Universität des Saarlandes
Fachrichtung 6.2 – Informatik
J. Siekmann, S. Autexier, K. Fischer, E. Melis, C. Zinn
M. Schiller, E. Schulz, R.Vollmann

1st Theoretical Assignment in Artificial Intelligence (WS 2006/2007)
Solutions

Note: You need not hand these exercises in, and they are not graded. But bring your solutions to the tutorials on Friday 3rd. The questions in the minitests at the beginning of the tutorials will be based on these topics.

Exercise 1.1 **(10 P)**

1. Define the concept *machine*. (5 P)

Solution:

A machine is any device that transmits or modifies energy to perform or assist in the performance of tasks. Normally, it requires an input as trigger, transmits the modified energy to an output that performs the desired task. Additionally, a machine works according to physico-chemical laws and shows predictable and reproducable behavior.

---

2. Use this definition to discuss whether human beings can be considered as machines.(5 P)

Solution:

In some sense humans are machines. The brain works as reasoning device, we get triggering effects from the environment that represent the input to activate our reasoning process, then, our brain tries to find a solution to the given problem and executes the particular action as output. But what about emotions, desires and so on? Humans do not always act predictably (by other humans). The question whether humans are more than just machines depends on whether humans are considered to posess a free will or not (see also determinism[1]).

Exercise 1.2 **(15 P)**

1. Do you think that a pile of stones and toilet paper can think? (3 P)

Solution:

In principle, a pile of stones and toilet paper cannot think.

---

[1] http://en.wikipedia.org/wiki/Determinism

2. Briefly discuss whether computers can think. (6 P)

Solution:

The term "thinking" has been defined by humans, by introspection into their own thinking process. Therefore, computers often fall short of imitating human modes of reasoning. However, there are computer programs that engage in activities like problem solving, reasoning and decision making, and demonstrate that they can in principle perform these "thinking" activites. However, it is unclear whether running a computer program is sufficient to create the phenomenon of "consciousness", which humans experience when thinking ("I think, therefore I am" – Descartes)

---

3. Weizenbaum (*Computer Power and Human Reason*, Chapter 2, 1976) has shown how a construction made from stones and toilet paper can be used to simulate a Turing machine (and thus computation). Does this change your answers to 1 and 2? Discuss your answer. (6 P)

Solution:

No. This shows that thinking and intelligence do not depend on the material of the "machinery" on which they are run, whether this is human brain cells, a computer, or a more unorthodox arrangement.

Exercise 1.3 **(15 P)**

Briefly discuss the four different forms of agents illustrated in the lecture. Name the abilities of the particular type and give examples for problems it can/cannot solve. Can you imagine other forms of agents? Show advantages of this *new* form of agent.

Solution:

- *Simple reflex agent: for each input from the environment, there is a precisely defined output specifying how to react.*
  - *such an agent can: play chess*
  - *not suitable for: serving as a waiter in a restaurant*
- *State-based agent: simple reflex agent with internal states in order to choose an action.*
  - *such an agent can: clean the floor*
  - *not suitable for: planning a trip*
- *Goal-based agent: agent has some sort of goal information describing states which are desirable.*
  - *such an agent can: artificial player in computer games*
  - *not suitable for: diplomatic negotiations*
- *Utility-based agent: provided with some function that evaluates the utility of particular goals. This allows to resolve goal conflicts and to consider the relative importance of one goal related to the other goals.*

- *such an agent can: perform security management tasks*
- *such an agent can't: learn (unless it's a learning agent :-))*

- *New form of agents: social acting agents that try to maximize the utility of the group instead of maximizing the own utility. This newform of agents could provide solutions to social dilemmas (e.g. prisoners dilemma).*

Exercise 1.4 **(20 P)**

1. What is a *Turing test*? (5 P)
2. Give four questions you would ask in a Turing test. Discuss why you pick these questions. (10 P)
3. Discuss whether the Turing test is appropriate to decide if computers can think. (5 P)

Solution:

