

1. Generalization (Tổng quát hóa)

- Là khả năng suy luận từ một trạng thái sang các trạng thái khác tương tự.
- VD: Nếu ta biết cách lái 1 chiếc ô tô → ta không cần học lại từ đầu với ô tô khác hoặc lái trên đường khác.

Trong RL:

- Nếu ta cập nhật giá trị của một trạng thái sss, **các trạng thái tương tự s'** cũng nhận được ảnh hưởng.
- → Giúp **học nhanh hơn** (không cần thăm mọi trạng thái).

Ví dụ: Robot thu thập lon, nhiều vị trí khác nhau nhưng có **khoảng cách tương đương** tới lon → nên có giá trị gần nhau.

✓ 2. Discrimination (Phân biệt hóa)

- Là khả năng phân biệt giá trị giữa hai trạng thái có vẻ giống nhau nhưng thực ra rất khác nhau.
- VD:
 - Trạng thái A: lon cách 3m nhưng có bức tường.
 - Trạng thái B: lon cách 3m nhưng không có vật cản.
 - → Dù "sensor input" tương tự, robot cần phân biệt được hai trạng thái này vì giá trị hành động khác nhau.
 -

Monte Carlo gần giống supervised learning nhất

Vì target là return thực

TD dùng bootstrapping → target không cố định

Không giống SL nữa

RL cần online + khả năng xử lý dữ liệu phụ thuộc thời gian

Một số thuật toán SL không làm được điều này