W. Pedrycz, *Fuzzy computing for data mining*, Proceedings of the IEEE 87:1575–1600, 1999.
- [119] T. Hofmann, *Probabilistic latent semantic indexing*, Proceedings of the 22nd ACM-SIGIR International Conference on Research and Development in Information Retrieval, pages 50–57, 1999.
- [120] L. Hollink, A.Th. Schreiber, J. Wielemaker and B. Wielinga, *Semantic annotation of image collections*, In Proceedings of the K-CAP 2003 Workshop on Knowledge Markup and Semantic Annotation, Florida, 2003.
- [121] L. Hollink, G. Nguyen, G. Schreiber, J. Wielemaker, B. Wielinga and M. Worring, *Adding Spatial Semantics to Image Annotations*, In Proceedings of the 4th International Workshop on Knowledge Markup and Semantic Annotation at 3rd International Semantic Web Conference, November 2004.
- [122] J. I. Hong and J. A. Landay, *An infrastructure approach to context-aware computing*, in Human-Computer Interaction vol. 16, 2001.
- [123] I. Horrocks, U. Sattler and S. Tobies, *Reasoning with Individuals for the Description Logic*, CADE-2000, LNAI 1831, Springer-Verlag, pages 482–496, 2000.
- [124] J. Hunter, *Adding Multimedia to the Semantic Web Building an MPEG-7 Ontology*, 2001.

- [125] J. Hunter, J. Drennan, and S. Little, *Realizing the hydrogen economy through semantic web technologies*, IEEE Intelligent Systems Journal - Special Issue on eScience, 19:40-47, 2004.
- [126] IBM, *Marvel: Multimedia analysis and retrieval system*. <http://mp7.watson.ibm.com/>
- [127] A. Isaac and R. Troncy, *Designing and Using an Audio-Visual Description Core Ontology*, 2004.
- [128] L. Itti, C. Koch, E. Niebur, *A model of saliency-based visual attention for rapid scene analysis*, IEEE Trans. Pattern Anal. Mach. Intell, vol. 20, no. 11, pp. 1254-1259, 1998.
- [129] M. Jaeger, *Probabilistic reasoning in terminological logics*, Proceedings of KR-94, pp. 305-316, 1994.
- [130] A.K. Jain, M.N. Murty, P.J. Flynn, *Data clustering: a review*, ACM Computing Surveys, Vol.31, 265-322, 1999.
- [131] G. Jeh and J. Widom, *Scaling Personalized Web Search*, in Proc. of the 12<sup>th</sup> International World Wide Web Conference (WWW2003). Budapest, Hungary, May 20-24, pp. 271-279, 2003.
- [132] J. Jeon, V. Lavrenko and R. Manmatha, *Automatic image annotation and retrieval using cross-media relevance models*, In Proc. of SIGIR 2003, pages 119-126, 2003.
- [133] Y.-G. Jiang, W.-L. Zhao, C.-W. Ngo, *Exploring Semantic Concept Using Local Invariant Features*, In Asia-Pacific Workshop on Visual Information Processing, Beijing, China, 2006.
- [134] R. Jin, J. Y. Chai and L. Si, *Effective automatic image annotation via a coherent language model and active learning*, In Proc. of ACM Multimedia 2004, pages 892-899, 2004..
- [135] J. Jiten, B. Merialdo and B. Huet, *Semantic feature extraction with multi-dimensional hidden Markov model*, In Proc. of SPIE CMCAMR 2006, volume 6073, pages 211-221, 2006.
- [136] D. Kelly, J. Teevan, *Implicit feedback for inferring user preference*, SIGIR Forum vol. 32, no. 2, pp. 18-28, 2003.
- [137] H. Kim and P. Chan, *Learning Implicit User Interest Hierarchy for Context in Personalization*, In Proc. of the International Conference on Intelligent User Interfaces (IUI 2003), Miami, USA, pp. 101-108, January 12-15, 2003.
- [138] A. Kiryakov, B. Popov, I. Terziev, D. Manov, D. Ognyanoff, *Semantic Annotation, Indexing, and Retrieval*, Journal of Web Semantics vol. 2, no. 1, Elsevier, pp. 47-49, 2004.
- [139] G. Klir, B. Yuan, *Fuzzy Sets and Fuzzy Logic, Theory and Applications*, Prentice Hall, 1995.

- [140] G. Klyne, J. J. Carroll and B. McBride, *Resource Description Framework (RDF): Concepts and Abstract Syntax*, W3C Recommendation, February 2004.
- [141] A. Kobsa, *Generic User Modelling Systems*, User Modeling and User-Adapted Interaction vol. 11, pp. 49-63, 2001.
- [142] R. Kohavi, D. Sommerfield, *Feature Subset Selection Using the Wrapper Model: Overfitting and Dynamic Search Space Topology*, Proceedings of KDD-95, 1995.
- [143] M. Koelsch, V. Pavlovic, B. Kisacanin and T. S. Huang, *Special Issue on Vision for Human-Computer Interaction*, Computer Vision and Image Understanding, Volume 108, Issues 1-2, October-November 2007.
- [144] I. Kompatsiaris, V. Mezaris and M. G. Strintzis, *Multimedia content indexing and retrieval using an object ontology*, Multimedia Content and Semantic Web - Methods, Standards and Tools, Editor G. Stamou, Wiley, New York, NY, 2004.
- [145] D. H. Kraft, G. Bordogna, G. Pasi, *Information Retrieval Systems: Where is the fuzz?*, IEEE World Congress on Computational Intelligence; IEEE International Conference on Fuzzy Systems, 1998.
- [146] I. Kraraosil and J. Micbalek, *Fuzzy metrics and statistical metric spaces*, Kybernetika, 11(5), pp. 336-344, 1975.
- [147] T. Landauer, P. Foltz and D. Laham, *An Introduction to Latent Semantic Analysis*, Discourse Processes, 25, pp. 259-284, 1998.
- [148] P. Langley, *Elements of Machine Learning*, Morgan Kaufmann, 1996.
- [149] S. Lawrence, *Context in Web Search*, IEEE Data Engineering Bulletin vol. 23, no. 3, pp. 25-32, 2000.
- [150] S. Lazebnik, C. Schmid, J. Ponce, *Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories*, IEEE Conference on Computer Vision and Pattern Recognition, CVPR 06 - Volume 2, 2169-2178, 2006.
- [151] H.-S. Lee, *An optimal algorithm for computing the maxmin transitive closure of a fuzzy similarity matrix*, Fuzzy Sets & Systems vol. 123, pp. 129-136, 2001.
- [152] K.L. Lee, H.L. Chen, *An efficient computation method for the texture browsing descriptor of MPEG-7*, IVC, vol. 23, no. 5, pp. 479-489, May 2005.
- [153] D. Lewis, *Index, Context, and Content*, in Kanger, S. and Ohman, S. (Eds.), Philosophy and Grammar, Reidel Publishing, 1980.
- [154] W. S. Li and D. Agrawal, *Supporting web query expansion efficiently using multi-granularity indexing and query processing*, Data and Knowledge Engineering 35(3):239–257, 2000.
- [155] J. Li, A. Najmi and R. M. Gray, *Image classification by a two-dimensional hidden Markov model*, IEEE Trans. Signal Processing, 48(2):517-533, 2000.

- [156] W. Li and M. Sun, *Semi-supervised learning for image annotation based on conditional random fields*, In Proc. of CIVR 2006, pages 463-472, 2006.
- [157] T.-S. Lim, W.-Y. Loh, Y.-S. Shih, *A Comparison of Prediction Accuracy, Complexity, and Training Time of Thirty-three Old and New Classification Algorithms*, Machine Learning 40:203-229, 2000.
- [158] P. Lipson, E. Grimson, and P. Sinha, *Configuration based scene classification and image indexing*, Proc. IEEE International Conference on Computer Vision and Pattern Recognition, 1997.
- [159] F. Liu, C. Yu and W. Meng, *Personalized Web Search For Improving Retrieval Effectiveness*, IEEE Transactions on Knowledge and Data Engineering vol. 16, no. 1, pp. 28-40, January 2004.
- [160] J. Luo and A. Savakis, *Indoor vs outdoor classification of consumer photographs using low-level and semantic features*, In Proc. of IEEE Int. Conf. on Image Processing (ICIP01), 2001.
- [161] J. Luo, A. Singhal, W. Zhu, *Natural Object Detection in Outdoor Scenes Based on Probabilistic Spatial Context Models*, in Proceedings of IEEE International Conference on Multimedia and Expo, 2002.
- [162] K. MacLeod, *An application specific neural model for document clustering*, Proceedings of the 4th annual parallel processing symposium, vol. 1, pp. 5-16, 1990.
- [163] A. Maedche, B. Motik, N. Silva and R. Volz, *MAFRA - An Ontology MAPPING Framework in the Context of the SemanticWeb*, In Proc. of the Workshop on Ontology Transformation at ECAI2002, Lyon, France, July 2002.
- [164] B. S. Manjunath, J. R. Ohm, V. V. Vasudevan, A. Yamada, *Color and texture descriptors*, Special Issue on MPEG-7, IEEE Trans. on Circuits and Systems for Video Technology, 11/6, 703-715, June 2001.
- [165] D. Martin, C. Fowlkes, J. Malik, *Learning to detect natural image boundaries using local brightness, color and texture cues*, IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, no. 5, pp. 530-549, 2004.
- [166] A. Mathes, *Folksonomies - Cooperative Classification and Communication Through Shared Metadata*, Computer Mediated Communication - LIS590CMC, Graduate School of Library and Information Science, University of Illinois Urbana-Champaign, 2004.
- [167] J. McCarthy, *Notes on Formalizing Context*, in Proc. of the 13th International Joint Conference on Artificial Intelligence (IJCAI 1993), Chambéry, France, pp. 81-98, August-September 1993.
- [168] J. F. McCarthy and T. D. Anagnost, *MusicFX: An Arbiter of Group Preferences for Computer Supported Collaborative Workouts*, Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work (pp. 363-372). Seattle, WA, USA, 1998.

