on a chain graph with $n+1$ vertices are $\le n^2$, and thus, $\mathbf{E}\{T\} \le n^2$. The probability that we declare no solution when in fact there is one is

$$\mathbf{P}\left\{\min_{t \le n^3} N_t > 0\right\} = \mathbf{P}\{T > n^3\} \le \frac{\mathbf{E}\{T\}}{n^3} \le \frac{1}{n}.$$

This is an example of a Monte Carlo algorithm, i.e., an algorithm that returns the correct answer with high probability.

In fact, the probability of an error drops off exponentially quickly. To see this, consider the random walk for the first $2n^2$ time units. Then consider the random walk for the next $2n^2$ time units, and so forth. In this manner, the n^3 steps are partitioned into $n/2$ such random walks. To commit an error (reporting no solution when in fact a solution exists) implies that each of these $n/2$ random walks fails to find a solution. The probability of such an individual failure, p, is bounded by $\mathbf{E}\{T\}/2n^2 < 1/2$. Thus, the probability of an overall error is bounded by $p^{n/2} \le 1/2^{n/2}$. The total expected complexity is bounded by

$$n^3 km = O(mn^3).$$

Bibliographic remarks. The randomized algorithm for 2-SAT is due to Papadimitriou (1991). The exercise on random 2-SAT is due to Goerdt (1992) and Chvatal and Reed (1992).

§8.9. The Metropolis chain → To simulate nuclear reactions!

Consider a finite set V of states, and assume that we wish to randomly generate a state X according to a distribution π:

$$\mathbf{P}\{X = i\} = \pi_i \,, i \in V.$$

something we only know $c\pi_i$, c unk

A typical request may be to generate X uniformly over V. This is not difficult when V is the set $\{1, \ldots, k\}$, or V is a simple structure, but it is quite cumbersome when V is huge, and in fact, the entire set V cannot even be counted or kept in memory. In such situations, the Markov Chain Monte Carlo (MCMC) method explicitly constructs an irreducible aperiodic Markov chain (X_t), started at any $X_0 \in V$, with the property that π is the limiting distribution of X_t as $t \to \infty$. One then returns X_t for sufficiently large t as an "approximation" of X.

1º Markov Chain Monte Carlo method!

First, construct a connected graph $G = (V, E)$ on V in which each vertex has low degree. Let d_i be the degree of vertex i. The nature of the connected graph is usually determined from the problem at hand—examples follow. Having fixed G, we define the transition probabilities as follows:

$$p_{i,j} = \begin{cases} \frac{1}{d_i} \min\left(\frac{\pi_j d_i}{\pi_i d_j}, 1\right) & \text{if } (i,j) \in E\,; \\ 0 & \text{if } (i,j) \notin E,\, i \ne j\,; \\ \underbrace{1 - \sum_{\ell : \ell \sim i} \frac{1}{d_i} \min\left(\frac{\pi_\ell d_i}{\pi_i d_\ell}, 1\right)}_{\text{leftover}} & \text{if } i = j. \leftarrow \text{To make it } \text{acyclic. aperiod} \end{cases}$$

Here $\ell \sim i$ denotes that ℓ is a neighbor of i. The Markov chain thus defined on a connected graph is called the Metropolis chain. Started at any vertex, one step in this chain is equivalent to the following 2-step transition: given X_t, let Y_t be a neighbor of X_t picked uniformly at random from all neighbors. Then set

$$X_{t+1} = \begin{cases} Y_t & \text{with probability } \min\left(\frac{\pi_{Y_t} d_{X_t}}{\pi_{X_t} d_{Y_t}}, 1\right) \\ X_t & \text{otherwise.} \end{cases}$$

← To be sure it is aperiodic

Construct it so the chain is reversible.

So π_i is a (scaled) solution

Start the random walk in a (any) state and wait a while.

The computational cost of a transition is thus the cost of computing both the degree and the π value at both X_t and Y_t, and the cost of generating a random neighbor. It is important to be able to identify neighbors and count them.

