Comparative Evaluation of Dynamic Programming on Frozen Lake Environments

Yash Bankar
Roll No: 24034017

Comparative Evaluation of DP on Frozen Lake Environments

Objective

We evaluate and compare the performance of dynamic programming (value iteration) on three environments:

- **FrozenLake-v1**: The standard 4x4 stochastic environment provided by Gymnasium.
- **FrozenLake-Custom-v0**: A smaller or deterministic custom map.
- **FrozenLake-Expanded-v0**: A larger or more complex variation.

Metrics Compared

Metric	**Description**
Iterations	Number of iterations required for value iteration to converge.
Time (s)	Time taken for convergence.
Avg Reward	Average reward per episode over 100 evaluations.
Avg Episode Length	Mean number of steps taken per episode.
Value Variance	Variance in value function estimates across all states.

Table 1: Evaluation Metrics for Comparison

Results

Environment	Iterations	Time (s)	Avg Reward	Avg Ep. Length	Value Varian
FrozenLake-v1	324	0.0382	0.767	42.94	0.08362
FrozenLake-Custom-v0	7	0.0010	1.000	6.00	0.17942
FrozenLake-Expanded-v0	11	0.0010	1.000	10.00	0.16982

Table 2: Performance Metrics Across Environments

Visual Comparison

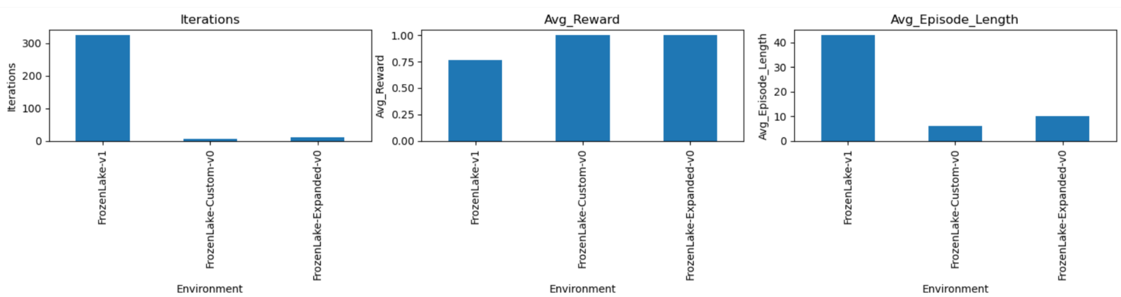

Figure 1: Bar plots comparing Iterations, Average Reward, and Episode Length across environments

Observations

- **FrozenLake-v1** required significantly more iterations to converge, likely due to its stochastic (slippery) nature.
- **Custom environments**, being deterministic, converged in fewer iterations and achieved optimal rewards consistently.
- Episode lengths were shortest in simpler maps, suggesting quicker goal completion.
- Value variance was higher in the custom environments, indicating sharper value differences across states.

Conclusion

Dynamic programming algorithms like value iteration are highly sensitive to environment complexity. Deterministic maps with fewer pitfalls lead to faster convergence and more efficient policies. In contrast, the stochastic nature of `FrozenLake-v1` increases convergence time and reduces reward consistency.