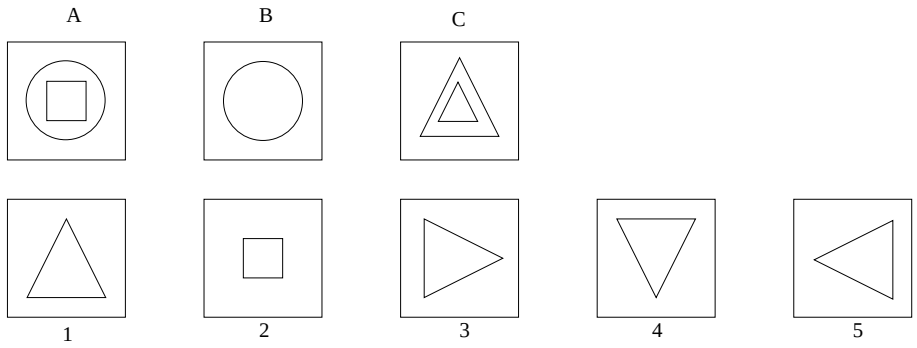
1. *An imitation game, where a human can ask questions to find out if its interlocutor is a person or a computer.* *(5 P)*
2. *Again much freedom. Good strategies are:*
   - *Asking arithmetic questions difficult for persons to see how quickly the answer comes (e.g., what is* 123×456*?).*
   - *Asking questions about feelings and emotions (e.g., How did you sleep tonight? What do you feel when you listen to Bach?).*
   - *Asking questions about behavior typical for humans (e.g., When did you go to the dentist last?)*
   - *Asking a conundrum*[2] *(a puzzling question, sometimes taking the form of a riddle), asking a rhetorical question or a making a joke. An example for a conundrum is: "If a sock becomes so darned that none of the originial yarn is there, then is it still the same sock?"*

   (10 P)
3. *It's a good argument for the Turing test that for now there don't exist computerprograms which could pass it (see also Jabberwacky and Loebner-Preis*[3]*). It's nevertheless important to dicuss the pros and cons.*

Exercise 1.5 **(30 P)**

Consider the following analogy test:

[2] http://en.wikipedia.org/wiki/Conundrum
[3] http://de.wikipedia.org/wiki/Loebner-Preis



1. Give a symbolic representation for each of the figures A-C, 1-5. Assume that a database exists with the prototypical objects `SQUARE, CIRCLE` and `TRIANGLE`. (15 P)

   What kind of problems do you encounter when applying the technique presented in the lecture to specify the objects and relations? (5 P)

2. Give a symbolic specification for each of the following transformations

   $A \rightarrow B, C \rightarrow 1, \ldots, C \rightarrow 5$. (5 P)

3. Generalise the symbolic specifications for each pair of transformation rules

   $(A \rightarrow B, C \rightarrow 1), \ldots, (A \rightarrow B, C \rightarrow 5)$
   Which figure do you obtain using the schema described in the lecture? Which figure would you have chosen? (5 P)

Solution:

Part 1

A
```
(FIGURE A (CONSISTS-OF P1 P2)
        (RELATIONS
         (INSIDE P2 P1))
        (SIM P1 CIRCLE (TRANS K 1 0 K))
        (SIM P2 SQUARE (TRANS K 0.5 0 K))
)
```

B
```
(FIGURE B (CONSISTS-OF P3)
        (SIM P3 CIRCLE (TRANS K 1 0 K))
```

C
```
(FIGURE C (CONSISTS-OF P4 P5)
        (RELATIONS (INSIDE P5 P4)
        (SIM P4 TRIANGLE (TRANS K 1 0 K))
        (SIM P5 TRIANGLE (TRANS K 0.5 0 K)))
```

1
```
(FIGURE 1 (CONSISTS-OF P6)
        (SIM P6 TRIANGLE (TRANS K 1 0 K))
```

2
```
(FIGURE 2 (CONSISTS-OF P7)
        (SIM P7 SQUARE (TRANS K 0.5 0 K))
```

3
```
(FIGURE 3 (CONSISTS-OF P8)
        (SIM P8 TRIANGLE (TRANS K 1 pi/2 K))
```

4
```
(FIGURE 4 (CONSISTS-OF P9)
        (SIM P9 TRIANGLE (TRANS K 1 pi K))
```

```
5 (FIGURE 5 (CONSISTS-OF P10)
        (SIM P10 TRIANGLE (TRANS K 1 -pi/2 K))
```

- *After assuming that the objects are already contained in the knowledge base, the main problem is to identify the translation. Do we need to rotate the circle? Is the square inside the circle or does the circle enclose the square?*

