Reinforcement Learning

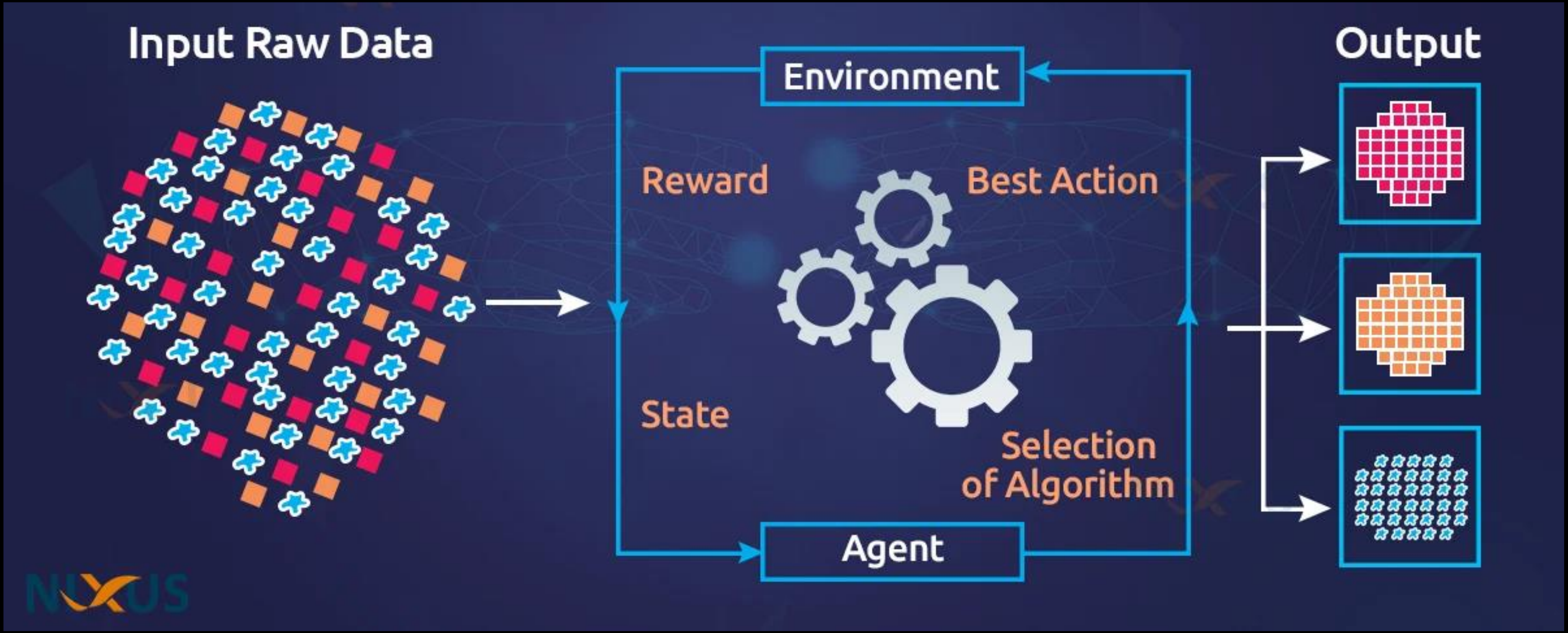
Asmaa Elbadrawy
PhD, Lecturer
IFT Program, ASU

Arizona State University

Supervised Learning simulates passive learning from a teacher.

