

文章编号: 1005-3026(2005) 10-0949-04

中国象棋计算机博弈系统评估函数的自适应遗传算法实现

王 骄, 王 涛, 罗艳红, 徐心和
(东北大学 信息科学与工程学院, 辽宁 沈阳 110004)

摘 要: 使用自适应遗传算法解决中国象棋计算机博弈问题。将博弈问题分解为搜索引擎、走法生成、评估函数和开局库四大模块, 然后将自适应遗传算法引入到评估函数中, 通过锦标赛算法对评估函数中的参数组合进行自动调整和优化。设计并开发了基于上述方法的离线自学习系统, 实验结果证明提高了程序的棋力。

关 键 词: 中国象棋计算机博弈; 博弈树; 评估函数; 锦标赛算法; 自适应遗传算法
中图分类号: TP 181 文献标识码: A

在过去的半个世纪里, 世界各地的学者花费了大量的心血对于计算机博弈——包括奥赛罗、checker、国际象棋进行研究。这是因为计算机博弈是人工智能的一块试金石, 各种搜索算法、模式识别及智能方法在计算机博弈中都可以得到广泛的应用。在长时间的研究中, 涌现出大量令人震惊的成果, 1997 年“深蓝”战胜卡斯帕罗夫的比赛就在全世界范围内引发了震动。在以上的三种棋类中, 计算机的水平都已经达到了世界冠军的水平。

近年来, 中国象棋计算机博弈也逐渐引起众多学者的关注, 这是因为中国象棋是世界上历史最为悠久的棋类, 而且它的空间复杂性和树的复杂性都要高于以上的三种棋类^[1]。

1 中国象棋计算机博弈问题描述

中国象棋计算机博弈可以分解为 4 个主要部分^[2]: 搜索引擎、走法生成模块、评估函数和开局库。

几乎所有的棋类问题, 都可以用博弈树来描述。博弈树^[3]是把计算机和用户所有可能走法和局面罗列出来的一颗树。红黑双方交替地按合理走法把树展开, 树的每一个节点都表示某一个特定局面。根节点表示的是当前需要计算的局面, 中间节点表示的是对弈过程中的某一个局面, 叶子节点是树的最底端, 表示可以推导的局面。叶子节

点和根节点之间的最大距离, 称为搜索深度。整个博弈树描述的是从当前局面出发, 包含所有可能的对弈过程的搜索树。中国象棋计算机博弈问题也就转化为寻求最佳路径的问题。

对于树中的每一个节点来说, 红黑双方都会从子节点中选择最有利于自己的分枝。因为博弈树中值的传递是由下至上的, 这就要求对叶子节点表示的局面必须有一个极为准确的打分。对于局面最为准确的估计莫过于已经分出胜负的情况, 即建立在叶子节点分出胜负的完全博弈树。中国象棋的完全博弈树大概有 10^{144} 个节点, 建立这个博弈树已经远远超出了当代计算机的处理能力。惟一的解决方法就是让博弈树扩展到计算机运算可以接受的深度, 然后对没有分出胜负的叶子节点给出一个最为准确的打分, 表示此局面下取得胜利的可能性。

$$Value = ComputerValue - UserValue$$

如果 Value 为正数, 说明局面对电脑有利, 为负则对用户有利。

Value 这个打分是由评估函数计算得到的。而按照中国象棋的走法规则生成合理走法将树展开, 是走法生成模块的功能。搜索引擎则是尽可能缩小树的规模, 避免一切冗余的计算。

开局库是独立于博弈树之外的模块^[4], 开局库存储了大量的专家棋谱, 如果根节点的局面在

开局库可以查询到,则提取开局库的对应走法而不必展开树。可以避免在开局时由于搜索深度的不足而带来战略上的失误。

评估函数是模式识别和智能算法应用最为广泛的领域。不管多么复杂的评估函数,都可以表示为一个多项式。评估函数一般来说必须包括 5 个方面的要素^[1],分别是固定子力值、棋子位置值、棋子灵活度值、威胁与保护值、动态调整值。每一方面的值又是由许多参数值构成的,例如固定子力值就包括 7 类棋子的值。即使最简单的评估函数也有 20 多个参数,将这些值线性地组合在一起得到最终的评估值,通常依赖于编程者的棋类知识,这就带来很大的局限性,通过手动调整很难达到全局最优。

