REASONING UNDER UNCERTAINTY: CERTAINTY THEORY

Table of Content

- Introduction
- Certainty Theory
- Definition
- Certainty Theory: Values Interpretation
- Certainty Theory: Representation
- Certainty Factor Propagation
- Advantages and Disadvantages

Introduction

- Very important topic since many expert system applications involve uncertain information
- Uncertainty can be defined as the lack of adequate information to make decision.
- Can be a problem because it may prevent us from making the best decision or may even cause bad decision to be made

Introduction

- A number of theories have been devised to deal with uncertainty such as:
  - **Certainty theory**
  - **Bayesian probability**
  - **Zadeh fuzzy theory**
  - Dempster-Shafer theory
  - Hartley theory based on classical sets
- Dealing with uncertainty requires reasoning under uncertainty

Introduction

- Very important topic since many expert system applications involve uncertain information
- Uncertainty can be defined as the lack of adequate information to make decision.
- Can be a problem because it may prevent us from making the best decision or may even cause bad decision to be made

Introduction

- The type of uncertainty may be caused by problems with data such as:
  - Data might be missing or unavailable
  - Data might be present but unreliable or ambiguous
  - The representation of the data may be imprecise or inconsistent
  - Data may just be user's best guest
  - Data may be based on defaults and defaults may have exceptions

Introduction

- Alternatively, the uncertainty may be caused by the represented knowledge, since it might:
  - Represent best guesses of the experts that are based on plausible of statistical associations they have observed
  - Not be appropriate in all situations

Introduction

- Reasoning under uncertainty involves 3 important issues:
  - How to represent uncertain data
  - How to combine two or more pieces of uncertain data
  - How to draw inference using uncertain data

Certainty Theory

- Grew out of MYCIN
- Relies on defining judgmental measures of belief rather than adhering to strict probabilities estimates.
- Expert make judgment when solving problems
- Problem information maybe incomplete. Yet experts learn to adapt to the situation and continue to reason about the problem intelligently

Certainty Theory

- Principle features of MYCIN is managing inexact information
- Inexact information has significant in medical domain because of time constraints such as emergency room
- Majority of the rules used in medicine is inexact

Certainty Theory

- Some uncertain phrases:
  - probably
  - it is likely that
  - it almost seems certain that

Certainty Theory

- Uncertain evidence is given CF or certainty factor value ranging from -1 to 1.
- Negative values degree of disbelief
- Positive values degree of belief
- \- Range of CF values

false	Possibly False	Unknown	Possible True	True
-1		0		1
Measures of disbelief				Measures of bilief

Certainty Theory: Definitions

1. **Measures of Belief (MB)**

   Number that reflects the measure of increased belief in a hypothesis H based on evidence E

   $$0 \leq MB \leq 1$$

2. **Measures of Disbelief (MD)**

   Number that reflects the measure of increase disbelief in a hypothesis H based on evidence E

   $$0 \leq MD \leq 1$$

Certainty Theory: Definitions

3. **Certainty Factor**

   Number that reflects the net level of belief in a hypothesis given available information

$$CF = MB - MD$$

$$-1 \leq CF \leq 1$$

Certainty Theory: Values Interpretation

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

Certainty Theory: Representation

UNCERTAIN EVIDENCE

- Representing uncertain evidence

  CF (E) = CF (It will probably rain today)

  = 0.6
- Similar to P(E), however, CF are not probabilities but measures of confidence
- Degree to which we belief evidence is true
- In ES written as exact term but add CF value to it

  Example: It will rain today **CF 0.6**

Certainty Theory : Representation

- Basic structure of rule in certainty model
  IF E THEN H CF (Rule)
- CF (Rule) = level of belief of H given E
- Given that E is true, we believe H according to:
  CF (H, E) = CF (Rule)

Certainty Theory : Representation

- Example:

IF There are dark cloud - E

THEN It will rain - H **CF = 0.8**

- This rule reads:

  "If there are dark clouds then it will almost certainly rain"

- CF (H, E) similar P (H|E)

Certainty Factor Propagation (C.F.P)

CFP FOR SINGLE PREMISE RULE

Concerned with establishing the level of belief in rule conclusion (H) when the available evidence (E) contained in the rule premise is uncertain.

CF (H, E) = CF (E) X CF (RULE)

Example

From previous rule:

CF (E) = 0.5

C.F.P. is

CF (H, E) = 0.5 x 0.8

= 0.4 #

In words: It maybe raining

Certainty Factor Propagation (C.F.P)

For same rule, with negative evidence for rule premise to the degree of CF (E) = -0.5

C.F.P is

CF (H, E) = -0.5 * 0.8

= -0.4 #

In words: It maybe not raining.

