CogSysI Lecture 9: Non-Monotonic and Human Reasoning

Intelligent Agents

WS 2006/2007

Part III: Reasoning and Inference

III.9 Non-Monotonic and Human Reasoning
Reasoning under Uncertainty

- Deduction in classical logic: based on the assumption that knowledge about the properties of the environment (set of axioms/rules) and perception (set of facts) is perfect!

- Uncertainty:
  - Rule *sore throat* → *flu* holds only with a certain probability.
  - Fact *The car is in working order* can be asserted only with some probability.
Approaches to Reasoning under U.

- Qualitative (logical) approach: non-monotonic reasoning
- Quantitative (probabilistic) approach: e.g. fuzzy logic
- Semi-Qualitative approach: Bayes nets
Bayes Nets

- Rules: Data $D \rightarrow$ Hypothesis $H$
- Bayes Theorem: 
  \[ p(H|D) = \frac{p(D|H) \cdot p(H)}{p(D)} \]

- Conditional probability of $H$ given $D$
- Apriori probabilities: $p(D)$, $p(H)$ and $p(D|H)$
- Example: Sore throat $\rightarrow$ flu
  
  \[ p(\text{Sore throat}) = 0.3 \] (probability that a person has a sore throat)
  
  \[ p(\text{flu}) = 0.4 \] (probability that a person has the flu)
  
  \[ p(\text{Sore throat}|\text{flu}) = 0.6 \] (probability that a person has a sore throat if he/she has the flu)

  With $p = 0.45$ we can infer that someone has the flu if he/she reports a sore throat

- Bayes Nets: directed acyclic graph (DAG) with concepts as nodes. A node has an outgoing edge to another node if it represents a precondition.
Non-Monotonic Reasoning

- A family of formal frameworks devised to capture inference of everyday life.

- Reasoners draw conclusions tentatively, reserving the right to retract them in the light of further information.

- Example: From “Tweety is a bird” we can follow “Tweety can fly” using the rule $\forall x \text{bird}(x) \rightarrow \text{flies}(x)$.
  The new information “Tweety is a penguin” brings us to a revision of the former conclusion. (“Tweety cannot fly”).

- Non-Monotonicity: Inference can be “undone” by new information.

- Some frameworks:
  - Default Logic (Reiter, 1980)
  - Truth Maintenance Systems (Doyle, 1980)
  - Modal Logic (dynamic logic, epistemic logic)
Default Logic

A default theory is a pair \((W, \Delta)\)

\(W\) is a “world description”, a set of first order formulas representing the “strict” or background information.

\(\Delta\) is a set of defaults, representing revisable information.

Inference rules in \(\Delta\) have the form: \(\gamma : \theta/\tau\).

\(\gamma\), \(\theta\), and \(\tau\) are also first order formulas with

- \(\gamma\): pre-requisite (conjunction of literals)
- \(\theta\): justification/consistency assumptions
- \(\tau\): default (revisable conclusion)

Interpretation: “If \(\gamma\) is known and if there is no evidence that \(\theta\) might be false, then \(\tau\) can be inferred.”

Justification is based on the closed world assumption: If \(\neg\theta\) cannot be derived from \(W\) then assume that it holds.
The Tweety Example

Theory:

\[ W = \{ \]
\[ \text{bird(Tweety),} \]
\[ \text{bird(Paul),} \]
\[ \text{penguin(Tweety),} \]
\[ \forall x \text{ penguin}(x) \rightarrow \neg \text{flies}(x) \} \]
\[ \Delta = \{ d_1 \} \text{ with } d_1 :: \text{bird}(x) : \text{flies}(x)/\neg \text{flies}(x) \]

From \( \gamma = \text{bird(Paul)} \) can be derived \( \text{flies(Paul)} \)

From \( \gamma = \text{bird(Tweety)} \) cannot be derived \( \text{flies(Tweety)} \) because there is evidence that \( \neg \text{flies(Tweety)} \)
Remarks on Default Logic

