Math 832, Assignment #3: Solutions

1. Let $T_1 = \inf\{n > 0 : X_n = 1\}$, $T_2 = \inf\{n > T_1 : X_n = 2\}$, $T_3 = \inf\{n > T_2 : X_n = 1\}$, $T_4 = \inf\{n > T_3 : X_n = 2\}$, etc. Then we can define (random) maps $F_1, F_2, \ldots$ from $\{1,2\}$ to itself by putting $F_1(1) = X_{T_1+1}$, $F_1(2) = X_{T_2+1}$, $F_2(1) = X_{T_3+1}$, $F_2(2) = X_{T_4+1}$, etc. By the strong Markov property, the (function-valued) random variables $F_1, F_2, \ldots$ are independent of each other, and it's easy to see that for all k, $\mathrm{Prob}(F_k(1) = 2) = p$ and $\mathrm{Prob}(F_k(2) = 1) = q$, so the common distribution of the random maps $F_1, F_2, \ldots$ respects the Markov kernel $p(\cdot,\cdot)$. The probability that F_k collapses $\{1,2\}$ to a singleton is $(1-p)q + p(1-q) > 0$, so with probability 1, there exists a smallest integer m and $X^* \in \{1,2\}$ such that $F_1 \circ F_2 \circ \ldots \circ F_n$ collapses $\{1,2\}$ to the singleton $\{X^*\}$ for all $n \geq m$. Once we have determined m and X^*, we can simply let $N = \inf\{n \geq 2m : X_n = X^*\}$.

Note: The same idea works for irreducible Markov chains with any fixed (finite) number of states, as long as the number of states is known ahead of time.

2. For any finite string $x = (x_1, \ldots, x_k)$ of $+1$'s and -1's, let $E(x) = \sum_{i=1}^{k} -x_i x_{i+1}$, so that for $s = (s_1, \ldots, s_n)$ and any $1 \leq m \leq n$ we have $E(s) = E(s_1, \ldots, s_m) + E(s_m, \ldots, s_n)$. Let $Z = \sum_s q^{-E(s)}$. Then we have

$$\mathrm{Prob}(X_1 = x_1, \ldots, X_m = x_m) = \frac{1}{Z} q^{-E(x_1,\ldots,x_m)} \sum_{x_{m+1},\ldots,x_n} q^{-E(x_m,\ldots,x_n)},$$

where the sum on the right can be written as $C(x_m)$ (since it does not depend on $x_1, \ldots, x_{m-1}$). Likewise, for any x'_m we have

$$\mathrm{Prob}(X_1 = x_1, \ldots, X_{m-1} = x_{m-1}, X_m = x'_m) = \frac{1}{Z} q^{-E(x_1,\ldots,x_{m-1},x'_m)} C(x'_m).$$

Dividing the first equation by the second, and using the fact that the energy difference $E(x_1, \ldots, x_{m-1}, x_m) - E(x_1, \ldots, x_{m-1}, x'_m)$ depends only on x_{m-1}, x_m, and x'_m, we find that

$$\frac{\mathrm{Prob}(X_1 = x_1, \ldots, X_{m-1} = x_{m-1}, X_m = x_m)}{\mathrm{Prob}(X_1 = x_1, \ldots, X_{m-1} = x_{m-1}, X_m = x'_m)}$$

doesn't depend on $x_1, \ldots, x_{m-2}$. It follows that

$$\mathrm{Prob}(X_m = x_m | X_1 = x_1, \ldots, X_{m-1} = x_{m-1})$$

is independent of $x_1, \ldots, x_{m-2}$, as was to be shown.

(In general, the Boltzmann distribution for 1-D statistical mechanics models with finite-range interactions are Markov.)

3. (a) The case $n = 0$ is trivial, so take $n \geq 1$. For fixed a, b, let Ω' be the set of all sequences in Ω with $X_{-n} = a$, $X_n = b$, and with no two consecutive 1's. Then it is easy to see that the uniform distribution on Ω' satisfies the conditional probability specification given in the problem. Conversely, any probability distribution μ on Ω that satisfies the specification must satisfy $\mu(\Omega') = 1$, and must have the property that $\mu(s) = \mu(s')$ whenever $s, s' \in \Omega'$ differ in only a single position. But this implies that μ is constant on Ω', since any two strings in Ω' can be reached from one another by a sequence of single-site mutations (to see this, note that the string $a, 0, 0, \ldots, 0, 0, b$ can be reached from any string by successively converting 1's to 0's). Hence μ must be the uniform distribution on Ω'.

(b) $\mathrm{Prob}(X_{m+1} = x_{m+1} | X_m = x_m, \ldots, X_1 = x_1) = N_1/N_0$ where N_0 (resp. N_1) is the number of sequences $x'_{m+1}, \ldots, x_n$ (resp. $x_{m+2}, \ldots, x_n$) such that $(x_m, \ldots, x_n)$ (resp. $(x_{m+1}, \ldots, x_n)$) contains no two consecutive 1's. Since N_0 depends only on x_m and N_1 depends only on x_{m+1}, $\mathrm{Prob}(X_{m+1} = x_{m+1} | X_m = x_m, \ldots, X_1 = x_1)$ is independent of $x_1, \ldots, x_{m-1}$, so $X_{-n}, \ldots, X_n$ is indeed a Markov chain.

(In one dimension, a Markov random field is always a Markov chain.)

