Seoul
National
University

Markov Chains & Random Walks

Name: Chong-kwon Kim

SCONE
Lab.

Stochastic Process

- A **stochastic process** X = {X(t): t ∈T} is a collection of random variables
  - Trace how the value of a **random variable** changes as the **time** flows
  - Example: Average daily temperature,
    # received cacao-talk messages/day
    # empty bins as balls are thrown to n bins

- **Space (State)**
  - Values of rv
  - Discrete
    - Discrete state process
  - Continuous

- Time
  - Discrete
    - Discrete time process
    - X_t as X(t)
  - Continuous

Markov Chain

- Definition: Markov Chain

  A discrete time process X_1, X_2, … is a Markov chain if

  $$\Pr(X_t = a_t \mid X_{t-1} = a_{t-1}, X_{t-2} = a_{t-2}, \cdots, X_0 = a_0)$$
  $$= \Pr(X_t = a_t \mid X_{t-1} = a_{t-1})$$

  Markov property
  Memoryless property

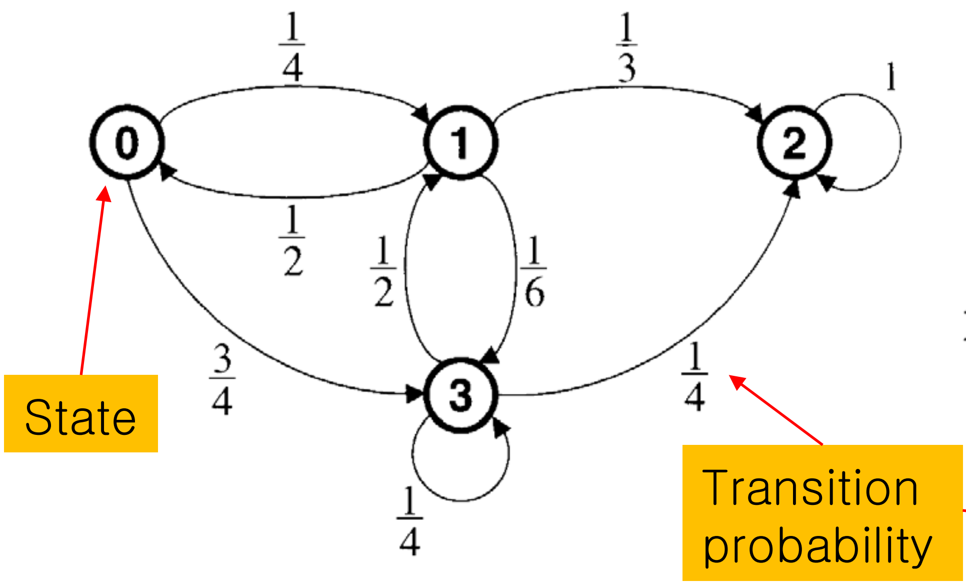
- Transition probability

  $$P_{i,j} = \Pr(X_t = j \mid X_{t-1} = i)$$

- Transition matrix

$$\mathbf{P} = \begin{pmatrix} P_{0,0} & P_{0,1} & \cdots & P_{0,j} & \cdots \\ P_{1,0} & P_{1,1} & \cdots & P_{1,j} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \ddots \\ P_{i,0} & P_{i,1} & \cdots & P_{i,j} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \ddots \end{pmatrix}$$

Markov Chain – Directed Graph



$$\mathbf{P} = \begin{bmatrix} 0 & 1/4 & 0 & 3/4 \\ 1/2 & 0 & 1/3 & 1/6 \\ 0 & 0 & 1 & 0 \\ 0 & 1/2 & 1/4 & 1/4 \end{bmatrix}$$

Transition from state 0 to state 3 in exactly three steps

- 0–1–0–3: $\frac{1}{4} \cdot \frac{1}{2} \cdot \frac{3}{4} = 3/32$
- 0–1–3–3: $\frac{1}{4} \cdot 1/6 \cdot \frac{1}{4} =$
- 0–3–1–3: $\frac{3}{4} \cdot \frac{1}{2} \cdot 1/6 =$
- 0–3–3–3: $\frac{3}{4} \cdot \frac{1}{4} \cdot \frac{1}{4} =$

