Knowledge-Based Systems and Deductive Databases

Wolf-Tilo Balke
Christoph Lofi
Institut für Informationssysteme
Technische Universität Braunschweig
http://www.ifis.cs.tu-bs.de
10 Expert Systems

10.1 Expert Systems
10.2 Heuristic Reasoning
10.3 Fuzzy Reasoning
10.4 Case-Based Reasoning
10.1 Expert Systems

• **Expert Systems** have been the main application of A.I. in the early 80ties

• **Idea:** Create a system which can draw **conclusions** and thus support people in difficult decisions
  – Simulate a **human expert**
  – Extract knowledge of experts and just cheaply **copy** it to all places you might need it
10.1 Expert Systems

- **Expert Systems** were supposed to be especially useful in
  - **Medical diagnosis**
    - ...used to be a failure
    - Currently, has its come-back in specialized areas
  - **Production and machine failure diagnosis**
    - Works quite well
  - **Financial services**
    - Widely used
• Usually, three **user groups** are involved when maintaining and using an expert system
  
  – **End Users**: The group that actually uses the system for problem solving assistance
    • e.g. young and/or general doctors, field users deploying complex machinery, ...
  
  – **Domain Experts**: Are those experts whose knowledge is to be “extracted”
    • e.g. highly-skilled specialist doctors, engineers of complex machinery, ...
  
  – **Knowledge Engineers**: Assist the domain experts in representing knowledge in a formally usable form, e.g. representing it as rules
10.1 Expert Systems

- **Common architecture** of an expert system
  - **User Interface**: Usually based on a question-response dialog
  - **Inference Engine**: Tries to deduce an answer based on the knowledge base and the problem data
  - **Explanation System**: Explains to the user why a certain answer was given or question asked
  - **Knowledge Base**: Set of rules and base facts
  - **Problem Data**: Facts provided for a specific problem via user interface
10.1 Expert Systems

• **Building** an expert system has several steps

  – Building up the **knowledgebase** needs the extraction of knowledge in the form of rules and beliefs from domain experts
    • For complex domains it is almost impossible
  – Deciding for a suitable **reasoning technique**
    • This part is usually well-understood
  – Designing an **explanation facility**
    • Automatically generating sensible explanations or even arguments for derived facts is a major problem
    • Often only the proof tree is returned…
10.1 Expert Systems

- The actual way of performing deduction in expert systems may differ
  - Often Prolog/Datalog-based logic programming engines build the core
  - Heuristic approaches, like MYCIN
  - Fuzzy approaches
  - Case-based reasoning
10.2 MYCIN

- **MYCIN**
  - Developed 1970 at Stanford University, USA
  - **Medical expert system** for treating **infections**
    - Diagnosis of infection types and recommended antibiotics (antibiotics names usually end with ~mycin)
  - Around 600 rules (also supporting **uncertainty**)
  - MYCIN was treated as a success by the project team…
    - Experiments showed good results, especially with rare infections
  - … but was never used in practice
    - Too clumsy
    - Technological constraints
• **Design considerations**
  
  – **Uncertain reasoning** is necessary
  
  • There is no complete and doubt-free data in medicine
  
  – However, most known approaches for uncertain reasoning had some severe drawbacks
  
  • No real distinction between **doubt**, **lack of knowledge** and **absence of belief**
  
  • As seen in last lecture: You very often end up with confidence intervals of $[0, 1]$, i.e. **deductions are useless**
  
  • A lot of **additional facts or rules** are necessary to reliably use uncertain reasoning
10.2 MYCIN

• MYCIN pioneered the idea of **certainty factors** for uncertain deduction

  – **Certainty factors**: the relative change of belief in some hypothesis facing a given observation

  – **MYCIN** is a **heuristic** system
    
    • Rules provided by experts are heuristic rules (i.e. are usually correct, but not always)
    
    • Also, there are additional heuristics involved by making certain assumptions (like the underlying model or independence of observations)
• **MYCIN** example rule

