人工智能概论复习资料

24 级 AI 学组

Xi'an Jiaotong University

绪论

一．人工智能的定义与目标

1. 人工智能正在引领第四次工业革命——智能化
2. 人的智能——解决问题、从经验中学习、运用知识适应新情况的能力
3. 人工智能是制造智能机器的科学和工程学；与使用计算机来了解人类智能的类似任务有关，但是人工智能不必将自己局限于生物学上可观察的方法——人工智能专家 John McCarthy
4. 人类智能与人工智能的区别：物质载体和运动方式不同、进化途径和本质属性不同、思维方式和实现机制不同

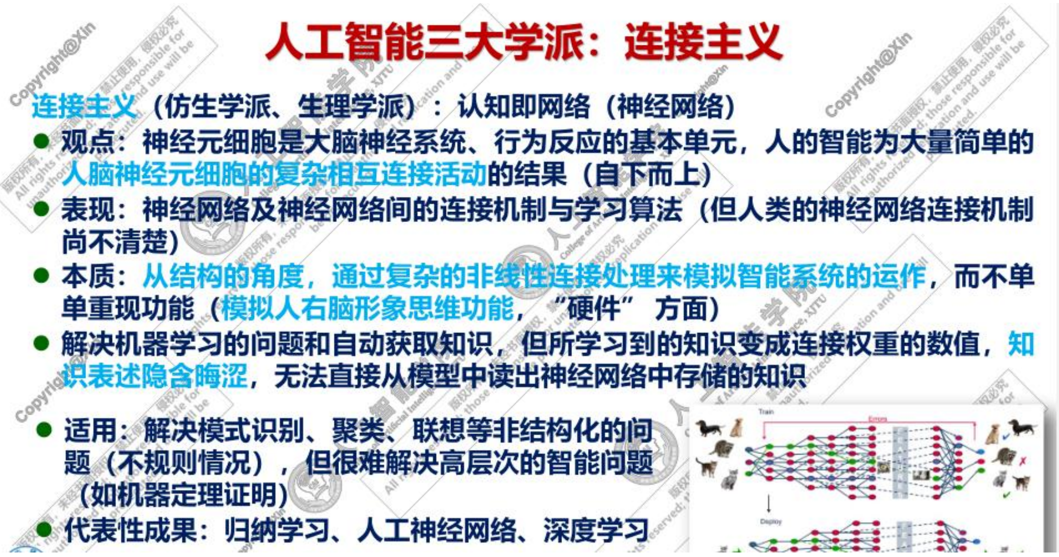
二．人工智能的发展历程

1. 早期种子：逻辑与推理（符号学派）：亚里士多德的三段论，莱布尼茨的逻辑演绎系统，布尔代数，弗雷格的谓词演算
2. **第一次爆发（1956-1974）**：出现字符识别程序、通用问题求解器、感知机；麦卡锡研制表处理语言 LISP；1956 年是人工智能元年，达特茅斯会议上 John McCarthy 首次提出人工智能这一术语，会议聚集了麦卡锡、明斯基、纽厄尔、西蒙等“AI 之父”。
3. **第一次寒冬（1974-1980）**：机器翻译失败；Lighthill 报告批评；Minsky & Papert《感知机》一书指出单层神经网络的局限性（无法解决 XOR 问题）
4. **第二次爆发（1980-1987）**：专家系统、神经网络复苏
5. **第二次寒冬（1987-1993）**：专家系统知识库构建困难、维护成本高、推理能力脆弱；LISP 机器市场崩溃，PC 革命兴起
6. **第三次爆发（1993 至今）**：机器学习和深度学习兴起：“深蓝”、AlexNet、AlphaGo……语音识别、计算机视觉、专家知识系统等新分支获得了应用

三．人工智能三大学派

1. 符号主义

2. 连接主义



3. 行为主义

四．人工智能三大类型

1. **弱人工智能**：为一项特定任务而设计和培训的人工智能系统，但并不意味着其效率低下或类似问题（目前所有 AI 都是弱智能）
2. **强人工智能**：具有认知能力的系统，当系统面临不熟悉的任务时，其足够智能地找到解决方案；旨在复制人类智能的许多（理想情况下）能力
3. 自我意识人工智能

五．往年真题

2024 年真题 （选择）人工智能元年、三大学派、两大分类

2023 年真题 （选择）两大分类、三大学派、三次爆发

知识图谱

基本知识:

1】抽象层次&加工程度（升序）：数据-信息-知识-知识图谱

2】知识分类：陈述性/过程性、概念/事实/规则

3】知识图谱为符号主义学派发展成果

4】知识图谱过程描述：从非/半结构化数据（图像、网页等）抽取信息、构建结构化数据（三元组）

5】知识图谱优势：关联链接不同来源、类型、结构知识，知识体系更广、更深且可扩充

6】本体：知识库知识的概念模板、位于模式层

7】实体：知识图谱中最基本元素

8】实体关系：详见知识表示中的语义网络表示法

9】**知识图谱基本单位**：**三元组**（实体 1-关系-实体 2/实体-属性-属性值）

10】知识图谱构成：多类型节点（概念、实体）+多类型边（概念、实体属性/概念、实体关系）构成的多关系图

11】知识图谱存储：RDF（三元组，不包含属性）/图数据库（属性图）

12】**知识图谱技术结构**:

简版：<u>信息抽取-知识表示-知识融合-知识加工-知识更新</u>

详细版:

1. 原始数据转化为三元组形式
2. 对三元组表示的数据知识融合（实体对齐&结合数据模型），形成标准数据表示
3. 按照一定推理规则，产生隐含知识
4. 质量评估，符合要求知识进入知识图谱

相关说明:

1）信息提取（非结构化数据源抽取结构化信息）:

包括：实体+关系+属性提取

2）知识表示（使计算机可存储计算）

包括：传统、深度学习表示法

3）知识融合（消歧、去重、剔错，提高知识质量）:

包括：1°实体链接：实体消歧（同名实体产生歧义+共指消解（多指称对应同实体）

4）知识加工（结构化、网络化知识体系）

包括：1° 本体构建（形式化明确定义概念&联系）：人工/自动

2° 知识推理（发现新知识&新关系）：基于逻辑/图

3°质量评估：舍弃置信度低的知识

5）知识更新（更新数据层和模式层）

包括：全面更新（从零重建）&增量更新（往现有里面增添）

13】**知识图谱逻辑结构（模式层&数据层）**:

1）**模式层**：存储提炼的知识集合（即节点为本体概念、边为概念关系）

2）**数据层**：存储客观事实（三元组）（即节点为实体、边为实体关系、属性）

14】知识图谱应用：智能搜索、智能问答、人物关系图、概念图谱、风险分析、可视化决策支持、推荐系统等

总结：

1】采用本体知识表示，揭示实体关系

2】基本组成单元为三元组（实体-关系-实体/实体-属性-属性值）

知识表示

一．概念汇总：

1. 知识由概念组成，概念是构成人类知识和思维体系的基本单元。
2. 概念主要组成：概念名，概念的内涵（事物对象所特有的属性，反映概念的本质属性，通常用命题来表示），概念的外延（满足概念内涵表示的对象构成的经典几何，反映具体事物对象（实例）的范围）。
3. 内涵越多，外延越少，反之亦然（反比关系）。
4. 从浅入深：DATA -> INFORMATION -> KNOWLEDGE -> WISDOM
5. 知识类型：结构性知识（事物间的关系），过程性知识（规则，策略），陈述性知识（概念，事实），启发式知识（与专业知识，经验等有关），元知识。
6. 知识的特性：相对正确性（与条件有关），不确定性，可表示性与可利用性。
7. 知识表示：将人类的知识形式化（符号）或模型化（结构），以便机器识别与理解。
8. 知识符号化过程与方法包括：1.用给定的知识结构，按一定原则、组织表示知识。2，解释该知识的含义。
9. 基于符号的知识表示方法：逻辑表示，产生式表示（生产规则），语义网络表示，框架表示。
10. 知识表示方法也可粗略地分为：叙述（陈述）性表示（静态特性，显性知识，容易表达，不可被计算机直接执行）， 过程性表示（动态特性，隐性知识，难以表达，可被计算机直接执行）。

二．谓词逻辑（数理逻辑）表示法：

1. 数理逻辑主要包括：命题逻辑，谓词逻辑。
2. 命题：一个能确定真或假意义的陈述句（不能是疑问句、命令句、感叹句等， 不含变量， 悖论不是命题）。
3. 命题逻辑是形式逻辑的基本组成部分，数理逻辑的基础。
4. 命题真假与条件相关。(1 + 1 = 10)
5. 5 个常用命题联结词中，非（not）是一元联结词，其余是二元联结词。
6. 命题 p 与 q 在所有情况下有相同的真假结果，则 p 与 q 逻辑上等价，用 p≡q 表示（汪老师提示您：不能用等号=）
7. 常见的推理规则

- 常见的推理规则：
  - 假言推理（modus ponens） $\alpha \to \beta, \alpha \Rightarrow \beta$
  - 与消解（and-elimination） $\alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_n \Rightarrow \alpha_i \ (1 \le i \le n)$
  - 与导入（and-introduction） $\alpha_1, \alpha_2, \cdots, \alpha_n \Rightarrow \alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_n$
  - 双重否定（double-negation elimination） $\neg\neg\alpha \Rightarrow \alpha$
  - 单项消解或单项归结（unity resolution） $\alpha \vee \beta, \neg\beta \Rightarrow \alpha$
  - 消解或归结（resolution） $\alpha \vee \beta, \neg\beta \vee \gamma \Rightarrow \alpha \vee \gamma$

8. 命题逻辑的局限：苏格拉底三段论。简单命题并不是逻辑推理的最终基本单元（但是命题逻辑的最基本单位），需要将命题分离成主语和谓词 → 谓词逻辑
9. 在命题逻辑的基础上，将原子命题分解成个体、谓词、量词以表达个体与总体的内在联系和数量关系。
10. 一元谓词 student(x)：通常表示个体的属性，多元谓词 father(x, y)：通常表示个体间的关系。
11. 函数与谓词的区别：

    - 函数中个体变量用个体常量（定义域）代入后结果仍是个体（值域），无真值
    - 谓词中个体变量用个体常量（定义域）代入后变为命题，具有确定的真值

    例子：father(x, y)是谓词（表示“x 是 y 的父皇”这个命题），而 father(y)则是函数（表示“y 的父皇”这个人）。
12. 量词：注意量词的辖域 –> 约束元：在辖域内且与量词中同名的变元；自由元：不受约束的变元。
13. 高阶谓词逻辑：如若谓词中的某个个体本身是一阶谓词，称为二阶谓词例：Works(engineer(Smith), IBM)。
14.
    - 原子谓词公式（原子公式）：若 $P(x_1, x_2, \cdots, x_n)$ 是n元谓词，$t_1, t_2, \cdots, t_n$ 是项，则 $P(t_1, t_2, \cdots, t_n)$ 为原子谓词公式
    - 谓词公式（合式公式）：由逻辑联结词和原子公式构成的用于陈述事实的复杂语句

三．产生式表示方法

1. 产生式表示法成为是许多成功的专家系统的第一选择的知识表示方法，产生式系统是专家系统的核心部分。
2. 蕴含式是产生式的特例。
3. 产生式表示中有：规则的表示，事实的表示。
4. 确定性规则的表示：

    $P \to Q$

    IF P THEN Q
5. 不确定性规则的表示：

6. 确定性事实的表示：一般使用三元组表示

7. 不确定性事实的表示：一般使用四元组表示

8. 有个巴科斯范式
9. 产生式系统求解：匹配规则 -> 所得结论作为新事实放入综合数据库 -> 重新匹配新规则

四．语义网络表示法

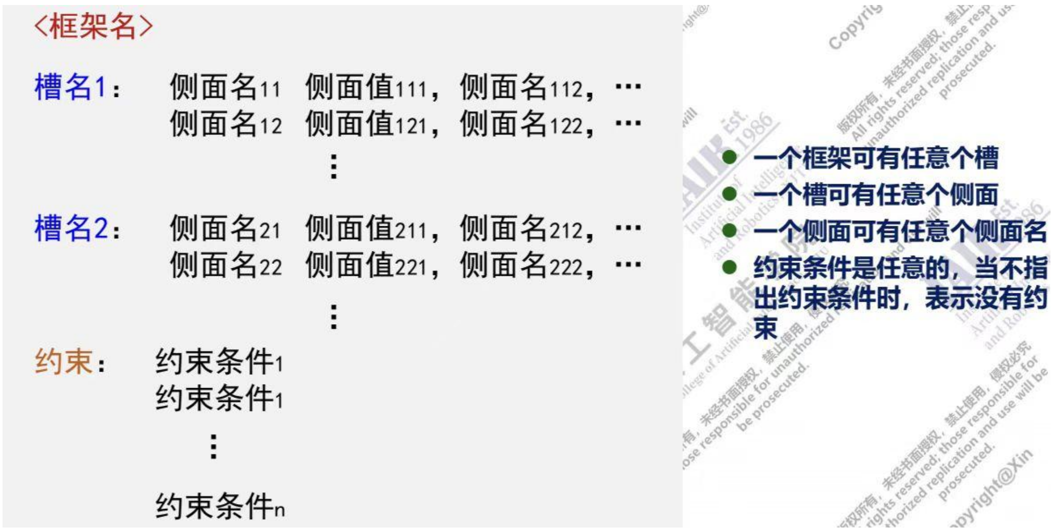
1. 语义基元：最基本的语义单元，用三元组表示：（节点一，弧，节点二）。
2. 语义基元的图表示（基本网元）

3. 常用的基本语义关系：1.实例关系“ISA”(is a) 2.分类关系”AKO”(a kind of) 3.成员关系”A-Member-of” 4.属性关系”Have”/”Can”/”Age” 5.包含关系”Part-of” 6.时间关系”Before”/”After” 7.位置关系“Located-prep.” 8.相近关系”Similar to” 9.推论关系”Reasoning to”
4. 几种知识表示的比较

五．框架表示方法

1. 框架的结构实例：

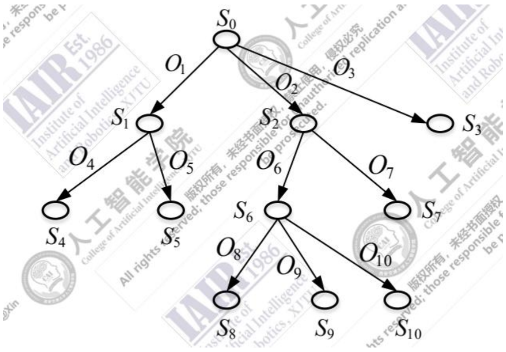
2. 一般的框架表示形式（有约束条件）：



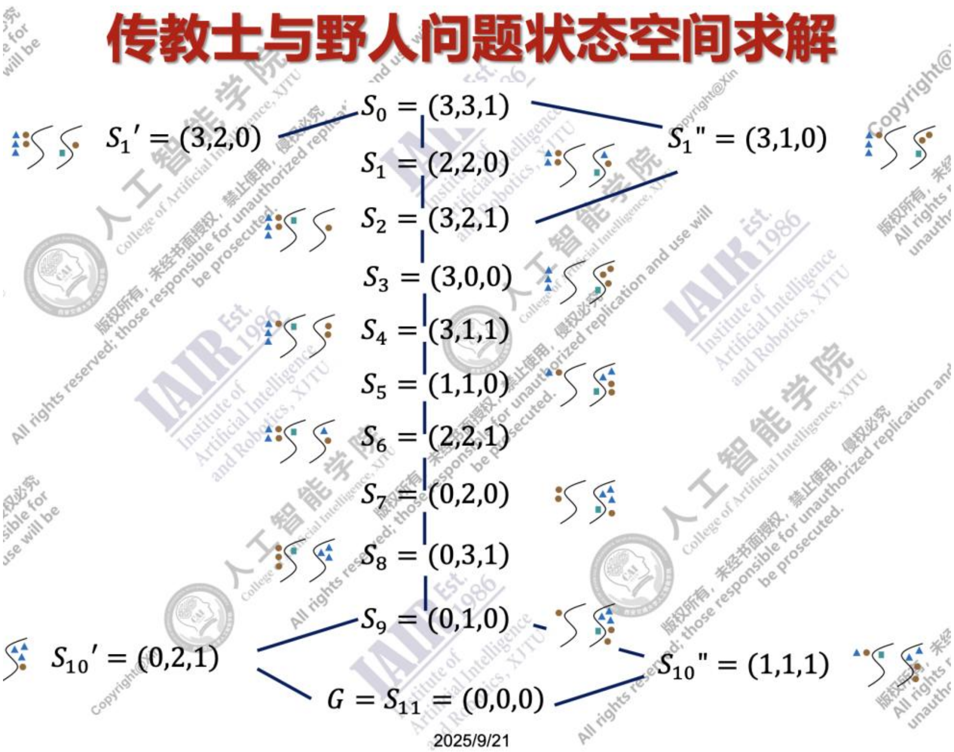
六．状态空间表示方法

1. 状态空间一般用四元组表示（每一个都表示一个集合）

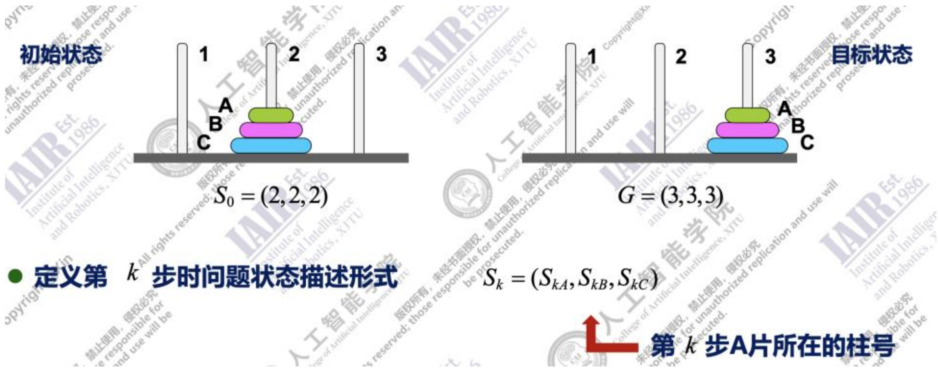
2. 状态空间也可用有向图描述，图的节点表示状态 S，图的弧表示相邻状态间变换的操作算子 O，从一个状态转移到另一状态的操作算子序列等价于有向图中的某一路径。



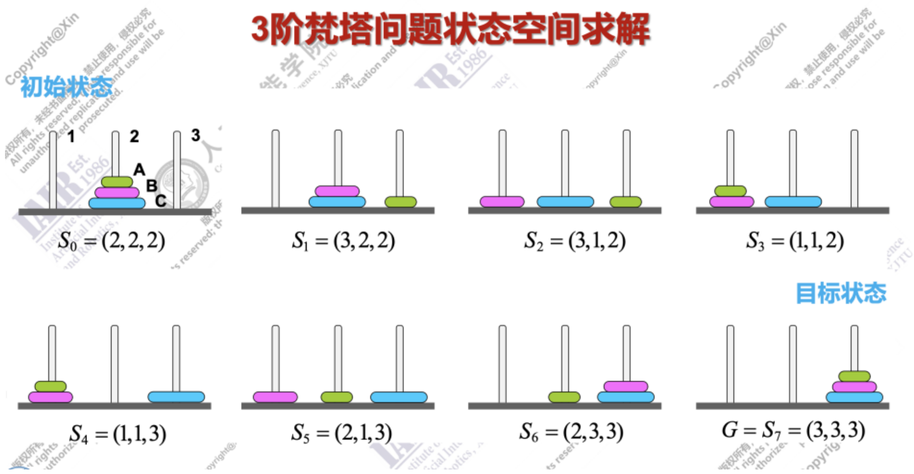
3. 例子：八数码问题（详情见"搜索求解"章节），推销员问题（TSP 问题，有限时间内给出问题的全部状态图十分困难），汉诺塔问题，传教士与与野人问题（S(x, y, z)状态，其中 x，y，z 分别表示在左岸的传教士，野人，船数，然后根据"人类启发式算法"（试出答案即可）得出结果）。



4. 汉诺塔问题求解：先表示状态空间



再画出状态转移



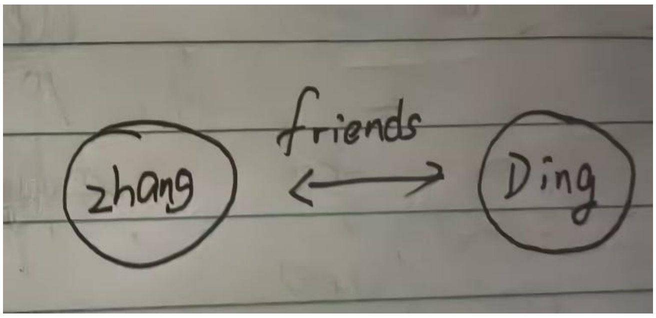
七．知识表示存在的问题

1. 人是如何表示知识，存储知识的？
2. 目前计算机中的知识处理也有三大难题："Clear-Cut"问题（如何清晰准确地将一个概念从其所属的更大范围中"切割"出来），现实知识难以转化为计算机可理解的知识，缺乏常识性知识。

