TDT4136 Introduksjon til kunstig intelligens
Assignment 1

Odd André Owren

September 2019

1. Two definitions used by Merriam-Webster's dictionary[2] which was not used in the lecture is:

I. "A branch of computer science dealing with the simulation of intelligent behavior in computers"

II. "The capability of a machine to imitate intelligent human behavior"

The final definition I've found is from Encyclopedia Britannica[1] and is stated as "The ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings."

2. The Turing test was developed by Alan Turing in 1950 with the intention of testing an intelligent machine's ability to behave equivalent to a human.
The test is conducted by having an intelligent system and a human answer a given set of questions from an interrogator, who is also human. The interrogator makes use of the responses from the machine and the human to evaluate which one is human and which one is a machine. If the interrogator is unable to tell the machine from the human, the machine is said to have passed the Turing test.

3. Thinking rationally is the act of making a choice that is perceived to be the most beneficial. This is usually based on a combination of situation, environmental factors and knowledge one already has. This knowledge is often informal and hard to translate to an algorithmic approach for a computer.
Acting rationally is related to acting through one's own beliefs to reach a goal. There is usually a connection between thinking rationally and acting rationally, but this is far from the case in all situations. Humans do have reflexes and other biologic reactions to the environment which gives us no time to think about how to act, but more often than not we act rationally. Thus acting rationally is not absolutely dependent on thinking rationally.

4. *The theory of reference* gives us a description of how to give real world objects and instances links to logic for use in e.g. computer systems.

5. Rationality is defined as an action or thinking that is based on or agreeable to reason. Rationality should not only be concerned with whether an action is reasonable or not, but also factor in the most optimal solution to a given task or situation.

6a. Yes. The robot, as well as humans, are not able to take all possible situations into the equation, and will most likely not consider the possibility of this happening as the chance of this is minimal. It would rather be a case of paranoia if one were to control every little factor when making a rational choice.

6b. The initial choice of crossing the road on a green light (given that the robot has a green light, and not the crossing cars) could be labelled as a rational choice. Although, it is very little rational to not reevaluate a choice when the robot has the actions to perceive a change of environment such that crossing the road would not be rational.

7a. It could be. It is rational for the robot to suck if an area is dirty, and it is also rational to move to another area if the area is clean. As the robot is a simple reflex agent, it would most likely follow the simple logic of moving on when an area is clean, and could end up with a heavy penalty if not factored in properly.

7b. As in the previous task, the simple logic is somewhat rational. In this case, when all areas have been cleaned, the robot could end up stationary as it is penalized by moving, and it's internal state would tell it that all areas are clean, so it would be rational not to move.

7c. In such a small environment we could assume that the agent would be rational, as there are few areas to keep track of at the same time.

```
def agent_vacuum():
    if wall_right == dirty:
        if position == right:
            suck()
        else:
            move_right()
    else if wall_left == dirty:
        if position == left:
            suck()
        else:
            move_left()
```

8. The environment described is:
- Partially observable: Agent does not have access to all states in the environment
- Single agent: Acting alone
- Deterministic: Next state is a function of current state and action by the agent

- Episodic: Agent perceives and acts upon it, then moves on
- Dynamic: The environment is changing as the agent is performing actions
- Discrete: The environment has a discrete amount of states and actions

9a. Simple reflex agents are fairly easy to implement, but is often very dumb. An environment needs to be very limited for a simple reflex agent to be rational.

9b. Model-based reflex agents keeps a model of how the environment may be changing. In fairly simple environments, this is a good approach to deal with partial observability, but if the model isn't good enough it will become unreliable, as the agent makes an educated guess as to how the environment is.

9c. Goal-based agents uses planning and searching to find an approach as to reach it's goals. This agent will be more generalized and therefore better suited to withstand noise and unforeseen factors in the environment, but is more complex.

9d. Utility-based agents applies searching and planning and uses a utility-function to find an optimal set of actions to solve a problem. A utility-based agent will be able to adjust it's choices through new information from the environment. This agent will again be a lot more complex than previously discussed agents, but is effective in more complex environments.

References

[1] B.J. Copeland. *Artificial intelligence*. URL: `https://www.britannica.com/technology/artificial-intelligence`. (accessed: 02.09.2019).

[2] Merriam-Webster. *Merriam Webster — Definition of Artificial Intelligence by Merriam-Webster*. URL: `https://www.merriam-webster.com/dictionary/artificial%20intelligence`. (accessed: 02.09.2019).