Task 1: Prototypes, Literature, Gaps, and Research Questions

Comparing Techniques for Handling Knowledge Degradation in Modular Reinforcement Learning Under Rapid Task Switching with Noise Injection

Research Questions:

1. *How can modular architectures, utilizing distinct submodules for task-specific processing, effectively generalize across tasks when exposed to different noise types (e.g., stochastic noise, deliberate misinformation), and what strategies can these architectures employ to maintain stability and prevent catastrophic forgetting during rapid task switching?*
2. *How can techniques such as knowledge distillation, regularization, or dynamic memory modules be compared in their ability to handle the degradation of learned knowledge in modular RL agent(s) when faced with rapid task switching across varying noise conditions, including deliberate misinformation and stochastic noise?*

Environment Selection:

1. **Acrobot-v1 (Control Environment):** "Acrobot-v1 is a classic control problem with a discrete action space, suitable for demonstrating rapid task switching and analyzing RL agent behavior under noise. This environment facilitates the implementation of modular RL techniques like knowledge distillation and regularization, particularly in scenarios where noise complicates control.
2. **BipedalWalker-v3 (Any Category Environment):** BipedalWalker-v3's environment provides a more complex continuous action space and involves higher-dimensional state features. It challenges the RL agent with tasks that require coordination and adaptability, making it suitable to explore generalization across tasks and how modular submodules can maintain stability under noisy conditions. The complexity allows for testing how modular architectures can generalize across tasks, especially when faced with stochastic noise or deliberate misinformation.

Brief Summary of Prototypes

- **Heavy Noise & Reward Injection Model:** The Heavy Noise & Reward Injection implementation for Acrobot-v1, though still in progress, has shown that parameter noise provides better exploration than Action Space Noise, with ongoing adjustments needed to optimize hyperparameters for favorable results Currently achieving rewards between -1000 to -8000, the environment has not yet been solved but shows progress in handling the challenging conditions of the task-switching setup. The current comparison models do not directly inject noise; rather, they test under noisy conditions to examine robustness. This has led to some interesting observations about the importance of gradual goal modifications and parameter tuning in ensuring stable learning. As part of ongoing development, multi-agent testing and noise injection between agents to further evaluate and enhance the system's resilience.
- **Comparison Model:** The comparison model consistently achieves rewards around -100, close to the solved state of -85. Further optimization is needed to improve stability under noisy conditions and rapid task switching. Simple task switching generalizes faster with smaller batch sizes but fails to solve the problem completely.
- **Failed Prototypes**: Attempts with a Q-Table approach for MountainCarContinuous-v0 and a QLearning approach for Acrobot-v1 failed, showing poor convergence in complex environments, with rewards below expected thresholds. Task switching in Acrobot proved more rigid and generalized compared to MountainCar.

Core Papers & Explanations on the Weaknesses/Gaps of the Papers

Each of the selected studies has not adequately addressed the challenges presented by dynamic noise or rapid task switching. To the best of my knowledge, current papers tackle:

1. **Modular Deep Learning (Scherer & Carter, 2023)**

   - **Weakness**: The study by Scherer and Carter effectively uses modular architectures to mitigate issues like task interference, thus preventing catastrophic forgetting. However, it mainly focuses on standard multi-task reinforcement learning settings, where different tasks are trained separately without interference. The paper does not delve into scenarios involving environmental noise or deliberate misinformation. Such noise is common in real-world applications and can challenge the stability of learning systems.
   - **Gap**: Due to the lack of investigation into noisy scenarios and deliberate misinformation, there is a gap in understanding how these modular architectures can maintain stability when exposed to such challenging conditions, especially during rapid task switching.
   - **Reason for Selection**: This paper provides foundational insights into modular architectures for multi-task RL, which are crucial for developing systems capable of handling task interference and preventing catastrophic forgetting.

