|  |  |
| --- | --- |
| A close up of a logo  Description automatically generated | **A close up of a sign  Description automatically generated** |

**Προτεινόμενα Θέματα Διπλωματικής Εργασίας**

Στασινός Κωνσταντόπουλος

Α. **Μαντεύοντας τα άγνωστα περιεχόμενα απομακρυσμένων endpoints από αποσπασματικές απαντήσεις**

Στο αποκεντρωμένο περιβάλλον του Linked Open Data cloud και του Σημασιολογικού Ιστού, τα συστήματα επεξεργασίας ερωτημάτων δεν έχουν τα λεπτομερή και ενήμερα μεταδεδομένα που απολαμβάνουν οι συμβατικές βάσεις δεδομένων. Για να συνδυάσουμε την δυναμική και αποκεντρωμένη φύση του Web με την αποδοτικότητα των κεντρικοποιημένων συστημάτων επεξεργασίας, χρειαζόμαστε ένα αποδοτικό και αυτόνομο τρόπο να δημιουργούμε και να συντηρούμε μεταδεδομένα που επιτρέπουν σε ένα σύστημα επεξεργασίας να αποφασίζει ποια απομακρυσμένα endpoints μπορεί να περιέχουν σχετικά δεδομένα και ποιο είναι το βέλτιστο πλάνο εκτέλεσης. Επιπλέον, θέλουμε να το πετύχουμε αυτό χωρίς να επιβαρύνουμε τα endpoints με διερευνητικά ερωτήματα, αλλά εξάγοντας συμπεράσματα από τις αποκρίσεις στα ερωτήματα ενός παραγωγικού workload.

Η διπλωματική θα μελετήσει μεθόδους που γενικεύουν τις αποκρίσεις ενός SPARQL endpoint, και έτσι συνθέτουν μεταδεδομένα που προσεγγίζουν τις παραμέτρους που χρειάζεται ένας query optimizer: τα είδη και τα μεγέθη των δεδομένων που παρέχει το endpoint, και το σε τι βαθμό τα δεδομένα ενός endpoint ενώνονται με τα δεδομένα ενός άλλου. Η διπλωματική θα βασιστεί σε προηγούμενες σχετικές εργασίες[[1]](#footnote-1),[[2]](#footnote-2) και θα δώσει στον φοιτητή την ευκαιρία για πειραματισμό με τεχνολογίες αιχμής στην μηχανική μάθηση, την ικανοποίηση περιορισμών, και την διαχείριση κατανεμημένων δεδομένων.

Εκτός από την θεωρητική μελέτη, η διπλωματική περιλαμβάνει και πειραματική επαλήθευση της προτεινόμενης μεθόδου, αξιοποιώντας δημόσια endpoints και υπάρχον ερευνητικό πρωτότυπο για federated query processing.[[3]](#footnote-3)

Η διπλωματική απαιτεί γνώσεις προγραμματισμού σε Java. Η συγγραφή της διπλωματικής πρέπει να γίνει στην αγγλική και σε LaTeX, και πρέπει να προετοιμασθεί και υποβληθεί σχετικό επιστημονικό άρθρο.

Β. **Η αναπαράσταση γνώσης ως υπόβαθρο του διαχωρισμού αιτιακών και συμπτωματικών συσχετισμών**

Οι τεχνολογίες που σήμερα αναφέρονται ως «τεχνητή νοημοσύνη» βασίζονται στην βέλτιστη προσαρμογή μοντέλων στα δεδομένα χωρίς να εξετάζουν αν οι συσχετισμοί που προκύπτουν είναι αιτιακοί, συν-αιτιακοί, ή συμπτωματικοί. Αυτή, όμως, η προσέγγιση στερεί από την τεχνητή νοημοσύνη την δυνατότητα να εκμεταλλεύται γνώση του κόσμου για να ξεχωρίζει spurious correlations.

Η διπλωματική θα μελετήσει παραδείγματα από spurious correlations[[4]](#footnote-4) και από επιτυχημένα πειράματα στατιστικής μηχανικής μάθησης από τη βιβλιογραφία, και θα αναζητήσει σχετική δημόσια διαθέσιμη σημασιολογική γνώση του κόσμου. Στο θεωρητικό επίπεδο, θα εξερευνήσει περιορισμούς και μετα-κανόνες που προκύπτουν μηχανιστικά από την σημασιολογική γνώση και θα μπορούσαν να απορρίψουν τα spurious correlations χωρίς να απορρίπτουν τα ορθά.

Εκτός από την θεωρητική μελέτη, η διπλωματική περιλαμβάνει και πειραματική επαλήθευση με χρήση κατάλληλης μεθόδου τεχνητής νοημοσύνης που να συνδυάζει αριθμητικούς και σημασιολογικούς κανόνες. Για παράδειγμα, μπορεί να αξιοποιηθεί το constraint satisfaction για να βρεθεί η επικάλυψη (ή απουσία επικάλυψης) των στατιστικών προβλέψεων και των θεωρητικών ή σημασιολογικών περιορισμών.[[5]](#footnote-5) Η μέθοδος που θα προταθεί, θα αξιολογηθεί σε κατάλληλο πεδίο εφαρμογής. Ένα παράδειγμα είναι η κατανόηση των συσχετισμών ανάμεσα στα δεδομένα αισθητήρων από τη χημική βιομηχανία,[[6]](#footnote-6) και συγκεκριμένα σε χημικές διαδικασίες με πολλαπλά αλληλοεξαρτώμενα στάδια. Άλλο παράδειγμα είναι η ανάλυση μεταφορικής χρήσης της γλώσσας, όπου στατιστικές ή συντακτικές ενδείξεις ότι κάποια φράση είναι μεταφορική φιλτράρονται με βάση την σημασιολογική συμβατότητα της ιδιότητας που μεταφέρεται. Τρίτο παράδειγμα είναι η μηχανική όραση, και η επιβεβαίωση ότι ένα οπτικό feature είναι inherent ή συμπτωματικό.