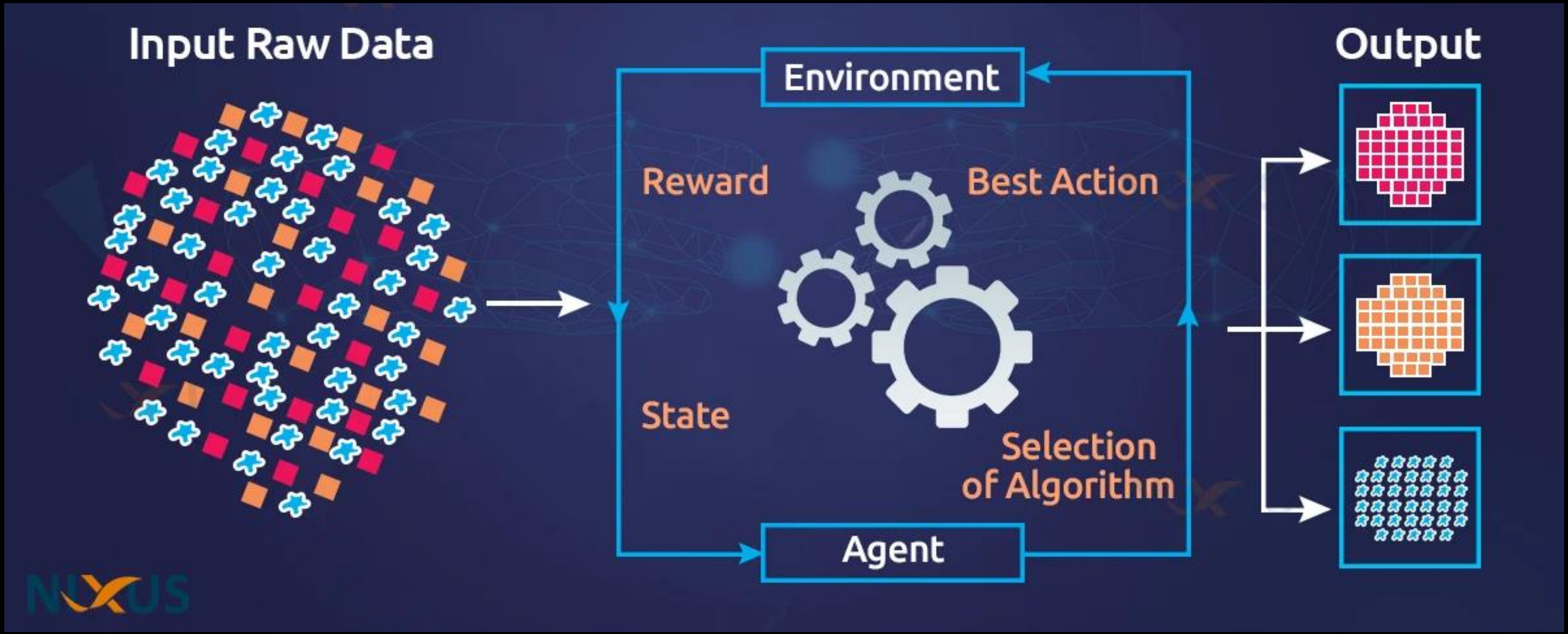
Reinforcement Learning (RL) simulates **active** learning from experience.

In RL, the agent interacts with the world and periodically receives reward or punishment.



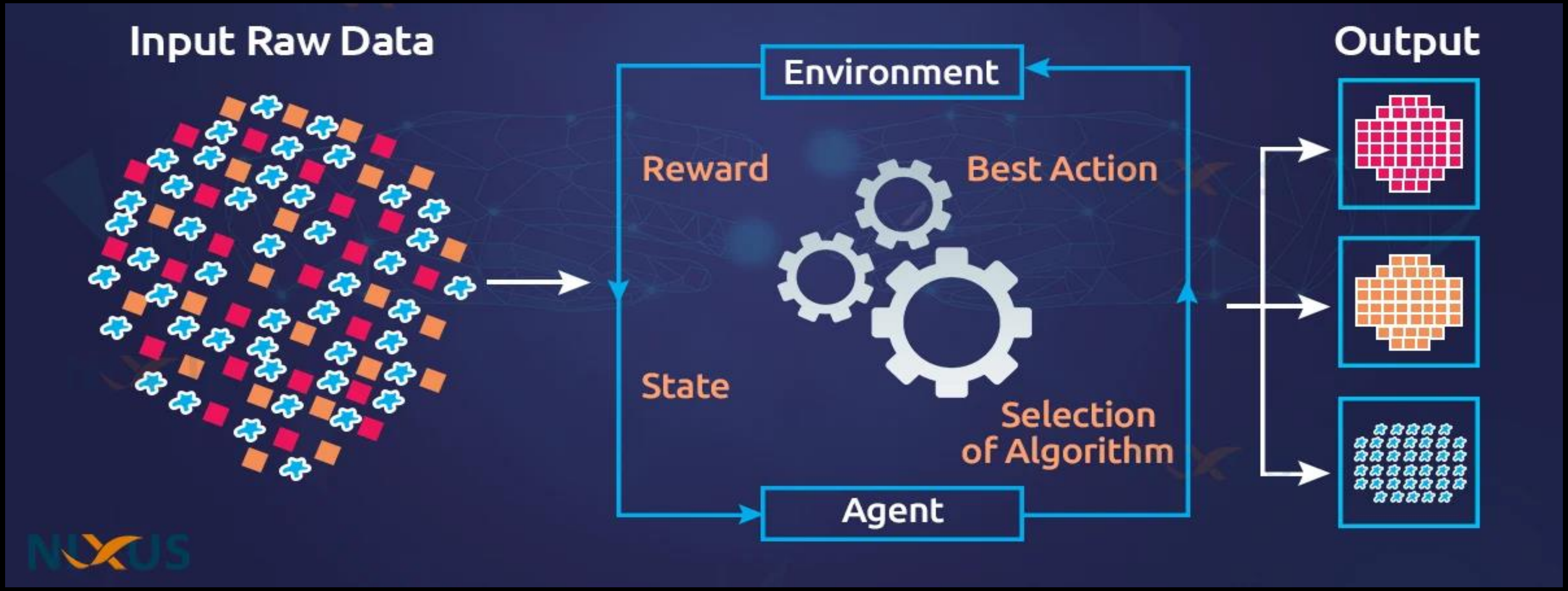
Source: https://nixustechnologies.com/reinforcement-learning-in-machine-learning/

