Neurosymbolic Active Goal Recognition in Partially Observable Environments

Extended Abstract

Chenyuan Zhang*
Monash University
Melbourne, Australia
chenyuan.zhang@monash.edu

Sukai Huang*
Monash University
Melbourne, Australia
sukai.huang@monash.edu

Hamid Rezatofighi
Monash University
Melbourne, Australia
hamid.rezatofighi@monash.edu

Mor Vered
Monash University
Melbourne, Australia
mor.vered@monash.edu

Buser Say
Monash University
Melbourne, Australia
buser.say@monash.edu

ABSTRACT

Active goal recognition, despite its importance for human–AI interaction and autonomous systems, has received relatively limited attention. Unlike passive goal recognition, which infers an actor's intent from observations alone, active goal recognition allows an observer to select informative actions to reduce uncertainty about the actor's goal. Building upon prior work in symbolic active goal recognition under POMDP settings, this paper introduces a neurosymbolic framework that addresses two key limitations. First, we extend the modeling capacity to account for heterogeneous actor behaviors, moving beyond the hand-crafted actor behaviour assumption. Second, we integrate neural models into the active goal recognition framework in two complementary ways: (i) by replacing actor models with Vision Language Models (VLMs) trained from data, and (ii) by employing reinforcement learning to train the observer over belief maps, thereby enabling adaptive decision-making beyond symbolic observer policy. Experiments on the grid-world domain show that our neurosymbolic approach achieves comparative performance over state-of-the-art symbolic methods. These results highlight the promise of neurosymbolic methods for robust active goal recognition in complex, uncertain environments.

KEYWORDS

Goal Recognition; Reinforcement Learning; Neurosymbolic Agent; POMDP

ACM Reference Format:

Chenyuan Zhang, Sukai Huang, Hamid Rezatofighi, Mor Vered, and Buser Say. 2026. Neurosymbolic Active Goal Recognition in Partially Observable Environments: Extended Abstract. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 3 pages. https://doi.org/10.65109/NCCT4251

*These authors contributed equally to this work.

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). https://doi.org/10.65109/NCCT4251

1 PROBLEM AND CONTRIBUTIONS

Goal recognition (GR) infers an actor's goal from a sequence of observed actions and is widely studied in planning and agent modelling [2, 4, 5]. Most GR methods assume a *passive* observer that only interprets actions after they happen and cannot change what it observes. In many multiagent settings this is unrealistic, because an observer can often move, act, or query the environment to obtain more informative observations. Active Goal Recognition (AGR) captures this by allowing the observer to take actions to reduce uncertainty about the actor's hidden goal, for instance by repositioning to get a clearer view of the actor in a cluttered workspace [1, 3, 6]. Recent work [6] models AGR as a POMDP, where the observer maintains beliefs and plans actions to gain information. While effective, these approaches typically rely on the *rational actor* assumption, that the actor follows (near-)optimal plans. However, people (and agents) can reach the same goal via very different, sometimes exploratory, behaviours (e.g., one person goes straight to the coffee machine while another checks their email first). Moreover, in complex domains computing optimal policies is expensive. Therefore, a reliable symbolic actor model may be unavailable. These challenges motivate approaches that can represent heterogeneous behaviours and learn actor models from data.

We make three contributions: (a) **Typed probabilistic AGR (T-PAGR):** extend probabilistic AGR to jointly infer the actor's goal and a latent *behavior type* that parameterizes action likelihoods; (b) **Neuro-AGR actor modeling:** replace hand-crafted actor likelihoods with a VLM model trained from trajectories, enabling data-driven heterogeneous behavior modeling. (c) **Neuro-AGR observer policy:** train an observer with reinforcement learning (RL) on belief maps, which replaces hand-coded policy and enables adaptive information gathering. Figure 1 presents the overall framework of our work and we will briefly introduce main components in the following sections.

