Information Theory

Mohamed Hamada

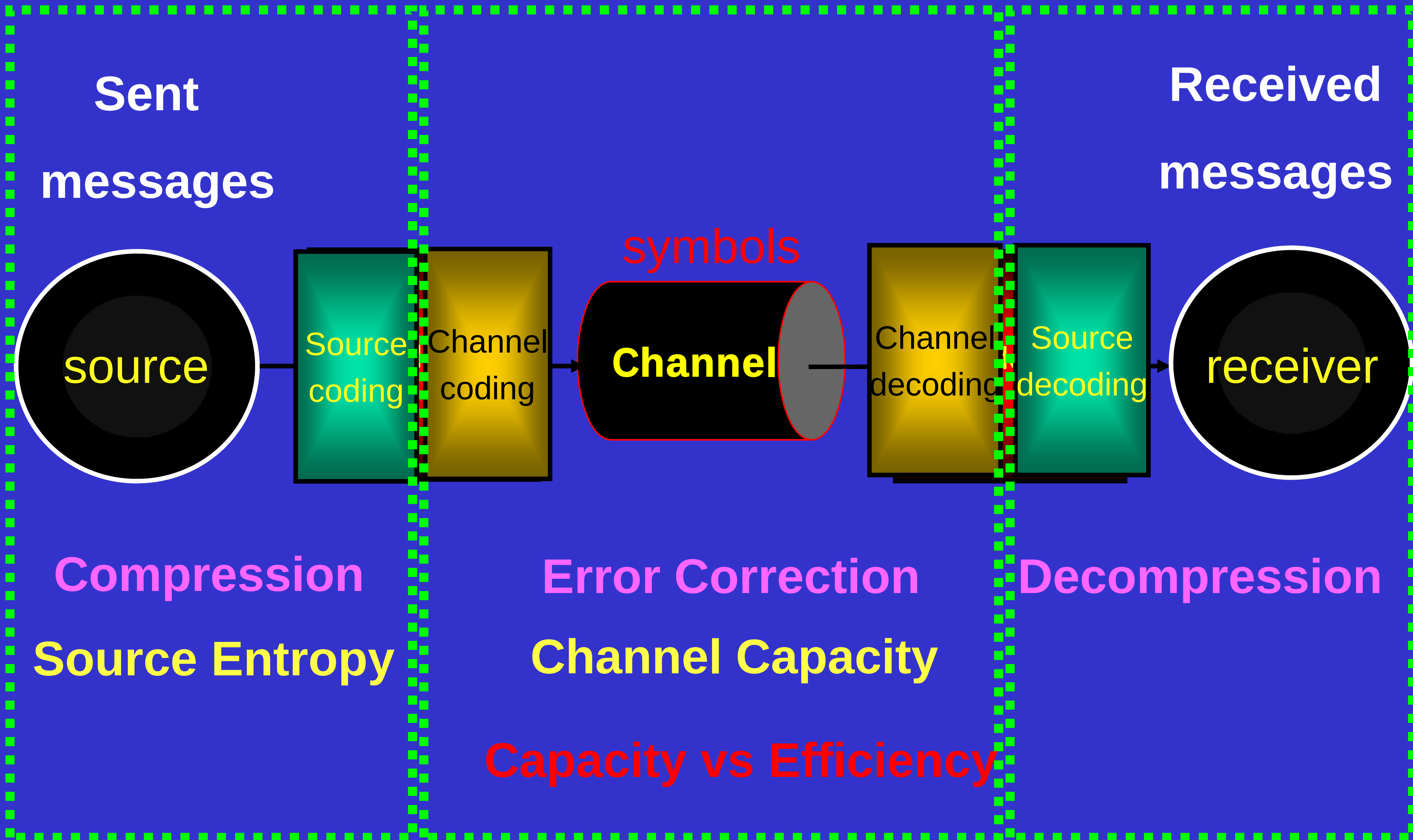
Software Engineering Lab
The University of Aizu

