Journal of Bioinformatics and Computational Biology
Vol. 7, No. 6 (2009) 1013–1029

Imperial College Press
www.icpress.co.uk

INFERENCE OF GENE REGULATORY NETWORKS USING BOOLEAN-NETWORK INFERENCE METHODS

GRAHAM J. HICKMAN

School of Biosciences, University of Nottingham
Sutton Bonington, LE12 5RD, United Kingdom
stxgh@nottingham.ac.uk

T. CHARLIE HODGMAN

Centre for Plant Integrative Biology
Multidisciplinary Centre for Integrative Biology
School of Biosciences, University of Nottingham
Sutton Bonington, LE12 5RD, United Kingdom
charlie.hodgman@nottingham.ac.uk

Received 8 June 2009
Revised 14 August 2009
Accepted 15 August 2009

The modeling of genetic networks especially from microarray and related data has become an important aspect of the biosciences. This review takes a fresh look at a specific family of models used for constructing genetic networks, the so-called Boolean networks. The review outlines the various different types of Boolean network developed to date, from the original Random Boolean Network to the current Probabilistic Boolean Network. In addition, some of the different inference methods available to infer these genetic networks are also examined. Where possible, particular attention is paid to input requirements as well as the efficiency, advantages and drawbacks of each method. Though the Boolean network model is one of many models available for network inference today, it is well established and remains a topic of considerable interest in the field of genetic network inference. Hybrids of Boolean networks with other approaches may well be the way forward in inferring the most informative networks.

Keywords: Boolean; inference; genetic networks.

1. Introduction

Biological systems and the processes that occur within them, be they genetic, metabolic or otherwise can be considered as biological networks.[1–3] Modern advances in technology such as the microarray allow a vast amount of information on these natural biological networks to be collected, covering thousands of genes[4] under various conditions and over different periods of time or even species.[5]

By using this wealth of information combined with modern computational techniques, biological networks can be reconstructed from the expression data and be modeled *in silico*.[6]

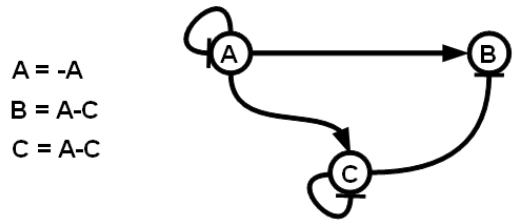
There are many advantages to be gained from the reconstruction and *in silico* modeling of biological systems. The researcher may use these networks to gain a greater understanding of how the biological systems operate and how they are regulated.[3,7] The modeling of biological networks can be seen as part of the move towards a holistic view of biology, incorporating a diverse range of data to build a model.[8,9] Once constructed and verified, the researcher can go on to use these network models to provide leads for further laboratory analysis as well as to predict biological responses.[10,11] Perhaps avoiding lengthy wet laboratory work on the organism to achieve a similar result, this way of modeling forms part of the "integrative systems biology" approach to the biological sciences.[12]

One arm of modeling is the reconstruction of genetic networks from experimental data such as time-series expression data from microarrays,[13,14] as well as chromatin immunoprecipitation data,[15] and using the information gained from experimental work to determine the interactions between the genes represented in the data.[16]There are many inference methods that can be used to construct these genetic networks including single value decomposition,[17,18] neural networks,[19] and differential equations.[20]

In addition to genetic inference methods there are a range of genetic network models that can be used to represent genetic networks. These can include but are not limited to Bayesian networks,[21–24] symbolic models,[25,26] and nonlinear vector autoregressive models.[27,28] Another such group for representing genetic networks are the so-called Boolean networks.

A Boolean network as demonstrated in Fig. 1, consists of a set of nodes connected with edges; each node corresponds to a gene and is a Boolean variable.[29] These Boolean variables have one of two values like "true or false" or, in describing the activity of genes, "on and off" (activated or inactivated). For the purposes of computational modeling and in the terms of a programming language the two variables can be represented as 0 and 1.[10] Within the Boolean network the state of the nodes (Boolean variables) and the transitions they can make are determined by the states of the other nodes (variables) in the network and the Boolean logic functions governing each node.[3]

Though Boolean networks have been a topic of considerable interest in the area of network inference, few reviews are available to draw conclusions together. The earliest reviews such as S. Huang (1999)[30] appeared in the 1990s, with reviews covering the development of inference methods and networks for genetic network modeling over the following years.[7,8] Of the latest reviews, Hecker *et al.* (2009) and Karlebach & Shamir (2008) provide only a general overview of the Boolean network model and the methods to infer the topology of Boolean networks.[31,32] Bornholdt (2008) produced a detailed review of the Boolean network model, with only minor detail regarding inference methods.[33] In consequence, this review provides a description of the different categories of Boolean networks, a study of



T			T+1		
A	B	C	A	B	C
0	0	0	1	0	0
1	0	0	0	1	1
0	1	1	1	0	0
1	0	0	0	1	1

Fig. 1. Representation of a simple Boolean network with three nodes, five edges and the functions for each node (i.e. 0 or 1 for "off" and "on"). Arrows and bars correspond to activation and inhibition, respectively. From an initial state where no genes are being expressed (shown on the first row at time T), the network input-output table shows that over successive time steps ($T- > T+1$, the $T+1$ becomes T on the next row), the network will repetitively cycle node B on and off. All other initial states lead to this same cycle (shown below).

the methods used to infer such networks and prospects for their role in network inference generally.