八．2024 年真题

"小张和小丁是朋友"的命题逻辑表示，产生式表示，语义网络表示？

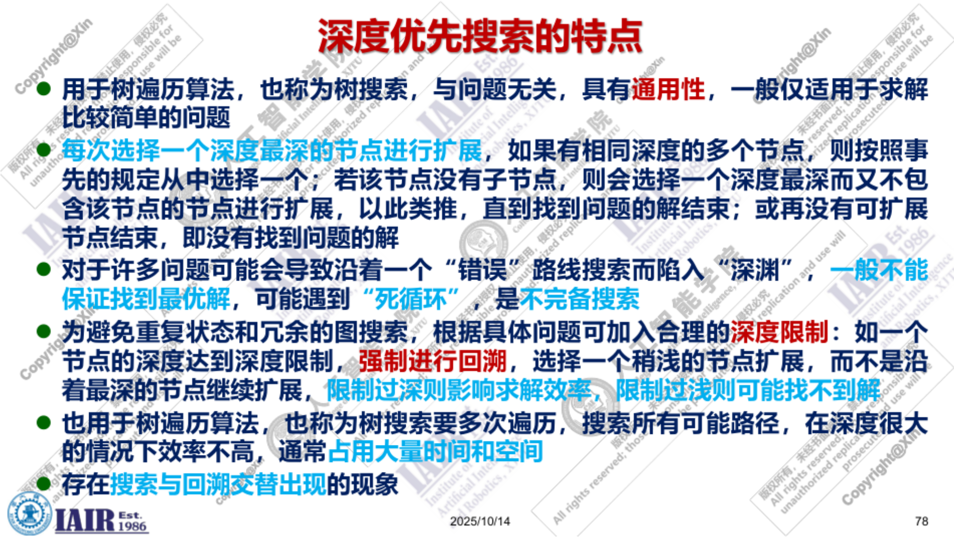
A. 命题逻辑表示：P，其中 P 表示"小张是小丁的朋友"
B. 谓词逻辑表示（感觉原题可能是这个）：friends(zhang, ding)
C. 产生式表示：（friends, zhang, ding）或 "IF zhang AND ding THEN friends"
D. 语义网络表示（节点，关系，节点）：(zhang, friends, ding) 或

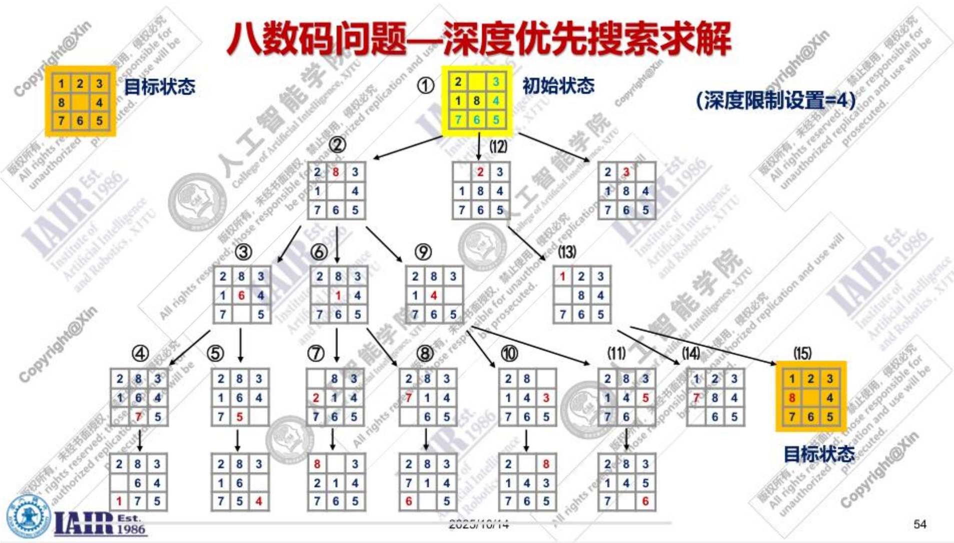


搜索求解

一．盲目搜索（无信息搜索）

1. 深度优先搜索（DFS）



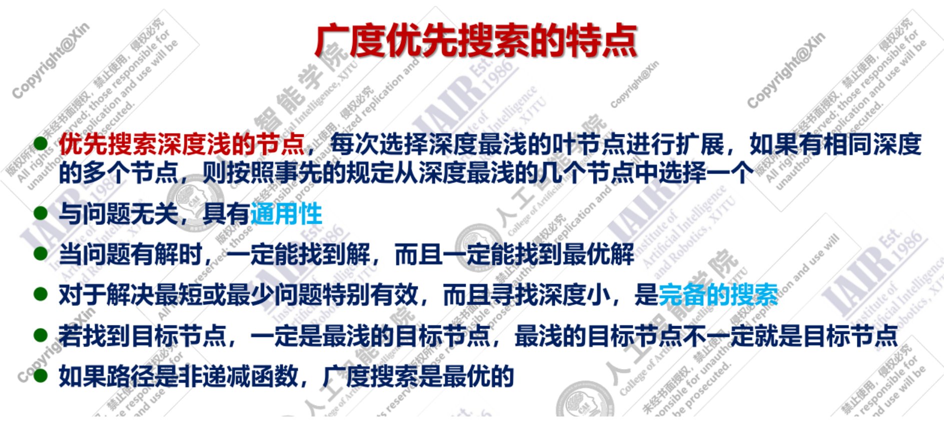


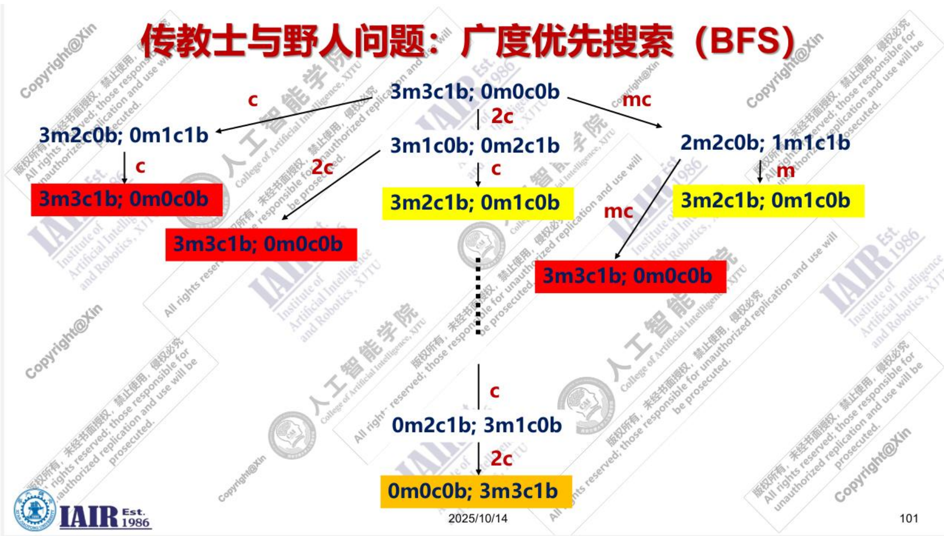
DFS 特例：回溯搜索

特点：不保留完整树结构、适用于没有明确动态规划和递归解法的问题

例子：N 皇后问题

2. 广度优先搜索（BFS）





3. DFS 与 BFS 的对比

深度优先搜索与广度优先搜索比较

	深度优先（DFS）	广度优先（BFS）
完备性	不一定（若解不在某分支，且该分支又是个无穷分支，则不可能得到解）	完备（在分支因子b有限情况下）
最优性	不具备	● 最优（如果路径代价是是节点深度的非递增函数的情况下）； ● 不一定最优（通常情况下）
时间复杂性	b^m（如果m>b时间复杂度会很高，但如果解决方案密集，可能比BFS快得多）	b^d
空间复杂性	b^m	b^d

b：分支因子；　d：解的深度
m：搜索树的最大深度　l：深度限制

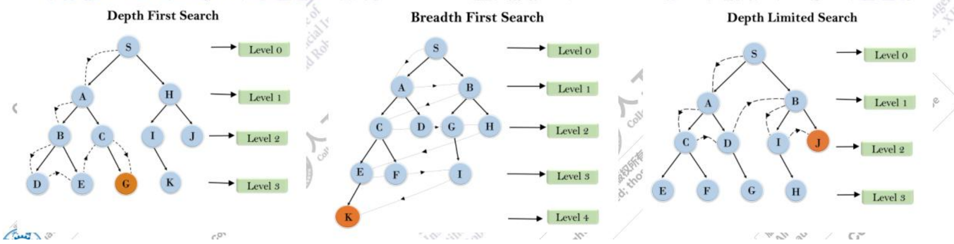
二．启发式搜索（有信息搜索）

1. 一致代价搜索 / Dijkstra 算法

   以广度优先搜索为基础，每次扩张代价最小的节点

一致代价搜索（Uniform-cost Search）/Dijkstra算法

- 一致代价搜索：总是扩展路径消耗最小的节点N，N点的路径消耗等于前一节点N-1的路径消耗加上N-1到N节点的路径消耗
- 一致代价搜索是在广度优先搜索上进行扩展的，使用优先级队列并在边缘中的状态发现更小代价的路径时引入的额外的检查，也称为代价一致搜索
- 一致代价搜索实际上是在广度优先搜索的基础上进行扩展，如果每一步的代价全部相等，则相当于广度优先



2. 爬山搜索

   每次随机选择一个位置，与邻节点比较并前往更大值。（局部最优解，能否得到全局最优解取决于初值）

3. 贪婪搜索

   A. 启发函数和评价函数

      启发函数 h(n)评估状态与目标状态的接近程度（越小表示越接近）

      已付出代价记为 g(n)

      评价函数 f(n)=g(n)+h(n)评价节点的重要程度

   B. 贪婪最佳优先搜索

      令 f(n)=h(n)，即只考虑与目标状态的接近程度，优先扩展离目标状态最近的节点

      1) 局部择优，不是最优搜索
      2) 对错误起点敏感，可能有死循环

3) 不完备

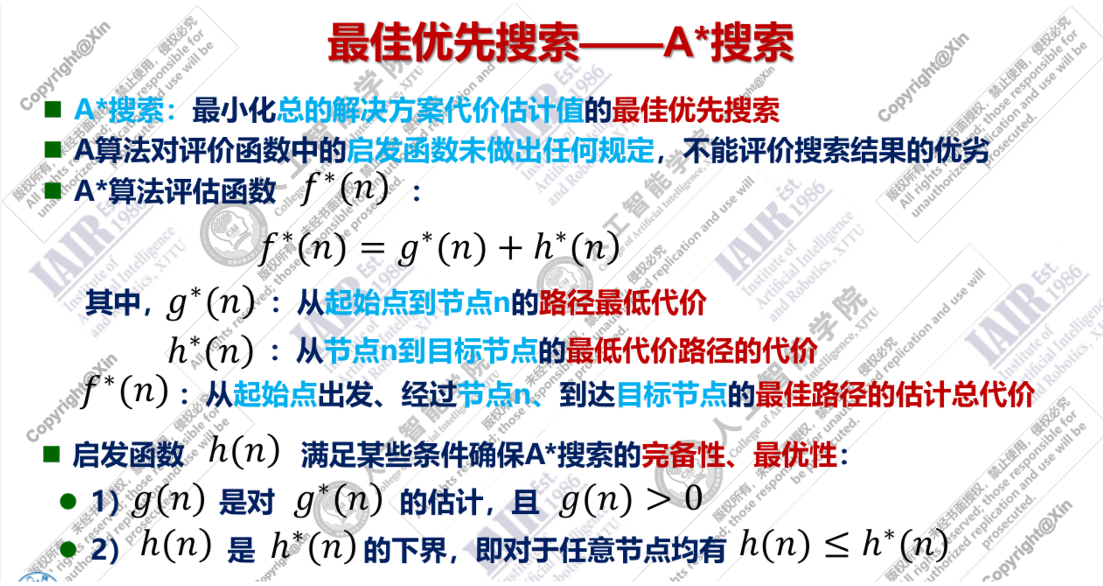
4) 时间/空间复杂度为 O(b^m)

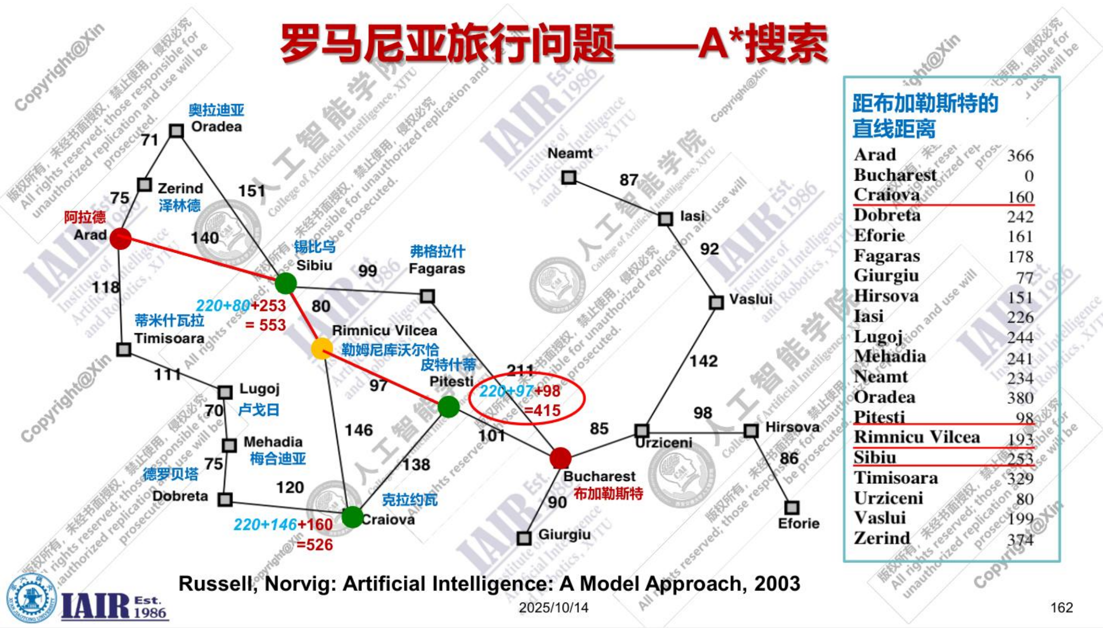
C. A 搜索

令 f(n)=g(n)+h(n)，每次选择具有最小评价函数值的节点扩展

D. A*搜索

对 A 搜索的启发函数 h(n)作了规定：

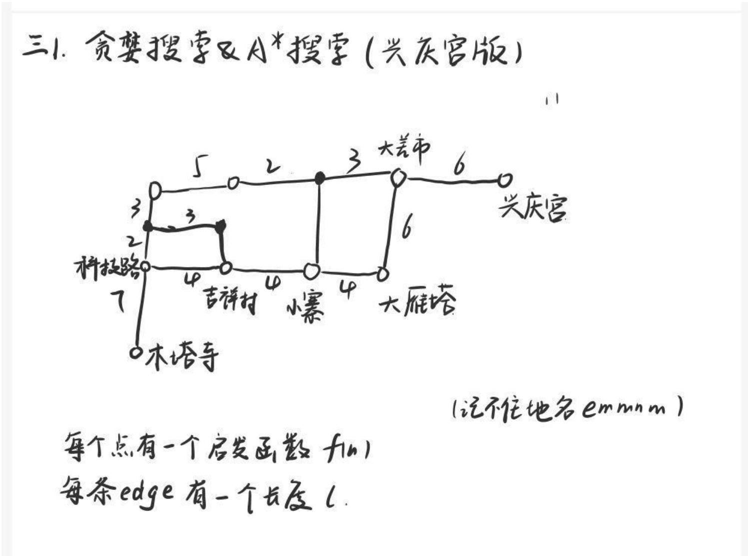




2024 年真题

（选择）盲目搜索包含：<u>深度优先搜索和广度优先搜索</u>

（简答）贪婪搜索和 A*搜索



图并不复杂，可以参考 ppt 上类似题目的求解过程

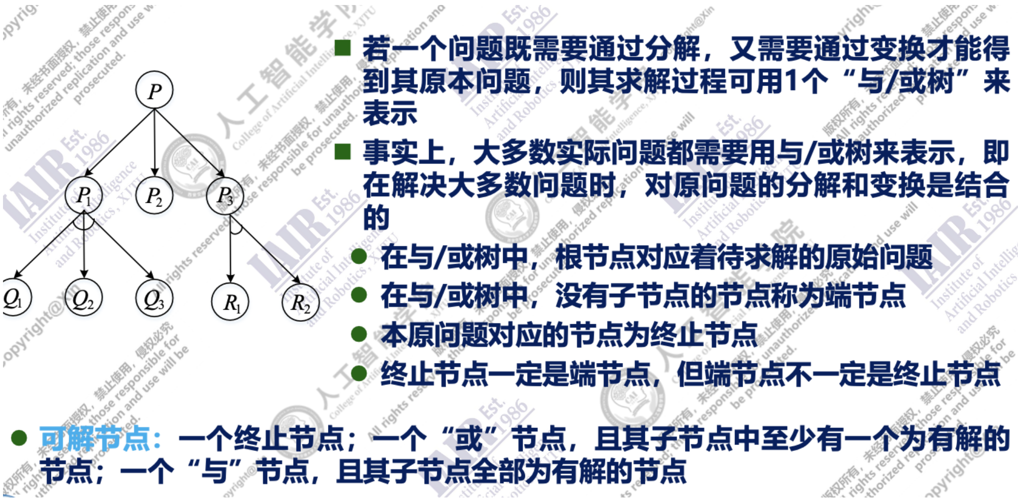
博弈搜索

一．博弈简介

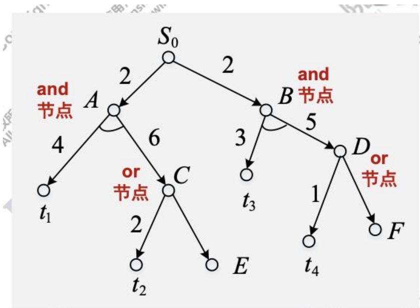
1. 人工智能领域的“博弈”：有完整信息的，确定性的、轮流行动的，两个参与者的零和（Zero-sum）游戏（总值相等而符号相反的效用函数的对立导致对抗）。
2. 博弈的现实分类：1.同时博弈 2.相继博弈 3.边界博弈 4.规则博弈 5.策略性博弈
3. 根据博弈者数量分类：单人、两人、多人博弈（单人博弈是退化的博弈）
4. 按照博弈先后顺序分类：静态博弈（参与者同时行动，或者后者不知道前者是如何行动的；例：石头剪刀布，罚点球时的守门员与罚球手）、动态博弈（参与者有先后顺序；例：下棋，拍卖会）。
5. 根据博弈中的得益分类：零和博弈（利益始终对立）、常和博弈（利益是对立的且是竞争关系）、变和博弈（合作利益存在）。
6. 根据博弈中得益的信息分类：完全信息博弈（各方都了解所有博弈方各种情况下的得益）、不完全信息博弈（至少部分博弈方不完全了解）。
7. 根据博弈过程的信息：完美信息博弈（各方对每轮进程完全了解）、不完美信息博弈（至少部分博弈方不完全了解）。
8. 根据博弈者之间是否合作：合作博弈（集体利益最大化）、非合作博弈（个人利益最大化）

二．问题归约表示（与或树）

1. 本源问题：指无法（或者不需要）再分解，且可直接解答的问题。
2. 与或树相关的概念：



3. 与或树代价的求解：



- 若节点 n 为终止节点，则其代价 $h(n)=0$
- 若节点 n 为或节点，且其子节点为 $n_1, n_2, \cdots, n_k$，则其代价

$$h(n)=\min_{1\le i\le k}\{c(n,n_i)+h(n_i)\}$$

其中，$c(n,n_i)$ 节点 n 到其子节点 n_i 的代价

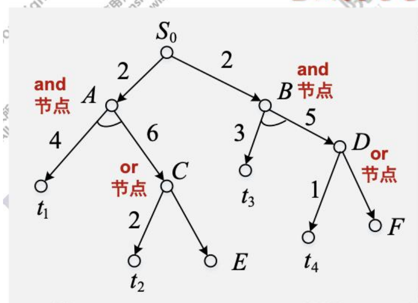
- 若节点 n 为与节点，且其子节点为 $n_1, n_2, \cdots, n_k$，则其代价

$$h(n)=\sum_{i=1}^{k}\{c(n,n_i)+h(n_i)\} \quad \text{或} \quad h(n)=\max_{1\le i\le k}\{c(n,n_i)+h(n_i)\}$$

