

ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΣ ΥΠΟΛΟΓΙΣΤΩΝ ΜΕ

C++

(2^η Εργασία)



ΜΠΟΓΔΑΝΟΣ ΜΙΧΑΗΛ (3100123)
ΜΑΓΚΟΣ ΡΑΦΑΗΛ-ΓΕΩΡΓΙΟΣ(3100098)

Αρχείο instance.h

Στο αρχείο αυτό έχουμε την δήλωση της κλάσης Instance.

Ιδιωτικά μέλη της κλάσης Instance:

- **unsigned featuresNum** : Μεταβλητή unsigned int η οποία θα περιέχει τον αριθμό των keywords του κάθε μηνύματος Instance.
- **bool cat**: Μεταβλητή boolean για τοποθέτηση της κατηγορίας του μηνύματος (true για spam, false για ham).
- **string filename**: Μεταβλητή string στην οποία αποθηκεύεται το αρχείο που περιέχει το μήνυμα που βρίσκεται στο Instance.
- **string* _features**: Δείκτης σε string για δυναμική καταχώρηση των keywords.
- **unsigned* _featureID**: Δείκτης σε unsigned int για δυναμική καταχώρηση του ID των keywords.
- **unsigned* _frequency**: Δείκτης σε unsigned int για δυναμική καταχώρηση της συχνότητας εμφάνισης των keywords.

Δημόσια μέλη της κλάσης Instance:

- **Instance(unsigned features=0)** : Δήλωση του κατασκευαστή αντικειμένων Instance. Ο κατασκευαστής χρησιμοποιεί την προεπιλεγμένη τιμή 0 ώστε να μπορεί να χρησιμοποιηθεί και ως προεπιλεγμένος.
- **unsigned getNumberOfFeatures() const** : Μέθοδος που επιστρέφει την μεταβλητή featuresNum (η μέθοδος δεν μπορεί να μεταβάλλει το αντικείμενο που την κάλεσε).
- **Instance(const Instance& original)** : Κατασκευαστής αντιγράφου που παίρνει ως όρισμα μια αναφορά σε αντικείμενο Instance (χωρίς να επιτρέπει αλλαγές στο όρισμα-Instance).
- **~Instance()** : καταστροφέας που αποδεσμεύει τον χώρο μνήμης που δέσμευσαν οι _features , _frequency , _featureID .
- **void setCategory(bool category)** : Θέτει την ιδιωτική μεταβλητή cat (που δείχνει την κατηγορία του Instance) ίση με το όρισμα.

