

# Reconstructing causal protein signaling networks from data

Joris Mooij

August 14, 2013

## Abstract

Proposal for a 3 months (6 ec) research project for Research B.

## 1 Description

Proteins play a vital role in the human body. In this project, the student will analyze protein concentration data measured in cells of the human immune system [1] in order to build a causal model of how various proteins interact in these cells. This has potentially important repercussions as malfunctioning of this particular protein signaling system is known to cause cancer. Understanding the complex dynamics of these proteins is traditionally done by performing many costly and labor-intensive experiments. The idea of this project is to investigate how modern *causal discovery* methods perform in reconstructing the protein signaling network from data gathered under various experimental conditions (see also Figure 1).

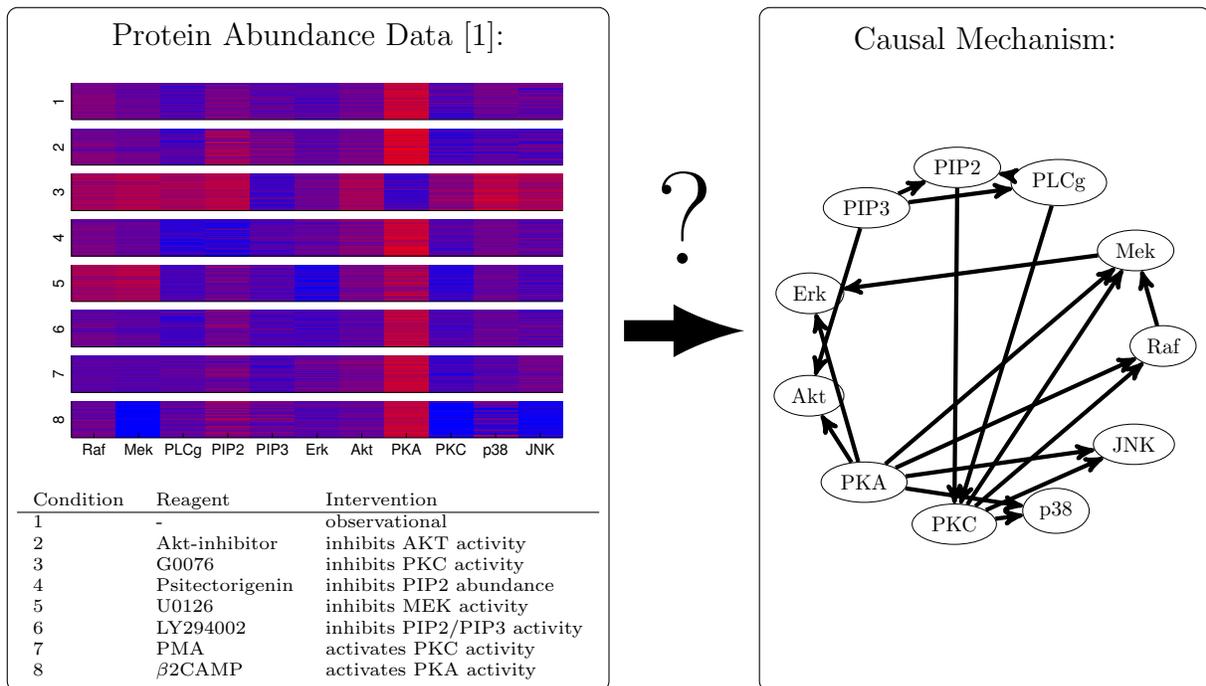


Figure 1: Schematic illustration of the task at hand: given experimental data and background knowledge concerning the experiments, reconstruct the causal structure of the mechanism that generated the data (in other words, infer how the expression of each protein influences the expressions of other proteins).

## 1.1 Type of work

The work to be performed is a combination of theory (30%), programming (40%) and experiments (30%). The student first familiarizes him-/herself with the current state-of-the-art causal discovery method [2] for this type of data, implements an improved version of the method in C++, runs it on the biological data and analyzes the results.

Possible improvements of the current method include optimization (the current method is quite slow), taking into account censoring (the measurement data are cut off at some threshold which is currently ignored), using more background knowledge (about the “sign” of the interventions, i.e., whether it is an activator or inhibitor), and implementing a method to compare the results with previous results obtained by other methods.

## 1.2 Supervisor

dr. Joris Mooij  
Institute for Computing and Information Sciences  
Intelligent Systems  
Telephone: +31 (0)24 3652354  
Email: [j.mooij@cs.ru.nl](mailto:j.mooij@cs.ru.nl)  
Web: [www.jorismooij.nl](http://www.jorismooij.nl)

## References

- [1] Karen Sachs, Omar Perez, Dana Pe’er, Douglas A. Lauffenburger, and Garry P. Nolan. Causal protein-signaling networks derived from multiparameter single-cell data. *Science*, 308(5721):523–529, 2005.
- [2] Joris M. Mooij and Tom Heskes. Cyclic causal discovery from continuous equilibrium data. In Ann Nicholson and Padhraic Smyth, editors, *Proceedings of the 29th Annual Conference on Uncertainty in Artificial Intelligence (UAI-13)*, pages 431–439. AUAI Press, 2013.