非精确性推理

中科大 自动化系　郑志刚

2018.11

提纲

- 非精确性推理概述
- 主观贝叶斯方法
- 可信度方法

§1.非精确性推理概述

做不精确推理的原因有多种：

- 推理所需的信息不完备：竞争双方不知道对方信息
- 背景知识不足：疑难病症的机理
- 多种原因导致同一结果：疾病的诊断
- 信息描述模糊：目击者对嫌疑犯的描述
- 信息中含有噪声：做假帐，虚假统计报表，采集数据当中的噪声（雷达、声纳/化验）等
- 规则是模糊的：定性描述，如“如果刑事犯罪猖獗，就应加大打击力度”等
- 推理能力不足：天气预报的计算
- 解决方案不唯一：多个方案如何选优的问题

不精确推理的必要性

- 研究非精确性推理的必要性
  - 现实世界中大多数知识和应用都是非精确的
  - AI是对人的智能的模拟
  - 有必要研究这种非精确性的知识的表示和运用
- 非精确性推理
  - 对非精确性知识的运用与处理
  - 从非精确的初始证据出发，通过反复应用非精确性知识，最后推出非精确但近于合理的结论

- 常用的非精确性推理技术
  - 确定性理论
  - 证据理论
  - 主观贝叶斯方法
  - 模糊理论

非精确性推理中的几个基本问题

- 一般推理系统中的普遍问题
  - 推理方式,推理策略
  - 例如:归结推理, 线形输入策略
- 非精确性推理中的特殊问题
  - 非精确性的表示和度量
  - 非精确性的传递和推算
  - 条件部分的非精确性确定
  - 非精确性的匹配算法及阈值选择

非精确性的表示和度量

- 采用何种方式表示非精确性的问题
  - 确定性方式,模糊方式,概率方式
- 描述2种非精确性
  - 静态知识(事实或者结论)的非精确性
  - 动态知识(规则)的非精确性
- 非精确性的表示方法应与推理方式相关,相适应

非精确性的传递和推算

- 推理的步骤
  - 前提->中间结论->中间结论 ->最终结论
- 非精确性的传递和推算
  - 非精确性如何在前提，中间结论和最终结论中传递

条件部分非精确性的确定

- 规则的前提部分通常是简单前提的复杂合取，析取形式
- 如何由简单前提的非精确性计算出复杂前提的非精确性

非精确性的匹配算法与阈值选择

- 证据与可应用的规则的前提条件部分是部分匹配的
- 如何计算出2者匹配的程度
- 判断该规则是否可以被运用

不精确推理将数值计算引入推理过程

- 继续使用逻辑联结词
- 真假值概率化，以表示某种可靠程度
- 在推理的前提和结论之间建立概率公式（解释）
- 前提称为证据，结论称为假设

应用：专家系统中的推理网络 (PROSPECTOR地矿系统 / MYCIN诊断系统)

§2. Bayes概率推理

2.1 Bayes概率

1. 主观Bayes主义：将概率推理与归纳逻辑的解释联系起来，现实世界的一些因果关系可以形成一种信念(归纳过程)，它并非在所有场合下都正确，因而可称为部分信念。表示这种信念的最好方法是概率方法，对概率的解释有若干种，其中Bayes创立了主观解释，也称为主观Bayes主义。其要点是：概率是个人的一种合理置信度，每个人的估计(概率)虽然各不相同，但应该满足概率的基本规律和其他某些客观规律，因而是合理的。

2.推理过程：在该方法中，推理规则表示为
if E(前提/证据) then H(结论/假设)(LS, LN)
规则强度用LS/LN表示
其不精确推理过程是：根据证据E的概率P(E)，利用规则的LS和LN，把结论的先验概率P(H)更新为后验概率P(H|E)，因而也称为概率传播。

3. Bayes概率

令假设H的先验概率为P(H)，P(H)是指定的。

证据E为真时H的条件概率为P(H|E)，条件概率P(H|E)可看作以一定概率成立的E→H的推理规则。

Bayes概率服从如下公理：

(1)$0 \le P(H) \le 1$；

(2)$P(H=T)=1$；

(3)$P(H \vee G)=P(H)+P(G)$，其中H和G互斥；特别有$P(H)+P(\neg H)=1$，并且一般地有 $\sum_i P(H_i \mid E)=1$，即证据E下H_i的全体构成了一切可能的假设，其中任意H_i和H_j $(i \ne j)$互斥。

Bayes理论中经常用到2个公式，即条件概率公式和逆概率公式。

4. 条件概率公式

$$P(H\&G)=P(H|G)P(G)=P(G|H)P(H) \qquad (1)$$

该公式指明了H和G两个假设之间存在着相关性。如果H和G相互独立，则

$$P(H|G)=P(H)$$

$$P(G|H)=P(G)$$

$$P(H\&G)=P(H)P(G) \qquad (2)$$

一个假设的先验概率可以表示为两个假设的概率，其中后面一个假设遍历各种可能性且各个可能性之间是独立的，则有(全概率公式)

$$P(H)=\sum_i P(H \& G_i) = \sum_i P(H|G_i)P(G_i) \qquad (3)$$

其解释为：假设H的概率等于从各种证据提供的信息所推出的条件概率之和。证据满足独立性和完全性。

5. 逆概率公式

从条件概率公式可得

$$P(H \mid G) = \frac{P(G \mid H) \times P(H)}{P(G)} \tag{4}$$

如果有多个假设$H_1 \sim H_n$，则上式化为

$$P(H_i \mid G) = \frac{P(G \mid H_i)P(H_i)}{\sum_{j=1}^{n} P(G \mid H_j)P(H_j)} \tag{5}$$

逆概率公式不仅仅是条件概率公式的一个简单变形，实际上很有用处。如果某个条件概率不便计算，则可以先计算其逆概率，而后算出所要的条件概率。

例子：假设P(肺炎)=1/10000，而P(咳嗽)=1/10，求P(肺炎|咳嗽)；

直接求P(肺炎|咳嗽)可能比较困难，但统计P(咳嗽|肺炎)可能比较容易(因为要上医院)，统计得90%的肺炎患者都咳嗽，那么，则

P(肺炎|咳嗽)=

★

$$\frac{P(\text{咳嗽}\mid\text{肺炎})\times P(\text{肺炎})}{P(\text{咳嗽})}=\frac{0.9\times 0.0001}{0.1}=0.0009$$