2. Boolean Network Categories

2.1. *Random boolean networks*

The classical model of the Boolean network in genetic modeling is the Random Boolean Network (RBN). These were proposed by Kauffman for the modeling of genetic regulatory networks;[34] they have also been referred to as Kauffman networks or N-K models.[10] RBNs are referred to as such because the connections between the nodes and the logic functions that govern them are initially assigned randomly before becoming fixed or quenched during operation.[10] The characteristics of the RBN are that the connections and functions assigned to each of the nodes are randomly selected from a group of all possible networks and function combinations. The number possible network combinations a RBN can have is represented in Eq. (1),[10] where N is the number of nodes and K the number of connections between the nodes. RBNs are an example of discrete dynamic systems where the states of the Boolean variables and time are discrete, though multi-state and hybrid models have been developed to compensate for this feature.[35]

$$(2^{2^k} N/(N-K))^N. \tag{1}$$

Many limitations have been identified with the original model; firstly the huge number of possible network combinations requires a lot of computational power,

which generally means that only a small representation of the whole network can be practically studied. The huge size of the network possibilities can also increase variance and make statistical analysis of the network difficult.[10] Other complexity-related problems identified with inferring Boolean networks in general include the consistency problem, which tries to find a network that fits all observations,[36] and the best-fit extension problem, when an inference method attempts to determine a network with inconsistent data with a minimum number of errors, though these problems were later determined to be polynomial-time solvable.[37] Despite the limitations, RBNs remain popular as network models and have been applied to understanding *Drosophila* embryonic development[3,38] and the yeast transcriptional network.[39]

2.2. *Asynchronous and temporal boolean networks*

A range of Boolean networks have been developed in addition to the classical RBN. These networks have been designed to overcome some of the perceived limitations of the classical RBN model and to build in new functions. One such group are the asynchronous Boolean networks which were designed as a way of avoiding the synchronous nature of the classical Boolean network.[40] Asynchronous Boolean networks update individual nodes randomly and asynchronously. As a result and as demonstrated in Fig. 2, they can avoid cyclic attractors though they may still contain point attractors (see Sec. 3 below). These networks may also be referred to as serial networks and synchronous (classical) networks as parallel networks; the attractors found within the network can vary depending upon the update method.[41]

It has been noted that asynchronous networks require greater computing resources to process than synchronous networks, owing to the increased complexity of the network. A hybrid synchronous–asynchronous model has been proposed which allows the determination of the asynchronous network properties more efficiently than an asynchronous model alone.[42]

Temporal Boolean networks are another variation, where the state of the network is not just determined by the previous state but the states preceding the previous state as well. This model is designed to take into account gene activity that takes place over more than one time period.[43]

2.3. *Probabilistic boolean networks*

Probabilistic Boolean Networks (PBNs), a stochastic extension of the classical RBN model, are one of the latest additions to the Boolean family, being introduced by Shmulevich *et al.* in 2001.[44] PBNs have been described as a collection of Boolean networks,[45,46] and have been developed as a means of overcoming the deterministic nature of the classical Boolean networks.[47] By referring to the example network of Fig. 1, Fig. 3 depicts the network possibilities (and associated probabilities) available for a PBN. Every node may influence itself and every other node by some probability. In this way, PBNs are designed to deal with uncertainty.[44]

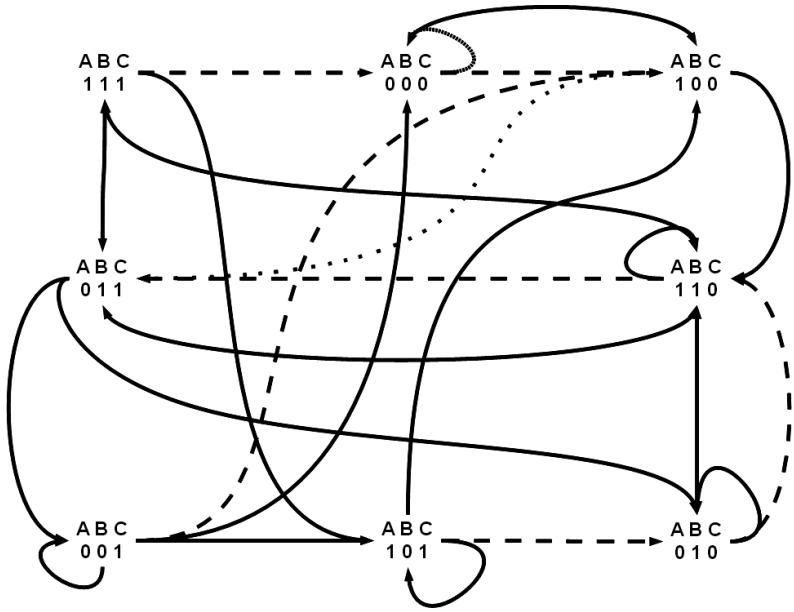


Fig. 2. Network wiring diagram demonstrating synchronous and asynchronous update methods. Each node represents a different state of the same network, i.e. the different combinations of a gene being expressed or not. Dashed lines represent state transitions possible in the synchronous network and solid lines the asynchronous network. While it may appear complex at first glance it is a representation of the same $N = 3$ network shown in Fig. 1. The dotted line between states 011 and 100 represents the cyclic attractor found within the synchronous network, circumvented by the additional state transitions possible in the asynchronous network.

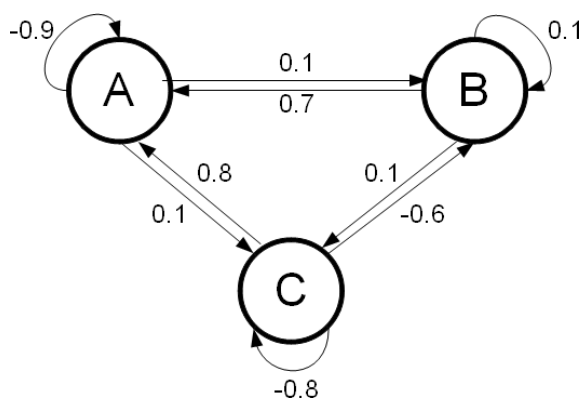


Fig. 3. Representation of a PBN, where the functions governing each node are selected from a number of possible functions depending on probability. In this case, negative values represent repression. This essentially allows the PBN over time to exist as a number of different Boolean networks.

