第3章 不确定性知识系统

现实世界中的大多数问题是不精确、非完备的。因此，人工智能需要研究不确定性的推理方法，以满足客观问题的需求。

3.1 不确定性推理概述

3.1.1 不确定性推理的含义

3.1.2 不确定性推理的基本问题（表示、匹配、更新、合成）

3.1.3 不确定性理的类型

3.2 可信度推理

3.3 主观Bayes推理

3.4 证据理论

3.5 模糊推理

3.6 概率推理

3.1.1 不确定性推理的含义

什么是不确定性推理

不确定性推理是指建立在不确定性知识和证据基础上的推理，泛指除精确推理以外的其它各种推理问题，包括不完备、不精确知识的推理，模糊知识的推理 等。

不确定性推理过程实际上是一种从不确定的初始证据出发，通过运用不确定性知识，最终推出具有一定不确定性但却又是合理或基本合理的结论的思维过程。

为什么要采用不确定性推理

所需知识不完备、不精确

所需知识描述模糊

多种原因导致同一结论

解题方案不唯一

3.1.2 不确定性推理的基本问题

不确定性的表示

(1) 知识的不确定性的表示

考虑因素：问题的描述能力，推理中不确定性的计算

含义：知识的确定性程度，或动态强度

表示：用概率，[0,1]，0接近于假，1接近真

用可信度，[-1,1]，大于0值越大越接近于真

(2) 证据不确定性的表示

证据的类型：按证据组织：基本证据，组合证据

按证据来源：初始证据，中间结论

表示方法：概率，可信度，模糊集等

基本证据：常与知识表示方法一致，如概率，可信度，模糊集等

组合证据：组合方式：析取的关系，合取的关系。

计算方法：基于基本证据

最大/最小方法，概率方法 等。

3.1.2 不确定性推理的基本问题

不确定性的匹配

含义

不确定的前提条件与不确定的事实匹配

问题

前提是不确定的，事实也是不确定的

方法

设计一个计算相似程度的算法，给出相似的限度

标志

相似度落在规定限度内为匹配，否则为不匹配

3.1.2 不确定性推理的基本问题

不确定性的更新　不确定性结论的合成

4. 不确定性的更新

主要问题

① 如何利用证据和知识的不确定性去更新结论的不确定性

② 在整个推理过程中，如何把初始证据的不确定性传递给最终结论

解决方法

对①，不同推理方法的解决方法不同

对②，不同推理方法的解决方法基本相同，即把当 前结论及其不确定性作为新的证据放入综合数据库，依次传递，直到得出最终结论

5. 不确定性结论的合成

含义：多个不同知识推出同一结论，且不确定性程度不同

方法：视不同推理方法而定

3.1.3 不确定性推理的类型

- 不确定性推理
  - 非数值方法
    - 框架推理
    - 语义网络推理
    - 常识推理
    - …
  - 数值方法
    - 概率类方法
      - 概率推理
      - 基于概率
        - 可信度方法
        - 主观Bayes方法
        - 证据理论
    - 模糊方法

第3章 不确定性知识系统

3.1 不确定性推理概述

3.2 可信度推理

3.3 主观Bayes方法

3.4 证据理论

3.5 模糊推理

3.6 概率推理

- 3.6.1 贝叶斯网络的概念及理论
- 3.6.2 贝叶斯网络推理的概念和类型
- 3.6.3 贝叶斯网络的精确推理
- 3.6.4 贝叶斯网络的近似推理

3.6.1 贝叶斯网络的概念及理论

1. 贝叶斯网络的定义

贝叶斯网络是由美国加州大学的珀尔（J.Pearl）于1985年首先提出的一种模拟人类推理过程中因果关系的不确定性处理模型。它是概率论与图论的结合，其拓扑结构是一个有向无环图。

定义3.17 设$X=\{ X_1, X_2, \ldots, X_n \}$是任何随机变量集，其上的贝叶斯网络可定义为BN=（B_S，B_P）。其中：

① B_S是贝叶斯网络的结构，即一个定义在X上的有向无环图。其中的每一个节点X_i都唯一地对应着X中的一个随机变量，并需要标注定量的概率信息；每条有向边都表示它所连接的两个节点之间的条件依赖关系。

② B_p为贝叶斯网络的条件概率集合，$B_p=\{ P(X_i \mid par(X_i)) \}$。其中，$par(X_i)$表示$X_i$的所有父节点的相应取值，$P(X_i \mid par(X_i))$是节点$X_i$的一个条件概率分布函数。

从以上定义可以看出，贝叶斯网络中的弧是有方向的，且不能形成回路，因此图有始点和终点。

3.6.1 贝叶斯网络的概念及理论

1. 贝叶斯网络的定义

例3.21 假设学生在碰见难题和遇到干扰时会产生焦虑，而焦虑又可导致认知迟缓和情绪波动。请用贝叶斯网络描述这一问题。

解： 图3.4是对上述问题的一种贝叶斯网络描述。在该图中，大写英文字母A、D、I、C和E分别表示节点“产生焦虑”、“碰见难题”、“遇到干扰”、“认知迟缓”和“情绪波动”，并将各节点的条件概率表置于相应节点的右侧。

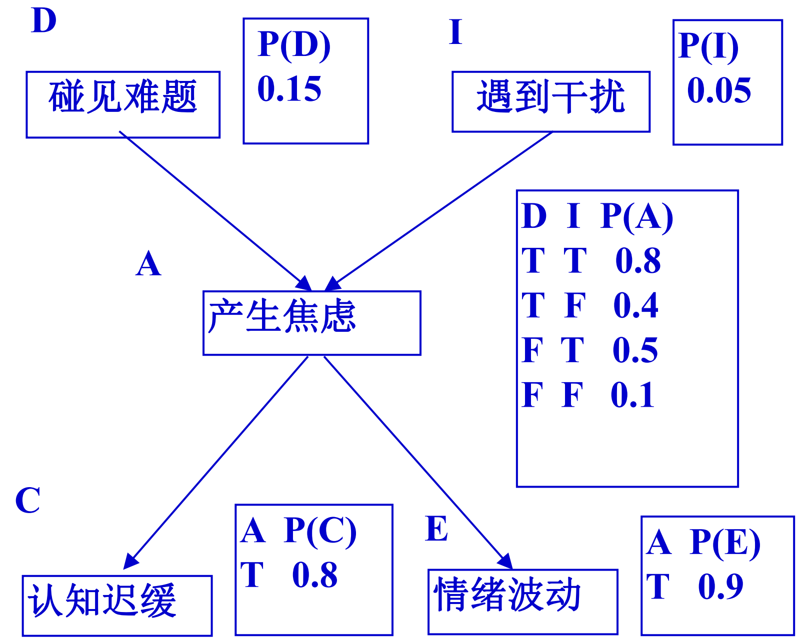


图3.4 关于学习心理的贝叶斯网络

其中，小写英文字母a、d、i、c和e分别表示布尔变量A、D、I、C和E取逻辑值为“True”，¬a、¬d、¬i、¬c和¬e来表示布尔变量A、D、I、C和E取逻辑值为“False”。

贝叶斯网络中每个节点的概率表就是该节点与其父节点之间的一个局部条件概率分布，由于节点D和I无父节点，故取其先验概率来填充。

3.6.1 贝叶斯网络的概念及理论

2. 贝叶斯网络的全联合概率分布表示

全联合概率分布也称为联合概率分布，是概率的合取形式，其定义为

定义3.18 设$X=\{ X_1, X_2, \ldots, X_n \}$为任何随机变量集，其全联合概率分布是指当对每个变量取特定值x_i（$i=1,2,\ldots,n$）时的合取概率，即

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

其简化表示形式为$P(x_1,x_2,\ldots,x_n)$。

由全联合概率分布，再重复使用乘法法则

$$P(x_1,x_2,\ldots,x_n)= P(x_n \mid x_{n-1},x_{n-2},\ldots,x_1)P(x_{n-1},x_{n-2},\ldots,x_1)$$

可以得到如下全联合概率分布表示：

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

$$= P(x_n \mid x_{n-1},x_{n-2},\ldots,x_1)P(x_{n-1}\mid x_{n-2}, x_{n-3},\ldots,x_1)\ldots P(x_2|x_1)P(x_1)$$

$$= \prod_{i=1}^{n} P(x_i \mid x_{i-1}, x_{i-2},\ldots,x_1)$$

这个恒等式对任何随机变量都是成立的，该式亦称为链式法则。

3.6.1 贝叶斯网络的概念及理论

2. 贝叶斯网络的全联合概率分布表示

根据贝叶斯网络的定义，对于节点变量X_i，其取值x_i的条件概率仅依赖于X_i的所有父节点的影响。按照前面的假设，我们用$par(X_i)$表示X_i的所有父节点的相应取值，$P(X_i \mid par(X_i))$是节点X_i的一个条件概率分布函数，则对X的所有节点，应有如下联合概率分布：

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

这个公式就是贝叶斯网络的联合概率分布表示。【条件概率有条件】

可见，贝叶斯网络的联合概率分布要比全联合概率分布简单得多，一个重要原因是其具有局部化特征。所谓局部化特征，是指每个节点只受到整个节点集中少数别的节点的直接影响，而不受这些节点外的其它节点的直接影响。

贝叶斯网络是一种线性复杂度的方法。原因是一个节点仅受该节点的父节点的直接影响，而不受其它节点的直接影响。

3.6.1 贝叶斯网络的概念及理论

3. 贝叶斯网络的条件依赖关系表示

从贝叶斯网络的局部化特征可以看出，贝叶斯网络能实现简化计算的最根本基础是条件独立性，即一个节点与它的祖先节点之间是条件独立的。下面从网络拓扑结构去定义下面两个等价的条件独立关系的判别准则：

(1) 给定父节点，一个节点与非其后代的节点之间是条件独立的。

例如，在图3.4所示的贝叶斯网络中，给定父节点“产生焦虑”的取值（即T或F），节点“认知迟缓”与非其后代节点“碰见难题”和节点“遇到干扰”之间是条件独立的。同样，节点“情绪波动”与非其后代节点“碰见难题”和节点“遇到干扰”之间也是条件独立的。

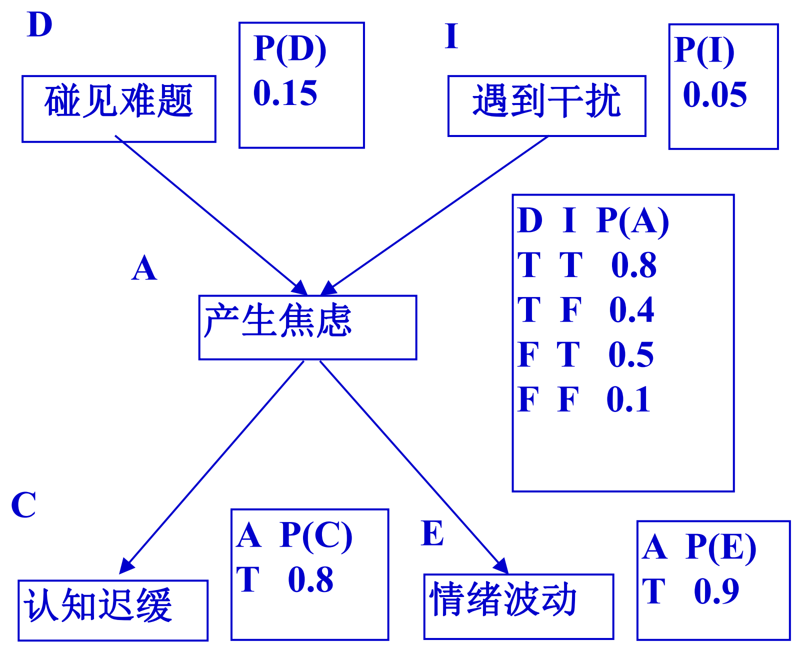


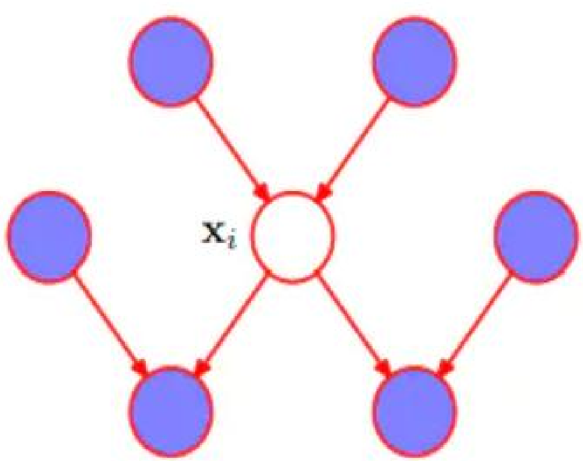
图3.4 关于学习心理的贝叶斯网络

3.6.1 贝叶斯网络的概念及理论

3. 贝叶斯网络的条件依赖关系表示

(2) 给定一个节点，该节点与其父节点、子节点和子节点的父节点一起构成了一个马尔科夫覆盖（**Markov blanket**)，则该节点与马尔科夫覆盖以外的所有节点之间都是条件独立的。

若给定节点“碰见难题”，其马尔科夫覆盖包括“产生焦虑”、“遇到干扰”。此时，节点“碰见难题”与处于马尔科夫覆盖以外的那些节点，如节点“认知迟缓”和节点“情绪波动”之间条件独立。



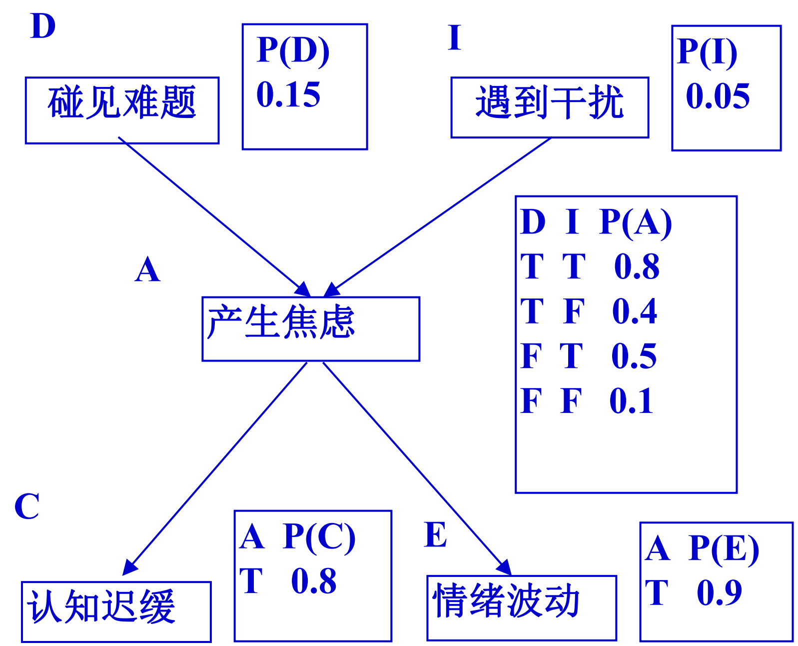


图3.4 关于学习心理的贝叶斯网络

为了针对某一结点研究其条件独立性，我们可以引入**马尔科夫毯**（Markov blanket）的概念。

考虑一个联合概率分布 $p(x_1,\dots,x_D)$ ，它由一个具有 D 个结点的有向图表示。考虑变量 x_i 对应的结点上的条件概率分布，其中条件为所有剩余的变量 $x_{j\neq i}$ ，我们可以进行概率分解

$$
\begin{aligned}
p(x_i|x_{j\neq i}) &= \frac{p(x_1,\dots,x_D)}{\int p(x_1,\dots,x_D)d\,x_i} \\
&= \frac{\prod_k p(x_k|pa_k)}{\int \prod_k p(x_k|pa_k)d\,x_i}
\end{aligned}
$$

对于这个式子，我们现在观察到任何与 x_i 没有函数依赖关系的因子都可以提到 x_i 的积分外，从而在分子与分母之间消去。剔除掉多余的因子后，剩下的因子可以分为两部分：

- 结点 x_i 本身的条件概率分布 $p(x_i|pa_i)$ ，条件概率分布 $p(x_i|pa_i)$ 依赖于 x_i 的父结点
- 条件概率分布 $p(x_k|pa_k)$ ，其中 x_k 是 x_i 的子结点。条件概率分布 $p(x_k|pa_k)$ 依赖于 x_i 的子结点以及同父结点（co-parents），即那些对 应于 x_k （而不是 x_i ）的父结点的变量。

3.6.1 贝叶斯网络的概念及理论

4. 贝叶斯网络的构造

依据贝叶斯网络的联合概率分布表示，其构造过程如下：

(1) 首先建立不依赖于其它节点的根节点，并且根节点可以不止一个。

(2) 加入受根节点影响的节点，并将这些节点作为根节点的子节点。此时，根节点已成为父节点。

(3) 进一步建立依赖于已建节点的子节点。重复这一过程直到叶节点为止。

(4) 对每个根节点，给出其先验概率；对每个中间节点和叶节点，给出其条件概率表。

例如，图3.4所示贝叶斯网络的构建过程如下：

(1) 先建立根节点“碰见难题”和“遇到干扰”；

(2) 加入受根节点影响节点“产生焦虑”，并将其作为两个根节点的子节点。

(3) 进一步加入依赖于已建立节点“产生焦虑”的子节点“思维迟缓”和“情绪波动”。由于这两个新建节点已为叶节点，故节点构建过程终止。

(4) 对每个根节点，给出其先验概率；对每个中间节点和叶节点，给出其条件概率表。

3.6.1 贝叶斯网络的概念及理论

5. 贝叶斯网络的简单应用示例

例3.22 对例3.21所示的贝叶斯网络，若假设已经产生了焦虑情绪，但实际上并未碰见难题，也未遇到干扰，请计算认知迟缓和情绪波动的概率。

解： 令相应变量的取值分别为：

$$a, \neg d, \neg i, c, e$$

其中，无否定符号表示变量取值为True，有否定符号表示变量取值为False，则按贝叶斯网络的联合概率分布表示

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

有：

$$P(c\wedge e\wedge a\wedge \neg d\wedge \neg i)$$

$$=P(c \mid a)P(e \mid a)P(a \mid \neg d\wedge \neg i)P(\neg d)P(\neg i)$$

$$=0.8\times 0.9\times 0.1\times 0.85\times 0.95$$

$$=0.05814$$

即所求的概率为0.05814。

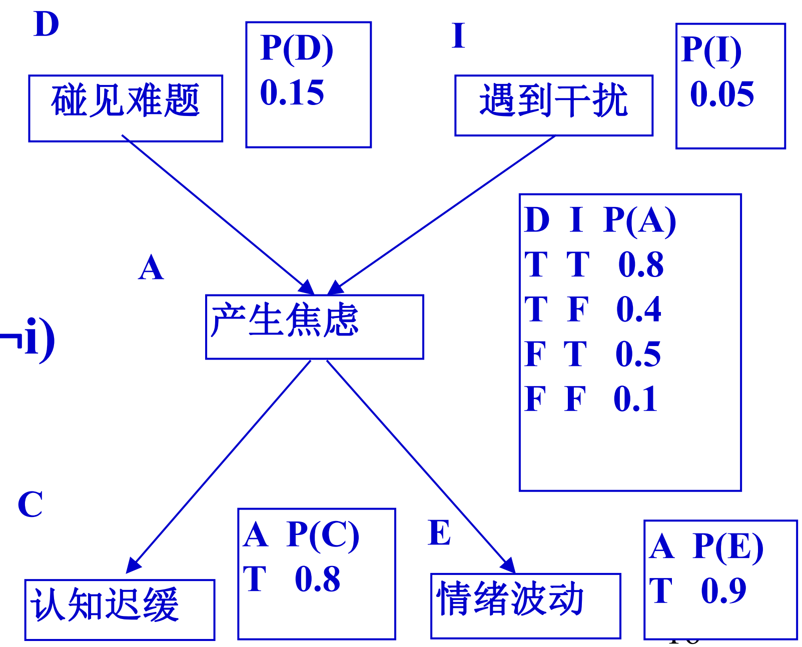


图3.4 关于学习心理的贝叶斯网络

3.6.2 贝叶斯网络推理的概念及类型

1. 贝叶斯网络推理的概念

贝叶斯网络推理是指利用贝叶斯网络模型进行计算的过程，其基本任务就是要在给定一组证据变量观察值的情况下，利用贝叶斯网络计算一组查询变量的后验概率分布。

假设，用X表示某查询变量，E表示证据变量集$\{E_1, E_2, \ldots, E_n\}$，s表示一个观察到的特定事件，Y表示一个非证据变量（亦称隐含变量）集$\{y_1, y_2, \ldots, y_m\}$，则全部变量的集合$V=\{X\}\cup E\cup Y$，其推理就是要查询后验概率$P(X|s)$。

例如，在例3.21所示的贝叶斯网络中，若已观察到的一个事件是“认知迟缓”和“情绪波动”，现在要询问的是“遇到干扰”的概率是多少。这是个贝叶斯网络推理问题，其查询变量为I，观察到的特定事件$s=\{c, e\}$，即求$P(I | c, e)$。

