Unveiling Human Utility: Approaches and Challenges

Du Jiang
School of Intelligence Science and Technology
Peking University
dujiang@pku.edu.cn

Abstract

This essay examines prominent methods for understanding human utility. Each method offers unique insights, from leveraging preferences to learning from human feedback and observing behaviors. Challenges in generalization, data intensity, and interpretational complexities are identified. Balancing these considerations is crucial in constructing robust computational frameworks for human utility estimation, fostering AI systems aligned with human values and preferences.

1 Introduction

Human decision-making is inherently guided by utility, representing individuals' subjective preferences and values. However, quantifying and representing human utility poses a significant challenge due to its inherent subjectivity, variability across individuals, and the absence of a universally meaningful unit of measure. To address this complex task, several crucial methods stand out in learning and representing human utility, each offering distinct advantages and challenges.

2 Approaches and Challenges

2.1 Preference-Based Reinforcement Learning

One prominent approach involves PBRL[7] (preference-based reinforcement learning). The method leverages observed preferences to discern underlying utility functions. This approach capitalizes on collecting data related to choices, rankings, or comparisons made by individuals. By doing so, it adapts to the unique preferences of different individuals, allowing for personalized utility representations. PBRL offers a structured means to learn utility without the need for explicitly defined utility functions.

This kind of approach can accommodate diverse and individualistic preferences, offering personalized utility representations. Also, the method provides a systematic framework to understand utility without requiring explicit utility function specifications. It can incorporate various sources of preference data, enabling a comprehensive understanding of individual choices.

However, reliance on collected preference data might restrict the generalizability of learned utility functions to broader contexts. Capturing the entirety of human preferences and utility might be challenging through limited preference-based data collection.

2.2 Deep Reinforcement Learning from Human Preferences

Currently, some approaches utilize deep learning techniques to learn from human-provided feedback or demonstrations[2, 3]. They aim to map human preferences or demonstrations to a utility function, allowing algorithms to learn from human behavior and preferences.

When trained with diverse and comprehensive data, deep reinforcement learning models can generalize across similar tasks or scenarios. The models can dynamically adjust and learn from feedback or demonstrations, improving accuracy over time. These algorithms can assimilate various forms of human-provided data such as demonstrations, rankings, or preferences, and thus help integration with diverse data types.

The most significant disadvantage of such approaches is data intensity. Requirements of significant amounts of human-provided feedback or demonstrations make the methods resource-intensive and time-consuming. Also, the performance might suffer if the training data lacks diversity or fails to represent the full spectrum of human preferences.

2.3 Inferring Human Utilities from Observational Data

Observational inference involves extracting information about human utilities from observed behaviors, actions, or choices made by individuals[8]. It aims to uncover implicit preferences by analyzing behavioral patterns.

The method avails of non-intrusive data collection. It can derive insights from observed behavior without explicitly requiring individuals to state their preferences. Observational data can reveal hidden or implicit aspects of human preferences that might not be explicitly stated.

Nevertheless, inferring utility from observed behavior requires careful analysis and interpretation, which can be subject to biases or misinterpretations. Also, behavioral observations might not capture the full richness and complexity of human utility functions, leading to incomplete representations.

2.4 Active Preference-Based Learning of Reward Functions

This method employs an iterative process where the algorithm actively seeks informative data points to refine learned utility functions[6, 1]. By prioritizing informative samples, it aims to efficiently collect data and enhance the accuracy of the learned utility function.

Prioritizing informative samples reduces the need for extensive data collection, focusing on the most valuable information. Also, the method allows for refinement over time, continually enhancing the utility function's accuracy. By selecting the most informative data points, the method maximizes the utility of collected data, saving resources.

While more efficient, it might still demand substantial human input or computational resources for selecting informative samples. Efficiency hinges on the quality and informativeness of the selected data points, requiring careful selection strategies.

2.5 Bayesian Psychology and Human Rationality

The Bayesian approach incorporates probabilistic models to represent uncertainty in human preferences[5, 4]. It provides a framework for capturing and updating beliefs about human utilities over time based on observed data.

The method utilizes probabilistic models to capture the uncertainty inherent in human preferences, allowing for nuanced representation. Adaptive learning facilitates continuous updates of beliefs about human utilities as new data becomes available. Moreover, the flexible framework allows for the incorporation of prior knowledge and the iterative refinement of beliefs.

As for the defects, Bayesian approaches might pose scalability challenges when applied to large datasets due to computational demands, and the choice of prior distributions could significantly impact the learned utility functions, potentially introducing biases.

3 Discussion

Each of these approaches offers valuable insights into understanding and representing human utility. Combining these methods could potentially yield a more comprehensive understanding of diverse utility functions across individuals. However, challenges persist in balancing data collection, generalization, and computational efficiency. Developing a robust computational framework for human utility estimation requires navigating these challenges while embracing the strengths of each approach. Ultimately, a holistic understanding of human utility will pave the way for more effective AI systems that align with human values and preferences.

References

[1] Erdem Biyik, Nicolas Huynh, Mykel Kochenderfer, and Dorsa Sadigh. Active Preference-Based Gaussian Process Regression for Reward Learning. In *Proceedings of Robotics: Science and Systems*, Corvalis, Oregon, USA, July 2020. doi: 10.15607/RSS.2020.XVI.041. 2

[2] Paul Christiano, Jan Leike, Tom Brown, Miljan Martic, Shane Legg, and Dario Amodei. Deep reinforcement learning from human preferences. 06 2017. 1

[3] Changyeon Kim, Jongjin Park, Jinwoo Shin, Honglak Lee, Pieter Abbeel, and Kimin Lee. Preference transformer: Modeling human preferences using transformers for RL. In *The Eleventh International Conference on Learning Representations*, 2023. URL `https://openreview.net/forum?id=Peot1SFDX0`. 1

[4] Hima Lakkaraju, Jure Leskovec, Jon Kleinberg, and Sendhil Mullainathan. *A Bayesian Framework for Modeling Human Evaluations*, pages 181–189. 06 2015. ISBN 978-1-61197-401-0. doi: 10.1137/1.9781611974010.21. 2

[5] Mike Oaksford and Nick Chater. Précis of bayesian rationality: The probabilistic approach to human reasoning. *Behavioral and Brain Sciences*, 32(1):69–84, 2009. doi: 10.1017/S0140525X09000284. 2

[6] Dorsa Sadigh, Anca D. Dragan, S. Shankar Sastry, and Sanjit A. Seshia. Active preference-based learning of reward functions. In *Robotics: Science and Systems*, 2017. URL `https://api.semanticscholar.org/CorpusID:12226563`. 2

[7] Christian Wirth, Riad Akrour, Gerhard Neumann, and Johannes Fürnkranz. A survey of preference-based reinforcement learning methods. *Journal of Machine Learning Research*, 18(136):1–46, 2017. URL `http://jmlr.org/papers/v18/16-634.html`. 1

[8] Yixin Zhu, Chenfanfu Jiang, Yibiao Zhao, Demetri Terzopoulos, and Song-Chun Zhu. Inferring forces and learning human utilities from videos. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016. 2