Algorithmic Decision-Making in Difficult Scenarios

Christopher B Rauch[1], Ursula Addison[2], Michael Floyd[3], Prateek Goel[1], Justin Karneeb[3], Ray Kulhanek[2], Othalia Larue[2], David Ménager[2], Mallika Mainali[1], Matthew Molineaux[2], Adam Pease[2], Anik Sen[1], JT Turner[3], Rosina Weber[1]

[1]Drexel University
[2]Parallax Advanced Research
[3]Knexus Research Corporation

cr625@drexel.edu, uaddison@parallaxresearch.org, michael.floyd@knexusresearch.com, pg427@drexel.edu, justin.karneeb@knexusresearch.com, ray.kulhanek@parallaxresearch.org, othalia.larue@parallaxresearch.org, david.menager@parallaxresearch.org, mm5579@drexel.edu, matthew.molineaux@parallaxresearch.org, adam.pease@parallaxresearch.org, as5867@drexel.edu, jt.turner@knexusresearch.com, rw37@drexel.edu

Abstract

We present an approach to algorithmic decision-making that emulates key facets of human decision-making, particularly in scenarios marked by expert disagreement and ambiguity. Our system employs a case-based reasoning framework, integrating learned experiences, contextual factors, probabilistic reasoning, domain-specific knowledge, and the personal traits of decision-makers. A primary aim of the system is to articulate algorithmic decision-making as a human-comprehensible reasoning process, complete with justifications for selected actions.

Introduction

The In the Moment (ITM) project, funded by DARPA's Defense Sciences Office, focuses on developing artificial intelligence systems that reflect the key personal attributes of trusted decision-makers and that are capable of complex decision-making in high-pressure scenarios characterized by expert disagreement, uncertainty, and resource constraints.

We have designed a system that uses algorithms to make decisions based on input data and predefined rules and models. This algorithmic decision-maker automates the creation of human-like arguments, selecting decisions supported by justifications that closely align with the attributes of trusted decision-makers. This is facilitated by a set of analytical tools that emulate aspects of human decision-making. The initial focus of the demonstration system will be on small military unit medical triage scenarios in austere environments, with plans to expand its applicability to mass casualty scenarios and other challenging decision domains.

Case-Based Reasoning and Analogy

Analogical reasoning is a cognitive process that uses similarities between different situations or concepts to make predictions or offer explanations. This method enables individuals to derive inferences and establish connections based on prior experiences and knowledge. Its utility extends across various cognitive tasks, including problem-solving, moral decision-making, and commonsense reasoning (Forbus and Hinrichs 2017).

Case-based reasoning incorporates elements of analogical reasoning by adapting solutions from past similar cases to address new problems. The process entails identifying a new problem, finding analogous cases, discerning distinctions, and modifying prior solutions to fit the current problem (Richter and Weber 2013).

Explainable Case-based Reasoning with Counterfactuals

Explainable Case-Based Reasoning (ECBR) enhances learning by representing examples as similar prior cases with justifications derived from numerical, probabilistic, and logical analyses based on domain knowledge. These analyses, produced by algorithms emulating human-like processes, aim to increase understandability and transparency. The integration of counterfactual cases, as outlined by Byrne (2019), further aids in understanding by presenting alternative decision-making scenarios validated through rule-based criteria. Warren, Keane, and Byrne (2022) contribute by introducing a method that boosts explanations with categorical transformations, converting continuous features into categorical ones. This approach is supported by findings by Warren, Smyth, and Keane (2022), indicating that people learn more effectively from counterfactuals with categorical features.

Analyzing Decisions

More complex decision-making processes are often characterized by information scarcity and the need for rapid, accurate outcomes with temporal and computational constraints (Woike, Katsikopoulos, and Martignon 2011). Our approach includes a Monte Carlo simulation to model human future-oriented thinking, heuristic rules to emulate how humans use heuristics to analyze decisions quickly, and an event-based diagnosis tool to simulate human conceptual inference processes and uncertainty management. These methods help further analyze present and past cases to provide plausible justifications for a given decision.

