# Heuristiscs-based High-level Strategy in Multi-Agent Environment ⋆


**Péter Gasztonyi** * **István Harmati** **

* *Department of Control Engineering and Information Technology*
*Budapest University of Technology and Economics,*
*Magyar Tudósok krt. 2/B422, Budapest, Hungary H-2117*
*(e-mail: peat@ peat.hu).*
** *Department of Control Engineering and Information Technology*
*Budapest University of Technology and Economics,*
*Magyar Tudósok krt. 2/B422, Budapest, Hungary H-2117*
*(e-mail: harmati@ iit.bme.hu).*



**Abstract:** In this paper, a high-level strategy concept is presented for robot soccer, based on low level heuristic inference methods, rather than explicit rule-based strategy. During tactical positioning, no strict role set is assigned for the agents, instead a fitting point of the role-space is selected dynamically. The algorithm for this approach applies fuzzy logic. We compute fields-of-quality, regarding some relevant aspects of the scenario, and integrate them into one field for each player, according to given strategic parameters (as weights). These fields will be the base of the players' decision of positioning. A significance order is also set up for the players, and their relevant location is derived from the decision-field, through subtractive clustering, in order of their significance. If an agent is in a position to manipulate the ball, an appropriate action is being selected for it. The simulation and experiments prove that the proposed approach can be efficient in dynamically changing environment or against opponents of different strategies.



Keywords: robotic soccer; fuzzy logic; subtractive clustering; multiagent system.


## 1. INTRODUCTION

The control and motion-planning of multiagent systems typically call for modular managing, thus the robot soccer strategies follow the same approach. Most solutions define about 3 levels for the modular control system, and these levels are responsible for different scope of phisical movement or strategical positioning. A typical realization is described in [4] and [7]. These systems operate with deterministic rule based strategies, while others try to optimize the role assignment dinamically [2]. However, these algorithms are outworn by newer conceptions, involving holistic approach [5] or interpreting the roles less strictly [3]. These approaches let the agents to take their tasks according to global states, or to define their positions in a more transitional way ('area of responsibility').

We intend to eliminate explicite role assignment, but a hierarchical control system is needed to control the agents' movements. Similarly to the generic concepts, we follow the undermentioned partition of the control logics to layers as seen in Figure 1.

The specific objective of the layers are the following:

- *High Level Strategy [HLS]*: coordinating the strategy; setting up actual strategic goals; distributing area and responsibility to players

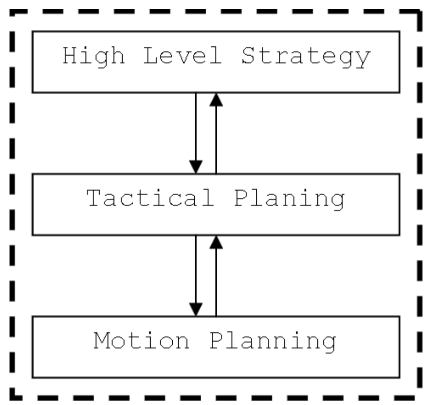


Fig. 1. The modular control system for robot soccer

- *Tactical Planning [TP]*: dynamic path planning for agents (handling joined movements); carrying out basic skills and special actions
- *Motion Planning [MP]*: low level path planning for the given robot type, with no obstacles supposed in space; generating control signs.

The objective of high level strategies is to provide optimal dislocation for the team in the dynamic environment, beside the obstructive behaviour of other team's agents on the same field. Optimal dislocation of the agents has a contribution to the dominance of the team, and assists to achieve own objectives - and to have the opponent fail theirs. However, since the optimal solution of the problem


⋆ The research was partly supported by the Hungarian Science Research Fund under grant OTKA K 71762 and by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

is not available, this paper aims to integrate some known partial heuristic solutions related to the subject. Taking simplifications, e.g. using fuzzy logic, which fits well to the problem, can be effective in the solution.

In Section 2, the problem is shown in details, and the used configuration is specified along with the outline of the strategy logic system. After this overview, the main concept of the solution is presented in Section 3, including the action-selection function of the agents. Finally, based on the above introduced knowledge, some aspect of the team coordinaton is discussed in Section 4, and finally, we draw the conclusions.

