Autonomy By Design

Kennis en Redeneren Theorie

Naam : Hussin Almoustafa
Studentnummer : 1776495

February 21, 2023

1 Opdracht 1 Agent-based simulatie

1.1 2.4

For each of the following assertions, say whether it is true or false and support your answer with examples or counterexamples where appropriate.

1. An agent that senses only partial information about the state cannot be perfectly rational.

False. In a partially observable environment, an agent may not have enough information to make a perfectly rational decision, since it lacks a complete description of the state of the world. For example, consider a game of chess where one player can only see the pieces on their half of the board. In this case, the agent cannot make perfectly rational moves, since it does not have complete information about the state of the game.

2. There exist task environments in which no pure reflex agent can behave rationally.

True. A pure reflex agent cannot act logically in some task settings, such as chess, since it lacks the ability to reason about the state of the world. In the game of chess, for instance, a pure reflex AI that only makes random moves cannot be regarded as reasonable.

3. There exists a task environment in which every agent is rational.

False. Because the criteria of rationality vary depending on the particular task environment and performance metric being used, there is no task environment in which every agent is rational. For instance, in the game of chess, various agents with different tactics could exist; depending on the specific job environment, some of these agents' strategies may be more reasonable than others.

4. The input to an agent program is the same as the input to the agent function.

False. The input for an agent function may differ from the input for an agent program. While the present percept is often the input to the agent function, the input to the agent program may also contain other internal and external variables, such as sensor readings and historical actions.

5. Every agent function is implementable by some program/machine combination.

True. Any program or machine combination can implement any agent function, provided the software is made to follow the rules outlined by the agent function. This is so because all a program is is a collection of instructions that a machine can follow to generate the expected goal.

6. Suppose an agent selects its action uniformly at random from the set of possible actions. There exists a deterministic task environment in which this agent is rational.

False. A uniformly random agent can sometimes be rational in a deterministic job setting. A uniformly random agent, on the other hand, makes judgments based on chance and cannot be said to be acting rationally. In a deterministic environment, the agent must make decisions based on the state of the world.

7. It is possible for a given agent to be perfectly rational in two distinct task environments.

True. If an agent's behaviour is appropriate for each task environment and the performance metrics are different, it is possible for the agent to behave exactly rationally in two different task contexts. So long as it uses a new strategy for each game, a perfectly rational chess player may likewise be completely rational in the game of checkers.

8. Every agent is rational in an unobservable environment.

False. Since an agent needs knowledge of the state of the environment in order to make rational judgments, not all agents are rational in an unobservable environment. An agent may not have all the information necessary to make a rational choice in an unobservable environment, and so cannot be regarded as fully rational.

9. A perfectly rational poker-playing agent never loses.

False. Even the most logical poker player could lose if they come up against someone who is more skilled or lucky. A poker game's outcome is determined by a variety of variables, including both players' strategies and the luck of the cards dealt.

1.2 2.5

For each of the following activities, give a PEAS description of the task environment and characterize it in terms of the properties listed in Section

- Playing soccer.
- Exploring the subsurface oceans of Titan.
- Shopping for used AI books on the Internet.
- Playing a tennis match.
- Practicing tennis against a wall.
- Performing a high jump.

- Knitting a sweater.
- Bidding on an item at an auction.

Agent Type	**Performance Measure**	**Environment**	**Actuators**	**Sensors**
Playing soccer	Winning or losing the game, scoring goals	A soccer field with two teams of players	Legs, feet	Eyes, ears

Agent Type	**Performance Measure**	**Environment**	**Actuators**	**Sensors**
Exploring the subsurface oceans of Titan	Collecting data about the subsurface ocean and its environment	The subsurface ocean of Titan and its environment	A probe or robot	Various sensors, such as cameras and scientific instruments

Agent Type	**Performance Measure**	**Environment**	**Actuators**	**Sensors**
Shopping for used AI books on the Internet	Finding the best deal on used AI books	Online bookstores, Internet	Browsing, clicking, typing	Computer screen, keyboard

Agent Type	**Performance Measure**	**Environment**	**Actuators**	**Sensors**
Playing a tennis match	Winning or losing the match, scoring points	A tennis court with two players	Arms, racket	Eyes, ears

