Modeling mechanistic biological networks: An advanced Boolean approach

T. Handorf [1]*, E. Klipp [1]

[1]Department of Theoretical Biophysics, Humboldtuniversität zu Berlin, Invalidenstr. 42, 10115 Berlin, Germany

ABSTRACT

Motivation: The understanding of the molecular sources for diseases like cancer can be significantly improved by computational models. Recently, Boolean networks have become very popular for modeling signaling and regulatory networks. However, such models rely on a set of Boolean functions which are in general not known. Unfortunately, while detailed information on the molecular interactions becomes available in large scale through electronic databases, the information on the Boolean functions does not become available simultaneously and has to be included manually into the models, if at all known.

Results: We propose a new Boolean approach which can directly utilize the mechanistic network information available through modern databases. The Boolean function is implicitly defined by the reaction mechanisms. Special care has been taken for the treatment of kinetic features like inhibition. The method has been applied to a signaling model combining the Wnt and MAPK pathway.

Availability: A sample C++ implementation of the proposed method is available for Linux and compatible systems through http://code.google.com/p/libscopes/wiki/Paper2011

Contact: handorf@physik.hu-berlin.de

1 INTRODUCTION

An important focus in Systems biology are emergent properties of biological systems, properties which arise from the system as a whole and which cannot be explained by looking at individual components alone. A common theme in this evolving field are biological networks which describe complex relations between biological entities. Modeling these networks uncovers the systems behavior as a whole based on the properties and interactions of the contained compounds.

An important subset of such networks describes mechanistic processes among biochemical species. Such processes play an important role in cellular metabolism, signal transductions and gene regulation. Computational models of these networks have been investigated for a long time. A first category of models concentrate on single pathways, i.e. functionally related parts of the networks, with the use of differential equations (Rapoport *et al.*, 1974; Rizzi *et al.*, 1997; Wolf and Heinrich, 2000; Teusink *et al.*, 2000; Heinrich *et al.*, 2002; Schoeberl *et al.*, 2002; Zi and Klipp, 2007; Kofahl and

Wolf, 2010). Such models usually yield high quality predictions of the system dynamics with quantitative information on the molecule concentrations. Those models, however, require accurate kinetic parameters.

The emergence of large scale electronic databases like KEGG (Kanehisa *et al.*, 2010), Reactome (Matthews *et al.*, 2009) or the Pathway Interaction Database (PID)(Schaefer *et al.*, 2009) opened new directions in modeling biological networks. For large models it is often unfeasible to obtain all the necessary kinetic parameters. In such cases so called structural methods are applied. These include elementary flux modes (Schuster and Hilgetag, 1994; Schuster *et al.*, 2000), flux balance analysis (Bonarius *et al.*, 1997; Edwards and Palsson, 2000) and Petri nets (Reddy *et al.*, 1996; Genrich *et al.*, 2001).

A further structural method which became very popular in particular for signaling and gene-regulatory networks are Boolean networks (Kauffman, 1969, 1986; de Jong, 2002; Handorf *et al.*, 2005; Fisher and Henzinger, 2007; Saez-Rodriguez *et al.*, 2009; Samaga *et al.*, 2009). In a Boolean network, each compound is represented by a network node and has an associated Boolean value. This value either represents the existence of a compound, i.e. it is *TRUE* if the compound is present in a significant concentration and *FALSE* if not, or it represents the activity. In the network, nodes are connected by edges which represent biochemical interactions. The Boolean value of a node is calculated in each step of the algorithm through a Boolean function which depends on the Boolean values of all inbound nodes.

Usually, while the topology of the network is known (e.g. from interaction databases or qualitative interaction studies), the Boolean function is unclear. One approach is to estimate this function by adapting the model to experimental data as shown in Saez-Rodriguez *et al.* (2009). In Handorf *et al.* (2005) we introduced a different Boolean approach for metabolic networks, the concept of Scopes, which circumvents this problem. Here, a (metabolic) reaction is active if all of its substrates are present (i.e. *TRUE*). Subsequently all compounds become *TRUE* for which at least one producing reaction is active. Hence, the Boolean function is clearly defined by the topology and is represented by the conjunction (AND) of all substrates of a reaction.

In signaling and gene-regulatory networks the situation is more complex. In fact, there exist different views on these networks. A phenomenological view (Figure 1a) represents experimental findings, like A and C activate and B and D inactivate protein E. This