- 若节点 n 为非终止节点的端节点，则其代价 $h(n)=\infty$

4. 与节点有两种代价计算方式（sum/max），要注意审题，判断使用哪种。
5. 例题：

与或树求解的代价计算示例



- 终止节点：t_1, t_2, t_3, t_4
- 非终止端节点：E, F
- 边上旁的数字为该边的代价

$$h(t_1)=h(t_2)=h(t_3)=h(t_4)=0$$
$$h(E)=h(F)=\infty$$

■ 左侧子树：

- 按与点求和代价计算：

$$h(C)=2,\ h(A)=12,\ h(S_0)=14$$

- 按与点最大值代价计算：

$$h(C)=2,\ h(A)=8,\ h(S_0)=10$$

■ 右侧子树：

- 按与点求和代价计算：

$$h(D)=1,\ h(B)=9,\ h(S_0)=11$$

- 按与点最大值代价计算：

$$h(D)=1,\ h(B)=6,\ h(S_0)=8$$

6. 博弈树是一种特殊的与或树，特点有：
   1) 博弈的初始状态是初始节点
   2) 博弈中 or 节点（Max 方）和 and 节点(Min 方)逐层交替出现
   3) 整个博弈过程始终站在某一方立场上（Max 方或者 Min 方）。
7. 博弈树的子树搜索完成后的返回值是针对该子树的局部最优值，但并不一定是全局最优值。

三．极小极大搜索过程

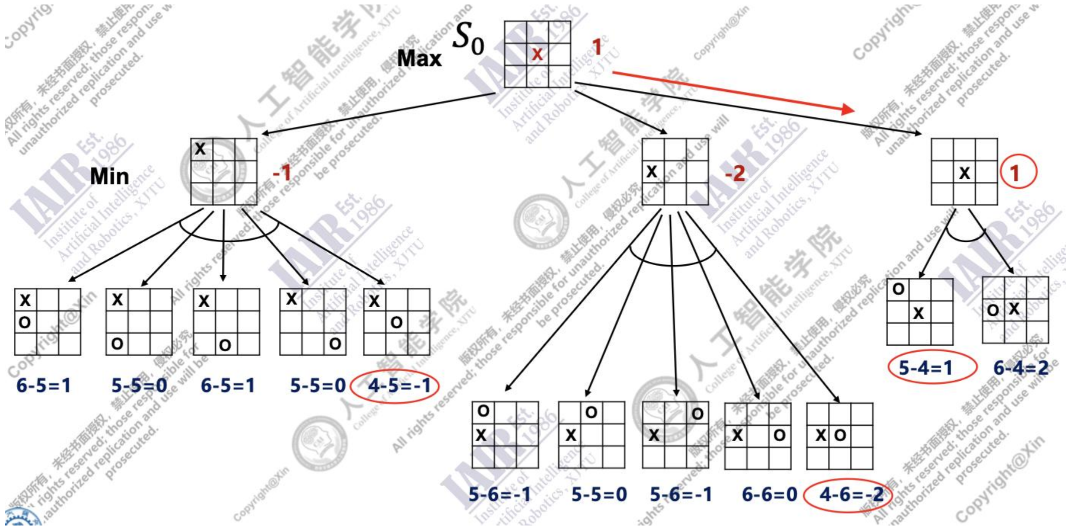
1. 生成 Max 方的 k 步博弈树
2. 定义评估函数，规定：

■ 估价函数 $e(p)$ 的规定：

- 1）若格局p是Max方获胜，则 $e(p)\rightarrow +\infty$
- 2）若格局p是Min方获胜，则 $e(p)\rightarrow -\infty$
- 3）对任何一方都不是获胜，则 $e(p)=e_{\max}(p)-e_{\min}(p)$

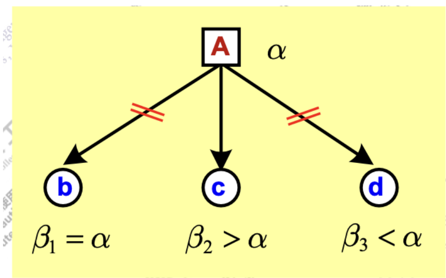
3. 用评估函数计算各叶节点的得分

4. 用极小极大运算自下而上逐层计算节点的得分（其中在 Max 处取最大值，Min 处取最小值）
5. 在根节点处结束操作
6. 例：（k=2 时的井字棋搜索过程）

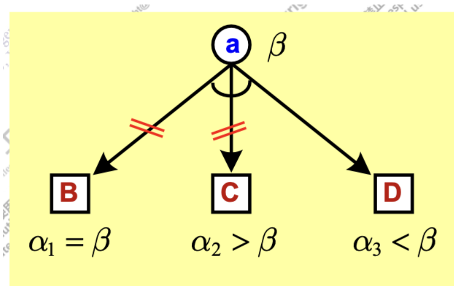


四． α-β 剪枝

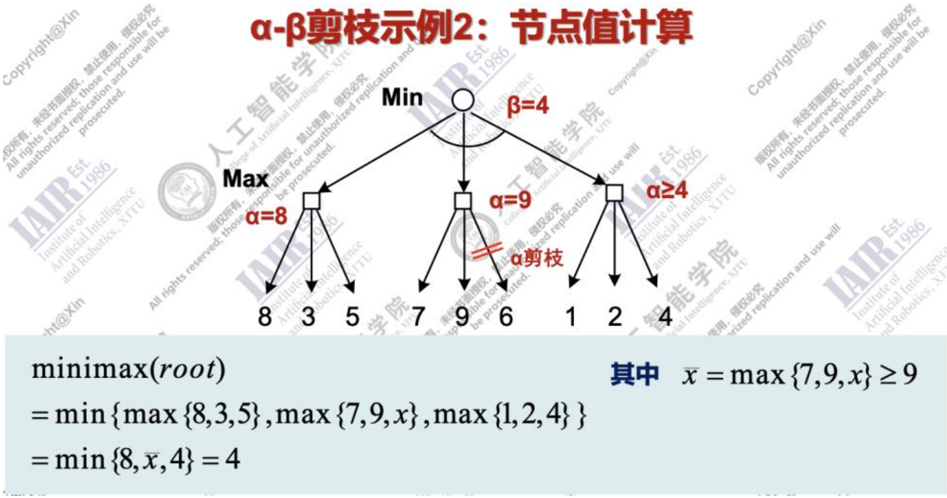
1. 基本思想：
   - **对于一个"或"（or）节点（Max方），取当前子节点中最大的倒推值作为其都推值的下界值，称为α（α ≥ 该最大值）**
   - **对于一个"与"（and）节点（Min方），取当前子节点中最小的倒推值作为其都推值得上界值，称为β（β ≤ 该最大值）**
   - **α：Max方可以搜索到的最好值，β：Min方可以接受的最坏值**
2. α剪枝规则：子节点（and 节点）的β值<=该节点（or 节点）的α值时可以直接删去该子节点。



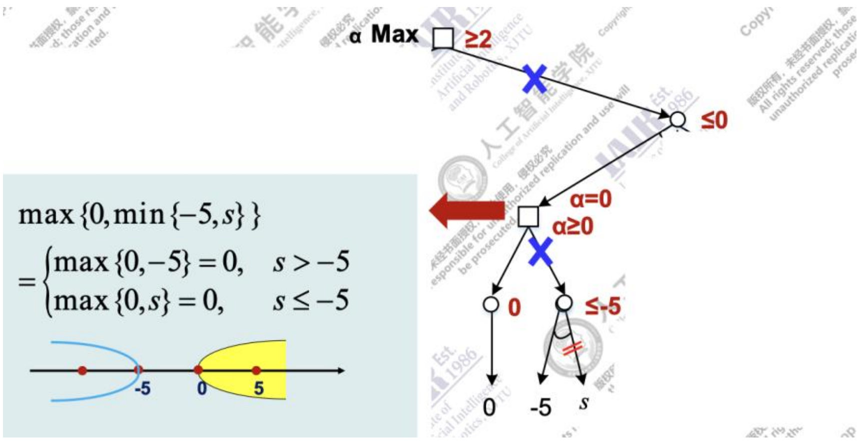
3. β剪枝规则：子节点（or 节点）的α值>=该节点（and 节点）β值时可以直接删去该子节点。



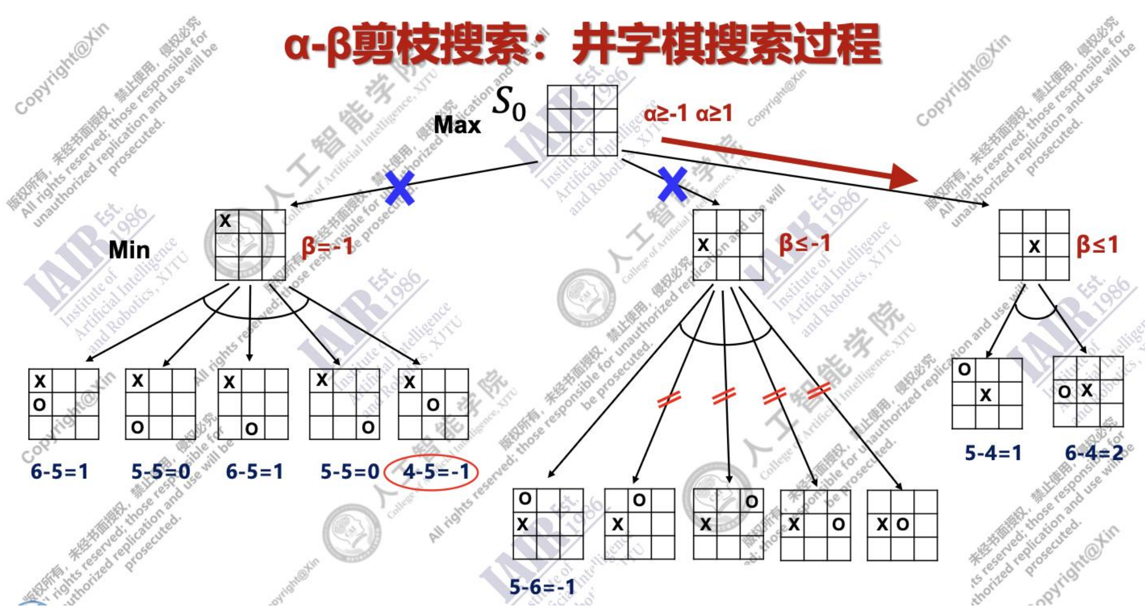
4. 可以用 max 与 min 函数来表示：设未知量为 x，发现 x 并不影响最终结果即可把 x 所在子树剪掉。



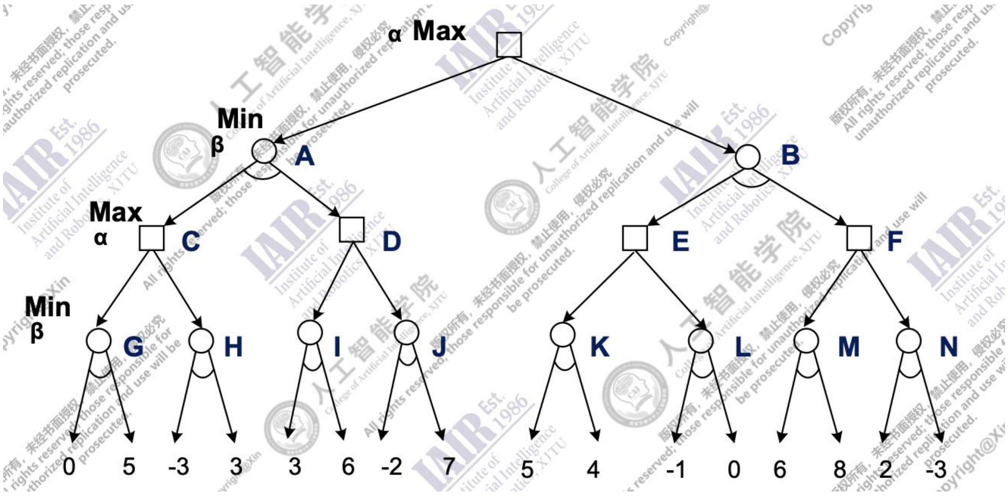
5. 每一次对节点的α或β值进行更改后，要从下到上更新所有节点的值。
6. 相关节点值计算验证：对该节点的子节点中的未知数用字母表示，然后分类讨论取值，并在数轴上表示。

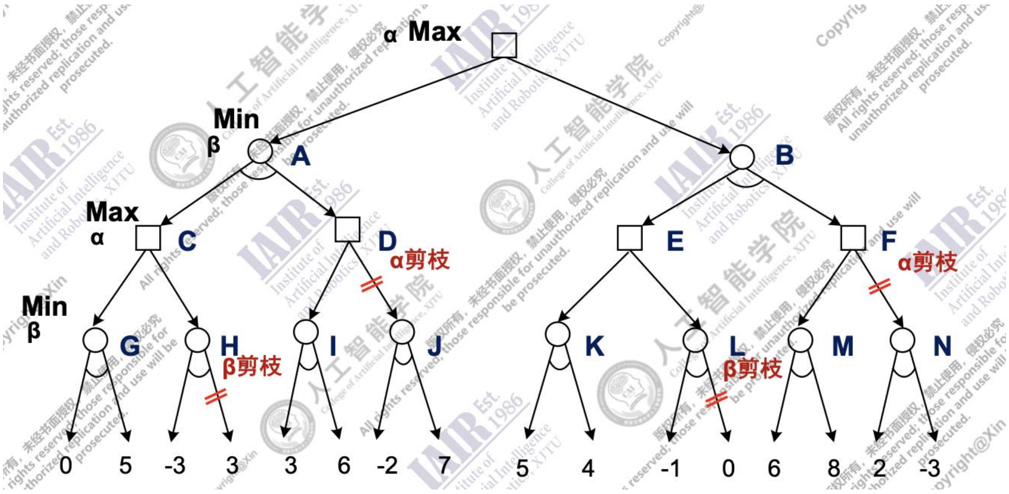


7. 例子：井字棋的 α-β 剪枝：



8. 例题：（课堂练习）





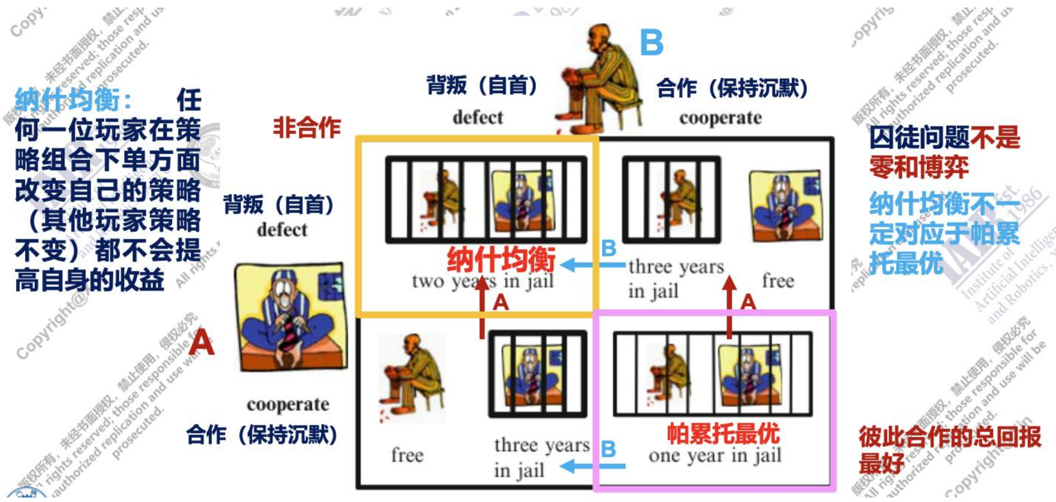
五．非合作博弈

1. 非合作博弈的分类以及对应的均衡概念：

2. 纳什均衡的概念：纳什均衡是在非合作博弈条件下，可能会形成的一种均衡状态，当参与人选定的策略组成纳什均衡后，就会形成一个平衡的局面，在这个平衡的局面中，任何一个参与人单方面地改变自己的策略，只

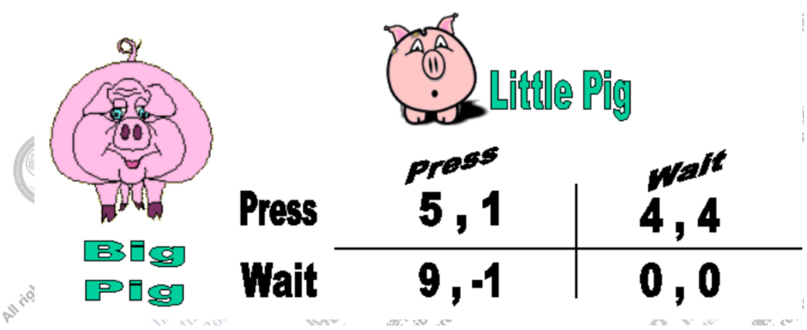
能是自己的收益下降（或不变），绝不可能使自己的收益增加

3. 并不是所有的非合作博弈都有纳什均衡，有可能没有均衡状态，而有纳什均衡存在的博弈也不见得只有一个均衡，可能存在两个或者多个均衡。
4. 囚徒困境：



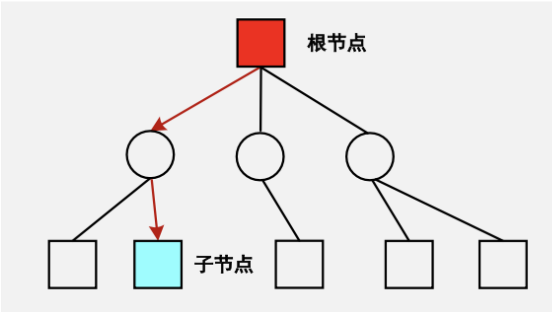
收益矩阵：可以看到，不管在 B 做出什么选择的情况下，A 选择背叛操作能够获得的收益是最大的（对 B 也是同理）。但是我们作为上帝视角可以看到，彼此合作的总回报才是最好的（帕累托最优）。

5. 智猪博弈：不管大猪怎么选，小猪总是要选择等待才会获得最大的收益（但对大猪不是这样，所以不是纳什均衡）。

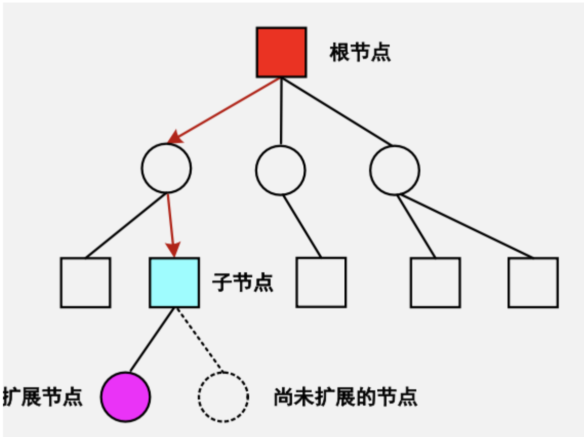


六．蒙特卡洛树搜索

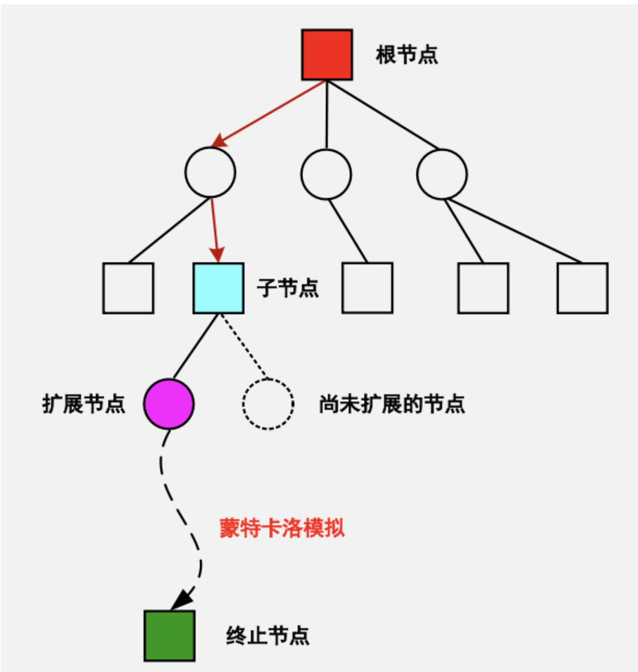
1. 不同于普通的树搜索，蒙特卡洛树搜索边模拟、边探索、边调整。
2. 既有树木搜索的精度，又具有随机抽样的普遍性。
3. 蒙特卡洛树搜索基于两个基本概念：
   - **可以使用随机模拟近似动作的真实值**
   - **这些模拟值可以有效地用于将策略调整为最佳优先策略**
4. 蒙特卡洛树搜索步骤：
   a) 选择：自上而下找到最急迫需要可扩展节点（非终节点且未被访问就是可拓展的节点；利用置信上限（UCT）公式从上到下找到“最急迫”节点）



