20th RoboCup International Symposium, Leipzig, Germany, July 2016.

Simulation Based Selection of Actions for a Humanoid Soccer-Robot

Heinrich Mellmann, Benjamin Schlotter, and Christian Blum

Department of Computer Science, Adaptive Systems Group
Humboldt-Universität zu Berlin, Germany
{mellmann,schlottb,blum}@informatik.hu-berlin.de
http://naoth.de

Abstract. This paper introduces a method for making fast decisions in a highly dynamic situation, based on forward simulation. This approach is inspired by the decision problem within the RoboCup domain. In this environment, selecting the right action is often a challenging task. The outcome of a particular action may depend on a wide variety of environmental factors, such as the robot's position on the field or the location of obstacles. In addition, the perception is often heterogeneous, uncertain, and incomplete. In this context, we investigate forward simulation as a versatile and extensible yet simple mechanism for inference of decisions. The outcome of each possible action is simulated based on the estimated state of the situation. The simulation of a single action is split into a number of simple deterministic simulations – *samples* – based on the uncertainties of the estimated state and of the action model. Each of the samples is then evaluated separately, and the evaluations are combined and compared with those of other actions to inform the overall decision. This allows us to effectively combine heterogeneous perceptual data, calculate a stable decision, and reason about its uncertainty. This approach is implemented for the kick selection task in the RoboCup SPL environment and is actively used in competitions. We present analysis of real game data showing significant improvement over our previous methods.

1 Introduction

A highly dynamic environment requires a robot to make decisions quickly and with limited information. In the RoboCup scenario, the robot that is in possession of the ball needs to take action as quickly as possible before the opponent players get a chance to interfere. However, the particular situation might be very complex and many aspects like the robot's position on the field as well as the positions of the ball and obstacles need to be taken into account. This makes inferring a decision a complicated task. In this work we propose an inference method based on forward simulation to handle this complexity and ensure short reaction times at the same time. We focus in particular on the RoboCup scenario where the robot has to choose the best kick from several different possibilities, which provides the motivation for our approach.

In the RoboCup community there have already been several attempts to implement similar methods. In particular [3, 4] and [1] focus on a very similar task – the selection of the optimal kick. In [3], a probabilistic approach is used to describe the kick selection problem which is then solved using Monte Carlo simulation. In [4], the kick is chosen to maximize a proposed heuristic *game situation score* which reflects the goodness of the situation. In [1], the authors use an instance based representation for the kick actions and employ Markov decision process as an inference method. Internal forward simulation has already been successfully used as an inference method in robotics. In [2], the authors investigate navigation of robots in a dynamic environment. They use a simulation approach to envision movements of other agents and pedestrians to enable avoiding dynamic obstacles while moving towards a goal. In [5] a pancake baking robot is planning its actions using a full physical simulation of the outcome of possible actions.

For an effective decision, data from heterogeneous sources (e.g., visual percepts, ultrasound) needs to be combined. Often different filtering/modeling techniques are used for state estimation, which can make inference of decisions a difficult task. In particular, representation of uncertainty is problematic. As we will show, the simulation based approach can handle it easily.

The intuition behind a simulation-based approach is to *imagine* (or simulate) what would happen as the result of the execution of a particular action and then choose the action with the optimal (imagined/simulated) outcome. A potential issue with this approach is that the quality of the decision depends on the quality of the simulation, i.e., the model of the environment. For example in [5, 6], the robots use complete fine-grained physical simulations for their decision-making. In contrast, we argue that the simulation itself can be quite coarse. To compensate for errors in the simulation, it is executed a number of times with varying initial conditions sampled according to the estimated state of the situation. Each of these realizations is evaluated individually and the overall decision for an action is then based on the distribution of the particular evaluations of the simulation. This is repeated for all possible actions (kicks) and the action with the best outcome distribution is chosen for execution.

We evaluate our approach based on labeled video and log data from real RoboCup competitions. The results show a significant improvement in comparison to our previous method.