Transition Probability

- Let $p_i(t)$ be the probability that X_t = i
- Let a vector $\bar{p}(t) = (p_0(t), p_1(t)\ , p_2(t)\ , \cdots)$ be a distribution of states at time t

$$p_i(t) = \sum_j p_j(t-1) \cdot P_{j,i}$$

$$\bar{p}(t) = \bar{p}(t-1) \cdot P$$

- m-step transition probability

$$P_{i,j}^{(m)} = \Pr(X_{t+m} = j \,|\, X_t = i\ )$$

- Let $P^{(m)}$ be the m-step transition matrix
  - (i, j) component of $P^{(m)}$ $\left(P_{i,j}^{(m)}\right)$ is $P_{i,j}^{m}$
- Show that $P^{(m)} = P^m$

m-time multiplication of P

SAT Problem – Preliminary

- **Boolean expression (formula)**
  - Expression built from variables using AND(^, *), OR(v, +) and NOT(−) operators
    - Precedence order: NOT→AND→OR
  - A formula $(x+\bar{y}) \cdot (x+y)$ is TRUE if x=T
  - Variable (x, y) is called literal and (x+y), $(x+\bar{y})$ are called clause
    - A clause is OR of literals
  - CNF(Conjunctive Normal Form): AND of Clauses

- **SAT (Satisfiability) problem**
  - Are there T/F assignments (Truth assignment) to variables(literals) that make the formula TRUE?
  - $(x+\bar{y})*(x+y)$
  - $(x+y)*(x+\bar{y})*(\bar{x}+y)*(\bar{x}+\bar{y})$

2-SAT Problem

- K-SAT
  - Each clause has exactly k literals

SAT is NP-Hard
But, 2-SAT is Polynomial, O(n^3)

- 2-SAT
  - Example
    - $(x \lor \neg y) \land (x \lor y) \land (y \lor z) \land (\neg x \lor \neg z) \rightarrow (x+\bar{y}) \cdot (x+y) \cdot (y+z) \cdot (\bar{x}+\bar{z})$

Start with
x=T
y=T
z=T

$(\bar{x}+\bar{z})$ = F
Change x=T→F
y=T
z=T

$(x+\bar{y})$ = F
x=F
Change y=T→F
z=T

(x+y) = F
Change x=F→T
y=F
z=T

$(\bar{x}+\bar{z})$ = F
x=T
y=F
Change z=T→F

(y+z) = F
x=T
Change y=F→T
z=F

Done

Probabilistic 2-SAT Algorithm

SCONE Lab.

Algorithm
1. Start w/ arbitrary truth assignment
2. Repeat up to $2mn^2$ before a solution S is found
   (a) Choose an unsatisfied clause randomly
   (b) Choose one literal randomly and switch its value
3. If found, return the solution, ow the formula is unsatisfiable

Papadimitriou. On selecting a satisfying truth assignment. IEEE FOCS, 1991

A Monte Carlo algorithm that may give incorrect answer
m: controls the error probability

- Notations
  - S be a satisfying assignment
  - Ai: Truth assignment after i-th change
  - Xi: # variables in Ai that are identical to S

- Obviously, $\Pr(X_{i+1} = 1 \mid X_i = 0) = 1$

Probabilistic 2-SAT Algorithm

Algorithm
1. Start w/ arbitrary truth assignment
2. Repeat up to $2mn^2$ before a solution S is found
   (a) Choose an unsatisfied clause randomly
   (b) Choose one literal randomly and switch its value
3. If found, return the solution, ow the formula is unsatisfiable

Papadimitriou. On selecting a satisfying truth assignment. IEEE FOCS, 1991

A Monte Carlo algorithm that may give incorrect answer
m: controls the error probability

Unsatisfiable	➔	Unsatisfiable
Satisfiable	➔	Find solution in 2mn^2 steps OR Unsatisfiable

Probability of failure? ⇨ **MC to derive prob.**

2-SAT Algorithm – Analysis

- Notations
  - S be a satisfying assignment
  - Ai: Truth assignment after i-th change
  - Xi: # variables in Ai that are identical to S
- Obviously, $\Pr(X_{i+1} = 1 \mid X_i = 0) = 1$
- Suppose now that $1 \le X_i \le n-1$
  - Consider Ai (where Xi = j) and a unsatisfied clause
  - One or both literals in the clause have different assign'ts between Ai and S

Why?