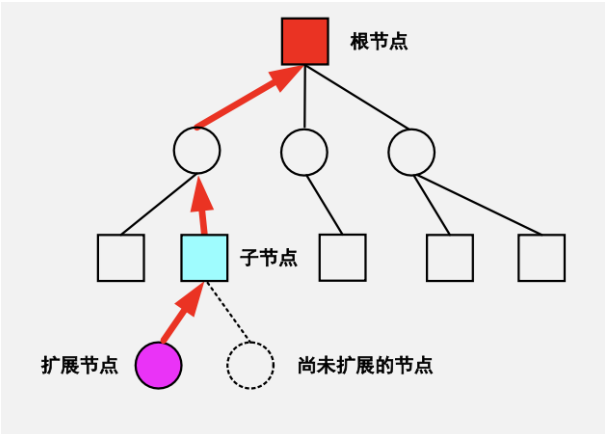
b) 扩展：在选定的节点上随机扩展一个或多个未被扩展的子节点以扩展搜索树。



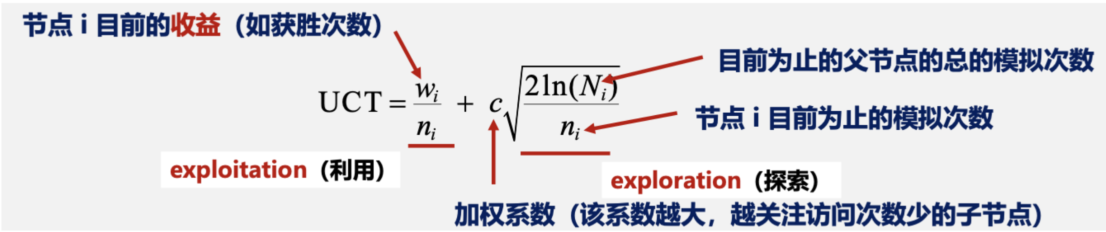
c) 模拟：在新节点上进行多次蒙特卡洛方法模拟，根据结果评估扩展节点的值函数。（蒙特卡洛模拟先随机抽样，再计算在目标区域内出现的次数）



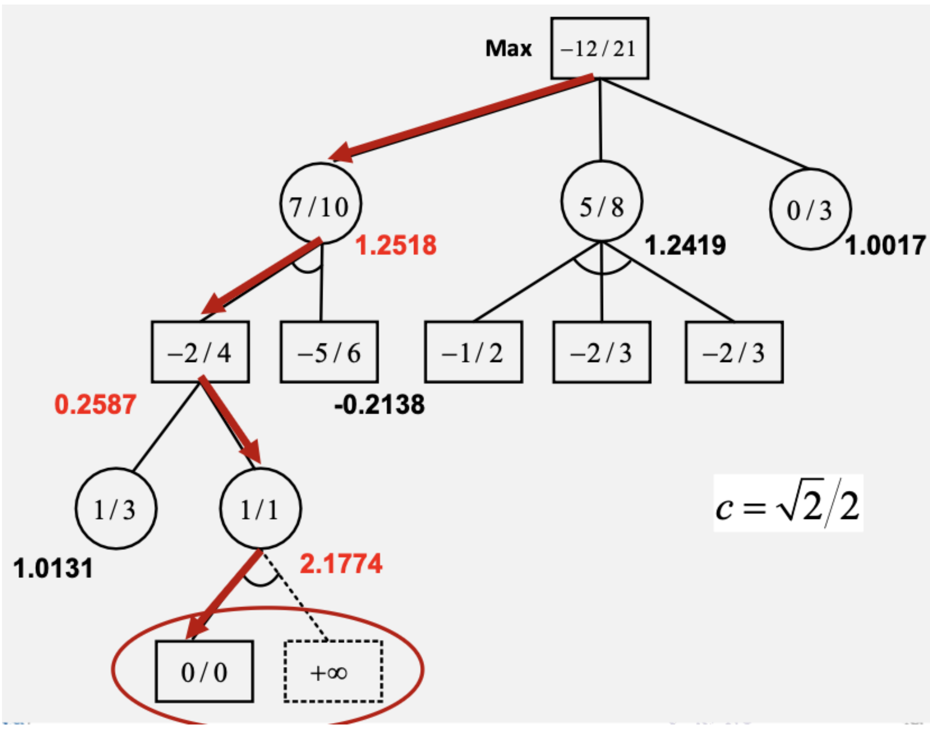
d) 回溯：根据模拟结果，向上更新路径上的节点。



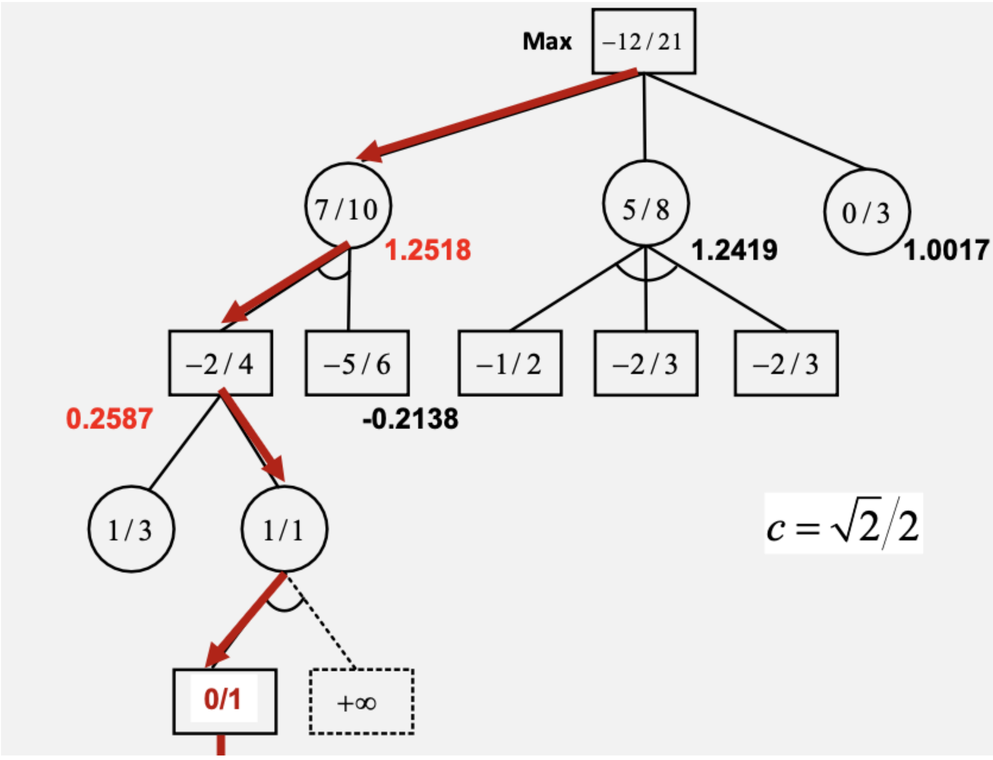
5. UCT（Upper Confidence Bounds for Trees）的计算公式（综合考虑当前胜率 exploitation、潜在胜率 exploration）：



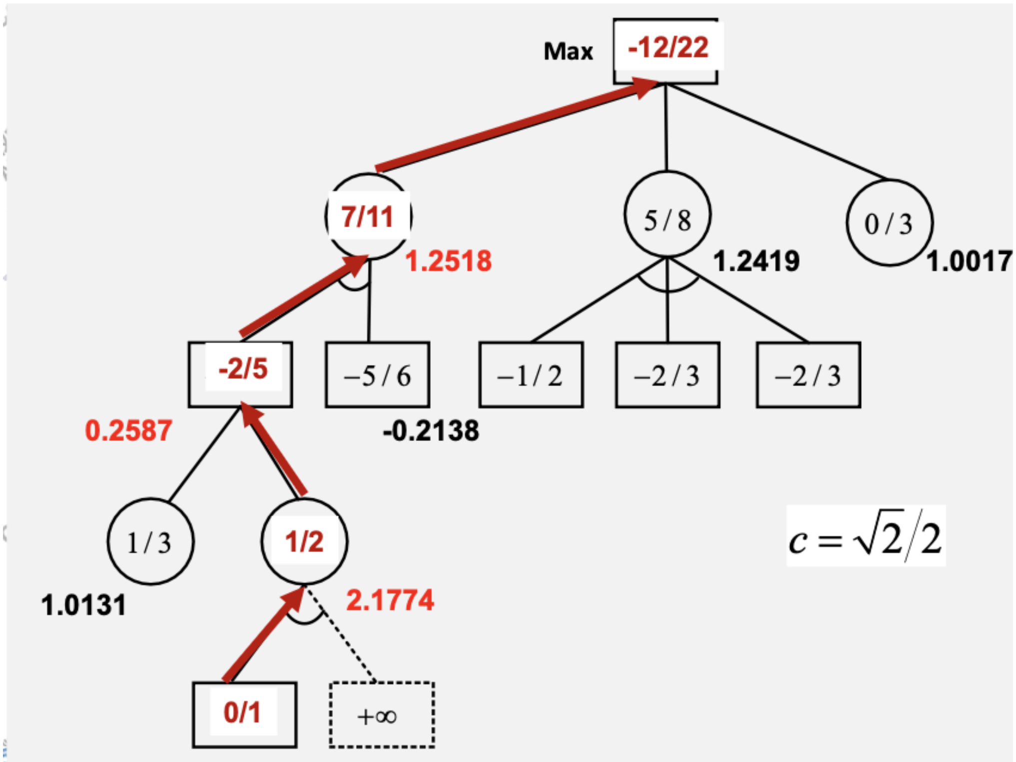
6. 搜索实例：
   a) 先找到 UCT 最大的未拓展节点并拓展其子节点



   b) 再对其子节点进行蒙特卡洛模拟（该题模拟了一次）

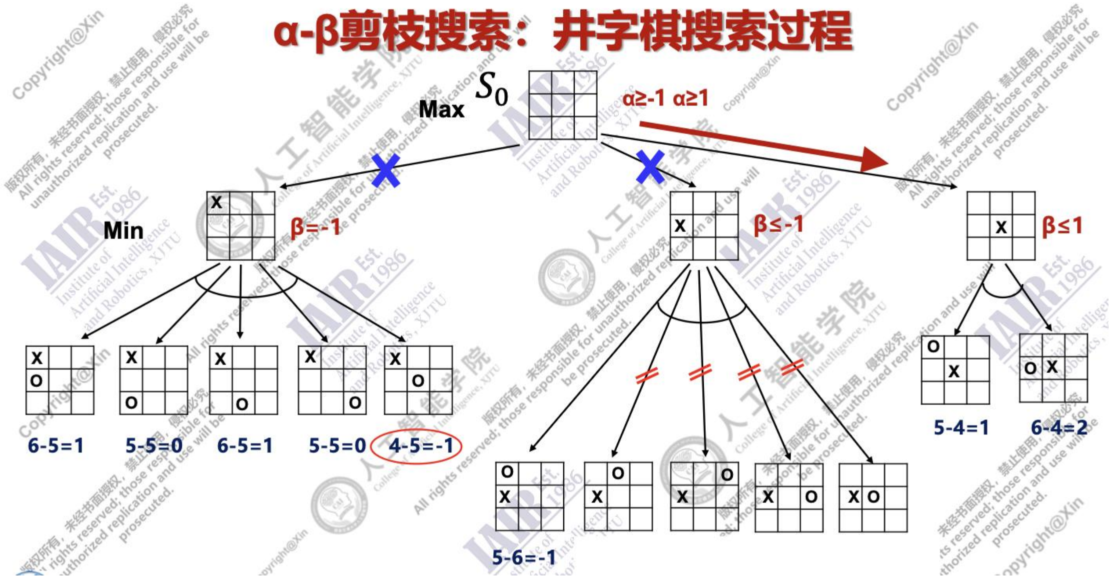


c) 再向上进行更新回溯



七．2024 年真题

1. （选择）囚徒困境不是零和博弈（一般来说是变和博弈）。
2. （简答）用极大极小化搜索求一个节点的取值，用$\alpha-\beta$剪枝对四层的博弈树进行剪枝（PPT 原图，改数字）：



想做练习的话可以看“四-8”，有课上的题目与答案。

确定性推理

一、推理基本概念

1. 推理构成：形式（前提结论逻辑关系）+内容（前提结论具体内容）
2. 逻辑形式有效性：可以从真前提（内容为真）得到真结论（但形式有效本身并不保证结论为真）

   逻辑形式可靠性：推理有效+推理前提全部为真
3. **推理分类**：

<u>按知识确定性</u>：

1）确定性推理=自然推理（直接通过逻辑规则推导）+归结演绎（永真性->不可满足性）

2）不确定推理=似然推理（概率论）+模糊推理（模糊逻辑）

<u>按逻辑基础</u>：

1）演绎推理（普遍->特殊）

2）归纳推理（特殊->普遍）

3）默认推理（知识不完全，无否定证据则认为结论成立，按照默认认识进行推理，遇到矛盾再回溯修改）

<u>按推理方向</u>：

1）正向推理（数据驱动）

2）逆向推理（目标驱动）

3）混合推理（先正后逆/先逆后正）

4）双向推理（同时正逆）

<u>按过程单调性</u>：单调（推理时一定是在增强/靠近结论）/非单调（可能发生冲突导致知识库需修改）

二、推理逻辑基础

1. **命题逻辑**：
   1) 命题：陈述句+真值唯一
   2) 原子命题：不包含联结词（五大）
   3) 文字：命题&其否定（后面有谓词逻辑可以扩展为谓词&其否定）
   4) 范式：析取/合取范式（不唯一），主析取/合取范式（唯一）

命题逻辑局限性：最小单位过大，无法处理整体&部分关系

2. **谓词逻辑**（分解拓展命题为个体和谓词，并增加量词）：
   1) 谓词：

      作用：表示个体属性&个体关系

      构成：谓词符号+个体符号（可以为函数）
   2) 量词（任意/存在）对应有量词辖域（如多量词出现，后在前辖域，这影响化简过程），约束/自由变元（影响化简），换名规则（子句集化简）
   3) 辖域收扩：只需要记住只有$\forall$对$\wedge$，$\exists$对$\vee$有分配律，其他都没有

4) 原子谓词公式：单谓词，无联结词

5) 谓词公式真值可能：永真/可满足（包含永真）/永假（不可满足）

6) 逻辑推理：（题目：已知每位航天员或者在太空空间站或者在地面，且并不是每位航天员都在太空空间站，请证明有些航天员在地面。）

解：设 $astronaut(x)$ **：** x **是航天员，** $onground(x)$ **：** x **在地面，**

$inspace(x)$ **：** x **在空间空间站**

前提： $(\forall x)(astronaut(x) \to ongroud(x) \vee inspace(x))$,

$\neg(\forall x)(astronaut(x) \to inspace(x))$

结论： $(\exists x)(astronaut(x) \wedge ongroud(x))$

$$\neg(\forall x)(astronaut(x) \to inspace(x)) \Rightarrow (\exists x)\neg(astronaut(x) \to inspace(x))$$

$$\Rightarrow (\exists x)\neg(\neg astronaut(x) \vee inspace(x))$$

$$\Rightarrow (\exists x)(astronaut(x) \wedge \neg inspace(x))$$

$\Rightarrow astronaut(a) \wedge \neg inspace(a)$ **a是航天员，a不在太空空间站**

$\Rightarrow$ $\neg inspace(a)$ **航天员a不在太空空间站**

$$(\forall x)(astronaut(x) \to ongroud(x) \vee inspace(x))$$

$$\Rightarrow astronaut(a) \to ongroud(a) \vee inspace(a)$$

$\Rightarrow$ $ongroud(a) \vee inspace(a)$ **航天员a在太空空间站或地面**

$$\left.\begin{array}{l} \neg inspace(a) \\ ongroud(a) \vee inspace(a) \end{array}\right\} \Rightarrow ongroud(a)$$ **航天员a在地面**

$\Rightarrow$ $astronaut(a) \wedge ongroud(a)$

a是航天员，a在地面

$\Rightarrow$ $(\exists x)(astronaut(x) \wedge ongroud(x))$

某些航天员在地面

三、自然演绎推理

1. 演绎推理为必然性推理，不产生新知识（只产生隐性知识）
2. 演绎推理内容：
   1) 三段论：大前提（一般性知识）+小前提（个别实例特殊性知识）=>针对实例特殊判断（即 A->B∧B->C=>A->C）
   2) 假言推理（演绎）：基本思想：大前提（条件句：如果（条件）就（结果））+小前提（条件）=>结果

      分类：

      a) 充分条件假言推理：肯定前件=>肯定后件，否定后件=>否定前件
      b) 必要条件假言推理：肯定后件=>肯定前件，否定前件=>否定后件
      c) 充分必要条件假言推理：肯定/否定前件<=>肯定/否定后件
3. **自然演绎**（从已知事实出发，按照命题/谓词逻辑规则进行推理，得到结论）：

组成：1)假言推理（P,P->Q=>Q）2)拒取式推理（¬Q,P->Q=>¬P）

3)假言三段论（P->Q，Q->R=>P->R）

四. 归纳演绎推理

分类:

1. 完全归纳:考查了每一个实例(所有对象/子类),是必然性推理

2. 不完全归纳(或然性推理):只考查了部分实例,包括简答枚举+科学归纳法(基于简单枚举,加入科学分析)

五. 归结演绎推理

1. 用途:定理证明&问题求解
2. 基础知识:
   1) 归纳演绎流程:谓词公式->前束范式->斯科伦范式->子句集->扩充子句集(结论证明:加入结论否定/问题求解:加入问题否定∨答案容器)->归结直到 NIL(空子句)
   2) 谓词公式范式:1° 前束范式(量词(全称和存在,不能有否定)只出现在最前面,母式无量词,联结词不能有单向/双向蕴含) 2° 斯科伦范式(在前束范式基础上,再消去所有存在量词,且将母式化成析取范式)
   3) **前束范式化简求解步骤**(从复合谓词公式开始):

Step1:消去→&↔

Step2:减少¬(即双重否定等价肯定)

Step3:¬深入(化出去的量词不能包含否定,使用¬∀等价于∃,¬∃等价于∀)

Step4:约束易名(所有约束变量都尽量不同,一般都有)(只要是一个量词辖域中出现与其同名的都需要进行更换)

Step5:辖域扩张(即量词前移)

注:这样化简出来的前束范式未必唯一,化简时需要注意量词之间的辖域关系

4) **斯科伦范式化简求解步骤**(从前束范式开始):

Step1:消去全部存在量词:

1)若存在量词左侧无任何全称量词,直接使用一个未使用过的个体常量在谓词公式中进行替代,并删去对应存在量词

2)若存在量词左侧有全称量词,则构建一个斯科伦函数(变量是其左侧全部全称量词)(注意多个斯科伦函数名不能相互冲突,且也不能与现存函数/谓词等冲突)

Step2:将母式化为合取范式(易漏)

5) 文字:原子谓词公式&其否定,互补文字:P(正文字)&¬P(负文字)
6) 子句:文字析取式
7) 子句集:字句合取式(默认其中全部变元都受全称量词约束,因此外面量词省略,花括号括起来,里面析取式之间用逗号隔开)
8) 空子句(NIL):无文字子句,永假(无法被任何解释满足)
9) **子句集化简求解步骤**(从斯科伦范式开始):

Step1:直接省略掉母式前的全部全称量词,并将母式改写成子句集形式(花括

号扩起，相邻子句之间使用逗号进行分隔）

Step2:重命名（子句集中任两个子句不含相同名称变量）

■ **将下列谓词公式化为子句集**

$$(\forall x)((\neg P(x)\vee\neg Q(x))\rightarrow(\exists y)(S(x,y)\wedge Q(x)))\wedge(\forall x)(P(x)\vee B(x))$$

- **1）消去蕴含符号** $\alpha\rightarrow\beta\equiv\neg\alpha\vee\beta$

$$\Leftrightarrow(\forall x)(\neg(\neg P(x)\vee\neg Q(x))\vee(\exists y)(S(x,y)\wedge Q(x)))\wedge(\forall x)(P(x)\vee B(x))$$

- **2）把否定符号移到每个谓词前面** $\neg(\neg\alpha)\equiv\alpha$

$$\Leftrightarrow(\forall x)((P(x)\wedge Q(x))\vee(\exists y)(S(x,y)\wedge Q(x)))\wedge(\forall x)(P(x)\vee B(x))$$

