COMPSCI 614: Randomized Algorithms with Applications to Data Science

Prof. Cameron Musco

University of Massachusetts Amherst. Spring 2024.

Lecture 18

Logistics

- Problem Set 4 is due Monday 4/22 at 11:59pm.
- No quiz this week.
- Office hours on Zoom after class.

Summary

Last Class: $A^TS^TSA \approx_\epsilon A^TA$

- Leverage score intuition. "uniqueness" of rows, cluster structure
- Connection to spectral graph sparsification — L Laplacian, $B^TB = L$
- Connection to effective resistances in electrical networks. **Note:** I am not going to finish this full derivation – see Lecture 17 slides if you are interested. Spielman spectral graph theory.

Summary

Last Class:

- Leverage score intuition.
- Connection to spectral graph sparsification
- Connection to effective resistances in electrical networks. **Note:** I am not going to finish this full derivation – see Lecture 17 slides if you are interested.

Today:

- New unit: Markov Chains.
- Markov chain based algorithms for 2-SAT and 3-SAT.

(NP-hard)

Markov Chain Definition

- A discrete time stochastic process is a collection of random variables $X_0, X_1, X_2, \ldots,$
- A discrete time stochastic process is a Markov chain if is it memoryless:

$$\Pr(X_t = a_t | X_{t-1} = a_{t-1}, \ldots, X_0 = a_0) = \Pr(X_t = a_t | X_{t-1} = a_{t-1})$$
$$= P_{a_{t-1}, a_t}.$$

Question: In a Markov chain, is X_t independent of $X_{t-2}, X_{t-3}, \ldots, X_0$?

(only when condition on X_{t-1}

Flipping coins, X_t = # heads up to time t

$X_t = X_{t-1} + 1$ w.p. $1/2$

$= X_{t-1}$ w.p. $1/2$

$\boxed{X_t \geq X_{t-j} \text{ for any } j > 0}$

Transition Matrix

A Markov chain $X_0, X_1, \ldots$ where each X_i can take $\boxed{m \text{ possible values}}$, is specified by the transition matrix $P \in [0,1]^{m \times m}$ with

$$\underline{P_{j,k}} = \Pr(\underline{X_{i+1} = k} | \underline{X_i = j}).$$

P

.5	.5	0	0
.25	.25	.5	0
0	1	0	0
0	.5	.5	0

rows sum to 1.

Transition Matrix

A Markov chain $X_0, X_1, \ldots$ where each X_i can take m possible values, is specified by the transition matrix $P \in [0,1]^{m \times m}$ with

$$P_{j,k} = \Pr(X_{i+1} = k | X_i = j).$$

Let $q_i \in [0,1]^{1 \times m}$ be the distribution of X_i. Then $q_{i+1} = q_i P$.

$[0 \;\; .1 \;\; .2 \;\; 0 \;\; .7]$

P

.5	.5	0	0
.25	.25	.5	0
0	1	0	0
0	.5	.5	0

Transition Matrix

A Markov chain $X_0, X_1, \ldots$ where each X_i can take m possible values, is specified by the transition matrix $P \in [0,1]^{m \times m}$ with

$$P_{j,k} = \Pr(X_{i+1} = k | X_i = j).$$

Let $q_i \in [0,1]^{1 \times m}$ be the distribution of X_i. Then $q_{i+1} = q_i P$.

q_0

1	0	0	0

P

.5	.5	0	0
.25	.25	.5	0
0	1	0	0
0	.5	.5	0

=

q_1

.5	.5	0	0

Transition Matrix

A Markov chain $X_0, X_1, \ldots$ where each X_i can take m possible values, is specified by the transition matrix $P \in [0,1]^{m \times m}$ with

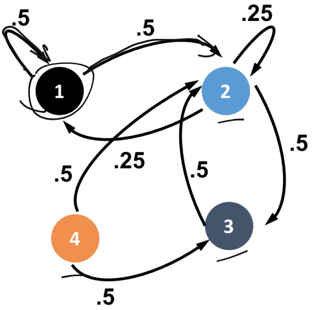
$$P_{j,k} = \Pr(X_{i+1} = k | X_i = j).$$

Let $q_i \in [0,1]^{1 \times m}$ be the distribution of X_i. Then $q_{i+1} = q_i P$.

$$\overset{\mathbf{q_1}}{\begin{bmatrix} .5 & .5 & 0 & 0 \end{bmatrix}} \quad \begin{array}{r} .5 \\ + \\ .5 \\ \\ \\ \end{array} \overset{\mathbf{P}}{\begin{bmatrix} .5 & .5 & 0 & 0 \\ .25 & .25 & .5 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & .5 & .5 & 0 \end{bmatrix}} = \overset{\mathbf{q_2}}{\begin{bmatrix} .375 & .375 & .25 & 0 \end{bmatrix}}$$

Graph View

Often viewed as an underlying state transition graph. Nodes correspond to possible values that each X_i can take.

