人工智能

- 博弈与搜索

马　超
武 汉 大 学

chaoma@whu.edu.cn

提 纲

- 博弈论
  - 博弈论概述
  - 博弈论发展历史
  - 博弈问题分类
  - 博弈问题实例
- 搜索策略
  - 极小极大搜索算法
  - α-β搜索法

博弈论（Game Theory）：概述

我们是我们无法了解的事物。

——美国作家爱默生皮尤

博弈论

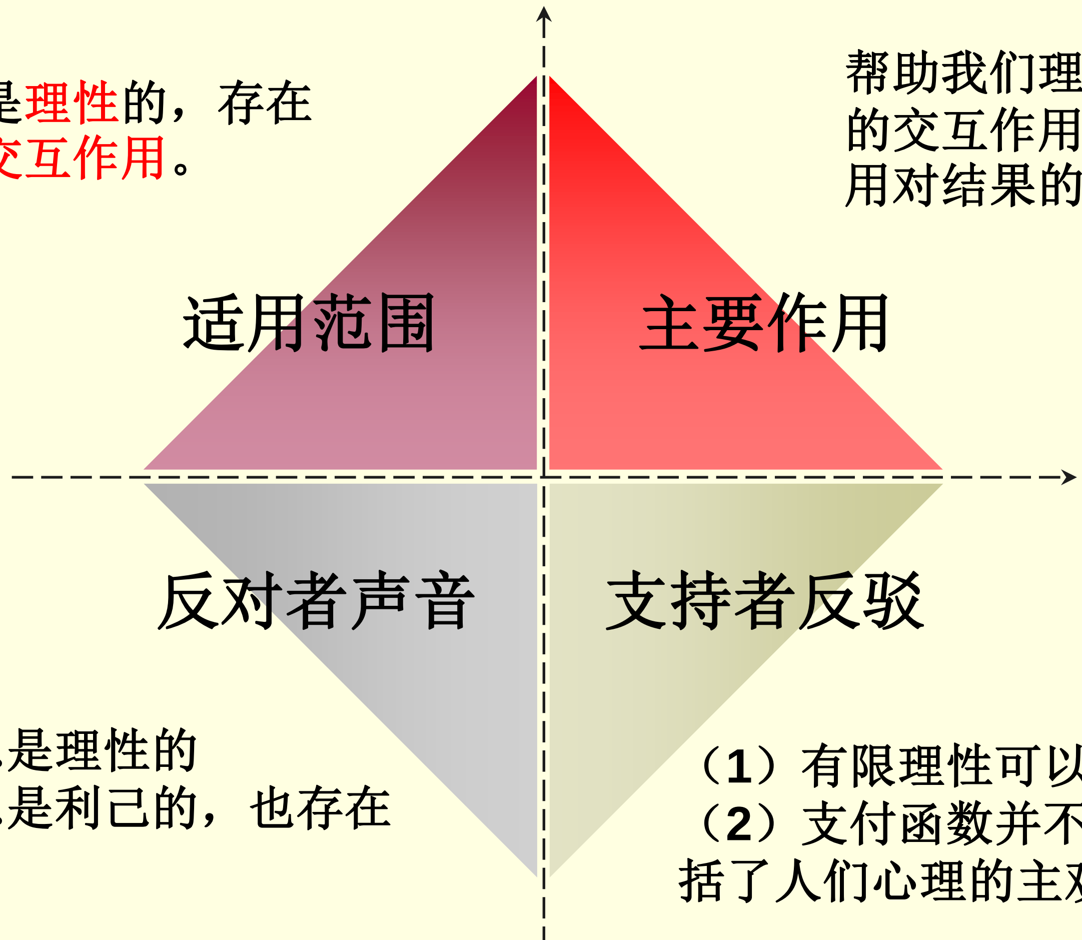
模型：它提供了一个研究复杂社会现象的框架，且将其简化为可供分析的模型。

方法：它提供了从这些与人们如何采取行动及人们该如何采取行动相关的模型中，总结命题的方法。

博弈论（Game Theory）：概述

参与人是理性的，存在行为的交互作用。

帮助我们理解人们行为的交互作用，及这种作用对结果的影响。



（1）人们并不总是理性的
（2）人们并不总是利己的，也存在利他行为

（1）有限理性可以被引入博弈模型中
（2）支付函数并不只是以金钱衡量，包括了人们心理的主观因素

博弈论提供了对于人们行为背后逻辑的洞察，还处于发展中。

博弈论（Game Theory）：概述

实际模型

- 实际模型旨在呈现真实世界，是对真实世界的准确描述，实际模型可能很复杂。

虚拟模型

- 虚拟模型是对真实模型的简化，并不一定是对真实模型的准确描述。我们使用虚拟模型尽可能从本质上模仿真实世界，并非复制真实世界。

构建虚拟模型

- 决定将哪些因素加入到虚拟模型中，就如同打包行李去大学，你没有办法把所有需要的东西都塞进行李箱。
- 从本质上模仿，就是把最关键的因素加入模型，忽略次要因素。有些时候，忽略次要因素更有助于研究关键因素之间的关系。

提 纲

- 博弈论
  - 博弈论概述
  - 博弈论发展历史
  - 博弈问题分类
  - 博弈问题实例
- 搜索策略
  - 极小极大搜索算法
  - α-β搜索法

博弈论（Game Theory）：发展历史

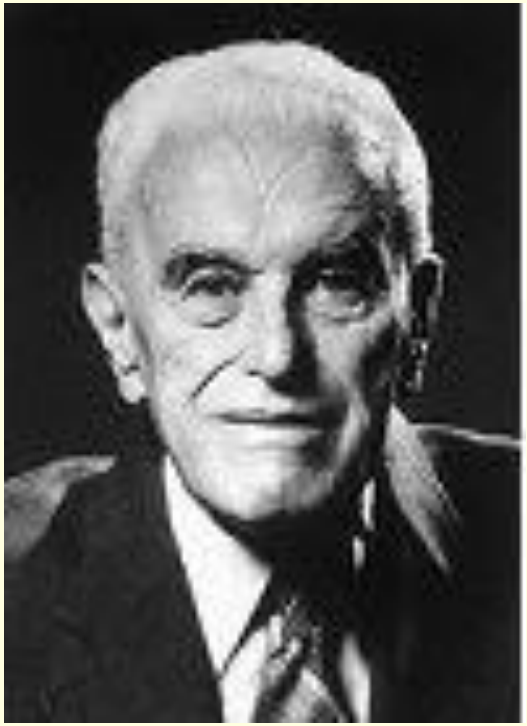
- 1838年库诺特（Cournot）寡头竞争模型
- 1883年伯川德（Bertrand）寡头竞争模型
- 1944年冯-诺依曼和摩根斯坦发表
  - 《博弈论和经济行为》
- 1950年纳什（Nash）提出了纳什均衡的概念
  - “Equilibrium points in n-person games”
  - “Non-cooperative games”
  - “The bargaining problem”
- 1965年泽尔腾（Selten）提出了子博弈精炼纳什均衡的概念

