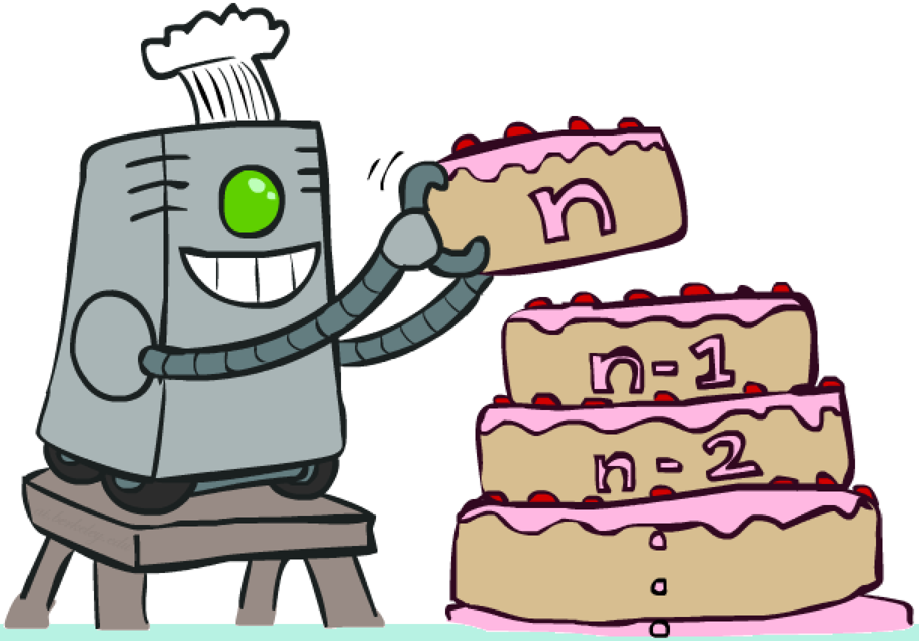
Advanced Topics in AI

Value Iteration



Instructor: Prof. Dr. techn. Wolfgang Nejdl

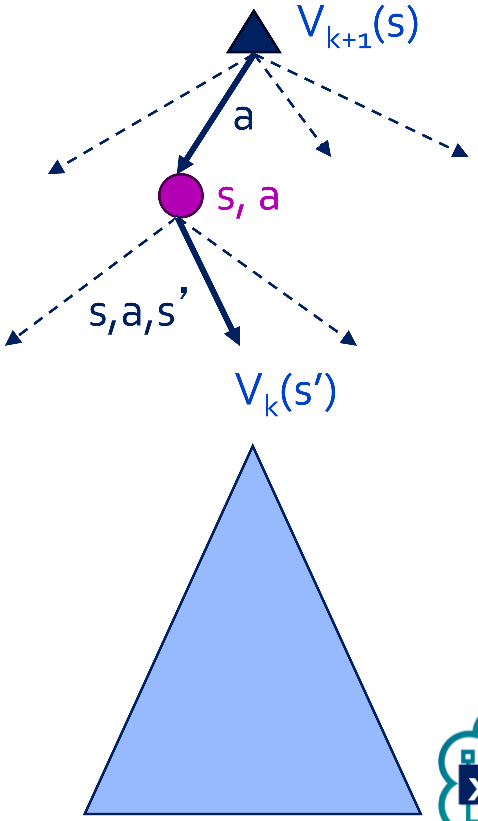
Leibniz University Hannover

Co-financed by the Connecting Europe Facility of the European Union

Value Iteration

- Start with $V_0(s) = 0$: no time steps left means an expected reward sum of zero
- Given vector of $V_k(s)$ values, do one ply of expectimax from each state:

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

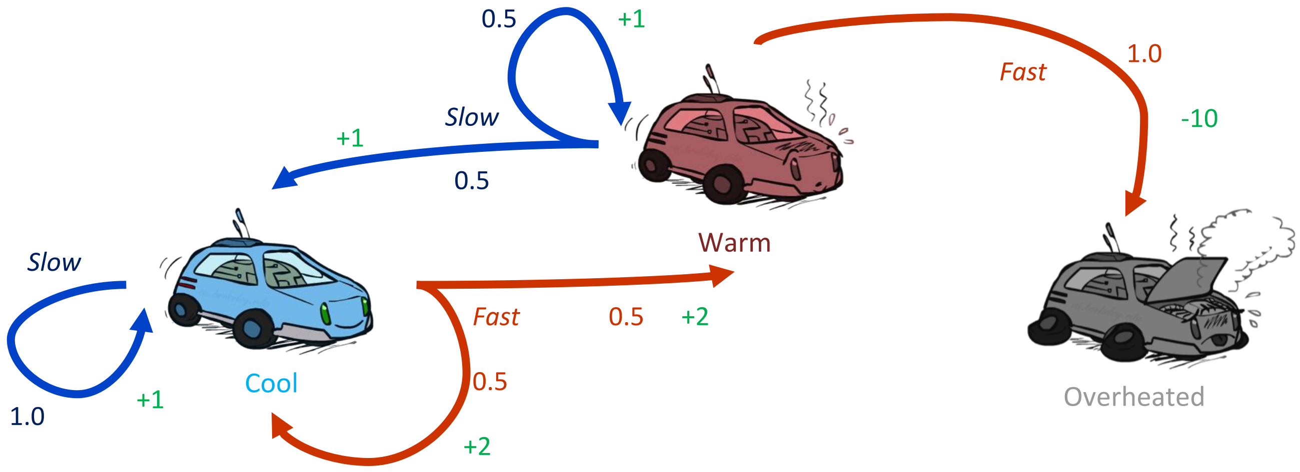


- Repeat until convergence, which yields V^*
- Complexity of each iteration: $O(S^2A)$
- Theorem: will converge to unique optimal values
  - Basic idea: approximations get refined towards optimal values
  - Policy may converge long before values do

Co-financed by the Connecting Europe Facility of the European Union

Example: Value Iteration

V_2			
V_1	S: 1 F: 0.5*2+0.5*2=2		
V_0	0	0	0



Assume no discount!

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

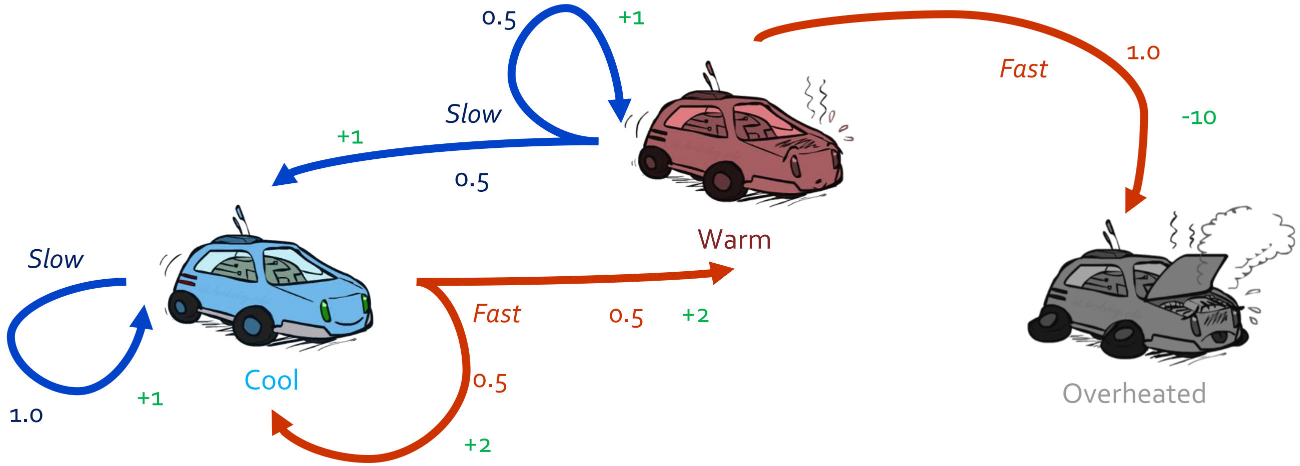
Co-financed by the Connecting Europe Facility of the European Union

