Analyzing Learned Markov Decision Processes using Model Checking for Providing Tactical Advice in Professional Soccer

Maaike Van Roy[1], **Wen-Chi Yang**[1], **Luc De Raedt**[1] and **Jesse Davis**[1]

[1]KU Leuven, Department of Computer Science; Leuven.AI

{firstname.lastname}@kuleuven.be

Abstract

Markov models are commonly used to model professional sports matches as they enable modelling the various actions players may take in a particular game state. In this paper, our objective is to reason about the goal-directed policies these players follow. Concretely, we focus on soccer and propose a novel Markov decision process (MDP) that models the behavior of the team possessing the ball. To reason about these learned policies, we employ techniques from probabilistic model checking. Our analysis focuses on defense, where a team aims to minimize its risk of conceding a goal (i.e., its opponent scores). Specifically, we analyze the MDP in order to gain insight into various ways an opponent may generate dangerous situations, that is, ones where the opponent may score a goal, during a match. Then, we use probabilistic model checking to assess how much a team can lower its chance of conceding by employing different ways to prevent these dangerous situations from arising. Finally, we consider how effective the defensive strategies remain once the offensive team adapts to them. We provide multiple illustrative use cases by analyzing real-world event stream data from professional soccer matches in the English Premier League.

1 Introduction

Markov models are one of the most commonly used modelling formalisms in AI. Due to their ability to model dynamic environments and decision making, Markov models have been used to tackle a variety of different real-world problems. One surprising area where Markov models have had a real-world impact is in analyzing professional sports matches, where they have been applied to American football [Goldner, 2012], basketball [Cervone *et al.*, 2016; Sandholtz and Bornn, 2018; Sandholtz and Bornn, 2020; Wang *et al.*, 2018], ice hockey [Routley and Schulte, 2015; Schulte *et al.*, 2017] and soccer [Hirotsu and Wright, 2002; Rudd, 2011; Singh, 2019; Van Roy *et al.*, 2021b; Yam, 2019]. In the context of sports, Markov models enable modelling the various actions players can take in particular game states. This can yield insight into areas such as predicting match outcomes [Dong *et al.*, 2015], evaluating decisions such as substitutions [Hirotsu and Wright, 2002] and shooting behavior [Van Roy *et al.*, 2021b], and rating players [Cervone *et al.*, 2016; Routley and Schulte, 2015; Rudd, 2011; Singh, 2019; Yam, 2019] which is useful for player acquisition.

Our objective is to reason about the goal-directed policies that players follow which could help support the tactical planning of coaches. Specifically, we consider soccer and take a defensive perspective where a team's objective is to minimize its chance of conceding a goal. To provide insight on how to achieve this defensive objective, it is necessary to understand how an opponent may generate dangerous situations, thus, it is necessary to reason about their goal-directed policies. For example, one will need to reason about which areas of the pitch the opponent uses to generate shooting opportunities from.

Our key insight is that probabilistic model checking techniques provide an avenue for reasoning about such goal-directed policies. These techniques provide formal guarantees about the probabilities of certain behaviors occurring in a Markov model. Moreover, probabilistic model checking can allow us to reason about the efficacy of certain actions that could disrupt your opponent's attack. For example, it can reason about how forcing an opponent *to avoid reaching areas of pitch* would affect their probability of generating a shot. This information can be used by coaches to aid in their planning.

Our framework for tackling this problem involves a two-step process that combines learning a Markov Decision Process (MDP) from data and using probabilistic model checking to reason about the learned policy. In the first step we learn a team-specific MDP to model an opponent's offensive behavior. The MDP's state space consists of locations on the pitch and the actions involve moving between these states or shooting on goal. We learn the observed policy of a team using event stream data, which is a common source of data about professional soccer matches that records information such as the location and time of on-the-ball events like passes and shots. In the second step we show how to use the probabilistic model checker PRISM [Kwiatkowska *et al.*, 2011] to reason about the learned MDP in a number of different ways. First, it allows for reasoning about how an opponent may generate a scoring opportunity. Second, it allows for reasoning about the effect certain defensive strategies would have on reducing the chance of conceding a goal. Furthermore, we can estimate how effective these strategies remain even if the opponent were to adapt to them.

The model provides actionable explanations on the behavior of opponents and can be used to provide invaluable insights to coaches when they are planning their strategies against them. To summarize, this paper makes the following contributions:

- it suggests a novel and real-world application that involves verifying properties of learned models in the context of sports;
- it proposes a novel MDP for modelling the tactical behavior of professional soccer teams;
- it shows how the probabilistic model checker PRISM can reason about the learned policies;
- it provides a number of illustrative examples of tactical insights about teams in the English Premier League.

2 Preliminaries

We provide some background on Markov models, their application to soccer, and probabilistic model checking.

