



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ Η/Υ & ΠΛΗΡΟΦΟΡΙΚΗΣ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Αυτόματη Παραγωγή και Αξιολόγηση Ασκήσεων και Χρήση Παιχνιδοποίησης σε Ευφυή Συστήματα Διδασκαλίας

Φωτεινή Γριβοκωστοπούλου

Επιβλέπων: Ιωάννης Χατζηλυγερούδης, Αναπληρωτής
Καθηγητής

Πάτρα, Ιούλιος 2016



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

UNIVERSITY OF PATRAS
SCHOOL OF ENGINEERING
DEPARTMENT OF COMPUTER ENGINEERING AND INFOMATICS

DOCTORAL DISSERTATION

Automated Generation and Assessment of Exercises and Use of Gamification in Intelligent Tutoring Systems

Foteini Grivokostopoulou

Supervisor: Ioannis Hatzilygeroudis, Associate

Professor

Patra, July 2016

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
ΠΟΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ Η/Υ & ΠΛΗΡΟΦΟΡΙΚΗΣ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

**Αυτόματη Παραγωγή και Αξιολόγηση Ασκήσεων και
Χρήση Παιχνιδοποίησης σε Ευφυή Συστήματα
Διδασκαλίας**

Φωτεινή Γριβοκωστοπούλου

Επταμελής Εξεταστική Επιτροπή:

Ιωάννης Χατζηλυγερούδης, Αναπληρωτής Καθηγητής, Τμήμα Μηχανικών Η/Υ & Πληροφορικής του Πανεπιστημίου Πατρών, (Επιβλέπων)

Σπυρίδων Λυκοθανάσης, Καθηγητής, Τμήμα Μηχανικών Η/Υ & Πληροφορικής του Πανεπιστημίου Πατρών, (Μέλος της Τριμελούς Επιτροπής)

Κυριάκος Σγάρμπας, Αναπληρωτής Καθηγητής, Τμήμα Ηλεκτρολόγων Μηχανικών & Τεχνολογίας Υπολογιστών του Πανεπιστημίου Πατρών (Μέλος της Τριμελούς Επιτροπής)

Μαρία Βίρβου, Καθηγήτρια, Τμήμα Πληροφορικής του Πανεπιστημίου Πειραιώς

Βασίλειος Μεγαλοοικονόμου, Καθηγητής, Τμήμα Μηχανικών Η/Υ & Πληροφορικής του Πανεπιστημίου Πανεπιστήμιο Πατρών

Αθανάσιος Τσακαλίδης, Καθηγητής, Τμήμα Μηχανικών Η/Υ & Πληροφορικής του Πανεπιστημίου Πατρών

Δημήτριος Καλλές, Επίκουρος Καθηγητής, Σχολή Θετικών Επιστημών και Τεχνολογίας του Ελληνικού Ανοικτού Πανεπιστημίου (Ε.Α.Π)

Ευχαριστίες

Ολοκληρώνοντας την εκπόνηση της διδακτορικής διατριβής μου θα ήθελα να ευχαριστήσω μια σειρά από ανθρώπους που με βοήθησαν καθ' όλη την διάρκεια της πορείας μου.

Καταρχάς, ευχαριστώ θερμά τον επιβλέποντα της διδακτορικής μου διατριβής, τον Αναπληρωτή Καθηγητή κ. Ιωάννη Χατζηλυγερούδη, που απετέλεσε τον καθοδηγητή μου σε όλα τα στάδια της διατριβής, για την υποστήριξη που μου παρείχε, την πολύτιμη βοήθεια, την αμέριστη συμπαράσταση και την εμπιστοσύνη που μου έδειξε όλα αυτά τα χρόνια. Θα ήθελα επίσης να ευχαριστήσω θερμά τον Καθηγητή κ. Σπυρίδωνα Λυκοθανάση και τον Αναπληρωτή Καθηγητή κ. Κυριάκο Σγάρμπα για τη συμβουλευτική υποστήριξη που παρείχαν κατά την εκπόνηση της διατριβής. Επίσης, ευχαριστώ θερμά τα μέλη της επταμελούς συμβουλευτικής επιτροπής κ. Αθανάσιο Τσακαλίδη, κ. Μαρία Βίρβου, κ. Δημήτριο Καλλέ και κ. Βασίλειο Μεγαλοοικονόμου για την τιμή που μου έκαναν.

Επιπλέον, θα ήθελα να ευχαριστήσω τους συνεργάτες μου και φίλους μου στην Ομάδα Τεχνητής Νοημοσύνης, Ισίδωρο Περίκο και Κώστα Κόβα, για την άριστη συνεργασία που είχαμε όλα αυτά τα χρόνια.

Τέλος, οφείλω ένα μεγάλο ευχαριστώ στην οικογένεια μου και στους φίλους μου για την υποστήριξη που μου παρείχαν καθ' όλη την διάρκεια των σπουδών μου.

Φωτεινή Γριβοκωστοπούλου

Ιούλιος 2016

Στον άνθρωπο που είναι πάντα δίπλα μου
και με στηρίζει

Περιεχόμενα

Περίληψη	19
Abstract	23
1 Εισαγωγή	25
1.1 Εισαγωγικά στοιχεία.....	25
1.2 Συνεισφορά της διατριβής	28
1.3 Διάρθρωση της διατριβής	30
2 Ευφυή Συστήματα Διδασκαλίας Τεχνητής Νοημοσύνης	33
2.1 Εισαγωγή	33
2.2 Σχετικές Εργασίες	34
2.3 Σημασιολογικός Ιστός.....	36
2.3.1 Τεχνολογίες Σημασιολογικού Ιστού και Οντολογίες	39
2.4 Ευφύες Σύστημα Διδασκαλίας Τεχνητής Νοημοσύνης.....	41
2.4.1 Αρχιτεκτονική Συστήματος.....	41
2.4.2 Εκπαιδευτική Διαδικασία.....	44
2.4.3 Αξιολόγηση Εκπαιδευόμενου.....	46
2.4.4 Οντολογία AITS.....	48
2.4.5 Διδασκαλία Αλγορίθμων Αναζήτησης με το Σύστημα AITS	50
2.4.6 Πειραματικά Αποτελέσματα συστήματος AITS.....	56
2.5 Ευφυή Συστήματα Διδασκαλίας Λογικής.....	61
2.5.1 Βασικές Έννοιες	62
2.5.2 Ευφύες Σύστημα Μετατροπής Κατηγορηματικής Λογικής σε Προτασιακή Μορφή	65
2.5.3 Ευφύες Σύστημα Μετατροπής Προτάσεων Κατηγορηματικής Λογικής σε Ισοδύναμες	85
2.6 Σύστημα Αυτόματης Απόδειξης Θεωρημάτων.....	93
2.6.1 Βασικές Έννοιες Αποδεικτικής Θεωρίας	95
2.6.2 Στρατηγικές Ελέγχου Εξαγωγής Συμπερασμάτων.....	97
2.6.3 Αποδείκτης Θεωρημάτων EX-ACT-P	98
2.6.4 Αλληλεπιδραστικός Αποδείκτης Θεωρημάτων	104
2.6.5 Πειραματικά Αξιολόγηση	109
3 Ημιαυτόματη Παραγωγή Ασκήσεων	113
3.1 Εισαγωγή	113

3.2	Σχετικές Εργασίες	114
3.3	Ημιαυτόματη Παραγωγή ασκήσεων λογικής	116
3.3.1	Μηχανισμός Αυτόματης Δημιουργίας Ασκήσεων ΚΛΠΤ	116
3.3.2	Πειραματικά Αποτελέσματα	118
3.4	Ημιαυτόματη Παραγωγή Αλληλεπιδραστικών Ασκήσεων για αλγορίθμους	120
3.4.1	Βασικές Έννοιες	120
3.4.2	Μηχανισμός Αυτόματης Δημιουργίας Ασκήσεων Αλγορίθμων	122
3.4.3	Πειραματικά Αποτελέσματα	125
4	Αυτόματος Προσδιορισμός Δυσκολίας Ασκήσεων	127
4.1	Εισαγωγή	127
4.2	Σχετική Έρευνα	128
4.3	Αυτόματος Προσδιορισμός Δυσκολία Ασκήσεων Μετατροπής Φυσικής Γλώσσας σε Κατηγορηματική Λογική	130
4.3.1	Μηχανισμός Αυτομάτου Προσδιορισμού Δυσκολίας Ασκήσεων ΦΓ σε ΚΛΠΤ	131
4.3.2	Πειραματικά Αποτελέσματα	140
4.4	Αυτόματος Προσδιορισμός Δυσκολία Ασκήσεων Μετατροπής Κατηγορηματικής Λογικής σε Προτασιακή Λογική	144
4.4.1	Προσδιορισμός δυσκολίας με βάση την δομή της πρότασης ΚΛΠΤ (FOL-DEA)	145
4.4.2	Προσδιορισμός δυσκολίας με βάση τα βήματα της διαδικασίας μετατροπής (PROC-DEA)	147
4.4.3	Πειραματικά αποτελέσματα των δύο προσεγγίσεων	149
4.4.4	Συνδυαστική προσέγγιση μέσω ευφυούς συστήματος κανόνων	152
4.4.5	Πειραματικά Αποτελέσματα Συστήματος Προσδιορισμού Δυσκολίας	156
4.5	Νευρο-ασαφείς και Νευρο-συμβολικές προσεγγίσεις για τον προσδιορισμό της δυσκολίας ασκήσεων αλγορίθμων αναζήτησης	157
4.5.1	Προσδιορισμός Παραμέτρων Ασκήσεων	158
4.5.2	Νευρο-ασαφής Προσέγγιση	159
4.5.3	Νευρωνικοί κανόνες	163
4.5.4	Πειραματικά Αποτελέσματα	164
5	Αυτόματη Αξιολόγηση Απαντήσεων	167
5.1	Εισαγωγή	167
5.2	Σχετικές Εργασίες	168
5.3	Μηχανισμός Αυτόματης Αξιολόγησης Απαντήσεων	171
5.4	Μηχανισμός Αυτόματης Αξιολόγησης Απαντήσεων σε Ασκήσεις Λογικής	172
5.4.1	Γενικός Μηχανισμός Αυτόματης Αξιολόγησης Απαντήσεων	172

5.4.2	Αυτόματη Αξιολόγηση Απαντήσεων σε Ασκήσεις Μετατροπής ΦΓ σε ΚΛΠΤ	174
5.4.3	Μηχανισμός Αυτόματης Αξιολόγησης Απαντήσεων σε Ασκήσεις Μετατροπής ΚΛΠΤ σε ΠΜ	180
5.4.4	Πειραματική Αξιολόγηση Μηχανισμού Αυτόματης Αξιολόγησης Απαντήσεων σε Ασκήσεις Λογικής	186
5.5	Μηχανισμός Αυτόματης Αξιολόγησης Απαντήσεων σε Ασκήσεις Αλγορίθμων Αναζήτησης	194
5.5.1	Γενική Μεθοδολογία Αυτόματης Αξιολόγησης Απαντήσεων	194
5.5.2	Μηχανισμός Ανίχνευσης Λαθών	196
5.5.3	Μοντελοποίηση Απάντησης.....	198
5.5.4	Σημαντικότητα των Λαθών.....	200
5.5.5	Μηχανισμός Αυτόματης Βαθμολόγησης	201
5.5.6	Πειραματική Αξιολόγηση Μηχανισμού Αυτόματης Βαθμολόγησης Απαντήσεων	205
6	Εκπαιδευτικά Παιχνίδια	211
6.1	Εισαγωγή	211
6.2	Σχετικές Εργασίες	213
6.3	Παιδαγωγικές Προσεγγίσεις και μαθησιακοί στόχοι	214
6.4	Εκμάθηση Αλγορίθμων με το Παιχνίδι Edu-PacMan	215
6.4.1	Εκπαιδευτικές δραστηριότητες του παιχνιδιού	215
6.4.2	Πειραματική Αξιολόγηση Edu-Pacman	219
6.5	Εκμάθηση Ικανοποίησης περιορισμών με το Εκπαιδευτικό Παιχνίδι Coloring-Map-Game	224
6.5.1	Το πεδίο της Ικανοποίησης Περιορισμών.....	224
6.5.2	Εκπαιδευτικές δραστηριότητες.....	227
6.5.3	Πειραματικά Αποτελέσματα	229
6.6	Παιχνιδοποίηση σε Εικονικό Περιβάλλον Μάθησης	232
7	Συμπεράσματα και Μελλοντικές Κατευθύνσεις	237
	Βιβλιογραφία	Error! Bookmark not defined.

Κατάλογος Εικόνων

Εικόνα 1.1 Βασική Αρχιτεκτονική Ευφρών Συστημάτων Διδασκαλίας	27
Εικόνα 2.1 Διασπρωματομένη προσέγγιση του Σημασιολογικού Ιστού	37
Εικόνα 2.2 Αρχιτεκτονική Συστήματος Διδασκαλίας Τεχνητής Νοημοσύνης	42
Εικόνα 2.3 Δεντρική Αναπαράσταση του Εκπαιδευτικού Υλικού	45
Εικόνα 2.4 Γραφική Απεικόνιση της Οντολογίας.....	48
Εικόνα 2.5 Γραφική Απεικόνιση της κλάση Theory	49
Εικόνα 2.6 Δεντρική Αναπαράσταση των Αλγορίθμων Αναζήτησης	51
Εικόνα 2.7 Οπτικοποίηση της λειτουργίας του A^* σε ένα παράδειγμα	53
Εικόνα 2.8 Άσκηση αξιολόγησης στον αλγόριθμο A^*	54
Εικόνα 2.9 Άσκηση εξάσκησης πάνω στον αλγόριθμο αναζήτησης A^*	55
Εικόνα 2.10 Ανάδραση του συστήματος μετά την απάντηση του εκπαιδευόμενου	56
Εικόνα 2.11 Η διαδικασία της πειραματικής μελέτης για το σύστημα AITS	57
Εικόνα 2.12 Πειραματικά αποτελέσματα pretest και posttest πάνω στους αλγορίθμους αναζήτησης	59
Εικόνα 2.13 Αποτελέσματα Ερωτηματολογίου.....	60
Εικόνα 2.14 Η Αρχιτεκτονική του Συστήματος FOL to CF.....	67
Εικόνα 2.15 Δεντρική Αναπαράστασης μιας έκφρασης ΚΛΠΤ.....	71
Εικόνα 2.16 Συντακτική Ανάλυση της έκφρασης ΚΛΠΤ	71
Εικόνα 2.17 Βοήθεια του συστήματος FOLtoCF	74
Εικόνα 2.18 Κύκλος Learning Analytics	75
Εικόνα 2.19 Δημιουργία νέας άσκησης.....	76
Εικόνα 2.20 Γραφική Αναπαράσταση των εκπαιδευόμενων ανά επίπεδο δυσκολίας της άσκησης.....	77
Εικόνα 2.21 Τύποι των λαθών των εκπαιδευόμενων	78
Εικόνα 2.22 Οι απαντήσεις των εκπαιδευόμενων ανά επίπεδο δυσκολίας	78
Εικόνα 2.23 Διαδικασία Πειραματικής Αξιολόγησης του Συστήματος FOLtoCF	80
Εικόνα 2.24 Μέση τιμή για κάθε ομάδα για το Pretest και το Posttest	81
Εικόνα 2.25 Αποτελέσματα χρόνου επίλυση του Posttest.....	83
Εικόνα 2.26 Χρόνος εξοικείωσης των φοιτητών με το Σύστημα.....	83
Εικόνα 2.27 Αποτελέσματα Ερωτηματολογίου για το Σύστημα FOLtoCF	84
Εικόνα 2.28 Αρχιτεκτονική του Ευφρούς Συστήματος Μετατροπής Κατηγορηματικής Λογικής σε Ισοδύναμες	87
Εικόνα 2.29 Αλληλεπίδραση με το σύστημα FOL-EQU.....	89
Εικόνα 2.30 Μέση τιμή για κάθε ομάδα για το Pretest και το Posttest	92
Εικόνα 2.31 Διαδικασία Απόδειξης μιας Πρότασης.....	95
Εικόνα 2.32 Κατηγοριοποίηση Στρατηγικών Ελέγχου	98

Εικόνα 2.33. Αρχιτεκτονική αποδείκτη θεωρημάτων EX-ACT-P.....	99
Εικόνα 2.34 Αποδείκτης Θεωρημάτων EX-ACT-P.....	103
Εικόνα 2.35 Μηχανισμός Αλληλεπίδρασης i-ACT-P.....	105
Εικόνα 2.36 Πρώτο είδος βοήθειας Error-specific feedback.....	106
Εικόνα 2.37 Δεύτερο είδος βοήθειας Procedural feedback.....	107
Εικόνα 2.38 Αλληλεπιδραστικός Αποδείκτης Θεωρημάτων i-ACTP.....	109
Εικόνα 2.39 Πειραματικά αποτελέσματα pretest και posttest πάνω στους αλγορίθμους αναζήτησης ...	111
Εικόνα 2.40 Ο μέσος αριθμός των λαθών της πειραματικής ομάδας και της ομάδας ελέγχου	111
Εικόνα 3.1 Μηχανισμός Παραγωγής Ασκήσεων Λογικής	117
Εικόνα 3.2 Παραγωγή FOL εκφράσεων	118
Εικόνα 3.3 Παραδείγματα δυαδικών δέντρων	121
Εικόνα 3.4 Πλήρες δυαδικό δέντρο	121
Εικόνα 3.5 Παράδειγμα παραγωγής μια άσκησης του HC αλγορίθμου.....	124
Εικόνα 3.6 Παράδειγμα παραγωγής μια άσκησης του αλγορίθμου A^*	125
Εικόνα 3.7 Παράδειγμα Αξιολόγηση Ασκήσεων σχετικά με την ποιότητα.....	126
Εικόνα 4.1. Αρχιτεκτονική του συστήματος προσδιορισμού της δυσκολίας ασκήσεων της διαδικασίας NLtoFOL	132
Εικόνα 4.2. Βασική Δομή του Βασισμένου στην ΚΛΠΤ ΕΣ.....	134
Εικόνα 4.3. Παράδειγμα Κανόνα Έμπειρου Συστήματος βασισμένο στην ΚΛΠΤ	135
Εικόνα 4.4. Αρχιτεκτονική Ευφυούς Συστήματος προσδιορισμού δυσκολίας (DEIS) FOLtoCF ασκήσεων.	154
Εικόνα 4.5. Διάγραμμα συμπερασμού.....	154
Εικόνα 4.6. Παράδειγμα προσδιορισμού τιμών παραμέτρων δυσκολίας μέσω του FOL-Analyser και του FOLtoCF Analyser	155
Εικόνα 4.7 Μεταβλητές Εισόδου.....	159
Εικόνα 4.8 Αρχιτεκτονική ANFIS.....	160
Εικόνα 4.9 Μοντέλο ANFIS προσδιορισμού Δυσκολίας Ασκήσεων	161
Εικόνα 4.10 Συναρτήσεις Συμμετοχής για την μεταβλητή πλήθος των κόμβο του δέντρου.....	162
Εικόνα 4.11 (α) Ένα νευρωνικό κανόνα ως μια μονάδα adaline (b) Συνάρτηση ενεργοποίησης (c) Λεκτική αναπαράσταση νευρωνικού κανόνα	164
Εικόνα 5.1. Αρχιτεκτονική μηχανισμού αυτόματης αξιολόγησης απαντήσεων	171
Εικόνα 5.2 Γενικός Μηχανισμός Αξιολόγησης Απαντήσεων	173
Εικόνα 5.3 Παράδειγμα δεντρικής αναπαράστασης μιας έκφρασης ΚΛΠΤ	174
Εικόνα 5.4 Αυτόματη Αξιολόγηση Ασκήσεων ΦΓ σε ΚΛΠΤ.....	180
Εικόνα 5.5 Αυτόματη Αξιολόγηση Ασκήσεων ΚΛΠΤ σε ΠΜ.....	186
Εικόνα 5.6 Διάγραμμα διασποράς μεταξύ διδάσκοντα και αυτόματου μηχανισμού.....	187
Εικόνα 5.7 Παραδείγματα βαθμολογημένων ασκήσεων από διδάσκοντα και το σύστημα.....	188

<i>Εικόνα 5.8 Παραδείγματα χαμηλών επιδόσεων βαθμολογημένων ασκήσεων από διδάσκοντα και σύστημα.....</i>	<i>189</i>
<i>Εικόνα 5.9 Παραδείγματα μεσαίων επιδόσεων βαθμολογημένων ασκήσεων από διδάσκοντα και σύστημα.....</i>	<i>189</i>
<i>Εικόνα 5.10 Παραδείγματα καλών επιδόσεων βαθμολογημένων ασκήσεων από διδάσκοντα και σύστημα.....</i>	<i>190</i>
<i>Εικόνα 5.11 Διάγραμμα διασποράς μεταξύ διδάσκοντα και αυτόματου μηχανισμού.....</i>	<i>192</i>
<i>Εικόνα 5.12 Παραδείγματα βαθμολογημένων ασκήσεων από διδάσκοντα και σύστημα.....</i>	<i>192</i>
<i>Εικόνα 5.13 Απαντήσεις των εκπαιδευόμενων στο ερωτηματολόγιο.....</i>	<i>194</i>
<i>Εικόνα 5.14 Γενική Μεθοδολογία Αυτόματης Αξιολόγησης Απαντήσεων.....</i>	<i>195</i>
<i>Εικόνα 5.15 Παράδειγμα αλληλεπιδραστικής άσκησης εφαρμογής της κατά πλάτος αναζήτησης.....</i>	<i>204</i>
<i>Εικόνα 5.16 Διάγραμμα διασποράς automatic marker σε σχέση με human marker.....</i>	<i>206</i>
<i>Εικόνα 5.17 Διαγράμματα διασποράς για (α) CorpusA, (β) CorpusB, (γ) CorpusC.....</i>	<i>209</i>
<i>Εικόνα 6.1 Αναθεωρημένη Ταξινόμια Bloom.....</i>	<i>215</i>
<i>Εικόνα 6.2 Παράδειγμα οπτικοποίησης.....</i>	<i>216</i>
<i>Εικόνα 6.3 Αλληλεπίδραση με τον χρήστη και καθορισμός επόμενης κίνησης.....</i>	<i>217</i>
<i>Εικόνα 6.4 Οπτικοποίηση αλγόριθμου Κατά Βάθος αναζήτησης.....</i>	<i>217</i>
<i>Εικόνα 6.5 Ανάλυση της λειτουργίας του αλγόριθμου μέσω της κίνησης του φαντάσματος.....</i>	<i>218</i>
<i>Εικόνα 6.6 Πειραματική διαδικασία του edu-Pacman Game.....</i>	<i>220</i>
<i>Εικόνα 6.7 Αποτελέσματα Pretest και Posttest για κάθε ομάδα.....</i>	<i>221</i>
<i>Εικόνα 6.8 Χρόνος Επίλυσης Posttest.....</i>	<i>222</i>
<i>Εικόνα 6.9 Παράδειγμα Χρωματισμού Χάρτη.....</i>	<i>226</i>
<i>Εικόνα 6.10 Γραφος Περιορισμών.....</i>	<i>226</i>
<i>Εικόνα 6.11 Το παιχνίδι coloring Map Game.....</i>	<i>228</i>
<i>Εικόνα 6.12 Αποτελέσματα Pretest και Posttest.....</i>	<i>230</i>
<i>Εικόνα 6.13 Αποτελέσματα Ερωτηματολογίου.....</i>	<i>232</i>
<i>Εικόνα 6.14 Παιχνιδοποίηση σε εικονικό Περιβάλλον VR-ALG.....</i>	<i>234</i>

Κατάλογος Πινάκων

Πίνακας 2.1 Προτάσεις AITS με βάση το σκορ στο θέμα	47
Πίνακας 2.2 Προτάσεις AITS με βάση το βαθμό στην ενότητα.....	47
Πίνακας 2.3 Ιδιότητες της κλάσης Theory.....	49
Πίνακας 2.4 Ιδιότητες της Κλάσης Examples.....	49
Πίνακας 2.5 Περιγραφικά μέτρα pretest και posttest στους αλγορίθμους αναζήτησης.....	58
Πίνακας 2.6 Αποτελέσματα της ανάλυσης των ερωτηματολογίων του AITS.....	60
Πίνακας 2.7 Περιγραφικά Μέτρα Pretest και Posttest.....	81
Πίνακας 2.8 t-Test Results.....	82
Πίνακας 2.9 Περιγραφικά Μέτρα Pretest και Posttest του FOL-EQU	91
Πίνακας 2.10 Απαντήσεις Ερωτηματολογίου για το σύστημα FOL-EQU.....	92
Πίνακας 2.11 Κατηγορίες λαθών	105
Πίνακας 2.12 Αποτελέσματα pretest και posttest πάνω στην απόδειξη θεωρημάτων.....	110
Πίνακας 3.1 Αξιολόγηση της ποιότητας των παραγόμενων ΚΛΠΤ.....	119
Πίνακας 3.2 Αξιολόγηση της καταλληλότητας των παραγόμενων ΚΛΠΤ.....	119
Πίνακας 3.3 Αξιολόγηση Ασκήσεων σε σχέση με την ποιότητά τους	125
Πίνακας 3.4 Αξιολόγηση Ασκήσεων σε σχέση με τη καταλληλότητά τους.	125
Πίνακας 4.1 Κανόνες για τον προσδιορισμό της δυσκολίας από το Βασισμένο στην ΚΛΠΤ Έμπειρο Σύστημα.	134
Πίνακας 4.2 Παραδείγματα προσδιορισμού του επιπέδου δυσκολίας από το Βασισμένο στην ΚΛΠΤ Έμπειρο Σύστημα.....	136
Πίνακας 4.3 Κανόνες προσδιορισμού επιπέδου δυσκολίας από το Βασισμένο στην ΦΓ ΕΣ.	139
Πίνακας 4.4 Πειραματικά αποτελέσματα μηχανισμού εκτίμησης δυσκολίας ασκήσεων	141
Πίνακας 4.5 Πειραματικά αποτελέσματα εκτίμησης δυσκολίας για κάθε μια κατηγορία	142
Πίνακας 4.6 Υπολογισμός Cohen's Kappa	143
Πίνακας 4.7 Κανόνες της Βασισμένης στην ΚΛΠΤ (FOL-DEA) προσέγγισης.....	145
Πίνακας 4.8 Κανόνες της προσέγγισης PROC-DEA.	148
Πίνακας 4.9 Παραδείγματα προτάσεων ΚΛΠΤ και τα αντίστοιχα επίπεδα δυσκολίας των δυο μεθόδων	149
Πίνακας 4.10 Πειραματικά αποτελέσματα της προσέγγισης FOL-DEA.....	149
Πίνακας 4.11 Πειραματικά αποτελέσματα για την προσέγγιση FOL-DEA για κάθε κατηγορία.....	150
Πίνακας 4.12 Πειραματικά αποτελέσματα της προσέγγισης PROC-DEA.....	150
Πίνακας 4.13 Πειραματικά αποτελέσματα για την προσέγγιση PROC -DEA για κάθε κατηγορία.....	150
Πίνακας 4.14 Cohen's kappa για τις προσεγγίσεις FOL-DEA και PROC-DEA	151
Πίνακας 4.15 Παραδείγματα μη σωστών προσδιορισμών Δυσκολίας Ασκήσεων.....	151

Πίνακας 4.16 Συνδυαστική προσέγγιση για τον προσδιορισμό των επιπέδων δυσκολίας	152
Πίνακας 4.17 Πειραματικά αποτελέσματα DEIS.....	156
Πίνακας 4.18 Μετρικές Αξιολόγησης του DEIS.....	156
Πίνακας 4.19 Πειραματικά Αποτελέσματα νευροασαφή μοντέλου	165
Πίνακας 4.20 Πειραματικά Αποτελέσματα Νευροκανόνων	165
Πίνακας 4.21. Cohen's kappa μετρική για την νευροασαφή και νευροσυμβολική προσέγγιση	166
Πίνακας 5.1 Κατηγοριοποίηση λαθών	175
Πίνακας 5.2 Σχήματα Βαθμολόγησης Ασκήσεων Μετατροπής ΦΓ σε ΚΛΠΤ.....	176
Πίνακας 5.3 Βάρη Λαθών για τις ατομικές εκφράσεις.....	177
Πίνακας 5.4. Βάρη λαθών για τους ποσοδείκτες.	178
Πίνακας 5.5 Βάρος λάθους σχετικής θέσης (order error) ποσοδεικτών.....	179
Πίνακας 5.6. Κατηγοριοποίηση Λαθών ΚΛΠΤ σε ΠΜ.....	182
Πίνακας 5.7. Σχήμα βαθμολόγησης	184
Πίνακας 5.8 Το βάρος κάθε είδους λάθους στις ατομικές εκφράσεις	184
Πίνακας 5.9 Πειραματικά αποτελέσματα μηχανισμού αυτόματης αξιολόγησης.....	190
Πίνακας 5.10 Πειραματικά αποτελέσματα μηχανισμού αυτόματης αξιολόγησης.....	191
Πίνακας 5.11 Πειραματικά αποτελέσματα μηχανισμού αυτόματης αξιολόγησης.....	193
Πίνακας 5.12 Πειραματικά αποτελέσματα μηχανισμού αυτόματης αξιολόγησης.....	193
Πίνακας 5.13 Παραδείγματα βαθμολογημένων απαντήσεων από το σύστημα	204
Πίνακας 5.14 Αποτελέσματα Συσχέτισης.....	207
Πίνακας 5.15 Πειραματικά αποτελέσματα μηχανισμού αυτόματης αξιολόγησης.....	209
Πίνακας 5.16 Πειραματικά αποτελέσματα μηχανισμού αυτόματης αξιολόγησης.....	210
Πίνακας 6.1 Πειραματικά αποτελέσματα Posttest για το Pacman.....	221
Πίνακας 6.2: Αποτελέσματα ερωτηματολογίου.	222
Πίνακας 6.3 Πειραματικά Αποτελέσματα Posttest	231
Πίνακας 6.4 Ερωτήσεις ερωτηματολογίου.	231

Περίληψη

Τα τελευταία χρόνια, η ραγδαία ανάπτυξη της τεχνολογίας και του διαδικτύου δεν θα μπορούσε να αφήσει ανεπηρέαστο τον τομέα των συστημάτων ηλεκτρονικής μάθησης (e-learning). Η χρήση του διαδικτύου άλλαξε τον τρόπο με τον οποίο παρέχεται η μάθηση στους εκπαιδευόμενους. Μια βασική και δημοφιλής κατηγορία εκπαιδευτικών συστημάτων είναι τα Ευφυή Συστήματα Διδασκαλίας-ΕΣΔ (Intelligent Tutoring Systems-ITSs), κύριο χαρακτηριστικό των οποίων είναι η παροχή εξατομικευμένης μάθησης στους εκπαιδευόμενους, με χρήση μεθόδων τεχνητής νοημοσύνης και κατάλληλων παιδαγωγικών προσεγγίσεων.

Στα πλαίσια της παρούσας διδακτορικής διατριβής, αναπτύξαμε νέες μεθοδολογίες και μηχανισμούς στο τομέα των ευφύων εκπαιδευτικών συστημάτων. Πιο συγκεκριμένα, αναπτύχθηκαν μεθοδολογίες για την ημιαυτόματη παραγωγή ασκήσεων, τον αυτόματο προσδιορισμό της δυσκολίας ασκήσεων και την αυτόματη αξιολόγηση ασκήσεων σε ευφυή συστήματα διδασκαλίας. Επίσης, για να βοηθήσουμε τους εκπαιδευόμενους στην καλύτερη κατανόηση σύνθετων εννοιών του πεδίου της τεχνητής νοημοσύνης, αναπτύξαμε συνιστώσες για ευφυή συστήματα διδασκαλίας που μοντελοποιούν σύνθετες διαδικασίες της τεχνητής νοημοσύνης, χρησιμοποιώντας ευφυείς τεχνικές. Οι μεθοδολογίες και οι μηχανισμοί που αναπτύχθηκαν έχουν ως στόχο να επιτύχουν αποτελεσματικότερη μάθηση και παροχή εξατομικευμένης διδασκαλίας, καθώς επίσης και να βοηθήσουν τους διδάσκοντες στη διδασκαλία.

Δύο σημαντικές λειτουργίες που συμβάλλουν στην καλύτερη μάθηση σ' ένα ευφρές σύστημα διδασκαλίας είναι η δυνατότητα πειραματισμού με διαδικασίες καθώς και η οπτικοποίηση διαδικασιών και αποτελεσμάτων. Προς αυτή την κατεύθυνση, δημιουργήθηκαν συνιστώσες ευφούς συστήματος διδασκαλίας που παρέχουν δυνατότητες οπτικοποίησης των διαδικασιών, όπως είναι η οπτικοποίηση της λογικής αποδεικτικής διαδικασίας και η οπτικοποίηση του τρόπου λειτουργίας των αλγορίθμων αναζήτησης. Σε αυτό το πλαίσιο, δημιουργήσαμε εκπαιδευτικά συστήματα απόδειξης θεωρημάτων που επιτρέπουν τον πειραματισμό με διάφορους συνδυασμούς στρατηγικών ελέγχου της διαδικασίας απόδειξης και παρέχουν αντίστοιχες οπτικοποιήσεις, καθώς και την παρέμβαση του χρήστη στη διαδικασία της απόδειξης.

Μια σημαντική επίσης συνιστώσα της εκπαιδευτικής διαδικασίας και των εκπαιδευτικών συστημάτων αποτελεί η παραγωγή νέου εκπαιδευτικού υλικού. Η δυνατότητα του εμπλουτισμού ενός ευφούς εκπαιδευτικού συστήματος συνεχώς και αυτόματα με νέες, διαφορετικές και πρωτότυπες ασκήσεις συμβάλλει στην καλύτερη προσαρμογή του εκπαιδευτικού περιεχομένου στις ανάγκες του κάθε εκπαιδευόμενου. Για τον σκοπό αυτό, αναπτύξαμε νέες προσεγγίσεις για την ημιαυτόματη παραγωγή ασκήσεων πάνω στην λογική ως γλώσσα αναπαράστασης και στους αλγορίθμους αναζήτησης, τα οποία αποτελούν σημαντικά αντικείμενα του πεδίου της τεχνητής νοημοσύνης, και

τα οποία περιέχουν σύνθετες έννοιες και πολύπλοκες διαδικασίες. Οι μηχανισμοί που αναπτύχθηκαν παράγουν προτάσεις λογικής πρώτης τάξεως, που αποτελούν εκφράσεις συντακτικά και σημασιολογικά ορθές, καθώς και διαδραστικές ασκήσεις που χρησιμοποιούν γραφικές αναπαραστάσεις της λειτουργίας των αλγορίθμων αναζήτησης.

Ένας άλλος σημαντικός παράγοντας στην εκπαιδευτική διαδικασία είναι ο προσδιορισμός της δυσκολίας των ασκήσεων. Ο προσδιορισμός του επιπέδου δυσκολίας αποτελεί μια πολύπλοκη και πολυδιάστατη διαδικασία, η οποία είναι άρρηκτα συνδεδεμένη με το πεδίο που εφαρμόζεται και είναι δύσκολο να προσδιοριστεί ένα γενικό μοντέλο γι' αυτήν. Για τον σκοπό αυτό δημιουργήσαμε διαφορετικά συστήματα προσδιορισμού της δυσκολίας ασκήσεων, που χρησιμοποιούν διαφορετικές προσεγγίσεις και τεχνικές για τα πεδία της λογικής και των αλγορίθμων αναζήτησης: α) ένα σύνθετο έμπειρο σύστημα, βασισμένο σε κανόνες, που χρησιμοποιεί στοιχεία και από την πρόταση λογικής και από την πρόταση φυσικής γλώσσας για την εκτίμηση της δυσκολίας ασκήσεων μετατροπής προτάσεων φυσικής γλώσσας σε προτάσεις λογικής πρώτης τάξεως, β) ένα σύνθετο έμπειρο σύστημα, βασισμένο σε κανόνες, που χρησιμοποιεί στοιχεία και από την πολυπλοκότητα μιας πρότασης λογικής και από τη σύνθεση των απαιτούμενων βημάτων, για την εκτίμηση της δυσκολίας ασκήσεων μετατροπής προτάσεων λογικής πρώτης τάξεως σε προτασιακή μορφή, γ) δύο υβριδικά έμπειρα συστήματα, που χρησιμοποιούν νεύρο-ασαφείς και νεύρο-συμβολικούς κανόνες αντίστοιχα, για τον αυτόματο προσδιορισμό του επίπεδο δυσκολίας ασκήσεων σχετικών με αλγορίθμους αναζήτησης, και βασίζονται σ' ένα σύνολο παραμέτρων που εξάγονται από τα χαρακτηριστικά της άσκησης καθώς και της σωστής απάντησης.

Τα τελευταία χρόνια, υπάρχει έντονο ενδιαφέρον για τον σχεδιασμό και την ανάπτυξη συστημάτων και μηχανισμών για την αυτόματη αξιολόγηση των εκπαιδευόμενων σε ασκήσεις και εκπαιδευτικές δραστηριότητες σε διάφορους τομείς. Η αξιολόγηση των απαντήσεων των εκπαιδευόμενων και ο προσδιορισμός των επιδόσεών τους είναι απαραίτητη σε ένα εκπαιδευτικό σύστημα για να γίνει πιο αποτελεσματική η διαδικασία της μάθησης. Για το σκοπό αυτό αναπτύξαμε γενικές μεθοδολογίες, όπου με βάση το είδος των λαθών του εκπαιδευόμενου, την μοντελοποίηση του τύπου της απάντησής του καθώς και έναν μηχανισμό για την βαθμολόγηση των ασκήσεων, αξιολογούν αυτόματα, με ακρίβεια και με συνέπεια τις απαντήσεις τους ανάλογα με το πεδίο εφαρμογής. Πιο συγκεκριμένα δημιουργήθηκαν μηχανισμοί για την αυτόματη αξιολόγηση απαντήσεων σε ασκήσεις μετατροπής προτάσεων φυσικής γλώσσας σε προτάσεις λογικής πρώτης τάξεως καθώς και ασκήσεων μετατροπής προτάσεων λογικής πρώτης τάξεως σε προτασιακή μορφή, που στηρίζονται στη δεντρική αναπαράσταση των προτάσεων. Επίσης, δημιουργήθηκε μηχανισμός για την αυτόματη αξιολόγηση απαντήσεων σε ασκήσεις αλγορίθμων αναζήτησης, βασισμένος στην αναπαράσταση των απαντήσεων ως αλφαριθμητικές ακολουθίες.

Για να προσελκύσουμε το ενδιαφέρον των εκπαιδευόμενων, να δημιουργήσουμε νέα κίνητρα ώστε να γίνει η μάθηση πιο ευχάριστη και διασκεδαστική, σχεδιάσαμε και αναπτύξαμε εκπαιδευτικά παιχνίδια για την εκμάθηση εννοιών της τεχνητής νοημοσύνης. Η μάθηση μέσω εκπαιδευτικών παιχνιδιών (game-based learning) έχει ως στόχο να διδάξει μέσα από το περιβάλλον των παιχνιδιών στους εκπαιδευόμενους έννοιες και διαδικασίες τεχνητής νοημοσύνης, ενώ παράλληλα παραμένει ένα αλληλεπιδραστικό και διασκεδαστικό παιχνίδι. Πιο συγκεκριμένα, δημιουργήσαμε μια εκπαιδευτική εκδοχή του γνωστού παιχνιδιού Pacman, το Edu-Pacman, εμπλουτισμένο με κατάλληλες οπτικοποιήσεις, μέσω του οποίου μπορούν οι εκπαιδευόμενοι να διδαχθούν διαδικασίες, όπως αλγόριθμοι αναζήτησης, οπισθοδρόμηση κλπ. Επίσης, δημιουργήσαμε μια εκπαιδευτική εκδοχή του παιχνιδιού του χρωματισμού χαρτών, μέσω του οποίου μπορούν να διδαχθούν διαδικασίες αλγορίθμων ικανοποίησης περιορισμών.

Τα πειραματικά αποτελέσματα που προέκυψαν από την χρήση των ευφών συστημάτων και των εκπαιδευτικών παιχνιδιών στους εκπαιδευόμενους ήταν ιδιαίτερα ικανοποιητικά. Επίσης, οι μηχανισμοί που αναπτύχθηκαν αξιολογήθηκαν και τα πειραματικά αποτελέσματα έδειξαν υψηλή ακρίβεια και αξιοπιστία.

Abstract

In recent years, the rapid development of technology and the internet could not leave unaffected the e-learning domain. The use of the internet has changed the way in which learning is provided to learners. A basic and popular category of educational systems is Intelligent Tutoring Systems (ITSs), the main feature of which is to provide personalized learning for learners, using artificial intelligence methods and suitable pedagogical approaches.

In the context of this PhD thesis, we have developed new methodologies in the field of intelligent tutoring systems. In particular, methodologies for the semiautomatic generation of exercises, the automatic determination of the difficulty of exercises and the automatic assessment of answers to exercises in intelligent tutoring systems have been developed. Moreover, to help learners better understand complex concepts of the field of artificial intelligence, we developed components for intelligent tutoring systems that model complex processes of artificial intelligence using intelligent techniques and visualizations. The methodologies and mechanisms developed are designed to result in effective learning for learners, by providing personalized teaching, as well as help tutors in teaching.

Two important functions that contribute in better learning in an intelligent tutoring system is the possibility for experimentation with processes as well as the visualization of the processes and their results. To this direction, we created components for intelligent tutoring systems that provide possibilities for visualization of processes, like the visualization of logical inferences and the visualization of the functions of search algorithms. In this context, we created educational systems for theorem proving that allow experimentation with different combinations of strategies for controlling the logical inference process and provide corresponding visualizations and also allow intervention of the learner in the inference process.

Another important function of the educational process and educational systems is the generation of new educational material. The ability of enriching an intelligent educational system continuously and automatically with new, different and original exercises, contributes in a better adaptation of the educational content to the needs of each learner. For this purpose, we developed new approaches for the semi-automatic production of exercises on logic as a knowledge representation language and search algorithms, which are important areas of artificial intelligence field and contain complex concepts and complicated processes. The developed mechanisms produce sentences of first order logic, which are syntactically and semantically correct, for use in exercises on logic, as well as interactive exercises that use graphical representations of functions of search algorithms.

Another important factor in the educational process is the determination of the difficulty levels of exercises. The determination of the level of difficulty of an exercise is a complex and multidimensional process that is closely related to the field it is applied to and is difficult to produce a

general model. For this purpose, we developed different systems for automatic determination of the levels of difficulty of exercises, which use different approaches and techniques, for the fields of logic and search algorithms: a) a compound rule-based expert system, which uses elements from the logical formula and the natural language sentence, for estimating the difficulty levels of exercises related to conversions of natural language sentences to logical formulas, b) a compound rule-based expert system, which uses elements from both the complexity of the logical formula and the composition of the required steps, for estimating the difficulty levels of exercises related to conversions of logical formulas to clausal forms, c) two hybrid expert systems, which use neuro-fuzzy and neuro-symbolic rules respectively, for the automatic estimation of the difficulty levels of exercises related to search algorithms, and are based on a set of parameters extracted from the characteristics of the exercise as well as characteristics of the correct answer.

In recent years, there is great interest in the design and development of systems and mechanisms for the automatic evaluation of the learners in exercises and training activities in various fields. The evaluation of learners' responses and their performances assessment is essential in an educational system, to make the learning process more effective. For this purpose, we developed general methodologies that based on the types of learners' errors, the modeling of the answer types and a mechanism for marking exercises, evaluate their answers automatically, accurately and consistently, in a way depending on the domain. More specifically, we created mechanisms for automatic assessment of the answers to exercises on converting first order logic formulas into clausal forms, which are based on tree representations of the formulas. Also, we created a mechanism for automatic assessment of answers to exercises on search algorithms, based on representation of answers as sequences of strings.

To attract the interest of learners, to create new incentives and to make learning more enjoyable and entertaining, we designed and developed educational games for learning concepts of artificial intelligence. Learning through an educational game (game-based learning) aims to teach learners through the game environment about concepts and processes of artificial intelligence, while remaining an interactive and entertaining game. More specifically, we created an educational version of the Pacman game, called Edu-Pacman, enriched with suitable visualizations, through which processes like search algorithms, backtracking etc. can be taught. Also, we created an educational version of the map coloring game, through which processes related to constraint satisfaction algorithms can be taught.

The experimental results collected from the use of the above intelligent mechanisms and systems and educational games by learners were particularly satisfactory. Moreover, the mechanisms developed were systematically evaluated and the experimental results show high accuracy and reliability.

1

Εισαγωγή

1.1 Εισαγωγικά στοιχεία

Τα τελευταία χρόνια η εξέλιξη της τεχνολογίας και η εξάπλωση του διαδικτύου έχουν αλλάξει τον τρόπο με τον οποίο το εκπαιδευτικό περιεχόμενο και η διαδικασία της μάθησης παρέχεται στους εκπαιδευόμενους. Η χρήση του διαδικτύου αποτελεί ένα κύριο μέσο για την εκπαίδευση, και προσφέρει νέες δυνατότητες για να γίνει η διαδικασία μάθησης πιο αποτελεσματική. Τα Ευφυή Συστήματα Διδασκαλίας (ΕΣΔ) (Intelligent Tutoring Systems - ITSs) αποτελούν μια γενιά εκπαιδευτικών συστημάτων που περιέχουν ευφυείς και παιδαγωγικές μεθόδους και έχουν ως στόχο να αυξήσουν την αποτελεσματικότητά της μάθησης και της διδασκαλίας. Τα ευφυή συστήματα διδασκαλίας είναι πολυδιάστατα συστήματα που χρησιμοποιούνται για να προσφέρουν εξατομικευμένη μάθηση στους εκπαιδευόμενους και αποτελούν τη δημοφιλέστερη κατηγορία εκπαιδευτικών συστημάτων. Επίσης, έχουν εξελιχθεί σε ένα βασικό μέσο παροχής εκπαίδευσης και συμβάλουν αποτελεσματικά στην βελτίωση της μάθησης και της κατανόησης (Aleven et al., 2009; VanLehn, 2006; Woolf, 2010). Ένα από τα κύρια χαρακτηριστικά των ΕΣΔ είναι ότι μπορούν να προσαρμόζουν την εκπαιδευτική διαδικασία στις ανάγκες του κάθε εκπαιδευόμενου με σκοπό να μεγιστοποιήσουν τα μαθησιακά αποτελέσματα. Αυτό επιτυγχάνεται κυρίως με την χρήση μεθόδων τεχνητής νοημοσύνης για να αναπαραστήσουν το παιδαγωγικό μοντέλο, τη γνώση του πεδίου διδασκαλίας, το μοντέλο του μαθητή και τις διαδικασίες αξιολόγησης (Shute & Zapata-Rivera, 2010).

Μερικοί βασικοί ορισμοί που έχουν δοθεί για τα ευφυή συστήματα διδασκαλίας είναι:

Τα ευφυή συστήματα διδασκαλίας (Intelligent Tutoring Systems) είναι συστήματα που προσφέρουν καινοτόμες και αποτελεσματικές εκπαιδευτικές δραστηριότητες οι οποίες μπορούν να

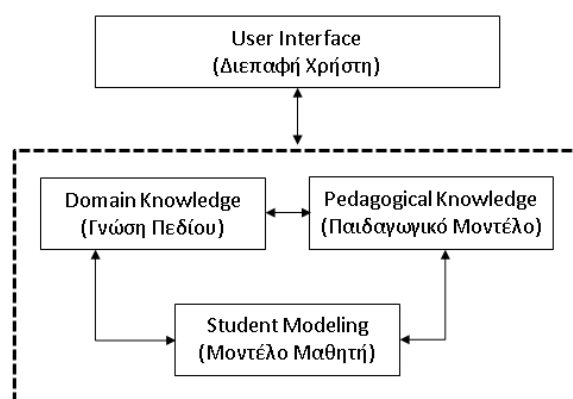
προσαρμόζονται στις ιδιαιτερότητες και τα χαρακτηριστικά των εκπαιδευόμενων (Brusilovsky, 1998).

Τα ευφυή συστήματα διδασκαλίας (Intelligent Tutoring Systems) αποτελούν ένα διεπιστημονικό πεδίο που ερευνά το πώς μπορούν να σχεδιαστούν εκπαιδευτικά συστήματα τα οποία να παρέχουν εκπαίδευση προσαρμοσμένη στις ανάγκες των εκπαιδευόμενων, όπως κάνουν οι διδάσκοντες. Η έρευνα στο πεδίο αυτό έχει σχεδιάσει με επιτυχία τεχνικές και συστήματα που παρέχουν προσαρμοστική υποστήριξη και την επίλυση προβλημάτων σε πολλά πεδία. Υπάρχουν, ωστόσο, πολλές εκπαιδευτικές δραστηριότητες που μπορούν να επωφεληθούν από την εξατομικευμένη υποστήριξη, όπως η μελέτη παραδειγμάτων, η εξερεύνηση διαδραστικών προσομοιώσεων και η μάθηση μέσω εκπαιδευτικών παιχνιδιών. Η παροχή εξατομικευμένης υποστήριξης στα πλαίσια αυτών των εκπαιδευτικών δραστηριοτήτων θέτει μοναδικές προκλήσεις, καθώς απαιτεί τα ευφυή εκπαιδευτικά συστήματα να μοντελοποιούν και να προσαρμόζονται στην μαθησιακή συμπεριφορά των εκπαιδευόμενων, τις δεξιότητες και τις νοητικές καταστάσεις οι οποίες συχνά δεν είναι τόσο δομημένες και καλά-ορισμένες όπως είναι στις παραδοσιακές εκπαιδευτικές δραστηριότητες επίλυσης προβλημάτων (Connati, 2009).

Επίσης, ένας άλλος ορισμός είναι:

Τα ευφυή συστήματα διδασκαλίας (Intelligent Tutoring Systems) είναι υπολογιστικά περιβάλλοντα μάθησης τα οποία ενσωματώνουν υπολογιστικά μοντέλα από τις γνωστικές επιστήμες, τις επιστήμες μάθησης, της υπολογιστικής γλωσσολογίας, της τεχνητή νοημοσύνης, τα μαθηματικά και άλλους τομείς. Επίσης, τα ευφυή συστήματα διδασκαλίας αναλύουν και παρακολουθούν τις καταστάσεις των εκπαιδευόμενων με μεγάλη λεπτομέρεια, μια διαδικασία που ονομάζεται μοντέλο μαθητή (student modeling) (Graesser et al., 2012).

Ένα Ευφυνές Σύστημα Διδασκαλίας (Intelligent Tutoring System-ITS) αποτελείται από τέσσερα βασικά μέρη: το πεδίο γνώσης (domain Knowledge), το μοντέλο χρήστη (user modeling), το παιδαγωγικό μοντέλο (Pedagogical Module) και τη διεπαφή χρήστη (user interface). Τα τέσσερα βασικά μέρη παρουσιάζονται στην εικόνα 1.1. Η γνώση του πεδίου περιέχει το εκπαιδευτικό υλικό που παρουσιάζεται στους εκπαιδευόμενους του συστήματος. Το Μοντέλο Μαθητή (Student modeling) αποτελεί μια βασική μονάδα ενός ευφυούς συστήματος που καταγράφει πληροφορίες και χαρακτηριστικά σχετικά με τον εκπαιδευόμενο όπως πληροφορίες που αφορούν την ικανότητα μάθησης, τις προτιμήσεις του εκπαιδευόμενου, το εκπαιδευτικό υλικό που έχει μελετήσει κ.ά.



Εικόνα 1.1 Βασική Αρχιτεκτονική Ευφυών Συστημάτων Διδασκαλίας

Το παιδαγωγικό μοντέλο διδασκαλίας διαθέτει μεθόδους για να αναπροσαρμόζεται η παρουσίαση του εκπαιδευτικού υλικού σύμφωνα με τα δεδομένα του μοντέλου του μαθητή καθώς επίσης, διαθέτει διάφορες στρατηγικές οι οποίες μπορούν να καθορίσουν τον τρόπο οργάνωσης του εκπαιδευτικού περιεχομένου.

Ένα από τα αρχικά οράματα, για τα ευφυή συστήματα διδασκαλίας ήταν τα συστήματα να γίνουν τόσο διορατικά όσο ο διδάσκοντας και να χρησιμοποιούν μερικές από τις ίδιες παιδαγωγικές και διδακτικές στρατηγικές που χρησιμοποιεί ο διδάσκοντας-ειδικός (Merrill et al., 1992; Lepper et al., 1993; McArthur et al., 1990). Τα τελευταία χρόνια, τα ευφυή εκπαιδευτικά συστήματα παρέχουν αρκετές δυνατότητες στους εκπαιδευόμενους και στους διδάσκοντες. Στην διεθνή βιβλιογραφία υπάρχουν αρκετά συστήματα με διαφορετικές δυνατότητες, όπως συστήματα τα οποία μπορούν να παρέχουν υποστήριξη σε κάθε βήμα της διαδικασία της σκέψης του εκπαιδευόμενου (VanLehn et al., 2005 ; Anderson et al., 1995), συστήματα με τα οποία ο εκπαιδευόμενος μπορεί να αλληλεπιδράσει και να επικοινωνήσει σε φυσική γλώσσα (Nye et al., 2014), συστήματα που περιέχουν σύνθετο μοντέλο διδασκαλίας και παιδαγωγικές στρατηγικές (Heffernan & Koedinger, 2002; Khachatryan et al., 2014) καθώς και συστήματα που αναγνωρίζουν και τις διάφορες συναισθηματικές καταστάσεις των εκπαιδευόμενων (D’Mello et al., 2010 ; Arroyo et al., 2014).

Υπάρχουν πολλά εκπαιδευτικά συστήματα, τα οποία χρησιμοποιούνται από εκατοντάδες ως χιλιάδες εκπαιδευόμενους κάθε χρόνο, παρέχοντας στατιστικά για τις επιδόσεις τους και την αποτελεσματικότητα στην μάθηση, όπως τα SQL-Tutor (Mitrovic & Ohlsson, 1999), ALEKS (Craig et al., 2013), Cognitive Tutor (Pane et al., 2013) και ASSISTments (Koedinger et al., 2010).

1.2 Συνεισφορά της διατριβής

Σ' αυτήν την ενότητα συνοψίζεται η ερευνητική συνεισφορά της παρούσας διατριβής. Βασικός σκοπός της διατριβής είναι η ανάπτυξη μηχανισμών που μπορούν να συμβάλλουν στην εξέλιξη των ευφών συστημάτων διδασκαλίας (ΕΣΔ) ώστε να γίνει πιο αποτελεσματική η διαδικασία της μάθησης. Ως πεδίο εφαρμογής των μηχανισμών μας χρησιμοποιήσαμε τα ευφυή συστήματα διδασκαλίας τεχνητής νοημοσύνης. Για την επίτευξη αυτών των στόχων αναπτύχθηκαν τα παρακάτω:

- Ανάπτυξη ευφών συστημάτων διδασκαλίας για την κατανόηση εννοιών της τεχνητής νοημοσύνης, που μοντελοποιούν σύνθετες διαδικασίες οι οποίες δυσκολεύουν τους εκπαιδευόμενους να τις κατανοήσουν. Πιο συγκεκριμένα, αναπτύχθηκε το ευφές συστήματα διδασκαλίας Τεχνητής Νοημοσύνης (AITS) το οποίο παρέχει εξατομικευμένη μάθηση και διδάσκει διάφορες έννοιες Τ.Ν. όπως οι αλγόριθμοι αναζήτησης, η ικανοποίηση περιορισμών, και μέσω γραφικών αναπαραστάσεων και οπτικοποιήσεων των αλγορίθμων αναζήτησης να γίνει καλύτερη η κατανόηση των εννοιών από τους εκπαιδευόμενους. Επίσης αναπτύχθηκαν ευφυή συστήματα διδασκαλίας που μοντελοποιούν σύνθετες έννοιες της λογικής όπως η διαδικασία μετατροπής Κατηγορηματικής Λογικής Πρώτης Τάξης σε Προτασιακή Μορφή καθώς και η δημιουργία ισοδύναμων εκφράσεων λογικής. Αξίζει να σημειωθεί ότι στην διεθνή βιβλιογραφία δεν υπάρχουν αντίστοιχα συστήματα.
- Παροχή δυνατότητας πειραματισμών και αντίστοιχων οπτικοποιήσεων των διαδικασιών και αποτελεσμάτων σε μη εκπαιδευτικούς μηχανισμούς για μάθηση μέσω εμπειρικής χρήσης. Πιο συγκεκριμένα, δημιουργήθηκαν δύο εκπαιδευτικές εκδοχές ενός υπάρχοντος αποδείκτη θεωρημάτων, δηλ. ενός συστήματος επίλυσης προβλημάτων μέσω αναπαράστασης σε λογική πρώτης τάξης. Η μία εκδοχή επιτρέπει τον πειραματισμό με διάφορους συνδυασμούς στρατηγικών ελέγχου της διαδικασίας απόδειξης και αντίστοιχες οπτικοποιήσεις. Η δεύτερη εκδοχή επιτρέπει την παρέμβαση του χρήστη στη διαδικασία απόδειξης με αντίστοιχη οπτικοποίηση των αποτελεσμάτων. Αξίζει να σημειωθεί ότι στην διεθνή βιβλιογραφία δεν υπάρχουν αντίστοιχα συστήματα.
- Ανάπτυξη μηχανισμών για την αυτόματη παραγωγή εκπαιδευτικού υλικού για τον εμπλουτισμό των ευφών συστημάτων διδασκαλίας και ανανέωση του εκπαιδευτικού υλικού. Πιο συγκεκριμένα, αναπτύχθηκαν α) μηχανισμός παραγωγής προτάσεων λογικής πρώτης τάξεως, που αποτελούν εκφράσεις συντακτικά και σημασιολογικά ορθές, για χρήση σε ασκήσεις λογικής και β) μηχανισμός παραγωγής διαδραστικών

ασκήσεων που χρησιμοποιούν γραφικές αναπαραστάσεις λειτουργίας αλγορίθμων αναζήτησης. Οι μηχανισμοί αυτοί λαμβάνουν υπόψη μια σημαντική παράμετρο που είναι ο προσδιορισμός του επιπέδου της δυσκολίας των παραγόμενων ασκήσεων. Στην διεθνή βιβλιογραφία δεν υπάρχουν αναφορές συστημάτων για την αυτόματη/ημιαυτόματη παραγωγή ασκήσεων στον τομέα της λογικής και γενικά για την παραγωγή αλληλεπιδραστικών ασκήσεων καθώς και να λαμβάνεται υπόψη το επίπεδο δυσκολίας της άσκησης για την δημιουργία της.

- Ανάπτυξη μηχανισμών για τον αυτόματο προσδιορισμό της δυσκολίας ασκήσεων που βασίζονται σε συνδυαστικές προσεγγίσεις και υβριδικούς μηχανισμούς. Πιο συγκεκριμένα, δημιουργήθηκαν α) σύνθετο έμπειρο σύστημα, βασισμένο σε κανόνες, που χρησιμοποιεί στοιχεία της σωστής απάντησης της άσκησης (πρόταση λογικής) και της εκφώνησης της άσκησης (πρόταση φυσικής γλώσσας) για την εκτίμηση της δυσκολίας ασκήσεων μετατροπής προτάσεων φυσικής γλώσσας σε προτάσεις λογικής πρώτης τάξεως, β) σύνθετο έμπειρο σύστημα, βασισμένο σε κανόνες, που χρησιμοποιεί στοιχεία και από την πολυπλοκότητα μιας πρότασης λογικής και από τη σύνθεση των απαιτούμενων βημάτων για την εκτίμηση της δυσκολίας ασκήσεων μετατροπής προτάσεων λογικής πρώτης τάξεως σε προτασιακή μορφή, γ) υβριδικά έμπειρα συστήματα, που χρησιμοποιούν νέυρο-ασαφείς και νέυρο-συμβολικούς κανόνες αντίστοιχα, για τον αυτόματο προσδιορισμό του επίπεδο δυσκολίας ασκήσεων πάνω σε αλγορίθμους αναζήτησης, και βασίζονται σ' ένα σύνολο παραμέτρων που εξάγονται από τα χαρακτηριστικά της άσκησης καθώς και της σωστής απάντησης.
- Σχεδίαση και εισαγωγή γενικής μεθοδολογίας/μοντέλου για αξιολόγηση απαντήσεων σε ασκήσεις, όπου οι απαντήσεις αποτελούν εκφράσεις κειμενικής μορφής. Με βάση το γενικό μοντέλο, δημιουργήθηκαν μηχανισμοί για την αυτόματη αξιολόγηση απαντήσεων σε ασκήσεις μετατροπής προτάσεων φυσικής γλώσσας σε προτάσεις λογικής πρώτης τάξεως καθώς και ασκήσεις μετατροπής προτάσεων λογικής πρώτης τάξεως σε προτασιακή μορφή.
- Σχεδίαση γενικής μεθοδολογίας/μοντέλου για την αξιολόγηση απαντήσεων σε ασκήσεις όπου οι απαντήσεις αποτελούν ακολουθίες στοιχείων. Πιο συγκεκριμένα, αναπτύχθηκε μια μεθοδολογία και μηχανισμοί αξιολόγησης όπου για την αξιολόγηση των απαντήσεων βασίζονται στην ομοιότητα μεταξύ της απάντησης του εκπαιδευόμενου και της σωστής απάντησης, στην μοντελοποίηση της απάντησης του, στην εις βάθος ανάλυση των λαθών, στον προσδιορισμό του είδους των λαθών που πραγματοποιούνται από τον εκπαιδευόμενο καθώς επίσης και στο μηχανισμό για την αυτόματη

βαθμολόγηση των απαντήσεων, που αξιολογεί τις απαντήσεις ανάλογα με το πεδίο εφαρμογής. Με βάση το μοντέλο αυτό, δημιουργήθηκε μηχανισμός για την αυτόματη βαθμολόγηση ασκήσεων σχετικών με αλγορίθμους αναζήτησης. Στην διεθνή βιβλιογραφία, οι υπάρχουσες εργασίες περιορίζονται στην αυτόματη αξιολόγηση απαντήσεων ασκήσεων διαφόρων γλωσσών προγραμματισμού, ωστόσο δεν υπάρχει κάποια γενική μεθοδολογία.

- Σχεδίαση και ανάπτυξη εκπαιδευτικών παιχνιδιών με βάση παιδαγωγικές προσεγγίσεις για την εκμάθηση και την κατανόηση των αλγορίθμων αναζήτησης και της ικανοποίησης περιορισμών με πιο αποτελεσματικό, φυσικό και διασκεδαστικό τρόπο. Πιο συγκεκριμένα, δημιουργήθηκαν α) μια εκπαιδευτική εκδοχή του γνωστού παιχνιδιού Pacman, το Edu-Pacman, εμπλουτισμένο με κατάλληλες οπτικοποιήσεις, μέσω του οποίου μπορούν να διδαχθούν διαδικασίες, όπως αλγόριθμοι αναζήτησης, οπισθοδρόμηση κλπ, β) μια εκπαιδευτική εκδοχή του παιχνιδιού του χρωματισμού χαρτών, μέσω του οποίου μπορούν να διδαχθούν παίζοντας διαδικασίες των αλγορίθμων ικανοποίησης περιορισμών, οπισθοδρόμησης κλπ.

1.3 Διάρθρωση της διατριβής

Η παρούσα διατριβή δομείται σε επτά κεφάλαια. Στο παρόν κεφάλαιο γίνεται μια σύντομη παρουσίαση της διάρθρωσης των κεφαλαίων.

Στο 2^ο κεφάλαιο παρουσιάζονται τα ευφυή συστήματα διδασκαλίας που σχεδιάστηκαν και αναπτύχθηκαν για το μάθημα της τεχνητής νοημοσύνης και έχουν ως στόχο να βοηθήσουν τους εκπαιδευόμενους στην κατανόηση και την εκμάθηση σύνθετων και πολύπλοκων εννοιών της τεχνητής νοημοσύνης.

Στο 3^ο κεφάλαιο παρουσιάζονται προσεγγίσεις για την ημιαυτόματη παραγωγή ασκήσεων πάνω στο πεδίο της λογικής και των αλγορίθμων αναζήτησης. Οι μηχανισμοί που αναπτύχθηκαν δημιουργούν ορθά δομημένες ασκήσεις και παράγουν ασκήσεις πάνω στη λογική που αποτελούν εκφράσεις συντακτικά και σημασιολογικά ορθές καθώς επίσης και αλληλεπιδραστικές ασκήσεις που αποτελούν δεντρικές αναπαραστάσεις για τους αλγορίθμους αναζήτησης.

Στο 4^ο κεφάλαιο παρουσιάζονται μεθοδολογίες για τον αυτόματο προσδιορισμό της δυσκολίας των ασκήσεων. Οι μεθοδολογίες που αναπτύχθηκαν βασίζονται σε σύνθετα έμπειρα

συστήματα και υβριδικούς μηχανισμούς για τον αυτόματο προσδιορισμό του επίπεδο δυσκολίας των ασκήσεων πάνω στη λογική και τους αλγορίθμους αναζήτησης.

Στο 5^ο κεφάλαιο παρουσιάζεται μια γενική μεθοδολογία για την αυτόματη ανάλυση και αξιολόγηση των απαντήσεων των εκπαιδευόμενων σε ασκήσεις. Η μεθοδολογία αυτή με βάση την ομοιότητα μεταξύ της απάντησης του εκπαιδευόμενου και της σωστής απάντησης και το είδος των λαθών καθώς και σ' έναν μηχανισμό βαθμολόγησης όπου με συνέπεια και με ακρίβεια αξιολογεί τις απαντήσεις των εκπαιδευόμενων σε ασκήσεις. Η μεθοδολογία αξιολόγησης εφαρμόστηκε για την αυτόματη αξιολόγηση απαντήσεων πάνω στην λογική και στους αλγορίθμους αναζήτησης.

Στο 6^ο κεφάλαιο παρουσιάζονται εκπαιδευτικά παιχνίδια που σχεδιάστηκαν για να προσελκύσουμε το ενδιαφέρον των εκπαιδευόμενων πάνω σε έννοιες της τεχνητής νοημοσύνης και να συμβάλλουν ώστε να γίνει η μάθηση πιο ευχάριστη και διασκεδαστική. Τα εκπαιδευτικά παιχνίδια έχουν ως στόχο να βοηθήσουν τους εκπαιδευόμενους την κατανόηση εννοιών της Τ.Ν.

Τέλος, στο κεφάλαιο 7 συνοψίζονται τα συμπεράσματα της εργασίας και οι μελλοντικές κατευθύνσεις της έρευνας.

2

Ευφυή Συστήματα Διδασκαλίας Τεχνητής Νοημοσύνης

2.1 Εισαγωγή

Η Τεχνητή Νοημοσύνη αποτελεί ένα σημαντικό πεδίο της επιστήμης, που διδάσκεται σε πολλά τμήματα, αλλά κυρίως στα τμήματα της επιστήμης των υπολογιστών, και περιέχει πολλές συνθέτες έννοιες και πολύπλοκες διαδικασίες, οι οποίες δημιουργούν δυσκολίες στους εκπαιδευόμενους να τις κατανοήσουν και να τις εφαρμόσουν. Για να βοηθήσουμε τους εκπαιδευόμενους στην εκμάθηση εννοιών της τεχνητής νοημοσύνης και τους διδάσκοντες στη διδασκαλία του μαθήματος σχεδιάσαμε και αναπτύξαμε το ευφρές *Σύστημα Διδασκαλίας Τεχνητής Νοημοσύνης-AITS (Artificial Intelligence Tutoring System-AITS)*. Το AITS έχει ως στόχο να βοηθήσει τους εκπαιδευόμενους μέσω διαφόρων αλληλεπιδραστικών δραστηριοτήτων να μελετήσουν θέματα της τεχνητής νοημοσύνης, όπως η αναπαράσταση γνώσης, οι αλγόριθμοι αναζήτησης, η ικανοποίηση περιορισμών κ.ά.. Το AITS είναι ένα προσαρμοστικό ευφρές σύστημα, που προσαρμόζει το εκπαιδευτικό περιεχόμενο και παρέχει κατάλληλη καθοδήγηση ανάλογα με τις ανάγκες του κάθε εκπαιδευόμενου με βάση το προφίλ του και το επίπεδο γνώσεων του. Στα πλαίσια του συστήματος, σχεδιάστηκαν και αναπτύχθηκαν μέθοδοι ανάλυσης και αξιολόγησης των απαντήσεων των εκπαιδευόμενων καθώς και τεχνικές για τον ακριβή καθορισμό του επιπέδου δυσκολίας των ασκήσεων, με στόχο τον προσδιορισμό του γνωστικού επιπέδου των εκπαιδευόμενων και την προσαρμογή του εκπαιδευτικού υλικού στις ανάγκες τους. Για την ανάπτυξη του συστήματος χρησιμοποιήθηκαν τεχνολογίες *σημασιολογικού ιστού (semantic web)* και πιο συγκεκριμένα οντολογίες, για να αναπαραστήσουν την γνώση του πεδίου και τα μοντέλα των εκπαιδευόμενων.

Ένα σημαντικό πεδίο της αναπαράστασης γνώσης είναι η λογική. Η λογική περιέχει σύνθετες έννοιες και διαδικασίες που δυσκολεύουν τους εκπαιδευόμενους να τις κατανοήσουν και για το σκοπό αυτό μοντελοποιήσαμε σύνθετες διαδικασίες και αναπτύξαμε ευφυή συστήματα για τη διδασκαλία τους. Πιο συγκεκριμένα, αναπτύξαμε τα ευφυή αλληλεπιδραστικά συστήματα FOLtoCF και FOL-EQU για την διδασκαλία της μετατροπής Κατηγορηματικής Λογικής Πρώτης Τάξης (ΚΛΠΤ) σε Προτασιακή Μορφή (ΠΜ) και την διδασκαλία λογικά ισοδύναμων εκφράσεων ΚΛΠΤ. Το σύστημα FOLtoCF είναι ένα ευφύες σύστημα που μοντελοποιεί τα βήματα της διαδικασίας μετατροπής εκφράσεων ΚΛΠΤ σε ΠΜ και έχει ως στόχο να βοηθήσει τους εκπαιδευόμενους στην αποτελεσματική και εις βάθος εκμάθηση της διαδικασίας. Επίσης, το σύστημα FOL-EQU είναι ένα ευφύες σύστημα που αναπτύχθηκε για την εκμάθηση των λογικών ισοδυναμιών και προσφέρει στους εκπαιδευόμενους την δυνατότητα να μελετήσουν και να μάθουν τους κανόνες των λογικών ισοδυναμιών. Οι εκπαιδευόμενοι μέσω του συστήματος μπορούν να μετατρέψουν εκφράσεις ΚΛΠΤ βήμα προς βήμα στις αντίστοιχες ισοδύναμες. Επίσης, για να βοηθήσουμε τους εκπαιδευόμενους να κατανοήσουν τη διαδικασία της απόδειξης θεωρημάτων και τις στρατηγικές επίλυσης, καθώς και να επιλύσουν προβλήματα αυτόματα ή αλληλεπιδραστικά, σχεδιάσαμε και αναπτύξαμε τον αυτόματο αποδείκτη θεωρημάτων EX-ACTP και τον αλληλεπιδραστικό αποδείκτη i-ACTP. Ο EX-ACTP είναι ένας αυτόματος αποδείκτης θεωρημάτων που βασίζεται στην λογική πρώτης τάξης και συνάγει με αυτόματο τρόπο από ένα σύνολο λογικών προτάσεων την αλήθεια μιας δεδομένης λογικής πρότασης, ενώ ο αλληλεπιδραστικός αποδείκτης θεωρημάτων i-ACTP πραγματοποιεί απόδειξη θεωρημάτων με αλληλεπιδραστικό τρόπο και έχει ως στόχο να βοηθήσει τους εκπαιδευόμενους στην κατανόηση των μεθόδων της αποδεικτικής θεωρίας και την παροχή κατάλληλης ανάδρασης. Επίσης, στα ευφυή συστήματα που αναπτύξαμε, ενσωματώσαμε μηχανισμούς για την αυτόματη παραγωγή ασκήσεων, τον αυτόματο προσδιορισμό της δυσκολίας των ασκήσεων καθώς και την αυτόματη αξιολόγηση των απαντήσεων των εκπαιδευόμενων σε ασκήσεις.

2.2 Σχετικές Εργασίες

Στην διεθνή βιβλιογραφία υπάρχουν αρκετά εκπαιδευτικά συστήματα στο πεδίο των αλγόριθμων. Πιο συγκεκριμένα, το εκπαιδευτικό σύστημα PATHFINDER (Sanchez-Torrubia et al., 2009) είναι ένα σύστημα που προσφέρει τη δυνατότητα στους εκπαιδευόμενους να μελετήσουν και να εφαρμόσουν αλγόριθμους, όπως τον αλγόριθμο Dijkstra, σε διάφορες ασκήσεις και προβλήματα. Επίσης, βασικό χαρακτηριστικό του συστήματος είναι ότι

προσφέρει τη δυνατότητα στους εκπαιδευόμενους να δουν μέσω οπτικοποιήσεων τα βήματα του αλγορίθμου, τον κώδικα του καθώς και να ενημερώνονται για πιθανά λάθη κατά την εφαρμογή τους. Στην εργασία (Kordaki et al., 2008), παρουσιάζεται ένα εκπαιδευτικό σύστημα το οποίο διδάσκει τα βασικά θέματα των αλγορίθμων και των δομών δεδομένων. Το σύστημα απευθύνεται κυρίως σε αρχάριους και δίνει την δυνατότητα στους εκπαιδευόμενους να εφαρμόσουν τους αλγορίθμους σε πολλές καταστάσεις και προβλήματα, εμπνευσμένα από καταστάσεις της καθημερινότητας. Επίσης, το εκπαιδευτικό σύστημα TRAKLA2 (Nikander et al., 2004) έχει ως στόχο να βοηθήσει τους εκπαιδευόμενους να μελετήσουν και να μάθουν να εφαρμόζουν τους αλγορίθμους αναζήτησης. Στα πλαίσια του συστήματος, οι εκπαιδευόμενοι μπορούν να δουν τον τρόπο λειτουργίας των αλγορίθμων μέσω οπτικοποιήσεων, καθώς και να λύσουν ασκήσεις με την βοήθεια του συστήματος.

Στο πεδίο της λογικής, υπάρχουν διάφορα εκπαιδευτικά συστήματα τα οποία χρησιμοποιούνται για την διδασκαλία της λογικής και την εκμάθηση λογικών διαδικασιών. Το εκπαιδευτικό σύστημα ProofWeb (Hendriks et al., 2010) αποτελεί ένα χαρακτηριστικό παράδειγμα συστήματος που αναπτύχθηκε για να βοηθήσει τους εκπαιδευόμενους να μάθουν τη διαδικασία της συλλογιστικής και της εξαγωγής συμπερασμών και βασίζεται στον αποδείκτη θεωρημάτων Coq (Bertot & Castéran, 2013). Οι εκπαιδευόμενοι μέσω του συστήματος ProofWeb μαθαίνουν να εφαρμόζουν τους κανόνες συλλογιστικής σε διάφορα λογικά προβλήματα με την βοήθεια και την καθοδήγηση του συστήματος. Σε αυτή την κατεύθυνση, πολλά από τα υπάρχοντα συστήματα διδασκαλίας λογικής παρέχουν στρατηγικές για την απόδειξη θεωρημάτων καθώς και καθοδήγηση στους εκπαιδευόμενους, όπως το AProS (Sieg, 2007) και το Pandora (Boda et al., 2007). Επίσης, το εκπαιδευτικό σύστημα KRRT (Alonso et al., 2007) έχει ως στόχο να βοηθήσει τους εκπαιδευόμενους να μάθουν την κατηγορηματική λογική ως γλώσσα αναπαράστασης γνώσης. Οι εκπαιδευόμενοι καλούνται να μεταφράσουν προτάσεις φυσικής γλώσσας σε κατηγορηματική λογική πρώτης τάξης (ΚΛΠΤ). Για κάθε πρόταση καλούνται να προσδιορίσουν την έκφραση ΚΛΠΤ που μεταφέρει το σημασιολογικό περιεχόμενό της και το σύστημα ελέγχει τη σύνταξη της έκφρασης ΚΛΠΤ και ενημερώνει τον εκπαιδευόμενο αν είναι σωστή ή λανθασμένη σε συντακτικό επίπεδο. Το IDEAS (Lodder et al., 2008) είναι ένα εκπαιδευτικό εργαλείο που βοηθάει τους εκπαιδευόμενους να επεξεργαστούν εκφράσεις προτασιακής λογικής. Οι εκπαιδευόμενοι με τη βοήθεια του εργαλείου μπορούν να μετατρέψουν εκφράσεις προτασιακής λογικής σε διαζευκτική κανονική μορφή (disjunctive normal form - DNF) εφαρμόζοντας κανόνες λογικής ισοδυναμίας. Το Logic-ITA (Yacef, 2005) είναι ένα εκπαιδευτικό σύστημα στο πεδίο της λογικής, το οποίο παρέχει στους εκπαιδευόμενους ένα περιβάλλον εξάσκησης και εκμάθησης δημιουργίας τυπικών αποδείξεων (formal proofs).

Επίσης, το σύστημα παρέχει στον διδάσκοντα μια σειρά από εργαλεία για να διαχειριστεί το εκπαιδευτικό υλικό του συστήματος και να καθορίσει τις εκπαιδευτικές προδιαγραφές και τους μαθησιακούς στόχους. Στην ίδια κατεύθυνση, το εκπαιδευτικό σύστημα P-Logic (Lukins et al., 2002) έχει ως στόχο να βοηθήσει τους εκπαιδευόμενους να μελετήσουν τη προτασιακή λογική και την απόδειξη θεωρημάτων. Το σύστημα καταγράφει όλες τις ενέργειες των εκπαιδευόμενων, τις αναλύει και εντοπίζει διαδικασίες στις οποίες αντιμετωπίζουν οι εκπαιδευόμενοι ιδιαίτερες δυσκολίες. Επίσης, το SAToulouse (Gasquet et al., 2011) σχεδιάστηκε για να βοηθήσει τους εκπαιδευόμενους στην χρήση και την κατανόηση της προτασιακής λογικής και στην αναπαράσταση διάφορων προβλημάτων. Οι εκπαιδευόμενοι καλούνται να αναπαραστήσουν σε προτασιακή μορφή δύσκολα και πολύπλοκα προβλήματα, όπως το Sudoku, και μέσω ενός αποδείκτη (SAT-prover) να τα επιλύσουν αυτόματα.

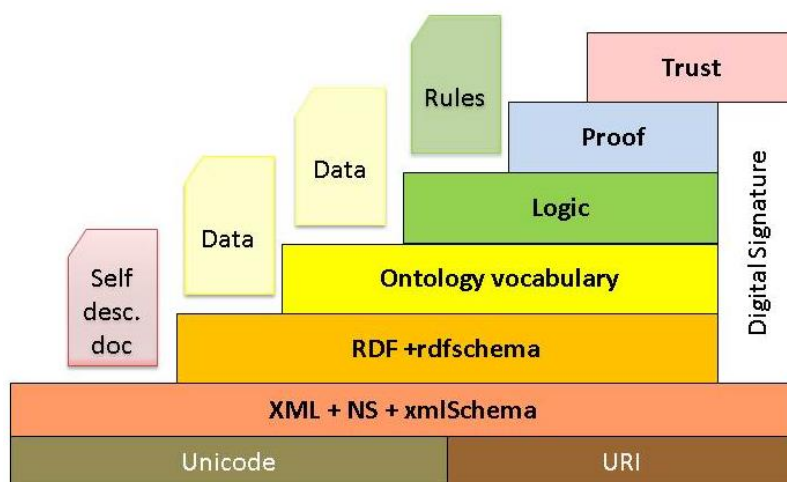
Αξίζει να σημειωθεί ότι κανένα από τα παραπάνω συστήματα δεν προσφέρει τη δυνατότητα εκμάθησης με συστηματικό τρόπο της διαδικασίας μετατροπής εκφράσεων κατηγορηματικής λογικής σε προτασιακή μορφή καθώς και την επεξεργασία και την μετατροπή προτάσεων ΚΛΠΤ σε ισοδύναμες μορφές. Επίσης, τα συστήματα αυτόματης απόδειξης δεν παρέχουν τη δυνατότητα στον χρήστη να επιλέξει τη στρατηγική επίλυσης για την επίλυση του προβλήματος και την εξαγωγή συμπερασμάτων. Επίσης, οι αλληλεπιδραστικοί αποδεικτές δεν παρέχουν ως δυνατότητα στους εκπαιδευόμενους, την εκμάθηση των στρατηγικών επίλυσης και τη βηματική απόδειξη θεωρημάτων μέσω διαφορετικών στρατηγικών. Επιπλέον, τα παραπάνω συστήματα δεν προσφέρουν μηχανισμούς προσδιορισμού της δυσκολίας των ασκήσεων και προσαρμογής του συστήματος με βάση τις ανάγκες των εκπαιδευόμενων καθώς και μηχανισμούς αυτόματης αξιολόγησης και δημιουργίας ασκήσεων.

2.3 Σημασιολογικός Ιστός

Ο σημασιολογικός ιστός (Semantic Web) αποτελεί μια επέκταση του σημερινού ιστού και έχει ως σκοπό την αυτοματοποίηση των λειτουργιών και των εφαρμογών του διαδικτύου. Ο σημασιολογικός ιστός περιλαμβάνει τεχνικές που υπόσχονται να βελτιώσουν τον υπάρχοντα παγκόσμιο ιστό και να καλύψει τα κενά και τις αδυναμίες του. Ως στόχους έχει τη βελτιωμένη αναζήτηση, την εκτέλεση σύνθετων διεργασιών και την εξατομίκευση της πληροφορίας σύμφωνα με τις ανάγκες του κάθε χρήστη. Ο Tim Berners-Lee, ο εμπνευστής του Παγκόσμιου Ιστού, δημιουργός και καθοδηγητής, ορίζει τον σημασιολογικό ιστό ως εξής (Berners-Lee et al., 2001):

“The Semantic Web is an extension of the current Web in which information is given well-defined meaning, better enabling computers and people to work in cooperation. ... a web of data that can be processed directly and indirectly by machines”

Η ανάπτυξη του σημασιολογικού Ιστού εξελίσσεται σταδιακά, με το κάθε βήμα να δημιουργεί ένα επίπεδο (layer) πάνω σε κάποιο άλλο. Ο σημασιολογικός ιστός είναι οργανωμένος σε μία σειρά επιπέδων. Τα επίπεδα αυτά δεν ανταποκρίνονται αυστηρά στην έννοια της αρχιτεκτονικής λογισμικού, αλλά αποτελούν περισσότερο τεχνολογικά επίπεδα λειτουργικότητας. Η ανάπτυξη του σημασιολογικού ιστού ακολουθεί ένα ιεραρχικό σύστημα επιπέδων, όπου κάθε “ανώτερο” επίπεδο επεκτείνει το “κατώτερό” του. Ο λόγος της ύπαρξης αρκετών διαφορετικών επιπέδων βοηθάει τους ερευνητές να αναπτύσσουν και να επεξεργάζονται συγκεκριμένα στοιχεία του. Η επιλογή αυτή βοηθάει στο να μην υπάρχουν αποκλίσεις σχετικά με την ανάπτυξη των προτύπων από τις επιστημονικές ομάδες. Επομένως, αν ο σημασιολογικός ιστός αντιμετωπιζόταν ως ένα ενιαίο επιστημονικό πεδίο δεν θα υπήρχε κατακερματισμός των εργασιών (εξειδίκευση) σε όλα τα επίπεδα και η ανάπτυξή του θα ήταν πιο αργή. Στην εικόνα 2.1, παρουσιάζεται ο «πύργος επιπέδων» του σημασιολογικού ιστού του Tim Berners-Lee, ο οποίος περιγράφει τα κύρια επίπεδα της σχεδίασης και του οράματος του σημασιολογικού ιστού.



Εικόνα 2.1 Διαστρωματωμένη προσέγγιση του Σημασιολογικού Ιστού

Όπως αναφέρθηκε και πιο πάνω ο σημασιολογικός ιστός αποτελεί εξέλιξη και επέκταση του παγκόσμιου ιστού, άρα χρησιμοποιεί την ήδη υπάρχουσα δομή του παγκόσμιου ιστού. Όπως παρουσιάζεται στην εικόνα 2.1, το καθένα από τα επίπεδα εκτελεί και μια ξεχωριστή λειτουργία.

Επίπεδο 1: Στο πρώτο και κατώτερο επίπεδο έχουμε το διαχωρισμό της πληροφορίας σε δύο μέρη. Στο πρώτο μέρος έχουμε το φυσικό επίπεδο επεξεργασίας της πληροφορίας, στο οποίο χρησιμοποιείται η κωδικοποίηση Unicode. Η κωδικοποίηση Unicode είναι ένα διεθνώς καταχωρημένο πρότυπο κωδικοποίησης χαρακτήρων, το οποίο χρησιμοποιείται για την αποθήκευση όλων των χαρακτήρων όλων των γλωσσών του πλανήτη, συμπεριλαμβανομένων και των ιδεογραμμάτων, των σημείων στίξης, των μαθηματικών συμβόλων και γενικά οποιουδήποτε χαρακτήρα μπορεί να παράγει το πληκτρολόγιο. Στο δεύτερο μέρος, έχουμε το λογικό επίπεδο, το οποίο χρησιμοποιεί τη διευθυνσιοδότηση των εγγράφων με βάση το πρωτόκολλο URI. Το πρωτόκολλο URI εξασφαλίζει ότι το όνομα κάθε ιστοσελίδας θα είναι μοναδικό.

Επίπεδο 2: Στο επίπεδο αυτό χρησιμοποιείται η γλώσσα σήμανσης XML, η οποία επιτρέπει στο χρήστη τη συγγραφή δομημένων εγγράφων, με λεξιλόγιο που ορίζει ο ίδιος ο χρήστης.

Επίπεδο 3: Κύριο ρόλο στο επίπεδο αυτό έχει το μοντέλο RDF, το οποίο χρησιμοποιείται για τη συγγραφή απλών δηλώσεων για τα αντικείμενα—πόρους του ιστού. Επίσης, το RDF σχήμα, το οποίο βασίζεται στη δομή του RDF και προσφέρει τη δυνατότητα ιεραρχικής οργάνωσης των αντικειμένων του ιστού, χρησιμοποιώντας κλάσεις, ιδιότητες και σχέσεις μεταξύ των αντικειμένων.

Επίπεδο 4: Οι οντολογίες (ontologies) αποτελούν επέκταση του RDF σχήματος και προσφέρουν τη δυνατότητα αναπαράστασης πιο περίπλοκων σχέσεων μεταξύ των αντικειμένων του ιστού. Για τη σχεδίαση οντολογιών έχουν χρησιμοποιηθεί, το RDF Schema και οι γλώσσες οντολογιών DAML + OIL και OWL, κυρίως η δεύτερη, που έχει καταστεί και πρότυπο της W3C (World Wide Web Consortium).

Επίπεδο 5: Στο επίπεδο της λογικής (logic) δίνεται η δυνατότητα για εξαγωγή συμπερασμάτων με τη χρήση κανόνων λογικής. Επιπλέον, επιτρέπει την αυτοματοποιημένη εξαγωγή συμπερασμάτων βάσει των δομών και των πληροφοριών που υπάρχουν στην οντολογία.

Επίπεδο 6: Το επίπεδο απόδειξης (proof) δίνει τη δυνατότητα, μέσω μιας μηχανής λογικής ανάλυσης και εξαγωγής συμπερασμάτων, τα δεδομένα που δημιουργήθηκαν από το συμπέρασμα να μπορούν να οδηγήσουν πίσω στις υποθέσεις που τα δημιούργησαν. Δηλαδή, επιτρέπει τη δημιουργία εξηγήσεων για τα συμπεράσματα στα οποία οδηγηθήκαμε από το επίπεδο της λογικής.

Επίπεδο 7: Το επίπεδο της εμπιστοσύνης (trust) εξασφαλίζει ότι η πληροφορία που δημιουργείται και αποστέλλεται μέσω του σημασιολογικού ιστού είναι αξιόπιστη. Αυτό γίνεται και με τη βοήθεια των ψηφιακών υπογραφών (digital signatures), που καθιστούν το επίπεδο αυτό το σημαντικότερο όλης της διαδικασίας.

2.3.1 Τεχνολογίες Σημασιολογικού Ιστού και Οντολογίες

Για να μπορέσει να χρησιμοποιηθεί στην πράξη ο σημασιολογικός ιστός θα πρέπει να δημιουργηθούν καινούργια εργαλεία που να αξιοποιούν τις δυνατότητές του και να μπορούν να γίνουν αποδεκτά από τους χρήστες.

Μεταδεδομένα (Metadata)

Ένα από τα πρόβλημα του Ιστού σήμερα είναι πως το περιεχόμενο του είναι δομημένο με τέτοιο τρόπο ώστε να γίνεται περισσότερο κατανοητό από τους ανθρώπους και λιγότερο από τα προγράμματα. Οι πληροφορίες σε μία ιστοσελίδα παρουσιάζονται με τρόπο ικανοποιητικό για τον άνθρωπο, αλλά η επεξεργασία τους από προγράμματα δεν είναι εύκολη. Υπάρχουν μεταδεδομένα, που δημιουργούνται από τη δομή της γλώσσας HTML και τις προτιμήσεις των χρηστών, αλλά και μεταδεδομένα που σχετίζονται με αντικείμενα, ιδέες, διαδικασίες καθώς και τις μεταξύ τους σχέσεις. Τα μεταδεδομένα δηλαδή κατέχουν μέρος της σημασίας-νοήματος των αποθηκευμένων δεδομένων. Η αποθήκευση μέρους της πληροφορίας των δεδομένων εξηγεί και το όνομα του ιστού ως σημασιολογικός ιστός. Ο σημασιολογικός ιστός χρησιμοποιεί τις γλώσσες XML και RDF, οι οποίες, εκτός από τις πληροφορίες μορφοποίησης που αναπαριστούν, περιέχουν και πληροφορίες σχετικά με τη σημασία των δεδομένων που εμφανίζονται.

Οντολογίες

Η οντολογία ασχολείται με το τι υπάρχει ή καλύτερα με την έννοια της ύπαρξης, ενώ ορισμένες φορές κατά μια γενικότερη έννοια πραγματεύεται τη μελέτη του τι μπορεί να υπάρχει. Ο πιο γνωστός ορισμός για την οντολογία, στην επιστήμη των υπολογιστών, πάνω στον οποίο στηρίχτηκαν και άλλοι ορισμοί, δόθηκε από τον Gruber (Gruber, 1995) και είναι ο ακόλουθος:

“Μια οντολογία είναι μια τυπική (formal), ρητή (explicit) προδιαγραφή μιας διαμοιρασμένης (shared) εννοιολογικής αναπαράστασης (conceptualization).”

Ο δεύτερος πιο δημοφιλής ορισμός διατυπώθηκε από τον Guarino (Guarino, 1995), ο οποίος προσπαθεί να απομονώσει την οντολογία από την φιλοσοφία:

“Η οντολογία είναι ένα σύνολο λογικών αξιωμάτων που δίνονται για να αποδώσουν το επιδιωκόμενο νόημα ενός λεξικού”

Η επιλογή της γλώσσας είναι ένα από τα βασικότερα στάδια για την ανάπτυξη μίας οντολογίας. Η εκφραστικότητα της γλώσσας αποδεικνύεται ως το σημαντικότερο χαρακτηριστικό, καθώς προσδιορίζει το είδος των σχέσεων και των εννοιών που θα αναπαρασταθούν. Αν και υπάρχουν αρκετές αναγνωρίσιμες από την μηχανή γλώσσες, ο σχεδιαστής της οντολογίας πρέπει να λάβει υπόψη του το βαθμό εκφραστικότητας της κάθε γλώσσας και να επιλέξει ανάλογα με τις ανάγκες του (Gómez-Pérez & Corcho, 2002). Από τα παραπάνω είναι φανερό ότι χρειαζόμαστε μια γλώσσα για να αναπαραστήσουμε μία οντολογία σε έναν ηλεκτρονικό υπολογιστή. Οι κύριες διαφορές ανάμεσα στις γλώσσες αναπαράστασης οντολογιών έχουν να κάνουν με την εκφραστικότητα της κάθε γλώσσας και τους υπολογιστικούς πόρους που είναι διαθέσιμοι. Οι πρώτες γλώσσες αναπαράστασης οντολογιών που αναπτύχθηκαν ήταν οι KIF και OKBC. Οι δύο αυτές γλώσσες δεν παρείχαν τη δυνατότητα στο χρήστη να αναπαριστά οντολογίες στο Σημασιολογικό Ιστό. Το μειονέκτημά τους αυτό σε συνδυασμό με την ανάπτυξη της XML οδήγησαν στη δημιουργία νέων βελτιωμένων γλωσσών, οι οποίες είναι κατάλληλες για την αναπαράσταση της γνώσης στο Σημασιολογικό Ιστό.

Η Semantic Web Rule Language (SWRL) αποτελεί την πρόταση του W3C σχετικά με τη χρήση κανόνων σε συνδυασμό με οντολογίες. Η SWRL προήλθε από τον συνδυασμό της γλώσσας OWL με τη RuleML (Rule Markup Language). Η SWRL είναι μια γλώσσα κανόνων που επεκτείνει την OWL DL με μια μορφή κανόνων διατηρώντας όμως την συμβατότητα προς τα πίσω με την υπάρχουσα σημασιολογία της OWL και επίσης επιτρέπει τη χρήση κανόνων τύπου Horn στα αξιώματα που προσφέρει η OWL.

Antecedent $\rightarrow$ Consequent

Το αριστερό μέλος των κανόνων (Antecedent) ονομάζεται σώμα (body) ενώ το δεξί μέλος (Consequent) ονομάζεται κεφαλή (head). Οι κανόνες ορίζονται ως εξής:

Κάθε φορά που οι συνθήκες που προδιαγράφονται από το Antecedent ισχύουν, θα πρέπει να ισχύουν και οι συνθήκες που προδιαγράφονται από το Consequent.

Ο αριθμός των ατόμων στην κεφαλή και στο σώμα ενός κανόνα μπορεί να είναι μηδενικός. Ένας κανόνας που δεν έχει ούτε ένα άτομο στο σώμα του ερμηνεύεται πάντοτε ως αληθής από οποιαδήποτε ερμηνεία. Ένας κανόνας με μηδενικό πλήθος ατόμων στην κεφαλή του δεν

μπορεί να επιφέρει καμία αλλαγή στην βάση γνώσης του συστήματος. Στην SWRL οι μεταβλητές που βρίσκονται στο σώμα και στην κεφαλή του κανόνα συμβολίζονται με το αγγλικό ερωτηματικό. Η ερμηνεία του παραδείγματος:

$$has_father(?a,?b) \wedge has_married(?b,?c) \rightarrow female(?c) \wedge has_mother(?a,?c)$$

είναι:

“Αν ο *a* έχει πατέρα τον *b* και ο *b* έχει παντρευτεί την *c* τότε η *c* είναι γυναίκα και μητέρα του *a*”.

Αξίζει να σημειωθεί πως ένας κανόνας με πολλαπλά άτομα στην κεφαλή του ισοδυναμεί με πολλαπλούς κανόνες που έχουν το ίδιο σώμα με τον αρχικό κανόνα και ένα άτομο στην κεφαλή τους. Για παράδειγμα ο παρακάτω κανόνας :

$$has_father(?a,?b) \wedge has_married(?b,?c) \rightarrow female(?c) \wedge has_mother(?a,?c)$$

είναι ισοδύναμος με τους επόμενους δύο κανόνες:

$$has_father(?a,?b) \wedge has_married(?b,?c) \rightarrow female(?c)$$

$$has_father(?a,?b) \wedge has_married(?b,?c) \rightarrow has_mother(?a,?c)$$

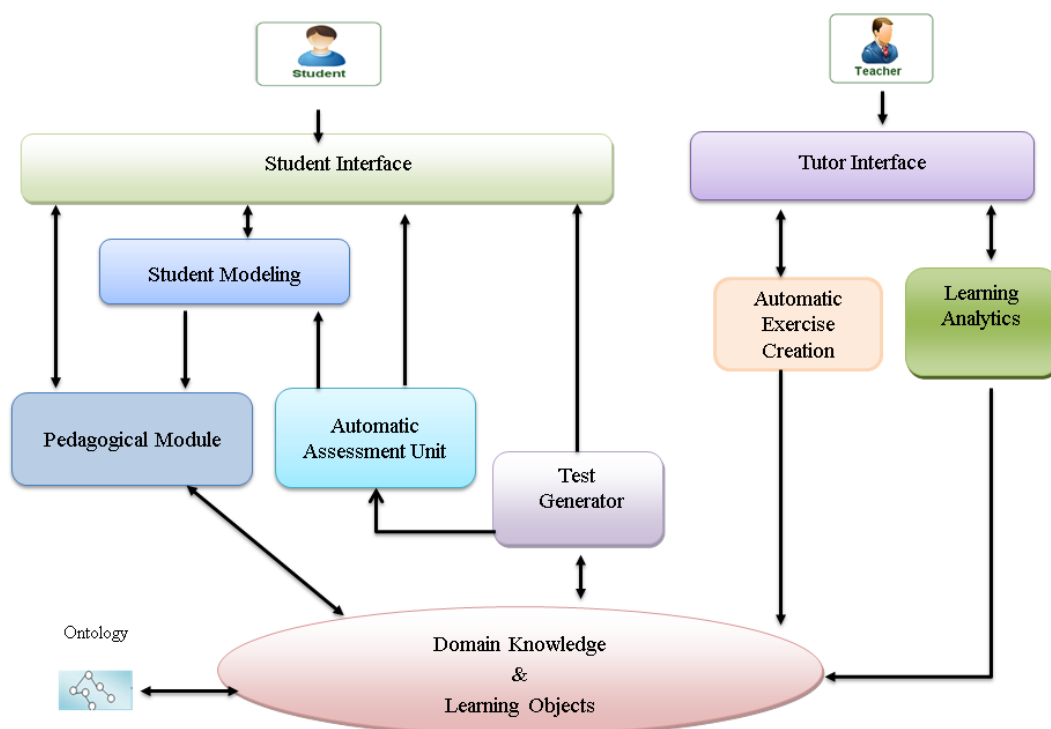
2.4 Ευφύες Σύστημα Διδασκαλίας Τεχνητής Νοημοσύνης

Το ευφύες σύστημα AITS έχει ως στόχο να βοηθήσει τους εκπαιδευόμενους στην εκμάθηση θεμάτων της τεχνητής νοημοσύνης όπως είναι η αναπαράσταση γνώσης, οι αλγόριθμοι αναζήτησης, η ικανοποίηση περιορισμών μέσω διαφόρων αλληλεπιδραστικών εκπαιδευτικών διαδικασιών και να προσφέρει προσαρμοζόμενη μάθηση στις ανάγκες των εκπαιδευόμενων. Το σύστημα σχεδιάστηκε και αναπτύχθηκε με βάση τεχνολογίες σημασιολογικού ιστού (Grivokostopoulou et al, 2014d) όπως είναι οι οντολογίες οι οποίες χρησιμοποιούνται για να αναπαρασταθεί η γνώση του πεδίου της τεχνητής νοημοσύνης που διδάσκει το σύστημα, οι περιορισμοί και οι διασυνδέσεις μεταξύ των εννοιών. Επίσης, χρησιμοποιήθηκαν κανόνες SWRL για να γίνει η αξιολόγηση των εκπαιδευόμενων με βάση τις επιδόσεις τους στις ασκήσεις και τα τεστ.

2.4.1 Αρχιτεκτονική Συστήματος

Η αρχιτεκτονική του ευφυούς Συστήματος Διδασκαλίας Τεχνητής Νοημοσύνης (AITS) παρουσιάζεται στην εικόνα 2.2 και αποτελείται από τις ακόλουθες μονάδες: Διεπαφή Εκπαιδευόμενου (Student Interface), Μοντελοποίηση Μαθητή (Student Modeling), Παιδαγωγική Μονάδα (Pedagogical Module), Μονάδα Αυτόματης Αξιολόγησης (Automatic

Assessment Unit), Γεννήτρια Τεστ (Test Generator), Διεπαφή Διδάσκοντα (Tutor Interface), Αυτόματη Δημιουργία Ασκήσεων (Automatic Exercise Creation), Στοιχεία Ανάλυσης Μάθησης (Learning Analytics) και Γνώση Πεδίου και Μαθησιακά Αντικείμενα (Domain Knowledge & Learning Objects). Με τις μονάδες αυτές αλληλεπιδρά τόσο ο εκπαιδευμένος, για να μελετήσει και να μάθει έννοιες σχετικές με το πεδίο της Τεχνητής Νοημοσύνης, όσο και ο διδάσκοντας, για την διαχείριση της διδασκαλία και της ανάλυσης των μαθησιακών δεδομένων των εκπαιδευόμενων.



