

ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΩΝ ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

Επεξεργασία Φυσικής Γλώσσας Απαλλακτική εργασία

ΟΜΑΔΑ ΑΝΑΠΤΥΞΗΣ:

ΒΑΣΙΛΕΙΟΣ ΖΑΡΤΗΛΑΣ ΠΑΠΑΧΑΡΑΛΑΜΠΟΥΣ - Π17168

Εκπαιδευτής:

Καθηγητής Παναγιωτόπουλος Θέμης

Πειραιάς Ιούλιος 2020

ПЕРІЕХОМЕНА

Θέμα 1:
1) Tokenize (Τμηματοποίηση)θ
2) Normalisation (Κανονικοποίηση)6
3) Recognising Entities (Αναγνώριση Οντοτήτων)
4) Dependency Parsing (Ανάλυση Εξαρτήσεων)
5) Generation (Δημιουργία)
Θέμα 2:10
(a)10
(b)1
Θέμα 3:
Θέμα i:
Υλοποίηση ενός ChatBot σε Java13

Θέμα 1:

Αναζητείστε στο διαδίκτυο ένα σύγχρονο θέμα επεξεργασίας φυσικής γλώσσας που σας έκανε εντύπωση, ή σας κίνησε το ενδιαφέρον από τις διαλέξεις του προσκεκλημένου ομιλητή κατά την διάρκεια του μαθήματος. Αναπτύξτε το θέμα ή την εφαρμογή σε 10 το πολύ σελίδες.

Αρχικά ας ορίσουμε την είναι η επεξεργασία φυσικής γλώσσας. Η επεξεργασία Φυσικής Γλώσσας είναι για τους υπολογιστές ένας τρόπος ανάλυσης, κατανόησης και άντλησης πληροφορίας από την ανθρώπινη γλώσσα με ένα έξυπνο και χρήσιμο τρόπο. Με τη χρήση της, οι προγραμματιστές μπορούν να οργανώνουν και να δομούν την υπάρχουσα γνώση για την εκτέλεση διαφόρων καθηκόντων.

Η Επεξεργασία Φυσικής Γλώσσας χρησιμοποιείται για την ανάλυση κειμένου και επιτρέπει τις μηχανές να κατανοήσουν πώς μιλούν οι άνθρωποι. Αυτή η αλληλεπίδραση ανθρώπου-υπολογιστή επιτρέπει τη δημιουργία εφαρμογών πραγματικού κόσμου, όπως η αυτόματη περίληψη, η ανάλυση συναισθήματος, η ανάλυση θεμάτων, η αναγνώριση λέξεων και μέρος του λόγου, η ανάλυση σχέσεων-λέξεων, και πολλά άλλα. Η Επεξεργασία Φυσικής Γλώσσας χρησιμοποιείται συνήθως για την ανάλυση κειμένου, την αυτόματη μετάφραση, και την αυτοματοποιημένη απάντηση ερωτήσεων

Το σύγχρονο θέμα που θα μας απασχολήσει είναι τα bot υποστήριξης πελατών.

Τι είναι ένα chatbot;

Ένα συμβατικό chatbot απαντά σε βασικά ερωτήματα πελατών και αιτήματα ρουτίνας με έτοιμες απαντήσεις.

Ένα αρνητικό των chatbots είναι ότι δεν μπορούν να αναγνωρίσουν ούτε ξεχωρίζουν πλήθος από ερωτήσεις, γιαυτό τον λόγο τα chatbots είναι πλέον εξοπλισμένα με τεχνολογίες τεχνητής νοημοσύνης και μηχανικής μάθησης για να μπορέσουν να ξεπεράσουν αυτά τα προβλήματα.

Τα chatbots πλέον είναι σε αρκετά καλό δρόμο, εκτός από την κατανόηση και την σύγκριση των μηνυμάτων των χρηστών, μπορούν επίσης

να δημιουργήσουν μόνα τους απαντήσεις σε ερωτήσεις χωρίς προαπαντημένες απαντήσεις.

Για παράδειγμα, το Reply.ai (<u>www.reply.ai</u>) έχει δημιουργήσει ένα προσαρμοσμένο chatbot που λειτουργεί με ML (Machine Learning) για να παρέχει υποστήριξη σε πελάτες. Τέλος, σύμφωνα με την εταιρεία, ένας μέσος οργανισμός μπορεί να ελέγξει σχεδόν το 40% των εισερχόμενων μηνυμάτων του μέσω της υποστήριξης που προσφέρει το εργαλείο της.

