Probability and Computation: Problem Sheet 3 Solutions

You are encouraged to submit your solutions at student reception or by emailing them to jas289 by 2pm Friday 14th of February

Question 1 (Schöning: tighter analysis)**.** *Use the following version of Schöning's Algorithm:*

> *(1) Start with a random truth assignment.*
>
> *(2) Repeat up to $3n$ times, terminating if all clauses are satisfied:*
>
> *(a) Choose an arbitrary clause that is not satisfied*
>
> *(b) Choose one of it's literals UAR and switch the variables value.*
>
> *(3) If a valid solution is found* `return` *it. O/W* `return` *unsatisfiable*

(i) Fix some satisfying assignment α. Let A_k be the event that the random assignment from step (1) disagrees with α on exactly k literals/variables. What is $\mathbf{P}[\, A_k \,]$?

(ii) Let P_k be the probability that we make $\leq k$ incorrect steps within our first $3k$ steps. Prove

$$P_k \geq \binom{3k}{k}\left(\frac{2}{3}\right)^k\left(\frac{1}{3}\right)^{2k}.$$

(iii) Recall Stirling's inequality $\sqrt{2\pi} \leq \frac{n!}{n^{n+1/2}e^{-n}} \leq e$, and show that $P_k \geq \frac{2^{-k}}{3\sqrt{k}}$.

(iv) Show that if a solution exists, Schöning's Algorithm succeeds with probability at least $\left(\frac{3}{4}\right)^n/(3\sqrt{n})$

(v) Deduce a bound on the time to find a solution w.h.p. using Schöning's Algorithm as above.

Solution: Part (i): Consider some arbitrary correct satisfying assignment α. Then since the random initial truth assignment x was uniformly random each literal is the same as in α with probability $1/2$ and each independently. So number which disagree is thus binomially distributed $\mathrm{Bin}(n, 1/2)$ and thus,

$$\mathbf{P}[\, A_k \,] = \binom{n}{k}\frac{1}{2^n}.$$

Part (ii): The probability P_k we make $\leq k$ incorrect steps within our first $3k$ steps is at least the probability we make exactly k incorrect steps in the first $3k$, which is

$$\binom{3k}{k}\left(\frac{2}{3}\right)^k\left(\frac{1}{3}\right)^{2k}.$$

Part (iii): We have

$$P_k \geq \binom{3k}{k}\left(\frac{2}{3}\right)^k\left(\frac{1}{3}\right)^{2k} \geq \frac{\sqrt{6\pi k}\left(\frac{3k}{e}\right)^{3k}}{e\sqrt{k}\left(\frac{2k}{e}\right)^{2k}\cdot e\sqrt{k}\left(\frac{k}{e}\right)^{k}}\cdot\frac{2^k}{3^{3k}} \geq \frac{2^{-k}}{3\sqrt{k}}$$

Where above we used the exact form of Stirling's ineq: $\sqrt{2\pi} \leq \frac{n!}{n^{n+1/2}e^{-n}} \leq e$.

Part (iv): By the previous two parts $\mathbf{P}[\,S|A_k\,] \geq 2^{-k}/3\sqrt{k}$, where S is the event of success. Hence,

$$\begin{aligned}\mathbf{P}[\,S\,] &= \sum_{k=0}^{n}\mathbf{P}[\,S|A_k\,]\,\mathbf{P}[\,A_k\,]\\ &\geq 2^{-n} + \sum_{k=1}^{n}\frac{2^{-k}}{3\sqrt{k}}\binom{n}{k}\frac{1}{2^n}\\ &\geq \frac{2^{-n}}{3\sqrt{n}}\sum_{k=0}^{n}\binom{n}{k}\frac{1}{2^k}\\ &= \frac{2^{-n}}{3\sqrt{n}}\left(1+\frac{1}{2}\right)^n = \left(\frac{3}{4}\right)^n/(3\sqrt{n}).\end{aligned}$$

Part (iv): By the boosting lemma this is $O\big(n\cdot\sqrt{n}\left(\frac{4}{3}\right)^n\cdot\log n\big) = O\big(n^{1.33334}\big)$ since each run of the algorithm takes $O(n)$ time.