Εικόνα 2.2 Αρχιτεκτονική Συστήματος Διδασκαλίας Τεχνητής Νοημοσύνης

Η *διεπαφή εκπαιδευόμενου (SI)* έχει σχεδιαστεί ώστε να καθιστά το εκπαιδευτικό σύστημα εύκολο στην χρήση. Αρχικά, ο χρήστης εγγράφεται στο σύστημα καταχωρώντας τις απαιτούμενες προσωπικές πληροφορίες όπως, για παράδειγμα, το όνομα, το έτος σπουδών κ.ά. Μετά την εγγραφή, κάθε εκπαιδευόμενος μπορεί ανά πάσα στιγμή να εισέλθει και να χρησιμοποιήσει το σύστημα.

Το *μοντέλο μαθητή*, που δημιουργείται μέσω της μονάδας *μοντελοποίηση μαθητή (SM)*, έχει σχεδιαστεί να αναπαριστά και να αποθηκεύει πληροφορίες σχετικά με τους εκπαιδευόμενους. Πιο συγκεκριμένα, αποθηκεύει αναλυτικές πληροφορίες για τον κάθε εκπαιδευόμενο όπως,

προσωπικές πληροφορίες, ενέργειες που έχει πραγματοποιήσει ο εκπαιδευόμενος στο σύστημα, οι επιδόσεις του, το επίπεδο γνώσεων σε κάθε έννοια κ.α. Με βάση το μοντέλο μαθητή, λαμβάνονται οι παιδαγωγικές αποφάσεις μέσω της παιδαγωγικής μονάδας και γίνονται οι κατάλληλες ενέργειες προσαρμογής της εκπαιδευτικής διαδικασίας στις ανάγκες του κάθε εκπαιδευόμενου.

Η *γεννήτρια τεστ* είναι μια μονάδα παραγωγής ασκήσεων, που χρησιμοποιείται για να παράγει προσαρμοσμένα τεστ με ένα συγκεκριμένο αριθμό ασκήσεων. Αποτελείται από ένα *σύστημα βασισμένο σε κανόνες (rule-based system)* το οποίο υλοποιεί τη λήψη αποφάσεων σχετικών με το επίπεδο δυσκολίας των ασκήσεων που συμπεριλαμβάνονται στο τεστ, προκειμένου αυτό να είναι προσαρμοσμένο στο επίπεδο γνώσης και στις ανάγκες του κάθε εκπαιδευόμενου (Hatzilygeroudis et al., 2006). Στην συνέχεια, δημιουργούνται μια σειρά από τεστ, που περιέχουν ασκήσεις διάφορων τύπων, ανάλογα με το θεματικό που εξετάζεται.

Η *διεπαφή διδάσκοντα (TI)* έχει σχεδιαστεί ώστε να προσφέρει την δυνατότητα στους διδάσκοντες να αλληλεπιδρούν με το σύστημα και να έχουν πρόσβαση σε εργαλεία και μηχανισμούς για τη διαχείριση του εκπαιδευτικού υλικού του συστήματος και την επίβλεψη και την ανάλυση των επιδόσεων των εκπαιδευόμενων.

Ο μηχανισμός για την *αυτόματη δημιουργία ασκήσεων* σχεδιάστηκε με στόχο να βοηθήσει τον διδάσκοντα στη δημιουργία νέου εκπαιδευτικού υλικού. Μέσω του μηχανισμού, ο διδάσκοντας μπορεί να προσδιορίσει μια σειρά από παραμέτρους της άσκησης και με βάση αυτές να παραχθούν αυτόματα μια σειρά από κατάλληλες ασκήσεις ανάλογα με το πεδίο της εφαρμογής.

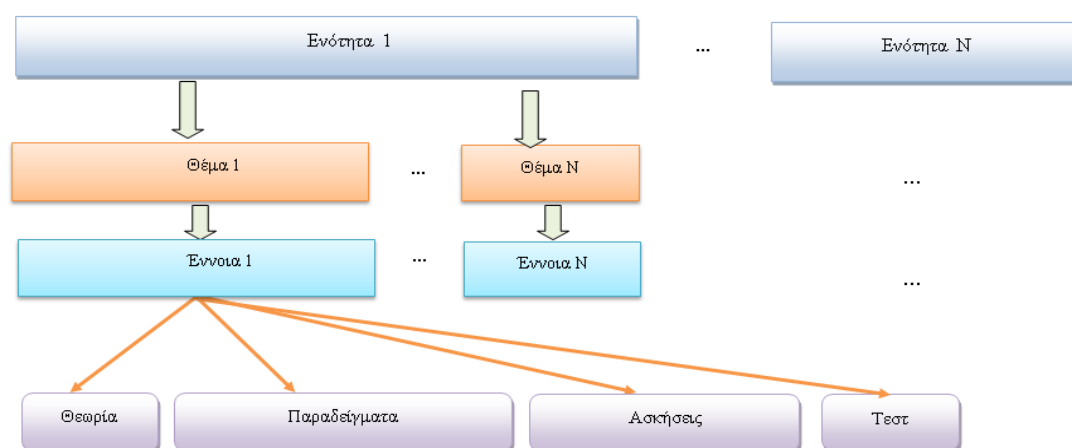
Η *μονάδα αυτόματης αξιολόγησης* έχει ως στόχο την ανάλυση και την αυτόματη αξιολόγηση των απαντήσεων των εκπαιδευόμενων σε ασκήσεις και εκπαιδευτικές δραστηριότητες καθώς και να παρέχει στους εκπαιδευόμενους άμεσα, με ακρίβεια και με συνέπεια την ανάλυση των επιδόσεών τους και των λαθών που έγιναν, χωρίς να χρειάζεται να παρεμβαίνει ο διδάσκοντας.

Η μονάδα *στοιχεία ανάλυσης μάθησης* αναλύει τα δεδομένα σχετικά με τις ενέργειες των εκπαιδευόμενων στο σύστημα και έχει ως στόχο να εξάγει χρήσιμες πληροφορίες από αυτές. Από την ανάλυση της διαδικασίας μάθησης των εκπαιδευόμενων, εξάγονται στατιστικές πληροφορίες σχετικές με τα συνηθισμένα λάθη που πραγματοποιούν οι εκπαιδευόμενοι, καθώς και οι έννοιες και οι διαδικασίες που τους δημιουργούν ιδιαίτερες δυσκολίες. Επίσης, οπτικοποιούνται τα αποτελέσματα που προκύπτουν με στόχο να βοηθηθεί ο διδάσκοντας στον προσδιορισμό και την καλύτερη κατανόηση των δυσκολιών που αντιμετωπίζουν οι εκπαιδευόμενοι. Επιπλέον, η μονάδα βασίζεται σε τεχνικές εξόρυξης γνώσης προκειμένου να

πραγματοποιήσει προβλέψεις σχετικά με τις επιδόσεις των εκπαιδευόμενων (Grivokostopoulou et al., 2014c). Με βάση τις εξαγόμενες πληροφορίες, ο διδάσκοντας και το σύστημα μπορούν να αποκτήσουν μια πιο αναλυτική περιγραφή των χαρακτηριστικών των εκπαιδευόμενων και να γίνει καλύτερη προσαρμογή των εκπαιδευτικών διαδικασιών στις ανάγκες και στις επιδόσεις τους.

2.4.2 Εκπαιδευτική Διαδικασία

Το εκπαιδευτικό περιεχόμενο του συστήματος AITS περιλαμβάνει τις εξής βασικές θεματικές περιοχές: «Αναπαράσταση Γνώσης και Συλλογιστική», «Αλγόριθμοι Αναζήτησης» και «Ικανοποίηση Περιορισμών» καθώς επίσης υπάρχει η δυνατότητα για εύκολη επέκταση του εκπαιδευτικού υλικού και σε άλλες έννοιες. Το εκπαιδευτικό περιεχόμενο του συστήματος αναπαριστάται ως δενδρικές δομές, όπου η ρίζα κάθε δέντρου αποτελεί μια *ενότητα* (π.χ. Αλγόριθμοι Αναζήτησης, Αναπαράσταση γνώσης, Ικανοποίηση Περιορισμών, κλπ), η κάθε ενότητα χωρίζεται σε *θέματα* και στην συνέχεια το κάθε θέμα περιέχει διάφορες *έννοιες*. Με τον τρόπο αυτό, κάθε ενότητα περιέχει με μια σειρά από θέματα και κάθε θέμα ασχολείται με μια σειρά από έννοιες. Κάθε έννοια περιέχει *θεωρία*, *παραδείγματα*, *ασκήσεις* και *τεστ* σχετικά με την έννοια. Στην εικόνα 2.3 παρουσιάζεται η δεντρική αναπαράσταση του εκπαιδευτικού υλικού. Η ιεραρχία της δεντρικής αναπαράστασης γίνεται με την εισαγωγή προτεραιοτήτων για τα στοιχεία του δέντρου (ενότητες, θέματα, έννοιες) που καθορίζονται αρχικά από τον διδάσκοντα και μπορούν να τροποποιηθούν κατάλληλα. Ο σχεδιασμός του εκπαιδευτικού περιεχομένου του συστήματος είναι απόλυτα επεκτάσιμος και τροποποιήσιμος και ο διδάσκοντας έχει την δυνατότητα να το διαχειριστεί ελεύθερα, ανάλογα με τις εκπαιδευτικές ανάγκες. Επίσης, το πλήθος των ενοτήτων εξαρτάται από τα χαρακτηριστικά του κάθε θέματος, για παράδειγμα, το θέμα Αναπαράσταση Γνώσης και Συλλογιστική αποτελείται από τρία κύρια θέματα την «Προτασιακή Λογική», την «Κατηγορηματική λογική» και την «Συλλογιστική» και κάθε ένα θέμα έχει πολλές έννοιες.



Εικόνα 2.3 Δεντρική Αναπαράσταση του Εκπαιδευτικού Υλικού

Στα πλαίσια της αλληλεπίδρασης των εκπαιδευόμενων με το σύστημα, ο κάθε εκπαιδευόμενος έχει την δυνατότητα να μελετήσει την θεωρία, τα παραδείγματα και τις ασκήσεις και να πραγματοποιήσει τεστ σε κάθε έννοια. Πιο συγκεκριμένα, το σύστημα παρέχει την δυνατότητα στον εκπαιδευόμενο να μελετήσει τις έννοιες των θεμάτων που έχει ασχοληθεί, αλλά δεν μπορεί να παραβιαστεί η ιεραρχία που έχει οριστεί αρχικά. Για παράδειγμα, δεν μπορεί ο εκπαιδευόμενος να μελετήσει μία έννοια αν δεν έχει μελετήσει νωρίτερα όλες τις προαπαιτούμενες. Ο σκοπός της ιεραρχίας είναι να βοηθήσει τον εκπαιδευόμενο να μελετάει και να μαθαίνει τις έννοιες του μαθήματος σταδιακά, συμβάλλοντας έτσι στην βελτίωση του επιπέδου γνώσης του. Πιο συγκεκριμένα, στο επίπεδο της έννοιας ο εκπαιδευόμενος μπορεί να αλληλεπιδράσει με διάφορα είδη ασκήσεων που παρέχονται από το σύστημα, όπως αλληλεπιδραστικές, πολλαπλής επιλογής κ.ά., ανάλογα με το θεματικό πεδίο. Μετά από την υποβολή μιας απάντησης του εκπαιδευόμενου, το σύστημα αναλύει και αξιολογεί την απάντηση του και σε περίπτωση λανθασμένης απάντησης παρέχονται κατάλληλες επεξηγήσεις, ανάλογα με το είδος της άσκησης. Στην συνέχεια, το σύστημα προτείνει στον εκπαιδευόμενο να αλληλεπιδράσει με μια νέα άσκηση κατάλληλου επιπέδου δυσκολίας, με βάση τη δυσκολία της άσκησης που προηγουμένως απάντησε σωστά και τις επιδόσεις του. Όταν ολοκληρώσει τη μελέτη μίας έννοιας, μπορεί να πραγματοποιήσει ένα προσαρμοσμένο τεστ σχετικό με την έννοια αυτή. Στην συνέχεια, μέσω του μηχανισμού αυτόματης αξιολόγησης, γίνεται αυτόματα η βαθμολόγηση του τεστ και παρέχεται η βαθμολογία του και κατάλληλη ανάδραση σχετικά με τα λάθη που έγιναν.

2.4.3 Αξιολόγηση Εκπαιδευόμενου

Όπως έχει ήδη αναφερθεί, μια βασική λειτουργία των ευφυών εκπαιδευτικών συστημάτων είναι η προσαρμοστικότητα, και πιο συγκεκριμένα ο τρόπος με το οποίο το εκπαιδευτικό υλικό και οι εκπαιδευτικές διαδικασίες προσαρμόζονται στον κάθε εκπαιδευόμενο. Για τον σκοπό αυτό, μια βασική παράμετρος του AITS αποτελεί το επίπεδο γνώσεων και η επίδοση των εκπαιδευόμενων. Στα πλαίσια του συστήματος, η αξιολόγηση ενός εκπαιδευόμενου είναι πολυδιάστατη και αφορά μια σειρά θεμάτων και εννοιών. Πιο συγκεκριμένα, γίνεται αναλυτική αξιολόγηση των εκπαιδευόμενων σε διάφορα επίπεδα όπως, στο επίπεδο του περιεχομένου, στο επίπεδο εννοιών και σε όλο το εκπαιδευτικό υλικό που έχει μελετήσει και σε όλες τις ενότητες που το αποτελούν. Η αξιολόγηση γίνεται με την διαδικασία προσαρμοστικών τεστ καθώς και μέσω ασκήσεων που παρέχονται ανάλογα με το θεματικό. Όταν ο εκπαιδευόμενος μελετήσει τα θέματα που απαρτίζουν μία ενότητα, τότε μπορεί να αξιολογηθεί στο επίπεδο της ενότητας αυτής. Ο προσδιορισμός του επιπέδου γνώσεων του εκπαιδευόμενου γίνεται με βάση τους επιμέρους βαθμούς σε κάθε έννοια και το επίπεδο του κατηγοριοποιείται σε πέντε κατηγορίες: “excellent”, “very good”, “good”, “medium” και “low”. Το σύστημα, κατά την διάρκεια που ο εκπαιδευμένος μελετάει και μαθαίνει μια έννοια, του παρέχει την δυνατότητα να αξιολογηθεί σε ασκήσεις. Για τον προσδιορισμό του επιπέδου δυσκολίας των ασκήσεων, έχει αναπτυχθεί μηχανισμός ο οποίος καθορίζει το επίπεδο δυσκολίας τους ανάλογα με τις επιδόσεις του εκπαιδευόμενου. Αρχικά οι πρώτες ασκήσεις που θα λάβει ο εκπαιδευόμενος θα είναι χαμηλού επιπέδου δυσκολίας (επίπεδο δυσκολίας very easy, βλ. παρακάτω) και στην συνέχεια, ανάλογα με το επίπεδο του εκπαιδευόμενου και τις απαντήσεις του, το σύστημα προσφέρει ασκήσεις κατάλληλου επιπέδου δυσκολίας με σκοπό την βελτίωση του επιπέδου του και προτρέπει στον εκπαιδευόμενο να ασχοληθεί με ασκήσεις μεγαλύτερης δυσκολίας με στόχο την αύξηση της βαθμολογίας και του επιπέδου του.

Αρχικά, το σύστημα προσδιορίζει για κάθε απάντηση του εκπαιδευόμενου τον επιμέρους βαθμό με βάση την ανάλυση της απάντησής του, καθώς και την δυσκολία της άσκησης (Hatzilygeroudis et al., 2006). Οι ασκήσεις που προσφέρει το σύστημα έχουν αναλυθεί και κατηγοριοποιηθεί σε 5 επίπεδα δυσκολίας τα οποία είναι: “very easy”, “easy”, “medium”, “difficult” και “advanced”.

Προσδιορισμός επιπέδου γνώσεων του εκπαιδευόμενου σε Θέμα

Ο προσδιορισμός του επιπέδου γνώσεων του εκπαιδευόμενου σ’ ένα θέμα γίνεται με βάση το βαθμό στο θέμα (ST), που υπολογίζεται ως εξής:

$$ST = \frac{\sum_{i=1}^n w_i * SC_i}{n}$$

όπου n αναπαριστά τον συνολικό αριθμό των εννοιών που αποτελείται το θέμα, w_i είναι το βάρος της έννοιας i , όπως έχει οριστεί από τον διδάσκοντα, και SC_i είναι ο βαθμός ανά έννοια. Ο βαθμός του εκπαιδευόμενου σε μια έννοια προκύπτει από το μέσο όρο όλων των βαθμών του στις ασκήσεις και τις εκπαιδευτικές δραστηριότητες της συγκεκριμένης έννοιας. Για να βαθμολογηθεί ο εκπαιδευόμενος σε κάποιο θέμα, πρέπει να έχει μελετήσει όλες τις έννοιες που το αποτελούν. Στην συνέχεια, το σύστημα προτείνει στον εκπαιδευόμενο τις επόμενες ενέργειες, ανάλογα με την επίδοσή του σε κάποιο θέμα, όπως παρουσιάζονται στην πίνακα 2.1

Πίνακας 2.1 Προτάσεις AITS με βάση το σκορ στο θέμα

Σκορ στο Θέμα (ST)	Προτάσεις AITS
$ST < 50$	Να γίνει μελέτη του θέματος ξανά.
$50 \leq ST \leq 80$	Μελέτη του θέματος ή εκμάθηση του επόμενου θέματος.
$ST > 80$	Εκμάθηση επόμενου θέματος.

Προσδιορισμός επιπέδου γνώσεων εκπαιδευόμενου σε Ενότητα

Ο προσδιορισμός του επιπέδου γνώσεων του εκπαιδευόμενου σε μια ενότητα γίνεται με βάση το βαθμό στην ενότητα (SS). Ο υπολογισμός του βασίζεται στην βαθμολογία του εκπαιδευόμενου στο κάθε θέμα και υπολογίζεται από τον ακόλουθο τύπο:

$$SS = \frac{\sum_{i=1}^k wt_i * ST_i}{k}$$

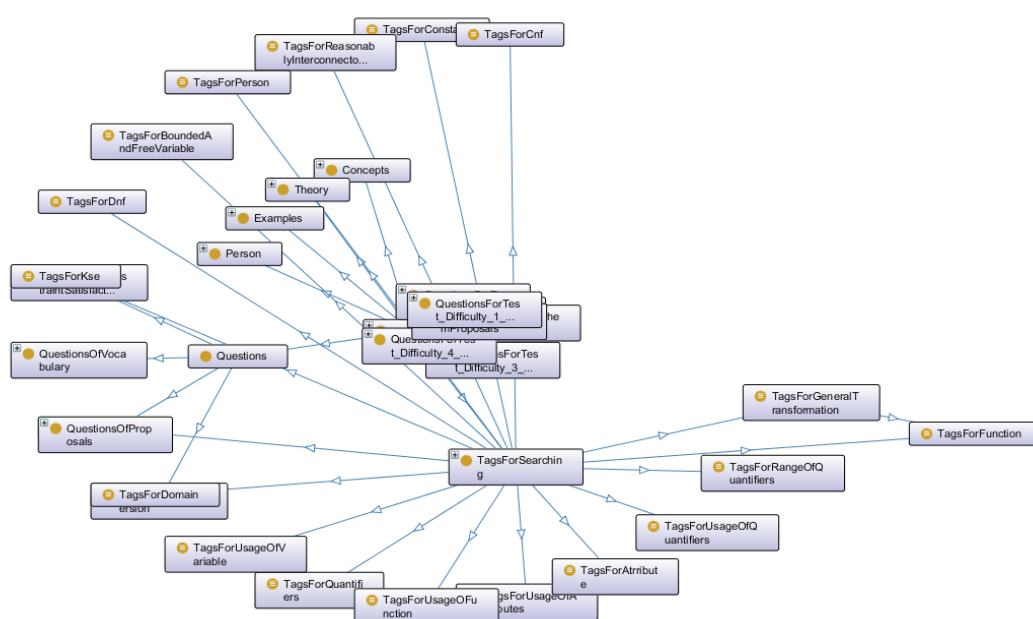
όπου k αναπαριστά τον αριθμό των θεμάτων που αποτελείται η ενότητα, wt_i είναι το βάρος του θέματος i , που ορίζεται αρχικά από τον διδάσκοντα, και ST_i είναι ο βαθμός στο θέμα i . Στην συνέχεια, το σύστημα προτείνει στον εκπαιδευόμενο επόμενες ενέργειες, ανάλογα με το βαθμό του σε κάποια ενότητα όπως παρουσιάζονται στον πίνακα 2.2

Πίνακας 2.2 Προτάσεις AITS με βάση το βαθμό στην ενότητα

Σκορ στην Ενότητα (SS)	Προτάσεις AITS
$SS < 50$	Να γίνει μελέτη ξανά της ενότητας.
$50 \leq SS \leq 80$	Μελέτη του Θέματος ή εκμάθηση επόμενης ενότητας.
$SS > 80$	Εκμάθηση επόμενης ενότητας.

2.4.4 Οντολογία AITS

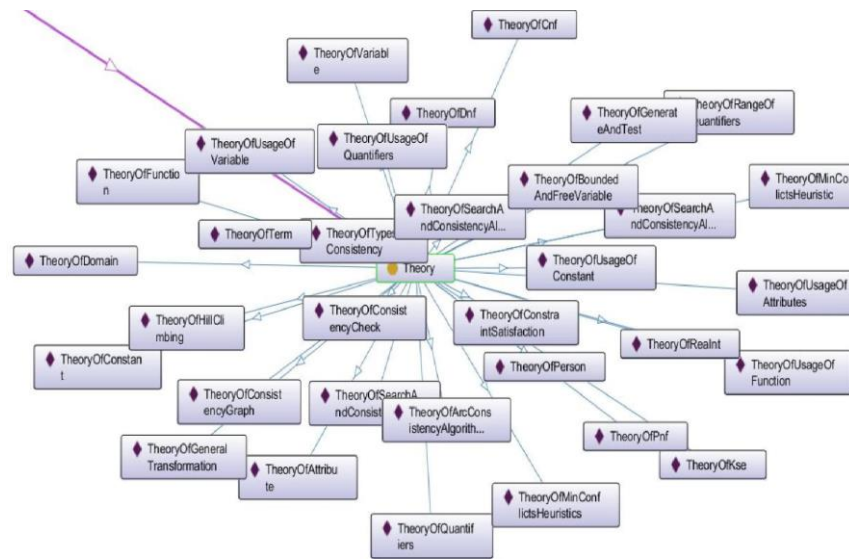
Η υλοποίηση της οντολογίας και των σημασιολογικών κανόνων SWRL, βασίστηκε στο εργαλείο ανοιχτού κώδικα Protégé. Η οντολογία του συστήματος περιλαμβάνει όλη την απαραίτητη πληροφορία σχετικά με τον χρήστη και το εκπαιδευτικό υλικό καθώς και αντίστοιχους κανόνες SWRL. Στην εικόνα 2.4 παρουσιάζεται μια γραφική απεικόνιση της οντολογίας.



Εικόνα 2.4 Γραφική Απεικόνιση της Οντολογίας

Στην συνέχεια παρουσιάζονται μια σύντομη περιγραφή μερικών κλάσεων και των αντίστοιχων ιδιοτήτων τους.

Κλάση Theory: Η κλάση αυτή διατηρεί το θεωρητικό υλικό του συστήματος και περιέχει όλα τα αντικείμενα της θεωρίας της κάθε έννοιας. Περιέχει το υλικό που θα μελετήσει ο εκπαιδευόμενος προκειμένου να κατανοήσει και να επιλύσει τις ασκήσεις. Στην εικόνα 2.5 παρουσιάζεται η γραφική απεικόνιση της κλάσης theory και αντίστοιχα στον πίνακα 2.3 οι ιδιότητες της κλάσης.



Εικόνα 2.5 Γραφική Απεικόνιση της κλάση Theory

Πίνακας 2.3 Ιδιότητες της κλάσης Theory

Ιδιότητα	Είδος Ιδιότητα	Περιγραφή
TheoryBelongToConcept	data, string	Η έννοια στην οποία αναφέρεται η συγκεκριμένη θεωρία.
TheoryHasAnnouncement	data, string	Προσδιορίζει την εκφώνηση της θεωρίας.
TheoryHasPriority	data, int	Η προτεραιότητα για κάθε θέμα.
TheoryHasThem	data, string	Το θέμα στο οποίο ανήκει η συγκεκριμένη θεωρία.

Κλάση Examples: Η κλάση αυτή περιέχει τα παραδείγματα που σχετίζονται με τις έννοιες που πραγματεύεται η θεωρία. Κάθε φορά που ο εκπαιδευόμενος διαβάζει τη θεωρία μίας έννοιας αποκτά πρόσβαση και στο αντίστοιχο παράδειγμα.

Πίνακας 2.4 Ιδιότητες της Κλάσης Examples

Ιδιότητα	Είδος Ιδιότητα	Περιγραφή
ExampleBelongToConcept	data, string	Η έννοια στην οποία ανήκει το συγκεκριμένο παράδειγμα.
ExampleHasAnnouncement	data, string	Προσδιορίζει το παράδειγμα

Η *Κλάση Questions* περιέχει τις ασκήσεις τις οποίες καλείται να επιλύσει ο εκπαιδευόμενος και έχει οργανωθεί ανάλογα με το θέμα της κάθε ενότητας. Για παράδειγμα, υπάρχουν οι

κλάσεις *QuestionsOfAlgorithms*, *QuestionsOfConstraintSatisfactionProblems* που περιέχουν ασκήσεις που σχετίζονται με την έννοια των αλγορίθμων αναζήτησης και της ικανοποίησης περιορισμών. Επίσης, υπάρχουν κλάσεις ανά θέμα, όπως η κλάση *QuestionsForThemHeuristic*, η κλάση *QuestionsForThemCSP*. Η κάθε κλάση είναι η υπερκλάση των κλάσεων που περιέχουν τεστ για κάθε θεματική ενότητα. Τα τεστ που θα λάβει ο κάθε εκπαιδευόμενος καθορίζονται από κανόνες SWRL με βάση το επίπεδο του.

Η *Κλάση QuestionsForTest*. Η κλάση αυτή είναι υπερκλάση ενός συνόλου κλάσεων στις οποίες αποθηκεύονται τα τελικά τεστ στα οποία θα εξεταστεί ο εκπαιδευόμενος ανάλογα με το επίπεδό του.

Επίσης, εκτός από την οντολογία του συστήματος για την λειτουργία της μονάδας αξιολόγησης έχουν αναπτυχθεί κανόνες SWRL. Στην συνέχεια παρουσιάζεται ένα παραδείγμα κανόνα για τον προσδιορισμό του βαθμού δυσκολίας.

```
QuestionsOfConstraintSatisfactionProblems(?p),QuestionHasDegreeOfDifficulty(?p,?d),  
equal(?d, 5) -> QuestionsForThemCSPWithDifficulty4_5(?p).
```

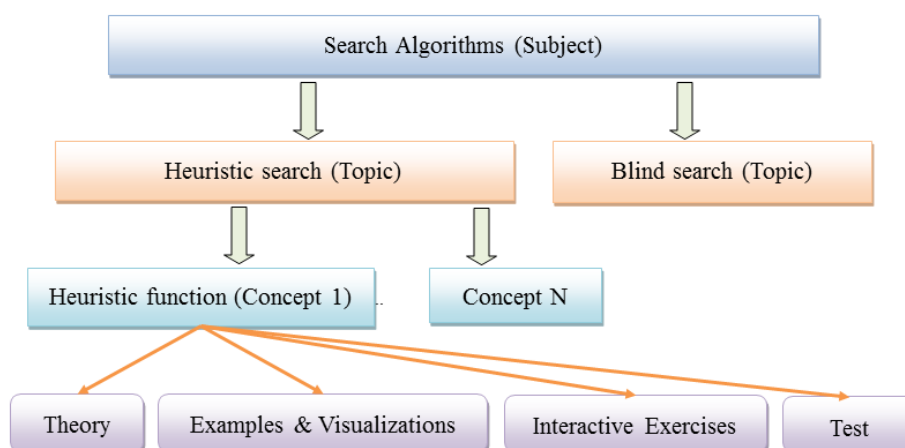
Σκοπός του παραπάνω κανόνα είναι να αντιστοιχήσει τις ασκήσεις βαθμού δυσκολίας 5 (advanced) στην αντίστοιχη ομάδα-κλάση ασκήσεων, η οποία συλλέγει ασκήσεις του θέματος «Ικανοποίησης Περιορισμών» με βαθμό δυσκολίας 4 ή 5.

2.4.5 Διδασκαλία Αλγορίθμων Αναζήτησης με το Σύστημα AITS

Στην τεχνητή νοημοσύνη, ένα θεμελιώδες πεδίο είναι οι αλγόριθμοι αναζήτησης, το οποίο αποτελεί μέρος σε κάθε μάθημα τεχνητής νοημοσύνης (Russel & Norvig, 2003). Οι αλγόριθμοι αναζήτησης βρίσκουν εφαρμογή σε διάφορα πεδία και η εις βάθος κατανόηση και η εφαρμογή τους από τους εκπαιδευόμενους είναι πολύ σημαντική. Επίσης, από την πλευρά του διδάσκοντα, η διδασκαλία των αλγορίθμων είναι ιδιαίτερα απαιτητική καθώς χρειάζεται να επεξηγηθεί αναλυτικά ο τρόπος λειτουργίας τους, τα βήματα τους και οι καταστάσεις από τις οποίες διέρχεται. Επομένως, στα πλαίσια της διδασκαλίας των αλγορίθμων, ο διδάσκοντας χρειάζεται, εκτός από την επίλυση των προβλημάτων αναζήτησης στον πίνακα, να παρέχει στους εκπαιδευόμενους αναλυτικές επεξηγήσεις και γραφικές αναπαραστάσεις για να τους βοηθήσει στην καλύτερη κατανόηση του τρόπου λειτουργίας των αλγορίθμων (Baecker, 1998). Επίσης, από τη πλευρά των εκπαιδευόμενων, το πεδίο των αλγορίθμων περιέχει πολυσύνθετες διαδικασίες που είναι δύσκολες να τις κατανοήσουν σε βάθος, δεδομένου ότι

περιλαμβάνουν σύνθετες έννοιες και συμβολισμούς (Shabanah et al., 2010). Επίσης, όταν οι εκπαιδευόμενοι μαθαίνουν νέες έννοιες, όπως οι ευρετικοί αλγόριθμοι, είναι δύσκολο να τις κατανοήσουν χωρίς την ύπαρξη και την διασύνδεση τους με κατάλληλα παραδείγματα. Σε αυτή την κατεύθυνση, η λειτουργία της οπτικοποίησης των αλγορίθμων, όπως και η αλληλεπιδραστική εφαρμογή των αλγορίθμων σε ασκήσεις και εκπαιδευτικές διαδικασίες, έχει ως στόχο να βοηθήσει τους εκπαιδευόμενους να συνδέσουν τις έννοιες και τις διαδικασίες του πεδίου με συγκεκριμένα παραδείγματα και να τις εφαρμόσουν με την βοήθεια του συστήματος. Στόχος, είναι να συμβάλουν στο να κατανοήσουν οι εκπαιδευόμενοι τη λειτουργία των αλγορίθμων εις βάθος και ταχύτερα από ότι μέσα από την απλή περιγραφή τους σε κείμενο και από τις στατικές αναπαραστάσεις (Shabanah et al., 2010)

Το θεματικό πεδίο των αλγορίθμων ακολουθεί την γενική δομή του περιεχομένου του AITS και περιέχει δυο σημαντικά θέματα, όπως παρουσιάζονται στη εικόνα 2.6.



Εικόνα 2.6 Δεντρική Αναπαράσταση των Αλγορίθμων Αναζήτησης

Η ενότητα "Theory" αποτελείται από το θεωρητικό μέρος των εννοιών των αλγορίθμων, ενώ η ενότητα "Examples & Visualizations" αποτελείται από παραδείγματα που παρουσιάζουν και επεξηγούν μέσω κατάλληλων γραφικών αναπαραστάσεων τον τρόπο λειτουργίας των αλγορίθμων σε διάφορα προβλήματα. Η ενότητα "Interactive Exercises" περιλαμβάνει αλληλεπιδραστικές ασκήσεις που καλούνται να επιλύσουν οι εκπαιδευόμενοι βήμα προς βήμα εφαρμόζοντας και καθορίζοντας κάθε φορά μια συγκεκριμένη ενέργεια με βάση τον αλγόριθμο αναζήτησης. Οι αλληλεπιδραστικές ασκήσεις είναι δύο ειδών, ασκήσεις εξάσκησης και ασκήσεις αξιολόγησης. Οι ασκήσεις εξάσκησης είναι αλληλεπιδραστικές ασκήσεις στις οποίες οι εκπαιδευόμενοι καλούνται να εφαρμόσουν τους αλγορίθμους αναζήτησης σε διάφορα προβλήματα και καταστάσεις με την βοήθεια και την καθοδήγηση του

συστήματος. Στον αντίποδα, οι *ασκήσεις αξιολόγησης* χρησιμοποιούνται για να εξετάσουν την πρόοδο των εκπαιδευόμενων και την κατανόηση των αντίστοιχων εννοιών.

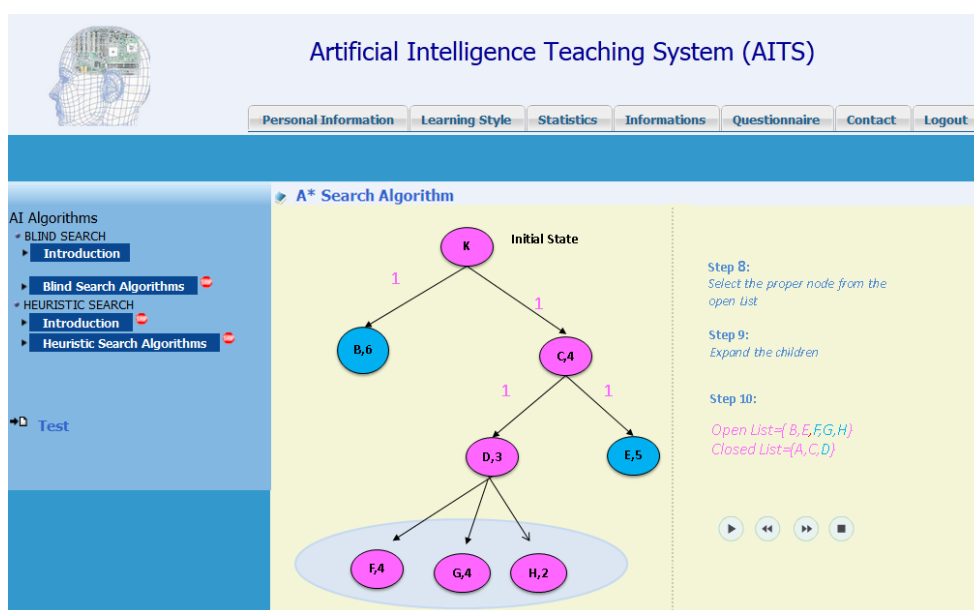
Η αξιολόγηση των απαντήσεων των εκπαιδευόμενων είναι χρήσιμη τόσο για τους εκπαιδευόμενους όσο και για το σύστημα. Πιο συγκεκριμένα, το σύστημα μέσω της ανάλυσης των επιδόσεων των εκπαιδευόμενων μπορεί να προσδιορίσει καλύτερα τις δεξιότητες και τις επιδόσεις τους και να προσαρμόσει κατάλληλα την εκπαιδευτική διαδικασία στις ανάγκες τους. Επίσης, η αυτοαξιολόγηση μπορεί να βοηθήσει τους εκπαιδευόμενους να εντοπίσουν ελλείψεις γνώσεων και αδυναμίες σε συγκεκριμένες έννοιες και να τις μελετήσουν περαιτέρω. Στα πλαίσια του συστήματος, προσδιορίζεται αυτόματα το επίπεδο δυσκολίας όλων των ασκήσεων μέσω των μηχανισμών που διαθέτει το σύστημα (βλ. κεφ. 4) όλες οι απαντήσεις των εκπαιδευόμενων αξιολογούνται αυτόματα από τον μηχανισμό αυτόματης αξιολόγησης (Grivokostopoulou et al., 2016a) (βλ. κεφ. 5).

Οπτικοποίηση των αλγορίθμων

Οι εκπαιδευόμενοι στο AITS μπορούν να μελετήσουν τις θεωρητικές πτυχές ενός αλγορίθμου μαζί με κατάλληλες εξηγήσεις και απεικονίσεις του αλγορίθμου σε διάφορες καταστάσεις και προβλήματα. Οι απεικονίσεις αποτελούν μια πολύ καλή προσέγγιση για να επεξηγήσουν αναλυτικά τη λειτουργία ενός αλγορίθμου (Hundhausen et al., 2002). Οι απεικονίσεις των αλγορίθμων μπορούν να βοηθήσουν τους εκπαιδευόμενους να κατανοήσουν βαθύτερα τον τρόπο που λειτουργίας του αλγορίθμου, καθώς και τα βήματα του αλγορίθμου με βάση παραμέτρους, όπως είναι η ευρετική συνάρτηση και το κόστος (Hansen et al., 2002; Naps et al., 2002).

Το AITS κατά την διάρκεια των οπτικοποιήσεων, παρέχει κατάλληλες επεξηγήσεις για τις ενέργειες και τις αποφάσεις του αλγορίθμου. Πιο συγκεκριμένα, αναλύει τον τρόπο που υπολογίστηκαν τιμές, όπως το ευρετικό, το κόστος κ.ά., καθώς και την επιλογή του κατάλληλου κόμβου που θα επισκεφτεί. Επίσης, ένα σημαντικό χαρακτηριστικό των οπτικοποιήσεων στο AITS είναι ότι σχεδιάστηκαν με βάση την προσέγγιση της ενεργητικής μάθησης (active learning). Σε αυτή την κατεύθυνση, οι εκπαιδευτικές δραστηριότητες σχεδιάστηκαν έτσι ώστε να εμπλέκουν τον χρήστη όσο το δυνατόν περισσότερο, για να αποφεύγεται η παθητική παρακολούθηση και να τον κάνουν να συμμετέχει και να σκέφτεται εις βάθος τις λειτουργίες του κάθε βήματος του αλγορίθμου, καθώς και την εφαρμογή τους. Οι αρχές της ενεργητικής μάθησης αναφέρουν ότι όσο περισσότερο οι εκπαιδευόμενοι ενεργούν στις εκπαιδευτικές δραστηριότητες, τόσο μεγαλύτερη θα είναι η νοητική προσπάθεια τους και τα μαθησιακά αποτελέσματά τους.

Σε αυτό το πλαίσιο, κατά την διάρκεια μιας οπτικοποίησης, η βηματική αναπαράσταση της εφαρμογής ενός αλγορίθμου σε μία άσκηση μπορεί να σταματήσει σε κάποιο τυχαίο βήμα και να ζητήσει από το εκπαιδευόμενο να προσδιορίσει τις τιμές των παραμέτρων του αλγορίθμου. Πιο συγκεκριμένα, η οπτικοποίηση μπορεί να ζητήσει από τον εκπαιδευόμενο να προσδιορίσει και να δικαιολογήσει την επόμενη ενέργεια που θα γίνει από τον αλγόριθμο. Για παράδειγμα, μια οπτικοποίηση μπορεί να σταματήσει και να ζητήσει από τον εκπαιδευόμενο να προσδιορίσει τον κατάλληλο επόμενο κόμβο που θα επισκεφτεί ο αλγόριθμος και σε περίπτωση σωστής επιλογής, μπορεί ζητηθεί από τον εκπαιδευόμενο να απαντήσει ερωτήσεις πολλαπλής επιλογής δικαιολογώντας την συγκεκριμένη απόφαση. Σε περίπτωση λανθασμένης απάντησης, παρέχονται άμεσα κατάλληλες εξηγήσεις στον εκπαιδευόμενο. Στην εικόνα 2.7, παρουσιάζεται ένα στιγμιότυπο της βηματικής οπτικοποίησης της λειτουργίας του αλγορίθμου A*.

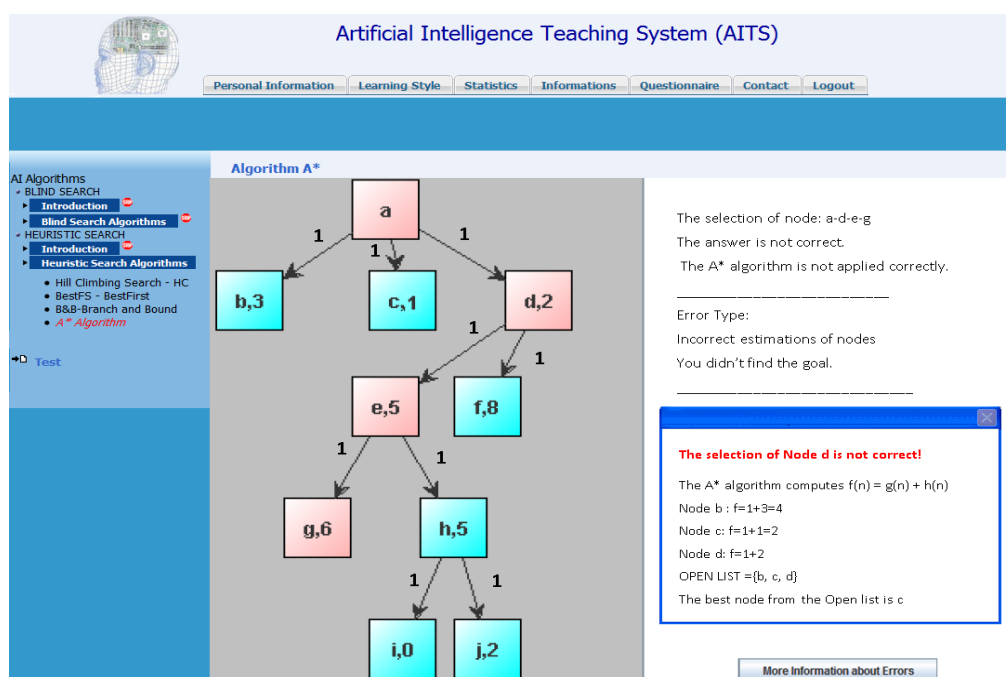


Εικόνα 2.7 Οπτικοποίηση της λειτουργίας του A* σε ένα παράδειγμα

Αλληλεπιδραστικές ασκήσεις

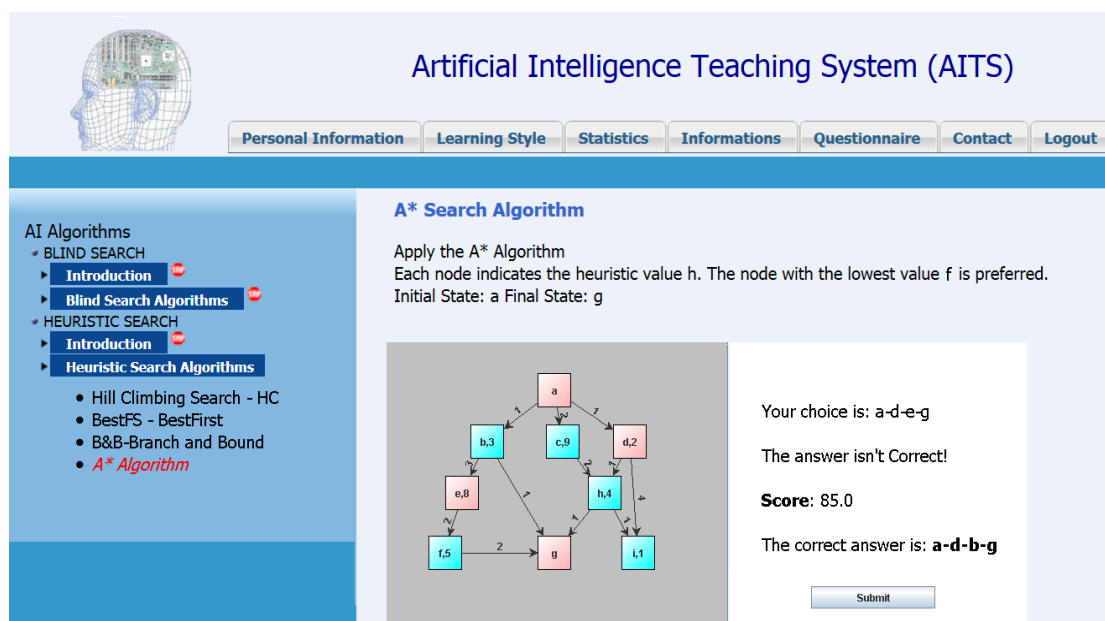
Το σύστημα, όπως αναφέραμε και παραπάνω, προσφέρει την δυνατότητα στους εκπαιδευόμενους να επιλύσουν αλληλεπιδραστικές ασκήσεις, οι οποίες αφορούν ασκήσεις εξάσκησης και ασκήσεις αξιολόγησης (Grivokostoroulou & Hatzilygeroudis, 2013c). Στις ασκήσεις εξάσκησης, καλούνται να επιλύσουν ένα πρόβλημα εφαρμόζοντας κάποιο συγκεκριμένο αλγόριθμο αναζήτησης και προσδιορίζοντας βήμα προς βήμα όλους τους

κόμβους που χρειάζεται να επισκεφτεί ο αλγόριθμος για να βρει τη λύση. Επίσης, στα πλαίσια μια άσκησης μπορεί να ζητηθεί από τον εκπαιδευμένο να εφαρμόσει και να προσδιορίσει όλη την λύση, ξεκινώντας από την ρίζα του δέντρου ή ένα μέρος της (ξεκινώντας από κάποιον ενδιάμεσο κόμβο) είτε να καθορίσει τον επόμενο κόμβο που θα επισκεφθεί ο αλγόριθμος. Στην εικόνα 2.8 παρουσιάζεται ένα παράδειγμα άσκησης εξάσκησης πάνω στον A* αλγόριθμο, καθώς επίσης η απάντηση του εκπαιδευόμενου, η βοήθεια του συστήματος και η αντίστοιχη αξιολόγηση που προσδιορίστηκε από το σύστημα.



Εικόνα 2.8 Άσκηση αξιολόγησης στον αλγόριθμο A*

Επίσης, το σύστημα παρέχει αλληλεπιδραστικές ασκήσεις αξιολόγησης που έχουν ως στόχο να αξιολογήσουν το επίπεδο γνώσης και κατανόησης του εκπαιδευόμενου. Οι ασκήσεις αξιολόγησης μπορούν να χρησιμοποιηθούν από τους εκπαιδευόμενους για να αξιολογήσουν τις επιδόσεις, καθώς και από το σύστημα για να προσδιοριστεί το επίπεδο γνώσεων τους, να παραχθεί εξατομικευμένη βοήθεια και να προσαρμοστεί το εκπαιδευτικό υλικό στο επίπεδο του και τις ανάγκες τους (Aleven et al., 2010; Rosa & Eskenazi, 2013). Επίσης, στις ασκήσεις αξιολόγησης παρέχεται βοήθεια και ανάδραση, αφού ολοκληρώσουν και αποστέλλουν τις απαντήσεις τους οι εκπαιδευόμενοι. Στην συνέχεια, οι απαντήσεις αναλύονται, προσδιορίζονται τα λάθη που έγιναν και παρέχεται ανάλυση των λαθών και η βαθμολόγηση κάθε απάντησης. Στην εικόνα 2.9, παρουσιάζεται μια άσκηση αξιολόγησης για τον αλγόριθμο A* καθώς και η αντίστοιχη βαθμολόγηση της απάντησης.



The screenshot shows the AITS web application. At the top, there's a header with a brain icon and the title "Artificial Intelligence Teaching System (AITS)". Below the header is a navigation bar with links: Personal Information, Learning Style, Statistics, Informations, Questionnaire, Contact, and Logout. The main content area is divided into two sections. On the left, under "AI Algorithms", there's a list of search algorithms: BLIND SEARCH (Introduction, Blind Search Algorithms), HEURISTIC SEARCH (Introduction, Heuristic Search Algorithms), and a list of specific algorithms: Hill Climbing Search - HC, BestFS - BestFirst, B&B-Branch and Bound, and A* Algorithm. The right section is titled "A* Search Algorithm" and contains instructions: "Apply the A* Algorithm. Each node indicates the heuristic value h. The node with the lowest value f is preferred. Initial State: a Final State: g". Below the text is a search tree diagram. The root node is 'a' (h,3). It has two children: 'b' (h,3) and 'c' (h,9). Node 'b' has two children: 'd' (h,8) and 'e' (h,8). Node 'c' has one child: 'f' (h,4). Node 'd' has one child: 'g' (h,5). Node 'e' has one child: 'h' (h,1). Node 'f' has one child: 'i' (h,1). The diagram shows the path from 'a' to 'b' to 'd' to 'g' highlighted. On the right side of the interface, there's a feedback section: "Your choice is: a-d-e-g", "The answer isn't Correct!", "Score: 85.0", and "The correct answer is: a-d-b-g". There is a "Submit" button at the bottom.

Εικόνα 2.9 Άσκηση εξάσκησης πάνω στον αλγόριθμο αναζήτησης A*

Ανάδραση και βοήθεια

Ένα βασικό χαρακτηριστικό του συστήματος είναι ότι παρέχει άμεσα αναλυτική ανάδραση και βοήθεια στους εκπαιδευόμενους. Η βοήθεια και η ανάδραση του συστήματος είναι ένας από τους βασικότερους παράγοντες, που συμβάλλουν στην αποτελεσματική μάθηση των εκπαιδευόμενων (Perikos et al., 2016; Narciss, 2008). Σε αυτή την κατεύθυνση, το σύστημα παρέχει την δυνατότητα στους εκπαιδευόμενους να ζητήσουν βοήθεια για να προσδιορίσουν τη σωστή απάντηση. Πιο συγκεκριμένα, αν η απάντηση του εκπαιδευόμενου δεν είναι σωστή, το σύστημα παράγει κατάλληλη βοήθεια και μπορεί να την λάβει ο εκπαιδευόμενος οποτεδήποτε το επιθυμεί. Επίσης, ο μηχανισμός ανάδρασης του AITS χρησιμοποιείται για να παρέχει άμεσα αναλυτική βοήθεια και καθοδήγηση στους εκπαιδευόμενους κατά την διάρκεια των εκπαιδευτικών δραστηριοτήτων τους.

The screenshot shows the AITS system interface. On the left is a sidebar with navigation links: 'Algorithms', 'BLIND SEARCH', 'Introduction', 'Blind Search Algorithms', 'HEURISTIC SEARCH', 'Introduction', 'Heuristic Search Algorithms', and 'Test'. The main area is titled 'Online Test' and contains 'Exercise 1: Select nodes following the Hill Climbing algorithm. Initial State: A'. Below this is a search tree diagram with nodes labeled a, b, c, d, e, f, g, h, i and their heuristic values. The tree shows a path from node a (4) to b (3) to e (8) to f (5) to g (2) to i (0). To the right of the tree is a 'Feedback-Score' section. It shows the same tree with some nodes highlighted in red. The feedback text says: 'Answer: a(2)-d(2)-h(4)-g(2). The Answer is not correct! Errors: Node H: The Node is wrong. Node G: The Node G is wrong. It isn't goal. The heuristic value of the goal is 0!'. Below this, it says: 'The Hill-Climbing select the child with the best heuristic value. The Node H is not correct. Node I has the best heuristic value. Score: 25'.

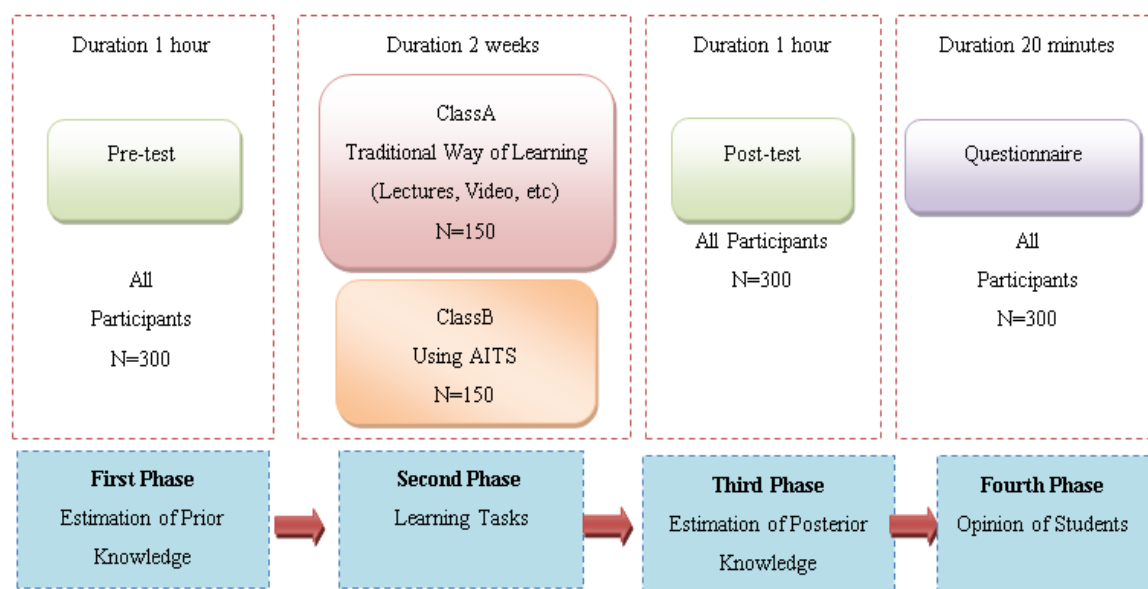
Εικόνα 2.10 Ανάδραση του συστήματος μετά την απάντηση του εκπαιδευόμενου

Πιο συγκεκριμένα, στις ασκήσεις εξάσκησης, παρέχεται η δυνατότητα να ζητηθεί βοήθεια από τον εκπαιδευόμενο πριν και μετά την υποβολή μιας απάντησης. Επομένως, ο εκπαιδευόμενος μπορεί να ζητήσει βοήθεια ενώ μελετάει και αλληλεπιδρά με μια άσκηση και πριν υποβάλλει μια απάντηση σε αυτήν. Μια απάντηση σε μια αλληλεπιδραστική άσκηση εξαρτάται κυρίως από το είδος και τα χαρακτηριστικά της άσκησης και γενικά αφορά τον προσδιορισμό των βημάτων ενός αλγορίθμου που προσπαθεί να λύσει ένα πρόβλημα. Αυτό μπορεί να αφορά τον προσδιορισμό όλων των κόμβων της λύσης, ενός μέρος αυτών ή ακόμη και μόνο του επόμενου κόμβου που θα επισκεφτεί ο αλγόριθμος. Η βοήθεια του συστήματος λειτουργεί σε δυο επίπεδα. Στο πρώτο επίπεδο, το σύστημα αναλύει την απάντηση του εκπαιδευόμενου και τον ενημερώνει σχετικά με την ορθότητα της απάντησής του. Στη συνέχεια, το σύστημα αναγνωρίζει τα ακριβή λάθη που έγιναν και παρέχονται κατάλληλες υποδείξεις στον εκπαιδευόμενο και καθοδήγηση για να διορθώσει τα λάθη του. Επίσης, το σύστημα, μέσω του αυτόματου μηχανισμού αξιολόγησης που διαθέτει, προσδιορίζει και ενημερώνει τον εκπαιδευόμενο για την βαθμολογία του σε κάθε άσκηση.

2.4.6 Πειραματικά Αποτελέσματα συστήματος AITS

Για να μελετήσουμε την αποτελεσματικότητα του συστήματος AITS σχεδιάστηκε και πραγματοποιήθηκε μια πειραματική μελέτη στα πλαίσια του μαθήματος της τεχνητής νοημοσύνης. Οι συμμετέχοντες στην μελέτη ήταν 300 προπτυχιακοί φοιτητές (άνδρες και γυναίκες) που παρακολουθούσαν το μάθημα της τεχνητής νοημοσύνης. Όλοι οι συμμετέχοντες ήταν φοιτητές στο 4^ο και 5^ο έτος σπουδών και η ηλικία τους ήταν από 21 έως 24 χρονών (M=22.5). Για να αξιολογήσουμε την αποτελεσματικότητα της μάθησης των εκπαιδευόμενων μέσω του συστήματος, χρησιμοποιήσαμε τη μέθοδο pretest-posttest, δηλ. τη

χρήση ενός τεστ πριν από το πείραμα και ενός τέστ μετά το πείραμα. Δημιουργήσαμε δύο ομάδες φοιτητών. Η πρώτη ομάδα (ClassA) αποτελείτο από 150 φοιτητές (80 άνδρες και 70 γυναίκες) και η δεύτερη ομάδα (ClassB) από 150 φοιτητές (77 άνδρες και 73 γυναίκες). Οι φοιτητές κάθε ομάδας επιλέχθηκαν τυχαία από ένα σύνολο 250 φοιτητών της τάξης.



Εικόνα 2.11 Η διαδικασία της πειραματικής μελέτης για το σύστημα AITS

Στα πλαίσια της μελέτης, η ομάδα ClassA (ομάδα ελέγχου-control group) μελέτησε το πεδίο των αλγορίθμων με βάση την παραδοσιακή προσέγγιση και δεν χρησιμοποίησε το σύστημα AITS. Οι φοιτητές παρακολούθησαν τις διαλέξεις του μαθήματος, μελέτησαν τις σημειώσεις και τα παραδείγματα και συζήτησαν τα αποτελέσματα με τον διδάσκοντα. Αντίθετα, οι φοιτητές της ομάδας ClassB (ομάδα πειραματισμού-experimental group) μελέτησαν το πεδίο των αλγορίθμων και επέλυσαν διάφορους τύπους ασκήσεων με την βοήθεια του συστήματος.

Η πειραματική μελέτη περιελάμβανε τέσσερις φάσεις: τη διεξαγωγή τεστ πριν την πειραματική μελέτη (pretest), τη “φάση της μάθησης”, την διεξαγωγή του τελικού τεστ (posttest) και την τελική φάση “συμπλήρωση ερωτηματολογίου”, όπως απεικονίζονται στην εικόνα 2.11. Αρχικά, οι φοιτητές και των δυο ομάδων συμμετείχαν στο pretest. Στη συνέχεια η ομάδα ClassA μελέτησε με τον παραδοσιακό τρόπο τους αλγορίθμους αναζήτησης, ενώ η ομάδα ClassB μέσω αλληλεπίδρασης με το AITS. Μετά την ολοκλήρωση της φάσης της μελέτης, διάρκειας 2 εβδομάδων, και οι δύο ομάδες συμμετείχαν σ’ ένα τελικό τεστ, το posttest. Οι ασκήσεις που περιελάμβαναν το pretest και το posttest ήταν ιδίου επιπέδου

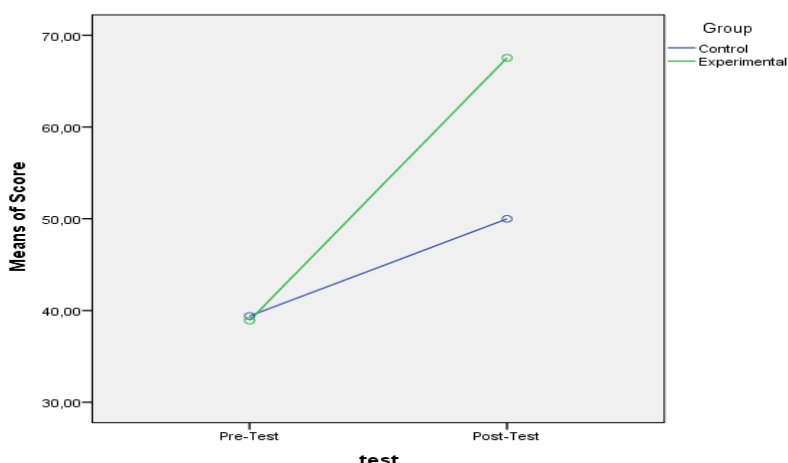
δυσκολίας. Μετά την συμμετοχή στο τελικό τεστ οι συμμετέχοντες συμπλήρωσαν ερωτηματολόγια.

Για να αναλύσουμε τις επιδόσεις των εκπαιδευόμενων αρχικά πραγματοποιήθηκε στατιστικός έλεγχος υποθέσεων και πιο συγκεκριμένα έγινε σύγκριση των μέσων τιμών των δυο ανεξάρτητων δειγμάτων *t-test (Independent Sample t-test)* για τις επιδόσεις των εκπαιδευόμενων στο pretest. Τα αποτελέσματα έδειξαν ότι η μέση τιμή και η τυπική απόκλιση στο pretest ήταν 39.4 και 10.74 για την ομάδα ClassA ($M = 39.4$, $SD = 10.74$), και 38.9 και 11 για την ClassB ($M = 38.9$, $SD = 11.0$) αντίστοιχα. Επίσης, η τιμή p (επίπεδο σημαντικότητας) ήταν $p = 0.691$ (> 0.05), $t = 0.398$ και το *effect size d* (Cohen, 1988) ήταν 0.046. Από τα αποτελέσματα φαίνεται ότι δεν υπάρχει στατιστικά σημαντική διαφορά μεταξύ των επιδόσεων των δυο ομάδων στο pretest. Συνεπώς, οι ομάδες ClassA και ClassB είχαν το ίδιο επίπεδο γνώσεων στις έννοιες των αλγορίθμων αναζήτησης πριν από την έναρξη της διαδικασίας μάθησης. Στον Πίνακα 2.5 παρουσιάζονται οι τιμές των περιγραφικών μέτρων των δύο ομάδων στο pretest και το posttest αντίστοιχα.

Πίνακας 2.5 Περιγραφικά μέτρα pretest και posttest στους αλγορίθμους αναζήτησης

Group	Pretest		Posttest	
	M	SD	M	SD
ClassA (control)	39.4	10.74	49.98	11.18
ClassB (experimental)	38.9	11.00	67.54	14.06

Τα αποτελέσματα δείχνουν ότι οι φοιτητές της ClassB, ενώ στο pretest είχαν μέση τιμή 38.9, αυτή αυξήθηκε στο posttest στο 67.54, ενώ η μέση τιμή της ομάδας ClassA από 39.4 στο pretest αυξήθηκε στο 49.98 στο posttest. Επομένως, τα αποτελέσματα δείχνουν μια αύξηση της μέσης τιμής για την ομάδα ClassB, η οποία είναι αρκετά μεγαλύτερη από την μέση τιμή της ομάδας ClassA στο posttest. Για να προσδιοριστεί η αποτελεσματικότητα της μάθησης, πραγματοποιήσαμε *ανάλυση διακύμανσης (ANOVA)* με επαναλαμβανόμενες μετρήσεις για να ελέγξουμε τη διαφορά μεταξύ των δύο ομάδων (control, experimental). Τα αποτελέσματα δείχνουν μια σημαντική διαφορά στην απόδοση της μάθησης με $F(1, 298) = 235.32$, $p < 0.001$ και το μέγεθος της επίδρασης *effect size* = 0.44. Στην εικόνα 2.12 παρουσιάζονται οι επιδόσεις της κάθε ομάδας στο pretest και το posttest αντίστοιχα.



Εικόνα 2.12 Πειραματικά αποτελέσματα pretest και posttest πάνω στους αλγορίθμους αναζήτησης

Επίσης, υπολογίσαμε το *κέρδος μάθησης (learning gains)* ως τη διαφορά *posttest – pretest*, που εκφράζει τη βελτίωση της μάθησης του εκπαιδευόμενου μετά την φάση της μάθησης. Στην συνέχεια, έγινε ανάλυση ANOVA για το *απλό κέρδος μάθησης (simple learning gains)* και έδειξε σημαντική διαφορά μεταξύ των παραμέτρων, $F(1, 298)=235.32$, $p < 0.001$, $MS_{Error}=103.87$. Επιπροσθέτως, υπολογίσαμε την *κανονικοποιημένη μάθηση (normalized learning)* ως εξής:

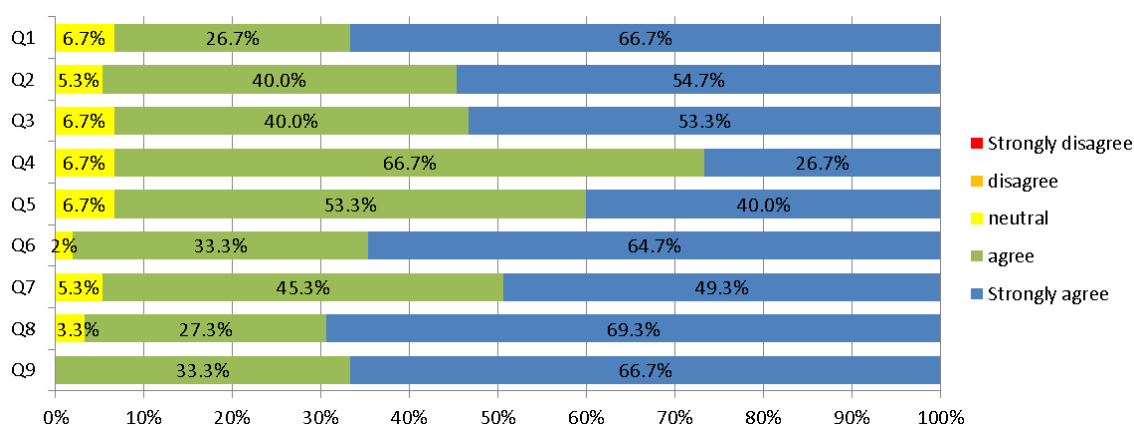
$$\frac{posttest - pretest}{1 - pretest}$$

Η ανάλυση ANOVA που έγινε στα *κανονικοποιημένα κέρδη (normalized gains)* έδειξε σημαντικές διαφορές μεταξύ των παραμέτρων, $F(1, 298)=91.28$, $p < 0.001$, $MS_{Error}=282.83$. Επίσης, από τα αποτελέσματα προκύπτει ότι η απόδοση των φοιτητών της ομάδας ClassB, που χρησιμοποίησαν το AITS, ήταν καλύτερη από της ομάδας ClassA.

Επίσης, μετά το posttest, οι φοιτητές και των δύο ομάδων συμπλήρωσαν ένα ερωτηματολόγιο, για να διατυπώσουν τις εμπειρίες τους και την γνώμη τους σχετικά με την επίδραση που είχε το σύστημα στην διαδικασία μάθησής τους. Το ερωτηματολόγιο της ομάδας ClassB είχε 12 ερωτήσεις. Οι απαντήσεις στις ερωτήσεις Q1-Q9 ήταν τύπου Likert scale (1: καθόλου, 5: παρά πολύ) και υπόλοιπες τρεις ερωτήσεις ήταν ανοικτού τύπου, για να εκφράσουν οι φοιτητές τις εμπειρίες και την γνώμη τους. Στον πίνακα 2.6 παρουσιάζονται οι ερωτήσεις και στην εικόνα 2.13 τα πειραματικά αποτελέσματα του ερωτηματολογίου για την ClassB.

Πίνακας 2.6 Αποτελέσματα της ανάλυσης των ερωτηματολογίων του AITS

Ερώτηση	
Q1	Βοήθησε το AITS να καταλάβετε καλύτερα τους αλγορίθμους αναζήτησης;
Q2	Η ανάδραση και η καθοδήγηση του συστήματος ήταν ικανοποιητική.
Q3	Πόσο βοήθησε η αυτόματη αξιολόγηση των απαντήσεων στην κατανόηση των λαθών και των αλγορίθμων αναζήτησης;
Q4	Νιώθετε ικανοποιημένος με τους βαθμούς του αυτόματου μηχανισμού αξιολόγησης;
Q5	Η βαθμολογία ήταν ακριβής;
Q6	Πως θα αξιολογούσατε την συνολική μαθησιακή σας εμπειρία με το σύστημα;
Q7	Νιώθετε με την χρησιμοποίηση του συστήματος πιο σίγουροι στην επίλυση ασκήσεων αλγορίθμων αναζήτησης;
Q8	Θα συνιστούσατε το σύστημα στους συμφοιτητές σας και θα θέλατε να ενσωματωθεί στο μάθημα;
Q9	Οι οπτικοποιήσεις με βοήθησαν να καταλάβω πιο αποτελεσματικά τον τρόπο λειτουργίας των αλγορίθμων.



Εικόνα 2.13 Αποτελέσματα Ερωτηματολογίου

Τα αποτελέσματα των ερωτηματολογίων είναι πολύ ικανοποιητικά. Τα αποτελέσματα των απαντήσεων των φοιτητών στις ερωτήσεις Q1-Q9 δείχνουν ότι η πλειοψηφία των φοιτητών βρήκε την αλληλεπίδραση με το σύστημα ευχάριστη και ότι το σύστημα τους βοήθησε να μάθουν καλύτερα τους αλγορίθμους αναζήτησης. Επίσης, το 93% των φοιτητών βρήκαν την αξιολόγηση που παρέχεται από το σύστημα να είναι πολύ ακριβής και το 94% των φοιτητών απάντησαν ότι η οπτικοποίηση των αλγορίθμων τους βοήθησε να καταλάβουν καλύτερα τον τρόπο λειτουργίας τους. Μετά την ανάλυση των απαντήσεων των φοιτητών προσδιορίστηκε ο δείκτης αξιοπιστίας Cronbach's alpha (Cronbach, 1951). Οι τιμές του δείκτη κυμαίνονται

από 0 μέχρι 1 και αποτελεί έναν από τους πιο γνωστούς και ευρέως χρησιμοποιούμενους δείκτες εσωτερικής συνάφειας (Hatcher, 1994). Ο δείκτης αξιοπιστίας ήταν πολύ καλός ($\alpha = 0.78$) για τους φοιτητές της ομάδας ClassB.

2.5 Ευφυή Συστήματα Διδασκαλίας Λογικής

Η λογική αποτελεί ένα πολύ σημαντικό επιστημονικό πεδίο της τεχνητής νοημοσύνης (TN) και αποτελεί και ένα δύσκολο πεδίο για τη διδασκαλία, που περιέχει σύνθετες έννοιες και διαδικασίες. Πολλοί διδάσκοντες αναγνωρίζουν ότι η λογική περιλαμβάνει σύνθετα θέματα και έννοιες που δυσκολεύουν τους εκπαιδευόμενους να τα κατανοήσουν. Μια βασική όψη της λογικής, στα πλαίσια της TN, αποτελεί η *Αναπαράσταση Γνώσης και η Συλλογιστική* (*Knowledge Representation & Reasoning*) μέσω της λογικής. Μια κύρια γλώσσα αναπαράστασης γνώσης και συλλογιστικής είναι η Κατηγορηματική Λογική Πρώτης Τάξης (ΚΛΠΤ), που διδάσκεται σχεδόν σε κάθε μάθημα TN (Barwise & Etchemendy, 2002). Η διδασκαλία και η εκμάθηση της ΚΛΠΤ ως γλώσσα αναπαράστασης γνώσης και συλλογιστικής περιλαμβάνει πολλές πτυχές.

Σημαντικές πτυχές της λογικής, για τη χρήση της ως γλώσσας αναπαράστασης και συλλογιστικής, είναι η διαδικασία μετάφρασης της φυσικής γλώσσας σε ΚΛΠΤ, μια διαδικασία που ονομάζεται *φορμαλισμός* φυσικής γλώσσας, καθώς και η διαδικασία μετατροπής ΚΛΠΤ σε Προτασιακή Μορφή (ΠΜ), που χρησιμοποιείται σε διαδικασίες όπως η συλλογιστική. Επίσης, η έννοια της ισοδυναμίας είναι ένα βασικό θέμα στη λογική. Δύο ή περισσότερες εκφράσεις ΠΜ ή ΚΛΠΤ χαρακτηρίζονται ως λογικά ισοδύναμες, αν έχουν το ίδιο λογικό περιεχόμενο. Η διαχείριση εκφράσεων ΚΛΠΤ και ο προσδιορισμός των ισοδύναμων εκφράσεων είναι μια πολύπλοκη και σύνθετη διαδικασία η οποία δημιουργεί ιδιαίτερες δυσκολίες στους εκπαιδευόμενους να την κατανοήσουν και να την εφαρμόσουν σωστά.

Για να βοηθήσουμε τους εκπαιδευμένους να κατανοήσουν τις διαδικασίες της λογικής σχεδιάσαμε και αναπτύξαμε δυο ευφυή συστήματα διδασκαλίας, το σύστημα *FOLtoCF* (Grivokostopoulou et al., 2014b) και το σύστημα *FOL-EQU* (Grivokostopoulou et al., 2013). Πιο συγκεκριμένα, το *FOLtoCF* είναι ένα ευφύες σύστημα για την εκμάθηση της διαδικασίας μετατροπής εκφράσεων ΚΛΠΤ σε ΠΜ. Επίσης, το σύστημα μέσω των *εργαλείων συγγραφής* (*authoring tools*) που διαθέτει, έχει ως στόχο να βοηθήσει τον διδάσκοντα στην διαχείριση του εκπαιδευτικού υλικού του συστήματος. Το *FOL-EQU* είναι ένα ευφύες σύστημα, που έχει ως στόχο να βοηθήσει τους εκπαιδευόμενους να κατανοήσουν τις λογικές ισοδυναμίες

και να μάθουν να εφαρμόζουν τους κανόνες ισοδυναμίας και να δημιουργούν λογικά ισοδύναμες εκφράσεις.

2.5.1 Βασικές Έννοιες

Η λογική παρέχει έναν τρόπο για την αποσαφήνιση και την τυποποίηση της διαδικασίας της ανθρώπινης σκέψης και προσφέρει μια σημαντική και εύχρηστη μεθοδολογία για την αναπαράσταση και επίλυση των προβλημάτων. Η ανάγκη χρήσης μιας αυστηρά ορισμένης γλώσσας, με τη μαθηματική έννοια, προήλθε από την ακαταλληλότητα της χρήσης της φυσικής γλώσσας σε υπολογιστικά συστήματα. Σε αντίθεση με τη φυσική γλώσσα, η λογική προσφέρει μια σαφή, ακριβή και απλή στη σύνταξη γλώσσα, καθώς και τη δυνατότητα παραγωγής νέας γνώσης από την ήδη υπάρχουσα. Η λογική ορίζεται σαν τη μελέτη της σωστής εξαγωγής συμπερασμάτων. Μια κατ' ελάχιστον απαίτηση, για σωστή εξαγωγή συμπερασμάτων, είναι η διατήρηση της αλήθειας, δηλαδή η απαίτηση από αληθείς υποθέσεις - προτάσεις να εξάγονται αληθή συμπεράσματα - προτάσεις.

Υπάρχουν διάφορα είδη λογικής και δύο βασικά είδη είναι η *Προτασιακή Λογική* (*Propositional Logic*) και η *Κατηγορηματική Λογική* (*Predicate Logic*). Η Προτασιακή Λογική αποτελεί την απλούστερη μορφή λογικής και αποτελεί την βάση για περισσότερο πολύπλοκες λογικές προτάσεις. Μια λογική πρόταση χαρακτηρίζεται είτε ως *αληθής* (*true*) είτε ως *ψευδής* (*false*), μπορεί δηλαδή να έχει δύο λογικές τιμές. Η προτασιακή λογική περιλαμβάνει τα εξής στοιχεία: (α) Ένα σύνολο από προτασιακά σύμβολα (P, Q, R , κλπ), τα οποία εκφράζουν απλά γεγονότα και ονομάζονται *άτομα* (*atoms*) ή *ατομικοί τύποι* (*atomic formulas*), που μπορούν να συνδέονται μεταξύ τους με τη χρήση *διασυνδετικών* (*connectives*), (β) Τα *λογικά διασυνδετικά*, που είναι: η άρνηση “ $\neg$ ”, η διάζευξη “ $\vee$ ” (λογικό «ή»), η σύζευξη “ $\wedge$ ” (λογικό «και»), η συνεπαγωγή “ $\Rightarrow$ ” και η διπλή συνεπαγωγή “ $\Leftrightarrow$ ” (λογική ισοδυναμία).

Η *Κατηγορηματική Λογική Πρώτης Τάξης-ΚΛΠΤ* (*First Order Predicate Logic-FOL*) αποτελεί επέκταση της προτασιακή λογικής και αντιμετωπίζει το πρόβλημα της μη προσπελασιμότητας των στοιχείων των γεγονότων της προτασιακή λογικής. Επίσης, μια σημαντική επέκταση σε σχέση με την προτασιακή λογικής είναι η ύπαρξη μεταβλητών. Επομένως, η ΚΛΠΤ επεκτείνει την προτασιακή λογική εισάγοντας τις έννοιες: *όροι* (*terms*), *κατηγορήματα* (*predicates*) και *ποσοδείκτες* (*quantifiers*).

Για τον καθορισμό των διαφόρων στοιχείων της ΚΛΠΤ χρειάζεται να διευκρινιστεί η παρακάτω έννοια:

Πεδίο ορισμού D ονομάζεται το σύνολο όλων των οντοτήτων που σχετίζονται με τη γνώση που θέλουμε να αναπαραστήσουμε.

Πιο συγκεκριμένα, η ΚΛΠΤ περιλαμβάνει τα ακόλουθα στοιχεία:

- Ένα σύνολο από σταθερές (*constants*), όπου μια σταθερά χρησιμοποιείται για να δηλώσει ένα συγκεκριμένο στοιχείο του πεδίου ορισμού D .
- Λογικές σταθερές *true* (T) και *false* (F).
- Ένα σύνολο από μεταβλητές (*variables*), όπου μια μεταβλητή χρησιμοποιείται για να δηλώσει ένα υποσύνολο στοιχείων του πεδίου ορισμού, χωρίς να τα καθορίζει.
- Ένα σύνολο από συναρτησιακά σύμβολα (*function symbols*), όπου ένα συναρτησιακό σύμβολο n τάξης παριστάνει μια αμφιμονοσήμαντη απεικόνιση από το D^n στο D .
- Ένα σύνολο από κατηγορήματα (*predicates*). Ένα κατηγορημα n ορισμάτων παριστάνει μια αντιστοίχιση από το D^n στο $\{T, F\}$ και εκφράζει ένα συσχετισμό μεταξύ των n στοιχείων του D . Αν τα n στοιχεία συσχετίζονται μεταξύ τους με τον τρόπο που δηλώνει το κατηγορημα, τότε αυτό παίρνει την τιμή *true*, διαφορετικά παίρνει την τιμή *false*.
- Τα λογικά διασυνδετικά της ΚΛΠΤ είναι ίδια με εκείνα της προτασιακής λογικής και είναι: η άρνηση “ $\neg$ ”, η διάζευξη “ $\vee$ ” (λογικό «ή»), η σύζευξη “ $\wedge$ ” (λογικό «και»), η συνεπαγωγή “ $\Rightarrow$ ” και η διπλή συνεπαγωγή “ $\Leftrightarrow$ ” (λογική ισοδυναμία).
- Δύο ποσοδείκτες: ο καθολικός ποσοδείκτης “ $\forall$ ” (*universal quantifier*) και ο υπαρξιακός ποσοδείκτης “ $\exists$ ” (*existential quantifier*).

Μια ατομική έκφραση ή ένα άτομο στην ΚΛΠΤ είναι μια έκφραση της μορφής $P(t_1, \dots, t_n)$, όπου P ένα κατηγορημα n ορισμάτων και $t_1, \dots, t_n$ είναι όροι. Ένας όρος (*term*) μπορεί να είναι μια σταθερά ή μια μεταβλητή ή μια συνάρτηση. Αν f είναι ένα συναρτησιακό σύμβολο n τάξης και $t_1, \dots, t_n$ είναι όροι, τότε $f(t_1, \dots, t_n)$ είναι όρος.

Μια ορθά δομημένη έκφραση (*well formed formula-wff*) στην ΚΛΠΤ ορίζεται ως εξής:

- Ένα άτομο είναι μια *wff*.
- Αν P είναι μια *wff*, τότε $\neg P$ είναι μια *wff*.
- Αν P και Q είναι *wffs*, τότε $(P \vee Q)$, $(P \wedge Q)$, $(P \Rightarrow Q)$ και $(P \Leftrightarrow Q)$ είναι *wffs*.
- Αν P είναι μια *wff* και x μια ελεύθερη μεταβλητή στην P , τότε $(\forall x)P$, $(\exists x)P$ είναι *wff*.

- v. Το σύνολό των *wffs* δημιουργείται μόνο από πεπερασμένο αριθμό εφαρμογών των (i), (ii), (iii).

Κάθε ποσοδείκτης έχει μια *εμβέλεια* (scope), η οποία είναι η έκφραση στην οποία εφαρμόζεται ο ποσοδείκτης.

Μια μεταβλητή λέγεται *δεσμευμένη* (bound) σε μια έκφραση, αν και μόνο αν μια εμφάνιση της μέσα στην έκφραση βρίσκεται μέσα στην εμβέλεια ενός ποσοδείκτη που την προσδιορίζει. Μια μεταβλητή που δεν είναι δεσμευμένη λέγεται *ελεύθερη* (free). Για παράδειγμα, στην έκφραση $(\forall x) Q(x,y)$ η μεταβλητή x είναι δεσμευμένη, ενώ η y είναι ελεύθερη.

Μια ορθά δομημένη έκφραση είναι σε *Συζευκτική Κανονική Μορφή-ΣΚΜ* (Conjunctive Normal Form-CNF) όταν αποτελείται σύζευξη διαζεύξεων, π.χ. $(A \vee B) \wedge (\neg A \vee G)$, ενώ είναι σε *Διαζευκτική Κανονική Μορφή-ΔΚΜ* (Disjunctive Normal Form-DNF) όταν αποτελείται από διάζευξη συζεύξεων, π.χ. $(A \wedge B) \vee (A \wedge G)$.

Μια έκφραση είναι σε *Κανονική Μορφή Prenex-KMP* (Prenex Normal Form-PNF) αν και μόνο αν έχει τη μορφή: $((Q_1x_1)(Q_2x_2) \dots (Q_nx_n)P) \dots$ όπου $Q_i, i = 1, \dots, n$ είναι καθολικός ($\forall$) ή υπαρξιακός ($\exists$) ποσοδείκτης και η P είναι μια έκφραση που δεν περιέχει ποσοδείκτες. Αποδεικνύεται ότι για κάθε έκφραση ΚΛΠΤ υπάρχει μια ισοδύναμη έκφραση σε *KMP* και με τον όρο «ισοδύναμη» εννοούμε ότι οι δυο εκφράσεις παίρνουν τις ίδιες λογικές τιμές (Τ ή F) κάτω από οποιαδήποτε ερμηνεία.

Η *Προτασιακή Μορφή-ΠΜ* (Clausal Form-CF) είναι μια απλοποιημένη εκδοχή της ΚΛΠΤ, η οποία αποτελείται από «απλοποιημένες» προτάσεις. Κάθε πρόταση αποτελείται από 'στοιχεία'. Ένα *στοιχείο* (literal) είναι είτε μια ατομική έκφραση (θετικό στοιχείο) είτε η άρνηση μιας ατομικής έκφρασης (αρνητικό στοιχείο). Μια *πρόταση* (clause) είναι ένα σύνολο από στοιχεία, το οποίο αναπαριστά τη διάζευξή τους, πρόκειται δηλαδή για μια διάζευξη στοιχείων. Υπάρχουν αρκετά διαφορετικά είδη προτάσεων, βάσει του πλήθους και του είδους των στοιχείων τους. Μια *κενή πρόταση* (empty clause) είναι μια πρόταση που δεν περιέχει κανένα στοιχείο. Μια πρόταση που περιέχει ένα ακριβώς ένα στοιχείο ονομάζεται *μοναδιαία* (unit clause), διαφορετικά ονομάζεται *μη μοναδιαία*. Γενικότερα, μια πρόταση που περιέχει n στοιχεία, ονομάζεται *n-πρόταση*. Όταν όλα τα στοιχεία μιας πρότασης είναι θετικά, ονομάζεται *θετική*, ενώ αν όλα τα στοιχεία της είναι αρνητικά ονομάζεται *αρνητική*. Μια πρόταση που περιέχει τόσο θετικά όσο και αρνητικά στοιχεία ονομάζεται *μεικτή*, ενώ μια πρόταση που περιέχει το πολύ ένα θετικό στοιχείο λέγεται *τύπου Horn*. Ένα ζευγάρι από

στοιχεία λέγεται *συμπληρωματικό* αν το ένα στοιχείο είναι θετικό, το άλλο αρνητικό και οι ατομικές τους εκφράσεις είναι ίδιες. Επίσης, μια έκφραση βρίσκεται σε ΠΜ αν μπορεί να γραφεί σαν σύζευξη από προτάσεις. Επειδή η σύζευξη και η διάζευξη είναι πράξεις προσεταιριστικές, αντιμεταθετικές και ανακλαστικές, μια έκφραση θεωρείται συχνά σαν ένα σύνολο από προτάσεις.

2.5.2 Ευφύες Σύστημα Μετατροπής Κατηγορηματικής Λογικής σε Προτασιακή Μορφή

Η διαδικασία μετατροπής εκφράσεων ΚΛΠΤ σε ΠΜ αποτελεί μια διαδικασία η οποία έχει μοντελοποιηθεί σε βήματα. Ωστόσο, στην διεθνή βιβλιογραφία δεν υπάρχουν συστήματα που να χρησιμοποιούν τη μοντελοποίηση αυτή για τη διδασκαλία της διαδικασίας. Τα υπάρχοντα συστήματα μετατρέπουν αυτόματα μια έκφραση ΚΛΠΤ σε ΠΜ χωρίς να παρέχεται η δυνατότητα στον χρήστη να μελετήσει και να εφαρμόσει τα βήματα της μετατροπής. Για τον σκοπό αυτό, σχεδιάσαμε και αναπτύξαμε το Ευφύες Εκπαιδευτικό Σύστημα FOLtoCF (Grivokostopoulou et al., 2014; Grivokostopoulou et al., 2012c), για την εκμάθηση της διαδικασίας μετατροπής ΚΛΠΤ σε ΠΜ το οποίο μοντελοποιεί τα βήματα της διαδικασίας και έχει ως στόχο να βοηθήσει τους εκπαιδευόμενους στην κατανόηση της διαδικασίας και την παροχή εξατομικευμένης μάθησης. Επίσης, το σύστημα αποτελεί ένα χρήσιμο εργαλείο για τον διδάσκοντα, καθώς του προσφέρει την δυνατότητα να δημιουργεί νέο εκπαιδευτικό υλικό και να παρακολουθεί τις δραστηριότητες και την πρόοδο των εκπαιδευόμενων μέσω των εργαλείων συγγραφής (*authoring tools*) που διαθέτει. Τα βήματα της διαδικασίας της μετατροπής ΚΛΠΤ σε ΠΜ που μοντελοποιήθηκαν από το σύστημα είναι τα εξής:

Βήμα 1ο: Απαλοιφή των διασυνδετικών “ $\Rightarrow$ ” και “ $\Leftrightarrow$ ”, χρησιμοποιώντας τις ισοδυναμίες:

$$P \Rightarrow Q \equiv \neg P \vee Q$$

$$P \Rightarrow Q \equiv (P \Rightarrow Q) \wedge (Q \Rightarrow P)$$

Βήμα 2ο: Περιορισμός της εμβέλειας των αρνήσεων, χρησιμοποιώντας τις ακόλουθες ισοδυναμίες:

$$\neg \neg P \equiv P$$

$$\neg(P \vee Q) \equiv \neg P \wedge \neg Q$$

$$\neg(P \wedge Q) \equiv \neg P \vee \neg Q$$

$$\neg(\forall x) P \equiv (\exists x) \neg P$$

$$(\neg \exists x) P \equiv (\forall x) \neg P$$

Βήμα 3ο: Μετονομασία μεταβλητών που δεσμεύονται από διαφορετικούς ποσοδείκτες, έτσι ώστε κάθε μεταβλητή να δεσμεύεται από έναν μόνο ποσοδείκτη και στην συνέχεια μετατροπή της έκφρασης σε *Κανονική Μορφή Prenex (KMP)*.

Βήμα 4ο: Απαλοιφή υπαρξιακών ποσοδεικτών (*Skolemisation*) και καθολικών ποσοδεικτών:

Θεωρούμε την ακόλουθη έκφραση σε κανονική μορφή Prenex

$$((Q_1x_1) (Q_2x_2) \dots ((Q_nx_n) F) \dots).$$

Η διαδικασία είναι η ακόλουθη:

- Αν αριστερά από ένα υπαρξιακό ποσοδείκτη Q_r , $1 \leq r \leq n$, εμφανίζονται καθολικοί ποσοδείκτες, έστω $Q_{s_1}, \dots, Q_{s_m}$, $1 \leq s_1 < s_2 \dots < s_m < r$, των οποίων η εμβέλεια εμπεριέχεται στην εμβέλεια του υπαρξιακού ποσοδείκτη, τότε αντικαθιστούμε κάθε εμφάνιση της μεταβλητής x_r στην F με μια *συνάρτηση Skolem* (sk_{-f_r}), της οποίας τα ορίσματα είναι οι μεταβλητές αυτών των καθολικών ποσοδεικτών, και διαγράφουμε το $(Q_r x_r)$
- Αν Q_r , $1 \leq r \leq n$ είναι ένας υπαρξιακός ποσοδείκτης και πριν από αυτόν δεν εμφανίζεται κανένας καθολικός ποσοδείκτης, τότε αντικαθιστούμε κάθε εμφάνιση της μεταβλητής με μια *σταθερά skolem* (sk_{-c_r}) και διαγράφουμε το $(Q_r x_r)$
- Μετά την απαλοιφή των υπαρξιακών ποσοδεικτών (*Skolemisation*) απαλείφουμε όλους τους καθολικούς ποσοδείκτες.
- Αν δεν εμφανίζεται κανένας υπαρξιακός ποσοδείκτης τότε απαλείφουμε τους καθολικούς ποσοδείκτης.

Βήμα 5ο: Μετατροπή της έκφρασης σε ΣΚΜ (CNF) χρησιμοποιώντας την ισοδυναμία:

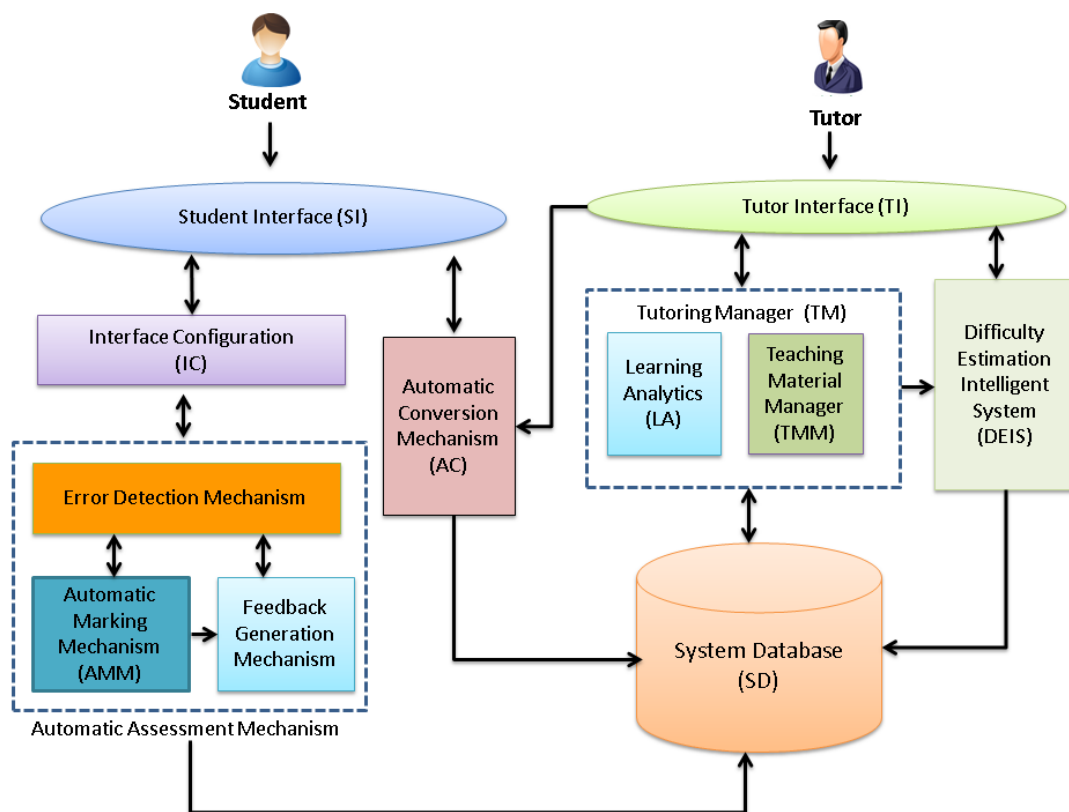
$$(P \wedge Q) \vee R \equiv (P \vee R) \wedge (Q \vee R)$$

Βήμα 6ο: Απαλοιφή των διασυνδετικών και προσδιορισμός του συνόλου των προτάσεων που προέκυψαν καθώς και μετονομασία των μεταβλητών έτσι ώστε μια μεταβλητή να μην εμφανίζεται σε περισσότερες από μια προτάσεις.

2.5.2.1 Αρχιτεκτονική Συστήματος FOLtoCF

Η αρχιτεκτονική του συστήματος FOLtoCF παρουσιάζεται στην εικόνα 2.14 και αποτελείται από μονάδες με τις οποίες αλληλεπιδρά ο εκπαιδευόμενος και μονάδες που αλληλεπιδρά ο διδάσκοντας. Πιο συγκεκριμένα, οι βασικές μονάδες με τις οποίες αλληλεπιδρά ο εκπαιδευόμενος είναι: η *Διεπαφή Εκπαιδευόμενου (Student Interface-SI)*, η *Διαμόρφωση Διεπαφής (Interface Configuration -IC)*, ο *Μηχανισμός Αυτόματης Μετατροπής (Automatic Conversion Mechanism -AC)*, ο *Μηχανισμός Αυτόματης Αξιολόγησης (Automatic Assessment*

Mechanism-AA). Αντίστοιχα οι μονάδες που υποστηρίζουν τις ανάγκες του διδάσκοντα είναι: η *Διεπαφή Διδάσκοντα (Tutor Interface-TI)*, ο *Διαχειριστής Διδασκαλίας (Tutoring Manager)*, το *Ευφρές Σύστημα Προσδιορισμού Δυσκολίας (Difficulty Estimation Intelligent System - DEIS)*, καθώς και η *Βάση Συστήματος (System Database-SD)* που αποθηκεύονται όλες πληροφορίες του συστήματος.



Εικόνα 2.14 Η Αρχιτεκτονική του Συστήματος FOL to CF

Η *Διεπαφή Εκπαιδευόμενου (SI)* έχει σχεδιαστεί και αναπτυχθεί ώστε να καθιστά τη χρήση του συστήματος εύκολη και ελκυστική για τον εκπαιδευόμενο. Αρχικά, ο χρήστης εγγράφεται στο σύστημα καταχωρώντας τις απαιτούμενες προσωπικές πληροφορίες, όπως το όνομα, το έτος σπουδών και το φύλο. Μετά την εγγραφή, ο εκπαιδευόμενος μπορεί ανά πάσα στιγμή να εισέλθει και να χρησιμοποιήσει το σύστημα. Η μονάδα *Διαμόρφωση Διεπαφής (IC)* ρυθμίζει την διεπαφή του εκπαιδευόμενου σύμφωνα με τις ανάγκες και τα χαρακτηριστικά της τρέχουσας μαθησιακής δραστηριότητας, όπως για παράδειγμα το τρέχον βήμα της διαδικασίας μετατροπής, καθώς και της παρεχόμενης βοήθειας. Επομένως, η διεπαφή του εκπαιδευόμενου είναι *δυναμική* και αναπροσαρμόζεται στις ανάγκες της τρέχουσας δραστηριότητας. Ο *Μηχανισμός Αυτόματης Μετατροπής (AC)* μετατρέπει αυτόματα μια έκφραση ΚΛΠΤ σε ΠΙΜ και καθορίζει τα βήματα της διαδικασίας που θα εφαρμοστούν

καθώς και την αντίστοιχη έξοδο του κάθε βήματος. Επίσης, ο μηχανισμός αυτός αποτελεί μέρος του *Αυτόματου Αποδείκτη Θεωρημάτων* (Grivokostopoulou & Hatzilygeroudis, 2010; Hatzilygeroudis et al., 1994) για την αυτόματη μετατροπή εκφράσεων ΚΛΠΤ σε ΠΜ, διεργασία που είναι απαραίτητη για την μετέπειτα εξαγωγή συμπερασμάτων. Ο *Μηχανισμός Αυτόματης Αξιολόγησης* (AA) αποτελείται από δυο επιμέρους μηχανισμούς, τον *Μηχανισμό Αναγνώρισης Λαθών* (*Error Detection mechanism-ED*) και τον *Μηχανισμό Παραγωγής Ανάδρασης* (*Feedback Generation Mechanism- FG*). Αρχικά, ο μηχανισμός αναγνώρισης λαθών αναγνωρίζει το είδος των λαθών που έγιναν και ο μηχανισμός παραγωγής ανάδρασης παρέχει εξατομικευμένη ανάδραση με βάση τις ανάγκες του εκπαιδευόμενου, με στόχο να τον βοηθήσει να ολοκληρώσει κάθε βήμα της διαδικασίας. Στην συνέχεια, μέσω της μονάδας αυτόματη βαθμολόγησης, αξιολογούνται αυτόματα οι απαντήσεις του εκπαιδευόμενου πάνω σε ασκήσεις ή τεστ και προσδιορίζεται αυτόματα η βαθμολογία του.

Ο *Διαχειριστής Διδασκαλίας* (TM) αποτελείται από δυο επιμέρους μονάδες, τον *Διαχειριστή Εκπαιδευτικού Υλικού* (*Teaching Material Manager- TM*) και την μονάδα Ανάλυσης Στοιχείων Μάθησης (*Learning Analytics-LA*). Η μονάδα της διαχείρισης εκπαιδευτικού υλικού αναπτύχθηκε για να βοηθήσει το διδάσκοντα στην διαχείριση του εκπαιδευτικού υλικού του συστήματος, όπως για παράδειγμα να επεξεργαστεί τις ασκήσεις και τα παραδείγματα που είναι διαθέσιμα στο σύστημα. Η μονάδα ανάλυσης στοιχείων μάθησης χρησιμοποιείται για να αναλύει τις ενέργειες των εκπαιδευόμενων και να παρέχει πληροφορίες στον διδάσκοντα σχετικά με τις επιδόσεις τους. Επίσης, παρέχει γενικές πληροφορίες σχετικά με τη διαδικασία της μάθησης του κάθε εκπαιδευόμενου, και παρέχει στατιστικές πληροφορίες σχετικά με συνηθισμένα λάθη που κάνουν οι εκπαιδευόμενοι, με το επίπεδο γνώσεων τους, με το χρόνο επίλυσης των ασκήσεων κ.ά. Επιπρόσθετα, διαθέτει ένα μηχανισμό για την οπτικοποίηση των αντίστοιχων αποτελεσμάτων, με στόχο να βοηθήσει τον διδάσκοντα να προσδιορίσει και να κατανοήσει καλύτερα τις δυσκολίες που αντιμετωπίζουν οι εκπαιδευόμενοι. Το *Ευφρές Σύστημα Προσδιορισμού Δυσκολίας* (DIES) έχει ως στόχο να αναλύει και να προσδιορίζει αυτόματα, με ακρίβεια και με συνέπεια το επίπεδο δυσκολίας των ασκήσεων χωρίς την παρέμβαση του διδάσκοντα και να συμβάλει στην καλύτερη προσαρμογή των εκπαιδευτικών διαδικασιών στους εκπαιδευόμενους. Τέλος, στην *βάση δεδομένων* (SD) αποθηκεύονται όλα τα απαραίτητα δεδομένα σχετικά με τον εκπαιδευόμενο και καταγράφονται πληροφορίες κατά την διάρκεια της αλληλεπίδρασης. Επίσης, αποθηκεύεται το εκπαιδευτικό υλικό.

2.5.2.2 Αλληλεπίδραση Χρήστη με το Σύστημα

Αρχικά, το σύστημα παρέχει την δυνατότητα στον χρήστη να εγγραφεί στο σύστημα εισάγοντας προσωπικές πληροφορίες όπως έτος, σπουδών, φύλο κ.α.. Στην συνέχεια, ο χρήστης μπορεί να εισέλθει στο σύστημα και να αλληλεπιδράσει με αυτό και ανάλογα με το είδος του χρήστη φοιτητή ή διδάσκοντα και να έχει πρόσβαση σε διαφορετικές πληροφορίες. Στην περίπτωση ενός χρήστη-εκπαιδευόμενου, αφού εισέλθει στο σύστημα μπορεί να δει το εκπαιδευτικό υλικό, το σύνολο όλων διαθέσιμων ασκήσεων (εκφράσεων ΚΛΠΤ) και το αντίστοιχο επίπεδο δυσκολίας τους και στην συνέχεια έχει την δυνατότητα να επιλέξει μια άσκηση για επίλυση ή να δημιουργήσει μια νέα έκφραση. Στην περίπτωση που επιλέξει να δημιουργήσει μια νέα άσκηση, τότε εισάγει την ΚΛΠΤ έκφραση που επιθυμεί και στην συνέχεια η έκφραση αναλύεται μέσω της μονάδας Formula Analyser (Grivokostopoulou et al., 2014b), προσδιορίζεται το επίπεδο δυσκολίας της (Grivokostopoulou et al., 2014a, Grivokostopoulou et al., 2012b) και η νέα πρόταση είναι πλέον διαθέσιμη. Έπειτα, ο εκπαιδευόμενος επιλέγοντας μια άσκηση ξεκινάει η διαδικασία μετατροπής από ΚΛΠΤ σε ΠΜ και καλείται να εφαρμόσει βήμα-βήμα τη διαδικασία της μετατροπής και να προσδιορίσει την έκφραση που αντιστοιχεί σε κάθε βήμα ή σε περίπτωση που δεν εφαρμόζεται το τρέχον βήμα να συνεχίσει στο επόμενο βήμα. Επίσης, το σύστημα μέσω του μηχανισμού ανάδρασης που διαθέτει παρέχει σε κάθε βήμα της διαδικασίας διάφορα είδη βοήθειας όπως για παράδειγμα ενημερώνει τον εκπαιδευόμενο για λάθη του στο τρέχον βήμα της διαδικασίας και επίσης παρέχει αναλυτικές πληροφορίες σχετικά αυτά.