Monte Carlo Tree Search

We employ simulation and a Monte Carlo tree search variant to emulate human risk analysis in scenarios with imperfect information and non-deterministic actions. Each simulation involves selecting actions, asserting unknown ground truths, and generating states. After multiple iterations, we analyze potential future states based on their likelihood and value (Browne et al. 2012). This iterative method mirrors human deliberation on possibilities (Yao, Zhao, and Liu 2022) and provides an explanatory mechanism through simulated roll-outs to justify decisions.

Heuristic Decision Making

Taking inspiration from human search strategies used in information parsing, heuristic decision-making recognizes that decision-makers may not have the time, resources, or cognitive abilities to gather and process all available information before making a decision (Mosier 2009). By applying domain-specific heuristic methods, algorithmic decision-makers can potentially better align with human decision-making processes and provide more realistic and effective solutions in general domains (Gigerenzer and Todd 1999) and, in particular, medical decision-making (Marewski and Gigerenzer 2012).

Event-Based Diagnosis

Our solution's event-based diagnosis component employs the Hybrid Event Memory System (HEMS) (Ménager, Choi, and Robins 2022), (Ménager and Choi 2023) to facilitate reasoning under uncertainty. HEMS, a psychologically plausible model, constructs a probabilistic experience taxonomy using Bayesian networks to represent episodes and schemas. Episodes, organized at the leaf level, and hierarchical schemas provide a structure for classifying new instances and learning incrementally. In decision-making scenarios, HEMS utilizes these elements to describe various aspects like casualties, demographic data, and actions. It excels in retrieving relevant information and conducting probabilistic inference for unobserved elements, thereby aiding in hypothesis evaluation and contextual understanding.

Justifications

The proposed algorithmic decision-maker compares possible decisions in a human-like way, presenting a novel form of justification supported by arguments. These arguments consist of a situation context, a set of factors that argue for or against a decision, and a set of weights that represent the importance of each factor in the specific context of the proposed decision.

Exposing these processes with human-like outputs and justifications enables decision-makers to improve their understanding of the applicable problem-solving methods employed (Floyd and Aha 2017). This is not possible with machine learning-based approaches such as neural networks that make decisions in a fundamentally different way and require post hoc explanation (Kenny and Keane 2021).

Key Decision-Maker Attributes

Our research suggests that analytic outcomes are perceived as subjectively significant by decision-makers based on their unique values and priorities (Shortland, Alison, and Moran 2019). This is particularly true in problem-solving scenarios lacking clear solutions, where the decision-maker's specific traits must align with those of the algorithm for the outcome to be deemed acceptable. Such alignment makes algorithmic recommendations more acceptable than those based solely on similarities to past cases (Kohn et al. 2021). Since direct insight into decision-makers' mental processes is impossible, we use inferred weighted mental process combinations to deduce these traits. The aim is to predict how these characteristics might influence future decisions, comparing them to past decisions that closely match the decision-maker's processes and are most relevant to the current situation.

Concluding Remarks

Our research introduces a method for algorithmic decision-making that seeks to replicate key elements of human decision processes, especially in situations where experts disagree and there is no clear answer. This approach utilizes explainable case-based reasoning, drawing on past experiences, contextual elements, and tools that use probabilistic reasoning, domain expertise, and characteristics of the decision-makers to create justifications that resemble human thought processes. This strategy allows for the examination of current and previous cases, offering reasoned explanations for the decisions made by the algorithm.

A crucial aspect of our study is the emphasis on attributes of the decision-maker, acknowledging the personalized significance of outcomes to individuals. Understanding how these attributes might impact future choices helps to ensure that our algorithmic system aligns with human experts.

Acknowledgments

This research was conducted as part of the In the Moment (ITM) project, supported by the Defense Advanced Research Projects Agency (DARPA) under contract number HR001122S0031.