Πριν λοιπόν προχωρήσουμε και να δείξουμε τα <u>5 κύρια</u> βήματα για να φτιάξουμε ένα chatbot μέσω της NPL (Natural Language Processing-Επεξεργασία Φυσικής Γλώσσας) θα μιλήσουμε για το "γιατί" να φτιάξουμε ένα chatbot μέσω της Επεξεργασίας Φυσικής Γλώσσας.

Γιατί λοιπόν να φτιάξουμε ένα chatbot και μάλιστα μέσω της Επεξεργασίας Φυσικής Γλώσσας;

Πλέον τα Chatbots γίνονται όλο και πιο συνηθισμένα και είναι ένα ισχυρό εργαλείο για να προσελκύει διαδικτυακούς επισκέπτες όπου θα αλληλεπιδρά μαζί τους στη φυσική τους γλώσσα σαν να συνομιλούν μεταξύ τους δύο άνθρωποι. Παλιότερα οι ιστοσελίδες και γενικότερα οι εταιρείες είχαν ανθρώπινο δυναμικό για να αλληλεπιδρούν με τους πελάτες, για να λύνουν τις απορίες τους αλλά και να απαντούν σε οποιοδήποτε ερώτημα τους. Στις μέρες μας λοιπόν αυτό είναι σχεδόν αδύνατο μέχρι και ακατόρθωτο εφόσον οι ιστοσελίδες έχουν υψηλή επισκεψιμότητα όπου φτάνει μέχρι και εκατομμύρια επισκέπτες για μεγάλες ιστοσελίδες. Συμπεραίνουμε λοιπόν ότι για να υπάρχει δυναμικό προσωπικό ολόκληρο το εικοσιτετράωρο και καθημερινά θα ήταν αδύνατο. Επίσης οι εταιρείες θα είχα τεράστια έξοδα για την εκπαίδευση του προσωπικού αλλά και για την μισθοδοσία του. Έτσι τα chatbots θα έλυναν το ζήτημα αυτό ενεργώντας όλο το εικοσιτετράωρο και θα εξυπηρετούσε τους επισκέπτες του ιστότοπου χωρίς ανθρώπινη βοήθεια και παρουσία.

Αλλά ας περάσουμε στο πραγματικό ερώτημα, "γιατί να χρησιμοποιήσετε την Επεξεργασία Φυσικής Γλώσσας;"

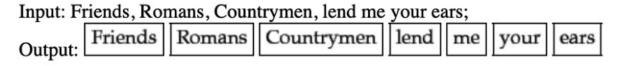
Η Επεξεργασία Φυσικής Γλώσσας είναι αυτή που επιτρέπει στα chatbots να κατανοούν τα μηνύματα των χρηστών/πελατών και να ανταποκρίνονται κατάλληλα. Για παράδειγμα όταν ένα chatbot λαμβάνει το μήνυμα "Γεια",

είναι η επεξεργασία φυσικής γλώσσας που του επιτρέπει να γνωρίζει ότι έχετε στείλει έναν τυπικό χαιρετισμό, όπου με την σειρά του το chatbot έχει την δυνατότητα μέσω της Τεχνητής Νοημοσύνης (AI – Artificial Intelligence) να ανταποκριθεί και να απαντήσει. Σε αυτή την περίπτωση το chatbot πιθανότατα θα απαντήσει πίσω επίσης με έναν χαιρετισμό. Όμως, χωρίς επεξεργασία φυσικής γλώσσας ένα chatbot δεν μπορεί να κάνει την ουσιαστική διάκριση μεταξύ των απαντήσεων "Γεια" και "Αντίο". Έτσι λοιπόν σε ένα chatbot χωρίς επεξεργασία φυσικής γλώσσας το "Γεια" και το "Αντίο" δεν θα είναι τίποτα περισσότερο από μηνύματα του χρήστη που βασίζονται σε κείμενο. Η Επεξεργασία Φυσικής Γλώσσας συμβάλλει στο να αναλύει τα μηνύματα του χρήστη που βασίζονται σε κείμενο έτσι ώστε η Τεχνητή Νοημοσύνη (AI) να μπορέσει να ανταποκριθεί όσο καλύτερα γίνεται.