3.6.2 贝叶斯网络推理的概念及类型

2. 贝叶斯网络推理的类型

步骤：首先确定各相邻节点之间的初始条件概率分布；然后对各证据节点取值；接着选择适当推理算法对各节点的条件概率分布进行更新；最终得到推理结果。

类型：分为精确推理和近似推理两大类。

精确推理

是一种可以精确地计算查询变量的后验概率的一种推理方法。它的一个重要前提是要求贝叶斯网络具有单连通特性，即任意两个节点之间至多只有一条无向路径连接。

当贝叶斯网络的两个节点之间存在有多于一条无向路径相连时，它是多连通的。其复杂度是指数级的，应该采用近似推理方法。

近似推理

是一种在不影响推理正确性的前提下，通过适当降低推理精确度来提高推理效率的一类方法。常用的近似推理算法主要有马尔科夫链蒙特卡洛算法等。

3.6.3 贝叶斯网络的精确推理

贝叶斯网络精确推理的主要方法包括基于枚举的算法、基于变量消元的算法和基于团树传播的算法等。其中，最基本的方法是基于枚举的算法，它使用全联合概率分布去推断查询变量的后验概率：

$$P(X \mid s) = \alpha P(X, s) = \alpha \sum_{Y} P(X, s, Y)$$

其中，X表示查询变量；s表示一个观察到的特定事件；Y表示隐含变量集{y_1，y_2，…，y_m}；*α*是归一化常数，用于保证相对于X所有取值的后验概率总和等于1。

为了对贝叶斯网络进行推理，可利用贝叶斯网络的概率分布公式

$$\mathbf{P}(\mathbf{x_1}, \mathbf{x2}, \ldots, \mathbf{x_n}) = \prod_{i=1}^{n} \mathrm{P}(\mathrm{x_i} \mid \mathrm{par}(\mathrm{X_i}))$$

将上式中的*P(X, s, Y)*改写为条件概率乘积的形式。这样，就可通过先对Y的各个枚举值求其条件概率乘积，然后再对各条件概率乘积求总和的方式去计算查询变量的条件概率。下面看一个精确推理的简单例子。

3.6.3 贝叶斯网络的精确推理

例3.23 以例3.21所示的贝叶斯网络为例，假设目前观察到的一个事件s = { c, e }，求在该事件的前提下碰见难题的概率P(D | c, e)是多少？

解： 按照精确推理算法，该询问可表示为：

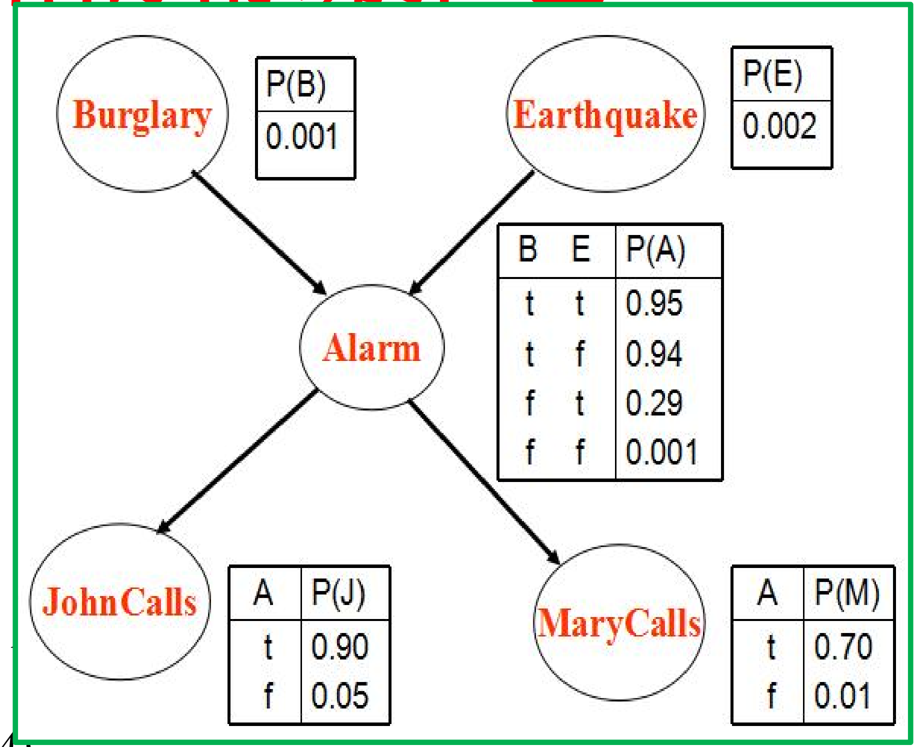
$$P(D|c,e)=\alpha P(D,c,e)=\alpha\sum_{I}\sum_{A}P(D,I,A,c,e)$$

其中，α是归一化常数，D取d和¬d，应用贝叶斯网络的概率分布公式：

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

先对D的不同取值d和¬d分别进行处理。

3.6.3 贝叶斯网络的精确推理



当D取值d时，有

$$P(d\mid c,e)=\alpha\sum_{I}\sum_{A}P(d,I,A,c,e)$$

$$=\alpha\sum_{I}\sum_{A}P(d)P(I)P(A\mid d,I)P(c\mid A)P(e\mid$$

$$=\alpha P(d)\sum_{I}P(I)\sum_{A}P(A\mid d,I)P(c\mid A)P(e\mid A)$$

$$=\alpha P(d)[P(i)(P(a\mid d,i)P(c\mid a)P(e\mid a)+P(\neg a\mid d,i)P(c\mid\neg a)P(e\mid\neg a))+$$
$$P(\neg i)(P(a\mid d,\neg i)P(c\mid a)P(e\mid a)+P(\neg a\mid d,\neg i)P(c\mid\neg a)P(e\mid\neg a))]$$

=α×0.15×[0.05×(0.8×0.8×0.9 + 0.2×0.2×0.1)+
0.95×(0.4×0.8×0.9 + 0.6×0.2×0.1)]
=α×0.15×[0.05×0.58+0.95×0.33] =α×0.15×0.314
=α×0.047

3.6.3 贝叶斯网络的精确推理

当D取值¬d时，有

$$P(\neg d \mid c,e)=\alpha\sum_{I}\sum_{A}P(\neg d,I,A,c,e)$$

$$=\alpha\sum_{I}\sum_{A}P(\neg d)P(I)P(A\mid\neg d,I)P(c\mid A)P(e\mid A)$$

$$=\alpha P(\neg d)[P(i)(P(a\mid\neg d,i)P(c\mid a)P(e\mid a)+P(\neg a\mid\neg d,i)P(c\mid\neg a)P(e\mid\neg a))+$$
$$P(\neg i)(P(a\mid\neg d,\neg i)P(c\mid a)P(e\mid a)+P(\neg a\mid\neg d,\neg i)P(c\mid\neg a)P(e\mid\neg a))]$$

=α×0.85×[0.05×(0.5×0.8×0.9 +0.5×0.2×0.1)+
0.95×(0.1×0.8×0.9 + 0.9×0.2×0.1)]
=α×0.85×[0.05×0.37+ 0.95×0.09] =α×0.85×0.104
=α×0.088

取α=1/(0.047+0.088)=1/0.135。因此有

P(D | c, e)= α(0.047 0.088)=(0.348, 0.652)

即在思维迟缓和情绪波动都发生时，遇到难题的概率是P(d | c, e)= 0.348,不是因为遇到难题的概率是P(¬d | c, e)= 0.652。

3.6.4 贝叶斯网络的近似推理

大规模多连通BN的精确推理计算上是不实际的，须通过近似推理来解决。

后验概率计算的主要近似方法：

马尔可夫链蒙特卡罗（MCMC）方法

马尔科夫链蒙特卡洛（即MCMC）算法是目前使用较广的一种贝叶斯网络近似推理方法。它通过对前一个状态作转移来生成下一个问题状态，通过对某个隐变量进行随机采样来实现对随机变量的条件分布估计。

变分法(Variational method)

……

一个简单例子：一个硬币，它的正反面分布未知，通过多次抛硬币（实际是对未知分布进行采样），然后把正面、反面的次数归一化，可估计出分布.

3.6.4 贝叶斯网络的近似推理

下边介绍的Gibbs采样是马尔科夫链蒙特卡洛（即MCMC）的一个最常用的特殊形式。

例3.24 我们知道，学习情绪会影响学习效果。假设有一个知识点，考虑学生在愉快学习状态下对该知识点的识记、理解、运用的情况，得到了如图3.5所示的多连通贝叶斯网络。如果目前观察到一个学生不但记住了该知识，并且还可以运用该知识，询问这位学生是否理解了该知识。

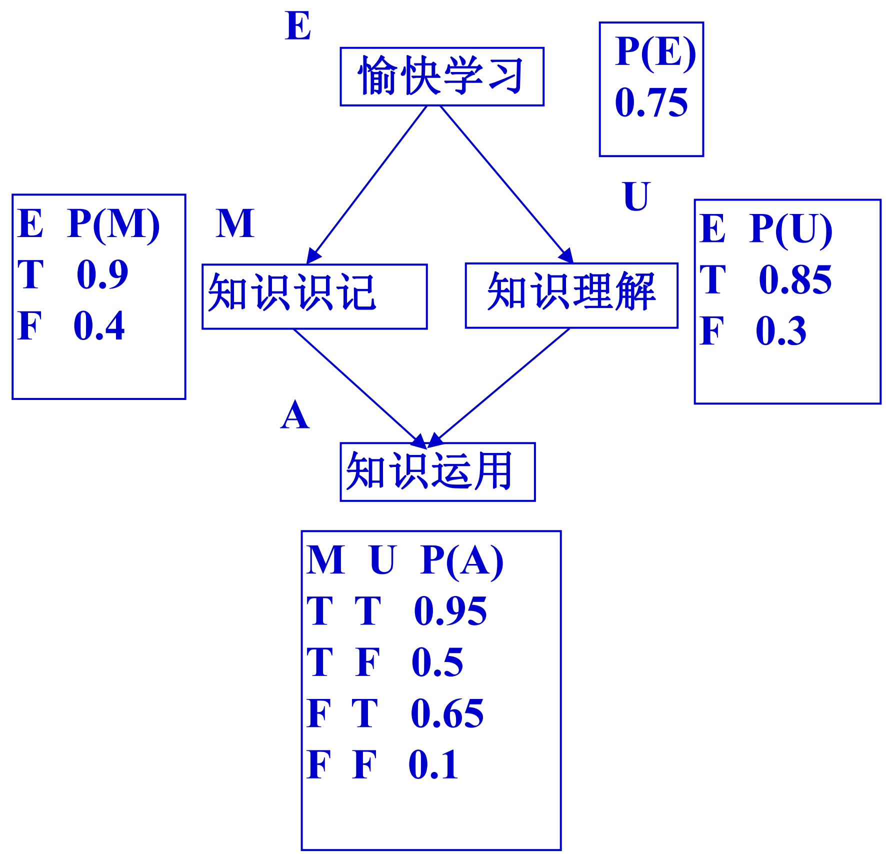


图3.5 关于愉快学习的贝叶斯网络

3.6.4 贝叶斯网络的近似推理

解：为解决这一问题，令E、M、U和A分别表示布尔变量节点“愉快学习”、“知识识记”、“知识理解”和“知识运用”，e、m、u和a分别表示这些变量取值为“True”，各节点边上的表格为相应节点的条件概率表。

本例的询问为P(U| m, a)。应用MCMC算法的推理步骤如下：

(1) 将“知识识记”节点M和“知识运用”节点A作为证据变量，并保持它们的观察值不变；

(2) 将“愉快学习”节点E和“知识理解”节点U作为隐变量，并进行随机初始化。假设，取值分别为e和$\neg$u，问题的初始状态为{ e, m, $\neg$u, a}；

(3) 反复执行如下步骤：

① 对隐变量E进行采样，由于E的马尔科夫覆盖仅包含节点M和U，按照变量M和U的当前值进行采样 P(E|m, $\neg$u)，若采样得到$\neg$e，则生成下一状态{$\neg$e, m, $\neg$u, a }；

② 对隐变量U进行采样，由于U的马尔科夫覆盖包含节点E、M和A，按照变量E、M和A的当前值进行采样P(U|$\neg$e, m, a)，若采样得到u，则生成下一状态{$\neg$e, m, u, a }。

3.6.4 贝叶斯网络的近似推理

在上述过程中，每次采样需要两步。以对隐变量E的采样为例，步骤如下：

第一步，先依据该隐变量的马尔科夫覆盖所包含的变量的当前值，计算该隐变量的条件概率；

例如，对图3.5所给出的问题，在初始状态下，对随机变量E进行采样，第一步可计算$P(E \mid m, \neg u)$的概率。即

$$P(e \mid m, \neg u)= P(e, m, \neg u) / P(m, \neg u)$$

$$=P(e)P(m \mid e)P(\neg u \mid e) / [P(e)P(m \mid e)P(\neg u \mid e) + P(\neg e)P(m \mid \neg e)P(\neg u \mid \neg e)]$$

$$=(0.75 \times 0.9 \times \underline{0.3}) / [0.75 \times 0.9 \times \underline{0.3} + 0.25 \times 0.4 \times \underline{0.3}]$$

$$=0.2025/0.2325=0.8710$$

$$P(\neg e \mid m, \neg u)=0.1290$$

故 $P(E \mid m, \neg u)=(0.8710,0.1290)$

3.6.4 贝叶斯网络的近似推理

第二步，根据$P(E \mid m, \neg u) = (0.8710, 0.1290)$进行采样，确定状态。

其基本方法是，生成一个随机数 $r \in [0,1]$，若$r<0.8710$，则E取e；否则，E取$\neg$e。

假设产生的随机数r=0.46，有0.46<0.8710，则E取e，转移到下一状态$\{e, m, \neg u, a\}$。

（4）重复上述步骤，直到所要求的重复次数N。若U为true, false的次数分别为n_1, n_2，则查询解为：

$$\text{Normalize}(\langle n_1, n_2\rangle) = \langle n_1/N, n_2/N\rangle$$

若上述过程访问了20个U=true的状态和60个U = false的状态，则所求查询的解为<0.25, 0.75>。

3.6.4 贝叶斯网络的近似推理

$$q_i(x \to x') = q_i((x_i, \bar{x}_i) \to (x_i', \bar{x}_i)) = P(x_i' \mid \bar{x}_i, e)$$

故此， $P(x_i' \mid \bar{x}_i, e)$ 也被称为转移概率，但它不是指**T--> F**或**F-->T**（必须变掉）的概率, ***x*'**可为 任何值，包括***x*'=*x***。

Stuart Russell, Peter Norvig. *Artificial Intelligence: A Modern Approach*，*3rd Ed.*

周志华《机器学习》

中的描述是正确的。

为什么**Gibbs**采样可行，可见**Russell**的书第**3**版**14.5.2**节。

MCMC算法描述

function Gibbs-Ask(X, e, BN, N) **return** P(X | e)

local variables: N[X], //关于查询变量X的向量计数，初值0

Z，//BN中的非证据变量集

x， //BN的当前状态

利用Z中变量的随机值来初始化x；

for $j = 1$ to N **do**

for each Z_i in Z **do**

给出Z_i的马尔可夫覆盖mb(Z_i)，并通过对P(Z_i |mb(Z_i))

采样来设置Z_i的值;

N(x) ← N(x) + 1; //x是当前状态x中的查询变量X的值

return Normalize(N[X]) //对N[X]进行归一化

第3章 不确定性知识系统

3.1 不确定性推理概述

3.2 可信度推理

3.2.1 可信度的概念

3.2.2 可信度推理模型

3.2.2 可信度推理的例子

3.3 主观Bayes推理

3.4 证据理论

3.5 模糊推理

3.6 概率推理

3.2.1 可信度的概念

可信度是指人们根据以往经验对某个事物或现象为真的程度的一个判断，或者说是人们对某个事物或现象为真的相信程度。

例如，沈强昨天没来上课，理由是头疼。就此理由，只有以下两种可能：一是真的头疼了，理由为真；二是没有头疼，理由为假。但就听话人而言，因不能确切知道，就只能某种程度上相信，即可信度。

可信度具有一定的主观性，较难把握。但对某一特定领域，让该领域专家给出可信度还是可行的。因此，可信度方法不失为一种实用的不确定性推理方法。

3.2.2 可信度推理模型

知识不确定性的表示

表示形式：

在CF模型中，知识是用产生式规则表示的，其一般形式为：

IF E THEN H (CF(H, E))

其中E是知识的前提条件；H是知识的结论；CF(H, E)是知识的可信度。

说明：

① E可以是单一条件，也可以是复合条件。例如：

$E=(E_1 \text{ OR } E_2) \text{ AND } E_3 \text{ AND } E_4$

② H可以是单一结论，也可以是多个结论

③ CF是知识的静态强度，CF(H, E)的取值为[-1, 1]，表示当E为真时，证据对H的支持程度，其值越大，支持程度越大。

例子：

IF 发烧 AND 流鼻涕 THEN 感冒 (0.8)

表示当某人确实有“发烧”及“流鼻涕”症状时，则有80%的把握是患了感冒的可信度是0.8。

3.2.2 可信度推理模型

可信度的定义

在CF模型中，把CF(H, E)定义为

CF(H, E)=MB(H, E)-MD(H, E)

式中MB称为信任增长度，表示因证据E的出现，使结论H为真的信任增长度，MB(H, E)定义为

$$MB(H,E)=\begin{cases}1, & \text{若} P(H)=1\\ \dfrac{\max\{P(H\mid E),P(H)\}-P(H)}{1-P(H)}, & \text{否则}\end{cases}$$

MD称为不信任增长度，MD(H, E)定义为

$$MD(H,E)=\begin{cases}1, & \text{若} P(H)=0\\ \dfrac{\min\{P(H\mid E),P(H)\}-P(H)}{-P(H)}, & \text{否则}\end{cases}$$

3.2.2 可信度推理模型

可信度的性质

MB和MD的关系

当MB(H, E)>0时，有P(H|E)>P(H)，即E的出现增加了H的概率

当MD(H, E)>0时，有P(H|E)<P(H)，即E的出现降低了H的概率

根据前面对CF(H, E)可信度、MB(H, E)信任增长度、MD(H, E)不信增长度的定义，可得到CF(H, E)的计算公式：

$$CF(H,E)=\begin{cases} MB(H,E)-0=\dfrac{P(H\mid E)-P(H)}{1-P(H)} & \text{若}P(H\mid E)>P(H) \\ 0 & \text{若}P(H\mid E)=P(H) \\ 0-MD(H,E)=-\dfrac{P(H)-P(H\mid E)}{P(H)} & \text{若}P(H\mid E)<P(H) \end{cases}$$

分别解释CF(H,E)>0，CF(H,E)=0，CF(H,E)<0

3.2.2 可信度推理模型

可信度的性质

(1) 互斥性

对同一证据，它不可能既增加对H的信任程度，又同时增加对H的不信任程度，这说明MB与MD是互斥的。即有如下互斥性：

当MB(H, E)>0时，MD(H, E)=0

当MD(H, E)>0时，MB(H, E)=0

(2) 值域

$$0 \le MB(H,E) \le 1, \quad 0 \le MD(H,E) \le 1, \quad -1 \le CF(H,E) \le 1$$

(3) 典型值

当CF(H,E)=1时，有P(H/E)=1，它说明由于E所对应证据的出现使H为真。此时，MB(H, E)=1，MD(H, E)=0。

当CF(H,E)= -1时，有P(H/E)=0，说明由于E所对应证据的出现使H为假。此时，MB(H, E)=0，MD(H,E)=1。

当CF(H,E)= 0时，有MB(H, E)=0、MD(H, E)=0。前者说明E所对应证据的出现不证实H；后者说明E所对应证据的出现不否认H。

3.2.2 可信度推理模型

可信度的性质

(4)对H的信任增长度等于对非H的不信任增长度

根据MB、MD定义及概率性质有（P(¬H|E)<P(¬H)时，另外情况类似):

$$MD(\neg H,E)=\frac{P(\neg H\,|\,E)-P(\neg H)}{-P(\neg H)}=\frac{(1-P(H\,|\,E))-(1-P(H))}{-(1-P(H))}$$

$$=\frac{-P(H\,|\,E)+P(H)}{-(1-P(H))}=\frac{-(P(H\,|\,E)-P(H))}{-(1-P(H))}$$

$$=\frac{P(H\,|\,E)-P(H)}{1-P(H)}=MB(H,E)\quad(\text{信任增长度})$$

再根据CF的定义和MB、MD的互斥性有

CF(H,E)+CF(¬H,E)

=(MB(H,E)-MD(H,E))+(MB(¬H,E)-MD(¬H,E))

=(MB(H,E)-0)+(0-MD(¬H,E)) **(由互斥性，另外情况类似)**

=MB(H,E)-MD(¬H,E)=0

它说明：

(1)对H的信任增长度等于对非H的不信任增长度

(2)对H的可信度与非H的可信度之和等于0

(3)可信度不是概率，概率满足P(H)+P(¬H)=1 和 0≤P(H),P(¬H)≤ 1，而可信度CF不满足

3.2.2 可信度推理模型

可信度的性质

(5)对同一前提E，若支持若干个不同(不相容）的结论H_i(i=1,2,…,n)，则

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

因此，如果发现专家给出的知识有如下情况

$CF(H_1, E)=0.7$，$CF(H_2, E)=0.4$

则因0.7+0.4=1.1>1为非法，应进行调整或规范化。

最后，需要指出，**CF(H,E)**的值一般由领域专家直接给出，而不是计算得到。

3.2.2 可信度推理模型

证据不确定性的表示

基本证据

表示方法：用可信度，取值范围也为[-1,1]。例如，CF(E)，其含义：

CF(E)=1，证据E肯定为真

CF(E)=-1，证据E肯定为假

CF(E)=0，对证据E一无所知

0<CF(E)<1，证据E以CF(E)程度为真

-1<CF(E)<0，证据E以CF(E)程度为假

否定证据

$$CF(\neg E)=-CF(E)$$

组合证据

合取：$E=E_1 \text{ AND } E_2 \text{ AND} \ldots E_n$时，若已知$CF(E_1)$，$CF(E_2)$，…，则

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

析取：$E=E_1 \text{ OR } E_2 \text{ OR} \ldots E_n$时，若已知$CF(E_1)$，$CF(E_2)$，…，则

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

3.2.2 可信度推理模型

不确定性的更新

CF模型中的不确定性推理实际上是从不确定的初始证据出发，不断运用相关的不确定性知识，逐步推出最终结论和该结论可信度的过程。而每一次运用不确定性知识，都需要由证据的不确定性和知识的不确定性去计算结论的不确定性。

不确定性的更新公式

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

若CF(E)<0，则

$$CF(H)=0$$

即该模型没考虑E为假对H的影响。

若CF(E)=1，则

$$CF(H)=CF(H,E)$$

即规则强度CF(H,E)实际上是在E为真时，H的可信度

3.2.2 可信度推理模型

结论不确定性的合成

当有多条知识支持同一个结论，且这些知识的前提相互独立，结论的可信度又不相同时，可利用不确定性的合成算法求出结论的综合可信度。

设有知识：IF E_1 THEN H (CF(H, E_1))

IF E_2 THEN H (CF(H, E_2))

则结论H 的综合可信度可分以下两步计算：

(1) 分别对每条知识求出其CF(H)。即

$CF_1(H)=CF(H, E_1)\times \max\{0, CF(E_1)\}$

$CF_2(H)=CF(H, E_2)\times \max\{0, CF(E_2)\}$

(2) 用如下公式求E_1与E_2对H的综合可信度

$$CF(H)=\begin{cases} CF_1(H)+CF_2(H)-CF_1(H)\times CF_2(H) & \text{若}CF_1(H)\geq 0 \\ & \text{且}CF_2(H)\geq 0 \\ CF_1(H)+CF_2(H)+CF_1(H)\times CF_2(H) & \text{若}CF_1(H)<0 \\ & \text{且}CF_2(H)<0 \\ \dfrac{CF_1(H)+CF_2(H)}{1-\min\{|CF_1(H)|,|CF_2(H)|\}} & \text{若}CF_1(H)\text{与} \\ & CF_2(H)\text{异号} \end{cases}$$

除容易计算外，如此计算有一些好的性质，如：

1、CF(H)∈[−1, 1]

2、符号相反时，互相削弱

3、合并后的CF为单调函数

3.2.3 可信度推理的例子

例3.1 设有如下一组知识：

r_1：IF E_1 THEN H (0.9)

r_2：IF E_2 THEN H (0.6)

r_3：IF E_3 THEN H (-0.5)

r_4：IF E_4 AND (E_5 OR E_6) THEN E_1 (0.8)

已知：$CF(E_2)=0.8$，$CF(E_3)=0.6$，$CF(E_4)=0.5$，$CF(E_5)=0.6$, $CF(E_6)=0.8$

求：CF(H)=?