Η διπλωματική απαιτεί γνώσεις προγραμματισμού για την προσαρμογή δεδομένων και κανόνων στο συντακτικό που απαιτεί ο solver που θα αξιοποιηθεί. Η συγγραφή της διπλωματικής πρέπει να γίνει στην αγγλική και σε LaTeX, και πρέπει να προετοιμασθεί και υποβληθεί σχετικό επιστημονικό άρθρο.

Γ. **Segmentation στερεοσκοπικού σήματος για την αναγνώριση ανθρώπων και άλλων εμποδίων**

Ένα σημαντικό κομμάτι της ρομποτικής όρασης είναι το segmentation, ο διαχωρισμός της συνολικής σκηνής στα διακριτά αντικείμενα τα οποία μετά θα αναγνωριστούν, κατηγοριοποιηθούν, και συσχετισθούν. Στο στερεοσκοπικό σήμα, το segmentation βασίζεται στις αποστάσεις μεταξύ των σημείων ή σε γεωμετρικά χαρακτηριστικά. Αυτή η προσέγγιση επωφελείται από την παρατήρηση της εξέλιξης της σκηνής στο χρόνο, καθώς τα κινούμενα αντικείμενα μπορεί να είναι πιο ευδιάκριτα σε κάποια frames και να ενισχύσουν την ακρίβεια του segmentation και σε πιο δύσκολα, δυσδιάκριτα frames. Από την άλλη, η παρατήρηση κινούμενων αντικειμένων προσθέτει κάποια αβεβαιότητα σχετικά με το data association, τον συσχετισμό των σημείων του ίδιο φυσικού αντικειμένου σε διαδοχικά frames.

Σκοπός της διπλωματικής είναι η μελέτη και προσαρμογή μεθόδων από το reinforcement learning ή παρόμοιων μεθόδων που αξιοποιούν delayed feedback, ώστε να αναπτυχθεί μέθοδος που αξιοποιεί τα ευδιάκριτα frames για βελτίωση της απόδοσης στα δυσδιάκριτα χωρίς να απαιτείται χειρωνακτική επισημείωση.

Εκτός από την θεωρητική μελέτη, η διπλωματική περιλαμβάνει και πειραματική επαλήθευση σε κατάλληλα σενάρια όπως ο διαχωρισμός και χαρακτηρισμός ανθρώπων που δημιουργούν μία συστάδα με άλλους ανθρώπους και έπιπλα, και σε επόμενη χρονική στιγμή χωρίζουν. Η διπλωματική μπορεί να βασισθεί στο σχετικό σύστημα που έχει αναπτυχθεί στο εργαστήριο[[7]](#footnote-7),[[8]](#footnote-8) και να το επεκτείνει ή μπορεί να δοκιμάσει τελείως νέα προσέγγιση στο πρόβλημα.

Η διπλωματική απαιτεί γνώσεις προγραμματισμού σε Python. Η συγγραφή της διπλωματικής πρέπει να γίνει στην αγγλική και σε LaTeX, και πρέπει να προετοιμασθεί και υποβληθεί σχετικό επιστημονικό άρθρο. Κατά την πειραματική της φάση, η διπλωματική θα απαιτήσει συχνή παρουσία στο ΕΚΕΦΕ «Δημόκριτος», εργάσιμες ημέρες και ώρες.

1. K Zamani, A Charalambidis, S Konstantopoulos, N Zoulis, E Mavroudi, Workload-Aware Self-Tuning Histograms for the Semantic Web. *Trans. Large-Scale Data and Knowledge-Centered Systems* 28, Sep 2016. <https://zenodo.org/record/159131> [↑](#footnote-ref-1)
2. Ε Κατηφόρης, *Query Optimization on Distributed Databases*. Πτυχιακή εργασία, Τμ. Πληροφορικής και Τηλεπικοινωνιών, ΕΚΠΑ, 2018, <https://pergamos.lib.uoa.gr/uoa/dl/object/2800246> [↑](#footnote-ref-2)
3. Semagrow, <https://github.com/semagrow> [↑](#footnote-ref-3)
4. <http://www.tylervigen.com/spurious-correlations> [↑](#footnote-ref-4)
5. Luc De Raedt, Tias Guns, Siegfried Nijssen, Constraint Programming for Data Mining and Machine Learning. In *Proceedings* *Twenty-Fourth AAAI Conference on Artificial Intelligence.* 2010. [↑](#footnote-ref-5)
6. Σε συνεργασία με μεγάλη χημική βιομηχανία. Σε περίπτωση επιλογής αυτού του θέματος, θα απαιτηθεί η υπογραφή non-disclosure agreement σχετικά με τα δεδομένα, αλλά η μέθοδος που θα αναπτυχθεί θα μπορεί να περιγραφεί ελεύθερα. [↑](#footnote-ref-6)
7. K Zamani, G Stavrinos, and S Konstantopoulos, Detecting and measuring human walking in laser scans. In Proceedings of the 10th Hellenic Conference on AI (SETN 2018). <https://doi.org/10.1145/3200947.3201026> [↑](#footnote-ref-7)
8. Prototype: <https://github.com/roboskel/hpr> , Demo: <https://vimeo.com/219075548> [↑](#footnote-ref-8)