Markov Model for Evaluating Soccer Decision Making

Molly McCann
Advisor: Dr. Najmeh Salehi

25 April 2024

Contents

1 Introduction

Markov Decision Processes (MDPs) are mathematical models used in many fields, including inventory control, maintenance, manufacturing and telecommunication, to represent decision-making problems in a stochastic environment [1]. MDPs are named after the Russian mathematician Andrey Markov and are an extension of Markov chains. However, MDPs contain additional components that allow for decision making and rewards for actions. MDPs have various applications in the real world, including finance, agriculture, and sports. In this study, we will explore the application of MDPs in the field of soccer. Soccer is one of the world's most popular and dynamic team sports. It is characterized by its fluidity, continuous action, and the need for both individual and team strategies. On the other hand, MDPs provide a powerful framework for modeling and optimizing decision-making. MDPs offer a systematic way to capture and address the complexities of the underlying process, making it a suitable tool to analyze soccer. Therefore, MDPs can provide a valuable tool for coaches, players, and researchers seeking to enhance performance and strategic insights in the sport.

2 Markov Chains

A Markov Process, or Markov chain, is a stochastic model used to describe a sequence of events or states where the probability of transitioning from one state to another depends only on the current state and is independent of the past states. This property is called the Markov property or memory-less property. Moreover, a Markov chain consists of the following elements: states, the state space, transition matrix, and the trajectory [2]. Next, we define and discuss each of these components.

2.1 States and State-space

While working with Markov chains, let $\{X_t$, t = 0, 1, 2,...$\}$ be a stochastic process. The state of a Markov chain at time t is the value of X_t. For example, if $X_t = i$, then the process is said to be in state i at time t. The state-space, S, is the set of all possible states. To summarize, if S is a countable set, $\{i, j, k, ...\}$, each $i \in S$ is called a state and S is called the state-space. Recalling the Markov property we can say X_{t+1} depends on the current state X_t, but not $X_{t-1}, ..., X_1, X_0$.

2.2 Transition Matrix

Before defining what a transition matrix is, there are a few basic definitions that are useful to understanding its components. In Markov chains, vectors consist of positive entries that add up to 1. These vectors are called probability vectors. A square matrix whose rows are probability vectors is called a stochastic matrix. The transition matrix P is a stochastic matrix whose rows represent X_t and whose columns represent X_{t+1}. In other words, assuming that the process is currently in state i, a transition matrix shows the probability that the next state will be state j. These transition probabilities of the Markov chain are shown as:

$$p_{ij} = \mathbb{P}(X_{t+1} = j \mid X_t = i) \quad \text{for } i, j \in S, \quad t = 0, 1, 2....$$

Thus, the transition matrix itself is $P = [p_{ij}]$. It lists all the possible states in the state space, S.

Before attempting to create a transition matrix, it is useful to create a diagram that displays the probability of each state given the initial state. As seen in Figure 1, the probability that the process will remain in the same state must be considered, but like all probabilities in a Markov process, its probability does not have to be non-zero. The corresponding transition matrix for this diagram is displayed in Figure 2. Note that the possible states are all listed across the rows and columns, and the probabilities of transitioning from one state to another is within the matrix. After each iteration of the Markov process, the values of the probabilities in the transition

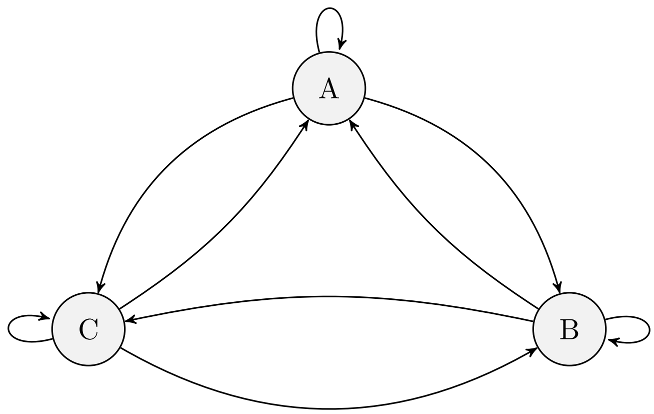

Figure 1: An example of a Markov transition diagram with 3 states.

$$\begin{array}{c} \\ A \\ B \\ C \end{array}
\begin{array}{c}
\begin{array}{ccc} \quad A \quad & \quad B \quad & \quad C \quad \end{array} \\
\begin{bmatrix}
\mathbb{P}(A \mid A) & \mathbb{P}(B \mid A) & \mathbb{P}(C \mid A) \\
\mathbb{P}(A \mid B) & \mathbb{P}(B \mid B) & \mathbb{P}(C \mid B) \\
\mathbb{P}(A \mid C) & \mathbb{P}(B \mid C) & \mathbb{P}(C \mid C)
\end{bmatrix}
\end{array}$$

Figure 2: A transition matrix with 3 possible states.

matrix change based on the previous iteration. The following example will apply the newly learned knowledge of states, state-space, and transition matrices to a soccer situation.

