



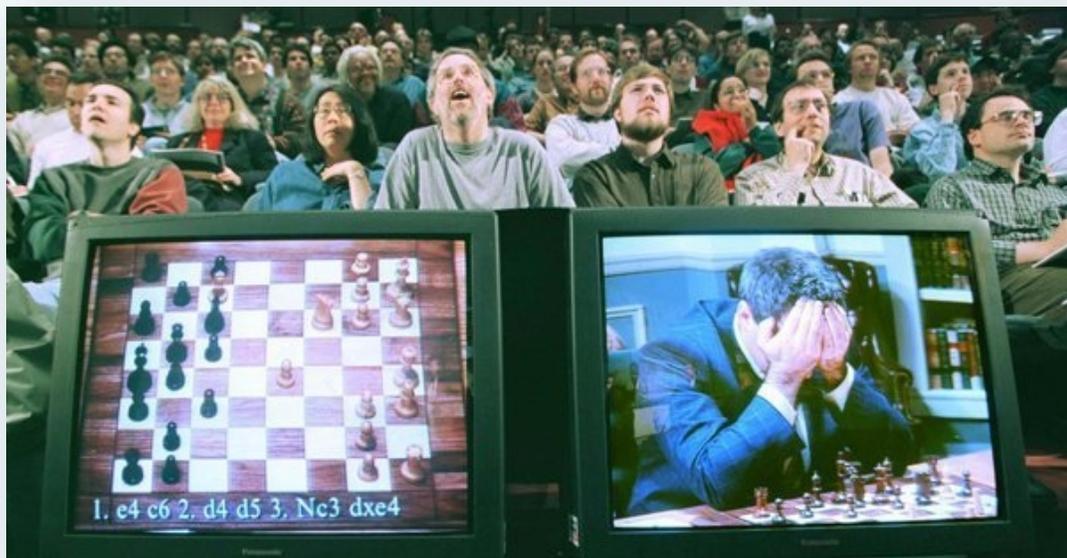
人工智能与信息社会

基于决策树和搜索的智能系统：从国际象棋到围棋

陈斌 北京大学 gischen@pku.edu.cn

深蓝的算法

- › AlphaBeta剪枝
- › 残局库（增加搜索深度）
- › 人类对局开局库



AlphaBeta之于围棋

> GNU Go

> 业余5~10级左右

> CGOS上, GNU Go被当作基准分数

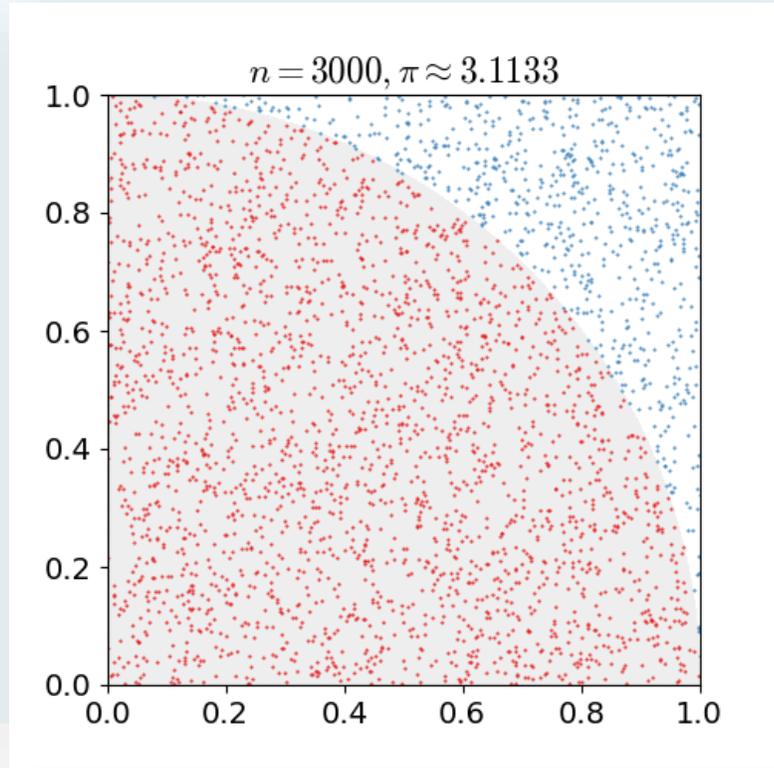
```
White (0) has captured 0 pieces
Black (X) has captured 0 pieces

  A B C D E F G H J K L M N O P Q R S T
19 . . . . . . . . . . . . . . . . . . 19
18 . . . . . . . . . . . . . . . . . . 18
17 . . . . . . . . . . . . . . . . . . 17
16 . . . + . . . . . + . . . . . + . . . 16
15 . . . . . . . . . . . . . . . . . . 15
14 . . . . . . . . . . . . . . . . . . 14
13 . . . . . . . . . . . . . . . . . . 13
12 . . . . . . . . . . . . . . . . . . 12
11 . . . . . . . . . . . . . . . . . . 11
10 . . . + . . . . . + . . . . . + . . . 10
 9 . . . . . . . . . . . . . . . . . . 9
 8 . . . . . . . . . . . . . . . . . . 8
 7 . . . . . . . . . . . . . . . . . . 7
 6 . . . . . . . . . . . . . . . . . . 6
 5 . . . . . . . . . . . . . . . . . . 5
 4 . . . + . . . . . + . . . . . + . . . 4
 3 . . . . . . . . . . . . . . . . . . 3
 2 . . . . . . . . . . . . . . . . . . 2
 1 . . . . . . . . . . . . . . . . . . 1
  A B C D E F G H J K L M N O P Q R S T

black(1): ■
```