- **3）变量标准化，重新命名变元**

$$\Leftrightarrow(\forall x)((P(x)\wedge Q(x))\vee(\exists y)(S(x,y)\wedge Q(x)))\wedge(\forall w)(P(w)\vee B(w))$$

- **4）消去存在量词，引入Skolem函数**

$$\Leftrightarrow(\forall x)((P(x)\wedge Q(x))\vee(S(x,f(x))\wedge Q(x)))\wedge(\forall w)(P(w)\vee B(w))$$

- **5）化为前束形**

$$\Leftrightarrow(\forall x)(\forall w)(((P(x)\wedge Q(x))\vee(S(x,f(x))\wedge Q(x)))\wedge(P(w)\vee B(w)))$$

- **6）化为标准形** $P\wedge(Q\vee R)\equiv(P\wedge Q)\vee(P\wedge R)$

$$\Leftrightarrow(\forall x)(\forall w)(((Q(x)\wedge P(x))\vee(Q(x)\wedge S(x,f(x))))\wedge(P(w)\vee B(w)))$$

$$\Leftrightarrow(\forall x)(\forall w)((Q(x)\wedge(P(x)\vee S(x,f(x))))\wedge(P(w)\vee B(w)))$$

$$\Leftrightarrow(\forall x)(\forall w)(Q(x)\wedge(P(x)\vee S(x,f(x)))\wedge(P(w)\vee B(w)))$$

仅全称量词 **简单析取式（或）** **合取范式（与）**

- **7）略去全称量词，导出母式**

$$Q(x)\wedge(P(x)\vee S(x,f(x)))\wedge(P(w)\vee B(w))$$

- **8）子句集变量标准化，使不同子式中的变元不同名**

$$Q(x)\wedge(P(y)\vee S(y,f(y)))\wedge(P(w)\vee B(w))$$

- **9）消去合取词，用子句集表示**

$$\{Q(x),P(y)\vee S(y,f(y)),P(w)\vee B(w)\}$$

可以看出，子句集也未必唯一

10) 子句集不可满足：对任意解释，子句集中子句都不能同时取得真值

11) 定理: G 为谓词公式，S 为其对应子句集，则在不可满足性上二者等价（G 不可满足<=>S 不可满足）

12) 海伯伦域:

构造方法：H_0 为 S（上文子句集）中所有个体常项构成集合（若无，往该集合中随意增加任一个个体常量）

$H_{i+1}=H_i\cup${S 中函数在 H_i 上全部实例}（函数变元遍取 H_i 全部元素）

13) 子句集中一旦出现 NIL，因为子句之间合取的关系，子句集必然是不可满足的

14) 归结原理（归结树展示）：

1. 命题逻辑：C1，C2 为子句集中两子句，C1 中 L1 与 C2 中 L2 为互补文字，则可将 L1 从 C1 中剔除，将 L2 从 C2 中剔除，将 L1,L2 剩余部分析取合成为 C12（C12 实为 C1&C2 推论）
2. 谓词逻辑：

   1° 若谓词无变量，退化同 1】

   2° 否则，就利用子句集中所有变元都受全称量词约束性质，为了进行上述归结操作对变量进行合适赋值，化为 1】

   【最一般合一】：达成合一效果（置换，使得谓词公式表达一致），且赋值操作数最少
3. 归结原理具体应用（在以上基础知识背景下）

应用 1（归结反演证明结论）：

Step1: 将前提表为谓词公式 P

Step2: 将待证明结论表为谓词公式 Q

Step3: 将{P，¬Q}（将永真转为不可满足证明）化为子句集 S

Step4: 运用以上归结原理（14），每次归结产生结果放入子句集中

Step5: 归结直到 NIL 出现，停止，归结反演证明结论成功

■ **利用归结原理如下推论证明:**
“有些学生喜欢任一教师。没有任一学生喜欢任一庸师。所以没有庸师的教师”

证明：1）定义谓词，并将前提用谓词公式表示

$student(x)$ **：x 是学生**

$teacher(x)$ **：x 是教师**

$quacks(x)$ **：x 是庸师**

$like(x,y)$ **：x 喜欢 y**

$$C_1:\ (\exists x)(student(x)\wedge(\forall y)(teacher(y)\rightarrow like(x,y)))$$

$$C_1:\ (\forall x)(student(x)\rightarrow(\forall y)(quacks(y)\rightarrow\neg like(x,y)))$$

$$G:\ (\forall x)(teacher(x)\rightarrow\neg quacks(x))$$

$$\neg G:\ \neg(\forall x)(teacher(x)\rightarrow\neg quacks(x))$$

3）将前提和结论化为子句集

$$C_1:\ (\exists x)(student(x)\wedge(\forall y)(teacher(y)\rightarrow like(x,y)))$$
$$\Leftrightarrow(\exists x)(\forall y)(student(x)\wedge(\neg teacher(y)\vee like(x,y)))$$
$$\Leftrightarrow(\forall y)(student(a)\wedge(\neg teacher(y)\vee like(a,y)))$$

$$C_2:\ (\forall x)(student(x)\rightarrow(\forall y)(quacks(y)\rightarrow\neg like(x,y)))$$
$$\Leftrightarrow(\forall x)(student(x)\rightarrow(\forall y)(\neg quacks(y)\vee\neg like(x,y)))$$
$$\Leftrightarrow(\forall x)(\forall y)(\neg student(x)\vee(\neg quacks(y)\vee\neg like(x,y)))$$

$$\neg G:\ \neg(\forall x)(teacher(x)\rightarrow\neg quacks(x))$$
$$\Leftrightarrow\neg(\forall x)(\neg teacher(x)\vee\neg quacks(x))\Leftrightarrow teacher(b)\wedge quacks(b)$$

4）得到 $\{C_1, C_2, \neg G\}$ 子句集

$$\{student(a), \neg teacher(y) \vee like(a,y), \neg student(x) \vee \neg quacks(y) \vee \neg like(x,y), teacher(b), quacks(b)\}$$

$$\Rightarrow \{student(a), \neg teacher(y) \vee like(c,y), \neg student(x) \vee \neg quacks(z) \vee \neg like(x,z), teacher(b), quacks(d)\}$$

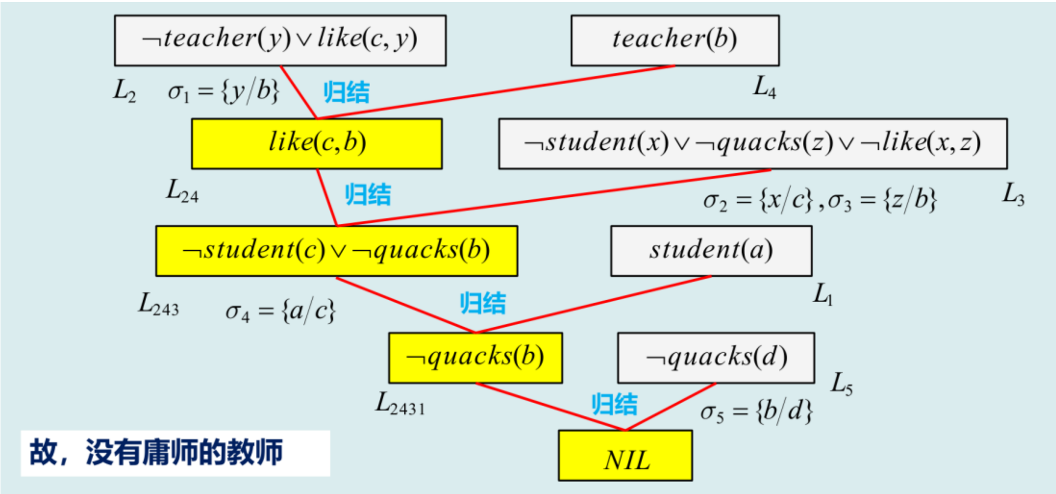
$L_1:\ student(a)$

$L_2:\ \neg teacher(y) \vee like(c,y)$

$L_3:\ \neg student(x) \vee \neg quacks(z) \vee \neg like(x,z)$

$L_4:\ teacher(b)$

$L_5:\ quacks(d)$



故，没有庸师的教师

应用 2（归结反演求解问题）（思想和应用 1 一样，Answer(x)仅为了省去回退而存在）：

Step1：将前提表为谓词公式 F，并化为子句集 S

Step2：将需要求解问题形式表示为 Q（x）（如需要求解谁是小偷，可以设置为 Q(x)：x 为小偷），得到¬Q(x)∨Answer(x)

Step3：将¬Q(x)∨Answer(x)加入 S，得到扩充子句集 S'

Step4：对 S'使用归结原理

Step5：最后得到结果只剩 Answer(个体常量）这个个体常量就是答案（可以不唯一，重复多次使用归结原理）

- **已知A、B、C三人中有个人从不说真话，也有人从不说假话。有人向这三个人分别提出同一问题："谁是说谎者？"，三个人回答如下：**

A答："B和C都是说谎者"；

B答："A和C都是说谎者"；

C答："A和B中至少有一个是说谎者"。

求证谁是老实人、谁是说谎者？

证明：1）定义谓词，并将已知前提谓词公式表示

$T(x)$ **：x说真话**

- **A答："B和C都是说谎者"**
  - **A说真话：** $A_1:\ T(A) \rightarrow (\neg T(B) \wedge \neg T(C))$
  - **A说假话：** $A_2:\ \neg T(A) \rightarrow (T(B) \vee T(C))$

- **B答：“A和C都是说谎者”**
  - **B说真话：** $B_1: \; T(B) \to (\neg T(A) \wedge \neg T(C))$
  - **B说假话：** $B_2: \; \neg T(B) \to (T(A) \vee T(C))$
- **C答：“A和B中至少有一个是说谎者”**
  - **C说真话：** $C_1: \; T(C) \to (\neg T(A) \vee \neg T(B))$
  - **C说假话：** $C_2: \; \neg T(C) \to (T(A) \wedge T(B))$

2）应用等价关系，将上述谓词公式范式化

$$A_1: \; T(A) \to (\neg T(B) \wedge \neg T(C)) \Leftrightarrow \neg T(A) \vee (\neg T(B) \wedge \neg T(C))$$
$$\Leftrightarrow \underline{(\neg T(A) \vee \neg T(B))} \wedge \underline{(\neg T(A) \vee \neg T(C))}$$

$$A_2: \neg T(A) \to (T(B) \vee T(C)) \Leftrightarrow T(A) \vee (T(B) \vee T(C))$$
$$\Leftrightarrow \underline{T(A) \vee T(B) \vee T(C)}$$

$$B_1: T(B) \to (\neg T(A) \wedge \neg T(C)) = \neg T(B) \vee (\neg T(A) \wedge \neg T(C))$$
$$\Leftrightarrow (\neg T(B) \vee \neg T(A)) \wedge \underline{(\neg T(B) \vee \neg T(C))}$$

$$B_2: \; \neg T(B) \to (T(A) \vee T(C)) \Leftrightarrow T(B) \vee (T(A) \vee T(C))$$
$$\Leftrightarrow T(B) \vee T(A) \vee T(C)$$

$$C_1: \; T(C) \to (\neg T(A) \vee \neg T(B)) \Leftrightarrow \underline{\neg T(C) \vee \neg T(A) \vee \neg T(B)}$$

$$C_2: \; \neg T(C) \to (T(A) \wedge T(B)) \Leftrightarrow T(C) \vee (T(A) \wedge T(B))$$
$$\Leftrightarrow \underline{(T(C) \vee T(A))} \wedge \underline{(T(C) \vee T(B))}$$

3）将上述谓词公式化为子句集

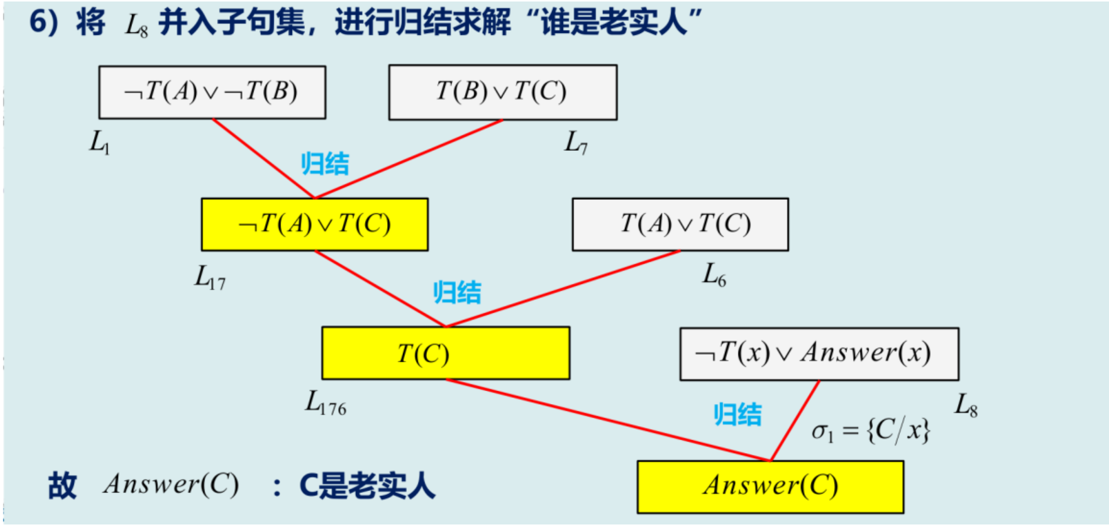
$L_1: \; \neg T(A) \vee \neg T(B)$
$L_2: \; \neg T(A) \vee \neg T(C)$
$L_3: \; T(A) \vee T(B) \vee T(C)$
$L_4: \; \neg T(B) \vee \neg T(C)$
$L_5: \; \neg T(C) \vee \neg T(A) \vee \neg T(B)$
$L_6: \; T(C) \vee T(A)$
$L_7: \; T(C) \vee T(B)$

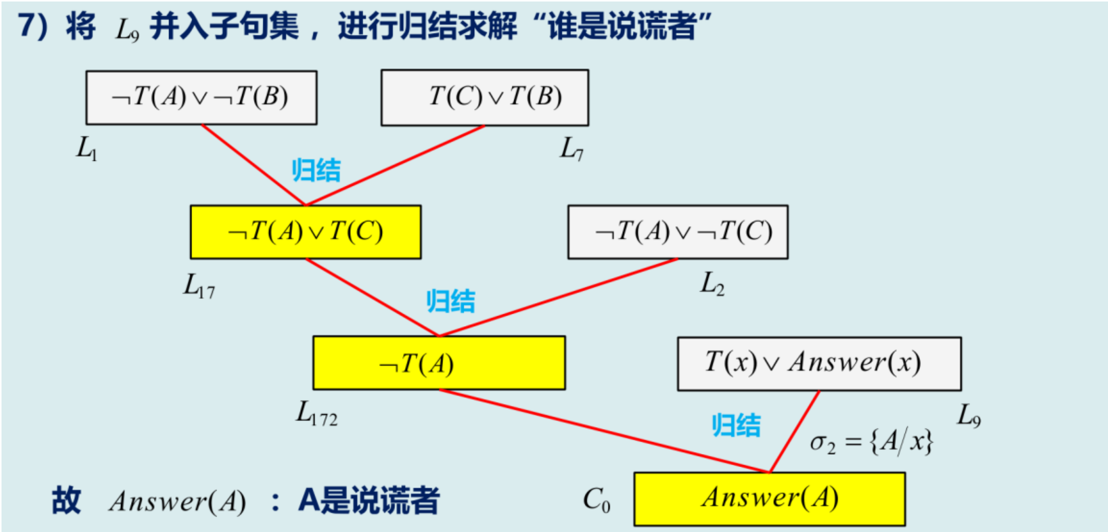
4）将“谁是老实人”表示为谓词公式，并否定其后再与谓词 $Answer$ 构成析取式

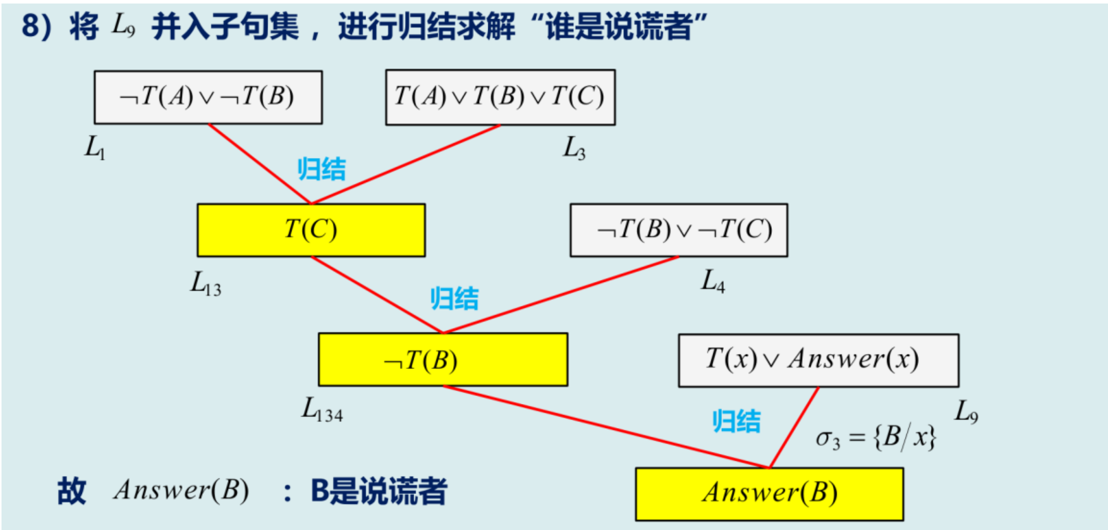
$$L_8: \; \neg T(x) \vee Answer(x)$$

5）将“谁是说谎者”表示为谓词公式，并否定其后再与谓词 $Answer$ 构成析取式

$$L_9: \; T(x) \vee Answer(x)$$







六. 2024 年真题:

选择题:

1】下列四个哪个不对（）

A. P—＞Q 相关（未回忆具体内容）

B. 给一个谓词公式，说它不是前束范式。（未回忆具体公式）

C. 小明在北京记为 P，在西安记为 Q，那么在西安或北京记为 PVQ

D. 推理方向包括前向，逆向，混合和双向。

解析：C 应使用排斥或而不是相容或（小明不可能既在北京又在西安）

简答题：

3. 化简为斯科伦范式。

$$(\forall x)((\neg P(x) \vee \neg Q(x)) \rightarrow (\exists y)(S(x,y) \wedge Q(x))) \wedge (\forall x)(P(x) \vee B(x))$$

略答：3. 解：$(\forall x)((\neg P(x)\vee\neg Q(x))\rightarrow(\exists y)(S(x,y)\wedge Q(x)))\wedge(\forall x)(P(x)\vee B(x))$

$\Leftrightarrow(\forall x)(\neg(\neg P(x)\vee\neg Q(x))\vee(\exists y)(S(x,y)\wedge Q(x)))\wedge(\forall x)(P(x)\vee B(x))$ （蕴涵消除）

$\Leftrightarrow(\forall x)((P(x)\wedge Q(x))\vee(\exists y)(S(x,y)\wedge Q(x)))\wedge(\forall x)(P(x)\vee B(x))$ （德摩根律）

$\Leftrightarrow(\forall x)(\exists y)((P(x)\wedge Q(x))\vee(S(x,y)\wedge Q(x)))\wedge(\forall z)(P(z)\vee B(z))$ （辖域扩张）

$\Leftrightarrow(\forall x)(\exists y)((P(x)\vee S(x,y)\wedge Q(x)))\wedge(Q(x)\vee(S(x,y)\wedge Q(x))))\wedge(\forall z)(P(z)\vee B(z))$ （分配律）

$\Leftrightarrow(\forall x)(\exists y)((P(x)\vee S(x,y))\wedge Q(x))\wedge(\forall z)(P(z)\vee B(z))$ （运用两次吸收律）

$\Leftrightarrow(\forall x)(\exists y)(\forall z)(((P(x)\vee S(x,y))\wedge Q(x))\wedge(P(z)\vee B(z)))$ （辖域扩张 化为前束范式）

$\Leftrightarrow(\forall x)(\exists y)(\forall z)((P(x)\vee S(x,y))\wedge Q(x)\wedge(P(z)\vee B(z)))$ （母式化为合取范式）

$\Leftrightarrow(\forall x)(\forall z)((P(x)\vee S(x,f(x)))\wedge Q(x)\wedge(P(z)\vee B(z)))$ （引入Skolem函数 消去存在量词得到斯科伦范式）

