

基于领域知识的多智能体分层协作规划方法

张晓勇, 吴敏, 彭军

ZHANG Xiao-yong, WU Min, PENG Jun

中南大学 信息科学与工程学院, 长沙 410083

School of Information Science and Engineering, Central South University, Changsha 410083, China

E-mail: zxyong@gmail.com

ZHANG Xiao-yong, WU Min, PENG Jun. Multi-agent layered cooperation planning method with domain knowledge. *Computer Engineering and Applications*, 2009, 45(3): 7-10.

Abstract: To solve the cooperation problem in RoboCup, a multi-agent layered cooperation planning method with domain knowledge is proposed. Agent's initial behavior plan can be established according to the information about environment. The plan can be adjusted by considering the agent's domain knowledge and analysis of effect of behavior to obtain a consistent behavior plan sequence that enable agent to make a high-speed decision and realize a MAS cooperation for limited communication.

Key words: Multi-Agent System(MAS); cooperation planning; domain knowledge; RoboCup

摘要: 针对复杂动态 RoboCup 仿真环境下的多智能体协作问题, 提出一种使用领域知识的多智能体分层协作规划方法。在根据系统提供的环境状态信息建立智能体初始行为规划的基础上, 通过逐步填入智能体的状态和行为特征属性知识, 分析行动实现的效果, 对初始规划行动进行调整, 为智能体制定协调一致的行为规划序列, 并获得相应的协作策略信息, 使其在实际对抗比赛中能针对当前环境快速做出决策, 实现通信受限的情况下多智能体的协作。

关键词: 多智能体系统; 协作规划; 领域知识; 机器人世界杯

DOI: 10.3778/j.issn.1002-8331.2009.03.002 **文章编号:** 1002-8331(2009)03-0007-04 **文献标识码:** A **中图分类号:** TP301

1 引言

多智能体系统(Multi-Agent Systems, MAS)是由多个相互作用的自治智能体形成的集合, 其研究重点在于智能体与智能体之间的交互, 即智能体为了联合采取行动或求解问题, 如何协调具有不同目标的智能体的知识、意图、规划和行为, 以对其资源进行合理安排, 最大程度地实现系统总体的目标^[1]。RoboCup(The Robot Soccer World Cup)机器人足球仿真系统是一个典型的多智能体系统, 它提供了一个实时异步、有噪声的多智能体环境, 用来模拟实际足球队员的比赛^[2]。为了赢得比赛, 不仅需要单个智能体具有灵巧快速的反应能力, 更需要为多个智能体制定一套灵活的协作方法来进行有效的球队协作, 才能完成球队的比赛任务——射门得分。

由于各 RoboCup 仿真参赛队伍的底层工作已经相对成熟, 目前对多智能体的研究主要集中在高层策略领域, 制定并实施合理的高层协作策略是仿真球队致胜的关键。而规划是一种重要的问题求解技术, 它从某个特定的问题状态出发, 寻求一系列行为动作并建立一个操作序列, 直到求得目标状态为止。在多智能体协作策略问题中, 可以通过对经常出现的状态、

重要的状态或典型的状态提前做出行为规划, 则在实时动态环境中球员就能按照预先规划快速做出行为决策, 来实现更强的配合和对抗。

然而传统的规划方法一般是基于单个智能体的, 当环境中还有其他智能体时, 智能体只是简单地将它们包含在它的环境模型中, 而不改变其基本规划过程, 降低了规划的性能, 无法满足 MAS 的需求^[3]。而在复杂、动态、不确定的多智能体环境中, 由于受限的通信带宽和环境噪声的影响, 每个智能体并不能准确地感知环境, 随时可能出现感知信息错误、观察失败等问题, 根据环境信息做出的行为决策会影响多智能体的协作效率。这就要求多智能体间具有灵活的团队协作规划能力, 使每个智能体能够知道在当前状态下对于最优的联合规划它应采取的动作。