2.1 Markov Models

A Markov model describes a probabilistic process for transitioning between various states in a system. A Markov Reward Process (MRP) is a Markov model where transitions may have an associated reward or cost. More formally, an MRP is a tuple $\langle \mathcal{S}, P, R, \gamma \rangle$ where $\mathcal{S}$ is the set of states, $P : \mathcal{S} \times \mathcal{S} \to [0, 1]$ is the transition function, $R : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ is a reward function associated with every transition and $\gamma \in [0, 1]$ is a discount factor.

A Markov Decision Process (MDP) [Bellman, 1957; Howard, 1960] extends an MRP to allow transitions between states to depend on the action taken in a state. Formally, an MDP is a tuple $\langle \mathcal{S}, \mathcal{A}, P, R, \gamma \rangle$ where $\mathcal{S}$ is the set of states, $\mathcal{A}$ is the set of actions, $P : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to [0, 1]$ is the transition function, $R : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to \mathbb{R}$ is the reward function, and $\gamma \in [0, 1]$ is a discount factor. For an MDP, a policy $\pi : S \to dist(A)$ can be defined which specifies the probability distribution over actions given a certain state. Together with the MDP, it defines the behavior of an agent.

2.2 Markov Models for Soccer

In soccer, Markov models have most commonly been used to assign a value to each on-the-ball action a player performs during a match, which is a key soccer analytics task [Decroos *et al.*, 2019; Liu *et al.*, 2020; Singh, 2019; Van Roy *et al.*, 2020]. For this, an MRP is typically used to model the in-game behavior of teams [Rudd, 2011; Singh, 2019]. These approaches divide a soccer match into possessions, where each possession is a sequence of consecutive on-the-ball actions carried out by the same team. Each action transitions the game from one state to another in the MRP. Players perform actions with the intention of arriving at game states where they have a higher chance of scoring. The players perform actions until one of two absorbing states is reached: (i) a goal is scored and they receive a reward of 1 or (ii) the possession ends (e.g., a turnover occurs, a shot is taken and missed, etc.) and they receive a reward of 0. The value of a non-absorbing state s is then the probability of eventually scoring from s, which can be obtained using the standard dynamic programming approach. The values of non-absorbing states are then used to value all on-the-ball actions.

Time	Action	Start	End	Result	Player	Team
20m49s	pass	(57.4, 49.1)	(45.9, 49.1)	success	İ. Gündoğan	Manchester City
20m49s	dribble	(45.9, 49.1)	(40.6, 43.0)	success	K. De Bruyne	Manchester City
20m51s	pass	(40.6, 43.0)	(31.8, 47.3)	fail	K. De Bruyne	Manchester City
20m51s	interception	(31.8, 47.3)	(31.8, 47.3)	success	D. Lovren	Liverpool

Figure 1: Four actions in event stream format recorded from Manchester City versus Liverpool on 14/1/2018.

The probabilities in the model are learned from historical event stream data, which is collected by human annotators. This type of data describes all on-the-ball actions during a soccer game. For each on-the-ball action, a number of attributes are recorded such as the type of the action (e.g., pass, dribble, shot), the start and end locations of the action, a timestamp, the player who performed the action, etc. Additionally, it records if the action succeeded or failed. For example, a pass is successful if it reaches and is controlled by a teammate. An example of an unsuccessful pass is one that is intercepted by the opposing team. Figure 1 illustrates a part of a possession sequence of Manchester City as it is recorded in the event stream data format.

2.3 Probabilistic Model Checking

Probabilistic model checkers (e.g., PRISM [Kwiatkowska *et al.*, 2011] and STORM [Hensel *et al.*, 2020]) verify whether a probabilistic system satisfies a specific property. A property describes a set of desired system behavior over time, and is specified by a temporal logic formula. As both the system and the property are mathematically formulated, probabilistic verification can provide rigorous, quantitative guarantees.

In this paper, we focus on the reachability related properties that are supported in the PRISM model checker to reason about how an opponent reaches a dangerous situation. Specifically, we use the probabilistic reachability property in PCTL* (which subsumes probabilistic computational tree logic, PCTL, and linear temporal logic, LTL) for MRPs. Such properties evaluate to either true or false in a state, and are constructed using the classical logical operators ($\wedge$, $\vee$, $\neg$), the temporal components α $\mathtt{U}$ β (α holds until β holds), α $\mathtt{U}^{\leq \mathtt{k}}$ β where $\mathtt{k}$ is a positive integer (α holds until β holds and β will hold within $\mathtt{k}$ steps), $\mathtt{X}$ α (α holds in the next step), and $\mathtt{P}_{\bowtie \mathtt{p}}[\alpha]$ where $\bowtie \in \{<, >, \leq, \geq\}$ (α holds with a probability $\bowtie$ $\mathtt{p}$). Additionally, the commonly used temporal component $\mathtt{F}$ α (α will eventually hold) can be defined by the until operator: $\mathtt{F}\ \alpha \equiv \mathtt{true}\ \mathtt{U}\ \alpha$. We will specifically look at quantitative properties of the form $\mathtt{P}_{=?}[\alpha]$ which query the probability that α holds. This means that a quantitative property evaluates to a real number in [0,1].

