人工智能导论

腾讯云人工智能特色班课程

主讲人：高 灿，致腾楼936

（davidgao@szu.edu.cn）

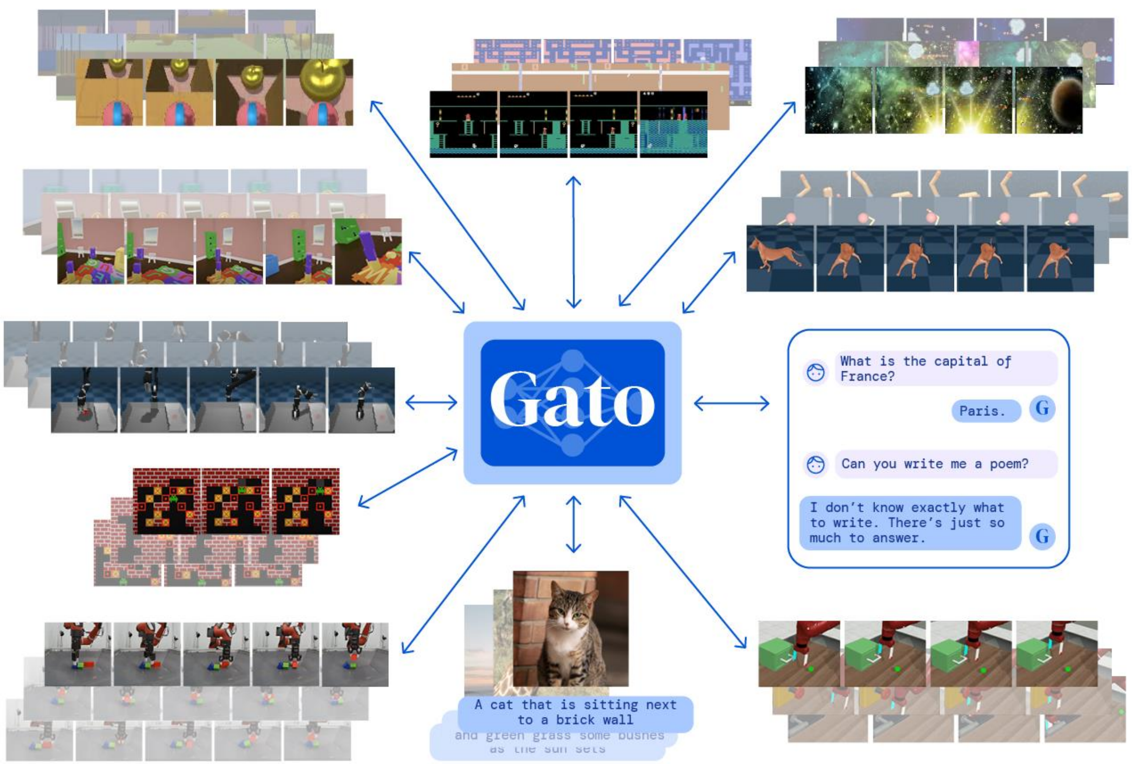
时 间：周三晚上11-12节 致理楼L1-306（理论）

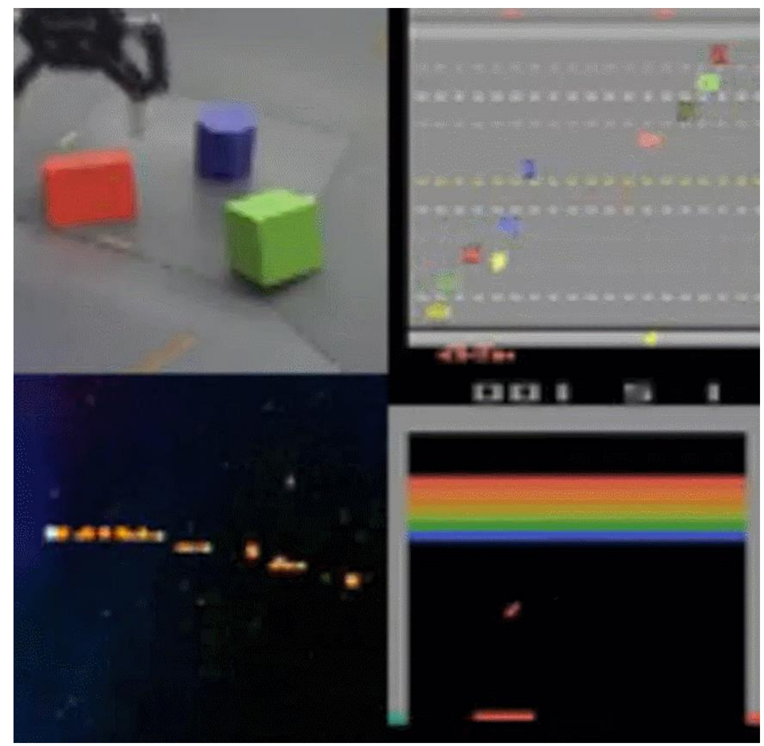
周三晚上13-14节(单) 致腾楼318（实验）

深圳大学 计算机与软件学院
College of Computer Science and Software Engineering of Shenzhen University

DeepMind「通才」AI智能体Gato来了，多模态、多任务，受大语言模型启发

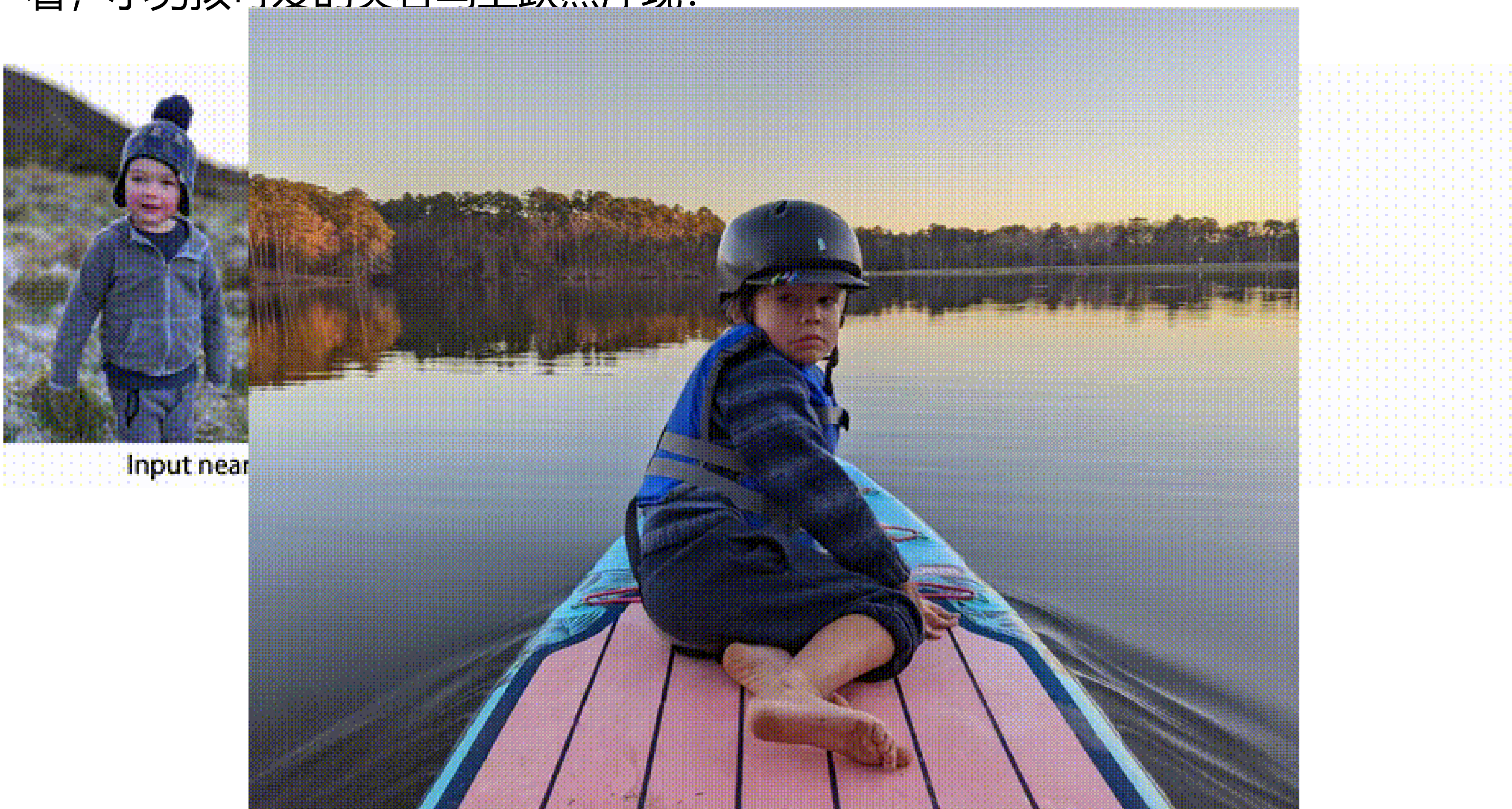
在写文章、画图之后，AI 大模型现在又同时有了打游戏的能力。不禁在想，DeepMind 的智能体 Gato 未来还能玩出哪些花活？受大规模语言建模的启发，Deepmind 应用类似的方法构建了一个单一的「通才」智能体 **Gato**，它具有多模态、多任务、多具身（embodiment）特点。





只需2张照片就能2D变3D，这个AI能自己脑补蜡烛吹灭过程，一作二作均是华人 | CVPR 2022

2张废片啪地一合！错过的精彩瞬间立刻重现，还能从2D升到3D效果。
看，小男孩可爱的笑容马上跃然浮现：



AI为人类开药方：准确预测9000名癌症患者适用药物！成果登上Nature子刊，出自华人团队

只需一个AI，9808名癌症患者对药物的临床反应，全能预测。而且结果和临床观察表现一致。这就是由纽约市立大学Lei Xie团队带来的最新成果CODE-AE（context-aware deconfounding autoencoder）。

Article | Open Access | Published: 17 October 2022

A context-aware deconfounding autoencoder for robust prediction of personalized clinical drug response from cell-line compound screening

