Setting parameters for biological models with ANIMO

Stefano Schivo
Formal Methods and Tools
Faculty of EEMCS
University of Twente
Enschede, The Netherlands
s.schivo@utwente.nl

Jetse Scholma Marcel Karperien Janine N. Post
Developmental BioEngineering
MIRA Institute for Biomedical Technology and Technical Medicine
University of Twente
Enschede, The Netherlands
{j.scholma, h.b.j.karperien, j.n.post}@utwente.nl

Jaco van de Pol Rom Langerak*
Formal Methods and Tools
Faculty of EEMCS
University of Twente
Enschede, The Netherlands
{j.c.vandepol, r.langerak}@utwente.nl

ANIMO (Analysis of Networks with Interactive MOdeling) is a software for modeling biological networks, such as e.g. signaling, metabolic or gene networks. An ANIMO model is essentially the sum of a network topology and a number of interaction parameters. The topology describes the interactions between biological entities in form of a graph, while the parameters determine the speed of occurrence of such interactions.

When a mismatch is observed between the behavior of an ANIMO model and experimental data, we want to update the model so that it explains the new data. In general, the topology of a model can be expanded with new (known or hypothetical) players, and enable it to match experimental data. However, the unrestrained addition of new parts to a model causes two problems: models can become too complex too fast, to the point of being intractable, and too many parts marked as "hypothetical" or "not known" make a model unrealistic. Even if changing the topology is normally the easier task, these problems push us to try a better parameter fit as a first step, and resort to modifying the model topology only as a last resource.

In this paper we show the support added in ANIMO to ease the task of expanding the knowledge on biological networks, concentrating in particular on the parameter settings.

1 Introduction

The investigation of biological processes relies on computational support on a daily basis. This happens not only because of the extremely large amount of data generated in the *-omics era*, but also because many processes are simply too complex to be understood by the human brain alone. For this reason, systems biology has become more and more important in the last several years.

A single biological network such as a signaling or gene network, may involve up to hundreds of different players. As it would be very difficult to understand the dynamic behavior of such networks just by looking at their static representations, many tools were built to help the biologists in that sense. ANIMO (Analysis of Networks with Interactive MOdeling) is one of such tools. Its primary objective is to let the expert biologists work directly on the formalization of their knowledge, supporting the generation of new insights on the studied processes. An ANIMO model is formed by two main parts: the network topology and the parameters. The topology describes which biological components are included in the

*Corresponding author

Submitted to:
SynCoP 2014

model, and which are the interactions we want to represent. The parameters define the rate of occurrence of such interactions, which are described based on simplified kinetic formulae.

Proteins are normally expressed at different concentrations in different individuals of the same species, and yet the overall behavior of their biological networks does not differ significantly. This phenomenon has led to the notion that biological networks are inherently *robust*. In modeling terms, this means that most of the parameters of a model can vary inside a certain interval without influencing the qualitative behavior of the whole network. The interactive approach of ANIMO is based on the assumption of robustness, as our tool is mainly aimed at the development of network models with a focus on the topology. Ideally, the biologist can "play" with the topology of a network, working more towards matching the qualitative behavior of experimental data, rather than precisely reproducing it. However, it is not our intention to concentrate exclusively on the network topology: in many cases a better parameter choice can improve the behavior of a network more than the addition of new components. Indeed, making an unnecessarily complex model could reduce its usefulness both in terms of analysis performances and closeness to reality. The first problem is simply due to the complexity of a network, which would require more and more computational resources to be analysed[1]. Realism of network models is more related to their ultimate usefulness: a model that explains a particular behavior very well but contains many nodes marked as "unknown" has little applicability, as its connection with known processes is very loose. Therefore, it is desirable that a better parameter set is regularly sought for during the design cycle of a complex biological network model. Some support for parameter choice was already provided in the first versions of ANIMO. We present here an extended set of tools aimed at achieving a closer fit between ANIMO models and experimental data. A guideline on how to use these tools to get the best results will also be presented as an ideal workflow. Thanks to the better awareness on parameter choice gained through this new extension of ANIMO, the biologists will be able to judge more easily which are the most promising topologies for a network, and thus drive the experimental research more efficiently.