解：由r_4得到：

$CF(E_1)=CF(E1,E')\times max\{0, CF(E'))$

$=0.8\times max\{0, CF(E_4 \ AND \ (E_5 \ OR \ E_6))\}$

$=0.8\times max\{0, min\{CF(E_4), CF(E_5 \ OR \ E_6)\}\}$

$=0.8\times max\{0, min\{CF(E_4), max\{CF(E_5), CF(E_6)\}\}\}$

$=0.8\times max\{0, min\{CF(E_4), max\{0.6, 0.8\}\}\}$

$=0.8\times max\{0, min\{0.5, 0.8\}\}$

$=0.8\times max\{0, 0.5\} = 0.4$

由r_1得到：$CF_1(H)=CF(H, E_1)\times max\{0, CF(E_1)\}$

$=0.9\times max\{0, 0.4\} = 0.36$

由r_2得到：$CF_2(H)=CF(H, E_2)\times max\{0, CF(E_2)\}$

$=0.6\times max\{0, 0.8\} = 0.48$

由r_3得到：$CF_3(H)=CF(H, E_3)\times max\{0, CF(E_3)\}$

$=-0.5\times max\{0, 0.6\} = -0.3$

根据结论不精确性的合成算法，$CF_1(H)$和$CF_2(H)$同号，有：

$$\begin{aligned} CF_{1,2}(H) &= CF_1(H)+CF_2(H)-CF_1(H)\times CF_2(H) \\ &= 0.36+0.48-0.36\times 0.48 \\ &= 0.84-0.17=0.67 \end{aligned}$$

$CF_{12}(H)$和$CF_3(H)$异号，有：

$$\begin{aligned} CF_{1,2,3}(H) &= \frac{CF_{1,2}(H)+CF_3(H)}{1-\min\{|CF_{1,2}(H)|, |CF_3(H)|\}} \\ &= \frac{0.67-0.3}{1-\min\{0.67, 0.3\}} = \frac{0.37}{0.7} \\ &= 0.53 \end{aligned}$$

即综合可信度为$CF(H)=0.53$

Comments on Certainty Factor Inference

...... Given these difficulties, how can truth-functional systems be made useful in practice? The answer lies in restricting the task and in carefully engineering the rule base so that undesirable interactions do not occur. ......

Heckerman (1986) has shown that under these circumstances, a minor variation on certainty-factor inference was exactly equivalent to Bayesian inference. In other circumstances, certainty factors could yield disastrously incorrect degrees of belief through overcounting of evidence. As rule sets became larger, ... practitioners found that the certainty factors of many other rules had to be "tweaked" when new rules were added.

--Variations and extensions of CF

-- Other systems borrow the idea of CF

第3章 不确定性知识系统

3.1 不确定性推理概述

3.2 可信度推理

3.3 主观Bayes推理

- **3.3.1 主观Bayes方法的概率论基础**
- **3.3.2 主观Bayes方法的推理模型**
- **3.3.3 主观Bayes推理的例子**

3.4 证据理论

3.5 模糊推理

3.6 概率推理

3.3.1 主观Bayes方法的概率论基础

1. 全概率公式

定理3.1 设事件$A_1,A_2,\ldots,A_n$满足：

(1)任意两个事件都互不相容，即当$i\neq j$时，有$A_i\cap A_j=\Phi$ $(i=1,2,\ldots,n;\ j=1,2,\ldots,n)$；

(2) $P(A_i)>0\ (i=1, 2, \ldots ,n)$;

(3) $D=\bigcup_{i=1}^{n} A_i$ （随机试验的样本空间D的一个划分）

则对任何事件B有下式成立：

$$P(B)=\sum_{i=1}^{n} P(A_i)\times P(B\mid A_i)$$

该公式称为全概率公式，它提供了一种计算P(B)的方法。

3.3.1 主观Bayes方法的概率论基础

2. Bayes公式

定理3.2 设事件$A_1,A_2,\ldots,A_n$满足定理3.1规定的条件，则对任何事件B有下式成立：

$$P(A_i \mid B)=\frac{P(A_i)\times P(B/A_i)}{\sum_{j=1}^{n}P(A_j)\times P(B/A_j)} \qquad i=1,2,\cdots,n$$

该定理称为Bayes定理，上式称为Bayes公式。

其中，$P(A_i)$是事件A_i的先验概率，$P(B|A_i)$是在事件A_i发生条件下事件B的条件概率；$P(A_i|\mathrm{B})$是在事件B发生条件下事件A_i的条件概率。

如果把全概率公式代入Bayes公式，则有：

$$P(A_i \mid B)=\frac{P(A_i)\times P(B\mid A_i)}{P(B)} \qquad i=1,2,\cdots,n$$

即

$$P(A_i \mid B)\cdot P(B)=P(B\mid A_i)\cdot P(A_i) \qquad i=1,2,\cdots n$$

这是Bayes公式的另一种形式。

Bayes公式给出了用逆概率$P(B|A_i)$求原概率$P(A_i|B)$的方法。

3.3.2 主观Bayes方法的推理模型

1. 知识不确定性的表示

表示形式：在主观Bayes方法中，知识是用产生式表示的，其形式为：

IF E THEN (LS, LN) H

其中，(LS, LN)用来表示该知识的知识强度，LS(充分性度量)和LN(必要性度量)的表示形式分别为：

$$LS=\frac{P(E\mid H)}{P(E\mid \neg H)}$$

$$LN=\frac{P(\neg E\mid H)}{P(\neg E\mid \neg H)}=\frac{1-P(E\mid H)}{1-P(E\mid \neg H)}$$

LS和LN的含义：由本节前面给出的的Bayes公式可知：

$$P(H\mid E)=\frac{P(E\mid H)\times P(H)}{P(E)}$$

$$P(\neg H\mid E)=\frac{P(E\mid \neg H)\times P(\neg H)}{P(E)}$$

3.3.2 主观Bayes方法的推理模型

1. 知识不确定性的表示

两式相除得：

$$\frac{P(H\mid E)}{P(\neg H\mid E)}=\frac{P(E\mid H)}{P(E\mid \neg H)}\bullet\frac{P(H)}{P(\neg H)} \qquad (3.1)$$

为讨论方便，下面引入几率函数

$$O(X)=\frac{P(X)}{1-P(X)} \qquad O(X)=\frac{P(X)}{P(\neg X)} \qquad (3.2)$$

可见，X的几率等于X出现的概率与X不出现的概率之比，P(X)与O(X)的变化一致，且有：

P(X)=0 时有 O(X)=0

P(X)=1 时有 O(X)=+∞

即把取值为[0,1]的P(X)放大为取值为[0,+∞]的O(X)

3.3.2 主观Bayes方法的推理模型

1. 知识不确定性的表示

把(3.2)式代入(3.1)式有：

$$P(A_i \mid B) = \frac{P(A_i) \times P(B / A_i)}{\sum_{j=1}^{n} P(A_j) \times P(B / A_j)} \qquad i = 1,2,\cdots,n$$

$$O(H \mid E) = \frac{P(E \mid H)}{P(E \mid \neg H)} \bullet O(H)$$

再把LS代入此式，可得：

$$O(H \mid E) = LS \bullet O(H) \qquad \textbf{(3.3)}$$

$$P(H \mid E) = \frac{LS \bullet P(H)}{(LS - 1) \bullet P(H) + 1} \qquad (3.5)$$

同理可得到关于LN的公式：

$$O(H \mid \neg E) = LN \bullet O(H) \qquad \textbf{(3.4)}$$

$$P(H \mid \neg E) = \frac{LN \bullet P(H)}{(LN - 1) \bullet P(H) + 1} \qquad (3.6)$$

式(3.3)和(3.4)就是修改的Bayes公式。可见：

当E为真时可用（3.3）计算O(H|E)　　【不意味着P(E)=1才能用】

当E为假时可用（3.4）计算O(H|¬E)

3.3.2 主观Bayes方法的推理模型

1. 知识不确定性的表示

LS的性质：

当LS>1时，O(H|E)>O(H)，说明E支持H，LS越大，E对H的支持越充分。当LS→∞时，O(H|E)→∞，即P(H|E)→1，表示由于E的存在将导致H为真。

当LS=1时，O(H|E)=O(H)，说明E对H没有影响。

当LS<1时，O(H|E)<O(H)，说明E不支持H。

当LS=0时，O(H|E)=0，说明E的存在使H为假。

LN的性质：

当LN>1时，O(H|¬E)>O(H)，说明¬E支持H，即由于E的不出现，增大了H为真的概率。并且，LN越大，¬E对H为真的支持就越强。当LN→∞时，O(H|¬E)→∞，即P(H|¬E)→1，表示由于¬E的存在将导致H为真。

当LN=1时，O(H|¬E)=O(H)，说明¬E对H没有影响。

当LN<1时，O(H|¬E)<O(H)，说明¬E不支持H，即由于¬E的存在，使H为真的可能性下降，或者说由于E不存在，将反对H为真。当LN→0时O(H|¬E) →0，即LN越小，E的不出现就越反对H为真，这说明H越需要E的出现。

当LN=0时，O(H|¬E)=0，说明¬E的存在（即E不存在）将导致H为假。

3.3.2 主观Bayes方法的推理模型

1. 知识不确定性的表示

LS与LN的关系

由于E和 ¬E不会同时支持或同时排斥H，因此只有下述三种情况：

① $LS>1$ 且 $LN<1$

② $LS<1$ 且 $LN>1$

③ $LS=LN=1$

证①：

$LS>1 \Leftrightarrow P(E|H)/P(E|\neg H)>1 \Leftrightarrow P(E|H) >P(E|\neg H)$

$\Leftrightarrow 1-P(E|H) < 1-P(E|\neg H)$

$\Leftrightarrow P(\neg E|H) < P(\neg E|\neg H)$

$\Leftrightarrow P(\neg E|H) /P(\neg E|\neg H) <1$

$\Leftrightarrow LN <1$

同理可证②、③，证明略

3.3.2 主观Bayes方法的推理模型

2. 证据不确定性的表示

基本证据的表示：

在主观Bayes方法中，基本证据E的不精确性是用其概率或几率来表示的。概率与几率之间的关系为：

$$O(E)=\frac{P(E)}{1-P(E)}=\begin{cases}0 & \text{当}E\text{为假时}\\ \infty & \text{当}E\text{为真时}\\ (0,\infty) & \text{当}E\text{非真也非假时}\end{cases}$$

在实际应用中，除了需要考虑证据E的先验概率与先验几率外，往往还需要考虑在当前观察下证据E的后验概率或后验几率。

以概率情况为例，对初始证据E，用户可以根据当前观察S将其先验概率P(E)更改为后验概率P(E|S)，即相当于给出证据E的动态强度。

3.3.2 主观Bayes方法的推理模型

2. 证据不确定性的表示

组合证据不确定性的计算：

证据的基本组合方式只有合取和析取两种。

当组合证据是多个单一证据的合取时，例

$$E=E_1 \text{ AND } E_2 \text{ AND } \dots \text{ AND } E_n$$

如果已知在当前观察S下，每一单一证据E_i有概率$P(E_1|S)$, $P(E_2|S)$, … ,$P(E_n|S)$，则

$$P(E|S)=\min\{ P(E_1|S), P(E_2|S), \dots ,P(E_n|S)\}$$

当组合证据是多个单一证据的析取时，例

$$E=E_1 \text{ OR } E_2 \text{ OR } \dots \text{ OR } E_n$$

如果已知在当前观察S下，每一单一证据E_i有概率$P(E_1|S)$, $P(E_2|S)$, … ,$P(E_n|S)$，则

$$P(E|S)=\max\{ P(E_1|S), P(E_2|S), \dots ,P(E_n|S)\}$$

3.3.2 主观Bayes方法的推理模型

3. 不确定性的更新

不确定性更新需要使用杜达等人给出的公式：

P(H|S)=P(H|E)×P(E|S)+P(H|¬E)×P(¬E|S)　　(3.7)

下面分四种情况讨论：

(1)P(E|S)=1

当P(E|S)=1时，P(¬E|S)=0。由(3.7)式和(3.5)式可得

$$P(H \mid S) = P(H \mid E) = \frac{LS \times P(H)}{(LS-1) \times P(H) + 1}$$

(2)P(E|S)=0

当P(E|S)=0时，P(¬E|S)=1。由(3.7)式和(3.6)式可得

$$P(H \mid S) = P(H \mid \neg E) = \frac{LN \times P(H)}{(LN-1) \times P(H) + 1}$$

(3)P(E|S)=P(E)

当P(E|S)=P(E)时，表示E与S无关。由(3.7)式和全概率公式可得

$$\begin{aligned} P(H|S) &= P(H|E) \times P(E|S) + P(H|\neg E) \times P(\neg E|S) \\ &= P(H|E) \times P(E) + P(H|\neg E) \times P(\neg E) = P(H) \end{aligned}$$

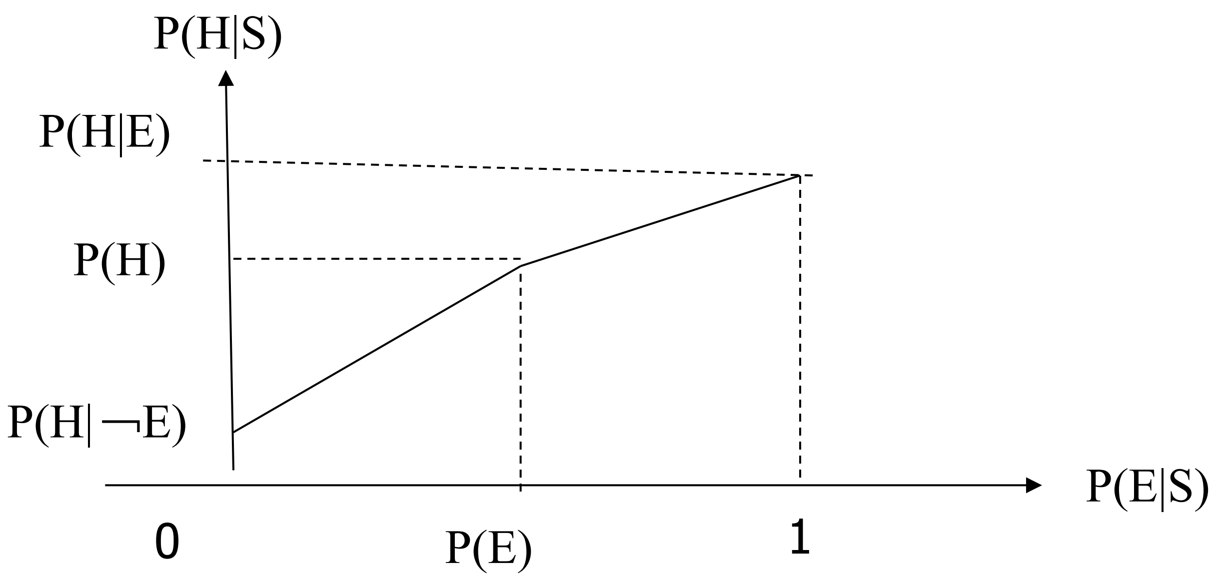
3.3.2 主观Bayes方法的推理模型

3. 不确定性的更新

(4) P(E/S)为其它值

上面已经得到了P(E|S)的3个特殊值：0，P(E),1；它们分别对应的3个值为P(H|¬E),P(H),P(H|E)。由此构造的分段线性插值函数为：

$$P(H|S)=\begin{cases} P(H|\neg E)+\dfrac{P(H)-P(H|\neg E)}{P(E)}\times P(E|S), & 若 0\le P(E|S)<P(E) \\ P(H)+\dfrac{P(H|E)-P(H)}{1-P(E)}\times\left[P(E|S)-P(E)\right], & 若 P(E)\le P(E|S)\le 1 \end{cases} \tag{3.8}$$



3.3.2 主观Bayes方法的推理模型

4. 结论不确定性的合成

假设有n条知识都支持同一结论H，并且这些知识的前提条件分别是n个相互独立的证据E_1、E_2、…、E_n，而每个证据所对应的观察又分别是S_1、S_2、…、S_n。在这些观察下，求H的后验概率的方法是：首先对每条知识分别求出H的后验几率$O(H|S_i)$，然后利用这些后验几率并按下述公式求出在所有观察下H的后验几率：

$$O(H \mid S_1, S_2, \ldots, S_n) = \frac{O(H \mid S_1)}{O(H)} \bullet \frac{O(H \mid S_2)}{O(H)}, \ldots, \frac{O(H \mid S_n)}{O(H)} \bullet O(H) \quad \textbf{(3.9)}$$

3.3.3 主观Bayes推理的例子

例3.2 设有规则

r_1: IF E_1 THEN (2, 0.0001) H_1

r_2: IF E_1 AND E_2 THEN (100, 0.001) H_1

r_3: IF H_1 THEN (200, 0.01) H_2

已知：$P(E_1)=P(E_2)=0.6$

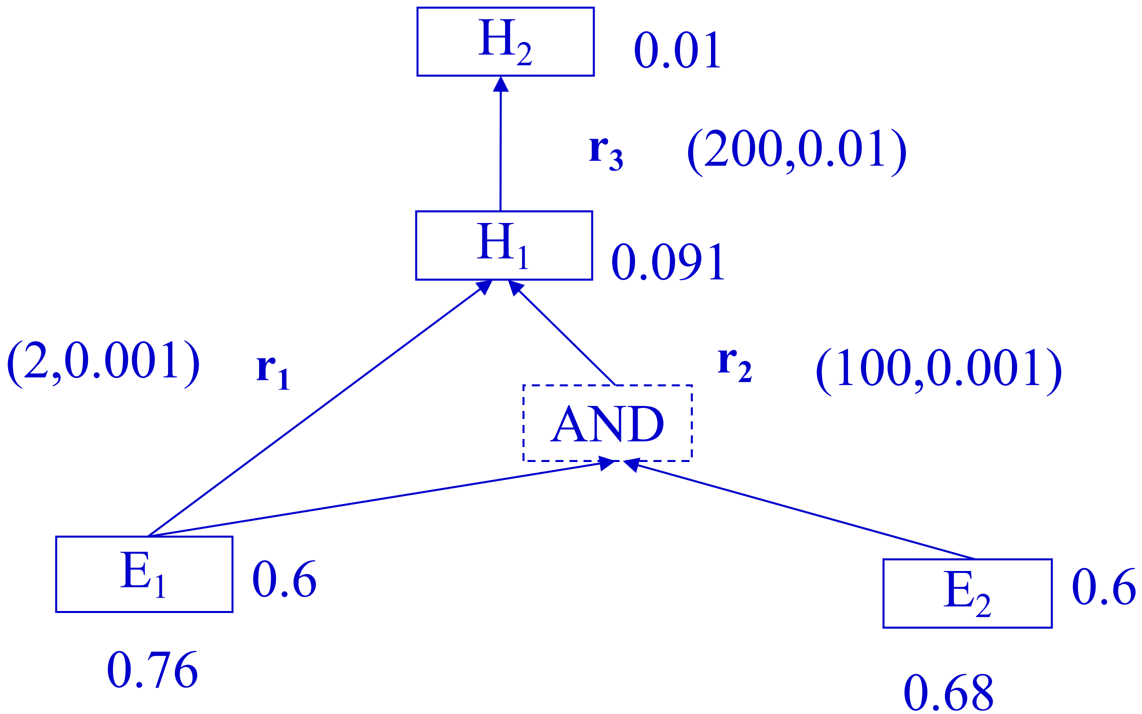
$P(H_1)=0.091$，$P(H_2)=0.01$

用户回答: $P(E_1|S_1)=0.76$, $P(E_2|S_2)=0.68$

求：$P(H_2|S_1,S_2)=?$

3.3.3 主观Bayes推理的例子

解：由已知知识得到的推理网络如下图所示。



P(H|S)=P(H|E)×P(E|S)+P(H|¬E)×P(¬E|S) **(3.7)**

$$O(H \mid S_1, S_2, \ldots, S_n) = \frac{O(H \mid S_1)}{O(H)} \bullet \frac{O(H \mid S_2)}{O(H)}, \ldots, \frac{O(H \mid S_n)}{O(H)} \bullet O(H)$$