也可以将(4)式写成

$$P(H \mid E) = [\frac{P(E \mid H)}{P(E)}]P(H) \qquad (4')$$

这说明先验概率P(H)可以通过方括号部分（作为修正因子）修正为后验概率P(H|E)（证据E为真时H的后验概率）。

在上面的例子中，医生认为一个人得肺炎的可能性为万分之一，一旦发现患者咳嗽，就将调整为万分之九。

将E看作证据，先验概率P(E)越小，且H为真时E的条件概率P(E|H)越大，则修正因子所起作用越大。

在上例中，如果P(咳嗽)=0.0001，P(咳嗽|肺炎)=0.9999，而P(肺炎)不变，则P(肺炎|咳嗽)=0.9999，远远超过原来的万分之九。

当有n个互相独立的证据，则有公式

$$P(H \mid E_1 \& \cdots \& E_n) = \frac{\prod_{i=1}^{n} P(E_i \mid H)}{\prod_{i=1}^{n} P(E_i)} \times P(H) \qquad (6)$$

(5.6)式可以写成递推公式形式：

$$P(H \mid E_1 \& \cdots \& E_{m+1}) = \frac{P(E_{m+1} \mid H)}{P(E_{m+1})} \times P(H \mid E_1 \& \cdots \& E_m) \qquad (7)$$

上式说明：随着新证据的不断获得，从证据少时的后验概率推出证据多时的后验概率，且每一步都是把上一步的后验概率视为新证据到来时的先验概率。

2.2 几率和似然比

引入相对量度

1. [定义1]几率：$O(H)=\frac{P(H)}{P(\neg H)}=\frac{P(H)}{1-P(H)}$ (8)

称为H的几率或先验几率，取值范围[0,∞)。

由此反过来有 $P(H)=\frac{O(H)}{1+O(H)}$ (9)

2. [定义2]条件几率：$O(H|E)=\frac{P(H|E)}{P(\neg H|E)}$ (10)

例子：O(晴天|冬天早晨有雾)=4.2，如果冬天早晨有雾，则该天为晴天的可能性是非晴天可能性的4.2倍。

由(8)(10)和条件概率公式可以推得后验几率和先验几率的关系：

$$O(H|E)=\frac{P(E|H)}{P(E|\neg H)}\times O(H) \quad (11)$$

3. [定义3]似然比：

$LS=P(E|H)/P(E|\neg H)$

$LN=P(\neg E|H)/P(\neg E|\neg H)$ (12)

- LS表示E为真对H的支持程度，充分性度量
- LN表示E为假对H的支持程度，必要性度量

则可得下述关系：

$O(H|E)=LS*O(H)$ $O(H|\neg E)=LN*O(H)$ (13)

有多个证据独立时，其公式可

$$O(H \mid E_1 \& \cdots \& E_n) = \prod_{i=1}^{n} LS(E_i \mid H) \times O(H) \qquad (14)$$

4. 对LS和LN的约束

对于LS和LN有如下约束要求：二者都是非负的，并且满足

$$\begin{cases} LS > 1 \& LN < 1 \\ LS < 1 \& LN > 1 \\ LS = 1 \& LN = 1 \end{cases}$$

5. LS和LN的应用

当P(E)=1时，利用LS来将先验几率O(H)更新为后验几率O(H|E)；当P(¬E)=1时，利用LN来更新几率。

在专家系统PROSPECTOR(一个用于探矿的ES)中同时应用了LS和LN，分别表示正面证据和反面证据的支持，称为充分因子和必要因子。并将LS、LN附着在每条规则之上。

当LS很大，说明证据成立时假设成立的可能性很大，否则LS<1说明E排斥H；LN很小，说明证据不成立时假设不成立的可能性很大。LS和LN之值接近1时说明证据成立或不成立对于结论是否成立影响不大。一般情况下，LS和LN不是根据定义计算出来的，而是给定的。

应用举例(1)

- 例子：评职称的概率
  - 设某副教授X要评正教授，现有4个指标，却有8人参与竞争 / 投票前夕，X作了如下预测：如果不考虑评委因素，则成功概率=1/2，此相当于先验几率O(H)=1；如果考虑评委因素，则情况如下：
  - 校评委共15人，其中5人来自其他竞争者所在系，4人与X素有微隙，尤其是其中2人兼具来自其他竞争者所在系，对X的成功构成了极大威胁，但聊以自慰的是评委中有5位老朋友，估计会投X的票

应用举例(2)

- 为此，X定义了如下的似然比：
  - LS(评委Y1出席|X评上)=1/2　　Y1来自其他竞争者所在系，同时令LN=2 (LS*LN=1)
  - LS(评委Y2出席|X评上)=1/4　　Y2与X素有微隙，同时令LN=4
  - LS(评委Y3出席|X评上)=1/8　　Y3来自其他竞争者所在系兼与X素有微隙，同时令LN=8
  - LS(评委Y4出席|X评上)=4Y4是X的老朋友，同时令LN=1/4
  - LS(评委Y5出席|X评上)=1　Y5不属于以上情况，LN=1

应用举例(3)

- 若15人全体出席，且假定各条件互相独立，则按公式(5.14)，X评上的后验概率是：

$$O(X评上|15人出席)= (\frac{1}{2})^3 \times (\frac{1}{4})^2 \times (\frac{1}{8})^2 \times (4)^5 \times (1)^3 \times 1 = \frac{1}{8}$$

- 根据(5.9)式，换为概率P=O/(1+O)=1/9，X评上的希望不大

应用举例(4)