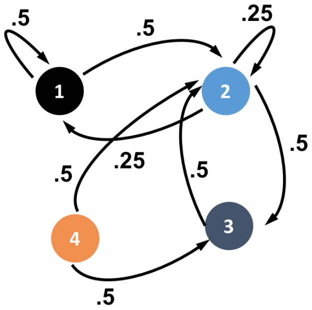


P

.5	.5	0	0
.25	.25	.5	0
0	1	0	0
0	.5	.5	0

Graph View

Often viewed as an underlying state transition graph. Nodes correspond to possible values that each X_i can take.



P

.5	.5	0	0
.25	.25	.5	0
0	1	0	0
0	.5	.5	0

The Markov chain is irreducible if the underlying graph consists of single strongly connected component. – I can get from state to any other state.

Motivating Example: Find a satisfying assignment for a 2-CNF formula with n variables.

conjunctive normal form

$(x_1 \vee \bar{x}_2 \vee x_4)$

$$(x_1 \vee \bar{x}_2) \wedge (\bar{x}_1 \vee \bar{x}_3) \wedge (x_1 \vee x_2) \wedge (x_4 \vee \bar{x}_3) \wedge (x_4 \vee \bar{x}_1)$$

solvable in polynomial time

2-SAT

Motivating Example: Find a satisfying assignment for a 2-CNF formula with n variables.

$$(x_1 \vee \bar{x}_2) \wedge (\bar{x}_1 \vee \bar{x}_3) \wedge (x_1 \vee x_2) \wedge (x_4 \vee \bar{x}_3) \wedge (x_4 \vee \bar{x}_1)$$

A simple 'local search' algorithm:

1. Start with an arbitrary assignment.
2. Repeat $2mn^2$ times, terminating if a satisfying assignment is found:
    - Chose an arbitrary unsatisfied clause.
    - Pick one of the variables in the clause uniformly at random, and switch the assignment of the variable.
3. If a valid assignment is not found, return that the formula is unsatisfiable.

2-SAT

Motivating Example: Find a satisfying assignment for a 2-CNF formula with n variables.

$$(x_1 \vee \bar{x}_2) \wedge (\bar{x}_1 \vee \bar{x}_3) \wedge (x_1 \vee x_2) \wedge (x_4 \vee \bar{x}_3) \wedge (x_4 \vee \bar{x}_1)$$

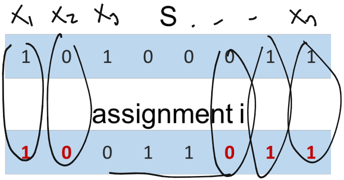
A simple 'local search' algorithm:

1. Start with an arbitrary assignment.
2. Repeat $2mn^2$ times, terminating if a satisfying assignment is found:
    - Chose an arbitrary unsatisfied clause.
    - Pick one of the variables in the clause uniformly at random, and switch the assignment of the variable.
3. If a valid assignment is not found, return that the formula is unsatisfiable.

Claim: If the formula is satisfiable, the algorithm finds a satisfying assignment with probability $\geq 1 - 2^{-m}$.

Randomized 2-SAT Analysis

Fix a satisfying assignment S. Let $X_i \leq n$ be the number of variables that are assigned the same values as in S, at step i.



$X_i = 5$

- $X_{i+1} = X_i \pm 1$ since we flip one variable in an unsatisfied clause.

Randomized 2-SAT Analysis

Fix a satisfying assignment S. Let $X_i \leq n$ be the number of variables that are assigned the same values as in S, at step i.

X_i is increasing over time.

if $X_i = n$ we are done.

