

Συνέδρια της Ελληνικής Επιστημονικής Ένωσης Τεχνολογιών Πληροφορίας & Επικοινωνιών στην Εκπαίδευση

Τόμ. 1 (2010)

5ο Συνέδριο Διδακτική της Πληροφορικής



Συνθετική Εργασία Μεταπτυχιακών Φοιτητών Πληροφορικής σε Υπολογιστικά Σημασιολογικά Λεξικά στα πλαίσια του μαθήματος «Υπολογιστική Γλωσσολογία»

Π. Μπλίτσας, Μ. Γρηγοριάδου

Βιβλιογραφική αναφορά:

Μπλίτσας Π., & Γρηγοριάδου Μ. (2023). Συνθετική Εργασία Μεταπτυχιακών Φοιτητών Πληροφορικής σε Υπολογιστικά Σημασιολογικά Λεξικά στα πλαίσια του μαθήματος «Υπολογιστική Γλωσσολογία». *Συνέδρια της Ελληνικής Επιστημονικής Ένωσης Τεχνολογιών Πληροφορίας & Επικοινωνιών στην Εκπαίδευση*, 1, 445–450. ανακτήθηκε από <https://eproceedings.epublishing.ekt.gr/index.php/cetpe/article/view/5175>

Συνθετική Εργασία Μεταπτυχιακών Φοιτητών Πληροφορικής σε Υπολογιστικά Σημασιολογικά Λεξικά στα πλαίσια του μαθήματος «Υπολογιστική Γλωσσολογία»

Π. Μπλίτσας¹, Μ. Γρηγοριάδου²

¹Π.Μ.Σ. “Βασική και Εφαρμοσμένη Γνωστική Επιστήμη”
Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών
pblitsas@di.uoa.gr

²Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών
gregor@di.uoa.gr

Περίληψη

Στην παρούσα εργασία παρουσιάζεται συνθετική εργασία εμπλουτισμού της ελληνικής βάσης του υπολογιστικού σημασιολογικού λεξικού WordNet με όρους του τομέα της Επιστήμης Υπολογιστών που δόθηκε σε μεταπτυχιακούς φοιτητές Πληροφορικής με σκοπό την εξοικείωσή τους με τέτοιου είδους λεξικά, την καταγραφή δυσκολιών που εμφανίστηκαν κατά την διαδικασία και προτάσεων διόρθωσης, και την επαφή τους με γνωστικό μοντέλο κατανόησης κειμένου. Η εργασία πραγματοποιήθηκε στα πλαίσια του μαθήματος «Υπολογιστική Γλωσσολογία» και στηρίχθηκε στο μοντέλο κατανόησης κειμένου των Baudet & Denhière, έτσι ώστε, σε αντίθεση με όλα τα υφιστάμενα Wordnets, να περιγράφονται με πιο αποτελεσματικό τρόπο, όχι μόνο οι στατικές σχέσεις μεταξύ των όρων που εκφράζουν τεχνικές οντότητες μέσα στο λεξικό, αλλά και εκείνες οι σχέσεις που αναφέρονται στα γεγονότα που συμβαίνουν κατά τη διάρκεια των λειτουργιών των μονάδων ενός υπολογιστικού συστήματος, καθώς και οι σκοποί που αυτές οι λειτουργίες ή/και οι μονάδες επιτελούν.

Λέξεις κλειδιά: μοντέλο κατανόησης κειμένου, υπολογιστική γλωσσολογία, υπολογιστικά λεξικά

Abstract

In the present work, an enrichment task of the Greek computational semantic dictionary WordNet database on the domain of Computer Science is presented, which was given to postgraduate students of Informatics, in order for them to get familiar with this kind of computational dictionaries, to write down difficulties appeared during the procedure and propositions for improvements of the tool, as well as to get in touch with a cognitive text comprehension model. This work took place in the frame of the “Computational Linguistics» subject and was based on Baudet & Denhière Text Comprehension Model, so that, unlike all the existing WordNets, this extended WordNet could describe in a more efficient way, except for the static relations among the terms expressing technical entities, more relations referring to events that happen during the operations of the computer units and to goals that these operations or/and units have to achieve.

