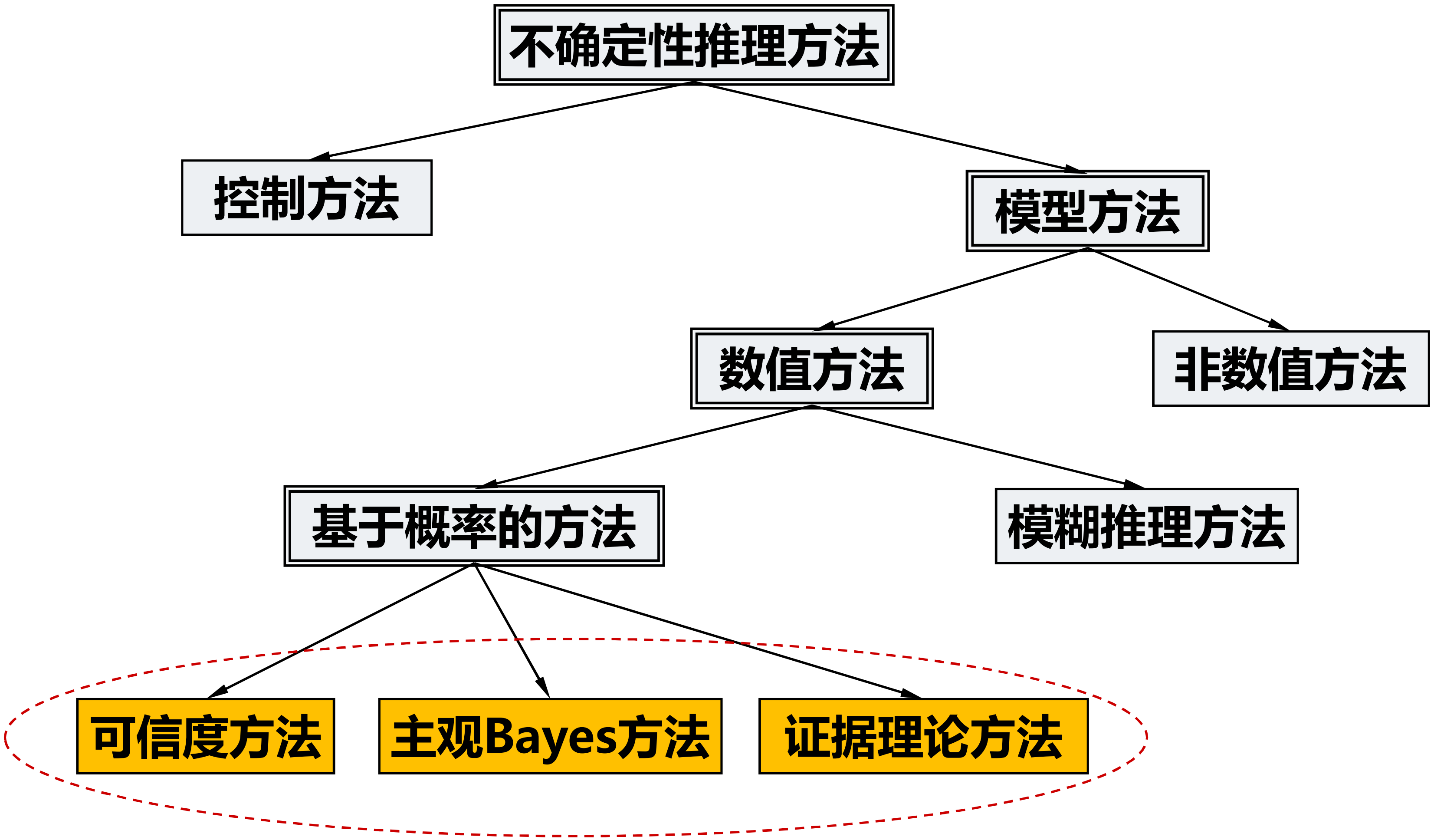
人工智能概论-不确定性推理

赵亚伟
zhaoyw@ucas.ac.cn
中国科学院大学 大数据分析技术实验室
2018.6.2

问题：现实世界中有些结论具有不确定性，如何解决不确定推理的问题?

答：概率（主观或客观的）

概述-分类



不确定性推理方法
控制方法
模型方法
数值方法
非数值方法
基于概率的方法
模糊推理方法
可信度方法
主观Bayes方法
证据理论方法

目录

概述

- **不精确来自人类的主观认识与客观实际之间存在的差异，是一种客观现象。**
  - 很多原因导致同一结果
  - 推理所需的信息不完备
  - 背景知识不足
  - 信息描述模糊
  - 信息中含有噪声
  - 规划是模糊的
  - 推理能力不足
  - 解题方案不唯一

在人类的知识和思维行为中，精确性只是相对的，不精确性才是绝对的。知识工程需要各种适应不同类的不精确性特点的不精确性知识描述方法和推理方法。

概述Cont.

- **不确定性(uncertainty)就是一个命题(亦即所表示的事件)的真实性不能完全肯定，而只能对其为真的可能性给出某种估计。**
- **例如：**
  - 如果乌云密布\电闪雷鸣，则可能要下暴雨。
  - 如果头痛发烧，则大概是患了感冒。
  - 小王是个高个子。
  - 张三和李四是好朋友。
  - 如果向左转，则身体就向左稍倾。

不确定性类型

- **现实中，不确定性的类型主要包括：**
  - 随机性
  - 模糊性
  - 不完全性 （对事物认识不足）
  - 不一致性

    随着推理的进行，原来成立的，变的不那么成立了

不确定性推理

- **推理**：从已知**事实**出发，运用相关**知识(或规则)**逐步推出**结论**或者证明某个**假设**成立或不成立的思维过程。
- **证据**：已知事实是推理过程的出发点即推理中使用的知识，我们把它称为**证据**。
- **不确定性推理**：是指从不确定性的初始证据出发，通过运用不确定性的知识，最终推理出具有一定程度的不确定性，但又是合理或者似乎合理的结论的思维过程。
- **不确定性推理**是研究复杂系统不完全性和不确定性的有力工具。

不确定性推理：不确定性

- 不确定性推理中有三种不确定性：
  - 知识（规则）的不确定性
  - 证据的不确定性
  - 结论的不确定性
- 三种不确定性均有相应的表示方式和度量标准

在专家系统中知识的不确定性一般是由领域专家给出的，通常是一个数值——知识的静态强度

- 用户在求解问题时提供的初始证据。
- 在推理中用前面推出的结论作为当前推理的证据。

由前两个不靠谱的“不确定性”推导出来的结论一定也具有不确定性

知识不确定性的表示

知识不确定性的表示

- 知识不确定性的表示方式是与不确定性推理方法密切相关的一个问题。在选择知识的不确定性表示时，通常需要考虑以下两个方面的因素：
  - **要能够比较准确地描述问题本身的不确定性**
  - **便于推理过程中不确定性的计算**
- 一般将这两个方面的因素结合起来综合考虑。 知识的不确定性通常为一个数值，也称为**知识的静态强度**。
- **在实际应用中，知识的不确定性是由领域专家给出的**

知识不确定性的表示Cont.

- 知识的**静态强度**可以是该知识在应用中成功的概率，也可以是该知识的可信程度等。
  - 如果用知识在应用中成功的概率来表示静态强度，则其取值范围为[0，1]，该值越接近于1，说明该知识越接近于“真”；其值越接近于0，说明该知识越接近于“假”。
  - 如果用知识的可信度来表示静态强度，则可其取值范围为[-1，1]（如MYCIN等专家系统），当该值大于0时，值越大说明知识越接近于“真”，当其值小于0时，值越小说明知识越接近于“假”。

证据不确定性的表示

- **证据的不确定性的表示**
  - 推理中的证据有两种来源：
    - 一种是**用户在求解问题时所提供的初始证据**，如病人的症状、检查结果等；
    - 另一种是**在推理中得出的中间结果**，即把当前推理中所得到的中间结论放入综合数据库，并作为以后推理的证据来使用。
  - 一般来说，证据的不确定性表示应该与知识的不确定性表示保持一致，以便推理过程能对不确定性进行统一处理。
  - 证据的不确定性**可以用概率来表示**，也**可以用可信度**等来表示，其意义与知识的不确定性类似。

结论不确定的表示

- 由不确定性的知识和证据，推导出来的结论也具有不确定性，这种不确定性一般用不确定“程度”来描述
- 结论一般采用 **结论+（可信）程度** 的方式进行表示，主要包括：
  - 最大最小法
  - 概率法
  - 有界法

