

Лекция 12

*Модели представления и обработки
неопределенных знаний.*

Коэффициенты уверенности Шортлиффа

Проблема представления неопределенных знаний

Экспертным знаниям, как правило, *присуща неопределенность*.

В инженерии знаний принято выделять различные *типы неопределенности* знаний (*НЕ-факторы*):

- неполнота;
 - неточность;
 - нечеткость;
- С другой стороны, источники неопределенности могут быть разделены на две большие категории:
 - недостаточно полное знание *о предметной области*;
 - недостаточная информация *о конкретной ситуации*;
 - Для представления и обработки неопределенных знаний в ЭС предложены различные формальные модели, в частности:
 - коэффициенты уверенности Шортлиффа;
 - теория свидетельств Демпстера-Шефера;
 - Байесовские сети доверия;
 - нечеткая логика и теория возможностей
 -

Неопределенность в продукционных БЗ

- Правила в БЗ в общем случае могут иметь сколь угодно сложную логическую структуру антецедента. Каждое *свидетельство в антецеденте может характеризоваться своей степенью неопределенности* (оценками уверенности);
- Кроме того, сами *правила* (отношение следования) также *могут быть взвешены оценками уверенности*;
- Правила в БЗ могут содержать в консеквенте одни и те же выводы;
- Т. о. *обобщенная схема обработки неопределенностей* в продукционных базах знаний имеет следующий вид:

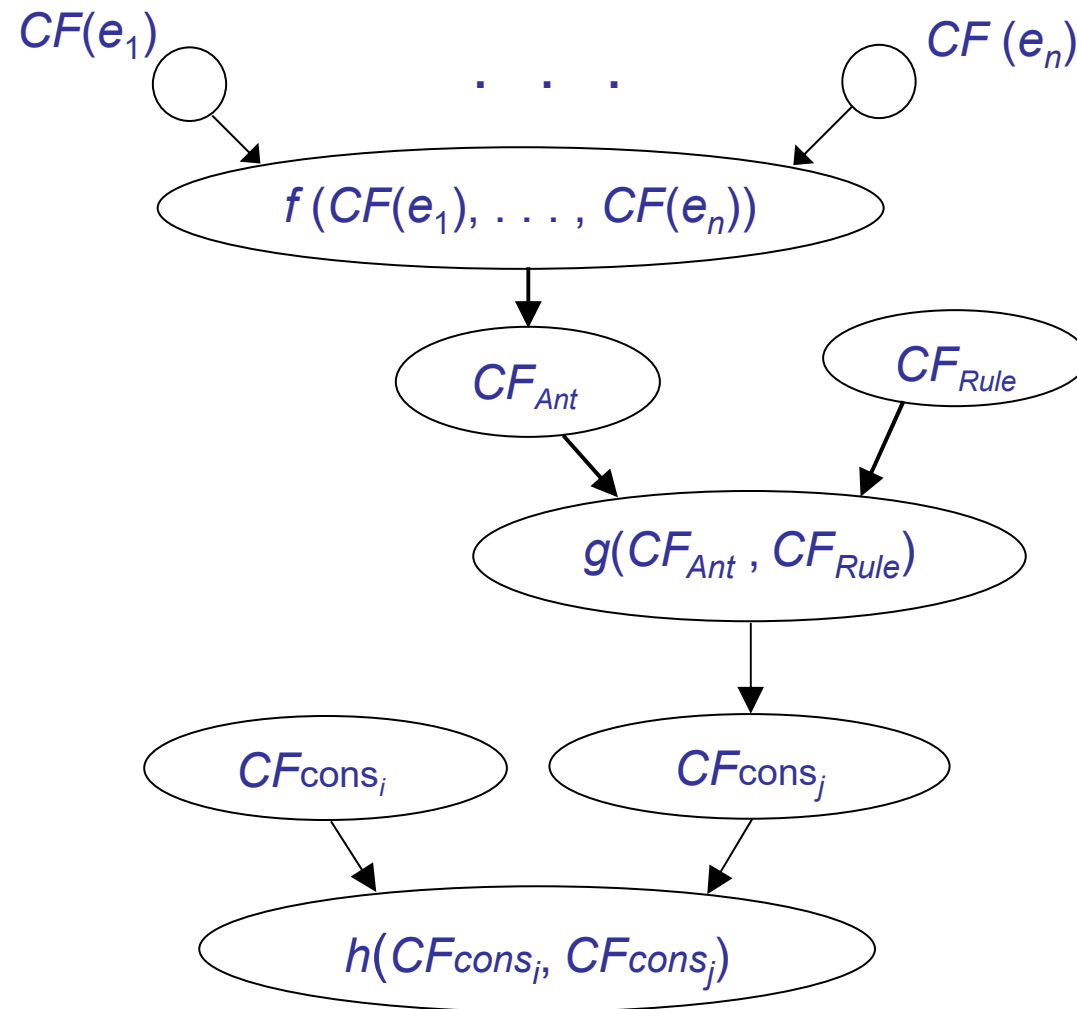
$$\underset{0,5}{A} \& \underset{0,6}{B} \xrightarrow{0,4} C$$