Keywords: *computational dictionaries, computational linguistics, text comprehension model*

1. Εισαγωγή

Ένας από τους βασικούς στόχους του μαθήματος της «Υπολογιστικής Γλωσσολογίας» του Π.Μ.Σ. του τμήματος Πληροφορικής & Τηλεπικοινωνιών είναι η εξοικείωση των φοιτητών με τη Σημασιολογία Λέξεων (*Lexical Semantics*), τα Σημασιολογικά Λεξικά τύπου Wordnet και την Άρση της Σημασιολογικής Αμφισημίας (*Word Sense Disambiguation*). Η παρούσα εργασία λοιπόν αποσκοπεί στην επαφή των μεταπτυχιακών φοιτητών με υπολογιστικά λεξικά τύπου WordNet και την εξοικείωσή τους με τον Visdic Editor, καθώς και στην επαφή τους με γνωστικό μοντέλο κατανόησης κειμένου. Η συνθετική εργασία που τους ανατέθηκε ήταν ο εμπλουτισμός του ελληνικού WordNet με όρους που σχετίζονται με τον τομέα της Επιστήμης Υπολογιστών, και ειδικότερα με σύγχρονους όρους Δικτύων Η/Υ και Βάσεων Δεδομένων, και σημασιολογικές σχέσεις μεταξύ τους, προκειμένου το λεξικό αυτό να μπορεί να χρησιμοποιηθεί σε περιπτώσεις αυτών των περιοχών της Επιστήμης Υπολογιστών. Ακόμη και αν, σε πολλές περιπτώσεις, υπάρχουν όροι σχετικοί με την επιστήμη των υπολογιστών στη βάση του ήδη υπάρχοντος ελληνικού WordNet, αυτοί έχουν μια εντελώς διαφορετική ερμηνεία από εκείνη που θα έπρεπε να έχουν στον τομέα αυτό. Έτσι, η χρήση τους σε ένα τέτοιο τομέα δεν είναι δυνατή. Επιπλέον, ένα μέρος από τις σημασιολογικές σχέσεις που

απαιτήθηκαν δεν υποστηρίζεται ακόμα και από το αγγλικό WordNet. Οι σχέσεις αυτές καθορίζουν μη τυποποιημένα χαρακτηριστικά γνωρίσματα, όπως "έχει χαρακτηριστικό", "έχει λειτουργία", κλπ.

Οι φοιτητές έπρεπε κατά τη διάρκεια του εμπλουτισμού αυτού να εξετάσουν σε βάθος τα χαρακτηριστικά που ήθελαν να προσδώσουν σε κάθε όρο. Η μεθοδολογία εισαγωγής των όρων παίζει ιδιαίτερα σημαντικό ρόλο, διότι ο κίνδυνος των «σημασιολογικών κυκλικών αναφορών» είναι πολύ μεγάλος (ενότητα 3). Μια αντιφατική και ελλιπής σημασιολογική ανάλυση θα μπορούσε να οδηγήσει σε κυκλικές αναφορές μεταξύ των νέων όρων και των ήδη υπαρχόντων. Σε μια τέτοια περίπτωση, η σημασιολογική βάση θα οδηγούσε σε εσφαλμένες λογικές σχέσεις μεταξύ των όρων, και κατά συνέπεια θα υπήρχε σημαντικός αντίκτυπος στην κατανόηση της σημασιολογικής οντολογίας. Ως εκ τούτου, ήταν αναγκαίο το ελληνικό Wordnet να εμπλουτιστεί με περισσότερα synsets και σχέσεις μεταξύ τους μετά από προσεκτική ανάλυση, υπό την προοπτική της «συγχώνευσης» του με όλα τα παρόμοια έργα εμπλουτισμού στον τομέα της Επιστήμης των Υπολογιστών (Kremizis et al., 2007). Στην παρούσα εργασία ο εμπλουτισμός του ελληνικού WordNet, όσον αφορά στις σχέσεις μεταξύ των όρων, ακολούθησε το γνωστικό μοντέλο κατανόησης κειμένου των Baudet και Denhière (1992).

Στη 2^η ενότητα παρουσιάζονται το «αναγνώσιμο από μηχανή» - machine readable λεξικό WordNet και ο VisDic Editor, ενώ στην 3^η ενότητα περιγράφονται οι διδακτικοί στόχοι της συνθετικής εργασίας, η διαδικασία εκπόνησης του εμπλουτισμού του ελληνικού WordNet με βάση το μοντέλο κατανόησης κειμένου των Baudet & Denhière, που ακολουθήθηκε από τους μεταπτυχιακούς φοιτητές, καθώς και παραδείγματα του εμπλουτισμένου πια WordNet. Τέλος, στην 4^η ενότητα συζητούνται προτάσεις βελτίωσης των υπαρχόντων εργαλείων από τους φοιτητές καθώς και προβλήματα που επισημάνθηκαν από τους ίδιους κατά τη διαδικασία της συνθετικής εργασίας.