The remainder of the paper is structured as follows. In the next section we discuss the action selection problem within the RoboCup domain. The main part of the paper consists of Section 3 and Section 4 where we describe the simulation and the evaluation-decision processes respectively. Our experimental findings are discussed in Section 5. Finally we conclude our findings in Section 6.

2 Action Selection Task in Robot Soccer

Consider the situation where a robot approaches the ball and needs to choose the right action from a fixed number of possibilities. In this study we assume the

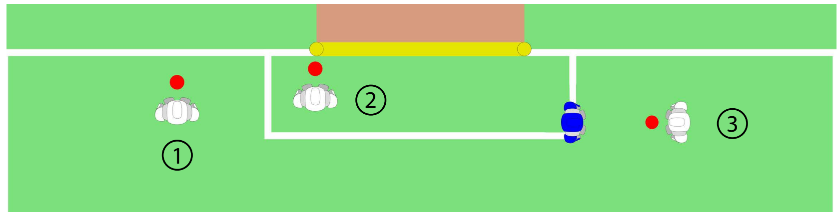

Fig. 1. Depictions of three different situations in which the best decision is not clear. The white robot is the robot having to take a decision on which (kick) action to perform while the blue robot is an opponent. The ball is depicted in red.

following possible actions: four different kicks, namely kick right, left, forward short (dribbling) and forward long, and a turn around the ball towards the opponent goal. The last option is to fall back in case no kick is possible, or if no kick would improve the situation.

To make an optimal decision different factors need to be taken into account. In our scenario we include estimated position of the robot on the field, position of the ball relatively to the robot, and obstacles in direct proximity. Each of these factors is modeled by a different probabilistic algorithm. We refer to the collective state estimated by these models as *situation state estimation.*

We approach this task using forward simulation. The outcome of each of the five actions is simulated using an estimated state of the situation, evaluated, and compared. The outcome of an action is described by the resulting position of the ball. Therefore, we need to model the interaction between the executing robot and the ball, the dynamics of the ball motion, and its possible interactions with the environment. In the following section these models will be discussed in detail.

3 Stochastic Forward Simulation

To be able to make decisions the robot needs an estimation of the state of the situation around it. In our case this state consists of the robot's position on the field, position of the ball relative to the robot, positions of the teammates and obstacles in close proximity. These particular aspects are usually estimated using various filtering techniques. In our case different independent probabilistic filters are involved, in particular particle filter for self localization and multi-hypothesis extended Kalman filter for the ball.

The task of the simulation process is to predict the state of the situation in case of the execution of a given action, e.g., kick. To do so, we need models for the effect of the action on the state of the situation, for the dynamics of particular objects and for interactions between the objects.

In general, an exhaustive physical simulation is a complicated and resource consuming process. To reduce complexity we make several assumptions. We focus only on simulating aspects involved in the action, i.e., the motion of the ball and its potential collision with obstacles and goals. We furthermore assume that all objects excluding the ball remain static.Though this is obviously not true, the velocity of the ball is usually much higher than that of the robots, which makes it a viable assumption in this case. To model collisions with obstacles, especially, goals we assume a fully nonelastic collision, where the ball's trajectory ends at the point of contact. With these assumptions we need to define the *dynamic model of the ball* and the *model for the effect of the kick on the ball*, which we discuss in the following two sections.

3.1 Ball Dynamics

To describe the dynamics of the ball motion we use a simple *rolling resistance* model which leads us to the following motion equation:

$$d(t) = -\frac{1}{2} \cdot g \cdot c_R \cdot t^2 + v_0 \cdot t \tag{1}$$

where $d(t)$ is the distance the ball has rolled after the time $t > 0$, c_R is the rolling resistance coefficient and v_0 is the initial velocity of the ball after the kick. By solving $d'(t) = 0$ and putting the result in eq. (1) the maximal rolling distance, i.e., the stopping distance of the ball, can readily be determined as

$$d_{max} = \frac{v_0^2}{2c_R \cdot g}. \tag{2}$$

The parameters v_0 and c_R of this model have to be determined experimentally. It should be noted that v_0 depends mainly on the particular kick motion and c_R depends mainly on the particular carpet of the field, since the ball remains the same. Thus, v_0 has to be estimated once for each kick motion and c_R once for each particular carpet.