2】（1）小李的老师是王老师

（2）小李和小张是同学

（3）若 x,y 为同学，则 x 的老师为 y 的老师

将以上三信息用谓词逻辑表示，并归结反演求出小张的老师是谁。

4. 解：设谓词 $T(x,y)$ 表示 x 的老师是 y　$C(x,y)$ 表示：x 和 y 是同学

设个体常项：小李表示为 L、小张表示为 Z，王老师表示为 W

则：(1) 可表示为 $T(L,W)$ ①　(2) 可表示为 $C(L,Z)$ ②

(3) 可表示为 ~~$(\forall x)(\forall$~~ $(\forall x)(\exists m)(\forall y)((C(x,y)\wedge T(x,m))\rightarrow T(y,m))$ ③

将 ③ 化为子句集 S：~~③ $\Leftrightarrow(\forall x)(\exists m)(\forall y)((C(x,y)\wedge T(x,m))\rightarrow T(y,m))$ (蕴~~

③ $\Leftrightarrow(\forall x)(\exists m)(\forall y)(\neg(C(x,y)\wedge T(x,m))\vee T(y,m))$ （消除蕴涵）

$\Leftrightarrow(\forall x)(\exists m)(\forall y)(\neg C(x,y)\vee\neg T(x,m)\vee T(y,m))$ （德摩根律）

$\Leftrightarrow(\forall x)(\forall y)(\neg C(x,y)\vee\neg T(x,f(x))\vee T(y,f(x)))$ （引入Skolem函数 消去存在量词得到斯科伦范式）

$\Leftrightarrow S=\{\neg C(x,y)\vee\neg T(x,f(x))\vee T(y,f(x))\}$ （化为子句集）

要求解的问题为 $T(Z,t)$ 则将 $\neg T(Z,t)\vee Answer(t)$、①、② 并入 S 形成扩充子句集 $S'=$
$\{\neg C(x,y)\vee\neg T(x,f(x))\vee T(y,f(x)),\neg T(Z,t)\vee Answer(t),T(L,W),C(L,Z)\}$

下面对 S' 进行归结求解问题：

$\boxed{\neg C(x,y)}\vee\neg T(x,f(x))\vee T(y,f(x))$　　$\boxed{C(L,Z)}$

$x=L,y=Z$

$\boxed{\neg T(L,f(L))}\vee T(Z,f(L))$　　$\boxed{T(L,W)}$

$f(L)=W$

$\boxed{T(Z,W)}$　　$\boxed{\neg T(Z,t)}\vee Answer(t)$

$t=W$

$Answer(W)$

因此，归结反演得出小张的老师为王老师

不确定性推理

似然推理（概率论）

一、概率推理

1. 全概率公式

若$\{B_i\}$为完备事件组（互不相交且构成全体样本空间），则有

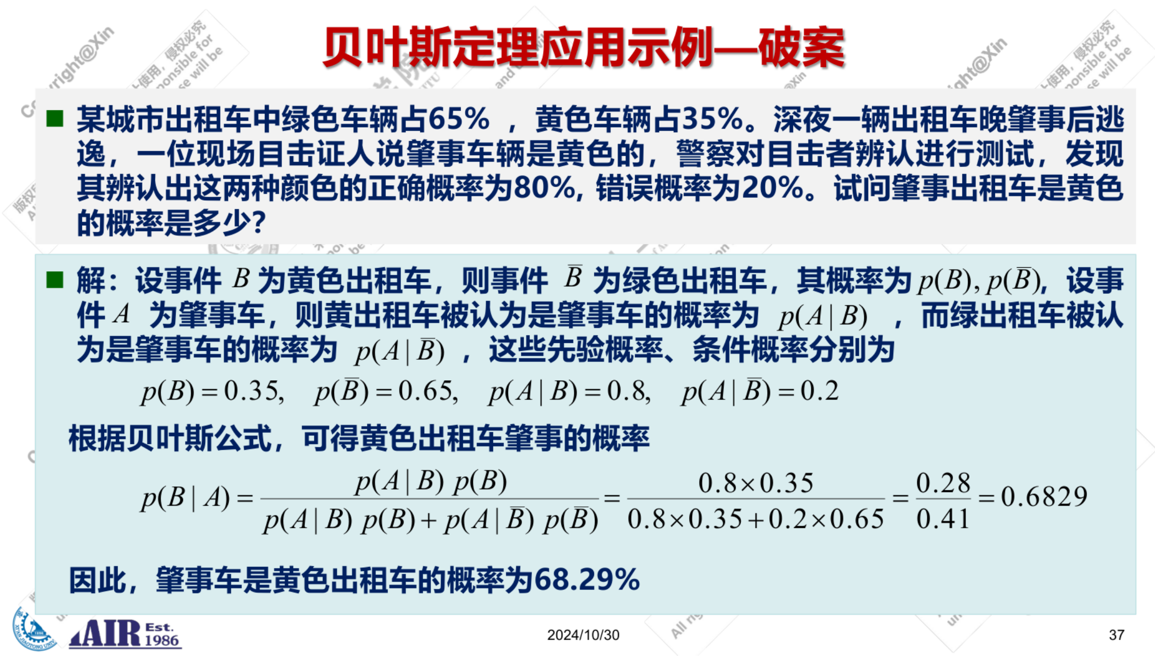
$$P(A) = \sum_i P(A|B_i)P(B_i)$$

2. 贝叶斯公式

根据先验概率和条件概率，可以得到后验概率：

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A)} = \frac{P(A|B_i)P(B_i)}{\sum_j P(A|B_j)P(B_j)}$$

应用：



3. 概率方法不确定性推理

3.1 单证单论

产生式形式：$\text{IF} \quad E \quad \text{THEN} \quad H$

假设H为某一假设，E为某一证据，则根据贝叶斯公式，有

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)}$$

其中，$P(E|H)$是结论出现时证据出现的概率，称为似然概率。

3.2 单证多论

$$P(H_i|E) = \frac{P(E|H_i)P(H_i)}{\sum_k P(E|H_k)P(H_k)}$$

3.3 多证多论

$$P(H_i|E_1, E_2, \dots, E_m) = \frac{P(E_1|H_i)P(E_2|H_i)\dots P(E_m|H_i)P(H_i)}{\sum_k P(E_1|H_k)P(E_2|H_k)\dots P(E_m|H_k)P(H_k)}$$

概率推理的示例1

H_1, H_2, H_3　　　　E

$$p(H_1) = 0.2, \quad p(H_2) = 0.5, \quad p(H_3) = 0.3$$

$$p(E \mid H_1) = 0.1, \quad p(E \mid H_2) = 0.3, \quad p(E \mid H_3) = 0.7$$

求在该证据下3个结论发生的概率

- **解：单证据多结论概率计算公式**

$$p(H_1 \mid E) = \frac{p(E \mid H_1)p(H_1)}{\sum_{k=1}^{3} p(E \mid H_k)p(H_k)} = \frac{0.1 \times 0.2}{0.1 \times 0.2 + 0.3 \times 0.5 + 0.7 \times 0.3} = 0.0526$$

$$p(H_2 \mid E) = \frac{p(E \mid H_2)p(H_2)}{\sum_{k=1}^{3} p(E \mid H_k)p(H_k)} = \frac{0.15}{0.38} = 0.3947$$

$$p(H_3 \mid E) = \frac{p(E \mid H_3)p(H_3)}{\sum_{k=1}^{3} p(E \mid H_k)p(H_k)} = \frac{0.21}{0.38} = 0.5526$$

二、主观贝叶斯方法

1. 基本思想

在主观贝叶斯方法中，通过使用一组充分性度量LS和必要性度量LN来表示知识的不确定性，其具体的产生式形式为：

$$\text{IF} \quad E \quad \text{THEN} \quad (LS, LN) \quad H \quad (p(H))$$

其中，LS为充分性度量，表示E对H的支持程度，取值区间$[0, +\infty)$；LN为必要性度量，表示$\neg E$对H的支持程度，取值区间$[0, +\infty)$。$p(H)$为没有任何证据下H的先验概率，由专家给出。

2. 几率

信任度：$p(H|E)$表示在有证据E的情况下，某种假设H为真的信任度

几率定义为某种事件H为真的概率与其为假的概率之比：

$$O(H) = \frac{p(H)}{p(\neg H)} = \frac{p(H)}{1 - p(H)}$$

反过来有

$$p(H) = \frac{O(H)}{1 + O(H)}$$

几率实质上表示证据的不确定性，将概率从$[0, 1]$映射到$[0, +\infty)$。

3. 充分性因子（LS）

充分性因子定义为

$$LS = \frac{O(H|E)}{O(H)} = \frac{p(E|H)}{p(E|\neg H)}$$

则有

$$O(H|E) = LS \cdot O(H)$$

- **当 $LS > 1$ 时，$O(H|E) > O(H)$ ，说明证据 E 支持结论 H ，LS 越大，支持越充分**
- **当 $LS \to \infty$ 时，$O(H|E) \to \infty$ ，即 $p(H|E) = 1$，说明证据 E 将导致结论 H 为真**
- **当 $LS = 1$ 时，$O(H|E) = O(H)$，说明证据 E 对结论 H 没有影响**
- **当 $LS < 1$ 时，$O(H|E) < O(H)$，说明证据 E 不支持结论 H**
- **当 $LS = 0$ 时，$O(H|E) = 0$ ，说明证据 E 的存在使结论 H 为假**

4. 必要性因子（LN）

必要性因子定义为

$$LN = \frac{O(H|\neg E)}{O(H)} = \frac{p(\neg E|H)}{p(\neg E|\neg H)}$$

则有

$$O(H|\neg E) = LN \cdot O(H)$$

- **当 $LN > 1$ 时，$O(H|\neg E) > O(H)$，说明证据 $\neg E$ 支持结论 H ，LN 越大，支持越充分**
- **当 $LN \to \infty$ 时,$O(H|\neg E) \to \infty$ ，即 $p(H|\neg E) = 1$ ，说明证据 $\neg E$ 的存在将导致结论 H 为真**
- **当 $LN = 1$ 时，$O(H|\neg E) = O(H$ ，说明证据 $\neg E$ 对结论 H 没有影响**
- **当 $LN < 1$ 时，$O(H|\neg E) < O(H)$ ，说明证据 $\neg E$ 不支持结论 H，即证据不存在，导致结论为真的可能性降低**
- **当 $LN = 0$ 时，$O(H|\neg E) = 0$ ，说明证据 $\neg E$ 的存在使结论 H 为假**

5. LS和LN的关系

- 证据存在 E 和证据不存在 $\neg E$ 不会同时支持或同时排斥结论 H，根据数学意义上的取值范围，充分性度量 LS 与必然性度量 LN 间仅存在3种情况：

$$\begin{cases} LS>1 \quad \text{and} \quad LN<1 \\ LS<1 \quad \text{and} \quad LN>1 \\ LS=LN=1 \end{cases}$$

- 当证据 E 对结论 H 支持越强烈，LS 就越大，而 LN 就越小，反之亦然
- 虽然已有LS、LN的计算公式，但实际上领域专家并不一定真按公式计算规则的LS、LN，而往往是凭经验给出
- 领域专家根据经验所提供的LS、LN通常不满足这一理论上的限制，常常在承认E支持H（即LS>1）的同时却否认E反对H（即LN<1）

6. 组合证据可信度计算

已知当前观察S下每个单一证据E_i的概率为$p(E_i|S)$，则可用最大最小法计算组合证据的可信度。

对于合取证据$E=E_1\wedge E_2\wedge...\wedge E_n$，可信度为

$$p(E|S)=\min\{p(E_1|S),p(E_2|S),\ldots,p(E_n|S)\}$$

对于析取证据$E=E_1\vee E_2\vee...\vee E_n$，可信度为

$$p(E|S)=\max\{p(E_1|S),p(E_2|S),\ldots,p(E_n|S)\}$$

对于非证据有

$$p(\neg E|S)=1-p(E|S)$$

7. 主观贝叶斯方法推理

将先验概率$p(H)$更新为后验概率$p(H|E)$时分3种不同情况：证据肯定存在、证据不存在或者证据不确定：

a. 证据确定为真 (p(E)=1)

$$p(H|E)=\frac{LS\cdot p(H)}{(LS-1)\cdot p(H)+1}$$

b. 证据确定为假 (p(E)=0)

$$p(H|\neg E)=\frac{LN\cdot p(H)}{(LN-1)\cdot p(H)+1}$$

c. 证据不确定 (杜达公式)

当证据E非真非假时，结论H依赖于证据E，而E基于当前观察S的可信度为$p(E|S)\in[0,1]$，则$p(H|S)$可按杜达公式计算：

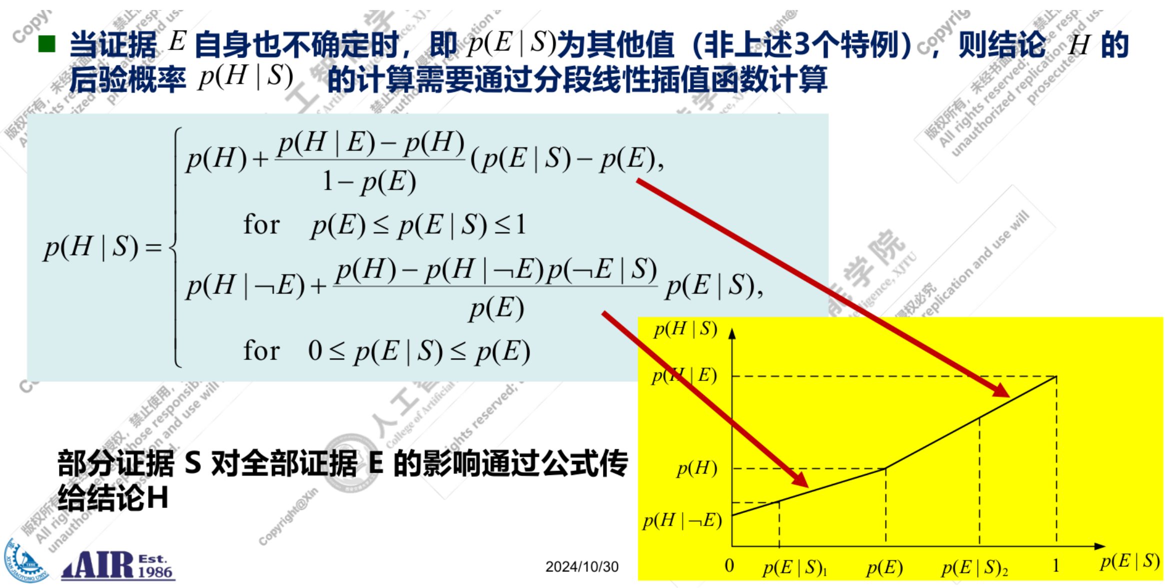
$$p(H|S)=p(H|E)\cdot p(E|S)+p(H|\neg E)\cdot p(\neg E|S)$$

当$p(E|S)$为0或1时，杜达公式分别简化为前两种情况。

当$p(E|S)=p(E)$时，杜达公式简化为全概率公式：

$$p(H|S)=p(H|E)\cdot p(E)+p(H|\neg E)\cdot p(\neg E)$$

当$p(E|S)$不为$0,1$或$p(E)$时，$p(H|S)$需要用分段线性插值计算：



多证据推理的结论合成

如果有多条证据$E_1, E_2, ..., E_m$支持同一结论H，且各证据相互独立，则后验几率计算公式为

$$O(H|E_1,E_2,...,E_n)=\frac{O(H|E_1)}{O(H)}\frac{O(H|E_2)}{O(H)}\cdots\frac{O(H|E_n)}{O(H)}\cdot O(H)$$

- **某天气预报的专家系统中，某地的下雨先验概率为** $p(H)=0.03$
  - **若刮西北风3~4级（设为E1），则会下雨，其充分性因子、必要性因子分别为12，1**
  - **若空气湿度为10%~90%（设为E2），则会下雨，其充分性因子、必要性因子分别为23，1**
  - **若前一天下雨（设为E3），则会下雨，其充分性因子、必要性因子分别为76，1**

当3个证据必然发生时，试求下雨的概率

- **解：利用已知条件可得如下产生式知识表示**

$$R_1:\ \text{IF}\ E_1\ \text{THEN}\ (12,1)\ H\ (p(H))$$
$$R_2:\ \text{IF}\ E_2\ \text{THEN}\ (23,1)\ H\ (p(H))$$
$$R_3:\ \text{IF}\ E_3\ \text{THEN}\ (76,1)\ H\ (p(H))$$

其中 $LS_1=12, LN_1=1, LS_2=23, LN_2=1, LS_3=76, LN_3=1$

- **证据E1（刮西北风）发生时，结论（即下雨）的概率可计算为**

$$p(H \mid E_1) = \frac{LS_1 \cdot p(H)}{(LS_1 - 1) \cdot p(H) + 1} = \frac{12 \times 0.03}{(12-1) \times 0.03 + 1} = 0.2707$$

- **在证据E1的基础上，证据E2（湿度）也发生时，结论（即下雨）的概率可计算为**

$$p(H \mid E_1, E_2) = \frac{LS_2 \cdot p(H \mid E_1)}{(LS_2 - 1) \cdot p(H \mid E_1) + 1} = \frac{23 \times 0.2707}{(23-1) \times 0.2707 + 1} = 0.8951$$

- **在证据E1和E2的基础上，证据E3（前一天下雨）也发生时，结论（即下雨）的概率可计算为**

$$p(H \mid E_1, E_2, E_3) = \frac{LS_3 \cdot p(H \mid E_1, E_2)}{(LS_3 - 1) \cdot p(H \mid E_1, E_2) + 1} = \frac{76 \times 0.8951}{(76-1) \times 0.8951 + 1} = 0.9985$$

三、可信度方法

1. 可信度因子模型

格式：IF E THEN H $(CF(H,E))$

- **可信度因子** $\mathrm{CF}(H,E)$ **的定义为信任与不信任的差，即**

$$\mathrm{CF}(H,E) = MB(H,E) - MD(H,E), \quad \mathrm{CF}(H,E) \in [-1,1]$$

- **MB (measure belief) 信任度量：表示证据E对前提结论H的信任增长度**
- **MD (measure disbelief) 不信任度量：表示证据E对前提结论H的不信任增长度**

$$MB(H,E) = \begin{cases} 1, & p(H) = 1 \\ \dfrac{\max\{p(H \mid E), p(H)\} - p(H)}{1 - p(H)}, & p(H) \neq 1 \end{cases}$$

$if\ MB(H,E) > 0,\ p(H \mid E) > p(H)$

证据E的出现会增加结论H的信任程度

$$MD(H,E) = \begin{cases} 1, & p(H) = 0 \\ \dfrac{\min\{p(H \mid E), p(H)\} - p(H)}{-p(H)}, & p(H) \neq 0 \end{cases}$$

$if\ MD(H,E) > 0,\ p(H \mid E) < p(H)$

证据E的出现会增加结论H的不信任程度

MB和MD具有互斥性，即二者至少有一个为0，则CF取值区间为$[-1,1]$。
实际应用中，CF通常由专家给出。

可信度和概率的区别：

$$CF(H,E) + CF(\neg H,E) = 0$$
$$p(H|E) + p(\neg H|E) = 1$$

若 $\mathrm{CF}(H,E) > 0$，$\mathrm{CF}(H,E)$ 的值越大，则证据E支持结论H越真
若 $\mathrm{CF}(H,E) < 0$，$\mathrm{CF}(H,E)$ 的值越小，则证据E支持结论H越假
若 $\mathrm{CF}(H,E) = 0$，则证据E的出现与否与H无关

若对于同一个证据E存在多个互不相容的结论$H_1, H_2, \ldots, H_n$，则有

$$\sum_{i=1}^{n} CF(H_i, E) \leq 1$$

2. 证据的不确定性表示

证据的不确定性用可信度因子表示为$CF(E)$.

对于多个证据$E_1, E_2, ..., E_n$，其合取证据$E = E_1 \wedge E_2 \wedge ... \wedge E_n$的可信度因子为

$$CF(E) = \min\{CF(E_1), CF(E_2), ..., CF(E_n)\}$$

其析取证据$E = E_1 \vee E_2 \vee ... \vee E_n$的可信度因子为

$$CF(E) = \max\{CF(E_1), CF(E_2), ..., CF(E_n)\}$$

3. 可信度推理

结论H的可信度$CF(H)$可按下列公式计算：

$$CF(H) = CF(H, E) \cdot \max\{0, CF(E)\}$$

当证据某种程度为真时，结论的可信度为知识的可信度；当证据某种程度为假时，知识不能使用；即在推理中未考虑假证据对结论的影响。

P.S. 如果一个证据E_i题目中未提到观察到时，默认$CF(E_i) = 0$.

