FUNDAMENTALS OF ARTIFICIAL INTELLIGENCE
Exercise 1: Intelligent Agents – Solutions Winter semester 2023/2024
Sebastian Mair

1 Presented Problems

Problem 1.1: Rational Agents

Problem 1.1.1: Which actions does a rational agent select?

Solution: A rational agent selects an action that is expected to maximize its performance measure, given the prior percept sequence and its built-in knowledge.

Problem 1.1.2: Which of these games would a rational agent always win or draw at and why? Is it physically possible to build such an agent?

a. Poker

Solution: No – environment is only partially observable, i.e. you cannot know which action you should take as you do not know the hand of the other players. Apart from that, even if you knew the hand of the other players, it could happen that you never win due to the randomness of the game. As you have to pay the blinds from time to time, you would eventually loose. Thus, a win or draw cannot be guaranteed for poker.

b. Tic-Tac-Toe (Noughts and Crosses)

Solution: Yes – environment is deterministic and fully observable. In Tic-Tac-Toe there is an algorithm that guarantees a win or a draw no matter how your opponent plays. A rational agent would make use of this algorithm. See: `xkcd.com/832`

c. Chess

Solution: Yes[1] – environment is deterministic and fully observable; however, the state space of chess is enormous (10^{43} board positions, 10^{120} possible games). A winning/drawing strategy has not yet been explicitly found and thus it is currently computationally infeasible to construct a rational agent using current techniques.

Problem 1.2: Intelligent Agents

Problem 1.2.1: Suggest performance measures for each of the agents, and which type of agent should be used.

a. GPS route guidance

Solution: Performance measure: e.g. time to reach goal, fuel expended to reach goal, speed limits on roads, avoiding toll routes and low bridges (for Heavy Goods Vehicles), motorways avoided/preferred, traffic jams avoided etc.
Type of agent: Utility-based Agent

b. Kettle (*Wasserkocher*)

Solution: Performance measure: Is water boiled?
Type of agent: Simple Reflex Agent

[1] Since chess has not yet been solved, it is not known whether a winning strategy exists for each player. What we do know, according to `Zermelo's theorem`, is that either a winning strategy exists for one player, or both sides can force a draw at least. See also: `http://en.wikipedia.org/wiki/Solving_chess`

c. Bi-directional Escalator on Munich Underground

Solution: Performance measure: does escalator not stop while transporting someone? Waiting time if escalator is empty, waiting time at the bottom might be weighted stronger.
Type of agent: Model-based Reflex Agent (partial observability is handled by internal state)

Problem 1.2.2: (adapted from *Russell & Norvig*, q. 2.9) Consider the Vacuum Cleaner environment from the lecture notes (slide 11, lecture 2), in which the agent's performance measure awards 3 points after each operation, if a floor was dirty before and is clean afterwards, and penalises 1 point for each movement during operation. It can only perceive the room it is in, e.g. $[A, \text{Dirty}]$. Initially, it is distributed uniformly whether a floor is dirty or not. The floor that the agent starts in is also chosen uniformly.

a. Can a simple reflex agent be rational for this environment?