为了解决这个参数优化的问题,本文在中国象棋计算机博弈领域引入了遗传算法。通过离线自主学习训练得到最佳参数组合。

2 基于自适应遗传算法的象棋评估函数

有学者已经把遗传算法引入了评估函数^[5],试图学习棋子之间内部的动态联系,但效果不是很理想。本文仅仅用遗传算法做参数优化,并且为了加强遗传算法的局部搜索能力,在进化过程中采用了自适应遗传算法。

2.1 遗传算法

遗传算法(Genetic Algorithm, GA)是近些年发展起来的一种崭新的全局优化算法,由 Holland 教授首次提出^[6]。它借用了生物遗传学的观点,通过选择、遗传、变异等作用机制,实现个体适应性的提高,体现了自然界中“物竞天择、适者生存”的进化过程。遗传算法吸引了大批的研究者,迅速推广到优化、搜索、机器学习等方面。

传统的爬山法^[7]是通过参数小范围的试探来寻求最优解,并且受初始采样值取值范围和单方向搜索的限制,很容易陷入局部最优。

模拟退火是一种基于蒙特卡罗迭代求解的启发式随机搜索算法,它试图通过模拟高温物体退火过程的方法,寻找优化问题的全局最优解或近似全局最优解。它可以避开局部最优点,但是致命的缺点是速度太慢,运行时间太长。

遗传算法同时使用多组搜索信息,具有很好的全局搜索能力,将重点集中在性能较高的地方,具有很快的搜索速度;而且遗传算法的鲁棒性也明显优于前两种算法,最有可能在象棋中取得成功。

2.2 适应度函数的计算——锦标赛算法

众所周知,遗传算法在进化搜索中基本不需要外部信息,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函数来进行搜索。一般而言,适应度函数是由目标函数变换而成的。但是象棋的评估比较复杂,很难像一般优化问题那样找到真实准确的目标函数,因此本文设计了一个专门用于棋类优化问题的适应度函数计算方法——锦标赛算法。

首先,要对自己开发的计算机博弈软件进行拓展,拓展后的程序包含了两个评估函数,但是搜索引擎、走法生成都是共用的。两个评估函数通过读取不同的参数组进行比赛,决出胜负。这样,就可以通过竞赛来确定下棋双方的适应度;在相同适应度的初始状态下,赢了对适应度进行加运算(奖赏),输了对适应度进行减运算(惩罚),和了则不操作。这样就解决了适应度函数不好确定的问题,实现了通过适应度函数来决定个体的优劣程度,体现了自然进化中的优胜劣汰原则。

遗传算法处理的对象主要是个体^[8],个体包括一组染色体串,每一个染色体对应于评估函数中的一个参数值。将染色体串解码到评估函数中即可求得评估值。染色体则是一个基因数组,基因的长短决定编码的精度。

至于染色体串的编码方式,本文选用的是通用的二进制编码方式,由于对解的精度要求不高,其编码和解码都可以快速简单地实现,而且二进制编码进行交叉变异操作时也比其他编码更加简便方便。

2.3 加快优化进程的相关举措

(1) 由于加入开局库,会使比赛时间变得很长;但如果没有开局库,对下棋过程的质量又会有很大影响,不能真正反映个体的好坏。所以本文对初始的比赛环境做了改动,初始状态是由包括许多红黑局面相当的中局数组构成,在进化了规定数量的代数以后,会自动修改比赛环境。

(2) 由于计算机之间的比赛中经常遇到死循环,所以本文也对循环加入了一些规则进行判断,避免了这种情况的出现。

(3) 由于象棋中的胜负是依据将或帅的存在与否进行判断,所以将和帅的值默认为无穷大,并不参与进化。

以上这些做法都有利于提高比赛的精彩程度,更快地选出真正优秀的个体,加快进化的速度。

2.4 遗传操作过程

(1) 锦标赛选择

遗传算法可采用的选择方法很多,有轮盘赌选择法、局部选择法、截断选择法和锦标赛选择法,本文采用了锦标赛选择法和精英选择策略^[9]。每两个个体之间进行先后手互换的两场比赛,取出适应度最高的一些精英作为下一代的父个体。锦标赛模式的缺点在于速度过慢。为了加快速度,本文对于标准锦标赛做了改进。将包含 m 个个体的种群随机分成 n 组,在组内进行锦标赛训练,每个组的冠军为筛选出的最佳,对 n 个冠军做交叉和变异操作产生 $m-n$ 个新个体,填满并形成一个新种群。做了这样的优化后,以包含 60 个个体的种群为例,在保证优胜劣汰原则的基础上,可以节省约 75% 的时间。

