Uncertainty management in rule-based expert systems

- Introduction
- Certainty Factor
- Probability Theory

(C) Copyright 2022 by Dr.Abeer El-Korany

What is uncertainty?

- Information can be incomplete, inconsistent, uncertain, or all three. In other words, information is often unsuitable for solving a problem.
- **Uncertainty** is defined as the lack of the exact knowledge that would enable us to reach a perfectly reliable conclusion.
- Classical logic permits only exact reasoning. It assumes that perfect knowledge always exists and the ***law of the excluded middle***

(C) Copyright 2022 by Dr.Abeer El-Korany

Introduction

Reasoning under uncertainty and with inexact knowledge

- frequently necessary for real-world problems
- heuristics
  - ways to mimic heuristic knowledge processing
  - methods used by experts
- empirical associations
  - experiential reasoning
  - based on limited observations
- probabilities
  - objective (frequency counting)
  - subjective (human experience)
- reproducibility
  - will observations deliver the same results when repeated

(C) Copyright 2022 by Dr.Abeer El-Korany

example

- Because patients who are diagnosed with the same sickness may experience different level of comforts or pain, it is good practice to incorporate an ***uncertainty*** degree
  - If an old man is a smoker, then he is a high risk patient.
  - If a middle man is a non-smoker, then he is a low risk patient.
  - If a young woman is a non-smoker, then she is a low risk patient.

(C) Copyright 2022 by Dr.Abeer El-Korany

Quantification of ambiguous and imprecise terms on a time-frequency scale

Ray Simpson (1944)		*Milton Hakel* (1968)	
Term	*Mean value*	*Term*	*Mean value*
Always	99	Always	100
Very often	88	Very often	87
Usually	85	Usually	79
Often	78	Often	74
Generally	78	Rather often	74
Frequently	73	Frequently	72
Rather often	65	Generally	72
About as often as not	50	About as often as not	50
Now and then	20	Now and then	34
Sometimes	20	Sometimes	29
Occasionally	20	Occasionally	28
Once in a while	15	Once in a while	22
Not often	13	Not often	16
Usually not	10	Usually not	16
Seldom	10	Seldom	9
Hardly ever	7	Hardly ever	8
Very seldom	6	Very seldom	7
Rarely	5	Rarely	5
Almost never	3	Almost never	2
Never	0	Never	0

(C) Copyright 2022 by Dr.Abeer El-Korany

Sources of uncertain knowledge

- **Weak implications .** Domain experts and knowledge engineers have the painful task of establishing concrete correlations between IF (condition) and THEN (action) parts of the rules.

- Expert systems need to have the ability to handle vague associations, for example by accepting the degree of correlations as numerical certainty factors.

(C) Copyright 2022 by Dr.Abeer El-Korany

- **Imprecise language.** Our natural language is ambiguous and imprecise. We describe facts with such terms as *often* and *sometimes*, *frequently* and *hardly ever*.

- It can be difficult to express knowledge in the precise IF-THEN form of production rules.

- However, if the meaning of the facts is quantified, it can be used in expert systems.

(C) Copyright 2022 by Dr.Abeer El-Korany

- **Unknown data.**
  - When the data is incomplete or missing, the only solution is to accept the value "unknown" and proceed to an approximate reasoning with this value.

- **Combining the views of different experts.**
  - Large expert systems usually combine the knowledge and expertise of a number of experts.
  - Unfortunately, experts often have contradictory opinions and produce conflicting rules.
  - To resolve the conflict, the knowledge engineer has to attach a weight to each expert and then calculate the composite conclusion.

(C) Copyright 2022 by Dr.Abeer El-Korany

Forms of Inexact Knowledge

- **uncertainty**(truth not clear)
  - probabilistic models, multi-valued logic (true, false, don't know,...), certainty factor theory
- **incomplete knowledge** (lack of knowledge)
  - P true or false [not known (??? defaults)]
- **defaults, beliefs** (assumptions about truth)
  - assume P is true, as long as there is no counter-evidence
  - assume P is true with Certainty Factor
- **contradictory knowledge** (true and false)
  - inconsistent fact base; somehow P and ¬P true

(C) Copyright 2022 by Dr.Abeer El-Korany

Inexact Knowledge - Example

Person A walks on Campus towards the bus stop. A few hundred yards away A sees someone and is quite sure that it's his next-door neighbor B who usually goes by car to the University. A screams B's name.

