STP: skills, tactics, and plays for multi-robot control in adversarial environments

B Browning*, **J Bruce**, **M Bowling**, and **M Veloso**
Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

The manuscript was received on 15 September 2003 and was accepted after revision for publication on 18 November 2004.

DOI: 10.1243/095965105X9470

Abstract: In an adversarial multi-robot task, such as playing robot soccer, decisions for team and single-robot behaviour must be made quickly to take advantage of short-term fortuitous events. When no such opportunities exist, the team must execute sequences of coordinated team action that increases the likelihood of future opportunities. A hierarchical architecture, called STP, has been developed to control an autonomous team of robots operating in an adversarial environment. STP consists of *skills* for executing the low-level actions that make up robot behaviour, *tactics* for determining what skills to execute, and *plays* for coordinating synchronized activity among team members. The STP architecture combines each of these components to achieve autonomous team control. Moreover, the STP hierarchy allows for fast team response in adversarial environments while carrying out actions with longer goals. This article presents the STP architecture for controlling an autonomous robot team in a dynamic adversarial task that allows for coordinated team activity towards long-term goals, with the ability to respond rapidly to dynamic events. Secondly, the subcomponent of skills and tactics is presented as a generalized single-robot control hierarchy for hierarchical problem decomposition with flexible control policy implementation and reuse. Thirdly, the play techniques contribute as a generalized method for encoding and synchronizing team behaviour, providing multiple competing team responses, and for supporting effective strategy adaptation against opponent teams. STP has been fully implemented on a robot platform and thoroughly tested against a variety of unknown opponent teams in a number of RoboCup robot soccer competitions. These competition results are presented as a mechanism to analyse the performance of STP in a real setting.

Keywords: multi-robot coordination, autonomous robots, adaptive coordination, adversarial task

1 INTRODUCTION

To achieve a high performance, autonomous multi-robot teams operating in dynamic adversarial environments must address a number of key challenges. The team must be able to coordinate the activities of each team member towards long-term goals but also be able to respond in real time to unexpected situations. Here, real time means responding at least as fast as the opponent. Moreover, the team needs to be able to adapt its response to the actions of the opponent. At an individual level, the robots must be able to execute sequences of complex actions leading towards long-term goals but also respond in real time to unexpected situations. Secondly, each robot must have a sufficiently diverse behaviour repertoire and be able to execute these behaviours robustly even in the presence of adversaries so as to make a good team strategy viable. Although these contrasting demands are present in multi-robot problems [**1**, **2**] and single-robot problems [**3–5**], the presence of adversaries compounds the problem significantly. If these challenges are not addressed for a robot team operating in a dynamic environment, the team performance will be degraded. For adversarial environments, where a team's weaknesses are actively exploited by good opponents, the team performance will degrade significantly.

The sheer complexity of multi-robot teams in adversarial tasks, where the complexity is essentially exponential in the number of robots, creates another

* *Corresponding author: Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA15213, USA. email: brettb@cs.cmu.edu*

significant challenge to the developer. Thus, control policy reuse across similar subproblems, as well as hierarchical problem decomposition, are necessary to make efficient use of developer time and resources.

Addressing all these challenges in a coherent seamless control architecture is an unsolved problem, to date. In this paper, a novel architecture, called STP, is presented for controlling a team of autonomous robots operating in a task-driven adversarial environment. STP consists of three main components, namely *skills*, *tactics*, and *plays*, built within a larger framework providing real-time perception and action generation mechanisms. Skills encode low-level single-robot control algorithms for executing a complex behaviour to achieve a short-term focused objective. Tactics encapsulate what the robot should do, in terms of executing skills, to achieve a specific long-term goal. Plays encode how the team of robots should coordinate their execution of tactics in order to achieve the team's overall goals.

The authors believe that STP addresses many of the challenges to multi-robot control in adversarial environments. Concretely, STP provides three key contributions. Firstly, it is a flexible architecture for controlling a team of robots in a dynamic adversarial task that allows for both coordinated actions towards long-term goals, and fast response to unexpected events. Secondly, the skills and tactics component can be decoupled from plays and supports hierarchical control for individual robots operating within a dynamic team task, potentially with adversaries. Finally, the play-based team strategy provides a generalized mechanism for synchronizing team actions and providing for a diversity of team behaviours. Additionally, plays can be effectively used to allow for strategy adaptation against opponent teams. STP has been fully implemented and extensively validated within the domain of RoboCup robot soccer [**6**]. In this paper, the development of STP within the domain of RoboCup robot soccer is detailed, evidence of its performance in real competitions with other teams is provided, and how the techniques apply to more general adversarial multi-robot problems is discussed.

This article is structured as follows. In the following section, the problem domain of RoboCup robot soccer within which STP has been developed is described. Section 3 presents an overview of the STP architecture and its key modules, leading to a detailed description of the single-robot components of skills and tactics in section 4 and team components of plays in section 5. Section 6 describes the performance of STP in RoboCup competitions against a variety of unknown opponent teams and discusses how STP can be improved and applied to other adversarial problem domains. Finally, section 7 presents related approaches to STP, and section 8 concludes the paper.

2 THE ROBOT SOCCER PROBLEM

The STP architecture is applicable to an autonomous robot team performing a task in an adversarial dynamic domain. To explore this problem concretely, RoboCup robot soccer is selected as the test-bed domain. More specifically, the Small-Size League (SSL), a division within the RoboCup initiative, is chosen. In this section, the SSL robot soccer problem is concretely defined together with the challenges that it poses. This section also details the specific test-bed, the CMDragons system, used to validate the STP architecture to provide a backdrop for the ensuing sections.

2.1 RoboCup robot soccer Small-Size League

RoboCup robot soccer is a world-wide initiative designed to advance the state of the art in robot intelligence through friendly competition, with the eventual goal of achieving human-level playing performance by 2050 [**6**]. RoboCup consists primarily of teams of autonomous robots competing against one another in games of soccer, together with an associated symposium for research discussion. There are a number of different leagues within RoboCup, which are designed to focus on different parts of the overall problem: developing intelligent robot teams. This article is primarily focused on the SSL.

An SSL game consists of two teams of five robots playing soccer on a $2.8\,\text{m} \times 2.3\,\text{m}$ field with an orange golf ball [**7**]. Each team must be *completely* autonomous for the duration of the game, which typically lasts for two 10 min halves. Here, autonomy means that there are no humans involved in the decision-making cycle while the game is in progress. The teams must obey rules that are like those of the Fédération Internationale de Football Association as dictated by a human referee. An assistant referee translates referee commands into a computer-usable format, which is transmitted to each team via RS-232 using a standardized protocol, via a computer running the *RefBox* program [**7**]. Figure 1 shows the general set-up as used by many teams in the SSL. The SSL is designed to focus on team autonomy. Therefore, global vision via overhead cameras and off-field computers, which can communicate with the robots via wireless radio, are allowed to be used.

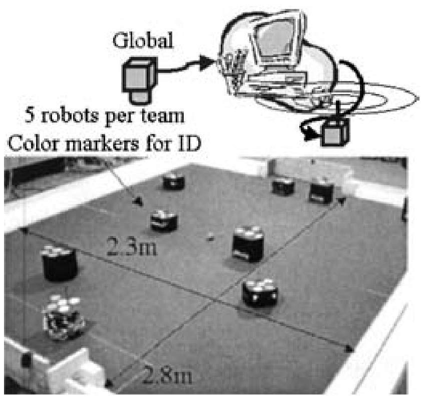

Fig. 1 An overview of the CMDragons small-size robot soccer team

SSL robot soccer involves many research issues. Examples of some of the research challenges include the following:

(a) building *complete* autonomous control systems for a dynamic task with a high performance;
(b) team control in a dynamic environment, and response to an unknown opponent team;
(c) behaviour generation given real sensor limitations of occlusion, uncertainty, and latency;
(d) fast navigation and ball manipulation in a dynamic environment, which are real-world sensors;
(e) fast robust low-latency vision, with easy-to-use calibration routines;
(f) robust high-performance robots with specialized mechanisms for ball manipulation.

A typical SSL game is highly dynamic, where ball speeds of 3 to 4 m/s and robots speeds of 1–2 m/s are common. With such speeds in a small environment, it becomes critical for information to be translated into action quickly in order for the team to be responsive to sudden events in the world. For example, if a robot kicks a ball at 3.5 m/s, a latency of 100 ms means that the ball will have moved over 35 cm before the robots could possibly respond to the observation that the ball had been kicked. High speed of motion and latency impact on control in the following ways.

1. Vision, tracking, and modelling algorithms must compromise between the need to filter noise and to detect unexpected events in minimum time.
2. Prediction mechanisms are required to compensate for latency for effective control.
3. Team and single robot control must adapt quickly to dynamic changes.

The last point means that all control decisions need to be recalculated as often as possible to allow the system to react quickly to unexpected events. As a rough guide, the CMDragons system [**8**] recalculates everything for each frame at a rate of 30 Hz. Typically, high-level decisions change at a slower rate than low-level decisions. For an approximate guide, a play typically lasts 5–30 s, while a tactic may operate over a time frame of 1–30 s, and a skill may operate over a 300 ms–5 s time frame. However, any decision at any level can be switched in the minimum time of one frame period (33 ms) to respond to any large-scale dynamic change.

2.2 The CMDragons

Figure 2 shows the major components of the control system developed for our CMDragons SSL team. This architecture is the result of a long series of developments since RoboCup 1997 [**8–12**]. Figure 3 shows the robot team members. As shown, the architecture consists of a number of modules beginning with vision and tracking, the STP architecture, navigation and motion control, and finally the robot control software and hardware. Each of the non-STP components is briefly described in the following paragraphs to provide the context for later discussions.