2.5.2.3 Μηχανισμός Αναγνώρισης Λαθών

Η διαδικασία αναγνώρισης λαθών είναι μια θεμελιώδης διεργασία στα Ευφυή Συστήματα Διδασκαλίας-ΕΣΔ (Intelligent Tutoring Systems-ITS). Η αναγνώριση λαθών αποτελεί ένα σημαντικό κίνητρο για τους εκπαιδευόμενους, και συμβάλει ώστε να γίνει η κατανόηση και η μάθηση πιο αποτελεσματική (Hirashima, 1998). Ο *Μηχανισμός Αναγνώρισης Λαθών* (Error Detection Mechanism- ED) του συστήματος εντοπίζει και αναγνωρίζει αυτόματα τα λάθη που έγιναν κατά τη διάρκεια των βημάτων της διαδικασίας μετατροπής. Πιο συγκεκριμένα, όταν ο εκπαιδευόμενος υποβάλλει μια απάντηση για ένα βήμα της διαδικασίας, ο μηχανισμός εντοπίζει και αναγνωρίζει το πλήθος των λαθών που έγιναν και στην συνέχεια προσδιορίζει τους τύπους των λαθών με βάση ένα σχήμα κατηγοριοποίησης λαθών. Κατά την διάρκεια της αλληλεπίδρασης του εκπαιδευόμενου με το σύστημα FOLtoCF, ο μηχανισμός λαθών διαχειρίζεται την απάντηση του για το κάθε βήμα ως εξής:

1. Ελέγχει αν το τρέχον βήμα της διαδικασίας είναι εφαρμόσιμο:

- 1.1 Αν είναι εφαρμόσιμο και ο εκπαιδευόμενος προσδιόρισε σωστά ότι εφαρμόζεται, τότε καλείται να εισάγει την απάντηση για το τρέχον βήμα της ΚΛΠΤ σε ΠΜ διαδικασίας και στη συνέχεια πηγαίνει στο 2.
- 1.2 Αν δεν είναι εφαρμόσιμο και ο εκπαιδευόμενος προσδιόρισε σωστά ότι δεν εφαρμόζεται το βήμα, τότε καλείται να προσδιορίσει το επόμενο βήμα της ΚΛΠΤ σε ΠΜ διαδικασίας και πηγαίνει στο 1.
- 1.3 Αν απάντηση του εκπαιδευόμενου στον προσδιορισμό της εφαρμογής του τρέχοντος βήματος είναι λανθασμένη, τότε ενημερώνεται ότι η απάντηση του δεν είναι σωστή και μπορεί να δει τη σωστή απάντηση μαζί με κατάλληλες επεξηγήσεις.
2. Ενεργοποιείται ο *Αναλυτής Εκφράσεων (Formula Analyzer)* για να αναλύσει την απάντηση του και να δημιουργήσει την αντίστοιχη δεντρική δομή.
3. Συγκρίνεται η απάντηση του εκπαιδευόμενου για το τρέχον βήμα με τη σωστή απάντηση:
 - 3.1 Αν όλα τα μέρη της απάντησης είναι σωστά, τότε ενημερώνεται ο εκπαιδευόμενος και πηγαίνει στο 1
 - 3.2 Διαφορετικά, ενεργοποιείται ο μηχανισμός αναγνώρισης λαθών.
4. Προσδιορίζονται οι τύποι των λαθών που έγιναν και ενημερώνεται ο εκπαιδευόμενος για τα αντίστοιχα λάθη.
5. Ο εκπαιδευόμενος μπορεί να δώσει μια νέα απάντηση για το τρέχον βήμα και πηγαίνει στο 2

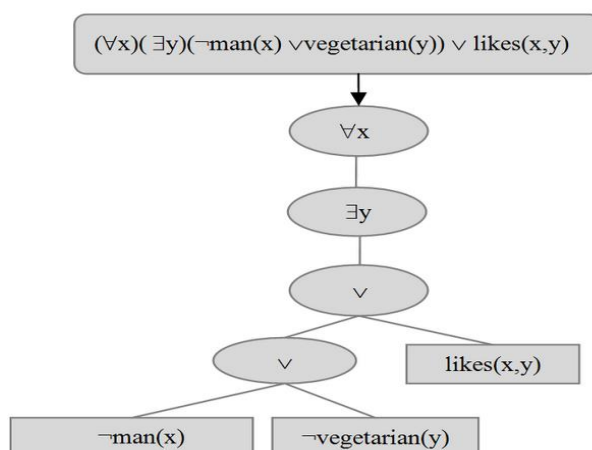
Ο *Αναλυτής Εκφράσεων (Formula Analyzer)* αναλύει τις απαντήσεις των εκπαιδευόμενων σε κάθε βήμα και για κάθε βήμα πραγματοποιεί συντακτική ανάλυση του τρέχοντος τύπου της έκφρασης και δημιουργείται η αντίστοιχη δεντρική δομή. Ο μηχανισμός λειτουργεί σε δύο στάδια. Αρχικά, δημιουργείται η δεντρική δομή για την τρέχουσα έκφραση και αναλύεται η έκφραση στα δομικά της μέρη, τους ποσοδείκτες (εάν υπάρχουν), τα διασυνδετικά και τις ατομικές εκφράσεις. Πιο συγκεκριμένα, οι ατομικές εκφράσεις αναπαριστούν τα φύλλα του δέντρου ενώ οι εσωτερικοί κόμβοι του δέντρου είναι τα διασυνδετικά. Επίσης, κάθε ποσοδείκτης ανάλογα με την εμβέλειά του, συνδέεται με τους κατάλληλους κόμβους του δέντρου. Στο δεύτερο στάδιο της διαδικασίας, γίνεται η ανάλυση των ατομικών εκφράσεων και των ποσοδεικτών (αν υπάρχουν) για την τρέχουσα πρόταση.

Επομένως, ο μηχανισμός για κάθε ατομική έκφραση προσδιορίζει τον αριθμό και τον τύπο των όρων της έκφρασης. Επίσης, για κάθε ποσοδείκτη (αν υπάρχει) καθορίζεται ο τύπος και η μεταβλητή που συνδέεται με αυτό.

Για παράδειγμα, θεωρούμε την ακόλουθη απάντηση του εκπαιδευόμενου για το δεύτερο βήμα της διαδικασίας το οποίο αφορά τον περιορισμό της εμβέλειας της άρνησης:

$$“(\forall x)(\exists y)(\neg \text{man}(x) \vee \neg \text{vegetarian}(y)) \vee \text{likes}(x, y)”.$$

Το σύστημα αρχικά αναγνωρίζει ότι η έκφραση είναι πρόταση ΚΛΠΤ και στην συνέχεια δημιουργεί την αντίστοιχη δεντρική μορφή της πρότασης και γίνεται η ανάλυση των ατομικών έκφρασεων, όπως παρουσιάζονται στις εικόνες 2.15 και 2.16 αντίστοιχα. Στην συνέχεια, γίνεται σύγκριση της δενδρικής δομής της απάντησης του εκπαιδευόμενου και της σωστής απάντησης ώστε να αναγνωριστούν πιθανά λάθη και προσδιορίζονται ο τύπος/(οι) των λαθών που έγιναν.



Εικόνα 2.15 Δεντρική Αναπαράσταση μιας έκφρασης ΚΛΠΤ

Analysing Atomic Expressions:	
Number of Atomic Expression:3	
man	predicate
vegetarian	predicate
likes	predicate
man(x)	->man arguments: 1 variable x
vegetarian(y)	->vegetarian arguments: 1 variable y
likes(x,y)	->likes arguments: 2 variable x variable y

Εικόνα 2.16 Συντακτική Ανάλυση της έκφρασης ΚΛΠΤ

Για την κατηγοριοποίηση των λαθών βασιστήκαμε στα βήματα της διαδικασίας FOLtoCF καθώς και σε πειραματικά αποτελέσματα. Οι βασικές κατηγορίες των λαθών σε κάθε βήμα της διαδικασίας είναι οι ακόλουθες:

- Quantifier Error:
 - (a) Universal Error: Ο καθολικός ποσοδείκτης που δεσμεύεται από μια συγκεκριμένη μεταβλητή δεν έχει αφαιρεθεί σωστά.
 - (b) Skolemization error: Ο καθολικός ποσοδείκτης που δεσμεύεται από μια συγκεκριμένη μεταβλητή δεν έχει αφαιρεθεί σωστά.
 - (c) Order Error: Η θέση του ποσοδείκτη δεν είναι σωστή.
- Argument Error:
 - (a) Type Error: Ο τύπος των ορισμάτων δεν προσδιορίστηκε σωστά (π.χ. μια μεταβλητή αντί μιας συνάρτησης, όπως “eats(y, x)” αντί “eats(sk-f(x), x)”
 - (b) Cardinality Error: Ο αριθμός των ορισμάτων της ατομικής έκφρασης δεν προσδιορίστηκε σωστά. Π.χ. γράφεται “loves(x, y, sk-f(x))” αντί “loves(x, sk-f(x))”.
- Connective Error:
 - (a) Scope error: Η εμβέλεια της άρνησης δεν είναι σωστή.
 - (b) Connective Error: Ένα διασυνδετικό (πχ. “^”) χρησιμοποιείται στην θέση κάποιου άλλου διασυνδετικού (πχ. “v”)
 - (c) Application Error: Η εφαρμογή της ισοδυναμίας δεν είναι σωστή (π. χ. $\neg(P \vee Q) \rightarrow \neg P \vee \neg Q$ αντί $\neg P \wedge \neg Q$).
 - (d) Structural Error: Η έκφραση δεν αποτελεί μια έκφραση σε ΠΜ

Για παράδειγμα, θεωρούμε την ακόλουθη ΚΛΠΤ έκφραση:

$$“(\forall x) (\exists y) (human(x) \wedge food(y) \Rightarrow eats(x, y))”$$

και για το βήμα 4 της διαδικασίας η απάντηση του εκπαιδευόμενου είναι:

$human(x) \wedge food(sk-f(y)) \wedge eats(x, y)$

ενώ η αντίστοιχη σωστή απάντηση είναι:

$\neg human(x) \vee \neg food(sk-f(y)) \vee eats(x, sk-f(y)).$

Αρχικά, ο Αναλυτής Έκφράσεων (Formula Analyzer) εντοπίζει και αναγνωρίζει δυο τύπους των λαθών που έγιναν από τον εκπαιδευόμενο. Το πρώτο λάθος που αναγνωρίζεται είναι ένα *argument error*, πιο συγκεκριμένα ένα *type error*, διότι το δεύτερο όρισμα του κατηγορήματος “eats” δεν προσδιορίστηκε σωστά. Το δεύτερο λάθος που αναγνωρίζεται στην απάντηση του εκπαιδευόμενου είναι ένα *connective error*, και πιο συγκεκριμένα ένα *application error*, διότι η συνεπαγωγή δεν αφαιρέθηκε σωστά.

2.5.2.4 Παροχή Ανάδρασης

Η ανάδραση ή ανατροφοδότηση (feedback) αποτελεί μια βασική πτυχή της εκπαιδευτικής διαδικασίας. Η ανάδραση καθορίζεται ως κάθε μήνυμα και πληροφορία που παρουσιάζει το σύστημα στον εκπαιδευόμενο μετά από αίτημα του ή μετά από απόφαση του συστήματος σχετικά με την βελτίωση της απόδοσης του εκπαιδευόμενου (Mory, 2004). Η ανάδραση, για να ενισχύει την κατανόηση και την απόκτηση γνώσης του εκπαιδευόμενου, θα πρέπει να παρέχει κατάλληλες πληροφορίες επαλήθευσης (verification) και λεπτομερούς επεξεργασίας (elaboration). Η επαλήθευση αφορά την επιβεβαίωση του κατά πόσον η απάντηση του εκπαιδευόμενου είναι σωστή ή όχι, ενώ η επεξεργασία αναλύει συγκεκριμένα λάθη και καθοδηγεί τον εκπαιδευόμενο προς την σωστή απάντηση (Shute et al., 2008). Το επίπεδο της επεξεργασίας και επαλήθευσης καθορίζει την ποσότητα των παρεχόμενων πληροφοριών της ανάδρασης. Μελέτες σχετικές με την αποτελεσματικότητα έχουν δείξει ότι η παροχή βήμα προς βήμα ανάδρασης στους μαθητές μπορεί να οδηγήσει σε καλύτερα μαθησιακά αποτελέσματα και βαθύτερη κατανόηση των εννοιών που μελετήθηκαν (Corbett & Anderson, 2001). Επίσης, τα ευφυή συστήματα διδασκαλίας που προσφέρουν βήμα προς βήμα ανάδραση, μπορούν να είναι τόσο αποτελεσματικά όσο ο διδάσκοντας (VanLehn, 2011). Για την παροχή κατάλληλης ανάδρασης στο σύστημα FOLtoCF βασιστήκαμε στην εργασία (Perikos et al., 2016), όπου παρουσιάζεται ένας μηχανισμός για την αυτόματη παραγωγής ανάδρασης σε ένα εκπαιδευτικό σύστημα. Πιο συγκεκριμένα, ο μηχανισμός για την παραγωγή κατάλληλης ανάδρασης για τον εκπαιδευόμενο αποτελείται από δύο διεργασίες, την εσωτερική διεργασία (inner loop) και την εξωτερική (outer loop) (VanLehn, 2006). Ο κύριος ρόλος της εσωτερικής διεργασίας είναι να παρέχει ανάδραση στον εκπαιδευόμενο ως προς τις ενέργειές του κατά τη διάρκεια των μαθησιακών δραστηριοτήτων στο σύστημα, ενώ ο ρόλος της εξωτερικής είναι να προσδιορίζεται η κατάλληλη εκπαιδευτική διεργασία προσαρμοσμένη στο επίπεδο γνώσεων του εκπαιδευόμενου. Ο εσωτερικός βρόχος του συστήματος είναι υπεύθυνος για την ανάλυση και την αξιολόγηση της απάντησης του εκπαιδευόμενου και την παροχή των κατάλληλων μηνυμάτων ανάδρασης. Συνεπώς, η παροχή βοήθειας πραγματοποιείται από την εσωτερική διεργασία του συστήματος.

Ο μηχανισμός προσφέρει δύο είδη βοηθειών και υποδείξεων. Το πρώτο είδος βοήθειας αφορά στατικά μηνύματα και υποδείξεις που καθορίζονται από τον εκπαιδευτικό για σε κάθε βήμα της διαδικασίας. Επομένως, κάθε βήμα της διαδικασίας συνδέεται με δύο επίπεδα υποδείξεων. Το πρώτο επίπεδο αποτελείται από στατικά μηνύματα, παρέχοντας γενικές υποδείξεις σχετικά με τον στόχο για τον τρέχον βήμα, πως να κινηθεί ο εκπαιδευόμενος για να επιτευχθεί ο στόχος και επίσης προσφέρει σχετικές υπενθυμίσεις για

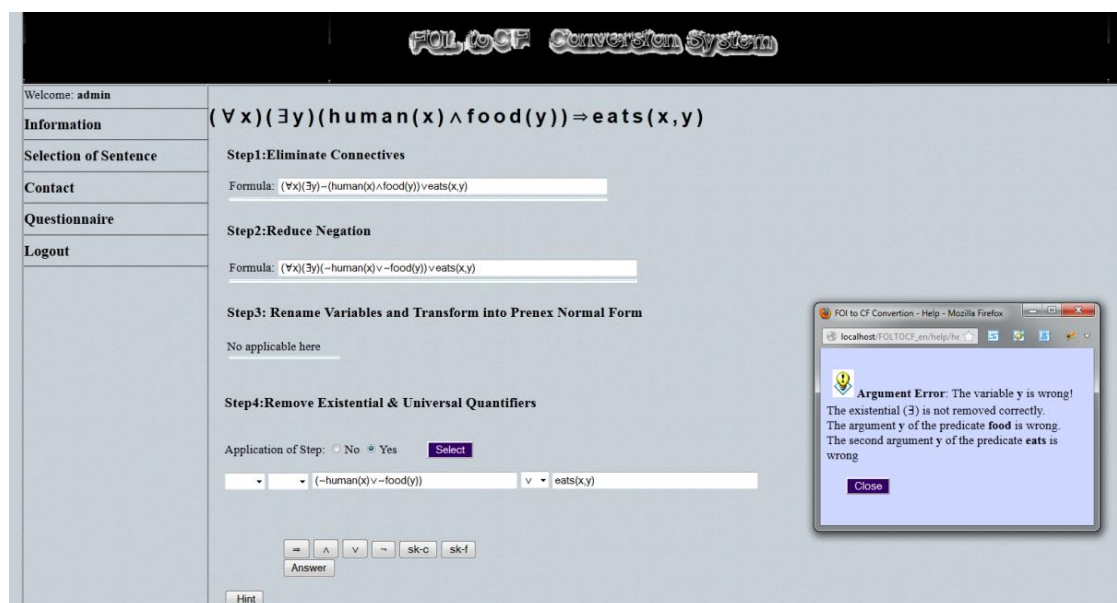
το θεωρητικό πλαίσιο (π.χ. ορισμοί, ισοδυναμίες, σύνταξη κ.ά.) που σχετίζεται με το αντίστοιχο βήμα. Στο δεύτερο επίπεδο των στατικών μηνυμάτων ενημερώνεται ο εκπαιδευόμενος για τη σωστή απάντηση του τρέχοντος βήματος.

Το δεύτερο είδος βοήθειας σχετίζεται με υποδείξεις και μηνύματα βοήθειας που καθορίζονται από την ανάλυση των απαντήσεων του εκπαιδευόμενου. Μετά τον προσδιορισμό των λαθών, ο μηχανισμός δημιουργεί βοήθειες προσαρμοσμένες στα λάθη που έγιναν. Τα μηνύματα αυτά εμφανίζονται μετά από σχετικό αίτημα του εκπαιδευόμενου για βοήθεια. Στην περίπτωση, όπου η απάντηση του εκπαιδευόμενου είναι εσφαλμένη, δημιουργούνται κατάλληλες υποδείξεις, οι οποίες είναι διαθέσιμες μέσω της επιλογής της βοήθειας.

Στη εικόνα 2.17 παρουσιάζεται η έκφραση ΚΛΠΤ:

$$“(\forall x)(\exists y)(human(x) \wedge food(y) \Rightarrow eats(x, y))”$$

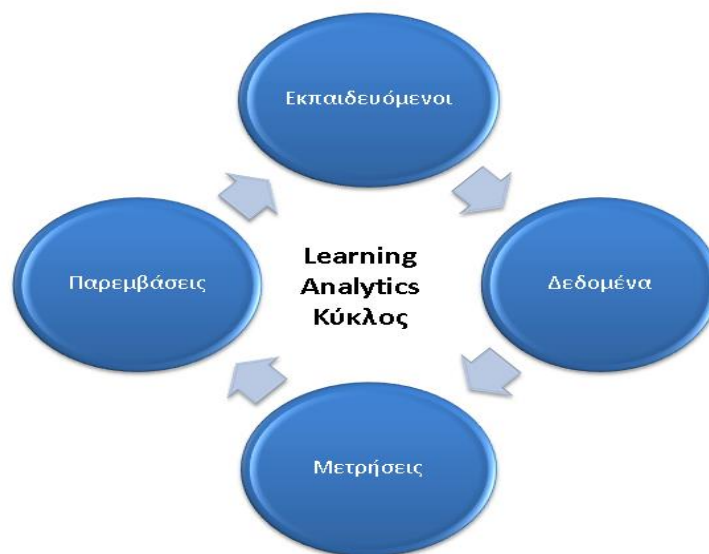
και η απάντηση του εκπαιδευόμενου για το τέταρτο βήμα της διαδικασίας μετατροπής. Η απάντηση του εκπαιδευόμενου αναλύθηκε και αναγνωρίστηκε ότι δεν είναι σωστή και προσδιορίστηκαν τα λάθη που έγιναν και παράχθηκε η κατάλληλη βοήθεια. Πιο συγκεκριμένα, μια βοήθεια αφορά το γεγονός ότι η μεταβλητή “y” δεν έχει αντικατασταθεί σωστά από μια μεταβλητή Skolem και συνεπώς δεν είναι σωστή.



Εικόνα 2.17 Βοήθεια του συστήματος FOLtoCF

2.5.2.5 Αλληλεπίδραση Διδάσκοντα με το Σύστημα

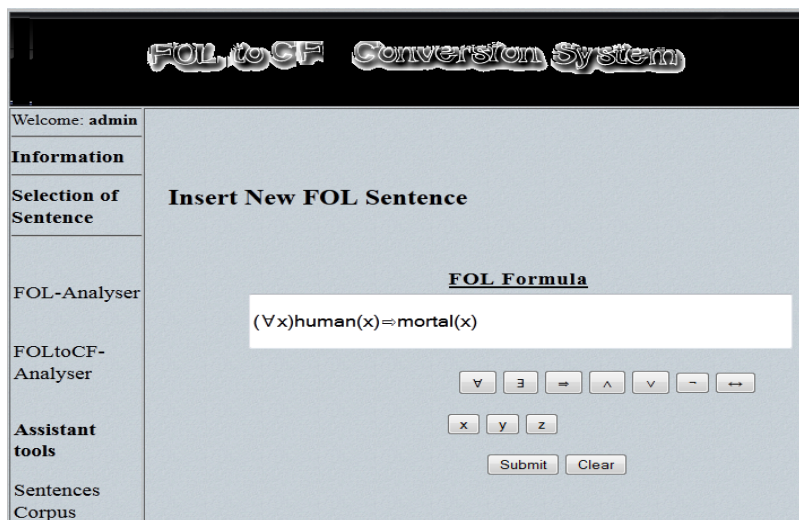
Η χρήση εργαλείων *Ανάλυσης Στοιχείων Μάθησης (Learning Analytics)* σε ένα εκπαιδευτικό σύστημα μπορεί να παρέχει ισχυρά μέσα στους διδάσκοντες για την ανάλυση της αποτελεσματικότητας των εκπαιδευτικών διαδικασιών και την βελτίωση τους, καθώς και την έμμεση ενίσχυση των επιδόσεων των μαθητών τους (Dyckhoff et al., 2012). Ο βασικός κύκλος της *Ανάλυσης Στοιχείων Μάθησης* φαίνεται στην εικόνα 2.18 (Clow, 2012). Η αρχή του κύκλου αναπαριστά τους εκπαιδευόμενους, οι οποίοι αλληλεπιδρούν με ένα σύστημα ή παίρνουν μέρος σε ένα μάθημα - MOOC (Massive Open Online Course). Το επόμενο βήμα αναπαριστά την συλλογή των δεδομένων που αφορούν τους μαθητές, όπως για παράδειγμα δημογραφικές πληροφορίες, αποτελέσματα αξιολόγησης τους κ.ά. Το τρίτο βήμα αφορά την επεξεργασία των εν λόγω δεδομένων, ώστε να γίνουν κατάλληλες μετρήσεις και αναλύσεις. Επίσης, οι αναλύσεις χρησιμοποιούνται για να οδηγήσουν σε μία ή περισσότερες παρεμβάσεις που έχουν επίδραση στην διαδικασία μάθησης των εκπαιδευόμενων.



Εικόνα 2.18 Κύκλος Learning Analytics

Στα πλαίσια του συστήματος FOLtoCF, αναπτύχθηκαν εργαλεία για την *διαχείριση διδασκαλίας*, που προσφέρουν την δυνατότητα στον διδάσκοντα να διαχειριστεί εύκολα το διδακτικό υλικό του συστήματος και να εισάγει εύκολα νέες ασκήσεις ή να τροποποιήσει υπάρχουσες. Πιο συγκεκριμένα, ο διδάσκοντας για να δημιουργήσει μια νέα άσκηση αρκεί να εισάγει την έκφραση ΚΛΠΤ που επιθυμεί, όπως φαίνεται στην εικόνα 2.19. Στην συνέχεια, γίνεται η ανάλυση της έκφρασης (εκφώνησης της άσκησης) και προσδιορίζεται αυτόματα η

σωστή απάντηση από το σύστημα και γίνεται ο προσδιορισμός του επιπέδου δυσκολίας της άσκησης.



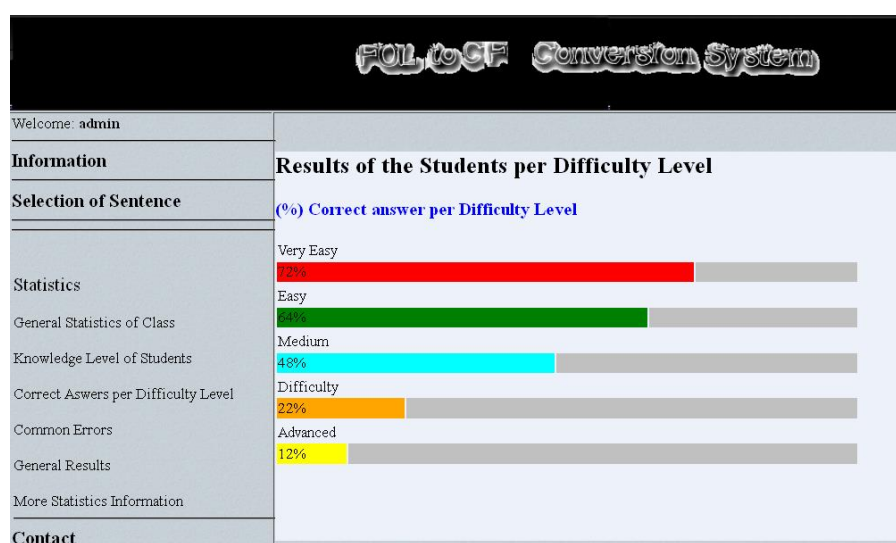
Εικόνα 2.19 Δημιουργία νέας άσκησης

Η μονάδα *Ανάλυσης Στοιχείων Μάθησης (Learning Analytics)* σχεδιάστηκε και αναπτύχθηκε ώστε να βοηθήσει τον διδάσκοντα στην παρακολούθηση των εκπαιδευτικών δραστηριοτήτων και την πρόοδο των εκπαιδευόμενων. Αρχικά, όλη η αλληλεπίδραση των εκπαιδευόμενων και οι ενέργειές τους στο σύστημα καταγράφονται και αναλύονται με στόχο να εξαχθούν χρήσιμες πληροφορίες για την μαθησιακή τους επίδοση και την απόκτηση γνώσης. Επίσης, παρέχεται στο διδάσκοντα η δυνατότητα παρακολούθησης της προόδου των εκπαιδευόμενων μέσω οπτικοποίησης των δεδομένων των δραστηριοτήτων τους, με σκοπό τη βελτιστοποίηση της μάθησης. Πιο συγκεκριμένα, το σύστημα παρέχει στατιστικά στοιχεία που αφορούν τόσο τον εκπαιδευόμενο όσο και τις ασκήσεις, όπως τα συχνότερα λάθη που έγιναν, τον χρόνο για κάθε δοκιμασία και άσκηση, το συνολικό χρόνο της αλληλεπίδρασης με το σύστημα, τις βοήθειες που ζήτησαν από το σύστημα κ.ά. Επίσης, το σύστημα παρέχει πληροφορίες σχετικές με τις ασκήσεις όπως, ο αριθμός των υποδείξεων του ζητηθήκαν για κάθε άσκηση, ο αριθμός των προσπαθειών που πραγματοποίησαν οι εκπαιδευόμενοι για να διορθώσουν τα λάθη τους και να ολοκληρώσουν την άσκηση. Επιπλέον, το σύστημα οπτικοποιεί τα δεδομένα και με χρήση γραφικών αναπαραστάσεων. Μερικά στατιστικά που παρέχει το σύστημα στον διδάσκοντα είναι τα ακόλουθα:

- Το ποσοστό των σωστών απαντήσεων στις ασκήσεις ανά επίπεδο δυσκολίας ανά εκπαιδευόμενο.
- Ο αριθμός λαθών που έγιναν ανά άσκηση ανά εκπαιδευόμενο.
- Ο τύπος των λαθών ανά άσκηση ανά επίπεδο δυσκολίας ανά εκπαιδευόμενο.

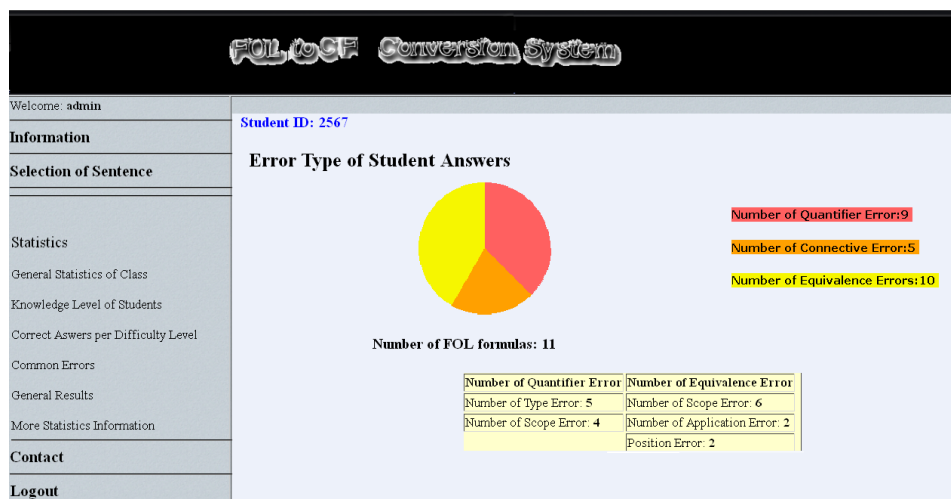
- Τα αιτήματα για βοήθεια ανά άσκηση και ανά επίπεδο δυσκολίας.
- Ο αριθμός των προσπαθειών του εκπαιδευόμενου ανά άσκηση ανά επίπεδο δυσκολίας.
- Ο μέσος χρόνος που απαιτείται ανά άσκηση ανά επίπεδο δυσκολίας ανά σπουδαστή.
- Το ποσοστό των λαθών που έγιναν ανά επίπεδο δυσκολίας ανά εκπαιδευόμενο.
- Το ποσοστό των σωστών απαντήσεων στις ασκήσεις ανά επίπεδο δυσκολίας για όλους τους εκπαιδευόμενους
- Το ποσοστό των λαθών που έγιναν ανά επίπεδο δυσκολίας για όλους τους εκπαιδευόμενους.
- Η ελάχιστη και η μέγιστη επίδοση των εκπαιδευόμενων ανά άσκηση ανά επίπεδο δυσκολίας.

Στην εικόνα 2.20 παρουσιάζεται ένα γράφημα που αναπαριστά το ποσοστό των σωστών απαντήσεων για τους εκπαιδευόμενους σε σχέση με το επίπεδο δυσκολίας των ασκήσεων. Για παράδειγμα, οι εκπαιδευόμενοι απάντησαν σωστά στο 72% των ασκήσεων που ήταν επιπέδου δυσκολίας “very easy”, απάντησαν σωστά στο 64% των ασκήσεων επιπέδου “easy” και μόλις στο 12% των ασκήσεων επιπέδου “advanced”.



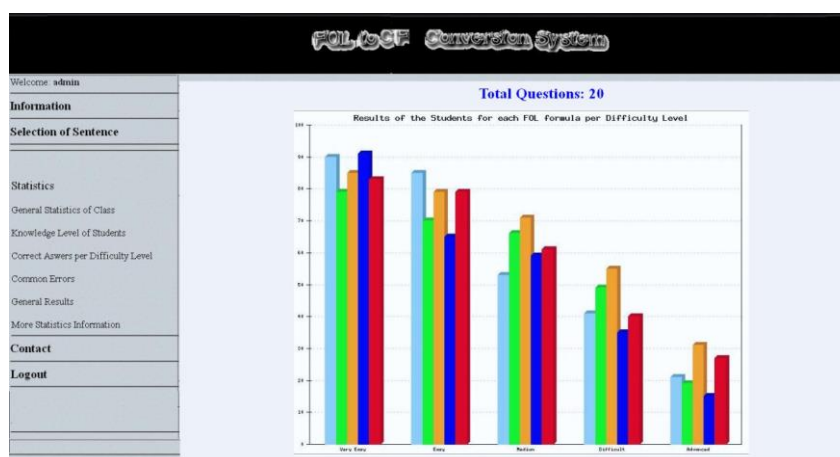
Εικόνα 2.20 Γραφική Αναπαράσταση των εκπαιδευόμενων ανά επίπεδο δυσκολίας της άσκησης

Επίσης, όπως αναφέρθηκε παραπάνω, το σύστημα FOLtoCF μπορεί να εντοπίσει και να αναγνωρίσει τον αριθμό και τον τύπο λαθών που έγιναν και μέσω των εργαλείων που διαθέτει ο διδάσκοντας, να γίνει παρουσίαση στατιστικών πληροφοριών σχετικών με τον τύπο λαθών που έγιναν. Στην εικόνα 2.21 παρουσιάζεται ο αριθμός και ο τύπος λαθών που έγιναν από τον εκπαιδευόμενο και αντίστοιχα στην εικόνα 2.22 παρουσιάζονται οι επιδόσεις των εκπαιδευόμενων για κάθε επίπεδο δυσκολίας των ασκήσεων.



Εικόνα 2.21 Τύποι των λαθών των εκπαιδευόμενων

Επίσης, ο διδάσκοντας χρησιμοποιώντας τις γραφικές παραστάσεις μπορεί να ανιχνεύσει ευκολότερα μαθησιακά προβλήματα που παρουσιάστηκαν κατά την διαδικασία, για παράδειγμα, ένα πολύ χαμηλό αριθμό σωστών απαντήσεων σε ένα συγκεκριμένο βήμα της διαδικασίας FOLtoCF.



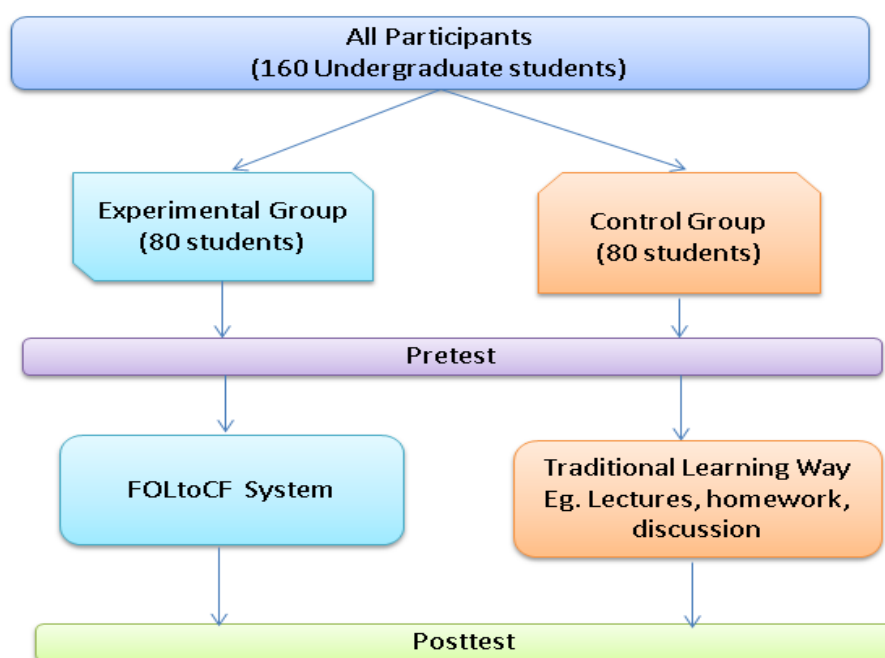
Εικόνα 2.22 Οι απαντήσεις των εκπαιδευόμενων ανά επίπεδο δυσκολίας.

Επομένως, τα δεδομένα που λαμβάνονται από το εργαλείο μαθησιακής ανάλυσης μπορούν να χρησιμοποιηθούν από το διδάσκοντα για να αναπροσαρμοστεί κατάλληλα το εκπαιδευτικό υλικό, ανάλογα με το επίπεδο των εκπαιδευόμενων, όπως για παράδειγμα παρέχοντας πιο πολλά παραδείγματα σε είδη ασκήσεων που αντιμετωπίζουν δυσκολίες. Επίσης, με βάση τις επιδόσεις των εκπαιδευόμενων, αναπροσαρμόζεται το συνολικό επίπεδο δυσκολίας των ασκήσεων και παράγεται το κατάλληλο είδος βοήθειας. Επομένως, το εργαλείο μαθησιακής ανάλυσης είναι χρήσιμο για τον διδάσκοντα και μπορεί στην συμβάλλει στη βελτίωση της εκπαιδευτικής διαδικασίας.

2.5.2.6 Πειραματικά Αποτελέσματα για το Σύστημα FOLtoCF

Για να μελετήσουμε την αποτελεσματικότητα των εκπαιδευτικών δυνατοτήτων του συστήματος FOLtoCF σχεδιάστηκε και πραγματοποιήθηκε μια πειραματική μελέτη. Στην πειραματική μελέτη επιλέχθηκαν τυχαία να συμμετέχουν 160 προπτυχιακοί φοιτητές (άνδρες-γυναίκες) που παρακολουθούσαν το μάθημα της τεχνητής νοημοσύνης. Αρχικά, για τις ανάγκες της μελέτης, δημιουργήσαμε δύο ομάδες φοιτητών, όπου κάθε μια περιείχε 80 φοιτητές, και επιλέχθηκε τυχαία η μια από τις δύο να είναι η *ομάδα πειράματος* (*experimental group*), ενώ η άλλη η *ομάδα ελέγχου* (*control group*). Πιο συγκεκριμένα, στα πλαίσια της μελέτης, η *ομάδα ελέγχου* (*control group*) χρησιμοποίησε για την μάθηση τον παραδοσιακό τρόπο διδασκαλίας και μελέτησε την διαδικασία της μετατροπής εκφράσεων ΚΛΠΤ σε ΠΜ μέσω διαλέξεων, σημειώσεων και παραδειγμάτων του μαθήματος, ενώ η πειραματική ομάδα (*experimental group*) μελέτησε μέσω αλληλεπίδρασης με το σύστημα FOLtoCF. Η διάρθρωση της πειραματικής μελέτης παρουσιάζεται στην εικόνα 2.23.

Αρχικά, δόθηκε ένα τεστ (pretest) και στις δυο ομάδες φοιτητών (*experimental* και *control*) για να αξιολογηθούν οι επιδόσεις των φοιτητών στην διαδικασία μετατροπής ΚΛΠΤ σε ΠΜ πριν την διεξαγωγή του πειράματος. Η διάρκεια του pretest ήταν 35 λεπτά και όλοι οι φοιτητές κλήθηκαν να μετατρέψουν 10 εκφράσεις ΚΛΠΤ σε ΠΜ. Στην συνέχεια, οι απαντήσεις των φοιτητών αξιολογήθηκαν από τον διδάσκοντα και βαθμολογήθηκαν. Έπειτα, η πειραματική ομάδα (*experimental group*) μελέτησε παραδείγματα και μετέτρεψε βηματικά προτάσεις ΚΛΠΤ σε ΠΜ μέσω του συστήματος FOLtoCF. Αντίθετα, η ομάδα ελέγχου (*Control group*) μελέτησε και έκανε ασκήσεις χρησιμοποιώντας το υλικό των βιβλίων και των διαλέξεων. Η διάρκεια του πειράματος ήταν μια εβδομάδα και για τις δυο ομάδες και οι ασκήσεις που μελέτησαν οι δυο ομάδες ήταν οι ίδιες ή ισοδύναμες. Στην συνέχεια, όλοι οι συμμετέχοντες έλαβαν ένα τελικό τεστ (posttest), όπου κλήθηκαν να μετατρέψουν 10 εκφράσεις ΚΛΠΤ σε ΠΜ και η διάρκεια του τεστ ήταν 30 λεπτά. Επίσης, τα δυο τεστ (pretest και posttest) ήταν του ίδιου επίπεδου δυσκολίας. Τέλος, οι φοιτητές της πειραματικής ομάδας (*experimental group*), που μελέτησαν μέσω του συστήματος FOLtoCF, συμπλήρωσαν ένα ερωτηματολόγιο για να εκφράσουν την γνώμη τους, τις εμπειρίες τους καθώς και να αξιολογήσουν τις δυνατότητες του συστήματος.



Εικόνα 2.23 Διαδικασία Πειραματικής Αξιολόγησης του Συστήματος FOLtoCF

Για να συγκρίνουμε την απόδοση των φοιτητών στις δυο ομάδες πραγματοποιήσαμε στατιστικό έλεγχο υποθέσεων (hypothesis testing).

Μηδενική Υπόθεση H_0 : Δεν υπάρχει διαφορά στις επιδόσεις των εκπαιδευόμενων των δυο ομάδων.

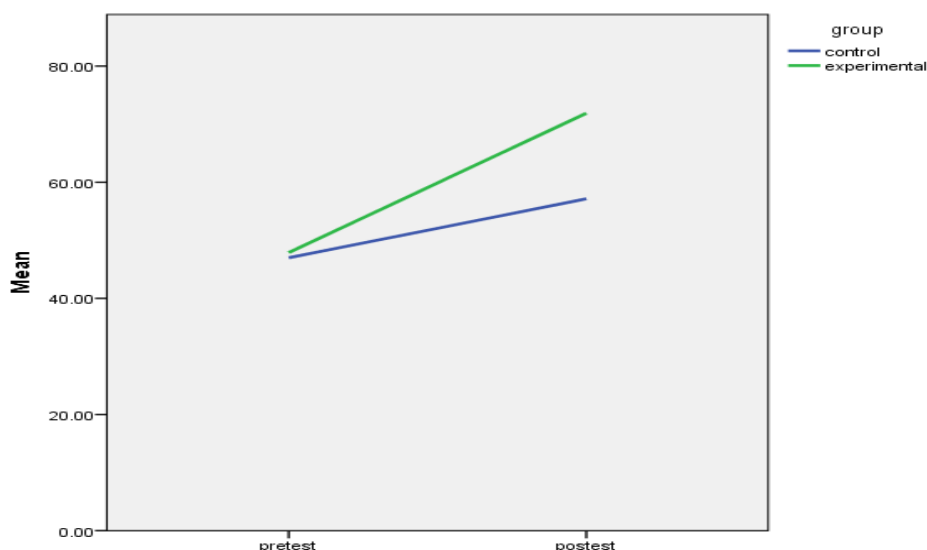
Εναλλακτική Υπόθεση H_1 : Είναι η «θετική» εκδοχή της μηδενικής υπόθεσης, δηλαδή «Υπάρχει διαφορά στις επιδόσεις των εκπαιδευόμενων των δυο ομάδων»

Πιο συγκεκριμένα, έγινε σύγκριση των μέσων τιμών των δυο ανεξάρτητων δειγμάτων *t-test* (*Independent Samples t-test*) για τις επιδόσεις των εκπαιδευόμενων στο pretest, με δεδομένο ότι τα δυο δείγματα ακολουθούν κανονική κατανομή και οι διασπορές δεν απέχουν πολύ. Από τα αποτελέσματα των εκπαιδευόμενων στο pretest, προκύπτει ότι η μέση τιμή και τυπική απόκλιση ήταν 47.88 και 13.84 για την πειραματική ομάδα και 46.99 και 15.12 αντίστοιχα για την ομάδα ελέγχου. Από τα αποτελέσματα προκύπτει ότι η τιμή p (επίπεδο σημαντικότητας) ήταν $p = 0.215$ ($p > 0.05$) και το effect size $d = 0.061$ (Cohen 1988). Επομένως, από τα αποτελέσματα φαίνεται ότι αποδεχόμαστε την μηδενική υπόθεση ($\alpha = 0.05$) και ότι οι δύο ομάδες των φοιτητών είχαν το ίδιο επίπεδο γνώσεων πριν την διεξαγωγή του πειράματος.

Πίνακας 2.7 Περιγραφικά Μέτρα Pretest και Posttest

	Group	Mean	SD	SE
Pretest	Experimental	47.88	13.84	1.55
	Control	46.99	15.12	1.69
Posttest	Experimental	71.89	14.94	1.67
	Control	57.15	13.66	1.53

Στον πίνακα 2.7 παρουσιάζονται τα περιγραφικά μέτρα και για τις δυο ομάδες και τα αντίστοιχα αποτελέσματα από τις δυο πειραματικές δοκιμές. Από τα αποτελέσματα του πίνακα προκύπτει ότι η μέση τιμή της πειραματικής ομάδας (experimental) αυξήθηκε από 47.88 σε 71.89 για την δοκιμασία του posttest, ενώ η μέση τιμή της ομάδας ελέγχου (control group) αυξήθηκαν από 46.99 σε 57.15. Όπως φαίνεται στην εικόνα 2.24 η μέση τιμή του posttest για την πειραματική ομάδα ήταν υψηλότερη από την μέση τιμή του posttest της ομάδας ελέγχου.



Εικόνα 2.24 Μέση τιμή για κάθε ομάδα για το Pretest και το Posttest

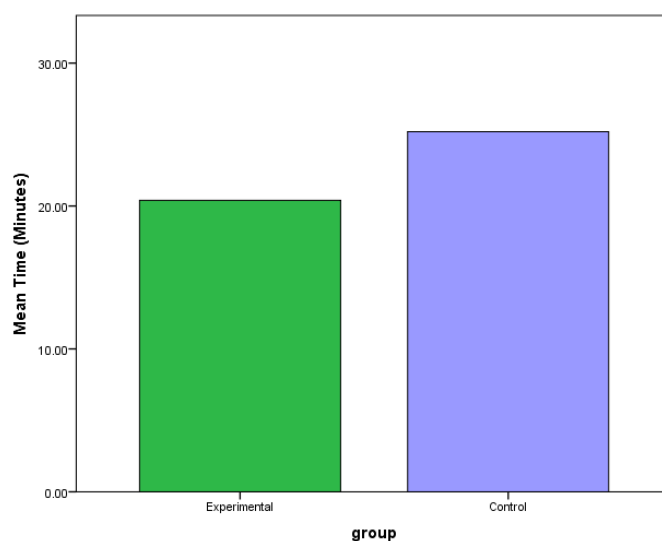
Όπως παρουσιάζεται στον Πίνακα 2.8, γίνεται έλεγχος των στατιστικών υποθέσεων στην δεύτερη και στην τρίτη στήλη του πίνακα (Levene's Test for Equality of Variances), δίνεται η τιμή της στατιστικής συνάρτησης ελέγχου (F) και το p-value (Sig.) του ελέγχου του Levene test για την ισότητα διασπορών. Η μηδενική υπόθεση ελέγχου είναι ότι οι διασπορές των δύο

πληθυσμών είναι ίσες και συνεπώς στην περίπτωση αυτή δεν απορρίπτεται η μηδενική υπόθεση και μπορούμε να συμπεράνουμε ότι έχει νόημα η υπόθεση της ισότητας των διασπορών. Με βάση τα πειραματικά αποτελέσματα του posttest, προκύπτει ότι οι φοιτητές της πειραματικής ομάδας έδωσαν περισσότερες σωστές απαντήσεις (προτάσεις σε ΠΜ) από την ομάδα ελέγχου. Επομένως, τα αποτελέσματα δείχνουν ότι η πειραματική ομάδα είχε καλύτερες επιδόσεις και κατανόησε καλύτερα τη διαδικασία μετατροπής προτάσεων ΚΛΠΤ σε ΠΜ από την ομάδα ελέγχου.

Πίνακας 2.8 *t-Test Results*

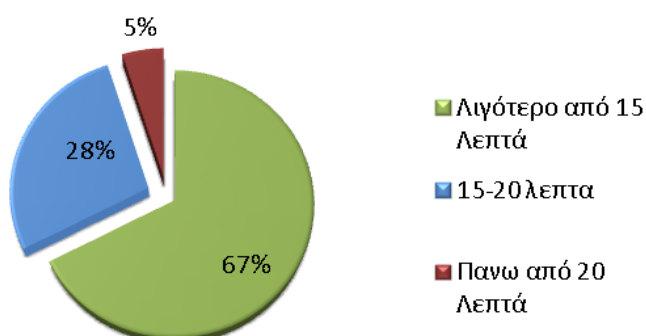
Equality of Variance		Levene's test for equality of variances		t-Test for mean					
		Sig.		T	Df	Sig. (2-tailed)	MD	95% Interval Difference	Confidence of the
								Lower	Upper
Pretest	Equal	1.55	0.21	0.387	158	0.69	2.29	-3.64	5.41
	Unequal			-0.751	156.7	0.69	2.29	-3.64	5.41
Posttest	Equal	0.64		6.513	158	0.00	2.26	10.26	19.21
	Unequal	0.255		6.513	156.74	0.00	2.26	19.26	19.21

Στην εικόνα 2.25 τα αποτελέσματα δείχνουν ότι η πειραματική ομάδα χρειάστηκε περισσότερο χρόνο για το posttest, που κυμαίνονταν από 10 με 30 λεπτά, με μέσο χρόνο τα 20 λεπτά, ενώ για την ομάδα ελέγχου ο χρόνος εξοικείωσης με το σύστημα κυμαινόταν από 20 σε 30 λεπτά με μέσο χρόνο τα 25 λεπτά. Επομένως, ο μέσος χρόνος για την πειραματική ομάδα ήταν 20 λεπτά και ήταν λιγότερος από το μέσο χρόνο της ομάδας ελέγχου (25 λεπτά).



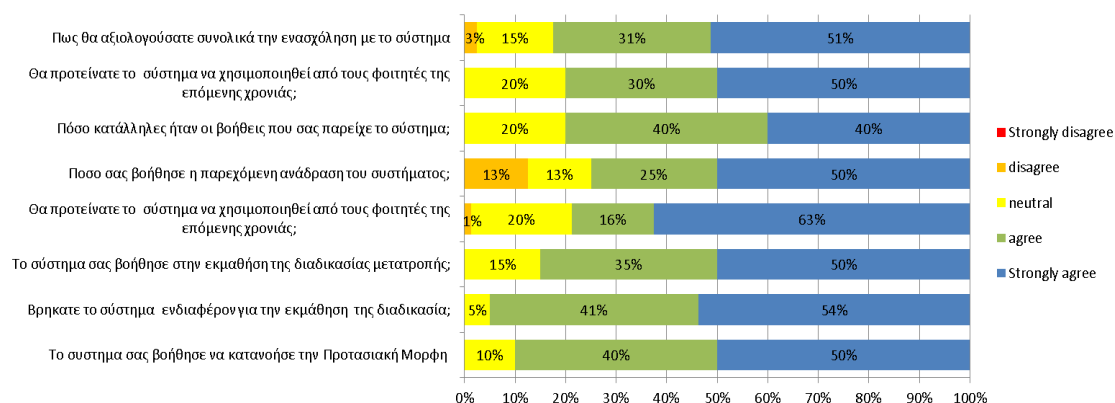
Εικόνα 2.25 Αποτελέσματα χρόνου επίλυση του Posttest

Στην συνέχεια, οι φοιτητές της πειραματικής ομάδας, που αλληλεπίδρασαν με το σύστημα FOLtoCF, συμπλήρωσαν ένα ερωτηματολόγιο. Το ερωτηματολόγιο περιείχε ερωτήσεις για να αξιολογήσουν τις εκπαιδευτικές και λειτουργικές συνιστώσες του συστήματος, την ευχρηστία του, να αποτυπώσουν την γνώμη τους και τις εμπειρίες τους με το σύστημα, καθώς και τις απόψεις τους σχετικά με την επίδραση του συστήματος στην εκμάθηση της διαδικασίας. Το ερωτηματολόγιο περιλάμβανε 11 ερωτήσεις. Οι ερωτήσεις 1-8 ήταν τύπου Likert scale (1: καθόλου, 5: παρά πολύ), ενώ οι ερωτήσεις 9-10 ήταν ανοικτού τύπου, όπου οι φοιτητές διατύπωσαν την γνώμη τους για το σύστημα καθώς και πιθανά προβλήματα που αντιμετώπισαν. Επίσης, η ερώτηση 11 αναφερόταν στο χρόνο που χρειάστηκαν οι εκπαιδευόμενοι για να εξοικειωθούν με το σύστημα και περιελάμβανε τρεις πιθανές απαντήσεις: α) λιγότερο από 15 λεπτά β) 15-20 λεπτά και γ) περισσότερα από 20 λεπτά. Στην εικόνα 2.26 παρουσιάζονται τα αποτελέσματα των απαντήσεων σχετικά με το χρόνο εξοικειώσεως, όπου το 95% των φοιτητών χρειάστηκε λιγότερο από 20 λεπτά και μόνο το 5% πάνω από 20 λεπτά.



Εικόνα 2.26 Χρόνος εξοικείωσης των φοιτητών με το Σύστημα

Στην εικόνα 2.27 παρουσιάζονται τα αποτελέσματα του ερωτηματολογίου που δείχνουν ότι οι φοιτητές βρήκαν ευχάριστη και ενδιαφέρουσα την εμπειρία χρήσης του συστήματος FOLtoCF. Οι απαντήσεις των φοιτητών δείχνουν ότι το 85% απάντησε ότι το σύστημα τους βοήθησε στην εκμάθηση της διαδικασίας μετατροπής εκφράσεων ΚΛΠΤ σε ΠΜ και ότι το σύστημα τους καθοδήγησε κατάλληλα κατά τα βήματα της διαδικασίας και τους παρείχε κατάλληλη βοήθεια (feedback) και υποδείξεις για τα λάθη τους.



Εικόνα 2.27 Αποτελέσματα Ερωτηματολογίου για το Σύστημα FOLtoCF

Για τον έλεγχο της αξιοπιστίας, υπολογίστηκε ο δείκτης Cronbach's α (alpha) ή δείκτης εσωτερικής συνάφειας (internal consistency coefficient) (Cronbach, 1951). Οι τιμές του δείκτη κυμαίνονται από 0 μέχρι 1 και αποτελεί έναν από τους πιο γνωστούς και ευρέως χρησιμοποιούμενους δείκτες εσωτερικής συνάφειας (Hatcher, 1994). Όταν η τιμή του δείκτη Cronbach Alpha είναι μεγαλύτερη του 0.80 τότε έχουμε πολύ υψηλή αξιοπιστία στη μέτρηση.

Στα πλαίσια του πειράματός μας, ο συντελεστής αξιοπιστίας του ερωτηματολογίου υπολογίστηκε ότι είναι 0.98 ($\alpha=0.98$), δηλαδή πολύ μεγαλύτερος από το 0.8, το οποίο δείχνει ότι το ερωτηματολόγιο έχει πολύ υψηλή αξιοπιστία. Γενικά, τα πειραματικά αποτελέσματα έδειξαν ότι το σύστημα βοήθησε τους φοιτητές να κατανοήσουν καλύτερα την διαδικασία της μετατροπής ΚΛΠΤ σε ΠΜ, η οποία είναι μια διαδικασία που δυσκολεύει τους περισσότερους φοιτητές. Επίσης, προκύπτει ότι η βηματική προσέγγιση και η παροχή ανάδρασης σε κάθε βήμα της διαδικασίας βοήθησε τους φοιτητές και συντέλεσε ώστε να γίνει η εκπαιδευτική διαδικασία πιο αποτελεσματική.

2.5.3 Ευφυές Σύστημα Μετατροπής Προτάσεων Κατηγορηματικής Λογικής σε Ισοδύναμες

2.5.3.1 Λογικές Ισοδυναμίες

Στη λογική δύο εκφράσεις p και q είναι λογικά ισοδύναμες αν έχουν το ίδιο λογικό περιεχόμενο. Πιο συγκεκριμένα, μπορούμε να ορίσουμε ότι δύο εκφράσεις p και q είναι λογικά ισοδύναμες (*logically equivalent*) εάν οι πίνακες αληθείας είναι ίδιοι κάτω από οποιαδήποτε ερμηνεία. Η έννοια της λογικής ισοδυναμίας εμφανίζεται τόσο στην Προτασιακή Λογική (ΠΛ) όσο και στην Κατηγορηματική Λογική Πρώτης Τάξης (ΚΛΠΤ). Στην ΚΛΠΤ υπάρχουν επιπλέον κανόνες σε σχέση με την ΠΜ που αφορούν τους ποσοδείκτες. Οι βασικές λογικές ισοδυναμίες στην ΚΛΠΤ είναι οι εξής:

Μοναδιαίος Νόμος (Idempotent Law):

$$p \vee p \leftrightarrow p$$

$$p \wedge p \leftrightarrow p$$

Νόμος της διπλής άρνησης (Double Negation Law): δηλώνει αν μία πρόταση p είναι αληθής, τότε δεν ισχύει ότι η πρόταση p δεν είναι αληθής. Ο νόμος ισχύει και αντίστροφα.

$$p \leftrightarrow \neg \neg p$$

Αντιμεταθετικός Νόμος (Commutative Law): Οι λογικές πράξεις ($\vee$, $\wedge$) έχουν την αντιμεταθετική ιδιότητα:

$$p \vee q \leftrightarrow q \vee p$$

$$p \wedge q \leftrightarrow q \wedge p$$

Προσεταιριστικός νόμος (Associative Law): Όταν σε μία πρόταση υπάρχουν δύο λογικές πράξεις, δύο λογικές διαζεύξεις ($\vee$) ή δυο λογικές συζεύξεις ($\wedge$), τότε η σειρά με την οποία γίνονται οι πράξεις δεν επηρεάζει το τελικό αποτέλεσμα. Οπότε, έχουμε:

$$(p \vee q) \vee r \leftrightarrow p \vee (q \vee r)$$

$$(p \wedge q) \wedge r \leftrightarrow p \wedge (q \wedge r)$$

Επιμεριστικός Νόμος (Distributive Law): Επιμερισμός ως προς τη σύζευξη και την διάζευξη αντίστοιχα ως εξής:

$$(p \vee q) \vee r \leftrightarrow p \vee (q \vee r)$$

$$(p \wedge q) \wedge r \leftrightarrow p \wedge (q \wedge r)$$

Νόμοι De Morgan (De Morgan's Laws) : Δηλώνουν ότι «Η άρνηση μίας διάζευξης ισούται με τη σύζευξη των αρνήσεων» και αντίστοιχα «η άρνηση μίας σύζευξης ισούται με τη διάζευξη των αρνήσεων». Οι νόμοι αναπαριστούνται ως εξής:

$$\neg(p \vee q) \leftrightarrow \neg p \wedge \neg q$$

$$\neg(p \wedge q) \leftrightarrow \neg p \vee \neg q$$

Απαλοιφή Συνεπαγωγής (Implication Elimination): Αφαιρεί την συνεπαγωγή από την πρώτη προϋπόθεση.

$$(p \Rightarrow q) \leftrightarrow (\neg p \vee q)$$

Απαλοιφή Διπλής Συνεπαγωγής (Biconditional Elimination): Η διπλή συνεπαγωγή αφαιρείται με την χρήση της παρακάτω ισοδυναμίας:

$$(p \Leftrightarrow q) \leftrightarrow (p \Rightarrow q) \wedge (q \Rightarrow p)$$

Νόμοι De Morgan για Ποσοδείκτες (De Morgan's Laws): Οι νόμοι De Morgan επεκτείνονται για να μπορούν να εφαρμοστούν σε ποσοδείκτες πάνω στην ΚΛΠΤ ως εξής:

$$\forall x P(x) \leftrightarrow \neg \exists x \neg P(x)$$

$$\neg \exists x P(x) \leftrightarrow \forall x \neg P(x)$$

Επιμεριστικός Νόμος για Ποσοδείκτες (Distributive Law): Η επιμεριστική ιδιότητα ισχύει και για τους ποσοδείκτες στην ΚΛΠΤ:

$$\forall x (P(x) \wedge Q(x)) \leftrightarrow \forall x P(x) \wedge \forall x Q(x)$$

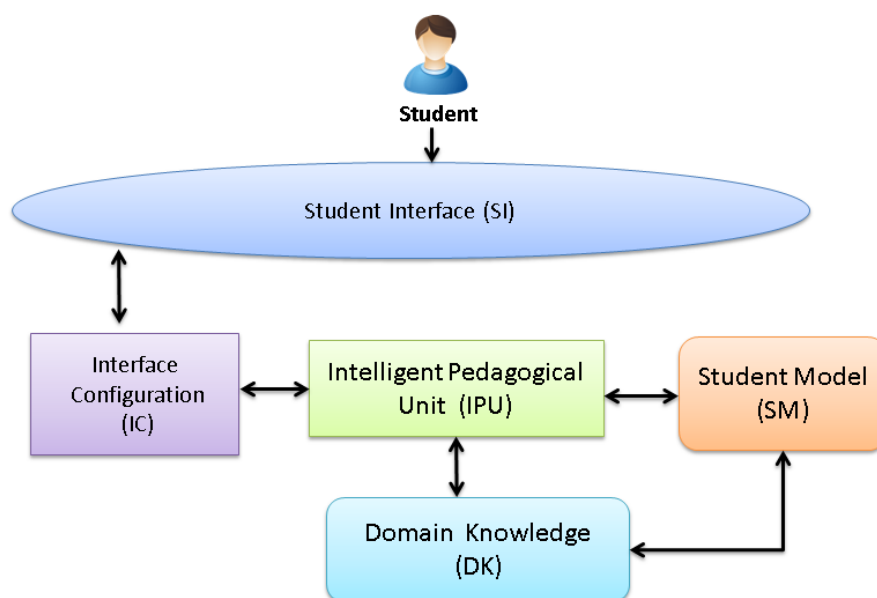
$$\exists x (P(x) \vee Q(x)) \leftrightarrow \exists x P(x) \vee \exists x Q(x)$$

2.5.3.2 Αρχιτεκτονική Συστήματος

Η αρχιτεκτονική του συστήματος ακολουθεί την κλασσική αρχιτεκτονική ενός Ευφυούς Συστήματος Διδασκαλίας (ΕΣΔ) και περιλαμβάνει τις ακόλουθες κύριες μονάδες: τη Γνώση Πεδίου (*Domain Knowledge*), το *Μοντέλο Μαθητή (Student Model-SM)*, την *Διεπαφή Χρήστη*

(*Student Interface-SI*), την *Διαμόρφωση Διεπαφής (Interface Configuration-IC)* και την *Ευφυή Παιδαγωγική Μονάδα (Intelligent Pedagogical Unit-IPU)* (εικόνα 2.28).

Η *γνώση Πεδίου* περιέχει το εκπαιδευτικό υλικό του συστήματος διδασκαλίας και αποτελεί τη βάση γνώσης του συστήματος. Κυρίως επικεντρώνεται στην παροχή βοήθειας στους εκπαιδευόμενους, για να μάθουν να δημιουργούν ισοδύναμες εκφράσεις ΚΛΠΤ και να κατανοήσουν τους κανόνες ισοδυναμίας. Αρχικά, ο εκπαιδευόμενος μέσω του *διεπαφής* *χρήστη* εγγράφεται στο σύστημα και καλείται να καταχωρήσει τις απαιτούμενες προσωπικές πληροφορίες, όπως το όνομα του, το φύλο, το έτος σπουδών κ.ά. Ο σχεδιασμός της διεπαφής έγινε με τέτοιο τρόπο ώστε να είναι απλή και εύκολη στην χρήση από τους εκπαιδευόμενους. Η *διαμόρφωση διεπαφής* ρυθμίζει την διεπαφή του εκπαιδευόμενου σύμφωνα με τις ανάγκες και τα χαρακτηριστικά της τρέχουσας εκπαιδευτικής δραστηριότητας καθώς και της παρεχόμενης βοήθειας. Επομένως, η διεπαφή του εκπαιδευόμενου είναι δυναμική και αναπροσαρμόζεται στις ανάγκες της τρέχουσας εκπαιδευτικής διαδικασίας. Το *μοντέλο μαθητή* αναπτύχθηκε για την καταγραφή και των πληροφοριών που σχετίζονται με τον εκπαιδευόμενο, όπως είναι οι μαθησιακές ικανότητές του, η πρόοδός του κατά την διαδικασία μάθησης και τα προσωπικά του στοιχεία.



Εικόνα 2.28 Αρχιτεκτονική του Ευφυούς Συστήματος Μετατροπής Κατηγορηματικής Λογικής σε Ισοδύναμες

Η *ευφυής παιδαγωγική μονάδα* αποτελεί το παιδαγωγικό μοντέλο και σχεδιάστηκε με τέτοιο τρόπο ώστε, να εκμεταλλεύεται τις παιδαγωγικές τεχνικές για την επίτευξη των

εκπαιδευτικών στόχων του συστήματος. Η μονάδα αυτή είναι υπεύθυνη να καθοδηγήσει τον εκπαιδευόμενο στην κατανόηση των λογικών ισοδυναμιών, παρέχοντας την κατάλληλη βοήθεια και ανάδραση κατά την διάρκεια της μάθησης. Επίσης, αναγνωρίζει τα λάθη που γίνονται και παρέχει κατάλληλη ανάδραση με βάση τα λάθη αυτά. Πιο συγκεκριμένα, η μονάδα διαχειρίζεται τις ενέργειες του εκπαιδευόμενου για κάθε άσκηση ισοδυναμίας ως εξής:

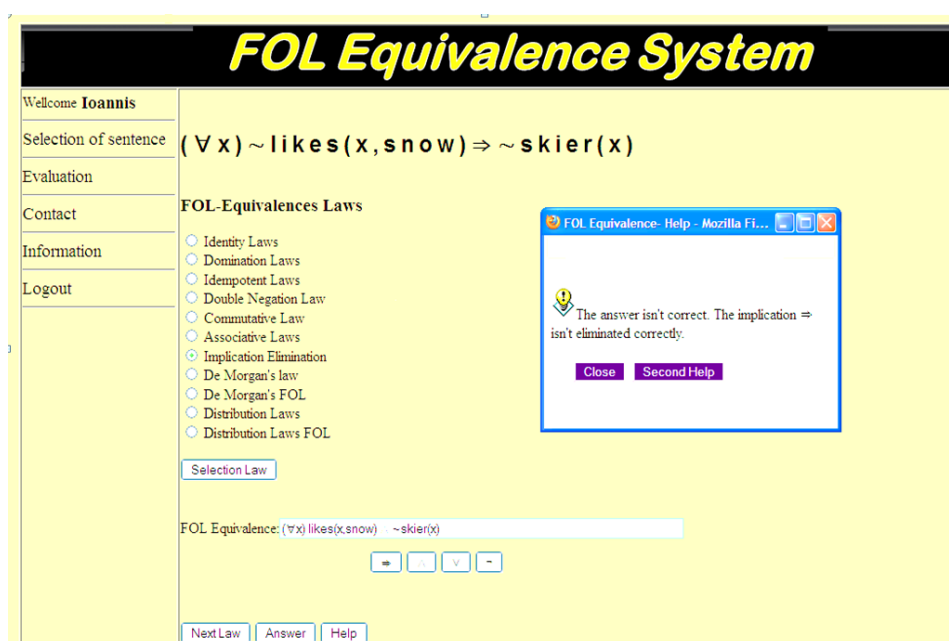
1. Αρχικά, ο εκπαιδευόμενος επιλέγει έναν κανόνα ισοδυναμίας για να τον εφαρμόσει σε μια έκφραση ΚΛΠΤ.
2. Ελέγχει αν μπορεί να εφαρμοστεί ο τρέχον κανόνας ισοδυναμίας:
 - Εάν εφαρμόζεται, τότε επιτρέπει στον εκπαιδευόμενο να εισάγει την απάντηση για τον τρέχοντα κανόνα.
 - Διαφορετικά, ενημερώνει τον εκπαιδευόμενο ότι ο επιλεγμένος κανόνας δεν εφαρμόζεται και παρέχει το κατάλληλο μήνυμα και τη δυνατότητα να επιλέξει ένα νέο κανόνα.
3. Ελέγχει την απάντηση για τον τρέχοντα επιλεγμένο κανόνα.
 - Αν η απάντηση είναι σωστή, ενημερώνει τον εκπαιδευόμενο και πηγαίνει στο βήμα1.
 - Διαφορετικά:
 - (α) Γίνεται προσδιορισμός των λαθών που έγιναν
 - (β) Παρέχεται κατάλληλη ανάδραση με βάση το είδος των λαθών και τα χαρακτηριστικά του αντίστοιχου κανόνα ισοδυναμίας
 - (γ) Επιτρέπεται στον εκπαιδευόμενο να εισάγει μια νέα απάντηση για τον τρέχοντα κανόνα και να πηγαίνει στο βήμα1.

2.5.3.3 Αλληλεπίδραση Συστήματος με τον Χρήστη

Ο εκπαιδευόμενος όταν συνδεθεί στο σύστημα έχει την δυνατότητα να προσπελάσει τις διαθέσιμες ασκήσεις και να επιλέξει την άσκηση-πρόταση που επιθυμεί να μετατρέψει σε ισοδύναμη μορφή. Η διαδικασία μετατροπής γίνεται βηματικά και σε κάθε βήμα καλείται ο εκπαιδευόμενος να επιλέξει και να εφαρμόσει έναν κανόνα ισοδυναμίας. Σε κάθε βήμα της διαδικασίας παρέχεται βοήθεια από το σύστημα την οποία μπορεί να ζητήσει ο εκπαιδευόμενος όποτε το επιθυμεί.

Αρχικά, ο εκπαιδευόμενος καλείται να επιλέξει τον κατάλληλο κανόνα ισοδυναμίας που μπορεί να εφαρμοστεί στην έκφραση ΚΛΠΤ από την λίστα όλων των κανόνων ισοδυναμίας. Στην περίπτωση που ο κανόνας δεν μπορεί να εφαρμοστεί, το σύστημα παρέχει κατάλληλη

βοήθεια και επεξηγήσεις με βάση μια γενική μεθοδολογία παραγωγής ανάδρασης (Perikos et al., 2016). Στην περίπτωση που ο κανόνας που επιλέχθηκε μπορεί να εφαρμοστεί, τότε ο εκπαιδευόμενος καλείται να προσδιορίσει τη νέα ισοδύναμη έκφραση ΚΛΠΤ. Στην συνέχεια, το σύστημα αναλύει την απάντηση και τον ενημερώνει αν η απάντηση είναι σωστή ή λανθασμένη και παρέχει τις κατάλληλες επεξηγήσεις. Επίσης, το σύστημα ενημερώνει τον εκπαιδευόμενο, μετά από κάθε επιλογή ενός κανόνα ισοδυναμίας, σχετικά με όλους τους προηγούμενους κανόνες που έχει εφαρμόσει μέχρι την τρέχουσα κατάσταση. Για παράδειγμα, θεωρούμε την ακόλουθη έκφραση ΚΛΠΤ, $(\forall x) \neg \text{likes}(x, \text{snow}) \Rightarrow \neg \text{skier}(x)$, όπου ο εκπαιδευόμενος έχει επιλέξει να εφαρμόσει τον κανόνα ισοδυναμίας *implication elimination* και η απάντηση του είναι, $(\forall x) \neg \text{likes}(x, \text{snow}) \wedge \neg \text{skier}(x)$. Στην συνέχεια, η απάντησή αυτή αναλύεται από το σύστημα και ανιχνεύεται ότι η συνεπαγωγή δεν αφαιρέθηκε σωστά και δημιουργούνται κατάλληλες υποδείξεις και επεξηγήσεις, όπως παρουσιάζεται στην εικόνα 2.29.



Εικόνα 2.29 Αλληλεπίδραση με το σύστημα FOL-EQU

Ως παράδειγμα θεωρούμε την ακόλουθη έκφραση ΚΛΠΤ:

$$(\forall x) \neg \text{likes}(x, \text{snow}) \Rightarrow \neg \text{skier}(x)$$

Στόχος είναι να εφαρμοστούν οι κατάλληλοι κανόνες ισοδυναμίας από τον εκπαιδευόμενο για να προκύψει η ισοδύναμη έκφραση

$$\neg (\exists x) (\neg \text{likes}(x, \text{snow}) \wedge \text{skier}(x))$$

Ένα παράδειγμα κανόνων που μπορούν να εφαρμοστούν για να καταλήξουμε στην έκφραση στόχος είναι οι εξής:

- (1) *Απαλοιφή Συνεπαγωγής*: $(\forall x) \neg \neg likes(x, snow) \vee \neg skier(x)$
- (2) *Νόμος της διπλής άρνησης*: $(\forall x) likes(x, snow) \vee \neg skier(x)$
- (3) *Νόμος De Morgan Ποσοδείκτες*: $\neg(\exists x) \neg(likes(x, snow) \vee \neg skier(x))$
- (4) *Νόμος De Morgan*: $\neg(\exists x)(\neg likes(x, snow) \wedge \neg skier(x))$
- (5) *Νόμος διπλής άρνησης*: $\neg(\exists x)(\neg likes(x, snow) \wedge skier(x))$

2.5.3.4 Πειραματικά Αξιολόγηση Συστήματος FOL-EQU

Για την αξιολόγηση του συστήματος πραγματοποιήσαμε μια πειραματική μελέτη κατά την διάρκεια του μαθήματος «Τεχνητή Νοημοσύνη». Στην πειραματική μελέτη συμμετείχαν 100 φοιτητές (άνδρες και γυναίκες). Αρχικά, οι φοιτητές χωρίστηκαν τυχαία σε δύο ομάδες των 50 φοιτητών η κάθε μια, την πειραματική ομάδα (experimental group) και την ομάδα ελέγχου (control group). Η πειραματική ομάδα χρησιμοποίησε το σύστημα και οι φοιτητές της ομάδας ασχολήθηκαν με ασκήσεις πάνω στις ισοδυναμίες, ενώ η ομάδα ελέγχου δεν χρησιμοποίησε το σύστημα, αλλά μελέτησε σχετικό υλικό, όπως διαλέξεις, παραδείγματα και ασχολήθηκε με ασκήσεις πάνω στις λογικές ισοδυναμίες.

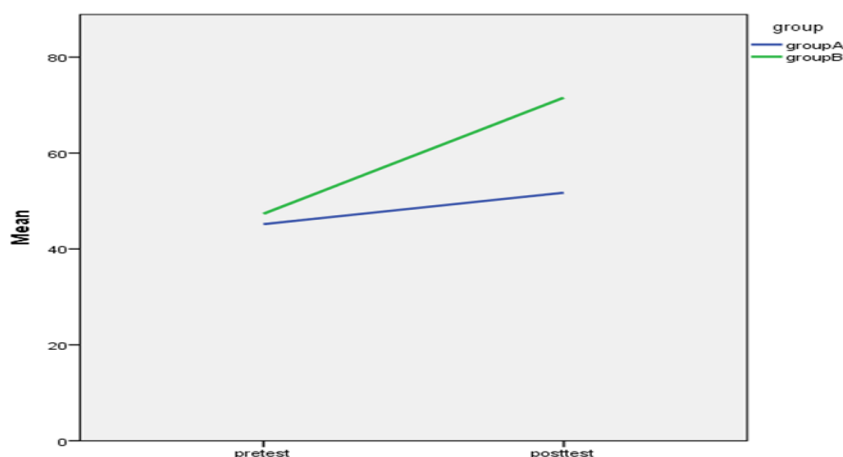
Αρχικά, όλοι οι φοιτητές απάντησαν σε ένα pretest πάνω στις λογικές ισοδυναμίες της ΚΛΠΤ. Το τεστ περιελάμβανε 15 εκφράσεις ΚΛΠΤ και οι φοιτητές κλήθηκαν να παρέχουν τις ισοδύναμες εκφράσεις σε 50 λεπτά. Στην συνέχεια, οι φοιτητές της πειραματικής ομάδας χρησιμοποίησαν το σύστημα FOL-EQU για να μελετήσουν για μια εβδομάδα, ενώ οι φοιτητές της ομάδας ελέγχου μελέτησαν με τον παραδοσιακό τρόπο για μια εβδομάδα και καθημερινά ενημέρωναν για την πρόοδο τους και για τις απορίες τους τον διδάσκοντα. Έπειτα, όλοι οι φοιτητές (και των δυο ομάδων) πήραν μέρος στο τελικό τεστ (posttest), το οποίο περιλάμβανε 15 εκφράσεις-ασκήσεις ΚΛΠΤ και είχε διάρκεια 50 λεπτά. Το pretest και το posttest περιελάμβαναν ασκήσεις που είχαν το ίδιο επίπεδο δυσκολίας και η κλίμακα της βαθμολογίας ήταν από 0 έως 100.

Για να αναλύσουμε την απόδοση, έγινε σύγκριση των μέσων τιμών των δυο ανεξάρτητων δειγμάτων *t-test* (*Independent Samples t-test*). Πιο συγκεκριμένα, η μέση τιμή (mean) και τυπική απόκλιση (standard deviation) για το pretest ήταν 45.18 και 14.72 για την πειραματική ομάδα και 47.34 και 14.01 αντίστοιχα για την ομάδα ελέγχου. Το *p-value* (επίπεδο σημαντικότητας) ήταν $0.567 > 0.05$ και $t = 0.46$, δηλαδή φαίνεται ότι δεν απορρίπτεται η μηδενική υπόθεση H_0 με επίπεδο σημαντικότητας $\alpha = 0.05$. Επομένως, από τα

αποτελέσματα φαίνεται ότι οι δύο ομάδες των φοιτητών είχαν το ίδιο επίπεδο γνώσεων πριν την διεξαγωγή του πειράματος και ότι οι δύο ομάδες των φοιτητών είχαν στατιστικά ισοδύναμες ικανότητες πριν από το πείραμα και επομένως είχαν περίπου το ίδιο επίπεδο γνώσης πριν την διεξαγωγή του πειράματος. Τα περιγραφικά στατιστικά μέτρα από την αξιολόγηση των επιδόσεων κατά την διαδικασία της μάθησης των φοιτητών παρουσιάζονται στον πίνακα 2.29. Στην εικόνα 2.30 παρουσιάζονται οι μέσες τιμές για τα pretest και posttest των δυο ομάδων. Τα αποτελέσματα δείχνουν ότι η μέση τιμή του pretest της πειραματικής ομάδας ήταν υψηλότερη από την μέση τιμή του pretest για την ομάδα ελέγχου. Ο στατιστικός έλεγχος *Levene's test* επιβεβαίωσε την ισότητα των διακυμάνσεων για τις ομάδες στο pretest ($F = 0.330$, $p = 0.567$) και από τα αποτελέσματα στο posttest ($F=3.016$, $p=0.00<0.05$) φαίνεται ότι υπάρχει μια σημαντική διαφορά μεταξύ των δύο ομάδων. Πιο συγκεκριμένα, οι φοιτητές είχαν κατανοήσει καλύτερα τους κανόνες ισοδυναμίας και δημιούργησαν σωστές ισοδύναμες εκφράσεις για περισσότερες εκφράσεις-ασκήσεις ΚΛΠΤ σε σχέση με την ομάδα ελέγχου.

Πίνακας 2.9 Περιγραφικά Μέτρα Pretest και Posttest του FOL-EQU

	Group	N	Mean	SD	SE
Pretest	Control Group	50	45.18	14.73	2.08
	Experimental Group	50	47.34	14.01	1.98
Posttest	Control Group	50	51.74	18.17	2.57
	Experimental Group	50	71.56	15.43	2.18



Εικόνα 2.30 Μέση τιμή για κάθε ομάδα για το Pretest και το Posttest

Στο δεύτερο μέρος της πειραματικής μελέτης, οι εκπαιδευόμενοι της πειραματικής ομάδας, που αλληλεπιδράσαν με το FOL-EQU, κλήθηκαν να συμπληρώσουν ένα ερωτηματολόγιο για να αξιολογήσουν διάφορες πτυχές του συστήματος. Πιο συγκεκριμένα, το ερωτηματολόγιο περιείχε ερωτήσεις για να αξιολογήσουν τις εκπαιδευτικές και τις λειτουργικές συνιστώσες του συστήματος, την ευχρηστία του και να διατυπώσουν την γνώμη τους και τις εμπειρίες τους με το σύστημα καθώς και τις απόψεις τους σχετικά με την επίδραση του συστήματος στην εκμάθηση των κανόνων ισοδυναμίας. Το ερωτηματολόγιο περιελάμβανε εννέα ερωτήσεις και τα αντίστοιχα αποτελέσματα παρουσιάζονται στον Πίνακα 2.10. Πιο συγκεκριμένα, οι ερωτήσεις Q1-Q6 βασίστηκαν στην κλίμακα Likert (1: καθόλου, 5: πάρα πολύ). Οι ερωτήσεις Q7-Q8 ήταν ερωτήσεις ανοικτού τύπου και αφορούσαν τα θετικά και τα αρνητικά σημεία του συστήματος, πιθανά προβλήματα που αντιμετώπισαν καθώς και βελτιώσεις που μπορούν να γίνουν. Επίσης, η ερώτηση 9 αναφερόταν στον χρόνο που χρειάστηκαν για να εξοικειωθούν με το σύστημα και είχε τρεις πιθανές απαντήσεις: α) λιγότερο από 15 λεπτά β) 15-30 λεπτά και γ) περισσότερο από 30 λεπτά. Οι απαντήσεις των εκπαιδευόμενων δείχνουν ότι το 72% των φοιτητών χρειάστηκε λιγότερο από 15 λεπτά και μόνο το 12% χρειάστηκε περισσότερο από 30 λεπτά.

Πίνακας 2.10 Απαντήσεις Ερωτηματολογίου για το σύστημα FOL-EQU

Q	Ερωτήσεις	Απαντήσεις (%)				
		1	2	3	4	5
1	Πως θα αξιολογούσατε την εμπειρία εκμάθησης με το σύστημα;	0	0	20	28	52
2	Πόσο σας βοήθησε το σύστημα να κατανοήσετε τις λογικές ισοδυναμίες;	0	0	18	32	50
3	Πόσο σας βοήθησε η ανάδραση που παρείχε το σύστημα;	0	4	12	36	48
4	Βρήκατε την διεπαφή του συστήματος εύκολη στην χρήση;	0	0	28	36	36
5	Πόσο η βοήθεια του συστήματος σας βοήθησε να αντιμετωπίσετε και να διορθώσετε τα λάθη σας;	0	2	14	34	50
6	Αποκτήσατε καλή αυτοπεποίθηση στην μετατροπή λογικών εκφράσεων σε ισοδύναμη μορφή;	0	4	16	38	42

Τα αποτελέσματα των απαντήσεων των φοιτητών στις ερωτήσεις Q1-Q6 δείχνουν ότι η πλειοψηφία βρήκε ευχάριστη την αλληλεπίδραση με το σύστημα και το 82% των φοιτητών

πιστεύουν ότι το σύστημα τους βοήθησε να μάθουν καλύτερα τους κανόνες ισοδυναμίας. Επίσης, το 84% των φοιτητών βρήκε την ανάδραση που παρέχεται από το σύστημα να είναι κατάλληλη και ότι τους βοήθησε στην εκμάθηση και τη δημιουργία ισοδύναμων εκφράσεων ΚΛΠΤ.

2.6 Σύστημα Αυτόματης Απόδειξης Θεωρημάτων

Στο τομέα της τεχνητής νοημοσύνης, η Αυτοματοποιημένη Απόδειξη θεωρημάτων (Automated Theorem Proving - ATP) είναι το πεδίο εκείνο που ασχολείται με την ανάπτυξη συστημάτων που αποδεικνύουν ότι μια υπόθεση (*conjecture*) είναι *λογική συνεπαγωγή* ενός συνόλου προτάσεων (*axioms* ή *hypotheses*). Τα συστήματα ATP χρησιμοποιούνται σε μία ευρεία περιοχή εφαρμογών όπως στα μαθηματικά, στην ανάπτυξη και επαλήθευση λογισμικού (software) και στα συστήματα που βασίζονται στη γνώση (Knowledge Based System) κ.ά.

Η γλώσσα στην οποία είναι γραμμένα τα αξιώματα, οι υποθέσεις και τα συμπεράσματα (γενικά γνωστά ως τύποι – formulae) είναι ένα είδος λογικής, συχνότερα η κλασσική ΚΛΠΤ, αλλά και μη κλασσικές λογικές καθώς και λογικές υψηλότερης τάξης. Αυτές οι γλώσσες επιτρέπουν την ακριβή αποτύπωση της απαραίτητης πληροφορίας σε προτάσεις, ώστε να μπορεί να την διαχειριστεί ένα σύστημα ATP. Η τυποποίηση αυτή αποτελεί τη δύναμη των συστημάτων ATP, καθώς δεν υπάρχει ασάφεια στη δήλωση μιας πρότασης, όπως συχνά συμβαίνει στη φυσική γλώσσα. Οι χρήστες χρειάζεται να περιγράψουν το πρόβλημα με ακρίβεια και αυτή η διαδικασία από μόνη της μπορεί να οδηγήσει σε μεγαλύτερη κατανόηση του προβλήματος και μέσω της αυτής μπορεί ο χρήστης να τυποποιήσει το πρόβλημα κατάλληλα για να υποβληθεί σε ένα σύστημα ATP.

Στην διεθνή βιβλιογραφία υπάρχουν αρκετά συστήματα ATP που στηρίζονται πάνω στην λογικής πρώτης τάξης, όπως τα SETHEO (Moser et al., 1997), SPASS, (Dahn et al., 1997), iProver (Korovin et al., 2008) και Vampire (Riazanov & Voronkov, 2002). Επίσης, υπάρχουν και αυτά που βασίζονται σε λογικές υψηλότερης τάξης, όπως είναι το HOL (Hurd, 1999 ; Gordon & Melham, 1993). Τα πεδία στα οποία έχουν χρησιμοποιηθεί επιτυχώς τα συστήματα ATP περιλαμβάνουν την Λογική, τα Μαθηματικά, την Επιστήμη των Υπολογιστών, την Μηχανική και τις Κοινωνικές Επιστήμες.

Τα συστήματα ATP είναι πάρα πολύ ισχυρά προγράμματα, ικανά να παράγουν κάθε πιθανό συμπέρασμα από μια αρχική βάση γνώσης. Εξ' αιτίας αυτού του εύρους, οι εφαρμογές και οι εκτελέσεις χρειάζεται πολλές φορές να καθοδηγούνται από έναν ειδικό στον

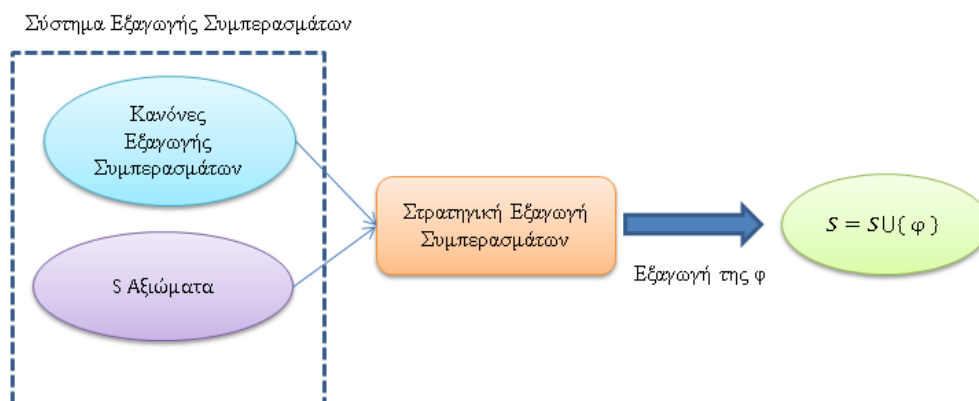
τομέα του προβλήματος, προκειμένου να λυθεί το πρόβλημα σε έναν εύλογο χρόνο. Επομένως, τα συστήματα ATP, παρά την ονομασία που τους έχει δοθεί, συχνά χρησιμοποιούνται από ειδικούς με αλληλεπιδραστικό τρόπο. Η αλληλεπίδραση μπορεί να είναι σε πολύ λεπτομερές επίπεδο, όπου ο χρήστης καθοδηγεί τις αναφορές που γίνονται από το σύστημα ή σε ένα πολύ υψηλότερο επίπεδο όπου ο χρήστης αποφασίζει ενδιάμεσα λήμματα που θα πρέπει να αποδειχθούν βαδίζοντας προς την απόδειξη του συμπεράσματος. Τα συστήματα ATP είναι επομένως μια τεχνολογία πολύ κατάλληλη σε περιπτώσεις όπου ένας ειδικός μπορεί να αλληλεπιδρά με ένα ισχυρό εργαλείο, ώστε να λύσει ενδιαφέροντα και δύσκολα προβλήματα του γνωστικού του πεδίου. Ένα σύστημα ATP είναι ικανό να παράγει σαν συμπεράσματα κάποια άγνωστα επακόλουθα της βάσης γνώσης, τα οποία να μην είναι καθόλου προφανή, ακόμη δηλαδή και να παράγει νέα γνώση.

Στη διεθνή βιβλιογραφία έχουν αρκετά συστήματα για ανάπτυξη τυπικών αποδείξεων (formal proofs) με τη συνεργασία ανθρώπου και συστήματος που ονομάζονται *αλληλεπιδραστικοί αποδείκτες θεωρημάτων* ή *εργαλεία υποστήριξης απόδειξης* (interactive theorem provers ή proof assistants) (Barendregt & Geuvers, 2001). Παραδείγματα εργαλείων υποστήριξης αποδείξεων σε ΚΛΠΤ καθώς και σε υψηλότερης τάξης λογικές, είναι Coq (Berramla et al., 2015), Matita (Asperti et al., 2007), Isabelle (Billingsley & Robinson, 2007) και ProofWeb (Kaliszyk et al., 2008).

Ωστόσο, δεν υπάρχουν εργαλεία για να βοηθήσουν τους εκπαιδευόμενους στην εκμάθηση και τους διδάσκοντες στη διδασκαλία της συλλογιστικής και της εξαγωγής συμπερασμάτων. Για να βοηθήσουμε τους φοιτητές να κατανοήσουν την αποδεικτική θεωρία και τις στρατηγικές ελέγχου, σχεδιάσαμε και αναπτύξαμε έναν αυτόματο αποδείκτη θεωρημάτων, τον EX-ACT-P (Grivokostopoulou & Hatzilygeroudis, 2010) ως επέκταση ενός υπάρχοντος αποδείκτη, του ACT-P (Hatzilygeroudis & Reichgelt, 1994), που βασίζεται στην λογική πρώτης τάξης και συνάγει με αυτόματο τρόπο από ένα σύνολο λογικών προτάσεων την αλήθεια μιας δεδομένης λογικής πρότασης, με την δυνατότητα χρήσης διαφορετικών στρατηγικών ελέγχου της διαδικασίας απόδειξης. Ο EX-ACT-P επιτρέπει τον συστηματικό πειραματισμό με διάφορες πτυχές της διαδικασίας απόδειξης του ACT-P. Επίσης, σχεδιάσαμε και αναπτύξαμε ένα αλληλεπιδραστικό αποδείκτη θεωρημάτων, τον i-ACTP, ως δεύτερη επέκταση του ACT-P, που έχει σαν στόχο να βοηθήσει τον εκπαιδευόμενο στην εκμάθηση των στρατηγικών ελέγχου και στην κατανόηση της διαδικασίας εξαγωγής συμπερασμάτων, αλληλεπιδρώντας με το σύστημα. Ο i-ACTP αναγνωρίζει τα λάθη του εκπαιδευόμενου και παρέχει κατάλληλη ανάδραση.

2.6.1 Βασικές Έννοιες Αποδεικτικής Θεωρίας

Μια βασική έννοια της αποδεικτικής θεωρίας των λογικών γλωσσών είναι η έννοια της *εξαγωγιμότητας* (*derivability*) σε σχέση με ένα σύνολο κανόνων εξαγωγής συμπερασμάτων. Μια πρόταση φ λέμε ότι είναι *εξαγωγή* (*derivable*) από ένα σύνολο υποθέσεων S αν το φ ανήκει στο S ή είναι αποτέλεσμα της εφαρμογής κάποιου κανόνα εξαγωγής συμπεράσματος σε εκφράσεις του S ή εξαγωγίμες από το S και συμβολίζεται ως $S \vdash \varphi$. Οι προτάσεις του S ονομάζονται *υποθέσεις* (*premises*) ή *αξιώματα* (*axioms*) και η πρόταση φ συμπέρασμα (*conclusion*) ή *θεώρημα* (*theorem*). Ορίζουμε ως *απόδειξη* (*proof*) μιας πρότασης φ από ένα σύνολο προτάσεων S μια ακολουθία προτάσεων, τέτοια ώστε η φ είναι η τελευταία πρόταση στην ακολουθία και κάθε πρόταση στην ακολουθία είναι είτε ένα αξίωμα, είτε μια πρόταση από το S , είτε μια πρόταση που παρήχθη από προηγούμενη (ες) πρόταση (εις) με την εφαρμογή κάποιου συμπερασματικού κανόνα. Μια πρόταση φ μπορεί να αποδειχθεί από ένα σύνολο προτάσεων S αν υπάρχει μια απόδειξη της φ από το S και συμβολίζεται ως $S \models \varphi$. Επίσης, η απόδειξη μιας πρότασης φ αναφέρεται ως *συνεπαγωγή* (*deduction*) ή *εξαγωγή* (*derivation*).



Εικόνα 2.31 Διαδικασία Απόδειξης μιας Πρότασης

Τα αξιώματα (υποθέσεις) και οι κανόνες εξαγωγής συμπερασμάτων αποτελούν ένα σύστημα *εξαγωγής συμπερασμάτων* (*inference system*) ή σύστημα συλλογισμού (*reasoning system*). Για την υλοποίηση της αποδεικτικής θεωρίας είναι απαραίτητη η ύπαρξη μιας στρατηγικής που θα ορίζει με ποιο τρόπο πρέπει να εφαρμόζονται οι κανόνες στις υποθέσεις για την εξαγωγή συμπεράσματος. Ο συνδυασμός μιας στρατηγικής και ενός συστήματος συλλογισμού συνιστούν μια *διαδικασία απόδειξης* (*proof procedure*). Στην εικόνα 2.31 παρουσιάζεται η διαδικασία της απόδειξης μιας πρότασης φ .

Η εγκυρότητα των συμπερασμάτων μιας διαδικασίας απόδειξης εξασφαλίζεται από την απαίτηση για λογική συνεπαγωγή των συμπερασμάτων από τις υποθέσεις. Η λογική συνεπαγωγή όμως εξαρτάται εξ' ολοκλήρου από τους κανόνες εξαγωγής συμπερασμάτων. Κανόνες που εξασφαλίζουν αυτό ονομάζονται *ορθοί* (*sound*) και η αντίστοιχη διαδικασία ορθή. Μια διαδικασία εξαγωγής συμπερασμάτων λέγεται *ορθή* (*sound*) αν και μόνο αν κάθε έκφραση που μπορεί να εξαχθεί από ένα σύνολο προτάσεων S με βάση αυτή τη διαδικασία συνεπάγεται λογικά από το S και αναπαρίσταται ως εξής:

$$\forall S, \forall \varphi, S \models \varphi \text{ ανν } S \vdash \varphi$$

Η ορθότητα είναι ένα χαρακτηριστικό που καθιστά μια διαδικασία εξαγωγής συμπερασμάτων χρήσιμη, αφού αποτρέπει την εξαγωγή λανθασμένων συμπερασμάτων από λογικά ευσταθείς υποθέσεις.

Επίσης, μια άλλη απαίτηση είναι να μπορούμε να εξάγουμε όλα τα εξαγωγήματα συμπεράσματα που εξασφαλίζουν αυτό. Οι αντίστοιχες διαδικασίες ονομάζονται *πλήρεις*. Μια διαδικασία εξαγωγής συμπεράσματος λέγεται *πλήρης* (*complete*), αν και μόνο αν κάθε έκφραση που συνεπάγεται λογικά από το σύνολο S είναι εξαγωγήμα από το S με βάση αυτή τη διαδικασία.

Η πληρότητα εξασφαλίζει ότι η συγκεκριμένη διαδικασία εξαγωγής συμπεράσματος θα έχει σαν αποτέλεσμα την εξαγωγή όλων των εκφράσεων που μπορούν να προκύψουν από τις αρχικές υποθέσεις, μηδενίζοντας την πιθανότητα παράλειψης κάποιου χρήσιμου συμπεράσματος.

Ο ACT-P χρησιμοποιεί ως κανόνα εξαγωγής συμπερασμάτων την αρχή της επίλυσης (*resolution principle*). Ο κανόνας αυτός παράγει μια νέα πρόταση από δύο υπάρχουσες προτάσεις και εξασφαλίζει ορθότητα και πληρότητα (μαζί με τον απλό κανόνα της παραγοντοποίησης, ο οποίος μετασχηματίζει μια πρόταση και χρειάζεται σε ορισμένες περιπτώσεις). Ο κανόνας της επίλυσης χρησιμοποιείται σε συνδυασμό με τη διαδικασία της αντίφασης της επίλυσης (*resolution refutation*) για την εξαγωγή συμπερασμάτων (θεωρημάτων) από ένα σύνολο προτάσεων (αξιώματα).

Ο σχεδιασμός του ACT-P έχει γίνει με βάση την αρχιτεκτονική των meta-level συστημάτων (Hatzilygeroudis & Reichgelt, 1994). Τα συστήματα αυτά αποτελούνται από δύο επίπεδα, το επίπεδο του αντικειμένου (*object-level*), που αφορά τη γνώση πεδίου και την εξαγωγή συμπερασμάτων από αυτή, και το μετα-επίπεδο (*meta-level*), το οποίο αφορά τον έλεγχο της διαδικασίας εξαγωγής συμπερασμάτων. Πιο συγκεκριμένα, υπάρχει ένας σκελετός βασισμένος στην αρχή της επίλυσης και τη διαδικασία της αντίφασης της επίλυσης, στον

οποίο συγκεκριμένα βήματα του προγραμματιστικού κύκλου επιτρέπεται να προγραμματίζονται από τον χρήστη, με τον επαναπρογραμματισμό συγκεκριμένων συναρτήσεων, που ελέγχουν τη διαδικασία παραγωγής νέων προτάσεων.

2.6.2 Στρατηγικές Ελέγχου Εξαγωγής Συμπερασμάτων

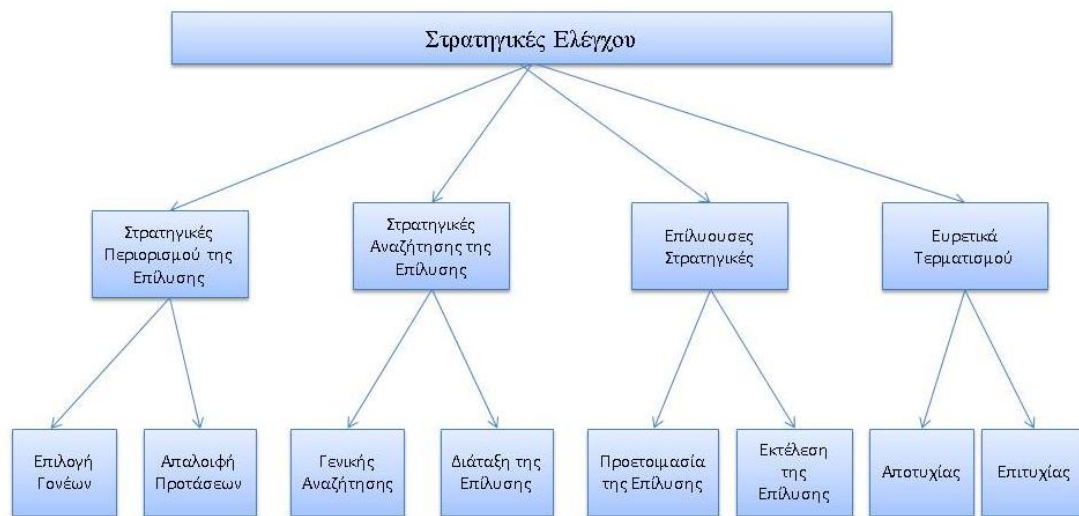
Οι στρατηγικές ελέγχου της διαδικασίας εξαγωγής συμπερασμάτων έχουν σκοπό την ελαχιστοποίηση των άχρηστων συμπερασμάτων και συνεπώς τη μείωση του υπολογιστικού χρόνου. Οι στρατηγικές ελέγχου της αντίφασης της επίλυσης μπορούν να κατηγοριοποιηθούν σε τέσσερις κύριες κατηγορίες (Hatzilygeroudis & Reichgelt, 1994): *στρατηγικές περιορισμού επίλυσης, στρατηγικές αναζήτησης επίλυσης, επιλύουσες στρατηγικές και ευρετικά τερματισμού*. Οι κατηγορίες αυτές και οι υποκατηγορίες τους παρουσιάζονται στην εικόνα 2.32.