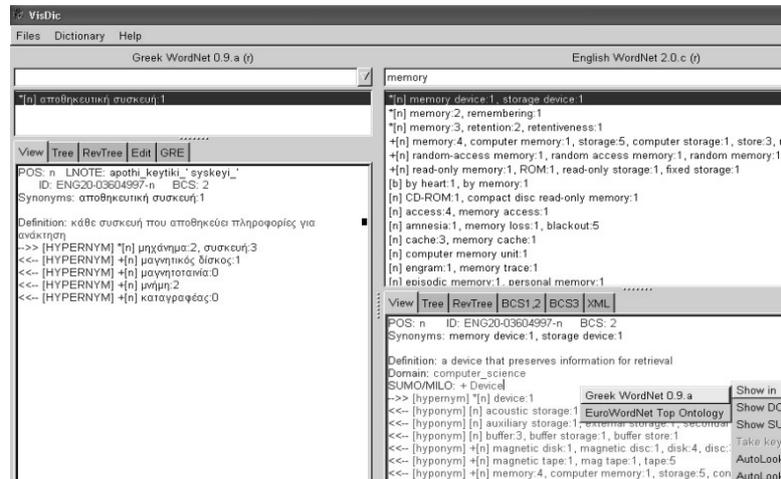
2. Wordnets και VisDic Editor

Το πρώτο WordNet (<http://wordnet.princeton.edu/wordnet/>) δημιουργήθηκε στο Πρίνσετον και ήταν στα αγγλικά. Αμέσως, αναπτύχθηκαν άλλα WordNets όπως το BalkaNet και το EuroWordNet. Στην πραγματικότητα, το WordNet είναι μια μεγάλη αγγλική λεξικολογική βάση δεδομένων με βάση τις έννοιες των λέξεων και τις σημασιολογικές σχέσεις που υπάρχουν μεταξύ τους. Ουσιαστικά, ρήματα, επίθετα και επιρρήματα ομαδοποιούνται σε σύνολα συνωνύμων (synsets), καθένα από τα οποία εκφράζει κάτι το μοναδικό. Ένα synset είναι ένα σύνολο λέξεων, που σε δεδομένο περιβάλλον η μία λέξη μπορεί να χρησιμοποιηθεί στη θέση μιας άλλης. Με τη βοήθεια των σημασιολογικών σχέσεων, η σημασία των λέξεων σε μια γλώσσα μπορεί να συνδέονται μεταξύ τους, σχηματίζοντας ένα δίκτυο των εννοιών του WordNet. Ένα άλλο σημαντικό χαρακτηριστικό του Wordnet είναι ο διαχωρισμός των εννοιών σε πεδία (τομείς). Μια λέξη μπορεί να ανήκει σε διάφορες synsets σε πολλούς τομείς. Κάθε synset σε κάθε τομέα, έχει τη δική του ερμηνεία και σημασιολογικές συσχετίσεις υπερωνυμίας, υπωνυμίας, ολωνυμίας ή μερωνυμίας με άλλες έννοιες. Για παράδειγμα, η λέξη "μνήμη" παρουσιάζεται σε χωριστά synsets στην Ψυχολογία & στην Επιστήμη Υπολογιστών.

Οι βασικές σημασιολογικές σχέσεις που υποστηρίζονται από το wordnet είναι: (i) *Συνώνυμο (Synonym)*: Το X είναι συνώνυμο του Y, όταν το Y μπορεί να αντικατασταθεί από το X χωρίς σημασιολογικό και εννοιολογικό λάθος, π.χ. ο «δήμος» είναι συνώνυμο της «πόλης». (ii) *Υπώνυμο/Υπερώνυμο (Hyponym/Hypernym)*: Το X είναι υπέρωνυμο του Y (ή με άλλα λόγια το Y είναι υπώνυμο του X), όταν το Y είναι ένα είδος του X, π.χ. ο "σκύλος" είναι υπώνυμο του «ζώου» ή αλλιώς ο σκύλος είναι είδος ζώου, (iii) *Μερώνυμο/Ολώνυμο (Meronym/Olonym)*: Το X είναι μερώνυμο του Y (ή με άλλα λόγια το Y είναι ολώνυμο του X), όπως ο "προφυλακτήρας" είναι μερώνυμο του "αυτοκινήτου", δηλ. ο προφυλακτήρα είναι μέρος του αυτοκινήτου, (iv) *Συγγένεια (Familiarity)*: Το Y είναι συγγενής του X, όταν τα X και Y μοιράζονται το ίδιο υπέρωνυμο, π.χ. ο "Λύκος" είναι συγγενής του "σκύλου" και ο "σκύλος" είναι συγγενής του "λύκου", αφού και τα δυο έχουν υπέρωνυμο το "ζώο".

