本章主要内容

1. 符号学派简介
2. 知识表示与推理的基本概念
3. 谓词逻辑的知识表示与推理
4. 产生式的知识表示与推理
5. 不确定知识表示与推理
6. 语义网络与知识图谱

产生式

- “产生式”由美国数学家波斯特（E.POST）在1934年首先提出，它根据串代替规则提出了一种称为波斯特机的计算模型，模型中的每条规则称为产生式。
- 许多专家系统用它来表示知识
- 产生式通常用于表示事实、规则以及它们的不确定性度量，适合于表示事实性知识和规则性知识。

产生式

- （1）确定性规则知识的产生式表示

▪ 基本形式：　IF　*P*　THEN　*Q*

　　或者：　$P \rightarrow Q$

　P 是产生式的前提，*Q* 是一组结论或操作。

　如果前提*P*被满足，则可得到结论*Q*或执行*Q*所规定的操作。

例如：

$$\underbrace{r_1:}_{\text{编号}}\ \underbrace{\text{IF　动物会飞　AND　会下蛋}}_{P}\ \underbrace{\text{THEN　该动物是鸟}}_{Q}$$

产生式

- （2）不确定性规则知识的产生式表示

- 基本形式：　IF　*P*　THEN　*Q*　（置信度）

　　或者：　$P \rightarrow Q$（置信度）

例如：IF　　发烧　　THEN　　　感冒　（0.6）

例如在专家系统MYCIN中：

　　IF　本微生物的染色斑是革兰氏阴性

　　　　本微生物的形状呈杆状

　　　　病人是中间宿主

　　Then 该微生物是绿脓杆菌，置信度是0.6

产生式

- （3）确定性事实性知识的产生式表示

▪ 三元组表示：　（对象，属性，值）

　　　　或者：　（关系，对象1，对象2）

例：　老李年龄是40岁：　　（*Li*，*age*，40）

　　　老李和老王是朋友：　（*friend*，*Li*，*Wang*）

产生式

- （4）不确定性事实性知识的产生式表示

▪ 四元组表示：（对象，属性，值，置信度）

或者：（关系，对象1，对象2，置信度）

例：老李年龄很可能是40岁：（*Li*，*age*，40，0.8）

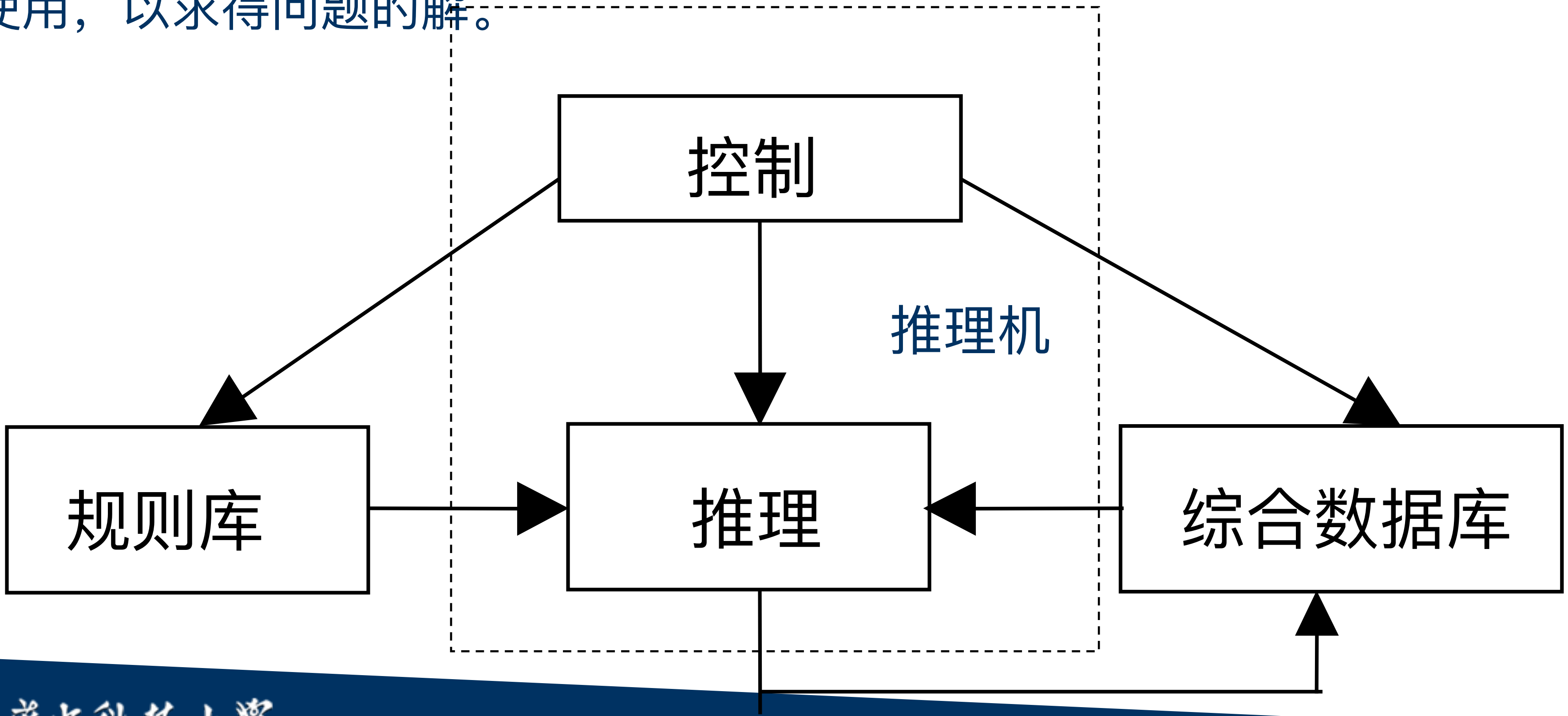
老李和老王不大可能是朋友：（*friend*，*Li*，*Wang*，*0.1*）

产生式和蕴含式的区别

- 产生式与谓词逻辑中的蕴含式基本形式相同
- 蕴含式是产生式的一种特殊情况
  - 除逻辑蕴含外，产生式还包括各种操作、规则、变换、算子、函数等。例如，“如果炉温超过上限，则立即关闭风门”是一个产生式，但不是蕴含式。产生式的外延很广。
  - 蕴含式只能表示精确知识，而产生式不仅可以表示精确的知识，还可以表示不精确知识。蕴含式的匹配总要求是精确的。产生式匹配可以是精确的，也可以是不精确的，只要按某种算法求出的相似度落在预先指定的范围内就认为是可匹配的。
- 产生式的形式描述及语义：
  - 巴科斯范式BNF（Backus normal form）