Οι *στρατηγικές περιορισμού επίλυσης* σχετίζονται με τη δημιουργία του χώρου αναζήτησης και χρησιμοποιούνται για να αυξήσουν την απόδοση περιορίζοντας το μέγεθος του χώρου, είτε διαγράφοντας συγκεκριμένες προτάσεις, είτε καθορίζονται αυστηρά τα είδη των προτάσεων που επιτρέπεται να επιλυθούν μεταξύ τους και διακρίνονται στις εξής υποκατηγορίες: *στρατηγικές επιλογής γονέων και στρατηγικές απαλοιφής προτάσεων*.

Οι *στρατηγικές αναζήτησης επίλυσης* σχετίζονται με τον τρόπο που εξετάζεται ο χώρος αναζήτησης και διακρίνονται σε *στρατηγικές γενικής αναζήτησης* και *στρατηγικές διάταξης επίλυσης*.

Οι *επιλύουσες στρατηγικές* καθορίζουν τον τρόπο εφαρμογής της αρχής της επίλυσης στις προτάσεις που επιλύονται, χρησιμοποιούνται κατά κύριο λόγο σε συνδυασμό με μια στρατηγική επιλογής γονέων και διακρίνονται σε *στρατηγικές προετοιμασίας επίλυσης* και *στρατηγικές εκτέλεσης επίλυσης*.

Τα *ευρετικά τερματισμού* καθορίζουν πότε θα πρέπει να σταματήσει η αναζήτηση για την εύρεση μιας απόδειξης και διακρίνονται σε μεθόδους που αναγνωρίζουν την *αποτυχία* και την *επιτυχία*.



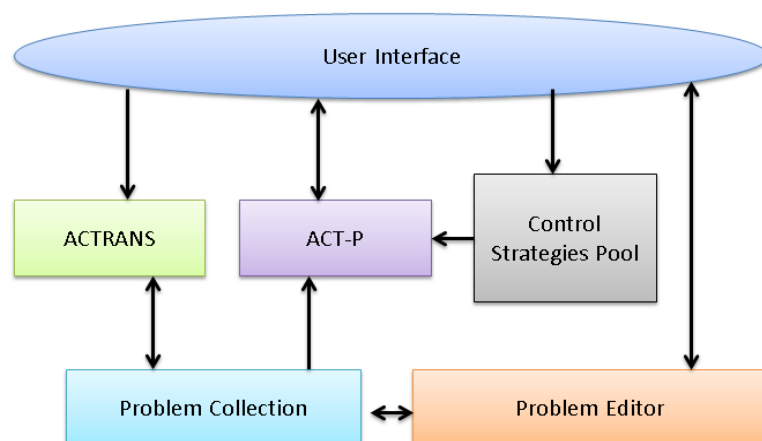
Εικόνα 2.32 Κατηγοριοποίηση Στρατηγικών Ελέγχου

2.6.3 Αποδείκτης θεωρημάτων EX-ACT-P

Ο EX-ACT-P είναι ένα σύστημα αυτόματης απόδειξης θεωρημάτων βασισμένο στη λογική πρώτης τάξης. Ο κύριος μηχανισμός λειτουργίας του EX-ACT-P (Extended-A Configurable Theorem-Prover) βασίζεται στον ACT-P (A Configurable Theorem-Prover).

2.6.3.1 Αρχιτεκτονική Αποδείκτη Θεωρημάτων

Η αρχιτεκτονική του αποδείκτη θεωρημάτων EX-ACT-P, που παρουσιάζεται στην εικόνα 2.33 αποτελείται από έξι κύριες συνιστώσες: *Διεπαφή χρήση (User interface)*, *Αυτόματος Μετασχηματιστής Κώδικα (ACTTRANS)*, *Επεξεργαστής Προβλημάτων (Problem editor)*, *ACT-P*, *Συλλογή Προβλημάτων (Problem Collection)*, *Δεξαμενή Στρατηγικών Ελέγχου (Control Strategies Pool)*.



Εικόνα 2.33. Αρχιτεκτονική αποδείκτη θεωρημάτων EX-ACT-P

Ο *Επεξεργαστής Προβλημάτων (Problem Editor)* χρησιμοποιείται για να κωδικοποιεί ένα πρόβλημα στη γλώσσα του αποδείκτη θεωρημάτων και παρέχει την δυνατότητα στον χρήστη να δημιουργεί το δικό του πρόβλημα και στην συνέχεια το πρόβλημα να προστίθεται στη συλλογή προβλημάτων και να είναι διαθέσιμο για πειραματισμό. Ο *ACTRANS (Automated Code TRANSformer)* χρησιμοποιείται για τον μετασχηματισμό των προβλημάτων της βιβλιοθήκης TPTP (Sutcliffe & Suttner, 1998), που περιέχει προβλήματα ειδικά για συστήματα ATP, στη γλώσσα του αποδείκτη θεωρημάτων. Στη συνέχεια, τα προβλήματα μπορούν να αποθηκεύουν και να είναι διαθέσιμα για επίλυση (απόδειξη). Επίσης, ο χρήστης μέσω της μονάδας *δεξαμενή στρατηγικών ελέγχου* μπορεί να επιλέξει ένα κατάλληλο συνδυασμό στρατηγικών ελέγχου που διαθέτει ο αποδείκτης και να αποτελέσει την γενική στρατηγική επίλυσης. Η μονάδα *ACT-P* αποτελεί την κύρια μονάδα για την παραγωγή αποδείξεων. Επίσης, η διεπαφή του συστήματος αποτελεί το περιβάλλον αλληλεπίδρασης του χρήστη με τον αποδείκτη.

2.6.3.2 Ο Μηχανισμός του ACT-P

Η κατηγοριοποίηση των στρατηγικών ελέγχου που παρουσιάστηκε προηγουμένως αποτελεί τη βάση για τον προγραμματιστικό κύκλο του ACT-P. Ο κύριος μηχανισμός του ACT-P, αυτός της απόδειξης θεωρημάτων, χωρίζεται σε δυο φάσεις. Στην πρώτη, οι λογικοί τύποι που κωδικοποιούν την γνώση του προβλήματος εισάγονται στο σύστημα. Οι λογικοί τύποι μπορεί να βρίσκονται σε μορφή ΚΛΠΤ ή σε μορφή οποιουδήποτε βήματος μετατροπής ανάμεσα σε αυτή και την προτασιακή μορφή. Αυτόματα μετατρέπονται σε προτασιακή μορφή, για να αποθηκευτούν στην βάση γνώσης ως αξιώματα. Η δεύτερη φάση της

εκτέλεσης του ACT-P, χωρίζεται με τη σειρά της σε δυο στάδια. Στη διάρκειά τους, η βάση γνώσης εμπλουτίζεται σταδιακά με τις προτάσεις που παράγονται. Το πρώτο από τα προαναφερθέντα στάδια είναι αυτό της *προετοιμασίας* που εκτελείται μόνο μία φορά. Ακολουθείται από αλληπάλληλες επαναλήψεις του δευτέρου σταδίου (γνωστό ως *κύριος κύκλος απόδειξης*), μέχρι να βρεθεί λύση (παραγωγή κενής πρότασης) ή να προκύψει αδιέξοδο (αδυναμία παραγωγής νέων προτάσεων). Σε πολλές στρατηγικές, ο διαχωρισμός μεταξύ αρχικοποίησης και επανάληψης δεν είναι καθόλου εύκολος.

Ένας κρίσιμος παράγοντας, που καθορίζει την ταχύτητα του ACT-P, είναι ο χρόνος εύρεσης των προτάσεων που μπορούν να επιλυθούν με μια δεδομένη πρόταση. Για κάθε πρόταση-αριστερό γονέα και για κάθε ένα από τα στοιχεία της τα οποία επιτρέπεται να επιλυθούν, χρειάζεται να βρεθούν όλες οι προτάσεις-δεξιοί γονείς που περιέχουν στοιχεία τα οποία επιλύονται με τα αντίστοιχα του εν λόγω αριστερού γονέα. Αυτή είναι μια υπολογιστικά χρονοβόρα διαδικασία. Η λύση στο πρόβλημα αυτό είναι η διατήρηση μιας δομής δεδομένων αποκλειστικά για την αποθήκευση των στοιχείων που εμφανίζονται σε προτάσεις. Μ' αυτόν τον τρόπο, όλες οι εμφανίσεις στοιχείου σε διάφορες προτάσεις βρίσκονται συγκεντρωμένες. Η συγκεκριμένη δομή δεδομένων που επιλέχθηκε για τον αποδείκτη θεωρημάτων είναι το δέντρο διακρίσεων και η μέθοδος δεικτοδότησης των στοιχείων που το χρησιμοποιεί ονομάζεται *discrimination-tree indexing* (McCune, 1992). Το δέντρο διακρίσεων μοιάζει με *tree*, στο οποίο εισάγονται μόνο οι συνδυασμοί συμβόλων που εμφανίζονται στην πράξη, ενώ οι υπόλοιποι αγνοούνται.

2.6.3.3 Στρατηγικές Ελέγχου EX-ACT-P

Η παραμετροποίηση της αποδεικτικής διαδικασίας υλοποιήθηκε μέσω των στρατηγικών ελέγχου. Οι στρατηγικές που υλοποιήθηκαν χωρίστηκαν σε 5 βασικές κατηγορίες.

Η κατηγορία *επιλογής γονέων* (*parent selection*) περιλαμβάνει στρατηγικές για την επιλογή των προτάσεων που επιτρέπεται να ενοποιηθούν:

Unit Resolution: Στη στρατηγική αυτή πρέπει τουλάχιστον ένας απ' τους γονείς να είναι μια μοναδιαία πρόταση.

P1 Resolution: Στη στρατηγική αυτή ακριβώς ένας από τους δύο γονείς πρέπει να είναι θετική πρόταση, δηλαδή να μην περιέχει κανένα στοιχείο με άρνηση.

NI Resolution: Η στρατηγική αυτή είναι η συμπληρωματική της P1 resolution, αφού θέτει τον περιορισμό ακριβώς ένας από τους γονείς να είναι αρνητική πρόταση, δηλαδή όλα τα στοιχεία της να έχουν άρνηση.

Hyperresolution: Η στρατηγική αυτή είναι μια πιο αποδοτική εκδοχή της P1 resolution. Ουσιαστικά γίνεται ταυτόχρονη επίλυση πολλών ατόμων μιας πρότασης.

Unit-Resulting Resolution: Η στρατηγική αυτή προσπαθεί να επιλύσει περισσότερες από δύο προτάσεις ταυτόχρονα και να οδηγήσει στην παραγωγή μοναδιαίων προτάσεων.

Input Resolution: Στη στρατηγική αυτή πρέπει μία απ' τις προτάσεις - γονείς να ανήκει στο σύνολο των αρχικών προτάσεων, δηλαδή των αξιωμάτων.

Linear Resolution: Η στρατηγική αυτή είναι μια γενίκευση της input resolution.

Linear Input Resolution: Είναι μια παραλλαγή της επίλυσης εισόδου (linear resolution), όπου ο δεξιός γονέας πρέπει να ανήκει στις αρχικές υποθέσεις, ενώ ο αριστερός είναι η πιο πρόσφατη επιλύουσα. Η παραλλαγή αυτή της γραμμικής επίλυσης είναι πλήρης μόνο για προτάσεις τύπου Horn.

Οι *στρατηγικές απαλοιφής (elimination strategies)* στοχεύουν στην εύρεση και διαγραφή ή αντικατάσταση προτάσεων που αποτελούν ταυτολογία ή συνοψίζονται από άλλες και περιλαμβάνει τις ακόλουθες στρατηγικές:

Tautology Elimination: Ταυτολογία (tautology) ονομάζεται μια πρόταση που περιέχει ένα ζευγάρι από συμπληρωματικά στοιχεία. Οι ταυτολογίες μπορούν να διαγραφούν από ένα σύνολο προτάσεων, γιατί δεν επηρεάζουν καθόλου το αν το σύνολο είναι ικανοποιήσιμο ή όχι.

Pure-literal Elimination: Ένα στοιχείο του οποίου το συμπλήρωμα δεν εμφανίζεται σε καμία πρόταση του συνόλου των αρχικών προτάσεων καλείται καθαρό (pure). Μια πρόταση που περιέχει ένα καθαρό στοιχείο είναι άχρηστη στην πορεία της εξαγωγής της κενής πρότασης Έτσι, μπορούμε να τη διαγράψουμε, μειώνοντας το πλήθος των δυνατών επιλύσεων και, επομένως, τον υπολογιστικό χρόνο.

Subsumption Elimination: Μια πρόταση C συνοψίζει (subsumes) μια πρόταση D αν και μόνο αν υπάρχει ένα σύνολο αντικαταστάσεων σ τέτοιο ώστε η εφαρμογή του στη

$C: C_S \subseteq D$. Έχει αποδειχθεί ότι οι προτάσεις που συνοψίζονται μπορούν να διαγραφούν χωρίς να δημιουργηθεί κανένα πρόβλημα. Υπάρχουν δύο τύποι σύνοψης κατά τη διάρκεια μιας απόδειξης: όταν μια νέα πρόταση συνοψίζεται από μια ήδη υπάρχουσα τότε λέμε ότι έχουμε *forward subsumption*, ενώ όταν μια νέα πρόταση συνοψίζει μια παλιά έχουμε *backward subsumption*.

Οι *στρατηγικές διάταξης επίλυσης* (*resolution ordering strategies*) περιλαμβάνουν τις εξής στρατηγικές:

Fewest Literal Preference: Σαν κριτήριο για τη διάταξη των υποψήφιων για επίλυση προτάσεων λαμβάνεται το μήκος τους, δηλαδή, το πλήθος των στοιχείων που περιέχουν. Όσο μικρότερο είναι το μήκος μιας πρότασης, τόσο καλύτερη θεωρείται.

Unit preference: Στη στρατηγική αυτή κριτήριο για τη διάταξη είναι το αν μια πρόταση που παράγεται είναι μοναδιαία ή όχι. Αυτό είναι λογικό, αφού η επίλυση των μοναδιαίων προτάσεων θα οδηγήσει (πιθανόν) γρηγορότερα στην εξαγωγή της κενής πρότασης.:-

Τέλος, οι *στρατηγικές εκτέλεσης επίλυσης* (*resolution execution strategies*) περιλαμβάνουν τις ακόλουθες στρατηγικές:

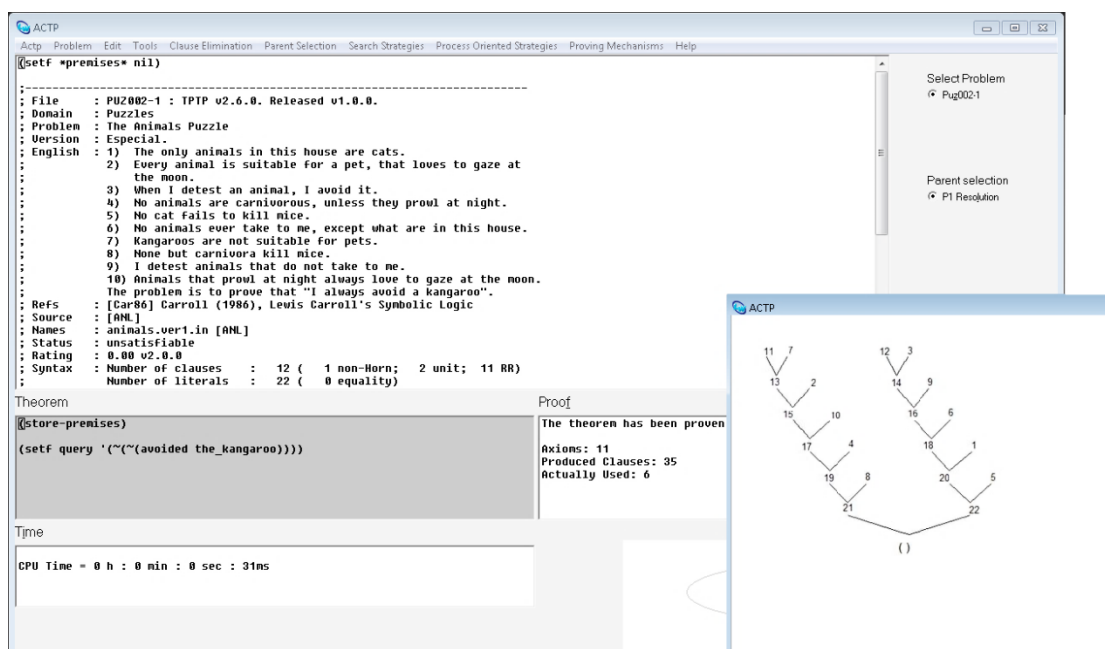
Ordered Resolution: Στη στρατηγική αυτή χρησιμοποιείται η έννοια της *διατεταγμένης πρότασης* (*ordered clause*), σύμφωνα με την οποία μια πρόταση θεωρείται σαν μια ακολουθία από στοιχεία και όχι σαν ένα σύνολο στοιχείων.

OI- Resolution: Η στρατηγική αυτή χρησιμοποιεί την έννοια της *διατεταγμένης σημασιολογικής σύγκρουσης* (*ordered semantic clash*).

2.6.3.4 Αλληλεπίδραση με τον EX-ACT-P

Αρχικά, οι χρήστες μέσω της διεπαφής μπορούν-να δημιουργήσουν ένα πρόβλημα, δηλαδή ένα σύνολο προτάσεων ΚΛΠΤ που περιγράφουν ένα πρόβλημα, μέσω του 'επεξεργαστή προβλημάτων' και στη συνέχεια επιλέγοντας την κατάλληλη στρατηγική ή συνδυασμό στρατηγικών να πραγματοποιήσουν επίλυση του προβλήματος. Επίσης, παρέχεται η δυνατότητα σ' ένα χρήστη να επιλέξει ένα πρόβλημα για επίλυση από την βιβλιοθήκη TPTP και στην συνέχεια μέσω του ACTTRANS να γίνει η μετάφραση στη γλώσσα του αποδείκτη θεωρημάτων.

Κατόπιν, ο χρήστης εισάγει το προς απόδειξη θεώρημα, ενεργοποιείται ο ACT-P και παράγεται αυτόματα μια απόδειξη του προβλήματος. Μετά την απόδειξη, ο χρήστης έχει την δυνατότητα να δει όλες τις απαραίτητες πληροφορίες για την επίλυση του συγκεκριμένου προβλήματος, όπως τον αριθμό των αξιωμάτων του προβλήματος, τον αριθμό των προτάσεων που χρησιμοποιήθηκαν κατά την διαδικασία της επίλυσης, τον χρόνο επίλυσης κ. α. Επίσης, ο χρήστης μπορεί να επιλέξει να δει βήμα προς βήμα τη διαδικασία απόδειξης, δίνοντας την δυνατότητα μέσω γραφικής απεικόνισης να δει τις προτάσεις που συνδυάστηκαν για να γίνει επίλυση του προβλήματος. Στην εικόνα 2.34, παρουσιάζεται η επίλυση ενός προβλήματος καθώς και η γραφική αναπαράσταση που αναπαριστά των συνδυασμό των προτάσεων που χρειάστηκαν για την επίλυση του προβλήματος.



Εικόνα 2.34 Αποδείκτης Θεωρημάτων EX-ACT-P

Επομένως, μέσω του EX-ACTP, οι εκπαιδευόμενοι μπορούν να πειραματιστούν με νέες πιθανές ασκήσεις-προβλήματα, που σχετίζονται με την αναπαράσταση γνώσης και συλλογιστική και να δοκιμάσουν διάφορους συνδυασμούς στρατηγικών επίλυσης. Επίσης, οι εκπαιδευόμενοι μπορούν να εξασκηθούν στο να επιλύσουν το ίδιο πρόβλημα με διαφορετικές στρατηγικές και να συγκρίνουν τα βήματα της απόδειξης, να δουν τις διαφορές μεταξύ των στρατηγικών επίλυσης και των βημάτων μέσω γραφικής απεικόνισης. Επίσης, παρέχεται η δυνατότητα να επιλύσουν άμεσα σύνθετα προβλήματα που δεν μπορούν να επιλυθούν στο χαρτί, λόγω μεγάλου αριθμού προτάσεων που απαιτούνται για την επίλυση.

2.6.4 Αλληλεπιδραστικός Αποδείκτης Θεωρημάτων

Ο αλληλεπιδραστικός αποδείκτης θεωρημάτων i-ACTP (interactive ACT-P) βασίζεται στον ACT-P για να πραγματοποιεί αποδείξεις θεωρημάτων με αλληλεπιδραστικό τρόπο. Ο i-ACTP έχει ως στόχο να αποτελέσει ένα χρήσιμο εργαλείο για τον διδάσκοντα στην διδασκαλία της εξαγωγής συμπερασμάτων και να βοηθήσει τους εκπαιδευόμενους στην κατανόηση των βασικών πτυχών της απόδειξης θεωρημάτων. Πιο συγκεκριμένα, ο αλληλεπιδραστικός αποδείκτης θεωρημάτων έχει ως στόχο να βοηθήσει στην κατανόηση των διαφόρων στρατηγικών επίλυσης και του τρόπου λειτουργίας της διαδικασίας της αντίφασης της επίλυσης, παρέχοντας ανάδραση (feedback) κατά την διάρκεια της αλληλεπίδρασης. Μια σημαντική διαφορά του αυτόματου αποδείκτη θεωρημάτων από τον αλληλεπιδραστικό είναι ότι ο πρώτος κατά την διαδικασία της επίλυσης επιλέγει τις προς επίλυση προτάσεις με βάση κάποια κριτήρια που έχουν προκαθοριστεί. Ο χρήστης δεν ακολουθεί ένα συγκεκριμένο τρόπο επιλογής, όπως ένα αυτόματο σύστημα απόδειξης θεωρημάτων, αλλά έχει ελευθερία επιλογής με βάση δικά του κριτήρια. Οπότε, ο αλληλεπιδραστικός αποδείκτης δεν εξάγει μια συγκεκριμένη λύση για την επίλυση ενός προβλήματος, αλλά με βάση τις επιλογές του χρήστη σε κάθε βήμα μπορεί προκύπτουν διαφορετικές λύσεις.

2.6.4.1 Μηχανισμός Αλληλεπιδραστικού Αποδείκτη Θεωρημάτων

Ο μηχανισμός του αλληλεπιδραστικού αποδείκτη βασίστηκε στις βασικές αρχές του αυτόματου αποδείκτη θεωρημάτων ACT-P (Grivokostopoulou & Hatzilygeroudis, 2010). Η αρχικοποίηση γίνεται με τον ίδιο τρόπο, δηλαδή ο κύριος προγραμματιστικός κύκλος και ο έλεγχος για τον τερματισμό είναι ίδιοι. Η κύρια διαφορά βρίσκεται στην επιλογή του βήματος που θα εκτελεστεί σε κάθε κύκλο. Πιο συγκεκριμένα, στον αυτόματο αποδείκτη η επιλογή γίνεται με βάση τη διατεταγμένη λίστα agenda, ενώ στον αλληλεπιδραστικό προσδιορίζεται από τον χρήστη. Η agenda είναι μια δομή που περιέχει τα βήματα που μπορεί να εκτελέσει ο αποδείκτης σε μια συγκεκριμένη χρονική στιγμή. Επίσης, ο μηχανισμός πραγματοποιεί έλεγχο της επιλογής του χρήστη. Στην περίπτωση που η επιλογή είναι επιτρεπτή με βάση την στρατηγική έλεγχου, τότε ο μηχανισμός πραγματοποιεί το βήμα και προωθείται στον επόμενο κύκλο, ενώ σε διαφορετική περίπτωση παρέχεται στον χρήστη κατάλληλη ανάδραση.

Δημιουργείται ένας κύκλος που εκτελείται σε βήματα και όταν ολοκληρωθούν όλοι οι απαιτούμενοι κύκλοι, ο αποδείκτης θεωρημάτων σταματάει τη διαδικασία.

Ο κάθε κύκλος αποτελείται από τα ακόλουθα βήματα:

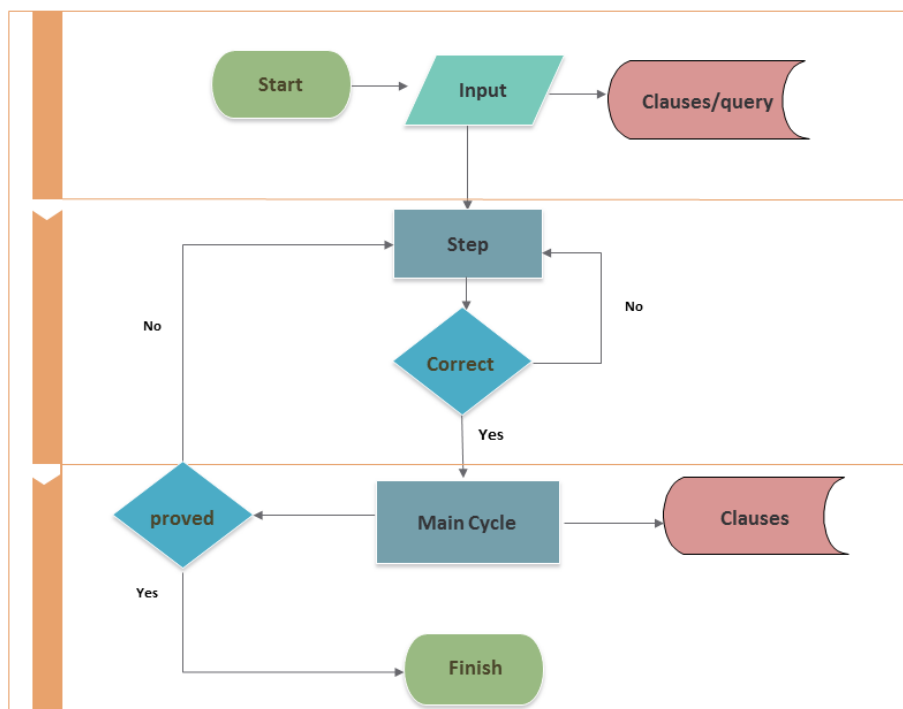
1. Επιλογή των προτάσεων που θα συμμετέχουν στον τρέχοντα κύκλο.

2. Έλεγχος της τρέχουσας επιλογής:

2.1 Αν η επιλογή του χρήστη είναι λανθασμένη, ενημερώνεται ο χρήστης και προσφέρεται το κατάλληλο είδος ανάδρασης.

2.2 Αν η επιλογή είναι σωστή, τότε γίνεται επίλυση των προτάσεων και επιστρέφει στο βήμα 1

Στην εικόνα 2.35 παρουσιάζεται ο μηχανισμός του αλληλεπιδραστικού αποδείκτη θεωρημάτων.



Εικόνα 2.35 Μηχανισμός Αλληλεπίδρασης i-ACT-P

2.6.4.2 Παροχή Ανάδρασης

Στον αλληλεπιδραστικό αποδείκτη θεωρημάτων, ο χρήστης προσδιορίζει τις προτάσεις που θα συνδυαστούν και σε περίπτωση λανθασμένων επιλογών το σύστημα ενημερώνει τον χρήστη και παρέχει κατάλληλη ανάδραση. Πιο συγκεκριμένα, όταν οι προτάσεις που επιθυμεί να συνδυάσει είναι λανθασμένες με βάση τους κανόνες της στρατηγικής επίλυσης, παρέχεται κατάλληλη ανάδραση και ενημερώνεται ο χρήστης για την λανθασμένη επιλογή του.

Στον Πίνακα 2.11 παρουσιάζονται οι βασικές στρατηγικές ελέγχου και οι αντίστοιχες κατηγορίες λαθών που αναγνωρίζονται από τον αποδείκτη.

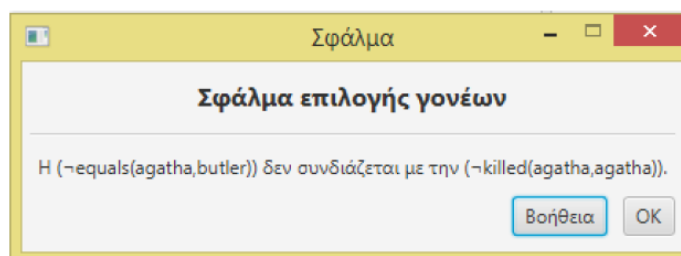
Πίνακας 2.11 Κατηγορίες λαθών

Στρατηγική Ελέγχου	Κατηγορία Λαθών.
Στρατηγικές απαλοιφής προτάσεων	Λάθος διαγραμμένη πρόταση
Στρατηγικές επιλογής γονέων	Λάθος επιλογή γονέων
Στρατηγικές Αναζήτησης	Λάθος Αναζήτησης (κατά βάθος ή πλάτος)

Πιο συγκεκριμένα, με βάση της κατηγορίες λαθών, αναγνωρίζονται τα λάθη και παρέχονται τα κατάλληλα είδη ανάδρασης:

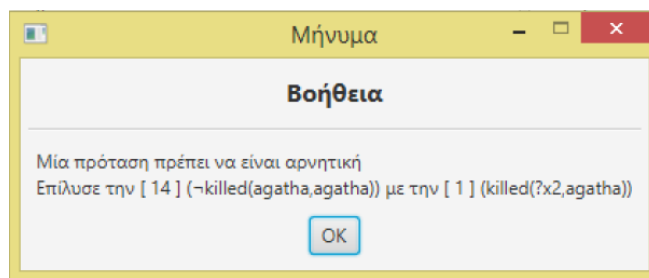
- 1) *Error-specific feedback* : αναλύονται και παρέχονται υποδείξεις για τα λάθη.
- 2) *Procedural feedback*: παρέχονται υποδείξεις για το πώς να ενεργήσει ο εκπαιδευόμενος για την διεκπεραίωση του τρέχοντος βήματος

Το πρώτο είδος ανάδρασης που παρέχει ο i-ACTP είναι το *error-specific feedback*, όπου ενημερώνεται ο εκπαιδευόμενος για την λανθασμένη επιλογή και του παρέχεται το κατάλληλο μήνυμα. Για παράδειγμα, στην περίπτωση που ο χρήστης επιλέξει για επίλυση δυο προτάσεις που δεν ενοποιούνται, το σύστημα ενημερώνει τον χρήστη για την λάθος επιλογή και του παρέχει σχετικό μήνυμα, όπως παρουσιάζεται στη εικόνα 2.36.



Εικόνα 2.36 Πρώτο είδος βοήθειας Error-specific feedback

Το δεύτερο είδος βοήθειας που παρέχει ο i-ACTP είναι το *procedural feedback*, όπου κατά την διαδικασία επίλυσης ενός προβλήματος, ενημερώνει τον χρήστη ότι υπάρχει κάποιο λάθος στις επιλογές του και του παρέχεται η επιλογή να ζητήσει βοήθεια. Η βοήθεια που παρέχει το σύστημα περιλαμβάνει μια σύντομη περιγραφή της σωστής επίλυσης, όπως για παράδειγμα μια περιγραφή της σωστής χρήσης της στρατηγικής που παραβιάστηκε, καθώς και υποδείξεις για τη σωστή εφαρμογή της στρατηγικής. Στην εικόνα 2.37, παρουσιάζεται η βοήθεια που δίνεται από τον αποδείκτη σε λανθασμένη επιλογή του χρήστη για την N1 επίλυση.



Εικόνα 2.37 Δεύτερο είδος βοήθειας Procedural feedback

2.6.4.3 Αλληλεπίδραση με το Χρήστη

Ο χρήστης μέσω της διεπαφής τους συστήματος μπορεί να εισάγει ένα πρόβλημα εκφρασμένο σε ΚΛΠΤ. Επίσης, μπορεί να επιλέξει να χρησιμοποιήσει ένα πρόβλημα από την βιβλιοθήκη TPTP. Τότε, ο αποδείκτης, μέσω του συστήματος μετατροπής ΚΛΠΤ σε ΠΜ, μετατρέπει αυτόματα όλες τις προτάσεις σε ΠΜ. Στην συνέχεια, μέσω του μεταφραστή ACTRANS, μετατρέπει τις προτάσεις στην γλώσσα του αποδείκτη.

Στην συνέχεια παρουσιάζεται ένα πρόβλημα από τη βιβλιοθήκη TPTP, εκφρασμένο σε ΚΛΠΤ.

Φυσική Γλώσσα:

“Someone who lives in Dreadbury Mansion killed Aunt Agatha. Agatha, the butler, and Charles live in Dreadbury Mansion, and are the only people who live therein. A killer always hates his victim, and is never richer than his victim. Charles hates no one that Aunt Agatha hates. Agatha hates everyone except the butler. The butler hates everyone not richer than Aunt Agatha. The butler hates everyone Aunt Agatha hates. No one hates everyone. Agatha is not the butler.”

Ερώτηση: *Agatha killed herself*

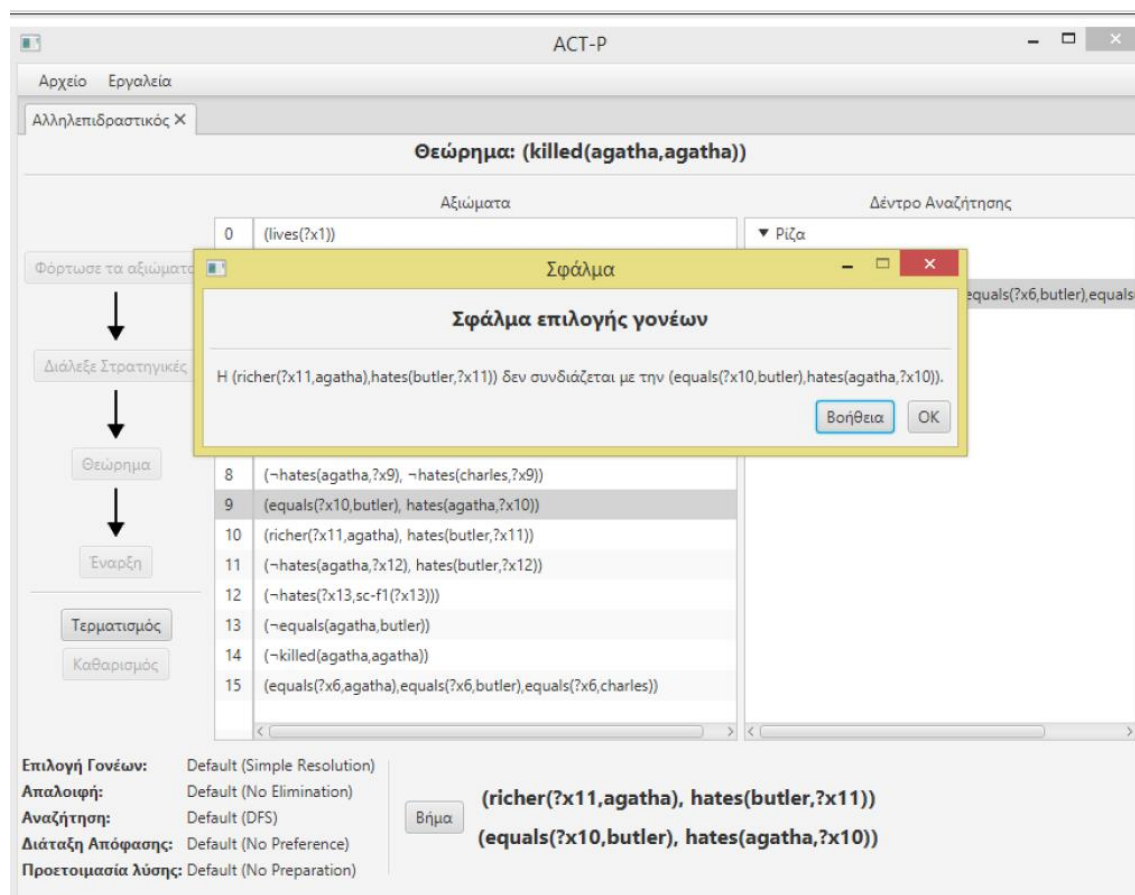
Μετάφραση σε ΚΛΠΤ:

$$\begin{aligned}
 &(\exists ?x) \text{ lives}(?x) \wedge \text{killed}(?x, \text{agatha}) \\
 &\text{lives}(\text{Agatha}) \\
 &\text{lives}(\text{butler}) \\
 &\text{lives}(\text{Charles}) \\
 &(\forall ?x) \text{ lives}(?x) \Rightarrow (\text{equals}(?x, \text{agatha}) \vee \text{equals}(?x, \text{butler}) \vee \text{equals}(?x, \text{charles})) \\
 &((\forall ?x ?y) \text{ killed}(?x, ?y) \Rightarrow \text{hates}(?x, ?y)
 \end{aligned}$$

$$\begin{aligned}
& ((\forall ?x ?y)) \text{ killed}(?x, ?y) \Rightarrow \neg \text{richer}(?x, ?y) \\
& (\forall ?x) \text{ hates}(\text{agatha}, ?x) \Rightarrow \neg \text{hates}(\text{charles}, ?x) \\
& (\forall ?x) \neg \text{equals}(?x, \text{butler}) \Rightarrow \text{hates}(\text{agatha}, ?x) \\
& (\forall ?x) \neg \text{richer}(?x, \text{agatha}) \Rightarrow \text{hates}(\text{butler}, ?x) \\
& (\forall ?x) \text{ hates}(\text{agatha}, ?x) \Rightarrow \text{hates}(\text{butler}, ?x) \\
& (\forall ?x)(\exists ?y) \neg \text{hates}(?x, ?y) \\
& \neg \text{equals}(\text{agatha}, \text{butler})
\end{aligned}$$

Προς Απόδειξη Πρόταση: killed(Agatha, Agatha)

Το παραπάνω πρόβλημα θα πρέπει να μετατραπεί σε ΠΜ για να μπορεί να διεξαχθεί η απόδειξη και γίνεται μέσω του μηχανισμού που διαθέτει ο αποδείκτης. Αρχικά, ο εκπαιδευόμενος εισάγει την προς απόδειξη έκφραση και στην συνέχεια επιλέγει μέσω των διαθέσιμων στρατηγικών, μια στρατηγική επίλυσης για το πρόβλημα. Μετά, επιλέγει τις προς επίλυση προτάσεις και με βάση τη στρατηγική απόδειξης ξεκινά η αλληλεπίδραση με τον αποδείκτη. Στην εικόνα 2.38, παρουσιάζεται η επίλυση ενός προβλήματος μέσω του αποδείκτη και η λανθασμένη επιλογή του χρήστη.



Εικόνα 2.38 Αλληλεπιδραστικός Αποδείκτης Θεωρημάτων i-ACTP

2.6.5 Πειραματικά Αξιολόγηση

Για την αξιολόγηση του αποδεικτικού θεωρημάτων πραγματοποιήσαμε μια πειραματική μελέτη κατά την διάρκεια του μαθήματος «Τεχνητή Νοημοσύνη». Στην πειραματική μελέτη συμμετείχαν 120 φοιτητές (άνδρες και γυναίκες). Αρχικά, δημιουργήσαμε δυο ομάδες φοιτητών, που είχαν 60 φοιτητές η κάθε ομάδα: την πειραματική ομάδα (groupA) και την ομάδα ελέγχου (groupB). Η επιλογή των φοιτητών έγινε τυχαία. Αρχικά, όλοι οι φοιτητές απάντησαν σ' ένα pretest που περιείχε ασκήσεις πάνω στην απόδειξη θεωρημάτων. Το τεστ περιελάμβανε 10 ασκήσεις για επίλυση και η διάρκεια του ήταν μια ώρα. Τα προβλήματα που εξασκήθηκαν οι φοιτητές ήταν προβλήματα τα οποία μπορούσαν να επιλυθούν στο χαρτί.

Στην συνέχεια, οι φοιτητές της πειραματικής ομάδας χρησιμοποίησαν τον i-ACTP στη μελέτη τους, ενώ οι φοιτητές της ομάδας ελέγχου μελέτησαν και έλυσαν προβλήματα με τον παραδοσιακό τρόπο στο χαρτί. Έπειτα, όλοι οι φοιτητές και των δυο ομάδων έλαβαν μέρος στο τελικό τεστ, το posttest, το οποίο περιελάμβανε 10 ασκήσεις για επίλυση και η διάρκεια του ήταν μια ώρα. Το pretest και το posttest περιελάμβαναν ασκήσεις πάνω

στις ισοδυναμίες που είχαν το ίδιο επίπεδο δυσκολίας και η βαθμολογία των τεστ κυμαινόταν από 0 έως 100.

Για να αναλύσουμε τις επιδόσεις των μαθητών σχεδιάστηκε και έγινε μια t-test ανάλυση στις επιδόσεις τους στο pretest. Θεωρούμε τη μηδενική υπόθεση: δεν υπάρχει καμία διαφορά μεταξύ των αποδόσεων των φοιτητών των δύο ομάδων στο pretest. Τα αποτελέσματα δείχνουν ότι η μέση τιμή και η τυπική απόκλιση στο pretest ήταν 40.21 και 8.98 για το GroupA ($M = 40.21$, $SD = 8.98$), και 41.38 και 9.54 για GroupB ($M = 41.38$, $SD = 9.54$) αντίστοιχα. Επίσης, η τιμή p (επίπεδο σημαντικότητας) ήταν $p = 0.492$ ($p > 0.05$) και $t = 0.69$ και το effect size d ήταν 0.12. Συνεπώς, από τα αποτελέσματα φαίνεται ότι οι ομάδες GroupA και GroupB δεν είχαν στατιστικά σημαντική διαφορά στο επίπεδο γνώσεων σχετικά με την απόδειξη θεωρημάτων και τις στρατηγικές επίλυσης πριν τη πειραματική μελέτη. Στον πίνακα 2.12 παρουσιάζονται η μέση τιμή και η τυπική απόκλιση των δύο ομάδων στο pretest και το posttest.

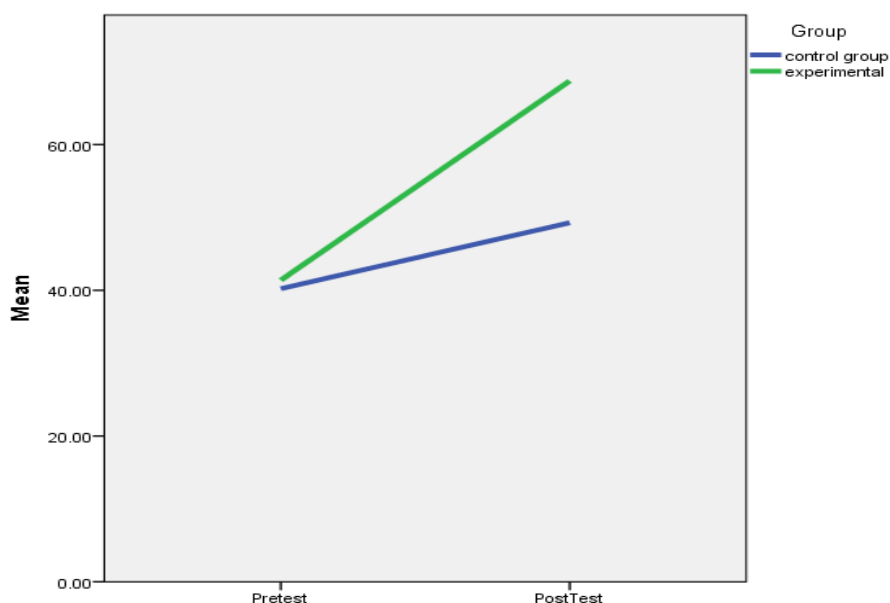
Πίνακας 2.12 Αποτελέσματα pretest και posttest πάνω στην απόδειξη θεωρημάτων

Group	Pretest		Posttest	
	M	SD	M	SD
GroupA (control)	40. 21	8. 98	49. 28	10. 82
GroupB (experimental)	41. 38	9. 54	68. 73	13. 7

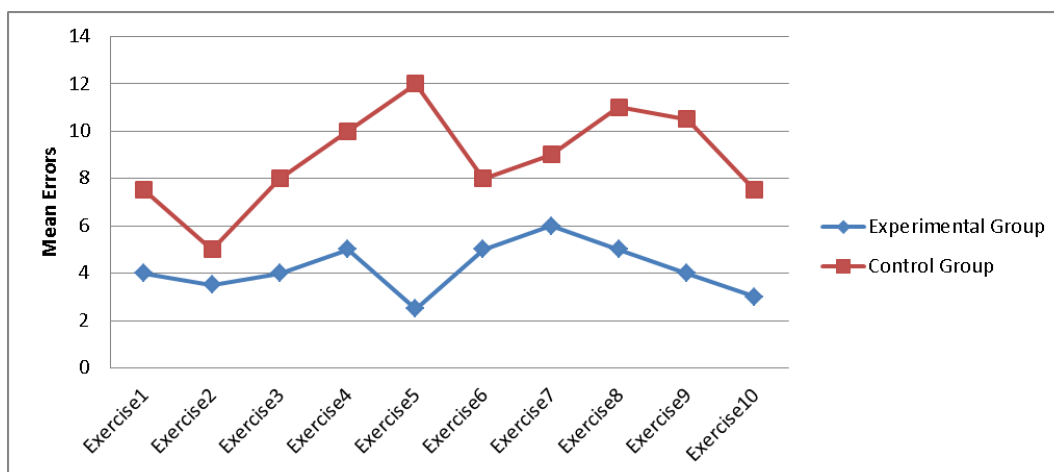
Τα αποτελέσματα δείχνουν ότι οι φοιτητές της ομάδας GroupB, ενώ στο pretest είχαν μια μέση τιμή 41.38, αυτή αυξήθηκε στο posttest στο 68.73, ενώ η ομάδα GroupA είχε μέση τιμή 40.21 στο pretest που αυξήθηκε στο 49.28 στο posttest. Επομένως, τα αποτελέσματα δείχνουν μια μεγάλη αύξηση της μέσης τιμής για την ομάδα GroupB, η οποία είναι αρκετά μεγαλύτερη από την μέση τιμή της ομάδας groupA στο posttest.

Για να προσδιοριστεί η αποτελεσματικότητα της μάθησης, πραγματοποιήσαμε μια ανάλυση συνδιακύμανσης (ANCOVA) για να ελέγξουμε τη διαφορά μεταξύ των δύο ομάδων, χρησιμοποιώντας τα αποτελέσματα του pretest ως συμμεταβλητή και τις βαθμολογίες στο posttest ως εξαρτημένες μεταβλητές. Από τα αποτελέσματα της ανάλυσης ANCOVA φαίνεται ότι οι διαφορές στις βαθμολογίες στο posttest είναι στατιστικά και σημαντικά διαφορετικές μεταξύ των δύο ομάδων ($F = 7.57$ $p = 0.00 < 0.05$). Από τα αποτελέσματα προκύπτει επίσης ότι η απόδοση των φοιτητών της πειραματικής ομάδας, που χρησιμοποίησαν τον αλληλεπιδραστικό αποδείκτη, ήταν καλύτερη από εκείνη της ομάδας

ελέγχου. Επιπλέον, στην εικόνα 2.39 παρουσιάζονται οι επιδόσεις της κάθε ομάδας στο pretest και το posttest, ενώ στην εικόνα 2.40 παρουσιάζεται η ανάλυση του πλήθους των λαθών των δυο ομάδων .



Εικόνα 2.39 Πειραματικά αποτελέσματα pretest και posttest πάνω στους αλγορίθμους αναζήτησης



Εικόνα 2.40 Ο μέσος αριθμός των λαθών της πειραματικής ομάδας και της ομάδας ελέγχου

Τα αποτελέσματα προκύπτει ότι ο μέσος αριθμός λαθών της ομάδας ελέγχου ήταν μεγαλύτερος από τον μέσο αριθμό λαθών των φοιτητών της πειραματικής ομάδας. Πιο συγκεκριμένα, ο μέσος αριθμός λαθών της πειραματικής ομάδας κυμαινόταν από 3 έως 6, ενώ ο μέσος όρος λαθών που έγιναν από την ομάδα ελέγχου κυμαινόταν από 5 έως 12 λάθη σε κάθε άσκηση.

3

Ημιαυτόματη

Παραγωγή

Ασκήσεων

3.1 Εισαγωγή

Η δημιουργία ασκήσεων αποτελεί μια σημαντική συνιστώσα της εκπαιδευτικής διαδικασίας και των εκπαιδευτικών συστημάτων. Η παραγωγή νέου εκπαιδευτικού υλικού αποτελεί μια σημαντική διεργασία, που είναι αναγκαία για τον εμπλουτισμό των εκπαιδευτικών συστημάτων με νέες ασκήσεις, αλλά ταυτόχρονα αποτελεί μια χρονοβόρα διαδικασία για τον διδάσκοντα. Η αυτοματοποίηση αυτής της διαδικασίας προσφέρει μεγάλη βοήθεια στον διδάσκοντα. Με αυτό τον τρόπο, παρέχεται η δυνατότητα σε ένα εκπαιδευτικό σύστημα να εμπλουτίζεται συνεχώς και αυτόματα με νέες διαφορετικές ασκήσεις, δίνοντας την δυνατότητα στους εκπαιδευόμενους να αλληλεπιδρούν και να εξασκούνται με διαφορετικές ασκήσεις και εκπαιδευτικές δραστηριότητες. Επομένως, η αυτόματη δημιουργία ασκήσεων μπορεί να συμβάλει στην καλύτερη προσαρμογή του εκπαιδευτικού περιεχομένου στις ανάγκες του κάθε εκπαιδευόμενου, όπως για παράδειγμα η παροχή κατάλληλων ασκήσεων κλιμακούμενης δυσκολίας ανάλογα με τις επιδόσεις τους.

Στα πλαίσια της ενότητας αυτής, παρουσιάζουμε νέες προσεγγίσεις για την ημιαυτόματη παραγωγή ασκήσεων σχετικών με τη λογική και τους αλγορίθμους αναζήτησης. Αρχικά, γίνεται αναφορά σε μια προσέγγιση για την δημιουργία ασκήσεων λογικής και πιο συγκεκριμένα για την δημιουργία συντακτικά και σημασιολογικά ορθών προτάσεων-ασκήσεων ΚΛΠΤ με διαφορετικά επίπεδα δυσκολίας. Οι παραγόμενες ασκήσεις μπορούν να χρησιμοποιηθούν σε εκπαιδευτικά συστήματα λογικής. Στην συνέχεια, παρουσιάζουμε μια προσέγγιση για την ημι-αυτόματη δημιουργία αλληλεπιδραστικών ασκήσεων αλγορίθμων

(Grivokostopoulou & Hatzilygeroudis, 2015). Η δημιουργία των ασκήσεων περιλαμβάνει δύο στάδια. Για την δημιουργία των ασκήσεων, απαιτείται η εισαγωγή τιμών για μια σειρά από παραμέτρους σε κάθε στάδιο, που σχετίζονται με την πολυπλοκότητα του δέντρου εφαρμογής ενός αλγορίθμου (πρώτο στάδιο) και το συγκεκριμένο πρόβλημα (δεύτερο στάδιο). Στην συνέχεια, γίνεται η δημιουργία των αντίστοιχων ασκήσεων και προσδιορίζονται οι λύσεις τους. Επίσης, παρέχεται η δυνατότητα από το σύστημα, μέσω των εργαλείων διαχείρισης εκπαιδευτικού υλικού, να μπορούν να τροποποιηθούν κατάλληλα οι παραγόμενες ασκήσεις.

3.2 Σχετικές Εργασίες

Στην διεθνή βιβλιογραφία, οι μηχανισμοί παραγωγής ασκήσεων, ανάλογα με τον βαθμό αυτοματισμού τους, μπορούν να ταξινομηθούν σε τρεις βασικές κατηγορίες (Cable et al., 2013)

(α) *Αυτόματοι παραγωγοί ασκήσεων (automatic exercise generators)*. Χρησιμοποιούν μια προσέγγιση για αυτόματη δημιουργία ασκήσεων, χωρίς καμία παρέμβαση του χρήστη. Το πλεονέκτημα είναι ότι δημιουργείται άμεσα ένας μεγάλος αριθμός ασκήσεων, αλλά ωστόσο δεν προσαρμόζεται στις ανάγκες του εκπαιδευτικού. Ως εκ τούτου, οι εκπαιδευτικοί δεν μπορούν να προσαρμόσουν τις ασκήσεις στις συνήθειες τους, ούτε στις ιδιαιτερότητες των μαθητών τους.

(β) *Χειροκίνητοι παραγωγοί ασκήσεων (manual exercise generators)*. Η χειροκίνητη παραγωγή ασκήσεων γίνεται μέσω των *εργαλείων συγγραφής (authoring tools)* από τον εκπαιδευτικό. Το κύριο πλεονέκτημα τους είναι ότι ο διδάσκοντας έχει την απόλυτη ελευθερία τόσο ως προς το πεδίο εφαρμογής όσο και ως προς το εκπαιδευτικό περιεχόμενο της άσκησης. Ωστόσο, ο διδάσκοντας πρέπει να καθορίσει πλήρως τις ασκήσεις και τις λύσεις τους, η οποία είναι μια επίπονη και χρονοβόρα διαδικασία.

(γ) *Ημιαυτόματοι παραγωγοί ασκήσεων (semi-automatic exercise generators)*. Αποτελούν μια ενδιάμεση κατηγορία ανάμεσα στις δυο παραπάνω προσεγγίσεις. Οι ημιαυτόματες γεννήτριες ασκήσεων αρχικά κατασκευάζουν τα δομικά μέρη της άσκησης και στην συνέχεια επιτρέπουν στο χρήστη να επέμβει στη διαδικασία δημιουργίας, καθορίζοντας μια σειρά από παραμέτρους σχετικά με τις ασκήσεις που θέλει να δημιουργήσει. Οι ημιαυτόματες γεννήτριες έχουν τις ίδιες δυνατότητες, όπως και οι αυτόματες γεννήτριες, δηλαδή να δημιουργούν γρήγορα ένα μεγάλο αριθμό ασκήσεων και να παρέχουν μια ευέλικτη λύση, επιτρέποντας στους εκπαιδευτικούς να ρυθμίσουν τις παραμέτρους των παραγόμενων ασκήσεων.

Η αυτόματη δημιουργία ασκήσεων πραγματοποιείται άμεσα και είναι κατάλληλη για τη δημιουργία μεγάλου πλήθους απλών ασκήσεων. Το κύριο μειονέκτημα των αυτόματων μεθόδων έγκειται στον ακριβή καθορισμό των παιδαγωγικών παραμέτρων των ασκήσεων που θα δημιουργηθούν καθώς και στην επαλήθευση της ποιότητας των παραγόμενων ασκήσεων. Αντιθέτως, η χειροκίνητη παραγωγή ασκήσεων προσφέρει στον διδάσκοντα πλήρη ευελιξία, αλλά είναι πολύ χρονοβόρα. Ιδιαίτερα, στις περιπτώσεις όπου απαιτείται ο διδάσκοντας να καθορίσει πάρα πολλές παραμέτρους των ασκήσεων, μπορεί να επηρεαστεί αρνητικά η ποιότητα των παραγόμενων ασκήσεων (Psyché et al., 2014).

Η ημιαυτόματη παραγωγή ασκήσεων συνδυάζει τα πλεονεκτήματα της αυτόματης και της χειροκίνητης παραγωγής ασκήσεων όπου, με μικρή επέμβαση του διδάσκοντα, μπορούν να δημιουργηθούν άμεσα οι κατάλληλες ασκήσεις. Το βασικότερο είναι ότι ο διδάσκοντας εμπλέκεται στην φάση της δημιουργίας τόσο όσο χρειάζεται για να διασφαλιστεί η ποιότητα και η καταλληλότητα των παραγόμενων ασκήσεων. Στα πλαίσια της εργασίας μας και με βάση τα ιδιαίτερα χαρακτηριστικά του πεδίου της λογικής και των αλγορίθμων, επιλέχθηκε να ακολουθηθεί μια ημιαυτόματη προσέγγιση παραγωγής ασκήσεων, η οποία είναι ιδιαίτερα κατάλληλη και αποτελεσματική.

Στην διεθνή βιβλιογραφία οι υπάρχουν αρκετές εργασίες που αφορούν την αυτόματη δημιουργία ερωτήσεων πολλαπλής επιλογής και συμπλήρωσης κενού χρησιμοποιώντας κυρίως τεχνικές επεξεργασίας φυσικής γλώσσας (Hoshino & Nakagawa, 2005; Heilman & Smith, 2010; Malafeev, 2014; Goto et al., 2010). Επίσης, υπάρχουν εργασίες για την δημιουργία ασκήσεων πάνω σε θέματα προγραμματισμού και παραγωγής κώδικα (Magdalenic et al., 2013; Radošević et al., 2010), ενώ υπάρχουν λίγες αναφορές για παραγωγή αλληλεπιδραστικών ασκήσεων.

Το CLAIRe (Cablé et al., 2013) είναι μια πλατφόρμα για την δημιουργία ασκήσεων με ημιαυτόματο τρόπο για τις ανάγκες της τριτοβάθμιας εκπαίδευσης. Το σύστημα περιλαμβάνει μια γεννήτρια ασκήσεων αυτό-αξιολόγησης, που επιτρέπει την παραγωγή ασκήσεων ημιαυτόματα καθώς και την αυτόματη αξιολόγηση των απαντήσεων του φοιτητή. Στην εργασία (Cristea & Tuduce, 2005) παρουσιάζεται ένα σύστημα για την εκμάθηση κυκλωμάτων και την παραγωγή ασκήσεων που αφορούν κυκλώματα εναλλασσόμενου ρεύματος. Πιο συγκεκριμένα, παράγονται κυκλώματα διαφορετικής πολυπλοκότητας με βάση τις επιθυμητές ιδιότητες, μέσω των αντίστοιχων τιμών τους. Επίσης, στην εργασία (Olney et al., 2011) παρουσιάζεται ένας μηχανισμός για την παραγωγή γραφικών αναπαραστάσεων ασκήσεων σχετικών με χάρτες εννοιών (Concepts map). Η μέθοδος που αναπτύχθηκε εξάγει αυτόματα αναπαραστάσεις γνώσης από ένα βιβλίο και τις χρησιμοποιεί

για να δημιουργήσει εννοιολογικούς χάρτες. Επίσης, αξίζει να σημειωθεί ότι από τις παραπάνω εργασίες, πολύ λίγες συνδυάζουν την αυτόματη δημιουργία ασκήσεων σε συνδυασμό με την δυσκολία της άσκησης ή/και τον αυτόματο προσδιορισμό της δυσκολίας.

3.3 Ημιαυτόματη Παραγωγή ασκήσεων λογικής

Η παραγωγή ασκήσεων λογικής αποτελεί μια συνθέτη και πολύπλοκη διαδικασία για την οποία απαιτείται ο προσδιορισμός πολλών συνιστωσών σε συντακτικό και σημασιολογικό επίπεδο. Για τον σκοπό αυτό, σχεδιάσαμε και αναπτύξαμε έναν μηχανισμό παραγωγής ασκήσεων λογικής, ο οποίος έχει ως στόχο να βοηθήσει τον διδάσκοντα στην δημιουργία νέων ασκήσεων και τον εμπλουτισμό των εκπαιδευτικών συστημάτων με νέες ασκήσεις λογικής. Ο κύριος στόχος του μηχανισμού είναι να προσφέρει έναν τρόπο άμεσης και εύκολης δημιουργίας νέων ασκήσεων ΚΛΠΤ, οι οποίες στην συνέχεια μπορούν να χρησιμοποιηθούν στα εκπαιδευτικά συστήματα και για την διδασκαλία θεμάτων λογικής

3.3.1 Μηχανισμός Αυτόματης Δημιουργίας Ασκήσεων ΚΛΠΤ

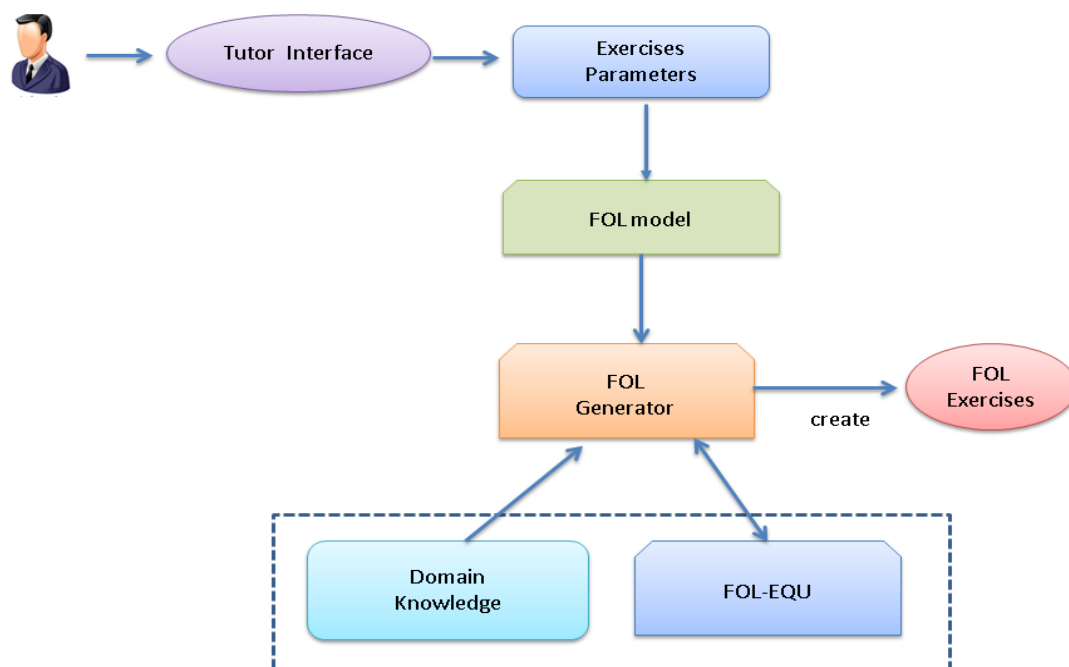
Ο μηχανισμός για την παραγωγή μιας νέας άσκησης ΚΛΠΤ λειτουργεί σε δυο στάδια: Αρχικά, στο πρώτο στάδιο ο διδάσκοντας καλείται να καθορίσει τις βασικές παραμέτρους της άσκησης, οι οποίες είναι οι ακόλουθες:

- Πλήθος ατομικών εκφράσεων
- Πλήθος σταθερών
- Πλήθος συναρτήσεων
- Πλήθος των ποσοδεικτών
- Ύπαρξη συνεπαγωγών

Ο μηχανισμός εξετάζει αν οι παράμετροι που έχουν εισαχθεί είναι έγκυροι (valid) και αν με βάση τις παραμέτρους αυτές μπορούν να παραχθούν ορθά δομημένες εκφράσεις (well-formed formula) ΚΛΠΤ. Έπειτα, ο μηχανισμός προσδιορίζει τα πρότυπα εκφράσεων ΚΛΠΤ (patterns), που είναι κατάλληλα και αντιστοιχούν στις παραμέτρους που προσδιορίστηκαν. Για κάθε σύνολο παραμέτρων ο μηχανισμός προσδιορίζει ένα πρότυπο έκφρασης ΚΛΠΤ. Τα πρότυπα που προσδιορίστηκαν εισάγονται στο μηχανισμό παραγωγής ΚΛΠΤ (FOL generator) για να παραχθούν οι κατάλληλες εκφράσεις με βάση το αρχικό πρότυπο.

Ο μηχανισμός παραγωγής ΚΛΠΤ είναι ο μηχανισμός όπου παράγει μια νέα έκφραση με βάση τις παραπάνω παραμέτρους, τη βάση γνώσης (domain knowledge) και τον μηχανισμό αυτόματης δημιουργίας λογικών ισοδυναμιών. Η βάση γνώσης του πεδίου αποτελείται από εκφράσεις ΚΛΠΤ που έχουν καθοριστεί από τον ειδικό και είναι αποθηκευμένες στο σύστημα. Ο μηχανισμός αυτόματης δημιουργίας ισοδύναμων εκφράσεων λογικής βασίζεται στους κανόνες ισοδυναμίας της κατηγορηματικής λογικής και δοθείσας μιας έκφρασης ΚΛΠΤ εφαρμόζει τους δυνατούς κανόνες ισοδυναμίας και παράγει νέες εκφράσεις. Με αυτό τον τρόπο, η βάση γνώσης δεν είναι στατική, αλλά εμπλουτίζεται συνεχώς, με νέες εκφράσεις, που παράγονται από τον μηχανισμό. Ο μηχανισμός δέχεται σαν είσοδο ένα σύνολο παραμέτρων και παράγει ως έξοδο εκφράσεις ΚΛΠΤ χωρίς την παρέμβαση του διδάσκοντα. Οι εκφράσεις βασίζονται στις προτιμήσεις του δημιουργού και παρέχεται η δυνατότητα να τροποποιηθούν, ανάλογα με τις τρέχουσες εκπαιδευτικές ανάγκες και με βάση τις προτιμήσεις του διδάσκοντα. Στην εικόνα 3.1 παρουσιάζεται ο τρόπος λειτουργίας του μηχανισμού παραγωγής ασκήσεων.

Η παραγωγή ασκήσεων σε συνδυασμό με την αυτόματη παραγωγή βοήθειας μπορούν να αποτελέσουν σημαντικές λειτουργικές συνιστώσες για να γίνει πιο αποτελεσματική η μάθηση και να προσαρμόζεται στις ανάγκες του κάθε εκπαιδευόμενου.



Εικόνα 3.1 Μηχανισμός Παραγωγής Ασκήσεων Λογικής

Στην εικόνα 3.2 παρουσιάζονται οι παράμετροι που προσδιορίστηκαν από τον διδάσκοντα καθώς και το σύνολο των εκφράσεων ΚΛΠΤ που παράχθηκαν από τον μηχανισμό. Ο

διδάσκοντας προσδιορίζει ότι επιθυμεί την δημιουργία μιας έκφρασης ΚΛΠΤ που να αποτελείται από δυο ατομικές εκφράσεις, ένα ποσοδείκτη και να είναι εύκολη (“easy”).

Automatic Generation FOL Exercises

Number of Atomic Expression: 2

Number of constant: -

Number of functions: -

Number of Quantifiers: 1

Implication: -

Negation: -

Difficulty Level: Easy

- ☒ $(\exists x)(\text{human}(x) \wedge \text{smart}(x))$
- ☐ $(\exists x)(\text{human}(x) \wedge \sim \text{smart}(x))$
- ☒ $(\forall x)(\text{human}(x) \Rightarrow \sim \text{smart}(x))$
- ☐ $\sim(\exists x)(\text{human}(x) \Rightarrow \sim \text{smart}(x))$
- ☐ $\sim(\exists x)(\text{human}(x) \Rightarrow \text{smart}(x))$
- ☐ $(\exists x)(\text{pet}(x) \wedge \text{dog}(x))$
- ☐ $(\exists x)(\text{cube}(y) \wedge \sim \text{Tet}(y))$
- ☐ $(\forall x)(\text{cube}(y) \Rightarrow \sim \text{Tet}(y))$
- ☐ $\sim(\exists x)(\sim \text{cube}(y) \vee \text{Tet}(y))$
- ☐ $(\exists x)(\text{pet}(x) \wedge \text{dog}(x))$
- ☐ $(\exists x)(\text{pet}(x) \wedge \text{cat}(x))$
- ☐ $(\exists x)\text{number}(x) \wedge \text{even}(x)$
- ☐ $(\exists x)\text{number}(x) \wedge \sim \text{even}(x)$
- ☐ $(\forall x)(\text{dog}(x) \Rightarrow \sim \text{cat}(x))$
- ☐ $(\forall z)\text{fruit}(z) \Rightarrow \text{healthy}(z)$
- ☐ $(\exists z)\text{fruit}(z) \wedge \text{healthy}(z)$
- ☐ $(\exists z)\text{fruit}(z) \wedge \sim \text{healthy}(z)$
- ☐ $(\exists x)\text{dog}(z) \wedge \sim \text{black}(x)$
- ☐ $(\forall x)\text{cat}(x) \Rightarrow \text{mammal}(x)$

Generate Clear

Εικόνα 3.2 Παραγωγή FOL εκφράσεων

3.3.2 Πειραματικά Αποτελέσματα

Για να αξιολογηθεί η απόδοση του μηχανισμού παραγωγής ασκήσεων σχεδιάστηκε και διεξήχθη μια πειραματική μελέτη. Αρχικά, στα πλαίσια της μελέτης δημιουργήθηκε ένα σύνολο από 100 ασκήσεις – εκφράσεις ΚΛΠΤ που είχαν παραχθεί αυτόματα από το σύστημα και οι αντίστοιχες παράμετροι δημιουργίας τους. Στην συνέχεια, τρεις διδάσκοντες βαθμολόγησαν την ποιότητα και την καταλληλότητα των παραγόμενων ασκήσεων για να εισαχθούν σε ένα εκπαιδευτικό σύστημα. Οι παραγόμενες ασκήσεις βαθμολογήθηκαν από τους διδάσκοντες στην κλίμακα από 1 έως 5. Στους πίνακες 3.1 και 3.2 παρουσιάζονται τα αποτελέσματα της αξιολόγησης των ασκήσεων από τους τρεις διδάσκοντες, όσον αφορά την ποιότητα και την καταλληλότητα των παραγόμενων εκφράσεων ΚΛΠΤ.

Πίνακας 3.1 Αξιολόγηση της ποιότητας των παραγόμενων ΚΛΠΤ

Tutors	N	Mean	SD
Tutor1	100	4.65	0.45
Tutor2	100	4.81	0.35
Tutor3	100	4.79	0.58

Πίνακας 3.2 Αξιολόγηση της καταλληλότητας των παραγόμενων ΚΛΠΤ

Tutors	N	Mean	SD
Tutor1	100	4.33	0.58
Tutor2	100	4.51	0.55
Tutor3	100	4.55	0.52

Οι εκφράσεις ΚΛΠΤ που παράχθηκαν αυτόματα από τον μηχανισμό αξιολογήθηκαν για την ποιότητα τους σε συντακτικό και σημασιολογικό επίπεδο. Από τα αποτελέσματα της αξιολόγησης προκύπτει ότι οι βαθμολογήσεις των αξιολογητών ήταν ιδιαίτερα ικανοποιητικές και οι μέσες τιμές για τον κάθε διδάσκοντα ήταν 4.65, 4.81 και 4.79 αντίστοιχα. Επίσης, οι αξιολογητές βαθμολόγησαν τις εκφράσεις ως προς την καταλληλότητα τους να εισαχθούν σε ένα εκπαιδευτικό σύστημα και να δοθούν ως ασκήσεις στους εκπαιδευόμενους. Τα αποτελέσματα της αξιολόγησης ήταν πολύ ικανοποιητικά ($M=4.33, SD=0.58$, $M=4.51$ $SD=0.55$, $M=4.55$ $SD=0.52$).

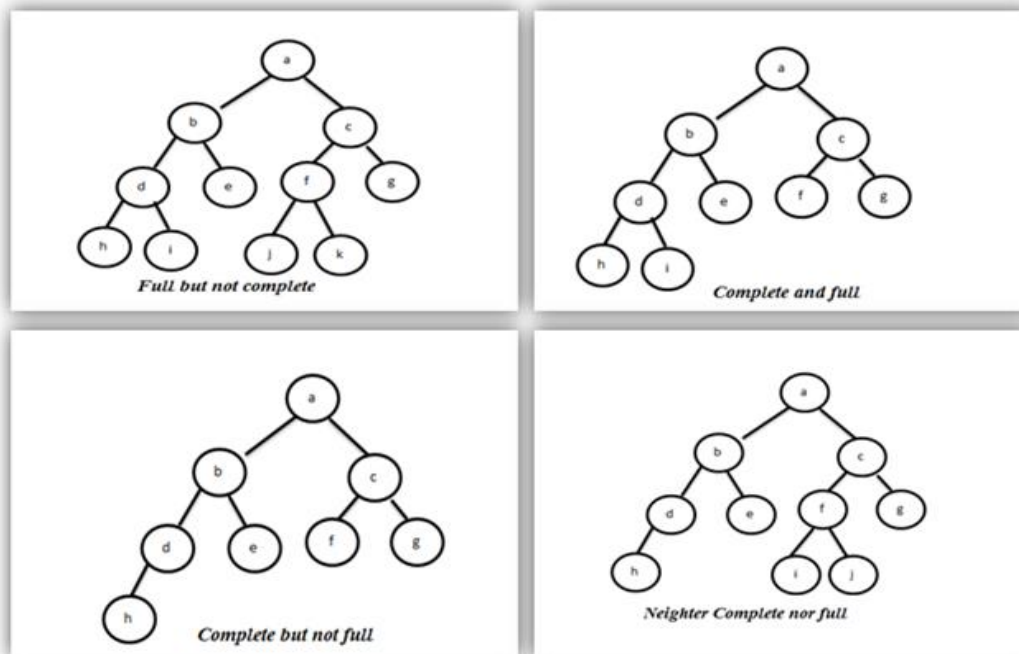
Τα αποτελέσματα της μελέτης δείχνουν ότι η προσέγγιση για την παραγωγή ασκήσεων λογικής μπορεί να δημιουργεί αυτόματα εκφράσεις ΚΛΠΤ, οι οποίες είναι σημασιολογικά ορθές. Η σημασιολογική ορθότητα αποτελεί το βασικότερο και πολυπλοκότερο κριτήριο που πρέπει να πληρείται κατά την δημιουργία εκφράσεων λογικής προκειμένου να μπορούν να χρησιμοποιηθούν σε λογικές διαδικασίες σε εκπαιδευτικά συστήματα. Συνεπώς, η προσέγγιση της ημι-αυτόματης παραγωγής ασκήσεων αποτελεί μια σημαντική συνεισφορά στα ευφυή εκπαιδευτικά συστήματα, παρέχοντας αυτόματα, άμεσα και με συνέπεια πρωτότυπες ασκήσεις και συμβάλλοντας στην αποτελεσματικότερη μάθηση.

3.4 Ημι-αυτόματη Παραγωγή Αλληλεπιδραστικών Ασκήσεων για αλγορίθμους

3.4.1 Βασικές Έννοιες

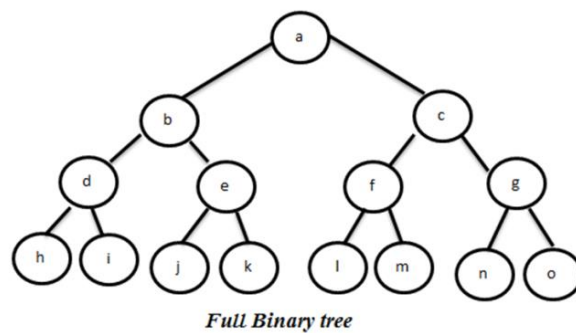
Για την επίλυση προβλημάτων αναζήτησης είναι απαραίτητο να γίνει η περιγραφή τους ως προβλήματα αναζήτησης καθώς και να προσδιοριστούν οι βασικές αρχές που διέπουν την αναζήτηση της λύσης. Πιο συγκεκριμένα, θα πρέπει να γίνει περιγραφή του χώρου καταστάσεων του προβλήματος, ο προσδιορισμός της αρχικής κατάστασης, της τελικής κατάστασης καθώς και ο προσδιορισμός των κατάλληλων ενεργειών για την εξεύρεση της λύσης και η επιλογή του κατάλληλου αλγορίθμου αναζήτησης για την επίλυση του προβλήματος. Η αναπαράσταση ενός προβλήματος ως πρόβλημα αναζήτησης αφορά και τον καθορισμό της αναπαράστασής του ως δέντρο ή γράφο. Η έννοια του δέντρου είναι βασική στα προβλήματα αναζήτησης. Ένα δέντρο είναι ένας μη κατευθυνόμενος γράφος στον οποίο οποιεσδήποτε δύο κορυφές συνδέονται με ένα και μόνο απλό μονοπάτι. Κάθε συνεκτικός γράφος χωρίς κύκλους είναι ένα δέντρο και ένα γενικό δέντρο αναζήτησης περιέχει όλα τα δυνατά μονοπάτια που μπορούν να ακολουθήσουν οι αλγόριθμοι αναζήτησης. Βασικά είδη δέντρων είναι τα ακόλουθα.

Δυαδικά Δέντρα (binary tree): Ένα δυαδικό δέντρο t είναι μία περασμένη συλλογή στοιχείων. Όταν το δυαδικό δέντρο δεν είναι κενό, τότε έχει μία ρίζα και τα υπόλοιπα στοιχεία (αν υπάρχουν) επιμερίζονται σε δύο δυαδικά δένδρα που ονομάζονται το αριστερό και το δεξιό υποδέντρο του t . Τα δυαδικά δένδρα παρουσιάζουν τις παρακάτω δύο βασικές ιδιότητες: α) το σχήδιο ενός δυαδικού δένδρου με n στοιχεία ($n > 0$) έχει ακριβώς $n-1$ ακμές και β) ένα δυαδικό δέντρο ύψους h ($h \geq 0$) έχει τουλάχιστον h και το πολύ $2^h - 1$ στοιχεία. Στις εικόνες 3.3 και 3.4 παρουσιάζονται οι διάφορες μορφές ενός δυαδικού δένδρου.



Εικόνα 3.3 Παραδείγματα δυαδικών δέντρων

Ένα δυαδικό δέντρο ύψους h είναι *πλήρες (full)* αν περιέχει ακριβώς $2^h - 1$ στοιχεία. *Συμπληρωμένο (complete)* δυαδικό δένδρο ύψους h είναι εκείνο το δυαδικό δένδρο, το οποίο προκύπτει από ένα πλήρες δυαδικό δένδρο ύψους h εάν διαγράψουμε k στοιχεία με αρίθμηση $2^h - i$, για $1 \leq i \leq k, (k \geq 0)$.



Εικόνα 3.4 Πλήρες δυαδικό δέντρο

3.4.2 Μηχανισμός Αυτόματης Δημιουργίας Ασκήσεων Αλγορίθμων

Η δημιουργία νέου εκπαιδευτικού υλικού και ειδικότερα νέων διαδραστικών ασκήσεων στη θεματική περιοχή των αλγορίθμων αναζήτησης είναι μια πολύπλοκη διαδικασία, η οποία απαιτεί από τον διδάσκοντα υψηλό επίπεδο δεξιοτήτων, όπως για παράδειγμα να δημιουργεί κώδικα σε κάποια γλώσσα προγραμματισμού. Επίσης, απαιτείται από τον διδάσκοντα να επιλύει τις ασκήσεις που δημιουργεί και να καθορίζει τις σωστές απαντήσεις.

Ο μηχανισμός αυτόματης δημιουργίας ασκήσεων (Exercise Generator-EGM) έχει αναπτυχθεί με στόχο να βοηθήσει τον διδάσκοντα να δημιουργήσει αυτόματα νέες ασκήσεις και λειτουργεί σε δυο στάδια. Στο πρώτο στάδιο ο διδάσκοντας αρχικά καθορίζει τις παραμέτρους της άσκησης και κατόπιν ο μηχανισμός σχεδιάζει και αναπτύσσει με βάση τις παραμέτρους και τις κατάλληλες γραφικές αναπαραστάσεις δέντρων.

Στο δεύτερο στάδιο της διαδικασίας, γίνεται η διαμόρφωση των ασκήσεων, με βάση επιπλέον παραμέτρους που έχει προσδιορίσει ο διδάσκοντας, όπως το είδος του αλγορίθμου αναζήτησης, η ευρετική συνάρτηση (όπου απαιτείται), ο αρχικός και τελικός κόμβος της άσκησης καθώς επίσης και το επίπεδο δυσκολίας της άσκησης. Η παραγωγή των ασκήσεων βασίζεται σε πρότυπα (pattern based approach) ασκήσεων τα οποία έχουν προκαθοριστεί και όταν μια νέα άσκηση δημιουργείται, τότε προσδιορίζεται το κατάλληλο πρότυπο και διαμορφώνεται σύμφωνα με τις παραμέτρους της άσκησης που καθορίστηκαν από τον διδάσκοντα. Μετά τη δημιουργία νέων ασκήσεων, γίνεται προσδιορισμός της δυσκολίας τους με βάση την προσέγγιση που παρουσιάζεται στην συνέχεια στην ενότητα 4.5 (Grivokostopoulou et al., 2015; Grivokostopoulou et al., 2016b) και αντίστοιχα γίνεται καθορισμός της κατάλληλης ανάδρασης με βάση τον μηχανισμό αυτόματης παραγωγής βοήθειας (Perikos et al., 2016).

Αρχικά, στο πρώτο στάδιο, ο διδάσκοντας καθορίζει τις τιμές των παραμέτρων της άσκησης (παράμετροι εισόδου) και ο μηχανισμός δημιουργεί κατάλληλες γραφικές αναπαραστάσεις των δέντρων. Οι παράμετροι που καθορίζονται από τον διδάσκοντα και είναι οι εξής:

- Ο αριθμός των κόμβων του δέντρου
- Ο μέγιστος αριθμός παιδιών για κάθε κόμβο.
- Το μέγιστο βάθος του δέντρου
- Ο βασικός τύπος του δέντρου:

- Πλήρες Δέντρο (Complete Tree):
- Δυαδικό Δέντρο (Binary Tree) :
 - Συμπληρωμένο δυαδικό δέντρο (Complete Binary Tree)
 - Πλήρες δυαδικό δέντρο (Full Binary Tree)
- Τυχαίο: κάθε κόμβος έχει ένα τυχαίο και αυθαίρετο αριθμό παιδιών, σύμφωνα με τον μέγιστο αριθμό που καθορίστηκε.

Ο μηχανισμός εξετάζει αν οι παράμετροι που εισήγαγε ο διδάσκοντας είναι ορθοί και μπορεί να παραχθεί ένα έγκυρο είδος δέντρου με τις κατάλληλες συνδέσεις μεταξύ των κόμβων. Σε αντίθετη περίπτωση, ο μηχανισμός ενημερώνει τον διδάσκοντα ότι οι επιλογές δεν είναι σωστές και καλείται να εισάγει νέες παραμέτρους. Στην συνέχεια, ο μηχανισμός με βάση τις τιμές που ορίζονται από τον διδάσκοντα, δημιουργεί αυτόματα μια σειρά από γραφικές απεικονίσεις. Πιο συγκεκριμένα, ο EGM παράγει αυτόματα κώδικα που υλοποιεί την δομή του δέντρου και παράγει μια σειρά από γραφικές απεικονίσεις των δένδρων με βάση τις παραμέτρους που προσδιορίστηκαν. Ο μηχανισμός παράγει διάφορα είδη δεντρικών αναπαραστάσεων, με βάση τις προδιαγραφές που καθορίστηκαν, και ο διδάσκοντας μπορεί να επιλέξει μια από αυτές να προσδιορίσει επιπρόσθετες πληροφορίες για την άσκηση. Πιο συγκεκριμένα, στο δεύτερο στάδιο, ο διδάσκοντας καλείται να καθορίσει επιπλέον πρόσθετες παραμέτρους οι οποίες είναι οι ακόλουθες:

- Το είδος του αλγορίθμου που θα χρησιμοποιηθεί για την επίλυση της άσκησης.
- Ο αρχικός κόμβος και τελικός κόμβος (στόχος)
- Επιλογή ευρετικής συνάρτησής και συνάρτηση κόστους (όπου απαιτείται). Ο διδάσκοντας στους αλγορίθμους ευρετικής αναζήτησης μπορεί να επιλέξει ως ευρετική μια από τις διαθέσιμες που παρέχει ο μηχανισμός ή να ορίσει μια δική του συνάρτηση. Οι ευρετικές συναρτήσεις που παρέχονται από τον μηχανισμό είναι:

- Η απόσταση Manhattan (Manhattan Distance) που υπολογίζεται ως εξής:

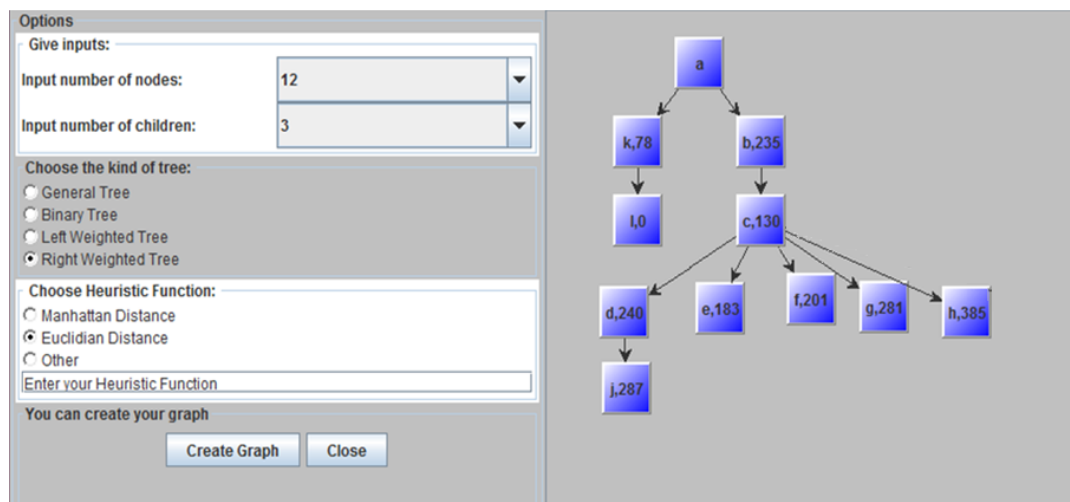
$$Md(S, F) = |X_S - X_F| + |Y_S - Y_F|$$

- Η ευκλείδεια απόσταση (Euclidian Distance) που υπολογίζεται ως εξής

$$d(S, F) = \sqrt{(X_S - X_F)^2 + (Y_S - Y_F)^2}$$

Στην συνέχεια, ο μηχανισμός με βάση την δεντρική δομή που προσδιορίστηκε στο πρώτο στάδιο της διαδικασίας, και με τις παραπάνω παραμέτρους, δημιουργεί την άσκηση. Στην

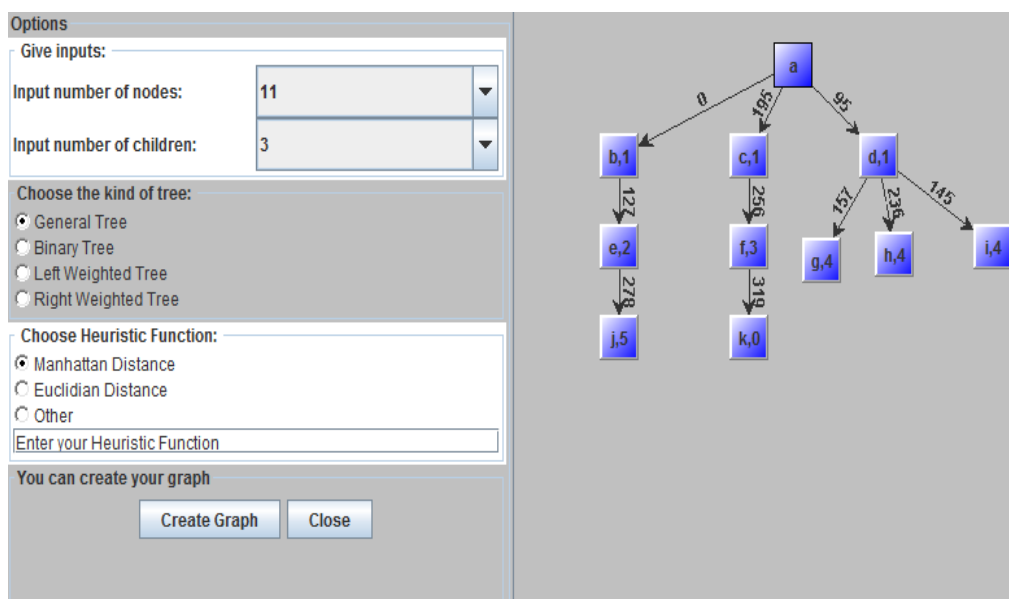
εικόνα 3.5 παρουσιάζεται μια άσκηση που δημιουργήθηκε από τον μηχανισμό για τον αλγόριθμο αναρρίχησης λόφου (Hill Climbing-HC) με ευρετική συνάρτηση την ευκλείδεια απόσταση.



Εικόνα 3.5 Παράδειγμα παραγωγής μια άσκησης του HC αλγορίθμου

Στην γραφική αναπαράσταση του δέντρου, ο κόμβος i έχει ευρετική τιμή μηδέν, αναπαριστώντας με αυτόν τον τρόπο ότι πρόκειται για τον στόχο του αλγορίθμου δηλαδή την τελική κατάσταση.

Στην εικόνα 3.6 παρουσιάζεται μια άσκηση που δημιουργήθηκε για τον αλγόριθμο A^* , όπου έχουν προσδιοριστεί επιπλέον χαρακτηριστικά για κάθε κόμβο, όπως το πραγματικό κόστος και η τιμή της ευρετικής συνάρτησης. Ο διδάσκοντας έχει επιλέξει ως δέντρο μια τυχαία αναπαράσταση “General Tree” και ως ευρετική συνάρτηση την Manhattan Distance.



Εικόνα 3.6 Παράδειγμα παραγωγής μια άσκησης του αλγορίθμου A*

3.4.3 Πειραματικά Αποτελέσματα

Για να αξιολογηθεί η απόδοση του μηχανισμού παραγωγής ασκήσεων σχεδιάστηκε και διεξήχθη μια πειραματική μελέτη. Αρχικά, στα πλαίσια της μελέτης δημιουργήθηκε ένα σύνολο από 150 ασκήσεις, που είχαν παραχθεί αυτόματα από το σύστημα μαζί με τις παραμέτρους δημιουργίας τους. Στην συνέχεια, τρεις διδάσκοντες βαθμολόγησαν την ποιότητα και την καταλληλότητα των παραγόμενων ασκήσεων για να εισαχθούν σε ένα εκπαιδευτικό σύστημα. Οι παραγόμενες ασκήσεις βαθμολογήθηκαν από τους διδάσκοντες στην κλίμακα από 1 έως 5. Στους πίνακες 3.3 και 3.4 παρουσιάζονται τα αποτελέσματα της αξιολόγησης των ασκήσεων από τους τρεις διδάσκοντες, όσον αφορά τη ποιότητα και την καταλληλότητα των παραγόμενων αλληλεπιδραστικών ασκήσεων.

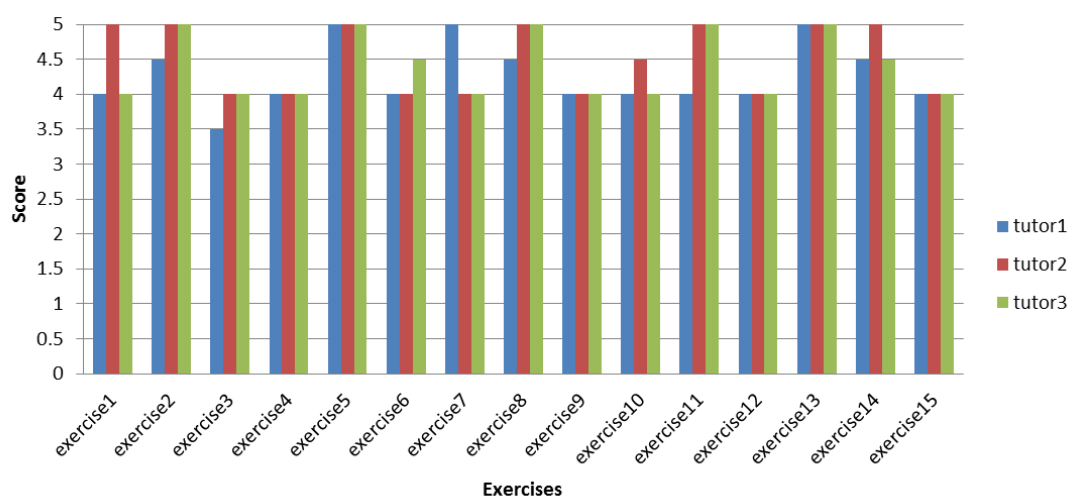
Πίνακας 3.3 Αξιολόγηση Ασκήσεων σε σχέση με την ποιότητά τους

Tutors	N	Mean	SD
Tutor1	150	4.39	0.69
Tutor2	150	4.53	0.61
Tutor3	150	4.56	0.59

Πίνακας 3.4 Αξιολόγηση Ασκήσεων σε σχέση με τη καταλληλότητά τους.

Tutors	N	Mean	SD
Tutor1	150	4.65	0.49
Tutor2	150	4.75	0.46
Tutor3	150	4.74	0.47

Από τα αποτελέσματα της αξιολόγησης προκύπτει ότι οι βαθμολογήσεις των αξιολογητών ήταν ιδιαίτερα ικανοποιητικές και οι μέσες τιμές για τον κάθε διδάσκοντα ήταν 4.39, 4.53 και 4.56 αντίστοιχα. Επίσης, οι αξιολογητές αξιολόγησαν τις ασκήσεις ως κατάλληλες να εισαχθούν σε ένα εκπαιδευτικό σύστημα και να δοθούν ως ασκήσεις στους εκπαιδευόμενους. Τα αποτελέσματα της αξιολόγησης ήταν πολύ ικανοποιητικά ($M=4.65$ $SD=0.49$, $M=4.75$ $SD=0.46$, $M=4.74$ $SD=0.47$). Στην εικόνα 3.7 παρουσιάζονται μερικά παραδείγματα αλληλεπιδραστικών ασκήσεων, όπως βαθμολογήθηκαν αντίστοιχα και από τους τρεις διδάσκοντες, απ' όπου προκύπτει ότι οι βαθμολογίες ήταν πολύ ικανοποιητικές.



Εικόνα 3.7 Παράδειγμα Αξιολόγηση Ασκήσεων σχετικά με την ποιότητα

4

Αυτόματος Προσδιορισμός Δυσκολίας Ασκήσεων

4.1 Εισαγωγή

Ο προσδιορισμός του επιπέδου δυσκολίας των ασκήσεων είναι ένας πολύ σημαντικός παράγοντας στην εκπαιδευτική διαδικασία καθώς συμβάλει στην καλύτερη προσαρμογή των εκπαιδευτικών δραστηριοτήτων στον κάθε εκπαιδευόμενο. Ο προσδιορισμός του επιπέδου δυσκολίας αποτελεί μια πολύπλοκη και χρονοβόρα διαδικασία η οποία είναι άρρηκτα συνδεδεμένη με το πεδίο που εφαρμόζεται, εξαρτάται από τα χαρακτηριστικά και τις λειτουργικές διαδικασίες του και είναι δύσκολο να μοντελοποιηθεί. Στις περισσότερες περιπτώσεις, γίνεται από τον διδάσκοντα ο οποίος βασίζεται στην εμπειρία του και την αίσθηση του, για να εκτιμήσει την δυσκολία μιας άσκησης χωρίς όμως να βασίζεται σε κάποιο μηχανισμό συλλογιστικής για την εξαγωγή και υποστήριξη των εκτιμήσεών του. Επομένως, ο προσδιορισμός της δυσκολίας των ασκήσεων από τον διδάσκοντα είναι προσεγγιστικός και σε ορισμένες περιπτώσεις μπορεί να καταστεί ασυνεπής (Alexandrou-Leonidou & Philippou, 2005; Van de Watering & Van der Rijt, 2006; Verdu et al., 2012).

Η ανάπτυξη μεθοδολογιών και συστημάτων για τον αυτόματο προσδιορισμό της δυσκολίας των ασκήσεων αποτελεί μια σημαντική διεργασία που μπορεί να βοηθήσει τους διδάσκοντες και ταυτόχρονα να αποτελέσει ένα χρήσιμο μηχανισμό στα ευφυή σύστημα διδασκαλίας. Η ενσωμάτωση ενός αυτόματου μηχανισμού προσδιορισμού της δυσκολίας σε ένα ευφύες εκπαιδευτικό σύστημα είναι σημαντική και ιδιαίτερα απαραίτητη στην εκπαιδευτική διαδικασία καθώς θεωρείται απαραίτητο, οι ασκήσεις που παρέχονται στους εκπαιδευόμενους να είναι προσαρμοσμένες στο επίπεδο γνώσεών τους. Ωστόσο, η παροχή ασκήσεων από το σύστημα οι οποίες είναι χαμηλότερου επιπέδου δυσκολίας από το επίπεδο γνώσεών του εκπαιδευόμενου πολλές φορές μπορεί να οδηγήσει στο να χάσει οποιαδήποτε

αίσθηση και διάθεση μάθησης. Από την άλλη πλευρά, η παροχή ασκήσεων με επίπεδο δυσκολίας υψηλότερο από το επίπεδο γνώσεων του και της απόδοσης του, μπορεί να απογοητεύσει τον εκπαιδευόμενο και να έχει αρνητική επίδραση στην μάθηση του (Leung & Li, 2007). Επομένως, για να μεγιστοποιήσουμε την αποτελεσματικότητα των μαθησιακών διαδικασιών, θα πρέπει να παρέχονται ασκήσεις και εκπαιδευτικές δραστηριότητες κατάλληλου επιπέδου στους εκπαιδευόμενους, που να είναι προσαρμοσμένες στο επίπεδο των γνώσεών τους. Συνεπώς, σε ένα εκπαιδευτικό σύστημα ο αυτόματος, ο ακριβής και συνεπής προσδιορισμός του επιπέδου δυσκολίας των ασκήσεων είναι ένας πολύ σημαντικός παράγοντας.

Στην διεθνή βιβλιογραφία, υπάρχουν λίγες αναφορές εκπαιδευτικών συστημάτων που διαθέτουν μηχανισμούς για την αυτόματη εκτίμηση του επιπέδου δυσκολίας των ασκήσεων και των τεστ. Στα περισσότερα εκπαιδευτικά συστήματα η εκτίμηση της δυσκολίας διεξάγεται χειροκίνητα από τον διδάσκοντα-εμπειρογνώμονα και ο προσδιορισμός της δυσκολίας των ασκήσεων και των δραστηριοτήτων μάθησης εξαρτάται από τις έννοιες των ασκήσεων που συμμετέχουν και απαιτούνται εξειδικευμένες μέθοδοι ανάλυσης.

Στα πλαίσια της ενότητας αυτής, παρουσιάζουμε νέες μεθοδολογίες για τον αυτόματο προσδιορισμό της δυσκολίας των ασκήσεων στο πεδίο της λογικής και των αλγορίθμων αναζήτησης. Αρχικά, παρουσιάζουμε μια προσέγγιση για τον προσδιορισμό της δυσκολίας ασκήσεων πάνω στην μετατροπή φυσικής γλώσσας σε ΚΛΠΤ (Perikos et al., 2011b), η οποία βασίζεται κυρίως στην εις βάθος επεξεργασία της εκφώνησης της άσκησης καθώς και στην ανάλυση των χαρακτηριστικών της σωστής απάντησης. Στην συνέχεια, παρουσιάζουμε ένα ευφυές σύστημα που σχεδιάστηκε και αναπτύχθηκε για τον προσδιορισμό της δυσκολίας ασκήσεων πάνω στην μετατροπή εκφράσεων ΚΛΠΤ σε ΠΜ (Grivokostopoulou et al., 2014a; Grivokostopoulou et al., 2012b). Η προσέγγιση αυτή βασίζεται κυρίως στην δομή της εκφώνησης της άσκησης καθώς και σε χαρακτηριστικά της διαδικασίας που απαιτείται για την επίλυση της άσκησης. Τέλος, εισάγουμε μια νευροασαφή προσέγγιση για τον προσδιορισμό της δυσκολίας ασκήσεων στο πεδίο των αλγορίθμων αναζήτησης (Grivokostopoulou et al., 2015; Grivokostopoulou et al., 2016b). Ο νευροασαφής μηχανισμός που αναπτύχθηκε δέχεται ως είσοδο μια σειρά από χαρακτηριστικά και παραμέτρους των ασκήσεων και με βάση αυτά προσδιορίζει το επίπεδο δυσκολίας τους.

4.2 Σχετική Έρευνα

Στη διεθνή βιβλιογραφία, οι μεθοδολογίες που χρησιμοποιούνται για τον προσδιορισμό της δυσκολίας των εκπαιδευτικών διαδικασιών και των ασκήσεων μπορούν να ταξινομηθούν σε

τρεις βασικές κατηγορίες μεθόδων, οι οποίες είναι: οι στατιστικές, οι ευρετικές και οι μαθηματικές (Conejo et al., 2014). Οι *στατιστικές μέθοδοι* προκειμένου να προσδιορίσουν το επίπεδο δυσκολίας των ασκήσεων βασίζονται στην ανάλυση των επιδόσεων των εκπαιδευόμενων σε αυτές. Στις *ευρετικές μεθόδους* η δυσκολία καθορίζεται από τον διδάσκοντα-εμπειρογνώμονα ενώ οι *μαθηματικές μέθοδοι* βασίζονται στην ανάλυση μιας εκπαιδευτικής διαδικασίας και προσδιορίζουν το επίπεδο δυσκολίας της με βάση τα χαρακτηριστικά και τις παραμέτρους της. Επίσης, οι μέθοδοι προσδιορισμού δυσκολίας μπορούν να διακριθούν σε εκείνες που βασίζονται στην ανάλυση των *χαρακτηριστικών της άσκησης* (*exercise-based*) και σε εκείνες που βασίζονται στην ανάλυση των *επιδόσεων των μαθητών* (*student-based*) στις ασκήσεις. Οι ευρετικές και οι μαθηματικές μέθοδοι είναι βασισμένες στην άσκηση (*exercise-based*), ενώ οι στατιστικές μέθοδοι ακολουθούν την βασισμένη στον μαθητή (*student-based*) προσέγγιση.