3.2 Kick-Action Model

The result of a kick can be described by the likelihood of the final ball location after its execution, i.e., positions where the ball is expected to come to a halt eventually. These positions can be estimated based on the dynamic model of the ball as described in Section 3.1 and the intended direction of the kick. We assume the direction of the ball motion α and the initial velocity v_0 of the kick behaving according to the Gaussian distribution. With this, the outcome of a kick action can be described as a tuple of initial velocity v_0, direction α, and corresponding standard deviations σ_v and σ_α:

$$a = (v_0, \alpha, \sigma_v, \sigma_\alpha) \in \mathbb{R}_+ \times [-\pi, \pi) \times \mathbb{R}_+ \times [-\pi, \pi) \tag{3}$$

We predict the outcome of an action by sampling from the Gaussian distributions:

$$predict(a) := (d_{max}(\epsilon_v), \epsilon_\alpha) \in \mathbb{R}_+ \times [-\pi, \pi) \tag{4}$$

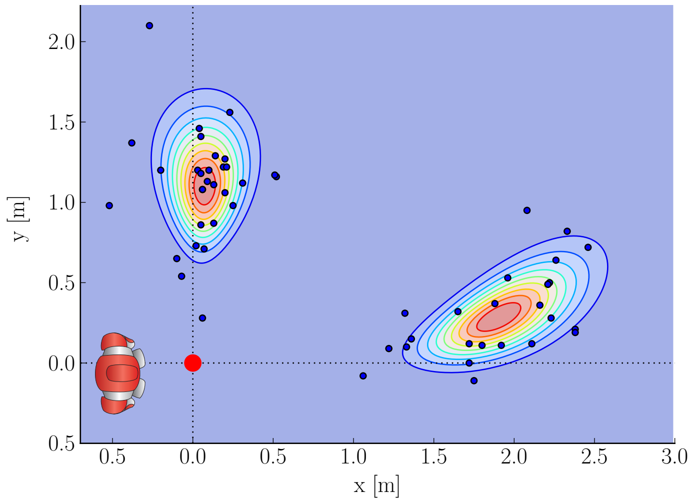

Fig. 2. Kick action model: distributions of the possible ball positions after a sidekick and the long kick forward with the right foot. Blue dots illustrate experimental data.

where $\epsilon_v \sim N(v, \sigma_v)$ and $\epsilon_\alpha \sim N(\alpha, \sigma_\alpha)$. Note that the function $predict(\cdot)$ is non-deterministic. Figure 2 illustrates the resulting likelihood for the final ball positions for a kick forward and a sidekick left. The parameters are estimated empirically.

3.3 Simulating the Consequences of an Action

The action is simulated a fixed number of times. The resulting ball position of one simulation is referred to as a *sample*. The positions of the samples are generated according to the model introduced in Section 3.2. The algorithm checks for possible collisions with the goal box and if there are any, the kick distance gets shortened appropriately. Collisions with the obstacle model are handled the same way.

A *hypothesis* for the action $a \in \mathcal{A}$ is defined as a set of $n \in \mathbb{N}$ samples drawn from the model distribution of an action a as described in Section 3.2.

$$\mathcal{H}_a := \{p_i | p_i = predict(a), i = 1 \dots n\} \subset \mathbb{R}_+ \times [-\pi, \pi) \tag{5}$$

4 Action Selection

The implementation of the simulation algorithm is divided into three main steps: simulate the consequences for all actions, evaluate the consequences, and decide the best action. In this section we discuss these components in the case of the kick selection as described in the Section 3.