博弈论（Game Theory）：发展历史

- 1967—1968年海萨尼（Harsanyi）
  - 贝叶斯纳什均衡的概念
- 1975—1991年泽尔腾（1975）、Kreps和Wilson（1982）、Fudenberg和Tirole（1991）
  - 精炼贝叶斯纳什均衡的概念
- 1994年纳什、海萨尼和泽尔腾
  - 诺贝尔经济学奖

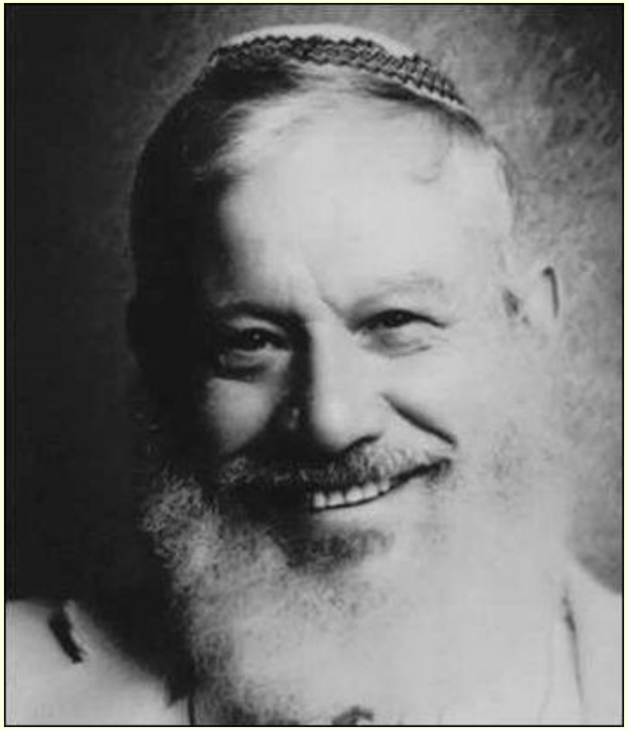
博弈论（Game Theory）：发展历史

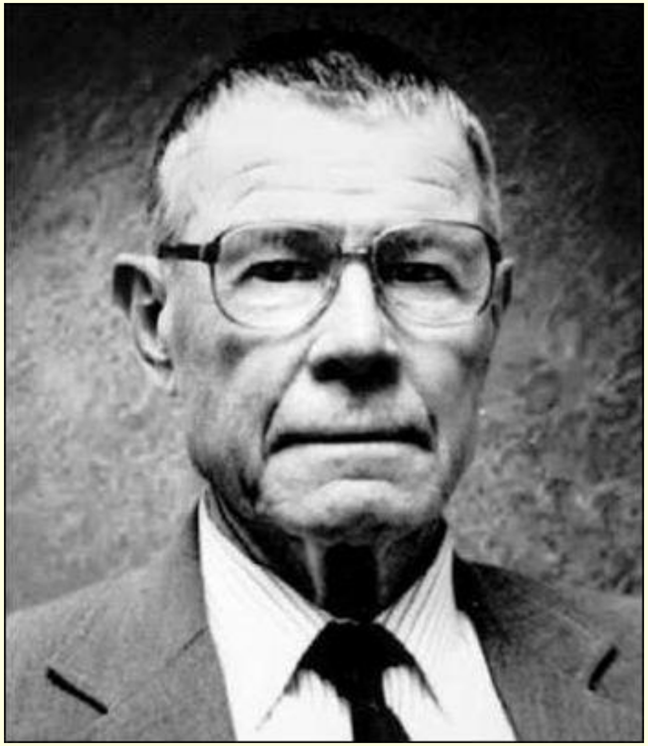
- 1994年纳什、海萨尼和泽尔腾获诺贝尔经济学奖


博弈论（Game Theory）：发展历史

- 2005年奥曼和谢林诺贝尔经济学奖





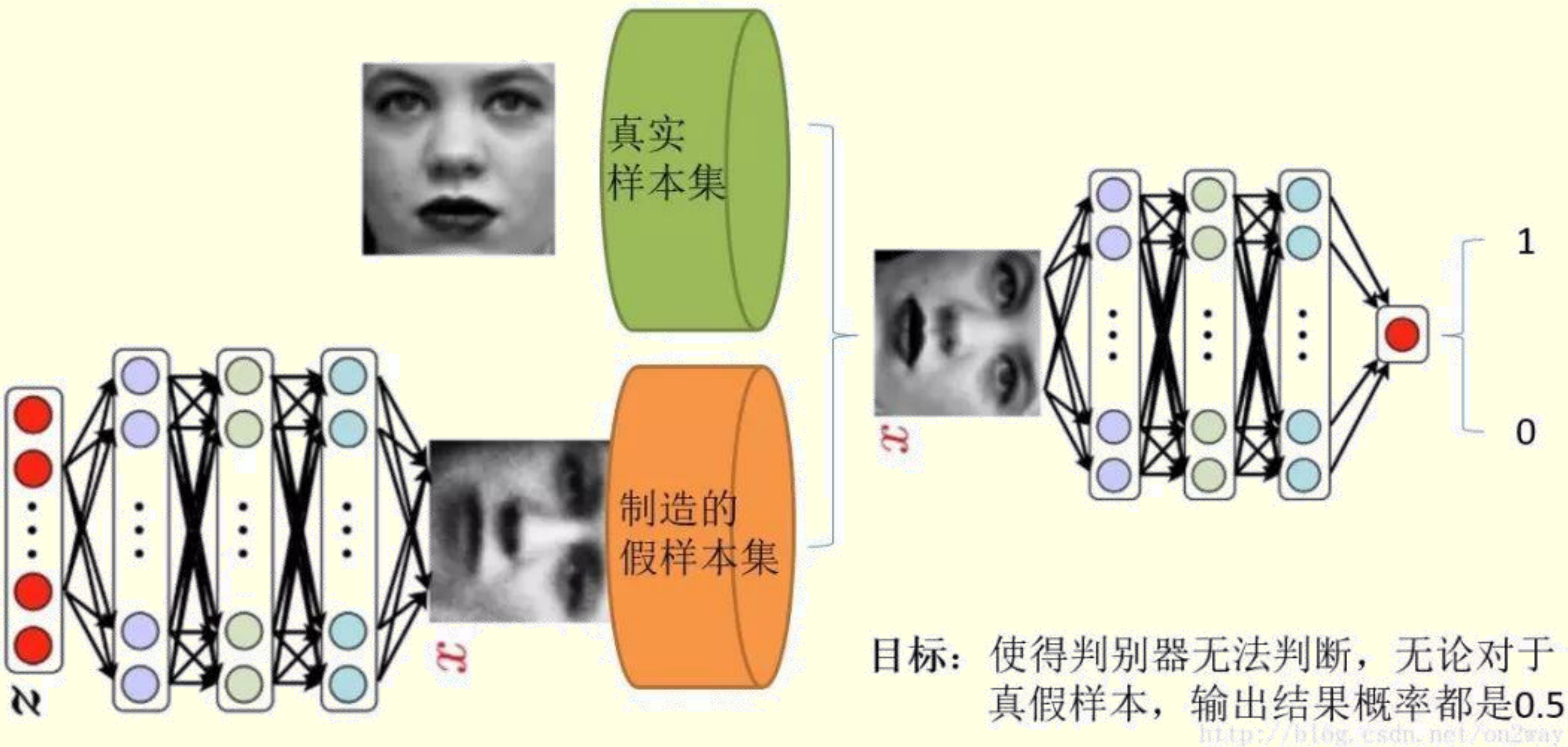
“奥曼从数学的角度,谢林从经济学的角度用博弈论重塑了对人类社会相互行为影响的分析框架”

