Knowledge Engineering

Semester 2, 2004-05

Michael Rovatsos
`mrovatso@inf.ed.ac.uk`

Lecture 7 – Model-Based Reasoning
4th February 2005

Where are we?

Last time . . .

- we discussed further issues in ontologies
  - Semantic networks
  - Description logics
  - Reasoning with default information

Today . . .

- Model-Based Reasoning Systems

Model-Based Reasoning

- So far, discussion focussed on general KR&R principles
- But what is their practical use?
- Discuss Model-Based Reasoning (MBR) as a "case study" in designing practical reasoning systems
- Basic idea: use a model of the system as a "simulation" of it to conduct reasoning about its behaviour
- Describe system in terms of its components and the interactions between them

---

Model-Based Reasoning

- Can be used in two ways:
  1. diagnosis (detection of faults)
  2. prediction of behaviour (for design & configuration)
- Here: Restriction to diagnostic tasks
- Interaction between predicted behaviour and actual observations ➡ identify system components that failed
- Particular challenge: identifying multiple simultaneous faults

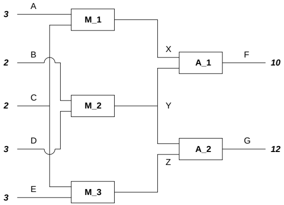
General Diagnostic Engine

- General Diagnostic Engine (GDE): a MBR engine intended to locate and isolate multiple simultaneous faults
- Assumptions:
  - Faults are in components, not in interconnections (unless these are defined as components)
  - Device representation is faithful
  - Faults are not intermittent
- Will look at extended example rather than precise algorithm

---

Example

Circuit of adders A_i and multipliers M_j, inputs A-E and outputs F, G



Minimal Candidates

- Basic problem: F should be 12 but is 10
- Treat input/output values (e.g. $A = 3$) as facts and statements facts like "M_1 is working" (written as M_1) as assumptions
- Can generate further facts under assumptions give:
  1. $X = 6\{M_1\}$
  2. $Y = 6\{M_2\}$
  3. $Z = 6\{M_3\}$
  4. $Z = 6\{M_2, A_2\}$ (from 2. and $G = 12$)
  5. $X = 4\{M_2, A_1\}$ (from 2. and $F = 10$)
  6. $Y = 4\{M_1, A_1\}$ (from 1. and $F = 10$)
  7. $Z = 8\{M_1, A_1, A_2\}$ (from 6. and $G = 12$)

Minimal Candidates

- Contradiction btw. 1. and 5. ➡ not all of M_1, M_2 and A_1 are working (same conflict caused by 6.)
- Conflict btw. 7. and 3. ➡ not all of M_1, A_1, A_2, M_3 are working
- At least one of $\{M_1, M_2, A_1\}$ and at least one of $\{M_1, M_3, A_1, A_2\}$ are faulty
- Set of **minimal candidates**: $\{A_1\}$, $\{M_1\}$, $\{A_2, M_2\}$, $\{M_2, M_3\}$ (minimal sets of components that would explain both assertions)

➡ Attention should focus on A_1 and M_1 ➡ measure X (measurement becomes a new fact and process continues)

Candidate Discrimination

- Problem with above procedure: generates too many possible faults
- How to identify best measurements to distinguish between candidates?
- Recall that new predictions are stored as statements $x = v\{e_1, \ldots, e_m\}$ where v is the value of x warranted by the minimal set of environments $\{e_1, \ldots, e_m\}$
- Any measurement that contradicts a predicted value is a conflict for the supporting environments
- In previous example: $X = 4$ vs. $X = 6$ resulted in one of $\{A_1\}$, $\{M_1\}$, $\{A_2, M_2\}$, $\{M_2, M_3\}$ being faulty