Example: Value Iteration

V_2

V_1 2 S: 0.5*1+0.5*1=1 F: -10

V_0 0 0 0



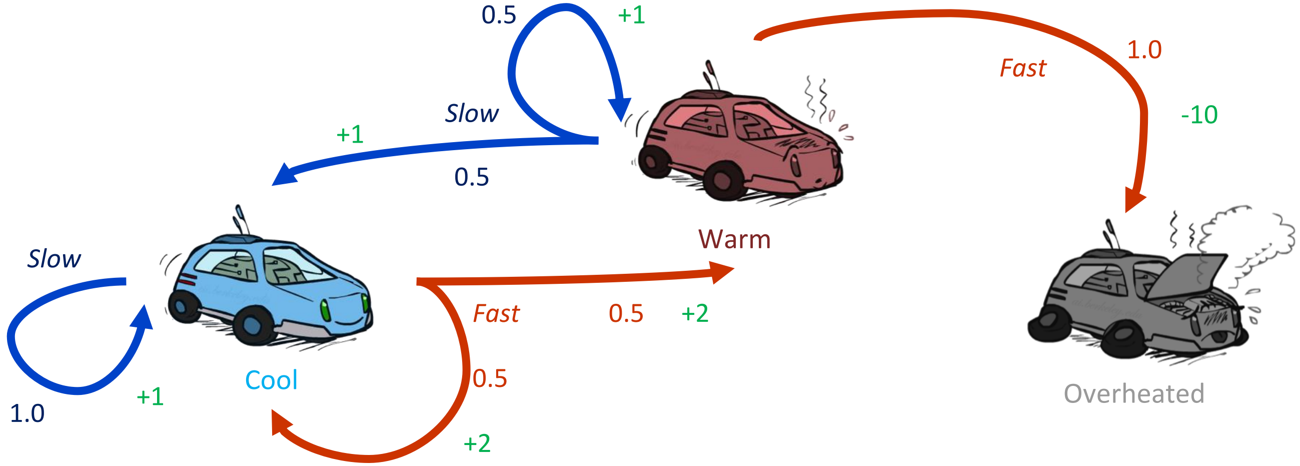
Assume no discount!

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

Co-financed by the Connecting Europe Facility of the European Union

Example: Value Iteration

V_2			
V_1	2	1	0
V_0	0	0	0



Assume no discount!

