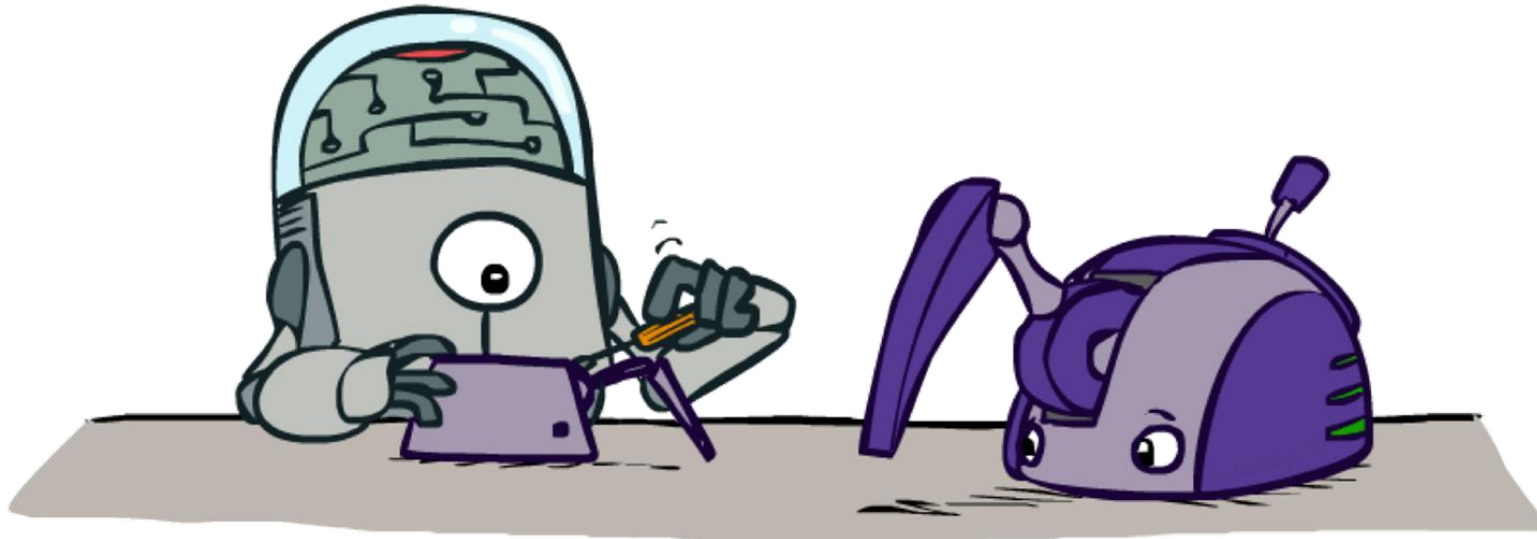


# Advanced Topics in AI

## Model-Based Reinforcement Learning



Instructor: Prof. Dr. techn. Wolfgang Nejdl

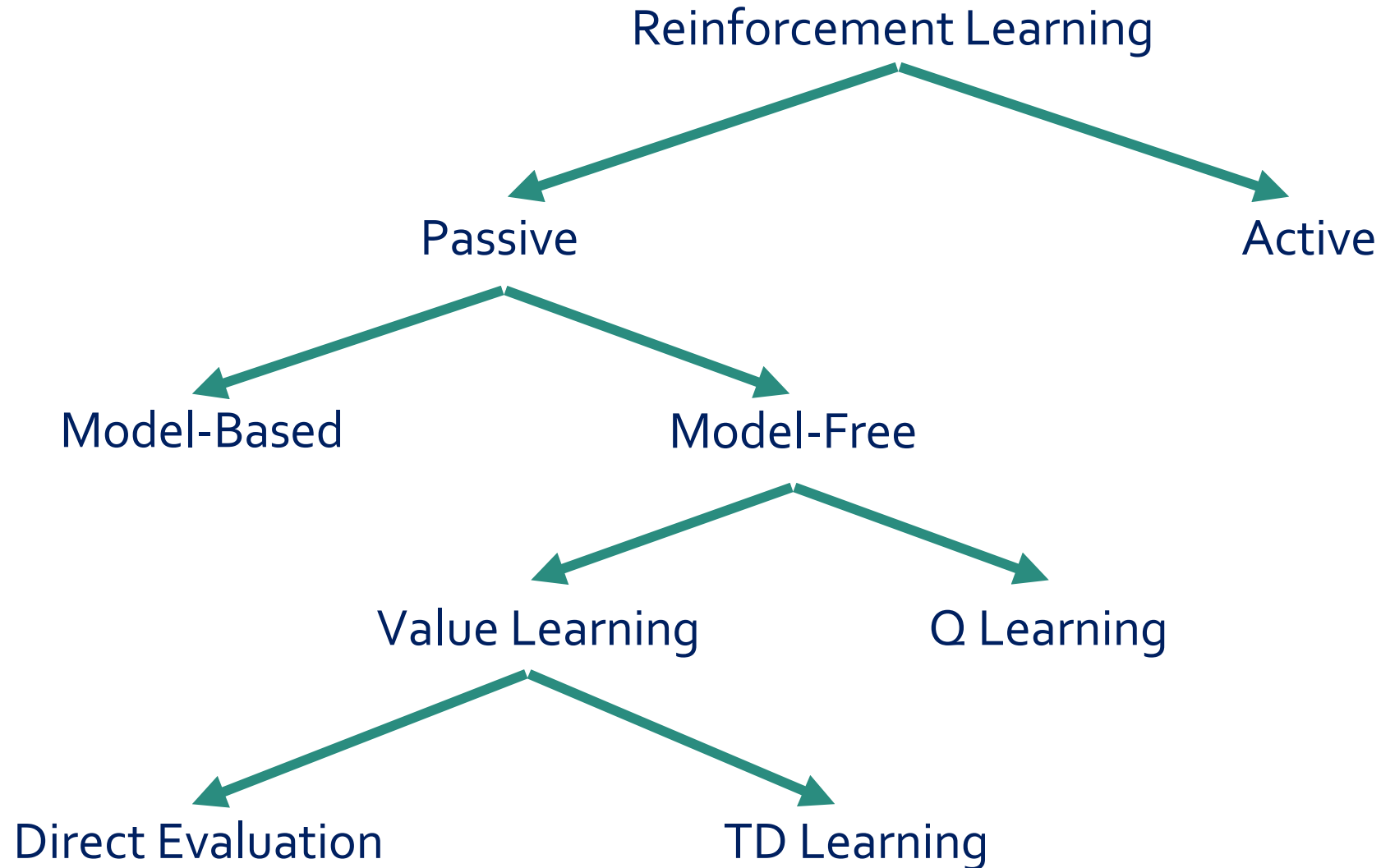
Leibniz University Hannover

[These slides were created by Dan Klein and Pieter Abbeel for CS188 Intro to AI at UC Berkeley. All materials are available at <http://ai.berkeley.edu>.]



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