- [169] K. Menger, *Statistical metrics*, Proc. Nat. Acad. Set, 28, pp. 535-537, 1942.
- [170] V. Mezaris, I. Kompatsiaris and M.G. Strintzis, *An Ontology Approach to Object-based Image Retrieval*, In Proc. IEEE Int. Conf. on Image Processing (ICIP03), Barcelona, Spain, Sept. 2003.
- [171] V. Mezaris, I. Kompatsiaris, N. V. Boulgouris and M. G. Strintzis, *Real-time compressed domain spatiotemporal segmentation and ontologies for video indexing and retrieval*, IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Audio and Video Analysis for Multimedia Interactive Services, 14(5):606-621, May 2004.
- [172] A. Micarelli and F. Sciarrone, *Anatomy and Empirical Evaluation of an Adaptive Web-Based Information Filtering System*, User Modelling and User-Adapted Interaction vol. 14, no. 2-3, pp. 159-200, February 2004.
- [173] R. Milanese, *Detecting salient regions in an image: from biology to implementation*, PhD Thesis, University of Geneva, Switzerland, 1993.
- [174] C. Millet, I. Bloch, P. Hede, P.-A. Moellic, *Using relative spatial relationships to improve individual region recognition*, In Proc. of 2nd European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology, EWIMT 2005, 2005.
- [175] T. M. Mitchell, *Machine Learning*, McGraw-Hill Companies, Inc., 1997.
- [176] K. Miyajima and A. Ralescu, *Spatial organization in 2D images*, Proc. of the Third IEEE Conference on: Fuzzy Systems, 1994.
- [177] S. Miyamoto, *Fuzzy Sets in Information Retrieval and Cluster Analysis*, Kluwer Academic Publishers, Dordrecht / Boston / London, 1990.
- [178] F. Monay, P. Quelhas, J.-M. Odobez, and D. Gatica-Perez., *Integrating co-occurrence and spatial contexts on patchbased scene segmentation*, In CVPR, Beyond Patches Workshop, New York, NY, June 17-22, 2006.
- [179] O. J. Morris, M. J. Lee, A. G. Constantinides, *Graph theory for image analysis: An approach based on the shortest spanning tree*, IEEE Proceedings, Vol. 133, pp. 146-152, 1986.
- [180] MPEG-7: Visual experimentation model (XM) version 10.0. ISO/IEC/JTC1/SC29/WG11, Doc. N4062, 2001.
- [181] Ph. Mylonas, M. Wallace and S. Kollias, *Using k-nearest neighbor and feature selection as an improvement to hierarchical clustering*, 3rd Hellenic Conference on Artificial Intelligence, Samos, Greece, 2004.
- [182] Ph. Mylonas and Y. Avrithis, *Context modeling for multimedia analysis and use*, in Proc. of the 5<sup>th</sup> International and Interdisciplinary Conference on Modeling and Using Context (CONTEXT '05), Paris, France, July 2005.
- [183] Ph. Mylonas, Th. Athanasiadis and Y. Avrithis, *Improving image analysis using a contextual approach*. In Proc. of 7th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), Seoul, Korea, 2006.

- [184] Ph. Mylonas, Th. Athanasiadis and Y. Avrithis, *Image Analysis Using Domain Knowledge and Visual Context*, In Proc. of 13th International Conference on Systems, Signals and Image Processing (IWSSIP 2006), Budapest, Hungary, 2006.
- [185] Ph. Mylonas and Y. Avrithis, *Using Multiple Domain Visual Context in Image Analysis*, 8th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2007), Santorini, Greece, 6-8 June 2007.
- [186] Ph. Mylonas, E. Spyrou, and Y. Avrithis, *Enriching a context ontology with mid-level features for semantic multimedia analysis*, 1st Workshop on Multimedia Annotation and Retrieval enabled by Shared Ontologies, co-located with SAMT 2007, Genova, Italy, December 2007.
- [187] Ph. Mylonas, E. Spyrou and Y. Avrithis, *High-Level Concept Detection based on Mid-level Semantic Information and Contextual Adaptation*, 2nd International Workshop on Semantic Media Adaptation and Personalization (SMAP 2007), London, United Kingdom, 17-18 December 2007.
- [188] Ph. Mylonas, E. Spyrou and Y. Avrithis, *Using Visual Context and Region Semantics for High-Level Concept Detection*, Multimedia Tools and Applications, Springer, 2008, accepted for publication.
- [189] F. Nack, A. Lindsay, *Everything You Wanted to Know About MPEG-7: Part 1*, IEEE Multimedia 6(3):65–77, 1999.
- [190] F. Nack, A. Lindsay, *Everything You Wanted to Know About MPEG-7: Part 2*, IEEE Multimedia 6(4):64-73, 1999.
- [191] M. Naphade and T. S. Huang, *A factor graph framework for semantic indexing and retrieval in video*, CVPR Workshop on Content-based Image and Video Retrieval, 2000.
- [192] R.M. Naphade and T.S. Huang, *A probabilistic framework for semantic video indexing, filtering, and retrieval*, IEEE Trans. on Multimedia, Vol. 3, No. 1, March 2001.
- [193] M. R. Naphade and J. R. Smith, *A hybrid framework for detecting the semantics of concepts and context*, In Proc. of CIVR 2003, pages 196–205, 2003.
- [194] E. Nuutila, *Efficient Transitive Closure Computation in Large Digraphs*, Acta Polytechnica Scandinavica, Mathematics and Computing in Engineering Series No. 74, Helsinki, 124 pages, 1995.
- [195] M. O'Connor, D. Cosley, J. A. Konstan and J. Riedl, *PolyLens: A Recommender System for Groups of Users*, Proceedings of the 7th European Conference on Computer Supported Cooperative Work (pp. 199-218). Bonn, Germany, 2001.
- [196] A. Oliva, A. Torralba, *Modeling the shape of the scene: A holistic representation of the spatial envelope*, Int. J. Comp. Vis., vol. 42, pp. 145-175, 2001.

- [197] A. Opelt, A. Pinz and A. Zisserman, *Incremental learning of object detectors using a visual shape alphabet*, In Proc. of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '06), Washington, DC, USA, 2006.
- [198] W. Osberger, A.J. Maeder, *Automatic Identification of Perceptually Important Regions in an Image*, in: Proceedings of IEEE International Conference on Pattern Recognition, 1998.
- [199] S.E. Palmer, *The effects of contextual scenes on the identification of objects*, Memory and Cogn., vol. 3, pp. 519-526, 1975.
- [200] K. Petridis, S. Bloehdorn, C. Saathoff, N. Simou, S. Dasiopoulou, V. Tzouvaras, S. Handschuh, Y. Avrithis, I. Kompatsiaris and S. Staab, *Knowledge Representation and Semantic Annotation of Multimedia Content*, IEE Proceedings on Vision Image and Signal Processing, Special issue on Knowledge-Based Digital Media Processing, Vol. 153, No. 3, pp. 255-262, June 2006.
- [201] V. A. Petrushin, L. Khan (Eds.), *Multimedia Data Mining and Knowledge Discovery*, 2007, XXVI.
- [202] B. Popov, A. Kiryakov, D. Ognyanoff, D. Manov and A. Kirilov, *KIM - A Semantic Platform for Information Extraction and Retrieval*, Journal of Natural Language Engineering vol. 10, no. 3-4, 2004, pp. 375-392.
- [203] P. Purdom, *A transitive closure algorithm*, BIT 10:76-94, 1970.
- [204] B. Rajagopalan and A. Deshmukh, *Evaluation of Online Personalization Systems: A Survey of Evaluation Schemes and A Knowledge-Based Approach*, Journal of Electronic Commerce Research vol. 6, no. 2, pp. 112-122, May 2005.
- [205] K. Rapantzikos, Y. Avrithis, S. Kollias, *On the use of spatiotemporal visual attention for video classification*, Proceedings of International Workshop on Very Low Bitrate Video Coding (VLBV '05), Sardinia, Italy, September 2005.
- [206] R.A. Rensink, J.K. O'Regan, J.J. Clark, *To see or not to see: The need for attention to perceive changes in scenes*, Psychol. Sci., vol. 8, pp. 368-373, 1997.
- [207] J. Rocchio, *Relevance feedback in Information Retrieval*, The SMART system - experiments in automatic document processing, Prentice Hall, 1971.
- [208] J. Rocchio, *Relevance feedback information retrieval*, in Salton, G. (Ed.), The Smart Retrieval System-Experiments in Automatic Document Processing, Prentice-Hall, Kansas City, MO, 1971, pp. 313-323.
- [209] U. Rutishauser, D. Walther, C. Koch, P. Perona, *Is attention useful for object recognition?*, ICCVPR, 2004
- [210] E. Saber, A.M. Tekalp, R. Eschbach, and K. Knox, *Automatic image annotation using adaptive colour classification*, CVGIP: Graphical Models and Image Processing, vol. 58, pp. 115-126, 1996.

