

分层技术在机器人足球中的应用研究

柯立堃,程家兴

(安徽大学 计算机学院,安徽 合肥 230039)

摘 要:多智能体学习是多智能体系统和机器学习等研究领域的交叉。由于多智能体系统的复杂性,研究者们采用基于分层的机器学习的方法来解决。设计了一个基于分层技术的决策子系统,在此基础上,使用机器学习的方法对分层之后的决策学习和技能学习进行了研究实现。最后,指出了当前机器人足球中分层学习技术存在的不足,为以后机器人足球研究又增加了一个新的研究课题。

关键词:分层技术;分层学习;决策系统;机器学习

中图分类号:TP242.6

文献标识码:A

文章编号:1673-629X(2007)03-0073-04

Research on Hierarchical Technology in Robotic Soccer

KE Li-kun, CHENG Jia-xing

(Department of Computer, Anhui University, Hefei 230039, China)

Abstract: MAS learning is a domain intercrossed between MAS and machine learning. Because of the inherent complexity of MAS, there are much interests in using machine learning techniques to handle it. Design a decision-making system based on hierarchical technologies, and research on the hierarchical learning in Robotic soccer using machine learning. Finally, point to the deficiency existing in the hierarchical learning technologies nowadays, which increase a new research task in the research of Robotic soccer.

Key words: hierarchical technologies; hierarchical learning; decision-making system; machine learning

0 引言

仿真机器人足球(RoboCup Simulation)是多主体系统(MAS, Multi-agent Systems)研究的一个标准问题,同时包含了动态、实时、不确定环境中的合作与对抗,已被国际人工智能界公认为人工智能新的挑战问题。对仿真机器人足球系统的研究,目前包括系统组织结构、多智能体结构及协调机制研究、智能体技能学习和对手模型预测等内容,这也是今后一段时期内 RoboCup 仿真比赛的研究方向。

学习是多智能体系统的重要特征之一,多智能体学习可以看作是多智能体系统和机器学习等研究领域的交叉。当任务和环境变得复杂时,完全依靠程序员的手工编程实现 Agent 基本行为的设计和组织工作,变得非常繁重甚至是不可能的。近年来, MAS 的学习成为一个新的研究热点。该研究方向的一个关键是用

学习技术增强 Agent 的智能,也就是 Agent 自己解决问题的能力。MAS 学习的内容包括个人技术、多 Agent 之间的合作技术和对手的合作与对抗模式等。

机器学习就足球机器人动作的选择是机器人足球比赛中的一项关键技术,其选择方法的优劣直接影响比赛结果。文献[1]中介绍了多种传统的机器学习方法^[1]。而分层学习是机器学习的一种,在机器人足球中也取得了广泛应用。笔者研究了分层决策系统和分层学习在仿真机器人足球中的实现情况,分析了在分层学习环境下踢球机器人的动作选择问题。

1 分层决策子系统的设计

决策系统的任务是根据比赛场上的态势作出是进攻还是防守的决定,给自己球队的每个机器人分配角色,并且协调不同机器人之间的合作。比赛成功与否很大程度上决定于策略系统的性能和可靠性。

文中模仿文献[2]中的基于两层结构的决策模型,设计一种基于三层模型的决策子系统的决策模型,如图1所示。这种模型也是一种分层递阶控制的结构,是一种单智能体模式下采用集中控制方法,即用一个全局控制器对机器人进行统一控制^[3]。三层分别是策

收稿日期:2006-07-06

基金项目:国家自然科学基金(60273043);安徽省高校学科拔尖人才基金(05025102);安徽省自然科学基金(050420204)