Q: Which forms of inexact knowledge and reasoning are involved here?

default - A wants to take a bus

belief, (un)certainty - it's the neighbor B

probability, default, uncertainty - the neighbor goes home by car

default - A wants to get a lift

default - A wants to go home

(C) Copyright 2022 by Dr.Abeer El-Korany

Examples of Inexact Knowledge

Person A walks on Campus towards the bus stop. A few hundred yards away A sees someone and is quite sure that it's his next-door neighbor B who usually goes by car to the University. A screams B's name.

Fuzzy - *a few hundred yards*

define a mapping from "#hundreds" to 'few', 'many', ...

not uncertain or incomplete but graded, vague

Probabilistic - *the neighbor usually goes by car*

probability based on measure of how often he takes car;

calculates always $p(F) = 1 - p(\neg F)$

Belief - *it's his next-door neighbor B*

"reasoned assumption", assumed to be true

Default - *A wants to take a bus*

assumption based on commonsense knowledge

(C) Copyright 2022 by Dr.Abeer El-Korany

Uncertainty in Individual Rules

- likelihood of evidence
  – for each premise
  – for the conclusion
  – combination of evidence from multiple premises

(C) Copyright 2022 by Dr.Abeer El-Korany

Certainty factors theory and evidential reasoning

- A **certainty factor** (*cf*), a number to measure the expert's belief.
- The maximum value of the certainty factor is, say, +1.0 (definitely true) and the minimum −1.0 (definitely false).
- For example, if the expert states that some evidence is almost certainly true, a *cf* value of 0.8 would be assigned to this evidence.

(C) Copyright 2022 by Dr.Abeer El-Korany

Uncertain terms and their interpretation in MYCIN

Term	*Certainty Factor*
Definitely not	−1.0
Almost certainly not	−0.8
Probably not	−0.6
Maybe not	−0.4
Unknown	−0.2 to +0.2
Maybe	+0.4
Probably	+0.6
Almost certainly	+0.8
Definitely	+1.0

(C) Copyright 2022 by Dr.Abeer El-Korany

Certainty Factor Theory

- Certainty Factor CF of Hypothesis H
  - ranges between -1 (denial of H) and +1 (confirmation of H)
  - allows the ranking of hypotheses
- Based on measures of belief MB and measures of disbelief MD
- MB is expressing the belief that H is true
- MD is expressing the belief that H is not true
- MB is not 1-MD - it's not like probabilities
- Experts determine values for MB, MD of H based on given evidence E $\rightarrow$ subjective

(C) Copyright 2022 by Dr.Abeer El-Korany

Certainty Factor

- certainty factor CF
  - ranges between -1 (denial of the hypothesis H) and +1 (confirmation of H)
    - A value of near +1 represents strong confidence
    - A value near zero represents a lack of evidence for or against the hypothesis
    - A value of near –1 represents a lack of confidence
  - allows the ranking of hypotheses
  - "A hypothesis is a logical supposition, a reasonable guess, an educated conjecture. It provides a tentative explanation for a phenomenon under investigation." (Leedy and Ormrod, 2001).
- Difference between belief and disbelief

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

(C) Copyright 2022 by Dr.Abeer El-Korany

Certainty Factor In RBS

- In expert systems with certainty factors, the knowledge base consists of a set of rules that have the following syntax:

  IF <evidence>
  THEN <hypothesis> {*cf* }

  where *cf* represents belief in hypothesis *H* given that evidence *E* has occurred.

Certainty Factor In RBS(cont.)

- **The certainty factor assigned by a rule is propagated through the reasoning chain.**
- **This involves establishing the net certainty of the rule consequent when the Evidence (E) in the rule antecedent is uncertain:**

$$cf\,(H,E) = cf\,(E) \times cf$$

For example,

IF sky is clear
THEN the forecast is sunny {*cf* 0.8}

and the current certainty factor of *sky is clear* is 0.5**, then the overall confidence of the whole rule is calculated as follow:**

$$cf\,(H,E) = 0.5 \cdot 0.8 = 0.4$$

This result value can be interpreted as ***"It may be sunny"*.**

(C) Copyright 2022 by Dr.Abeer El-Korany

Manipulating Confidence Factors

- An expert assigns a confidence for each rule in the rule base
  - These values are used to calculate new values as the system progresses towards a solution
  - The content of the rule itself is much more significant than the CF assigned
- Combining CF for premises using AND and OR

CF(P1 and P2) = MIN(CF(P1), CF(P2)), and
CF(P1 or P2) = MAX(CF(P1), CF(P2)).