The Markov chain is irreducible since G is connected and $p_{i,j} > 0$ for all $(i,j) \in E$ (this requires that all $\pi_i > 0$, something we will assume throughout). The Markov chain is not necessarily aperiodic, so the aperiodicity needs to be verified case by case. Finally, we must show that π is indeed the stationary distribution. We do this by proving that the Metropolis chain is time-reversible:

$$\pi_i p_{i,j} = \pi_j p_{j,i}$$

for all i, j. This is trivial for $i = j$, so we assume $i \neq j$. If i and j are not neighbors, then $p_{i,j} = p_{j,i} = 0$, and so we are fine too. Thus, consider $(i,j) \in E$. Assume without loss of generality that

$$\pi_j d_i \geq \pi_i d_j .$$

Then

$$\pi_i p_{i,j} = \frac{\pi_i}{d_i} = \pi_j \times \frac{1}{d_j} \times \frac{\pi_i d_j}{\pi_j d_i} = \pi_j p_{j,i},$$

as required.

When $\pi_i = \pi_j$ $\forall_{ij}$

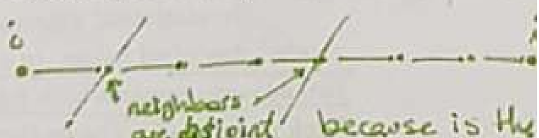
For the uniform distribution on S, we note that $X_{t+1} = Y_t$, a uniform random neighbor, with probability $\min(1, d_{X_t}/d_{Y_t})$. Put differently, $p_{i,j} = 1/\max(d_i, d_j)$ if i is a neighbor of j and 0 otherwise. If in addition, all degrees are equal, as in a regular graph, then $X_{t+1} = Y_t$, and the Metropolis chain coincides with a simple random walk on the graph G. It is quite interesting that for all Metropolis chains with uniform limit distribution, $\sup_i \mathbf{E}\{T_i\} = O(n^2 \log n)$, and that this bound is tight as it is reached for the cobweb graph [see the exercises]. To prove this, recall the proof of the upper bound for the cover time in random walks. We note that for an edge (i,j), if π_i denotes the stationary probability of state i,

$P_{ij} = \frac{1}{\max(d_i, d_j)}$

$$\mathbf{E}\{T_{iji}\} \leq \frac{1}{\pi_i \mathbf{P}\{i \to j \text{ in one step}\}} = n \max(d_i, d_j) .$$

Let i and j be arbitrary nodes and consider the shortest path from i to j. By looking at the expected time to go from i to j along this path, and back, we note that

$$\mathbf{E}\{T_{iji}\} \leq n \sum_{(k,\ell)} \max(d_k, d_\ell) \leq 2n \sum_k d_k$$



because is the shortest path b/w i, j

where the first sum is over all edges (k,ℓ) on the path, and the last sum is over all vertices of the path. If we split the sum into three parts by taking nodes on the path in groups modulo 3 (the first, 4th, 7th, and so on are in group one; the second, fifth, and so on in group 2; and the third, sixth, and ninth nodes are in group 3, for example), then the sum of the degrees in each group is not more than $n-1$, because within a group, any neighbor of a node u in that group cannot be a neighbor of another node v in the same group. Indeed, if neighbors were shared, then we could have found a shorter path from i to j passing through that neighbor. Thus,

$$\mathbf{E}\{T_{ij}\} \leq \mathbf{E}\{T_{iji}\} \leq 6n(n-1) < 6n^2 .$$

Therefore, by Matthews' inequality,

$$\sup_i \mathbf{E}\{T_i\} \leq 6n^2 H_n .$$

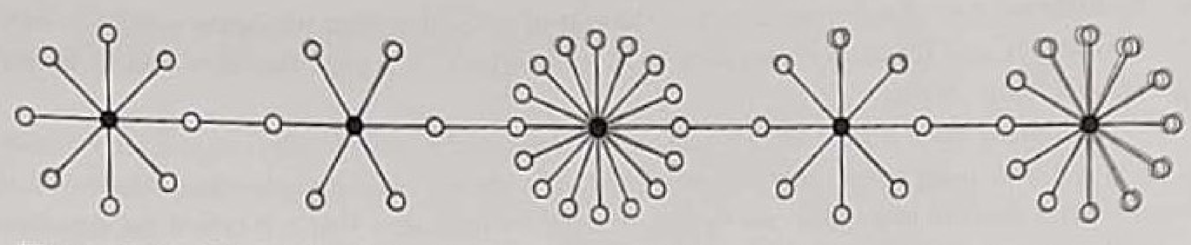


Figure 8.18. Illustration of the argument above. The leftmost black node is i, and the rightmost black node is j. Each third node on the shortest path from i to j is darkened. The neighborhoods of those nodes are necessarily non-overlapping.

