

# Chapter 5: Inexact Reasoning

Expert Systems: Principles and  
Programming, Fourth Edition

# Objectives

---

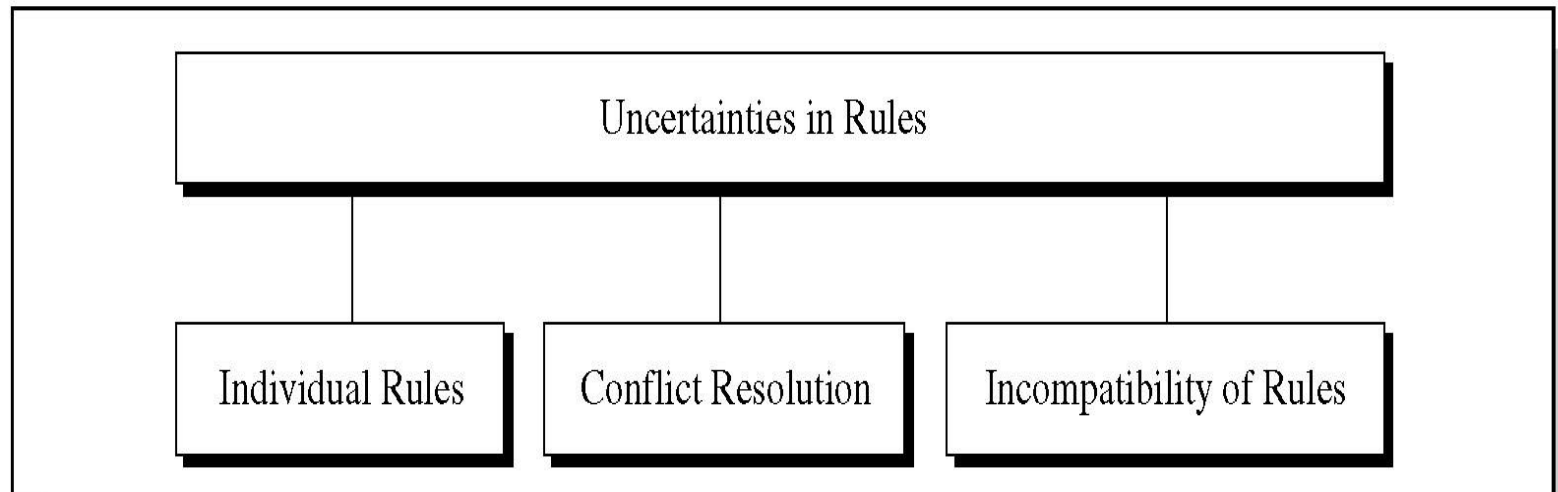
- Explore the sources of uncertainty in rules
- Analyze some methods for dealing with uncertainty
- Learn about the Dempster-Shafer theory
- Learn about the theory of uncertainty based on fuzzy logic
- Discuss some commercial applications of fuzzy logic