Di He, Qiao Liu, You Wu & Lei Xie

Nature Machine Intelligence (2022) | Cite this article

Metrics

量子位

目录

1. 基于知识的智能体

1. 基于知识的智能体

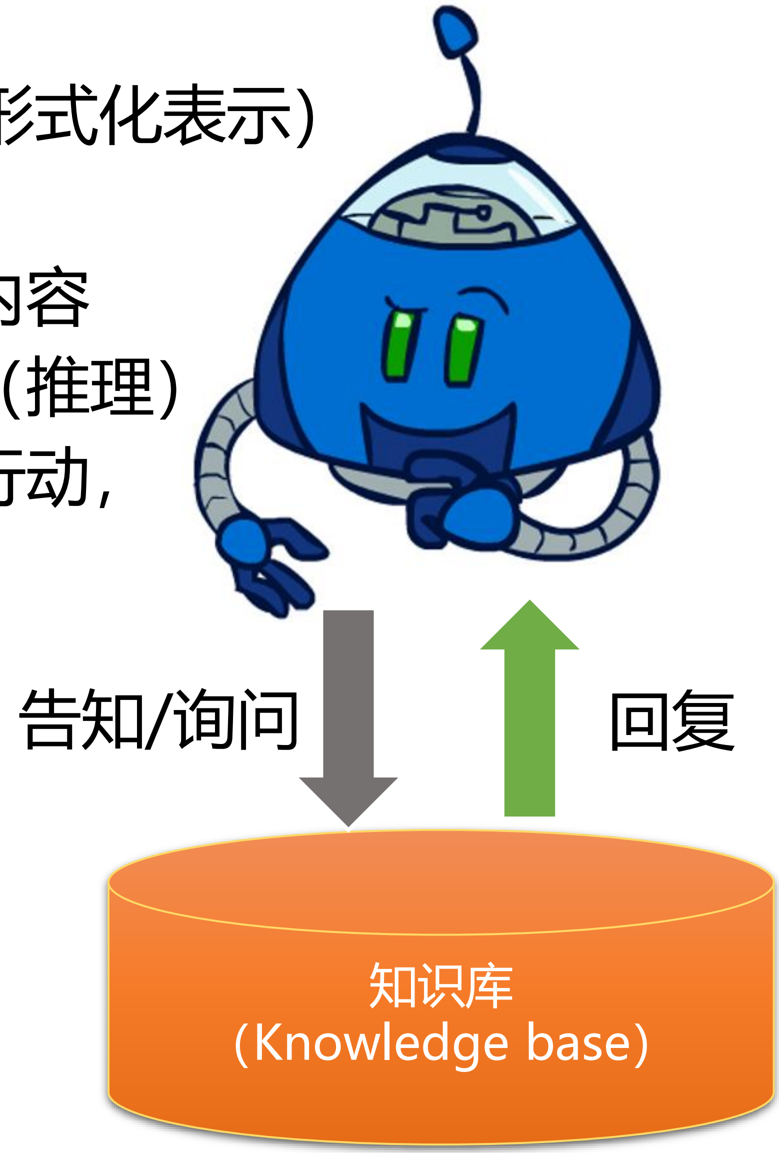
知识库KB = 语句α的集合（知识形式化表示）

陈述性方法构建智能体

- 智能体告诉知识库它感知到的内容
- 询问知识库应该执行什么行动（推理）
- 智能体告诉知识库它所选择的行动，并执行该行动

智能体需：

- 表示状态和行动
- 集成新的感知
- 更新问题的内在状态表示
- 推导问题的隐含特性
- 推导适当的行动



1. 基于知识的智能体

实例：怪物世界

评价

金子：+1000, 死亡：-1000

每步：-1, 射击：-10

环境（4*4）

靠近怪兽房间（非对角）有臭气；

靠近无底洞的房间有微风；

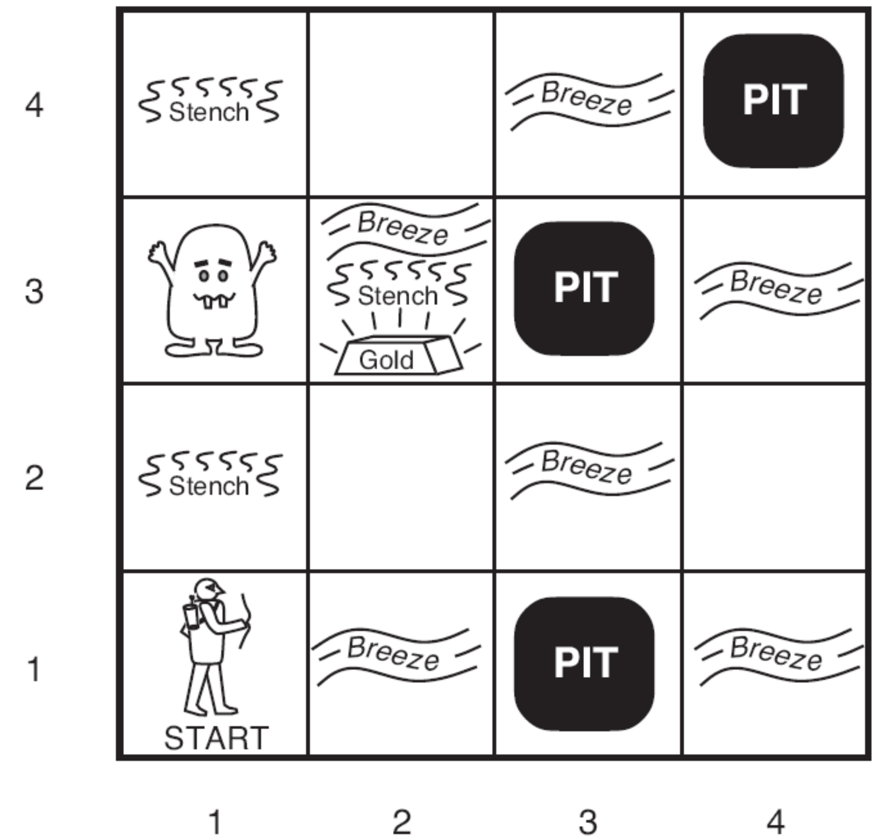
在金子房间金光闪闪；

如果碰到怪物，可以射死它；

只有一支箭；

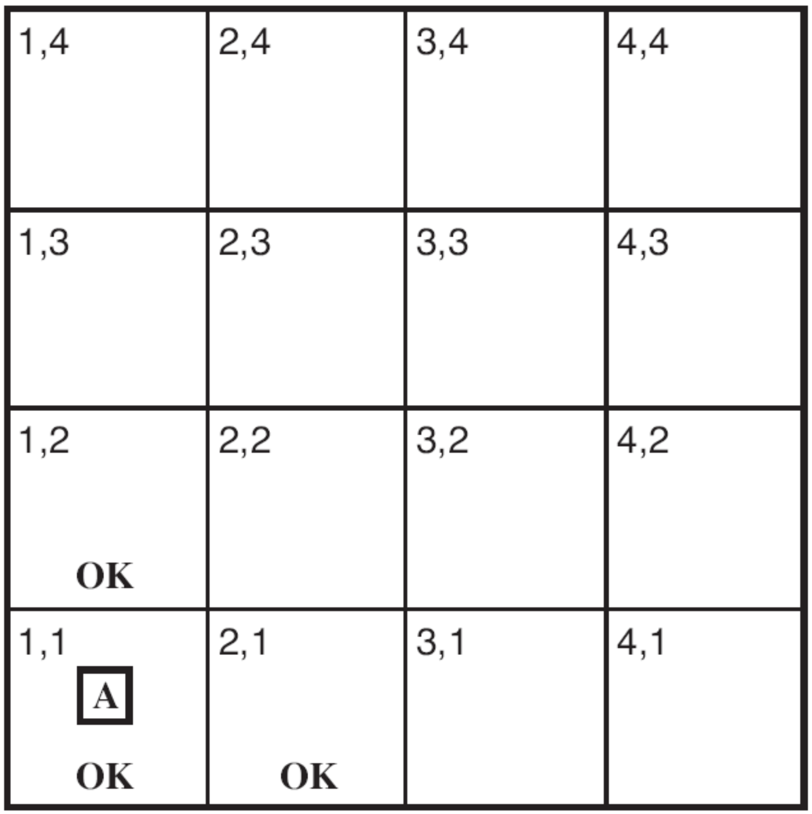
感知器：恶臭，微风，闪光

行动: 左转，右转，向前，射击，抓取，释放



1. 基于知识的智能体

实例：怪物世界

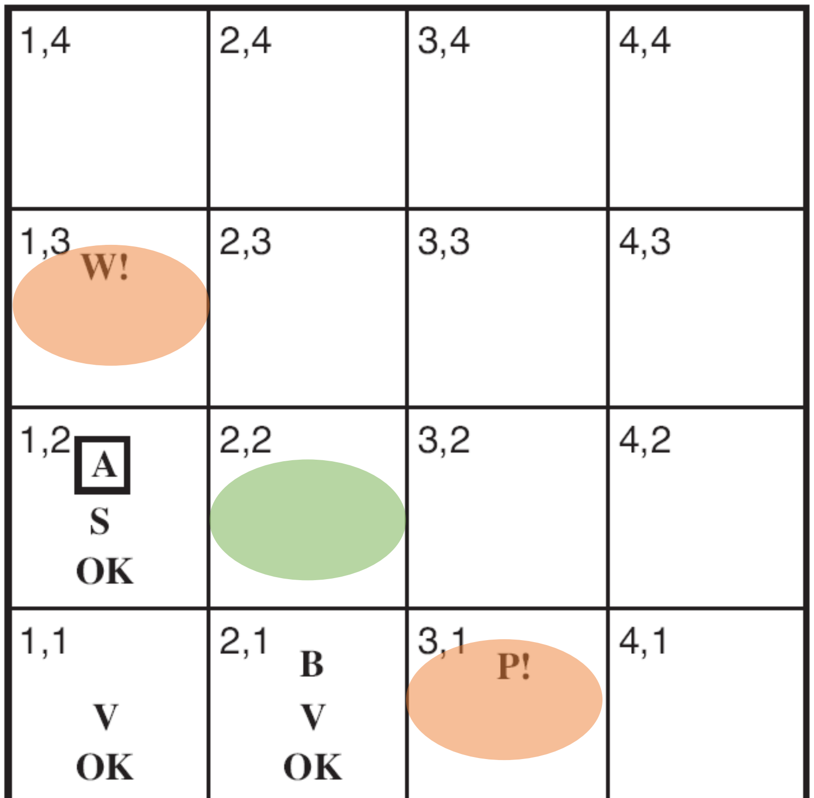


A = 智能体
B = 微风
G = 闪光
OK = 安全房间
P = 无底洞
S = 臭气
V = 已访问
W = 怪物

[1, 1]: 安全-[无，无，无]

1. 基于知识的智能体

实例：怪物世界



A = 智能体
B = 微风
G = 闪光
OK = 安全房间
P = 无底洞
S = 臭气
V = 已访问
W = 怪物

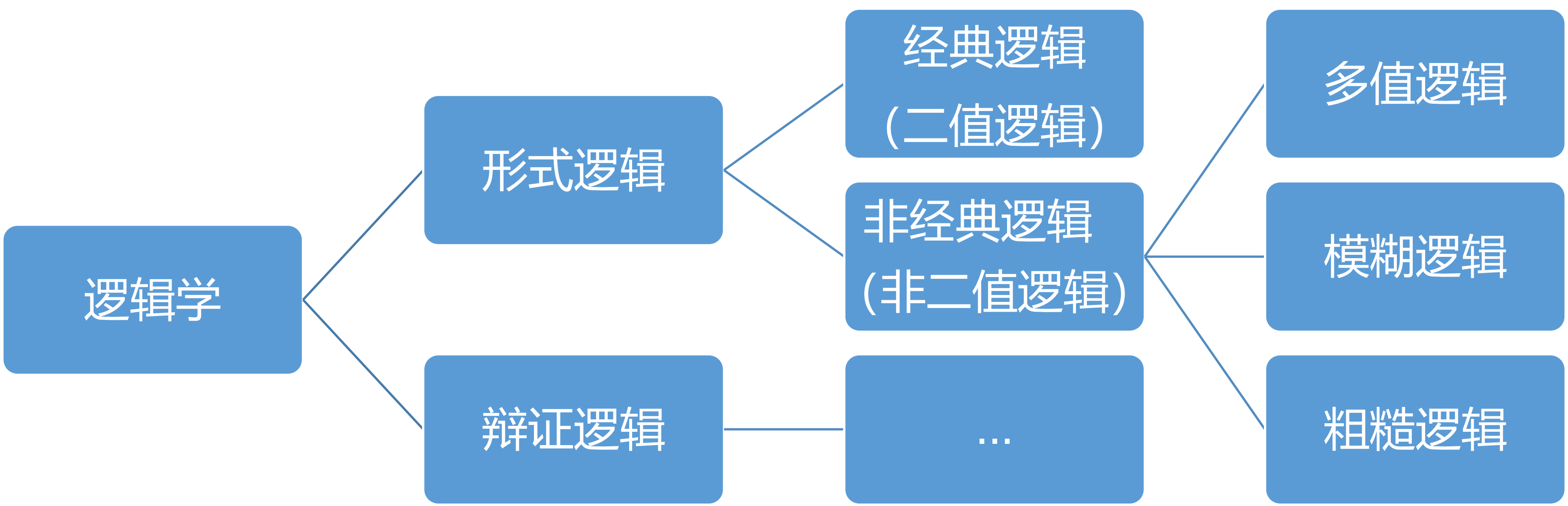
[1, 2]: 风险-[臭气，无，无] … [2, 3]: 风险-[臭气，微风，闪光]

智能体根据可用信息得出结论-逻辑推理

2. 命题逻辑

逻辑（Logic）是外来词，指的是思维的规律和规则

- 狭义上逻辑既指思维的规律，也指研究思维规律的学科即逻辑学
- 广义上逻辑泛指规律，包括思维规律和客观规律



https://baike.baidu.com/item/%E9%80%BB%E8%BE%91/543?fr=aladdin

2. 命题逻辑

实例：逻辑推理

2. 命题逻辑

逻辑定义（人工智能领域）：是一种操作事实以便得出真实结论的形式化系统-《人工智能-一种现代化方法》

- “The tool for distinguishing between the true and the false” – Averroes (12世纪).

语法：构造有效语句的规则

- 例：$x + 2 \geq y$ （√）； $\geq x2y +$ (×)

语义：语句的意义

- 语义定义了语句的真值
- 例：当$x = 5$ 且$y = 7$时，$x + 2 \geq y$ 为真

2. 命题逻辑

命题（Proposition）是一个非真即假的陈述句，是推理的基本要素

- 陈述句
- 可判断真值且真值唯一

例1：1）深圳是一线城市

2）这里的风景真美啊！

3）你吃饭了吗？

例2：1）深圳大学是一所高等教育学府；

2） 1+1=10

3）火星上有生命

命题一般用大写英文符号表示，P，Q，R

P：雪是白色的，Q：2+2=3， R：1+x > y

2. 命题逻辑

语法与语义：

原子(Atomic)命题: 一个表示真假断言的命题符号P（T或F）

否定(Negation): 设P是命题，P的否定也是命题，记为¬P

符号¬称为**否定连接词**，简称**否定词**，称¬P为P的否定式

P：爱因斯坦是物理学家，¬ P：爱因斯坦不是物理学家;

P：黑板是黑的，¬ P：黑板不是黑的;

P	¬ P
T	F
F	T

2. 命题逻辑

析取(Disjunction): 设P、Q是命题，P与Q的析取也是一个命题，记为P ∨ Q，称为**析取式**

符号 ∨ 称为**析取连接词**，简称**析取词**，称P 、Q为**析取子句**

P：他在教室，Q：他在寝室， P ∨ Q：他**要么**在教室**要么**在寝室

P：开关坏了，Q：灯泡坏了， P ∨ Q：开关坏了**或者**灯泡坏了

P	Q	P ∨ Q
T	T	T
T	F	T
F	T	T
F	F	F

一真为真
两假为假

表示两个命题至少其一成立，日常汉语中“或者”， “不是...就是...”，“要么...要么...”等可用析取符号 ∨ 表示。

2. 命题逻辑

合取(Conjunction): 设P、Q是命题，P与Q的合取也是一个命题，记为P ∧ Q，称为**合取式**

符号 ∧ 称为**合取连接词**，简称**合取词**，称P 、Q为**合取子句**

P：今天下雨，Q：明天下雨， P ∧ Q：今天和明天**都**下雨；

P：他跑得快，Q：他跳得高， P ∧ Q：他**不仅**跑得快**而且**跳得高；

P	Q	P ∧ Q
T	T	T
T	F	F
F	T	F
F	F	F

两真为真
一假为假

表示两个命题同时成立，日常汉语中“都”， “并且”，“不仅…而且…” ,” 即…又…”等可用合取符号 ∧ 表示。

2. 命题逻辑

蕴涵(Implication): 设P、Q是命题， P蕴涵Q也是一个命题，记为$P \to Q$

符号 $\to$ 称为**蕴涵连接词**，简称**蕴涵取词**，称P为蕴涵式的**前件**，Q为蕴涵式的**后件**

P：该动物是只鸟，Q：它有翅膀， $P \to Q$：**只要**该动物是只鸟，它**就**有翅膀

P：函数f(x)可导，Q：函数f(x)连续， $P \to Q$：**若**函数f(x)可导，**则**函数f(x)连续

P：他是不想当厨子的裁缝，Q：他不是好司机， $P \to Q$：**如果**他是不想当厨子的裁缝，**则**他不是好司机。

P：明天下雨，Q：地球以外有生命， $P \to Q$：**如果**明天下雨，**那么**地球以外有生命

表示P是Q的充分条件，或者说Q是P的必要条件。日常汉语中 "只要…就…"， "如果..则…" 等可用蕴涵$P \to Q$表示。

2. 命题逻辑

数理逻辑中的"蕴涵"与日常所说的"蕴涵"意义并不完全一致。

日常用语中的蕴涵前件和后件一定是**意义上有关连**的两个句子。

数理逻辑中**内容毫不相干**的命题可以构成一个新命题，其真值完全由其构成命题的真假所确定。

P	Q	$P \to Q$
T	T	T
T	F	F
F	T	T
F	F	T

两真为真
前假为真

2. 命题逻辑

等价(Biconditional): 设P、Q是命题，P等价于Q也是一个命题，记为P ↔ Q

符号 ↔ 称为**等价连接词**，简称**等价词**，称P为等价式的**左端**，Q为等价式的**右端**

P：三角形等边，Q：三角形等角， P ↔ Q：三角形等边**当且仅当**三角形等角

P	Q	P ↔ Q
T	T	T
T	F	F
F	T	F
F	F	T

两同为真
两异为假

表示P与Q的互为充分必要条件。日常汉语中“当且仅当”，“除非不…否则…” 等可以写成P ↔ Q的形式

2. 命题逻辑

真值表：

P	Q	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \rightarrow Q$	$P \leftrightarrow Q$
T	T	F	T	T	T	T
T	F	F	F	T	F	F
F	T	T	F	T	T	F
F	F	T	F	F	T	T

$\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$称为**逻辑接连词**(connectives)

2. 命题逻辑

命题与复合命题：

单个命题称为**原子命题**

包含连接词的命题称为**复合命题或复合命题公式**

组成复合命题的称为**成分命题**

复合命题公式的真假值有赖于成分命题的真假值，而不注重成分命题的内容。

2. 命题逻辑

命题符号化：利用连接词将日常语言转化成数理逻辑中的形式化命题的过程

命题符号化是运用数理逻辑解决实际问题的基本出发点

例1：小张既聪明又勤奋，所以他的成绩好

解：P：小张聪明

Q：小张勤奋

R：小张的成绩好

符号化的结果为：$(P \wedge Q) \rightarrow R$

例2：小王总是在图书馆看书，除非他病了或者图书馆不开门

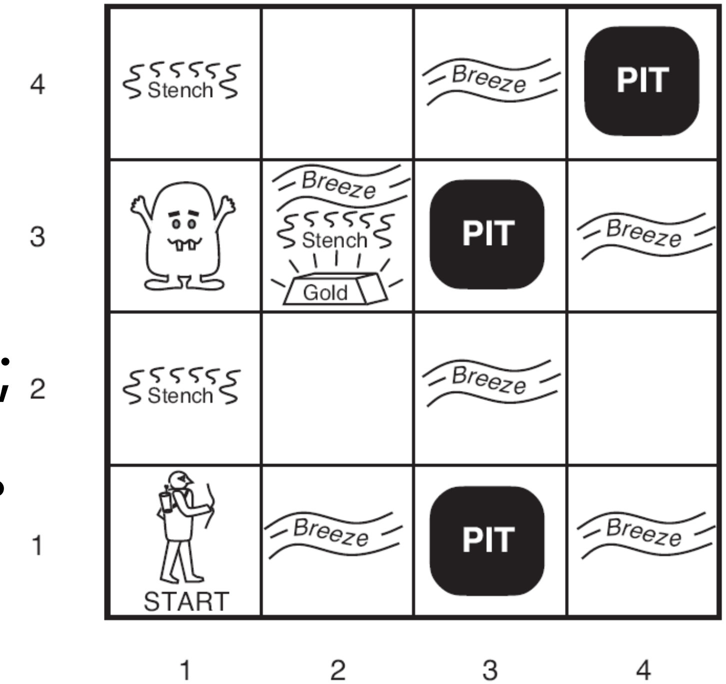
符号化的结果为：$\neg (P \vee \neg Q) \rightarrow R$

2. 命题逻辑

命题符号化-怪物世界

怪物世界的知识库:

- 如果$[x,y]$中有无底洞，则$P_{x,y}$为真;
- 如果$[x,y]$中有怪兽，则$W_{x,y}$为真;
- 如果$[x,y]$中感知到微风，则$B_{x,y}$为真;
- 如果$[x,y]$中感知到臭气，则$S_{x,y}$为真。



根据当前状态信息，能否推出$\neg P_{1,2}$

- R_1: $\neg P_{1,1}$
- R_2: $B_{1,1} \leftrightarrow (P_{1,2} \vee P_{2,1})$
- R_3: $\neg B_{1,1}$

2. 命题逻辑

逻辑等价性

等价性：如果命题P和Q在任何情况下都有**相同的真值**，则称P和Q是**逻辑等价**的，记为$P \Leftrightarrow Q$

主要逻辑等价式：

(1) 交换律

$$P \vee Q \Leftrightarrow Q \vee P$$

$$P \wedge Q \Leftrightarrow Q \wedge P$$

(2) 结合律

$$(P \vee Q) \vee R \Leftrightarrow P \vee (Q \vee R)$$

$$(P \wedge Q) \wedge R \Leftrightarrow P \wedge (Q \wedge R)$$

2. 命题逻辑

主要逻辑等价式：

(3) 分配律

$$P \vee (Q \wedge R) \Leftrightarrow (P \vee Q) \wedge (P \vee R)$$

$$P \wedge (Q \vee R) \Leftrightarrow (P \wedge Q) \vee (P \wedge R)$$

(4) 得摩根律

$$\neg(P \vee Q) \Leftrightarrow \neg P \wedge \neg Q$$

$$\neg(P \wedge Q) \Leftrightarrow \neg P \vee \neg Q$$

(5) 双重否定律

$$\neg\neg P \Leftrightarrow P$$

(6) 吸收律

$$P \vee (P \wedge Q) \Leftrightarrow P$$

$$P \wedge (P \vee Q) \Leftrightarrow P$$

2. 命题逻辑

主要逻辑等价式：

(7) 否定律

$$P \vee \neg P \Leftrightarrow T$$

$$P \wedge \neg P \Leftrightarrow F$$

(8) 连接词化归律 $P \rightarrow Q \Leftrightarrow \neg P \vee Q$

(9) 逆否律 $P \rightarrow Q \Leftrightarrow \neg Q \rightarrow \neg P$

2. 命题逻辑

永真蕴涵： 如果命题逻辑P和Q使公式$P \rightarrow Q$永真，则称公式P永真蕴涵Q，记为$P \Rightarrow Q$，称Q为P的逻辑结论，P为Q的前提

永真蕴涵式:

(1) 假言推理

$$P, P \rightarrow Q \Rightarrow Q$$

(2) 拒取式推理

$$\neg Q, P \rightarrow Q \Rightarrow \neg P$$

(3) 假言三段论

$$P \rightarrow Q\ , Q \rightarrow R \Rightarrow P \rightarrow R$$

3. 基于逻辑的确定性推理

推理：从初始证据出发，按某种策略不断运用知识库中的已知知识，逐步推出结论的过程

演绎推理（Deductive Reasoning）：一般到个别

大前题：M-P（M是P）

小前题：S-M（S是M）　　**苏格拉底三段论**

结论：S-P（S是P）

归纳推理（Inductive Reasoning）：个别到一般

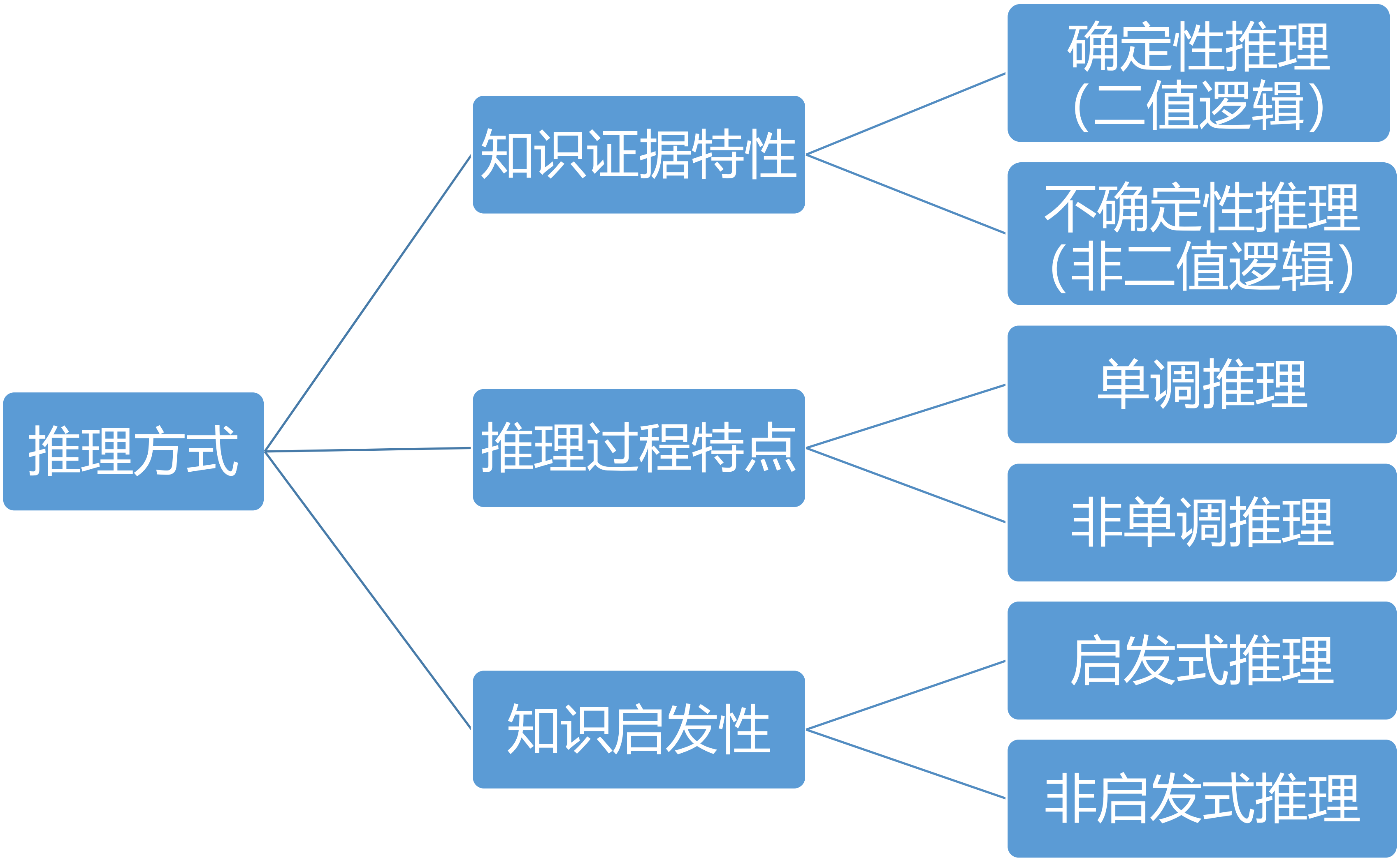
具体：铜、铁、铝、金等金属能导电

一般：金属能导电

默认推理（Default Reasoning）

知识不完全情况下，默认某些条件成立（可能被推翻）

3. 基于逻辑的确定性推理



3. 基于逻辑的确定性推理

- 推理过程
  - 正向推理
  - 逆向推理
  - 混合推理
  - 双向推理

正向推理	逆向推理	混合推理	双向推理
知识⇒结论	假设⇒验证	知识、假设、验证⇒结论	知识⇒结论 ⇑⇓ 假设⇒验证

冲突消解（Conflict Resolution）

3. 基于逻辑的确定性推理

自然演绎推理：从一组已知为真的事实出发，直接运用经典逻辑中的推理规则推出结论的过程

推理基础：逻辑等价式+永真蕴涵

推理规则：

P规则：推理的任何步骤都可引入前提

T规则：可将永真蕴涵式引入推理

逻辑等价：$P \leftrightarrow Q \Rightarrow (P \to Q) \wedge (Q \to P)$

合取消除：$P \wedge Q \Rightarrow P, Q$

假言推理：$P, P \to Q \Rightarrow Q$

拒取式推理：$\neg Q, P \to Q \Rightarrow \neg P$

假言三段论：$P \to Q\, , Q \to R \Rightarrow P \to R$

注意： $Q, P \to Q \Rightarrow P$　　　$\neg P, P \to Q \Rightarrow \neg Q$

3. 基于逻辑的确定性推理

基于知识库的确定性推理

初始状态：初始知识库

行动：由应用于语句的所有推理规则组成，需匹配推理规则的前件

结果：行动的结果是将推理规则的下半部分的语句加入知识库

目标：包含要证明或推理的语句状态

3. 基于逻辑的确定性推理

例 1：命题逻辑自然演绎推理

已知：$A, B, A \to C, B \wedge C \to D, D \to Q$

结论：Q

1、 A (P规则)

2、 $A \to C$ (P规则)

3、 C (T规则，假言推理)

4、 B (P规则)

5、 $B \wedge C \to D$ (P规则)

6、 D (T规则，假言推理)

7、 $D \to Q$ (P规则)

8、 Q (T规则，假言推理)

3. 基于逻辑的确定性推理

例 2：怪物世界推理

已知： R_1: $\neg P_{1,1}$, R_2: $B_{1,1} \boxed{\leftrightarrow} (P_{1,2} \vee P_{2,1})$, R_3: $\neg B_{1,1}$

结论：$\neg P_{1,2}$

1、R_2: $B_{1,1} \boxed{\leftrightarrow} (P_{1,2} \vee P_{2,1})$ (P规则)

2、R_4: $(B_{1,1} \rightarrow (P_{1,2} \vee P_{2,1})) \wedge ((P_{1,2} \vee P_{2,1}) \rightarrow B_{1,1})$ (逻辑等价)

3、R_5: $(P_{1,2} \vee P_{2,1}) \rightarrow B_{1,1}$ (合取消除)

4、R_3: $\neg B_{1,1}$ (P规则)

5、R_6: $\neg(P_{1,2} \vee P_{2,1})$ (拒取式推理)

6、R_7: $\neg P_{1,2} \wedge \neg P_{2,1}$ (De Morgan定律)

7、R_8: $\neg P_{1,2}$ (合取消除)

3. 基于逻辑的确定性推理

归结反演推理

归结原理：设$C_1 = L \vee \alpha$与$C_2 = \neg L \vee \beta$是知识库中的任意两个析取子句，则可从C_1和C_2中分别消去L和$\neg L$，并归结成**一个新子句**$C_{12} = (\alpha \vee \beta)$。$C_{12}$称为$C_1$和$C_2$的**归结式**，$C_1$和$C_2$称为$C_{12}$的**亲本(父辈)子句**

命题假言推理：亲本子句 P $\quad$ $\neg P \vee Q$ $\underline{(P \rightarrow Q)}$

归结式 Q

合并：亲本子句 $P \vee Q$ $\quad$ $\neg P \vee Q$

归结式 $Q \vee Q = Q$

验证：三段论的正确性$P \rightarrow Q\,,\, Q \rightarrow R \Rightarrow P \rightarrow R$

3. 基于逻辑的确定性推理

归结反演推理

子句(Clause)：任何命题或命题的**析取式**

空子句：不包含任何命题的子句，NIL表示

空子句是永假的，不可满足的（**归结反演的目标**）

合取范式：以子句的合取式表示的语句称为合取范式(**CNF**)

归结原理可应用于析取子句，将知识库所有语句转换成合取范式$P_1 \wedge P_2 \wedge \ldots \wedge P_n$，其所有析取子句为**子句集**$S=\{P_1, P_2, \ldots, P_n\}$

例：求$P \leftrightarrow (Q \vee R)$的子句集

1. 消等价 $(P \rightarrow (Q \vee R)) \wedge ((Q \vee R) \rightarrow P)$
2. 去蕴含 $(\neg P \vee Q \vee R) \wedge (\neg(Q \vee R) \vee P)$
3. 移否定 $(\neg P \vee Q \vee R) \wedge ((\neg Q \wedge \neg R) \vee P)$
4. 合取式 $(\neg P \vee Q \vee R) \wedge (\neg Q \vee P) \wedge (\neg R \vee P)$
5. 子句集$\{\neg P \vee Q \vee R, \neg Q \vee P, \neg R \vee P\}$

3. 基于逻辑的确定性推理

试一试：求下列命题公式的子句集

$[T \to [(P \to Q) \wedge \neg (R \to S)]]$

1. 去蕴含 $[\neg T \vee [(\neg P \vee Q) \wedge \neg (\neg R \vee S)]]$
2. 移否定 $[\neg T \vee [(\neg P \vee Q) \wedge (R \wedge S)]]$
3. 合取式 $(\neg T \vee \neg P \vee Q) \wedge (\neg T \vee R) \wedge (\neg T \vee S)$
4. 子句集 $\{\neg T \vee \neg P \vee Q, \neg T \vee R, \neg T \vee S\}$

3. 基于逻辑的确定性推理

归结反演推理

推理基础：

(1) Q为$P_1, P_2, P_3, \ldots, P_n$的逻辑结论，当且仅当$(P_1 \wedge P_2 \wedge P_3 \wedge \ldots \wedge P_n) \wedge \neg Q$是不可满足的

(2)命题公式不可满足的充要条件是其子句集不可满足

归结反演：应用归结原理问题证明的过程

一般步骤：

(1) 已知前提事实命题公式集表示F

(2) 结论以命题公式否定式表示$\neg Q$

(3) 命题公式集$\{F, \neg Q\}$化为**子句集**S

(4) 对子句集S中的子句进行**归结**，并将结果并入到S

(5) 反复进行归结，若出现空子句，结论得证

3. 基于逻辑的确定性推理

归结反演推理

算法描述：

function PL-RESOLUTION(KB, α) **returns** *true* or *false*
  inputs: KB, the knowledge base, a sentence in propositional logic
    α, the query, a sentence in propositional logic

  $clauses \leftarrow$ the set of clauses in the CNF representation of $KB \wedge \neg\alpha$
  $new \leftarrow \{\ \}$
  loop do
    for each C_i, C_j **in** $clauses$ **do**
      $resolvents \leftarrow$ PL-RESOLVE(C_i, C_j)
      if $resolvents$ contains the empty clause **then return** *true*
      $new \leftarrow new \cup resolvents$
    if $new \subseteq clauses$ **then return** *false*
    $clauses \leftarrow clauses \cup new$

子句集

归结

空子句

3. 基于逻辑的确定性推理

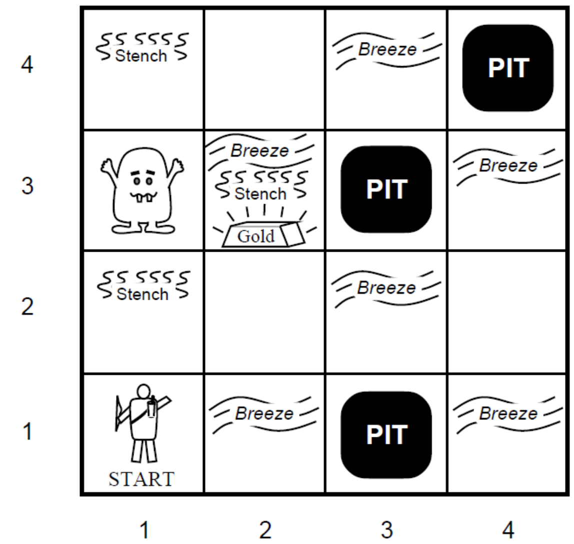
归结反演推理-怪物世界

智能体位于[1,1]，无微风，相关的知识库为：

KB=R_2: $B_{1,1} \leftrightarrow (P_{1,2} \vee P_{2,1}) \wedge R_3$: $\neg B_{1,1}$

证明：α，也即$\neg P_{1,2}$

将$KB \wedge \neg \alpha$转换为CNF的子句集



$\boxed{\neg P_{2,1} \vee B_{1,1}}$ $\boxed{\neg B_{1,1} \vee P_{1,2} \vee P_{2,1}}$ $\boxed{\neg P_{1,2} \vee B_{1,1}}$ $\boxed{\neg B_{1,1}}$ $\boxed{P_{1,2}}$

3. 基于逻辑的确定性推理

命题逻辑局限性:

命题逻辑限定原子命题是不能细分的整体

例1：（三段论）凡人都是要死的，苏格拉底是人，苏格拉底是要死的

P：凡人都是要死的

Q：苏格拉底是人

R：苏格拉底要死

命题逻辑：（P ∧ Q）→ R ?

例2：P：小王是学生；

Q：小李是学生；

解1：

凡人都要死的：$(\forall x)P(x)$

苏格拉底是人：$x = Socrates$

苏格拉底是要死的：$P(Socrates)$

解2：

凡人都要死的：$(\forall x)(Human(x) \rightarrow Death(x))$

苏格拉底是人：$Human(Socrates)$

苏格拉底是要死的：$Death(Socrates)$

命题逻辑描述能力和泛化能力有限

需对命题进一步分析-**谓词逻辑**

4. 基于概率的不确定推理

不确定性是普遍现象

- 知识认识不够
- 背景知识不足
- 信息描述含糊
- 信息含有噪声
- 信息不完整
- 信息不一致
- 规划是模糊的
- 推理能力不足
- 方案不唯一

在人类的知识和思维行为中，**精确性**只是相对的，**不精确性**才是绝对的。知识工程需要适应不同特点的不确定性知识表示和推理方法。

随机性、含糊性、不精确性、不完整性、不一致性

4. 基于概率的不确定推理

不确定性推理（Reasoning with uncertainty）：是一种建立在非经典逻辑基础上的基于不确定性知识的推理，它从**不确定性的初始证据**出发，通过运用**不确定性知识**，推出具有**一定程度**不确定性的和合理的或近乎合理的结论。

不确定性的研究内容

1）不确定性的表示

（1）知识不确定性的表示

（2）证据不确定性的表示

（3）结论不确定性的表示

2）不确定性的度量

不确定性的程度量化

他不可能/有可能/很大可能一定/赢得比赛。

他0%、30%、 85%、100%赢得比赛

4. 基于概率的不确定推理

- 不确定性推理
  - 概率论
    - 可信度方法
    - 证据理论
    - 贝叶斯理论
  - 集合论
    - 模糊集
    - 阴影集
    - 粗糙集
  - …
    - …

5. 贝叶斯理论

概率论基础

随机现象：

在一定条件下，必然发生或必然不发生的现象，称为**确定现象**

在一定条件下，可能出现这个结果，也可能出现那个结果，而且不能事先确定出现哪一个结果的现象，称为**随机现象**

随机现象结果不可预知，但有其规律可循(统计规律)

随机试验：

观察一定条件下发生的随机现象称为**随机试验**

1）可重复试验；

2）所有结果事先明确，且不止一个；

3）每次试验结果不可预知

5. 贝叶斯理论

样本空间、样本点以及随机事件：

样本空间是指试验的所有可能结果的集合

样本点是指样本空间中的单个结果

随机事件是指样本空间的子集

他们三者的关系应该是： 样本点⊂随机事件⊂样本空间

例：抛两枚硬币观察它们的正反面情况

样本空间：{(正, 正)，(正, 反)，(反, 正)，(反, 反)}

样本点：(正，反)，随机事件：第一枚硬币为正面

5. 贝叶斯理论

频率与概率

随机事件A在n次实验中发生m次，比值m/n为随机事件A的**频率**

随机事件A发生可能性大小的数值称为随机事件A发生的**概率**，记为$P(A)$

频数、频率和概率区别：

频数是指在多次试验中某个事件出现的次数

频率是某个事件在整体实验中出现的次数占整体试验次数的比例，因此在试验过程中频率的值会改变，如果试验次数多的话，频率可能会在概率周围浮动；

概率是某个事件的客观出现的可能性，是一个固定值不因试验次数改变而改变；

频率和概率：频率是客观**试验所得**事件真实发生的比例，而概率是客观现实分析所得的事件发生可能性的**固定值**。

5. 贝叶斯理论

随机变量(离散和连续)

定义：设随机试验的样本空间是Ω={e}，X=X(e)是定义在样本空间Ω上的实值单值函数，称X=X(e)为随机变量

随机变量将随机试验中的**事件**转换为**数字**的一个抽象

不确定性的**随机变量**表示：

R=明天下雨?

T=温度冷还是热?

D=开车上班需要多长时间?

L=小明在哪里?