# Reinforcement Learning Taxonomy



# Model-Based Reinforcement Learning

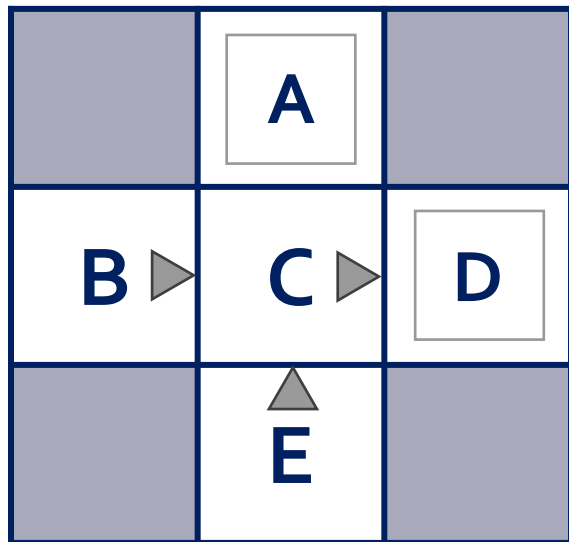
- Model-Based Idea:
  - Learn an approximate model based on experiences
  - Solve for values as if the learned model were correct
- Step 1: Learn empirical MDP model
  - Count outcomes  $s'$  for each  $s, a$
  - Normalize to give an estimate of  $\hat{T}(s, a, s')$
  - Discover each  $\hat{R}(s, a, s')$  when we experience  $(s, a, s')$
- Step 2: Solve the learned MDP
  - For example, use value iteration, as before