- 但是，当又有消息说，一位来自其他竞争者所在系兼与X素有微隙的评委A不能出席，而代之以一位态度中立的评委 / 此时，X又作了一番推测：

  LS(¬A出席|X评上)=LN(A出席|X评上)=8

  LS(中立评委|X评上)=1

则在原结果的基础上乘上上述因子，使得后验几率=1，即后验概率=1/2，X评上的前景大大改观 ★

2.3 证据不确定下的概率推理

实际应用中，证据(在上例中即Y是否出席)是不能完全确定的，此时PROSPECTOR系统对LS和LN的计算就采用了新公式。

1. 证据的不确定性：0<P(E)<1，此时有观察到的证据(观察)S，如何根据P(E|S)求出P(H|S)。

但是不存在严格估算P(E|S)的方法，因为S影响E的机制难以弄清，属于是常识推理的范围

2. 线性插值公式：以此从P(E|S)计算P(H|S)

$$P(H|S)=\begin{cases} P(H)+\dfrac{P(H|E)-P(H)}{1-P(E)}[P(E|S)-P(E)] & \text{当}P(E)\le P(E|S)\le 1 \\ P(H|\neg E)+\dfrac{P(H)-P(H|\neg E)}{P(E)}P(E|S) & \text{当}0\le P(E|S)<P(E) \end{cases}$$

此式称为EH公式，参见图5.1示意

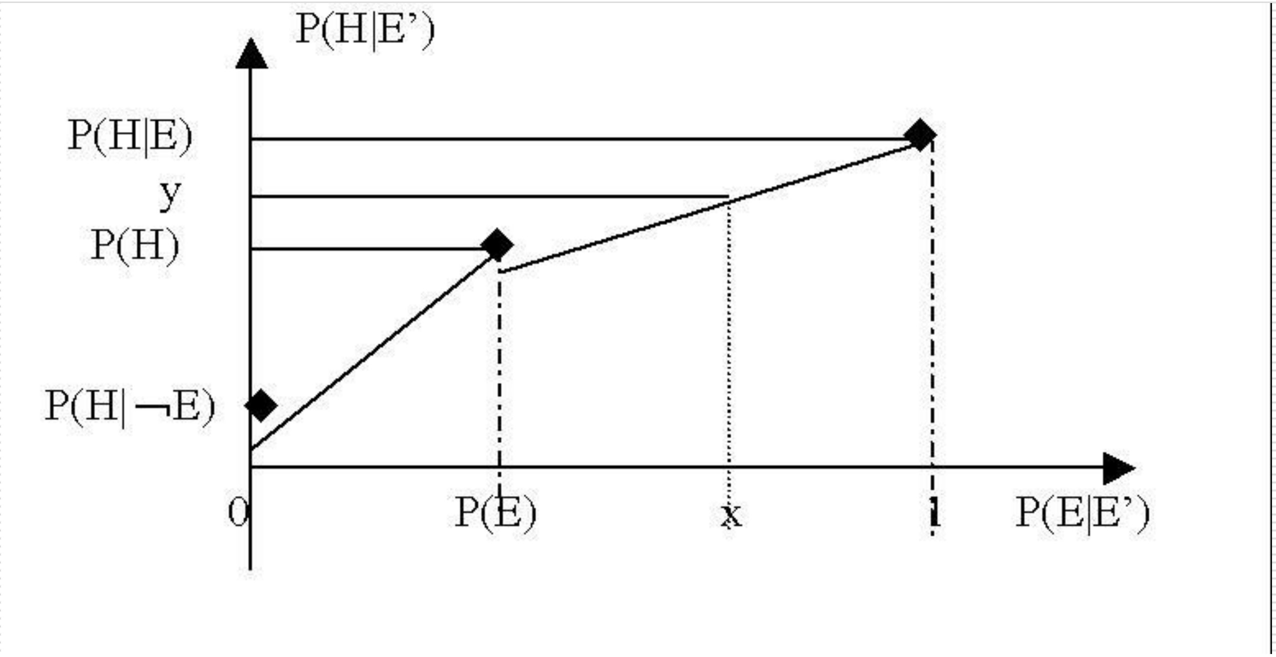


图 1

3. 多个独立证据下假设的求解

设有独立证据E_1, … , E_n的观察为S_1,…,S_n且有规则$E_1 \rightarrow H$, …, $E_n \rightarrow H$。假定分别得到的假设H的后验几率为

$O(H|S_1)$, …, $O(H|S_n)$则有公式如下

$$O(H \mid S_1 \,\&\, \ldots \,\&\, S_n) = \frac{O(H \mid S_1)}{O(H)} \times \cdots \times \frac{O(H \mid S_n)}{O(H)} \times O(H)$$

4. 证据的合取

$$P(E_1 and E_2 and \cdots and E_n \mid S) = \min\{P(E_1 \mid S), P(E_2 \mid S), \cdots, P(E_n \mid S)\}$$

5. 证据的析取

$$P(E_1 or E_2 or \cdots or E_n \mid S) = \max\{P(E_1 \mid S), P(E_2 \mid S), \cdots, P(E_n \mid S)\}$$

6. 方法的不足：混合使用概率和几率；当证据确定时，使用LS/LN，而不确定时又使用插值公式，存在不一致的地方

2.4 推理网络中的例子

设有一个推理网络如下图所示(图2)

