Homework 8 Solutions

Note: These solutions are not necessarily model answers. Rather, they are designed to be tutorial in nature, and sometimes contain a little more explanation than an ideal solution. Also, bear in mind that there may be more than one correct solution. The maximum total number of points available is 44.

1. (a) Successively multiplying the row vector $(1, 0, 0, 0)$ by P three times, we get the sequence of vectors $\frac{1}{10}(0, 3, 1, 6)$, $\frac{1}{100}(58, 16, 22, 4)$ and $\frac{1}{1000}(74, 348, 192, 386)$. Thus the probability of being in state 4 after 3 steps starting in state 1 is 0.386. [Note that there is no need to compute the matrix product P^3.] *2pts*

A surprising number of people wrote out all paths of length 3 leading from state 1 to state 4 and summed their probabilities. While this is correct, it shows a lack of understanding of linear algebra and its connection with Markov chains. You should be sure that you understand how to solve this problem using matrix-vector mulitiplication.

(b) Multiplying the row vector $\frac{1}{4}(1, 1, 1, 1)$ by P three times as in part (a), we end up with the vector $\frac{1}{4000}(1003, 1210, 1009, 778)$. Thus the probability of being in state 4 after 3 steps starting in the uniform distribution is $\frac{778}{4000} = 0.1945$. *2pts*

The same comment as in part (a) applies to this part.

(c) The stationary distribution is $\pi = (0.2346, 0.3048, 0.2631, 0.1975)$ (to 4 decimal places), by solving for $\pi\mathbf{P} = \pi$ using a linear algebra package. [Note that π is a *left* (row) eigenvector, not the more commonly used right (column) eigenvector. The corresponding right eigenvector with eigenvalue 1 of *any* stochastic matrix is always the uniform distribution!] *2pts*

Quite a lot of students just took a large enough power of P so that it "looked as if it had converged". As in parts (a) and (b), this indicates a lack of understanding: the point here is that π is the unique (right) eigenvector of P with eigenvalue 1, and you can find this eigenvector using any linear algebra package.

(d) By numerical computation, we see that the variation distance $\|p_1^t - \pi\| := \frac{1}{2}\sum_y |P_{1,y} - \pi_y|$ has the value 0.00135 at $t = 20$ and the value 0.000989 at $t = 21$. Therefore the desired solution for a threshold of 0.001 is $t = 21$. *2pts*

A surprising number of students got the wrong answer here—for reasons that were not entirely clear. Note that this just involves taking large enough powers of P until the above distance drops below 0.001. This is a simple computation using any linear algebra package that supports matrix multiplication.

2. (a) It follows from the coupon collecting problem that for all v, $C_v(G) = \Theta(n \log n)$. (We are trying to hit all the vertices, and at each step, we hit each vertex with equal probability, except for the minor detail that we never hit the same vertex twice in succession. It is easy to check that this small change makes no difference to the asymptotics of coupon collecting.) *2pts*

(b) The lollipop graph has n vertices and $O(n^2)$ edges, so by the general bound in MU, Lemma 7.15, the cover time is bounded above by $O(|V| \cdot |E|) = O(n^3)$. *2pts*

(c) We know from the analysis of random walk on the line graph (MU, pp. 173–174) that $H_{b,a} = \Theta(n^2)$, so we have the lower bound $C_b(G) = \Omega(n^2)$. (Note that the existence of the clique is irrelevant until the walk reaches a for the first time.) For the upper bound, the same observation tells us that the expected number of steps to reach a from b (and thus along the way to visit all the vertices on the "tail") is $O(n^2)$; to complete the upper bound, we need to show that the walk, starting at a, will visit all vertices of the clique in expected time at most $O(n^2)$. Starting at a, with probability $\frac{1}{n/2}$ the walk moves right onto the tail, in which case the time is still bounded by the overall cover time $C(G) = O(n^3)$. With *4pts* probability $\frac{n/2-1}{n/2}$ the walk moves left onto the clique. At this point, we could essentially use the $\Theta(n \log n)$ cover time for $K_{n/2}$ from part (a), with the following caveat: every time the walk hits a, there is a small chance that it will make an excursion onto the tail. However, in $O(n \log n)$ steps on the clique the expected number of times the walk hits a is only $O(\log n)$, and whenever this happens, with probability $\frac{1}{n/2}$ it makes an expected $O(n)$ steps on the tail before returning to a. Putting all this together, we get

$$C_b(G) \leq H_{b,a} + \frac{1}{n/2} C(G) + \frac{n/2-1}{n/2}(\Theta(n \log n) + O(\log n) \cdot \frac{1}{n/2} \cdot O(n^2)) = O(n).$$