- **bool getCategory() const** : Επιστρέφει την τιμή της ιδιωτικής bool μεταβλητής cat(η μέθοδος δεν μπορεί να μεταβάλλει το αντικείμενο που την κάλεσε).
- **void setFileName(const string& fileName)** : Αποθηκεύει το όρισμα(αναφορά σε string,που δεν μπορεί να αλλάξει αυτό στο οποίο αναφέρεται) στην ιδιωτική μεταβλητή filename.
- **string getFileName() const** : Επιστρέφει την τιμή της ιδιωτικής μεταβλητής filename(η μέθοδος δεν μπορεί να μεταβάλλει το αντικείμενο που την κάλεσε).
- **void setFeature(unsigned i, const string& feature, unsigned featureID, unsigned frequency)** : Θέτει τους δυναμικά δεσμευμένους χώρους _features, _featureID και _frequency στην θέση i ίσους με feature(αναφορά σε string που δεν μπορεί να αλλάξει αυτό στο οποίο αναφέρεται) , featureID και frequency αντιστοίχως.
- **string getFeature(unsigned i) const** : Επιστρέφει την string τιμή από τον δυναμικό πίνακα στη θέση i (η μέθοδος δεν μπορεί να μεταβάλλει το αντικείμενο που την κάλεσε).
- **unsigned getFeatureID(unsigned i) const** : Επιστρέφει την unsigned int τιμή από τον δυναμικό πίνακα στη θέση i (η μέθοδος δεν μπορεί να μεταβάλλει το αντικείμενο που την κάλεσε).
- **unsigned getFrequency(unsigned i) const** : Επιστρέφει την unsigned int τιμή από τον δυναμικό πίνακα στη θέση i (η μέθοδος δεν μπορεί να μεταβάλλει το αντικείμενο που την κάλεσε).
- **Instance& operator=(const Instance& right)** : Υπερφορτώνει τον τελεστή = , ώστε στο αριστερό μέρος του να εκχωρούνται τα περιεχόμενα του αντικειμένου Instance που βρίσκεται στο δεξιό μέρος.Εάν τα περιεχόμενα αριστερά και δεξιά του τελεστή είναι ίσα τότε επιστρέφεται κατ'αναφορά το αριστερό μέρος με χρήση της δεσμευμένης λέξης *this που δείχνει στη διεύθυνση μνήμης του αντικειμένου που κάλεσε την συνάρτηση(δηλαδή δεν γίνεται καμία αλλαγή).Αλλιώς,διαγράφονται οι δυναμικοί πίνακες του αριστερού αντικειμένου Instance,οι μεταβλητές cat,filename και featuresNum γίνονται ίσες με τις αντίστοιχες του αντικειμένου right και οι δυναμικοί πίνακες του right αναθέτουν την τιμή τους στους αντίστοιχους δυναμικού πίνακες του αριστερού αντικειμένου με χρήση βρόγχου for(από 0 έως το μέγεθος των πινάκων του right).
- **friend ostream& operator<<(ostream& out, const Instance& inst)** : Φίλια συνάρτηση (έχει πρόσβαση στα ιδιωτικά μέλη του Instance αντικειμένου) που υπερφορτώνει τον τελεστή εξόδου.
- **friend istream& operator>>(istream& in, Instance& inst)** : Φίλια συνάρτηση (έχει πρόσβαση στα ιδιωτικά μέλη του Instance αντικειμένου) που υπερφορτώνει τον τελεστή εισόδου.