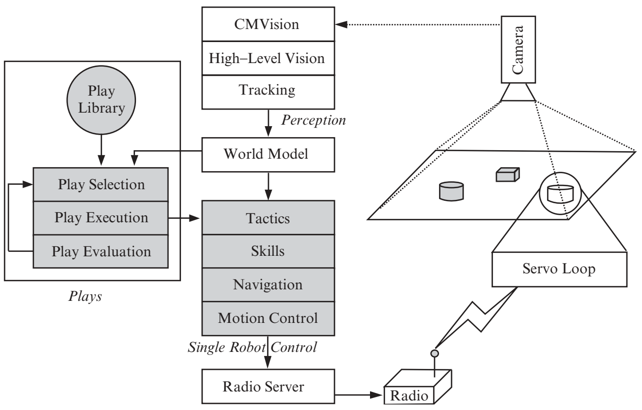

Fig. 2 Overview of the CMDragons team architecture

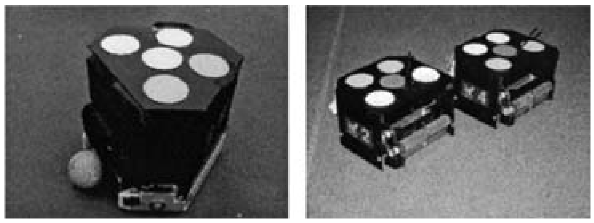

Fig. 3 The CMDragons robots. The robot on the left is an OmniBot, while the robots on the right are DiffBots. Each robot fits within a cylinder 18 cm in diameter and 15 cm tall

Information passes through the entire system synchronized with incoming camera frames at 30 Hz. Thus a new frame arrives, vision and tracking are run on the processed frame, and the resulting information is fed into the world model. The STP architecture is executed, followed by navigation and motion control. The resulting motion command is sent to the robot and the robot executes the command with local control routines.

2.2.1 Perception

Vision is the primary means of perception for the CMDragons team. Everything in the SSL is colour coded (see Fig. 3), making colour vision processing algorithms a natural choice. The ball is orange and the field is green carpet with white lines and white angled walls. Each robot is predominantly black with a yellow or blue circular marker in its centre. Depending upon who wins the toss of the coin before the game, one team uses yellow markers while the other uses blue. Each robot typically has another set of markers arranged in some geometric pattern that uniquely identifies the robot and its orientation. Knowledge of an opponent's additional markers is usually not available before a game.

In the CMDragons team, images from the camera arrive at a frame rate of 30 Hz into an off-field computer. For reference purposes, most of the system described here runs on a 2.1 GHz AMD Athlon XP 2700+ system, although a 1.3 GHz processor was used previously without any difficulties. Using our fast colour vision library, CMVision [**13**], coloured blobs are extracted from each image. The colours are identified on the basis of prior calibration to produce a threshold mapping from pixel values to symbolic colour. With knowledge of each robot's unique marker layout, high-level vision finds each robot in the image and determines its position and orientation. The position of the ball and each opponent robot is also found. Orientation for opponents cannot be found owing to the lack of advance knowledge on their marker layout. The world position of each object is then determined via a parametric camera model learned during game set-up. Full details of the vision algorithms can be found in reference [**14**].

Filtered position and velocity information is derived using a set of independent extended Kalman–Bucy filters (EKBFs) for each object [**8**]. As velocity information cannot be derived from each camera image alone, and there is too much noise for frame differentials to be effective, the EKBFs are used to nullify the effects of both noise and intermittency (missing data). Additionally, the EKBFs provide a prediction mechanism through forward modelling, which is useful for overcoming latency. In summary, the full vision and tracking module provides estimates of each robot location and orientation, each opponent location, and ball location, with velocities for all 11 objects. Taken together, these estimates provide the robot's belief state about the state of the world.

2.2.2 World model belief state

All beliefs about the state of the world, where the robots are, etc., are encapsulated in a world belief model. In short, the world model acts as a centralized storage mechanism for beliefs for all layers of the control architecture to use. The belief model contains the following:

(a) all perceptual information obtained from the tracker (e.g. robot positions and velocities);
(b) game state information derived from the received referee commands;
(c) opponent modelling information derived from statistical models of observed opponent behaviour;
(d) high-level predicates derived from the perceived state, such as which team has possession of the ball, is in attack, and is in a particular role.

Each high-level predicate is a Boolean function of the tracker and/or game state belief. To account for noise, each Boolean function incorporates empirically determined hysteresis to prevent undue oscillation at the decision boundary. These predicates, due to their Boolean nature, provide a symbolic representation that is often more useful for making decisions than the raw belief models, e.g. deciding whether to run an attacking play or a defensive play.

2.2.3 Navigation and motion control action interface

The STP architecture consists of team control and individual robot control. Following the common technique of hybrid hierarchical control [**15**, **16**], lower modules were developed for obstacle-free navigation and motion control. Essentially, these modules provide resources to the robot for generating actions in the world. The resources provided are obstacle-free navigation, motion control, and direct robot commands. Figure 4 shows the control hierarchy.

The navigation module generates a near-optimal obstacle-free path to the goal location using the

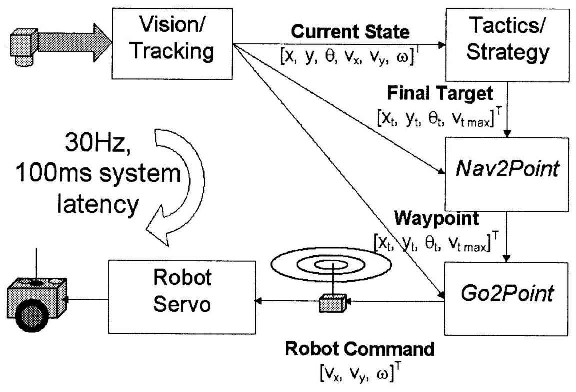
Vision/ Tracking
Current State
$[x, y, \theta, v_x, v_y, \omega]^T$
Tactics/ Strategy
Final Target
$[x_t, y_t, \theta_t, v_{t\,max}]^T$
30Hz, 100ms system latency
Nav2Point
Waypoint
$[x_t, y_t, \theta_t, v_{t\,max}]^T$
Robot Servo
Go2Point
Robot Command
$[v_x, v_y, \omega]^T$

Fig. 4 The CMDragons control architecture based on way-point primitives

beliefs stored in the world model. Based on this path, the motion control module calculates a trajectory to achieve a short-term target way point on the path that does not collide with obstacles. Using this trajectory, a velocity command is issued to the robot hardware to execute.

Because of the dynamic nature of robot soccer, both navigation and motion control are recalculated each frame, for each robot. This places strict computational limitations on each of these modules. A fast randomized path planner [**17**] was developed and implemented on the basis of the rapidly exploring random trees (RRT) algorithm [**18**]. Similarly, a trapezoid-based near-optimal motion control algorithm for quickly generating robot motion commands was developed [**8**].

2.2.4 Robot hardware

Each robot is an omnidirectional platform capable of spinning while driving in any direction. Each robot is equipped with a ball manipulation device that includes a solenoid actuated 'kicker' and a motorized 'dribbler'. The kicker moves an aluminium plate to contact with the ball, propelling it at speeds of around 3.5–4 m/s. The dribbler is a rubber-coated bar that is mounted horizontally at ball height and connected to a motor. As the bar spins against a ball, it causes the ball to spin backwards against the robot, thereby allowing the robot to move around effectively with the ball. Each robot has an on-board processor and runs local velocity-based servo loops using integrated encoder feedback and standard proportional–integral–derivative (PID) control techniques [**19**]. Additionally, the robot is equipped with a frequency-modulated radio receiver which it uses to receive movement commands from the external computer.

3 THE STP ARCHITECTURE

This section overviews the STP architecture, leading to a detailed discussion of skills, tactics, and plays.

3.1 Goals

The presence of an opponent has many, sometimes subtle, effects on *all* levels and aspects of control. Generating robust behaviour that responds to the actions of the opponent is a significant challenge. The challenges for team control are as follows.

1. A temporally extended sequence of coordinated activities must be executed among team members towards some longer-term goal while simultaneously responding as a team to unexpected events both fortuitous and disastrous.
2. There must be the ability to respond as a team to the capabilities, tactics, and strategies of the opponent.
3. Robust behaviour must be executed despite sensor limitations and world dynamics.
4. A modular compact architecture must be provided with facilities for easily configuring team play, and for analysing the performance of the decision-making process.

The first and second goals are direct impacts from controlling a team of robots in an adversarial environment. It is desirable for the team control architecture to generate robust behaviour that increases the chance of future opportunities against the opponent. Whenever such opportunities arise, whatever the cause, the team must take advantage of this opportunity immediately. Conversely, if an opportunity arises for the opponent, the team must respond quickly and intelligently to minimize the damage that the opponent can cause. Such responsive behaviour must occur throughout the architecture. Building a responsive team while overcoming the usual limitations of real-world sensors, such as latency, noise, and uncertainty, is the major goal of the STP framework.

In robot soccer, robust development is a significant issue. Many teams have gone through bad experiences caused by poor development procedures or facilities. Thus, a good architecture is compact and modular such that changes in one module have a minimal impact on the operation of another module. Given the number of parameters in a complex team architecture, the ability to reconfigure those parameters easily and to analyse the performance of different parameter settings is extremely useful to the development cycle.

3.2 Skills, tactics, and plays

The STP architecture was developed to achieve the goals of responsive adversarial team control. The key component of STP is the division between single-robot behaviour and team behaviour. In short, team behaviour results from executing a coordinated sequence of single-robot behaviours for each team member. *Plays*, *tactics*, and *skills*, and how they interact for a team of N robots, are now defined.