Cox 等人提出的多智能体规划协作算法^[4], 采用规划空间搜索每个智能体规划状态, 智能体之间共享各自每个规划步, 在搜索过程中不断修正智能体的每一步规划, 动态减少了搜索空间的大小, 但该算法会导致智能体搜索状态呈指数增加, 进而影响到搜索效率。Dimopoulos 等人将经典规划扩展到多智能体环境中, 提出了一种分支界限算法^[5], 利用该算法组织协调各

基金项目: 国家杰出青年科学基金(the National Natural Science Foundation of China under Grant No.60425310)。

作者简介: 张晓勇(1980-), 男, 博士生, 主要研究领域为计算机应用、多智能体系统; 吴敏(1963-), 男, 博士, 博士生导师, 主要研究领域为智能系统与机器人技术、网络系统与企业管理、先进控制与自动化技术; 彭军(1967-), 女, 博士, 博士生导师, 主要研究领域为多智能体系统、计算机应用。

收稿日期: 2008-08-21 **修回日期:** 2008-10-31

个智能体的规划,在智能体之间形成一致、无冲突的规划,使具有不同能力的多个智能体互相协助,实现各自目标任务,这种方法仅适用于多智能体协作的一种特殊情况——协助,而在完全合作的 MAS 中存在局限性。

主要研究 RoboCup 中多智能体协作规划问题,通过在求解问题之前暂时删去某些细节,把复杂的问题分解为一些比较小的子问题,在问题求解过程中再逐步填入根据观察和识别所获得的领域知识,逐层细化智能体的内部状态和行为属性值,最终获得各智能体合理的动作规划序列,确保智能体针对当前形势快速行动的同时,实现智能体间有效的协作。通过在 RoboCup 仿真比赛中的应用实践,验证该协作规划方法具有一定的推理规划能力,不仅能够提高单个智能体的反应速度,还可以促进多智能体的协作。

2 使用领域知识的多智能体协作规划方法

多智能体的协作规划主要针对多智能体行为的管理,以实现智能体策略的选取和行为的输出。RoboCup 机器人足球仿真比赛是一个通信受限的多智能体环境,只提供了一个单通道、低带宽的通信机制,在一定的仿真周期内,球员间只能传递有限的消息,而且信息还存在延时。即使收到了通信信息,也是在前段时间的信息,可能已不适用当前的环境。所以,需要寻求一种适用于 RoboCup 机器人足球仿真环境、不依赖于通信的多智能体协作规划方法。领域知识^[9]是指一个专门领域内重要的问题或概念以及这些问题和概念之间的相关关系。将领域知识应用于多智能体系统的协作规划中,采用预先已知的系统领域知识对不同情况下智能体的行动策略进行提前规划,使智能体针对当前环境快速进行策略选取,促进多智能体协作的实现。

该协作规划方法的结构如图 1 所示,可分成两个阶段:(1)利用文献[7]中状态预测算法构建多智能体的初始行为规划轮廓;(2)结合智能体状态和动作选择的领域知识进行协作策略的规划,主要包括协作主体的确定和协作方案的选择,最终实现智能体动作序列的输出。

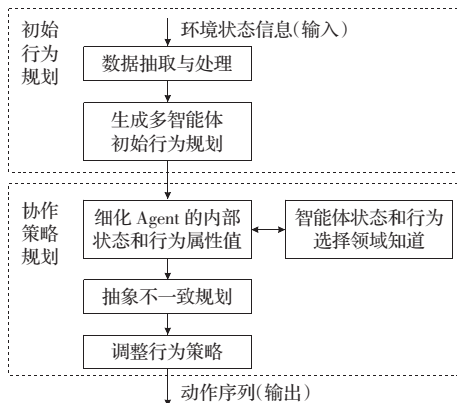


图 1 使用领域知识的多智能体协作规划过程

在构建多智能体初始行为规划轮廓时,智能体通过与动态环境的交互,获得与自身相关的内部环境状态信息,充分考虑其它智能体对环境的影响,采用重要度函数和信息处理方法,对协作所需信息进行筛选和处理,提取出有用的信息,根据状态预测算法对智能体的行为进行预测,建立多智能体初始行为规划序列。

