第14讲 不确定性推理

第14讲 不确定性推理

14.1 贝叶斯推理

- 贝叶斯方法思想
- 贝叶斯规则
- 贝叶斯网络推理

14.2 模糊逻辑推理

- 模糊逻辑
- 隶属函数和隶属度
- 模糊推理

概率基础

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

Bayes公式

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

联合概率分布

$P(T,W)$

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

边缘概率分布

$P(T)$

T	P
hot	0.5
cold	0.5

$P(W)$

W	P
sun	0.6
rain	0.4

条件概率分布

$P(W|T)$

$P(W|T=hot)$

W	P
sun	0.8
rain	0.2

$P(W|T=cold)$

W	P
sun	0.4
rain	0.6

14.1 贝叶斯推理-应用贝叶斯规则

- 因果 $P(\text{effect}|\text{cause})$

- 诊断

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

- 条件独立的完全联合分布为：

$$\mathbf{P}(Cause, Effect_1, Effect_2, \cdots, Effect_n) = \mathbf{P}(Cause)\prod_i \mathbf{P}(Effect_i \mid Cause)$$

14.1 贝叶斯推理-例：拼写检查

在网上搜索，当不小心输入一个不存在的单词时，搜索引擎会提示你是不是要输入某一个正确的单词，比如当你在Google中输入“Julw”时，系统会猜测你的意图：是不是要搜索“July”，如下图所示：

14.1 贝叶斯推理-例：拼写检查

如果把拼写正确的情况记做c（代表correct），拼写错误的情况记做w（代表wrong），那么“拼写检查”要做的事情就是：在发生w的情况下，试图推断出c。换言之：已知w，然后在若干个备选方案中，找出可能性最大的那个c，也就是求P(c|*w*) 的最大值。根据贝叶斯定理：

$$P(c|w) = P(w|c) * \frac{P(c)}{P(w)}$$

由于对于所有备选的c来说，对应的都是同一个w，所以它们的*P(w)*是相同的，因此我们只要最大化$P(w|c) * P(c)$

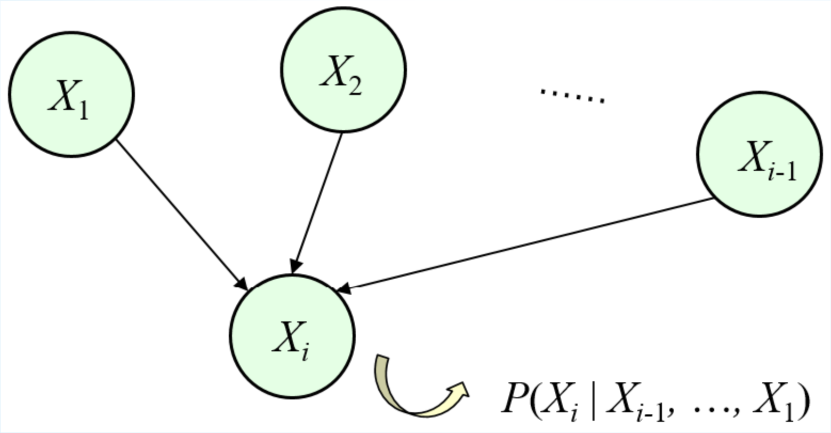
*P(c)*表示某个正确的词的出现"概率"，它可以用"频率"代替。如果我们有一个足够大的文本库，那么这个文本库中每个单词的出现频率，就相当于它的发生概率。

P(*w*|*c*)表示在试图拼写c的情况下，出现拼写错误w的概率。

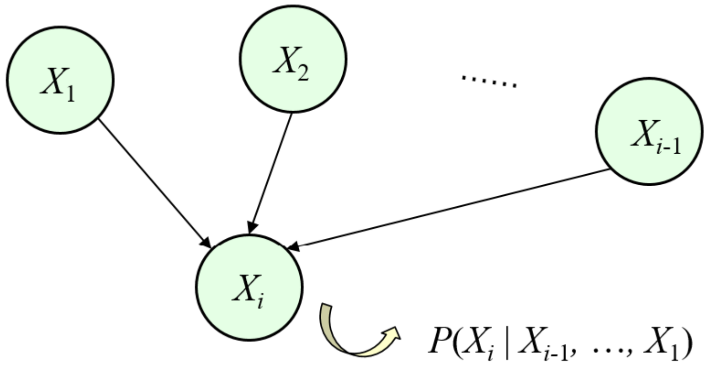
14.1 贝叶斯推理-贝叶斯网络

简单的因果关系可采用贝叶斯规则，而复杂的关系则需要基于贝叶斯网络。

贝叶斯网络 (Bayesian network)，又称信念网络(Belief Network)，或有向无环图模型(directed acyclic graphical model)，是一种概率图模型，由Judea Pearl首先提出。是一种模拟人类推理过程中因果关系的不确定性处理模型，其网络拓扑结构是一个有向无环图(DAG)。



14.1 贝叶斯推理-贝叶斯网络



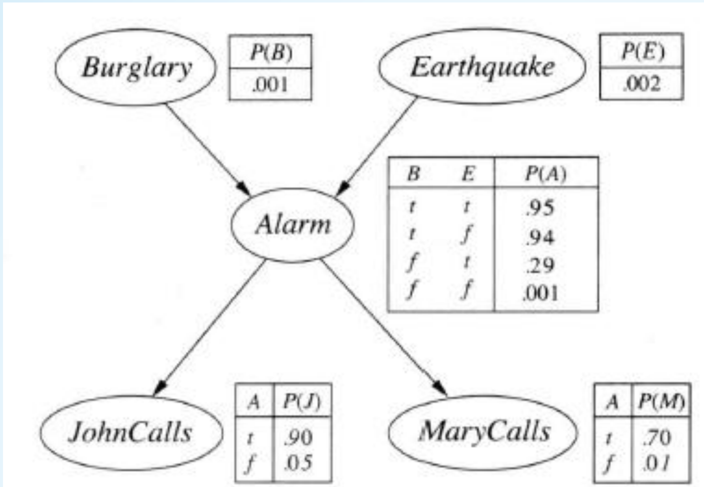
定义：贝叶斯网络是一种采用有向非循环图表示随机变量及其条件相关性的网络。包含三个要素：

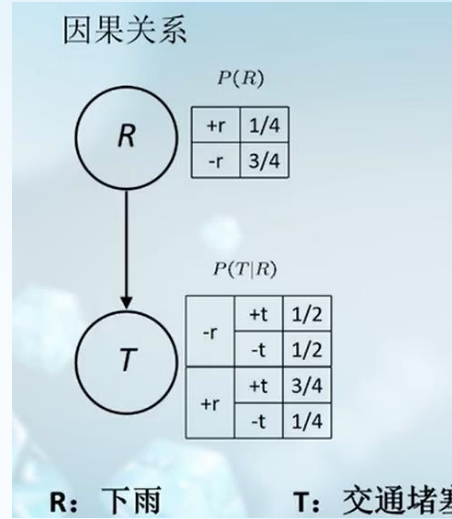
1）一组节点，每个节点对应于一个随机变量；

2）这些节点之间的有向连接；

3）每个节点在给定父节点下的条件概率分布，即 $\boldsymbol{P}(X_i|Parents(X_i))$。

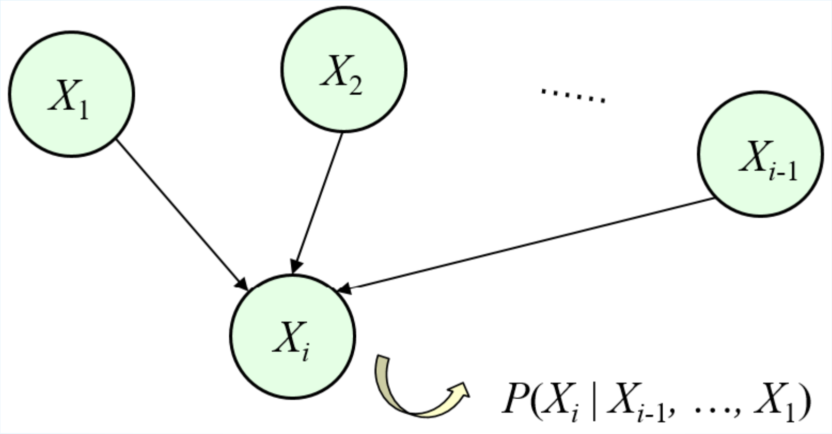
14.1 贝叶斯推理- 例：警报网络





- 拓扑信息
- 条件概率表 CPT(conditional probability table)：
  - ✓ 用于离散随机变量
  - ✓ 每一行包含了节点的每个取值对于一个条件事件的条件概率
  - ✓ 每行概率和为1，对于布尔变量省略一列
  - ✓ 具有k个布尔父节点的布尔变量，CPT中有？ 2^k个可独立指定的概率