If the organism 1) stains grampos 2) has coccus shape 3) grows in chains
then there is a suggestive evidence of 0.7 that it is *streptococcus*

– i.e. the expert stating this rule would strongly strengthen his/her belief in streptococcus when given the observations 1-3
• MYCIN example

--------PATIENT-1--------
1) Patient's name: FRED SMITH
2) Sex: MALE
3) Age: 55
4) Have you been able to obtain positive cultures from a site at which Fred Smith has an infection? YES

--------INFECTION-1--------

5) What is the infection? PRIMARY-BACTEREMIA
6) Please give the date when signs of INFECTION-1 appeared. 5/5/75
The most recent positive culture associated with the primary bacteremia will be referred to as:

--------CULTURE-1--------

7) From what site was the specimen for CULTURE-1 taken? BLOOD
8) Please give the date when this culture was obtained. 5/9/75
The first significant organism from this blood culture will be called:

--------ORGANISM-1--------

9) Enter the identity of ORGANISM-1. UNKNOWN
10) Is ORGANISM-1 a rod or coccus (etc.)? ROD
11) The gram stain of ORGANISM-1: GRAMNEG
• The certainty factor model is further based on measures of belief and disbelief
  – Certainty factor can be computed by combining belief and disbelief measures
  – Both are treated individually, i.e. increasing belief does not decrease disbelief automatically
• The informal definitions of \textit{disbelief} and \textit{belief} are as follows

- \textbf{Measure of belief} for hypothesis $h$ given the observation $E$
  
  $\text{MB}(h|E) = x$ means “In the light of evidence $E$, one’s belief that $h$ is true increases by $x$”

- \textbf{Measure of disbelief} for hypothesis $h$ given the observation $E$
  
  $\text{MD}(h|E) = x$ means “In the light of evidence $E$, one’s disbelief that $h$ is true increases by $x$”

- Belief and disbelief are normalized to $[0,1]$
Examples:

- $\text{MB}(\text{canFly}(x) | \text{isBird}(x)) = 0.8$
  - “Knowing that $x$ is a bird, my belief that $x$ can fly increases strongly by 0.8”

- $\text{MD}(\text{canFly}(x) | \text{isBiggerThan}(x, 2.00\text{m})) = 0.9$
  - “Knowing that $x$ is bigger than 2.00m, my disbelief that $x$ can fly increases strongly by 0.9”

- $\text{MD}(\text{canFly}(x) | \text{isBird}(x)) = 0.1$
  - “Knowing that $x$ is a bird, my disbelief that $x$ can fly increases by 0.1”
    - Could be a chicken, or penguin, or whatever
The **certainty factor** is finally the difference of belief and disbelief for a given pair hypotheses and observation:

- \[ CF(h|E) := MB(h|E) - MD(h|E) \]
- Thus, certainty factors are within \([-1, 1]\)
- A certainty factor describes the change of belief when a given fact/observation is known
  - It is thus a **relative measurement** combining belief and disbelief
10.2 Certainty Factors

– A **positive certainty factor** means that after learning a fact, my belief into something increases
  • The fact “confirms” the hypotheses
  • For negative certainty, the disbelief increases

– If only **certainty factors** are used for knowledge modeling, one can extract the according **belief** and **disbelief** values directly
  • This approach is used in MYCIN

\[
MB(\ldots) = \begin{cases} 
0 & \text{if } CF(\ldots) < 0 \\
CF(\ldots) & \text{if } CF(\ldots) \geq 0 
\end{cases} \quad MD(\ldots) = \begin{cases} 
-CF(\ldots) & \text{if } CF(\ldots) < 0 \\
0 & \text{if } CF(\ldots) \geq 0 
\end{cases}
\]
10.2 Certainty Factors

• Also note that \( CF(h|E) + CF(\neg h|E) \leq 1 \)
  – They are not probabilities! i.e. known equality
    \( P(h|E) + P(\neg h|E) = 1 \) does not hold for certainty factors

