

## ΠΡΟΣ

- 1) Όλα τα μέλη ΔΕΠ του Τμήματος Επιστήμης Υπολογιστών
- 2) Τους εκπροσώπους των Μεταπτυχιακών φοιτητών του Τμήματος Επιστήμης Υπολογιστών
- 3) Την Επταμελή Εξεταστική Επιτροπή
- 4) Όλα τα μέλη της Πανεπιστημιακής Κοινότητας

**Πρόσκληση σε Δημόσια Παρουσίαση της Διδακτορικής Διατριβής της  
κα. Ζωγραφιστού Δήμητρας**

**Doctoral Dissertation Defense**

**Mrs. Dimitra Zografistou**

Την Τρίτη, 26/11/2019 και ώρα 13:00 στην αίθουσα Τηλεδιάσκεψης Κ206 του Τμήματος Επιστήμης Υπολογιστών του Πανεπιστημίου Κρήτης στο Ηράκλειο, θα γίνει η δημόσια παρουσίαση και υποστήριξη της Διδακτορικής Διατριβής της υποψηφίας διδάκτορος του Τμήματος Επιστήμης Υπολογιστών κα. Ζωγραφιστού Δήμητρας με θέμα:

***“ArgQL: Μία τυπική και δομημένη γλώσσα υποβολής επερωτήσεων σε επιχειρηματολογικούς διαλόγους”***

***“ArgQL: Querying Argumentative Dialogues using a Formal, Structured Language”***

## ΠΕΡΙΛΗΨΗ

Μπροστά στην ραγδαία εξέλιξη των κοινωνικών δικτύων που βιώνουμε την τελευταία δεκαετία και της αδιαμφισβήτητης επικράτησής τους ανάμεσα στα διαφορετικά μέσα επικοινωνίας, το πρόβλημα της ανάκτησης ωφέλιμης πληροφορίας από διαλογικά δεδομένα είναι τώρα πιο κρίσιμο από ποτέ. Μέχρι πρόσφατα, οι προκλήσεις που συνοδεύονταν με το συγκεκριμένο πρόβλημα αντιμετωπιζόνταν ως ένα επιπλέον πεδίο εφαρμογής των υπολογιστικών μοντέλων που έχουν αναπτυχθεί σε τομείς όπως αυτόν του Σημασιολογικού Ιστού, της Ανάκτησης Πληροφορίας της Εξόρυξης Δεδομένων κλπ. Παρόλα αυτά, υπάρχουν κάποιες πληροφοριακές ανάγκες που σχετίζονται με το πρόβλημα της αναζήτησης

πληροφορίας σε διαλόγους, οι οποίες είναι κοινές για όλους τους τύπους διαλόγων, ανεξαρτήτως του ευρύτερου πλαισίου τους ή του στόχου που μπορεί να έχουν, πχ. που αφορούν τη δομή των διαφορετικών απόψεων και των πιθανών συσχετίσεων μεταξύ τους, οι οποίες θα μπορούσαν να σταθούν σαν μία ανεξάρτητη περιοχή προς μελέτη. Η απομόνωση αυτών των αναγκών και η ένταξή τους στον προσδιορισμό μίας τυπικής γλώσσας, η οποία θα έχει σχεδιαστεί αποκλειστικά γι' αυτό το σκοπό, αποτελεί μία ερευνητική κατεύθυνση, η οποία έχει μέχρι τώρα μελετηθεί ελάχιστα.

Παρακινούμενοι από αυτή την έλλειψη, σε αυτή την εργασία παρουσιάζουμε την ArgQL (Γλώσσα Επερωτήσεων Επιχειρηματολογίας), μία υψηλού επιπέδου, δηλωτική γλώσσα για την υποβολή ερωτημάτων σε διαλογικά δεδομένα. Η ArgQL προσφέρει μία απλή και σχετιζόμενη με διαλόγους ορολογία για την σύνταξη ερωτημάτων στο συγκεκριμένο τομέα, τα οποία, με τις υπάρχουσες γλώσσες επερωτήσεων, θα ήταν αρκετά δύσκολο να εκφραστούν. Η θεωρία που θεμελιώνει το μοντέλο δεδομένων υιοθετεί κάποια από τα επικρατέστερα σημασιολογικά μοντέλα, που έχουν οριστεί στην περιοχή της Υπολογιστικής Επιχειρηματολογίας. Συνεπώς, τα δεδομένα στα οποία απευθύνεται η γλώσσα αποτελούνται από γράφους διασυνδεδεμένων και δομημένων επιχειρημάτων και η ArgQL δίνει τη δυνατότητα της πλοήγησης σε τέτοιους γράφους. Παρουσιάζουμε λοιπόν τις τυπικές προδιαγραφές τις γλώσσας, οι οποίες περιλαμβάνουν τον ορισμό των βασικών της δομών, το συντακτικό της καθώς επίσης και το σημασιολογικό μοντέλο, βάσει του οποίου καθορίζεται η αποτίμηση των δομών αυτών με στοιχεία του μοντέλου δεδομένων.

