A Survey on Coordination Methodologies for Simulated Robotic Soccer Teams

Fernando Almeida*‡, Nuno Lau†‡, Luís Paulo Reis§¶
falmeida@di.estv.ipv.pt, lau@det.ua.pt, lpreis@fe.up.pt

*DI/IPV - Department of Informatics, Polytechnic Institute of Viseu, Viseu, Portugal
†DETI/UA - Electronics, Telecommunications and Informatics Department, University of Aveiro, Aveiro, Portugal
‡IEETA - Institute of Electronics and Telematics Engineering of Aveiro, Aveiro, Portugal
§DEI/FEUP - Department of Informatics Engineering, Faculty of Engineering, University of Porto, Porto, Portugal
¶LIACC - Artificial Intelligence and Computer Science Laboratory, University of Porto, Porto, Portugal

***Abstract*—Multi-agent systems (MAS) are a research topic with ever-increasing importance. This is due to their inherently distributed organization that copes more naturally with real-life problems whose solution requires people to coordinate efforts.**

One of its most prominent challenges consists on the creation of efficient coordination methodologies to enable the harmonious operation of teams of agents in adversarial environments. This challenge has been promoted by the Robot World Cup (RoboCup) international initiative every year since 1995.

RoboCup provides a pragmatic testbed based on standardized platforms for the systematic evaluation of developed MAS coordination techniques. This initiative encompasses a simulated robotic soccer league in which 11 against 11 simulated robots play a realistic soccer game that is particularly suited for researching coordination methodologies.

This paper presents a comprehensive overview of the most relevant coordination techniques proposed up till now in the simulated robotic soccer domain.

***Index Terms*—Coordination methodologies, MAS, simulated robotic soccer, RoboCup.**

I. INTRODUCTION

The development of efficient methodologies (e.g. languages, models) for MAS coordination in adversarial environments is one of the most interesting scientific challenges promoted by the RoboCup [33] and is mainly supported by its soccer simulation leagues. The main goal of coordination mechanisms in these leagues is to adequately control a team of players and an optional coach to win matches against adversary teams.

Soccer is an inherently coordinated game in which team fitness directly relates to how well players can synchronize to perform tasks (e.g. passing). However, team coordination can be complex to achieve, mostly due to the multitude of variables (e.g. players and ball positions) players must consider to make the best decision at each instant. Moreover, measuring its success quantitatively is difficult as it doesn't necessarily relate to the final match score (e.g. a team might play better than the opposite but still lose), thus more data must be considered to perform an accurate assessment (e.g. ball possession).

The rest of the paper is organized as follows. Section II describes the RoboCup initiative and its physical soccer simulator. Section III presents a general definition of coordination and its related issues in the robotic soccer domain. Sections IV, V, VII and VI provide a discussion of developed techniques for simulated robotic soccer organized in different perspectives. Section VIII addresses the lessons learned from the survey.

II. ROBOCUP: A TESTBED FOR COORDINATION

RoboCup was designed to meet the requirements of handling real complexities in a restricted world and provides standard challenges in a common platform to foster Artificial Intelligence and Intelligent Robotics research [17].

Its most pragmatic goal is to develop a team of fully autonomous humanoid robot soccer players capable of winning a soccer game against the winner of the World Cup by 2050. This ambition although difficult to achieve, will surely drive significant technological breakthroughs while trying [33].

The main focus of RoboCup is Robotic Soccer (RoboCup-Soccer), although other application domains exist focusing on different scopes like disaster rescue, robotics education for young students and human assistance on everyday life tasks.

The RoboCupSoccer domain has 5 leagues [11]: there is a virtual (Simulation League) and several hardware (Small-Size, Medium-Size, Standard Platform and Humanoid) leagues.

This paper focuses on the RoboCupSoccer 2D Simulation League (RoboCupSoccer2D) although other simulation subleagues (3D, 3D Development and Mixed Reality) exist. This league enables a virtual soccer match between 2 teams of 11 simulated agents each with an optional online coach using a physical soccer simulation system. Agents have an environment-aware body and can act autonomously to perform reactive or pro-active actions in an individual or sociable manner, although interaction is highly constrained as described in Section III. The environment is partially observable through non-symbolic sensors, stochastic, sequential, dynamic and multi-agent without centralized control [11].

This league presents 3 strategic research challenges for multi-agent interaction [33]:

- Multi-agent learning of individuals (e.g. ball interception) and teams (e.g. adapt player positioning to opponents);
- Teamwork to enable to real-time planning, replanning and execution of tasks in a dynamic adversary environment;
- Agent modelling to reason about others (e.g. intentions).

TABLE I
LIST OF SOCCER SERVER CORE ACTIONS BY CATEGORY

Category	Actions
Movement	Dash*, Turn, Move
Ball control	Kick, Catch, Tackle
Perception control	Turn neck, Change view, Attention to
Communication	Point to, Say
Match information	Score

*Dash impacts players stamina which is continuously assessed through their energy (liveness), effort (movement efficiency) and recovery (energy renewal rate)

Soccer Server is an open-source client/server physical soccer simulation system [36][7] used in RoboCupSoccer2D. It uses well defined protocols to enable communication between clients (players and coaches) and itself to manage connections, gather world perceptions and control clients actions.

Firstly, all clients connect to the server and sending introductory initialization data to which the server replies with the current simulation settings (e.g. player characteristics). These settings can be tweaked in order to enhance the simulation.

During the match, each team can have an online coach that receives global error-free information about world objects and all the messages sent from the players and the referee. All communication is done exclusively via the server and coach-to-players communication is highly restricted.

