Learning Goal Recognition Models using Human Examples

Leonardo Amado
PUCRS, Brazil
leonardo.amado@edu.pucrs.br

Reuth Mirsky
Bar Ilan University, Israel
The University of Texas at Austin, USA
mirskyr@cs.biu.ac.il

Felipe Meneguzzi
University of Aberdeen, Scotland
PUCRS, Brazil
felipe.meneguzzi@pucrs.br

***Abstract*—Most approaches for goal recognition rely on specifications of the possible dynamics of the actor in the environment when pursuing a goal. However, encoding these dynamics requires careful design by a human expert, a design which is often not robust to noise at recognition time. In this paper, we present our recent framework that combines learning and goal recognition to alleviate the need for careful, manual domain design. This framework consists of two main stages: Offline learning of policies for each potential goal, and online inference. In this short paper, we focus on the first stage of learning the needed policies, and propose an approach to use behavioral cloning to elicit these policies. The aim of this extended abstract is to share the new goal recognition framework with the Human-Interactive Robot Learning (HIRL) community and obtain feedback on the best practices to promoting its implementation using human examples.**

I. Introduction

Goal recognition (GR) is a key task in artificial intelligence, where a *recognizer* infers the goal of an *actor* based on a sequence of observations [10]. Real-world scenarios for goal recognition applications include elderly care [7], traffic control [13], and even military planning [12]. Most of these applications require the recognizer to use images as input, thus needing sophisticated algorithms capable of dealing with noise and spurious data. A common approach to enable the robot to perceive and infer the person's goal in this situation consists of a pipeline of activity recognition from raw images and translation into actions for a symbolic GR algorithm (Figure 1 top). Once the raw images are processed into observations, a goal recognizer further processes a sequence of these observations into a goal or a distribution of goals. This process might include crafting elaborate domain theories, multiple planner executions in real-time, intricate domain optimizations, or any combination of these tasks [8], [14]. One of the main limitations of this approach is the high cost of constructing these domain theories. This construction requires a deliberate design and accurate specification of domain dynamics, which is usually a process done manually by an expert. Moreover, some recognizers require costly online computations, such as multiple planner executions. These computations can hinder the recognizer's real-time inference ability, especially when observations are processed incrementally and the goal of the actor is re-evaluated often throughout the plan execution [3].

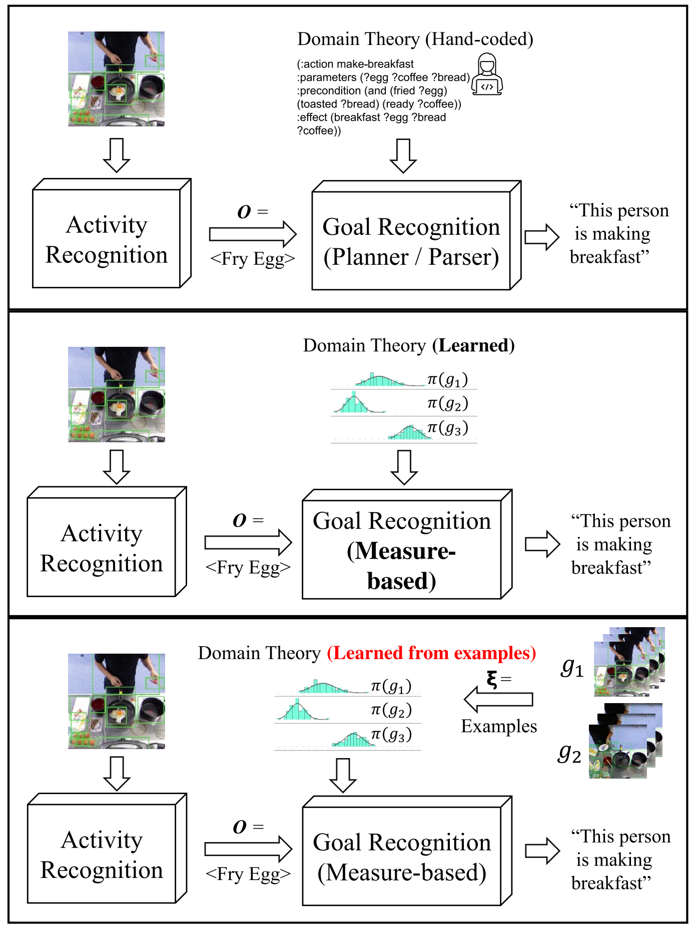

Fig. 1: A comparison of existing model-based approaches for goal recognition (top), the intermediate proposed solution already implemented (middle), and our proposed framework (bottom). The key changes are presented in bold and the new proposed changes are presented in red.

The goal of this short paper is to propose an alternative approach for goal recognition which fully relies on user-generated examples that will be used to learn the domain theory instead of the manually-crafted descriptions of a domain expert. We first outline our current intermediate framework in which the domain theory is replaced with a set of policies

(Figure 1 middle), and then we elaborate on our next plans for extending this framework to use learning from examples instead of model-free reinforcement learning (Figure 1 bottom). Our aim in this publication is sharing this progress with the Human-Interactive Robot Learning (HIRL) community and getting their feedback on the best practices to use to promote these ideas further.

II. Goal Recognition as Reinforcement Learning

The Goal Recognition as Reinforcement Learning (RG as RL) framework aims to address the limitations discussed earlier by replacing manually crafted representations and online executions with model-free Reinforcement Learning (RL) techniques. This framework performs efficient and robust GR without the need to craft a domain model and without any planner or parser executions during recognition [1].