---

*to whom correspondence should be addressed
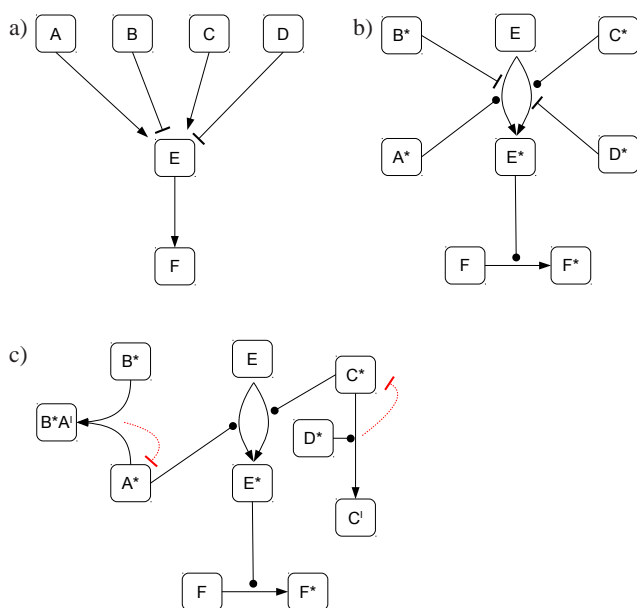
Fig. 1. Different representations of signaling or regulatory processes: The three views shall represent the same mechanistic process. a) network of activating and inactivating proteins. b) intermediate representation with reactions. c) fully mechanistic representation, the dotted inhibition arrows represent additional markings which carry the information on inhibitions. (*) denote active proteins and (I) inactivated proteins in b) and c). The different views roughly equal the SBGN Activity Flow (a) and SBGN Process Description (c) (Novre *et al.*, 2009).

view, which is similar to the SBGN[1] Activity Flow (Novre *et al.*, 2009), is used for classical Boolean models as indicated above and requires a Boolean function which converts the states of the proteins A to D into the state of protein E.

Clearly, there exist molecular mechanisms behind the experimental observations which can be described by biochemical reactions. A view which is particularly popular in differential equation models is shown in Figure 1b where reactions have substrates and products as in metabolic network but can also be influenced by a set of activators or inhibitors.

However, this view is still not fully mechanistic. Reactions are never directly activated or inhibited by a compound. The only agent that effects the reaction rate is the catalyst. Other modifications are indirect by influencing the activities of the substrates or the catalyst (Figure 1c). Within this fully mechanistic view, which is similar to the SBGN Process Description, the Boolean function is again clear: Its a conjunction (AND) of all substrates and the catalyst.

This type of mechanistic network information can be retrieved from modern databases like Reactome or PID. With the present method we are able to automatically create a functional Boolean network model including the Boolean function for signaling or gene-regulatory pathways.

Simulation of the resulting Boolean model allows to explore potential qualitative behavior of the network upon stimulation. Obviously, such a Boolean network cannot reproduce behavior which results from the kinetics of the participating biochemical

---

[1] Systems Biology Graphical Notation

reactions. This holds in particular for inhibition which in the fully mechanistic view is only indirectly included (inhibition occurs due to competition for the same substrate, see methods).

This is in contrast to manually built classical Boolean networks which are derived from experimental observations and therefore naturally include inhibitory effects. By including phenomenological data, the classical Boolean networks are actually not purely topological anymore. Including an inhibitory interaction already assumes that the compound concentrations or the kinetic properties of the participating reactions are in certain suitable ranges. The fully mechanistic view however does not contain implicit kinetic information. Hence, this type of information has to be added to the automatically created network as discussed in the methods section.

The proposed Boolean approach has been applied to a signaling and gene-regulatory network describing the interaction between the Wnt pathway and the MAP kinase cascade. The targets of the two pathways, β-catenin and ERK, are known initiators for differentiation and proliferation and are also of great interest in diseases like cancer. In Kim *et al.* (2007) a model of these two interacting pathways has been published. It includes crosstalk at the signaling and gene-regulatory level.

Our model could reproduce their findings. We also manually implemented a classical Boolean network using CellNetAnalyzer (Klamt *et al.*, 2007) and compared it to our model.

2 APPROACH

We implemented the network model of crosstalk between the Wnt and MAPK pathways presented in Kim *et al.* (2007), as shown in Figure 2. To this end, we imported the reactions from the Reactome database (see supplemental section S4). The imported network was further refined to match the model of Kim *et al.* as not all utilized reactions are in the database, yet. It should be noted that this is not a mandatory step in general. The method described here is able to directly use the data from the database if the contained information already sufficiently describes the analyzed system.

Furthermore, kinetic features of the system, like inhibition, have been identified and included into the model as described in the method section.

In the refined model, the crosstalk between the WNT and MAPK pathways is represented through, first, an unknown protein X which is transcribed in response to the transcription factor β-catenin and which activates cRaf upstream of ERK, second, an inhibition of GSK3β by active ERK which in turn inhibits the destruction complex of β-catenin in the Wnt pathway and third, a direct stimulation of RAS by the Wnt signal. The known inhibitory interactions in the pathways were incorporated as mentioned before. The supplemental Figure S6 shows the complete network.

In order to compare the result of our scopes method to the result of the classical Boolean approach, we manually built an activity flow model of the participating species. This model (shown in Figure 3) is much less complex than the model created from the databases since it omits many mechanistic details and requires the manual definition of the network logic. Model simulations were performed using synchronous Boolean simulation in the "odefy" module of CellNetAnalyzer.

