

ΘΕΜΑΤΑ ΔΙΠΛΩΜΑΤΙΚΩΝ ΕΡΓΑΣΙΩΝ
Εργ. Συστημάτων Βάσεων Γνώσεων & Δεδομένων

Database supported Reasoning System (DBRS): ΣΥΣΤΗΜΑ ΑΠΟΘΗΚΕΥΣΗΣ ΚΑΙ ΔΙΑΧΕΙΡΙΣΗΣ ΜΕΓΑΛΟΥ ΟΓΚΟΥ ΣΗΜΑΣΙΟΛΟΓΙΚΩΝ ΣΧΗΜΑΤΩΝ

ΠΛΗΡΟΦΟΡΙΕΣ: Γ. Λιαγούρης, 210 772 1436, liagos@dblab.ece.ntua.gr

ΠΕΡΙΛΗΨΗ: Η διπλωματική εργασία στοχεύει στη βελτίωση και επέκταση του συστήματος DBRS που συνδυάζει ένα Σχεσιακό Σύστημα Διαχείρισης Βάσεων Δεδομένων (RDBMS) και μια Μηχανή Συλλογιστικής Ανάλυσης (Reasoner), προκειμένου να αποθηκεύει και να διαχειρίζεται οντολογίες (ontologies) μεγάλου όγκου και υψηλής εκφραστικότητας (expressivity).

ΑΤΟΜΑ: 1-2

ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ: Java/RDFS/OWL/SQL/PostgreSQL

ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ: Ο όρος **Σημασιολογικός Ιστός (Semantic Web)** αναφέρεται στο όραμα εμπλουτισμού του υπάρχοντος **Συντακτικού Ιστού (Syntactic Web)** με **σημασιολογική πληροφορία (semantics)**, δηλαδή με πληροφορία που περιγράφει τα ίδια τα δεδομένα που υπάρχουν, οργανώνονται στο διαδίκτυο, παρουσιάζονται και μεταφέρονται μέσω αυτού. Η σημασιολογική πληροφορία, ή αλλιώς **μεταδεδομένα (metadata)**, στοχεύει στο να καταστήσει τους πόρους του διαδικτύου (Web resources) προσπελάσιμους από αυτοματοποιημένες διαδικασίες, δηλαδή από αλγοριθμικές διαδικασίες που δεν απαιτούν καθόλου ή τουλάχιστον περιορίζουν σε ένα βαθμό την “ανθρώπινη παρέμβαση”.

Κλειδί στην προσπάθεια υλοποίησης των στόχων του Σημασιολογικού Ιστού είναι η κατασκευή και διαχείριση **οντολογιών (ontologies)**, οι οποίες αποτελούν το λεξιλόγιο περιγραφής των μεταδεδομένων. Βασική προϋπόθεση για τη χρήση οντολογιών στα πλαίσια του Σημασιολογικού Ιστού είναι η εξασφάλιση της **λογικής τους συνέπειας (logical consistency)**, αλλά και η δυνατότητα αποδοτικής **εξαγωγής υπονοούμενης (implicit)** πληροφορίας από αυτές μέσω μιας λογικής διαδικασίας. Οι δύο αυτές λειτουργίες επιτελούνται μέχρι στιγμής από τις **Μηχανές Συλλογιστικής Ανάλυσης (Reasoners)**, που όμως λειτουργούν αποκλειστικά στην κύρια μνήμη (main memory). Ωστόσο, στις πραγματικές εφαρμογές (real-world applications), ο όγκος και η πολυπλοκότητα των οντολογιών τείνουν να αυξάνονται. Αποτέλεσμα αυτής της αύξησης είναι η αδυναμία ανάλυσης των οντολογιών, λόγω πολυπλοκότητας που συνεπάγεται υψηλό κόστος σε κατανάλωση κύριας μνήμης, αλλά ακόμα και η αδυναμία, σε ορισμένες περιπτώσεις, απλής φόρτωσής τους (loading) σε αυτήν λόγω όγκου.

Στο εργαστήριο έχει ήδη αναπτυχθεί το σύστημα DBRS, το οποίο συνδυάζει ένα Σχεσιακό Σύστημα Διαχείρισης Βάσεων Δεδομένων (RDBMS) και μια Μηχανή Συλλογιστικής Ανάλυσης. Η Βάση Δεδομένων του DBRS οργανώνει και αποθηκεύει τη σημασιολογική πληροφορία σε δευτερεύουσα μνήμη (σκληρό δίσκο), ενώ η Μηχανή Συλλογιστικής Ανάλυσης να είναι απαραίτητη για την εκτέλεση των **επαγωγικών διαδικασιών (inference procedures)** που προϋποθέτει η ολοκληρωμένη διαχείριση σημασιολογικών σχημάτων υψηλής εκφραστικότητας.

Στόχος της προτεινόμενης διπλωματικής είναι η βελτίωση και επέκταση του συστήματος DBRS, καθώς και η σύγκρισή του με άλλα υπάρχοντα συστήματα παρόμοιας “φιλοσοφίας”. Επομένως, η διπλωματική θα ασχοληθεί με τα παρακάτω θέματα:

ΘΕΜΑΤΑ ΔΙΠΛΩΜΑΤΙΚΩΝ ΕΡΓΑΣΙΩΝ
Εργ. Συστημάτων Βάσεων Γνώσεων & Δεδομένων

1. Μελέτη και εξοικείωση με το σύστημα DBRS που έχει αναπτυχθεί στο εργαστήριο.
2. Μελέτη και εξοικείωση με γλώσσες περιγραφής σημασιολογικών σχημάτων, όπως οι RDFS και OWL.
3. Μελέτη και εξοικείωση με φορμαλισμούς αναπαράστασης γνώσης (Knowledge Representation), καθώς και με αλγορίθμους συλλογιστικής ανάλυσης (Reasoning). Έμφαση θα δοθεί στην Περιγραφική Λογική (Description Logic -DL) και στην Tableau διαδικασία.
4. Υλοποίηση τεχνικών εξαγωγής υπονοούμενης γνώσης με χρήση Τεχνολογίας Βάσεων Δεδομένων (triggers, stored procedures, procedural languages) στα πλαίσια του DBRS.
5. Μελέτη τύπων ερωτήσεων σε σημασιολογική πληροφορία και βελτίωση των μεθόδων αποτίμησής τους από το DBRS.
6. Μελέτη υπαρχόντων συστημάτων οργάνωσης και αποθήκευσης σημασιολογικών σχημάτων (π.χ. Jena, KAON2), με σκοπό την αξιολόγηση της απόδοσής τους και τη σύγκριση με το DBRS.

ΑΝΑΦΟΡΕΣ:

DBRS: <http://www.dbnet.ece.ntua.gr/pubs/uploads/DIPL-2008-9.pdf>

OWL: <http://www.w3.org/TR/owl2-profiles/>

RDFS: <http://www.w3.org/TR/rdf-schema/>

Jena: <http://jena.sourceforge.net/>

KAON2: <http://kaon2.semanticweb.org/>