作者简介:柯立堃(1984-),男,安徽黄山人,硕士研究生,研究方向为仿真机器人足球研究;程家兴,教授,博士生导师,研究领域为智能计算、时空规划、机器人足球等。

略协调层、运动规划层和基本动作层。

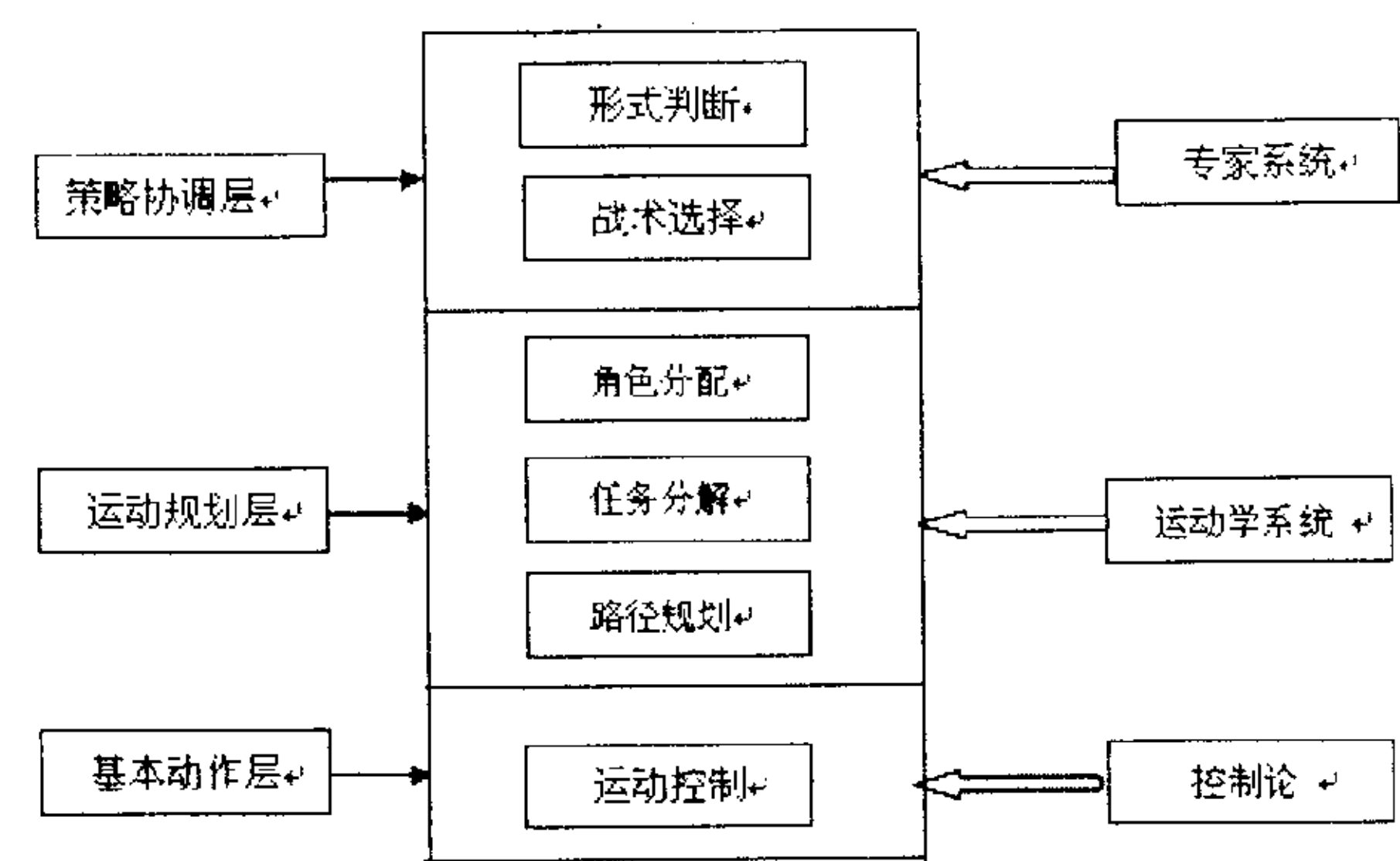


图 1 基于三层模型的决策子系统

下面对这三层结构分别做介绍。

(1)策略协调层是整个决策控制系统的主导思想,由人工智能起控制作用。它作为推理机的规则发生器,处理高层信息,用于机器推理、决策、学习和记忆操作。协调层确定各机器人的目标位置和行为,分析场上的态势,确定我方的基本对策与队形,选择合适的策略战术。该层主要采用专家系统,其推理是基于判断规则的,也就是 IF-THEN 方式。

(2)运动规划层是将协作意图进行详细地分解,将协调层的任务落实到每个机器人,进行角色分配,并给出最佳的任务序列,完成各个机器人运动的目标、轨迹、速度的计算。通过个体的动作和运动方式来体现群体协作的整体意图。该层决定机器人的运动轨迹,其设计要求规划的速度要快,同时应尽可能平滑连贯。

(3)基本动作层是分层结构的底层,是和运动学有关的运动控制层。本层将完成运动规划层产生的阶段性目标和具体的行动指令对应起来,生成一定的运动指令,包括基本动作,如:转动、射门、加速等。本层要求具有较高的精度和较低的智能,它按控制论进行控制,对相关过程执行适当的控制作用。