In soccer, a possession sequence can start anywhere on the field. Consequently, when mapping locations on the field to states in the MRP, a property can be evaluated for any initial location or state. We will use $\mathtt{P}_{f=?}[prop]$ to indicate that a property $prop$ is evaluated for a state f. This then returns the probability that $prop$ holds in f.

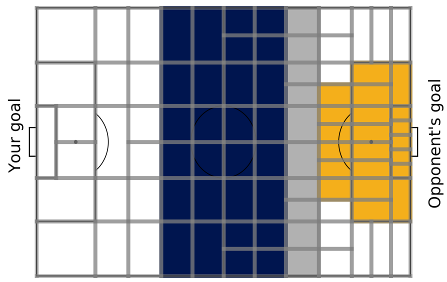


Figure 2: The gray lines denote the field states used in the MDP. A team attacks from left to right. The partitioning is more fine-grained near the opponent's goal and more coarse-grained in defensive half of the pitch. The colored regions denote zones used in the verification queries: yellow denotes the area where most shots are taken from, gray denotes the final third entry region, and dark blue denotes the middle third of the pitch.

3 MDP for Soccer

Our goal is to design a discrete soccer MDP for analyzing and verifying aspects of a team's tactical or goal-directed behavior. Most existing Markov models for soccer are MRPs [Rudd, 2011; Singh, 2019] that consider only a probability distribution over two actions: move or shoot. However, this is too limiting from a tactical perspective. We propose a soccer MDP that generalizes a soccer MRP with the following actions: *shoot* and *move_to* with an intended destination. This allows explicitly reasoning about a team's actions, policy, and tactical behavior in various situations. Formally, the soccer MDP is a tuple $\langle \mathcal{S}, \mathcal{A}, P, R, \gamma \rangle$ where $\mathcal{S}$, $\mathcal{A}$, P and R are defined as follows, and γ is a discount factor.

$\mathcal{S} = \mathcal{FS} \cup \mathcal{E}$ where $\mathcal{FS}$ is the set of 89 field states shown in Figure 2 and $\mathcal{E} = \{lp, ng, gs\}$ denotes the set of absorbing states or end states, where lp signifies loss of possession (i.e., a failed move), ng a failed shot (i.e., no goal), and gs a successful shot (i.e., the goal state). We deviate from the use of a uniform grid because our expanded action set creates issues of data sparsity. The analysis needs to occur on the team level because each team employs different tactical styles and there is limited data per team.[1] Our partitioning of the field states uses a fine-grained partitioning where chances of scoring are higher (i.e., in front of the opponent's goal), and a more coarse-grained partitioning where goal scoring chances are lower (i.e., on a team's own half). This ensures sufficient data in each state while being fine-grained enough to capture important differences between locations.

$\mathcal{A} = \{move\_to(f) \mid f \in \mathcal{FS}\} \cup \{shoot\}$ is the set of actions that allow for *where* to move the ball to.

$P : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow [0, 1]$ is the transition function and is defined for both the absorbing states, $\mathcal{E}$, and the field states, $\mathcal{FS}$. States $\mathcal{E}$ are absorbing, and so the set of possible actions for such states is empty. For field states, the transition probabilities are defined as follows:

- $P(f, move\_to(f'), f')$ denotes the probability of successfully moving to f' from f; if the move fails, the result is a loss of possession (lp), thus $P(f, move\_to(f'), lp) = 1 - P(f, move\_to(f'), f')$;
- $P(f, shoot, gs)$ denotes the probability of scoring a goal from f if the *shoot* action has been selected; if the shot fails to reach the goal, the result is ng, thus $P(f, shoot, ng) = 1 - P(f, shoot, gs)$;
- $P(f, a, f') = 0$ in all other cases.

$R : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow \mathbb{R}$ is the reward function. The reward is 1 only when $R(f, shoot, gs) = 1$ (with $f \in \mathcal{FS}$), and 0 otherwise.

The long-term value of a state $s \in \mathcal{S}$ in this MDP is then given by the following value function [Bellman, 1966]:

$$V_\pi(s) = \sum_{a \in \mathcal{A}} \pi(a|s) \sum_{s' \in \mathcal{S}} P(s, a, s')(R(s, a, s') + \gamma V_\pi(s'))$$

with π the policy to be followed. A policy is defined as the probability distribution over actions for each possible state:

$$\pi(a|s) = Pr[A = a|S = s]$$

We will write $V(s)$ instead of $V_\pi(s)$ when reasoning about an MRP whose policy π is fixed.