# Uncertainty and Rules

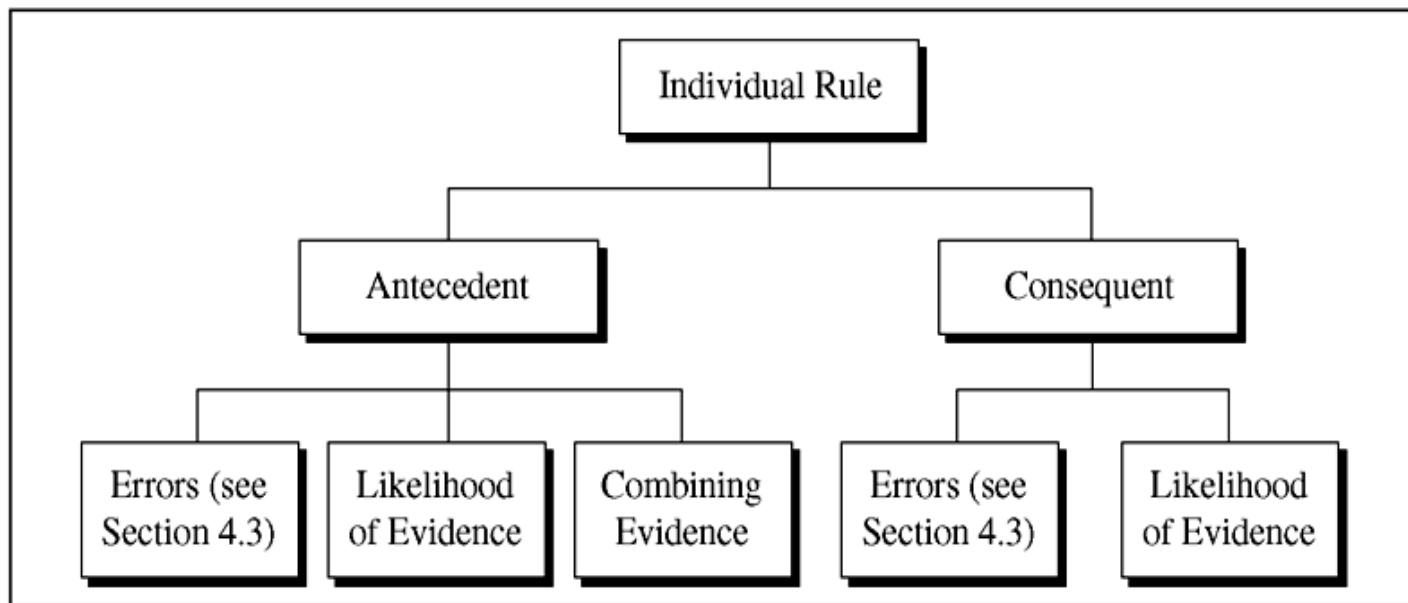
---

- We have already seen that expert systems can operate within the realm of uncertainty.
- There are several sources of uncertainty in rules:
  - Uncertainty related to individual rules
  - Uncertainty due to conflict resolution
  - Uncertainty due to incompatibility of rules

