# Assignment 1 - AI Fundamentals and Intelligent Agents

TDT4136 Introduction to Artificial Intelligence

Mads Adrian Simonsen

September 7, 2021

1. i) "AI is the theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages." - Oxford Languages / Google

   ii) "AI is the capacity of a computer, robot, or other programmed mechanical device to perform operations and tasks analogous to learning and decision making in humans, as speech recognition or question answering." - Random House Webster's Unabridged Dictionary / Dictionary.com

   iii) "AI is the branch of computer science involved with the design of computers or other programmed mechanical devices having the capacity to imitate human intelligence and thought." - Random House Webster's Unabridged Dictionary / Dictionary.com

2. The Turing test is a method which tests whether a machine can behave like a human when communicating via typed messages. The test is conducted by an interrogator having a conversation with the machine in five minutes. The interrogator then have to guess whether they were communicating with a human or a machine. If the interrogator guesses they were talking with a human at least 30% of the time, the machine passes the test.

3. *Thinking rationally* builds on logic, which is a series of statements that are either true or false, which is then used to conclude whether a new statement is true or false. This is a deterministic approach, but the world we live in is non-deterministic filled with uncertainty. Rational thought uses statistical tools to fill this uncertainty gap. However, rational thought is missing an important tool to make an impact to the world; acting. That is, we need an agent that acts according to the conclusions that were made by rational thought. *Acting rationally* is impacting the world by taking action that coincides with rational

thinking, meaning the agent acts to achieve the best outcome, or when there is uncertainty, best expected outcome.

As good as it may sound, thinking rationally is not an absolute condition for acting rationally. Sometimes, it is simply not feasible to compute the best possible outcome because the problem is too complex and time demanding. Therefore we must deal with *limited rationality* instead, e.g. act to achieve a local minimum rather than the more optimal global minimum for a cost function.

4. *Rationality* is a combination of thinking and acting rationally as described above. Loosely speaking, rationality seeks to "do the right thing". What "the right thing" is depends on what task the agents are given, and what their goal is. A more formal definition is
"Rationality is the quality of being based on or in accordance with reason or logic." - Oxford Languages / Google

5. Aristotle makes a connection between knowledge and action. He argues that actions are justified by a logical connections between goals, and knowledge of the outcome. Furthermore, Aristotle says that we do not discuss the end result, but the means. If there is a chain of steps needed, we start with the last one taking one step backwards, and look at the possible states and what means are needed to reach the end result. For each state, we take another step backwards and look for possible states and the means needed to go from the previous state to the current state. We repeat the process until we reach the beginning state. Then we have found a solution and can carry it out starting at the beginning state. If there are no routes that lead to the beginning state, we stop and conclude that the end result can not be reached.

   In 1959 (2300 years later), this algorithm was implemented by Newell and Simon in their *General Problem Solver* program. This program could in principle solve any problems that can be expressed as a set of well-formulated formulas (WWFs) or Horn clauses, that has the form of a directed graph with one or more sources and sinks.

6. a) We have no evidence that the robot is not rational as the robot crossing the street does not have "look-up" in its action portfolio, and therefore can not receive percepts regarding the helicopter above.

   b) Assuming the road is such that it is possible to have a clear sight such that the robot could receive percepts about the incoming car when performing the action "look-left-look-right", and additionally that the robot's prior knowledge of the environment includes that there is a positive probability that drivers could cross the street on red light, we conclude that the robot is irrational.

7. a) A simple reflex agent cannot be rational for the vacuum environment, because it can only receive percepts in the current square it is located in, and forgets all previous perceps. So if the last square was just cleaned, the agent will forget this, and will be returning to it sooner than needed, assuming the agent has a 0.5 probability of moving left and 0.5 probability of moving right, meaning an expected value of 2 time steps between moves.

   b) A reflex agent with state can however be rational. The agent might have some knowledge of the location's probability distributions of becoming dirty after some amount of time, so the agent can have an internal state that counts the number of times the agent has stayed in the same square, and use this number to decide if it should move to the adjacent square.

   c) With the environment being fully observable, the percept sequence of the agent gives no extra information. Thus, a simple reflex agent is sufficient to be rational.

```
function REFLEX-VACUUM-AGENT(statusA, statusB, location)
  status = {A : statusA, B : statusB}

  if status[location] == dirty
    return Suck
  else
    if status[!location] == dirty
      if location == A
        return Right
      else
        return Left
      end
    else
      return Nothing
    end
  end
end
```

8. The vacuum cleaner environment in Figure 2.2 in the textbook is:

   i) **nondeterministic**, as there is an uncertainty of when a square fill become dirty;

   ii) **episodic**, as the vacuum cleaner agent only cares about the current environment being clean regardless of past or future decisions;

   iii) **static**, since the agent is making short and simple computations, the environment does not have time to change before the agent has decided for an action;

iv) **discrete**, even though the amount of dirt increases continuously, the vacuum cleaner agent will receive percepts about the location being dirty once the location has reached a certain threshold of dirt;

v) **known**, as the agent knows what happens when it moves right or left, or suck.

9. a) **Simple reflex agents**
*Advantages:* As simple reflex agents does keep track of their percept sequence, they save a lot of memory. They are fast in their actionm and they are generally easy to implement.
*Disadvantages:* The agent is dependent on the environment of being fully observable. When the environment is partially observable, simple reflex agents will generally have problems of making correct decisions and can potentially get stuck in infinite loops.

b) **Model-based reflex agents**
*Advantages:* The agent can maintain its internal state which depends on its percept history, and therefore store information of some of the unobservable part of the current environment.
*Disadvantages:* Generally harder to implement, and can take longer time to deliberate decisions, meaning we could have a dynamic environment.

c) **Goal-based agents**
*Advantages:* When the agent has a goal, it can compute an optimal sequence of actions needed to reach the goal, instead of single steps that seem closer to the goal, e.g. following a GPS might get you faster and in one piece rather than blindly following a compass.
*Disadvantages:* Even harder to implement, and can be very computationally demanding, and sometimes infeasible to obtain the optimal choice.

d) **Utility-based agents**
*Advantages:* The agent takes into account the cost of reaching the goal, as it wants to maximize the utility function, e.g. the passenger that wishes to go to destination A, might want to avoid toll stations and is happy with taking a longer route.
*Disadvantages:* Much harder to implement, computationally demanding, and sometimes infeasible to obtain the optimal choice.