## 2. PRELIMINARIES

In our framework, the mobile agents are realized by differential driven robots. These robots are described with their transition equations:

$$\begin{aligned} \dot{x} &= \frac{r}{2}(u_r + u_l)cos\phi \\ \dot{y} &= \frac{r}{2}(u_r + u_l)sin\phi \\ \dot{\phi} &= \frac{r}{L}(u_r - u_l) \end{aligned} \tag{1}$$

where $x$ and $y$ are the coordinates, $\phi$ is the orientation of the agent, while $u$ is the angular speed of a wheel, $r$ is the wheel diameter and $L$ is the distance between the wheels.

As we do not have extra ball manipulator devices onboard, the simplified exterior of the robots is a symmetric cylindered case. However, this type of cover is easy to apply and also to simulate; it affects the logic rather badly, due to expecting a quite precise maneuvering from the agents - as we will see this problem a bit more detailed in the next section. This realization results in a soccer team whose members try to bump into the ball such a way that the ball is preferred to be driven into the opposite team's cage - which goal cage is a hole in the boarding of the squared field. As of the similarities with the physics of the billiards game, this realization is referred to as 'billiard ball model'. According to the simulation, the 3D cylindrical robots are projected into a 2D circle, and the collisions are simulated in the 2D plane, too.

One team can consist arbitrary number of robots, but the algorithm is implemented for a typical of 3..5 agents. These robots have a maximum angular velocity for their wheels, thus limiting both the linear and turning speeds. They also have a weight parameter, which is counted during collision calculations. As part of the billiard ball model, these calculations are detailed in [1]. According to the current model, the ball friction is constant: the ball loses 0.8% of its current speed in every simulation cycle.

The modular arhitecture of the team-coordination logic is based on the following considerations:

- The drafting of the main goal includes the following: a) score goals, b) more than the opponent. This suggest a parallel realization of two independent algorithms for the different aims (scoring and defending), which are fundamentally conquering, but can also have a validity attribute according to the current game situation, so in the given scenario, one can decide whether to attack or defend. Furthermore, these layers (and so the calculations) might be connected on the recognition that the one team's offence chances are the other's threats.
- The "scoring" also requires a special action, namely, manipulating the ball in a desired direction (let it be called a "kick"). A kick is a detached action from the class of different types of positional-purpose moving, because its execution strongly affected not only by the intention, but also the current scenario. Thus, it has to be calculated separately and at a higher priority level (before other calculations).
- When not actually trying to score ("kicking to the goal"), but simply positioning, an appropriate logic is needed for setting up an order among the agents of the team and decide the targets themselves.
- An interchangeable interface is recommended for the agents' control and communication, as to support not fully homogeneous teams, or two homogeneous teams of different type of robots, too. In these cases, the motion control of certain robot types will be different, of course, but remains compatible to be built into the team.
- Besides these internal physical constraints, one might face other limitations, e.g. the presence of other objects - which might also be a subject to handle.

## 3. FUNDAMENTALS OF THE HLS LAYER

The basic concept of the HLS in subject is to avoid explicit fix rule-based strategy and force improvisative play in order to be able to adapt to different opponents or to those who are dinamically changing their strategies. The main idea is to take the ball trajectory design as the primary object (instead of positioning the players directly, according to a certain formation or any heuristics). This can be done by positioning the players, of course, as to form the vertices of the ball trajectory in question - but as from our aspect, the available reasonable ball trajectories, in theory, will not depend on any predefined formation.

One could say that the ball trajectory can not be planned due to the chaotic nature of the game, which comes from the indeterministic nature of the opponent's moves. In our case, however, the ball's dynamics are much faster than players' moves, thus these trajectories can be planned for a short while, indeed.
Since the players have limited speeds and they have to bump against the ball to kick it, the required (potential) vertices are spread across the field in limited ranges, but this aspect also makes possible to combine some different quality measurement into one decision-field. Different possible trajectories come from different reasons, and the integration to a common field is performed through the following *strategic parameters*, which weight or affect the individual property-fields:

- **Shooting Rate**: the preference of shooting on goal.
- **Pressure**: the preference for agents to act on the opponent's side.
- **Possession**: the preference of safe kick sequences (where the possibility of the opponent will touch the ball is minimal).