不确定性的度量

- **采用不同的数据和方法来度量确定性的程度，在确定度量方法及其范围时，需要注意：**
  - 度量要充分表达相应知识及证据不确定性程度。
  - 度量范围的指定应便于领域专家及用户对不确定性的估计。
  - 度量要便于对不确定性的传递和计算，对结论算出的不确定性度量不能超出度量规定范围。
  - 度量的确定是直观的，同时应有相应理论基础（可解释）。

不确定性的推理：匹配算法

- **推理是一个不断运用知识的过程，为了找到所需要的知识，需要在这一过程中用知识的前提条件与证据进行匹配，只有匹配成功的知识才有可能被应用**
- **不确定性匹配算法：**
  - 设计一个用来计算匹配双方相似程度的方法，在指定一个相似的限度（阈值），用来度量匹配是否在指定限度内

核心：相似度计算

不确定性的推理：更新算法

- **推理过程中的知识不确定性的动态累积和传递，这一个过程称之为更新，更新算法包括：**
  - （**1**）在每一步推理中，如何把证据及知识的不确定性传递给结论。
  - （**2**）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。

核心：不确定性的传递计算

不确定性推理法的类型

- **关于不确定性推理的类型有多种不同的分类方法，如果按照是否采用数值来描述非精确性，可将其分为数值方法和非数值方法两大类型。**
  - **数值方法**是一种用数值对非精确性进行定量表示和处理的方法。
  - **非数值方法**是指除数值方法以外的其他各种对不确定性进行表示和处理的方法，如非单调推理等。

不确定性推理法的类型：数值方法

- **对于数值方法，又可按其所依据的理论分为两种类型**
  - 一类是基于概率论的有关理论发展起来的方法，称为**基于概率的模型**，如主观**Bayes**方法、证据理论、可信度方法等；
  - 另一类是基于模糊逻辑理论发展起来的可能性理论方法，称为**模糊推理**。

不确定问题的数学模型表示

- **不确定问题的数学模型表示的3方面问题**
  - 表示问题：

    表达要清楚。表示方法规则不仅仅是数，还要有语义描述。
  - 计算问题：

    不确定性的传播和更新。也是获取新信息的过程。
  - 语义问题：

    将各个公式解释清楚。

目录

- □ 不确定推理
- □ 概率推理
- □ 主观贝叶斯方法
- □ 可信度方法
- □ 证据理论
- □ 小结

知识体系

- **概率推理**
  - 概率论
  - 产生式规则
  - 朴素**Bayes**方法（重点）

事件概率

- 在一定条件下，可能发生也可能不发生的试验结果叫做**随机事件**，简称**事件**
- 事件发生的可能性大小是事件本身固有的一种客观属性，称这种事件发生的可能性大小为**事件的概率**
- 令A表示一个事件，概率记为P（A），P(A)具有一些基本性质，如
  - $0 \le P(A) \le 1$，必然事件D的P(D)=1，不可能事件Φ的P(Φ)=0
  - 事件A_1，A_2，…，A_k是两两互斥事件，则$P(\cup_{i=1}^{k} A_i) = P(A_1) + P(A_2) + \ldots + P(A_k)$
  - $P(\bar{A}) = 1 - P(A)$
  - 其他从略

概率推理方法

- **产生式规则**
  - **IF E THEN H_i**
    - E：前提条件，H_i：结论 (Hypothesis，假设即结论)
  - **Bayes定理：$P(E|H_i) \rightarrow P(H_i|E)$**
    - $P(H_i|E)$ ：在证据E出现的条件下，结论H_i成立的确定性程度（条件概率）

例如：

E：咳嗽，H_i：支气管炎，

条件概率$P(H_i|E)$：统计咳嗽的人中有多少是患支气管炎的。

逆概率$P(E|H_i)$：统计患支气管炎的人中有多少人是咳嗽的。

Bayes公式

□ 单证据情况

结论H_i 成立时所对应的证据出现的条件概率

结论H_i的先验概率

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

例子

H_1, H_2, H_3：**结论**，E：**证据。**

已知： $P(H_1)=0.3,\ P(H_2)=0.4,\ P(H_3)=0.5$ 先验概率

$P(E|H_1)=0.5,\quad P(E|H_2)=0.3,\quad P(E|H_3)=0.3$

求： $P(H_1|E),\quad P(H_2|E),\quad P(H_3|E)$ 后验概率

解：$$P(H_1 \mid E)=\frac{P(H_1)P(E \mid H_1)}{P(H_1)P(E \mid H_1)+P(H_2)P(E \mid H_2)+P(H_3)P(E \mid H_3)}$$

$$=\frac{0.3\times 0.5}{0.3\times 0.5+0.4\times 0.3+0.5\times 0.4}$$

$$=0.32$$

同理可得：$P(H_2 \mid E)=0.26$，$P(H_3 \mid E)=0.43$

Bayes公式

多证据情况

- 多个证据$E_1, E_2, \ldots, E_m$，多个结论$H_1, H_2, \ldots, H_n$，且每个证据都以一定程度支持结论。
- 扩充后的公式：

$$P(H_i|E_1E_2\cdots E_m) = \frac{P(E_1|H_i)\,P(E_2|H_i)\cdots P(E_m|H_i)P(H_i)}{\sum_{j=1}^{n} P(E_1|H_j)P(E_2|H_j)\cdots P(E_m|H_j)P(H_j)}$$

$i=1, 2, \ldots, n$

例子

已知

$P(H_1)=0.4$, $P(H_2)=0.3$, $P(H_3)=0.3$, $P(E_1 \mid H_1)=0.5$,

$P(E_1 \mid H_2)=0.6$, $P(E_1 \mid H_3)=0.3$, $P(E_2 \mid H_1)=0.7$,

$P(E_2 \mid H_2)=0.9$, $P(E_2 \mid H_3)=0.1$

求： $P(H_1 \mid E_1E_2)$， $P(H_1 \mid E_1E_2)$， $P(H_1 \mid E_1E_2)$ ？

解：

$$P(H_1|E_1E_2)=\frac{P(E_1|H_1)P(E_2|H_1)P(H_1)}{P(E_1|H_1)P(E_2|H_1)P(H_1)+P(E_1|H_2)P(E_2|H_2)P(H_2)+P(E_1|H_3)P(E_2|H_3)P(H_3)}$$

$$=\frac{0.5\times0.7\times0.4}{0.5\times0.7\times0.4+0.6\times0.9\times0.3+0.3\times0.1\times0.3}$$

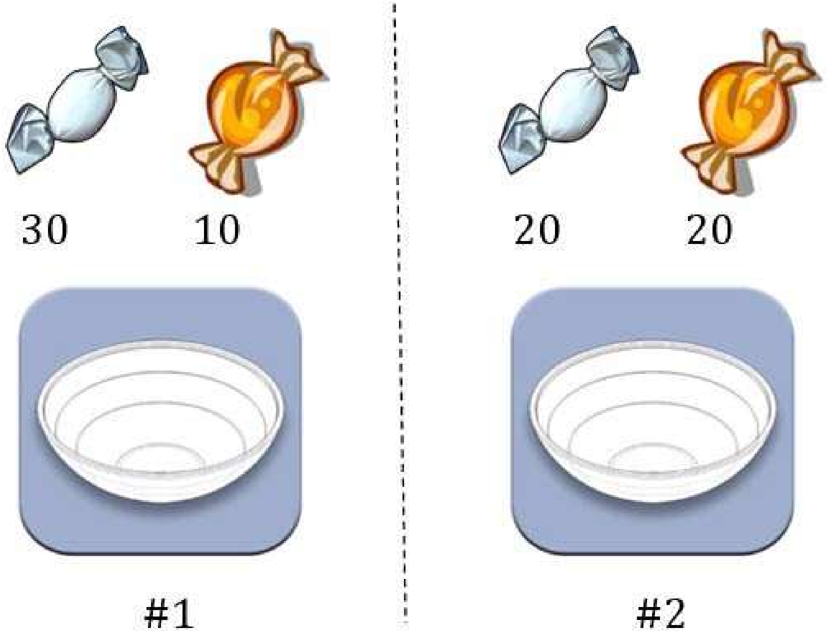
$$=0.45$$

同理可得： $P(H_2|E_1E_2)=0.52$

$P(H_3|E_1E_2)=0.03$

例子：水果糖问题

- 两个一模一样的碗，一号碗有**30**颗水果糖和**10**颗巧克力糖，二号碗有水果糖和巧克力糖各**20**颗。现在随机选择一个碗，从中摸出一颗糖，发现是水果糖。请问这颗水果糖来自一号碗的概率有多大？



- 我们假定，**H1**表示一号碗，**H2**表示二号碗。由于这两个碗是一样的，所以**P(H1)=P(H2)**，也就是说，在取出水果糖之前，这两个碗被选中的概率相同。因此，**P(H1)=0.5**，我们把这个概率就叫做“**先验概率**”，即没有做实验之前，来自一号碗的概率是**0.5**。（故，也称**事前概率**（事件发生之前的概率））

