Laboratoire Bordelais de Recherche en Informatique
Unité Mixte de Recherche CNRS (UMR 5800)

CORDI/S PhD project proposal

INRIA team: GRAVITÉ

Author of the proposal: Guy Melançon

Title of the proposal: Visual and Interactive Exploration of Semantic Networks

Keywords (max 5-6): visual analytics, semantic networks, visual and dynamic queries

Scientific context (10-15 lines): The PhD project takes place within the ANR TANGUY project headed by Thalès Communications and in collaboration with Xerox (XRCE Grenoble) and the law firm FIDAL. The project aims at designing a methodology to support the incremental construction and easy use of semantic networks. The development of a (prototype) system should help law experts to deal with large document collections from which potential facts have to be identified, extracted, organized and ultimately validated or annotated.

Starting from valid pieces of information extracted from text document (Xerox expertise), our goal is to offer adapted mechanisms to visualize, navigate and annotate all this information. The construction of a network of facts should follow from semantic proximity between pieces of information or from users' knowledge. As a consequence, the interactive navigation of the network can exploit available results in graph visualization.

Goal (10-15 lines): Information Visualization can be described as the process of mapping non spatial data into relevant visual forms to support analytical reasoning. The goal is to provide the user with a readable and coherent visual representation of the data, equipped with relevant interactions, which becomes a real challenge when dealing with data dynamically evolving with time.

The data we face within TANGUY can be characterized as heterogeneous, uncertain, incomplete and evolving in time. The semantic networks we must deal with is made of pieces computed through natural language analysis, together with informal pieces input by users and categorized in an ad hoc manner. The visualization should help visually query relevant information based on context, help users reason about it and at the same time enrich the network itself – readable drawings and relevant interactions must be designed and validated.

- **Project** (10-15 lines): The project basically consists of two parts. As a first task, the PhD is expected to get involved with all of the partners in order to grasp the understanding and potential uses of semantic networks in the TANGUY context. Efforts will be devoted to help final users (law experts) to specify their needs in terms of possible interactions and system functionalities. This phase requires leveraging user knowledge and know-how induced from experience or emerging from the organization, turning it into system specifications (GUI, graphics, interactions). Ths phase will most probably be accomplished through participative design.
- The second part of the project will exploit the available or to-be-invented methods and techniques from graph visualization in order to design fulfilling the users need. We can expect the system to astutely combine techniques from graph mining (analysis) together with graph drawing (graphical representation as well as interactions) adapted to the tasks conducted by law experts as identified and modelled through participative design workshops. The two parts of the project will interact and evolve through several iterations where prototypes will serve as a proof-of-concept for the project.

Skills. Graph (data) mining; graph drawing and visualisation; human-computer interaction; certified skills developing with C++. Prior knowledge of Tulip is a plus.

351, cours de la Libération
F-33405 Talence Cedex
Tél.: +33(0)5 4000 69 00
Fax: +33(0)5 4000 66 69

Laboratoire Bordelais de Recherche en Informatique
Unité Mixte de Recherche CNRS (UMR 5800)

Advisors: David Auber, Guy Melançon

References (max 5):

351, cours de la Libération
F-33405 Talence Cedex

Tél.: +33(0)5 4000 69 00
Fax: +33(0)5 4000 66 69