Email: hamada@u-aizu.ac.jp
URL: http://www.u-aizu.ac.jp/~hamada

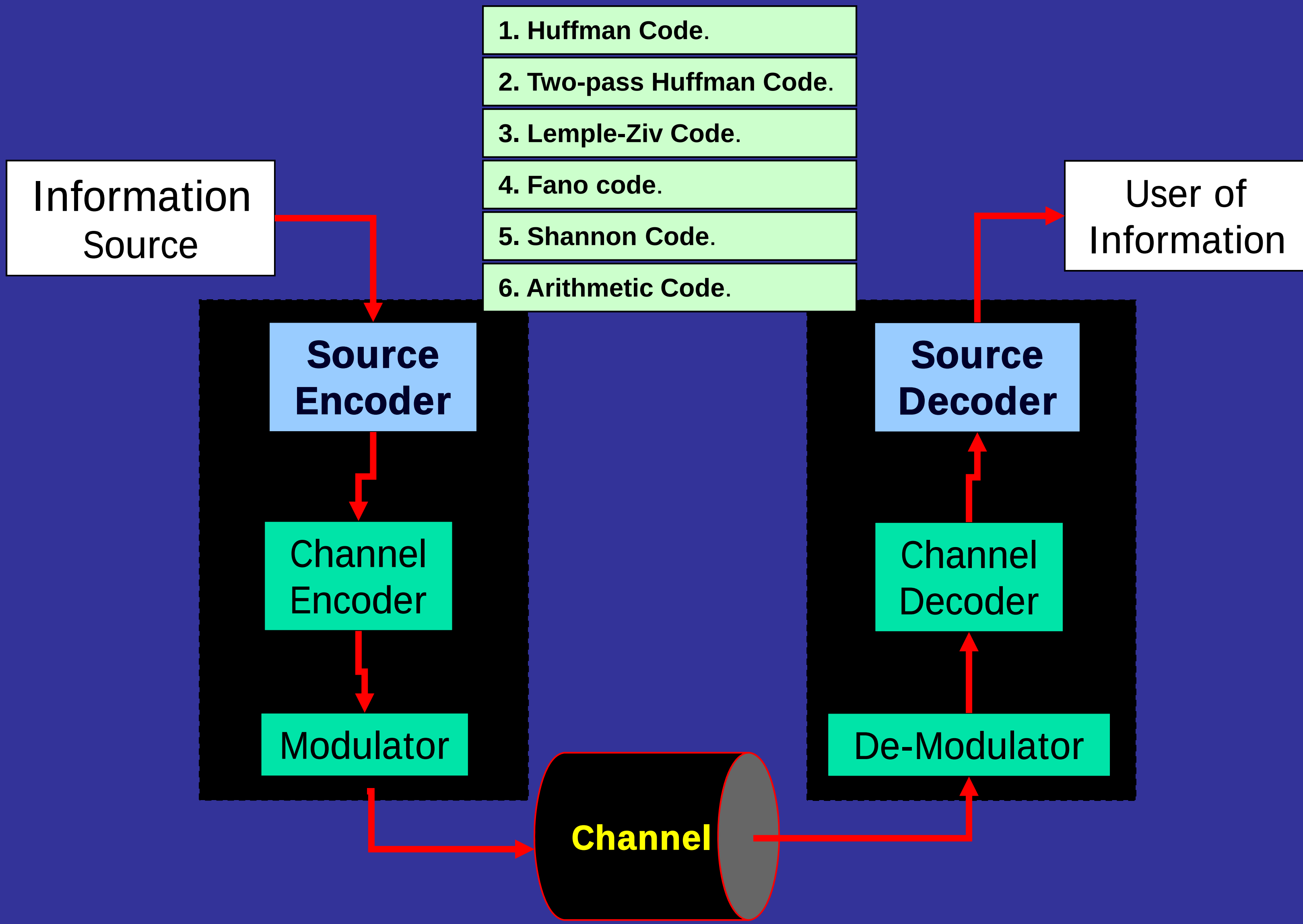
Today's Topics

- Memoryless information source.
- Information source with memory:
  - stochastic process
  - Markov Process
  - ergodic process

