Let then $C = (V_1, V_2)$ be a cut in $G$, and divide the edges in $G$ corresponding as

$$\begin{aligned} E_1 &= \{\langle i,j\rangle \in E : i,j \in V_1\}, \\ E_2 &= \{\langle i,j\rangle \in E : i,j \in V_2\}, \\ E_C &= \{\langle i,j\rangle \in E : i \in V_1 \wedge j \in V_2\}. \end{aligned}$$

Consider the spin glass state $\sigma$ determined as

$$S_i = \begin{cases} +1, & \text{if } i \in V_1, \\ -1, & \text{if } i \in V_2. \end{cases}$$

For this,

$$\begin{aligned} H(\sigma) &= -\sum_{\langle ij\rangle} J_{ij} S_i S_j = \sum_{\langle ij\rangle \in E} S_i S_j \\ &= \sum_{\langle ij\rangle \in E_1} S_i S_j + \sum_{\langle ij\rangle \in E_2} S_i S_j + \sum_{\langle ij\rangle \in E_C} S_i S_j \\ &= |E_1| + |E_2| - |E_C| \\ &= |E| - 2|E_C| \\ &= |E| - 2w(C). \end{aligned}$$

Conversely, given any spin glass state $\sigma$, one obtains a cut $C$ satisfying $w(C) = \frac{1}{2}|E| - \frac{1}{2}H(\sigma)$.

Thus, graph cuts and spin glass states correspond one-to-one, with $w(C) \propto -H(\sigma)$, and minimising one is equivalent to maximising the other.

The result means that the SK spin glass ground state problem is in a sense "universal" difficult problem, i.e. it contains as special cases all the $\sim$2000 other known NP-complete problems.

For $J_{ij} > 0$ and arbitrary $\bar{h}$ the problem reduces to network flow, and can be solved in polynomial time. For planar $G$ and $\bar{h} = 0$ the problem also has a polynomial time algorithm (Fisher 1966 (2-D lattices), Barahona 1982). However, for planar $G$ with $\bar{h} \neq 0$, and for 3-D lattices the problem is NP-complete (Barahona 1982). It is also NP-complete for every other nonplanar crystal lattice graph (Istrail 2000). Thus, the dimensionality of the system is not crucial to the complexity of the ground state problem; the key is rather the planarity of the interconnection graph.

## 6.3 Neural Networks

John Hopfield proposed, in an influential paper in 1982, to use the SK model as a basis for "neural associative memories". The idea is to create an $N$-site SK

system whose local potential minima correspond to a set of $N$-bit vectors to be stored. These local minima are also stable states of the system's deterministic (0-temperature) "Glauber dynamics". When such a system is initialised at a state which is "close" to one of the stored stable states, the dynamics (presumably) tends to return it to the nearby local minimum. Thus small perturbations in the stable states tend to get corrected, and the system has "error-correcting" or "associative" capabilities.

More precisely, the deterministic dynamics of such a system is as follows: at a given discrete time instant, a randomly (or in a round-robin manner) chosen site $k$ is updated according to the local rule:

$$\begin{aligned} S'_k &= \operatorname{sgn} \underbrace{\left( \sum_{\langle kj \rangle} J_{kj} S_j + h_k \right)}_{(\bigstar)} \\ &= \begin{cases} +1, & \text{if } (\bigstar) > 0, \\ -1, & \text{if } (\bigstar) < 0, \\ S_k, & \text{if } (\bigstar) = 0, \end{cases} \end{aligned}$$

It can be seen that each time a site changes state, the value of $H(\sigma)$ decreases: Assume $S'_k \neq S_k$. Consider

$$\begin{aligned} H(\sigma') - H(\sigma) &= -\sum_{\langle ij \rangle} J_{ij} S'_i S'_j - \sum_i h_i S'_i \\ &\quad + \sum_{\langle ij \rangle} J_{ij} S_i S_j + \sum_i h_i S_i \\ &= -\sum_{\langle kj \rangle} J_{kj} S'_k S_j + \sum_{\langle kj \rangle} J_{kj} S_k S_j - h_k (S'_k - S_k) \\ &= -\underbrace{(S'_k - S_k)}_{\blacktriangle} \underbrace{\left( \sum_{\langle kj \rangle} J_{kj} S_j + h_k \right)}_{\blacktriangledown} \\ &< 0, \end{aligned}$$

where ▼ and ▲ have the same sign.