- 再假定，**E**表示水果糖，所以问题就变成了在已知**E**的情况下，来自一号碗的概率有多大，即求**P(H1|E)**。我们把这个概率叫做“**后验概率**”，即在**E**事件发生之后，对**P(H1)**的修正（故，也称**事后概率**）。

根据条件概率公式，得到 $P(H_1|E)=P(H_1)\frac{P(E|H_1)}{P(E)}$

已知，$P(H_1)$等于0.5，$P(E|H_1)$为一号碗中取出水果糖的概率，等于0.75，那么求出P(E)就可以得到答案。根据全概率公式，

$$P(E)=P(E|H_1)P(H_1)+P(E|H_2)P(H_2)$$

$$P(E)=0.75\times0.5+0.5\times0.5=0.625$$

将数字代入原方程，得到

$$P(H1|E)=0.5\times\frac{0.75}{0.625}=0.6$$

这表明，来自一号碗的概率是0.6。也就是说，取出水果糖之后，H_1事件的可能性得到了增强。

例子：预测问题

- **已知一些数据，如下表。现有一条记录：**
  X=(AGE=“<30”,INCOME=medium,STUDENT=yes,CREDIT_RAT=fair)
- **类别(预测)属性为BAYS_COMPU, 问: 其X的BAYS_COMPU值为yes还是no?**

ID	AGE	INCOME	STUDENT	CREDIT_RAT	BAYS_COMPU
1	<30	High	no	Fair	no
2	<30	High	no	Excellent	no
3	30-40	High	no	Fair	yes
4	>40	Medium	no	Fair	yes
5	>40	Low	yes	Fair	yes
6	>40	Low	yes	Excellent	no
7	30-40	Low	yes	Excellent	yes
8	<30	Medium	no	Fair	no
9	<30	High	yes	Fair	yes
10	>40	Medium	yes	Fair	yes
11	<30	Medium	yes	Excellent	yes
12	30-40	Medium	no	Excellent	yes
13	30-40	Medium	yes	Fair	yes
14	>40	Low	no	Excellent	no

b_k： 已知属性

A_i： 类别属性

计算过程

贝叶斯分类器的训练就是计算先验概率、条件概率等值，当一条新的记录出现，可以根据这些取值计算新记录的属性类别（后验概率），实现预测。

- **计算先验概率 $P(A_i)$**
  P(BAYS_COMP=yes)=9/14
  P(BAYS_COMP=no)=5/14
- **计算条件概率 $P(b_k|A_i)$**
  P(AGE=‘<30’|BAYS_COMP=yes)=2/9
  P(AGE=‘<30’|BAYS_COMP=no)=3/5
  P(INCOME=‘Medium’|BAYS_COMP=yes)=5/9
  P(INCOME=‘Medium’|BAYS_COMP=no)=1/5
  同样计算STUDENT=yes,CREDIT_RAT=fair的条件概率（略）
- **计算 $\Pi(P(b_k|A_i))$**
  P(A|BAYS_COMP=yes)=2/9*5/9
  P(A|BAYS_COMP=no)=3/5*1/5
- **计算后验概率 $\Pi(P(b_k|A_i))P(A_i)$（省略了STUDENT,CREDIT_RAT，正常应计算)**
  P(A|BAYS_COMP=yes)* P(BAYS_COMP=yes)=2/9*5/9*9/14=0.079
  P(A|BAYS_COMP=no)* P(BAYS_COMP=no)=3/5*1/5*5/14=0.042
- **二者除以相同的P(B)，不计算了。**
- **结论：由于0.079>0.042，因此，X的BAYS_COMPU的取值为“yes”**

优缺点分析

- 优点：较强的理论背景和良好的数学特征，当证据及结论都彼此独立时计算的复杂度比较低。

- 缺点：要求给出结论$\boldsymbol{H_i}$ 的先验概率$\mathbf{P}(\boldsymbol{H_i})$ 及证据$\boldsymbol{E_j}$的条件概率 $\mathbf{P}(\mathbf{E_j}|\mathbf{H_i})$。 $\mathbf{H_i}$的独立性很难保证。

目录

- 不确定推理
- 概率推理
- **主观贝叶斯方法**
- 可信度方法
- 证据理论
- 小结

主观贝叶斯方法

概述

- 在**Prospector**的探矿系统（**1981**年，斯坦福研究所**SRI**）的研究过程中提出的。
- 原有贝叶斯公式只考虑**E**出现对**H**的影响，没有考虑**E**不出现的影响。
- 直接根据贝叶斯公式进行推理计算简单明了，但是它要求结论H_i相互独立，实际上难以保证。而且$P(E|\ H_i)$的计算通常比较困难。
- 所以在求解不确定性推理问题时，还不能直接使用贝叶斯公式，而是使用对其经过改进的主观贝叶斯公式

内容体系

- **主观Bayes方法**
  - 知识不确定性的表示
    - 充分性因子**LS**
    - 必要性因子**LN**
    - 几率 **O(X)**
  - 证据不确定性的表示
    - 可信度**C(E|S)**
    - **EH**公式
    - **CP**公式
  - 推理
    - 已知**P(E)**、**LS**、**LN**、**P(H)**先验概率等，求后验概率**P(H|E)**
    - 已知**C(E|S)**、**LS**、**LN**、先验几率，求后验几率**O(H|S)**

主观贝叶斯方法：知识不确定性的表示

- **在主观Bayes方法中，知识是用产生式规则表示的，具体形式为**

  IF E THEN (LS，LN) H (P(H))

- **其中**
  - （1）**E** 是该知识的前提条件。它既可以是一个简单条件，也可以是复合条件。
  - （2）**H** 是结论。**P(H)**是 **H** 的先验概率，它指出在没有任何证据情况下的结论 **H** 为真的概率，即 **H** 的一般可能性。其值由领域专家根据以往的实践及经验给出。
  - （3）**(LS, LN)**为规则强度。其值由**领域专家**给出。**LS**、**LN**相当于知识的**静态强度**。

主观的体现

知识的静态强度

- 规则的**充分性度量LS** ：表示E为真时，对H的影响。(规则成立的充分性)
  - 取值范围

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

H为真时**E**出现的概率除以**H**为假时**E**出现的概率

- 规则的**必要性度量LN** ：表示E为假时，对H的影响。(规则成立的必要性)
  - 取值范围

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

H为真时**E**不出现的概率除以**H**为假时**E**不出现的概率

证据E肯定存在的情况

- 证据E肯定存在，即P(E)=1时：
  - 由基本贝叶斯公式，可得：

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

$$P(\sim H/E)=\frac{P(E|\sim H)\times P(\sim H)}{P(E)}=P(E/\sim H)\times P(\sim H)$$

  - 两式相除得：

$$\frac{P(H/E)}{P(\sim H/E)}=\frac{P(E/H)}{P(E/\sim H)}\times\frac{P(H)}{P(\sim H)}=LS\times\frac{P(H)}{1-P(H)}$$

几率函数及更新

- 几率（Odds）函数O(X)：单调性与P(X)相同，(0, +∞)

$$O(X) = \frac{P(X)}{P(\sim X)} = \frac{P(X)}{1 - P(X)} \quad P(X) = \frac{O(X)}{1 + O(X)}$$

- 采用规则的充分性LS计算

$$\frac{P(H/E)}{P(\sim H/E)} = LS \times \frac{P(H)}{1 - P(H)}$$

$$O(H/E) = LS \times O(H)$$

O(X)与LN的关系

E肯定出现的情况下，H的先验几率更新为后验几率的公式

LS越大，O(H|E)就越大，且P(H|E)也越大，说明E对H的支持越强
当LS→∞时， O(H|E) →∞，P(H|E) →1，说明E的存在导致H为真
因此，说明E对H是充分的，称LS为充分性因子

证据E肯定不存在的情况

- **证据E肯定不存在，即P(E)=0时：**
  - **P(～E)=1**，由贝叶斯公式，可得：

$$P(H/\sim E)=\frac{P(\sim E\,|\,H)\times P(H)}{P(\sim E)}=P(\sim E/H)\times P(H)$$

$$P(\sim H/\sim E)=\frac{P(\sim E\,|\sim H)\times P(\sim H)}{P(\sim E)}=P(\sim E/\sim H)\times P(\sim H)$$

  - 两式相除得：

$$\frac{P(H/\sim E)}{P(\sim H/\sim E)}=\frac{P(\sim E/H)}{P(\sim E/\sim H)}\times\frac{P(H)}{P(\sim H)}=LN\times\frac{P(H)}{1-P(H)}$$