Τα 5 κύρια βήματα για ένα Chatbot

1) Tokenize (Τμηματοποίηση)

Το Tokenization είναι η τεχνική για το τεμαχισμό του κειμένου σε κομμάτια, που ονομάζονται τμηματοποίηση και ταυτόχρονα αφαιρούμε ορισμένους χαρακτήρες, όπως τα σημεία στίξης. Αυτές οι τμηματοποιήσεις αντιπροσωπεύουν γλωσσικά το κείμενο.



Τμηματοποίηση μιας Πρότασης

2) Normalisation (Κανονικοποίηση)

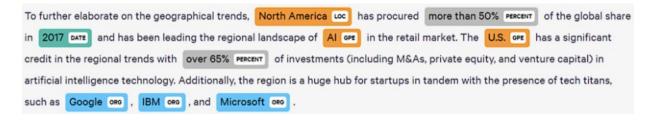
Η κανονικοποίηση επεξεργάζεται το κείμενο έτσι ώστε να ανακαλύψει τα συνηθισμένα ορθογραφικά λάθη που ενδέχεται να αλλάξουν την επιδιωκόμενη σημασία του αιτήματος του χρήστη. Ένα πολύ καλό παράδειγμα που εκτελεί κανονικοποίηση είναι το παρακάτω tweet.

Original: @user3419 nay lol y u say dat?&wat u doing 2day?

Post-normalization: No, why did you say that? What you doing today?

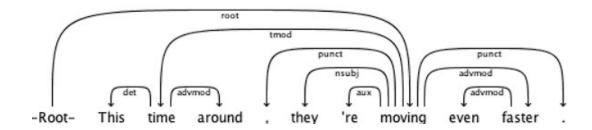
3) Recognising Entities (Αναγνώριση Οντοτήτων)

Αυτό το βήμα βοηθά το chatbot να προσδιορίσει πιο πράγμα μιλάει π.χ. είναι αντικείμενο ή χώρα ή αριθμός ή διεύθυνση του χρήστη. Παρατηρήστε στο παρακάτω παράδειγμα πως η Google, η IBM και η Microsoft είναι διακεκριμένα με μπλε χρώμα ως οργανώσεις (Organizations-ORG).



4) Dependency Parsing (Ανάλυση Εξαρτήσεων)

Σε αυτό το βήμα χωρίζουμε την πρόταση σε ουσιαστικά, ρήματα, αντικείμενα, σε κοινές φράσεις και σημεία στίξης. Αυτή η τεχνική βοηθά το μηχάνημα να εντοπίσει φράσεις και αυτό με τη σειρά του αντιλαμβάνεται τι θέλουν να πουν οι χρήστες.



5) Generation (Δημιουργία)

Τέλος έχουμε την δημιουργία της απόκρισης. Τα παραπάνω βήματα εμπίπτουν στο NLU (Natural Language Understanding- Κατανόηση Φυσικής Γλώσσας). Αυτά λοιπόν τα βήματα βοηθούν το chatbot να κατανοήσει τη σημασία της πρότασης που του δίνει ο χρήστης.

Το συγκεκριμένο βήμα (Δημιουργία) εμπίπτει στο NLG(Natural Language Generation-Δημιουργία Φυσικής Γλώσσας) όπου λαμβάνει σαν είσοδο την έξοδον των προηγούμενων βημάτων της Κατανόησης Φυσικής Γλώσσας και δημιουργεί έναν αριθμό προτάσεων όπου έχουν το ίδιο νόημα.

Οι προτάσεις που δημιουργούνται είναι παρόμοιες από την άποψη των παρακάτω παραδειγμάτων:

• Η σειρά των λέξεων:

"το φως της κουζίνας" είναι παρόμοιο με το "φως στην κουζίνα"

• Ενικός/Πληθυντικός:

"το φως της κουζίνας" είναι παρόμοιο με το "τα φώτα της κουζίνας"

• Ερώτηση:

"κλείστε την πόρτα" είναι παρόμοιο με "σας πειράζει να κλείσετε την πόρτα;"

• Αρνητικότητα:

"ενεργοποιήστε την τηλεόραση" είναι παρόμοιο με το "Μην ενεργοποιήσετε την τηλεόραση στις 21:00"

• Ευγένεια:

"ενεργοποιήστε την τηλεόραση" είναι παρόμοιο με "θα μπορούσατε να ενεργοποιήσετε την τηλεόραση"

Εν κατακλείδι,με βάση το πλαίσιο της ερώτησης του χρήστη, το chatbot μπορεί να απαντήσει με μία από τις παραπάνω επιλογές και ο χρήστης θα ήταν ικανοποιημένος. Σε πολλές περιπτώσεις λοιπόν, οι χρήστες δεν μπορούν να κάνουν διάκριση μεταξύ του chatbot και του ανθρώπου.