然后,利用智能体所感知到的状态信息和行为规划的领域

知识,逐步细化初始行为规划中智能体的内部状态和动作属性值,分析在当前环境下智能体执行某个动作的条件是否满足,抽象出不一致的规划,将分析结果转换成规则,根据规则所提供的信息确定协作主体及其选取的行为策略,对各智能体的行为过程进行优化,获得智能体正确地实现任务的动作规划序列,从而实现有效的协作。

3 协作规划方法在 RoboCup 中的实现

第 2 章只给出了一个通用的多智能体协作规划框架,本章将其应用到 RoboCup 机器人足球仿真球员的行为规划中,根据各球员具有的状态和行为规划领域知识对其行为过程进行规划,在实际对抗比赛中球员就可根据预先定义好的行为策略对当前环境快速做出行为决策,在行动上达成共识,从而实现一种不依赖于通信的显式的多智能体协作。

3.1 RoboCup 中领域知识的层次化描述

在使用领域知识的协作规划方法中,在对智能体进行行为规划前,有必要先 RoboCup 领域中智能体的状态和行为知识进行处理。本文采用层次化方法对这些知识进行归纳概括,从中提取出相关的特征属性,建立这些属性的有序层次集合,形成一种更清晰的知识关系,为后续逐层调整智能体的初始行为规划提供更有利的信息和依据。

定义 1 智能体的状态特征集 F 。

RoboCup 仿真比赛环境是一个复杂、动态的多智能体环境,有许多种可能的状态,而且每一个状态都有它们相应的特征属性,使它们区别于其它的状态,即 $F=\{feature1, feature2, \dots\}$ 。将这些特征属性分成以下 3 大类,来对球员状态知识进行描述:

(1)空间位置类 $S \in F$

该类用来描述球或球员智能体位于比赛场地的哪个特定空间区域。设 S 为整个球场所有区域的集合:

$S=\{\text{后场防守区 } 1, \text{ 中场配合区 } 2(2a \text{ 左中场}, 2b \text{ 右中场}), \text{ 前场射门区 } 3(3a, 3b, \dots), \text{ 禁区 } 4, \text{ 球门区 } 5, \dots\}$

球员在不同区域活动时有不同的动作选择。

(2)速度类 $V \in F$

对球员和球的运动速度特征定义如下:

$V=\{\text{当前智能体是静止的 still、在移动 moving}\{\text{速度是快速的 fast、中速的 medium、慢速的 slow}\}, \dots\}$

(3)与球相关的方向类 D

该类用来描述球员和球之间的相对位置与相对速度的关系,定义为:

$D=\{\text{球员控球 has-ball, 球员位于球的左边 left, 球员朝球方向运动 Incoming, 球正远离球员 Moving-away}, \dots\}$ 。

上述只是对单个状态的描述,也可以对状态进行组合,把单个的状态叫做基本状态,几个状态的集合叫做组合状态,如组合状态 $\{\text{Incoming, fast}\}$ 就表示球快速朝球员方向运动,可简称为 Incoming-fast。在 RoboCup 这样一个多智能体复杂动态的环境中,一般使用组合状态来对智能体当前状态进行描述。

定义 2 智能体角色特征集 R 。

在 RoboCup 机器人足球比赛中,每个球员都充当一定的角色,角色不同所承担的责任也就不同。球员在做出行为决策时,可以根据自己所扮演的角色进行行为推理及规划,排除不属于自身角色的行为动作,从而减少自身所选取的行为和其他球员产生冲突的可能性。可见,球员所充当的角色与其行为规

划息息相关,体现了球员间较高层次上的协作。因此,引入角色概念,将其作为球员行为规划时需考虑的一个重要的属性。所有球员可以充当的角色用集合 R 来描述,定义如下:

$$R = \{\text{守门员 } r_1, \text{前锋球员 } r_2 \{ \text{左前锋 } r_{21}, \text{中锋 } r_{22}, \text{右前锋 } r_{23} \}, \text{中场球员 } r_3 \{ r_{31}, r_{32}, \dots \}, \dots\}$$