The simulator provides a set of players with distinguished capabilities (heterogeneous players) from which the coach must build a team to play a soccer match. During the match players receive tailored multimodal sensor information (aural, vision and body) according to their standpoint. This information is received through messages (hear, see and sense body) sent regularly from the simulator, that can be inaccurate (e.g. vision accuracy varies inversely with objects distance). Based on these perceptions, players can act upon the world to inflict changes in it using the core actions depicted in Table I.

Also during the match, a referee (automated or human) can make rulings that change the play mode (e.g. free-kick) and are immediately relayed to all clients. The human referee is used to judge situations driven by player's intentions (e.g. player obstruction) which are still difficult to evaluate automatically.

The simulation executes in discrete time steps (cycles). Throughout each step players can take actions, restricted in number and by play mode (e.g. one kick per cycle), that will be applied to objects (players and the ball) at the end of the step. The next step is simulated by applying only the allowed actions to the state information (e.g. update objects positions) and eventually by solving conflicting situations (e.g. several players might kick the ball simultaneously).

Some of the research developed has shown that robotic soccer [1] and consequently RoboCup [35][34] can be used effectively to study MAS and coordination techniques in particular. In most cases these techniques can be generalized to other domains [6] (e.g. network routing [53]).

III. COORDINATION PROBLEMS IN SIMULATED ROBOTIC SOCCER

Robotic Soccer is an instance of Periodic Team Synchronization (PTS) domains [52] in which players have sporadic opportunities to communicate fully in a safe offline situation (e.g. in the locker-room) while being able to act autonomously in real-time with little or no communication.

One of the most important tasks for players is to select and initiate an appropriate (possibly cooperative) behavior in a given context, using (or not) knowledge from past experiences in order to help their team to win. Good coordination methodologies can help achieve this goal, although their success is still highly dependent on players individual abilities (low-level skills) to execute adequate competitive decisions.

The coordination difficulties enforced by the simulator are:

- Many multimodal information can be sensed at once, making it difficult to process;
- Environment's unpredictability makes it difficult to predict future states;
- Clients can't rely on message reception due to communication unreliability;
- Low-bandwidth makes it difficult to convey meaningful knowledge in messages;
- Uncertainty in perceived world information may lead to conflicting behaviors between agents [39], due to invalid state knowledge representations.

More specifically the simulated robotic soccer domain presents researchers with the following types of challenges:

- Perception: Where, when and how should players use their vision? To whom should they listen to? How to estimate information of others?
- Communication: What, when and how should players exchange information? How should exchanged information be used?
- Action: Which action should the player perform that is best for the team? How to evaluate different types of actions (e.g. pass vs dribble)? How to execute a given elementary (e.g. kick) or compound action (e.g. dribble)?
- Coordination: How to structure coordination dependencies between players? With whom should a player coordinate his actions? How should actions be coordinated with others? How to adapt coordination in real-time? How can the coach be used to coordinate team players?

The answer to some of these questions and others more specific will be discussed in the remaining sections.

IV. TECHNOLOGIES FOR COORDINATION

A. Coordination by Communication

Sharing pertinent world information can be useful to achieve team coordination. In earlier Soccer Server versions communication constraints were relaxed and allowed the transmission of long messages. This extremely permissive condition motivated the development of techniques that relied on sharing lots of meaningful information about the world's state knowledge among teammates to make better informed decisions.

Currently, message size is restricted to a minimum and poses a new challenge that requires the cautious selection of pertinent information to convey at each instant. To circumvent the previous constraint an Advanced Communications framework [42] was proposed in which a player maintains a communicated world state (separated from his perceived world state) using only information from teammates, without any prediction or perception information of his own. By comparing both worlds, a player assesses the interest of items of his perceived world state to his teammates and selects the most useful information (e.g. objects positions) to share. Information utility metrics were based on domain-specific heuristics but were later extended to accommodate the current situation and estimated teammate's knowledge [12].

Other techniques were proposed that use little or no communication by adding knowledge assumptions (e.g. Locker-Room Agreements discussed in Section VI-A) to reason over players intentions based on assigned roles [20] (combined with Coordination Graphs discussed in Section VII-A), offline learned prediction models [54] and player's beliefs [38][16] to adapt to their actions.

The trend in this domain will be towards little or no communication due to the constraints mentioned in Section III and also because communication introduces an overhead and delay that can degrade the player performance. The combination of implicit coordination with beliefs exchange yields better performance with communication loss than explicit coordination with intentions communication alone [16]. The exchange of beliefs among teammates allows a more coherent and complete global belief about the world. This global belief can then be used to predict players utilities and adapt actions to players predicted intentions to achieve the best (joint) action. As state estimation accuracy reaches an acceptable upper bound it will eventually replace explicit communication.

B. Coordination by Intelligent Perception

The smart usage of player sensors can be an efficient way to leverage coordination with other players, by collecting the most valuable information at each instant.

During the match players can assume three types of visualizations. These are chonse using a strategic looking mechanism based on their internal world state information and the current match situation [42]:

- Ball-centered: look at the ball to react quickly to its sudden velocity changes (e.g. kick by a player);
- Active: look at the target location of a desired action (e.g. a pass to perform);
- Strategic: look at a strategic location to improve the world's state accuracy (e.g. find an open space for a pass).

The usefulness of the information gathered using the previous approaches is different and can be classified based on its intended usage scope, validity over time and motivation for player behavior in future actions as depicted in Table II.

Ultimately, this information can be combined to enhance the player's world state accuracy and empower better decisions.

