5.3 Metabolic Networks

5.4 Bayesian Networks

Let $G = (V, E)$ be a directed acyclic graph. We assume that the vertices $i \in V$ ($1 \leq i \leq n$) represent for example genes and correspond to random variables x_i. For each y_i we define a conditioned probability

$$P(x_i|\text{parent}(x_i)) = P(x_i|P_a(x_i)) \tag{5.29}$$

and the joint probability distribution

$$P(x_1, ..., x_n) = \prod_{i=1}^{n} P(x_i|P_a(x_i)) \tag{5.30}$$

We define a parameter θ to be the set of conditioned probabilities

Example 5.4.1 (Bayesian Network)
From the graph shown in figure 5.6 representing the Bayesian network we can read off the joint probability distribution

$$P(x_1, ..., x_n) = P(x_1)P(x_2)P(x_3|x_1, x_2)P(x_4|x_2)P(x_5|x_3, x_4) \tag{5.31}$$

and θ is the set of local conditional probabilities.

Let x be distributed as $x \sim f_\theta(x)$ with parameter θ. Our aim is to estimate the parameter θ, given that there are iid observations $\{x_i\}$ of the random variable x. We start the discussion with an example.

Example 5.4.2 (Unfair dice)
Let $\{x_i\}$ be a sequence of results of rolling a dice such that $x_i = 1$ denotes head and $x_i = 0$ tail. The dice is unfair so that we have for the probability that the random variable takes on the result $x_i = 1$ is $P(x_i = 1) \neq 1/2$. Define the parameter $\theta = P(x_i = 1)$. We need an estimate for θ. Of course, we know that we are dealing with a binomial distribution (see section 8.2) and

$$P(x = 1) = p, \quad P(x = 0) = q, \quad q = 1 - p \tag{5.32}$$

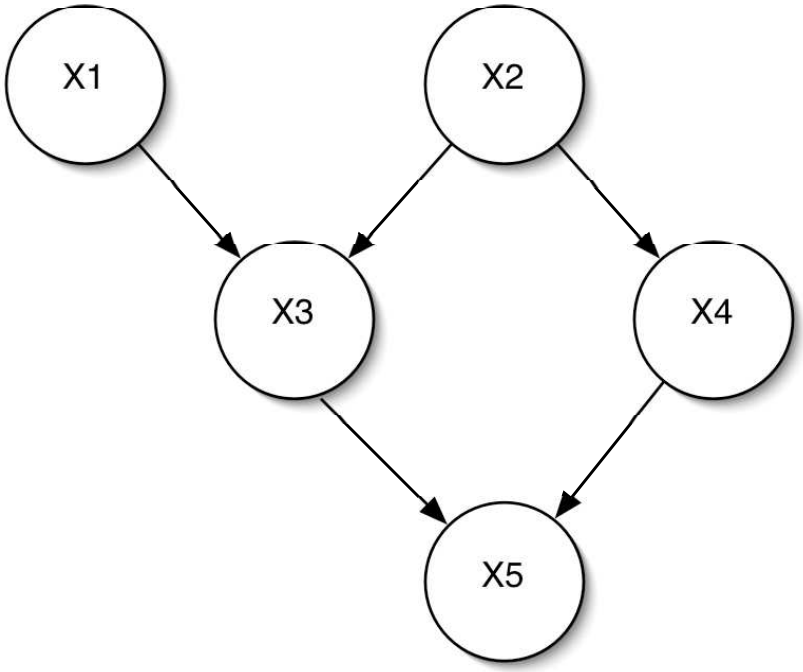


Figure 5.6: A simple Bayesian network

How many successful realizations k do we have after we have tried n-times? This is given by the binomial distribution (recall that the reverse question gives the negative binomial distribution, see appendix).
Without any prior information we can use the maximum likelihood method to estimate the parameter θ. Let $x_1, ..., x_n$ denote the observations, with $H + T = n$ where H is the number of trails we have obtained head, and T the number of trails we have obtained tail. Then

