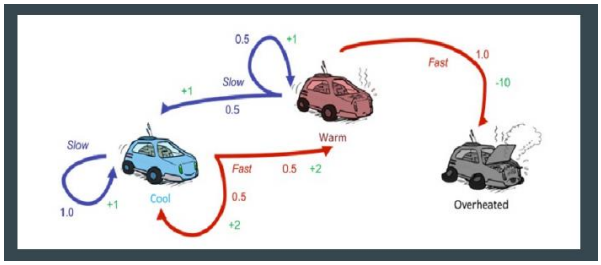
Lecture 2 Assignment



Probability and Reward table:

S	a	S'	P(S'|a,S')	R(S,a,S')
Cool	Fast	Warm	0,5	2
Cool	Fast	Cool	0,5	2
Cool	Slow	Cool	1	1
Warm	Slow	Warm	0,5	1
Warm	Slow	Cool	0,5	1
Warm	Fast	Overheated	1	-10

Optimal Value Table

	Cool	Warm	Overheated
V0	0,0	0,0	0,0
V1	2,0	1,0	-10,0
V2	3,5	2,5	-10,0
V3	5,0	4,0	-10,0

In the first iteration, all values are 0. For overheated state, all values will remain -10 starting iteration 1, it will not change because it can change from state.

For the second Iteration:

Cool State:

$$Q_1(C,F) = T(C,F,C)[R(C,F,C) + \gamma V_0(C)] + T(C,F,W)[R(C,F,W) + \gamma V_0(W)]$$
$$Q_1 = 0.5[2 + 1 * 0] + 0.5[2 + 1 * 0] = 2$$

$$Q_2(C,S) = T(C,S,C)[R(C,S,C) + \gamma V_0(C)]$$
$$Q_2 = 1[1 + 1 * 0] = 1$$

$$V_1(C) = \max\big(Q_1(C,F), Q_2(C,S)\big) = 2$$

Warm State:

$$Q_1(W,F) = T(W,F,O)[R(E,F,O) + \gamma V_0(O)]$$
$$Q_1 = 1[-10 + 1 * 0] = -10$$

$$Q_2(W,S) = T(W,S,C)[R(W,S,C) + \gamma V_0(C)] + T(W,S,W)[R(W,S,W) + \gamma V_0(W)]$$
$$Q_2 = 0.5[1 + 1 * 0] + 0.5[1 + 1 * 0] = 1$$

$$V_1(W) = \max\big(Q_1(W,F), Q_2(W,S)\big) = 1$$

Lecture 2 Assignment

For the second iteration:

Cool State:

$$Q_1(C,F) = T(C,F,C)[R(C,F,C) + \gamma V_1(C)] + T(C,F,W)[R(C,F,W) + \gamma V_1(W)]$$
$$Q_1 = 0.5[2 + 1 * 2] + 0.5[2 + 1 * 1] = 3.5$$

$$Q_2(C,S) = T(C,S,C)[R(C,S,C) + \gamma V_1(C)]$$
$$Q_2 = 1[1 + 1 * 2] = 3$$

$$V_2(C) = \max\big(Q_1(C,F), Q_2(C,S)\big) = 3.5$$

Warm State:

$$Q_1(W,F) = T(W,F,O)[R(E,F,O) + \gamma V_1(O)]$$
$$Q_1 = 1[-10 + 1 * -10] = -20$$

$$Q_2(W,S) = T(W,S,C)[R(W,S,C) + \gamma V_1(C)] + T(W,S,W)[R(W,S,W) + \gamma V_1(W)]$$
$$Q_2 = 0.5[1 + 1 * 2] + 0.5[1 + 1 * 1] = 2.5$$

$$V_2(W) = \max\big(Q_1(W,F), Q_2(W,S)\big) = 2.5$$

For the third iteration:

Cool State:

$$Q_1(C,F) = T(C,F,C)[R(C,F,C) + \gamma V_2(C)] + T(C,F,W)[R(C,F,W) + \gamma V_2(W)]$$
$$Q_1 = 0.5[2 + 1 * 3.5] + 0.5[2 + 1 * 2.5] = 5$$

$$Q_2(C,S) = T(C,S,C)[R(C,S,C) + \gamma V_1(C)]$$
$$Q_2 = 1[1 + 1 * 3.5] = 4.5$$

$$V_3(C) = \max\big(Q_1(C,F), Q_2(C,S)\big) = 5$$

Warm State:

$$Q_1(W,F) = T(W,F,O)[R(E,F,O) + \gamma V_2(O)]$$
$$Q_1 = 1[-10 + 1 * -10] = -20$$

$$Q_2(W,S) = T(W,S,C)[R(W,S,C) + \gamma V_1(C)] + T(W,S,W)[R(W,S,W) + \gamma V_1(W)]$$
$$Q_2 = 0.5[1 + 1 * 3.5] + 0.5[1 + 1 * 2.5] = 4$$

$$V_3(W) = \max\big(Q_1(W,F), Q_2(W,S)\big) = 4$$