TABLE II
COMPARISON OF DIFFERENT VISUALIZATION APPROACHES

Approach	Usage scope	Information validity period	Target behavior
Ball-centered	Individual	Short	Reactive
Active	Individual or Collective	Short to Medium	Reactive or Deliberative
Strategic	Collective	Medium to Long	Deliberative

V. POSITIONING

A. Coordination for General Positioning

The selection of a good position to move into during the match is a challenging task for players due to the unpredictable behavior of other players and the ball. The likelihood of collaboration in a soccer match is directly related to the adequacy of a player's position (e.g. open pass lines for attack).

During a match, at most one player can carry the ball at each instant. For this reason, players will spend most of their time without the ball and trying to figure out where to move.

The first positioning techniques proposed allowed players to situate themselves in an anticipated useful way for the team in two different contexts [48]:

- Opponent marking: player moves next to a given opponent rather than staying at his default home position;
- Ball-dependent: player adjusts his location, within a given movement range, based on the ball's current position;
- Strategic Positioning using Attraction and Repulsion [48] (SPAR): player tries to maximize the distance to all players and minimize the distance to the opponent goal, the active teammate and the ball. This algorithm enables players to anticipate the collaborative needs of their teammates by positioning themselves to open pass lines for the teammate with the ball.

The previous techniques are rather reactive and demand fast responses from players according to the target object behavior. This leads to quickly wearing out stamina because the current match situation ins't adequately considered. To solve these issues, techniques were proposed that distinguish between active (e.g. ball possession) and strategic match situations [42]:

- Simple Active Positioning: players always assume an active and non-strategic position (e.g. ball recovery);
- Active Positioning with Static Formation: extends the previous so that players can return to their default home position in the static formation, if there isn't a good enough active action to perform;
- Simple Strategic Positioning: uses only one situation and one dynamic formation;
- Situation Based Strategic Positioning [44] (SBSP): defines team strategy as a set of player roles (defining their behavior) and a set of tactics composed of several formations. Each formation is used for a different strategic situation and assigns each player a default spatial positioning and a role. Contrarily to SPAR, it allows the team to have completely diverse but suitable shapes (e.g. compact for defending) for different situations and teammates to have different positional behaviors;
- Delaunay Triangulation (DT): similar in idea to SBSP, it divides the soccer field into triangles according to training data [2] and builds a map from a focal point (e.g. ball position) to a desirable positioning of each player. It

also allows the use of constraints to fix topological relations between different sets of training data to compose more flexible team formations, Unsupervised Learning Methods (e.g. Growing Neural Gas) to cope with large or noisy datasets and Linear Interpolation methods (e.g. Goraud Shading) to circumvent unknown inputs. Despite its simplicity, DT has a good approximation accuracy, is locally adjustable, fast running, scalable and can reproduce results for identical training data. On the other hand it requires much memory to store all training data and has a high cost to maintain its consistency.

Another task addressed in a soccer match is the dynamic (or flexible) positioning of team players that consists on switching players positions within a formation [48] to improve the team's performance (e.g. save player's energy for quicker responses). However, if misused it can increase player's movement (e.g. player moves across the field to occupy its new position).

The methods proposed to aid players weigh the cost/benefit ratio for deciding to switch positions are based on:

- Role Exchange: continuously assesses the usefulness of exchanging positions based on tactical gains [42] (e.g. distance to a strategic position, adequacy of next versus current position and coverage of important positions). It extends previous work that used flexible player roles with protocols for switching among them [52] to accommodate the exchange of players positions and types in the formation and has been used in conjunction with SBSP;
- Voronoi Cells: distributes players across the field and uses Attraction Vectors to reflect players' tendency towards specific objects based on the current match situation and players' roles [8]. It claims to have solved a few restrictions in SBSP (e.g. obligation to use home positions and fixed number of players for each role);
- Partial (Approximate) Dominant Regions [31]: divides the field into regions based on the players time of arrival (similar to a Voronoi diagram based on the distance of arrival), each of which shows an area that players can reach faster than others. It has been used for marked teammates to find a good run-away position.

B. Defensive Coordination

The main goal of a defending team, without ball possession, is to stop the opponent's team attack and create conditions to launch their own. In general, defensive behaviors (e.g. marking) involve positioning decisions (e.g. move to intercept the ball). Defensive positioning is an essential aspect of the game, as players without the ball will spend most of their time moving somewhere rather than trying to intercept it.

Collaborative defensive positioning has been described as a multi-criteria assignment problem where n defenders are assigned to m attackers, each defender must mark at most one attacker and each attacker must be marked by no more than one defender [23]. The Pareto Optimality principle was used to improve the usefulness of the assignments by simultaneously minimizing the required time to execute an action and the threat prevented by taking care of an attacker [24]. Threats are considered preemptive over time and are prevented using a heuristic-criterion that considers:

- Angular size of own goal from the opponent's location;
- Distance from the opponent's location to own goal;
- Distance between the ball and opponent's location.

This technique can achieve good performances while balancing gracefully the costs and rewards involved in defensive positioning, but it doesn't seem to deal adequately with uneven defensive situations:

- Outnumbered defenders shouldn't mark specific attackers but rather position themselves in a way that difficults their progression towards to the goal's center;
- Outnumbered attackers: more than one defender should mark an attacker (e.g. ball owner) pursuing a strategy to quickly intercept the ball or compel the opponent to make a bad decision and lose the ball.

Marking consists on guarding an opponent to prevent him from advancing the ball towards the goal, making a pass or getting the ball. Its goal is to seize the ball and start an attack.

The opponent to mark can be chosen by the player (e.g. closest opponent), by the team captain following a preset algorithm (e.g. as part of the Locker-Room Agreement [48] discussed in Section VI-A), using matching algorithms [47] or Fuzzy Logic [46]. Choosing the opponent to mark based only on its proximity isn't suitable as it disregards relevant information (e.g. teammates nearby) and will lead to poor decisions. Also, the use of a fixed centralized mediator (e.g. coach) to assign opponents to teammates although faster to compute has a negative impact in players autonomy. With the exception of PTS periods, this approach isn't robust enough due to the communication constraints mentioned in Section III and because it provides a single point of failure.

