Note: These notes are based on material from James Aspnes textbook (see reference below) and [1]

1 The Consensus Problem

The set of processes are divided into *good*: follow the protocol; and *bad*: may deviate from protocol.

In the *Consensus Problem*, every process starts with a single input bit. We require that all good processes terminate and also:

- **Agreement:** All non-faulty processes decide the same value
- **Validity:** The value decided must equal the input bit of some good process

1.1 Problem Model

We assume *synchronous communication*: there is some known bound on how long it takes any message to get from a sender to a receiver. In particular, an upper bound (Δ) on the message transit time is known to all processors. Hence, the processors can operate in rounds of duration Δ.

We will assume that up to f processors can fail. When a process fails, one or more of its outgoing messages are lost in the round it fails in and no processes ever hear from it in subsequent rounds.

2 Some Probability

Lemma 1. *(Linearity of Expectation) Given a set of random variables $X_1, \ldots X_n$,*

$$E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i)$$

Proof: We first prove this for two random variables X and Y.

$$\begin{aligned} E(X+Y) &= \sum_{x\in X}\sum_{y\in Y}(x+y)Pr(X=x, Y=y) \\ &= \sum_{x\in X}\sum_{y\in Y} x\cdot Pr(X=x, Y=y) + \sum_{y\in Y}\sum_{x\in X} y\cdot Pr(X=x, Y=y) \\ &= \sum_{x\in X} x\cdot Pr(X=x) + \sum_{y\in Y} y\cdot Pr(Y=y) \\ &= E(X)+E(Y) \end{aligned}$$

The general result for n random variables now follows by induction. □

Lemma 2. *(Markov's Inequality) Let X be a random variable that only takes on nonnegative values (i.e. $X \geq 0$ always). Then for any $\lambda > 0$,*

$$Pr(X \geq \lambda) \leq \frac{E(X)}{\lambda}.$$

Proof: Assume not. Then for some value $\lambda > 0$, $Pr(X \geq \lambda) > \frac{E(X)}{\lambda}$. If this is true, then the expected value of X can be bounded as:

$$\begin{aligned} E(X) &\geq \sum_{i \geq \lambda} iPr(X = i) \\ &\geq \sum_{i \geq \lambda} \lambda Pr(X = i) \\ &= \lambda Pr(X \geq \lambda) \\ &> \lambda \frac{E(X)}{\lambda} \\ &= E(X) \end{aligned}$$

But this sequence of inequalities implies that $E(X) > E(X)$, which is clearly a contradiction. □

2.1 Chernoff Bounds

The following important bound only works for independent random variables. We prove it for 0/1-valued random variables: the r.v.'s only take on the values 0 or 1, and we prove an upper bound. The lemma generalizes easily to get a lower bound.

Lemma 3. *(Chernoff bounds) Let $X_1, \ldots, X_n$ be independent 0/1-valued random variables and let $p_i = E(X_i)$, where $0 \leq p_i < 1$ for all i. Then the sum $X = \sum_i X_i$, which has mean $\mu = E(X) = \sum_i p_i$ satisfies*

$$Pr(X \geq (1 + \delta)\mu) \leq (c_\delta)^\mu,$$

where $c_\delta = \frac{e^\delta}{(1+\delta)^{1+\delta}}$.

Proof: Consider an arbitrary positive constant t, to be set later, and consider the random variable e^{tX}. (If $X = 2$, say, this rv is e^{2t}.). A nice property of this random variable is the following:

$$\begin{aligned} E(e^{tX}) &= E\left(e^{t\sum_i X_i}\right) \\ &= E\left(\prod_{i \in [1,n]} e^{tX_i}\right) \\ &= \prod_{i \in [1,n]} E\left(e^{tX_i}\right) \end{aligned}$$

The last inequality holds since the X_i random variables are independent, and hence so are the e^{tX_i} random variables; and since $E(XY) = E(X)E(Y)$ if X and Y are independent. Note that

$$E(e^{tX_i}) = (1 - p_i) + p_i e^t.$$

Thus, we have:

$$\begin{aligned} \prod_{i \in [1,n]} E(e^{tX_i}) &= \prod_{i \in [1,n]} [1 + p_i(e^t - 1)] \\ &\leq \prod_{i \in [1,n]} e^{p_i(e^t - 1)} \\ &\leq e^{\mu(e^t - 1)} \end{aligned}$$

In the above, the second step holds by the inequality $1 + x \leq e^x$ (via Taylor expansion of e. Recall that $e^x = 1 + x + x^2/2! + x^3/3! + \ldots$). Markov's inequality says that for any positive r.v. Y, and any $\lambda > 0$, $Pr(Y \geq \lambda) \leq E(Y)/\lambda$. In the following, $Y = e^{tX}$, and $\lambda = e^{t(1+\delta)\mu}$.

$$\begin{aligned} Pr(X \geq (1+\delta)\mu) &= Pr(e^{tX} \geq e^{t(1+\delta)\mu}) \\ &\leq \frac{e^{\mu(e^t-1)}}{e^{t(1+\delta)\mu}} && \text{Markov's} \\ &\leq e^{\mu((e^t-1)-t(1+\delta))} \end{aligned}$$

This holds for any positive t, and is minimized when $t = \ln(1+\delta)$ (to see this, differentiate to get the minimum). This gives the lemma statement. □

Using a symmetric argument, we can bound the probability of deviation below the mean. Combining the results and using some approximations gives the following extremely useful lemma.

