

网页新闻贴吧知道音乐图片视频地图文库百科

百度首页 登录 注册

Baidu 百科

UCT算法

进入词条

全站搜索

帮助

声明：百科词条人人可编辑，词条创建和修改均免费，绝不存在官方及代理商付费代编，请勿上当受骗。详情>>

首页 分类 特色百科 用户 权威合作 手机百科

个人中心

UCT算法

本词条由“科普中国”百科科学词条编写与应用工作项目 审核。

UCT算法（Upper Confidence Bound Apply to Tree），即上限置信区间算法，是一种博弈树搜索算法，该算法将蒙特卡罗树搜索(Monte—Carlo Tree Search，MCTS)方法与UCB公式结合，在超大规模博弈树的搜索过程中相对于传统的搜索算法有着时间和空间方面的优势。

中文名	上限置信区间算法	提出者	Levente Kocsis与Csaba Szepesvári
外文名	Upper Confidence Bound Apply to Tree	提出时间	2006
别 称	树图置信算法	应用学科	计算机科学
		适用领域范围	计算机博弈

目录	<div><div>1 提出</div><div>2 研究现状</div><div>3 基本思想<ul style="list-style-type: none">树内选择策略</div></div> <div><div>• 缺省仿真策略</div><div>• 仿真结果回传</div><div>4 算法优势</div><div>5 应用</div><div>6 领域发展</div></div> <div><div>• 深蓝——蛮算硬汉</div><div>• 浪潮天梭——以一敌五</div><div>• 沃森——全才学霸</div></div>
----	--

提出

早些年，计算机博弈对于棋类游戏的研究集中在基于模式识别和专家系统的方法上(最典型的是基于静态评估函数的仅一B博弈树方法)，并在国际象棋、中国象棋等项目中获得了成功。但是对于围棋类的项目，传统的方法一直无法取得满意的结果。在2000年前后，世界上最高水平的计算机围棋软件的棋力还比不上人类的业余初段。

围棋作为一种特殊的完备信息博弈，由于其本身的非完备特性，也成为了蒙特卡罗方法应用的载体之一。但是，由于围棋的复杂度高，且极具欺骗性，对计算机程序提出了巨大的挑战。为了处理如此众多的可能情况，人工智能专家已经设计出一些算法，来限制搜索的范围，但它们都无法在大棋盘的比赛中战胜实力稍强的人类棋手。2006年秋季，两位匈牙利研究人员报告了一种新算法，它的胜率比现有最佳算法提高了5%，能够在小棋盘的比赛中与人类职业棋手抗衡。这种被称为UCT(UpperConfidenceboundsappliedtoTrees)的算法，是匈牙利国家科学院计算机与自动化研究所（位于布达佩斯）的列文特·科奇什(LeventeKocsis)与加拿大阿尔伯塔大学（UniversityofAlberta，位于埃德蒙顿）的乔鲍·塞派什瓦里(CsabaSzepesvári)合作提出的，是著名的蒙特卡罗方法(MonteCarlomethod)的扩展应用。^[1]

研究现状

法国南巴黎大学的数学家西尔万·热利(SylvainGelly)与巴黎技术学校的王毅早(YizaoWang，音译)将UCT集成到一个他们称之为MoGo的程序中。该程序的胜率竟然比先前最先进的蒙特卡罗扩展算法几乎高出了一倍。2007年春季，MoGo在小棋盘的比赛中击败了实力强劲的职业棋手，在大棋盘比赛中也击败了实力稍弱的业余棋手，充分展示了能力。热利认为UCT易于实现，并有进一步完善的空间。科奇什预言，10年以后，计算机就能攻克最后的壁垒，终结人类职业棋手对围棋的统治。

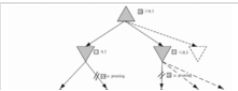
哈工大深圳研究生院智能计算中心也是国内最早探索非完备信息博弈的机构之一。经过3年的研究，已经取得了一定成功，并在此基础上开发了一套拥有完整人机交互界面并具有较强智能水平的四国军棋博弈系统。^[1]

基本思想

由于围棋有很大的分支系数（BranchingFactor），传统搜索技术一直没有在计算机围棋领域取得有意义的成果，2006年UCT算法改变了这种局面。UCT算法是一种特殊的蒙特卡罗搜索算法，它由树内选择策略、缺省仿真策略和仿真结果回传三部分组成。

树内选择策略

如图1所示，传统搜索技术都有搜索深度d这个参数。当搜索达到深度d时，搜索算法从评估函数取得评估值，而搜索算法负责找到让评估值最大的分支。这种在同一深度d获取评估值的方式，让搜索深度d，分支系数b，与搜索树叶子节点数N的关系为



V百科

往期回顾

汉字

微科普

最难的汉字，你会几个？

【科普】

网络流行语

第九期

权威合作编辑

“科普中国”百科科学词条编写...