2.2.1 Example

The following example of a transition matrix is a small scale application to soccer based on the larger process that will later be modeled using MDPs. Since this application is on a small scale, the field has been broken down into thirds to condense the number of possible states. Figure 7 displays a diagram of the soccer pitch during the first half, operating under the condition that team 1 attempts to score in the goal on the right side of the pitch. Each team will have 4 states which means the resulting transition matrix contains 8 states total. The states are represented by the following:

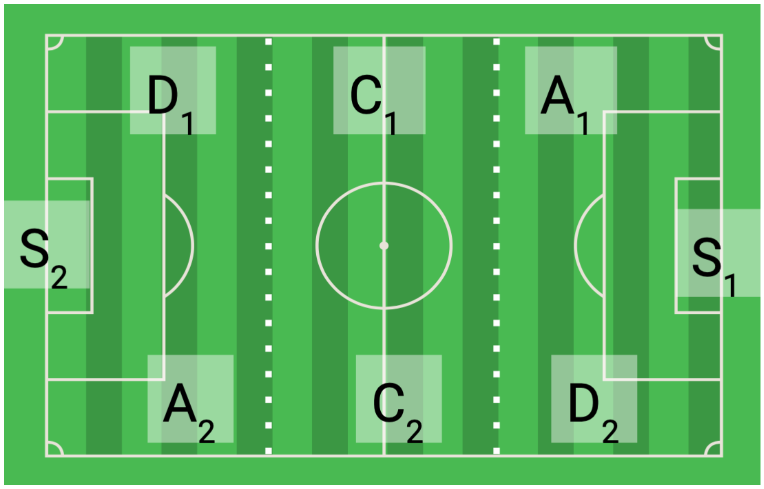

Figure 3: The zonal and possession breakdown of the pitch.

- D_1 - defensive third for team 1
- C_1 - center third for team 1
- A_1 - attacking third for team 1
- S_1 - scored goal for team 1
- D_2 - defensive third for team 2
- C_2 - center third for team 2
- A_2 - attacking third for team 2
- S_2 - scored goal for team 2

Placing the current state on the rows and the next state as the columns, a transition matrix is formed as a base model of Markov processes in soccer. The values of the probabilities in the matrix are arbitrary with the exception of the first 2 rows. This is because the rules of soccer require that once a goal is scored, the ball returns to the center line for a kick off, starting with the possession of the other team. For example, if team 1 scores a goal (S_1) then the ball must return to the center line for a team 2 kickoff (C_2). Thus, the probability that the next state will be C_2 when the current state is S_1 is always 1 and vice versa.

There are also a few other values within this transition matrix that are influenced by the physical properties of soccer match. For instance, if the current state is D_1, meaning team 1 has the ball in

$$
\begin{array}{c} \\ S_1 \\ S_2 \\ A_1 \\ A_2 \\ C_1 \\ C_2 \\ D_1 \\ D_2 \end{array}
\begin{array}{c}
\begin{array}{cccccccc} S_1 & S_2 & A_1 & A_2 & C_1 & C_2 & D_1 & D_2 \end{array} \\
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0.25 & 0.05 & 0.35 & 0.2 & 0.05 & 0.1 \\
0.05 & 0 & 0.05 & 0.2 & 0.25 & 0.2 & 0.2 & 0.05 \\
0.05 & 0 & 0.15 & 0.1 & 0.25 & 0.15 & 0.2 & 0.1 \\
0 & 0 & 0.15 & 0.2 & 0.2 & 0.2 & 0.1 & 0.15 \\
0.15 & 0 & 0.05 & 0.2 & 0.2 & 0.15 & 0.1 & 0.15 \\
0 & 0.1 & 0.35 & 0 & 0.2 & 0.2 & 0.05 & 0.1
\end{bmatrix}
\end{array}
$$

Figure 4: Example of a transition probability matrix for soccer.

their defensive third, then the only way the next state can be S_2 is if team 1 scores an own goal. This is because team 2 must gain possession of the ball first (which requires another state) before a team member can shoot the ball.

2.3 Trajectory

A trajectory of a Markov chain is a particular set of values for X_t where $t = 0, 1, 2, ...$ The trajectory is generally referred to as $s_0, s_1, s_2, ...$, meaning that $X_0 = s_0, X_1 = s_1, X_2 = s_2, ...$ It essentially describes the path that the process takes through the states. The trajectory follows the Markov property because only its most recent point affects what will occur next. In the following example, we apply this term to a soccer oriented situation that will set up the basic comprehension necessary for the later use of MDPs.

2.3.1 Example

Using the matrix from 2.2.1, the following example shows a possible trajectory that may occur during in a soccer match. Say the initial state is C_1, meaning team 1 has the ball in the center third of the field. Then, the ball is intercepted or taken from team 1 but is still within the center third of the field. The state has now transitioned from C_1 to C_2. Now that team 2 has the ball,

$$C_1 \longrightarrow C_2 \longrightarrow A_2 \longrightarrow A_2 \longrightarrow D_1$$

Figure 5: An example trajectory during a small segment of a match.

they pass forward to their attacking third. Once in the attacking third team 2 connects a pass with their teammate who then takes a shot, but misses and the ball goes out to the left of the goal. Because the ball went out on the goal line and the last person to touch the ball was a member of team 2, the ball becomes a goal kick for team 1. In other words, this series of events ends in state D_1. As seen in Figure 5, the entire trajectory for these events that have taken place in this soccer match would follow the trajectory C_1 to C_2 to A_2 to A_2 to D_1. Notice that A_2 is repeated when mapping out the trajectory. This is because there is a state correlated with each action/event that takes place. Thus, when there are two consecutive events within the same third of the field, the trajectory will show the state of every event.

3 Eigenvalues and Eigenvectors

Eigenvalues are a concept from linear algebra that play a crucial role in various mathematical and scientific applications. Eigenvalues, sometimes referred to as characteristic roots, are a set of scalars that are associated with a linear system of equations [3]. Each eigenvalue has a corresponding eigenvector.

Eigenvalues and eigenvectors have various applications in fields such as physics, engineering, computer science, and more. They are particularly useful in solving systems of linear equations, analyzing dynamic systems, and understanding the behavior of linear transformations. Knowledge of eigenvalues and eigenvectors will assist in the implementation of MDP models on decision making in soccer. Next, we will learn how to solve for eigenvalues and eigenvectors to better understand how they function; however, their computation requires many steps and is better calculated by imputing into computers [3].