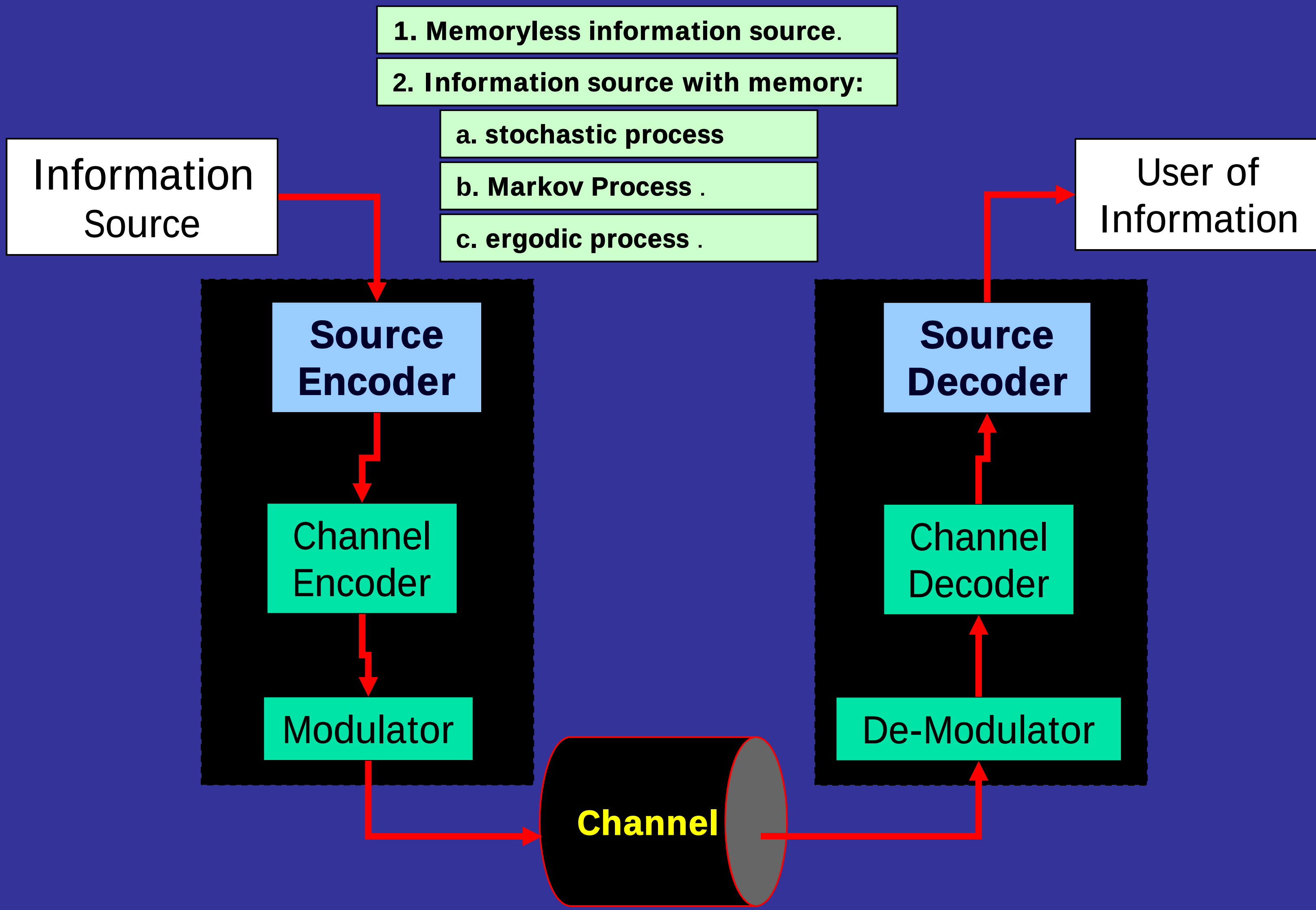
INFORMATION TRANSFER ACROSS CHANNELS



Digital Communication Systems



Digital Communication Systems



Information Source

1. Memoryless information source.

2. Information source with memory: stochastic process.

3. Information source with memory: Markov Process.

4. Information source with memory: ergodic process.

Information Source

1. Memoryless information source.

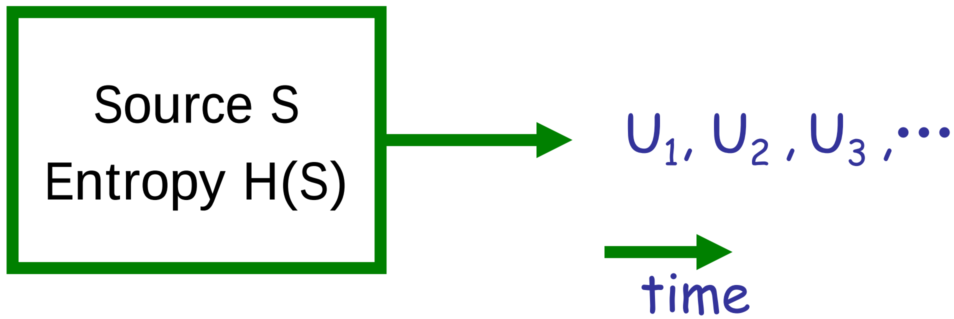
2. Information source with memory: stochastic process.

3. Information source with memory: Markov Process.

4. Information source with memory: ergodic process.

1. Memoryless information source.

Memoryless source S: generates independent symbols



- Example 1: throw dice 600 times, what do you expect?
- Example 2: throw coin 100 times, what do you expect?

The output of the current experment does not depend on the output of any of the previus expermints

$$P(x=a|x1x2.....xn) = P(x=a)$$

Information Source

1. Memoryless information source.

2. Information source with memory: stochastic process.

3. Information source with memory: Markov Process.

4. Information source with memory: ergodic process.

2. Information source with memory: stochastic process.

- **Stochastic process: is any sequence of random variables from some probability space.**

- **According to the dictionary "Stochastic" comes from a Greek word meaning "to aim at" .**

- **Stochastic, from the Greek** "stochos" or "goal", means of, relating to, or characterized by conjecture and randomness.
  A **Stochastic process is one whose behavior is non-deterministic** in that the next state of the environment is partially but not fully determined by the previous state of the environment.

- **Example: count of thrown a die.**

Information Source

1. Memoryless information source.

2. Information source with memory: stochastic process.

3. **Information source with memory: Markov Process.**

4. Information source with memory: ergodic process.

Andrei Andreyevich Markov

Markov is particularly remembered for his study of Markov chains, sequences of random variables in which the future variable is determined by the present variable but is independent of the way in which the present state arose from its predecessors. This work launched the theory of stochastic processes.

Born: 14 June 1856 in Ryazan, Russia
Died: 20 July 1922 in Petrograd (now St Petersburg), Russia

Sources with memory

General source: |S| states

$S \in \{1, 2, \bullet\bullet\bullet, |S|\}$

Output

$U_1, U_2, \bullet\bullet\bullet, U_L$ subscript = time

Markov source: |S| states

New state = f(old state, output)

$S \in \{1, 2, \bullet\bullet\bullet, |S|\}$

Output

$U_1, U_2, \bullet\bullet\bullet, U_L$ subscript = time

Markov information source (Process) (Chain)

It is an information source with memory in which the probability of a symbol occurring in a message will depend on a finite number of preceding symbols

Markov process is (a simple form of) dependent stochastic process in which the probability of an outcome of a trial depends on the outcome of the immediately preceding trial, i.e.

$$P(X_n = j \mid X_1, X_2, \ldots X_{n-1}) = P(X_n = j \mid X_{n-1})$$

Markov information source (Process) (Chain)

Let an experiment has a finite number of n possible outcomes

$A = \{ a_1, ..., a_n \}$ called *states*.

For a Markov process we specify a table of probabilistic associated with transitions from any state to another.

This is called *probability transition matrix T*.

For a Markov information source S = (A, T), T has the form:

$$
T = \begin{array}{c} \\ a_1 \\ a_2 \\ \\ a_n \end{array}
\begin{array}{c} \begin{array}{cccc} a_1 & a_2 & \dots\dots\dots & a_n \end{array} \\
\begin{pmatrix}
p(a_1|a_1) & p(a_2|a_1) & \dots\dots\dots & p(a_n|a_1) \\
p(a_1|a_2) & p(a_2|a_2) & \dots\dots\dots & p(a_n|a_2) \\
 & & .\ .\ .\ . & \\
p(a_1|a_n) & p(a_2|a_n) & \dots\dots\dots & p(a_n|a_n)
\end{pmatrix} \end{array}
$$

Transition Diagram

Shannon Diagram:

Each node represents the currents information source output

Example : **Consider the *Nikkei* stock index at certain day has the following information source:**

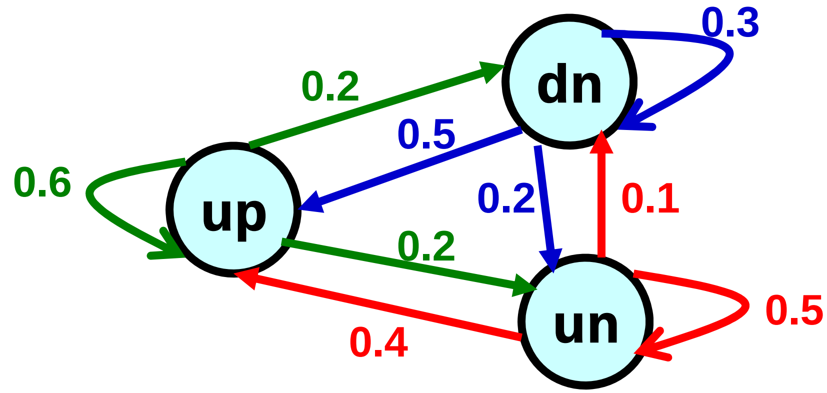
Index	up	down	unchanged
Probability	0.5	0.2	0.3