A Neural Network trained with a back-propagation algorithm that uses a linear transfer function was proposed to decide the type of marking to perform based on the distance from the player to ball, the number of opponents and teammates within the player's field of view (FoV) and the distance from the player to his own goal [46]. The output accuracy of this method could be improved by considering other relevant information that lies outside the player's FoV (e.g. nearby opponents behind the player).

Aggressive marking behavior can also be learned using a NeuroHassle policy [14] based on a neural network trained with a back-propagation variant of the Resilient Propagation (RPROP) reinforcement learning technique.

C. Offensive Coordination

To improve position selection during offensive situations (e.g. the team owns the ball) players should find the best reachable position to receive a pass or score a goal.

The Pareto Optimality Principle was applied to enable systematic decision-making regarding offensive positioning [25] based on the following set of partially conflicting criteria for simultaneous optimization [41]:

- Players must preserve formation and open spaces;

- Attackers must be open for a direct pass, keep an open path to the opponent's goal and stay near the opponent's offside line to be able to penetrate the defense;
- Non-attackers should create chances to launch the attack.

A Simultaneous Perturbation Stochastic Approximation (SPSA) combined with a RPROP learning technique (RSPSA) was proposed to Overcome the Opponent's Offside Trap (OOOT) by coordinated passing and player movements [13]. The receiver of the OOOT pass should start running into the correct direction at the right point in time, preferably being positioned right before the offside line while running at its maximal velocity when the pass is executed.

VI. TEAM COORDINATION

A. Coordination for Strategic Actions

In real soccer, team strategies are rehearsed during mundane training of team players and applied during a match. The same strategies are often used in matches, but for some opponents they must be swapped to adapt to their unexpected behavior.

Strategies typically consist on a set of tactics composed by formations that map a strategic position and a distinguished role to each player to guide his behavior.

To deal with the challenges of PTS domains a Locker Room Agreement (LRA), based in the definition of a flexible team structure (consisting of roles, formations and set-plays), can be used for players to consent on globally accessible environmental cues as triggers for changes in strategy [48]. Team strategies are communicated with a timestamp for players to recognize changes and always keep the most recent ones to disseminate to others. The team's formation can be either static or change dynamically during the match on team synchronization opportunities (e.g. kick-in) or via triggered-communication where one teammate (e.g. team captain) makes a decision and broadcasts it to his teammates.

Set-plays are predefined plans for structuring a team's behavior depending on the situation. A high-level generic and flexible framework that defines a language for set-play definition, management and execution was proposed in [29]. A set-play involves players' references (individual or role based) and steps (states of execution) that can have conditions to be carried out. Each step is lead by the ball carrying player (in charge of making the most important decisions) and can have several transitions (possibly with conditions) for subsequent steps. The main transition of a step defines a list of directives consisting of actions that should (or not) be performed. The execution of a set-play requires a tight synchronization between all participants to enable a successful cooperation. To cope with the simulator communication restrictions, only the lead player is allowed to send messages. This technique could be improved to achieve implicit coordination through a kind of belief state exchange, because the player that owns ball decides when to start the set-play and informs the involved parties. From that moment on and while the set-play follows its default path, communication among players could be dropped until a deviation is decided by the ball owner because all involved parties know the steps.

Another method proposed for high-level coordination and description of team strategies is Hierachical Task Network (HTN) planning [37] which is to be embedded in each player. It combines high level plans (making use of previous domain knowledge to speed up the planning process) with reactive basic operators, so that players can pursue a global strategy while staying reactive to changes in the environment. This method separates the expert knowledge specified as team strategies from the player implementation making it easier to maintain. The objective of HTN is to perform tasks which can be either complex or primitive. Complex tasks are expanded into subtasks until they become primitive.

B. Hierarchical Coordination

In real life soccer, natural hierarchical relations exist among different team members and imply a leadership connotation (e.g. a coach instructs strategy to players).

A coach and trainer are privileged agents used to advise players during online games and offline work out (training) situations respectively. The need of communication from coach to players motivated the definition of coaching languages.

CLang [7] is the standard coaching language used in RoboCup since 2001 to promote a new RoboCup competition focused only on coaching techniques, but it lacks the ability to specify a team's complete behavior with sufficient detail.

Coach Unilang [43] was proposed to enable the communication of behavioral changes to players during games using different kinds of strategic information (e.g. instructions, statistics, opponent's information and definitions) based on real soccer concepts. Players can ignore received messages, interpret them as orders (must be used and will replace knowledge) or as advices (can be used with a given trust level).

Strategy Formalization Language [32] extends CLang by representing team behavior in a human-readable format easily modifiable in real-time by abstracting low-level concepts.

The main coaching techniques developed make use of:

- Neural Networks (previously trained with adequate data) to recognize opponent's team formation and provide appropriate counter formation to players [55];
- Matching Algorithms that continuously builds a table that assigns a preliminary opponent to mark to each teammate and briefs all players periodically [47].

The ability to recognize tactics and formations used by opponent teams reveals part of their strategy and can be used to implement counter strategies. To address this opportunity training techniques make use of:

- Sequential Pattern Data Mining using Unsupervised Symbolic Learning of Prediction Rules for situations and behavior during matches [26];
- Triangular Planar Graphs to build topological structures for discovering tactical behavior patterns [40].