- [211] M. Sahami, S. Yusufali and M. Q. Wang Baldonado, *Real-Time Full-Text Clustering of Networked Documents*, AAAI/IAAI, pages 845, 1997.
- [212] P. Salembier, J. R. Smith, *MPEG-7 Multimedia Description Schemes*, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 11, No. 6, 2001.
- [213] G. Salton and M. McGill, *Introduction to Modern Information Retrieval*, McGraw-Hill, New York, 1983.
- [214] G. Salton and M. McGill, *Introduction to Modern Information Retrieval*, McGraw-Hill, Inc., 1986.
- [215] E. Sanchez, *Fuzzy Logic and the Semantic Web*, Elsevier Science Inc., New York, NY, USA, 2006.
- [216] L. Sanghoon, M.M. Crawford, *Unsupervised classification using spatial region growing segmentation and fuzzy training*, In Proc. of the IEEE International Conference on Image Processing, Thessaloniki, Greece, 2001.
- [217] S. Santini, *Exploratory Image Databases: Content-based Retrieval*, Academic, New York, 2001.
- [218] B. Saux and G. Amato, *Image classifiers for scene analysis*, In Proc. of International Conference on Computer Vision and Graphics, 2004.
- [219] B. Schilit, N. Adams and R. Want, *Context-Aware Computing Applications*, in Proceedings of the 1st IEEE Workshop on Mobile Computing Systems and Applications, Santa Cruz, CA, US, December, 1994.
- [220] H. Schütze and C. Silverstein, *Projections for efficient document clustering*, SIGIR '97: Proceedings of the 20th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 74-81, Philadelphia, Pennsylvania, United States, ACM Press, New York, NY, USA, 1997.
- [221] P.G. Schyns, A. Oliva, *From blobs to boundary edges: evidence for time and spatial scale dependent scene recognition*, Psychol. Sci., vol. 5, pp. 195-200, 1994.
- [222] F. Sebastiani, *Machine learning in automated text categorization*, ACM Comput Surv 34(1):1-47, 2002.
- [223] R. Seidel, *On the All-Pairs-Shortest-Path problem in unweighted undirected graphs*, J. Computer & System Sciences 51(3):400-403, 1995.
- [224] X. Shen, B. Tan, C. Zhai, *Context-sensitive information retrieval using implicit feedback*, in Proc. of the 28<sup>th</sup> annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR 2005), pp. 43-50, 2005.
- [225] J. Shi, J. Malik, *Normalized cuts and image segmentation*, IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 8, pp. 888-905, 2000.

- [226] N. Simou, V. Tzouvaras, Y. Avrithis, G. Stamou and S. Kollias, *A Visual Descriptor Ontology for Multimedia Reasoning*, Proc. of Workshop on Image Analysis for Mul-timedia Interactive Services (WIAMIS '05), Montreux, Switzerland, April 13-15, 2005.
- [227] E. Simoudis, *Reality check for data mining*, IEEE Expert, 11(5), 1992.
- [228] A. Singhal, J. Luo and W. Zhu, *Probabilistic spatial context models for scene content understanding*, in Proc. IEEE Int. Conf. Computer Vision Pattern Recognition (CVPR), Madison, WI, pp. 235-241, 2003.
- [229] A. Singhal and J. Luo, *Hybrid approach to classifying sky regions in natural images*, Proc. SPIE Conference on Electronic Imaging, 2003.
- [230] J. Sivic, *Efficient visual search of images and videos*, PhD thesis, University of Oxford, 2006.
- [231] A. F. Smeaton, P. Over and W. Kraaij, *Evaluation campaigns and TRECVID*, In Proceedings of the 8th ACM International Workshop on Multimedia Information Retrieval, Santa Barbara, California, USA, October 26 - 27, 2006.
- [232] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, *Content-based image retrieval at the end of the early years*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(12), pp. 1349 - 1380, 2000.
- [233] J. R. Smith and C.-S. Li, *Decoding image semantics using composite region templates*, in Proc. IEEE Int. Workshop on Content-based Access of Image and Video Database, 1998.
- [234] C. Snoek, M. Worring, D. Koelma and A. Smeulders, *A Learned Lexicon-Driven Paradigm for Interactive Video Retrieval*, IEEE Trans. on Multimedia, 9(2): 280-292, 2007.
- [235] F. Souvannavong, B. Merialdo and B. Huet, *Region-based video content indexing and retrieval*, In Proc. of 4th International Workshop on Content-Based Multimedia Indexing, Riga, Latvia, 2005.
- [236] F. Souvannavong, L. Hohl, B. Mérialdo and B. Huet, *Structurally Enhanced Latent Semantic Analysis for Video Object Retrieval*, Special Issue of the IEE Proceedings on Vision, Image and Signal Processing , Vol. 152(6), pp. 859-867, 2005.
- [237] E. Spyrou, H. Le Borgne, T. Mailis, E. Cooke, Y. Avrithis, and N. O'Connor, *Fusing MPEG-7 visual descriptors for image classification*, In Proc. of International Conference on Artificial Neural Networks ICANN 05, Warsaw, Poland, September 2005.
- [238] E. Spyrou, G. Koumoulos, Y. Avrithis and S. Kollias, *Using Local Region Semantics for Concept Detection in Video*, 1st International Conference on Semantics And digital Media Technology (SAMT 2006), Athens, Greece, 2006.

- [239] E. Spyrou and Y. Avrithis, *A Region Thesaurus Approach for High-Level Concept Detection in the Natural Disaster Domain*, In Proc of the 2nd International Conference on Semantics And digital Media Technologies (SAMT), 2007.
- [240] E. Spyrou, P. Kapsalas, G. Tolias, Ph. Mylonas, Y. Avrithis et al., *The COST292 experimental framework for TRECVID 2007*, 5th TRECVID Workshop, Gaithersburg, USA, November 2007.
- [241] E. Spyrou, Ph. Mylonas and Y. Avrithis, *Semantic Multimedia Analysis based on Region Types and Visual Context*, 4th IFIP Conference on Artificial Intelligence Applications & Innovations (AIAI), Athens, Greece, 19-21 September 2007.
- [242] E. Spyrou, G. Tolias, Ph. Mylonas and Y. Avrithis, *A Semantic Multimedia Analysis Approach Utilizing a Region Thesaurus and LSA*, 9th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2008), Klagenfurt, Austria, 7-9 May 2008.
- [243] S. Staab and R. Studer, *Handbook on ontologies, international handbooks on information systems*, Heidelberg: Springer-Verlag, 2004.
- [244] G. Stoilos, G. Stamou, V. Tzouvaras, J.Z. Pan, I. Horrocks, *Fuzzy OWL: Uncertainty and the Semantic Web*, International Workshop of OWL: Experiences and Directions, Galway, 2005.
- [245] G. Stoilos, G. Stamou, V. Tzouvaras, J.Z. Pan and I. Horrocks, *The Fuzzy Description Logic f-SHIN*, International Workshop on Uncertainty Reasoning For the Semantic Web (2005), 2005.
- [246] U. Straccia, *Reasoning within Fuzzy Description Logics*, Journal of Artificial Intelligence Research, volume 14, pages 137-166, 2001.
- [247] T.F. Syeda-Mahmood, *Data and model-driven selection using colour regions*, Int. J. Comput. Vision 21(1), pp. 9-36, 1997.
- [248] Y. Sun, H. Zhang, L. Zhan and M. Li, *MyPhotos - A System for Home Photo Management and Processing*, ACM Multimedia Conf. (ACMMM 2002), Juan-les-Pins, France, Dec. 1-6 2002.
- [249] M. Sushmita, A. Tinku, *Data Mining: Multimedia, Soft Computing, and Bioinformatics*, October 2003.
- [250] R. Tansley, C. Bird, W. Hall, P. Lewis, and M. Weal, *Automating the linking of content and concept*, In Proc. ACM Int. Multimedia Conf. and Exhibition (ACM MM-2000), Oct./Nov. 2000.
- [251] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, Academic Press, 1998.
- [252] L.-E. Thorelli, *An algorithm for computing all paths in a graph*, BIT 6:347-349, 1966.