更新

- 采用规则的充分性LS和几率函数O(H)计算

$$\frac{P(H/\sim E)}{P(\sim H/\sim E)} = LN \times \frac{P(H)}{P(\sim H)}$$

$$O(H/\sim E) = LN \times O(H)$$

O(X)与LN的关系

E肯定不出现的情况下，H的先验几率更新为后验几率的公式

$$\left.\begin{aligned} O(H/E) &= LS \times O(H) \\ O(H/\sim E) &= LN \times O(H) \end{aligned}\right\}$$

修正的贝叶斯公式

LN反映了~H的出现对H的支持程度
当LN=0时，将使 O(H|~E)=0，说明E的不存在导致H为假
因此，说明E对H是必要的，称LS为必要性因子

小结：LS与LN的取值与意义

$$LS=\begin{cases}1 & O(H|E)=O(H) & E\text{对}H\text{没影响}\\ >1 & O(H|E)>O(H) & E\text{支持}H\\ <1 & O(H|E)<O(H) & E\text{不支持}H\end{cases}$$

$$LN=\begin{cases}1 & O(H|\sim E)=O(H) & \sim E\text{对}H\text{没影响}\\ >1 & O(H|\sim E)>O(H) & \sim E\text{支持}H\\ <1 & O(H|\sim E)<O(H) & \sim E\text{不支持}H\end{cases}$$

，<u>**且必须满足**</u>：<u>**对LS、LN赋值时**</u>

<u>**LS、LN**≥ 0 。</u>

<u>**LS, LN**不能同时 ＞ 1 或 ＜ 1</u>

<u>**LS, LN**可同时＝**1**</u>

证据不确定性的表示

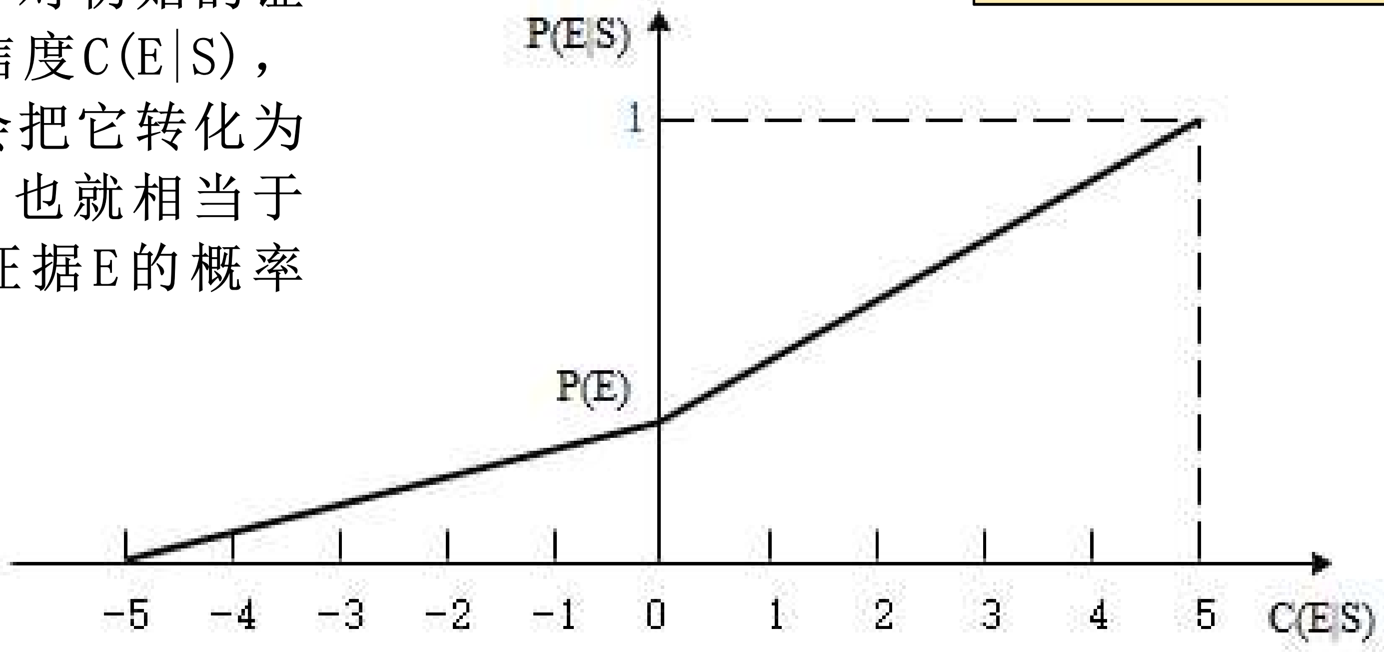
- 主观贝叶斯方法中证据不确定性也是用概率表示的。对于初始证据E，用户根据观察S给出P(E|S)，相当于动态强度。
- 由于难于给出P(E|S)，因而在具体应用系统中往往采用变通方法，如在Prospector中引进了可信度的概念，让用户在 -5 ~ +5 之间的11个整数中根据实际情况选一个数作为初始证据的可信度C(E|S)
  - C(E|S) = -5，表示在观察S下证据E肯定不存在，即P(E|S) = 0
  - C(E|S) = 0，表示S与E无关，即P(E|S) = P(E)
  - C(E|S) = 5，表示在观察S下证据E肯定存在，即P(E|S) = 1
  - C(E|S) 为其他数时，与P(E|S)的对应关系可通过线性插值得到

C(E|S) 与 P(E|S)的对应关系

$$P(E|S)=\begin{cases}\dfrac{C(E|S)+P(E)\times(5-C(E|S))}{5}, 若 0\le C(E|S)\le 5\\[2ex] \dfrac{P(E)\times(C(E|S)+5)}{5}, 若 -5\le C(E|S)\le 0\end{cases}$$

C(E|S)由用户指定

只要用户对初始的证据的可信度C(E|S)，系统就会把它转化为P(E|S)，也就相当于给出了证据E的概率P(E|S)



证据不确定：P(E|S) = 1 、0 或 P(E)时

- 当证据不确定（~E）时，采用杜达（Duda）等人证明的下列公式计算后验概率

$$P(H|S) = P(H|E)P(E|S)+P(H|\sim E)\ P(\sim E|S)$$

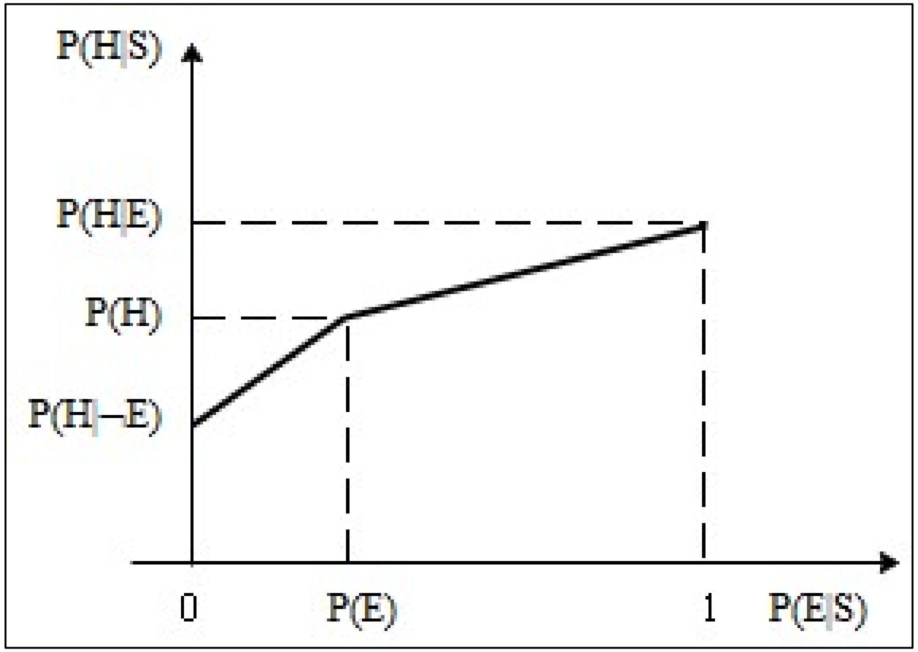
  - 当P(E|S) = 1时， P(~ E|S) = 0， P(H|S) = P(H|E)
  - 当P(E|S) = 0时， P(~ E|S) = 1， P(H|S) = P(H| ~ E)
  - 当P(E|S) = P(E)时，

$$P(H|S) = P(H|E)P(E|S)+P(H|\sim E)\ P(\sim E|S)$$
$$=P(H|E)P(E)+P(H|\sim E)\ P(\sim E)$$
$$=P(HE)+P(H\ \sim E)$$
$$=P(H)$$

证据不确定：P(E|S) 为其他值时

- P(E|S)为其他值时，用线性插值得 EH 公式：

$$P(H/S)=\begin{cases} P(H/\neg E)+\dfrac{P(H)-P(H/\neg E)}{P(E)}\times P(E/S), 若 0\le P(E/S)<P(E) \\ P(H)+\dfrac{P(H/E)-P(H)}{1-P(E)}\times[P(E/S)-P(E)], 若 P(E)\le P(E/S)\le 1 \end{cases}$$



