Georg-August-Universität
Göttingen

Probability Models II

Markov Chain and Its Applications

Wenzhong Li
lwz@nju.edu.cn

Markov Chain

- *Stochastic process*: a stochastic process

  $X = \{X(t) : t \in T\}$

  is a collection of random variables on a index set T.
  - If t represent time, this process models X changing over time
- A *discrete-time Markov chain* is a sequence of random variables {X_n, n=1,2,…}, X_n Є {0,1,2,…}, with the *Markov property*, namely that, the next state depends only on the current state and not on the past.

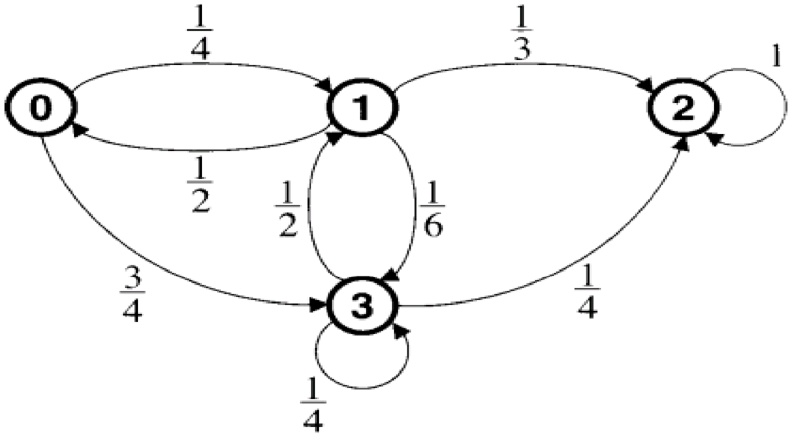
$$Pr(X_{n+1} = i_{n+1} | X_1 = i_1, X_2 = i_2, \cdots, X_n = i_n) = Pr(X_{n+1} = i_{n+1} | X_n = x_n)$$

Representation of Markov Chain

- (1) Transition matrix
  - Let P_{ij} represent the probability that the process will make transitions from state I to state j,

$$P_{ij} \geqslant 0, \qquad i,j \geqslant 0; \qquad \sum_{j=0}^{\infty} P_{ij} = 1, \qquad i = 0, 1, \ldots$$

  - Let $P=[P_{ij}]$ denote the matrix of one-setp transition probability
- (2) State diagram
  - Markov chain can also be represented by a directed, weighted graph D=(V,E,w)
  - The set of the graph is the set of states of the chain
  - There is a directed edge (i,j) Є E if and only if P_{ij},>0
  - The weight $w(i,j) = P_{ij}$

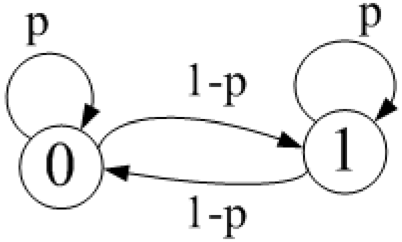


$$\mathbf{P} = \begin{bmatrix} 0 & 1/4 & 0 & 3/4 \\ 1/2 & 0 & 1/3 & 1/6 \\ 0 & 0 & 1 & 0 \\ 0 & 1/2 & 1/4 & 1/4 \end{bmatrix}$$

Example: A Communication System

- Consider a communications system that transmits the digits 0 and 1.
- Each digit transmitted must pass through several stages
- At each stage there is a probability p that the digit entered will be unchanged when it leaves.
- Let X_n denote the digit entering the nth stage, then $\{X_n, n = 0, 1, \ldots\}$ is a two-state Markov chain