• This actually means
  – If some evidence supports an hypothesis, this **does not mean** that the negation is supported in the inverse manner
    • E.g., no reliable statements regarding the negation
10.2 Certainty Factors

• How are belief factors and certainty factors related to probability?
  – We will need a formalization in order to derive valid rules for combination and chaining of rules
  – For understanding and modeling knowledge and rules, the informal definition is usually used
    • I.e. the quantified change of belief when a given fact/observation is discovered
    • Assumption: The formal model matches the intended semantics of the informal definition
• **Measure of belief**

\[
MB(h|E) = \begin{cases} 
\frac{\max(P(h|E), P(h)) - P(h)}{1-P(h)} & \text{if } P(h) \neq 1 \\
1 & \text{otherwise}
\end{cases}
\]

– This means

• Is 0 if \( P(h|E) \leq P(h) \), i.e. the evidence does not increase the probability of the hypothesis

• Otherwise, is the increase in probability when giving a certain evidence in proportion to the uncertainty (improbability) of the hypothesis alone
10.2 Certainty Factors

- Definitions for measure of disbelief and certainty factor are analogously
  
  - Assumption: These statistical notation does represent a fuzzy concept of human increase of belief

\[
MD(h|E) = \begin{cases} 
\frac{P(h) - \min(P(h|E), P(h))}{P(h)} & \text{if } P(h) \neq 0 \\
1 & \text{otherwise}
\end{cases}
\]

\[
CF(h|E) = \begin{cases} 
\frac{P(h|E) - P(h)}{1 - P(h)} & \text{if } P(h|E) \geq P(h), P(h) \neq 1 \\
\frac{P(h|E) - P(h)}{P(h)} & \text{if } P(h) \geq P(h|E), P(h) \neq 0
\end{cases}
\]
10.2 Certainty Factors

• These definitions heavily rely on various a priori probabilities and conditional probabilities
  – Those are usually not known and / or cannot be determined
  – A user-provided certainty factor (based on informal definitions) thus proxies for all those probabilities
    • “Given observation E, my belief into h decreases by 0.3” thus implicitly contains information on $P(h|E)$, $P(h)$ and their relation
10.2 Certainty Factors

• So finally, the simplest form of rules using certainty factors is

  – IF a THEN h WITH CF(h|a)
  – Thus, we can have **confirming rules** (positive CF) or **disconfirming rules** (negative CF)

  – Based on this rule type, some simple operations may be defined
    • Chaining
    • Parallel Combination
10.2 Certainty Factors

• Cognitive user load using different models
  – **Strict reasoning:**
    “If there are black dots on teeth, then this is caries.”
    • Easy, but too restrictive and thus often leads to wrong rules
  – **Probabilistic reasoning:**
    “If there are black dots on teeth, then this is caries with a probability of 0.82.”
    • Absolute statement on probabilistic frequencies
    • Lots of statistical evaluation necessary to determine all needed a-priori and conditional probabilities
  – **Certainty factors:**
    “If there are black dots on teeth, then this is a strong positive (0.8) evidence for caries.”
    • Relative statement on strength of evidence
    • No absolute statistics necessary
10.2 Certainty Factors

• **Rule chaining**
  
  – Chain rules consecutively, e.g.
    
    - IF e THEN a WITH CF(a|e)
    - IF a THEN h WITH CF(h|a)
    - ⇒ IF e THEN h WITH CF(h|e)