➔ $\Pr(X_{i+1} = j+1 \mid X_i = j) \ge 1/2$

➔ $\Pr(X_{i+1} = j-1 \mid X_i = j) \le 1/2$

- The transition probabilities are not fixed

➔We fix them pessimistically and define a Markov chain based on Yi (reflects pessimistic case)

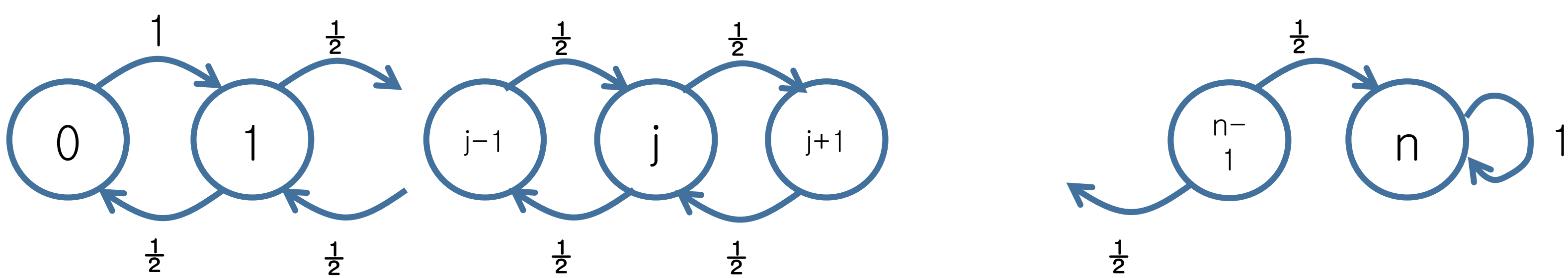
$\Pr(Y_{i+1} = 1 \mid Y_i = 0) = 1$

$\Pr(Y_{i+1} = j+1 \mid Y_i = j) = 1/2$

$\Pr(Y_{i+1} = j-1 \mid Y_i = j) = 1/2$

2-SAT Algorithm – Analysis

- Define a Markov Chain
  - State: # variables matched in Ai and S



  - Define
    - Zj: Random variable, # steps required to reach state n starting from state j
    - h_j: Expectation of Zj

$$E[Z_j] = \tfrac{1}{2}(1+E[Z_{j-1}]) + \tfrac{1}{2}(1+E[Z_{j+1}])$$

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

2-SAT Algorithm – Analysis

SCONE Lab.

- From a set of equations

$$h_n = 0$$
$$h_0 = h_1 + 1$$
$$h_j = \frac{h_{j-1}}{2} + \frac{h_{j+1}}{2} + 1, \text{ for } 0 < j \le (n-1)$$

Inductively show that $h_j = h_{j+1} + 2j + 1$

Finally,

$h_0 = h_1 + 1 = h_2 + 1 + 3 = \sum_{i=0}^{n-1} 2i + 1 = n^2$

2–SAT Algorithm – Analysis

- Claim
  - The probability of failure is at most $(\frac{1}{2})^m$

- Proof
  - Consider the algorithm as m repetitions of $2n^2$ steps
  - Each repetition starts at a certain state j ($\neq$ n)
    - Z: Random variable of # steps to reach the state n
    - $h_j \leq n^2$
    - Pr (Z > $2n^2$) $\leq n^2/2n^2 = \frac{1}{2}$

      ➔ Prob. of not finding a solution after $2n^2$ steps is at most 1/2

  ➔ Pr(All m repetitions fail) $\leq (\frac{1}{2})^m$

Now, you can read Papadimitriou's paper!!

3-SAT Problem

- 2-SAT is Polynomial
- 3-SAT Polynomial also?

