Challenges in RL
Environment: size, action space, and state space

THE UNIVERSITY OF
SOUTHERN MISSISSIPPI

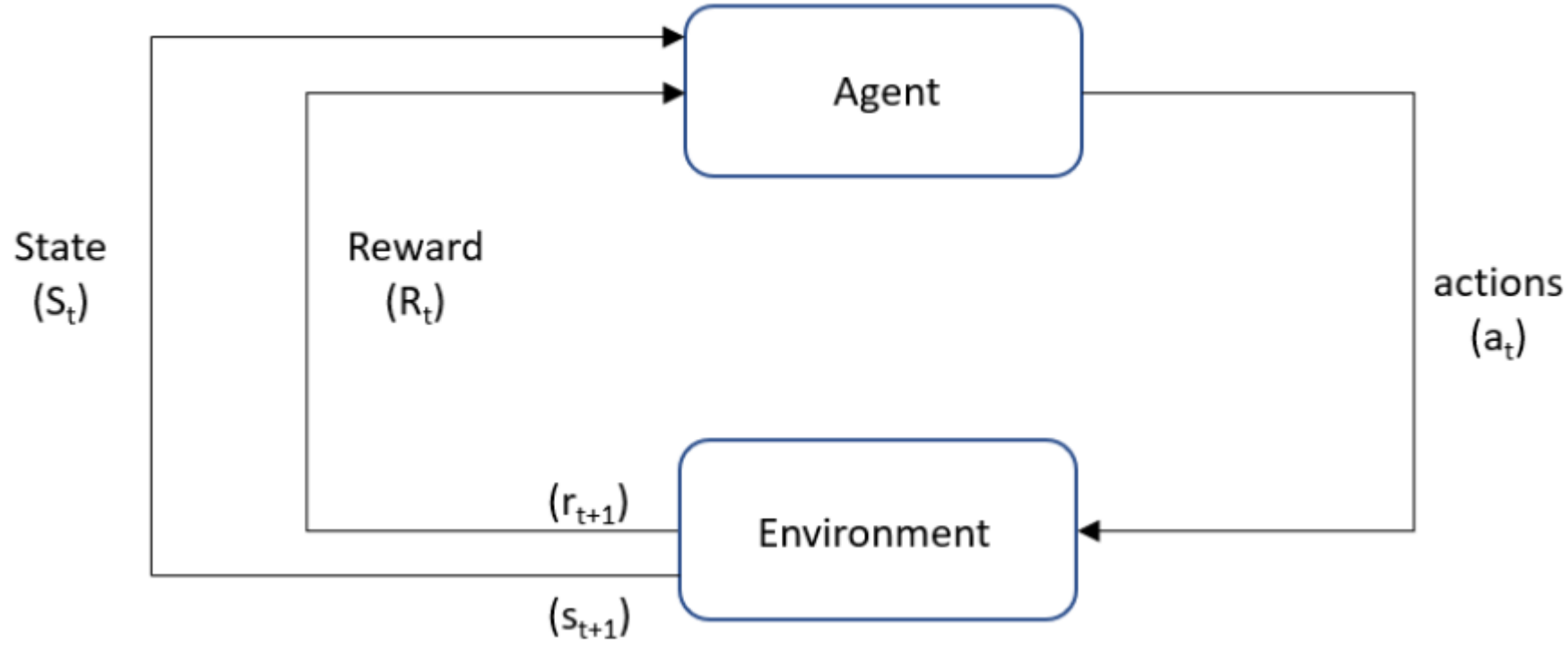
Challenges in RL

- In Q-learning, a Q-table is implemented:
  - From the robot in the grid example:
    - 3x3x4 = 36 total states -> The robot can be at any location in any orientation
    - 4 different states per 3x3 grid
      - Actions = Left, right, fwd. Location = (orientation) up/down/left/right
      - The state of the world is the robot's position + orientation

source	Target		
			
		(1,1, facing up)	
		(0,1,facing up)	

Challenges in RL

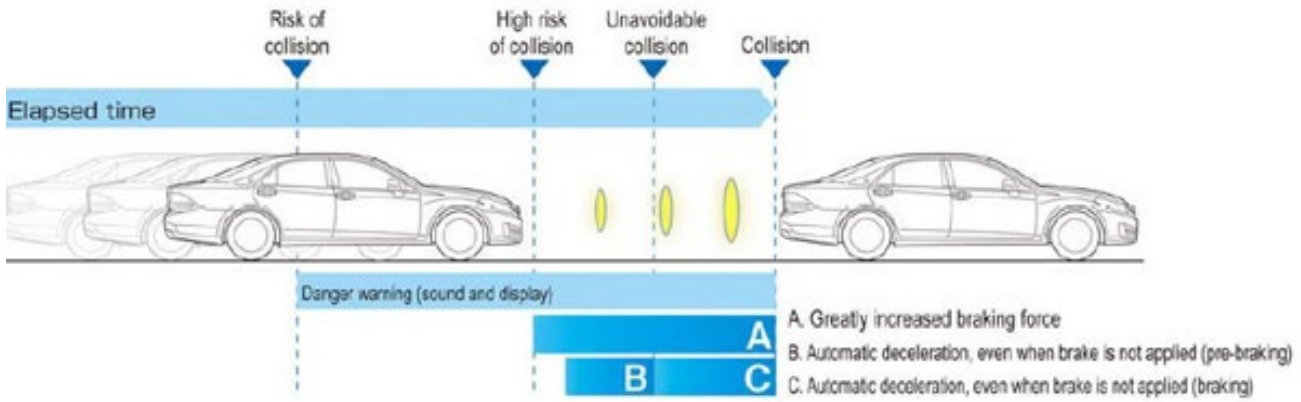
- Moving from (0,1) to (1,1) and obtaining a Q-value:
  - Select an action from the pool
  - Execute selected action
  - Approximate a new q-value: $Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha \left[R_{t+1} + \gamma \max_a Q\,(S_{t+1}, a)\right]$
  - Save Q-values in table:
    - Here, the complexity increases:
      - Q-Table needs to contain:
        - Combination of: actions, state, and what q-values were given for the combination of (action, state) pair
        - The Q-table is now an n-dimensional list.