Examples: ① Consider the following problem: generate a random labeled connected graph with n vertices and e edges, such that each such graph is equally likely. We label the nodes $1, \ldots, n$. If e is small compared to $(1/2)n \log n$, the following strategy just does not seem efficient enough: keep generating independent copies of G_{ne} until we obtain a connected graph. Indeed, the expected number of iterations before halting is $1/\mathbf{P}\{G_{ne}$ is connected$\}$, and this is just too large.

We take this simple example to illustrate the Markov chain method. Create a connected graph $\mathcal{G}$ in which each node corresponds to a different connected graph G_{ne}. The edges of $\mathcal{G}$ connect those vertices that differ by at most two edges. Now take a starting node on $\mathcal{G}$, and begin a Metropolis walk with uniform limit distribution. Note that this requires a transition from node (graph) u to v with probability $1/\max(d_u, d_v)$, and thus requires knowing how many neighbors in $\mathcal{G}$ both u and v have. Note that the set of neighbors of u can be dynamically maintained. It can also be obtained by adding each possible edge that is not an edge of u (for a total of $(\binom{n}{2} - e)$ edges, and then removing each possible edge of u that will not cause disconnection (this number is at least $e - n$ and at most e). After a sufficiently long time, this Metropolis walk results in a nearly uniform random graph from $\mathcal{G}$.

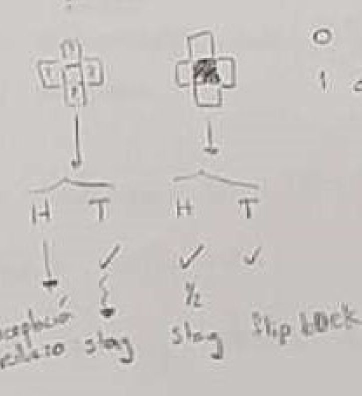
Example ②

§8.10. The hard core model

In statistical physics, a number of quite complex models have emerged such as the "hard core model", described below. In a graph $G = (V, E)$, we assign values 0 and 1 to each vertex and ask that no two adjacent vertices both take the value 1. Let S be the subset of all feasible graphs with this property. Each member of S can be thought of as a binary vector from $\{0,1\}^n$ where $n = |V|$. Assume that we are interested in generating a uniformly distributed member of S. The MCMC methodology here consists in creating a Markov chain on S with the following properties: we start at the safe value $(0, 0, 0, \ldots, 0)$. The possible neighbors of $v \in S$ are identified by flipping the i-th bit, $1 \le i \le n$. It is easy to check whether flipping the i-th bit leads to a vertex of S. If $X_t = v$, then we generate a random index $Z \in \{1, \ldots, n\}$, which is like picking a random vertex of V. In X_{t+1}, all bits agree with those of X_t except possibly the Z-th. The Z-th bit of X_{t+1} is one if and only if a fair coin toss comes up heads and all neighbors of Z in G have the value 0 (so that $X_{t+1} \in S$).

The irreducibility of the chain is easily established by noting that the zero state can be reached from any state in S, and that any state of S, by reversing this, can be reached from the zero state. So, all states of S communicate. The aperiodicity follows by noting that we have $\mathbf{P}\{X_{t+1} = X_t | X_t = v\} \ge 1/2$

0 ⇒ T or H and not "acceptable to flip"
1 ⇒ H and "acceptable to flip"



for all $v \in S$. Finally, we must establish that the stationary distribution is the uniform distribution. To see this, we resort once again to time-reversibility. Consider $u, v \in S$. Then defining

$$\pi_u = \pi_v = \frac{1}{|S|},$$

we see that, if d_u denotes the degree of u,

$$p_{u,v} = \begin{cases} \frac{1}{2n} & \text{if } u \sim v \\ 0 & \text{if } u \not\sim v \\ 1 - \frac{d_u}{2n} & \text{if } u = v. \end{cases}$$

When u, v are not adjacent, this is trivial. When $u \sim v$, and they differ only in the k-th component, then all neighbors of k in G must have the value 0, so u and v agree on all those 0-values. A transition is made with the given probability, $1/(2n)$. But then, we have

$$\pi_u p_{u,v} = \pi_v p_{v,u} = \begin{cases} \frac{1}{2n|S|} & \text{if } u \sim v \\ 0 & \text{if } u \not\sim v. \end{cases}$$

This establishes the time-reversibility, and thus, the uniform distribution is the unique stationary distribution.