- **RiskTake**: a real value between -1..+1, that affects the default general parameters of actions, e.g. the range of safe reach for an agent.
- **ActiveSubtractRadii**: the radius parameter of subtractive clustering, when deriving *positioning* target points from a decision field. Thus, it affects the admissible proximity for agents.
- **PassingSubtractRadii**: the similar radius parameter when deriving a target for a *kick*.
- **ActiveSubtractPoints**: the number of target points to be returned when running the subtractive function. It is desired to calculate more points than the number of team members, as advanced distribution functions can benefit from this.

The 'good places' (of where to kick) can be determined through the obtained quality field, and a secondary forthcoming target can also be chosen to refine the actual action and prepare an 'indirect pass', that is, to aim the kick not in the *actual* direction of a teammate, but to a point where it shall be at the time when the ball reaches it. This procedure can apply only to the players who have a reasonable chance to perform the kick in a limited time, considering both the technical part of the kick and the opponents' dislocation. Once the action is fixed, the other agents can move according to this decision. In case of strictly distributed logics (i.e. when no communication is present between the agents), the action is treated as fixed by the agents, when the kick is performed by the actual kicker. Other main fields can be constructed of 'where to move' for the players, which is derived from the 'where would be the ball in best place' field, through a simple time-cost function of players to move there.

In Figure 2, one can notice the small spot as the ball, and the two teams, consisting all filled black and coloured members, respectively. The upper-right coloured agent is allocated to manipulate the ball, thus its target sign is shown accordingly. Then the coloured left player is best to support the kicker, while the lower agent is about to move back closer to its own goal. These allocations are inferenced from a calculated field for the expected moment of the kick, which tells the value of passing to the given points of the field, from the expected kicking point in that time. The extracting algorithm discretizes the field to some points, representing good target places. This can involve more possible good places, and more distribution algorithms can be used to allocate targets to players. A sample quality field for passing, and a subtractive allocation procedure is shown through Figure 3-9.

The formula of one subtraction step for one discretized grid point *P=(x,y)* of the field is as follows:

$$Q(P) = \max(0, Q(P) - Q_{max} \cdot exp(-0.5 \cdot \quad (\frac{(x_{max} - x, y_{max} - y)}{\sigma})^2)) \quad (2)$$

where $(x_{max} - x, y_{max} - y)$ is the distance of $P$ from the location of the maximum value of $Q_{max}$ in the field, while sigma is a constant parameter.

Although it has been mentioned in Section 2, that the one team's offence chances are the other's threats, thus they could be treated as the dual pairs of each other,