With the transition matrix T:

$$T = \begin{array}{c} \\ up \\ dn \\ un \end{array} \begin{array}{ccc} up & dn & un \\ \hline \end{array}$$

	up	dn	un
up	0.6	0.2	0.2
dn	0.5	0.3	0.2
un	0.4	0.1	0.5

The transition diagram is:



Transition Diagram

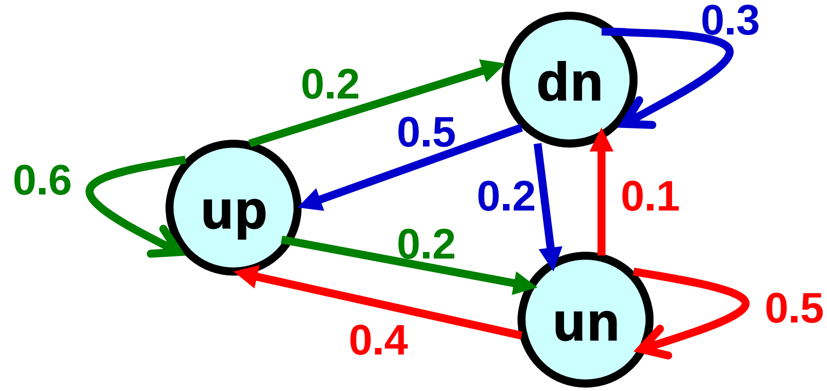
Example : **Consider the *Nikkei* stock index at certain day has the following information source:**

Index	up	down	unchanged
Probability	0.5	0.2	0.3

With the transition matrix T:

$$T = \begin{array}{c} \\ up \\ dn \\ un \end{array} \begin{array}{c} \begin{array}{ccc} up & dn & un \end{array} \\ \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.5 & 0.3 & 0.2 \\ 0.4 & 0.1 & 0.5 \end{pmatrix} \end{array}$$

The transition diagram is:



- **What is the probability of 5 consecutive up days?**

Sequence is up-up-up-up-up

P(up,up,up,up,up) = $0.5 \times (0.6)^4 = 0.0648$

Transition Diagram

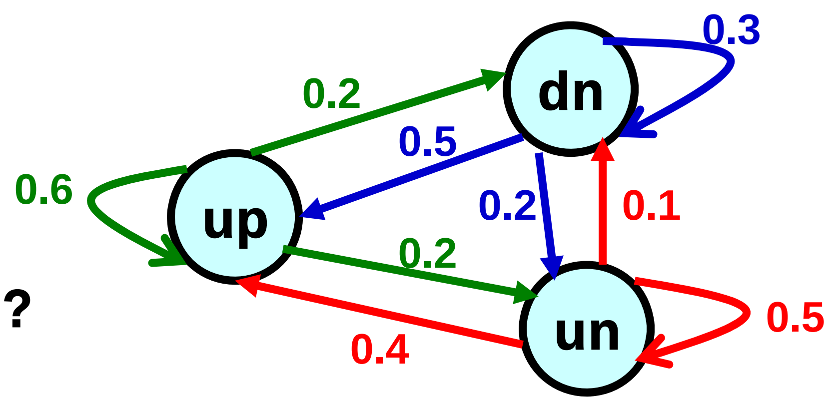
Example : **Consider the *Nikkei* stock index at certain day has the following information source:**

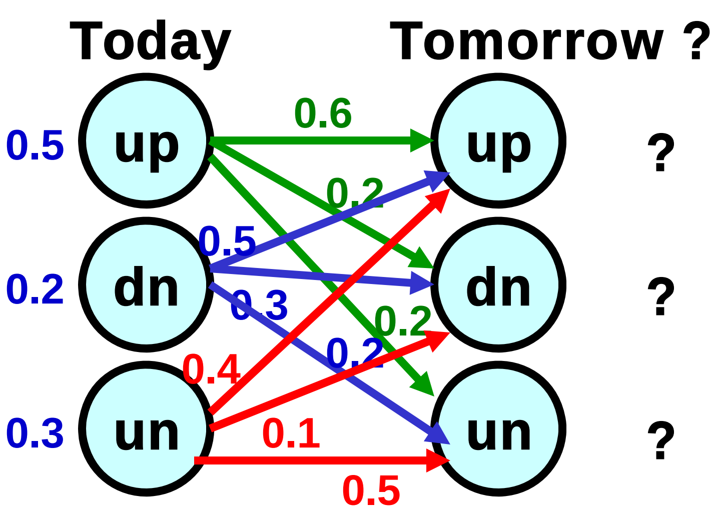
Index	up	down	unchanged
Probability	0.5	0.2	0.3

With the transition matrix T:

$$
\begin{array}{c} \\ up \\ dn \\ un \end{array}
\begin{array}{c} \begin{array}{ccc} up & dn & un \end{array} \\ \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.5 & 0.3 & 0.2 \\ 0.4 & 0.1 & 0.5 \end{pmatrix} \end{array}
$$

The transition diagram is:

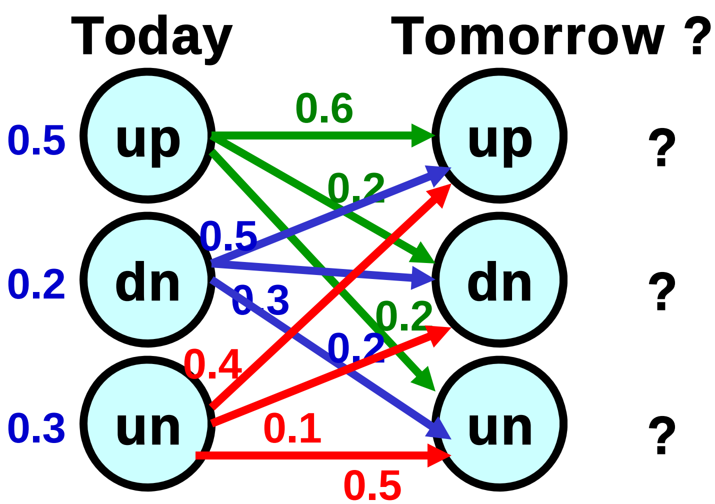


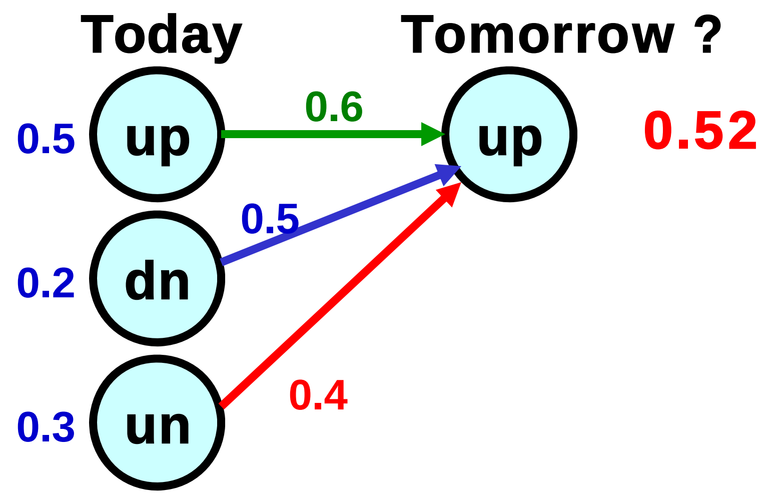


Transition Diagram

Example : **Consider the *Nikkei* stock index at certain day has the following information source:**

Index	up	down	unchanged
Probability	0.5	0.2	0.3





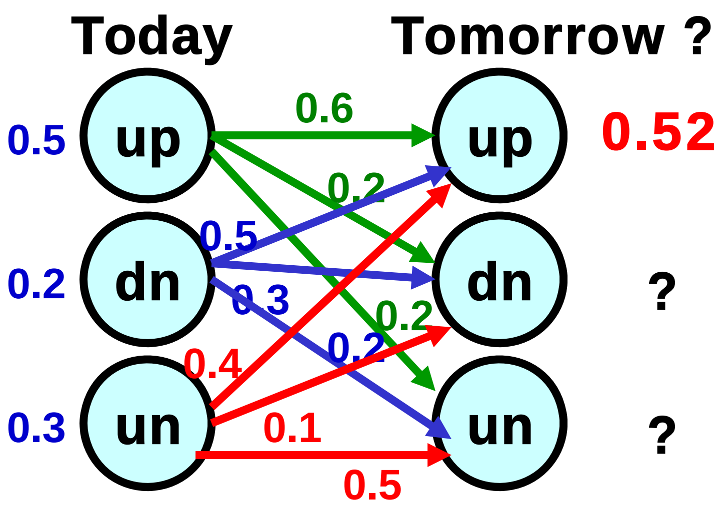
Tomorrow: p(up) = 0.5 x 0.6 + 0.2 x 0.5 + 0.3 x 0.4

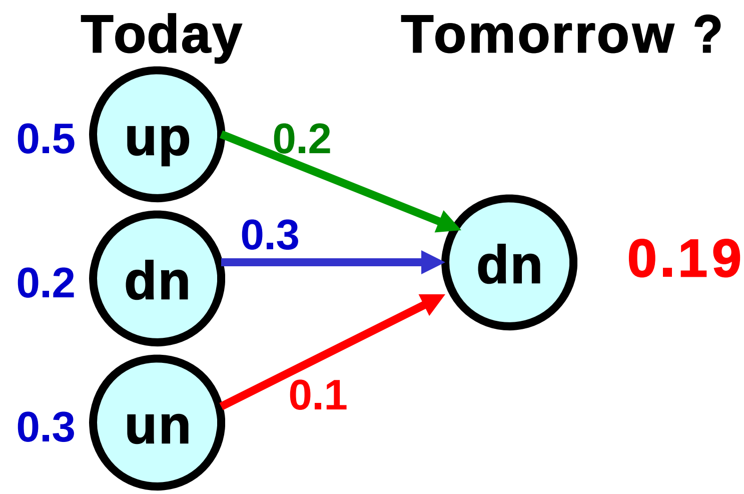
= **0.52**

Transition Diagram

Example : **Consider the *Nikkei* stock index at certain day has the following information source:**

Index	up	down	unchanged
Probability	0.5	0.2	0.3





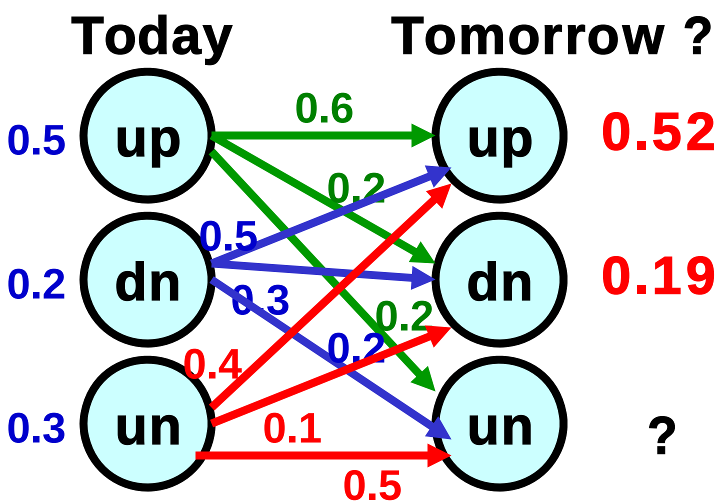
Tomorrow: p(dn) = 0.5 x 0.2 + 0.2 x 0.3 + 0.3 x 0.1

