Brief Comparison between Monte Carlo, Dynamic Programming and Q-Learning

Eng. Métwalli

11/20/2024

1 Introduction

Approaches:

- **Monte Carlo (MC):** Learns by averaging rewards after completing full tasks (episodes). Monte Carlo methods estimate the value of states based on observed returns from complete episodes. These methods do not require knowledge of the environment's rules but need tasks with a clear start and end.
- **Dynamic Programming (DP):** Solves problems step-by-step but requires knowing all the rules of the environment in advance. Dynamic Programming calculates the value of states by iteratively solving the Bellman equation, using full knowledge of the environment.
- **Q-Learning:** Learns without needing to know the environment by trial-and-error, updating knowledge after each action. Q-Learning is a model-free method where the agent learns by trial-and-error, updating Q-values after each action.

2 Comparison of Methods

Feature	**Monte Carlo**	**Dynamic Programming**	**Q-Learning**
Model Requirement	No (works without rules)	Yes (needs full rules)	No (works without rules)
Update Timing	After finishing a task	Step-by-step	After each action
Type of Learning	From full tasks	Uses known rules	Learns through trial-and-error
Bootstrapping	No	Yes	Yes
Policy	Explicit or implicit	Explicit	Implicit
Exploration	Requires exploration	No need (rules are known)	Explores automatically

Table 1: Comparison of RL methods