Cognitive Agentic AI: Probabilistic Novelty Detection for Continual Adaptation in HRI

Khashayar Ghamati[1,*], Farshid Amirabdollahian[1], Diego Resende Faria[2] and Abolfazl Zaraki[1,*]

Abstract— **Adapting to novel tasks in human-robot interaction (HRI) is crucial for long-term autonomy, yet remains a major challenge for autonomous agents deployed in unpredictable open-world settings. This paper introduces CAPA-AI, a novel framework that integrates probabilistic novelty detection with continual post-deployment adaptation achieved via transfer learning to address this challenge. The framework's novelty detection component employs conditional probability and the Jaccard Index to identify unfamiliar tasks by quantifying their deviation from the agent's knowledge base of previously learned tasks. Upon detecting a novel task, the agent utilises transfer learning to repurpose prior knowledge and update its models without retraining from scratch. We detail the design of CAPA-AI, including an isolated learning phase for initial skill acquisition and the construction of a dynamic knowledge base. The complete system was deployed on a social robot in real-world HRI scenarios to evaluate its performance. Experimental results demonstrated that the agent accurately detects novel tasks and adapts to them, achieving adaptation and novelty detection accuracies of 80% and 89%, respectively. These findings underscore the efficacy of the proposed approach and highlight a significant step towards robust open-world deployment of AI agents in HRI, where continuous adaptation and the safe handling of unforeseen tasks are essential.**

I. INTRODUCTION

Developing AI agents capable of effectively adapting to novel environments and managing unforeseen circumstances remains an enduring challenge, primarily due to the inherent complexity and unpredictability of real-world interactions. Adaptation within robotics and AI spans a broad spectrum of tasks, ranging from simple gestures—such as a social robot waving goodbye when a user departs—to sophisticated social behaviours that accurately interpret and respond to subtle human intentions. Equipping such an agent, typically comprising multiple integrated AI models, necessitates extensive training and an in-depth understanding of the diverse scenarios it might encounter. Nonetheless, given the inherent uncertainty of real-world environments, anticipating and preparing for every conceivable scenario is fundamentally impractical. Development generally commences with the training of individual models, subsequently integrating them into a cohesive agent capable of robust real-world interactions [1].

AI agents' adaptive capabilities can be conceptualised as comprising two distinct phases: initial training and post-deployment adaptation. In the initial phase, the model undergoes rigorous training via simulations or extensive datasets, exposing it to diverse learning scenarios. This training employs varied strategies aimed at optimising performance while mitigating prevalent issues such as overfitting and underfitting. Nevertheless, despite thorough pre-deployment training, effectively updating an AI model's knowledge during the subsequent post-deployment phase remains notably challenging. Introducing new information in operational environments, particularly in the absence of suitable supervision, can trigger substantial issues such as catastrophic forgetting [2]. Consequently, this deterioration in performance exacerbates the risks associated with overfitting and underfitting.

Furthermore, real-world interactions inherently involve significant uncertainty: sensor data can be noisy, environments may be only partially observable, and sudden changes can render previously acquired knowledge obsolete. An agent must therefore assess its confidence accurately and recognise when it encounters an unforeseen or novel state. Without the capacity to estimate uncertainty, an agent risks acting recklessly in unfamiliar or ambiguous situations. Contemporary safe Reinforcement Learning (RL) methodologies attempt to mitigate this risk by estimating the potential danger associated with actions, thereby avoiding those deemed excessively uncertain or unsafe [3]. Yet, reliably managing unforeseen circumstances remains an intricate challenge. This uncertainty is intrinsically linked to the necessity for continual adaptation, wherein the agent identifies when its current policy or model is no longer sufficient and subsequently updates itself accordingly.

Addressing these complexities necessitates the development of robust methods enabling AI agents to continuously adapt their internal models in response to novel post-deployment scenarios—a critical yet relatively underexplored aspect of artificial intelligence research [4]. This requirement is especially pressing in HRI contexts, where maintaining human safety is of paramount importance.

In view of these considerations, the present study explicitly investigates the post-deployment behaviour of AI models, aiming to enable autonomous learning and effective adaptation in unfamiliar environments without prior explicit context. Our research highlights the essential first step in this adaptive process: accurately discerning whether encountered tasks represent genuinely novel or previously experienced scenarios. To achieve this, we propose an AI agent named Cognitive Agentic Probabilistic Adaptation AI (CAPA-AI), incorporating a Novelty Detection (ND) module alongside a pre-trained RL model. The ND module facilitates the

[1] K. Ghamati*, F. Amirabdollahian, and A. Zaraki* are with the School of Physics, Engineering and Computer Science (SPECS), and Robotics Research Group of the University of Hertfordshire, Hatfield, UK.

[2] DR. Faria is with the School of Science, Loughborough University, UK. {k.ghamati, a.zaraki}@herts.ac.uk, https://ghamati.com/CAPA-AI

effective identification and adaptation to novel situations. The performance of the proposed agent is evaluated through both simulated and real-world experiments involving a humanoid.

The subsequent sections presents the detailed design of our proposed agent and elaborate upon our novel methodology for novelty detection, which constitutes the core component underpinning the agent's adaptive capabilities.