Question 2. *Consider the following Markov Chains*

$$A = \begin{pmatrix} 0 & 1/9 & 2/9 & 2/3\\ 1/7 & 1/7 & 5/7 & 0\\ 2/9 & 5/9 & 0 & 2/9\\ 3/5 & 0 & 1/5 & 1/5 \end{pmatrix}$$

$$B = \begin{pmatrix} 0 & 1/2 & 0 & 1/2\\ 0 & 0 & 2/3 & 1/3\\ 0 & 1/3 & 1/3 & 1/3\\ 1/3 & 1/3 & 1/3 & 0 \end{pmatrix}$$

$$C = \begin{pmatrix} 0 & 2/3 & 0 & 1/3 & 0 & 0\\ 2/5 & 0 & 0 & 3/5 & 0 & 0\\ 0 & 0 & 0 & 0 & 2/3 & 1/3\\ 1/4 & 3/4 & 0 & 0 & 0 & 0\\ 0 & 0 & 2/3 & 0 & 0 & 1/3\\ 0 & 0 & 1/3 & 0 & 1/3 & 1/3 \end{pmatrix}$$

(i) *Which of the above are irreducible?*

(ii) *Which of the above are reversible?*

(iii) *Calculate the stationary distribution of the reversible irreducible chain(s) above.*

(iv) *In lectures we showed that any finite irreducible chain has a unique stationary distribution. Give an example of a (finite :p) reducible chain with more than one stationary distribution.*

Question 3. *Recall that an undirected weighted graph $G = (V, E, w)$ is an undirected graph with a weight function $w : E \to \mathbb{R}_+$ which is positive and symmetric, that is for any $ij \in E$, $w(ij) = w(ji) > 0$.*

(i) *Let P be a Markov chain. Show that P is reversible if and only if P is a simple random walk on an undirected weighted graph G.*

(ii) *Show that a simple random walk on an undirected graph $G = (V, E, w)$ has stationary distribution*

$$\pi(x) = \frac{\sum_{xy\in E} w(xy)}{2\sum_{e\in E} w(e)}, \qquad \textit{for all } x \in V.$$

(iii) *Given a reversible Markov chain P, show that P is irreducible if and only if the associated undirected weighted graph G is connected.*

Solution: *Part (i):* We start showing that a random walk on an undirected graph is always reversible, i.e., it satisfies the detailed balance condition: $\pi(u)P(u,v) = \pi(v)P(v,u)$ for any $u, v \in V$. For ease of notation define $d(v) = \sum\limits_{z \in V} w(v,z)$, this is the weighted analogue of the degree. Let u, v be arbitrary vertices. Then,

$$\pi(u)P(u,v) = \frac{d(u)}{\sum_{z\in V} d(z)} \cdot \frac{w(u,v)}{d(u)} = \frac{d(v)}{\sum_{z\in V} d(z)} \cdot \frac{w(v,v)}{d(v)} = \pi(v)P(v,u)$$

where the second equation follows from the fact that, since G is undirected, $w(u,v) = w(v,u)$.
We now show the reverse implication: we are given a transition matrix P on Ω with stationary distribution π such that $\pi(u)P(u,v) = \pi(v)P(v,u)$ for any $u, v \in \Omega$, and we want to show we can construct an undirected weighted graph $G = (V, E, w)$ such that P is the transition matrix of a random walk on G. First of all, we choose $\Omega = V$. Then, we construct the weight function $w \colon V \to \mathbb{R}_{\geq 0}$ as $w(u,v) = \pi(u)P(u,v)$, and we set $E = \{\{u,v\} \colon w(u,v) > 0\}$. We need to show that w is a proper weight function. Clearly, w is nonnegative and strictly positive exactly on E. Moreover,

$$w(u,v) = \pi(u)P(u,v) = \pi(v)P(v,u) = w(v,u),$$

since P is reversible. Finally,

$$\frac{w(u)}{\sum_{z\in V} w(u,z)} = \frac{\pi(u)P(u,v)}{\sum_{z\in V} \pi(u)P(u,z)} = \frac{P(u,v)}{\sum_{z\in V} P(u,z)} = P(u,v),$$

where the last equality follows from the fact that each row of P sum up to 1 (by definition of transition matrix). Hence, we have shown that a random walk on G has transition matrix P.

Part (ii): Let $\pi \colon V \to \mathbb{R}$ such that $\pi(u) = \frac{d(u)}{\sum_{z\in V} d(z)}$. Then, π is stationary for P. To prove this, we just need to check that $\pi P = \pi$. Let u be an arbitrary vertex. Then,

$$(\pi P)(u) = \sum_{v\in V} \pi(v)P(v,u) = \sum_{v\in V} \frac{d(v)P(v,u)}{\sum_{z\in V} d(z)} \sum_{v\in V} \frac{w(v,u)}{\sum_{z\in V} d(z)} = \frac{d(v)}{\sum_{z\in V} d(z)} = \pi(v).$$

where the third equality follows from $P(v,u) = w(v,u)/d(v)$.