14.1 贝叶斯推理-贝叶斯网络的语义



贝叶斯网络语义：可以表示完全联合概率分布

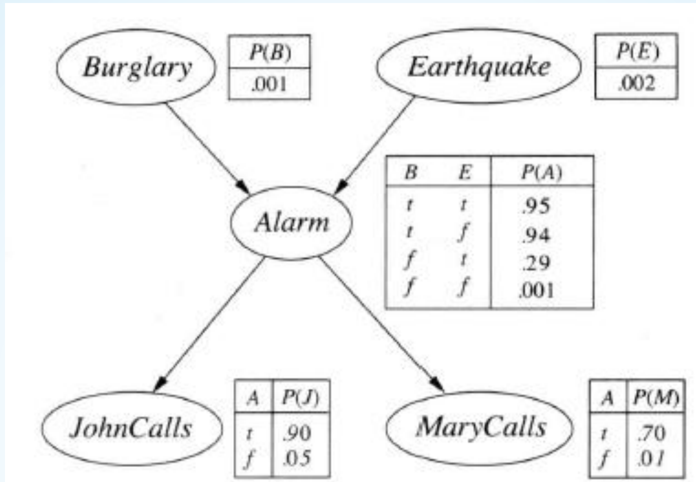
n个随机变量的联合概率分布（Joint probability distribution）：

$$\boldsymbol{P}(X_1, \ldots, X_n) = \boldsymbol{P}(X_n | X_{n-1}, \ldots, X_1) \ldots \ldots \boldsymbol{P}(X_2 | X_1) P(X_1)$$

$$= \prod_{i=1}^{n} \boldsymbol{P}(X_i | X_{i-1}, \ldots, X_1)$$

亦称为一组随机变量的链式规则（Chain rule）。

14.1 贝叶斯推理-贝叶斯网络语义的应用



报警器响了，但既没盗贼也没地震，两人都打电话的概率：

$$P(j,m,a,\neg b,\neg e) = P(j|\,a)P(m|\,a)P(a|\,\neg b \wedge \neg e)P(\neg b)P(\neg e) = 0.90 \times 0.70 \times 0.001 \times 0.999 \times 0.998 = 0.000628$$

报警器没响，既没盗贼也没地震，两人都没打电话的概率？

14.1 贝叶斯推理 – 概率推理

概率推理: 从已知概率分布计算目标概率分布。

通常基于联合概率求解条件概率。

- P (on time | no reported accidents) $= 0.90$
- 表示，给定一个证据，一个智能体对于某事件发生的置信度有多少。

随着新证据出现, 概率发生变化:

- P (on time | no accidents, 5 a. m.) $= 0.95$
- P (on time | no accidents, 5 a. m., raining) $= 0.80$
- 观察到新的证据后, 置信度随之更新。

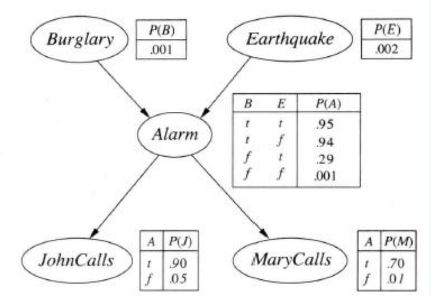
14.1 贝叶斯推理

概率推理系统的基本任务是要在给定某个已观察到的事件（**证据变量**）后，计算一组**查询变量**的后验概率分布。

- ✓ Q表示查询变量；
- ✓ ***E***表示证据变量集；
- ✓ ***H***表示既非证据也非查询变量集——隐藏变量。

这样，全部变量集合是Q∪***E***∪***H***。

典型的查询是询问后验概率***P***(Q|***E***)。



例：两人都打电话，出现盗贼的概率是？

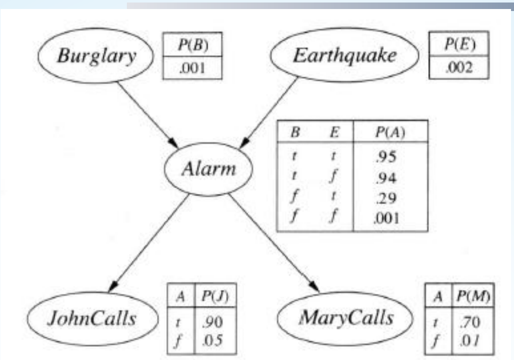
$P(B|+j,+m)$

查询变量：B

证据变量：J，M

隐藏变量：A，E

14.1 贝叶斯推理-贝叶斯网络中的枚举概率推理



消除隐藏变量，得到查询和证据变量的联合分布，然后归一化。

$$\begin{aligned} P(B \mid +j,+m) \propto_B\ & P(B,+j,+m) \\ &= \textstyle\sum_{e,a} P(B,e,a,+j,+m) \\ &= \textstyle\sum_{e,a} P(B)P(e)P(a \mid B,e)P(+j \mid a)P(+m \mid a) \end{aligned}$$