II. Foundations of Continual Adaptation

A. Adaptation Challenges and Identified Gaps

Adapting AI models to diverse environments represents an active and rapidly progressing research domain, with significant implications across numerous applications. Continuous advancements and innovation remain crucial to enhancing model adaptability and generalisation, ensuring the effective deployment of AI systems within complex and dynamically evolving scenarios. Effective generalisation across diverse contexts necessitates AI frameworks capable of adapting seamlessly to generalised tasks—such as those encountered in radio communication systems—through centralised training mechanisms combined with decentralised data generation and management [5], [6]. Furthermore, techniques such as Transfer Learning (TL) and randomisation considerably augment model robustness, empowering agents to dynamically adapt to shifting environments [7], [8], [9].

Under these frameworks, AI models can leverage precise, real-time data, enabling sophisticated adaptive decision-making, particularly in contexts characterised by limited data availability. Approaches such as supervised learning, reinforcement learning, and multimodal learning have demonstrated substantial efficacy in customising adaptive strategies to heterogeneous groups, thereby achieving productive synergies and effectively managing inherent trade-offs [10], [11].

Despite rigorous training protocols, updating AI models post-deployment poses significant challenges, notably the phenomenon of catastrophic forgetting [2]. Catastrophic forgetting refers to models' propensity to lose previously acquired knowledge when integrating new information without adequate supervisory mechanisms [12]. This issue is especially pronounced during transitions from structured, supervised training environments to potentially unsupervised, real-world interactions, considerably increasing the risk of performance deterioration [13].

Extensive research underscores the necessity for robust continual adaptation strategies to mitigate knowledge loss during post-deployment model updates [14]. Thus, the development of resilient methodologies for continual adaptation represents a significant open research challenge, necessitating innovative approaches to facilitate autonomous learning within dynamic real-world settings [15]. The imperative for real-time learning and adaptive capabilities becomes particularly evident in HRI scenarios, where ensuring human safety remains paramount [16]. Real-time adaptation inherently carries risks, as agents may inadvertently demonstrate unsafe behaviours when confronted with unfamiliar circumstances [17]. Consequently, environmental feedback mechanisms become indispensable, serving not only to validate adaptive learning efficacy but also to ensure adherence to safety constraints [18]. Ultimately, the overarching ambition remains the development of autonomous systems capable of self-adaptation without persistent external supervision [19], a goal that highly relies on the agent's capability to identify and track novel events in a dynamic environment.

B. Novelty Detection in Artificial Intelligence

Effective ND is a prerequisite for successful adaptation in AI models, particularly within open-world environments, where adapting to unforeseen data is essential [20]. Novelty detection capabilities are critical for continual learning paradigms, enabling AI systems to incrementally assimilate new knowledge and maintain robust performance in response to environmental shifts [21]. This capability is particularly significant in class-incremental learning—a specialised domain within continual learning—where distinguishing accurately between known and unknown inputs ensures effective adaptation to novel tasks [22].

Historically, ND methods have predominantly focused on offline scenarios. However, adaptive environments characterised by continuous data inflow introduce substantial complexities into ND processes. The intricate relationship between ND and adaptive learning is particularly significant, as each considerably influences the performance of the other, highlighting the need for integrated frameworks capable of concurrently addressing both challenges [22].

Considering these intricate dynamics, this study introduces an advanced AI agent featuring a pre-trained RL model specifically engineered to capitalise effectively on environmental feedback, adhering rigorously to the constraints inherent in continual adaptation scenarios. The proposed agent utilises a sophisticated probabilistic approach to assess the novelty of incoming observations in relation to previously encountered situations. Through this methodology, the agent effectively leverages TL techniques, dynamically and efficiently updating its foundational knowledge [7], [23]. These innovations equip AI agents with enhanced capabilities to discern and respond accurately to novel scenarios, substantially improving their autonomous learning and adaptive capacities. These advancements aim to bridge the existing gap between theoretical potential and practical deployment, enhancing the adaptability, intelligence, and operational safety of AI systems within real-world environments.

III. Development of the Cognitive Agentic Model

In this section, we provide a detailed explanation of CAPA-AI model, presenting its mechanisms for detecting novelty and adapting to similar scenarios. Figure 1 presents the architectural outline of our Agentic AI, illustrating an independently trained RL model actively interacting within real-world environments. Such interactions enable the agent not only to respond effectively to familiar tasks but also, importantly, to refine its decisions based on feedback received.

Prior to executing any given task, the previously established Reinforcement Learning-based Attention Model

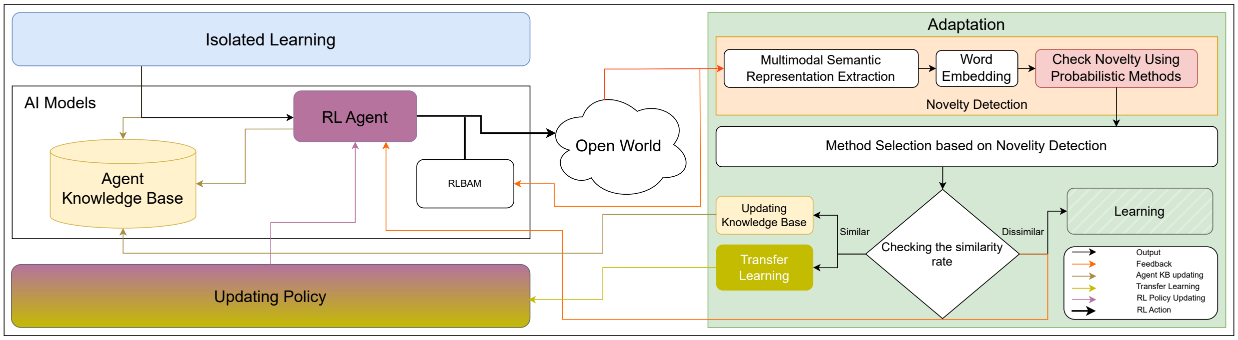


Fig. 1: **CAPA-AI Agentic AI Architecture** An RL agent is pre-trained using Isolated Learning, while the Agent Knowledge Base is concurrently established to store task-related data for Novelty Detection. During interaction with the open world, environmental feedback is processed by the Adaptation module to assess novelty. If a novel task exhibits sufficient similarity to previously learned tasks, the agent adapts using Transfer Learning, leveraging its existing knowledge base. Otherwise, feedback is passed to the RL agent to select an appropriate action.

