

# 基于态势评估的足球机器人决策模型

娄云峰<sup>1</sup>, 陈斌<sup>2</sup>

(1. 西北工业大学 计算机学院, 陕西 西安 710072;  
2. 西北工业大学 电子信息学院, 陕西 西安 710072)

**摘要:** 为了提高足球机器人中策略执行的有效性和针对性, 提出了一种基于赛场正向态势评估和反向执行效果评估的有限状态自动机分层决策模型。该模型将赛场决策分为战术策略和具体任务, 通过正向态势评估信息和任务执行效果的反馈信息进行综合量化, 基于有限状态自动机建立策略和任务的转换关系, 从而做出适合当前态势的战术策略。经仿真实验表明, 该决策模型在很大程度上提高了策略决策的有效性和针对性, 使得球队的整体攻防能力得到了相应的改善。

**关键词:** 足球机器人; 决策; 态势评估; 分层有限状态自动机

中图分类号: TP39

文献标识码: A

文章编号: 1673-629X(2013)09-0099-04

doi: 10.3969/j.issn.1673-629X.2013.09.025

## Decision Making Model Based on State Estimation in Robot Soccer

LOU Yun-feng<sup>1</sup>, CHEN Bin<sup>2</sup>

(1. College of Computer, Northwestern Polytechnical University, Xi'an 710072, China;  
2. College of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China)

**Abstract:** In order to improve validity and pertinence of strategies in robot soccer, put forward a hierarchical finite state automaton decision model based on match state estimations and performing effects. This model synthesizes and quantitates the information into tactical strategy and specific tasks, through state estimations and feedback of performing effects, can establish tactics conversion relationship based on finite state automaton, accordingly make a tactic suiting current states. The simulations indicate that the decision model improves validity and pertinence of tactics to a great extent, enabling that whole abilities of attacking and defending for the soccer team attain corresponding melioration.

**Key words:** robot soccer; decision-making; state estimation; hierarchical FSM

## 0 引言

足球机器人是人工智能的一项丰硕成果, 其理想的目标是战胜人类足球世界冠军<sup>[1]</sup>。随着机器人技术的发展, 对机器人的底层控制已经较为成熟和精准, 而上层决策已成为机器人足球的关键技术, 仿真比赛中尤为明显。因此, 在今后的发展过程中, 能否将人类足球中的先进战术、策略理念灌输给机器人及其决策“大脑”将是比赛胜负的关键, 也是其实现理想的必经之路。

在现代的人类足球比赛中, 强队之间往往同时准备好几套战术, 在比赛前主教练则会根据不同的对手选择相应的战术, 比赛中也会根据对方打法的特点、执行当前战术的效果以及场上态势的变化而改变战术策

略。这就是人类比赛中先进的战术理念, 即通过对场上基本信息和状态的分析总结, 扬长避短地修订己方的战术策略。但是, 目前基于这种思想进行决策的研究并不多见, 大多是以场上的简单信息和基本处理通过各种方法进行决策<sup>[2-4]</sup>, 例如采用基于直觉模糊决策<sup>[5]</sup>, 强化学习<sup>[6-7]</sup>, 以及模糊推理<sup>[8]</sup>等决策方法, 但是这些决策没有很好地融合执行效果的反馈信息, 这样便导致战术变换欠缺灵活性, 不具有明显的有效性和针对性。文中正是基于人类这种先进的足球理念, 结合 SimuroSot5v5 仿真平台, 将场上局势信息和状态进行分析提炼, 并将机器人个体执行当前战术任务的效果作为反馈, 将正向评估和反向评估的信息综合量化后, 通过有限状态自动机转换模型做出适合当前态势的策略, 从而提高对抗竞技中己方战术策略的针对

收稿日期: 2012-12-12

修回日期: 2013-03-15

网络出版时间: 2013-05-09

基金项目: 西北工业大学研究生创业种子基金 (Z2011142)