Στην βιβλιογραφία, μια βασική τεχνική που χρησιμοποιείται στις βασισμένες στον μαθητή προσεγγίσεις είναι η *Item Response Theory* (IRT) (Cheng et al., 2008; Embertson & Reise, 2000). Πιο συγκεκριμένα, η IRT είναι μια τεχνική που βασίζεται στην εκτίμηση της πιθανότητας ένας εκπαιδευόμενος να απαντήσει σωστά μια άσκηση. Ένα εκπαιδευτικό σύστημα που βασίζεται στην IRT ξεκινά με ασκήσεις μέτριας δυσκολίας. Μια σωστή απάντηση του εκπαιδευόμενου θα έχει ως αποτέλεσμα να του δοθεί στην συνέχεια μια πιο δύσκολη άσκηση, ενώ στην περίπτωση μια λανθασμένης απάντησης θα του δοθεί στην συνέχεια μια λιγότερο δύσκολη άσκηση. Η αυτόματη και ακριβής εκτίμηση του επιπέδου δυσκολίας των ασκήσεων θα μπορούσε να συνεισφέρει στα συστήματα που εφαρμόζουν την IRT τεχνική για την επιλογή των κατάλληλων ασκήσεων προς τους εκπαιδευόμενους. Στην εργασία (Verdu et al., 2012), παρουσιάζεται ένα ασαφές έμπειρο σύστημα (*genetic fuzzy expert system*) που βασίζεται σε κανόνες για να ταξινομεί την δυσκολία των ασκήσεων. Αρχικά, ο εκπαιδευτικός καθορίζει το επίπεδο δυσκολίας των ασκήσεων και στην συνέχεια ένα ασαφές έμπειρο σύστημα αναπροσαρμόζει το αρχικό επίπεδο δυσκολίας της άσκησης με βάση τις επιδόσεις των εκπαιδευόμενων. Στην εργασία (Koutsojannis et al., 2007), παρουσιάζεται μια υβριδική προσέγγιση για τον προσδιορισμό του επιπέδου δυσκολίας ασκήσεων. Πιο συγκεκριμένα, αρχικά ο εμπειρογνώμονας καλείται να δώσει μια σειρά από κανόνες για τον υπολογισμό της δυσκολία και στην συνέχεια χρησιμοποιείται ένας γενετικός αλγόριθμος για να καθορίσει τη δυσκολία των ασκήσεων σύμφωνα με τις απαντήσεις και τις επιδόσεις των μαθητών. Στην εργασία (Kunichika et al., 2002), παρουσιάζεται μια υπολογιστική μέθοδος εκτίμηση δυσκολίας ερωτήσεων σε ένα σύστημα για την εκμάθηση της αγγλικής γλώσσας. Η δυσκολία προσδιορίζεται με βάση την πολυπλοκότητα των ερωτήσεων καθώς και δεδομένα που εξάγονται από την επεξεργασία της φυσικής γλώσσας

και την ανάλυση του συντακτικού δέντρου των προτάσεων. Στην εργασία (Barker-Plummer et al., 2009), οι συγγραφείς μελετούν τη δυσκολία της μετάφρασης προτάσεων φυσικής γλώσσας σε ΚΛΠΤ. Το κύριο χαρακτηριστικό της προσέγγισης αυτής είναι ότι προσπαθεί να προσδιορίσει την δυσκολία της μετατροπής αναλύοντας τις απαντήσεις και τις επιδόσεις των μαθητών. Για το σκοπό αυτό οι συγγραφείς χρησιμοποιούν στην εργασία τους ένα σύνολο από είκοσι προτάσεις φυσικής γλώσσας και μελετούν τις απαντήσεις των εκπαιδευόμενων σε αυτές και με βάση τις απαντήσεις των εκπαιδευόμενων γίνεται υπολογισμός δύο μετρικών. Η πρώτη μετρική αφορά το ποσοστό των εκπαιδευόμενων που έκαναν λάθος σε μια πρόταση, ενώ η δεύτερη μετρική προσδιορίζει τον αριθμό των προσπαθειών που πραγματοποιήθηκαν σε μία άσκηση ώστε να βρεθεί η σωστή απάντηση. Τέλος, με βάση αυτές τις δύο μετρικές κάθε μια άσκηση χαρακτηρίζεται Εύκολη ή Δύσκολη ως προς το να γίνει κάποιο λάθος, καθώς επίσης και Εύκολη ή Δύσκολη ως προς το να επιλυθεί το λάθος. Επίσης, με την περαιτέρω ανάλυση των απαντήσεων των μαθητών, παρουσιάζεται στην εργασία (Barker-Plummer et al., 2008) μια ταξινόμηση των τύπων των λαθών που έγιναν από τους μαθητές κατά τη μετάφραση προτάσεων φυσικής γλώσσας σε ΚΛΠΤ.

4.3 Αυτόματος Προσδιορισμός Δυσκολία Ασκήσεων Μετατροπής Φυσικής Γλώσσας σε Κατηγορηματική Λογική

Η διδασκαλία της Κατηγορηματική Λογικής Πρώτης τάξης (ΚΛΠΤ) ως γλώσσα αναπαράστασης και συλλογισμού έχει πολλές πτυχές. Μια σημαντική πτυχή είναι η μετατροπή φυσικής γλώσσας σε ΚΛΠΤ, μια διαδικασία που ονομάζεται *λογικός φορμαλισμός* (formalization) της φυσικής γλώσσας. Για να βοηθήσουμε τους εκπαιδευόμενους να κατανοήσουν την διαδικασία της μετατροπής, αναπτύχθηκε το ευφύες σύστημα NLtoFOL (Hatzilygeroudis & Perikos, 2009 ; chronopoulos et al., 2010), το οποίο έχει ως στόχο να βοηθήσει τους εκπαιδευόμενους να κατανοήσουν και να μάθουν να εφαρμόζουν την διαδικασία μετατροπής βηματικά με βάση τη διαδικασία SIP (Hatzilygeroudis, 2007).

Σε αυτή την ενότητα παρουσιάζουμε ένα μηχανισμό που αναπτύχθηκε για τον αυτόματο προσδιορισμό του επίπεδου δυσκολίας των ασκήσεων στην μετατροπή Φυσικής Γλώσσας (ΦΓ) σε ΚΛΠΤ. Αρχικά, πραγματοποιείται σημασιολογική ανάλυση της εκφώνησης της άσκησης (πρόταση ΦΓ) και εις βάθος ανάλυση της σωστής απάντησης (έκφραση ΚΛΠΤ) με στόχο να εξαχθούν χαρακτηριστικά και παράμετροι της άσκησης. Στην συνέχεια, προσδιορίζεται το επίπεδο δυσκολίας της άσκησης με βάση τα εξαγόμενα χαρακτηριστικά μέσω ευφύων συστημάτων. Επομένως, ο μηχανισμός προσδιορίζει την δυσκολία των

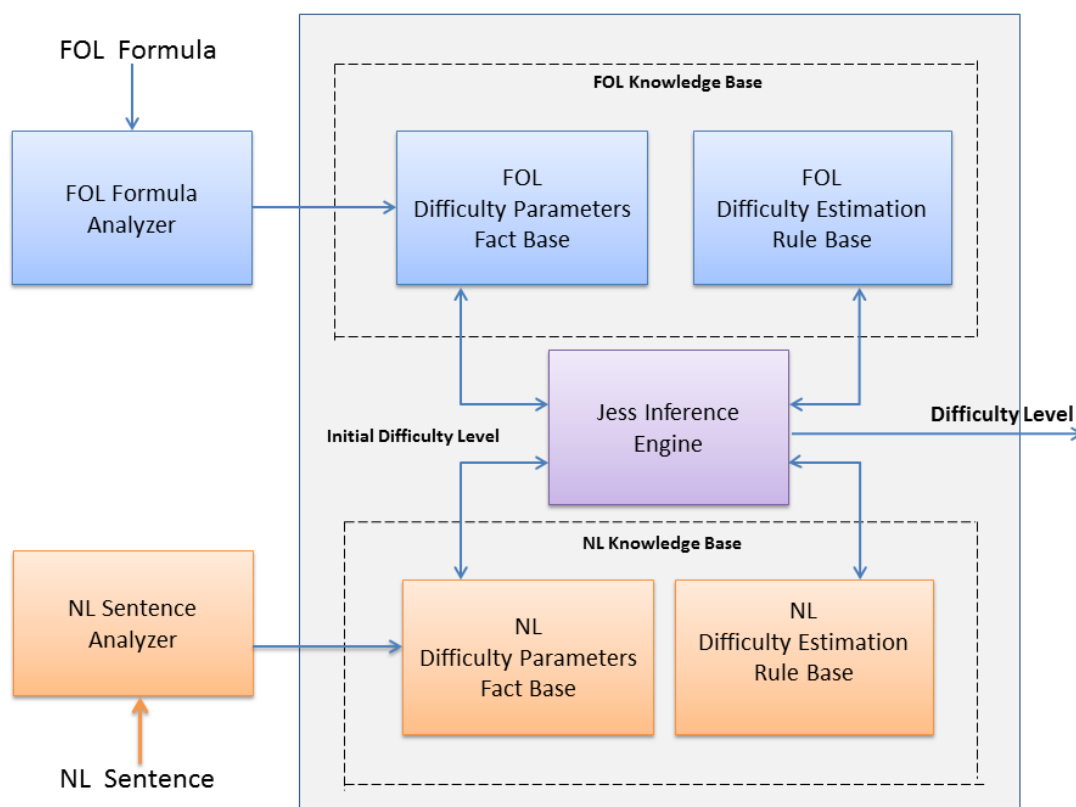
ασκήσεων με βάση τόσο την δομή και την σημασιολογία της εκφώνησης όσο και την πολυπλοκότητα της απάντησης. Ο στόχος του μηχανισμού είναι να προσδιορίσει αυτόματα, με ακρίβεια και με συνέπεια το επίπεδο δυσκολίας ασκήσεων χωρίς την παρέμβαση του διδάσκοντα.

4.3.1 Μηχανισμός Αυτομάτου Προσδιορισμού Δυσκολίας Ασκήσεων ΦΓ σε ΚΛΠΤ

Αρχικά, ο μηχανισμός προσδιορισμού της δυσκολίας των ασκήσεων προσδιορίζει το αρχικό επίπεδο δυσκολίας της άσκησης με βάση την δομή και τη πολυπλοκότητα της σωστής απάντησης. Για τον αρχικό προσδιορισμό της δυσκολίας, ο μηχανισμός βασίζεται σε παραμέτρους της έκφρασης ΚΛΠΤ, όπως είναι το πλήθος, το είδος και η σειρά των ποσοδεικτών, το πλήθος των συνεπαγωγών και το πλήθος των διαφορετικών διασυνδετικών. Στην συνέχεια, γίνεται ανάλυση της πρότασης φυσικής γλώσσας και εξάγονται κατάλληλα χαρακτηριστικά που συμβάλουν σημαντικά στον προσδιορισμό της πολυπλοκότητας της έκφρασης ΚΛΠΤ και κατ' επέκταση της δυσκολία του φορμαλισμού. Επομένως, ο τελικός προσδιορισμός της δυσκολίας της άσκησης γίνεται με βάση τις σημασιολογικές πτυχές της φυσικής γλώσσας και τη δομή της έκφρασης ΚΛΠΤ. Το επίπεδο δυσκολίας των ασκήσεων κατηγοριοποιείται σε πέντε επίπεδα τα οποία είναι τα εξής : “very easy”, “easy”, “medium”, “difficult” και “advanced”.

4.3.1.1 Αρχιτεκτονική του Συστήματος

Η αρχιτεκτονική του συστήματος εκτίμησης δυσκολίας απεικονίζεται στην Εικόνα 4.1 και αποτελείται από πέντε βασικές μονάδες: *Αναλυτής Εκφράσεων ΚΛΠΤ (FOL Analyzer)*, *Αναλυτής Προτάσεων ΦΓ (NL Analyzer)*, τη *Βάση Γνώσης ΚΛΠΤ (FOL knowledge base)*, τη *Βάση Γνώσης ΦΓ (NL knowledge base)* καθώς και το *Μηχανισμό Εξαγωγής Συμπερασμάτων Jess (Jess Inference Engine)*.



Εικόνα 4.1. Αρχιτεκτονική του συστήματος προσδιορισμού της δυσκολίας ασκήσεων της διαδικασίας NLtoFOL

Αρχικά, το σύστημα δέχεται ως είσοδο τη σωστή απάντηση μιας άσκησης, δηλαδή μια έκφραση ΚΛΠΤ, η οποία αναλύεται μέσω της μονάδας *Αναλυτής Εκφράσεων ΚΛΠΤ* (Perikos et al., 2011a). Η μονάδα αυτή αναλύει την απάντηση και προσδιορίζει τις τιμές των παραμέτρων που σχετίζονται με την πολυπλοκότητα της άσκησης (Πίνακας 4.1) και στην συνέχεια εισάγει τα αντίστοιχα γεγονότα στην *Βάση Γεγονότων Παραμέτρων Δυσκολίας ΚΛΠΤ* (*FOL Difficulty Parameters Fact Base*). Από την άλλη πλευρά, η μονάδα *Αναλυτής Προτάσεων ΦΓ* δέχεται ως είσοδο την εκφώνηση σε φυσική γλώσσα και προσδιορίζει τις τιμές των παραμέτρων που σχετίζονται με τη σημασιολογία της φυσικής γλώσσας και εισάγει τα αντίστοιχα γεγονότα στη *Βάση Γεγονότων Παραμέτρων Δυσκολίας ΦΓ* (*NL Difficulty Parameters Fact Base*).

Το σύστημα εκτίμησης δυσκολίας αποτελείται από δύο έμπειρα συστήματα (ΕΣ) υλοποιημένα σε Jess (Friedman-Hill, 2003). Το πρώτο ΕΣ, το *Βασισμένο στην ΚΛΠΤ* (*FOL-Based*), εξάγει συμπεράσματα για το επίπεδο δυσκολίας χρησιμοποιώντας την βάση γνώσης που σχετίζεται με την πολυπλοκότητα της έκφρασης ΚΛΠΤ. Το δεύτερο ΕΣ, το *Βασισμένο στη ΦΓ* (*NL-Based*), βασίζεται σε παραμέτρους που προέκυψαν από την ανάλυση της δομής

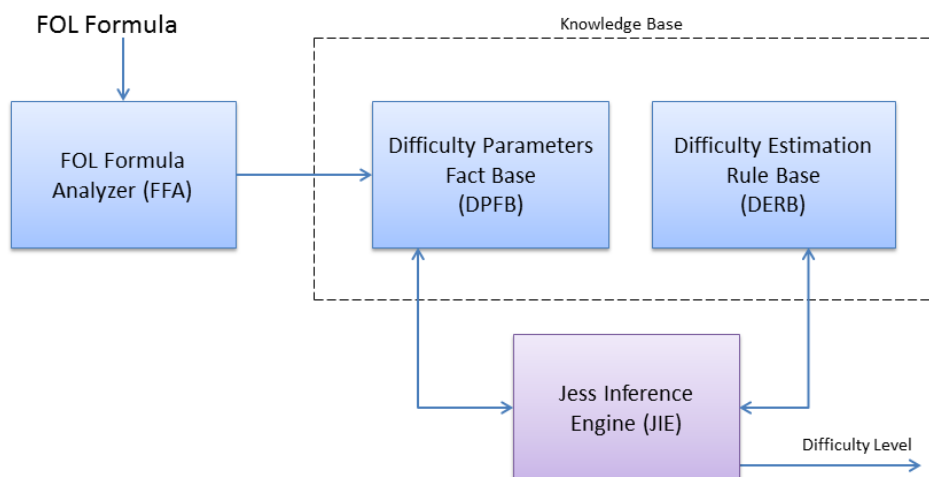
και της σημασιολογίας της φυσικής γλώσσας καθώς και στον αρχικό προσδιορισμό της δυσκολίας του πρώτου έμπειρου συστήματος και με βάση αυτά προσδιορίζει το τελικό επίπεδο δυσκολίας της άσκησης.

4.3.1.2 Έμπειρο Σύστημα προσδιορισμού δυσκολίας με βάση την έκφραση ΚΛΠΤ

Όπως αναφέρθηκε προηγουμένως, το Βασισμένο στην ΚΛΠΤ έμπειρο σύστημα εμπλέκει παραμέτρους που σχετίζονται με την συντακτική δομή της έκφρασης ΚΛΠΤ για τον προσδιορισμό του επιπέδου δυσκολίας των ασκήσεων και αποτελεί ένα έμπειρο σύστημα κανόνων (rule-based expert system) υλοποιημένο σε Jess (Εικόνα 4.1). Πιο συγκεκριμένα, το έμπειρο σύστημα αποτελείται από την *Βάση Γεγονότων Παραμέτρων Δυσκολίας (Difficulty Parameters Fact Base)*, η οποία διαχειρίζεται τις τιμές των παραμέτρων και την *Βάση Κανόνων Εκτίμησης Δυσκολίας (Difficulty Estimation Rule Base)*, η οποία αποτελείται από τους κανόνες προσδιορισμού της δυσκολίας με βάση τις τιμές των παραμέτρων. Αρχικά, όταν μια νέα έκφραση ΚΛΠΤ εισέλθει στο σύστημα αναλύεται από τον *Αναλυτή Εκφράσεων ΚΛΠΤ*, προσδιορίζονται οι τιμές των παραμέτρων της και αποθηκεύονται ως γεγονότα στην Βάση Γεγονότων. Έπειτα, το έμπειρο σύστημα δίνει ως έξοδο τον αρχικό προσδιορισμό του επιπέδου της δυσκολίας της άσκησης.

Για την ανάπτυξη του Βασισμένου στην ΚΛΠΤ έμπειρου συστήματος συνεργαστήκαμε με διδάσκοντες-ειδικούς που διέθεταν εμπειρία στη διδασκαλία της λογικής ως γλώσσας αναπαράστασης γνώσης. Στα πλαίσια της συνεργασίας αυτής καθορίστηκαν ποιοι παράμετροι των ασκήσεων και σε ποιο βαθμό είναι σημαντικοί για τον προσδιορισμό της δυσκολίας τους. Μετά από διεξοδική μελέτη, καταλήξαμε στις ακόλουθες παραμέτρους:

- Το πλήθος και ο τύπος των ποσοδεικτών.
- Το πλήθος των συνεπαγωγών.
- Το πλήθος των διαφορετικών διασυνδετικών ($\Rightarrow, \vee, \wedge, \neg$).



Εικόνα 4.2. Βασική Δομή του Βασισμένου στην ΚΛΠΤ ΕΣ

Το Βασισμένο στην ΚΛΠΤ ΕΣ προσδιορίζει το επίπεδο δυσκολίας της διαδικασίας της μετατροπής με βάση τις παραπάνω παραμέτρους. Μια έκφραση ΚΛΠΤ, σε αντίθεση με μια πρόταση ΦΓ, έχει ελεγχόμενη και αυστηρή σύνταξη κάτι που καθιστά ευκολότερη στην ανάλυση της και σύμφωνα με τους ειδικούς αποτελεί μια καλή επιλογή για την εξόρυξη σημαντικών χαρακτηριστικών γνωρισμάτων. Στο πίνακα 4.1, παρουσιάζεται ένα σύνολο από κανόνες του συστήματος οι οποίοι βασίζονται στην έκφραση ΚΛΠΤ.

Πίνακας 4.1 Κανόνες για τον προσδιορισμό της δυσκολίας από το Βασισμένο στην ΚΛΠΤ Έμπειρο Σύστημα.

#	Πλήθος των συνεπαγωγών ($\Rightarrow$)	Πλήθος των Ποσοδεικτών ($\exists \forall$)	Καθολικός Ποσοδείκτης ($\forall$)	Υπαρξιακός Ποσοδείκτης ($\exists$)	Πλήθος των διαφορετικών διασυνδετικών ($\wedge \vee$)	Επίπεδο Δυσκολίας
1	<2	0	*	*	*	Very Easy
2	≥ 2	0	*	*	*	Easy
3	0	1	*	*	<3	Easy
4	1		yes	no	≤ 1	Easy
5	1 ^d	>1 ^d	no	yes	$\geq 3^d$	Medium
6	0	>1 ^d	*	*	$\geq 3^d$	Medium
7	1	*	yes	no	2	Medium
8	1	≤ 2	yes	no	≥ 3	Medium
9	≥ 2	1	*	*	<2	Medium
10	1	*	yes	yes	*	Difficult

11	1	>2	yes	no	>=3	Difficult
12	≥ 2	1	*	*	>=2	Difficult
13	≥ 2	>1	*	*	*	Advanced

^dThese rule premises are combined with disjunction

Στην εικόνα 4.3 παρουσιάζεται ένα παράδειγμα κανόνα του συστήματος σε Jess και η αντίστοιχη επεξήγησή του.

```
(defrule rule6
  ?result <- (fol (imp 0) (quant?q) (unv ?u)
    (exst ?e) (difcon ?d) (eba ?a))
    (or (test(> ?q 1)) (test(>= ?d 3)))
=>
  (retract ?result)
  (printout output "Medium" crlf))
```

Επεξήγηση:
 ΕΑΝ ο αριθμός των συνεπαγωγών είναι 0 (?imp=0)
 ΚΑΙ
 (υπάρχουν περισσότεροι από 1 ποσοδείκτες (?q>1)
 Ή
 υπάρχουν 3 ή περισσότερα διασυνδετικά (?d≥3))
 ΤΟΤΕ το αποτέλεσμα είναι Medium

Εικόνα 4.3. Παράδειγμα Κανόνα Έμπειρου Συστήματος βασισμένο στην ΚΛΠΤ

Για παράδειγμα θεωρούμε την ακόλουθη άσκηση που αποτελείται από την εκφώνηση της άσκησης δηλαδή την πρόταση σε φυσικής γλώσσας και από την αντίστοιχη σωστή απάντηση σε ΚΛΠΤ

Εκφώνηση
(Πρόταση ΦΓ)

Every city has a dog catcher that has been bitten by any
dog living in the city

Σωστή Απάντηση
(Εκφραση ΚΛΠΤ)

$(\forall x) \text{city}(x) \Rightarrow ((\exists y) \text{dog-catcher}(y, x) \wedge (\forall z) (\text{dog}(z) \wedge \text{lives-in}(z, x)) \Rightarrow \text{has-bitten}(z, y))$

Μετά από την ανάλυση της έκφρασης ΚΛΠΤ προκύπτει ότι ο συνολικός αριθμός των διασυνδετικών είναι δύο, ο συνολικός αριθμός των ποσοδεικτών είναι τρεις, η έκφραση περιλαμβάνει καθολικό και υπαρξιακό ποσοδείκτη και ο αριθμός των διαφορετικών διασυνδετικών είναι δύο. Με βάση τις παραμέτρους αυτές προκύπτει ότι μπορεί να ενεργοποιηθεί ο κανόνας 13 (Πίνακας 4.1) και επομένως το επίπεδο δυσκολίας της άσκησης εκτιμάται ως “advanced”.

Επίσης, στον Πίνακα 4.2 παρουσιάζονται μερικά παραδείγματα ασκήσεων και τα αντίστοιχα επίπεδα δυσκολίας όπως προσδιορίστηκαν από το Βασισμένο στην ΚΛΠΤ ΕΣ.

Πίνακας 4.2 Παραδείγματα προσδιορισμού του επιπέδου δυσκολίας από το Βασισμένο στην ΚΛΠΤ Έμπειρο Σύστημα

Πρόταση ΦΓ	Έκφραση ΚΛΠΤ	$\Rightarrow$	$\exists \forall$	$\forall$	$\exists$	$\wedge \vee$	Κανόνας	Επίπεδο Δυσκολίας
John forgets his coffee when he is late.	late(John) $\Rightarrow$ forgets(John, coffee-of(John))	1	0	No	No	1	Rule1	Very Easy
All politicians are crooks.	$(\forall x)$ politician(x) $\Rightarrow$ crook(x)	1	1	Yes	No	1	Rule4	Easy
Apples and oranges are fruits.	$(\exists x)$ apple(x) $\vee$ orange(x)	0	1	No	Yes	1	Rule3	Easy
All romans are either loyal to Caesar or they hate him	$(\forall x)$ (roman(x) $\Rightarrow$ (loyal(x,Caesar) $\vee$ hates(x, Caesar)))	1	1	Yes	No	2	Rule7	Medium
There is a singer that everyone loves some of their songs.	$(\exists x)$ (singer(x) $\wedge$ $(\exists z)$ (sings(x,z) $\wedge$ loves(y,z))))	0	3	Yes	Yes	1	Rule6	Medium
A table that expands is useful for a small room	$(\forall x)$ table(x) $\wedge$ expands(x) $\Rightarrow$ (small(room) $\Rightarrow$ useful(x))	2	1	Yes	No	2	Rule12	Difficult
Some butcher likes all vegetarians.	$(\exists x)$ butcher(x) $\wedge$ $(\forall y)$ (vegetarian(y) $\Rightarrow$ likes(x,y))	1	2	Yes	Yes	2	Rule10	Difficult
Gary loves any dog who is larger than every cat.	$(\forall x)$ ((dog(x) $\wedge$ $(\forall y)$ (cat(y) $\Rightarrow$ larger(x,y))) $\Rightarrow$ loves(gary, x))	2	2	Yes	No	2	Rule13	Advanced

4.3.1.3 Σημασιολογικά Χαρακτηριστικά της Φυσικής Γλώσσας

Όπως αναφέρθηκε παραπάνω, ο προσδιορισμός της δυσκολίας των ασκήσεων βασίζεται και στην ανάλυση της εκφώνησης της άσκησης, δηλαδή στα χαρακτηριστικά της φυσικής γλώσσας. Πιο συγκεκριμένα, το Βασισμένο στη ΦΓ έμπειρο σύστημα αναλύει τη πρόταση ΦΓ και εξάγει μια σειρά από χαρακτηριστικά τα οποία οργανώνονται σε δύο κατηγορίες. Η πρώτη κατηγορία περιλαμβάνει χαρακτηριστικά τα οποία εξάγονται αποκλειστικά και μόνο από τη ανάλυση της φυσικής γλώσσας, ενώ η δεύτερη κατηγορία περιλαμβάνει χαρακτηριστικά που οι τιμές τους προσδιορίζονται αναλύοντας την πρόταση ΦΓ και συγκρίνοντας την με την αντίστοιχη έκφραση ΚΛΠΤ. Τα χαρακτηριστικά της δεύτερης κατηγορίας προσδιορίζουν την πολυπλοκότητα της πρότασης και αποτελούν ένα μέτρο της επιπρόσθετης επεξεργασίας που απαιτείται από τον εκπαιδευόμενο για να την μεταφράσει σε ΚΛΠΤ. Τα χαρακτηριστικά γνωρίσματα που εξάγονται προκύπτουν από τη μορφοσυντακτική ανάλυση της φυσικής γλώσσας και είναι τα εξής:

- *Word order matching*: Αναφέρεται στην αναδιάταξη που γίνεται στην σειρά των λέξεων ανάμεσα στη πρόταση φυσικής γλώσσας και την έκφραση ΚΛΠΤ. Ο αριθμός των αναδιατάξεων αποτελεί μια σημαντική παράμετρο πολυπλοκότητας.
- *Anaphoric Connectives*: Σχετίζεται με τη ύπαρξη αναφορικών εκφράσεων. Όσο μεγαλύτερος ο αριθμός τους τόσο αυξάνεται η πολυπλοκότητα της πρότασης.
- *Special words or phrases*: Αναφέρεται στην ύπαρξη συγκεκριμένων λέξεων που μπορούν να αυξήσουν την πολυπλοκότητα της άσκησης λόγω της δύσκολης σημασιολογίας τους. Χαρακτηριστικά παραδείγματα λέξεων είναι οι ακόλουθες {only, exactly one, unless, provided that, as long as, while}. Για παράδειγμα η ακόλουθη πρόταση φυσικής γλώσσας “Only cats chase mice”, αν και είναι σύντομη και φαίνεται απλή, θεωρείται δύσκολο να μετατραπεί σε ΚΛΠΤ, διότι θα πρέπει οι εκπαιδευόμενοι να κατανοήσουν την σημασιολογία της και να την ερμηνεύσουν ως “If something chases some mouse then it is a cat”.
- *Negating keywords*: Αναφέρεται στον αριθμό συγκεκριμένων λέξεων κλειδιών, όπως {not, no, neither, none, except, never}, που υποδηλώνουν κάποιο είδος άρνησης.
- *Quantifier Mismatch*: Αναφέρεται σε ορισμένες λέξεις-κλειδιά σε μια πρόταση, που συνδέονται με την χρήση συγκεκριμένων ποσοδεικτών. Για παράδειγμα λέξεις όπως “something”, “some”, “a” υπονοούν την χρήση του υπαρξιακού ποσοδείκτη “∃”. Όμως υπάρχουν περιπτώσεις όπου μπορεί να απαιτείται η χρήση του καθολικού ποσοδείκτη. Για παράδειγμα, στη ακόλουθη πρόταση φυσικής γλώσσας : “A banana

is both a fruit and a weapon”, παρότι η ύπαρξη της λέξης “a” υπονοεί τη χρήση του υπαρξιακού ποσοδείκτη, ωστόσο θα πρέπει να χρησιμοποιηθεί ο καθολικός ποσοδείκτης, και η σωστή έκφραση ΚΛΠΤ είναι: $(\forall x) (banana(x) \Rightarrow (fruit(x) \wedge weapon(x)))$

- *Connective Mismatch*: Αναφέρεται σε ορισμένες λέξεις-κλειδιά σε μια πρόταση, που συνδέονται με τη χρήση συγκεκριμένων διασυνδετικών. Ορισμένες λέξεις-κλειδιά σε μια πρόταση φυσικής γλώσσας υποδεικνύουν ότι θα πρέπει να χρησιμοποιηθούν συγκεκριμένα διασυνδετικά. Για παράδειγμα, προτάσεις που περιέχουν λέξεις όπως “and” ή “both” υποδεικνύουν ότι θα πρέπει να χρησιμοποιήσουμε το λογικό διασυνδετικό “Λ”. Υπάρχουν περιπτώσεις όπου λέξεις της φυσική γλώσσα μπορεί να υπονοούν τη χρήση συγκεκριμένου διασυνδετικού, ενώ η σημασιολογία της πρότασης να είναι διαφορετική και στην ΚΛΠΤ έκφραση να χρησιμοποιηθεί το κατάλληλο διασυνδετικό. Για παράδειγμα, στην ακόλουθη πρόταση φυσικής γλώσσας: “Apples and oranges are fruits” παρότι η ύπαρξη της λέξης “and” υπονοεί τη χρήση του διασυνδετικού “Λ”, ωστόσο λόγω της σημασιολογίας της πρότασης θα πρέπει να χρησιμοποιηθεί το διασυνδετικό “V”, και η σωστή έκφραση ΚΛΠΤ είναι: $(\forall x)(apple(x) \vee orange(x)) \Rightarrow fruit(x)$.

4.3.1.4 Προσδιορισμός Τελικού Επιπέδου Δυσκολία Ασκήσεων

Για τον τελικό προσδιορισμό του επιπέδου δυσκολίας μιας άσκησης, ο μηχανισμός βασίζεται και στις παραμέτρους τις σχετικές με την έκφραση ΚΛΠΤ και στα χαρακτηριστικά της πρότασης ΦΓ. Πιο συγκεκριμένα, στο τελικό βήμα, εάν είναι απαραίτητο, λόγω των επιπλέον δυσκολιών που προκύπτουν από την σημασιολογία της πρότασης ΦΓ, αναθεωρείται το αρχικό επίπεδο δυσκολίας της άσκησης. Στον Πίνακα 4.3 παρουσιάζονται οι κανόνες για τον προσδιορισμό της δυσκολίας των ασκήσεων με βάση την έκφραση ΚΛΠΤ και τα σημασιολογικά χαρακτηριστικά της πρότασης ΦΓ. Η πρώτη στήλη του πίνακα αναπαριστά την αρχική εκτίμηση του επιπέδου δυσκολίας της άσκησης από το Βασισμένο στην ΚΛΠΤ ΕΣ, ενώ οι υπόλοιπες στήλες αφορούν τα χαρακτηριστικά που σχετίζονται με τη σημασιολογία της πρότασης ΦΓ. Στην συνέχεια, με βάση τα γεγονότα εισόδου ενεργοποιείται ο κατάλληλος κανόνας και γίνεται αναπροσδιορισμός του αρχικού επιπέδου δυσκολίας, σε διαφορετική περίπτωση το τελικό επίπεδο δυσκολίας είναι το αρχικό. Στον Πίνακα 4.3, το σύμβολο “*” δηλώνει ότι η τιμή της συγκεκριμένης παραμέτρου δεν παίζει ρόλο, το SUM είναι το άθροισμα των τιμών των τεσσάρων αριθμητικών χαρακτηριστικών και ο εκθέτης “d”

δηλώνει ότι οι παράμετροι αυτοί συνδυάζονται διαζευκτικά μεταξύ τους, ενώ με τις υπόλοιπες συζευκτικά.

Πίνακας 4.3 Κανόνες προσδιορισμού επιπέδου δυσκολίας από το Βασισμένο στην ΦΓ ΕΣ.

Αρχικό Επίπεδο FOL- Based	Word order matching	Anaphoric Connectives	Negating Keywords	Special words/ phrases	Quantifier Mismatch	Connective Mismatch	Τελικό Επίπεδο Δυσκολίας
Difficult	*	>2	*	*	*	*	Advanced
Difficult	SUM >3				*	*	Advanced
Medium	*	>2	*	*	*	*	Difficult
Medium	*	*	>1	*	*	*	Difficult
Medium	SUM >3				*	*	Advanced
Easy	0	>1	*	*	yes	*	Medium
Very Easy	>2 ^d	>2 ^d	*	*	*	yes ^d	Easy

Στην εικόνα 4.4, παρουσιάζεται ένα παράδειγμα κανόνα προσδιορισμού του τελικού επιπέδου δυσκολίας με τις κατάλληλες επεξηγήσεις.

```
(defrule rule18
  ?result <- (fn1 (foldifeas) (ord ?o) (ana ?a)
               (neg ?n) (spe ?s) (qin yes) (cin ?c))
  (test(> ?a 1))
=>
  (retract ?result)
  (printout output "Medium" crlf))
```

Επεξήγηση:

ΕΑΝ η δυσκολία που προσδιορίστηκε από το Βασισμένο στην ΚΛΠΤ ΕΣ είναι Easy (foldifeas)

ΚΑΙ

υπάρχει quantifier mismatch (qin=yes)

ΚΑΙ

τα anaphoric connectives είναι περισσότερα από 1
(?a>1)

ΤΟΤΕ

Η τελική δυσκολία είναι Medium

Εικόνα 4.4. Παράδειγμα Κανόνα και οι επεξηγήσεις του Βασισμένου στη ΦΓ Έμπειρου Συστήματος.

4.3.2 Πειραματικά Αποτελέσματα

Για να αξιολογήσουμε την απόδοση, την ακρίβεια και την αξιοπιστία της μεθοδολογίας που αναπτύχθηκε πραγματοποιήσαμε μια πειραματική μελέτη σε ένα σύνολο ασκήσεων στο σύστημα NLtoFOL. Πιο συγκεκριμένα, δημιουργήσαμε ένα σύνολο από 160 διαφορετικούς τύπους ασκήσεων, όπου για κάθε άσκηση υπήρχε η πρόταση φυσικής γλώσσας και η αντίστοιχη μετάφραση της σε ΚΛΠΤ. Αρχικά, διδάσκοντες-εμπειρογνώμονες εκτίμησαν το επίπεδο δυσκολίας όλων των ασκήσεων. Έπειτα, δόθηκε ως είσοδος το σύνολο όλων των ασκήσεων στο σύστημα για να προσδιοριστεί αυτόματα το επίπεδο δυσκολίας τους. Για να αξιολογήσουμε τον μηχανισμό αυτό, έγινε σύγκριση του επιπέδου δυσκολίας των ασκήσεων που εκτιμήθηκε από τους διδάσκοντες και του επιπέδου δυσκολίας που προσδιορίστηκε από τον αυτόματο μηχανισμό. Πιο συγκεκριμένα, για την αξιολόγηση της ακρίβειας και της ποιότητας των επιδόσεων του αυτόματου μηχανισμού προσδιορισμού δυσκολίας αναλύσαμε τα αποτελέσματα και υπολογίσαμε μια σειρά από μετρικές αξιολόγησης. Δεδομένου ότι η

κλάση εξόδου είναι μια μεταβλητή πολλαπλών τιμών (*multiple class output variable*) υπολογίστηκαν οι μετρικές μέση ορθότητα (*average accuracy*), ακρίβεια (*precision*), ανάκληση (*recall*) και *F-measure* (Sokolova & Lapalme, 2009). Η μέση ορθότητα (*average accuracy*) υπολογίζεται ως εξής:

$$Average\ Accuracy = \frac{\sum_{i=1}^l \frac{tp_i + tn_i}{tp_i + fn_i + fp_i + tn_i}}{l}$$

όπου tp , fn , fp και tn είναι ορίζονται τα True Positive, False Negative, False Positive, και True Negative αντίστοιχα και l είναι η κλάση (Sokolova & Lapalme, 2009).

Η ακρίβεια (*precision*) υπολογίζεται από τον λόγο των στιγμιότυπων που ανήκουν στην κλάση και σωστά ταξινομήθηκαν σε αυτή ως εξής:

$$Precision = \frac{\sum_{i=1}^l tp_i}{\sum_{i=1}^l (tp_i + fp_i)}$$

Η ανάκληση (*recall*) υπολογίζεται από τον λόγο των στιγμιότυπων που κατηγοριοποιήθηκαν σωστά στην κλάση προς όλο το πλήθος των στιγμιότυπων που ανήκαν την κλάση ως εξής:

$$Recall = \frac{\sum_{i=1}^l tp_i}{\sum_{i=1}^l (tp_i + fn_i)}$$

Η μετρική *F-measure* συνδυάζει την ακρίβεια και την ανάκληση ως εξής:

$$F - measure = \frac{2 \times prec \times rec}{prec + rec}$$

Τα αποτελέσματα που υπολογίστηκαν παρουσιάζονται στον ακόλουθο πίνακα.

Πίνακας 4.4 Πειραματικά αποτελέσματα μηχανισμού εκτίμησης δυσκολίας ασκήσεων

Average accuracy	0.95
Precision	0.95
Recall	0.95
F-measure	0.95

Τα αποτελέσματα δείχνουν μια πολύ καλή απόδοση του μηχανισμού προσδιορισμού της δυσκολίας των ασκήσεων. Πιο συγκεκριμένα, ο μηχανισμός στο 95% των περιπτώσεων προσδιόρισε σωστά το επίπεδο δυσκολίας της άσκησης κατατάσσοντας στην κατάλληλη κατηγορία. Στην συνέχεια, πραγματοποιήσαμε μια πιο λεπτομερή ανάλυση των επιδόσεων

του μηχανισμού και εξετάσαμε την απόδοσή του σε κάθε μια κατηγορία δυσκολίας. Τα αποτελέσματα που προέκυψαν παρουσιάζονται στον πίνακα 4.5.

Πίνακας 4.5 Πειραματικά αποτελέσματα εκτίμησης δυσκολίας για κάθε μια κατηγορία

Class	Precision	Recall	F-Measure
VeryEasy	1	1	1
Easy	1	1	1
Medium	0.89	0.96	0.92
Difficult	0.94	0.84	0.89
Advanced	1	1	1

Τα αποτελέσματα δείχνουν ότι το σύστημα φαίνεται να έχει πολύ καλή απόδοση σε όλες τις κατηγορίες των ασκήσεων. Πιο συγκεκριμένα, τα αποτελέσματα δείχνουν ότι για τις κατηγορίες “very easy”, “easy” και “advanced” το σύστημα προσδιόρισε με απόλυτη ακρίβεια το επίπεδο δυσκολίας των ασκήσεων αυτών. Επίσης το σύστημα είχε πολύ καλή επίδοση στον προσδιορισμό του επιπέδου δυσκολίας των ασκήσεων στις κατηγορίες “Medium” και “Advanced”. Αξίζει να σημειωθεί ότι, οι μόνες περιπτώσεις όπου το σύστημα δεν προσδιόρισε με ακρίβεια το επίπεδο δυσκολίας ασκήσεων είναι σε ορισμένες περιπτώσεις όπου το επίπεδο δυσκολίας ήταν “Medium” και το χαρακτηρίστηκε “Difficult” καθώς και το αντίστροφο. Οι παραπάνω περιπτώσεις αποτυχίας τους συστήματος μπορεί να οφείλεται και σε ενδογενή δυσκολία να γίνει απόλυτος διαχωρισμός ασκήσεων σε μέτριες προς δύσκολες ασκήσεις.

Στην συνέχεια, υπολογίσαμε το στατιστικό δείκτη Cohen’s Kappa (Cohen, 1960), για την αποτίμηση της απόδοσης του συστήματος και για τον προσδιορισμό της αξιοπιστίας του βαθμού συμφωνίας μεταξύ του διδάσκοντα και του αυτόματου μηχανισμού προσδιορισμού της δυσκολίας. Ο δείκτης Cohen’s Kappa ορίζεται ως ένα μέτρο για τον προσδιορισμό της αξιοπιστίας της συμφωνίας ανάμεσα σε δύο αξιολογητές. Θεωρούμε ότι δυο αξιολογητές ταξινόμησαν N στοιχεία σε k αλληλο-αποκλειόμενες κατηγορίες. Τα αποτελέσματα παρουσιάζονται σε ένα πίνακα γειτνίασης $k \times k$. Το p_{ij} , όπου $i, j = 1, 2, \dots, k$ ορίζεται ως εξής:

$$p_{ij} = \frac{N_{ij}}{N}$$

όπου N_{ij} ο αριθμός των περιπτώσεων που, ενώ από τον αξιολογητή A (Rater A) κατηγοριοποιήθηκαν στην κατηγορία i , από τον αξιολογητή B (Rater B) κατηγοριοποιήθηκαν στην κατηγορία j . Για τον υπολογισμό του k συμπληρώνεται ο Πίνακας 4.6, όπου

$$p_{.j} = \sum_{i=1}^k p_{ij}$$

$$p_{i.} = \sum_{j=1}^k p_{ij}$$

Πίνακας 4.6 Υπολογισμός Cohen's Kappa

RATER A	RATER B				Total
	1	2	...	k	
1	p_{11}	p_{12}	...	p_{1k}	$p_{1.}$
2	p_{21}	p_{22}	...	p_{2k}	$p_{2.}$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$
n	p_{n1}	p_{n2}	...	p_{nn}	$p_{n.}$
Total	$p_{.1}$	$p_{.2}$	...	$p_{.n}$	1

Οι διαγώνιοι p_{ii} αντιπροσωπεύουν το ποσοστό της παρατηρηθείσας συμφωνίας μεταξύ των αξιολογητών σε κάθε κατηγορία. Το συνολικό ποσοστό της παρατηρούμενης συμφωνίας είναι:

$$p_o = \sum_{i=1}^k p_{ii}$$

και το συνολικό ποσοστό της αναμενόμενης συμφωνίας είναι:

$$p_e = \sum_{i=1}^k p_{i.} \cdot p_{.i}$$

Η μετρική Kappa, που αντιπροσωπεύει τον βαθμό συμφωνίας των αξιολογητών, υπολογίζεται ως εξής:

$$k = \frac{p_o - p_e}{1.0 - p_e}$$

Στην περίπτωση που υπάρχει πλήρη συμφωνία, η μετρική k θα είναι 1. Γενικά, ο βαθμός συμφωνίας με βάση την μετρική k μπορεί να χαρακτηριστεί ως :

Μη συμφωνία	για $k < 0$
Ισχνή συμφωνία	για k μεταξύ 0.01-0.20

Καλή συμφωνία	για k μεταξύ 0.21-0.40
Μέτρια συμφωνία	για k μεταξύ 0.41-0.60
Ισχυρή συμφωνία	για k μεταξύ 0.61-0.80
Σχεδόν τέλεια συμφωνία	για k μεταξύ 0.81-0.99

Στην περίπτωση μας ήταν $n=5$ και $1 \rightarrow$ Very Easy, $2 \rightarrow$ Easy, $3 \rightarrow$ Medium, $4 \rightarrow$ Difficult, $5 \rightarrow$ Advanced. Στα πλαίσια της μελέτη μας, η μετρική k χρησιμοποιήθηκε για να υπολογίσει τον βαθμό αξιοπιστίας της συμφωνίας μεταξύ των εκτιμήσεων δυσκολίας του διδάσκοντα και του συστήματος στο σύνολο των 160 ασκήσεων. Τα αποτελέσματα της μελέτης έδειξαν ότι υπήρξε μια σχεδόν τέλεια συμφωνία (Viera & Garret, 2005) μεταξύ εκπαιδευτών και συστήματος, δεδομένου ότι η μετρική ήταν $k = 0.851$ (confidence interval 95%, $p < 0.0005$). Τα αποτελέσματα επιβεβαιώνουν την πολύ καλή απόδοση του συστήματος. Επιπλέον, στις περισσότερες περιπτώσεις, τα χαρακτηριστικά που εξάγονται από την ανάλυση της δομής και της γραμματικής σύνταξη της φυσικής γλώσσας συμβάλουν στην εκτίμηση της πολυπλοκότητας της άσκησης μετατροπής.

4.4 Αυτόματος Προσδιορισμός Δυσκολία Ασκήσεων Μετατροπής Κατηγορηματικής Λογικής σε Προτασιακή Λογική

Η διαδικασία μετατροπής εκφράσεων Κατηγορηματικής Λογικής (ΚΛΠΤ) σε Προτασιακή μορφή (ΠΜ), όπως έχει ήδη αναφερθεί στο κεφάλαιο 2, είναι μια βηματική διαδικασία και για την εκμάθηση της αναπτύχθηκε το ευφές σύστημα FOLtoCF (Grivokostopoulou et al., 2014b; Grivokostopoulou et al., 2012c). Στα πλαίσια αυτής της ενότητας παρουσιάζουμε δυο μεθόδους για τον προσδιορισμό της δυσκολίας ασκήσεων μετατροπής εκφράσεων ΚΛΠΤ σε Προτασιακή μορφή. Στόχος των μεθόδων αυτών είναι να γίνει προσδιορισμός με αυτόματο τρόπο και με ακρίβεια του επιπέδου δυσκολίας των ασκήσεων, χωρίς την παρέμβαση του διδάσκοντα, καθώς και να παρέχει στους εκπαιδευόμενους ασκήσεις κατάλληλου επιπέδου. Οι δυο αυτές μέθοδοι βασίζονται σε διαφορετικού είδους προσεγγίσεις. Πιο συγκεκριμένα, η *Βασισμένη στην ΚΛΠΤ μέθοδος* (FOL-DEA: FOL-based Difficulty Estimation Approach) προσδιορίζει τη δυσκολία των ασκήσεων με βάση τη δομή και τα χαρακτηριστικά της έκφρασης ΚΛΠΤ, ενώ η *Βασισμένη στη Διαδικασία μέθοδος* (PROC-DEA: PROCess-based Difficulty Estimation Approach) βασίζεται στην πολυπλοκότητα των βημάτων της διαδικασίας μετατροπής της έκφρασης ΚΛΠΤ σε ΠΜ.

4.4.1 Προσδιορισμός δυσκολίας με βάση την δομή της πρότασης ΚΛΠΤ (FOL-DEA)

Αρχικά, έγινε ανάλυση των προτάσεων ΚΛΠΤ και προσδιορίστηκαν τα χαρακτηριστικά που συμβάλουν στον καθορισμό του επιπέδου δυσκολίας, σε συνεργασία με διδάσκοντες-εμπειρογνώμονες. Έτσι, καταλήξαμε στα ακόλουθα χαρακτηριστικά:

- Το πλήθος, το είδος, η σειρά και η θέση των ποσοδεικτών.
- Το πλήθος των συνεπαγωγών.
- Ο αριθμός των διαφορετικών συνδετικών $\{\wedge, \vee, \Rightarrow, \Leftrightarrow\}$
- Ο τύπος της άρνησης:
 - Άρνηση πριν από μια ομάδα ατομικών εκφράσεων.
 - Άρνηση πριν από ένα ποσοδείκτη.
- Αν αριστερά από την συνεπαγωγή υπάρχει ομάδα ατομικών εκφράσεων.

Στο πίνακα 4.7 παρουσιάζεται το σύνολο των κανόνων που προέκυψαν μετά τη συνεργασία με τους εμπειρογνώμονες για τον καθορισμό του επιπέδου δυσκολίας. Ο μηχανισμός κατατάσσει το επίπεδο δυσκολίας των ασκήσεων σε μία από τις παρακάτω πέντε κατηγορίες: α) Very-Easy (VE), β) Easy (E), γ) Medium (M), δ) Difficult (D) και ε) Advanced (A). Πιο συγκεκριμένα, κάθε στήλη του πίνακα αναπαριστά μια παράμετρο. Η PNF προσδιορίζει αν η έκφραση είναι σε Κανονική Μορφή Prenex, ενώ η LHS αν υπάρχει ομάδα ατομικών εκφράσεων αριστερά της συνεπαγωγής. Το «Υ» σημαίνει «Ναι», το «Ν» «Όχι», το «*» σημαίνει «οποιαδήποτε τιμή» και το «-» σημαίνει «αδιάφορο τιμής». Για παράδειγμα, ας θεωρήσουμε την έκφραση $\langle (\forall x)(\exists y)(human(x) \wedge food(y)) \Rightarrow eats(x,y) \rangle$. Μετά από την ανάλυση της προέκυψε ότι ο συνολικός αριθμός των ποσοδεικτών είναι δύο, υπάρχει ένας καθολικός και ένας υπαρξιακός ποσοδείκτης και η έκφραση είναι σε PNF μορφή. Επίσης, στην έκφραση υπάρχει μόνο μια συνεπαγωγή, δύο διαφορετικά διασυνδετικά ($\wedge, \Rightarrow$) και μια ομάδα ατομικών εκφράσεων αριστερά της συνεπαγωγής. Με βάση τις παραπάνω παραμέτρους ενεργοποιείται ο κανόνας 19 (πίνακας 4.7) και το επίπεδο δυσκολίας της άσκησης προσδιορίζεται ως “medium”.

Πίνακας 4.7 Κανόνες της Βασισμένης στην ΚΛΠΤ (FOL-DEA) προσέγγισης

Rules	Num of $\forall, \exists$	$\forall$	$\exists$	Order $\forall\exists$	PNF	Num Of $\Rightarrow$	Num of Different connecti ves	$\neg$ in front of group	$\neg$ in front of $\forall, \exists$	Group at LHS of $\Rightarrow$	Difficulty Level
1.	0	N	N	-	-	0	≤ 2	*	-	-	VE
2.	≤ 2	N	Y	-	*	0	< 1	*	*	-	VE
3.	1	N	Y	-	*	0	≥ 1	*	N	-	VE
4.	≤ 2	Y	N	-	*	0	< 1	*	*	-	VE
5.	1	Y	N	-	*	0	≥ 1	*	N	-	VE
6.	1	N	Y	-	*	0	≥ 1	*	Y	-	E
7.	1	Y	N	-	*	0	≥ 1	*	Y	-	E
8.	≥ 2	Y	Y	*	*	0	≥ 0	*	*	-	E
9.	≥ 2	Y	N	-	*	0	≥ 1	N	N	-	E
10.	2	N	Y	-	*	0	≥ 1	N	N	-	E
11.	≤ 2	Y	N	-	*	1	≥ 1	N	N	*	E
12.	≤ 2	N	Y	-	*	1	≥ 1	N	N	*	E
13.	0	N	N	-	-	1	≥ 0	*	-	Y	E
14.	1	Y	N	-	*	1	≥ 1	Y	Y	*	M
15.	1	N	Y	-	*	1	≥ 0	*	Y	*	M
16.	2	N	Y	*	*	1	≥ 1	N	*	*	M
17.	≥ 2	Y	N	-	Y	1	≥ 1	*	*	*	M
18.	≥ 2	Y	Y	*	*	1	≥ 1	*	*	N	M
19.	2	Y	Y	Y	*	1	≥ 1	*	N	Y	M
20.	3	N	Y	-	*	1	≥ 1	N	N	*	M
21.	3	Y	N	-	N	1	≥ 1	N	Y	*	M
22.	> 2	N	Y	-	*	1	≥ 1	*	Y	*	D
23.	≥ 2	Y	N	-	*	2	≥ 1	*	*	*	D
24.	> 2	Y	Y	*	*	1	≥ 1	*	*	*	D

25.	≥ 2	Y	Y	*	*	>1	≥ 1	*	*	*	A
26.	≥ 2	Y	N	-	*	>2	≥ 1	*	*	*	A
27.	≥ 2	N	Y	N	Y	>1	≥ 1	Y	Y	Y	A

4.4.2 Προσδιορισμός δυσκολίας με βάση τα βήματα της διαδικασίας μετατροπής (PROC-DEA)

Η δεύτερη μέθοδος που αναπτύξαμε βασίστηκε σε ενδογενή χαρακτηριστικά της διαδικασίας μετατροπής ΚΛΠΤ σε ΠΜ και ονομάζεται *Βασισμένη στη Διαδικασία (PROC-DEA: Process-based difficulty estimation approach)*. Τα χαρακτηριστικά που προσδιορίστηκαν, σε συνεργασία με εμπειρογνώμονες, για να μοντελοποιηθεί η διαδικασία είναι, ο αριθμός και η πολυπλοκότητα των βημάτων της διαδικασίας που απαιτούνται να εφαρμοστούν. Πιο συγκεκριμένα, προσδιορίστηκε η πολυπλοκότητα κάθε βήματος με βάση τον αριθμό των υποβημάτων που χρειάζεται να εφαρμοστούν. Τα βήματα και τα υποβήματα της διαδικασίας είναι τα εξής:

Βήμα 1: Απαλοιφή των συνεπαγωγών.

Βήμα 2: Περιορισμό της εμβέλεια των αρνήσεων.

Βήμα 3: Μετατροπή σε Κανονική Μορφή Prenex (PNF).

Βήμα 4: Απαλοιφή Ποσοδεικτών.

Βήμα 4α: Χρησιμοποίηση συναρτήσεων Skolem.

Βήμα 4β: Χρησιμοποίηση σταθερών Skolem.

Βήμα 4γ: Απαλοιφή Καθολικών Ποσοδεικτών.

Βήμα 5: Μετατροπή σε Συζευκτική Κανονική Μορφή (CNF).

Βήμα 6: Εξαγωγή προτάσεων και μετονομασία μεταβλητών.

Στον πίνακα 4.8 παρουσιάζονται οι κανόνες που σχεδιάστηκαν για την μέθοδο PROC-DEA. Το επίπεδο δυσκολίας κατηγοριοποιείται στις ίδιες πέντε κατηγορίες, όπως παραπάνω, και τα σύμβολα έχουν την ίδια σημασιολογία με αυτά του Πίνακα 4.7. Στο συμβολισμό “Step1/Times”, το “Times” αναφέρεται στον αριθμό εκτελέσεων του Βήματος 1.

Για παράδειγμα, θεωρούμε την ακόλουθη έκφραση ΚΛΠΤ: “ $(\forall x)(\exists y)(\text{human}(x) \wedge \text{food}(y)) \Rightarrow \text{eats}(x,y)$ ”. Οι τιμές των παραμέτρων του Πίνακα 4.8, για την μετατροπή της σε ΠΜ, έχουν ως εξής: Το Step1 της διαδικασίας εφαρμόζεται για να απαλοιφή της συνεπαγωγή ($\Rightarrow$) μία φορά και επομένως η τιμή της αντίστοιχης παραμέτρου είναι “Y/1”. Το Step2 έχει

τιμή “Υ”, καθώς μετά την απαλοιφή της συνεπαγωγής προκύπτει άρνηση (–) σε ομάδα ατόμων και η εμβέλεια της χρειάζεται να περιοριστεί. Το Step3 δεν εφαρμόζεται και η τιμή του είναι “N”. Επίσης, όσον αφορά την εφαρμογή του Step4 της διαδικασίας, το υποβήμα Step4a έχει ως τιμή “Υ” και επομένως το Step4b έχει “N” και Step4c “Υ”. Τέλος, το Step5 δεν εφαρμόζεται και έχει τιμή “N”, ενώ το Step6 εφαρμόζεται και έχει τιμή “Υ”. Με βάση τις παραπάνω παραμέτρους για την συγκεκριμένη πρόταση ενεργοποιείται ο κανόνας 21 και το σύστημα προσδιορίζει το επίπεδο δυσκολίας ως “Medium”.

Πίνακας 4.8 Κανόνες της προσέγγισης PROC-DEA.

Κανόνας	Step1/ Times	Step2	Step3	Step4a	Step4b	Step4c	Step5	Step6	Επίπεδο δυσκολίας
1.	N	*	N	N	N	N	*	*	VE
2.	N	N	*	N	Y	N	*	*	VE
3.	N	N	N	N	N	Y	*	*	VE
4.	N	Y	*	N	*	Y	N	N	VE
5.	N	N	Y	N	N	Y	N	*	VE
6.	N	N	Y	Y	N	Y	N	N	VE
7.	N	N	Y	Y	N	Y	N	Y	E
8.	N	Y	Y	N	N	Y	*	Y	E
9.	N	N	Y	N	N	Y	Y	*	E
10.	Y	N	N	N	N	Y	Y	Y	E
11.	N	*	N	N	Y	N	*	*	E
12.	Y/1	*	N	N	Y	N	N	*	E
13.	Y/1	*	N	N	N	Y	N	*	E
14.	Y/1	Y	Y	N	Y	N	N	*	E
15.	Y/1	N	Y	N	N	Y	N	*	E
16.	N	*	N	Y	N	Y	N	N	E
17.	N	N	Y	N	N	Y	N	Y	E
18.	N	N	Y	N	N	Y	Y	Y	M
19. e	Y/1	N	Y	N	Y	Y	N	Y	M
20.	Y/1	N	Y	Y	N	Y	N	*	M
21.	Y/1	Y	*	*	N	Y	*	Y	M
22.	Y/1	N	Y	N	Y	Y	N	*	M
23.	Y/1	Y	Y	*	N	Y	N	Y	D
24.	Y/1	N	N	Y	N	Y	N	*	D
25.	Y/2	*	*	*	N	Y	*	*	D
26.	Y/2	N	Y	Y	N	Y	N	N	A
27.	Y/2	Y	Y	*	Y	Y	*	*	A
28.	Y>2	Y	*	*	*	Y	*	*	A

Στην συνέχεια, παρουσιάζονται μερικά παραδείγματα εκφράσεων ΚΛΠΤ και τα αντίστοιχα επίπεδα δυσκολίας που προσδιορίστηκαν με βάση τις δύο παραπάνω προσεγγίσεις FOL-DEA και PROC-DEA καθώς και οι κανόνες που ενεργοποιήθηκαν σε κάθε περίπτωση.

Πίνακας 4.9 Παραδείγματα προτάσεων ΚΛΠΤ και τα αντίστοιχα επίπεδα δυσκολίας των δυο μεθόδων

Έκφραση ΚΛΠΤ	FOL-DEA	Κανόνας FOL-DEA	PROC-DEA	Κανόνας PROC-DEA
$\neg \text{cat}(\text{merilyn}) \wedge \neg \text{cat}(\text{jenny})$	Very Easy	1	Very Easy	1
$(\exists x) \text{pet}(x) \wedge \text{dog}(x)$	Very Easy	3	Very Easy	2
$(\forall x) (\text{car}(x) \wedge \text{fast}(x) \Rightarrow \text{expensive}(x))$	Easy	11	Easy	13
$\neg (\forall x) (\neg \text{man}(x) \vee \text{smart}(x))$	Easy	6	Easy	11
$(\exists x) \neg (\text{cat}(x) \vee \text{black}(x)) \vee \text{dog}(x)$	Very Easy	3	Easy	11
$(\forall x) (\exists y) (\text{human}(x) \wedge \text{food}(y)) \Rightarrow \text{eats}(x, y)$	Medium	19	Medium	21
$(\forall x) \text{city}(x) \rightarrow ((\exists y) \text{dog-catcher}(y, x) \wedge (\forall z)$ $(\text{dog}(z) \wedge \text{lives-in}(z, x)) \rightarrow \text{has-bitten}(z, y))$	Advanced	25	Advanced	28

4.4.3 Πειραματικά αποτελέσματα των δύο προσεγγίσεων

Για να αξιολογήσουμε την απόδοση και την ακρίβεια των δύο προσεγγίσεων FOL-DEA και PROC-DEA, χρησιμοποιήσαμε ένα σύνολο από 110 εκφράσεις ΚΛΠΤ, οι οποίες είχαν διαφορετικές συντακτικές δομές. Αρχικά, το σύνολο όλων των εκφράσεων αξιολογήθηκε από διδάσκοντες-εμπειρογνώμονες για να προσδιοριστεί το επίπεδο δυσκολίας της κάθε άσκησης. Έπειτα, το σύνολο όλων των εκφράσεων ΚΛΠΤ δόθηκε ως είσοδος στο σύστημα για προσδιοριστούν τα επίπεδα δυσκολίας τους από αυτό. Για την αξιολόγηση της ακρίβειας και της ποιότητας του αυτόματου μηχανισμού προσδιορισμού δυσκολίας συγκρίναμε την εκτίμηση των διδασκόντων και του συστήματος και υπολογίσαμε τις αντίστοιχες μετρικές αξιολόγησης. Τα πειραματικά αποτελέσματα της προσέγγισης FOL-DEA παρουσιάζονται στους πίνακες 4.10 και 4.11 αντίστοιχα.

Πίνακας 4.10 Πειραματικά αποτελέσματα της προσέγγισης FOL-DEA

Average accuracy	0.77
Precision	0.78
Recall	0.77

F-measure	0.77
-----------	------

Πίνακας 4.11 Πειραματικά αποτελέσματα για την προσέγγιση FOL-DEA για κάθε κατηγορία

	Precision	Recall	F-Measure
Very Easy	0.77	0.77	0.77
Easy	0.66	0.70	0.68
Medium	0.72	0.81	0.76
Difficulty	0.93	0.70	0.80
Advanced	1	1	1

Στους πίνακες 4.12 και 4.13 παρουσιάζονται τα πειραματικά αποτελέσματα της προσέγγισης PROC-DEA.

Πίνακας 4.12 Πειραματικά αποτελέσματα της προσέγγισης PROC-DEA

Average accuracy	0.82
Precision	0.83
Recall	0.82
F-measure	0.82

Πίνακας 4.13 Πειραματικά αποτελέσματα για την προσέγγιση PROC -DEA για κάθε κατηγορία

	Precision	Recall	F-Measure
Very Easy	1	0.81	0.90
Easy	0.74	0.93	0.82
Medium	0.73	0.73	0.73
Difficulty	0.79	0.75	0.77
Advanced	1	0.91	0.95

Τα αποτελέσματα της μελέτης δείχνουν μια πολύ καλή απόδοση και των δύο προσεγγίσεων. Από το σύνολο των 110 εκφράσεων ΚΛΠΤ, η FOL-DEA προσέγγιση προσδιόριζε σωστά την δυσκολία των 85, ενώ η PROC-DEA προσδιόρισε σωστά την δυσκολία 92 εκφράσεων. Η προσέγγιση FOL-DEA προσδιόρισε σωστά το επίπεδο δυσκολίας στο 77% των περιπτώσεων, ενώ η προσέγγιση PROC-DEA προσδιόρισε σωστά το 82% των περιπτώσεων. Παρατηρώντας τα αποτελέσματα, η προσέγγιση FOL-DEA εμφανίζεται να έχει καλύτερη απόδοση στην εκτίμηση της δυσκολίας ασκήσεων που ανήκουν στις κατηγορίες “Difficult” και “Advanced”, ενώ η PROC-DEA φαίνεται να είναι καλύτερη στην εκτίμηση ασκήσεων που ανήκουν στις κατηγορίες “Very Easy” και “Easy”. Επιπροσθέτως, έγινε υπολογισμός της μετρικής Cohen’s kappa k η οποία υπολογίζει τον βαθμό συμφωνίας των εκτιμήσεων της δυσκολίας μεταξύ του διδάσκοντα και του συστήματος στο σύνολο των 110 ασκήσεων για τις δύο προσεγγίσεις FOL-DEA και PROC-DEA. Τα αποτελέσματα παρουσιάζονται στον πίνακα 4.14.

Πίνακας 4.14 Cohen’s kappa για τις προσεγγίσεις FOL-DEA και PROC-DEA

	Automatic Rater FOL-DEA	Automatic Rater PROC-DEA
Human Rater	0.705	0.709

Τα αποτελέσματα της μετρικής k δείχνουν μια καλή συμφωνία μεταξύ του διδάσκοντα-εμπειρογνώμονα και του συστήματος και για τις δύο προσεγγίσεις. Στην συνέχεια, στον πίνακα 4.15 παρουσιάζονται ενδεικτικές μη σωστές εκτιμήσεις της δυσκολίας ασκήσεων για κάθε προσέγγιση.

Πίνακας 4.15 Παραδείγματα μη σωστών προσδιορισμός Δυσκολίας Ασκήσεων

ΚΛΠΤ	FOL-DEA	PROC-DEA	Ειδικός
$(\exists x) \neg (a(x) \vee b(x)) \vee c(x)$	Very easy	Easy	Easy
$(\forall x)(\forall y)(a(x) \wedge b(y)) \Rightarrow (\exists z)(\exists w)(c(z, x) \wedge d(w, y))$	Advanced	Difficult	Advanced

4.4.4 Συνδυαστική προσέγγιση μέσω ευφυούς συστήματος κανόνων

Με βάση τα παραπάνω πειραματικά αποτελέσματα που προέκυψαν για να βελτιώσουμε περαιτέρω την ακρίβεια των προσεγγίσεων σχεδιάσαμε και αναπτύξαμε ένα ευφύες σύστημα για τον προσδιορισμό της δυσκολίας (Difficulty estimation intelligent system-DEIS) της διαδικασίας ΚΛΠΤ σε ΠΜ, με βάση τον συνδυασμό των δυο προσεγγίσεων FOL-DEA και PROC-DEA. Πιο συγκεκριμένα, γίνεται συνδυασμός των εξόδων των δυο προσεγγίσεων για να προσδιοριστεί η τελική εκτίμηση της δυσκολίας. Η συνδυασμένη προσέγγιση υλοποιείται μέσω ενός ευφυούς συστήματος κανόνων το οποίο δέχεται ως είσοδο μια έκφραση ΚΛΠΤ και έχει ως έξοδο το επίπεδο δυσκολίας της άσκησης. Το ευφύες σύστημα έχει ως στόχο να να παρέχει με συνέπεια και μεγαλύτερη ακρίβεια το επίπεδο δυσκολίας.

Η συνδυαστική προσέγγιση περιέχει τα αποτελέσματα των δυο προσεγγίσεων. Ο συνδυασμός βασίστηκε στα αποτελέσματα των δυο προσεγγίσεων και πιο συγκεκριμένα στο γεγονός ότι η προσέγγιση FOL-DEA έχει καλύτερα αποτελέσματα στην εκτίμηση της δυσκολίας ασκήσεων στις κατηγορίες “Difficult” και “Advanced”, ενώ αντίστοιχα η προσέγγιση PROC-DEA έχει καλύτερα αποτελέσματα στις κατηγορίες “Very Easy” και “Easy”. Για να βελτιωθεί περαιτέρω η ακρίβεια, δημιουργήσαμε ένα τρίτο σύνολο κανόνων για την συνδυαστική προσέγγιση. Στον πίνακα 4.16 παρουσιάζεται το σύνολο των κανόνων που προέκυψαν από πειραματικά αποτελέσματα και προσδιορίζουν το τελικό επίπεδο δυσκολίας. Οι κανόνες αυτοί αναπαριστούν τις περιπτώσεις εκείνες όπου οι δυο προσεγγίσεις δίνουν ως έξοδο διαφορετικό επίπεδο δυσκολίας, ενώ στις περιπτώσεις που δίνουν ως έξοδο το ίδιο συμπέρασμα, το επίπεδο δυσκολίας δεν αναπροσαρμόζεται αλλά είναι παραμένει το ίδιο.

Πίνακας 4.16 Συνδυαστική προσέγγιση για τον προσδιορισμό των επιπέδων δυσκολίας

Κανόνας	Προσδιορισμός δυσκολίας από FOL-DEA	Προσδιορισμός δυσκολίας από PROC-DEA	Τελικός προσδιορισμός Δυσκολίας
1	Very easy	Easy	Easy
2	Easy	Very Easy	Very Easy
3	Easy	Medium	Medium
4	Medium	*	Medium

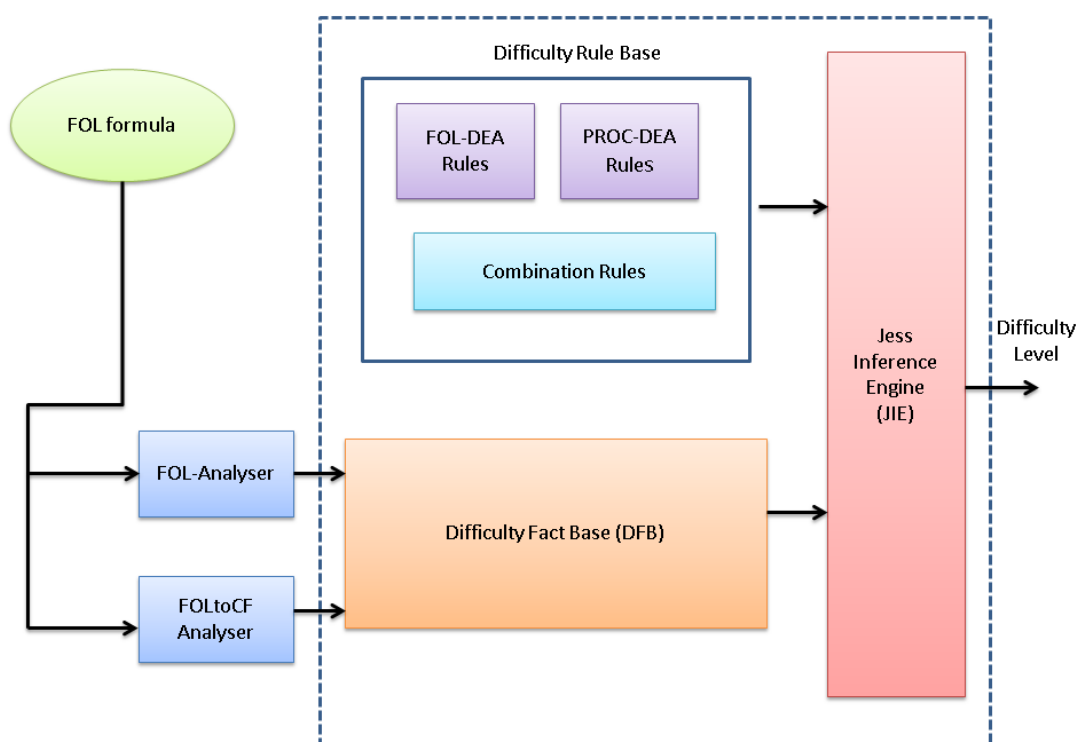
5	Difficulty	*	Difficulty
6	Advanced	Difficulty	Advanced

4.4.4.1 Αρχιτεκτονική και τρόπος λειτουργίας

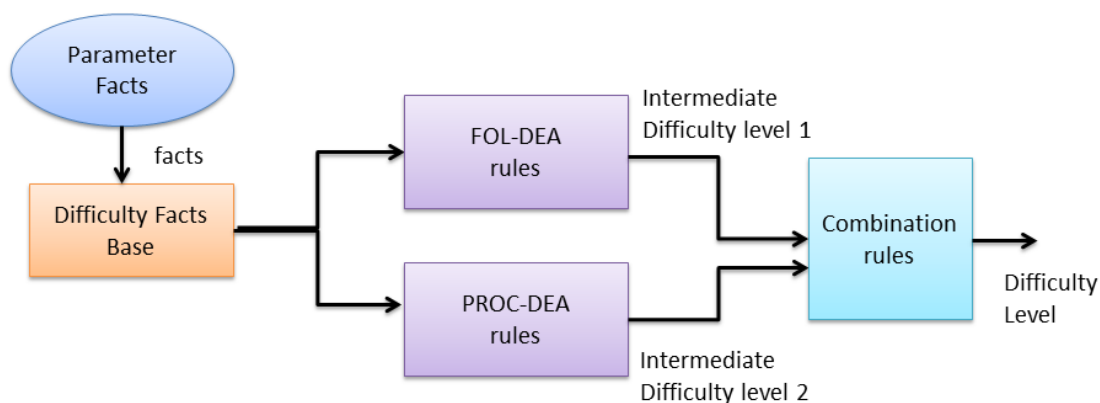
Η αρχιτεκτονική του ευφυούς συστήματος προσδιορισμού δυσκολίας (*Difficulty Estimation Intelligent System-DEIS*) αναπαριστάται στην εικόνα 4.5. Το σύστημα αποτελείται από τις ακόλουθες μονάδες: τον *μηχανισμό εξαγωγής συμπερασμάτων* (*Inference engine JIE*), που είναι υπεύθυνος για την εξαγωγή των συμπερασμάτων, την βάση γεγονότων δυσκολίας (*Difficulty Fact Base, DFB*) και την βάση κανόνων (*Difficulty Rule Base-DRB*). Πιο συγκεκριμένα, η βάση κανόνων αποτελείται από τρεις ομάδες (modules) κανόνων: FOL-DEA (οι κανόνες του πίνακα 4.7), PROC-DEA (οι κανόνες του πίνακα 4.8) και οι *συνδυαστικοί κανόνες* (*combination rules*), που είναι οι κανόνες του πίνακα 4.16. Επίσης, υπάρχει ένας επιπρόσθετος κανόνας στην περίπτωση ίδιου συμπεράσματος και από τις δυο προσεγγίσεις.

Η διαδικασία προσδιορισμού του επιπέδου δυσκολίας είναι η ακόλουθη:

1. Αρχικά, γίνεται εισαγωγή της έκφρασης ΚΛΙΠΤ.
2. Ο FOL Analyser αναλύει και προσδιορίζει τις τιμές των παραμέτρων της προσέγγισης FOL-DEA και γίνεται εισαγωγή των αντίστοιχων γεγονότων στην βάση γεγονότων (DFB).
3. Ο FOLtoCF Analyser αναλύει και προσδιορίζει τις τιμές των παραμέτρων της προσέγγισης PROC-DEA με βάση τα βήματα της διαδικασίας μετατροπής και εισάγονται τα αντίστοιχα γεγονότα στην DFB.
4. Ενεργοποιείται η Jess Inference Engine (JIE) και προσδιορίζεται το επίπεδο δυσκολίας ως εξής:
 - 4.1 Εξάγεται ένα ενδιάμεσο επίπεδο δυσκολίας με βάση τα γεγονότα και τους κανόνες της προσέγγισης FOL-DEA.
 - 4.2 Εξάγεται ένα ενδιάμεσο επίπεδο δυσκολίας βασισμένο στα γεγονότα και τους κανόνες της προσέγγισης PROC-DEA
 - 4.3 Αν δύο προσεγγίσεις έχουν ως έξοδο το ίδιο επίπεδο δυσκολία, τότε η τελική έξοδος του συστήματος αποτελεί το κοινό επίπεδο δυσκολίας. Σε αντίθετη περίπτωση, εξάγεται το επίπεδο δυσκολίας με βάση τους *συνδυαστικούς κανόνες*.



Εικόνα 4.4. Αρχιτεκτονική Ευφυούς Συστήματος προσδιορισμού δυσκολίας (DEIS) FOLtoCF ασκήσεων.



Εικόνα 4.5. Διάγραμμα συμπερασμού

Στην εικόνα 4.6 παρουσιάζεται το *διάγραμμα ροής συμπερασμού (inference flow)* του ευφυούς συστήματος κανόνων.

Στην εικόνα 4.7 παρουσιάζεται ένα παράδειγμα έκφρασης ΚΛΠΤ, $(\forall x)(\exists y)(human(x) \wedge food(y)) \Rightarrow eats(x,y)$, για την οποία έχουν προσδιοριστεί οι τιμές των παραμέτρων δυσκολίας της προσέγγισης FOL-DEA, μέσω του FOL-Analyser (βήμα 2 της παραπάνω

διαδικασίας), και της προσέγγισης PRO-DEA, μέσω του FOLtoCF Analyser (βήμα 3 της διαδικασίας).

FOL-Analyser	FOL to CF Analyser
FOL formula: $(\forall x)(\exists y)(\text{human}(x) \cap \text{food}(y) \Rightarrow \text{eats}(x,y))$	FOL formula: $(\forall x)(\exists y)(\text{human}(x) \cap \text{food}(y) \Rightarrow \text{eats}(x,y))$
Difficulty Estimation Parameters:	Difficulty Estimation Parameters :
Number of Quantifiers: 2	Step1: yes Times: 1
Type of Quantifiers:	Step2: no
$\forall$ = yes	Step3: no
$\exists$ = yes	Step4a: no
order $\forall \exists$: no	Step4b: yes
PNF: yes	Step4c: yes
Number of Implication: 1	Step5: no
Number of Different connectives: 2	Step6: yes
Negation in front of Quantifiers: no	
Negation in front of group: no	
Group at LHS of $\Rightarrow$: no	

Εικόνα 4.6. Παράδειγμα προσδιορισμού τιμών παραμέτρων δυσκολίας μέσω του FOL-Analyser και του FOLtoCF Analyser

Στην συνέχεια παρουσιάζονται παραδείγματα κανόνων των δύο προσεγγίσεων, καθώς και ένας συνδυαστικός κανόνας.

Κανόνας FOL-DEA:

```
(defrule rule
  ?df1 <- (fol {q<=2} (un no) (ex yes) (ord nil)
    (pnf yes|no) (imp 0) {con < 1}
    (n_group yes|no) (n_quant yes|no) (group nil)
=>
  (retract ?df1)
  (print-level-message "very-easy"))
```

Κανόνας PROC-DEA

```
(defrule rule
  ?df2 <- (Step (Step1 no) (numb 0) (Step2 yes| no)
    (Step3 no) (Step4a no) (Step4b no)
    (Step4c no) (Step5 yes|no) (Step6 yes|no))
=>
  (retract ?df2)
  (print-level-message "very-easy"))
```

Συνδυαστικός κανόνας:

```
(defrule rule
```

```
?df <- (difficulty (df1 easy) (df2 very-easy))
=>
(retract ?df)
(print-level-message "very-easy")
```

4.4.5 Πειραματικά Αποτελέσματα Συστήματος Προσδιορισμού Δυσκολίας

Για να αξιολογήσουμε την απόδοση του ευφυούς συστήματος προσδιορισμού δυσκολίας, χρησιμοποιήσαμε τον ίδιο σύνολο των 110 προτάσεων που χρησιμοποιήσαμε και στις δύο προηγούμενες προσεγγίσεις και υπολογίσαμε τις ίδιες μετρικές αξιολόγησης. Τα αποτελέσματα της πειραματικής μελέτης παρουσιάζονται στους πίνακες 4.17 και 4.18. Από το σύνολο των 110 διαφορετικών ΚΛΠΤ εκφράσεων που χρησιμοποιήθηκαν στα πλαίσια της μελέτης, το DEIS εκτίμησε σωστά το επίπεδο δυσκολίας των 99 εκφράσεων. Επομένως, η γενική ορθότητα (accuracy) του DEIS είναι 0.9, η οποία είναι καλύτερη από τις δυο άλλες προσεγγίσεις. Επίσης, από τα αποτελέσματα του DEIS (πίνακας 4.17), είναι σαφές ότι η συνδυασμένη προσέγγιση έχει καλύτερα αποτελέσματα για όλες τις μετρικές και τις κλάσεις.

Πίνακας 4.17 Πειραματικά αποτελέσματα DEIS

Average accuracy	0.90
Precision	0.92
Recall	0.90
F-measure	0.90

Πίνακας 4.18 Μετρικές Αξιολόγησης του DEIS

	Precision	Recall	F-Measure
Very Easy	1	0.88	0.94
Easy	0.87	0.96	0.91
Medium	0.78	0.96	0.86

Difficulty	1	0.71	0.83
Advanced	1	1	1

Τα αποτελέσματα δείχνουν ότι το DEIS βελτιώνει την απόδοση τους σε όλες τις κατηγορίες και έχει μια καθολικά καλύτερη απόδοση σε σύγκριση με τις δυο επιμέρους προσεγγίσεις, FOL-DEA και PROC-DEA. Πιο συγκεκριμένα, το DEIS παρουσιάζει μια άριστη ή σχεδόν τέλεια απόδοση στις κατηγορίες “Very Easy”, “Difficulty”, “Advanced” και επίσης παρουσιάζει μια πολύ καλή απόδοση στις κατηγορίες “Easy” και “Medium”, όπου βελτιώνει αισθητά την απόδοση των δύο επιμέρους συστημάτων. Επίσης, υπολογίστηκε η μετρική Cohen’s Kappa για την αξιοπιστία της συμφωνίας μεταξύ του προσδιορισμού της δυσκολίας από τον διδάσκοντα και του DEIS. Τα αποτελέσματα έδειξαν ότι υπάρχει σχεδόν τέλεια συμφωνία στον προσδιορισμό της δυσκολίας μεταξύ του διδάσκοντα και του DEIS καθώς η μετρική kappa ήταν $k = 0.976$ (95% confidence interval, $p < .0005$). Επίσης, από τα αποτελέσματα της μετρικής kappa για το DEIS δείχνουν ότι βελτιώθηκε αρκετά η συμφωνία μεταξύ του διδάσκοντα και του συστήματος σε σχέση με τις άλλες προσεγγίσεις και είναι σχεδόν τέλεια.

Τέλος, το βασικότερο χαρακτηριστικό του DEIS είναι ότι συνδυάζει με αποτελεσματικό τρόπο τις δυο προσεγγίσεις, ώστε να διατηρηθεί η αρχική καλή επίδοσή του μηχανισμού σε ορισμένες κατηγορίες ασκήσεων και να βελτιωθεί η απόδοσή του στις κατηγορίες που είχαν χαμηλότερη ακρίβεια. Η χρήση ευφυούς συστήματος που συνδυάζει τα αποτελέσματα των επιμέρους συστημάτων φαίνεται να αποτελεί μια καλή τεχνική για την βελτίωση της απόδοσής τους.

4.5 Νευρο-ασαφείς και Νευρο-συμβολικές προσεγγίσεις για τον προσδιορισμό της δυσκολίας ασκήσεων αλγορίθμων αναζήτησης

Στο πεδίο της τεχνητής νοημοσύνης (TN), ένα βασικό πεδίο αποτελούν οι αλγόριθμοι αναζήτησης. Για να βοηθήσουμε τους εκπαιδευόμενους στην εκμάθηση των αλγορίθμων αναζήτησης, αναπτύχθηκε το Ευφύες Σύστημα Διδασκαλίας Τεχνητής Νοημοσύνης (AITS) το οποίο παρουσιάστηκε αναλυτικά στο Κεφάλαιο 2. Μια βασική συνιστώσα των εκπαιδευτικών συστημάτων είναι να προσαρμόζουν τις εκπαιδευτικές δραστηριότητες και τις ασκήσεις στο επίπεδο των γνώσεων και των επιδόσεων των εκπαιδευόμενων. Για το σκοπό

αυτό, ο προσδιορισμός του επίπεδου δυσκολίας των ασκήσεων είναι ιδιαίτερα σημαντικός και απαραίτητο να γίνεται με ακρίβεια και με συνέπεια.

Σε αυτή την ενότητα παρουσιάζουμε δύο προσεγγίσεις, μια νευρο-ασαφή (neuro-fuzzy) και μια νευρο-συμβολική προσέγγιση (neuro-symbolic), η οποία βασίζεται στους νευροκανόνες (neurules), για την εκτίμηση της δυσκολίας ασκήσεων σχετικών με το πεδίο των αλγορίθμων αναζήτησης. Οι δύο αυτές προσεγγίσεις, έχουν ως στόχο να παρέχουν με αυτόματο τρόπο, με συνέπεια και με αντικειμενικότητα το επίπεδο δυσκολίας των ασκήσεων. Αρχικά, σε συνεργασία με τους διδάσκοντες-εμπειρογνώμονες έγινε ο καθορισμός των χαρακτηριστικών που επηρεάζουν την πολυπλοκότητα των ασκήσεων. Στην συνέχεια, έγινε ανάλυση των ασκήσεων και εξήχθησαν οι κατάλληλες παράμετροι και έπειτα με βάση τις παραμέτρους αυτές έγινε εκτίμηση της δυσκολίας των ασκήσεων από την νευρο-ασαφή και η νευρο-συμβολική προσέγγιση.

4.5.1 Προσδιορισμός Παραμέτρων Ασκήσεων

Για τον προσδιορισμό της δυσκολίας των ασκήσεων, ένας βασικός παράγοντας είναι ο ακριβής προσδιορισμός των παραμέτρων που συμβάλουν στην πολυπλοκότητα και τη δυσκολία των ασκήσεων αλγορίθμων. Για τον σκοπό αυτό ζητήθηκε η συμβολή διδασκόντων-ειδικών με εμπειρία στο πεδίο της Τεχνητής Νοημοσύνης και των αλγορίθμων αναζήτησης. Στην συνέχεια, οι ασκήσεις αλγορίθμων αναζήτησης αναλύθηκαν και εξάχθηκαν οι κατάλληλες παράμετροι που σχετίζονται με χαρακτηριστικά της άσκησης. Επίσης, λόγω της φύσης των ασκήσεων εξάχθηκαν παράμετροι που σχετίζονται με την τοπολογία και με την πολυπλοκότητα των δεντρικών δομών, που εφαρμόζονται στους αλγορίθμους αναζήτησης. Οι παράμετροι που προσδιορίστηκαν είναι οι ακόλουθοι:

- Οι παράμετροι εισόδου: το πλήθος των κόμβων του δέντρου (Number of nodes), ο μέσος όρος των παιδιών των κόμβων (Average number of the children), το βάθος του δέντρου (Depth of the tree), το μήκος της λύσης (Solution's length), δηλαδή το πλήθος των κόμβων που περιέχει το μονοπάτι της λύσης, που ενώνει την αρχική κατάσταση με μία τελική κατάσταση (στόχος).
- Οι παράμετροι εξόδου: Το επίπεδο δυσκολίας της άσκησης κατηγοριοποιείται στις ακόλουθες λεκτικές τιμές: “very easy”, “easy”, “medium”, “difficulty”, “advanced”.

Στον πίνακα 4.7 παρουσιάζονται οι μεταβλητές εισόδου και το αντίστοιχο πεδίο τιμών που παίρνουν.

Εικόνα 4.7 Μεταβλητές Εισόδου

Μεταβλητή	Πεδίο Τιμών
Number of nodes	3-40
Average number of the children	2-6
Depth of the tree	3-10
Solution's length	1-40

Κάθε μία από τις παραπάνω μεταβλητές συσχετίζεται με διαφορετικές πτυχές της πολυπλοκότητας της άσκησης. Η μεταβλητή ‘πλήθος των κόμβων’ αναπαριστά το συνολικό πλήθος των κόμβων που περιέχει η δεντρική αναπαράσταση της άσκησης. Η μεταβλητή ‘μέσος όρος των παιδιών’ των κόμβων αντιπροσωπεύει και μοντελοποιεί τις δυνατές επιλογές μετάβασης που μπορούν να γίνουν από τον εκπαιδευόμενο σε κάθε κόμβο της άσκησης και παίρνει συνεχείς τιμές από 2 έως και 6. Η μεταβλητή ‘βάθος του δέντρου’ αναπαριστά το μέγιστο βάθος του δέντρου και παίρνει ακέραιες τιμές από 3 έως 10. Επίσης, η μεταβλητή ‘μήκος της λύσης’ αναπαριστά το πλήθος των κόμβων από την αρχική κατάσταση μέχρι την τελική κατάσταση, δηλαδή το μονοπάτι (path) της λύσης, και παίρνει ακέραιες τιμές με μέγιστη τιμή 40. Η παράμετρος αυτή αναπαριστά τα βήματα που ακολούθησε ο εκπαιδευόμενος για να προσδιορίσει τη λύση και συσχετίζεται με την πολυπλοκότητα του γραφήματος του δέντρου και επίσης σε συνδυασμό και με άλλες παραμέτρους, μπορεί να αποτελεί ένα σημαντικό παράγοντα για τον προσδιορισμό του επιπέδου δυσκολίας των ασκήσεων.

4.5.2 Νευρο-ασαφής Προσέγγιση

Τα νευρο-ασαφή συστήματα συνδυάζουν την τεχνολογία των νευρωνικών δικτύων και της ασαφούς λογικής. Σκοπός του συνδυασμού αυτού είναι η αξιοποίηση των πλεονεκτημάτων και των δύο μεθόδων και κυρίως της υπολογιστική δύναμης των νευρωνικών δικτύων και της επικοινωνίας υψηλού επιπέδου με τον χρήστη μέσω των λεκτικών μεταβλητών που χρησιμοποιεί η ασαφής λογική. Τα νευρο-ασαφή συστήματα συνδυάζουν τις τεχνικές των νευρωνικών δικτύων και των συστημάτων ασαφούς συμπερασμού. Το πιο γνωστό ίσως μοντέλο τέτοιου συνδυασμού είναι το ANFIS (Adaptive Neuro-Fuzzy Inference System). Ο στόχος του σχεδιασμού τέτοιων συστημάτων δεν είναι η μοντελοποίηση του ίδιου του προβλήματος κατασκευάζοντας ένα αναλυτικό μαθηματικό μοντέλο, αλλά η αξιοποίηση της γνώσης και της εμπειρίας ενός ειδικού. Το ANFIS συνδυάζοντας πτυχές των νευρωνικών δικτύων με πτυχές των ασαφών συστημάτων συμπερασμού (Fuzzy Inference Systems-FIS),

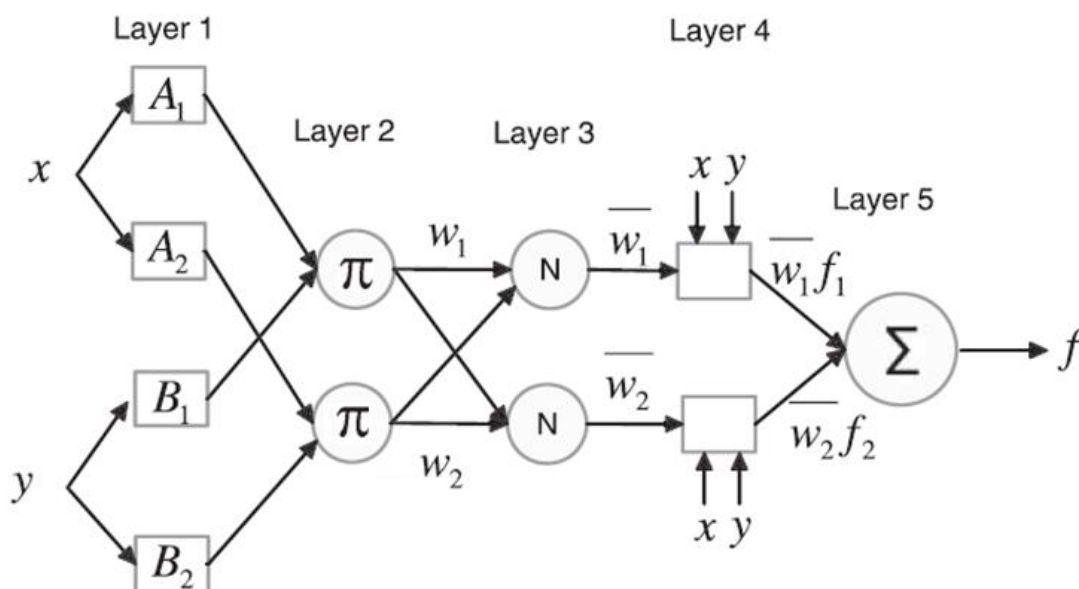
συνδυάζει το πλεονέκτημα της εύκολης εφαρμογής με την ικανότητα εκπαίδευσης (Jang, 1993). Ένα FIS προσομοιώνει τη συμπεριφορά κανόνων *if-then* μέσω της γνώσης των ειδικών-εμπειρογνομώνων ή με τη βοήθεια μίας διαθέσιμης βάσης δεδομένων.

Για λόγους απλότητας, υποθέτουμε ότι ένα ασαφές σύστημα συμπερασμού έχει δύο εισόδους, x , y , και μία έξοδο f . Η βάση των κανόνων περιέχει δύο ασαφείς κανόνες *if-then* για το ασαφές μοντέλο Takagi και Sugeno, οι οποίοι εκφράζονται ως εξής:

Rule1: *if x is A_1 and y is B_1 then $f_1 = p_1x + q_1y + r_1$*

Rule2: *if x is A_2 and y is B_2 then $f_2 = p_2x + q_2y + r_2$*

όπου A_i και B_i είναι τα ασαφή σύνολα, f_i είναι οι έξοδοι που ορίζονται εντός της ασαφούς περιοχής από τον ασαφή κανόνα, και p_i , q_i , r_i είναι οι παράμετροι που καθορίζονται κατά τη διάρκεια της εκπαίδευσης (training).

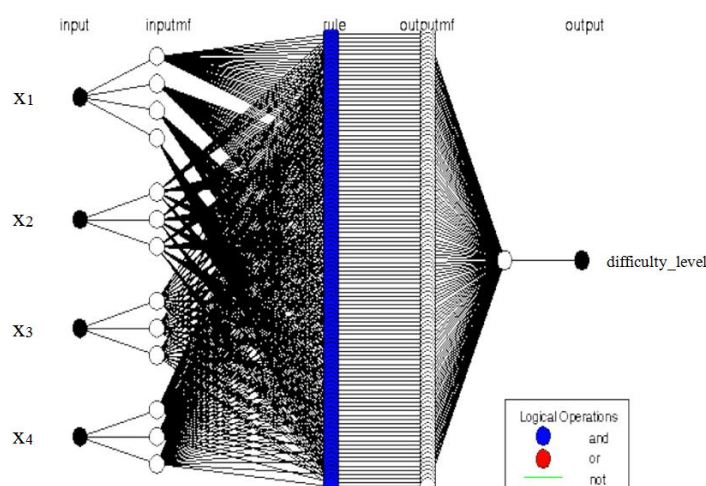


Εικόνα 4.8 Αρχιτεκτονική ANFIS

Στην εικόνα 4.8, οι κυκλικοί είναι οι σταθεροί κόμβοι, ενώ οι τετράγωνοι κόμβοι με παραμέτρους είναι οι προσαρμοστικοί κόμβοι. Η αρχιτεκτονική του προσαρμοστικού νευρο-ασαφούς συστήματος συμπερασμού χωρίζεται σε 5 επίπεδα (Jang, 1993).

Ο νευρο-ασαφής μηχανισμός προσδιορισμού της δυσκολίας παίρνει ως είσοδο ειδικά χαρακτηριστικά των ασκήσεων και δίνει ως έξοδο το επίπεδο δυσκολίας των ασκήσεων.

Το μοντέλο ANFIS που δημιουργήσαμε αποτελείται από πέντε επίπεδα και από τέσσερις μεταβλητές εξόδου x_1, x_2, x_3, x_4 , που αναπαριστούν τον πλήθος των κόμβων, το μέσο όρο των παιδιών, το βάθος του δέντρου και το μήκος της λύσης αντίστοιχα καθώς επίσης και μια μεταβλητή εξόδου που αναπαριστά το επίπεδο δυσκολίας (*difficulty_level*) όπως παρουσιάζεται στην εικόνα 4.9.



Εικόνα 4.9 Μοντέλο ANFIS προσδιορισμού Δυσκολίας Ασκήσεων

Στην περίπτωση μας, το επίπεδο 1, που είναι το επίπεδο εισόδου του νευρο-ασαφούς μοντέλου, έχει 4 νευρώνες που αντιπροσωπεύουν τις τέσσερις μεταβλητές εισόδου. Το επίπεδο εισόδου στέλνει τις εξωτερικές ξερές τιμές (*crisp values*), απ' ευθείας στο επόμενο στρώμα (επίπεδο 2), το οποίο είναι το επίπεδο ασαφοποίησης, και οι *crisp* τιμές από το πρώτο επίπεδο μετασχηματίζονται στις κατάλληλες λεκτικές ασαφείς τιμές (*low, average ή high*). Κάθε κόμβος είναι προσαρμοστικός με τη λειτουργία του υπολογισμού του βαθμού των συναρτήσεων συμμετοχής κάθε εισόδου.

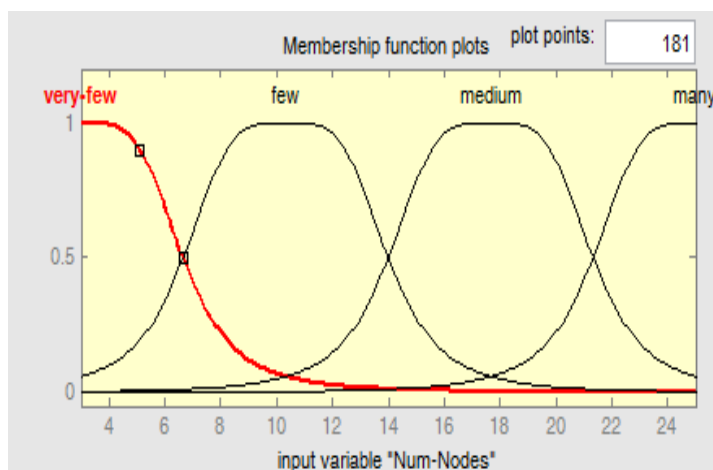
Οι νευρώνες του επιπέδου 2 αναπαριστούν ένα ασαφές σύνολο για κάθε μία από τις λεκτικές μεταβλητές εισόδου. Κάθε κόμβος στο επίπεδο είναι ένας σταθερός (*fixed*) κόμβος και έχει ως ρόλο του την εκτέλεση του πολλαπλασιασμού των σημάτων εξόδου του πρώτου επιπέδου. Η κάθε έξοδος αυτού του επιπέδου αντιστοιχεί στην ισχύ ενεργοποίησης (*firing strength*) του κάθε κανόνα.

Οι νευρώνες του επιπέδου 3 (επίπεδο ασαφών κανόνων) αναπαριστούν τους ασαφείς κανόνες (*fuzzy rules*) ($R_1, \dots, R_{120}$) και κάθε κόμβος υπολογίζει το λόγο του βαθμού ενεργοποίησης (*firing strength*). Οι έξοδοι αυτού του επιπέδου ονομάζονται κανονικοποιημένοι βαθμοί ενεργοποίησης (*normalized firing strengths*).

Στο επίπεδο 4 (έξοδος επιπέδου συναρτήσεων συμμετοχής) κάθε κόμβος είναι ένας προσαρμοστικός κόμβος και δημιουργεί έναν απλό γραμμικό συνδυασμό των εισόδων του συστήματος και ενός συνόλου παραμέτρων με τις εξόδους του επιπέδου 3, και μετά υπολογίζει τη συνεισφορά του κάθε κανόνα προς τη συνολική έξοδο. Οι παράμετροι σε αυτό το επίπεδο ορίζονται ως επακόλουθοι παράμετροι (consequent parameters).

Το επίπεδο 5 (επίπεδο ασαφοποίησης) αποτελείται από ένα μόνο σταθερό κόμβο και υπολογίζει τη συνολική έξοδο σαν το συνολικό άθροισμα όλων των εισερχόμενων σημάτων – συναρτήσεων.

Στην εικόνα 4.10, παρουσιάζονται οι συναρτήσεις συμμετοχής (membership functions) για την παράμετρο εισόδου που αναπαριστά το πλήθος των κόμβων σε μια δεντρική δομή. Η παράμετρος εισόδου μπορεί να διακριθεί σε τέσσερα ασαφή σύνολα (λεκτικές τιμές), τα οποία είναι: “very-few”, “few”, “medium” και “many”.



Εικόνα 4.10 Συναρτήσεις Συμμετοχής για την μεταβλητή πλήθος των κόμβο του δέντρου

Στο μοντέλο ANFIS οι έξοδοι σε κάθε επίπεδο αποτελούν μέρος των εισόδων του επόμενου επιπέδου. Η εκπαίδευση του ANFIS βασίζεται σε έναν υβριδικό αλγόριθμο εκμάθησης (Kakar et al., 2005 ; Jang, 1993). Ο αλγόριθμος περιλαμβάνει έναν προς τα εμπρός έλεγχο (forward pass) και ένα προς τα πίσω πέρασμα (backward pass). Στον προς τα εμπρός έλεγχο ο αλγόριθμος χρησιμοποιεί την μέθοδο των ελαχίστων τετραγώνων για να αναγνωρίσει τις μεταβλητές συμπερασμού παραμέτρους του επιπέδου 4. Ο υβριδικός αλγόριθμος έχει αποδειχθεί ότι είναι πολύ αποτελεσματικός σε μοντέλα ANFIS. Στο πλαίσιο της μελέτης μας, το μοντέλο ANFIS εκπαιδεύτηκε με βάση το σύνολο δεδομένων των ασκήσεων που προσδιορίστηκε η δυσκολία τους από τους εμπειρογνώμονες. Μετά την εκπαίδευση του, το νευρο-ασαφές μοντέλο είναι πλήρες και μπορεί να χρησιμοποιηθεί για την ταξινόμηση νέων ασκήσεων.