S

1	0	1	0	0	0	1	1

assignment i

1	**0**	0	1	1	**0**	**1**	**1**

- $X_{i+1} = X_i \pm 1$ since we flip one variable in an unsatisfied clause.
- $\Pr(X_{i+1} = X_i + 1) \geq$ 1/2
- $\Pr(X_{i+1} = X_i - 1) \leq$ 1/2

we flip variable in unsatisfied clause

$x_1 = 0$ $x_2 = 0$

$$(x_1 \vee \bar{x}_2) \wedge (\bar{x}_1 \vee \bar{x}_3) \wedge \boxed{(x_1 \vee x_2)} \wedge (x_4 \vee \bar{x}_3) \wedge (x_4 \vee \bar{x}_1)$$

S: $x_1 = 1$ $x_2 = 0$ $x_2 = 1$

Coupling to a Markov Chain

The number of correctly assigned variables at step i, X_i, obeys

$$\Pr(X_{i+1} = X_i + 1) \geq \frac{1}{2} \quad \text{and} \quad \Pr(X_{i+1} = X_i - 1) \leq \frac{1}{2}.$$

Coupling to a Markov Chain

The number of correctly assigned variables at step i, X_i, obeys

$$\Pr(X_{i+1} = X_i + 1) \underset{?}{\overset{=}{\geq}} \frac{1}{2} \quad \text{and} \quad \Pr(X_{i+1} = X_i - 1) \underset{?}{\overset{=}{\leq}} \frac{1}{2}.$$

Is $X_0, X_1, X_2, \ldots$ a Markov chain? — NO!

(would be a Markov chain if the above equations were equalities

Coupling to a Markov Chain

The number of correctly assigned variables at step i, X_i, obeys $X_i = 2$

$$\Pr(X_{i+1} = X_i + 1) \geq \frac{1}{2} \quad \text{and} \quad \Pr(X_{i+1} = X_i - 1) \leq \frac{1}{2}.$$

$x_1 = 1$, $x_2 = 0$, $x_3 = 1$, $x_4 = 0$

$$(x_1 \vee x_2) \wedge (x_3 \vee x_4)$$

S: 1 1 1 1

Is $X_0, X_1, X_2, \ldots$ a Markov chain?

Define a Markov chain $Y_0, Y_1, \ldots$ such that $Y_0 = X_0$ and:

$$\Pr(Y_{i+1} = 1 | Y_i = 0) = 1$$
$$\Pr(Y_{i+1} = j+1 | Y_i = j) = 1/2 \text{ for } 1 \leq j \leq n-1$$
$$\Pr(Y_{i+1} = j-1 | Y_i = j) = 1/2 \text{ for } 1 \leq j \leq n-1$$
$$\Pr(Y_{i+1} = n | Y_i = n) = 1.$$

Coupling to a Markov Chain

The number of correctly assigned variables at step i, X_i, obeys

$$\Pr(X_{i+1} = X_i + 1) \geq \frac{1}{2} \quad \text{and} \quad \Pr(X_{i+1} = X_i - 1) \leq \frac{1}{2}.$$

Is $X_0, X_1, X_2, \ldots$ a Markov chain?

$X_0 \cdot X_1 \cdot\cdot X_2$

Define a Markov chain $Y_0, Y_1, \ldots$ such that $Y_0 = X_0$ and:

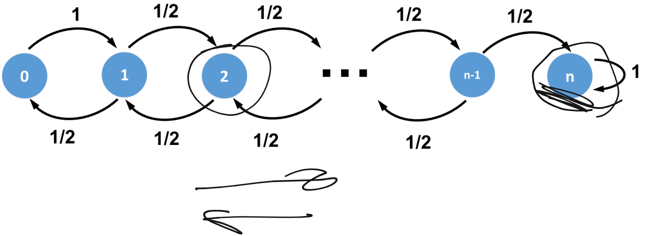
$Y_0 \quad Y_1 \cdots$

$$\begin{aligned}
\Pr(Y_{i+1} = 1 | Y_i = 0) &= 1 \\
\Pr(Y_{i+1} = j + 1 | Y_i = j) &= 1/2 \text{ for } 1 \leq j \leq n - 1 \\
\Pr(Y_{i+1} = j - 1 | Y_i = j) &= 1/2 \text{ for } 1 \leq j \leq n - 1 \\
\Pr(Y_{i+1} = n | Y_i = n) &= 1.
\end{aligned}$$

- Our algorithm terminates as soon as $X_i = n$. We expect to reach this point only more slowly with Y_i. So it suffices to argue that $Y_i = n$ with high probability for large enough i.
- Formally could use a coupling argument (will see later on).

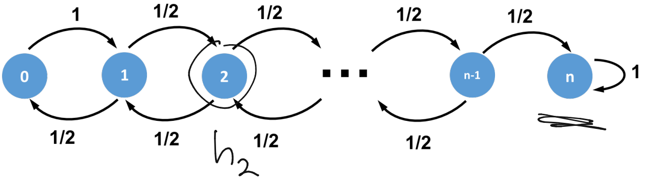
Simple Markov Chain Analysis

Want to bound the expected time required to have $\mathbf{Y}_i = n$.