Το EuroWordNet έχει παράγει WordNets για διάφορες ευρωπαϊκές γλώσσες (<http://www.illc.uva.nl/EuroWordNet/>). Για το έργο αυτό αναπτύχθηκε ο VisDic Editor, ο οποίος αποτελεί εργαλείο συγγραφής σημασιολογικών λεξικών τύπου wordnet σε XML μορφή. Το ελληνικό wordnet αναπτύχθηκε στο Πανεπιστήμιο της Πάτρας από ομάδα γλωσσολόγων, με τη συμβολή του Πανεπιστημίου Αθηνών. Ο VisDic Editor (<http://nlp.fi.muni.cz/projekty/visdic/>) είναι μια γραφική

εφαρμογή (Horak & Smrž, 2004), η οποία χρησιμοποιείται για την αναζήτηση όρων και σημασιολογικών σχέσεων μεταξύ τους, την τροποποίηση λεξικών αποθηκευμένων σε XML (eXtensible Markup Language) μορφή, και την περιήγηση στα BalkaNet wordnets (Σχήμα 1).



Σχήμα 1: Παράδειγμα αναζήτησης του όρου «αποθηκευτική συσκευή» στο ελληνικό (αριστερά) και αγγλικό (δεξιά) WordNet μέσω VisDic editor

3. Εκπόνηση Συνθετικής Εργασίας στο Ελληνικό Σημασιολογικό Λεξικό WordNet

Βασικοί διδακτικοί στόχοι της συνθετικής εργασίας που ανατέθηκε στους φοιτητές ήταν:

- Η εξοικείωση τους με τα υπολογιστικά λεξικά, τους τρόπους αξιοποίησής τους και τις σχέσεις μεταξύ των εννοιών που αυτά υποστηρίζουν.
- Η εξοικείωσή τους με την πλατφόρμα του Visdic Editor και τον τρόπο εισαγωγής νέων εννοιών και σχέσεων μέσω της πλατφόρμας αυτή για το ελληνικό WordNet.

Αρχικά, πραγματοποιήθηκε τρίωρο φροντιστήριο ώστε να έρθουν οι φοιτητές σε επαφή με τα προς χρήση εργαλεία και το γνωστικό μοντέλο κατανόησης κειμένου, το οποίο θα αποτελέσει και την διαφοροποίηση του σημασιολογικού περιεχομένου το οποίο θα πρέπει να ενσωματώσουν στα ήδη υπάρχοντα λεξικά. Στη συνέχεια τους προτάθηκαν συγκεκριμένα κεφάλαια από τα βιβλία Πληροφορικής «Η επιστήμη των υπολογιστών: Μια ολοκληρωμένη παρουσίαση» (Brookshear, 2006) και «Εισαγωγή στην επιστήμη των υπολογιστών» (Fogouzan, 2003) ως βασική πηγή άντλησης όρων της Επιστήμης των Υπολογιστών με σκοπό τον εμπλουτισμό του ελληνικού WordNet με νέες έννοιες που αφορούν στις «Βάσεις Δεδομένων» και τα «Δίκτυα Υπολογιστών». Οι φοιτητές πραγματοποίησαν το έργο που τους ανατέθηκε σε τρία στάδια.

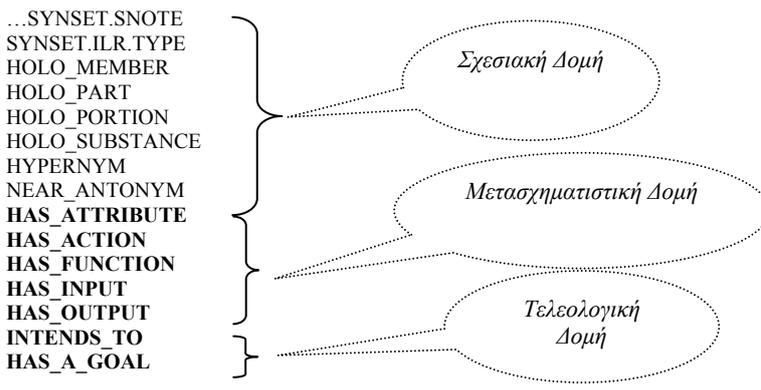
Στο πρώτο στάδιο, αναγνώρισαν και κατανόησαν τα κείμενα εκείνα που αφορούσαν τα συγκεκριμένα γνωστικά αντικείμενα του τομέα της Επιστήμης Υπολογιστών. Μέσα στα κείμενα εντόπισαν τους όρους, τους ορισμούς και τις σχέσεις που πρέπει να χρησιμοποιηθούν κατά τον εμπλουτισμό του σημασιολογικού λεξικού. Κατά τη διάρκεια αναζήτησης των όρων-ορισμών μέσα στα κείμενα, προέκυψε η ανάγκη εύρεσης νέων όρων που δεν εμφανίζονται σ' αυτά. Για το λόγο αυτό χρησιμοποιήθηκε το «Λεξικό της Κοινής Ελληνικής Γλώσσας» που εκδίδεται από το Ινστιτούτο Νεοελληνικών Σπουδών του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης, ως πρόσθετη πηγή, προκειμένου να αντληθούν οι όροι που δεν βρέθηκαν και να καλυφθούν τα "σημασιολογικά κενά".