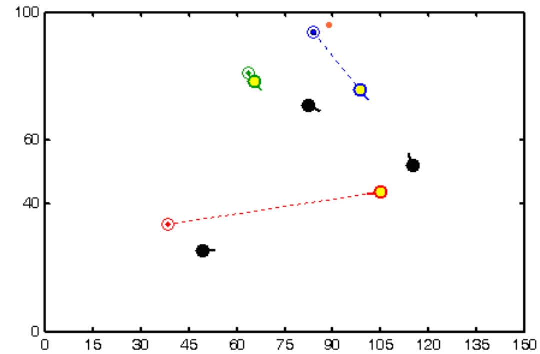


Fig. 2. Screenshot of a scenario in simulation

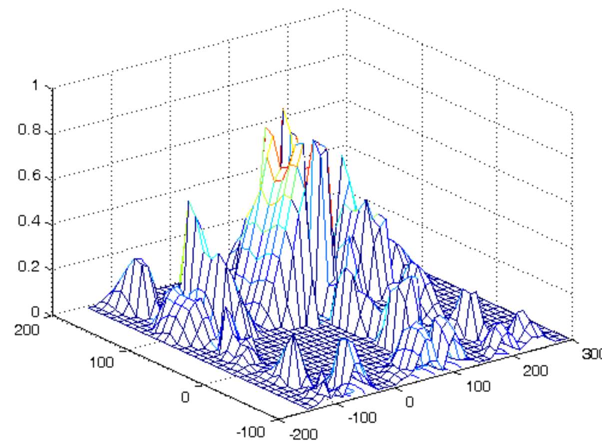


Fig. 3. The initial decision map for the situation

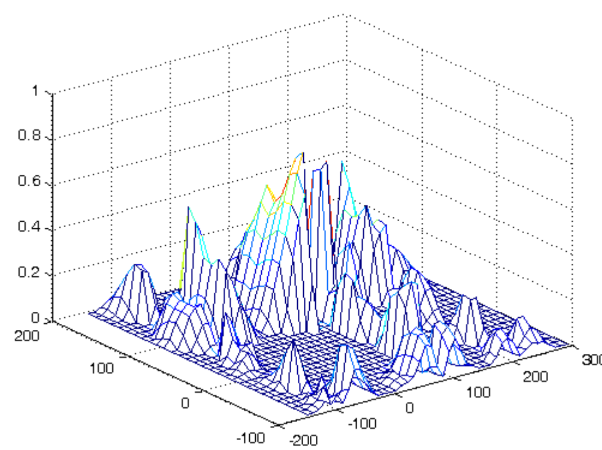


Fig. 4. Decision map after the first subtraction step

despite in the realization, this method was shown to be wasting too much of resources, therefore simplifications had to be applied in the defence logic. The connection between the offense logic and the simplified defense logic remains the recognition that the ball, if untouched, moves following a straight line. Thus, it can only change its direction if a player kicks it or it is bumped into the wall. The effect of collisions with the wall is calculated through