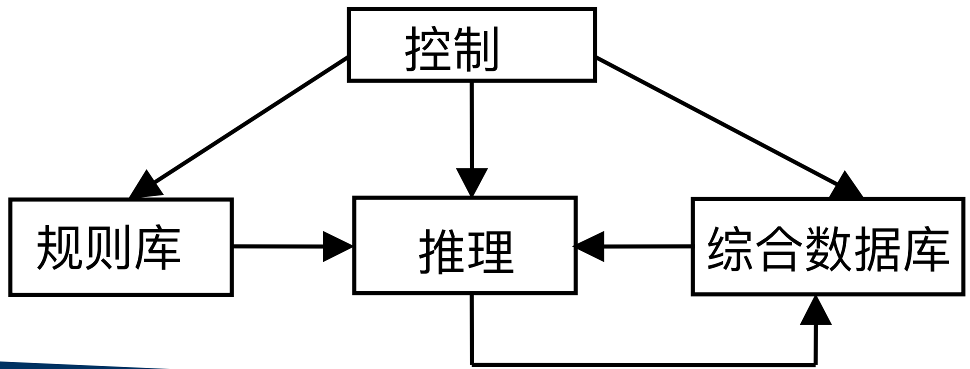
产生式系统

•一组产生式放在一起，互相配合，一个产生式的结论可以供另一个作为事实使用，以求得问题的解。



产生式系统

- 规则库：用于描述相应领域内知识的产生式集合。
- 综合数据库（事实库、上下文、黑板等）：一个用于存放问题求解过程中各种当前信息的数据结构。
- 控制系统（推理机）：由一组程序组成，负责整个产生式系统的运行，实现对问题的求解。
  - 推理
  - 冲突消解
  - 执行规则
  - 检查推理终止条件



产生式系统例子

•例如：动物识别系统——识别虎、金钱豹、斑马、长颈鹿、鸵鸟、企鹅、信天翁七种动物的产生式系统。







产生式系统例子

r_1： IF 该动物有毛发 THEN 该动物是哺乳动物

r_2： IF 该动物有乳房 THEN 该动物是哺乳动物

r_3： IF 该动物有羽毛 THEN 该动物是鸟

r_4： IF 该动物会飞 AND 会下蛋 THEN 该动物是鸟

r_5： IF 该动物吃肉 THEN 该动物是食肉动物

r_6： IF 该动物有犬齿 AND 有爪 AND 眼盯前方
THEN 该动物是食肉动物

r_7： IF 该动物是哺乳动物 AND 有蹄
THEN 该动物是有蹄类动物

r_8： IF 该动物是哺乳动物 AND 是反刍动物
THEN 该动物是有蹄类动物

规则库

产生式系统例子

r_9: IF 该动物是哺乳动物 AND 是食肉动物 AND 是黄褐色
AND 身上有暗斑点 THEN 该动物是金钱豹

r_{10}: IF 该动物是哺乳动物 AND 是食肉动物 AND 是黄褐色
AND 身上有黑色条纹 THEN 该动物是虎

r_{11}: IF 该动物是有蹄类动物 AND 有长脖子 AND 有长腿
AND 身上有暗斑点 THEN 该动物是长颈鹿

r_{12}: IF 该动物是有蹄类动物 AND 身上有黑色条纹
THEN 该动物是斑马

r_{13}: IF 该动物是鸟 AND 有长脖子 AND 有长腿 AND 不会飞
AND 有黑白二色 THEN 该动物是鸵鸟

r_{14}: IF 该动物是鸟 AND 会游泳 AND 不会飞
AND 有黑白二色 THEN 该动物是企鹅

r_{15}: IF 该动物是鸟 AND 善飞 THEN 该动物是信天翁

产生式系统例子

- 设已知初始事实存放在综合数据库中：

该动物身上有：暗斑点，长脖子，长腿，乳房，蹄，哺乳动物

r_1：IF 该动物有毛发 THEN 该动物是哺乳动物

[redacted]

r_3：IF 该动物有羽毛 THEN 该动物是鸟

r_4：IF 该动物会飞 AND 会下蛋 THEN 该动物是鸟

r_5：IF 该动物吃肉 THEN 该动物是食肉动物

r_6：IF 该动物有犬齿 AND 有爪 AND 眼盯前方 THEN 该动物是食肉动物

r_7：IF 该动物是哺乳动物 AND 有蹄 THEN 该动物是有蹄类动物

r_8：IF 该动物是哺乳动物 AND 是反刍动物 THEN 该动物是有蹄类动物

√ ->

"该动物是哺乳动物"

加入综合数据库

产生式系统例子

- 设已知初始事实存放在综合数据库中：

该动物身上有：暗斑点，长脖子，长腿，乳房，蹄，哺乳动物，有蹄类动物

r_1： IF 该动物有毛发 THEN 该动物是哺乳动物

[redacted]

r_3： IF 该动物有羽毛 THEN 该动物是鸟

r_4： IF 该动物会飞 AND 会下蛋 THEN 该动物是鸟

r_5： IF 该动物吃肉 THEN 该动物是食肉动物

r_6： IF 该动物有犬齿 AND 有爪 AND 眼盯前方

THEN 该动物是食肉动物

[redacted]

THEN 该动物是有蹄类动物

r_8： IF 该动物是哺乳动物 AND 是反刍动物

THEN 该动物是有蹄类动物

√ ->

“该动物是哺乳动物”

加入综合数据库

√ ->

“该动物是有蹄类动物”

加入综合数据库

产生式系统例子

该动物身上有：暗斑点，长脖子，长腿，乳房，蹄，哺乳动物，有蹄类动物

r_9：IF 该动物是哺乳动物 AND 是食肉动物 AND 是黄褐色
AND 身上有暗斑点 THEN 该动物是金钱豹

r_{10}：IF 该动物是哺乳动物 AND 是食肉动物 AND 是黄褐色
AND 身上有黑色条纹 THEN 该动物是虎

r_{11}：IF 该动物是有蹄类动物 AND 有长脖子 AND 有长腿
AND 身上有暗斑点 THEN 该动物是长颈鹿

r_{12}：IF 该动物是有蹄类动物 AND 身上有黑色条纹
THEN 该动物是斑马

r_{13}：IF 该动物是鸟 AND 有长脖子 AND 有长腿 AND 不会飞
AND 有黑白二色 THEN 该动物是鸵鸟

r_{14}：IF 该动物是鸟 AND 会游泳 AND 不会飞
AND 有黑白二色 THEN 该动物是企鹅

r_{15}：IF 该动物是鸟 AND 善飞 THEN 该动物是信天翁

匹配成功！

长颈鹿

产生式系统的特点

- 不断从规则库中选择可用规则与综合数据库中的已知事实进行匹配
- 规则的每一次成功匹配都使综合数据库中增加新的内容，朝着问题解决的方向前进，称为推理，是专家系统中的核心内容

优点：

- 自然性
- 模块性
- 有效性
- 清晰性

缺点：

- 效率不高
- 不能表达结构性知识

适合产生式表示的知识：

- 领域知识间关系不密切，不存在结构关系，如化学方面。
- 经验性及不确定性的知识，且相关领域中对这些知识没有严格、统一的理论，如医疗论断。
- 领域问题的求解过程可被表示为一系列相对独立的操作，且每个操作可被表示为一条或多条产生式规则。

本章主要内容

1. 符号学派简介
2. 知识表示与推理的基本概念
3. 谓词逻辑的知识表示与推理
4. 产生式的知识表示与推理
5. 不确定知识表示与推理
6. 语义网络与知识图谱

本章主要内容

- 不确定知识表示与推理

➢ 不确定知识表示与推理的概念

➢ CF模型

➢ DS证据理论

➢ 模糊推理

不确定性知识表示与推理

- 现实世界中客观上存在的随机性、模糊性
- 我们对事物的认识往往是不精确、不完全的，具有一定程度的不确定性
- 反映到知识和证据上，就形成了不确定性知识及不确定性证据
- 在信息不完善、不精确的情况下，如何推理?

