

# Sub4All: Υποσύστημα Αναγνώρισης Φωνής

ΙΕΛ

Τελικός Έλεγχος Sub4All, 8 Μαΐου 2008

1

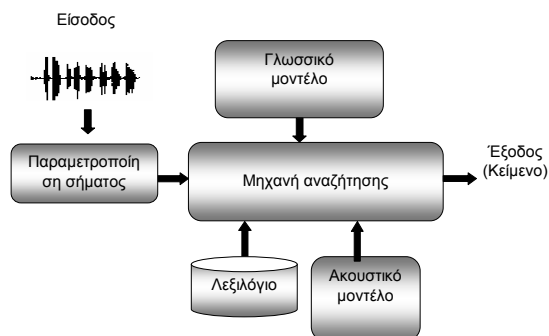
## Υποσύστημα Αναγνώρισης Φωνής

- Συνοπτική περιγραφή συστήματος
  - Εξαγωγή παραμέτρων
  - Φωνητική Μετεγγραφή Ελληνικών
  - Ακουστική Μοντελοποίηση
  - Στατιστική γλωσσική μοντελοποίηση
  - Μηχανή αναζήτησης
- Αξιολόγηση του Συστήματος

Τελικός Έλεγχος Sub4All, 8 Μαΐου 2008

2

## Συνοπτική περιγραφή συστήματος



- Large Vocabulary Recognition (LVR)
- Young (1996) + Voice-in, IEA
- Το εν λόγω σύστημα έχει χρησιμοποιηθεί ως πυρήνας του προϊόντος

Ηλεκτρονικός Λογονόμος  
Τελικός Έλεγχος Sub4All, 8 Μαΐου 2008

3

## Αξιολόγηση του Συστήματος

- Η αναγνώριση φωνής χρησιμοποιείται για την αντιστοίχιση και ευθυγράμμιση συνδέοντας ένα κείμενο με το εισερχόμενο ηχητικό σήμα σηματοδοτώντας τα χρονικά όρια των λέξεων.
- Το κείμενο μπορεί να περιέχει αποσπάσματα που δεν υπάρχουν στο ηχητικό σήμα ή αντιστρόφως, καθώς και άγνωστες λέξεις για το σύστημα αναγνώρισης.
- Προσπαθήσαμε να δημιουργήσουμε μια αντιστοίχιση και ευθυγράμμιση ανθεκτική σε τέτοιου είδους «ατέλειες» χωρίς απαιτούμενη ανθρώπινη παρέμβαση.
- Σημαντικό ρόλο σε αυτή την προσπάθεια παίζουν η ταχύτητα και η ακρίβεια της αναγνώρισης φωνής.

Τελικός Έλεγχος Sub4All, 8 Μαΐου 2008

4

## Αξιολόγηση του Συστήματος

- Η ταχύτητα της διαδικασίας αναγνώρισης φωνής ορίζεται από τα πόσα δευτερόλεπτα ηχητικού σήματος επεξεργάζονται ανά δευτερόλεπτο του χρόνου του επεξεργαστή.
- Η ακρίβεια της αναγνώρισης φωνής εκφράζεται ως ποσοστό λάθους αναγνωρισμένων λέξεων (word error rate- WER).

Τελικός Έλεγχος Sub4All, 8 Μαΐου 2008

5

## Αξιολόγηση του Συστήματος

- Η ευθυγράμμιση του παραχθέντος κειμένου από το σύστημα αναγνώρισης φωνής με το δεδομένο κείμενο των διαλόγων βασίζεται στη γνωστή μεθοδολογία της δυναμικής στρέβλωσης του χρόνου DTW (Dynamic Time Warping).
- Η DTW χρησιμοποιεί στοιχεία του δυναμικού προγραμματισμού ώστε να επιλύσει το πρόβλημα της βέλτιστης αντιστοίχισης δύο κειμένων με ταυτόσημο περιεχόμενο αλλά λεκτικές διαφορές.
- Στην παρούσα εργασία ο αλγόριθμος που χρησιμοποιήθηκε αντιμετώπισε με επιτυχία τόσο τις περιπτώσεις λανθασμένων λέξεων όσο και τις περιπτώσεις λέξεων που προστέθηκαν ή αφαιρέθηκαν με αυθαίρετο τρόπο.

Τελικός Έλεγχος Sub4All, 8 Μαΐου 2008

6

## Αξιολόγηση του Συστήματος

- Πιο συγκεκριμένα το ποσοστό επιτυχίας του συστήματος αναγνώρισης φωνής κυμάνθηκε στο 70% περίπου.
- Η χρησιμοποίηση του DTW επέτρεψε την αποτελεσματική εκμετάλλευση αυτού του αποτελέσματος ώστε να αντιστοιχιστούν κατάλληλοι χρόνοι εκφοράς και στο υπόλοιπο 30% των λέξεων που δεν υποστηρίζεται από την αναγνώριση φωνής.
- Μετρήθηκε ότι η ακρίβεια των εκτιμηθέντων χρόνων εκφοράς αυτού του 30% των λέξεων ξεπέρασε τα 200 msec.
- Έτσι οι μικρές ανακρίβειες στην εκτίμηση του χρόνου εκφοράς αυτών των λέξεων περνούν απαρατήρητες στον χρήστη του συστήματος.