Agent Type	**Performance Measure**	**Environment**	**Actuators**	**Sensors**
Practicing tennis against a wall	Improving tennis skills	A tennis court with a wall	Arms, racket	Eyes

Agent Type	**Performance Measure**	**Environment**	**Actuators**	**Sensors**
Performing a high jump	Clearing a high bar	A track and field with a high jump bar	Legs, body	Eyes

Agent Type	**Performance Measure**	**Environment**	**Actuators**	**Sensors**
Knitting a sweater	Creating a knitted sweater	A place to knit, such as a chair or couch	Hands	Pattern, yarn

Agent Type	Performance Measure	Environment	Actuators	Sensors
Bidding on an item at an auction	Winning the item at a certain price or not winning it	An auction site or room	Bidding, typing	Computer screen, keyboard

2 Opdracht 2 Propositielogica-agent

2.1 7.8

Which of the following are correct?

1. $False \models True$

True

2. $True \models False$

False

3. $(A \wedge B) \models (A \Leftrightarrow B)$

True

4. $A \Leftrightarrow B \models A \vee B$

False

5. $A \Leftrightarrow B \models \neg A \vee B$

True

6. $(A \wedge B) \Rightarrow C \models (A \Rightarrow C) \vee (B \Rightarrow C)$

True

7. $(C \vee (\neg A \wedge \neg B)) \equiv ((A \Rightarrow C) \wedge (B \Rightarrow C))$

True

8. $(A \vee B) \wedge (\neg C \vee \neg D \vee E) \models (A \vee B)$

True

9. $(A \vee B) \wedge (\neg C \vee \neg D \vee E) \models (A \vee B) \wedge (\neg D \vee E)$

True

10. $(A \vee B) \wedge \neg(A \Rightarrow B) is satisfiable$

True

11. $(A \Leftrightarrow B) \wedge (\neg A \vee B)$is satisfiable

True

12. $(A \Leftrightarrow B) \Leftrightarrow C$ has the same number of models as $(A \Leftrightarrow B)$for any fixed set of proposition symbols that includes A, B, C

True

2.2 Deel 2

Voorbeeld en ontologie van een agent met behulp van figuur 8.1 uit het boek van Norvig en Russel.

Let's look at an example of an agent that analyzes a restaurant domain using first-order logic. The agent's ontology comprises relations like "likes," "serves," and "isLocatedIn," as well as objects like "customers," "restaurants," and "dishes." Epistemological commitments held by the agent include assertions like "Customer X enjoys Dish Y," "Restaurant Z serves Dish Y," and "Restaurant Z is in City W."

The AI can articulate increasingly complicated beliefs and draw conclusions from them using first-order logic. To illustrate the idea that "If a customer enjoys a meal, and the restaurant serves that dish, then the consumer would like to go to that restaurant," the agent might use the following example:

$((likes(x, y) serves(z, y) likesToGo(x, z)))xyz$

The agent can then create new beliefs by using logic and inference rules, such as "If Customer X enjoys Dish Y and Restaurant Z serves Y, then Customer X would prefer to go to Restaurant Z." The agent's initial assumption and the logical principle of modus ponens lead to this inference.

By making inferences and coming to conclusions based on its ontology and beliefs, the agent can describe and reason about complicated domains using first-order logic in this manner. For many AI applications, the agent must be able to reason systematically and rigorously, which is made possible by the use of formal languages like first-order logic.

3 Opdracht 3 Predikaatlogica-agent

3.1 8.10

This exercise uses the function MapColor and predicates In(x,y) , Borders(x,y) , and Country(x) , whose arguments are geographical regions, along with constant symbols for various regions. In each of the following we give an English sentence and a number of candidate logical expressions. For each of the logical expressions, state whether it (1) correctly expresses the English sentence; (2) is syntactically invalid and therefore meaningless; or (3) is syntactically valid but does not express the meaning of the English sentence.

3.2 1. Paris and Marseilles are both in France

.

a. In$(Paris \wedge Marseilles, France)$.

(2) Syntactically invalid. Cannot use conjunction inside a term.

b. In $(Paris, France) \wedge In(Marseilles, France)$
(1) Correct.

c. $In(Paris, France) \vee In(Marseilles, France)$
(3) Incorrect. Disjunction does not express “both.