Designer of an RL learning agent must provide a reward function.



Source: https://nixustechnologies.com/reinforcement-learning-in-machine-learning/

RL works very good in simulated environments (games) as there is abundance of opportunities to gain experience.



Source: https://nixustechnologies.com/reinforcement-learning-in-machine-learning/

Preliminaries

A Transition Model

It is a model that describes the result of taking a certain action in a certain state. Usually, the result is reaching another state.

$T(s, a, s')$ **denotes the transition when agent in state *s*, executes action *a*, and reaches state *s'*.**

A Reward Function

$R(s, a, s')$ denotes the reward the agent receives when it reaches state *s'* from state *s* by executing action *a*.

A Solution Policy

π denotes a solution policy that specifies, for each state s, what is the best action to take.

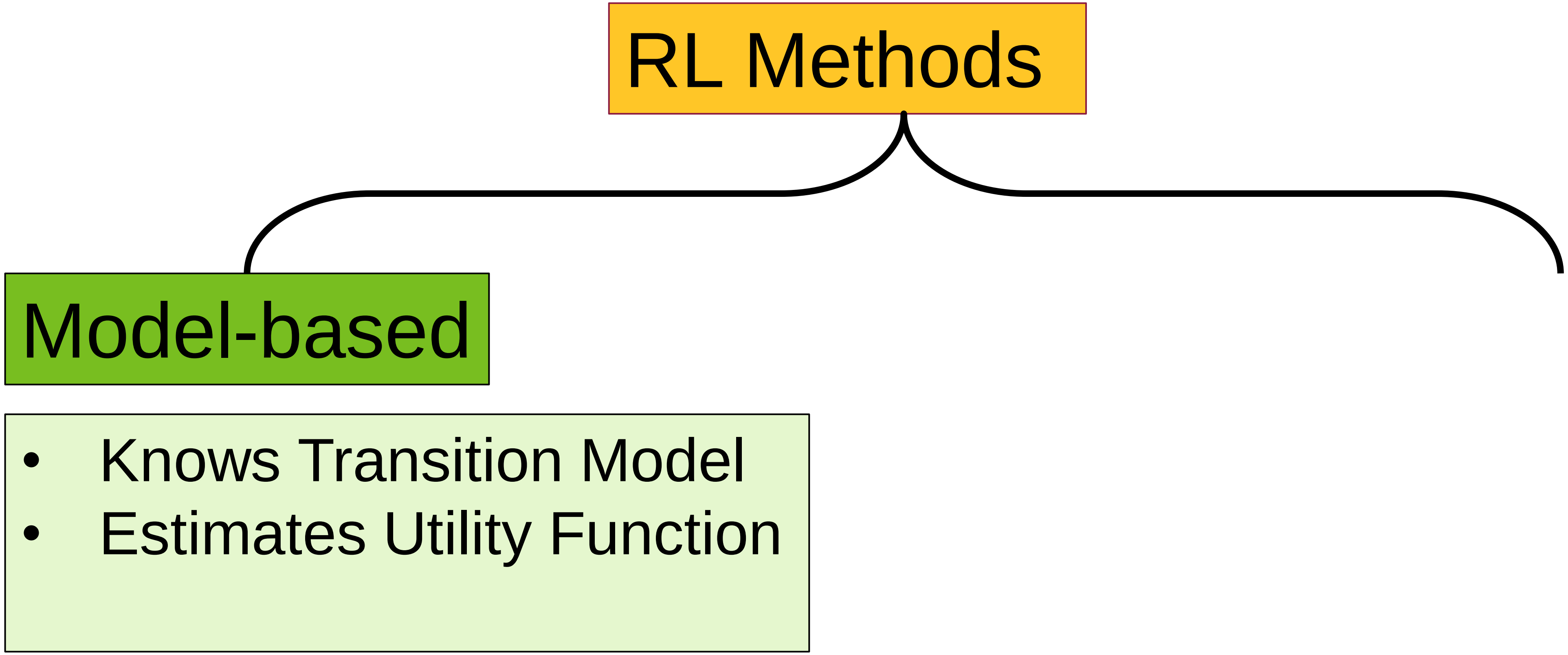
$\pi(s)$ denotes the action recommended by the policy π for state s.