Part 2

```
(TRANSFORMATION A-->B (REMOVE X2 ((INSIDE X2 X1)
                                  (SIM X2 SQUARE (TRANS K 0.5 0 K))))
                     (MATCH X1 FROM ((INSIDE X2 X1)
                                     (SIM X1 CIRCLE (TRANS K 1 0 K)))
                               TO ((SIM X1 CIRCLE (TRANS K 1 0 K)))
                               WITH (TRANS K 1 0 K)))

(TRANSFORMATION-C-->1 (REMOVE X2 ((INSIDE X2 X1)
                                  (SIM X2 TRIANGLE (TRANS K 0.5 0 K))))
                      (MATCH X1 FROM ((INSIDE X2 X1)
                                      (SIM X1 TRIANGLE (TRANS K 1 0 K)))
                                TO ((SIM X1 TRIANGLE (TRANS K 1 0 K)))
                                WITH (TRANS K 1 0 K)))

(TRANSFORMATION-C-->2 (REMOVE X2 ((INSIDE X2 X1)
                                  (SIM X2 TRIANGLE (TRANS K 0.5 0 K))))
                      (REMOVE X1 ((INSIDE X2 X1)
                                  (SIM X1 TRIANGLE (TRANS K 1 0 K))))
                      (ADD X3 ((SIM X3 SQUARE (TRANS K 0.5 0 K)))))

(TRANSFORMATION-C-->3 (REMOVE X2 ((INSIDE X2 X1)
                                  (SIM X2 TRIANGLE (TRANS K 0.5 0 K))))
                      (MATCH X1 FROM ((INSIDE X2 X1)
                                      (SIM X1 TRIANGLE (TRANS K 1 0 K)))
                                TO ((SIM X1 TRIANGLE (TRANS K 1 pi/2 K)))
                                WITH (TRANS K 1 pi/2 K)))

(TRANSFORMATION-C-->4 (REMOVE X2 ((INSIDE X2 X1)
                                  (SIM X2 TRIANGLE (TRANS K 0.5 0 K))))
                      (MATCH X1 FROM ((INSIDE X2 X1)
                                      (SIM X1 TRIANGLE (TRANS K 1 0 K)))
                                TO ((SIM X1 TRIANGLE (TRANS K 1 pi K)))
                                WITH (TRANS K 1 pi K)))

(TRANSFORMATION-C-->5 (REMOVE X2 ((INSIDE X2 X1)
                                  (SIM X2 TRIANGLE (TRANS K 0.5 0 K))))
                      (MATCH X1 FROM ((INSIDE X2 X1)
                                      (SIM X1 TRIANGLE (TRANS K 1 0 K)))
                                TO ((SIM X1 TRIANGLE (TRANS K 1 -pi/2 K)))
                                WITH (TRANS K 1 -pi/2 K)))
```

Part 3

```
(GEN-A->B+C->1 (REMOVE X2 ((INSIDE X2 X1)
                           (SIM X2 Q (TRANS K 0.5 0 K))))
               (MATCH X1 FROM ((INSIDE X2 X1)
                               (SIM X1 R (TRANS K 1 0 K)))
                     TO ((SIM X1 R (TRANS K 1 0 K)))
                     WITH (TRANS K 1 0 K)))

(GEN-A->B+C->2 (REMOVE X2 ((INSIDE X2 X1)
                           (SIM X2 Q (TRANS K 0.5 0 K)))))

(GEN-A->B+C->3 (REMOVE X2 ((INSIDE X2 X1)
                           (SIM X2 Q (TRANS K 0.5 0 K)))))

(GEN-A->B+C->4 (REMOVE X2 ((INSIDE X2 X1)
                           (SIM X2 Q (TRANS K 0.5 0 K)))))

(GEN-A->B+C->5 (REMOVE X2 ((INSIDE X2 X1)
                           (SIM X2 Q (TRANS K 0.5 0 K)))))
```

The first generalisation consists of the most concordances. Therefore, the algorithm would select the first figure as solution. We would also choose the first figure as solution.

Exercise 1.6 **(10 P)**

Consider a human, who understands only English, equipped with a English rule book that includes instructions on how to manipulate Chinese symbols, and sheets of blank papers. The person sits in a room with a small opening to the outside. Through the opening appear slips of paper with (for our human) indecipherable Chinese symbols. The human finds matching symbols in the rule book, and, by following the instructions, eventually writes Chinese symbols on a paper slip. Then, he passes back the paper slip to the outside world. The first paper slip was a Chinese question, and the second paper slip is the appropriate Chinese answer.

Discuss controversely if the described system can be considered as understanding Chinese.
(pros: 5P, cons: 5P)

Solution:

Note: This is Searle's famous Chinese room *argument*[4].

Contra: *The person is operating on the level of the syntax of Chinese writings, but has no idea about the semantics of the Chinese characters (and neither do the rule book nor the papers convey the meaning of the writings).* *(5 P)*

Pro: *The system as a whole has the understanding of Chinese, even though its subparts (the person, the book, the papers) don't.* *(5 P)*

---

[4]The "Chinese room" is formulated in "Minds, Brains, and Programs," by John R. Searle, The Behavioral and Brain Sciences, vol. 3., Cambridge University Press