Translating Disease Networks to Executable Boolean Models
M2 Internship position for 2018/2019

Introduction: Maps of molecular interactions concerning a specific disease have emerged as a useful and intuitive way of describing the mechanisms of the disease in a systematic way. Based on information mining, human curation and expert advice, they summarize current knowledge about biological pathways in a standard representation, both human and readable machine [1].
These disease-specific maps can serve as templates for visualization and analysis of omics datasets, or can be analyzed in terms of the underlying network structure. However, their static nature provides a relatively limited understanding of the emerging behavior of the system under different conditions. Computational modeling can reveal dynamic network properties through in silico simulations and perturbations and can be used to test and predict assumptions. In order to address the lack of kinetic data and the large size of the biological pathways described in a disease-specific map, Boolean modeling can be used to study the qualitative dynamic behavior of the system.
The construction of a map of molecular interactions and a dynamic model are two tasks with different purposes that are today mainly performed in a completely independent way. On the one hand, it is a question of forming a knowledge base in the form of a map, and on the other of defining an abstraction of the system that captures dynamic behaviors of interest. Yet these two constructs share a lot of information, including the mode of influence (e.g., activation or inhibition) or especially the topology of the network.
However, there is currently no automatic pipeline to transition from a static representation to a dynamic model based on the common properties of these two formalisms.

Objectives: In this project we want to establish a complete pipeline allowing the automated construction of large Boolean logical models, from interaction maps. We rely on existing tools [2-6] for the construction of the map, its visualization and division into modules and the analysis and simulation of the logic models obtained and we want to develop the missing brick, CaSQ, in order to to bridge the gap between static and dynamic representations.
The idea is to define simplification rules and logical formulas for the Boolean model according to the topology and the annotations of the starting map (e.g. a degradation does not have to appear in the logic model unless it is catalyzed). Other modes of translation, notably based on the generic converters between models of the type Process Diagram and Activity Flow developed within the framework of the consortium Disease Maps [1], will be considered. The semantics envisioned for translation are based on the notions of activity present in SBGN [7] and in CellDesigner [3], on the difference between reactants and modulators of a reaction (and the type of modulation), etc.

The project will benefit from the progress that has already been made concerning the inference of a large-scale signaling model for rheumatoid arthritis fibroblasts from a molecular map [http://www.colomoto.org/events/2018-eccb/ECCB18_W6_Niarakis.pdf], where a preliminary version of the pipeline was proposed and the CaSQ tool (CellDesigner files as SBML-qual) began to be developed [https://gitlab.inria.fr/soliman/sbgnpd2sbmlq].

The current project will be focused on the further development of CaSQ with implementation of standards-related functionalities (SBGN-PD and SBGN-AF) and the benchmarking of the

pipeline using the large scale disease map for RA developed by GenHotel but also other molecular maps of the Disease Maps consortium.
The resulting logical model will be analyzed using tools of the CoLoMoTo consortium, namely Cell Collective and GINsim [5-6].

The evaluation of the coherence of the proposed RA Boolean model will be tested against published data. Inconsistencies will be progressively fixed through appropriate modifications of logical rules and/or addition/deletion of interactions/components until the model is tuned. Systematic testing of different initial conditions and stimuli could further lead to predictions regarding the outcomes of specific perturbations, as single or combined effects, as well as the identification of novel potential pharmacological intervention points.

Supervision
Dr. Anna Niarakis, (MCF, UEVE, Paris Saclay) is a member of The Disease Maps Project, (http://disease-maps.org/rheumatoidarthritis) where she is responsible for the construction of a detailed molecular map concerning RA and a member of the CoLoMoTo (Consortium for Logical Models and Tools) (http://www.colomoto.org/events/2018-eccb/workshop.html).
The student will benefit from the labs' expertise in RA, the progress that has been made concerning the pipeline and CaSQ and also from the close collaboration and co-supervision of Dr. Sylvain Soliman (CR, EPI Lifeware, Inria Saclay) responsible for the tool development.

Prerequisites:
Ability to work in an interdisciplinary team
Basic bioinformatics skills (familiarity with Python and R is highly appreciated)
Interest in computational systems biology
Good command of the English language

CVs with a motivation letter and at least one reference should be sent to the following address: anna.niaraki@univ-evry.fr with the indication: Application for the M2 post in the subject.
The internship will start February 2019 until July 2019.

Host team :
GenHotel-Evry -EA3886 - Laboratoire de Recherche Européen pour la Polyarthrite Rhumatoïde, Université Evry Val d'Essonne, Paris-Saclay (UEVE-UPSaclay)
2, rue Gaston Crémieux
CP5727 91057 EVRY-GENOPOLE cedex
tel : (33) 01 60 87 45 72

References –Proposed Bibliography
[1] M. Ostaszewski et al. "Community-driven roadmap for integrated disease maps". In : *Briefings in Bioinformatics* (2018). doi : 10.1093/bib/bby024.
[2] F. Fages et al. "Influence Systems vs Reaction Systems". In : Ezio Bartocci, Pietro Lio, Nicola Paoletti, editor, CMSB'16: *Proceedings of the fourteenth international conference on Computational Methods in Systems Biology*, *Lecture Notes in Computer Science,* (2016), vol 9859, p. 98.

[3] Y. Matsuoka et al. “Modeling and Simulation Using CellDesigner”. In : *Transcription Factor Regulatory Networks : Methods and Protocols* (2014). doi : 10.1007/978-1-4939-0805-9_11.

[4] P. Gawronet et al. “MINERVA—a platform for visualization and curation of molecular interaction network”. In : *npj Systems Biology and Applications* (2016), vol. 2, p. 16020, doi : 10.1038/npjsba.2016.20
[5] T. Helikar et al. “The Cell Collective : Toward an open and collaborative approach to systems biology”. In : *BMC Systems Biology* (2012). doi : 10.1186/1752-0509-6-96.
[6] A. Naldi et al. “Logical modelling of regulatory networks with GINsim 2.3”. In : *Biosystems* (2009). doi : 10.1016/j.biosystems.2009.04.008.
[7] N. Le Novere et al. “The Systems Biology Graphical Notation”. In : *Nature Biotechnology* (2009). doi: 10.1038/nbt.1558.
[8] V. Singh et al. “Computational Systems Biology Approach for the Study of Rheumatoid Arthritis : From a Molecular Map to a Dynamical Model”. In : *Genomics and Computational Biology* (2017). doi :10.18547/gcb.2018.vol4.iss1.e100050.