4.1 Evaluation

The samples of each hypothesis are individually evaluated by two different systems. First, each sample $h \in \mathcal{H}_a$ is assigned a label

$$label(h) \in \mathcal{L} := \{\text{INFIELD}, \text{OUT}, \text{GOALOPP}, \text{GOALOWN}, \text{COLLISION}\} \quad (6)$$

based on where on the field it is, e.g., inside the field, inside the own goal, outside the field etc. These labels reflect the corresponding discrete rules of the game.

In the second step, all samples labeled $INFIELD$ are evaluated by a scalar potential field encoding the team strategy. An example of a potential field used in our experiments is described closer in section 4.3.

4.2 Decision

The overall decision has to take into account the trade-off between possible risks,e.g., ball leaving the field, and possible gains, e.g., scoring a goal, weighted by the chances of their occurrence. The estimation of those risks and gains can be done based on the individual ratings of the particular simulation results, i.e., samples. The likelihood of the occurrence of an event marked by a label $\lambda \in \mathcal{L}$ within a hypothesis $\mathcal{H}_a$ can be estimated as

$$p(\lambda|a) := \frac{|\{h \in \mathcal{H}_a | label(h) = \lambda\}|}{|\mathcal{H}_a|}. \quad (7)$$

For instance, the likelihood for scoring a goal with the action a can be written as $p(\text{GOALOPP}|a)$.

In our experiments we use a minimal two step decision process, whereby the actions that are *too risky* are discarded in the first step and the one with the highest gain is selected in the second. More precisely, we call an action *too risky* if there is a high chance for kicking the ball out of the field or scoring own goal. The set of actions with acceptable risk can be defined as:

$$\mathcal{A}_{acc} := \{a \in \mathcal{A} | p(\text{INFIELD} \cup \text{GOALOPP}|a) \geq T_0 \wedge p(\text{GOALOWN}|a) \leq T_1\} \quad (8)$$

with fixed thresholds T_0 and T_1 (in our experiments we used $T_0 = 0.85$ and $T_1 = 0$). Note that the cases indicated by OUT and $COLLISION$ are treated the same by this rule. From this set the actions with the highest likelihood of scoring a goal are selected

$$\mathcal{A}_{goal} := \text{argmax} \left\{ p(\text{GOALOPP}|a) | a \in \mathcal{A}_{acc} \right\}. \quad (9)$$

In case that $\mathcal{A}_{goal}$ is empty the default action is always turn around the ball. In case $\mathcal{A}_{goal}$ contains more than one possible action, the best action is selected randomly from the set of actions with the maximal strategic value based on the potential field

$$a_0 \in \text{argmax}\{value(a) | a \in \mathcal{A}_{goal}\} \quad (10)$$

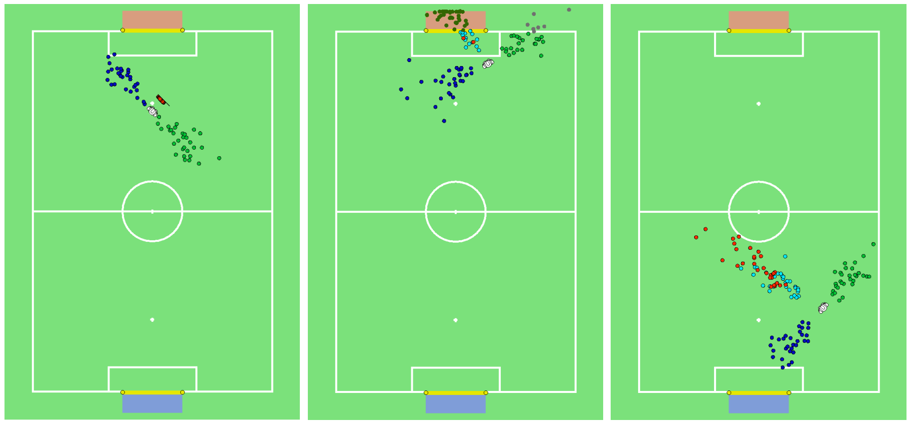

Fig. 3. Three examples for kick simulations. Each possible kick direction is simulated with 30 samples (different colors correspond to different kicks). Left: the short and long kicks are shortened due to collision with an obstacle. Middle: long kick is selected as the best action since it has the most samples result in a goal. Right: the best action is sidekick to the right – the other kicks are more likely to end up in a dangerous position for the own goal according to the potential field.