- [253] S. Thorpe, D. Fize, C. Marlot, *Speed of processing in the human visual system.* Nature, vol. 381, pp. 520-522, 1996.
- [254] A. Torralba and P. Sinha, *Statistical context priming for object detection*, Proc. International Conference on Computer Vision, 2001.
- [255] A. Torralba, *Contextual priming for object detection*, Int. J. Comp. Vis., vol. 53, pp. 169-191, 2003.
- [256] A. Torralba, *Modelling global scene factors in attention*, vol. 20, no. 7, pp. 1407-1418, Jul 2003.
- [257] A. Torralba, *Contextual influences on saliency*, *Neurobiology of attention*, Eds. L. Itti, G. Rees, J. Tsotsos, Academic Press Inc. (London) Ltd, 2005.
- [258] N. Tsapatsoulis, M. Wallace and S. Kasderidis, *Improving the Performance of Re-source Allocation Networks through Hierarchical Clustering of High Dimensional Data*, Proceedings of the International Conference on Artificial Neural Networks (ICANN), Istan-bul, Turkey 2003.
- [259] G. Tsechpenakis, G. Akrivas, G. Andreou, G. Stamou and S. Kollias, *Knowledge-Assisted Video Analysis and Object Detection*, Proceedings of European Symposium on Intelligent Technologies, Hybrid Systems and their Implementation on Smart Adaptive Systems (Eunite02), Albufeira, Portugal, September 2002.
- [260] B.L. Tseng, L. Ching-Yung, J.R. Smith, *Using MPEG-7 and MPEG-21 for personalizing video*, Multimedia, IEEE Volume 11, Issue 1, Page(s):42 - 52, Jan-March 2004.
- [261] C. Tsinaraki, P. Polydoros, S. Christodoulakis, *Interoperability support for Ontology-based Video Retrieval Applications*, CIVR 2004.
- [262] A. Vailaya and A. Jain, *Detecting sky and vegetation in outdoor images*, Proc. SPIE, vol. 3972, January 2000.
- [263] D. Vallet, Ph. Mylonas, M. A. Corella, J. M. Fuentes, P. Castells and Y. Avrithis, *A Semantically-Enhanced Personalization Framework for Knowledge-Driven Media Services*, Proc. of IADIS International Conference on WWW / Internet (ICWI '05), Lisbon, Portugal, October 19-22, 2005.
- [264] J. van Eijck, *The proper treatment of context in NL*, In Monachesi, P. (Ed.), Computational Linguistics in the Netherlands 1999. Selected Papers form the 10<sup>th</sup> CLIN Meeting. Utrecht Institute of Linguistics OTS, 2000.
- [265] C. J. van Rijsbergen, *Information Retrieval*, Butterworths and Co, 1979.
- [266] N. Voisine, S. Dasiopoulou, V. Mezaris, E. Spyrou, Th. Athanasiadis, I. Kompatziaris, Y. Avrithis and M. G. Strintzis, *Knowledge-assisted video analysis using a genetic algorithm*, In Proc. of 6th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2005), April 13-15, 2005.

- [267] M. Wallace, G. Stamou, *Towards a Context Aware Mining of User Interests for Consumption of Multimedia Documents*, Proceedings of the IEEE International Conference on Multimedia and Expo (ICME), Lausanne, Switzerland, 2002.
- [268] M. Wallace, G. Akrivas, Ph. Mylonas, Y. Avrithis, S. Kollias, *Using context and fuzzy relations to interpret multimedia content*, Proceedings of the 3rd International Workshop on Content-Based Multimedia Indexing (CBMI), IRISA, Rennes, France, September 2003.
- [269] M. Wallace, K. Karpouzis, G. Stamou, G. Moschovitis, C. Schizas and S. Kollias, *The Electronic Road: Personalized Content Browsing*, IEEE Multimedia, Multimedia Content Modeling and Personalization, Vol. 10, No.4, p. 49-59, 2003.
- [270] M. Wallace and S. Kollias, *Soft Attribute Selection for Hierarchical Clustering in High Dimensions*, Proceedings of the International Fuzzy Systems Association World Congress(IFSA), Istanbul, Turkey, June-July 2003.
- [271] D. Walther, U. Rutishauser, C. Koch, P. Perona, *On the usefulness of attention for object recognition*, ECCV, pp. 96-103, 2004.
- [272] J.Z. Wang, J. Li, G. Wiederhold, *SIMPLICITY: Semantic-sensitive Integrated Matching for Picture Libraries*, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 23, No.9, 947-963, 2001.
- [273] H.S. Warren, *A modification of Warshall's algorithm for the transitive closure of binary relations*, Comm. ACM 18(4):218-220, 1975.
- [274] S. Warshall, *A theorem on Boolean matrices*, J. ACM 9(1):11-12, 1962.
- [275] S. Weiss and N. Indurkhy, *Predictive data mining: a practical guide*, Morgan Kaufmann, 1998.
- [276] R. W. White, J. M. Jose, C. J. van Rijsbergen and I. Ruthven, *A simulated study of implicit feedback models*, in Proc. of the European Conference on Information retrieval (ECIR 2004), pp. 311-326, 2004.
- [277] J. Wielemaker, A.Th. Schreiber, B. Dubbeldam and B.J. Wielinga, *Ontology-based photo annotation*, IEEE Intelligent Systems, May/June 2001.
- [278] P. Willett, *Recent trends in hierachic document clustering: a critical review*, Inf Process Manag 24(5):577-597, 1988.
- [279] R. Wilkinson and M. Wu, *Evaluation Experiments and Experience from the Perspective of Interactive Information Retrieval*, Proc. of the 3<sup>rd</sup> Workshop on Empirical Evaluation of Adaptive Systems, in conjunction with the 2<sup>nd</sup> International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems, Eindhoven, The Netherlands, pp. 221-230, August 2004.
- [280] I. Witten and E. Frank, *Data mining: practical machine learning tools and techniques with Java implementations*, Morgan Kaufmann, 2000.

- [281] R.R. Yager, *Intelligent control of the hierarchical agglomerative clustering process*, IEEE Transactions on Systems, Man and Cybernetics, Part B 30(6): 835-845, 2000.
- [282] R. Yan, M.-Y. Chen and A. Hauptmann, *Mining relationship between video concepts using probabilistic graphical models*, In Proc. of ICME 2006, pages 301-304, 2006.
- [283] C. Yang, M. Dong and F. Fotouhi, *Region based image annotation through multipleinstance learning*, In Proc. of ACM Multimedia 2005, pages 435-438, 2005.
- [284] A. Yoshitaka, S. Kishida, M. Hirakawa, and T. Ichikawa, *Knowledge-assisted content-based retrieval for multimedia databases*, IEEE Multimedia, 1(4):12-21, Winter 1994.
- [285] J. Yuan, J. Li and B. Zhang, *Exploiting spatial context constraints for automatic image region annotation*, In Proceedings of the 15th international conference on Multimedia, pp. 595-604, Augsburg, Germany, 2007.
- [286] L.A. Zadeh, *Fuzzy Sets*, Information and Control 8:228-353, 1965.
- [287] R. Zhao, W.I. Grosky, *Narrowing the Semantic Gap-Improved Text-Based Web Document Retrieval Using Visual Features*, IEEE Transactions on Multimedia 4(2), 2002.
- [288] J. Zhao, Y. Shimazu, K. Ohta, R. Hayasaka, Y. Matsushita, *An Outstandingness Oriented Image Segmentation and its Applications*, in: Proceedings of the International Symposium on Signal Processing and its Applications, 1996.
- [289] DARPA Agent Markup Language Homepage <http://www.daml.org/>
- [290] EXIF, JEITA CP-3451, Exchangeable image file format for digital still cameras: Exif Version 2.2, Japan Electronics and Information Technology Industries Association, April 2002.
- [291] FaceBook, <http://www.facebook.com/>
- [292] Flickr, <http://www.flickr.com/>
- [293] Flixster, <http://www.flixster.com/>
- [294] IEEE LTSC, P1484.1 Architecture and Reference Model, <http://ltsc.ieee.org/index.html>
- [295] ISO/IEC, 21000-7:2004, Information Technology-Multimedia Framework (MPEG-21)Part 7: Digital Item Adaptation, 2004.
- [296] ISO/IEC FDIS 15938-5, *Information Technology - Multimedia Content Description Interface: Multimedia Description Schemes*, ISO/IEC JTC1/SC29/WG11/M4242, pp. 442-448, October 2001.
- [297] ISO/IEC FDIS 15938-5, ISO/IEC JTC 1/SC 29 M 4242, *IT v MCDI v Part 5: MDS*, pp. 442-448, October 2001.