由于球场形势是实时、动态、连续的,球员不可能自始至终都属于一个固定的角色。定义一个角色值函数 $\mu(r_i) \in [0, 1]$, 来描述场上球员属于某一角色 r_i 的程度。如在 RoboCup 中某一个球员属于前卫的角色值为 0.3, 属于前锋的角色值为 0.7, 那么在当前形势下,该球员扮演的角色就为前锋,主要任务为选择合适的时机射门。

定义 3 智能体可选的行为动作集 A 。

在 RoboCup 中为了达到预期的目标,智能体必需对相应的状态做出反映,采取并执行合适的行为动作,使环境进入下一个更接近目标的状态。智能体的行为动作可以是单个的,也可以是一系列的动作序列;既可以是底层的,也可以是高层的。也就是说,它既可以是移动到某一个位置、将球踢到某一个点等这样一些简单的动作;也可以是先找球,再截球,然后带球,最后再传球的动作序列;既可以是 kick、turn、dash 这些底层的动作,也可以是传球、带球、射门、盯人这些高层配合策略。这些单个的或一系列的、底层的或高层的行为构成了球员智能体对于相应状态采取的行为动作集 A 。

3.2 协作规划策略的确定

在 RoboCup 中对多个智能体进行行为规划时,并不需要对系统总体目标做出完全的规划,而只要生成近期的动作序列就可以了,这是因为仿真环境是不断变化的,很多情况无法预料,长期的规划很可能因为情况的变化而失去意义。因此,利用 3.1 节描述的 RoboCup 中层次化领域知识仅对智能体短期的行为进行规划。

在球员智能体根据系统提供的环境状态信息建立初始行为规划的基础上,逐步填入与行为规划相关的状态和行为特征属性知识,对初始规划行动进行分解,分析归纳行动实现的情况,生成协作策略的值规则,最终确定智能体每一步如何行动。

对于简单问题,根据角色就可以确定智能体的行为规划,而对于 RoboCup 中多智能体协作这样复杂的问题,需要将角色特征和状态特征相结合,组成一个属性对 $r_i.feature$, 来表示智能体在特定角色下的状态特征变量,作为规划的前提条件,用来控制整个规划过程的进行。当智能体属于某个角色且其状态变量的逻辑值为真时,智能体才能执行某个动作,使环境进入下一个更接近目标的状态,依次类推,引起一系列动作序列的执行,直到达到目标状态,则整个规划过程结束,从而得到智能体行为规划的值规则。这里给出一个值规则的示例如下:

$$\text{If } (r_1.feature1 \wedge r_2.feature2) \Rightarrow X$$

其含义为:当智能体属于角色 r_1 的状态特征 $feature1$ 且属于角色 r_2 的状态特征 $feature2$ 都为真时,该智能体可以成功执行动作规划 X 。

在实际 RoboCup 仿真比赛中就可以根据上述生成的值规则所包含的策略信息,对智能体进行行为规划,采取并执行合适的动作,协作完成目标任务。

在一场足球比赛中,有效的进攻是球队取胜的关键,很难想象一个球员从后场开始带球,长途跋涉最终实现射门成功。

RoboCup 仿真环境与真实环境一样,球队必须通过有组织的进攻,才能实现射门得分这个团队总的目标任务。下面通过我方从左边路进攻协作射门的实例,来说明应用使用领域知识的协作规划方法对球员协作行为进行规划的过程。

当前我方边锋 11 号队员控球,从边路带球到对方底线后,根据提出的协作规划方法对它和中锋 9 号队员下一步的行动计划进行规划,以实现下底传中射门。整个协作规划的过程如图 2 所示。