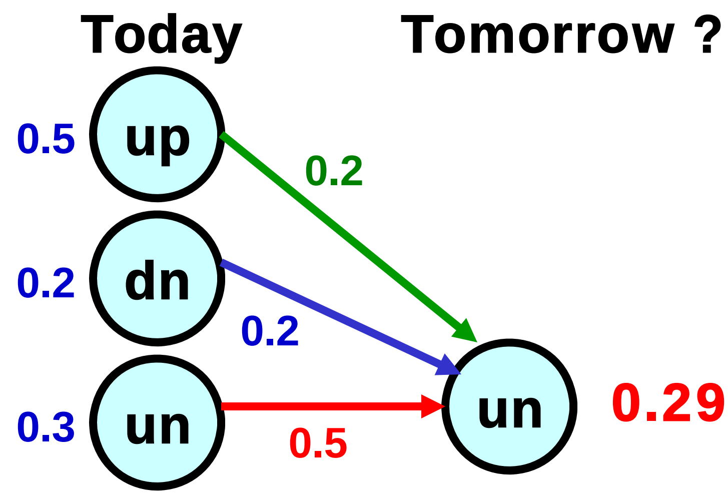
= **0.19**

Transition Diagram

Example : **Consider the *Nikkei* stock index at certain day has the following information source:**

Index	up	down	unchanged
Probability	0.5	0.2	0.3





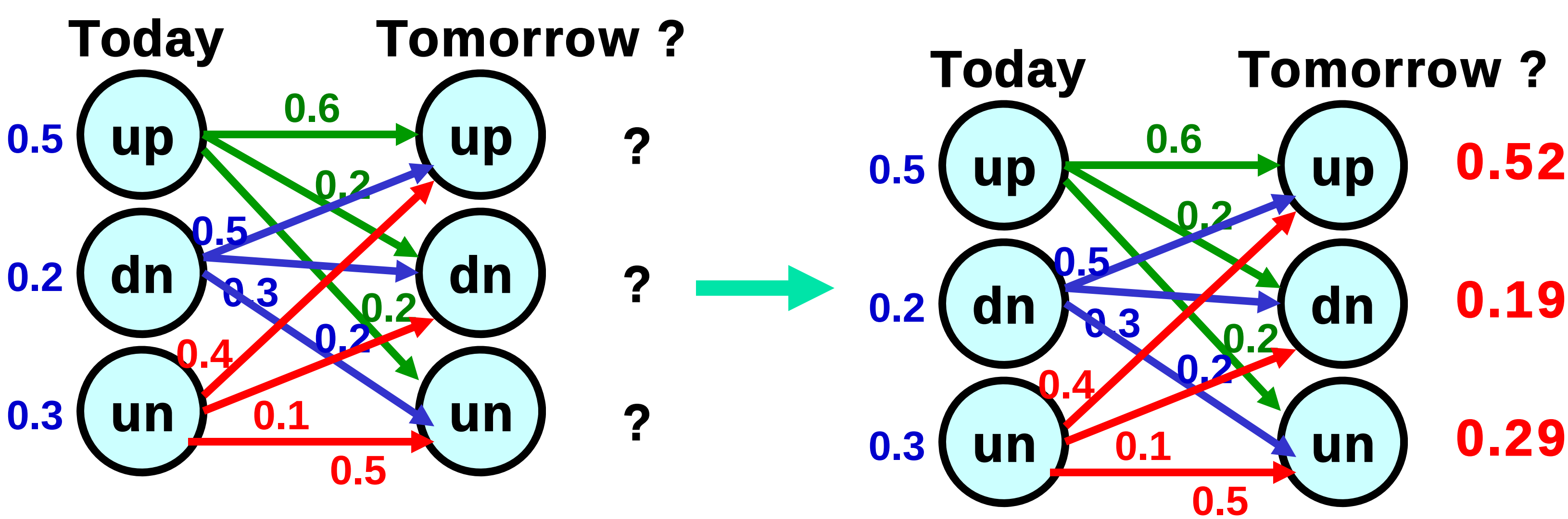
Tomorrow: p(un) $= 0.5 \times 0.2 + 0.2 \times 0.2 + 0.3 \times 0.5$

$= 0.29$

Transition Diagram

Example : **Consider the *Nikkei* stock index at certain day has the following information source:**

Index	up	down	unchanged
Probability	0.5	0.2	0.3

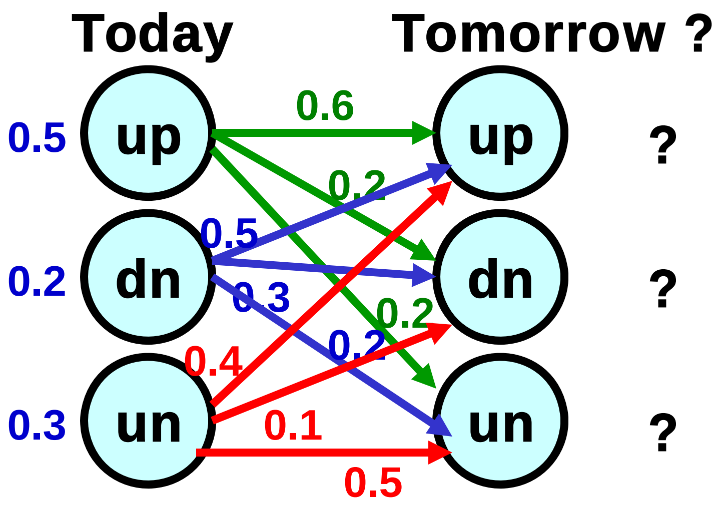


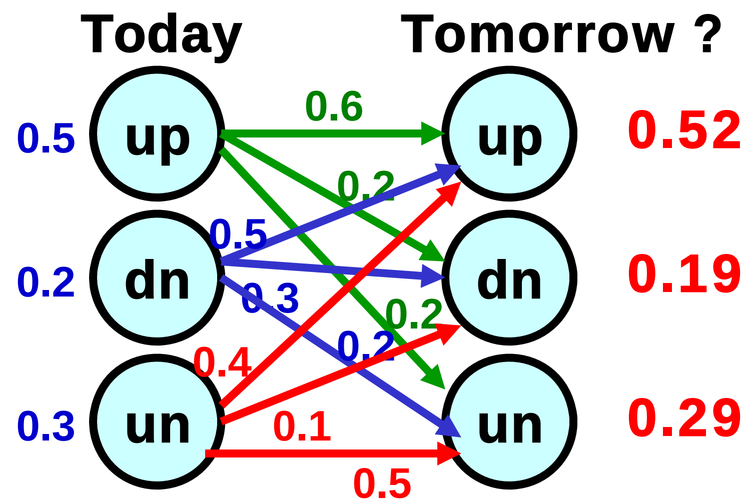
Tomorrow: p(up) = 0.5 x 0.6 + 0.2 x 0.5 + 0.3 x 0.4 = **0.52**