3.3 2. There is a country that borders both Iraq and Pakistan

.

a. $\exists c Country(c) \wedge Border(c, Iraq) \wedge Border(c, Pakistan)$
(1) Correct.

b. $\exists c Country(c) \Rightarrow [Border(c, Iraq) \wedge Border(c, Pakistan)]$
(3) Incorrect. Use of implication in existential.

c. $[\exists c Country(c)] \Rightarrow [Border(c, Iraq) \wedge Border(c, Pakistan)]$
(2) Syntactically invalid. Variable c used outside the scope of its quantifier.

d.$\exists c Border(Country(c), Iraq \wedge Pakistan)$
(2) Syntactically invalid. Cannot use conjunction inside a term.

3.4 3. All countries that border Ecuador are in South America

a. $\forall c Country(c) \wedge Border(c, Ecuador) \Rightarrow In(c, SouthAmerica)$
(1) Correct.

b. $\forall c Country(c) \Rightarrow [Border(c, Ecuador) \Rightarrow In(c, SouthAmerica)]$
(1) Correct

c. $\forall c[Country(c) \Rightarrow Border(c, Ecuador)] \Rightarrow In(c, SouthAmerica)$
(3) Incorrect

d. $\forall c Country(c) \wedge Border(c, Ecuador) \wedge In(c, SouthAmerica)$
(3)Incorrect

3.5 4. No region in South America borders any region in Europe

.

a. $\neg[\exists c, d In(c, SouthAmerica) \wedge In(d, Europe) \wedge Borders(c, d)]$
(1) Correct

b. $\forall c, d[In(c, SouthAmerica) \wedge In(d, Europe)] \Rightarrow \neg Borders(c, d)]$
(1) Correct

c. $\neg\forall c In(c, SouthAmerica) \Rightarrow \exists d In(d, Europe) \wedge \neg Borders(c, d)$
(3) Incorrect

d. $\forall c In(c, SouthAmerica) \Rightarrow \forall d In(d, Europe) \Rightarrow \neg Borders(c, d)$
(1) Correct

3.6 5. No two adjacent countries have the same map color

a. $\forall x, y \neg Country(x) \vee \neg Country(y) \vee \neg Borders(x, y) \vee \neg(MapColor(x) = MapColor(y))$
(1) Correct

b. $\forall x, y(Country(x) \wedge Country(y) \wedge Borders(x, y) \wedge \neg(x = y)) \Rightarrow \neg(MapColor(x) = MapColor(y))$
(1) Correct

c. $\forall x, y Country(x) \wedge Country(y) \wedge Borders(x, y) \wedge \neg(MapColor(x) = MapColor(y))$
(3) Incorrect

d. $\forall x, y(Country(x) \wedge Country(y) \wedge Borders(x, y)) \Rightarrow MapColor(x \neq y)$
(2) Syntactically invalid

3.7 Ontologie in predikaatlogica voor Wumpus World

A formal definition of a domain that outlines the concepts and connect ons between them is called an ontology. In this instance, the Wumpus World's ontology comprises details on the places in the world, the things that are found there, and the acts that agents are capable of.

The following predicates are part of the Wumpus World ontology:

"At(Ag, L)" denotes the presence of the agent "Ag" at the specified location.

Pit(L) indicates that there is a pit at location "L."

"Wumpus(L)" indicates the presence of a Wumpus at location "L."

"Glitter(L)" denotes the presence of glitter at location "L." (gold).

"Breeze(L)" indicates that there is a well close by because a breeze is present at place "L".

"Stench(L)" denotes the presence of a smell at place "L," indicating the presence of a Wumpus nearby.

These predicates allow knowledge and information about the Wumpus World to be stated in a formal language, which makes it simple for an agent to process and evaluate this knowledge and information. Moreover, logical rules that characterize an agent's actions in the Wumpus World can be created using these predicates.

3.8 Beschrijf de toegevoegde waarde van predikaatlogica ten opzichte van propositielogica

Compared to propositional logic, predicate logic is a more flexible and comprehensive logic system. Predicate logic also deals with arguments about objects and relations between them, whereas propositional logic only considers the validity or falsity of simple statements (propositions).