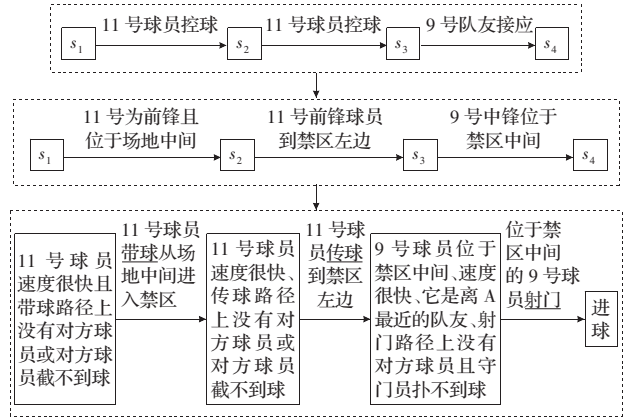


图 2 前锋球员 11 号、9 号协作射门的规划

图 2 中协作规划可分成 3 个阶段:首先,根据获得的状态信息,对当前场上形势进行总体把握;其次,考虑球员的角色和当前在场上的位置这两个特征属性,排除不属于自身角色的行为动作;最后,考虑球员的运球速度、行动时可能选择的路径、对手位置和速度等特征属性值,确定智能体采取的行动,并指定行动的路径,这样就得到当前环境下 11 号和 9 号队员的行为规划序列:11 号队员带球→11 号队员传球→9 号队员射门,前一个动作的正确实现是后一个动作实现的前提,从而实现了球员间进攻策略的规划。图 3 给出了实际比赛中行为规划的示意图。



图 3 局部协作射门的规划过程示意图
(黑色为我方球员;灰色为对方球员;●为球)

上述整个协作规划的步骤具体描述如下:

- (1)我方 11 号控球队员,它要进行行为决策,必须获取与其相关的对方球员和队友的策略信息:对方中前卫前一刻控球,可知其处于逐渐远离球的状态,位于它后面;对方左后卫离球很近,虽然它与球移动方向相同,但相对球来说,它是静止的,因为它的移动仅仅是在防守,并没有直接对球进行拦截;对方右后卫位于球后一段距离处,无法成功拦截球。在以上 3 个主要前提下可规划出 11 号队员从左边路进攻的第一步策略——带球;
- (2)带球经过一段时间后,对方右后卫离球很近,此时 11 号队员被其紧逼,因此,它应该把球传给周围没有敌人的中锋

9号队友,即第二步策略——执行传球动作;

(3)当9号队友接到球以后,球的速度很快,其他队员,如我方左右边锋、对方各个后卫队员与球之间都有一段距离,在这个形势下9号队员最佳策略就是去寻找一条合适的射门路径,然后成功破门得分,至此整个协作规划过程结束。

协作规划结束后,可得到智能体进行行为选择的值规则,表1给出了部分值规则,在实际比赛中当球员智能体所处的环境状态与上述场景相类似时,就可以使智能体直接根据它们所包含的策略信息快速做出决策,实现射门得分。

表1 边路进攻中下底传中射门的策略信息

If(对方中前卫.Moving-away-slow \wedge 对方左后卫.still \wedge 对方右后卫.be-hind \wedge ...)	→我方边锋11号.带球
If(对方右后卫.Immediate \wedge ...)	→我方边锋11号.传球
If(球.fast \wedge 我方中锋.Moving-away \wedge 我方右前锋.Short-distance \wedge ...)	→我方中锋9号.射门
If(对方位于右方的中后卫.Left \wedge 对方位于右方的中后卫.Moving-away-fast \wedge 对方右后卫.Long-distance \wedge ...)	→我方中锋9号.射门成功

4 仿真比赛结果

为了验证提出的使用领域知识的协作规划方法的合理性和有效性,将其应用于CSU_YunLu仿真球队中,从历年参加RoboCup仿真比赛的球队中挑选了一些不同风格球队作为对抗球队,根据比赛情况来评估该协作方法在动态、不确定环境中的表现及整体性能。表2给出了部分比赛结果。

