Assignment 1: AI Fundamentals

## Theoretical questions:

1. Artificial Intelligence does not have a single definition. In the book we are shown eight different definitions that have four different approaches. Three of these are:
   1. “[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning ...” (Bellman, 1978)
   2. “The study of the computations that make it possible to perceive, reason, and act.” (Winston, 1992)
   3. “AI . . . is concerned with intelligent behavior in artifacts.” (Nilsson, 1998)
2. The Turing Test is a test designed by Alan Turing used to determine whether an agent is intelligent or not. This test is conducted using a human interrogator who asks questions to two entities, a computer and a human. If the interrogator is unable to tell which answer is from the human and which answer is from the computer then the computer has passed the test.
3. Thinking rationally has to do with the thoughts going on inside someones head while acting rationally has to do with what actions are actually carried out. It is likely that if an entity thinks rationally then it will also act rationally, however it is not necessary to think rationally to act rationally.
4. Rationality means to draw conclusions using logic. The definition of rationality is as follows: “the quality of being based on or in accordance with reason or logic.”
5. Aristotle’s argument about the connection between knowledge and action is that actions are justified by goals and knowledge of an action’s outcome; if you know that an action’s outcome will bring you closer to your goal then you will perform that action. He suggests that you start at and end and try to figure out how you can get there. Once you’ve found this action you must figure out how to get to the point where you can perform this action. This creates an algorithm which eventually will bring you to where you are and will provide you with steps on how to achieve your goal. Newell and Simon were the first to implement this system in their GPS (general problem solver). This was intended to work as a universal problem solver machine, however it could not solve any real-world problems because of the combinatorial explosion.
6. Considering this robot:
   1. As the robot is unable to look upwards it cannot perceive the helicopter. Since it cannot perceive the helicopter it does not expect that crossing the road would do anything other than maximize its performance measure. This means that the robot is rational.
   2. As the road has a green light, the robot can expect cars to cross. Choosing to cross while the cars have a green light is the same as choosing an action that is not expected to maximize the performance measure. Hence the robot is not rational.
7. Considering:
   1. A simple reflex agent would choose to suck if the square it is on is dirty and would choose to move if the square it is on is clean. As the robot does not know whether the other square is clean or not it would risk moving when not necessary, which would cost it a performance point. This would not be maximizing its performance measure and hence it cannot be rational.
   2. A reflex agent with state would know which its percept sequence. If the agent does not know whether the other square is dirty or not then it can not be rational as it could risk moving when not necessary and hence waste points.
   3. If the simple reflex agent can perceive the status of both locations then it can be rational. This is because it can choose to move if the square it is on is clean and the other is dirty, and otherwise it would avoid movement. This would maximize it’s performance measure. Agent function:
      1. [A, Clean], [B, Clean] -> do nothing
      2. [A, Dirty], [B, Clean] -> suck
      3. [A, Clean], [B, Dirty] -> move
      4. [A, Dirty], [B, Dirty] -> suck

It would do the same actions if it was on square B, but opposite.

1. Properties:
   1. Observable: Partially. The cleaner only knows the state of the square it is on
   2. Agents: Single. There is only one agent, the cleaner.
   3. Deterministic: yes. The action made by the cleaner is the action that happens.
   4. Episodic: Sequential. If the cleaner has cleaned a square then that affects that square for some time (until It potentially gets dirty again)
   5. Static: Dynamic. Dirt reappears after some time, no matter what the cleaner deliberates.
   6. Discrete: Discrete. The environment has discrete happenings such as dirt being there vs dirt not being there, the cleaner sucking vs the cleaner not sucking and so on.
2. The different kinds:
   1. Simple reflex agents: The advantage of these is that they are simple; these are easy to make and very predictable. The disadvantage is that they are very limited and require that the correct action can be based on the current percept – which is not necessarily the case.
   2. Model-based reflex agents: These agents are rather simple but can do more than the simple reflex agents because they can store information about earlier percepts. Because of this, they can store knowledge about the rest of the world even if it is only partially observable, which is a big advantage over the simple reflex agent. The disadvantage is that they are still limited to reflexes and hence rather simple.
   3. Goal-based agents: These agents can do more than the reflex agents because they are able to consider which actions will make them reach their goal; they are flexible and can perform more complicated tasks that require consideration. The disadvantage is that they are less efficient than a reflex agent.
   4. Utility-based agents: These agents can consider more factors than simply reaching a goal; they can look at how efficient they reach the goal or even have several goals at once. This makes them perform their task better and makes them capable of doing tasks that consider more factors, i.e. are more realistic. The disadvantage is that they are more complicated and require longer to make.