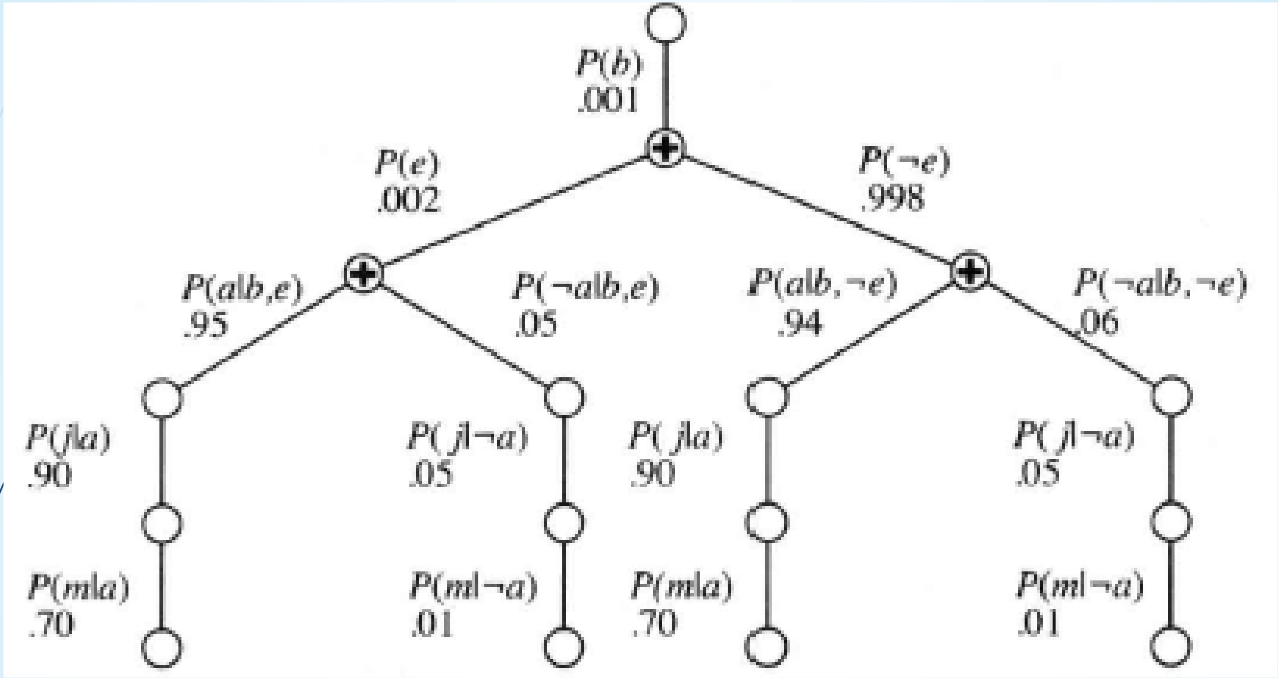
4项加和，每一个子项涉及5个元素的连乘。

$$\begin{aligned} &P(B,+j,+m) \\ &= P(B)P(+e)P(+a \mid B,+e)P(+j \mid +a)P(+m \mid +a) + P(B)P(+e)P(-a \mid B,+e)P(+j \mid -a)P(+m \mid -a) \\ &\quad P(B)P(-e)P(+a \mid B,-e)P(+j \mid +a)P(+m \mid +a) + P(B)P(-e)P(-a \mid B,-e)P(+j \mid -a)P(+m \mid -a) \end{aligned}$$

$P(b \mid +j,+m)=\alpha 0.00059224$ $\quad$ $P(\neg b \mid +j,+m)=\alpha 0.0014919$

$$P(B \mid +j,+m)=\langle 0.284, 0.716\rangle$$

14.1 贝叶斯推理-贝叶斯网络中的精确推理



先计算所有隐藏变量a下的后验概率，再计算隐藏变量e的所有情况下的后验概率。

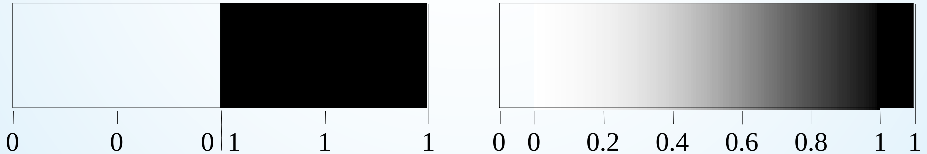
14.2 模糊推理-模糊逻辑(Fuzzy logic)

$\stackrel{?}{=}$

14.2 模糊推理-模糊逻辑(Fuzzy logic)

人的思维和认知，个子高、温度适中等，没有明确界限，模糊概念

二值逻辑有其局限性，比如 “个子高”

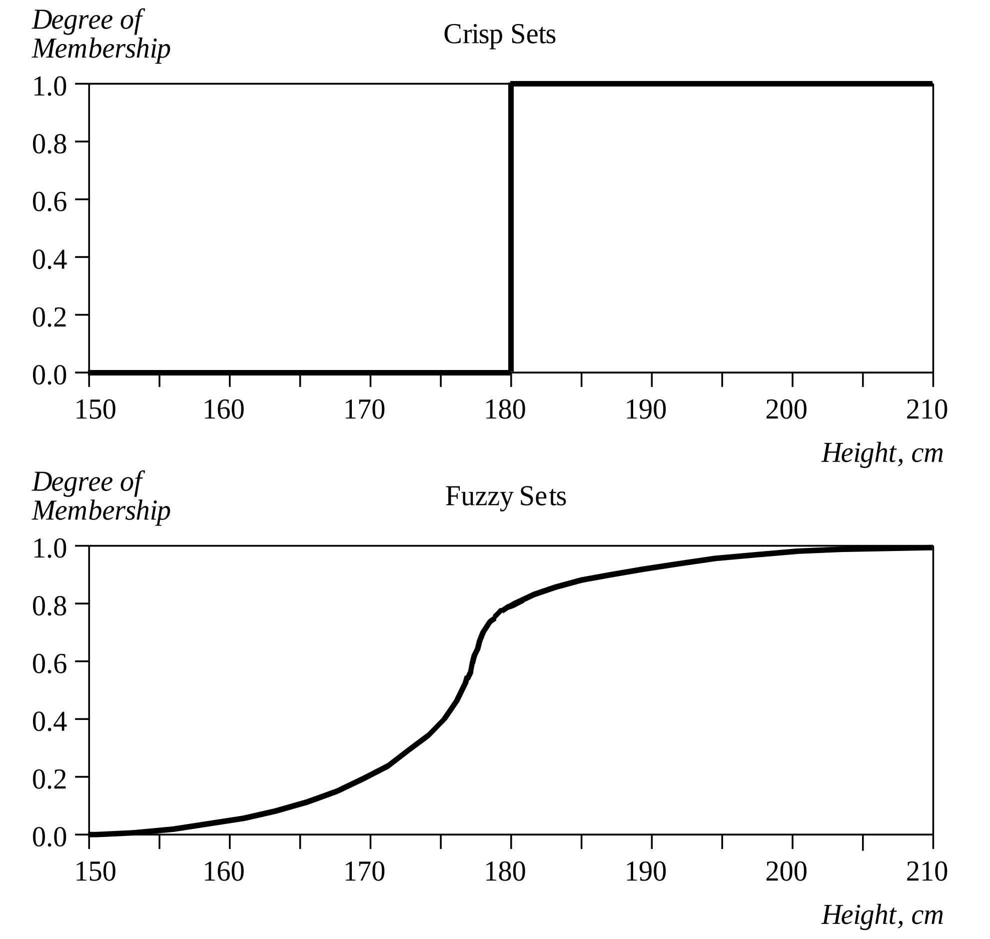


(*a*) Boolean Logic. (*b*) Multi-valued Logic.

14.2 模糊推理-模糊集合

Name	Height, cm	Degree of Membership	
		Crisp	*Fuzzy*
Chris	208	1	1.00
Mark	205	1	1.00
John	198	1	0.98
Tom	181	1	0.82
David	179	0	0.78
Mike	172	0	0.24
Bob	167	0	0.15
Steven	158	0	0.06
Bill	155	0	0.01
Peter	152	0	0.00

14.2 模糊推理-模糊集合



14.2 模糊推理-隶属函数和隶属度

- 若对论域U中任一元素u，都存在数$\tilde{A}(u) \in [0,1]$与之对应，则称$\tilde{A}$为U上的模糊子集，$\tilde{A}(u)$称为u对$\tilde{A}$的**隶属度**。当u在U中变动时，$\tilde{A}(u)$就是一个函数，称为$\tilde{A}$的**隶属函数**。
- $\tilde{A}(u)$越接近于1，表示u属于$\tilde{A}$的程度越高；
- $\tilde{A}(u)$越接近于0，表示u属于$\tilde{A}$的程度越低。

- **语言值** A：个子高

  $\tilde{A}(x_1) = 1.0, \tilde{A}(x_2) = 1.0, \tilde{A}(x_3) = 0.98, \dots, \tilde{A}(x_{10}) = 0.00$
- **模糊子集**$\tilde{A}$:

$$\tilde{A} = \frac{1.0}{x_1} + \frac{1.0}{x_2} + \frac{0.98}{x_3} + \cdots + \frac{0.00}{x_{10}}$$