with strategic values defined as

$$value(a) := \int_{\Omega} p(x|a) \cdot potential(x)\ \mathrm{d}x = \frac{1}{n}\sum_{i=0}^{n} potential(x_i) \tag{11}$$

Figure 3 illustrates several situations with the corresponding simulated hypotheses and their evaluations.

4.3 Potential Field

A potential field assigns a value to each position of the ball inside the field. The values reflect the static strategy of the game and are used to compare possible ball positions in terms of their strategic value. For instance, the position a meter away in front of the opponent goal is obviously much better than the one in front of the own goal. In our experiments we use the following potential field:

$$P(x) = \underbrace{x^T \cdot \nu_{\text{opp}}}_{\text{linear slope}} - \underbrace{N(x|\mu_{\text{opp}}, \Sigma_{\text{opp}})}_{\text{opponent goal attractor}} + \underbrace{N(x|\mu_{\text{own}}, \Sigma_{\text{own}})}_{\text{own goal repulsor}}, \tag{12}$$

where $N(\cdot|\mu, \Sigma)$ is the normal distribution with mean μ and covariance Σ. It consists of three different parts: the linear slope points from the own goal towards the opponent goal and is modeling the general direction of attack; the exponential repulsor $N(x|\mu_{\text{own}}, \Sigma_{\text{own}})$ prevents kicks towards the center in front of own goal; and $N(x|\mu_{\text{opp}}, \Sigma_{\text{opp}})$ creates an exponential attractor towards the opponent goal.

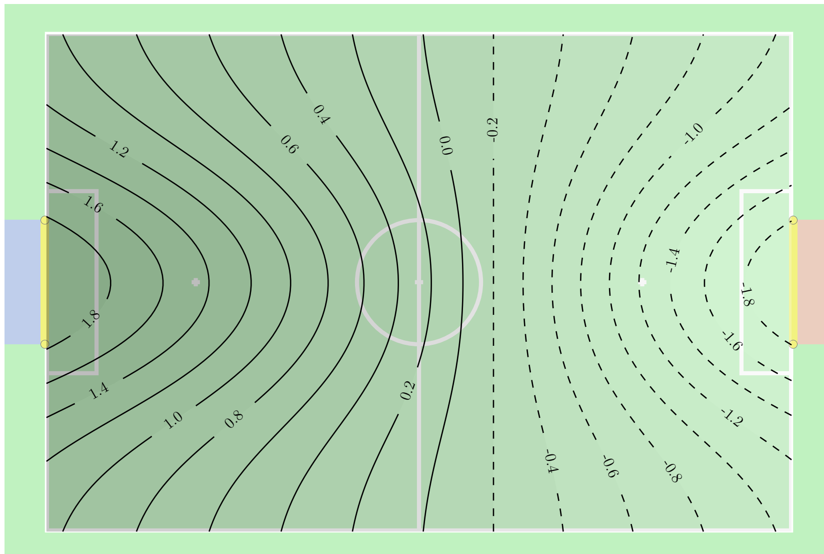

Fig. 4. Strategic potential field evaluating ball positions. Own goal is on the left (blue).

The configuration used in our experiments is

$$\nu_{\text{opp}} = (-1/x_{\text{opp}}, 0)^T \tag{13}$$

with $x_{\text{opp}} = 4.5$ being the x-position of the opponent goal and

$$\mu_{\text{own}} = (-4.5, 0) \qquad \mu_{\text{opp}} = (4.5, 0) \tag{14}$$

$$\Sigma_{\text{own}} = \begin{pmatrix} 3.375^2 & 0 \\ 0 & 1.2^2 \end{pmatrix} \qquad \Sigma_{\text{opp}} = \begin{pmatrix} 2.25^2 & 0 \\ 0 & 1.2^2 \end{pmatrix} \tag{15}$$

for the repulsor and attractor respectively. All parameters are of unit m. Figure 4 illustrates the resulting potential field.

