

# Multi-Agent 层次协作模型在 RoboCup 中的应用

余 斌<sup>1,2</sup>, 徐 怡<sup>1</sup>, 李龙澍<sup>1</sup>, 李学俊<sup>1</sup>

(1. 安徽大学 计算机科学与技术学院, 安徽 合肥 230039;

2. 安庆师范学院 计算机与信息学院, 安徽 安庆 246011)

**摘 要:** RoboCup 是多 Agent 系统研究的一个很好的平台。建立一支成功的机器人足球队需要很多领域的知识, 合理的模型结构和 Agent 之间的协调与协作是 RoboCup 比赛中赢球的关键所在, 协调与协作是多 Agent 系统研究的重要课题。针对 RoboCup 仿真比赛的多智能体协作问题, 分析了目前在 RoboCup 中的几个典型多智能体协作模型, 提出一种三层的 Multi-Agent 层次协作模型, 它包括全局层、局部层和个体层。并通过实验证明该模型是合理的、有效的。

**关键词:** RoboCup; 多智能体系统; 协调; 协作

**中图分类号:** TP242.6

**文献标识码:** A

**文章编号:** 1673-629X(2006)06-0180-04

## Application of Multi-Agent Layered Collaboration Model in RoboCup Simulation Match

YU Bin<sup>1,2</sup>, XU Yi<sup>1</sup>, LI Long-shu<sup>1</sup>, LI Xue-jun<sup>1</sup>

(1. School of Computer Science & Technology, Anhui Univ., Hefei 230039, China;

2. School Computer & Information, Anqing Normal College, Anqing 246011, China)

**Abstract:** Robocup is an excellent platform for Multi-Agent. Various fields knowledges make up of a success robocup team. Reasonable model-structure and the correspondency & collaboration are the key to prevail in the robocup matches, so they are the important problem in the Multi-Agent researchs. Aiming at the problem of the Multi-Agent collaboration in robocup emulational matches, this paper analyses some classic Multi-Agent collaboration model in robocup, and provides a three layers model for Multi-Agent collaboration, which concludes the world layer, the local layer and the individual layer. We have experimented to prove that the three layers model is reasonable and effective.

**Key words:** RoboCup; multi-agent; correspondency; collaboration

### 0 引言

RoboCup(RobotWorldCup, 机器人足球比赛)是多 Agent 系统(Multi-Agent System, MAS)一个特殊的研究领域。机器人足球赛是由 Agent 组成的两支足球队进行足球比赛, RoboCup 是一个存在协商、协作和对抗的 MAS 问题。RoboCup 的主要目的就是通过提供一个标准的、易于评价的比赛平台, 促进 MAS 的研究与发展。

多 Agent 系统研究基于理性 Agent 的假设, 与协调一组可能预先存在的自主 Agent 的智能行为有关, 研究重点在于 Agent 与 Agent 之间的交互, 即 Agent 为了联合采取行动或求解问题, 如何协调各自的知识、目标、策略和规划。

MAS 的特点:

- 1) 每个 Agent 具有有限的信息和有限的能力;
- 2) 系统控制是分布式的;
- 3) 数据是集中式的;
- 4) 计算是异步的。

在 RoboCup 之前, 国际象棋一直被认为是 AI 的标准问题, 国际象棋是回合制、静态的、集中控制和封闭的。而机器人足球则是动态、实时、分布控制、开放的, 所以 RoboCup 和 MAS 的特点很吻合。在 RoboCup 仿真比赛的 SoccerServer 中规定: Player 可以具有有限的预定义信息; Player 的通信只能通过 SoccerServer 完成, 并具有范围限制; Player 的体力有限; 由 Server 传给 Player 的环境信息是有噪声的; 比赛处于一个动态和实时的环境当中。为了完成比赛任务, 每个队员, 即一个 Agent, 必须具有如下能力:

- a. 个人技术。对应于人类足球队员的个人能力。
- b. 决策能力根据比赛的实际情况, 进行实时决策, 以决定下一步的动作。
- c. 合作能力。可以和其他 Agent 合作完成一定的子