3.3.3 主观Bayes推理的例子

(1) 计算$O(H_1|S_1)$

先把$P(H_1)$更新为E_1下的后验概率$P(H_1|E_1)$　　**根据（3.5）**

$$P(H_1 \mid E_1) = \frac{LS_1 \times P(H_1)}{(LS_1 - 1) \times P(H_1) + 1} = \frac{2 \times 0.091}{(2-1) \times 0.091 + 1}$$
$$= 0.167$$

由于$P(E_1|S_1)=0.76>P(E_1)$，使用(3.8)式的后半部分，得$P(H_1|S_1)$为：

$$P(H_1 \mid S_1) = P(H_1) + \frac{P(H_1 \mid E_1) - P(H_1)}{1 - P(E_1)} \times (P(E_1 \mid S_1) - P(E_1))$$
$$= 0.091 + \frac{(0.167 - 0.091)}{1 - 0.6} \times (0.76 - 0.6)$$
$$= 0.121$$

$$O(H_1 \mid S_1) = \frac{P(H_1 \mid S_1)}{1 - P(H_1 \mid S_1)} = \frac{0.121}{1 - 0.121} = 0.138$$

3.3.3 主观Bayes推理的例子

(2) 计算$O(H_1|(S_1\ \text{AND}\ S_2))$

由于r_2的前件是E_1、E_2的合取关系，且已知

$P(E_1|S_1)=0.76$，$P(E_2|S_2)=0.68$，

即$P(E_2|S_2)<P(E_1|S_1)$。按合取取最小的原则，这里仅考虑E_2对H_1的影响，即把计算$O(H_1|(S_1\ \text{AND}\ S_2))$的问题转化为计算$O(H_1|S_2)$的问题。 **[推理稍不严谨]**

把H_1的先验概率$P(H_1)$更新为在E_2下的后验概率$P(H_1|E_2)$ **[$P(E|S)=P(E_2|S_2)$]**

$$P(H_1 \mid E_2) = \frac{LS_2 \times P(H_1)}{(LS_2 - 1)\times P(H_1) + 1} = \frac{100 \times 0.091}{(100-1)\times 0.091 + 1} = 0.909$$

又由于$P(E_2|S_2)=0.68>P(E_2)$，还使用(3.8)式的后半部分，得$P(H_1|S_2)$为：

$$P(H_1 \mid S_2) = P(H_1) + \frac{P(H_1 \mid E_2) - P(H_1)}{1 - P(E_2)} \times (P(E_2 \mid S_2) - P(E_2))$$

$$= 0.091 + \frac{(0.909 - 0.091)}{1 - 0.6} \times (0.68 - 0.6) = 0.255$$

$$O(H_1 \mid S_2) = \frac{P(H_1 \mid S_2)}{1 - P(H_1 \mid S_2)} = \frac{0.255}{1 - 0.255} = 0.342$$

3.3.3 主观Bayes推理的例子

(3) 计算$O(H_1|S_1, S_2)$

先将H_1的先验概率转换为先验几率

$$O(H_1)=\frac{P(H_1)}{1-P(H_1)}=\frac{0.091}{1-0.091}=0.1$$

再根据合成公式计算H_1的后验几率

$$O(H_1\mid S_1,S_2)=\frac{O(H_1\mid S_1)}{O(H_1)}\times\frac{O(H_1\mid S_2)}{O(H_1)}\times O(H_1)$$

$$=\frac{0.138}{0.1}\times\frac{0.342}{0.1}\times 0.1=0.472$$

然后再将后验几率转换为后验概率

$$P(H_1\mid S_1,S_2)=\frac{O(H_1\mid S_1,S_2)}{1+O(H_1\mid S_1,S_2)}=\frac{0.472}{1+0.472}=0.321$$

3.3.3 主观Bayes推理的例子

(4) 计算$P(H_2|S_1,S_2)$

对r_3，H_1相当于已知事实，H_2为结论。将H_2的先验概率$P(H_2)$更新为在H_1下的后验概率$P(H_2|H_1)$

$$P(H_2 \mid H_1) = \frac{LS_3 \times P(H_2)}{(LS_3 - 1) \times P(H_2) + 1} = \frac{200 \times 0.01}{(200-1) \times 0.01 + 1} = 0.669$$

由于$P(H_1|S_1,S_2) = 0.321 > P(H_1)$，仍使用(3.8)式的后半部分，得到在当前观察$S_1$、$S_2$下$H_2$的后验概率$P(H_2|S_1,S_2)$

$$P(H_2 \mid S_1,S_2) = P(H_2) + \frac{P(H_2 \mid H_1) - P(H_2)}{1 - P(H_1)} \times [P(H_1 \mid S_1,S_2) - P(H_1)]$$

$$= 0.1 + \frac{0.669 - 0.01}{1 - 0.091} \times (0.321 - 0.091) = 0.177$$

可以看出，H_2的先验概率是0.01，通过r_1、r_2、r_3及初始证据进行推理，最后推出H_2的后验概率为0.177，相当于概率增加了16倍多。

3.3.4 主观Bayes推理的特性

主要优点：

- 主观贝叶斯方法的大部分公式是在概率基础上（近似）推导出来的，有较强的理论基础；
- 规则的LS和LN由领域专家根据实践经验给出，避免了大量的数理统计工作；
- 从推理过程可以看出，它实现了不确定性的逐级传递，是一种比较实用而又灵活的不确定性推理方法。

主要缺点：

它要求领域专家给出规则的同时给出P(H)，有时比较困难的。

第3章 不确定性知识系统

3.1 不确定性推理概述

3.2 可信度推理

3.3 主观Bayes方法

3.4 证据理论

3.4.1 证据理论的形式化描述

3.4.2 证据理论的推理模型

3.4.3 推理实例

3.4.4 证据理论推理的特征

3.5 模糊推理

3.6 概率推理

3.4.1 证据理论的形式化描述

1. 概率分配函数

DS理论处理的是集合上的不确定性问题，为此需要先建立命题与集合之间的一一对应关系，以把命题的不确定性问题转化为集合的不确定性问题。

(1) 幂集

设Ω为样本空间，且Ω中的每个元素都相互独立，则由Ω的所有子集构成的幂集记为2^{Ω}。

当Ω中的元素个数为N时，则其幂集2^{Ω}的元素个数为2^N。

例3.3 设Ω={红，黄，白}，求Ω的幂集2^{Ω}。

解： Ω的幂集可包括如下子集：

$A_0=\Phi$,　　A_1={红},　　A_2={黄},　　A_3={白},

A_4={红，黄},　　A_5={红，白},　　A_6={黄，白},　　A_7={红，黄，白}

其中，Φ表示空集，空集也可表示为{}。上述子集的个数正好是$2^3=8$

3.4.1 证据理论的形式化描述

1. 概率分配函数

(2) 一般的概率分配函数

定义3.1 设函数m：$2^{\Omega}\to[0, 1]$，且满足

$$m(\Phi)=0$$

$$\sum_{A\subseteq\Omega} m(A)=1$$

则称m是2^{Ω}上的概率分配函数，m(A)称为A的基本概率数。

例3.4 对例3.3所给出的有限集Ω，若定义2^{Ω}上的一个基本函数m：

m({}, {红}, {黄}, {白}, {红，黄}, {红，白}, {黄，白}, {红，黄，白})

=(0, 0.3, 0, 0.1, 0.2, 0.2, 0, 0.2)

请说明该函数满足概率分配函数的定义。

解：(0, 0.3, 0, 0.1, 0.2, 0.2, 0, 0.2)分别是幂集2^{Ω}中各个子集的基本概率数。显然m满足

$$m(\Phi)=0$$

$$\sum_{A\subseteq\Omega} m(A)=1$$

即满足概率分配函数的定义。

3.4.1 证据理论的形式化描述

1. 概率分配函数

对一般概率分配函数的说明

(1) 概率分配函数的作用是把Ω的任一子集映射为[0，1]上的一个数m(A)

当A⊂Ω，且A由单个元素组成时，则m(A)表示对A的精确信任度；

当A⊂Ω、A≠Ω，且A由多个元素组成时，m(A)也表示对A的精确信任度，但却不知道这部分信任度该分给A中哪些元素；

当A=Ω时，则m(A)表示不知道该如何分配的部分。

例如，对上例所给出的有限集Ω及基本函数m，当

A={红}时，有m(A)=0.3，它表示对命题“x是红色”的精确信任度为0.3。

B= {红，黄}时，有m(B)=0.2，x是红或黄的信任度0.2，但不知道怎样分。

C=Ω={红，黄，白}时，有m(Ω)=0.2，同样不知道该怎样分配。

(2) 概率分配函数不是概率 (是对各子集的信任分配)

例如，在例3.3中，m符合概率分配函数的定义，但

m({红})+m({黄})+m({白})=0.3+0+0.1=0.4<1

因此m不是概率，因为概率P要求：P(红)+P(黄)+P(白)=1

3.4.1 证据理论的形式化描述

1. 概率分配函数

(3) 一个特殊的概率分配函数

设$\Omega=\{s_1,s_2,\ldots,s_n\}$，m为定义在$2^\Omega$上的概率分配函数，且m满足

$$(1) m(\{s_i\}) \geq 0 \qquad \text{对任何} s_i \in \Omega$$

$$(2) \sum_{i=1}^{n} m(\{s_i\}) \leq 1$$

$$(3) m(\Omega) = 1 - \sum_{i=1}^{n} m(\{s_i\})$$

$$(4) \text{当} A \subset \Omega \text{且} |A| > 1 \text{或} |A| = 0 \text{时}, m(A) = 0$$

其中，|A| 表示命题A所对应的集合中的元素个数。

该概率分配函数的特殊性：

① 只有当子集中的元素个数为1时，其概率分配数才有可能大于0；

② 当子集中有多个或0个元素，且不等于全集时，其概率分配数均为0；

③ 全集Ω的概率分配数按(3)计算。

3.4.1 证据理论的形式化描述

1. 概率分配函数

例3.5 设Ω={红，黄，白}，有如下概率分配函数

m({}，{红}，{黄}，{白}，{红，黄，白})

=(0, 0.6, 0.2, 0.1, 0.1)

其中：m({红，黄})=m({红，白})=m({黄，白})=0，可见，m符合上述概率分配函数的定义。

(4) 概率分配函数的合成

在实际中，由于证据来源不同，对同一幂集，可能得到不同的概率分配函数，对前述特殊概率分配函数，有

定义3.2 设m_1和m_2是2^Ω上的基本概率分配函数，它们的正交和 $m = m_1 \oplus m_2$ 定义为

$$m(\{s_i\}) = K^{-1} \times [m_1(s_i) \times m_2(s_i) + m_1(s_i) \times m_2(\Omega) + m_1(\Omega) \times m_2(s_i)]$$

其中：

$$K = m_1(\Omega) \times m_2(\Omega) + \sum_{i=1}^{n} [m_1(s_i) \times m_2(s_i) + m_1(s_i) \times m_2(\Omega) + m_1(\Omega) \times m_2(s_i)]$$

其它文献中与此定义式未必相同，可能更general，可能等价，可能改进

3.4.1 证据理论的形式化描述

2. 信任函数和似然函数

根据上述特殊的概率分配函数，可构造其信任函数和似然函数。

定义3.3 对任何命题A$\subseteq$Ω,其信任函数(Belief function)为

$$Bel(A)=\sum_{s_i\in A} m(\{s_i\})$$

$$Bel(\Omega)=\sum_{B\subseteq\Omega} m(B)=\sum_{i=1}^{n} m(\{s_i\})+m(\Omega)=1$$

信任函数也称为下限函数，表示对A的总体信任度。

定义3.4 对任何命题A $\subseteq$Ω,其似然函数(plausibility function/似真)为

$$\begin{aligned}Pl(A)&=1-Bel(\neg A)=1-\sum_{s_i\in\neg A} m(\{s_i\})=1-[\sum_{i=1}^{n} m(\{s_i\})-\sum_{s_i\in A} m(\{s_i\})]\\&=1-[1-m(\Omega)-Bel(A)]\\&=m(\Omega)+Bel(A)\end{aligned}$$

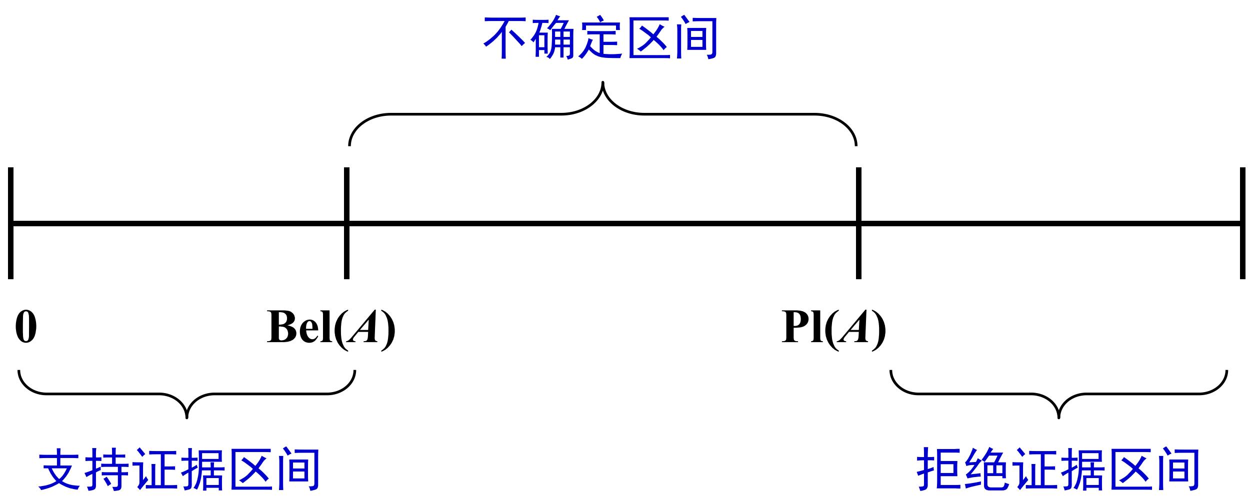
$$Pl(\Omega)=1-Bel(\neg\Omega)=1-Bel(\Phi)=1$$

似然函数也称为上限函数，表示对A的非假信任度。可以看出，对任何命题A $\subseteq$Ω、 B$\subseteq$Ω都有 【非Ω】

Pl(A)-Bel(A) = Pl(B)-Bel(B) = m(Ω)

3.4.1 证据理论的形式化描述

2. 信任函数和似然函数



信任度是对假设信任程度的下限估计—**悲观估计**；

似然度是对假设信任程度的上限估计—**乐观估计**。

Suppose you pick a coin, which might or might not be fair. Now suppose an expert testifies with 90% certainty that the coin is fair (i.e., he is 90% sure that P(Heads) =0.5). Then Dempster—Shafer theory gives Bel(Heads)= 0.9 x 0.5 = 0.45 and likewise Bel(¬Heads) = 0.45. There is still a 10 percentage point "gap" that is not accounted for by the evidence.

3.4.1 证据理论的形式化描述

2.信任函数和似然函数

例3.6 设Ω={红，黄，白}，概率分配函数

m({}，{红}，{黄}，{白}，{红，黄，白})

=(0, 0.6, 0.2, 0.1, 0.1)

A={红，黄}，求m(Ω)、Bel(A)和Pl(A)的值。

解：m(Ω)=1-[m({红})+m({黄})+m({白})]

=1-(0.6+0.2+0.1)=0.1

Bel({红，黄})=m({红})+m({黄})=0.6+0.2=0.8

Pl({红，黄})=m(Ω)+Bel({红，黄})=0.1+0.8=0.9

或

Pl({红，黄})=1-Bel(¬{红，黄})=1-Bel({白})=1-0.1=0.9

3.4.1 证据理论的形式化描述

3. 类概率函数

定义3.5 设Ω为有限域，对任何命题A $\subseteq$Ω，命题A的类概率函数为

$$f(A)=Bel(A)+\frac{|A|}{|\Omega|}\times[Pl(A)-Bel(A)]$$

其中，|A|和|Ω|分别是A及Ω中元素的个数。

类概率函数f(A)具有如下性质：

(1) $\sum_{i=1}^{n} f(\{s_i\})=1$

证明：

因为

$$f(\{s_i\})=Bel(\{s_i\})+\frac{|\{s_i\}|}{|\Omega|}\bullet[Pl(\{s_i\})-Bel(\{s_i\})]$$

$$=m(\{s_i\})+\frac{1}{n}\bullet m(\Omega) \qquad (i=1,2,\dots,n)$$

故

$$\sum_{i=1}^{n} f(\{s_i\})=\sum_{i=1}^{n}[m(s_i)+\frac{1}{n}\bullet m(\Omega)]$$

$$=\sum_{i=1}^{n} m(\{s_i\})+m(\Omega)=1$$

3.4.1 证据理论的形式化描述

3. 类概率函数

(2)对任何 $A \subseteq \Omega$,有Bel(A)≤f(A)≤Pl(A)

证明：根据 f(A)定义 $f(A) = Bel(A) + \frac{|A|}{|\Omega|} \times [Pl(A) - Bel(A)]$

因

$$Pl(A) - Bel(A) = m(\Omega), \qquad \frac{|A|}{|\Omega|} \geq 0$$

故

$$Bel(A) \leq f(A)$$

又因 $\frac{|A|}{|\Omega|} \leq 1$**，即**

$$f(A) \leq Bel(A) + Pl(A) - Bel(A)$$

所以

$$f(A) \leq Pl(A)$$

3.4.1 证据理论的形式化描述

3. 类概率函数

(3) 对任何$A \subseteq \Omega$，有f(¬A)=1-f(A)

证明：因为

$$f(\neg A) = Bel(\neg A) + \frac{|\neg A|}{|\Omega|} \times [Pl(\neg A) - Bel(\neg A)]$$

$$Bel(\neg A) = \sum_{s_i \in \neg A} m(\{s_i\}) = 1 - \sum_{s_i \in A} m(\{s_i\}) - m(\Omega) = 1 - Bel(A) - m(\Omega)$$

$$|\neg A| = |\Omega| - |A|$$

$$Pl(\neg A) - Bel(\neg A) = m(\Omega)$$

故

$$f(\neg A) = 1 - Bel(A) - m(\Omega) + \frac{|\Omega| - |A|}{|\Omega|} \times m(\Omega)$$

$$= 1 - Bel(A) - m(\Omega) + m(\Omega) - \frac{|A|}{|\Omega|} \times m(\Omega)$$

$$= 1 - [Bel(A) + \frac{|A|}{|\Omega|} \times m(\Omega)] = 1 - f(A)$$

3.4.1 证据理论的形式化描述

3. 类概率函数

根据以上性质，可得如下推论

(1) f(Φ)=0

(2) f(Ω)=1

(3) 对任何 $A \subseteq \Omega$ ，有0≤f(A)≤1

例3.7 设Ω={红，黄，白}，概率分配函数
m({}，{红}，{黄}，{白}，{红，黄，白})=(0，0.6，0.2，0.1，0.1)
若A={红，黄}，求f(A)的值。

解：

$$f(A) = Bel(A) + \frac{|A|}{|\Omega|} \times [Pl(A) - Bel(A)]$$

$$= m(\{红\}) + m(\{黄\}) + \frac{2}{3} \times m(\{红，黄，白\})$$

$$= 0.6 + 0.2 + \frac{2}{3} \times 0.1 = 0.87$$

3.4.2 证据理论的推理模型

1. 知识不确定性的表示

表示形式：

IF E THEN $H=\{h_1, h_2, \ldots, h_n\}$ $CF=\{c_1, c_2, \ldots, c_n\}$

其中：E为前提条件，它既可以是简单条件，也可以是用合取或析取词连接起来的复合条件；

H是结论，它用样本空间中的子集表示，$h_1, h_2, \ldots, h_n$ 是该子集中的元素；

CF是可信度因子，用集合形式表示。该集合中的元素$c_1, c_2, \ldots, c_n$ 用来指出$h_1, h_2, \ldots, h_n$ 的可信度，c_i与h_i一一对应。

并且，c_i应满足如下条件：

$$c_i \geq 0 \qquad i = 1,2,\ldots,n$$

$$\sum_{i=1}^{n} c_i \leq 1$$

3.4.2 证据理论的推理模型

2. 证据不确定性的表示

不确定性证据E的确定性用CER(A)表示，CER(A) ∈[0, 1]