- [298] ISO/IEC JTC1/SC29/WG1 N1646R (2000) JPEG 2000 Part I. Final Committee Draft Version 1.0.
- [299] ISO/IEC JTC1/SC29/WG11, “MPEG-7 Context, Objectives and Technical Roadmap, (v.12)”, Doc. N2861, July 1999.
- [300] ISO/IEC JTC1/SC29/WG11 (2000) Text of ISO/IEC CD 15938-2 Information technology – Multimedia content description interface – Part 2: Description definition language.
- [301] Java, <http://java.sun.com/>
- [302] KAON, <http://kaon.semanticweb.org/>
- [303] LinkedIn, <http://www.linkedin.com/>
- [304] Merriam-Webster Online Dictionary <http://www.m-w.com/>
- [305] MPEG-7 Visual Experimentation Model (XM) (2001), Version 10.0, ISO/IEC/JTC1/SC29/WG11, Doc. N4062.
- [306] MySpace, <http://www.myspace.com/>
- [307] Plaxo, <http://www.plaxo.com/>
- [308] Principal components analysis. <http://obelia.jde.aca.mmu.ac.uk/multivar/pca.htm>
- [309] SlideShare, <http://www.slideshare.net/>
- [310] TeleInformatics System for Continuous Collection, Processing, Diffusion of Material for Teacher Training in Special Education, <http://www.image.ntua.gr/spero/>
- [311] UCI Machine Learning Repository, <http://www.ics.uci.edu/~mlearn/MLRepository.html>
- [312] XML 1.0 (Fourth Edition), W3C Recommendation, 16 August 2006, Tim Bray, Jean Paoli, C. M. Sperberg-McQueen, Eve Maler, Franc,ois Yergeau eds. <http://www.w3.org/XML>
- [313] XML Schema Part 0: Primer, W3C Working Draft, Sept 2000 <http://www.w3.org/TR/xmlschema-0>
- [314] W3C, <http://www.w3.org/>
- [315] W3C, <http://ontoedit.com/>
- [316] W3C, RDF, <http://www.w3.org/RDF/>
- [317] W3C, RDF, <http://www.w3.org/TR/rdf-schema/>
- [318] W3C, RDF Reification, [http://www.w3.org/TR/rdf-schema/#ch\\_reificationvocab](http://www.w3.org/TR/rdf-schema/#ch_reificationvocab)

[319] W3C Recommendation, OWL Web Ontology Language Reference, <http://www.w3.org/TR/owl-ref/>, 10 February 2004

[320] Youtube, <http://www.youtube.com/>





# Βιογραφικό Σημείωμα

Ο Φοίβος-Απόστολος Μυλωνάς γεννήθηκε στις 17 Ιανουαρίου 1978 στην Αθήνα και επί του παρόντος είναι Ερευνητής στο Εργαστήριο Ψηφιακής Επεξεργασίας Εικόνας, Βίντεο και Συστημάτων Πολυμέσων της Σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Η/Υ του Εθνικού Μετσόβιου Πολυτεχνείου. Είναι Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών από το Εθνικό Μετσόβιο Πολυτεχνείο (2001), καθώς και κάτοχος Διπλώματος Ειδίκευσης (Master) στα Προηγμένα Πληροφοριακά Συστήματα από το Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών (2003). Από τον Σεπτέμβριο του 2001 συνεργάζεται με το Εργαστήριο Ψηφιακής Επεξεργασίας Εικόνας Βίντεο και Συστημάτων Πολυμέσων της Σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Η/Υ. Από τον Οκτώβριο του 2002 εργάζεται υπό την επίβλεψη του Καθηγητή Ε.Μ.Π. κ. Στέφανου Κόλλια, με στόχο την απόκτηση διδακτορικού τίτλου στην περιοχή της προσαρμοζόμενης πρόσβασης σε πολυμεσικό περιεχόμενο με βάση το εννοιολογικό πλαίσιο και το προφίλ του χρήστη.

Ανάμεσα στα ερευνητικά του ενδιαφέροντα βρίσκονται η ανάκτηση πληροφορίας με βάση το περιεχόμενο, η αναπαράσταση και ανάλυση του εννοιολογικού πλαισίου, η εξόρυξη δεδομένων, η καθοδηγούμενη από τη γνώση ανάλυση πολυμέσων, ζητήματα που άπτονται θεμάτων προσωποποίησης περιεχομένου, προσαρμογής χρηστών, μοντελοποίησης και εξαγωγής προφίλ χρηστών με χρήση ασαφών, βασισμένων σε οντολογίες παραμέτρων γνώσης. Έχει δημοσιεύσει εργασίες του σε 20 διεθνή περιοδικά και κεφάλαια βιβλίων, έχει συγγράψει 30 ερευνητικά άρθρα τα οποία και δημοσιεύθηκαν σε πρακτικά διεθνών συνεδρίων και ημερίδων, είναι κριτής των διεθνών περιοδικών “Signal Processing”, “Neurocomputing”, “Multimedia Tools and Applications”, “IEEE Transactions on Circuits and Systems for Video Technology”, “Journal of Networks” και “Journal of Visual Communication and Image Representation” και εώς σήμερα έχει εμπλακεί ενεργά στη διοργάνωση 11 διεθνών συνεδρίων. Είναι μέλος της IEEE από το 1999, της ACM από το 2001, του Τεχνικού Επιμελητηρίου Ελλάδας από το 2001 και της Ελληνικής Ένωσης Μηχανολόγων και Ηλεκτρολόγων Μηχανικών από το 2002.

Στο πλαίσιο των ερευνητικών δραστηριοτήτων του συγκαταλέγεται η έκδοση 8 βιβλίων και περιοδικών. Ειδικότερα:

- Guest Editor
  - Springer-Verlag: Semantic Media Adaptation and Personalization  
M. Angelides, P. Mylonas, M. Wallace (eds.)  
**ACM/Springer Multimedia Systems Magazine**  
Volume 13, Number 2 / August, 2007 [ISSN: 0942-4962]
  - Springer-Verlag: Special Issue on Semantic Media Adaptation and Personalization  
M. Angelides, P. Mylonas, M. Wallace (eds.)  
**Multimedia Tools and Applications**  
December 2008 [ISSN: 1380-7501]
  - Springer-Verlag: Special Issue on Multimedia Semantics, Adaptation & Personalization  
M. Angelides, P. Mylonas, M. Wallace (eds.)  
**Signal, Image and Video Processing**  
December 2008 [ISSN: 1863-1703]
- Edited Books
  - IEEE CRC Press: Advances in Semantic Media Adaptation and Personalization, Volume 2  
M. Wallace, M. Angelides, P. Mylonas (eds.)
  - Springer-Verlag: Advances in Semantic Media Adaptation and Personalization  
M. Wallace, M. Angelides, P. Mylonas (eds.)  
**Studies in Computational Intelligence**, Vol. 93
  - IEEE Computer Society: Semantic Media Adaptation & Personalization  
P. Mylonas, M. Wallace, M. Angelides (eds.)  
SMAP 2008 IEEE Proceedings
  - IEEE Computer Society: Semantic Media Adaptation & Personalization  
P. Mylonas, M. Wallace, M. Angelides (eds.)  
SMAP 2007 IEEE Proceedings
  - IEEE Computer Society: Semantic Media Adaptation & Personalization  
P. Mylonas, M. Wallace, M. Angelides (eds.)  
SMAP 2006 IEEE Proceedings

Ο κ. Μυλωνάς συμμετέχει ως Κριτής (Reviewer) στα ακόλουθα Διεθνή Περιοδικά:

- Eurasip - **Signal Processing**, [ISSN: 0165-1684]
- Elsevier - **Neurocomputing**, [ISSN: 0925-2312]
- Academy Publisher - **Journal of Networks**, [ISSN: 1796-2056]
- IEEE - **IEEE Transactions on Circuits and Systems for Video Technology**, [ISSN: 1051-8215]

- Elsevier - **Journal of Visual Communication and Image Representation**, [ISSN: 1047-3203]
- Springer - **Multimedia Tools and Applications**, [ISSN: 1380-7501]

Κατά τη διάρκεια εκπόνησης της διατριβής του, ο κ. Μυλωνάς εργάστηκε ως ερευνητής σε 7 ερευνητικά και αναπτυξιακά έργα:

- **WeKnowIt** - Emerging, Collective Intelligence for personal, organisational and social use, 2008-2011
- **MESH** - Multimedia Semantic Syndication for Enhanced News Services, 2005-2008
- **COST292** - Semantic Multimodal Analysis of Digital Media, 2005-2008
- **visualAsset** - Multimedia document analysis, information extraction and management system, GSRT-EPAN, 2004-2006
- **aceMedia** - Integrating Knowledge, Semantics and Content for User-Centred Intelligent Media Services, FP6-001765, 2004-2007
- **SPERO** - TeleInformatics System for Continuous Collection, Processing, Diffusion of Material for Teacher Training in Special Education, Leonardo, 2002-2004
- **FAETHON** - Unified Intelligent Access to Heterogeneous Audiovisual Content, IST , 2001-2003

ενώ υποστήριξε εργαστηριακά μαθήματα του Τομέα Τεχνολογίας Πληροφορικής και Υπολογιστών υπό την επίβλεψη του Καθηγητή Ε.Μ.Π. Σ. Κόλλια:

- Γραφική με Υπολογιστές [2007, 2006, 2005, 2004, 2003, 2002]
- Ψηφιακή Επεξεργασία Εικόνας και Ανάλυση Video [2007, 2006, 2005, 2004, 2003, 2002]

Ο κ. Μυλωνάς ενεπλάκη έως σήμερα ενεργά στη διοργάνωση 11 συνεδρίων και ημερίδων:

- 3rd Int. Workshop on Semantic Media Adaptation & Personalization  
General Chair, Prague, Czech Republic, December 15-16, 2008
- 1st International Workshop on Collective Semantics: Collective Intelligence & the Semantic Web (CISWeb 2008), in conjunction with the 5th European Semantic Web Conference (ESWC-08)  
Program Committee member, Tenerife, Spain, June 1, 2008
- 2nd Int. Workshop on Semantic Media Adaptation and Personalization  
General Chair, London, United Kingdom, December 17-18, 2007
- 1st Workshop on Multimedia Annotation and Retrieval enabled by Shared Ontologies (MAReSO), in conjunction with the 2nd International Conference on Semantics And digital Media Technologies (SAMT2007)  
Program Committee member, Genova, Italy, December 5, 2007
- Int. Workshop on Ubiquitous and Decentralized User Modeling, in conjunction with the 11th International Conference on User Modeling (UM 2007)  
Workshop Organizer, Corfu, Greece, June 26, 2007

- 1st International Conference on Semantic and Digital Media Technologies  
Special Session Organizer & Local Arrangements, Athens, Greece, December 6-8, 2006
- 1st Int. Workshop on Semantic Media Adaptation and Personalization  
General Chair, Athens, Greece, December 4-5, 2006
- Closing Workshop of Research and Technology Training Network on Multimedia Knowledge Discovery and Management (MULTI-MINE '05)  
Local Arrangements, Athens, Greece, June 28, 2005
- SPERO Workshop: TeleInformatics System for Continuous Collection, Processing, Diffusion of Material for Teacher Training in Special Education  
Local Arrangements, Athens, Greece, November 5, 2004
- Workshop on Intelligent Access to Audiovisual Content (WIAAC)  
Local Arrangements, Athens, Greece, April 3, 2003.
- International Workshop on Very Low Bitrate Video Coding (VLBV '01)  
Local Arrangements, Athens, Greece, October 11-12, 2001.

