Follow my Advice: Assume-Guarantee Approach to Task Planning with Human in the Loop

Georg Friedrich Schuppe, Ilaria Torre, Iolanda Leite and Jana Tumova

KTH Royal Institute of Technology and Chalmers University of Technology

Introduction

- When robots and humans operate in the same space, being able to guarantee safety is critical.
- **BUT:** It is hard to provide guarantees for **task satisfaction** and **safety** due to the unknown behaviour of the human actor.
- Since provable guarantees are difficult to obtain unconditionally, we take an assume-guarantee perspective. Along with guarantees on the robot's task satisfaction, we compute the weakest sufficient assumptions on the human's behavior.

Method

- We do not aim to explicitly represent the human's states, goals or intention through a model.
- Instead, we abstract the influence of the human's actions on the environment relevant to the robot's task through a set of atomic propositions and their changes.
- Given a robot represented as a labelled MDP $\mathcal{M}$ and a set of propositions controlled by the human actor Σ_h we create a labelled stochastic game

$$\widetilde{G} = ((\widetilde{S}, \widetilde{E}), Act \cup \{\epsilon\}, \tilde{\delta}, s_0, AP_r, L)$$

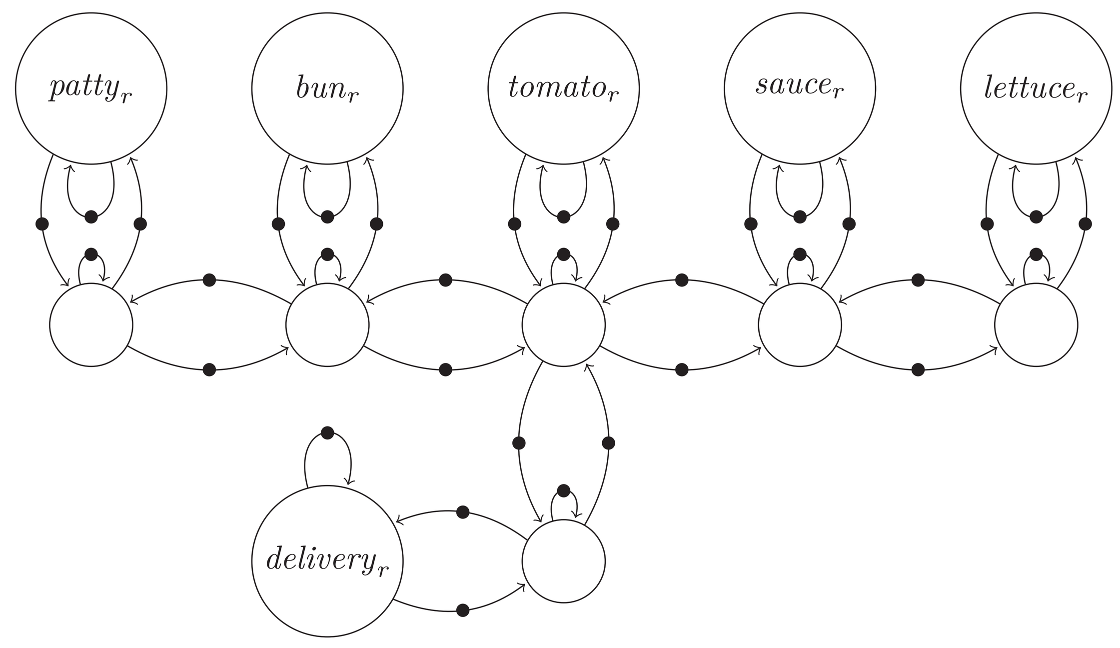
- In the next step, we introduce the specification of the robot into the game. First, we translate ϕ into a *deterministic finite automaton (DFA)*. We then construct a stochastic game with a reachability objective that is determined from the accepting states of the DFA. In this game, finding an almost-surely winning strategy translates to satisfying the specification ϕ.

$$\widehat{G} = ((\widehat{S}, \widehat{E}), Act \cup \{\epsilon\}, \hat{\delta}, \hat{s}_0, AP_r, \widehat{L})$$

- We generate sufficient assumptions split into two parts: Safety assumptions and fairness assumptions expressed as a set of player-2 edges from $\widehat{G}$.
- Safety assumptions $E_s \subseteq E_2$ are edges that player 2 can never take.
- Fairness assumptions $E_l \subseteq E_2$ of edges that need to be chosen fairly (i.e. infinitely many times upon infinitely many visits to their outgoing state).
- If the robot can find an almost-surely winning strategy in $\widetilde{G}$ for ψ, it satisfies ϕ regardless of what the human does. If not, additional assumptions are necessary to offer guarantees.