In a first step we investigated the effect of the Wnt stimulus on the two target proteins ERK and β-catenin for the MAPK and the Wnt
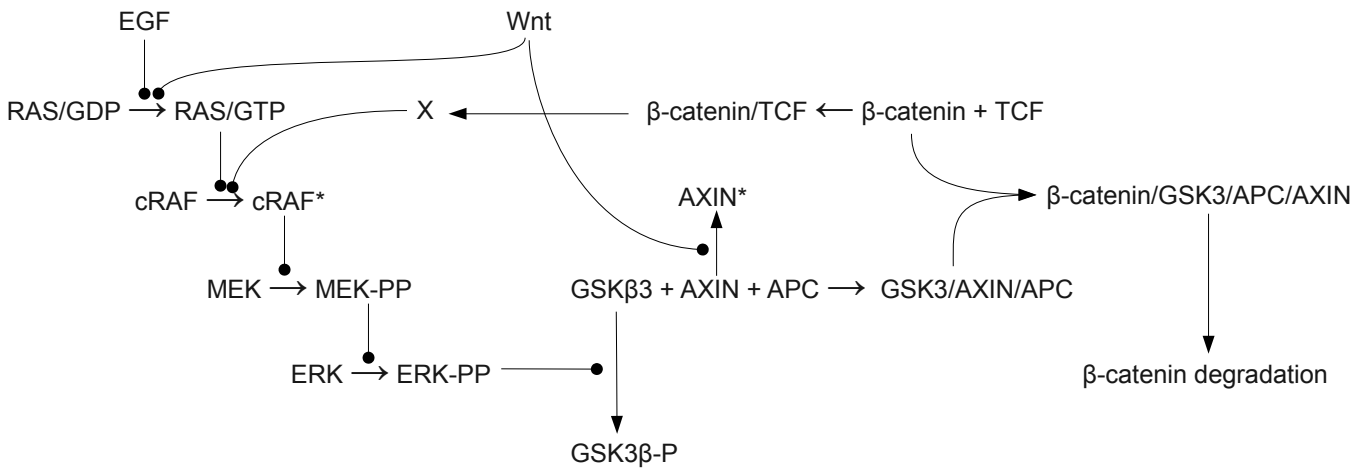
Fig. 2. Crosstalk of the Wnt and Erk pathways. Modified from Kim *et al.*. Crosstalk is represented through, first, an unknown protein X which is transcribed in response to the transcription factor β-catenin and which activates cRaf upstream of ERK, second, an inhibition of GSK3β by active ERK which in turn inhibits the destruction complex of β-catenin in the Wnt pathway and third, a direct stimulation of RAS by the Wnt signal.

pathway. Figure 4 shows similar behavior for our method and the classical Boolean approach as provided by CellNetAnalyzer. Before the Wnt stimulus is applied the system is usually assumed to be off in the absence of crosstalk. The positive feedback loop introduced with the crosstalk, however, leads to a periodic activity of the two target proteins in both models. Once the Wnt stimulus is applied, both, ERK and β-catenin become continuously active. If the Wnt stimulus is removed (not shown for CellNetAnalyzer) both proteins stay active which is again due to the positive feedback loop.

We performed an attractor analysis for both models which yielded results consistent with the above observations. Please see section S3 in the supplement for further details.

Kim *et al.* reported similar results using a model of differential equations. The effect of the feedback loop is actually dependent on the kinetic parameters and in particular on whether the signal
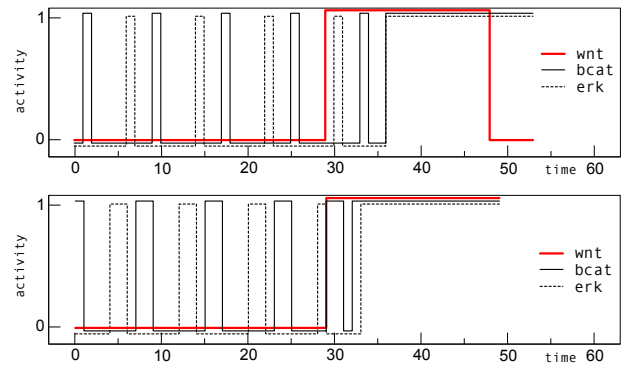

Fig. 4. Response of β-catenin and ERK to WNT stimulus: (**upper**) results for the scope method; (**lower**) results for CellNetAnalyzer. Here, time denotes the step number of the Boolean simulation which is an abstract time which can roughly considered as a monotonous function of the real time
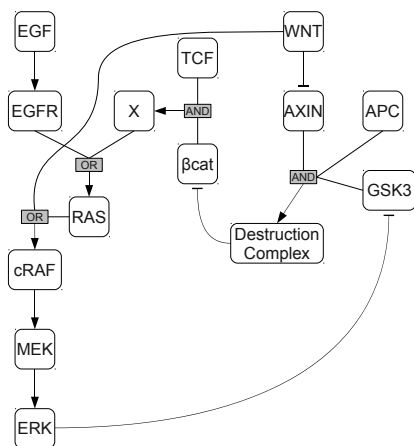
Fig. 3. Activity flow diagram of the Wnt and MAPK pathway and their crosstalk. The Boolean functions have been defined manually in order to reproduce the known behavior of the components. Logical operators have been indicated in the Figure if they were not clear from the topology alone.