随机变量的**样本空间**：

R∈{真，假}，T∈{冷，热}

D∈[0, ∞]，L∈{所有可能的地方}

5. 贝叶斯理论

随机变量的概率分布：

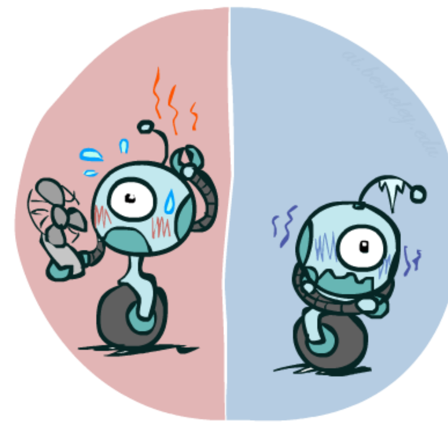
分配每个可能取值一个概率

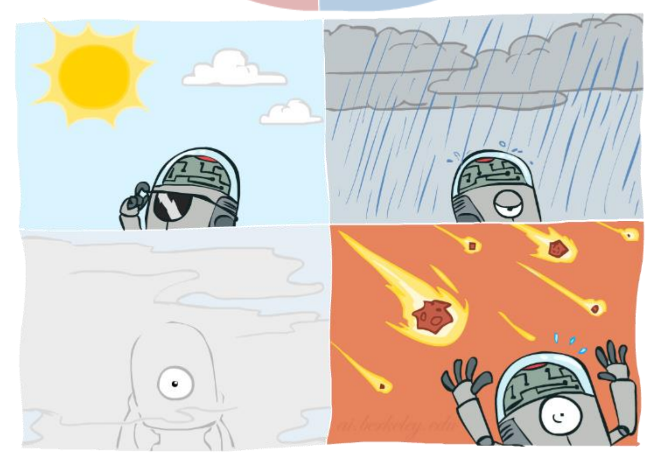
$P(T)$

T	P
hot	0.5
cold	0.5

$P(W)$

W	P
sun	0.6
rain	0.1
fog	0.3
meteor	0.0





概率分布是所有可能取值的概率表

$P(W = rain) = 0.1 \quad \forall x \;\; P(X = x) \geq 0 \quad \sum_x P(X = x) = 1$

5. 贝叶斯理论

联合分布：

一组随机变量的联合分布

$$P(X_1 = x_1, X_2 = x_2, \ldots X_n = x_n)$$

$$P(x_1, x_2, \ldots x_n)$$

约束： $P(x_1, x_2, \ldots x_n) \geq 0$

$$\sum_{(x_1, x_2, \ldots x_n)} P(x_1, x_2, \ldots x_n) = 1$$

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

d 个值域的 n 个随机变量分布大小为 (n^d)

5. 贝叶斯理论

联合分布：

事件：一组结果的集合

$$P(E) = \sum_{(x_1 \ldots x_n) \in E} P(x_1 \ldots x_n)$$

任何事件的概率可通过联合分布计算

- 天气为hot和sun的概率
- 天气为hot的概率
- 天气为hot或sun的概率

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

5. 贝叶斯理论

联合分布：

例：已知两随机变量的分布，求如下事件的概率

P(+x, +y) ?

$$P(X, Y)$$

X	Y	P
+x	+y	0.2
+x	-y	0.3
-x	+y	0.4
-x	-y	0.1

P(+x) ?

P(-y 或 +x) ?

5. 贝叶斯理论

边缘分布：

边缘分布是去除联合分布中变量的子表

边缘化：累加去除变量后的行

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

$$P(t) = \sum_s P(t, s)$$

$$P(s) = \sum_t P(t, s)$$

$P(T)$

T	P
hot	0.5
cold	0.5

$P(W)$

W	P
sun	0.6
rain	0.4

$$P(X_1 = x_1) = \sum_{x_2} P(X_1 = x_1, X_2 = x_2)$$

5. 贝叶斯理论

边缘分布：

例：已知联合分布，求边缘分布

$P(X,Y)$

X	Y	P
+x	+y	0.2
+x	-y	0.3
-x	+y	0.4
-x	-y	0.1

$$P(x)=\sum_y P(x,y)$$

$$P(y)=\sum_x P(x,y)$$

$P(X)$

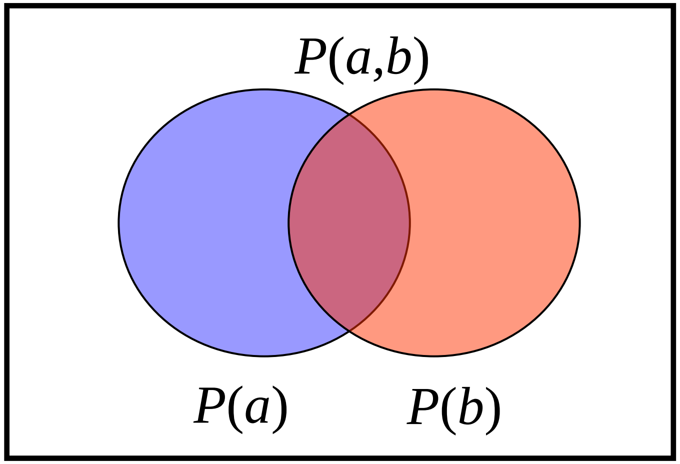
X	P
+x	
-x	

$P(Y)$

Y	P
+y	
-y	

5. 贝叶斯理论

条件概率： 事件b发生的条件下，事件a发生的概率



$$P(a|b) = \frac{P(a,b)}{P(b)}$$

$$P(T,W)$$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

$$P(W=s|T=c) = \frac{P(W=s,T=c)}{P(T=c)} = \frac{0.2}{0.5} = 0.4$$

$$= P(W=s,T=c) + P(W=r,T=c)$$
$$= 0.2 + 0.3 \quad = 0.5$$

5. 贝叶斯理论

条件概率：事件b发生的条件下，事件a发生的概率

例：已知联合分布，求条件概率

$$P(X,Y)$$

X	Y	P
+x	+y	0.2
+x	-y	0.3
-x	+y	0.4
-x	-y	0.1

$$P(a|b) = \frac{P(a,b)}{P(b)}$$

P(+x | +y) ?

P(-x | +y) ?

P(-y | +x) ?

5. 贝叶斯理论

条件分布：给定某些变量固定值的情况下，变量的概率分布

条件分布

$P(W|T)$

$$P(W|T = hot)$$

W	P
sun	0.8
rain	0.2

$$P(W|T = cold)$$

W	P
sun	0.4
rain	0.6

联合分布

$$P(T, W)$$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

5. 贝叶斯理论

乘法规则：

已知事件概率和条件概率，推导联合概率

$$P(x|y) = \frac{P(x,y)}{P(y)} \Longleftrightarrow P(y)P(x|y) = P(x,y)$$

$P(W)$

R	P
sun	0.8
rain	0.2

$P(D|W)$

D	W	P
wet	sun	0.1
dry	sun	0.9
wet	rain	0.7
dry	rain	0.3

$\Longleftrightarrow$

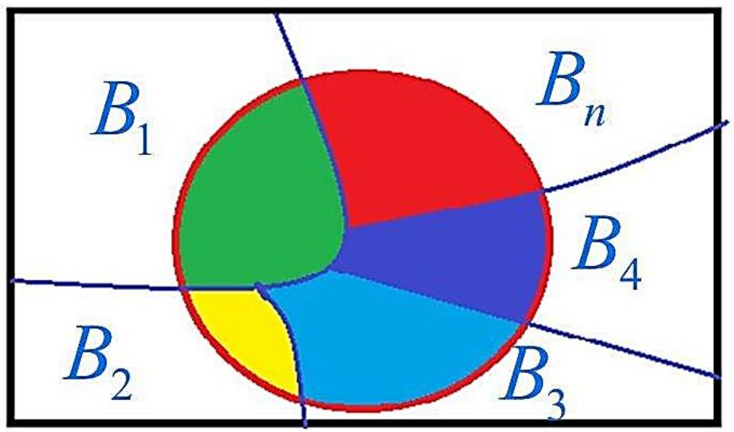
$P(D,W)$

D	W	P
wet	sun	0.08
dry	sun	0.72
wet	rain	0.14
dry	rain	0.06

5. 贝叶斯理论

全概率公式:

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



链式规则:

$$P(x_1, x_2, x_3) = P(x_1)P(x_2|x_1)P(x_3|x_1, x_2)$$

$$P(x_1, x_2, \ldots x_n) = \prod_i P(x_i|x_1 \ldots x_{i-1})$$

5. 贝叶斯理论

贝叶斯规则：

$$P(A,B) = P(A|B)P(B)$$
$$P(A,B) = P(B|A)P(A)$$

$$P(A|B) = \frac{P(B|A)}{P(B)}P(A)$$

后验概率　　调整因子　　先验概率

这是我的规则!

5. 贝叶斯理论

正向概率与逆向概率：

已知信息⇒事件信息

假设桶中有N个白球，M个黑球，伸手进去摸一次，摸出黑球的概率是多大？--**正向概率**

未知信息⇐观察信息

事先并不知道桶中黑白球的比例，随机摸出一个（或好几个）球，观察球的颜色之后，那么可以对袋子里面的黑白球的比例作出什么样的推测？ --**逆向概率**

5. 贝叶斯理论

先验概率与后验概率：

$$P(A|B) = \frac{P(B|A)}{P(B)} P(A)$$

逆向概率求解：

- 计算各种不同假设的可能性（后验概率）
- 计算最可能的假设（模型比较）

贝叶斯理论是机器学习的核心方法之一，在众多领域有广泛的应用

- 语音识别：根据麦克风的音频波形数据Y推测语音信息X
- 文字识别：根据扫描的图像数据Y推测用户手写体的文字X
- 邮件过滤：根据收到的邮件文本Y推测邮件的种类X
- 医疗诊断：根据症状Y推测疾病的类型X

5. 贝叶斯理论

先验概率与后验概率：

例：一所学校里面有 60% 的男生，40% 的女生。男生总是穿长裤，女生则一半穿长裤一半穿裙子。在校园中从背影看到一个穿长裤的学生，求该学生是男生的概率（源自维基百科）

解法1(公式法)：A：性别，B：类型

P(A=男生) = 0.6，P(A=女生) = 0.4

P(B=长裤) = 0.8，P(B=裙子) = 0.2

P(B=长裤|A=男) = 1.0，P(B=长裤|A=女) = 0.5

P(A=男|B=长裤) = 【P(B=长裤|A=男) / P(B=长裤)】* P(A=男)

0.75 = 1.0 / 0.8 * 0.6

后验概率 = 调整因子 * 先验概率

P(A=女|B=长裤) = 【P(B=长裤|A=女) / P(B=长裤)】* P(A=女)

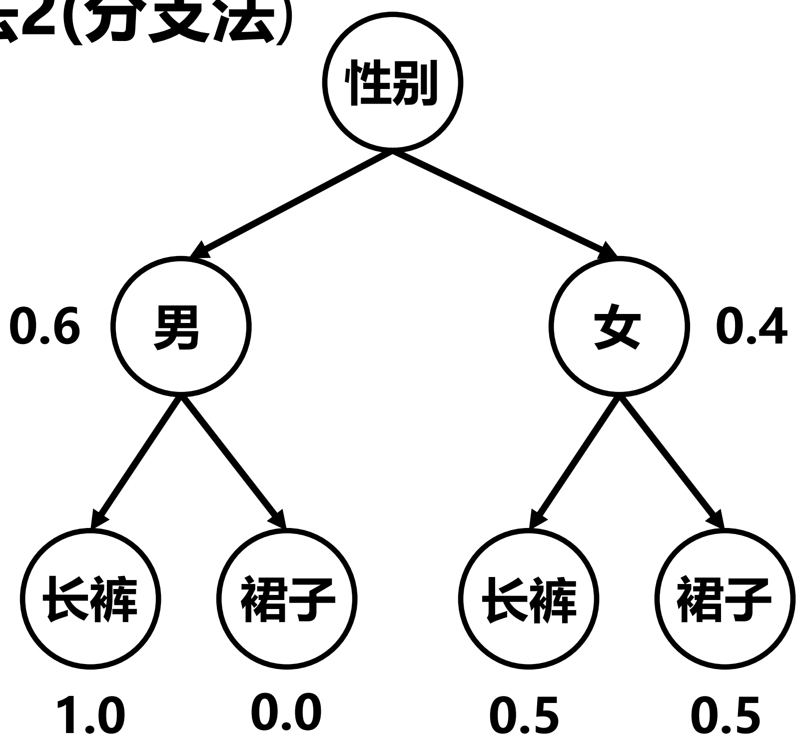
= 1 - P(A=男|B=长裤) = 0.25

5. 贝叶斯理论

先验概率与后验概率：

例：一所学校里面有 60% 的男生，40% 的女生。男生总是穿长裤，女生则一半穿长裤一半穿裙子。在校园中从背影看到一个穿长裤的学生，求该学生是男生的概率（源自维基百科）

解法2(分支法)



P(长裤)=0.6*1.0+0.4*0.5=0.8,
P(B=长裤|A=男) =1.0,
P(A=男)=0.6

P(A=男|B=长裤)
= 0.6*1/0.8 =0.75

5. 贝叶斯理论

贝叶斯推理：

$$P(\text{cause}|\text{effect}) = \frac{P(\text{effect}|\text{cause})P(\text{cause})}{P(\text{effect})}$$

疾病　症状

例：M：脑膜炎， S：头痛，

已知： $P(+m) = 0.0001 \quad P(+s|+m) = 0.8 \quad P(+s|-m) = 0.01$

$$P(+m|+s) = \frac{P(+s|+m)P(+m)}{P(+s)}$$

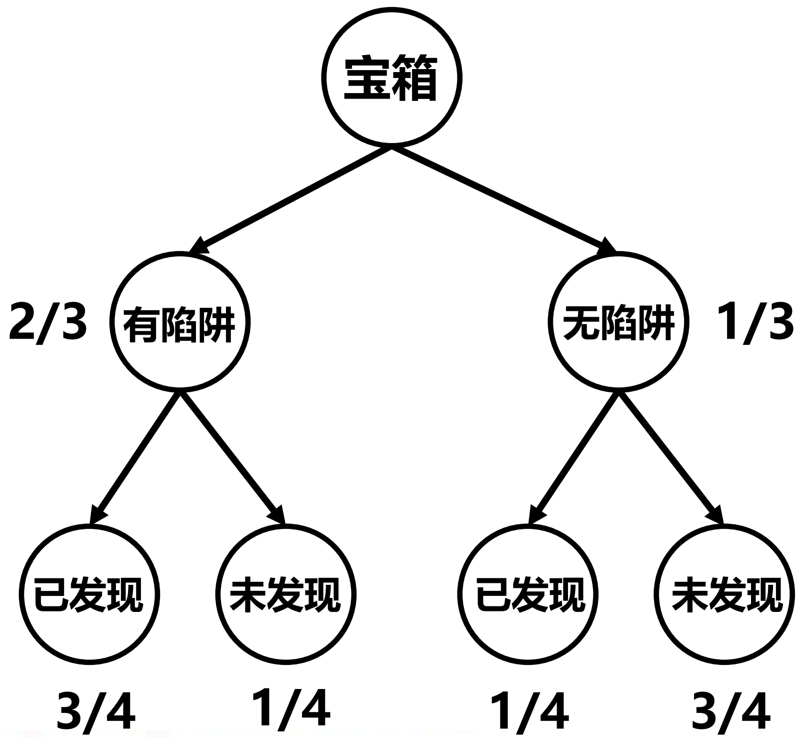
$$= \frac{P(+s|+m)P(+m)}{P(+s|+m)P(+m) + P(+s|-m)P(-m)}$$

$$= \frac{0.8 \times 0.0001}{0.8 \times 0.0001 + 0.01 \times 0.999}$$

5. 贝叶斯理论

贝叶斯推理：

例：在某个角色扮演游戏中，玩家打倒怪物能获得宝箱，但宝箱有2/3的概率有陷阱（掉血）。玩家可以施放魔法来检测陷阱，但有1/4的概率魔法失效。假设玩家打倒怪物获得宝箱，并施放魔法，求玩家打开宝箱中了陷阱的概率。



P(有陷阱|未发现)=?