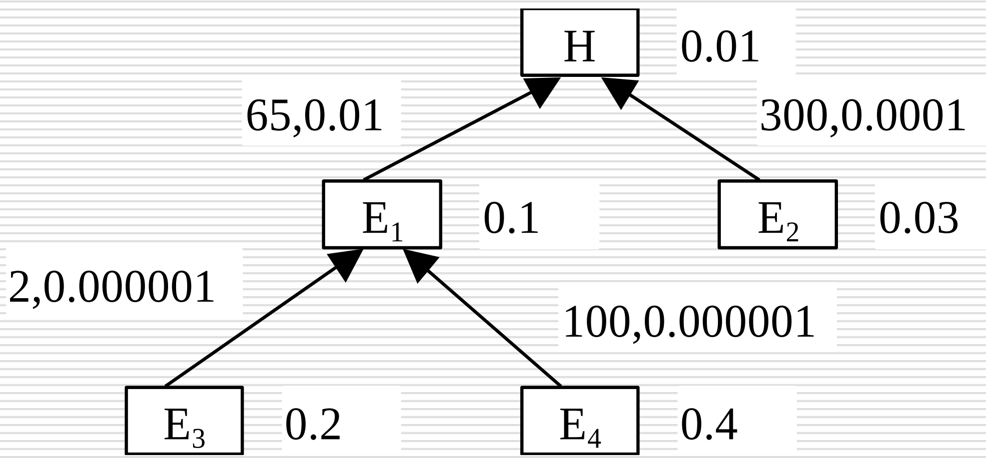


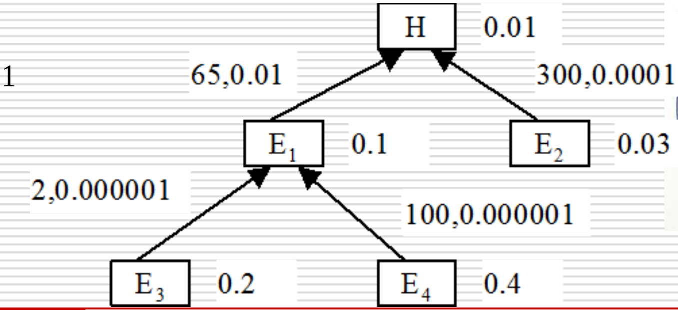
图2

图中的数值为：节点的先验概率；规则的LS/LN值。并且给出初始证据在各自观察下的概率值：

$P(E_3|S_1)=0.7$　　$P(E_4|S_2)=0.6$　　$P(E_2|S_3)=0.02$

求假设H的后验概率$P(H|S_1\&S_2\&S_3)$

$$P(H|S)=\begin{cases} P(H)+\dfrac{P(H|E)-P(H)}{1-P(E)}[P(E|S)-P(E)] & \text{当}P(E)\le P(E|S)\le 1 \\ P(H|\neg E)+\dfrac{P(H)-P(H|\neg E)}{P(E)}P(E|S) & \text{当}0\le P(E|S)<P(E) \end{cases}$$



求解过程:

(1)根据P(E3|S1)计算P(E1|S1)

因P(E3|S1)=0.7>P(E3)=0.2用EH公式前半部分。

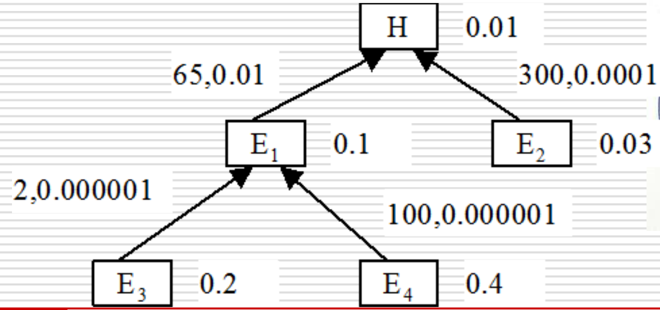
先求P(E1|E3)，从O(E1|E3)=LS*O(E1)代入 O(X)=P(X)/[1 – P(X)]可得

$$P(E_1|E_3)=\frac{LS\times P(E_1)}{1-P(E_1)+LS\times P(E_1)}$$

按此式计算得：P(E1|E3)=(2*0.1)/(1-0.1+2*0.1)=0.1818

则P(E1|S1)=0.1+[(0.1818-0.1)/(1-0.2)]*(0.7-0.2)=0.151125

$$P(H|S)=\begin{cases} P(H)+\dfrac{P(H|E)-P(H)}{1-P(E)}[P(E|S)-P(E)] & \text{当}P(E)\le P(E|S)\le 1 \\ P(H|\neg E)+\dfrac{P(H)-P(H|\neg E)}{P(E)}P(E|S) & \text{当}0\le P(E|S)<P(E) \end{cases}$$

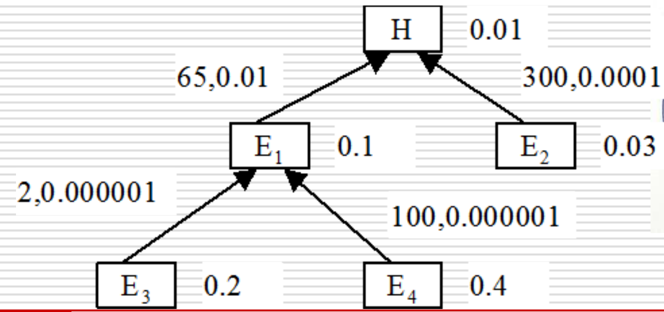


(2)根据P(E4|S2)计算P(E1|S2)，类似于(1)可得

P(E1|E4)=(100*0.1)/(1-0.1+100*0.1)=10/10.9=0.9174311

因P(E4|S2)=0.6>P(E4)=0.4，故仍然利用EH公式前半部分

则P(E1|S2)=0.1+[(0.9174311-0.1)/(1-0.4)]*(0.6-0.4)=0.372477



(3)由P(E1|S1)和P(E4|S2)计算P(E1|S1&S2)，先计算 O(E1|S1&S2)

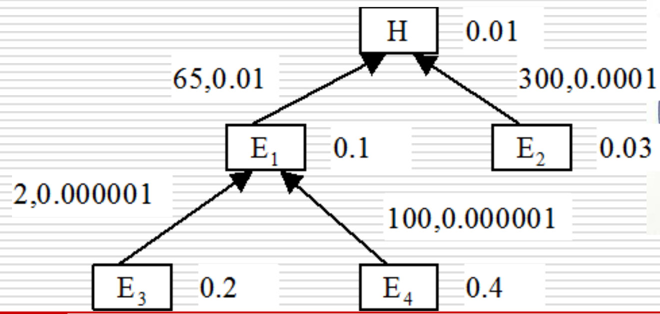
因O(E1)=0.1/(1-0.1)=0.1111111，
O(E1|S1)=0.151125/(1-0.151125) =0.1780297，
O(E1|S2)=0.372477/(1-0.372477)=0.593567

故O(E1|S1&S2)

=[O(E1|S1)/O(E1)]*[O(E1|S2)/O(E1)]*O(E1)=0.9510532

则
P(E1|S1&S2)=O(E1|S1&S2)/(1+O(E1|S1&S2))=0.4874563

$$P(H|S)=\begin{cases} P(H)+\dfrac{P(H|E)-P(H)}{1-P(E)}[P(E|S)-P(E)] & \text{当} P(E)\le P(E|S)\le 1 \\ P(H|\neg E)+\dfrac{P(H)-P(H|\neg E)}{P(E)}P(E|S) & \text{当} 0\le P(E|S)<P(E) \end{cases}$$