  – CF(h|e) = MB(h|e) - MD(h|e) can be computed from it’s components as follows
    
    - MB(h|e) = MB(h|a) * max(0, CF(a|e))
    - MD(h|e) = MD(h|a) * max(0, CF(a|e))
    - Thus, chaining is essentially a simple multiplication
• Parallel combination
  – Combining multiple rules for the same hypothesis
    • IF \( e \) THEN \( h \) WITH \( CF(h|e_1) \)
    • IF \( a \) THEN \( h \) WITH \( CF(h|e_2) \)
  – Parallel combination should be undefined when both certainty factors are opposing with maximal certainty
• The combined certainty factor can be computed independently by determining the belief and disbelief values

\[ MB(h|e_1,e_2) = \begin{cases} 
0 & \text{if } MD(h|e_1,e_2) = 1 \\
MB(h|e_1) + MB(h|e_2) - MB(h|e_1) \times MB(h|e_2) & \text{otherwise}
\end{cases} \]

\[ MD(h|e_1,e_2) = \begin{cases} 
0 & \text{if } MB(h|e_1,e_2) = 1 \\
MD(h|e_1) + MD(h|e_2) - MD(h|e_1) \times MD(h|e_2) & \text{otherwise}
\end{cases} \]
10.2 Certainty Factors

**Example:**

- If there is are **black dots** on teeth, my **belief** in **caries increases** moderately (0.5).
- If the x-ray shows **no damage** to the **adamantine**, then my **belief** in caries **decreases** strongly (-0.9).

  - $\text{CF(caries|dots)} = 0.5$, $\text{CF(caries|noDamage)} = -0.9$
  - $\text{CF(caries|dots,noDamage)} = ?$
  
  - $\text{MB(caries|dots)} = 0.5$, $\text{MD(caries|dots)} = 0$
  - $\text{MB(caries|noDamage)} = 0$
  - $\text{MD(caries|noDamage)} = 0.9$

- $\text{MB(caries|dots, noDamage)} = 0.5 + 0 - 0.5*0 = 0.5$
- $\text{MD(caries|dots, noDamage)} = 0 + 0.9 - 0*0.9 = 0.9$
- $\text{CF(caries|dots, noDamage)} = -0.4$
10.2 Certainty Factors

- If the **gum is red**, my belief in **periodontitis** increases moderately (0.5).
- If the there **are loose teeth**, my belief in **periodontitis** increases slightly (0.3).
  
  - **CF** (periodontitis|redGum) = 0.5,
  
    - **CF** (periodontitis|looseTeeth) = 0.3
  
    - **CF** (periodontitis| redGum, looseTeeth) = ?

  - **MB** (periodontitis|redGum) = 0.5,
  
    - **MD** (periodontitis|redGum) = 0
  
    - **MB** (periodontitis|looseTeeth) = 0.3
  
    - **MD** (periodontitis|looseTeeth) = 0

- **MB** (periodontitis|rg, lt) = 0.5 + 0.3 - 0.5*0.3 = 0.65
  
  - **MD** (periodontitis|rg, lt) = 0 + 0 - 0*0 = 0
  
  - **CF** (periodontitis| redGum, looseTeeth) = 0.65
How did the actual MYCIN system work?

- Only **confirming** or **disaffirming** rules with certainty factors

- For each patient, a **predefined set of standard facts** has to be provided
  - Like age, general, general condition, common facts on the sample, etc
  - These are used to rule out all completely unrealistic conclusions
10.2 MYCIN

--------PATIENT-1--------
1) Patient's name: FRED SMITH
2) Sex: MALE
3) Age: 55
4) Have you been able to obtain positive cultures from a site at which Fred Smith has an infection? YES

--------INFECTION-1--------
5) What is the infection? PRIMARY-BACTEREMIA
6) Please give the date when signs of INFECTION-1 appeared. 5/5/75
The most recent positive culture associated with the primary bacteremia will be referred to as:

--------CULTURE-1--------
7) From what site was the specimen for CULTURE-1 taken? BLOOD
8) Please give the date when this culture was obtained. 5/9/75
• After that, the systems switches to a **backward-chaining** approach
  – Most promising rules are selected, and the system tries to prove each of them
    • Discard all rules with known false premises
    • Prefer rules with high certainty factors
  – Missing information is requested from the user in a dialog-style interaction

The first significant organism from this blood culture will be called:

```
--------ORGANISM-1--------
9) Enter the identity of ORGANISM-1. **UNKNOWN**
10) Is ORGANISM-1 a rod or coccus (etc.)? **ROD**
11) The gram stain of ORGANISM-1: **GRAMNEG**
```
Finally, the system will present all possible deductions to the user along with their certainty factors.