A team's policy and transition model can typically be learned from historical data by estimating all probabilities with simple counts. However, in this soccer MDP, this is complicated by the chosen action space. Calculating the number of times an action intended to move the ball from state f to f' requires knowing both the successful and unsuccessful move actions from state f to f'. However, for unsuccessful actions, the intended end location is unknown. Thus, in order to estimate the probabilities, one must first identify the intended end locations of failed movement actions (i.e., dribbles, passes, crosses). We solve this problem by using the approach of Van Roy et al. [2021a]. The approach uses a gradient boosted trees ensemble to predict the intended end location of actions based on several characteristics of the actions and what happened before those actions. Afterwards, the team-specific policy and transition model can be obtained in the standard way by counting the occurrences of actions in each state.

4 Reasoning about Soccer MDPs

To reason about a team's goal-directed behavior, we first apply the above mentioned methodology to learn a team-specific policy and transition model of the underlying MDP. We use real-world event stream data from the 2017/18 and 2018/19 English Premier League (EPL) seasons, provided by StatsBomb, and create a team-specific MDP for each of the 17 teams that played in both seasons. The policy together with the MDP can be transformed into an MRP that probabilistic model checkers such as PRISM can analyze. That is, we analyze the MRP obtained by fixing a team's policy in the soccer

[1] Top flight teams play between 34 and 38 league games in a season and perform between 500 and 2000 actions per game. Data more than a season old may not be relevant: playing styles vary over time due to changes in playing and management personnel.

MDP. By checking different properties against the MRP, we gain insight into how the team reacts to different situations.

In soccer, a team's objective is to score goals, which can be specified as a probabilistic reachability property. Checking this property against the team's MRP produces a value function that assigns to each location a probability of scoring, and hence represents the team's threat level. Formally,

$$V(f) = \mathtt{P}_{f=?}[\mathtt{F}\ gs] \tag{1}$$

The threat level can, of course, be used by the opponent for reasoning about the effect of possible defensive tactics which is the focus of this paper. We now show how one can reason about possible strategies for reducing the opponent's scoring opportunities by analyzing the opponent's MRP (i.e., MDP with fixed policy) and forcing them to avoid certain critical locations on the field. Specifically, we look at the crucial locations for generating shots and buildup play, and end by assessing how effective these defensive strategies remain after the offensive team adapts to them.

4.1 Shot Suppression

Roughly 91% of shots arise from the yellow shaded region ($shot\_locs$) in Figure 2. One approach to ultimately help reduce your chance of conceding, is to *directly* suppress the shots your opponent takes from this favored region. However, once the ball reaches this region, your opponent has already created a dangerous situation. Therefore, another approach is to limit the number of times your opponent reaches this favored region to *indirectly* suppress shots. We now discuss both approaches.

Indirect shot suppression

To reason about how likely your opponent is to reach $shot\_locs$ from a location f, consider the following query:

$$\mathtt{P}_{f=?}[\mathtt{F}\ shot\_locs] \tag{2}$$

Suppose we could prevent an opponent from ever entering $f' \in non\_shot$, where $non\_shot = \mathcal{FS} \setminus shot\_locs$. We can reason about the effect of this counterfactual policy that forces the opposing team to avoid state f' (i.e., the opposing team can never enter state f') on this probability using the following query:

$$\mathtt{P}_{f=?}[\neg f'\ \mathtt{U}\ shot\_locs] \tag{3}$$

Both queries can be combined to reason about the effect of avoiding f' on ever reaching $shot\_locs$ from f. However, a sequence can begin in any state. Therefore, to compute the average reduction in your opponent's probability of reaching $shot\_loc$ when forced to avoid f', we must sum over all locations in non_shot in numerator and denominator:

$$1 - \frac{\sum_{f \in non\_shot} \mathtt{P}_{f=?}[\neg f'\ \mathtt{U}\ shot\_locs]}{\sum_{f \in non\_shot} \mathtt{P}_{f=?}[\mathtt{F}\ shot\_locs]} \tag{4}$$

By computing Formula 4 for all f', we can measure each state's percent decrease in the probability of reaching $shot\_locs$ when prevented from entering f'. This gives an indication of f''s importance for indirectly suppressing shots.

Figures 3a and 3b illustrate the effect of indirect shot suppression for Manchester City and Burnley. Manchester City's most important states lie centrally with the flanks having a smaller effect. An opponent that prevents them from entering their most important state would decrease Manchester City's chance of reaching the common shooting locations by almost 20%. Tactically, Manchester City is a possession-based team that gradually builds up their attack. Often, their top creator is midfielder Kevin De Bruyne, who tends to operate on the center-left, and is known for his well-timed through balls. For Burnley, their most important states lie centrally, just outside the shooting locations, and on the right side of the penalty box. Burnley is known for their frequently crossing the ball into the penalty box and also have one of the better cross accuracy percentages of teams in the EPL. Preventing them from reaching their top location decreases their chance of reaching the shooting locations by just over 9%.