Στο δεύτερο στάδιο, πρόσθεσαν χειρονακτικά τα νέα και αναγκαία είδη σημασιολογικών σχέσεων με βάση το μοντέλο κατανόησης κειμένου. Σύμφωνα με το μοντέλο, το άτομο που διαβάζει ένα κείμενο, κατασκευάζει τη μικροδομή και τη μακροδομή του. Ειδικότερα σε ένα τεχνικό επιστημονικό κείμενο, εμφανίζονται οι πιο κάτω γνωστικές κατηγορίες αντικειμένων: (1) *Άτομο*: Ο όρος χρησιμοποιείται για τις οντότητες που συμμετέχουν στην αναπαράσταση γνώσης. (2) *Κατάσταση*: Ο όρος περιγράφει μια στατική κατάσταση όπου δεν υπάρχει μεταβολή κατά τη διάρκεια μιας χρονικής περιόδου. (3) *Γεγονός*: Ο όρος περιγράφει μια δράση (όχι από άνθρωπο) που προκαλεί αλλαγές. Ένα τεχνικό κείμενο που απεικονίζει ένα τεχνικό σύστημα, το οποίο περιλαμβάνει ένα σύνολο συνδεδεμένων

ατόμων, αναφέρεται σε μονάδες του συστήματος, οι οποίες, με τη σειρά τους, είναι σύνολα οντοτήτων με συγκεκριμένη λειτουργία. Συγκεκριμένα, κατά την ανάγνωση οικοδομούνται δύο βασικές δομές: *Μικροδομή* και *Μακροδομή*. Η Μικροδομή περιλαμβάνει δύο υπο-δομές: Τη *Σχεσιακή Δομή* (Relational Structure - RS), η οποία αποτελεί την οντολογία των μονάδων που χαρακτηρίζονται από τις τιμές των χαρακτηριστικών τους, και των στατικών σχέσεων μεταξύ τους, και τη *Μετασχηματιστική Δομή* (Transformational Structure - TrS), η οποία αποτελεί την περιγραφή της ακολουθίας των γεγονότων, που εκτελούνται από τις μονάδες και των αιτιακών/χρονικών σχέσεων μεταξύ τους. Η Μακροδομή, εκτός από τη μικροδομή, περιλαμβάνει την *Τελεολογική Δομή* (Teleological Structure - TeS) των στόχων και υποστόχων των λειτουργιών των μονάδων του συστήματος. Η αρχική κατάσταση ενός τεχνικού συστήματος μεταβάλλεται, ώστε να επιτευχθεί μια τελική κατάσταση και να ικανοποιηθεί ένας προκαθορισμένος στόχος. Κάθε μονάδα που ενσωματώνεται στο σύστημα έχει αναπτυχθεί προκειμένου να ικανοποιήσει το δικό της υποσκοπό. Η προσθήκη πραγματοποιήθηκε με την βοήθεια του VisDic Editor. Πιο αναλυτικά, δημιουργήθηκε το XML αρχείο που περιέχει τις τελικές σχέσεις εμπλουτισμού. Δημιουργήθηκε καινούριος gre_db κατάλογος μέσω του VisDic. Αντιγράφηκαν στον νέο κατάλογο τα αρχεία του καταλόγου που περιέχονται στο ελληνικό WordNet, καθώς και το εμπλουτισμένο .inf αρχείο, το οποίο περιέχει τα νέα σύνολα συνωνύμων. Νέες εγγραφές προστέθηκαν με το όνομα του κάθε νέου τύπου σχέσης. Εδώ πρέπει να σημειωθεί ότι οι ήδη υπάρχουσες σχέσεις του WordNet αφορούν στη σχεσιακή δομή του μοντέλου. Παραδείγματα τέτοιων τύπων σχέσεων σχετικών με το μοντέλο εμφανίζονται στον Πίνακα 1, καθώς και ενδεικτικό εμπλουτισμένο αρχείο σχέσεων στο Σχήμα 2. Οι σχέσεις "has_function" και "has_action", καθώς και οι σχέσεις "has_a_goal" και "intends_to" θεωρούνται συνώνυμες.