By being able to cope with a degree of uncertainty, PBNs gain several advantages over the classical Boolean model. They are not deterministic, can deal with noise,[8] have greater flexibility, and include the ability to incorporate known biological information.[44] While providing the improvements noted above, PBNs retain all the advantages of the classical Boolean network model, such as its conceptually simple binary rule–based modeling system.[47]

PBNs have been found to be comparable with dynamic Bayesian networks,[44,48] though when inference methods are compared directly, PBNs were found to perform worse in terms of identifying interactions between genes but were faster to run.[49] A problem for PBNs as with the earlier RBNs is the large number of possible connections and function combinations, making the amount of processing power and time required very demanding on computing resources for large numbers (20+) of genes. This problem is not helped by the added complexity of probabilistic Boolean networks.[50,51]

3. Network Dynamics

An important feature of Boolean networks is that they are dynamic and that the running network can enter into an *attractor state*.[6] These are states where the network progresses to a point where it will then cycle through a set of stable states, this being called a *cyclic attractor*.[52] A biological example could be the gene expression changes seen in circadian rhythms. In addition to cyclic attractors, singleton or small attractors are possible. These are stable states that a Boolean network can reach when the network results in a single state, such as erythrocyte formation.[52] An example of what the network topology for both cyclic and singleton attractors could look like can be found in Fig. 4.

Other features impacting the dynamics of Boolean networks include positive and negative feedback loops. The identification of features such as attractor states and feedback loops is important because it yields information on the dynamics of the network and provides information on the regulation of the genes represented in the network and cells' expression states.[53] For example, networks containing greater

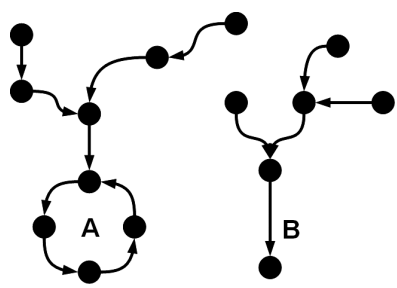


Fig. 4. Representations of cyclic and singleton attractors, each node represents a network state (the different combination of 0 or 1 as shown in Fig. 2. Network A progresses to a cyclic attractor where the network cycles though a series of states (as occurs in Fig. 1), while network B progresses to a point attractor or single state.

numbers of independent negative-feedback loops have been found to demonstrate more random behavior.[54]

A variety of inference algorithms have been developed for finding singleton attractors within Boolean networks. Several examples of algorithms which enable singleton attractors to be identified without looking at all the possible states of the network have been published.[52] These algorithms, such as the basic recursive algorithm which identifies singleton attractors by looking at the gene activity profiles, were found to be simple and capable but only applicable to RBNs with less than a hundred genes. There are now many different algorithms available for identifying attractor states within Boolean networks.[55]

Work has also been conducted into the dynamics of PBNs.[50] Their nondeterministic nature allows them to cycle through a number of possible attractors, unlike RBNs.[56] Algorithms have been built to identify the steady-state distributions of PBNs,[57] using Monte Carlo and later power-simulation methods.[58]

4. Inference Methods

A variety of different computer algorithms have been developed to infer genetic networks in the form of Boolean networks and identify features within them, such as attractor states.[55,59] One aspect of all Boolean-network inference is that the experimental data from, for example, time-series microarray experiments must first be processed to binary, since the Boolean functions of the networks can only deal with binary data.[60]

4.1. *REVEAL*

Boolean-network inference was started with REVEAL (REVerse Engineering ALgorithm). It was developed and proposed in 1998 by Shoudan Liang and written in the C programming language.[61] The algorithm infers RBNs from large volumes of expression data. The algorithm operates via a stepwise process that uses mutual information analysis of the state transition (expression) data to determine the connections between the nodes (genes), as shown in Fig. 5. The algorithm was initially only able to function with synchronous Boolean networks, though it was envisioned to be expanded to multi-state networks. The REVEAL algorithm was also found to have difficulties in dealing with the noise in biological systems.[7] Other disadvantages of the algorithm were that it worked best with only a low number of inputs (connections) per gene (K being from 1–3, which is not so far from biological reality) and had a maximum practical gene (node) number of around 30–50. The reason for this limitation is that the algorithm operates by brute force, examining all possible input configurations to determine the correct input for each gene, with more than three inputs the number of combinations becomes too large to compute with even a small number of genes. Initially larger numbers of connections and genes were expected to be handled with increased processing and multi-threading capabilities.[61]