收稿日期: 2005-09-22

作者简介: 余 斌(1981-), 男, 安徽安庆人, 硕士研究生, 研究方向为智能软件; 李龙澍, 教授, 博导, 研究方向为智能软件体系结构和组件技术。

目标。

d. 学习能力。包括在线学习(比赛过程中通过学习来判断对方的行为)和离线学习(主要用于学习个人技术)。

e. 通讯能力。与本队的其他队员实现有限的信息交换(要受到 SoccerServer 的限制)并不被对方破译。

Agent 感知环境的信息主要是通过视觉的方式,听觉信息的传递受到通讯带宽的严重限制,因此 Agent 之间的信息交互不能简单地通过传统的 Agent 通讯语言,如 KQML<sup>[1]</sup>来实现,需要重新设计。系统的鲁棒性是至关重要的,每个 Agent 的决策不能受到任何环境因素的限制。

## 1 仿真比赛的球队结构

在 RoboCup 中,每个 Agent 的基本能力都是一致的,最基本的动作也是一样的,但是如何组合它们,成为较高层次的组合动作,成为 Agent 球队结构的核心内容。有效的组合动作将降低问题搜索空间。参加 RoboCup 比赛的各国机器人球队结构差异很大,如使用 BDI 模型、反应式结构,但用得最多的是基于层次结构的混合结构模型<sup>[2]</sup>。

### 1.1 基于 BDI 模型的慎思结构

在一个动态和不确定的域中,如 RoboCup 比赛,建立一个统一的、稳定的环境模型是必要的。通过一个固定的环境模型,短期的错误信息需要被更正,不精确的信息需要重新进行估算,通过推理获得信息。为了能够实现最终的目标,每个 Agent 需要在其内部维护一个稳定的环境模型,含有一个稳定的信念(belief)[Hans-Dieterburkhardt et al, 1998]。有多支球队使用 BDI 模型,其中最著名的 ATHumboldtTeam 曾获得 1998 年比赛的亚军。ATHumboldtTeam 的结构特点是按照 BDI 模型的结构建立全队的规划过程。

- 1) Belief: 环境模型;
- 2) Desires: 从固定目标库中选择的目标;
- 3) Intention: 表现为两个阶段的规划过程。

### 1.2 反应结构

从传感器进入的数据同时传给多个行为模块,如果某些模块的条件满足,则这些行为就不会有输出。但究竟哪一个被执行,或全部执行,需要由 Arbiter 模块进行判决。一般常用的算法有:静态优先级算法、动态优先级算法和混合算法。选择哪个算法需要由具体应用环境而定。

### 1.3 层结构

层结构一般可分为两层或三层:通讯层、控制层和决策层。Sohota 发展了一个决策技术<sup>[3]</sup>,称之为:reactivedeliberation,它在多个事先定义好的行为之间进行选择。Stone & Veloso 建立了层学习模式,首先学习低级技术,然后按照层的顺序,越来越高的层技术和行为被加入。此系统与 Sohota 系统的最大差别在于 Stone 没有使用定义好的行为,所以,系统适合整队比赛,而 Sohota 的技术主要用在一对一的情况。但 Stone 的层学习有一个重要的问题,由于学习的过程是从低级到高级逐步进行,每步学

习都是在低一层的基础上进行的,因此,当某一层学习产生误差时,误差会通过层之间进行传递,使整个系统出现决策失误。Stone 的论文中,Agent 拥有三种不同的状态空间:worldstate, locker-roomagreement 和 internalstate。Agent 还有两种不同的行为:internal(内部动作)和 external(外部动作)。worldstate 反映了环境的状态,表现为 Agent 的 belief。外部动作改变 worldstate,感知器捕捉 worldstate 的改变。locker-roomagreement 是队伍内部协调的方法。对于球队中的每一个球员,locker-roomagreement 都是一样的。它定义了复杂的队伍结构和内部 Agent 通信协议。internalstate 定义了 Agent 的内部变量,它存储了以前的和当前的状态信息。对于 long-term 的决策动作,internalstate 是维系它们在前后时间段连续的状态标量。内部动作根据 internalstate, worldstate 和 locker-roomagreement 来更新 Agent 的当前 internalstate。外部动作参考 worldstate 和 internalstate,选择执行的单位动作。动作影响真实模型,改变 Agent 未来的感知和预测的 worldstate。

