

PhD in System Biology and Bioinformatics

Title: Building signaling pathways by reasoning on pieces of biological knowledge

Supervised by: Pr. Christine Froidevaux (LRI Bioinformatics group CNRS and AMIB group INRIA, University Paris Sud)

Location: LRI, building 490, Univ. Paris Sud, Orsay)

Contact: Christine.Froidevaux@lri.fr

In collaboration with: Anne Poupon (INRA BIOS Tours, on methodology and biological knowledge formalization validation).

General context: The work will be done in the context of the ASAM project (INRIA/AMIB - INRA/BIOS) which aims to help the understanding of signaling pathways involving G protein-coupled receptors (excellent targets in pharmacogenomics). In this work we are only interested in the influence graph (static model) and do not take into account the dynamics of the network. Large amounts of experiments are available in this context while globally interpreting all the results remains a very challenging task for biologists. As a solution, ASAM proposes to design a Knowledge Base (KB) containing expert rules to interpret the various experimental results and semi automatically construct signaling pathways (represented as networks).

The initial knowledge base is built by the scientists from experimental data and information contained in scientific literature. It is expressed in a logical setting (first-order logic, mainly logic programming).

The goal to achieve during the PHD is two folds:

- 1- Construct and analyze signaling pathways by reasoning on the expert knowledge provided by the scientists. Systems for automated deduction for consequence finding or answer set programming, as well as constraint-based settings will be investigated. The formalization will enable us to (i) establish the consistency or the inconsistency of the model (ii) suggest updates of the knowledge base (model revision) in order to ensure consistency (iii) query properties of the network.
- 2- Guide scientists in the choice of new experiments to be conducted by diagnosing missing interactions in the network. The completeness of the network will be evaluated in relation with the scientists. Discovering missing experiments (initial facts) should be performed automatically by building hypotheses (abductive reasoning).

Depending on the interest and background of the applicant quality of the data obtained based on its provenance could also be studied. Interestingly, ASAM intends to help users evaluating the quality of the Knowledge Base: each part of the network (data or relationship between data) may be associated with quality information depending on its provenance (data obtained by various experiments, each being of various qualities).

Signaling pathways based on two receptors (for which knowledge has been collected and formalized), namely FSH and EGF, will serve as the basis of this work.

The SOLAR system developed at the National Institute of Informatics (Tokyo), which is a system for consequence-finding could be used and extended for tasks 1 and 2, in collaboration with Katsumi Inoue's group in Japan.

Prerequisite: No prerequisite in biology is required. Interests in genomics together with motivation for Artificial Intelligence approaches to Bioinformatics are expected. Basic skills in logic (logic programming) are needed. Depending on the applicant's profile emphasis can be put on Machine Learning methods to acquire signaling networks.

References :

Hidetomo Nabeshima, Koji Iwanuma, Katsumi Inoue and Oliver Ray, SOLAR: An automated deduction system for consequence finding, *AI Communications* 23, 183–203, 2010.

Martin Gebser, Carito Guziolowski, Mihail Ivanchev, Torsten Schaub, Anne Siegel, Sven Thiele, Philippe Veber: Repair and Prediction (under Inconsistency) in Large Biological Networks with Answer Set Programming. *KR (Knowledge reasoning)*, 497-507, AAAI Press, 2010.

Domitille Heitzler, Pascale Crépieux, Anne Poupon, Frédérique Clément, François Fages and Eric Reiter, Towards a systems biology approach of G protein-coupled receptor signalling: challenges and expectations, *C R Biologies* 332(11):947-57, 2009.

Marie Agier, Christine Froidevaux, Jean-Marc Petit, Yoan Renaud, Jef Wijsen, On Armstrong-compliant logical query languages, 4th. Int. Workshop on Logic in Databases (LID 2011), Co-located with EDBT & ICDT 2011, Uppsala.