The gist of this work is that we replace the manually-constructed domain theory with a set of *policies*, such that each policy is representative for a different goal (Figure 1 middle). This framework consists of two main stages:

1) Learning a set of policies, such that for each potential goal the actor might pursue $g \in G$, there is a respective policy π_g. Ideally, π_g is representative of the actor's behavior when aiming to achieve g.
2) Inferring the goal of an actor given a sequence of observations, $\mathrm{O} = \langle s_0, a_0, s_1, a_1, \ldots \rangle$. Once the policies $\{\pi_g\}_{g \in \mathcal{G}}$ are known, and given an observation sequence O, the inferred goal g^* is the one that minimizes the measured distance (Distance) between its respective policy and the observations, as defined in Equation 1.

$$g^* = \arg\min_{g \in \mathcal{G}} \mathrm{Distance}(\pi_g, \mathrm{O}) \tag{1}$$

There are many ways to implement this framework, where the key decisions that will affect the resulting algorithm are: the learning approach used to elicit an accurate set of policies; and the choice of the *Distance* measure to compare the observation sequence and the policies. In this paper, we focus on the first decision – learning an accurate set of policies, but instead of using traditional RL algorithms, we propose to learn these policies from human examples.

III. Learning from Human Examples

While there are many ways to learn from human examples – demonstrations [2], [15], reward shaping [16], [17], evaluative feedback [6], [11], corrections [5] – in this work we focus on *Imitation Learning*, which is a problem setup in which an agent is trained to perform a task from demonstrations by learning a mapping between observations and actions [9]. Imitation learning is our first choice, as it does not require the human to be aware of the action space used to model the environment or even to be aware of the learner observing it, which enables the actor to adopt more naturalistic behaviors rather than using pedagogical instructions or needing to model the learner's knowledge. This ability is especially important when learning a model for keyhole goal recognition [10], as a key assumption in this problem is that the actor is not aware of the observer and does not try to assist or hide its goals.

One of the main approaches to solve the imitation learning problem is behavioral cloning [4], in which the learner tries to replicate the policy of the human based on observations, and it can complement or even fully replace learning a policy from the environment. As seen in Figure 1 (bottom), behavioral cloning will be used separately for each goal $g \in G$ to learn a set of policies π_g. Then, these policies will be used for goal recognition in a similar fashion to GR as RL, where an observation sequence is matched with the most likely goal the actor is trying to pursue.

As this is an ongoing work, we would love to participate in the HIRL workshop so we can get expert human feedback that will enable us to learn better policies, which we will then use to accomplish our research goals.

References

[1] Anonymous. Anonymized. In *Anonymous*, Anonymous 2022.
[2] Brenna D Argall, Sonia Chernova, Manuela Veloso, and Brett Browning. A survey of robot learning from demonstration. *Robotics and autonomous systems*, 57(5):469–483, 2009.
[3] Masataro Asai and Alex Fukunaga. Classical Planning in Deep Latent Space: Bridging the Subsymbolic-Symbolic Boundary. In *Proc. of the AAAI Conf. on Artificial Intelligence*, 2018.
[4] Michael Bain and Claude Sammut. A framework for behavioural cloning. In *Machine Intelligence 15*, pages 103–129, 1995.
[5] Andreea Bobu, Marius Wiggert, Claire Tomlin, and Anca D Dragan. Feature expansive reward learning: Rethinking human input. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, pages 216–224, 2021.
[6] Yuchen Cui, Qiping Zhang, Alessandro Allievi, Peter Stone, Scott Niekum, and W Bradley Knox. The empathic framework for task learning from implicit human feedback. *arXiv preprint arXiv:2009.13649*, 2020.
[7] Christopher W Geib. Problems with intent recognition for elder care. In *Proc. of the AAAI-02 Workshop "Automation as Caregiver*, pages 13–17, 2002.
[8] Christopher W Geib and Robert P Goldman. A probabilistic plan recognition algorithm based on plan tree grammars. *Artificial Intelligence*, 173(11):1101–1132, 2009.
[9] Ahmed Hussein, Mohamed Medhat Gaber, Eyad Elyan, and Chrisina Jayne. Imitation learning: A survey of learning methods. *ACM Computing Surveys (CSUR)*, 50(2):1–35, 2017.
[10] Henry A Kautz and James F Allen. Generalized plan recognition. In *AAAI*, volume 86, page 5, 1986.
[11] W Bradley Knox and Peter Stone. Interactively shaping agents via human reinforcement: The tamer framework. In *Proceedings of the fifth international conference on Knowledge capture*, pages 9–16, 2009.
[12] Jean Oh, Felipe Meneguzzi, Katia Sycara, and Timothy J Norman. An agent architecture for prognostic reasoning assistance. In *IJCAI*, 2011.
[13] David V. Pynadath and Michael P. Wellman. Accounting for context in plan recognition, with application to traffic monitoring. *CoRR*, abs/1302.4980, 2013.
[14] Miguel Ramírez and Hector Geffner. Probabilistic plan recognition using off-the-shelf classical planners. In *Twenty-Fourth AAAI Conference on Artificial Intelligence*, 2010.
[15] Stefan Schaal et al. Learning from demonstration. *Advances in neural information processing systems*, pages 1040–1046, 1997.
[16] Halit Bener Suay and Sonia Chernova. Effect of human guidance and state space size on interactive reinforcement learning. In *2011 Ro-Man*, pages 1–6. IEEE, 2011.
[17] Ana C Tenorio-Gonzalez, Eduardo F Morales, and Luis Villasenor-Pineda. Dynamic reward shaping: training a robot by voice. In *Ibero-American conference on artificial intelligence*, pages 483–492. Springer, 2010.