

模式的极大似然估计在围棋博弈中的应用

尚鹏飞, 李文峰, 刘知青

北京邮电大学软件学院, 北京 (100876)

E-mail: shangpf520@gmail.com

摘要: 提出一种基于模式的相对频率的极大似然估计方法来提高计算机围棋博弈水平。按照模式模型从大量的专家棋谱中抽取围棋模式, 通过计算每一个模式相对频率的极大似然估计从而生成围棋博弈模式库。还使用了 Katz 回退算法来解决模式的稀疏问题从而提高模式的利用率。实验结果证明了方法的有效性。

关键词: 模式; 相对频率; 极大似然估计; Katz 回退算法

中图分类号: TP181

1. 引言

计算机围棋博弈一直以来都是人工智能的重要分支, 虽然经过四十年的研究与发展已经取得了许多重大的进步, 其中包括^[1,2,4,7,6], 但是就目前的计算机围棋博弈水平来看, 计算机还不足以击败职业围棋选手。其主要的原因就是: 其他计算机棋类博弈的方法不适合围棋。其他计算机棋类博弈使用的方法是利用一个高效的算法来搜索博弈树并对当前的棋局进行评估。而对于围棋由于它的搜索空间过大所以搜索整个对弈树是不现实的, 并且要得到一个全局的评估函数很困难。目前比较流行的解决计算机围棋博弈的方法是带指导的蒙特卡罗算法。使用几个简单的规则例如: 不要填自己的眼, 不要落在棋盘角上等等来改进随机的蒙特卡罗算法^[11]。在蒙特卡罗算法中使用从专家棋谱中抽取的模式进行模拟, 所以模式的好坏也会影响对弈的结果。现在应用比较广泛的方法是使用机器学习算法来评估模式^[10]。本文也将采用基于模式的方法, 着重介绍使用相对频率的极大似然估计方法来分析模式, 从而提高计算机围棋的博弈水平。我们将在第二节讨论模式的定义和表示, 第三节介绍相对频率的极大似然估计以及在模式中的应用, 第四节解决数据稀疏问题, 第五节给出这些方法在实际对弈中的结果和分析以及以后的工作。

2. 模式的定义和表示

在围棋博弈中对弈双方的最终目的是利用棋子占据更多的区域, 所以棋子与其周围环境之间的联系的重要性要强于棋子在棋盘上具体位置的重要性。就像自然语言中的词一样, 通常关心词以及词出现的上下文环境而并不关心这个词出现的具体位置, 称这样的词法分析为上下文相关分析, 同样称这样的模式为上下文相关模式。对于给定的任意一个在棋谱上的棋子 M , 定义 M 的上下文相关模式 CP_M 为: 该棋子 M 信息的集合, 它包含了棋子 M 本身, 棋子 M 周围其他棋子以及棋子 M 周围棋盘的信息(一般为棋盘边界和空点组成)。显然这个定义是二维空间上的定义而不是经常在自然语言处理中使用的一维时间上的定义。虽然围棋对弈也具有一维时间上的顺序但二维空间上的信息要比在一维时间上的信息重要^[8], 这是因为许多围棋对弈中的落子顺序是可交换的。

二维空间上的模式是有一定范围的, 虽然通常这个范围并不是固定的并且可以是任意形状的, 但是固定范围和形状的模式更容易数字化, 更容易让计算机处理。把固定范围和形状并且不含任何棋谱信息的模式叫做模式模型^[9]并使用符号 $TP(x, y)$ 来表示, 其中 x 表示 $MAX(Hx, Hy)$, y 表示 $MIN(Hx, Hy)$, Hx 和 Hy 分别表示模式模型中的点到其中心点所在坐标的

x 轴距离和 y 轴距离。把 $TP(x, y)$ 叫做 x 乘 y 阶模式模型，显然高阶模式模型的所覆盖棋盘的范围比低阶模式模型的大。在本文所提及的实验中使用的都是如图 1 所示的模式模型来抽取模式的。

虽然不能直接评估这些模式并给每一个模式赋一个值来表示它的有效性，但是可以使用^[3]中所描述的方法从专家棋谱中抽取出这些模式并统计出每一个模式的频率，称该频率为模式的绝对频率。模式的绝对频率可以从统计学的角度衡量一个模式的好坏，有很多的方法都是基于模式的绝对频率进行分析并得到了一些很有启发性的结论如^[3,8]。

