Faculteit Wiskunde en Informatica Vrije Universiteit, Amsterdam Joris Hulstijn	**Re-Exam Knowledge-based Systems** Wednesday, 15th of August, 13.30 – 16.30

name:
student number:

Assessment:

Exercise 1	Exercise 2	Exercise 3	Exercise 4	Exercise 5	additional
10 +10	10 +10	10 +10	12 +8	10	10

Exercise 1: General

a. Applying knowledge-based systems in practice has *advantages* as well as *disadvantages*. Name at least four advantages and three disadvantages of the use of knowledge-based systems.

b. Development of a knowledge-based system proceeds in a number of *phases*. Name the five development phases for knowledge-based systems. Describe the *end-product* of each phase, and their most characteristic *activities*.

Exercise 2: Representation

a. There are roughly three types of spatial representation: *nearest first*, *Quadtree* and *hierarchical* representation. What are the distinguishing characteristics of these spatial representation types?

b. After your yearly check-up, the doctor has bad news and good news. The bad news is that you tested positive for some serious disease. The test is 99% accurate (i.e. the probability of testing positive given that you have the disease is 0.99, as is the probability of testing negative if you don't have the disease). The good news is that the disease is rare: it occurs for only 1 in every 10.000 people.
Why is this good news? In other words: what is the chance that you actually have the disease?

Exercise 3: Classification and Data Abstraction

a. In classification raw data often first need to be combined and abstracted from. Name four types of *data abstraction*. Give an example of each type.

Consider the following classification scheme (fig 1). First raw data (R_i) are combined according to logical operators and transferred into abstract data (D_i), which are then used for hierarchical classification into a number of solution classes (S_i). Sub-classes with a common ‘parent’ are *mutually exclusive*.

Indicate for each of the following vectors of data, whether the solution classes ($S_1 \ldots S_8$) are *matching*, *consistent* or *inconsistent* with the data.

1. R = (? ? 0 1 1 ? ?)
2. R = (1 1 0 0 1 10)

Fig 1: classification using data abstraction

Exercise 4: Scheduling

Making the daily work schedule for postmen and women is a complex task[1]. Postmen and women are assigned to a particular mail area. There are the following conditions. Every mail area must be done. Every postman or woman with a regular contract must be assigned some area. The other areas are divided among so called `temporary employees'. Postmen and women specialise in their 'own area'; in this way the job is done quicker and better. One area may have several such 'specialists'.

a. Explain that making a mail area schedule can be seen as an instance of a configuration task. Make use of a drawing to sketch how the specification and configuration space, and the various knowledge sources and models are filled in.

In practice, making a mail area schedule proceeds in a number of steps. No backtracking is needed between steps.

step 0. *Make a list of all mail areas, along with the specialist for each area. make a list of all employees with a regular contract, and all temporary employees who are available, i.e. not ill or off duty.*
step 1. *For each area, see if there are specialists with a regular contract. If so, take the first. If not, continue.*
step 2. *For each remaining postman or woman with a regular contract, assign an arbitrary area.*
step 3. *For each remaining area, see if there are specialists among the temporary staff. if so, take the first. If not, take a temporary staff member. If there are none left, start making phone calls to employees off duty to ask for overtime.*

b. Argue whether this configuration method is an example of MCF1, MCF2 of MCF3.

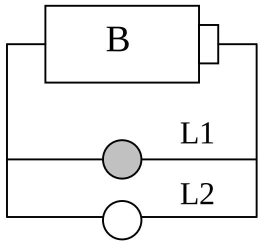
---

[1] Thanks to Marije Geldof, Ruud Schellekens, Nanja Smets en Gertjan Wijnalda for the example.

Exercise 5: Diagnosis

Consider the following electric circuit, with a battery B and two lights L1 and L2 in parallel. We predict that L1 and L2 are both on. However, we observe that L1 is off, but that L2 is on! We use the GDE method, in which there are no specific fault models.

a. What are the conflict sets for these observations?
b. According to the GDE method, what are the minimal covering sets of broken parts that would explain these observations?
c. Is this outcome in accordance with the expected behaviour of batteries?
d. Would a specific fault model for batteries change this outcome? If so, how? If not, why not?



Observations:

L1 off

L2 on