Thus, since the value of $H(\sigma)$ is lower bounded by

$$H(\sigma) \geq -\sum_{\langle ij \rangle} |J_{ij}| - \sum_i |h_i|,$$

the system converges eventually to a local minimum of its Hamiltonian.

How should one then craft the interaction coefficients so that a given set of patterns become stable states of the system's dynamics? This can in principle be done in various ways, of which Hopfield proposed the following adaptation of "Hebb's rule":[4]

Consider first a single pattern $\sigma = (S_1, \ldots, S_N) \in \{+1, -1\}^N$ and choose $J = \sigma\sigma^T - I = [S_i S_j]_{ij} - I, h = 0$. Then the dynamics operates as follows:

$$\operatorname{sgn}(J\sigma) = \operatorname{sgn}\left((\sigma\sigma^T - I)\sigma\right) = \operatorname{sgn}\left((||\sigma||^2 - 1)\sigma\right) = \sigma,$$

i.e. $\sigma$ is a stable state of the dynamics.

Given then a (smallish) set of patterns $\sigma_1, \ldots, \sigma_m$, choose

$$J = \sum_{p=1}^{m} \sigma_p \sigma_p^T - mI \qquad \left(\text{or normalised } J = \frac{1}{m}\sum_p \sigma_p \sigma_p^T - I\right).$$

If the patterns are random, independent identically distributed bit vectors, and there are only $m \ll N$ of them, they are "almost orthogonal", and we may approximate:

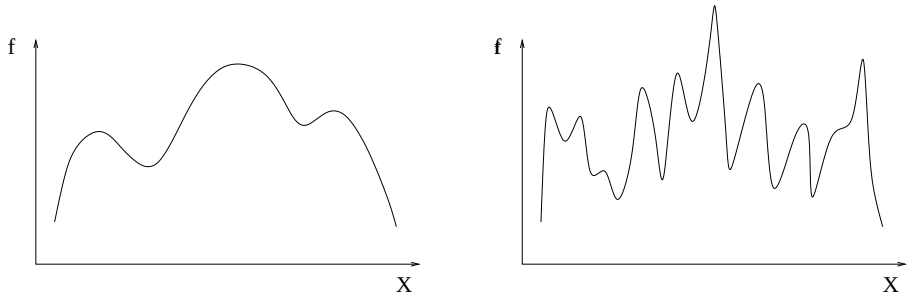
$$\begin{aligned}
\operatorname{sgn}(J\sigma_k) &= \operatorname{sgn}\left(\left(\sum_{p=1}^{m} \sigma_p \sigma_p^T - mI\right)\sigma_k\right) \\
&= \operatorname{sgn}\left(\underbrace{(||\sigma_k||^2 - m)\sigma_k}_{\text{"signal"}} + \underbrace{\sum_{p\neq k} \overbrace{(\sigma_p^T \sigma_k)}^{\approx 0}\sigma_p}_{\text{"noise"}}\right) \\
&= \sigma_k,
\end{aligned}$$

"with high probability".

This analysis has been performed rigorously many times under different assumptions, and the number of patterns $m$ that can be reliably stored has been estimated under different criteria. Typically, the "reliable" storage capacity comes out as $m \approx 0.14N \ldots 0.18N$.

The deterministic Glauber dynamics of SK spin glasses has also other computationally interesting features. One can e.g. show that convergence to a stable state

[4] In a 1949 book, D. O. Hebb suggested as a basic mechanism of neuronal memory that simultaneous activity reinforces the interconnections between neurons. Physiologically this suggestion is still controversial, but mathematically the idea has been used as a basis of several learning mechanisms in artificial neural networks.




(a) $K \ll N$ (b) $K \approx N$

Figure 4: A smooth (a) and a rugged (b) NK fitness landscape.

can require a number of spin flips that is exponential in $N$ (A. Haken et al. ca. 1989), and that one can in fact embed arbitrary computations in the dynamics (Orponen 1995). (More precisely, determining whether a given "output spin" is $+1$ or $-1$ in the local minimum reached from a given initial state is a "PSPACE-complete" problem.)