Question 4. *Recall that a probability vector (distribution) is a non-negative real vector whose elements sum to* 1*. A stochastic matrix is a real square matrix, where each row is a probability vector. Observe every Stochastic matrix gives rise to a Markov chain and vice versa.*

(i) Let $\nu \in \mathbb{R}^n_+$ be a probability vector and $M \in \mathbb{R}^{n\times n}_+$ be a stochastic matrix. Show that νM is a probability vector.

A doubly stochastic matrix is a real square matrix, where each row and column is a probability vector.

(ii) Prove that the uniform distribution is stationary for any Markov chain whose transition matrix is doubly stochastic.

Question 5. *Show that if P is the transition matrix of a reversible Markov chain then the matrix P^t also defines a reversible Markov chain.*

Solution: Assuming that $\pi(x)P_{x,y} = \pi(y)P_{y,x}$ we aim to prove

$$\pi(x)P^t_{x,y} = \pi(y)P^t_{y,x}, \tag{1}$$

by induction. Observe that $\pi(x)P_{x,y} = \pi(y)P_{y,x}$ is the base case. Lets assume (1) holds, then

$$\begin{aligned}
\pi(x)P(x,y)^{t+1} &= \pi(x) \sum_{z\in V} P(x,z)^t \cdot P(z,y) \\
&= \pi(x) \sum_{z\in V} \frac{\pi(z)}{\pi(x)} P(z,x)^t \cdot \frac{\pi(y)}{\pi(z)} P(y,z) \\
&= \pi(y) \sum_{z\in V} P(y,z) \cdot P(z,x)^t \\
&= \pi(y)P(y,x)^{t+1}.
\end{aligned}$$

BONUS SOLUTION after typing this I found another proof typed by Luca last year:

The easiest way is probably to use the fact that P is reversible if and only if $\langle Pf, g\rangle_\pi = \langle f, Pg\rangle_\pi$ for any $f, g \in V \to \mathbb{R}$. Hence, we need to show that, for arbitrary $f, g \in V \to \mathbb{R}$, $\langle P^t f, g\rangle_\pi = \langle f, P^t g\rangle_\pi$:

$$\langle P^t f, g\rangle_\pi = \langle P^{t-1} f, Pg\rangle_\pi = \langle P^{t-2} f, P^2 g\rangle_\pi = \cdots = \langle f, P^t g\rangle_\pi$$

where at each step we have applied the reversibility of P.

Question 6. *Consider the Complete graph K_n which is the graph on n vertices where each pair of vertices is connected by an edge. Let $x, y \in V$ where $x \neq y$.*

(i) Show that $\mathbf{E}_x\left[\tau_y^+\right] = n - 1$.

(ii) What is the distribution of τ_y^+?

Question 7. *State j is accessible from state i if, for some integer $n \geq 0$, $P_{i,j}^n > 0$. If two states i and j are accessible from each other, we say that they communicate and we write $i \sim j$. Prove that communicating relation $\sim$ defines an equivalence relation.*

Question 8. *Prove rigorously the claim made in lecture that the expected time for RAND 2-SAT to find a given solution is at most the hitting time $h(0, n)$ of the random walk on a path.*

Solution: (Credit: Dmitros Los) Let $h'(i)$ be the expected time for Rand-2-SAT to find our favourite solution α starting from i literals in agreement with α. A conceptually simple proof that $h(0)$ is at most the hitting time of n from 0 by the SRW on a path is to show directly that

$$h'(k) \leq h'(k+1) + 2k + 1,$$

by induction. The reason is that then

$$\begin{aligned} h'(0) &\leq h'(1) + 1 \leq (h'(2) + 2 + 1) + 1 \\ &\leq h'(3) + 2 \cdot 2 + 1 + 2 + 1 + 1 \\ &\leq \cdots \\ &\leq 2\sum_{k=1}^{n-1} k + n \\ &= n^2, \end{aligned}$$

which just so happens to be $h(0)$, the hitting time of n from the random walk started at 0.

(Base case) For $k = 0$, we have $h'(0) = h'(1) + 1 = h'(1) + 2 \cdot 0 + 1$.

(Inductive step) Assume that it holds for k , then for $k + 1$ we have

$$\begin{aligned} h'(k+1) &\leq \frac{1}{2}(h'(k) + h'(k+2)) + 1 \\ &\leq \frac{1}{2}(h'(k+1) + 2k + 1 + h'(k+2)) + 1 \end{aligned}$$

Rearranging this yields

$$h'(k+1) \leq 2k + 1 + h'(k+2) + 2 = h'(k+2) + 2(k+1) + 1,$$

as claimed.
