Example: Value Iteration
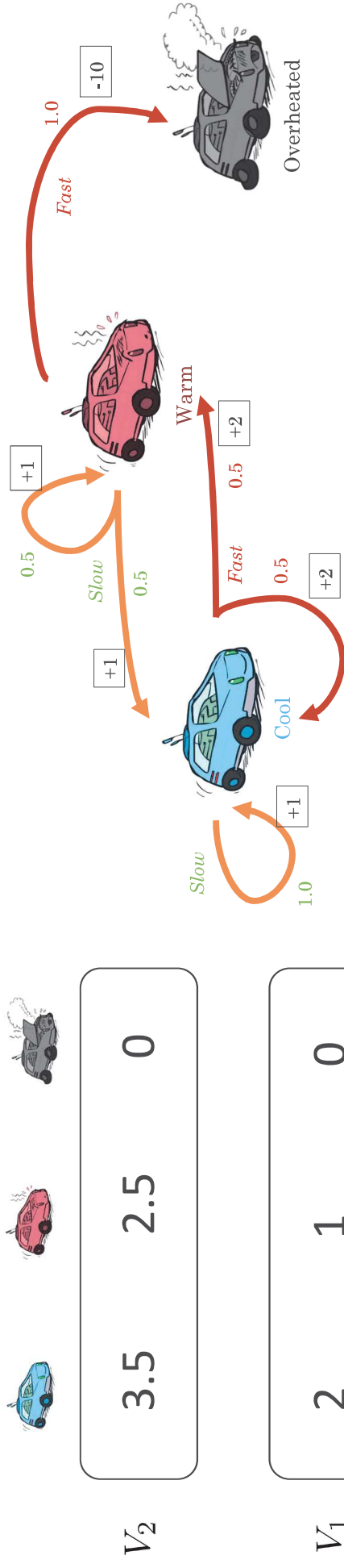

Assume no discount!

$$V_{k+1}(s) \leftarrow \max_a \sum_{s'} T(s, a, s') \left[R(s, a, s') + \gamma V_k(s') \right]$$

	🚗	🚗	🚗
V_2	3.5	2.5	0
V_1	2	1	0
V_0	0	0	0

Problems with Value Iteration

- Value iteration repeats the Bellman updates:

$$V_{k+1}(s) \leftarrow \max_a \sum_{s'} T(s,a,s') \left[R(s,a,s') + \gamma V_k(s') \right]$$

s

a

s, a

s,a,s'

s'

- Problem 1: It's slow – O(S²A) per iteration

- Problem 2: The "max" at each state rarely changes

- Problem 3: The policy often converges long before the values