Θέμα 2:

(a)

```
thema2.pl

sentence(s(NP,VP)) --> noun_phrase(NP),verb_phrase(VP).
pp(pp(PREP,NP)) --> prep(PREP),noun_phrase(NP).
noun_phrase(np(D,N)) --> det(D),noun(N).
verb_phrase(vp(V,NP,PP)) --> verb(V),noun_phrase(NP),pp(PP).
det(det(the)) --> [the].
verb(verb(brought)) --> [brought].
noun(noun(waiter)) --> [waiter].
noun(noun(meal)) --> [meal].
noun(noun(table)) --> [table].
prep(prep(to)) --> [to].
```

M ϵ ?- sentence(X,N,[]).

Η γραμματική μας μπορεί να χρησιμοποιηθεί ως «**γεννήτορας προτάσεων**» που είναι συντακτικά ορθές, ανεξάρτητα αν βγάζουν νόημα, αυτό αφορά την σημασιολογία. Αλλά όπως βλέπου αναγνωρίζουμε την πρόταση στη πιο κάτω φωτογραφία:

```
File Edit Settings Run Debug Help

X = s(np(det(the), noun(waiter)), vp(verb(brought), np(det(the), noun(waiter)), pp(prep(to), np(det(the), noun(table))))),

N = [the, waiter, brought, the, waiter, to, the, table];

X = s(np(det(the), noun(waiter)), vp(verb(brought), np(det(the), noun(meal)), pp(prep(to), np(det(the), noun(waiter))))),

N = [the, waiter, brought, the, meal, to, the, waiter];

X = s(np(det(the), noun(waiter)), vp(verb(brought), np(det(the), noun(meal)), pp(prep(to), np(det(the), noun(meal))))),

N = [the, waiter, brought, the, meal, to, the, meal];

X = s(np(det(the), noun(waiter)), vp(verb(brought), np(det(the), noun(meal)), pp(prep(to), np(det(the), noun(table))))),

N = [the, waiter, brought, the, meal, to, the, table];

X = s(np(det(the), noun(waiter)), vp(verb(brought), np(det(the), noun(table)), pp(prep(to), np(det(the), noun(waiter))))),
```

```
thema2.pl
      sentence(s(NP,VP)) --> noun_phrase(NP), verb_phrase(VP).
      pp(pp(PREP,NP)) --> prep(PREP),noun_phrase(NP).
      noun phrase(np(D,N)) --> det(D),noun(N).
      verb_phrase(vp(V,NP,PP)) --> verb(V),noun_phrase(NP),pp(PP).
      det(det(the)) --> [the].
      verb(verb(brought)) --> [brought].
      noun(noun(waiter)) --> [waiter].
      noun(noun(meal)) --> [meal].
      noun(noun(table)) --> [table].
 10
      prep(prep(to)) --> [to].
SWI-Prolog -- d:/UniPcNotes/Code/EPEKSERGASIAFG/thema2.pl
                                                                 File Edit Settings Run Debug Help
Welcome to SWI-Prolog (threaded, 64 bits, version 8.2.1)
SWI-Prolog comes with ABSOLUTELY NO WARRANTY. This is free softwar
Please run ?- license. for legal details.
For online help and background, visit https://www.swi-prolog.org
For built-in help, use ?- help(Topic). or ?- apropos(Word).
?- sentence(S,[the,waiter,brought,the,meal,to,the,table],[]).
S = s(np(det(the), noun(waiter)), vp(verb(brought), np(det(the), n
oun(meal)), pp(prep(to), np(det(the), noun(table))))).
```

Απάντηση:

?-

S = s(np(det(the), noun(waiter)), vp(verb(brought), np(det(the), noun(meal)), pp(prep(to), np(det(the), noun(table))))).