(2) 均匀交叉

交叉方法很多,有单点交叉、多点交叉、顺序交叉、循环交叉等等,为了使交叉在便于操作的同时更加广义化,本文选用了均匀交叉,参数之间的间隔点作为潜在的交叉点。均匀交叉根据交叉率 p_c 随机地产生与参数个数等长的 0—1 掩码,掩码中的片断表明了哪个父个体向子个体提供变量值。通过这个掩码和选择的父个体一起确定了子个体。

(3) 变异

变异是指以等于变异率 p_m 的概率改变一个或几个基因,对于二进制串来说,就是根据变异率实现基因的 0—1 翻转。变异是一种局部随机搜索,与选择/交叉算子结合在一起,保证了遗传算法的有效性,使遗传算法具有局部的随机搜索能力。同时使得遗传算法保持种群的多样性,以防止出现非成熟收敛。

2.5 自适应遗传算法

在整个遗传算法实现的过程中,交叉率 p_c 和变异率 p_m 的选择是影响遗传算法行为和性能的关键所在,直接影响算法的收敛性。 p_c 越大,新个体产生的速度就越快,然而 p_c 过大时遗传模式被破坏的可能性也越大,使得具有高适应度的个体结构很快就会被破坏;但是如果 p_c 过小,会使搜索过程缓慢,以至停滞不前。对于变异率 p_m ,如果 p_m 过小,就不易产生新的个体结构;如果 p_m 取值过大,那么遗传算法就变成了纯粹的随机搜索算法。但是,目前还没有通用的能够一次性确定 p_c 和 p_m 的方法。针对不同的优化问题,需要反复通过实验来确定 p_c 和 p_m ,这是一件非常繁琐的工作。为此,本文引进了自适应遗传算法^[10](AdaptiveGA,AGA), p_c 和 p_m 能够随着适应度

自动改变。
 p_c 和 p_m 的计算公式如下:
$$p_c = \begin{cases} p_{c1} - \frac{(p_{c1} - p_{c2})(f' - f_{avg})}{f_{max} - f_{avg}}, & f' \geq f_{avg}; \\ p_{c1}, & f' < f_{avg} \end{cases}$$
$$p_m = \begin{cases} p_{m1} - \frac{(p_{m1} - p_{m2})(f_{max} - f)}{f_{max} - f_{avg}}, & f' \geq f_{max}; \\ p_{m1}, & f' < f_{max} \end{cases}$$

式中, $p_{c1}=0.9$, $p_{c2}=0.6$, $p_{m1}=0.1$, $p_{m2}=0.001$ 。 f_{max} 为群体中最大的适应度值; f_{avg} 为每代群体的平均适应度值; f' 为要交叉的两个个体中较大的适应度值; f 为要变异个体的适应度值。
这种自适应的 p_c 和 p_m 能够提供相对某个解的最佳 p_c 和 p_m ,在保持群体多样性的同时,保证了遗传算法的收敛性。

3 实验结果

为了测试进化的效果,设计了一个测试实验,实验是由 100 代、200 代、300 代、400 代、500 代、600 代的冠军和两个随机选手组成一个小组,进行循环赛,每两个对手之间都要进行互换先后手的两场比赛。胜者积分加 1,负者减 1。最后的积分如图 1 所示。

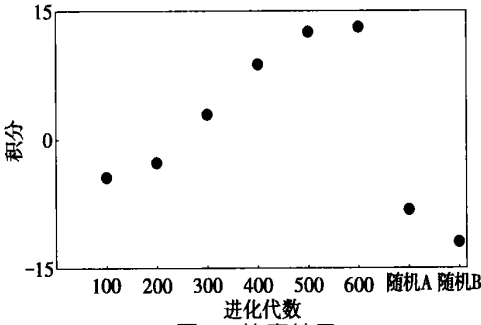


图 1 比赛结果
Fig. 1 Game results

如图 1 所示,得分随着代数的增加单调增加,而参加比赛的随机选手基本没有取胜的机会。在进化到 600 代时,本文给出了一部分训练的结果,此时的固定子力值如表 1 所示。

表 1 训练结果
Table 1 Training results

车	马	炮	士	象	兵
989	439	442	226	210	55

可以看到,固定子力值的大小排列顺序符合人的经验,车的值略大于马和炮的和,马和炮的值都约等于象和士的和,这都是象棋高手默认的规律。

用 600 代的参数组和完全用经验设定的参数

组的程序进行了比赛, 结果 600 代的胜率明显占优。

4 结 论

在中国象棋的评估函数中, 引入了自适应遗传算法优化参数组合, 它加强了传统遗传算法的局部搜索能力, 避免了传统的爬山法容易陷入局部最优的缺点, 而且不像模拟退火那样收敛过慢。它同时维护一组参数, 既可以继承已有参数中的优良内容, 又可以避免因局部振荡而失败, 并且具有很好的收敛速度。

