

# The Dyna Algorithm

## 1. khái niệm:

**Dyna** là một thuật toán **Reinforcement Learning** kết hợp cả học từ trải nghiệm (**model-free**) và lập kế hoạch từ mô hình (**model-based**), được đề xuất bởi **Richard Sutton**.

**Dyna** là sự kết hợp giữa:

- **Trải nghiệm thật:** học trực tiếp từ môi trường (như Q-learning),
- **Mô hình môi trường:** học mô hình để mô phỏng trạng thái và phần thưởng,
- **Lập kế hoạch nội bộ (planning):** dùng mô hình để "tưởng tượng" các bước đi và cập nhật Q-values mà không cần tương tác thật.

## 2. Thuật toán:

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