4.4 Kick Selection Visualization

Figure 5 illustrates the decisions made by the algorithm depending on the robot's position on the field with the ball in front of the robot for three different fixed orientations of the robot. Since the simulation it stochastic, the decision is repeated 20 times for each cell on the field.

5 Quantitative Analysis in Real Game Situations

In general evaluation of decision algorithms is difficult because they tend to behave differently in the isolated environment of the lab than under real conditions, e.g., during a soccer competition. In this section we present analysis of the simulation based action selection using human labeled combined video and log data from real games.

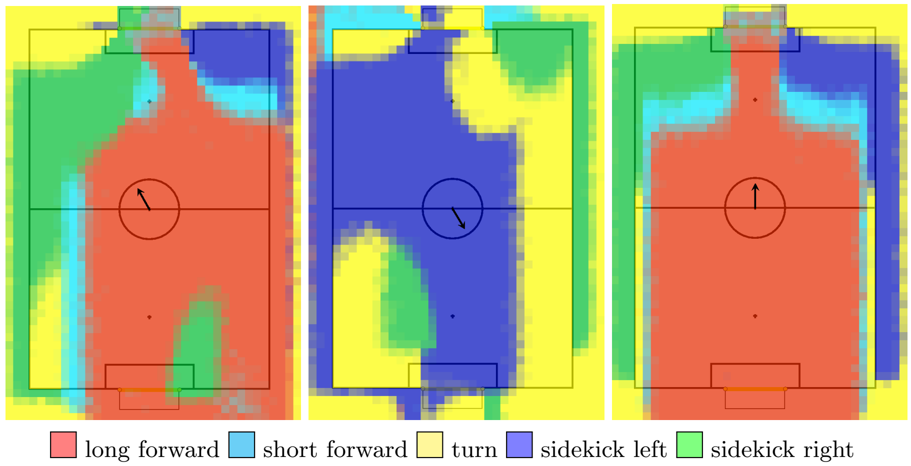

Fig. 5. Resulting decisions based on different positions of the ball on the field and three different orientations of the robot. Different colors correspond to differnt decisions. The orientation of the robot is indicated by the arrow. Own goal is at the bottom, opponent at the top.

5.1 Methodology

Evaluation of algorithms in real robot soccer competition conditions is a challenging task. This is mainly because in a real game many factors affect the performance of the robot in a particular situation, e.g., robot executes a wrong kick because it is not localized correctly. To minimize the influence of side factors on the evaluation we need to observe what actually happened and the internal state of the robot at the same time.

For this purpose we recorded videos overlooking the whole field of the games during RoboCup competitions in 2015 alongside with log files recorded by each of the robots. Video recordings provide a ground truth of the situation while log data recorded by the robots provides the corresponding internal state. The log files contain perceptions and the behavior decision tree for every cognition cycle (33 ms). This allows us to extract the situations when the robot took a decision to kick.

The logs have been synchronized with video files and the extracted kick actions manually labeled. The labeling procedure has been performed with the help of the interface which had been designed specifically for this purpose. Figure 6 illustrates an example of a labeling session for the first half of the game with the team *NaoDevils* at the RoboCup 2015.

The labeling criteria consist of 15 distinct boolean labels in three categories: technical execution of the kick, e.g., robot did miss the ball; situation model (was the estimation of robots position on the field and the ball correct?); result of the action and strategic improvement of the situation (ball left the field, was moved closer to the opponent goal etc.).