博弈论（Game Theory）：发展历史

- 《美丽心灵》- 2001年
  - 约翰·福布斯·纳什
  - 1928年出生于西弗吉尼亚州
  - 1948年毕业于卡内基工学院
  - 1950年获得普林斯顿大学博士学位
    - 博士论文中提出"*纳什均衡*"
  - 1958年被《财富》评为天才数学家
  - … … ？！
  - 1994年获诺贝尔经济学奖
  - 2015年因车祸去世

博弈论（Game Theory）：发展历史

- 2014年Goodfellow设计了对抗生成网络（GAN）



提 纲

- 博弈论
  - 博弈论概述
  - 博弈论发展历史
  - 博弈问题分类
  - 博弈问题实例
- 搜索策略
  - 极小极大搜索算法
  - α-β搜索法

博弈问题：分类

- 零和博弈
  - 指参与博弈的各方，在严格竞争下，一方的收益必然意味着另一方的损失，博弈各方的收益和损失相加总和永远为“零”，双方不存在合作的可能。
- 非零和博弈
  - 指参与博弈的各方，在竞争与合作并存的情况下，自己的所得并不与他人的损失的大小相等，博弈各方的收益或损失相加总和不为“零”，双方存在合作的可能。

博弈问题：分类

	静态	动态
完全信息	完全信息静态博弈 纳什均衡 纳什（1950，1951）	完全信息动态博弈 子博弈精炼纳什均衡 泽尔腾（1965）
不完全信息	不完全信息静态博弈 贝叶斯纳什均衡 海萨尼（1967－1968）	不完全信息动态博弈 精炼贝叶斯纳什均衡 泽尔腾（1975） Kreps和Wilson（1982） Fudenberg和Tirole （1991）

博弈问题：分类

- 纳什均衡
  - 非合作博弈均衡
  - 一个策略组合被称为纳什平衡，当每个博弈者的平衡策略都是为了达到自己期望收益的最大值，与此同时，其他所有博弈者也遵循这样的策略
  - 纯策略
    - 任何一种情况下玩家的确定性决策
  - 纯策略纳什均衡
    - 所有玩家均采用纯策略
  - 混合策略纳什均衡
    - 至少一个玩家不采用纯策略（决策的不确定性）

博弈问题：分类

- 子博弈精炼纳什均衡
  - TBA

博弈问题：分类

- 贝叶斯纳什均衡
  - TBA

博弈问题：分类

- 精炼贝叶斯纳什均衡
  - TBA

提 纲

- 博弈论
  - 博弈论概述
  - 博弈论发展历史
  - 博弈问题分类
  - 博弈问题实例
- 搜索策略
  - 极小极大搜索算法
  - α-β搜索法

博弈问题实例：囚徒困境

B / A	坦白	抵赖
坦白	无期，无期	释放，死刑
抵赖	死刑，释放	一年，一年

- 非合作博弈
  - 囚徒A和B被分隔在不同的房间（防止串供）
  - A和B非法侵入事实确凿（一年刑期）
  - 审问A和B是否还有其他更严重的犯罪行为

博弈问题实例：囚徒困境

- 如何利用囚徒困境 - 沃伦·巴菲特提出的政治策略
  - 一位富翁提出一项对竞选活动融资的法案，他承诺，如果这一法案立法不成，他会向最有力支持这一法案的政党提供10亿美元。

民主党 / 共和党	支持	反对
支持	法案通过	法案不通过，共和党拿到10亿美元
反对	法案不通过，民主党拿到10亿美元	法案不通过

博弈问题实例：囚徒困境

- 如何跳出囚徒困境
  - 报复与惩罚，如：黑手党组织
  - “人质”方案，如：秦始皇、燕国太子丹
  - 忠诚文化：如：江湖义气、效忠文化
  - 建立共赢关系，如：双边贸易

博弈问题实例：吉诺维斯谋杀案

- 吉诺维斯谋杀案——市民责任的博弈
  - 1964年纽约市（皇后区的Kew花园），一个叫吉诺维斯的妇女被歹徒杀害，残忍的袭击持续了半个多小时，她一直在尖叫，很多人听到了她的尖叫，超过30人在命案现场，但没有人帮助她，也没有人报警。

张三 \ 李四	旁观	报警	
旁观	0，0	10，7	p
报警	7，10	7，7	1-p
	q	1-q	

博弈问题实例：吉诺维斯谋杀案

- 各参与人在不同策略下的预期收益：
  - 张三
    - 旁观：0*q + 10*(1-q) = 10 -10q (1)
    - 报警：7*q + 7*(1-q) = 7 (2)
    - 纳什均衡满足李四选择q使得张三在各策略之间的预期收益无差异，则使式(1)＝式(2)，解得q＝0.3，这是李四旁观的概率。

博弈问题实例：吉诺维斯谋杀案

- 各参与人在不同策略下的预期收益：
  - 李四
    - 旁观：0*p + 10*(1-p) = 10 -10p　(3)
    - 报警：7*p + 7*(1-p) = 7　(4)
    - 同理，使式(3)＝式(4)，解得p＝0.3，这是张三旁观的概率。

博弈问题实例：吉诺维斯谋杀案

- 至少有一个参与人报警的概率：
  - 有时候人们以随机的方式选择自己的行动，例如在本案例中张三和李四都会以0.3的概率选择旁观，以0.7的概率选择报警。由此可以计算各种情况的概率：
    - 两个人都报警的概率为0.7 × 0.7 = 0.49
    - 张三报警李四旁观的概率为0.7 × 0.3 = 0.21
    - 李四报警张三旁观的概率为0.3 × 0.7 = 0.21
    - 两个人都旁观的概率为0.3 × 0.3 = 0.09
    - 因此，警方最终得到报警的概率是：
      - 0.49 + 0.21 + 0.21 = 0.91