Lemma 4. *Let $X_1, \ldots X_n$ be independent Poisson trials such that $P(X_i = 1) = p_i$. Let $X = \sum_i X_i$ and $\mu = E(X)$. Then for $0 \leq \delta \leq 1$,*

$$Pr(|X - \mu| \leq \delta\mu) \leq 2e^{-\mu\delta^2/3}$$

2.2 Using Chernoff Bounds

Assume we flip a fair coin n times and let X be the number of heads. Note that $E(X) = n/2$. Then by Chernoff bounds, we have that:

$$Pr(|X - n/2| \leq \delta n/2) \leq 2e^{-n\delta^2/6}$$

Q: What is the smallest value of δ that still ensures that we have polynomially small probability?
A: To ensure this, need $2e^{-n\delta^2/6} \leq n^{-1}$, which means that $-n\delta^2/6 \leq -\ln n$.
How about $\delta = 1$: we get $-n1/6 \leq -\ln n$ which works
How about $\delta = 1/\sqrt{n}$: we get $-n(1/n)/6 = \Theta(1)$
How about $\delta = \sqrt{(\ln n)/n}$: we get $-n(\ln n)/n/6 = \Theta(-\ln n)$. That works!

3 Bitcoin Consensus

The algorithm below is a simplification of blockchain consensus. A round consists of a time period for mining (say 10 minutes) in which one solution is expected.

1. $C \leftarrow$ some initial chain
2. For $r \leftarrow 0 \ldots \infty$
   (a) Let x be the block I want to add to chain C
   (b) Repeat for this round
      i. Choose a random y
      ii. If $h(C, x, y) \leq D$ then: $C \leftarrow Cx$; Broadcast C along with y. End the round.
3. $C \leftarrow$ longest valid chain received this round

3.1 Analysis

Let G be the set of good processes, and B be the set of bad. For process i, let p_i be the expected number of puzzles solved by i in one round.

Let $\alpha = \sum_{i \in G} p_i$ and $\beta = \sum_{i \in B} p_i$.

Theorem 1. *Assume $\beta \leq 1/2(\alpha - \alpha^2)$. Then, after m rounds of Bitcoin Consensus, with probability at least $1 - O(1/m)$, all but at most the last $55(\alpha - \alpha^2) \log m$ blocks in the longest blockchain are valid.*

Proof: Since every good process adds at most one block per round, the probability that a good process adds a block in a round equals the expected number of blocks added (i.e. p_i). In the following, when we say "add" a block, we mean add a block to some chain, not necessarily the chain that eventually is the longest.

Then, inclusion-exclusion says that the probability that the good processes add at least one block in round r is

$$\sum_{i \in G} p_i - \sum_{i,j \in G; i \neq j} p_i p_j \geq \alpha - \alpha^2$$

(Do you see why $\sum_{i,j \in G; i \neq j} p_i p_j \leq \alpha^2$?)

Let X_r be an indicator r.v. that is 1 if the good processes add a block in round r. Then by linearity, the good processes add an expected $\sum_{r=1}^{m} E(X_r) = m(\alpha - \alpha^2)$ blocks in m rounds. Note that we pessimistically are assuming that if two or more good blocks are mined in the same round, that at most 1 is added.

Let Y_i be an indicator variable for the success of the i-th puzzle attempt by a bad process. Then, by linearity, the expected number of blocks added by bad processes in m rounds is $\sum_i E(Y_i) = m\beta$ (this last summation is over the number of rounds times the number of attempts per round).

Let X be the number of blocks added by the good and Y be the number of blocks added by the bad. By above, we'll assume that $E(X) = m(\alpha - \alpha^2)$ and $E(Y) = m\beta$. Both X and Y are the sum of $0-1$ independent random variables. So by Chernoff bounds, for any δ, $0 < \delta \leq 1$:

$$Pr(X \leq (1-\delta)m(\alpha - \alpha^2) \leq e^{-\delta^2 m(\alpha - \alpha^2)/2}$$
$$Pr(Y \geq (1+\delta)m\beta) \leq e^{-\delta^2 m\beta/3}$$

Let $k = m(\alpha - \alpha^2) = E(X)$. Then $E(Y) = m\beta \leq k/2$. Set $\delta = 1/3$ in the above and we get:

$$Pr(X \leq (2/3)k) \leq e^{-k/18}$$
$$Pr(Y \geq (2/3)k) \leq e^{-k/54}$$

By a union bound, with probability of failure at most $e^{-k/18} + e^{-k/54} \leq e^{-k/55}$, neither of these events happen, and then the number of good blocks is larger than the number of bad blocks. If $k \geq 55 \log m$, we can say that neither of these events happen with probability at most $1/m$.

Good processes never mine on bad (invalid) blocks. So if the number of good blocks is bigger than the number of bad blocks, then (1) the longest chain contains all of the X good blocks; and (2) all bad blocks in this longest chain must occur at the end, since otherwise they would never have been mined on – included in the chain – by a subsequent good miner.

How many invalid blocks can be at the end of this longest blockchain? Again, by the above, it can be at most $k = 55 \log n$ rounds until the number of good blocks catches up with the number of bad. □

4 References

- "Authenticated Algorithms for Byzantine Agreement" by Dolev and Strong. https://www2.imm.dtu.dk/courses/02220/2015/L12/DolevStrong83.pdf
- "Notes on Theory of Distributed System" by James Aspnes. https://www.cs.yale.edu/homes/aspnes/classes/465/notes.pdf

References

[1] Garay, J., Kiayias, A., and Leonardos, N. The bitcoin backbone protocol: Analysis and applications. *Journal of the ACM 71*, 4 (2024), 1–49.