Αρχείο instance.cpp

- **Instance::Instance(unsigned features)** : Υλοποίηση του κατασκευαστή των Instances. Η μεταβλητή features (όρισμα) αναθέτει την τιμή της στην ιδιωτική μεταβλητή featuresNum της κλάσης Instance. Ακολουθώντας δεσμεύονται δυναμικά οι απαραίτητοι χώροι με χρήση new και μέγεθος την τιμή της μεταβλητής featuresNum.
- **Instance::Instance(const Instance& original)** : Υλοποίηση του κατασκευαστή αντιγράφου. Οι μεταβλητές featuresNum, filename και cat γίνονται ίσες με τις ανάλογες μεταβλητές του ορίσματος. Ακολουθώντας δημιουργούνται οι δυναμικοί πίνακες με μέγεθος featuresNum και παίρνουν τις τιμές τους από το original με χρήση ενός βρόγχου for.
- **ostream & operator<<(ostream& out, const Instance& inst)** : Τυπώνει τα περιεχόμενα του αντικειμένου Instance στην ζητούμενη μορφή, με χρήση ενός βρόγχου for που διατρέχει τους δυναμικούς πίνακες του αντικειμένου και ανακτά τα κατάλληλα πεδία με χρήση μεθόδων get(getFeature, getFeatureID, getFrequency). Ακολουθώντας επιστρέφεται αναφορά του ostream out για μελλοντική χρήση του.
- **istream& operator>>(istream& in, Instance& inst)** : Χρήση ενός ρεύματος εισόδου που διαβάζει ένα αρχείο κειμένου. Κατά την υλοποίηση της συνάρτησης θεωρήθηκε ότι η ανάγνωση μπορεί να ξεκινάει είτε από την πρώτη γραμμή του αρχείου (< messagecollection messages = "... " >) για αυτόνομη χρήση ώστε να διαβάζει και να δημιουργεί αντικείμενα Instance είτε μόνο ενότητες <message></message> (για χρήση στην InstancePool). Συνεπώς, η πρώτη γραμμή αλλά και η επόμενη (κενή γραμμή) αγνοούνται. Όταν εντοπιστεί όμως η γραμμή στην οποία ορίζεται το μήνυμα (< message file = ".." category = ".." features = ".." >) αρχικά με κατάλληλο substr (από τη θέση του τελευταίου = και μήκος ίσο με το μήκος της γραμμής μείον την θέση εμφάνισης του τελευταίου " μείον 1) στην γραμμή αποσπάται ο αριθμός (αρχικά είναι string όμως με χρήση της μεθόδου atoi λαμβάνουμε την αριθμητική τιμή του string) των features και κατασκευάζεται ένα προσωρινό αντικείμενο Instance. Με παρόμοιο σκεπτικό βρίσκουμε από την γραμμή τα πεδία filename (substr από τη θέση του πρώτου = και μήκος τη θέση του δεύτερου " μείον τη θέση του πρώτου " μείον 1) και category (substr από τη θέση του = που βρίσκεται μετά από τους χαρακτήρες .txt " και μήκος από τη θέση εμφάνισης της λέξης features μείον τη θέση του = που βρίσκεται μετά από τους χαρακτήρες .txt " -5). Εάν το substr δώσει αποτέλεσμα την λέξη spam τότε το category γίνεται true, αλλιώς γίνεται false. Με χρήση της getline διαβάζουμε από το αρχείο την επόμενη γραμμή, η οποία θα περιέχει ενότητα feature token. Χρησιμοποιούμε έναν βρόγχο for που με κλήση της setFeature στο temp αντικείμενο τοποθετεί τη θέση, το feature, το featureID και το frequency, τα οποία έχουν αποσπαστεί με κατάλληλα substr. Ο βρόγχος ξεκινά από το 0 αυξάνοντας κατά 1 και σπάει μόλις φτάσει στο μέγεθος των features του temp.

Η μέθοδος ολοκληρώνεται επιτυχώς εάν η επόμενη γραμμή είναι η ετικέτα τέλους `</ message>` τότε το όρισμα `inst` με χρήση του `operator =` γίνεται ίσο με το `temp`. Έπειτα διαβάζουμε μια ακόμη γραμμή, η οποία θα είναι κενή, ώστε να την αγνοήσουμε σε μελλοντική ανάγνωση από το ίδιο `stream` (για παράδειγμα στην `InstancePool`). Εάν δεν υπάρχει ετικέτα τέλους τότε τυπώνεται μήνυμα σφάλματος και το προσωρινό αντικείμενο καταστρέφεται χωρίς πριν να έχει αποδώσει τα περιεχόμενά του στο όρισμα.

Αρχείο `instancepool.h`

Στο αρχείο αυτό έχουμε την δήλωση της κλάσης `InstancePool`.

Ιδιωτικά μέλη της κλάσης `InstancePool`:

➤ `vector<Instance> instances` : Χρήση `vector` για αποθήκευση των `Instance` αντικειμένων.

Δημόσια μέλη της κλάσης `InstancePool`:

- `InstancePool()` : Προεπιλεγμένος κενός κατασκευαστής. Ουσιαστικά με την κλήση του θα δημιουργηθεί ο ιδιωτικός `vector instances` που δεν θα περιέχει κανένα `Instance`.
- `~InstancePool()`: Η χρήση καταστροφέα σε αυτή την κλάση δεν ήταν απαραίτητη διότι και η κλάση `vector` αλλά και η κλάση `Instance` διαθέτουν καταστροφείς.
- `InstancePool(const InstancePool& original)` : Κατασκευαστής αντιγράφου που παίρνει ως όρισμα μια αναφορά σε αντικείμενο `InstancePool` (χωρίς να επιτρέπει αλλαγές στο όρισμα-`InstancePool`).
- `unsigned getNumberOfInstances() const` : Επιστρέφει το μέγεθος του `vector` (άρα το μέγεθος του `InstancePool`).
- `const Instance& operator[](unsigned index) const` : Υπερφόρτωση του τελεστή `[]`. Επιστρέφει το στοιχείο που βρίσκεται στη θέση `index` του `vector`.
- `InstancePool& operator=(const InstancePool& right)` : Υπερφόρτωση του τελεστή `=`. Εάν το αντικείμενο που κάλεσε την συνάρτηση (`this`, αριστερά του `=`) είναι το ίδιο με το αντικείμενο δεξιά του `=` τότε επιστρέφεται το πρώτο. Αλλιώς ο `vector` του `this` γίνεται ίσος με τον `vector` του `right` και επιστρέφεται το `this`.
- `friend istream& operator>>(istream& in, InstancePool& ip)` : Φίλια συνάρτηση (έχει πρόσβαση στα ιδιωτικά μέλη του `Instance` αντικειμένου) που υπερφορτώνει τον τελεστή εισόδου.
- `friend ostream& operator<<(ostream& out, const InstancePool& ip)` : Φίλια συνάρτηση (έχει πρόσβαση στα ιδιωτικά μέλη του `Instance` αντικειμένου) που υπερφορτώνει τον τελεστή εξόδου.

Αρχείο instancepool.cpp

- **InstancePool(const InstancePool& original)** : Υλοποίηση του κατασκευαστή αντιγράφου. Ουσιαστικά ο vector του αντικειμένου original χρησιμοποιείται για την κατασκευή του νέο αντικειμένου.
- **istream& operator>>(istream& in, InstancePool& ip)** : Αρχικά ο vector του ip αδειάζει τα περιεχόμενά του. Ακολουθώς ορίζεται ένα αντικείμενο Instance για προσωρινή χρήση. Χρησιμοποιούμε το ρεύμα εισόδου in για να διαβάσουμε την επόμενη γραμμή που θα είναι η < messagecollection messages = "..." > . Αποσπάμε τον αριθμό των μηνυμάτων με substr (από τη θέση του = +3 και με μήκος ίσο με το μήκος της γραμμής μείον τη θέση του τελευταίου ") και το μετατρέπουμε σε int με χρήση της μεθόδου atoi και διαβάζουμε την επόμενη γραμμή (κενή γραμμή). Στη συνέχεια με ένα βρόγχο for από 0 έως τον αριθμό των μηνυμάτων κάνουμε χρήση του υπερφορτωμένου τελεστή >> για αντικείμενα τύπου Instance στο προσωρινό Instance και το προσθέτουμε στον vector του ip. Μετά το τέλος του βρόγχου επιστρέφουμε το ρεύμα εισόδου.
- **ostream& operator<<(ostream& out, const InstancePool& ip)** : Τυπώνουμε αρχικά το πλήθος των μηνυμάτων που βρίσκουμε καλώντας την μέθοδο getNumberOfInstances() στο αντικείμενο ip. Ακολουθώς με ένα βρόγχο for από το 0 έως το πλήθος των μηνυμάτων κάνουμε χρήση του υπερφορτωμένου τελεστή [] της InstancePool και επιστρέφουμε ένα αντικείμενο Instance στο οποίο εφαρμόζουμε τον υπερφορτωμένο τελεστή << για να τυπώσουμε τα περιεχόμενά του.

