Generating Explanations for Temporal Logic Planner Decisions

Daniel Kasenberg*, Ravenna Thielstrom, and Matthias Scheutz

dmk@cs.tufts.edu @dkasenberg dkasenberg.github.io

Human Robot Interaction Lab
Tufts University
@hrilab hrilab.tufts.edu

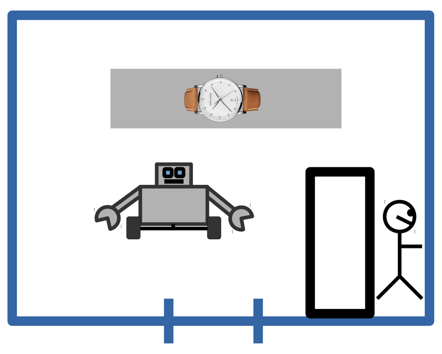
Introduction

We aim to develop artificial agents which can

- **Learn** interpretable objectives (through language and behavior) [1]
- **Behave** competently with respect to these objectives, *even when they conflict* [2]
- Explain their behaviors to human teammates in terms of these objectives (and correct objectives or world models if needed)

We have developed a system which constructs **explanations** for the behavior of a multi-objective **linear temporal logic (LTL) planning agent** operating in a **Markov decision process (MDP).**

Test scenario: ShopWorld



- Agent is a robot sent to go shopping for its user in a store selling a **watch**
- User wants the watch, but gives the robot insufficient money to afford it

Linear temporal logic (LTL) [5]

- A simple propositional logic encoding time

$$\phi ::= p \mid \neg\phi_1 \mid \phi_1 \vee \phi_2 \mid \phi_1 \wedge \phi_2 \mid \phi_1 \rightarrow \phi_2$$
$$\mid \mathbf{X}\phi_1 \mid \mathbf{G}\phi_1 \mid \mathbf{F}\phi_1 \mid \phi_1 \mathbf{U} \phi_2$$

where p a proposition, ϕ_1, ϕ_2 LTL statements.

- $\mathbf{X}\phi_1$: "in the next time step, ϕ_1"
- $\mathbf{G}\phi_1$: "in all present and future time steps, ϕ_1"
- $\mathbf{F}\phi_1$: "in some present or future time step, ϕ_1"
- $\phi_1\mathbf{U}\phi_2$: "ϕ_1 will be true until ϕ_2 becomes true"

LTL objectives in ShopWorld

"Leave the store while holding the watch"

$$\mathbf{F}(leftStore \wedge holding)$$

"Do not leave the store while holding anything which you have not bought"

$$\mathbf{G}\neg(leftStore \wedge holding \wedge \neg bought)$$

Multi-objective LTL planning

We define a multi-objective LTL planning problem as a tuple

$$\mathcal{P} = \langle \mathcal{M}, \mathbf{\Phi}, \mathbf{w}, \mathbf{z} \rangle$$

where

- $\mathcal{M} = \langle S, A, P, R, \gamma \rangle$ is a Markov decision process
- $\mathbf{\Phi} = \phi_1, \cdots, \phi_n$ a set of (syntactically safe/co-safe) LTL objectives
- $\mathbf{w}, \mathbf{z} \in \mathbb{R}^n$ contain the *weight* w_i and *priority* z_i respectively of ϕ_i, specifying preferences among objectives
  - Objectives w/ **same** priority → traded off using weights
  - Objectives w/ **different** priorities → higher-priority takes precedence (lexicographic ordering)

Basic solution approach:

- Compile each objective ϕ_i into a finite state machine (FSM) $M^{\phi_i} = \langle \Sigma^{\phi_i}, Q^{\phi_i}, \delta^{\phi_i}, q_0^{\phi_i}, F^{\phi_i} \rangle$ which accepts only on "good/bad prefixes" of ϕ_i
- Construct a new "product" MDP $\mathcal{M}^{\otimes}$ whose state space is $S \times Q^{\phi_1} \times \cdots \times Q^{\phi_n}$
- Construct reward functions $R^{\phi_1}, \cdots, R^{\phi_n}$ over $\mathcal{M}^{\otimes}$ such that $\mathcal{P}$ is a reward maximization problem (solvable, e.g., using value iteration)

Planning in ShopWorld

Given the LTL objectives above (and assuming the anti-shoplifting objective has higher priority), the agent performs only the single action *leaveStore.*

Explanation structures

Consider an agent who has acted according to trajectory τ. We wish to answer questions of the form "**Why** ψ?", where ψ is an arbitrary (safe/co-safe) LTL formula.

In response to such a question, we construct an explanation structure $\langle \Gamma, E, \tau', E' \rangle$, where

- $\Gamma \in \{\textsc{QueryFalse}, \textsc{NegQueryImpossible}, \textsc{AltQuery}\}$
- τ' is either a trajectory, or the empty set
- E contains one or more pairs $(\phi, \textsc{Evidence}(\tau, \phi))$ where
  - ϕ is an LTL statement
  - $\textsc{Evidence}(\tau, \phi) = \min\{|E| : E \subseteq \{0, \cdots, T\} \times L(\Pi);\ \tau \vDash E;\ \text{for all } \tau' \text{ s.t. } \tau' \vDash E, \tau' \notin \text{Traj}_{\checkmark}(\phi)\}$

e.g. in ShopWorld,

$$\textsc{Evidence}(\tau, \mathbf{F}(leftStore \wedge holding)) = \{(0, \neg holding), (1, \neg holding)\}$$

- E' is as E, but for τ'

Generating explanation structures

1. $\tau \vDash \psi$? If not, return
$$\langle \textsc{QueryFalse}, \{(\psi, \textsc{Evidence}(\tau, \psi))\}, \emptyset, \emptyset \rangle$$

e.g. **Why** $\mathbf{G}\neg leftStore$?
$$\langle \textsc{QueryFalse}, \{(\mathbf{G}\neg leftStore, \{(1, leftStore)\})\}, \emptyset, \emptyset \rangle$$

2. Does any possible trajectory τ' have $\tau' \vDash \neg\psi$? If not, return
$$\langle \textsc{NegQueryImpossible}, \emptyset, \emptyset, \emptyset \rangle$$

e.g. **Why** $\mathbf{G}\neg bought$?

