Artificial Intelligence Fundamentals and Intelligent Agents

1. My definition of AI is: A human creation that successfully completes complex tasks by learning and understanding.

   The definition of AI varies a lot, and often change over time due to new technological advances. Therefore, I favor the definition from the European Commission over Britannica and Cambridge. This definition is not restricted to computers or the human mind, but rather explains a general idea of how AI works.

   Britannica : “The ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings.”

   Cambridge Dictionary: “computer technology that allows something to be done in a way that is similar to the way a human would do it.” or “the study of how to make computers that have some of the qualities of the human mind, for example the ability to understand language, recognize pictures, solve problems, and learn.”

   European Commission's High-Level Expert Group on Artificial Intelligence: “AI systems act in the physical or digital dimension by perceiving their environment, processing and interpreting information and deciding the best action(s) to take to achieve the given goal. Some AI systems adapt their behavior by analyzing how the environment is affected by their previous actions.”

2. The Turing test is meant to test whether a computer is intelligent. It does this by having a human judge communicate with it digitally. The judge does not know it he/she is communicating with a human or computer. If the judge cannot decide if the computer he/she is communicating with is human or computer, the computer is according to the Turing test regarded as intelligent.

3. Rational thinking is just processing data to reach a result. This often leads to a rational action. However, a thinking creation can act rationally without thinking rationally. An example is human instincts which through evolution has made us act rationally without the need to think. If a human accidentally touches something hot and immediately retracts his/her hand, this is based on instincts and does not require the human to think rationally.
4. Rationality, according to the book, means doing "the right thing" given an agent's knowledge. It can also be defined as doing what creates a wanted outcome given an agent's knowledge. I think this is what the book means by "the right thing", but this can be misunderstood as doing something good. An agent that is programmed to do something that is considered bad can also be rational if the actions it performs is helping its goal.
5. Aristotle's claim require that an agent has relevant knowledge and a goal. He claims that one "predicts" the outcome of an action, and together with knowledge, decide how to act. Actions are results of knowledge and goals. He made an algorithm that says: To reach a goal, you can start at the goal and iterate necessary actions all the way to the beginning. Newell and Simon implemented these ideas in their GPS program. GPS stands for General Problem Solver and is meant to work as a universal problem solver. The input problem must be broken down to its most fundamental formulas. The goal is divided into subgoals all the way down, as Aristotle proposed. Even though it is called a general problem solver, it is not very general. It works on geometrical and logical problems, but not universally.
6.
    a. The robot never even had an opportunity to look up. Therefore, it cannot gain knowledge about the helicopter, and it arguably acted rationally based on its knowledge.
    b. The robot can look to left, where the car came from. If the robot has knowledge about the opportunity of cars coming, it is irrational. If it does not have this

knowledge, it is according to the book's definition not necessarily irrational because it is acting rational based on its knowledge.

7.

a. A simple reflex agent cannot be fully rational in this environment because it does not know if the other area is dirty or not. Therefore, some of its actions cannot be based on reason. Since it can only observe its current area, it would not know if it is worth taking a one-point penalty to cross to the other area.
b. A simple reflex agent with state can be rational based on its current knowledge, but it is not possible to know the state of an unobservable area. Its knowledge is limited, but based on that knowledge, it can act rational. The agent in a) does not have any of this necessary knowledge.
c. An agent that knows the state of both areas all the time can be rational given the correct algorithm. It will always know if it is worth crossing over to the other area and taking the penalty. One possible function for the vacuum cleaner:

```
if location = A:
        if statusA = Dirty then return Suck
        if statusB = Dirty then return Right

else if location = B:

        if statusB = Dirty then return Suck
        if statusA = Dirty then return Left
```

This does not lead to unnecessary travels between the A and B.

8. The environment is partially observable because the current tile is observable, but not the other. It is obviously single-agent because it is only one agent in the environment. It is also deterministic since there are no random elements or factors. However, it may not appear deterministic in an unobservable environment for the agent. It is sequential because every action depends on the last one. If the vacuum cleaner moves to B, it cannot vacuum A. The environment is continuously changing and is therefore dynamic,

but it is changing within certain rules and is also discrete. The agent knows the possible states of the environment, so the environment is known.

9.

a. When certain states require certain actions, a simple reflex agent can work fine. Simple reflex actions respond to states with given actions, but their intelligence is very limited. If the environment is not fully observable, the simple reflex agent may do some irrational things.
b. A model-based reflex agent is like a simple reflex agent but keeps an internal state of the unobserved parts of the environment. The state is based on earlier events. This is not bullet-proof since an agent cannot be sure about the unobserved.
c. Goal-based agents do not respond directly to environment states as the two above, but have a goal that they want to reach. This makes them more flexible when a state does not have one clear solution. One downside is that they are less efficient and must learn what to do in situations that are very easy for reflex agents.
d. Goal-based agents may complete the goal, but in an unwanted way. Utility-based agents solve this by implementing desirable ways of executing a task. This can for example make a car drive safer than a goal-based car. Utility-based agents are naturally much more complex than the others and does also have to learn what to do in simple situations where reflex agents thrive.