Tomorrow: p(dn) = 0.5 x 0.2 + 0.2 x 0.3 + 0.3 x 0.1 = **0.19**

Tomorrow: p(un) = 0.5 x 0.2 + 0.2 x 0.2 + 0.3 x 0.5 = **0.29**

Transition Diagram





Tomorrow: p(up) = 0.5 x 0.6 + 0.2 x 0.5 + 0.3 x 0.4 = 0.52

Tomorrow: p(dn) = 0.5 x 0.2 + 0.2 x 0.3 + 0.3 x 0.1 = 0.19

Tomorrow: p(un) = 0.5 x 0.2 + 0.2 x 0.2 + 0.3 x 0.5 = 0.29

***Nikkei* Tomorrow:** **We can write it as:**

$$(p(up), p(dn), p(un)) = (0.5, 0.2, 0.3) \times \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.5 & 0.3 & 0.2 \\ 0.4 & 0.1 & 0.5 \end{pmatrix}$$

Initial probability

Transition matrix

Transition Diagram

Tomorrow: p(up) = 0.5 x 0.6 + 0.2 x 0.5 + 0.3 x 0.4

Tomorrow: p(dn) = 0.5 x 0.2 + 0.2 x 0.3 + 0.3 x 0.1

Tomorrow: p(un) = 0.5 x 0.2 + 0.2 x 0.2 + 0.3 x 0.5

Nikkei **Tomorrow:**

$$P^1 = (p(up), p(dn), p(un)) = (0.5, 0.2, 0.3) \times \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.5 & 0.3 & 0.2 \\ 0.4 & 0.1 & 0.5 \end{pmatrix}$$

Put the initial probability as P^0, then the probability of tomorrow (i.e. after one-step) P^1 can by found by:

$$P^1 = P^0 \times T$$

where T is the transition matrix.

In General, the probability after r-setps (I.e. after r days in the Nikkei example) can be found by the formula:

$$P^r = P^0 \times T^r$$

Transition Diagram

Example : **Consider the *Nikkei* stock index at certain day has the following information source:**

Index	up	down	unchanged
Probability	0.5	0.2	0.3

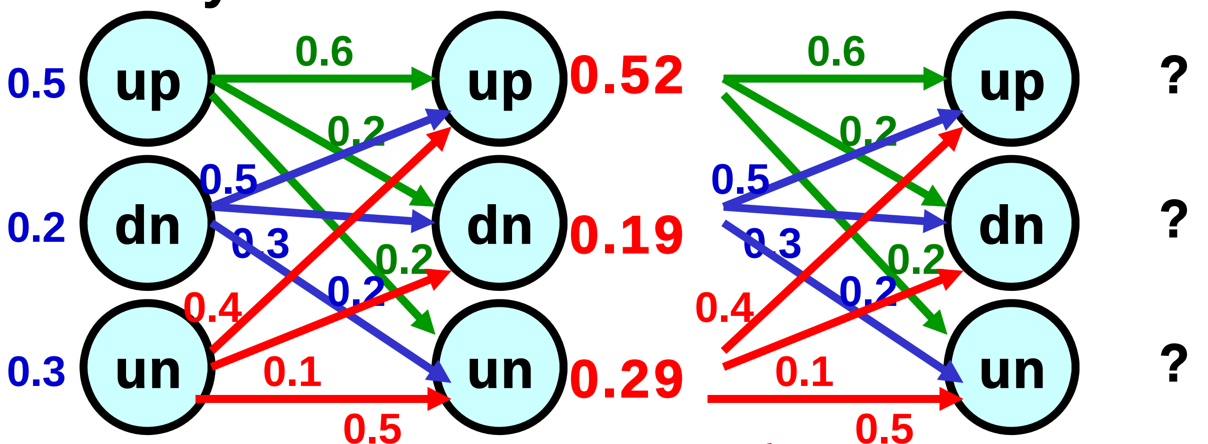
P^0

$$T = \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.5 & 0.3 & 0.2 \\ 0.4 & 0.1 & 0.5 \end{pmatrix}$$

T

Find the probability of the index after 2 days?

Today **Tomorrow** **After Tomorrow ?**



$$P^r = P^0 \times T^r$$

$$P^2 = P^0 \times T^2 = (0.5, 0.2, 0.3) \times \left\{ \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.5 & 0.3 & 0.2 \\ 0.4 & 0.1 & 0.5 \end{pmatrix} \times \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.5 & 0.3 & 0.2 \\ 0.4 & 0.1 & 0.5 \end{pmatrix} \right\}$$