After that, the user may *query* why the system deduced those facts (system *explains* the answers) – Mainly using the proof trees of each successful rule.

**Did you use RULE 163 to find out anything about ORGANISM-1?**

RULE163 was tried in the context of ORGANISM-1, but it failed because it is not true that the patient has had a genito-urinary tract manipulative procedure (clause 3).

**Why didn't you consider streptococcus as a possibility?**

The following rule could have been used to determine that the identity of ORGANISM-1 was streptococcus: RULE033

But clause 2 (“the morphology of the organism is coccus”) was already known to be false for ORGANISM-1, so the rule was never tried.
• Was MYCIN a success?
  – Partially…
  – During field evaluation, MYCIN deduced a correct treatment in 69% of all test infections
    • …which is a lot better than diagnoses by average non-specialist physicians
    • …but worse than diagnoses by infection specialists (~80% - however, those specialist often disagreed such that a real evaluation is not possible as there is no “gold” standard for infection treatments)
    • This result is very representative for most expert systems which perform worse than real experts, but usually better than non-experts
  – However, the system never made it into practice mainly due to legal and ethical issues
    • Who is responsible (and can be sued) in case of a mistake?
• **Vagueness** is in the nature of most expert decisions
  – Symptom for cavities: a person has **discolored** teeth. Not at all, slightly, very,…?! 

• The vagueness cannot only be modeled by an agent’s **belief** in a statement, but also directly
  – **Fuzzy set theory** (Lotfi Zadeh, 1965)
  – Expresses the degree of **possibility** (as opposed to probability)
  – Captures the idea of **linguistic variables**
• Crisp set membership degrees (1 or 0) are often insufficient for expressing vague concepts
  – Consider the set of discolored teeth in cavities diagnosis
    • Teeth with brown spots are discolored (1)
    • White teeth are not discolored (0)
    • What about yellowish teeth? Depends on the degree of stain! ([0, 1])
10.3 Fuzzy Reasoning

• A **fuzzy set** is defined by membership function $\mu$ mapping the universe $\Omega$ to the unit interval $[0,1]$

  – The normal **set operations** with the characteristic membership function can be easily extended for fuzzy sets
    
    • $(\mu_1 \cap \mu_2)(\omega) := \min\{\mu_1(\omega), \mu_2(\omega)\}$
    • $(\mu_1 \cup \mu_2)(\omega) := \max\{\mu_1(\omega), \mu_2(\omega)\}$
    • The complement of $\mu(\omega) := 1 - \mu(\omega)$

  – Some characteristics of **Boolean algebra** are preserved, others not
    
    • E.g., distributivity holds, but DeMorgan’s laws not…
10.3 Fuzzy Reasoning

• How can this be applied for reasoning?
  – We have **fuzzy facts** and can deduce **new (fuzzy) facts** from them
  – Back to our **toothache** example…
    • **Fact**: Tom has yellow stained teeth.
    • **Rule**: If a person has very discolored teeth, it is cavities.
    • **Does Tom have cavities..?!**

• Obviously we need to relate the **degree of staining** with the **premise** of our rule
  – Usually the degree is (linguistically) **discretized**
  – Different degrees of staining have different **possibilities** of relating to cavities
This leads to **possibility distributions**

- Only depend on the **possibility** that a case is described by a certain class
  - Nobody would state that somebody with white teeth has ‘discolored’ teeth
  - There is a possibility of 50% that a yellow stain would be considered as ‘discolored’

- Are somewhat similar to **probability** distributions, but not depending on observed cases
  - Possibility is an **upper limit** for probabilities

- **Possibility theory** introduced by L. Zadeh in 1978
A possibility distribution assigns the possibility of a characteristic to some measurable property.