---

Candidate Discrimination

- Cases after measurement:
  - $X = 4$, conflict with $\{M_1\}$ ➡ $\{M_1\}$ becomes new minimal candidate
  - $X = 6$, conflict with $\{A_1, M_2\}$ and $\{A_1, A_2, M_3\}$ ➡ new candidates $\{A_1\}$, $\{M_2, M_3\}$ and $\{A_2, M_2\}$
  - $X \neq 4$ and $X \neq 6$, conflict with $\{A_1, M_2\}$, $\{A_1, A_2, M_3\}$ and $\{M_1\}$ ➡ minimal candidates $\{A_1, M_1\}$, $\{M_1, M_2, M_3\}$, $\{A_2, M_1, M_2\}$
- In this simple example, X was identified beacuse more probable singletons $\{M_1\}$ and $\{A_1\}$ are differentiable with its measurement

---

Candidate Discrimination

- In general case: hypothesize over possible measurements
- Choose minimal entropy $\sum_i -p_i \log p_i$ where p_i is probability that i-th remaining candidate is culprit
- Let m the number of values for x_i, S_{ik} the number of candidates which require x_i to have value v_{ik} and U_i the set of candidates that will not be eliminated regardless of the value of x_i
- Choose x_i that minimises

$$\sum_{k=1}^{m}[p(S_{ik}) + \frac{p(U_i)}{m}]\log[p(S_{ik}) + \frac{p(U_i)}{m}] - p(U_i)\log\frac{1}{m}$$

- Relies on probabilities of failure for components to calculate $p(S_{ik})$ and $p(U_i)$

Simplified method

- Assume that all components fail independently with equal probability (strong assumption!)
- Consider only candidates with minimum number of elements $= N$
- Let c_{ik} number of candidates that predict value v_{ik} for variable x_i
- Choose x_i that minimises $\sum_k c_{ik} \log c_{ik}$
- Iteratively perform one-step lookahead for $N = 1$, $N = 2$, etc.

Example

- In our example, two single-component candidates: $\{M_1\}$, $\{A_1\}$ $(N = 1)$
- Possible measurements:
  - $X = 4$ ➡ M_1 faulty (since it predicts $X = 6$), A_1 not (it is part of environments $\{A_1, M_2\}$ and $\{A_1, A_2, M_3\}$
  - $X = 6$ ➡ A_1 faulty
  - $Y = 6$ or $Z = 6$ ➡ A_1 or A_2 faulty
  - Things like $Y = 4$ are ruled out in present consideration (its supporting environment would be $\{A_1, M_1\}$ (same for $Z = 8$)
- One component that predicts either value for X, two for the only possibly value for Y and Z
- Entropies $X: \; 1 \log 1 + 1 \log 1 = 0, \; Y/Z: \; 2 \log 2 = 1.4$

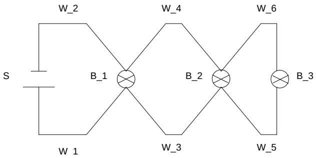
Introducing Fault Models

- GDE based on idea of "component is faulty if retraction of its correctness assumption is consistent with observations"
- But no knowledge of how components might fail
- Consider following example: If some bulbs in an electrical circuit are not lit, GDE would also consider that lit bulbs are faulty since they operate without power and battery is empty
- Logically consistent but counter-intuitive
- Solution: include explicit **fault models** such that if each of the known possible faults contradicts observations the component can't be faulty

---

Example

Observations: B_3 is lit while B_1 and B_2 are off



Minimal candidates: $\{B_1, B_2\}$, $\{S, B_3\}$, $\{S, W_5\}$, $\{W_2, W_5\}$ etc. (22 total)

Fault Models

- Only $\{B_1, B_2\}$ reasonable, otherwise wires would have to produce voltage or bulb lit without voltage
- But GDE would require further measurements ...
- Use following fault models
  - Bulb broken
  - Wire broken
  - Battery empty
- First one rules out all candidates in which B_3 occurs
- Since previous candidates were minimal, delete those with deleted elements
- B_3 is lit, so there is current ➡ eliminate all candidates with faulty battery or wires

Summary

- Model-based reasoning
- General Diagnostic Engine
- Candidate Discrimination
- Fault Models
- Next time: **Reasoning with Uncertainty**