在实际中，如果是初始证据，其确定性由用户给出；如果是中间证据，其确定性由推理得到。

3.4.2 证据理论的推理模型

3. 组合证据不确定性的表示

当组合证据是多个证据的合取时

$E=E_1\ \text{AND}\ E_2\ \text{AND}\ \ldots\ \text{AND}\ E_n$

则

$CER(E)=\min\{CER(E_1), CER(E_2), \ldots ,CER(E_n)\}$

当组合证据是多个证据的析取时

$E=E_1\ \text{OR}\ E_2\ \text{OR}\ \ldots\ \text{OR}\ E_n$

则

$CER(E)=\max\{CER(E_1), CER(E_2), \ldots ,CER(E_n)$

3.4.2 证据理论的推理模型

4. 不确定性的更新

设有知识 IF E THEN $H=\{h_1, h_2, \ldots, h_n\}$ $CF=\{c_1, c_2, \ldots, c_n\}$

则求结论H的确定性CER(H)的方法如下：

(1)求H的概率分配函数

$$m(\{h_1\},\{h_2\},\ldots,\{h_n\}) = (CER(E)\times c_1, CER(E)\times c_2, \cdots, CER(E)\times c_n)$$

$$m(\Omega) = 1 - \sum_{i=1}^{n} CRE(E)\times c_i$$

如果有两条或多条知识支持同一结论H，例：

IF E_1 THEN $H=\{h_1, h_2, \ldots, h_n\}$ $CF=\{c_{11}, c_{12}, \ldots, c_{1n}\}$

IF E_2 THEN $H=\{h_1, h_2, \ldots, h_n\}$ $CF=\{c_{21}, c_{22}, \ldots, c_{2n}\}$

则按正交和求CER(H)，即先求出：

$m_1=m_1(\{h_1\},\{h_2\},\ldots,\{h_n\})$

$m_2=m_2(\{h_1\},\{h_2\},\ldots,\{h_n\})$

然后再用公式 $m = m_1 \oplus m_2$ 求m_1和m_2的正交和，最后求得H的m。

3.4.2 证据理论的推理模型

4. 不确定性的更新

(2)求Bel(H)、Pl(H)及f(H)

$$Bel(H)=\sum_{i=1}^{n}m(\{h_i\})$$

$$Pl(H)=1-Bel(\neg H)$$

$$f(H)=Bel(H)+\frac{|H|}{|\Omega|}\times[Pl(H)-Bel(H)]=Bel(H)+\frac{|H|}{|\Omega|}\times m(\Omega)$$

3.4.2 证据理论的推理模型

4. 不确定性的更新

(3)求H的确定性CER(H)

按公式

CER(H)=MD(H|E') ×f(H)

计算结论H确定性。其中MD(H|E') 为知识的前提条件与外部输入的证据E'的匹配度

$$MD(H \mid E') = \begin{cases} 1 & \text{如果H所要求的证据都已出现在E'中} \\ 0 & \text{否则} \end{cases}$$

3.4.3 推理实例

例3.8 设有如下规则：

r_1：IF E_1 AND E_2 THEN $A=\{a_1, a_2\}$ $CF=\{0.3, 0.5\}$

r_2：IF E_3 THEN $H=\{h_1, h_2\}$ $CF=\{0.4, 0.2\}$

r_3：IF A THEN $H=\{h_1, h_2\}$ $CF=\{0.1, 0.5\}$

已知用户对初始证据给出的确定性为：

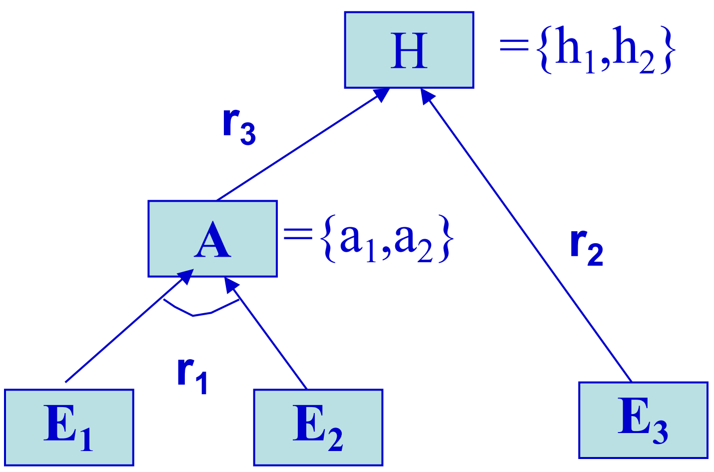
$CER(E_1)=0.8$ $CER(E_2)=0.6$

$CER(E_3)=0.9$

并假定Ω中的元素个数|Ω|=10

求：CER(H)=?

解：由给定知识形成的推理网络如右图所示：



3.4.3 推理实例

(1) 求CER(A)

由r_1:

$CER(E_1\ \ AND\ \ E_2)$

$=\min\{CER(E_1),\ CER(E_2)\}$

$=\min\{0.8,\ 0.6\} = 0.6$

$m(\{a_1\}, \{a_2\})=\{0.6\times0.3, 0.6\times0.5\} = \{0.18, 0.3\}$

$Bel(A)=m(\{a_1\})+m(\{a_2\})=0.18+0.3=0.48$

$Pl(A)=1-Bel(\neg A)=1-0=1$　　（隐含假设A所在的样本空间中其它元素的概率分配函数都为0）

$f(A)=Bel(A)+|A|/|\Omega|\bullet[Pl(A)-Bel(A)]$

$=0.48+2/10*[1-0.48]$

$=0.584$

故　$CER(A)=MD(A|E')\times f(A)=0.584$

3.4.3 推理实例

(2) 求CER(H)

由r_2得

$m_1(\{h_1\}, \{h_2\})=\{CER(E_3)\times 0.4, CER(E_3)\times 0.2\}$

$=\{0.9\times 0.4, 0.9\times 0.2\}$

$=\{0.36, 0.18\}$

$m_1(\Omega)=1-[m_1(\{h_1\})+m_1(\{h_2\})]$

$=1-[0.36+0.18]=0.46$

由r_3得

$m_2(\{h_1\}, \{h_2\})=\{CER(A)\times 0.1, CER(A)\times 0.5\}$

$=\{0.58\times 0.1, 0.58\times 0.5\}$

$=\{0.06, 0.29\}$

$m_2(\Omega)=1-[m_2(\{h_1\})+m_2(\{h_2\})]$

$=1-[0.06+0.29]=0.65$

3.4.3 推理实例

求正交和$m=m_1 \oplus m_2$

$$
\begin{aligned}
K=&m_1(\Omega)\times m_2(\Omega)\\
&+m_1(\{h_1\})\times m_2(\{h_1\})+m_1(\{h_1\})\times m_2(\Omega)+m_1(\Omega)\times m_2(\{h_1\})\\
&+m_1(\{h_2\})\times m_2(\{h_2\})+m_1(\{h_2\})\times m_2(\Omega)+m_1(\Omega)\times m_2(\{h_2\})\\
=&0.46\times 0.65\\
&+0.36\times 0.06+0.36\times 0.65+0.46\times 0.06\\
&+0.18\times 0.29+0.18\times 0.65+0.46\times 0.29\\
=&0.30+(0.02+0.23+0.03)+(0.05+0.12+0.13)\\
=&0.88
\end{aligned}
$$

$$
\begin{aligned}
m(h_1)&=\frac{1}{K}\bullet[m_1(\{h_1\})\bullet m_2(\{h_1\})+m_1(\{h_1\})\bullet m_2(\Omega)+m_1(\Omega)\bullet m_2(\{h_1\})]\\
&=\frac{1}{0.88}\times[0.36\times 0.06+0.36\times 0.65+0.46\times 0.06]\\
&=0.32
\end{aligned}
$$

3.4.3 推理实例

同理可得：

$$m(h_2)=\frac{1}{K}\times[m_1(\{h_2\})\times m_2(\{h_2\})+m_1(\{h_2\})\times m_2(\Omega)+m_1(\Omega)\times m_2(\{h_2\})]$$

$$=\frac{1}{0.88}\times[0.18\times 0.29+0.18\times 0.65+0.46\times 0.29]=0.34$$

故有： $\mathbf{m(\Omega)=1-[m(\{h_1\})+m(\{h_2\})]}$

$\mathbf{=1-[0.32+0.34] = 0.34}$

再根据m可得

$\mathbf{Bel(H)=m(\{h_1\})+m(\{h_2\}) = 0.32+0.34 = 0.66}$

$\mathbf{Pl(H)=m(\Omega)+Bel(H)=0.34+0.66=1}$

$$f(H)=Bel(H)+\frac{|H|}{|\Omega|}\bullet[Pl(H)-Bel(H)]=0.66+\frac{2}{10}\times(1-0.66)=0.73$$

CER(H)=MD(H|E')×f(H)=0.73

应用：推理

设有规则

(r1) IF 流鼻涕 THEN 感冒但非过敏性鼻炎 (0.9) 或 过敏性鼻炎但非感冒 (0.1)

(r2) IF 眼发炎 THEN 感冒但非过敏性鼻炎 (0.8) 或 过敏性鼻炎但非感冒 (0.05)

有事实

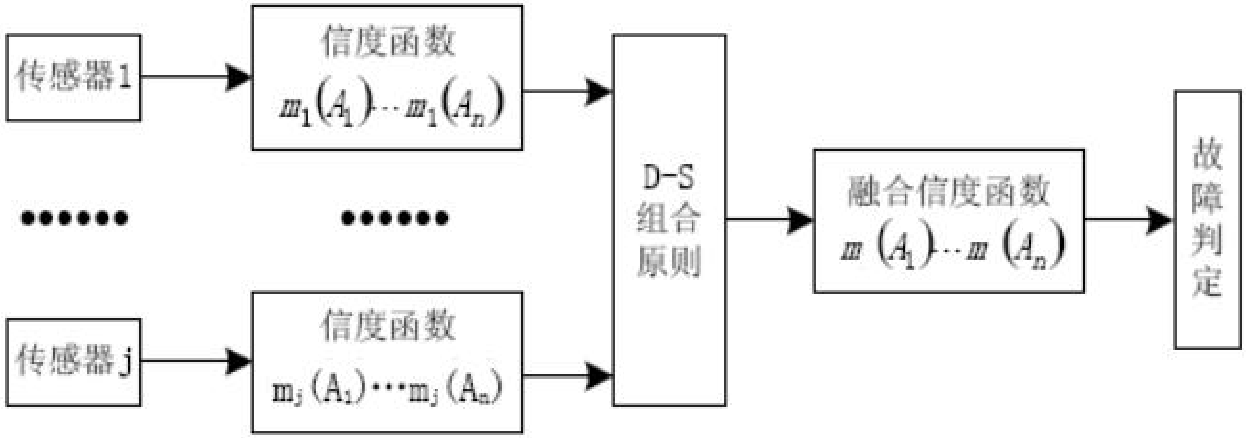
小王流鼻涕 (0.9)

小王眼发炎 (0.4)

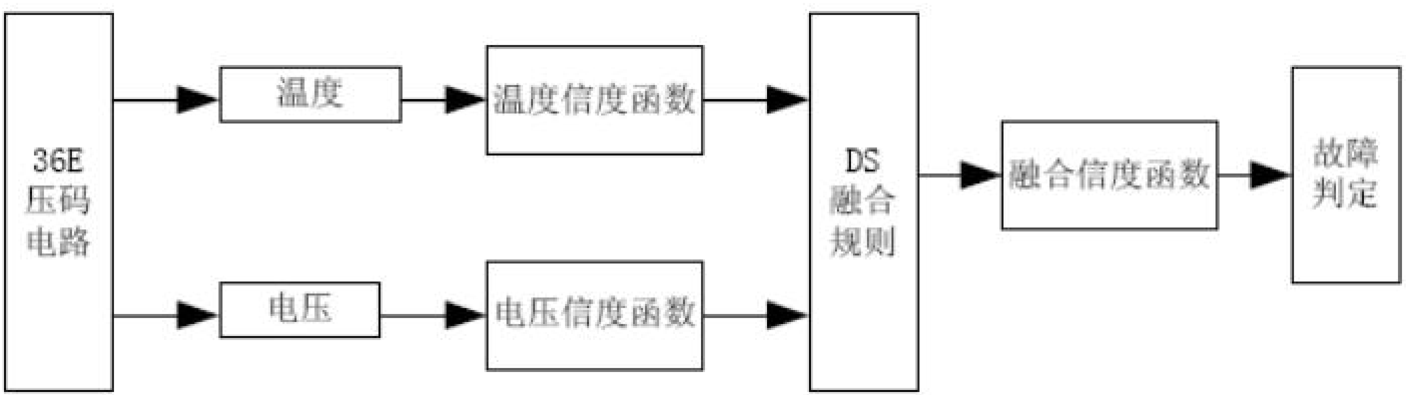
试诊断小王患的什么病？

计算 [Bel({h1}), Pl({h1})] & [Bel({h2}), Pl({h2})] | f(h1) & f(h2)

应用：信息融合



图中$m_j(A_1)\cdots\cdots m_j(A_n)$，表示传感器j测得的症状属于故障$A_i\cdots\cdots An$的信度函数$m(A_1),m(A_2)\cdots m(A_n)$是传感器融合后分配到各故障模式上的信度函数值。



3.4.4 证据理论推理的特性

主要优点：

- 能满足比概率更弱的公理系统，能处理由“不知道”所引起的不确定性；
- 由于Ω的子集可以是多个元素的集合，因而知识的结论部分可以是更一般的假设，便于领域专家从不同细节、不同层次上表达知识。

主要缺点：

- 要求Ω中元素满足互斥条件，实际中不易做到；
- 需要给出的概率分配函数多；
- 证据合成规则没有非常坚固的理论支撑；
- 计算比较复杂。

很多变形、很多改进、很多应用，现在仍是研究热点

第3章 不确定性知识系统

3.1 不确定性推理概述

3.2 可信度推理

3.3 主观Bayes方法

3.4 证据理论

3.5 模糊推理

3.6 概率推理

引言

大脑具有模糊思维功能

- **Russell：“传统逻辑都习惯于假定使用的是精确符号。因此它不适合于尘世生活，而仅仅适用于想象的天体存在物…。逻辑学较别的学科使我们更接近于天堂”**
- **Wiener：同计算机相比，人脑主要的一个优越性似乎是“能够掌握尚未完全明确的含糊概念”。**
- **精确方法的逻辑基础是传统的二值逻辑，要求命题要么真、要么假。在处理模糊概念和模糊命题时，将会在逻辑上产生悖论。**

- “秃头悖论”

首先，我们都同意以下两条公设

（1）存在秃头的人和非秃头的人

（2）若有n根头发的人秃，则有n+1根头发的人亦秃

由此便可导致

秃头悖论 所有人都秃

证明：用数学归纳法

(i) $n=0$的人显然是秃头

(ii) 假定 $n=k$的人是秃头

(iii)由公设(2), $n=k+1$的人也是秃头。

于是由数学归纳法知，对任意$n\geq0$，有n根头发的人都是秃头。从而所有人都是秃头。

- 年龄悖论
- 身高悖论
- 饥饱悖论
- ……

➢ 量与质是统一的，量的变化包含着质的变化。从头发根数来区分秃与不秃，其绝对的界限是没有的。但根数的加1减1又必须计较，在这微小的量变之中蕴含着质的差别，而这种差别只简单用“是” 与“非”这两个字是绝对不能刻画出来的。

➢ 需要刻画此类模糊现象

引言

1. 数学模型分类：

一、确定性数学模型

二、随机性数学模型

集合确定

三、模糊性数学模型

集合不确定

2. 随机性与模糊性的区别：

随机性：事件发生的可能性。

模糊性：概念外延不确定。

模糊数学产生、发展及现状

1965年，美国UC Berkeley大学控制论专家L.A. Zadeh(扎德)教授在《Information and Control》发表第一篇模糊数学论文《Fuzzy Sets》揭开了模糊数学诞生的序幕。Zadeh开创性提出了模糊集概念及其研究方法。


L.A. Zadeh

扎德，美国工程科学院院士。1921年2月生于前苏联巴库。1949年获哥伦比亚大学电机工程博士。后任伯克利加利福尼亚大学电机工程与计算机科学系教授。因发展模糊集理论的先驱性工作而获电气与电子工程师学会(IEEE)的教育勋章。

扎德在控制理论方面有重要贡献。1949年他在关于时变网络频率分析的博士论文中引入的时变变换函数的概念，后来成为线性时变系统分析的工具。1950～1952年他和J.拉加齐尼合作，推广了维纳预测理论，在设计有限存储滤波器和预测器中得到广泛应用。他们发展的采样控制系统的Z变换逼近，成为分析这类系统的标准方法。1953年他给出一种设计非线性滤波器的新的逼近方法。1963年他和C.A.德舍尔合著的《线性系统的状态空间理论》是该领域的经典著作。书中介绍的状态空间逼近已成为最优控制中的标准工具，广泛用于工业机器人和社会经济系统。

2017.9.6去世。

INFORMATION AND CONTROL 8, 338–353 (1965)

Fuzzy Sets*

L. A. ZADEH

Department of Electrical Engineering and Electronics Research Laboratory, University of California, Berkeley, California

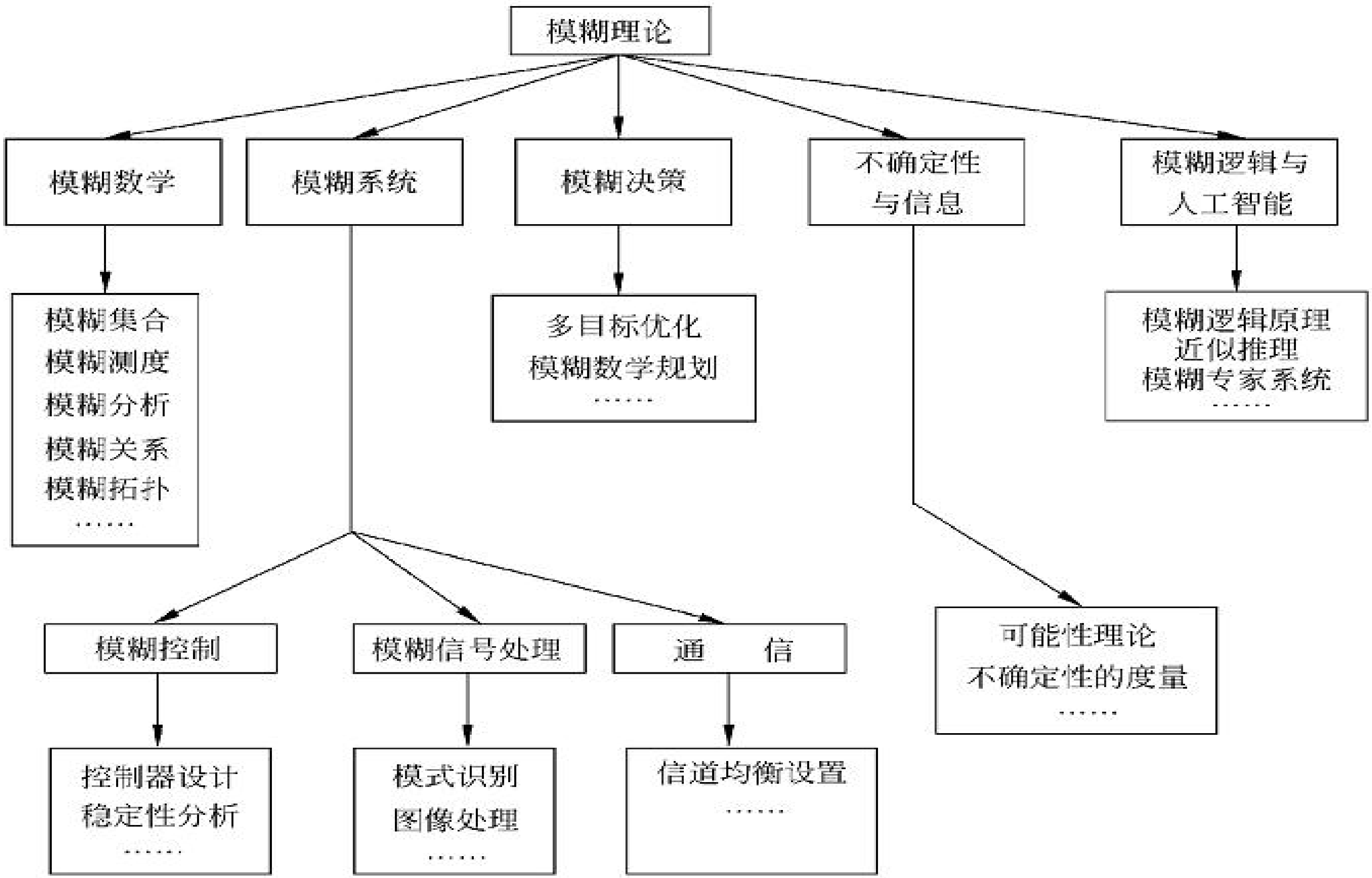
A fuzzy set is a class of objects with a continuum of grades of membership. Such a set is characterized by a membership (characteristic) function which assigns to each object a grade of membership ranging between zero and one. The notions of inclusion, union, intersection, complement, relation, convexity, etc., are extended to such sets, and various properties of these notions in the context of fuzzy sets are established. In particular, a separation theorem for convex fuzzy sets is proved without requiring that the fuzzy sets be disjoint.

