

Random Tabular Q-Planning

1. khái niệm

Random Tabular Q-Planning là một phương pháp lập kế hoạch trong học tăng cường, sử dụng mô hình để tạo ra các trải nghiệm giả lập nhằm cải thiện chính sách. Là một kỹ thuật lập kế hoạch trong đó agent **ngẫu nhiên chọn các cặp (trạng thái, hành động)** từ không gian đã biết, sau đó **cập nhật giá trị Q** bằng cách sử dụng mô hình mô phỏng của môi trường (model dynamics)

Phương pháp này kết hợp giữa:

- **Q-learning** trong dạng bảng (tabular Q-learning),
- Và **lập kế hoạch ngẫu nhiên** bằng cách sử dụng mô hình động học (model) mà agent đã học từ tương tác với môi trường.

$$Q(s, a) \leftarrow Q(s, a) + \alpha \left[r + \gamma \max_{a'} Q(s', a') - Q(s, a) \right]$$

Trong đó:

- s, a : Trạng thái và hành động được lấy mẫu ngẫu nhiên
- s' : Trạng thái tiếp theo dự đoán từ mô hình
- γ : Hệ số chiết khấu
- α : Learning rate
- r : Phần thưởng dự đoán từ mô hình