THALES

Internship Thalès / LRI – Université Paris-Sud

Model to filter data

Thales Research & Technology (TRT):

Dir./Dept.: Group « Sciences et Techniques de l'Information »

Service: LRASC, «Analysis and reasoning in Complex Systems » Laboratory

Contacts: Gaëlle Lortal, gaelle.lortal@thalesgroup.com, Claire Laudy, claire.laudy@thalesgroup.com

Université Paris-Sud :

Laboratory: Laboratoire de Recherche en Informatique (LRI)

Group: LaHDAK, Large-scale Heterogeneous DAta and Knowledge

Contact: Yue Ma, yue.ma@lri.fr

Duration: 6 months

Starting date: as soon as possible

Location: TRT (Palaiseau) and LRI-Université Paris-Sud (Orsay)

Description: The research group "Sciences et Techniques de l'Information" at Thales Research & Technology is composed of several labs, among which one, LRASC, is specialized in Analysis and Reasoning in Complex Systems.

One of the themes studied at LRASC is Information Fusion. Graph based information fusion has been previously studied within the lab, and due to massive information available in heterogeneous resources, filtering is now a necessary step in the fusion to get more pertinent information. A situation model is used in the fusion process to filter information coming from different sources and keep only the relevant pieces of information. This implies that a situation model is provided as an input of the system.

Expertise in the considered domain is needed to provide such a situation model. That is, knowledge engineers usually use end-users knowledge and expertise (through texts, documents and interviews) to provide it. However, when the domain is knowledge intensive, it becomes labor-intensive.

The aim of the internship is to propose a solution to support an intelligence operator to define the situation model and avoid, or minimize, to turn to knowledge engineers. End-users have difficulties to model their domain as such a model is implicit in automatic data processing. What they know are steps of their activity (high level concepts) and only parts of more or less generic models. End-users are unable to provide the whole situation model but elements of parts of the model with which they are familiar could be used to generate it.

In the use-cases considered in this internship, we are given a set of instances that should be kept after filtering by a situation model. The idea we propose is to abstract a situation model from these interesting examples as seeds. For this, a target language of situation models has to be defined in the first place. To benefit from advanced Ontology Based Data Access (OBDA) techniques, we plan to use OBDA query language as the target language, a well-known sub-class of SQL queries, called conjunctive queries (conjunctions of atoms). In this way, the implementation of the proposed approach will further allow reasoning over background knowledge and database data (encoded in

an ontology). The answers returned for a generated query will make explicit all the potential instances kept by the corresponding situation model. If a new instance kept by the situation model is unintended, it means a further revision of the situation model should be performed. In this way, end users can be assisted to valid the generated situation model.

The work realized during the internship will consist in a first exploration of the problem and may be continued for a PhD.

Mission:
The main objective of the internship is to propose an approach to discover a situation model for a knowledge intensive domain. The tasks to be accomplished are the following:
- Ontological modeling of an existing database
- Creation of an algorithm for generating situation models
- Evaluation of the generated situation model with data coming from use cases.

Funding: This Internship will be funded by the Thales Company.

Your skills:
Your profile: Information Science & Artificial Intelligence Engineering
Education Level: Master2 Research or equivalent
Mandatory skills: Logic, programming language (JAVA, Python, ...)
Other skills: Inductive Logic Programming, Ontology modeling language (OWL), query languages, reasoning mechanisms.
Skills acquired during the internship: Ontology building, Ontology-Based Data Access systems, Information Fusion.
Language: English/French

Contacts:
Claire Laudy : claire.laudy@thalesgroup.com
Gaëlle Lortal : gaelle.lortal@thalesgroup.com
Yue Ma : yue.ma@lri.fr