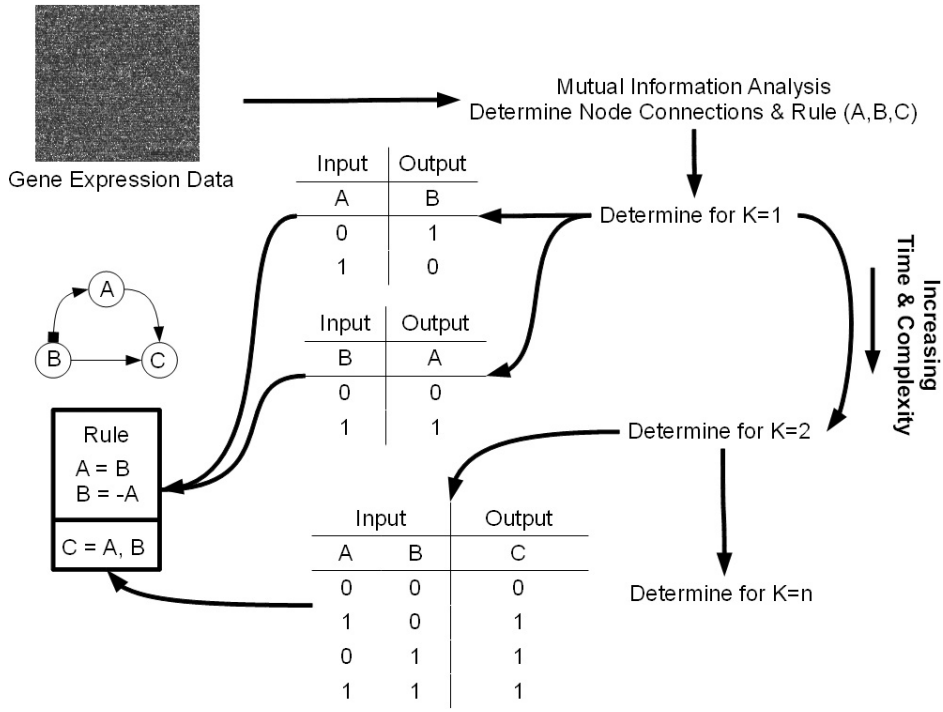


Fig. 5. Graphical representation of the REVEAL network inference method. The algorithm will initially attempt to explain the expression of each gene on the basis of a single connection ($K = 1$) to another gene. For the remaining genes, the process is repeated with the value of K incremented by one (i.e. $K = 2, 3$, etc.). In the figure elements A.B can be explained for $K = 1$ but element C cannot and requires an additional step.[61]

4.2. *BOOL*

A mathematically simpler algorithm than REVEAL was proposed by Akutsu *et al.* (1999).[62] This algorithm uses a similar principle to REVEAL in that it infers genetic networks from state transition information (input/output gene expression pairs) using a Boolean model. As with REVEAL, the algorithm demonstrated that it was possible to infer a Boolean network by using only a small amount of transition information, for example a network the size of 100,000 genes could be identified from as few as 100 state transition pairs.[62] The only problem with the network is that, although it is a simpler algorithm, it is inefficient and more time and memory intensive than even the original REVEAL algorithm. It was also still bound by the problems of REVEAL in that it is limited to only small networks with low numbers of inputs per gene.[62]

Akutsu and co-workers went on to develop the BOOL family of inference algorithms. BOOL-2 is capable of dealing with noisy data, but was found to require more expression-profile data to identify a network than BOOL and REVEAL (90–140

expression patterns for a 160-node network) in the presence of noise.[63] However, by being able to deal with noise, the algorithm was more suitable for application with real expression data.

An important issue surrounding the study of Boolean networks (and modeling in general) is the question of using real as opposed to artificially generated data when testing algorithms. Modeling should ideally be tested and verified using real data.[31] Hakamada *et al.* reconstructed networks using 21 genes from actual expression data in the study of *S. cerevisiae*.[64] Hakamada *et al.* went on to test a Boolean inference method against the known biological information found in the KEGG database.[65] The inference method was able to reconstruct many of the interactions identified from the KEGG database in addition to identifying known interactions not shown in KEGG but supported elsewhere.

4.3. *Methods reconstructing PBNs*

In addition to the classical RBNs, a range of inference methods have become available to generate newer Boolean network variations. Cotta and Troya (2004) presented a reverse-engineering approach for noisy, temporal Boolean networks using evolutionary algorithms and demonstrated it with a 32-node network.[66] Martin *et al.* have proposed an algorithm based on Akutsu *et al.* (2000), but instead of a single optimized network the algorithm generates several likely possibilities and can be considered as a simple PBN.[67]

Legacy software like REVEAL has now been used as a benchmark for the newer algorithms being developed. Zhao *et al.* (2006) have developed an inference algorithm combining a PBN model with the minimum descriptive length (MDL) principle.[68] The method incorporating MDL has demonstrated a speed improvement over REVEAL of more than tenfold for a network of 20 nodes, 30 edges and 100 time points. It was also easily scalable to larger networks with hundreds of nodes.[68] One study, however, has questioned the accuracy of the REVEAL and BOOL algorithms, describing the ability of these algorithms to recognize regulation between genes as no better than chance.[69] Further studies would resolve this matter, but other advances make it unnecessary.

4.4. *Improving inference methods*

Methods to improve the accuracy of Boolean inference algorithms have recently been made. One method of improvement is the incorporation of known biological data into networks.[70] Another method of improvement is to make Boolean inference algorithms more applicable to larger genetic networks: one proposed method based on chi-square tests has been developed, which reduced computing times by a factor of more than 70 in a 120-gene network (K being three).[71] Another method proposes two algorithms capable of reverse engineering networks and also reducing the number of plausible networks generated.[72] The improvement of inference methods

for PBNs has also been considered with the use of methods based on Monte Carlo Markov Chains for the improvement of network processing and approximation.[50,51]

4.5. *AIGNET*

Several algorithms that have been created for the inference of Boolean networks do not rely upon a single model or inference method; these are the hybrid systems. One such family of algorithms is AIGNET or Algorithms for Inference of Genetic NETworks.[73] AIGNET was proposed as a hybrid method for inferring genetic networks using algorithms generating first a static network in the form of a multi-level digraph then a dynamic model using S-systems, being power-law models consisting of sets of non linear differential equations, as shown in Fig. 6.[73]

The first stage generates a binary matrix from gene-expression matrix data. This matrix is used to identify genes which interact with each other. An accessibility matrix is then used to place interdependent genes into an equivalence set. The effects of indirect affection are then removed, and the resulting skeleton matrix is used to draw the static Boolean network. The multi-level digraph approach was found to be exceptionally fast at analyzing large genetic networks — less than one second for a 10,000-gene network (300 MHz processor) — but is incapable of identifying the order of interdependent genes in an equivalence set (genes that interact and regulate each other).[74] In order to overcome this limitation, S-systems have been applied to identify the network topology of genes in an equivalence set, using continuous expression data (i.e. numerical values for the transcript levels).[74]

