

Khái niệm về Model và Planning

1. Model là gì?

- Model lưu trữ kiến thức về **động lực học** (dynamics) của môi trường.
- Cho biết từ trạng thái và hành động hiện tại, ta có thể dự đoán được trạng thái kế tiếp và phần thưởng tương ứng.
- Nhờ đó, ta có thể **tưởng tượng kết quả hành động mà không cần thực sự thực hiện** trong môi trường thật.

2. Planning là gì?

- Planning là quá trình sử dụng model để **cải thiện chính sách** (policy).
- Cách làm là dùng trải nghiệm mô phỏng (simulated experience) từ model để cập nhật giá trị (value function).
- Nhờ mô phỏng này, ta tăng hiệu quả lấy mẫu (sample efficiency), tức là cần ít tương tác thật với môi trường mà vẫn học được chính sách tốt.

Hai loại model phổ biến

1. Sample model (Mô hình mẫu):

- Mô hình này sinh ra một kết quả thực tế dựa trên các xác suất ẩn của môi trường.
- Ví dụ: khi tung đồng xu, sample model sẽ sinh ra "ngửa" hoặc "sấp" ngẫu nhiên dựa trên xác suất 50% mỗi bên.
- Ưu điểm: tốn ít tính toán, chỉ tạo ra một kết quả cụ thể tại mỗi lần chạy mô phỏng.
- Ví dụ thực tế: Công ty CloudFlare dùng mẫu ngẫu nhiên từ hình ảnh bức tường đèn lava để mã hóa dữ liệu internet.

2. Distribution model (Mô hình phân phối):

- Mô hình này mô tả đầy đủ xác suất của mọi kết quả có thể xảy ra.

- Ví dụ: mô hình sẽ cho biết xác suất 50% "ngửa" và 50% "sấp" cho một đồng xu, và có thể mô tả xác suất của tất cả chuỗi tung 5 đồng xu (32 khả năng).
- Ưu điểm: có nhiều thông tin hơn, có thể sử dụng để sinh ra mẫu theo xác suất cụ thể.
- Nhược điểm: phức tạp hơn, đặc biệt khi không gian trạng thái hoặc hành động lớn.