当程序拥有进化参数的评估函数时棋力明显提高, 每当加入了一些新参数, 丰富了棋子之间联系的奖励和惩罚时, 都可以用自适应遗传算法进行优化, 而不是传统上的手工调整, 可以最大限度发挥评估函数的作用。这也为开发出高水平的计算机博弈程序打下了坚实的基础。

参考文献:

[1] Yen S J, Chen J C, Yang T N. Computer Chinese chess[J]. *ICGA Journal*, 2004, (3): 3—18.
[2] 王小春. PC 游戏编程[M]. 重庆: 重庆大学出版社, 2002. 1—27.

(Wang X C. *PC game programming*[M]. Chongqing: Chongqing University Press, 2002. 1—27.)
[3] Marsland T A. Computer chess and search[D]. Edmonton: University of Alberta, 1991.
[4] Hsu S C, Tsao K M. Design and implementation of an opening game knowledge-base system for computer Chinese chess[R]. *Bulletin of the College of Engineering, N T U*, 1991, (53): 75—86.
[5] Lorenz D, Markovitch S. Derivative evaluation function learning using genetic operators[A]. *Proceedings of the AAAI Fall Symposium on Games: Planing and Learning* [C]. New Carolina, 1993. 106—114.
[6] Holland J H. *Adaptation in nature and artificial system* [M]. Ann Arbor: The University of Michigan Press, 1975.
[7] 米凯利维茨 Z. 演化程序——遗传算法和数据编码的结合[M]. 北京: 科学出版社, 2000. 20—23.
(Michalewicz Z. *Genetic algorithm + data structure = evolution programs*[M]. Beijing: Science Press, 2000. 20—23.)
[8] 李敏强. 遗传算法的基本理论与应用[M]. 北京: 科学出版社, 2002. 1—15.
(Li M Q. *Genetic algorithm's base theory and application* [M]. Beijing: Science Press, 2002. 1—15.)
[9] 王小平, 曹立明. 遗传算法——理论、应用与软件实现[M]. 西安: 西安交通大学出版社, 2002. 195—210.
(Wang X P, Cao L M. *Genetic algorithm—theory, application and programming implement* [M]. Xi'an: Xi'an Jiaotong University Press, 2002. 195—210.)
[10] Angeline J, Pollack B. Competitive environments evolve better solution for complex tasks[A]. *Proceedings of the Fifth International Conference on Genetic Algorithms* [C]. Urbana-Champaign, 1993. 264—270.

Implementation of Adaptive Genetic Algorithm of Evaluation Function in Chinese Chess Computer Game System

WANG Jiao, WANG Tao, LUO Yan-hong, XU Xin-he
(School of Information Science & Engineering, Northeastern University, Shenyang 110004, China. Correspondent: WANG Jiao, E-mail: soldierwj@sina.com)

Abstract Adaptive Genetic Algorithm (AGA) is used to solve the problem of chinese chess game on computer. The game system is divided into four parts: searching engine, move generator, evaluation function, and opening book. Then, the AGA is brought into evaluation function to automatically adjust and optimize the parameters' combination of evaluation function by tournament algorithm. Based on the above methods, an off-time self-study system is designed. The experiment results show that the AGA improves the performance of the program developed.
Key words Chinese chess computer game; game tree; evaluation function; tournament algorithm; adaptive genetic algorithm
(Received February 1, 2005)

待发表文章
摘要 预报

热轧工艺中加热温度对 IF 钢组织性能的影响

张锦刚, 蒋奇武, 刘沿东, 左 良

选取 Ti-Nb-IF 钢热轧工艺中的加热温度为影响因素, 采用显微观察、TEM 二相粒子分析、织构分析等实验分析手段, 研究了 1 140 ℃和 1 214 ℃两种热轧加热温度制备的 Ti-Nb-IF 钢样品在随后的冷轧、退火和平整过程中显微组织演变和织构演变特征的影响及最终产品的力学性能。结果表明, 低温加热有利于 IF 钢屈强比的降低及深冲性能的提高, 加热温度为 1 140 ℃时, 其屈服强度可降低 30 MPa, 塑性应变比可提高 0.16。其主要原因为低温加热工艺保障 IF 钢热轧后产生粗大的二相粒子和细小铁素体晶粒, 在随后的冷轧和退火处理过程中产生分布均匀和强的 γ 再结晶织构。