CMPSCI 711: "Really Advanced Algorithms"

Lecture 10 – More Markov Chains and Coupling

Andrew McGregor

Last Compiled: February 26, 2009

Outline

First Attempt: Naive Algorithm

- An algorithm for 3-SAT:
  1. Pick arbitrary assignment
  2. Pick an unsatisfied clause and randomly flip the value assigned to one of the three variables
  3. Repeat Step 2 until there are no unsatisfied clauses.

First Attempt: Naive Algorithm

- An algorithm for 3-SAT:
  1. Pick arbitrary assignment
  2. Pick an unsatisfied clause and randomly flip the value assigned to one of the three variables
  3. Repeat Step 2 until there are no unsatisfied clauses.
- Let $x^{(t)}$ be the assignment at time t.

First Attempt: Naive Algorithm

- An algorithm for 3-SAT:
  1. Pick arbitrary assignment
  2. Pick an unsatisfied clause and randomly flip the value assigned to one of the three variables
  3. Repeat Step 2 until there are no unsatisfied clauses.
- Let $x^{(t)}$ be the assignment at time t.
- Let $X^{(t)} = n - \Delta(x^{(t)}, y)$ for satisfying assignment y:

$$X^{(t+1)} = X^{(t)} \pm 1 \quad \text{and} \quad \mathbb{P}\left[X^{(t+1)} = X^{(t)} + 1\right] \geq 1/3$$

First Attempt: Naive Algorithm

- An algorithm for 3-SAT:
  1. Pick arbitrary assignment
  2. Pick an unsatisfied clause and randomly flip the value assigned to one of the three variables
  3. Repeat Step 2 until there are no unsatisfied clauses.
- Let $x^{(t)}$ be the assignment at time t.
- Let $X^{(t)} = n - \Delta(x^{(t)}, y)$ for satisfying assignment y:

$$X^{(t+1)} = X^{(t)} \pm 1 \quad \text{and} \quad \mathbb{P}\left[X^{(t+1)} = X^{(t)} + 1\right] \geq 1/3$$

- How long until we terminate?

First Attempt: Doesn't work for 3-SAT...

- Let $M = (Y_0, Y_1, \ldots)$ be chain with states $\{0, 1, \ldots, n\}$ and

  $$P_{0,1} = 1,\ \ P_{n,n} = 1,\ \ P_{i,i+1} = \frac{1}{3} \text{ and } P_{i,i-1} = \frac{2}{3} \text{ for } i \in [n-1]$$

First Attempt: Doesn't work for 3-SAT…

- Let $M = (Y_0, Y_1, \ldots)$ be chain with states $\{0, 1, \ldots, n\}$ and

$$P_{0,1} = 1,\ P_{n,n} = 1,\ P_{i,i+1} = \frac{1}{3} \text{ and } P_{i,i-1} = \frac{2}{3} \text{ for } i \in [n-1]$$

- Let h_j be expected time to reach n when starting from j:

$$h_n = 0, \quad h_0 = h_1 + 1, \quad h_j = \frac{2h_{j-1}}{3} + \frac{h_{j+1}}{3} + 1 \text{ for } i \in [n-1]$$

First Attempt: Doesn't work for 3-SAT...

- Let $M = (Y_0, Y_1, \ldots)$ be chain with states $\{0, 1, \ldots, n\}$ and

$$P_{0,1} = 1,\ \ P_{n,n} = 1,\ \ P_{i,i+1} = \frac{1}{3} \text{ and } P_{i,i-1} = \frac{2}{3} \text{ for } i \in [n-1]$$

- Let h_j be expected time to reach n when starting from j:

$$h_n = 0, \quad h_0 = h_1 + 1, \quad h_j = \frac{2h_{j-1}}{3} + \frac{h_{j+1}}{3} + 1 \text{ for } i \in [n-1]$$

- Solving recurrence gives $h_j = 2^{n+2} - 2^{j+2} - 3(n-j)$.

First Attempt: Doesn't work for 3-SAT...

- Let $M = (Y_0, Y_1, \ldots)$ be chain with states $\{0, 1, \ldots, n\}$ and

$$P_{0,1} = 1, \; P_{n,n} = 1, \; P_{i,i+1} = \frac{1}{3} \text{ and } P_{i,i-1} = \frac{2}{3} \text{ for } i \in [n-1]$$

- Let h_j be expected time to reach n when starting from j:

$$h_n = 0, \quad h_0 = h_1 + 1, \quad h_j = \frac{2h_{j-1}}{3} + \frac{h_{j+1}}{3} + 1 \text{ for } i \in [n-1]$$

- Solving recurrence gives $h_j = 2^{n+2} - 2^{j+2} - 3(n-j)$.
- This is no better than trying all possible assignments!