Certainty Factor Propagation (C.F.P)

CFP FOR MULTIPLE RULE PREMISE

- For rules with more than one premises:
  - conjunctive rules (AND)
  - disjunctive rules (OR)

Certainty Factor Propagation (C.F.P)

CFP FOR MULTIPLE RULE PREMISE

- **Conjunctive Rule**

  IF EI AND E2 AND ... THEN H **CF** (RULE)

  CF (H, E1 AND E2 AND..) = min {CF (Ei)} * CF (RULE)

- min function returns minimum value of a set of numbers.

Certainty Factor Propagation (C.F.P)

Example:

IF The sky is dark

AND The wind is getting stronger

THEN It will rain CF = 0.8

Assume:

CF (the sky is dark) = 1.0

CF (the wind is getting stronger) = 0.7

CF (It will rain) = min {1.0,0.7} * 0.8

= 0.56 #

In words: It <u>probably</u> will rain

Certainty Factor Propagation (C.F.P)

CFP FOR MULTIPLE RULE PREMISE

- **Disjunctive Rule (OR)**

  IF E1 OR E2 OR ... THEN H **CF** (RULE)

  CF (H, E1 OR E2 OR..) = Max {CF (Ei)} * CF (RULE)

- Max function returns the maximum value of a set of numbers

Certainty Factor Propagation (C.F.P)

- Example:

  IF the sky is dark

  OR the wind is getting stronger

  THEN It will rain **CF = 0.9**

  CF (It will rain) = max {1.0, 0.7} * 0.9

  = 0.9 #

- **In words**: It almost certainly will rain

Certainty Factor Propagation (C.F.P)

CFP SIMILARLY CONCLUDED RULES

- For multiple rules that support a hypothesis (same hypothesis)

- Consider from 2 individuals:
  - weatherman
  - farmer

Certainty Factor Propagation (C.F.P)

Rule 1

IF the weatherman says it is going to rain (**E1**)

THEN It is going to rain (**H**)

CF (Rule 1) = 0.8

Rule 2

IF the farmer says it is going to rain (**E2**)

THEN It is going to rain (**H**)

CF (Rule 2) = 0.8

Certainty Factor Propagation (C.F.P)

- CF of both rules set to equal implying equal confidence in the 2 sources
- Naturally more confident in conclusion
- MYCIN team developed incrementally acquired evidence to combined belief and disbelief values by rules concluding the same hypothesis

Certainty Factor Propagation (C.F.P)

FORMULA FOR INCREMENTALLY ACQUIRED EVIDENCE

$CF_{combined}$ (CF1, CF2)

= CF1 + CF2 * (1 - CF1) **Both > 0**

$$= \frac{CF1 + CF2}{1 - \min\{|CF1|,|CF2|\}}$$ **One < 0**

= CF1 + CF2 * (1 + CF1) **Both < 0**

where,

CF1 = confidence in H established by one rule (RULE 1)

CF2 = confidence in it established by one rule (RULE 2)

CF1 = CF1 (H, E)

CF2 = CF2 (H, E)

Certainty Factor Propagation (C.F.P)

- There are 2 properties of the equations:
  - **Commutative:**
    - allow evidence to be gathered in any order
  - **Asymptotic:**
    - If more than one source confirm a hypothesis then a person will feels more confident. Incrementally add belief to a hypothesis as new positive evidence is obtained
    - To prevent the certainty value of the hypothesis exceeding 1

Certainty Factor Propagation (C.F.P)

Example

Consider Rain Prediction: Rule 1 and 2

Explore several cases

Case 1: Weatherman and Farmer Certain in Rain

CF (E1) = CF (E2) = 1.0

CF1 (H, E1) = CF (E1) * CF (RULE 1)

= 1.0 * 0.8 = 0.8

CF2 (H, E2) =CF (E2) * CF (RULE 2)

= 1. 0 * 0.8 = 0.8

C.F.P. for single Premise Rule

Certainty Factor Propagation (C.F.P)

Since both > 0,

$CF_{combine}$ (CF1, CF2) = CF1 + CF2 * (1 - CF1)

= 0.8 + 0.8 * (1 - 0.8)

= 0.96 #

CF supported by > 1 rule can be <u>incrementally increase</u> more confident

Certainty Factor Propagation (C.F.P)

<u>**Case 2**: Weatherman certain in rain, Farmer certain no rain</u>

CF (E1) = 1.0 CF (E2) = -1.0

CF1 (H, E1) = CF (E1) * CF (RULE 1)
= 1.0 * 0.8 = 0.8

CF2 (H, E2) = CF (E2) * CF (RULE 2)
= -1.0 * 0.8 = -0.8

Since either one < 0

$$CF_{combined}(CF1, CF2) = \frac{CF1 + CF2}{1 - \min\{|CF1|,|CF2|\}}$$

$$= \frac{0.8 + (-0.8)}{1 - \min\{0.8, 0.8\}}$$

= 0 #

CF set to unknown because one say “no” and the other one say “yes”