2 ANIMO models

The starting point of ANIMO is the traditional static representation of biological networks, which can be easily drawn and managed in softwares like Cytoscape. Indeed, ANIMO was implemented as a plug-in to Cytoscape, with the aim of adding dynamics to the static representation of biological networks, and thus allow for analysis on the behavior of such networks. The user interface of ANIMO can be seen in Figure **??**: at the center is the Network panel, where the network model is represented in the familiar nodes-edges form used in the domain of biology. Models in ANIMO are *activity-based*, in the sense that nodes have their activity as main property, and interactions among nodes change the activity of their targets. The concept of activity can be intended for example as a generic post-translational modification a molecule can undergo to change its function. In the case of a kinase, the phosphorylated state is usually interpreted as active. In the context of a gene network, the activity of a node standing for a gene represents its current transcriptional status. A basic type of analysis that can be performed with ANIMO is the generation of a simulation run, which is presented to the user in the form of a graph (to the right). The activity graphs generated by ANIMO show the variation in activity of selected nodes over the course of the simulation run. In addition to that, a slider placed under each graph allows the user to color the nodes in the Network panel depending on their activity level at any point during the course of the simulation: the Legend panel on the left links colors to activity.

[1] We also refer to the problem of *state space explosion*: when a model contains too many loosely coupled components, the set of its possible evolutions grows exponentially, to the point of making it impossible to apply some analysis techniques.

Figure 1: fig:small-example-feedback

Of the two main components of an ANIMO model, only the topology is immediately visible to the user in the Network panel; the parameters are accessed by double clicking the arcs representing node-node interactions. The dialog window that is shown for an interaction contains the details of the abstract reaction kinetic describing the interaction, together with the current value of its parameter k (see Fig. **??**). The unique parameter associated to any interaction in ANIMO is used as a scale factor to make the modelled reaction occur faster or slower. We also provide the user with pre-set values for k, encouraging an initial qualitative assignment of reaction rates as "slow", "fast" and so forth. This approach is based on the assumption that biological networks are inherently robust: once an acceptable set of parameters is found, a more precise parameter search will generally have little impact on the fit of a model. However, as can be seen in the small example in Figure **??**, robustness does not imply that any parameter choice will do. In particular, network motifs such as feedback loops make models more dependent on parameter settings. As biological networks tend to heavily rely on cross-talk, network topologies can rapidly become complex. Some manual parameter fitting is currently necessary for the more complex ANIMO models. This work can be slow and error-prone, taking away some of the user-friendliness for which ANIMO aims. We will show how ANIMO has been improved to make parameter choice easier for the user.

3 Support for parameter synthesis in ANIMO

When a model does not match experimental data, ANIMO offers three main tools to achieve a better fit: manual parameter editing, comparisons between different model versions, and the newly introduced parameter sweep.

3.1 Manual parameter editing

Double-clicking on an arrow in the Cytoscape representation of an interaction lets the user access a dialog that allows to change the approximated scenario and parameters for the selected interaction. The simplified scenarios were described in [**?**], and we explain them briefly here:

scenario 1

scen 2

scen 3

The choice for the parameter k can be made directly inserting a numeric value, or by choosing a preset among the proposed qualitative values *very slow, slow, medium, fast, very fast*. We advise our users to base their initial parameter choice on these qualitative presets. In Figure **??**, all parameters are initially set to *medium* (Fig. **??**a), while the peak in b was obtained by setting the parameter of $C \dashv B$ to *fast* and the parameter of $E \dashv D$ to *slow*.

3.2 Comparing model versions

Particularly large networks make it more difficult for the user to understand the effect of a change in network topology or interaction parameters. To overcome this difficulty, ANIMO allows the user to visually compare two versions of a model in terms of their simulation results. The Results panel in Figure **??**

shows the graph of selected node activities during the course of an ANIMO simulation. The button "Compare with..." allows to compare the current simulation data with another based on a possibly different version of the model. When the two simulations to be compared have been chosen, ANIMO produces a new graph plotting the difference between the two original simulations. The slider under the new graph allows to visualize also in the Network panel the changes in node activity in the whole network: Figure **??** shows the difference between the two versions of the model in Figure **??**. The difference was computed as `b - a`. Nodes colored in green in the Network panel are more active in version `b`, while red nodes are more active in version `a`.

By changing the title of a simulation, a user can keep track of different versions of their model, which can be also kept and saved with the model, to be used in future sessions or for sharing purposes. Should a user want to backtrack to a particular version of the network, the "Reset to this" button can be used. As an additional help to recall what a model version consisted of, the tooltipo of the "Reset to this" button shows an image of the network topology, taken in the moment when the selected simulation was begun. A click on the "Reset to this" button will reset the network topology and parameters at the ones used to generate the selected simulation. As the adaptation of a model may proceed on different paths, the "Compare with..." and "Reset to this" buttons provide the user with some help in tracking the changes to the model and selecting the most promising ones.