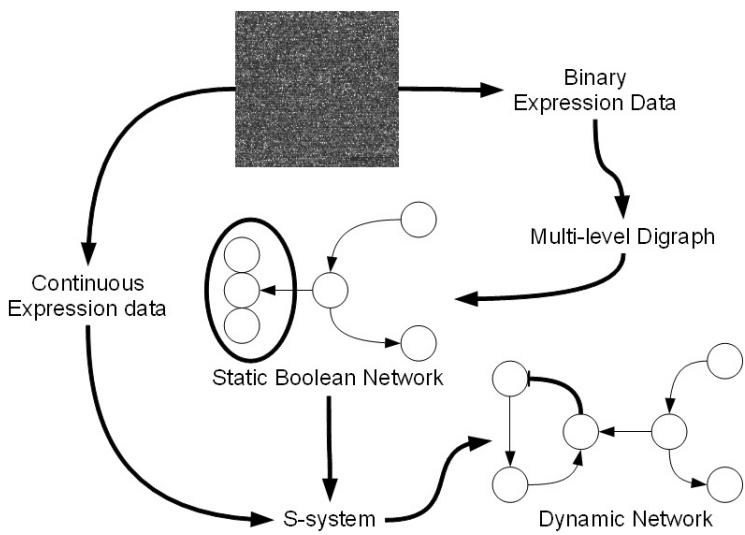


Fig. 6. Representations of the AIGNET genetic inference method. A hybrid system which uses both the Boolean network and S-system to reconstruct a genetic network. Hybrids like AIGNET can offer performance advantages in excess of each method individually.[73]

One should note that the S-system models were unable to process the large genetic networks efficiently, requiring substantial processing resources as the number of parameters increases.[75,76] Indeed, the AIGNET algorithm was able to process large genetic networks more efficiently than either of the methods applied separately,[77] and has been applied to experimental microarray data.[78] Later incarnations of the AIGNET algorithm have gone on to include Bayesian and threshold-test models.[79] Additionally the PEACE-1 algorithm has also been proposed as a method for solving the problem of equivalence created by the static Boolean network of AIGNET and generating kinetic information.[80] However, in one study, AIGNET incorrectly removed certain network connections in the special case of double-mutant knock-outs.[81]

5. Conclusions

Boolean networks are a simple and useful model for describing genetic regulatory systems, which have successfully been applied to several different organisms. From the initial Random Boolean Networks proposed by Kauffman in 1969, a wide variety of Boolean network models have been developed, along with a range of different inference methods which generate these models. A driver of Boolean network development has been to overcome some of the limitations found with the original RBNs, which can be described as simplistic, deterministic and synchronous. Early Boolean inference algorithms were also limited, able to process less than a hundred nodes (genes), and only a limited number of connections between each node. For larger genetic networks, processing the huge number of possible network combinations becomes difficult and resource intensive owing to the brute force methods employed.

In recent years, interest in the Boolean model has waned in favor of newer models and methods such as Bayesian networks and differential equations. These methods can avoid many of the noted limitations of the Boolean network.[8] Yet many of the above-mentioned limitations of the Boolean model have themselves been overcome. More sophisticated inference techniques and algorithms have also become available which allow Boolean networks to handle larger numbers of genes and shorten processing times. This ability to handle large numbers of genes is important if the researcher wishes to take advantage of the quantities of data made available by modern "omics" methods.

The simplicity of Boolean networks has been praised and the potential of the Boolean network model for the predictive modeling of incomplete networks should not be overlooked.[33] Their simplicity can be considered their greatest criticism and yet their great strength. Despite the introduction of alternatives, the Boolean network should remain as a useful representation for network inference, not just for early exploration, or when complexity is not a requirement,[82] but in concert with other network models. The studies of hybrid approaches like AIGNET have shown that individual models and inference methods used in isolation may not necessarily

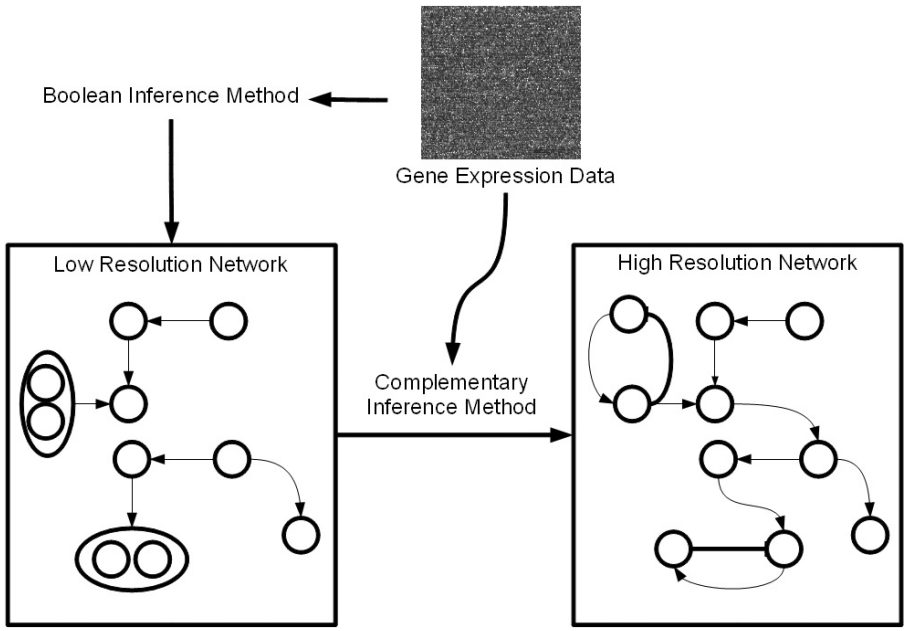


Fig. 7. Boolean inference algorithms can be used to create large low-resolution networks from partial data. When combined with a complementary inference method, a high-resolution network can be resolved. The use of multiple inference techniques can also allow the integration of more experimental data or knowledge at a later stage.

be the best approach to accurate network inference. Hybrid approaches, however, allow large datasets to be processed to a gross Boolean network topology in which much smaller subsets of the network may be refined by alternative methods such as Bayesian and S-systems, which give better resolution as they make use of a more extensive set of prior information (see Fig. 7).[83]

Acknowledgments

The authors would like to thank Z. Emmerson of the Nottingham Arabidopsis Stock Centre (NASC) for supplying the array images used in this review.