Certainty Factor Propagation (C.F.P)

Case 3: Weatherman and Farmer believe at different degrees that it is going to rain

CF (E1) = -0.8 CF (E2) = -0.6

CF1 (H, E1) = CF (E1) * CF (RULE 1)

= -0.8 x 0.8

= -0.64

CF2 (H, E2) = CF (E2) * CF (RULE 2)

= -0.6 * 0.8

= -0.48

Since both < 0

$CF_{combined}$ (CF1, CF2) = CF1 + CF2 * (1 + CF1)

= - 0.64 - 0.48 * (1 - 0.64)

= -0.81 #

Show incremental decrease when more than one source disconfirming evidence is found

Certainty Factor Propagation (C.F.P)

Case 4: Several Sources Predict Rain at the same level of belief but one source
predicts no rain

If many sources predict rain at the same level, CF(Rain) = 0.8, the CF value converge towards 1

$CF_{combined}$ (CF1, CF2 ..) 0.999 = CF_{old}
CF_{old} = collected old sources info.

If new source say negative

$CF_{new} = -0.8$

then,

$$CF_{combined}(CF_{old}, CF_{new}) = \frac{CF_{old} + CF_{new}}{1 - \min\{CF_{old}, CF_{new}\}} = \frac{0.999-0.8}{1 - 0.8} = 0.995$$

Shows single disconfirming evidence does not have major impact

Certainty Factor Propagation (C.F.P)

CERTAINTY PROPAGATION FOR COMPLEX RULES

Combination of conjunctive and disjunctive statement (“AND”, “OR”)

Example:

IF	E1	
AND	E2	
OR	E3	max
AND	E4	
THEN	H	

CF (H) = Max {min (E1, E2), min (E3, E4)} * CF (RULE)

CF Example Program

- To Illustrate CF through a set of rules, we consider a small problem of deciding whether or not I should go to a ball game.
- Assume the hypothesis “I should not go to the ball game”.

Rule 1

IF	the weather looks lousy	E1
OR	I am in a lousy mood	E2
THEN	I shouldn’t go to the ball game	H1 CF = 0.9

Rule 2

IF	I believe it is going to rain	E3
THEN	the weather looks lousy	E1 CF = 0.8

Rule 3

IF	I believe it is going to rain	E3
AND	the weatherman says it is going to rain	E4
THEN	I am in a lousy mood	E2 CF = 0.9

CF Example Program

Rule 4

IF	the weatherman says it is going to rain	E4
THEN	the weather looks lousy	E1 CF = 0.7

Rule 5

IF	the weather looks lousy	E1
THEN	I am in a lousy mood	E2 CF = 0.95

Assume the following CF:

I believe it is going to rain	CF(E3) = 0.95
Weatherman believes it is going to rain	CF(E4) = 0.85

Assume backward chaining is used with the goal of :

"I shouldn't go to the ball game", H1

Rule 1

IF	the weather looks lousy	E1
OR	I am in a lousy mood	E2
THEN	I shouldn't go to the ball game	H1 CF= 0.9

Rule 2

IF	I believe it is going to rain	E3
THEN	the weather looks lousy	E1 CF = 0.8

Rule 3

IF	I believe it is going to rain	E3
AND	the weatherman says it is going to rain	E4
THEN	I am in a lousy mood	E2 CF = 0.9

Rule 4

IF	the weatherman says it is going to rain	E4
THEN	the weather looks lousy	E1 CF = 0.7

Rule 5

IF	the weather looks lousy	E1
THEN	I am in a lousy mood	E2 CF = 0.95

Assume the following CF:

I believe it is going to rain	CF(E3) = 0.95
Weatherman believes it is going to rain	CF(E4) = 0.85

Advantages and Disadvantages

ADVANTAGES:

a. It is a simple computational model that permits experts to estimate their confidence in conclusion being drawn.
b. It permits the expression of belief and disbelief in each hypothesis, allowing the expression of the effect of multiple sources of evidence.
c. It allows knowledge to be captured in a rule representation while allowing the quantification of uncertainty.
d. The gathering of CF values is significantly easier than gathering of values for other methods such as Bayesian method.

Advantages and Disadvantages

DISADVANTAGES:

a. Non-independent evidence can be expressed and combined only by "chunking" it together within the same rule. With large quantities of evidence this is quite unsatisfactory.
b. CF values are unable to represent efficiently and naturally certain dependencies between uncertain beliefs.
c. CF of a rule is dependent on the strength of association between evidence and hypothesis. Thus, handling changes in the knowledge base is highly complex.