## 6.4 The NK Model

Introduced by Stuart Kauffman (ca. 1986) as a "tunable family of fitness landscapes".

A *fitness landscape* is a triple $\langle X, R, f \rangle$, where $X$ is the *configuration* (or *state*) *space*, $R \subseteq X \times X$ is a *neighbourhood relation* on $X$, and $f : X \to \mathbb{R}$ is a *fitness* (or *objective*) *function* .

A point $x \in X$ is a *local optimum* (of $f$ on $X$) if

$$f(y) \leq f(x) \quad \forall\, yRx$$

and a *global optimum* (*maximum*) if

$$f(y) \leq f(x) \quad \forall\, y \in X$$

Questions of the "ruggedness" of landscapes (correlation structure), number and height of local optima, sizes of "attraction basins" of local optima with respect to "hill-climbing" algorithms etc. are of great interest for natural landscapes.

In Kauffman's NK models, $X = A^N$ (usually just $X = \{0,1\}^N$) and $K$ is a tunable neighbourhood size parameter that influences the landscape characteristics, especially its ruggedness (cf. Figure 4).

The model can be seen as a toy model of "epigenetic interactions in chromosomes" — or also a generalisation of the spin glass model.

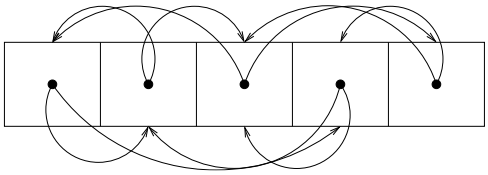


Figure 5: An NK interaction network with $N = 5$, $K = 2$.

In Kauffman's model, a *chromosome* is an $N$-vector of *loci* (*genes*, "positions"), each of which has a value from a set of *alleles* $A$ (usually just $A = \{0,1\}$). A "filled-in" chromosome $\alpha \in A^N$ is called a *genotype*.

The fitness of each gene $i \in \{1,\dots,N\}$ in a genotype $\alpha = (a_1,\dots,a_N) \in A^N$ depends on the allele $a_i$ and $K$ other alleles $a^i_1,\dots,a^i_K$ via some local fitness function $f^i(\alpha) = f^i(a_i; a^i_1,\dots,a^i_K)$, usually normalised so that $f^i(\alpha) \in [0,1]$. The total fitness of a genotype $\alpha \in A^N$ is the normalised sum of its genes' local fitnesses:

$$f(\alpha) = \frac{1}{N}\sum_{i=1}^{N} f^i(a_i; a^i_1,\dots,a^i_k) \quad \in [0,1].$$

Figure 5 illustrates an NK network with five loci and two "epigenetic interactions" per locus.

In Kauffman's versions of the model, the $K$ loci affecting locus $i$ can either be systematically selected as e.g. $i+1,\dots,i+K \pmod N$, or the chromosome can be simply "randomly wired". The $f^i$ are usually determined as randomly generated $2^{K+1}$-element "interaction tables".

From the spin glass perspective, e.g. a 1-D Ising model with $N$ spins can be seen as an $N2$ network where $f^i(S_i; S_{i-1}, S_{i+1}) = \frac{J}{2}(S_{i-1}S_i + S_iS_{i+1})$, and an SK spin glass with coefficients $J_{ij}$ and local fields $h_i$ as an $N(N-1)$ network where

$$f^i(S_i; \sigma \setminus \{S_i\}) = \frac{1}{2}\sum_{\langle ij\rangle} J_{ij}S_iS_j + h_iS_i.$$

Basic properties of the NK model, for binary alleles $A = \{0,1\}$ and varying values of $K$, include the following:

$K = 0$**:**

> If $f^i(0) \neq f^i(1)\ \forall i = 1,\dots,N$, then there is a unique global optimum, which is easily found by e.g. the obvious 1-locus mutation "hill-climbing" algorithm.

Expected length of the hill-climbing path is $N/2$. (Half of the alleles are "right" in the beginning, after that one allele gets fixed at each step.)