2 分层学习的应用实现

2.1 分层学习的概念

分层学习是机器学习的一种方法,它定义一系列建立一个杂系统分层次学习方案的准则,这些准则包括:直接从输入到输出的学习难以实现、可以给出从底层到高层的任务分解、机器学习探索训练数据或在不同层次产生适应性学习、一层的学习输出提供给另一层。分层学习法,将问题分解到若干行为层上,并在每一层上使用机器学习技术。学习到的每一层次(子任务)会直接影响下一层的学习。每一层上机器学习方法的选择取决于子任务。利用层次式的任务分解技术,多个机器学习模块可以组合产生更有效的行为^[4]。

这种分层学习的思想最先是由 Stone 提出的,并在各支球队中取得不错的效果。把分层学习的思想应用于机器人足球比赛中,使智能体首先学会带球、趟球、传球和射门等等这些基本动作,在学会的基础上,才让智能体学会什么时候去应用这些动作。

2.2 分层学习策略在球队中的应用

机器人足球是多 Agent 马尔可夫决策问题。机器人足球学习的任务是利用机器学习寻找到最优策略,可分解在两个层次上进行:技能层和策略层。两层都是采用分层学习和其他机器学习方法结合的途径得以实现。采用分层学习,使智能体在新的外界环境下学会采用这些基本技能,然后再采用高层的动作选取策略进行比赛。

技能层中所有的基本技能包括基本行为和基本动作。足球机器人的基本行为包括直行和旋转。基本动作是以基本行为作为基础,它主要包括射门、扫球、拦截、带球、传球、踢球、跑位、避障八个动作。这些基本技能都是用机器学习学习到的,每个技能行动持续一定周期并有明确定义的前提和目标。对技能行动的学习是一个动态优化问题,任务是找到对任何初始状态都具有最小累积代价的动作。

策略层的最终任务是赢得比赛,必须对进攻和防守这些团队行为进行学习。策略层的学习原则与基本技能层基本相同,即角色分工,阵型变化、战术库的更新,穿越穿球的实现等。此层的学习大多采用强化学习的方法来实现。强化学习过程中,若进球奖赏值为正,否则奖赏值为负。通过对导致进球或失球的传球轨迹进行学习,能够找到有助于进球的最短的传球路径。得到球的队员使用这一算法传球,没有得到球的队员对跑位进行学习。

这种层次式的任务分解,采用了分层学习同多种机器学习技术相结合的学习方式。这样可以发挥各种机器学习技术本身特有的优势,提高整体学习的效率和效果。我们将球员 Agent 的学习分为四个层次,如表 1 所示。

表 1 分层的学习系统

基本技能层	射门、转身等	神经网络,决策树
复杂技能层	传球、截球、带球等	基于神经网络的强化学习
合作策略层	战术选择、运作选择等	神经网络+产生式系统
对抗策略层	根据对手变换策略等	Q 学习+基于示例的学习

●基本技能层。

底层学习用来训练简单的技能,如射门、转身等。

神经网络能由连续的输入得到连续的输出,利用 BP 神经网络训练队员进行射门的学习。因为 BP 算法理论依据坚实,物理概念清晰,其误差反向传播;其次,

输入向量的维数不是很高,所以计算简捷,也比较简单。用BP神经网络进行射门学习的球队在比赛中能快速及时地进行射门。

对于转身,则采取决策树进行学习,因为决策树的方法可以有效地处理遗漏属性并忽略无关属性。一棵决策树可以看作一个对目标分类的划分和获取策略。在决策树算法中最常使用的是由Quinlan提出的ID3算法。用此算法可以很好地解决机器人Agent转身存在的问题。

由于底层涉及的都是一些静态的基本技能,一般不会随着比赛环境的变化而改变,因此都采用了离线学习的方式。

●复杂技能层。

这一层则对复杂的技能,尤其是涉及多个球员Agent之间进行合作的技能进行学习,例如截球、传球以及带球过人等。这些技能可以采用基于神经网络的强化学习方法学习得到。

强化学习是足球机器人中应用最为广泛的一种学习方法,它具有顺序和快速决策、延迟奖赏、处理噪音和随机能力、多智能体学习以及适应庞大的状态空间等特点。其目标就是发现最优策略,而最优策略是从输入状态到行动的映射使得回报的总和最大化。Q-learning^[5]是一种最常用的非模型的强化学习方法,它通过对动作状态对到期望回报的映射的动作——评价函数Q来解决非完全信息或马尔可夫问题。通过Q学习,机器人可学习到射门(Shoot)、传球(Push)、断球(Intercept)、守门(Goal)等各项基本技能。