Simple Markov Chain Analysis

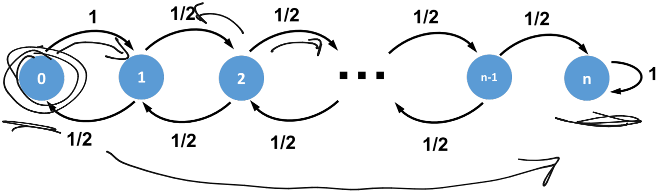
Want to bound the expected time required to have $\mathbf{Y}_i = n$.



Let h_j be the expected number of steps to reach n when starting at node j (i.e., the expected termination time when j variables are assigned correctly.)

Simple Markov Chain Analysis

Want to bound the expected time required to have $\mathbf{Y}_i = n$.



Let h_j be the expected number of steps to reach n when starting at node j (i.e., the expected termination time when j variables are assigned correctly.)

$$\begin{cases} h_n = 0 \\ h_0 = h_1 + 1 \\ h_j = \frac{h_{j-1}}{2} + \frac{h_{j+1}}{2} + 1 \text{ for } 1 \leq j \leq n-1 \end{cases}$$

Simple Markov Chain Analysis

Claim: $h_j = h_{j+1} + 2j + 1$.

Simple Markov Chain Analysis

Claim: $h_j = h_{j+1} + 2j + 1$. Can prove via induction on j.

Simple Markov Chain Analysis

Claim: $h_j = h_{j+1} + 2j + 1$. Can prove via induction on j.

- $h_0 = h_1 + 1$, satisfying the claim in the base case.

$h_{0+1} + 2 \cdot 0 + 1$

Simple Markov Chain Analysis

Claim: $h_j = h_{j+1} + 2j + 1$. Can prove via induction on j.

- $h_0 = h_1 + 1$, satisfying the claim in the base case.

$h_j + 2(j-1) + 1$ via inductive hypothesis

$$h_j = \frac{h_{j-1}}{2} + \frac{h_{j+1}}{2} + 1$$

$$h_j = \frac{h_j}{2} + (j-1) + \frac{1}{2} + \frac{h_{j+1}}{2} + 1$$

$$h_j = \frac{h_j}{2} + \frac{h_{j+1}}{2} + j + \frac{1}{2}.$$

$$\frac{h_j}{2} = \frac{h_{j+1}}{2} + j + \frac{1}{2}$$

$$\boxed{h_j = h_{j+1} + 2j + 1}$$

Claim: $h_j = h_{j+1} + 2j + 1$. Can prove via induction on j.

- $h_0 = h_1 + 1$, satisfying the claim in the base case.

$$h_j = \frac{h_{j-1}}{2} + \frac{h_{j+1}}{2} + 1$$
$$h_j = \frac{h_j}{2} + (j-1) + \frac{1}{2} + \frac{h_{j+1}}{2} + 1$$
$$h_j = \frac{h_j}{2} + \frac{h_{j+1}}{2} + j + \frac{1}{2}.$$

- Rearranging gives: $\underline{h_j = h_{j+1} + 2j + 1}$.

Claim: $h_j = h_{j+1} + 2j + 1$. Can prove via induction on j.

- $h_0 = h_1 + 1$, satisfying the claim in the base case.

$$h_j = \frac{h_{j-1}}{2} + \frac{h_{j+1}}{2} + 1$$
$$h_j = \frac{h_j}{2} + (j-1) + \frac{1}{2} + \frac{h_{j+1}}{2} + 1$$
$$h_j = \frac{h_j}{2} + \frac{h_{j+1}}{2} + j + \frac{1}{2}.$$

- Rearranging gives: $h_j = h_{j+1} + 2j + 1$.

So in total we have:

$$h_0 = h_1 + 1 = h_2 + 3 + 1 = \ldots = \sum_{j=0}^{n-1}(2j+1) \simeq$$

$h_1 = h_2 + 3$

Claim: $h_j = h_{j+1} + 2j + 1$. Can prove via induction on j.