Better Algorithm

Intuition: If the naive algorithm doesn't find a satisfying assignment early, it never will.

Better Algorithm

Intuition: If the naive algorithm doesn't find a satisfying assignment early, it never will.

A better algorithm for 3-SAT:

1. Pick random assignment.
2. Run naive algorithm for $3n$ steps or until formula is satisfied
3. Repeat Steps 1 and 2 until there are no unsatisfied clauses.

Better Algorithm: Analysis

- Let q_j be probability that naive algorithm finds a satisfying assignment in $3j$ steps when the initial assignment disagrees with a satisfying assignment is exactly j positions:

$$q_j \geq \mathbb{P}\left[2j \text{ of the steps increase the state }\right] \geq \binom{3j}{j}\left(\frac{2}{3}\right)^j\left(\frac{1}{3}\right)^{2j}$$

Better Algorithm: Analysis

- Let q_j be probability that naive algorithm finds a satisfying assignment in $3j$ steps when the initial assignment disagrees with a satisfying assignment is exactly j positions:

$$q_j \geq \mathbb{P}\left[2j \text{ of the steps increase the state }\right] \geq \binom{3j}{j}\left(\frac{2}{3}\right)^j\left(\frac{1}{3}\right)^{2j}$$

- Stirling's Formula: $\sqrt{2\pi m}(m/e)^m \leq m! \leq 2\sqrt{2\pi m}(m/e)^m$.

Better Algorithm: Analysis

- Let q_j be probability that naive algorithm finds a satisfying assignment in $3j$ steps when the initial assignment disagrees with a satisfying assignment is exactly j positions:

$$q_j \geq \mathbb{P}\left[2j \text{ of the steps increase the state }\right] \geq \binom{3j}{j}\left(\frac{2}{3}\right)^{j}\left(\frac{1}{3}\right)^{2j}$$

- Stirling's Formula: $\sqrt{2\pi m}(m/e)^m \leq m! \leq 2\sqrt{2\pi m}(m/e)^m$.
- Bound $q_j \geq c2^{-j}/\sqrt{j}$ for $c = \sqrt{3}/\sqrt{64\pi}$ and $q_0 = 1$.

Better Algorithm: Analysis

- Let q_j be probability that naive algorithm finds a satisfying assignment in $3j$ steps when the initial assignment disagrees with a satisfying assignment is exactly j positions:

$$q_j \geq \mathbb{P}\left[2j \text{ of the steps increase the state }\right] \geq \binom{3j}{j}\left(\frac{2}{3}\right)^j\left(\frac{1}{3}\right)^{2j}$$

- Stirling's Formula: $\sqrt{2\pi m}(m/e)^m \leq m! \leq 2\sqrt{2\pi m}(m/e)^m$.
- Bound $q_j \geq c2^{-j}/\sqrt{j}$ for $c = \sqrt{3}/\sqrt{64\pi}$ and $q_0 = 1$.
- Find satisfying assignment in $3n$ steps with probability at least

$$q = \sum_{j=0}^{n} q_j \binom{n}{j} 2^{-n} \leq 0.75^n c/\sqrt{n}$$

Better Algorithm: Analysis

- Let q_j be probability that naive algorithm finds a satisfying assignment in $3j$ steps when the initial assignment disagrees with a satisfying assignment is exactly j positions:

$$q_j \geq \mathbb{P}\left[2j \text{ of the steps increase the state }\right] \geq \binom{3j}{j}\left(\frac{2}{3}\right)^j\left(\frac{1}{3}\right)^{2j}$$

- Stirling's Formula: $\sqrt{2\pi m}(m/e)^m \leq m! \leq 2\sqrt{2\pi m}(m/e)^m$.
- Bound $q_j \geq c2^{-j}/\sqrt{j}$ for $c = \sqrt{3}/\sqrt{64\pi}$ and $q_0 = 1$.
- Find satisfying assignment in $3n$ steps with probability at least

$$q = \sum_{j=0}^{n} q_j \binom{n}{j} 2^{-n} \leq 0.75^n c/\sqrt{n}$$

- Repeating naive algorithm $2/q = O((4/3)^n\sqrt{n})$ times fails with prob. $1/2$

Outline

Gamblers Ruin

- You and a friend repeatedly toss an unbiased coin: If "heads" you give your friend \$1 and if "tails" your friend gives you \$1.

