

The Value of Real-Time Automated Explanations in Stochastic Planning (Abstract Reprint)

Claudia V. Goldman¹, Ronit Bustin², Wenyan Qi³, Zhengyu Xing³, Rachel McPhearson-White⁴ and Sally Rogers⁴

¹ Hebrew University Business School, Hebrew University of Jerusalem

² Toga Networks, a Huawei Company, Tel Aviv, Israel

³ General Motors (China) Investment Co., Ltd., Shanghai, China

⁴ General Motors, Warren Technical Center, Warren, MI, USA

Abstract Reprint. This is an abstract reprint of a journal article by [Goldman *et al.*, 2025].

and Sally Rogers. The value of real-time automated explanations in stochastic planning. *Artificial Intelligence*, 343:104323, 2025.

Abstract

Recently, we are witnessing an increase in computation power and memory, leading to strong AI algorithms becoming applicable in areas affecting our daily lives. We focus on AI planning solutions for complex, real-life decision-making problems under uncertainty, such as autonomous driving. Human trust in such AI-based systems is essential for their acceptance and market penetration. Moreover, users need to establish appropriate levels of trust to benefit the most from these systems. Previous studies have motivated this work, showing that users can benefit from receiving (handcrafted) information about the reasoning of a stochastic AI planner, for example, controlling automated driving maneuvers. Our solution to automating these hand-crafted notifications with explainable AI algorithms, XAI, includes studying: (1) what explanations can be generated from an AI planning system, applied to a real-world problem, in real-time? What is that content that can be processed from a planner's reasoning that can help users understand and trust the system controlling a behavior they are experiencing? (2) when can this information be displayed? and (3) how shall we display this information to an end user? The value of these computed XAI notifications has been assessed through an online user study with 800 participants, experiencing simulated automated driving scenarios. Our results show that real time XAI notifications decrease significantly subjective misunderstanding of participants compared to those that received only a dynamic HMI display. Also, our XAI solution significantly increases the level of understanding of participants with prior ADAS experience and of participants that lack such experience but have non-negative prior trust to ADAS features. The level of trust significantly increases when XAI was provided to a more restricted set of the participants, including those over 60 years old, with prior ADAS experience and non-negative prior trust attitude to automated features.

References

[Goldman *et al.*, 2025] Claudia V. Goldman, Ronit Bustin, Wenyan Qi, Zhengyu Xing, Rachel McPhearson-White,