表2 采用协作规划方法后的比赛情况

对抗球队(OPP)	比分(OPP:CSU_YunLu)	我方平均控球率/(%)
中科大 2005	6:1	39.3
UvA_Base2005	1:5	67.5
上海大学 2006	0:12	80.6
MRB2004	0:8	77.2
厦门大学 2004	4:2	52.4
FCPortugal2005	7:0	36.8

由表2可知,CSU_YunLu仿真球队除了与中科大、FCPortugal这两支国内外强队的对抗中得到的比赛结果不是很理想外,其余比赛结果还是比较好的,在整个比赛过程中球队具有相对较高的控球率,从整体结果均衡来看,仍达到了预期目标。

为了更客观地反映所提出的协作规划方法的效果,下面通过两个具体例子来分析说明其有效性。

例1 积极进攻时小范围内的协作射门。

处于对方前场的我方3个前锋9号、10号、11号球员智能体通过局部配合最终目标固定设置为将球射入对方球门,通过观测我方三个球员生成的行动计划过程以及执行情况来评测该协作策略的中长期规划以及动态执行能力。当球被对抗球队的球员截下或是完成目标时强制比赛结束,共进行了100次比赛训练。

比赛结束时将球被拦截前球距离对方球门中点的距离以日志形式输出到文件中,来作为对球员完成目标情况的评估。经分析发现,在绝大多数情况下,我方3个球员可以根据行动计划将球推进到距离球门27m以内的位置去,即可以射门的距离,可见本文提出的协作策略能够产生出正确的行为规划以支持其完成目标,并且能够在一定程度上对执行期间所遇到非预期的情况进行计划调整,具备相当的推理规划能力。

例2 积极进攻时团队协作射门

团队协作射门是大范围内的配合,需要大量的中场和前锋球员的共同配合才能完成这个团队决策,如图4所示。



图4 团队协作射门的过程

将采用了本文提出的协作规划方法的CSU_YunLu球队分别与中国科技大学和UvA_Base各进行15场比赛,通过这个实验来测试该协作策略的采用对整个比赛结果的影响,图5、图6给出了改进前后CSU球队和不同对手比赛的射门成功率对比曲线图,从图中曲线数据的对比可以看出采用了新的协作策略的CSU球队射门成功率有了明显提高。

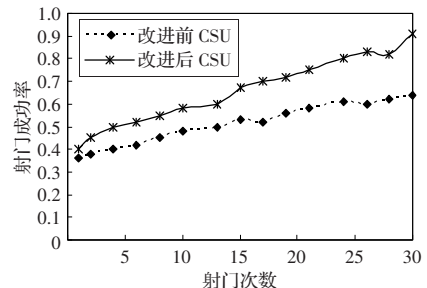


图5 改进前后的CSU与中科大比赛的射门成功率对比

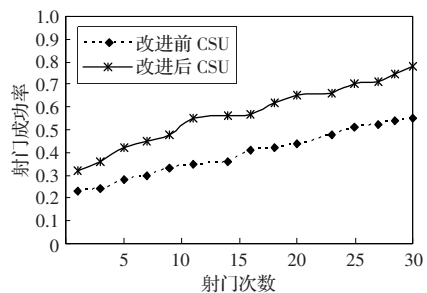


图6 改进前后的CSU与UvA_Base比赛的射门成功率对比

5 结论

针对复杂动态RoboCup仿真环境下多智能体协作问题,提出了使用领域知识的多智能体协作规划方法。智能体根据系统提供的环境状态信息首先建立初始行为规划,再在协作问题求解过程中逐步考虑与智能体行为规划相关的状态和行为特征属性领域知识,不断分析其行动实现的情况,最终获得智能体有序的动作规划,并提取规划中协作策略信息形成值规则,使智能体在实际对抗比赛中能根据这些规则快速进行行为选择,实现了多智能体在通信受限的情况下有效的协作。将该协作规划方法应用到RoboCup仿真球队进攻协作中,其提前规划特性满足了在进攻中球员对快速决策的需求,取得了较满意的协作效果。