↓

不确定性推理

主要介绍基于概率论的方法

不确定性推理

- 不确定性推理：从不确定性的初始证据出发，通过运用不确定性的知识，最终推出具有一定程度的不确定性但却是合理或者近乎合理的结论的思维过程。
- 一些基本问题：
  - 不确定性的表示与量度
  - 不确定性匹配算法及阈值的选择
  - 组合证据不确定性的算法
  - 不确定性的传递算法
  - 结论不确定性的合成

不确定性推理的基本问题

- 不确定性的表示与量度
  - 知识不确定性的表示
  - 证据不确定性的表示
  - 不确定性的度量
- 不确定性匹配算法及阈值的选择
  - 知识与数据库中已知证据进行匹配时，由于知识和证据都具有不确定性，如果知识所要求的不确定性程度与证据实际具有的不确定性知识不同，怎么才算匹配成功?
  - 设计一个算法来计算双方相似的程度
  - 设定阈值来衡量多大程度可称之为匹配

不确定性推理的基本问题

- 组合证据不确定性的算法
  - 多个简单条件可通过逻辑运算构成复合条件
  - 匹配时对应即为组合证据，其不确定性需要方法来计算
- 不确定性的传递算法
  - 每一步推理中，如何把证据及知识的不确定性传递给结论
  - 在多步推理中，如何把初始证据的不确定性传递给最终结论
- 结论不确定性的合成
  - 用不同知识进行推理得到了相同结论，但得到的不确定程度不同，需要算法来进行处理

本章主要内容

- 不确定知识表示与推理

➢ 不确定知识表示与推理的概念

➢ CF模型

➢ DS证据理论

➢ 模糊推理

可信度方法

- 1975年肖特里菲（E. H. Shortliffe）等人在确定性理论（Theory of Confirmation）的基础上，结合概率论等提出的一种不确定性推理方法
- 是许多专家系统的基础
- 可信度：根据经验对一个事物或现象为真的相信程度
- 可信度方法带有较大的主观性和经验性，但直观、简单，且有效
- 基本方法：C-F模型（ The Certainty-Factor Model ）

C-F模型

- 1. 知识不确定性的表示
  - 产生式规则表示：

    IF　E　THEN　H　(*CF(H,E)*)

  - 可信度因子*CF(H,E)*：该条知识的可信度
  - *CF(H,E)*在[-1, 1]上取值，由领域专家直接给出
    - ➢证据的出现增加结论H为真的可信度，则$CF(H,E)>0$
    - ➢证据的出现增加结论H为假的可信度，则$CF(H,E)<0$
    - ➢证据的出现与H无关，则$CF(H,E)=0$

例：IF　头痛　AND　流涕　THEN　感冒　(0.7)

C-F模型

- 2. 证据不确定性的表示

$$CF(E)=0.6\text{：}\quad E\text{的可信度为}0.6$$

  - $CF(E)$也在[-1, 1]上取值
  - 对初始证据，如肯定为真则取1，肯定为假则取-1
  - 某种程度为真则取$0 < CF(E) < 1$，某种程度为假则取 $-1 < CF(E) < 0$
  - 如果暂未获得相关的观察，可看作无关，取0

C-F模型

- 3. 组合证据不确定性的算法
  - 当组合证据是多个单一证据的合取（与）时，可信度取最小值
  - 当组合证据是多个单一证据的析取（或）时，可信度取最大值

$$E = E_1 \quad \text{AND} \quad E_2 \quad \text{AND} \quad \cdots \quad \text{AND} \quad E_n$$

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

$$E = E_1 \quad \text{OR} \quad E_2 \quad \text{OR} \quad \cdots \quad \text{OR} \quad E_n$$

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

C-F模型

- 4. 不确定性的传递算法
  - 从不确定的初始证据出发，最终推出结论并求出其可信度
  - 结论H的可信度：

$$CF(H) = \underbrace{CF(H,E)}_{\text{知识的可信度}} \times \underbrace{\max\{0, CF(E)\}}_{\text{证据的可信度}}$$

  - 当证据以某种程度为假时，即CF(E)<0，则CF(H)=0
  - 当证据为真，即CF(E)=1，则CF(H)= CF(H, E)

模型中没有考虑证据为假的程度对结论H所产生的影响

当证据为真时，结论H的可信度就是知识的可信度

C-F模型

- 5. 结论不确定性的合成算法
  - 如果由多条不同知识推出了相同结论，但可信度不同，需要合成算法求出综合可信度
  - 考虑两两合成，有知识如下：

IF	E_1	THEN	H	$(CF(H, E_1))$
IF	E_2	THEN	H	$(CF(H, E_2))$

  - 分别对每条知识求出结论可信度CF(H)：

$$CF_1(H)=CF(H, E_1)\times\max\{0, CF(E_1)\}$$
$$CF_2(H)=CF(H, E_2)\times\max\{0, CF(E_2)\}$$

  - 用下述公式求出E_1和E_2对H的综合可信度$CF_{12}(H)$：

C-F模型

$$CF_{1,2}(H)=\begin{cases} CF_1(H)+CF_2(H)-CF_1(H)CF_2(H) & 若CF_1(H)\geq 0,\ CF_2(H)\geq 0 \\ CF_1(H)+CF_2(H)+CF_1(H)CF_2(H) & 若CF_1(H)<0,\ CF_2(H)<0 \\ \dfrac{CF_1(H)+CF_2(H)}{1-\min\{|CF_1(H)|,|CF_2(H)|\}} & 若CF_1(H)与CF_2(H)异号 \end{cases}$$

- 例： 如果头痛，则感冒（0.8）
- 如果发热，则感冒（0.9）
- 如果头痛且咳嗽，则发热（0.7）
- 妈妈观察到小王头痛的可信度是0.6，咳嗽的可信度是0.8，问小王是否感冒？