博弈问题实例：吉诺维斯谋杀案

- 假设参与人数为N：
  - 假设张三认为其他人旁观的概率为q，则：
    - 张三选择旁观：$0\times q^{N}-1+10(1-q^{N-1})=10-10\ q^{N-1}$
    - 张三选择报警：$7\times q^{N}-1+7(1-q^{N-1})=7$
    - 那么，均衡状态下张三的最优混合策略应使得上面两式相等:
      - $10-10\ q^{N-1}=7\rightarrow q^{*}=0.3^{1/(N-1)}$
  - 在吉诺维斯谋杀案中，N＝38的情况下
    - 每个人都将以0.97（即就是 $0.3^{1/(38-1)}$ ）的概率选择袖手旁观，那么无人报警的概率就是0.29(即就是0.97^{38})。

博弈问题实例：吉诺维斯谋杀案

起初他们追杀共产主义者，我没有说话，

　　因为我不是共产主义者；

接着他们追杀犹太人，我没有说话，

　　因为我不是犹太人；

后来他们追杀工会成员，我没有说话，

　　因为我不是工会成员；

此后，他们追杀天主教徒，我没有说话，

　　因为我不是天主教徒；

最后他们奔我而来，却再没有人站起来为我说话了。

——马丁·尼莫拉

（德国新教教士，此诗铭刻于波士顿犹太人屠杀纪念碑）

博弈问题实例：吉诺维斯谋杀案

- 策略选择提示：
  - 在博弈中，有时需要以某种概率随机地选择自己的行动，即采取混合策略
  - 有时候，混合策略的后果是低效率的，尤其是大规模群体需要协调行动的时候
- 如何避免？

提 纲

- 博弈论
  - 博弈论概述
  - 博弈论发展历史
  - 博弈问题分类
  - 博弈问题实例
- 搜索策略
  - 概述
  - 极小极大搜索算法
  - α-β搜索法

搜索策略：概述

- 60年代，已经出现若干博弈程序
- 1997年， IBM公司设计的深蓝（deep blue）计算机与国际象棋大师卡斯帕罗夫对弈，取得了三胜二和一负的好成绩
- 2016年，DeepMind所设计的AlphaGo以4:1战胜韩国围棋天王李世石
  - 状态空间相对国际象棋或中国象棋巨大
    - 每个点有黑、白、无三种状态，共361个空格
  - 局势评估复杂度高
  - 目标变化多样

搜索策略：概述

- 如何表示博弈问题的状态，博弈过程和博弈取胜的知识，这是目前人类仍在探讨之中的问题。要提高博弈问题求解程序的效率，应作到如下两点：
  - 改进生成过程，使之只生成好的走步，如按棋谱的方法生成下一步；
  - 改进测试过程，使最好的步骤能够及时被确认。

搜索策略：概述

- 要达到上述目的有效途径是使用启发式方法引导搜索过程，使其只生成可能赢的走步。而这样的博弈程序应具备：
  - 一个好的（即只产生可能赢棋步骤的）生成过程；
  - 一个好的静态估计函数。

提 纲

- 博弈论
  - 博弈论概述
  - 博弈论发展历史
  - 博弈问题分类
  - 博弈问题实例
- 搜索策略
  - 概述
  - 极小极大搜索算法
  - α-β搜索法

极小极大搜索算法：主要思路

- 极小极大搜索策略是考虑双方对弈若干步之后，从可能的步中选一步相对好的走法来走，即在有限的搜索深度范围内进行求解；

- 为此要定义一个静态估计函数f，以便对棋局的势态作出优劣估计。这个函数可根据棋局优劣势态的特征来定义。

极小极大搜索算法：主要思路

- 主要规定：
  - MAX代表程序方
  - MIN代表对手方
  - P 代表一个棋局（即一个状态）
- f(P)的大小由棋局势态的优劣来决定:
  - 有利于MAX的势态， f(P)取正值
  - 有利于MIN的势态， f(P)取负值
  - 势态均衡， f(P)取零

极小极大搜索算法：主要思路

- 使用静态函数进行估计必须以下述两个条件为前提：
  - 双方都知道各自走到什么程度、下一步可能做什么
  - 不考虑偶然因素影响
- 在这个前提下，博弈双方必须考虑：
  - 如何产生一个最好的走步
  - 如何改进测试方法，能尽快搜索到最好的走步。

极小极大搜索算法：主要思路

- MINMAX的基本步骤是：
  - 当轮到MIN走步的结点时，MAX应考虑最坏的情况（因此，f(P)取极小值）；
  - 当轮到MAX走步的结点时，MAX应考虑最好的情况（因此，f(P)取极大值）；
  - 当评价往回倒推时，相应于两位棋手的对抗策略，不同层上交替的使用步骤一、二两种方法向上传递倒推值。

极小极大搜索算法：实例

- 在九宫格棋盘上两位选手轮流在棋盘上摆各自的棋子，每次一枚，谁先取得三子一线的结果就取胜：
  - 设程序方MAX的棋子用X表示
  - 对手方MIN的棋子用O表示

X_5	O_2	O_4
X_7	X_1	
X_3		O_6

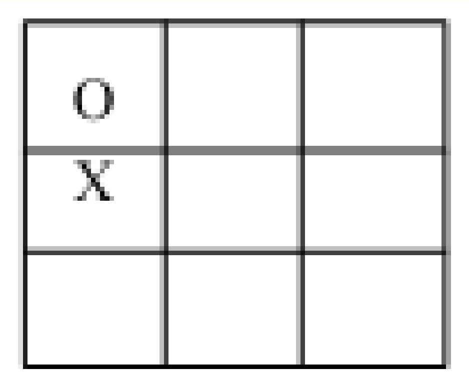
极小极大搜索算法：实例

静态估计函数为:

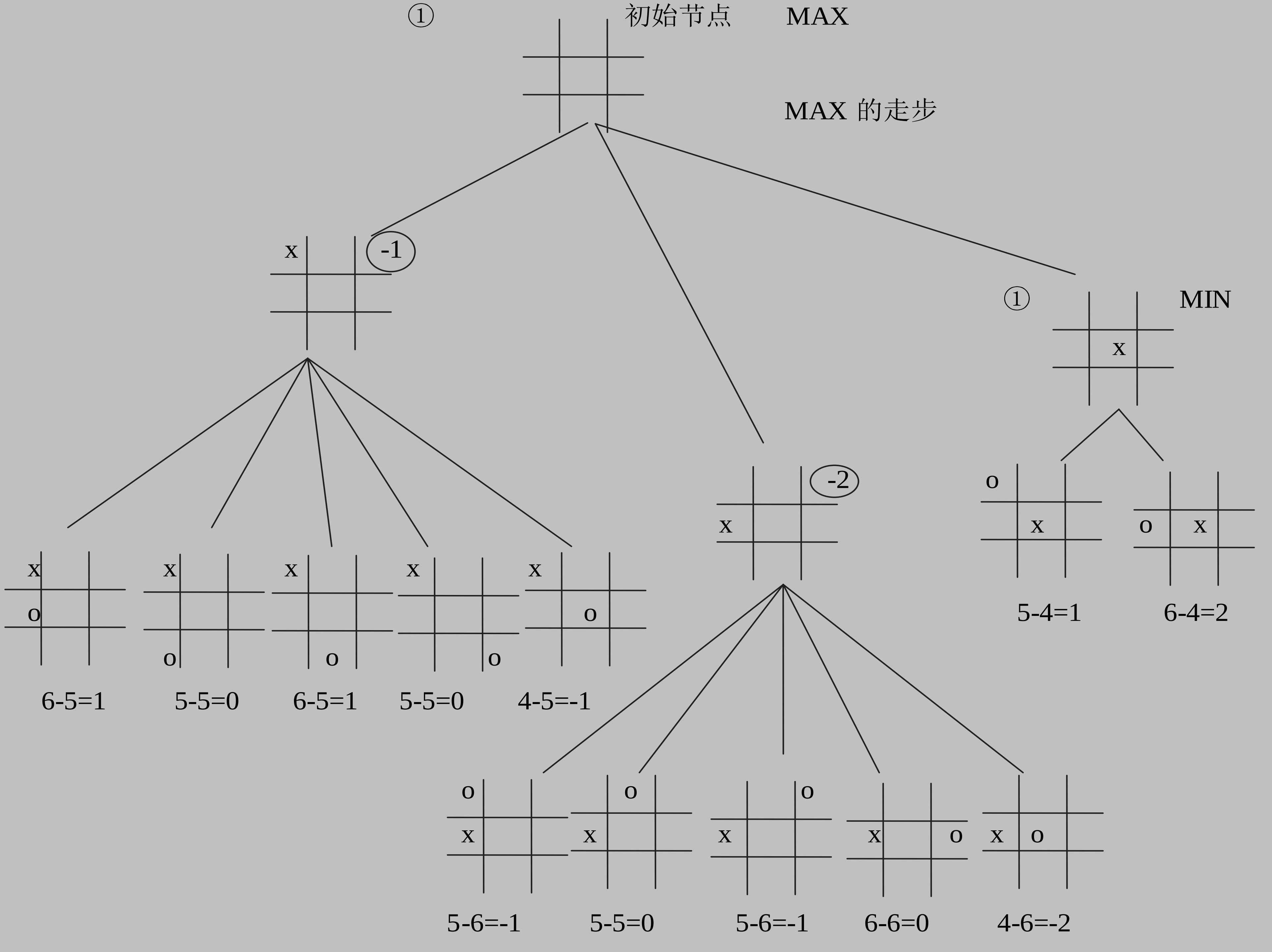
$$
f(p) = \begin{cases} +\ \infty & \text{当}p\text{为MAX赢} \\ -\ \infty & \text{当}p\text{为MIN赢} \\ \text{全部空格放X后三字成一线的总数）} & \\ \quad -\text{（全部空格放O后三字成一线的总数）} & \end{cases}
$$