- $h_0 = h_1 + 1$, satisfying the claim in the base case.

$$h_j = \frac{h_{j-1}}{2} + \frac{h_{j+1}}{2} + 1$$
$$h_j = \frac{h_j}{2} + (j-1) + \frac{1}{2} + \frac{h_{j+1}}{2} + 1$$
$$h_j = \frac{h_j}{2} + \frac{h_{j+1}}{2} + j + \frac{1}{2}.$$

- Rearranging gives: $h_j = h_{j+1} + 2j + 1$.

So in total we have:

$$h_0 = h_1 + 1 = h_2 + 3 + 1 = \ldots = \sum_{j=0}^{n-1}(2j+1) = n^2.$$

$\llcorner$ expected time to find satisfying assignment $\leq n^2$

Simple Markov Chain Analysis

Upshot: Consider the Markov chain $Y_0, Y_1, \ldots$, and let i^* be the minimum i such $Y_{i^*} = n$. Then $\mathbb{E}[i^*] \leq n^2$.

Simple Markov Chain Analysis

Upshot: Consider the Markov chain $Y_0, Y_1, \ldots$, and let i^* be the minimum i such $Y_{i^*} = n$. Then $\mathbb{E}[i^*] \le n^2$.

- Thus, by Markov's inequality, with probability $\ge 1/2$, our 2-SAT algorithms finds a satisfying assignment within $2n^2$ steps.

$Pr(i^* \geq 2n^2) \leq 1/2$

$O(n^2)$ steps fairly intuitive

↳ n^2 step walk on infinite line, s.d. of my position is exactly n

$O(n)$ $O(n)$

Upshot: Consider the Markov chain $Y_0, Y_1, \ldots$, and let i^* be the minimum i such $Y_{i^*} = n$. Then $\mathbb{E}[i^*] \leq n^2$.

- Thus, by Markov's inequality, with probability $\geq 1/2$, our 2-SAT algorithms finds a satisfying assignment within $2n^2$ steps.
- Splitting our $2mn^2$ total steps into m periods of $2n^2$ steps each, we fail to find a satisfying assignment in all m periods with probability at most $1/2^m$.

$2n^2$ $2n^2$ . . . $2n^2$

m periods

Prob at least $1 - 1/2^m$ I find sat assignmt.

More Challenging Problem: Find a satisfying assignment for a 3-CNF formula with n variables.

$$(x_1 \vee \bar{x}_2 \vee \bar{x}_3) \wedge (\bar{x}_1 \vee \bar{x}_3 \vee x_4) \wedge (x_1 \vee x_2 \vee \bar{x}_3).$$

More Challenging Problem: Find a satisfying assignment for a 3-CNF formula with n variables.

$$(x_1 \vee \bar{x}_2 \vee \bar{x}_3) \wedge (\bar{x}_1 \vee \bar{x}_3 \vee x_4) \wedge (x_1 \vee x_2 \vee \bar{x}_3).$$

- 3-SAT is famously NP-hard. What is the naive deterministic runtime required to solve 3-SAT?

$\boxed{2^n}$ time to check all possible assignments

$2^n \cdot \text{poly}(n)$

3-SAT

More Challenging Problem: Find a satisfying assignment for a 3-CNF formula with n variables.

$$(x_1 \vee \bar{x}_2 \vee \bar{x}_3) \wedge (\bar{x}_1 \vee \bar{x}_3 \vee x_4) \wedge (x_1 \vee x_2 \vee \bar{x}_3).$$

- 3-SAT is famously NP-hard. What is the naive deterministic runtime required to solve 3-SAT?
- The current best known runtime is $O(1.307^n)$ [Hansen, Kaplan, Zamir, Zwick, 2019].
- Will see that our simple Markov chain approach gives an $O(1.3334^n)$ time algorithm.

$\left(\frac{4}{3}\right)^n \cdot \text{poly}(n) = O(1.3334^n)$

$2^{10} \simeq 1000$ $\quad 2^{40} \simeq 1$ trillion.

$\left(\frac{4}{3}\right)^{40} = 100K$

More Challenging Problem: Find a satisfying assignment for a 3-CNF formula with n variables.

$$(x_1 \vee \bar{x}_2 \vee \bar{x}_3) \wedge (\bar{x}_1 \vee \bar{x}_3 \vee x_4) \wedge (x_1 \vee x_2 \vee \bar{x}_3).$$

$(x_1 \vee x_2 \vee \bar{x}_3 \vee x_4)$

- 3-SAT is famously NP-hard. What is the naive deterministic runtime required to solve 3-SAT?
- The current best known runtime is $O(1.307^n)$ [Hansen, Kaplan, Zamir, Zwick, 2019].
- Will see that our simple Markov chain approach gives an $O(1.3334^n)$ time algorithm.

  stronger hypotheses than $P \neq NP$
- Note that the exponential time hypothesis conjectures that $O(c^n)$ is needed to solve 3-SAT for some constant $c > 1$. The strong exponential time hypothesis conjectures that for $k \to \infty$, solving k-SAT requires $O(2^n)$ time.