A *play* P is a fixed team plan which consists of a set of applicability conditions, termination conditions, and N roles, one for each team member. Each role defines a sequence of tactics $T^1, T^2, \ldots$ and associated parameters to be performed by that role in the ordered sequence. Assignment of roles to team members is performed dynamically at run time. Upon role assignment, each robot i is assigned its tactic T_i to execute from the current step of the sequence for that role. Tactics, therefore, form the action primitives for plays to influence the world. The full set of tactics can be partitioned into active tactics and non-active tactics. Active tactics are those involved with ball manipulation. There is only one active tactic among the roles per step in the sequence. The successful completion of the active tactic is used to trigger the transition to the next step in the sequence for *all* roles in the play. Plays are discussed in greater detail in section 5.

A *tactic* T encapsulates a single-robot behaviour. Each robot i executes its own tactic as created by the current play P. A tactic T_i determines the skill state machine SSM_i to be executed by the robot i. If the tactic is active, it also contains evaluation routines to determine whether the tactic has completed. If the skill state machine differs from that executed previously, then execution begins at the first skill in the state machine, i.e. S_i. If the skill state machine did not change, then execution continues at the last skill transitioned to. The tactic T_i also sets parameters $SParams_i$ to be used by the executing skill S_i. Thus, skills form the action primitives for tactics.

A *skill* S is a focused control policy for performing some complex action. Each skill is a member of one, or more, skill state machines $SSM_1, SSM_2, \ldots$. Each skill S determines what skill it transitions to S' based upon the world state, the time skill S has been executing for, and the executing tactic for that robot. The executing tactics may reset, or change and reset, the executing skill state machine. Each skill can command the robot to perform actions either directly, through motion control, or through navigation. If commanded through navigation, navigation will generate an intermediate, obstacle free way point for motion control which will then generate a command to send to the robot.

Both skills and tactics must evaluate the world state, in sometimes complex ways, to make useful decisions. For example, some tactics determine the best position to move to in order to receive a pass. Alternatively, some defensive tactics evaluate which opponent robot might move to receive a pass and where to go to prevent the opponent from achieving this goal. To prevent unnecessary duplication, and to modularize the architecture more, these evaluations are extracted into an evaluation module which is usable by both tactics and skills. Tactics, skills, and evaluations are detailed in section 4.

Plays, tactics, and skills, form a hierarchy for team control. Plays control the team behaviour through tactics, while tactics encapsulate individual robot behaviour and instantiate actions through sequences of skills. Skills implement the focused control policy for actually generating useful actions. Table 1 shows the main execution algorithm for the STP architecture. The clear hierarchical arrangement of plays for team control, tactics for single-robot behaviour, and skills for focused control are shown.

4 TACTICS AND SKILLS FOR SINGLE-ROBOT CONTROL

Single-robot control in the STP architecture consists of tactics and skills. Tactics provide the interface for team control via plays, while skills provide the mechanisms for generating behaviour in a compact reusable way. First, tactics are described in greater depth, followed by skills, and finally the evaluation module.

4.1 Tactics

Tactics are the topmost level of single-robot control. Each tactic encapsulates a single-robot behaviour.

Table 1 The main STP execution algorithm

```
Process STP Execution
 1. CaptureSensors()
 2. RunPerception()
 3. UpdateWorldModel()
 4. $P \leftarrow$ ExecutePlayEngine()
 5. for each robot $i \in \{1, \ldots, N\}$
 6.     $(T_i, TParams_i) \leftarrow$ GetTactic($P, i$)
 7.     $(SSM_i, SParams_i) \leftarrow$ ExecuteTactic($T_i, TParams_i$)
 8.     if NewTactic($T_i$) then
 9.         $S_i \leftarrow SSM_i(0)$
10.     $(command_i, S_i') \leftarrow$ ExecuteStateMachine($SSM_i, S_i, SParams_i$)
11.     $robot\_command_i \leftarrow$ ExecuteRobotControl($command_i$)
12.     SendCommand($i, robot\_command_i$)
```

Each tactic is parametrized allowing for more general tactics to be created which are applicable to a wider range of world states. Through parametrization a wider range of behaviour can be exhibited through a smaller set of tactics, making play design easier. Table 2 provides the list of tactics that were implemented for robot soccer. The meaning of each tactic should be reasonably obvious from the tactic name.

During execution, one tactic is instantiated per robot. A tactic, as determined by the executing play, is created with the parameters defined for the play. That tactic then continues to execute until the play transitions to the next tactic in the sequence. As described above, each tactic instantiates action through the skill layer. In short, the tactics determine which skill state machine will be used and sets the parameters for executing those skills. Example parameters include target way points, target points to shoot at, opponents to mark, and so on. Different tasks may use many of the same skills but provide different parameters to achieve the different goals of the tactic. The shooting and passing tactics are good examples. The skills executed by the two are very similar, but the resulting behaviour can be quite different due to the different parameter assignments. Finally, each tactic may store any local state information it requires to execute appropriately.

Table 2 List of tactics with their accepted parameters

Active Tactics
shoot (Aim | Noaim | Deflect ⟨*role*⟩)
steal [⟨*coordinate*⟩]
clear
active_def [⟨*coordinate*⟩]
pass ⟨*role*⟩
dribble_to_shoot ⟨*region*⟩
dribble_to_region ⟨*region*⟩
spin_to_region ⟨*region*⟩
receive_pass
receive_deflection
dribble_to_position ⟨*coordinate*⟩ ⟨*theta*⟩
position_for_start ⟨*coordinate*⟩ ⟨*theta*⟩
position_for_kick
position_for_penalty
charge_ball
Non-Active Tactics
position_for_loose_ball ⟨*region*⟩
position_for_rebound ⟨*region*⟩
position_for_pass ⟨*region*⟩
position_for_deflection ⟨*region*⟩
defend_line ⟨*coordinate*-1⟩ ⟨*coordinate*-2⟩ ⟨*min-dist*⟩ ⟨*max-dist*⟩
defend_point ⟨*coordinate*-1⟩ ⟨*min-dist*⟩ ⟨*max-dist*⟩
defend-lane ⟨*coordinate*-1⟩ ⟨*coordinate*-2⟩
block ⟨*min-dist*⟩ ⟨*max-dist*⟩ ⟨*side-pref*⟩
mark ⟨*orole*⟩ (ball | our_goal | their_goal | shot)
goalie
stop
velocity ⟨*vx*⟩ ⟨*vy*⟩ ⟨*vtheta*⟩
position ⟨*coordinate*⟩ ⟨*theta*⟩

Table 3 shows the algorithm for the shoot tactic used to kick the ball at the goal or towards teammates for one-shot deflections at the goal. Not shown are the conditioning of the tactic decision tree on the parameters specified by the active play. In this case, the play can only disable deflection decisions. The tactic consists of evaluating the options of shooting directly at the goal, or shooting to a teammate to deflect or kick at goal in a so-called one-shot pass. Each option is assigned a score which, loosely, defines the likelihood of success. Much of the operation of determining the angles to shoot at and generating the score is pushed into the evaluation module, described in section 4.3.

The tactic (indeed nearly all tactics) make use of additive hysteresis in the decision-making process. Hysteresis is a necessary mechanism to prevent debilitating oscillations in the selected choice from frame to frame. Each action in the shoot tactic, as with any other tactic, takes a non-negligible period of time to perform that is substantially greater than a single-decision cycle at 30 Hz. With the dynamics of the environment further complicated by occlusion, noise, and uncertainty, it is often the case that two or more choices will oscillate over time in terms of its score. Without hysteresis, there will be corresponding oscillations in the action chosen. The end result is often that the robot will oscillate between distinctly different actions and effectively be rendered immobile. The physical manifestation of this behaviour, ironically, is that the robot appears to ‘twitch’ and to be ‘indecisive’. In most robot domains, such oscillations will degrade performance. In adversarial domains such as robot soccer, where it is important to carry out an action *before* the opponent can respond, such oscillations completely destroy the robot’s actions. Hysteresis provides a usable, easily understandable mechanism for preventing such oscillations and is used pervasively throughout the STP architecture.

4.2 Skills

Most tactics require the execution of a sequence of recognizable skills, where the actual sequence may depend upon the world state. An example skill sequence occurs when a robot tries to dribble the ball to the centre of the field. In this case, the robot, firstly, will go to the ball, secondly, will get the ball on to its dribbler, thirdly will turn the ball around if necessary, and then, finally, will push the ball towards the target location with the dribbler bar spinning. A

Table 3 Algorithm for the shoot tactic for shooting at goal directly or by one-shot passes to teammates. Each action is evaluated and assigned a score. The action with the best score better than the score for the previously selected action is chosen and its target passed to the running skill. The skill state machine used is the *Moveball* state machine

```
Tactic Execution shoot(i):
 1. bestscore ← 0
 2. (score, target) ← evaluation.aimAtGoal()
 3. if (was kicking at goal) then
 4.     score ← score + HYSTERESIS
 5. $SParam_i$ ← setCommand(MoveBall, target, KICK_IF_WE_CAN)
 6. bestscore ← score

 7. foreach teammate j do
 8.     if (evaluation.deflection(j) > THRESHOLD) then
 9.         (score, target) ← evaluation.aimAtTeammate(j)
10.         if (was kicking at player j) then
11.             score ← score + HYSTERESIS
12.         if (score > bestscore) then
13.             $SParam_i$ ← setCommand(MoveBall, target, KICK_IF_WE_CAN)
14.             bestscore ← score

15. if (No target found OR score < THRESHOLD) then
16.     target ← evaluation.findBestDribbleTarget()
17.     $SParam_i$ ← SetCommand(MoveBall, target, NO_KICK)
```

different sequence would be required if the ball were against the wall, or in the corner. Additional skills would be executed, such as pulling the ball off the wall, in order to achieve the final result.