VII. LOCAL COORDINATION

A. Coordination for Action Selection

Deciding what the player should do at a given moment in a soccer game is critical. Player's individual decision should depend on the actions performed (or expected) of other players and balance their risks and rewards. However, these dependencies can change rapidly in dynamic environment as a result of the continuously changing state, thus efficient and scalable methods must be developed to solve this issue.

The action selection mechanisms proposed make use of:

- An idealized world model combined with observed player's state information to predict the best action [50];
- An option-evaluation architecture for different actions with comparable probabilistic scores [49];
- Player roles and a measurement opponents interference in the current situation using a multi-layer perceptron [18];
- Coordination Graphs (CGs) [19] where each node represents a player and its edges (possibly directed) define dependencies between nodes that have to coordinate their actions. This approach is based on the assumption that in most situations only a few players (typically nearby) need to coordinate their actions, while the remaining are capable of acting individually. To solve coordination dependencies in CGs algorithms like Variable Elimination (VE) [17], Max-Plus (MP) [21] and Simulated Annealing (SA) [9] were proposed. VE requires communication to always find an optimal solution but only upon termination and with a high computational cost (due to its action enumeration behavior for neighbors). MP solves VE high computational cost and makes the solution available at anytime, but it can only find near optimal solutions (except for tree-structured CGs) and restricts coordination to pairs of players. SA improves MP being able to work without communication and not restricting coordination between pairs, but it can only find approximate solutions with an associated confidence;
- Fuzzy logic and bidirectional neural networks to determine the odds and priorities of action selection based on human knowledge [57];
- Case-Based Reasoning to explicitly distinguish between controllable and uncontrollable indexing features, corresponding to players positions [45].

B. Coordination for Behavior Acquisition

Teams often use flexible (to some extent) predefined strategies set on the LRA. However they can prove fruitless, when playing against opponents that exhibit incompatible behaviors. Modelling the opponent's behavior thus becomes a necessity to allow convenient adaptation. However, as most players' are unseen for quite some time this task becomes a challenge.

With adequate models of players behavior, a player can improve his world model accuracy and consequently make better decisions by anticipating collaborative needs of teammates (e.g. open a line of pass).

Machine learning techniques have been proposed to address the issue of player adaptation to unforeseen situations [3][1].

Layered learning [48] has been proposed to enable learning low-level skills and ultimately use them to train higher-level skills that can involve coordination. The highest layer of the previous approach uses a Team-Partitioned Opaque-Transition Reinforcement Learning (TPOT-RL) technique to allow team players to learn effective policies and thus cooperate to achieve a specific goal. This technique divides the learning task among teammates, using coarse action-dependent features and gathers rewards directly from environmental observations. It is particularly suitable for this domain which presents huge state spaces (most of them hidden) and limited training opportunities.

Policy gradient RL was proposed to coordinate decision making between a kicker and a receiver in free-kicks [30][15].

Two other important subtasks of a soccer game, Keepaway and Breakaway, have been used to study specific behavioral coordination issues. Keepaway is a game situation where one team (the keepers), tries to maintain ball possession within a limited region, while the opposing team (the takers) attempt to gain possession. Breakaway is another game situation with the purpose of the attackers trying to score goals against defenders. RL techniques have proven its their usefulness to improve decision-making in these tasks [28][51]. The recognition of the potential for RL techniques, lead to the proposal of the following methods to accelerate them:

- Preference Knowledge-Based Kernel Regression (KBKR) to give advice about preferred actions [28];
- Heuristic Accelerated Reinforcement Learning (HARL): using predefined heuristic information based on Minimax-Q [4] and Q-Learning [6];
- Case Based-HARL: heuristics are derived from a case base using Q-Learning [5].

C. Ball Passing Coordination

Passing is a crucial skill in soccer and it reflects the cooperative nature of the game. Without sophisticated passing skills, it will be difficult for a team to win a match. The number of passing possibilities for the ball carrying player can be overwhelming and thus efficient methods must be employed for real-time decision-making.

The main criteria used to decide where to pass the ball are:

- Tactical value of the pass destination;
- Chance of opponent intercepting the pass;
- Confidence on the receiver's position and interception;
- Location and orientation upon ball reception;
- Situations originated if the ball is intercepted;
- Passing travel distance;
- Initial and final player congestion on pass execution;
- Chance of providing a shoot opportunity.

Instead of relying on the previous predefined criteria that embeds the passing strategy, this strategy can be learned using Q-Learning [27].

To balance the implicit risks and gains of the previous criteria with the costs and real-time constraints of adequate decision-making developed techniques apply a weighted sum based on the player's type [42], Fuzzy logic [46] and the Pareto Optimality Principle [22].

To improve the efficiency of the previous position searching methods, a Rational Passing Decision based on Regions [56] classification (e.g. tactical, dominant, passable and falling) was proposed. Each region captures qualitative knowledge of passing in a natural and efficient way. This technique has a low computational complexity, allows the player to decide rationally without precise information and balances success and reward of passing. However, these pros depend highly on the regions characteristics, specifically their dimension.

Voronoi Diagrams [10] were proposed to limit the number of possible meaningful passes, but are unable to find (or learn) the selection of an optimal pass.

VIII. CONCLUSION

Since the start of the RoboCup initiative, several coordination techniques were proposed that tackle core MAS coordination issues in simulated robotic soccer.

The majority of these techniques has dealt with the problem of adequate player positioning, due to its impact on the successful execution of other actions (e.g. passing) during a match. Also many of presented techniques are interdependent (e.g. CG and VE) and rely heavily on coordination technologies. In general, positioning techniques have evolved from reactive to more deliberative approaches, meaning that players now put the team's goals in front of his own because it is the only way for successful coordination to be achieved. Due to its complexity, this problem as been studied in more narrower scopes (e.g. defensive and offensive situations like opponent marking and ball passing respectively) with good results. However, situations where the number of teammates and opponents is uneven still don't seem to be adequately addressed by any of these.