$$\underset{0,8}{D} \& \underset{0,7}{E} \xrightarrow{0,6} C$$

- Фактор определенности свидетельств, связанных конъюнкцией $A \& B$:
$$CF_{Ant} = \min (CF [A], CF [B])$$
- Фактор определенности свидетельств, связанных дизъюнкцией $A \vee B$:
$$CF_{Ant} = \max (CF [A], CF [B])$$
- Фактор определенности консеквента:

$$CF_{cons} = CF_{Ant} \cdot CF_{Rule}$$

Обобщенная схема обработки неопределенных знаний в продукционных системах



Коэффициенты уверенности Шортлиффа

- Один из первых подходов к формализации и обработке неопределенных знаний был предложен **Шортлиффом** при разработке системы **MYCIN**
- Неопределенность представляется *коэффициентами уверенности* (КУ), принимающими значение в интервале $[-1; +1]$.
 - 1 соответствует абсолютной ложности утверждения,
 - +1 – абсолютной истинности,
 - 0 – полная неопределенность.

Вычисление коэффициента уверенности

- Рассмотрим правило вида:

Если *<Свидетельство- X >*, то *<Вывод- A >*.

- Пусть $P(A)$ – априорная вероятность гипотезы A ;
 $P(A/X)$ – апостериорная вероятность гипотезы A при наличии свидетельства X .
- Значение CF коэффициента уверенности в выводе A при наличии свидетельства X определяется следующим образом:

$$CF[A, X] = \begin{cases} (P(A/X) - P(A)) / (1 - P(A)), & \text{если } P(A/X) \geq P(A); \\ P(A/X) - P(A) / P(A), & \text{если } P(A/X) < P(A). \end{cases}$$

- Первая формула соответствует случаю, когда X подтверждает A , а вторая – когда X опровергает A .
- Антеcedенты правил обычно имеют сложную логическую структуру и представляют собой формулы, составленные из отдельных свидетельств с использованием операций конъюнкции и дизъюнкции. В этом случае при вычислении КУ составного антеcedента используются следующие формулы:
- для $X \& Y$: $CF_{and} = \min (CF[X], CF[Y])$;
- для $X \vee Y$: $CF_{or} = \max (CF[X], CF[Y])$.

Вычисление коэффициента уверенности

- Помимо неопределенности *исходных свидетельств* может иметь место неполная уверенность эксперта в *справедливости* некоторого *правила вывода*.
- В этом случае КУ приписываются самому правилу (импликативной связке). Тогда коэффициент уверенности CF_{Cons} для заключения, полученного по некоторому правилу, определяется выражением:

$$CF_{Cons} = CF_{Ant} * CF_{Rule} ,$$

- где – CF_{Ant} и CF_{Rule} соответственно КУ антецедента и собственно правила;
* – операция умножения.
- Различные правила в БЗ могут одновременно подтверждать (опровергать) одну и ту же гипотезу Н. В этом случае каждое правило рассматривается как независимое свидетельство.

Вычисление коэффициента уверенности

- Обозначим $CF1 = CF[H:E1]$ – коэффициент уверенности в гипотезе H при наличии свидетельства E1,
- аналогично $CF2 = CF[H:E2]$ для свидетельства E2,
- $CF1,2 = CF[H:E1,E2]$ – коэффициент уверенности в H при наличии свидетельств E1 и E2. Тогда имеют место следующие правила комбинирования свидетельств:

$$CF1,2 = \begin{cases} CF1 + CF2 - CF1 * CF2, & \text{если } CF1 > 0 \text{ и } CF2 > 0; \\ CF1 + CF2 + CF1 * CF2, & \text{если } CF1 < 0 \text{ и } CF2 < 0; \\ (CF1 + CF2) / (1 - \min(|CF1|, |CF2|)), & \text{если } CF1 \text{ и } CF2 \text{ имеют разные знаки} \end{cases}$$

- Данные формулы обладают следующими важными свойствами:
- 1. Симметричностью свидетельств, т.е. независимостью от порядка их получения.
- 2. По мере накопления подтверждающих (опровергающих) свидетельств значение КУ смещается к определенности (-1 или +1).

Вычисление коэффициентов уверенности. Пример

- Пусть ЭС включает два следующих правила (с одинаковым заключением):

1. Если (X не смотрит CNN) [CF = 0,8]
И (X не верит версии о российских хакерах), [CF = 0,75]
То (X будет голосовать за Трампа); [CF = 0,9]
2. Если (X не нравится Х.Клинтон) [CF = 0,4]
ИЛИ (X за ограничение миграции), [CF = 0,6]
То (X будет голосовать за Трампа) [CF = 0,7]

В правиле 1 (свидетельства связаны конъюнкцией) значение коэффициента уверенности антецедента в целом:

$$CF1_{Ant} = \min(0,8; 0,75) = 0,75$$

Значение КУ вывода по этому правилу есть:

$$CF1 = CF1_{Ant} * CF1_{Rule} = 0,75 * 0,9 = \mathbf{0,675}$$

- Для правила 2 (свидетельства связаны дизъюнкцией):

$$CF2_{Ant} = \max(0,4; 0,6) = 0,6,$$

$$CF2 = CF2_{Ant} * CF2_{Rule} = 0,6 * 0,7 = \mathbf{0,42}$$

Т.к. CF1 и CF2 имеют одинаковые знаки по правилу комбинирования свидетельств имеем:

$$CF1,2 = 0,675 + 0,42 - 0,675 * 0,42 = 1,095 - 0,2835 \approx \mathbf{0,81}$$