Θέμα 3:

```
D:\UniPcNotes\Code\NPL\thema3p17168.pl - Sublime Text (UNREGISTERED)
File Edit Selection Find View Goto Tools Project Preferences Help
       thema3p17168.pl
       expression(Value) --> number(Value).
       expression(Value) --> number(X), [+], expression(V), {Value is X+V}.
      expression(Value) --> number(X), [-], expression(V), {Value is X-V}. expression(Value) --> number(X), [*], expression(V), {Value is X*V}. expression(Value) --> number(X), [/], expression(V), {V\=0, Value is X/V}.
       expression(Value) --> left parenthesis, expression(Value), right parenthesis.
       left_parenthesis --> ['('].
       right_parenthesis --> [')'].
 11
       number(X) --> digit(X).
 12
       number(Value) --> digit(X),number(Y),{numberofdigits(Y,N),Value is X*2^N+Y}.
 13
 14
      digit(0) --> [0].
      digit(1) --> [1].
 17
      numberofdigits(Y,1) :-
 18 Z is Y/2, Z<1,!
     numberofdigits(Y,N) :-
       Z is (Y - mod(Y,2))/2,
       numberofdigits(Z,N1), N is N1+1.
```

Στο θέμα αυτό πήραμε τον κώδικα που είχαμε ως ένδειξη και τον προσαρμόσαμε έτσι ώστε να ταιριάζει στις απαιτήσεις μας. Αρχικά κρατήσαμε μόνο τα δύο ψηφία 0 και 1 με το

```
digit(0) --> [0]. και
```

digit(1) --> [1]. στη συνέχει έχουμε αλλάξει τον αριθμό 10 όπου υπήρχε στον κώδικα μας με το ψηφίο 2 για το δυαδικό σύστημα, αυτό έγινε για τις γραμμές 20 και 12.

Επίσης στην γραμμή 5 στην διαίρεση διορθώσαμε το $\{V/=0, Value is X/V\}$ σε $\{V\setminus=0, Value is X/V\}$.

Το αποτέλεσμα:

```
SWI-Prolog -- d:/UniPcNotes/Code/NPL/thema3p17168.pl
File Edit Settings Run Debug Help
Welcome to SWI-Prolog (threaded, 64 bits, version 8.2.1)
SWI-Prolog comes with ABSOLUTELY NO WARRANTY. This is free software.
Please run ?- license. for legal details.
For online help and background, visit https://www.swi-prolog.org
For built-in help, use ?- help(Topic). or ?- apropos(Word).
?- expression(B1,['(',1,1,+,'(',1,1,-,1,1,')',')'],[]).
B1 = 3 .
?- expression(B2,['(',1,0,*,'(',1,1,*,1,0,')',')'],[]).
B2 = 12
?- expression(B3,['(',1,0,-,'(',1,1,-,1,0,')',')'],[]).
B3 = 1.
?- expression(B4,['(',1,0,-,'(',1,1,*,1,0,')',')'],[]).
?- expression(B5,['(',1,0,-,'(',1,1,/,1,0,')',')'],[]).
B5 = 0.5.
```

Θέμα i:

Υλοποίηση ενός ChatBot σε Java.

Στο θέμα i επέλεξα να φτιάξω ένα μικρό chatbot εφόσον με απασχόλησε και στο θέμα 1. Ελπίζω να το βρείτε ενδιαφέρον. Το chatbot γράφτηκε σε java.

Ακολουθεί ο κώδικάς με φωτογραφικό υλικό: (Ο κώδικας παραδόθηκε και στο φάκελο mychatbot)

Main.java:

```
ChatBot.java
Main.java
       import java.util.Set;
       import java.util.Scanner;
       import static java.lang.System.exit;
       public class Main {
           public static void main(String args[]) {
               ChatBot mybot = new ChatBot( name: "MariaBot");
               System.out.println("Welcome. I am " + mybot.name);
               System.out.println("What is your name? ");
               Scanner scanner = new Scanner(System.in);
               String username = scanner.nextLine();
               System.out.println("Hello " + username + "." + " " + "Ask me a question please!");
               mybot.addQuestion( erotisi: "old you", apantisi: "Never ask a lady for her age!!");
               mybot.addQuestion( erotisi: "capital Cyprus", apantisi: "Nicosia!!");
               mybot.addQuestion( erotisi: "hello", apantisi: "Hello to you!!");
               mybot.addQuestion( erotisi: "capital Greece", apantisi: "Athens!!");
               mybot.addQuestion( erotisi: "weather today", apantisi: "Sunny...I suppose ");
               mybot.addQuestion( erotisi: "how are you", apantisi: "I feel very well");
               mybot.addQuestion( erotisi: "bye", apantisi: "See you soon");
```

ChatBot.java

ΠΑΡΑΔΕΙΓΜΑ ΕΚΤΕΛΕΣΗΣ CHATBOT:

