Enhancing Teleoperation in Dynamic Environments: A Novel Shared Autonomy Framework Leveraging Multimodal Language Models

Guillaume Lorthioir[1,2], Mehdi Benallegue[1,2], Rafael Limon Cisneros[1,2], Ixchel Ramirez[1]

I. INTRODUCTION

Robots often lack full autonomy, necessitating human teleoperation, such as in surgery or hazardous interventions, and extending to embodying human presence remotely, as seen in projects like the ANA Avatar Xprize [1]. However, challenges such as communication delays, obstructions, complex interfaces, and operator expertise persist, affecting teleoperation efficacy. Shared autonomy enhances teleoperation by providing robots with partial autonomy, assisting teleoperators in task completion—for example, inferring operator goals and autonomously positioning grippers for object manipulation, expediting processes. Yet, current shared autonomy approaches face difficulties in dynamic environments. This paper introduces a novel shared autonomy framework addressing this challenge.

The methodology employs a Multimodal Language Model (MLM) to dynamically analyze the robot's environment, identifying potential tasks for the operator. Subsequently, goal recognition using a Hidden Markov Model (HMM) predicts the operator's most probable task, refining robot motion accordingly. This ongoing research emphasizes leveraging MLM for dynamic environmental adaptation. Recent advancements in Large Language Models (LLM) and MLM have substantially enhanced reasoning capabilities, with Google researchers demonstrating the feasibility of issuing high-level commands to robots, translated into low-level commands executed by the robot [2].

This framework consists of three components: MLM-based dynamic environment identification and goal generation, goal recognition using HMM, and assistance, determining when and to what extent the operator should be assisted with the robot.

II. MAIN METHOD

A. Interpreting the environment

Current shared autonomy frameworks face challenges in adapting to dynamic environments, often constrained by predefined settings where robots possess precise object knowledge. However, introducing new objects, relocating existing ones, or adjusting the robot's position can hinder these frameworks' effectiveness. To tackle this, we propose leveraging MLM.

Initially, the robot can use video feedback to identify objects in its vicinity. Real-time object detection algorithms like YOLO8 can assist in this process. Building on [2] approach, we combine these visuals with high-level action keywords familiar to the robot (e.g., ”grab,” ”pour,” ”move to”). This fusion dynamically generates user goals, denoted as $G = g_1, g_2, ..., g_n$. Each goal g_i correlates with a high-level action tied to an object (e.g., ”grab the bottle”). Refinement of these goals incorporates knowledge of the robot's current state and preconditions, filtering achievable goals from unattainable ones (e.g., ”pour the bottle” requires first ”grabbing the bottle”), thus optimizing computational efficiency.

This method enables the generation of potential goals, empowering users within the shared autonomy framework across diverse scenarios. To manage computational load, we address the timing of goal updates. Updating after goal achievement or when the environment's object count changes significantly by more than N from the prior step offers viable strategies.

In summary, by integrating MLM with object recognition, we enhance shared autonomy frameworks' adaptability, enabling efficient goal generation and dynamic assistance in varied contexts. Strategic goal updating further optimizes computational resources, ensuring practical implementation.

B. Goal recognition

When potential goals have been identified, our framework needs to predict which goal is the most likely to be pursued by the user in order to assist them. User interaction with the robot for task execution involves a discrete goal set $G = g_1, g_2, ..., g_n$, with n goals. Actions, such as controller input, eye tracking, and robot motion, are observed. The observation vector $\Theta_t = \theta_t^1, \theta_t^2, ..., \theta_t^k$ at time t varies in size (k). Real-time inference of the user's current goal g^* is attempted based on these observations. Our observations, sourced directly from teleoperation system feedback, are presumed to be complete. However, user actions may change due to personal will, potentially leading to goal shifts or indecision during task execution. Inspired by [3], we adopt a Bayesian filtering approach within a Hidden Markov Model (HMM) for goal recognition. HMM hidden states represent achievable goals (G) alongside an ”Undecided” state, denoted as $Undecided \in G$ indicating the absence of a specific goal pursuit. Observations (O) encompass all possible observation vectors. Figure 1 illustrates this HMM, where user transitions within upper states (G) directly influence observation likelihoods. The Forward Algorithm coupled with the law of total probability and the chain rule can be utilized to model a probability distribution over G, reflecting uncertainty in candidate goals. We first consider

[1]National Institute of Advanced Industrial Science and Technology, Japan
[2]Centre National de la Recherche Scientifique, France

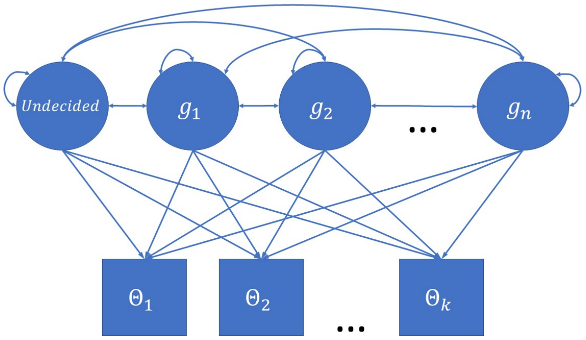

Fig. 1. HMM diagram illustrating our goal recognition problem. Upper states represent different states in G and their transitions, influencing the lower observation set O.