  $\left(1+\frac{1}{\log n}\right)^n$

Randomized 3-SAT Algorithm

1. Start with an arbitrary assignment.
2. Repeat *m* times, terminating if a satisfying assignment is found:
    - Chose an arbitrary unsatisfied clause.
    - Pick one of the variables in the clause uniformly at random, and switch the assignment of the variable.
3. If a valid assignment is not found, return that the formula is unsatisfiable.

As in the 2-SAT setting, let X_i be the number of correctly assigned variables at step i. We have:

$$\Pr(X_i = \underline{X_{i-1}} + 1) \geq 1/3$$
$$\Pr(X_i = X_{i-1} - 1) \leq 2/3$$

correct: 0 0 0

$(x_1 \vee x_2 \vee x_3)$

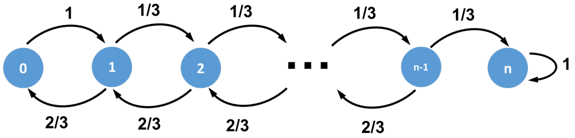
S: 1 0 0

Randomized 3-SAT Analysis

As in the 2-SAT setting, let X_i be the number of correctly assigned variables at step i. We have:

$$\Pr(X_i = X_{i-1} + 1) \geq$$
$$\Pr(X_i = X_{i-1} - 1) \leq$$

Define the coupled Markov chain $Y_0, Y_1, \ldots$ as before, but with $Y_i = Y_{i-1} + 1$ with probability $1/3$ and $Y_i = Y_{i-1} - 1 = 2/3$.



Randomized 3-SAT Analysis

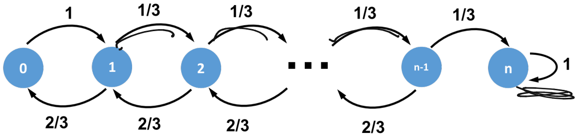
As in the 2-SAT setting, let X_i be the number of correctly assigned variables at step i. We have:

$$\Pr(X_i = X_{i-1} + 1) \geq$$
$$\Pr(X_i = X_{i-1} - 1) \leq$$

Define the coupled Markov chain $Y_0, Y_1, \ldots$ as before, but with $Y_i = Y_{i-1} + 1$ with probability $1/3$ and $Y_i = Y_{i-1} - 1 = 2/3$.

$1/2, 1/2$

termination time was $O(n^2)$



How many steps do you expect are needed to reach $Y_i = n$?

$\mathrm{poly}(n)? \Rightarrow P = NP$ $\qquad$ $\exp(n)$

Randomized 3-SAT Analysis

Letting h_j be the expected number of steps to reach n when starting at node j,

$$h_n = 0$$

$$h_0 = h_1 + 1$$

$$h_j = \frac{2h_{j-1}}{3} + \frac{h_{j+1}}{3} + 1 \text{ for } 1 \leq j \leq n - 1$$

Randomized 3-SAT Analysis

Letting h_j be the expected number of steps to reach n when starting at node j,

$$h_n = 0$$

$$h_0 = h_1 + 1$$

$$h_j = \frac{2h_{j-1}}{3} + \frac{h_{j+1}}{3} + 1 \text{ for } 1 \leq j \leq n-1$$

2(j-1)

- We can prove via induction that $h_j = \underline{h_{j+1}} + \underline{2^{j+2}} - 3$ and in turn, $h_0 \underline{= 2^{n+2} - 4 - 3n}$.

Letting h_j be the expected number of steps to reach n when starting at node j,

$$h_n = 0$$
$$h_0 = h_1 + 1$$
$$h_j = \frac{2h_{j-1}}{3} + \frac{h_{j+1}}{3} + 1 \text{ for } 1 \leq j \leq n-1$$

- We can prove via induction that $h_j = h_{j+1} + 2^{j+2} - 3$ and in turn, $h_0 = 2^{n+2} - 4 - 3n$.
- Thus, in expectation, our algorithm takes at most $\approx 2^{n+2}$ steps to find a satisfying assignment if there is one.

