Exam Knowledge Systems
August 14th 2003

Questions	points
question 1: Knowledge Systems	10+10
question 2: Time & Space	10+10
question 3: Uncertainty	10+10
question 4: Classification	5+5+5+5
question 5: Diagnosis	5+5+5+5
TOTAL:	100
mark = points / 10	

Good luck!

Question 1: Knowledge System

a. What is meant by "fidelity of a knowledge representation" ? Illustrate with an example how "fidelity" is related to the goal?

b. What is meant by a canonical representation form? Illustrate this with an example.

ANSWER:

a. Whether the statements in the model are true in reality (correct).

b. A unique representation for each meaning.

For instance: the representation of set in an alphabetic ordered list. The sets: {a,b,c}, {c,b,a}, {a,c,b} have the same canonical form [a,b,c].

Question 2: Time & Space

a. We have incomplete knowledge about the time an action (e.g. studying the material for KS) has taken place. How can you express this incompleteness in a point-based time representation and in an interval based time representation?

b. What is the problem with "nearest-first search" algorithm when objects are non-uniformly distributed in the space? Which representation method can solve this problem?

ANSWER

a.

Point based: the action X happens between point t1 and t2.

Interval based: the action X of duration d1 happens in the interval of i1([Ts,Te]).

b.

Some of the regions will be full. This results in expensive exhaustive search At the other hand some of the regions will be small. Problem too many small regions.

Solution: quad tree (only splitting up these regions that have more then x objects.)

Question 3: Uncertainty

This question is about the "certainty factor model".

a. Given:

observation A, with certainty factor 0.7
observation B, with certainty factor 0.9
observation C, with certainty factor - 0.8
observation D, with certainty factor 0.2
rule: if A then E, certainty factor: 0.8
rule: if B then E, certainty factor: 0.9
rule: if C then F, certainty factor: 0.8
rule: if D then F, certainty factor: 0.6
rule: if (E or F) then G, certainty factor:0.7

What is the certainty factor of G? Make clear how you got the value of the certainty factor of G.

b. Give two disadvantages and two advantages of the "certainty factor model".

ANSWER:

a.

cf(A) = 0.7

cf(B) = 0.9

cf(C) = - 0.8

cf(D) = 0.2

cf(if A then E) = 0.8

cf(if B then E) = 0.9

cf(if C then F) = 0.8

cf(if D then F) = 0.6

cf(if (E or F) then G) = 0.7

cf(C | {if A then E, if B then E}) =

cf(A).cf(if A then E) = 0.7 * 0.8 = 0.56

cf(B).cf(if B then E) = 0.9 * 0.9 = 0.81

$x > 0, y > 0$: x+y-xy = 0.7 * 0.8 + 0.9 * 0.9 - (0.7 * 0.8 * 0.9 * 0.9) = 0.9164

cf(C | {if C then F, if D then F}) =

cf(C).cf(if C then F) = max(0,-0.8) * 0.8 = 0

cf(D).cf(if D then F) = 0.2 * 0.6 = 0.12

$x = 0, y > 0$: x+y = 0 + 0.12 = 0.12

(x+1)/(1-min(abs(x),abs(y)))

cf(G) = cf(E or F) * cf(if (E or F) then G) = 0.9164 * 0.7 = 0.64148

cf(E or F) = max(cf(E), cf(F)) = max(0.9164, 0,12) = 0.9164

cf(G) = 0.64148

b.

Advantages:

- Nicely fits the rule-based paradigm, no essential change in representation fomalism necessary.

- Easy to implement.

- Some practical successes.

- Propagation rules at first sight not unreasonable.

Disadvantages: - Interpretation, meaning of CF unclear.

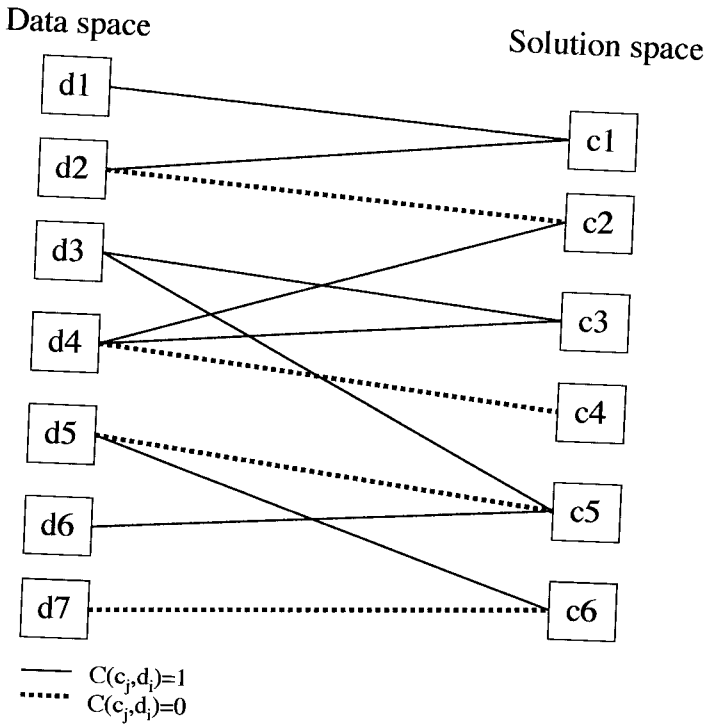
- No decision theory associated with CF.

- Propagation rules cannot be justified.

- Successes of CF can be explained to be largely independent of CF model.

Question 4: Classification

Given the following classification knowledge:



Answer the following questions with the **given observations**:
$\{d1 = 0, d2 = 1, d3 = 1, d4 = ?, d5 = 0, d6 = 1, d7 = 0\}$.

a. Which classes are inconsistent with the given observations, and which classes are consistent with the given observations.

b. Which data elements are irrelevant for the class c1?

c. Which classes are a "positive coverage" given the observations?

d. Can a hierarchical classification method be useful for the given classification knowledge? If yes, give then the hierarchical solutionspace. If no, show why.

ANSWER:

a. consistent are: c3, c4, c5

inconsistent are c1, c2, c6

b. d3 - d7, have no line, and are therefore irrelevant.

c. c3 (because of d3=1), c5 (all relevant data for c5 is obtained). These are the consistent classes, which have at least one relevant observation that has been obtained.

d. Yes.

Abstract solution for c3+c5 with d3, and abstract solution for c2+c3 with d4. Remove: C(c2,d4)=1, C(c3,d3)=1, remove C(c3,d4)=1, remove C(c5,d3)=1.

C(c1,d1)=1

C(c1,d2)=1

C(c2,d2)=0

C(c4,d4)=1

C(c5,d5)=0

C(c5,d6)=1

C(c6,d5)=1

C(c6,d7)=0

C(c3+c5,d3)=1

C(c2+c3,d4)=1

Question 5: Diagnosis

a. What is meant by model based diagnosis?

b Give a concrete model based system, and motivate why this system is a model based system.

c Give an advantage of a model based system.

d. Why is the single fault assumption not very realistic? Motivate your answer.

ANSWER:

a. There is a model from the artefact that has to be diagnosed. For instance a causal model.

b. GDE. GDE uses a structural model, and behaviour models for each type of component.

c. Reusability of diagnostic method. Reusability of models of the artefact.

d. interdependent faults:

fault1 $->$ fault2

cause $->$ fault1 and fault2

End exam