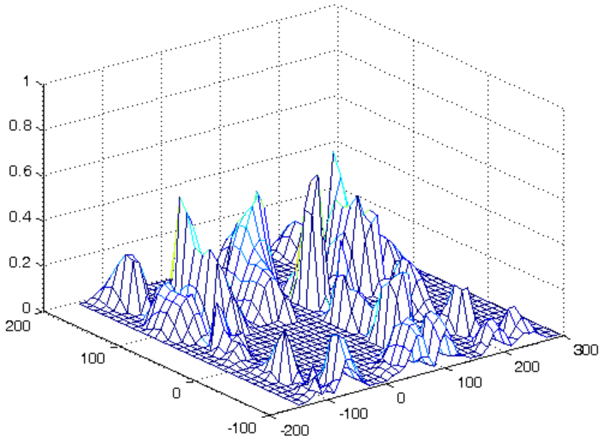


Fig. 5. Decision map after the second subtraction step

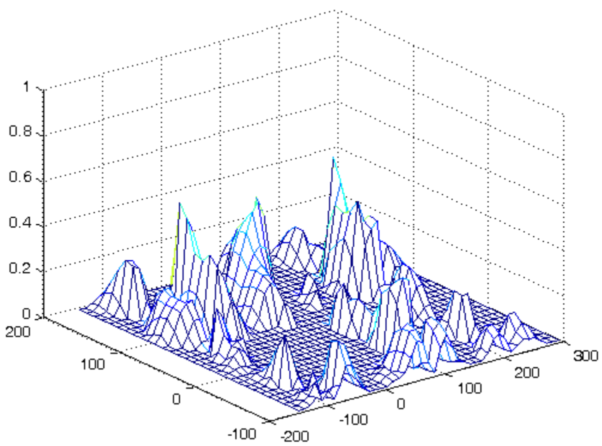


Fig. 6. Decision map after the third subtraction step

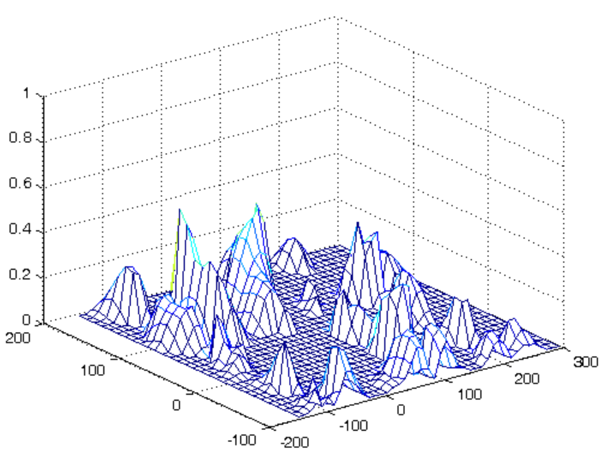


Fig. 7. Decision map after the fourth subtraction step

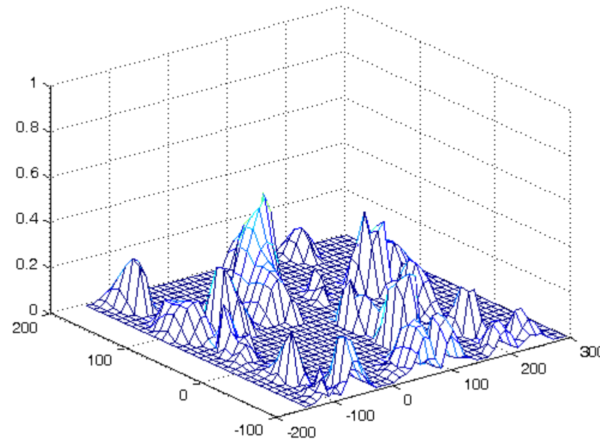


Fig. 8. Decision map after the fifth subtraction step

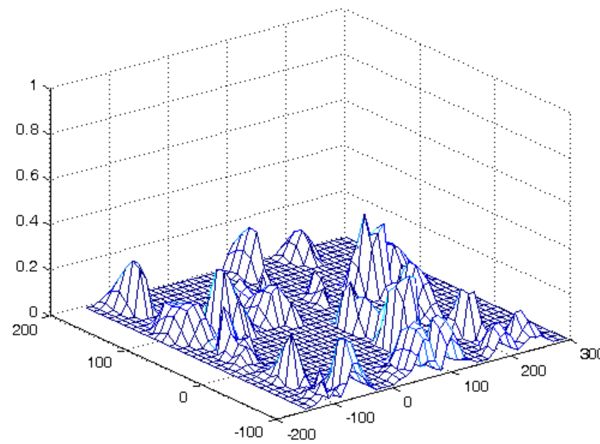


Fig. 9. Decision map after the sixth subtraction step

an appropriate reflection transformation (depending on physics modeling). Altogether, if a self player is clearly able to kick the ball before the opponent (or a "mirror instance") reaches it, then that action will be chosen by the algorithm. The only real threat to be prepared to is if an opponent agent is likely to manipulate ball before a self teammate could do the same. Assuming that it is going to kick the ball as soon as possible (which is likely the most efficient action), the defence can be prepared to close the possible ways from the kicking point to the goal cage or to other potential kickers of the opponent. This likelihood (the possibility that the opponent reaches the ball first) is linked to a parameter which partly describes the seriousness of the current threat - the other part is that *where* in the field is that situation. Figure 10 shows an example of localizing these threats of possible opponent kicks.

## 4. TEAM COORDINATION

The purpose of positioning is to create a condition in which the likelihood of accomplishing the currently ordered task is the possible maximum. Positioning in itself is not

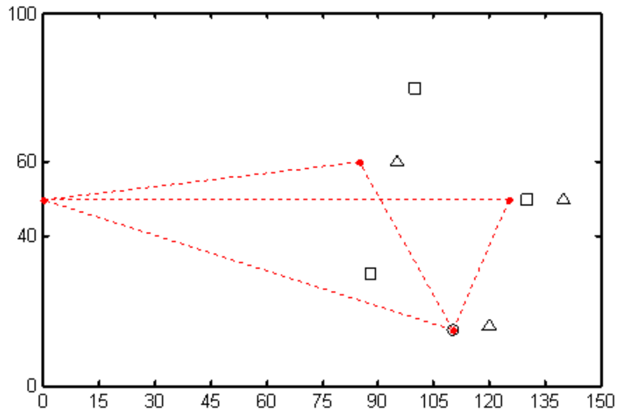


Fig. 10. Localizing threats of opponent kicks. Squares are teammates, triangles are opponents, and the ball is currently at (110,15). Dashed lines are possible paths of ball to own goal.

necessary to be treated as a skill - indeed, we position to be in position to perform a real skill in the future.