(C) Copyright 2022 by Dr.Abeer El-Korany

Certainty Factor In RBS(cont.)

The CF values of the premise is computed together

Given the following rule R1:

If A and B then C *(CF = 0.6)*

And suppose the following evidence :

CF(A) *(0.5)*

CF(B) *(0,7)*

CF(A and B) = min(CF(A),CF(B)) = *(0.5)*

CF of the R1 = CF * CF(A and B) = 0.6 * 0.5 = *(0.3)*

(C) Copyright 2022 by Dr.Abeer El-Korany

Certainty Factor In RBS(cont.)

- For conjunctive rules such as

IF <evidence E_1>
⋮
AND <evidence E_n>
THEN <hypothesis H> {cf}

the certainty of hypothesis *H*, is established as follows:

$$cf\,(H, E_1 \cap E_2 \cap \ldots \cap E_n) = \min\,[cf\,(E_1), cf\,(E_2), \ldots, cf\,(E_n)] \times cf$$

For example,

IF sky is clear
AND the forecast is sunny
THEN the action is 'wear sunglasses' {*cf* 0.8}

and the certainty of *sky is clear* is 0.9 and the certainty of the *forecast of sunny* is 0.7, then

$$cf\,(H, E_1 \cap E_2) = \min\,[0.9, 0.7] \cdot 0.8 = 0.7 \cdot 0.8 = 0.56$$

(C) Copyright 2022 by Dr.Abeer El-Korany

Certainty Factor In RBS(cont.)

- **For disjunctive rules such as**

IF <evidence E_1>

⋮

OR <evidence E_n>

THEN <hypothesis H> $\{cf\}$

the certainty of hypothesis *H*, is established as follows:

$$cf\,(H, E_1 \cup E_2 \cup \ldots \cup E_n) = max\,[cf\,(E_1), cf\,(E_2), \ldots, cf\,(E_n)] \times cf$$

For example,

IF sky is overcast

OR the forecast is rain

THEN the action is 'take an umbrella' {*cf* 0.9}

and the certainty of *sky is overcast* is 0.6 and the certainty of the *forecast of rain* is 0.8, then

$$cf\,(H, E_1 \cup E_2) = max\,[0.6, 0.8] \cdot 0.9 = 0.8 \cdot 0.9 = 0.72$$

(C) Copyright 2022 by Dr.Abeer El-Korany

Calculating CF of the same rule

- certainty factors that support the same conclusion
- combining antecedent evidence
  - use of premises with less than absolute confidence
    - $E_1 \wedge E_2 = \min(CF(H, E_1), CF(H, E_2))$
    - $E_1 \vee E_2 = \max(CF(H, E_1), CF(H, E_2))$
    - CF(not E1) = - CF(E1)

(C) Copyright 2022 by Dr.Abeer El-Korany

Combining Certianty factor in Rule Chain Example1

The CF-values are chained together with the rule firing

Consider the following two rules:

R1- IF A THEN B CF1 = 0.7

R2- IF B THEN C CF2 = 0.3

And input is A with *CF=0.5*

CF(A)=0.5 , Cf (R1) = (0.7) => CF(B)=0.35

CF(B)=0.35, Cf (R2) = (0.3) => CF(C) = 0.105

(C) Copyright 2022 by Dr.Abeer El-Korany

Combining Certianty factor in Rule Chain Example2: <u>Confirming the same hypothesis</u>

- When the same consequent is obtained as a result of the execution of two or more rules, the individual certainty factors of these rules must be merged to give a combined certainty factor for a hypothesis.

  Suppose the knowledge base consists of the following rules:

  Rule 1: IF *A* is *X*
  THEN *C* is *Z* {*cf* 0.8}

  Rule 2: IF *B* is *Y*
  THEN *C* is *Z* {*cf* 0.6}

  What certainty factor should be assigned to object *C* having value *Z* if both *Rule* 1 and *Rule* 2 are fired?

(C) Copyright 2022 by Dr.Abeer El-Korany

Confirming the same hypothesis(cont.)

- Common sense suggests that, if we have two pieces of evidence (*A* is *X* and *B* is *Y*) from different sources (*Rule 1* and *Rule 2*) supporting the same Hypothesis (*C* is *Z*), then the confidence in this hypothesis should increase and become stronger than if only one piece of evidence had been obtained.

