Shyam Sunder Rajasekaran

Multi-Agent Reinforcement Learning

Environment Configuration

- **Grid Structure**: 10×10 grid with randomly placed obstacles (1-6) and a single target
- **Agent Observations**: Each agent observes a local 3×3 window centered on its position
- **Communication Channel**: Agents can exchange a continuous scalar value (-1 to 1)
- **Action Space**: 5 discrete actions (Up, Down, Left, Right, Stay)
- **Reward Structure**: +10 when both agents reach the target simultaneously, otherwise 0
- **Termination Conditions**: Success (both agents at target) or maximum steps (50) reached

Network Architecture

Each agent utilizes a Deep Q-Network (DQN) with the following architecture:

- **Input Layer**: 10 neurons (9 for the 3×3 observation window + 1 for the communication signal)
- **Hidden Layers**: Two fully-connected hidden layers with 64 neurons each and ReLU activation
- **Output Branches**:
    - Action branch: 5 outputs representing Q-values for each action
    - Communication branch: 1 output with tanh activation (constraining values to [-1, 1])

Training Hyperparameters

The agents were trained using the following hyperparameters:

Parameter	Value	Description
Batch Size	64	Number of transitions per optimization step
Discount Factor (γ)	0.99	Weight for future rewards
Learning Rate	0.001	Step size for Adam optimizer

Epsilon Start	1.0	Initial exploration rate
Epsilon End	0.1	Final exploration rate
Epsilon Decay	0.995	Multiplicative factor for epsilon decay
Replay Buffer Size	10,000	Capacity of experience replay buffer
Target Network Update	Every 10 episodes	Frequency of target network updates
Hidden Layer Size	64 neurons	Width of hidden layers
Maximum Episodes	1,000	Training termination ceiling

Early stopping was implemented with a threshold of 95% success rate over the last 100 episodes.

Training Results

Training performance was evaluated using two key metrics:

1. **Episode Rewards**: Cumulative reward per episode
2. **Success Rate**: Moving average (window=100) of target acquisition success