4.5.3 Νευρωνικοί κανόνες

Οι νευρωνικοί κανόνες (Neuralrules) είναι ένα είδος υβριδικών κανόνων. Κάθε νευρωνικός κανόνας μπορεί να θεωρηθεί ως μία νευρωνική μονάδα adaline (Εικόνα 4.11.α). Οι εισόδοι C_i ($i = 1 \dots n$) της μονάδας είναι οι συνθήκες του κανόνα. Σε κάθε συνθήκη C_i ανατίθεται ένας αριθμός sf_i , που καλείται *παράγοντας σπουδαιότητας* (significance factor) και αντιστοιχεί στο βάρος της αντίστοιχης εισόδου της μονάδας adaline. Επιπλέον, σε κάθε κανόνα ανατίθεται ένας αριθμός sf_0 , που καλείται *πολωτικός παράγοντας* (bias factor), και αντιστοιχεί στο βάρος της εισόδου πόλωσης της μονάδας. Κάθε είσοδος παίρνει τιμή από το σύνολο διακριτών τιμών

$$C_i = \begin{cases} 1, & \text{αλήθης συνθήκη} \\ -1, & \text{ψευδής συνθήκη} \\ 0, & \text{άγνωστη συνθήκη} \end{cases}$$

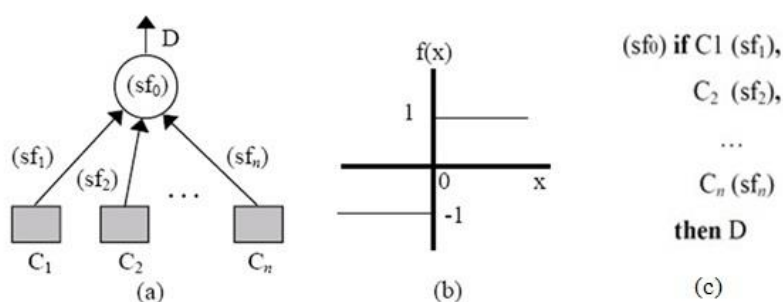
Η έξοδος D , η οποία αναπαριστά το συμπέρασμα (απόφαση) του κανόνα, υπολογίζεται από τους ακόλουθους τύπους:

$$D = f(\mathbf{a}), \mathbf{a} = sf_0 + \sum_{i=1}^n sf_i C_i$$

όπου $\mathbf{a}$ είναι η τιμή ενεργοποίησης (activation value) και $f(x)$ είναι η συνάρτηση ενεργοποίησης (activation function), η οποία είναι μία συνάρτηση κατωφλίου:

$$f(a) = \begin{cases} 1, & \text{αν } a \geq 0 \\ -1, & \text{διαφορετικά} \end{cases}$$

Συνεπώς, η έξοδος μπορεί να πάρει μια από τις δυο τιμές “-1” και “1”, αναπαριστώντας την αποτυχία ή την επιτυχία ενός κανόνα.



Εικόνα 4.11 (α) Ένα νευρωνικό μονάδα adaline (β) Συνάρτηση ενεργοποίησης (γ) Λεκτική αναπαράσταση νευρωνικού μονάδας

Η λεκτική αναπαράσταση του νευρωνικού μονάδας παρουσιάζεται στην εικόνα 4.11c, όπου C_i είναι οι συνθήκες και η έξοδος D αναπαριστά το συμπέρασμα (απόφαση) του κανόνα. Ο συντελεστής σπουδαιότητας (significance factor) sf_i αναπαριστά το βάρος της αντίστοιχης εισόδου της μονάδας adaline.

Οι νευρωνικοί κανόνες μπορούν να δημιουργηθούν από εμπειρικά δεδομένα μέσω μιας καλά ορισμένης διαδικασίας (Hatzilygeroudis & Prentzas, 2010). Οι παραγόμενοι νευρωνικοί κανόνες αποτελούν μια ολοκληρωμένη βάση κανόνων, η οποία μπορεί να χρησιμοποιηθεί για την εξαγωγή συμπερασμάτων και ακόμη παράγουν εξηγήσεις μέσα από μια καλά καθορισμένη διαδικασία εξαγωγής συμπερασμάτων (inference process) (Hatzilygeroudis & Prentzas, 2015).

4.5.4 Πειραματικά Αποτελέσματα

Μια πειραματική μελέτη διεξήχθη και για τις δύο προσεγγίσεις με στόχο την αξιολόγηση της απόδοσής τους. Για τη μελέτη αξιολόγησης, χρησιμοποιήσαμε ένα σύνολο δεδομένων από 240 διαφορετικές ασκήσεις σε αλγόριθμους αναζήτησης. Για τις ανάγκες της μελέτης, δύο διδάσκοντες, με εμπειρία στην διδασκαλία του μαθήματος της τεχνητής νοημοσύνης, κλήθηκαν να προσδιορίσουν τη δυσκολία των ασκήσεων. Για κάθε άσκηση, οι διδάσκοντες προσδιόρισαν το κατάλληλο επίπεδο δυσκολίας της άσκησης και στη συνέχεια δημιουργήθηκε ένα σύνολο δεδομένων. Το σύνολο δεδομένων, περιλάμβανε τιμές των παραμέτρων δυσκολίας της άσκησης και το αντίστοιχο επίπεδο δυσκολίας, που προσδιορίστηκε από τους διδάσκοντες-εμπειρογνώμονες. Για τις ανάγκες του πειράματος, τα δύο τρίτα του συνόλου δεδομένων χρησιμοποιήθηκαν για εκπαίδευση (training), δηλαδή για την παραγωγή των μοντέλων των δύο προσεγγίσεων, και το υπόλοιπο ένα τρίτο του συνόλου δεδομένων χρησιμοποιήθηκε για την αξιολόγηση των μοντέλων. Για την αξιολόγηση της απόδοσης των μηχανισμών, δεδομένου ότι αποτελούν μοντέλα κατηγοριοποίησης,

χρησιμοποιήσαμε τις αντίστοιχες μετρικές αξιολόγησης. Δεδομένου ότι η κλάση εξόδου είναι μια μεταβλητή πολλαπλών τιμών (*multiple class output variable*), υπολογίστηκαν οι μετρικές μέση ορθότητα (*average accuracy*), ακρίβεια (*precision*), ανάκληση (*recall*) και *F-measure* (Sokolova & Lapalme, 2009).

Αρχικά, παρουσιάζονται τα αποτελέσματα απόδοσης της νευρο-ασαφούς προσέγγισης. Τα αποτελέσματα της αξιολόγησης της απόδοσης του ANFIS παρουσιάζονται στον Πίνακα 4.19. Η αξιολόγηση αφορά τη σύγκριση των αποτελεσμάτων του νευρο-ασαφούς συστήματος με τα επίπεδα δυσκολίας που καθορίστηκαν από τους διδάσκοντες-εμπειρογνώμονες.

Πίνακας 4.19 Πειραματικά Αποτελέσματα νευροασαφή μοντέλου

Average accuracy	0.85
Precision	0.84
Recall	0.85
F-measure	0.84

Τα αποτελέσματα δείχνουν ότι ο μηχανισμός έχει πολύ καλή απόδοση και περίπου στο 85% των περιπτώσεων ο νευρο-ασαφής ταξινομητής ήταν σε θέση να προσδιορίσει το σωστό επίπεδο δυσκολίας των ασκήσεων.

Πίνακας 4.20 Πειραματικά Αποτελέσματα Νευρωνικών

Average accuracy	0.78
Precision	0.84
Recall	0.78
F-measure	0.81

Τα αποτελέσματα του νευρωνικών παρουσιάζονται στον πίνακα 4.20 και δείχνουν μια απόδοση στο 78%.

Και τα δύο μοντέλα έδειξαν μια πολύ καλή απόδοση. Η καλύτερη απόδοση επετεύχθη με το νευρο-ασαφές μοντέλο, όπου η ορθότητα και η ακρίβεια ήταν καλύτερες σε σχέση με την προσέγγιση των νευρωνικών στην πειραματική μελέτη. Επίσης, οι δυο προσεγγίσεις είχαν μια πολύ καλή ακρίβεια η οποία ήταν 0.84 και για τις δύο προσεγγίσεις. Στα πλαίσια της μελέτη μας, η μετρική Cohen's Kappa χρησιμοποιήθηκε για να υπολογίσει τον βαθμό

αξιοπιστίας της συμφωνίας μεταξύ των εκτιμήσεων δυσκολίας του διδάσκοντα και του συστήματος στο σύνολο των 240 ασκήσεων. Τα αποτελέσματα παρουσιάζονται στον Πίνακα 4.21.

Πίνακας 4.21. Cohen's kappa μετρική για την νευροασαφή και νευροσυμβολική προσέγγιση

	Neuro-Fuzzy	Neurules
Expert Human Rater	0.807	0.722

Τα αποτελέσματα της μελέτης έδειξαν ότι υπήρξε μια σχεδόν τέλεια συμφωνία (Viera et al. 2005) μεταξύ του διδάσκοντα-εμπειρογνώμονα και της νευρο-ασαφούς προσέγγισης, δεδομένου ότι η μετρική $k = 0.807$ (confidence interval 95%, $p < 0.0005$) δείχνει πολύ ισχυρή συμφωνία. Επίσης, η συμφωνία μεταξύ του διδάσκοντα-εμπειρογνώμονα και της νευρο-συμβολικής προσέγγισης ήταν ελάχιστα χαμηλότερη και η μετρική Kappa ήταν $k = 0.722$ (confidence interval 95%, $p < 0.0005$) δείχνει επίσης μια ισχυρή συμφωνία.

5

Αυτόματη

Αξιολόγηση

Απαντήσεων

5.1 Εισαγωγή

Τα τελευταία χρόνια, υπάρχει έντονο ενδιαφέρον για την ανάπτυξη συστημάτων και μηχανισμών για αυτόματη αξιολόγηση των εκπαιδευόμενων σε ασκήσεις και εκπαιδευτικές δραστηριότητες σε διάφορους τομείς. Η αξιολόγηση των απαντήσεων των εκπαιδευόμενων και, ως εκ τούτου, ο προσδιορισμός των επιδόσεών τους είναι αναγκαία για να γίνει πιο αποτελεσματική η διαδικασία της μάθησης. Η διαδικασία της αυτόματης αξιολόγησης απαντήσεων σε συνδυασμό με την παροχή της κατάλληλης ανάδρασης και βοήθειας από το σύστημα μπορούν να παρέχουν στον εκπαιδευόμενο προσαρμοσμένη ανάδραση στις εκπαιδευτικές του ανάγκες και να ενισχύσουν της απόκτηση γνώσης (Clark, 2012; Nicol & Macfarlane-Dick, 2006; Heffernan & Heffernan, 2014). Επίσης, η αυτόματη αξιολόγηση είναι απαραίτητη για να ενημερώνει το *μοντέλο του εκπαιδευόμενου (student model)* και να παρέχει τις κατάλληλες πληροφορίες σχετικά με το επίπεδο των γνώσεων του εκπαιδευόμενου καθώς και να εντοπίζει μαθησιακά κενά και εκπαιδευτικές δραστηριότητες στις οποίες ο εκπαιδευόμενος ή η τάξη αντιμετωπίζουν ιδιαίτερες δυσκολίες. Μελέτες από την γνωστική και εκπαιδευτική ψυχολογία αναφέρουν ισχυρές συσχετίσεις μεταξύ της αυτοαξιολόγησης και των μαθησιακών αποτελεσμάτων, επισημαίνοντας ότι οι εκπαιδευόμενοι οι οποίοι έχουν επίγνωση του επιπέδου τους και της επίδοσής τους έχουν καλύτερα μαθησιακά αποτελέσματα (Chi et al., 1989; Long & Alevan, 2013; Winne & Hadwin, 1998). Ωστόσο, η σχεδίαση μεθόδων για την αυτόματη αξιολόγηση των επιδόσεων των εκπαιδευόμενων είναι μια διαδικασία πολύπλοκη και άμεσα εξαρτώμενη και προσαρμοσμένη στα χαρακτηριστικά του

πεδίου που διδάσκεται και συνεπώς απαιτείται γνώση των βασικών αρχών του πεδίου, των εννοιών και των περιορισμών που ισχύουν.

Στα πλαίσια αυτού του κεφαλαίου παρουσιάζουμε μια νέα γενική μεθοδολογία για την αυτόματη ανάλυση και αξιολόγηση των απαντήσεων των εκπαιδευόμενων σε ασκήσεις. Η μεθοδολογία που αναπτύξαμε για την αυτόματη βαθμολόγηση βασίζεται στο προσδιορισμό του είδους των λαθών που πραγματοποιούνται από τον εκπαιδευόμενο, στην μοντελοποίηση του τύπου της απάντησής του καθώς και σε ένα γενικό μηχανισμό βαθμολόγησης, που αξιολογεί τις ασκήσεις ανάλογα με το πεδίο εφαρμογής. Πιο συγκεκριμένα, η μεθοδολογία αξιολόγησης εφαρμόστηκε στο πεδίο της Τεχνητής Νοημοσύνης όπου αναπτύξαμε μηχανισμούς για την αυτόματη αξιολόγηση ασκήσεων σχετικών με τη λογική, ως γλώσσας αναπαράστασης και συλλογισμού, και τους αλγορίθμους αναζήτησης. Στόχος της αυτόματης αξιολόγησης είναι να παρέχει στους εκπαιδευόμενους άμεσα, με ακρίβεια και με συνέπεια την ανάλυση των επιδόσεών τους πάνω σε ασκήσεις και εκπαιδευτικές δραστηριότητες χωρίς να χρειάζεται να παρεμβαίνει ο διδάσκοντας. Επομένως, σε αντίθεση με την βαθμολόγηση από τον διδάσκοντα, η οποία δεν γίνεται άμεσα, απαιτεί χρόνο και δεν είναι πάντα συνεπής, η αυτόματη αξιολόγηση προσφέρει άμεσα, με ακρίβεια και συνέπεια την ανάλυση των απαντήσεων των εκπαιδευόμενων, την βαθμολογία τους καθώς και την παροχή κατάλληλης ανάδρασης (feedback) προσαρμοσμένη στα λάθη που έγιναν.

5.2 Σχετικές Εργασίες

Η αυτόματη αξιολόγηση των εκπαιδευόμενων αποτελεί μια σημαντική ερευνητική περιοχή που έχει ελκύσει το ενδιαφέρον των ερευνητών, καθώς θεωρείται αναγκαίο τα εκπαιδευτικά συστήματα διδασκαλίας να ενσωματώνουν μηχανισμούς αυτόματης αξιολόγησης των ενεργειών και των επιδόσεων των εκπαιδευόμενων σε ασκήσεις και σε εκπαιδευτικές δραστηριότητες. Στο πεδίο της αναπαράστασης γνώσης, τα συστήματα KRRT (Alonso et al., 2007) και FITS (Alonso et al., 2006) έχουν ως στόχο να βοηθήσουν τους φοιτητές να μάθουν την χρήση της ΚΛΠΤ ως γλώσσας αναπαράστασης γνώσης. Ωστόσο, δεν διαθέτουν μηχανισμούς για την αυτόματη ανάλυση και αξιολόγηση των απαντήσεων των φοιτητών και η μόνη ανάδραση που παρέχουν σχετίζεται με το αν η απάντηση είναι σωστή ή λανθασμένη. Στην εργασία (Barker-Plummer et al., 2012), οι συγγραφείς πραγματοποιούν μια ανάλυση των απαντήσεων των εκπαιδευόμενων σε ασκήσεις μετατροπής προτάσεων ΦΓ σε εκφράσει ΚΛΠΤ, χρησιμοποιώντας και υπολογίζοντας την *απόσταση μετασχηματισμού* (edit distance) μεταξύ των λανθασμένων απαντήσεων και των αντίστοιχων σωστών, ως μέτρο προσδιορισμού της ορθότητας μιας απάντησης. Η εργασία ωστόσο εστιάζει στον γενικό

προσδιορισμό της ορθότητας μιας απάντησης, χωρίς να παρέχει μηχανισμό αναλυτικής βαθμολόγησης, παροχής ανάδρασης στον εκπαιδευόμενο και επεξήγησης των λαθών.

Το σύστημα TRAKLA2 (Nikander et al., 2004) αναπτύχθηκε για την αυτόματη αξιολόγηση απαντήσεων σε ασκήσεις αλγορίθμων και παρέχει στους εκπαιδευόμενους την δυνατότητα να λύσουν ασκήσεις και να λάβουν άμεσα την αξιολόγηση των απαντήσεων τους.

Στο πεδίο της επιστήμης των υπολογιστών και ειδικότερα στον προγραμματισμό έχουν αναπτυχθεί αρκετά συστήματα για την αυτόματη αξιολόγηση ασκήσεων (Douce et al., 2005; Alemán, 2011; Ala-Mutka, 2005). Τα συστήματα αυτά ενσωματώνουν μηχανισμούς για την αυτόματη αξιολόγηση με στόχο την άμεση βαθμολόγηση των απαντήσεων του εκπαιδευόμενου. Τα βασικότερα από αυτά τα συστήματα είναι τα: ASSYST (Jackson & Usher, 1997), Boss (Joy et al., 2005), GAME (Blumenstein et al., 2008), CourseMarker (Higgins et al., 2005), AutoLEP (Wang et al., 2011) και Autograder (Helmick, 2007).

Το εκπαιδευτικό σύστημα ASSYST (Jackson & Usher, 1997) είναι ένα πολύ γνωστό σύστημα αυτόματης βαθμολόγησης και ένα από τα πρώτα συστήματα που διαθέτει διεπαφή για τον χρήστη. Το ASSYST ενημερώνει τους εκπαιδευόμενους για τις περιπτώσεις στις οποίες έχουν επιτύχει ή αποτύχει, αλλά δημιουργεί επίσης μια βαθμολογία με βάση το μέτρο κυκλωματικής πολυπλοκότητας (*cyclomatic complexity*) (McCabe, 1976). Επίσης, αναλύει την απάντηση του εκπαιδευόμενου, η οποία είναι ένα πρόγραμμα, και με βάση μια σειρά συντακτικών και λειτουργικών κριτηρίων καθορίζει αν τα κριτήρια ικανοποιούνται και στην συνέχεια αξιολογεί τη λειτουργία του προγράμματος με ένα σύνολο προκαθορισμένων δεδομένων. Επίσης, παρέχει ένα περιβάλλον και για τους μαθητές, για να υποβάλουν τις δικές δοκιμές και να αξιολογηθούν στην συνέχεια αυτόματα από το σύστημα. Το σύστημα CourseMaster (Higgins et al., 2005) είναι ένα παράδειγμα ενός συστήματος που παρέχει ένα ευέλικτο σύστημα αξιολόγησης, που έχει την δυνατότητα και τους μηχανισμούς να αξιολογεί τις εργασίες προγραμματισμού των φοιτητών σε διάφορες γλώσσες προγραμματισμού, όπως η C και η C++. Το σύστημα παρέχει αυτόματα ανάδραση στους μαθητές και αναφορές στους εκπαιδευτές σχετικά με τις επιδόσεις των μαθητών τους. Η μέθοδος αξιολόγησης βασίζεται σε συντακτικές αναλύσεις του προγράμματος καθώς και σε λειτουργικές και δυναμικές δοκιμές εκτέλεσης των ασκήσεων των εκπαιδευόμενων. Το σύστημα AutoLEP (Tiantian et al., 2009) παρέχει ένα περιβάλλον για την εκμάθηση της γλώσσας προγραμματισμού C, που ενσωματώνει ένα μηχανισμό αυτόματης αξιολόγησης των απαντήσεων που βασίζεται σε στατική συντακτική ανάλυση και δυναμικές δοκιμές των προγραμμάτων. Η αξιολόγηση βασίζεται στην ομοιότητα των λύσεων των εκπαιδευόμενων και των σωστών λύσεων. Πιο

συγκεκριμένα, το σύστημα αρχικά αναλύει και δημιουργεί μια δενδρική αναπαράσταση της απάντησης του εκπαιδευόμενου και στην συνέχεια τη συγκρίνει με μια σειρά σωστών πρότυπων προγραμμάτων. Με βάση την ομοιότητα μεταξύ των απαντήσεων των εκπαιδευόμενων και των σωστών λύσεων παρέχει κατάλληλες υποδείξεις για αναγνώριση και αντιμετώπιση των λαθών. Το σύστημα QuizPACK (Brusilovsky & Sosnosvsky, 2005) είναι ένα ακόμη παράδειγμα συστήματος που αξιολογεί τις απαντήσεις και τις επιδόσεις των εκπαιδευόμενων στον προγραμματισμό. Το σύστημα έχει την δυνατότητα να δημιουργεί παραμετροποιημένες ασκήσεις για τη γλώσσα C και να αξιολογεί αυτόματα την ορθότητα των απαντήσεων των εκπαιδευόμενων.

Επίσης, έχουν αναπτυχθεί συστήματα για την αυτόματη αξιολόγηση απαντήσεων σε ασκήσεις διαγραμμάτων, όπως στην εργασία (Higgins & Bligh, 2006), όπου γίνεται αυτόματη αξιολόγηση διαγραμμάτων οντοτήτων συσχετίσεων (entity relationship diagrams) με βάση τα χαρακτηριστικά των διαγραμμάτων (commonality variation). Επίσης, στην εργασία (Thomas et al., 2008) παρουσιάζεται ένα εργαλείο για την αξιολόγηση διαγραμματικών ασκήσεων, όπως ασκήσεις σε διαγράμματα Συσχετίσεων-Οντοτήτων (ERDs) και σε Unified Modeling Language (UML), όπου αναλύεται και υπολογίζεται η ομοιότητα της απάντησης του εκπαιδευόμενου με τη σωστή απάντηση και ο αλγόριθμος βαθμολόγησης βασίζεται σε χαρακτηριστικά των διαγραμμάτων.

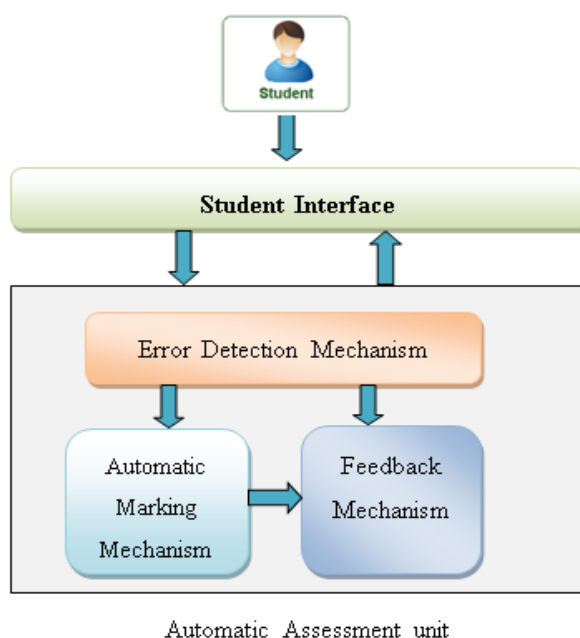
Οι μέθοδοι που βασίζονται στην ομοιότητα γραφημάτων αποτελούν μια καλή προσέγγιση και χρησιμοποιούνται αρκετά συχνά για να αναλύσουν και να αξιολογήσουν τις απαντήσεις των εκπαιδευόμενων. Στην εργασία (Naudé et al., 2010), οι συγγραφείς χρησιμοποιούν μέτρα ομοιότητας γραφημάτων για να αξιολογήσουν τον πηγαίο κώδικα του προγράμματος, συγκρίνοντας τη δομική ομοιότητα μεταξύ της απάντησης του εκπαιδευόμενου και τις ήδη βαθμολογημένες ασκήσεις. Επίσης, στην εργασία (Stajduhar & Maus, 2015), οι συγγραφείς αξιολογούν SQL ασκήσεις συγκρίνοντας την ομοιότητα μεταξύ της απάντησης του εκπαιδευόμενου και της σωστής, με μεθόδους όπως η ευκλείδεια απόσταση και η απόσταση Levenshtein. Τα αποτελέσματα δείχνουν ότι οι μετρήσεις αποστάσεων μπορούν να συμβάλλουν στη συνολική ακρίβεια της μεθόδου αξιολόγησης.

Αξίζει να σημειωθεί ότι οι παραπάνω εργασίες για την αυτόματη αξιολόγηση χρησιμοποιούν μετρικές ομοιότητας, για την εκτίμηση της ομοιότητας μεταξύ της απάντησης του μαθητή και της σωστής απάντησης, ωστόσο αυτό αποτελεί το μόνο και το βασικό κριτήριο αξιολόγησης. Επομένως, δεν βασίζονται σε περαιτέρω ανάλυση των λαθών, όπως ο προσδιορισμός λαθών απροσεξίας, και δεν υπάρχει κάποιο είδος μοντελοποίησης της απάντησης, με βάση κάποια συγκεκριμένα χαρακτηριστικά ανάλογα με το πεδίο εφαρμογής.

Η μεθοδολογία και οι μηχανισμοί αξιολόγησης που αναπτύξαμε και παρουσιάζουμε στο παρόν κεφάλαιο βασίζονται στην ομοιότητα μεταξύ των απαντήσεων, στην εις βάθος ανάλυση των λαθών και στην μοντελοποίηση της απάντησης του εκπαιδευόμενου. Επίσης, σε συνδυασμό με έναν αλγόριθμο για τον προσδιορισμό της βαθμολογίας παρέχετε στον εκπαιδευόμενο η βαθμολογία του και αναλυτικότερη ανάδραση με βάση τα λάθη του. Η μεθοδολογία που αναπτύξαμε μπορεί να προσφέρει νέες δυνατότητες αξιολόγησης σε υπάρχοντα συστήματα και να δώσει νέες κατευθύνσεις στον χώρο της αυτόματης ανάλυσης και αξιολόγησης των εκπαιδευόμενων.

5.3 Μηχανισμός Αυτόματης Αξιολόγησης Απαντήσεων

Σε αυτή την ενότητα παρουσιάζουμε τον γενικό μηχανισμό για την αυτόματη αξιολόγηση απαντήσεων. Κατά τη διάρκεια της αλληλεπίδρασης με το σύστημα, ο εκπαιδευόμενος δίνει μια απάντηση σε μια άσκηση και αυτή προωθείται στον μηχανισμό αυτόματης αξιολόγησης (*Automatic Assessment*). Η γενική αρχιτεκτονική του μηχανισμού αυτόματης αξιολόγησης αποτελείται από τις εξής τρεις βασικές μονάδες: τον *μηχανισμό αναγνώρισης λαθών* (*Error Detection Mechanism*), τον *μηχανισμό αυτόματης βαθμολόγησης* (*Automatic Marking Mechanism*) και τον *μηχανισμό ανάδρασης* (*Feedback Mechanism*), όπως παρουσιάζεται στην εικόνα 5.1.



Εικόνα 5.1. Αρχιτεκτονική μηχανισμού αυτόματης αξιολόγησης απαντήσεων

Ο μηχανισμός αναγνώρισης λαθών χρησιμοποιείται για να αναλύει την απάντηση του εκπαιδευόμενου, να ανιχνεύει τα λάθη που έγιναν και να κατηγοριοποιεί την απάντηση του.

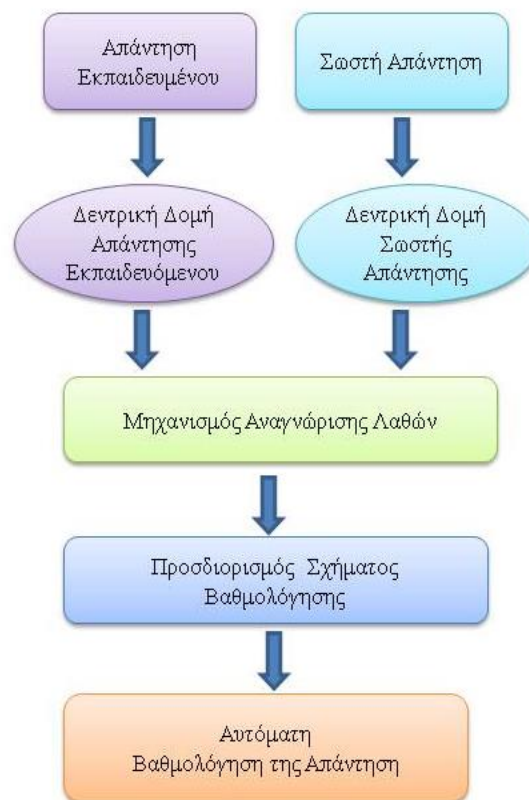
Στην συνέχεια, αλληλοεπιδρά με τον μηχανισμό αυτόματης βαθμολόγησης, ο οποίος υπολογίζει την βαθμολογία της απάντησης του εκπαιδευόμενου στις ασκήσεις. Τέλος, ο μηχανισμός ανάδρασης (feedback Mechanism) χρησιμοποιείται για να παρέχει άμεση ανάδραση στο εκπαιδευόμενο σχετικά με την βαθμολογία και τα λάθη που πραγματοποιήθηκαν στις απαντήσεις.

5.4 Μηχανισμός Αυτόματης Αξιολόγησης Απαντήσεων σε Ασκήσεις Λογικής

5.4.1 Γενικός Μηχανισμός Αυτόματης Αξιολόγησης Απαντήσεων

Η διαδικασία αξιολόγησης απαντήσεων σε ασκήσεις λογικής είναι μια πολυσύνθετη διαδικασία όπου απαιτείται η αξιολόγηση διαφορετικών συνιστωσών των απαντήσεων των εκπαιδευόμενων σε συντακτικό και σημασιολογικό επίπεδο. Συνήθως, κατά τη διδασκαλία της λογικής ως γλώσσας αναπαράστασης γνώσης στα πλαίσια ενός μαθήματος τεχνητής νοημοσύνης, οι εκπαιδευόμενοι καλούνται, ανάμεσα σε άλλες, να επιλύσουν και ασκήσεις πάνω στη μετατροπή προτάσεων ΦΓ σε ΚΛΠΤ, καθώς και ασκήσεις πάνω στη μετατροπή ΚΛΠΤ σε Προτασιακή Μορφή (ΠΜ). Στη συνέχεια ο διδάσκοντας καλείται να αξιολογήσει τις απαντήσεις τους. Η αξιολόγηση των απαντήσεων από τον διδάσκοντα απαιτεί αρκετό χρόνο για να πραγματοποιηθεί και η ενημέρωση των εκπαιδευόμενων για τις επιδόσεις τους δεν γίνεται άμεσα, αλλά συνήθως μετά από μέρες. Επομένως, μεσολαβεί αρκετός χρόνος από την επίλυση της άσκησης μέχρι τον προσδιορισμό της σωστής απάντησης, γεγονός που πολλές φορές δημιουργεί προβλήματα στην μετέπειτα κατανόηση των λαθών. Επίσης, συνήθως λόγω του μεγάλου πλήθους απαντήσεων, δεν είναι δυνατόν ο διδάσκοντας να ενημερώνει κάθε εκπαιδευόμενο αναλυτικά για τα λάθη του, αλλά συνήθως παρέχει τη σωστή απάντηση και τα πιο συχνά λάθη που έγιναν σε μια τάξη. Συνεπώς, ένας μηχανισμός αυτόματης αξιολόγησης των απαντήσεων θα αποτελούσε ένα χρήσιμο εργαλείο για τον διδάσκοντα και τους εκπαιδευόμενους. Για τον σκοπό αυτό, αναπτύχθηκε ένας μηχανισμός αυτόματης αξιολόγησης που αναλύει αυτόματα και αξιολογεί τις απαντήσεις των εκπαιδευόμενων σε ασκήσεις λογικής. Στόχος είναι να βοηθήσει τους εκπαιδευόμενους στην άμεση, αναλυτική και συνεπή αξιολόγηση των απαντήσεών τους και ταυτόχρονα να αποτελέσει ένα χρήσιμο εργαλείο για τον διδάσκοντα, παρέχοντας του την δυνατότητα να ενημερώνεται άμεσα για τις επιδόσεις των εκπαιδευόμενων και τα λάθη που έγιναν.

Ο γενικός μηχανισμός σχεδιάστηκε και αναπτύχθηκε για να αξιολογεί αυτόματα και με ακρίβεια ασκήσεις σχετικά με τη μετατροπή ΦΓ σε ΚΛΠΤ (Perikos et al., 2012) και τη μετατροπή ΚΛΠΤ σε ΠΜ (Grivokostopoulou et al. 2012a; Grivokostopoulou et al. 2012b) και ακολουθεί την γενική αρχιτεκτονική που παρουσιάστηκε στην εικόνα 5.1. Αρχικά, ο μηχανισμός δέχεται σαν είσοδο την απάντηση του εκπαιδευμένου, η οποία είναι μια έκφραση σε ΚΛΠΤ ή σε ΠΜ, η οποία αναλύεται και δημιουργείται η δεντρική δομή της. Στο επόμενο βήμα της διαδικασίας, γίνεται σύγκριση των δεντρικών δομών της απάντησης του και της σωστής απάντησης και προσδιορίζονται τα είδη των λαθών. Στην συνέχεια, με βάση την δομή της έκφρασης ΚΛΠΤ ή ΠΜ προσδιορίζεται το κατάλληλο σχήμα βαθμολόγησης. Τέλος, ενεργοποιείται ο μηχανισμός αυτόματης βαθμολόγησης, όποιος βαθμολογεί την απάντηση του εκπαιδευμένου με βάση τα παραπάνω αποτελέσματα. Στην εικόνα 5.2 παρουσιάζεται η διαδικασία του μηχανισμού για την αυτόματη αξιολόγηση απαντήσεων.



Εικόνα 5.2 Γενικός Μηχανισμός Αξιολόγησης Απαντήσεων

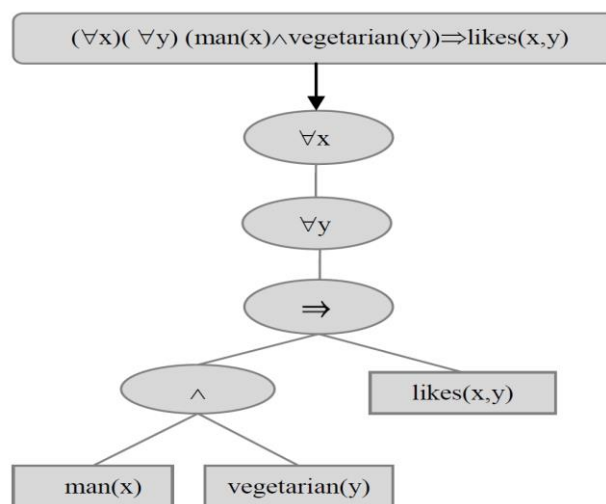
5.4.2 Αυτόματη Αξιολόγηση Απαντήσεων σε Ασκήσεις Μετατροπής ΦΓ σε ΚΛΠΤ

Σε αυτή την ενότητα παρουσιάζουμε τον μηχανισμό που αναπτύχθηκε για την αυτόματη αξιολόγηση απαντήσεων σε ασκήσεις μετατροπής ΦΓ σε ΚΛΠΤ και γίνεται αναλυτική περιγραφή της διαδικασίας αξιολόγησης και των αντίστοιχων μονάδων.

5.4.2.1 Μηχανισμός Αναγνώρισης Λαθών

Όπως παρουσιάστηκε παραπάνω, στη γενική αρχιτεκτονική του μηχανισμού (εικόνα 5.1), μια βασική μονάδα του συστήματος είναι ο *μηχανισμός αναγνώρισης λαθών*, ο οποίος αναλαμβάνει να αναγνωρίζει και να αναλύει τα λάθη που έγιναν στην απάντησή του εκπαιδευόμενου. Αρχικά, ο μηχανισμός αναλύει την απάντηση του εκπαιδευόμενου, η οποία είναι μια έκφραση ΚΛΠΤ, και μέσω του Αναλυτή Εκφράσεων ΚΛΠΤ (FOL Analyzer) προσδιορίζεται η δεντρική του αναπαράσταση. Έπειτα, γίνεται σύγκριση της δεντρικής αναπαράστασης της απάντησης του και της σωστής απάντησης, ώστε να προσδιοριστούν τα είδη των λαθών. Πιο συγκεκριμένα, όταν ο εκπαιδευόμενος απαντήσει σε μια άσκηση, η απάντησή του (έκφραση ΚΛΠΤ) αποτελεί την είσοδο στο μηχανισμό αναγνώρισης λαθών, που την διαχειρίζεται ως εξής:

1. Παραγωγή (μέσω FOL Analyzer) των δεντρικών αναπαραστάσεων της απάντησης του εκπαιδευόμενου και της σωστής απάντησης.
2. Σύγκριση ατομικών εκφράσεων των δέντρων (φύλλα δέντρων).
3. Σύγκριση των διασυνδεδεμένων των δέντρων (εσωτερικοί κόμβοι).
4. Σύγκριση των ποσοδεικτών των δέντρων.



Εικόνα 5.3 Παράδειγμα δεντρικής αναπαράστασης μιας έκφρασης ΚΛΠΤ

Αν σε κάποιο από τα βήματα 2-4 της παραπάνω διαδικασίας βρεθεί αναντιστοιχία μεταξύ της σωστής απάντησης και της απάντησης του εκπαιδευόμενου, χρησιμοποιείται ο *μηχανισμός αναγνώρισης λαθών* για να προσδιοριστούν οι αντίστοιχοι τύποι λαθών που έγιναν. Ο μηχανισμός αναγνώρισης λαθών βασίζεται στην ανάλυση της δεντρικής αναπαράστασης της απάντησης του εκπαιδευόμενου και ο προσδιορισμός των τύπων των λαθών γίνεται με βάση το *σχήμα κατηγοριοποίησης λαθών*, που έχει καθοριστεί με βάση το πεδίο εφαρμογής.

Στην εικόνα 5.3 παρουσιάζεται ένα παράδειγμα δεντρικής αναπαράστασης, για την έκφραση ΚΛΠΤ: “ $(\forall x)(\forall y) (man(x) \wedge vegetarian(y)) \Rightarrow likes(x,y)$ ”.

Το σχήμα κατηγοριοποίησης λαθών βασίστηκε στην ανάλυση των λαθών των εκπαιδευόμενων και σε γενικές αρχές και περιορισμούς του πεδίου. Τα πιθανά είδη των λαθών μοντελοποιήθηκαν σε τρεις κύριες κατηγορίες (Perikos et al., 2012) και παρουσιάζονται στον ακόλουθο πίνακα:

Πίνακας 5.1 Κατηγοριοποίηση λαθών

Κατηγορία Λάθους		Είδος/Τύπος Λάθους
Ατομικές Εκφράσεις	Predicate Error	Number error
		Person error
		Semantics error
	Term Error	Type error
		Semantics error
	Argument Error	Cardinality error
		Order error
Διασυνδετικά		Connective error
		Scope error
Ποσοδείκτες		Type error
		Scope error
		Order error

5.4.2.2 Μηχανισμός Αυτόματης Βαθμολόγησης

Ο μηχανισμός αυτόματης βαθμολόγησης αναπτύχθηκε με στόχο να βαθμολογεί αυτόματα τις απαντήσεις των εκπαιδευόμενων και βασίζεται στην ανάλυση όλων των δομικών μερών της απάντησης. Πιο συγκεκριμένα, προκειμένου να γίνει με ακρίβεια η βαθμολόγηση των απαντήσεων, διαχωρίζεται σε τρεις επιμέρους βαθμολογήσεις:

- a) Βαθμολόγηση της δημιουργίας των ατομικών εκφράσεων
- b) Βαθμολόγηση του προσδιορισμού των διασυνδεδετικών μεταξύ των ατομικών εκφράσεων
- c) Βαθμολόγηση της επιλογής ποσοδεικτών.

Ανάλογα με τα χαρακτηριστικά της κάθε άσκησης προκύπτουν διαφορετικές συνθέσεις των επί μέρους βαθμολογιών για τον προσδιορισμό της συνολικής βαθμολογίας. Για παράδειγμα, ο καθορισμός των ατομικών εκφράσεων είναι απαραίτητο να πραγματοποιείται σε κάθε άσκηση, ενώ ο καθορισμός ποσοδεικτών ή διασυνδεδετικών ενδέχεται να μην είναι πάντα αναγκαίος στα πλαίσια μιας άσκησης. Σε κάθε περίπτωση, οι επιμέρους βαθμολογίες συμβάλουν με διαφορετικό βάρος στον προσδιορισμό της τελικής βαθμολογίας. Η βαθμολόγηση των ασκήσεων παρουσιάζεται στον Πίνακα 5.2, όπου προσδιορίζονται τέσσερα διαφορετικά σχήματα σύνθεσης για τον υπολογισμό της τελικής βαθμολογίας, ανάλογα με τη δομή της πρότασης ΚΛΠΤ.

Πίνακας 5.2 Σχήματα Βαθμολόγησης Ασκήσεων Μετατροπής ΦΓ σε ΚΛΠΤ

Σχήμα	Ατομικές εκφράσεις	Διασυνδεδετικά	Ποσοδείκτες	Σύνθεση Βαθμολογίας	Παράδειγμα έκφρασης ΚΛΠΤ
SS1	Yes	Yes	Yes	35-35-30	$(\forall x) (\text{dog}(x) \wedge \text{loves}(x, \text{master-of}(x))) \Rightarrow \text{loves}(\text{maria}, x)$
SS2	Yes	Yes	No	55-45-0	$\text{dog}(\text{pluto}) \wedge \text{dog}(\text{jack})$
SS3	Yes	No	Yes	65-0-35	$(\exists x) \text{hope}(x)$
SS4	Yes	No	No	100-0-0	$\text{man}(\text{socrates})$

Πιο συγκεκριμένα, ο μηχανισμός βαθμολόγησης λειτουργεί σε βήματα ως εξής:

1. Προσδιορισμός του κατάλληλου σχήματος βαθμολόγησης
2. Υπολογισμός των επιμέρους βαθμολογιών με βάση τα λάθη που πραγματοποιήθηκαν
3. Συνδυασμός των επιμέρους βαθμολογιών για τον υπολογισμό του τελικού βαθμού.

Με βάση τα δομικά μέρη της σωστής απάντησης, επιλέγεται το κατάλληλο σχήμα βαθμολόγησης της απάντησης του εκπαιδευόμενου. Στην συνέχεια προσδιορίζεται η βαθμολογία κάθε δομικού μέρους της απάντησης, με κριτήριο την βαρύτητα των αντίστοιχων λαθών. Κάθε είδος λάθους έχει συνδεθεί με το αντίστοιχο βάρος, το οποίο αντιπροσωπεύει την σημασία του λάθους και το οποίο επηρεάζει την τελική βαθμολογία.

Αρχικά, γίνεται προσδιορισμός της βαθμολογίας των ατομικών εκφράσεων. Η βαθμολογία για κάθε ατομική έκφραση υπολογίζεται από τον τύπο (1), ενώ η βαθμολογία συνολικά για τις ατομικές εκφράσεις υπολογίζεται από τον τύπο (2).

$$AS_i = \begin{cases} \left(maxscore * (1 - \sum_{j=1}^k w_j) \right), k \geq 1 \\ maxscore, k=0 \end{cases} \quad (1)$$

$$Atom_Score = \frac{\sum_{i=1}^n AS_i}{n} \quad (2)$$

όπου n είναι ο αριθμός των ατομικών εκφράσεων που προσδιορίστηκαν, AS_i αναπαριστά την βαθμολογία της ατομικής έκφρασης i , w_j αναπαριστά το βάρος του λάθους που διαπιστώθηκε και το k αναπαριστά το πλήθος των λαθών που έγιναν στις ατομικές εκφράσεις. Το $maxscore$ αναπαριστά τον μέγιστο βαθμό στην εκάστοτε κλίμακα βαθμολόγησης. Τα βάρη των λαθών για τις ατομικές εκφράσεις παρουσιάζονται στον πίνακα 5.3.

Πίνακας 5.3 Βάρη Λαθών για τις ατομικές εκφράσεις

Predicate Error			Term Error		Argument Error	
Number Error	Person Error	Semantics error	Type error	Semantics error	Cardinality error	Order error
25%	25%	25%	20%	5%	30%	20%

Πιο συγκεκριμένα, για μια έκφραση ΚΛΠΤ οι κατηγορίες λαθών που μπορούν να γίνουν ταυτόχρονα είναι από την κατηγορία Predicate Error μπορεί να γίνει μόνο ένας τύπος λάθους, ενώ από τους κατηγορίες Term Error και Argument Error μπορούν να γίνουν ταυτόχρονα και οι δυο τύποι λαθών.

Στην συνέχεια, γίνεται ο υπολογισμός της βαθμολογίας των διασυνδετικών των ατομικών εκφράσεων ως εξής:

$$Connective_Score = \frac{Connective_Number - Connective_Errors}{Connective_number} \times maxscore$$

όπου *Connectives_Number* είναι το πλήθος των διασυνδετικών της σωστής απάντησης και *Connective_Errors* αναπαριστά το πλήθος των λανθασμένων διασυνδετικών στην απάντηση.

Τέλος, υπολογίζεται η βαθμολογία για κάθε ποσοδείκτη ως εξής:

$$QS_i = \begin{cases} (maxscore * (1 - \sum_{j=1}^k qw_j)), k \geq 1 \\ maxscore, k=0 \end{cases} \quad \text{για } 1 \leq i \leq n$$

όπου το QS_i αναπαριστά την βαθμολογία για τον ποσοδείκτη i , qw_j είναι το βάρος του λάθους που έγινε στον προσδιορισμό του ποσοδείκτη i και το k αναπαριστά το πλήθος των λαθών. Τα βάρη για κάθε κατηγορία λάθους για τους ποσοδείκτες παρουσιάζονται στον πίνακα 5.4.

Πίνακας 5.4. Βάρη λαθών για τους ποσοδείκτες.

Scope Error	Type Error
30%	70%

Αξίζει να σημειωθεί ότι, κατά τη διαδικασία της μετατροπής, εκτός από τον προσδιορισμό του λάθους τύπου (type error) και του λάθους της εμβέλειας των ποσοδεικτών (scope error), χρειάζεται να προσδιοριστεί και η σχετική θέση των ποσοδεικτών μέσα στην έκφραση. Σε αντίθεση με τους τύπους των λαθών που παρουσιάζονται στον Πίνακα 5.4, στα οποία κάθε είδος λάθους επηρεάζει την ορθότητα και την βαθμολογία για ένα συγκεκριμένο ποσοδείκτη, ο εντοπισμός ενός λάθους του τύπου σχετικής θέσης (order error) επηρεάζει τη συνολική βαθμολογία των ποσοδεικτών. Επομένως, σε αυτή την περίπτωση η συνολική βαθμολογία των ποσοδεικτών υπολογίζεται από τον ακόλουθο τύπο:

$$Quantifier_Score = \frac{\sum_{i=1}^n QS_i}{n} - error_{order} * w_{order} * maxscore$$

Όπου *Quantifier_Score* είναι η συνολική βαθμολογία για τους ποσοδείκτες, n αναπαριστά το πλήθος των ποσοδεικτών, $error_{order}$ αναπαριστά την ύπαρξη του λάθους σχετικής θέσης

(order error) και w_{order} αναπαριστά το βάρος του λάθους σχετικής θέσης, όπως ορίζεται στον πίνακα 5.5, το οποίο καθορίστηκε από πειραματικά αποτελέσματα.

Πίνακας 5.5 Βάρος λάθους σχετικής θέσης (order error) ποσοδεικτών.

Order Error
50%

Ο τελικός βαθμός της απάντησης υπολογίζεται από τον ακόλουθο τύπο:

$$Overall_Score = \frac{\sum_{i=1}^3 ET_i * sw_i}{maxscore}$$

Όπου ET_i αναπαριστά την βαθμολογία της κάθε επιμέρους κατηγορία και sw_i είναι η αντίστοιχο βάρος της κάθε κατηγορίας στον τελικό βαθμό σύμφωνα με το σχήμα βαθμολόγησης (Πίνακας 5.2).

Για παράδειγμα θεωρούμε την ακόλουθη πρόταση Φυσικής Γλώσσας “All humans eat some food” η οποία μετατρέπεται σε ΚΛΠΤ ως εξής:

“(∀ x) (∃ y) human(x) ∧ food(y) ⇒ eats(x, y)”

και την απάντηση του εκπαιδευόμενου:

“(∀ x)(∀ y)human(x) ∧ food(y) ∧ eats(x, y)”

Η έκφραση ΚΛΠΤ αναλύεται και προσδιορίζεται ότι το κατάλληλο σχήμα βαθμολόγησης είναι το σχήμα $SS1$ όπως φαίνεται στον Πίνακα 5.2. Επομένως, η δημιουργία των ατομικών εκφράσεων έχει συντελεστή 35, ο προσδιορισμός των διασυνδετικών έχει συντελεστή 35 και ο προσδιορισμός των ποσοδεικτών έχει συντελεστή 30. Για την σωστή μετατροπή χρειάζεται να προσδιοριστούν τρεις ατομικές εκφράσεις και ο εκπαιδευόμενος έχει προσδιορίσει σωστά τις δυο από αυτές και έχει πραγματοποιήσει ένα λάθος τύπου *number error* στην ατομική έκφραση $eats(y)$. Η βαθμολογία για κάθε ατομική έκφραση και η συνολική βαθμολογία των ατομικών εκφράσεων είναι αντίστοιχα: $AS_1=100$, $AS_2=100$, $AS_3=75$ $Atom_score = (100 + 100 + 75)/3 = 91$. Επίσης, τα διασυνδετικά δεν προσδιορίστηκαν σωστά και πιο

συγκεκριμένα υπάρχει ένα λάθος στο διασυνδεδετικό ($\wedge$) αντί του προσδιορισμού της συνεπαγωγής ($\Rightarrow$) και η βαθμολογία των διασυνδεδετικών είναι $Connective_Score = ((2-1)/2) * 100 = 50$. Επιπλέον, οι αντίστοιχες βαθμολογίες για τους ποσοδείκτες είναι $QS_1 = 100$, $QS_2 = 30$ και η συνολική βαθμολογία στην κατηγορία των ποσοδεικτών είναι $Quantifier_Score = (100+30)/2 - 0 * 50 = 65$. Το συνολικό σκορ για την απάντηση του εκπαιδευόμενου είναι $Overall_Score = (91 * 35 + 50 * 35 + 65 * 30) / 100 = 68.9$.

NL to FOL Conversion System				
en Welcome isidoros	Automatic Marking Tool			
Selection of sentence				
Score				
Contact				
Information				
Questionnaire				
Logout				
	Natural Language	Student Answer	Score	Error Type
	Apples are delicious	$(\forall x) \text{apples}(x) \wedge \text{delicious}(x)$	61.5	Connective Error: The $\wedge$ is wrong
	Pluto and Jack are dogs	$\text{dog}(\text{pluto}) \vee \text{dog}(\text{jack})$	55	Connective Error: The $\vee$ is wrong
	All gardener like the sun	$\text{gardener}(x) \Rightarrow \text{like}(\text{sun})$	65	Quantifier Error: The quantifier is not determined Argument Error-Cardinality and Order Error: The number and order of arguments is wrong
	Some numbers are even	$(\exists x) \text{numbers}(x) \vee \text{even}(x)$	60.6	Predicate Error: Number Error The word 'numbers' is wrong. It is not in singular number Connective Error: The $\vee$ is wrong
	All human like some drink	$(\exists y)(\forall x) (\text{human}(x) \wedge \text{drink}(y)) \Rightarrow \text{likes}(x,y)$	80	Quantifier Error-Order Error: The order of quantifiers wrong
	All men like all vegetarians	$(\forall x)(\forall y) (\text{men}(x) \wedge \text{vegetarian}(y)) \Rightarrow \text{likes}(x,y)$	94	Predicate Error: Number Error The word 'men' is wrong. It is not in singular number

Εικόνα 5.4 Αυτόματη Αξιολόγηση Ασκήσεων ΦΓ σε ΚΛΠΤ

Στην εικόνα 5.4 παρουσιάζονται μερικά παραδείγματα απαντήσεων των εκπαιδευόμενων όπως βαθμολογήθηκαν από το μηχανισμό, τα λάθη που έγιναν καθώς και κατάλληλες επεξηγήσεις τους.

5.4.3 Μηχανισμός Αυτόματης Αξιολόγησης Απαντήσεων σε Ασκήσεις Μετατροπής ΚΛΠΤ σε ΠΜ

Σε αυτή την ενότητα παρουσιάζουμε τον μηχανισμό που αναπτύχθηκε για την αυτόματη βαθμολόγηση ασκήσεων μετατροπής Κατηγορηματικής Λογικής Πρώτης Τάξεως (ΚΛΠΤ) σε Προτασιακή Μορφή (ΠΜ). Το σύστημα βασίζεται στην γενική αρχιτεκτονική του μηχανισμού αξιολόγησης (εικόνα 5.1). Στη συνέχεια παρουσιάζονται αναλυτικά οι λειτουργίες κάθε μονάδας και οι αντίστοιχες διεργασίες.

5.4.3.1 Μηχανισμός Αναγνώρισης Λαθών

Ο μηχανισμός αναγνώρισης λαθών (*Error Detection Mechanism*) είναι μια βασική μονάδα του συστήματος που αναγνωρίζει και προσδιορίζει τους τύπους των λαθών που έγιναν στην

απάντηση του εκπαιδευόμενου, η οποία είναι μια έκφραση σε μορφή ΣΚΜ- Συζευκτική Κανονική Μορφή (CNF-Conjunctive Normal Form), που είναι ουσιαστικά το τελευταίο βήμα πριν την ΠΜ (CF: Clause Form). Αρχικά, γίνεται ανάλυση της έκφρασης ΣΚΜ της σωστής απάντησης και της απάντησης του εκπαιδευόμενου με χρήση του εργαλείου CF-Analyzer, το οποίο αναλύει τις εκφράσεις και σχηματίζει τις αντίστοιχες δεντρικές δομές. Έπειτα, γίνεται σύγκριση της απάντησης του εκπαιδευόμενου και της σωστής απάντησης, μέσω των δεντρικών δομών τους, για να αναγνωριστούν τα αντίστοιχα λάθη. Ο μηχανισμός αναγνώρισης λαθών διαχειρίζεται την απάντηση του εκπαιδευόμενου ως εξής:

1. Δημιουργία δεντρικών δομών των εκφράσεων ΣΚΜ της σωστής απάντησης και της απάντησης του εκπαιδευόμενου (CF-Analyzer)
2. Αν η απάντηση πρέπει να είναι σε πλήρη ΣΚΜ (δηλ. να περιέχει το συνδετικό «ΚΑΙ»), τότε
 - Αν είναι σε πλήρη ΣΚΜ, θεωρείται ότι είναι δομικά σωστή.
 - Διαφορετικά, προσδιορίζεται το δομικό λάθος που έγινε.
3. Έλεγχος των διασυνδετικών της απάντησης. Για κάθε διασυνδετικό
 - Αν υπάρχει αναντιστοιχία μεταξύ του διασυνδετικού της απάντησης και της σωστής, τότε προσδιορίζεται ο τύπος του λάθους.
 - Διαφορετικά, το διασυνδετικό θεωρείται σωστό.
4. Έλεγχος των ατομικών εκφράσεων. Για κάθε άτομο
 - Αν υπάρχει αναντιστοιχία μεταξύ της απάντησης και της σωστής, τότε προσδιορίζεται ο τύπος του λάθους (πχ. argument error).
 - Διαφορετικά, η ατομική έκφραση θεωρείται σωστή.
5. Έλεγχος ύπαρξης ποσοδεικτών
 - Αν υπάρχει ποσοδείκτης, τότε προσδιορίζεται ο τύπος του λάθους.
 - Διαφορετικά, θεωρείται ότι δεν υπάρχουν λάθη στους ποσοδείκτες.

Ο προσδιορισμός του τύπου ενός λάθους βασίζεται σε ένα σχήμα κατηγοριοποίησης των λαθών που προέκυψε από τις αρχές και τους περιορισμούς του πεδίου καθώς και από την διεξοδική ανάλυση των λαθών των εκπαιδευόμενων. Τα λάθη, στην περίπτωση της μετατροπής από ΚΛΠΤ σε ΠΜ, έχουν κατηγοριοποιηθεί σε τρεις κατηγορίες, όπου κάθε κατηγορία περιλαμβάνει ένα σύνολο από διαφορετικούς τύπους λαθών, όπως παρουσιάζεται στον πίνακα 5.6. Κάθε αναντιστοιχία στην απάντηση αναλύεται, διασυνδέεται με την κατάλληλη κατηγορία λαθών και πιο συγκεκριμένα με τον αντίστοιχο τύπο λάθους. Στην

συνέχεια, το πλήθος και τα είδη (τύποι) των λαθών αποτελούν την είσοδο για τον μηχανισμό αυτόματης βαθμολόγησης.

Πίνακας 5.6. Κατηγοριοποίηση Λαθών ΚΛΠΤ σε ΠΜ

Κατηγορία λάθους	Είδος λάθους	Επεξήγηση
Quantifier Error	Universal Error	Ο καθολικός ποσοδείκτης (εξ) δεν έχει (έχουν) αφαιρεθεί σωστά.
	Skolemization error	Ο υπαρξιακός ποσοδείκτης (εξ) δεν έχει (έχουν) αφαιρεθεί σωστά.
Argument Error	Type error	Ο τύπος ενός ορίσματος δεν είναι σωστός (π.χ. έχει χρησιμοποιηθεί μεταβλητή αντί για συνάρτηση, δηλαδή "loves(y,x)" αντί για "loves(sk-f(x),x)")
	Cardinality error	Ο αριθμός των ορισμάτων μιας ατομικής έκφρασης είναι λάθος. (π.χ. "loves(x,y,sk-f(x))" αντί για "loves(x,sk-f(x))")
Connective error	Structural error	Η δομή της πρότασης είναι λάθος, δεν είναι σε ΣΚΜ.
	Connective Error	Έχει χρησιμοποιηθεί λάθος διασυνδετικό ή έχει παραληφθεί κάποια άρνηση. (π.χ. "&" αντί για "v")
	Scope of Negation	Η εμβέλεια της άρνησης δεν είναι σωστή.

5.4.3.2 Μηχανισμός Αυτόματης Βαθμολόγησης Ασκήσεων

Σε αυτή την ενότητα, γίνεται παρουσίαση του μηχανισμού αυτόματης βαθμολόγησης απαντήσεων σε ασκήσεις μετατροπής προτάσεων από ΚΛΠΤ σε ΠΜ. Ο μηχανισμός αυτός βασίζεται στη δομή της εκφώνησης της άσκησης, δηλαδή στην έκφραση ΚΛΠΤ και τον τύπο λαθών. Πιο συγκεκριμένα, στην περίπτωση που η απάντηση του εκπαιδευομένου είναι σωστή, τότε λαμβάνει την μέγιστη βαθμολογία, ενώ σε διαφορετική περίπτωση η βαθμολόγηση της άσκησης γίνεται με βάση την δομή της έκφρασης ΚΛΠΤ καθώς και τα είδη των λαθών.

Ο μηχανισμός αυτόματης βαθμολόγησης δεν βαθμολογεί την απάντηση του εκπαιδευόμενου ως σωστή ή λανθασμένη, αλλά παρέχει λεπτομερή ανάλυση των λαθών και ακριβή βαθμολογία. Ο υπολογισμός της βαθμολογίας γίνεται ως εξής:

1. Προσδιορισμός του σχήματος βαθμολόγησης
2. Υπολογισμός των επιμέρους βαθμολογιών, με βάση τα αντίστοιχα λάθη
3. Υπολογισμός της τελικής βαθμολογίας της πρότασης.

Για να αξιολογηθεί με ακρίβεια η απάντηση, έγινε καταμερισμός της συνολικής βαθμολογίας σε τέσσερις επιμέρους βαθμολογίες, που αντιστοιχούν στις τέσσερις κύριες κατηγορίες λαθών. Επομένως, για τον υπολογισμό της τελικής βαθμολογίας θα πρέπει να υπολογιστούν οι ακόλουθες επιμέρους βαθμολογίες:

- Βαθμολογία για τις ατομικές εκφράσεις (AS).
- Βαθμολογία για τα διασυνδεδετικά (CS).
- Βαθμολογία για τους καθολικούς ποσοδείκτες (US).
- Βαθμολογία για τους υπαρξιακούς ποσοδείκτες (ES).

Ο προσδιορισμός της τελικής βαθμολογίας βασίζεται στις επιμέρους βαθμολογίες και στην δομή της έκφρασης ΚΛΠΤ. Επίσης, κάθε επιμέρους βαθμολογία συμβάλλει με διαφορετικό βάρος στην διαμόρφωση της τελικής βαθμολογίας, σύμφωνα με το σχήμα βαθμολόγησης (SS). Τα σχήματα βαθμολόγησης καθώς και η αντίστοιχη διάρθρωση της βαθμολόγησης (βάρη συμμετοχής) παρουσιάζονται στον πίνακα 5.7. Για παράδειγμα, το σχήμα βαθμολόγησης SS_3 αντιστοιχεί σε εκφράσεις ΚΛΠΤ που δεν περιέχουν υπαρξιακούς ποσοδείκτες και οι αντίστοιχες τιμές βαρών των επιμέρους βαθμολογιών είναι: 30, 0, 60 και 10.

Αρχικά, μηχανισμός αυτόματης βαθμολόγησης υπολογίζει την βαθμολογία κάθε ατομικής έκφρασης (AS_i) ως εξής:

$$AS_i = \begin{cases} maxscore * (1 - \sum_{j=1}^k w_j), & \text{αν } k \geq 1 \\ maxscore, & \text{αν } k = 0 \end{cases} \quad (1)$$

Όπου k αναπαριστά τον πλήθος των λαθών που έγιναν και w_j αναπαριστά το βάρος του λάθους j και εξαρτάται από τον είδος του λάθους j , σύμφωνα με τον Πίνακα 5.8. Κατόπιν, υπολογίζεται η βαθμολογία όλων των ατομικών εκφράσεων:

$$AS = \frac{\sum_{i=1}^n AS_i}{n} \quad (2)$$

όπου n είναι ο αριθμός των ατόμων της σωστής απάντησης.

Πίνακας 5.7. Σχήμα βαθμολόγησης

Σχήμα	$\forall$	$\exists$	Connectives	Ατομικές Εκφράσεις	Διάρθρωση σχήματος βαθμολόγησης (sw_i)	Παράδειγμα ΚΛΠΤ έκφρασης
SS_1	No	No	No	Yes	0-0-0-100	dog(pluto)
SS_2	No	No	Yes	Yes	0-0-70-30	dog(pluto) $\wedge$ cat(lousi)
SS_3	Yes	No	Yes	Yes	30-0-60-10	$(\forall x) (\text{human}(x) \Rightarrow \text{mortal}(x))$
SS_4	Yes	Yes	No	Yes	30-50-0-20	$(\forall x) (\exists y) \text{ loves}(x, y)$
SS_5	Yes	Yes	Yes	Yes	10-40-35-15	$(\forall x) (\exists y) (\text{human}(x) \wedge \text{food}(y)) \Rightarrow \text{eats}(x, y)$
SS_6	No	Yes	Yes	Yes	0-30-30-40	$(\exists x) \text{ red}(x) \wedge \text{apple}(x)$
SS_7	No	Yes	No	Yes	0-60-0-40	$(\exists x) \text{ hope}(x)$
SS_8	Yes	No	No	Yes	60-0-0-40	$(\forall x) \text{ cube}(x)$

Πίνακας 5.8 Το βάρος κάθε είδους λάθους στις ατομικές εκφράσεις

Ατομικές Εκφράσεις	
Argument Error	
Type Error	Cardinality
70%	30%

Στην συνέχεια, προσδιορίζεται η επιμέρους βαθμολογία για τα διασυνδεδετικά της έκφρασης (CS) ως εξής:

$$CS = \frac{\text{Connective_Num} - \text{Connective_Errors}}{\text{Connective_Num}} * \text{maxscore}$$

όπου το *Connectives_Num* αναπαριστά το πλήθος των διασυνδετικών που χρειάζεται να προσδιοριστούν στην έκφραση ΠΜ από τον εκπαιδευόμενο και το *Connective_Errors* αναπαριστά τον αριθμό των λαθών. Στην περίπτωση που δεν υπάρχουν λάθη, το CS παίρνει την μέγιστη τιμή της βαθμολογίας (*maxscore*), που στην συγκεκριμένη περίπτωση είναι 100, δηλ. $CS = 100$.

Έπειτα, ο μηχανισμός υπολογίζει την βαθμολογία για τους καθολικούς ποσοδείκτες (*US*) και υπαρξιακούς ποσοδείκτες (*ES*), οι οποίες υπολογίζονται ως εξής:

$$US = \begin{cases} 0 & \text{αν } k \geq 1 \\ \text{maxscore} & \text{αν } k = 0 \end{cases}$$

$$ES = \begin{cases} 0 & \text{αν } k \geq 1 \\ \text{maxscore} & \text{αν } k = 0 \end{cases}$$

όπου k αναπαριστά τον αριθμό των λαθών που έγιναν για κάθε είδος ποσοδείκτη. Η τελική βαθμολογία της απάντησης υπολογίζεται ως εξής:

$$CF_Score = \frac{CS * sw_1 + US * sw_2 + ES * sw_3 + AS * sw_4}{\text{maxscore}}$$

όπου sw_1, sw_2, sw_3, sw_4 είναι οι αντίστοιχες τιμές (βάρη) του σχήματος βαθμολόγησης (βλ. Πίνακα 6.7) και *maxscore* είναι ο μέγιστος βαθμός στην κλίμακα αξιολόγησης.

Για παράδειγμα, θεωρούμε την ακόλουθη πρόταση ΚΛΠΤ: “ $(\forall x)(\exists y)(\text{human}(x) \wedge \text{food}(y)) \Rightarrow \text{eats}(x, y)$ ”. Στην συνέχεια παρουσιάζεται η αντίστοιχη πρόταση σε ΠΜ καθώς και η απάντηση του εκπαιδευόμενου.

Σωστή μετατροπή CF

“ $\{\neg \text{human}(x) \vee \neg \text{food}(\text{sk-f}(x)) \vee \text{eats}(x, \text{sk-f}(x))\}$ ”

Απάντηση Εκπαιδευόμενου
(Λανθασμένη)

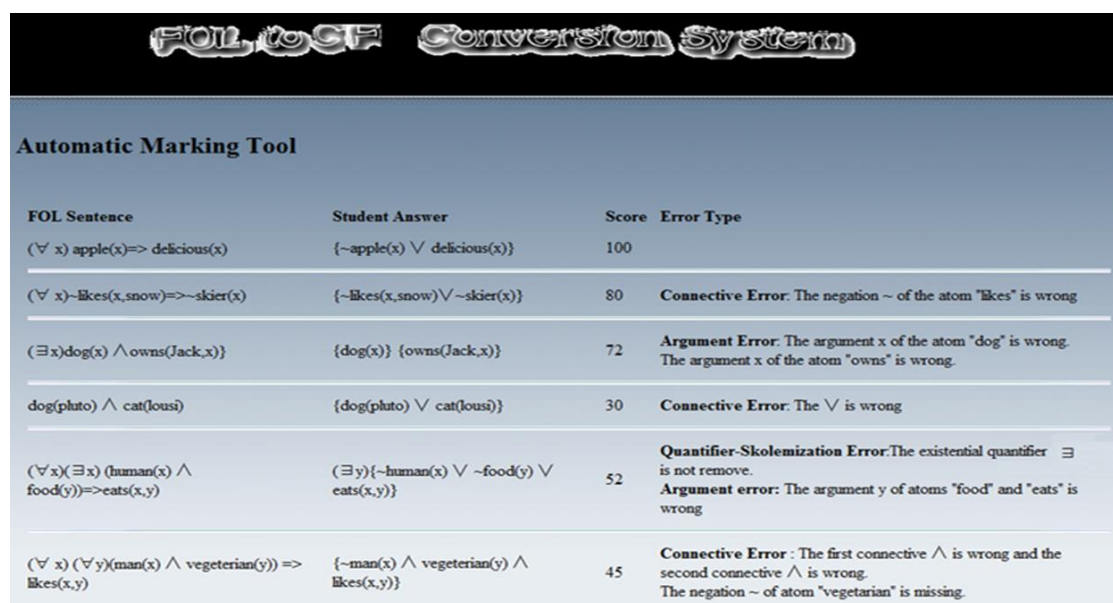
“ $(\exists y)\{\neg \text{human}(x) \vee \neg \text{food}(y) \vee \text{eats}(x, y)\}$ ”

Από την ανάλυση της έκφρασης προκύπτει ότι η απάντηση δεν είναι σωστή και το σχήμα βαθμολόγησης είναι το SS_5 . Για την παραπάνω απάντηση, έχουμε ότι οι ατομικές εκφράσεις δεν προσδιορίστηκαν σωστά, και υπάρχει ένα λάθος τύπου *argument error*. Επομένως, οι επιμέρους βαθμολογίες για κάθε ατομική έκφραση είναι: $AS_1=100$, $AS_2=100-70=30$, $AS_3=100-70=30$ και το $AS = (100+30+30)/3 = 53.3$. Τα διασυνδετικά της έκφρασης έχουν

προσδιοριστεί σωστά και $CS = 100$. Στην έκφραση έχει αφαιρεθεί σωστά ο καθολικός ποσοδείκτης, άρα $US=100$, και δεν έχει αφαιρεθεί ο υπαρξιακός ποσοδείκτης, άρα έχουμε $ES = 0$, και πιο συγκεκριμένα έχουμε ένα λάθος τύπου *Skolemization error*. Επομένως, η τελική βαθμολογία διαμορφώνεται ως εξής:

$$CF_Score = (53.3 * 15 + 100 * 35 + 0 * 40 + 100 * 10) / 100 = 52.9$$

Στην εικόνα 5.5 παρουσιάζονται μερικά παραδείγματα ασκήσεων, η βαθμολόγηση των απαντήσεων καθώς και ανάλυση των λαθών.



FOL Sentence	Student Answer	Score	Error Type
$(\forall x) \text{apple}(x) \Rightarrow \text{delicious}(x)$	$\{\sim \text{apple}(x) \vee \text{delicious}(x)\}$	100	
$(\forall x) \sim \text{likes}(x, \text{snow}) \Rightarrow \sim \text{skier}(x)$	$\{\sim \text{likes}(x, \text{snow}) \vee \sim \text{skier}(x)\}$	80	Connective Error: The negation $\sim$ of the atom "likes" is wrong
$(\exists x) \text{dog}(x) \wedge \text{owns}(\text{Jack}, x)$	$\{\text{dog}(x)\} \{\text{owns}(\text{Jack}, x)\}$	72	Argument Error: The argument x of the atom "dog" is wrong. The argument x of the atom "owns" is wrong.
$\text{dog}(\text{phluto}) \wedge \text{cat}(\text{lousi})$	$\{\text{dog}(\text{phluto}) \vee \text{cat}(\text{lousi})\}$	30	Connective Error: The $\vee$ is wrong
$(\forall x)(\exists x) (\text{human}(x) \wedge \text{food}(y)) \Rightarrow \text{eats}(x, y)$	$(\exists y) \{\sim \text{human}(x) \vee \sim \text{food}(y) \vee \text{eats}(x, y)\}$	52	Quantifier-Skolemization Error: The existential quantifier $\exists$ is not remove. Argument error: The argument y of atoms "food" and "eats" is wrong
$(\forall x)(\forall y)(\text{man}(x) \wedge \text{vegetarian}(y)) \Rightarrow \text{likes}(x, y)$	$\{\sim \text{man}(x) \wedge \text{vegetarian}(y) \wedge \text{likes}(x, y)\}$	45	Connective Error: The first connective $\wedge$ is wrong and the second connective $\wedge$ is wrong. The negation $\sim$ of atom "vegetarian" is missing.

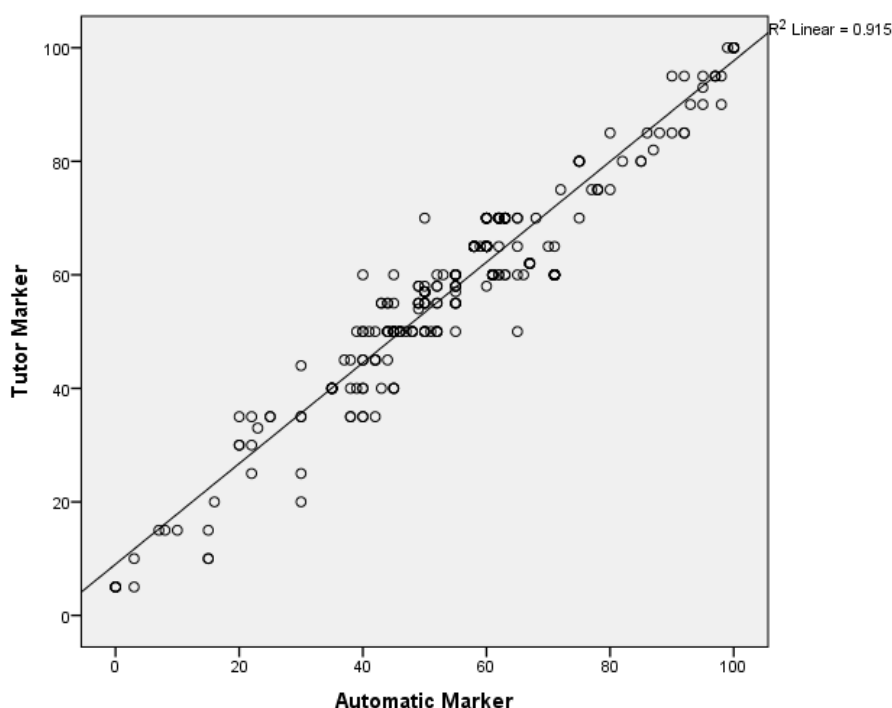
Εικόνα 5.5 Αυτόματη Αξιολόγηση Ασκήσεων ΚΛΠΤ σε ΠΜ

5.4.4 Πειραματική Αξιολόγηση Μηχανισμού Αυτόματης Αξιολόγησης Απαντήσεων σε Ασκήσεις Λογικής

5.4.4.1 Πειραματική αξιολόγηση μηχανισμού σε ασκήσεις μετατροπής ΦΓ σε ΚΛΠΤ

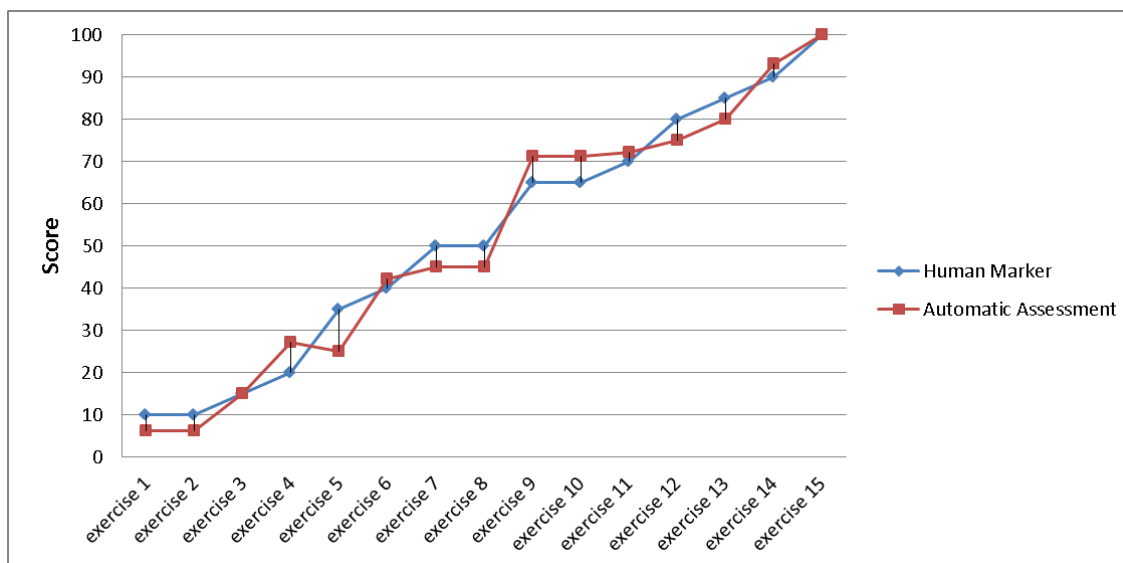
Στα πλαίσια της πειραματικής αξιολόγησης της απόδοσης του μηχανισμού, συλλέξαμε και χρησιμοποιήσαμε διάφορες απαντήσεις φοιτητών σε ασκήσεις μετατροπής ΦΓ σε ΚΛΠΤ. Πιο συγκεκριμένα, συλλέξαμε τυχαία 270 απαντήσεις φοιτητών και δημιουργήθηκε ένα σύνολο με ασκήσεις διαφόρων επιπέδων δυσκολίας και τις αντίστοιχες απαντήσεις των φοιτητών. Αρχικά, το σύνολο όλων των απαντήσεων δόθηκε να αξιολογηθεί από διδάσκοντα (ανθρώπινος αξιολογητής-human marker) και στην συνέχεια βαθμολογήθηκε από τον μηχανισμό αυτόματης βαθμολόγησης. Για να ελέγξουμε αν υπάρχει σχέση ανάμεσα στις

αξιολογήσεις του διδάσκοντα και του μηχανισμό αυτόματης βαθμολόγησης, υπολογίσαμε τον συντελεστή συσχέτισης Pearson r . Η συσχέτιση Pearson απεικονίζει το βαθμό της γραμμικής σχέσης μεταξύ δύο μεταβλητών και κυμαίνεται από -1 μέχρι +1. Μια συσχέτιση ίση με +1 σημαίνει ότι υπάρχει μια τέλεια θετική γραμμική σχέση μεταξύ των μεταβλητών. Στην εικόνα 5.6 παρουσιάζεται το *διάγραμμα διασποράς* (scatterplot), το οποίο δείχνει ότι υπάρχει μια πολύ ισχυρή, θετική συσχέτιση με $r = 0.957$ ($p < 0.001$) και $R^2 = 0.915$ μεταξύ των βαθμών του διδάσκοντα και του αυτόματου μηχανισμού βαθμολόγησης.



Εικόνα 5.6 Διάγραμμα διασποράς μεταξύ διδάσκοντα και αυτόματου μηχανισμού

Στην συνέχεια, στην εικόνα 5.7 παρουσιάζονται παραδείγματα ασκήσεων όπως βαθμολογήθηκαν από το διδάσκοντα και αντίστοιχα από το μηχανισμό αυτόματης βαθμολόγησης.

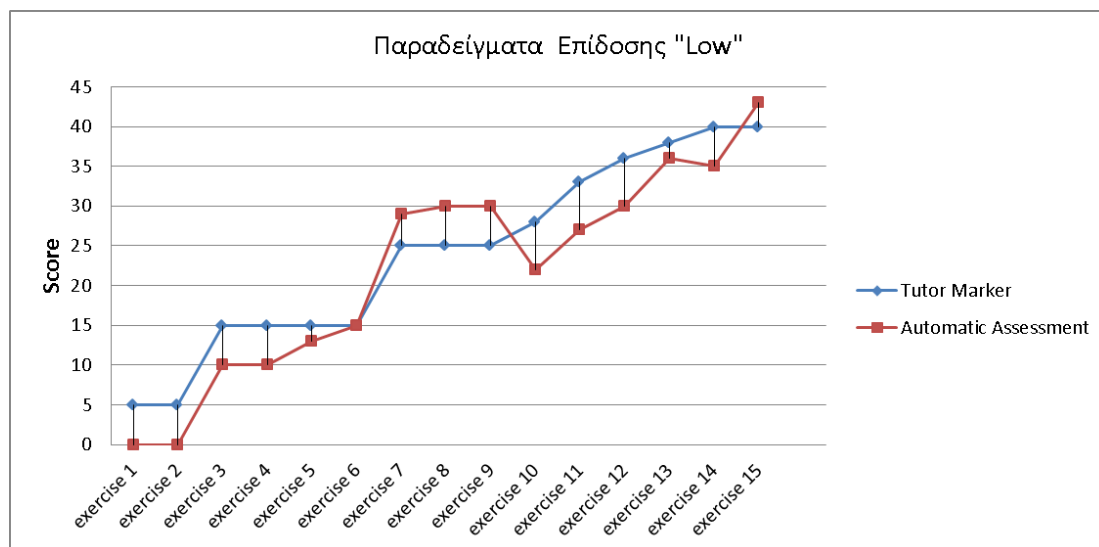


Εικόνα 5.7 Παραδείγματα βαθμολογημένων ασκήσεων από διδάσκοντα και το σύστημα

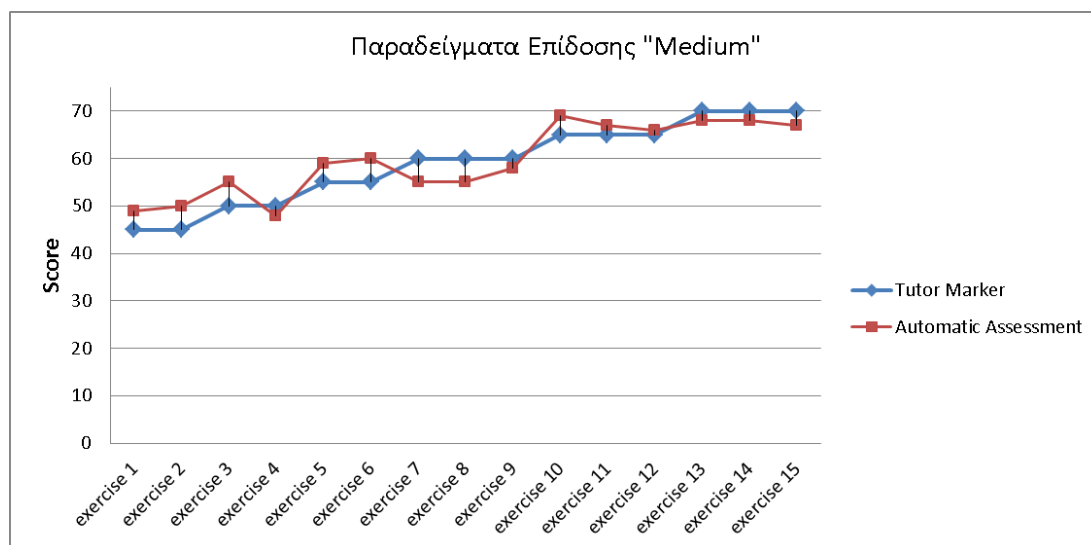
Για να γίνει η αξιολόγηση της απόδοσης του μηχανισμού πιο αναλυτική, κατηγοριοποιήσαμε τις απαντήσεις των εκπαιδευόμενων σε τρεις κατηγορίες: καλές, μέτριες και χαμηλού επιπέδου απαντήσεις. Πιο συγκεκριμένα, οι απαντήσεις των εκπαιδευόμενων με βάση την βαθμολόγηση του διδάσκοντα κατατάχτηκαν στις εξής κατηγορίες :

- Low: Περιλαμβάνει απαντήσεις που βαθμολογήθηκαν από 0 έως και 45
- Medium: Περιλαμβάνει απαντήσεις που βαθμολογήθηκαν από 45 και πάνω έως και 75
- Good: Περιλαμβάνει απαντήσεις που βαθμολογήθηκαν από 75 και πάνω έως 100

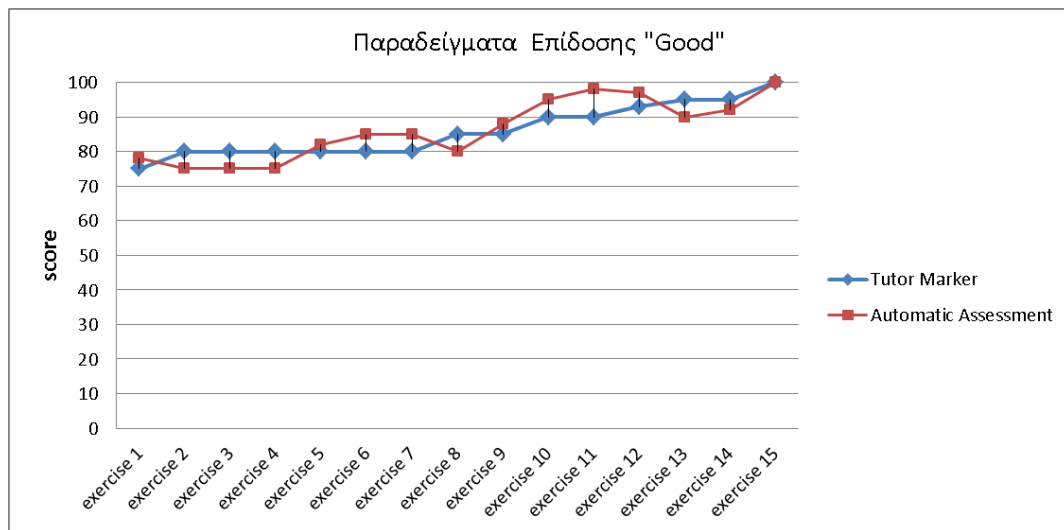
και στην συνέχεια μελετήσαμε την απόδοση του μηχανισμού και την συμφωνία του με τον διδάσκοντα σε κάθε μια κατηγορία. Στις εικόνες 5.8, 5.9 και 5.10 παρουσιάζονται μερικά ενδεικτικά παραδείγματα για κάθε κατηγορία από απαντήσεις εκπαιδευόμενων σε τεστ και αντίστοιχα πώς βαθμολογήθηκαν οι απαντήσεις από τον διδάσκοντα και τον μηχανισμό αυτόματης βαθμολόγησης.



Εικόνα 5.8 Παραδείγματα χαμηλών επιδόσεων βαθμολογημένων ασκήσεων από διδάσκοντα και σύστημα.



Εικόνα 5.9 Παραδείγματα μεσαίων επιδόσεων βαθμολογημένων ασκήσεων από διδάσκοντα και σύστημα



Εικόνα 5.10 Παραδείγματα καλών επιδόσεων βαθμολογημένων ασκήσεων από διδάσκοντα και σύστημα

Στην συνέχεια, για να αποτιμήσουμε την απόδοση του μηχανισμού υπολογίστηκαν οι μετρικές αξιολόγησης: μέση ορθότητα (*average accuracy*), ακρίβεια (*precision*), ανάκληση (*recall*) και *F-measure* δεδομένου ότι έχουμε έξοδο πολλαπλών κλάσεων. Τα αποτελέσματα της ανάλυσης παρουσιάζονται στον ακόλουθο Πίνακα 5.9.

Πίνακας 5.9 Πειραματικά αποτελέσματα μηχανισμού αυτόματης αξιολόγησης

Average accuracy	0.89
Precision	0.93
Recall	0.90
F-measure	0.91

Τα αποτελέσματα δείχνουν ότι ο μηχανισμός αυτόματης αξιολόγησης έχει πολύ καλή απόδοση και προσδιόρισε ορθά το 89% των απαντήσεων κατατάσσοντας τα στην κατάλληλη κατηγορία. Επίσης, παρατηρώντας τα αποτελέσματα του Πίνακα 5.10, προκύπτει ότι υπάρχει καλύτερη συμφωνία μεταξύ της βαθμολόγησης του συστήματος και του ειδικού στις ασκήσεις με μεσαίες και καλές επιδόσεις, ενώ στην περίπτωση των χαμηλών επιδόσεων φαίνεται ο διδάσκοντας βαθμολογεί με υψηλότερους βαθμούς από τον μηχανισμό αυτόματης αξιολόγησης.

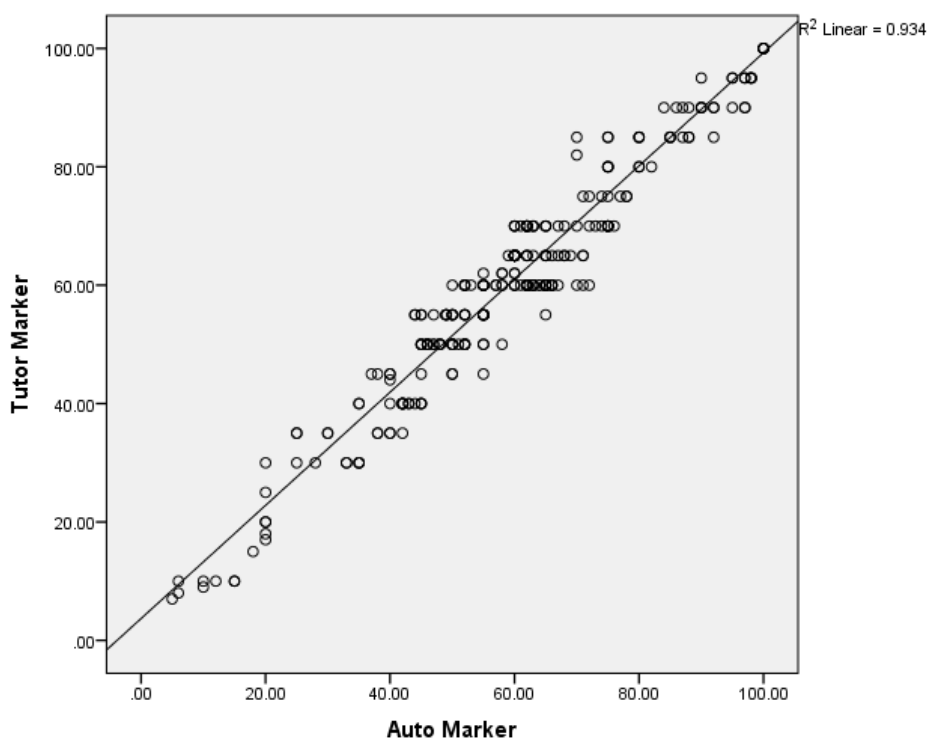
Πίνακας 5.10 Πειραματικά αποτελέσματα μηχανισμού αυτόματης αξιολόγησης

	Precision	Recall	F-Measure
Low	0.93	0.66	0.77
Medium	0.87	0.98	0.92
Good	1	0.94	0.97

Επιπροσθέτως, γίνεται υπολογισμός της συμφωνίας μεταξύ των κατηγοριών των βαθμολογιών του εκπαιδευτή και του αυτόματου μηχανισμού με βάση την μετρική Cohen's Kappa statistic. Η μετρική kappa υπολογίστηκε στο σύνολο των 270 απαντήσεων των εκπαιδευόμενων για να προσδιορίσει αν υπήρχε και σε ποιο βαθμό συμφωνία μεταξύ του διδάσκοντα και του αυτόματου μηχανισμού αξιολόγησης. Τα αποτελέσματα έδειξαν ότι υπάρχει μια σχεδόν τέλεια συμφωνία μεταξύ του εκπαιδευτή και του αυτόματου μηχανισμού αξιολόγησης καθώς $k = 0.858$ (95% confidence interval, $p < .0005$).

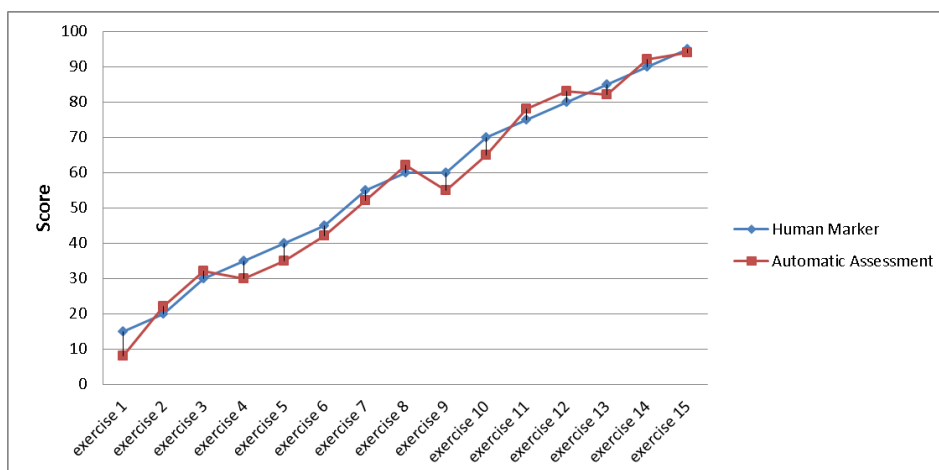
5.4.4.2 Πειραματική αξιολόγηση μηχανισμού σε ασκήσεις μετατροπής ΚΛΠΤ σε ΠΜ

Στην ενότητα αυτή παρουσιάζεται η πειραματική αξιολόγηση του μηχανισμού αυτόματης αξιολόγησης ασκήσεων σε ασκήσεις μετατροπής εκφράσεων ΚΛΠΤ σε ΠΜ. Στα πλαίσια της πειραματικής μελέτης έγινε τυχαία συλλογή απαντήσεων φοιτητών σε ασκήσεις διαφόρων επιπέδων δυσκολίας και δημιουργήθηκε ένα σύνολο που περιέχει 285 απαντήσεις και οι οποίες βαθμολογήθηκαν από τον διδάσκοντα και έπειτα από το σύστημα. Στην συνέχεια, έγινε υπολογισμός του συντελεστή συσχέτισης Pearson r , προκειμένου να προσδιορίσουμε αν υπάρχει σχέση ανάμεσα στις αξιολογήσεις του διδάσκοντα και του μηχανισμού αυτόματης βαθμολόγησης. Στην εικόνα 5.11, παρουσιάζεται το διάγραμμα διασποράς (scatterplot), το οποίο δείχνει ότι υπάρχει μια πολύ ισχυρή, θετική συσχέτιση με $r = 0.966$ ($p < 0.001$ και $R^2 = 0.934$) μεταξύ των βαθμών του διδάσκοντα και του αυτόματου μηχανισμού βαθμολόγησης.



Εικόνα 5.11 Διάγραμμα διασποράς μεταξύ διδάσκοντα και αυτόματου μηχανισμού

Στην εικόνα 5.12, παρουσιάζονται παραδείγματα ασκήσεων όπως βαθμολογήθηκαν από το διδάσκοντα και από το μηχανισμό αυτόματης βαθμολόγησης.



Εικόνα 5.12 Παραδείγματα βαθμολογημένων ασκήσεων από διδάσκοντα και σύστημα

Για να γίνει η αξιολόγηση της απόδοσης του μηχανισμού πιο αναλυτική, κατηγοριοποιήσαμε τις απαντήσεις των εκπαιδευόμενων σε τρεις κατηγορίες: low, medium και good, όπως και στην προηγούμενη ενότητα. Στην συνέχεια, υπολογίσαμε τις ίδιες μετρικές. Τα αποτελέσματα παρουσιάζονται στους Πίνακες 5.11 και 5.12.

Πίνακας 5.11 Πειραματικά αποτελέσματα μηχανισμού αυτόματης αξιολόγησης

Average accuracy	0.93
Precision	0.92
Recall	0.93
F-measure	0.92

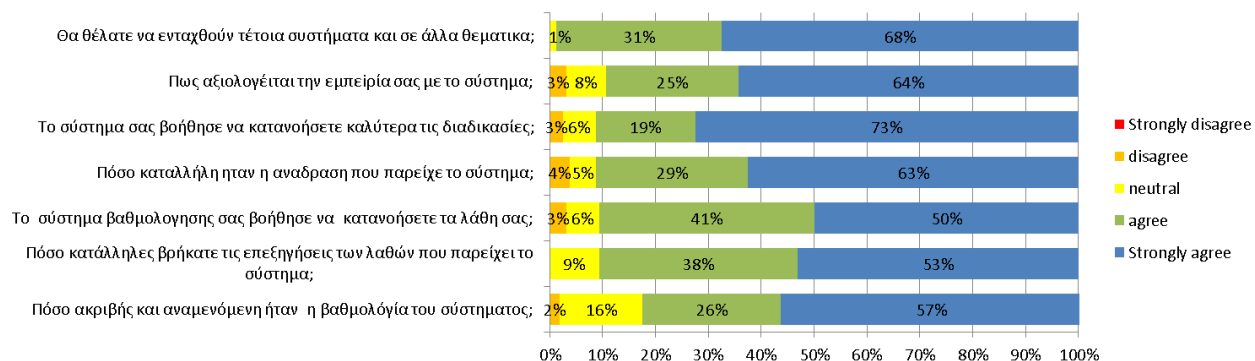
Τα αποτελέσματα του Πίνακα 5.11 δείχνουν ότι ο μηχανισμός αυτόματης αξιολόγησης έχει πολύ καλή απόδοση προσδιόρισε ορθά στο 92.6% των απαντήσεων την κατάλληλη κατηγορία. Παρατηρώντας τα αποτελέσματα του Πίνακα 5.12, προκύπτει ότι υπάρχει πολύ καλή συμφωνία μεταξύ της βαθμολόγησης του συστήματος και του ειδικού σε όλες τις κατηγορίες.

Πίνακας 5.12 Πειραματικά αποτελέσματα μηχανισμού αυτόματης αξιολόγησης

	Precision	Recall	F-Measure
Low	0.94	0.87	0.90
Medium	0.93	0.96	0.94
Good	0.90	0.87	0.89

Επιπροσθέτως, έγινε υπολογισμός της συμφωνίας μεταξύ των κατηγοριών των βαθμολογιών του εκπαιδευτή και του αυτόματου μηχανισμού με βάση την μετρική Cohen's Kappa (Cohen 1960). Τα αποτελέσματα έδειξαν ότι υπάρχει μια ισχυρή συμφωνία μεταξύ του εκπαιδευτή και του αυτόματου μηχανισμού αξιολόγησης, καθώς $k = .768$ (95% confidence interval, $p < .0005$).

Στην τελευταία φάση της πειραματικής μελέτης αξιολόγησης του μηχανισμού, ζητήθηκε από τους εκπαιδευόμενους που χρησιμοποίησαν το εκπαιδευτικό σύστημα να συμπληρώσουν ένα ερωτηματολόγιο και να εκφράσουν την γνώμη τους σχετικά με τις επιδόσεις του μηχανισμού. Πιο συγκεκριμένα, δόθηκε σε ένα σύνολο 160 φοιτητών (άνδρες και γυναίκες) ένα ερωτηματολόγιο το οποίο περιελάμβανε 8 ερωτήσεις, όπου οι ερωτήσεις 1 έως 7 ήταν ερωτήσεις με απαντήσεις σε κλίμακα Likert (1: καθόλου, 5: παρά πολύ), ενώ η ερώτηση 8 ήταν ανοιχτού τύπου, ώστε να εκφράσουν την γνώμη τους, την εμπειρία τους, και οτιδήποτε άλλο κρίνουν απαραίτητο σχετικά με το μηχανισμό



Εικόνα 5.13 Απαντήσεις των εκπαιδευόμενων στο ερωτηματολόγιο

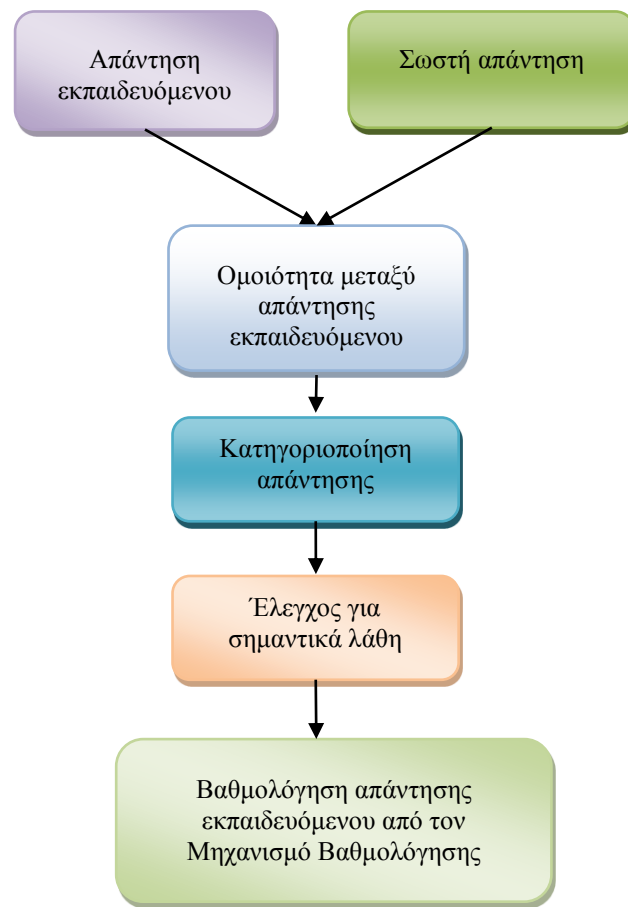
Τα αποτελέσματα (βλ. εικόνα 5.13) έδειξαν ότι το μεγαλύτερο μέρος των φοιτητών έδωσαν θετικές απαντήσεις σχετικά με το σύστημα αυτόματης βαθμολόγησης. Οι απαντήσεις δείχνουν ότι πάνω από το 80% των φοιτητών βρήκε την βαθμολόγηση των ασκήσεων να είναι ακριβή και αναμενόμενη. Επίσης, πάνω από το 85% των φοιτητών βρήκε την παροχή ανάδρασης κατάλληλη από το σύστημα και χρήσιμη για την κατανόηση λαθών.