Direct shot suppression

Besides limiting a team's chance of reaching locations from which they shoot, reducing the number of shots your opponent takes also helps reduce your chance of conceding. To reason about how a team generates shots, consider the query:

$$\mathtt{P}_{f=?}[\mathtt{F}\ (shot\_locs\ \mathtt{U}\ (gs \vee ng))] \tag{5}$$

which gives the probability of a sequence starting in f eventually reaching a state in $shot\_locs$ after which it eventually reaches either gs or ng. That is, it computes the unrestricted probability of a sequence starting in f resulting in a shot from $shot\_locs$. We can also reason about the effect of forcing the opposing team to avoid state f' on the probability of shooting using the following query:

$$\mathtt{P}_{f=?}[\neg f'\ \mathtt{U}\ (shot\_locs\ \mathtt{U}\ (gs \vee ng))] \tag{6}$$

By combining these two queries, we can compute the percent decrease in your opponent's probability of shooting if you can force them to avoid entering location f':

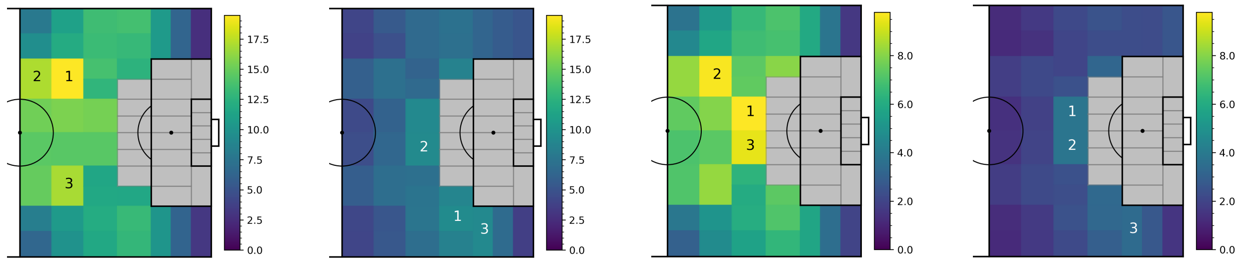
$$1 - \frac{\sum_{f \in non\_shot} \mathtt{P}_{f=?}[\neg f'\ \mathtt{U}\ (shot\_locs\ \mathtt{U}\ (gs \vee ng))]}{\sum_{f \in non\_shot} \mathtt{P}_{f=?}[\mathtt{F}\ (shot\_locs\ \mathtt{U}\ (gs \vee ng))]} \tag{7}$$

By computing this for all f', we can measure each state's importance for directly suppressing shots.

Figures 3c and 3d illustrate the effect of direct shot suppression for Manchester City and Burnley. Again, Manchester City's most important states lie centrally and Burnley's most important states tend to lie more to the right side of the penalty box. An opponent that prevents Manchester City (Burnley) from entering their most important state would decrease Manchester City's (Burnley's) chance of shooting by almost 10% (4%).

4.2 Movement Suppression

The final third entry and middle third regions (resp. gray and dark blue regions in Figure 2) are critical for building up an attack. The following query reasons about reducing the opponent's chance of scoring in these states (reg), and thus reducing their threat during buildup. Specifically, for a state $f \in reg$, it queries a set of states $area \subseteq \mathcal{FS} \setminus reg$ where a team should prevent its opponent from entering in order to



(a) Indirect (MC) (b) Indirect (B) (c) Direct (MC) (d) Direct (B)

Figure 3: The percent decrease in reaching ((a) and (b)) and shooting from ((c) and (d)) the common shot locations (in gray) for Manchester City (MC) and Burnley (B). Yellow shading indicates states with a larger decrease, whereas dark blue shading indicates a smaller decrease. The three states with the largest decrease are labeled in each figure. Manchester City experiences the biggest decreases in the (deep) central areas with less impact on the flanks. In contrast, Burnley experiences large decreases in the central areas near the box and on the right flank.

decrease the opponent's chance of scoring by at least x percentage points.

$$\forall f \in reg : (\mathtt{P}_{f=?}[\mathtt{F}\ gs] - \mathtt{P}_{f=?}[\neg area\ \mathtt{U}\ gs]) \geq x \tag{8}$$

The first $\mathtt{P}_{f=?}$ is simply the chance of scoring in state f (i.e., Equation 1) whereas the second $\mathtt{P}_{f=?}$ is the chance of scoring while avoiding $area$. Reasoning about which subsets of the states would form good candidate areas can be addressed using the query:

$$area = \{f'' \in \mathcal{FS} \setminus reg \mid \mathtt{P}_{f''=?}[\mathtt{F}\ f'] \geq b\} \tag{9}$$

This query forms a cluster of states around a state $f' \in \mathcal{FS} \setminus reg$ by finding all other states f'' such that the probability of reaching f' when starting from f'' is greater than b.