设各项证据为：

头痛E_1 $CF(E_1)=0.6$ 感冒H

发热E_2

咳嗽E_3 $CF(E_3)=0.8$

知识可写为：

1. IF E_1 THEN H (0.8)
2. IF E_2 THEN H (0.9)
3. IF $E_1 \wedge E_3$ THEN E_2 (0.7)

C-F模型

已知规则和证据：

r_1: IF E_1 THEN H (0.8)

r_2: IF E_2 THEN H (0.9)

r_3: IF $E_1 \wedge E_3$ THEN E_2 (0.7)

$CF(E_1)=0.6$；$CF(E_3)=0.8$

分别对每条规则求出结论可信度CF(H)：

由r_1得到：$CF_1(H) = CF(H, E_1)\times\max\{0, CF(E_1)\} = 0.8\times0.6 = 0.48$

由r_3得到：$CF(E_2) = CF(E2, E1\wedge E3)\times\max\{0, CF(E1\wedge E3)\}$

$= 0.7\times\min\{CF(E_1), CF(E_3)\} = 0.7\times0.6 = 0.42$

由r_2得到：$CF_2(H) = CF(H, E_2)\times\max\{0, CF(E_2)\}= 0.9\times0.42 = 0.378$

求出结论的综合可信度：

$$CF_{12}(H) = CF_1(H)+CF_2(H)-CF_1(H)\,CF_2(H)$$
$$= 0.48+0.378-0.48\times0.378 = 0.67656$$

确实比较有可能感冒了！

$$CF_{1,2}(H) = \begin{cases} \boxed{CF_1(H)+CF_2(H)-CF_1(H)CF_2(H) \quad 若CF_1(H)\ge 0 \quad CF_2(H)\ge 0} \\ CF_1(H)+CF_2(H)+CF_1(H)CF_2(H) \quad 若CF_1(H)<0 \quad CF_2(H)<0 \\ \dfrac{CF_1(H)+CF_2(H)}{1-\min\{|CF_1(H)|,|CF_2(H)|\}} \quad 若CF_1(H) \quad CF_2(H) \end{cases}$$

本章主要内容

- 不确定知识表示与推理

➢ 不确定知识表示与推理的概念

➢ CF模型

➢ DS证据理论

➢ 模糊推理

D-S证据理论的发展历史

- Dempster在1967年的文献《多值映射导致的上下文概率》[1] 中提出上、下概率的概念。
- 在文献《贝叶斯推理的一般化》[2] 中进一步探讨了不满足可加性的概率问题以及统计推理的一般化问题。
- Shafer在Dempster研究的基础上提出了证据理论，把Dempster合成规则推广到更为一般的情况，并于1976年出版《证据的数学理论》[3]

主要参考文献

- [1] Dempster, A. P. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 1967, 38(2): 325-339. 【提出证据理论的第一篇文献】
- [2] Dempster, A. P. Generalization of Bayesian Inference. *Journal of the Royal Statistical Society*. Series B 30, 1968:205-247.
- [3] Shafer, G. ***A Mathematical Theory of Evidence***. Princeton University Press, 1976. 【证据理论的第一本专著，标志其正式成为一门理论】
- [4] Barnett, J. A. Computational methods for a mathematical theory of evidence. In: *Proceedings of 7th International Joint Conference on Artificial Intelligence(IJCAI-81)*, Vancouver, B. C., Canada, Vol. II, 1981: 868-875. 【第一篇将证据理论引入AI领域的标志性论文】

为什么需要证据理论

- 应用：多分类器融合、不确定性推理、专家意见综合、多准则决策、模式识别、综合诊断等领域

- 各个领域遇到的不确定性证据和证据冲突问题
  - 自动驾驶领域中，多传感器信息融合问题
  - 网络安全入侵检测领域中，多源日志和多源告警的融合决策问题
  - 数据分析中，不同的分类器的输出结果的综合决策问题
  - Etc.

为什么需要证据理论

- 举例：发生抢劫案，警方判定罪犯肯定是嫌疑人A、B、C中的一个，但不知道是哪一个。
  - 两个证人张三、李四只是看到了部分过程，有不同的判断，用概率表示。
  - 共三种情况：A作案，B作案，C作案，具体如下：

假设	张三认为	李四认为
A作案	0.86	0.02
B作案	0.13	0.90
C作案	0.01	0.08

为什么需要证据理论

- **DS证据理论用途**：根据不同证人提供的概率，给出每种假设的综合概率，起到了不同数据源数据融合的作用。

- 比如通过DS理论综合得出结果如下（B的嫌疑更大）：

假设	综合概率
A作案	0.1274
B作案	0.8667
C作案	0.0059

证据理论基本概念

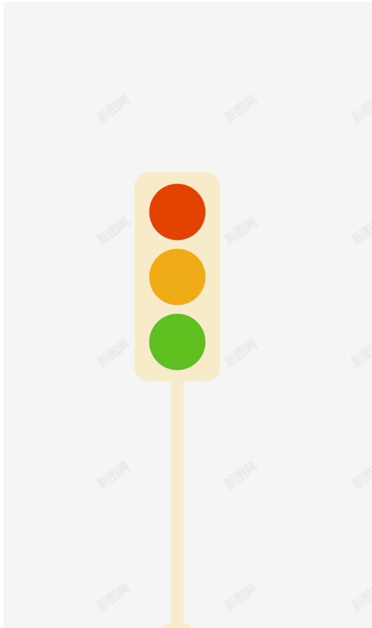
识别框架

- 状态集合X：例：$X = \{a, b, c\}$
  - X的幂集
    - ➢ 原集合中所有的子集（包括全集和空集）构成的集族
    - ➢ 记为：2^X：$\{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{a, c\}, \{a, b, c\}\}$
- 基本概率分配(Basic Probability Assignment, BPA)：
  - 给2^X的每一个元素（假设）分配一个概率（测度）
  - 称为mass函数，表示相信的程度(a degree of belief)
  - 满足：$\sum_{A \subseteq X} m(A) = 1, m(\emptyset) = 0$　　使得m(A)>0的A称为**焦元**（Focal elements）

证据理论基本概念

- 例：交通信号灯，可能状态：X: {Green, Yellow, Red}
  - 假设：2^X: $\{\varnothing,\{G\},\{Y\},\{R\},\{G,Y\},\{Y,R\},\{G,R\},\{G,Y,R\}\}$

Hypothesis	Mass
Null	0
Green	0.15
Yellow	0.25
Red	0.35
Green or Yellow	0.04
Yellow or Red	0.06
Green or Red	0.05
Any	0.1



证据理论基本概念

•信任函数(Belief function)，记为Bel

$$Bel(A)=\sum_{B\subseteq A} m(B)\,,\forall A\in 2^X$$

- 如何理解$\boldsymbol{B\subseteq A}$：A是集合，B是A的所有子集，$\boldsymbol{B}\in 2^X$
- It is the amount of belief that directly supports either the given hypothesis (A) or a more specific one (B, subset of A)
- 信任函数表达了客观概率的下限，下限函数

证据理论基本概念

•似真(然)函数(Plausibility function)，记为Pl

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

$B\cap A\neq\emptyset$ ：与集合A有公共元素

- 所有与假说A存在交集的mass函数之和
- 1减去所有与A不存在交集的mass函数之和
- 似真函数表示客观概率的上限，it “could possibly be the true state of the system” up to that value, because **there is only so much evidence** that contradicts that hypothesis.