It is easier, however, to assign certain roles to players, which means that we discretize the responsibility and many attributes for the players according to multiple components. By tying up a player to a single role, we simply lose the overlapped information between the real situation's demands and the possible actions, determined by that certain role. This is because the actual situation can transit to another where completely different role assignment may occur. In this case, sharp changes in roles would have to be carried out to some players, which is clearly not the most economic way of coordinating the players and the team. Instead, we want to take every opportunitiy and the relevant threats into consideration by their appropriate importance in every moment. This effort can be fulfilled by the earlier described use of properly weighted multi-purpose positioning, which results in smoother transitions between the defined roles. Moreover, if the defined selectable purposes of positioning do not correspond directly to a role set, rather are components of conventional 'roles', then the agents are available to pick the most fitting continuous point of the purpose-space for the actual situation.
This aspect implies even the omission of a fixed goalie role, e.g. as if the conditions allow to involve all players in an offence, the most efficient dislocation will be chosen, considering the appointed strategic and tactical parameters.

Once the position is solved this way, if one agent has the chance to manipulate the ball, a direct skill action (kick) is called. If more agents has the opportunity to kick, more of them may be investigated for a target and for a secondary purpose to refine the first aim. The kicker selection method uses fuzzy algorithm, which fits well to the problem and can be easily expanded. Fuzzy rules for the kicker selection may look like:

- if *distanceToBall* is **near** and *relativeSpeed* is **approaching** and *potentialShootAngle* is **front** then *playerIsShooter* is **high**,
- if *distanceToBall* is **medium** or *relativeSpeed* is **invariant** then *playerIsShooter* is **medium**,

where

$$\mu_{distanceToBall} = \frac{distanceToBall_{min}}{distanceToBall} \quad (3)$$

and

$$\mu_{potentialShootAngle} = cos(\frac{potentialShootAngle}{2}) \quad (4)$$

and *relativeSpeed* is the ball velocity vector projected to the direction towards the player, while *potentialShootAngle* is the estimated angle between the ball's moving direction and the direction from the expected kick-point to the related target.

When the potential kickers are selected with their available targets along with the secondary options, these choices will logically form a set of precedence graph of actions. If these nodes are ranked, the best action-chain can be found, through an appropriate weight-function, which is affected by the actual values of the strategical parameters - like risk-taking, shooting-rate, etc.

The selected action-chain determines the kicker and its chosen target, along with the secondary intention concerning a possible next kick, which would refine the execution of the forthcoming one. As a consequence of the billiard ball model, proper calculations are needed for the optimal motion planning for a correct kick: a) the main option is whether to kick the ball in one touch or try to stop it first (which would result in 2 kicks indeed) b) the second task is to plan the vertices for a safe bump, i.e. to be in the right position in the right time, avoiding any collisions with the ball before the planned one. Detailed considerations and calculations on this problem can be found in [1].
As the kick action is a well defined maneuver, it is executed at the TP level. Additionally, for the time-cost estimations on HLS level, the kick action function should have an alternative interface for these quick-response queries.

In case of no kicker can be selected clearly, e.g. due to the opponent's proximity to the ball, 'some agent of the self team should take more part in the defensive tasks', that is, to gain more affinity to parry the detected threats. This fuzzy phrase can be easily implemented with the utilization of some properties describing the situation in the field. The affinities in question are described with two parameters: *maxAff*, which determines the maximum affinity in the team, and *maxExp*, an exponential which affects the distribution ratio of the individual affinity values among the agents. With these parameters, the individual defensive affinities are calculated by:

$$defAff(1) = maxAff \quad (5)$$

for the agent that can be allocated to the most serious threat by the lowest cost, in the given situation, and

$$defAff(j) = maxAff \cdot exp(maxExp \cdot \frac{j-1}{M-1}) \quad (6)$$

for the other agents, where M is the number of members in a team.
The properties, as the inputs of the fuzzy inference system, are:

- *pressure*: the current value of the formerly mentioned strategic parameter,

- *kicklead*: an estimated lead in simulation time cycles for the self players to reach the ball before the opponent,
- *sumcritic*: a 'danger-meter' value, calculated from the properties of the detected threats,
- *average_q_active*: a 'chance-meter' value, calculated from the quality of possible targets, those are derived from the decision maps.