I. INTRODUCTION

- **1973年，扎德教授又提出模糊逻辑(Fuzzy Logic)的理论，并积极倡导将模糊理论向人工智能方向发展。模糊逻辑的研究虽然时间还不长，但在智能模拟和智能控制等领域却已有了飞快的发展。**
- **1974年，印度裔英国学者马德尼(E.H.Mamdani)首先将模糊理论用于锅炉和蒸汽机的控制，并实验成功，开创了模糊控制的新领域。**
- **80年代后期以来，在日本采用模糊控制技术的家电产品大量上市，模糊技术在图像识别、自动控制、市场预测、人工智能等领域普遍应用，掀起了一股模糊热。日本、美国和我国都成功地研制出了智能化的模糊推理机，这表现了模糊理论的强大生命力和伟大意义。**

- 另一方面，模糊理论在学术界也得到了普遍的认同和重视。1992年，IEEE召开了第一届关于模糊系统　的国际会议（FUZZ—IEEE），并决定以后每年举行一次。1993年IEEE创办了专刊IEEE Transaction on Fuzzy System。
- 当前，模糊理论和应用正向深度和广度进一步发展，发展的速度越来越快，研究成果大量涌现，已经成为世界各国高科技竞争的重要领域之一。

模糊技术的研究热点



模糊理论
模糊数学
模糊系统
模糊决策
不确定性
与信息
模糊逻辑与
人工智能
模糊集合
模糊测度
模糊分析
模糊关系
模糊拓扑
……
多目标优化
模糊数学规划
……
模糊逻辑原理
近似推理
模糊专家系统
……
模糊控制
模糊信号处理
通 信
可能性理论
不确定性的度量
……
控制器设计
稳定性分析
……
模式识别
图像处理
……
信道均衡设置
……

学术期刊

- 模糊数学与系统
- Fuzzy Sets and Systems
- IEEE Trans. Fuzzy Systems
- IEEE Trans. Cybernetics
- Information Sciences

3.5.1 模糊集及其运算

1. 模糊集的定义

普通集合只能表示清晰概念

$$\forall u \in U, A \subset U \Rightarrow u \in A \text{或} u \overline{\in} A$$

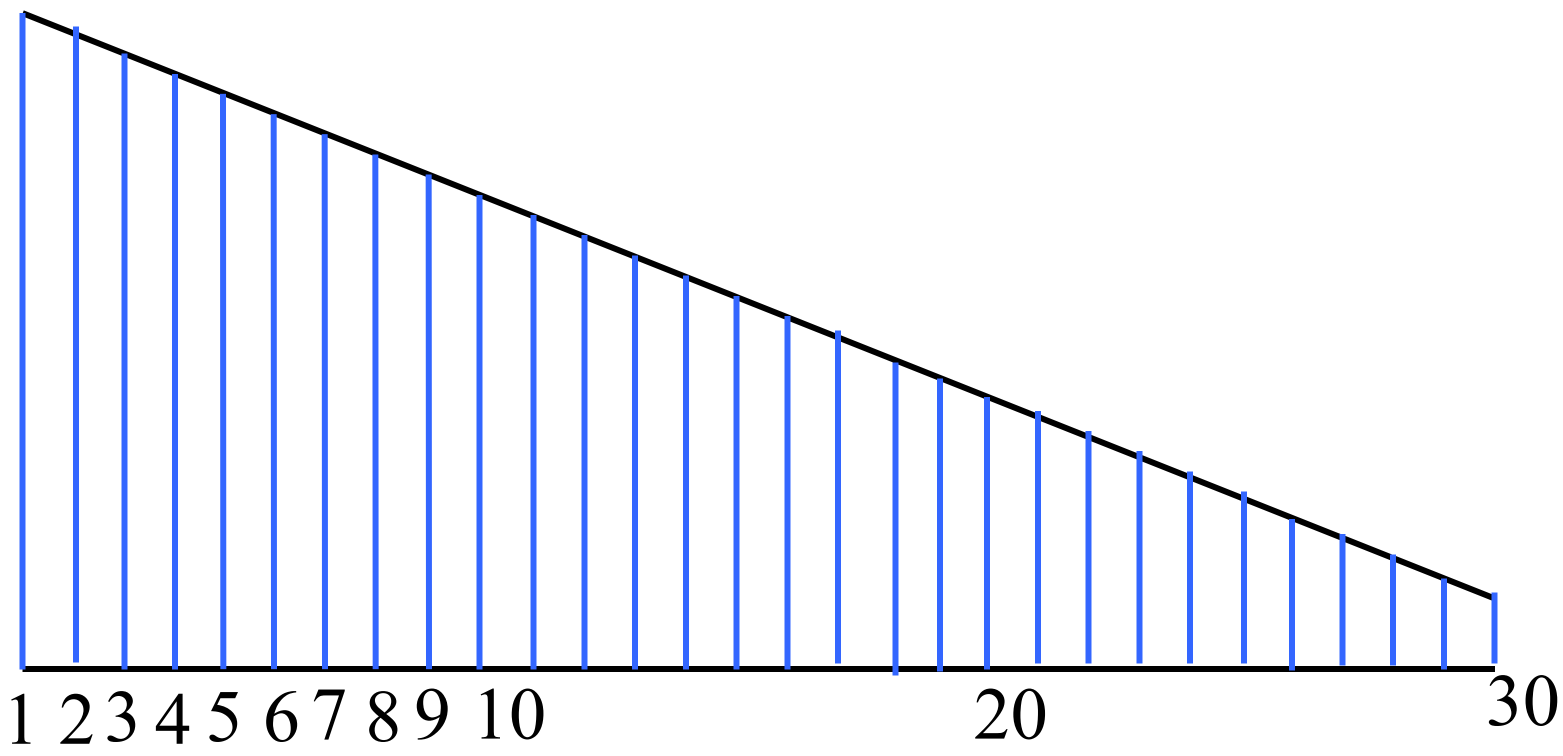
子集A由映射 $C_A: \quad U \to \{0,1\}$

因此特征函数 $C_A(u) = \begin{cases} 1 & u \in A \\ 0 & u \overline{\in} A \end{cases}$

只能表达“非此即彼”的清晰概念（现象）。
不能表达“亦此亦彼”的模糊概念（现象）。

例 从图的30条线段中，选出“长的线段”。

从左起，第1条是属于“长线段”，第2，3...条越靠右的线段作为“长线段”的**资格(membership)**就越降低，第30条线段根本不能作为“长线段”的成员。应属于“短线段”



3.5.1 模糊集及其运算

1. 模糊集的定义

定义3.8 设U是给定论域，$\mu_F(u)$是把任意u∈U映射为[0, 1]上某个实值的函数，即

$$\mu_F(u): \mathrm{U} \to [0, 1]$$

$$\mathrm{u} \to \mu_F(u)$$

称$\mu_F(u)$为定义在U上的一个隶属函数(membership function)，由$\mu_F(u)$（对所有 u∈U ）所构成的集合

$$F=\{\ \mu_F(u) \mid u \in U\ \}$$

则称F为U上的一个模糊集，$\mu_F(u)$称为u对F的隶属度。

说明：

① 模糊集F由隶属函数μ_F来刻画的，μ_F把U中的每一个元素u都映射为[0, 1]上的一个值 $\mu_F(u)$ 。

② $\mu_F(u)$的值表示u隶属于F的程度，其值越大，表示u隶属于F的程度越高。当$\mu_F(u)$仅取0和1时，模糊集F便退化为一个普通集合。

3.5.1 模糊集及其运算

1. 模糊集的定义

例3.9 设论域U={20, 30, 40, 50, 60}给出的是年龄，请确定一个刻画模糊概念“年轻”的模糊集F。

解：由于模糊集是用其隶属函数来刻画的，因此需要先求出描述模糊概念“年轻”的隶属函数。假设对论域U中的元素，其隶属函数值分别为：

$$\mu_F(20)=1, \mu_F(30)=0.8, \mu_F(40)=0.4,$$
$$\mu_F(50)=0.1, \mu_F(60)=0$$

则可得到刻画模糊概念“年轻”的模糊集

F={ 1, 0.8, 0.4, 0.1, 0}

说明其含义。

模糊隶属度函数确定方法：
模糊统计、专家打分、三分法等。带有主观色彩，但要符合实际

3.5.1 模糊集及其运算

2. 模糊集的表示

(1) 离散且为有限论域的表示方法

设论域 $U=\{u_1, u_2, \ldots, u_n\}$为离散论域，则其模糊集可表示为：

$$F=\{ \mu_F(u_1), \mu_F(u_2), \ldots, \mu_F(u_n) \}$$

为了能够表示出论域中的元素与其隶属度之间的对应关系，扎德引入了一种模糊集的表示方式：先为论域中的每个元素都标上其隶属度，然后再用“+”号把它们连接起来，即

$$F= \mu_F(u_1)/u_1+ \mu_F(u_2)/u_2 +\ldots + \mu_F(u_n)/u_n$$

也可写成

$$F=\sum_{i=1}^{n}\frac{\mu_F(u_i)}{u_i}$$

其中，$\mu_F(u_i)$为u_i对F的隶属度；“$\mu_F(u_i)/ u_i$”不是相除关系，只是一个记号；“+”也不是算术意义上的加，只是一个连接符号。

3.5.1 模糊集及其运算

2. 模糊集的表示

在上述表示方法中，当某个u_i对F的隶属度=0时，可省略不写。例如，前面例3.9的模糊集F可表示为：

F= 1/20+ 0.8/30+ 0.4/40+ 0.1/50

有时，模糊集也可写成如下两种形式：

单点形式

$$F=\{\mu_F(u_1)/u_1, \mu_F(u_2)/u_2, \cdots, \mu_F(u_n)/u_n\}$$

序偶形式

$$F=\{(\mu_F(u_1),u_1),(\mu_F(u_2),u_2),\cdots,(\mu_F(u_n),u_n)\}$$

3.5.1 模糊集及其运算

2. 模糊集的表示

(2) 连续论域的表示方法

如果论域是连续的，则其模糊集可用一个实函数来表示。例如，扎德以年龄为论域，取U=[0, 100]，给出了“年轻”与“年老”这两个模糊概念的隶属函数

$$\mu_{年老}(u)=\begin{cases}0 & 当0\le u\le 50\\ [1+(\frac{5}{u-50})^2]^{-1} & 当50< \mathrm{u}\le 100\end{cases}$$

$$\mu_{年轻}(u)=\begin{cases}1 & 当0\le u\le 25\\ [1+(\frac{u-25}{5})^2]^{-1} & 当25< \mathrm{u}\le 100\end{cases}$$

(3) 一般表示方法

不管论域U是有限的还是无限的，是连续的还是离散的，扎德又给出了一种类似于积分的一般表示形式：

$$F=\int_{u\in U}\mu_F(u)/u$$

这里的记号不是数学中的积分符号，也不是求和，只是表示论域中各元素与其隶属度对应关系的总括。

3.5.1 模糊集及其运算

3. 模糊集的运算

定义3.9 设F、G分别是U 上的两个模糊集，对任意u∈U，都有

$$\mu_F(u) = \mu_G(u)$$

成立，则称F等于G，记为F=G。

定义3.10 设F、G分别是U上的两个模糊集，对任意u∈U，都有$\mu_F(u) \le \mu_G(u)$成立，则称F含于G，记为F ⊆ G。

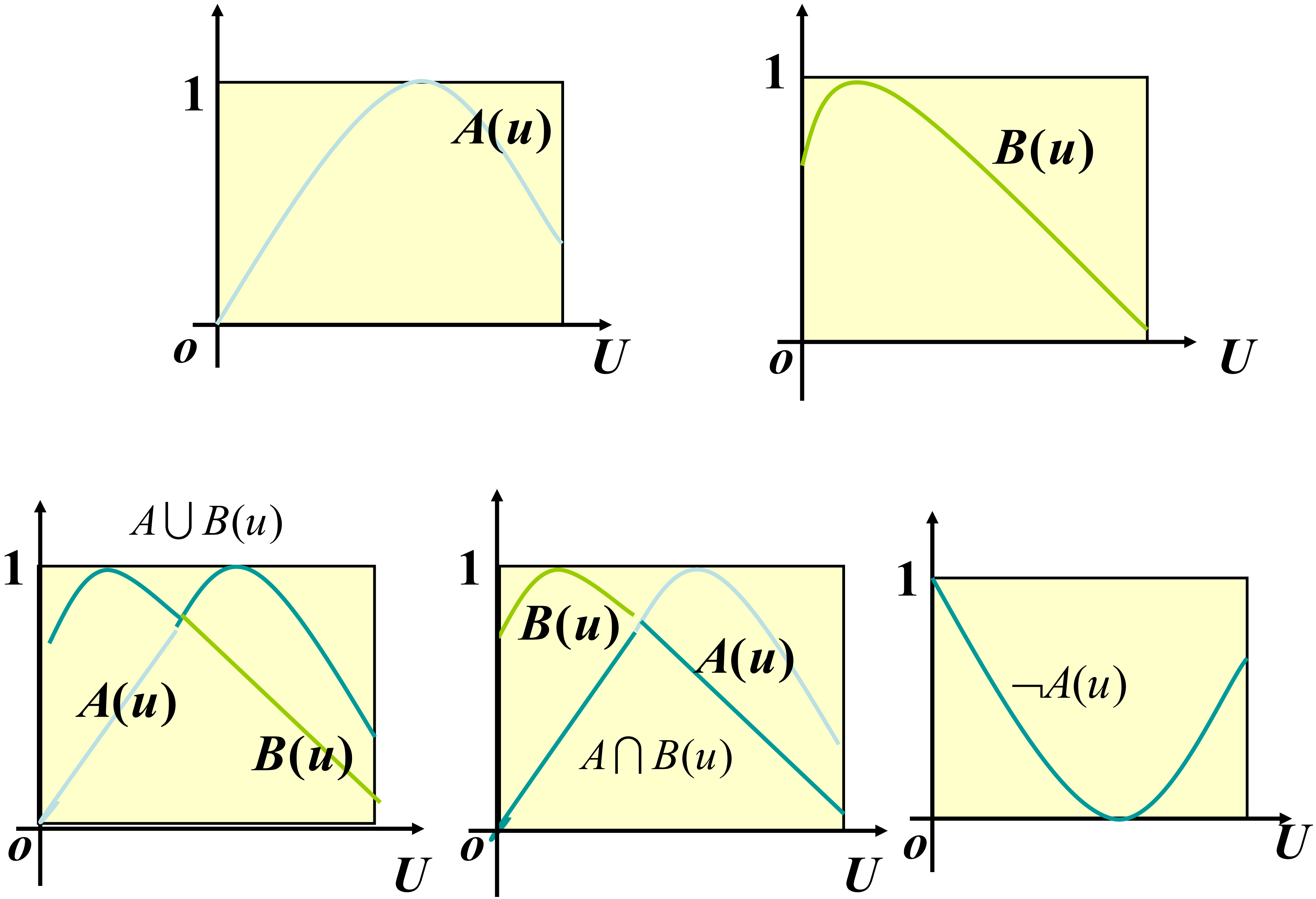
定义3.11 设F、G分别是U上的两个模糊集，则F∪G、F∩G分别称为F与G的并集、交集，它们的隶属函数分别为：

$$F \bigcup G : \mu_{F \cup G}(u) = \max_{u \in U}\{\mu_F(u), \mu_G(u)\}$$

$$F \bigcap G : \mu_{F \cap G}(u) = \min_{u \in U}\{\mu_F(u), \mu_G(u)\}$$

定义3.12 设F为U上的模糊集，称¬F为F的补集，其隶属函数为：

$$\neg F : \mu_{\neg F}(u) = 1 - \mu_F(u)$$



1
A(u)
o
U
1
B(u)
o
U
A∪B(u)
1
A(u)
B(u)
o
U
1
B(u)
A(u)
A∩B(u)
o
U
1
¬A(u)
o
U

$(F(U),\cup,\cap,\neg)$不满足互补律，即

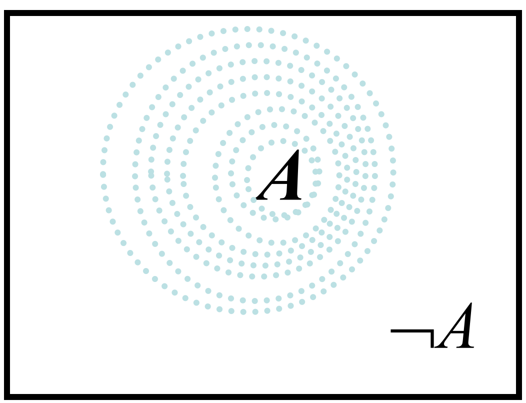
$$A\cup\neg A\neq U,\qquad A\cap\neg A\neq\Phi$$

说明$A\cup\neg A$不能覆盖U, A与$\neg A$相互交迭。但是有

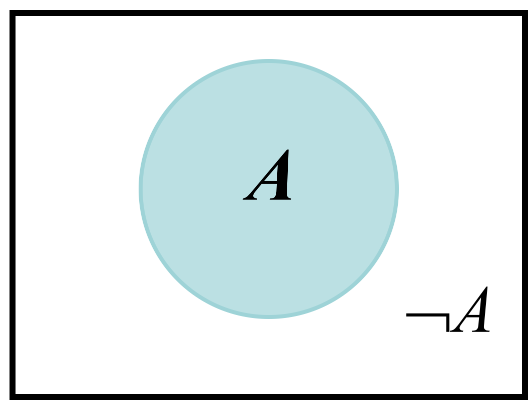
$$\forall A\in F(U), A(u)\cup\neg A(u)\geq\frac{1}{2}(\forall u\in U)$$

$$\forall A\in F(U), A(u)\cap\neg A(u)\leq\frac{1}{2}(\forall u\in U)$$

其原因是F集没有明确的边界。



$(F(U),\cup,\cap,\neg)$



$(P(U),\cup,\cap,\neg)$

3.5.1 模糊集及其运算

3. 模糊集的运算

例3.10 设U={1,2,3}，F和G分别是U上的两个模糊集（小，大），且

$F=1/1+0.6/2+0.1/3$

$G=0.1/1+0.6/2+1/3$

则 $F\cup G=(1\vee 0.1)/1+(0.6\vee 0.6)/2+(0.1\vee 1)/3=1/1+0.6/2+1/3$

$F\cap G=(1\wedge 0.1)/1+(0.6\wedge 0.6)/2+(0.1\wedge 1)/3=0.1/1+0.6/2+0.1/3$

$\neg F=(1-1)/1+(1-0.6)/2+(1-0.1)/3=0.4/2+0.9/3$

从这个例子可以看出，两个模糊集之间的运算实际上就是逐点对隶属函数作相应的运算。

F集的截集

F集是由隶属函数确定的。而其中包含哪些元素无法确定。即F集的边界是模糊的。但在实际问题中对于模糊现象常常要做出不模糊的判决。因此，需要把F集和普通集联系起来。这个桥梁就是 $\lambda-$截集。

例1 在一次“优胜者”的选拔考试中，10位应试者及其成绩如下表

应试者	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
成绩	100	92	35	68	82	25	74	80	40	55

按“择优录取”原则挑选。

设F集A表示“优胜者”，有

$$A=\frac{1}{x_1}+\frac{0.92}{x_2}+\frac{0.35}{x_3}+\frac{0.68}{x_4}+\frac{0.82}{x_5}+\frac{0.25}{x_6}+$$

$$\frac{0.74}{x_7}+\frac{0.80}{x_8}+\frac{0.4}{x_9}+\frac{0.55}{x_{10}}$$

择优录取实际上就是将F集转化为普通集。

先确定阈值 $\lambda(0\le\lambda\le 1)$,再将隶属度 $A(x_i)\ge\lambda$的元素挑选出来。因此，当 $\lambda=0.7,\ 0.9$时有

$$A_{0.7}=\{x_1,x_2,x_5,x_7,x_8\}$$

$$A_{0.9}=\{x_1,x_2\}$$

一般情况下，给出A_λ定义如下：

定义1　设$A \in F(U), \lambda \in [0,1]$,记

$$\text{(1)} \qquad A_{\lambda} = \{u \mid u \in U, A(u) \geq \lambda\}$$

称A_{λ}为A的一个$\lambda-$截集,λ称为阈值（或置信水平　）。

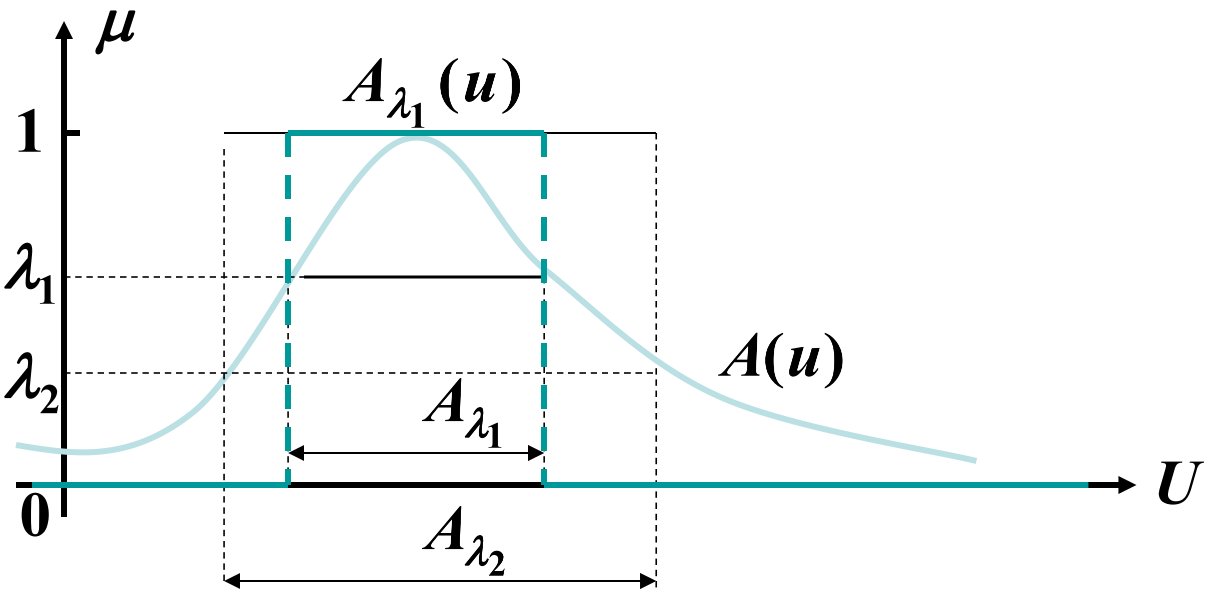
$$\text{(2)} \qquad A_{\underset{\bullet}{\lambda}} = \{u \mid u \in U, A(u) > \lambda\}$$

称$A_{\underset{\bullet}{\lambda}}$为$A$的一个$\lambda-$强截集。

由定义知A_{λ}是一个普通集。

$\forall u \in U$，当$A(u) \geq \lambda$时，$u \in A_{\lambda}$. 当$A(u) < \lambda$时，$u \notin A_{\lambda}$.

