不确定性与不确切性

不确定：不能确定真伪

不确切：描述模糊，不够严格精确

确定－确切性信息：现在在下50ml以上的雨

不确定－确切性信息：明天可能下50ml以上的雨

确定－不确切性信息：现在在下大雨

不确定－不确切性信息：明天可能下大雨

不确定性知识

设 $C(S)$ 为命题 S 的信度，二元组 $(S, C(S))$ 就可作为不确定性命题的一种表示形式，将不确定性产生式规则 $A \to B$ 表示为 $(A \to B, C(A \to B))$ 或 $A \to (B, C(B|A))$，其中 $C(B|A)$ 为在 A 为真时 B 为真的信度，如

若头痛且发烧，则患了感冒 (0.8)

↓

头痛且发烧 ⟶ (感冒, 0.8)

不确定性推理＝符号推演＋信度计算

确定性理论

在专家系统基础上引入置信因子（可信度）$CF(H)$，取值范围为 $[-1, 1]$，如：If A then B (0.7)

1、不确定性度量

采用 CF (certainty factor，可信度) 作为不确定性度量

对于规则 (H, E) 来说

$$CF(H,E) = \begin{cases} \dfrac{P(H|E) - P(H)}{1 - P(H)} & P(H|E) > P(H) \\ 0 & P(H|E) = P(H) \\ \dfrac{P(H|E) - P(H)}{\ldots} & P(H|E) < P(H) \end{cases}$$

$P(H)$ $\quad P(H|E) < P(H)$

E是前提 H是结论 $P(H)$为H先验概率

2. 前提事实总CF值计算

$$CF(E_1 \wedge E_2 \wedge \cdots \wedge E_n) = \min\{CF(E_1), CF(E_2), \cdots, CF(E_n)\}$$

$$CF(E_1 \vee E_2 \vee \cdots \vee E_n) = \max\{CF(E_1), CF(E_2), \cdots, CF(E_n)\}$$

其中E_i是与规则前提各条件匹配的事实

3. 推理结论CF值计算

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

E是前提对应的事实 $CF(H, E)$是规则中结论可信度

即结论可不可信，与规则本身靠谱与否有关，也与前提的可信度有关

4、重复结论的CF值计算

若同一结论被两条规则推出，得两个可信度$CF(H_1)$和$CF(H_2)$

则最终$CF(H)$为

$$CF(H) = \begin{cases} CF(H_1) + CF(H_2) - CF(H_1)CF(H_2) & CF(H_1) \geq 0 \text{ 且 } CF(H_2) \geq 0 \\ CF(H_1) + CF(H_2) + CF(H_1)CF(H_2) & CF(H_1) < 0 \text{ 且 } CF(H_2) < 0 \\ CF(H_1) + CF(H_2) & \text{else} \end{cases}$$

例、规则：① If A then B (0.9)

② If B and C then D (0.8)

③ If A and C then D (0.7)

④ If B or D then E (0.6)

事实：A　$CF(A)=0.8$

　　　C　$CF(C)=0.9$

求：$CF(E)$

由①：$CF(B)=0.72$

由②：$CF(D)_1 = 0.8 \times \min\{0.72, 0.9\}$

$= 0.576$

由③：$CF(D)_2 = 0.7 \times \min\{0.8, 0.9\}$

$= 0.56$

$\therefore CF(D) = 0.576 + 0.56 - 0.576 \times 0.56$

$= 0.81344$

由④：$CF(E) = 0.6 \times \max\{0.72, 0.81344\}$

$= 0.488064$

不确定性知识的表示

程度化元组：(<对象>,<属性>,(<属性值>,<程度>))

如：张三比较胖

(张三，体型，(胖，0.9))

程度化谓词　P_d(<对象$_1$，对象$_2$，对象$_3$，…，对象$_n$>)

或 dP(<对象$_1$，对象$_2$，对象$_3$，…，对象$_n$>)

P为谓词，d为程度

P_d为下标表示法　dP为乘法表示法

如：$friends_{1.15}$(张三，李四)