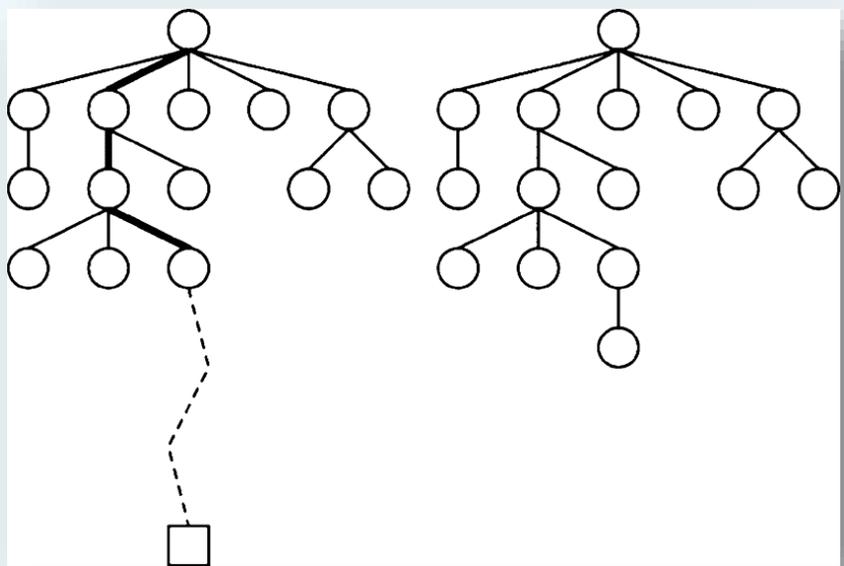
蒙特卡洛方法

- › 通过随机采样计算得到近似结果
- › 一种通用的计算方法



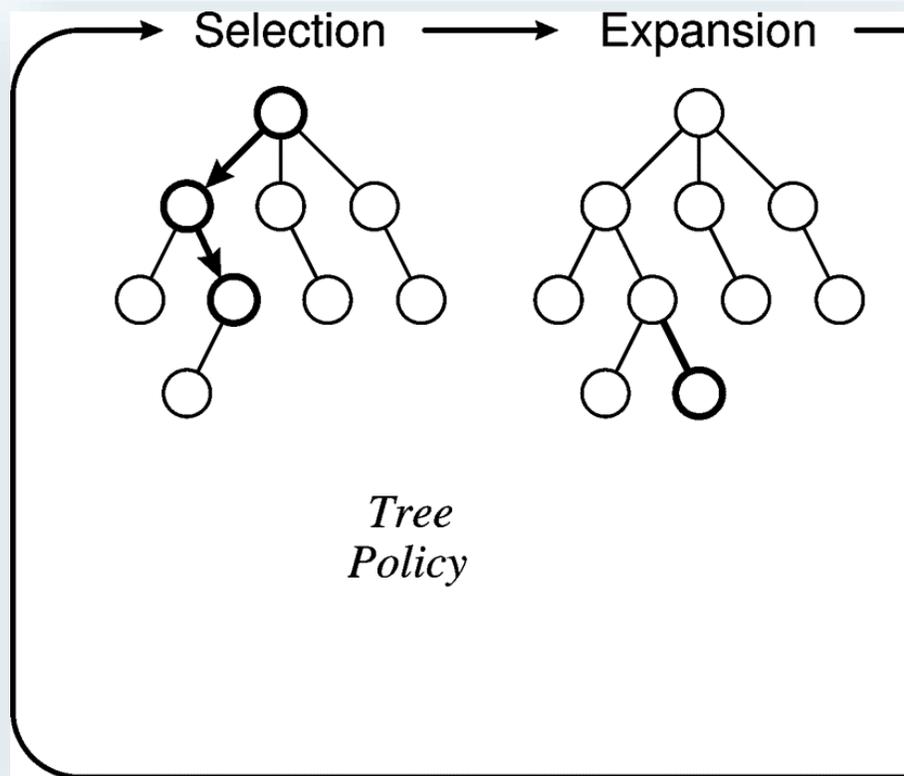
蒙特卡洛树搜索 (MCTS)

