Inferring the Decisions from Utility

Jinfan He
Peking University
2200012979@stu.pku.edu.cn

Abstract

Utility is deeply intertwined with the human decision-making process. However, because it is recessive, it is often difficult to measure directly. In this paper, we first introduce the utility-related hypothesis, then we introduce a model to explain the human decision-making process that fits the existing hypothesis, and propose some improvements based on cognitive cost theory. We conjecture the process by which humans learn the utility of others.Then we introduce the current general approach to learning utilities and point out some of the problems with this approach. Finally, we propose a prediction mechanism for the possible future decisions of others.

1 Introduction

1.1 Utility and utilitarianism

Utility is defined as the intrinsic human assessment of the value of things, as distinct from their actual value. Different individuals may attribute vastly different utilities to the same thing, and this utility is implicit, making measuring someone else's utility an extremely challenging task. However, we do need methods to gauge the utility of others. According to utilitarianism, humans follow the principle of maximizing expected utility in the decision-making process. If we aim to analyze someone else's decision-making process, we first need to understand their expected utility.

1.2 Bounded rationality theory

However, utilitarianism is not the sole explanation for human decision-making. Herbert Simon proposed the theory of bounded rationality[3], suggesting that humans follow a principle of satisfice in decision-making, where choices that achieve a certain level of utility are considered equivalent. But this can also be interpreted as comparing the cognitive or time costs required for these choices against the anticipated benefits.

2 Models of human decision-making

A general model for human decision-making is as follows[2]:

$$P(c|r) = \frac{exp(\beta \cdot r(c))}{\sum_{c' \in C} exp(\beta \cdot r(c'))}$$

$$\max_{P} \mathbf{E}_{c \sim P}[r(c)]$$

Here, c represents a choice, r(c) represents the reward of choice c, P(c|r) represents the probability distribution P of humans on the decision set C, and β signifies the level of rationality when making decisions. When $\beta \sim \infty$, humans naturally choose the option with the highest expected value. The lower the β, the less concerned humans are with the outcome. When $\beta = 0$, the decision-making process is entirely random.

However, this procedure does not account for the cognitive cost of estimation. We know that humans follow the principle of achieving tasks with minimal cognitive cost. We can think of the human decision-making process as a sequence of decisions, where before each decision there is an estimate of the additional expected benefit versus cognitive and time costs. If the required cognitive and time costs exceed the anticipated benefits, a choice is made and the decision process ends. This process is entirely rational, hence we can consider β as infinite. Here, any irrational decision-making by humans is due to either an excessively high cognitive cost or a significantly different utility compared to others.

3 Learning human utility

Now that it is known how to predict other people's decisions through utility, the question becomes how to understand other people's utility and predict the cognitive and temporal costs. The simplest idea is to assume that other people have the same utility as we do, turning the process into a mental simulation - using your value function and decision-making process to model other people's decisions. However, it is clear that everyone has different utilities, leading to significant errors in such predictions. Thus, we aim to learn the utility of others and construct a value assessment system for each individual, which requires substantial cognitive resources. Thus, templates are used to group similar people, assuming they have approximately similar value assessment systems.

Utility is influenced by many factors, such as individual physical needs, mental needs, and cognitive understanding of things. However, these factors are still challenging to perceive. Therefore, common methods rely on inferring possible utilities through external behavioral patterns, like inverse reinforcement learning (IRL) [1]. IRL alternates between two processes: one phase uses action trajectories to infer a hidden reward function, and the other uses reinforcement learning based on the inferred reward function to learn an imitation policy. The basic principle of IRL is to ensure that any action decision different from the action trajectory incurs the maximum possible loss.

However, multiple reward functions may fit a single action trajectory and a reward function may also produce multiple optimal action trajectories, so it may need to consider multiple action trajectories to determine a reward function. Also, the learned reward function is less transferable, as different beliefs in different situations yield drastically different reward functions. For example, water in a desert would have a higher value than in the usual case. Essentially, the learned reward function often represents a person's representation under the influence of physical needs, mental needs, and cognition in a given situation, and is not invariant.

4 Future Direction

So, can we still predict other people's decisions through utility? This may require learning more fundamental information, such as inferring the underlying beliefs, personality traits, physical states of others, then further understanding how other people's utility is affected by these factors, and finally understanding to what extent other people's decision-making processes are influenced by cognitive costs-how deeply they contemplate decision-making choices.

5 Conclusion

This essay begins by introducing the definition of utility and highlighting its crucial role in predicting human decision-making processes. It covers two theories, utilitarianism and bounded rationality, and presents a decision model that can explain both. Recognizing that bounded rationality may be due to cognitive cost, we proposed a new latent decision model. Regarding the learning process of utility, an initially simple idea is suggested: transfer one's own utility and decision mechanism to others, akin to mental simulation. However, due to the significant error caused by individual differences, a template model is proposed to reduce the cognitive resources needed to learn the utility of others. Subsequently, Inverse Reinforcement Learning (IRL) is introduced, which learns utility through external behavioral patterns. Potential issues with this approach are also analyzed. Finally, conjectures are made about possible future mechanisms for predicting other people's decisions based on utility.

References

[1] Stuart Russell Andrew Y Ng. Algorithms for inverse reinforcement learning. *ICML*, 1:2, 2000. 2

[2] Smitha Milli Hong Jun Jeon and Anca Dragan. Reward-rational (implicit) choice: A unifying formalism for reward learning. *Advances in Neural Information Processing Systems*, 33:4415–4426, 2020. 1

[3] Herbert A Simon. A behavioral model of rational choice. *The quarterly journal of economics*, 69 (1):99–118, 1955. 1