证据理论基本概念

- 例：交通信号灯，可能状态：X: {Green, Yellow, Red}
  - 假设：2^X: $\{\varnothing,\{G\},\{Y\},\{R\},\{G,Y\},\{Y,R\},\{G,R\},\{R,Y,R\}\}$

Hypothesis	Mass	Belief	Plausibility
Null	0		
Green	0.15		
Yellow	0.25		
Red	0.35		
Green or Yellow	0.04		
Yellow or Red	0.06		
Green or Red	0.05		
Any	0.1		

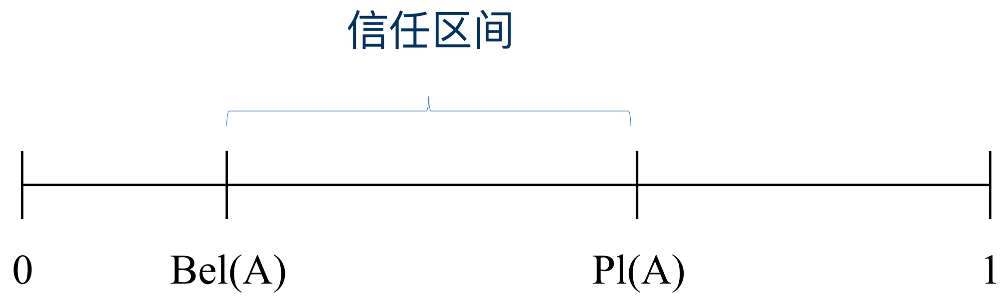
$$Bel(A) = \sum_{B \subseteq A} m(B)\,, \forall A \in 2^X$$

$$Pl(A) = \sum_{B | B \cap A \neq \emptyset} m(B)$$

$$= 1 - Bel(\neg A)$$

证据理论基本概念

- 信任区间：对假设A的信任区间$\left[Bel(A),\ Pl(A)\right]$



- 表示对某个假设的确认程度 (一种区间不确定)

Dempster合成规则

- 例子：共三种情况：A作案，B作案，C作案
  - 令m_1, m_2分别标记两组mass函数

假设	张三认为	李四认为
A作案	0.86	0.02
B作案	0.13	0.90
C作案	0.01	0.08

  - 如何得到?

假设	综合概率
A作案	0.1274
B作案	0.8667
C作案	0.0059

Dempster合成规则

•Dempster合成规则 (Dempster's combinational rule) 也称证据合成公式，其定义如下：

- 对于$\forall A \subseteq X$上的两组mass函数$m1$，$m2$，其对应的Dempster合成规则为：

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

- 其中，K为归一化常数

$$K = \sum_{B \cap C \neq \varnothing} m_1(B) \cdot m_2(C) = 1 - \sum_{B \cap C = \varnothing} m_1(B) \cdot m_2(C)$$

Dempster合成规则

假设	张三认为	李四认为
Adam作案	0.86	0.02
Ben作案	0.13	0.90
Clark作案	0.01	0.08

•解：首先计算归一化常数K

$$K = \sum_{B \cap C \neq \varnothing} m_1(B) \cdot m_2(C)$$

$$= m_1(Adam) \cdot m_2(Adam) + m_1(Ben) \cdot m_2(Ben) + m_1(Clark) \cdot m_2(Clark)$$

$$= 0.86 \times 0.02 + 0.13 \times 0.90 + 0.01 \times 0.08 = 0.135$$

思考：为什么不考虑如{Adam，Ben}这样的状态？

Dempster合成规则

•解：其次，关于Adam的组合mass函数：

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

$$= \frac{1}{K} \cdot m_1(\{Adam) \cdot m_2(\{Adam\})$$

$$= \frac{1}{0.135} \times 0.86 \times 0.02 = 0.1274$$

Dempster合成规则

- 解：关于Ben的组合mass函数：

$$m_1 \oplus m_2(\{\text{Ben}\}) = \frac{1}{K} \sum_{B \cap C = \{Ben\}} m_1(B) \cdot m_2(C)$$

$$= \frac{1}{K} \cdot m_1(\{Ben) \cdot m_2(\{Ben\})$$

$$= \frac{1}{0.135} \times 0.13 \times 0.90 = 0.8667$$

Dempster合成规则

• 解：关于Clark的组合mass函数：

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

$$= \frac{1}{K} \cdot m_1(\{Clark\}) \cdot m_2(\{Clark\})$$

$$= \frac{1}{0.135} \times 0.08 \times 0.01 = 0.0059$$

Dempster合成规则

- 合成后的mass函数m_{12}

假设	综合概率
Adam作案	0.1274
Ben作案	0.8667
Clark作案	0.0059

- 对于这个简单的实例而言，对于Adam, Ben, Clark的组合mass函数，再求信任函数、似然函数：
  - Bel({Adam}) = 0.1274， Pl({Adam})= 0.1274
  - Bel({Ben}) =0.8667， Pl({Ben})= 0.8667
  - Bel({Clark}) = 0.0059， Pl({Clark})= 0.0059

思考：为什么Bel和Pl相同?

*广义贝叶斯理论的解释

- 当mass函数m中的所有焦元都是单点集（即单个假设集），且这些焦元都满足Bayes独立条件时，Dempster证据合成公式就退化为Bayes公式，所以，
  - ♦ Bayes公式是Dempster证据合成公式的特例。
- 反过来说，
  - ♦ Dempster证据合成公式是Bayes公式的广义化。

条件独立式多随机变量的重要概念。定义如下，若

$$p(a|b,c) = p(a|c). \tag{1}$$

我们就说，a 与 b 关于 c 条件独立。条件独立的符号表示如下：

$$a \perp\!\!\!\perp b \mid c$$

Dempster合成规则

n个mass函数的Dempster合成规则

- 对于∀A⊆Θ，识别框架Θ上的有限个mass函数$m_1, m_2, ..., m_n$的Dempster合成规则为：

$$(m_1 \oplus m_2 \oplus \cdots \oplus m_n)(A) = \frac{1}{K} \sum_{A_1 \cap A_2 \cap \cdots \cap A_n = A} m_1(A_1) \cdot m_2(A_2) \cdots m_n(A_n)$$

- 其中

$$K = \sum_{A_1 \cap \cdots \cap A_n \neq \varnothing} m_1(A_1) \cdot m_2(A_2) \cdots m_n(A_n)$$

$$= 1 - \sum_{A_1 \cap \cdots \cap A_n = \varnothing} m_1(A_1) \cdot m_2(A_2) \cdots m_n(A_n)$$