through the crosstalk is strong enough for sustained activation. For a standard parameter set they reported that the activity stayed in a low state without the stimulus, switched to active with the stimulus and returned to the low state after the removal of the stimulus. However, by varying the synthesis rate of β-catenin or the phosphatase activity for ERK, the system showed an increased activity already before the stimulus and a sustained activity after stimulus removal. This parameter dependent distinction cannot be made with the Boolean approaches.

At that point it should be noted that the oscillations in the off-state observed in the Boolean systems are artifacts of the Boolean formalism. In fact, in a differential equation model they would not correspond to a stable limit cycle. It is a mere expression of that an initial activation in the cycle is passed around without being attenuated or amplified. Differential equation models could show oscillations of this type transiently, in particular if the initial activation time is shorter than the round-trip time.
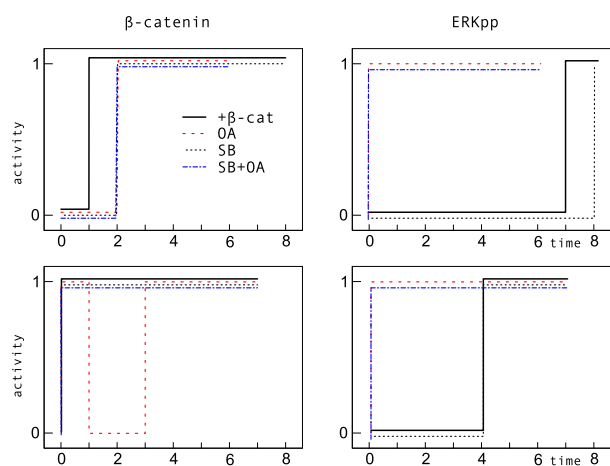
Fig. 5. Response of β-catenin (**left**) and ERK (**right**) to the different interferences (β-catenin overexpression (+β-**cat**), ERK-phosphatase inhibition with okadaic acid (**OA**), GSK-3β inhibition with SB216763 (**SB**) and both (**SB+OA**)). The (**upper**) graph shows the results for the scope method, the (**lower**) for CellNetAnalyzer. The different curves have been shifted slightly to avoid overlap.

Kim *et al.* argued that the activating potential of the reported feedback loop may be important in particular in cancer, indicating that a slight variation of the kinetic rates may result in a persistent activation of the proliferation related Wnt and MAPK pathways even in the absences of the corresponding stimuli. They further examined their model by a set of interference experiments which they experimentally validated. First, β-catenin has been overexpressed (+β-**cat**), second, active ERK levels have been increased by phosphatase inhibition using okadaic acid (**OA**), third, GSK3β has been inhibited by SB216763 (**SB**), and fourth, the effect of ERK activation and GSK3β inhibition has been studied in combination (**SB+OA**).

We repeated the experiments using our method and compared it with the results from the classical Boolean approach (Figure 5). The results from the two Boolean methods are similar and actually identical with respect to the temporal order of the responses to the different interferences when allowing that two effects reported at different times with one method may occur at the same time in the other method. One exception is the short initial activation of β-catenin for **OA** in CellNetAnalyzer which is due to the unstimulated oscillation discussed earlier.

The observed temporal order is in line with the results from Kim *et al.*. With β-catenin overexpression β-catenin comes up first while ERK activation shows a long delay. OA treatment shows fast ERK response and a delayed β-catenin activation. SB has a similar effect as β-catenin overexpression while the combination SB+OA shows the fastest effects of the two single interferences.

3 METHOD

The purpose of the proposed method is to model mechanistic reaction networks as provided by modern electronic database using a novel Boolean approach. The method is based on the previously published concept of scopes. Scopes have been developed for metabolic reaction networks. It is

a step-wise algorithm. Initially a set of seed compounds is defined which is the set of initially available compounds. In each following step all active reactions are determined by testing whether all substrates of an reaction are available. Then, all products of the active reactions are added to the set of available compounds. This process yields a set of compounds (the scope) that are synthesizable from the seed compounds.

Here we adapt the concept to signaling networks. The challenge is to include the effects of catalysts, activators and inhibitors. As discussed before, in the fully mechanistic view, towards which modern databases evolve, actually only catalysts remain. This can be included in the method by simply requiring the availability of all substrates and the catalyst for the execution of a reaction. We call this the *static mode* of our method which is equivalent to the previously defined Scopes.