No, But, there is an algorithm whose Average is Polynomial

3-SAT Algorithm

1. Start w/ arbitrary truth assignment
2. Repeat up to m before a solution S is found
   (a) Choose randomly an unsatisfied clause
   (b) Choose one literal randomly and switch its value
3. If found, return the solution, ow the formula is unsatisfiable

U. Schöning, A probabilistic algorithm for k-SAT and constraint satisfaction problems, IEEE FOCS, 1999.

3-SAT Problem

- Like 2-SAT, define
  - Ai and Xi

    ➔ $\Pr(X_{i+1} = j+1 \mid X_i = j) \geq 1/3$

    ➔ $\Pr(X_{i+1} = j-1 \mid X_i = j) \leq 2/3$

  - Also define pessimistic Yi as before

    $Y_0 = X_0$

    $\Pr(Y_{i+1} = 1 \mid Y_i = 0) = 1$

    $\Pr(Y_{i+1} = j+1 \mid Y_i = j) = 1/3$

    $\Pr(Y_{i+1} = j-1 \mid Y_i = j) = 2/3$

    $h_n = 0$

    $h_0 = h_1 + 1$

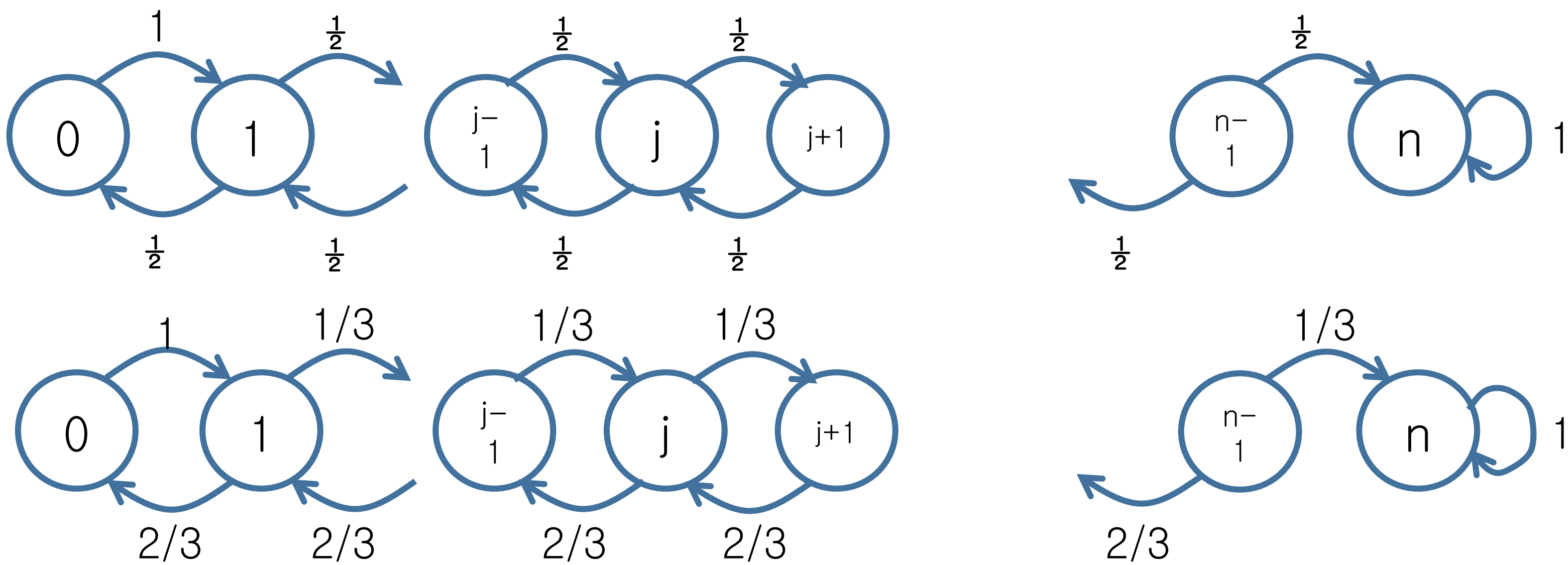
    $h_j = \frac{2 \cdot h_{j-1}}{3} + \frac{h_{j+1}}{3} + 1$, for $0 \leq j \leq (n-1)$

  ➔ $h_j = 2^{n+2} - 2^{j+2} - 3(n-j)$

$\boldsymbol{h_j = \Theta(2^n)}$

3-SAT Algorithm

- Strategies for 2-SAT and 3-SAT



Observations
1. #correct variables from random truth assignment is ?
2. From the initial state, 3-SAT becomes worse as more steps are taken