- Somebody has ‘discolored’ teeth.
- Somebody has ‘very discolored’ teeth.
10.3 Fuzzy Reasoning

• An important feature is the ability to define **hedges**
  – Provide operations to maintain close ties to **natural language**, and allow for the **mathematical generation** of fuzzy statements
  – The initial definition of hedges is a **subjective** process

• A simple example may transform the statement ‘teeth are stained.’ to ‘teeth are **very** stained.’
  – The hedge ‘very’ is usually defined as \( \mu_{\text{very}}(\omega) := (\mu(\omega))^2 \)
    • Thus, if \( \mu_{\text{stained}}(\text{Tom}) = 0.8 \), then \( \mu_{\text{very\_stained}}(\text{Tom}) = 0.64 \)
  – Other common hedges are ‘more or less’, typically \( \text{SQRT}(\mu(\omega)) \), ‘somewhat’, ‘rather’, ‘sort of’, ...
• Still, possibility distributions have to be linked to determining the **truth of conclusions**

  – Idea is a **conditional** possibility distribution

  • possibility(yellow stains | teeth are very discolored) = truth(teeth are very discolored | yellow stains)

  • The first part uses the fuzzy **membership function** describing the classes of all stains that are considered as discolored
10.3 Fuzzy Reasoning

• Now let’s turn to the problem of reasoning
  – Consider the general case with fuzzy sets $A, A'$ over $\Omega_1$ and $B$ over $\Omega_2$
    • Fact: $X$ is $A'$.
    • Rule: if $X$ is $A$, then $Y$ is $B$.
  – Depending on the connection between $A$ and $A'$ the inference will result in the conclusion $Y$ is $B'$ with a fuzzy set $B'$ over $\Omega_2$

• How can this be calculated?
  – Encode each piece of information by possibility measures corresponding to suitable fuzzy sets
10.3 Fuzzy Reasoning

• If knowledge from two or more facts with respective possibility distributions has to be combined...
  – Then first the facts have to be aggregated (using min, max, …)
  – Secondly, the aggregated possibility distribution has to be established (corresponding e.g., to the conjunction of the facts)
10.3 Fuzzy Reasoning

• The actual **inference process** applying rules in fuzzy expert systems has usually four steps
  
  – Fuzzification
  – Inference
  – Composition
  – Defuzzification

– Called **Mamdani-style** fuzzy inference introduced by Ebrahim Mamdani of London University, 1975
10.3 Fuzzy Reasoning

• **Fuzzification**
  – Membership functions defined on the input variables are applied to their actual values, to determine the **degree of truth for each rule premise**

• **Inference**
  – The **truth value** for the premise of each rule is computed, and **applied to the conclusion part** of each rule
    • Either cut the consequent membership function at the level of the antecedent truth value (**clipping**)
    • Or adjust the consequent membership function by multiplying all its membership degrees by the antecedent truth value (**scaling**)
  – This results in one **fuzzy subset** to be assigned to each output variable for each rule
10.3 Fuzzy Reasoning

• **Composition (or aggregation)**
  
  – **Unification** of the outputs of all rules
  
  – All of the fuzzy subsets assigned to each output variable are *combined together* to form a single fuzzy subset for each output variable

• **Defuzzification**
  
  – It may be useful to just examine the *fuzzy subsets* that are the result of the composition process
  
  – More often, this fuzzy value needs to be converted to a single *crisp value* (called defuzzification)
10.3 Fuzzy Reasoning

- Example: ‘Tom’s teeth have yellow stains’
  - Fuzzification: to what degree are they ‘slightly discolored’, ‘discolored’, ‘very discolored’, …?
  - Inference: apply all inputs to the fuzzy rules and calculate the degrees of the conclusion
    - ‘if teeth are slightly discolored, then cavities is unlikely’, …, ‘if teeth are very discolored, then cavities is almost sure’
    - This leads to a possibility distribution for a diagnosis of cavities
10.3 Fuzzy Reasoning

- **Composition**: aggregate all membership degrees for the different conclusions using ‘\( \bigcup \)’

- **Defuzzification**: there are several defuzzification methods, but probably the most popular one is the centroid technique
  
  - It finds the point where a vertical line would slice the aggregate set into two equal masses (centre of gravity)
Case-based reasoning (CBR) is a methodology for solving problems by utilizing previous experiences. It is not really a formal reasoning process, but relies on heuristics to arrive at conclusions.