14.2 模糊推理-隶属函数和隶属度

- 语言值 A：个子高

  $\widetilde{A}(x_1)=1.0, \widetilde{A}(x_2)=1.0, \widetilde{A}(x_3)=0.98, \dots, \widetilde{A}(x_{10})=0.00$

- 模糊子集$\widetilde{A}$:

$$\widetilde{A}=\frac{1.0}{x_1}+\frac{1.0}{x_2}+\frac{0.98}{x_3}+\cdots+\frac{0.00}{x_{10}}$$

- 语言值 B：性格开朗

  $\widetilde{B}(x_1)=1.0, \widetilde{B}(x_2)=0.1, \widetilde{B}(x_3)=0.5, \dots, \widetilde{B}(x_{10})=1.0$

- 模糊子集$\widetilde{B}$:

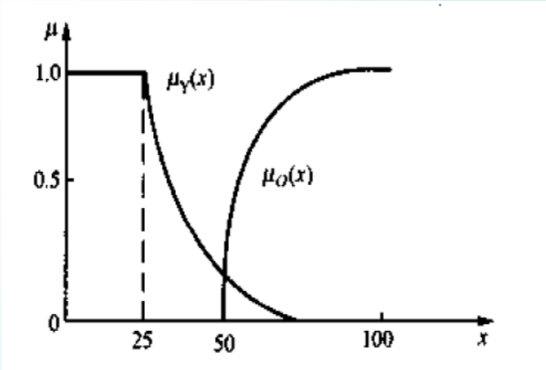
$$\widetilde{B}=\frac{1.0}{x_1}+\frac{0.1}{x_2}+\frac{0.5}{x_3}+\cdots+\frac{1.0}{x_{10}}$$

14.2 模糊推理-隶属函数和隶属度

- 以年龄为论域，设X=[0,100]，并设A表示模糊集合“年老”，B表示模糊集合“年轻”，则两者的隶属函数分别为：

$$\mu_A(x)=\begin{cases}0 & 0\le x\le 50\\ \dfrac{1}{1+(5/x-50)^2} & 50\le x\le 100\end{cases}$$

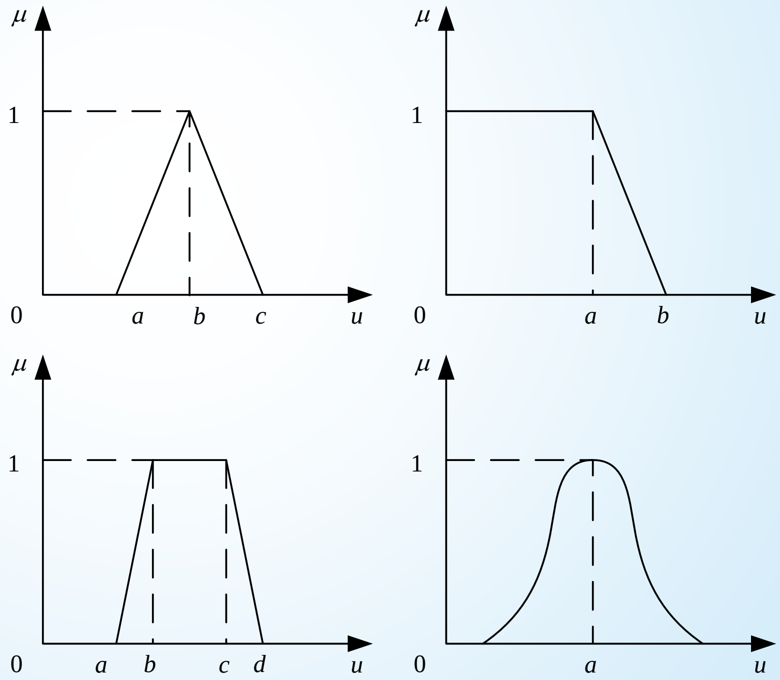
$$\mu_B(x)=\begin{cases}1 & 0\le x\le 25\\ \dfrac{1}{1+(x-25/5)^2} & 25\le x\le 100\end{cases}$$