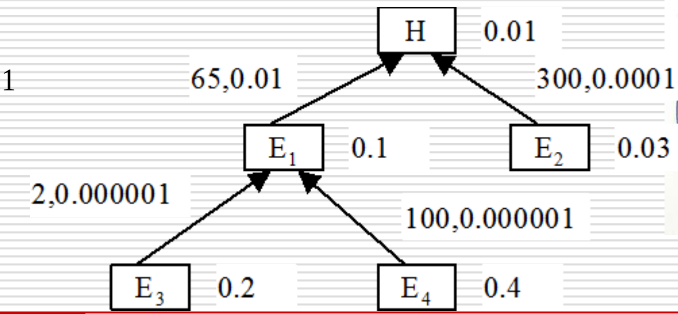


(4)根据P(E1|S1&S2)计算P(H|S1&S2)，因 P(E1|S1&S2)>P(E1)=0.1，所以仍然使用EH公式前半部分

先计算P(H|E1)=LS*P(H)/[1-P(H)+LS*P(H)]=65*0.01/(1-0.01+65*0.01) =0.3963414

再计算P(H|S1&S2)=P(H)+[(P(H|E1)-P(H))/(1-P(E1))]*(P(E1|S1&S2)-P(E1))=0.01+[(0.3963414-0.01)/(1-0.1)]*(0.3963414-0.1)=0.1763226

$$P(H|S)=\begin{cases} P(H)+\dfrac{P(H|E)-P(H)}{1-P(E)}[P(E|S)-P(E)] & \text{当}P(E)\le P(E|S)\le 1 \\ P(H|\neg E)+\dfrac{P(H)-P(H|\neg E)}{P(E)}P(E|S) & \text{当}0\le P(E|S)<P(E) \end{cases}$$



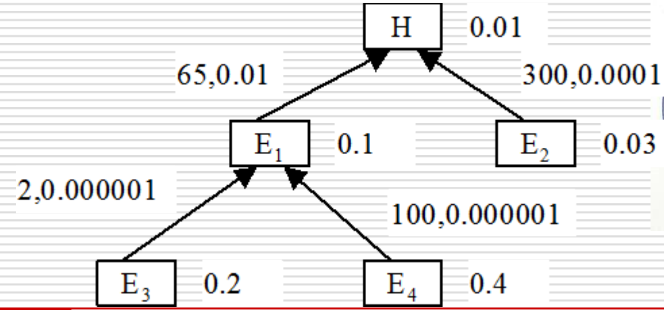
(5)根据P(E2|S3)计算P(H|S3)，因为 P(E2|S3)=0.02<P(E2)=0.03，故使用EH公式的后半部分

需先计算P(H|¬E2)，由O(H|¬E2)=LN*O(H)及 O(X)=P(X)/[1 – P(X)]可得

$$P(H|\neg E_2)=\frac{LN\times P(H)}{1-P(H)+LN\times P(H)}$$

则P(H|¬E2)=(0.0001*0.01)/(1-0.01+0.0001*0.01)=0.000001

进一步求P(H|S2)=P(H|¬E2)+[(P(H)-P(H|¬E2))/P(E2)]*P(E2|S2) =0.000001+[(0.01-0.000001)/0.03]*0.02=0.006667



(6)根据独立证据E1和E2，计算P(H|S1&S2&S3)

先计算后验几率，因为

O(H)=0.01/(1-0.01)=0.010101

O(H|S1&S2)=0.1763226/(1-0.1763226)=0.2140675

O(H|S3)=0.006667/(1-0.006667)=0.00671174

则

O(H|S1&S2&S3)=(0.2140675/0.010101)*(0.00671174/0.010101) *0.010101=0.142239

最后得到后验概率

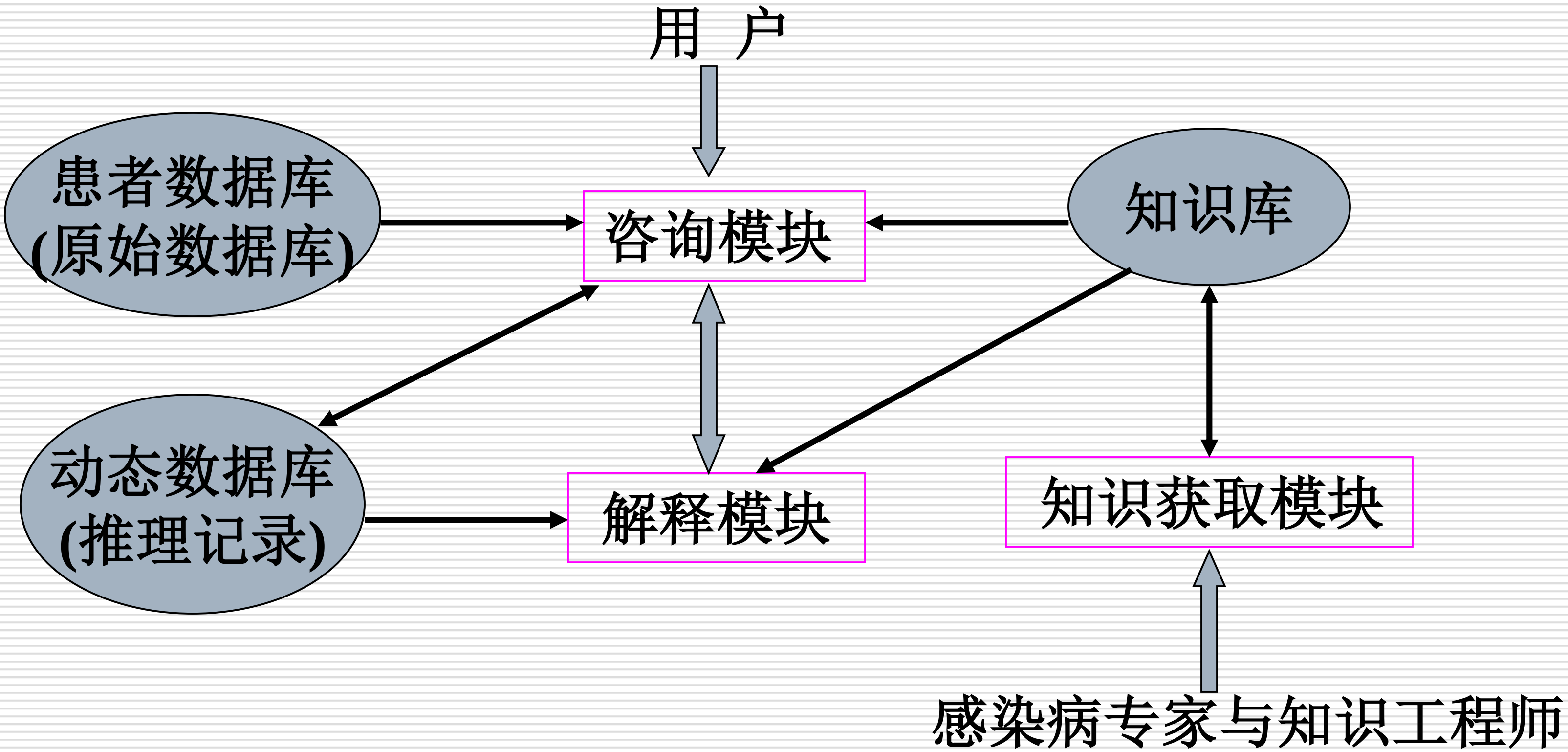
P(H|S1&S2&S3)=0.142239/(1+0.142239)=0.1245264