Neighbouring genotypes $\alpha$, $\alpha'$ are always highly correlated, as necessarily $|f(\alpha) - f(\alpha')| \leq 1/N$.

$1 \leq K < N-1$**:**

For $K = 1$, a global optimum can still be found in polynomial time. For $K \geq 2$, global optimisation is NP-complete. However, for adjacent affecting loci $(i \curvearrowleft i+1, \ldots, i+K)$, the problem can be solved in time $O(2^K N)$ (Weinberger).

$K = N-1$**:**

Neighbouring genotypes are totally uncorrelated.

$\Rightarrow$ Probability that a given genotype $\alpha$ is a local optimum is equal to the probability that $\alpha$ has the highest rank within its 1-mutant neighbourhood. This probability is equal to $1/(N+1)$.

$\Rightarrow$ The expected number of local optima is $2^N/(N+1)$.

The expected number of improvement steps for 1-mutant hill-climbing to hit a local optimum is proportional to $\log_2 N$ (each improvement step typically halves the rank of the genotype within the neighbourhood).

The expected waiting time for finding an improvement step is proportional to $N$.

# 7 Random Graphs

## 7.1 The Erdős-Rényi Model(s)

Two closely related "uniform" random graph models introduced in 1959 by P. Erdős & A. Rényi and E. N. Gilbert.

Consider the family $\mathcal{G}_n$ of all (labelled, undirected) graphs on $n$ nodes. Denote $N = \binom{n}{2}$; then $|\mathcal{G}_n| = 2^N$.

Define the following two probability spaces

[Erdős & Rényi:] $\mathcal{G}(n,M)$ = all $G \in \mathcal{G}_n$ with exactly $M \leq N$ edges, taken with uniform probability, i.e.

$$\Pr(G_M = H) = \begin{cases} \binom{N}{M}^{-1}, & \text{if } H \text{ has } M \text{ edges} \\ 0; & \text{otherwise.} \end{cases}$$

[Gilbert:] $\mathcal{G}(n,p)$ = all $G \in \mathcal{G}_n$, taken so that each edge has occurrence probability $p$, $0 \leq p \leq 1$, independently of the other edges, i.e.

$$\Pr(G_p = H) = p^M(\underbrace{1-p}_{q})^{N-M}, \text{ if } H \text{ has } M \text{ edges.}$$

These spaces are in a precise sense "close" if $M \sim pN$, and are often both referred to (unfairly to Gilbert) as the "Erdős-Rényi random graph model", or alternatively as the $\mathcal{G}(n,M)$ and $\mathcal{G}(n,p)$ random graph models.

Let $\Omega_n, n = 0,1,2,\ldots$ be a sequence of probability spaces of $n$-node graphs. Say that *almost every* (*a.e.*) graph in $\Omega_n$ has property $Q$ if

$$\Pr(G \in \Omega_n \text{ has } Q) \to 1, \text{ as } n \to \infty.$$

Conversely, *almost no* graph in $\Omega_n$ has property $Q$ if a.e. graph in $\Omega_n$ has property $\neg Q$, i.e.

$$\Pr(G \in \Omega_n \text{ has } Q) \to 0, \text{ as } n \to \infty.$$

**Theorem 7.1** *Let $H$ be a fixed graph and $p$ a constant, $0 < p < 1$. Then a.e. $G \in \mathcal{G}(n,p)$ contains an induced copy of $H$.*

Remark: an "induced copy" means here a subset of nodes whose induced subgraph is isomorphic to $H$.

*Proof.* Let $k = |H| =$ number of nodes in $H$. Then a graph $G$ with $n = |G| \geq k$ nodes can be partitioned into $\lfloor n/k \rfloor$ disjoint sets of $k$ nodes (with some left over). For each of these sets, the probability that it forms an induced copy of $H$ is $r > 0$. (Precisely, $r = \frac{k!}{|\mathrm{Aut}(H)|} p^{e(H)} q^{\binom{k}{2} - e(H)}$.)

