



Θεματικές Διπλωματικών Εργασιών

Τρουμπούκης Αντώνης

A. Αναπαράσταση και επεξεργασία προτιμήσεων στον Σημασιολογικό Ιστό

Συνοπτική περιγραφή:

Οι προτιμήσεις παίζουν πολύ σημαντικό ρόλο στη λήψη αποφάσεων τόσο σε πολύπλοκα προβλήματα βελτιστοποίησης, όσο και σε απλά προβλήματα της καθημερινής ζωής. Ως εκ τούτου, δεν αποτελεί έκπληξη το ότι οι προτιμήσεις έχουν μελετηθεί σε πολλά επιστημονικά πεδία (φιλοσοφία, οικονομικά, ψυχολογία, κ.ο.κ.). Βασικός στόχος της μελέτης των προτιμήσεων στην Πληροφορική είναι η δημιουργία φορμαλισμών και συστημάτων που θα δίνουν τη δυνατότητα στους χρήστες αφενός να αναπαριστούν με σαφήνεια και εκφραστικότητα τις προτιμήσεις τους, και αφετέρου να θέτουν πιο αποτελεσματικά ερωτήματα και να λαμβάνουν πιο περιεκτικές απαντήσεις, οι οποίες θα ανταποκρίνονται καλύτερα στα πραγματικά τους ενδιαφέροντα. Ειδικό ενδιαφέρον¹ παρουσιάζει η χρήση των προτιμήσεων στην περιοχή του Σημασιολογικού Ιστού (Semantic Web), όπου η ελεύθερη και κατανεμημένη διασύνδεση των διασυνδεδεμένων δεδομένων (Linked Data) παρουσιάζει περισσότερες προκλήσεις ως προς την προσπέλαση χρήσιμης πληροφορίας από ότι σε παραδοσιακά συστήματα βάσεων δεδομένων: η μορφή των δεδομένων στον Σημασιολογικό Ιστό είναι γενικά ετερογενής και οι πηγές συνήθως διαφέρουν ως προς την αξιοπιστία, την ποιότητα και τη συνάφεια ως προς την επιθυμία του χρήστη. Σκοπός της εργασίας είναι η ανάπτυξη μεθόδων και τεχνολογιών για την μοντελοποίηση και επεξεργασία προτιμήσεων στην περιοχή του Σημασιολογικού Ιστού. Ενδεικτικές κατευθύνσεις στις οποίες μπορεί να εστιάσει το θέμα είναι: η βελτιστοποίηση ήδη υπάρχοντων συστημάτων επερωτήσεων (όπως η γλώσσα

¹ O. Pivert, O. Slama, V. Thion: SPARQL extensions with preferences: a survey. SAC 2016: 1015-1020

επερωτήσεων SPREFQL²), η ανάπτυξη νέων γλωσσών επερωτήσεων με προτιμήσεις, ή η ανάπτυξη νέων οντολογιών για την μοντελοποίηση προτιμήσεων πάνω σε διασυνδεδεμένα δεδομένα.

B. Αναπαράσταση προτιμήσεων μεταξύ συνόλων με λογικούς φορμαλισμούς

Συνοπτική περιγραφή:

Οι προτιμήσεις πάνω σε σύνολα αντικειμένων είναι μια σημαντική κατηγορία προτιμήσεων που εμφανίζεται αρκετά συχνά σε πρακτικές εφαρμογές. Για παράδειγμα, «θέλω να αγοράσω ένα σύνολο από τρία βιβλία, από τα οποία το ένα προτιμώ να είναι αστυνομικό μυθιστόρημα, και η συνολική τιμή των τριών βιβλίων να είναι η ελάχιστη δυνατή» ή «θέλω να νοικιάσω ένα σύνολο από πέντε ταινίες, αλλά θα προτιμούσα ολοκληρωμένες σειρές ταινιών». Ενώ οι προτιμήσεις μεταξύ αντικειμένων έχουν μελετηθεί αρκετά στην βιβλιογραφία του λογικού προγραμματισμού³, δεν έχει γίνει το ίδιο και για τις προτιμήσεις μεταξύ συνόλων. Ένας αρκετά κομψός και εκφραστικός τρόπος για την μοντελοποίηση προτιμήσεων μεταξύ συνόλων είναι η αναπαράσταση τους με την χρήση του λογικού προγραμματισμού υψηλής τάξης⁴, αν και η προσέγγιση αυτή παρουσιάζει μειονεκτήματα στον χρόνο εκτέλεσης των προγραμμάτων, καθώς η αναζήτηση για τα προτιμότερα σύνολα γίνεται στην ουσία με generate-and-test τρόπο. Σκοπος της εργασίας είναι η ανάπτυξη μιας νέας, αποδοτικής, προσέγγισης για την αναπαράσταση προτιμήσεων μεταξύ συνόλων, καθώς και η σύγκριση της προσέγγισης αυτής με την αντίστοιχη του λογικού προγραμματισμού υψηλής τάξης, τόσο στο από εκφραστικής όσο και πρακτικής άποψης. Μια πιθανή κατεύθυνση είναι η χρήση κάποιου συστήματος λογικού προγραμματισμού με περιορισμούς (όπως π.χ. το σύστημα ECLiPSe⁵), καθώς τα συστήματα αυτά προσφέρουν την εναλλακτική μέθοδο constraint-and-generate, όπου η χρήση περιορισμών αποκόπτει κόμβους από το δέντρο αναζήτησης και συνεπώς μειώνει κατα πολύ τον χρόνο εκτέλεσης του προγράμματος.

² A. Troumpoukis, S. Konstantopoulos, A. Charalambidis: An Extension of SPARQL for Expressing Qualitative Preferences. International Semantic Web Conference 2017: 711-727

³ J. Delgrande, T. Schaub, H. Tompits, K. Wang: A Classification and Survey of Preference Handling Approaches in Nonmonotonic Reasoning. Comput. Intell. 20(2): 308-334 (2004)

⁴ A. Charalambidis, P. Rondogiannis, A. Troumpoukis: Higher-order logic programming: An expressive language for representing qualitative preferences. Sci. Comput. Program. 155: 173-197 (2018)

⁵ The ECLiPSe Constraint Programming System <http://www.eclipseclp.org/>

Γ. Τεχνικές μετασχηματισμού γεωχωρικών επερωτήσεων σε κατανεμημένες βάσεις στον Σημασιολογικό Ιστό

Συνοπτική περιγραφή:

Τα γεωχωρικά διασυνδεδεμένα δεδομένα (Geospatial Linked Data)⁶ είναι ένας αναδύομενος τομέας του Σημασιολογικού Ιστού με αυξανόμενο ενδιαφέρον για την έρευνα και τη βιομηχανία. Τα δεδομένα αυτά αναπαριστούν αντικείμενα, γεγονότα ή άλλες οντότητες σε συνδυασμό με την τοποθεσία τους πάνω ή κοντά στην επιφάνεια της Γης. Υπάρχουν, ωστόσο, διάφορες τεχνολογίες Σημασιολογικού Ιστού όπου δεν έχουν φτάσει ακόμα σε ιδιαίτερα υψηλό επίπεδο ωριμότητας, όπως για παράδειγμα η τεχνολογία των federated συστημάτων επερωτήσεων, τα οποία συνδυάζουν δεδομένα από πολλές απομακρυσμένες γεωχωρικές βάσεις. Σε αυτό το πλαίσιο, πρόσφατα⁷, επιβεβαιώθηκε πειραματικά ότι ένας απλός μετασχηματισμός μπορεί να επιταχύνει την διαδικασία του υπολογισμού μέχρι και 2 τάξεις μεγέθους σε επερωτήσεις που αφορούν ζευγάρια οντοτήτων τα οποία έρχονται από διαφορετικές βάσεις και οι γεωμετρίες τους βρίσκονται εντός μίας καθορισμένης απόστασης μεταξύ τους. Σκοπός της εργασίας είναι η ανάπτυξη παρόμοιων τεχνικών μετασχηματισμού και για άλλες πράξεις που συνδυάζουν δεδομένα από απομακρυσμένες γεωχωρικές βάσεις. Για το πειραματικό μέρος, υπάρχει η δυνατότητα αξιοποίησης υπάρχουσας υποδομής του εργαστηρίου για federated query processing⁸.

⁶ https://www.w3.org/community/geosemweb/wiki/Main_Page

⁷ A. Troumpoukis, S. Konstantopoulos, N. Prokopaki-Kostopoulou, A Geospatial Join Optimization for Federated GeoSPARQL Querying, GeoLD2022.

⁸ Semagrow, <http://github.com/semagrow>