$$\hat{\theta} = \text{argmax} f_\theta(y_i) \quad i = 1, ...n \tag{5.33}$$
$$= \text{argmax} \theta^H (1-\theta)^T \tag{5.34}$$
$$= \frac{H}{H+T} \tag{5.35}$$

which corresponds to the empirical frequency of the sample for the event H. Even if we have prior information in the sense that we have a previous set of experiments where the number of heads was 5 and the number of tails was 10, we would

obtain

$$\hat{\theta} = \frac{5+H}{5+H+10+T} \tag{5.36}$$

which is irrelevant if the number of total experiments $n = T + H$ goes to infinity. But now suppose we are not sure that we measured 5 heads and 10 tails and can only tell that presumably the probability with which we obtained head was $1/4$. For this uncertainty we assume a prior distribution for $theta$ $P(\theta)$

$$P(\theta) = \beta(5,10) = \frac{\Gamma(5)\Gamma(10)}{\Gamma(5+10)} \tag{5.37}$$

where we have assumed that the prior is distributed as a *beta distribution*

$$\beta(p,q) \sim \frac{\Gamma(5)\Gamma(10)}{\Gamma(5+10)} = \int_0^1 x^{p-1}(1-x)^{q-1}dx \tag{5.38}$$

Assume that we make new observations. We use the Bayesian law to compute the posterior distribution

$$P(\theta|x) = \frac{P(x|\theta)P(\theta)}{P(x)} \tag{5.39}$$

$$= \frac{P(x|\theta)P(\theta)}{\int P(x|\theta)P(\theta)d\theta} \tag{5.40}$$

where $P(x|\theta)$ is the likelihood function. In the new experiment we find 50 heads and 50 tails so that we obtain for the likelihood function

$$P(x|\theta) = \theta^{50}(1-\theta)^{50} \tag{5.41}$$

from which we obtain the posterior distribution

$$P(\theta|x) = \frac{\theta^{59}(1-\theta)^{69}}{\int \theta^{59}(1-\theta)^{69}d\theta} \tag{5.42}$$

Note that the posterior function is the same as the prior function. In such a case we call the distribution conjugate. Thus the beta distribution is the conjugate to the binomial distribution.

The result of a Bayesian analysis is a posterior distribution and not a single value! This distribution can be used to make predictions.

5.5 Boolean Networks

Boolean networks are one of the simplest models for the behaviour of genomic networks. A *Boolean network* [213, 214] $G(V, F)$ is defined by a set of nodes $V = \{x_1, ..., x_N\}$ (corresponding to genes) and a list of Boolean functions $F = (f_1, ..., f_k)$ corresponding to the edges. In these models, gene expression is quantized to just two levels: ON and OFF (1 or 0). The state of a node (gene) is completely determined by the values of other nodes at time t by means of underlying logical Boolean functions. The model is represented in the form of directed graphs. Each x_i represents the state (expression) of gene i, where $x_i = 1$ represents the fact that gene i is expressed and $x_i = 0$ means it is not expressed.
The time evolution of the network is described by

$$x(t+1) = f_i(x_{j_1}(t), ..., x_{j_{k_i}}(t)) \tag{5.43}$$

where k_i denotes the connectivity of the node i. For the average connectivity we have

$$\langle K \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i \quad . \tag{5.44}$$

The list of Boolean functions F represents the rules of regulatory interactions between genes. The tuple (G, F) is one realization. In the *quenched model* one realizations is chosen at kept fixed for all times. In the *annealed model* a new realization is chosen at random at very time step.
Any given gene transforms its inputs (regulatory factors that bind to it) into an output, which is the state or expression of the gene itself. All genes are assumed to update synchronously or asynchronously in accordance with the functions assigned to them and this process is then repeated. Despite the simplicity of the model, they can capture a number of essential features of real genomic networks.
Asynchronous random boolean networks (ARBN) incorporate all the cases in which at each time point a single node is selected in order to be updated. The node to be updated can be chosen at random or according to a deterministic rule based:

- clock scheme [215, 216],

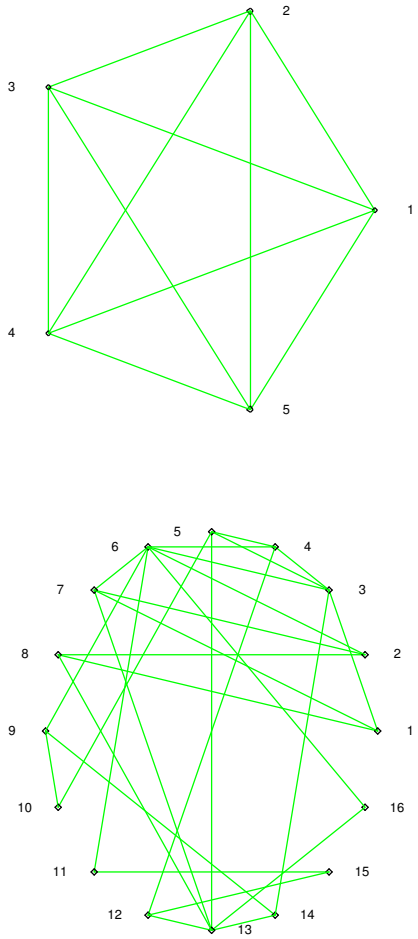
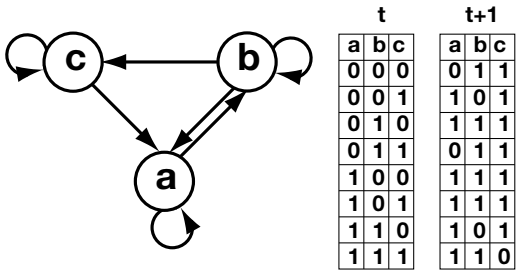


Figure 5.7: A random network with 5 nodes and 10 edges and one with 16 nodes and 28 edges



t			t+1		
a	b	c	a	b	c
0	0	0	0	1	1
0	0	1	1	0	1
0	1	0	1	1	1
0	1	1	0	1	1
1	0	0	1	1	1
1	0	1	1	1	1
1	1	0	1	0	1
1	1	1	1	1	0

Figure 5.8: A sample boolean network

- cyclic scheme [217],
- random independent scheme [218], and
- random order scheme [218]

Consider the example shown in figure 5.8. The assignment of values to nodes made in the truth table fully describes the state of the model at any given time. The change of model state over time is fully defined by the functions in F. If we assign initial values to the nodes, all further states are determined. Consider the time evolution of the states as a trajectory. Since the number of possible states is finite, all trajectories eventually end up in single *steady states*, or a *cycle of steady states*. In this context we define an *attractor* of a trajectory as a single steady state, or a cycle at the end of the trajectory. The *basin of attraction* for a specific attractor is a set of all trajectories leading to it. States in gene networks are often characterized by stability in the sense that small changes in value of a few nodes do not change the attractor.
At each time point t the system can be in one of the 2^N possible states. We have 2^{2^k} possible logic functions per node which yield

$$\text{possible} \quad \text{nets} = N_F = \left(\frac{2^{2^k} N!}{(N-k)!} \right)^N \tag{5.45}$$

The geometry of the network strongly influences the behaviour. We can distinguish between

- irregular networks

- regular networks like lattices where for example we have $|x_{j_1}(t), ..., x_{j_{k_i}(t)}| = 2d$ neighbours on a simple hypercubic lattice

From the statistical mechanics point of view we have to specify an ensemble of the coupling functions, i.e. each must be assigned a weight. For this we can consider the following options

- equal weight ensemble
- magnetization biased
- forcing functions
- additive functions