Hence $C_b(G) = \Theta(n^2)$, as required.

Note: Technically, we should use Wald's Equation to justify the above argument. Wald's Equation (MU Section 13.3) says that $\mathrm{E}[\sum_{i=1}^{T} X_i] = \mathrm{E}[T]\mu$*, where the* X_i *are iid r.v.'s with mean* μ *and* T *is a (random) stopping time. We are implicitly using this fact above when we bound the total time spent on excursions from* a *into the tail. The same applies to our analysis in part (d) below. We did not require you to explicitly appeal to Wald's Equation for this problem.*

A lot of students did not consider that when the walk reaches vertex a*, there is only a* $2/n$ *chance that it then goes into the tail. (They were able to compensate for this error by arguing that the expected time spent on the tail starting from the vertex to the right of* a *is only* $O(n)$*.)*

(d) As in part (c), we consider the two cases for the first step of the walk starting at a: *4pts*

- Conditioned on the walk moving right at a (which occurs with probability $\frac{1}{n/2}$), then with probability at least $1-\frac{2}{n}$ the walk will reach a again before reaching b. This follows from the Gambler's Ruin Problem (as discussed in lecture and section; see also MU, Section 7.2.1): this is the probability that the first player wins in a fair gambling situation between two players whose initial capitals are $\frac{n}{2}-1$ dollars and 1 dollar respectively. The walk will then take another $H_{a,b}$ steps to get to b. Hence, conditioned on moving right at a, the expected number of steps to reach b is at least $(1-\frac{2}{n})H_{a,b}$.
- Conditioned on the walk moving left at a (which occurs with probability $\frac{n/2-1}{n/2}$), then it will take an expected $\Omega(n)$ steps within the clique before returning to a. This follows from looking at a geometric distribution with mean $\frac{1}{n/2-1}$, which is the probability of reaching a from any vertex on the clique in one step. Hence, conditioned on moving left at a, the expected number of steps to reach b is at least $H_{a,b} + \Omega(n)$.

Combing the two cases, we obtain

$$H_{a,b} \geq \frac{1}{n/2}\Big(1-\frac{2}{n}\Big)H_{a,b} + \frac{n/2-1}{n/2}\Big(H_{a,b} + \Omega(n)\Big) = H_{a,b} - \frac{4}{n^2}H_{a,b} + \frac{n/2-1}{n/2}\Omega(n),$$

as required. Rearranging the terms yields: $\frac{4}{n^2}H_{a,b} \geq \Omega(n)$ and thus $H_{a,b} = \Omega(n^3)$.

(e) Clearly, $C(L_n) \geq H_{a,b} = \Omega(n^3)$ (using part (d)). Combined with the upper bound from part (b), this gives $C(L_n) = \Theta(n^3)$. *1pt*

(f) False. For a counterexample, take G to be the lollipop graph, and G' to be the complete graph on n vertices. Then $\Theta(n \log n) = C(G') < C(G) = \Theta(n^3)$. *1pt*

(g) False. Take G to be the line graph on n vertices, and G' to be the lollipop graph. Then $\Theta(n^3) = C(G') > C(G) = \Theta(n^2)$. *1pt*

3. (a) To see that the process is irreducible, note that from any configuration we can reach a configuration in which all the k particles sit consecutively on the cycle (by successively moving particles clockwise until they hit the next particle around the cycle). By a symmetrical argument, we can then reach any other configuration from this configuration. *3pts*

Quite a lot of students failed to specify how to move from any configuration to any other. Just saying "the process is irreducible" is not sufficient here (since it's definitely not obvious!).

