



Εθνικόν και Καποδιστριακόν
Πανεπιστήμιον Αθηνών

Τμήμα Πληροφορικής και Τηλεπικοινωνιών

Φωνητικές Διεπαφές Χρήστη-Τεχνολογίες Φωνής

Ενότητα 7: Ευχρηστία Φωνητικών Διεπαφών
Χρήστη

Γεώργιος Κουρουπέτρογλου

koupe@di.uoa.gr



Περιεχόμενα ενότητας

Μέθοδοι προσδιορισμού φωνητικών διεπαφών
χρήστη

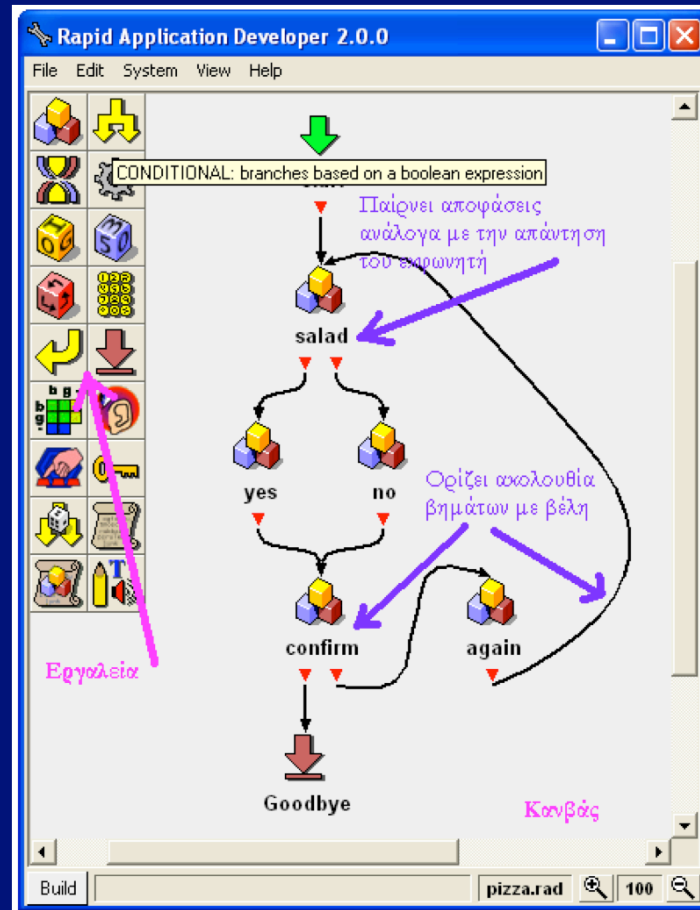
Developing Spoken Dialogue Systems Using the CSLU Toolkit

