ECE 559BH Distributed Network Algorithms **SPRING 2006**
PROBLEM SET 4 **Due Tuesday, March 28**

Byzantine agreement problem, LT codes

Reading: See website.

Problems to be handed in:

1. Byzantine agreement (S. Yang)
The expected number of rounds for the randomized algorithm ByzGen to reach agreement is bounded by a constant (three or four, depending on definition) as long as $t < n/8$, but the algorithm uses common randomness–all nodes see the same coin flip in each round. Specify a variation of ByzGen which uses private, but not common, randomness, and reaches agreement in a constant number of rounds if $t \leq \sqrt{n}$. (Hint: What if each nodes flips a fair coin, and reports the result to all neighbors?)

SOLUTION 1 Let ModByzGen be a modified version of ByzGen. Each round of ModByzGen requires two rounds of communication, first one in which votes are sent, and a second in which all nodes send the results of a random coin flip. The nodes execute ByzGen as before, except they take a majority vote of the reported coin flips and use that as the coin flip variable of ByzGen. Say the coin flip mechanism works in the round if all reliable processors use the majority vote of coin flips. If $t \leq \sqrt{t}$ then the coin flip mechanism works and produces a 1 with probability at least $Q(1) - \epsilon$, and a 0 with probability at least $Q(1) - \epsilon$, for any fixed $\epsilon > 0$ and n large enough. Note that if the majority votes of coin flips don't all agree, no harm is done. The probability all nodes report common values in the next round is thus at least $Q(1) - \epsilon$. The mean number of rounds until convergence is bounded above by $\frac{2}{Q(1)-\epsilon} + 2$.

SOLUTION 2 Let ModByzGen be a modified version of ByzGen. Each round of ModByzGen requires two rounds of communication, first one in which votes are sent, and a second in which all nodes send the results of a random coin flip. Let $G = 0.8n$. The nodes execute ByzGen as before, except instead of using L and H, a node i takes its threshold to be $\tau_i = \frac{n_i}{3} + 0.5$, where n_i is the number of reported 1's that i received from the coin toss part of the round. Note that the range of these thresholds is at most $t/3$.

We say that the faulty processors *foil* a threshold x if they cause tally to exceed x for at least one good node and tally to be no more than x for at least one good node. Given the tally's, the set of thresholds that could foil a tally is an interval $[A, B]$ of length at most $\sqrt{n}$. If Z represents the number of coin flips of good processors, then unless Z is in some interval of length at most $6\sqrt{n}$, either all the thresholds are less than A or all the thresholds are greater than B. The probability that Z falls outside an interval of length $6\sqrt{n}$ is greater than or equal to the probability if falls outside the centered interval of that length. By the central limit theorem, the probability that Z differs from its mean by at least $3\sqrt{n}$ converges to $2Q(3) \approx 10\%$ as $n \to \infty$, so that for all large n, this probability is at least 0.5. Thus if $\epsilon < 2Q(3)$, for all large n, there is at least chance *epsilon* that all nodes have the same vote at the end of a round. Thus, the mean number of rounds is at most $\frac{1}{\epsilon} + 1$.

2. Bounding a martingale (can be used for a minor variation of a part of Luby's proof)
Let $\boldsymbol{\mathcal{F}} = (\mathcal{F}_t : t \in \mathbb{Z}_+)$ be a filtration. Suppose X_0 is an integrable, $\mathcal{F}_0$ measurable random variable, $m = (m_t : t \in \mathbb{Z}_+)$ is an adapted process with values in the positive integers, and for each $t \geq 0$, V_t is an $\mathcal{F}_{t+1}$ measurable random variable such that, given $\mathcal{F}_t$, its conditional distribution is binomial with parameters m_t and $\frac{1}{m_t}$. That is, $(V_t|\mathcal{F}_t) \sim \text{Binom}(m_t, \frac{1}{m_t})$. Let $X_t = X_0 + V_0 + \cdots + V_{t-1} - t$.
(a) Show that X is a martingale relative to $\boldsymbol{\mathcal{F}}$.
(b) Show that if $V \sim \text{Binom}(n, \frac{1}{n})$ for some integer $n \geq 1$, then $E[e^{\alpha(V-1)}] \leq e^{\alpha^2}$ for $|\alpha| \leq 1$.
(c) Show that $E[e^{\alpha(X_t - X_0)}] \leq e^{\alpha^2 t}$ for $|\alpha| \leq 1$.
(d) Using a Chernoff bound, show that for $\lambda \geq 0$, $P\{|X_t - X_0| \geq \lambda\} \leq \begin{cases} 2e^{-\frac{\lambda}{2}} & \text{if } t \leq \frac{\lambda}{2} \\ 2e^{-\frac{\lambda^2}{4t}} & \text{if } t \geq \frac{\lambda}{2} \end{cases}$

SOLUTION (a) Clearly X is adapted and $E[|X_t|] < \infty$ for all t. Also,
$E[X_{t+1} - X_t|\mathcal{F}_t] = E[V_t|\mathcal{F}_t] - 1 = \frac{m_t}{m_t} - 1 = 0$. Thus, X is a martingale relative to $\boldsymbol{\mathcal{F}}$.