P(未发现) = 2/3*1/4+1/3*3/4

= 5/12

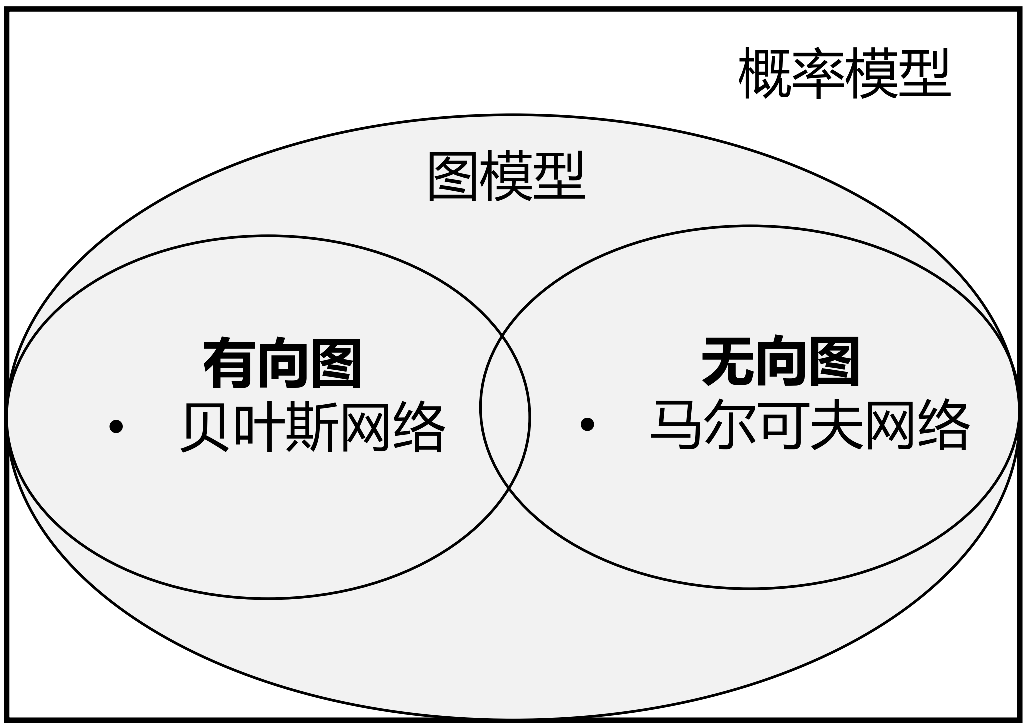
P(未发现|有陷阱)=1/4

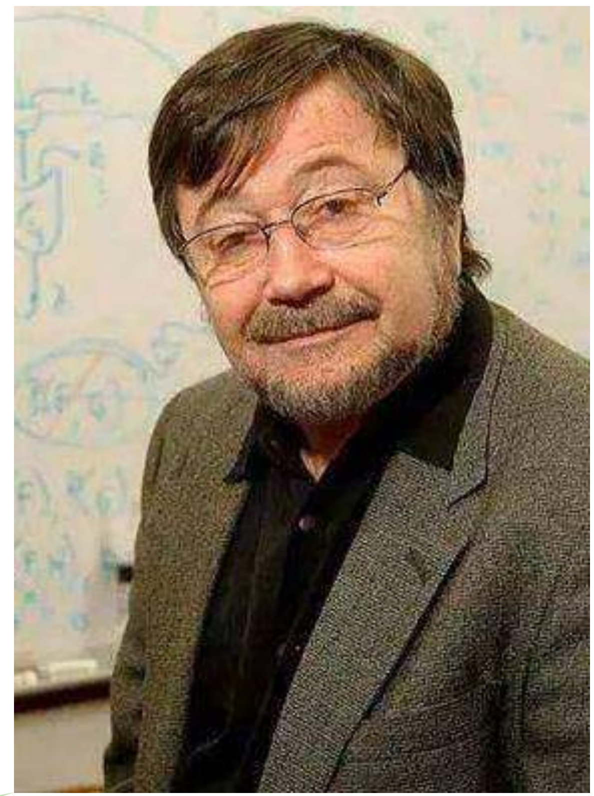
P(有陷阱)=2/3

P(有陷阱|未发现)=2/5

5. 贝叶斯理论

贝叶斯网络(Bayesian network)，又称信念网络(Belief Network)，或有向无环图模型(directed acyclic graphical model)，是一种常用的概率图模型之一





Hi, Bayes Again

贝叶斯网络之父 Judea Pearl
2011年图灵奖获得者

5. 贝叶斯理论

贝叶斯网络是为了处理人工智能中**不确定性问题**而发展起来的一种不确定性推理方法

贝叶斯网络是一种系统地描述**随机变量之间关系**的工具，有可效地进行**概率推理**(Probabilistic inference)

理论上，概率推理只需要获得**随机变量的联合概率分布**，但变量众多时，联合概率分布复杂度成**指数级**增长

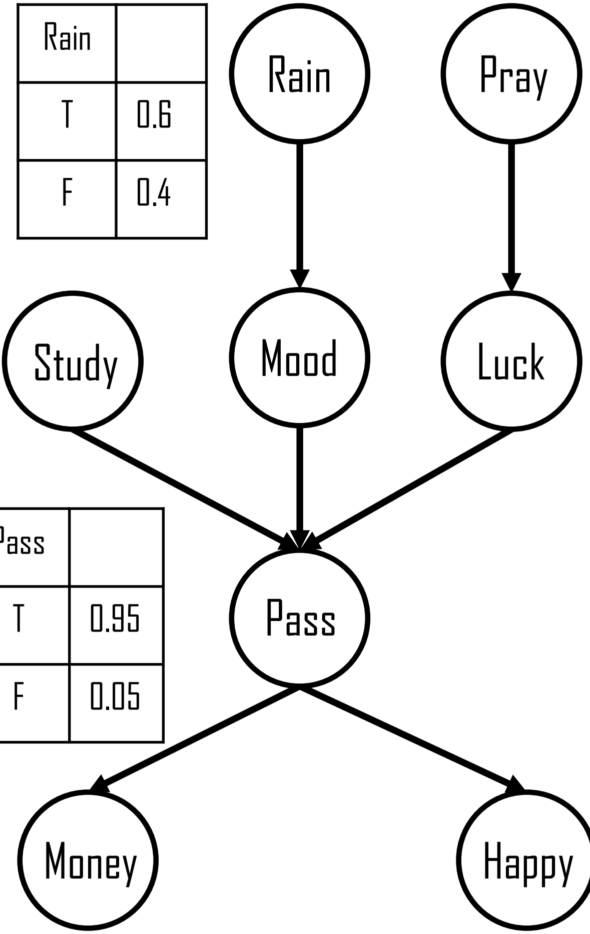
贝叶斯网络将复杂的联合概率分布**分解成一系列相对简单的模块**，降低概率推理的复杂度，从而应用于大型问题

应用领域：医疗诊断、智能决策、数据挖掘、文本分析

贝叶斯网络的优点：**可视化，可解释，精确性**

5. 贝叶斯理论

P(Rain, Pray, Study, Mood, Luck, Pass, Money, Happy)



Rain	
T	0.6
F	0.4

Pray	
T	1.0
F	0

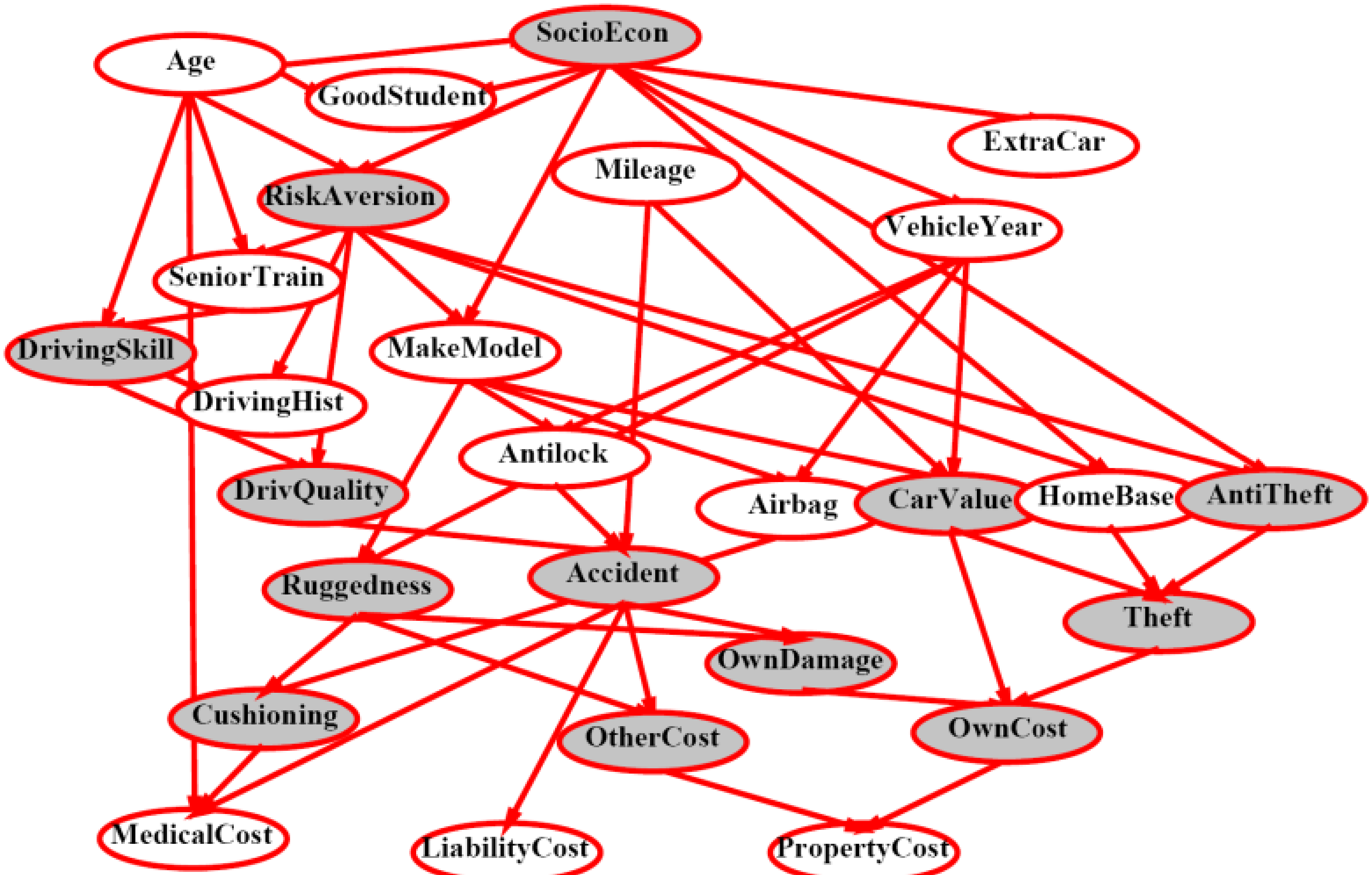
Study	
T	0.65
F	0.35

Luck	
T	1.0
F	0

Pass	
T	0.95
F	0.05

Bayesian Network for Fun

5. 贝叶斯理论



Bayesian Network-Insurance

5. 贝叶斯理论

理论基础：

链式规则

如何推导?

$$
\begin{aligned}
P(X_1, X_2, \ldots X_n) &= P(X_1)P(X_2|X_1)P(X_3|X_1, X_2)\ldots \\
&= \prod_{i=1}^{n} P(X_i|X_1, \ldots, X_{i-1})
\end{aligned}
$$

贝叶斯公式

$$P(A|B) = \frac{P(B|A)}{P(B)}P(A)$$

条件独立

$$P(x_i|x_1 \cdots x_{i-1}) = P(x_i|parents(X_i))$$

5. 贝叶斯理论

两变量独立：

x：天气，y：膝盖痛

x：天气，y：牙痛

x：牙痛，y：膝盖痛

数学描述： $\forall x, y : P(x, y) = P(x)P(y)$

联合分布是两个变量分布的简单乘积

另一种形式 $\forall x, y : P(x|y) = P(x)$

变量独立： 常见假设

经验联合分布：近似独立

贝叶斯分类器：变量独立

5. 贝叶斯理论

两变量独立：

$P(T)$

T	P
hot	0.5
cold	0.5

$P_1(T,W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

$P_2(T,W)$

T	W	P
hot	sun	0.3
hot	rain	0.2
cold	sun	0.3
cold	rain	0.2

$P(W)$

W	P
sun	0.6
rain	0.4

5. 贝叶斯理论

条件独立：

x：熬夜，y：迟到，z：赖床

x：塞车，y：打伞，z：下雨

数学描述： $\forall x, y, z : P(x, y|z) = P(x|z)P(y|z)$

给定变量z，变量x条件独立于y

也即当z发生时，x发生与否与y发生与否是无关的

另一种形式 $\forall x, y, z : P(x|z, y) = P(x|z)$

$$P(\text{Traffic}, \text{Rain}, \text{Umbrella}) =$$
$$P(\text{Rain})P(\text{Traffic}|\text{Rain})P(\text{Umbrella}|\text{Rain}, \text{Traffic})$$
$$P(\text{Rain})P(\text{Traffic}|\text{Rain})P(\text{Umbrella}|\text{Rain})$$

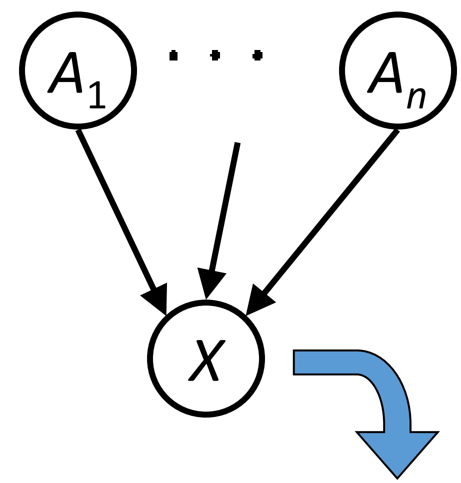
5. 贝叶斯理论

贝叶斯网络 = 图拓朴结构 + 局部条件概率

结点：变量（有限取值）

边：变量之间的直接影响（概率分布）

- 一组结点V，一个节点表示一个变量X
- 有向、无环图BN
- 每个结点都有条件概率分布



联合概率分布⇒条件概率分布

$$P(x_1, x_2, \ldots x_n) = \prod_{i=1}^{n} P(x_i | parents(X_i))$$

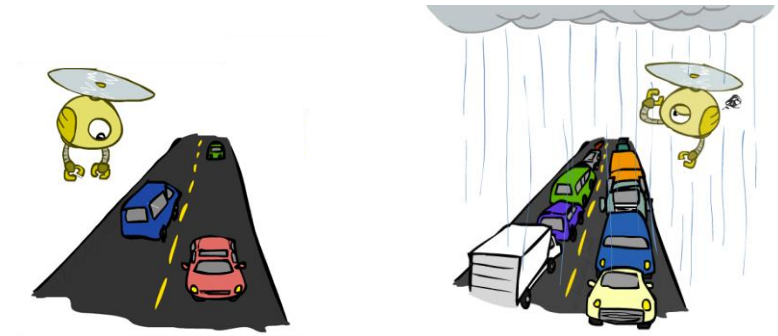
$$P(X | A_1 \ldots A_n)$$

5. 贝叶斯理论

两变量之间关系

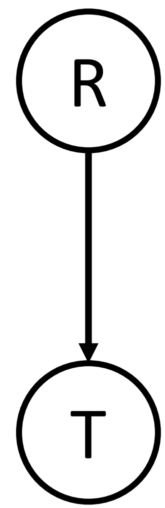
R: 下雨

T: 交通事故



网络1： 独立

网络2： 下雨会引起交通事故



$P(R)$

+r	1/4
-r	3/4

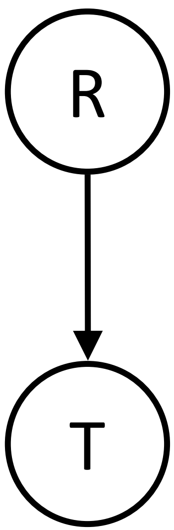
$P(T|R)$

+r	+t	3/4
	-t	1/4
-r	+t	1/2
	-t	1/2

5. 贝叶斯理论

两变量之间关系

$$P(y)P(x|y) = P(x,y)$$



$P(R)$

+r	1/4
-r	3/4

$P(T|R)$

+r	+t	3/4
	-t	1/4
-r	+t	1/2
	-t	1/2

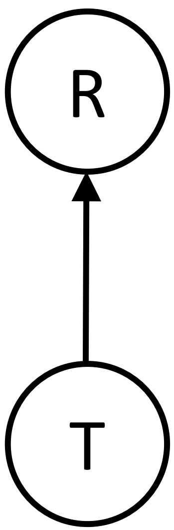
$P(T,R)$

+r	+t	3/16
+r	-t	1/16
-r	+t	6/16
-r	-t	6/16

5. 贝叶斯理论

两变量之间关系

$$P(x|y) = \frac{P(x,y)}{P(y)}$$



$P(T)$

+t	9/16
-t	7/16

$P(R|T)$

+t	+r	1/3
	-r	2/3
-t	+r	1/7
	-r	6/7

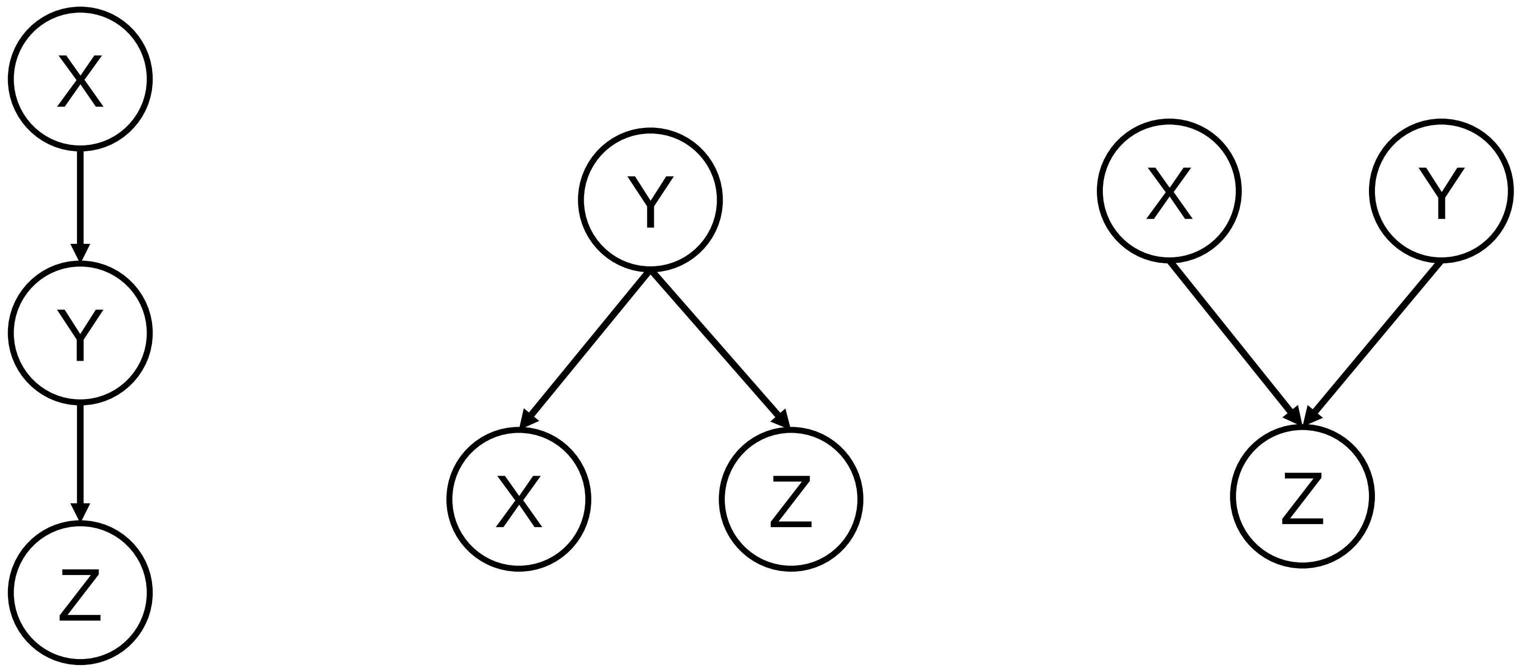
$P(T,R)$

+r	+t	3/16
+r	-t	1/16
-r	+t	6/16
-r	-t	6/16

逆关系-贝叶斯

5. 贝叶斯理论

三变量之间关系（图分割）



a) 顺连　　b) 分连　　c) 汇连

5. 贝叶斯理论

三变量之间关系

X: 低气压　　Y: 下雨　　Z: 塞车

a) 顺连(因果链)