神经网络作为一种监督学习方法,在机器人足球中不单应用于射门的学习,还可应用于对踢球、控球、截球等单一技能的学习。好多球队中就使用了BP神经网络来让球员学习截球的技能,还可以通过BP神经网络学习射门和截球的动作,同时利用BP神经网络^[6]研究了存在对手的情况下球员之间进行传球等协作行为。

在所采用的强化学习中仍然使用了函数估计和强化学习相结合的技术,以处理连续状态空间的问题。除了截球技巧需要队员在线学习外,这类学习过程大多在离线状态下进行。

●合作策略层。

一旦足球机器人队员学会了足球技巧和动作之后,它还需要进一步学习如何根据赛场上的状况进行决策,即学习如何选择下一个合适的动作。在前两层里,假设Agent在固定的队形里,并且不交换位置。但是这个队形在与一个特定对手比赛时并不一定是最好的。笔者认为团队可以有多种不同的比赛队形,并且

每个队员也可以灵活地交换位置,这样团队的整体性能会更好。在这里,采用把产生式系统和神经网络方法有机结合的办法来实现。队员在比赛中在线调整这一层的合作策略。

在球队代码中,采用产生式系统和BP算法相结合的办法进行动作选择和战术选择。

IF 球场信息满足产生式系统的某条规则

THEN {

利用产生式系统进行动作选择,

输出相应的动作编号

}

ELSE {

把球场信息当作神经网络的输入变量,

利用BP神经网络算法进行动作选择,

输出相应的动作编号

}

在情况比较简单时,产生式系统是最直接、最简单也是比较实用的方法,然而,对于复杂情况难以进行准确的判断。BP网络在本质上是一种输入到输出的映射,它能够学习大量的输入与输出之间的精确数学表达式,只要用已知的样本对BP网络加以训练,网络就具有输入输出对之间映射能力。因此,对于比赛场上比较复杂的情况下,BP网络的性能优于产生式系统,但有时在情况比较简单时,性能不如产生式系统好。两者结合能互相弥补各自的缺陷,更及时、更合理地做出决策。

●对抗策略层。

在现实中,球队会面对各种不同的对手。通过训练球队与许多不同的对手比赛。就可以学到若干不同的策略。这一层采用基于示例的学习进行离线学习。它能够迅速将球队当前的对手进行建模,并与过去曾比赛过的最相似的对手匹配,从而采用以前成功使用过的策略。在这个策略层,战术的自动生成、角色的自动发现以及对于对手战术的自动分析都得通过学习得到。对这些方面的研究还很欠缺。

我们采用Q学习和基于示例的学习(Case Based Learning, CBL)结合学习的方法^[7]。基于示例的学习是一种增量式的机器学习方法,该方法适合较难发现规律性的知识且因果关系难以用精确模型表示的领域。案例库中包含两种类型的案例:规则案例和自学习案例。在这种结合的学习模型中,足球机器人在初始的行为学习时,案例库中只存在规则案例,机器人通过Q学习选择符合规则案例约束的动作执行,并将学习结果产生新案例保存。随着学习的进行,如果案例库中存在与当前状态匹配的案例,则机器人选择该案

例指定的动作执行,否则自动生成新案例保存。

2.3 当前分层技术存在的问题

智能体从个体行为到团队行为的四个层次进行学习,为其它团队行为和对抗行为建立新的学习层次也是一个有意义的研究课题。目前的分层学习存在着一个很重要的问题,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。这些问题留待以后的研究中解决。

对于单个智能体,分层学习相对比较简单。整个学习过程分三步进行:首先,训练智能体学会基本技能;然后训练智能体学会在一定的外部环境下选用合适的动作;最后,在模拟真实比赛的环境中,训练智能体在不断改变的外部环境中选取合适的动作。这种学习方法非常有效地解决了单个智能体在未知环境下的学习问题^[5]。相对单个智能体,多智能体在分层学习时存在角色的分工、多智能体之间的协调等一系列问题,实现起来困难得多。虽然多智能体的分层学习在当前的机器人足球中已经有了相当多的应用,但还没有得到很好的实现。