作者简介: 娄云峰 (1987-), 男, 吉林长春人, 硕士研究生, 研究方向为智能决策、分布式计算。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20130509.1100.055.html>

性和有效性。

### 1 决策模型

#### 1.1 态势评估

态势评估<sup>[9]</sup>是指根据场上的各种基本环境状态信息,通过信息融合的方法和技术对场上所关心的态势信息进行分析和整合,再交给决策子系统作为战术策略的决策依据。诸如:将球的位置、速度、角度,各机器人位置和速度等基本环境信息,通过目前已有的态势分析算法及技术<sup>[10-11]</sup>提炼出球的控制权、局势对谁有利、对方薄弱区域、我方重点防守区域等影响战术决策的复合信息。

除了正向融合信息的态势评估之外,文中还将机器人个体的任务执行效果作为反馈信息进行评估和量化,即反向评估。诸如:前锋角色的机器人在某种战术下执行射门任务后,将执行效果经过评估量化后记录在决策系统的“黑板”上,这些反馈信息的评估结果同样作为决策系统中重要的决策依据。

#### 1.2 决策框架的设计

目前,决策框架的基本形式也较为成熟<sup>[6-8]</sup>,但考虑到决策中心要利用赛场基本信息、态势评估信息以及执行反馈信息作为决策系统的输入,依赖有限状态自动机的转换关系进行决策。因此,可得决策系统框架如图 1 所示。

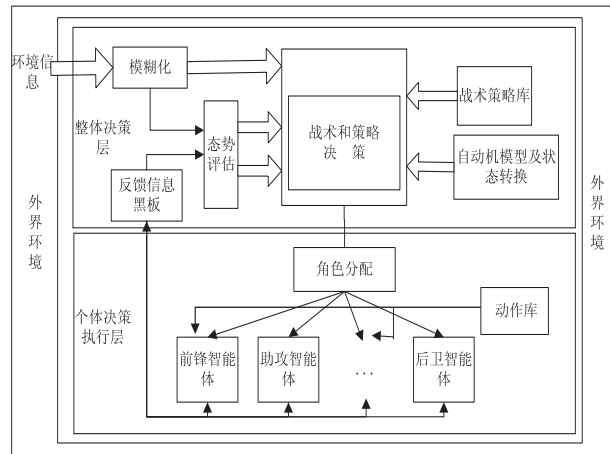


图 1 足球机器人决策框架

正如图上所示,正向态势评估结果和反馈信息的综合量化以及有限状态机中各策略状态之间的转换关系是整个决策系统的核心,而战术策略库和动作库可以根据实战经验和相应策略的需求来建立。

### 2 有限状态自动机形式化决策

#### 2.1 分层决策

为使决策目标更加清晰,决策结果对场上局势更具有针对性。根据决策内容和决策目标的不同将决策

子系统分为两层有限状态自动机决策模型<sup>[12]</sup>(如图 2 所示)。上层为战术策略决策,其主要作用相当于场边“主教练”,根据场上局势的变化,经过分析和评估后,将重新制定适合当前局面的整体战术策略,以当前局势信息的正向评估为主要决策条件;下层决策则是在某一个战术策略中,针对己方采用的战术为每个队员制定和部署任务,这一层决策会有较多的不同任务内容,历史执行效果和态势的正向评估都将作为该层决策的主要依据,这样对于多次执行效果不好的任务或动作将不再执行,执行效果好的任务或动作将会被优先选择执行,这样便有利于对任务或动作进行最优的决策和部署。

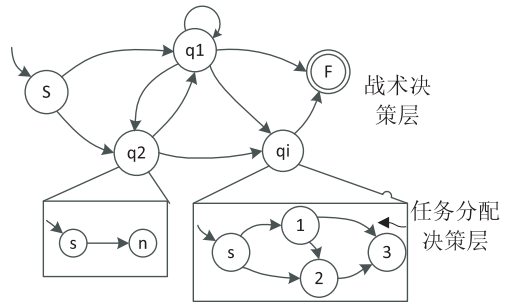


图 2 有限状态自动机分层决策模型

无论是战术策略还是任务部署决策模型,都可以建立其有限状态自动机(FSM)决策的形式化描述为五元组<sup>[9]</sup>:  $M = (Q, \Sigma, \delta, q_0, F)$ 。其中:

$Q$  为决策状态集合,  $Q = \{q_0, q_1, \dots, q_n\}$ ,  $\forall q_i \in Q$ ,  $q_i$  称为决策状态。在上层决策中即为战术策略集合,比如:全攻、全守、攻防兼备、重点防守、重点进攻等全局性的战术策略状态;而在下层的任务决策过程中,可以针对某一战术执行不同的具体任务,比如在重点进攻中可以设计为边路传中进攻、二过一进攻、门前进攻等具体的任务状态;

$\Sigma$  为态势评估综合量化指标集合,即决策模型的输入,也可以理解为战术策略转移的触发条件,  $\Sigma = (O_{ad}, H_i)$ ,  $O_{ad}$  为正向评估的综合量化结果,  $H_i$  为反向评估结果的综合量化结果。其具体评估方式和量化方法后文将进行详述;

$\delta$  为战术策略转移函数,即  $\delta: Q * \Sigma \rightarrow Q$ ,这需要根据足球比赛的理念和实际的比赛经验进行合理的设计;

$q_0$  为初始策略状态,根据实际比赛经验可设定为中场开球战术;

$F$  为终态,即决策的终止状态,根据实际比赛经验可设定为己方进球或者被进球。

#### 2.2 评估信息的综合量化

上述 FSM 决策模型中,态势评估综合量化指标  $\Sigma$  作为该决策模型中的关键输入,其包含的内容主要分

为基本信息  $I$ , 评估信息  $O$  和历史反馈信息  $H$ , 因此  $\Sigma = f(I, O, H)$ 。

为有效利用态势信息  $O$ , 将其分为进攻性和防守性两类信息进行分类评估, 进攻性评估信息  $O_a = \{O_{a1}, O_{a2}, \dots, O_{an_a}\}$ , 防守性评估信息  $O_d = \{O_{d1}, O_{d2}, \dots, O_{dn_d}\}$ ,  $O_{ai}, O_{di}$  均为评估后等级划分的结果, 且  $O_{ai}, O_{di} \in [0, 1]$ 。取值越大则表示越符合该评估项。例如对于评估项  $O_{ai}$  = “己方控球程度”, 若  $O_{ai} = 0.9$ , 则代表评估认为当前球在己方控制之内; 若  $O_{ai} = 0.3$ , 则表明当前球基本不在己方控制之内。将进攻性评估和防守性评估的结果进行综合量化即可得到正向评估结果  $O_{ad}$ , 计算方法如下:

$$O_{ad} = \frac{\sum_i O_{ai}}{n_a} / \frac{\sum_i O_{di}}{n_d} \quad (1)$$

反馈信息  $H = \{H_{11} \dots H_{ij} \dots H_{nm_a}\}$ ,  $H_{ij} \in [0, 1]$ ,  $H_{ij}$  表示为在第  $i$  种战术策略中的第  $j$  种任务的历史执行效果。其中  $n$  为战术策略的个数,  $m_i$  为第  $i$  种战术策略中的任务执行方案数。当  $H_{ij}$  值越大时, 则表明执行效果越好, 否则, 执行效果越差。初始值可设定为 0.5,  $H_{ij}$  更新方法为:

$$H_{ij} = (H_{ij} + \text{new}H_{ij}) / 2 \quad (2)$$

对于某一战术下的所有任务方案的执行反馈结果可综合为该战术总体的执行效果  $H_i$ , 其值反映了该战术策略在比赛中的总体效果, 其计算方法为:

$$H_i = (\sum_j H_{ij}) / m_i \quad (3)$$

因此, 综合上述三种评估信息即可得到自动机状态转换向量  $\Sigma = (O_{ad}, H_i)$ ,  $O_{ad}$  表明了当前攻防态势的综合信息,  $H_i$  则表明了对于某一策略状态执行的历史效果累计情况, 根据实际的比赛经验或专家意见可分别设定不同策略状态转换之间阈值  $G_0$  和  $G_H$ 。当某一状态执行效果较差(低于  $G_H$ ), 可根据  $O_{ad}$  的变化趋势及其关心的评估项迁移到其他的策略状态来继续执行, 从而可以改善当前赛场上的不利局势。对于下层的具体任务分配决策层, 状态转换向量可以细化为  $\Sigma = (O_a, O_d, H_{ij})$ , 其转换条件则更看重该任务执行的效果  $H_{ij}$  和具体的某正向评估指标  $O_i$ 。

### 2.3 决策过程

根据上述分层 FSM 决策模型和评估信息的处理过程, 可设计整个决策过程的关键步骤(如图3)为:

Step1. 初始化决策系统中的关键参数, 包括  $Q$ 、 $H_{ij}$ 、 $\delta$  以及  $G_H$  和  $G_0$  等;

Step2. 获取场上关键的基本信息  $I$ 。包括球和各机器人的位置、速度、方向角等基本状态信息;

Step3. 利用  $I$  通过态势评估计算当前局势进攻性和防守性的评估结果  $O_a$  和  $O_d$ ;

Step4. 将  $O_a$  和  $O_d$  通过式(1) 综合计算出  $O_{ad}$ , 通过式(3) 综合量化出战术策略的反馈评估信息  $H_i$ ;

Step5. 将  $O_{ad}$  和  $H_i$  作为上层战术策略决策的输入, 基于比赛经验或专家意见设计出战术策略转换模型及其转换关系, 以  $O_{ad}$  条件为主, 进行策略状态转换, 即上层战术策略决策;

Step6. 将  $O_a$ 、 $O_d$  和  $H_{ij}$  作为下层具体任务制定和分配决策的输入, 以各任务的历史执行效果  $H_{ij}$  为主, 进行任务决策转换;

Step7. 根据机器人自身的任务和角色, 在动作库中选择合适的动作完成当前自身的任务;

Step8. 当整体任务状态完成时, 将整体任务的执行进行效果评估得到  $\text{new}H_{ij}$ , 根据式(2) 更新相应的历史反馈信息  $H_{ij}$ ;

Step9. 每一次决策重复执行 Step2 到 Step8, 直到进球或者被进球时决策退出, 并重置初始状态, 等待下一次决策开始。

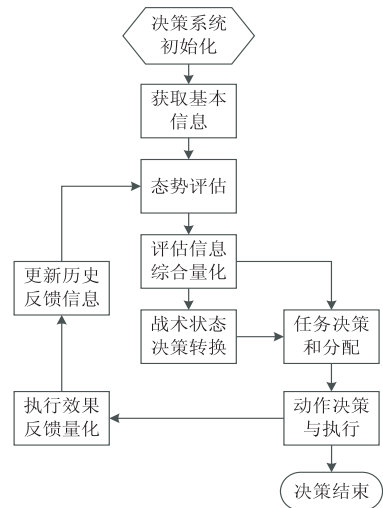


图3 决策执行流程图

### 3 仿真实验

为了验证文中提出的决策模型在整体战术策略决策上的有效性和针对性, 在 SimuroSot5v5 仿真平台上进行了仿真实验。实验采用加入了该决策模型的策略系统和改进前原策略系统进行比赛。

通过 C++ 语言按照上述基于态势评估和分层自动机的决策模型编写策略文件(即动态链接库), 其中, 蓝方 Stratage\_New.dll 就是增加了基于态势评估和分层自动机的决策文件, 而黄方 Stratage\_Old.dll 是原始的决策策略, 即通过对场上信息的简单处理, 采用 if……else……结构进行推理决策。通过与平台之间的数据通信, 获取球场上各个机器人的状态, 包括位置、角度等基本信息, 经过决策输出给平台己方机器人的左右轮速即可完成仿真比赛。