3.1 Solving for Eigenvalues and Eigenvectors

Consider a square matrix P of size $n \times n$. An eigenvector of P is a non-zero vector $\mathbf{v}$ such that when P operates on $\mathbf{v}$, the result is a scaled version of $\mathbf{v}$. This scaling factor is known as the eigenvalue associated with $\mathbf{v}$. Formally, if $\mathbf{v}$ is an eigenvector of P with eigenvalue π, then:

$$P\mathbf{v} = \pi\mathbf{v}$$

Here, π represents the eigenvalue and $\mathbf{v}$ is the corresponding eigenvector. The process of finding the eigenvalues of a matrix is fairly straightforward. First, find the size of the matrix P for which the eigenvalues and eigenvectors are being solved for. Determine the identity matrix of the same order. Then, using the identity matrix, calculate the matrix $P - I\pi$, where π is a scalar multiplier. Figure 4 shows the process for finding the matrix $P - I\pi$ using arbitrary letters as the values inside matrix P. After finding the matrix $P - I\pi$, find its determinant and set the equation equal to zero, i.e. $\det(P - \pi I) = 0$. From the equation created in the previous step, calculate all possible values of π, which are the eigenvalues of the matrix P.

The process of calculating the eigenvectors of a matrix has more complications involved in comparison to finding eigenvalues. Using the eigenvalues found with the above technique, return back to the equation $P - I\pi$. However, instead of using π as a scalar multiplier, the eigenvalues (all possible values of π) are inserted into the equation. Then, put the matrix or matrices that result from plugging the eigenvalue(s) into reduced row echelon form (RREF). After finding the RREF matrix, reparameterize the values in the matrix according to the free variables. The equation that results from the the reparameterization will lead to the eigenvector that corresponds with that specific eigenvalue. These steps are easier understood using an example.

$$P = \begin{bmatrix} A & B & C \\ D & E & F \\ G & H & K \end{bmatrix}, \quad I\pi = \begin{bmatrix} \pi & 0 & 0 \\ 0 & \pi & 0 \\ 0 & 0 & \pi \end{bmatrix} \rightarrow \quad P - I\pi = \begin{bmatrix} A-\pi & B & C \\ D & E-\pi & F \\ G & H & K-\pi \end{bmatrix}$$

Figure 6: The process for finding the matrix $P - I\pi$.

3.1.1 Example

The following example uses the steps specified above to solve for the eigenvalues and their corresponding eigenvectors of the given matrix, P.

$$\begin{array}{c c} & \begin{array}{ccc} A & B & C \end{array} \\ \begin{array}{c} A \\ B \\ C \end{array} & \begin{bmatrix} 1/5 & 3/5 & 1/5 \\ 4/5 & 1/5 & 0 \\ 2/5 & 2/5 & 1/5 \end{bmatrix} \end{array}$$

Using a 3x3 identity matrix and the scalar multiplier π, the matrix $P - I\pi$ is created. Then its determinant is found and set equal to zero. After solving for π using the above equation, it is clear that matrix P has three eigenvalues. These are represented by π_1, π_2, and π_3.

$$(\tfrac{1}{5} - \pi)^3 + 0 + (\tfrac{1}{5} \times \tfrac{4}{5} \times \tfrac{2}{5}) + (\tfrac{1}{5}^2 \times (\tfrac{1}{5} - \pi)) - 0 - ((\tfrac{1}{5} - \pi) \times \tfrac{4}{5} \times \tfrac{3}{5}) = 0$$

$$\text{eigenvalues: } \pi_1 = \frac{\sqrt{2}-1}{5}, \ \pi_2 = \frac{-\sqrt{2}-1}{5}, \ \pi_3 = 1$$

The next step is to find the matrices that correspond with $P - I\pi_1$, $P - I\pi_2$, and $P - I\pi_3$. Once these are found, they are row reduced and the values within are reparameterized according to the free variables. Once this is done, the eigenvectors corresponding with the three eigenvalues above are found.

$$\begin{bmatrix} \frac{1}{5} - \frac{\sqrt{2}-1}{5} & \frac{3}{5} & \frac{1}{5} \\ \frac{4}{5} & \frac{1}{5} - \frac{\sqrt{2}-1}{5} & 0 \\ \frac{2}{5} & \frac{2}{5} & \frac{1}{5} - \frac{\sqrt{2}-1}{5} \end{bmatrix} \xrightarrow{\text{eigenvector}} \begin{bmatrix} \frac{7\sqrt{2}+10}{2} \\ -4\sqrt{2}-6 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} \frac{1}{5} + \frac{\sqrt{2}-1}{5} & \frac{3}{5} & \frac{1}{5} \\ \frac{4}{5} & \frac{1}{5} + \frac{\sqrt{2}-1}{5} & 0 \\ \frac{2}{5} & \frac{2}{5} & \frac{1}{5} + \frac{\sqrt{2}-1}{5} \end{bmatrix} \xrightarrow{\text{eigenvector}} \begin{bmatrix} \frac{-7\sqrt{2}+10}{2} \\ 4\sqrt{2}-6 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} \frac{1}{5} - 1 & \frac{3}{5} & \frac{1}{5} \\ \frac{4}{5} & \frac{1}{5} - 1 & 0 \\ \frac{2}{5} & \frac{2}{5} & \frac{1}{5} - 1 \end{bmatrix} \xrightarrow{\text{eigenvector}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Note that, this is the definition of the "right eigenvectors", which is the standard way of calculating eigenvectors. But, for Markov chains, we are interested in the "left eigenvector" of eigenvalue of one ($\pi = 1$). This eigenvector corresponds to the stationary distribution of the Markov chain. For a square matrix P, a left eigenvector is characterized by

$$\psi P = \pi \psi$$

Where ψ is a left eigenvector of P. It can be shown that $\pi = 1$ is always an eigenvalue of stochastic matrices, setting $\pi = 1$ leads to:

$$\psi P = \psi$$

Solving for ψ leads to the stationary distribution of the corresponding Markov Chain. Using this formula, the stationary distribution of the transition probability matrix given in Figure 4 are calculated and shown in Table 1.