$$\mathbf{P} = \left\| \begin{matrix} p & 1-p \\ 1-p & p \end{matrix} \right\|$$

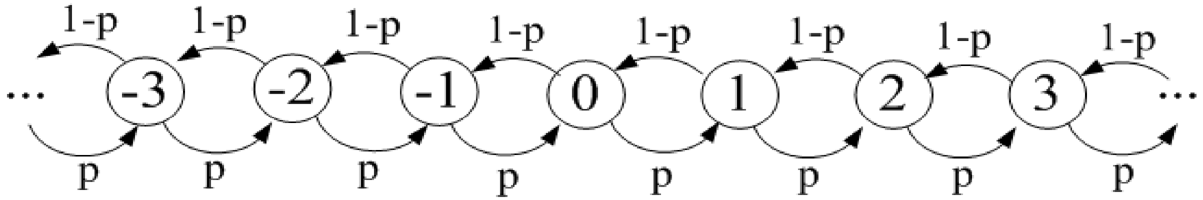


More Examples

- Random Walk
  - A Markov chain whose state space is given by the integers i = 0, $\pm$1, $\pm$2, ... is said to be a random walk if, for some number 0 < p < 1,

$$P_{i,i+1} = p = 1 - P_{i,i-1}, \qquad i = 0, \pm 1, \ldots$$

  - It models an individual walking on a straight line who at each point of time either takes one step to the right with probability p or one step to the left with probability 1 − p.

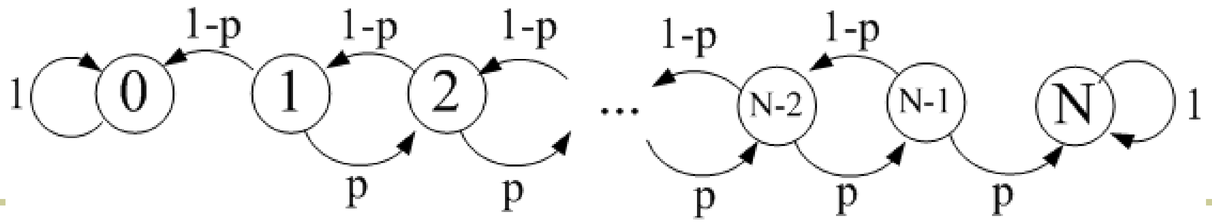


- A Gambling Model
  - A gambler at each game will wins $1 with probability p and losses $1 with probability 1-p.
  - The gambler will quits playing when he goes broke or he attains a fortune of $N.

$$P_{i,i+1} = p = 1 - P_{i,i-1}, \qquad i = 1, 2, \ldots, N-1,$$

$$P_{00} = P_{NN} = 1$$

  - State 0 and N are called *absorbing states* since once entered they are never left



Chapman-Kolmogorov Equations

- n-step transition probability
  - ○ P_{ij}^{n} : the probability that a process in state i will be in state j after n additional transitions

$$P_{ij}^{n} = P\{X_{n+k} = j \mid X_k = i\}, \quad n \geqslant 0,\ i,j \geqslant 0$$

- Chapman–Kolmogorov equations

$$P_{ij}^{n+m} = \sum_{k=0}^{\infty} P_{ik}^{n} P_{kj}^{m} \quad \text{for all } n, m \geqslant 0,\ \text{all } i, j$$

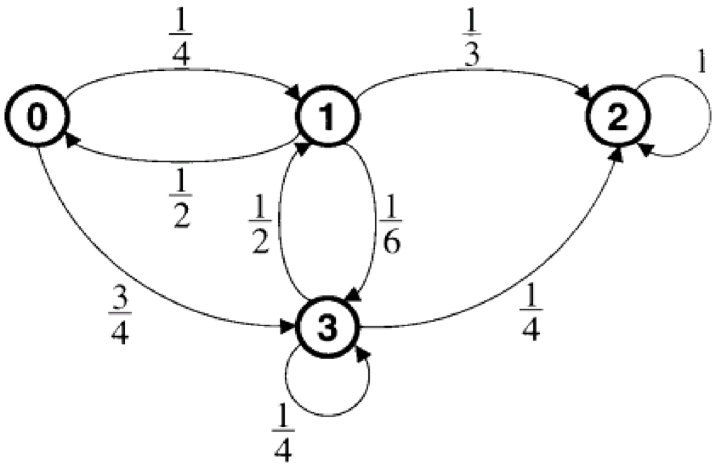
  - ○ It provides a method for computing these n-step transition probabilities
  - ○ $P_{ik}^{n} P_{kj}^{m}$ represents the probability that starting in i the process will go to state j in n + m transitions through a path which takes it into state k at the nth transition
- If we let $\mathbf{P}^{(n)}$ denote the matrix of n-step transition probabilities