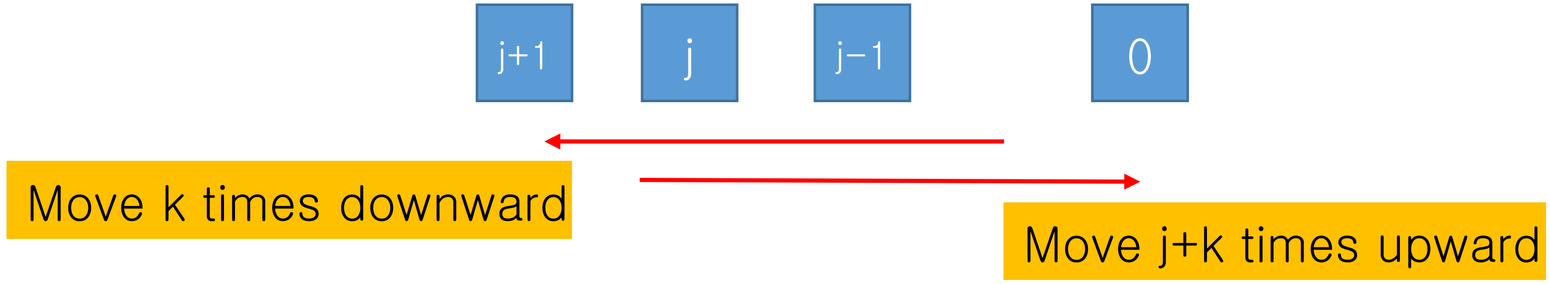
Instead of one long steps, run multiple short runs with different initial assignments

3-SAT Algorithm

Modified 3-SAT Algorithm

1. Repeat up to m times, before satisfied
   (a) Start w/ arbitrary truth assignment
   (b) Repeat up to $3n$ times, before satisfied
      i) Choose an unsatisfied clause randomly
      ii) Choose one literal randomly and switch its value
2. If found, return the solution, ow the formula is unsatisfiable

- Analysis
  - Define j : # incorrect variables
  - q_j= Pr(Reach to S after correcting j incorrect variables)



3-SAT Algorithm – Analysis

- Analysis
  - The probability to reach S, even if k downward moves are included, is at least

$$\binom{j+2k}{k}\left(\frac{2}{3}\right)^k\left(\frac{1}{3}\right)^{j+k},\ k = 0, 1, 2, \cdots, j$$

$$\Rightarrow q_j \geq \min_{k=0,1,\ldots,j}\left\{\binom{j+2k}{k}\left(\frac{2}{3}\right)^k\left(\frac{1}{3}\right)^{j+k}\right\}$$

$$\geq \binom{3j}{j}\left(\frac{2}{3}\right)^j\left(\frac{1}{3}\right)^{2j}$$

Apply Sterling's Formula

$$\sqrt{2\pi m}\left(\frac{m}{e}\right)^m \leq m! \leq 2\sqrt{2\pi m}\left(\frac{m}{e}\right)^m$$

$$\geq \frac{c}{\sqrt{j}}\left(\frac{27}{4}\right)^j\left(\frac{2}{3}\right)^j\left(\frac{1}{3}\right)^{2j}$$

$$\geq \frac{c}{\sqrt{j}}\frac{1}{2^j}$$

$c=\sqrt{3}/8\sqrt{\pi}$

  - $q \geq \sum_{j=0}^{n} \Pr(j \text{ mismatches in a random assignment}) \cdot q_j$

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

$q \geq \frac{c}{\sqrt{n}}\left(\frac{3}{4}\right)^n$

3-SAT Algorithm – Analysis

- Analysis
  - With one random assignment, reach to S with the probability at least q
  - Repeat with new random assignment until SUCCESS

    ➔ Geometric distribution with parameter q

    ➔ # trials until SUCCESS = 1/q

Note: q ≥ $\frac{c}{\sqrt{n}}(\frac{3}{4})^n$

  - Each repetition requires at most 3n steps
  - Expected # steps until SUCCESS is O($n^{3/2}(\frac{4}{3})^n$)

b repetitions of $2 \cdot \frac{3}{c} \cdot n^{3/2}(\frac{4}{3})^n$ step batch
Pr(Failure) $\leq 2^{-b}$