“科普中国”是为我国科普信息化建设塑造的全...

什么是权威编辑

查看编辑版本

资源提供

中国电子学会

中国电子学会（Chinese Instit...

提供资源类型：内容

什么是资源合作

词条统计

浏览次数：7039次

编辑次数：1次历史版本

最近更新：2016-03-13

创建者：clei词条

1 人脸识别 算法

2 WebRTC

3 人脸识别

4 编程学习入门

5 1万元存一年多

6 人工翻译

7 linux红帽认

8 服务机器人

9 电商分销系统

10 美国移民

11 人脸识别系统

12 大数据如何学

13 大数据分析培

14 投资理财公司

15 粤港车牌

16 车牌识别系统

17 成都实木家具

18 excel培训

19 海南文昌房价

20 智能大棚

21 机器人编程入

22 人工智能培

百科数字博物馆
国宝居然会说话？

第1页 共4页

2018/1/2 上午10:20

$N = b^d$ 。UCT 算法与传统搜索技术的最大区别在于不同的分支可以有不同的搜索深度，如图 2 所示。UCT 算法在不同的深度获取评估值。对于最有“希望”求解问题的分支，UCT 算法的搜索深度可以很深（远大于 d ），而对于“希望”不大的分支，其搜索深度可以很浅（远小于 d ）。当最有“希望”求解问题的分支数量远少于“希望”不大的分支数量时，UCT 算法就可以把搜索资源有效地用于最有“希望”求解问题的分支，从而获得比传统搜索算法更深的有效深度 d' 。这个具有神奇力量的“希望”是由树内选择策略计算的。

从根节点开始，对于每一个非叶子节点 n 的孩子 $n_i \in \text{ch}(n)$ ，树内选择策略计算一个评估值 v_i ，并根据 v_i 值选择一个孩子节点进行下一步选择，直到到达叶子节点。如果当前节点为最大节点（Max Node）时，树内选择策略选择 v_i 值最大的孩子进行下一步选择；如果当前节点为最小节点（Min Node）时，树内选择策略选择 v_i 值最小的孩子进行下一步选择。 v_i 的计算公式是：

$$v_i = v_i + c \cdot \sqrt{\frac{2 \ln(\sum_i T_i)}{T_i}}$$

式中： v_i 是以节点 n_i 为根节点的子树的所有仿真结果的平均值，反映了根据目前仿真结果观测到的节点 n_i 能提供的回报值的期。 T_i 是节点 n_i 的访问次数，也是节点 n_i 被树内选择策略选中的次数。 $\sum_i T_i$ 是节点 n 的访问次数。 c 是一个手工设定的常数。 c 的作用是平衡 UCT 算法的利用需求（exploitation）和探索需求（exploration）。

缺省仿真策略

当搜索到达叶子节点时，UCT 算法执行扩展操作（Expansion）：把此叶子节点允许的所有合法下一步产生的子节点，作为新的叶子节点加入到搜索树中，并正确初始化其 v 值和 T 值。UCT 算法并没有使用额

外的评估函数来获取新叶子节点的评估 v 值，而是使用缺省仿真策略来继续搜索直到游戏进入结束状态。此时，棋盘上每一个位置都有明确的归属，黑白赢还是白方赢可以很容易地计算出来，叶子结点的评估值就是当黑方胜时为 1，白方赢为 0。最简单的缺省仿真策略就是在所有的合法下一步中，均匀地随机选择下一步。用随机策略作为缺省仿真策略产生的程序棋力不高，因此大多数棋力不错的程序都采用了更加复杂的缺省仿真策略。这些实际使用的缺省仿真策略，考虑了气（liberty）、形（pattern）、定式（joseki）、攻击模式