Πίνακας 1: Παραδείγματα νέων τύπων σχέσεων εμπλουτισμού

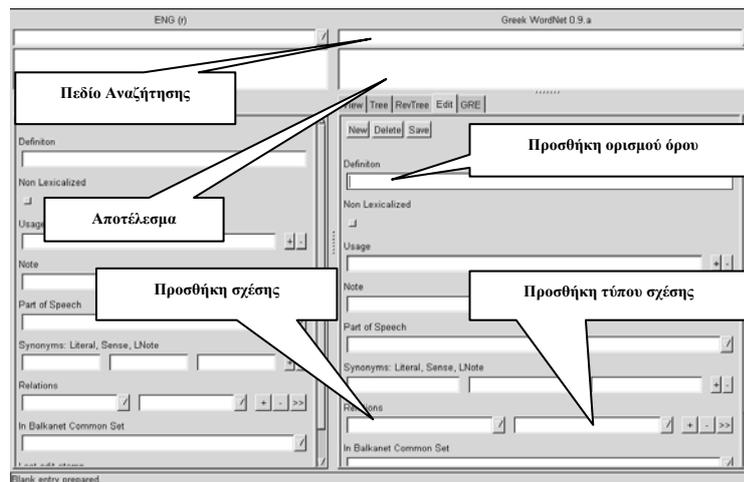
Δομή	Τύπος σχέσης	Παράδειγμα
RS	has_attribute (έχει ιδιότητα)	X has_attribute Y π.χ. κύρια μνήμη has_attribute ταχύτητα
TrS	has_function (έχει λειτουργία)	X has_function ή has_action Y π.χ. κύρια μνήμη has_function αποθήκευση
	has_action (έχει δράση)	
	has_input (έχει είσοδο) has_output (έχει έξοδο)	X as_input ή has_output Y π.χ. πρόσθεση has_input αριθμός
TeS	has_a_goal (έχει σκοπό)	X has_a_goal ή intends_to Y
	intends_to (σκοπεύει να)	π.χ. κύρια μνήμη has_a_goal προσωρινή αποθήκευση.



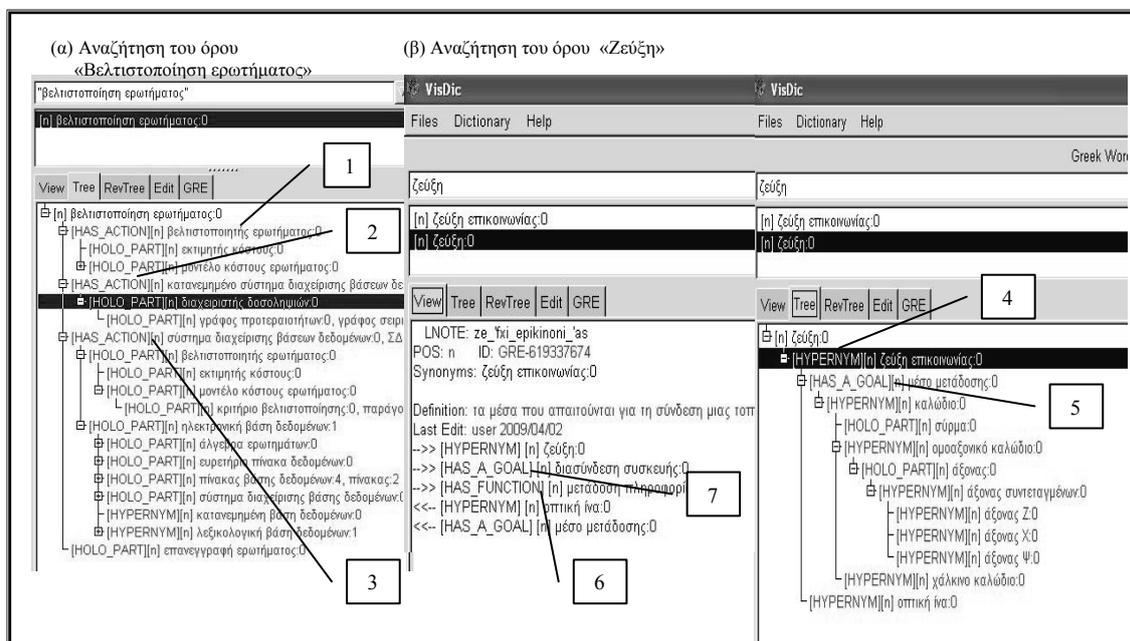
Σχήμα 2: Παράδειγμα εγγραφής των νέων σχέσεων στο .inf αρχείο