In other work by the present authors, a hierarchical behaviour-based architecture was developed, where behaviours form a state machine with transitions conditioned on the observed state and internal state [**20**]. Although no use is made of the hierarchical properties of the approach here, use is made of the state machine properties to implement the sequence of skills that make up each tactic. Each skill is treated as a separate behaviour and forms a unique state in the state machine. In contrast with tactics, which execute until the play transitions to another tactic, each skill transitions to itself or another skill at each time step.

Each skill consists of three components: sensory processing, command generation, and transitions. Sensory processing consists of using or generating the needed sensory predicates from the world model. Commonly used sensors are generated once per frame, ahead of time, to prevent unnecessary duplication of effort. Command generation consists in determining the action for the robot to perform. Commands can be instantiated through the navigation module or motion control. In some cases, commands are sent directly to the robot. Transitions define the appropriate next skill that is relevant to the execution of the tactic. Each skill can transition to itself or to another skill. Transitions are conditioned on state variables set by the tactics or state machine variables, such as the length of time that the active skill has been running. This makes it possible to use the same skill in multiple sequences. A skill can be used for different tactics, or in different circumstances for the same tactic, thereby allowing for skill reuse and the minimizing of code duplication.

Table 4 shows our algorithm for the driveToGoal skill used to drive the ball towards the desired target, which is continually adjusted by the tactic as execution cycles. The skills first determines which skill it will transition to. If no skill is found, it transitions to itself. The decision tree shows conditioning on the active state machine, MoveBall in this case, and conditioning upon the active tactic. Decisions are also made using high-level predicates, e.g. *ball_on_front*, derived from the tracking data by the world model. References to the world are not shown to aid clarity.

4.3 Evaluation module

There are numerous computations about the world that need to be performed throughout the execution of plays, tactics, and skills in order to make good decisions. Many of these computations are evaluations of different alternatives and are often used numerous times. Aim evaluation is a good example, as the

Table 4 The DriveToGoal skill which attempts to push the ball towards the desired direction to kick. The transitions decision tree, which includes conditioning on the active tactic, the active state machine, and predicates derived from the world model, is shown. The command generation calculations are simplified here to aid clarity but require a number of geometric calculations to determine the desired target point

```
Skill Execution DriveToGoal($i$):
1. if ($SSM_i$ = MoveBall AND ball_on_front AND can_kick AND shot_is_good) then
2.     Transition(Kick)
3. if (ball_on_front AND ball_is_visble) then
4.     Transition(GotoBall)
5. if (robot_distance_from_wall < THRESHOLD AND $robot_s tuck$) then
6.     Transition(SpinAtBall)

   Command generation
7. $command_i$.navigate ← true
8. $command_i$.target ← calculateTarget()
```

same evaluation of alternatives is called at least 24 times during a single cycle of execution! All these evaluations are combined into a single module. There are three classes of evaluations that occur: aiming, defence, and target positions.

4.3.1 Aim evaluations

Aim evaluations determine the best angle for the robot to aim towards to kick the ball through a specified line segment while avoiding a list of specified obstacles. Using the world model, the aim evaluations determine the different open angles to the target. It then chooses the largest open angle with additive hysteresis if the last chosen angle, assuming that there is one, is still a valid option. The use of a line segment as the target allows the same evaluation to be used for aiming at the opponent's goal, for opponents aiming at the goal of the present authors' team, as well as for passes and deflections to teammates or from opponents to their teammates.

4.3.2 Defensive evaluations

Defensive evaluations determine where the robot should move so that it best defends a specified point or line segment. Although similar to target position evaluations, the technique used is quite different. There are a number of different variations of defensive evaluations for defending lines, points, or defending along a given line. Each evaluation uses similar techniques, but the point chosen and hence the behaviour generated vary and are useful in different situations.

The most commonly used defensive evaluation is line defences. For line defences, the evaluation attempts to blend between choosing a defensive point that is good if the ball could be kicked at any angle from its observed location and a point to intercept the ball if it were to remain moving at its current velocity. First a linear Gaussian is created to describe the desirability of each point on the line for defending against a static kick. The Gaussian is centred on the point that, when accounting for the robot size, equalizes the time that it would take for the robot to move to block a shot at either end of the defended segment. A second linear Gaussian is generated by predicting the ball motion forwards in time to where it crosses the defended line, and calculating the corresponding tracker uncertainty projected on to this line. Essentially, the faster the ball is kicked, the more certain its crossing point, which results in a much narrower taller Gaussian. In addition, obstacles along the trajectory can also add substantial uncertainty into the interception Gaussian. When these two Gaussian functions are multiplied, the result represents a smooth blending between the two alternatives. Generally, the static kick Gaussian dominates but, as the ball is kicked more rapidly towards the defence line, the interception Gaussian pushes the defender to intercept the current trajectory. Such a smooth shift is desirable to avoid having to develop techniques for deciding between intercepting or defending, and the corresponding hysteresis that would be required.

4.3.3 Target position evaluation

The final type of evaluation determines the best target position to achieve a given task. Examples include the best position to receive a deflection, the best position to acquire a loose ball, and the best location to dribble towards to get a shot at goal or to pass to a teammate. In each case, there is a range of competing criteria that the evaluation ideally would optimize that can often be represented as an objective function of some kind. For example, to receive a pass for a shot at goal the robot needs to get into a position that gives it a clear shot at goal, a reasonable deflection angle so that it has an opportunity to receive and kick the ball, and a clear shot to its teammate with ball possession. Clearly, an objective function could be written to describe this problem, and attempts could be made to find the optimal solution. This approach is problematic due to the computational constraint that only a fraction of the processor is available for this task, and it needs to be repeated many times during a single execution cycle. Additionally, the dynamics of the environment, combined with sensing noise, mean that the optimal point will invariably be unstable over time. Thus, the robot will never stabilize, which is essential for situations such as receiving a pass as its teammate needs a steady target. In many cases, however, if near-optimal values are considered, reasonably stable sets form over extended periods. Thus, an evaluation method is required with low computational requirements to find quasi-static near-optimal locations.

A sample-based approach to this problem has been taken. For each evaluation, a series of points are generated randomly drawn uniformly from the region of space of interested specified in the evaluation call. The objective function is evaluated at each point, and the best value is recorded. If there was a point chosen previously, its value is calculated and if it is within α standard deviations of the score of the best point, for some defined α, it is again selected as the target point. If there was no target point previously, or it is no longer α optimal, the best point is chosen as the target. Thus, the α value imparts a hysteresis effect as used in the other evaluations.

5 PLAYS FOR MULTI-ROBOT TEAM CONTROL

The final component of the STP architecture are plays. Plays form the highest level in the control hierarchy, providing strategic level control of the entire team. The strategic team problem involves selecting each robot's behaviour in order to achieve team goals, given a set of tactics, which are effective and parametrized individual robot behaviours. Team strategy is built around the concept of a *play* as a team plan, and the concept of a *playbook* as a collection of team plans. First the goals for the design of a team strategy system are explored and then how plays and playbooks achieve these goals is investigated.

5.1 Goals

Obviously the main criterion for a team strategy system is performance. A single monolithic team strategy that maximizes performance, however, is impractical. In addition, there is not likely to be a single optimal strategy independent of the adversary. Instead of focusing directly on team performance, a set of six simpler goals that the present authors believe are more practical and lead to strong overall team performance are as follows:

(a) coordinated team behaviour;
(b) temporally extended sequences of action (deliberative);
(c) inclusion of special-purpose behaviour for certain circumstances;
(d) ease of human design and augmentation;
(e) ability to exploit short-lived opportunities when they occur (reactive);
(f) on-line adaptation to the specific opponent.

The first four goals require plays to be able to express complex, coordinated, and sequenced behaviour among teammates. In addition, the language must be human readable to make play design and modification simple. These goals also require a powerful system capable of executing the complex behaviours that the plays describe. The fifth goal requires the execution system also to recognize and exploit opportunities that are not explicitly described by the current play. Finally, the sixth goal requires the system to alter its overall behaviour over time. Note that the strategy system requires both deliberative and reactive reasoning. The dynamic environment makes a strictly deliberative system unlikely to be able to carry out its plan, but the competitive nature often requires explicitly deliberative sequences of actions in order to create scoring opportunities.

First the novel play language, together with the coupled execution system, are introduced. Then how playbooks can provide multiple alternative strategies for playing against the unknown opponent is described.

5.2 Play specification

A play is a multi-agent plan, i.e. a joint policy for the entire team. The definition of a play, therefore, shares many concepts with classical planning. A play consists of four main components:

(a) applicability conditions;
(b) termination conditions;
(c) roles;
(d) sequence of tactics to execute per role.

Applicability conditions specify when a play can be executed and are similar to planning operator preconditions. Termination conditions define when execution is stopped and are similar to an operator's effects, although they include a number of possible outcomes of execution. The roles describe the actual behaviour to be executed in terms of individual robot tactics. The execution details can include a variety of optional information that can help to guide the play execution system. Now each of these components is looked at individually.

5.2.1 Applicability conditions

The conditions for a play's applicability can be defined as any logical formula of the available state predicates. The conditions are specified as a logical disjunctive normal form (DNF) using the APPLICABLE keyword, with each disjunct specified separately. In the example play in Table 5, the play can only be executed from a state where the offense predicate is true. The offense predicate is actually a fairly complex combination of the present and past possession of the ball and its present and past position on the field. Predicates can be easily added and Table 6 lists the current predicates used by our system. Note that predicates can also take parameters, as in the case of ball_x_gt *X*, which checks whether the ball is over the distance *X* down field.

Like preconditions in classical planning, applicability conditions restrict when a play can be executed. By constraining the applicability of a play, special-purpose plays can be designed for very specific circumstances. An example of such a play is shown in Table 7. This play uses the ball_in_their_corner predicate to constrain the play to be executed only when the ball is in a corner near the opponent's goal. The play explicitly involves dribbling the ball out of the corner to obtain a better angle for a shot at goal. Such a play only really makes sense when initiated from the play's applicability conditions.