(c) $\mathrm{Prob}(X_0 = 1) = N'/N$, where N is the total number of strings beginning with $X_{-n} = a$ and ending with $X_n = b$, and N' is the number of strings satisfying $X_{-n} = a$, $X_0 = 1$, and $X_n = b$. N is the (a, b)th entry of M^{2n} (where

$$M = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

is the "possibility matrix" for transitions in the Markov chain), and N' is equal to the $(a, 1)$th entry of M^n times the $(1, b)$th entry of M^n. Write $M = LDL^{-1}$ with

$$\begin{aligned} L &= \begin{pmatrix} d & e \\ f & g \end{pmatrix} = \begin{pmatrix} \tau & -1 \\ 1 & \tau \end{pmatrix}, \\ D &= \begin{pmatrix} \tau & 0 \\ 0 & -\tau^{-1} \end{pmatrix}, \text{ and} \end{aligned}$$

$$L^{-1} = \begin{pmatrix} d' & e' \\ f' & g' \end{pmatrix} = \begin{pmatrix} \tau/(\tau^2+1) & 1/(\tau^2+1) \\ -1/(\tau^2+1) & \tau/(\tau^2+1) \end{pmatrix},$$

with $\tau = (1+\sqrt{5})/2$. The (a,b) entry of $M^{2n} = LD^{2n}L^{-1}$ is equal to τ^{2n} times the (a,b) entry of

$$\begin{pmatrix} dd' & de' \\ fd' & fe' \end{pmatrix}$$

plus τ^{-2n} times a multiplier that does not depend on n; the latter term becomes negligible compared to the former as $n \to \infty$. Similarly, the $(a,1)$ entry of M^n times the $(1,b)$ entry of M^n is equal to τ^{2n} times the (a,b) entry of the 2-by-2 matrix

$$\begin{pmatrix} de' \\ fe' \end{pmatrix} \begin{pmatrix} fd' & fe' \end{pmatrix} = \begin{pmatrix} de'fd' & de'fe' \\ fe'fd' & fe'fe' \end{pmatrix}$$

plus a negligible smaller term. Taking the ratio, we find that $\mathrm{Prob}(X_0 = 1)$ converges to $e'f = 1/(\tau^2+1) \approx .2764$ as $n \to \infty$. (Note: other forms of this answer are possible, such as $(\tau-1)/(2\tau-1)$.)

It's worth noticing that the boundary conditions a, b asymptotically made no difference in this example. That's true fairly generally for 1-dimensional systems (leaving aside phenomena associated with periodicity). But 2- (and higher-) dimensional systems are very different in this regard: boundary conditions can persist in having an effect on the behavior of the system at the origin even as the boundary goes out to infinity.

4. (a) Stationarity of the measure μ is equivalent to the condition that for all j, $\sum_{i \neq j} \mu(i) r_{i,j} = \mu(j) \sum_{i \neq j} r_{j,i}$. That is, μ (expressed as a row-vector) should be in the null-space of the matrix M whose (i,j)th entry is $r_{i,j}$ for $i \neq j$, and whose (j,j)th entry is $-\sum_{i \neq j} r_{j,i}$. Equivalently, μ is fixed by the matrix $I+M$ (where I is the identity). All row-sums in M vanish, so all row-sums in $I+M$ equal 1. Wouldn't it be nice if $I+M$ were actually a stochastic matrix? Well, we can make it so: for, we may assume without loss of generality that the rates $r_{i,j}$ satisfy $\sum_{i \neq j} r_{j,i} < 1$ for all j (if they don't, rescale all the rates so that they do). Then $I+M$ is a stochastic matrix, so that the existence of a vector μ with non-negative entries satisfying $(\mu)(I+M) = \mu$ is a consequence of our theorem concerning the existence of stationary measures for Markov chains.

(b) Suppose it is possible to partition S into two subsets A and A^c such that for all $i \in A$ and $j \in A^c$, $r_{i,j} = r_{j,i} = 0$. Then part (a), applied to the Markov process obtained by using the same rates $r_{i,j}$ on the restricted state-space A, assures us that there exists a stationary probability measure for the original Markov process that assigns measure 1 to A and measure 0 to A^c. Similarly, there is another stationary probability measure that assigns measure 0 to A and measure 1 to A^c. Thus, in this case the stationary distribution is not unique.

Conversely, suppose that the stationary measure is not unique. Then the transition kernel associated with the matrix $I + M$ constructed in part (a) is not irreducible (because we know from the theory of Markov chains that if the kernel were irreducible, the stationary measure μ would be unique); hence, by the structure theorem for Markov chains, there must exist two or more non-communicating (non-empty) components $S_1, S_2, S_3, \ldots$ such that $p_{i,j} = 0$ whenever i, j belong to distinct components. Letting $A = S_1$ and $A^c = S_2 \cup S_3 \cup \ldots$ (for instance), we see that for all $i \in A$ and $j \in A^c$, $r_{i,j} = r_{j,i} = 0$.

We conclude that the stationary measure is unique if and only if no such partition $S = A \dot{\cup} A^c$ exists. Equivalently, the stationary measure is unique if and only if, for all $i, j \in S$, there exists a sequence $k_0, k_1, k_2, \ldots, k_s$ such that all the rates r_{k_0,k_1}, r_{k_1,k_2}, $\ldots$, r_{k_{s-1},k_s} are positive, and such that either $k_0 = i$ and $k_s = j$ or $k_0 = j$ and $k_s = i$.

5. Since all rates are equal, it is easy to check that the uniform distribution on the state-space is stationary.

The "moral" of this problem is that for many purposes, continuous-time evolution is more mathematical tractable than discrete-time evolution. (Compare the results of this problem with the results of our two earlier problems on lords-and-ladies.)