Τελικός Έλεγχος Sub4All, 8 Μαΐου 2008

7

## Πειραματικά αποτελέσματα

### ΑΡΧΙΚΟ ΚΕΙΜΕΝΟ

[Πέραν τούτου, μιλώ όχι για αναθεώρηση όλης της ερμηνευτικής διάταξης,] αλλά μόνο για αναθεώρηση, η οποία αφορά τη συγκεκριμενοποίηση της πλειοψηφίας που απαιτείται και δη την αναφορά ρητώς στην παράγραφο 3 του άρθρου 28.

ΑΝΔΡΕΑΣ ΛΟΒΕΡΔΟΣ: Κύριε Υπουργέ, με συγχωρείτε πάρα πολύ.

ΠΡΟΕΔΡΟΣ (Ιωάννης Τραγάκης): Σας παρακαλώ, κυριε Λοβέρδο, να μην διακόπτετε.

ΑΝΔΡΕΑΣ ΛΟΒΕΡΔΟΣ: Λέγονται πράγματα που δεν είναι έτσι, κύριε Πρόεδρε.

### ΜΕΤΑΣΧΗΜΑΤΙΣΜΕΝΟ ΚΕΙΜΕΝΟ

[Πέραν τούτου, μιλώ όχι για αναθεώρηση όλης της ερμηνευτικής διάταξης,] αλλά μόνο για αναθεώρηση, η οποία αφορά τη συγκεκριμενοποίηση της πλειοψηφίας που απαιτείται και δη την αναφορά ρητώς στην παράγραφο τρία του άρθρου είκοσι οκτώ,

ΑΝΔΡΕΑΣ ΛΟΒΕΡΔΟΣ: Κύριε Υπουργέ, με συγχωρείτε πάρα πολύ.

ΠΡΟΕΔΡΟΣ (Ιωάννης Τραγάκης): Σας παρακαλώ, κυριε Λοβέρδο, να μην διακόπτετε.

ΑΝΔΡΕΑΣ ΛΟΒΕΡΔΟΣ: Λέγονται πράγματα που δεν είναι έτσι, κύριε Πρόεδρε.

Τελικός Έλεγχος Sub4All, 8 Μαΐου 2008

8

## Πειραματικά αποτελέσματα

### ΔΕΙΓΜΑ ΤΗΣ ΕΞΟΔΟΥ ΤΟΥ ΣΥΣΤΗΜΑΤΟΣ ΑΝΑΓΝΩΡΙΣΗΣ ΦΩΝΗΣ

πρέπει [0,1231][68070,0 , 560,0]  
να [0,0486][68630,0 , 290,0]  
τροποποιήσουμε [0,1508][68920,0 , 790,0]  
ρητά [0,1748][69710,0 , 430,0]  
τη [0,3120][70140,0 , 1000,0]  
διάταξη [0,1283][71140,0 , 560,0]  
που [0,0874][71700,0 , 150,0]  
λέει [0,1213][71850,0 , 370,0]  
ότι [0,1456][72220,0 , 1150,0]  
τα [0,0406][73370,0 , 90,0]  
πολιτικά [0,0921][73460,0 , 490,0]  
κόμματα [0,1016][73950,0 , 430,0]  
έχουν [0,2793][74380,0 , 310,0]  
δικαίωμα [0,1901][74690,0 , 780,0]  
χρηματοδότησης [0,4588][75470,0 , 870,0]  
όπου [0,1320][76340,0 , 170,0]

Τελικός Έλεγχος Sub4All, 8 Μαΐου 2008

9

## Πειραματικά αποτελέσματα

### ΑΠΟΤΕΛΕΣΜΑ ΣΥΣΤΗΜΑΤΟΣ ΕΥΘΥΓΡΑΜΜΙΣΗΣ (Alignment)

| Εκφωνηθείσα Λέξη   | Αρχή εκφώνησης (msec) | Τέλος εκφώνησης (msec) |
|--------------------|-----------------------|------------------------|
| αλλά               | 1250                  | 1490                   |
| μόνο               | 1560                  | 1800                   |
| για                | 1800                  | 1930                   |
| αναθεώρηση,        | 1960                  | 2560                   |
| η                  | 2560                  | 2602                   |
| οποία              | 2602                  | 2810                   |
| αφορά              | 3020                  | 3320                   |
| τη                 | 4090                  | 4210                   |
| συγκεκριμενοποίηση | 4450                  | 5530                   |
| της                | 5530                  | 5650                   |
| πλειοψηφίας        | 5650                  | 6230                   |
| που                | 6230                  | 6351                   |
| απαιτείται         | 6351                  | 6756                   |
| και                | 6756                  | 6878                   |
| δη                 | 6878                  | 6959                   |
| την                | 6959                  | 7080                   |
| αναφορά            | 7160                  | 7580                   |
| ρητώς              | 7750                  | 8050                   |
| στην               | 8120                  | 8360                   |
| παράγραφο          | 8360                  | 8830                   |
| τρία               | 8830                  | 9070                   |
| του                | 9070                  | 9160                   |
| άρθρου             | 9640                  | 10000                  |
| είκοσι             | 10000                 | 10254                  |
| οκτώ,              | 10254                 | 10424                  |

Τελικός Έλεγχος Sub4All, 8 Μαΐου 2008

10