References

1. Tong AHY, Lesage G, Bader GD, Ding H, Xu H, Xin X, Young J, Berriz GF *et al.*, Global mapping of the yeast genetic interaction network, *Science* **303**:808–813, 2004.
2. Guimerà R, Amaral ALN, Functional cartography of complex metabolic networks, *Nature* **433**(7028):895–900, 2005.
3. Albert R, Boolean modeling of genetic regulatory networks, *Lec Notes Phys* **650**:459–481, 2004.
4. D'Ambrosio C, Gatta L, Bonini S, The future of microarray technology: Networking the genome search, *Allergy* **60**(10):1219–1226, 2005.

5. Borque G, Sankoff D, Improving gene network inference by comparing expression time-series across species, developmental stages or tissues, *J Bioinform Comput Biol* **2**(4):765–783, 2004.
6. D'haeseleer P, Liang S, Somogyi R, Genetic network inference: From co-expression clustering to reverse engineering, *Bioinformatics* **16**(8):707–726, 2000.
7. van Someren EP, L. F. A Wessels, Backer E, Reinders MJT, Genetic network modeling, *Pharmacogenomics* **3**(4):507–525, 2002.
8. Markowetz F, Spang R, Inferring cellular networks — a review, *BMC Bioinform* **8** Suppl. 6:S5:1–17, 2007.
9. Imoto S, Higuchi T, Goto T, Tashiro K, Kuhara S, Miyano S, Combining microarrays and biological knowledge for estimating gene networks via Bayesian networks, *Proc IEEE Comput Soc Bioinform Conf* 104–113, 2003.
10. Gershenson C, Introduction to random Boolean networks, *Workshop and Tutorial Proceedings, Ninth International Conference on the Simulation and Synthesis of Living Systems (ALife IX)*, pp. 160–173, 2004.
11. Imoto S, Savoie CJ, Aburatani S, Kim S, Tashiro K, Kuhara S, Miyano S, Use of gene networks for identifying and validating drug targets, *J Bioinform Comput Biol* **1**(3):459–474, 2003.
12. Albert R, Network inference, analysis, modeling in systems biology, *Plant Cell* **19**(11):3327–3338, 2007.
13. Hanai T, Hamada H, Okamoto M, Application of bioinformatics for DNA microarray data to bioscience, bioengineering and medical fields, *J Biosci Bioeng* **101**(5):377–384, 2006.
14. Ma PCH, Chan KCC, An effective data mining technique for reconstructing gene regulatory networks from time series expression data, *J Bioinform Comput Biol* **5**(3):651–668, 2007.
15. Zhao W, Serpedin E, Dougherty ER, Recovering genetic regulatory networks from chromatin immunoprecipitation and steady-state microarray data, *EURASIP J Bioinform Syst Biol*, **2008**:248747, doi:10.1155/2008/248747, 2008.
16. Slonim DK, From patterns to pathways: gene expression data analysis comes of age, *Nat Genet* **32** Suppl:502–508, 2002.
17. P. D'Haeseleer, Wen X, Fuhrman S, Somogyi R, Linear modeling of mRNA expression levels during CNS development and injury, *Pac Symp Biocomput* **4**:41–52, 1999.
18. Yeung MK, J. Tegnér, Collins JJ, Reverse engineering gene networks using singular value decomposition and robust regression, *Proc Natl Acad Sci USA* **99**(9):6163–6168, 2001.
19. Keedwell E, Narayanan A, Savic D, Modeling gene regulatory data using artificial neural networks, *Proc International Joint Conf Neural Networks (IJCNN* 2002), pp. 183–189, 2002.
20. Chen T, He HL, Church GM, Modeling gene expression with differential equations, *Pac Symp Biocomput* **4**:29–40, 1999.
21. Friedman N, Linial M, Nachman I, Pe'er D, Using Bayesian networks to analyze expression data, *J Comput Biol* **7**(3,4):601–620, (2000)
22. Murphy K, Mian S, Modeling gene expression data using dynamic Bayesian networks, *Technical Report*, Berkeley, 1999.
23. Imoto S, Sunyong K, Goto T, Aburatani S, Tashiro K, Kuhara S, Miyano S, Bayesian network and nonparametric heteroscedastic regression for nonlinear modeling of genetic networks, *J Bioinform Comput Biol* **1**(2):231–52, 2003.
24. Tamada Y, Bannai H, Imoto S, Katayama T, Kanehisa M, Miyano S, Utilizing evolutionary information and gene expression data for estimating gene networks with Bayesian network models, *J Bioinform Comput Biol* **3**(6):1295–1313, 2005.