(b) The process is aperiodic because some of the configurations have self-loops (i.e., have a non-zero probability of not changing in one step). These are any configurations in which two or more particles sit consecutively on the cycle. *2pts*

(c) We claim that the stationary distribution (which by parts (a) and (b) and the Fundamental Theorem (MU, Theorem 7.7) must be unique) is the uniform distribution over all $\binom{n}{k}$ configurations. To verify this, fix any configuration σ. Let g be the number of (maximal) consecutive groups of particles in σ. Then the number of (non-self-loop) transitions out of σ is exactly g (one for each group, corresponding to the "head" particle of the group moving clockwise one step). Similarly, the number of (non-self-loop) transitions *into* σ is also g (corresponding to the "tail" particle of each group having moved clockwise one step to its present position). Since all these transitions have the same probability $\frac{1}{k}$ (and since self-loops contribute the same transition probability into and out of σ), it follows that the sum of transition probabilities into σ is equal to the sum of transition probabilities out of σ, which is exactly 1. This means that the Markov chain is "doubly stochastic", so as we saw in class its stationary distribution must be uniform. *4pts*

4. (a) Fix X_t, Y_t. Suppose the positions of card c in X_t and Y_t are j and j' respectively. Then only the cards in positions i, j, j' may be affected by the move. If $j = j'$, then clearly d_t remains unchanged. If $j \neq j'$, then the move creates a new matched card (namely, card c at position i). At the same time, we may lose at most one existing match (from the cards that are at position i before the move). Hence, d_t never increases with t. *3pts*

(b) d_t decreases (by at least 1) precisely when there is no match at position i *and* the card c is not already matched up. The probability that we pick a position i whose cards differ in X_t and Y_t is $\frac{d_t}{n}$. Similarly, this is also the probability that we pick a card c that is unmatched. Since these events are independent, the probability that d_t decreases is $(\frac{d_t}{n})^2$. *3pts*

(c) If the distance is d_t, then by part (b) the time until d_t decreases is a geometric r.v. with parameter $(\frac{d_t}{n})^2$, so its expectation is $(\frac{n}{d_t})^2$. Hence, the expected number of steps T until $X_T = Y_T$ is at most $\sum_{d_t=1}^{n}(\frac{n}{d_t})^2 \leq n^2 \sum_{i=1}^{\infty} \frac{1}{i^2} = \frac{\pi^2}{6}n^2$. *3pts*

(d) Applying Markov's inequality to the statement from part (c), we deduce that for any choice of initial states X_0, Y_0, the probability that starting from X_0, Y_0 we take more than $\frac{cn^2}{\varepsilon}$ steps until $X_T = Y_T$ is at most ε; i.e., for $T = \frac{cn^2}{\varepsilon}$ we have $\Pr[X_T \neq Y_T | X_0, Y_0] \leq \varepsilon$. Hence, by the Coupling Lemma (MU, Lemma 12.2), the mixing time is bounded by $\tau(\varepsilon) \leq \frac{cn^2}{\varepsilon}$. *3pts*

5. **[Optional: not graded]** My strategy is simply to perform exactly the same procedure as you, except that I start with some arbitrary card (say, the first card). The key observation is that if our two cards should ever coincide, then our two processes will remain identical at all future times, and thus I will identify your card correctly. Now each time I finish a count and look at the next (random) card, there is a positive probability that its value will be d, the remainder of your current count. Thus if we were to continue this process indefinitely (with an infinite deck), coupling would occur in finite time with probability 1. In the finite setting, with 52 cards and ten choices for the first card, it turns out that there is a pretty good chance (around 70%) that coupling will have occurred before the end of the deck is reached. This can be a fun trick to try on friends; the fact that you sometimes lose gives them an incentive to keep trying to win. The precise success probability for this trick is still an open problem; those interested in reading the state of the art on it may refer to "The Kruskal Count," by J. Lagarias, E. Rains and R.J. Vanderbei, http://arxiv.org/abs/math/0110143. *0pts*