Table 5 A simple example of a play

```
PLAY Naive Offense

APPLICABLE offense
DONE aborted !offense

ROLE 1
  shoot A
  none
ROLE 2
  defend_point {−1400 250} 0 700
  none
ROLE 3
  defend_lane {B 0  −200} {B 1175  −200}
  none
ROLE 4
  defend_point {−1400  −250} 0 1400
  none
```

Table 6 List of state predicates

Play predicates	
offense	our_kickoff
> defense	their_kickoff
> their_ball	our_freekick
> our_ball	their_freekick
> loose_ball	our_penalty
> ball_their_side	their_penalty
> ball_our_side	ball_x_gt *X*
ball_midfield	ball_x_lt *Y*
ball_in_our_corner	ball_absy_gt *Y*
ball_in_their_corner	ball_absy_lt *Y*
nopponents_our_side *N*	

5.2.2 Termination conditions

Termination conditions specify when the play's execution should stop. Just as applicability conditions are related to operator preconditions in classical planning, termination conditions are similar to operator effects. Unlike classical planning, however, there is too much uncertainty in execution to know the exact outcome of a particular play. The termination conditions list possible outcomes and associate a *result* with each possible outcome. The soccer domain itself defines a number of stopping conditions, e.g. the scoring of a goal or the awarding of a penalty shot. The play's termination conditions are in addition to these and allow for play execution to be stopped and a new play initiated even when the game itself is not stopped.

Termination conditions, such as applicability conditions, use logical formulae of state predicates. In

Table 7 A special-purpose play that is only executed when the ball is in an offensive corner of the field

```
PLAY Two Attackers, Corner Dribble 1

APPLICABLE offense in_their_corner
DONE aborted !offense
TIMEOUT 15

ROLE 1
   dribble_to_shoot { R { B 1100 800 } { B 700 800 } 300}
   shoot A
   none
ROLE 2
   block 320 900  −1
   none
ROLE 3
   position_for_pass { R { B 1000 0 } { B 700 0 } 500 }
   none
ROLE 4
   defend_line { −1400 1150 } { −1400  −1150 } 1100 1400
   none
```

addition to specifying a conjunction of predicates, a termination condition also specifies the result of the play if the condition becomes true. In the play specification, they are delineated by the DONE keyword, followed by the result, and then the list of conjunctive predicates. Multiple DONE conditions can be specified and are interpreted in a disjunctive fashion. In the example play in Table 5, the only terminating condition, beside the default soccer conditions, is if the team is no longer on offence (! is used to signify negation). The play's result is then 'aborted'.

The results for plays are as follows: succeeded, completed, aborted, and failed. These results are used to evaluate the success of the play for the purposes of reselecting the play later. This is the major input to the team adaptation system, which is described later. Roughly speaking, the results of succeeded and failed to mean that a goal was scored, or some other equivalently valuable result, such as a penalty shot, are used. The completed result is used if the play was executed to completion. For example, in the play in Table 5, if a robot was able to complete a shot, even if no goal was scored, the play is considered completed. In a defensive play, switching to offence may be a completed result in the DONE conditions. The aborted result is used when the play was stopped without completing.

Besides DONE conditions, there are two other ways in which plays can be terminated. The first is when the sequence of behaviours defined by the play are executed. As mentioned above, this gives the play the completed result. This will be described further when we examine the play execution system. The second occurs when a play runs for a long time with no other termination condition being triggered. When this occurs the play is terminated with an aborted result and a new play is selected. This allows the team to commit to a course of action for a period of time but recognizes that in certain circumstances a particular play may not be able to progress any further.

5.2.3 Roles

As plays are multi-agent plans, the main component are the roles. Each play has four roles, one for each non-goalie robot on the field. A role consists of a list of behaviours for the robot to perform in sequence. In the example play in Table 5, there is only a single behaviour listed for each role. These behaviours will simply be executed until one of the termination conditions apply. In the example play in Table 7, the first role has two sequenced behaviours. In this case the robot will dribble the ball out of the corner. After the first tactic finishes, the robot filling that role will switch to the shoot tactic and try to manipulate the ball towards the goal.

Sequencing also requires coordination, which is a critical aspect of multi-agent plans. Coordination in plays requires all the roles to transition simultaneously through their sequence of behaviours. For example, consider the more complex play in Table 8. In this play, one player is assigned to pass the ball to another player. Once the pass behaviour is completed, *all* the roles transition to their next behaviour, if one is defined. So, the passing player will switch to a mark behaviour, and the target of the pass will switch to a behaviour to receive the pass, after which it will switch to a shooting behaviour.

Roles are not tied to any particular robot. Instead, they rely on the play execution system to carry out this role assignment. The order of the roles presented in the play act as hints to the execution system for filling the roles. Roles are always listed in order of priority. The first role is always the most important and usually involves some manipulation of the ball. This provides the execution system with the knowledge needed to select robots to perform the roles and also for role switching when appropriate opportunities present themselves.

(a) *Tactics in roles.* The different behaviours that can be specified by a role are the individual robot tactics that were discussed in section 4.1. As mentioned, these tactics are highly parametrized behaviours. For example, the defend_point tactic

Table 8 A complex play involving sequencing of behaviours

```
PLAY Two Attackers, Pass

APPLICABLE offense
DONE aborted !offense

OROLE 0 closest_to_ball

ROLE 1
  pass 3
  mark 0 from_shot
  none
ROLE 2
  block 320 900 −1
  none
ROLE 3
  position_for_pass { R { 1000 0 } { 700 0 } 500 }
  receive_pass
  shoot A
  none
ROLE 4
  defend_line { −1400 1150} {−1400 −1150} 1000 1400
  none
```

takes a point on the field and minimum and maximum ranges. The tactic will then position itself between the point and the ball, within the specified range. By allowing for this large degree of parametrization the different behaviours can be combined into a nearly infinite number of play possibilities. The list of parameters accepted by the different tactics is shown in Table 2.

(b) *Coordinate systems.* Many of the tactics take parameters in the form of 'coordinates' or 'regions'. These parameters can be specified in a variety of coordinate systems allowing for added flexibility in specifying plays in general terms. Coordinates can be specified either as absolute field position or as ball-relative field positions. In addition, the positive y axis can also be specified to depend on the side of the field that the ball is on, the side of field that the majority of the opponents are on, or even a combination of these two factors. This allows tremendous flexibility in the specification of the behaviours used in plays. Regions use coordinates to specify non-axis aligned rectangles as well as circles. This allows, for example, a single play to be general with respect to the side of the field and position of the ball.

5.2.4 Execution details

The rest of the play specification are execution details, which amount to providing hints to the execution system about how to execute the play. These optional components are timeout and opponent roles. The timeout overrides the default amount of time a play is allowed to execute before aborting the play and selecting a new play.

Opponent roles allow robot behaviours to refer to opponent robots in defining their behaviour. The play in Table 8 is an example of this. The first role switches to marking one of the opponents after it completes the pass. The exact opponent that is marked depends upon which opponent was assigned to opponent role 0. Before the teammate roles are listed, opponent roles are defined by simply specifying a selection criteria for filling the role. The example play uses the closest_to_ball criterion, which assigns the opponent closest to the ball to fill that role and consequently to be marked following the pass. Multiple opponent roles can be specified and they are filled in turn using the provided criterion.

5.3 Play execution

The play execution module is responsible for actually instantiating the play into real robot behaviour; i.e. the module must interpret a play by assigning tactics to actual robots. This instantiation consists of key decisions: role assignment, role switching, sequencing tactics, opportunistic behaviour, and termination.

Role assignment uses tactic-specific methods for selecting a robot to fill each role, in the order of the role's priority. The first role considers all four field robots as candidates to fill the role. The remaining robots are considered to fill the second role, and so on. Role switching is a very effective technique for exploiting changes in the environment that alter the effectiveness of robots fulfilling roles. The play executor handles role switching using the tactic-specific methods for selecting robots, using a bias towards the current robot filling the role. Sequencing is needed to move the entire team through the sequence of tactics that make up the play. The play executor monitors the current *active player*, i.e. the robot whose role specifies a tactic related to the ball (see Table 2). When the tactic succeeds, the play is transitioned to the next tactic in the sequence of tactics, for *each* role. Finally, opportunistic behaviour accounts for changes in the environment where a very basic action would have a valuable outcome. For example, the play executor evaluates the duration of time and potential success of each robot shooting immediately. If an opportunistic behaviour can be executed quickly enough and with a high likelihood of success, then the robot immediately switches its behaviour to take advantage of the situation. If the opportunity is then lost, the robot returns to executing its role in the play.

The play executor algorithm provides basic behaviour beyond what the play specifies. The play executor, therefore, simplifies the creation of plays, since this basic behaviour does not need to be considered in the design of plays. The executor also gives the team robustness to a changing environment, which can cause a play's complex behaviour to be no longer necessary or to require some adjustment to the role assignment. It also allows for fairly complex and chained behaviour to be specified in a play, without fear that short-lived opportunities will be missed.

The final consideration of play execution is termination. How plays specify their own termination criteria, either through predicates or a timeout, have already been described. The executor checks these conditions and also checks whether the play has completed its sequence of behaviours, as well as checking incoming information from the referee. If the final active tactic in the play's sequence of tactics completes, then the play is considered to have completed and is terminated. Alternatively, the game may be stopped by the referee to declare a penalty, to award a free kick, to award a penalty kick, to declare a score, and so on. Each of these conditions terminates the play but also may effect the determined outcome of the play. Goals are always considered successes or failures, as appropriate. Penalty kicks are also considered play successes and failures. A free kick for the present authors' team deems the play as completed, while a free kick for the opponent sets the play outcome to aborted. Play outcomes are the critical input to the play selection and adaptation system.