例如，P的格局为：

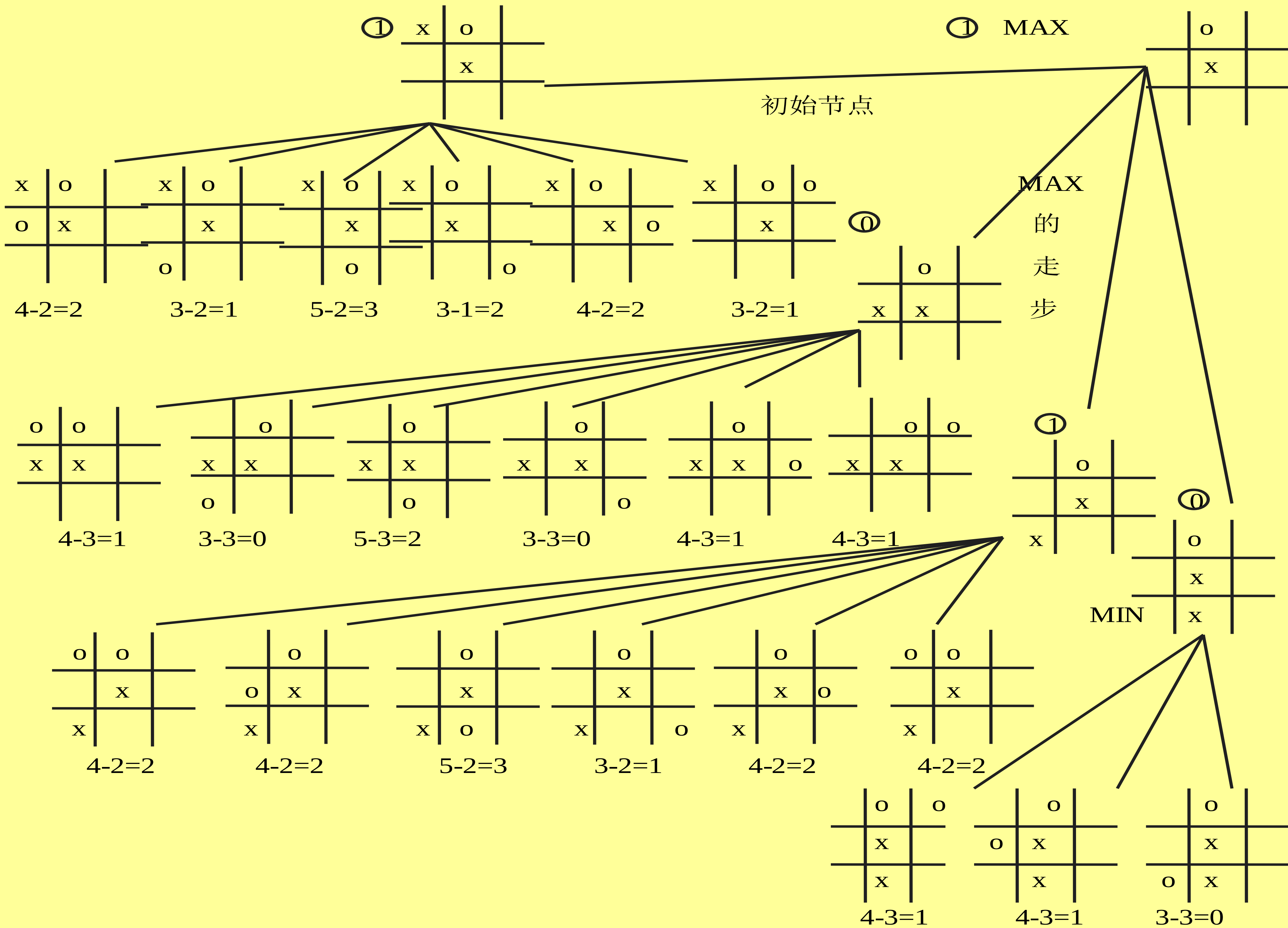
极小极大搜索算法：实例



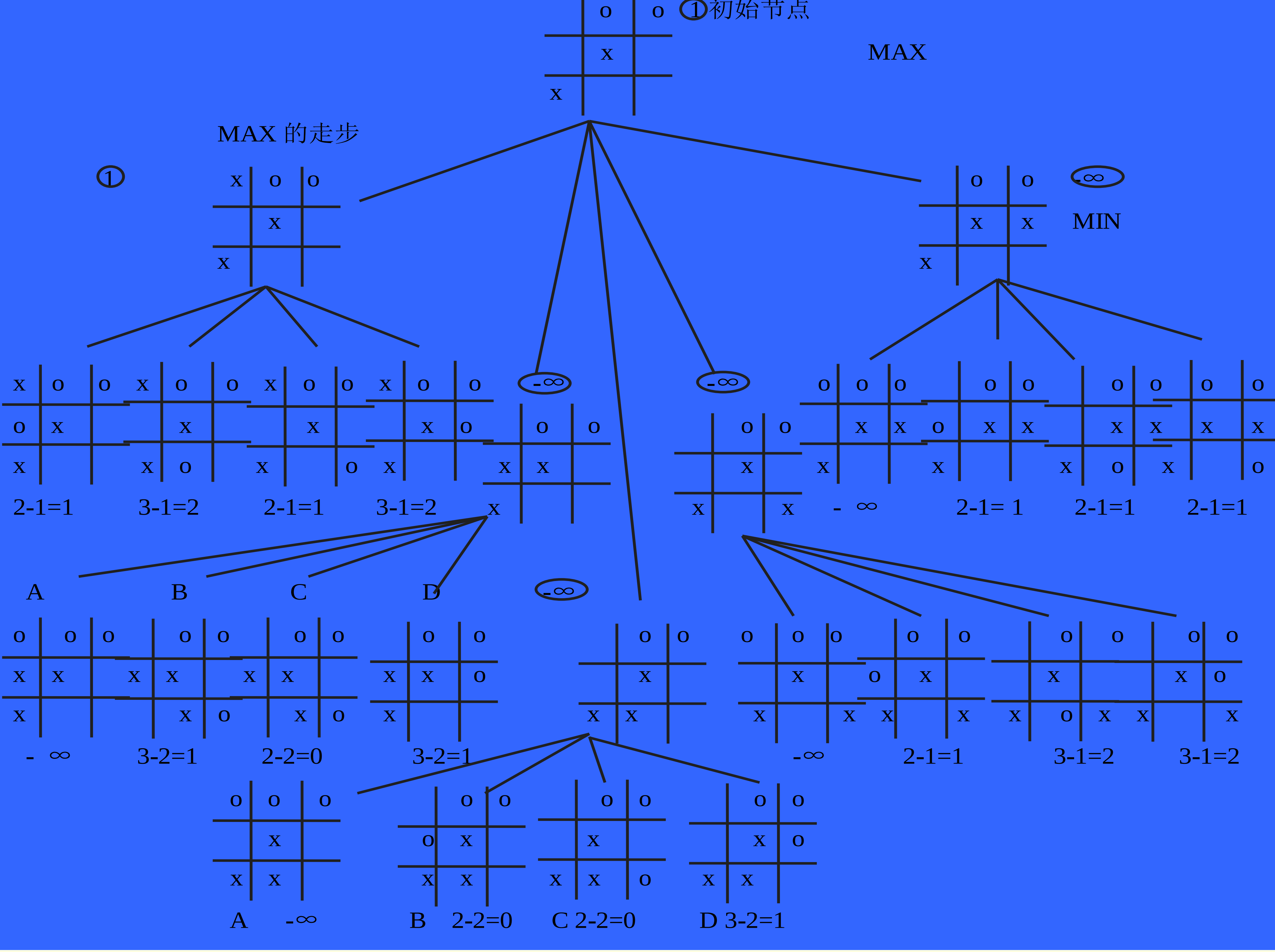
- 则可得f(p)=5 – 6 = -1。
- 现在考虑走两步的搜索过程，即算法中K=2。利用棋盘对称性条件，则MAX走第一步棋调用算法产生搜索树如下图所示。



①
初始节点
MAX
MAX 的走步
-1
①
MIN
-2
6-5=1
5-5=0
6-5=1
5-5=0
4-5=-1
5-4=1
6-4=2
5-6=-1
5-5=0
5-6=-1
6-6=0
4-6=-2



1
MAX
初始节点
4-2=2
3-2=1
5-2=3
3-1=2
4-2=2
3-2=1
0
MAX
的
走
步
4-3=1
3-3=0
5-3=2
3-3=0
4-3=1
4-3=1
1
0
MIN
4-2=2
4-2=2
5-2=3
3-2=1
4-2=2
4-2=2
4-3=1
4-3=1
3-3=0