Gamblers Ruin

- You and a friend repeatedly toss an unbiased coin: If "heads" you give your friend \$1 and if "tails" your friend gives you \$1.
- You can't afford to loose more than $\$\ell_1$ and your friend can't afford to loose more than $\$\ell_2$ dollars and when either happens, the game terminates.

Gamblers Ruin

- You and a friend repeatedly toss an unbiased coin: If “heads” you give your friend \$1 and if “tails” your friend gives you \$1.
- You can't afford to loose more than $\$\ell_1$ and your friend can't afford to loose more than $\$\ell_2$ dollars and when either happens, the game terminates.
- What's the probability you win $\$\ell_2$ dollars?

Gamblers Ruin Analysis

- Consider the Markov chain $(X_0, X_1, \ldots)$ where X_t is your profit at time t and $X_0 = 0$. Then

$$X_t \in \{-\ell_1, -\ell_1 + 1, \ldots, 0, \ldots, \ell_2\}$$

$$P_{-\ell_1,-\ell_1} = 1 = P_{\ell_2,\ell_2} \quad P_{i,i+1} = P_{i,i-1} = 1/2 \text{ for } -\ell_1 < i < \ell_2$$

Gamblers Ruin Analysis

- Consider the Markov chain $(X_0, X_1, \ldots)$ where X_t is your profit at time t and $X_0 = 0$. Then

$$X_t \in \{-\ell_1, -\ell_1 + 1, \ldots, 0, \ldots, \ell_2\}$$

$$P_{-\ell_1,-\ell_1} = 1 = P_{\ell_2,\ell_2} \quad P_{i,i+1} = P_{i,i-1} = 1/2 \text{ for } -\ell_1 < i < \ell_2$$

- We want to calculate $q = \lim_{t\to\infty} \mathbb{P}[X_t = \ell_2]$.

Gamblers Ruin Analysis

- Consider the Markov chain $(X_0, X_1, \ldots)$ where X_t is your profit at time t and $X_0 = 0$. Then

$$X_t \in \{-\ell_1, -\ell_1 + 1, \ldots, 0, \ldots, \ell_2\}$$

$$P_{-\ell_1,-\ell_1} = 1 = P_{\ell_2,\ell_2} \quad P_{i,i+1} = P_{i,i-1} = 1/2 \text{ for } -\ell_1 < i < \ell_2$$

- We want to calculate $q = \lim_{t\to\infty} \mathbb{P}[X_t = \ell_2]$.
- Note that $\mathbb{E}[X_t] = 0$ and

$$\lim_{t\to\infty} \mathbb{E}[X_t] = \ell_2 q - \ell_1(1-q)$$

Gamblers Ruin Analysis

- Consider the Markov chain $(X_0, X_1, \ldots)$ where X_t is your profit at time t and $X_0 = 0$. Then

$$X_t \in \{-\ell_1, -\ell_1 + 1, \ldots, 0, \ldots, \ell_2\}$$

$$P_{-\ell_1,-\ell_1} = 1 = P_{\ell_2,\ell_2} \quad P_{i,i+1} = P_{i,i-1} = 1/2 \text{ for } -\ell_1 < i < \ell_2$$

- We want to calculate $q = \lim_{t\to\infty} \mathbb{P}[X_t = \ell_2]$.
- Note that $\mathbb{E}[X_t] = 0$ and

$$\lim_{t\to\infty} \mathbb{E}[X_t] = \ell_2 q - \ell_1(1 - q)$$

- Hence $q = \ell_1/(\ell_1 + \ell_2)$.

Outline

Motivating Example: Shuffling

- You're given a pack of m cards.
- Draw a random card, place it at top, and repeat.
- How many repeats until they're "reasonably shuffled"?

Motivating Example: Shuffling

- You're given a pack of m cards.
- Draw a random card, place it at top, and repeat.
- How many repeats until they're "reasonably shuffled"?

Consider the process as a Markov Chain with

Motivating Example: Shuffling

- You're given a pack of m cards.
- Draw a random card, place it at top, and repeat.
- How many repeats until they're "reasonably shuffled"?

Consider the process as a Markov Chain with

- States: All $n = m!$ arrangements of the cards.

Motivating Example: Shuffling

- You're given a pack of m cards.
- Draw a random card, place it at top, and repeat.
- How many repeats until they're "reasonably shuffled"?

Consider the process as a Markov Chain with

- States: All $n = m!$ arrangements of the cards.
- Transition Probability: $P_{ij} = 1/m$ if arrangement j can be formed from arrangement i by moving a card to the top.