Randomized 3-SAT Analysis

Letting h_j be the expected number of steps to reach n when starting at node j,

$$h_n = 0$$
$$h_0 = h_1 + 1$$
$$h_j = \frac{2h_{j-1}}{3} + \frac{h_{j+1}}{3} + 1 \text{ for } 1 \leq j \leq n-1$$

- We can prove via induction that $h_j = h_{j+1} + 2^{j+2} - 3$ and in turn, $h_0 = 2^{n+2} - 4 - 3n$.
- Thus, in expectation, our algorithm takes at most $\approx 2^{n+2}$ steps to find a satisfying assignment if there is one.
- Is this an interesting result? NO. we already have the 2^n algo.

Modified 3-SAT Algorithm

Key Idea: If we pick our initial assignment uniformly at random, we will have $\mathbb{E}[X_0] = n/2$. With very small, but still non-negligible probability, X_0 will be much larger, and our random walk will be more likely to find a satisfying assignment.

Modified 3-SAT Algorithm

Key Idea: If we pick our initial assignment uniformly at random, we will have $\mathbb{E}[X_0] = n/2$. With very small, but still non-negligible probability, X_0 will be much larger, and our random walk will be more likely to find a satisfying assignment.

Modified Randomized 3-SAT Algorithm:

Repeat m times, terminating if a satisfying assignment is found:

1. Pick a uniform random assignment for the variables.
2. Repeat $3n$ times, terminating if a satisfying assignment is found:
    - Chose an arbitrary unsatisfied clause.
    - Pick one of the variables in the clause uniformly at random, and switch the assignment of the variable.

If a valid assignment is not found, return that the formula is unsatisfiable.

Consider a single random assignment with $X_0 = n - j$. I.e., we need to correct j variables to find a satisfying assignment.

$E(j) = \frac{n}{2}$

Modified 3-SAT Analysis

Consider a single random assignment with $X_0 = n - j$. I.e., we need to correct j variables to find a satisfying assignment.

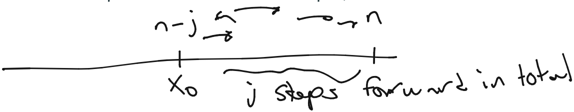
Let q_j be a lower bound on the success probability in this case. Since $j \leq n$ and since we run the search process for $3n$ steps,

$$\begin{aligned} q_j &= \Pr[X_{3n} = n] \\ &\geq \Pr[X_{3j} = n] \end{aligned}$$

Modified 3-SAT Analysis

Consider a single random assignment with $X_0 = n - j$. I.e., we need to correct j variables to find a satisfying assignment.

Let q_j be a lower bound on the success probability in this case. Since $j \leq n$ and since we run the search process for $3n$ steps,



$$
\begin{aligned}
q_j &= \Pr[X_{3n} = n] \\
&\geq \Pr[X_{3j} = n] \\
&\geq \Pr[\text{take exactly } 2j \text{ steps forward and } j \text{ steps back in } 3j \text{ steps}]
\end{aligned}
$$

Modified 3-SAT Analysis

Consider a single random assignment with $\mathsf{X}_0 = n - j$. I.e., we need to correct j variables to find a satisfying assignment.

Let q_j be a lower bound on the success probability in this case. Since $j \leq n$ and since we run the search process for $3n$ steps,

$$\begin{aligned} q_j &= \Pr[\mathsf{X}_{3n} = n] \\ &\geq \Pr[\mathsf{X}_{3j} = n] \\ &\geq \Pr[\text{take exactly } 2j \text{ steps forward and } j \text{ steps back in } 3j \text{ steps}] \\ &= \binom{3j}{j}\left(\frac{2}{3}\right)^{j} \cdot \left(\frac{1}{3}\right)^{2j}. \end{aligned}$$

Modified 3-SAT Analysis

Consider a single random assignment with $X_0 = n - j$. I.e., we need to correct j variables to find a satisfying assignment.