Center for Spoken Language Understanding

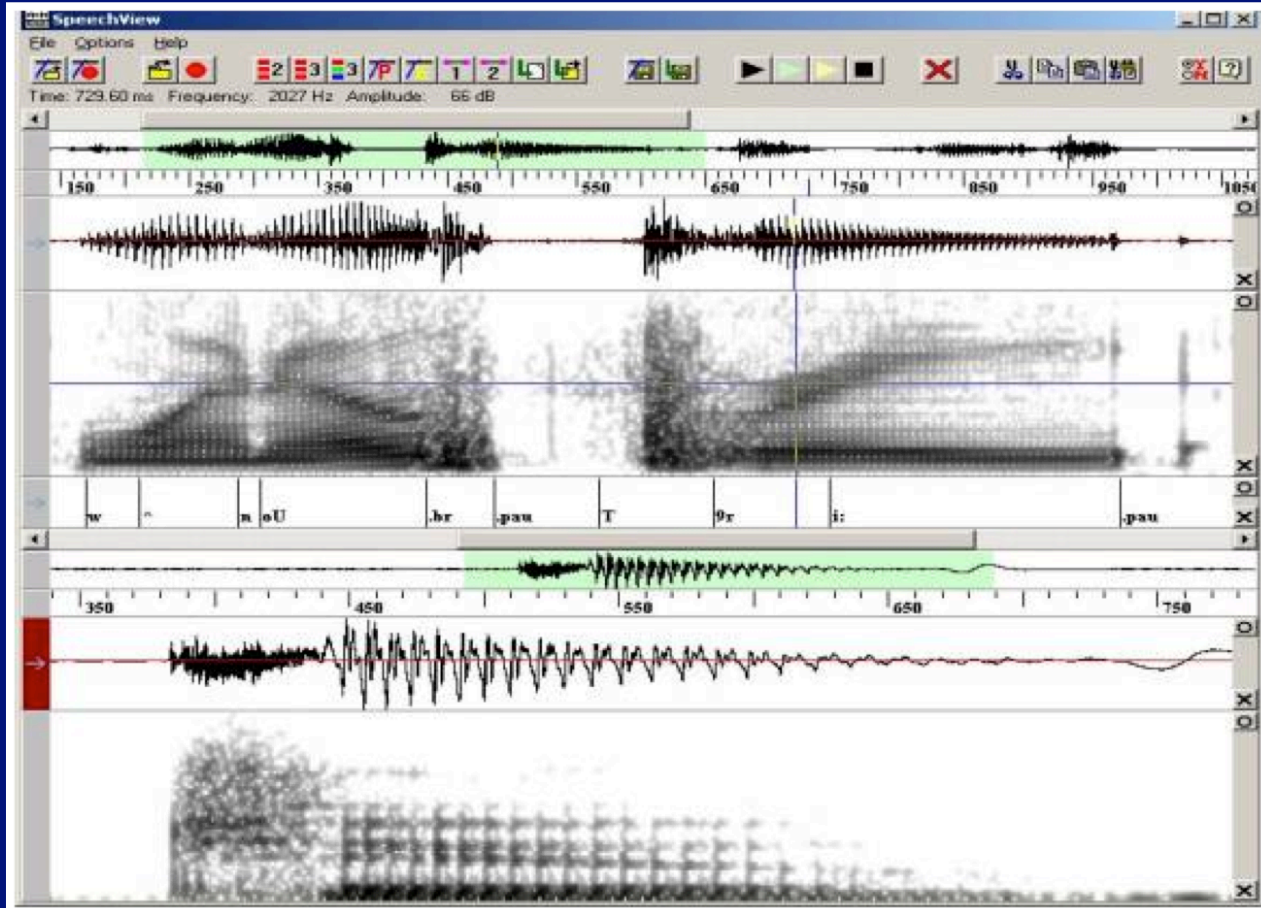
- RAD (Rapid Application Developer):
 - a graphically based authoring environment
 - supports directed dialogues using a finite state-based dialogue model
- Speech Viewer
- Baldi Sync



Περιβάλλον του RAD



Speech Viewer



Baldi Sync

bsync - Baldi with natural speech

File Options Help

Text to align with:
Parakalo, doste enan arithmo apo to miden eos to enea

TTS English

Labels Align

arakalo doste enan arithmo apo to miden eos to enea

ph @ k a l o u . p a u d s t e n a n @ n e a r i t h m o a p o . p a u t o m i d e n e o s t o e n e a .