Τέλος, η ερευνητική του πορεία περιλαμβάνει ένα **Βραβείο από το Θωμαείδιο Ίδρυμα για την πρόοδο των τεχνών και των επιστημών (2005)** και μία μικρής έκτασης **Υποτροφία Έρευνας** από το Ινστιτούτο Επικοινωνίας και Υπολογιστικών Συστημάτων του Εθνικού Μετσόβιου Πολυτεχνείου (2005).





# Παράρτημα Α

□

TABLE 3.2 SOME CLASSES OF FUZZY INTERSECTIONS ( $t$ -NORMS)

Reference	Formula $i(a, b)$	Decreasing generator $f(a)$	Parameter range	As parameter converges to 0	As parameter converges to 1 or $-1$	As parameter converges to $\infty$ or $-\infty$
Dombi [1982]	$\left\{ 1 + \left[ \left( \frac{1}{a} - 1 \right)^{\lambda} + \left( \frac{1}{b} - 1 \right)^{\lambda} \right]^{\frac{1}{\lambda}} \right\}^{-1}$	$\left( \frac{1}{a} - 1 \right)^{\lambda}$	$\lambda > 0$	$i_{\min}(a, b)$	$\frac{ab}{a + b - ab}$ when $\lambda = 1$	$\min(a, b)$
Frank [1979]	$\log_s \left[ 1 + \left( \frac{s^a - 1}{s - 1} \right) \left( s^b - 1 \right) \right]$	$-\ln \left( \frac{s^a - 1}{s - 1} \right)$	$s > 0, s \neq 1$	$\min(a, b)$	$\max(0, a + b - 1)$	
Hannacher [1978]	$\frac{ab}{r + (1 - r)(a + b - ab)}$	$-\ln \left( \frac{a}{r + (1 - r)a} \right)$	$r > 0$	$\frac{ab}{a + b - ab}$	$ab$ when $r = 1$	$i_{\min}(a, b)$
Schweizer & Sklar 1 [1963]	$\{\max(0, a^p + b^p - 1)\}^{\frac{1}{p}}$	$1 - a^p$	$p \neq 0$	$ab$	$\max(0, a + b - 1)$ , when $p = 1$ ; $\frac{ab}{a + b - ab}$ , when $p \rightarrow \infty$ ;	$i_{\min}(a, b)$ as $p \rightarrow \infty$ ;
Schweizer & Sklar 2	$1 - [(1 - a)^p + (1 - b)^p]^{1/p}$ $-(1 - a)^p(1 - b)^p]^{1/p}$	$\ln[1 - (1 - a)^p]^{1/p}$	$p > 0$	$i_{\min}(a, b)$	$ab$ when $p = 1$	$\min(a, b)$
Schweizer & Sklar 3	$\exp(-(\ln a)^p + (\ln b)^p)^{\frac{1}{p}}$	$ \ln a ^p$	$p > 0$	$i_{\min}(a, b)$	$ab$ when $p = 1$	$\min(a, b)$
Schweizer & Sklar 4	$\frac{ab}{[a^p + b^p - a^p b^p]^{\frac{1}{p}}}$	$a^{-p} - 1$	$p > 0$	$ab$	$\frac{ab}{a + b - ab}$ , when $p = 1$	$\min(a, b)$
Yager [1980]	$1 - \min \left\{ 1, [(1 - a)^w + (1 - b)^w]^{\frac{1}{w}} \right\}$	$(1 - a)^w$	$w > 0$	$i_{\min}(a, b)$	$\max(0, a + b - 1)$ when $w = 1$	$\min(a, b)$
Dubois & Prade [1980]	$\frac{ab}{\max(a, b, \alpha)}$		$\alpha \in [0, 1]$	$\min(a, b)$	$ab$ when $\alpha = 1$	
Weber [1983]	$\max \left( 0, \frac{a + b + \lambda ab - 1}{1 + \lambda} \right)$	$\frac{1}{\lambda} \ln[1 + \lambda(1 - a)]$	$\lambda > -1$	$\max(0, a + b - 1)$	$i_{\min}(a, b)$ as $\lambda \rightarrow -1$ ; $\max[0, (a + b + ab - 1)/2]$ when $\lambda = 1$ .	$ab$
Yu [1985]	$\max[0, (1 + \lambda)(a + b - 1) - \lambda ab]$	$\frac{1}{\lambda} \ln \frac{1 + \lambda}{1 + \lambda a}$	$\lambda > -1$	$\max(0, a + b - 1)$	$\max[0, 2(a + b - ab/2 - 1)]$ when $\lambda = 1$ .	$i_{\min}(a, b)$

TABLE 3.3 SOME CLASSES OF FUZZY UNIONS ( $\mu$ -CONORMS)

Reference	Formula $u(a, b)$	Increasing generator $g(a)$	Parameter range	As parameter converges to 0	As parameter converges to 1 or $-1$	As parameter converges to $\infty$ or $-\infty$
Dombi [1982]	$\left\{1 + \left[\left(\frac{1}{a} - 1\right)^\lambda + \left(\frac{1}{b} - 1\right)^\lambda\right]^{-\frac{1}{\lambda}}\right\}^{-1}$	$\left(\frac{1}{a} - 1\right)^{-\lambda}$	$\lambda > 0$	$u_{\max}(a, b)$	$\frac{a + b - 2ab}{1 - ab}$ when $\lambda = 1$	$\max(a, b)$
Frank [1979]	$1 - \log_s \left[1 + \frac{(s^{1-a}-1)(s^{1-b}-1)}{s-1}\right]^{-\frac{1}{\lambda}}$	$-\ln \left(\frac{s^{1-a}-1}{s-1}\right)$	$s > 0, s \neq 1$	$\max(a, b)$	$a + b - ab$ as $s \rightarrow 1$	$\min(1, a + b)$
Hamacher [1978]	$\frac{a+b+(r-2)ab}{r+(r-1)ab}$	$-\ln \left(\frac{1-a}{r+(1-r)(1-a)}\right)$	$r > 0$	$\frac{a+b-2ab}{1-ab}$	$a + b - ab$ when $r = 1$	$u_{\max}(a, b)$
Schweizer & Sklar 1 [1963]	$1 - \{\max(0, (1-a)^p + (1-b)^p - 1)\}^{\frac{1}{p}}$	$1 - (1-a)^p$	$p \neq 0$	$a + b - ab$	$\frac{a+b-2ab}{1-ab}$ , when $p = -1$	$u_{\max}(a, b)$ as $p \rightarrow \infty$ ; $u_{\min}(a, b)$ as $p \rightarrow -\infty$ .
Schweizer & Sklar 2	$[a^p + b^p - a^p b^p]^{\frac{1}{p}}$	$\ln[1 - a^p]^{\frac{1}{p}}$	$p > 0$	$u_{\max}(a, b)$	$a + b - ab$ when $p = 1$	$u_{\max}(a, b)$
Schweizer & Sklar 3	$1 - \exp(-( \ln(1-a) ^p +  \ln(1-b) ^p)^{\frac{1}{p}})$	$ \ln(1-a) ^p$	$p > 0$	$u_{\max}(a, b)$	$a + b - ab$ when $p = 1$	$u_{\max}(a, b)$
Schweizer & Sklar 4	$1 - \frac{(1-a)(1-b)}{[(1-a)^p + (1-b)^p - (1-a)^p(1-b)^p]^{\frac{1}{p}}}$	$(1-a)^{-p} - 1$	$p > 0$	$a + b - ab$	$\min\left(1, \frac{a+b}{1-ab}\right)$ when $p = 1$	$\max(a, b)$
Yager [1980e]	$\min\left[1, (a^w + b^w)^{\frac{1}{w}}\right]$	$a^w$	$w > 0$	$u_{\max}(a, b)$	$\min(1, a + b)$ when $w = 1$	$u_{\max}(a, b)$
Dubois & Prade [1980]	$1 - \frac{(1-a)(1-b)}{\max((1-a), (1-b), \alpha)}$	$\alpha \in [0, 1]$		$u_{\max}(a, b)$	$a + b - ab$ when $\alpha = 1$	
Weber [1983]	$\min\left(1, a + b - \frac{\lambda}{1-\lambda}ab\right)$	$\frac{1}{\lambda} \ln \frac{1+\lambda}{1+\lambda(1-a)}$	$\lambda > -1$	$\min(1, a + b)$	$u_{\max}(a, b)$ as $\lambda \rightarrow -1$ ; $\min(1, (a + b - ab)/2)$ when $\lambda = 1$ .	$a + b - ab$ as $\lambda \rightarrow -1$ ;
Yu [1985]	$\min(1, a + b + \lambda ab)$	$\frac{1}{\lambda} \ln(1 + \lambda a)$	$\lambda > -1$	$\min(1, a + b)$	$\min(1, a + b + ab)$ when $\lambda = 1$ .	$u_{\max}(a, b)$