(RLBAM) [24] is utilised to pinpoint salient features during interactions in real-world scenarios. This model efficiently directs the agent's focus towards key aspects pertinent to the primary task or environmental context, thereby significantly reducing the collection of noise and irrelevant data. When the ND module recognises an unfamiliar task, the agent engages TL methodologies to assimilate the new task, conditional upon its similarity to prior experiences. Subsequently, the RL policy undergoes updates incorporating strategies derived from the most analogous tasks. Concurrently, the Agent's Knowledge Base (AKB), which stores data from each encountered scenario, is updated to include information about this newly assimilated task, thus enhancing future ND capabilities.

The agent's framework is divided into two primary phases. The first phase involves constructing the AKB and pre-training the RL model through the Isolated Learning (IL) method, introducing initial tasks to the agent using a dedicated dataset. The second phase is the Adaptation stage, activated upon agent deployment. During the second phase, incoming observations are processed by the ND module. If the similarity between the new observation and existing AKB data exceeds 90%, the agent leverages its pre-trained RL model for decision-making. Conversely, if the similarity ranges between 60% and 89%, indicating a novel yet reasonably familiar situation or tasks, the agent employs TL to adapt accordingly. Thus, the ND module is integral to our AI agent's functionality; effective recognition of novelty empowers the agent to adapt proactively, fostering efficient learning and superior performance in subsequent interactions. Without reliable ND, indiscriminate processing of observations—without validation against existing knowledge—would result in unreliable outputs when facing tasks beyond the agent's existing understanding.

Upon detecting novelty, the subsequent essential step involves enabling the agent to adapt effectively. Our research outlines two distinct adaptation pathways: firstly, employing the AKB to facilitate adaptation if the novel task sufficiently mirrors previously encountered tasks; secondly, initiating exploratory behaviours to gather additional data in scenarios significantly different (with similarity below 60%). Within this study, we primarily emphasise the first approach, utilising the agent's existing knowledge base for task assimilation, while exploratory behaviours for unfamiliar scenarios are reserved for future research efforts.

A notable advantage of our proposed model lies in its environmental independence. Specifically, an RL agent pre-trained in a given environment can efficiently generalise its operation across diverse environments through the integrated use of ND and TL, rapidly adapting to new tasks encountered. To evaluate this capability, we utilised three distinct datasets: one dedicated to IL for initial construction of both the RL agent and AKB, and two additional datasets to assess the agent's adaptability through ND. Ultimately, the agent was deployed onto the ARI humanoid within a real-world context, with performance evaluations focusing on its ability to accurately respond to both familiar and novel, yet similar, tasks. The following subsections delve deeper into the architecture, particularly emphasising the ND module due to its central role within our proposed framework.

A. Isolated Learning

IL within AI and RL denotes the process of training an agent in an environment without direct interaction with other agents or external stimuli beyond pre-defined experiences. In our approach, we utilise IL to pre-train an RL agent using a dataset or simulations encompassing a limited and well-understood set of situations. This method is specifically adopted because accurately estimating novelty and effectively using similar prior actions require a robust foundational model exhibiting high reliability and accuracy. Such a model serves as the critical baseline during TL; therefore, it must demonstrate acceptable performance. In essence, integrating a reliable and potent base model and subsequently updating it for novel tasks directly influences the success of adaptation.

To facilitate the pre-training phase of the RL agent, we employ the RHM dataset [25], which comprises 26,804 video

Task	Robot Verbal Actions
Drinking	That looks delicious! What are you having?
	Enjoying a little break with a drink?
	Is that your favorite snack?
	How does that taste?
	Staying hydrated is important!
	That must be refreshing!
	Taking a quick energy boost?
	Do you prefer something sweet or savory?
	Looks like a good meal—hope you're enjoying it!
	Is this your usual choice?

TABLE I: Example set of categorised verbal responses generated by the RL agent for the Drinking task. The agent selects contextually appropriate dialogue to foster natural interactions within HRI scenarios.

clips spanning 14 distinct human activities. In line with AI literature conventions, each activity within the dataset is referred to as a *task*. The included tasks consist of walking, stretching, carrying objects, cleaning, closing cans, drinking, lifting objects, opening cans, putting down, reaching, sitting down, stairs down, stairs up, and standing up. Correspondingly, we constructed 14 distinct sets of robot's verbal actions aligned with each task, with each set comprising ten diverse sentences relevant to the particular activity. The agent is tasked with selecting the correct sentence set corresponding to a given activity and then choosing one sentence from the set to be employed as robot dialogue. For instance, Table I exemplifies one such set of the task *Drinking*, illustrating the ten related sentences. During operation, once the RL agent correctly identifies the relevant task set, it subsequently selects one sentence at random from this set. Consequently, the RL agent is explicitly trained to perform accurate task-set selections.

For the training process, we adopted proximal policy optimisation (PPO), a well-established Actor-Critic algorithm recognised for its effectiveness in model training [24]. The parameters of PPO have been meticulously tuned using the Optuna library[1], resulting in optimal values, including a policy learning rate of 0.000407, a value learning rate of 6.662×10^{-5}, a hidden layer size of 64 with a single layer, and a clip epsilon of 0.202. The discount factor γ was optimised to 0.911, the lambda parameter λ to 0.889, and both the entropy coefficient and value function coefficient were set to 0.0418. Training was conducted over 500 episodes, at which point the reward function converged to a stable value. Concurrently, we developed the AKB intended to store task-specific data utilised for ND. The details on the ND algorithm are provided in subsection III-B.1.

B. Adaptation

Interaction in real-world environments inherently involves uncertainty, as anticipating every possible event is impractical. It is therefore crucial, especially within HRI guidelines, that an agent is equipped to firstly identify novel occurrences and subsequently adapt by learning from these events to ensure safe decision-making. To address this challenge, we propose an agent architecture featuring a structured pipeline that systematically processes incoming observations. Initially, the pipeline evaluates these observations to detect novelty. Depending on the outcome of this novelty assessment, the observations are then routed to the appropriate subsequent layer, either for further learning or immediate decision-making. As illustrated in Figure 1 (highlighted in green), our model specifically handles the incoming feedback generated from RL interactions, enabling effective ND and adaptation. The following sub-section elaborates further on the adaptation capabilities of the proposed system.

1) Novelty Detection: To effectively detect novelty within our environment, we first process incoming data to extract essential information required to evaluate the current task context. Our approach begins by converting visual inputs, specifically frames extracted from video sequences, into meaningful textual descriptions. We utilised a vision-language large language model, LLaVA [26], to perform this conversion. Carefully designed prompts guide LLaVA to generate comprehensive environmental context descriptions for each selected video frame.

To ensure these descriptions remain concise and contextually relevant, we subsequently refine LLaVA's textual outputs using DeepSeek [27], a self-hosted language model specialised in contextual keyword extraction. DeepSeek distils the descriptions, selectively retaining keywords that capture essential contextual information while discarding irrelevant details, such as general descriptive phrases or prepositional terms. This refinement process produces a focused set of contextually pertinent keywords that accurately represent the current state of the environment.

These distilled keywords form the foundational input for our novelty detection framework. By analysing changes or deviations in these extracted keywords over time, our system can reliably identify and flag novel or previously unseen events within the monitored environment.

Our method employs probabilistic techniques to analyse keywords for ND, structured into two distinct phases. Firstly, during the IL phase, we extract frames from video clips in the RHM dataset, processing these through our ND pipeline. For each specific task, we compile a list of associated words. For instance, in the case of the *Drinking* task, based on frames processed by LLaVA and DeepSeek during pre-training, the agent generates a keyword list that includes terms such as *Drinking*, *Coffee*, *Stairs*, *Chair*, *Cup*, *Glass*, *Alcohol*, *Wine*, *Bar*, and *Beer*. Each word's frequency is assessed to determine its relevance. Specifically, *Drinking* emerges as the most frequent word, appearing 1900 times, and thus represents the task's centre of mass, or in probabilistic terms, the expected value. Following *Drinking*, *Coffee* appears 1400 times, *Stairs* 1200 times, *Chair* 800 times, and *Cup* and *Glass* each appear 600 times. Less frequent terms include *Alcohol* with 300 occurrences, *Wine* with 200, and *Bar* and *Beer*, each with 100 occurrences, indicating their decreasing relevance to the focal context of the *Drinking* task.

We also utilise this pipeline during real-world interactions following the agent's deployment. In this phase, frames extracted from the robot camera's captured video replace the

[1]https://optuna.org

dataset's video inputs. Subsequently, LLaVA and DeepSeek process these frames to extract textual data, converting the environmental context into descriptive words. These extracted words are then compared against the existing AKB, established during the IL phase, to identify instances of novelty. The following subsection outlines our proposed approach to novelty detection.

2) Assessing Novelty Using a Probabilistic Approach: This section is pivotal to the study, as ND serves as an essential prerequisite for the agent to initiate adaptation. The agent carefully analyses the extracted keywords, utilising conditional probability alongside the Jaccard Index [28] to compute the posterior probability. This calculated probability assesses whether newly encountered words are associated with any of the pre-trained tasks. Should the posterior probability exceed a predefined threshold ranging between 0.6 (60%) and 0.89 (89%), TL is employed to transfer actions from a related pre-trained task to the novel situation. To compute the posterior probability, each keyword is treated as a random variable. In every interaction, the sample space encompasses all potential words, covering both those already existing within the AKB and newly encountered terms. A uniform distribution is assumed, with its Probability Mass Function (PMF) employed to facilitate probability calculations. Following each interaction, the AKB undergoes updates, integrating new keywords from learned tasks and incrementally enhancing the frequency counts for recurring words. This updating mechanism is critical for accurately determining likelihood functions. ND evaluation at each interaction is performed using Eq. 1.

$$P(T_i \mid w_1, w_2, \ldots, w_n) \propto P(T_i) \cdot \prod_{j=1}^{n} P(w_j \mid T_i) \tag{1}$$