h



Αντιστοίχιση συμβόλων IPA σε WordBet

IPA	WordBet	π.χ.		IPA	WordBet	π.χ.		IPA	WordBet	π.χ.	
ɪ	i	Beat		οι	>i	Boy		ʌ	ʌ	Thin	
ɪ	I	Bit		αΥ	αU	House		σ	s	Sit	
E	E	Get		οΥ	οU	Boat		Σ	S	Shoe	
ə	@	Bat		π	pe ph	Pie		η	h	Hall	
↔	&c	Sofa		τ	te th	Tie		ω	v	Veil	
υ	u	Boot		κ	ke kh	Cry		Δ	D	They	
Υ	U	Book		β	be b	Button		ζ	z	Zero	
θ	^	The		δ	de d	Dye		Z	Z	Beige	
ο	>	Bought		γ	ge g	Guy		δZ	dZc dZ	Judge	
A	A	Father		μ	m	Me		λ	l	Long	
εΠ	3r	Bird		ν	n	No		©	9r	Run	
™	&r	Teacher		N	N	Sing		φ	j	Yes	
ει	ei	Buy		φ	f	Fine		ω	w	Went	
αI	αI	Buy									
■ Φωνήεντα			■ Δίφθογγοι			■ Σύμφωντα					

Βασικά Αντικείμενα RAD