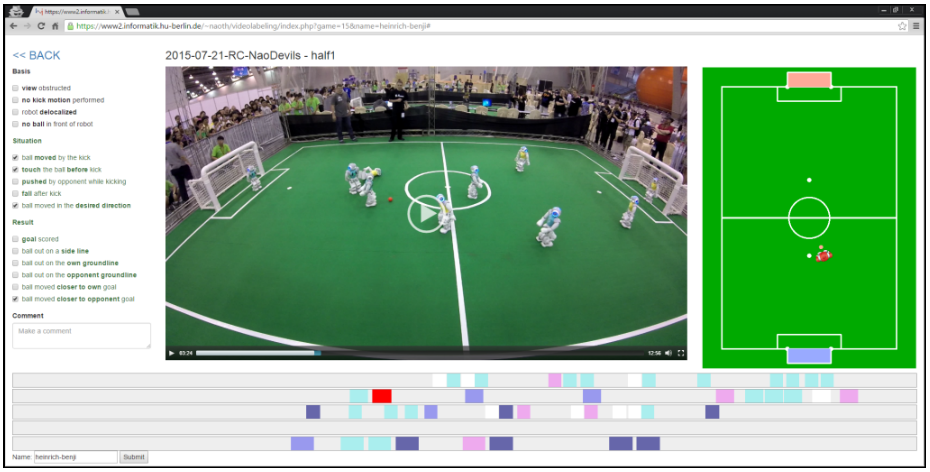

Fig. 6. Illustration of the labeling interface used to collect data regarding the quality of the kicks. At the bottom are time lines for each of the robots. Different actions are represented by buttons on the time line with different colors. On the right the robots estimated state is visualized, i.e., estimation of its position, ball model and obstacles. On the left are three categories of labels capturing the quality of the action.

5.2 Data Set

For our analysis we took a look at the games our team has played in two different competitions in 2015 – the *German Open 2015* (GO15) and *RoboCup 2015* (RC15). In both competitions our robots performed well – we reached the third place at the German Open and quarter finals at the RoboCup. At the GO15 we used our previous solution for action selection based on a manually adjusted heuristic decision tree and a potential field indicating the best direction towards the goal while at the RC15 the presented simulation based approach had been employed.

From GO15 a total of five game halves have been analyzed with: *ZKnipsers* (two halves, preliminaries); *HULKS* (first half, preliminaries); and *Nao Devils* (two halves, game for the 3rd place). And from RC15 we analyzed three complete games with: *RoboCanes* (two halves, preliminaries); *Nao Devils* (two halves, intermediate round); and *HTWK* (two halves, quarter finals). The selection of the games depends largely on the availability of the videos and log data.

5.3 Results

To single out the effect of the kick selection we focus on kicks where the robot was well *localized* (so it knew what it was doing) and kicks where executed successfully, i.e., the ball went in the intended direction and did not collide with opponent. In short: *successful* kicks are the ones which comply with our action

Algorithm	New		Old	
Total number of kicks	163		196	
Robot was localized	150	(92.02 %)	165	(84.18 %)
Successful execution	93	(57.06 %)	153	(78.06 %)
Failed execution	70		43	
Failed: opponent interference	33	(47.14 %)	14	(32.56 %)
Failed: technical failure	37	(52.86 %)	29	(67.44 %)
Successful execution + Localized	86	(52.76 %)	131	(66.84 %)
+1	67	**(77.91 %)**	88	(67.18 %)
0	15	**(17.44 %)**	39	(29.77 %)
-1	4	(4.65 %)	4	(3.05 %)
Out at opponent goal line	1	(**1.16 %**)	8	(6.11 %)

Table 1. Analysis results of video material. The new algorithm shows a higher rate of strategic improvements (+1) and a lower rate of mediocre kicks (0). It is also about 5 times less likely to kick out at the opponent field line.

model as described in section 3.2. The top part of the table 1 illustrates the numbers of the successful and failed kicks.

Our analysis has also revealed that a high percentage of the actions fail due to various reasons. The main reasons appear to be failure in the technical execution, e.g., the robot trips and doesn't kick the ball properly, and interference by opponent players. Both aspects are not part of the simulation and require further investigation. The table 1 (**Failed execution**) summarizes the rates of the failed kicks split in these two cases. The higher opponent interference in the case of the new approach can be explained by the more challenging opponent teams at the RC15.