1.15 friends(张三，李四)
张三和李四是好朋友

程度化规则：即程度化语言值+产生式规则
单条件程度化规则的一般形式为：

(<对象>, <特征>, (<语言值>, <程度>))

↓

(<对象>, <特征>, (<语言值>, <程度>))

或 $(A, d) \rightarrow (B, f(d))$

其中 $d = c_A(x)$ 是规则前件语言值A的程度
函数值 $f(d)$ 是规则后件语言值B的程度
函数 $f(d)$ 是原规则 $A(x) \rightarrow B(y)$ 的伴随程度函数

如：(香蕉，颜色，(黄，0.7)) → (香蕉，生熟，(熟，0.9))
香蕉只要有些黄就比较熟了
$d=0.7$　$f(d)=0.9$ (这里只接给定)

(严重(鼻塞)，d_x) ∧ ((头)很痛，d_y) ∧ (高烧，d_z)
→ (重感冒，$1.2(0.3d_x + 0.2d_y + 0.5d_z)$)
如果严重鼻塞、头很痛且发高烧则患了重感冒

程度化框架　含有程度化语言值的框架

如：框架名：<大枣>
　　类属：(<干果>, 0.8)
　　形状：(<圆>, 0.7)
　　颜色：(<红>, 1.0)

程度化语义网　含有程度化语言值的语义网

如：狗 $\xrightarrow{类属}$ (食肉动物，0.9)

$\downarrow$ 嗅觉

(灵敏，1.5)

基于模糊集合与模糊关系的模糊推理

模糊集合：设 U 是一个论域，U 到区间 $[0,1]$ 的一个映射 $\mu: U \to [0,1]$，就确定了 U 的一个模糊子集 A。映射 μ 称为 A 的隶属函数，记为 $\mu(u)$。对于任意 $u \in U$，$\mu_A(u) \in [0,1]$ 称为 u 属于模糊子集 A 的程度，简称隶属度。如 $U=\{$所有人$\}$，$A=\{$高个子的人$\}$，

模糊集合 A 一般可以表示为

$$A=\{\mu_A(u_1)/u_1, \mu_A(u_2)/u_2, \cdots\}$$

或 $A=\mu_A(u_1)/u_1+\mu_A(u_2)/u_2+\cdots$

或 $A=\int_{u \in U} \mu_A(u)/u$

模糊关系：集合 $U_1, U_2, \cdots, U_n$ 的笛卡尔积集 $U_1 \times U_2 \times \cdots \times U_n$ 的一个模糊子集 R 称为 $U_1, U_2, \cdots, U_n$ 间的一个模糊关系。特别地，U^n 的一个模糊子集称为 U 上的 n 元模糊关系。

笛卡尔积：$U_1 \times U_2=\{(a_i, b_i)\}$ $a_i \in U_1$, $b_i \in U_2$

例、$U=\{1,2,3\}$，则 U 上远大于模糊关系

可表示为：

$R_{>>} = 0.1/(1,2) + 0.4/(1,3) + 0.1/(2,3)$

成矩阵

$$\begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & 0 & 0.1 & 0.4 \\ 2 & 0 & 0 & 0.1 \\ 3 & 0 & 0 & 0 \end{array}$$

模糊集合的运算

$\mu_{A\cap B}(x) = \min(\mu_A(x), \mu_B(x))$

$\mu_{A\cup B}(x) = \max(\mu_A(x), \mu_B(x))$

补集 $\mu_{A'}(x) = 1 - \mu_A(x)$

模糊关系的合成

设 $R_1 = (s_{il})_{n\times k}$　$R_2 = (t_{lj})_{k\times m}$ 则

$R = R_1 \circ R_2 = (r_{ij})_{n\times m}$

$r_{ij} = \bigvee_{l=1}^{k} (s_{il} \wedge t_{lj})$

即对 R_1 第 i 行和 R_2 第 j 列对应元素先做交运算，再对所有结果做并运算（先取小，再取大）

如：$R_1 = \begin{bmatrix} 0.1 & 0.6 & 0.3 \\ 0.4 & 0.7 & 0.9 \\ 0.5 & 0.8 & 1 \end{bmatrix}$　$R_2 = \begin{bmatrix} 0.1 & 0.4 \\ 1 & 0.9 \\ 0.7 & 0.8 \end{bmatrix}$