Besides positioning, other techniques were proposed to cope with the remaining player's actions (e.g. marking).

Coordination technologies have evolved a lot since the start of RoboCup mostly due to added functionalities and constraints in the latest simulator releases. Although the use of communication and intelligent perception can assist team coordination through the sharing of pertinent world information and enhance the player's world state accuracy respectively, the simulator constraints discourage relying solely on them.

Team strategies are usually very complex and are typically embedded into players knowledge prior to a game (e.g. using LRA). The strategic approaches have also evolve from fixed policies to more flexible and dynamic policies that are based on real-time match information and previous opponent knowledge. Coaching was used to tweak team strategy mostly by giving advices to players and allow a quicker adaptation to opponent's behavior. Training methods have been used as a foundation to build into team members effective knowledge that can accelerate team coordination during real-time match situations (e.g. learning opponent behavior).

Action selection and behavior acquisition must rely on a good understanding of what can be achieved by intelligent perception and communication techniques.

Machine learning techniques (e.g. Q-Learning) were successfully used for behavior acquisition and adaptive coordination when faced with unpredicted constraints or situations. Due to their high computational cost and thus unfeasibility for real-time decision making, acceleration techniques must be used to increase their efficiency and make them adequate for online usage (e.g. HARL, KBKR). It can be argued that machine learning techniques can be more accurate than hand-coding rule-based (possibly conditional) techniques.

In order to succeed, a good coordination methodology should always consider the following aspects:

- Incorporate past knowledge (e.g. using LRA) to accelerate initial decisions for usual situations, driven from direct human expertise or by offline learned prediction models. This knowledge can be tailored for specific opponents;
- Knowledge should be adaptable according to opponent behavior in real-time;
- Use alternative techniques to complement and replace technologies based on communication and perception.

ACKNOWLEDGMENT

This work was financially supported by Polythecnic Institute of Viseu under a PROFAD scholarship.

REFERENCES

[1] A. Agah and K. Tanie, 'Robots Playing to Win: Evolutionary Soccer Strategies', in *IEEE ICRA*, volume 1, pp. 632–637, Albuquerque, NM, USA, (1997). IEEE.

[2] H. Akiyama and I. Noda, 'Multi-Agent Positioning Mechanism in the Dynamic Environment', in *RoboCup 2007: Robot Soccer World Cup XI*, eds., U. Visser, F. Ribeiro, T. Ohashi, and F. Dellaert, volume 5001 of *LNAI*, 377–384, Springer, Berlin, (2008).

[3] T. Andou, 'Refinement of Soccer Agents' Positions using Reinforcement Learning', in *RoboCup-97: Robot Soccer World Cup I*, ed., H. Kitano, volume 1395 of *LNAI*, 373–388, Springer-Verlag, Berlin, (1998).

[4] R. Bianchi, C. Ribeiro, and A. Costa, 'Heuristic Selection of Actions in Multiagent Reinforcement Learning', in *IJCAI-07*, pp. 690–696, Hyderabad, India, (2007). Morgan Kaufmann Publishers Inc.

[5] R. Bianchi, R. Ros, and R. Mantaras, 'Improving Reinforcement Learning by Using Case Based Heuristics', in *Case-Based Reasoning Research and Development*, eds., L. McGinty and D. Wilson, volume 5650 of *LNAI*, 75–89, Springer, Seattle, WA, (2009).

[6] L. Celiberto and J. Matsuura, *Robotic Soccer: The Gateway for Powerful Robotic Applications*, volume 2 of *Proceedings of ICINCO-2006*, IST, IC&C, Setubal, 2008.

[7] M. Cheny, K. Dorer, E. Foroughi, F. Heintz, Z. Huangy, S. Kapetanakis, K. Kostiadis, J. Kummeneje, J. Murray, I. Noda, O. Obst, P. Riley, T. Stevens, Y. Wangy, and X. Yiny, *RoboCup Soccer Server Users Manual*, For Soccer Server Version 7.07 and later, The RoboCup Federation, 2003.

[8] H. Dashti, N. Aghaeepour, S. Asadi, M. Bastani, Z. Delafkar, F. Disfani, S. Ghaderi, S. Kamali, S. Pashami, and A. Siahpirani. Dynamic Positioning based on Voronoi Cells (DPVC), July 2005 2005.

[9] J. Dawei and W. Shiyuan, 'Using the Simulated Annealing Algorithm for Multiagent Decision Making', in *RoboCup 2006: Robot Soccer World Cup X*, eds., G. Lakemeyer, E. Sklar, D. Sorrenti, and T. Takahashi, volume 4434 of *LNAI*, 110–121, Springer, Berlin, (2007).

[10] H. Endert, T. Karbe, J. Krahmann, and F. Trollmann. Dainamite - Team Description, 2009.

[11] RoboCup Federation. RoboCup: Overview, 01-10-2010 2010.

[12] R. Ferreira, L. Reis, and N. Lau, 'Situation Based Communication for Coordination of Agents', in *Scientific Meeting of the Portuguese Robotics Open*, eds., L. Reis, A. Moreira, E. Costa, P. Silva, and J. Almeida, pp. 39–44, Porto, (2004). FEUP Edições.

[13] T. Gabel and M. Riedmiller. Brainstormers 2D - Team Description, 2009.

[14] T. Gabel, M. Riedmiller, and F. Trost, 'A Case Study on Improving Defense Behavior in Soccer Simulation 2D: The Neurohassle Approach', in *RoboCup 2008: Robot Soccer World Cup XII*, eds., L. Iocchi, H. Matsubara, A. Weitzenfeld, and C. Zhou, volume 5399 of *LNCS*, 61–72, Springer, Berlin, (2009).