$$\mathbf{P}^{(n)} = \mathbf{P}^{(n-1+1)} = \mathbf{P}^{n-1} \cdot \mathbf{P} = \mathbf{P}^{n}$$

Example

- For example, calculate the probability of going from 0 to state 3 in exactly three steps.
- (1) use the graph, there are four paths:
  - 0-1-0-3, 0-1-3-3, 0-3-1-3, 0-3-3-3
  - Probability is 3/32+1/96+1/16+3/64=41/192
- (2) use transition matrix:

$$\mathbf{P}^3 = \begin{bmatrix} 3/16 & 7/48 & 29/64 & 41/192 \\ 5/48 & 5/24 & 79/144 & 5/36 \\ 0 & 0 & 1 & 0 \\ 1/16 & 13/96 & 107/192 & 47/192 \end{bmatrix}.$$



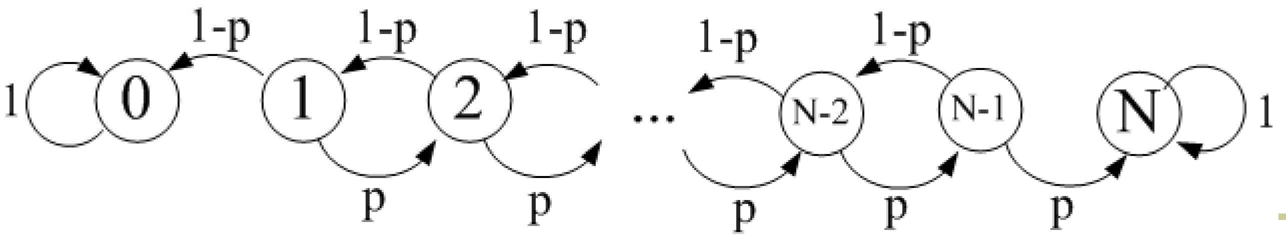
$$\mathbf{P} = \begin{bmatrix} 0 & 1/4 & 0 & 3/4 \\ 1/2 & 0 & 1/3 & 1/6 \\ 0 & 0 & 1 & 0 \\ 0 & 1/2 & 1/4 & 1/4 \end{bmatrix}$$

Classification of States

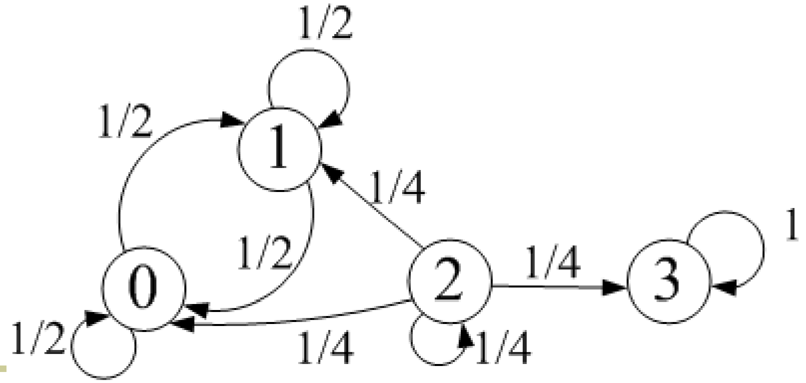
- A first step in analyzing the long-term behavior of a Markov chain is to classify its states
- *Accessible*
  - State j is said to be *accessible* from state i if $P_{ij}^n > 0$ for some n≥0.
  - This implies that state j is accessible from state i if and only if, starting in i, it is possible that the process will ever enter state j.
  - Since if j is not accessible from i, then

$$P\{\text{ever enter } j|\text{start in } i\} = P\left\{\bigcup_{n=0}^{\infty}\{X_n = j\}\middle| X_0 = i\right\}$$