In the equal weight ensemble we assign every coupling function the same weight $1/N_F$.
In the magnetization bias situation the probability of occurrence of a coupling function is p if the result is 0 and $1-p$ if the result is 1.
In the forcing function ensemble the value of the function is determined if one of the arguments $m \in \{1, ..., k\}$ takes on a predetermined value. For example $x_m = 0$. The value of the function is not determined if $x_m = 1$.
And for the additive function ensemble we have

$$x_i(t+1) = \Theta(f_i(t)) \tag{5.46}$$

and

$$f(t) = h + \sum_{j=1}^{N} c_{ij} x_j(t) \tag{5.47}$$

where h is a bias.
Now let us look at the response of the system if we change conditions. Let

$$\Sigma_0 = \{x_1(0), ..., x_N(0)\} \tag{5.48}$$

denote the starting state and

$$\hat{\Sigma_0} = \{\hat{x_1}(0), ..., \hat{x_N}(0)\} \tag{5.49}$$

another starting state which differs from Σ_0 only in a few nodes. To describe the difference we introduce the *Hamming distance*

$$D(t) = \sum_{i=1}^{N} (x_i(t) - \hat{x}(t))^2 \tag{5.50}$$

to describe the difference in the evolution of the network, given two initial conditions. Clearly, if the network is such that any local discrepancy remains localized, then we do expect the Hamming distance to remain finite in the thermodynamic limit. If, on the other hand the discrepancy (or damage) can be propagated to almost every vertex, then we expect the Hamming distance to diverge.
Assume that we are dealing with the uniform distribution for the coupling functions. On average, a change of a single vertex will affect the argument of k functions. Hence $kD(0)$ functions are affected. Each of these is affected with a probability $1/2$ so that $D(1) = kD(0)$ and in general

$$D(t) = \left(\frac{k}{2}\right)^t D(0) \tag{5.51}$$

We can thus distinguish three phases

- chaotic phase for $k_c > 2$
- frozen for $k_c < 2$
- critical for $k_c = 2$

In the last case the state of the system will be dominated by fluctuations.
Let us now look at the case where

$$f_i = \begin{cases} 0 & \text{with probability } p \\ 1 & \text{with probability } 1-p \end{cases} \tag{5.52}$$

For a given p and a given connectivity k we will have critical values

$$k_c(p) \quad \text{and} \quad p_c(k) \tag{5.53}$$

Let

$$a(t) \equiv 1 - D(t)/N \tag{5.54}$$

be the probability that two vertices have the same value in Σ_t and $\hat{\Sigma}$. The probability that the arguments of the functions f_i have the same value is given by

$$\rho_k = [a(t)]^k \tag{5.55}$$

The overlap is the same in the next time step if the arguments of the coupling functions are identical. This occurs with probability ρ_k. The overlap is also the same if the arguments of the coupling functions are also different, which occurs with probability $1 - \rho_k$ but the values are nevertheless the same, which occurs with probability $2p(1-p)$. Altogether we have

$$\begin{aligned} a(t+1) &= 1 - (1-\rho_k)2p(1-p) && (5.56)\\ &= 1 - \frac{1-[a(t)]^k}{k_c} && (5.57) \end{aligned}$$

where we have used

$$\begin{aligned} k_c &= \frac{1}{2p(1-p)} && (5.58)\\ p_c &= \frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{1}{k}} && (5.59) \end{aligned}$$

Clearly the recursion relation has a trivial fixpoint given by

$$a^* = 1 \tag{5.60}$$

What kind of fixpoint is this? To answer this question we perturb the fixpoint slightly with δa_t. Then coming from below we have

$$1 - \delta a_{t+1} = 1 - \frac{1-[1-\delta a_t]^k}{k_c} \tag{5.61}$$

yielding

$$\delta a_{t+1} \approx \frac{k}{k_c} \delta a_t \tag{5.62}$$

The trivial fixpoint is unstable if $k/k_c > 1$. The resulting phase diagram can be seen in figure 5.9.