Clearly, in signaling and gene-regulatory networks, dynamical effects, and in particular inhibition, are especially important for the cellular functions of these networks. As argued before, in the fully mechanistic view, inhibition is indirect. It occurs due to reactions where an inhibitor catalyzes a modification of a compound, thereby leading to a depletion of the active form of the compound. To distinguish inhibiting reactions from other non-inhibiting reactions which use the same compound as substrate, an inhibitory flag has to be set to mark that the reaction is inhibitory to a specific substrate.

As an example we consider the reaction $A^* + B^* \rightarrow A^I B^*$ from Figure 1c which depletes A^* and thereby inactivates it. This effect is described by an "inhibition flag" to these reactions which is visualized by tee-shaped arrows in Figures 1c and S6(supplement). This flag essentially describes kinetic information. It means that for example in case of Michaelis-Menten kinetics the K_m-value is sufficiently low and the V^{max}-value is high enough to cause a low enough concentration of A^* such that other reactions depending on A^* are not significantly activated.

For the *dynamic mode* of the method this flag is used to include inhibitory effects into the Boolean network approach. Therefore, we assign to each compound node two Boolean variables, one indicating the presence of the compound (as in the classical approach) and the second indicating its depletion.

We further support the second paradigm of inhibition, the direct inhibition of a reaction by an inhibitor (cf. Figure 1b). Although databases like Reactome encourage their curators to use the fully mechanistic view, if known, this paradigm will still be around for the next time. The activity flow scheme (cf. Figure 1a) is not supported as this is perfectly covered by classic Boolean approaches.

Apart from inhibition, there are also other dynamical features which may influence the qualitative behavior of the automatically generated Boolean model. In particular, depending on the actual implementation in the database, a simple phosphorylation may become an amplifying feedback loop. For example in our test case, phosphorylated MEK (a kinase) binds to its target ERK, phosphorylates it and dissociates. Once activated, phospho-MEK will continuously activate ERK as it is recycled. In biology such an amplification may of course happen, but its quantitative effect strongly depends on the kinetic parameters, in particular on how fast phospho-MEK is dephosphorylated and, hence, deactivated.

For our example network, we use an instantaneous deactivation of phospho-MEK after ERK activation. This is implemented by connecting the dissociation reaction to a new compound which cannot be used for further target activation.

The method is implemented as follows: The second inhibitory paradigm is dealt with by requiring all reaction inhibitors to be absent. Admittedly, this is not the only Boolean function possible here. However, as the inclusion of this type of inhibition is only a concession to the current data situation and the usage of this type is already very sparse in e.g. the Reactome database we assume this to be a practical solution.

The first inhibition paradigm, the depletion inhibition, is more complicated. As mentioned, an additional Boolean variable is introduced representing the depletion state of a compound. This variable is set to *TRUE* if at least one outgoing inhibitory reaction is active. All non-inhibitory
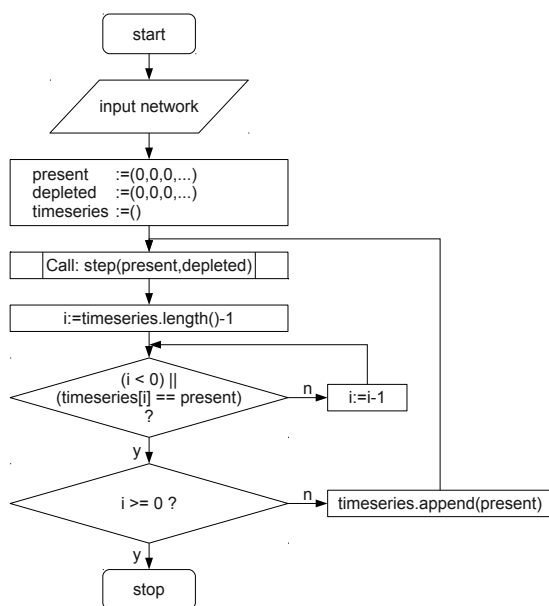
Fig. 6. General algorithm. "present" and "depleted" are arrays which contain presence and depletion variables (1/0 for *TRUE* or *FALSE*) for each compound. "timeseries" is an array containing the state for each time step.

reactions are only active if all of their substrates and the catalyst are available and none of these is depleted. The algorithm is summarized in Figure 6. The implementation of a single iteration step is shown in Figure 7.

The algorithm also features depletion propagation. This is important for example if two forms of a protein can be considered in quasi steady state, i.e. they can be fast and reversibly converted into one another. If, for example, the second form is depleted then also the first form needs to be depleted and, hence, inactivated even if form two is originally produced from form one. This information is included into the reaction converting form one into form two by indicating that the depletion state of the product is passed on to the substrate.

A step by step walk-through of the algorithm on an example network is provided in the supplement, section S1. The method can be reformulated as a classical Boolean network by introducing further "virtual" nodes. This is discussed in the supplement, section S2.

