

重力四子棋 实验报告

计 76 陈之杨 2017011377

1 实验内容

实现一个重力四子棋的对弈策略。

2 实现算法

笔者实现了基于蒙特卡洛方法的信心上限树搜索算法。对于待决策状态，以之为根建立信心上限树，树中每一个结点的子节点对应某一种落子方案的后续状态。容易发现，深度为奇数的结点对应我方决策的状态，深度为偶数的结点对应对方决策的状态。

对于每一次搜索，从根节点出发，如果有可拓展的结点，则选择新拓展的结点，否则按照信心评分选择最高的子节点。如此重复直至到达一个叶子结点。评分规则为：

$$\frac{Q(v)}{N(v)} + c \sqrt{\frac{2 \ln(N(v_f))}{N(v)}}.$$

其中， v 为评分结点， v_f 为 v 的父亲结点。 $N(v)$ 为 v 通过蒙特卡洛方法模拟的次数， $Q(v)$ 为 v 在模拟中的得分（胜计 1 分，负计 -1 分，平计 0 分）， c 为可调参数。需要注意的是，相邻两层结点的决策者是不同的，故评分时 $Q(v)$ 对于 v_f 应当取相反数处理。

对于选中的叶子结点，使用蒙特卡洛方法随机模拟双方的落子过程直至结束。用评分更新该叶子结点至根结点路径上的所有结点。同样地，相邻层结点的评分应当互为相反数。

重复这样的搜索过程直至接近时限，选择根节点评分最高的孩子结点作为当前状态的决策。注意此时评分规则里的 c 应当置 0 处理，因为之后不会再进行模拟。

3 测试结果

笔者与 100.d11 进行了 20 轮对战测试，最终胜率为 50%。