5.4 Playbook and play selection

Plays define a team plan. A playbook is a collection of plays and, therefore, provides for a whole range of possible team behaviours. Playbooks can be composed in a number of different fashions. For example, it could be ensured that for all possible game states there exists a single applicable play. This makes play selection simple since it merely requires executing the one applicable play from the playbook. A more interesting approach is to provide multiple applicable plays for various game states. This adds a play selection problem but also adds alternative modes of play that may be more appropriate for different opponents. Multiple plays also give options from among which adaptation can select. In order to support multiple applicable plays, a playbook also associates a weight with each play. This weight corresponds to how often the play should be selected when applicable.

Play selection, the final component of the strategy layer, then amounts to finding the set of applicable plays and selecting one based on the weights. Specifically, if $p_{1\ldots k}$ are the plays whose applicability condition are satisfied, and w_i is their associated weight, then p_j is selected with probability

$$\Pr(p_j|\mathbf{w}) = \frac{w_j}{\sum_{i=1}^{k} p_i}$$

Although these weights can simply be specified in the playbook and left alone, they also are the parameters that can be adapted for a particular opponent. A weighted experts algorithm (e.g. randomized weighted majority [**21**] and Exp 3 [**22**]) tailored to our specific domain is used to adapt the play weights during the course of the game. The weight changes were based on the outcomes from the play execution. These outcomes include obvious results such as goals and penalty shots, as well as the plays' own termination conditions and timeout factors. These outcomes are used to modify the play weights so as to minimize the play selection regret, i.e. the success that could have been achieved if the optimal play had been known in advance less the actual success achieved. This adaptation has been described elsewhere in more detail [**23**].

5.5 Achieving our goals

The play-based strategy system achieves all six goals that were set out in section 5.1. Sequences of synchronized actions provide a mechanism for coordinated team behaviour, as well as deliberative actions. Applicability conditions allow for the definition of special-purpose team behaviour. The play execution system handles moments of opportunity to allow for the team to have a reactive element. Incorporating all this into a human-readable text format makes adding and modifying plays quite easy. Finally, the ability to assign outcomes to the execution of plays is also the key capability used to adapt the weights used in play selection, achieving the final goal of a strategy system.

6 RESULTS AND DISCUSSION

RoboCup competitions provide a natural method for testing and evaluating techniques for single-robot and team control against a range of unknown opponents of varying capabilities and strategies. Indeed, this is a major focus for the competitions. The STP architecture has been evolved through feedback from competitions. Here the results derived from the RoboCup 2003 competition are mainly reported on but they also include anecdotal results from the following:

(a) RoboCup 2003, held in July in Padua, Italy; international competition with 21 competitive teams; CMDragons finished fourth (see http://www.robocup2003.org);
(b) RoboCup American Open 2003, held in May in Pittsburgh, USA; regional competition open to American continent teams; included ten teams from USA, Canada, Chile, and Mexico; CMDragons won first place (see http://www.americanopen03.org);
(c) RoboCup 2002, held in June in Fukuoka, Japan; international competition with 20 competitive teams; CMDragons were quarter-finalists (see http://www.robocup2002.org).

6.1 Team results

Overall, the STP architecture achieves the goals outlined in section 3.1. Using it, the present authors' team is able to respond quickly to unexpected situations while carrying out coordinated actions that increase the likelihood of future opportunities. The system is able to execute complex plays involving multiple passes and dribbling; however, because of the risk of losing the ball, real game plays do not exceed dribbling with one pass for a deflection on goal or a one-shot pass on goal. A one-shot pass is where one robot passes to another, which then takes a shot on goal. Indeed, such one-shot passes were responsible for a number of goals. Figure 5 shows an example from the game against ToinAlbatross from Japan.

The STP architecture is responsive to opportunistic events, both fortuitous and negative. Figure 6 shows an example of an opportunistic event occurring during an attacking manoeuvre against RoboDragons from Japan. The result was a goal that would not have occurred had the architecture persisted with its team plan. It is interesting to note that the whole episode occurs in less than 1 s. Figure 7 shows the effectiveness of dynamic role switching during a play, which results in smoother execution of the play.

The architecture is modular and reconfigurable. As an example of this aspect, at the RoboCup 2003 competition the playbook used by the team during the round-robin phase was completely rewritten.

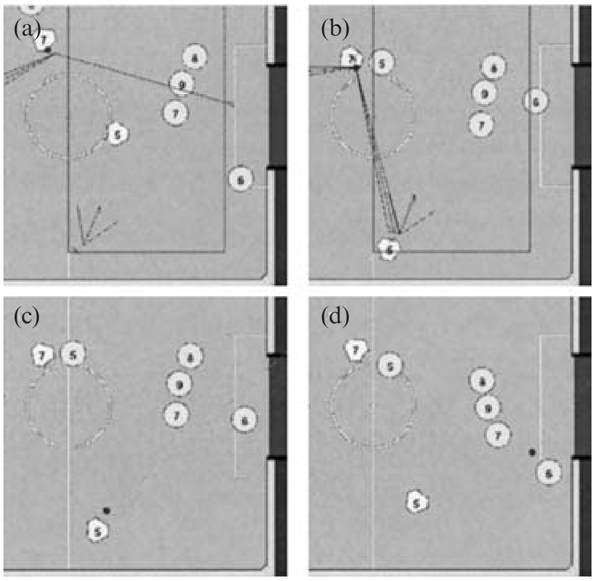

Fig. 5 Example of a deflection goal against ToinAlbatross from Japan. The dark lines show debugging output from the tactics and the light line shows the tracked ball velocity. (a) The shooting robot is unable to take a shot; robot 5 begins to move to a good deflection point. (b) The kicker is lined up and its target zone is on robot 5. (c), (d) The kick and resulting deflection employed to score a goal. The entire sequence takes less than 1 s.

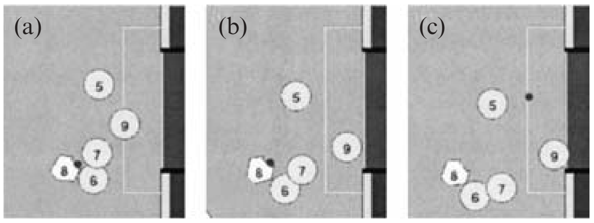

Fig. 6 Example of opportunism leading to a goal. A log sequence for RoboCup 2003 against RoboDragons is shown. (a) The robot gets the ball. (b), (c) Unexpectedly, a gap opens on goal. The robot moves and shoots to score. The entire sequence takes 15 frames, or 0.5 s.

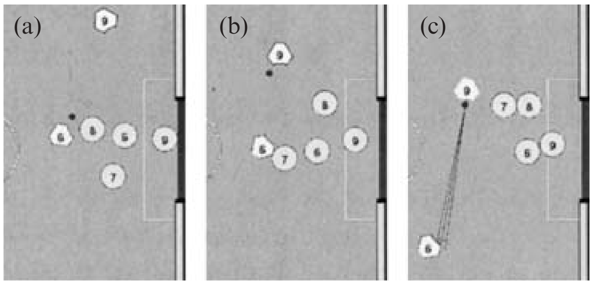

Fig. 7 Example of role switching. (a) Here the first robot is the active player, but the ball rolls too fast away from it. (b), (c) The second player smoothly takes over this task, while the first player moves out to receive a deflection. Taken from the game against RoboDragons

Modularity helps in making changes while minimizing the impact on the rest of the system. Reconfigurability is achieved through the play language, and use of configuration files to specify parameters for tactics and skills.

To demonstrate the need for different plays, and implicitly the need for different tactics to enable the implementation of a range of different plays, we compared the results of the play weights after the first half for two different games. Tables 9 and 10 show the weights at the end of the first half for the game against ToinAlbatross from Japan and Field Rangers from Singapore respectively. The weights and selection rates indicate the successfulness of each play. Different strategies are required

Table 9 Offense weights at the end of the first half for game against ToinAlbatross

Play	Weight	Selection	Selection (%)
o1_deep_stagger	0.021	6	10.3
o1_points_deep	2.631	11	19.0
o2_deflection_deep	0.280	40	69.0
o1_points_deep_deflections	0.015	1	1.7

Table 10 Offensive weights at the end of the first half for game against Field Rangers

Play	Weight	Selection	Selection (%)
o1_deep_stagger	1.080	23	50.00
o1_points_deep	0.098	2	4.35
o2_deflection_deep	1.123	17	36.96
o1_points_deep_deflections	0.657	4	8.70

to play effectively against the different styles of each opponent. The different in play weights clearly shows this. Therefore the conclusion is drawn that a diversity of tactics (and correspondingly a diversity of plays) is a useful tool for adversarial environments.

6.2 Single-robot results

Figure 8 shows a sequence of frames captured from the log of the game against RoboDragons from Japan. The robot shown is executing the shoot tactic and progresses through a series of skills determined by the progression of the world state. Given different circumstances, say if the ball were against the wall or in the open, the sequence of executed skills would be different. As with the play opportunism, the entire sequence occurs in only a few seconds.