（attack patterns）、防守模式（defense patterns）等一些重要的围棋基本概念。此外，为了获得更加可靠的评估值，每次到达叶子节点，UCT 算法可以使用 k 次仿真的平均结果作为评估 v 。 k 的值可以是几百或者几千。

仿真结果回传

当叶子节点通过仿真获得新的 v 值和 T 值时，UCT 算法通过结果回传更新搜索路径上的所有内部节点（internal nodes）的 v 值和 T 值。其公式为：

$$T_i = \sum_i T_i$$
$$v = \frac{\sum_i v_i T_i}{T}$$

即父节点的访问次数 T 是所有子节点访问次数 T_i 之和，观测回报值 v 是所有子节点观测回报值 v_i 的加权平均，权重为子节点被访问的比例 $\frac{T_i}{T}$ 。结果回传从叶子节点开始，沿搜索路径逐级向上更新，直到根节点。^[2]

算法优势

一、UCT 的工作模式是时间可控的

我们可以在算法执行过程中的任何时间突然终止算法，UCT 算法可以返回一个差不多理想的结果。当然如果给与更为充分的时间的话，算法结果会非常逼近实际的最优值。但是这一点在 alpha-beta 搜索中是绝对行不通的。图 3-3 展现了如果我们突然中断 alpha-beta 搜索程序，某些处于根节点之下的第一层的节点都还没有被计算，这样的结果此时 alpha-beta 搜索程序的返回结果很有可能和实际地最优解相去甚远。而图 3-2 是一个典型的中断时的 UCT 搜索树的情况，我们可以通过两幅图片的比较看出 UCT 算法的时间可控性的原因。当然，传统的搜索算法也有一些优化策略来增加其时间可控性，例如设置初始搜索深度。但不可否认的是，UCT 的任意时间终止特性更加强大，这可以使它应用于特定的博弈系统中的时候能够非常方便的控制走步产生的最大时间。这是传统的搜索算法所无法比拟的。

二、UCT 具有更好的鲁棒性

这是因为它使用一种平滑的方式处理搜索过程中的不确定性。在每个节点，其计算值取决于它的搜索节点序列上的所有子节点的计算值，其值是一个经过平滑的最大值的估计值。这样，由于每个子节点的计算过程都经过重新的抽样计算，不会因为个别严重偏离事实的抽样结果而对最终的结果产生致命性的影响。同时，由于算法在确定计算的节点序列时，依赖于第一层子节点的估值以及该估值的可信度。也就是说，如果某一个子节点有一个远大于其他子节点的估值，那么相对于其他节点，它就有机会更为频繁的被搜索和计算，UCT 算法对这个“最大值”节点会进行最为频繁的搜索。在非完备信息博弈问题的条件下，与蒙特卡罗抽样算法相结合的 UCT 算法中，频繁的搜索意味着频繁的抽样，而大量的抽样次数正式保证估计值能够尽量逼近真值的唯一途径。同时，如果有两个子节点拥有相近的估值并大于其他节点，那么 UCT 算法会均衡的对两个节点进行搜索。这样的搜索方式有一个优势，那就是 UCT 算法最后的作为搜索结果的节点以及次优节点一定是经过多次抽样的具有较高估值可信度的节点。

三、在 UCT 搜索算法的过程中，博弈树以一种非对称的形式动态扩展出来

这样做有两个好处。首先，传统的博弈树扩展方式，仍然以 alpha-beta 搜索树为例，每向下扩展一层都意味着博弈书规

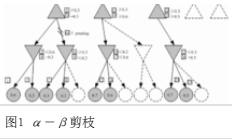


图1 α-β剪枝

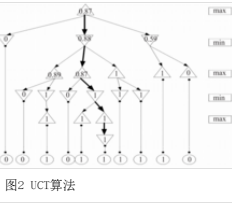


图2 UCT算法

分享

百科数字博物馆
国宝居然会说话？



模的指数型增长以及搜索时间的指数型增加。对于内存和 CPU 性能都有限的个人电脑来说，这一问题有的情况下是致命的。而在 UCT 算法搜索过程中，每次对于更深一层的扩展仅局限于搜索序列的最后一个节点。这样的 UCT 算法可以在扩展节点的同时不断的动态释放计算过的节点内存，使得算法运行的时间复杂性和空间复杂性可以被更好的控制。其次，正因为上述特性，对于较好的作为被选候补的节点，算法往往可以进行更为深入的搜索，这点可以在图一中体现出来，同时，这种非对称性扩展完全是在算法的执行过程中自动进行的。因此，和传统的博弈树算法相比较，UCT算法有着其独有的优势，特别是当博弈树规模非常大的时候。UCT算法首次应用的围棋博弈系统，以及本文即将讨论的四国军棋博弈系统都属此例。因此，UCT搜索算法在本系统中的使用是切合实际的。^[1]