25. Aiguier M, Gaston C, Le Gall P, Longuet D, Touil A, A temporal logic for input output symbolic transition systems, *Proc 12th Asia-Pac Softw Eng Conf*, pp. 43–50, 2005.
26. Mateus D, Gallois J, Comet J, Le Gall P, Symbolic modeling of genetic regulatory networks, *J Bioinform Comput Biol* **5**(2b):627–640, 2007.
27. Mukhopadhyay ND, Chatterjee S, Causality and pathway search in microarray time series experiment, *Bioinformatics* **23**(4):442–449, 2007.
28. Fujita A, Sato JR, Garay-Malpartida HM, Sogayar MC, Ferreira CE, Miyano S, Modelling nonlinear gene regulatory networks from time series gene expression data, *J Bioinform Comput Biol* **6**(5):961–979, 2008.
29. de Jong H, Modeling and simulation of genetic regulatory systems: A literature review, *J Comput Biol* **9**(1):67–103, 2002.
30. Huang S, Gene expression profiling, genetic networks, cellular states: An integrating concept for tumorigenesis and drug discovery, *J Mol Med* **77**(6), 469–480, 1999.
31. Hecker M, Lambeck S, Toepfer S, van Someren E, Guthkea R, Gene regulatory network inference: Data integration in dynamic models — a review, *BioSystems* **96**(1):86–103, 2009.
32. Karlebach G, Shamir R, Modeling and analysis of gene regulatory networks, *Nat Rev Mol Cell Biol* **9**(10):770–80, 2008.
33. Bornholdt S, Boolean network models of cellular regulation: Prospects and limitations, *J R Soc Interface* **5**(Suppl), S85–94, 2008.
34. Kauffman SA, Metabolic stability and epigenesis in randomly constructed genetic nets, *J Theor Biol* **22**(3):437–467, 1969.
35. Laubenbacher R, Stigler B, A computational algebra approach to the reverse engineering of gene regulatory networks, *J Theor Biol* **229**(4):523–537, 2003.
36. Lahdesmaki H, Shmulevich I, Yli-Harja O, On learning gene regulatory networks under the Boolean network model, *Mach Learn* **52**(1):147–167 2003.
37. Shmulevich I, Yli-Harja O, Astolam J, Inference of genetic regulatory networks under the bestfit extension paradigm, *Proceedings of the IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing (NSIP-01)*, Maryland, Baltimore, (2001b).
38. Chaves M, Albert R, Sontag ED, Robustness and fragility of Boolean models for genetic regulatory networks, *J Theor Biol* **235**(3):431–449, 2005.
39. Kauffman SA, Peterson C, Samuelsson B, Troein C, Random Boolean network models and the yeast transcriptional network, *Proc Natl Acad Sci USA* **100**(25):14796–14799, 2003.
40. Gershenson C, Classification of random Boolean networks, *Artif Life* **8**:1–8, 2002.
41. Goles E, Salinas L, Comparison between parallel and serial dynamics of Boolean networks, *Theor Comput Sci* **396**:247–253, 2008.
42. Garg A, Di Cara A, Xenarios I, Mendoza L, De Micheli G, Synchronous versus asynchronous modeling of gene regulatory networks, *Bioinformatics* **20**(17):1917–1925, 2008.
43. Silvescu A, Honavar V, Temporal Boolean network models of genetic networks and their inference from gene expression time series, *Complex Systems* **13**(1):54–70, 2001.
44. Shmulevich I, Dougherty ER, Kim S, Zhang W, Probabilistic Boolean networks: A rule-based model for gene regulatory networks, *Bioinformatics* **18**(2):261–274, 2001.
45. Ching W, Zhang S, Jiao Y, Akutsu T, Wong AS, Optimal finite-horizon control for probabilistic Boolean networks with hard constraints, *The First International Symposium on Optimization and Systems Biology (OSB'07)*:21–28, 2007.
46. Marshall S, Yu L, Xiao Y, Dougherty ER, Inference of a probabilistic Boolean network from a single observed temporal sequence, *EURASIP J Bioinform Syst Biol* **2007**:32454, doi: 10.1155/2007/32454, 2007.

47. Shmulvich I, Dougherty ER, Zhang W, From Boolean to probabilistic Boolean networks as models of genetic regulatory networks, *Proc IEEE* **90**(11):1778–1792, 2002.
48. Lähdesmäki H, Hautaniemi S, Shmulevich I, Yli-Harja O, Relationships between probabilistic Boolean networks and dynamic Bayesian networks as models of gene regulatory networks, *Signal Processing* **86**(4):814–834, 2006.
49. Li P, Zhang C, Perkins EJ, Gong P, Deng Y, Comparison of probabilistic Boolean network and dynamic Bayesian network approaches for inferring gene regulatory networks, *BMC Bioinformatics* **8**(Suppl 7):S13, 2007.
50. Shmulevich I, Gluhovsky I, Hashimoto RF, Dougherty ER, Zhang W, Steady-state analysis of genetic regulatory networks modeled by probabilistic Boolean networks, *Comp Funct Genomics* **4**(6)L:601–608, 2003.
51. Ching W, Ng MM, Fung ES, Akutsu T, On construction of stochastic genetic networks based on gene expression sequences, *Int J Neural Syst* **15**(4):297–310, 2005.
52. Zhang S, Hayashida M, Akutsu T, Ching W, Ng MK, Algorithms for finding small attractors in Boolean networks, *EURASIP J Bioinform Syst Biol* **2007**:20180, 2007.
53. Wuensche A, Genomic regulation modeled as a network with basins of attraction, *Pac Symp Biocomput* **3**:89–102, 1998.
54. Sontag E, Veliz-Cuba A, Laubenbacher R, Jarrah AS, The effect of negative feedback loops on the dynamics of Boolean networks, *Biophysical Journal* **95**(2):518–526, 2008.
55. Akutsu T, Hayashida M, Tamura T, Algorithms for inference, analysis and control of Boolean networks, *Lect Notes Comput Sci* **5147**:1–15, 2008.
56. Brun M, Dougherty ER, Shmulevich I, Steady-state probabilities for attractors in probabilistic Boolean networks, *Signal Processing* **85**:1993–2013, 2005.
57. Qian X, Dougherty ER, On the long-run sensitivity of probabilistic Boolean networks, *J Theor Biol* **257**(4):560–577, 2008.
58. Ching W, Zhang S, Ng MK, Akutsu T, An approximation method for solving the steady-state probability distribution of probabilistic Boolean networks, *Bioinformatics* **23**(12):1511–1518, 2007.
59. Ideker TE, Thorsson V, Karp RM, Discovery of regulatory interactions through perturbation: inference and experimental design, *Pac Symp Biocomput* **5**:302–313, 2000.
60. Hakamada K, Hanai T, Honda H, Kobayashi T, Identifying genetic network using experimental time series data by Boolean algorithm, *Genome Informatics* **12**:272–273, 2001.
61. Liang S, Fuhrman S, Somogyi R, REVEAL, a general reverse-engineering algorithm for inference of genetic network architectures, *Pac Symp Biocomput* **3**:18–29, 1998.
62. Akutsu T, Miyano S, Kuhara S, Identification of genetic networks from a small number of gene expression patterns under the Boolean network model, *Pac Symp Biocomput* **4**:17–28, 1999.
63. Akutsu T, Miyano S, Kuhara S, Inferring qualitative relations in genetic networks and metabolic pathways, *Bioinformatics* **16**(8):727–734, 2000.
64. Hakamada K, Hanai T, Honda H, Kobayashi T, Identifying genetic network using experimental time series data by Boolean algorithm, *Genome Informatics* **12**:272–273, 2001.
65. Hakamadaa K, Hanai T, Hondaa H, Kobayashia T, A preprocessing method for inferring genetic interaction from gene expression data using Boolean algorithm, *J Biosci Bioeng* **98**(6):457–463, 2004.
66. Cotta C, Troya JM, Reverse engineering of temporal Boolean networks from noisy data using evolutionary algorithms, *Neurocomputing* **62**:111–129, 2004.
67. Martin S, Zhang Z, Martino A, Faulon J, Boolean dynamics of genetic regulatory networks inferred from microarray time series data, *Bioinformatics* **23**(7):866–874, 2007.