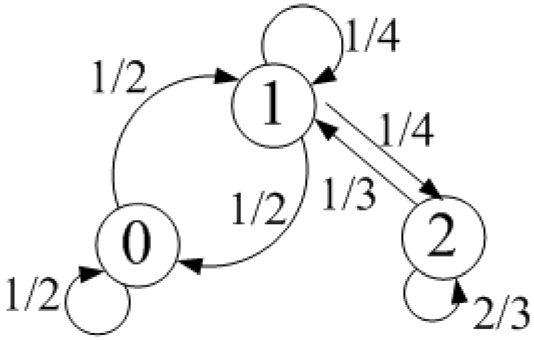
$$\leqslant \sum_{n=0}^{\infty} P\{X_n = j|X_0 = i\} = \sum_{n=0}^{\infty} P_{ij}^n = 0$$



- *Communicate*
  - Two states i and j that are accessible to each other are said to *communicate*, write as i↔j
  - Any state communicates with itself since
$$P_{ii}^0 = P\{X_0 = i | X_0 = i\} = 1$$
  - Properties:
    - (i) *reflexive*: State i communicates with state i, all i 0.
    - (ii) *symmetric*: If state i communicates with state j, then state j communicates with state i.
    - (iii) *transitive*: If state i communicates with state j, and state j communicates with state k, then state i communicates with state k.



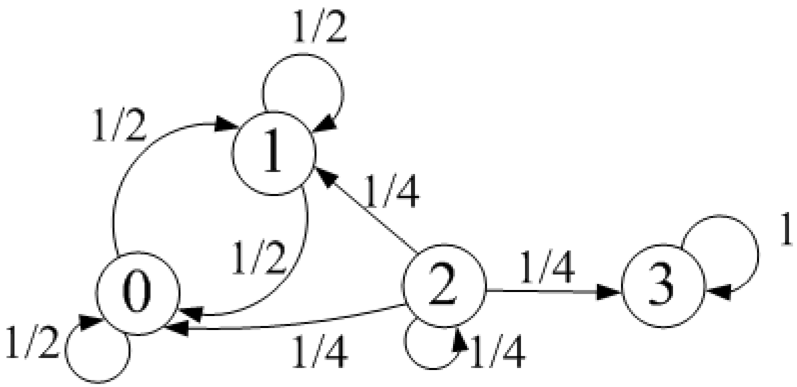
- *Class*
  - Two states that communicate are said to be in the same *class*.
  - Any two classes of states are either identical or disjoint
  - The concept of communication divides the state space up into a number of separate classes
- *Irreducible*
  - The Markov chain is said to be *irreducible* if there is only one class, that is, if all states communicate with each other.
  - Lemma: A finite Markov chain is irreducible if and only if its graph representation is a strongly connected graph



Example

- Consider a Markov chain consisting of the four states 0, 1, 2, 3 and having transition probability matrix

$$\mathbf{P} = \begin{Vmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 1 \end{Vmatrix}$$

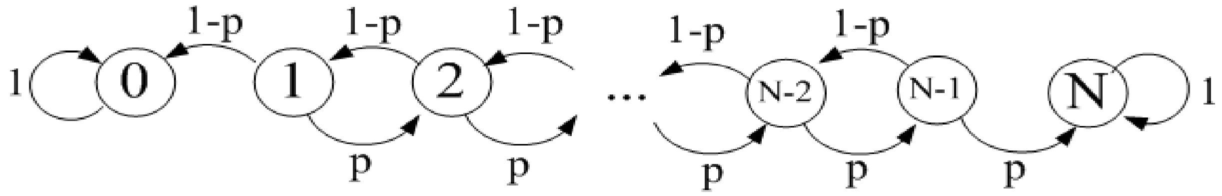


- The classes of this Markov chain are {0, 1}, {2}, and {3}.