# Figure 5.1 Major Uncertainties in Rule-Based Expert Systems

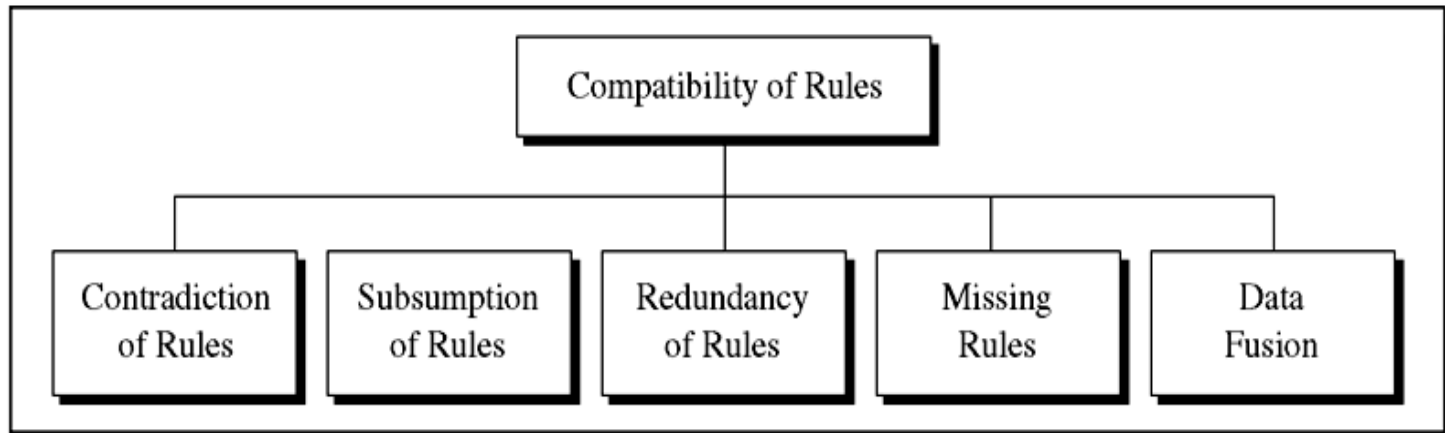


# Figure 5.2 Uncertainties in Individual Rules

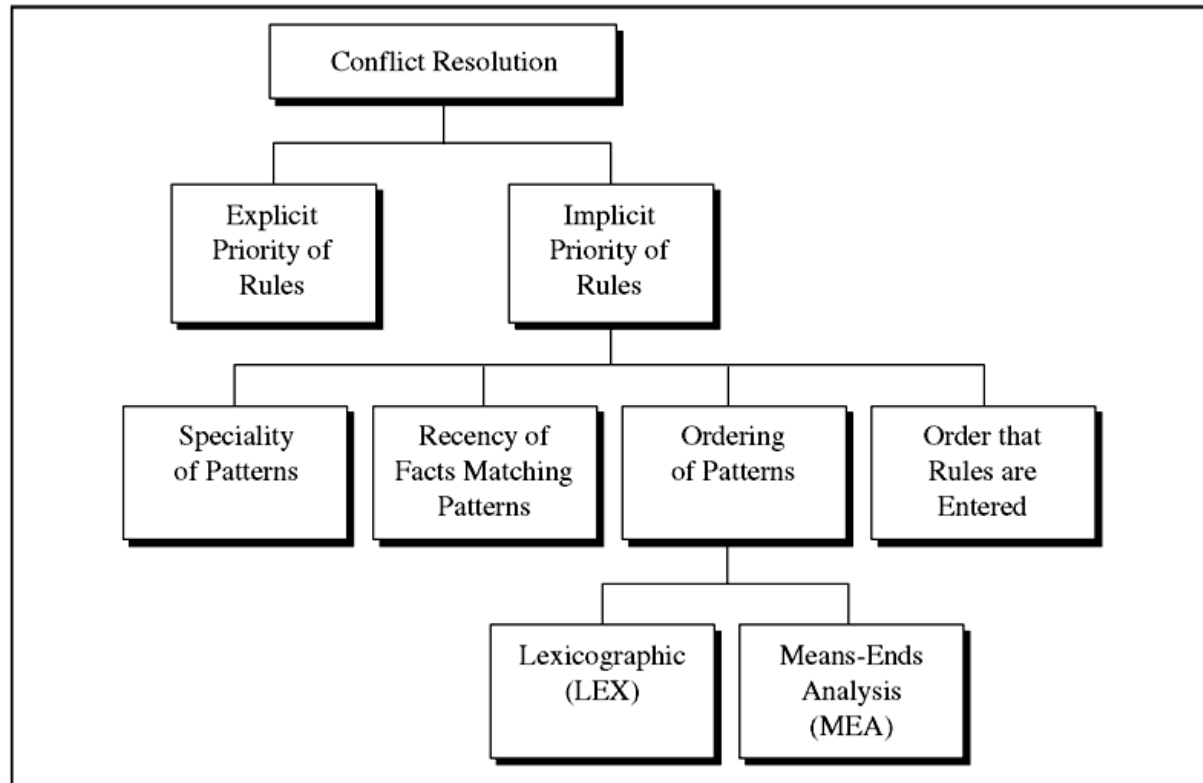


# Figure 5.3 Uncertainty Associated with the Compatibilities of Rules

---



# Figure 5.4 Uncertainty Associated with Conflict Resolution



# Goal of Knowledge Engineer

---

- The knowledge engineer endeavors to minimize, or eliminate, uncertainty if possible.
- Minimizing uncertainty is part of the verification of rules.
- Verification is concerned with the correctness of the system's building blocks – rules.



# Verification vs. Validation

---

- Even if all the rules are correct, it does not necessarily mean that the system will give the correct answer.
- Verification refers to minimizing the local uncertainties.
- Validation refers to minimizing the global uncertainties of the entire expert system.
- Uncertainties are associated with creation of rules and also with assignment of values.

# Ad Hoc Methods

---

- The ad hoc introduction of formulas such as fuzzy logic to a probabilistic system introduces a problem.
- The expert system lacks the sound theoretical foundation based on classical probability.
- The danger of ad hoc methods is the lack of complete theory to guide the application or warn of inappropriate situations.