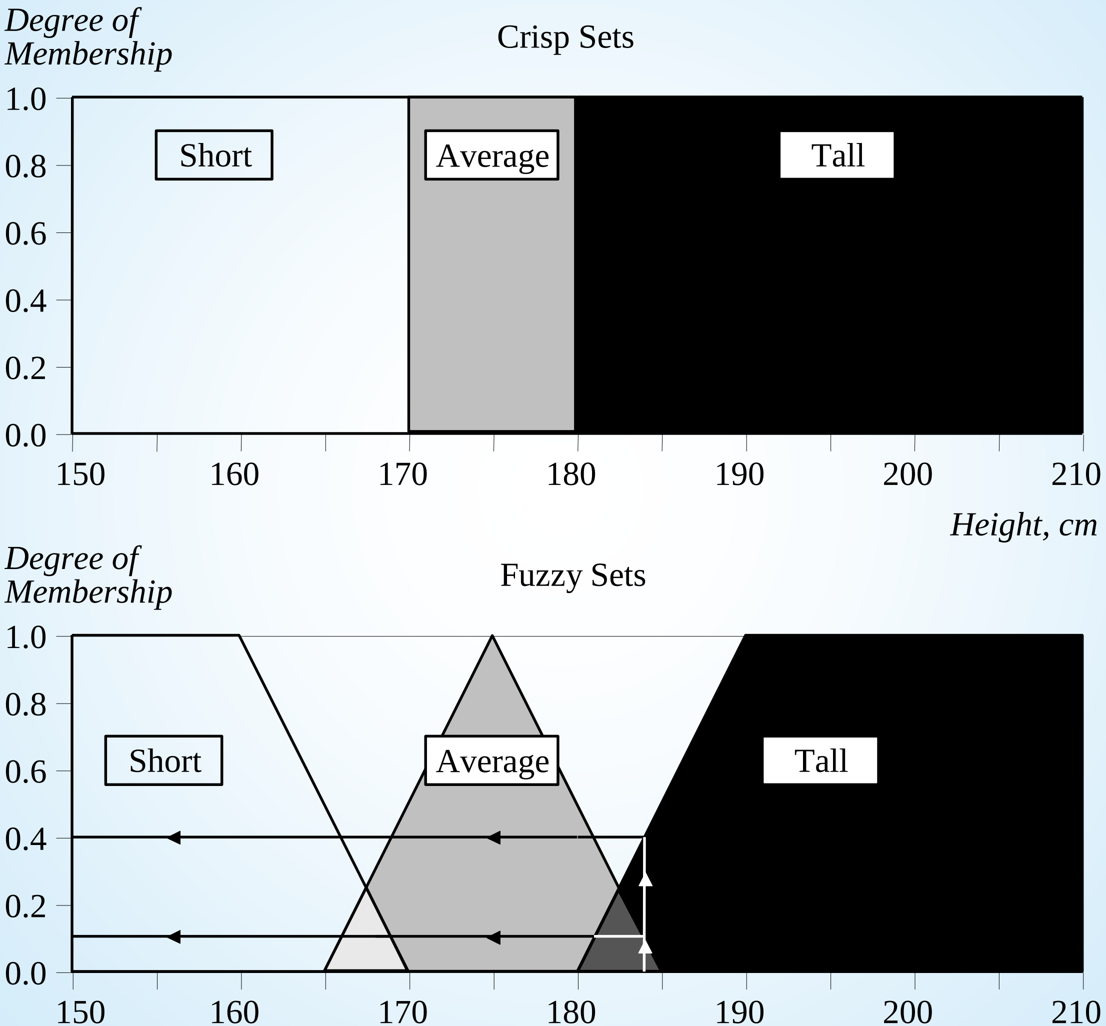
14.2 模糊推理-隶属函数和隶属度

常用的隶属度函数有三角形、梯形、高斯型、非对称型等。

例如：温度适宜



14.2 模糊推理-隶属函数和隶属度

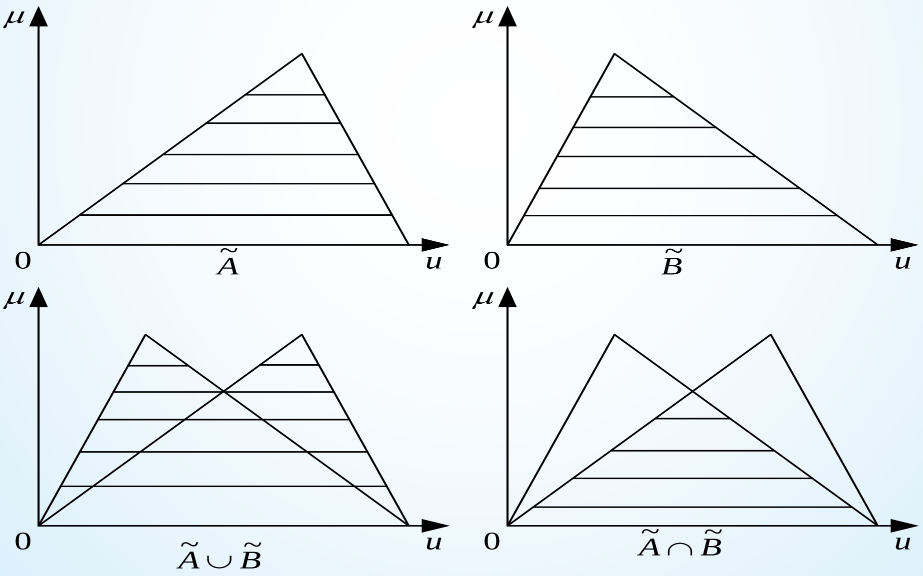


14.2 模糊推理-模糊集合的运算

并集(Union)：$(\tilde{A} \cup \tilde{B})(u) = \max\{\tilde{A}(u), \tilde{B}(u)\} = \tilde{A}(u) \vee \tilde{B}(u)$

交集(Intersection)：$(\tilde{A} \cap \tilde{B})(u) = \min\{\tilde{A}(u), \tilde{B}(u)\} = \tilde{A}(u) \wedge \tilde{B}(u)$

补集(Complement)：$\tilde{A}^{C}(u) = 1 - \tilde{A}(u)$



14.2 模糊推理-模糊集合的运算

例1：论域U={爷gp，奶gm，爸f，妈m，李四ls}

$$\tilde{A} = \text{"men"} = \frac{1}{gp} + \frac{0}{gm} + \frac{1}{f} + \frac{0}{m} + \frac{0}{ls}$$

$$\tilde{B} = \text{"young"} = \frac{0.1}{gp} + \frac{0.2}{gm} + \frac{0.9}{f} + \frac{1}{m} + \frac{1}{ls}$$

$$\tilde{A} \cap \tilde{B} = \text{"youngmen"} = \frac{0.1}{gp} + \frac{0}{gm} + \frac{0.9}{f} + \frac{0}{m} + \frac{0}{ls}$$

$$\tilde{A} \cup \tilde{B} = \text{"youngormen"} = \frac{1}{gp} + \frac{0.2}{gm} + \frac{1}{f} + \frac{1}{m} + \frac{1}{ls}$$

14.2 模糊推理-模糊集合的运算

例2：设A表示接近于1的实数，B表示接近于2的实数，隶属函数分别为：

$$\mu_{\tilde{A}}(x)=\frac{1}{1+(x-1)^2},\mu_{\tilde{B}}(x)=\frac{1}{1+(x-2)^2}$$

$\tilde{A}\cup\tilde{B}$表示接近于1或接近于2的实数；

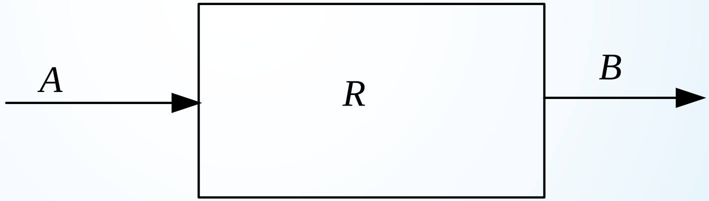
$\tilde{A}\cap\tilde{B}$表示既接近于1又接近于2的实数。

$$\mu_{\tilde{A}\cup\tilde{B}}(x)=\begin{cases}\dfrac{1}{1+(x-1)^2}, & x\le 1.5\\ \dfrac{1}{1+(x-2)^2}, & x>1.5\end{cases}$$

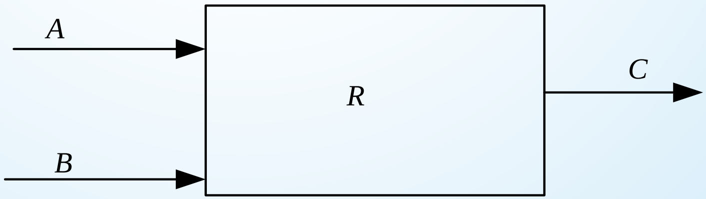
14.2 模糊推理-模糊推理

- 前向推理（模糊蕴含关系implication）
  - ✓ 前提1：如果x为A，则y为B
  - ✓ 前提2：x为A1
  - ✓ 结论：　y为B1

二元推理：



三元推理：



14.2 模糊推理-模糊推理的图示角度

Mamdani模糊推理

以一个简单的两输入一输出问题为例，包含如下三条规则

Rule: 1	Rule: 1
IF x is A3	IF project_funding is adequate
OR y is B1	OR project_staffing is small
THEN z is C1	THEN risk is low
Rule: 2	Rule: 2
IF x is A2	IF project_funding is marginal
AND y is B2	AND project_staffing is large
THEN z is C2	THEN risk is normal
Rule: 3	Rule: 3
IF x is A1	IF project_funding is inadequate
THEN z is C3	THEN risk is high

14.2 模糊推理-模糊推理的图示角度

模糊推理的步骤：

- Fuzzification of the input variables (输入变量模糊化)
- Rule evaluation（规则评判，即推理）
- Aggregation of the rule outputs（合成）
- Defuzzification（解模糊化）