Figure 4 shows the results for Manchester City and Liverpool. To reduce Manchester City's chance of scoring from each of the final third entry states by at least 10%, the crucial areas to avoid lie around the center and left side of the field which is where creators like Kevin De Bruyne, Leroy Sané and Raheem Sterling operate the most. Decreasing their chance of scoring by at least 1% in each state in the middle third of the field can be done by forcing them to avoid the center of their defensive third. To reduce Liverpool's chance of scoring from each of the final third entry states by at least 10%, the crucial areas to avoid are a mirrored version of those of Manchester City. Specifically, the middle and right side of the field should be avoided, which are exactly those locations where Liverpool's attacking wing-back Trent Alexander-Arnold and "Player of the Season 2017/2018" Mohamed Salah operate. Decreasing their chance of scoring by at least 1% in each state in the middle third of the field can be done in the same way as for Manchester City, and thus by forcing them to avoid the center of their defensive third.

4.3 Evaluating the Effect of Adapting the Policy

Until now, we have explored the effect of forcing a team to avoid certain $areas$ on its chance of scoring. In practice, if an opponent enacts such a strategy, a team will eventually react and adapt their old policy π towards a new one π'.

A simple new policy π' will stop trying to reach locations in $area$, so $\pi'(move\_to(f')|f) = 0$ for all $f' \in area$ and the lost probability mass will be redistributed over all other states $f'' \in \mathcal{FS} \setminus area$ as follows:

$$\pi'(move\_to(f'')|f) = \frac{\pi(move\_to(f'')|f)}{1 - \sum_{f' \in area} \pi(move\_to(f')|f)} \tag{10}$$

By fixing the new policy π' in the team's MDP, we can reason about the effect on the chances of scoring while adapting to being forced to avoid $area$ as:

$$1 - \frac{\sum_{f \in \mathcal{FS}} V_{\pi'}(f)}{\sum_{f \in \mathcal{FS}} V_{\pi}(f)} \tag{11}$$

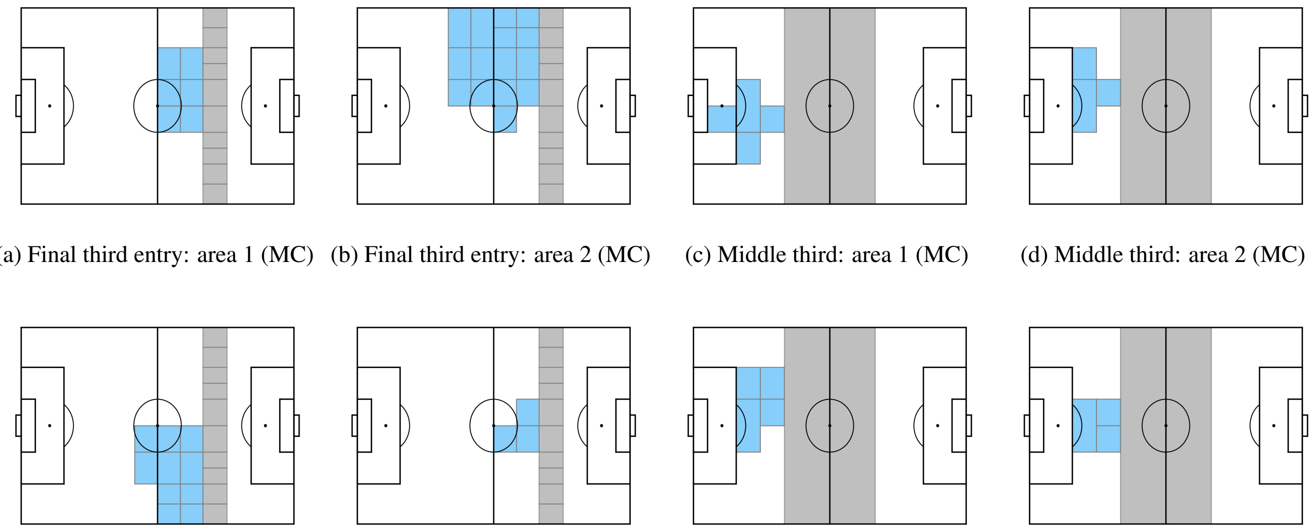
We illustrate the effect of Manchester City and Liverpool adapting their policies to an opponent's strategy. Forcing Manchester City to avoid the blue area in Figure 4a will decrease their chance of scoring by 16.9%. However, if they adjust their policy, the decrease is reduced to 3.9%. For Liverpool, forcing them to avoid the blue area in Figure 4e will decrease their chance of scoring by 12.3%. When they adjust their policy, this decrease is reduced to 4%. While the decrease is less impressive in the latter case for both teams, this still represents a reasonable reduction, certainly given that adapting one's strategy is hard.

5 Related Work

While Markov models have many applications in sports, the most prominent use is to objectively quantify a player's contributions during a match. The intuition is that a Markov model enables assessing how much a player's action increases her team's chance of scoring in the near future. Such models have begun to have a significant impact in professional soccer, where clubs (e.g., Liverpool)[2] and companies are employing them to help in areas such as player acquisition.