Table 1: Stationary distribution

State	S_1	S_2	A_1	A_2	C_1	C_2	D_1	D_2
ψ	0.036	0.011	0.155	0.122	0.243	0.208	0.119	0.106

4 Markov Decision Processes

A Markov Decision Process (MDP) provides a mathematical framework for modeling decision making when an outcome is partly random and partly under the control of the decision maker. Finding the solution to an MDP allows for an optimal rule or policy to be derived in order to inform the decision maker on what is the best suited decision to make in order to achieve the desired outcome. MDPs are used for a variety of purposes and disciplines including optimization problems in manufacturing, economics, and sports.

4.1 Markov Reward Processes

A Markov Reward Process provides a framework for modeling and analyzing dynamic scenarios in which entities transition between different states over time, such as a game character navigating through levels, a robot completing tasks, or the fluctuation of stock market values. The primary objective is to assess the quality of these transitions by evaluating the rewards associated with them.

A Markov Reward Process can be likened to a soccer player's journey during a match. In this analogy, the soccer player's states represent different positions and situations on the field, such as being in possession of the ball, attempting a pass, or defending against an opponent. As the game progresses, the player transitions between these states, making decisions and taking actions.

The rewards in this context could symbolize the outcomes of these actions. For instance, successfully scoring a goal or creating a goal-scoring opportunity might yield positive rewards, while losing possession or conceding a goal might lead to negative rewards.

Analyzing the player's performance within the Markov Reward Process framework helps assess how effective their movements and decisions are during the game, shedding light on the overall quality of their play and strategic choices.

Let's define the following notations:

States: A set of states $\mathcal{S} = \{S_1, S_2, \ldots, S_n\}$

Transition Probabilities: The probability of transitioning from state S_i to state S_j and denoted by, P_{ij} .

Immediate Rewards: The immediate reward upon entering state S_i and denoted by, R_i .

Discount Factor: A discount factor for future rewards and denoted by, γ .

4.1.1 Transition Probability Matrix

The transition probability matrix P describes the probabilities of transitioning between states and is defined as follows:

$$P = \begin{bmatrix} P_{11} & P_{12} & \ldots & P_{1n} \\ P_{21} & P_{22} & \ldots & P_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ P_{n1} & P_{n2} & \ldots & P_{nn} \end{bmatrix}$$

where P_{ij} represents the probability of transitioning from state S_i to state S_j, and $\sum_j P_{ij} = 1$ for all i.

4.1.2 Expected Total Reward

In the context of a Markov Reward Process, we define the expected total reward as the expected sum of rewards over an infinite time horizon. This can be represented using the following equation:

$$G_i = \sum_{t=0}^{\infty} \gamma^t R_{S_{t+1}}$$

Where:

G_i is the expected total reward starting from state S_i,

γ is the discount factor,

$R_{S_{t+1}}$ is the immediate reward upon entering state S_{t+1} at time $t + 1$.

Now, let's illustrate this concept with a soccer example. Consider a soccer player's performance in a match:

Suppose the player starts in state S_i, representing their current position on the field. The discount factor γ accounts for the importance of future rewards. The immediate reward $R_{S_{t+1}}$ could be a value assigned to a successful action, such as scoring a goal or making a successful pass.

Using this framework, we can calculate the expected total reward for the player, starting from their current position, by considering all possible future states and rewards as the match progresses. This can be done using the Bellman equation. The Bellman equation depicts the relation between the expected total reward for a state and the expected total reward for the states that follow. The Bellman equation is expanded below:

$$G_i = R_i + \gamma \sum_j P_{ij} G_j$$

Where:

G_i is the expected total reward starting from state S_i,

R_i is the immediate reward upon entering state S_i,

γ is the discount factor,

P_{ij} is the transition probability from state S_i to state S_j.

The Bellman equation is relevant to Markov Reward Processes because it expresses the value of a state as the sum of the immediate reward obtained when entering that state and the discounted value of all the possible successor states. This allows for a recursive-like calculation of the value of states which is a key part of the decision-making process.

4.2 Markov Decision Processes

Now that we have delved into the Markov reward process in the previous section, we are poised to take the next step and explore the fascinating realm of Markov decision processes (MDPs). MDP is a powerful mathematical framework for modeling decision-making scenarios where outcomes result from a blend of randomness and the choices made by a decision maker. It provides a mathematical framework for modeling decision making in situations where outcomes are partly random and partly under the control of a decision maker.

A Markov decision process is a Markov reward process with decisions. It is an environment in which all states are Markov.

Definition. A Markov Decision Process is a tuple $< \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma >$, where

- $\mathcal{S}$ is a finite set of states,
- $\mathcal{A}$ is a finite set of actions,
- $\mathcal{P}$ is a state transition probability matrix,

$$\mathcal{P}^a_{ss'} \ = \ \mathbb{P}[S_{t+1} = s' | S_t = s, A_t = a], \tag{1}$$

- $\mathcal{R}$ is a reward function, $\mathcal{R}^a_s \ = \ \mathbb{P}[R_{t+1} | S_t = s, A_t = a]$,
- γ is a discount factor, $\gamma \in [0, 1]$.
- State Space (S): The state space of an MDP contains all possible states of a process. An individual state represents a specific situation within the system. When applied to soccer, a state could represent the position of the players or the position of the ball on the field.
- Action Space (A): Each state has a set of actions that are available to the decision maker. The action space contains all possible actions that the decision maker can choose when interacting

with the system. For example, in a soccer match, some basic actions might include dribble, pass, and shoot.