### 3 应对措施

应对措施需要从软件市场、软件分析、终端安全加固和用户安全意识等方面来进行:

1) 软件市场需要增强对软件的安全审查,例如谷歌的 Bouncer 服务<sup>[6]</sup>在保护 Android 用户免受恶意软件侵害的同时,为开发者提供比苹果应用程序审查政策更自由的政策。

2) 软件分析<sup>[7]</sup>: 针对应用软件可以进行静态分析<sup>[8]</sup>和动态分析。静态分析是针对软件的源代码或二进制程序进行数据流、控制流等方面的分析,例如 woodpecker 可以检测程序是否存在能力泄露问题<sup>[9]</sup>, stoway 则可以检查软件是否符合最小特权要求<sup>[10]</sup>等。动态分析将待检查软件安装到测试机上并运行,然后使用软件各项功能,并同时检测其对资源的调用,例如 taintdroid<sup>[11]</sup>可以跟踪程序是否泄露用户的敏感信息。

3) 终端安全加固: 可以从内核层、Android 框架层来进行,例如 NSA 推出了 SELinux 的 Android 适配版本,Android 框架层存在 Apex 等多个项目对原有的权限机制进行改进<sup>[12]</sup>。

4) 用户安全意识: Android 软件安装时会给出程序所需危险权限的提示,但很多用户并不理解,或者不顾安全风险安装来自非官方市场的软件。因此,提高安全意识是保证终端安全的重要因素。

### 4 结束语

针对智能移动终端的安全威胁进入了快速增长时期,除了用户安全意识和终端安全加固手段之外,还需要加强立法和服务提供商的安全监控等方面的研究。

#### 参考文献:

- [1] 黄海峰. 2011 年安卓手机木马激增六大安全厂商共话移动终端安全趋势[J]. 通信世界, 2012(9): 32-33.
- [2] 网 秦. 2011 年中国大陆地区手机安全报告[R/OL]. 2012-01-12. <http://www.netqin.com/upload/File/baogao/>

20120112. pdf.

- [3] 宋 杰, 党李成, 郭振朝, 等. Android OS 手机平台的安全机制分析和应用研究[J]. 计算机技术与发展, 2010, 20(6): 152-155.
- [4] 蒋绍林, 王金双, 张 涛, 等. Android 安全研究综述[J]. 计算机应用与软件, 2012, 29(10): 205-210.
- [5] Enck W. Defending Users against Smartphone Apps: Techniques and Future Directions[C]//Proc. of the 7th International Conference on Information Systems Security. Berlin: Springer-Verlag, 2011: 49-70.
- [6] 谷 歌. Android and Security[R/OL]. 2012. <http://googlemobile.blogspot.com/2012/02/android-and-security.html>.
- [7] 丁丽萍. Android 操作系统的安全性分析[J]. 信息安全, 2012(3): 28-31.
- [8] 童振飞. Android 恶意软件静态检测方案的研究[D]. 南京: 南京邮电大学, 2012.
- [9] Grace M, Zhou Y, Wang Z, et al. Systematic detection of capability leaks in stock Android smartphones[C]//Proceedings of the 19th Annual Symposium on Network and Distributed System Security. Hilton San Diego Resort & Spa, San Diego, California: [s. n. ], 2012.
- [10] Felt A P, Chin E, Hanna S, et al. Android permissions demystified[C]//Proceedings of the 18th ACM Conference on Computer and Communications Security. [s. l. ]: ACM, 2011: 627-638.
- [11] Enck W, Gilbert P, Chun B G, et al. TaintDroid: an information-flow tracking system for realtime privacy monitoring on smartphones[C]//Proceedings of the 9th USENIX conference on operating systems design and implementation. San Francisco: USENIX Association, 2010: 1-6.
- [12] Nauman M, Khan S, Zhang X. Apex: extending android permission model and enforcement with user-defined runtime constraints[C]//Proceedings of the 5th ACM Symposium on Information, Computer and Communications Security. [s. l. ]: ACM, 2010: 328-332.