In propositional logic, truth values are assigned to propositions, but specific objects or their qualities cannot be described. For instance, claims regarding individual people, places, or times are not feasible in propositional logic; one can only make general statements like "the sun is shining" or "it is raining."

Predicate logic allows for the formulation of arguments about certain objects and the connections among them. Predicate logic, for instance, allows for the usage of phrases like "John is a teacher" and "John's house is bigger than Mary's house." Because of this, predicate logic is far more suited to expressing complicated concepts and the connections between them.

Predicate logic's ability to express quantification is a significant bonus. Predicate reasoning allows one to make claims like "all dogs are animals" or "at least one white dog exists." Predicate logic can thus be used to articulate generic claims and construct logical arguments.

Predicate logic, as opposed to propositional logic, has the advantage of being able to explain distinct objects and the connections between them, as well as be able to express quantification and generalization. Predicate logic is a considerably more robust and all-encompassing logic system as a result of this than propositional logic.

4 Opdracht 4 Onzekerheid

4.1 13.1 Quantifying Uncertainity

Show from first principles that $P(ab \wedge a) = 1$

$$P(a|a \wedge b) = \frac{P(a \wedge a \wedge b)}{P(a \wedge b)}$$

$$= \frac{P(a \wedge b)}{P(a \wedge b)}$$

$$= 1$$

Using the **Product Rule**

4.2 13.4 Quantifying Uncertainity

Would it be rational for an agent to hold the three beliefs

$$P(A) = 0.4, P(B) = 0.3, and P(A \vee B) = 0.5$$

?

If so, what range of probabilities would be rational for the agent to hold for $A \wedge B$? Make up a table like the one in Figure 13.2, and show how it supports your argument about rationality. Then draw another version of the table where

$$P(A \vee B) = 0.7$$

.

Explain why it is rational to have this probability, even though the table shows one case that is a loss and three that just break even. (Hint: what is Agent 1 committed to about the probability of each of the four cases, especially the case

that is a loss?)

Because they are against the probability axioms, the beliefs $P(A) = 0.4$, $P(B) = 0.3$, **and** $P(AB) = 0.5$ **are not rational for an agent to have. The agent specifically thinks that** $P(AB)$ **is** 0.5, **which denotes that there is at least a** 0.5 **chance that either** A **or** B **will occur. It is not feasible for both** A **and** B **to have probabilities greater than 0.5 since the total of the probabilities of** A **and** B **cannot be more than 1.**

We can create a table showing the potential probabilities for A and B, assuming the agent is rational and adheres to the probability axioms, to better understand this:

	A=0	**A=1**
B=0	**0.3**	**0.2**
B=1	**0.1**	**0.4**

In this table, the agent's belief that $P(A) = 0.4$ **and** $P(B) = 0.3$ **is represented in the first column and row. The belief is that** $P(A \vee B) = 0.5$ **is represented in the sum of the last two entries, which is the probability of either** A **or** B **happening.**

To find the range of probabilities that would be rational for the agent to hold for $A \wedge B$, **we can look at the entries in the table where both** A **and** B **are 1. In this case, the agent is committed to** $P(A \wedge B)$ **being at least 0.1, since that's the probability of both A and B happening. The maximum possible value of** $P(A \wedge B)$ **is 0.4, which is the probability of** A **happening on its own.**

Therefore, a rational agent could hold any belief in the range $0.1 \leq P(A \wedge B) \leq 0.4$, **as long as they also believe that** $P(A) = 0.4, P(B) = 0.3$, **and** $P(A \vee B) = 0.5$.

Let's now explore the scenario in which $P(AB) = 0.7$. **We can create a new table to illustrate why this might be reasonable:**

	A=0	**A=1**
B=0	**0.0**	**0.3**
B=1	**0.4**	**0.3**

The likelihood of either A **or** B **occurring is shown in this table as the sum of the last two entries. According to the agent, this value is** $P(AB) = 0.7$. **The agent continues to hold that the values shown in the first column and row,** $P(A) = 0.4$ **and** $P(B) = 0.3$, **are accurate.**

As there is no item where both A and B are 1, this table demonstrates that the agent is committed to $P(AB)$ being 0.0. A or B must have a higher probability than they did in the previous table, according to the agent, who also thinks that the sum of the last two entries is 0.7. As it yields the lowest probability, the agent specifically thinks that $P(A) = 0.4$ and $P(B) = 0.3$ is the least likely situation for $A \vee B$.