§3 可信度方法

3.1 可信度方法的原则

- **MYCIN系统是第一个采用了不确定推理逻辑的专家系统，在20世纪70年代非常有名。**
- **这个系统提出该确定性方法时遵循了下面的原则：**
  （1） 不采用严格的统计理论。使用的是一种接近统计理论的近似方法。
  （2） 用专家的经验估计代替统计数据
  （3） 尽量减少需要专家提供的经验数据，尽量使少量数据包含多种信息。
  （4） 新方法应适用于证据为增量式地增加的情况。
  （5） 专家数据的轻微扰动不影响最终的推理结论。

MYCIN 概述



MYCIN系统结构图

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。

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。

3.2. 知识不确定性的表示:

表示形式：

在C-F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。

例子：

IF 发烧 AND 流鼻涕 THEN 感冒 (0.8)

说明：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。

说明：

(1) E可以是单一条件，也可以是复合条件。例如：

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

(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。

(4) CF(H, E)可以理解为规则的可信度

3.3.可信度的定义与性质

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)

MB: 信任增长度，**MB(H, E)**定义为**:**

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

MD:不信任增长度，**MB(H, E)**定义为**:**

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

MB和MD的关系:

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

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

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

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

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

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

可信度的性质:

- 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

- 值域

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

- **典型值**

(1) 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。

(2) 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。

(3)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。

(4) 对H的信任增长度等于对非H的不信任增长度

$$MD(\neg H, E) = MB(H, E)$$

$$CF(H,E)+CF(\neg H,E) = 0$$

- 对**H**的信任增长度等于对非**H**的不信任增长度
- 对**H**的可信度与非**H**的可信度之和等于**0**
- 可信度不是概率**:**P(H|E)+P(¬H|E)=1

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

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

若：专家给出的知识有如下情况

$CF(H_1, E)=0.7$, $CF(H_2, E)=0.4$

非法，应进行调整或规范化

3.4. 证据不确定性的表示

证据（E）不确定性的表示：

证据的不确定性也是用可信度来表示的，其取值范围也为[-1,1]

若E为初始证据，其值由用户给出。

若E为中间结论，其值可通过计算得到。

不确定性的含义：

对E，其可信度CF(E)的含义如下：

CF(E)=1，证据E肯定它为真

CF(E)=-1，证据E肯定它为假

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

4. 否定证据不确定性的计算

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

5. 组合证据不确定性的计算

“合取”与“析取”两种基本情况。

合取:

当组合证据是多个单一证据的组合时

即 $E=E_1$ AND E_2 AND … AND E_n时，若已知$CF(E_1)$，$CF(E_2)$，…，$CF(E_n)$，则

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

析取:

当组合证据是多个单一证据的析取时

即$E=E_1$ OR E_2 OR … OR E_n时，若已知$CF(E_1)$，$CF(E_2)$，…，$CF(E_n)$，则

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

3.5. 不确定性推理

- 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性知识，逐步推出最终结论和该结论可信度的过程。
- 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: CF(H)=0 即该模型没考虑E为假对H的影响。

若CF(E)=1: 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

3.6. 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

$$
CF(H)=\begin{cases}
CF_1(H)+CF_2(H)-CF_1(H)\times CF_2(H) & \text{若} CF_1(H)\geq 0 \\
 & \text{且} CF_2(H)\geq 0 \\
CF_1(H)+CF_2(H)+CH_1(H)\times CF_2(H) & \text{若} CF_1(H)<0 \\
 & \text{且} CF_2(H)<0 \\
\dfrac{CF_1(H)+CF_2(H)}{1-\min\{|CF_1(H)|,|CF_2(H)|\}} & \text{若} CF_1(H)\text{与} \\
 & CF_2(H)\text{异号}
\end{cases}
$$

例子

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：

$$CF(E_1)=0.8\times\max\{0, CF(E_4 \text{ AND } (E_5 \text{ OR } E_6))\}$$
$$= 0.8\times\max\{0, \min\{CF(E_4), CF(E_5 \text{ OR } E_6)\}\}$$
$$=0.8\times\max\{0, \min\{CF(E_4), \max\{CF(E_5), CF(E_6)\}\}\}$$
$$=0.8\times\max\{0, \min\{CF(E_4), \max\{0.6, 0.8\}\}\}$$
$$=0.8\times\max\{0, \min\{0.5, 0.8\}\}$$
$$=0.8\times\max\{0, 0.5\} = 0.4$$

例子

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_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：