Finally, the fuzzy rules to determine the *maxAff* and *maxExp* parameters:

- if *pressure* is **high** and *kicklead* is **high** then *most_defensive* is **verylow**
- if *pressure* is **low** then *most_defensive* is **mid**
- if *kicklead* is **none** then *most_defensive* is **high**
- if *kicklead* is **high** then *most_defensive* is **none**
- if *sumcritic* is **high** then *most_defensive* is **high**
- if *pressure* is **low** then *affinity_distribution* is **split**
- if *kicklead* is **none** then *affinity_distribution* is **equalized**
- if *kicklead* is **high** then *affinity_distribution* is **split**
- if *sumcritic* is **high** then *affinity_distribution* is **balanced**
- if *sumcritic* is **low** then *affinity_distribution* is **split**
- if *average_q_active* is **poor** then *most_defensive* is **high**
- if *average_q_active* is **poor** then *affinity_distribution* is **balanced**
- if *average_q_active* is **promising** and *sumcritic* is **high** then *affinity_distribution* is **split**.

Then the output *most_defensive* is binded to *maxAff*, and *affinity_distribution* to *maxExp*.

## 5. CONCLUSION

Heuristic-based robot soccer strategy have been proposed in the paper. The heuristics rely on distinguished potential fields that provides features for the inference system realized by fuzzy expert. The concept is very flexible and results smooth positioning. In contrast to role based approaches, there is no assigned role to the agents, they are able to accomplish different task in different extent, although the compatibility with role based approach can be easily guaranteed. In this way, both interleaving and fully discretized roles cab be achieved. One real challenge is to find an optimal tuning algorithm for the strategic parameters, which would assure the proper tracking speed of a varying tactic used by an advanced opponent. A possible improvements can be achieved by reinforcment learning technique.

## REFERENCES


[1] GASZTONYI, P.: *Magasszintü csapatstratégiák fejlesztése robotfocihoz*, M.Sc. Thesis, 2006, Budapest University of Technology and Economics, Hungary

[2] SNG, H.L., SEN GUPTA, G., MESSOM, C.H.: *Strategy for Collaboration in Robot Soccer*, IEEE DELTA 2002, New Zealand, 2002, pp 347-351, ISBN: 0-7695-1453-7

[3] TEWS, A., WYETH,G.: *Multi-Robot Coordination in the Robot Soccer Environment*, Uneversity of Queensland, Australia

[4] BOWLING, M.: *Multiagent Learning in the Presence of Agents with Limitations*, Ph.D. Thesis, Carnegie Mellon University, Pittsburgh, PA, USA, CMU-CS-03-118

[5] JENNINGS, A., SHAH, CH., JANI, H, UPADHYA, M., KAPADIA, M.: *Team Yarra: holistic soccer*, RMIT University, Melbourne, Australia

[6] PETER STONE, MANUELA VELOSO: *Task Decomposition and Dynamic Role Assignment for RealTime Strategic Teamwork*, Computer Science Department, Carnegie Mellon University

[7] APARÍCIO, M.: *Design and Implementation of a Population of Cooperative Autonomous Robots*, M.Sc. Thesis, 2000, Universidade Técnica de Lisboa, Portugal

[8] LANTOS B.: *Fuzzy Systems and Genetic Algorithms*, Müegyetemi Kiadó, 2002, ISBN: 963 420 706 5

[9] WANG, T., LIU, J., JIN, X.: *Minority Game Strategies for Dynamic Multi-Agent Role Assignment*, Hong Kong Baptist University, Kowloon Tong, Hong Kong

[10] RUSSEL, S. J., NORVIG, P.: *Artificial Intelligence: A Modern Approach*, Prentice Hall, 2002

[11] TATLIDEDE, U., KAPLAN, K., KÖSE, H., AKIN, H. L.: *Reinforcement Learning for Multi-Agent Coordination in Robot Soccer Domain*, Bogazici University, Istanbul, Turkey, 2005