

*This homework is due at 9:30am on April 5, 2024. You can write directly in the available space below, or in a separate file.*

Name: \_\_\_\_\_

The homework focuses on the manuscript *The perils of trial-and-error reward design: misdesign through overfitting and invalid task specifications* by S. Booth et al. in AAAI, 2023. You can find it on the course schedule.

1. Provide a **summary** of the paper by S. Booth et al. You must include clearly labeled sections (each with a bold-faced header) named **Motivation**, **Contribution**, **Methodology**, and **Conclusion**. Note that you should provide at most one page of content this and subsequent questions, so your review will be an abbreviated version compared to Fong's recommendations. (10 pts)



2. Provide a **critique** of the paper by S. Booth et al. You must include at least two, and no more than three, critiques. Each critique must be clearly identified (e.g., prepending the critique with “First critique:” and “Second critique:”) and must be well reasoned using objective, as opposed to subjective, criteria. *You cannot include examples that the authors already mention in the paper.*  
(10 pts)

- 3.** Provide a **synthesis** of the paper by S. Booth et al. You must include one, and no more than two, idea that describes how the work could be further developed. *You cannot include examples that the authors already mention in the paper.* (10 pts)