每取定一个λ值，由F集A得到一个普通集A_λ.当λ值从1减少到0时，A_λ逐渐向外扩展。因此，F集A对应着一族集合 A_λ.λ的增减引起A_λ的收缩和膨胀。F集A是一个具有游移边界的 集合。



分解定理

（分解定理I）　设$A \in F(U)$,则

$$A = \bigcup_{\lambda \in [0,1]} (\lambda A_\lambda)$$

例如,设F集$A=\frac{0.5}{u_1}+\frac{0.6}{u_2}+\frac{1}{u_3}+\frac{0.7}{u_4}+\frac{0.3}{u_5}$,

取$\lambda-$截集,我们得到:

$$A_1=\{u_3\}$$

$$A_{0.7}=\{u_3,u_4\}$$

$$A_{0.6}=\{u_2,u_3,u_4\}$$

$$A_{0.5}=\{u_1,u_2,u_3,u_4\}$$

$$A_{0.3}=\{u_1,u_2,u_3,u_4,u_5\}$$

将$\lambda-$截集写成F集的形式，例如

$$A_{0.7}=\frac{1}{u_3}+\frac{1}{u_4}$$

于是按数乘 F集的定义，得

$$1A_1 = \frac{1}{u_3}$$

$$0.7A_{0.7} = \frac{0.7}{u_3} + \frac{0.7}{u_4}$$

$$0.6A_{0.6} = \frac{0.6}{u_2} + \frac{0.6}{u_3} + \frac{0.6}{u_4}$$

$$0.5A_{0.5} = \frac{0.5}{u_1} + \frac{0.5}{u_2} + \frac{0.5}{u_3} + \frac{0.5}{u_4}$$

$$0.3A_{0.3} = \frac{0.3}{u_1} + \frac{0.3}{u_2} + \frac{0.3}{u_3} + \frac{0.3}{u_4} + \frac{0.3}{u_5}$$

应用分解定理I构成原来的F集

$$A=\bigcup_{\lambda\in[0,1]}\lambda A_\lambda=1A_1\cup 0.7A_{0.7}\cup 0.6A_{0.6}\cup 0.5A_{0.5}\cup 0.3A_{0.3}$$

$$=\frac{1}{u_3}\cup\left(\frac{0.7}{u_3}+\frac{0.7}{u_4}\right)\cup\left(\frac{0.6}{u_2}+\frac{0.6}{u_3}+\frac{0.6}{u_4}\right)\cup$$

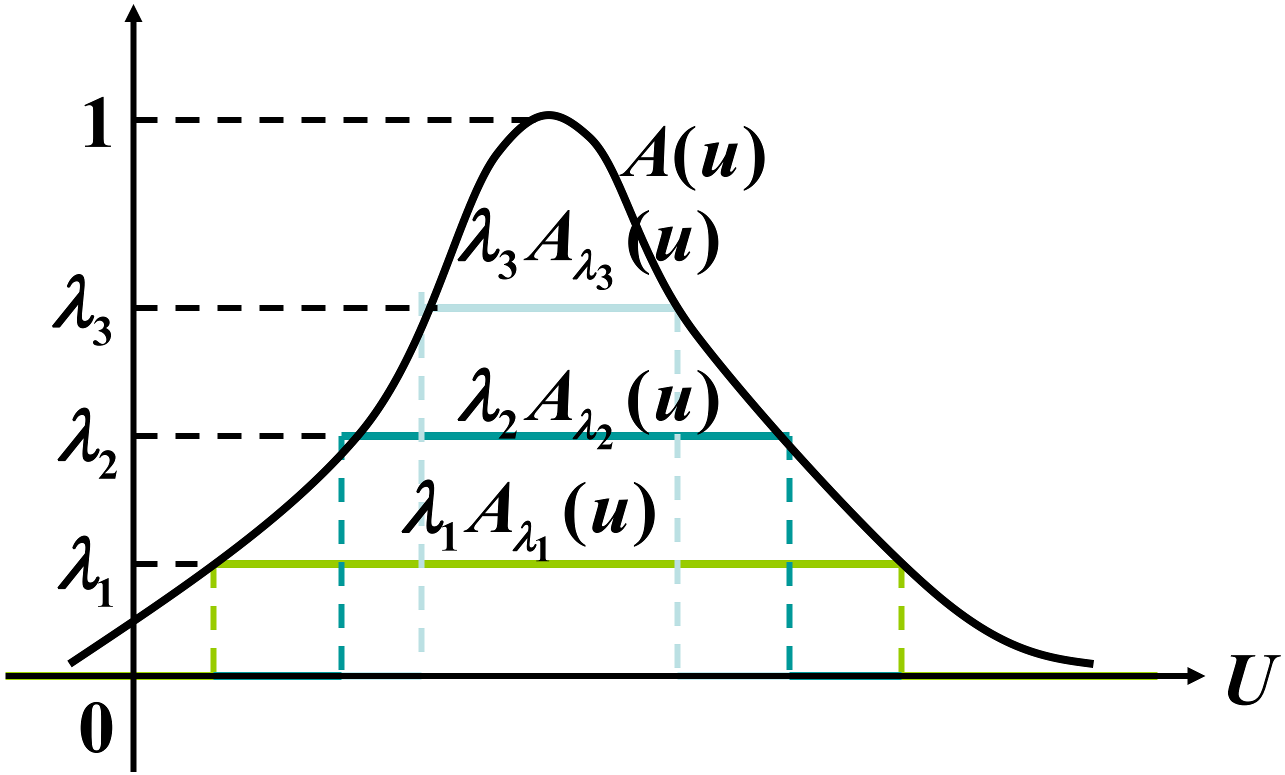
$$\left(\frac{0.5}{u_1}+\frac{0.5}{u_2}+\frac{0.5}{u_3}+\frac{0.5}{u_4}\right)\cup\left(\frac{0.3}{u_1}+\frac{0.3}{u_2}+\frac{0.3}{u_3}+\frac{0.3}{u_4}+\frac{0.3}{u_5}\right)$$

$$=\frac{0.3\vee 0.5}{u_1}+\frac{0.3\vee 0.5\vee 0.6}{u_2}+\frac{0.3\vee 0.5\vee 0.6\vee 0.7\vee 1}{u_3}+$$

$$\frac{0.3\vee 0.5\vee 0.6\vee 0.7}{u_4}+\frac{0.3}{u_5}$$

$$= \frac{0.5}{u_1} + \frac{0.6}{u_2} + \frac{1}{u_3} + \frac{0.7}{u_4} + \frac{0.3}{u_5}$$

分解定理的直观表示如图所示



3.5.2 模糊关系及其运算

1. 模糊关系的定义

设V与W是两个普通集合，V与W的笛卡尔乘积为

$V\times W=\{(v,w)|$ 任意 $v\in V$，任意 $w\in W\}$ 所有可能序偶(v,w)构成

所谓从V到W的关系R，是指V×W上的一个子集，即

$$R\subseteq V\times W$$

记为

$$V\xrightarrow{R}W$$

对于V×W中的元素(v,w)，若(v,w)∈R，则称v与w有关系R；若(v,w) $\notin$ R，则称v与w没有关系。

例3.11 设V={1班，2班，3班}，W={男队，女队}

则V×W中有6个元素，即

V×W ={(1班，男队)，(2班，男队)，(3班，男队)，(1班，女队)，(2班，女队)，(3班，女队)}

其中，每个元素是一代表队。假设要进行一种双方对垒的循环赛，则每一个赛局都是V×W中的一个子集，它构成了V×W上的一个关系。

3.5.2 模糊关系及其运算

1. 模糊关系的定义

定义3.13 设F_i是$U_i(i=1,2,\ldots,n)$上的模糊集，则称

$$F_1\times F_2\times\cdots\times F_n=\int_{u_1\times u_2\times\cdots\times u_n}(\mu_{F_1}(u_1)\wedge\mu_{F_2}(u_2)\wedge\cdots\wedge\mu_{F_n}(u_n))/(u_1,u_2,\cdots,u_n)$$

为$F_1,F_2,\ldots,F_n$的笛卡尔乘积，它是$U_1\times U_2\times\ldots\times U_n$上的一个模糊集。

定义3.14 在$U_1\times U_2\times\ldots\times U_n$上的一个n元模糊关系R是指以$U_1\times U_2\times\ldots\times U_n$为论域的一个模糊集，记为

$$R=\int_{U_1\times U_2\times\cdots\times U_n}\mu_R(u_1,u_2,\cdots,u_n)/(u_1,u_2,\cdots,u_n)$$

3.5.2 模糊关系及其运算

1. 模糊关系的定义

例3.12 设有一组学生 U={u_1,u_2}={秦学，郝玩}，一些在计算机上的活动V={v_1,v_2,v_3}={编程，上网，玩游戏}

并设每个学生对各种活动的爱好程度分别为 $\mu_F(u_i, v_j)$ i=1,2；j=1,2,3，即

$$\mu_R(\text{秦学},\text{编程})=0.9,\ \mu_R(\text{秦学},\text{上网})=0.4,\quad \mu_R(\text{秦学},\text{玩游戏})=0.1,$$
$$\mu_R(\text{郝玩},\text{编程})=0.2,\ \mu_R(\text{郝玩},\text{上网})=0.5,\quad \mu_R(\text{郝玩},\text{玩游戏})=0.8$$

则U×V上的模糊关系R为

$$R=\begin{bmatrix} 0.9 & 0.4 & 0.1 \\ 0.2 & 0.5 & 0.8 \end{bmatrix}$$

3.5.2 模糊关系及其运算

2. 模糊关系的合成

定义3.15 设R_1与R_2分别是U×V与V×W上的两个模糊关系，则R_1与R_2的合成是从U到W的一个模糊关系，记为 $R_1 o R_2$ 。其隶属函数为

$$\mu_{R_1 \circ R_2}(u,w) = \vee\{\mu_{R_1}(u,v) \wedge \mu_{R_2}(v,w)\}$$

其中，∧和∨分别表示取最小和取最大。

例3.13 设有以下两个模糊关系

$$R_1 = \begin{bmatrix} 0.4 & 0.5 & 0.6 \\ 0.8 & 0.3 & 0.7 \end{bmatrix} \qquad R_2 = \begin{bmatrix} 0.7 & 0.9 \\ 0.2 & 0.8 \\ 0.5 & 0.3 \end{bmatrix}$$

则R_1与R_2的合成是

$$R = R_1 \circ R_2 = \begin{bmatrix} 0.5 & 0.5 \\ 0.7 & 0.8 \end{bmatrix}$$

其方法是把R_1的第i行元素分别与R_2的第j列的对应元素相比较，两个数中取最小者，然后再在所得的一组最小数中取最大的一个，并以此数作为$R_1 o R_2$的元素R(i,j)。

3.5.2 模糊关系及其运算

3. 模糊变换

定义3.16 设$F=\{\mu_F(u_1),\mu_F(u_2),..,\mu_F(u_n)\}$是论域U上的模糊集，R是U×V上的模糊关系，则FoR=G 称为模糊变换。一般形式见p97

例3.14 设F=(1, 0.6, 0.2)

$$R=\begin{bmatrix} 1 & 0.5 & 0 & 0 \\ 0.5 & 1 & 0.5 & 0 \\ 0 & 0.5 & 1 & 0.5 \end{bmatrix}$$

则$G= FoR=\{(1\wedge 1)\vee(0.6\wedge 0.5)\vee(0.2\wedge 0),$

$(1\wedge 0.5)\vee(0.6\wedge 1)\vee(0.2\wedge 0.5),$

$(1\wedge 0)\vee(0.6\wedge 0.5)\vee(0.2\wedge 1),$

$(1\wedge 0)\vee(0.6\wedge 0)\vee(0.2\wedge 0.5)\}$

$=\{1, 0.6, 0.5, 0.2\}$

3.5.3 模糊知识表示

1. 模糊命题的描述

模糊命题是基于模糊逻辑，利用模糊谓词、模糊量词、模糊修饰语等来对模糊命题的模糊性进行描述的。

模糊谓词

设x∈U，F为模糊谓词，即U中的一个模糊关系，则模糊命题可表示为

x is F

其中的模糊谓词F可以是大、小、年轻、年老、冷、暖、长、短等。

模糊量词

模糊逻辑中使用的模糊量词，如极少、很少、几个、少数、多数、大多数、几乎所有等。这些模糊量词可以很方便地描述类似于下面的命题：

大多数成绩好的学生学习都很刻苦。

很少有成绩好的学生特别贪玩。

3.5.3 模糊知识表示

1. 模糊命题的描述

模糊修饰语

设m是模糊修饰语，x是变量，F为模糊谓词，则模糊命题可表示为 x is mF，模糊修饰语也称为程度词，常用的程度词有“很”、“非常”、“有些”、“绝对”等。

模糊修饰语的四种主要运算：

① 求补 表示否定，如“不”、“非”等，其隶属函数的表示为

$$\mu_{\text{非}F}(u) = 1 - \mu_F(u) \qquad u \in [0,1]$$

② 集中 表示“很”、“非常”等，其效果是减少隶属函数的值：

$$\mu_{\text{非常}F}(u) = \mu_F{}^2(u) \qquad u \in [0,1]$$

③ 扩张 表示“有些”、“稍微”等，其效果是增加隶属函数的值：

$$\mu_{\text{有些}F}(u) = \mu_F{}^{\frac{1}{2}}(u) \qquad u \in [0,1]$$

3.5.3 模糊知识表示

1. 模糊命题的描述

④ 加强对比 表示“明确”、“确定”等，其效果是增加0.5以上隶属函数的值，减少0.5以下隶属函数的值：

$$\mu_{\text{确实}\ F}(u)=\begin{cases}2\mu_F^{\ 2}(u) & \text{若 } 0\le \mu_F(u)\le 0.5\\ 1-2(1-\mu_F(u))^2 & \text{若 } 0.5<\mu_F(u)\le 1\end{cases}$$

在以上4种运算中，集中与扩张用的较多。例如，语言变量“真实性”取值“真”和“假”的隶属函数定义为：

$$\mu_{\text{真}}(u)=u \qquad u\in[0,1]$$

$$\mu_{\text{假}}(u)=1-u \qquad u\in[0,1]$$

则“非常真”、“有些真”、“非常假”、“有些假”可定义为

$$\mu_{\text{非常真}}(u)=u^2 \qquad u\in[0,1] \qquad\qquad \mu_{\text{非常假}}(u)=(1-u)^2 \qquad u\in[0,1]$$

$$\mu_{\text{有些真}}(u)=u^{\frac{1}{2}} \qquad u\in[0,1] \qquad\qquad \mu_{\text{有些假}}(u)=(1-u)^{\frac{1}{2}} \qquad u\in[0,1]$$

3.5.3 模糊知识表示

2. 模糊知识的表示方式

在扎德的推理模型中，产生式规则的表示形式是

IF x is F THEN y is G

其中：x和y是变量，表示对象；F和G分别是论域U和V上的模糊集，表示概念。

模糊推理中所用的证据是用模糊命题表示的，其一般形式为

x is F'

3.5.4 模糊概念的匹配

1. 语义距离

语义距离用于刻划两个模糊概念之间的差异。这里主要讨论汉明距离。

离散论域：设U={u_1, u_2, … , u_n}是一个离散有限论域，F和G分别是论域U上的两个模糊概念的模糊集，则F和G的汉明距离定义为

$$d(F,G)=\frac{1}{n}\times\sum_{i=1}^{n}\left|\mu_F(u_i)-\mu_G(u_i)\right|$$

连续论域：如果论域U是实数域上的某个闭区间[a, b]，则汉明距离为

$$d(F,G)=\frac{1}{b-a}\int_a^b\left|\mu_F(u)-\mu_G(u)\right|d(u)$$

例3.15 设论域U={-10, 0, 10, 20, 30}表示温度，模糊集

F= 0.8/-10+0.5/0+0.1/10

G=0.9/-10+0.6/0+0.2/10

分别表示“冷”和“比较冷”，则

d(F,G)=0.2×(|0.8-0.9|+|0.5-0.6|+|0.1-0.2|)=0.2×0.3=0.06

即F和G的汉明距离为0.06。

求出距离后，可转化为匹配度

3.5.4 模糊概念的匹配

2. 贴近度

设F和G分别是论域 $U=\{u_1, u_2, \ldots, u_n\}$ 上的两个模糊概念的模糊集，则它们的贴近度定义为

$$(F, G)=(1/2) * (F \cdot G +(1-F \odot G))$$

其中：

$$F \cdot G = \bigvee_U (\mu_F(u_i) \wedge \mu_G(u_i))$$

$$F \odot G = \bigwedge_U (\mu_F(u_i) \vee \mu_G(u_i))$$

称F·G为内积，F⊙ G为外积。

例3.16 设论域U及其上的模糊集F和G如上例所示，则

$F \cdot G=(0.8 \wedge 0.9) \vee (0.5 \wedge 0.6) \vee (0.1 \wedge 0.2) \vee (0 \wedge 0) \vee (0 \wedge 0)$

$=0.8 \vee 0.5 \vee 0.1 \vee 0 \vee 0=0.8$

$F \odot G= (0.8 \vee 0.9) \wedge (0.5 \vee 0.6) \wedge (0.1 \vee 0.2) \wedge (0 \vee 0) \wedge (0 \vee 0)$

$=0.9 \wedge 0.6 \wedge 0.2 \wedge 0 \wedge 0 =0$

$(F, G)=0.5 \times (0.8+(1-0))=0.5 \times 1.8=0.9$

即F和G的贴近度为0.9。

模糊集的贴近度

贴近度是对两个F集接近程度的一种度量。

定义1　设$A,B,C\in F(U)$,若映射

$$N:\quad F(U)\times F(U)\to[0,1]$$

满足条件：

① $N(A,B)=N(B,A)$;

② $N(A,A)=1,N(U,\Phi)=0$;

③ 若$A\subseteq B\subseteq C$,则$N(A,C)\le N(A,B)\wedge N(B,C)$.