证据理论的冲突：Zadeh悖论

- 如果证据相互矛盾：

假设	张三认为	李四认为	合成
Adam作案	0.99	0	0.00
Ben作案	0.01	0.01	1.00
Clark作案	0	0.99	0.00

- 这显然违背人们的直觉！

证据理论的冲突：Zadeh悖论

- 增加数据，消解Zaheh悖论：

假设	张三认为	李四认为	合成
{Adam}	0.98	0	0.49
{Ben}	0.01	0.01	0.015
{Clark}	0	0.98	0.49
Θ ={Adam, Ben, Clark}	0.01	0.01	0.005

- 大家可以自行套公式计算一下

为什么需要证据理论

特性：

- 证据理论基于人们对客观世界的认识，根据人们掌握的证据和知识，对不确定性事件给出不确定性度量。
  - 证据理论使得不确定性度量更切近人们的习惯，易于使用。
- 由于在证据理论中需要的先验数据比概率推理理论中的更为直观、更容易获得，再加上Dempster合成公式可以综合不同专家或数据源的知识或数据，这使得证据理论在专家系统、信息融合等领域中得到了广泛应用。

基于证据理论的推理

- 基于证据理论的不确定性推理的步骤：

  （1）建立问题的样本空间Θ

  （2）由经验给出，或者由随机性规则和事实的信度度量计算基本概率分配函数(mass函数)

  （3）将不同来源的mass函数进行组合

  （4）计算所关心子集的信任函数值Bel(A)或似然函数值Pl(A)

  （5）由信任函数值或似然函数值得出结论

本章主要内容

- 不确定知识表示与推理

➢ 不确定知识表示与推理的概念

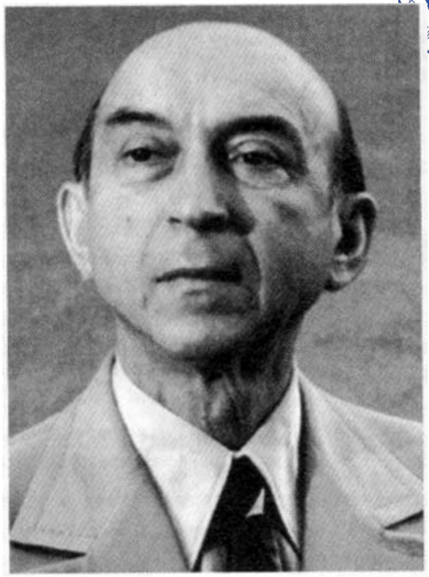
➢ CF模型

➢ DS证据理论

➢ 模糊推理

模糊推理方法

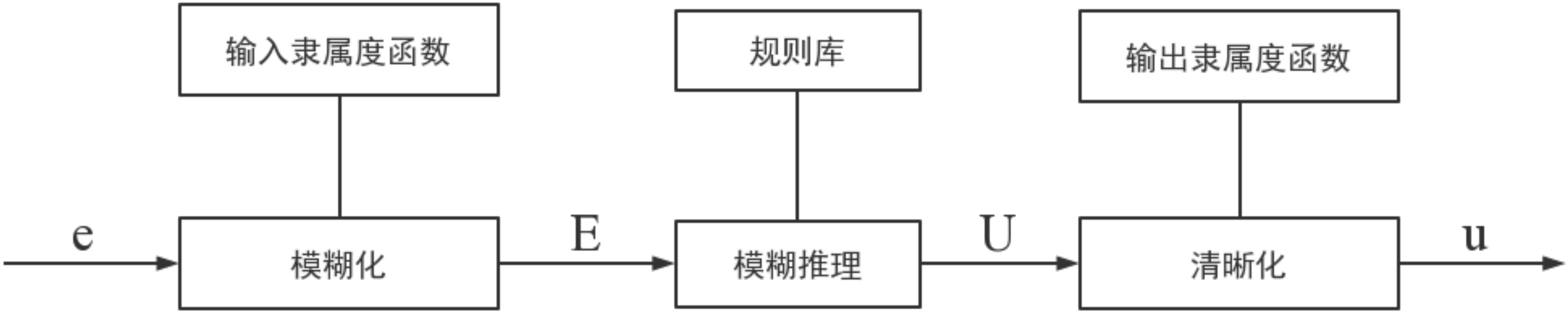
- 1965年，美国L.A.Zadeh发表了“fuzzy set”的论文，首先提出了模糊理论
- 模糊控制以模糊数学为基础，运用语言规则知识表示和计算机技术，由模糊推理进行决策，是人工智能的一个重要分支



Lotfi A. Zadeh

HUAZHONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

模糊推理过程



模糊集合和隶属函数

- **模糊集合是经典集合的扩充**
  - 经典集合中，元素a和集合A的关系只有两种：a属于A或a不属于A，即只有两个真值“真”/“假”。
  - 模糊逻辑中，引入**隶属度**的概念，描述介于真/假中间的过程
  - 模糊逻辑给集合中每一个元素赋予一个介于0和1之间的实数，描述其属于一个集合的强度，该实数称为**元素属于一个集合的隶属度**。集合中所有元素的隶属度全体构成集合的**隶属函数**。

$$u_A(x) = \begin{cases} 1 & x \in A \\ (0,1) & x\text{属于}A\text{的程度} \\ 0 & x \notin A \end{cases}$$

- 其中A称为模糊集合，由0、1及$u_A(x)$ 构成。
- $u_A(x)$表示元素x属于模糊集合A的程度取值范围为[0,1]，称 $u_A(x)$为x属于模糊集合A的隶属度。