Here, $P(T_i)$ represents the prior probability, and $P(w_j \mid T_i)$ is the likelihood function, expressing the probability that w_j belongs to T_i. Our experiments demonstrate that, due to the dynamic update of the AKB, the number of words per task varies. As the PMF of a uniform distribution is computed by dividing the number of occurrences of a word by the total number of words in a task (Eq. 4), tasks that are performed more frequently see their probabilities decrease over time, leading to potentially unreliable results. To address this, we apply the Jaccard Index as a prior probability, neutralising the effect of uneven word list sizes. The revised computation using the Jaccard Index as the prior probability is given in Eq. 2.

$$P(T_i \mid w_1, w_2, \ldots, w_n) \propto J(T_i, W) \cdot \left(\prod_{j=1}^{n} P(w_j \mid T_i) \right) \tag{2}$$

The Jaccard Index $J(T_i, W)$ is calculated using Eq. 3, where $\text{intersect}(T_i, W)$ denotes the number of new words present in T_i, and $\text{union}(T, W)$ is the total number of unique words, combining the agent's knowledge base and newly encountered words.

$$J(T_i, W) = \frac{|\text{intersect}(T_i, W)|}{|\text{union}(T, W)|} \tag{3}$$

To compute the posterior probability using the chain rule, we first calculate the Jaccard Index (Eq. 3). Then, utilising Bayes' rule, we transform $P(T_j \mid w_j)$ into $P(w_j \mid T_j)$ as shown in Eq. 5. Here, $P(w_j \mid T_i)$ is computed according to Eq. 4, and $P(T_i)$, being equal for all tasks, is computed via Eq. 6.

$$P(w_j \mid T_i) = \frac{\text{Total frequency of } w_j \text{ in task } T_i}{\text{Total frequency of all words in task } T_i} \tag{4}$$

$$P(T_i \mid w_j) = \frac{P(w_j \mid T_i) \cdot P(T_i)}{\sum_j P(w_j \mid T_j) \cdot P(T_j)} \tag{5}$$

$$P(T_i) = \frac{1}{\text{Number of Tasks}} \tag{6}$$

Upon evaluating Eq. 2, the posterior probability obtained is forwarded to the method selection module, wherein it is assessed for novelty relative to the agent's existing knowledge base. Should the observation be classified as novel, yet exhibit a similarity ranging between 60% and 90% with a previously learnt task, TL is employed. This mechanism enables the agent to leverage actions associated with the known task, thereby enhancing its initial approach to the novel task. Consequently, the RL agent is positioned advantageously to make informed decisions, facilitating effective action selection and promoting policy refinement driven by subsequent feedback. Thus, the RL policy undergoes appropriate updates, applying insights from the previously acquired task policy to the newly identified task.

IV. Proof of Concept Evaluation

We conducted an evaluation of the proposed CAPA-AI framework in two distinct phases: firstly, within a simulated environment, and subsequently, in real-world HRI scenarios. In the first stage, following the pre-training of the RL agent and the establishment of the initial version of the AKB using a comprehensive dataset, we evaluated the agent's performance with two distinct and separate datasets. In the second stage, the agent's effectiveness was assessed through a practical HRI scenario in which experimenters interacted directly with a social robot. Detailed descriptions of the evaluation methodologies and findings are presented in the following sections.

A. Isolated Learning Using the RHM Dataset

Utilising the RHM dataset [25], we undertook IL to pretrain the agent and establish its foundational knowledge base. Although our model targets scenarios involving two experimenters, the original RHM dataset frequently features frames with only a single person. To mitigate this limitation and ensure comprehensive coverage of two-person scenarios, we employed the method proposed by [24], leveraging randomisation and generalisation techniques to synthesise additional states featuring a second individual. Concurrently, each frame underwent processing within our pipeline to convert visual information into task-specific textual descriptors, thereby enriching the agent's knowledge base.

B. Evaluation Using Two Distinct Datasets

Upon completion of the IL stage, our framework underwent evaluation against two separate datasets: Eatsense, which includes 16 tasks [29], and Polar, encompassing 9 tasks [30]. The evaluation commenced with the extraction of frames from videos corresponding to each task, subsequently processing them through our established pipeline. Post-processing, the ND module assessed whether incoming frames represented novel tasks when compared against the existing AKB. If novelty was detected, TL was initiated to update both the RL agent's policy and its knowledge base. Conversely, frames demonstrating similarity rates exceeding 90% to previously learned tasks prompted updates to the RL agent's policy without invoking TL, due to the high confidence in task familiarity.

C. Preliminary Human-Robot Interaction Test

To further validate the agent's performance, we implemented it on the ARI robot across various real-world scenarios comprising familiar and novel tasks. Accuracy was measured based on the ratio of correct decisions against the total tasks presented. Figure 4 illustrates three distinct experimental scenarios.

In the first scenario, an experimenter performed the familiar action of drinking water, accurately identified by CAPA-AI (Figure 4a). The second scenario introduced a novel action—using a tissue—categorised by the agent as *Cleaning*, demonstrating effective novelty adaptation (Figure 4b). A third scenario involved lying down, with sufficient similarity to the known task *Sitting Down*, thus prompting the corresponding adaptive response (Figure 4c).

Each figure illustrates the comprehensive processing workflow of observations collected during interactions. Frames from ARI's head camera were processed via CAPA-AI's perception pipeline employing LLaVA and DeepSeek models to extract relevant keywords.

V. Results and Discussion

The performance of the proposed CAPA-AI framework was rigorously evaluated through both simulation-based and real-world experiments, demonstrating its capacity to adaptively respond to dynamic task variations.

A. Simulation and Dataset Evaluation

In the initial evaluation stage, we assessed ND accuracy by calculating the proportion of correctly identified novel tasks relative to the total number of processed frames across simulation environments and benchmark datasets, specifically RHM, EatSense, and Polar. For instance, frames corresponding to the *Eating* task were recognised as novel yet contextually similar to the previously learned *Drinking* task. This similarity triggered the TL module to reuse behavioural policies initially acquired for *Drinking*, thereby demonstrating effective task generalisation capabilities.

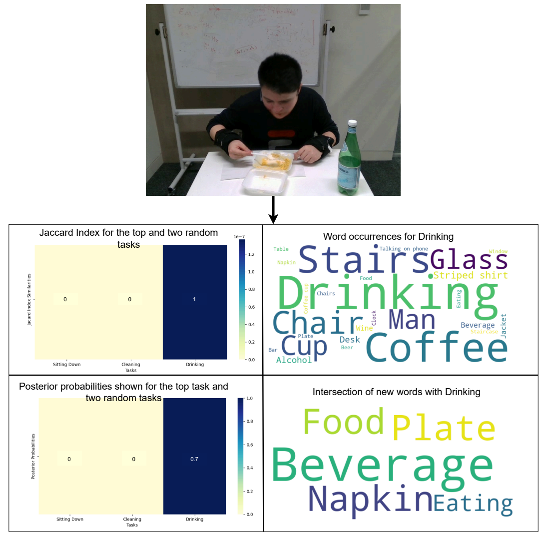


Fig. 2: Result of Novelty Detection for a sampled frame

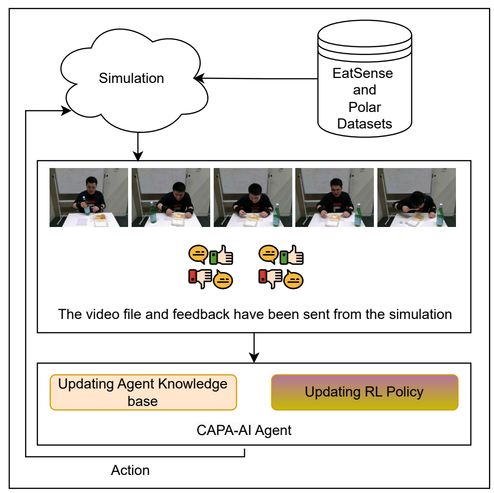
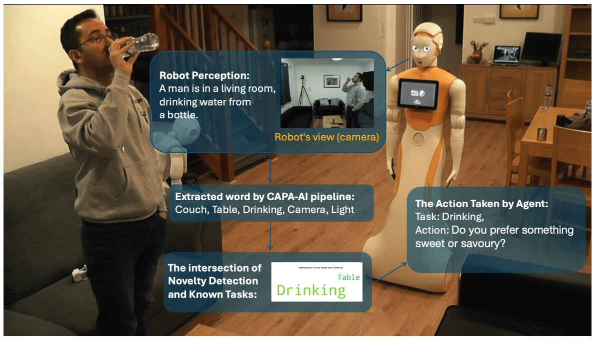


Fig. 3: Showing how Evaluating the CAPA-AI using Polar and EatSense datasets

Figure 3 illustrates the evaluation pipeline. Initially, the agent undergoes pre-training to establish the AKB and the RL model. Subsequently, a simulation environment built using Gymnasium[2] was configured to process video clips sourced from the EatSense and Polar datasets. Following this, frames were extracted using a scene-change detection algorithm, combined with feedback from the last executed action, and passed to the CAPA-AI module. CAPA-AI then estimated novelty and refined the RL policy based on the received environmental feedback.

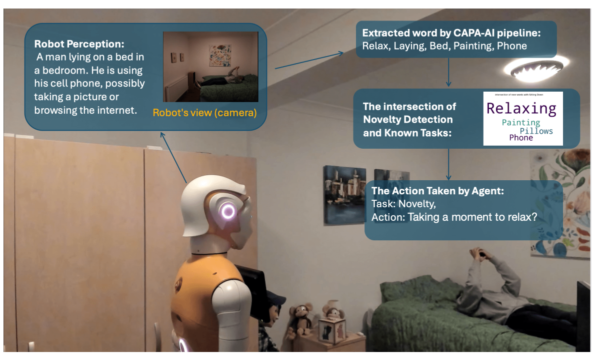
A critical advantage of employing RL in this context is its ability to leverage environmental feedback to improve ND accuracy dynamically. Specifically, if the ND module estimates the similarity between a novel and known task

[2]https://gymnasium.farama.org/



(a) Task: Drinking, CAPA-AI Decision: Drinking

(b) Task: Using Tissue, **CAPA-AI Decision: Novelty, Similar Task: Cleaning**



(c) Task: Laying, **CAPA-AI Decision: Novelty, Similar Task: Sitting Down**

Fig. 4: **Real-world results** Three samples of CAPA-AI decision base on the real task.

within a range of 60% to 89%, but this estimate proves incorrect, subsequent environmental feedback from executing an action related to the known task will rectify the misclassification. This interactive feedback loop enables the agent to adapt continuously, enhancing its performance during task interactions.

Additionally, Figure 2 presents an illustrative example where the ND module accurately identified task similarity. Here, the Jaccard Index demonstrated a 100% overlap in linguistic descriptors with the *Drinking* task, resulting in posterior probabilities indicating a 70% likelihood of task equivalence. Despite environmental variability, the shared linguistic descriptors significantly facilitated accurate detection and effective adaptation, culminating in an overall novelty detection and adaptation accuracy of 89% during the simulation stage.