Παράδειγμα Χρήσης

Δεδομένα εισόδου

```
output - Σημειωματάριο
Αρχείο Επεξεργασία Μορφή Προβολή Βοήθεια
< messagecollection messages = "289" >

< message file = "spam/spmsga1.txt" category = "spam" features = "26" >
< feature token = "!" id = "0" freq = "3" />
< feature token = "remove" id = "4" freq = "1" />
< feature token = "your" id = "6" freq = "2" />
< feature token = "you" id = "7" freq = "5" />
< feature token = "removed" id = "10" freq = "1" />
< feature token = "@" id = "13" freq = "1" />
< feature token = "receive" id = "14" freq = "1" />
< feature token = "our" id = "16" freq = "2" />
< feature token = "mailing" id = "18" freq = "1" />
< feature token = "hundreds" id = "21" freq = "1" />
< feature token = "business" id = "24" freq = "1" />
< feature token = "%" id = "29" freq = "1" />
< feature token = "save" id = "32" freq = "1" />
< feature token = "$" id = "35" freq = "4" />
< feature token = "com" id = "43" freq = "1" />
< feature token = "every" id = "44" freq = "2" />
< feature token = "yourself" id = "46" freq = "1" />
< feature token = "company" id = "55" freq = "1" />
< feature token = "dollars" id = "61" freq = "1" />
< feature token = "earn" id = "75" freq = "1" />
< feature token = "(" id = "77" freq = "2" />
< feature token = "here" id = "83" freq = "1" />
< feature token = "customers" id = "85" freq = "1" />
< feature token = ")" id = "86" freq = "2" />
< feature token = "make" id = "91" freq = "1" />
< feature token = "do" id = "98" freq = "1" />
< / message>

< message file = "spam/spmsga102.txt" category = "spam" features = "34" >
< feature token = "!" id = "0" freq = "59" />
< feature token = "free" id = "3" freq = "1" />
< feature token = "remove" id = "4" freq = "4" />
< feature token = "your" id = "6" freq = "20" />
< feature token = "you" id = "7" freq = "22" />
< feature token = "money" id = "9" freq = "3" />
< feature token = "@" id = "13" freq = "1" />
< feature token = "receive" id = "14" freq = "1" />
< feature token = "our" id = "16" freq = "6" />
< feature token = "get" id = "17" freq = "1" />
< feature token = "mailing" id = "18" freq = "2" />
< feature token = "today" id = "22" freq = "2" />
< feature token = "internet" id = "26" freq = "1" />

Ln 22, Col 37
```

```
output - Σημειωματάριο
Αρχείο Επεξεργασία Μορφή Προβολή Βοήθεια
< feature token = "days" id = "30" freq = "1" />
< feature token = "save" id = "32" freq = "1" />
< feature token = "$" id = "35" freq = "24" />
< feature token = "offer" id = "36" freq = "1" />
< feature token = "now" id = "39" freq = "5" />
< feature token = "sales" id = "40" freq = "1" />
< feature token = "best" id = "41" freq = "1" />
< feature token = "com" id = "43" freq = "1" />
< feature token = "ll" id = "45" freq = "1" />
< feature token = "cash" id = "50" freq = "1" />
< feature token = "over" id = "53" freq = "3" />
< feature token = "guarantee" id = "59" freq = "1" />
< feature token = "service" id = "62" freq = "1" />
< feature token = "#" id = "64" freq = "2" />
< feature token = "check" id = "69" freq = "3" />
< feature token = "reply" id = "73" freq = "1" />
< feature token = "ever" id = "79" freq = "1" />
< feature token = "guaranteed" id = "80" freq = "1" />
< feature token = "zip" id = "87" freq = "1" />
< feature token = "make" id = "91" freq = "1" />
< feature token = "do" id = "98" freq = "6" />
< / message>

< message file = "ling/5-1274msg1.txt" category = "ham" features = "6" >
< feature token = "university" id = "2" freq = "1" />
< feature token = "linguistics" id = "5" freq = "1" />
< feature token = "@" id = "13" freq = "1" />
< feature token = "de" id = "47" freq = "6" />
< feature token = "(" id = "77" freq = "2" />