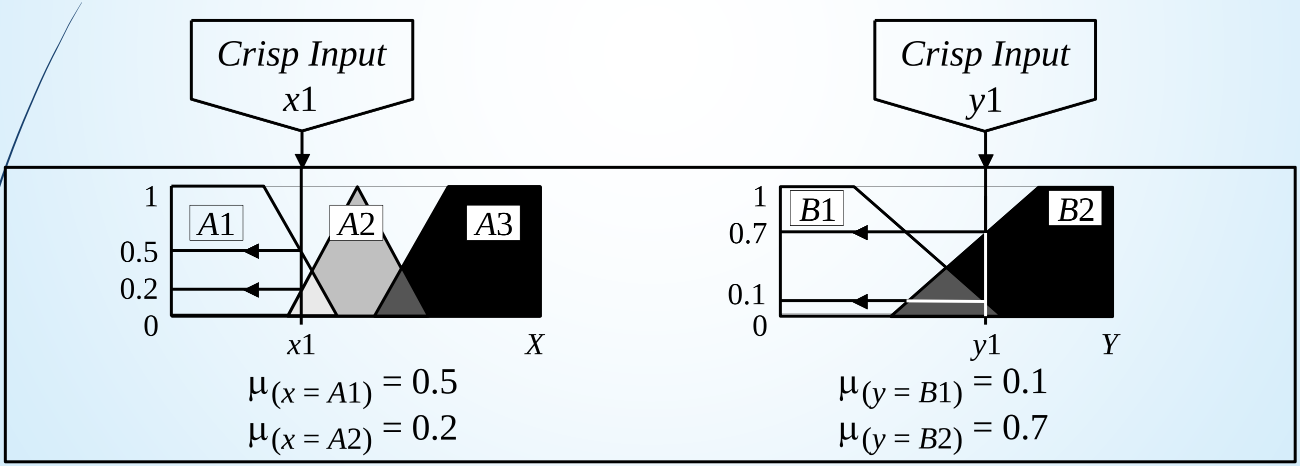
典型方法：

- Mamdani
- Sugeno

14.2 模糊推理-模糊推理的图示角度

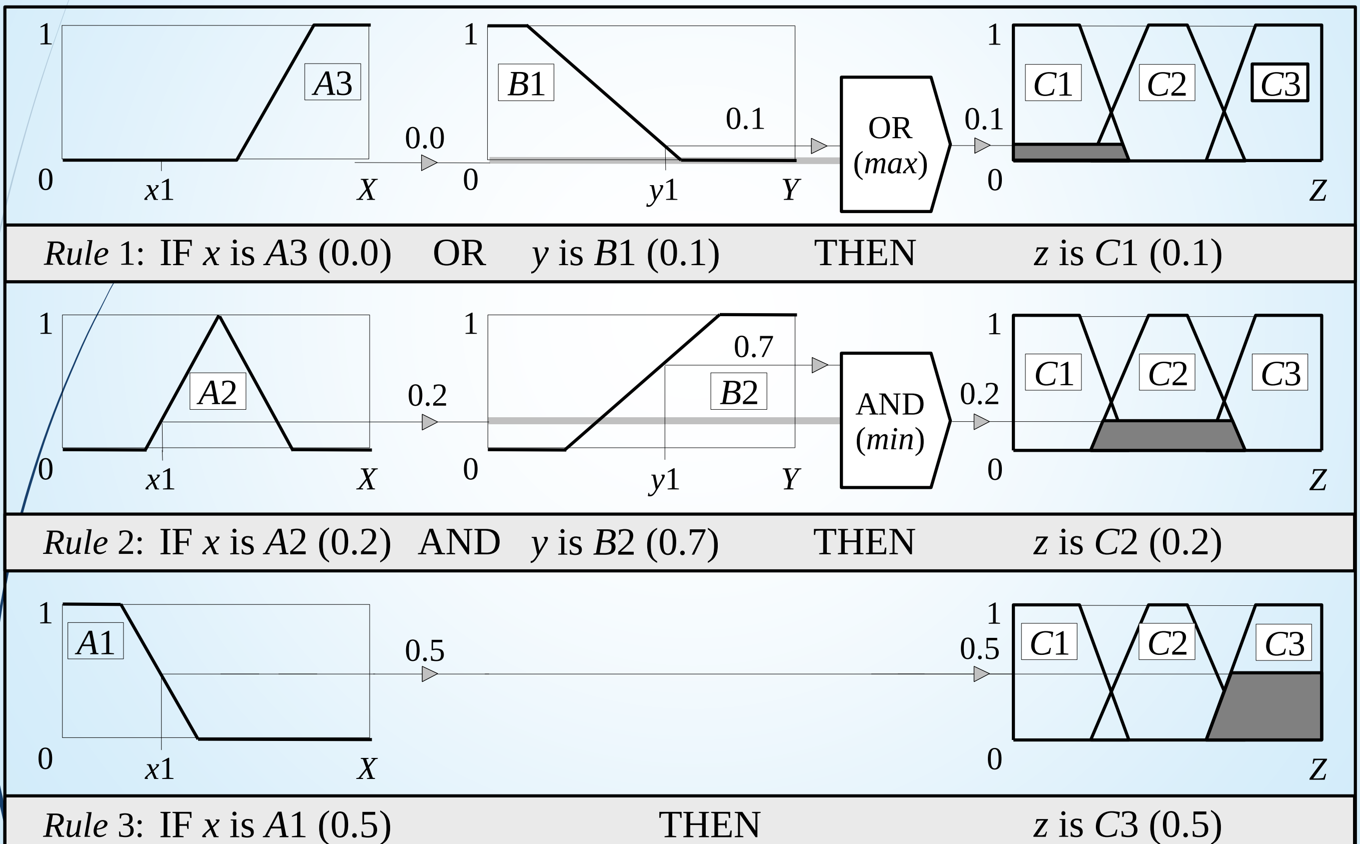
Mamdani模糊推理第一步：Fuzzification

第一步是给定输入x1和y1（例如project funding和project staffing），并且根据每种可能的模糊集合确定这些输入的隶属度



14.2 模糊推理-模糊推理的图示角度

Mamdani模糊推理第二步：Rule Evaluation



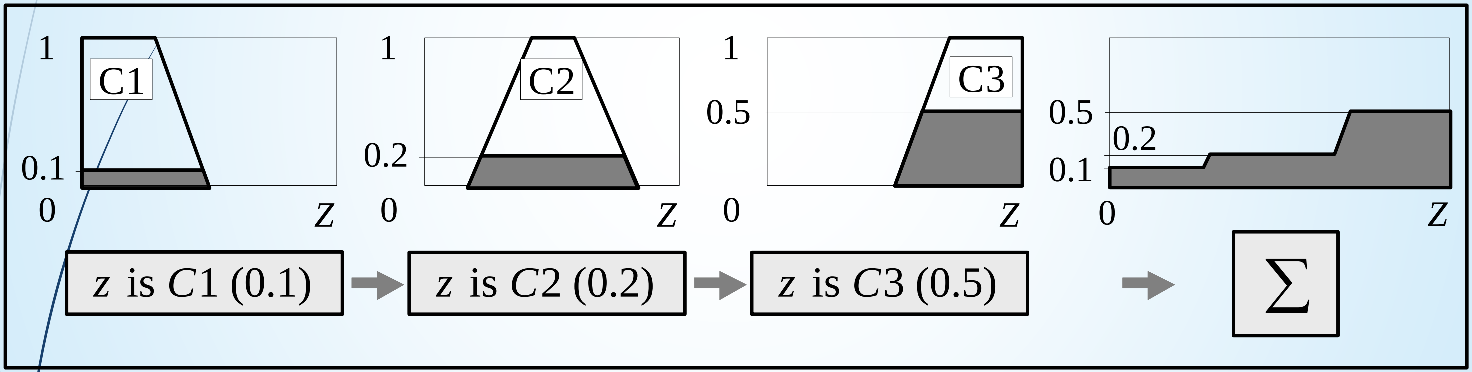
14.2 模糊推理-模糊推理的图示角度

- Mamdani模糊推理第三步：Aggregation

- ✓ 合成是将所有规则的所输出结果统一的过程；
- ✓ 先对所有规则下隶属函数的输出结果进行裁剪或缩放，然后将它们合并到一个单一的模糊集合中；
- ✓ 合成过程的输入是被裁剪或缩放的结果所属的隶属函数的列表，输出是一个满足任一输出变量的模糊集合。