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

Co-financed by the Connecting Europe Facility of the European Union

Example: Value Iteration

V_2

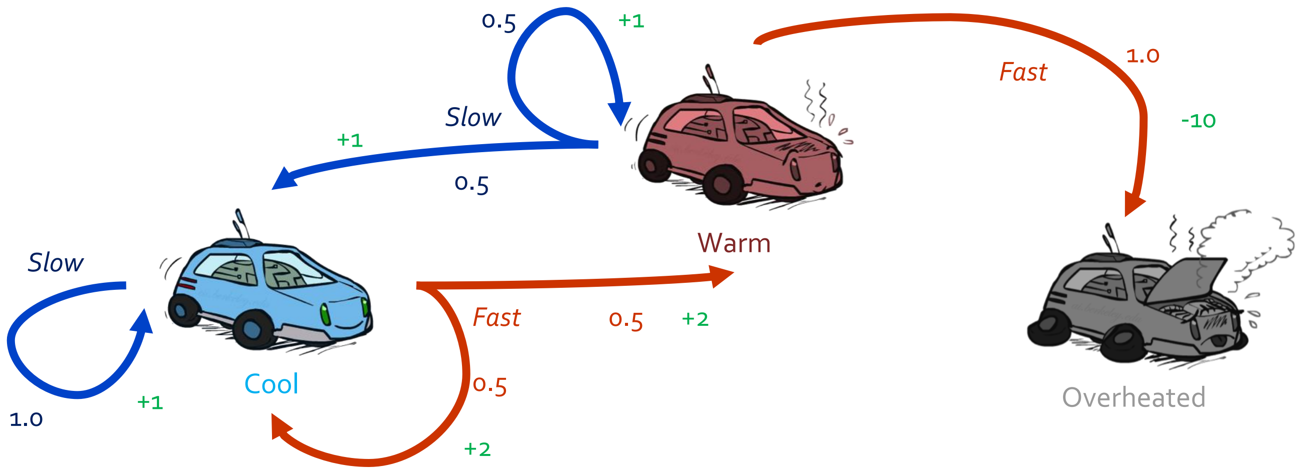
S: 1+2=3
F: 0.5*(2+2)+0.5*(2+1)=3.5

V_1

2	1	0

V_0

0	0	0



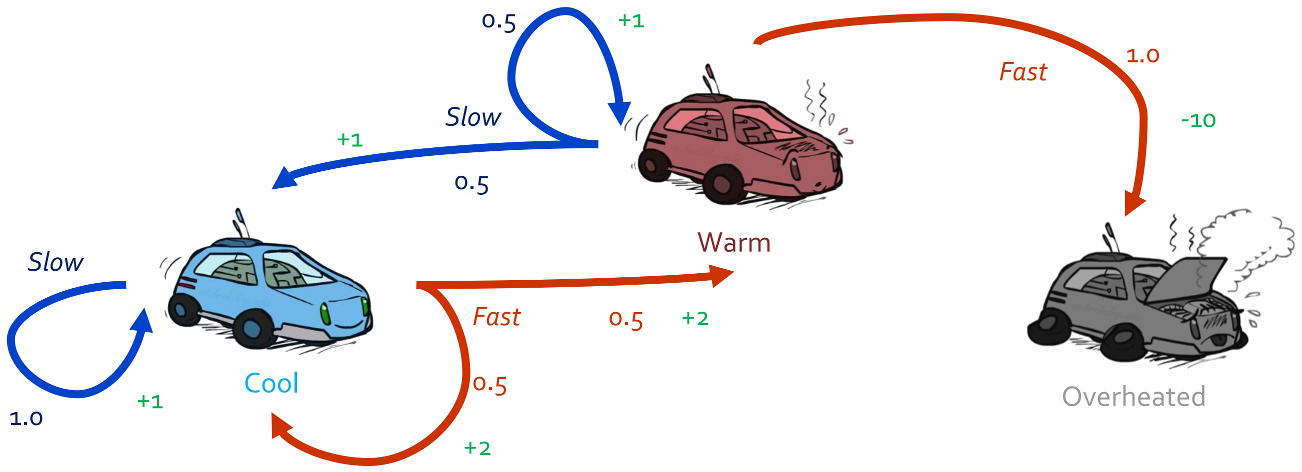
Assume no discount!

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

Co-financed by the Connecting Europe Facility of the European Union

Example: Value Iteration

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



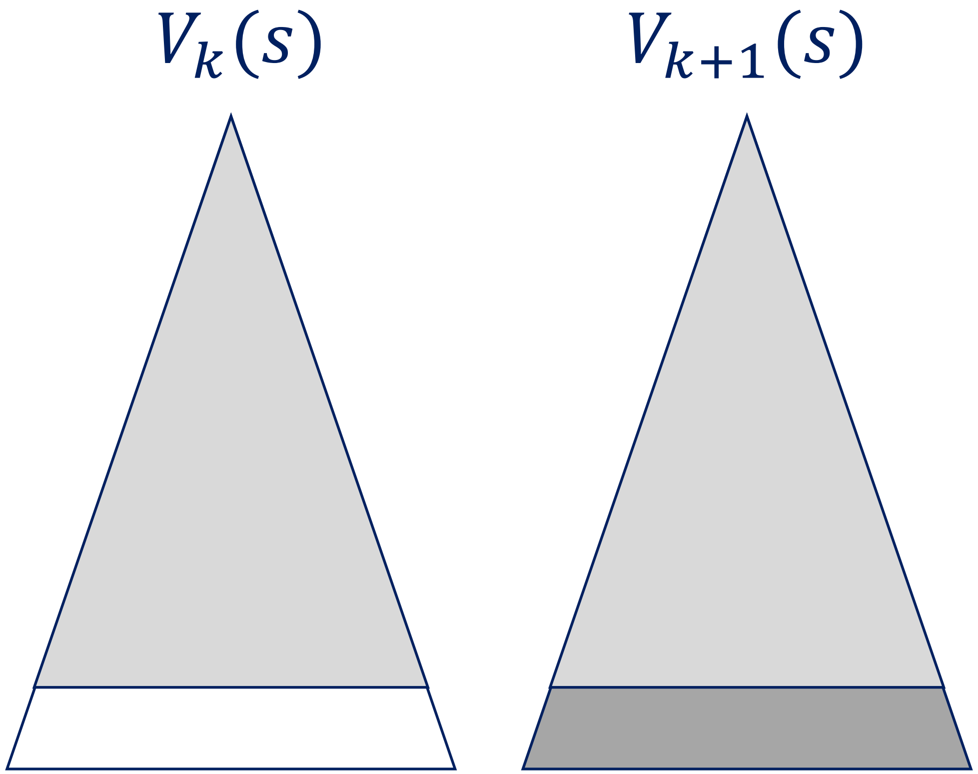
Assume no discount!