华中科技大学 HUAZHONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

模糊集合的表示方法

- 与经典集合不同，模糊集合不仅要列出属于集合的元素，还要注明元素属于集合的隶属度
- 当论域中元素数目有限时，模糊集合A的数学描述为：

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

其中，$\mu_A(x)$为元素属于模糊集A的隶属度，X是元素x的论域。

模糊集合的表示方法

- 论域离散且元素数有限：Zadeh表示法

$$A=\{\mu_A(x_1)/x_1,\mu_A(x_2)/x_2,\cdots,\mu_A(x_n)/x_n\}$$

- 论域连续或元素数无限

$$A=\int_{x\in U}\mu_A(x)/x$$

- 序偶表示法

$$A=\{(\mu_A(x_1),x_1),(\mu_A(x_2),x_2),\cdots,(\mu_A(x_n),x_n)\}$$

- 向量表示法

$$A=\{\mu_A(x_1),\mu_A(x_2),\cdots,\mu_A(x_n)\}$$

模糊集合的表示方法

- 例：设论域U={张三，李四，王五}，评语为“学习好”。设三个人学习成绩总评分是：张三95分，李四90分，王五85分，三人都学习好，但又有差异。

- 采用隶属函数$u_A(x)=x/100$，由三个人的成绩可知三人“学习好”的隶属度分别为：

$$u_A(\text{张三})=0.95,\quad u_A(\text{李四})=0.90,\quad u_A(\text{王五})=0.85$$

- 用“学习好”这一模糊子集A可以表示为：$A=\{0.95, 0.90, 0.85\}$

（1）模糊集合的包含关系

- 若 $\mu_A(x) \geq \mu_B(x)$ ，则 $A \supseteq B$

（2）模糊集合的相等关系

- 若 $\mu_A(x) = \mu_B(x)$ ，则 $A = B$

（3）模糊集合的交并补运算

∨ 表示取大运算
∧ 表示取小运算

① 交运算(intersection) $A \cap B$

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

② 并运算(union) $A \cup B$

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

③ 补运算(complement) $\overline{A}$ 或者 A^C

$$\mu_{\overline{A}}(x) = 1 - \mu_A(x)$$

模糊集合的运算

(4) 模糊集合的代数运算

① 代数积： $\mu_{AB}(x) = \mu_A(x)\mu_B(x)$

② 代数和： $\mu_{A+B}(x) = \mu_A(x) + \mu_B(x) - \mu_{AB}(x)$

③ 有界和： $\mu_{A\oplus B}(x) = \min\{1, \mu_A(x) + \mu_B(x)\} = 1 \wedge [\mu_A(x) + \mu_B(x)]$

④ 有界积： $\mu_{A\otimes B}(x) = \max\{0, \mu_A(x) + \mu_B(x) - 1\} = 0 \vee [\mu_A(x) + \mu_B(x) - 1]$

模糊关系

- 普通关系描述两个集合中的元素之间是否有关系
- 模糊关系描述两个模糊集合中的元素之间的关联程度

模糊关系的定义 ：

A、B 为两个模糊集合，模糊关系用叉积(cartesian product)表示：

$$R : A \times B \to [0,1]$$

叉积常用最小算子运算：

$$\mu_{A\times B}(a,b) = \min\{\mu_A(a), \mu_B(b)\}$$

若A、B 为离散模糊集，其隶属函数分别为：

$$\mu_A = [\mu_A(a_1), \mu_A(a_2), \cdots, \mu_A(a_n)], \qquad \mu_B = [\mu_B(b_1), \mu_B(b_2), \cdots, \mu_B(b_n)]$$

则其叉积运算： $\mu_{A\times B}(a,b) = \mu_A^T \circ \mu_B$

两个模糊向量的叉积类似于向量乘积，只是乘法运算用取小运算代替

模糊关系

- 当论域为有限时，可以用模糊矩阵来表示模糊关系

例：某地区人的身高论域X={140,150,160,170,180}（单位：cm），体重论域Y={40,50,60,70,80}（单位：kg）。

身高与体重的模糊关系表

R (X \ Y)	40	50	60	70	80
140	1	0.8	0.2	0.1	0
150	0.8	1	0.8	0.2	0.1
160	0.2	0.8	1	0.8	0.2
170	0.1	0.2	0.8	1	0.8
180	0	0.1	0.2	0.8	1

从X到Y的一个模糊关系R，用模糊矩阵表示：

$$R=\begin{bmatrix} 1 & 0.8 & 0.2 & 0.1 & 0 \\ 0.8 & 1 & 0.8 & 0.2 & 0.1 \\ 0.2 & 0.8 & 1 & 0.8 & 0.2 \\ 0.1 & 0.2 & 0.8 & 1 & 0.8 \\ 0 & 0.1 & 0.2 & 0.8 & 1 \end{bmatrix}$$

模糊关系的合成

- 设模糊关系 $Q \in X \times Y, R \in Y \times Z$，则模糊关系 $S \in X \times Z$ 称为Q与R的合成。
- S等于模糊矩阵Q与R的叉乘
- 模糊矩阵的合成常用计算方法：
  - 最大-最小合成法：写出矩阵乘积QR中的每个元素，然后将其中的乘积运算用取小运算代替，求和用最大运算代替
  - 最大-代数积合成法：写出矩阵乘积QR中的每个元素，然后将其中的求和运算用取大运算代替，乘积运算不变

模糊关系的合成

$$X=\{x_1,x_2x_3,x_4\},\quad Y=\{y_1,y_2,y_3\},\quad Z=\{z_1,z_2\}$$

$$Q\in X\times Y,\quad R\in Y\times Z,\quad S\in X\times Z,\quad \text{求}S。$$

最大-最小合成法：
其中的乘积运算用取小运算代替，求和用最大运算代替

$$Q=\begin{bmatrix}0.5 & 0.6 & 0.3\\ 0.7 & 0.4 & 1\\ 0 & 0.8 & 0\\ 1 & 0.2 & 0.9\end{bmatrix}$$

$$R=\begin{bmatrix}0.2 & 1\\ 0.8 & 0.4\\ 0.5 & 0.3\end{bmatrix}$$

$$S=Q\circ R=\begin{bmatrix}0.5 & 0.6 & 0.3\\ 0.7 & 0.4 & 1\\ 0 & 0.8 & 0\\ 1 & 0.2 & 0.9\end{bmatrix}\circ\begin{bmatrix}0.2 & 1\\ 0.8 & 0.4\\ 0.5 & 0.3\end{bmatrix}$$

$$=\begin{bmatrix}(0.5\wedge0.2)\vee(0.6\wedge0.8)\vee(0.3\wedge0.5) & (0.5\wedge1)\vee(0.6\wedge0.4)\vee(0.3\wedge0.3)\\ (0.7\wedge0.2)\vee(0.4\wedge0.8)\vee(1\wedge0.5) & (0.7\wedge1)\vee(0.4\wedge0.4)\vee(1\wedge0.3)\\ (0\wedge0.2)\vee(0.8\wedge0.8)\vee(0\wedge0.5) & (0\wedge1)\vee(0.8\wedge0.4)\vee(0\wedge0.3)\\ (1\wedge0.2)\vee(0.2\wedge0.8)\vee(0.9\wedge0.5) & (1\wedge1)\vee(0.2\wedge0.4)\vee(0.9\wedge0.3)\end{bmatrix}$$

$$=\begin{bmatrix}0.6 & 0.5\\ 0.5 & 0.7\\ 0.8 & 0.4\\ 0.5 & 1\end{bmatrix}$$

模糊推理

- 模糊知识表示

  （<对象>，<属性>，（<属性值>，<隶属度>））

  - 例：张三比较胖：（张三，体型，（胖，0.9））

- 扩充到产生式规则、语义网络等中

  - 例：如果患者有些头疼并且发高烧 则 他患了重感冒

（患者，症状，（头疼，0.3））∧（患者，症状，（发烧，0.9））→（患者，疾病，（感冒，0.9））

许多模糊规则可以表示为从条件论域到结论论域的模糊关系矩阵R。通过条件模糊向量与模糊关系 R 的合成进行模糊推理，得到结论的模糊向量，然后采用“清晰化”方法将模糊结论转换为精确量。