Furthermore, the integration of the Jaccard Index into the posterior probability computation significantly influences outcomes when task similarity falls below 100%. This methodology aids in moderating the influence of task-specific keyword counts, ensuring the posterior probability reflects genuine keyword relevance and avoids potential overfitting or underfitting by appropriately scaling the significance of keyword intersections between novel and known tasks.

B. Real-World HRI Evaluation

To assess the real-world applicability of the proposed framework, the CAPA-AI system was initially pre-trained using the RHM dataset and subsequently deployed on the ARI robot during a 30-minute HRI session. This session incorporated a variety of randomly generated activity scenarios designed to evaluate the agent's adaptive capabilities in realistic conditions. The ND module continuously processed incoming data streams in real-time, assessing task familiarity using probabilistic measures. Adaptation responses were strategically determined by these probability scores: values ranging between 60% and 89% initiated TL-based adaptation, whereas scores exceeding 90% prompted immediate responses through direct RL.

Under these experimental conditions, the CAPA-AI framework successfully achieved an adaptation accuracy of 80%. The agent consistently demonstrated robustness, effectively managing several environmental complexities, including ambiguous task descriptions, significant background noise, and sensor variability.

All experiments were performed using a high-performance workstation equipped with an NVIDIA GeForce RTX 4090 GPU (24 GB GDDR6X), ensuring efficient real-time processing of Robot Operating System (ROS) messages from the ARI robot. This computational setup enabled rapid task-switching capabilities and seamless adaptation, essential for dynamic and interactive real-world scenarios.

C. Observations and Challenges

While the overall performance confirms the adaptability and robustness of CAPA-AI, several key challenges emerged. Firstly, as the vocabulary in the knowledge base grows, the number of likelihood functions required for posterior estimation increases, raising computational complexity. Secondly, semantically related descriptors occasionally led to false-positive task matches. Although training on diverse datasets mitigated this effect, future work must further enhance contextual disambiguation and consider hierarchical task representations to improve precision and scalability.

VI. Conclusion

This study proposed CAPA-AI, a novel agentic framework designed to support continual adaptation in open-ended, real-world environments. The system integrates probabilistic novelty detection with transfer learning mechanisms, enabling autonomous agents to identify task novelty and repurpose prior knowledge where appropriate.

Comprehensive simulation-based experiments, supported by deployment on the ARI robot, demonstrated effective novelty detection (89% accuracy) and robust adaptation under real-world HRI conditions (80% accuracy). These findings highlight the practical utility and resilience of CAPA-AI in dynamic and unpredictable environments.

Nevertheless, the framework presents some limitations. Increasing knowledge base size elevates computational demands due to the proliferation of likelihood evaluations. Additionally, semantic ambiguity in task descriptors can impair accuracy during task recognition. While partially mitigated through dataset diversity, these challenges require further attention.

Future work will aim to address these limitations by: (i) improving computational efficiency through model optimisation, (ii) enhancing task disambiguation via hierarchical or contextual modelling, and (iii) investigating unsupervised learning approaches to enable adaptation in entirely dissimilar and previously unseen tasks.

In summary, CAPA-AI offers a robust, scalable, and context-aware approach to novelty detection and continual task adaptation, providing a strong foundation for long-term autonomous operation in real-world settings.

VII. Acknowledgment

ChatGPT (OpenAI, San Francisco, CA, USA) was utilised to assist in drafting and refining sections of the manuscript. The authors have reviewed and approved all AI-assisted content to ensure accuracy and integrity.

References

[1] A. Zaraki, M. Khamassi, L. J. Wood, G. Lakatos, C. Tzafestas, F. Amirabdollahian, B. Robins, and K. Dautenhahn, "A novel reinforcement-based paradigm for children to teach the humanoid kaspar robot," *International Journal of Social Robotics*, vol. 12, pp. 709–720, 2020.

[2] S. Mazumder and B. Liu, "Lifelong and continual learning dialogue systems," *arXiv preprint arXiv:2211.06553*, 2022.

[3] R. Inamdar, S. K. Sundarr, D. Khandelwal, V. D. Sahu, and N. Katal, "A comprehensive review on safe reinforcement learning for autonomous vehicle control in dynamic environments," *e-Prime-Advances in Electrical Engineering, Electronics and Energy*, p. 100810, 2024.

[4] G. Kim, C. Xiao, T. Konishi, Z. Ke, and B. Liu, "Open-world continual learning: Unifying novelty detection and continual learning," *Artificial Intelligence*, vol. 338, p. 104237, 2025.

[5] J. Aldrich, D. Garlan, C. Kästner, C. Le Goues, A. Mohseni-Kabir, I. Ruchkin, S. Samuel, B. Schmerl, C. S. Timperley, M. Veloso *et al.*, "Model-based adaptation for robotics software," *IEEE software*, vol. 36, no. 2, pp. 83–90, 2019.

[6] K. Ghamati, M. Banitalebi Dehkordi, and A. Zaraki, "Towards ai-powered applications: The development of a personalised llm for hri and hci," *Sensors*, vol. 25, no. 7, p. 2024, 2025.

[7] B. Liu, S. Mazumder, E. Robertson, and S. Grigsby, "Ai autonomy: Self-initiated open-world continual learning and adaptation," *AI Magazine*, vol. 44, no. 2, pp. 185–199, 2023.

[8] A. Zaraki, D. Mazzei, M. Giuliani, and D. De Rossi, "Designing and evaluating a social gaze-control system for a humanoid robot," *IEEE Transactions on Human-Machine Systems*, vol. 44, no. 2, pp. 157–168, 2014.

