

# CSE 203A: Randomized Algorithms

Spring 2026

**Lecture 13:** Generalization to RSAT

**Date:** 27 Apr 2026

**Instructor:** Daniel M. Kane

**Scribe:** Zhuohan Zhou

## Lecture Overview

In this lecture, we reviewed the Random Walk algorithm for the 3-SAT (and  $k$ -SAT) problem and introduced the formal mathematical framework of **Markov Chains** to analyze such state-transition processes. We established the conditions under which a Markov Chain converges to a stationary (steady-state) distribution, specifically focusing on finite, connected, and aperiodic chains. We used linear algebra, specifically **eigenvalue analysis** of transition matrices, to understand the long-term behavior of state distributions and bound the maximum eigenvalues.

## 1 Introduction

In this lecture, we transition from analyzing specific randomized algorithms (like the one for 3-SAT) to establishing a generalized framework for any process that moves from state to state with certain probabilities. By modeling these processes as Markov Chains, we can use transition matrices and their eigenvalues to rigorously prove whether an algorithm will eventually converge to a desired steady state, which is a fundamental concept for analyzing random walks on graphs.

## 2 Brief Reminders about Previous Topics

### Random Walk Algorithm for $k$ -SAT

---

**Algorithm 1:** Random Walk Algorithm for  $k$ -SAT

---

**Input:** CNF Formula with  $n$  variables and clause size  $k$

**Output:** A satisfying assignment  $X$ , or Failure

Sample a uniformly random assignment  $X \in \{0, 1\}^n$

**for**  $i = 1$  **to**  $3n$  **do**

**if**  $X$  *satisfies all clauses* **then**

**return**  $X$

    Find an arbitrary violated clause  $C$

    Pick a variable  $x_j$  uniformly at random from the variables in  $C$

    Flip the value of  $x_j$  in  $X$

**return** Failure

---

### Why does this work better than global random changes?

If there is a perfect satisfying assignment  $Y$ , and our current assignment  $X$  violates a specific clause, then  $X$  and  $Y$  *must differ* in at least one of the inputs to that specific clause. By restricting our random flip to only the variables within that violated clause, we have a  $1/k$  chance (for  $k$ -SAT, e.g.,

1/3 for 3-SAT) of directly correcting an error. If  $X$  is already reasonably close to  $Y$ , flipping globally would rarely hit an error, but this "local flip" acts as a targeted radar, significantly increasing the probability of moving closer to  $Y$ .

### Probability of reaching the solution:

If we assume  $X$  is  $\alpha n$  steps away (distance) from the solution  $Y$ , the probability that all of the next  $\alpha n$  steps correct errors directly can be approximated as:

$$P \sim \frac{\binom{n}{\alpha n}}{2^n} \left(\frac{1}{k}\right)^{\alpha n}$$

Using Stirling's formula to approximate the binomial coefficient, this expands to:

$$\approx \frac{\left(\left(\frac{1}{\alpha}\right)^\alpha \left(\frac{1}{1-\alpha}\right)^{1-\alpha}\right)^n}{2^n} \left(\frac{1}{k^\alpha}\right)^n$$

Since  $\alpha$  is small,  $1 - \alpha$  is close to 1, meaning  $\frac{1}{1-\alpha} > 1$ . To simplify and find a lower bound, we can drop this term, yielding:

$$> \frac{1}{2^n} \left(\frac{1}{k\alpha}\right)^{\alpha n}$$

To optimize this probability, we can set  $\alpha = \frac{1}{2k}$ . Substituting this into our lower bound gives:

$$> \frac{1}{2^n} \left(\frac{1}{k \cdot \frac{1}{2k}}\right)^{\frac{n}{2k}} = \frac{1}{2^n} (2)^{\frac{n}{2k}} = 2^{-n} \cdot 2^{\frac{n}{2k}} = 2^{-n(1-\frac{1}{2k})}$$

This represents the lower bound on the probability of reaching the solution in one full run. The expected time (number of attempts) to find the solution is proportional to the inverse of this probability.

Therefore, the overall time complexity of this randomized algorithm for  $k$ -SAT is approximately:

$$\text{Time} \approx 2^{n(1-\frac{1}{2k})}$$

## 3 Main Lecture Content

### 3.1 Markov Chains: Definition and Transition Rules

A Markov Chain is a sequence of random variables  $X_0, X_1, X_2, \dots$  that satisfies the **Markov Property** (memorylessness) — the next state depends only on the current state, not on the sequence of events that preceded it:

$$Pr(X_{n+1}|X_n, X_{n-1}, \dots, X_0) = Pr(X_{n+1}|X_n)$$

We assume these are **Time-Independent (Homogeneous)**, meaning the transition probabilities do not change over time:

$$Pr(X_{t+1} = i|X_t = j) = Pr(X_{s+1} = i|X_s = j)$$

An intuitive example is a random walk on a directed graph  $G$ . The random variables  $X_i$  represent the vertices. If  $X_n = v$ , then  $X_{n+1}$  is a uniformly random neighbor of  $v$ .