- State: What it's observed
- State space: All possible states for a system
- Action space: set of actions. Has to be finite.

Challenges in RL

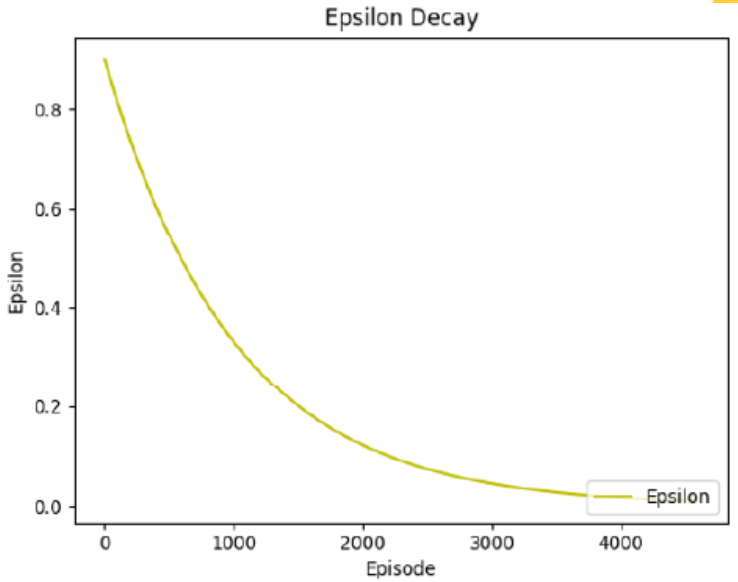
- To train a RL agent to create a model that will predict stopping parameters to prevent side collision:

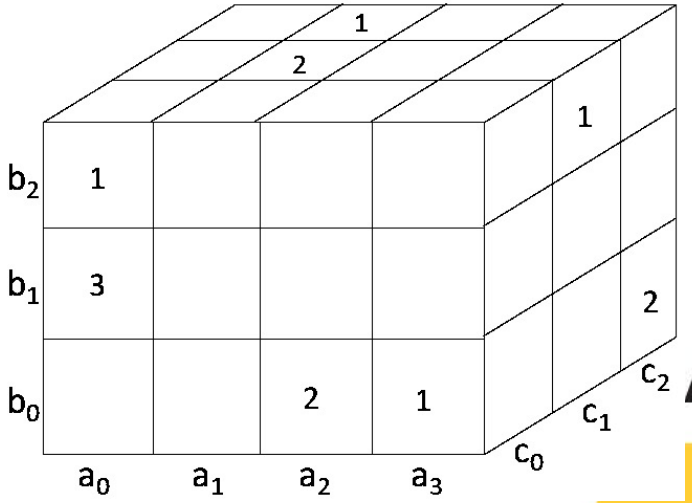


Some Actions	Other factors
Increase breaking force	Speed of the vehicle, Distance to next obstacle, Feedback from ABS system(corrections on-the-fly if needed) Increase sampling speed, etc.
Automatic deceleration (no brake, engine deceleration)	
Automatic deceleration(brakes, engine deceleration)	

Challenges in RL

- Training the agent:
  - Combination of (action, state):
    - Increases as the agent explores
  - Growth of the Q-table becomes “stable” after training
    - Related to the explore-exploit ratio
    - However, at this point, the Q-table will be extremely big
  - Training is complete: now save and deploy the model:
    - What are some problems?





Challenges in RL

- If using Q-learning:
  - For the model to make an effective response to avoid a collision:
    - Needs to look for the proper Q-value:
      - Needs to match sampled scenario with Q-value
  - The model saves parameters according to what it was configured to save:
    - Defined limits

Challenges in RL

- Some problems:
  - Action space and state space are large
  - The environment contains “other factors”
  - Regardless of the approach (SARSA, Q-Learning):
    - Problems with generalization
    - Generalization:
      - the ability of the same developed model to predict given multiple environments.
      - the ability of the same developed model to identify irregularities and make corrections “on the fly”
- Can excel in:
  - Industry Automation, why?
  - Trading and Finance, why?

Challenges in RL

- To remedy these challenges:
  - Deep Learning (DL)
    - Increases the processing scope
      - Instead of prediction, an approximation is calculated
    - Can handle the larger environment
      - “other factors”
  - Deep Reinforcement Learning (DRL):
    - More powerful than a DL approach by itself
    - Implements RL Q-learning to increase the processing scope

THE UNIVERSITY OF
SOUTHERN MISSISSIPPI®