- *Recurrent* and *transient*
  - ○ For any state i we let f_i denote the probability that, starting in state i, the process will ever re-enter state i.
    - If $f_i = 1$, state i is said to be *recurrent*
    - if $f_i < 1$, state i is said to be *transient*
  - ○ If state i is recurrent then starting in state i, the process will re-enter state i again and again and again—infinitely often.
  - ○ If state i is transient, starting in state i, the probability that the process will be in state i for exactly n time periods equals

$$f_i^{n-1}(1-f_i),\ n \geqslant 1$$

which is, the number of time periods that the process will be in state i has a geometric distribution with finite mean $1/(1 - f_i)$.



- More about recurrent:
- State i is recurrent if and only if, starting in state i, the expected number of time periods that the process is in state i is infinite.
- Theorem: State i is

$$\text{recurrent if } \sum_{n=1}^{\infty} P_{ii}^{n} = \infty,$$

$$\text{transient if } \sum_{n=1}^{\infty} P_{ii}^{n} < \infty$$

- Recurrence is a class property.
  - Corollary: If state i is recurrent, and state i communicates with state j, then state j is recurrent.
  - It also implies that transience is a class property. If state i is transient and communicates with state j, then state j must also be transient.

Example

- Consider the Markov chain having states 0, 1, 2, 3, 4 and

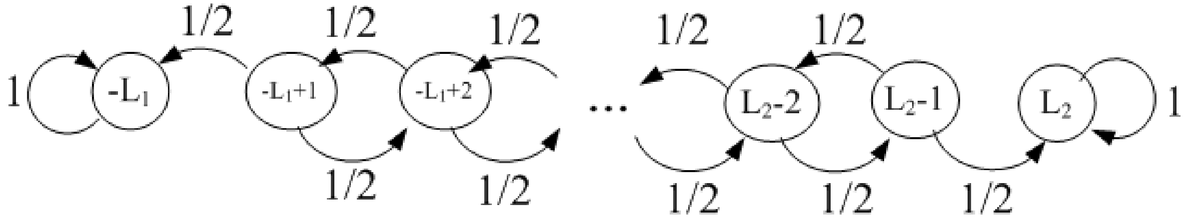
$$\mathbf{P} = \begin{Vmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & 0 & \frac{1}{2} \end{Vmatrix}$$

- Determine the recurrent state.
- Solution: This chain consists of the three classes {0, 1}, {2, 3}, and {4}. The first two classes are recurrent and the third transient.

Example: The Gambler's Ruin

- In a fair game, there are two players.
- Each round a player wins \$1 with probability 1/2, or loses \$1 with probability 1/2
- The state of the system at time t is the number of dollars won by player 1. The initial state is 0.
- Let L_1 (L_2) be the money player 1 (player 2) have
- The game ends when it reached on of the two states $-L_1$ or L_2. At this point, one player is ruined.
- What is the probability that player 1 wins L_2 dollars before losing L_1 dollars?

- We can formulate the problem as a Markov chain with two absorbing, recurrent states.



- Clearly $-L_1$ and L_2 are recurrent states, and all other states are transient.
- Let P_i^t be the probability that after t steps, the chain is at state i
- For $-L_1<i<L_2$, state i is transient and so

$$\lim_{t\to\infty} P_i^t = 0$$

- Let q be the probability that the game ends with player 1 winning L_2 dollars, so the chain was absorbed into state L_2 (Similar for state $-L_1$)

$$\lim_{t\to\infty} P^t_{L_2} = q, \lim_{t\to\infty} P^t_{-L_1} = 1 - q,$$

- For a fair game, the expected gain in each step is 0. Let W^t be the gain player 1 after t steps. Then $E[W^t]=0$ for any t.
- Thus $$E[W^t] = \sum_{i=-L_1}^{L_2} iP^t_i = 0$$
- And $$\lim_{t\to\infty} E[W^t] = L_2 q - L_1(1-q) = 0$$
- We have $$q = \frac{L_1}{L_1 + L_2}$$
- Thus, the probability of winning (or losing) is proportional to the amount of money a player is willing to lose (or win).