A Utility (Value) Function

U denotes the performance of the agent measured as the sum of all rewards earned from current state onwards until reaching the goal state.

Reinforcement Learning Methods

RL Methods



RL Methods

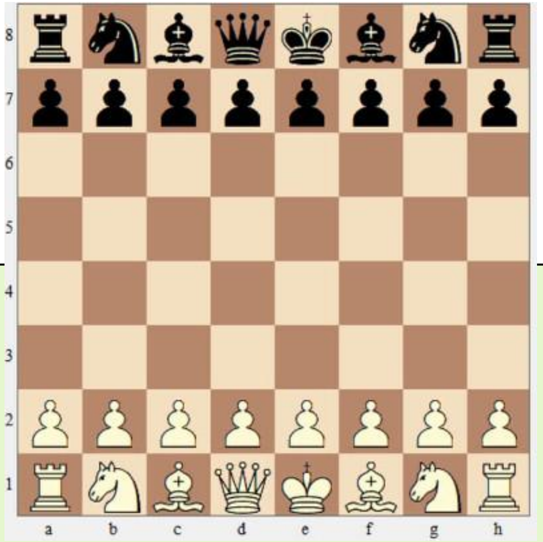
Model-based

- Knows Transition Model
- Estimates Utility Function

RL Methods

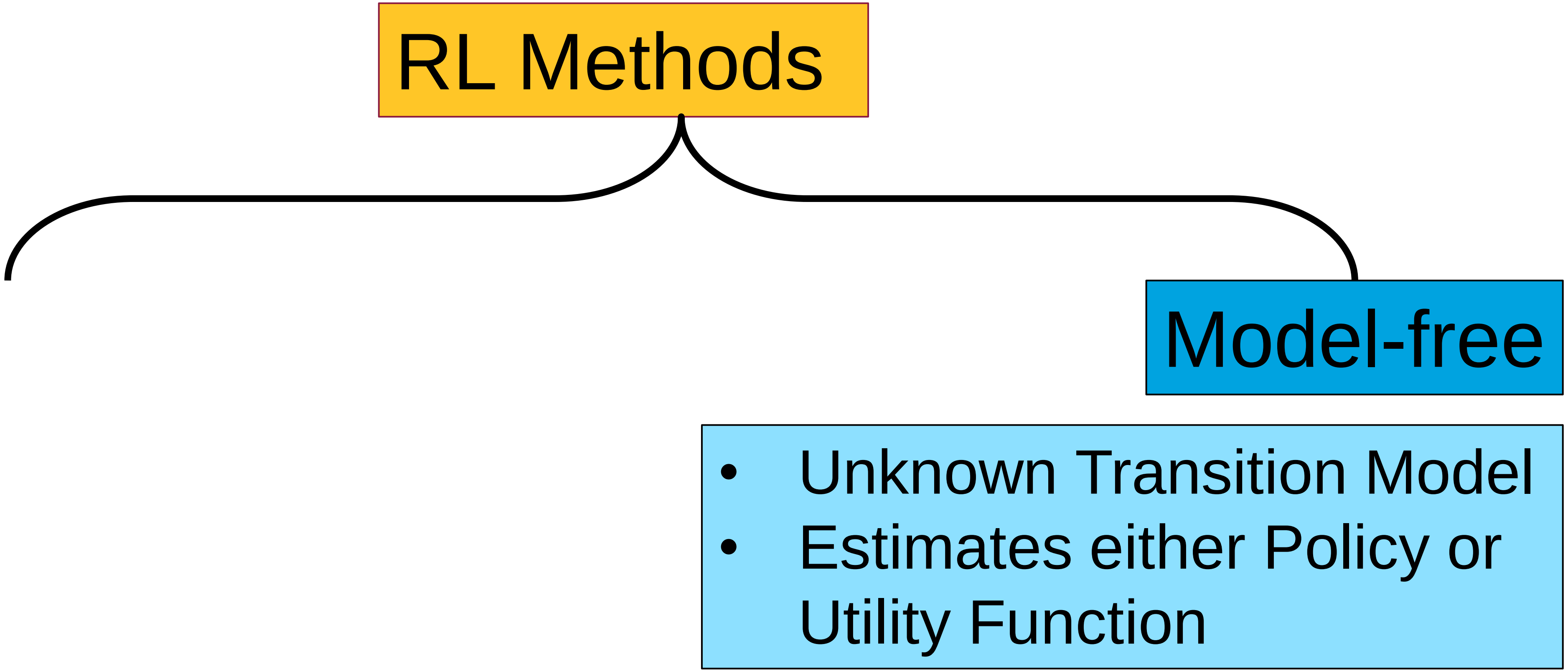
Model-based

- Knows Transition Model
- Estimates Utility Function



Chess-Playing Agent:

- Knows result of each action at each state (Transition Model)
- Estimate utility of each state (sum over future expected reward)



RL Methods

Model-free

- Unknown Transition Model
- Estimates either Policy or Utility Function

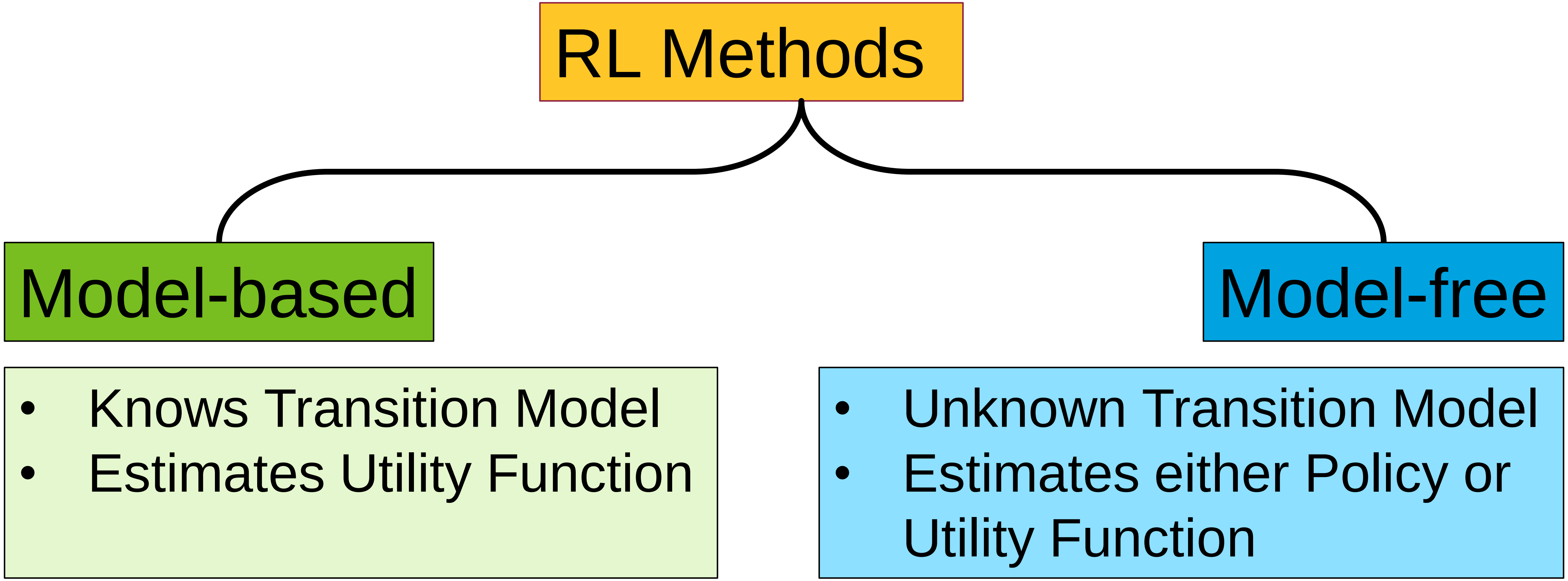
RL Methods

Rover Explorer Agent

- Being in a certain state, agent does not know what next state a certain action would take it to.
- Estimates either policy(best action to take for each state), or utility (future reward of being in a current state).

Model-free

- Unknown Transition Model
- Estimates either Policy or Utility Function



RL Methods

Model-based

- Knows Transition Model
- Estimates Utility Function

Model-free

- Unknown Transition Model
- Estimates either Policy or Utility Function

RL Applications

Gaming Applications:
Chess Playing Agent

Exploration Tasks: Robo-Vac exploring a room for the first time

Minimizing Energy Costs

Personalized Recommendations

RL Agents can make use of domain knowledge to speed up the learning process.

Exploration Tasks: Robo-Vac

Detect room boundaries first.

Minimizing Energy Costs

Season?

Location?

Personalized Recommendations

Exclude certain categories?