X与Z是否独立?
- 低压导致下雨导致塞车
- 高压无雨不塞车

概率计算
- $P(+y \mid +x) = 1$,
- $P(-y \mid -x) = 1$,
- $P(+z \mid +y) = 1$,
- $P(-z \mid -y) = 1$

5. 贝叶斯理论

三变量之间关系

X: 低气压　　Y: 下雨　　Z: 塞车

a) 顺连(因果链)

给定Y，X与Z是否独立?

$$P(z|x,y) = \frac{P(x,y,z)}{P(x,y)}$$

$$= \frac{P(x)P(y|x)P(z|y)}{P(x)P(y|x)}$$

$$= P(z|y)$$

$$P(x,y,z) = P(x)P(y|x)P(z|y)$$

证据阻塞了变量之间的影响

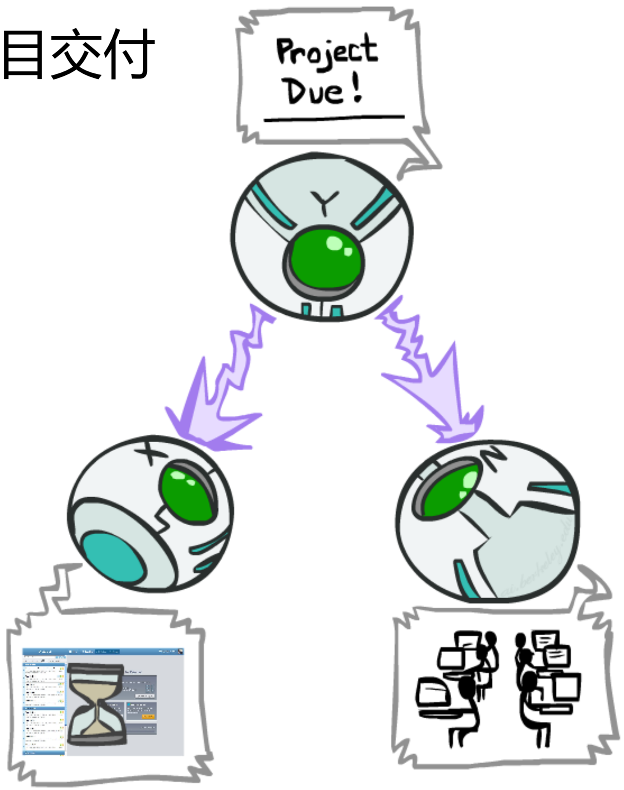
5. 贝叶斯理论

三变量之间关系

Y: 项目交付

X: 讨论群

Z: 实验室



b) 分连(共同原因)

X与Z是否独立?

- 项目导致讨论热烈
- 和实验室爆满

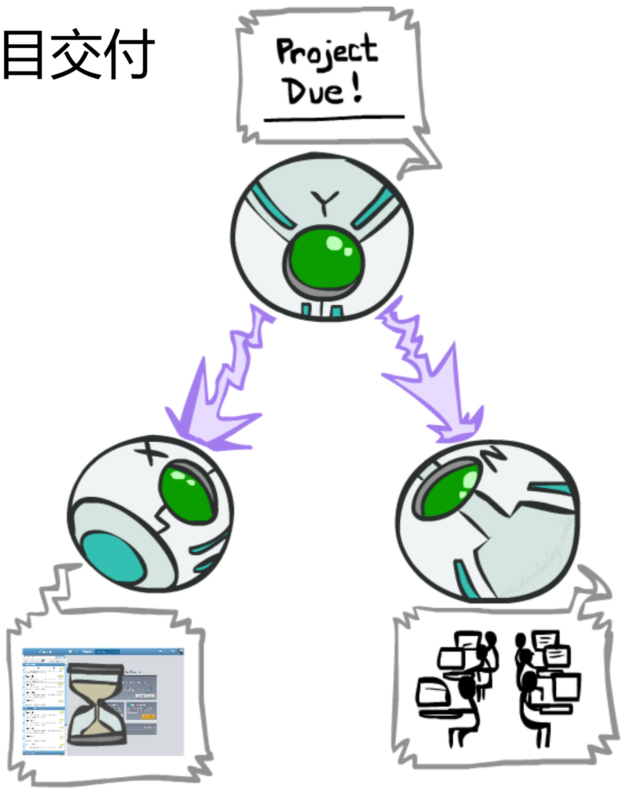
概率计算

- $P(+x \mid +y) = 1$,
- $P(-x \mid -y) = 1$,
- $P(+z \mid +y) = 1$,
- $P(-z \mid -y) = 1$

5. 贝叶斯理论

三变量之间关系

Y: 项目交付



X: 讨论群

Z: 实验室

给定Y，X与Z是否独立?

$$P(z|x,y) = \frac{P(x,y,z)}{P(x,y)}$$

$$= \frac{P(y)P(x|y)P(z|y)}{P(y)P(x|y)}$$

$$= P(z|y)$$

b) 分连(共同原因) $P(x,y,z) = P(y)P(x|y)P(z|y)$

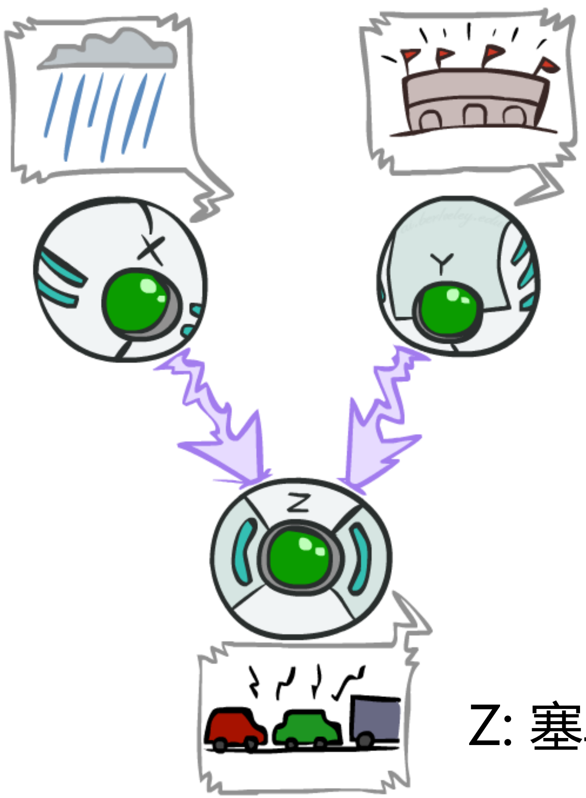
观测到的原因阻塞了变量之间的影响

5. 贝叶斯理论

三变量之间关系

X: 下雨

Y: 球赛



Z: 塞车

c) 汇连(多因一果)

X与Y是否独立?

- 下雨和球赛引起塞车
- 但下雨和球赛不相关

给定Z，X与Y是否独立?

- 看到塞车，就会讨论下雨和球赛哪个是造成结果的主要原因

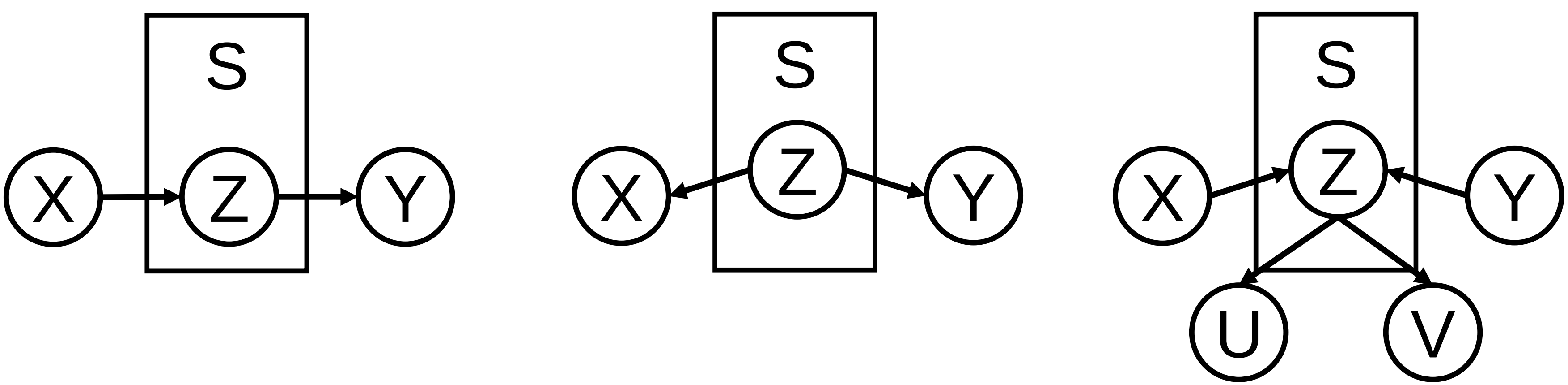
观测到的结果激活原因变量之间的影响

5. 贝叶斯理论

多变量之间关系-阻塞

设S为一节点集合，X和Y是不在S中的两个节点，考虑X和Y之间的一条通路，如果满足下面条件之一，则称X和Y被S所**阻塞**

- 有一个在S中的顺连节点;
- 有一个在S中的分连节点;
- 有一个汇连节点Z，它和它的后代节点均不在S中



a) 顺连　　**b) 分连**　　**c) 汇连**

5. 贝叶斯理论

图分割与变量独立

如果X和Y之间的所有通路都被S阻塞，则说有向分割(Directed separate)X和Y，简称**d-separate**，**d-分割**

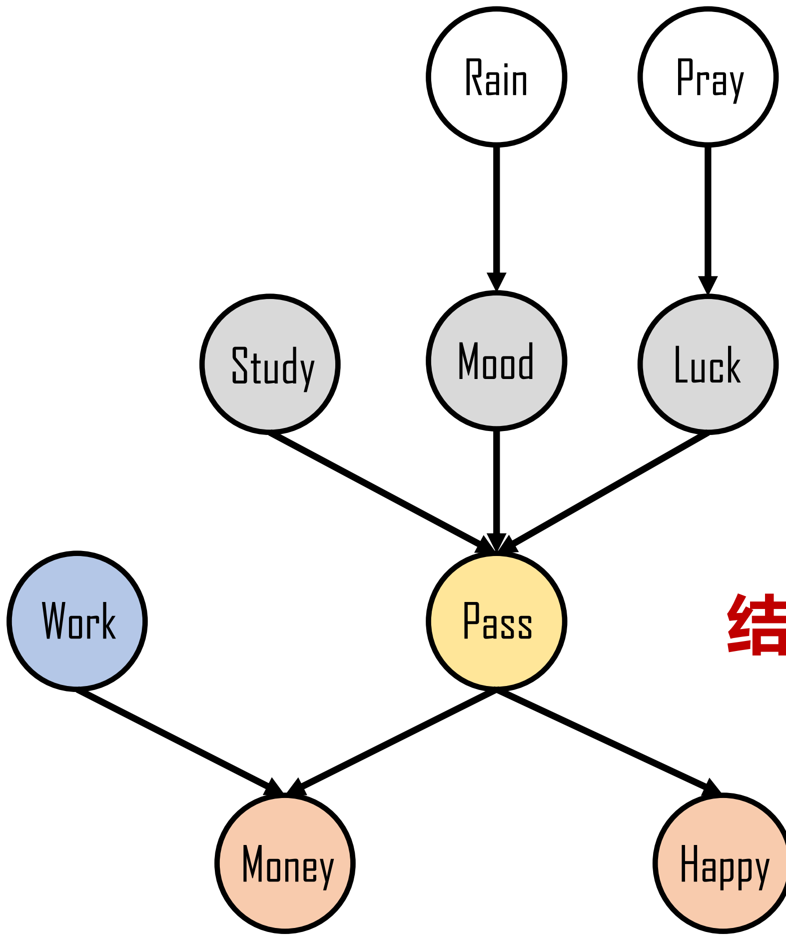
如果S能d-分割X和Y，则X和Y在给定S时条件独立

定理**（整体马尔可夫性）**：设X和Y为贝叶斯网BN中的两个变量结点，S为BN中不包含X和Y的节点集合，如果S d-分割X和Y，那么X和Y在给定Z时条件独立

d-分割是图论的概念，而条件独立是概率论的内容，定理揭示了贝叶斯网络中图论和概率之间的关系

定义**（马尔可夫覆盖）**：给定一个结点X，其在贝叶斯网络中的马尔可夫覆盖（Markov Blanket）包括其**父节点**、**子节点**以及**子节点的父节点**