Because the agent is committed to $P(AB) = 0.7$, they must alter their views about the probabilities of A and B in accordance with this probability, even though the table shows one instance that is a loss and three that just break even. In order to meet the requirement that $P(AB) = 0.7$, the agent is specifically willing to accept the trade-off between the probability loss for the scenario where A and B both occur and the probability gains for the scenario where just A or only B occur.

To put it another way, the agent is prepared to accept a certain amount of probability in the event that A and B both occur in order to raise the chance that at least one of A or B occurs, which is the desired outcome denoted by $P(AB) = 0.7$. The agent can be said to have a reasonable belief in $P(AB) = 0.7$ if they think the trade-off is worthwhile.

Because there is no entry in the table for the scenario where both A and B occur, it should be noted that the agent is not bound to any certain probability in this situation. The agent's conviction that P(AB)=0.7 is determined only by the trade-off between the other scenarios and the probability loss for the scenario in which both A and B occur.

4.3 Een reeks van states in de Wumpus World met daarbij de kennis met betrekking tot onzekerheid, geüpdate aan de hand van de stelling van Bayes

Below I give an example of a series of states in the Wumpus World and how the knowledge regarding uncertainty can be updated using Bayes' theorem.

Step 1: The agent is in location (1,1) and has no information about the world. The agent has the following prior probabilities:

$P(W1) = \mathbf{0.2}$ **(probability there is a Wumpus in the world)**

$P(P1) = 0.1$ (probability there is a kernel in the world)
$P(G1) = 0.1$ (probability there is a gold piece in the world)

Step 2: The agent senses a stink, which means there may be a Wumpus in an adjacent location. The agent has the following likelihoods:

$P(S1|W1) = 0.6$ (probability of detecting odor if there is a Wumpus in the adjacent location)
$P(S1|\neg W1) = 0.05$ (probability of observing a smell if there is no Wumpus in the adjacent location)

Now the agent can calculate the probability that there is a Wumpus in the world given the perception of the stench:

$P(W1|S1) = P(S1|W1) * P(W1)/(P(S1|W1) * P(W1) + P(S1|\neg W1) * P(\neg W1))$
= 0.6 * 0.2 / (0.6 * 0.2 + 0.05 * 0.8)
= 0.789

The probability that there is no Wumpus in the world, given the perception of the stench, is then:

$P(\neg W1|S1) = 1$ - $P(W1|S1)$
= 1 - 0.789
= 0.211

Step 3: The agent moves to location (2,1) and feels no smell or draft, meaning there is no wumpus or pit in an adjacent location. The prior probabilities remain unchanged.

Step 4: The agent moves to location (2,2) and senses a draft, which means there may be a pit in an adjacent location. The agent has the following likelihoods:

$P(B2|P2) = 0.6$ (probability of detecting a breeze if there is a pit in the adjacent location)
$P(B2|\neg P2) = 0.05$ (probability of detecting a breeze if there is no pit in the adjacent location)

Now the agent can calculate the probability that there is a kernel in the world given the observation of the breeze:

$P(P2|B2) = P(B2|P2) * P(P2)/(P(B2|P2) * P(P2) + P(B2|\neg P2) * P(\neg P2))$
= 0.6 * 0.1 / (0.6 *0.1 + 0.2 * 0.9) = 0.24

Here $P(P2)$ is the initial probability that there is a kernel in $P2$, so 0.1, and $P(\neg P2)$ is the initial probability that there is no kernel in $P2$, so 0.9.

$P(B2|P2)$ is the probability that there is a breeze in $P2$ given that there is a pit in $P2$, this is 0.6.

$P(B2|\neg P2)$ is the probability that there is a breeze in $P2$ given that there is no pit in $P2$, this is 0.2.

With this information, the agent can therefore conclude that there is a 0.24 probability that there is a pit in $P2$, given the observation of the breeze in $P2$.