Υπο-Γενικό Συνθήκη διάλογος

Αναγνώριση Χαρακτήρων και ψηφίων

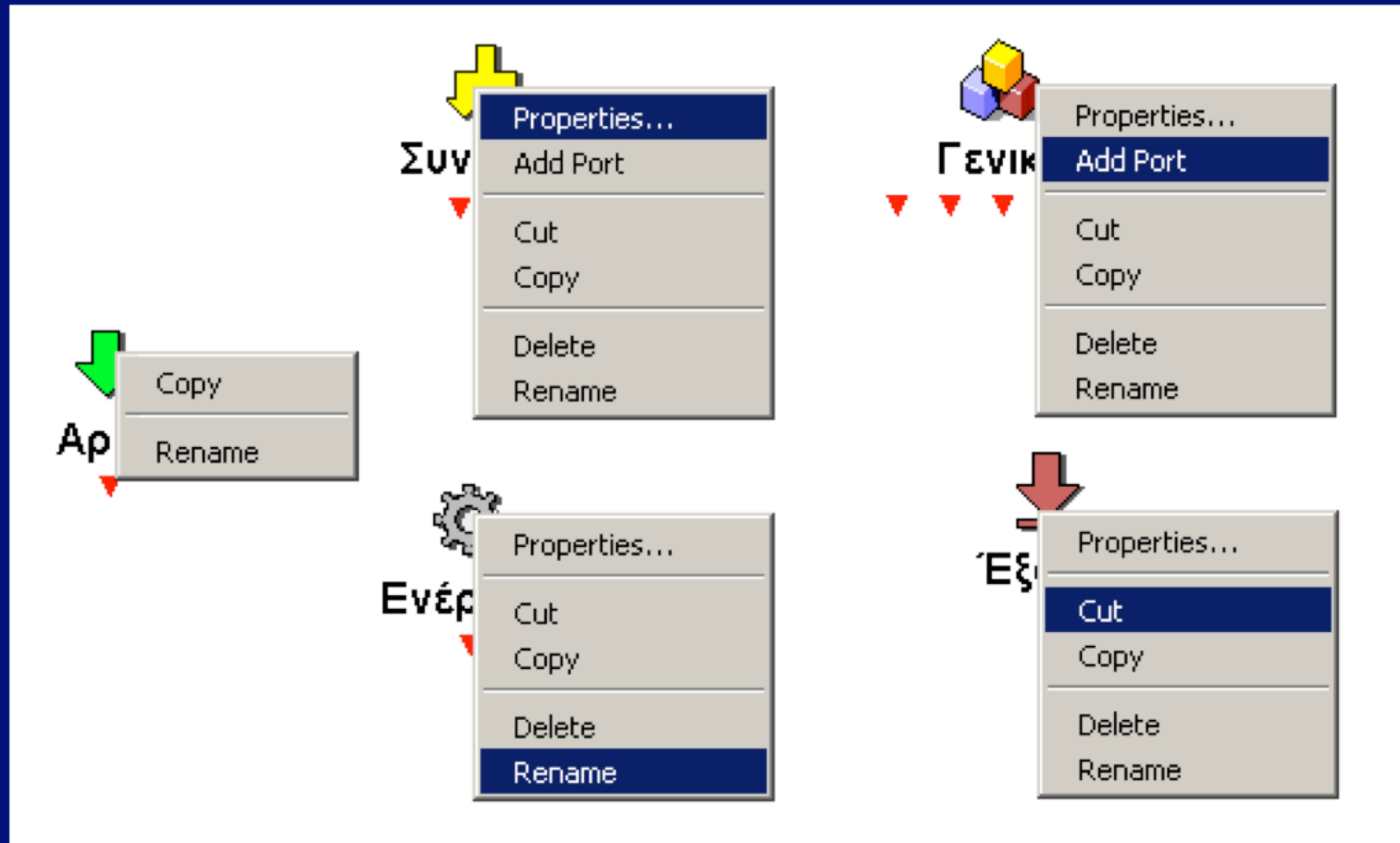


Αναγνώριση ψηφίων

Είσοδος DTFM

Τέλος Διαλόγου

Ρυθμίσεις αντικειμένων RAD



Ρυθμίσεις Γενικού Αντικειμένου

RAD prompt: input_num

TTS | Markup | Recorded | Recognition | DTMF | Misc | On Enter | On Exit | Tucker-Maxon

Prompt

Parakalo, doste enan arithmo apo to miden eos to enea

TTS <- Save... Load...

File

C:/rad/application/prompts/mess_01.sob

Edit... Select... Play Clear

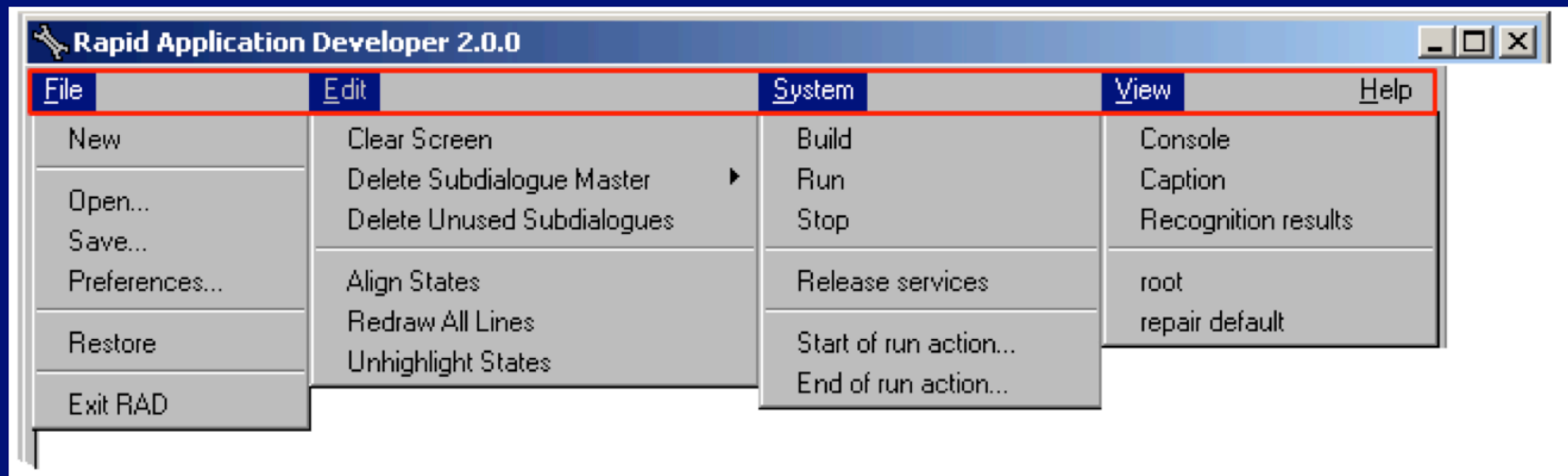
Captioning

Show captioning Save geometry

Use: recorded Face:

OK Cancel

Μενού επιλογών της RAD



Γενικές Προτιμήσεων στην RAD

Global Preferences

General | Audio | Recog/DTMF | TTS | Dictionaries | Editing | About | Tucker-Maxon

Display

Animated Face: NullFace

Captioning Highlight Active Console VU meter

Recognition Results Trace Canvas

Misc

Repair Run Loop

Barge-in

Backup

Paths

Beep File: C:/Program Files/CSLU/Toolkit/2.0/apps/rad/endbeep.wav Browse...

Default Repair: C:/Program Files/CSLU/Toolkit/2.0/apps/rad/Repair/default1.rad Browse...

Data Capture & Playback

Capture directory: C:/Documents and Settings/user/.rad/data

Playback from file: live Browse...

Set Default Get Default OK Cancel



Ρυθμίσεις Μικροφώνου στις Γενικές Προτιμήσεις στην RAD

Audio

Audio Device

Type	Hostname	Name	Devid	
audio	local	*	*	Use Default

Audio Parameters

Beep before recognition

Maximum Record Duration: 5 s

Leading Silence Duration: 5000 ms

Trailing Silence Duration: 300 ms

Record Backoff: 150 ms

Voice Detection Threshold: 38 %stdev Calibrate

Master Volume > Microphone > Advanced > Microphone Boost

Εγχειρίδιο Χρήσης (1 από 7)

ΠΡΟΛΟΓΟΣ

1. Ανάπτυξη διαλογικού συστήματος χρησιμοποιώντας το CSLU Toolkit
2. Γενικές οδηγίες για το RAD
3. Μεταγλώττιση & εκτέλεση εφαρμογής
4. Προβλήματα αναγνώρισης
5. Εντολές (prompts)

Εγχειρίδιο Χρήσης (2 από 7)

6. Λεξιλόγιο αναγνώρισης

6.1 Πρόσθεση νέας πόρτας και σύνδεση με νέες καταστάσεις

6.2 Διακλάδωση

6.3 Επιβεβαίωση

7. Υποδιάλογοι

7.1 Δημιουργία υποδιαλόγου

7.2 Επιστροφή από υποδιάλογο

7.3 Αποθήκευση υποδιαλόγου

Εγχειρίδιο Χρήσης (3 από 7)

8. Αναγνώριση ψηφίων – Αναγνώριση αλφαριθμητικών – DTMF είσοδος
 - 8.1 Αναγνωριστής ψηφίων
 - 8.2 Αναγνωριστής αλφαριθμητικών
 - 8.3 DTMF είσοδος
9. Εναπομένουσα διάρκεια σιγής
10. Γραμματική αναγνώρισης
 - 10.1 Καθορισμός γραμματικής αναγνώρισης
 - 10.2 Σύνθετοι κανόνες γραμματικής αναγνώρισης

Εγχειρίδιο Χρήσης (4 από 7)

11. Έξοδος ομιλίας στο RAD

11.1 Τροποποίηση του ρυθμού (rate) και του τόνου (pitch) της ομιλίας

12. Χρήση Tcl/TK εντολών στο RAD

12.1 Χρήση Tcl κώδικα για επεξεργασία αναγνωρισμένων λέξεων

13. Προσθήκη νέας προφοράς και δημιουργία δυναμικού αναγνωριστή

13.1 Προσθήκη προφοράς

13.2 Δυναμική αναγνώριση

14. Σύνδεση μιας εφαρμογής με βάση δεδομένων χρησιμοποιώντας το TclODBC



Εγχειρίδιο Χρήσης (5 από 7)

- Tutorial 1: Basic functions in RAD: The Pizza application.
- Tutorial 2: Developing basic functions for a Student Information System.
- Tutorial 3: The default repair sub-dialogue.
- Tutorial 4: Sub-dialogues.
- Tutorial 5: Digit recognition.
- Tutorial 6: DTMF input.