4,4	4,3	4,2	4,1	4,0	4,1	4,2	4,3	4,4
4,3	3,3	3,2	3,1	3,0	3,1	3,2	3,3	4,3
4,2	3,2	2,2	2,1	2,0	2,1	2,2	3,2	4,2
4,1	3,1	2,1	1,1	1,0	1,1	2,1	3,1	4,1
4,0	3,0	2,0	1,0	0,0	1,0	2,0	3,0	4,0
4,1	3,1	2,1	1,1	1,0	1,1	2,1	3,1	4,1
4,2	3,2	2,2	2,1	2,0	2,1	2,2	3,2	4,2
4,3	3,3	3,2	3,1	3,0	3,1	3,2	3,3	4,3
4,4	4,3	4,2	4,1	4,0	4,1	4,2	4,3	4,4

图 1 模式模型图，同一模式模型使用同一灰度表示，灰度大的模式模型包含灰度小的模式模型。

3. 相对频率的最大似然估计

绝对频率虽然有许多比较好的属性，但却很少使用^[3]。因为它只说明了模式的常用性，并没有说明模式的重要性。例如：如图 2 所示的 3*3 模式在使用的 16067 盘专家棋谱中一共出现了 39248 次也就是说当出现图 3 所示的情况时专家一共在锚点处(即模式中间的那个位置)落子 39248 次。但是我们想知道的是在图 3 所示的情况下落在锚点处是否要比落在棋盘的其他位置处都要好。所以使用模式的相对频率来表示这个概念。模式的相对频率值越大表示在当前的环境下落在锚点处就越好，模式的相对频率值越小表示在当前的环境下落在锚点处就越不好。

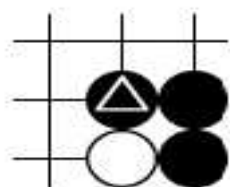


图 2 3*3 模式

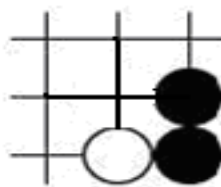


图 3 3*3 模式去掉锚点

定义模式的相对频率 $PR(CP_M)$ 为：

$$P_R(CP_M) = \frac{C_A(CP_M)}{C_N(CP_M)} \quad (1)$$

其中 $C_A(CP_M)$ 表示：模式 CP_M 的绝对频率； $C_N(CP_M)$ 表示：模式 CP_M 的所有可能出现的频率，它是使用如图 4 所示的算法来统计模式 CP_M 的 $C_N(CP_M)$ 。当算法是基于一定数量的训练数据时，得到的就是相对频率的最大似然估计，它选择的参数值对于训练数据给出了最高的概率^[5]。也就是说在所有可能出现的情况中，已经出现的情况具有最大的概率。

```

for each move M' in a game record loop
  for each playable move M on board loop
    Play M on board;
    if CP_M not in database then
      Insert CP_M into database;
      C_N(CP_M) := 1;
    else
      C_N(CP_M) ++;
    end if;
    Remove M on board;
  end loop;
  Play M' on board;
end loop;

```

图 4 计算 $C_N(CP_M)$ 的算法

4. 数据稀疏问题

实验中抽取到的模式数量如下表 1 所示，随着模式模型范围的增大，模式数量也急剧地增加。但是对于计算机围棋博弈在使用这些模式，特别是使用大范围的模式模型抽取出来的模式时，就会出现数据稀疏问题。

表 1 模式的绝对频率和所有可能出现的频率

	$C_A(CP_M)$	$C_N(CP_M)$
(1,0)	48	48
(1,1)	2,839	3,518
(2,0)	102,778	190,077
(2,1)	1,135,299	11,481,703
(2,2)	1,508,482	20,952,691
(3,0)	1,737,229	23,473,259
(3,1)	3,410,931	64,798,472
(3,2)	3,993,697	137,876,342
(3,3)	2,555,283	6,763,923,765

为了解决这个问题我们引入 Katz 回退算法^[5]，该算法的主要思想是：在进行模式的匹配时若高阶模式模型不能匹配就回退到低阶模式模型上去，该算法表述如下：

$$P_{(TP(x, y))} = \begin{cases} (1 - d) P_{(TP(x, y))} & P_{(TP(x, y))} > K \\ \alpha P_{(TP(x, y-1))} & P_{(TP(x, y))} \leq K \text{ 且 } y \geq 1 \\ \alpha P_{(TP(x-1, y+x-1))} & P_{(TP(x, y))} \leq K \text{ 且 } y < 1 \\ 0 & x = 0 \end{cases} \quad (2)$$

其中的 d 表示模式的折扣，即将一些概率转移到低阶模式上去。 α 表示正规化因子，以确保在折扣中预留的概率会分配到以回退方法估计的模式中去。表示模式频率的阈值。在我们的实验中取 $d = 0, \alpha = 1, k = 0$ 。