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：

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

$CF_{12}(H)$和$CF_3(H)$异号，有：

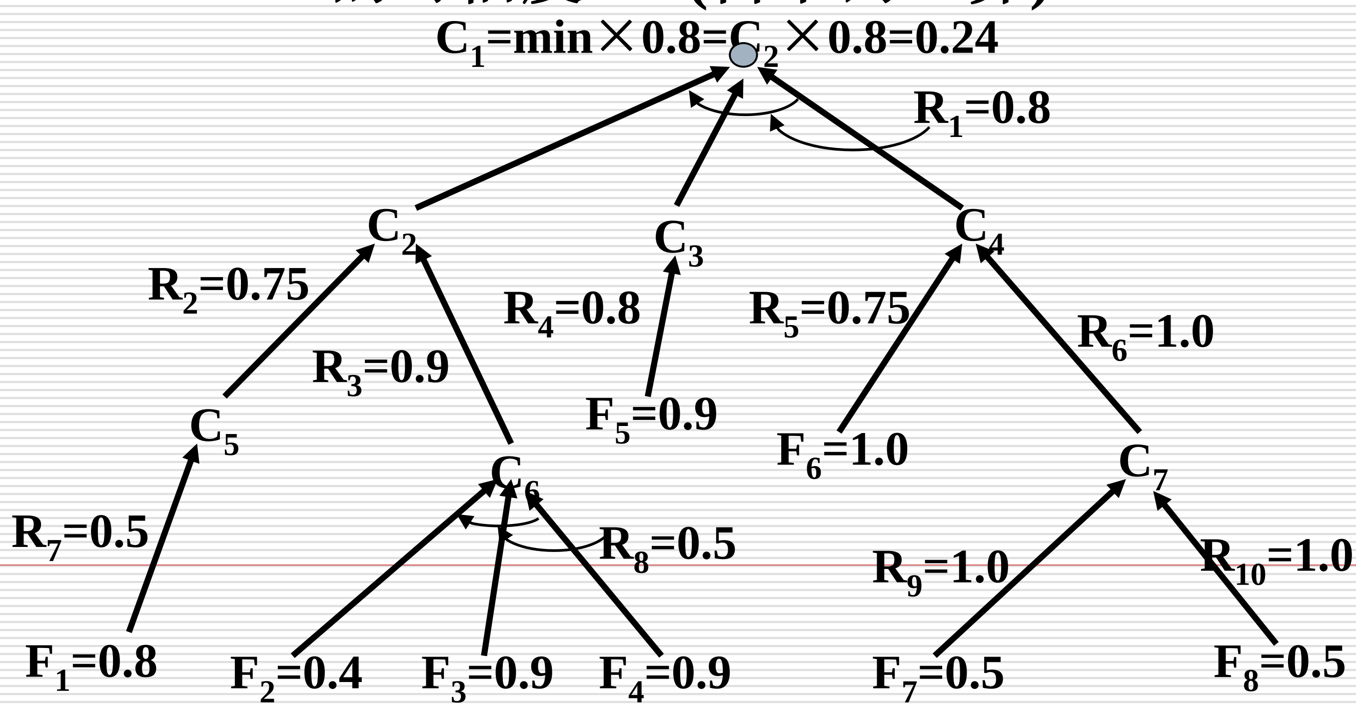
$$
\begin{aligned}
CF_{1,2,3}(H) &= \frac{CF_{1,2}(H)+CF_3(H)}{1-\min\{|CF_{1,2}(H)|,|CF_3(H)|\}} \\
&= \frac{0.67-0.3}{1-\min\{0.67,0.3\}} = \frac{0.37}{0.7} \\
&= 0.53
\end{aligned}
$$

