Markov Decision Process and Dynamic Programming
Final Report

Navneet Kashyap

Results Summary

Here's a detailed overview of the results comparing Policy Iteration (PI) and Value Iteration (VI) across different FrozenLake environments:

Performance Comparison

Environment	Algorithm	Iterations	Time (s)	Avg. Reward	Avg. Length
CustomFrozenLake-v1	PI	3	0.0032	1.00	8.0
(5×5, no slip)	VI	–	0.0007	1.00	8.0
CustomFrozenLake-v1-slip	PI	3	0.0460	0.576	53.7
(5×5, slippery)	VI	–	0.0266	0.523	54.3
ExpandedFrozenLake-v1	PI	3	0.0061	1.00	6.0
(4×4 + key, no slip)	VI	–	0.0007	1.00	6.0
ExpandedFrozenLake-v1-slip	PI	4	0.0458	1.00	25.2
(4×4 + key, slippery)	VI	–	0.0140	0.998	25.6
FrozenLake-v1	PI	3	0.0257	0.713	45.5
(Default 4×4)	VI	–	0.0192	0.741	44.8

Table 1: Comparison of PI vs. VI across various FrozenLake setups.

Key Insights

1. **Perfect Performance in Deterministic Maps:**

   - Both PI and VI reached 100% success (reward = 1.00) in all non-slippery environments.
   - VI ran 3–10× faster than PI.

2. **Slippery Environments:**

   - The expanded "key" map stayed near perfect (99.8–100%).
   - The custom 5×5 slippery map averaged only 52–58% success.

- The original 4×4 slippery map saw 71–74% success.

3. **Path Efficiency:**

   - Deterministic maps all yielded optimal 6–8-step routes.
   - The expanded slippery map still only took 25 steps on average.
   - The original slippery map required 45 steps—longest of all.

4. **Algorithmic Trade-offs:**

   - VI consistently cuts computation time by 30–70%.
   - Both methods converge in just 3–4 iterations.
   - Overall performance (reward/length) is very similar.

5. **Impact of Map Design:**

   - Adding a key mechanic doesn't hurt performance under slippage.
   - Larger 5×5 grid showed more performance degradation when slippery.
   - Original FrozenLake proved most challenging despite smaller size (on the basis of average length)

Conclusion

The DP algorithms demonstrated awesome performance in deterministic environments, with VI being computationally faster. In stochastic environments, performance varied based on map design, with the expanded key-based environment showing very less drop in reward to slippage. The original FrozenLake environment required the most length on average and thus proved more challenging for the agent despite being smaller.