

基于退火进化算法的机器人足球策略研究

杜秀全,程家兴

(安徽大学 计算机学院,安徽 合肥 230039)

摘要:针对一般遗传算法存在的一个显著的问题:“早熟收敛”与“快速收敛”之间的矛盾,解决早熟收敛的基本思想就是保持群体中个体的多样性,而模拟退火接受准则(即 Metropolis 准则)可以解决这方面的问题,可以避免搜索陷入局部极值,确保找到问题的整体最优解。因此在遗传算法中引入退火操作,定义了一个足球机器人的动作集合,根据赛场上的实际情况为足球机器人分配角色与任务,然后利用退火进化算法为足球机器人选择合适的动作,采用了一种高效的适应性评价方法,实验表明,应用退火进化算法的仿真足球机器人动作准确、效果更好。

关键词:遗传算法;模拟退火;足球机器人;Metropolis

中图分类号:TP18

文献标识码:A

文章编号:1673-629X(2008)02-0101-03

Study on Soccer Robot Strategy Based on Annealing Evolution Algorithm

DU Xiu-quan, CHENG Jia-xing

(Computer School, Anhui University, Hefei 230039, China)

Abstract: Aim at existence the prominent problem of the general genetic algorithm: inconsistency of between “precocious convergence” and “fast convergence”, the basic thought that the solution precocious convergence is maintains in individual multiplicity of the community, moreover simulation annealing accepts rule (namely the Metropolis rule) can solve this problem, may avoid searching falls into the partial extremum, insure to find the overall optimal solution of the problem. So introduce annealing operation in the genetic algorithm, a set of soccer robot actions is defined, roles and tasks are assigned to every soccer robot according to the situation of the game, then make use of annealing evolution algorithm for the soccer robot to choose appropriate actions, used one highly effective adaptability estimate method. Experiments show that the soccer robots applied annealing evolution algorithm own more accurate actions than others.

Key words: genetic algorithm; simulated annealing; soccer robot; Metropolis

0 引言

机器人足球世界杯(RoboCup, Robot World Cup)是国际上为促进分布式人工智能、智能机器人技术研究与发展而举行的大型比赛和学术活动。RoboCup以多智能体系统和分布式人工智能为主要研究背景,其主要目的就是通过提供一个标准的、易于评价的比赛仿真平台,检验各种智能机器人技术,促进多智能体系统与分布式人工智能的研究与发展。仿真足球队的设计是一个非常复杂的问题,每个球员是一个Agent^[1],如只考虑22名Agent在场上的位置、速度以

及身体朝向等三个因素,场上的状态就有 10^{198} 种之多^[2],再加上球的位置、速度以及过去的状态,场上的状态数就会更多。对于如此庞大的一个状态空间,仅仅依靠人的经验进行手工编码来处理Agent的所有行为是不可能的。因此,很多研究者都尝试应用机器学习的方法^[2]来解决这个复杂的问题。

设计了一种用于足球机器人动作规则的退火进化算法,算法引入了模拟退火的思想,给出了一种适应性评价函数,仿真实验表明,算法的收敛快,全局收敛性都有了明显的提高,应用该算法的仿真足球机器人动作灵活、准确、效果好。

1 机器人动作设计

在应用退火进化算法为足球机器人进行动作选择时,机器人的动作粒度的选择非常重要,如果动作的粒度过小,则搜索的空间大,并且难于评价,相反,如果动作的粒度过大,动作集合难于设计^[3]。

收稿日期:2007-05-24

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);安徽大学研究生创新项目(20073056)

作者简介:杜秀全(1982-),男,安徽全椒人,硕士研究生,研究方向为智能计算理论与应用、优化方法、机器人仿真足球;程家兴,教授,博导,研究方向为智能计算与优化方法等。

为了方便评价,将机器人的动作分为 2 人层次进行描述:

1) 总体描述:包括动作目的、总的方式、持续时间、是否有球员配合、动作的类型等。

2) 详细描述:包括动作实现的细节、规则及在以后若干周期的速度、加速度、速度差等等。

笔者认为赛场中的机器人动作可以分为以下类型:

- (1) 直线型:向前或向后直线运动;
- (2) 旋转型:向右或向左旋转;
- (3) 弧线型:向前或向后,向右或向左的弧线运动。

2 模拟退火算法

模拟退火算法^[4]最初的思想由 Metropolis 在 1953 年提出,Kirkpatrick 在 1983 年成功地将其应用在组合最优化问题中。

模拟退火算法来源于固体退火原理,将固体加温至充分高,再让其徐徐冷却,加温时,固体内部粒子随温度升变为无序状,内能增大,而徐徐冷却时粒子渐趋有序,在每个温度都达到平衡态,最后在常温时达到基态,内能减为最小。根据 Metropolis 准则,粒子在温度 T 时趋于平衡的概率为 $e^{-\Delta E/(kT)}$,其中 E 为温度 T 时的内能, ΔE 为其改变量, k 为 Boltzmann 常数。用固体退火模拟组合优化问题,将内能 E 模拟为目标函数值 f ,温度 T 演化成控制参数 t ,即得到解组合优化问题的模拟退火算法:由初始解 i 和控制参数初值 t 开始,对当前解重复“产生新解 \rightarrow 计算目标函数差 \rightarrow 接受或舍弃”的迭代,并逐步衰减 t 值,算法终止时的当前解即为所得近似最优解,这是基于蒙特卡罗迭代求解法的一种启发式随机搜索过程。退火过程由冷却进度表(Cooling Schedule)控制,包括控制参数的初值 t 及其衰减因子 Δt 、每个 t 值时的迭代次数 L 和停止条件 S 。

2.1 模拟退火算法的模型

模拟退火算法可以分解为解空间、目标函数和初始解三部分。

模拟退火的基本思想:

- (1) 初始化:初始温度 T (充分大),初始解状态 S (是算法迭代的起点),每个 T 值的迭代次数 L ;
- (2) 对 $k = 1, \dots, L$ 做第(3)至第(6)步;
- (3) 产生新解 S' ;
- (4) 计算增量 $\Delta t' = C(S') - C(S)$,其中 $C(S)$ 为评价函数;
- (5) 若 $\Delta t' < 0$ 则接受 S' 作为新的当前解,否则以

概率 $\exp(-\Delta t'/T)$ 接受 S' 作为新的当前解;

(6) 如果满足终止条件则输出当前解作为最优解,结束程序。终止条件通常取为连续若干个新解都没有被接受时终止算法;

(7) T 逐渐减少,且 $T \rightarrow 0$,然后转第(2)步。

2.2 利用模拟退火的 Metropolis 准则进行数学描述

对于一个最小的组合优化问题 (S, f) , S 是所有可行解的集合,构成解空间; f 是目标函数。在搜索问题最优解的过程中,假设当前解为 i , j 是 i 邻域中的一个解,则在温度参数为 t 时, j 替代 i 作为当前解的概率 $A_{ij}(t)$ 可定义如下:

$$A_{ij}(t) = \begin{cases} 1, & \text{如果 } f(j) \leq f(i) \\ \exp\left[-\frac{f(j) - f(i)}{t}\right], & \text{如果 } f(j) > f(i) \end{cases}$$

由于模拟退火具有概率突跳的双向搜索能力,既容易跳出局部极值的陷阱,又能确保搜索的全局优化性。如果将模拟退火引入到遗传算法策略中,可以丰富遗传算法的搜索行为,避免出现早熟收敛,同时又可以利用遗传算法本身强大的并行全局搜索能力。

3 退火进化算法描述

退火进化算法是将模拟退火的 Metropolis 准则引入到遗传算法的群体更新阶段,通过模拟退火的 Metropolis 准则自适应地调整遗传算法进化过程的选择压力,在保证群体多样性的同时,又能在进化的后期,逐渐加快收敛速度,克服遗传算法应用中的早熟收敛现象,提高搜索效率,得到全局最优解。退火进化算法的流程如图 1 所示。

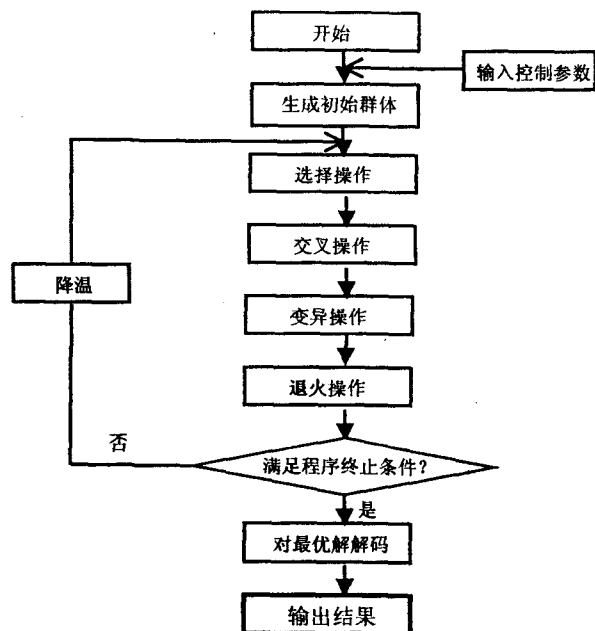


图 1 退火进化算法

模拟退火中的 Metropolis 接受准则引入到遗传算法的群体更新策略中,即在用新群体更新旧群体时,对新群体中每个个体施行退火操作,决定进入下一代的个体,实现新旧群体的替换。基于退火进化算法的机器人足球描述如下:

1)设置遗传算法以及退火过程中涉及到的各控制参数、包括群体规模、交叉概率、变异概率、初始温度和退火迭代次数等。

2)初始化程序中的各个变量。

3)生成初始群体。

4)选择被复制而直接进入下一代或参与交叉操作的个体。

5)对选择的个体按照交叉概率实施交叉操作,生成新群体,并采用择优原则,即如果生成的新个体目标函数值大于当前最优解,则新个体替换为当前最优解,否则最优解仍保持不变^[5]。

6)对新群体中的所有个体按照变异概率进行变异操作,同样采用择优原则:对变异操作生成的新群体中所有个体实施退火操作,实现新旧群体的更新。

判断是否满足程序终止条件。若不满足程序终止条件,则按照定义的降温方式更新温度参数 T 。如果满足程序终止条件,则终止进化,输出最优解。

4 适应度函数的设计

遗传算法^[6]在进化搜索中基本上不用外部信息,仅以适应度函数 fitness function 为依据,利用种群中每个个体的适应度值来进行搜索。因此 fitness function 选取直接影响到遗传算法的收敛速度以及能否找到最优解。通过适应度函数来决定个体的优劣程度,它体现了自然进化中的优胜劣汰原则。适应度函数设计直接影响到遗传算法的性能,适应度函数设计不当有可能造成遗传算法的欺骗性问题。

针对在机器人足球比赛中,目标是将球踢进对方的球门,也就是说球离对方球门的距离越近越好,所以将球离对方球门的距离的变化作为评价函数,定义如下:

$$R = \Delta d$$

其中, Δd 表示球离对方球门距离的变化,前进为正,后退为负。

5 仿真实验

在 FIRA 仿真的环境下,对该动作规则算法进行了实验。为了与参考文献[4]比较,进行了两个实验:截球实验和射门实验。各项实验进行 100 次,统计结果如表 1 所示,实际比赛成绩与实验结果表明,采用了退火进化算法后,策略子系统截球能力、射门能力都得到了提高。

表 1 实验成功率统计

实验类型	试验次数	球速 $\text{cm} \cdot \text{s}^{-1}$	成功次数	成功率
截球	100	0.4~1.48	91	0.91
射门	100	0.4~1.48	84	0.84

6 结论

多 Agent 协作从 20 世纪 80 年代就被提出,已经有一些探索性的研究,但总的来说,多智能体协作技术还不成熟,在协作时机、协作过程、协作机理、协作协议、协作稳定性等方面的研究还远远不够。应进一步深入研究多智能体的协作模型和协作方法,建立具有集体理性、能完成复杂任务的、稳定的协作智能体。如何训练 Agent 以及加强 Agent 相互之间的协作是 Multi-Agent 系统研究的一个重要课题,文中针对遗传算法的缺点和退火算法的优点,有效地将两者结合在一起,通过实验验证了该算法的可行性,提出的方法避免了传统遗传算法出现的早熟收敛,提高了搜索效率。

参考文献:

- [1] Zhang Bo, Cai Qingsheng, Chen Xiaoping, et al. An Agent Team for RoboCup Simulator League[C]//Proceedings of the 3rd World Congress on Intelligent Control and Automation. 合肥:中国科学技术大学出版社,2000.
- [2] 杨增光,李龙澍. 决策树学习在 RoboCup 仿真球队中的应用研究[J]. 系统仿真学报,2004,16(4):653-656.
- [3] 刘 钊,陈建勋. 基于自适应遗传算法的足球机器人策略设计[J]. 哈尔滨工业大学学报,2005,37(7):7-9.
- [4] 杨学星,丁海军. 基于模拟退火算法和遗传算法的图像降噪研究[J]. 计算机工程与应用,2006(4):79-80.
- [5] 毕惟红,任红民,吴庆标. 一种新的遗传算法最优保存策略[J]. 浙江大学学报:理学版,2006(1):32-35.
- [6] 金 玺,郭 巧,黄 鸿,等. 在机器人足球中利用遗传算法进行多智能体学习[J]. 机器人技术与应用,2003(5):45-47.

(上接第 100 页)

propagation[J]. ACM Transactions on Graphics,2005,24(3):861-868.

[10] Fattal R, Lischinski D, Werman M. Gradient domain high dynamic range compression[J]. ACM Transactions on Graphics,

2002,21(3):249-256.

[11] P'erez P, Gangnet M, Blake A. Poisson image editing[J]. ACM Transactions on Graphics,2003,22(3):313-318.