[9] M. Shahabian Alashti, K. Ghamati, H. Samani, and A. Zaraki, "Towards memory-driven agentic ai for human activity recognition," in *Social Robotics: 17th International Conference, ICSR 2025, Naples, Italy, September 10-12, 2025, Proceedings 17*. Springer, 2025, pp. 1–8.

[10] S.-M. Cheong, K. Sankaran, and H. Bastani, "Artificial intelligence for climate change adaptation," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 12, no. 5, p. e1459, 2022.

[11] L. Cominelli, D. Mazzei, M. Pieroni, A. Zaraki, R. Garofalo, and D. De Rossi, "Damasio's somatic marker for social robotics: preliminary implementation and test," in *Biomimetic and Biohybrid Systems: 4th International Conference, Living Machines 2015, Barcelona, Spain, July 28-31, 2015, Proceedings 4*. Springer, 2015, pp. 316–328.

[12] J. Kirkpatrick, R. Pascanu, N. Rabinowitz, J. Veness, G. Desjardins, A. A. Rusu, K. Milan, J. Quan, T. Ramalho, A. Grabska-Barwinska *et al.*, "Overcoming catastrophic forgetting in neural networks," *Proceedings of the national academy of sciences*, vol. 114, no. 13, pp. 3521–3526, 2017.

[13] H. Li, P. Barnaghi, S. Enshaeifar, and F. Ganz, "Continual learning using bayesian neural networks," *IEEE transactions on neural networks and learning systems*, vol. 32, no. 9, pp. 4243–4252, 2020.

[14] J. J. Lee, S. I. Lee, and H. Kim, "Continual learning for instance segmentation to mitigate catastrophic forgetting," in *2021 18th International SoC Design Conference (ISOCC)*. IEEE, 2021, pp. 85–86.

[15] A. Gupta, S. Savarese, S. Ganguli, and L. Fei-Fei, "Embodied intelligence via learning and evolution," *Nature communications*, vol. 12, no. 1, p. 5721, 2021.

[16] S. Gu, A. Kshirsagar, Y. Du, G. Chen, J. Peters, and A. Knoll, "A human-centered safe robot reinforcement learning framework with interactive behaviors," *Frontiers in Neurorobotics*, vol. 17, p. 1280341, 2023.

[17] S. Di Falco and F. M. Vieider, "Environmental adaptation of risk preferences," *The Economic Journal*, vol. 132, no. 648, pp. 2737–2766, 2022.

[18] Z. Chen and B. Liu, *Lifelong machine learning*. Morgan & Claypool Publishers, 2018.

[19] A. Ayub, Z. Francesco, P. Holthaus, C. L. Nehaniv, and K. Dautenhahn, "Continual Learning through Human-Robot Interaction – Human Perceptions of a Continual Learning Robot in Repeated Interactions," *International Journal of Social Robotics (**Springer**)*, 2025.

[20] R. Stern, W. Piotrowski, M. Klenk, J. de Kleer, A. Perez, J. Le, and S. Mohan, "Model-based adaptation to novelty for open-world ai," in *Proceedings of the ICAPS Workshop on Bridging the Gap Between AI Planning and Learning*, 2022.

[21] M. A. Pimentel, D. A. Clifton, L. Clifton, and L. Tarassenko, "A review of novelty detection," *Signal processing*, vol. 99, pp. 215–249, 2014.

[22] R. Aljundi, D. O. Reino, N. Chumerin, and R. E. Turner, "Continual novelty detection," in *Conference on Lifelong Learning Agents*. PMLR, 2022, pp. 1004–1025.

[23] M. B. Dehkordi, A. Zaraki, and R. Setchi, "Feature extraction and feature selection in smartphone-based activity recognition," *Procedia Computer Science*, vol. 176, pp. 2655–2664, 2020.

[24] K. Ghamati, A. Zaraki, and F. Amirabdollahian, "Ari humanoid robot imitates human gaze behaviour using reinforcement learning in real-world environments," in *2024 IEEE-RAS 23rd International Conference on Humanoid Robots (Humanoids)*. IEEE, 2024, pp. 653–660.

[25] M. Bamorovat Abadi, M. R. Shahabian Alashti, P. Holthaus, C. Menon, and F. Amirabdollahian, "Rhm: Robot house multi-view human activity recognition dataset," in *ACHI 2023: The Sixteenth International Conference on Advances in Computer-Human Interactions*. IARIA, 2023.

[26] H. Liu, C. Li, Q. Wu, and Y. J. Lee, "Visual instruction tuning," *Advances in neural information processing systems*, vol. 36, pp. 34 892–34 916, 2023.

[27] D. Guo, D. Yang, H. Zhang, J. Song, R. Zhang, R. Xu, Q. Zhu, S. Ma, P. Wang, X. Bi *et al.*, "Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning," *arXiv preprint arXiv:2501.12948*, 2025.

[28] R. Real, "Tables of significant values of jaccard's index of similarity," *Miscel· lania Zoologica*, pp. 29–40, 1999.

[29] M. A. Raza, L. Chen, L. Nanbo, and R. B. Fisher, "Eatsense: human centric, action recognition and localization dataset for understanding eating behaviors and quality of motion assessment," *Image and Vision Computing*, vol. 137, p. 104762, 2023.

[30] W. Ma and S. Liang, "Polar: Posture-level action recognition dataset," in *2019 6th International Conference on Systems and Informatics (ICSAI)*. IEEE, 2019, pp. 427–433.