应用

- 2006年第一个基于UCT算法的围棋程序MoGo在9×9棋盘上达到了专业棋手的水平。
- 2009年8月基于UCT的开源程序Fuego在9×9棋盘上战胜了周俊勋九段。
- 2012年6月计算机围棋程序Zwn19S在19×19棋盘上达到了KGS的6段评级。
- 2013年3月围棋软件CrazyStone在受让四子的情况下。战胜日本棋手石田芳夫九段，其棋力已达到业余五、六段的水平。

领域发展

深蓝——蛮算硬汉

- 1997年，美国超级计算机“深蓝”以2胜1负3平战胜了当时世界排名第一的国际象棋大师卡斯帕罗夫。“深蓝”的运算能力当时在全球超级计算机中居第259位，每秒可运算2亿步。
- 在今天看来，“深蓝”还算不上足够智能，主要依靠强大的计算能力穷举所有路数来选择最佳策略：“深蓝”靠硬算可以预判12步，卡斯帕罗夫可以预判10步，两者高下立现。
- 比赛中，第二局的完败让卡斯帕罗夫深受打击，他的斗志和体力在随后3局被拖垮，在决胜局中仅19步就宣布放弃。

浪潮天梭——以一敌五

- 2006年，中国超级计算机“浪潮天梭”同时迎战柳大华、张强、汪洋、徐天红、朴风波5位中国象棋特级大师。在2局制的博弈中，浪潮天梭以平均每步棋27秒的速度，每步66万亿次的棋位分析与检索能力，最终以11：9的总比分险胜。
- 张强坦承：“以往和人比赛，到了最后时刻就是意志和心态的对决了，看谁能坚持到最后，谁能不犯错误。但是计算机没有这样的问题。”
- 从那场比赛开始，象棋软件蓬勃发展，人类棋手逐渐难以与之抗衡。

沃森——全才学霸

- 2011年，“深蓝”的同门师弟“沃森”在美国老牌智力问答节目《危险边缘》中挑战两位人类冠军。虽然比赛时不能接入互联网搜索，但“沃森”存储了2亿页的数据。“沃森”可以在3秒内检索数百万条信息并以人类语言输出答案，还能分析题目线索中的微妙含义、讽刺口吻及谜语等。“沃森”还能根据比赛奖金的数额、自己比对手落后或领先的情况、自己擅长的题目领域来选择是否要抢答某一个问题。“沃森”最终轻松战胜两位人类冠军，展示出的自然语言理解能力一直是人工智能界的重点课题。^[3]

参考资料

1. 基于UCT算法的非完备信息多人军棋博弈系统 . 中国知网. 2008[引用日期2016-03-12]
2. 封闭域UCT算法的实现及其性能测试 . 万方数据. 2014-06[引用日期2016-03-12]
3. 那些年，那些人机大战 . 凤凰网. 2016-03-09[引用日期2016-03-12]

词条标签：科学

分享

猜你关注

- 生姜快速减肥法
咋样能快速减肥
- 快速减腹部
生日送啥礼物好
- 快速的减肥法
餐饮人脸属性分析技术
- 生日送什么东西好
学android要多久
- 快速减肥产品
电脑人脸识别锁



- 👤 新手上路
- 📖 我有疑问
- 🗉 投诉建议

- 成长任务

编辑规则
- 编辑入门

百科术语
- 我要质疑

参加讨论
- 在线客服

意见反馈
- 举报不良信息

投诉侵权信息
- 未通过词条申诉

封禁查询与解封

©2017 Baidu [使用百度前必读](#) | [百科协议](#) | [百度百科合作平台](#) | 京ICP证030173号 
[京公网安备11000002000001号](#)

分享











百科数字博物馆
国宝居然会说话？