$R = R_1 \circ R_2 = \begin{bmatrix} 0.6 & 0.6 \\ 0.7 & 0.8 \\ 0.8 & 0.8 \end{bmatrix}$

以 R_{11} 为例　$R_{11} = (0.1 \wedge 0.1) \vee (0.6 \wedge 1) \vee (0.3 \wedge 0.7)$

$= 0.1 \vee 0.6 \vee 0.3 = 0.6$

模糊推理

例：张三，李四和王五三位同学对数学、编程和英语的掌握程度如下 (模糊关系)

	数学	编程	英语
张三	0.1	0.6	0.3
李四	0.4	0.7	0.9
王五	0.5	0.8	1

数学，编程和英语对学术和工程能力的贡献如下（模糊关系）：

	工程	学术
数学	0.1	0.4
编程	1	0.9
英语	0.7	0.8

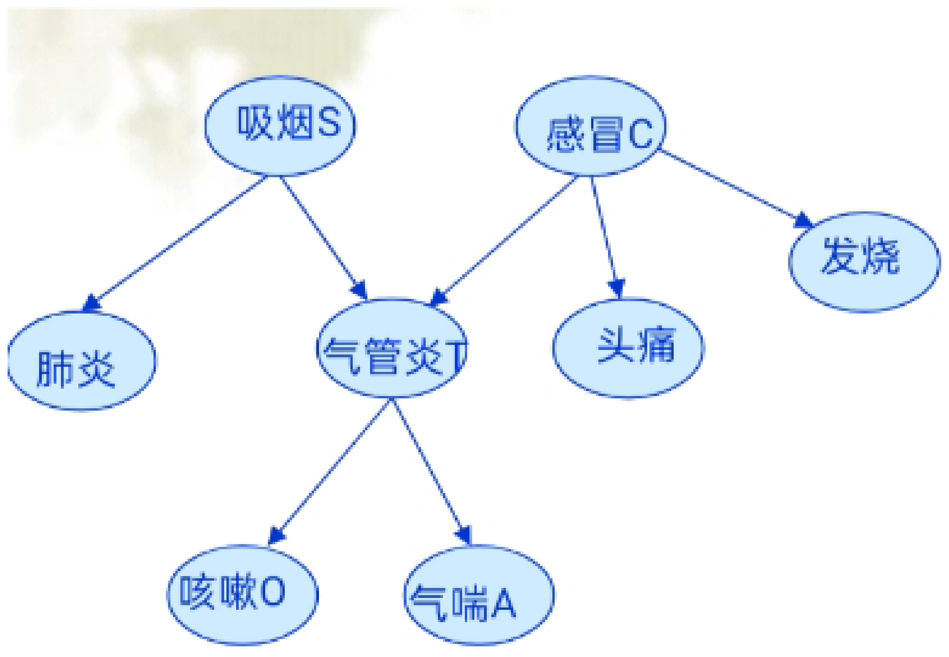
张三、李四和王五谁更适合工程，谁更适合学术？

$R = R_1 \circ R_2 =$

	工程	学术
张三	0.6	0.6
李四	0.7	0.8
王五	0.8	0.8

贝叶斯网络

以随机变量为节点，以条件概率为节点间关系强度的有向无环图。节点表示随机变量，有向边描述了相关变量的关系，而且每个节点附一个条件概率表以刻画相关变量对该变量的影响，条件概率可视为节点之间的关系强度。有向边出发端节点称为因节点，终点端称为果节点



P(s)=0.4 P(¬s)=0.6
P(c)=0.8 P(¬c)=0.2
P(T|S,C)=0.35
P(T|¬S,C)=0.25
P(T|S,¬C)=0.011
P(T|¬S,¬C)=0.002
P(O|T)=0.85
P(O|¬ T)=0.15
P(A|T)=0.5，
P(A| ¬T)=0.1

贝叶斯公式：

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