[15] H. Igarashi, K. Nakamura, and S. Ishihara, 'Learning of Soccer Player Agents using a Policy Gradient Method: Coordination between Kicker and Receiver during Free Kicks', in *IJCNN*, ed., X. He, H.and Xu, pp. 46–52, Hong Kong, (2008). IEEE.

[16] M. Isik, F. Stulp, G. Mayer, and H. Utz, 'Coordination without Negotiation in Teams of Heterogeneous Robots', in *RoboCup 2006: Robot Soccer World Cup X*, eds., G. Lakemeyer, E. Sklar, D. Sorrenti, and T. Takahashi, volume 4434 of *LNAI*, 355–362, Springer, Berlin, (2007).

[17] W. Jin, W. Tong, W. Xiao, and M. Xiangping, 'Multi-Robot Decision Making based on Coordination Graphs', in *ICMA*, pp. 2393–2398, (2009).

[18] H. Kim, H. Shim, M. Jung, and J. Kim. Action Selection Mechanism for Soccer Robot, 1997.

[19] J. Kok, M. Spaan, and N. Vlassis, 'Multi-Robot Decision Making using Coordination Graphs', in *11th ICAR*, eds., A. Almeida and U. Nunes, pp. 1124–1129, Coimbra, Portugal, (2003).

[20] J. Kok, M. Spaan, and N. Vlassis, 'Non-Communicative Multi-Robot Coordination in Dynamic Environments', *Robotics and Autonomous Systems*, **50**(2-3), 99–114, (2005).

[21] J. Kok and N. Vlassis, 'Using the Max-Plus Algorithm for Multiagent Decision Making in Coordination Graphs', in *RoboCup 2005: Robot Soccer World Cup IX*, eds., A. Bredenfeld, A. Jacoff, I. Noda, and Y. Takahashi, volume 4020 of *LNAI*, 359–360, Springer, Berlin, (2005).

[22] V. Kyrylov, 'Balancing Gains, Risks, Costs, and Real-Time Constraints in the Ball Passing Algorithm for the Robotic Soccer', in *RoboCup 2006: Robot Soccer World Cup X*, eds., G. Lakemeyer, E. Sklar, D. Sorenti, and T. Takahashi, volume 4434 of *LNAI*, 304–313, Springer, Berlin, (2007).

[23] V. Kyrylov and E. Hou, 'While the Ball in the Digital Soccer is Rolling, where the Non-Player Characters should go in a Defensive Situation?', in *Future Play*, eds., B. Kapralos, M. Katchabaw, and J. Rajnovich, pp. 90–96, Toronto, Canada, (2007). ACM.

[24] V. Kyrylov and Eddie Hou, 'Pareto-Optimal Collaborative Defensive Player Positioning in Simulated Soccer', in *RoboCup 2009: Robot Soccer World Cup XIII*, eds., J. Baltes, M. Lagoudakis, T. Naruse, and S. Shiry, volume 5949 of *LNAI*, Springer, Berlin, (2010).

[25] V. Kyrylov and S. Razykov, 'Pareto-Optimal Offensive Player Positioning in Simulated Soccer', in *RoboCup 2007: Robot Soccer World Cup XI*, eds., U. Visser, F. Ribeiro, T. Ohashi, and F. Dellaert, volume 5001 of *LNAI*, 228–237, Springer, Berlin, (2008).

[26] A. Lattner, A. Miene, U. Visser, and O. Herzog, 'Sequential Pattern Mining for Situation and Behavior Prediction in Simulated Robotic Soccer', in *9th RoboCup International Symposium*, eds., A. Lattner, A. Miene, U. Visser, and O. Herzog, Osaka, Japan, (2005).

[27] X. Li, W. Chen, J. Guo, Z. Zhai, and Z. Huang, 'A New Passing Strategy based on Q-Learning Algorithm in RoboCup', in *ICCSSE*, volume 1, pp. 524–527. IEEE, (2008).

[28] R. Maclin, J. Shavlik, L. Torrey, T. Walker, and E. Wild, 'Giving Advice about Preferred Actions to Reinforcement Learners via Knowledge-Based Kernel Regression', in *AAAI-05 and IAAI-05*, eds., M. Veloso and S. Kambhampati, pp. 819–824, Pittsburgh, Pennsylvania, (2005). AAAI Press / The MIT Press.

[29] L. Mota and L. Reis, 'Setplays: Achieving Coordination by the Appropriate use of Arbitrary Pre-Defined Flexible Plans and Inter-Robot Communication', in *ROBOCOMM-2007*, pp. 1–7, Athens, (2007). IEEE Press.

[30] K. Nakamura and H. Igarashi, 'Learning of Decision Making at Free Kicks using Policy Gradient Methods', in *Robotics and Mechatronics*, (2005).

[31] R. Nakanishi, K. Murakami, and T. Naruse, 'Dynamic Positioning Method Based on Dominant Region Diagram to Realize Successful Cooperative Play', in *RoboCup 2007: Robot Soccer World Cup XI*, eds., U. Visser, F. Ribeiro, T. Ohashi, and F. Dellaert, volume 5001 of *LNAI*, 488–495, Springer, Berlin, (2008).

[32] A. Nie, A. Hönemann, A. Pegam, C. Rogowski, L. Hennig, M. Diedrich, P. Hügelmeyer, S. Buttinger, and T. Steffens, 'ORCA - Osnabrueck RoboCup Agents Project', Technical report, Institute of Cognitive Science, (2004).