References

Browne, C. B.; Powley, E.; Whitehouse, D.; Lucas, S. M.; Cowling, P. I.; Rohlfshagen, P.; Tavener, S.; Perez, D.; Samothrakis, S.; and Colton, S. 2012. A Survey of Monte Carlo Tree Search Methods. *IEEE Transactions on Computational Intelligence and AI in Games*, 4(1): 1–43.

Byrne, R. M. 2019. Counterfactuals in Explainable Artificial Intelligence (XAI): Evidence from Human Reasoning. In *IJCAI*, 6276–6282.

Floyd, M. W.; and Aha, D. W. 2017. Using explanations to provide transparency during trust-guided behavior adaptation. *AI Communications*, 30(3-4): 281–294.

Forbus, K. D.; and Hinrichs, T. 2017. Analogy and Qualitative Representations in the Companion Cognitive Architecture. *AI Magazine*, 38(4): 34–42.

Gigerenzer, G.; and Todd, P. M. 1999. Fast and frugal heuristics: The adaptive toolbox. In *Simple heuristics that make us smart*, Evolution and cognition, 3–34. New York, NY, US: Oxford University Press. ISBN 978-0-19-512156-8.

Kenny, E. M.; and Keane, M. T. 2021. Explaining Deep Learning using examples: Optimal feature weighting methods for twin systems using post-hoc, explanation-by-example in XAI. *Knowledge-Based Systems*, 233(C).

Kohn, S. C.; de Visser, E. J.; Wiese, E.; Lee, Y.-C.; and Shaw, T. H. 2021. Measurement of Trust in Automation: A Narrative Review and Reference Guide. *Frontiers in Psychology*, 12.

Marewski, J. N.; and Gigerenzer, G. 2012. Heuristic decision making in medicine. *Dialogues in Clinical Neuroscience*, 14(1): 77–89.

Mosier, K. L. 2009. Searching for coherence in a correspondence world. *Judgment and Decision Making*, 4(2): 154–163.

Ménager, D. H.; and Choi, D. 2023. Hybrid Event Memory as a Case Base for State Estimation in Cognitive Agents. In *Case-Based Reasoning Research and Development: 31st International Conference, ICCBR 2023, Aberdeen, UK, July 17–20, 2023, Proceedings*, 134–149. Berlin, Heidelberg: Springer-Verlag. ISBN 978-3-031-40176-3.

Ménager, D. H.; Choi, D.; and Robins, S. K. 2022. A Hybrid Theory of Event Memory. *Minds and Machines*, 32(2): 365–394.

Richter, M. M.; and Weber, R. O. 2013. *Case-Based Reasoning: A Textbook*. Berlin, Heidelberg: Springer. ISBN 978-3-642-40166-4 978-3-642-40167-1.

Shortland, N. D.; Alison, L. J.; and Moran, J. M. 2019. *Conflict: How Soldiers Make Impossible Decisions*. Oxford University Press. ISBN 978-0-19-062345-6.

Warren, G.; Keane, M. T.; and Byrne, R. M. J. 2022. Features of Explainability: How users understand counterfactual and causal explanations for categorical and continuous features in XAI. ArXiv:2204.10152 [cs].

Warren, G.; Smyth, B.; and Keane, M. T. 2022. "Better" Counterfactuals, Ones People Can Understand: Psychologically-Plausible Case-Based Counterfactuals Using Categorical Features for Explainable AI (XAI). In Keane, M. T.; and Wiratunga, N., eds., *Case-Based Reasoning Research and Development*, volume 13405, 63–78. Springer International Publishing. ISBN 978-3-031-14922-1 978-3-031-14923-8.

Woike, J. K.; Katsikopoulos, K. V.; and Martignon, L. 2011. Categorization with Limited Resources: A Family of Simple Heuristics. In *Crossref*. Oxford University Press. ISBN 978-0-19-974428-2.

Yao, F.; Zhao, H.; and Liu, H. 2022. Human-Robot Collaboration using Monte Carlo Tree Search. In *2022 International Conference on Automation, Robotics and Computer Engineering (ICARCE)*, 1–3.