Limiting Probabilities

- Observation

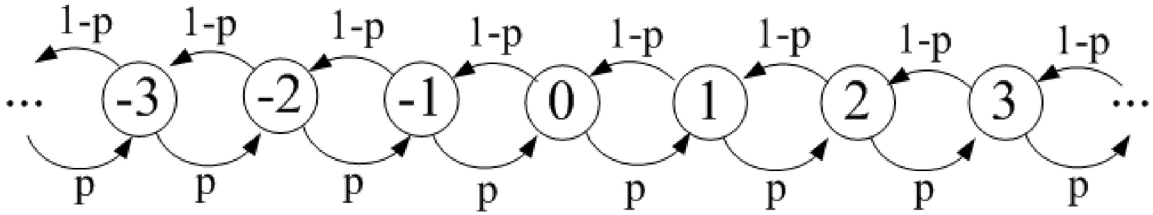
$$\mathbf{P} = \begin{Vmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{Vmatrix} \qquad \mathbf{P}^{(4)} = \begin{Vmatrix} 0.5749 & 0.4251 \\ 0.5668 & 0.4332 \end{Vmatrix}$$

$$\mathbf{P}^{(8)} = \begin{Vmatrix} 0.572 & 0.428 \\ 0.570 & 0.430 \end{Vmatrix}$$

- Each of the rows has almost identical entries
- P_{ij}^{n} is converging to some value (as n →∞) that is the same for all i
- There seems to exist a limiting probability that the process will be in state j after a large number of transitions, and this value is independent of the initial state

More Properties

- *Period*
  - State i is said to have period d if whenever n is not divisible by d, $P_{ii}^{n} = 0$, and d is the largest integer with this property.
  - For instance, in the random walk case, it has period 2, since for any n if odd $P_{ii}^{n} = 0$



- *Periodic* and *aperiodic*
  - A state with period 1 is said to be *aperiodic*, otherwise, it is *periodic*
  - Periodicity is a class property

- *Positive recurrent*
  - If state i is recurrent, and starting from i, the expected time until the process returns to state i is finite, then it is said to be *positive recurrent*
  - In a finite-state Markov chain all recurrent states are positive recurrent
- *Ergodic*
  - Positive recurrent, aperiodic states are called *ergodic*

Stationary Distribution

- Theorem
  - For an irreducible ergodic Markov chain $\lim_{n\to\infty} P_{ij}^n$ exists and is independent of i.
  - Let $\pi_j = \lim_{n\to\infty} P_{ij}^n, \quad j \geqslant 0$
  - Then π_j is the unique nonnegative solution of

$$\pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij}, \quad j \geqslant 0,$$

$$\sum_{j=0}^{\infty} \pi_j = 1$$

  - In another word, let $\pi = (\pi_1, \cdots, \pi_n)$
  - We can obtain the stationary distribution by solving

$$\pi = \pi P$$

$$\sum_i \pi_i = 1$$

  - π_j will also equal the long-run proportion of time that the Markov chain is in state j.

Example

- Consider the Markov chain

$$\mathbf{P} = \begin{Vmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{Vmatrix}$$

- In the long run, what proportion of time is the process in each of the three states?
- Solution

$$\begin{aligned} \pi_0 &= 0.5\pi_0 + 0.3\pi_1 + 0.2\pi_2, \\ \pi_1 &= 0.4\pi_0 + 0.4\pi_1 + 0.3\pi_2, \\ \pi_2 &= 0.1\pi_0 + 0.3\pi_1 + 0.5\pi_2, \end{aligned}$$

$$\pi_0 + \pi_1 + \pi_2 = 1$$

- Solving yields

$$\pi_0 = \frac{21}{62}, \quad \pi_1 = \frac{23}{62}, \quad \pi_2 = \frac{18}{62}$$