< feature token = ")" id = "86" freq = "2" />
< / message>

< message file = "ling/5-1275msg1.txt" category = "ham" features = "14" >
< feature token = "language" id = "1" freq = "1" />
< feature token = "university" id = "2" freq = "8" />
< feature token = "linguistics" id = "5" freq = "3" />
< feature token = "@" id = "13" freq = "3" />
< feature token = "languages" id = "15" freq = "2" />
< feature token = "edu" id = "25" freq = "2" />
< feature token = "linguistic" id = "31" freq = "3" />
< feature token = "$" id = "35" freq = "9" />
< feature token = "department" id = "63" freq = "1" />
< feature token = "(" id = "77" freq = "14" />
< feature token = "theory" id = "78" freq = "5" />
< feature token = ")" id = "86" freq = "14" />

Ln 22, Col 37
```

```
output - Σημειωματάριο
Αρχείο Επεξεργασία Μορφή Προβολή Βοήθεια

< feature token = "make" id = "91" freq = "1" />
</ message>

< message file = "ling/5-1279msg1.txt" category = "ham" features = "8" >
< feature token = "university" id = "2" freq = "1" />
< feature token = "linguistics" id = "5" freq = "1" />
< feature token = "you" id = "7" freq = "1" />
< feature token = "@" id = "13" freq = "1" />
< feature token = "our" id = "16" freq = "1" />
< feature token = "(" id = "77" freq = "3" />
< feature token = ")" id = "86" freq = "3" />
< feature token = "do" id = "98" freq = "2" />
</ message>

< message file = "ling/5-1279msg2.txt" category = "ham" features = "9" >
< feature token = "!" id = "0" freq = "1" />
< feature token = "linguistics" id = "5" freq = "1" />
< feature token = "you" id = "7" freq = "1" />
< feature token = "@" id = "13" freq = "1" />
< feature token = "edu" id = "25" freq = "1" />
< feature token = "internet" id = "26" freq = "2" />
< feature token = "best" id = "41" freq = "1" />
< feature token = "(" id = "77" freq = "1" />
< feature token = ")" id = "86" freq = "1" />
</ message>

< message file = "ling/5-1279msg3.txt" category = "ham" features = "7" >
< feature token = "university" id = "2" freq = "1" />
< feature token = "your" id = "6" freq = "1" />
< feature token = "you" id = "7" freq = "1" />
< feature token = "english" id = "11" freq = "3" />
< feature token = "department" id = "63" freq = "1" />
< feature token = "(" id = "77" freq = "5" />
< feature token = ")" id = "86" freq = "8" />
</ message>

< message file = "ling/5-1282msg3.txt" category = "ham" features = "22" >
< feature token = "language" id = "1" freq = "1" />
< feature token = "university" id = "2" freq = "1" />
< feature token = "linguistics" id = "5" freq = "1" />
< feature token = "you" id = "7" freq = "1" />
< feature token = "money" id = "9" freq = "1" />
< feature token = "english" id = "11" freq = "2" />
< feature token = "@" id = "13" freq = "1" />
< feature token = "receive" id = "14" freq = "1" />

Ln 22, Col 37
```

```
output - Σημειωματάριο
Αρχείο Επεξεργασία Μορφή Προβολή Βοήθεια

< feature token = "languages" id = "15" freq = "1" />
< feature token = "our" id = "16" freq = "6" />
< feature token = "papers" id = "19" freq = "4" />
< feature token = "edu" id = "25" freq = "2" />
< feature token = "linguistic" id = "31" freq = "5" />
< feature token = "$" id = "35" freq = "6" />
< feature token = "over" id = "53" freq = "1" />
< feature token = "check" id = "69" freq = "1" />
< feature token = "reply" id = "73" freq = "1" />
< feature token = "(" id = "77" freq = "9" />
< feature token = "theory" id = "78" freq = "10" />
< feature token = "here" id = "83" freq = "1" />
< feature token = ")" id = "86" freq = "9" />
< feature token = "analysis" id = "99" freq = "3" />
</ message>

</ messagecollection >

Ln 22, Col 37
```