Problem Formulation

We model the robot's capabilities through a finite labelled MDP



The robot is given a high-level task formulated as an LTL$_f$ specification:

$$\begin{aligned}\phi_g =\,& F\ buns_r \wedge F\ patty_r \wedge F\ lettuce_r\\ &\wedge F(ketchup_r \wedge ketchup_h) \wedge F\,tomato_r\\ &\wedge G\ \neg(buns_r \wedge buns_h) \wedge G\ \neg(patty_r \wedge patty_h)\\ &\wedge G\ \neg(lettuce_r \wedge lettuce_h)\\ &\wedge G\ \neg(tomato_r \wedge tomato_h)\end{aligned}$$

Can we find a strategy reacting to the human actor that provably satisfies the specification? If not, can we provide conditional guarantees? What are the assumptions on the human actor?

From Assumption to Advice

During runtime, we communicate made assumptions to human actors. In our study, we used pictograms as a simple way of non-verbal communication.

Safety assumptions turn into safety advice:

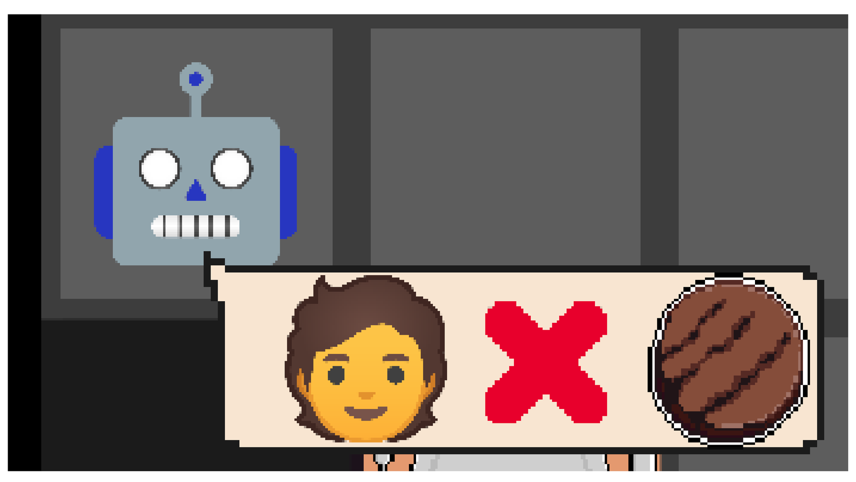


Figure 1: Safety Advice: Don't reach for the patties!

Fairness assumptions turn into fairness advice:

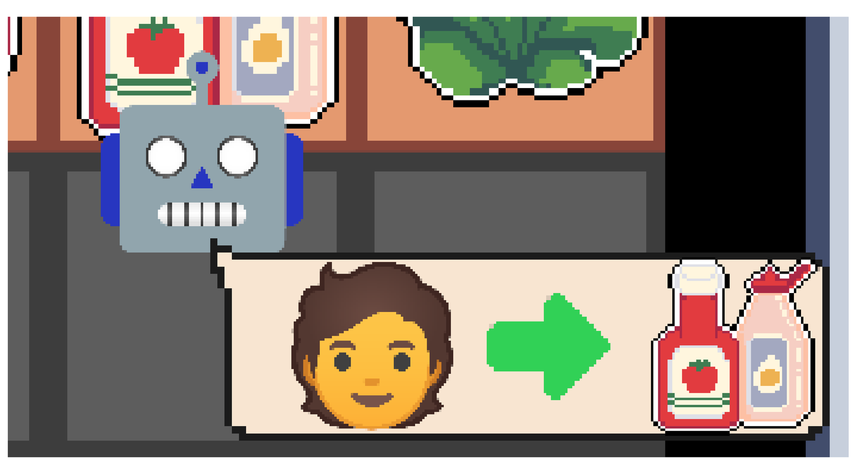


Figure 2: Fairness Advice: I am waiting for help with the sauce!

Contributions

- We propose to take an *assume-guarantee* perspective to correct-by-design task planning with a human in the loop.
- Along with guarantees on the robot's task satisfaction, we synthesize assumptions on the human's behavior.
- We conduct an online user study to compare our method to two baselines.

Study and Results

We conducted an online user study with 109 participants to compare our approach to two baselines:

- *NoAdvice*: The robot does not issue any advice.
- *NextMove*: The robot communicates to the user exactly which move to take in every step.

- In our approach the robot is perceived as safer, more intelligent and more compliant compared to the NextMove-condition.
- Our approach leads to less safety violations than not communicating at all.

Acknowledgement

This work is partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by Knut and Alice Wallenberg Foundation and the Swedish Research Council (VR) (project no. 2017-05102). This research has been carried out as part of the Vinnova Competence Center for Trustworthy Edge Computing Systems and Applications at KTH Royal Institute of Technology. The authors are also affiliated with Digital Futures.