Στο τρίτο στάδιο, οι φοιτητές εμπλούτισαν χειρονακτικά το ελληνικό WordNet με νέους όρους σχετικά με τα γνωστικά αντικείμενα. Κατέγραψαν τις λέξεις/όρους που δεν υπήρχαν στο ελληνικό WordNet μέσω της διεπαφής του VisDic που φαίνεται στο Σχήμα 3. Για κάθε λεξικό υπάρχει ένα ανοικτό πλαίσιο, το οποίο περιέχει κατάλογο όρων που εμφανίζονται μετά την αναζήτηση. Το όνομα του λεξικού εμφανίζεται στην κορυφή του κάθε πλαισίου. Στο πεδίο της αναζήτησης όρου στο ελληνικό WordNet, προστέθηκαν οι νέοι όροι επιλέγοντας το κουμπί "New", και συνδέθηκαν με άλλους όρους. Στο πεδίο "Definition", προστέθηκαν οι ορισμοί των νέων όρων. Στο πεδίο "Part of

speech", καταγράφηκε το μέρος του λόγου των νέων όρων (π.χ. ρήμα, επίθετο). Τέλος, στο πεδίο "Relations", προστέθηκαν οι σχέσεις μεταξύ του νεοεισαχθέντα όρου με άλλους όρους.



Σχήμα 3: Προσθήκη νέων όρων μέσω του VisDic Editor



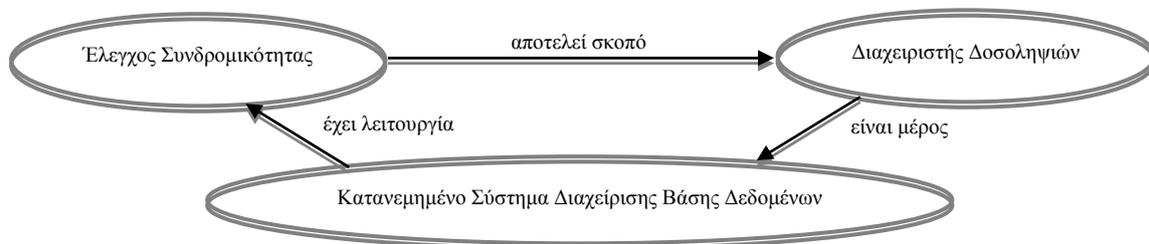
Σχήμα 4: Παραδείγματα αναζήτησης όρων στα πεδία των Βάσεων Δεδομένων και Δικτύων Υπολογιστών

Στο Σχήμα 4α φαίνεται το αποτέλεσμα της αναζήτησης του όρου "Βελτιστοποίηση ερωτήματος" του γνωστικού αντικείμενου των Βάσεων Δεδομένων στην καρτέλα "Δέντρο-Tree" του VisDic Editor. Το παράδειγμα 1 περιγράφει τη σχέση μετασχηματιστικής δομής "βελτιστοποιητής ερωτήματος-has_action-βελτιστοποίηση ερωτήματος". Το παράδειγμα 2 περιγράφει τη σχέση μετασχηματιστικής δομής "κατανεμημένο σύστημα διαχείρισης δεδομένων-has_action- βελτιστοποίηση ερωτήματος". Το παράδειγμα 3 περιγράφει την μερωνυμική σχέση σχεσιακής δομής "ηλεκτρονική βάση δεδομένων-holopart-κατανεμημένο σύστημα δεδομένων διαχείρισης". Αντίστοιχα, στο Σχήμα 4β φαίνεται το αποτέλεσμα της αναζήτησης του όρου "ζεύξη" του γνωστικού αντικείμενου των Δικτύων Υπολογιστών, μέσω της καρτέλας "Προβολή-View" και της καρτέλας "Δέντρο-Tree" του VisDic Editor. Το παράδειγμα 4 περιγράφει την υπερωνυμική σχέση σχεσιακής δομής "ζεύξη επικοινωνίας-hypernym-ζεύξη". Το παράδειγμα 5 περιγράφει τη σχέση τελεολογικής δομής "μέσο μετάδοσης-has_a_goal-ζεύξη επικοινωνίας". Το παράδειγμα 6 περιγράφει τη σχέση μετασχηματιστικής δομής "ζεύξη επικοινωνίας-has_function-μετάδοση πληροφορίας". Το παράδειγμα 7 περιγράφει τη σχέση τελεολογικής δομής "διασύνδεση συσκευής-has_a_goal- ζεύξη". Στο πεδίο των Δικτύων H/Y το ελληνικό Wordnet εμπλουτίστηκε με εβδομήντα (70) νέα synsets και τροποποιήθηκαν εξήντα (60) από

τα ήδη υπάρχοντα, ενώ για το πεδίο των Βάσεων Δεδομένων οι φοιτητές εμπλούτισαν τα ήδη υπάρχοντα synsets με πενήντα (50) καινούρια και τροποποίησαν άλλα δεκαεννιά (19).

