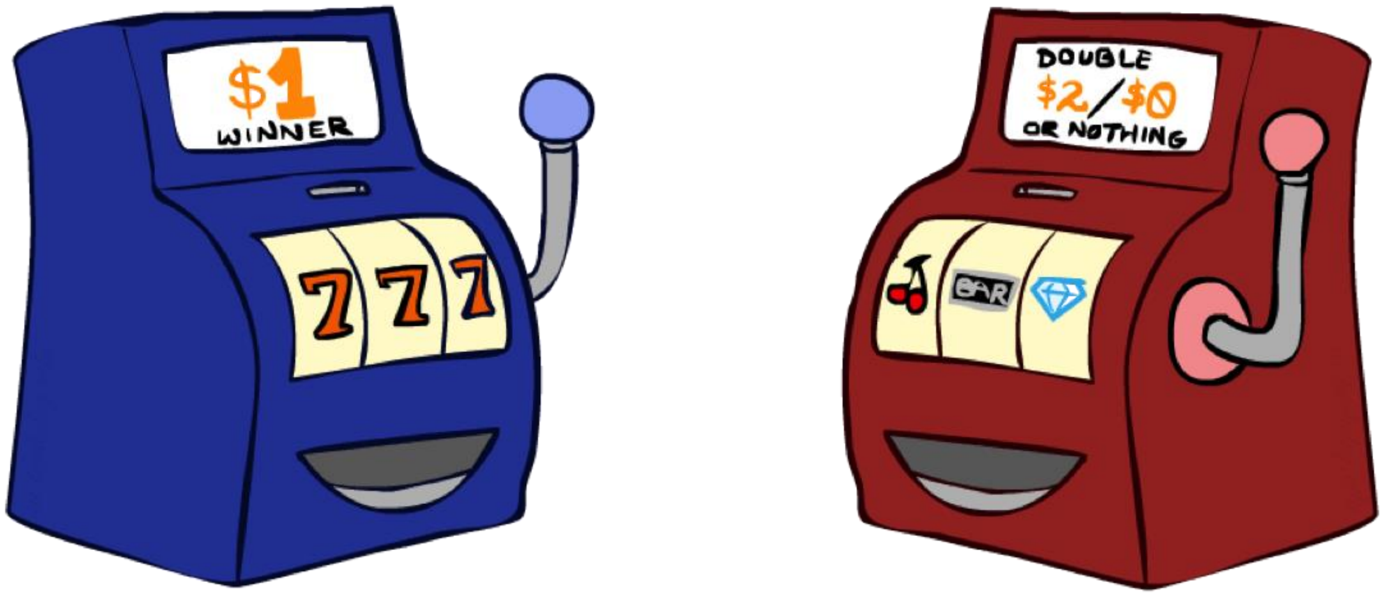
Advanced Topics in AI

Summary and Outlook



Instructor: Prof. Dr. techn. Wolfgang Nejdl

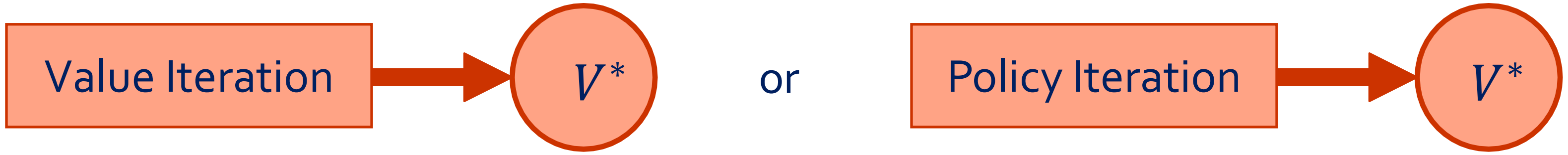
Leibniz University Hannover

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

[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.]

