ISLAMIC UNIVERSITY OF TECHNOLOGY

Organization of Islamic Cooperation

Board Bazar, Gazipur

Markov Chains

MATH 4741

1st December, 2022

**1. Applications of Markov Decision Processes (MDPs)**

**Example 1: Chatbots**

**Use Case**: Chatbots used by business on their social media platforms make use of MDPs to decide upon the most appropriate response to what the customer has asked.

**Importance**: This specific application is essential in modern society, where shopping via social media is commonplace. Having an automated bot that can respond to common questions reduces the amount of pressure that real custom service agents must face.

**Appropriateness**: MDPs are appropriate in these cases since we have clearly specified states (the possible questions the customer can ask) along with clearly specified actions that should be taken from each state (what the chatbot’s response should be). The chatbot can learn to take appropriate actions, i.e., give appropriate answers, since it will either be rewarded or punished based on the appropriateness of its answer to a given question during the training phase. The variety of questions that could be asked by customers along with the various ways in which a single question could be phrased makes MDPs more appropriate that a simple state-based system.

**Example 2: Traffic Control**

**Use Case**: MDPs can be used to control traffic lights at intersections. Depending on a variety of information, such as the amount of time a light has been red, the amount of traffic on the road that is being held by a red light, etc., the lights can be controlled.

**Importance**: The increase in traffic in densely populated cities makes traffic control an essential part of our lives. The job is difficult for humans to do manually, since the amount of information that must be considered to control traffic efficiently is too large and changes rapidly. Having an automated systems that controls the traffic based on all this information can improve the efficiency of traffic movement.

**Appropriateness**: The different traffic-related information can be combined to place the system at a specific state. Depending on the state, an action can be chosen to change the traffic lights. Since the system can be cleanly represented as a state diagram, MDPs are an appropriate way to solve the problem. The large amount of dynamic information going into such a system makes it inappropriate to have a fixed system that just changes states based on the information. A system that learns from the information, such as one trained using an MDP, will be able to do a much better job than one which is hardcoded.

**Example 3: Robots**

**Use Case**: Although robots that have general intelligence such as the one shown in [The Jetsons](https://thejetsons.fandom.com/wiki/Rosey) is still out of reach, robots that can accomplish simple tasks such as folding a towel have been created. These robots make use of MDPs.

**Importance**: The current simple version of the robots may not seem that useful, but they are an important steppingstone towards general intelligence in robots. If that is achieved, ideally, it would mean an end to menial chores for all of us.

**Appropriateness**: The main use case of MDPs and reinforcement learning is when we are unable to provide code to program a system or provide labelled data that the system can learn from. It is difficult to even describe the steps involved in a task as simple as ‘fold a towel’. This is exactly where an MDP is perfect. A variety of information regarding the robot, such as how high its arms are held, how strongly it is gripping the towel and so on can be used to create a state internally. The robot will then decide what changes to make and which new state to move to based on how it is being rewarded. The closer it gets to successfully folding the towel, the higher its reward, which is how it will learn to complete the task.

2.

Transition Matrix,

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | 0 | 1 | 2 |
|  | 0 | 0 | 1 | 0 |
| 1 | 0.25 | 0.5 | 0.25 |
| 2 | 0 | 1 | 0 |

Probability that the number of white balls in the first urn remains unchanged Probability that the system remains in the same state

3.