Example: A Model of Class Mobility

- A problem of interest to sociologists is to determine the proportion of society that has an upper-, middle- or lower-class occupation.
- In a family, we assume that the occupation of a child depends only on his or her parent’s occupation.
- The transition probability matrix is given by

$$\mathbf{P} = \begin{Vmatrix} 0.45 & 0.48 & 0.07 \\ 0.05 & 0.70 & 0.25 \\ 0.01 & 0.50 & 0.49 \end{Vmatrix}$$

- For instance, we suppose that the child of a middle-class worker will attain an upper-, middle-, or lower-class occupation with respective probabilities 0.05, 0.70, 0.25.
- What is the proportion of society in the long term?

- Solution

The limiting probabilities π_i thus satisfy

$$\pi_0 = 0.45\pi_0 + 0.05\pi_1 + 0.01\pi_2,$$
$$\pi_1 = 0.48\pi_0 + 0.70\pi_1 + 0.50\pi_2,$$
$$\pi_2 = 0.07\pi_0 + 0.25\pi_1 + 0.49\pi_2,$$
$$\pi_0 + \pi_1 + \pi_2 = 1$$

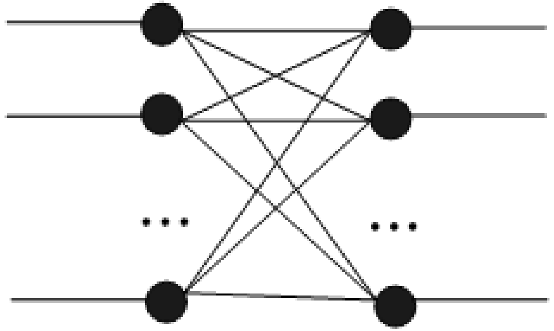
Hence,

$$\pi_0 = 0.07, \quad \pi_1 = 0.62, \quad \pi_2 = 0.31$$

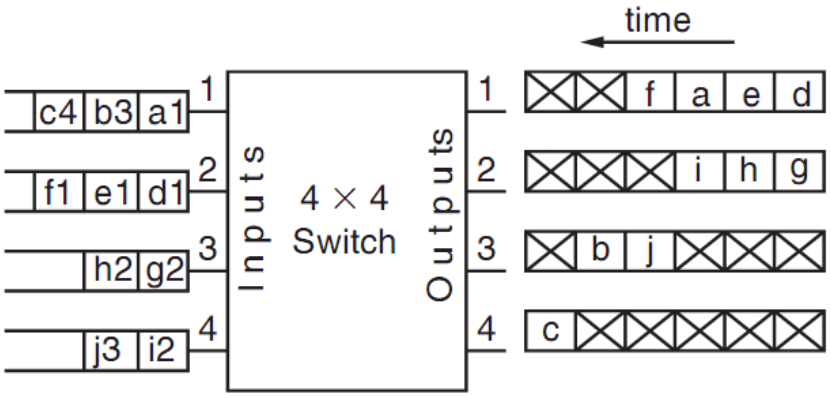
- So in the long run, 7 percent of its people in upper-class jobs, 62 percent of its people in middle-class jobs, and 31 percent in lower-class jobs.

Example: Achievable throughput in a Input-Queueing packet Switch

- High-speed packet switch is designed to handle fixed-length packets (cells)
- Consider such an N × N cell switch having N input ports and N output ports
- Time is slotted, and switching is synchronized to the slot boundaries
- The transmission rates on the input and output links be equal
- In each slot, up to one cell can be received on each of the N input links, and the cell can have any of the N output links as its destination.
- Destination conflicts: more than one cell with the same destination arrive at the inputs in a slot
- If destination conflicts happen, only one cell can be transmitted to the destination link



