Fast contextualization and analysis of logical networks with FALCON

Sebastien De Landtsheer De Landtsheer
Luana Presta
Philippe Lucarelli

2019/01/10

1 Introduction

FALCON is a Matlab toolbox for the contextualization of Bayesian Networks created by the Systems Biology group of the LSRU at the University of Luxembourg. There are three papers related to the toolbox, referenced below [1, 2, 3]. The toolbox is self-contained but you will need the Parallel Computing Toolbox and the Optimization Toolbox. you can type `ver` into Matlab to display the installed toolboxes on your machine.

2 Getting the toolbox

The toolbox is available on the Github repository *github.com/sysbiolux/FALCON*. There are two ways to access the files:

- Installing Github on your computer and cloning the repository.
- Downloading the toolbox as a folder and copying it in your Matlab folder.

In general, it is better to clone the repository as it keeps the online version linked to your local one. Therefore you can update your local version when updates are available and even upload your changes so they are available to others. For this exercise, it is ok to just download the files and copy them. If you plan of using the toolbox in the future, it is better to complete a proper installation with Github. The toolbox contains the necessary drivers and datafiles for this exercise. Please use the *master* version.

3 Simplified NFKB model

We created a simplified model of NFKB signaling, inspired by the previous ODE exercise. The network model is defined as a list of interactions in the file *NFKB-net.xlsx*. Open the file with Excel and look at the structure.

3.1 Questions

1. How many interactions does the model have? How many nodes?

4 Searching the parameters

Run the provided driver script *DriverFALCON_MISB2019.mat* (it will use the datafile *NFKB-data1.xlsx*) until the building of the model. Explore the model structure in the variable called *estim*. Then run the parameter identification.

4.1 Questions

2. What is the number of free parameters? What values can they take?
3. What is the final cost of the optimization? What are the units and what do these unit express?

5 Analyzing the results

5.1 Questions

4. What do you observe in the simulated node values during one simulation?
5. The variable *MeanStateValueAll* contains the predicted steady-state node values. Graph the predicted NFKB activity versus the concentration of IL-1. The indices of the nodes are stored in the variable *estim.state_names*.
6. Run the *FitEvol* function. Do the different optimizations converge to the same optimum? At the same speed? Do they start from the same point? Explain why this is the case.

6 Systems-level analyses

Run the resampling, parameter sensitivity and KO functions. This will take several minutes to complete. Observe the graphs popping up.

6.1 Questions

7. What parameters are the most/least constrained by the data?
8. Which edges are the most/least important for the fit of the model?
9. Which parameters are correlated to each other?

7 Comparing datasets

Change the data file used from *NFKB-data1.xlsx* to *NFKB-data2.xlsx* and to *NFKB-data3.xlsx* and optimize the model for these two datasets.

7.1 Questions

10. Compare the different costs for these three datasets.

8 Comparing models

Change the model file used from *NFKB-net.xlsx* to *NFKB-net-mod.xlsx* and fit it to the dataset *NFKB-data1.xlsx*.

8.1 Questions

11. What modification has been made? Is it sensible in the biological concept?
12. What are the costs for this model? Can we compare these values to the first optimization?
13. Describe your observation regarding the behavior of these two models.

9 Using regularization

Run the first part of the provided script *DriverFALCON_MISB2019_Reg.mat* (no regularization).

9.1 Questions

14. How many free parameters are there now? What values can they take?

Now run the provided script *DriverFALCON_MISB2019_Reg.mat* until the end (regularized part). It will need a few minutes to finish. Analyze the results.

15. Explore the variable *estim.RegMatrix.Groups*. What does it represent and how does it influence the results?
16. The variable *AICs* contains the AIC values for each regularization strength. Plot these values and indicate which regularization strength induces the best model parametrization.
17. Bringing back the model to what it represents, what can we conclude from this result?

References

[1] Sébastien De Landtsheer, Philippe Lucarelli, and Thomas Sauter. Using regularization to infer cell line specificity in logical network models of signaling pathways. *Frontiers in Physiology*, 9(MAY):1–13, 2018.

[2] Sébastien De Landtsheer, Panuwat Trairatphisan, Philippe Lucarelli, and Thomas Sauter. FALCON: a toolbox for the fast contextualization of logical networks. *Bioinformatics*, 33(June 2017):3431–3436, 2017.

[3] Greta Del Mistro, Philippe Lucarelli, Ines Müller, Sébastien De Landtsheer, Anna Zinoveva, Meike Hutt, Martin Siegemund, Roland E. Kontermann, Stefan Beissert, Thomas Sauter, and Dagmar Kulms. Systemic network analysis identifies XIAP and IκBα as potential drug targets in TRAIL resistant BRAF mutated melanoma. *npj Systems Biology and Applications*, 4(1):39, 2018.