5.5 Μηχανισμός Αυτόματης Αξιολόγησης Απαντήσεων σε Ασκήσεις Αλγορίθμων Αναζήτησης

Σε αυτή την ενότητα παρουσιάζουμε τον μηχανισμό αυτόματης αξιολόγησης απαντήσεων εκπαιδευόμενων σε αλληλεπιδραστικές ασκήσεις αλγορίθμων αναζήτησης (Grivokostopoulou et al., 2016a; Grivokostopoulou & Hatzilygeroudis, 2013a; Grivokostopoulou & Hatzilygeroudis, 2013b).

5.5.1 Γενική Μεθοδολογία Αυτόματης Αξιολόγησης Απαντήσεων

Στην εικόνα 5.14 παρουσιάζουμε τη γενική μεθοδολογία αξιολόγησης που ακολουθήσαμε και η οποία μπορεί να εφαρμοστεί και σε άλλα παρεμφερή γνωστικά πεδία.



Εικόνα 5.14 Γενική Μεθοδολογία Αυτόματης Αξιολόγησης Απαντήσεων

Η διαδικασία αποτελείται από τέσσερα στάδια τα οποία είναι τα εξής:

1. Καθορισμός της ομοιότητας μεταξύ της απάντησης του εκπαιδευόμενου και της σωστής απάντησης, μέσω *μηχανισμού ανίχνευσης λαθών*
2. Κατηγοριοποίηση των απαντήσεων σύμφωνα με *σχήμα κατηγοριοποίησης απαντήσεων* (*answer categorization scheme*)
3. Έλεγχος για *σημαντικά λάθη* (*important errors*)
4. Υπολογισμός βαθμολογίας μέσω *αυτόματου βαθμολογητή* (*automated marker*)

Για να μπορεί να εφαρμοστεί η παραπάνω διαδικασία και σε άλλα πεδία αρκεί να καθοριστούν κατάλληλα τα ακόλουθα:

- *Μετρική ομοιότητας* (*similarity metric*) μεταξύ απάντησης εκπαιδευόμενου και σωστής απάντησης
- *Σχήμα κατηγοριοποίησης απαντήσεων* (*answer categorization scheme*)

- *Αλγόριθμος αυτόματης βαθμολόγησης (automated marking algorithm)*

Το πλαίσιο αξιολόγησης μπορεί να εφαρμοστεί με ιδιαίτερη αποτελεσματικότητα σε πεδία όπου οι απαντήσεις των εκπαιδευόμενων στις ασκήσεις μπορούν να αναπαρασταθούν ως ακολουθίες στοιχείων (strings).

5.5.2 Μηχανισμός Ανίχνευσης Λαθών

Ο Μηχανισμός ανίχνευσης λαθών (Error Detection Mechanism-EDM) αποτελεί μια βασική μονάδα του συστήματος, η οποία είναι υπεύθυνη να εντοπίζει και να αναγνωρίζει αυτόματα τα λάθη του εκπαιδευόμενου σε αλληλεπιδραστικές ασκήσεις.

5.5.2.1 Αναπαράσταση Απαντήσεων

Οι ασκήσεις στους αλγόριθμους αναζήτησης που καλείται να λύσει ο εκπαιδευόμενος είναι ως εξής. Δίνεται το γράφημα ενός δέντρου αναζήτησης και ζητείται από τον εκπαιδευόμενο ξεκινώντας από την αρχική κατάσταση (ρίζα) να φτάσει σε μια τελική κατάσταση με βάση κάποιον αλγόριθμο αναζήτησης. Ο εκπαιδευόμενος πρέπει να επιλέξει τους κατάλληλους κόμβους-καταστάσεις με την κατάλληλη σειρά, ώστε να δώσει σωστή απάντηση. Οι κόμβοι-καταστάσεις ονοματίζονται με γράμματα της Αγγλικής αλφαβήτου, που καλούνται *ετικέτες (labels)*. Επομένως, η απάντηση του εκπαιδευόμενου αναπαρίσταται σαν μια ακολουθία κόμβων. Ένας κόμβος στην ακολουθία μπορεί να αναπαρίσταται με τρεις διαφορετικές μορφές, ανάλογα με τον αλγόριθμο αναζήτησης που χρησιμοποιήθηκε. Πιο συγκεκριμένα, η μία μορφή είναι να είναι ένα γράμμα της αλφαβήτου, π.χ. «B», στην περίπτωση χρήσης τυφλών αλγορίθμων. Η δεύτερη μορφή είναι να ακολουθείται το γράμμα από ένα ακέραιο αριθμό σε παρένθεση, ο οποίος αντιπροσωπεύει ή το κόστος της διαδρομής από την αρχική κατάσταση, αν χρησιμοποιείται π.χ. αλγόριθμος ελάχιστου κόστους, ή την τιμή του ευρετικού, αν χρησιμοποιείται αλγόριθμος που χρησιμοποιεί ως κριτήριο μόνο το ευρετικό, π.χ. «B(5)». Στην τρίτη μορφή το γράμμα συνοδεύεται από δύο ακέραιους αριθμούς σε παρένθεση, στην περίπτωση που χρησιμοποιείται αλγόριθμος που χρησιμοποιεί και το κόστος και το ευρετικό, όπως ο A^* , π.χ. «B(4, 3)».

Για παράδειγμα, η παρακάτω ακολουθία κόμβων: <A B C M N> θα μπορούσε να αποτελεί μια απάντηση σε μια άσκηση σχετική με τον αλγόριθμο αναζήτησης κατά βάθος όπου A, B, C, M και N αναπαριστούν τους κόμβους-καταστάσεις. Επίσης, για παράδειγμα, η ακολουθία <A(7) B(5) K(3) O(1) L(0)> θα μπορούσε να είναι μια απάντηση σε μια άσκηση αλγορίθμου *αναρρίχησης λόφου (hill climbing)* όπου τα A, B, K, O, L αντιπροσωπεύουν τους κόμβους και τα 7, 5, 3, 1 και 0 είναι αντίστοιχα οι τιμές της ευρετικής συνάρτησης. Τέλος,

μια ακολουθία της μορφής $\langle A(7, 3) B(5, 4) M(3, 5) W(1, 6) Q(0, 7) \rangle$ θα μπορούσε να είναι μια απάντηση σε μια άσκηση σχετική με τον αλγόριθμο A^* , όπου τα A, B, M, W και Q είναι οι κόμβοι, τα 7, 3, 5, 3, 1 και 0 είναι οι τιμές της ευρετικής συνάρτησης και τα 3, 4, 5, 6, 7 είναι οι τιμές κόστους του κάθε κόμβου.

5.5.2.2 Προσδιορισμός Ομοιότητας

Μια βασική λειτουργία του μηχανισμού ανίχνευσης λαθών είναι ο προσδιορισμός της ομοιότητας μεταξύ της ακολουθίας της απάντησης του εκπαιδευόμενου (*Student Answer-SA*) και της αντίστοιχης της σωστής απάντησης (*Correct Answer-CA*). Πιο συγκεκριμένα, για τον προσδιορισμό της ομοιότητας μεταξύ της *SA* και της *CA*, θεωρούμε τις δυο ακολουθίες ως *αλφαριθμητικά* (*strings*) και υπολογίζουμε την *απόσταση μετασχηματισμού* (*edit distance*) (Levenshtein, 1966) μεταξύ τους. Πιο συγκεκριμένα, συμβολίζουμε την *απόσταση μετασχηματισμού* μεταξύ της απάντησης του εκπαιδευόμενου και της σωστής απάντησης ως $d(SA, CA)$, η οποία ορίζεται ως το ελάχιστο πλήθος στοιχειωδών πράξεων που απαιτούνται για να μετασχηματιστεί η ακολουθία *SA* στην *CA*. Ορίζουμε τρεις στοιχειώδεις πράξεις μετασχηματισμού: *εισαγωγή* (*insert*), *διαγραφή* (*delete*) και *μετονομασία* (*relabeling*) ως εξής:

- *node insertion*: Εισαγωγή ενός νέου κόμβου
- *node deletion*: Διαγραφή ενός κόμβου
- *node relabeling*: Μετονομασία ενός κόμβου

Η *εισαγωγή κόμβων* (*node insertion*) και η *μετονομασία κόμβων* (*node relabeling*) διακρίνονται επιπλέον σε τρεις ειδικές περιπτώσεις:

- *Second_Node_Insertion (SCI)*: Εισαγωγή του δεύτερου κόμβου της ακολουθίας
- *Second_Node_Relabeling (SNR)*: Μετονομασία του δεύτερου κόμβου της ακολουθίας
- *Goal_Node_Relabeling (GNR)*: Μετονομασία του τελευταίου κόμβου (κόμβος στόχου) της ακολουθίας

Η διάκριση των παραπάνω ειδικών περιπτώσεων έγινε για να δοθεί κάποιο επιπλέον βάρος σε ορισμένους τύπους λαθών. Για παράδειγμα, όταν ένα λάθος στην απάντηση του εκπαιδευόμενου γίνεται στον δεύτερο κόμβο της ακολουθίας, στην πραγματικότητα αποτελεί την πρώτη επιλογή που έχει γίνει από τον εκπαιδευόμενο, άρα αποτελεί σημαντικό λάθος, καθώς επίσης και η επιλογή του κόμβου στόχου, που στην πραγματικότητα είναι η τελευταία επιλογή του. Επομένως, ο προσδιορισμός της ομοιότητας μεταξύ της απάντησης του

εκπαιδευόμενου (SA) και της σωστής απάντησης (CA), καθορίζεται από την απόσταση μετασχηματισμού, θεωρώντας ότι το κόστος για κάθε στοιχειώδη πράξη (εισαγωγή, διαγραφή, μετονομασία) είναι 1, ενώ για τις πράξεις των ειδικών περιπτώσεων (SCI, SNR, GNR) το κόστος είναι 2. Όσο μικρότερη η τιμή της απόστασης μετασχηματισμού τόσο πιο όμοιες είναι οι δύο ακολουθίες. Στην συνέχεια παρουσιάζονται μερικά παραδείγματα απαντήσεων φοιτητών και αντίστοιχα ο υπολογισμός της απόστασης μετασχηματισμού.

Παράδειγμα 1:

Για παράδειγμα θεωρούμε την ακόλουθη ακολουθία που έδωσε ένας φοιτητής ως απάντηση:

SA=<A B D E M H>

και την αντίστοιχη σωστή ακολουθία:

CA= <A C D E Z H>.

Για να υπολογίσουμε την ομοιότητα μεταξύ των δυο ακολουθιών, χρειάζεται να εφαρμοστεί ο τελεστής relabeling δυο φορές: ο κόμβος B να μετονομαστεί σε κόμβος C και ο κόμβος M να μετονομαστεί σε Z. Χρησιμοποιώντας την παραπάνω προσέγγιση, το κόστος της μετονομασίας (relabeling) του κόμβου B σε κόμβο C είναι 2, διότι αναφερόμαστε στην ειδική περίπτωση *Second_Node_Relabeling*, ενώ το κόστος της μετονομασίας (relabeling) από το M στο Z είναι 1. Επομένως, η απόσταση μετασχηματισμού είναι $d(SA, CA)=2+1=3$.

Παράδειγμα 2:

Θεωρούμε τις ακόλουθες δυο ακολουθίες:

SA=<A(0, 5) C(1, 4) D(2, 3) G(3, 4) F(3, 4), K(0, 7)> και

CA=<A(0, 5) C(1, 4) D(2, 3) M(3, 2) F(3, 4)>

Όπου SA είναι η ακολουθία της απάντηση ενός εκπαιδευόμενου σε μια άσκηση πάνω στον αλγόριθμο A* και CA είναι η ακολουθία της σωστής απάντησης. Για να γίνουν ταυτόσημες απαιτούνται οι ακόλουθες πράξεις: μετονομασία του G (3, 4) σε M (3, 2) και διαγραφή του κόμβου K (0, 7). Το κόστος κάθε πράξης είναι 1. Επομένως, η απόσταση μετασχηματισμού των δύο ακολουθιών είναι $d(SA, CA) = 1 + 1 = 2$.

5.5.3 Μοντελοποίηση Απάντησης

Για την καλύτερη αξιολόγηση των απαντήσεων των εκπαιδευόμενων αναπτύξαμε έναν μηχανισμό που μοντελοποιεί την απάντηση των εκπαιδευόμενων σε διάφορες κατηγορίες.

Πιο συγκεκριμένα, κατηγοριοποιήσαμε τις απαντήσεις των εκπαιδευόμενων με βάση δύο βασικές συνιστώσες την *πληρότητα* (*completeness*) και την *ακρίβεια* (*accuracy*) της απάντησης. Πιο συγκεκριμένα, ο μηχανισμός που αναπτύξαμε κατηγοριοποιεί την απάντηση του εκπαιδευόμενου και την χαρακτηρίζει ως πλήρη (*complete*), εάν όλα τα συστατικά μέρη της σωστής απάντησης εμφανίζονται στην απάντησή του και σε διαφορετική περίπτωση χαρακτηρίζεται ως *ελλιπής* (*incomplete*) ή *πλεονασματική* (*superfluous*) (Fiedler & Tsovaltzi, 2003) και (Gouli et al., 2006). Επίσης, μια απάντηση του εκπαιδευόμενου χαρακτηρίζεται ως *ακριβής* (*accurate*) στην περίπτωση που όλα τα συστατικά μέρη της απάντησης είναι σωστά, ενώ στην περίπτωση που κάποιο συστατικό μέρος της απάντησης δεν είναι σωστό η απάντηση χαρακτηρίζεται ως *ανακριβής* (*inaccurate*). Με βάση την παραπάνω προσέγγιση κατηγοριοποιήσαμε τις απαντήσεις σε πέντε κατηγορίες οι οποίες είναι οι ακόλουθες:

- ***Incomplete – Accurate (IncAcc)***: Όλοι οι κόμβοι της απάντησης του εκπαιδευόμενου που εμφανίζονται είναι σωστοί, αλλά είναι υποακολουθία της σωστής απάντησης και η απόσταση μετασχηματισμού είναι μεγαλύτερη από 0. Δηλαδή, η απάντηση του εκπαιδευόμενου περιλαμβάνει σωστούς κόμβους αλλά λιγότερους κόμβους από τη σωστή απάντηση.
- ***InComplete – InAccurate (IncIna)***: Οι κόμβοι της απάντησης του εκπαιδευόμενου είναι υποσύνολο αυτών της σωστής απάντησης και η απόσταση μετασχηματισμού είναι μεγαλύτερη από 0.
- ***Complete –Inaccurate (ComIna)***: Μόνο όταν ο απαιτούμενος αριθμός των κόμβων εμφανίζεται και η απόσταση μετασχηματισμού είναι μεγαλύτερη από 0. Δηλαδή η ακολουθία περιλαμβάνει σωστούς κόμβους, αλλά οι μεταβάσεις είναι λανθασμένες.
- ***Complete – Accurate (ComAcc)***: Όλοι οι κομβοί της απάντησης του εκπαιδευόμενου που εμφανίζονται είναι σωστοί και η απόσταση μετασχηματισμού είναι 0.
- ***Superfluous (SF)***: Οι κόμβοι που εμφανίζονται είναι περισσότεροι από τους απαιτούμενους και η απόσταση μετασχηματισμού είναι μεγαλύτερη από 0. Εδώ, διακρίνουμε δυο περιπτώσεις:
 - **SF -LG**: Ο τελευταίος κόμβος της ακολουθίας είναι η κατάσταση-στόχος
 - **SF-LNG**: Ο τελευταίος κόμβος της ακολουθίας δεν είναι η κατάσταση-στόχος, δηλαδή ο εκπαιδευόμενος συνέχισε την προσπέλαση κόμβων ενώ είχε φτάσει στο στόχο ή δεν κατάφερε να φτάσει στο στόχο.

Στην περίπτωση που η απάντηση του εκπαιδευόμενου χαρακτηρίζεται ως *complete* και *accurate*, έχει δώσει η σωστή απάντηση. Σε όλες τις άλλες περιπτώσεις, η απάντηση του εκπαιδευόμενου χαρακτηρίζεται ως εσφαλμένη και ο εκπαιδευόμενος θα πρέπει να πραγματοποιήσει τις κατάλληλες ενέργειες για τη δημιουργία της σωστής.

5.5.4 Σημαντικότητα των Λαθών

Ένας σημαντικός παράγοντας στην βαθμολόγηση της απάντησης του εκπαιδευόμενου είναι ο προσδιορισμός της σημαντικότητας των λαθών που πραγματοποιήθηκαν. Δεδομένου ότι όλα τα λάθη που γίνονται από τον εκπαιδευόμενο σε μια άσκηση δεν είναι της ίδιας σημασίας, είναι πολύ σημαντικό να γίνει σωστός διαχωρισμός του είδους του λάθους. Για παράδειγμα, υπάρχουν λάθη κατά την επίλυση μιας άσκησης τα οποία δεν οφείλονται σε έλλειψη γνώσης ή κατανόησης του αντικειμένου, αλλά γίνονται λόγω απροσεξίας ή έλλειψης συγκέντρωσης των εκπαιδευόμενων. Η απροσεξία μπορεί να συμβεί στις περιπτώσεις κατά τις οποίες ο εκπαιδευόμενος κάνει λάθη, παρότι γνωρίζει τη σωστή απάντηση και διαθέτει τις κατάλληλες δεξιότητες και τη γνώση (Herskovitz et al., 2013).

Αξίζει να σημειωθεί ότι στα εκπαιδευτικά συστήματα τα λάθη απροσεξίας αποτελούν μια πολύ συχνή περίπτωση λαθών και μπορεί να γίνουν συχνά ακόμη και από εκπαιδευόμενους που έχουν υψηλές δεξιότητες και επιδόσεις (San Pedro et al. 2011). Πράγματι, σε πολλές περιπτώσεις, οι εκπαιδευόμενοι που διαθέτουν τις κατάλληλες δεξιότητες μπορεί να παρασυρθούν ή να υπερεκτιμήσουν τις δεξιότητές τους και να κάνουν λάθη απροσεξίας ακόμη και σε εύκολες ασκήσεις (San Pedro et al., 2014). Σε αυτή την κατεύθυνση, ένα σημαντικό στοιχείο της μοντελοποίησης των απαντήσεων των εκπαιδευόμενων αποτελεί και ο προσδιορισμός των λαθών απροσεξίας (*careless errors*). Επομένως, είναι σημαντικό να αναγνωριστούν τα λάθη τα οποία οφείλονται όχι σε έλλειψη γνώσης και δεξιοτήτων αλλά σε απροσεξία του εκπαιδευόμενου και να βαθμολογηθούν με διαφορετικό τρόπο. Στην προσέγγισή μας θεωρούμε ότι η ύπαρξη ενός μόνο λάθους (*single error*) στην απάντηση του εκπαιδευόμενου, όπου η ακολουθία της απάντησης αποτελείται από πολλούς κόμβους μπορεί να οφείλεται σε απροσεξία του εκπαιδευόμενου. Με βάση τα παραπάνω κατηγοριοποιήσαμε τα λάθη απροσεξίας σε δυο κατηγορίες:

Τύπος I: Απαιτείται μόνο μια στοιχειώδης πράξη μετασχηματισμού (*insertion, deletion, relabeling*) ή από τις ειδικές περιπτώσεις (SNI, SNR, GNR) για να μετασχηματιστεί η απάντηση του εκπαιδευόμενου στην σωστή απάντηση.

Τύπος II: Απαιτείται μόνο η πράξη της μετονομασίας μεταξύ των δυο διαδοχικών κόμβων (δηλαδή έχει αλλάξει η σειρά των κόμβων μεταξύ τους) για να μετασχηματιστεί η απάντηση του εκπαιδευόμενου στην σωστή απάντηση.

Να σημειωθεί ότι ο *τύπος I* αφορά τις απαντήσεις εκείνες στις οποίες υπάρχει έλλειψη ενός κόμβου ή υπάρχει επιπλέον ένας κόμβος, στον οποίο θα πρέπει να εφαρμοστεί η πράξη του μετασχηματισμού. Ο *τύπος II* αφορά τις απαντήσεις εκείνες, στις οποίες έχουμε δύο κόμβους σε λανθασμένες θέσεις.

5.5.5 Μηχανισμός Αυτόματης Βαθμολόγησης

Σε αυτή την ενότητα, παρουσιάζουμε το *μηχανισμό αυτόματης βαθμολόγησης* (Automatic Marking Mechanism-AMM), ο οποίος αναπτύχθηκε για να βαθμολογεί αυτόματα και άμεσα τις απαντήσεις των εκπαιδευόμενων. Αρχικά, ο μηχανισμός, με βάση το μήκος της ακολουθίας, διακρίνει δύο βασικούς τύπους ασκήσεων, οι οποίες είναι η *απλή* (simple) και η *πολύπλοκη* (complex) αντίστοιχα. Πιο συγκεκριμένα, μια άσκηση χαρακτηρίζεται ως *απλή* αν η σωστή απάντηση αποτελείται από 6 ή λιγότερους κόμβους, ενώ στις υπόλοιπες περιπτώσεις χαρακτηρίζεται ως *πολύπλοκη*. Ο διαχωρισμός έγινε προκειμένου να καταστεί πιο ακριβής η αξιολόγηση κάθε ενός τύπου άσκησης και βασίστηκε σε εμπειρικά δεδομένα και πειραματικά αποτελέσματα. Στην συνέχεια, με βάση τον τύπο των ασκήσεων γίνεται η αξιολόγηση και η βαθμολόγηση των απαντήσεων των εκπαιδευόμενων.

Αρχικά, ο μηχανισμός αυτόματης βαθμολόγησης υπολογίζει την *απόσταση μετασχηματισμού* μεταξύ της απάντησης του εκπαιδευόμενου (SA_i) και της *σωστής απάντησης* (CA_i). Στις περιπτώσεις όπου ο τύπος της άσκησης είναι *απλή*, ο βαθμός του εκπαιδευόμενου υπολογίζεται από τον ακόλουθο τύπο:

$$SS_i = \begin{cases} 100 - 100 * \frac{d(SA_i, CA_i)}{n_c}, & \text{αν } d(SA_i, CA_i) < n_c \\ 0, & \text{αλλιώς} \end{cases}$$

όπου SS_i είναι η βαθμολογία της άσκησης i , $d(SA_i, CA_i)$ είναι η απόσταση μετασχηματισμού και n_c είναι ο αριθμός των κόμβων-καταστάσεων της σωστής απάντησης (CA_i).

Η αξιολόγηση και η βαθμολόγηση των ασκήσεων των εκπαιδευόμενων αποτελεί μια πολύπλοκη και επίπονη διαδικασία για τον αξιολογητή, ιδίως στις περιπτώσεις όπου θα πρέπει να αξιολογηθούν πολύπλοκες ασκήσεις. Ο μηχανισμός αυτόματης βαθμολόγησης

(AMM) για την βαθμολόγηση των απαντήσεων βασίζεται στον τύπο της απάντησης, στην απόσταση μετασχηματισμού καθώς και στην ανάλυση των τύπων των λαθών που πραγματοποιήθηκαν και προσδιορίστηκαν από τον μηχανισμό αναγνώρισης λαθών. Επομένως, ο μηχανισμός μοντελοποιεί και προσδιορίζει τον βαθμό κατανόησης των εκπαιδευόμενων σχετικά με τον τρόπο λειτουργίας των αλγόριθμων. Στην συνέχεια, παρουσιάζεται ο αλγόριθμος βαθμολόγησης, ο οποίος βασίστηκε σε πειραματικά αποτελέσματα, καθώς και στην προσομοίωση του τρόπου βαθμολόγησης του εκπαιδευτή.

1. Αν η άσκηση είναι *απλή*, τότε η βαθμολογία SSi υπολογίζεται μέσω του τύπου (1)

2. Αν η άσκηση είναι *πολύπλοκη*, τότε

2.1 Αν η απάντηση εκπαιδευόμενου (SAi) είναι τύπου ComIna, τότε

2.1.1 Αν η SAi έχει ένα μόνο λάθος (single-error) και είναι Type I, τότε

$$SSi = maxscore * (1 - (w1 * d(SAi, CAi))/nc)$$

2.1.2 Αν η SAi έχει ένα μόνο λάθος (single-error) και είναι Type II, τότε

$$SSi = maxscore * (1 - (w2 * d(SAi, CAi))/nc)$$

2.1.3 Αλλιώς, η SSi υπολογίζεται από τον τύπο (1)

2.2 Αν η απάντηση του εκπαιδευόμενου (SAi) είναι τύπου IncAcc, τότε

2.2.1 Αν η SAi έχει ένα μόνο λάθος (single-error) και είναι Type I, τότε

$$SSi = maxscore * (1 - (w2 * d(SAi, CAi))/nc)$$

2.2.2 Αλλιώς, η SSi υπολογίζεται από τον τύπο (1)

2.3 Αν η απάντηση του εκπαιδευόμενου (SAi) είναι τύπου IncIna, τότε

2.3.1 η SSi υπολογίζεται από τον τύπο (1)

2.4 Αν η απάντηση του εκπαιδευόμενου (SAi) είναι τύπου SF-LG, τότε

2.4.1 Αν η SAi έχει ένα μόνο λάθος (single-error) και είναι Type I, τότε

$$SSi = maxscore * (1 - w1/nc)$$

2.4.2 Αν η SAi έχει ένα μόνο λάθος (single-error) και είναι Type II, τότε

$$SSi = maxscore * (1 - (w2 * d(SAi, CAi))/nc)$$

2.4.3 Αλλιώς, η SSi υπολογίζεται μέσω του τύπου (1)

2.5 Αν η απάντηση του εκπαιδευόμενου (SAi) είναι τύπου SF-LNG, τότε

2.5.1 η SSi υπολογίζεται από τον τύπο (1)

2.6 Αν η απάντηση του εκπαιδευόμενου (SAi) είναι τύπου ComAcc, τότε

$$SSi = maxscore$$

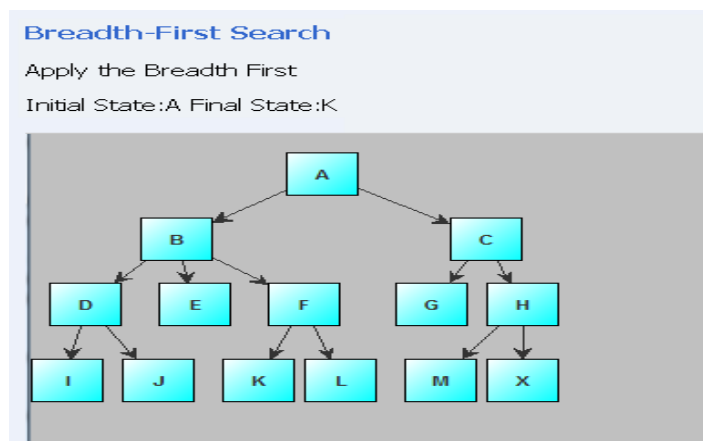
όπου ως $d(SAi, CAi)$ είναι η απόσταση μετασχηματισμού, nc είναι ο αριθμός των κόμβων-καταστάσεων της σωστής απάντησης (CAi) και $w1$ και $w2$ είναι παράμετροι (βάρη) που

αναπαριστούν την σημαντικότητα των σφαλμάτων. Για τον υπολογισμό των βαρών αυτών βασιστήκαμε σε πειραματικά αποτελέσματα και στην εμπειρία των διδασκόντων-εμπειρογνομόνων. Έτσι, καταλήξαμε στις τιμές: $w_1 = 0.6$ και $w_2 = 0.8$. Επίσης, ο προσδιορισμός των τιμών των βαρών μπορεί να γίνει αυτόματα και να εξατομικευτεί στον κάθε εκπαιδευτή παρέχοντας την δυνατότητα στον μηχανισμό να προσομοιώσει τον τρόπο βαθμολόγησης του κάθε εκπαιδευτή. Για το σκοπό αυτό θα χρειαστεί ο εκπαιδευτής να βαθμολογήσει ένα σύνολο απαντήσεων των εκπαιδευόμενων και με βάση την βαθμολογία που προσδιορίστηκε που δημιουργηθεί το κατάλληλο και εξατομικευμένο μοντέλο βαθμολόγησής του.

5.5.5.1 Παραδείγματα Αξιολόγησης Απαντήσεων

Για παράδειγμα, θεωρούμε την ακόλουθη περίπτωση, όπου η αρχική κατάσταση είναι ο κόμβος A και η σωστή απάντηση είναι: CA=< A B K E W Z M>, για τον αλγόριθμο αναζήτησης κατά βάθος (Depth First). Η απάντηση ενός φοιτητή είναι η εξής: SA=< A B K E L D M H F>. Αρχικά, ο μηχανισμός αναγνώρισης σφαλμάτων υπολογίζει την απόσταση μετασχηματισμού μεταξύ της SA και της CA. Οι απαιτούμενες ενέργειες που πρέπει να γίνουν είναι: δύο μετονομασίες ($W \rightarrow L$ και $D \rightarrow Z$) και δυο διαγραφές (H και F). Κάθε ενέργεια έχει κόστος 1. Επομένως, η απόσταση μετασχηματισμού είναι $d(AE, CA) = 2 * 1 + 2 * 1 = 2 + 2 = 4$. Επίσης, ο μηχανισμός αναγνώρισης λαθών αναλύει και κατατάσσει την απάντηση του εκπαιδευόμενου στην περίπτωση SF-LG. Έτσι, σύμφωνα με τον αλγόριθμο βαθμολόγησης έχουμε: $SS = 100 - 100 * (4/7) = 42$.

Έστω ένα δεύτερο παράδειγμα, όπου η αρχική κατάσταση είναι ο κόμβος A και η σωστή απάντηση είναι: CA = <A(8) K(5) M(3) L(2) D(2) C(1) N(0)>, για τον αλγόριθμο αναρρίχησης λόφου (Hill Climbing). Η απάντηση του εκπαιδευόμενου είναι η ακόλουθη: SA=<A(8) B(6) F(8) M(3) L(2) D(2) C(1) N(0)>. Αρχικά, ο μηχανισμός αναγνώρισης λαθών υπολογίζει την απόσταση μετασχηματισμού μεταξύ SA και CA. Οι απαιτούμενες πράξεις είναι: μετονομασία δεύτερου κόμβου (*second_node_relabeling*) ($B(6) \rightarrow K(5)$) και διαγραφή κόμβου (F(8)). Το κόστος της πρώτης πράξης είναι 2, ενώ της δεύτερης είναι 1. Έτσι, η απόσταση μετασχηματισμού είναι $d(AE, CA) = 2 + 1 = 3$. Επίσης, ο μηχανισμός κατηγοριοποιεί την απάντηση του εκπαιδευόμενου ως SF-LG. Επομένως, σύμφωνα με τον αλγόριθμο βαθμολόγησης υπολογίζεται ότι η τελική βαθμολογία ως $SS = 100 - 100 * (3/7) = 57,15$.



Εικόνα 5.15 Παράδειγμα αλληλεπιδραστικής άσκησης εφαρμογής της κατά πλάτος αναζήτησης

Στην εικόνα 5.15 παρουσιάζεται μία άσκηση για τον αλγόριθμο αναζήτησης κατά πλάτος (breadth first search), με σωστή απάντηση: A-B-C-D-E-F-G-H-I-J-K. Στον Πίνακα 5.13, παρουσιάζονται παραδείγματα απαντήσεων των εκπαιδευόμενων, η κατηγοριοποίηση της κάθε απάντησης καθώς και η αντίστοιχη βαθμολογία όπως υπολογίστηκε από τον μηχανισμό αυτόματης βαθμολόγησης.

Πίνακας 5.13 Παραδείγματα βαθμολογημένων απαντήσεων από το σύστημα

Απάντηση Εκπαιδευόμενου	Κατηγορία Απάντησης	Βαθμολογία
A-B-C-D-E-F-G-H-I-K-J	<i>ComIna</i>	78.1
A-C-B-D-F-E-G-H-I-J-K	<i>ComIna</i>	55
A-B-C-D	<i>IncAcc</i>	28
A-B-C-D-E-F-G-H-I-J	<i>IncAcc</i>	85
A-B-C-J-K-I	<i>IncIna</i>	28
A-B-C-D-E-G-H-F-L-M-X-K	<i>SF-LG</i>	55
A-B-C-D-E-G-H-F-L-M-X-J-K	<i>SF-LG</i>	55
A-B-C-D-E-F-H-G-J-I-L-M	<i>SF-LNG</i>	55
A-B-C-D-E-F-G-H-I-J-L-M	<i>SF-LNG</i>	73
A-X-M-L	<i>IncIna</i>	0
A-B-C-D-E-F-G-H-I-J-K	<i>ComAcc</i>	100

5.5.6 Πειραματική Αξιολόγηση Μηχανισμού Αυτόματης Βαθμολόγησης Απαντήσεων

Σε αυτή την ενότητα παρουσιάζουμε τα αποτελέσματα που προέκυψαν από την πειραματική αξιολόγηση του μηχανισμού αυτόματης βαθμολόγησης ασκήσεων. Κατά την διάρκεια της πειραματικής μελέτης, ο αυτόματος μηχανισμός βαθμολόγησης χρησιμοποιήθηκε για την αξιολόγηση των επιδόσεων των μαθητών και η συμπεριφορά του μελετήθηκε με βάση τις αντίστοιχες αξιολογήσεις των διδασκόντων ειδικών. Αρχικά, για την αξιολόγηση του μηχανισμού επιλέξαμε τυχαία 80 προπτυχιακούς φοιτητές (άνδρες και γυναίκες) του τμήματος μας, που είχαν παρακολουθήσει το μάθημα της τεχνητής νοημοσύνης και είχαν χρησιμοποιήσει το ευφύες εκπαιδευτικό σύστημα, στο οποίο έχει ενσωματωθεί ο μηχανισμός, και επίλυσαν ασκήσεις και τεστ ασκήσεων πάνω στους αλγορίθμους τυφλής και ευρετικής αναζήτησης. Κατά την διάρκεια της αλληλεπίδρασης των φοιτητών με το σύστημα, όλες οι ενέργειες και οι απαντήσεις τους καταγράφηκαν. Για την πειραματική μελέτη, αρχικά σχηματίστηκε ένα σύνολο από 400 απαντήσεις, το οποίο περιείχε απαντήσεις σε διάφορες ασκήσεις αλγόριθμων αναζήτησης. Στην συνέχεια, το σύνολο των απαντήσεων βαθμολογήθηκε από διδάσκοντες-ειδικούς στο αντικείμενο, που διέθεταν πολλά χρόνια εμπειρίας στη διδασκαλία αλγόριθμων αναζήτησης.

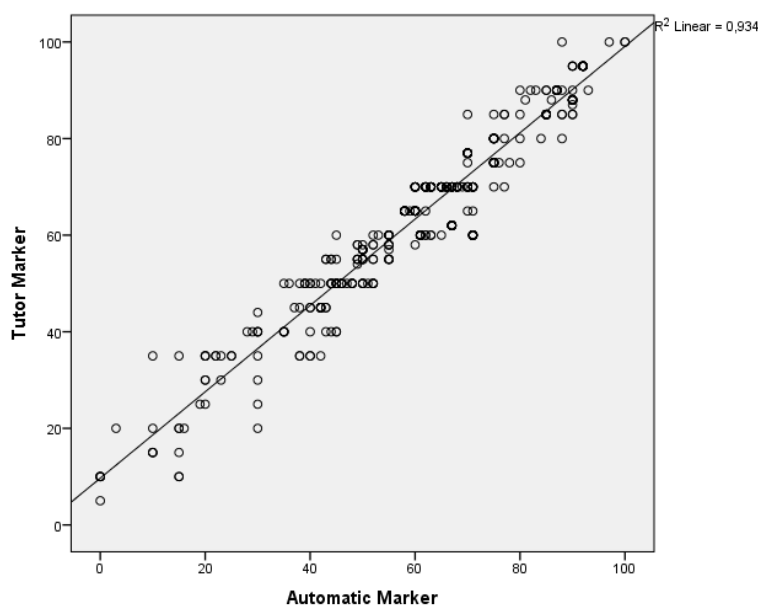
Μετά τη δημιουργία των δύο συνόλων δεδομένων, υπολογίσαμε το *συντελεστή συσχέτισης Pearson r* (Pearson correlation coefficient) για να ελέγξουμε αν υπάρχει σχέση μεταξύ των βαθμών των διδασκόντων-εμπειρογνομόνων και της βαθμολόγησης από το μηχανισμό αυτόματης βαθμολόγησης. Τα αποτελέσματα έδειξαν ότι υπάρχει πολύ ισχυρή, θετική συσχέτιση μεταξύ των βαθμών των διδασκόντων-εμπειρογνομόνων και της αυτόματης βαθμολόγησης του συστήματος, η οποία ήταν στατιστικά σημαντική ($r = 0.96, n = 400, p < 0.0005$).

Επίσης, χρησιμοποιήσαμε και *γραμμική παλινδρόμηση* (linear regression). Η γραμμική παλινδρόμηση είναι μια στατιστική τεχνική μοντελοποίησης, που ερευνά τη συσχέτιση μεταξύ μίας εξαρτημένης μεταβλητής και μιας ή περισσότερων ανεξάρτητων μεταβλητών. Η απλούστερη περίπτωση παλινδρόμησης είναι η *απλή γραμμική παλινδρόμηση* (simple linear regression), κατά την οποία χρησιμοποιούμε μόνο μια μεταβλητή x και μια δεύτερη μεταβλητή y , η οποία μπορεί να προσεγγιστεί ικανοποιητικά από μια γραμμική συνάρτηση της x , όπου x η ανεξάρτητη και y η εξαρτημένη μεταβλητή. Στην περίπτωση μας, θεωρούμε ως ανεξάρτητη μεταβλητή τη βαθμολογία που προσδιορίστηκε αυτόματα από το σύστημα (automated marker) και εξαρτημένη τη βαθμολογία των διδασκόντων-εμπειρογνομόνων (human marker).

Ο στόχος της παλινδρόμησης είναι να μας προσφέρει ένα τρόπο αξιολόγησης του μηχανισμού αυτόματης βαθμολόγησης των ασκήσεων. Η παλινδρόμηση δεν χρησιμοποιείται για να προβλέψει τις επιδόσεις των εκπαιδευόμενων. Αρχικά, ελέγξαμε ότι τα δεδομένα είναι κατάλληλα και πληρούν τις προϋποθέσεις που απαιτούνται ώστε η γραμμική παλινδρόμηση να δώσει ένα έγκυρο αποτέλεσμα. Για παράδειγμα, κοιτάζοντας το διάγραμμα διασποράς (Εικόνα 5.16), μπορούμε να δούμε ότι η προϋπόθεση για τις δύο μεταβλητές (automated marker, tutor marker) να είναι συνεχείς πληρείται: υπάρχει μια γραμμική σχέση μεταξύ των δύο μεταβλητών και δεν υπάρχουν σημαντικές ακραίες τιμές. Μια απλή γραμμική παλινδρόμηση (simple linear regression) στα δυο σύνολα απαντήσεων των εκπαιδευόμενων έδωσε στην εξίσωση:

$$y = 9.697 + 0.894x \text{ και } r^2 = 0.934$$

Τα αποτελέσματα του γραμμικού μοντέλου παλινδρόμησης έδειξαν ότι υπάρχει ισχυρή θετική σχέση (Dancey & Reidy, 2007) μεταξύ του μηχανισμού αυτόματης βαθμολόγησης (automated marker) και των διδασκόντων-εμπειρογνομόνων (human marker), αφού το r^2 υπολογίστηκε ότι είναι $r^2 = 0.967$, κάτι που επιπροσθέτως δείχνει ότι το 93.4% της συνολικής μεταβολής στην y μπορεί να εξηγείται από την γραμμική σχέση μεταξύ x και y . Αυτό σημαίνει ότι η ευθεία γραμμή της προσέγγισης της παλινδρόμησης αντιπροσωπεύει και υποστηρίζει περίπου 93.4% της μεταβλητότητας στο σύνολο δεδομένων. Το μοντέλο προσεγγίζει αρκετά καλά τα δεδομένα ($F = 5656$, $p < 0.0005$).



Εικόνα 5.16 Διάγραμμα διασποράς automatic marker σε σχέση με human marker

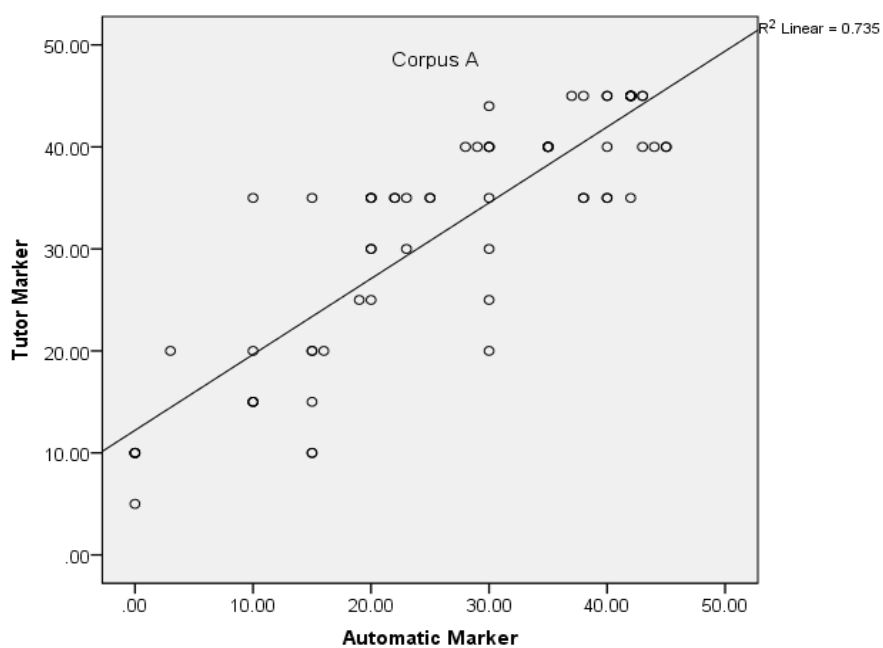
Στη συνέχεια, δημιουργήσαμε τρία σύνολα (corpuses), εκτός από το σύνολο των 400 απαντήσεων, με ονόματα CorpusA, CorpusB και CorpusC. Το CorpusA περιείχε μόνο τις απαντήσεις των εκπαιδευόμενων που αξιολογήθηκαν από τους διδάσκοντες και εμπίπτουν στις κατηγορίες “very low” και “low”, το CorpusB περιείχε μόνο τις απαντήσεις που αξιολογήθηκαν ως “medium” και “good” και το CorpusC μόνο τις απαντήσεις επιπέδου ‘excellent’. Πιο συγκεκριμένα, χαρακτηρίζουμε μια βαθμολογημένη απάντηση ως “very low” αν η βαθμολογία είναι μέχρι και 25, ‘low’ αν η βαθμολογία του είναι από 25 και πάνω μέχρι και 45, “medium”, αν είναι από 45 και πάνω μέχρι και 65, “good” εάν είναι από 65 και πάνω μέχρι και 85 και ‘excellent’, αν η βαθμολογία είναι μεγαλύτερη από 85. Ο σκοπός της δημιουργίας των τριών συνόλων είναι για να γίνει πιο αναλυτική αξιολόγηση των επιδόσεων του αυτόματου μηχανισμού αξιολόγησης ως προς το είδος και το επίπεδο βαθμολόγησης των απαντήσεων των φοιτητών.

Στον πίνακα 5.14 παρουσιάζονται οι συσχετίσεις μεταξύ των βαθμολογιών των διδασκόντων (human marker) και του αυτόματου μηχανισμού αξιολόγησης για τα τρία αυτά σύνολα απαντήσεων. Η συσχέτιση στο σύνολο απαντήσεων corpusC ήταν 0.896, και ήταν η υψηλότερη μεταξύ των συνόλων.

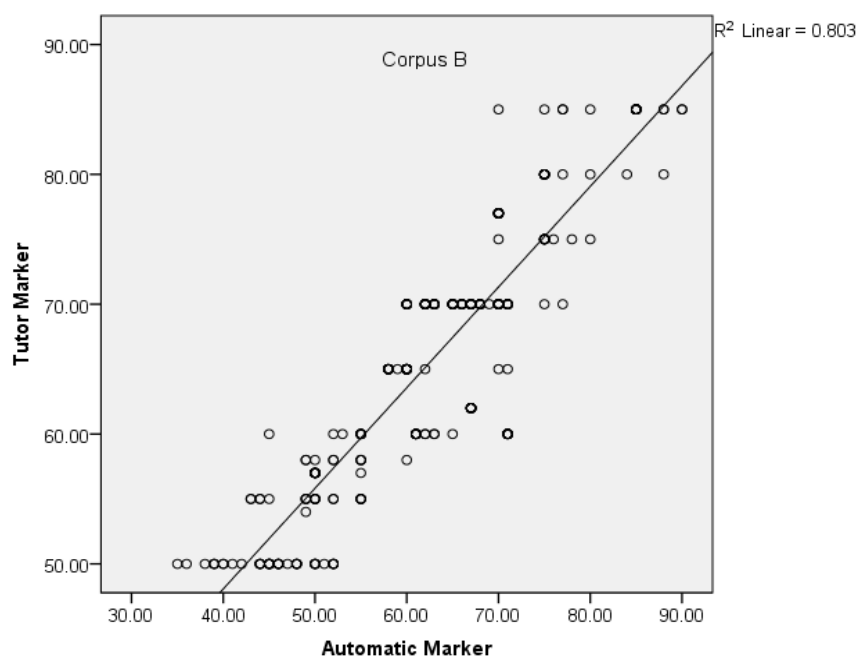
Πίνακας 5.14 Αποτελέσματα Συσχέτισης

Corpus	Correlation	Significance of correlation
Corpus A	0.858	1.98×10^{-20}
Corpus B	0.896	1.26×10^{-100}
Corpus C	0.629	7.49×10^{-7}

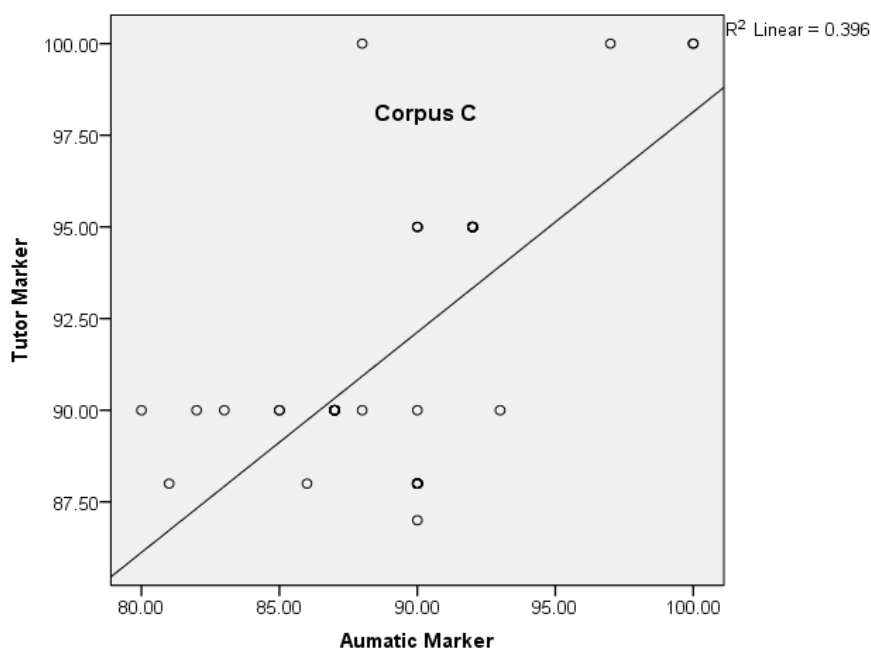
Επιπλέον, στις Εικόνες 5.17 (α, β, γ) παρουσιάζονται, τα διαγράμματα διασποράς, που προέκυψαν για τις τρεις κατηγορίες. Στην πραγματικότητα, στην εικόνα 5.17 b δείχνει μια καλύτερη συμφωνία μεταξύ της βαθμολόγησης του διδάσκοντα και του συστήματος για απαντήσεις των εκπαιδευόμενου που ήταν επιπέδου “medium” από ότι στα άλλα επίπεδα, δεδομένου ότι τα στοιχεία επικεντρωθεί περισσότερο στην περιοχή της αντίστοιχης γραμμής.



(α)



(β)



(γ)

Εικόνα 5.17 Διαγράμματα διασποράς για (α) CorpusA, (β) CorpusB, (γ) CorpusC

Δεύτερη Πειραματική Μελέτη

Ένα δεύτερο πείραμα σχεδιάστηκε για να αξιολογήσει το σύστημα αυτόματης βαθμολόγησης απαντήσεων. Στα πλαίσια του δεύτερου πειράματος, η αξιολόγηση των απαντήσεων και ο προσδιορισμός της βαθμολογίας τους από τον αυτόματο μηχανισμό αντιμετωπίστηκε ως πρόβλημα κατηγοριοποίησης. Με βάση αυτή την προσέγγιση, στόχος είναι να γίνει αξιολόγηση της απόδοσης του μηχανισμού στο να προσδιορίζει και να κατατάσσει μια απάντηση στην κατάλληλη κατηγορία απαντήσεων. Οι κατηγορίες απαντήσεων είναι οι “very low”, “low”, “medium”, “good” και “excellent” και για κάθε μια από τις 400 απαντήσεις των φοιτητών προσδιορίστηκε η κατηγορία της με βάση την βαθμολογία των εμπειρογνομόνων. Σε αυτή την κατεύθυνση, υπολογίστηκαν κατάλληλες μετρικές για να αποτιμήσουν την απόδοση του μηχανισμού. Αρχικά, η αξιολόγηση βασίστηκε στις γνωστές μετρικές : μέση ορθότητα (*average accuracy*), ακρίβεια (*precision*) και *F-measure*, δεδομένου ότι έχουμε έξοδο πολλαπλών κλάσεων. Τα αποτελέσματα της ανάλυσης παρουσιάζονται στον Πίνακα 5.15.

Πίνακας 5.15 Πειραματικά αποτελέσματα μηχανισμού αυτόματης αξιολόγησης

Average accuracy	0.83
Precision	0.84
F-measure	0.84

Τα αποτελέσματα δείχνουν ότι ο μηχανισμός αυτόματης αξιολόγησης έχει πολύ καλή απόδοση και προσδιόρισε ορθά για το 83% των απαντήσεων των εκπαιδευόμενων την κατάλληλη κατηγορία. Συνολικά προσδιόρισε ορθά την κατηγορία σε 332 από τις 400 απαντήσεις και στον Πίνακα 5.16 παρουσιάζονται οι αναλύσεις των επιδόσεων του μηχανισμού σε κάθε κατηγορία.

Πίνακας 5.16 Πειραματικά αποτελέσματα μηχανισμού αυτόματης αξιολόγησης

	Precision	Recall	F-Measure
Very Low	0.90	0.58	0.70
Low	0.72	0.55	0.62
Medium	0.78	0.95	0.84
Good	0.89	0.85	0.87
Excellent	0.92	1	0.96

Τα αποτελέσματα δείχνουν ότι ο μηχανισμός έχει πολύ καλή απόδοση, ιδιαίτερα στην κατηγορία απαντήσεων που είναι ‘good’ και ‘excellent’. Τέλος, γίνεται υπολογισμός της συμφωνίας μεταξύ των κατηγοριών των βαθμολογιών των εκπαιδευτών και του αυτόματου μηχανισμού με βάση την μετρική Cohen’s Kappa statistic. Η μετρική kappa υπολογίστηκε στο σύνολο των 400 απαντήσεων των εκπαιδευόμενων για να προσδιορίσει αν υπήρχε και σε ποιο βαθμό συμφωνία μεταξύ των εκπαιδευτών και του αυτόματου μηχανισμού αξιολόγησης. Τα αποτελέσματα έδειξαν ότι υπάρχει μια ισχυρή συμφωνία μεταξύ τους, καθώς $k = 0.707$ (95% confidence interval, $p < .0005$).

6

Εκπαιδευτικά Παιχνίδια

6.1 Εισαγωγή

Τα τελευταία χρόνια, η ραγδαία ανάπτυξη της εκπαιδευτικής τεχνολογίας έχει αναδείξει νέα μέσα που μπορούν να χρησιμοποιηθούν στην εκπαιδευτική διαδικασία. Χαρακτηριστικό παράδειγμα είναι τα εκπαιδευτικά παιχνίδια τα οποία αποτελούν ένα ταχύτατα αναπτυσσόμενο πεδίο και προσφέρουν ποικίλες δυνατότητες για σχεδιασμό καινοτόμων εκπαιδευτικών δραστηριοτήτων. Ένα από τα κύρια χαρακτηριστικά των εκπαιδευτικών παιχνιδιών είναι ότι μπορούν να προσφέρουν ευχάριστες και ελκυστικές εκπαιδευτικές δραστηριότητες στους χρήστες οι οποίες μπορούν να προσελκύουν το ενδιαφέρον τους και να δημιουργήσουν και να ενισχύσουν τα κίνητρα τους για μάθηση (Burguillo, 2010; Dickey, 2011; Ebner & Holzinger, 2007; Harris & Reid, 2005; Liu & Chu, 2010). Τα εκπαιδευτικά παιχνίδια αποτελούν ένα πλούσιο περιβάλλον μάθησης και προσφέρουν ποικίλους και αποτελεσματικούς τρόπους απόκτησης γνώσης (Cagiltay, 2007; Papastergiou, 2009; Tüzün et al., 2009). Η χρήση των παιχνιδιών στην εκπαίδευση ενισχύει την ενασχόλησή των εκπαιδευόμενων με εκπαιδευτικές δραστηριότητες και συμβάλλει στη βελτίωση των επιδόσεων τους (Hwang et al., 2012). Σε αυτή την κατεύθυνση, πολλές κατηγορίες εκπαιδευτικών παιχνιδιών έχουν αναπτυχθεί για να αντιμετωπίσουν προβλήματα μάθησης όπως για παράδειγμα το πρόβλημα της εγκατάλειψης της εκπαιδευτικής διαδικασίας και της έλλειψης κινήτρων μάθησης, και έχουν ως στόχο να γίνουν οι εκπαιδευτικές διαδικασίες πιο αποτελεσματικές και ευχάριστες προς τους εκπαιδευόμενους (Prensky, 2001; Yessad et al., 2010).

Ιδιαίτερα στις μέρες μας, υπάρχει έντονο ενδιαφέρον για την αξιοποίηση των παιχνιδιών ως εκπαιδευτικά εργαλεία με κύριο στόχο να βοηθήσουν τους διδάσκοντες στην

διδασκαλία καθώς και τους εκπαιδευόμενους στην εκμάθηση εννοιών με έναν ευχάριστο και διασκεδαστικό τρόπο (Wu et al., 2012). Στο γεγονός αυτό συνέβαλε ιδιαίτερα η τεχνολογική εξέλιξη και η ευρεία χρήση φορητών υπολογιστών και κινητών συσκευών. Η μάθηση μέσω εκπαιδευτικών παιχνιδιών (game-based learning) έχει ως στόχο να διδάξει μέσα από το περιβάλλον των παιχνιδιών στους χρήστες έννοιες και διαδικασίες καθώς επίσης και να δώσει την δυνατότητα στους χρήστες να εφαρμόσουν διαδικασίες που έχουν διδαχθεί, ενώ παράλληλα να παραμένουν οι εκπαιδευτικές δοκιμασίες αποτελεσματικές καθώς και το παιχνίδι να διατηρεί τον αλληλεπιδραστικό και διασκεδαστικό χαρακτήρα του (Prensky, 2001; Kapp, 2012). Το κύριο όφελος αυτής της κατηγορίας παιχνιδιών σε σύγκριση με τα συστήματα ηλεκτρονικής μάθησης (e-learning) είναι ότι αναπτύσσουν ισχυρά κίνητρα για μάθηση στους εκπαιδευόμενους μέσω του τεχνητού περιβάλλοντος που διαθέτουν, κάνοντας την μάθηση πιο ελκυστική και ευχάριστη για τους παίκτες.

Τα εκπαιδευτικά παιχνίδια μπορούν να χρησιμοποιηθούν για να διδάξουν έννοιες σε διάφορους τομείς και ιδιαίτερα στον τομέα της επιστήμης των υπολογιστών και να αποτελέσουν έναν αποτελεσματικό τρόπο εκπαίδευσης και απόκτησης γνώσεων (Sung&Hwang, 2013). Πολλά εκπαιδευτικά παιχνίδια έχουν χρησιμοποιηθεί με επιτυχία για τη διδασκαλία μαθημάτων της επιστήμης των υπολογιστών και της τεχνητής νοημοσύνης (Bayliss, 2007; Wong et al., 2010; Taylor, 2011; Sosnowski et al., 2013; DeNero & Klein, 2010) προσφέροντας καινοτόμες και αποτελεσματικές εκπαιδευτικές δραστηριότητες. Οι έρευνες έχουν δείξει ότι μέσα από τα εκπαιδευτικά παιχνίδια οι παίκτες αποκτούν ιδιαίτερα κίνητρα μελέτης, αυξάνοντας το ενδιαφέρον των εκπαιδευόμενων και συμβάλλοντας στο να γίνει πιο αποτελεσματική και πιο ευχάριστη η απόκτηση γνώσης και η ενασχόληση με εκπαιδευτικές διαδικασίες (Prensky, 2001).

Στα πλαίσια αυτής της ενότητας παρουσιάζουμε δυο εκπαιδευτικά παιχνίδια και ένα τρισδιάστατο εκπαιδευτικό περιβάλλον εικονικής πραγματικότητας που αναπτύχθηκαν για την εκμάθηση διαδικασιών της τεχνητής νοημοσύνης και πιο συγκεκριμένα για την εκμάθηση της λειτουργίας των αλγορίθμων αναζήτησης και της ικανοποίησης περιορισμών. Αρχικά, γίνεται παρουσίαση των παιδαγωγικών προσεγγίσεις στις οποίες βασίστηκε ο σχεδιασμός των παιχνιδιών και γίνεται προσδιορισμός των μαθησιακών τους στόχων. Στην συνέχεια, παρουσιάζουμε το εκπαιδευτικό παιχνίδι Edu-Pacman (Grivokostopoulou et al., 2016c) το οποίο βασίζεται στο κλασικό παιχνίδι Pacman και το οποίο αναπτύχθηκε με στόχο να βοηθήσει τους παίκτες να κατανοήσουν μέσω διαφόρων δραστηριοτήτων το τρόπο λειτουργίας των αλγορίθμων αναζήτησης. Έπειτα, παρουσιάζουμε το εκπαιδευτικό παιχνίδι Coloring Map Game (Hatzilygeroudis et al., 2012a ; Hatzilygeroudis et al., 2012b) που έχει ως στόχο να βοηθήσει τους παίκτες να μάθουν έννοιες της ικανοποίησης περιορισμών. Οι παίκτες μέσω του παιχνιδιού μαθαίνουν παίζοντας να εφαρμόζουν τους αλγορίθμους

ικανοποίησης περιορισμών καθώς και τις έννοιες της διατήρησης συνέπειας και της οπισθοδρόμησης. Επίσης, παρουσιάζουμε το τρισδιάστατο εκπαιδευτικό περιβάλλον VR-ALG το οποίο έχει στόχο να δώσει την δυνατότητα στους εκπαιδευόμενους να μελετήσουν και να εφαρμόσουν αλγορίθμους αναζήτησης σε διάφορες εκπαιδευτικές δραστηριότητες και δοκιμασίες.

6.2 Σχετικές Εργασίες

Η ανάπτυξη εκπαιδευτικών παιχνιδιών έχει προσελκύσει το έντονο ενδιαφέρον της εκπαιδευτικής κοινότητας και η χρήση και ενσωμάτωση των παιχνιδιών στην εκπαίδευση έχει αυξηθεί με ραγδαίους ρυθμούς σε όλες τις βαθμίδες. Στην διεθνή βιβλιογραφία υπάρχουν αρκετές αναφορές για την χρήση εκπαιδευτικών παιχνιδιών στο πεδίο της επιστήμης των υπολογιστών και της τεχνητής νοημοσύνης. Πιο συγκεκριμένα, στην εργασία (Levitin & Papalaskari, 2002) παρουσιάζεται η χρήση εκπαιδευτικών παιχνιδιών πάζλ ως μέσο για την διδασκαλία της σχεδίασης και της ανάλυσης διαφόρων ειδών αλγορίθμων. Στην εργασία (Markov et al., 2006) παρουσιάζεται μια επέκταση του παιχνιδιού N-puzzle για την διδασκαλία βασικών θεμάτων των αλγορίθμων αναζήτησης μέσω διαφόρων εκπαιδευτικών δραστηριοτήτων. Η εκπαιδευτική εφαρμογή SEPIA (Strategy Engine for Programming Intelligent Agents) (Sosnowski et al., 2013) είναι μια μηχανή στρατηγικής για τον προγραμματισμό ευφύων πρακτόρων. Πιο συγκεκριμένα, βασίζεται σε παιχνίδια στρατηγικής πραγματικού χρόνου, υποστηρίζει την ανάπτυξη ευφύων πρακτόρων και το βασικότερο χαρακτηριστικό είναι ότι μέσω των δραστηριοτήτων που προσφέρει μπορεί να βοηθήσει στην διδασκαλία θεμάτων σχεδιασμού ενεργειών (planning) της τεχνητής νοημοσύνης. Το εκπαιδευτικό παιχνίδι Ticket to Ride (Chang et al., 2008) σχεδιάστηκε και χρησιμοποιήθηκε για να βοηθήσει τους εκπαιδευόμενους να μάθουν διάφορα θέματα γράφων και αλγορίθμων όπως οι αλγόριθμοι του Kruskal, Prim και Dijkstra. Μέσω του παιχνιδιού, οι παίκτες μαθαίνουν να εφαρμόζουν τους παραπάνω αλγορίθμους σε διάφορα σενάρια του παιχνιδιού τα οποία σχετίζονται με την ένωση πόλεων. Το Robocode (Hartness, 2004) είναι ένα παιχνίδι προγραμματισμού σε Java όπου ο στόχος των παικτών είναι να σχεδιάσουν ρομπότ που θα ανταγωνιστούν ενάντια σε άλλα ρομπότ σε πεδίο μάχης. Ο παίκτης προγραμματίζει το ρομπότ του και καθορίζει με χρήση μεθόδων τεχνητής νοημοσύνης πώς να συμπεριφέρεται και να αντιδρά σε γεγονότα που συμβαίνουν στην μάχη. Το παιχνίδι έχει ως στόχο να βοηθήσει τους παίκτες να μάθουν προγραμματισμό σε Java με έναν ευχάριστο τρόπο καθώς μπορούν να μάθουν την ανάπτυξη των αλγορίθμων και το παιχνίδι είναι κατάλληλο τόσο για αρχάριους προγραμματιστές όσο και για προχωρημένους. Το παιχνίδι Wu's Castle (Eagle & Barnes 2008) είναι ένα δισδιάστατο παιχνίδι ρόλων όπου οι εκπαιδευόμενοι μαθαίνουν τον τρόπο λειτουργίας των δομών επανάληψης και των πινάκων με αλληλεπιδραστικό τρόπο και

μέσω κατάλληλων γραφικών αναπαραστάσεων. Επίσης, το παιχνίδι παρέχει άμεση ανάδραση και καθοδήγηση στους εκπαιδευόμενους και επίσης τους δίνει την δυνατότητα να πειραματιστούν σε ένα ασφαλές περιβάλλον.

6.3 Παιδαγωγικές Προσεγγίσεις και μαθησιακοί στόχοι

Ο σχεδιασμός των εκπαιδευτικών παιχνιδιών που αναπτύξαμε βασίστηκε σε *θεωρίες ανάπτυξης κινήτρων* (*motivational theories*) και σε *θεωρίες μάθησης* (*learning theories*). Στην διεθνή βιβλιογραφία υπάρχουν διάφορες προσεγγίσεις μάθησης που μπορούν να χρησιμοποιηθούν σε ένα παιχνίδι, όπως η μάθηση μέσω της πράξης (*learning by doing*), η μάθηση από τα λάθη (*learning from mistakes*), η μάθηση προσανατολισμένη στον στόχο (*goal-oriented learning*), τα παιχνίδια ρόλων (*roleplaying*) και η κατασκευαστική μάθηση (*constructivist learning*) (Burgos et al., 2007; Prensky, 2001). Πιο συγκεκριμένα στα πλαίσια του σχεδιασμού των εκπαιδευτικών παιχνιδιών, βασιστήκαμε στην θεωρία ανάπτυξης κινήτρων του Malone (Malone, 1981) η οποία δίνει έμφαση στο πώς να γίνουν οι εκπαιδευτικές διαδικασίες και ο σχεδιασμός των εκπαιδευτικών παιχνιδιών πιο διασκεδαστικοί και να παρέχουν κίνητρα στους εκπαιδευόμενους μέσω δραστηριοτήτων οι οποίες δημιουργούν προκλήσεις, θέτουν στόχους στους εκπαιδευόμενους, παρέχουν κατάλληλη ανάδραση, αβέβαιη έκβαση των ενεργειών των χρηστών, ενισχύουν την αυτοεκτίμηση, καλλιεργούν τις δεξιότητες, προκαλούν την περιέργεια και επίσης ενισχύουν την φαντασία (Malone & Lepper, 1987).

Οι μαθησιακοί στόχοι (*learning objective*) των παιχνιδιών έχουν ως στόχο να βοηθήσουν τους παίκτες να μάθουν και να κατανοήσουν καλύτερα την λειτουργία των αλγορίθμων αναζήτησης και της ικανοποίησης περιορισμών με έναν ευχάριστο και διασκεδαστικό τρόπο. Πιο συγκεκριμένα, στα πλαίσια των παιχνιδιών έχουν σχεδιαστεί διάφορες εκπαιδευτικές δραστηριότητες, οι οποίες προσφέρουν την δυνατότητα στους παίκτες αλληλεπιδρώντας με το παιχνίδι να κατανοήσουν τον τρόπο λειτουργία των αλγορίθμων και κυρίως να μάθουν να τους εφαρμόζουν σε διάφορα προβλήματα. Οι μαθησιακοί στόχοι του παιχνιδιού διαμορφώθηκαν με βάση την *αναθεωρημένη ταξινόμια του Bloom* (*Revised Bloom taxonomy*) (Anderson et al., 2001). Η ταξινόμια του Bloom (Bloom 1956) είναι μια κατηγοριοποίηση των διαφορετικών στόχων που οι διδάσκοντες θέτουν στους εκπαιδευόμενους και διαχωρίζει τους εκπαιδευτικούς στόχους σε τρεις τομείς: τον *γνωστικό* (*cognitive*), τον *συναισθηματικό* (*affective*) και το *ψυχοκινητικό* (*psychomotor*). Η αναθεωρημένη ταξινόμια του Bloom αποτελείται από 6 επίπεδα και ξεκινάει από τις χαμηλότερου επιπέδου δεξιότητες προς τις υψηλότερες. Τα έξι επίπεδα τη ταξινόμιας είναι:

Θυμάμαι, Κατανοώ, Εφαρμόζω, Αναλύω, Αξιολογώ, Δημιουργώ και παρουσιάζονται στην εικόνα 6.1.



Εικόνα 6.1 Αναθεωρημένη Ταξινόμια Bloom

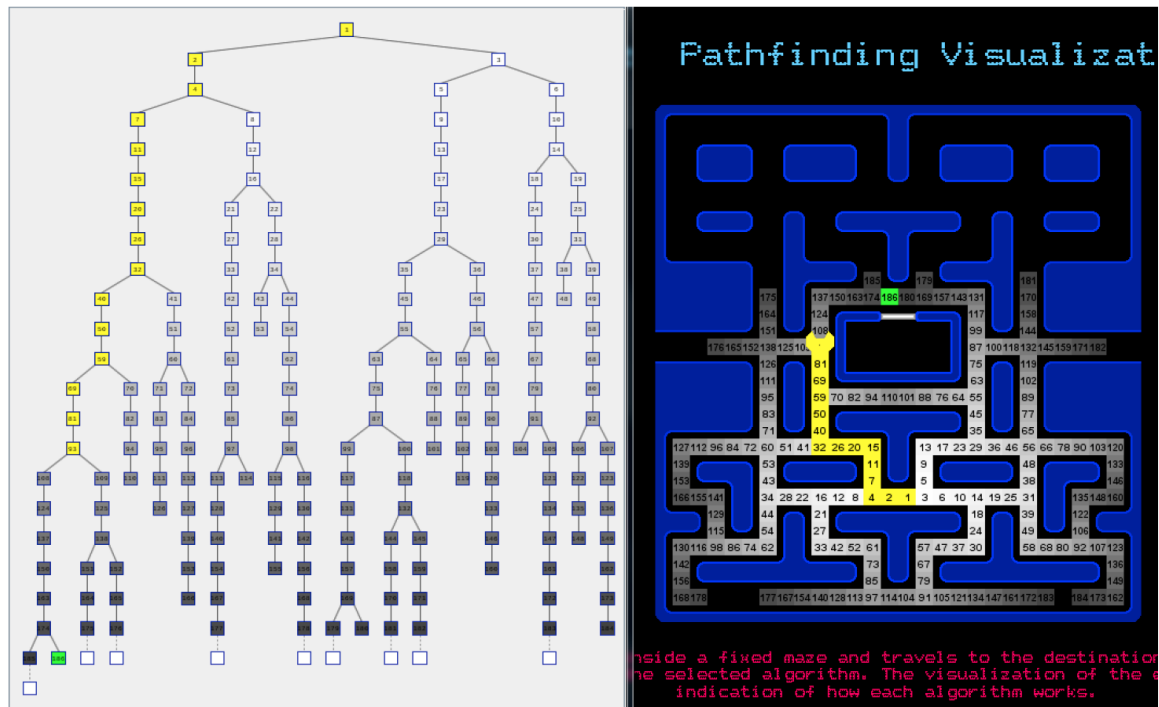
6.4 Εκμάθηση Αλγορίθμων με το Παιχνίδι Edu-PacMan

Το Edu-Pacman είναι ένα εκπαιδευτικό παιχνίδι που έχει τα βασικά χαρακτηριστικά του παιχνιδιού Pacman και συνδυάζει το παιχνίδι με την μάθηση. Πιο συγκεκριμένα, το παιχνίδι έχει ως στόχο να βοηθήσει τους παίκτες να μάθουν μέσω του παιχνιδιού τον τρόπο λειτουργίας των αλγορίθμων αναζήτησης με έναν ευχάριστο και διασκεδαστικό τρόπο. Το παιχνίδι βασίστηκε στο κλασσικό παιχνίδι Pacman και κύριος στόχος του είναι να ενισχύσει την κατασκευή γνώσης των παικτών και να συμβάλλει την καλύτερη κατανόηση των αλγορίθμων αναζήτησης. Οι παίκτες αλληλεπιδρώντας με το παιχνίδι μπορούν να δουν μέσω οπτικοποιήσεων τον τρόπο λειτουργίας των αλγορίθμων αναζήτησης και να εφαρμόσουν αλγόριθμους αναζήτησης σε διάφορες δοκιμασίες και λαβύρινθους του παιχνιδιού.

6.4.1 Εκπαιδευτικές δραστηριότητες του παιχνιδιού

Στα πλαίσια του εκπαιδευτικού παιχνιδιού, σχεδιάστηκαν μια σειρά από εκπαιδευτικές δραστηριότητες μέσω των οποίων ο παίκτης μελετάει και εφαρμόζει τους αλγόριθμους αναζήτησης σε δοκιμασίες και σε εκπαιδευτικές δραστηριότητες του παιχνιδιού. Αρχικά, ο παίκτης μπορεί να επιλέξει τον αλγόριθμο που επιθυμεί να μελετήσει και στην συνέχεια το παιχνίδι του προσφέρει την δυνατότητα στο παίκτη να δει το θεωρητικό μέρος του αλγορίθμου, τον ψευδοκώδικα, το διάγραμμα ροής, την εφαρμογή του μέσω οπτικοποιήσεων καθώς και να εφαρμόσει τον αλγόριθμο σε διάφορες εκπαιδευτικές καταστάσεις και δραστηριότητες. Η οπτικοποίηση των αλγορίθμων στα πλαίσια του παιχνιδιού γίνεται, παρουσιάζοντας με έναν εύληπτο, επεξηγηματικό και γραφικό τρόπο την εφαρμογή των

αλγορίθμων αναζήτησης σε διάφορους λαβύρινθους και καταστάσεις του παιχνιδιού (εικόνα 6.2). Πιο συγκεκριμένα, η οπτικοποίηση αναλύει και επεξηγεί με γραφικό τρόπο την λειτουργία των αλγορίθμων αναζήτησης. Στα πλαίσια του παιχνιδιού, η οπτικοποίηση μπορεί να δείξει τον τρόπο που θα κινηθεί ο Pacman καθώς εφαρμόζει έναν συγκεκριμένο αλγόριθμο προκειμένου να φτάσει στον στόχο του μέσα σε έναν λαβύρινθο.



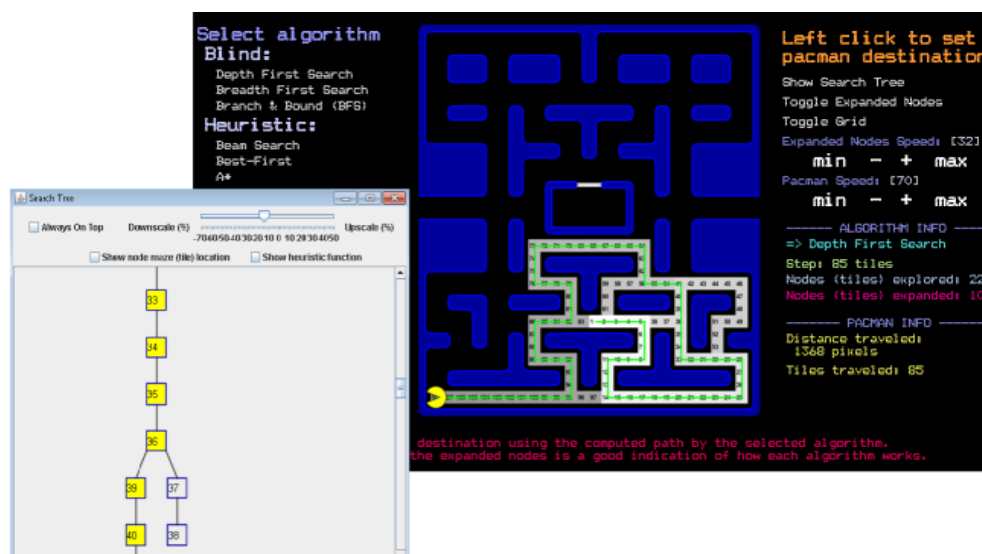
Εικόνα 6.2 Παράδειγμα οπτικοποίησης

Κατά την διάρκεια της αλληλεπίδρασης μπορεί να ερωτηθεί ο παίκτης να προσδιορίσει την επόμενη κίνηση σύμφωνα με το βήμα του αλγορίθμου (εικόνα 6.3). Στην περίπτωση που ο παίκτης απαντήσει σωστά, μπορεί να του ζητηθεί να αιτιολογήσει την τρέχουσα επιλογή του απαντώντας σε ερωτήσεις πολλαπλής επιλογής ενώ στην περίπτωση που ο παίκτης απαντήσει λάθος, παρουσιάζεται η σωστή απάντηση και κατάλληλες επεξηγήσεις. Το παιχνίδι παρέχει στους εκπαιδευόμενος τη δυνατότητα να επιλέξουν τον τύπο του αλγορίθμου που επιθυμούν να μελετήσουν και στη συνέχεια να δουν την εφαρμογή του στον λαβύρινθο. Στα πλαίσια της εφαρμογής, παρουσιάζονται επιπλέον επεξηγήσεις για τις ενέργειες που πραγματοποιούνται από τον αλγόριθμο, πληροφορίες για το κόστος, την ευρετική συνάρτηση κάθε κόμβου του λαβύρινθου καθώς και τα βήματα που χρειάστηκαν να γίνουν.



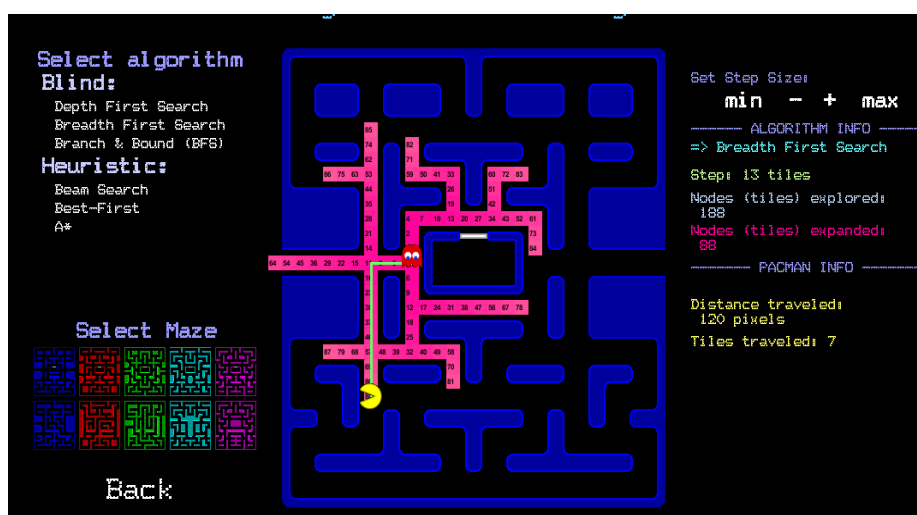
Εικόνα 6.3 Αλληλεπίδραση με τον χρήστη και καθορισμός επόμενης κίνησης.

Για παράδειγμα, στην εικόνα 6.4 παρουσιάζεται η εφαρμογή του αλγορίθμου κατά βάθος αναζήτησης, όπως εφαρμόζεται από τον Pacman προκειμένου να φτάσει στο στόχο (power-up) στην κάτω αριστερά γωνία του λαβυρίνθου ξεκινώντας από την προεπιλεγμένη θέση εκκίνησης και επίσης παρουσιάζεται η αντίστοιχη δεντρική αναπαράσταση του λαβυρίνθου. Η γραφική αναπαράσταση δείχνει το τρόπο που κινήθηκε ο Pacman σύμφωνα με ένα συγκεκριμένο αλγόριθμο. Οι κόμβοι του γραφήματος είναι τα τετράγωνα του λαβυρίνθου και ανάλογα με το είδος του αλγορίθμου παρουσιάζονται πρόσθετες πληροφορίες όπως ευρετικά και κόστος για να βοηθήσουν τους παίκτες να κατανοήσουν πως έγιναν οι αποφάσεις για μετακίνηση.



Εικόνα 6.4 Οπτικοποίηση αλγορίθμου Κατά Βάθος αναζήτησης

Επιπλέον, το παιχνίδι παρέχει διάφορα παραδείγματα για το πώς κινούνται τα φαντάσματα (ghosts) σύμφωνα με ένα συγκεκριμένο αλγόριθμο. Πιο συγκεκριμένα, το παιχνίδι μπορεί να εστιάσει σε συγκεκριμένες πτυχές που αφορούν την κίνηση ενός φαντάσματος μέσα στο λαβύρινθο σύμφωνα με ένα συγκεκριμένο αλγόριθμο. Στην εικόνα 6.5 παρουσιάζεται μια ενδεικτική κατάσταση όπου το φάντασμα έχοντας στόχο να φτάσει στον Pacman και εφαρμόζοντας το αλγόριθμο της κατά πλάτος αναζήτησης, έχει εξερευνήσει 188 πλακίδια, έχει επεκτείνει 88 πλακίδια και η απόσταση με τον Pacman είναι 13 πλακίδια. Η απόσταση μεταξύ του φαντάσματος και του Pacman υπολογίζεται ως η συντομότερη διαδρομή μεταξύ τους και αναπαρίσταται με πράσινη γραμμή. Το παιχνίδι καθώς κινούνται τόσο ο Pacman όσο και το φάντασμα μέσα στον λαβύρινθο απεικονίζει σε κάθε κατάσταση, τον αριθμό των πλακιδίων που το φάντασμα έχει διερευνήσει και επεκτείνει με βάση τον αλγόριθμο που υλοποιεί.



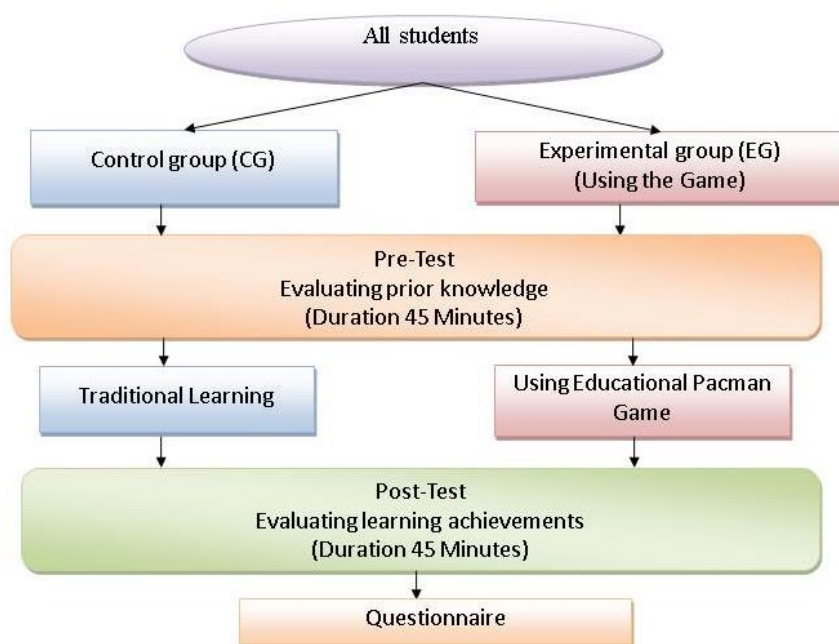
Εικόνα 6.5 Ανάλυση της λειτουργίας του αλγόριθμου μέσω της κίνησης του φαντάσματος.

Επίσης, αναπτύχθηκε μια ακόμη λειτουργία του παιχνιδιού που έχει στόχο να παρέχει μεγαλύτερες προκλήσεις στους παίκτες και μέσω του ανταγωνισμού να κεντρίσει περισσότερο το ενδιαφέρον τους. Σε αυτήν την λειτουργία ο παίκτης καλείται να επιλύσει προκαθορισμένες πίστες (λαβύρινθους) κάτω από διαφορετικές συνθήκες (π.χ. μέσα σε συγκεκριμένο χρόνο, με μικρό/μεγάλο πλήθος αντιπάλων-φαντασμάτων κ.α.), κάτι που δημιουργεί συνεχείς προκλήσεις και παρέχει κίνητρα στον παίκτη για επίτευξη υψηλότερης βαθμολογίας και καλύτερων επιδόσεων. Όταν ο παίκτης ολοκληρώνει ένα επίπεδο-δοκιμασία μεταβαίνει σε επόμενη δοκιμασία που είναι πιο περίπλοκη και έχει μεγαλύτερες προκλήσεις. Οι δοκιμασίες έχουν σχεδιαστεί με τέτοιο τρόπο ώστε να χρειάζεται να γίνει εφαρμογή συγκεκριμένου αλγορίθμου αναζήτησης και πραγματοποίηση κατάλληλων κινήσεων του Pacman από τον χρήστη προκειμένου να επιτευχθούν οι απαιτήσεις του επιπέδου. Οι γενικές απαιτήσεις της κάθε δοκιμασίας είναι ο Pacman να φτάσει σε ένα συγκεκριμένο στόχο μέσα

στο λαβύρινθο, που μπορεί να είναι ένα αντικείμενο (φρούτο, δύναμη power-up) και επίσης να έχει συλλέξει όλες τις τελείες του λαβυρίνθου εφαρμόζοντας συγκεκριμένο αλγόριθμο αναζήτησης. Ο παίκτης καλείται σε κάθε δοκιμασία ξεκινώντας από μια τυχαία θέση, να φτάσει στην επίτευξη του στόχου μετακινώντας τον Pacman με βάση το συγκεκριμένο αλγόριθμο που ζητάει το επίπεδο. Σε περίπτωση που ο χρήστης κάνει λάθος κίνηση, τότε τα φαντάσματα που βρίσκονται στο λαβύρινθο μπορεί να κινηθούν πιο γρήγορα προς τον Pacman και γενικότερα, ο μόνος τρόπος για να ολοκληρώσει ο παίκτης την δοκιμασία είναι να εφαρμόσει σωστά τον αλγόριθμο. Στα πλαίσια του παιχνιδιού, καθώς ο παίκτης μεταβαίνει σε επόμενες δοκιμασίες, οι δοκιμασίες γίνονται όλο και πιο περίπλοκες με βάση παραμέτρους όπως το είδος του λαβυρίνθου, ο αριθμός των φαντασμάτων στο λαβύρινθο, οι στόχοι που πρέπει να επιτευχθούν και επίσης η πολυπλοκότητα του αλγορίθμου και των παραμέτρων του. Το παιχνίδι παρέχει στους παίκτες στατιστικά στοιχεία, όπως τον χρόνο που χρειάστηκε για την επίλυση μιας δοκιμασίας, την βαθμολογία (score) κ.α.

6.4.2 Πειραματική Αξιολόγηση Edu-Pacman

Για να μελετήσουμε την αποτελεσματικότητα του εκπαιδευτικού παιχνιδιού Edu-Pacman πραγματοποιήθηκε μια πειραματική μελέτη στα πλαίσια του μαθήματος της τεχνητής νοημοσύνης. Σκοπός της μελέτης είναι να αξιολογήσουμε την αποτελεσματικότητα στην μάθηση των φοιτητών και την ανάπτυξη κινήτρων χρησιμοποιώντας το εκπαιδευτικό παιχνίδι Edu-Pacman για την εκμάθηση των αλγορίθμων αναζήτησης. Στην μελέτη συμμετείχαν 38 φοιτητές (άνδρες και γυναίκες) που παρακολουθούσαν το μάθημα της τεχνητής νοημοσύνης. Η διαδικασία που ακολουθήθηκε στην πειραματική μελέτη παρουσιάζεται αναλυτικά στην εικόνα 6.6. Όλοι συμμετέχοντες χωρίστηκαν τυχαία σε δύο ομάδες των 19 φοιτητών, οι οποίες ήταν η πειραματική ομάδα (experimental group - EG) και η ελέγχου ομάδα (control group - CG) αντίστοιχα.



Εικόνα 6.6 Πειραματική διαδικασία του edu-Pacman Game.

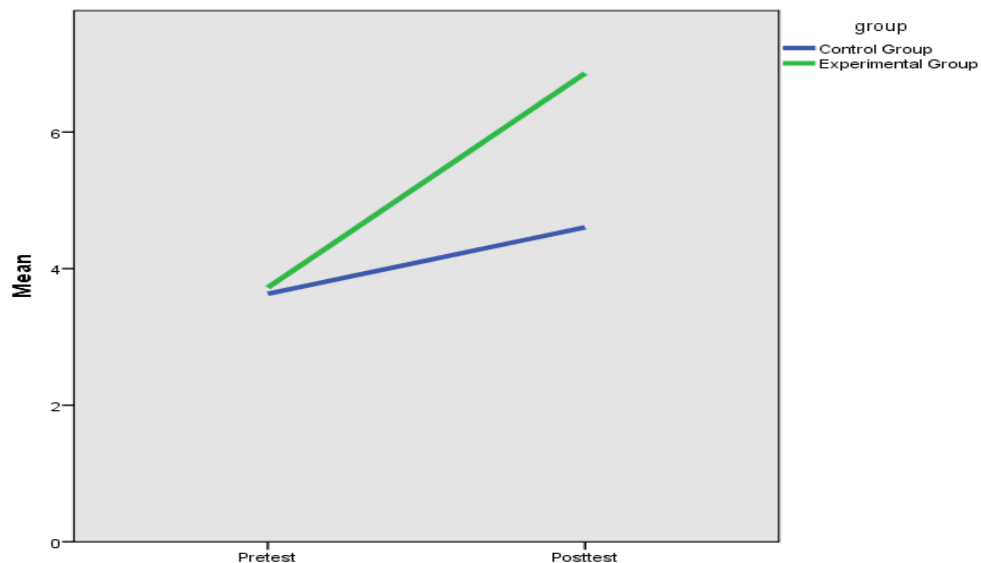
Αρχικά, οι φοιτητές και των δυο ομάδων συμμετείχαν σε ένα pretest που είχε ως στόχο να αξιολογήσει το επίπεδο γνώσεων και να ελέγξει αν και οι δύο ομάδες αρχικά είχαν το ίδιο επίπεδο γνώσεων πριν την διεξαγωγή της πειραματικής μελέτης. Το pretest περιείχε δώδεκα ασκήσεις σχετικά με τους αλγόριθμους αναζήτησης και η διάρκεια του τεστ ήταν 45 λεπτά. Μετά την διεξαγωγή του τεστ, οι επιδόσεις των εκπαιδευόμενων βαθμολογήθηκαν από τον διδάσκοντα στην κλίμακα από 0 έως 10 μονάδες. Στην συνέχεια, η πειραματική ομάδα (experimental group) αλληλεπίδρασε με το εκπαιδευτικό παιχνίδι edu-Pacman για δύο εβδομάδες ενώ η ομάδα ελέγχου (control group) χρησιμοποίησε τον παραδοσιακό τρόπο μάθησης και μελέτησε μέσω διαφανειών και παραδειγμάτων τις αντίστοιχες έννοιες. Μετά από την περίοδο των 2 εβδομάδων, οι φοιτητές και των δυο ομάδων συμμετείχαν στο τελικό posttest. Το posttest είχε ως σκοπό να αξιολογηθεί η επίδοση των εκπαιδευόμενων μετά την ενασχόλησή τους με το παιχνίδι και την συμμετοχή τους στις αντίστοιχες δραστηριότητες μάθησης του παιχνιδιού. Το posttest περιείχε δώδεκα ασκήσεις και το επίπεδο δυσκολίας του ήταν ίδιο με αυτό του pretest.

Αρχικά, έγινε μια μονόδρομη ανάλυση διασποράς (one-way Analysis of Variance - ANOVA) στις απαντήσεις των εκπαιδευόμενων στο pretest. Η μέση τιμή και η τυπική απόκλιση για την ομάδα ελέγχου (Control Group - CG) ήταν $M=3.63$ και $SD=0.57$ και αντίστοιχα για την πειραματική ομάδα (Experimental - EG) ήταν $M=3.70$ και $SD=0.62$ (εικόνα 6.6). Από τα αποτελέσματα φαίνεται ότι γίνεται δεκτή η μηδενική υπόθεση με επίπεδο σημαντικότητας $\alpha=0.05$ και επομένως θεωρούμε ότι δεν υπάρχει σημαντική διαφορά

μεταξύ των φοιτητών των ομάδων ($p = 0.648 > 0.05$, $F = 0.21$) και φαίνεται ότι οι δύο ομάδες είχαν ισοδύναμο επίπεδο γνώσεων πριν από τη χρήση του παιχνιδιού. Στη συνέχεια, πραγματοποιήσαμε μια ανάλυση συνδιακύμανσης (ANCOVA) για να ελέγξουμε τη διαφορά μεταξύ των δύο ομάδων, χρησιμοποιώντας τα αποτελέσματα pretest ως συμμεταβλητή και τις βαθμολογίες στο posttest ως εξαρτημένες μεταβλητές. Στον πίνακα 6.1 παρουσιάζονται τα περιγραφικά μέτρα που προέκυψαν για το posttest και στην εικόνα 6.7 παρουσιάζονται οι μέσες τιμές των δυο ομάδων και τα αποτελέσματα δείχνουν μεγάλη αύξηση της μέσης τιμής της πειραματικής ομάδας σε σχέση με την ομάδα ελέγχου.

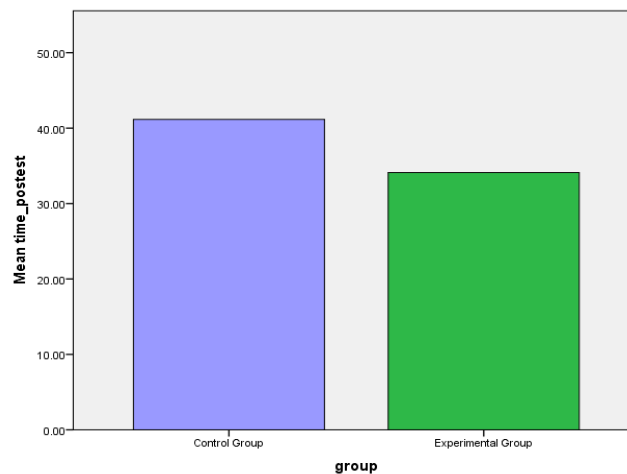
Πίνακας 6.1 Πειραματικά αποτελέσματα Posttest για το Pacman

Group	N	Mean	S. D.
Control Group (CG)	19	4.605	0.698
Experimental Group (EG)	19	6.861	0.782



Εικόνα 6.7 Αποτελέσματα Pretest και Posttest για κάθε ομάδα

Από τα αποτελέσματα της ανάλυσης ANCOVA φαίνεται ότι οι διαφορές στις βαθμολογίες στο posttest είναι στατιστικά και σημαντικά διαφορετικές μεταξύ των δύο ομάδων ($F = 83.14$, $p = 0.00 < 0.05$). Από τα αποτελέσματα προκύπτει ότι η απόδοση των φοιτητών της πειραματικής ομάδας, που χρησιμοποίησε το παιχνίδι Edu-Pacman, ήταν καλύτερη από εκείνη της ομάδας ελέγχου. Επίσης, στην εικόνα 6.8, παρουσιάζονται τα αποτελέσματα για τον χρόνο που χρειάστηκαν οι εκπαιδευόμενοι για την επίλυση των ασκήσεων του posttest με μέσο χρόνο τα 34.1 λεπτά για την ομάδα ελέγχου ενώ για την πειραματική ομάδα ο μέσος χρόνος ήταν 41.1.



Εικόνα 6.8 Χρόνος Επίλυσης Posttest

Στη συνέχεια, οι φοιτητές της πειραματικής ομάδας (experimental group) που χρησιμοποίησαν το παιχνίδι, κλήθηκαν να συμπληρώσουν ένα ερωτηματολόγιο για την αξιολόγηση της ευχρηστίας του παιχνιδιού, της βοήθειας στην κατανόηση των αλγορίθμων αναζήτησης καθώς και να εκφράσουν την γνώμη τους και τις εμπειρίες τους στα πλαίσια της αλληλεπίδρασης με το παιχνίδι. Το ερωτηματολόγιο περιελάμβανε 15 ερωτήσεις. Οι ερωτήσεις Q1-Q12 βασίστηκαν στην κλίμακα Likert (1: καθόλου, 5: πάρα πολύ) και οι ερωτήσεις Q13-Q15 ήταν ανοικτού τύπου και καλούσαν τους φοιτητές να εκφράσουν την γνώμη τους, να τονίσουν θετικά και αρνητικά σημεία του παιχνιδιού, προβλήματα που πιθανώς αντιμετώπισαν καθώς και να προτείνουν βελτιώσεις που θα ήθελαν να γίνουν στο παιχνίδι.

Πίνακας 6.2: Αποτελέσματα ερωτηματολογίου.

ΕΡΩΤΗΣΕΙΣ	ΑΠΑΝΤΗΣΕΙΣ (%)				
	1	2	3	4	5
1 Η ενασχόληση με το παιχνίδι ήταν διασκεδαστική	0	0	0	5.3	94.7
2 Η διεπαφή του παιχνιδιού είναι εύκολη στη χρήση.	0	0	0	10.5	89.5
3 Η ενασχόληση με το παιχνίδι σας έκανε πιο ενεργό στο μάθημα	0	0	0	21.0	78.9
4 Το παιχνίδι μου έδωσε κίνητρα για μελέτη.	0	0	10.5	10.5	78.9
5 Το παιχνίδι ενίσχυσε την συμμετοχή μου στο μάθημα.	0	0	5.3	15.8	78.9
6 Η ενασχόληση με το παιχνίδι με αύξησε το ενδιαφέρον μου για μελέτη.	0	0	5.3	15.8	78.9
7 Η ενασχόληση με το παιχνίδι είναι πιο διασκεδαστική από άλλους τρόπους μάθησης	0	0	0	10.5	89.5

8	Το παιχνίδι με βοήθησε να μάθω τους αλγόριθμους αναζήτησης.	0	0	0	5.3	94.7
9	Το παιχνίδι με βοήθησε να κατανοήσω εις βάθος τους αλγόριθμους αναζήτησης	0	0	5.3	10.5	84.2
10	Η οπτικοποίηση των αλγορίθμων με βοήθησε να καταλάβω τον τρόπο λειτουργίας τους	0	0	0	10.5	89.5
11	Το παιχνίδι μου παρείχε νέους τρόπους σκέψης σχετικά με τους αλγόριθμους αναζήτησης.	0	0	0	10.5	89.5
12	Θα συνιστούσατε το παιχνίδι στους συμφοιτητές σας και θα θέλατε να ενσωματωθεί στο μάθημα;	0	0	5.3	10.5	84.2

Μετά την ανάλυση των απαντήσεων των φοιτητών στο ερωτηματολόγιο, υπολογίστηκε η μετρική Cronbach για να προσδιοριστεί η αξιοπιστία του ερωτηματολογίου. Ο συντελεστής αξιοπιστίας Cronbach είναι $\alpha=0.97$, δηλαδή είναι μεγαλύτερος από 0.8 κάτι που σημαίνει ότι το ερωτηματολόγιο έχει υψηλή αξιοπιστία. Τα αποτελέσματα της ανάλυσης των απαντήσεων του ερωτηματολογίου δείχνουν ότι οι φοιτητές ήταν πολύ θετικοί και ότι η πλειοψηφία των φοιτητών (94.7%) βρήκαν σε μεγάλο βαθμό την εκμάθηση μέσω του παιχνιδιού να είναι διασκεδαστική και επίσης ένα μεγάλο ποσοστό (89.5%) βρήκε το παιχνίδι εύκολο στη χρήση. Επιπλέον, ένα μεγάλο μέρος των φοιτητών δήλωσε ότι το παιχνίδι τους έδωσε επιπλέον κίνητρα για μελέτη (78.9%), αύξησε την μελέτη τους (78.9%) και το ενδιαφέρον τους (78.9%) και έγιναν πιο ενεργητικοί στο μάθημα (78.9%). Επιπρόσθετα, όσον αφορά την αποτελεσματικότητα στην μάθηση μέσω του παιχνιδιού, οι εκπαιδευόμενοι δήλωσαν ότι το παιχνίδι τους βοήθησε να μάθουν πιο αποτελεσματικά (94.7%) και να κατανοήσουν εις βάθος το τρόπο εφαρμογής των αλγορίθμων αναζήτησης (84.2%). Η αλληλεπίδραση με το παιχνίδι κατά τη διάρκεια εφαρμογής των αλγορίθμων και η οπτικοποίηση τους, βοήθησε τους φοιτητές να κατανοήσουν καλύτερα τον τρόπο λειτουργίας σε πρακτικό επίπεδο (89.5%). Γενικά, τα πειραματικά αποτελέσματα έδειξαν ότι το παιχνίδι βοήθησε τους εκπαιδευόμενους να κατανοήσουν εις βάθος τους αλγόριθμους αναζήτησης και τον τρόπο λειτουργίας τους με έναν ευχάριστο και διασκεδαστικό τρόπο. Επιπλέον, το (89.5%) των φοιτητών δήλωσε ότι οι εκπαιδευτικές δραστηριότητες και οι δοκιμασίες του παιχνιδιού τους καλλιέργησαν ένα νέο τρόπο σκέψης της λειτουργίας των αλγορίθμων αναζήτησης. Τέλος, η πλειοψηφία των εκπαιδευόμενων (84.2%) πρότεινε να ενσωματωθεί το παιχνίδι στο μάθημα και να χρησιμοποιείται από τους φοιτητές.

6.5 Εκμάθηση Ικανοποίησης περιορισμών με το Εκπαιδευτικό Παιχνίδι Coloring-Map-Game

Στην ενότητα αυτή παρουσιάζουμε το εκπαιδευτικό παιχνίδι Coloring Map Game που σχεδιάστηκε για να βοηθήσει τους εκπαιδευόμενους να κατανοούν έννοιες της ικανοποίησης περιορισμών και των αλγορίθμων διατήρησης συνέπειας. Το Coloring Map Game είναι ένα παιχνίδι ενός παίκτη και για τον σχεδιασμό του βασιστήκαμε στο κλασικό πρόβλημα του χρωματισμού του χάρτη. Στα πλαίσια του παιχνιδιού οι παίκτες μαθαίνουν παίζοντας να εφαρμόζουν τους αλγορίθμους ικανοποίησης περιορισμών και τις έννοιες της διατήρησης συνέπειας και οπισθοδρόμησης σε διάφορες δοκιμασίες.