Using these models for tactical advice has received less attention. Some approaches have applied MRPs to analyze

[2]https://freakonomics.com/podcast/london-live/



(a) Final third entry: area 1 (MC) (b) Final third entry: area 2 (MC) (c) Middle third: area 1 (MC) (d) Middle third: area 2 (MC)

(e) Final third entry: area 1 (L) (f) Final third entry: area 2 (L) (g) Middle third: area 1 (L) (h) Middle third: area 2 (L)

Figure 4: Illustrates for Manchester City (top row) and Liverpool (bottom row) four areas (blue) to prevent them from reaching in order to decrease their chances of scoring in each final third entry state by at least 10% and in each middle third state by at least 1%.

set-pieces such as corners, free-kicks and throw-ins [Rudd, 2011], to identify where teams create value from by valuing their actions [Singh, 2019], and to determine the optimal times of substitutions and tactical changes [Hirotsu and Wright, 2002]. Other works use the optimal policy of an MDP and/or estimate the effect of changes to the policy to provide tactical advice [Sandholtz and Bornn, 2018; Sandholtz and Bornn, 2020; Van Roy *et al.*, 2021b; Wang *et al.*, 2018]. Our proposed MDP uses the same large action space for soccer as discussed in Van Roy et al. [2021b]. In contrast, our methods of providing tactical advice differ from the above mentioned approaches. Our work takes a *defensive* perspective to providing tactical advice, and aims to use probabilistic model checking to assess the effect of employing different defensive strategies.

In contrast to real-life soccer, MDP's have been extensively used in research on strategic reasoning and planning in (simulated) robot soccer. For example, an approach for providing coaching advice has been developed by Riley and Veloso [2004] and uses Q-learning to reason about a learned MDP. Additionally, the works of Ahmadi and Stone [2008] and Bai et al. [2012] investigated different automated action planning strategies for in-game decision making. Such automated planning strategies cannot immediately be used in real-life soccer as it is a highly unpredictable game. However, instead of finding the optimal plan, we aim to identify dangerous situations and, based on this, evaluate the effects of specific forced changes to a policy of a specific team. The resulting insights can immediately be used in the more general tactical planning by coaches.

Finally, applying verification techniques to sports models has not been extensively explored. Dong et al. [2015] have applied probabilistic model checking to a tennis MDP to predict the win probability and identify a player's best action to improve. Van Roy et al. [2021b] have applied probabilistic model checking to a soccer MDP to identify the best action regarding long-distance shooting. In contrast, our work is to the best of our knowledge the first to apply probabilistic model checking to inform a defensive game plan in soccer.

6 Conclusion

We have shown how machine learning techniques can learn a model that can be used to reason about goal-directed policies in the complex dynamic environment of professional soccer. We believe that our approach is also applicable to other environments. While there are no strong guarantees about the model's correctness as would be required in a verification context, it clearly supports reasoning about strategies and policies with respect to safety (i.e. reducing the chance of conceding). Furthermore, visualizing the results of the queries can help human soccer experts better understand the effects of potential strategies, which in turn contributes to trustworthy AI. From an application perspective, the proposed approaches can form a basis for future tactical analysis in sports.

Acknowledgments

This work was supported by the Research Foundation – Flanders under EOS No. 30992574 and the KU Leuven Research Fund (C14/17/070, C14/18/062). We thank StatsBomb for providing the event data used in this research.

References

[Ahmadi and Stone, 2008] Mazda Ahmadi and Peter Stone. Instance-based action models for fast action planning. In Ubbo Visser, Fernando Ribeiro, Takeshi Ohashi, and Frank Dellaert, editors, *RoboCup-2007: Robot Soccer World Cup XI*, volume 5001 of *Lecture Notes in Artificial Intelligence*, pages 1–16. Springer Verlag, Berlin, 2008.

[Bai *et al.*, 2012] Aijun Bai, Feng Wu, and Xiaoping Chen. Towards a principled solution to simulated robot soccer. In Xiaoping Chen, Peter Stone, Luis Enrique Sucar, and Tijn van der Zant, editors, *RoboCup-2012: Robot Soccer World Cup XVI*, volume 7500 of *Lecture Notes in Computer Science*, pages 141–153. Springer, 2012.

[Bellman, 1957] Richard Bellman. A markovian decision process. *Indiana University Mathematics Journal*, 6(4):679–684, 1957.

[Bellman, 1966] Richard Bellman. Dynamic programming. *Science*, 153(3731):34–37, 1966.

[Cervone *et al.*, 2016] Daniel Cervone, Alex D'Amour, Luke Bornn, and Kirk Goldsberry. A multiresolution stochastic process model for predicting basketball possession outcomes. *Journal of the American Statistical Association*, 111(514):585–599, 2016.