As for models in general, initial conditions are important. Setting only the stimulants (i.e. Wnt and EGF in the test case) to *TRUE* will not be sufficient. In fact, there are a lot of proteins in signaling pathways which are present either in their inactive or active forms. In that respect, the two approaches differ, as for our method inactive forms need to be initially provided (e.g. unphosphorylated MEK must be present), while for the classical approach the inactive forms are ignored. There is also a number of active forms which are assumed to be present in the beginning. Table 1 shows the initially present nodes for the two Boolean approaches for discussed WNT/MAPK model.

In principle, the information on initial presence or activity of compounds depends on prior biological knowledge. However, for larger networks, defining the initial conditions manually may be tedious. As proposed in Handorf *et al.* (2008) reasonable sets of initial conditions can be inferred automatically from the network topology. For this method in particular the above defined *static mode* will be important.

4 DISCUSSION

We proposed a new method for the automated generation of Boolean network models from curated mechanistic network databases. Here,
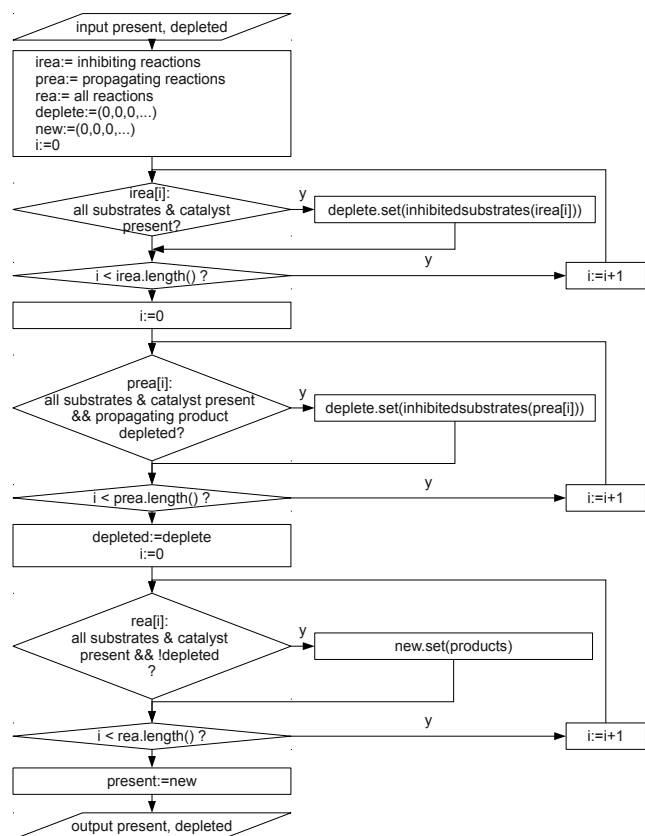
Fig. 7. Module "step": Calculate a single iteration. The algorithm iterates over all inhibiting reaction (irea), propagating reaction (prea) and all reactions (rea). Iterating over "irea" and "prea" create a new array "deplete" containing compounds which become depleted in the current step. When iterating over all reactions products of active reactions are set in the new array "new". "new" and "depleted" are set to "present" and "depleted" respectively. This ensures synchronous operation. Please note that the figure over-simplifies the algorithm a bit. To be exact: When iterating over "irea" and "prea" depleted substrates are only allowed if they are the "inhibiting" substrates. Further, in the loop over "rea" inhibiting reactions can be active if their inhibiting substrates are depleted and hence the products are activated. This is the desired behavior.

the Boolean functions are implicitly defined. Kinetic properties of the biological system, such as inhibitions or amplifications may influence the system behavior but are not automatically included when importing from these databases. This requires an additional review of the model and may require further refinements.

In fact, one should refrain from perceiving Boolean networks as purely topological. They contain indeed kinetic information as they are generally inferred from the observed dynamical system behavior, even if the exact kinetic parameters are not known and, hence, are not part of the network definitions. However, kinetic features will enter the Boolean network in terms of inhibitory or activatory edges.

As discussed, inhibition in mechanistic reaction networks, as available through the mentioned electronic databases, occurs through competion of different reactions for the same substrate or catalysts. Depending on the kinetic parameters, one reaction may

Table 1. Initial conditions

Scopes		CellNetAnalyzer
p21 RAS:GDP	MEK	AXIN
ATP	Gsk3beta	APC
DSH	PP2A	GSK3
β-catenin	CK1alpha	β-catenin
APC	Axin1	TCF
TCF	GTP	
cRafpp:14-3-3	LRP	
Frizzled receptors	ADP	
unidentified protein kinase		

Nodes initially set (*TRUE*) in the Wnt/MAPK pathway for our approach and in CellNetAnalyzer. These nodes will be held *TRUE* during the complete simulation. All other compounds are initially *FALSE* but may change during simulation.

deplete the substrate concentration significantly, thereby leading to an inactivation of other reactions. If such inhibitions are observed in a particular pathway, they can be added to the network in a subsequent functional curation step and considered by our proposed Boolean method.