Έξοδος του προγράμματος

```
Διαχειριστής: C:\Windows\System32\cmd.exe
C:\Users\Rafael\Desktop\Test>g++ example.cpp instance.h instance.cpp instancepool.h instancepool.cpp -o Assignment2.exe
C:\Users\Rafael\Desktop\Test>Assignment2 < output.txt
< messagecollection messages = "9" >

< message file = "spam/spmsga1.txt" category = "spam" features = "26" >
< feature token = "!" id = "0" freq = "3" />
< feature token = "remove" id = "4" freq = "1" />
< feature token = "your" id = "6" freq = "2" />
< feature token = "you" id = "7" freq = "5" />
< feature token = "removed" id = "10" freq = "1" />
< feature token = "@" id = "13" freq = "1" />
< feature token = "receive" id = "14" freq = "1" />
< feature token = "our" id = "16" freq = "2" />
< feature token = "mailing" id = "18" freq = "1" />
< feature token = "hundreds" id = "21" freq = "1" />
< feature token = "business" id = "24" freq = "1" />
< feature token = "%" id = "29" freq = "1" />
< feature token = "save" id = "32" freq = "1" />
< feature token = "$" id = "35" freq = "4" />
< feature token = "com" id = "43" freq = "1" />
< feature token = "every" id = "44" freq = "2" />
< feature token = "yourself" id = "46" freq = "1" />
< feature token = "company" id = "55" freq = "1" />
< feature token = "dollars" id = "61" freq = "1" />
< feature token = "earn" id = "75" freq = "1" />
< feature token = "(" id = "77" freq = "2" />
< feature token = "here" id = "83" freq = "1" />
< feature token = "customers" id = "85" freq = "1" />
< feature token = ")" id = "86" freq = "2" />
< feature token = "make" id = "91" freq = "1" />
< feature token = "do" id = "98" freq = "1" />
</ message>

< message file = "spam/spmsga102.txt" category = "spam" features = "34" >
< feature token = "!" id = "0" freq = "59" />
< feature token = "free" id = "3" freq = "1" />
< feature token = "remove" id = "4" freq = "4" />
< feature token = "your" id = "6" freq = "20" />
< feature token = "you" id = "7" freq = "22" />
< feature token = "money" id = "9" freq = "3" />
< feature token = "@" id = "13" freq = "1" />
< feature token = "receive" id = "14" freq = "1" />
< feature token = "our" id = "16" freq = "6" />
< feature token = "get" id = "17" freq = "1" />
< feature token = "mailing" id = "18" freq = "2" />
< feature token = "today" id = "22" freq = "2" />
< feature token = "internet" id = "26" freq = "1" />
```

```
Διαχειριστής: C:\Windows\System32\cmd.exe
< feature token = "days" id = "30" freq = "1" />
< feature token = "save" id = "32" freq = "1" />
< feature token = "$" id = "35" freq = "24" />
< feature token = "offer" id = "36" freq = "1" />
< feature token = "now" id = "39" freq = "5" />
< feature token = "sales" id = "40" freq = "1" />
< feature token = "best" id = "41" freq = "1" />
< feature token = "com" id = "43" freq = "1" />
< feature token = "ll" id = "45" freq = "1" />
< feature token = "cash" id = "50" freq = "1" />
< feature token = "over" id = "53" freq = "3" />
< feature token = "guarantee" id = "59" freq = "1" />
< feature token = "service" id = "62" freq = "1" />
< feature token = "#" id = "64" freq = "2" />
< feature token = "check" id = "69" freq = "3" />
< feature token = "reply" id = "73" freq = "1" />
< feature token = "ever" id = "79" freq = "1" />
< feature token = "guaranteed" id = "80" freq = "1" />
< feature token = "zip" id = "87" freq = "1" />
< feature token = "make" id = "91" freq = "1" />
< feature token = "do" id = "98" freq = "6" />
</ message>

< message file = "ling/5-1274msg1.txt" category = "ham" features = "6" >
< feature token = "university" id = "2" freq = "1" />
< feature token = "linguistics" id = "5" freq = "1" />
< feature token = "@" id = "13" freq = "1" />
< feature token = "de" id = "47" freq = "6" />