### 3.2 Transition Matrices

Let  $q_i^t = Pr(X_t = i)$  be the probability distribution over states at time  $t$ . We define the transition probability from state  $j$  to state  $i$  as  $P_{ji} = Pr(X_{n+1} = i | X_n = j)$ . Since probabilities must sum to 1, for any state  $j$ :

$$\sum_i P_{ji} = 1$$

We can compute the distribution at time  $t + 1$  using the Law of Total Probability:

$$q_i^{t+1} = \sum_j Pr(X_{t+1} = i | X_t = j) Pr(X_t = j) = \sum_j q_j^t P_{ji}$$

In matrix notation, if  $q^t$  is a row vector and  $P$  is the transition matrix:

$$q^{t+1} = q^t P \implies q^t = q^0 P^t$$

### 3.3 Eigenvalue Analysis of Long-Term Behavior

To understand  $q^t$  as  $t \rightarrow \infty$ , we analyze the eigenvalues of  $P$ . Suppose  $P$  has eigenvectors  $v_i$  with corresponding eigenvalues  $\lambda_i$ . We can express the initial distribution as a linear combination of eigenvectors:  $q^0 = \sum a_i v_i$ . Then the state at time  $t$  is:

$$q^t = \sum a_i \lambda_i^t v_i$$

For large  $t$ , this sum is dominated by the terms with the largest  $|\lambda_i|$ .

**What is the largest  $|\lambda|$ ?**

Using the  $L_1$  norm, we can prove that  $|\lambda_i| \leq 1$ :

$$|qP|_1 = \sum_i \left| \sum_j q_j P_{ji} \right| \leq \sum_{i,j} |q_j| P_{ji} = \sum_j \left( |q_j| \sum_i P_{ji} \right) = \sum_j |q_j| = |q|_1$$

Thus, multiplying by  $P$  does not increase the  $L_1$  norm, meaning any eigenvalue  $\lambda$  must satisfy  $|\lambda| \leq 1$ .

Furthermore, since  $\sum_i P_{ji} = 1$ , we know  $P\mathbf{1} = \mathbf{1}$  (where  $\mathbf{1}$  is the all-ones column vector). Thus,  $P$  has a right eigenvalue of exactly 1. Because  $(P - I)$  is non-invertible,  $P$  also has 1 as a left eigenvalue. The corresponding left eigenvector  $\pi$  satisfies:

$$\pi = \pi P$$

This  $\pi$  is called the **steady-state (or stationary) distribution**.

### 3.4 Convergence and Obstacles

Conjecture:  $q^t \rightarrow \pi$  as  $t \rightarrow \infty$ . However, there are three main obstacles to this convergence:

1. **Disconnected Subgraphs:** Leads to multiple steady states depending on the starting node.
2. **Periodic Behavior:** For a vertex  $v$ , it's only possible to return to  $v$  after some time that's a multiple of  $n$  (where  $n > 1$ ). For example, in a bipartite graph, you can only return to your starting side after an even number of steps, preventing convergence to a stable distribution.
3. **Technical issues for infinite Markov Chains** In an infinite state space, the process can be "transient" (drifting to infinity and never returning) or "null recurrent." In such cases, a stationary distribution  $\pi$  that sums to 1 may not exist, which is why the Fundamental Theorem specifically requires the chain to be **finite**.

#### Fundamental Theorem of Markov Chains:

For a finite, connected (irreducible), and aperiodic Markov Chain, from any initial state, the distribution  $q^t$  converges to a unique steady state  $\pi$ .

*Note on aperiodicity:* For an aperiodic chain, the set of possible return times has a Greatest Common Divisor (GCD) of 1. Because of this, there's always an integer linear combination of return times that sums to a sufficient length  $L$ , ensuring that eventually, the process doesn't get trapped in a cyclic bipartite-like oscillation.

### 3.5 Coupling Argument

There are some steady state  $\pi$ , and we have two copies of random walk:  $X, Y$ .

$X_0 \sim q^0, Y_0 \sim \pi$

$X$  &  $Y$  evolve independently until  $X_t = Y_t$ . Then they evolve in sync.

**Claim:** as  $t \rightarrow \infty, Pr(X_t = Y_t) \rightarrow 1$

#### Core Lemma (Meeting Probability):

$\exists N, \varepsilon > 0$  s.t. for any states  $x_t = a, y_t = b$ :

$$Pr(X_{N+t} = Y_{N+t} \mid X_t = a, Y_t = b) \geq \varepsilon$$

#### Intuition:

- **Connectivity:** Both can eventually reach some same vertex  $B$ .
- **Aperiodicity:** GCD of return times to  $B$  is 1  $\implies$  possible for  $X, Y$  to be at  $B$  at the *exact same time* (after  $N$  steps).
- **Exponential Decay:** Because of the Lemma, every  $N$  steps, there is  $\geq \varepsilon$  chance they join. Thus,  $Pr(\text{not met}) \leq (1 - \varepsilon)^{\lfloor t/N \rfloor} \rightarrow 0$  exponentially.
- Since  $Y_t$  is always in steady state  $\pi, q^t \rightarrow \pi$ .  $\pi$  is the unique steady state.

## 4 Summary

In this lecture, we showed why local random flips in violated clauses are significantly more efficient for SAT solving than global random flips. We then abstracted random state-transitions into the rigorous framework of Markov Chains. By analyzing the transition matrix  $P$ , we proved that its

eigenvalues are bounded by 1, and that a steady state  $\pi$  exists. Finally, we outlined the requirements—finiteness, connectivity, and aperiodicity—for a Markov process to reliably converge to this steady state.

## References

- [1] R. Motwani and P. Raghavan, *Randomized Algorithms*, Cambridge University Press, 1995.