3.3 Parameter sweep

In order to avoid introducing too many changes to the topology of a network, we recommend ANIMO users to be reasonably sure that a topology does not fit a given data set before modifying it. As this would entail an extensive parameter search which, even with the tools described in Sections 3.1 and 3.2, would require a considerable amount of time and effort, we have provided the latest version of ANIMO with a support for parameter sweeps. The idea is to let ANIMO explore the user-defined parameter space by computing a series of simulations based on the same network topology, when only the parameters are being changed. Comparing the results with a given experimental data series will let the user select the most promising parameter choices, eventually selecting the one that behaves the best.

4 Suggested ANIMO workflow

5 Conclusions

References

[1] R.J. van Glabbeek (2010): *The EPTCS bibliography style*. Available at `http://biblio.eptcs.org`.

[2] E.D.I. Thor & E. di Thor, editors (1987): *Title B1: A Book with editors*, second edition. *Series* 21, Publisher, Address, doi:10.4204/EPTCS. Available at `http://arxiv.org/abs/1009.3306`. Note.

[3] E.D.I. Thor & E. di Thor, editors (1987): *Title B2: An Inbook with editors*, second edition, type (of chapter) II, pp. 1–999. *Series* 443, Publisher, Address, doi:10.4204/EPTCS. Available at `http://arxiv.org/abs/1009.3306`. Note.

[4] E.D.I. Thor & E. di Thor, editors (1987): *Title B3: A Proceedings*. *Series* 443, Organization, Publisher, Address, doi:10.4204/EPTCS. Available at `http://arxiv.org/abs/1009.3306`. Note.

[5] A. Ut, H. Or & Co Author (1987): *Title A: An Article*. *Jounal* 443(21):4, doi:10.4204/EPTCS. Available at `http://arxiv.org/abs/1009.3306`. Note.

[6] A. Ut, H. Or & Co Author (1987): *Title B1: A Book with authors*, second edition. *Series* 443, Publisher, Address, doi:10.4204/EPTCS. Available at `http://arxiv.org/abs/1009.3306`. Note.

[7] A. Ut, H. Or & Co Author (1987): *Title B2: An Inbook with authors*, second edition, type (of chapter) II, pp. 1–999. *Series* 443, Publisher, Address, doi:10.4204/EPTCS. Available at `http://arxiv.org/abs/1009.3306`. Note.

[8] A. Ut, H. Or & Co Author (1987): *Title C1: An Inproceedings*. In E.D.I. Thor & E. di Thor, editors: *Booktitle*, *Series* 443, Organization, Publisher, Address, pp. 1–999, doi:10.4204/EPTCS. Available at `http://arxiv.org/abs/1009.3306`. Note.

[9] A. Ut, H. Or & Co Author (1987): *Title C2: An Incollection*. In E.D.I. Thor & E. di Thor, editors: *Booktitle*, second edition, type II, *Series* 443, Publisher, Address, pp. 1–999, doi:10.4204/EPTCS. Available at `http://arxiv.org/abs/1009.3306`. Note.

[10] A. Ut, H. Or & Co Author (1987): *Title R: A Techreport*. Type (of report) 21, Institution, Address, doi:10.4204/EPTCS. Available at `http://arxiv.org/abs/1009.3306`. Note.

[11] A. Ut, H. Or & Co Author (1987): *Title R1: A Techreport*. Technical Report 21, Institution, Address, doi:10.4204/EPTCS. Available at `http://arxiv.org/abs/1009.3306`. Note.

[12] A. Ut, H. Or & Co Author (1987): *Title R2: A PhDthesis*. Ph.D. thesis, School, Address, doi:10.4204/EPTCS. Available at `http://arxiv.org/abs/1009.3306`. Note.

[13] A. Ut, H. Or & Co Author (1987): *Title R3: A Masterthesis*. Master's thesis, School, Address, doi:10.4204/EPTCS. Available at `http://arxiv.org/abs/1009.3306`. Note.

[14] A. Ut, H. Or & Co Author (1987): *Title R4: A Manual*, second edition. Organization, Address, doi:10.4204/EPTCS. Available at `http://arxiv.org/abs/1009.3306`. Note.

[15] A. Ut, H. Or & Co Author (1987): *Title U: A Booklet*. Howpublished, Address, doi:10.4204/EPTCS. Available at `http://arxiv.org/abs/1009.3306`. Note.

[16] A. Ut, H. Or & Co Author (1987): *Title U: A Misc*. Howpublished, doi:10.4204/EPTCS. Available at `http://arxiv.org/abs/1009.3306`. Note.

[17] A. Ut, H. Or & Co Author (1987): *Title U: An Unpublished*, doi:10.4204/EPTCS. Available at `http://arxiv.org/abs/1009.3306`. Note.