the case of a single observation at each time step, where $\Theta_t = \theta_t$ is a vector of dimension one. We use the colon notation $\theta_{0:t}$ for the sequence of observation $\theta_0, ..., \theta_t$, we define $b_t(g_t)$ as the belief of the user pursuing goal $g \in G$ at time t, as follows:

$$b_t(g_t) = P(\theta_t|g_t) \sum_{g_{t-1} \in G} P(g_t|g_{t-1}) b_{t-1}(g_{t-1}) \quad (1)$$

Equation 1 is used to compute $b_t(g_t)$ in the context of a single observation at each time step. However, we aim to incorporate multiple observations, as a greater volume of data typically enhances goal prediction accuracy. We now address the case of $\Theta_t = \theta_t^1, \theta_t^2, ..., \theta_t^k$ being of dimension $k > 1$. We assume that the distinct observations $\theta_t^1, \theta_t^2, ..., \theta_t^k$ are conditionally independent given a goal g. Hence, we can express equation 1 as:

$$b_t(g_t) = (\prod_{\theta_t \in \Theta_t} P(\theta_t|g_t)) \sum_{g_{t-1} \in G} P(g_t|g_{t-1}) b_{t-1}(g_{t-1}) \quad (2)$$

Computing this for all $g \in G$ provides us with the estimated distribution b_t representing our belief regarding the user's goal across the set G at time t. To determine g_t^*, we simply select the goal g that maximizes the value of b_t.

In contrast to many shared autonomy methods, which rely on trajectory planning to approximate the likelihood value $P(\theta_t|g_t)$ and may incur computational overhead, we adopt a landmark heuristic inspired by [4]. In the pursuit of a goal, a landmark denotes a fact or action essential for its achievement. We employ the principle of Landmark Uniqueness, which involves assessing the occurrence of this landmark across various goals.

C. Assistance

Deciding when and how much a robot should assist a user is complex. Our goal recognition model isn't perfect. As Meneguzzi and Pereira [5] note, no existing approach can make flawless online predictions. Thus, we must consider the possibility of model errors leading to incorrect robot assistance. This could harm teleoperation performance instead of improving it. To assess our predictions, we gauge the confidence in them. Dragan et al. [6] proposed various methods for computing confidence. We opt to compute confidence based on the entropy of the probability distribution b_t, denoted as $conf$:

$$conf = 1 + \frac{\sum_{g_t \in G} b_t(g_t) log(b_t(g_t))}{H_{max}} \quad (3)$$

Here, H_{max} represents the maximum value that the entropy (the numerator) could reach. When b_t tends to resemble a uniform distribution, $conf$ approaches 0, indicating low confidence in our prediction, as no goal stands out as more probable than others. Conversely, when b_t assigns high probabilities to only a few goals, resulting in low entropy, $conf$ tends toward 1. This reflects the confidence that our goal recognition algorithm has identified only a few goals as likely. Then, we employ policy-blending, one of the most commonly used methods for shared autonomy [6]. It involves combining the trajectory desired by the user for the end-effector with the trajectory desired by the robot. We define u_b as the blend policy and obtain it as follows:

$$u_b = u_h(1 - \alpha) + u_r \alpha \quad (4)$$

Here, u_h represents the human policy and u_r denotes the robot policy, with $\alpha \in [0, 1]$ serving as a parameter controlling the weight of the human and the robot in the blended policy. The value of α is determined by the $conf$ metric detailed earlier. Specifically, $\alpha = conf$ if $conf \in [0.2, 0.8]$, 0 if $conf$ is smaller than that, and 0.8 otherwise.

III. Discussion

The framework introduced in this paper is an ongoing work aimed at addressing the challenge of dynamic environments for shared autonomy. As explained earlier, current shared autonomy approaches struggle with dynamic environments since the assistance the robot can provide to the operator is typically limited to a few actions and objects. This framework could enable dynamic adaptation to the environment by leveraging MLM, thus providing assistance to the operator in many more situations than current frameworks can accommodate. A fast algorithm for goal recognition, based on Hidden HMM and a landmark heuristic, is utilized to compute the most likely goal pursued by the user, and subsequently assists the operator based on the confidence in our prediction.

We posit that this framework could reduce the cognitive load on the operator and provide a more robust teleoperation method in cases of noisy or delayed signals. Furthermore, in future work, the framework could be adapted to facilitate cooperation between autonomous robots and humans. The main distinction would be that instead of collecting data from the operator, the robot would gather data from an observed human and then attempt to assist them to the best of its abilities.

References

[1] "Ana avatar xprize." `https://www.xprize.org/prizes/avatar`, 2023.

[2] D. Driess, F. Xia, M. S. Sajjadi, C. Lynch, A. Chowdhery, B. Ichter, A. Wahid, J. Tompson, Q. Vuong, T. Yu, *et al.*, “Palm-e: An embodied multimodal language model,” *arXiv preprint arXiv:2303.03378*, 2023.

[3] S. Jain and B. Argall, “Probabilistic human intent recognition for shared autonomy in assistive robotics,” *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 9, no. 1, pp. 1–23, 2019.

[4] R. F. Pereira, N. Oren, and F. Meneguzzi, “Landmark-based approaches for goal recognition as planning,” *Artificial Intelligence*, vol. 279, p. 103217, 2020.

[5] F. R. Meneguzzi and R. F. Pereira, “A survey on goal recognition as planning,” in *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI), 2021, Canadá.*, 2021.

[6] A. D. Dragan and S. S. Srinivasa, “A policy-blending formalism for shared control,” *The International Journal of Robotics Research*, vol. 32, no. 7, pp. 790–805, 2013.