Solving the Frozen Lake Problem with Value Iteration

The Frozen Lake problem is a classic grid world problem where an agent must navigate from a starting point to a goal while avoiding holes (dangerous states) on a frozen lake. The goal is to find the optimal policy that maximizes the agent's chance of reaching the goal safely.

Value Iteration Algorithm

Value iteration is an iterative algorithm used to compute the optimal value function and the optimal policy for a given Markov decision process (MDP). Here's the value iteration algorithm for solving the Frozen Lake problem:

Algorithm 1 Value Iteration

Initialize $V(s) = 0$ for all states s
Initialize $\epsilon > 0$ as the convergence threshold
Initialize Δ as a large value
Initialize iteration counter $k = 0$
while $\Delta > \epsilon$ **do**
    $\Delta \leftarrow 0$
    for each state s **do**
        $v \leftarrow V(s)$
        $V(s) \leftarrow \max_a \left(R(s,a) + \gamma \sum_{s'} P(s'|s,a)V(s')\right)$
        $\Delta \leftarrow \max(\Delta, |v - V(s)|)$
    end for
    Increment iteration counter: $k \leftarrow k + 1$
    Output current value function V and iteration number k
end while

In this algorithm:

- $V(s)$ is the value function for state s, representing the expected cumulative reward from state s onwards.
- $R(s, a)$ is the immediate reward for taking action a in state s.
- γ is the discount factor, determining the importance of future rewards.
- $P(s'|s, a)$ is the transition probability to state s' from state s after taking action a.
- ϵ is the convergence threshold, determining when to stop iterating.
- Δ is the maximum change in the value function across all states in an iteration.
- k is the iteration counter, indicating the number of iterations performed.

The algorithm iteratively updates the value function until it converges to the optimal value function $V^*(s)$, which represents the maximum expected cumulative reward from each state under the optimal policy. Each iteration improves the estimate of the optimal value function, leading to a better approximation of the optimal policy.

Solution

The solution to the Frozen Lake problem using value iteration involves applying the value iteration algorithm to compute the optimal value function $V^*(s)$ and the optimal policy $\pi^*(s)$ for the given MDP. The optimal policy $\pi^*(s)$ can be derived from the optimal value function $V^*(s)$ by selecting the action a in each state s that maximizes the expression inside the max operator in the Bellman optimality equation.