**EE422/CS421Introduction to Robotics**

**Spring 2025**

Group Members

Fadime Yaren Durmuş 200201003

Mustafa Kağan Aytaç 200202026

**Report – Presentation 2**

**Multi-Agent Reinforcement Learning and Battery-Aware Robotics**

Multi-Robot Systems (MRS) are evolving beyond mere task execution into systems that can learn, adapt, and collaborate in uncertain and dynamic environments. At the heart of this transformation lies Multi-Agent Reinforcement Learning (MARL), which enables robots to learn policies not only by interacting with the environment but also by considering the presence and actions of other agents in the system. Unlike single-agent reinforcement learning, MARL introduces the challenge and opportunity of emergent behaviors like cooperation, negotiation, and collective reasoning.

In MARL, agents aim to optimize long-term reward through trial-and-error interactions. One of the most fundamental structures is Independent Q-Learning, where each agent learns as if others are static. While simple, this method is often unstable in truly dynamic settings. More effective frameworks, such as Centralized Training with Decentralized Execution (CTDE), offer a balance between global awareness and local autonomy. Algorithms like MADDPG (Multi-Agent Deep Deterministic Policy Gradient), QMIX, and MAAC exemplify this approach, allowing agents to be trained with shared knowledge but operate using only local observations.

Communication further enhances coordination. Learned communication protocols, such as those found in CommNet and TarMAC, allow agents to share selective, learned representations. Attention-based systems, like MAAC, help agents determine which peers are relevant at any given moment. Additionally, graph-based MARL methods utilize Graph Neural Networks (GNNs) to structure inter-agent influence in dynamic interaction networks, enabling agents to learn through message-passing in complex environments.

One striking real world use case is the simulation of pedestrian groups. In MARL-Ped, pedestrian-like agents are trained to navigate without collisions by observing and mimicking human movement patterns. The result is an emergent swarm of agents that move fluidly, avoiding congestion and deadlock in crowded spaces. Such approaches not only demonstrate MARL's applicability in human-inspired environments but also pave the way for robots to operate seamlessly in populated areas like airports or shopping malls.

While learning collective behavior is critical, sustaining it requires energy. Battery charging becomes not a technical afterthought but an intelligent decision process. Robots need to learn when to continue a task and when to detour for energy. This introduces a second reinforcement learning problem: one rooted in foresight and prioritization.

In energy-aware robotics, reward shaping plays a foundational role. Agents receive penalties for energy depletion and rewards for timely charging. Safe reinforcement learning ensures that policies adhere to energy constraints, avoiding dangerous depletion. Multi-objective reinforcement learning treats task completion and battery conservation as parallel goals, requiring agents to trade short term gains for long-term sustainability. Hierarchical models allow a high-level controller to make charging decisions while delegating the implementation to lower-level controllers.

Auction-based charging algorithms also offer a practical solution when multiple agents must share limited charging stations. Methods like the Contract Net Protocol or Hungarian Auction are often coupled with MARL strategies, where agents learn to place bids based on urgency, task progress, and predicted future value. Charging thus becomes a distributed coordination problem, deeply embedded in the learning architecture.

Together, MARL and battery charging algorithms define not only how robots learn to act, but how they learn to survive. These systems are no longer just reactive; they’re proactive, strategic, and self-preserving. As robots increasingly learn to balance their roles as individuals and collaborators, their ability to reason about energy, time, and interaction moves us closer to truly autonomous, adaptive, and sustainable robotic ecosystems.