We tested our method on an example network describing the crosstalk of the Wnt and MAPK pathways. This network model contains several cases of the mentioned substrate depletion, for example the inhibition of β-catenin by binding to the destruction complex. We compared our method to the classical Boolean approach as well as to the results of Kim *et al.* (2007) who did a differential equation model with experimental validation.

The two Boolean approaches show similar results, as expected. Also, the Boolean approaches could reproduce many features shown by Kim *et al.*. In fact, the temporal order of the events was the same as in the more sophisticated differential equation model. Quantitative effects like concentrations or the strength of the feedback are of course beyond the scope of the Boolean approaches.

Clearly, the main advantage of the proposed method over the classical approach is the straight forward model definition using curated reaction network databases. This gain is even more important in the context of the fast growing amount of interactions available through these databases. In fact, while it was possible to create the Boolean logic of our sample network by hand, such a task will be tedious when talking about databases containing several thousand reactions.

Still, our method comes with a number of drawbacks which are common to all Boolean approaches. Concentrations and interaction strengths cannot be fully covered by *TRUE* and *FALSE* values. The usage of two Boolean variables per compound, i.e. inactive, active and activated but depleted allows us to model some concentration dependent kinetic effects like indirect inhibition by depletion. However, the general problem of the very crude representation of reality remains. For example antagonistic effects like phosphorylation and dephosphorylation will lead to intermediate compound activity which has to be mapped to either *TRUE* or *FALSE* in a biologically meaningful way.

Treatment of such effects is, however, still possible in Boolean approaches by choosing the right network wiring. For the above

mentioned phosphorylation/dephosphorylation cycles which are a common motif in signaling networks, the one or the other step is usually more important for the signal transduction. In case of unspecific phosphatases it can be assumed that their regulation occurs external to the pathway under investigation. Hence removing the dephosphorylation step will mostly yield the desired result. Furthermore, the function of the phosphatases especially in MAP kinase cascades is rather not to deactivate proteins but to keep the activity at a certain operation point which is optimal for the pathway function. Clearly, this kind of model refinement cannot be done in an automated fashion.

For future improvements of the automatic import of Boolean networks from databases, the curation of mechanistic databases should become sensitive for the interplay of topology and kinetics. The databases may indicate the kinetic mode a reaction usually operates in, e.g. when it is known that a reaction depletes the active form of a compound this could be marked by a "inhibition" flag.

With the size of the network also the number of errors can be expected to increase in the data. While our method assumes correct network data in first place, this can be used to actually verify the network information by comparing the simulation results to experimental results. With methods for model extension and prediction of initial conditions (Handorf *et al.*, 2008; Christian *et al.*, 2009) which are based on the concept of scopes, deficits of those networks may be identified and corrections can be proposed.

To summarize, the presented method facilitates usage of mechanistic electronic databases by defining large parts of the network logic from the topology. The difficulties of predicting the exact dynamical behavior of the biological system due to its kinetic parameters has been discussed and possible means of model refinement have been proposed. Hence, mechanistic databases and automatic import of these has the potential to push the development of Boolean network models of signaling and regulatory networks, although further review of the network data, in terms of a functional curation, may be necessary.

REFERENCES

Bonarius, H. P. J., Schmid, G., and Tramper, J. (1997). Flux analysis of underdetermined metabolic networks: The quest for the missing constraints. *Trends Biotech.*, **15**, 308–314.

Christian, N., May, P., Kempa, S., Handorf, T., and Ebenhöh, O. (2009). An integrative approach towards completing genome-scale metabolic networks. *Mol Biosyst*.

de Jong, H. (2002). Modeling and simulation of genetic regulatory systems: a literature review. *J Comput Biol*, **9**(1), 67–103.

Edwards, J. and Palsson, B. (2000). The escherichia coli mg1655 in silico metabolic genotype: Its definition, characteristics, and capabilities. *Proc Natl Acad Sci USA*, **97**, 5528–5533.

Fisher, J. and Henzinger, T. A. (2007). Executable cell biology. *Nat Biotechnol*, **25**(11), 1239–1249.

Genrich, H., Küffner, R., and Voss, K. (2001). Executable petri net models for the analysis of metabolic pathways. *Int J STTT*, **3**, 394–404.

Handorf, T., Ebenhöh, O., and Heinrich, R. (2005). Expanding metabolic networks: Scopes of compounds, robustness, and evolution. *J. Mol. Evol.*, **61**, 498–512.

Handorf, T., Christian, N., Ebenhöh, O., and Kahn, D. (2008). An environmental perspective on metabolism. *J Theor Biol*, **252**(3), 530–537.

Heinrich, R., Neel, B. G., and Rapoport, T. A. (2002). Mathematical models of protein kinase signal transduction. *Mol Cell*, **9**(5), 957–970.

Kanehisa, M., Goto, S., Furumichi, M., Tanabe, M., and Hirakawa, M. (2010). Kegg for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res*, **38**(Database issue), D355–D360.