CP公式

- **根据用户告知的 C(E|S)，求P(H|S)**
  - 将P(E|S) (P45) 代入上述EH公式，则得CP公式：

$$
P(H|S) = \begin{cases} P(H|\sim E) + [P(H) - P(H|\sim E)] \times \left[\frac{1}{5} C(E|S) + 1\right], \text{若} C(E|S) \le 0 \\ P(H) + [P(H|E) - P(H)] \times \frac{1}{5} C(E|S), \text{若} C(E|S) > 0 \end{cases}
$$

主观贝叶斯推理I：后验概率P(H|E)

- 主观**Bayes**方法的不精确推理过程就是根据**证据E的概率P(E)**，利用规则的**LS**和**LN**，把结论**H**的**先验概率P(H)**更新为**后验概率P(H|E)**的过程

- **推理计算 — 不确定性的传递：**

  已知规则**E→H**的**(LS,LN)**和**P(H)**、**P(E)**

  如何计算**P(H/E)**或**P(H/～E)**？

$$P(H)\xrightarrow{PE,LS,LN} P(H/E)\text{或}P(H/\neg E)$$

先验概率

后验概率?

主观贝叶斯推理I：后验概率P(H|E)

- **由已知LS、P(H)，通过下式可更新P(H)为P(H|E)**

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

如果针对E不出现情况，P(H|~E)更新则需要借助LN

$$P(H/\sim E)=\frac{LN\times P(H)}{(LN-1)\times P(H)+1}$$

- **证明：将LS式代入上式**

$$\mathrm{LS}=\frac{\mathrm{P(E\,|\,H)}}{\mathrm{P(E\,|\!\sim H)}}$$

由于

$$\frac{P(H/E)}{P(\sim H/E)}=\frac{P(E/H)}{P(E/\sim H)}\times\frac{P(H)}{P(\sim H)}=LS\times\frac{P(H)}{1-P(H)}$$

证明二者相等即可。

例子

- 设有如下知识：

r_1 : IF E_1 THEN (10,1) H_1 (0.03)

r_2 : IF E_2 THEN (20,1) H_2 (0.05)

r_3 : IF E_3 THEN (1,0.002) H_3 (0.3)

- 求：当证据 E_1, E_2, E_3 存在及不存在时，$P(H_i \mid E_i)$ 及 $P(H_i \mid \neg E_i)$ 的值各是多少？

- 解：

$$P(H_1|E_1)=\frac{LS_1\times P(H_1)}{(LS_1-1)\times P(H_1)+1}$$

$$=\frac{10\times 0.03}{(10-1)\times 0.03+1}$$

$$=0.24$$

$$\frac{P(H_1|E_1)}{P(H_1)}=\frac{0.24}{0.03}=8$$

$$P(H_2|E_2)=\frac{LS_2\times P(H_2)}{(LS_2-1)\times P(H_2)+1}$$

$$=\frac{20\times 0.05}{(20-1)\times 0.05+1}$$

$$=0.51$$

$$\frac{P(H_2|E_2)}{P(H_2)}=\frac{0.51}{0.05}=10.2$$

□ 解（续）：

$$P(H_3|\neg E_3)=\frac{LN_3\times P(H_3)}{(LN_3-1)\times P(H_3)+1}$$

$$=\frac{0.002\times 0.3}{(0.002-1)\times 0.3+1}$$

$$=0.00086$$

$$\frac{P(H_3)}{P(H_3|\neg E_3)}\approx 350$$

主观贝叶斯推理II：求P(H|S)

- **求P(H|S)**
  - 当采用初始证据进行推理时，通过提问用户得到**C(E|S)**，通过**CP**公式就可求得**P(H|S)**
  - 当采用推理过程中得到的中间结论作为证据进行推理时，通过**EH**公式可求得**P(H|S)**
- **如果有n条知识都支持同一结论H，而且每条知识的前提条件分别是n个相互独立的证据$E_1, E_2, \ldots, E_n$，而这些证据又分别与观察$S_1, S_2, \ldots, S_n$相对应。**
  - 首先对每条知识分别求出**H**的后验几率**O(H|S)**
  - 然后按下述公式求出所有观察下**H**的后验几率

$$O(H|S_1,S_2,\cdots,S_n)=\frac{O(H|S_1)}{O(H)}\times\frac{O(H|S_2)}{O(H)}\times\cdots\times\frac{O(H|S_n)}{O(H)}\times O(H)$$

例子：不确定传递算法

- **已知下列规则**

  R1： IF E_1 THEN (2, 0.000001) H_1

  R2： IF E_2 THEN (100, 0.000001) H_1

  R3： IF E_3 THEN (65, 0.01) H_2

  R4： IF E_4 THEN (300, 0.001) H_2

  且先验几率$O(H_1) = 0.1$，$O(H_2) = 0.01$，通过用户得到$C(E_1|S_1) = 3$，$C(E_2|S_2) = 1$，$C(E_3|S_3) = -2$。

  求：$O(H_2|S_1, S_2, S_3)$

- 解题思路
- 根据联合几率公式

$$O(H|S_1,S_2,\cdots,S_n)=\frac{O(H|S_1)}{O(H)}\times\frac{O(H|S_2)}{O(H)}\times\cdots\times\frac{O(H|S_n)}{O(H)}\times O(H)$$

- 可知，需要计算$O(H|S_i)$、$O(H)$，根据几率公式

$$O(X)=\frac{P(X)}{P(\sim X)}=\frac{P(X)}{1-P(X)}$$

- 可知，需要计算$P(H|S_i)$和$P(H)$，根据几率公式可知

$$P(H_1|E_1)=\frac{O(H_1|E_1)}{1+O(H_1|E_1)}=\frac{LS_1O(H_1)}{1+LS_1O(H_1)}=\frac{2\times0.1}{1+2\times0.1}=\frac{1}{6}$$

- 根据CP公式计算$P(H_1|S_1)$，最后计算$O(H_1|S_1)$
- 同理可求得$O(H_1|S_2)$、 $O(H_1|S_3)$，参考教材。

主观贝叶斯方法评价

- **主观Bayes方法的评价**
  - 主观**Bayes**方法的主要优点：
    - 具有较坚实的理论基础。
    - 知识的静态强度 **LS** 及**LN** 是由领域专家根据实践经验给出的，推出的结论有较准确的确定性。
    - 主观**Bayes**方法是一种比较实用且较灵活的不确定性推理方法。
  - 主观**Bayes**方法的主要缺点：
    - 要求领域专家在给出知识时，同时给出**H**的先验概率。
    - **Bayes**定理中关于事件独立性的要求使主观**Bayes** 方法的应用受到了限制。

目录

- 不确定推理
- 概率推理
- 主观贝叶斯方法
- 可信度方法
- 证据理论
- 小结

知识体系

- **可信度方法**
  - 基于可信度的不确定性表示
    - 知识的不确定性表示
    - 证据不确定表示
  - 可信度方法的推理算法
    - 组合证据的不确定算法
    - 不确定性传递算法
    - 多个独立证据退出同一假设的合成算法

可信度

- **可信度方法是美国斯坦福大学E.H.Shortliffe等人在确定性理论的基础上，结合概率论等提出的一种不确定性推理方法。1976年在专家系统MYCIN中首先应用，它是不确定推理方法中应用最早、且简单有效的方法之一。**
- **什么是可信度?**
  - 根据经验对一个事物或现象为真的相信程度称为可信度。
  - 可信度也称作确定性因子。用以度量知识和证据的不确定性。可信度具有较大的主观性和经验性。

知识不确定性的表示

- **在该模型中，知识是用产生式规则表示的，不确定性以可信度CF(H,E)表示。**
  - 一般形式：**IF　E　THEN　　H　(CF(H, E))**
  - 其中：
    - （**1**）**E**是知识的前提或称为证据，可以是命题的合取、析取组合等。
    - （**2**）结论**H**可为单一命题，也可以是复合命题。
    - （**3**）**CF(H,　E)**为确定性因子，简称可信度，用以量度规则的确定性（可信）程度。

信任增长度和不信任增长度

- **在MYCIN中**
  - **CF(H, E) = MB(H, E) - MD(H, E)**
- **MB(H, E)(Measure Belief)指信任增长度，表示因与E匹配的证据出现，使H为真的信任增长度。定义如下：**

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

- **MD(H, E)(Measure Disbelief)指不信任增长度，表示因与E匹配的证据出现，使H为真的不信任增长度。定义如下：**

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

MB、MD、CF性质

- **性质1：取值范围**
  - **$0 \le MB(H, E) \le 1$**
  - **$0 \le MD(H, E) \le 1$**
  - **$-1 \le CF(H, E) \le 1$**
- **性质2：互斥性**
  - **MB**和**MD**是互斥的。即：
    - 当**MB>0**时，**MD=0**
    - 当**MD>0**时，**MB=0**

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

MB、MD、CF性质Cont.