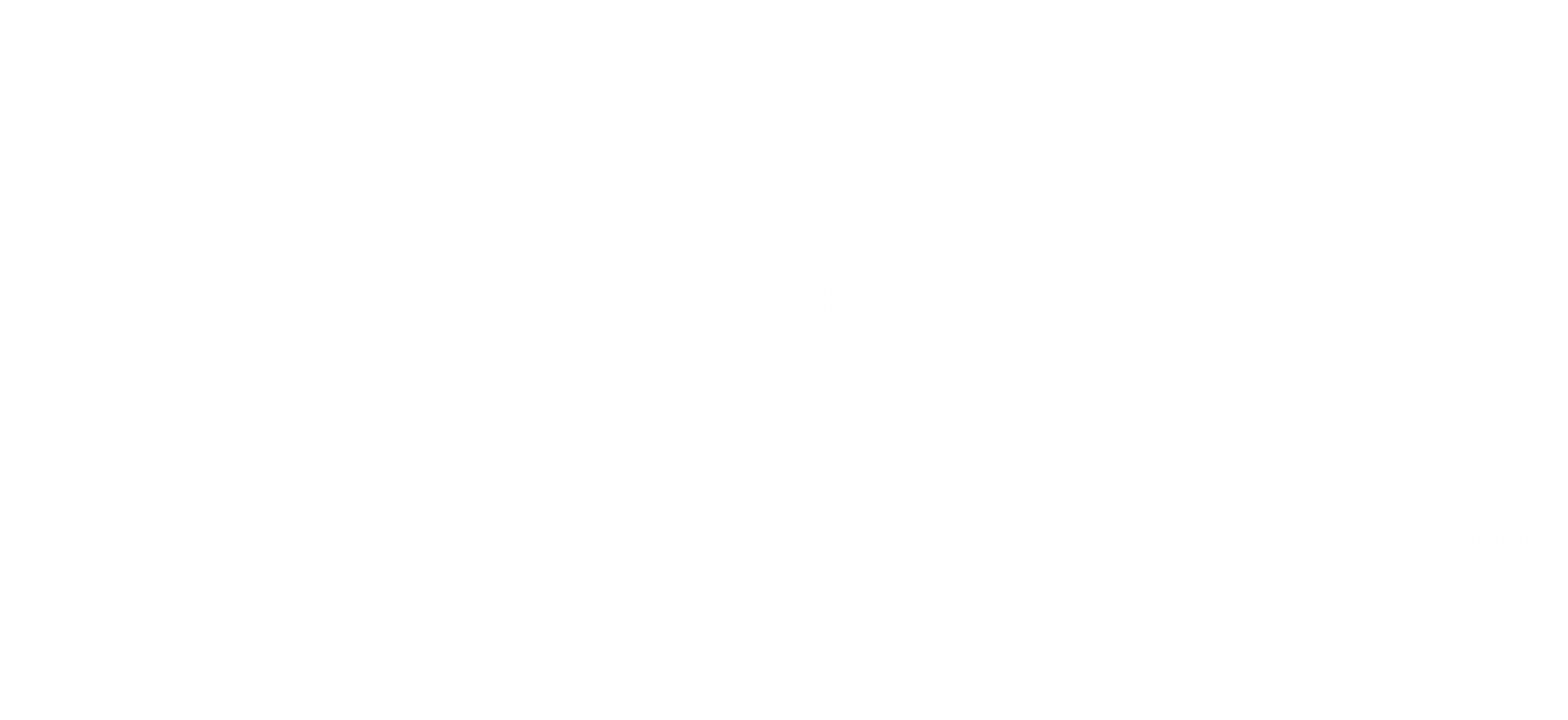
We can visualize the Markov chain as a sliding window over 3 days. If the current state shows the value of days 1, 2 and 3, the next state will show the value of days 2, 3 and 4.

The total number of states is .

Transition Matrix,

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | RRR | RRN | RNR | RNN | NRR | NRN | NNR | NNN |
|  | RRR | 0.8 | 0.2 | 0 | 0 | 0 | 0 | 0 | 0 |
| RRN | 0 | 0 | 0.4 | 0.6 | 0 | 0 | 0 | 0 |
| RNR | 0 | 0 | 0 | 0 | 0.6 | 0.4 | 0 | 0 |
| RNN | 0 | 0 | 0 | 0 | 0 | 0 | 0.4 | 0.6 |
| NRR | 0.6 | 0.4 | 0 | 0 | 0 | 0 | 0 | 0 |
| NRN | 0 | 0 | 0.4 | 0.6 | 0 | 0 | 0 | 0 |
| NNR | 0 | 0 | 0 | 0 | 0.6 | 0.4 | 0 | 0 |
| NNN | 0 | 0 | 0 | 0 | 0 | 0 | 0.2 | 0.8 |

Transition Diagram:



4.

Transition Matrix,

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | 1 | 2 |
|  | 1 | 0.7 | 0.3 |
| 2 | 0.6 | 0.4 |

where state represents the th coin being flipped.

Probability of a coin flipped on the third day after the initial flip (day 4) being coin 1 Probability of results of day 3 being heads

If the coin flipped on Monday comes up heads, we flip coin 1 on Tuesday.

Friday is 4 days after Monday.

Probability that a coin flipped on Friday being heads given that coin 1 is flipped on Tuesday

5.

Let be the probability that the next exam is of type given that the previous exam was of type .

Transition Matrix,

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | 0 | 1 | 2 |
|  | 0 | 0.8 | 0.1 | 0.1 |
| 1 | 0.6 | 0.2 | 0.2 |
| 2 | 0.4 | 0.3 | 0.3 |

Long Run Probabilities:

Ans: ths of the exams are of type 1, ths of the exams are of type 2 and ths of the exams are of type 3.

6.

Transition Matrix,

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | A | B |
|  | A | 0.6 | 0.4 |
| B | 0.3 | 0.7 |

where state is the trip being to the th zone.

Long Run Probabilities:

Let the random variable be the profit per trip.

Average profit per trip .