Thus, the probability that none of these sets forms an induced copy of $H$ is

$$(1-r)^{\lfloor n/k \rfloor} \to 0, \text{ as } n \to \infty. \square$$

Let $k, l \in \mathbb{N}$. Say that a graph $G = (V,E)$ has property $Q_{kl}$ if $\forall\, U, W, |U| \leq k, |W| \leq l, U \cap W = \varnothing$, $G$ contains a node $v \in V \setminus (U \cup W)$ such that $v$ is adjacent to all $u \in U$ and no $w \in W$ (cf. Figure 6).

**Lemma 7.2** *For every constant $p$, $0 < p < 1$, and all $k, l \in \mathbb{N}$, a.e. $G \in \mathcal{G}(n,p)$ has property $Q_{kl}$.*

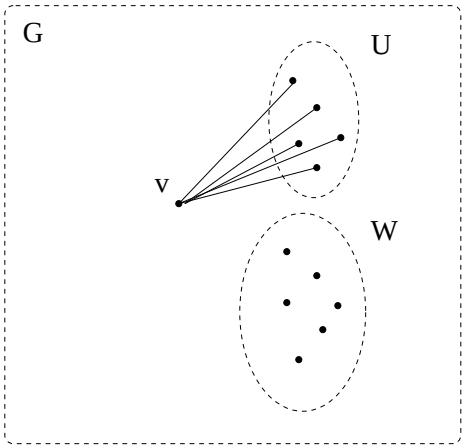



Figure 6: Property $Q_{kl}$.

*Proof.* For a fixed $U, W, v \in V \setminus (U \cup W)$, the probability that the condition is satisfied is

$$p^{|U|}q^{|W|} \geq p^k q^l$$

The events are independent for different $v$, so the probability that no appropriate $v$ exists is

$$\left(1 - p^{|U|}q^{|W|}\right)^{n-|U|-|W|} \leq \left(1 - p^k q^l\right)^{n-k-l}.$$

There are at most $n^{k+l}$ $(U,W)$-pairs to be considered, so the probability that some pair has no good $v$ is bounded by

$$n^{k+l}(\underbrace{1 - p^k q^l}_{<1})^{n-k-l} \to 0, \text{ as } n \to \infty.$$

Thus in a.e. $G \in \mathcal{G}(n,p)$ all $(U,W)$-pairs have some appropriate $v$. □

**Corollary 7.3** *Let $p$, $0 < p < 1$, be a constant. Then (i) a.e. $G \in \mathcal{G}(n,p)$ has minimum degree $\geq k$, for given constant $k$ (ii) a.e. $G \in \mathcal{G}(n,p)$ has diameter 2 (iii) a.e. $G \in \mathcal{G}(n,p)$ is $k$-connected for given constant $k$.*

*Proof.* (i) and (ii) are immediate.

(iii) In a.e. $G \in \mathcal{G}(n,p)$, no two nodes $u_1$, $u_2$ can be separated by a cutset of size $k-1$, because we may choose in Lemma 7.2 $U = u_1, u_2$, $W = w_1, \ldots, w_{k-1}$ for arbitrary $w_1, \ldots, w_{k-1}$, and obtain a path $u_1$—$v$—$u_2$ connecting $u_1$, $u_2$ and avoiding $w_1, \ldots, w_{k-1}$. □

**Corollary 7.4** *Let $\phi$ be any first-order sentence about graphs (i.e. quantification over nodes, relations $E(u,v)$ + identity). Then either $G \models \phi$ or $G \models \neg\phi$ for a.e. $G \in \mathcal{G}(n,p)$.*

*Proof.* Skipped. □

Thus, all the first-order properties of $\mathcal{G}(n,p)$ for fixed $p$ are easily captured. Things are more interesting when the number of nodes discussed and/or the probability $p$ depends on $n$.