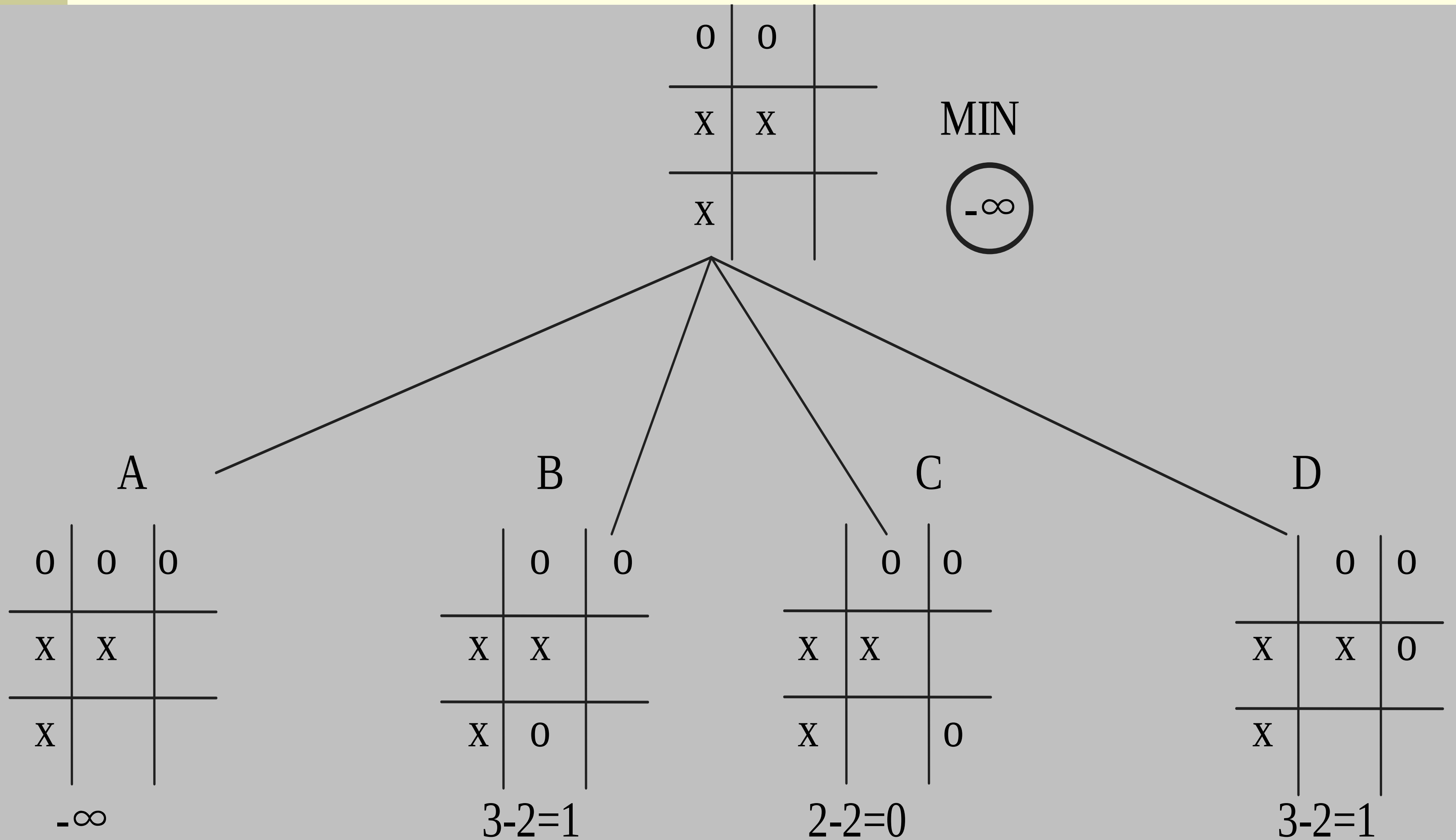


①初始节点
MAX
MAX 的走步
①
-∞
MIN
2-1=1
3-1=2
2-1=1
3-1=2
-∞
-∞
- ∞
2-1= 1
2-1=1
2-1=1
A
B
C
D
-∞
- ∞
3-2=1
2-2=0
3-2=1
-∞
2-1=1
3-1=2
3-1=2
A -∞
B 2-2=0
C 2-2=0
D 3-2=1

提纲

- 博弈论
  - 博弈论概述
  - 博弈论发展历史
  - 博弈问题分类
  - 博弈问题实例
- 搜索策略
  - 概述
  - 极小极大搜索算法
  - α-β搜索法

α-β搜索算法：实例



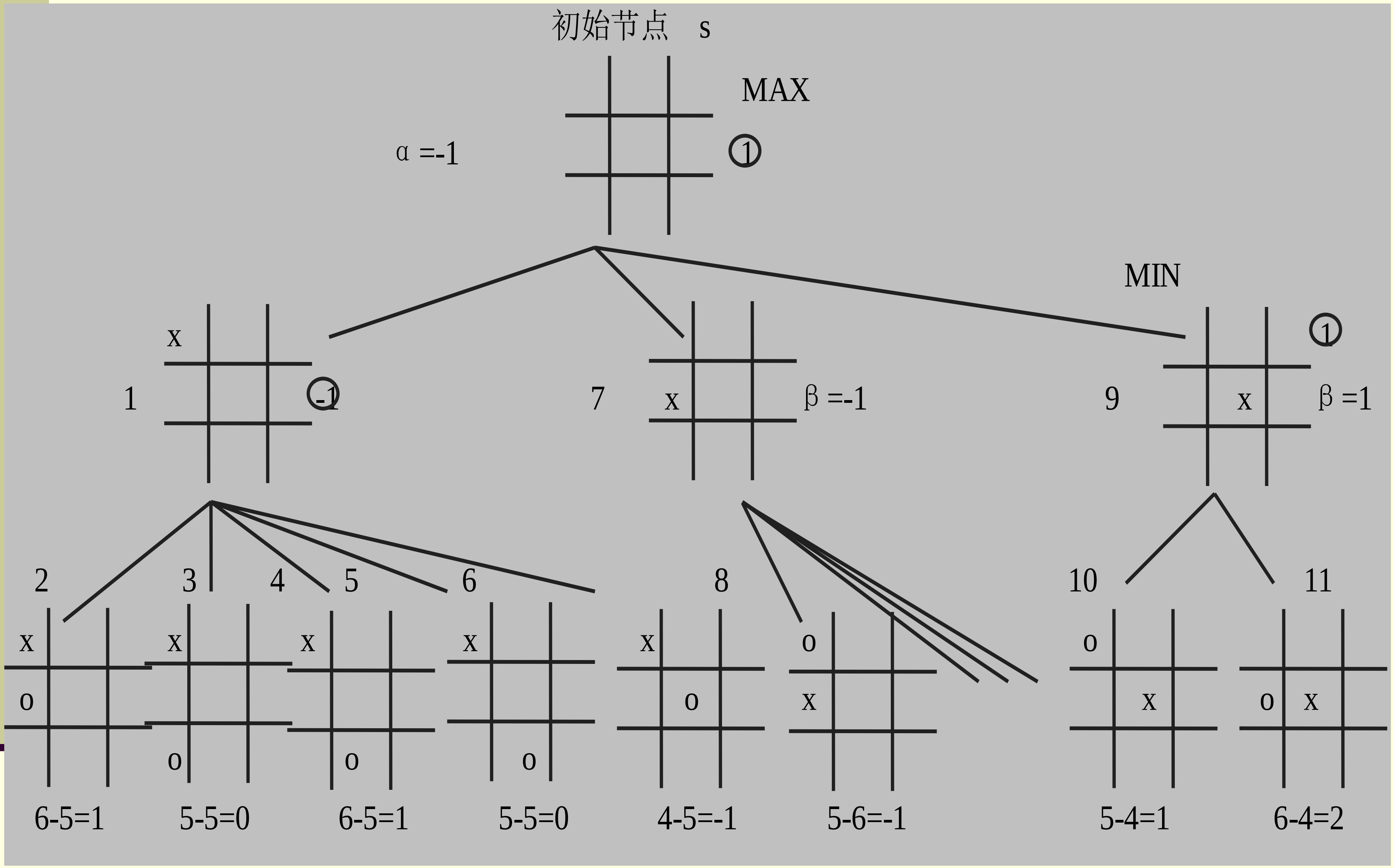
α-β搜索算法：实例

- 这时其中一个MIN节点要生成A,B,C,D四个节点，然后逐个计算其静态估计值，最后求得倒推值-∞，把它赋给这个结点。其实生成节点A后，如果马上进行静态估计，得知F(A)= -∞之后，就可以断定生成B,C,D以及进行估计是多余的，该MIN节点的倒推值一定是-∞。