- Transition Probabilities (P): Transitions from one state to a new state occurs when the decision maker selects an action in a given state. The chosen action effects the probability of moving from the current state to a specific next state. This probability is captured by the state transition function, denoted as P(s, s'). It represents the likelihood that taking action a in state s will result in state s' in the next time step.

- Immediate Rewards: An immediate reward refers to the benefit or cost obtained by the decision maker for taking a particular action in a specific state. This reward is received immediately upon transitioning from one state to another by executing a particular action. It is typically defined by the environment and is a crucial component used by the decision maker to learn and make decisions about which actions to take in different states. In soccer, the reward could represent a goal scored for (positive) or a goal scored against (negative).

- Markov Property: The Markov property states that the future state only depends on the current state and the chosen action. Because the transition between states does not depend on the past states or actions that lead up to the present moment, the Markov property is sometimes called the memoryless property.

- Policy Function (π): The policy function defines the strategy or behavior that the decision maker employs to discern which actions to take in different states. Formally, a policy is a mapping from states to actions, indicating what action the decision maker should choose when in a particular state. The goal is to find an optimal policy that maximizes the expected cumulative reward over time.

- Value Functions: Also a component of MDPs are value functions. These functions quantify the expected cumulative rewards achievable from a given state or state-action pair. The state-value function (V) represents the expected cumulative reward starting from a particular state s and following a certain policy π. The action-value function (Q) represents the expected cumulative reward starting from a particular state s, taking a specific action a, then following a certain policy π.

5 MDPs Applications in Soccer

At its core, an MDP (Markov Decision Process) is a mathematical model that represents decision-making problems within a stochastic (uncertain) environment. In the case of soccer, this environment includes various elements such as player positions, ball location, game score, and more. Player positioning is a key component of MDPs in soccer, as it plays a crucial role in determining optimal strategies, tactical planning, player development, and performance analysis. Understanding how players position themselves on the field is essential for coaches and analysts to make informed decisions and enhance team performance.

In this study, the aim is to determine the next action for our team based on the position of the ball and which team possesses the ball. To achieve this, we divide the field into three zones, and the state is determined by:

- The position of the ball.
- Which team has the ball, i.e., our team or the opponent.

5.1 Modeling Assumptions

In this section, we discuss the assumptions we make to be able to use MDP to model the game of soccer. The main assumption of this study are as follows:

- **Discrete time**: Even though soccer is a continuous time game, we assume that we can model it using a discrete time Markov chain process. We divide the state space into meaningfully distinct states and assume that the variation in the actual time it takes to transition between states does not significantly impact the analysis and can be ignored.

- **Infinite time horizon**: Even though soccer is played over a limited amount of time (90 minutes plus extra time), to simplify the analysis, we use an infinite time horizon.

- **Single Player Game**: Soccer is played by two teams with opposing objectives. On the other hand, MDPs aim to optimize decision making for one entity. In order to be able to model soccer using MDP, we fix the strategy of one of the teams and optimize the decisions for the other team.

5.2 States and Field Breakdown

For this model, the field is broken down into thirds: attacking, central, and defensive. The states of the model are the same as in the example 2.2.1, where possession and the location of the ball are taken into account. The 8 total states are as follows:

- D_1 - defensive third for team 1
- C_1 - center third for team 1
- A_1 - attacking third for team 1
- S_1 - scored goal for team 1
- D_2 - defensive third for team 2
- C_2 - center third for team 2
- A_2 - attacking third for team 2
- S_2 - scored goal for team 2

Note that when the ball is out of bounds for a throw in or corner kick, the model assumes that the ball is still within one of the three zones, although the possession and state have changed. In other words, to simplify the number of states, the model assumes that the ball is constantly within

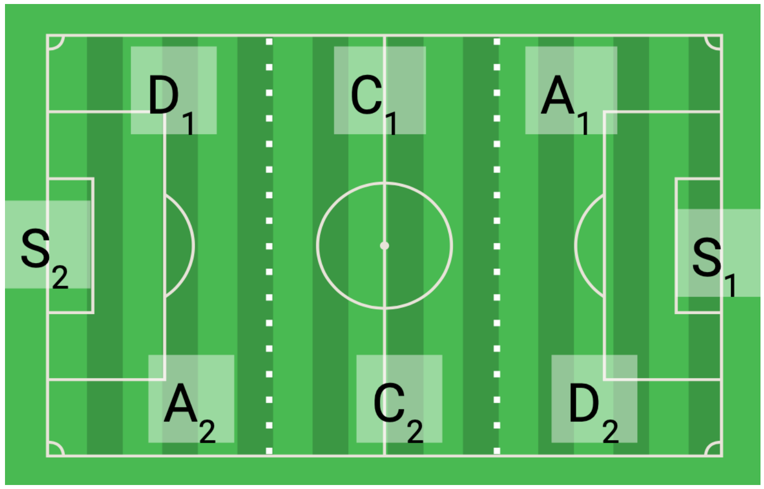

Figure 7: The zonal and possession breakdown of the pitch.

the area of the pitch.

5.3 Defining Actions

For this study, there are nine general actions that a player with the ball can decide to do. These actions include dribble, short pass, mid pass, long pass, short cross, mid cross, and shot. Below provides a definition and explanation of the possibilities of how each action is employed.