- Determining consistency of defaults is problematic
- Algorithmic approach: calculate extension (model) of a default theory
  Problem: extensions may not exist, are not unique
- To deal with non unique extensions, two strategies were proposed: a credulous and a bold strategy. In the first case, an inference is accepted if it is in some extension $E$, in the second case, it is accepted only if it is included in all extensions $E$.
- Determining whether a given default theory has an extension is highly intractable even for a simple subset of FOL, which only allows conjunctions of literals.
Fuzzy Logic

- Reasoning with expressions describing membership in fuzzy sets
- Membership function: $T(x) \in [0 \ldots 1]$ with 0 corresponding to logical “false” (not member) and 1 corresponding to logical “true” (member)
- Example: $Bird(x)$ can be defined such that it returns 0 for all kinds of fishes, 1 for “typical” birds (canary, sparrow), and smaller values for birds such as penguins or hens
- Problem: membership-function may depend on context (a soup bowl can be viewed as a cup if nothing else for pouring a drink in is available)
- Relation to prototype theory (Rosch)
Fuzzy Reasoning

- Standard rules for evaluating fuzzy truth
  \[ T(A \land B) = \min(T(A), T(B)) \]
  \[ T(A \lor B) = \max(T(A), T(B)) \]
  \[ T(\neg A) = 1 - T(A) \]

- For \( T(X) \in \{0, 1\} \) these rules correspond to the classical logical operations \textit{and}, \textit{or}, and \textit{not}
Human Reasoning

Everyday reasoning is necessarily non-monotonic: it is impossible to verify all premisses (knowledge not available, time constraints).

Three examples:
- Symbol-Distance Effect (Potts, 1972)
- Wason-Selection Task (Wason & Johnson-Laird, 1972)
- Syllogistic Reasoning with Mental Models (Johnson-Laird, 1983)
Symbol-Distance Effect

Presentation of assertions (in arbitrary sequence):
```
``Peter is taller than John’’
``‘‘John is taller than Mary’’
``‘‘Mary is taller than Bill’’
```

Task: Verification of a statement “x taller y”

Dependent variable: reaction time

Hypothesis: If humans solve this task with logical reasoning, then it holds, the reaction time will increase with the number of applications of the transitivity rule.

taller(x, y) ∧ taller(y, z) → taller(x,z)
Symbol-Distance Effect cont.

RT

Assumption

Data

taller(Peter, John)  taller(John, Mary)  taller(Mary, Bill)  taller(Bill, Sue)

inference steps

0 1 2 3

integrated mental representation:

Peter  John  Mary  Bill  Sue
Wason-Selection Task

“If of one side of a card is a vowel, then there is an even number on its other side”

Check this proposition by turning as few cards as possible.

A  D  4  7
Results:

<table>
<thead>
<tr>
<th></th>
<th>A, 4</th>
<th>A</th>
<th>A,7</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>46%</td>
<td>33%</td>
<td>4%</td>
<td>17%</td>
</tr>
</tbody>
</table>

**Modus ponens**: vowel → even, vowel $\models$ even

**Modus tollens**: vowel → even, $\neg$ even $\not\models$ $\neg$ vowel

- Subjects do not apply *modus tollens*
- Subjects might interpret implication as biconditional (checking A and 4)
- Subjects might prefer to check cards which are mentioned in the rule to be tested (“matching bias”)
- If context is given, subjects perform the correct tests! (92%) (Johnson-Laird et al., 1972; see Müsseler & Prinz, 2002)
Wason-Selection Task cont.

You work at a post office and need to check whether letters are stamped correctly: If an envelope is sealed it must have a 50 Lira stamp.
Mental Models

Syllogistic reasoning:
- Construction of an integrated internal representation of the premisses
- “Read out” the conclusion

Influence factors on performance (error rates, performance time)
- Number of possible models
- Sequence of presentation of premisses
Mental Models, Example

All squares are striped.
All striped objects have a bold margin.

All squares have a bold margin?

All squares are striped.
Some striped objects have a bold margin.

Some squares have a bold margin?