# Sources of Uncertainty

---

- Potential contradiction of rules – the rules may fire with contradictory consequents, possibly as a result of antecedents not being specified properly.
- Subsumption of rules – one rule is subsumed by another if a portion of its antecedent is a subset of another rule.

# Uncertainty in Conflict Resolution

---

- There is uncertainty in conflict resolution with regard to priority of firing and may depend on a number of factors, including:
  - Explicit priority rules
  - Implicit priority of rules
    - Specificity of patterns
    - Recency of facts matching patterns
    - Ordering of patterns
      - Lexicographic
      - Means-Ends Analysis
    - Ordering that rules are entered

# Uncertainty

---

- When a fact is entered in the working memory, it receives a unique timetag – indicating when it was entered.
- The order that rules are entered may be a factor in conflict resolution – if the inference engine cannot prioritize rules, arbitrary choices must be made.
- Redundant rules are accidentally entered / occur when a rule is modified by pattern deletion.

# Uncertainty

---

- Deciding which redundant rule to delete is not a trivial matter.
- Uncertainty arising from missing rules occurs if the human expert forgets or is unaware of a rule.
- Data fusion is another cause of uncertainty – fusing of data from different types of information.

# Certainty Factors

---

- Another method of dealing with uncertainty uses certainty factors, originally developed for the MYCIN expert system.

# Difficulties with Bayesian Method

---

- The Bayesian method is useful in medicine / geology because we are determining the probability of a specific event (disease / location of mineral deposit), given certain symptoms / analyses.
- The problem is with the difficulty / impossibility of determining the probabilities of these givens – symptoms / analyses.
- Evidence tends to accumulate over time.



# Belief and Disbelief

---

- Consider the statement:  
“The probability that I have a disease plus the probability that I do not have the disease equals one.”
- Now, consider an alternate form of the statement:  
“The probability that I have a disease is one minus the probability that I don’t have it.”

# Belief and Disbelief

---

- It was found that physicians were reluctant to state their knowledge in the form:

“The probability that I have a disease is one minus the probability that I don’t have it.”

- Symbolically,  $P(H|E) = 1 - P(H'|E)$ , where  $E$  represents evidence

# Likelihood of Belief / Disbelief

---

- The reluctance by the physicians stems from the likelihood of belief / disbelief – not in the probabilities.
- The equation,  $P(H|E) = 1 - P(H'|E)$ , implies a cause-and-effect relationship between  $E$  and  $H$ .
- The equation implies a cause-and-effect relationship between  $E$  and  $H'$  if there is a cause-and-effect between  $E$  and  $H$ .

# Measures of Belief and Disbelief

---

- The certainty factor, CF, is a way of combining belief and disbelief into a single number.
- This has two uses:
  1. The certainty factor can be used to rank hypotheses in order of importance.
  2. The certainty factor indicates the net belief in a hypothesis based on some evidence.

# Certainty Factor Values

---

- Positive CF – evidence supports the hypothesis
- $CF = 1$  – evidence definitely proves the hypothesis
- $CF = 0$  – there is no evidence or the belief and disbelief completely cancel each other.
- Negative CF – evidence favors negation of the hypothesis – more reason to disbelieve the hypothesis than believe it

# Threshold Values

---

- In MYCIN, a rule's antecedent CF must be greater than 0.2 for the antecedent to be considered true and activate the rule.
- This threshold value minimizes the activation of rules that only weakly suggest the hypothesis.
- This improves efficiency of the system – preventing rules to be activated with little or no value.
- A combining function can be used.

# Difficulties with Certainty Factors

---

- In MYCIN, which was very successful in diagnosis, there were difficulties with theoretical foundations of certain factors.
- There was some basis for the CF values in probability theory and confirmation theory, but the CF values were partly ad hoc.
- Also, the CF values could be the opposite of conditional probabilities.

# Dempster-Shafer Theory

---

- The Dempster-Shafer Theory is a method of inexact reasoning.
- It is based on the work of Dempster who attempted to model uncertainty by a range of probabilities rather than a single probabilistic number.



