



---

## Assignment 2

**Deadline:** 27 June 2021 at 23:59 (German time)

**Late Deadline:** 25 July 2021 at 23:59 (German time)

---

You can hand in this assignment as a group. Please don't forget to write all your names on your solution. Hand in your solution as a single pdf file in the dcms. If you hand in by the regular deadline, you get all the points that you earned on this assignment. If you hand in by the late deadline, your assignment will be corrected and 30% of the points that you got on this assignment will be deducted. You cannot hand in an assignment twice. You need 50% of all assignment points in order to get the exam admission. You can get 18 points in total of which you can get 9 points on this assignment.

---

Imagine that you read the following text by Max Mustermann in a newspaper:

Recently, autonomous systems that make decisions for us shifted into the focus of researchers, columnist, and politicians. Apparently, current autonomous deciders are rather opaque, and it is often demanded to make them 'explainable'. If an autonomous decider is explainable, we are supposed to be able to get the *real* reasons for its output in a comprehensible format at all times. But the demand for explainability is, to be frank, a stupid one, and we should not demand explainability from autonomous deciders. Why so? Human decisions are often taken to be the gold standard for a good decision. But advocates for explainability overlook that a human decision lacks precisely what they demand from autonomous deciders, namely explainability. Maybe, we can ask a human afterwards why they did what they did, but we can never be sure that they answer truthfully. We cannot even be sure that they know the real reasons for their action themselves. It could well be that their brain just made up some reason afterwards that is very different to the true, hidden, subconscious reason for the action in question. But if a human is not explainable, and we are more than happy with human decisions, then we also should not demand explainability from autonomous deciders. Everything else is hypocritical bollocks.

- (a) (3 points) Reconstruct the main argument from the text above as an extended standard form. Be charitable in your reconstruction.
- (b) (3 points) Give a plausible attack of this argument as a valid extended standard form. Additionally, state what exactly you are attacking.
- (c) (3 points) Briefly describe three different ways in which Max Mustermann could plausibly react to your attack if you were in a discussion with him. Assume that both of you adhere to the rules of a proper discussion.