CS174 **Final Exam Solutions** **Spring 2000**
J. Canny **May 17**

1. (20 points)

   (a) (4 points) What is the probability that a random permutation of $(1, \ldots, n)$ has *at least one* fixed point? Assume n is large.

   The number of fixed points in a random permutation has Poisson(1) distribution. Hence the probability of no fixed points is $1/e$ and the probability of at least one fixed point is $1 - 1/e$.

   (b) (4 points) If $\mathrm{Cov}(X, Y) = \mathrm{E}[(X - \overline{X})(Y - \overline{Y})]$ where $\overline{X} = \mathrm{E}[X]$, express $\mathrm{Cov}(X, Y)$ in terms of $\mathrm{E}[X]$, $\mathrm{E}[Y]$ and $\mathrm{E}[XY]$.

   Distributing the expression, we get:

   $$Cov(X, Y) = \mathrm{E}[XY - \overline{X}Y - X\overline{Y} + \overline{XY}] = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] - \mathrm{E}[X]\mathrm{E}[Y] + \mathrm{E}[X]\mathrm{E}[Y]$$

   And the last expression simplifies to $\mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]$.

   (c) (4 points) Suppose you draw a random card from a pack of n cards, look at it, replace it in the pack, and then reshuffle. Repeat m times. What is the expected value of m until you have seen all the cards?

   This is the coupon collector's problem, so we should expect to draw $n \ln n + O(n)$ cards.

   (d) (4 points) Suppose you want to bound the probability of more than $(3/4)n$ heads in n tosses of a fair coin. Out of Markov, Chebyshev and Chernoff, which bound would you use and why? Assuming n is large, you shouldnt need to compute anything to answer this question.

   Since all the tosses are independent, we can use Chernoff bounds which are the strongest bounds of the three.

   (e) (4 points) Suppose you draw m cards at random from a deck of n cards, with replacement. How large should n be *in terms of* m so that you have small probability (say less than half) of drawing the same card twice?

   This is the birthday paradox. So we should have $n > m^2$ to insure a probability of less than half of drawing the same card twice.

2. (20 points) Let $G_{n,p}$ be a random graph on n vertices with edge probability p. How large should p be to be almost certain that the graph has a 5-clique? You dont need to analyze variances. Assume the distribution of the number of cliques converges about its mean for large n.

   Consider a fixed set of 5 vertices. The probability that this given set is a 5-clique is p^{10} since there are 10 vertices in a 5-clique. The number of possible sets of 5 vertices in G is $\binom{n}{5} \approx n^5/120$. This means the expected number of 5-cliques in the graph as n gets large is

$\frac{n^5}{120}p^{10}$. If we assume that the distribution converges about the mean for large n, we just need to see what minimum value of p makes the expectation ≥ 1:

$$\frac{n^5}{120}p^{10} = 1 \leftrightarrow p = \frac{120^{\frac{1}{10}}}{\sqrt{n}}$$

3. (20 points) Recall that a n-vertex graph $G = (V, E)$ is a c-expander iff every subset $A \subset V$ of $|A| \leq n/2$ vertices satisfies:

$$|N(A)| \geq (1 + c)|A|$$

where $N(A)$ is the set of vertices in A and their neighbors. If a graph is a $1/2$-expander, give an upper bound on the path length between any two vertices in the graph. Include the constant factor in your answer.

Consider any two vertices u and v. Since G is an expander, the number of vertices reachable from u in k or fewer steps is $(1 + c)^k$ as long as this number does not exceed $n/2$. This means that in $\log_{1+c} n/2$ steps from u, we can reach half of the vertices of G. Similarly, in $\log_{1+c} n/2$ steps from v, we can reach half of the vertices of G. Since the two halves must have a vertex in common, there must be a path of length $\leq 2 \log_{1+c} n/2$ between u and v in G.

4. (20 points)

   (a) (5 points) How many elements are there in the multiplicative group $\mathbb{Z}_{49}^*$ of integers mod 49?

   The number of elements in $\mathbb{Z}_{49}^*$ is the number of elements relatively prime to 49 which is $\phi(49) = \phi(7^2) = 7(7 - 1) = 42$.

   (b) (5 points) The group $\mathbb{Z}_{49}^*$ is cyclic. How many *generators* does it have?

   For any generator g, g^c is a generator if c is relatively prime to order of the group. Hence the number of generators is $\phi(\phi(49)) = \phi(42) = \phi(2 \times 3 \times 7) = (2-1)(3-1)(7-1) = 12$.

   (c) (5 points) Let g be a generator of $\mathbb{Z}_{49}^*$. What is the smallest power m of g such that $g^m = 1 (\mathrm{mod}\ 49)$?

   By Fermat's theorem, $m \leq 42$ since for any element a, the theorem tells us that $a^{\phi(n)} = 1 \bmod n$. Since g is a generator, the powers of g must enumerate the whole group before enumerating 1, which means that $m \geq 42$. Together, the two statements tell us that $m = 42$.

   (d) (5 points) Is $\mathbb{Z}_{49}$ a field? How many roots does the equation $X^2 = 0 (\mathrm{mod}\ 49)$ have?

   $X^2 = 0 \bmod 49$ means that X^2 is a multiple of 49, which in turn means that X is a multiple of 7. There are 7 multiples of 7 within $\mathbb{Z}_{49}$, so there are 7 roots to this equation. $\mathbb{Z}_{49}$ is therefore not a field.

5. (20 points)

   (a) (10 points) Briefly describe a protocol for a zero-knowledge proof that a graph G has a hamiltonian cycle (a cycle that visits every vertex exactly once). You dont need to show that this protocol succeeds, just write it down.

   The prover chooses a random isomorphism of G, call it $G\prime$. Prover computes the adjacency matrix of $G\prime$ and bit-commits the isomorphism, the adjacency matrix and the hamiltonian cycle in $G\prime$. Verifier chooses randomly to see either the entire matrix and the isomorphism (to verify that G is isomorphic to $G\prime$), or the hamiltonian cycle and the corresponding edges in the adjacency matrix. This is similar to the ZKP of the k-clique problem.

   (b) (10 points) Explain how to make this proof non-interactive. Again, just give a protocol, dont show that it works.

   Let the prover make k independent versions of the proof above. Call them $P_1, \ldots, P_k$. Then, instead of getting random bits from the verifier, the prover computes $r = h(G, P_1, \ldots, P_k)$ where h is a public and secure hash function returning a k-bit number. Then the prover uses the i'th bit of r to choose his reponse to the P_i. This is very similar to problem 2 in the last homework (hw12).

6. (20 points)

   (a) (10 points) Let (n, e) be your RSA public key and (n, d) be your private key where $ed = 1 (\mathrm{mod}\ \ \phi(n))$. Give a zero-knowledge proof protocol so that you (prover) can demonstrate to a verifier that you know d. Hint: the verifier starts by sending a random number to the prover.

   Verifier asks the prover to digitally sign a random message. Alternatively, the verifier can encrypt a random message with the prover's public key and ask the prover to decrypt it.

   (b) (10 points) Given what you know about digital signatures, explain what might be dangerous for you the prover to follow the protocol above.

   You end up signing something you can't read, which could commit you to something against your will. In the alternative protocol, you could end up decrypting a message which was meant to remain confidential.