Review Paper01

Zixuan Xia

1 Overview of the Paper

This paper is mainly about the introduction of the Expected Utility Theory, which is based on an hyperthosis – the Expected Utility Hypothesis (EUH), a theoretical framework explaining how decision-makers make choices under uncertainty. Their central arguments are as follows:

- The Expected Utility Hypothesis provides a unified explanation for both risk-averse and risk-seeking behavior.
- The hypothesis is based on a set of axioms (such as completeness, transitivity, independence, and continuity) that ensure decision-makers can rank different choices and maximize expected utility accordingly.
- While the theory is influential, it is not without flaws, as several paradoxes (e.g., the St. Petersburg Paradox, the Allais Paradox) challenge its applicability in real-world decision-making.

2 Personal Understandings and Questions

2.1 Personal Understandings

The core viewpoints of the Expected Utility Theory is that, assume people are rational identities, they will usually make decisions based on utility.

The paper raises questions about the measurability of utility, particularly highlighting that early economists tended to believe that utility was "immeasurable," whereas modern economics attempts to measure it through experimental methods. This reminds me of the challenge of designing reward functions in reinforcement learning: while we can theoretically define a clear objective function, in practice, the validity and reasonableness of this function are often difficult to determine.

A simple example discussed in the previous reinforcement learning lecture illustrates this issue: Suppose a person has two commuting options—taking the train, which costs 20 CHF and takes 1 hour, or driving, which costs only 10 CHF and takes the same or even less time. Some may find the train more comfortable, as driving requires a high level of concentration and might negatively

impact their productivity for work or study afterward. Others may feel that driving while listening to music helps them relax.

Therefore, I believe that defining an appropriate and universally valid reward signal remains an open question in both economics and reinforcement learning. Published in 1952, this paper had a profound impact on economics and decision

theory. Today, the principles of the Expected Utility Hypothesis have permeated reinforcement learning (RL) and serve as one of its foundational concepts:

- In RL, an agent seeks to maximize a utility function, typically the expected reward. This aligns with the core principle of the Expected Utility Hypothesis, where individuals (or agents) make choices that maximize expected gains.
- **Policy optimization in reinforcement learning** is often grounded in maximizing expected cumulative reward, a concept closely related to expected utility theory.

2.2 Personal Questions

- The paper discusses multiple paradoxes and modifications to the Expected Utility Hypothesis. I am unclear on how these paradoxes are mathematically addressed, particularly in reinforcement learning, where human risk aversion needs to be formally modeled.
- How is the **independence axiom** specifically applied in reinforcement learning? Are there RL algorithms that explicitly violate this axiom?
- How does the measurability issue of the Expected Utility Hypothesis impact empirical research in economics? What insights does this offer for reward function design in reinforcement learning?