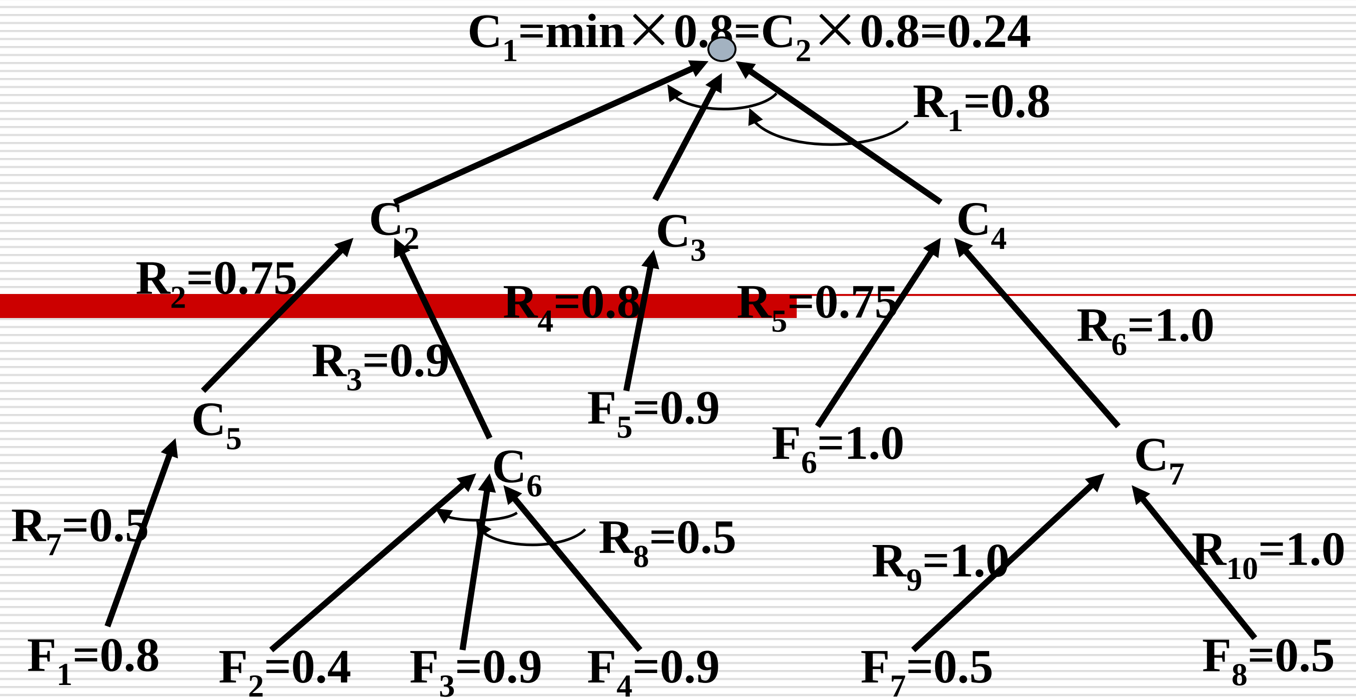
即综合可信度为CF(H)=0.53

MYCIN 不精确推理

不精确推理过程可以总结如下：

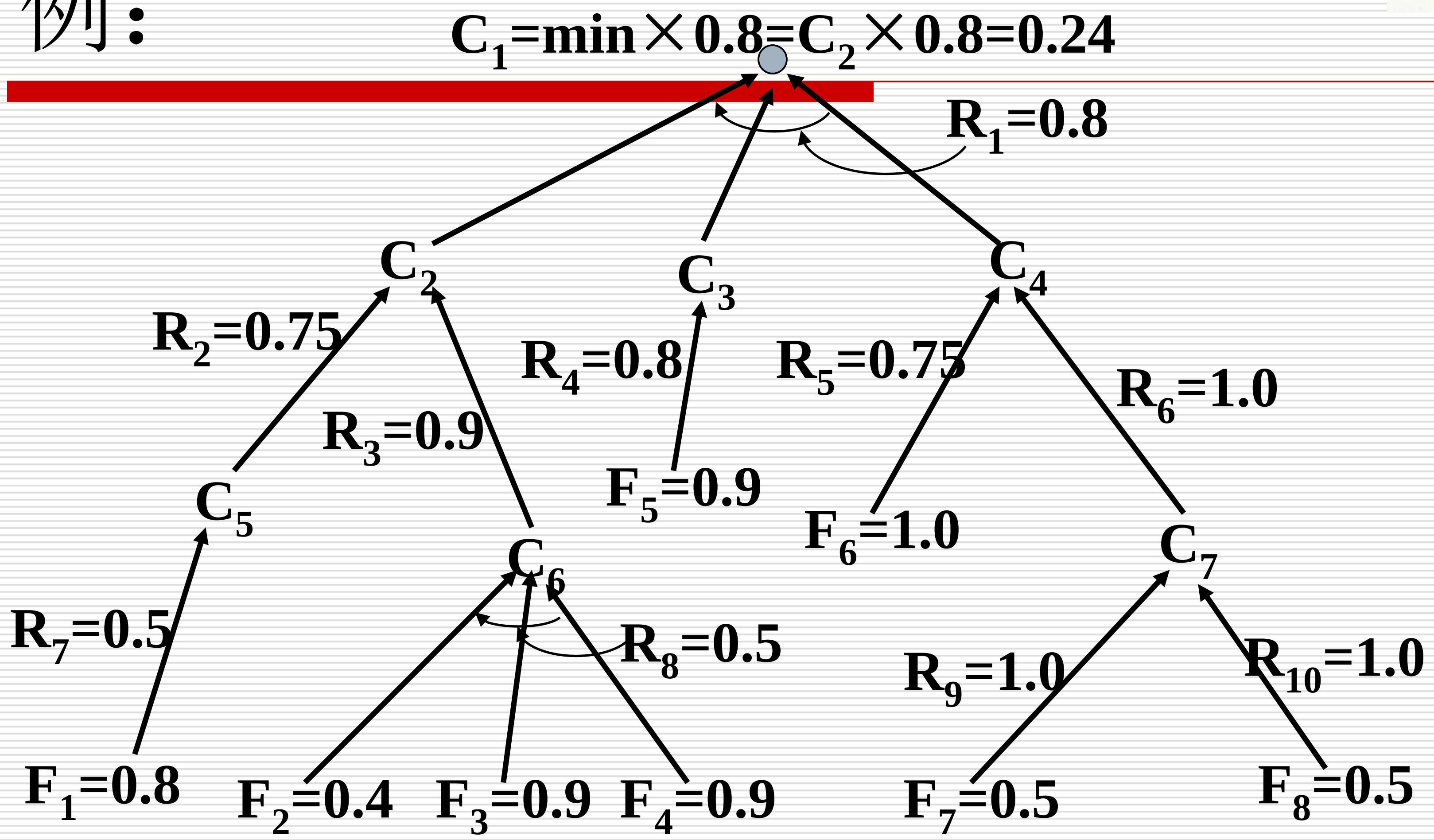
每条规则RULE和每项事实FACT各自都有一个确定的可信度(数值在[-1,1]闭区间内)，给了事实FACT的可信度F，按照规则RULE的可信度R，即可以如下地自下而上(从树叶到树根，前一层的C是后一层的F)计算出各层推断出结论CONCLUSION的可信度CF(自下而上算)：





- “与”节点处的结论可信度C=(推断规则的可信度 R)×(输入分支中的 min可信度 F或C)
- “或”节点处的结论可信度C=[(规则可信度R_1)与(输入分支1的可信度C_1)之乘积C_1R_1]+[(规则可信度R_2)与 (输入分支2的可信度C_2)之乘积C_2R_2]-(C_1R_1)×(C_2R_2)。
- 在推理过程中，一般还规定有一个统一的阈值，比方MYCIN系统是0.2；凡遇可信度≤阈值时，即置成0.0，表示谈不上可信不可信。所以在推理链上，凡遇C≤0.2者，置成C=0。

例：


C1=min×0.8=C2×0.8=0.24
R1=0.8
C2
C3
C4
R2=0.75
R3=0.9
R4=0.8
R5=0.75
R6=1.0
C5
C6
C7
F5=0.9
F6=1.0
R7=0.5
R8=0.5
R9=1.0
R10=1.0
F1=0.8
F2=0.4
F3=0.9
F4=0.9
F7=0.5
F8=0.5

其中: $C_5=0.8\times0.5=0.4$,

$C_6=\min\times0.5=0.4\times0.5=0.2\approx0$,

$C_2=0.4\times0.75=0.3$, $C_3=0.9\times0.8=0.72$

$C_7=0.5\times1.0+0.5\times1.0-$

$0.5\times1.0\times1.0\times0.5=0.75$.

$C_4=1.0\times0.75+0.75\times1.0-1.0\times0.75\times0.7\times1.0$

$=0.93$,

推理链上的可信度计算过程