□ 性质3：相互影响

- 若 $P(H|E) = 1$，即E为真则H为真时，则 $MB(H|E) = 1$，$MD(H|E) = 0$，$CF(H|E) = 1$
- 若 $P(H|E) = 0$，即E为真则H为假时，则 $MB(H|E) = 0$，$MD(H|E) = 1$，$CF(H|E) = -1$
- 若 $P(H|E) = P(H)$，即E对H没有影响，则 $MB(H|E) = 0$，$MD(H|E) = 0$，$CF(H|E) = 0$

□ 性质4：独立性原则

- 对于同一个证据E，若存在n个互不相容的假设Hi（i = 1, 2, …, n），则有

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

MB、MD、CF性质Cont.

性质5：对应关系

- 可行度CF(H|E)与概率P(H|E)有一定的对应关系，但又有区别，对于P(H|E)，有

$$P(H|E) + P(\sim H|E) = 1$$

- 而对于CF(H|E)，有

$$CF(H|E) + CF(\sim H|E) = 0$$

- 表明：如果一个证据对某个假设成立有利，则就必然对假设不成立不利，而且二者影响程度相同，方向相反

由公式知，计算CF(H, E)需要已知P(H)与P(H|E)，然而，在实际应用中这两个值很难获得，而是在建立规则库时由领域专家凭经验主观确定的。

证据不确定性的表示

- **证据E的不确定性用证据的可行度CF(E)表示**
  - 初始证据的可信度由用户在系统运行时提供
  - 中间结果由不精确推理算法求得
- **CF(E)的取值范围与CF(H|E)相同，即$-1 \le CF(E) \le 1$**
  - 当证据以某种程度为真时， $CF(E) > 0$
  - 当证据肯定为真时， $CF(E) = 1$
  - 当证据以某种程度为假时， $CF(E) < 0$
  - 当证据以某种程度为假时， $CF(E) = -1$
  - 当证据一无所知时， $CF(E) = 0$

可信度方法的推理算法

- 1、**组合证据不确定性的算法**

有时因为缺失一个很小的证据，就可能使一个罪犯逃脱法律制裁

  - (1) 当组合证据是多个单一证据的**合取**时，即：
    - E = E_1 AND E_2 AND … AND E_n
    - 则 $CF(E) = \min\{CF(E_1), CF(E_2) \ldots CF(E_n)\}$
  - (2) 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
    - 则 $CF(E) = \max\{CF(E_1), CF(E_2) \ldots CF(E_n)\}$

如果任意一个证据都可以对一个罪犯绳之以法，那么就要那个最有说服力的

可信度方法的推理算法Cont.

2、不确定性的传递算法

- 不确定性的传递算法定义如下：

$$CF(H) = CF(H,E) \times \max\,[0，CF(E)]$$

- 由上式可以看出：
  - （1）CF(E) ＜ 0时，CF(H)=0，说明该模型没有考虑证据为假时对结论H所产生的影响。
  - （2）CF(E) ＝ 1时，CF(H)=CF(H，E)，说明规则可信度CF(H, E)就是证据为真时的结论H的可信度。

可信度方法的推理算法Cont.

3、结论不确定性的合成算法

- 若由多条不同知识推出了相同的结论，但可信度不同，则可用合成算法求出综合的可信度。由于对多条知识的综合可通过两两的合成实现，所以下面只考虑两条知识的情况。
- 设有如下知识：

  IF E1 THEN H （CF(H,E1))

  IF E2 THEN H （CF(H,E2))

- 则结论**H**的综合可信度由两步算出：

  （**1**）首先分别对每一条知识求出**CF(H)**

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

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

  （**2**）求出E_1和E_2对**H**的综合影响所形成的可信度$CF_{1,2}(H)$

可信度方法的推理算法Cont.

$$\mathrm{CF1,2(H)}=\begin{cases}\mathrm{CF1(H)}+\mathrm{CF2(H)}-\mathrm{CF1(H)}\times\mathrm{CF2(H)} & \text{若}\ \begin{matrix}\mathrm{CF1(H)}\ge 0\\ \mathrm{CF2(H)}\ge 0\end{matrix}\\ \mathrm{CF1(H)}+\mathrm{CF2(H)}+\mathrm{CF1(H)}\times\mathrm{CF2(H)} & \text{若}\ \begin{matrix}\mathrm{CF1(H)}<0\\ \mathrm{CF2(H)}<0\end{matrix}\\ \dfrac{\mathrm{CF1(H)}+\mathrm{CF2(H)}}{1-\min\{|\mathrm{CF1(H)}|,|\mathrm{CF2(H)}|\}} & \text{若CF1(H)与CF2(H)异号}\end{cases}$$

例子

- 设有如下一组知识:
  - R1: IF E1 THEN H (0.8)
  - R2: IF E2 THEN H (0.6)
  - R3: IF E3 THEN H (-0.5)
  - R4: IF E4 AND (E5 OR E6) THEN E1 (0.7)
  - R5: IF E7 AND E8 THEN E3 (0.9)
- 已知: CF(E2)=0.8 CF(E4)=0.5 CF(E5)=0.6
- CF(E6)=0.7 CF(E7)=0.6 CF(E8)=0.9
- 求: 综合可信度CF(H) = ? （$CF_{1,2,3}(H)$）

例子

- **解：由R4:**

$CF(E_1)=0.7\times\max\{0,CF[E_4 \text{ AND}(E_5 \text{ OR } E_6)]\} = 0.35$

由**R5: $CF(E_3)=0.54$**

由**R1: $CF_1(H)=0.28$**

由**R2: $CF_2(H)=0.48$**

由**R3: $CF_3(H)=-0.27$**

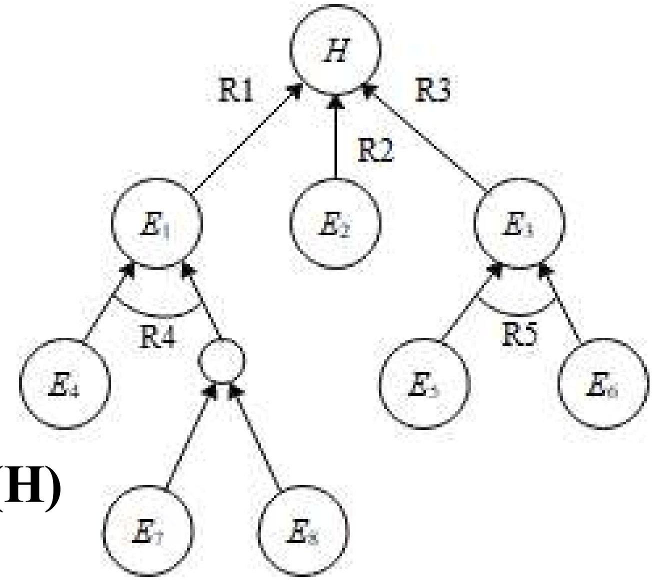
根据结论不确定性的合成算法得到:

$CF_{1,2}(H)= CF_1(H)+ CF_2(H)+ CF_1(H)\times CF_2(H)$

$= 0.63$

$$CF_{1,2,3} = \frac{CF_{1,2}(H) + CF_3(H)}{1 - \min\{|CF_{1,2}(H)|, |CF_3(H)|\}}$$

$$= 0.49$$



目录

证据理论（Evident Theory）

概述

- 诞生：又称D-S理论（Dempster - Shafer），源于20世纪60年代美国哈佛大学数学家A. P. Dempster在利用上、下限概率来解决多值映射问题方面的研究工作。自1967年起连续发表了一系列论文，标志着证据理论的正式诞生。
- 形成：Dempster的学生G. Shafer对证据理论做了进一步的发展，引入信任函数概念，形成了一套基于“证据”和“组合”来处理不确定性推理问题的数学方法，并于1976年出版了《证据的数学理论》，这标志着证据理论正式成为一种处理不确定性问题的完整理论。
- 适用领域：信息融合、专家系统、情报分析、法律案件分析、多属性决策分析，等等。

D-S理论的基本思想

- **证据理论是用集合表示命题的。**
  - 设**D**是变量***x***的样本空间，其中具有***n***个元素，在任一时刻变量***x***的取值都会落入某个子集，也就是说，**D**的任一子集**A**都对应于一个关于***x***的命题，该命题为"***x***的值在**A**中"，所以用集合**A**表示该命题。
- **证据理论基于概率分配函数计算信任函数值和似然函数值，实现推理过程**