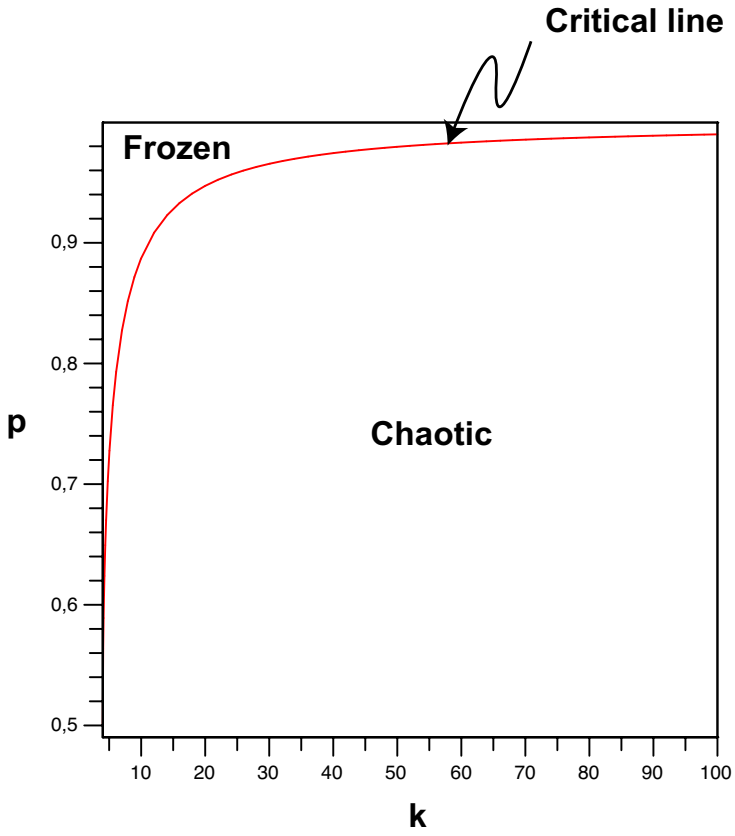


Figure 5.9: The phase diagram is only shown for $0 \leq p \leq 0.5$.

5.6 Scale-free Networks

So far we could attach a definite scale to the network. We shall now investigate networks that are scale free. For this we use the following connectivity distribution

$$P(k) = \frac{1}{\zeta(\gamma)} k^{-\gamma} \quad , \quad \gamma > 1 \quad . \tag{5.63}$$

with

$$\zeta(\gamma) = \sum_{k=1}^{\infty} k^{-\gamma} \tag{5.64}$$

being the Riemann Zeta-function.
The condition that $P(k)$ is normed requires that $\gamma > 1$. For the first moment we obtain

$$\langle k \rangle = \sum_{k=1}^{\infty} kP(k) = \begin{cases} \infty & \text{if } 1 < \gamma \leq 2 \\ \zeta(\gamma - 1)/\zeta(\gamma) & \text{if } \gamma > 2 \end{cases} \tag{5.65}$$

As done in the previous section we considered the case of the annealed model. Every element receives k inputs with probability $P(k)$

$$a(t+1) = G(a(t)) \tag{5.66}$$

and recall that $a(t) = 1 - D(t)/N$. The average probability that $k = 1, 2, ...$ controlling elements of the coupling function f are identical is

$$\mu(a) = \sum_{k=1}^{\infty} a^k P(k) \quad . \tag{5.67}$$

From this we can calculate the recursion function G

$$G(a) = 1 - 2p(1-p)[1 - \sum_{k=1}^{\infty} a^k P(k)] \tag{5.68}$$

For the fixpoint a^* of eq 5.66 we can again analyse the stability by considering $a^* + \delta a^*$. a^* is unstable iff

$$\begin{aligned} 1 &= \lim_{a\uparrow 1} \frac{dG(a)}{da} \\ &= 2p(1-p)\sum_{k=1}^{\infty} kP(k) \\ &= 2p(1-p)\langle k\rangle \quad . \end{aligned} \tag{5.69}$$

The phase transition is along this line because the fixpoint is stable for $\lim_{a\uparrow 1} \frac{dG(a)}{da} < 1$ and unstable for $\lim_{a\uparrow 1} \frac{dG(a)}{da} > 1$.