3. Solve the augmented LTL planning problem
$$\left\langle \mathcal{M}, (\mathbf{\Phi}, \neg\psi), \begin{bmatrix} \mathbf{w} \\ 1 \end{bmatrix}, \begin{bmatrix} \mathbf{z} \\ \max_i z_i + 1 \end{bmatrix} \right\rangle$$

execute solution policy → alternative trajectory τ'; return
$$\langle \textsc{AltQuery}, \{(\phi, \textsc{Evidence}(\tau, \phi)) : \tau \text{ violates } \phi\}, \tau', \{(\phi, \textsc{Evidence}(\tau', \phi)) : \tau' \text{ violates } \phi\} \rangle$$

e.g. **Why** $\mathbf{G}\neg(leftStore \wedge holding)$?

1. $\tau \vDash \mathbf{G}\neg(leftStore \wedge holding)$ ✓
2. $\exists \tau'$ s.t. $\tau' \nvDash \mathbf{G}\neg(leftStore \wedge holding)$ ✓
3. τ': *pickUp, leaveStore*

return $\langle \textsc{AltQuery}, \{(\mathbf{F}(leftStore \wedge holding(glasses)), \{(0, \neg holding(glasses)), (1, \neg holding(glasses))\})\}, \tau', \{(\mathbf{G}\neg(leftStore \wedge holding(glasses) \wedge \neg bought(glasses)), \{(2, leftStore), (2, holding(glasses)), (2, \neg bought(glasses))\})\} \rangle$

Explanatory dialogue in natural language

- We have integrated our explanation approach into DIARC, a robotic architecture, to facilitate **natural language** explanatory dialogue. [3,4]
- Specifications and queries in an object-oriented extension to LTL (*violation enumeration language; VEL*) allowing object quantification
- Utterance → VEL query →explanation structure → natural language response
- Example: ShopWorld with two objects (*glasses* and *watch*); agent can afford one
  - Buys the glasses, leaves the watch

Input utterance	VEL query	Explanation in memory	Output utterance
"Why didn't you buy anything?"	**Why** $\forall x.\mathbf{G}\neg bought(x)$?	$\langle \textsc{QueryFalse}, \{(\forall x.\mathbf{G}\neg bought(x), \{(2, bought(glasses))\})\}, \varnothing, \varnothing \rangle$	"I bought the glasses"
"Why didn't you buy everything?"	**Why** $\exists x.\mathbf{G}\neg bought(x)$?	$\langle \textsc{NegQueryImpossible}, \varnothing, \varnothing, \varnothing \rangle$	"It was impossible for me to buy everything"
"Why didn't you leave the store while holding everything?"	**Why** $\forall x.\mathbf{G}\neg(leftStore \wedge holding(x))$?	$\langle \textsc{AltQuery}, \{\{(\mathbf{F}(leftStore \wedge holding(watch)), \{(0, \neg holding(watch)), (1, \neg holding(watch)), (2, \neg holding(watch))\})\}, \varnothing\}, \tau', \{(\mathbf{G}\neg(leftStore \wedge holding(glasses) \wedge \neg bought(glasses)), \{(3, leftStore), (3, holding(glasses)), (3, \neg bought(glasses))\})\} \rangle$	"I could have left the store while holding everything, but that would have violated more important rules"
↪"How would you have done that?"			"I would have picked up the glasses, picked up the watch, bought the watch, and left the store"
↪"What rules would you have broken?"			"I would have left the store while holding the glasses, which I had not bought"
↪"How would that have been worse?"			"Leaving the store while holding the glasses which I had not bought is worse than not leaving the store while holding the watch"

Future work

- Incorporating explicit causal models (esp. in $\textsc{NegQueryImpossible}$ case)
- Tailoring explanations to interactant knowledge
- Adapting to stochastic environments
- Need to represent multiple trajectories or probability distribution
- Improving efficiency of planner
- Impractical for nontrivial domains
- Dropping assumption that agent has perfect knowledge of transition dynamics

Acknowledgments

This project was supported in part by NSF IIS grant 43520050 and by NASA grant C17-2D00-TU.

References

[1] Kasenberg, D., & Scheutz, M. (2018, April). **Norm conflict resolution in stochastic domains.** In Thirty-Second AAAI Conference on Artificial Intelligence.

[2] **Kasenberg, D.** and Scheutz, M., 2017, December. **Interpretable apprenticeship learning with temporal logic specifications.** In 2017 IEEE 56th Annual Conference on Decision and Control (CDC) (pp. 4914-4921). IEEE.

[3] Kasenberg, D., Roque, A., Thielstrom, R., Chita-Tegmark, M. and Scheutz, M., 2019. **Generating justifications for norm-related agent decisions.** In Proceedings of the 12th International Conference on Natural Language Generation (pp. 484-493).

[4] Kasenberg, D., Roque, A., Thielstrom, R. and Scheutz, M., 2019. **Engaging in Dialogue about an Agent's Norms and Behaviors.** arXiv preprint arXiv:1911.00229.

[5] Pnueli, A. (1977, September). **The temporal logic of programs.** In 18th Annual Symposium on Foundations of Computer Science (sfcs 1977) (pp. 46-57). IEEE.