## Παράρτημα Β

## 12.5.2.3 SemanticRelation CS semantics

key	keyFor	A key B if and only if B is a key for accessing A.	Video segment A key image B, which means that the image is a key for accessing the video segment.
annotation	annotationOf	A annotation B if and only if B is an annotation or description of A.	Image A annotation audio segment B, which means that the audio segment provides an annotation or description of the image.
appearance	appearanceIn	A appearance B if and only if B has an appearance in or is perceivable in A.	Object "car" appearance image B, which means that the car appears in the image.
reference	referenceOf	A reference B if and only if B is a reference of A and A does not appear in B.	Object "car's engine" reference video B, which means that the video is a reference of the car's engine
quality	qualityOf	A quality B if and only if B is a quality of one or more properties of A.	Object "fire engine" quality red color, which means that the color red is a quality of fire engines.
symbol	symbolOf	A symbol B if and only if B is a symbol of A.	Object "computers" symbol an image showing zeros and ones, which means that the image is a symbol of computers.
location	locationOf	A location B	In the description, "Mary receives a flower in the street," the semantic place

<i>Relation Name</i>	<i>Inverse Relation</i>	<i>Definition</i>	<i>Examples (Informative)</i>
		if and only if B is a location of A.	"receive" location event "street" means that the street is the location of the receiving event.
source	sourceOf	A source B  if and only if B is a source location of A.	In the example "John moved the box from the floor to table," the event "move" source event "floor" means that the floor is the source location of the moving event.
destination	destinationOf	A destination B  if and only if B is a destination location of A.	In the description, "Mary moved from California to Toronto," the event "move" destination semantic place "Toronto" means that Toronto is the destination location of the event move.
path	pathOf	A path B  if and only if B is a path of A.	In the description "The train travelled along the track," event "travel" pathOf semantic place "track" means that the track is the path of the event travel.
time	timeOf	A time B  if and only if B is a time of A.	In the description "Mary was born on Christmas," event "born" time "Christmas" means that Christmas is the time of the birth event.
depiction	depictionOf	A depiction B  if and only if B is a depiction of A.	Narrative world "glass with water" depiction object "painting," which means that the painting is a depiction of the glass with water.
representation	representationOf	A representation B  if and only if B is a representation of or symbolizes in some sense A.	Concept "liberty" representation object "Statue of Liberty," which means that the Statue of Liberty represents the concept of liberty.
context	contextFor	A context B  if and only if B is a context or surrounding contributing	Event "Rosa Parks' bus ride" context concept "civil rights struggle," which means that the civil rights struggle is the context for Rosa Parks' bus ride.

<i>Relation Name</i>	<i>Inverse Relation</i>	<i>Definition</i>	<i>Examples (Informative)</i>
circumstances for A			
interpretation	interpretationOf	A interpretation B if and only if B is an interpretation of A.	Event "policeman arresting a suspected killer" interpretation the concept "law enforcement," which means that law enforcement is an interpretation of the arrest event.
agent	agentOf	A agent B if and only if B is an agent of or performs or initiates A.	In the description "Peter gives the book to Mary," the event "give" agent object "Peter" means that Peter is the agent of the book.
patient	patientOf	A patient B if and only if B is affected by or undergoes the action of A.	In the description "Peter gives the book to Mary," the event "give" patient object "book" means that the book is the patient of the giving.
experiencer	experiencerOf	A experiencer B if and only if B is an experiencer of or passively perceives a stimulus from A.	In the description "John hates pizza," the event "hate" experiencer object "John" means that John is an experiencer of hate.
stimulus	stimulusOf	A stimulus B if and only if B is a stimulus of A.	In the description, "The boy sees the barking dog," the event "see" stimulus object "dog" means that the barking dog is the stimulus of the seeing.
causer	causerOf	A causer B if and only if B is a causer of A and B is not an agent of A	In the description "Lightening strikes the tree," the event "strike" causer object "lightening" means that lightening is the causer of the strike.
goal	goalOf	A goal B if and only if B is the goal or objective of A.	In the description "Mary receives a flower," the event "receive" goal object "Mary" means that the flower is the goal of the receiving.
beneficiary	beneficiaryOf	A beneficiary B	In the description "John bakes a cake for Mary," the event "bake" beneficiary

<i>Relation Name</i>	<i>Inverse Relation</i>	<i>Definition</i>	<i>Examples (Informative)</i>
		if and only if B is the beneficiary of A such as by some action of A.	object "Mary" means that Mary is the beneficiary of the baking.
theme	themeOf	A theme B  if and only if B is a theme or topic of A.	In the description "the boy talks about the dog," the event "talk" theme object "dog" means that the dog is the theme of the boy's talking.
result	resultOf	A result B,  if and only if B is a result of A.	In the description "John bakes a cake," the event "bake" result object "cake" means that the cake is the result of the baking.
instrument	instrumentOf	A instrument B  if and only if B is an instrument of or is employed by A.	In the description, "Mary cut the salami with a knife" the event "cut" instrument object "knife" means that that the knife is the instrument of the cutting.
accompanier	accompanierOf	A accompanier B  if and only if B is an accompanier or joint agent of A, but B is not an initiator of A.	In the description, "John travelled to France with Mary," the event "travel" accompanier object "Mary" means that Mary is the accompanier of the travelling.
combination	-	A0 combination A1, A2, ... An  if and only if  A0 is a combination of A1, A2, ... A3.	Object "amphibious vehicle" combination objects "car," "boat," and "plane".
summary	summaryOf	A summary B  if and only if B is a summary of A.	In the description "Jim got hurt, he broke his leg and arm," the event "broke his leg and arm" summary event "hurt" means that getting hurt is the summary of the breaking of his leg and arm.
specialization	generalization	A specialization B  if and only if B is a specialization in	Object "mammal" specialization object "dog," which means that dog is a specialization of a mammal.

<i>Relation Name</i>	<i>Inverse Relation</i>	<i>Definition</i>	<i>Examples (Informative)</i>
meaning of A.			
similar	similar	A similar B if and only if  B is similar to or synonymous with A.	Object "car" similar object "automobile".
opposite	opposite	A opposite B if and only if  B is opposite in meaning of A.	Color "white" opposite color "black".
example	exampleOf	A example B if and only if  B is an example of A.	Concept "former world leader" example object "Bill Clinton," which means that Bill Clinton is an example of a former world leader.
interchangeable	interchangeable	A interchangeable B if and only if  B is interchangeable with or can be substituted for A.	Object "car" interchangeable object "Cadillac," which means that the car and Cadillac are synonymous in the context of the description.
identifier	identifier	A identifier B if and only if  A = B.	Agent object "Maria, 25 years old" identifier Agent object "Maria, 1 year old."
part	partOf	A part B if and only if  B is a part of A, or in other words, if there exists a C such that (A union C) refines B. In other words, B is part of A if and only if A is a "meaningful" entity that is contained by B.	Semantic place "New York City" part semantic place "Manhattan," which means that Manhattan is part of New York City.
contrast	contrast	A contrast B if and only if  B is a contrast of A.	The sepia Segments in "The Wizard of Oz" contrast the color Segments.

<i>Relation Name</i>	<i>Inverse Relation</i>	<i>Definition</i>	<i>Examples (Informative)</i>	
property	propertyOf	A property B if and only if  B is a property of if is used to qualify A.	Object property "ripeness".	"banana". Concept
user	userOf	A user B if and only if  B is a user of A.	The concept metaphor "blossoming" user concept "maturing".	
component	componentOf	A component B if and only if  B is a functional component of A.	Object "car" component object "wheel," which means that the wheel is a functional component of the car.	
substance	substanceOf	A substance B if and only if  B is a substance or constituent of A, and B permeates A.	object "martini" substance object "gin," which means that gin is a substance of a martini.	
entailment	entailmentOf	A entailment B if and only if  B is a necessary condition for A to happen.	Event entailment event "marriage," which means that marriage is a necessary condition for divorce.	"divorce"
manner	mannerOf	A manner B if and only if  B is a manner or adverb of A.	Event "John runs" manner concept "quickly," which means that John runs quickly.	
state	stateOf	A state B if and only if  B is a state of or provides the parameters to describe the state of A.	In the description "The sunset is 90% orange," the semantic state event "sunset" state "90% orange".	
dependency	dependencyOn	A dependency B if and only if  B has a dependency on	Color B dependency color A by the formula A=0.2*B, which means that color A has a dependencyOn color	

<i>Relation Name</i>	<i>Inverse Relation</i>	<i>Definition</i>	<i>Examples (Informative)</i>
		A, or in other words, there is a formula that relates the parameters describing the state of B to the parameters describing the state of A.	B.
membershipFunction -	A membershipFunction B  if and only if  A is a relation, and B is a membership function that governs the strength of A.	The relation property between object "banana" and concept "ripeness" membershipFunction semantic state "cognitive color," which means that the cognitive color governs the strength of the ripeness of the banana.	

