ECE 566

Uncertainty

Three methods

- Certainty Factors
- Bayesian Inference
- Dempster Shafer Theory of Evidence

Certainty Factors

In MYCIN, every rule relevant to the goal is used, unless one of them succeeds with certainty. Many of the inferences have an inexact character.

If:

- The stain of the organism is Gram negative, and
- The morphology of the organism is rod,

Then:

The class of the organism is Enterio Bacteriaceae (0.8)←CF(rule)

CF(conclusion) = CF(premise) x CF(rule)

Ex: Suppose we have the following in Working Memory (WM):

GRAM = (Gramneg 1.0)

MORPH = (Rod 0.8)(Coccus 0.2)

Then

CF(premise) = min (certainty of individual premises in a rule)

= 0.8 (for the above rule's premise)

Thus

Certainty of ORGANISM = Enterobacteriaceae is:

0.8 x 0.8 = 0.64

Certainty factors are assigned by statisctical experience. They take values in the range [-1 1]

To combine the certainties in a hypothesis that can be derived from multiple sources:

- If the evidence regarding a hypothesis supports it between –0.2 to +0.2, it is in conclusive, it is abandoned.
- If we have two sources that support or disconfirm a hypothesis with different certainties X and Y:

$$\text{Combined CF} = \begin{cases} X + Y - XY & X,Y > 0.2 \\ X + Y + XY & X,Y < -0.2 \\ (X + Y)/(1 - \min(|X|, |Y|)) & \text{otherwise} \end{cases}$$

Thus if two pieces of information both confirm or both disconfirm the hypothesis, confidence in the hypothesis goes up (or down). If they conflict, the denominator dampens the effect.

Bayesian Inference

In this method, parameters are combined according to the rules of probability theory.

P(A | K) stands for a belief in A given a body of knowledge K.

Belief measures follow the three axioms of probability:

1. For any event A, $0 <= P(A) <= 1$
2. P(sure proposition) = 1
3. $P(A \vee B) = P(A) + P(B)$
   if A and B are mutually exclusive

Any event A can be written as the union of non-intersecting (mutually exclusive) components:

$P(A) = P(A \wedge B) + P(A \wedge \neg B)$

More generally, if B_i are mutually exclusive and exhaustive (i=1, 2, …, n) then:

$$P(A) = \sum_i P(A \wedge B_i)$$

Ex: What is the probability that if we roll two dice their outcomes will be equal:

$$P(A) = \sum_{i} P(A \wedge B_i)$$

A = Two dice are equal

B_i (i=1,2,...,6) = the event that the first die is *i*

$$P(A) = \sum_{i=1}^{6} P(A \wedge B_i) = \sum_{i=1}^{6} \frac{1}{36}$$

$$= 6 \times \frac{1}{36} = \frac{1}{6}$$

The basic conditional probability expression is :

$$P(A \wedge B) = P(A \mid B)\, P(B) \qquad (*)$$

$P(A \mid B) = P(A) \Rightarrow$ we say A and B are independent

$P(A \mid B \wedge C) = P(A \mid C) \Rightarrow$ we say A and B are conditionally independent given C.

Ex: In our last example, we can write:

$$P(A \wedge B_i) = P(\underbrace{A}_{\text{two dice are equal}} \mid B_i)\, P(B_i)$$

$$P(A \wedge B_i) = P\,(\text{dice 1 and 2 are } i \mid \text{the 1st die is } i\,)\; P(\text{the first die is } i)$$

$$= \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

Ex: What is the probability that the 1st die will be larger than the 2nd die?

X: 1st die's outcome

Y: 2nd die's outcome

A: X > Y

$$P(A) = \sum_{i=1}^{6} P(A \wedge X = i) = \sum_{i=1}^{6} P(X > Y \mid X = i)\, P(X = i)$$

$$= \frac{1}{6}\sum_{i=1}^{6} P(i > Y) = \frac{1}{6}\sum_{i=1}^{6}\sum_{j=1}^{i-1} P(Y = j)$$

$$= \frac{1}{6}\sum_{i=2}^{6} \frac{1}{6}(i-1) = \frac{1}{6}\left(\frac{1}{6} + \frac{2}{6} + \cdots + \frac{5}{6}\right) = \frac{5}{12}$$

Chain Rule:

Repeated application of the basic conditional probability expression in (*) gives:

$$
\begin{aligned}
P(E_1 \wedge E_2 \wedge E_3 \wedge \ldots \wedge E_n) = {} & P(E_n \mid E_{n-1} \wedge E_{n-2} \wedge \ldots \wedge E_1). \\
& P(E_{n-1} \mid E_{n-2} \wedge \ldots \wedge E_1). \\
& \vdots \\
& P(E_2 \mid E_1). \\
& P(E_1)
\end{aligned}
$$

Bayes' Formula

prior probability

$$\underbrace{P(H \mid e)}_{\text{posterior probability}} = \frac{P(e \mid H)\,\overbrace{P(H)}^{\text{prior probability}}}{P(e)}$$

posterior probability

$$P(e) = P(e \mid H)\,P(H) + P(e \mid \neg H)P(\neg H)$$

$$P(H \mid e) + P(\neg H \mid e) = 1$$

Ex: We have a couple of roulette tables, and one dice table. We hear someone yell twelve. How can we estimate the certainty that it came from a dice table or a roulette table.

We don't have

P(dice | 12) and/or P(roulette | 12)

We have

$$P(12 \mid \text{dice}) = \frac{1}{36}$$

$$P(12 \mid \text{roulette}) = \frac{1}{38}$$

$$P(\text{dice}) = \frac{1}{3}$$

$$P(\text{roulette}) = \frac{2}{3}$$

$$P(\text{dice} \mid 12) = \frac{P(12 \mid \text{dice})P(\text{dice})}{P(12)}$$

$$P(12) = P(12 \mid \text{dice})P(\text{dice}) + P(12 \mid \text{roulette})P(\text{roulette})$$

$$P(\text{dice} \mid 12) =$$

Complete this example as an exercise.

$$P(\text{roulette} \mid 12) = ?$$

Q. When can we simply write P(A) or P(A) instead of P(A|K) and leave out K from the problem?

A. When K remains constant, we don't need to explicate it. But whenever background information may undergo change, such that it affects our knowledge about the problem's outcome, we need to identify the assumptions that account for our belief.