对 IF A THEN B类型的模糊规则的推理

若已知输入为A，则输出为B；若现在已知输入为A'，则输出B'用合成规则求取为：$B' = A' \circ R$

其中A 到B模糊的关系R: $\mu_R(x,y) = \min[\mu_A(x), \mu_B(y)]$

控制规则库的N条规则有N个模糊关系： $R_1, R_2, \cdots, R_n$

对于整个系统的全部控制规则所对应的模糊关系R：

$$R = R_1 \cup R_2 \cup \cdots \cup R_n = \bigcup_{i=1}^{n} R_i$$

对 IF A THEN B类型的模糊规则的推理

- 例：已知输入的模糊集合A和输出的模糊集合B：

$$A = 1.0/a_1 + 0.8/a_2 + 0.5/a_3 + 0.2/a_4 + 0.0/a_5$$

$$B = 0.7/b_1 + 1.0/b_2 + 0.6/b_3 + 0.0/b_4$$

当输入为： $A' = 0.4/a_1 + 0.7/a_2 + 1.0/a_3 + 0.6/a_4 + 0.0/a_5$ 求输出B'

解： 求A到B的模糊关系R：

$$R = A \times B = \mu_A^{\mathrm{T}} \circ \mu_B = \begin{bmatrix} 1.0 \\ 0.8 \\ 0.5 \\ 0.2 \\ 0.0 \end{bmatrix} \circ [0.7 \quad 1.0 \quad 0.6 \quad 0.0]$$

对 IF A THEN B类型的模糊规则的推理

- 例：已知输入的模糊集合A和输出的模糊集合B：

$$A = 1.0/a_1 + 0.8/a_2 + 0.5/a_3 + 0.2/a_4 + 0.0/a_5$$

$$B = 0.7/b_1 + 1.0/b_2 + 0.6/b_3 + 0.0/b_4$$

当输入为： $A' = 0.4/a_1 + 0.7/a_2 + 1.0/a_3 + 0.6/a_4 + 0.0/a_5$ 求输出B'

解： 求A到B的模糊关系R：

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

对 IF A THEN B类型的模糊规则的推理

$$B' = A' \circ R = \boxed{\begin{bmatrix} 0.4 & 0.7 & 1.0 & 0.6 & 0.0 \end{bmatrix}} \circ \begin{bmatrix} 0.7 & 1.0 & 0.6 & 0.0 \\ 0.7 & 0.8 & 0.6 & 0.0 \\ 0.5 & 0.5 & 0.5 & 0.0 \\ 0.2 & 0.2 & 0.2 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix} = (0.7, 0.7, 0.6, 0.0)$$

$$\boxed{(0.4 \wedge 0.7) \vee (0.7 \wedge 0.7) \vee (1.0 \wedge 0.5) \vee (0.6 \wedge 0.2) \vee (0.0 \wedge 0.0)}$$

$\vee$ 表示取大运算

$\wedge$ 表示取小运算

模糊决策

- 模糊决策（模糊判决/解模糊/清晰化）：由模糊推理得到的结论或者操作是一个模糊向量，需要转化为确定值
  - 最大隶属度法：取隶属度最大的量作为推理结果
  - 加权平均判决法：将各隶属度作为权值进行加权平均
  - 中位数法：论域上把隶属函数曲线与横坐标转成的面积平分为两部分的元素称为模糊集的中位数

模糊决策

- 例： 设有模糊控制规则："如果温度低，则将风门开大"。

  设温度和风门开度的论域为{1，2，3，4，5}。

  "温度低"和"风门大"的模糊量：

  "温度低"= $1/1+0.6/2+0.3/3+0.0/4+0/5$

  "风门大" = $0/1+0.0/2+0.3/3+0.6/4+1/5$

  已知事实"温度较低"，可以表示为

  "温度较低"= $0.8/1+1/2+0.6/3+0.3/4+0/5$

  试用模糊推理确定风门开度。

模糊决策

解：（1）确定模糊关系 R

$$R = \begin{bmatrix} 1.0 \\ 0.6 \\ 0.3 \\ 0.0 \\ 0.0 \end{bmatrix} \circ \begin{bmatrix} 0.0 & 0.0 & 0.3 & 0.6 & 1.0 \end{bmatrix}$$

$$= \begin{bmatrix} 0.0 & 0.0 & 0.3 & 0.6 & 1.0 \\ 0.0 & 0.0 & 0.3 & 0.6 & 0.6 \\ 0.0 & 0.0 & 0.3 & 0.3 & 0.3 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix}$$

模糊决策

（2）模糊推理

$$B' = A' \circ R = \begin{bmatrix} 0.8 & 1.0 & 0.6 & 0.3 & 0.0 \end{bmatrix} \circ \begin{bmatrix} 0.0 & 0.0 & 0.3 & 0.6 & 1.0 \\ 0.0 & 0.0 & 0.3 & 0.6 & 0.6 \\ 0.0 & 0.0 & 0.3 & 0.3 & 0.3 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix}$$

$$= \begin{bmatrix} 0.0 & 0.0 & 0.3 & 0.6 & 0.8 \end{bmatrix}$$

（3）模糊决策

- ➢用最大隶属度法：风门开度为5
- ➢用加权平均法：(0.3*3+0.6*4+0.8*5)/(0.3+0.6+0.8)=4.29 风门开度为4

作业1

- 已知嫌疑人A、B、C三人中只有一人是罪犯，且罪犯一定会说谎，而好人一定说真话。设警察向嫌疑人A、B、C三人各自进行问询：
- A回答：B或C中至少有一人是罪犯；
- B回答：A是罪犯；
- C回答：A和我都是好人。

- 试用谓词逻辑推理证明：B是罪犯。

作业2

- 设有如下一组推理规则：

r1: IF E1 AND E2 THEN E3 (0.9)

r2: IF E3 OR E4 THEN E5 (0.8)

r3: IF E5 THEN H (0.7)

r4: IF E6 THEN H (0.9)

- 已知 CF(E1) = 0.6，CF(E2) = 0.5，CF(E4) = 0.4，CF(E6) = 0.8，求CF(H)。

作业3

- 警方判定罪犯肯定是嫌疑人A、B中的一个，但不知道是哪一个；
- 经过问询，两个证人张三、李四对他们认为的罪犯给出如下不同的判断：

假设	张三认为 m_1	李四认为 m_2
A作案	0.8	0.6
B作案	0.1	0.2
A或B作案	0.1	0.2

- 试使用D-S理论分析最有可能的罪犯。

作业4

- 设有模糊规则：如果天气冷，则调高空调温度。

- 设室温的论域为{0, 5, 10}，单位摄氏度，隶属度 [1.0, 0.8, 0.5]
- 调高空调论域为{1, 3, 5}，单位摄氏度，隶属度 [0.2, 0.5, 1.0]

- 已知事实"天气冷"={(0.1, 0), (0.6, 5), (0.2, 10)}

- 试用模糊推理确定空调应该怎么调。