We shall now analyse the trajectory through phase space in terms of *limit cycles* and *attractors*. For this we switch to the quenched RBN. In this model the coupling functions are independent of time. For every initial condition we can follow Σ_t which surely will reach a state that was visited before. In this case it will cycle. Thus the phase space $\Omega = 2^N$ can be partitioned into cycles and attractors. An attractor can be defined as set of points from the phase space $A_t \equiv \{\Sigma_t\} \subset \Omega$ which will be mapped onto itself $A_{t+1} = A_t = A_0$. A *basin of attraction* of an attractor A_0 is a subset of Ω with

$$\text{ex } T < \infty : \Sigma_T \in A_0 \tag{5.70}$$

Let $k = N$. Assume that we start with Σ_0 and follow the trajectory. Let q_t denote the probability that the random walk is still not closed after t steps and let p_t denote the probability that the walk terminates (closes) after exactly t steps. If the trajectory is open after t steps then $t+1$ different points were visited. Hence there are $t+1$ possibilities to terminate the walk in the next step and this happens with the probability $\rho_t = (t+1)/|\Omega|)$. Thus

$$p_{t+1} = \frac{t+1}{|\Omega|} q_t \quad . \tag{5.71}$$

The probability that the walk after $t+1$ steps is open is

$$q_{t+1} = q_t(1-\rho_t) = q_t\left(1 - \frac{t+1}{|\Omega|}\right) \tag{5.72}$$

with $q_0 = 1$. Due to the fact that the phase space is growing exponentially we can write

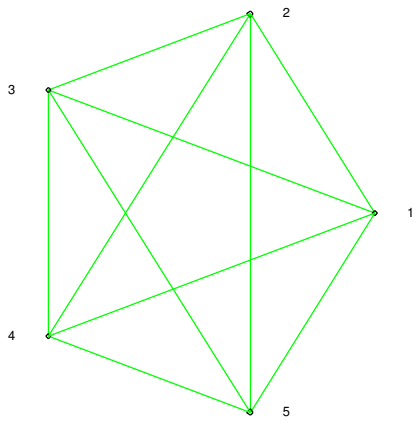


Figure 5.10: A random network with 5 nodes and 10 edges

$$q_t = \prod_{i=1}^{t}\left(1-\frac{i}{|\Omega|}\right) \approx \prod_{i=1}^{t} e^{-i/|\Omega|} \tag{5.73}$$

$$= e^{-\sum_i i/|\Omega|} = e^{-t(t+1)/2|\Omega|} \tag{5.74}$$

We can now calculate the average cycle length. Let $P(L)$ denote the probability that a given starting point is in the basin of attraction with cycle length L. The closing event happens with equal probability so that

$$P(L) = \sum_{t=L}^{|\Omega|} \frac{p_t}{t} \tag{5.75}$$

and for the expectation value

$$\langle L \rangle \approx |\Omega|^{1/2} \quad . \tag{5.76}$$

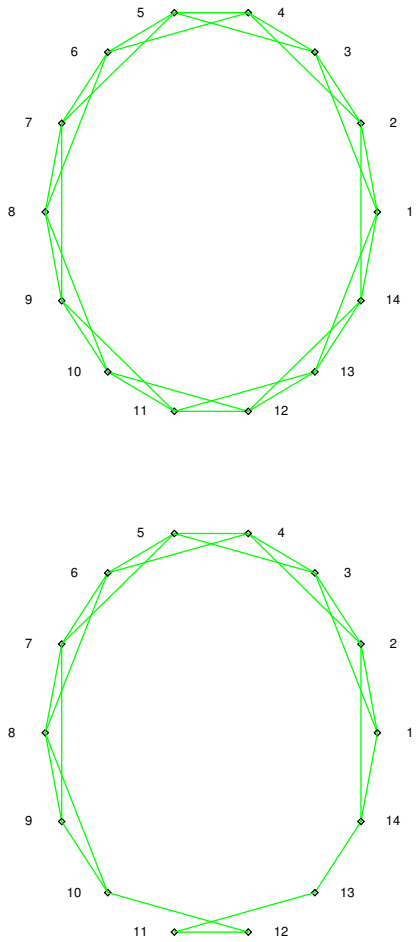


Figure 5.11: Generation of a simple small world random network