Given the wide range of world states that occur during a game, and the need to execute different skill sequences for different world states, it becomes difficult to analyse the performance of the skill state machine. Consequently, it becomes difficult to determine how to improve its performance for future games. A number of logging techniques have been developed to aid in this analysis. The logging techniques take three forms. During development and game play, statistics are recorded for the transitions

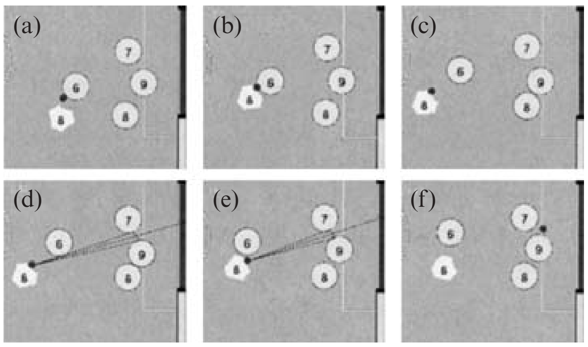

Fig. 8 An example shoot sequence taken from the RoboCup 2003 round-robin game of CMDragons'03 versus RoboDragons. (a), (b) The robot first executes the steal_ball skill. (c) This is followed by goto_ball. (d) Once the ball is safely on the robot's dribbler, it begins drive_to_goal to aim at the selected open shot at goal or to drive to a point where it can take the shot. (e) Upon being in position to take a good shot it kicks, leading to a scored goal

between skills as shown in Table 11 for the game against RoboDragons. During development, monitoring is also carried out for the presence of one node, and two-node loops, online. Thus, it can be quickly determined when skills transitions oscillate, or a skill fails to transition to another skill as appropriate.

6.3 Remaining issues

Based upon its performance in RoboCop competitions, the STP architecture provides many useful mechanisms for autonomously controlling a robot team in adversarial environments. There are issues that require further investigation in order to improve its overall capabilities, however.

Table 11 Robot log from RoboDragons game

Skill	Cause	Transition	Count	Count (%)
GotoBall	Command	Position	209	62.39
	GotAShot	Kick	3	0.90
	WithinDriveRange	DriveToGoal	67	20.00
	CanSteal	StealBall	33	9.85
	SpinOffWall	SpinAtBall	3	0.90
	CanBump	BumpToGoal	20	5.97
SteelBall	Command	Position	1	3.03
	BallAwayFromMe	GotoBall	14	42.42
	BallAwayFromOpp	GotoBall	18	54.55
DriveToGoal	CanKick	Kick	15	22.39
	BallTooFarToSide	GotoBall	52	77.61
BumpToGoal	Command	Position	1	5.00
	TargetTooFar	GotoBall	19	95.00
Kick	Command	Position	1	5.56
	BallNotOnFront	GotoBall	17	94.44
SpinAtBall	Command	Postion	1	33.33
	BallMoved	GotoBall	2	66.67
Position	Command	GotoBall	212	100.00

The greatest weakness of the current approach resides in the need to develop the skills and its corresponding state machine. The techniques and algorithms described here provided very useful tools for developing robot behaviour; however, development is still not a trivial process and much improvement can still be made. Each skill requires the development of a complex control algorithm, which is necessarily dependent upon the environment conditions and the capabilities of the robot hardware. Developing high-performance skills is a challenging process that requires creativity, knowledge of the robots capabilities, and large amounts of testing. Combining these skills into state machines is equally challenging. To do so, the decision tree to determine under what conditions a skill transitions to its counterpart must be accurately created. Loops caused by oscillations must be avoided, and it must be ensured that each transition occurs only in states for which the target skill can operate from. Finally, each skill typically requires a large number of parameters to define its behaviour and transition properties. Determining correct values for these parameters is a difficult and tedious process. Thus, future work will focus on easing the difficulties of skill development.

Another issue that needs further investigation is the dependence of skill execution on good sensor modelling. The unavoidable occurrence of occlusion, particularly during ball manipulation, has a severe impact on skill execution. Modelling the motion of the ball while it is occluded helps to reduce this impact but raises complications for when the ball modelling is incorrect. In particular, occasional observations of the ball may show inconsistencies with the modelled behaviour, causing the skills to change their mode of execution. Consequently, oscillations in output decisions occur, which detract from the performance of the skill. There is no easy solution to this problem, and it is an area of ongoing investigation.

7 RELATED WORK

There have been a number of investigations into control architectures for robot teams. Prime examples include Alliance [**1**], three-layer-based approaches [**16**] which build upon the single-robot versions (see, for example, reference [**15**]), or the more recent market-based approaches [**2**]. None of these architectures, however, has been applied to *adversarial* environments. As discussed throughout this article, adversarial environments create many novel challenges for team control that do not occur in non-adversarial domains. Within the domain of robot soccer, there have, naturally, been many varied approaches into single-robot and team control. Now the most relevant of these approaches is reviewed. Initially the focus is on teams that have demonstrated high levels of team cooperation and performance.

Beginning at single-robot control, there are a number of related approaches to the current work. In particular, the skills-based behaviour architecture employed was loosely inspired by the techniques used by the FU-Fighters team of Rojas and co-workers [**24**, **25**]. Their team is controlled by successive layers of reactive behaviours that operate at different characteristic time constants. There is a clear difference between a FU-Fighters' style approach and STP. Plays, although selected reactively, enable a team easily to execute sequences of actions that extend over a period of time. Moreover, with dynamic role switching, the team members may change their role assignments but still carry out the directives of the play as a whole. The state machine component of skills also contrasts against the purely reactive approach of FU-Fighters, whereby an extended sequence of actions can occur even in the presence of ball occlusion and noise.

The use of finite state machines for single-robot control is not a unique approach. Indeed, many researchers have investigated state machine approaches in a variety of contexts (see for example, references [**3**] and [**26**], or see reference [**4**] for more examples). The present authors' approach is unique, however, in that each skill is a state in the state machine sequence. The state sequence is a function of both the world and the delegating tactic. Finally, the active tactic *continually* updates the parameters used by the active skill as it modifies its decisions based on the world. For example, the TShoot tactic may switch its decision from shooting at one side of the goal to shooting at the other. The active skill, whatever it may be, will make a corresponding switch and perhaps transition to another skill depending upon the current situation. This combination of features makes the skill layer a unique approach.

At the team level, a number of teams use potential-field-based techniques for team control in the SSL (see, for example, references [**27**] to [**29**]). Potential-field-based team control is also popular outside the SSL, in the mid-size [**30**], simulation [**31**], and Sony AIBO legged leagues [**32**]. Potential fields are used to determine target field positions for moving or kicking. Essentially, the potential field value is determined for each cell in a grid covering the field. The shape of the potential field is formed by combining the usual attraction–repulsion operations common to potential field techniques [**33**, **34**]. Some teams also add to the potential field functions based on clear paths to the ball. This approach is similar to the use of evaluations described in section 4. The major difference occurs in the use of a sample-based approach to find a near-optimal point. The present authors found that a sample-based approach allows greater flexibility in defining the underlying objective function; additionally it avoids the issues of grid resolution and the computational effects of increasing the complexity of the evaluation function. Both techniques must use hysteresis or some similar mechanism.

Potential field techniques are also commonly used for navigation (see, for example, references [**35**] to [**37**]), although other reactive techniques are popular as well (see, for example, references [**38**] and [**39**]). Reactive navigation is quite successful in a dynamic and open environment but has been found by the present authors and others to be less effective in cluttered environments such as robot soccer (see, for example, reference [**40**]). Here fast planning-based approaches have been found to be much more powerful. See reference [**17**] for further discussion on this topic.

The Cornell Big Red team of D'Andrea *et al.* [**38**] utilizes a playbook approach that is similar to the use of plays described here. Their approach differs from that here, in that the playbook itself is a finite state machine where the plays are the states, rather than each play consisting of a set of states. As a result, the whole state machine is needed to have deliberative sequences of actions. The STP play-based approach, by encoding state transitions within plays, allows for multiple plays to be encoded to be operable for the same situations. As these multiple plays will utilize different sequences, it is reasonable to expect that the plays will have different degrees of effectiveness against different opponents. The STP approach, when combined with adaptation, allows for greater robustness of behaviour against a range of opponents because the best play to use in a given situation can be found from among a range of applicable plays.

8 CONCLUSIONS

In this article, the STP architecture has been presented for autonomous robot team control in adversarial environments. The architecture consists of plays for team control, tactics for encapsulating single-robot behaviour, and a skill state machine for implementing robot behaviour. The contributions of the STP architecture are to provide robust team coordination towards longer-term goals while remaining reactive to short-term opportunistic events. Secondly, the STP architecture is intended to provide team coordination that is responsive to the actions of the opponent team. Finally, the architecture is designed to be modular and to allow easy reconfiguration of team strategy and control parameters.

The STP architecture has been fully implemented in the small-size robot soccer domain and has been evaluated against a range of opponents of differing capabilities and strategies. Moreover, the techniques and algorithms employed have been evaluated across a number of international and regional competitions. In this article, the results based on these competitions that the present authors believe validate the STP approach have been presented.

Much work remains, however, to improve further the capabilities of play-based team control and skill-based single-robot behaviour. In particular, considerable future work is required to overcome the need to specify large numbers of parameters in order to gain high-performance skill execution. Future goals are to incorporate learning and adaptation at all levels in order to address this issue.

ACKNOWLEDGEMENTS

This research was sponsored by Grants DABT63-99-1-0013, F30602-98-2-013 and F30602-97-2-020. The information in this publication does not necessarily reflect the position of the funding agencies and no official endorsement should be inferred.

REFERENCES

1 **Parker, L.** Alliance: an architecture for fault-tolerant multi-robot cooperation. *IEEE Trans. Robotics Automn*, 1998, **14**(2), 220–240.
2 **Dias, M. B.** and **Stentz, A.** Opportunistic optimization for market-based multi-robot control. In Proceedings of (*IROS-2002*), September 2002.
3 **Brooks, R. A.** A robust layered control system for a mobile robot. *IEEE J. Robotics Automn*, March 1986, **2**(1), 14–23.
4 **Arkin, R. C.** *Behaviour-based Robotics*, 1998 (MIT Press, Cambridge, Massachusetts).

5 **Simmons, R., Fernandez, J., Goodwin, R., Koenig, S.** and **O'Sullivan, J.** Lessons learned from Xavier. *IEEE Robotics Automn Mag.*, June 2000, **7**(2), 33–39.