## 2 Multi-Agent 层次协作模型

协作模型的优劣直接影响多智能体系统的协作质量和是否能够完成一定的任务。在目前 RoboCup 仿真比赛中,已提出了多种多智能体协作模型,最典型的是 CMU-nited, FCPortugal 和 Tsinghuaeolus 等的模型,其协作的关键技术都各有特点。为了解决多智能体协作过程中反应性和智能性难以调和矛盾,文中提出一种三层的 Multi-Agent 层次协作模型,它包括全局层、局部层和个体层。通过实验证明,该模型是合理的、有效的。

通常来讲,多智能体的协作模型要包括以下 3 个方面:

- 1) 协作模型,它要求多智能体要在这个模型下进行协作;
- 2) 协作介质,它指出智能体通过何种媒介来进行协作;
- 3) 协作规则,它管理多智能体是如何来进行合作的。

为了使得智能体具有设计简单、反应速度快、可适应性强和智能度高的特点,提出了 Multi-Agent 层次协作模型。这个模型具有三层结构,它包括全局层(Global Layer)、局部层(Local Layer)和个体层(Individual Layer)。全局层主要是负责对手建模(Opponent Module)和确定阵形(Formation);局部层的任务是角色变换(Role)、策略选取(Tactics)和智能体之间的通讯(Communication Protocols);而个体层主要是负责智能体的行为方式。

## 3 全局层设计

全局层管理所有的智能体,它根据对手建模和场上状态来选择合适的阵形。为了提高智能体的可适应性,采用静态阵形和动态阵形相结合的策略。

### 3.1 静态阵形

通过定义智能体的静态位置,使智能体在不同的状态处在不同的位置上。在 RoboCup 中,静态阵形主要是球队的初始阵形,球员可以静态地知道自己所处的位置,如 4-3-3,3-3-4,3-4-3,4-4-2,3-2-5 以及 3-5-2 等。

### 3.2 动态阵形

在 RoboCup 仿真比赛中,环境和信息都是实时变化的,仅定义静态阵形是不够的,必须定义动态阵形,以很好地适应实时动态环境的要求。这里使用一种基于对手建模的动态阵形变换策略<sup>[4]</sup>。假设  $\Delta S$  为己方球队的得分与对方球队的得分之差,  $\Delta M$  为 6000 与当时仿真周期之差,则定义如下:

```
If  $\Delta S \in (1, 3)$  and  $\Delta M \in (3000, 6000)$ ,
Then keep the former formation;
Else switch 4-4-2 formation;
If  $\Delta S \in (4, +\infty)$ 
Then keep the former formation;
If  $\Delta S \in (-2, 0)$  and  $\Delta M \in (3000, 6000)$ ,
Then switch 4-3-3 formation;
Else switch 3-3-4 formation;
If  $\Delta S \in (-\infty, -3)$ 
Then switch 3-2-5 formation
```

很显然,比分和时间是两个影响阵形变化的主要的因素。比分的确定原则:我方落后,则采取主动进攻,转向主攻的阵形,如(3-3-4);我方领先,就保存优势,以防守反攻为主,转向主守的阵形,如(5-4-1)。时间的确定原则:如果时间过半,但是我方还处在输球状态,则要变换阵形;如果时间过半我方处在有利状态下,则不变阵形,或者变为偏重防守的阵形。

## 4 局部层设计

在 RoboCup 中,文中根据阵形将球队的所有智能体分为 3 个单元:进攻单元、防守单元和中间单元。进攻单元的智能体只需进攻,而不需要回到自己的半场上去防守;防守单元中的智能体只需要防守,不能够进入对方半场;中间单元则需要较为复杂的行为,该单元的智能体有时既需要参加进攻有时又要参加防守。