Εγχειρίδιο Χρήσης (6 από 7)

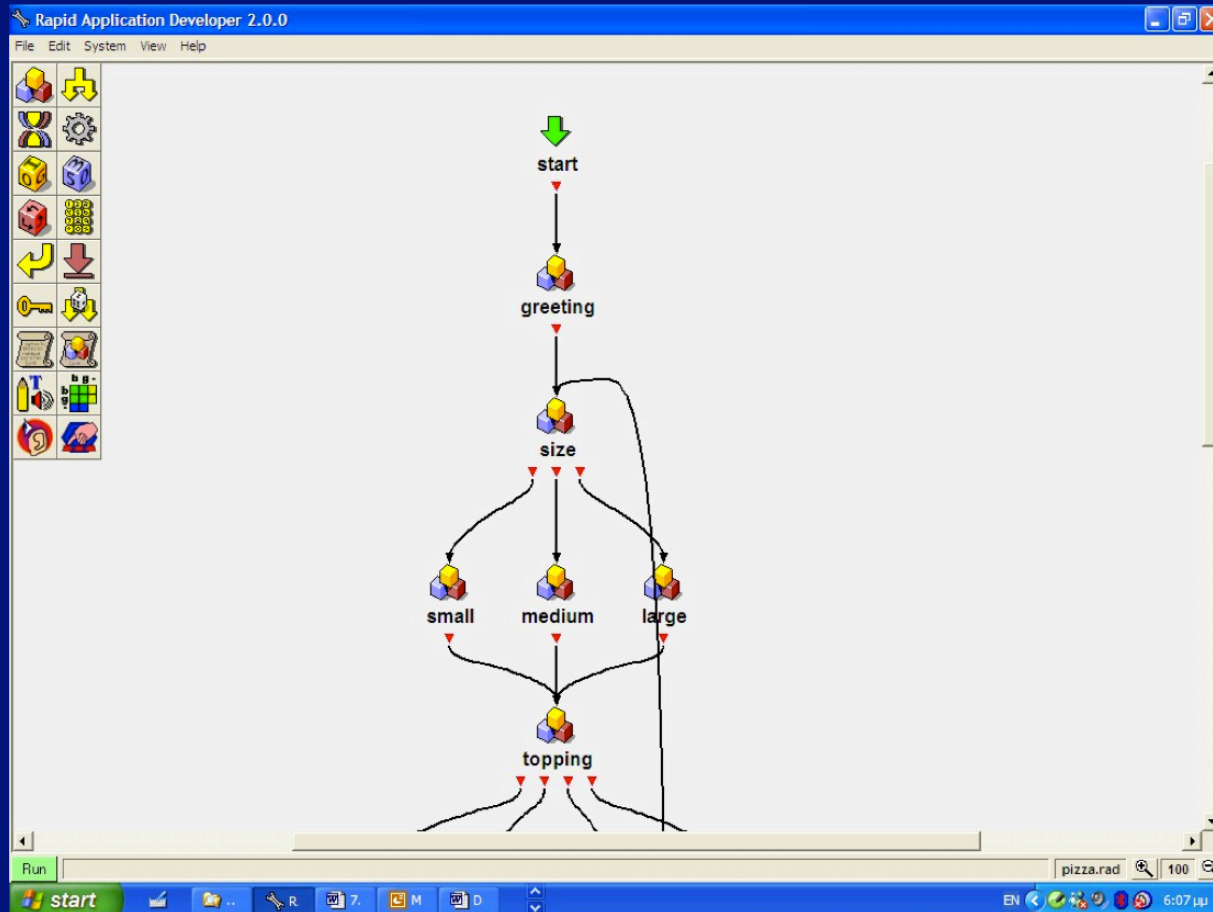
- Tutorial 7: Alpha-digit recognition.
- Tutorial 8: Recognition grammars.
- Tutorial 9: Speech output in RAD.
- Tutorial 10: Using Tcl in RAD.
- Tutorial 11: Improving recognition and creating a dynamic recogniser.
- Tutorial 12: Linking a RAD application to a database using TclODBC.