[Decroos *et al.*, 2019] Tom Decroos, Lotte Bransen, Jan Van Haaren, and Jesse Davis. Actions speak louder than goals: Valuing player actions in soccer. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1851–1861, New York, NY, USA, 2019. ACM.

[Dong *et al.*, 2015] Jin Song Dong, Ling Shi, Le Vu Nguyen Chuong, Kan Jiang, and Jing Sun. Sports strategy analytics using probabilistic reasoning. pages 182–185. IEEE, 2015.

[Goldner, 2012] Keith Goldner. A markov model of football: Using stochastic processes to model a football drive. *Journal of Quantitative Analysis in Sports*, 8:1–18, 2012.

[Hensel *et al.*, 2020] Christian Hensel, Sebastian Junges, Joost-Pieter Katoen, Tim Quatmann, and Matthias Volk. The probabilistic model checker storm, 2020.

[Hirotsu and Wright, 2002] Nobuyoshi Hirotsu and Mike Wright. Using a markov process model of an association football match to determine the optimal timing of substitution and tactical decisions. *Journal of the Operational Research Society*, 53(1):88–96, 2002.

[Howard, 1960] Ronald A. Howard. *Dynamic programming and Markov processes*. Wiley, London, 1960.

[Kwiatkowska *et al.*, 2011] Marta Kwiatkowska, Gethin Norman, and David Parker. PRISM 4.0: Verification of probabilistic real-time systems. In G. Gopalakrishnan and S. Qadeer, editors, *Proc. 23rd International Conference on Computer Aided Verification (CAV'11)*, volume 6806 of *LNCS*, pages 585–591. Springer, 2011.

[Liu *et al.*, 2020] Guiliang Liu, Yudong Luo, Oliver Schulte, and Tarak Kharrat. Deep soccer analytics: learning an action-value function for evaluating soccer players. *Data mining and knowledge discovery*, 2020.

[Riley and Veloso, 2004] Patrick Riley and Manuela Veloso. Advice generation from observed execution: Abstract Markov decision process learning. 2004.

[Routley and Schulte, 2015] Kurt Routley and Oliver Schulte. A markov game model for valuing player actions in ice hockey. In *Proceedings of the 31st Conference on Uncertainty in Artificial Intelligence*, pages 782–791, 2015.

[Rudd, 2011] Sarah Rudd. A Framework for Tactical Analysis and Individual Offensive Production Assessment in Soccer Using Markov Chains. In *New England Symposium on Statistics in Sports*, 2011.

[Sandholtz and Bornn, 2018] Nathan Sandholtz and Luke Bornn. Replaying the nba. In *Proceedings of the 12th Annual MIT Sloan Sports Analytics Conference*, 2018.

[Sandholtz and Bornn, 2020] Nathan Sandholtz and Luke Bornn. Markov decision processes with dynamic transition probabilities: An analysis of shooting strategies in basketball. *The annals of applied statistics*, 14(3):1122–1145, 2020.

[Schulte *et al.*, 2017] Oliver Schulte, Mahmoud Khademi, Sajjad Gholami, Zeyu Zhao, Mehrsan Javan, and Philippe Desaulniers. A markov game model for valuing actions, locations, and team performance in ice hockey. *Data Mining and Knowledge Discovery*, 31(6):1735–1757, 2017.

[Singh, 2019] Karun Singh. Introducing expected threat. https://karun.in/blog/expected-threat.html, 2019.

[Van Roy *et al.*, 2020] Maaike Van Roy, Pieter Robberechts, Tom Decroos, and Jesse Davis. Valuing on-the-ball actions in soccer: A critical comparison of xT and VAEP. In *Proceedings of the 2020 AAAI Workshop on AI in Team Sports*, 2020.

[Van Roy *et al.*, 2021a] Maaike Van Roy, Pieter Robberechts, Wen-Chi Yang, Luc De Raedt, and Jesse Davis. Learning a markov model for evaluating soccer decision making. In *RL4RealLife workshop at ICML 2021*, 2021.

[Van Roy *et al.*, 2021b] Maaike Van Roy, Pieter Robberechts, Wen-Chi Yang, Luc De Raedt, and Jesse Davis. Leaving goals on the pitch: Evaluating decision making in soccer. In *Proceedings of the 15th Annual MIT Sloan Sports Analytics Conference*, 2021.

[Wang *et al.*, 2018] Jiaxuan Wang, Ian Fox, Jonathan Skaza, Nick Linck, Satinder Singh, and Jenna Wiens. The advantage of doubling: A deep reinforcement learning approach to studying the double team in the nba. In *Proceedings of the 12th Annual MIT Sloan Sports Analytics Conference*, 2018.

[Yam, 2019] Derek Yam. Attacking contributions: Markov models for football. https://statsbomb.com/2019/02/attacking-contributions-markov-models-for-football/, 2019.