### 4.1 角色

进攻单元有两个角色:积极进攻和消极进攻;防守单元也有两个角色:消极防守和积极防守;中间单元中的两个角色:辅助进攻和辅助防守。很显然,角色主要有两种状态:进攻状态和防守状态。球员处于进攻状态时,如果是控球队员,则根据当前的情况处理球。当在射门范围而且有很好的进球机会时,则立即射门;否则,把球传给有更好机会的队友;如果没有队友接应,则自己带球作转移,一般是向守方队员密度小的方向转移,继续寻找机会。在防守状态中,很重要的一个概念是防守反攻,所以每时每

刻都应该留意球,判断是否能把球抢断下来。如果断球成功,则要分析当前的情况进行分别处理;如果断球失败,则要进行盯人防守策略。比赛时攻防状态的变换主要是由球被哪方所控制来进行的。当球在己方控制时,则进入进攻状态。

\* 进攻状态可由四种情况进入:

- 1) 己方队员开任意球、门球、界外球、角球等。
- 2) 队员在后防区域内截对方的传球或抢断下对方的带球,如果在此区域内则被对方队员再截回去的可能性大,或传球的成功率低,这时应将球踢出界外。
- 3) 队员在中场范围内得球。
- 4) 在对方的后防区域断它们的失误球和回传球。一旦球进入对方控制,立即切换到防守状态。

\* 防守状态可由三种情况进入:

- 1) 队员在对方后场丢球,中间单元的球员和进攻单元的球员进行防守状态,防守单元的球员要进行盯人防守。
- 2) 队员在中场丢球,前锋回防中场,辅助防守负责前去抢球;积极防守队员回防。
- 3) 队员在我方后场丢球,辅助防守队员与积极防守队员全力拼抢。

### 4.2 策略

采用的策略包括角色、方案、对手模型策略、队友模型策略及其通信协议。作为团队策略,它是由一系列的方案、产生方案的规则、角色、对手策略模型、队友策略模型及其通信协议组成,其中方案是由一系列的阵形、产生阵形的规则以及预设计划组成的,阵形则是由智能体在阵形中的位置决定。在 RoboCup 中,当球员做某个动作时,预测该动作产生的结果是非常重要的。因此,需要采取适当的预测策略。文中采用 PLOS(Predictive Locally Optimal Skills)和 IMBBOP(Ideal-Model Based Behavior Outcome Prediction)技术<sup>[5]</sup>。PLOS,也就是“预测性、近距离优化技术”。当球员做某个动作时,它会在时间和空间上分析做动作以后几个周期内外部环境的变化,并且预计未来几个周期可能的动作,根据分析结果对当前要做的动作进行优化处理。同时,根据自身的情况再做一次优化处理,最后得到本仿真周期所要做的最基本动作。当下一个动作周期到来的时候,如果周围环境和上周期所预测的十分接近,那么智能体将保持上个周期的行为模式,并且继续执行上个周期的行为。然而,如果周围环境和预测有本质的不同,所有的情况要重新进行计算,而不是继续执行上一次计划中的指令。IMBBOP 就是基于理想模型的行为输出预测,这个模型对对方的预测就是对方在理想的世界模型下的最优的决策。队员之间之所以能协作是由于有共同的目标,因此每个队员在决策时都要考虑全局利益。

文中提出的局部协作策略是一种基于角色的方法。整体队形为 433,角色有前锋、中场、后卫、守门员,每个队员根据自己角色的不同采取不同的策略,但有一个共同的目标:防守时(即对方持球)阻止对方进球或减少球进入我

方球门的危险区域;进攻时(即我方持球)尽量让球射入对方球门或让球进入对方球门禁区。

下面以防守为例,说明队员协作的具体运用。防守时有 3 个动作:盯人(Mark)、阻截(Block)和跑本位点(Formation)。盯人的目的是防止被盯的对手接到传球;阻截的目的是阻碍对方控球队员带球前进;本位点是每个队员根据自己角色的和当前球的位置得出的一个活动区域的重心,跑本位点是指球员站在本位点处防守,即进行区域防守。每个队员根据自己角色的不同有不同的防守区域,即不同本位点。每个队员对自己的防守区域都有个承诺:保证对方不会带球穿过我方的防守区域,同时它对其他队友也有个信念:队友一定不会让对手带球穿过队友的防守区域。正是每个防守队员基于这样的承诺和信念,球队在面对对方的进攻时能保持良好的队形,不会让对手轻易突破防线。但是,由于队员的能力有限并不能绝对保证能成功拦截。因此,队员在选择防守动作时,应能根据当前的环境信息和其他队员模型判断其成功履行承诺的概率。建立其他队员的模型是在比赛前训练出来的,训练的方法有再励学习、决策树、神经网络等。