6.5.1 Το πεδίο της Ικανοποίησης Περιορισμών

Τα *Προβλήματα Ικανοποίησης Περιορισμών* (Constraint Satisfaction Problems - CSPs) είναι μαθηματικά προβλήματα που ορίζονται ως ένα σύνολο μεταβλητών των οποίων η κατάσταση πρέπει να ικανοποιεί έναν αριθμό από περιορισμούς. Πιο συγκεκριμένα, στα προβλήματα ικανοποίησης περιορισμών είναι γνωστές μερικές ιδιότητες της τελικής κατάστασης και επιδιώκεται η εύρεση ενός πλήρους στιγμιότυπου της τελικής κατάστασης. Αυτή η κατηγορία προβλημάτων είναι αντικείμενο μελέτης στον τομέα της τεχνητής νοημοσύνης καθώς βρίσκεται εφαρμογές σε πολλά πεδία όπως ο χρονοπρογραμματισμός ενεργειών, η διαχείριση πόρων, η χρηματιστηριακή ανάλυση και άλλα. Σύμφωνα με τον Nadel (Nadel, 1990) ένας πολύ μεγάλος αριθμός προβλημάτων στον τομέα της τεχνητής νοημοσύνης και σε άλλους τομείς τις επιστήμης των υπολογιστών μπορεί να θεωρηθεί ως ειδική περίπτωση προβλήματος ικανοποίησης περιορισμών και να επιλυθεί ως τέτοιου είδους πρόβλημα. Τα προβλήματα ικανοποίησης περιορισμών αναπαριστώνται από ένα σύνολο μεταβλητών όπου κάθε μία από αυτές μπορεί να πάρει τιμές από ένα συγκεκριμένο πεδίο τιμών και από ένα σύνολο περιορισμών δηλαδή σχέσεων μεταξύ των μεταβλητών. Με βάση αυτήν την αναπαράσταση, ως λύση ορίζεται η ανάθεση τιμών στις μεταβλητές του προβλήματος, η οποία δεν παραβιάζει τους περιορισμούς. Η επίλυση αυτών των προβλημάτων δεν μπορεί να γίνει με τους κλασσικούς αλγόριθμους αναζήτησης σε πραγματικό χρόνο, λόγω του φαινομένου της *συνδυαστικής έκρηξης* (combinatorial explosion) για αυτό χρησιμοποιούνται αλγόριθμοι *ελέγχου συνέπειας* (consistency) ή κλασσικοί αλγόριθμοι σε συνδυασμό με άλλους αλγόριθμους. Υπάρχουν πολλά προβλήματα στην καθημερινή ζωή και σε διάφορους τομείς που μπορούν να αναπαρασταθούν και να επιλυθούν ως προβλήματα ικανοποίησης περιορισμών.

Ένα πρόβλημα ικανοποίησης περιορισμών (Constraint Satisfaction Problem - CSP) αναπαριστάται από:

- Ένα σύνολο n μεταβλητών $V_1, V_2, \dots, V_n$
- Ένα σύνολο n πεδίων τιμών $D_1, D_2, \dots, D_n$ που αντιστοιχούν σε κάθε μεταβλητή έτσι ώστε $V_i \in D_i$
- Ένα σύνολο σχέσεων (περιορισμών) $C_1, C_2, \dots, C_n$ μια σχέση μεταξύ των μεταβλητών του προβλήματος.

Ανάλογα με το πόσες μεταβλητές περιλαμβάνει ένας περιορισμός χαρακτηρίζεται ως: μοναδιαίος (unary), δυαδικός (binary) ή ανώτερης τάξης (higher order).

Αλγόριθμοι Ελέγχου Συνέπειας

Υπάρχουν τρεις βασικές κατηγορίες ελέγχου συνέπειας (Consistency check): οι αλγόριθμοι συνέπειας κόμβου (Node Consistency), οι αλγόριθμοι συνέπειας τόξου (Arc Consistency - AC) και οι αλγόριθμοι συνέπειας μονοπατιού (Path Consistency algorithms). Η βασική ιδέα όλων των αλγόριθμων ελέγχου συνέπειας είναι η απαλοιφή από τα αρχικά πεδία των μεταβλητών εκείνων των τιμών οι οποίες δεν μπορούν να συμμετέχουν στην τελική λύση. Αυτό επιτυγχάνεται με τον έλεγχο συνέπειας, δηλαδή με την διαγραφή από το πεδίο τιμών κάθε μεταβλητής όλων εκείνων των τιμών που παραβαίνουν τους περιορισμούς. Ο αλγόριθμος AC-3 είναι ο απλούστερος αλγόριθμος συνέπειας τόξου όπου ελέγχει όλους τους δυαδικούς περιορισμούς και διαγράφει τις αντίστοιχες τιμές που δεν ικανοποιούν τους περιορισμούς αυτούς από τα πεδία τιμών των μεταβλητών.

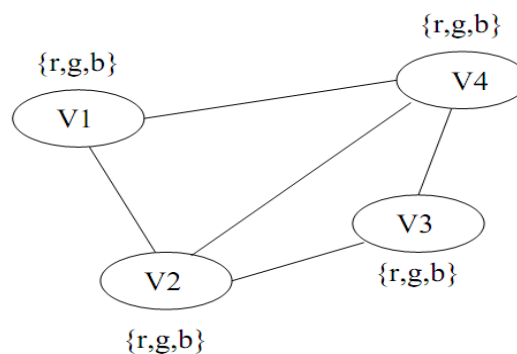
Το πρόβλημα χρωματισμού χάρτη (Color Map Problem) (Biggs et al., 1976) είναι ένα μαθηματικό πρόβλημα χρωματισμού ενός χάρτη χωρών, με ένα περιορισμένο αριθμό χρωμάτων, έτσι ώστε γειτονικές χώρες να μην έχουν το ίδιο χρώμα. Το πρόβλημα του χρωματισμού του χάρτη μπορεί να αναπαρασταθεί ως πρόβλημα ικανοποίησης περιορισμών. Πιο συγκεκριμένα, μεταβλητές του προβλήματος είναι η κάθε περιοχή του χάρτη και πεδία τιμών των μεταβλητών αυτών είναι τα διαθέσιμα χρώματα, ενώ αντίστοιχα οι περιορισμοί που θέτονται είναι δύο γειτονικές «μεταβλητές» (περιοχές στον χάρτη) να μην έχουν ταυτόχρονα την ίδια τιμή (χρώμα). Για παράδειγμα, στην Εικόνα 6.9 παρουσιάζεται ένα παράδειγμα του χρωματισμού του χάρτη ως πρόβλημα ικανοποίησης περιορισμών.



Εικόνα 6.9 Παράδειγμα Χρωματισμού Χάρτη

Ο χρωματισμός του χάρτη, ως πρόβλημα ικανοποίησης περιορισμών αναπαρίσταται ως εξής:

- Ένα σύνολο από μεταβλητές $V = \{V_1, V_2, V_3, V_4\}$ όπου κάθε μεταβλητή αναπαριστά μια περιοχή στον χάρτη.
- Ένα σύνολο τιμών του πεδίου D : κάθε πεδίο περιλαμβάνει και τις αντίστοιχες μεταβλητές που μπορεί να πάρει. Επομένως, έχουμε $D = \{D_1, D_2, D_3, D_4\}$ και ισχύει $D_1 = D_2 = D_3 = D_4 = \{\text{red, green, blue}\}$.
- Ένα σύνολο από περιορισμούς C : Κάθε περιορισμός αναπαριστάται ως μια σχέση μεταξύ των μεταβλητών του προβλήματος. Όπου, $C = \{C_{12}, C_{14}, C_{23}, C_{24}, C_{34}\}$, όπου $C_{ij} \equiv V_i \neq V_j$.



Εικόνα 6.10 Γραφος Περιορισμών

Επίσης, στην Εικόνα 6.10 παρουσιάζεται ο αντίστοιχος γράφος περιορισμών του παραπάνω προβλήματος του χρωματισμού του χάρτη. Ο γράφος περιορισμών περιέχει έναν κόμβο για κάθε μεταβλητή και μία ακμή ανάμεσα σε κάθε ζεύγος κόμβων για την οποία υπάρχει ένας περιορισμός μεταξύ των αντίστοιχων δύο μεταβλητών και κάθε ακμή αντιπροσωπεύει δύο γειτονικές περιοχές στο χάρτη. Για την επίλυση του προβλήματος γίνεται συνδυασμός του αλγορίθμου AC-3 και της αναζήτησης με *οπισθοδρόμηση* (*backtracking*) (Russel & Norvig, 2003).

6.5.2 Εκπαιδευτικές δραστηριότητες

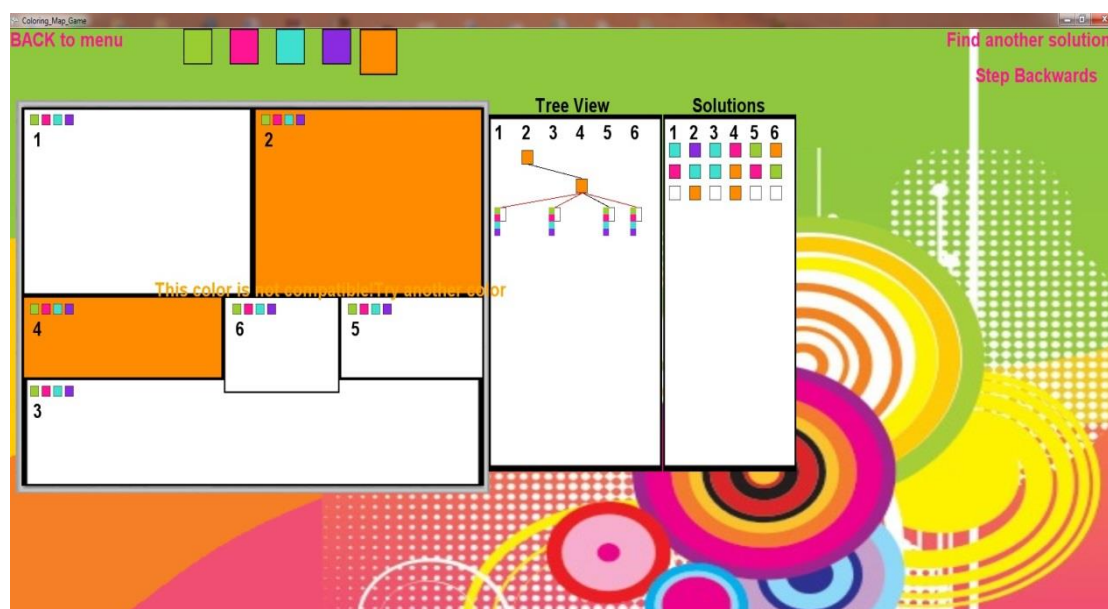
Για να βοηθήσουμε τους εκπαιδευόμενους να κατανοήσουν με φυσικό και πρακτικό την ικανοποίηση περιορισμών σχεδιάστηκε και αναπτύχθηκε το παιχνίδι Coloring Map Game το οποίο είναι ένα εκπαιδευτικό παιχνίδι που βασίζεται στο κλασσικό πρόβλημα του χρωματισμού του χάρτη. Ο κύριος στόχος είναι να προσφέρει στον εκπαιδευόμενο ένα ευχάριστο τρόπο όπου να εξοικειωθεί με τους βασικούς κανόνες που χαρακτηρίζουν τους αλγόριθμους ικανοποίησης περιορισμών, να κατανοήσει τις αρχές που διέπουν τα προβλήματα αυτά καθώς και να τον βοηθήσει στην εκμάθηση και την εφαρμογή των αλγορίθμων.

Για την δημιουργία του παιχνιδιού επιλέχθηκε το πρόβλημα του χρωματισμού του χάρτη κυρίως για δυο λόγους. Αρχικά, αποτελεί ένα κλασσικό παράδειγμα προβλήματος ικανοποίησης περιορισμών και είναι κατάλληλο για την αναπαράσταση του προβλήματος και την επίλυση μέσω αλγορίθμων ικανοποίησης περιορισμών. Επίσης, το πρόβλημα είναι κατάλληλο για την γραφική αναπαράσταση κάθε βήματος της αλγοριθμικής διαδικασίας με φυσικό και πρακτικό τρόπο. Επίσης, το Coloring Map Game στοχεύει να βοηθήσει τους εκπαιδευόμενους να αντιληφτούν και να κατανοήσουν καλύτερα μέσω του παιχνιδιού την έννοια της διάδοσης περιορισμών κατά την διάρκεια της επίλυσης προβλημάτων και εκπαιδευτικών δοκιμασιών. Στα πλαίσια του παιχνιδιού, οι παίκτες μπορούν να επιλέξουν τον αριθμό των περιοχών και τον αριθμό των χρωμάτων που είναι επαρκή για την επίλυση ή να επιλέξουν ένα από τα τρία διαθέσιμα επίπεδα δυσκολία (easy, medium, difficult) που καθορίζουν την πολυπλοκότητα του προβλήματος. Ο προσδιορισμός του επιπέδου δυσκολίας κάθε χάρτη βασίζεται στον αριθμό των χρωμάτων που χρειάζεται ο χάρτης για να επιλυθεί και τον μέγιστο αριθμό των περιοχών που συνορεύει κάθε περιοχή.

Αρχικά παρουσιάζεται το πρόβλημα του χάρτη και οι αντίστοιχες περιοχές που θα πρέπει να χρωματιστούν από τον παίκτη. Ο παίκτης ξεκινάει το παιχνίδι επιλέγοντας ένα χρώμα και την αντίστοιχη περιοχή που θέλει να χρωματίσει. Οι περιορισμοί που ισχύουν σύμφωνα με το πρόβλημα είναι ότι δυο γειτονικές περιοχές δεν πρέπει να έχουν χρωματιστεί με το ίδιο χρώμα, δεδομένου ότι κάθε περιοχή αναπαριστά μια μεταβλητή του προβλήματος. Επίσης, το παιχνίδι κατά την διάρκεια της αλληλεπίδρασης ενημερώνει και καθοδηγεί τον παίκτη με κατάλληλα μηνύματα για τις επιλογές και ο παίκτης μπορεί να ζητήσει οποιαδήποτε στιγμή βοήθεια πατώντας την κατάλληλη επιλογή.

Ο παίκτης μπορεί να ενημερωθεί για συγκεκριμένους περιορισμούς του προβλήματος και για τις αλλαγές που συμβαίνουν καθώς επιλέγει χρώματα για τις περιοχές και να προσομοιώνει με αυτό τον τρόπο τα βήματα του αλγορίθμου AC-3. Επίσης, για κάθε επιλογή του παίκτη,

δημιουργείται κατάλληλη δεντρική δομή που αναπαριστά τις περιοχές, τα βήματα χρωματισμού που έχουν πραγματοποιηθεί καθώς και τον τρόπο διάδοσης των περιορισμών. Επιπρόσθετα παρέχεται η δυνατότητα στο παίκτη να αναιρέσει κάποιον χρωματισμό περιοχής δηλαδή να κάνει βήματα προς τα πίσω και να μεταβεί σε προηγούμενες καταστάσεις της διαδρομής της λύσης. Η διαδικασία αυτή είναι γνωστή στους αλγόριθμους αναζήτησης ως οπισθοδρόμηση (Backtracking). Επομένως, ο παίκτης μέσω του παιχνιδιού μαθαίνει να εφαρμόζει παίζοντας τον τρόπο διάδοσης των περιορισμών καθώς και της έννοιας της οπισθοδρόμηση οι οποίες αποτελούν βασικά χαρακτηριστικά των προβλημάτων και των αλγορίθμων ικανοποίησης περιορισμών. Το παιχνίδι προσφέρει την δυνατότητα στον χρήστη να επιλύσει τον ίδιο πρόβλημα-χάρτη πολλές φορές και να βρει διαφορετικές λύσεις, μαθαίνοντας καλύτερα τον τρόπο λειτουργίας των αλγορίθμων ικανοποίησης περιορισμών. Στην εικόνα 6.11 παρουσιάζονται οι επιλογές των χρωμάτων του χρήστη καθώς και η αντίστοιχη δεντρική.



Εικόνα 6.11 Το παιχνίδι Coloring Map Game

Μια δεύτερη λειτουργία του παιχνιδιού έχει σχεδιαστεί για να παρέχει μεγαλύτερες προκλήσεις στους παίκτες και μέσω του ανταγωνισμού να κεντρίσει περισσότερο το ενδιαφέρον τους. Σε αυτήν την λειτουργία ο παίκτης καλείται να λύσει προκαθορισμένες πίστες (λαβύρινθους) κάτω από διαφορετικές συνθήκες (π.χ. μέσα σε συγκεκριμένο χρόνο, με λιγότερα/περισσότερα χρώματα κ.α.), κάτι που δημιουργεί συνεχείς προκλήσεις και παρέχει κίνητρα στον παίκτη για επίτευξη υψηλότερη βαθμολογία και καλύτερες επιδόσεις. Το παιχνίδι υπολογίζει το σκορ του παίκτη, όπου καθορίζεται από το επίπεδο δυσκολίας του προβλήματος, τις ενέργειες του καθώς και τον χρόνο που απαιτήθηκε για την επίλυσή του. Ο σκοπός του παιχνιδιού είναι να αξιολογηθεί η κατανόηση τόσο σε απλά όσο και πιο σύνθετα

προβλήματα και ταυτόχρονα να αποτελέσει κίνητρο μεταξύ των παικτών για το καλύτερο σκορ.

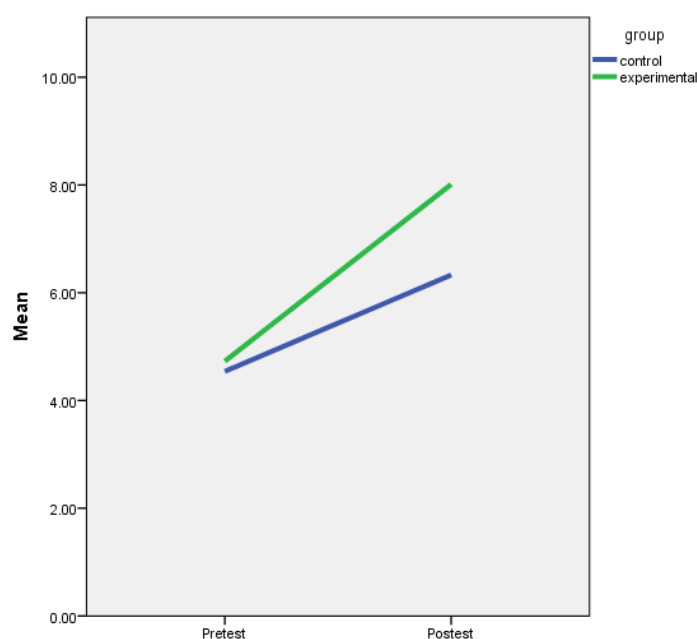
Το παιχνίδι του χρωματισμού του χάρτη μπορεί να χρησιμοποιηθεί αποτελεσματικά σε ένα μάθημα Τεχνητής Νοημοσύνης για τη διδασκαλία της επίλυσης των προβλημάτων ικανοποίησης περιορισμών και των αντίστοιχων αλγορίθμων κάνοντας την διαδικασία της μάθησης πιο ευχάριστη και ελκυστική. Πιο συγκεκριμένα, οι εκπαιδευόμενοι παίζοντας με το παιχνίδι, μπορούν να εξοικειωθούν με έννοιες όπως η διάδοση περιορισμών, η οπισθοδρόμηση, οι πολλαπλές διαδρομές επίλυσης κ. α. και να δουν μέσω του παιχνιδιού με φυσικό και πρακτικό τον τρόπο εφαρμογής τους πάνω στο χάρτη. Επίσης, παίζοντας και χρωματίζοντας περιοχές μιμούνται τα βήματα αλγορίθμων ικανοποίησης περιορισμών και της οπισθοδρόμησης (backtracking) με την αναίρεση προηγούμενων χρωματισμών δηλαδή βημάτων τους.

6.5.3 Πειραματικά Αποτελέσματα

Για να αξιολογήσουμε το εκπαιδευτικό παιχνίδι Coloring Map Game διεξήγαμε μια πειραματική μελέτη στην οποία συμμετείχαν προπτυχιακοί φοιτητές του τμήματος. Στην πειραματική μελέτη, συμμετείχαν 100 προπτυχιακοί φοιτητές (άντρες και γυναίκες) που συμμετείχαν στο μάθημα της τεχνητής νοημοσύνης. Όλοι οι φοιτητές χωρίστηκαν τυχαία σε δύο ομάδες των 50 φοιτητών, την πειραματική ομάδα (experimental group - EG) και την ομάδα ελέγχου (control group - CG) αντίστοιχα. Η πειραματική ομάδα χρησιμοποίησε το εκπαιδευτικό παιχνίδι, ενώ η ομάδα ελέγχου δεν έπαιξε με το παιχνίδι αλλά μελέτησε το εκπαιδευτικό υλικό του μαθήματος.

Σκοπός της πειραματικής μελέτης είναι να αξιολογηθούν οι εκπαιδευτικές δυνατότητες του παιχνιδιού και να αποτιμηθεί το ενδιαφέρον και η συμμετοχή των εκπαιδευόμενων. Αρχικά, ένα pretest αναπτύχθηκε για να αξιολογηθεί η προηγούμενη γνώση των εκπαιδευόμενων σχετικά με τις έννοιες της ικανοποίησης περιορισμών και των αλγορίθμων. Πιο συγκεκριμένα, το pretest περιείχε 15 ερωτήσεις και η διάρκεια του τεστ ήταν 50 λεπτά. Όλες οι επιδόσεις των εκπαιδευόμενων στο pretest βαθμολογήθηκαν από τον διδάσκοντα στην κλίμακα από 0 έως 10 μονάδες. Κατά τη διάρκεια της μαθησιακής δραστηριότητας, οι εκπαιδευόμενοι της πειραματικής ομάδας συμμετείχαν στις δραστηριότητες μάθησης μέσω του παιχνιδιού, ενώ η ομάδα ελέγχου χρησιμοποίησε τον παραδοσιακό τρόπο. Έπειτα, οι φοιτητές συμμετείχαν σε ένα posttest που περιείχε 15 ασκήσεις και είχε το ίδιο επίπεδο δυσκολίας με το pretest.

Ο στόχος της μελέτης ήταν να εξετάσει την αποτελεσματικότητα του παιχνιδιού όσον αφορά τη βελτίωση της εκμάθησης και των επιδόσεων των εκπαιδευόμενων. Για να αναλύσουμε τις επιδόσεις των φοιτητών πραγματοποιήθηκε μια t-test ανάλυση στις επιδόσεις τους στο pretest. Τα αποτελέσματα στο pretest ήταν ($M = 4.54$, $SD = 0.80$) για την ομάδα ελέγχου και ($M = 4.73$, $SD = 0.83$), για την πειραματική ομάδα αντίστοιχα. Επίσης, η τιμή p (επίπεδο σημαντικότητας) ήταν $p = 0.25$ ($p > 0.05$) και το effect sized= 0.23. Επομένως φαίνεται ότι οι δύο ομάδες είχαν ισοδύναμο επίπεδο γνώσεων σχετικά με τις έννοιες των αλγορίθμων ικανοποίησης περιορισμών πριν από την έναρξη της διαδικασίας της μάθησης. Στην εικόνα 6.12 παρουσιάζεται η μέση τιμή για το pretest και posttest αντίστοιχα και για τις δυο ομάδες και φαίνεται ότι η μέση τιμή στο pretest είναι πολύ μεγαλύτερη στην πειραματική ομάδα. Στη συνέχεια, πραγματοποιήσαμε μια ανάλυση συνδιακύμανσης (ANCOVA) για να ελέγξουμε τη διαφορά μεταξύ των δύο ομάδων χρησιμοποιώντας τα αποτελέσματα του pretest ως συμμεταβλητή και τις βαθμολογίες posttest ως εξαρτημένες μεταβλητές. Στον Πίνακα 6.3 παρουσιάζονται τα πειραματικά αποτελέσματα της ανάλυσης ANCOVA, στην οποία οι προσαρμοσμένες μέσες τιμές των βαθμολογιών posttest ήταν 8.01 για την πειραματική ομάδα και 6.33 για την ομάδα ελέγχου. Επιπλέον, φαίνεται ότι υπάρχει σημαντική διαφορά μεταξύ των δύο ομάδων ($F = 116.1$ και $p = 0.00 < 0.05$). Από τα αποτελέσματα προκύπτει ότι οι επιδόσεις των φοιτητών της πειραματικής ομάδας ήταν καλύτερες από εκείνες της ομάδας ελέγχου κάτι που σημαίνει ότι το εκπαιδευτικό παιχνίδι συνέβαλε θετικά στην κατασκευή της γνώσης και στην κατανόηση των διαδικασιών από τους εκπαιδευομένους.



Εικόνα 6.12 Αποτελέσματα Pretest και Posttest

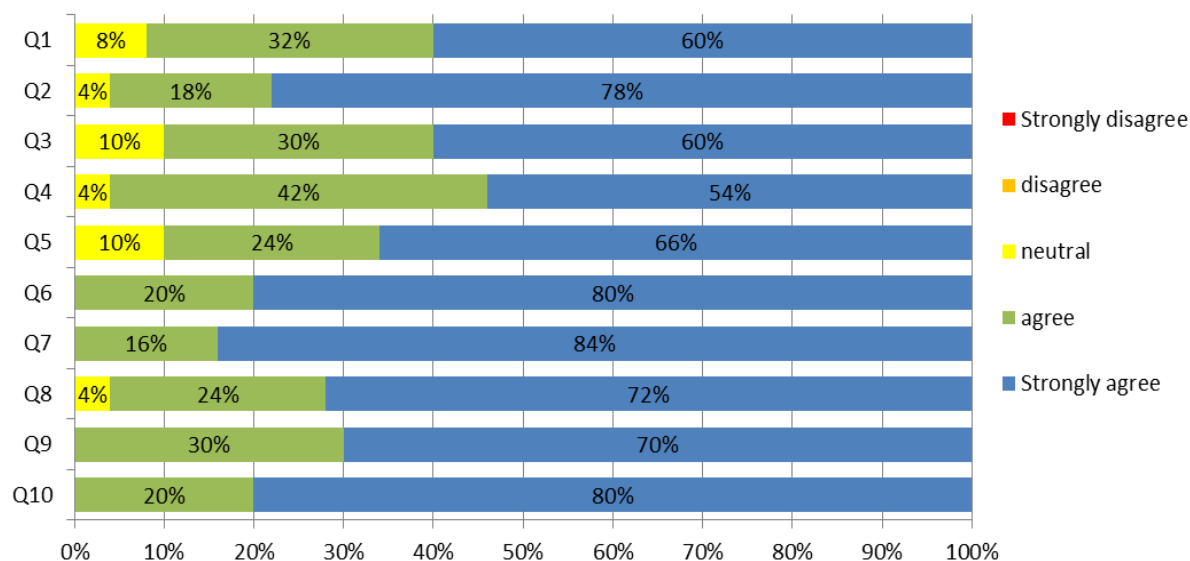
Πίνακας 6.3 Πειραματικά Αποτελέσματα Posttest

Group	N	Mean	S.D.	Adjusted mean
Experimental group	50	8.01	0.93	7.98
Control group	50	6.33	0.62	6.36

Οι φοιτητές της πειραματικής ομάδας (experimental group) που χρησιμοποίησαν το παιχνίδι, κλήθηκαν να συμπληρώσουν ένα ερωτηματολόγιο για την αξιολόγηση της ευχρηστίας του παιχνιδιού, του βαθμού που τους βοήθησε στην κατανόηση της ικανοποίησης περιορισμών και των αλγορίθμων καθώς και να εκφράσουν την γνώμη τους και τις εμπειρίες τους. Το ερωτηματολόγιο περιελάμβανε 14 ερωτήσεις. Οι ερωτήσεις Q1-Q10 βασίστηκαν στην κλίμακα Likert (1: καθόλου, 5: πάρα πολύ) και οι ερωτήσεις Q11-Qz14 ήταν ανοικτού τύπου και καλούσαν τους φοιτητές να εκφράσουν την γνώμη τους, να τονίσουν θετικά και αρνητικά σημεία του παιχνιδιού, προβλήματα που πιθανώς αντιμετώπισαν καθώς και να προτείνουν βελτιώσεις που θα ήθελαν να γίνουν στο παιχνίδι. Στον πίνακα 6.4 παρουσιάζονται οι ερωτήσεις του περιελάμβανε το ερωτηματολόγιο και στην εικόνα 6.13 τα αποτελέσματα της ανάλυσης των απαντήσεων των εκπαιδευόμενων.

Πίνακας 6.4 Ερωτήσεις ερωτηματολογίου.

ΕΡΩΤΗΣΕΙΣ
Q1 Η ενασχόληση με το παιχνίδι ήταν διασκεδαστική
Q2 Η διεπαφή του παιχνιδιού είναι εύκολη στη χρήση.
Q3 Η ενασχόληση με το παιχνίδι με έκανε πιο ενεργό στο μάθημα
Q4 Το παιχνίδι μου έδωσε κίνητρα για μελέτη
Q5 Το παιχνίδι ενίσχυσε την συμμετοχή μου στο μάθημα.
Q6 Η ενασχόληση με το παιχνίδι με αύξησε το ενδιαφέρον μου για μελέτη.
Q7 Η ενασχόληση με το παιχνίδι είναι πιο διασκεδαστική από άλλους τρόπους μάθησης
Q8 Το παιχνίδι βοήθησε στην κατανόηση της έννοια της ικανοποίησης περιορισμών.
Q9 Το παιχνίδι βοήθησε να κατανοήσω τους αλγόριθμους.
Q10 Θα συνιστούσατε το παιχνίδι στους συμμαθητές σας και θα θέλατε να ενσωματωθεί στο μάθημα;



Εικόνα 6.13 Αποτελέσματα Ερωτηματολογίου

Μετά την ανάλυση των απαντήσεων των φοιτητών στο ερωτηματολόγιο, υπολογίστηκε η μετρική Cronbach για να προσδιοριστεί η αξιοπιστία του ερωτηματολογίου. Ο συντελεστής αξιοπιστίας είναι $\alpha=0.971$, είναι μεγαλύτερος από 0.8 κάτι που σημαίνει ότι το ερωτηματολόγιο έχει υψηλή αξιοπιστία. Τα αποτελέσματα της ανάλυσης των απαντήσεων του ερωτηματολογίου δείχνουν ότι οι φοιτητές ήταν πολύ θετικοί και δείχνουν ότι η πλειοψηφία των φοιτητών (92%) βρήκαν σε μεγάλο βαθμό την εκμάθηση μέσω του παιχνιδιού να είναι διασκεδαστική και επίσης ένα μεγάλο ποσοστό (96%) βρήκε το παιχνίδι εύκολο στη χρήση. Επιπλέον, ένα μεγάλο μέρος των φοιτητών δήλωσε ότι το παιχνίδι τους έδωσε επιπλέον κίνητρα για μελέτη (90%), αύξησε την μελέτη τους (96%) και το ενδιαφέρον τους και συνέβαλε στο να γίνουν πιο ενεργοί στο μάθημα (90%). Επιπρόσθετα, όσον αφορά την αποτελεσματικότητα στην μάθηση μέσω του παιχνιδιού, οι εκπαιδευόμενοι δήλωσαν ότι το παιχνίδι τους βοήθησε να μάθουν πιο αποτελεσματικά και να κατανοήσουν την εφαρμογή των αλγορίθμων (96%). Γενικά, τα πειραματικά αποτελέσματα έδειξαν ότι το παιχνίδι βοήθησε τους εκπαιδευόμενους να κατανοήσουν εις βάθος τους αλγορίθμους αναζήτησης και τον τρόπο εφαρμογής τους με έναν ευχάριστο και διασκεδαστικό τρόπο. Τέλος, όλοι οι εκπαιδευόμενοι πρότειναν να ενσωματωθεί το παιχνίδι στο μάθημα και να χρησιμοποιείται από τους φοιτητές.

6.6 Παιχνιδοποίηση σε Εικονικό Περιβάλλον Μάθησης

Η τεχνολογική πρόοδος έχει αναδείξει νέες τεχνολογίες και μέσα πού μπορούν να χρησιμοποιηθούν στην εκπαιδευτική διαδικασία. Ένα παράδειγμα αποτελεί η εικονική πραγματικότητα που προσφέρει ποικίλες και καινοτόμες δυνατότητες σχεδιασμού εκπαιδευτικών δραστηριοτήτων. Η ανάπτυξη εκπαιδευτικών περιβαλλόντων τα οποία

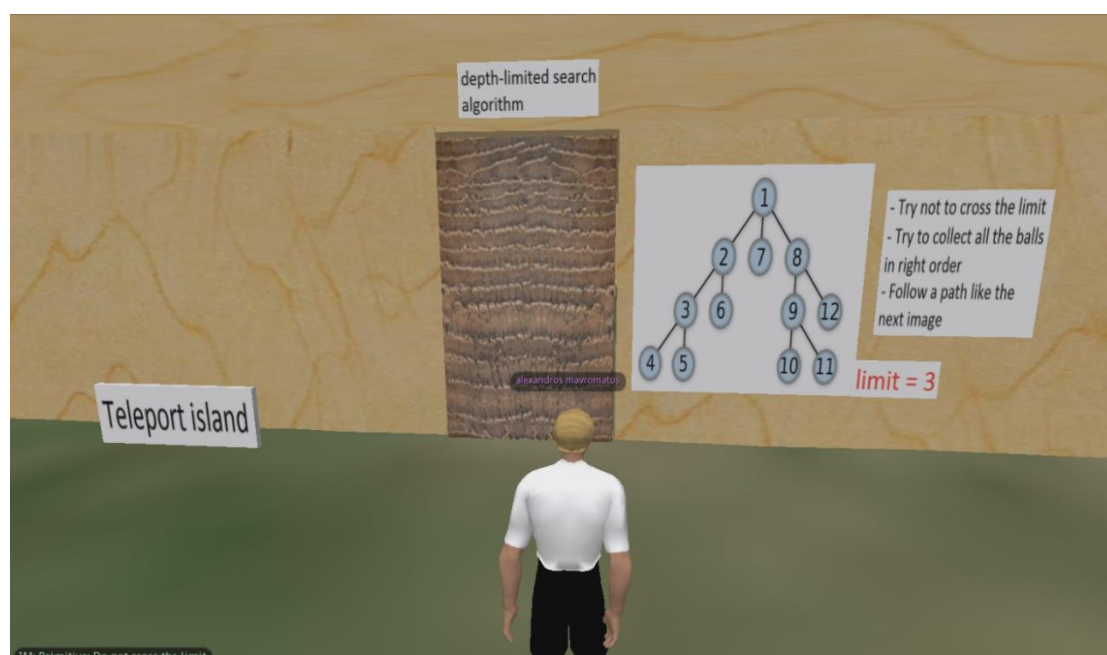
βασίζονται στην εικονική πραγματικότητα ανοίγουν νέους ορίζοντες στην εκπαίδευση κυρίως λόγω των εγγενών χαρακτηριστικών και των τρισδιάστατων (3D) προσομοιώσεων που είναι σε θέση να υποστηρίζουν και να προσφέρουν στους εκπαιδευόμενους (Allison et al., 2012).

Οι τρισδιάστατοι εικονικοί κόσμοι (3D Virtual Worlds) αποτελούν μια από τις δημοφιλέστερες κατηγορίες περιβαλλόντων εικονικής πραγματικότητας και μπορούν να συμβάλουν αποτελεσματικά στην βελτίωση της μάθησης και της κατανόησης των εκπαιδευόμενων (Gregory et al., 2016). Ένα κύριο χαρακτηριστικό των εικονικών κόσμων είναι ότι μπορούν να προσομοιώσουν και να αναπαραστήσουν καταστάσεις και διαδικασίες του φυσικού κόσμου καθώς και να προσφέρουν καινοτόμες εμπειρίες μάθησης στους εκπαιδευόμενους με σκοπό να μεγιστοποιήσουν τα μαθησιακά αποτελέσματα (Chesney et al., 2009; Stefan, 2016). Οι τρισδιάστατοι εικονικοί κόσμοι επιτρέπουν πολυδιάστατη αλληλεπίδραση των εκπαιδευόμενων με εκπαιδευτικές δραστηριότητες και δοκιμασίες και μπορούν να οδηγήσουν σε εποικοδομητική κατασκευή της γνώσης (DeFreitas et al., 2010). Στην διεθνή βιβλιογραφία έχουν χρησιμοποιηθεί με μεγάλη επιτυχία σε διάφορες εκπαιδευτικές περιοχές (Mikropoulos & Natsis, 2011) με στόχο να προσφέρουν στους μαθητές την δυνατότητα να εξερευνήσουν και να κατανοήσουν πώς πραγματοποιούνται και λειτουργούν διαδικασίες και κατασκευές μέσα από κατάλληλες προσομοιώσεις και οπτικοποιήσεις (Ku & Mahabaleshwar, 2011). Οι εκπαιδευτικές δυνατότητες των εικονικών κόσμων εξαρτώνται σε μεγάλο βαθμό από τα χαρακτηριστικά των εκπαιδευτικών δραστηριοτήτων και των σεναρίων που προσφέρουν στους εκπαιδευόμενους (Girvan & Savage, 2010). Επίσης, είναι βασικό σε ένα εκπαιδευτικό περιβάλλον οι εκπαιδευτικές δραστηριότητες να σχεδιάζονται με στόχο να ελκύουν το ενδιαφέρον των εκπαιδευόμενων, να τους παρέχουν κίνητρα μελέτης και να τους εμπλέκουν ενεργά σε δοκιμασίες με στόχο να μεγιστοποιήσουν την αποτελεσματικότητα των εκπαιδευτικών διαδικασιών.

Για την εκμάθηση των αλγορίθμων αναζήτησης σχεδιάστηκε και αναπτύχθηκε ένα τρισδιάστατο εκπαιδευτικό περιβάλλον το VR-ALG με στόχο να δώσει την δυνατότητα στους εκπαιδευόμενους να μελετήσουν και να εφαρμόσουν αλγορίθμους αναζήτησης σε διάφορες εκπαιδευτικές δραστηριότητες και δοκιμασίες. Αρχικά, ο εκπαιδευόμενος εγγράφεται στο περιβάλλον. Κατά τη διάρκεια της εγγραφής, ο εκπαιδευόμενος καλείται να καταχωρήσει τις απαιτούμενες προσωπικές πληροφορίες όπως το όνομα του, το φύλο, το έτος σπουδών κ. α.. Μετά την εγγραφή, μπορεί να συνδεθεί οποιαδήποτε στιγμή στον κόσμο με τον προσωπικό του λογαριασμό. Στον εικονικό κόσμο, όλοι οι χρήστες αναπαριστώνται με τα avatars και μπορούν να κινηθούν ελεύθερα στον κόσμο, να αλληλεπιδράσουν με κατασκευές και να συμμετάσχουν σε εκπαιδευτικές δραστηριότητες.

Ο VR-ALG εικονικός κόσμος σχεδιάστηκε και αναπτύχθηκε με τέτοιο τρόπο ώστε να προσφέρει ένα πλήρες εκπαιδευτικό περιβάλλον στους διδάσκοντες και τους

εκπαιδευόμενους. Μέσω του εικονικού κόσμου προσφέρεται η δυνατότητα στους εκπαιδευόμενους να μελετήσουν το θεωρητικό μέρος του πεδίου των αλγορίθμων καθώς και να εξασκηθούν στην εφαρμογή των αλγορίθμων αναζήτησης σε διάφορες καταστάσεις και δοκιμασίες. Οι εκπαιδευόμενοι στα πλαίσια των εκπαιδευτικών δοκιμασιών, καλούνται να εφαρμόσουν συγκεκριμένους αλγορίθμους αναζήτησης μέσα σε τρισδιάστατους λαβύρινθους στο εικονικό περιβάλλον. Πιο συγκεκριμένα, ο παίκτης καλείται να αναζητήσει την κατάλληλη έξοδο του λαβύρινθου, κινούμενος κάθε φορά με βάση συγκεκριμένο αλγόριθμο αναζήτησης. Με αυτό τον τρόπο, ο παίκτης πρέπει να ακολουθηθεί το αντίστοιχο μονοπάτι που θα ακολουθούσε ο αλγόριθμος αναζήτησης. Συγκεκριμένα, ο χρήστης θα πρέπει ολοκληρώνοντας το παιχνίδι να έχει κατανοήσει τον τρόπο λειτουργίας του εκάστοτε αλγόριθμου και πως εφαρμόζεται σε ένα δέντρο αναζήτησης καθώς επίσης και ανάλογα με το είδος του αλγορίθμου να έχει κατανοήσει και τα ιδιαίτερα χαρακτηριστικά του όπως πχ. ευρετική συνάρτηση, τα παιδιά που αναπτύσσονται κ.ά.



Εικόνα 6.14 Παιχνιδοποίηση σε εικονικό Περιβάλλον VR-ALG

Αρχικά ο παίκτης μαθαίνει πώς λειτουργεί ο κάθε αλγόριθμος διαβάζοντας έξω από κάθε λαβύρινθο βασικές οδηγίες για το πώς θα κινηθεί μέσα στον λαβύρινθο και επίσης μέσω διαφανειών να μελετήσει το υλικό ανάλογα με το αλγόριθμο. Στη συνέχεια προσπαθεί να εφαρμόσει τα βήματα του Αλγορίθμου μέσα στο λαβύρινθο προκειμένου να βρει την κατάλληλη έξοδο του. Επίσης, αφού έχει ολοκληρώσει ένα σύνολο από διασχίσεις σε διαφορετικούς αλγόριθμους και έχει πραγματοποιήσει ένα σύνολο από ενέργειες παρέχεται

στον εκπαιδευόμενο η δυνατότητα να απαντήσει ένα προσαρμοσμένο τεστ που περιλαμβάνει τους αλγόριθμους που έχει εφαρμόσει στις αντίστοιχες δοκιμασίες.

Ο κύριος στόχος του παιχνιδιού είναι να βοηθήσει τον εκπαιδευόμενο να πειραματιστεί και να μάθει σε ένα περιβάλλον που προσομοιώνει τον πραγματικό κόσμο και με φυσικό τρόπο τις έννοιες των αλγορίθμων και να βοηθήσει στην καλύτερη κατανόηση των εννοιών. Επομένως, ο παίκτης μέσα στο τρισδιάστατο κόσμο κατανοεί τον τρόπο λειτουργίας των αλγορίθμων αναζήτησης και με έναν ευχάριστο και διασκεδαστικό τρόπο μαθαίνει παίζοντας έννοιες των αλγορίθμων αναζήτησης.

7

Συμπεράσματα και Μελλοντικές Κατευθύνσεις

Στο Κεφάλαιο αυτό παρουσιάζεται μια γενική επισκόπηση της βασικής συνεισφοράς της διατριβής και παραθέτονται τα κύρια συμπεράσματα που προέκυψαν καθώς και οι κύριες κατευθύνσεις στις οποίες θα μπορούσε να εστιάσει η μελλοντική έρευνα. Η παρούσα διατριβή συνεισφέρει στο τομέα της τεχνητής νοημοσύνης και ειδικότερα στα Ευφυή Εκπαιδευτικά Συστήματα.

Αρχικά, μια συνιστώσα της έρευνας ήταν να μελετήσουμε μηχανισμούς και μεθοδολογίες για την παραγωγή ασκήσεων σε ευφυή εκπαιδευτικά συστήματα. Για τον σκοπό αυτό σχεδιάσαμε και αναπτύξαμε μηχανισμούς για την ημιαυτόματη παραγωγή ασκήσεων πάνω στο πεδίο της λογικής και των αλγορίθμων. Οι μηχανισμοί αυτοί παράγουν ορθά δομημένες ασκήσεις, με βάση ένα σύνολο παραμέτρων που προσδιορίζονται από τον διδάσκοντα και βασίζονται σε προσεγγίσεις βασισμένες σε πρότυπα (pattern-based). Επίσης μια σημαντική παράμετρος που λαμβάνεται υπόψη για την παραγωγή των ασκήσεων είναι το επίπεδο δυσκολία της άσκησης. Η ημιαυτόματη παραγωγή ασκήσεων παρέχει την δυνατότητα στους εκπαιδευόμενους να αλληλεπιδρούν με διαφορετικές και πρωτότυπες ασκήσεις και συμβάλλει στην αποτελεσματικότερη μάθηση. Τα πειραματικά αποτελέσματα αξιολόγησης των μηχανισμών παραγωγής ασκήσεων λογικής και αλληλεπιδραστικών ασκήσεων αλγορίθμων έδειξαν υψηλή ακρίβεια και συνέπεια στις παραγόμενες ασκήσεις. Επίσης, οι ασκήσεις που δημιουργήθηκαν ανταποκρινόντουσαν πλήρως στις απαιτήσεις των διδασκόντων και ήταν κατάλληλες για να εισαχθούν στο ευφύες σύστημα διδασκαλίας.

Στην συνέχεια, σχεδιαστήκαν και αναπτύχθηκαν οι μηχανισμοί και οι μεθοδολογίες για τον αυτόματο προσδιορισμό της δυσκολίας των ασκήσεων. Οι μηχανισμοί βασίστηκαν σε σύνθετα έμπειρα συστήματα και υβριδικούς μηχανισμούς όπως νεύρο-ασαφείς (neuro-fuzzy) και νεύρο-συμβολικές (neuro-symbolic) προσεγγίσεις. Οι μηχανισμοί αυτοί εφαρμόστηκαν σε ασκήσεις μετατροπής εκφράσεων Φυσικής γλώσσας σε ΚΛΠΤ, σε ασκήσεις μετατροπής εκφράσεων ΚΛΠΤ σε ΠΜ καθώς επίσης και σε ασκήσεις πάνω στους αλγορίθμους αναζήτησης. Τα πειραματικά αποτελέσματα της αξιολόγηση των μηχανισμών και των μεθοδολογιών έδειξαν υψηλή ακρίβεια και αξιοπιστία στον προσδιορισμό της δυσκολίας των ασκήσεων.

Επίσης, μια σημαντική συνιστώσα στα εκπαιδευτικά συστήματα είναι η αυτόματη αξιολόγηση των απαντήσεων των εκπαιδευόμενων, η οποία αποτελεί μια ιδιαίτερα αναγκαία καθώς και απαιτητική διαδικασία. Η αξιολόγηση των απαντήσεων των εκπαιδευόμενων και ως εκ τούτου, ο προσδιορισμός των επιδόσεών τους είναι αναγκαία για να γίνει πιο αποτελεσματική η διαδικασία της μάθησης. Για το σκοπό αναπτύχθηκε μια γενική μεθοδολογία που βασίζεται στην ομοιότητα μεταξύ της σωστής απάντησης και της απάντησης του εκπαιδευόμενου, στην μοντελοποίηση της απάντησης του εκπαιδευόμενου, στο είδος των λαθών που έγιναν, καθώς επίσης σε ένα μηχανισμό για την αυτόματη βαθμολόγηση των απαντήσεων, που αξιολογεί τις απαντήσεις ανάλογα με το πεδίο εφαρμογής. Τα πειραματικά αποτελέσματα της εφαρμογής της μεθοδολογίας σε ασκήσεις στην λογική και σε αλληλεπιδραστικές ασκήσεις πάνω στους αλγορίθμους αναζήτησης ήταν ιδιαίτερα ικανοποιητικά. Πιο συγκεκριμένα, προκύπτει ισχυρή συμφωνία μεταξύ της αξιολόγησης των διδασκόντων και των μηχανισμών αυτόματης αξιολόγησης. Αξίζει να σημειωθεί ότι, η πειραματική μελέτη διεξήχθη σε διαφορετικού τύπου απαντήσεις σε ασκήσεις όπου οι απαντήσεις ήταν κειμενικής μορφής, όπως εκφράσεις ΚΛΠΤ, εκφράσεις σε ΠΜ καθώς επίσης και έγινε αξιολόγηση και σε απαντήσεις σε διαδραστικές ασκήσεις πάνω στους αλγορίθμους αναζήτησης.

Μια επιπλέον συνεισφορά αποτελεί η ανάπτυξη ευφών συστημάτων τα οποία μοντελοποιούν και διδάσκουν διαδικασίες της Τεχνητής Νοημοσύνης οι οποίες είναι ιδιαίτερα πολύπλοκες και σύνθετες να τις κατανοήσουν οι εκπαιδευόμενοι. Το ευφές σύστημα Διδασκαλίας AITS παρέχει εξατομικευμένη μάθηση και διδάσκει διάφορες έννοιες Τ.Ν. όπως οι αλγόριθμοι αναζήτησης και η ικανοποίηση περιορισμών. Τα ευφών συστήματα διδασκαλίας FOLtoCF και FOL-EQU είναι ευφών συστήματα διδασκαλίας που μοντελοποιούν σύνθετες έννοιες της λογικής όπως τη διαδικασία μετατροπής

Κατηγορηματικής λογικής Πρώτης Τάξης σε Προτασιακή Μορφή και η δημιουργία ισοδύναμων εκφράσεων λογικής, τα οποία μέσω των μηχανισμών που διαθέτουν, συμβάλουν στην καλύτερη κατανόηση των εννοιών από τους εκπαιδευόμενους. Τα πειραματικά αποτελέσματα από την εφαρμογή των συστημάτων σε φοιτητές έδειξαν πολύ ικανοποιητικά αποτελέσματα και η μάθηση μέσω των συστημάτων έγινε πιο αποτελεσματική, άμεση και ευχάριστη.

Επίσης, μια άλλη συνιστώσα των ΕΣΔ που αναπτύχθηκαν είναι ότι παρέχουν την δυνατότητα πειραματισμών και οπτικοποιήσεων των διαδικασιών και αποτελεσμάτων σε μη εκπαιδευτικούς μηχανισμούς για μάθηση μέσω εμπειρικής χρήσης. Πιο συγκεκριμένα, δημιουργήθηκαν δύο εκπαιδευτικές εκδοχές ενός αποδείκτη θεωρημάτων. Η πρώτη εκδοχή του αποδείκτη θεωρημάτων EX-ACTP επιτρέπει τον πειραματισμό με διάφορους συνδυασμούς στρατηγικών ελέγχου της διαδικασίας απόδειξης και αντίστοιχες οπτικοποιήσεις. Η δεύτερη εκδοχή του i-ACTP επιτρέπει την παρέμβαση του χρήστη στη διαδικασία απόδειξης με αντίστοιχη οπτικοποίηση αποτελεσμάτων.

Επίσης, μια άλλη κατεύθυνση της διατριβής είναι η ανάπτυξη και σχεδιασμός εκπαιδευτικών παιχνιδιών για την εκμάθηση εννοιών της τεχνητής νοημοσύνης όπως οι αλγόριθμοι αναζήτησης και η ικανοποίηση περιορισμών. Οι εκπαιδευόμενοι μπορούν να μάθουν παίζοντας και να κατανοούν σύνθετες έννοιες της τεχνητής νοημοσύνης με έναν ευχάριστο, διασκεδαστικό και φυσικό τρόπο. Τα πειραματικά αποτελέσματα σε φοιτητές που αλληλεπιδράσαν μέσω των δυο παιχνιδιών έδειξαν ότι τα παιχνίδια συνέβαλλε να γίνει η μάθηση πιο αποτελεσματική και ταυτόχρονα τα παιχνίδια αποτέλεσαν ένα ευχάριστο και διασκεδαστικό μέσω διδασκαλίας.

Τα αποτελέσματα της έρευνας της διατριβής θέτουν περαιτέρω κατευθύνσεις στις οποίες θα μπορούσε να κινηθεί η μελλοντική έρευνα. Στην συνέχεια, παρουσιάζονται μερικές βασικές πιθανές μελλοντικές κατευθύνσεις. Μια μελλοντική κατεύθυνση στο πεδίο της ημιαυτόματη παραγωγή ασκήσεων είναι η γενίκευση του μηχανισμού και σε άλλα συναφή πεδία καθώς επίσης και η επέκταση του μηχανισμού παραγωγής ασκήσεων ΚΛΠΤ. Πιο συγκεκριμένα, μια πιθανή κατεύθυνση για τη επέκταση του μηχανισμού παραγωγής ΚΛΠΤ είναι η δημιουργία μιας γενικής οντολογία με ειδικά χαρακτηριστικά που θα αποτυπώνει τις σχέσεις μεταξύ των κατηγορημάτων, σε διάφορα πεδία ενδιαφέροντος, με στόχο να μπορούν να συνδεθούν κατάλληλα τα κατηγορήματα, να προσδιοριστούν οι κατάλληλοι ποσοδείκτες και τα διασυνδετικά, ώστε να δημιουργηθούν εκφράσεις σημασιολογικά ορθές και παράλληλα με διαφορετικά νοήματα και απλό πολλά πεδία.

Επίσης, μια πιθανή μελλοντική κατεύθυνση, για τον αυτόματο προσδιορισμό της δυσκολίας είναι να γίνει επέκταση του μηχανισμού λαμβάνοντας υπόψη και τις επιδόσεις των εκπαιδευόμενων για τον προσδιορισμό της δυσκολίας χρησιμοποιώντας τεχνικές εξαγωγής γνώσης (data mining) για την εξαγωγή της γνώσης των εκπαιδευόμενων.

Μια σημαντική μελλοντική κατεύθυνση, στο πεδίο της αυτόματη αξιολόγηση των επιδόσεων των εκπαιδευόμενων είναι η δυνατότητα προσαρμογής του μηχανισμού αξιολόγησης με βάση τα χαρακτηριστικά και τον τρόπο αξιολόγησης του κάθε διδάσκοντα. Πιο συγκεκριμένα, να παρέχεται η δυνατότητα στον μηχανισμό να μπορεί να προσομοιώνει τον τρόπο βαθμολόγησης του εκάστοτε αξιολογητή χρησιμοποιώντας τεχνικές μηχανικής μάθησης και συλλογιστικής βασισμένης σε περιπτώσεις (case-based reasoning). Με αυτό τον τρόπο θα μπορούσαν να δημιουργηθούν στερεότυπα αξιολόγησης που να προσημειώνουν τον τρόπο αξιολόγησης των εκπαιδευτικών όπως «ελαστικός» , «αυστηρός» κ.ά.

Στο πεδίο των εκπαιδευτικών παιχνιδιών μια πιθανή μελλοντική κατεύθυνση θα ήταν να αναπτυχθούν μηχανισμοί που να παράγουν αυτόματα ανάδραση (automatic feedback generation) με βάση τις ενέργειες των εκπαιδευόμενων και ανάλογα με τις ενέργειες τους μέσα στο περιβάλλον του παιχνιδιού να παράγεται αυτόματα η κατάλληλη ανάδραση. Επίσης, μια πιθανή κατεύθυνση είναι η δημιουργία μηχανισμών για την αυτόματη αξιολόγηση των επιδόσεων μέσα στο παιχνίδι όπου με βάση το σενάριο μάθησης και τις επιλογές που έγιναν κατά την διάρκεια των εκπαιδευτικών δραστηριοτήτων στα πλαίσια των παιχνιδιών να γίνεται και ο προσδιορισμός της βαθμολογίας.

Τέλος, μια ακόμη πιθανή κατεύθυνση τόσο για τα ευφυή εκπαιδευτικά συστήματα όσο και για τα εκπαιδευτικά παιχνίδια, αφορά την αναγνώριση της συναισθηματικής κατάστασης του εκπαιδευόμενου κατά την διάρκεια της μάθησης και της ενασχόλησης με εκπαιδευτικές διαδικασίες, μέσω της αναγνώρισης και της ανάλυσης των εκφράσεων του προσώπου καθώς και της επικοινωνίας πχ. με τους άλλους εκπαιδευόμενους.

Δημοσιεύσεις

Στην ενότητα αυτή παρουσιάζονται οι δημοσιεύσεις που προέκυψαν στα πλαίσια της διδακτορικής Διατριβής.

Δημοσιεύσεις σε Διεθνή Περιοδικά

- J1.** Foteini Grivokostopoulou, Isidoros Perikos, Ioannis Hatzilygeroudis, “An Educational System for Learning Search Algorithms and Automatically Assessing Student Performance”. *International Journal of Artificial Intelligence in Education*, 1-34, 2016, Springer (Impact Factor: 1.553)
- J2.** Isidoros Perikos, Foteini Grivokostopoulou, Ioannis Hatzilygeroudis, “Assistance and Feedback Mechanism in an Intelligent Tutoring System for Teaching Conversion of Natural Language into Logic”, *International Journal of Artificial Intelligence in Education*, 2016 Springer, (Impact Factor: 1.553)
- J3.** Foteini Grivokostopoulou, Ioannis Hatzilygeroudis, Isidoros Perikos, “Teaching assistance and automatic difficulty estimation in converting first order logic to clause form”, *Artificial Intelligence Review*, Volume 42, Issues 3, pp. 347-367, 2014, Springer (Impact Factor: 1.731)
- J4.** Isidoros Perikos, Foteini Grivokostopoulou, Konstantinos Kovas, Ioannis Hatzilygeroudis, “Automatic estimation of exercises' difficulty levels in a tutoring system for teaching the conversion of natural language into first order logic”. *Expert Systems*, 33(6), 569-580. (Impact Factor: 0.947)

Δημοσιεύσεις σε Βιβλία-Τόμους:

- B1.** Foteini Grivokostopoulou, Isidoros Perikos, Ioannis Hatzilygeroudis, “Difficulty estimation of exercises on tree-based search algorithms using neuro-fuzzy and neuro-symbolic approaches”, *Advances in Combining Intelligent Methods*, Springer 2016
- B2.** Foteini Grivokostopoulou, Ioannis Hatzilygeroudis, Isidoros Perikos, “An Intelligent System for Learning First Order Logic to Clause Form Conversion”, *Multicultural Awareness and Technology in Higher Education: Global Perspectives*, chapter 13, pp.265-284, IGI Global, 2014.

Δημοσιεύσεις σε Διεθνή Συνέδρια με Πλήρη Κρίση:

- C1.** Foteini Grivokostopoulou, Isidoros Perikos, Ioannis Hatzilygeroudis, “An Educational Game for Teaching Search Algorithms”. In *Proceedings of the 8th*

- International Conference on Computer Supported Education (CSEDU),179-3, pp. 129-136, 2016, Rome, Italy.
- C2.** Foteini Grivokostopoulou, Isidoros Perikos, Ioannis Hatzilygeroudis “Estimating the Difficulty of Exercises on Search Algorithms using a Neuro-Fuzzy Approach” In Tools with Artificial Intelligence (IEEE ICTAI), 2015 IEEE 27th International Conference on, pp. 866-872, 2015, Vietri sul mare, Italy.
- C3.** Foteini Grivokostopoulou, Ioannis Hatzilygeroudis, “Semi-Automatic Generation of Interactive Exercises Related to Search Algorithms” 10th International Conference on Computer Science & Education (IEEE ICCSE), pp. 33-37, 2015, Fitzwilliam College, Cambridge University, UK.
- C4.** Foteini Grivokostopoulou, Isidoros Perikos, Ioannis Hatzilygeroudis, “Utilizing Semantic Web Technologies and Data Mining Techniques to Analyze Students Learning and Predict Final Performance”, IEEE International Conference on Teaching, Assessment and Learning for Engineering (IEEE TALE), pp. 488-494, New Zealand.
- C5.** Foteini Grivokostopoulou, Isidoros Perikos, Ioannis Hatzilygeroudis, “Using Semantic Web Technologies in a Web Based System for Personalized Learning AI Course”, IEEE International Conference on Technology for Education (IEEE T4E), pp. 257-260, India, 2014.
- C6.** Foteini Grivokostopoulou, Ioannis Hatzilygeroudis “Automated Marking for Interactive Exercises on Heuristic Search Algorithms”, Proceedings of IEEE International Conference on Teaching, Assessment and Learning for Engineering (IEEE TALE), pp. 598-603, 2013, Bali, Indonesia.
- C7.** Foteini Grivokostopoulou, Ioannis Hatzilygeroudis, “Teaching AI Search Algorithms In A Web-Based Educational System” Proceedings of the IADIS International Conference e-Learning 2013, pp. 83-90, Prague, Czech Republic, 2013
- C8.** Foteini Grivokostopoulou, Ioannis Hatzilygeroudis. “An Automatic Marking System for Interactive Exercises on Blind Search Algorithms” International conference on Artificial Intelligence in Education (AIED), pp. 783-786, Memphis, USA.
- C9.** Foteini Grivokostopoulou, Isidoros Perikos, Ioannis Hatzilygeroudis “An Automatic Marking System for FOL to CF Conversions” Proceedings of IEEE International Conference on Teaching, Assessment and Learning for Engineering 2012(IEEE TALE) pp. H1A-7- H1A-9, 2012, Hong Kong.
- C10.** Foteini Grivokostopoulou, Isidoros Perikos, Ioannis Hatzilygeroudis “Assistant tools for teaching FOL to CF Conversion” Proceedings of 8th Artificial Intelligence Applications and Innovations(AIAI) Conference Volume 381/2012, pp. 306-315, 2012, Halkidiki, Greece.
- C11.** Ioannis Hatzilygeroudis, Foteini Grivokostopoulou, Isidoros Perikos, “Using Game-based Learning in Teaching CS Algorithms” Proceedings of IEEE

International Conference on Teaching, Assessment, and Learning for Engineering (IEEE TALE) 2012, pp. H2C-9 - H2C-12, 2012, Hong Kong.

C12.Foteini Grivokostopoulou, Isidoros Perikos, Ioannis Hatzilygeroudis “A Web-based Interactive System for Learning FOL to CF conversion” Proceedings of the IADIS International Conference e-Learning 2012, pp. 287-294, 2012, Lisbon, Portugal.

C13.Ioannis Hatzilygeroudis, Foteini Grivokostopoulou, Isidoros Perikos, “Teaching Aspects of Constraints Satisfaction Algorithms via a Game” Proceedings of Third AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI), pp. 2371-2372, 2012, Toronto, Canada.

C14.Isidoros Perikos, Foteini Grivokostopoulou, Ioannis Hatzilygeroudis “Automatic Marking of NL to FOL Conversions” In Proc. of 15th IASTED International Conference on Computers and Advanced Technology in Education (CATE), pp. 227-233, Napoli, Italy

C15.Isidoros Perikos, Foteini Grivokostopoulou, Ioannis Hatzilygeroudis, Konstantinos Kovas “Difficulty Estimator for Converting Natural Language into First Order Logic.” In Intelligent Decision Technologies (KES IDT), pp. 135-144, 2012, Athens, Greece.

C16.Foteini Grivokostopoulou, Ioannis Hatzilygeroudis “A Tool for Experimenting with a Theorem Prover” The Second International Conference on Advanced Cognitive Technologies and Applications (COGNITIVE) pp. 78-83, 2010, Lisbon, Portugal.

Εργασίες σε Workshops

W1.Foteini Grivokostopoulou, Isidoros Perikos, Ioannis Hatzilygeroudis, "An Intelligent Tutoring System for Teaching FOL Equivalence", In The First Workshop on AI-supported Education for Computer Science (AIEDCS 2013), pp. 20-29, Memphis, USA.

References

- Anderson, L. W., Krathwohl, D. R., & Bloom, B. S. (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. Allyn & Bacon.
- Ala-Mutka, K. M. (2005). A survey of automated assessment approaches for programming assignments. *Computer science education*, 15(2), 83-102.
- Alemán, J. L. F. (2011). Automated assessment in a programming tools course. *Education, IEEE Transactions on*, 54(4), 576-581.
- Aleven, V., McLaren, B. M., Sewall, J., & Koedinger, K. R. (2009). A new paradigm for intelligent tutoring systems: Example-tracing tutors. *International Journal of Artificial Intelligence in Education*, 19(2), 105-154.
- Aleven, V., Roll, I. D. O., McLaren, B. M., & Koedinger, K. R. (2010). Automated, unobtrusive, action-by-action assessment of self-regulation during learning with an intelligent tutoring system. *Educational Psychologist*, 45(4), 224-233
- Alexandrou-Leonidou, V., & Philippou, G. N. (2005). Teachers' beliefs About Students Development Of The Pre-Algebraic Concept Of Equation. *International Group For The Psychology Of Mathematics Education*, 41.
- Allison, C., Campbell, A., Davies, C. J., Dow, L., Kennedy, S., McCaffery, J. P. ., & Perera, G. I. U. S. (2012). Growing the use of Virtual Worlds in education: an OpenSim perspective. *Proceedings of the 2nd European Immersive Education Summit*.
- Alonso, J. A., Aranda, G. A., Martín. -Mateos, F. J. : FITS: Formalization with an Intelligent Tutor System. In: IV International Conference On Multimedia And Information And Communication Technologies In Education (2006)
- Alonso, J. A., Aranda, G. A., Martin-Matceos, F. J. : KRRT: Knowledge Representation and Reasoning Tutor. (2007). In: Moreno Diaz, R., Pichler, F., Quesada Arencibia, A. (eds.) EUROCAST 2007. LNCS, vol. 4739, pp. 400–407. Springer, Heidelberg
- Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The journal of the learning sciences*, 4(2), 167-207.
- Arroyo, I., Woolf, B. P., Burelson, W., Muldner, K., Rai, D., & Tai, M. (2014). A multimedia adaptive tutoring system for mathematics that addresses cognition, metacognition and affect. *International Journal of Artificial Intelligence in Education*, 24(4), 387-426.
- Asperti, A., Coen, C. S., Tassi, E., & Zacchiroli, S. (2007). User interaction with the Matita proof assistant. *Journal of Automated Reasoning*, 39(2), 109-139.
- Baecker, R. (1998). Sorting out sorting: A case study of software visualization for teaching computer science. *Software visualization: Programming as a multimedia experience*, 1, 369-381.

- Barendregt, H., & Geuvers, H. (2001). Proof-Assistants Using Dependent Type Systems. *Handbook of automated reasoning*, 2, 1149-1238.
- Barker-Plummer, D., Cox, R., Dale, R. (2009) Dimensions of Difficulty in Translating Natural Language into First-Order Logic. Educational Data Mining 2009: 2nd International conference on Educational Data Mining, Cordoba, Spain 220-228.
- Barker-Plummer, D., Cox, R., Dale, R., Etchemendy, J. (2008): An Empirical Study of Errors in Translating Natural Language into Logic. In: Proceedings of the 30th Annual Conference of the Cognitive Science Society, Washington, DC 505-510
- Barker-Plummer, D., Dale, R., & Cox, R. (2012). Using edit distance to analyse errors in a natural language to logic translation corpus. In Proceedings of the 5th international conference on educational data mining (pp. 134–141).
- Barwise, J., & Etchemendy J. (2002). Language, Proof and Logic, Center for the Study of Language and Information, 1st edn.
- Bayliss, J. D. (2007). The effects of games in CS1-3. In Microsoft Academic Days Conference on Game Development in Computer Science Education (pp. 59-63)
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific american*, 284(5), 28-37.
- Berramla, K., Deba, E. A., & Senouci, M. (2015, November). Formal validation of model transformation with Coq proof assistant. In *New Technologies of Information and Communication (NTIC), 2015 First International Conference on* (pp. 1-6). IEEE.
- Bertot, Y., & Castéran, P. (2013). *Interactive theorem proving and program development: Coq'Art: the calculus of inductive constructions*. Springer Science & Business Media.
- Biggs, N., Lloyd, E. K., & Wilson, R. J. (1976). Graph Theory, 1736-1936. Oxford University Press
- Billingsley, W., & Robinson, P. (2007). Student proof exercises using MathsTiles and Isabelle/HOL in an intelligent book. *Journal of Automated Reasoning*, 39(2), 181-218.
- Bloom, B. S. (1956). Taxonomy of educational objectives: The classification of educational goals.
- Blumenstein, M., Green, S., Fogelman, S., Nguyen, A., & Muthukkumarasamy, V. (2008). Performance analysis of GAME: A generic automated marking environment. *Computers & Education*, 50(4), 1203-1216.
- Boda, K., Ma, J., Sinnadurai, G., & Summers, A. (2007). Pandora: A reasoning Toolbox using natural Deduction Style. *Logic journal of the IGPL*, 15(4), 293-304.
- Brusilovsky, P. (1998, August). Adaptive educational systems on the world-wide-web: A review of available technologies. In Proceedings of Workshop "WWW-Based Tutoring" at 4th International Conference on Intelligent Tutoring Systems (ITS'98), San Antonio, TX.
- Brusilovsky, P., & Sosnovsky, S. (2005). Individualized exercises for self-assessment of programming knowledge: An evaluation of QuizPACK. *Journal on Educational Resources in Computing (JERIC)*, 5(3), 6.
- Burgos, D., Tattersall, C., & Koper, R. (2007). Re-purposing existing generic games and simulations for e-learning. *Computers in Human Behavior*, 23(6), 2656-2667.
- Burguillo, J. C. (2010). Using game theory and competition-based learning to stimulate student motivation and performance. *Computers & Education*, 55(2), 566-575.

- Cablé, B., Guin, N., & Lefevre, M. (2013, July). An authoring tool for semi-automatic generation of self-assessment exercises. In *Artificial Intelligence in Education* (pp. 679-682). Springer Berlin Heidelberg.
- Cagiltay, N. E. (2007). Teaching software engineering by means of computer-game development: Challenges and opportunities. *British Journal of Educational Technology*, 38(3), 405-415.
- Chang, W. C., Chiu, Y. D., & Li, M. F. (2008). Learning Kruskal's Algorithm, Prim's Algorithm and Dijkstra's Algorithm by board game. In *Advances in Web Based Learning-ICWL 2008* (pp. 275-284). Springer Berlin Heidelberg.
- Cheng, I., Shen, R., & Basu, (2008) A.. An algorithm for automatic difficulty level estimation of multimedia mathematical test items. In *Proceedings of the 8th IEEE International Conference on Advanced Learning Technologies*. Los Alamitos, CA: IEEE Computer Society 175–179.
- Chesney, T., Chuah, S. H., & Hoffmann, R. (2009). Virtual world experimentation: An exploratory study. *Journal of Economic Behavior & Organization*, 72(1), 618-635.
- Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive science*, 13(2), 145-182.
- Chronopoulos, T., Perikos, I., & Hatzilygeroudis, I. (2010, October). An example-tracing tutor for teaching NL to FOL conversion. In *IFIP International Conference on Artificial Intelligence Applications and Innovations* (pp. 170-178). Springer Berlin Heidelberg.
- Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review*, 24(2), 205-249.
- Clow, D. (2012, April). The learning analytics cycle: closing the loop effectively. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 134-138). ACM.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. 2nd edn. Hillsdale, New Jersey: L
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1), 37-46.
- Conati, C. (2009, July). Intelligent Tutoring Systems: New Challenges and directions. In *IJCAI* (Vol. 9, pp. 2-7).
- Conejo, R., Guzmán, E., Perez-De-La-Cruz, J. L., & Barros, B. (2014). An empirical study on the quantitative notion of task difficulty. *Expert Systems with Applications*, 41(2), 594-606.
- Corbett, A. T., & Anderson, J. R. (2001). Locus of feedback control in computer-based tutoring: Impact on learning rate, achievement and attitudes. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 245-252). ACM.
- Craig, S. D., Hu, X., Graesser, A. C., Bargagliotti, A. E., Sterbinsky, A., Cheney, K. R., & Okwumabua, T. (2013). The impact of a technology-based mathematics after-school program using ALEKS on student's knowledge and behaviors. *Computers & Education*, 68, 495-504
- Cristea, P., & Tuduce, R. (2005). Automatic generation of exercises for selftesting in adaptive e-learning systems. In *Exercises on AC Circuits', Third International Workshop on Authoring of Adaptive and Adaptable Educational Hypermedia*.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16 (3), 297 – 334.

- D'Mello, S., Lehman, B., Sullins, J., Daigle, R., Combs, R., Vogt, K., ... & Graesser, A. (2010, June). A time for emoting: When affect-sensitivity is and isn't effective at promoting deep learning. In *International Conference on Intelligent Tutoring Systems* (pp. 245-254). Springer Berlin Heidelberg.
- Dahn, B. I., Gehne, J., Honigmann, T., & Wolf, A. (1997, July). Integration of automated and interactive theorem proving in ILF. In *International Conference on Automated Deduction* (pp. 57-60). Springer Berlin Heidelberg.
- Dancey, C. P., & Reidy, J. (2007). *Statistics without maths for psychology*. Pearson Education.
- De Freitas, S., Rebolledo-Mendez, G., Liarokapis, F., Magoulas, G., & Poulouvassilis, A. (2010). Learning as immersive experiences: Using the four-dimensional framework for designing and evaluating immersive learning experiences in a virtual world. *British Journal of Educational Technology*, 41(1), 69-85.
- DeNero, J., & Klein, D. (2010). Teaching introductory artificial intelligence with pac-man. In *Proceedings of the Symposium on Educational Advances in Artificial Intelligence*.
- Dickey, M. D. (2011). Murder on Grimm Isle: The impact of game narrative design in an educational game-based learning environment. *British Journal of Educational Technology*, 42(3), 456-469.
- Douce, C., Livingstone, D., & Orwell, J. (2005). Automatic test-based assessment of programming: A review. *Journal on Educational Resources in Computing (JERIC)*, 5(3), 4.
- Dyckhoff, A. L., Zielke, D., Bültmann, M., Chatti, M. A., & Schroeder, U. (2012). Design and Implementation of a Learning Analytics Toolkit for Teachers. *Educational Technology & Society*, 15(3), 58-76.
- Eagle, M., & Barnes, T. (2008). Wu's castle: teaching arrays and loops in a game. *ACM SIGCSE Bulletin*, 40(3), 245-249.
- Ebner, M., & Holzinger, A. (2007). Successful implementation of user-centered game based learning in higher education: An example from civil engineering. *Computers & education*, 49(3), 873-890.
- Emberson, S. E., Reise, S. P. *Item Response Theory for Psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates (2000).
- Fiedler, A., & Tsovaltzi, D. (2003). Automating hinting in an intelligent tutorial dialog system for mathematics. *Knowledge Representation and Automated Reasoning for E-Learning Systems*, 23.
- Friedman-Hill, E. (2003). *Jess in Action: Rule-Based Systems in Java*. Manning Publications Company.
- Gasquet, O., Schwarzentruher, F., & Strecker, M. (2011). Satoulouse: the computational power of propositional logic shown to beginners. In *Tools for Teaching Logic* (pp. 77-84). Springer Berlin Heidelberg.
- Girvan, C., & Savage, T. (2010). Identifying an appropriate pedagogy for virtual worlds: A Communal Constructivism case study. *Computers & Education*, 55(1), 342-349.
- Gómez-Pérez, A., & Corcho, O. (2002). Ontology languages for the semantic web. *IEEE Intelligent systems*, 17(1), 54-60.
- Gordon, M. J., & Melham, T. F. (1993). Introduction to HOL A theorem proving environment for higher order logic.

- Goto, T., Kojiri, T., Watanabe, T., Iwata, T., & Yamada, T. (2010). Automatic generation system of multiple-choice cloze questions and its evaluation. *Knowledge Management & E-Learning: An International Journal (KM&EL)*, 2(3), 210-224.
- Gouli, E., Gogoulou, A., Papanikolaou, K. A., & Grigoriadou, M. (2006). An adaptive feedback framework to support reflection, guiding and tutoring. *Advances in web-based education: Personalized learning environments*, 178-202.
- Graesser, A. C., Conley, M. W., & Olney, A. (2012). Intelligent tutoring systems. In S. Graham & K. Harris (Eds.), *APA Educational Psychology Handbook: Vol. 3. Applications to Learning and Teaching* (pp. 451 – 473). Washington, DC: American Psychological Association
- Gregory, S., Lee, M. J., Dalgarno, B., & Tynan, B. (Eds.). (2016). *Learning in Virtual Worlds: Research and Applications*. Athabasca University Press.
- Grivokostopoulou, F., & Hatzilygeroudis, I. (2015). Semi-automatic generation of interactive exercises related to search algorithms. In *Computer Science & Education (ICCSE), 2015 10th International Conference on* (pp. 33-37). IEEE.
- Grivokostopoulou, F., & Hatzilygeroudis, I. (2013a). An Automatic Marking System for Interactive Exercises on Blind Search Algorithms. In *Artificial Intelligence in Education (AIED)* (pp. 783-786). Springer Berlin Heidelberg.
- Grivokostopoulou, F., & Hatzilygeroudis, I. (2013b). Automated marking for interactive exercises on heuristic search algorithms. In *Teaching, Assessment and Learning for Engineering (TALE), 2013 IEEE International Conference on* (pp. 598-603). IEEE
- Grivokostopoulou, F., & Hatzilygeroudis, I. (2013c). Teaching AI Search Algorithms in a Web-Based Educational System. *Proceedings of the IADIS International Conference e-Learning 2013*, pp. 83-90, 23 – 26 July, Prague, Czech Republic.
- Grivokostopoulou, F., & Hatzilygeroudis, I. (2010). A Tool for Experimenting with a Theorem Prover *Proceedings Of Second International Conference on Advanced Cognitive Technologies and Applications (COGNITIVE 2010)* pages pp. 78-83.
- Grivokostopoulou, F., Hatzilygeroudis, I., & Perikos, I. (2014a). Teaching assistance and automatic difficulty estimation in converting first order logic to clause form. *Artificial Intelligence Review*, 42(3), 347-367.
- Grivokostopoulou, F., Hatzilygeroudis, I., & Perikos, I. (2014b). An Intelligent System for Learning First Order Logic to Clause Form Conversion. *Multicultural Awareness and Technology in Higher Education: Global Perspectives: Global Perspectives*, 265.
- Grivokostopoulou, F., Perikos, I., & Hatzilygeroudis, I. (2016a). An Educational System for Learning Search Algorithms and Automatically Assessing Student Performance. *International Journal of Artificial Intelligence in Education*, 1-34.
- Grivokostopoulou, F., Perikos, I., & Hatzilygeroudis, I. (2016b). Difficulty estimation of exercises on tree-based search algorithms using neuro-fuzzy and neuro-symbolic approaches , *Advances in Combining Intelligent Methods*, Springer (to appear).
- Grivokostopoulou, F., Perikos, I., & Hatzilygeroudis, I. (2016c). An Educational Game for Teaching Search Algorithms. In *Proceedings of the 8th International Conference on Computer Supported Education* 179-3, pp. 129-136.
- Grivokostopoulou, F., Perikos, I., & Hatzilygeroudis, I. (2015). Estimating the Difficulty of Exercises on Search Algorithms Using a Neuro-Fuzzy Approach. In *Tools with Artificial Intelligence (ICTAI), 2015 IEEE 27th International Conference on* (pp. 866-872). IEEE.
- Grivokostopoulou, F., Perikos, I., & Hatzilygeroudis, I. (2014c). Utilizing semantic web technologies and data mining techniques to analyze students learning and predict final

- performance. In *Teaching, Assessment and Learning (TALE), 2014 International Conference on* (pp. 488-494). IEEE.
- Grivokostopoulou, F., Perikos, I., & Hatzilygeroudis, I. (2014d). Using Semantic Web Technologies in a Web Based System for Personalized Learning AI Course. In *Technology for Education (T4E), 2014 IEEE Sixth International Conference on* (pp. 257-260). IEEE.
- Grivokostopoulou, F., Perikos, I., & Hatzilygeroudis, I. (2013, June). An intelligent tutoring system for teaching fol equivalence. In *The First Workshop on AI-supported Education for Computer Science (AIEDCS 2013)* (p. 20-29).
- Grivokostopoulou, F., Perikos, I., & Hatzilygeroudis, I. (2012a). An automatic marking system for FOL to CF conversions. In *Teaching, Assessment and Learning for Engineering (TALE), 2012 IEEE International Conference on* (pp. H1A-7). IEEE.
- Grivokostopoulou, F., Perikos, I., & Hatzilygeroudis, I. (2012b). Assistant tools for teaching FOL to CF conversion. In *IFIP International Conference on Artificial Intelligence Applications and Innovations* (pp. 306-315). Springer Berlin Heidelberg.
- Grivokostopoulou, F., Perikos, I., & Hatzilygeroudis, I. (2012c). A web-based interactive system for learning FOL to CF conversion. In *Proceedings of the IADIS international conference e-learning* (pp. 287-294).
- Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing?. *International journal of human-computer studies*, 43(5), 907-928.
- Guarino, N. (1995). Formal ontology, conceptual analysis and knowledge representation. *International journal of human-computer studies*, 43(5), 625-640.
- Hansen, S., Narayanan, N. H., & Hegarty, M. (2002). Designing educationally effective algorithm visualizations. *Journal of Visual Languages & Computing*, 13(3), 291-317.
- Harris, K., & Reid, D. (2005). The influence of virtual reality play on children's motivation. *Canadian Journal of Occupational Therapy*, 72(1), 21-29.
- Hartness, K. (2004). Robocode: using games to teach artificial intelligence. *Journal of Computing Sciences in Colleges*, 19(4), 287-291.
- Hatcher, L. (1994). A step-by-step approach to using the SAS(R) system for factor analysis and structural equation modeling. Cary, NC: SAS Institute.
- Hatzilygeroudis, I. (2007): Teaching NL to FOL and FOL to CL Conversions. In: *Proceedings of the 20th International FLAIRS Conference*, Key West, FL, May 2007, AAAI Press, Menlo Park 309–314.
- Hatzilygeroudis, I., Grivokostopoulou, F., & Perikos, I. (2012, August). Using game-based learning in teaching CS algorithms. In *Teaching, Assessment and Learning for Engineering (TALE), 2012 IEEE International Conference on* (pp. H2C-9). IEEE.
- Hatzilygeroudis, I., Grivokostopoulou, F., & Perikos, I. (2012, July). Teaching aspects of constraint satisfaction algorithms via a game. In *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence* (pp. 2371-2372). AAAI Press.
- Hatzilygeroudis, I., Koutsojannis, C., Papavlasopoulos, C., & Prentzas, J. (2006, July). Knowledge-based adaptive assessment in a Web-based intelligent educational system. In *Sixth IEEE International Conference on Advanced Learning Technologies (ICALT'06)* (pp. 651-655). IEEE.
- Hatzilygeroudis, I., Perikos, I. (2009) A web-Based Interactive System for Learning NL to FOL Conversion. *New Directions in Intelligent Interactive Multimedia Systems and Services – 2. Studies in Computational Intelligence*. Springer-Verlag, Volume 226/2009, 297-307.

-
- Hatzilygeroudis, I., & Prentzas, J. (2015). Symbolic-neural rule based reasoning and explanation. *Expert Systems with Applications*, 42, 4595-4609.
- Hatzilygeroudis, I., & Prentzas, J. (2010). Integrated rule-based learning and inference. *IEEE Transactions on Knowledge and Data Engineering*, 22, 1549-1562.
- Hatzilygeroudis, I., & Reichgelt, H. (1994). ACT-P: a configurable theorem-prover. *Data & knowledge engineering*, 12(3), 277-296.
- Heffernan, N. T., & Heffernan, C. L. (2014). The ASSISTments ecosystem: building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education*, 24(4), 470-497.
- Heffernan, N. T., & Koedinger, K. R. (2002, June). An intelligent tutoring system incorporating a model of an experienced human tutor. In *International Conference on Intelligent Tutoring Systems* (pp. 596-608). Springer Berlin Heidelberg.
- Heilman, M., & Smith, N. A. (2010, June). Good question! statistical ranking for question generation. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 609-617). Association for Computational Linguistics.
- Helmick, M. T. (2007, June). Interface-based programming assignments and automatic grading of java programs. In *ACM SIGCSE Bulletin* (Vol. 39, No. 3, pp. 63-67). ACM.
- Hendriks, M., Kaliszyk, C., Van Raamsdonk, F., & Wiedijk F. (2010). Teaching logic using a state-of-the-art proof assistant. *Acta Didactica Napocensia*, 3(2), 35-48.
- Hershkovitz, A., de Baker, R. S. J., Gobert, J., Wixon, M., & Sao Pedro, M. (2013). Discovery With Models A Case Study on Carelessness in Computer-Based Science Inquiry. *American Behavioral Scientist*, 57(10), 1480-1499.
- Higgins, C. A., & Bligh, B. (2006). Formative computer based assessment in diagram based domains. *ACM SIGCSE Bulletin*, 38(3), 98-102.
- Higgins, C. A., Gray, G., Symeonidis, P., & Tsintsifas, A. (2005). Automated assessment and experiences of teaching programming. *Journal on Educational Resources in Computing (JERIC)*, 5(3), 5.
- Hirashima, T., Horiguchi, T., Kashiara, A., & Toyoda, J. (1998). Error-based simulation for error-visualization and its management. *International Journal of Artificial Intelligence in Education*, 9(1-2), 17-31.
- Hoshino, A., & Nakagawa, H. (2005, June). A real-time multiple-choice question generation for language testing: a preliminary study. In *Proceedings of the second workshop on Building Educational Applications Using NLP* (pp. 17-20). Association for Computational Linguistics.
- Hundhausen, C. D., Douglas, S. A., & Stasko, J. T. (2002). A meta-study of algorithm visualization effectiveness. *Journal of Visual Languages & Computing*, 13(3), 259-290.
- Hurd, J. (1999, September). Integrating gandalf and hol. In *International Conference on Theorem Proving in Higher Order Logics* (pp. 311-321). Springer Berlin Heidelberg.
- Hwang, G. J., H. Y., Hung, C. M., Huang, I., & Tsai, C. C. (2012). Development of a personalized educational computer game based on students' learning styles. *Educational Technology Research and Development*, 60(4), 623-638.
- Jackson, D., & Usher, M. (1997, March). Grading student programs using ASSYST. In *ACM SIGCSE Bulletin* (Vol. 29, No. 1, pp. 335-339). ACM.
-

- Jang, J. S. (1993). ANFIS: adaptive-network-based fuzzy inference system. *Systems, Man and Cybernetics, IEEE Transactions on*, 23(3), 665-685.
- Joy, M., Griffiths, N., & Boyatt, R. (2005). The boss online submission and assessment system. *Journal on Educational Resources in Computing (JERIC)*, 5(3), 2.
- Kakar, M., Nyström, H., Aarup, L. R., Nøttrup, T. J., & Olsen, D. R. (2005). Respiratory motion prediction by using the adaptive neuro fuzzy inference system (ANFIS). *Physics in medicine and biology*, 50(19), 4721.
- Kaliszyk, C., Van Raamsdonk, F., Wiedijk, F., Wupper, H., Hendriks, M., & de Vrijer, R. (2008). Deduction using the ProofWeb system. *report ICISR08016, Radboud University Nijmegen*.
- Kapp, K. M. (2012). *The gamification of learning and instruction: game-based methods and strategies for training and education*. John Wiley & Sons.
- Khachatryan, G. A., Romashov, A. V., Khachatryan, A. R., Gaudino, S. J., Khachatryan, J. M., Guarian, K. R., & Yufa, N. V. (2014). Reasoning Mind Genie 2: an intelligent tutoring system as a vehicle for international transfer of instructional methods in mathematics. *International Journal of Artificial Intelligence in Education*, 24(3), 333-382.
- Koedinger, K. R., McLaughlin, E. A., & Heffernan, N. T. (2010). A quasi-experimental evaluation of an on-line formative assessment and tutoring system. *Journal of Educational Computing Research*, 43(4), 489-510.
- Kordaki, M., Miatidis, M., & Kapsampelis, G. (2008). A computer environment for beginners' learning of sorting algorithms: Design and pilot evaluation. *Computers & Education*, 51(2), 708-723.
- Korovin, K. (2008, August). iProver—an instantiation-based theorem prover for first-order logic (system description). In *International Joint Conference on Automated Reasoning* (pp. 292-298). Springer Berlin Heidelberg.
- Koutsojannis, C., Beligiannis, G., Hatzilygeroudis, I., Papavlasopoulos, C., Prentzas, J. (2007) Using a hybrid AI approach for exercise difficulty level adaptation, *International Journal of Continuing Engineering Education and Life-Long Learning*, 17(4-5), 256-272.
- Ku, K., & Mahabaleshwarkar, P. S. (2011). Building interactive modeling for construction education in virtual worlds. *Journal of Information Technology in Construction*, 16, 189-208.
- Kunichika, H., Urushima, M., Hirashima, T., & Takeuchi (2002) A.. A computational method of complexity of questions on contents of English sentences and its evaluation. In *Proceedings of the International Conference on Computers in Education*. Auckland, New Zealand: IEEE Computer Society 97–101.
- Lepper, M. R., Woolverton, M., Mumme, D. L., & Gurtner, J. (1993). Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. *Computers as cognitive tools*, 1993, 75-105.
- Leung, E. W. C., & Li, Q. (2007). An experimental study of a personalized learning environment through open-source software tools. *Education, IEEE Transactions on*, 50(4), 331-337.
- Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. In *Soviet physics doklady* (Vol. 10, No. 8, pp. 707-710).
- Levitin, A., & Papalaskari, M. A. (2002). Using puzzles in teaching algorithms. In *ACM SIGCSE Bulletin* (Vol. 34, No. 1, pp. 292-296). ACM.

-
- Liu, T. Y., & Chu, Y. L. (2010). Using ubiquitous games in an English listening and speaking course: Impact on learning outcomes and motivation. *Computers & Education*, 55(2), 630-643.
- Lodder, J., Passier, H., & Stuurman, S. (2008, December). Using IDEAS in Teaching Logic, Lessons Learned. In *CSSE (5)* (pp. 553-556).
- Long, Y., & Aleven, V. (2013, July). Skill diaries: Improve student learning in an intelligent tutoring system with periodic self-assessment. In *Artificial intelligence in education* (pp. 249-258). Springer Berlin Heidelberg.
- Lukins, S., Levicki, A., & Burg, J. (2002, February). A tutorial program for propositional logic with human/computer interactive learning. In *ACM SIGCSE Bulletin* (Vol. 34, No. 1, pp. 381-385). ACM.
- Magdaleníć, I., Radošević, D., & Orehovački, T. (2013). Autogenerator: Generation and execution of programming code on demand. *Expert systems with applications*, 40(8), 2845-2857.
- Malafeev, A. (2014, April). Automatic Generation of Text-Based Open Cloze Exercises. In *International Conference on Analysis of Images, Social Networks and Texts_x000D_* (pp. 140-151). Springer International Publishing.
- Malone T. W. (1981) Towards a theory of intrinsically motivating instruction *Cognitive Science*, 5 (4), pp. 333–369
- Malone, T. W., & Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. *Aptitude, learning, and instruction*, 3(1987), 223-253.
- Markov, Z., Russell, I., Neller, T., & Zlatareva, N. (2006). Pedagogical possibilities for the N-puzzle problem. In *Frontiers in Education Conference, 36th Annual* (pp. 1-6). IEEE.
- McArthur, D., Stasz, C., & Zmuidzinas, M. (1990). Tutoring techniques in algebra. *Cognition and Instruction*, 7(3), 197–244.
- McCabe, T. J. (1976). A complexity measure. Page 407 of: ICSE '76: Proceedings of the 2nd international conference on Software engineering. Los Alamitos, CA, USA: IEEE Computer Society Press.
- McCune, W. (1992). Experiments with discrimination-tree indexing and path indexing for term retrieval. *Journal of Automated Reasoning*, 9(2), 147-167.
- Merrill, D. C., Reiser, B. J., Ranney, M., & Trafton, J. G. (1992). Effective tutoring techniques: A comparison of human tutors and intelligent tutoring systems. *The Journal of the Learning Sciences*, 2(3), 277-305.
- Mikropoulos, T. A. and Natsis, A., 2011. Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & Education*, 56(3), 769-780.
- Mitrovic, A., & Ohlsson, S. (1999). Evaluation of a constraint-based tutor for a database language.
- Mory, E. H. (2004). Feedback research revisited. *Handbook of research on educational communications and technology*, 2, 745-783.
- Moser, M., Ibens, O., Letz, R., Steinbach, J., Goller, C., Schumann, J., & Mayr, K. (1997). SETHEO and e-SETHEO-the CADE-13 systems. *Journal of Automated Reasoning*, 18(2), 237-246.
- Nadel, B. A. (1990). *Some applications of the constraint satisfaction problem*. Wayne State University, Department of Computer Science.
- Naps, T. L., Rößling, G., Almstrum, V., Dann, W., Fleischer, R., Hundhausen, C. .., & Velázquez-Iturbide, J. Á. (2002, June). Exploring the role of visualization and
-

- engagement in computer science education. In *ACM Sigcse Bulletin* (Vol. 35, No. 2, pp. 131-152). ACM.
- Narciss, S. (2008). Feedback strategies for interactive learning tasks. *Handbook of research on educational communications and technology*, 3, 125-144.
- Naudé, K. A., Greyling, J. H., & Vogts, D. (2010). Marking student programs using graph similarity. *Computers & Education*, 54(2), 545-561.
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in higher education*, 31(2), 199-218.
- Nikander, J., Korhonen, A., Seppälä, O., Karavirta, V., Silvasti, P., & Malmi, L. (2004). Visual algorithm simulation exercise system with automatic assessment: TRAKLA2. *Informatics in Education-An International Journal.*, (Vol 3_2), 267-288.
- Nye, B. D., Graesser, A. C., & Hu, X. (2014). AutoTutor and family: A review of 17 years of natural language tutoring. *International Journal of Artificial Intelligence in Education*, 24(4), 427-469.
- Olney, A. M., Cade, W. L., & Williams, C. (2011). Generating concept map exercises from textbooks. In *Proceedings of the 6th workshop on innovative use of NLP for building educational applications* (pp. 111-119). Association for Computational Linguistics.
- Pane, J. F., Griffin, B. A., McCaffrey, D. F., & Karam, R. (2013). Effectiveness of cognitive tutor algebra I at scale. *Educational Evaluation and Policy Analysis*, 0162373713507480.
- Papastergiou, M. (2009). Digital game-based learning in high school computer science education: Impact on educational effectiveness and student motivation. *Computers & Education*, 52(1), 1-12.
- Perikos, I., Grivokostopoulou, F., Hatzilygeroudis, I. (2016) Assistance and Feedback Mechanism in an Intelligent Tutoring System for Teaching Conversion of Natural Language into Logic, *International Journal of Artificial Intelligence in Education*, 2016 (To Appear)
- Perikos, I., Grivokostopoulou, F., Hatzilygeroudis, I., Kostas, K. (2016) Automatic estimation of exercises' difficulty levels in a tutoring system for teaching the conversion of natural language into first order logic". *Expert Systems*, 33(6), 569-580.
- Perikos, I., Grivokostopoulou, F., & Hatzilygeroudis, I. (2011a). Teaching assistant tools for NL to FOL Conversion. In *Proceedings of the IADIS International Conference e-Learning* (pp. 20-23).
- Perikos, I., Grivokostopoulou, F., Hatzilygeroudis, I., Kostas, K. (2011b). Difficulty Estimator for Converting Natural Language into First Order Logic. In *Intelligent Decision Technologies. Smart Innovation, Systems and Technologies Volume 10*(2011) 135-14
- Perikos, I., Grivokostopoulou, F., & Hatzilygeroudis, I. (2011c) Help Generation in a System for learning Natural Language to First Order Logic Conversion. In *Proceedings of the 14th IASTED International Conference on Computers and Advanced Technology in Education (CATE 2011)*.
- Perikos, I., Grivokostopoulou, F., & Hatzilygeroudis, I. (2012). Automatic marking of NL to FOL conversions. In *Proc. of 15th IASTED International Conference on Computers and Advanced Technology in Education (CATE), Napoli, Italy* (pp. 227-233).
- Prensky, M. (2001). Digital game-based learning.

- Psyché, V., Kalemjian, A., Mozes, J., Maurice, C., Poirier, P., Bourdeau, J., & Nkambou, R. (2014). GYM-Author: Generation Of Self-Learning Exercises In Philosophy.
- Radošević, D., Orehovački, T., & Stapić, Z. (2010, September). Automatic on-line generation of student's exercises in teaching programming. In Radošević, D., Orehovački, T., Stapić, Z.: " Automatic On-line Generation of Students Exercises in Teaching Programming", Central European Conference on Information and Intelligent Systems, CECIIS.
- Riazanov, A., & Voronkov, A. (2002). The design and implementation of VAMPIRE. *AI communications*, 15(2, 3), 91-110.
- Rosa, K. D., & Eskenazi, M. (2013). Self-Assessment in the REAP tutor: knowledge, interest, motivation, & learning. *International Journal of Artificial Intelligence in Education*, 21(4), 237-253.
- Russell, S., Norvig, P. (2003). Artificial Intelligence: a modern approach, 2nd edn. Prentice Hall, Upper Saddle River
- San Pedro, M. O. C. Z., d Baker, R. S., & Rodrigo, M. M. T. (2011). Detecting carelessness through contextual estimation of slip probabilities among students using an intelligent tutor for mathematics. In *Artificial Intelligence in Education* (pp. 304-311). Springer Berlin Heidelberg.
- San Pedro, M. O. Z., d Baker, R. S., & Rodrigo, M. M. T. (2014). Carelessness and affect in an intelligent tutoring system for mathematics. *International Journal of Artificial Intelligence in Education*, 24(2), 189-210.
- Sánchez-Torrubia, M. G., Torres-Blanc, C., & Lopez-Martinez, M. A. (2009). PathFinder: A visualization eMathTeacher for actively learning Dijkstra's algorithm. *Electronic Notes in Theoretical Computer Science*, 224, 151-158.
- Shabanah, S. S., Chen, J. X., Wechsler, H., Carr, D., & Wegman, E. (2010, April). Designing computer games to teach algorithms. In *Information Technology: New Generations (ITNG)*, 2010 Seventh International Conference on (pp. 1119-1126). IEEE.
- Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, 78(1), 153-189.
- Shute, V. J., & Zapata-Rivera, D. (2010). Educational measurement and intelligent systems. of the International Encyclopedia of Education. Oxford, UK: Elsevier Publishers.
- Sieg, W. (2007). The AProS project: Strategic thinking & computational logic. *Logic journal of the IGPL* 15(4), 359-368.
- Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427-437.
- Sosnowski, S., Ernsberger, T., Cao, F., & Ray, S. (2013). SEPIA: A Scalable Game Environment for Artificial Intelligence Teaching and Research. In Proceedings of the 27th AAAI Conference on Artificial Intelligence
- Stajduhar, I., & Mause, G. (2015, May). Using string similarity metrics for automated grading of SQL statements. In *Information and Communication Technology, Electronics and Microelectronics* , 2015 38th International Convention on (pp. 1250-1255). IEEE.
- Stefan, L. (2016, April). Virtual Worlds In Online Education And Training-An Evaluation Report. In *The International Scientific Conference eLearning and Software for Education* (Vol. 2, p. 453). " Carol I" National Defence University

- Sung, H. Y., & Hwang, G. J. (2013). A collaborative game-based learning approach to improving students' learning performance in science courses. *Computers & Education*, 63, 43-51.
- Sutcliffe, G., & Suttner, C. (1998). The TPTP problem library. *Journal of Automated Reasoning*, 21(2), 177-203.
- Taylor, M. E. (2011, August). Teaching reinforcement learning with mario: An argument and case study. In *Proceedings of the Second Symposium on Educational Advances in Artificial Intelligence* (pp. 1737-1742).
- Thomas, P., Smith, N., & Waugh, K. (2008). Automatically assessing graph-based diagrams. *Learning, Media and Technology*, 33(3), 249-267.
- Tiantian, W., Xiaohong, S., Peijun, M., Yuying, W., & Kuanquan, W. (2009, March). AutoLEP: An Automated Learning and Examination System for Programming and its Application in Programming Course. In *Education Technology and Computer Science, 2009. ETCS'09. First International Workshop on* (Vol. 1, pp. 43-46). IEEE.
- Tüzün, H., Yılmaz-Soylu, M., Karakuş, T., İnal, Y., & Kızılkaya, G. (2009). The effects of computer games on primary school students' achievement and motivation in geography learning. *Computers & Education*, 52(1), 68-77.
- Van de Watering, G., & Van der Rijt, J. (2006). Teachers' and students' perceptions of assessments: A review and a study into the ability and accuracy of estimating the difficulty levels of assessment items. *Educational Research Review*, 1(2), 133-147.
- VanLehn, K. (2006). The behavior of tutoring systems. *International journal of artificial intelligence in education*, 16(3), 227-265.
- VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197-221.
- Vanlehn, K., Lynch, C., Schulze, K., Shapiro, J. A., Shelby, R., Taylor, L., ... & Wintersgill, M. (2005). The Andes physics tutoring system: Lessons learned. *International Journal of Artificial Intelligence in Education*, 15(3), 147-204.
- Verdú, E., Verdú, M. J., Regueras, L. M., de Castro, J. P., & García, R. (2012). A genetic fuzzy expert system for automatic question classification in a competitive learning environment. *Expert Systems with Applications*, 39(8), 7471-7478.
- Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. *Fam Med*, 37(5), 360-363.
- Wang, T., Su, X., Ma, P., Wang, Y., & Wang, K. (2011). Ability-training-oriented automated assessment in introductory programming course. *Computers & Education*, 56(1), 220-226.
- Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. *Metacognition in educational theory and practice*, 93, 27-30.
- Wong, D., Zink, R., & Koenig, S. (2010). Teaching artificial intelligence and robotics via games. In *Proc. of the EAAI Symposium*.
- Woolf, B. P. (2010). Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning. Morgan Kaufman.
- Wu, W. H., Hsiao, H. C., Wu, P. L., Lin, C. H., & Huang, S. H. (2012). Investigating the learning-theory foundations of game-based learning: a meta-analysis. *Journal of Computer Assisted Learning*, 28(3), 265-279.
- Yacef, K. (2005). The Logic-ITA in the classroom: a medium scale experiment. *International Journal of Artificial Intelligence in Education*, 15(1), 41-62.

- Yessad, A., Labat, J. M., & Kermorvant, F. (2010). Segae: A serious game authoring environment. In *Advanced learning technologies (icalt), 2010 ieee 10th international conference on* (pp. 538-540). IEEE.