14.2 模糊推理-模糊推理的图示角度

Mamdani模糊推理第三步：Aggregation



14.2 模糊推理-模糊推理的图示角度

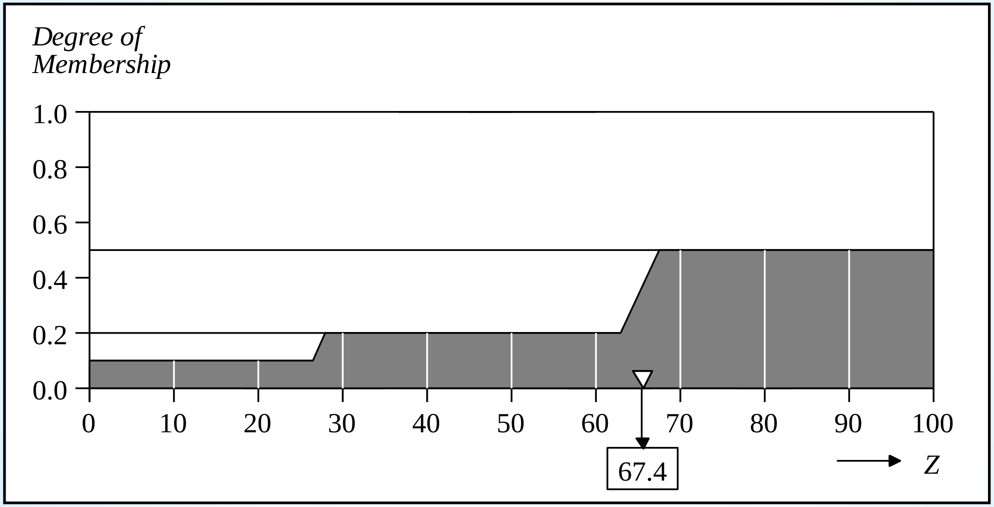
Mamdani模糊推理第四步：Defuzzification

- ✓ 模糊推理过程的最后一步是解模糊化；
- ✓ 模糊能够帮助我们评价这些规则，但是一个模糊系统的最终输出必须是一个准确的数；
- ✓ 解模糊化过程的输入是合成输出的模糊集合，输出是一个单个的数。

14.2 模糊推理-模糊推理的图示角度

Mamdani模糊推理第四步：Defuzzification

比如重心法：



$$\mathrm{COG} = \frac{(0+10+20)\times 0.1+(30+40+50+60)\times 0.2+(70+80+90+100)\times 0.5}{0.1+0.1+0.1+0.2+0.2+0.2+0.2+0.5+0.5+0.5+0.5} = 67.4$$

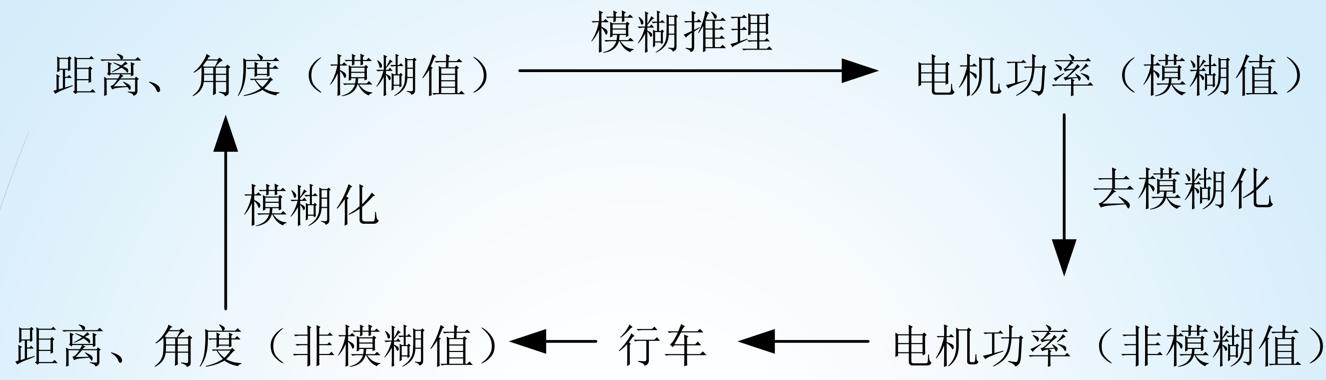
14.2 模糊推理-例：模糊推理实现控制

行车模糊控制系统

- ✓ 行车用来从船上装卸集装箱和在工厂车间里运送部件；
- ✓ 提高速度并且避免物体摆动；
- ✓ 控制的两个任务：定位和减震；
- ✓ 物体重量不同和距终点距离不同，行车被控过程为时变系统，采用常规控制有局限性。

14.2 模糊推理-例：模糊推理实现控制

14.2 模糊推理-例：模糊推理实现控制



1. 确定模糊控制系统的输入和输出变量；
2. 定义输入输出模糊集及其隶属函数；
3. 建立控制规则；
4. 模糊化；
5. 模糊推理；
6. 去模糊化。

14.2 模糊推理-例：模糊推理实现控制

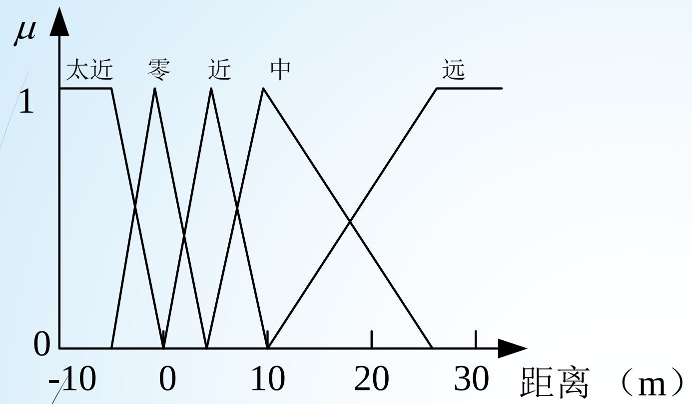
1. 确定模糊控制系统的输入和输出变量

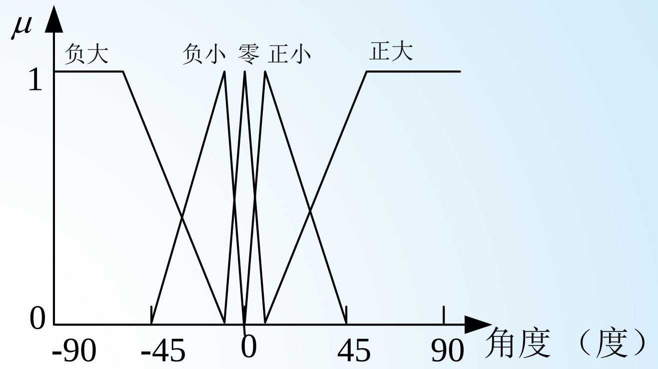
- 输入变量：物体距终点的距离；物体振荡的角度
- 输出变量：电机驱动功率

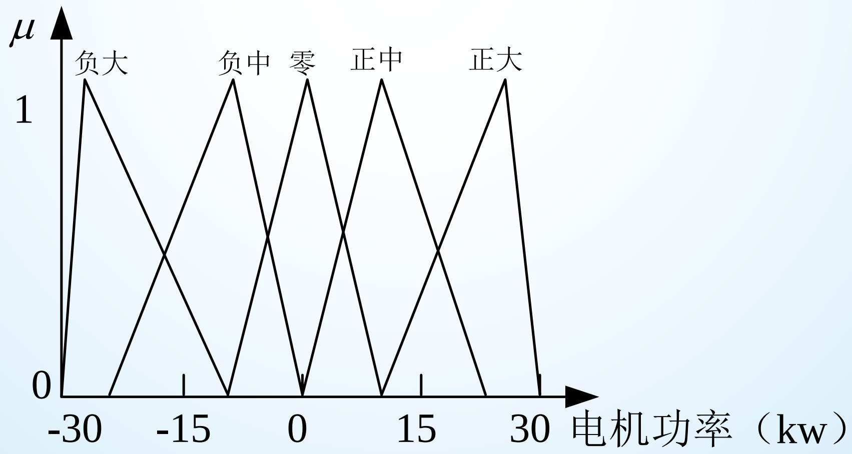
2. 定义输入输出模糊集及其隶属函数

- 距离：远、中、近、零、太近
- 角度：正大、正小、零、负小、负大
- 电机功率：正大、正中、零、负中、负大

14.2 模糊推理-例：模糊推理实现控制







14.2 模糊推理-例：模糊推理实现控制

3. 建立控制规则：25条规则

- 规则1：如果距离=中 并且角度=正小，则功率=正中
- 规则2：如果距离=中 并且角度=零，则功率=零
- 规则3：如果距离=远 并且角度=零，则功率=正中
- ……