Information Source

1. Memoryless information source.

2. Information source with memory: stochastic process.

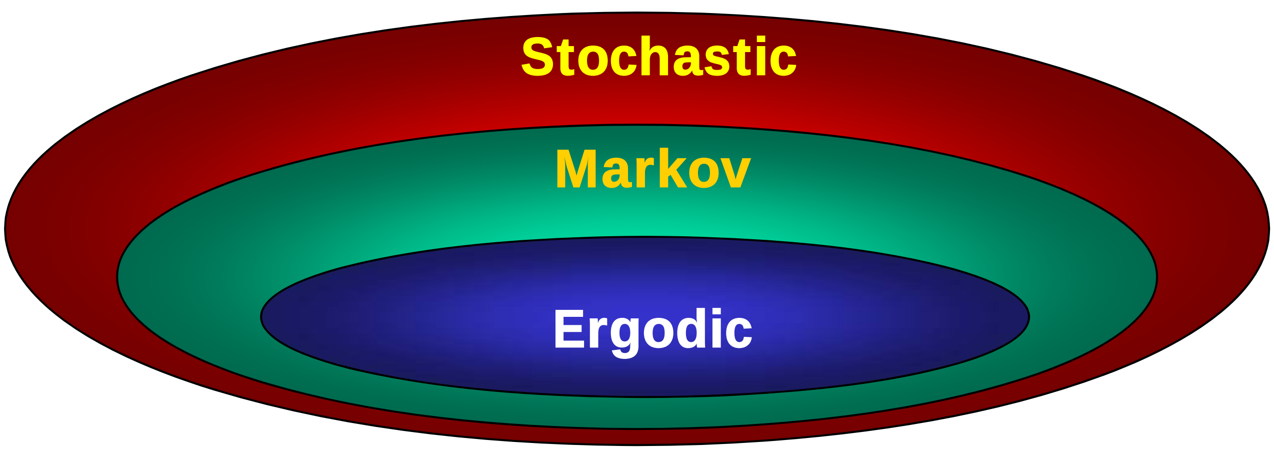
3. Information source with memory: Markov Process.

4. **Information source with memory: ergodic process.**

Ergodic Markov Process

A Markov chain is said to be *ergodic* if, after a certain finite number of steps, it is possible to go from any state to any other state with a nonzero probability.

The following figure shows the relationship between processes:



Ergodic Markov Process

To check that some Markov process is ergodic:

1. We check the transition diagram for this process

2. if, from any state, we can reach all other states, then the process is ergodic

3. if some state(s) are not reachable from any other state, then the process is NOT ergodic

Ergodic Markov Process

Example 1:

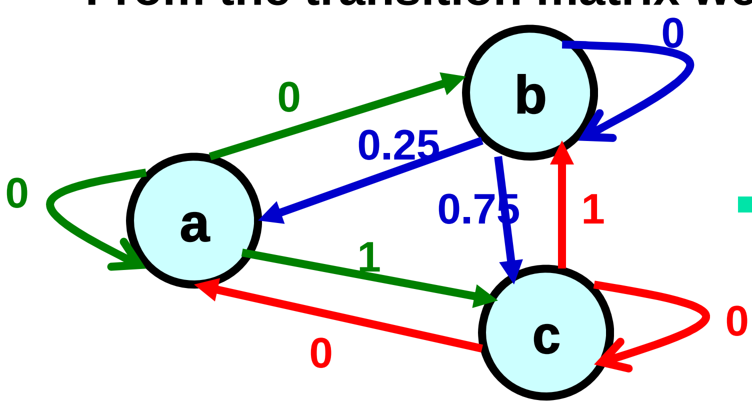
Consider the process with the following transition matrix

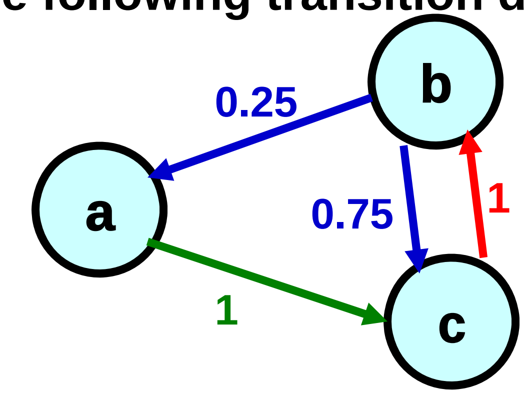
$$T = \begin{array}{c} \\ a \\ b \\ c \end{array} \begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{pmatrix} 0 & 0 & 1 \\ 0.25 & 0 & 0.75 \\ 0 & 1 & 0 \end{pmatrix} \end{array}$$

Is this an ergodic process?

Answer

From the transition matrix we can build the following transition diagram:





From a we can go b and c
From b we can go a and c
From c we can go a and b

The process is ergodic

Ergodic Markov Process

Example 2:

Consider the process with the following transition matrix

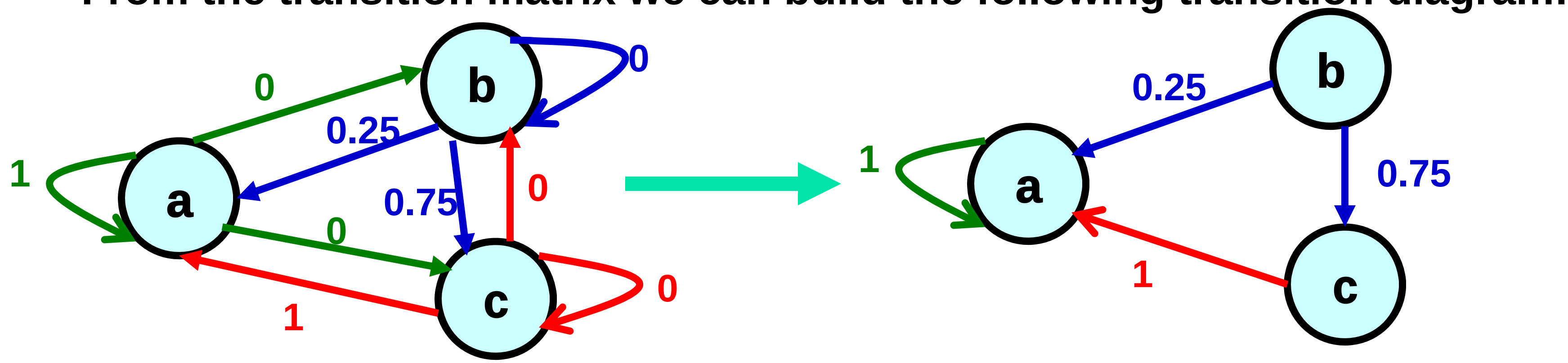
T =

	a	b	c
a	1	0	0
b	0.25	0	0.75
c	1	0	0

Is this an ergodic process?

Answer

From the transition matrix we can build the following transition diagram:



From a we can NOT go to b or c
From b we can go a and c
From c we can go only a

The process is NOT ergodic