In the lover part of the table 1 we summarize the evaluation of the the kick results according to the strategic improvement of the ball position as described in section 5.1. The separation used here is very rough: +1 corresponds to the cases where the strategic position of the ball was clearly improved by the action, e.g., it was moved closer towards the opponent goal; -1 was given when the ball moved towards own goal or away from the opponent goal; and 0 when no improvement was visible, e.g., ball moved along the middle line. The results show that the new approach results in a higher rate of improvements (+1) and a lower rate of mediocre kicks (0), while the rate of cases where the position of the ball worsened (-1) remained at a comparable level.

Another important factor is the *number of times the ball leaves the field* because it results in a tactical disadvantage as the ball is replaced into the field. The penalty is especially large when the ball leaves on the opponent goal line, since the ball is then reset to the middle line. In this case we can see a significant improvement with the new approach as only one kick (1.16 %) left the field at the opponent goal line in contrast to more than 6 % (8 kicks) with the old solution.

In summary, the data shows that the new approach performs more robustly than our previous solution. The new algorithm is about 5 times less likely to kick out at the opponent field line (decrease by 81 %) and 16 % more likely to kick towards the opponent goal.

6 Conclusions and Future Work

We presented and discussed an action selection algorithm based on forward simulation. We discussed its application in the scenario of kick selection for robot soccer. This kick selection algorithm was successfully implemented and used in RoboCup competitions. The three main advantages of the presented approach are easy implementation and extensibility. Experimental data collected in real RoboCup games has shown that the algorithm performs very well and is an improvement over the algorithm used by our team up to now.

Our current effort focuses in particular on stepwise extension to simulating the ball approach and more dynamic evaluation. For instance, the potential field might reflect the influence regions of the own teammates based on their position, which would favor the kicks towards these regions and enable emergent passing.

At the present state the implemented method is limited to the selection of the kicks only. We believe that the true potential of the forward simulation can only unfold if extended to all areas of decision making like role decision, passing, positioning etc.

References

1. Ahmadi, M., Stone, P.: RoboCup 2007: Robot Soccer World Cup XI, chap. Instance-Based Action Models for Fast Action Planning, pp. 1–16. Springer Berlin Heidelberg, Berlin, Heidelberg (2008), http://dx.doi.org/10.1007/978-3-540-68847-1_1
2. Bordallo, A., Previtali, F., Nardelli, N., Ramamoorthy, S.: Counterfactual reasoning about intent for interactive navigation in dynamic environments. In: Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on. pp. 2943–2950 (Sept 2015)
3. Dodds, R., Vallejos, P., Ruiz-del Solar, J.: Probabilistic kick selection in robot soccer. In: Robotics Symposium, 2006. LARS '06. IEEE 3rd Latin American. pp. 137–140 (Oct 2006)
4. Guerrero, P., Ruiz-Del-Solar, J., Díaz, G.: Robocup 2007: Robot soccer world cup xi. chap. Probabilistic Decision Making in Robot Soccer, pp. 29–40. Springer-Verlag, Berlin, Heidelberg (2008), http://dx.doi.org/10.1007/978-3-540-68847-1_3
5. Kunze, L., Beetz, M.: Envisioning the qualitative effects of robot manipulation actions using simulation-based projections. Artificial Intelligence pp. – (2015), http://www.sciencedirect.com/science/article/pii/S0004370214001544
6. Winfield, A.F.T., Blum, C., Liu, W.: Advances in Autonomous Robotics Systems: 15th Annual Conference, TAROS 2014, Birmingham, UK, September 1-3, 2014. Proceedings, chap. Towards an Ethical Robot: Internal Models, Consequences and Ethical Action Selection, pp. 85–96. Springer International Publishing, Cham (2014), http://dx.doi.org/10.1007/978-3-319-10401-0_8