Solution: No.
Intuitively, the simple reflex agent cannot remember what it did in prior time steps, and can therefore not tell whether it already cleaned the other floor or not. Let us prove this formally in the following.
A rational agent should maximise the expected performance measure given its percept sequence.
Let us investigate how a rational agent would behave:
Let us assume that the agent is starting in floor A. If A is Dirty, the agent should clean it, as this gives a reward of 3. Cleaning A only later would lead to a smaller expected performance measure for sure, as then a movement would be necessary, which is penalised by -1. Thus, the expected performance measure that our rational agent reaches so far is $0.5 * 3 = 1.5$.
Floor A is clean now. Note that the agent does not know whether B is clean or not, as there is no information about B in its percept sequence. Now, the agent can either stay in A forever, giving an expected performance measure of 0, or move to floor B at some point by applying action Right. A rational agent would do the latter; to see this, we calculate the expected performance measure for this case: Let us assume the agent stays in A for n time steps, and then moves to B, giving an immediate reward of -1. The chance of B being Dirty is 0.5; cleaning B would then give a reward of 3. After that, all floors are clean, and there cannot be gained a positive reward anymore; on the other hand, staying on B by applying NoOp all the time gives a reward of 0 for each step. From the percept sequence, the agent knows that all floors are clean, and stays therefore in B forever. Thus, the expected reward of moving to B after n time steps is $0 * n + (-1) + 0.5 * 3 = 0.5$, which is higher than the reward for staying in A forever.
Now, the expected performance measure that our rational agent reaches when starting in A is $1.5 + 0.5 = 2$. The same holds when starting in B (the explanations above work analogously), therefore, the overall expected performance measure is $0.5 * 2 + 0.5 * 2 = 2$.
Let us now assume, our rational agent would be a simple reflex agent. A simple reflex agent can - according to its definition - only consider the current percept for its decisions, but not the whole percept sequence. Therefore, we can define a simple reflex agent for the vacuum cleaner environment by specifying, which actions are taken for the 4 possible percepts $[A, \text{Dirty}]$, $[A, \text{Clean}]$, $[B, \text{Dirty}]$ and $[B, \text{Clean}]$. It still holds, that a rational agent always has to clean the current floor immediately if it is dirty, as explained above. So, our simple reflex agent chooses action *Suck* for the percepts $[A, \text{Dirty}]$ and $[B, \text{Dirty}]$. Now, the simple reflex agent only needs to define which action to choose for the remaining percepts $[A, \text{Clean}]$ and $[B, \text{Clean}]$, i.e. whether to move to the other floor or stay in the current one when seeing the corresponding percept. Table 1 and 2 show the expected performance measures for all possible strategies when starting in A resp. B; each cell in the tables corresponds to a certain strategy of the simple reflex agent, e.g. the upper left cell means that the agent chooses action *Left* for the percept $[B, \text{Clean}]$ and action *Right* for the percept $[A, \text{Clean}]$. Table 3 shows the resulting overall performance measures (i.e. when starting in A or B uniformly). It can be seen, that none of the four strategies can reach the performance measure of 2 that we calculated above for a rational agent. Therefore, a simple reflex agent cannot be rational for the vacuum cleaner environment.

b. What about a reflex agent with state?

Solution: Yes, it can be rational. As explained in the solution for a., the agent only needs to remember that it does not have to go back into a room once it has cleaned it. This can be realized by having a percept sequence longer than just the present.

Problem 1.3: Environments

		$[B, \text{Clean}]$	
		Left	NoOp
$[A, \text{Clean}]$	Right	$-\infty$	$0.5 * 3 + (-1) + 0.5 * 3 = 2$
	NoOp	$0.5 * 3 = 1.5$	$0.5 * 3 = 1.5$

Table 1: Expected performance measure when starting in A

		$[B, \text{Clean}]$	
		Left	NoOp
$[A, \text{Clean}]$	Right	$-\infty$	$0.5 * 3 = 1.5$
	NoOp	$0.5 * 3 + (-1) + 0.5 * 3 = 2$	$0.5 * 3 = 1.5$

Table 2: Expected performance measure when starting in B

		$[B, \text{Clean}]$	
		Left	NoOp
$[A, \text{Clean}]$	Right	$-\infty$	1.75
	NoOp	1.75	1.5

Table 3: Expected performance measure (overall)

Problem 1.3.1: Andrea says: "a game of billiards is deterministic: a player's action is determined by the state of the table and where the ball is."

Bernhard says: "A game of billiards is stochastic, as one player doesn't know what the other player will do."

Catherine says: "A game of billiards is stochastic because it is impossible to know exactly where the ball is and what the shape of the ball and the table are. When the player hits the ball, it might go somewhere else than intended."

Who do you agree with?

Solution: To Bernhard: We can regard the other players as other agents; then the environment is multi-agent and not necessarily stochastic (similar to the deterministic and multi-agent environment of chess, see *Russell & Norvig, 3rd ed.*, p. 43). If we regard the other players not as agents but as part of the environment, then the environment is stochastic.

