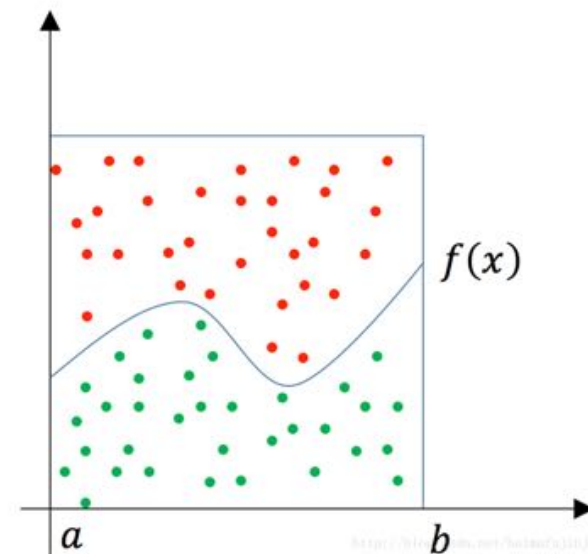
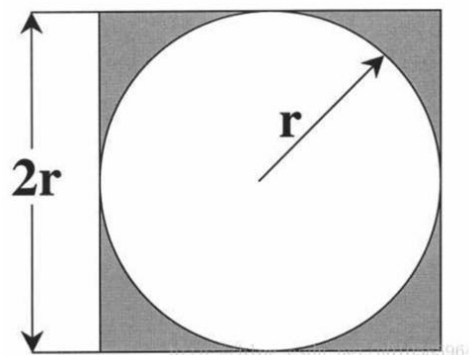


蒙特卡罗树搜索算法

赵耀

蒙特卡罗？

- ▶ 是指使用随机数（或更常见的伪随机数）来解决很多计算问题的方法。
- ▶ 蒙特卡罗是人名吗？ NO -----其实是一座很有名的赌城的名字
- ▶ 蒙特卡罗的经典应用
 - 投针实验求圆周率 π
 - 数学中求积分
- ▶ 蒙特卡罗的发展得益于计算能力的提升



适用场景

- ▶ 有些模型含有不确定的随机因素
- ▶ 有些模型难以定量，或者定量的代价太大

应用的步骤或过程

- ▶ 问题可以用概率表达，或者能用概率统计出某种特征，或是一种参数，这个特征或参数有助于获取问题的解
- ▶ 利用随机数或者伪随机数模拟一个过程，通过多次的实验，统计某个事件发生的概率。实验次数越多，统计值越准确。

蒙特卡罗方法应用于棋类游戏

- ▶ Alpha-Beta剪枝用于棋类时，由于时间限制，递归的深度有限；如要完成所有分支，搜索深度太深，时间复杂度太大；如果递归的深度浅，则评估不准确。且有搜索宽度的问题（可落子的范围大）。
- ▶ 蒙特卡罗的解决方式：下一步棋如何效果评估？多次模拟后续的下棋步骤，模拟时双方都随机下，直到决出胜负。比如模拟1000次，黑棋胜800次，白棋胜200次，则统计出此步棋（导致的当前棋局）黑棋胜率为0.8，白棋胜率为0.2，黑棋占优。

蒙特卡洛方法与蒙特卡洛树的区别

- ▶ 蒙特卡洛树搜索—博弈的思维

算法的基本步骤

► 具体步骤概括如下：

- 1.由当前局面建立根节点，生成根节点的全部子节点，分别进行模拟对局；
- 2.从根节点开始，进行最佳优先搜索；
- 3.选择最大胜率的子节点（改进：用UCB 公式计算每个子节点的 UCB 值，选择最大值的子节点）；
- 4.若此节点不是叶节点，则以此节点作为根节点，重复 2；
- 5.直到遇到叶节点，如果叶节点未曾经被模拟对局过，对这个叶节点模拟对局；否则为这个叶节点随机生成子节点，并进行模拟对局；
- 6.将模拟对局的收益（一般胜为 1 负为 0）按对应颜色更新该节点及各级祖先节点，同时增加该节点以上所有节点的访问次数；
- 7.回到 2，除非此轮搜索时间结束或者达到预设循环次数；
- 8.从当前局面的子节点中挑选平均收益最高的给出最佳着法。