α-β搜索算法：主要思路

- *α-β*剪枝法就是把生成后继和倒推值估计结合起来，及时剪掉一些无用分枝，以此来提高算法的效率。

- *α-β*剪枝法，采用有界深度优先策略进行搜索，当生成节点达到规定的深度时，就立即进行静态估计，而一旦某个非端节点有条件确定倒推值时，就立即赋值。



- 当生成到节点6后，节点1的倒推值可确定为-1。

α-β搜索算法：实例

- *这时对于初始节点S来说，虽然其他子节点尚未生成，但由于S属于极大层，所以可以推断它的倒推值不会小于-1。*
  - *我们定义极大层的这个下界值为α。因此S的α= - 1*
  - *S的α值为-1，说明的S倒推值不会比-1更小，但会不会比-1更大，还取决于其他后继节点倒推值。我们继续生成搜索树。*
  - *当第8个节点生成后，其估计值为-1，就可以断定节点7的倒推值不可能大于-1。*

α-β搜索算法：实例

- *定义极小层的这个上界值为β 。*
  - *因此现在可以确定节点7的β= - 1。*
  - *有了极小层的β值， 容易发现α≥β时， 节点7的其他子节点不必生成，因为S的极大值不可能比这个β值小，再生成无疑是多余的，因此可以进行剪枝*
  - *只要在搜索过程中记住倒推值的上下界并进行比较，当α≥β时就可以实现修剪操作。*

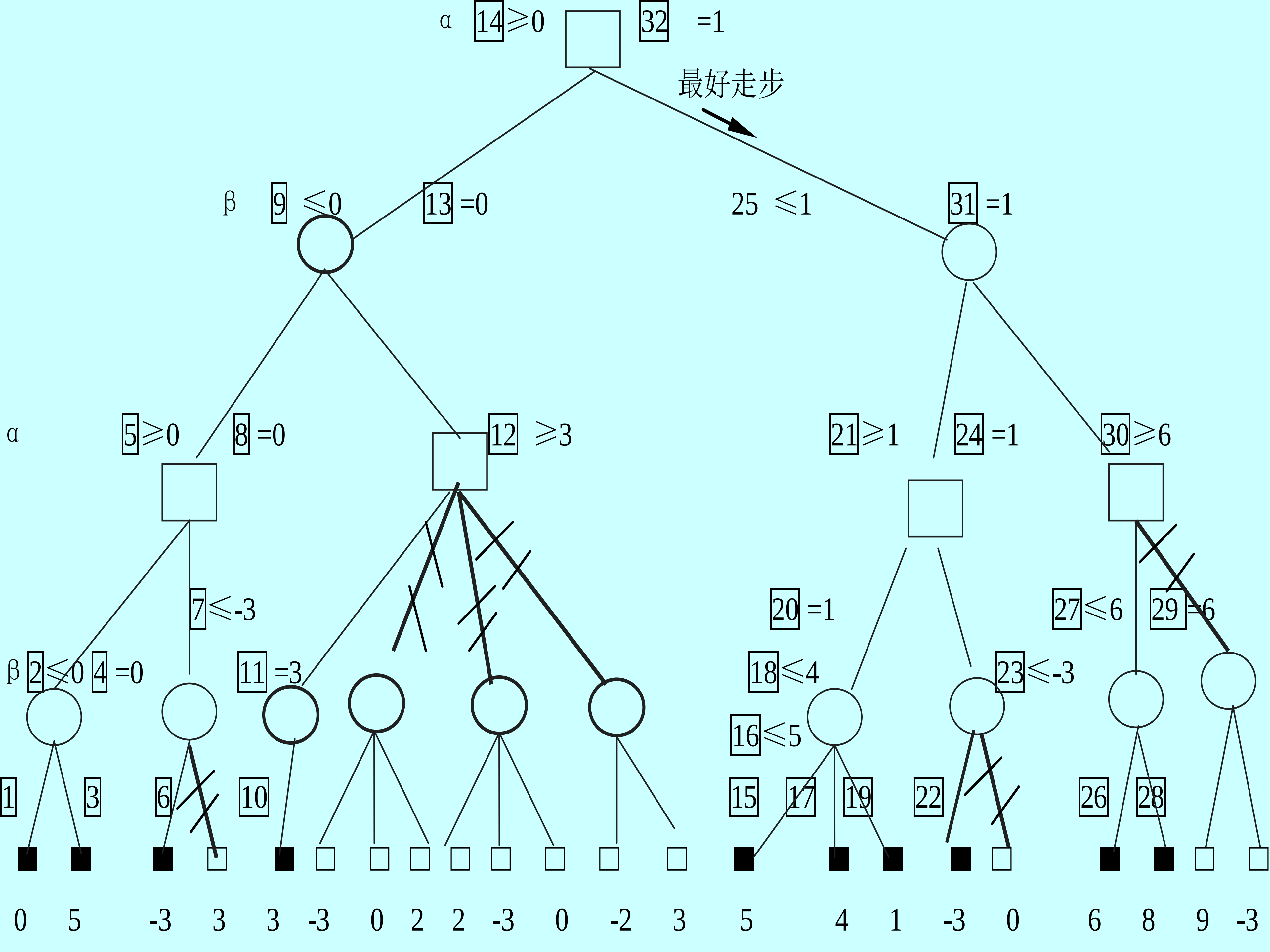
α-β搜索算法：实例

- α,β值还可以随时修正，但极大层的α倒推值下界永不下降，因为实际的倒推值取后继节点最终确定的倒推值中的最大者。

- 同理，极小层的倒推值上界β永不上升，因为实际倒推值取后继节点最终确定的倒推值中的最小者。

α-β搜索算法：剪枝规则

- *α剪枝：若任一极小值层节点的β值小于或等于它任一先辈极大值层节点的α值，即α（先辈层）≥β（后继层），则可中止该极小值层中这个节点以下的搜索。该节点最终的倒推值就确定为这个β值。*
- *β剪枝：若任一极大值层节点的α值大于或等于它任一先辈极小值层节点的β值，即α（后继层）≥β（先辈层），则可以中止该极大值层中这个节点以下的搜索。这个MAX节点的最终倒推值就确定为这个α值。 。*



α 14≥0
32 =1
最好走步
β 9 ≤0
13 =0
25 ≤1
31 =1
α
5≥0
8 =0
12 ≥3
21≥1
24 =1
30≥6
7≤-3
20 =1
27≤6
29 =6
β 2≤0 4 =0
11 =3
18≤4
23≤-3
16≤5
1
3
6
10
15
17
19
22
26
28
0 5 -3 3 3 -3 0 2 2 -3 0 -2 3 5 4 1 -3 0 6 8 9 -3

Q & A