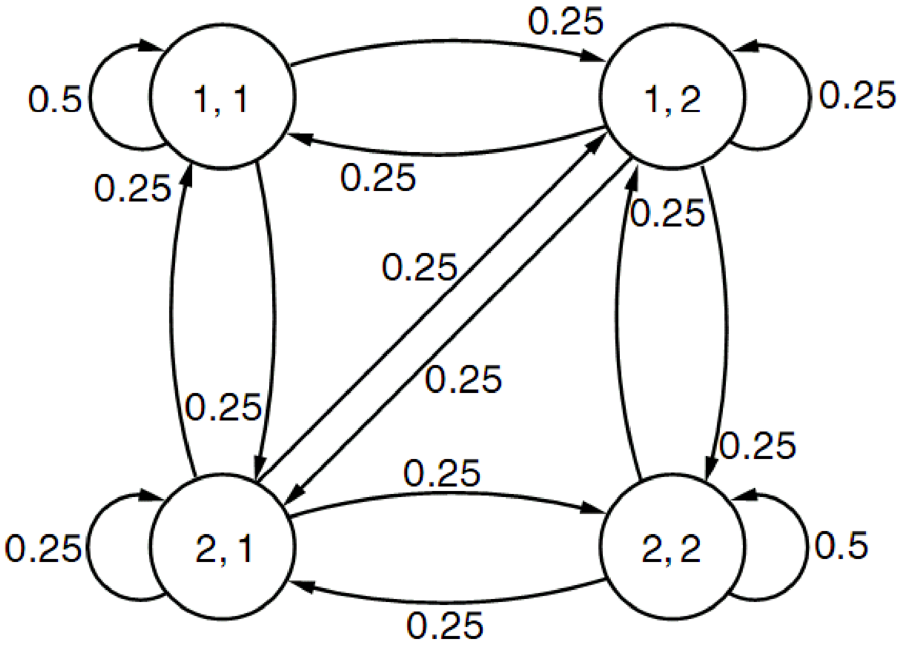
- Input-queued (IQ) switch
  - There is a first-in-first-out (FIFO) queue maintained at each input and an arriving cell is placed at the tail of this queue
  - The switch is capable of switching one cell to each output and one cell from each input
  - The *head-of-line (HOL) blocking* problem affects its achievable throughput
    - E.g., in the figure, although output 4 is free, packet c is blocked by its HOL packets, a and b, which are waiting for their outputs to become free.
  - Question: how low is this reduction in the throughput?



Analysis

- Consider an (N × N) IQ switch operating under the extreme condition of input saturation: Every time a cell is switched from the input to its output, it is immediately replaced by another cell.
- Each input cell chooses output port randomly and independently
- In the case of destination conflict, one of the cells is chosen randomly to be switched

- Consider the case for N = 2
  - Let $d_n^{(i)}$ be the destination of the cell in input i during slot n.
  - $d_n := [d_n^{(1)},\ d_n^{(2)}]$ can be called the state of the switch at the beginning of slot n
  - There are four possible states that the switch can be in: [1, 1], [1, 2], [2, 1], and [2, 2]
  - The next state only depends on the current state
  - The switch can be modeled as a four-state discrete time Markov chain



- Solving the Markov chain, we get the stationary probability of the switch being in each of the four states to be 0. 25.
- To obtain the throughput, we have
  - If the system is in state [1, 1] or [2, 2], only one of the cells can be switched to the output, and the throughput is 0.5 cell per port
  - If the switch is in either [1, 2] or [2, 1], both cells can be switched to their output and the throughput is 1 cell per port
- Thus the stationary saturation throughput is

$(0.25 \times 0.5 + 0.25 \times 0.5 + 0.25 \times 1.0 + 0.25 \times 1.0) = 0.75$ cells per port

- For larger N, the analysis is complicated since the number of states in the Markov chain will be N^N

Homework

- Write a computer simulation program to verify the result of the Gambler’s Ruin problem
- Paper reading
    - Simon Heimlicher and Kavé Salamatian, Globs in the Primordial Soup - The Emergence of Connected Crowds in Mobile Wireless Networks, Mobihoc 2010 (Best Paper)
    - Anthony J. Nicholson and Brian D. Noble, BreadCrumbs: Forecasting Mobile Connectivity, Mobicom 2008