UCT公式

- ▶ 蒙特卡洛树搜索通过迭代来一步步地扩展博弈树的规模，UCT 树是不对称生长的，其生长顺序也是不能预知的。它是根据子节点的性能指标导引扩展的方向，这一性能指标便是 UCB 值。它表示在搜索过程中既要充分利用已有的知识，给胜率高的节点更多的机会，又要考虑探索那些暂时胜率不高的兄弟节点，这种对于“利用”（Exploitation）和“探索”（Exploration）进行权衡的关系便体现在 UCT 着法选择函数的定义上，即子节点 N_i 的 UCB 值按如下公式计算：

$$\frac{W_i}{N_i} + \sqrt{\frac{C \times \ln N}{N_i}}$$

其中：

W_i ：子节点获胜的次数；

N_i ：子节点参与模拟的次数；

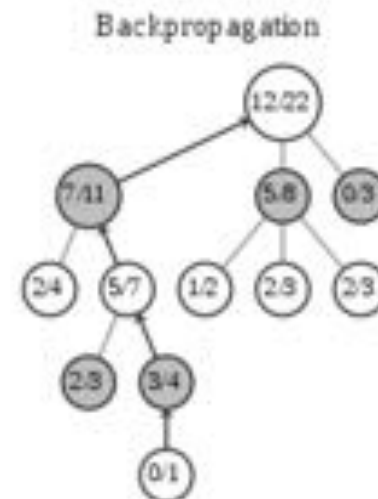
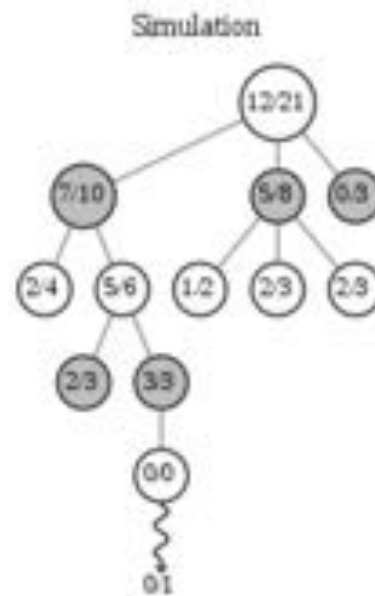
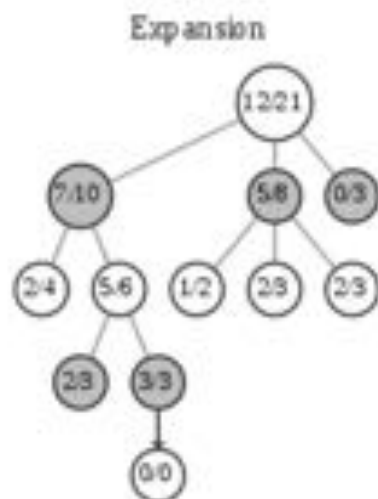
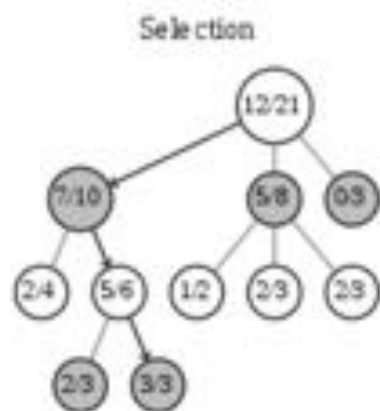
N ：当前节点参与模拟的次数

C ：加权系数。

可见 UCB 公式由两部分组成，其中前一部分就是对已有知识的利用，而后一部分则是对未充分模拟节点的探索。 C 小偏重利用；而 C 大则重视探索。需要通过实验设定参数来控制访问节点的次数和扩展节点的阈值。