Let q_j be a lower bound on the success probability in this case. Since $j \leq n$ and since we run the search process for $3n$ steps,

$$\begin{aligned} q_j &= \Pr[X_{3n} = n] \\ &\geq \Pr[X_{3j} = n] \\ &\geq \Pr[\text{take exactly } 2j \text{ steps forward and } j \text{ steps back in } 3j \text{ steps}] \\ &= \binom{3j}{j} \left(\frac{2}{3}\right)^j \cdot \left(\frac{1}{3}\right)^{2j}. \end{aligned}$$

Via Stirling's approximation, $\binom{3j}{j} \geq \frac{1}{\sqrt{j}} \cdot \frac{3^{3j-2}}{2^{2j-2}}$, giving:

$$\underline{q_j} \geq \frac{2^2}{3^2\sqrt{j}} \cdot \frac{3^{3j}}{2^{2j}} \cdot \frac{2^j}{3^{3j}} \approx \frac{1}{\sqrt{j} \cdot 2^j} \geq \boxed{\frac{1}{\sqrt{n} \cdot 2^j}.}$$

Modified 3-SAT Analysis

Our overall probability of success in a single trial is then lower bounded by:

$$q \geq \sum_{j=0}^{n} \Pr[X_0 = n - j] \cdot q_j$$

Modified 3-SAT Analysis

Our overall probability of success in a single trial is then lower bounded by:

$$
\begin{aligned}
q &\geq \sum_{j=0}^{n} \Pr[X_0 = n - j] \cdot q_j \\
&\geq \sum_{j=0}^{n} \underbrace{\binom{n}{j} \cdot \frac{1}{2^n}} \cdot \frac{1}{\sqrt{n} \cdot 2^j}
\end{aligned}
$$

Modified 3-SAT Analysis

Our overall probability of success in a single trial is then lower bounded by:

$$
\begin{aligned}
q &\geq \sum_{j=0}^{n} \Pr[X_0 = n - j] \cdot q_j \\
&\geq \sum_{j=0}^{n} \binom{n}{j} \cdot \frac{1}{2^n} \cdot \frac{1}{\sqrt{n} \cdot 2^j} \\
&\geq \frac{1}{\underline{\sqrt{n} \cdot 2^n}} \sum_{j=0}^{n} \binom{n}{j} \cdot \frac{1}{2^j}
\end{aligned}
$$

Modified 3-SAT Analysis

Our overall probability of success in a single trial is then lower bounded by:

$$
\begin{aligned}
q &\geq \sum_{j=0}^{n} \Pr[X_0 = n - j] \cdot q_j \\
&\geq \sum_{j=0}^{n} \binom{n}{j} \cdot \frac{1}{2^n} \cdot \frac{1}{\sqrt{n} \cdot 2^j} \\
&\geq \frac{1}{\sqrt{n} \cdot 2^n} \sum_{j=0}^{n} \binom{n}{j} \cdot \frac{1}{2^j} \\
&= \frac{1}{\sqrt{n} \cdot 2^n} \cdot \left(\frac{3}{2}\right)^n
\end{aligned}
$$

$\left(1 + \frac{1}{2}\right)^n = \sum_{j=0}^{n} \binom{n}{j} \frac{1}{2^j}$

Modified 3-SAT Analysis

Our overall probability of success in a single trial is then lower bounded by:

$$
\begin{aligned}
q &\geq \sum_{j=0}^{n} \Pr[X_0 = n - j] \cdot q_j \\
&\geq \sum_{j=0}^{n} \binom{n}{j} \cdot \frac{1}{2^n} \cdot \frac{1}{\sqrt{n} \cdot 2^j} \\
&\geq \frac{1}{\sqrt{n} \cdot 2^n} \sum_{j=0}^{n} \binom{n}{j} \cdot \frac{1}{2^j} \\
&= \frac{1}{\sqrt{n} \cdot 2^n} \cdot \left(\frac{3}{2}\right)^n = \underline{\frac{1}{\sqrt{n}} \cdot \left(\frac{3}{4}\right)^n}.
\end{aligned}
$$

Modified 3-SAT Analysis

Our overall probability of success in a single trial is then lower bounded by:

$$
\begin{aligned}
q &\geq \sum_{j=0}^{n} \Pr[X_0 = n - j] \cdot q_j \\
&\geq \sum_{j=0}^{n} \binom{n}{j} \cdot \frac{1}{2^n} \cdot \frac{1}{\sqrt{n} \cdot 2^j} \\
&\geq \frac{1}{\sqrt{n} \cdot 2^n} \sum_{j=0}^{n} \binom{n}{j} \cdot \frac{1}{2^j} \\
&= \frac{1}{\sqrt{n} \cdot 2^n} \cdot \left(\frac{3}{2}\right)^n = \frac{1}{\sqrt{n}} \cdot \left(\frac{3}{4}\right)^n.
\end{aligned}
$$

Thus, if we repeat for $m = O\left(\sqrt{n} \cdot \left(\frac{4}{3}\right)^n\right) = O(1.33334^n)$ trials, with very high probability, we will find a satisfying assignment if there is one.