3. Συζήτηση

Οι φοιτητές όχι μόνο αξιοποίησαν με τον καλύτερο δυνατό τρόπο το υλικό που τους παρασχέθηκε μέσω των κειμένων αλλά πρότειναν διάκριση επιπλέον σημασιολογικών σχέσεων, και βελτιώσεις στις λειτουργίες και τη χρηστικότητα των εργαλείων όπως (1) τον καθορισμό περισσότερων κατηγοριών σχέσεων υπό την έννοια της "μερωνυμίας", για παράδειγμα, παρατηρήθηκε διαφορά μεταξύ της κληρονομικότητας χαρακτηριστικών για μία έννοια X, η οποία αποτελεί υπώνυμο μιας άλλης έννοιας Y, από την περίπτωση που η έννοια X προκαλεί μία μεταβολή της συμπεριφοράς της Y, (2) μια σαφέστερη διάκριση μεταξύ των όρων που εκφράζουν αφηρημένες και συγκεκριμένες έννοιες, και τέλος (3) μεγαλύτερη φιλικότητα και προσαρμοστικότητα του εργαλείου επέκτασης και εμπλουτισμού του ελληνικού WordNet, ιδίως όσον αφορά στον τομέα της Επιστήμης των Υπολογιστών και το συγκεκριμένο γνωστικό μοντέλο κατανόησης κειμένου. Επιπλέον κατά τη διάρκεια επέκτασης του ελληνικού WordNet με νέες σχέσεις, οι φοιτητές παρατήρησαν κυκλικές διαδρομές στον νοητό γράφο των σχέσεων μεταξύ των όρων. Παράδειγμα αυτού του είδους "σημασιολογικής κυκλικής αναφοράς", η οποία μπορεί να θεωρηθεί και ως σημασιολογικό ανάλογο του φαινομένου του "κυκλικού ορισμού" (Namjoshi & Kurshan, 1999) φαίνεται στο Σχήμα 5.



Σχήμα 5: Παραδείγματα σημασιολογικής κυκλικής αναφοράς

Σε αυτό το παράδειγμα, δεν είναι σαφές εάν ο "έλεγχος συνδρομικότητας" είναι ο τελικός σκοπός του "Κατανεμημένου Συστήματος Διαχείρισης Βάσης Δεδομένων", λειτουργία του ίδιου λογισμικού ή, σε περίπτωση που είναι σημασιολογικά έγκυρο και των δύο προηγούμενων. Στην πραγματικότητα και σε πολλές περιπτώσεις, η σημασιολογική ανάλυση είναι ένα πολύ δύσκολο έργο, ιδίως στον χώρο της Επιστήμης Υπολογιστών, διότι είναι απαραίτητο να εξετάσουμε όλες τις σημασιολογικές σχέσεις μεταξύ των εννοιών που εκφράζονται από το λεξικό ορισμών.

Βιβλιογραφία

- Baudet, S. & Denhière, G. (1992). *Lecture, compréhension de texte et science cognitive*. Presses Universitaires de France, Paris.
- Brookshear, J. G. (2005). *Η επιστήμη των υπολογιστών: Μια ολοκληρωμένη παρουσίαση*. Εκδόσεις Κλειδάριθμος. ISBN 960-209-890-2.
- Forouzan, A. B. (2003). *Εισαγωγή στην επιστήμη των υπολογιστών*. Εκδόσεις Κλειδάριθμος. ISBN 960-209-707-8.
- Gruber, T. (1993). *Toward Principles for the Design of Ontologies Used for Knowledge Sharing*. Technical Report KSL 93-04. Stanford University, Knowledge Systems Laboratory. Revision.
- Horák, A. & Smrž, P. (2004). VisDic - Wordnet Browsing and Editing Tool. In *Proceedings of the 2nd International Wordnet Conference - GWC 2004*. Brno, Czech Republic: Masaryk University, 2003, pp. 136-141, ISBN 80-210-3302-9.
- Kremizis, A., Konstantinidi, I., Papadaki, M., Keramidas, G. & Grigoriadou, M. (2007). Greek WordNet extension in the domain of Psychology and Computer Science. In *Proceedings of the 8th Hellenic European Research Computer Mathematics and its Applications Conference (HERCMA 2007)*, Economical University, Athens, Greece, available in <http://www.aueb.gr/pympe/hercma/proceedings2007/>.
- Namjoshi, K. & Kurshan, R. (1999). Efficient Analysis of Cyclic Definitions. In *Proceedings of the 11th International Conference on Computer Aided Verification*, pp. 394-405. ISBN: 3-540-66202-2.