5. 实验及结果分析

实验中只使用了 1 乘 1 阶和 2 乘 2 阶模式，并且在对弈时只选取模式的相对概率的最大似然估计值在阈值以上的模式(该阈值是由围棋专家设定的)即通常所说的好点。在 9 路棋谱上与 gnugo3.6 对弈的结果如表 2 所示：

表 2 Lingo(即我们的引擎)和 Mogo(现世界最强棋)分别与 gnugo3.6 对弈的胜率。

	未使用模式	1 乘 1 阶模式	2 乘 2 阶模式	Katz 回退算法
Lingo	34.6%	52.8%	48.7%	60.6%
Mogo	38.2%	58.4%	50.8%	68.2%

对对弈结果的分析我们可以得到以下结论：使用模式的相对频率的最大似然估计能显著提高计算机围棋的博弈水平，在使用单一模式的情况下低阶模式比高阶模式效果要好，这主要的原因是高阶模式的数据稀疏造成的。在使用 Katz 回退算法将低阶模式和高阶模式结合起来使用的情况下计算机围棋的博弈水平要比使用单一低阶模式的情况下要高。

由于在实验中只是使用了 1 乘 1 阶和 2 乘 2 阶模式，在以后的工作中我们会提取更多的不同阶模式进行实验，另外还可以再定义另外一个阈值（同样由专家设定），在进行对弈前将相对概率的最大似然估计值在该值以下的模式即通常所说的坏点直接去掉，这样就可以缩小计算机围棋博弈时的搜索空间。

6. 总结

本文通过使用模式的相对频率的最大似然估计方法来提高计算机围棋博弈水平，并且使用 Katz 回退算法来解决模式的稀疏问题，并通过实验证明了该方法的有效性。在今后的工作中我们会去尝试使用其他方法来度量模式的好坏，从而进一步提高计算机围棋博弈水平，为人工智能领域做出贡献。

参考文献

- [1] D.B.Benson. Life in the game of GO [J]. Information Sciences, 1976, 10: 203-213.
- [2] E.Berlekamp and D. Wolfe. Mathematical Go End-games [J]. Nightmares for the Professional Go Player, 1994, 3: 78-84.
- [3] Zhiqing Liu. Automatic Acquisition of Contextual Patterns and Collocations in the Game of GO [J]. in the Proceedings of the 12th Annual Conference of the Chinese Association of Artificial Intelligence, 2008, 7: 32-52.
- [4] M. Boon. Pattern Matcher for Goliath [J]. Computers GO, 1998, 13: 89-90.
- [5] Christopher D and Manning Hinrich Sch tze. Foundations of Statistical Natural Language Processing [M] . U.S.A.: Publishing House of Electronics Industry, 2005.
- [6] T. Cazenave. Generation of patterns with external conditions for the game of Go [J]. Advances in Computer Games Conference, 1999, 10: 77-89.
- [7] K. Chen and Z. Chen. Static analysis of life and death in the game of Go [J]. Information Sciences, 1999, 121: 102-204.
- [8] Zhiqing Liu Qing Dou and Benjie Lu. Frequency Distribution of Contextual Patterns in the Game of GO [J]. accepted and appear in the Proceedings of the 6th International Conference on Computers and Games, Springer Lecture Notes in Computer Science, 2008, 5313: 20-46.
- [9] David Stern Ralf Herbrich and Thore Graepel. Bayesian pattern ranking for move prediction in the game of Go [J]. In Proceedings of the 23rd International Conference on Machine Learning, 2006, 173: 102-205.
- [10] Rémi Coulom. Computing “Elo Ratings” of Move Patterns in the Game of Go [J]. ICGA, 2007, 31(2): 198-208.
- [11] Sylvain Gelly Yizao Wang Rémi Munos and Olivier Teytaud. Modification of UCT with patterns in Monte-Carlo Go [J]. Technical Report RR-6062, 2006, 32: 30-56.

The Application of Patterns' MLE in the Game of Go

Pengfei Shang, Wenfeng Li, Zhiqing Liu

School of Software Engineering, Beijing University of Posts and Telecommunications, Beijing
(100876)

Abstract

This paper based on the relative frequency's MLE of patterns to improve the level of the computer in the game of Go. Patterns are extracted from the game records of experts in accordance with the pattern's templates, by calculating the relative frequency's MLE of each patterns in order to generate the patterns' library of the game of Go. Katz's backing-off algorithm is used to solve the sparse of patterns to enhance the utilization of the patterns.

Keywords: pattern; relative frequency; MLE; Katz's backing-off algorithm

作者简介：尚鹏飞，男，1981 年生，硕士研究生，主要研究方向是人工智能，自然语言处理和机器学习。