源于《计算机博弈原理与方法学概述》

一个网络常见图



UCB值计算:

当前根节点N = 21

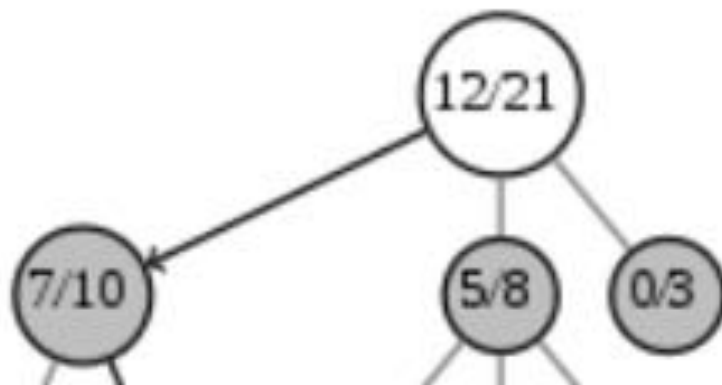
子节点1: $W_i = 7$ $N_i = 10$

当 $c = 10$

则子节点1: $UCB = 7/10 + \sqrt{10 * \log(21)/10} = 2.44$

子节点2: $UCB = 5/8 + \sqrt{10 * \log(21)/8} = 2.57$

子节点3: $UCB = 0/3 + \sqrt{10 * \log(21)/3} = 3.18$



当 $c = 1$

则子节点1: $UCB = 1.25$

子节点2: $UCB = 1.24$

子节点3: $UCB = 1.00$

c 值越大越偏向低胜率（所谓探索）

c 值越小越偏向高胜率（所谓利用）

反方节点

UCB值计算:

当前根节点 $N = 10$

子节点1: $W_i = 2$ $N_i = 4$

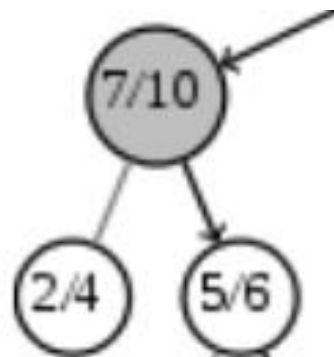
当 $c = 10$

则子节点1: $UCB = (1 - 2/4)$

$+ \sqrt{10 * \log(10)/4} = 2.89$

子节点2: $UCB = (1 - 5/6) +$

$\sqrt{10 * \log(10)/6} = 2.12$



当 $c = 1$

则子节点1: $UCB = 1.25$

子节点2: $UCB = 0.78$

反方节点

UCB值计算:

当前根节点N = 20

子节点1: $W_i = 7$ $N_i = 14$

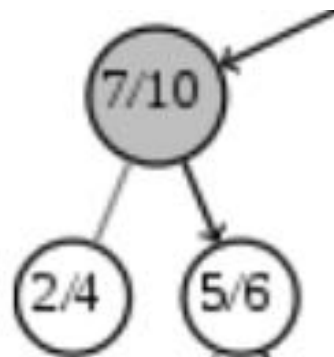
当 $c = 10$

则子节点1: $UCB = (1 - 2/4)$

$+ \sqrt{10 * \log(20)/14} = 1.96$

子节点2: $UCB = (1 - 5/6) +$

$\sqrt{10 * \log(20)/6} = 2.40$



当 $c = 1$

则子节点1: $UCB = 0.96$

子节点2: $UCB = 0.87$

随着模拟次数的增多, 子节点2
逐步获得模拟的机会

蒙特卡罗在五子棋上的模拟

- ▶ 注意：不能纯模拟，还是得能识别出活三冲四等棋局。权衡局势或结合棋局评估函数使用。

Gold: 120
Silver: 110
Bronze: 105

Base = Pass number – 3

For example, if 160 students pass the usability test and you get the 110th, you will get 76 points.

$$(110-3)/(160-3) = 68\% \quad 60\% < 68\% < 70\%$$

Others:

[illegible]