核心：
概率分配函数

证据理论概念

- **证据理论涉及的概念**
  - 概率分配函数
  - 信任函数
  - 似然函数
  - 信任函数与似然函数的关系
  - 概率分配函数的正交和（证据的组合）
  - 基于证据理论的不确定推理

证据

- **证据:**
  - 用集合**D**来表示：如**D**中的每个元素代表一种疾病。讨论一组疾病**A**发生的可能性时，**A**变成了单元（某些假设）的集合。**D**内元素$\mathbf{A}_i$间是互斥的，但$\mathbf{A}_i$中元素间是不互斥的。

概率分配函数

- 设 D 是变量 x 所有可能取值的集合，且 D 中的元素是互斥的，在任一时刻 x 都取且只能取 D 中的某一个元素为值，则称 D 为 x 的样本空间。

- 在证据理论中，D 的任何一个子集 A 都对应于一个关于 x 的命题，称该命题为“x 的值是在 A 中”。

- 设 x：所看到的颜色，D={红，黄，蓝}，

  则 A={红}：“x 是红色”；

  A={红，蓝}：“x 或者是红色，或者是蓝色”。

概率分配函数

- 设D为样本空间，领域内的命题都用D的子集表示，则**概率分配函数**（basic probability assignment function）定义如下：

定义1　设函数 M：$2^D \to [0, 1]$ （对任何一个属于D的子集A，命它对应一个数$M \in [0，1]$） 且满足

$$M(\Phi) = 0$$

$$\sum_{A \subseteq D} M(A) = 1$$

称M为2^D上的基本概率分配函数，$M(A)$为 A的基本概率数，表示对A的精确信任。

概率分配函数

几点说明：

（1）设样本空间D中有n个元素，则D中子集的个数为2^n个

2^D：D的所有子集。

（2）概率分配函数：把D的任意一个子集A都映射为[0，1]上的一个数$M(A)$，**事实就上是对D的各个子集进行信任分配**

$A \subset D$时，$M(A)$是对相应命题A不确定性的度量。

（3）概率分配函数与概率不同。

举例

- 设 D={红，黄，蓝}
  - 则其子集个数 $2^3=8$，具体为：
    - A={红}，A={黄}，A={蓝}，A={红，黄}，A={红，蓝}，A={黄，蓝}，A={红，黄，蓝}，A={ Φ }
  - A={红}，$M(A)$=0.3表示命题"x是红色"的分配的信任度是0.3
  - M（{红}）=0.3，M（{黄}）=0，M（{蓝}）=0.1，M（{红，黄}）=0.2，M（{红，蓝}）=0.2，M（{黄，蓝}）=0.1，M（{红，黄，蓝}）=0.1，M（Φ）=0
  - 但：M（{红}）+ M（{黄}）+ M（{蓝}）=0.4，所以，概率分配函数与概率不同

信任函数

定义2　命题的信任函数（belief function）Bel:

$$2^D \rightarrow [0,1] \text{且} \ Bel(A) = \sum_{B \subseteq A} M(B) \quad \forall A \subseteq D$$

$Bel(A)$：对命题A为真的总的信任程度，是A的所有子集的基本概率数之和

- 由信任函数及概率分配函数的定义推出：

$$Bel(\Phi) = M(\Phi) = 0$$

$$Bel(D) = \sum_{B \subseteq D} M(B) = 1$$

*Bel*类似于概率密度函数，表示A中所有子集的基本概率分配数值的和，用来表示对A的总信任度

设 ***D*** ={红，黄，蓝}

M（{红}）=**0.3**，***M***（{黄}）=**0**，***M***（{红，黄}）=**0.2**，

Bel({红，黄}) ＝ M({红}) ＋ M({黄}) ＋ M({红，黄}) =0.3 + 0.2 =0.5

似然函数

- **似然函数**（plansibility function）：**表示对A为非假（不否定A）的信任程度，它是所有与A相交的子集的基本概率数之和。似然函数又称为上限函数。**

定义3 似然函数 $Pl : 2^{D} \to [0, 1]$ 且

$Pl(A) = 1 - \mathrm{Bel}(\neg A)$ 对所有的 $A \subseteq D$

设 D ={红，黄，蓝}

M（{红}）=0.3，M（{黄}）=0，M（{红，黄}）=0.2，

Bel({红, 黄}) = M({红}) + M({黄}) + M({红, 黄})

=0.3 + 0.2 =0.5

Pl({蓝}) = 1- Bel(~{蓝}) = 1- Bel({红, 黄}) = 1 - 0.5 = 0.5

信任函数与似然函数的关系

- **信任函数与似然函数具有如下关系**
  - $Pl(A) \geq \text{Bel}(A)$　　非假 ≥ 真
- **证明**

因为 $\text{Bel(A)} + \text{Bel}(\neg\text{A}) = \sum_{B\subseteq A} M(B) + \sum_{C\subseteq \neg A} M(C)$

$$\leq \sum_{E\subseteq D} M(E) = 1$$

由于 $Pl(A) = 1- \text{Bel}(\neg A)$，所以

$Pl(A) - \text{Bel}(A) = 1- \text{Bel}(\neg A)- \text{Bel}(A)$

$= 1- (\text{Bel}(\neg A)+ \text{Bel}(A)) \geq 0$

$\therefore Pl(A) \geq \text{Bel}(A)$

Bel(A)：对 A 为真的信任程度
Pl(A)：对 A 为非假的信任程度
A (Bel(A) , Pl(A))：对 A 信任程度的下限与上限。

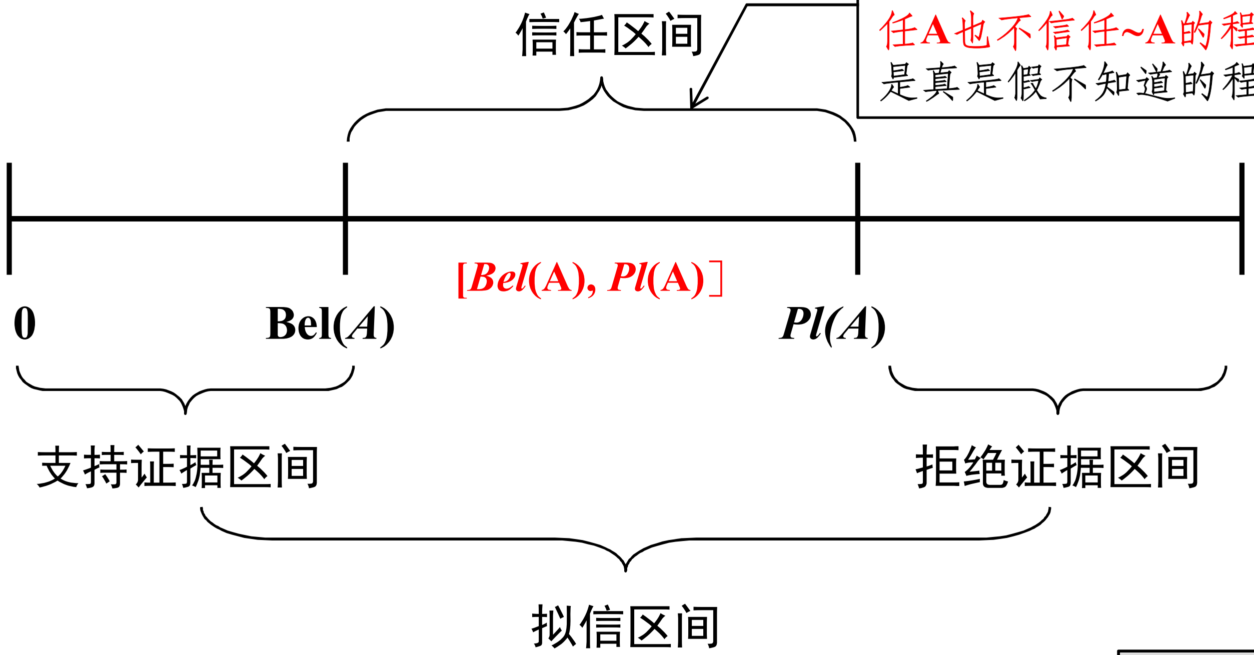
$0 \leq Bel(\text{A}) \leq Pl(\text{A}) \leq 1$ （Bel 是 Pl 的一部分）

说明

- Bel (A)表示对命题A为真的信任程度；
- Bel (~A)表示对命题~ A为真的信任程度，即表示A为假的信任程度；
- Pl (A)＝ 1 – Bel (~ A)表示对A为非假的信任程度。
- 可以看到：**A不为假并不代表A一定为真**，也就是说对A不为假的信任程度应该大于对A为真的信任程度。