5. 贝叶斯理论

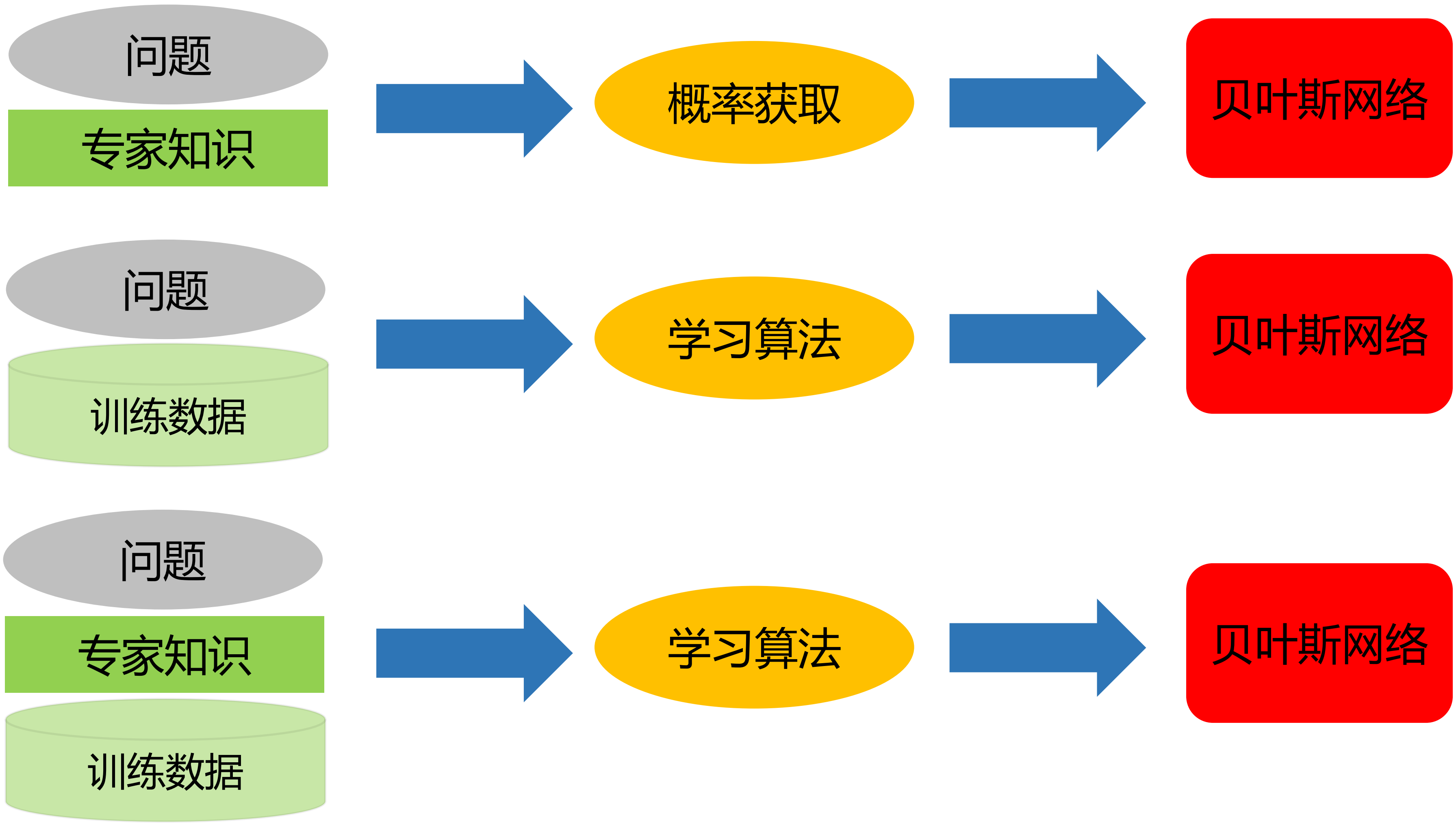


结点Pass的马尔可夫覆盖

Bayesian Network for Fun

5. 贝叶斯理论

贝叶斯网构建



5. 贝叶斯理论

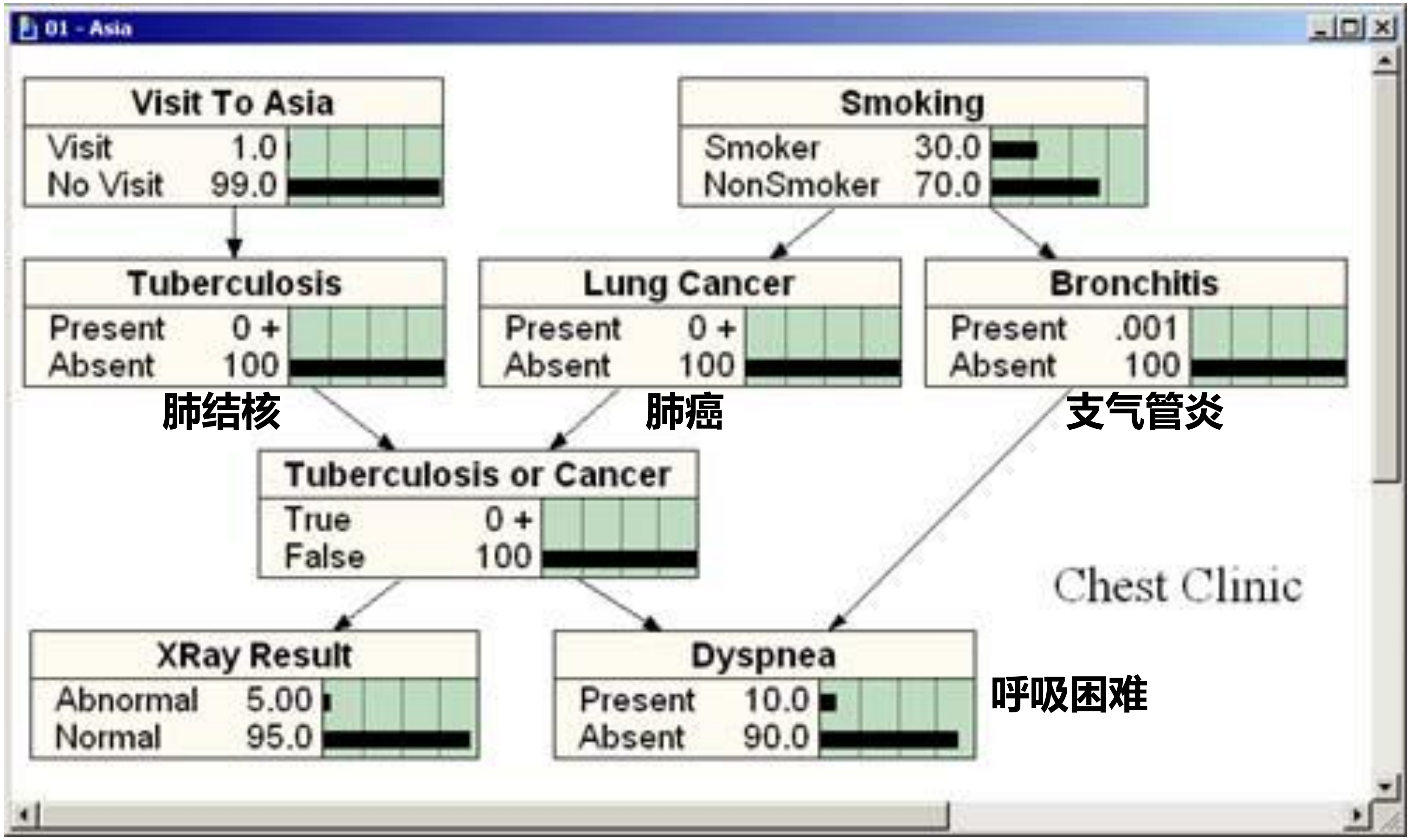
贝叶斯网络推理实例-肺部疾病诊断(源自www.norsys.com)

假想你是一名新毕业的医生，专攻肺部疾病。你决定建立一个胸部疾病诊所，主治肺病及相关疾病。课本知识中已经告诉你了肺癌、肺结核和支气管炎的发生比率以及这些疾病典型的临床症状、病因等，于是你就可以根据课本里的理论知识建立自己的Bayes网。如根据如下数据信息：

吸烟的人有30%.
每10万人中就就有70人患有肺癌
每10万人中就就有10人患有肺结核
每10万人中就就有800人患有支气管炎
10%人存在呼吸困难症状, 大部分人是哮喘、支气管炎和其他非肺结核、非肺癌性疾病引起

5. 贝叶斯理论

实例-肺部疾病诊断



贝叶斯网络（先验知识）

5. 贝叶斯理论

实例-肺部疾病诊断

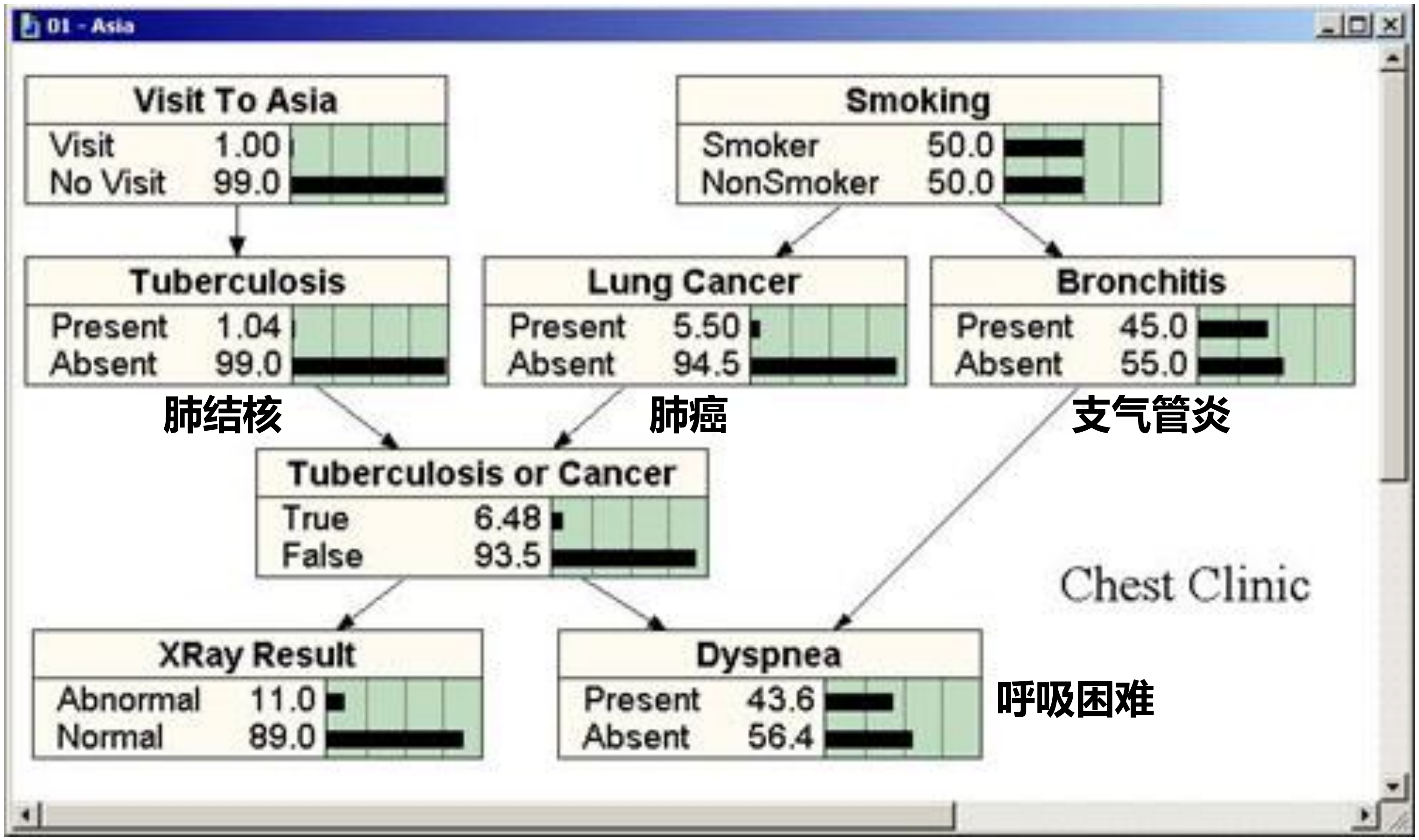
这样的一个BN模型对你意义不大，因为它没有用到来你诊所病人的案例数据，**不能反映真实病人**的情况。

当诊所诊治了数千病人后，会发现课本中所描述的情况与实际诊所数据显示的情况是完全不同的，**实际诊所数据**显示：

- 50%的病人吸烟
- 1%患有肺结核
- 5.5% 得了肺癌
- 45% 患有不同程度支气管炎

5. 贝叶斯理论

实例-肺部疾病诊断



贝叶斯网络（数据信息）

5. 贝叶斯理论

实例-肺部疾病诊断

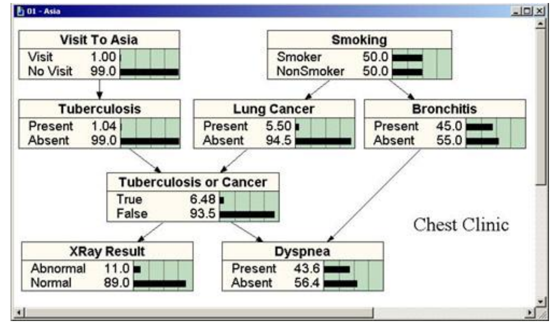
如何在日常诊断中用该BN模型?