(上接第 102 页)

- [6] 施卫强. 基于强化学习的足球机器人决策系统设计[D]. 长沙: 中南大学, 2007.
- [7] 王俊英, 李德华. 群决策专家权重自适应算法研究[J]. 计算机应用研究, 2011, 28(2): 532-535.
- [8] 李晓冰, 徐 扬. 基于直觉模糊推理的多属性群决策方法研究[J]. 计算机应用研究, 2012, 29(2): 533-535.
- [9] 唐进岭, 张著洪. 多项目多任务选择动态规划及其职能决策[J]. 计算机技术与发展, 2012, 22(9): 75-79.
- [10] 段 勇, 崔宝侠, 徐心和. 多智能体强化学习及其在足球机器人角色分配中的应用[J]. 控制理论与应用, 2009, 26(4): 371-376.
- [11] 陈建平, 杨宜民, 魏 良. 基于效用值的足球机器人系统任务分配方法[J]. 计算机工程, 2011, 37(1): 197-200.
- [12] Hopcroft J E, Motwani R, Ullman J D. Automata Theory, Languages, and Computation[M]. 3rd ed. Beijing: China Machine Press, 2008.

# 基于态势评估的足球机器人决策模型

作者: [娄云峰](#), [陈斌](#), [LOU Yun-feng](#), [CHEN Bin](#)  
作者单位: [娄云峰, LOU Yun-feng\(西北工业大学 计算机学院, 陕西 西安, 710072\)](#), [陈斌, CHEN Bin\(西北工业大学 电子信息学院, 陕西 西安, 710072\)](#)  
刊名: [计算机技术与发展](#)

ISTIC

英文刊名: [Computer Technology and Development](#)

年, 卷(期): 2013(9)

## 参考文献(12条)

1. [洪炳熔](#) [机器人足球比赛-发展人工智能的里程碑](#) 2000(04)
2. [Farahnakian F.](#) [Mozayani N](#) [Learning through Decision Tree in Simulated Soccer Environment](#) 2008
3. [Fan Bo.](#) [Pu Jiexin.](#) [Liu Gang](#) [Multi-agent Decision Making Based on Evidence Reasoning](#) 2009
4. [Ueda R.](#) [Arai T.](#) [Umeda K](#) [Real-time Decision Making with State-Value Function under Uncertainty of State Estimation - Evaluation with Local Maxima and Discontinuity](#) 2005
5. [雷英杰.](#) [王宝树.](#) [王毅](#) [基于直觉模糊决策的战场态势评估方法](#)[期刊论文]-[电子学报](#) 2006(12)
6. [施卫强](#) [基于强化学习的足球机器人决策系统设计](#)[学位论文] 2007
7. [王俊英.](#) [李德华](#) [群决策专家权重自适应算法研究](#)[期刊论文]-[计算机应用研究](#) 2011(02)
8. [李晓冰.](#) [徐扬](#) [基于直觉模糊推理的多属性群决策方法研究](#)[期刊论文]-[计算机应用研究](#) 2012(02)
9. [唐进岭.](#) [张著洪](#) [多项目多任务选择动态规划及其职能决策](#) 2012(09)
10. [段勇.](#) [崔宝侠.](#) [徐心和](#) [多智能体强化学习及其在足球机器人角色分配中的应用](#)[期刊论文]-[控制理论与应用](#) 2009(04)
11. [陈建平.](#) [杨宜民.](#) [魏良](#) [基于效用值的足球机器人系统任务分配方法](#)[期刊论文]-[计算机工程](#) 2011(01)
12. [Hopcroft J E.](#) [Motwani R.](#) [Ullman J D](#) [Automata Theory, Languages, and Computation](#) 2008

本文链接: [http://d.g.wanfangdata.com.cn/Periodical\\_wjz201309025.aspx](http://d.g.wanfangdata.com.cn/Periodical_wjz201309025.aspx)