5.3.1 Dribble

The dribble action is straight forward. A player can dribble within one state, dribble forwards- meaning dribble towards the goal that team is progressing towards, or dribble backwards- meaning dribble towards the goal that team is defending.

5.3.2 Short Pass

A short pass is a pass that stays within one third. To simplify the model an the amount of actions, a short pass includes a clearance, goal kick, free kick, throw in, or punt that originates and ends in the same third.

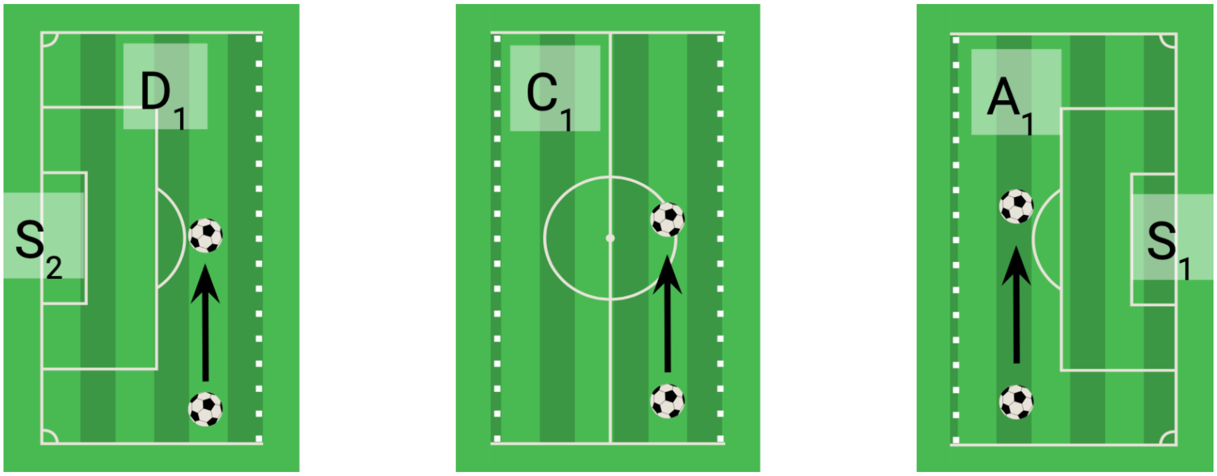

Figure 8: Three examples of a short pass.

5.3.3 Mid Pass

A mid pass is a pass that moves into an adjacent third. Like a short pass, a mid pass also includes a clearance, goal kick, free kick, throw in, or punt that originates in one third and ends in an adjacent third.

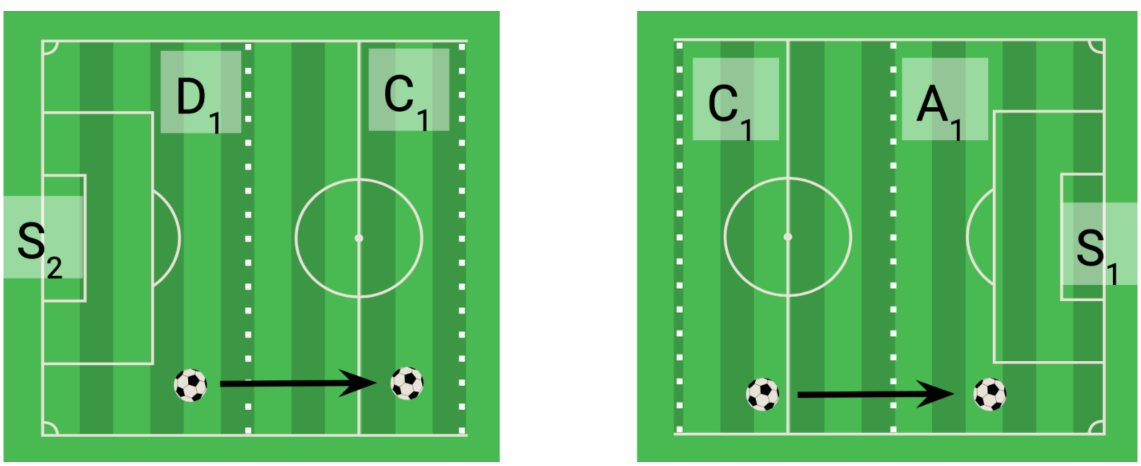

Figure 9: Two examples of a mid pass.

5.3.4 Long Pass (Long Ball)

A long pass or long ball as it is frequently called in soccer, is a a pass that begins in the defensive third of the field and ends in the attacking third of the field. This type of pass must goes through, but does not stop in the center third. In order to simplify the model, a long pass includes a clearance, goal kick, free kick, punt.

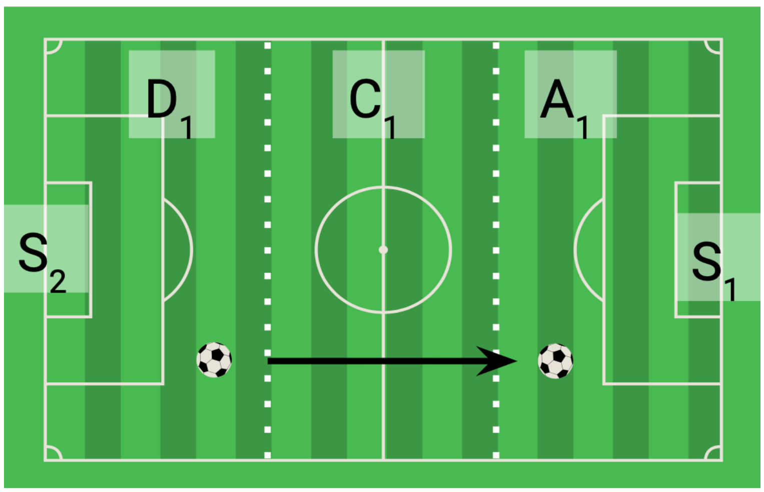

Figure 10: Example of a long pass/long ball.