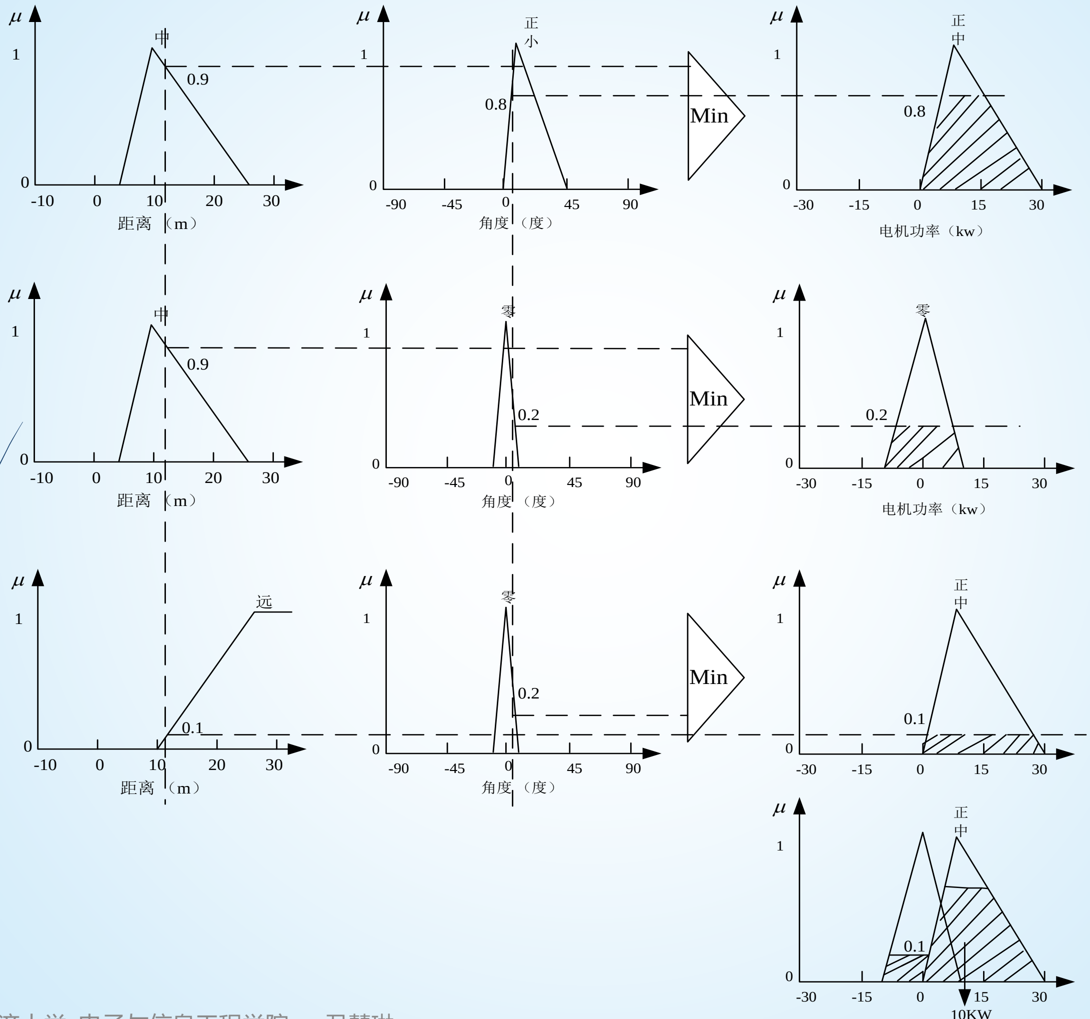
4. 模糊化

- 比如：设距离为12m，偏角为8度

5. 模糊推理

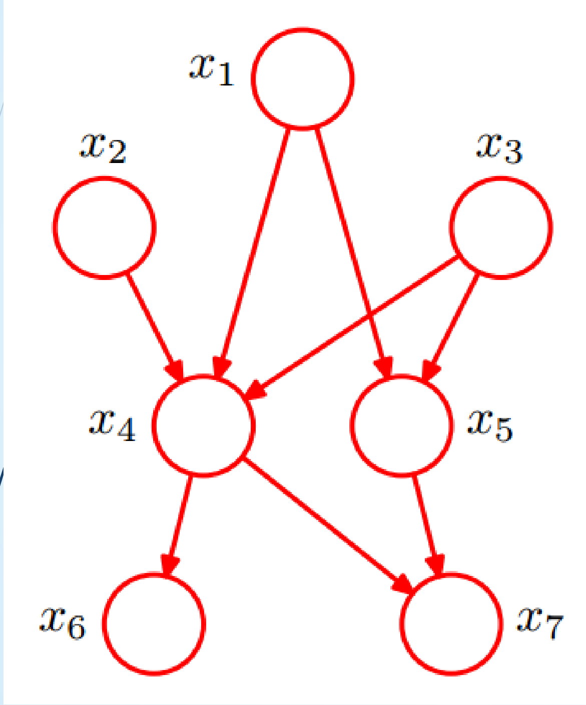
6. 去模糊化：面积重心法，得到电机功率

14.2 模糊推理-例：模糊推理实现控制



重点复习

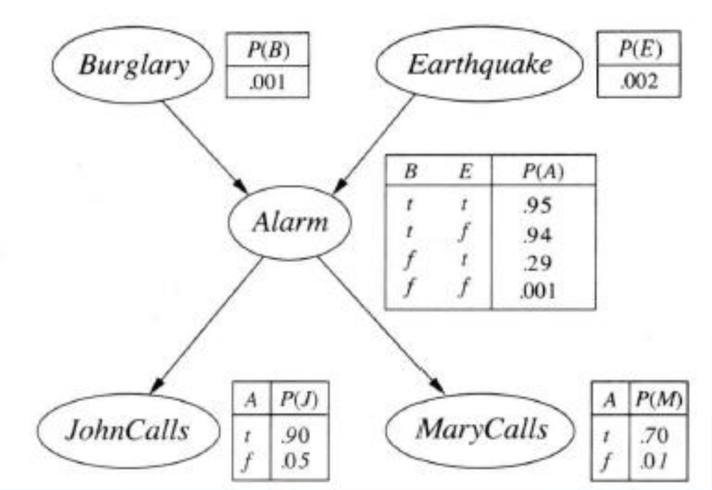
1. 根据下面贝叶斯网络，表示 $x_1, \dots, x_7$ 的完全联合概率分布。



$$p(x_1, \dots, x_7) =$$
$$p(x_1)p(x_2)p(x_3)p(x_4|x_1,x_2,x_3)$$
$$\mathrm{p}(x_5|x_1,x_3)p(x_6|x_4)p(x_7|x_4,x_5)$$

重点复习

2. 基于课堂讲的防盗贝叶斯网络例子，求报警器没响，既没盗贼也没地震，两人都打电话的概率。



重点复习

3. 设A表示接近于1的实数，B表示接近于2的实数，隶属函数分别为：

$$\mu_{\tilde{A}}(x) = \frac{1}{1+(x-1)^2}, \mu_{\tilde{B}}(x) = \frac{1}{1+(x-2)^2}$$

$\tilde{A} \cap \tilde{B}$表示既接近于1又接近于2的实数。

求 $\mu_{\tilde{A} \cap \tilde{B}}(x)$