68. Zhao W, Serpedin E, Dougherty ER, Inferring gene regulatory networks from time series data using the minimum description length principle, *Bioinformatics* **22**(17):2129–2135, 2006.
69. Wimberly FC, Heiman T, Glymour C, Ramsey J, Experiments on the accuracy of algorithms for inferring the structure of genetic regulatory networks from associations of gene expressions, 1: Algorithms using binary variables, *Proc IJCAI Bioinformatics Workshop 2003*, 2003.
70. Liu W, Lahdesmaki H, Dougherty ER, Shmulevich I, Inference of Boolean networks using sensitivity regularization, *EURASIP J Bioinform Syst Biol* **2008**:780541, doi: 10.1155/2008/780541, 2008.
71. Kim H, Lee JK, Park T, Boolean networks using the chi-square test for inferring large-scale gene regulatory networks, *BMC Bioinformatics* **8**:37, 2007.
72. Mehra S, Hu W, Karypis G, A Boolean algorithm for reconstructing the structure of regulatory networks, *Metab Eng* **6**(4):326–339, 2004.
73. Maki Y, Tominaga D, Okamoto M, Watanabe S, Eguchi Y, Development of a system for the inference of large scale genetic networks, *Pac Symp Biocomput* **6**:446–458, 2001.
74. Maki Y, Watanabe S, Eguchi Y, Tominaga D, Okamoto M, AIGNET: A system that infers large scale genetic networks, *Genome Informatics* **10**:253–254, 1999.
75. Maki Y, Ueda T, Okamoto M, Uematsu N, Inamura K, Uchida K, Takahashi Y, Eguchi Y, Inference of genetic network using the expression profile time course data of mouse P19 cells, *Genome Informatics* **13**:382–383, 2002.
76. Daisuke T, Horton P, Inference of scale-free networks from gene expression time series, *J Bioinform Comput Biol* **4**(2):503–514, 2006.
77. Arikawa Y, Watanabe S, Eguchi Y, Maki Y, Tominaga D, Okamoto M, AIGNET: Improvement of a system and application to the experimentally observed expression data, *Genome Informatics* **11**:291–292, 2000.
78. Arikawa Y, Takahashi Y, Watanabe S, Maki Y, Okamoto M, Eguchi Y, Inference of a gene network from the experimentally observed expression data by using AIGNET, *Genome Informatics* **12**:274–275, 2001.
79. Takahashi Y, Arikawa Y, Watanabe S, Maki Y, Aburatani S, Kuhara S, Eguchi Y, Comparison between two different genetic network inferring models on expression profiles in *S. cerevisiae*, *Genome Informatics* **13**:380–381, 2002.
80. Kikuchi S, Tominaga D, Arita M, Takahashi K, Tomita M, Dynamic modeling of genetic networks using genetic algorithm and s-system, *Bioinformatics* **19**(5):643–650, 2003.
81. Tringe SG, Wagner A, Ruby SW, Enriching for direct regulatory targets in perturbed gene-expression profiles, *Genome Biol* **5**(4):R29.1–14, 2004.
82. Schlitt T, Brazma A, Network inference, analysis, modeling in systems biology, *Plant Cell* **19**(11);3327–3338, 2007.
83. Yuan A, Chen G, Rotimi C, Genetic network analysis by quasi-Bayesian method, *J Bioinform Comput Biol* **7**(1):175–192, 2009.

Graham J. Hickman graduated from Birmingham University, UK, in 2007, having completed a degree in Biochemistry with Molecular and Cell Biology. After working in analytical chemistry at Bodycote Health Sciences for a year, he returned to full-time higher education in 2008 at Nottingham University, UK. Currently, Graham is developing novel silica surfaces by biomimetic methods and studying the effects of these surfaces on cancer cells using proteomic methods.

T. Charlie Hodgman's research career has almost exclusively involved cross-disciplinary working. For 15 years (1977–95, from ICI through several Departments in Cambridge), his research involved the simultaneous application of both laboratory and computational biological techniques to major research issues in plant, animal, medical and biotechnological sciences. Gaining an international reputation in Bioinformatics, he was recruited by GlaxoWellcome in 1995, where he both collaborated with laboratory scientists on specific pharmaceutical R&D projects and developed his own Bioinformatics R&D programme in what is now known as Systems Biology. In 2004, he moved to the University of Nottingham, UK, where he established the Multidisciplinary Centre for Integrative Biology. This centre currently has a £10 m research portfolio, which predominantly includes the BBSRC/EPSRC Centre for Plant Integrative Biology focusing on plant and crop research, though there are also animal and microbial projects. He has served on various UK and EU Research funding panels, and the steering committee of the pan-European ERA-SysBio. He has also contributed to the development of national strategies for the Research Councils, DTI and OST.