# Dempster-Shafer

---

1. The Dempster-Shafer theory assumes that there is a fixed set of mutually exclusive and exhaustive elements called environment and symbolized by the Greek letter  $\Omega$ :

$$\Omega = \{ \text{document}_1, \text{document}_2, \dots, \text{document}_N \}$$


# Dempster-Shafer

---

- The environment is another term for the universe of discourse in set theory.
- Consider the following:
  - ☐ = {rowboat, sailboat, destroyer, aircraft carrier}
- These are all mutually exclusive elements

# Dempster-Shafer

---

- Consider the question:  
“What are the military boats?”
- The answer would be a subset of  :  
 $\{\text{📄}_3, \text{📄}_4\} = \{\text{destroyer, aircraft carrier}\}$

# Dempster-Shafer

---

- Consider the question:








“What boat is powered by oars?”

- The answer would also be a subset of  :

$\{\text{rowboat}\} = \{\text{rowboat}\}$

This set is called a singleton because it contains only one element.

# Dempster-Shafer

- Each of these subsets of  is a possible answer to the question, but there can be only one correct answer.
- Consider each subset an implied proposition:
  - The correct answer is: { <sub>1</sub>, <sub>2</sub>, <sub>3</sub> }
  - The correct answer is: { <sub>1</sub>, <sub>3</sub> }
- All subsets of the environment can be drawn as a hierarchical lattice with  at the top and the null set  $\emptyset$  at the bottom

# Dempster-Shafer

---

- An environment is called a frame of discernment when its elements may be interpreted as possible answers and only one answer is correct.
- If the answer is not in the frame, the frame must be enlarged to accommodate the additional knowledge of element..

# Dempster-Shafer

---

## 2. Mass Functions and Ignorance

In Bayesian theory, the posterior probability changes as evidence is acquired. In Dempster-Shafer theory, the belief in evidence may vary. We talk about the degree of belief in evidence as analogous to the mass of a physical object – evidence measures the amount of mass.

# Dempster-Shafer

---

- Dempster-Shafer does not force belief to be assigned to ignorance – any belief not assigned to a subset is considered no belief (or non-belief) and just associated with the environment.
- Every set in the power set of the environment which has mass  $> 0$  is a focal element.
- Every mass can be thought of as a function:

$$m: P(\text{folder}) \rightarrow [0, 1]$$



# Dempster-Shafer

---

## 3. Combining Evidence

Dempster's rule combines mass to produce a new mass that represents the consensus of the original, possibly conflicting evidence

The lower bound is called the **support**; the upper bound is called the **plausibility**; the **belief measure** is the total belief of a set and all its subsets.

# Dempster-Shafer

---

4. The moving mass analogy is helpful to understanding the support and plausibility.
  - The support is the mass assigned to a set and all its subsets
  - Mass of a set can move freely into its subsets
  - Mass in a set cannot move into its supersets
  - Moving mass from a set into its subsets can only contribute to the plausibility of the subset, not its support.
  - Mass in the environment can move into any subset.

# Approximate Reasoning

---

- This is theory of uncertainty based on fuzzy logic and concerned with quantifying and reasoning using natural language where words have ambiguous meaning.
- Fuzzy logic is a superset of conventional logic – extended to handle partial truth.
- Soft-computing means computing not based on classical two-valued logics – includes fuzzy logic, neural networks, and probabilistic reasoning.

# Fuzzy Sets and Natural Language

---

- A discrimination function is a way to represent which objects are members of a set.
  - 1 means an object is an element
  - 0 means an object is not an element
- Sets using this type of representation are called “crisp sets” as opposed to “fuzzy sets”.
- Fuzzy logic plays the middle ground – like human reasoning – everything consists of degrees – beauty, height, grace, etc.