5 个体行为层

个体行为层定义智能体单独完成任务的行为,包括抢断、射门、清理球、传球、拦截、盘带、到定点、盯人、面向球等。这些行为都是预设的。

(上接第 179 页)

通过使用 Matlab SVM Toolbox 训练仿真比较,结果如表 1 所示。最终选择  $\epsilon = 0.05, \delta = 1.0, c = \infty$  作为实际测量模型参数,并用 VC++ 编制了相应的测试应用程序。投入运行后,测试误差  $< 0.1\%$ , 满足工艺设计精度  $< 0.3\%$  的要求,取得较好的测试效果。

表 1 训练仿真结果( $c = \infty$ )

序号	$\epsilon$	$\delta$	支持向量数(svs)	最大误差(%)
1	0.004	2.0	12	0.448
2	0.008	2.0	12	0.400
3	0.032	2.0	12	0.248
4	0.05	2.0	12	0.150
5	0.128	2.0	12	0.223
6	0.25	2.0	12	0.470
7	0.05	4.0	12	0.354
8	0.05	1.0	12	0.053
9	0.05	0.5	12	0.232
10	0.05	0.1	12	0.470

3 结束语

主要结论是:

6 实际结果与结论

实验中,笔者在自己的球队 Winlet2005(底层技术采用 UVA2003)的基础上加以改进,注重球员之间的配合,加快攻防转换的节奏,在与单结构如反应结构、慎思结构球队以及不注重协作能力的球队(Winlet2003)的比赛中战果显赫,在 100 场的较量中取得 83 胜、10 平、7 负的成绩。多次实验结果充分说明文中提出的 Multi-Agent 层次协作模型是合理的、有效的。

参考文献:

[1] Omicini A, Papadopoulos G A. Why Coordinate on Models and Language in AI[J]. Applied Artificial Intelligence, 2001, 15 (1): 1-11.

[2] Stone P. Layered Learning in Multi-agent Systems[D]. America: Carnegie Mellon University, 1998.

[3] Reis L P, Lau N, Oliveira E C. Situation Based Strategic Positioning for Coordinating Team of Homogeneous Agents[EB/OL]. <http://www.ieeta.pt/robocup/archive.htm>. 2002-05.

[4] 仵博, 吴敏, 曹卫华. 一种基于行为的 Multi-Agent 决策算法及其在 RoboCup 中的应用[J]. 机器人技术与应用, 2002(2): 46-50.

[5] Stone P, Riley P, Veloso M. Defining and Using Ideal Teammate And Opponent Agent Models. AT&T Labs[EB/OL]. <http://www.research.att.com/pstone>. 2001-05.

(1)在多传感器信息融合过程中,采用支持向量机的学习算法是可行的,可有效解决小样本、非线性、参数之间存在模糊关系的信息融合建模问题;

(2)该模型结构灵活,无需人工经验,可从实测数据集直接建立模型,具有一定的推广性;

(3)核函数的类型及相关参数的选择对融合精度有一定的影响,如何优化,有待做进一步的研究。

参考文献:

[1] 袁布儿, 杨东勇, 林毅. 多传感器信息融合及其在工业控制中的应用[J]. 浙江工业大学学报, 1999, 27(4): 281-286.

[2] 王军, 苏剑波, 席裕庚. 多传感器融合综述[J]. 数据采集与处理, 2004, 19(1): 72-77.

[3] Vapnik V. The nature of statistical learning theory[M]. New York: Springer-Verlag, 1995.

[4] Gunn S. Support Vector machine for classification and regression[Z]. Jsis report, Canada: Image Speech & Intelligent System Group, University of Southampton, 1998.

[5] Vanderbei R J. LOQO: An Interior Point Method for Quadratic Programming[R]. NJ: Princeton University, 1994.