- › 一种通过在决策空间中随机采样并根据结果构建决策树来寻找最优策略的方法
- › 决策树的构建：选择、扩张、模拟、反馈



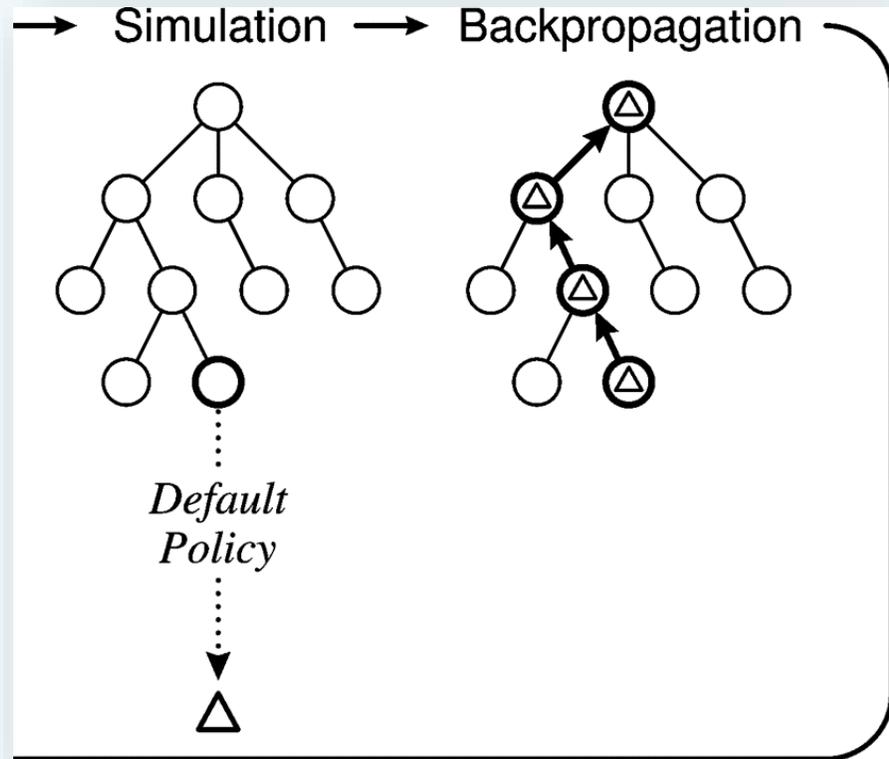
选择、扩张

- 树策略：从决策树中选择并创建新的叶节点

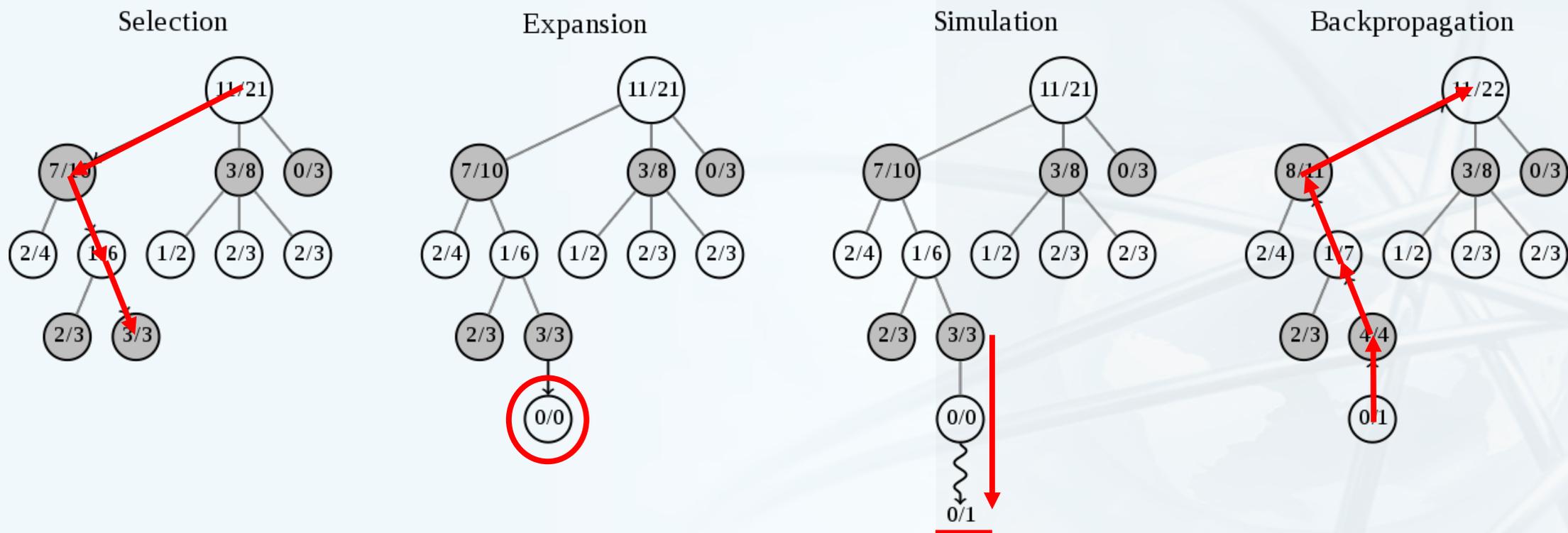


模拟、反馈

- 默认策略：从一个非中止状态不断进行游戏并得到一个价值评估



MCTS例子



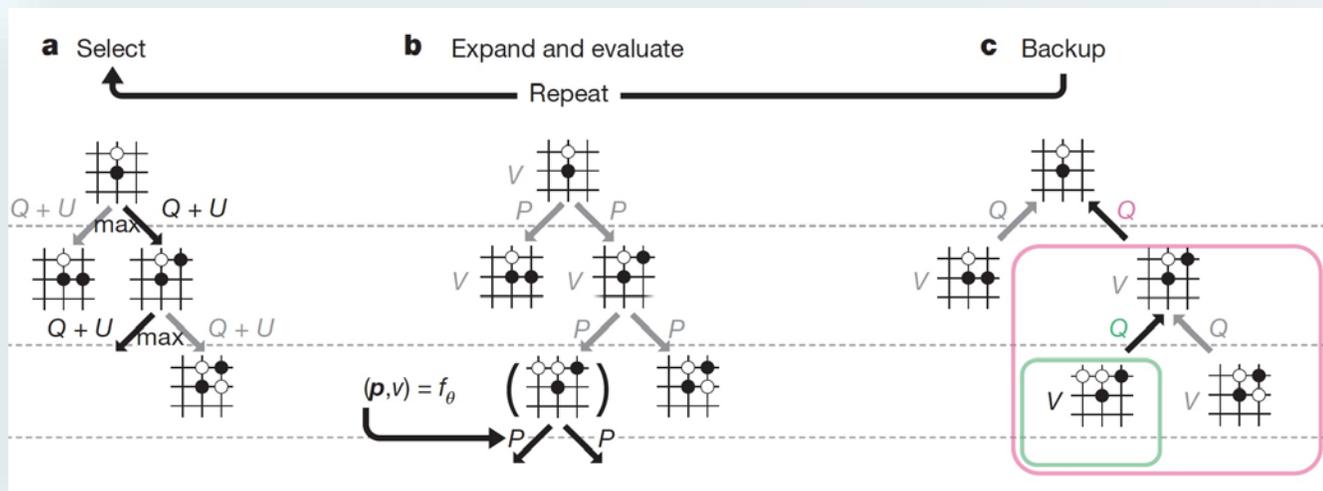
MCTS之于围棋

- › MoGo 第一个使用使用蒙特卡洛树搜索的围棋程序（2006年），在9×9的棋盘上击败了职业选手
- › DeepZenGo是AlphaGo之前最强的围棋程序之一，可以达到与职业棋士差距3~4子的水平



AlphaGo

- 在MCTS的基础上，通过神经网络进行训练得到更好的树策略和估值函数



结果比较

软件或人类	◆ BayesElo ◆
AlphaGo Zero (40 blocks版)	5422?
AlphaGo (Master版)	5231?
AlphaGo Zero (20 blocks版)	5022?
AlphaGo (Lee版)	4672?
朴廷桓	4592?
柯洁	4590?
井山裕太	4546?
李世石	4514?
DeepZenGo	4269
AlphaGo (Fan版, 176 GPU)	4122?
AlphaGo (Fan版, 48 CPU与8 GPU)	3862?
GNU Go	1800