# Fuzzy Sets and Natural Language

---

- In fuzzy sets, an object may partially belong to a set measured by the membership function – grade of membership.
- A fuzzy truth value is called a fuzzy qualifier.
- Compatibility means how well one object conforms to some attribute.
- There are many type of membership functions.
- The crossover point is where  $\mu = 0.5$

# Fuzzy Set Operations

---

- An ordinary crisp set is a special case of a fuzzy set with membership function  $[0, 1]$ .
- All definitions, proofs, and theorems of fuzzy sets must be compatible in the limit as the fuzziness goes to 0 and the fuzzy sets become crisp sets.

# Fuzzy Set Operations

<b>Set equality</b>	<b>Set Complement</b>
<b>Set Containment</b>	<b>Proper Subset</b>
<b>Set Union</b>	<b>Set Intersection</b>
<b>Set Product</b>	<b>Power of a Set</b>
<b>Probabilistic Sum</b>	<b>Bounded Sum</b>
<b>Bounded Product</b>	<b>Bounded Difference</b>
<b>Concentration</b>	<b>Dilation</b>
<b>Intensification</b>	<b>Normalization</b>

# Fuzzy Relations

- A relation from a set A to a set B is a subset of the Cartesian product:

$$A \times B = \{(a,b) \mid a \in A \text{ and } b \in B\}$$

- If X and Y are universal sets, then

$$R = \{ \mu_R(x, y) / (x, y) \mid (x, y) \in X \times Y \}$$



# Fuzzy Relations

---

- The composition of relations is the net effect of applying one relation after another.
- For two binary relations P and Q, the composition of their relations is the binary relation:

$$R(A, C) = Q(A, B) \circ P(B, C)$$

# Table 5.7 Some Applications of Fuzzy Theory

---

Control Algorithms  
Medical Diagnosis  
Decision Making  
Economics  
Engineering  
Environmental  
Literature  
Operations Research  
Pattern Recognition  
Psychology  
Reliability  
Security  
Science

# Table 5.8 Some Fuzzy Terms of Natural Language

---

tall  
hot  
low  
medium  
high  
very  
not  
little  
several  
few  
many  
more  
most  
about  
approximately  
left-winger

# Linguistic Variables

---

- One application of fuzzy sets is computational linguistics – calculating with natural language statements.
- Fuzzy sets and linguistic variables can be used to quantify the meaning of natural language, which can then be manipulated.
- Linguistic variables must have a valid syntax and semantics.

# Extension Principle

---

- The extension principle defines how to extend the domain of a given crisp function to include fuzzy sets.
- Using this principle, ordinary or crisp functions can be extended to work a fuzzy domain with fuzzy sets.
- This principle makes fuzzy sets applicable to all fields.

# Fuzzy Logic

---

- Just as classical logic forms the basis of expert systems, fuzzy logic forms the basis of fuzzy expert systems.
- Fuzzy logic is an extension of multivalued logic – the logic of approximate reasoning – inference of possibly imprecise conclusions from a set of possibly imprecise premises.

# Possibility and Probability and Fuzzy Logic

---

- In fuzzy logic, **possibility** refers to allowed values.
- **Possibility distributions** are not the same as **probability distributions** – frequency of expected occurrence of some random variable.

# Translation Rules

---

- Translation rules specify how modified or composite propositions are generated from their elementary propositions.
  1. Type I modification rules
  2. Type II composition rules
  3. Type III quantification rules
  4. Type IV quantification rules



# State of Uncertainty

## Commercial Applications

---

- There are two mountains – logic and uncertainty
- Expert systems are built on the mountain of logic and must reach valid conclusions given a set of premises – valid conclusions given that –
  - The rules were written correctly
  - The facts upon which the inference engine generates valid conclusions are true facts
- Today, fuzzy logic and Bayesian theory are most often used for uncertainty.

# Summary

---

- In this chapter, non-classical probability theories of uncertainty were discussed.
- Certainty factors, Dempster-Shafer and fuzzy theory are ways of dealing with uncertainty in expert systems.
- Certainty factors are simple to implement where inference chains are short (e.g. MYCIN)
- Certainty factors are not generally valid for longer inference chains.

# Summary

---

- Dempster-Shafer theory has a rigorous foundation and is used for expert systems.
- Fuzzy theory is the most general theory of uncertainty formulated to date and has wide applicability due to the extension principle.