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

Co-financed by the Connecting Europe Facility of the European Union

Convergence

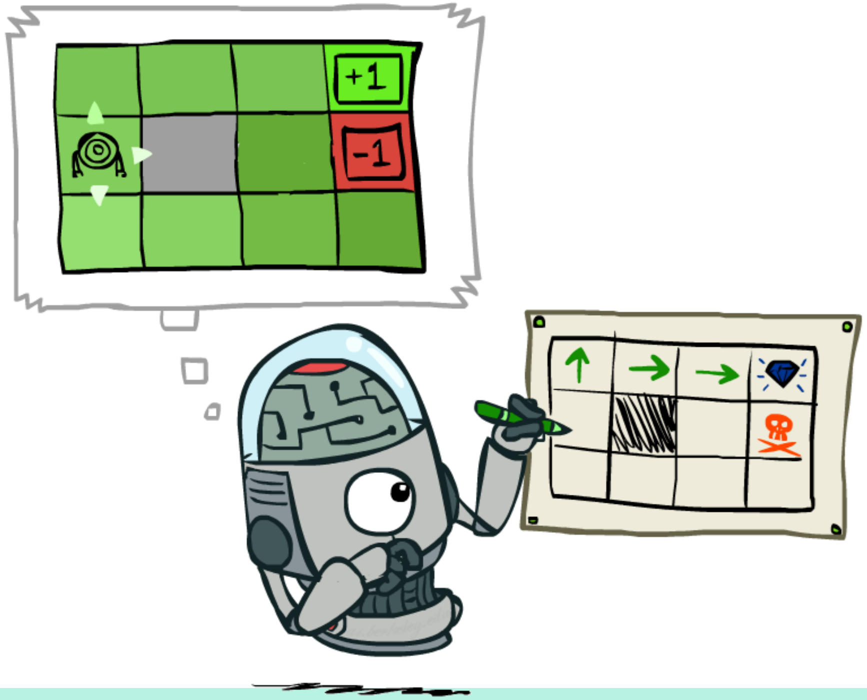
- How do we know the V_k vectors are going to converge? (assuming $0 < \gamma < 1$)
- Proof Sketch:
  - For any state V_k and V_{k+1} can be viewed as depth $k+1$ expectimax results in nearly identical search trees
  - The difference is that on the bottom layer, V_{k+1} has actual rewards while V_k has zeros
  - That last layer is at best all R_{MAX}
  - It is at worst R_{MIN}
  - But everything is discounted by γ^k that far out
  - So V_k and V_{k+1} are at most $\gamma^k \max |R|$ different
  - So as k increases, the values converge



Co-financed by the Connecting Europe Facility of the European Union

Advanced Topics in AI

Next: Policy-based methods



Instructor: Prof. Dr. techn. Wolfgang Nejdl

Leibniz University Hannover

Co-financed by the Connecting Europe Facility of the European Union