(and repeat as needed)



# Example: Model-Based RL

Input Policy  $\pi$



Assume:  $\gamma = 1$

Observed Episodes (Training)

Episode 1

B, east, C, -1  
C, east, D, -1  
D, exit, x, +10

Episode 2

B, east, C, -1  
C, east, D, -1  
D, exit, x, +10

Episode 3

E, north, C, -1  
C, east, D, -1  
D, exit, x, +10

Episode 4

E, north, C, -1  
C, east, A, -1  
A, exit, x, -10

Learned Model

$\hat{T}(s, a, s')$

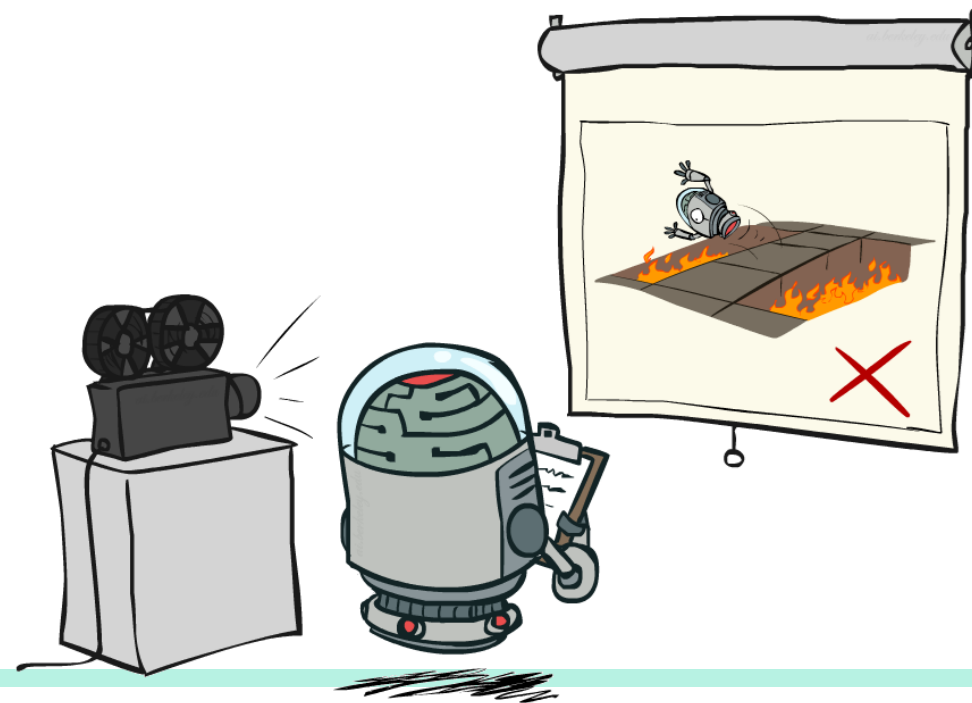
T(B, east, C) = 1.00  
T(C, east, D) = 0.75  
T(C, east, A) = 0.25  
...

$\hat{R}(s, a, s')$

R(B, east, C) = -1  
R(C, east, D) = -1  
R(D, exit, x) = +10  
...

# Advanced Topics in AI

## Next: Direct Evaluation



Instructor: Prof. Dr. techn. Wolfgang Nejdl

Leibniz University Hannover



[These slides were created by Dan Klein and Pieter Abbeel for CS188 Intro to AI at UC Berkeley. All materials are available at <http://ai.berkeley.edu>.]



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