(C) Copyright 2022 by Dr.Abeer El-Korany

An Example Calculation

- Suppose a rule in the knowledge base is:

  (P1 and P2) $\rightarrow$ R1 (.7) P3 R2 (.3)

- Suppose P1, P2, and P3 have CFs of 0.6, 0.4, 0.4, respectively

CF(P1(0.6) and P2(0.4)) = MIN(0.6,0.4) = 0.4.

The CF for R1 is 0.7 in the rule, so R1 is added to the set of case specific knowledge with the associated CF of $(0.7) \times (0.4) = 0.28$.

The CF for R2 is 0.3 in the rule, so R2 is added to the set of case specific knowledge with the associated CF of $0(.3) \times (0.4) = 0.12$.

(C) Copyright 2022 by Dr.Abeer El-Korany

Confirming the same hypothesis

- To calculate a combined certainty factor we can use the following equation:

$$cf(cf_1, cf_2) = \begin{cases} cf_1 + cf_2 \times (1 - cf_1) & \text{if } cf_1 > 0 \text{ and } cf_2 > 0 \\ \dfrac{cf_1 + cf_2}{1 - min\left[|cf_1|, |cf_1|\right]} & \text{If } cf_1 \text{ and } cf_2 \text{ are of different signs} \\ cf_1 + cf_2 \times (1 + cf_1) & \text{if } cf_1 < 0 \text{ and } cf_2 < 0 \end{cases}$$

where

cf_1 is the confidence in hypothesis *H* established by *Rule* 1;

cf_2 is the confidence in hypothesis *H* established by *Rule* 2.

(C) Copyright 2022 by Dr.Abeer El-Korany

- The certainty factors theory provides a *practical* alternative to Bayesian reasoning. The heuristic manner of combining certainty factors is different from the manner in which they would be combined if they were probabilities. The certainty theory is not “mathematically pure” but does mimic the thinking process of a human expert.

(C) Copyright 2022 by Dr.Abeer El-Korany

Combining Rules

- Suppose rules R1 and R2 support the result H
- Given the CF for R1 and R2, we calculate the CF for H as follows

$CF(R1) + CF(R2) - (CF(R1) \times CF(R2))$ when $CF(R1)$ and $CF(R2)$ are positive,

$CF(R1) + CF(R2) + (CF(R1) \times CF(R2))$ when $CF(R1)$ and $CF(R2)$ are negative,

and

$$\frac{CF(R1) + CF(R2)}{1 - MIN(\,|CF(R1)|\,,\,|CF(R2)|\,)}$$ otherwise,

- expert system shell will use certainty factors to search for the most likely solution

(C) Copyright 2022 by Dr.Abeer El-Korany

Stanford Certainty Factor Theory

- Remember the base rule for Certainty Factor **CF (H|E)** :

  CF (H|E) = MB(H|E) – MD(H|E) -1 ≤ CF(H) ≤ 1

- Integrate Certainty Factors into reasoning.
- <u>CF-value for H calculated using CFs of premises P in rule</u>

  CF(H) = CF(P1 <u>and</u> P2) = <u>min</u> (CF(P1),CF(P2))

  CF(H) = CF(P1 <u>or</u> P2) = <u>max</u> (CF(P1),CF(P2))

- <u>CF-value for H combined from different rules, experts, ...</u>

  CF(H) = CF1 + CF2 – CF1· CF2 if both **CF1,CF2** > 0

  CF(H) = CF1 + CF2 + CF1· CF2 if both **CF1,CF2** < 0

  $$CF(H) = \frac{CF1 + CF2}{1 - \min(|CF1|,|CF2|)} \quad \text{else}$$

(C) Copyright 2022 by Dr.Abeer El-Korany

Summary Reasoning and Uncertainty

- many practical tasks require reasoning under uncertainty
  - missing, inexact, inconsistent knowledge
- variations of probability theory are often combined with rule-based approaches
  - works reasonably well for many practical problems

(C) Copyright 2022 by Dr.Abeer El-Korany

Advantages and Problems of Certainty Factors

- Advantages
  - simple implementation
  - reasonable modeling of human experts' belief
    - expression of belief and disbelief
  - successful applications for certain problem classes
  - evidence relatively easy to gather
    - no statistical base required
- Problems
  - partially ad hoc approach
    - theoretical foundation through Dempster-Shafer theory was developed later
  - combination of non-independent evidence unsatisfactory
  - new knowledge may require changes in the certainty factors of existing knowledge
  - certainty factors can become the opposite of conditional probabilities for certain cases
  - not suitable for long inference chains

(C) Copyright 2022 by Dr.Abeer El-Korany