证据区间和不确定性

Pl(A) - *Bel*(A)表示既不为假、又不为真的信任程度或者说既不信任A也不信任~A的程度，即是对A是真是假不知道的程度。



只要不是真的就不信

信任度 *Bel* 是对假设信任程度的下限估计—悲观估计；

似然度 *Pl* 是对假设信任程度的上限估计—乐观估计。

只要不是假的就信，如Null也信

信任区间讨论

- **由此可知，[*Bel*(A), *Pl*(A)] 存在三个特殊的区间：**
  - [*Bel*(A), *Pl*(A)]=[1,1]，表示信任A为真；
  - [*Bel*(A), *Pl*(A)]=[0,0]，表示信任A为假；
  - [*Bel*(A), *Pl*(A)]=[0,1]，表示对A是真是假一无所知。
  - [*Bel*(A), *Pl*(A)]=[0.25,0.85]，表示对A为真的信任度为0.25，A为假的信任度为0.15。0.85-0.25=0.6表示对A不知道的程度。

概率分配函数的正交和

- 有时对同样的证据会得到两个**不同的概率分配函数**，例如，对样本空间**D={a，b}**，从不同的来源分别得到如下两个概率分配函数：
  - $M_1(\{a\})=0.3$　$M_1(\{b\})=0.6$　$M_1(\{a,b\})=0.1$　$M_1(\Phi)=0$
  - $M_2(\{a\})=0.4$　$M_2(\{b\})=0.4$　$M_2(\{a,b\})=0.2$　$M_2(\Phi)=0$
- 此时需要对它们进行组合，**Dempster**提出了一种组合方法，即对这两个概率分配函数进行**正交和**运算。

概率分配函数的正交和（证据的组合）

设 M_1 和 M_2 是两个概率分配函数，其正交和

$$M = M_1 \oplus M_2 \qquad M(\Phi) = 0$$

其中:

$$\mathrm{M(A) = K^{-1} \times \sum_{x \cap y = A} M_1(x) \times M_2(y)}$$

$$K = 1 - \sum_{x \cap y = \Phi} M_1(x) M_2(y) = \sum_{x \cap y \neq \Phi} M_1(x) M_2(y)$$

如果 $K \neq 0$，则正交和 M 也是一个概率分配函数；

如果 $K = 0$，则不存在正交和 M，即没有可能存在概率函数，称 M_1 与 M_2 矛盾。

概率分配函数的正交和

设 $M_1, M_2, \ldots M_n$ 是 n 个概率分配函数，则其正交和 $M = M_1 \oplus M_2 \oplus \ldots \oplus M_n$ 为

$$M(\Phi) = 0$$

$$M(A) = K^{-1} \sum_{\cap A_I = A} \prod_{1 \le i \le n} M_i(A_i)$$

其中:

$$K = \sum_{\cap A_i \ne \Phi} \prod_{1 \le i \le n} M_i(A_i)$$

例子：概率分配函数的正交和

- **设 D ={黑，白}，且设**

$$M_2(\{黑\},\{白\},\{黑,白\},\Phi) = (0.6,0.3,0.1,0)$$

$$M_1(\{黑\},\{白\},\{黑,白\},\Phi) = (0.3,0.5,0.2,0)$$

则：$K = 1 - \sum_{x \cap y = \Phi} M_1(x) M_2(y)$

$$= 1 - [M_1(\{黑\})M_2(\{白\}) + M_1(\{白\})M_2(\{黑\})]$$

$$= 1 - [0.3 \times 0.3 + 0.5 \times 0.6] \ = 0.61$$

$$M(\{黑\}) = K^{-1} \sum_{x \cap y = \{黑\}} M_1(x) M_2(y)$$

$$= \frac{1}{0.61}[M_1(\{黑\})M_2(\{黑\}) + M_1(\{黑\})M_2(\{黑,白\}) + M_1(\{黑,白\})M(\{黑\})]$$

$$= \frac{1}{0.61}[0.3 \times 0.6 + 0.3 \times 0.1 + 0.2 \times 0.6] = 0.54$$

概率分配函数的正交和

- 同理可得：

$$M(\{白\}) = 0.43$$

$$M(\{黑,白\}) = 0.03$$

- 组合后得到的概率分配函数：

$$M(\{黑\}), \{白\}, \{黑,白\}, \Phi) = (0.54, 0.43, 0.03, 0)$$

基于证据理论的不确定性推理

- **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**D**。
  - （2）由经验给出，或者由随机性规则和事实的信度度量算基本概率分配函数。
  - （3）计算所关心的子集的信任函数值、似然函数值。
  - （4）由信任函数值、似然函数值得出结论。
- 已知两元组（Bel(A), Pl(A)）可以表示证据的不确定性，同理，它也可以表示不确定的规则。
- 信任函数和似然函数都是基于**概率分配函数**定义的，随着概率分配函数的定义不同，会产生不同的应用模型。

例子

- **设有规则：**
  - （1）如果 流鼻涕 则 感冒但非过敏性鼻炎（**0.9**）或过敏性鼻炎但非感冒（**0.1**）。
  - （2）如果 眼发炎 则 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：**
  - （1）小王流鼻涕（**0.9**）。
  - （2）小王发眼炎（**0.4**）。
- **问：小王患的什么病？**

取样本空间：$D=\{h_1,h_2,h_3\}$

h_1 表示“感冒但非过敏性鼻炎”，

h_2 表示“过敏性鼻炎但非感冒”，

h_3 表示“同时得了两种病”。

取下面的基本概率分配函数：

$$M_1(\{h_1\})=0.9\times0.9=0.81$$

$$M_1(\{h_2\})=0.9\times0.1=0.09$$

$$M_1(\{h_1,h_2,h_3\})=1-M_1(\{h_1\})-M_1(\{h_2\})=1-0.81-0.09=0.1$$

$$M_2(\{h_1\})=0.4\times0.8=0.32$$

$$M_2(\{h_2\})=0.4\times0.05=0.02$$

$$M_2(\{h_1,h_2,h_3\})=1-M_2(\{h_1\})-M_2(\{h_2\})=1-0.32-0.02=0.66$$

将两个概率分配函数组合

$$K = 1/\{1-[M_1(\{h_1\})M_2(\{h_2\})+M_1(\{h_2\})M_2(\{h_1\})]\}$$
$$=1/\{1-[0.81\times 0.02+0.09\times 0.32]\}$$
$$=1/\{1-0.045\}=1/0.955=1.05$$

$$M(\{h_1\})=K[M_1(\{h_1\})M_2(\{h_1\})+M_1(\{h_1\})M_2(\{h_1,h_2,h_3\}$$
$$+M_1(\{h_1,h_2,h_3\})M_2(\{h_1\})]$$
$$=1.05\times 0.8258=0.87$$

$$M(\{h_2\})=K[M_1(\{h_2\})M_2(\{h_2\})+M_1(\{h_2\})M_2(\{h_1,h_2,h_3\}$$
$$+M_1(\{h_1,h_2,h_3\})M_2(\{h_2\})]$$
$$=1.05\times 0.0632=0.066$$

$$M(\{h_1,h_2,h_3\})=1-M(\{h_1\})-M(\{h_2\})=1-0.87-0.066=0.064$$

信任函数:

$$Bel(\{h_1\}) = M(\{h_1\}) = 0.87$$

$$Bel(\{h_2\}) = M(\{h_2\}) = 0.066$$

似然函数:

$$Pl(\{h_1\}) = 1 - Bel(\neg\{h_1\}) = 1 - Bel(\{h_2, h_3\})$$
$$= 1 - [M(\{h_2\} + M(\{h_3\}) = 1 - [0.066 + 0] = 0.934$$

$$Pl(\{h_2\}) = 1 - Bel(\neg\{h_2\}) = 1 - Bel(\{h_1, h_3\})$$
$$= 1 - [M(\{h_1\} + M(\{h_3\}) = 1 - [0.87 + 0] = 0.13$$

结论：小王可能是感冒了。

证据理论的优势和局限性

- **优势：**
  - 满足比Bayes概率理论更弱的条件，即不需要知道先验概率，具有直接表达“不确定”和“不知道”的能力。
- **局限性：**
  - 要求证据必须是独立的，而这有时不易满足；证据合成规则没有非常坚固的理论支持，其合理性和有效性还存在较大的争议；计算上存在着潜在的组合爆炸问题。

目录

- 不确定推理
- 概率推理
- 主观贝叶斯方法
- 可信度方法
- 证据理论
- 小结

小结

- **可信度方法:**
  - 证据、结论和知识的不确定性以可信度进行度量。
- **主观Bayes方法:**
  - 证据与结论的不确定性以概率形式度量，知识的不确定性以数值对（**LS,LN**）进行度量。
- **D-S理论:**
  - 证据与结论用集合表示，不确定性度量用信任函数与似然函数表示；知识的不确定性通过一个集合形式的可信度因子表示。

The End