5.3.5 Short Cross

In general, a cross is a horizontally driven ball that switches sides of the field, traditionally near the team's goal that they are progressing towards. A short cross is a cross that stays within one third of the field. The left image in Figure 11 is an example of a short cross.

5.3.6 Mid Cross

A mid cross is a cross that originates in one third and ends in the adjacent third. Again, this type of ball is done traditionally near the goal that the team is scoring on. The right image in Figure 11 is an example of a mid cross.

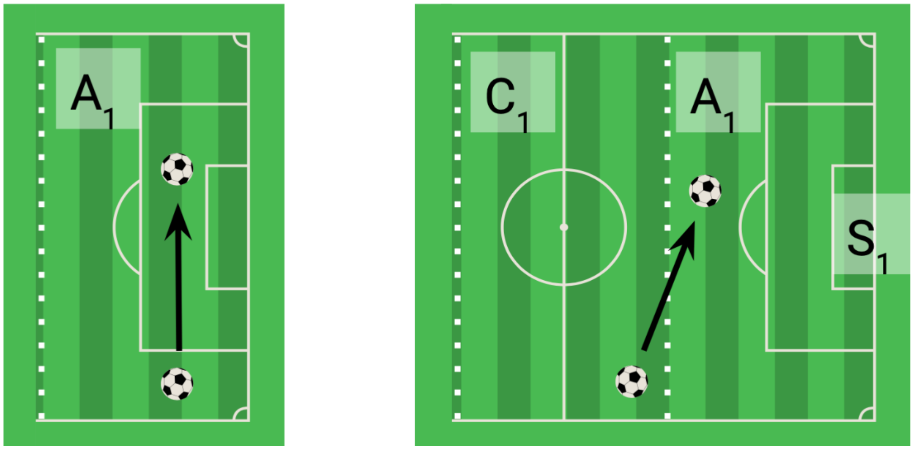

Figure 11: Two examples of a cross.

5.3.7 Short Out

A short out is a strategic move that occurs most commonly in the defensive third and occasionally in the central third. It is when a player purposely kicks the ball out of bounds. This is done when a player is under pressure and does not have many open teammates available to pass to. A short out originates in one third and ends within that same third (due to the assumption that the ball is still within the closest third when outside of the playing area boundary.)

5.3.8 Mid Out

A mid out is similar to a short out except for the fact that it originates in one third and ends in the adjacent third. Again, this action occurs mostly in the defensive third and occasionally in the central third of the field.

5.3.9 Shot

The shot action is again a straight forward action. It is when a player attempts to score a goal. This action becomes more complex when considering the transition probabilities.

5.4 Transition Probabilities

With every action comes the possibility of losing possession. This is taken into account by the transition probabilities. The transition probabilities that describe each possible outcome in each state were created using subject-matter expert (SME) estimations. A subject-matter expert is a person whose years experience with a subject has allowed them to accumulate great knowledge in a particular field. A degree or license are also typically considered qualifications for one to be a SME.

For every state, each outcome of the action needs to be considered and a probability assigned. These probabilities for each state must add up to one because each of these states will represent a new stochastic matrix. The following subsections show the normalized probabilities assigned to each outcome for 6 states. The states S_1 and S_2 are not included in this because there is only one possible outcome after a goal is scored- the ball returns to the center of the field and the other team begins with possession. Also note that intercepted means there has been a possession change, as the action that was performed has been intercepted by the other team.

5.4.1 Defensive Third for Team 1 (D_1)

action	comment	next state	probability
dribble		D_1	0.75
dribble	intercepted	A_2	0.25
dribble forward		C_1	0.73
dribble forward	intercepted	C_2	0.27
short pass		D_1	0.93
short pass	intercepted	A_2	0.07
mid pass forward		C_1	0.89
mid pass forward	intercepted	C_2	0.11
long pass		A_1	0.67
long pass	intercepted	D_2	0.33
short cross		D_1	1
mid cross		C_1	0.83
mid cross	intercepted	C_2	0.17
short out		A_2	1
mid out		C_2	1

5.4.2 Central Third for Team 1 (C_1)

action	comment	next state	probability
shot		S_1	0.20
shot	deflected	A_1	0.20
shot	saved/missed	D_2	0.40
shot	intercepted	D_2	0.20
dribble		C_1	0.65
dribble	intercepted	C_2	0.35
dribble forward		A_1	0.59
dribble forward	intercepted	D_2	0.41
dribble back		D_1	0.625
dribble back	intercepted	A_2	0.375
short pass		C_1	0.57
short pass	intercepted	C_2	0.43
mid pass forward		A_1	0.62
mid pass forward	intercepted	D_2	0.38
mid pass back		D_1	0.75
mid pass back	intercepted	A_2	0.25
short cross		C_1	0.67
short cross	intercepted	C_2	0.33
mid cross		A_1	0.625
mid cross	intercepted	D_2	0.375
short out		C_2	1
mid out forward		D_1	0.50
mid out back		A_2	0.50

5.4.3 Attacking Third for Team 1 (A_1)

action	comment	next state	probability
shot		S_1	0.22
shot	deflected	A_1	0.26
shot	saved/missed	D_2	0.35
shot	intercepted	D_2	0.17
dribble		A_1	0.55
dribble	intercepted	D_2	0.45
dribble back		C_1	0.625
dribble back	intercepted	C_2	0.375
short pass		A_1	0.5625
short pass	intercepted	D_2	0.4375
mid pass back		C_1	0.75
mid pass back	intercepted	C_2	0.25
short cross		A_1	0.57
short cross	intercepted	D_2	0.43