Motivating Example: Shuffling

- You're given a pack of m cards.
- Draw a random card, place it at top, and repeat.
- How many repeats until they're "reasonably shuffled"?

Consider the process as a Markov Chain with

- States: All $n = m!$ arrangements of the cards.
- Transition Probability: $P_{ij} = 1/m$ if arrangement j can be formed from arrangement i by moving a card to the top.
- Claim: The stationary distribution is $\pi_i = 1/n$ for all $i \in [n]$.

Mixing Time

Definition

Given two distributions over $[n]$, p and q, the variational distance is

$$|p - q| = \frac{1}{2} \sum_{i \in [n]} |p_i - q_i| = \max_{A \subset [n]} |p(A) - q(A)|$$

where $p(A) = \sum_{i \in A} p_i$ and $q(A) = \sum_{i \in A} q_i$.

Mixing Time

Definition

Given two distributions over $[n]$, p and q, the variational distance is

$$|p-q| = \frac{1}{2}\sum_{i\in[n]}|p_i - q_i| = \max_{A\subset[n]}|p(A)-q(A)|$$

where $p(A) = \sum_{i\in A} p_i$ and $q(A) = \sum_{i\in A} q_i$.

Definition

Let M be a Markov chain and $q^{i,(t)}$ be the distribution over states at time t if $X_0 = i$. Define $\tau_i(\epsilon) = \min\{t : |q^{i,(t)} - \pi| \leq \epsilon\}$. The mixing time of a chain is $\tau(\epsilon) = \max_{i\in[n]}\tau_i(\epsilon)$.

Mixing Time

Definition

Given two distributions over $[n]$, p and q, the variational distance is

$$|p - q| = \frac{1}{2} \sum_{i \in [n]} |p_i - q_i| = \max_{A \subset [n]} |p(A) - q(A)|$$

where $p(A) = \sum_{i \in A} p_i$ and $q(A) = \sum_{i \in A} q_i$.

Definition

Let M be a Markov chain and $q^{i,(t)}$ be the distribution over states at time t if $X_0 = i$. Define $\tau_i(\epsilon) = \min\{t : |q^{i,(t)} - \pi| \leq \epsilon\}$. The mixing time of a chain is $\tau(\epsilon) = \max_{i \in [n]} \tau_i(\epsilon)$.

Theorem

For the shuffling of a pack of m cards $\tau(\epsilon) \leq m \ln(m/\epsilon)$.

Coupling and Coupling Lemma

Definition

Consider a Markov chain M with transition matrix P. We say that $Z_t = (A_t, B_t)$ is a coupling for M if A_t and B_t belong to the same state space as M and

$$\mathbb{P}\left[A_{t+1} = i' | Z_t = (i,j)\right] = P_{ii'} \text{ and } \mathbb{P}\left[B_{t+1} = j' | Z_t = (i,j)\right] = P_{jj'}$$

Coupling and Coupling Lemma

Definition

Consider a Markov chain M with transition matrix P. We say that $Z_t = (A_t, B_t)$ is a coupling for M if A_t and B_t belong to the same state space as M and

$$\mathbb{P}\left[A_{t+1} = i' | Z_t = (i,j)\right] = P_{ii'} \text{ and } \mathbb{P}\left[B_{t+1} = j' | Z_t = (i,j)\right] = P_{jj'}$$

Lemma

Let $Z_t = (A_t, B_t)$ be a coupling for a Markov chain M on a state space $[n]$. Suppose there exists a t^* such that, for every $i \in [n]$,

$$\mathbb{P}\left[A_{t^*} \neq B_{t^*} | A_0 = i, B_0 \sim \pi\right] \leq \epsilon$$

where π is the stationary distribution of M. Then $\tau(\epsilon) \leq t^*$.

Proof of Coupling Lemma

Proof.

- Let $S \subset [n]$ and denote $\pi(S) = \sum_{i \in S} \pi_i$

$$
\begin{aligned}
\mathbb{P}[A_{t^*} \in S] &\geq \mathbb{P}[A_{t^*} = B_{t^*} \text{ and } B_{t^*} \in S] \\
&= 1 - \mathbb{P}[A_{t^*} \neq B_{t^*} \text{ or } B_{t^*} \notin S] \\
&\geq 1 - \mathbb{P}[A_{t^*} \neq B_{t^*}] - \mathbb{P}[B_{t^*} \notin S] \\
&\geq \mathbb{P}[B_{t^*} \in S] - \epsilon \\
&= \pi(S) - \epsilon
\end{aligned}
$$

□

Proof of Coupling Lemma

Proof.