Εγχειρίδιο Χρήσης (7 από 7)

Ο κώδικας των Tutorials 4-12 βρίσκεται στην ιστοσελίδα του μαθήματος στο e-class (στη θέση: Έγγραφα)

Οδηγίες εγκατάστασης του CSLUtoolkit βρίσκονται στην ιστοσελίδα του μαθήματος στο e-class (στη θέση: Έγγραφα)

Tutorial 1: Basic functions in RAD: The Pizza application



Tutorial 1: Basic functions in RAD

- Start and loading the pizza application:
File>Examples>pizza
- Compiling and running an application
Build>Run
- Prompts
- Recognition vocabulary:
add a pronunciation
Add a new word> new state>new linking
- Branching
- Verification



Άσκηση 7.1

Στο Tutorial 4:

- α) αλλάξτε το μήνυμα καλωσορίσματος ώστε να είναι στην ελληνική γλώσσα και
- β) προσθέσατε ένα νέο όνομα στη λίστα των φοιτητών που αναγνωρίζονται

Άσκηση 7.2 (1 από 2)

Συνδυάστε τα Tutorials 2,3,5 και 6 ώστε να αναπτύξετε μια εφαρμογή που θα υλοποιεί ένα τυπικό διάλογο όπως ο παρακάτω:

1. System: Welcome to the Student System Main Menu.
2. Please say your username.
3. User: Liz (ή
4. System: Thank you.
5. Now please say your four-digit pin.
6. User: 6 5 6 6 (οποιονδήποτε συνδυασμό 4 από τα 10 ψηφία) .

Άσκηση 7.2 (2 από 2)

7. System: Was that 6 5 6 6?
8. User: Yes.
9. System: The system provides details on students, courses and reports.
10. For students say students or press 1, for courses say courses or press 2, for reports say reports or press 3.

Στο βήμα 6 χρησιμοποιείστε το Εργαλείο Αναγνώρισης Ψηφίων και στο βήμα 10 είσοδο DTFM

Άσκηση 7.3 (1 από 4)

1. System: Welcome to the Student System Main Menu.
2. Please say your username.
3. User: Liz
4. System: Thank you.
5. Now please say your four-digit pin.
6. User: 6 5 6 6 (οποιοδήποτε συνδυασμό 4 από τα 10 ψηφία) .

Άσκηση 7.3 (2 από 4)

7. System: Was that 6 5 6 6?
8. User: Yes.
9. System: The system provides details on students, courses and reports.
10. For students say students or press 1, for courses say courses or press 2, for reports say reports or press 3.
11. User: Students.
12. System: This is Student Details.

Άσκηση 7.3 (3 από 4)

13. Say "View details" to view existing student details.
14. Say "Add details" to add new student details.
15. User: View details.
16. System: Viewing student details.
17. What is the student ID?
18. User: 96050918.
19. System: I have the following details: student John Scott, course code DK003, at stage 1.
20. Would you like any more information?
21. User: Courses.

Άσκηση 7.3 (4 από 4)

Παράδειγμα στοιχείων στη βάση δεδομένων:

	studentID	firstname	lastname	coursecode	stage
1	96050918	john	scott	dk003	1
2	96069783	david	wilson	dk005	award
3	96317159	rosemary	blackbum	dk003	1
4	96416781	jennifer	black	dk003	award

Άσκηση 8

- Υλοποιείστε την άσκηση 7.3 σε VoiceXML στην πλατφόρμα IBM Webshpere Voice Toolkit

