



# 基于深度强化学习和蒙特卡洛搜索树的连六棋 AI

江俊广<sup>1</sup>, 陈宇韶<sup>1</sup> and 郑洛成<sup>1</sup>

<sup>1</sup>Tsinghua University

## Summary

**Background** 我们一开始打算设计五子棋的 AI。然而五子棋本身规则具有一定缺陷。首先，实际比赛中五子棋先后手优势差距比较大。先手往往能够决定胜负。所以历史上各国的棋手们采取一系列方法限制先手的优势来达到平衡先后手的目的。但仍然不能够很好解决问题。同时容易导致规则过于复杂。Github 上的 AlphaZero-Gomoku 是基于无禁手规则五子棋 AI。但是在实际运行中出现的的问题是，后手人类比较难战胜先手 AI。后手 AI 也比较难战胜先手人类。所以我们考虑更改五子棋为连六棋，即达成六子获胜。特黑者第一千放一个黑子于棋盘上。之后双方轮流落子。每手放二个棋子于棋盘上。这样可以保证游戏的公平性。因为各方每次下完一手后，盘面都比对方多一子。因此开局可自然达成平衡的状态。这使得公平性大为提升。另外六子棋不具有先手优势。所以也不需要为了保持公平性。而制定一些额外的规则。理论上游戏所使用的棋盘可以是无限大的。对一般玩家而言，采用围棋的十九路棋盘即可。

## Methods

### 1 规则的改变和公平性的保证

连六棋改变了游戏规则。除了将获胜条件由五连子改为六连子之外，还有第二手及之后每一手投两个子。这样可以保证游戏的公平性。因为各方每次下完一手后，盘面都比对方多一子。因此开局可自然达成平衡的状态。这使得公平性大为提升。另外六子棋不具有先手优势。所以也不需要为了保持公平性。而制定一些额外的规则。

### 2 蒙特卡洛搜索树的调整

**调整原因** 连六棋每一回合可以走两个棋子。因此走法复杂度是相同规模的五子棋的走法复杂度的平方。尽管搜索树的深度变成了原先的一半。更大的问题是神经网络的输出表示变得比较复杂。原先用棋盘大小的概率矩阵表示。现在必须用 棋盘大小<sup>2</sup> 的概率矩阵表示。

**解决方法** 采用贪心决策。即每一回合的每步都选尽可能是当前局面最优的棋子。AI 在棋盘上选择一个棋子后。基于新的棋盘再选择一个棋子。尽管不能保证两步结合后一定是最优决策。但是不会太差。同时可以使得走法复杂度降低一个平方数量级。同时使得神经网络的输出较为简洁。

**具体修改** 修改蒙特卡洛搜索树的树节点。原先是父子结点代表两个对手。现在是爷爷和孙子节点代表两个对手。在从叶节点向上传递局面评估值的时候需要判断何时取反。

### 3 神经网络输入的调整——棋盘状态的表示

AlphaGo Zero 中，一共使用了 17 个二值特征平面来描述当前局面。其中前 16 个平面描述了最近 8 步对应的双方 player 的棋子位置。最后一个平面描述当前 player 对应的棋子颜色，即先后手

五子棋用“4 个的二值特征平面”表示。其中前两个平面分别表示当前 player 的棋子位置和对手 player 的棋子位置。有棋子的位置是 1。没棋子的位置是 0。然后第三个平面表示对手 player 最近一步的落子位置，也就是整个平面只有一个位置是 1，其余全部是 0。第四个平面，也就是最后一个平面表示的是当前 player 是不是先手 player。如果是先手 player 则整个平面全部为 1，否则全部为 0。

连六棋用 4 个二值特征平面表示。前两个平面分别表示当前 player 的棋子位置和对手 player 的棋子位置。有棋子的位置是 1。没棋子的位置是 0。第三个平面表示对手 player 最近一回合下的所有棋子。整个平面最多 2 个 1。其余全部是 0。第四个平面表示当前 player 在该回合已经下的棋子。整个平面最多 1 个 1。其余全部是 0。由于连六棋中，先后手并没有什么太大的区别。因此去除了五子棋中表示先后手的平面。

**Results** Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

**Conclusion** Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

## Figures & Tables

表 1: Caption for table one		
Heading1	Heading2	Heading3
Row1	0.1	.01
Row2	0.2	.02
Row3	0.3	0.03
Row4	0.4	0.04

图 1: Caption for figure one.

图 2: Caption for figure Two.