6 **Kitano, H., Asada, M., Kuniyoshi, Y., Noda, I.** and **Osawa, E.** RoboCup: the robot world cup initiative. In Proceedings of the First International Conference on *Autonomous Agents* (*Agents'97*) (Eds W. L. Johnson and B. Hayes-Roth), New York, 1997, pp. 340–347 (ACM Press, New York).

7 **Asada, M., Obst, O., Polani, D., Browning, B., Bonarini, A., Fujita, M., Christaller, T., Takahashi, T., Takokoro, S., Sklar, E.** and **Kaminka, G. A.** An overview of RoboCup-2002 Fukuoka/Busan. *AI Mag.*, Spring 2003, **24**(2), 21–40.

8 **Bruce, J., Bowling, M., Browning, B.** and **Veloso, M.** Multi-robot team response to a multi-robot opponent team. In Proceedings of the 2003 IEEE International Conference on *Robotics and Automation* (*ICRA'93*), Taiwan, May 2003 (in press).

9 **Veloso, M., Stone, P.** and **Han, K.** The CMUnited-97 robotic soccer team: perception and multiagent control. *Robotics Autonomous Systems*, 1999, **29**(2–3), 133–143.

10 **Veloso, M., Bowling, M., Achim, S., Han, K.** and **Stone, P.** CMUnited-98: a team of robotic soccer agents. In Proceedings of (*IAAI-99*), 1999.

11 **Veloso, M., Bowling, M.** and **Achim, S.** CMUnited-99: small-size robot team. In *RoboCup-99: Robot Soccer World Cup III* (Eds M. Veloso, E. Pagello and H. Kitano) 2000, pp. 661–662 (Springer-Verlag, Berlin).

12 **Browning, B., Bowling, M., Bruce, J., Balasubramanian, R.** and **Veloso, M.** Cmdragons'01—vision-based motion tracking and heterogeneous robots. In *RoboCup-2001: The Fifth RoboCup Competitions and Conferences* (Eds A. Birk, S. Coradeschi and S. Tadokoro), 2002 (Springer-Verlag, Berlin).

13 **Bruce, J., Balch, T.** and **Veloso, M.** Fast and inexpensive color image segmentation for interactive robots. In Proceedings of (*IROS-2000*), Japan, October 2000.

14 **Bruce, J.** and **Veloso, M.** Fast and accurate vision-based pattern detection and identification. In Proceedings of the 2003 IEEE International Conference on *Robotics and Automation* (*ICRA'03*), Taiwan, May 2003 (in press).

15 **Gat, E.** On three-layer architectures. In *Artificial Intelligence and Mobile Robots*, 1997 (MIT-AAAI Press).

16 **Simmons, R., Smith, T., Dias, M. B., Goldberg, D., Hershberger, D., Stentz, A.** and **Zlot, R.** A layered architecture for coordination of mobile robots. In *Multi-Robot Systems: From Swarms to Intelligent Automata*, 2002 (Kluwer, Dordrecht).

17 **Bruce, J.** and **Veloso, M.** Real-time randomized path planning for robot navigation. In Proceedings of (*IROS-2002*), Switzerland, October 2002 (in press).

18 **LaValle, S. M.** Rapidly-exploring random trees: a new tool for path planning. Technical Report 98-11, October 1998.

19 **Nise, N. S.** *Control Systems Engineering: Analysis and Design*, 1991 (Benjamin Cummings).

20 **Lenser, S., Bruce, J.** and **Veloso, M.** A modular hierarchical behavior-based architecture. In *RoboCup-2001: The Fifth RoboCup Competitions and Conferences* (Eds A. Birk, S. Coradeschi and S. Tadokoro), 2002 (Springer-Verlag, Berlin).

21 **Littlestone, N.** and **Warmuth, M.** The weighted majority algorithm. *Information and Computation*, 1994, **108**, 212–261.

22 **Auer, P., Casa-Bianchi, N., Freund, Y.** and **Schapire, R. E.** Gambling in a rigged casino: The adversarial multi-arm bandit problem. In Proceedings of the 36th Annual Symposium on *The Foundations of Computer Science*, Milwaukee, Wisconsin, 1995, pp. 322–331 (IEEE Computer Society Press, New York).

23 **Bowling, M., Browning, B.** and **Veloso, M.** Plays as effective multiagent plans enabling opponent-adaptive play selection. In Proceedings of the International Conference on *Automated Planning and Scheduling* (*ICAPS'04*), 2004 (in press).

24 **Rojas, R., Behnke, S., Liers, A.** and **Knipping, L.** FU-Fighters 2001 (global vision). In *RoboCup-2001: The Fifth RoboCup Competitions and Conferences* (Eds A. Birk, S. Coradeschi and S. Tadokoro), 2002 (Springer-Verlag, Berlin).

25 **Behnke, S.** and **Rojas, R.** A hierarchy of reactive behaviors handles complexity. In *Balancing Reactivity and Social Deliberation in Multi-agent Systems*, 2001, pp. 239–248 (Springer-Verlag, Berlin).

26 **Balch, T., Boone, G., Collins, T., Forbes, H., MacKenzie, D.** and **Santamaria, J.** Io, Ganymede and Callisto – a multiagent robot trash-collecting team. *AI Mag.*, 1995, **16**(2), 39–53.

27 **Kiat, N. B., Ming, Q. Y., Hock, T. B., Yee, Y. S.** and **Koh, S.** LuckyStar II. In *RoboCup-2001: The Fifth RoboCup Competitions and Conferences* (Eds A. Birk, S. Coradeschi and S. Tadokoro), 2002 (Springer-Verlag, Berlin).

28 **Wyeth, G., Ball, D., Cusack, D.** and **Ratnapala, A.** UQ RoboRoos: achieving power and agility in a small size robot. In *RoboCup-2001: The Fifth RoboCup Competitions and Conferences* (Eds A. Birk, S. Coradeschi and S. Tadokoro), 2002 (Springer-Verlag, Berlin).

29 **Nagasaka, Y., Murakami, K., Naruse, T., Takahashi, T.** and **Mori, Y.** Potential field approach to short term action planning in RoboCup F180 league. In *RoboCup-2002: Robot Soccer World Cup VI* (Eds P. Stone, T. Balch and G. Kraetzschmar), 2001, pp. 41–51 (Springer-Verlag, Berlin).

30 **Stancliff, S., Balasubramanian, R., Balch, T., Emery, R., Sikorski, K.** and **Stroupe, A.** CMU Hammerheads 2001 team description. In *RoboCup-2001: The Fifth RoboCup Competitions and Conferences* (Eds A. Birk, S. Coradeschi and S. Tadokoro), 2002 (Springer-Verlag, Berlin).

31 **Meyer, J., Adolph, R., Stephan, D., Daniel, A., Seekamp, M., Weinert, V.** and **Visser, U.** Decision-making and tactical behavior with potential fields. In *RoboCup-2002: Robot Soccer World Cup VI* (Eds R. Rojas, G. A. Kaminka, P. U. Lima), 2003, pp. 304–311 (Springer-Verlag, Berlin).

32 **Johansson, S. J.** and **Saffiotti, A.** Using the electric field approach in the RoboCup domain. In *RoboCup-2001: The Fifth RoboCup Competitions and Conferences* (Eds A. Birk, S. Coradeschi and S. Tadokoro), 2002 (Springer-Verlag, Berlin).

33 **Khatib, O.** Real-time obstacle avoidance for manipulators and mobile robots. *Int. J. Robotics Res.*, Spring 1986, **5**(1).

34 **Arkin, R. C.** Motor schema based navigation for a mobile robot. In Proceedings of the IEEE International Conference on *Robotics and Automation*, Raleigh, North Carolina, April 1987, pp. 264–271 (IEEE, New York).

35 **Damas, B. D., Lima, P. U.** and **Custodio, L. M.** A modified potential fields method for robot navigation applied to dribbling in robotic soccer. In *RoboCup*-2001: *The Fifth RoboCup Competitions and Conferences* (Eds A. Birk, S. Coradeschi and S. Tadokoro), 2002 (Springer-Verlag, Berlin).

36 **Emery, R., Balch, T., Shern, R., Sikorski, K.** and **Stroupe, A.** CMU Hammerheads team description. In *RoboCup-2000: Robot Soccer World Cup IV* (Eds P. Stone, T. Balch and G. Kraetzschmar), 2001, pp. 575–578 (Springer-Verlag, Berlin).

37 **Wyeth, G., Tews, A.** and **Browning, B.** UQ Robo-Roos: kicking on to 2000. In *RoboCup-2000: Robot Soccer World Cup IV* (Eds P. Stone, T. Balch and G. Kraetzschmar), 2001, pp. 555–558 (Springer-Verlag, Berlin).

38 **D'Andrea, R., Kalmar-Nagy, T., Ganguly, P.** and **Babish, M.** The Cornell RoboCup team. In *RoboCup-2000: Robot Soccer World Cup IV* (Eds P. Stone, T. Balch and G. Kraetzschmar), 2001, pp. 41–51 (Springer-Verlag, Berlin).

39 **Bowling, M.** and **Velos, M.** Motion control in dynamic multi-robot environments. In *RoboCup-99: Robot Soccer World Cup III* (Eds M. Veloso, E. Pagello and H. Kitano) 2000, pp. 222–230 (Springer-Verlag, Berlin).

40 **Weigel, T., Kliener, A., Diesch, F., Dietl, M., Gutmann, J.-S., Nebel, B., Stiegeler, P.** and **Szerbakowski, B.** CS Freiburg 2001. In *RoboCup-2001: The Fifth RoboCup Competitions and Conferences* (Eds A. Birk, S. Coradeschi and S. Tadokoro), 2002 (Springer-Verlag, Berlin).