Kauffman, S. (1986). Autocatalytic sets of proteins. *Journal of Theoretical Biol.*, **119**(1), 1–24.

Kauffman, S. A. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *J Theor Biol*, **22**(3), 437–467.

Kim, D., Rath, O., Kolch, W., and Cho, K.-H. (2007). A hidden oncogenic positive feedback loop caused by crosstalk between wnt and erk pathways. *Oncogene*, **26**(31), 4571–4579.

Klamt, S., Saez-Rodriguez, J., and Gilles, E. D. (2007). Structural and functional analysis of cellular networks with cellnetanalyzer. *BMC Syst Biol*, **1**, 2.

Kofahl, B. and Wolf, J. (2010). Mathematical modelling of wnt/-catenin signalling. *Biochem Soc Trans*, **38**(5), 1281–1285.

Matthews, L., Gopinath, G., Gillespie, M., Caudy, M., Croft, D., de Bono, B., Garapati, P., Hemish, J., Hermjakob, H., Jassal, B., Kanapin, A., Lewis, S., Mahajan, S., May, B., Schmidt, E., Vastrik, I., Wu, G., Birney, E., Stein, L., and D'Eustachio, P. (2009). Reactome knowledgebase of human biological pathways and processes. *Nucleic Acids Res*, **37**(Database issue), D619–D622.

Novre, N. L., Hucka, M., Mi, H., Moodie, S., Schreiber, F., Sorokin, A., Demir, E., Wegner, K., Aladjem, M. I., Wimalaratne, S. M., Bergman, F. T., Gauges, R., Ghazal, P., Kawaji, H., Li, L., Matsuoka, Y., Villger, A., Boyd, S. E., Calzone, L., Courtot, M., Dogrusoz, U., Freeman, T. C., Funahashi, A., Ghosh, S., Jouraku, A., Kim, S., Kolpakov, F., Luna, A., Sahle, S., Schmidt, E., Watterson, S., Wu, G., Goryanin, I., Kell, D. B., Sander, C., Sauro, H., Snoep, J. L., Kohn, K., and Kitano, H. (2009). The systems biology graphical notation. *Nat Biotechnol*, **27**(8), 735–741.

Rapoport, T. A., Heinrich, R., Jacobasch, G., and Rapoport, S. (1974). A linear steady-state treatment of enzymatic chains. a mathematical model of glycolysis of human erythrocytes. *Eur J Biochem*, **42**(1), 107–120.

Reddy, V., Liebman, M., and Mavrovouniotis, M. (1996). Qualitative analysis of biochemical reaction systems. *Comput. Biol. Med.*, **26**, 9–24.

Rizzi, M., Baltes, M., Theobald, U., and Reuss, M. (1997). In vivo analysis of metabolic dynamics in saccharomyces cerevisiae: Ii. mathematical model. *Biotechnol Bioeng*, **55**(4), 592–608.

Saez-Rodriguez, J., Alexopoulos, L. G., Epperlein, J., Samaga, R., Lauffenburger, D. A., Klamt, S., and Sorger, P. K. (2009). Discrete logic modelling as a means to link protein signalling networks with functional analysis of mammalian signal transduction. *Mol Syst Biol*, **5**, 331.

Samaga, R., Saez-Rodriguez, J., Alexopoulos, L. G., Sorger, P. K., and Klamt, S. (2009). The logic of egfr/erbb signaling: theoretical properties and analysis of high-throughput data. *PLoS Comput Biol*, **5**(8), e1000438.

Schaefer, C. F., Anthony, K., Krupa, S., Buchoff, J., Day, M., Hannay, T., and Buetow, K. H. (2009). Pid: the pathway interaction database. *Nucleic Acids Res*, **37**(Database issue), D674–D679.

Schoeberl, B., Eichler-Jonsson, C., Gilles, E. D., and Mller, G. (2002). Computational modeling of the dynamics of the map kinase cascade activated by surface and internalized egf receptors. *Nat Biotechnol*, **20**(4), 370–375.

Schuster, S. and Hilgetag, C. (1994). On elementary flux modes in biochemical reaction systems at steady state. *J. Biol. Syst.*, **2**, 165–182.

Schuster, S., Fell, D., and Dandekar, T. (2000). A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nature Biotechnology*, **18**, 326–332.

Teusink, B., Passarge, J., Reijenga, C. A., Esgalhado, E., van der Weijden, C. C., Schepper, M., Walsh, M. C., Bakker, B. M., van Dam, K., Westerhoff, H. V., and Snoep, J. L. (2000). Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? testing biochemistry. *Eur J Biochem*, **267**(17), 5313–5329.

Wolf, J. and Heinrich, R. (2000). Effect of cellular interaction on glycolytic oscillations in yeast: a theoretical investigation. *Biochem J*, **345 Pt 2**, 321–334.

Zi, Z. and Klipp, E. (2007). Constraint-based modeling and kinetic analysis of the smad dependent tgf-beta signaling pathway. *PLoS One*, **2**(9), e936.