Συμβουλευτείτε τα σύντομα tutorials για VoiceXML που βρίσκονται την ιστοσελίδα του μαθήματος στο e-class (στη θέση: Έγγραφα)

Οδηγίες εγκατάστασης του VoiceWebSphere βρίσκονται στην ιστοσελίδα του μαθήματος στο e-class (στη θέση: Έγγραφα)

Τέλος Ενότητας

Ευχρηστία Φωνητικών Διεπαφών Χρήστη

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στο πλαίσιο του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αθηνών**» έχει χρηματοδοτήσει μόνο την αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «**Εκπαίδευση και Δια Βίου Μάθηση**» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Σημειώματα

Σημείωμα Ιστορικού Εκδόσεων Έργου

Το παρόν έργο αποτελεί την έκδοση 1.0.

Σημείωμα Αναφοράς

Copyright Εθνικών και Καποδιστριακών Πανεπιστημίων Αθηνών 2015, Γεώργιος Κουρουπέτρογλου 2015. Γεώργιος Κουρουπέτρογλου.
«Φωνητικές Διεπαφές Χρήστη-Τεχνολογίες Φωνής. Ευχρηστία Φωνητικών Διεπαφών Χρήστη». Έκδοση: 1.0. Αθήνα 2015. Διαθέσιμο από τη δικτυακή διεύθυνση: <http://opencourses.uoa.gr/courses/DI37/>.

Σημείωμα Αδειοδότησης

Το παρόν υλικό διατίθεται με τους όρους της άδειας χρήσης Creative Commons Αναφορά, Μη Εμπορική Χρήση Παρόμοια Διανομή 4.0 [1] ή μεταγενέστερη, Διεθνής Έκδοση. Εξαιρούνται τα αυτοτελή έργα τρίτων π.χ. φωτογραφίες, διαγράμματα κ.λ.π., τα οποία εμπεριέχονται σε αυτό και τα οποία αναφέρονται μαζί με τους όρους χρήσης τους στο «Σημείωμα Χρήσης Έργων Τρίτων».



[1] <http://creativecommons.org/licenses/by-nc-sa/4.0/>

Ως **Μη Εμπορική** ορίζεται η χρήση:

- που δεν περιλαμβάνει άμεσο ή έμμεσο οικονομικό όφελος από την χρήση του έργου, για το διανομέα του έργου και αδειοδόχο
- που δεν περιλαμβάνει οικονομική συναλλαγή ως προϋπόθεση για τη χρήση ή πρόσβαση στο έργο
- που δεν προσπορίζει στο διανομέα του έργου και αδειοδόχο έμμεσο οικονομικό όφελος (π.χ. διαφημίσεις) από την προβολή του έργου σε διαδικτυακό τόπο

Ο δικαιούχος μπορεί να παρέχει στον αδειοδόχο ξεχωριστή άδεια να χρησιμοποιεί το έργο για εμπορική χρήση, εφόσον αυτό του ζητηθεί.

Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)

μαζί με τους συνοδευόμενους υπερσυνδέσμους.

Σημείωμα Χρήσης Έργων Τρίτων (1 από 2)

Οι φωτογραφίες που περιέχονται στην παρουσίαση αποτελούν πνευματική ιδιοκτησία τρίτων. Απαγορεύεται η αναπαραγωγή, αναδημοσίευση και διάθεσή τους στο κοινό με οποιονδήποτε τρόπο χωρίς τη λήψη άδειας από τους δικαιούχους.

Σημείωμα Χρήσης Έργων Τρίτων (2 από 2)

Η δομή και οργάνωση της παρουσίασης, καθώς και το υπόλοιπο περιεχόμενο, αποτελούν πνευματική ιδιοκτησία του συγγραφέα και του Πανεπιστημίου Αθηνών και διατίθενται με άδεια Creative Commons Αναφορά Μη Εμπορική Χρήση Παρόμοια Διανομή Έκδοση 4.0 ή μεταγενέστερη.