Task 1: Prototypes, Literature, Gaps, and Research Questions

2. **Modular Networks Prevent Catastrophic Interference in Model-Based Multi-Task Reinforcement Learning (Bu & Malik, 2022)**
   - **Weakness**: Bu and Malik show that modular networks effectively isolate task-specific learning in stable environments, preventing task interference. However, the study lacks exploration of the impact of dynamic noise or rapid task switching, focusing only on controlled multi-task environments.
   - **Gap**: The absence of analysis in dynamic noisy environments and rapid task switching leaves a significant gap. It remains unclear how modular architectures maintain learned knowledge when faced with noise and switching conditions.
   - **Reason for Selection**: This paper addresses modular networks and their ability to prevent task interference, which directly aligns with the goal of improving modular RL stability under rapid task switching.
3. **Overcoming Reward Model Noise in Instruction-Guided Reinforcement Learning (Chekanov, 2022)**
   - **Weakness**: Chekanov's study addresses noisy reward signals in instruction-guided RL. However, the focus is primarily on noise at the reward signal level and does not extend to scenarios involving task-specific misinformation or rapid task switching between different goals. This limits the paper's applicability to modular systems that need to handle deliberate misinformation and ensure stability across rapidly changing tasks.
   - **Gap**: The absence of an exploration into task-specific misinformation during rapid switching highlights a gap in the understanding of how modular RL architectures can maintain stability and effectively adapt to sudden changes. This gap necessitates further exploration into strategies that could help modular systems cope with misinformation and environmental noise while switching between tasks.
   - **Reason for Selection**: This paper prvides insights into handling noisy reward signals, which is essential for understanding the challenges of robustness in modular RL systems, especially under dynamic noise conditions.

Explanation and Justification of Selected Gaps

The selected gaps in the literature focus on the stability and knowledge retention of modular RL systems under rapid task switching and dynamic noise conditions, including deliberate misinformation. These gaps were chosen due to the following limitations identified in the current literature:

1. **Current Focus on Standard Multi-Task Settings**: The existing studies focus primarily on stable, well-defined multi-task environments, where tasks are largely isolated, and the complexity of real-world conditions, such as noise and misinformation, is not accounted for. This results in a lack of understanding about how modular RL systems can handle frequent task switching and adapt to dynamic noise.
2. **Single-Agent Noise Robustness**: Although there has been work on making single-agent RL systems robust to noisy conditions, this research does not extend to modular architectures that need to learn from multiple tasks and adapt under both noise and rapid switching scenarios. The knowledge degradation and how to retain previously learned skills remain largely unexplored when these two challenges are combined.
3. **Gaps in Task-Specific Performance and Knowledge Degradation**: The literature lacks an in-depth exploration of task-specific performance of modular architectures under conditions involving both dynamic noise and rapid task switching. Furthermore, no studies have yet thoroughly investigated how knowledge retention can be ensured when agents must transition between tasks in noisy environments, which is crucial for avoiding catastrophic forgetting.

Addressing these gaps is critical to building resilient RL systems.

Explanation and Justification of Research Questions Grounded in the Gaps

Both research questions are crucial for ensuring modular systems retain acquired skills while adapting to new tasks in dynamic environments involving noise and misinformation

1. **Research Question 1** is grounded in the gap regarding the resilience of modular RL architectures in dynamic, noisy environments with frequent task switching. The question aims to investigate how task-specific processing can help generalize learning across tasks despite the challenges of stochastic noise and misinformation. Preventing catastrophic forgetting an area that existing literature has not adequately addressed through rapid task switching.
2. **Research Question 2** directly addresses the knowledge degradation gap. It compares different techniques (e.g., knowledge distillation, regularization) to determine their effectiveness in retaining learned knowledge during rapid task switching under noise and misinformation. This comparative approach aims to establish which strategy is most effective in maintaining stability in modular RL systems, a topic that has not been sufficiently explored.

These research questions are critical for building robust and adaptable modular RL systems capable of handling real-world complexity, where agents must navigate multiple tasks, noisy environments, and rapid transitions without losing previously gained skills.

Task 1: Prototypes, Literature, Gaps, and Research Questions

References

1. Scherer, C., & Carter, E. (2023). Modular Deep Learning. Retrieved from https://doi.org/10.48550/arXiv.2302.11529
2. Bu, W., & Malik, A. (2022). Modular networks prevent catastrophic interference in model-based multi-task reinforcement learning. Retrieved from https://doi.org/10.1007/978-3-030-95470-3_23
3. Xu, R., & Harris, M. (2022). Robust multi-agent reinforcement learning for noisy environments. Retrieved from https://doi.org/10.1007/s12083-021-01133-2
4. Dreisbach, G. (2014). Task-specific effects of reward on task switching. Psychological Research, 78(6), 733-746. Retrieved from https://doi.org/10.1007/s00426-014-0595-z
5. S.V. Chekanov (2022). Overcoming Reward Model Noise in Instruction-Guided Reinforcement Learning. Retrieved from https://doi.org/10.48550/arXiv.2202.11529