则称$N(A,B)$为F集A与B的贴近度。

3.5.5 模糊推理的方法

模糊推理实际上是按照给定的推理模式，通过模糊集合与模糊关系的合成来实现的。主要讨论：

1. 模糊关系的构造

2. 模糊推理的基本模式

3.5.5 模糊推理方法

1. 模糊关系的构造

模糊关系R_m

R_m是由扎德提出的一种构造模糊关系的方法。设F和G分别是论域U和V上的两个模糊集，则R_m定义为

$$R_m = \int_{U\times V} (\mu_F(u) \wedge \mu_G(v)) \vee (1-\mu_F(u))/(u,v)$$

其中，×号表示模糊集的笛卡尔乘积。

例3.17 设U=V={1，2，3}，F和G分别是U和V上的两个模糊集，且F=1/1+0.6/2+0.1/3， G=0.1/1+0.6/2+1/3，求U×V 上的 R_m

解：

$$R_m = \begin{bmatrix} 0.1 & 0.6 & 1 \\ 0.4 & 0.6 & 0.6 \\ 0.9 & 0.9 & 0.9 \end{bmatrix}$$

如：$R_m(2, 3) = (0.6 \wedge 1) \vee (1-0.6) = 0.6 \vee 0.4 = 0.6$

3.5.5 模糊推理方法

1. 模糊关系的构造

模糊关系R_c

R_c是由麦姆德尼(Mamdani)提出的一种构造模糊关系的方法。

设F和G分别是论域U和V上的两个模糊集，则R_c定义为

$$R_c = \int_{U \times V} (\mu_F(u) \wedge \mu_G(v)) /(u, v)$$

例：对例3.17所给出的模糊集

F=1/1+0.6/2+0.1/3,　G=0.1/1+0.6/2+1/3

其R_c为

$$R_c = \begin{bmatrix} 0.1 & 0.6 & 1 \\ 0.1 & 0.6 & 0.6 \\ 0.1 & 0.1 & 0.1 \end{bmatrix}$$

如R_c(3, 2):

$$R_c(3,2) = \mu_F(u_3) \wedge \mu_G(v_2) = 0.1 \wedge 0.6 = 0.1$$

3.5.5 模糊推理方法

1. 模糊关系的构造

模糊关系R_g

R_g是米祖莫托(Mizumoto)提出的一种构造模糊关系的方法。

设F和G分别是论域U和V上的两个模糊集，则R_g定义为

$$R_g = \int_{U\times V} (\mu_F(u) \to \mu_G(v)) /(u, v)$$

其中

$$\mu_F(u) \to \mu_G(v) = \begin{cases} 1 & \text{当}\mu_F(u) \le \mu_G(v)\text{时} \\ \mu_G(v) & \text{当}\mu_F(u) > \mu_G(v)\text{时} \end{cases}$$

例：对例3.17所给出的模糊集

F=1/1+0.6/2+0.1/3,　G=0.1/1+0.6/2+1/3

其R_g为

$$R_g = \begin{bmatrix} 0.1 & 0.6 & 1 \\ 0.1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

3.5.5 模糊推理方法

2. 模糊推理的基本模式

模糊假言推理(1/2)

设F和G分别是U和V上的两个模糊集，且有知识

IF x is F THEN y is G

若有U上的一个模糊集F'，且F可以和F'匹配，则可以推出y is G'，且G'是V上的一个模糊集。这种推理模式称为模糊假言推理，其表示形式为：

知识：IF x is F THEN y is G

证据：x is F'

结论：y is G'

在这种推理模式下，模糊知识

IF x is F THEN y is G

表示在F与G之间存在着确定的因果关系，设此因果关系为R。则有

G'=F'oR

其中的模糊关系R，可以是R_m、R_c或R_g中的任何一种。

3.5.5 模糊推理方法

2. 模糊推理的基本模式

模糊假言推理(2/2)

例3.18 对例3.17所给出的F、G，以及所求出的R_m，设有已知事实：

{x is 较小}

并设“较小”的模糊集为：

较小=1/1+0.7/2+0.2/3

求在此已知事实下的模糊结论。

解：本例的模糊关系R_m已在例3.17中求出，设已知模糊事实“较小”为F'，F'与R_m的合成即为所求结论G'。

$$G' = F' \circ R_m = \{1, 0.7, 0.2\} \circ \begin{bmatrix} 0.1 & 0.6 & 1 \\ 0.4 & 0.6 & 0.6 \\ 0.9 & 0.9 & 0.9 \end{bmatrix}$$ （首先计算F'和F的匹配度）

$=\{0.4, 0.6, 1\}$

即所求出的模糊结论G'为 G'=0.4/1+0.6/2+1/3

3.5.5 模糊推理方法

2. 模糊推理的基本模式

模糊拒取式推理(1/2)

设F和G分别是U和V上的两个模糊集，且有知识

IF x is F THEN y is G

若有V上的一个模糊集G'，且G'可以和G匹配，则可以推出x is F'，且F'是U上的一个模糊集。这种推理模式称为模糊拒取式推理，其表示形式为：

知识：IF x is F THEN y is G

证据：y is G'

结论：x is F'

在这种推理模式下，模糊知识

IF x is F THEN y is G

也表示在F与G之间存在着因果关系，设此关系为R，则有

F'=RoG'

其中的模糊关系R，可以是R_m、R_c或R_g中的任何一种。

3.5.5 模糊推理的方法

2. 模糊推理的基本模式

模糊拒取式推理(2/2)

例3.19 设F、G如例3.17所示，已知事实为｛ y is 较大｝且“较大”的模糊集为：较大=0.2/1+0.7/2+1/3，若已知事实与G匹配，以模糊关系Rc为例，在此已知事实下推出F’。

解：本例的模糊关系Rc已在前面求出(p101)，设模糊概念“较大”为G'，则 Rc与G'的合成即为所求的F'。

$$F' = R_c \circ G' = \begin{bmatrix} 0.1 & 0.6 & 1 \\ 0.1 & 0.6 & 0.6 \\ 0.1 & 0.1 & 0.1 \end{bmatrix} \circ \begin{bmatrix} 0.2 \\ 0.7 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.6 \\ 0.1 \end{bmatrix}$$

即所求出的F'为　　F'=1/1+0.6/2+0.1/3

3.5.5 模糊推理的方法

2. 模糊推理的基本模式

模糊假言三段论推理(1/3)

设F、G、H分别是U、V、W上的3个模糊集，且由知识

IF x is F THEN y is G

IF y is G THEN z is H

则可推出:

IF x is F THEN z is H

这种推理模式称为模糊假言三段论推理。

3.5.5 模糊推理方法

2. 模糊推理的基本模式

模糊假言三段论推理(2/3)

在模糊假言三段论推理模式下，模糊知识

r_1: IF x is F THEN y is G

表示在F与G之间存在着确定的因果关系，设此因果关系为R_1。

模糊知识

r_2: IF y is G THEN z is H

表示在G与H之间存在着确定的因果关系，设此因果关系为R_2。

若模糊假言三段论成立，则模糊结论

r_3: IF x is F THEN z is H

的模糊关系R_3可由R_1与R_2的合成得到。即

$$R_3=R_1 \circ R_2$$

这里的关系R_1、R_2、R_3都可以是前面所讨论过的R_m、R_c、R_g中的任何一种。

3.5.5 模糊推理方法

2. 模糊推理的基本模式

模糊假言三段论推理(3/3)

例3.20 设U=W=V={1, 2, 3}，E=1/1+0.6/2+0.2/3，F=0.8/1+0.5/2+0.1/3，G=0.2/1+0.6/2+1/3。按R_g求E×F×G上的关系R。

解：先求E×F上的关系R_1

$$R_1=\begin{bmatrix}0.8 & 0.5 & 0.1\\ 1 & 0.5 & 0.1\\ 1 & 1 & 0.1\end{bmatrix}$$

再求F×G上的关系R_2

$$R_2=\begin{bmatrix}0.2 & 0.6 & 1\\ 0.2 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}$$

最后求E×F×G上的关系R

$$R=R_1\circ R_2=\begin{bmatrix}0.2 & 0.6 & 0.8\\ 0.2 & 0.6 & 1\\ 0.2 & 1 & 1\end{bmatrix}$$

进展I: Takagi-Sugeno Model

T. TAKAGI AND M. SUGENO, Fuzzy Identification of Systems and Its Applications to Modeling and Control, IEEE TRANS. SYSTEMS, MAN, AND CYBERNETICS, VOL.15, NO. 1, 1985

$$\begin{aligned}&\text{Rule } i\text{: IF } x_1(k) \text{ is } M_{i1} \cdots \text{ and } x_n(k) \text{ is } M_{in}\\ &\qquad \text{THEN } \boldsymbol{x}(k+1) = \boldsymbol{A}_i\boldsymbol{x}(k) + \boldsymbol{B}_i\boldsymbol{u}(k)\end{aligned}$$

final output of the fuzzy system is inferred as follows:

$$\boldsymbol{x}(k+1) = \frac{\sum_{i=1}^{r} w_i(k)\{\boldsymbol{A}_i\boldsymbol{x}(k) + \boldsymbol{B}_i\boldsymbol{u}(k)\}}{\sum_{i=1}^{r} w_i(k)}$$

where

$$w_i(k) = \prod_{j=1}^{n} M_{ij}(x_j(k)).$$

$M_{ij}(x_j(k))$ is the grade of membership of $x_j(k)$ in M_{ij}.

进展II: Type-2 Fuzzy Systems

Jerry M. Mendel
http://sipi.usc.edu/~mendel/

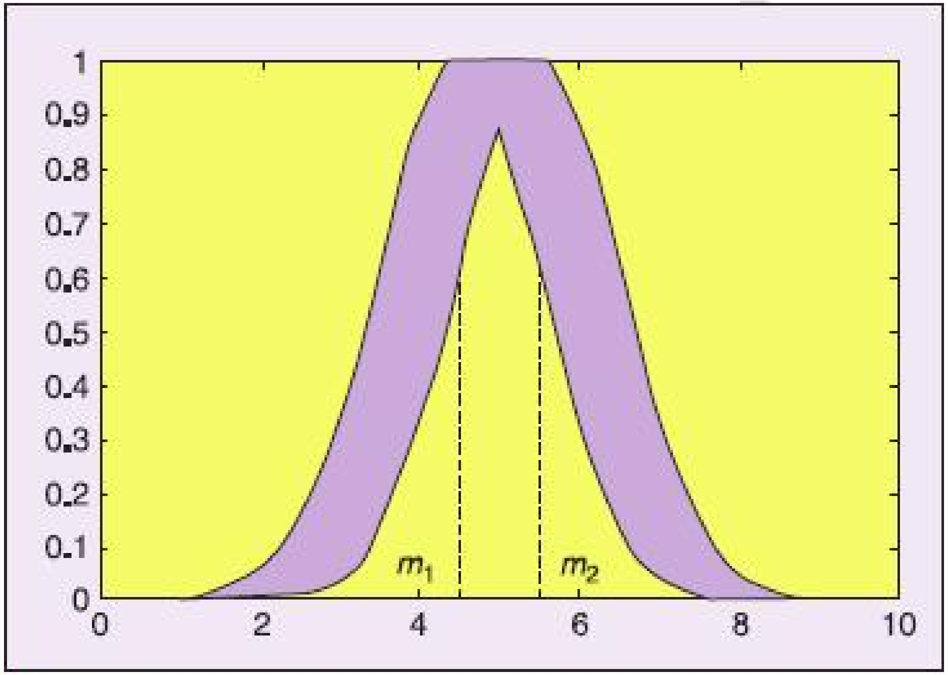
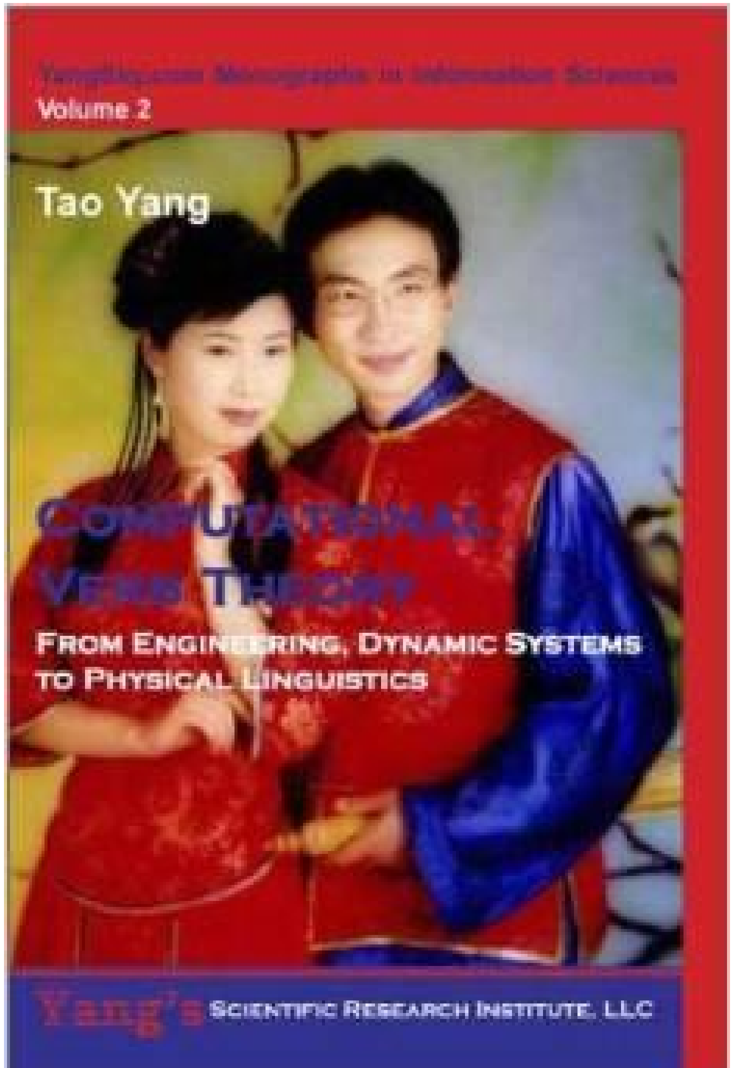


FIGURE 2 FOU for a Gaussian primary MF whose mean varies in the interval $[m_1, m_2]$ and whose standard deviation is a constant.

进展III: Computational Verb Theory



Tao Yang

www.yangsky.com

International Journal of Computational Cognition

进展IV: Generalized Information Theory

George J. Klir

钟义信

鲁晨光　http://survivor99.com/lcg/

……

提及并不代表认同

Li-Xin Wang vs Fei-Yue Wang

Li-Xin Wang received the Ph.D. degree in 1992 from the Department of Electrical Engineering – Systems, University of Southern California (won USC's Phi Kappa Phi's highest Student

of Science and Techno
HKUST to become an
real estate markets in
Fall 2013 by joining t
and Technology, Xian
across the wild land of

His research intere
microstructure, tradin
control.

Fei-Yue Wang (S'87–M'89–SM'94–F'03) received the Ph.D. degree in computer and systems engineering from Rensselaer Polytechnic Institute, Troy, NY, in 1990.

He joined the University of Arizona, Tucson, in 1990 and became a Professor and the Director of the Robotics Laboratory and Program in Advanced Research for Complex Systems. In 1999, he founded the Intelligent Control and Systems Engineering Center at the Chinese Academy of Sciences (CAS), Beijing, China, under the support of the Outstanding Oversea Chinese Talents Program. Since 2002, he has been the Director of the Key Laboratory on Complex Systems and Intelligence Science, CAS. He is currently the Vice President of the Institute of Automation, CAS. From 1995 to 2000, he was the Editor-in-Chief of the *International Journal of Intelligent Control and Systems* and the *World Scientific Series in Intelligent Control and Intelligent Automation*. His research interests include social computing, web science, and intelligent control.

Li-Xin Wang vs Fei-Yue Wang

标题 1–20	引用次数	发表年份
Adaptive fuzzy systems and control: design and stability analysis LX Wang Prentice-Hall, Inc.	3784	1994
A course in fuzzy systems LX Wang Prentice-Hall press, USA	[illegible]	[illegible]
Generating fuzzy rules b LX Wang, JM Mendel Systems, Man and Cybernet		
Fuzzy basis functions, u least-squares learning LX Wang, JM Mendel Neural Networks, IEEE Tran		
Stable adaptive fuzzy co LX Wang Fuzzy Systems, IEEE Trans		
Fuzzy systems are unive LX Wang Fuzzy Systems, 1992., IEEE		

A new look at type-2 fuzzy sets and type-2 fuzzy logic systems LX Wang IEEE Transactions on Fuzzy Systems 25 (3), 693-706	6	2017
Modeling stock price dynamics with fuzzy opinion networks LX Wang IEEE Transactions on Fuzzy Systems 25 (2), 277-301	5	2017
Fuzzy opinion networks: A mathematical framework for the evolution of opinions and their uncertainties across social networks LX Wang, JM Mendel IEEE Transactions on Fuzzy Systems 24 (4), 880-905	8	2016
Benefits of transformational behaviors for leaders: A daily investigation of leader behaviors and need fulfillment. K Lanaj, RE Johnson, SM Lee Journal of Applied Psychology 101 (2), 237	34	2016
Dynamical models of stock prices based on technical trading rules—Part III: Application to Hong Kong stocks LX Wang IEEE Transactions on Fuzzy Systems 23 (5), 1680-1697	~~9~~	2015
Dynamical models of stock prices based on technical trading rules—Part III: Application to Hong Kong stocks LX Wang IEEE Transactions on Fuzzy Systems 23 (5), 1680-1697	9	2015
Dynamical models of stock prices based on technical trading rules part I: the models LX Wang IEEE Transactions on Fuzzy Systems 23 (4), 787-801	16 *	2015

Li-Xin Wang vs Fei-Yue Wang

[Reduction and axiomization of covering generalized rough sets]
W Zhu, FY **Wang** - Information sciences, 2003 - Elsevier
This paper investigates some basic properties of covering generalized rough sets, and their comparison with the corresponding ones of Pawlak's rough sets, a tool for data mining. The focus here is on the concepts and conditions for two coverings to generate the same ...
被引用次数：583 相关文章 所有 3 个版本 引用

[Social computing: From social informatics to social intelligence] per
FY **Wang**, KM Carley, D Zeng... - Intelligent Systems, IEEE, 2007 - ieeexplore.ieee.org
Social computing is a cross-disciplinary research and application field with theoretical underpinnings including both computational and social sciences (see figure 1). To support social interaction and communication, it relies on communication; humancomputer ...
被引用次数：358 相关文章 所有 15 个版本 引用

[On three types of covering-based rough sets]
W Zhu, FY **Wang** - ... and Data Engineering, IEEE Transactions on, 2007 - ieeexplore.ieee.org
Abstract—Rough set theory is a useful tool for data mining. It is based on equivalence relations and has been extended to coveringbased generalized rough set. This paper studies three kinds of covering generalized rough sets for dealing with the vagueness and ...
被引用次数：289 相关文章 所有 7 个版本 引用

[Adaptive dynamic programming: an introduction]
FY **Wang**, H Zhang, D Liu - ... Intelligence Magazine, IEEE, 2009 - ieeexplore.ieee.org
Abstract: In this article, we introduce some recent research trends within the field of adaptive/approximate dynamic programming (ADP), including the variations on the structure of ADP schemes, the development of ADP algorithms and applications of ADP schemes. ...
被引用次数：281 相关文章 所有 4 个版本 引用

[Parallel control and management for intelligent transportation systems: Concepts, architectures, and applications]
FY **Wang** - Intelligent Transportation Systems, IEEE ..., 2010 - ieeexplore.ieee.org
Abstract—Parallel control and management have been pro-posed as a new mechanism for conducting operations of complex systems, especially those that involved complexity issues of both engineering and social dimensions, such as transportation systems. This paper ...
被引用次数：230 相关文章 所有 7 个版本 引用

作 业

3.10 证据理论

3.10 设有如下一组推理规则：

r_1: IF E_1 AND E_2 THEN $A=\{a\}$ (CF = {0.9})

r_2: IF E_2 AND (E_3 OR E_4) THEN $B=\{b_1, b_2\}$ (CF = {0.5, 0.4})

r_3: IF A THEN $H=\{h_1, h_2, h_3\}$ (CF = {0.2, 0.3, 0.4})

r_4: IF B THEN $H=\{h_1, h_2, h_3\}$ (CF = {0.3, 0.2, 0.1})

且已知初始证据的确定性分别为：$\mathrm{CER}(E_1)=0.6$，$\mathrm{CER}(E_2)=0.7$，$\mathrm{CER}(E_3)=0.8$，$\mathrm{CER}(E_4)=0.9$，假设$|\Omega|=10$，求$\mathrm{CER}(H)$。

3.14 模糊运算

3.14 设有论域 $U=\{u_1, u_2, u_3, u_4, u_5\}$，并设 F、G 是 U 上的两个模糊集，且有

$$F=0.9/u_1+0.7/u_2+0.5/u_3+0.3/u_4$$

$$G=0.6/u_3+0.8/u_4+1/u_5$$

请分别计算 $F\cap G$，$F\cup G$，$\neg F$。

3.16 模糊关系合成

3.17 模糊变换

3.16　设有如下两个模糊关系：

$$R_1=\begin{bmatrix}0.3 & 0.7 & 0.2\\ 1 & 0 & 0.4\\ 0 & 0.5 & 1\end{bmatrix},\qquad R_2=\begin{bmatrix}0.2 & 0.8\\ 0.6 & 0.4\\ 0.9 & 0.1\end{bmatrix}$$

请写出 $\boldsymbol{R}_1$ 与 $\boldsymbol{R}_2$ 的合成 $\boldsymbol{R}_1\circ\boldsymbol{R}_2$。

3.17　设 F 是论域 U 上的模糊集，$\boldsymbol{R}$ 是 $U\times V$ 上的模糊关系，F 和 $\boldsymbol{R}$ 分别为

$$F=\{0.4,0.6,0.8\},\qquad \boldsymbol{R}=\begin{bmatrix}0.1 & 0.3 & 0.5\\ 0.4 & 0.6 & 0.8\\ 0.6 & 0.3 & 0\end{bmatrix}$$

求模糊变换 $F\circ\boldsymbol{R}$。

3.20 模糊推理

3.20 设 $U=V=W=\{1,2,3,4\}$，且设有如下规则：

r_1: IF x is F THEN y is G

r_2: IF y is G THEN z is H

r_3: IF x is F THEN z is H

其中，F、G、H 的模糊集分别为

$$\begin{cases} F=1/1+0.8/2+0.5/3+0.4/4 \\ G=0.1/2+0.2/3+0.4/4 \\ H=0.2/2+0.5/3+0.8/4 \end{cases}$$

请分别对各种模糊关系验证满足模糊假言三段论的情况。

3.24 贝叶斯网络

3.24 设有如图 3.6 所示的贝叶斯网络，请计算报警铃响了但实际上并无盗贼入侵，也未发生地震，而李和张都打来电话的概率。

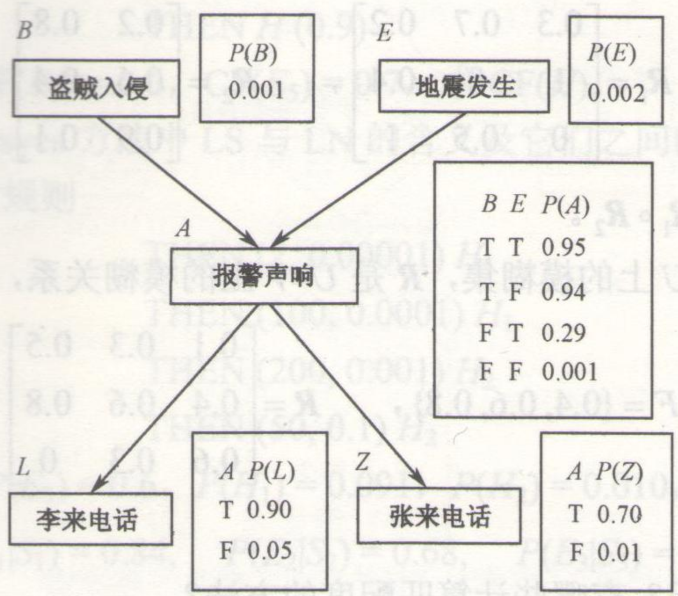


图 3.6 习题 3.24 的贝叶斯网络