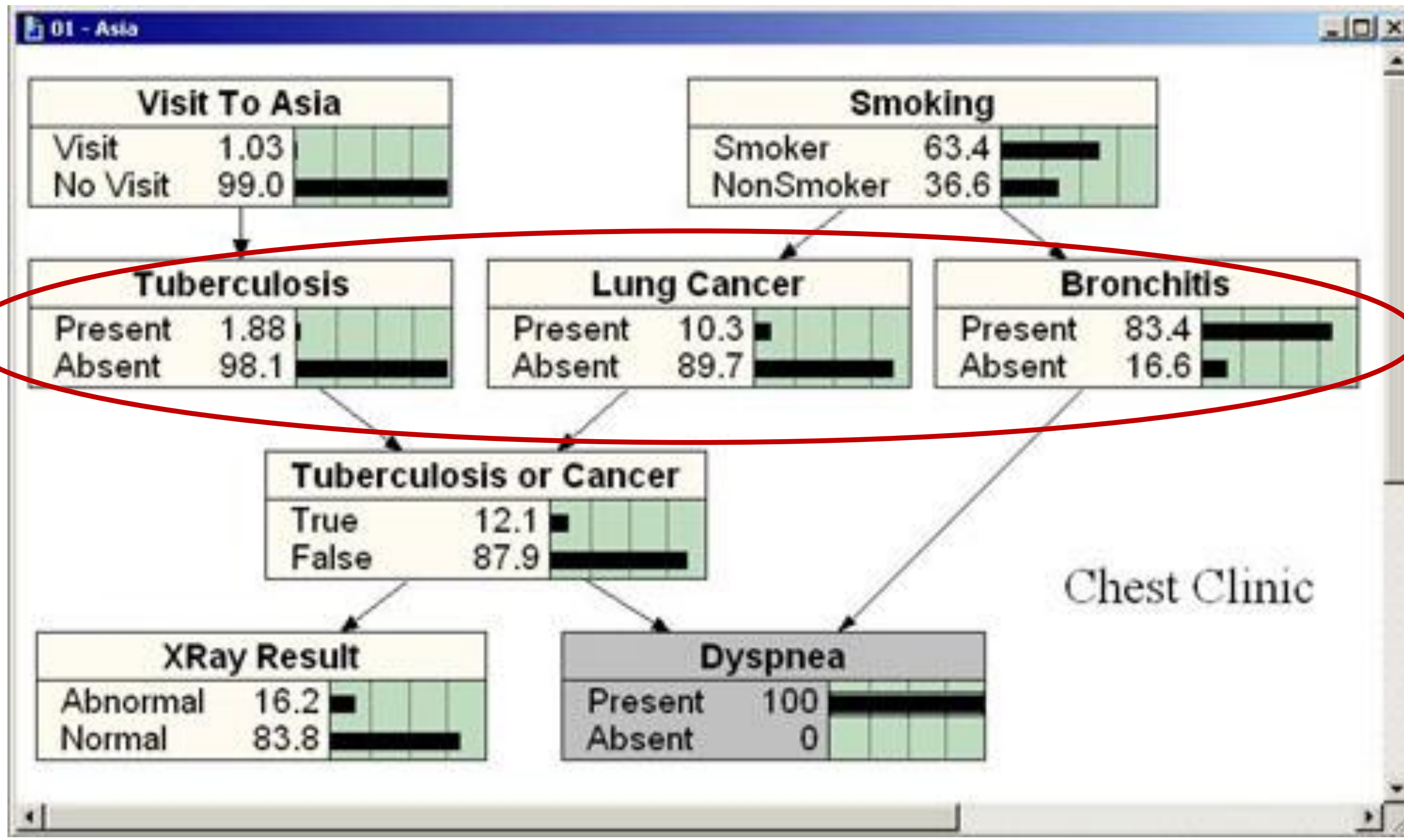
新患者（无信息）⇒咨询患者⇒ BN网络调整⇒推断结果

例：一个女病人进入诊所，她告诉我们她呼吸困难。将这个信息输入到网络，我们相信病人的信息，认为其存在100%呼吸困难。

5. 贝叶斯理论

实例-肺部疾病诊断

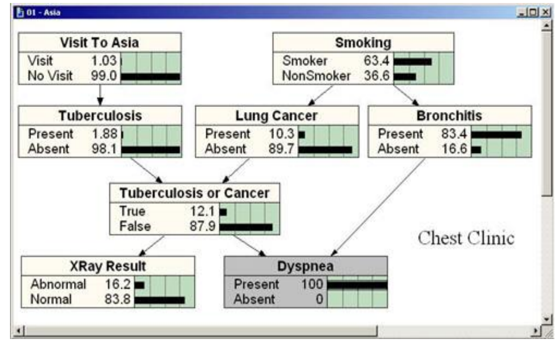


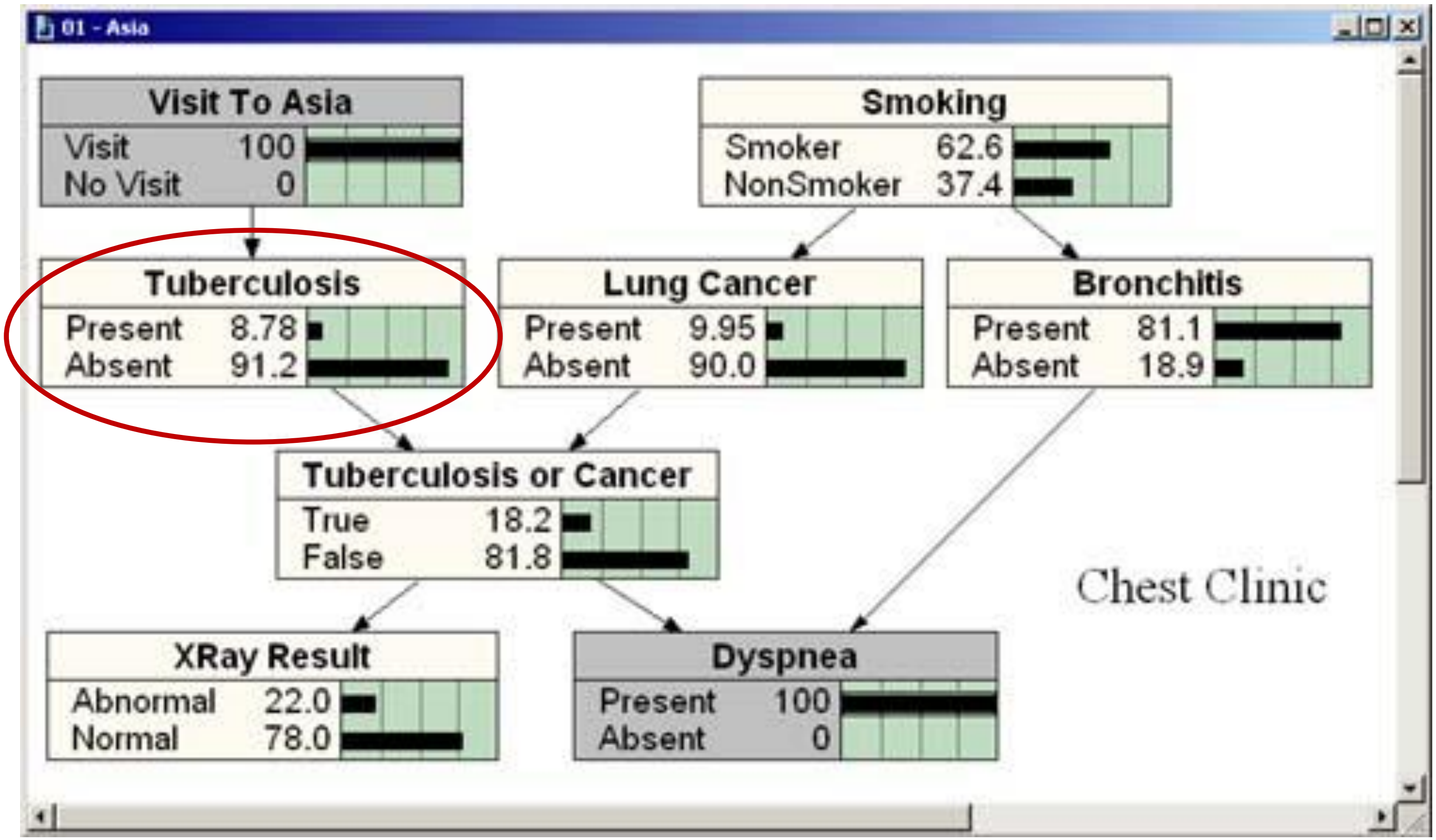


贝叶斯网络（100%呼吸困难）

5. 贝叶斯理论

实例-肺部疾病诊断

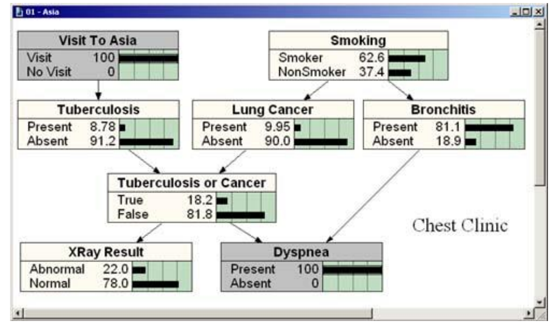


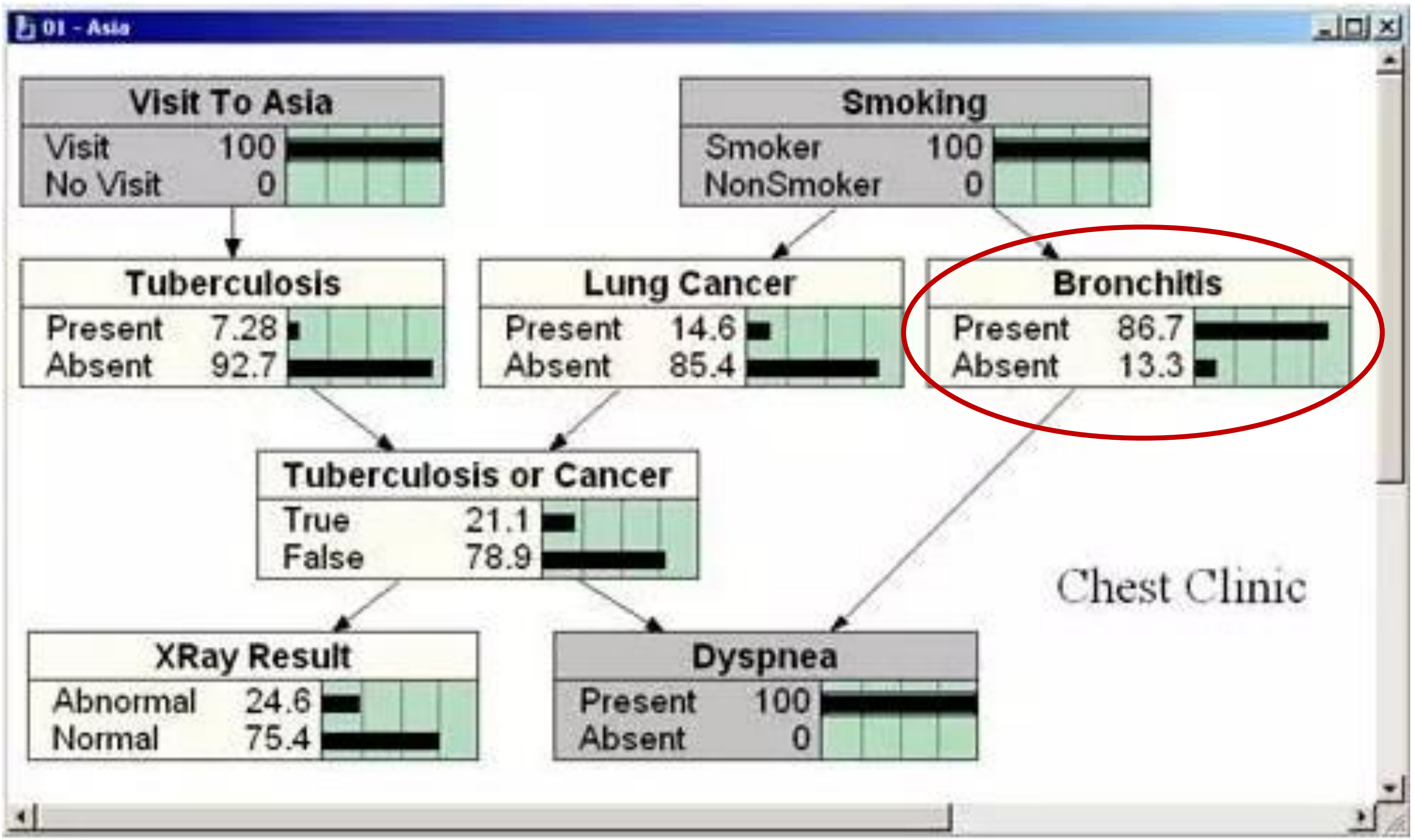


贝叶斯网络（去过亚洲）

5. 贝叶斯理论

实例-肺部疾病诊断

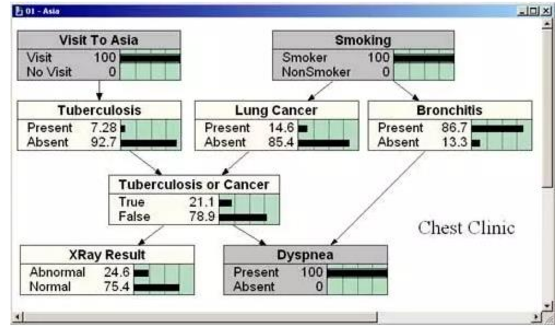


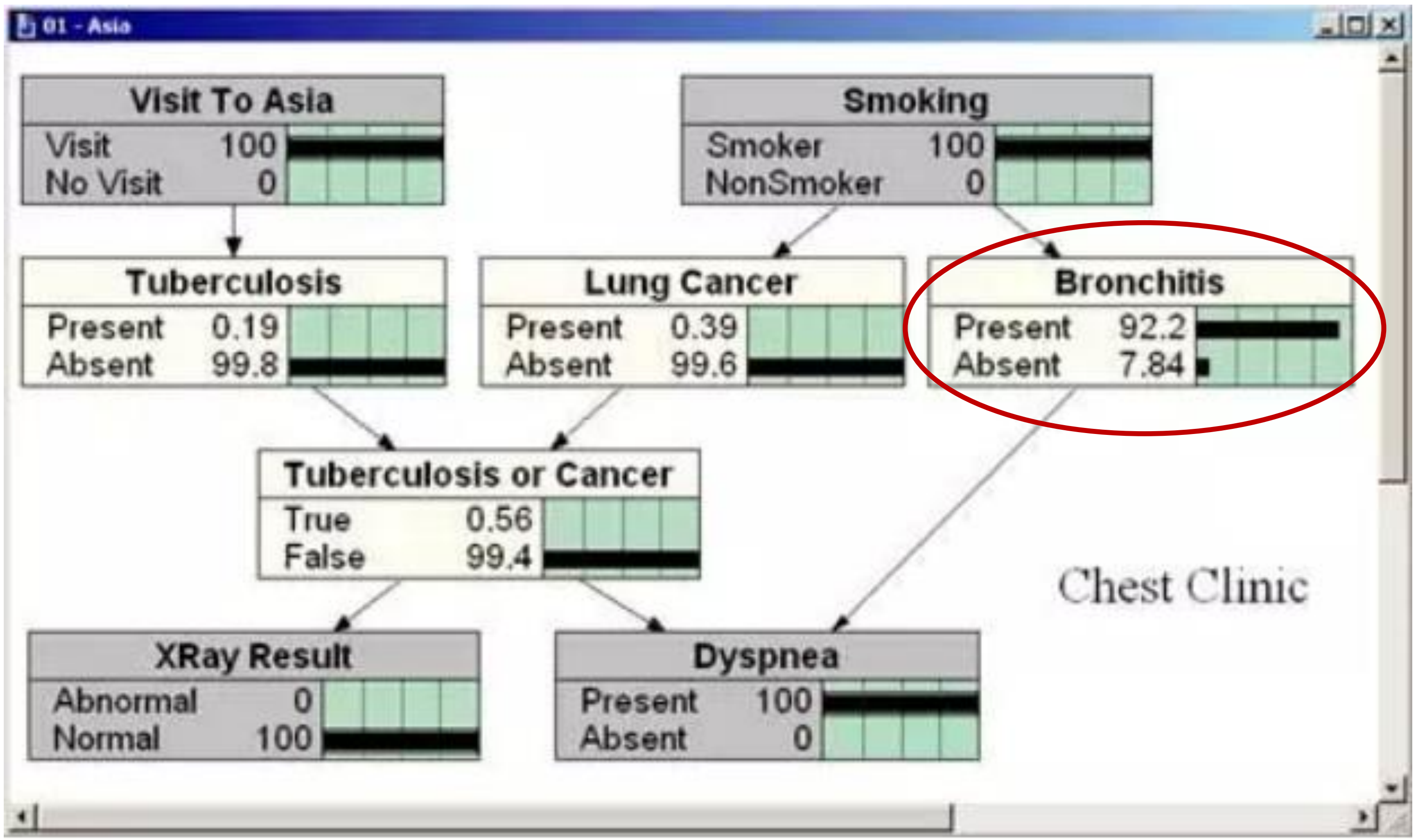


贝叶斯网络（吸烟）

5. 贝叶斯理论

实例-肺部疾病诊断

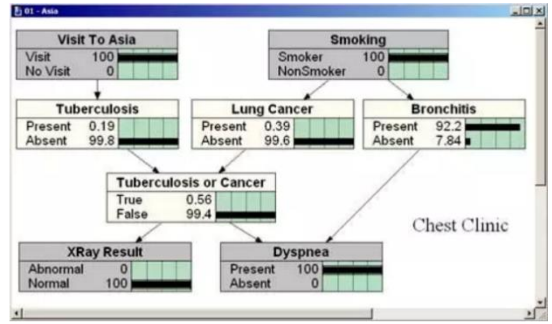


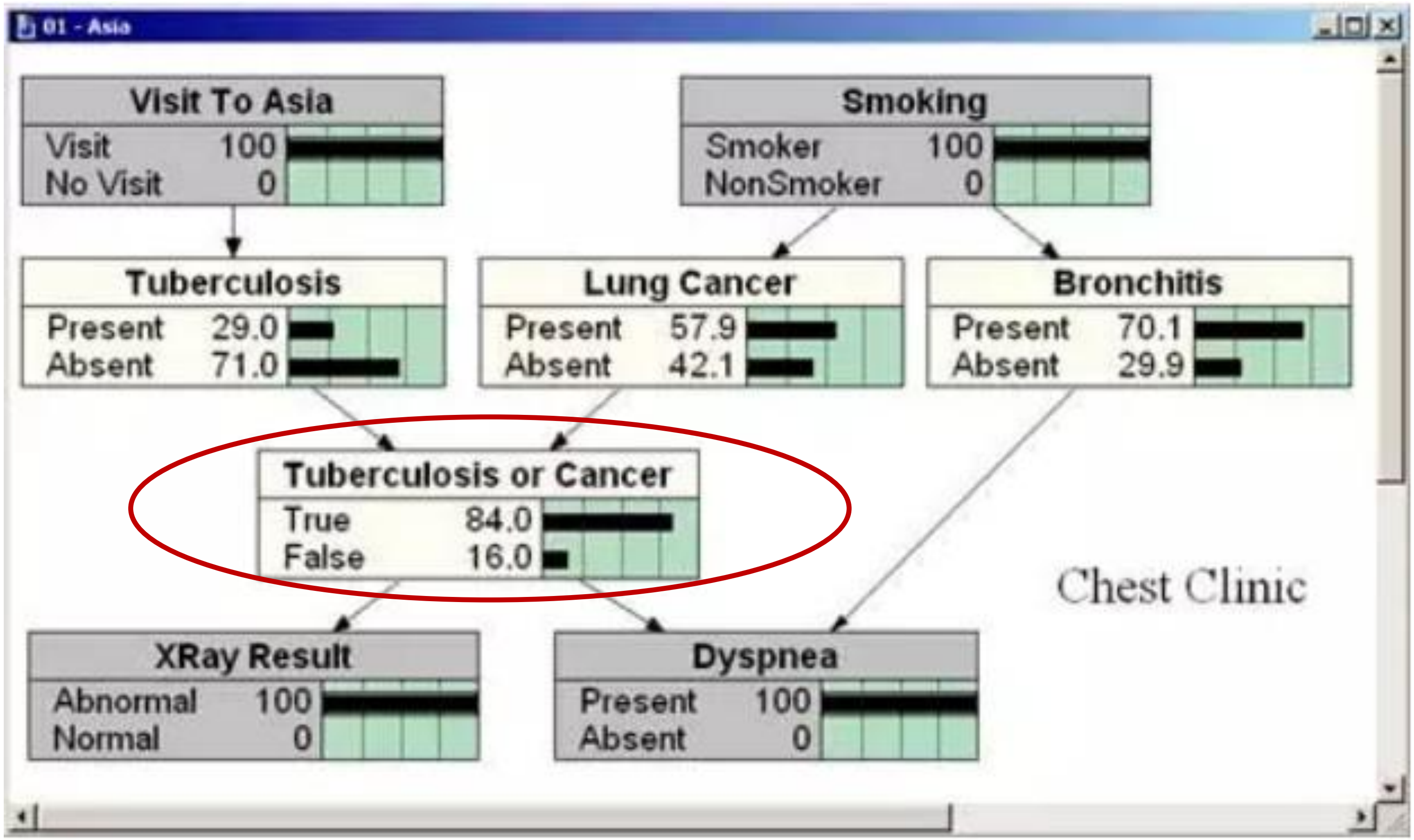


贝叶斯网络（X光正常）

5. 贝叶斯理论

实例-肺部疾病诊断





贝叶斯网络（X光不正常）

5. 贝叶斯理论

实例-肺部疾病诊断

贝叶斯网络是一个用严格的**数学方法**来模拟一个世界的方法

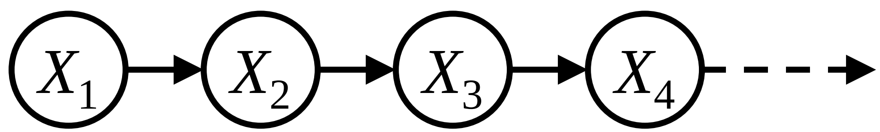
贝叶斯网络非常灵活，同时也是**计算效率非常高**的方法。

贝叶斯网络最强大之处在于从每个阶段结果获得新信息时，节点间的概率会自动调整，是个不断**自我学习**自我调整的过程

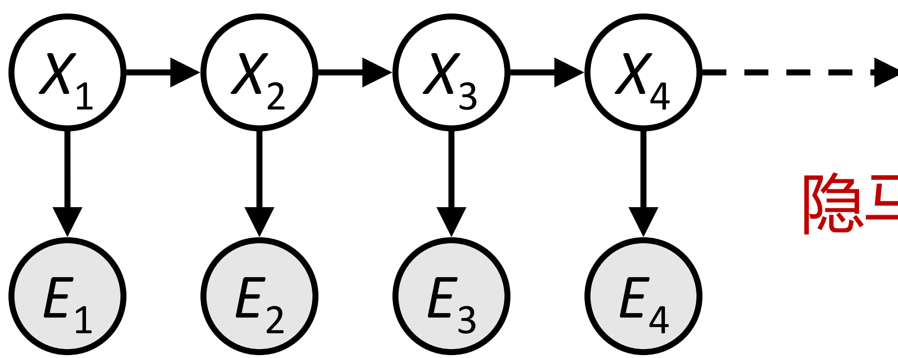
贝叶斯网络推理时只要个体信息足够完善足够完整，得到的诊断**结果就越精准**。

5. 贝叶斯理论

扩展内容-时间上的概率推理



$P(X_1)$ $P(X_t|X_{t-1})$ 马尔可夫模型



隐马尔可夫模型

6. 练习题

找一个逻辑推理题（或怪物世界的问题），尝试将该问题进行命题逻辑表示，并考虑如何进行推理。

6. 练习题

已知某种疾病的发病率是0.001，即1000人中会有1个人得病。现有一种试剂可以检验患者是否得病，它的准确率是0.99，即在患者确实得病的情况下，它有99%的可能呈现阳性。它的误报率是5%，即在患者没有得病的情况下，它有5%的可能呈现阳性。现有一个病人的检验结果为阳性，请问他确实得病的可能性有多大？如果误报率从5%降为1%，请问病人得病的概率会变成多少？

谢谢聆听 欢迎提问