[33] I. Noda, M. Asada, H. Matsubara, M. Veloso, and H. Kitano, 'RoboCup as a Strategic Initiative to Advance Technologies', in *IEEE ICSMC*, volume 6, pp. 692–697, Tokyo, Japan, (1999). IEEE Press.

[34] I. Noda, H. Matsubara, K. Hiraki, and I. Frank, 'Soccer Server: A Tool for Research on Multi-Agent Systems', *Applied Artificial Intelligence*, **12**(2-3), 233–250, (1998).

[35] I. Noda and P. Stone, 'The RoboCup Soccer Server and CMUnited Clients: Implemented Infrastructure for MAS research', *Autonomous Agents and Multi-Agent Systems*, **7**(1-2), 101–120, (2003).

[36] I. Noda, S. Suzuki, H. Matsubara, M. Asada, and H. Kitano. Overview of RoboCup-97, 1998.

[37] O. Obst and J. Boedecker, 'Flexible Coordination of Multiagent Team Behavior using HTN Planning', in *RoboCup 2005: Robot Soccer World Cup IX*, eds., I. Noda, A. Jacoff, A. Bredenfeld, and Y. Takahashi, 521–528, Springer, Berlin, (2006).

[38] E. Pagello, A. D'Angelo, F. Montesello, F. Garelli, and C. Ferrari, 'Cooperative Behaviors in Multi-Robot Systems through Implicit Communication', *Robotics and Autonomous Systems*, **29**(1), 65–77, (1999).

[39] J. Penders, 'Conflict-based Behaviour Emergence in Robot Teams', in *Conflicting Agents: Conflict Management in Multi-Agent Systems*, Multiagent Systems, Artificial Societies, and Simulated Organizations: International Book Series, 169–202, Kluwer Academic Publishers, Norwell, (2001).

[40] F. Ramos and H. Ayanegui, 'Discovering Tactical Behavior Patterns supported by Topological Structures in Soccer Agent Domains', in *AAMAS-2008*, eds., L. Padgham, D. Parkes, J. Müller, and S. Parsons, volume 3, pp. 1421–1424, Estoril, Portugal, (2008). IFAAMAS.

[41] S. Razykov and V. Kyrylov, 'While the Ball in the Digital Soccer is Rolling, where the Non-Player Characters should go if the Team is Attacking?', in *Future Play*, Ontario, Canada, (2006). ACM.

[42] L. Reis, *Coordination in Multi-Agent Systems: Applications in University Management and Robotic Soccer*, Phd, 2003.

[43] L. Reis and N. Lau, 'Coach UNILANG - A Standard Language for Coaching a (Robo)Soccer Team', in *RoboCup 2001: Robot Soccer World Cup V*, eds., A. Birk, S. Coradeschi, and S. Tadokoro, volume 2377 of *LNAI*, 183–192, Springer, Berlin, (2002).

[44] L Reis, N. Lau, and E. Oliveira. Situation Based Strategic Positioning for Coordinating a Team of Homogeneous Agents, 2001.

[45] R. Ros, J. Arcos, R. de Mantaras, and M. Veloso, 'A Case-based Approach for Coordinated Action Selection in Robot Soccer', *Artificial Intelligence*, **173**(9-10), 1014–1039, (2009).

[46] M. Simões, B. Silva, A. Cerqueira, and L. Silva. Bahia2D - Team Description, 2009.

[47] F. Stolzenburg, J. Murray, and K. Sturm, 'Multiagent Matching Algorithms with and without Coach', *Decision Systems*, **15**(2-3), 215–240, (2006).

[48] P. Stone, *Layered Learning in Multi-Agent Systems*, Phd, 1998.

[49] P. Stone and D. McAllester, 'An Architecture for Action Selection in Robotic Soccer', in *AAMAS-06*, pp. 316–323, Montreal, Quebec, Canada, (2001). ACM.

[50] P. Stone, P. Riley, and M. Veloso, 'Defining and using Ideal Teammate and Opponent Agent Models', in *IAAI-00*, (2000).

[51] P. Stone, R. Sutton, and G. Kuhlmann, 'Reinforcement Learning for RoboCup Soccer Keepaway', *Adaptive Behavior*, **13**(3), 165–188, (2005).

[52] P. Stone and M. Veloso, 'Task Decomposition, Dynamic Role Assignment, and Low-Bandwidth Communication for Real-Time Strategic Teamwork', *Artificial Intelligence*, **110**(2), 241–273, (1999).

[53] P. Stone and M. Veloso. Team-Partitioned, Opaque-Transition Reinforcement Learning, 1999.

[54] F. Stulp, M. Isik, and M. Beetz, 'Implicit Coordination in Robotic Teams using Learned Prediction Models', in *ICRA*, IEEE ICRA, 1330–1335, IEEE, New York, (2006).

[55] U. Visser, C. Drucker, S. Hubner, E. Schmidt, and H. Weland, 'Recognizing Formations in Opponent Teams', in *RoboCup 2000: Robot Soccer World Cup IV*, eds., P. Stone, T. Balch, and G. Kraetzschmar, volume 2019 of *LNAI*, 391–396, Springer-Verlag, Berlin, (2001).

[56] X. Yuan and T. Yingzi, 'Rational Passing Decision Based on Region for the Robotic Soccer', in *RoboCup 2007: Robot Soccer World Cup XI*, eds., U. Visser, F. Ribeiro, T. Ohashi, and F. Dellaert, volume 5001 of *LNAI*, 238–245, Springer, Berlin, (2008).

[57] R. Zafarani and M. Yazdchi, 'A Novel Action Selection Architecture in Soccer Simulation Environment using Neuro-Fuzzy and Bidirectional Neural Networks', *International Journal of Advanced Robotic Systems*, **4**(1), 93–101, (2007).