To Andrea and Catherine: The environment is fully observable and in principle, you should be able to calculate exactly using the laws of physics, where the ball will go under a particular action. However, in practice and in the real world, the table may not be perfectly flat, the balls may be slightly damaged, the player might not be skilled enough to execute a perfect shot, etc. For example, one might say to oneself "I could go for that ball on the other side of the table, but I have a better chance of potting this ball closer to me." This shows that humans consider billiards an (at least partially) stochastic environment.
As such, it is really the agent's perspective (or model of the world) that matters. Perhaps one day we will build an amazing billiards robot that can calculate the exact trajectory of the ball and execute a perfect shot based on a deterministic environment model. Until then, billiards-playing robots will probably use a stochastic environment model. In robot football, see for example:
`sciencedirect.com/science/article/pii/S1474667017321079`

2 Additional Problems

Problem 1.4: Intelligent Agents

Problem 1.4.1: Suggest performance measures for each of the agents, and which type of agent should be used.

a. Bomb disposal agent

Solution: Performance measure: Bomb does not explode.
Type of agent: Goal-based, possibly utility-based if there is a time constraint (to evaluate two different actions in terms of their trade-off between effectiveness and time-efficiency).

b. Weather forecast

Solution: Performance measure: e.g. penalise false forecasts, reward true forecasts, maximise precision & recall for specific weather events.
Type of agent: Goal-based, Utility, Learning

c. Chess/strategy game on clock

Solution: Performance measure: win on time
Type of agent: Goal-based, Reflex Agent with state (due to time constraints, moves could be preprogrammed), Learning

Problem 1.4.2: Think of further agents. For each, propose a performance measure and decide which type of agent can be used.

Solution: You can discuss your solution with fellow students and unclear cases in moodle. However, note that often several solutions exist for the same problem.

Problem 1.4.3: (from *Russell & Norvig*, q. 2.2) Both the *Performance Measure* and the *Utility Function* measure how well an agent is doing. What is the difference between the two?

Solution: Performance Measure is a specification from the designer or programmer (external to the agent) to specify what the agent **should** do. If an agent always acts to maximise or achieve the Performance Measure, this is a **Rational Agent**.

Not all agents have a Utility Function, for example, reflex agents do not. The Utility Function is used internally by the robot itself to evaluate the best course of action is to **optimally** achieve/maximise the Performance Measure(s), given its perceived state.

Problem 1.5: Environments

Problem 1.5.1: (adapted from *Russell & Norvig* 3^{rd} *ed.*, q. 2.4) For each of the following activities, give a PEAS description of the task environment and characterise it in terms of the properties listed in slides 19-25 from Lecture 2.

1. Playing football.

Solution:

Performance Measure	Environment	Actuators	Sensors
Scoring more goals than the opponent, as few players sent off/yellow carded as possible	Football field	Legs, feet, head, chest, hands (for goalkeeper, or if you're Diego Maradona)	Eyes, ears, balance, pain receptors, touch (skin)

partially observable, multi-agent, stochastic, sequential, continuous, dynamic, unknown

2. Repair robot for subsea cables.

   Solution:

Performance Measure	Environment	Actuators	Sensors
Cable repaired, good quality of repair, minimum fuel expenditure, minimum time	The ocean	Propulsion system, robot arm for repairs	Camera, ultrasound, radar, accelerometer, GPS etc.

   partially observable, single-agent, stochastic, sequential, continuous, dynamic, unknown.

3. Price query for a product on the internet.

   Solution:

Performance Measure	Environment	Actuators	Sensors
The best priced product	Internet	Search queries	Automatic text analysis

   partially observable, single-agent, stochastic, episodic, continuous, dynamic, unknown

4. Bidding on an item at an auction.

   Solution:

Performance Measure	Environment	Actuators	Sensors
Obtaining the item (if wanted), price paid	Auction House / eBay	Placing a bid (vocally, by telephone, electronically)	Eyes, ears

   partially observable, multi-agent, deterministic (or stochastic, in the event of a tie, where two agents bid at the same time and the winner is decided by chance), episodic, continuous, dynamic, known.