(b) $E[e^{\alpha(V-1)}] = e^{-\alpha}(1+\frac{e^{\alpha}-1}{m})^m \le e^{e^{\alpha}-1-\alpha}$, where we used the fact $1+x \le e^x$ for all x. By the Taylor series expansion,

$$\begin{aligned} e^{\alpha}-1-\alpha &= \frac{\alpha^2}{2}+\frac{\alpha^3}{3!}+\cdots \\ &\le \alpha^2(\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots) \quad (\text{if } |\alpha| \le 1) \\ &= \alpha^2 \end{aligned}$$

(c) This is similar to the proof of the Azuma-Hoeffding inequality. Proof by induction. The bound holds for $t=0$. Suppose it holds for some $t \ge 0$. Then
$E[e^{\alpha(X_{t+1}-X_0)}] = E[E[e^{\alpha(X_{t+1}-X_0)}|\mathcal{F}_t]] = E[e^{\alpha(X_t-X_0)}E[e^{\alpha V_t}|\mathcal{F}_t]] \le E[e^{\alpha(X_t-X_0)}]e^{\alpha^2} \le e^{\alpha^2(t+1)}$. Thus the bound holds for $t+1$, and hence for all t by induction.
(d) For $\lambda > 0$ and $0 \le \alpha \le 1$, $P\{X_t - X_0 \ge \lambda\} \le E[e^{\alpha(X_t-X_0-\lambda)}] \le e^{-(\alpha\lambda-\alpha^2 t)}$. We next want to select $\alpha \in [0,1]$ to maximize $\alpha\lambda - \alpha^2 t$. Thus, take $\alpha = \begin{cases} 1 & \text{if } t \le \frac{\lambda}{2} \\ \frac{\lambda}{2t} & \text{if } t \ge \frac{\lambda}{2} \end{cases}$, which yields the bound as desired.

3. Reserved packets in the gross ripple
(This problem addresses the output symbols contributed by the term $\tau(k/R)$ in Luby's paper on LT codes.) Suppose that k is large and $R = 3\sqrt{k}\ln(3k/\delta)$ for δ fixed with $0 < \delta < 1$ (that is, R grows somewhat faster than $\sqrt{k}$). Let $t_o = k - R$ and $d = k/R$. Consider a special output symbol of degree d which has been reserved for use after t_o symbols have been decoded, if the special output symbol enters the gross ripple before t_o. (The special output symbol can be forced to leave the ripple early if some other output symbol connected to the same input symbol is decoded.) Let p denote the probability the special output symbol is in the gross ripple after t_o symbols have been decoded, assuming the decoding process successfully decodes at least t_o symbols.
(a) Argue that $p \approx e^{-1}$ based on a Poisson approximation.
(b) Find the exact value of p, and show it converges to e^{-1} as $k \to \infty$. (Hint: $\frac{d}{k} = \frac{o(1)}{d}$.)

SOLUTION (a) The output symbol is in the gross ripple at time t_0 if and only if it is initially connected to exactly one of the R symbols not decoded by time t_0. If the endpoints of the d edges coming out of the coded symbol were selected independently, then each edge would be connected to one of the last R input symbols with probability $\frac{k-t_0}{k} = \frac{1}{d}$, and so the number of such edges would have the binomial distribution, or approximately the Poisson distribution, with mean one. Thus number is thus exactly one with probability $\lambda e^{-\lambda}|_{\lambda=1} = e^{-1} \approx 0.368$.
(b) $p = \frac{d(k-t_0)(t_0)(t_0-1)\cdots(t_0-d+2)}{k^d}$. Thus, $(1-\frac{1}{d}-\frac{d}{k})^{d-1} \le p \le (1-\frac{1}{d})^{d-1}$. Since $1-\frac{1}{d}-\frac{d}{k} = 1-\frac{1+o(1)}{d}$, both the lower and upper bounds converge to e^{-1} as $k \to \infty$, so that $p \to e^{-1}$.

4. LT decoding with half degree one and half degree two packets
Consider LT decoding for $k >> 1$ input symbols based on $Poi(k)$ output symbols. Suppose each output symbol is initially degree one or two, with equal probability, independently of all other output symbols.
(a) What is the expected number of input symbols included in one or more output symbols?
(b) What fraction of input symbols are successfully decoded by the LT decoding algorithm (with high probability for large k)?
(c) What fraction of output symbols are not decoded?
(d) What fraction of output symbols are connected to two input symbols which have degree one? (Such output symbols are not decoded. Are these essentially all of the output symbols that are not decoded?)

SOLUTION (a) A given input symbol is connected to a given output symbol with probability $\frac{1}{2}\frac{1}{k}+\frac{1}{2}\frac{2}{k} = \frac{1.5}{k}$. The expected number of output symbols is k, so the expected degree of an input symbol is $\frac{3}{2}$. Thus, a given input symbol is included in one or more output symbols with probability exactly $1-e^{-1.5} \approx 0.77687$. This probability is also the expected fraction of input symbols included in at least one output symbol.
(b) Since $\beta(t) = \frac{t+t^2}{2}$ and $\beta'(t) = \frac{1}{2}+t$, the fraction of input symbols decoded is the positive root z^* of $\frac{1}{2}+t+\ln(1-t) = 0$.
(c) Output symbols with degree two connected to two uncoded packets will not be decoded. This is a limiting

fraction $\frac{(1-z^*)^2}{2}$ of all output packets (for k large).
(d) Consider an output symbol of degree two. Each of the two input symbols connected to it is included in no other output symbol with limiting probability $e^{-1.5}$. Thus, the desired fraction is $\frac{1}{2}(e^{-1.5})^2$. The answers to (c) and (d) are different, so these are not the only output symbols that are not decoded. (The algorithm gets stuck when there are no degree one output symbols left. The remaining configuration of degree two symbols can be arbitrary.)