两条证据支持同一结论时的综合可信度：

$$\mathrm{CF}_{12}(H) = \begin{cases} \mathrm{CF}_1(H) + \mathrm{CF}_2(H) - \mathrm{CF}_1(H) \cdot \mathrm{CF}_2(H), & \mathrm{CF}_1(H) \geq 0, \mathrm{CF}_2(H) \geq 0 \\ \mathrm{CF}_1(H) + \mathrm{CF}_2(H) + \mathrm{CF}_1(H) \cdot \mathrm{CF}_2(H), & \mathrm{CF}_1(H) < 0, \mathrm{CF}_2(H) < 0 \\ \dfrac{\mathrm{CF}_1(H) + \mathrm{CF}_2(H)}{1 - \min\{|\mathrm{CF}_1(H)|, |\mathrm{CF}_2(H)|\}}, & \mathrm{CF}_1(H) \times \mathrm{CF}_2(H) < 0 \end{cases}$$

对于多条证据，可按递归计算综合可信度。

- **设有如下一组不确定性推理的知识规则**

R1: IF E_1 AND E_2 THEN E_3 (1.0)

R2: IF E_3 OR E_4 THEN E_5 (0.8)

R3: IF E_5 THEN H (0.8)

R4: IF E_6 THEN H (0.9)

已知其中4个证据的可信度为 $\mathrm{CF}(E_1) = 0.7, \mathrm{CF}(E_2) = 0.5, \mathrm{CF}(E_4) = 0.4, \mathrm{CF}(E_6) = 0.8$

求结论的综合可信度 $\mathrm{CF}(H)$

- **解：1）求4条知识规则的可信度**

- **R1:**
$$\begin{aligned}\mathrm{CF}(E_3) &= \mathrm{CF}(R_1) \times \max\{0, \mathrm{CF}(E_1 \text{ AND } E_2)\} \\ &= \mathrm{CF}(R_1) \times \max\{0, \min\{\mathrm{CF}(E_1), \mathrm{CF}(E_2)\}\} \\ &= 1.0 \times \max\{0, \min\{0.7, 0.5\}\} \\ &= 1.0 \times \max\{0, 0.5\} \\ &= 1.0 \times 0.5 = 0.5\end{aligned}$$

IF E THEN H $(\mathrm{CF}(H, E))$

$\mathrm{CF}(H) = \mathrm{CF}(H, E) \times \max\{0, \mathrm{CF}(E)\}$

“与”组合证据

$\mathrm{CF}(E) = \min\{\mathrm{CF}(E_1), \mathrm{CF}(E_2)\}$

“或”组合证据

$\mathrm{CF}(E) = \max\{\mathrm{CF}(E_1), \mathrm{CF}(E_2)\}$

- **R2:**
$$\begin{aligned}\mathrm{CF}(E_5) &= \mathrm{CF}(R_2) \times \max\{0, \mathrm{CF}(E_3 \text{ OR } E_4)\} \\ &= \mathrm{CF}(R_2) \times \max\{0, \max\{\mathrm{CF}(E_3), \mathrm{CF}(E_4)\}\} \\ &= 0.8 \times \max\{0, \max\{0.5, 0.4\}\} \\ &= 0.8 \times \max\{0, 0.5\} \\ &= 0.8 \times 0.5 = 0.4\end{aligned}$$

- **R3：** $\mathrm{CF}_1(H)=\mathrm{CF}(H,E_5)\times\max\{0,\mathrm{CF}(E_5)\}$
  $=0.8\times\max\{0,0.4\}$
  $=0.8\times 0.4=0.32$

$\text{IF}\quad E\quad \text{THEN}\quad H\quad (\mathrm{CF}(H,E))$
$\mathrm{CF}(H)=\mathrm{CF}(H,E)\cdot\max\{0,\mathrm{CF}(E)\}$

- **R4：** $\mathrm{CF}_2(H)=\mathrm{CF}(H,E_6)\times\max\{0,\mathrm{CF}(E_6)\}$
  $=0.9\times\max\{0,0.8\}$
  $=0.9\times 0.8=0.72$

2）根据结论不确定性的合成公式，综合可信度计算为

$$\mathrm{CF}(H)=\mathrm{CF}_{12}(H)=\mathrm{CF}_1(H)+\mathrm{CF}_2(H)-\mathrm{CF}_1(H)\times\mathrm{CF}_2(H)\qquad \mathrm{CF}_1(H)\ge 0,\ \mathrm{CF}_2(H)\ge 0$$
$$=0.32+0.72-0.32\times 0.72$$
$$=0.8096$$

所以，结论H的综合可信度为 $\mathrm{CF}(H)=0.8096$

4. 带阈值的可信度推理

带阈值限度的知识不确定性

- 在CF模型中引入阈值限度来限制不合乎要求的证据及相应的知识的运用
- 带阈值限度的知识不确定性表示：

前提假设（hypothesis）（结论）

$$\text{IF}\quad E\quad \text{THEN}\quad H\quad (\mathrm{CF}(H,E),l)$$

证据（evidence）　可信度　阈值

- 知识可信度 $\mathrm{CF}(H,E)$ 的取值范围：
  $0\le \mathrm{CF}(H,E)\le 1$
- $\mathrm{CF}(H,E)=0\rightarrow p(H\mid E)=0$ 证据绝对否定结论
- $\mathrm{CF}(H,E)=1\rightarrow p(H\mid E)=1$ 证据绝对支持结论

- 阈值 l 的取值范围：
  $0<l\le 1$
- 证据的可信度满足阈值规定知识适用的条件：
  $\mathrm{CF}(E)\ge l$
  ➡ 知识规则才能够被使用

2024/10/30　96

带阈值限度的结论不确定性合成算法

- 设多条不同阈值限度的知识规则都有相同的结论，即

$$\text{IF } E_1 \text{ THEN } H\ (\mathrm{CF}(H,E_1),l_1),\quad \cdots,\quad \text{IF } E_n \text{ THEN } H\ (\mathrm{CF}(H,E_n),l_n)$$

- 若每条知识规则都满足，即 $\mathrm{CF}(E_i)\ge l_i$ ，则利用各自的结论可信度 $\mathrm{CF}_i(H)$ ，可计算结论的综合可信度 $\mathrm{CF}(H)$ ：

- **1）极大值法：** $\mathrm{CF}(H)=\max\{\mathrm{CF}_1(H),\mathrm{CF}_2(H),\cdots,\mathrm{CF}_n(H)\}$
- **2）加权求和法：** $\mathrm{CF}(H)=\dfrac{1}{\sum_{i=1}^{n}\mathrm{CF}(H,E_i)}\sum_{i=1}^{n}\mathrm{CF}(H,E_i)\cdot\mathrm{CF}(E_i)$
- **3）有限和法：** $\mathrm{CF}(H)=\min\left\{\sum_{i=1}^{n}\mathrm{CF}_i(H),1\right\}$
- **4）递推法：** $C_1=\mathrm{CF}(H,E_1)\cdot\mathrm{CF}(E_1)$
  $C_k=C_{k-1}+(1-C_{k-1})\mathrm{CF}(H,E_k)\cdot\mathrm{CF}(E_k)$

5. 加权的可信度推理

加权的不确定性推理

- 考虑各个子条件的重要性和独立性，每个子条件可分配不同的权重
- 知识不确定性的表示：

$$\text{IF } E_1(w_1) \wedge E_2(w_2) \wedge \cdots \wedge E_n(w_n) \text{ THEN } H\ (\mathrm{CF}(H,E), l\ )$$

  - 加权因子 w_i 的值由专家确定，其满足条件： $0 \le w_i \le 1,\quad \sum_{i=1}^{n} w_i = 1$

- 组合证据的可信度计算：

$$\mathrm{CF}(E) = \frac{1}{\sum_{i=1}^{n} w_i} \sum_{i=1}^{n} w_i \cdot \mathrm{CF}(E_i)$$

- 不确定性的传递算法（结论的可信度计算）：

$$\mathrm{CF}(H) = \mathrm{CF}(H,E) \cdot \mathrm{CF}(E)$$

2024/10/30 99

注意，存在阈值时，算出$CF(E)$后还要判断是否有$CF(E) \geq l$.

四、证据理论

1. 基本概念

识别框架W为变量x的所有可能取值组成的非空集合，其中的元素互不相容：

$$W = \{w_1, w_2, \ldots, w_n\}$$

$m(A)$为A的基本概率赋值，表示证据对命题A的信任程度。

幂集2^W为识别框架W的所有子集组成的集合，包含2^n个元素：

$$2^W = \{A | A \subseteq W\}$$

基本概率分配函数为识别框架的幂集2^W中的每一个子集A分配一个值$m(A)$，满足：

$$\sum_{A \subseteq W} m(A) = 1, \quad m(\emptyset) = 0$$

2. 信任函数和似然函数

对于对应识别框架W子集的命题A，其信任函数$Bel(A)$定义为：

$$Bel(A) = \sum_{B \subseteq A} m(B)$$

其似然函数$Pl(A)$定义为：

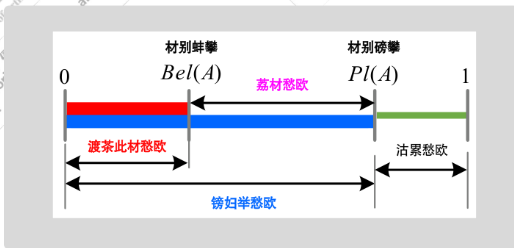
$$Pl(A) = 1 - Bel(\neg A) = \sum_{B \cap A \neq \emptyset} m(B)$$

信任函数和似然函数的关系：

$$0 \leq Bel(A) \leq Pl(A) \leq 1$$

$Bel(A)$表示置信下限，$Pl(A)$表示置信上限，$Pl(A)-Bel(A)$表示不确定性，$[Bel(A),Pl(A)]$表示置信区间。

- **在证据理论中，信任函数 $Bel(A)$ 和似然函数 $Pl(A)$ 综合表示证据的不确定性**
- **信任函数 $Bel(A)$ 和似然函数 $Pl(A)$ 分别表示对命题 A 信任程度的置信区间的下限和上限，即两元组 $[Bel(A),Pl(A)]$ 表示信任区间**
- **命题 A 的完全可信区间：$[0,Bel(A)]$**
- **命题“A 为真”不可怀疑区间：$[0,Pl(A)]$**



$[Bel(A),Pl(A)]=[1,1]$ ：**表示命题A为真**

$[Bel(A),Pl(A)]=[0,0]$ ：**表示命题A为假**

$[Bel(A),Pl(A)]=[0,1]$ ➡ $Bel(\neg A)=1-Pl(A)=0$ ：**对A完全无知，对非A也不信任**

$[Bel(A),Pl(A)]=[0.5,0.5]$ ：**表示无法完全确定命题A是否为真**

3. 德普斯特证据组合

设识别框架W有P组概率分配函数$m_1,m_2,\ldots,m_P$，则其组合的概率分配函数m为：

$$m(A)=(m_1\oplus\ldots\oplus m_P)(A)=\frac{1}{K}\sum_{\cap A_i=A}\prod_{i,p}m_p(A_i)$$

其中，$K=1-\sum_{\cap A_i=\emptyset}\prod_{i,p}m_p(A_i)$ 为归一化因子。

对于两组概率分配函数m_1和m_2，其组合公式为：

$$m(A)=(m_1\oplus m_2)(A)=\frac{1}{K}\sum_{B\cap C=A}m_1(B)m_2(C)$$

其中，$K=1-\sum_{B\cap C=\emptyset}m_1(B)m_2(C)$.

$K=0$时表示所给证据矛盾。

$$W=\{b,w\}$$

$$m_1(\{b\},\{w\},\{b,w\},\varnothing)=(0.3,0.5,0.2,0) \qquad b:black$$

$$m_2(\{b\},\{w\},\{b,w\},\varnothing)=(0.6,0.3,0.1,0) \qquad w:white$$

- **求2组概率分配函数 m_1,m_2 的正交和**

解：根据正交和公式

$$K=1-\sum_{\cap A_i=\varnothing}\prod_{\substack{1\le i\le n\\1\le p\le P}}m_p(A_i)=1-\sum_{x\cap y=\varnothing}m_1(x)m_2(y)$$

$$=1-(m_1(\{b\})m_2(\{w\})+m_1(\{w\})m_2(\{b\}))=1-(0.3\times0.3+0.5\times0.6)=0.61$$

$$m(\{b\})=\frac{1}{K}\sum_{x\cap y=\{b\}}m_1(x)m_2(y)$$

$$=\left(m_1(\{b\})m_2(\{b\})+m_1(\{b\})m_2(\{b,w\})+m_1(\{b,w\})m_2(\{b\})\right)/0.61$$

$$=(0.3\times0.6+0.3\times0.1+0.6\times0.2)/0.61=0.33/0.61=0.5410$$

$$m(\{w\})=\frac{1}{K}\sum_{x\cap y=\{w\}} m_1(x)m_2(y)$$
$$=\left(m_1(\{w\})m_2(\{w\})+m_1(\{w\})m_2(\{b,w\})+m_1(\{b,w\})m_2(\{w\})\right)/0.61$$
$$=(0.5\times0.3+0.5\times0.1+0.3\times0.2)/0.61=0.26/0.61=0.4262$$

$$m(\{b,w\})=\frac{1}{K}\sum_{x\cap y=\{b,w\}} m_1(x)m_2(y)=\frac{1}{0.61}\left(m_1(\{b,w\})m_2(\{b,w\})\right)$$
$$=\frac{1}{0.61}(0.2\times0.1)=\frac{0.02}{0.61}=0.0328$$

所以，经过证据组合后得到的概率分配函数为：

$$m(\{b\},\{w\},\{b,w\},\varnothing)=(0.5410,0.4262,0.0328,0)$$

4. 证据理论推理

基本的证据理论推理过程包括以下步骤：

1. 确定识别框架W，写出幂集对应每个证据的基本概率赋值$m_p(A_i)$；
2. 使用德普斯特证据组合计算综合的概率分配函数$m(A)$；
3. 根据要求计算信任函数$Bel(A)$和似然函数$Pl(A)$.

一般来说默认$m(\emptyset)=0$.
如果结论有且仅有一个，则$|A|\neq 1$时$m(A)=0$.

- **一起谋杀案有3个犯罪嫌疑人分别为Peter、Paul和Mary，2个目击证人为W1和W2，目击证人分别指证犯罪嫌疑人，得到两个mass函数 m1 和 m2如下**

	Peter	Paul	Mary
证人W1：$m_1(\cdot)$	0.86	0.13	0.01
证人W2：$m_2(\cdot)$	0.02	0.90	0.08

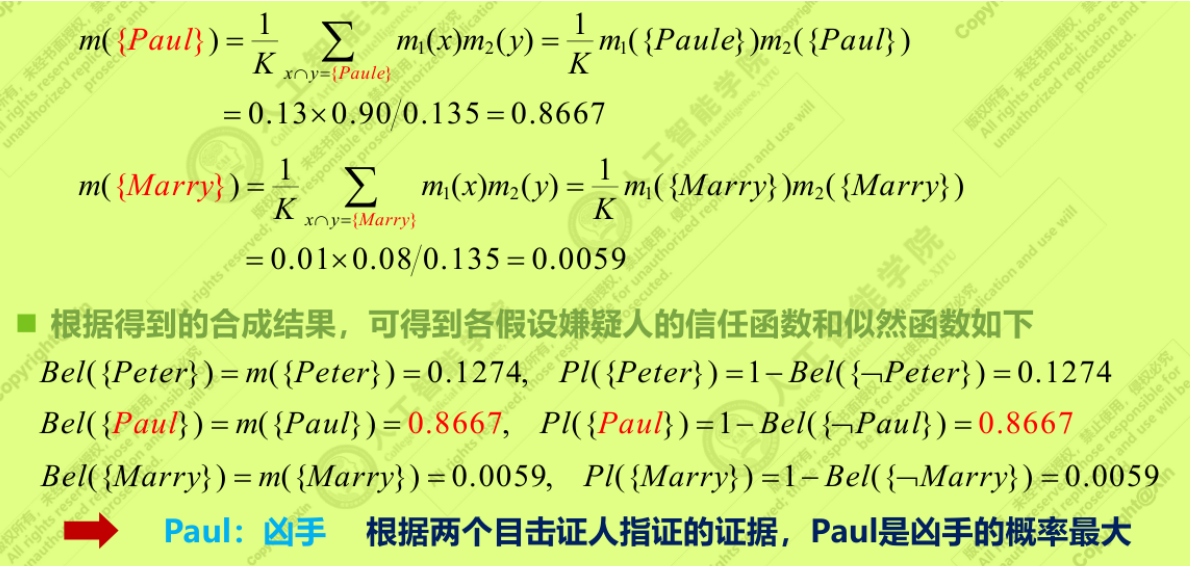
在该案件中罪犯只能有1个，根据3个目击者的证据确定3个嫌疑人中的哪一个是犯人？

- **解：该问题的识别框架为：** $W=\{Peter,Paul,Marry\}$ **，其基本概率分配函数分别为** $m(\{Peter\}),m(\{Paul\}),m(\{Marry\})$

根据证据规则，将上述2个不同规则得到的概率分配函数组合，可得到

$$K=1-\sum_{\cap A_i=\varnothing}\prod_{\substack{1\le i\le n\\1\le p\le P}} m_p(A_i)=1-\sum_{x\cap y=\varnothing} m_1(x)m_2(y)$$
$$=1-(m_1(\{Peter\})m_2(\{Paul\})+m_1(\{Peter\})m_2(\{Marry\})+m_1(\{Paul\})m_2(\{Marry\})$$
$$+m_2(\{Peter\})m_1(\{Paul\})+m_2(\{Peter\})m_1(\{Marry\})+m_2(\{Paul\})m_1(\{Marry\}))$$
$$=1-(0.86\times0.90+0.86\times0.08+0.13\times0.08+0.02\times0.13+0.02\times0.01+0.90\times0.01)$$
$$=1-0.865=0.135$$

$$m(\{Peter\})=\frac{1}{K}\sum_{x\cap y=\{Peter\}} m_1(x)m_2(y)=\frac{1}{K}\left(m_1(\{Peter\})m_2(\{Peter\})\right)$$
$$=0.86\times0.02/0.135=0.1274$$



5. Zadeh悖论

设识别框架$W=\{H_1,H_2,H_3\}$，有两组证据：

证据	H_1	H_2	H_3
$m_1(\cdot)$	0.99	0.00	0.01
$m_2(\cdot)$	0.00	0.99	0.01

使用德普斯特证据组合计算综合的概率分配函数m：

$$m(H_1)=0,\quad m(H_2)=0,\quad m(H_3)=1$$

Zadeh悖论说明德普斯特证据组合有时无法解决证据严重或完全冲突的情况，且概率分配函数的微小变化可能会导致组合证据的概率分配函数产生较大变化。

近似推理（模糊逻辑）

一、模糊逻辑

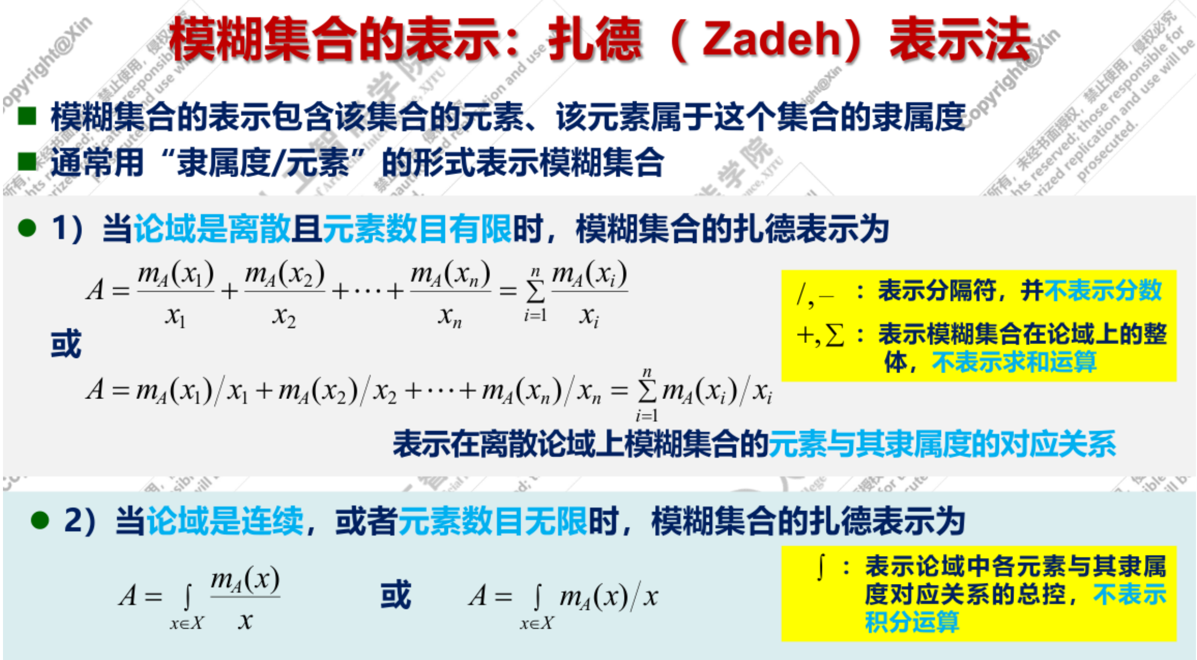
1. 模糊集合和隶属度

模糊集A是定义在某一论域X上的集合，其每个元素$x\in X$都具有一个隶属度$\mu_A(x)$，表示x属于模糊集A的程度，取值范围为$[0,1]$：

$$A=\{(x,\mu_A(x),x\in X)\}$$

模糊集A完全由其隶属函数$\mu_A(x)$唯一确定。当隶属度$\mu_A(x)$仅取0或1时，模糊集退化为普通集合。

模糊集可用Zadeh表示法表示：



2. 模糊集合的基本运算

模糊集合的运算实际上是逐点对隶属函数进行相应的运算。

设A和B为定义在同一论域X上的模糊集，则有：

- 相等：$A = B \iff \mu_A(x) = \mu_B(x), \forall x \in X$
- 包含：$A \subseteq B \iff \mu_A(x) \leq \mu_B(x), \forall x \in X$
- 并集：$\mu_{A\cup B}(x) = \max\{\mu_A(x), \mu_B(x)\}$
- 交集：$\mu_{A\cap B}(x) = \min\{\mu_A(x), \mu_B(x)\}$
- 补集：$\mu_{\neg A}(x) = 1 - \mu_A(x)$