< feature token = "(" id = "77" freq = "2" />
< feature token = ")" id = "86" freq = "2" />
</ message>

< message file = "ling/5-1275msg1.txt" category = "ham" features = "14" >
< feature token = "language" id = "1" freq = "1" />
< feature token = "university" id = "2" freq = "8" />
< feature token = "linguistics" id = "5" freq = "3" />
< feature token = "@" id = "13" freq = "3" />
< feature token = "languages" id = "15" freq = "2" />
< feature token = "edu" id = "25" freq = "2" />
< feature token = "linguistic" id = "31" freq = "3" />
< feature token = "$" id = "35" freq = "9" />
< feature token = "department" id = "63" freq = "1" />
< feature token = "(" id = "77" freq = "14" />
< feature token = "theory" id = "78" freq = "5" />
< feature token = ")" id = "86" freq = "14" />
< feature token = "syntax" id = "97" freq = "4" />
< feature token = "analysis" id = "99" freq = "1" />
</ message>
```

```
Διαχειριστής: C:\Windows\System32\cmd.exe
< message file = "ling/5-1277msg1.txt" category = "ham" features = "9" >
< feature token = "university" id = "2" freq = "1" />
< feature token = "you" id = "7" freq = "4" />
< feature token = "english" id = "11" freq = "1" />
< feature token = "@" id = "13" freq = "1" />
< feature token = "edu" id = "25" freq = "1" />
< feature token = "department" id = "63" freq = "1" />
< feature token = "(" id = "77" freq = "3" />
< feature token = ")" id = "86" freq = "3" />
< feature token = "make" id = "91" freq = "1" />
</ message>

< message file = "ling/5-1279msg1.txt" category = "ham" features = "8" >
< feature token = "university" id = "2" freq = "1" />
< feature token = "linguistics" id = "5" freq = "1" />
< feature token = "you" id = "7" freq = "1" />
< feature token = "@" id = "13" freq = "1" />
< feature token = "our" id = "16" freq = "1" />
< feature token = "(" id = "77" freq = "3" />
< feature token = ")" id = "86" freq = "3" />
< feature token = "do" id = "98" freq = "2" />
</ message>

< message file = "ling/5-1279msg2.txt" category = "ham" features = "9" >
< feature token = "!" id = "0" freq = "1" />
< feature token = "linguistics" id = "5" freq = "1" />
< feature token = "you" id = "7" freq = "1" />
< feature token = "@" id = "13" freq = "1" />
< feature token = "edu" id = "25" freq = "1" />
< feature token = "internet" id = "26" freq = "2" />
< feature token = "best" id = "41" freq = "1" />
< feature token = "(" id = "77" freq = "1" />
< feature token = ")" id = "86" freq = "1" />
</ message>

< message file = "ling/5-1279msg3.txt" category = "ham" features = "7" >
< feature token = "university" id = "2" freq = "1" />
< feature token = "your" id = "6" freq = "1" />
< feature token = "you" id = "7" freq = "1" />
< feature token = "english" id = "11" freq = "3" />
< feature token = "department" id = "63" freq = "1" />
< feature token = "(" id = "77" freq = "5" />
< feature token = ")" id = "86" freq = "8" />
</ message>

< message file = "ling/5-1282msg3.txt" category = "ham" features = "22" >
< feature token = "language" id = "1" freq = "1" />
< feature token = "university" id = "2" freq = "1" />
< feature token = "linguistics" id = "5" freq = "1" />
```

```
Διαχειριστής: C:\Windows\System32\cmd.exe
< feature token = "$" id = "35" freq = "6" />
< feature token = "over" id = "53" freq = "1" />
< feature token = "check" id = "69" freq = "1" />
< feature token = "reply" id = "73" freq = "1" />
< feature token = "(" id = "77" freq = "9" />
< feature token = "theory" id = "78" freq = "10" />
< feature token = "here" id = "83" freq = "1" />
< feature token = ")" id = "86" freq = "9" />
< feature token = "analysis" id = "99" freq = "3" />
</ message>

</ messagecollection >

C:\Users\Rafael\Desktop\Test>
```