Given graph $G$, denote:

| | | |
|---|---|---|
| independence number $\alpha(G)$ | = | size of the largest independent set in $G$, |
| clique number $\omega(G)$ | = | size of the largest clique in $G$, |
| chromatic number $\chi(G)$ | = | smallest number of colours needeed for colouring nodes in $G$ so that no two adjacent nodes get the same colour. |

**Lemma 7.5** *Given $n \geq k \geq 2$, random $G \in \mathcal{G}(n,p)$:*

$$\Pr(\alpha(G) \geq k) \leq \binom{n}{k} q^{\binom{k}{2}}.$$

*Proof.* Probability that given $k$-set of nodes in $G$ is independent is $q^{\binom{k}{2}}$. Total number of $k$-sets is $\binom{n}{k}$. □

**Theorem 7.6** *Let $p, 0 < p < 1$ and $\varepsilon > 0$ be constant. Then for a.e. $G \in \mathcal{G}(n,p)$:*

$$\chi(G) \geq \frac{\ln 1/q}{2+\varepsilon} \cdot \frac{n}{\ln n} = \Omega\left(\frac{n}{\ln n}\right) = \textit{large!}$$

*Proof.* By Lemma 7.5, for any fixed $n \geq k \geq 2$:

$$\begin{aligned}
\Pr(\alpha(G) \geq k) \;&\leq\; \binom{n}{k} q^{\binom{k}{2}} \leq n^k q^{\binom{k}{2}} \\
&= q^{k\frac{\ln n}{\ln q} + \frac{1}{2}k(k-1)} \\
&= q^{\frac{k}{2}[-\frac{2\ln n}{\ln 1/q} + k - 1]} \\
&\to 0 \text{ for } k \text{ large,}
\end{aligned}$$

i.e. when

$$\frac{k}{2}\left[-\frac{2\ln n}{\ln 1/q} + k - 1\right] \to \infty.$$

A sufficient condition for this to hold is that $k \geq k(n,\varepsilon) = (2+\varepsilon)\frac{\ln n}{\ln 1/q}$. Thus for large $n$, almost no graph $G \in \mathcal{G}(n,p)$ can have a colouring that would assign the same colour to $k(n,\varepsilon)$ or more nodes. Hence, a proper colouring of almost any $G \in \mathcal{G}(n,p)$ requires at least $\frac{n}{k(n,\varepsilon)} = \frac{\ln 1/q}{2+\varepsilon} \cdot \frac{n}{\ln n}$ colours. □

**Theorem 7.7** *Let $p$, $0 < p < 1$ be constant. Then for a.e. $G \in \mathcal{G}(n,p)$:*

$$\omega(G) \in \{d, d+1\},$$

*where $d = d(n,p)$ is the largest integer such that*

$$\binom{n}{d} p^{\binom{d}{2}} \geq \ln\, n.$$

*(This implies $d = 2\log_{1/p}(n) + O(\log\log n.)$.)* □

A *graph property* $Q$ is an isomorphism-closed family of graphs, i.e. if $G \in Q$ (or "$G$ has $Q$") and $G \approx G'$, then also $G' \in Q$.

A graph property is *monotone* if it is preserved under addition of edges, i.e. if $G = (V,E)$ and $G' = (V,E')$ are graphs such that $E \subseteq E'$ and $G$ has $Q$, then also $G'$ has $Q$.

A *threshold function* for the graph property Q is a function $t : \mathbb{N} \to \mathbb{R}$ such that

$$\Pr(G \in \mathcal{G}(n,p(n)) \text{ has } Q) \xrightarrow[n\to\infty]{} \begin{cases} 1, \text{ if } p \succ t \\ 0, \text{ if } p \prec t. \end{cases}$$

Notation:

$$p \succ t \Leftrightarrow \lim_{n\to\infty} \frac{p(n)}{t(n)} = \infty,$$
$$p \prec t \Leftrightarrow \lim_{n\to\infty} \frac{p(n)}{t(n)} = 0,$$
$$p \sim t \Leftrightarrow \lim_{n\to\infty} \frac{p(n)}{t(n)} = 1,$$
$$p \approx t \Leftrightarrow p(n) = \Theta(t(n)).$$

Denote: $P_n^Q(p) = \Pr(G \in \mathcal{G}(n,p) \text{ has } Q)$.

Then for monotone $Q$: $p_1 \leq p_2 \Rightarrow P_n^Q(p_1) \leq P_n^Q(p_2)\ \forall n$.

Denote: $p_n^Q(\alpha) =$ the smallest $p$ such that $P_n^Q(p) \geq \alpha$.