3. 模糊集合的代数运算

设A和B为定义在同一论域X上的模糊集，则有：

- 代数和：$\mu_{A+B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x) \cdot \mu_B(x)$
- 代数积：$\mu_{A\cdot B}(x) = \mu_A(x) \cdot \mu_B(x)$
- 有界和：$\mu_{A\oplus B}(x) = \min\{1, \mu_A(x) + \mu_B(x)\}$
- 有界积：$\mu_{A\odot B}(x) = \max\{0, \mu_A(x) + \mu_B(x) - 1\}$

P.S. 有时也用$\vee$表示取最大值运算，$\wedge$表示取最小值运算。

二、笛卡尔积与模糊关系

1. 笛卡尔积

对于两个非空集合X和Y，其笛卡尔积$X \times Y$定义为：

$$X \times Y = \{(x, y) | x \in X, y \in Y\}$$

2. 模糊关系

设两个离散模糊集合A, B的隶属函数分别为：

$$\mathbf{m}_A = [\mu_A(x_1), \mu_A(x_2), \dots, \mu_A(x_m)]$$
$$\mathbf{m}_B = [\mu_B(y_1), \mu_B(y_2), \dots, \mu_B(y_n)]$$

则其模糊关系$R = A \times B$的隶属函数为：

$$\mu_{A\times B} = \mathbf{m}_A^T \circ \mathbf{m}_B$$

元素的外积运算为取最小值运算：$x \circ y = \min\{x, y\} = x \wedge y$

■ **已知输入模糊集合 A 和输出模糊集合 B 分别为**

$$A = 1/a_1 + 0.8/a_2 + 0.5/a_3 + 0.2/a_4 + 0.0/a_4, \quad B = 0.7/b_1 + 1.0/b_2 + 0.6/b_3 + 0.0/b_4$$

求模糊集合 A 到 B 的模糊关系

■ **解：**

$$R = A \times B = \boldsymbol{m}_A^T \circ \boldsymbol{m}_B = \begin{bmatrix} 1.0 \\ 0.8 \\ 0.5 \\ 0.2 \\ 0.0 \end{bmatrix} \circ [0.7 \quad 1.0 \quad 0.6 \quad 0.0]$$

$$= \begin{bmatrix} 1.0\wedge 0.7 & 1.0\wedge 1.0 & 1.0\wedge 0.6 & 1.0\wedge 0.0 \\ 0.8\wedge 0.7 & 0.8\wedge 1.0 & 0.8\wedge 0.6 & 0.8\wedge 0.0 \\ 0.5\wedge 0.7 & 0.5\wedge 1.0 & 0.5\wedge 0.6 & 0.5\wedge 0.0 \\ 0.2\wedge 0.7 & 0.2\wedge 1.0 & 0.2\wedge 0.6 & 0.2\wedge 0.0 \\ 0.0\wedge 0.7 & 0.0\wedge 1.0 & 0.0\wedge 0.6 & 0.0\wedge 0.0 \end{bmatrix} = \begin{bmatrix} 0.7 & 1.0 & 0.6 & 0.0 \\ 0.7 & 0.8 & 0.6 & 0.0 \\ 0.5 & 0.5 & 0.5 & 0.0 \\ 0.2 & 0.2 & 0.2 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix}$$

3. 模糊关系的合成

模糊关系的合成可以类比矩阵乘法，将加法替换为取最大值运算，乘法替换为取最小值运算：

$$X = \{x_1, x_2, x_3, x_4\},\ Y = \{y_1, y_2, y_3\},\ Z = \{z_1, z_2\}$$

$$R_1 = X \times Y = \begin{bmatrix} 0.8 & 0.2 & 0.5 \\ 0.2 & 0.4 & 0.9 \\ 1.0 & 0.0 & 0.7 \end{bmatrix} \qquad R_2 = Y \times Z = \begin{bmatrix} 0.1 & 0.9 \\ 0.7 & 0.8 \\ 0.0 & 1.0 \end{bmatrix}$$

求模糊关系 R_1, R_2 **的合成** S

■ **解：根据模糊关系的合成运算规则，可得**

先取最小再取最大

$$S = R_1 \circ R_2 = \begin{bmatrix} 0.8 & 0.2 & 0.5 \\ 0.2 & 0.4 & 0.9 \\ 1.0 & 0.0 & 0.7 \end{bmatrix} \circ \begin{bmatrix} 0.1 & 0.9 \\ 0.7 & 0.8 \\ 0.0 & 1.0 \end{bmatrix}$$

$$= \begin{bmatrix} (0.8\wedge 0.1)\vee(0.2\wedge 0.7)\vee(0.5\wedge 0.0) & (0.8\wedge 0.9)\vee(0.2\wedge 0.8)\vee(0.5\wedge 1.0) \\ (0.2\wedge 0.1)\vee(0.4\wedge 0.7)\vee(0.9\wedge 0.0) & (0.2\wedge 0.9)\vee(0.4\wedge 0.8)\vee(0.9\wedge 1.0) \\ (0.1\wedge 0.1)\vee(0.0\wedge 0.7)\vee(0.9\wedge 0.0) & (1.0\wedge 0.9)\vee(0.0\wedge 0.8)\vee(0.7\wedge 1.0) \end{bmatrix}$$

$$= \begin{bmatrix} 0.1\vee 0.2\vee 0.0 & 0.8\vee 0.2\vee 0.5 \\ 0.1\vee 0.4\vee 0.0 & 0.2\vee 0.4\vee 0.9 \\ 0.1\vee 0.0\vee 0.0 & 0.9\vee 0.0\vee 0.7 \end{bmatrix} = \begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.9 \\ 0.1 & 0.9 \end{bmatrix}$$

三、模糊推理和决策

1. 模糊推理

已知输入为A，输出为B；给出输入A'，输出B'可用模糊关系合成计算：

$$B' = A' \circ R, \quad \text{where} \;\; R = A \times B$$

2. 模糊决策

得到的输出B'是一个模糊向量，需要将其转化为确定值才能应用。常用的决策方法如下。

a. 最大隶属度法

选择隶属度最大的元素对应的值作为推理结果；若有多个最大值，取其平均值。

b. 重心法（加权平均决策法）

计算隶属函数的重心作为推理结果：

$$b' = \frac{\sum\limits_i b_i \cdot \mu(b_i)}{\sum\limits_i \mu(b_i)}$$

c. 中位数法

将模糊集的中位数，即论域上将隶属函数曲线与横坐标围成的面积平分的元素值作为推理结果。

2024年真题

1.（概率推理）小明母亲去家长会的概率是0.75，母亲不去条件下父亲去的概率是0.8，母亲去条件下父亲去的概率是0.25，求小明父亲去条件下母亲去的概率。

设A表示母亲去家长会，B表示父亲去家长会，则有：
$P(A) = 0.75$，$P(B|\neg A) = 0.8$，$P(B|A) = 0.25$
应用贝叶斯公式，有

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\neg A)P(\neg A)} = \frac{0.25 \times 0.75}{0.25 \times 0.75 + 0.8 \times 0.25} \approx 0.4839$$

2.（多证据主观贝叶斯推理）$p(H) = 0.1$，LS分别为$15, 20, 24$，LN均为1，E_1, E_2, E_3均为必然发生，求$p(H|E_1, E_2, E_3)$。

由于E_1, E_2, E_3均为必然发生，则有

$$p(H|E_1) = \frac{LS_1 \cdot p(H)}{(LS_1 - 1) \cdot p(H) + 1} = 0.625$$

$$p(H|E_1, E_2) = \frac{LS_2 \cdot p(H|E_1)}{(LS_2 - 1) \cdot p(H|E_1) + 1} \approx 0.9709$$

$$p(H|E_1, E_2, E_3) = \frac{LS_3 \cdot p(E_1, E_2)}{(LS_3 - 1) \cdot p(E_1, E_2) + 1} \approx 0.9988$$

也可以先分别算出$O(H|E_i)$，再使用公式：

$$O(H|E_1, E_2, E_3) = \frac{O(H|E_1)}{O(H)} \frac{O(H|E_2)}{O(H)} \frac{O(H|E_3)}{O(H)} \cdot O(H)$$

最后将$O(H|E_1, E_2, E_3)$化为$p(H|E_1, E_2, E_3)$，结果相同。

3.（概率推理）黄绿出租车问题

参考PPT原题

4.（证据理论推理）一个星星的分类存在争议，经查阅A,B两篇资料，总结出以下mass函数。求三个种类的信任函数与似然函数，并说明该星最可能是哪种星。

	行星	矮行星	小行星
$m_1(\cdot)$	0.6	0.3	0.1
$m_2(\cdot)$	0.2	0.7	0.1

设识别框架W={行星, 矮行星, 小行星}，已知概率分配函数m_1和m_2，则根据德普斯特证据组合，有

$$K = 1 - \sum_{B \cap C = \emptyset} m_1(B) m_2(C) = 0.33$$

代入公式计算组合概率分配函数$m(A)$:

$$m(A) = (m_1 \oplus m_2)(A) = \frac{1}{K} \sum_{B \cap C = A} m_1(B) m_2(C)$$

解得

$$m(\text{行星}) \approx 0.353, \quad m(\text{矮行星}) \approx 0.618, \quad m(\text{小行星}) \approx 0.029$$

由于证据A, B只将mass分配给单一元素，因此信任函数和似然函数都等于概率分配函数:

$$\begin{aligned} Bel(\text{行星}) &= Pl(\text{行星}) \approx 0.353 \\ Bel(\text{矮行星}) &= Pl(\text{矮行星}) \approx 0.618 \\ Bel(\text{小行星}) &= Pl(\text{小行星}) \approx 0.029 \end{aligned}$$

矮行星的信任度和似然度最高，因此该星最可能是矮行星。

智能体

1. **智能体**：在环境中通过传感器感知环境，并通过效应器（或执行器）自主作用于环境
   智能：智能体在执行任务过程中学习的能力
2. **多智能体系统**：多个相互作用但各自自主智能体构成，各智能体相互通信、合作、竞争，各组成智能体可以完全异质，无全局数据或者控制
3. 智能体三组件：传感器、执行器、效应器
   传感器：感知环境，检测环境变化
   执行器：驱动效应器完成动作（e.g. 肌肉关节）
   效应器：采取行动，直接作用于环境（e.g. 手指、腿）
4. **智能体任务环境表示 PEAS**（一种对智能体属性进行分组的表示模型）
   P (Performance Measure - 性能度量)：智能体行为成功的客观指标
   E (Environment - 环境)：智能体周围的工作环境和条件
   A (Actuators - 效应器)：智能体通过其产生结果，传递动作输出给环境
   S (Sensors - 感知器)：智能体通过其接受输入，观察感知环境
   ——举例（自动驾驶汽车 PEAS）：
   P：速度、安全性、指定任务用时、乘客舒适度等
   E：道路、其他车辆、行人、交通信号灯、路标等
   A：油门、刹车、喇叭、信号灯等
   S：摄像头、GPS、雷达、里程表、车速表等
5. **理性智能体**：通过每个可能的感知序列提供的证据和智能体内置知识，能做到选择一个可以预期最大化改进其性能度量的动作
6. **全知智能体**：明确直到其行为的实际结果，做出动作完全正确（现实不可能）
   理性≠全知：理性是预期性能最大化，全知是实际结果正确。智能体可能因感知信息/先验知识有限或有误而无法做到全知，但这不影响其理性
   理性≠完美：理性是期望性能最大化；完美是实际性能最大化
   理性=自主、学习、探索：理性智能体应该是自主的，应该学习以弥补不完整或不正确的先验知识
   举例：理性智能体-国际象棋（组合爆炸，无法遇见所有可能结果）V.S. 全知智能体-井字棋
7. **智能体环境类型**
   1) 完全可观察&部分可观察
   完全可观察：智能体在每个时间点可感知/访问环境的完整状态（无需存储历史），否则称为部分可观察
   举例：国际象棋（完全）；自动驾驶（部分，无法知道其他司机意图）
   2) 确定性&随机性
   确定性：当前状态和智能体动作唯一完全确定下一状态
   举例：国际象棋（确定性）；自动驾驶（随机性，依赖其他车辆行为）

3) **竞争&合作**

4) **单智能体&多智能体**

5) **静态&动态**

静态：智能体执行操作时，环境不会改变

举例：交通路况（动态）；解数独（静态）

6) **离散&连续**

连续：可执行操作是连续的

举例：国际象棋（离散）；自动驾驶（连续--速度、角度等）

7) **片段&延续**

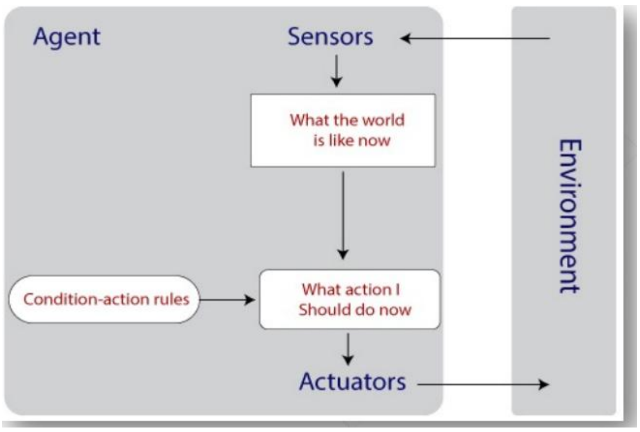
片段：智能体经验被划分为彼此独立子片段，后续片段不依赖前面片段动作

举例：图像分类（片段）；国际象棋、自动驾驶（延续）

8. **智能体结构**：
   1) **架构**：物理传感器、执行器等构成部件的物理布局
   2) **函数**：从感知序列到动作的映射
   3) **程序**：在物理架构上实现/执行智能体函数
9. 智能体各组件关系：
   1) 架构使来自传感器的感知可被程序使用
   2) 架构运行程序，将程序动作选择提供给执行器
10. **智能体函数和程序的输入**：

    函数：全部感知历史（数学理想）

    程序：当前感知（实际工程中通过维持一个内部状态近似替代全部感知历史）
11. 智能体形式：包括人类智能体、机器人智能体和软件智能体。
12. **智能体类型**：
    1) **简单反射智能体**



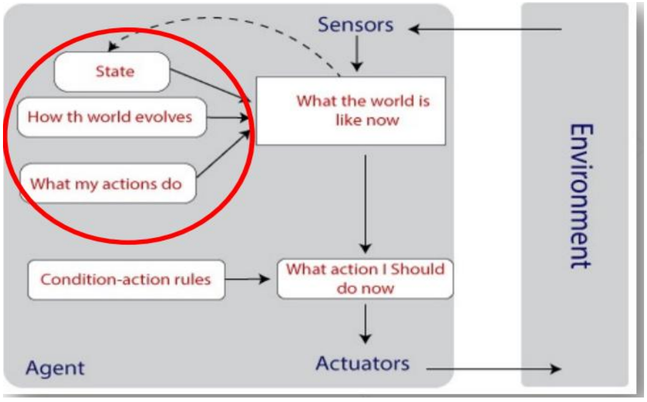
①. 只根据当前感知采取行动，完全不考虑感知历史其他部分

②. 基于条件-动作规则工作

③. 在完全可观察环境中必然成功

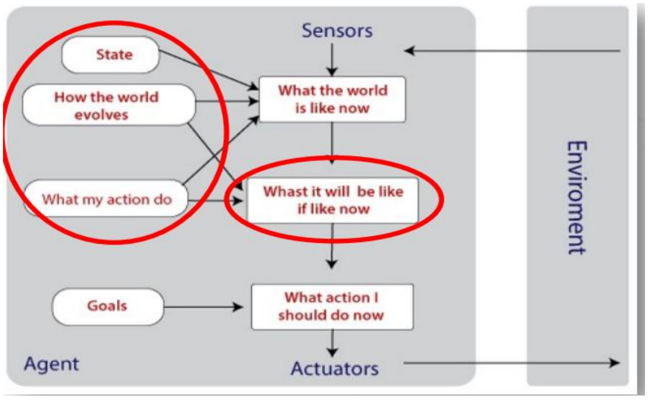
④. 构建方式：通用条件-行为规则解释器=>特定特务环境具体化

2) **基于模型的反射智能体**



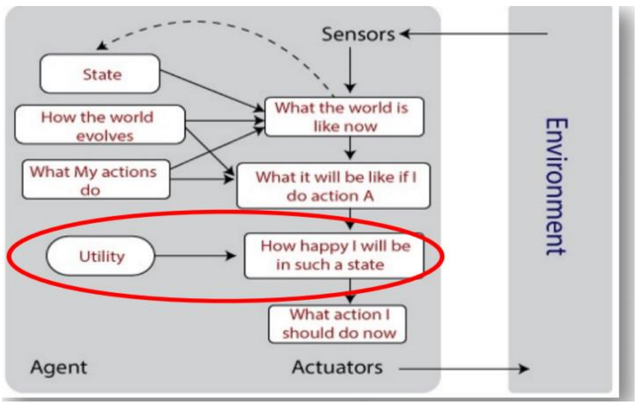
①. 内含“世界知识”模型，可查找条件与当前情况匹配规则执行动作
②. 使用内部状态跟踪（智能体行为怎么影响世界）
③. 需要的知识：世界如何独立于智能体演化+基于感知历史的当前状态表示

3) **基于目标的智能体**



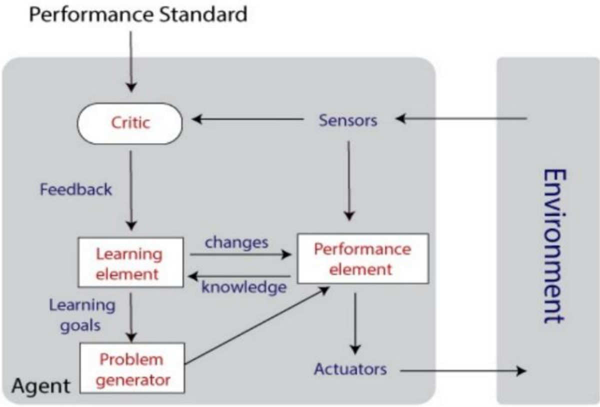
①. 相比基于模型的反射智能体，除了当前环境状态知识还加入了目标状态信息
②. 采取每个动作都是为了缩小与目标的距离

4) **基于效用的智能体**



①. 相比基于目标的反射智能体，除了当前环境状态知识，还增加了一个“效益测量”部件（也就是基于达到目标的最佳路径）
②. 使用效用函数量化目标
③. 基于世界模型、效用函数、感知序列选择期望效用最大化行动

5) **学习智能体**:



①. 从基本知识开始行动，依据奖励/惩罚等反馈及过去经验进行学习

②. 组成元件:

a) 评判元件：根据固定性能标准确定智能体性能度量

b) 性能元件：接受感知信息&决策外部动作

c) 学习元件：利用评判元件反馈，从环境中学习确定如何修改性能元件

d) 问题产生器：负责得到新的、有信息的经验，并建议探索性行为

13. 智能体应用：信息搜索、导航智能、医疗诊断、自动驾驶等

14. **2024 年真题**:

选择:

9. 自动驾驶 PEAS，里程表不是传感器。

解析：错误，由 PPT 关于 PEAS 自动驾驶的举例，里程表是传感器

10. 智能体分类（看原理图选择，ppt 原图）

解析：智能体共有五种分类，参照上面五张原理图熟悉