□



# Ευρετήριο

- agent, 42  
agglomerative, 73  
association rule learning, 66  
centroid, 70  
clusters, 71  
context-aware, 122  
DDL, 61  
Description Schemes, 41  
Digital Item, 42  
Digital Item Adaptation, 42  
divisive, 73  
DL, 187  
domain, 45, 52  
e-learning, 145  
e-questionnaire, 146, 147  
fall-out, 36  
folksonomies, 278  
fuzzy c-means, 6, 68  
information retrieval, xviii  
Internet, xvii  
intra-cluster variance, 75  
intranet, 29  
kNN, 6, 73, 78  
LSA, 33, 231, 273  
LSI, 33  
MPEG-21, 41, 42  
MPEG-7, 41  
MPEG-7 DDL, 42  
ontologies, 45  
OWL, xxiv, 63, 187, 194  
pixels, 207  
principal components analysis, 76  
RDF, xxiv, 190  
reasoning, 45  
reification, xxiv, 57, 197, 253  
Semantic Web, 46, 55  
singleton, 95  
social network, 279  
social profile, 279  
statement, 57  
stemming, 93  
support vector machines, 66  
tagging, 278  
tags, 279  
thesaurus, 62  
unsupervised, 73, 76  
user input, 122  
V-linkage, 76  
weak modifier, 144  
XML, 56, 60, 91  
XML Schema, 42, 61  
Ευκλείδια απόσταση, 79  
Οντολογία Εννοιολογικού Πλαισίου, 213  
Περιγραφικές Λογικές, 68  
Περιγραφική Λογική, 100  
Σημασιολογικός Ιστός, xxiv, 186  
Σχέδια Περιγραφής, 46  
έγγραφο, 85  
έννοια, 37, 125, 129, 193, 205  
έννοιες, 53, 54  
ένταση, 84, 95  
ένταση ταξινομικού εννοιολογικού πλαι-  
σίου, 83, 91  
ένωση, 127  
αβεβαιότητα, 8, 19, 45, 93, 121, 248

αδύναμος τροποποιητής, 171  
ακρίβεια, 35, 92, 162, 216, 229, 266,  
269, 271, 275  
αλήθεια εδάφους, 162  
αλληλεπίδραση, xviii, 1  
ανάκληση, 35, 92, 162, 216, 229, 266,  
269, 271, 275  
ανάκτηση πληροφορίας, 29, 125  
ανάλυση πρωτογενών συνιστώσων, 242  
ανίχνευση αντικειμένων, 27, 228, 229  
αναγνώριση προτύπων, 71  
ανατροφοδότηση, 39  
αντικείμενο ειδήσεων, 110  
αξιώμα συσχέτισης ασαφών συνόλων,  
207  
αποτελεσματικότητα, 35, 111  
απόλυτη αλήθεια, 110, 200, 209, 275  
απόσταση Mahalanobis, 79  
απόσταση minimax, 79  
απόσταση Minkowsky, 79  
ασαφής δυαδική σχέση, 131  
ασαφής σύνθεση, 158  
ασαφής τροποποιητής, 97  
ασαφοποίηση, 46, 140, 226  
  
βάση γνώσης, 100, 102, 116, 125, 159,  
162, 164, 167, 172, 202, 204,  
235, 237  
  
γενετικός αλγόριθμος, 76, 187, 211  
γλωσσικός φράκτης, 97  
γλώσσες περιγραφής, 100  
γνωστική όραση, 10, 186  
γραμμική πληθικότητα, 171  
  
δεδομένα εκπαίδευσης, 70  
διάνυσμα αναπαράστασης, 234  
  
είδος περιοχής, 12, 227  
ελεγχόμενο λεξιλόγιο, 55  
εννοιολογικό πλαίσιο, xviii, xxiii–xxv,  
4  
εντατικές έννοιες, 54  
εξαγωγή χαρακτηριστικών γνωρισμάτων,  
76  
εξόρυξη δεδομένων, 17, 71  
επεκτατικές έννοιες, 54  
επεξεργασία εικόνας, 71  
επικρότηση, 162

επιλογή χαρακτηριστικών γνωρισμάτων,  
71, 76  
ερώτημα, 138  
ερώτημα αναζήτησης, 30  
ερώτημα με χρήση παραδείγματος, 38  
ευαισθησία, 35, 111  
  
η κατάρα των διαστάσεων, 72  
  
θεματική κατηγορία, 130  
θεματική κατηγοριοποίηση εγγράφων, 6,  
85  
θεματική περιοχή, 52, 55, 193, 205  
θησαυρός, 62  
θησαυρός περιοχών, 234, 235  
  
ιδιομορφία, 35, 111  
ιδιότητες, 53, 54  
ιεραρχία ταξινόμησης, 82, 127  
ιεραρχική συσταδοποίηση, 71, 72  
ισορροπημένο φ-σκορ, 36  
  
κέρδος πληροφορίας, 76  
καταμερισμός, 72  
κατηγοριοποίηση, 70  
κοινή ερμηνεία, 84, 96, 129  
κριτήριο του Ward, 75  
  
λανθάνουσα σημασιολογική ανάλυση, 6,  
68, 114, 225  
λαονομία, 223  
  
μείωση διαστάσεων εισαγωγής, 71  
μερικώς διατακτική σχέση, 129  
μεταδεδομένα, 18, 94  
μετρική ομοιότητας με βάση το συνημί-  
τονο, 160  
μηχανική μάθηση, 15, 16, 71  
μοναδική αποσύνθεση τιμής, 68  
μονοσύνολο, 74, 84  
μοντελοποίηση χρήστη, 3, 120  
  
οντολογία, xxiv, 58, 179, 186  
οντολογίες, 52, 82, 211  
οπτικοί περιγραφείς, 211  
  
περιγραφέας, 41  
περιγραφείς, 61  
περιπλεκτικό ασαφές συμπλήρωμα, 83,  
90

- πιθανοτική λανθάνουσα σημασιολογική ανάλυση, 114
- πλέγμα, 225
- πλοιήγηση, 62
- προσωπική μετρική κατάταξης, 160
- προτίμηση, 125
- προφίλ, xxiv
- προφίλ χρήστη, 137
- πρωτότυπα στιγμιότυπα, 204
- πόροι, 56, 190
- ρητή συσχετιστική ανατροφοδότηση, 40
- σάκος λέξεων, 224
- σημασιολογική δεικτοδότηση, 38
- σημασιολογική οντότητα, 59
- σημασιολογική περιγραφή, 46
- σημασιολογικό κενό, 223
- σταθμισμένο άθροισμα, 213
- στιγμιότυπα, 53
- συμβολισμός αθροίσματος, 47
- συνεπαγωγή, 130
- συνεργατική σύσταση, 152
- συντακτικό ονομάτων και διευθύνσεων, 56
- συσταδιοποίηση, 70, 140, 233
- συσταδιοποίηση μέσου συνδέσμου, 75
- συσταδιοποίηση με βάση την πυκνότητα, 72
- συσταδιοποίηση με βάση το πλέγμα, 72
- συσταδιοποίηση μοναδιαίου συνδέσμου, 75
- συσταδιοποίηση πλήρους συνδέσμου, 75
- συστολή, 127
- συστροφικό συμπλήρωμα, 134
- συσχετιστική ανατροφοδότηση, 24, 40
- σχέδια κωδικοποίησης, 61
- σχέδια περιγραφής, 41, 61
- σχέδια περιγραφής πολυμέσων, 60
- σχέδια ταξινόμησης, 62
- σχέσεις, 53, 55
- σχέσεις ένταξης, 56
- σχέση, 129
- σχέση πολλών-με-πολλά, 91
- σχέση συμβατότητας, 188
- σχεδόν-διατακτική σχέση, 129
- σχεδόν-ταξινομική σχέση, 129
- σχετικότητα εννοιολογικού πλαισίου, 187
- σύνδεσμος, 86
- σύνθεση, 127
- σύνοδος αναζήτησης, 169
- ταξινομία, 55, 82, 128
- ταξινομική σχέση, 82, 127, 188
- ταξινομικό εννοιολογικό πλαίσιο, 6, 83, 89
- ταξινόμηση σκηνής, 27, 228, 229
- τρέχουσα κατάσταση προόδου, 227
- τυφλή συσχετιστική ανατροφοδότηση, 40
- τύποι περιοχής, 233
- τύπος περιοχής, xxv, 221, 247
- υπαγωγή, 54
- υπερφόρτωση πληροφοριών, 29
- υπηρεσία εξαγωγής συμπερασμάτων, 101
- υπονοούμενη συσχετιστική ανατροφοδότηση, 40
- υπόθεση αληθείας, 206
- φ-μετρική, 36
- χαρακτηρισμός, 91
- χωρικές διαστάσεις εισαγωγής, 6, 77
- όρος, 37, 62
- ύψος ασαφούς συνόλου, 47
- ύψος ταξινομικού εννοιολογικού πλαισίου, 83, 91

□