5.4.4 Defensive Third for Team 2 (D_2)

action	comment	next state	probability
dribble		D_2	0.60
dribble	intercepted	A_1	0.40
dribble forwards		C_2	0.625
dribble forwards	intercepted	C_1	0.375
short pass		D_2	0.71
short pass	intercepted	A_1	0.29
mid pass forwards		C_2	0.71
mid pass forwards	intercepted	C_1	0.29
long pass		A_2	0.625
long pass	intercepted	D_1	0.375
short cross		D_2	0.75
short cross	intercepted	A_1	0.25
mid cross		C_2	0.71
mid cross	intercepted	C_1	0.29
short out		A_1	1
mid out		C_1	1

5.4.5 Central Third for Team 2 (C_2)

action	comment	next state	probability
shot		S_2	0.10
shot	deflected	A_2	0.10
shot	saved/missed	D_1	0.40
shot	intercepted	D_1	0.40
dribble		C_2	0.55
dribble	intercepted	C_1	0.45
dribble forward		A_2	0.53
dribble forward	intercepted	D_1	0.47
dribble back		D_2	0.67
dribble back	intercepted	A_1	0.33
short pass		C_2	0.57
short pass	intercepted	C_1	0.43
mid pass forward		A_2	0.53
mid pass forward	intercepted	D_1	0.47
mid pass back		D_2	0.56
mid pass back	intercepted	A_1	0.44
short cross		C_2	0.50
short cross	intercepted	C_1	0.50
mid cross		A_2	0.63
mid cross	intercepted	D_1	0.37
short out		C_1	1
mid out forward		D_1	0.44
mid out back		A_1	0.56

5.4.6 Attacking Third for Team 2 (A_2)

action	comment	next state	probability
shot		S_2	0.11
shot	deflected	A_2	0.21
shot	saved/missed	D_1	0.32
shot	intercepted	D_1	0.36
dribble		A_2	0.55
dribble	intercepted	D_1	0.45
dribble back		C_2	0.60
dribble back	intercepted	C_1	0.40
short pass		A_2	0.52
short pass	intercepted	D_1	0.48
mid pass back		C_2	0.60
mid pass back	intercepted	C_1	0.40
short cross		A_2	0.56
short cross	intercepted	D_1	0.44

5.5 Rewards

The aim of this analysis is to optimize the decision making for Team 1. To achieve this, the reward function is directly tied to the score of the game from Team 1's point of view. In other words, whenever a transition to S_1 is made, i.e. Team 1 scores a goal, a reward of $+1$ is applied. In addition, whenever a transition to S_2 is made, i.e. Team 2 scores a goal, a reward of -1 is applied.

6 Iterative Solution Algorithm

In this section, we use the policy iteration algorithm to solve the MDP described in the previous sections. The policy iteration algorithm is an iterative procedure that can be used to find a solution for Markov decision processes. It involves iteratively modifying a policy until convergence to an optimal policy is achieved. Policy iteration consists of two main steps: policy evaluation and policy improvement. These steps are repeated iteratively until the policy converges to the optimal policy.

- **Policy Evaluation:** Given a policy π, evaluate its performance by estimating the value function $V(\pi)$. The value function represents the expected cumulative reward when following policy π. This step involves solving the Bellman equation for the current policy.

- **Policy Improvement:** Once the value function $V(\pi)$ is obtained, improve the policy by selecting actions that maximize expected rewards. Update the policy based on the current value function.

- **Iterative Process:** Repeat the policy evaluation and improvement steps until convergence criteria are met. Convergence is typically achieved when the policy no longer changes significantly between iterations.

7 Problem Results

In this section, Python programming language is used to implement the policy iteration algorithm to solve the MDP model described in section 5. As mentioned earlier, to avoid having to solve a game-theoretic model, we fix the strategy of Team 2 to the policy provided in Table 2:

Table 2: Fixed Policy For Team 2

State	**Action**
A2	Shot
C2	Dribble Forwards
D2	Mid Pass

For a discount factor of $\gamma = 0.9$, the optimal policy for Team 1 is presented in Table 3.

Table 3: Optimal Policy For Team 1

State	**Action**
A1	Shot
C1	Shot
D1	Mid Pass

Moreover, and the optimal policy values are given in Table 4.

Table 4: Optimal Policy Values

State	**Value**
A1	0.5837
D2	0.3645
A2	0.2795
D1	0.4660
S2	0.4998
C1	0.5554
C2	0.3299
S1	0.2969

Because the reward function is designed to be directly tied to the score, the value in each state can be interpreted as the average discounted number of goals scored in that state. Therefore, based on the results in this table, the average discounted goals scored in all states are positive numbers and we can deduce that Team 1 is stronger than Team 2. The optimal action when Team 1 is attacking (A1) is to shoot. Also when Team 1 has the ball in the center of the field (C1), the optimal decision is to shoot as well. The reason for this seemingly risky decision is because Team 2 is a weak opposition and even if Team 1 loses possession of the ball, Team 2 cannot effectively take advantage of the possession to score a goal.

References

[1] Henk C Tijms. *A first course in stochastic models.* John Wiley and sons, 2003.

[2] J. R. Norris. *Markov Chains.* Cambridge University Press, 1998.

[3] Hervé Abdi. The eigen-decomposition: Eigenvalues and eigenvectors. *Encyclopedia of measurement and statistics*, pages 304–308, 2007.