Ακολούθως, προτείνουμε μία μεθοδολογία για τη μετάφραση της ArgQL σε άλλες γλώσσες, και συγκεκριμένα, δείχνουμε την περίπτωση της RDF (Γλώσσα Περιγραφής Πηγών) και της συσχετιζόμενης με αυτήν γλώσσας επερωτήσεων, SPARQL. Προς αυτή την κατεύθυνση, ορίζουμε ένα RDF σχήμα το οποίο βασίζεται στο εννοιολογικό μοντέλο της οντολογίας AIF και παρουσιάζουμε την αντιστοιχία μεταξύ αυτής της αναπαράστασης και του δικού μας μοντέλου δεδομένων. Στη συνέχεια, χτίζουμε τη διαδικασία της μετάφρασης των επερωτήσεων, ως ένα σύνολο κανόνων που ορίζουν την αντιστοιχία μεταξύ των διαφορετικών δομών της ArgQL και των αντίστοιχων μοτίβων γράφου στην SPARQL. Οι ιδιότητες της ορθότητας και πληρότητας της διαδικασίας, επαληθεύονται αποδεικνύοντας την ισοδυναμία μεταξύ των αποτελεσμάτων για καθεμία από τις δύο γλώσσες, αναφορικά και με την τυπική σημασιολογία τους. Παρά την ορθότητα των κανόνων μετάφρασης, η ακριβής εφαρμογή τους οδηγεί σε μη βέλτιστες επερωτήσεις. Γι' αυτό το λόγο, πάνω σε αυτή τη μεθοδολογία, προτείνουμε κάποιες βελτιστοποιήσεις, οι οποίες επιτυγχάνουν μικρότερα σε μήκος και κατ' επέκταση πιο αποδοτικά ερωτήματα. Αναδεικνύουμε ακόμα την πρακτική πλευρά της ArgQL υλοποιώντας τη γλώσσα και δίνοντας της τη δυνατότητα να εκτελείται σε πραγματικές συλλογές δεδομένων. Έτσι, ένα από τα αποτελέσματα αυτής της δουλειάς, είναι μία εφαρμογή του ιστού, στην οποία μπορεί οποιοσδήποτε να δοκιμάσει να τρέξει τα δικά του ερωτήματα. Τέλος, διεξάγουμε μία πειραματική μελέτη, προκειμένου να αξιολογήσουμε την απόδοση των επερωτήσεων κάτω από διαφορετικές παραμέτρους εκτέλεσης.

Επιβλέπων: Καθηγητής, Δημήτριος Πλεξουσάκης

## ABSTRACT

In light of the rapid evolution of social media that we experience over the past decade, and their establishment as one of the main means of communication and dialogical exchange, the problem of extracting meaningful information from data residing in human dialogues is now more crucial than ever. Until recently, the challenges associated with this problem were being addressed as an application domain for the computational models of fields like the Semantic Web, Information Retrieval, Data Mining etc. However, there are some information requirements when searching in dialogues which are quite specific and common for all types of dialogues, regardless of their context or goal, e.g. concerning the structure of the different opinions and the correlations among them, which could stand as an autonomous area to study. Isolating those requirements and bringing them together in the specification of a formal language, designed exclusively for this purpose, is a research direction which has been given less attention.

Incited by this deficiency, in this thesis we introduce ArgQL (Argumentation Query Language), a high-level declarative language for querying dialogical data. ArgQL provides a simple and dialogue-related terminology to write queries in the domain, which in existing query languages would be quite difficult to express. The theory that founds the data model adopts some of the most prevailing semantics in the area of Computational Argumentation. As a result, the target data consist of graphs of interconnected, structured arguments and ArgQL allows for the navigation across such graphs. We present the formal specification of the language, including the definition of its main constructs, the concrete syntax, as well as the semantics that determine the evaluation of those constructs against the data model.

Subsequently, we propose a methodology to translate ArgQL into other languages and in particular we show the case of RDF and its associated query language, SPARQL. To this end, we define an RDF scheme based on the AIF conceptualization and we formalize the mappings between this and our data model. We then build the process of query translation, as a set of rules that define the correspondence between the different ArgQL constructs and the respective SPARQL graph patterns. The soundness and completeness of the process is verified by proving the equivalence between the matching data for each of these two languages, with respect to their formal semantics. Although correct, the conformation to the precise definition of the translation rules results to non-optimal queries. Therefore, on top of this methodology, we propose some optimization, that succeeds shorter and by extension more efficient queries. We also give prominence to the practical side of ArgQL and we implement the language, allowing for its execution on real data-sets. Thus, one of the outcomes of this work was an online query endpoint, where someone can test his own queries. Finally, we conduct an experimental study to evaluate the query performance under different execution parameters.

Supervisor: Professor, Dimitris Plexousakis