- Let $S \subset [n]$ and denote $\pi(S) = \sum_{i \in S} \pi_i$

$$
\begin{aligned}
\mathbb{P}[A_{t^*} \in S] &\geq \mathbb{P}[A_{t^*} = B_{t^*} \text{ and } B_{t^*} \in S] \\
&= 1 - \mathbb{P}[A_{t^*} \neq B_{t^*} \text{ or } B_{t^*} \notin S] \\
&\geq 1 - \mathbb{P}[A_{t^*} \neq B_{t^*}] - \mathbb{P}[B_{t^*} \notin S] \\
&\geq \mathbb{P}[B_{t^*} \in S] - \epsilon \\
&= \pi(S) - \epsilon
\end{aligned}
$$

- Similarly $\mathbb{P}[A_{t^*} \in [n] \setminus S] \geq \pi([n] \setminus S) - \epsilon$ and hence,

$$\mathbb{P}[A_{t^*} \in S] \leq \pi(S) + \epsilon$$

□

Proof of Coupling Lemma

Proof.

- Let $S \subset [n]$ and denote $\pi(S) = \sum_{i \in S} \pi_i$

$$
\begin{aligned}
\mathbb{P}[A_{t^*} \in S] &\geq \mathbb{P}[A_{t^*} = B_{t^*} \text{ and } B_{t^*} \in S] \\
&= 1 - \mathbb{P}[A_{t^*} \neq B_{t^*} \text{ or } B_{t^*} \notin S] \\
&\geq 1 - \mathbb{P}[A_{t^*} \neq B_{t^*}] - \mathbb{P}[B_{t^*} \notin S] \\
&\geq \mathbb{P}[B_{t^*} \in S] - \epsilon \\
&= \pi(S) - \epsilon
\end{aligned}
$$

- Similarly $\mathbb{P}[A_{t^*} \in [n] \setminus S] \geq \pi([n] \setminus S) - \epsilon$ and hence,

$$\mathbb{P}[A_{t^*} \in S] \leq \pi(S) + \epsilon$$

- Hence, $|\mathbb{P}[A_{t^*} \in S] - \pi(S)| \leq \epsilon$ and therefore $|\pi - q^{(t^*)}| \leq \epsilon$

□

Return to Shuffling (1/2)

Theorem

For the shuffling of a pack of m cards $\tau(\epsilon) \leq m \ln(m/\epsilon)$.

Return to Shuffling (1/2)

Theorem

For the shuffling of a pack of m cards $\tau(\epsilon) \leq m \ln(m/\epsilon)$.

- Let A and B be "copies" of the shuffling Markov chain M:
  - For A, pick a card at random from deck and place it on top.
  - For B, pick card "X" and place it on the top where "X" was the value of the card seen in A.

Return to Shuffling (1/2)

Theorem

For the shuffling of a pack of m cards $\tau(\epsilon) \leq m \ln(m/\epsilon)$.

- Let A and B be "copies" of the shuffling Markov chain M:
  - For A, pick a card at random from deck and place it on top.
  - For B, pick card "X" and place it on the top where "X" was the value of the card seen in A.
- Note that (A, B) are a coupling of M.

Return to Shuffling (2/2)

- The arrangement of cards in both chains is identical if we've picked each card at least once

Return to Shuffling (2/2)

- The arrangement of cards in both chains is identical if we've picked each card at least once
- After $t^* = m \ln m + cm$ steps, probability a specific card hasn't been picked is

$$(1 - 1/m)^{t^*} \leq \exp(-\ln m + c) \leq e^{-c}/m$$

Return to Shuffling (2/2)

- The arrangement of cards in both chains is identical if we've picked each card at least once
- After $t^* = m \ln m + cm$ steps, probability a specific card hasn't been picked is

$$(1 - 1/m)^{t^*} \leq \exp(-\ln m + c) \leq e^{-c}/m$$

- Hence, after t^* steps both chains have identical arrangements with probability at least $1 - e^{-c}$.

Return to Shuffling (2/2)

- The arrangement of cards in both chains is identical if we've picked each card at least once
- After $t^* = m \ln m + cm$ steps, probability a specific card hasn't been picked is

$$(1 - 1/m)^{t^*} \leq \exp(-\ln m + c) \leq e^{-c}/m$$

- Hence, after t^* steps both chains have identical arrangements with probability at least $1 - e^{-c}$.
- By coupling lemma, $\tau(\epsilon) \leq m \ln(m/\epsilon)$.