2 TYPED PROBABILISTIC AGR

We consider an *actor* acting toward a hidden goal $g \in G$ in a shared environment and an *observer* that can move and sense under partial observability. Following prior work [6], the observer's decision

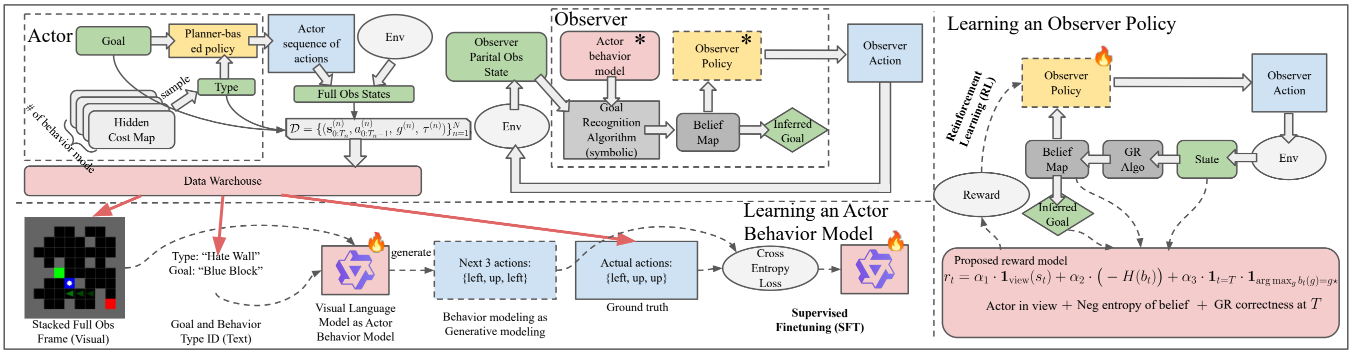

Figure 1: Neuro-AGR replaces the observer's internal symbolic models with learnable components. Top Left: Actor and observer architecture. Bottom Left: The learning-based actor behavior model based on a Vision Language Model (VLM) is trained on offline trajectories. Right: The observer policy can be learned via RL with a reward that encourages actor visibility and reduces goal uncertainty. This replaces hand-crafted, model-based strategies with a data-driven approach.

problem is modeled as a probabilistic POMDP whose hidden variables include the actor's state and goal. Under the *keyhole* assumption, the actor is unaware of the observer, so the actor's behavior does not depend on the observer's actions. The observer maintains a joint belief $j_t(s_t, g) = P(s_t, g \mid o_{0:t}, u_{0:t})$, updated recursively by (i) predicting the actor's next state using a goal-conditioned actor model and (ii) correcting using the observation likelihood $P(o_t \mid s_t, u_t)$. The resulting belief over goals $b_t(g) = \sum_{s_t} j_t(s_t, g)$ can be used to define a belief-based reward that encourages actions that sharpen the goal distribution.

To model heterogeneous actor behaviors, we extend PAGR with a *static* latent type $\tau \in \mathcal{T}$ that remains fixed within an episode and modulates the actor's policy. The observer does not observe τ and therefore performs joint inference over (s_t, g, τ) via a typed belief $j_t(s_t, g, \tau)$. Belief prediction uses a typed transition model $P_\theta(s_{t+1} \mid s_t, g, \tau)$ (reducing to standard PAGR when $|\mathcal{T}| = 1$), while the other parts of the algorithm remain unchanged.

3 NEURO-AGR: LEARNING THE ACTOR AND OBSERVER

Neuro-AGR is a modular observer architecture that keeps *goal inference symbolic* while making both the *actor model* and the *observer policy* replaceable learning-based components. The standard probabilistic AGR framework can be seen as a special case where both modules are hand-crafted and symbolic.

To model actor behavior, we fine-tune a vision–language model (VLM) on goal- and type-labeled trajectories to predict the actor's next action given a pixel-based state representation and text conditioning on (g, τ). The predicted action distribution is then integrated into the symbolic belief update. Because VLM inference is expensive and can be unreliable, we use a neurosymbolic fallback: the VLM is activated only when the actor is visible (avoiding costly enumeration over unobserved states), and its predictions are used only when confident; otherwise the system falls back to a lightweight symbolic distance-based predictor.

