MDP connu: solution offline

But

Calculer V^*, Q^*, π^*

Technique

Value / policy iteration

Évaluer une politique fixe π

Policy evaluation

MDP inconnu: Model-based

MDP inconnu: Mode-free

But

Technique

Calculer V^*, Q^*, π^*

VI/PI approx. MDP

Évaluer une politique fixe π

PE approx. MDP

But

Technique

Calculer V^*, Q^*, π^*

Q-learning

Évaluer une politique fixe π

Value learning