3 结束语

机器人足球领域,具有复杂、动态、实时、不确定性的特点,是机器学习应用的极佳领域。机器人足球赛中的各类机器学习方法已经引起了研究者的广泛关注。鉴于分层技术的重要性,目前各支球队都在积极

地开展对 RoboCup 中机器学习方法,建立了不同的基于分层结构的决策系统,成立了分层学习和多种机器学习技术相结合的机器学习系统,并将这一研究的初步结果应用到各自研制的仿真球队中,取得了不错的效果。

文中结合分层学习和机器学习的技术,设计了一个基于三层模型的决策子系统,使用机器学习的方法对分层后的决策和技能进行了研究与实现,并指出了当前机器人足球多智能体学习中存在的不足,为以后机器人足球研究又增加了一个新的研究课题。

参考文献:

- [1] 童亮,陆际联. 仿真机器人足球学习方法研究综述[J]. 计算机仿真学报,2004,21(6):1-5.
- [2] 赵逢达,孔令富,李贤善. 基于分层结构模型的机器人足球决策系统设计[J]. 哈尔滨工业大学学报,2005,37(7):933-935.
- [3] 杨佩陈,兆乾,陈世福. 机器学习在 RoboCup 中的应用研究[J]. 计算机科学学报,2003,30(6):118-121.
- [4] 李红娜,姚分喜,黄鸿. 分层增强学习在足球机器人比赛中的应用[J]. 计算机仿真学报,2005,22(6):145-147.
- [5] 周浦城,洪炳镨,郭耸. 基于强化学习的多机器人协作[J]. 计算机工程与应用,2005(28):10-13.
- [6] 洪炳镨,薄喜柱,韩学东. 基于人工神经网络的足球机器人分层学习研究[J]. 哈尔滨工业大学计算机工程与应用报,2004,36(7):859-861.
- [7] 柳在鑫,王进戈,王强. 基于 Q 学习与 CBL 结合的机器人足球行为学习研究[J]. 西华大学学报:自然科学版,2005,24(4):58-60.
- [8] 王强,王进戈,柳在鑫. 基于 Q 学习与 CBL 结合的机器人足球行为学习研究[J]. 甘肃科技,2004,20(10):83-85.
- [9] Witlox F, Tindemans H. The Application of Rough Sets Analysis in Activity-based Modeling: Opportunities and Constraints[J]. Expert Systems with Applications, 2004 (27): 585-592.
- [10] Tsumoto S. Mining Diagnostic Rules from Clinical Databases Using Rough Sets and Medical Diagnostic Model[J]. Information Sciences, 2004(162):65-80.
- [11] Li RenPu. Mining Classification Rules Using Rough Sets and Neural Networks[J]. European Journal of Operational Research, 2004(157):439-448.
- [12] Supriya K D, Krishna P R. Clustering Web Transactions Using Rough Approximation[J]. Fuzzy Sets and Systems, 2004 (148):131-138.
- [13] 胡可云,陆玉昌,石纯一. 粗糙集理论及其应用进展[J]. 清华大学学报:自然科学版,2001,41(1):64-68.
- [14] 蒋良孝,蔡之华,刘钊. 一种基于粗糙集的决策规则挖掘算法[J]. 微型机与应用,2004(3):7-8.
- [15] Li Xiangshen, Han Tongloh. Applying Rough Sets to Market Timing Decisions[J]. Decision Support Systems, 2004 (37): 583-597.
- [16] 肖智,张志恒,黄海生. 粗糙集理论在企业财务危机预测中的应用[J]. 决策参考,2004(3):48-53.
- [17] 吴国芳. 粗糙集理论在房产定价决策中的应用[J]. 技术经济,2004(6):11-13.
- [18] Chun Chehuang, Tzu Liangtseng. Rough Set Approach to Case-based Reasoning Application[J]. Expert Systems with Applications, 2004(26):369-385.
- [19] 刘云忠,宣慧玉,林国玺. 粗糙集理论在我国税收预测中的应用[J]. 系统工程理论与实践,2004(10):98-103.
- [20] 黄沛,李剑. 基于粗糙集的保险风险规则挖掘模型[J]. 系统工程,2002,20(5):34-39.
- [21] 黄沛,李剑. 基于粗糙集理论的续保规则挖掘模型[J]. 上海交通大学学报,2004,38(4):641-645.
- [22] 刘嘉,梁勇勇. 基于粗糙集理论的客户分类规则挖掘模

(上接第 72 页)