FUNDAMENTALS OF ARTIFICIAL INTELLIGENCE

Exercise 4: Logical Agents - Problems Winter Semester 2022/2023

Adrian Kulmburg

Problem 4.1: Model, satisfaction relation, and entailment

(*Taken from [3] question 7.4*) Which of the following statements are correct? Show its correctness by reasoning about the models satisfying each sentence.

1. $False \models True$
2. $True \models False$
3. $(A \wedge B) \models (A \Leftrightarrow B)$
4. $(A \Leftrightarrow B) \models (A \vee B)$
5. $(A \Leftrightarrow B) \models (\neg A \vee B)$

Problem 4.2: Validity, satisfiability, and unsatisfiability

(*Exercise is adapted from [3] question 7.10.*) Recall first the definition of *validity* and *satisfiability*.

Problem 4.2.1: Prove the following two metatheorems:

1. Sentence α is valid if and only if $\alpha \equiv True$,
2. Sentence α is unsatisfiable if and only if $\alpha \equiv False$.

Problem 4.2.2: Show whether each of the following sentences is valid, satisfiable, or unsatisfiable by using the two metatheorems above, the standard logical equivalences from the lecture, and the following three logical equivalences:

$$\alpha \vee \neg\alpha \equiv True \tag{1}$$
$$\alpha \wedge \alpha \equiv \alpha \tag{2}$$
$$\alpha \vee \alpha \equiv \alpha \tag{3}$$

1. $Smoke \Rightarrow Smoke$
2. $(Smoke \Rightarrow Fire) \Rightarrow (\neg Smoke \Rightarrow \neg Fire)$
3. $Smoke \vee Fire \vee \neg Fire$
4. $(Fire \Rightarrow Smoke) \wedge Fire \wedge \neg Smoke$

Problem 4.3: Knights and Knaves

(*Proof by inference rule.*) The following puzzle is taken from [1]. Suppose we are in on island with two types of inhabitants: "knights" who always tell the truth, and "knaves" who always lie.

> According to this problem, three of the inhabitants – A, B and C – were standing together in the garden. A stranger passed by and asked A, "Are you a knight or a knave?". A answered, but rather indistinctly, so the stranger could not make out what he said. The stranger then asked B, "What did A say?". B replied, "A said that he is a knave". At this point the third man, C, said "Don't believe B; he's lying!". The question is, what are B and C?

Model this logic puzzle by introducing three atomic propositions A, B, and C with intended interpretation that A, B, and C are knights.

Problem 4.3.1: How can you formalise the sentence *A says that B is a knight*?

Problem 4.3.2: Assume that *Remark* represents what a person says and that we can represent it using propositional logic[1]. Additionally, assume that P could either be A, B, or C. From the previous problem, can you generalise the method to model the sentence "person P says (or replies) *Remark*"?

Problem 4.3.3: Model the following facts which are taken from the puzzle:

1. B replies, "A said that he is a knave".
2. C says, "Don't believe B; he's lying!".

Problem 4.3.4: By using the following logical equivalences

$$(X \Leftrightarrow \neg X) \equiv False \tag{4}$$
$$(X \Leftrightarrow False) \equiv \neg X \tag{5}$$

and the following deduction (inference) rule

$$\frac{P \Leftrightarrow Q \qquad Q}{P}$$

deduce what B and C are.

Problem 4.4: Superman does not exist

(*Proof by resolution.*) The following text is taken from [2].

> If Superman were able and willing to prevent evil, he would do so. If Superman were unable to prevent evil, he would be impotent; if he were unwilling to prevent evil, he would be malevolent. Superman does not prevent evil. If Superman exists, he is neither impotent nor malevolent. Therefore, Superman does not exist.

Assume that we use the following variables and their meaning:

$$\begin{array}{rcl} A & : & \text{Superman is able to prevent evil.} \\ W & : & \text{Superman is willing to prevent evil.} \\ I & : & \text{Superman is impotent.} \\ M & : & \text{Superman is malevolent.} \\ P & : & \text{Superman prevents evil.} \\ E & : & \text{Superman exists.} \end{array}$$

[1] This is a mouthful way to say that *Remark* is a propositional logic sentence.

Problem 4.4.1: By using the variables defined above, and assuming that we want to prove "Superman does not exist", formalise the rest of the facts from the text.

Problem 4.4.2: Identify which sentences belong to the knowledge base KB, and which sentence we want to deduce. Recall the resolution approach for propositional logic. What shall we do next?

Problem 4.4.3: Prove diagramatically with the resolution approach that "Superman does not exist".

Problem 4.5: Completeness and soundness

Recall first the definition of *completeness* and *soundness*:

> **Completeness:** An inference algorithm is **complete** if and only if for every entailed sentence $KB \models \alpha$, the inference algorithm will always be able to derive it.
>
> **Soundness:** An inference algorithm is **sound** if and only if for every sentence it derives, it is guaranteed that the sentence is entailed $KB \models \alpha$.

Problem 4.5.1: Suppose that we have an inference algorithm which will always be able to derive any given sentence (regardless whether it is entailed or not). Would this inference algorithm be complete? Sound?

Problem 4.5.2: Suppose now that we have an inference algorithm which will always be unable to derive any sentence (regardless whether it is entailed or not). Would this inference algorithm be complete? Sound?

References

[1] R. Backhouse. *Algorithmic Problem Solving.* Wiley Publishing, 1st edition, 2011.

[2] D. Gries and F. B. Schneider. *A Logical Approach to Discrete Math.* Springer-Verlag New York, Inc., New York, NY, USA, 1993.

[3] S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach.* Prentice Hall series in artificial intelligence. Prentice Hall, 2010.