- Similar to case-based law systems using precedents...
- Or case analysis in medical treatments...
- Or repairing a car...
• Examples

– Cooking banana pancakes is like cooking normal pancakes… just throw in some bananas…

– Biomimicry: imitate nature to utilize natural effects for complex engineering tasks
  • E.g., how to cool houses in Africa without air-conditioning?
  • Idea: the same way termites build hives
10.4 Case-Based Reasoning

• **General operation**
  – Present the system with a problem
  – Search a case base for most similar problems
  – Return their solutions
• Cases are records of previous experiences
  – Problem specification
  – Relevant attributes of the environment
  – Applied solution
  – Benefit/success of the solution

• Representation needs to reflect all features necessary for retrieval
• **4-phase model** proposed by Agnar Aamodt and Enric Plaza in 1994
10.4 Case-Based Reasoning

• **Retrieve**
  – Given a target problem, retrieve cases from memory that are relevant to solving it

• **Reuse**
  – Map the solution from the previous case to the target problem

• **Revise**
  – Test the new solution in the real world (or a simulation) and, if necessary, revise

• **Retain**
  – Store successfully adapted experiences as a new case in memory
10.4 Case-Based Reasoning

- **Case retrieval** is the process of finding closest cases, i.e., most similar cases, to the current case
  - (Indexed) features of cases in the case base are compared to the features of the current case
  - **Syntactical** approaches vs. **semantic** approaches
  - The hardest part of the CBR process is defining a suitable **similarity measure**
    - Nearest neighbor retrieval, hierarchical browsing, knowledge-guided approaches, validated retrieval, …
    - Often a **semi-automatic** process
• **Case adaptation** translates the retrieved solution into a solution appropriate for the current problem
  
  – Applied in the reuse phase (basic adaptation) and in the revise phase (learning from failure)
  
  – Often a manual process needing deeper domain understanding
  
  – The **degree of success** (and thus the value for the case base) has to be measured
10.4 Case-Based Reasoning

• **Case base maintenance** is part of the retain phase
  – The larger the case base, the more of the **problem space** is covered, but too many cases will degrade system **performance**
  – Maintenance strategies are quite similar to **caching strategies**
  – Is a case really **necessary** for the case base?
    • How successful was its solution?
    • Are there already similar cases?
    • How often is a specific case used?
10.4 Case-Based Reasoning

• Comparison to **rule-based systems**

  – **Rule bases**…
  
  * Abstract knowledge in a set of production rules of the ‘If...Then...’-type
  * Have to be acquired **before** the system can be used
  * Applicable to a **large** set of general domains
  * Provide **proofs** for derived statements

  – **Case bases**…
  
  * Only state **specific characteristics** of previous cases plus solutions
  * Are built up **while** the system is used
  * Applicable only for **specific kinds** of domains
  * Provide **arguments** for derived statements
10.4 Case-Based Reasoning

• **When** to use case-based reasoning?
  
  – Does the domain have an underlying **model**?
    • Random factors cannot be captured…
  
  – Are there **exceptions** and **novel cases**?
    • Without them rules might be easier…
  
  – Do cases **recur**?
    • If not, there is no point in building a case base…
  
  – Is there **significant** benefit in adapting past solutions?
    • The reasoning process might be more expensive than actually solving the problem…
• The Semantic Web
  – Visions and benefits
  – Basic constructs
• Representing information