For decision making, instead of relying on hand-coded observer strategies, we learn the observer policy π_ϕ with reinforcement learning (PPO). The policy takes as input a compact *belief-map* image,

Table 1: Goal Recognition Average Convergence Rate (%) for Different Observer Configurations on the grid-world domain.

Size	Initial Distance (d_0) between Actor and Observer											
	Pure Symbolic			Hybrid W/O Threshold			Hybrid W/ Threshold ⋆			Reinforcement Learning		
	$d_0 = 3$	$d_0 = 5$	$d_0 = 7$	$d_0 = 3$	$d_0 = 5$	$d_0 = 7$	$d_0 = 3$	$d_0 = 5$	$d_0 = 7$	$d_0 = 3$	$d_0 = 5$	$d_0 = 7$
10	**38**	29	**24**	28	25	20	37	**31**	23	**38**	30	20
12	37	33	25	31	32	24	**40**	33	25	33	**37**	**27**
15	37	27	19	33	26	18	**39**	26	19	38	**30**	**23**

which summarizes the observer's current posterior over the actor's location and goal preferences and naturally handles partial observability. This representation also improves generalization because it abstracts away environment-specific visual details. We design the reward signal that encourages the actor in view and reducing entropy of the goal belief.

4 RESULTS

We report results on a test set of **1,535** task instances spanning all combinations of **grid sizes** (10×10, 12×12, 15×15), **initial distances** ($d_0 \in \{3, 5, 7\}$), and **four latent actor behavior types**. All learning-based components are trained on instances generated by the same procedure (with separate validation layouts), and we evaluate using the standard **convergence rate (CV)** metric (higher is better), which rewards earlier and stable concentration of belief on the true goal.

Table 1 shows that **naively** integrating the VLM actor model (*Hybrid W/O Threshold*) consistently **degrades** convergence relative to the *Pure Symbolic* baseline, indicating that low-confidence VLM predictions can corrupt the belief update. In contrast, **confidence-thresholded** integration (*Hybrid W/ Threshold*) is **robust** and yields the best performance when the actor is more often visible (typically $d_0 \in \{3, 5\}$), while the *RL observer policy* is competitive and achieves the top result in several harder settings (e.g., larger maps / d_0=7).

ACKNOWLEDGMENTS

This work is supported by the DARPA Assured Neuro Symbolic Learning and Reasoning (ANSR) program under award number FA8750-23-2-1016.

REFERENCES

[1] Peta Masters and Mor Vered. 2021. What's the Context? Implicit and Explicit Assumptions in Model-Based Goal Recognition. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*, Zhi-Hua Zhou (Ed.). International Joint Conferences on Artificial Intelligence Organization, 4516–4523. https://doi.org/10.24963/ijcai.2021/615 Survey Track.

[2] Miquel Ramırez and Hector Geffner. 2011. Goal recognition over POMDPs: Inferring the intention of a POMDP agent. In *IJCAI*. IJCAI/AAAI, 2009–2014.

[3] Maayan Shvo and Sheila A McIlraith. 2020. Active goal recognition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. 9957–9966.

[4] Mor Vered, Gal A Kaminka, and Sivan Biham. 2016. Online goal recognition through mirroring: Humans and agents. In *Annual Conference on Advances in Cognitive Systems 2016*. Cognitive Systems Foundation.

[5] Chenyuan Zhang, Charles Kemp, and Nir Lipovetzky. 2024. Human Goal Recognition as Bayesian Inference: Investigating the Impact of Actions, Timing, and Goal Solvability. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems*. 2066–2074.

[6] Chenyuan Zhang, Cristian Rojas Cardenas, Hamid Rezatofighi, Mor Vered, and Buser Say. 2025. Probabilistic Active Goal Recognition. In *Proceedings of the 22nd International Conference on Principles of Knowledge Representation and Reasoning*. 880–890. https://doi.org/10.24963/kr.2025/85