Composition of Signaling Pathway Models and its Application to Parameter Estimation

Geoffrey Koh[1], David Hsu[2], P.S. Thiagarajan[2]

[1] NUS Graduate School for Integrative Sciences and Engineering. Email: g0306427@nus.edu.sg

[2] Department of Computer Science, National University of Singapore. Email: {dyhsu, thiagu}@comp.nus.edu.sg

Abstract

To gain a systems-level understanding of signaling pathways, we need dynamical models that can capture the kinetics of the pathway elements and their interactions. Building these models requires both the network structure and rate parameters to be known. Estimating the parameters by fitting the model to data is a challenging problem due to the high-dimensional search space. Furthermore, model construction is an incremental process, due to new players being discovered and additional experimental data on the known players being made available. Hence we have developed a methodology that allows pathway models to be easily refined and expanded. We use the probabilistic graphical model known as ***factor graphs*** to model the individual pathway components. Upon composition and integration, the inference technique of belief propagation can be used to reconcile parameter estimates and obtain a globally consistent model.

Pathway Components

Pathway components can arise in the following ways:

- Pathway decomposition to reduce the search space for parameter estimation of large pathway models.
- Cross-talk interaction being discovered between different signaling pathways.
- Additional signaling elements being elucidated.

The dynamics of the pathway can be described by a system of Ordinary Differential Equations (ODE) where the *ith* equation has the form

$$\dot{\mathbf{x}}_i = f_i(\mathbf{x}(t), \mathbf{p})$$

x(t) is a vector-valued function describing the concentration levels of molecular species at time t and **p** is the set of pathway parameters.

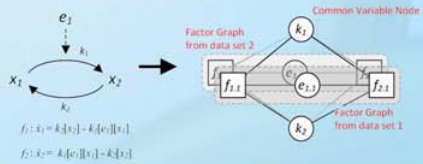
Belief Propagation

Loopy max-product belief propagation propagates the local constraints globally. Upon convergence, the variable nodes of the factor graph contain the maximum a posteriori distributions of the parameters. Assignment based on this distribution yields the most likely values for the parameters.

Composition and Integration

Each pathway component can be represented as a factor graph. New data sets on an existing pathway model can also be modeled as a separate factor graph.

The factor graphs are combined by fusing the new and existing graphs at their common variable nodes. Belief propagation will ensure that all the data / pathway components are taken into account when computing the final distributions for the parameter values.



Results

We tested this approach on a simplified model of the Akt-MAPK pathway

- 36 molecule species with 12 observed species
- 42 unknown parameters
- 4 sets of pseudo-experimental data to be introduced incrementally

Factor Graph

A factor graph is an undirected bipartite graph consisting of variable nodes and factor nodes. It captures the local dependencies between the parameters and enzymes.

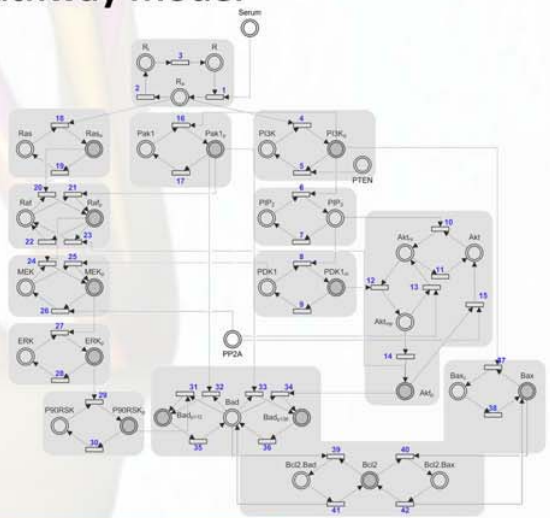
A simple reaction and its factor graph

The nodes are each associated with a probability distribution.

Variable Nodes - A probability distribution (belief) that is uniform for unknown parameters.

Factor Nodes - A joint distribution that captures the dependencies of the factor node on the variable nodes, as specified in the ODEs. The table is built by ***sampling*** the values of the parameters and scoring them against experimental data.

Pathway Model



	BP	SRES	GA
1 Dataset	**0.412**	0.483	61.96
2 Datasets	**1.548**	1.356	17.38
3 Datasets	**1.250**	3.020	263.55
4 Datasets	**0.203**	2.040	46.76

Performance comparison of **BP** (Belief Propagation), **SRES** (Evolutionary Strategies with Stochastic Ranking) and **GA** (Genetic Algorithms). Scores are the weighted mean square difference between experimental data and simulated profiles obtained using the estimated parameters. Smaller scores are better.

Conclusion

As signaling pathway models get larger, a modular approach for construction and parameter estimation where the components are elucidated separately and composed systematically is essential. We have adopted a probabilistic technique to pathway modeling which not only aids us for this purpose, but can also handle partial and noisy data well. Future works include improving the belief propagation algorithm, as well as implementing a "guided sampling" approach, so as to improve the results by focusing sampling on the more promising regions of the parameter search space.