Summary: MDP Algorithms

- So you want to….
  - Compute optimal values: use value iteration or policy iteration



  - Compute values for a particular policy: use policy evaluation

  - Turn your values into a policy: use policy extraction (one-step lookahead)

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

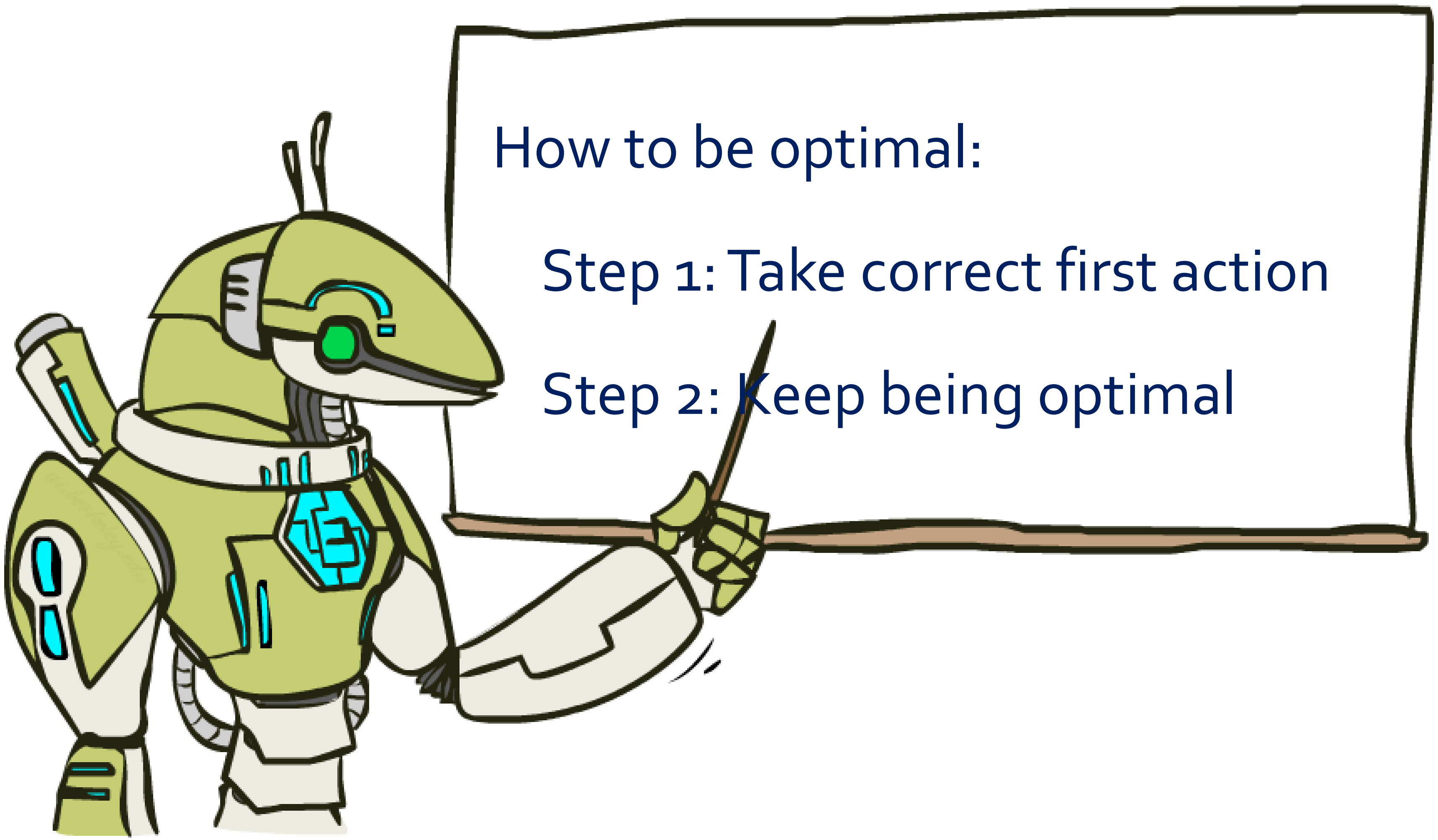
Summary: MDP Algorithms

- So you want to….
  - Compute optimal values: use value iteration or policy iteration
  - Compute values for a particular policy: use policy evaluation
  - Turn your values into a policy: use policy extraction (one-step lookahead)

- These all look the same!
  - They basically are – they are all variations of Bellman updates
  - They all use one-step lookahead expectimax fragments
  - They differ only in whether we plug in a fixed policy or max over actions

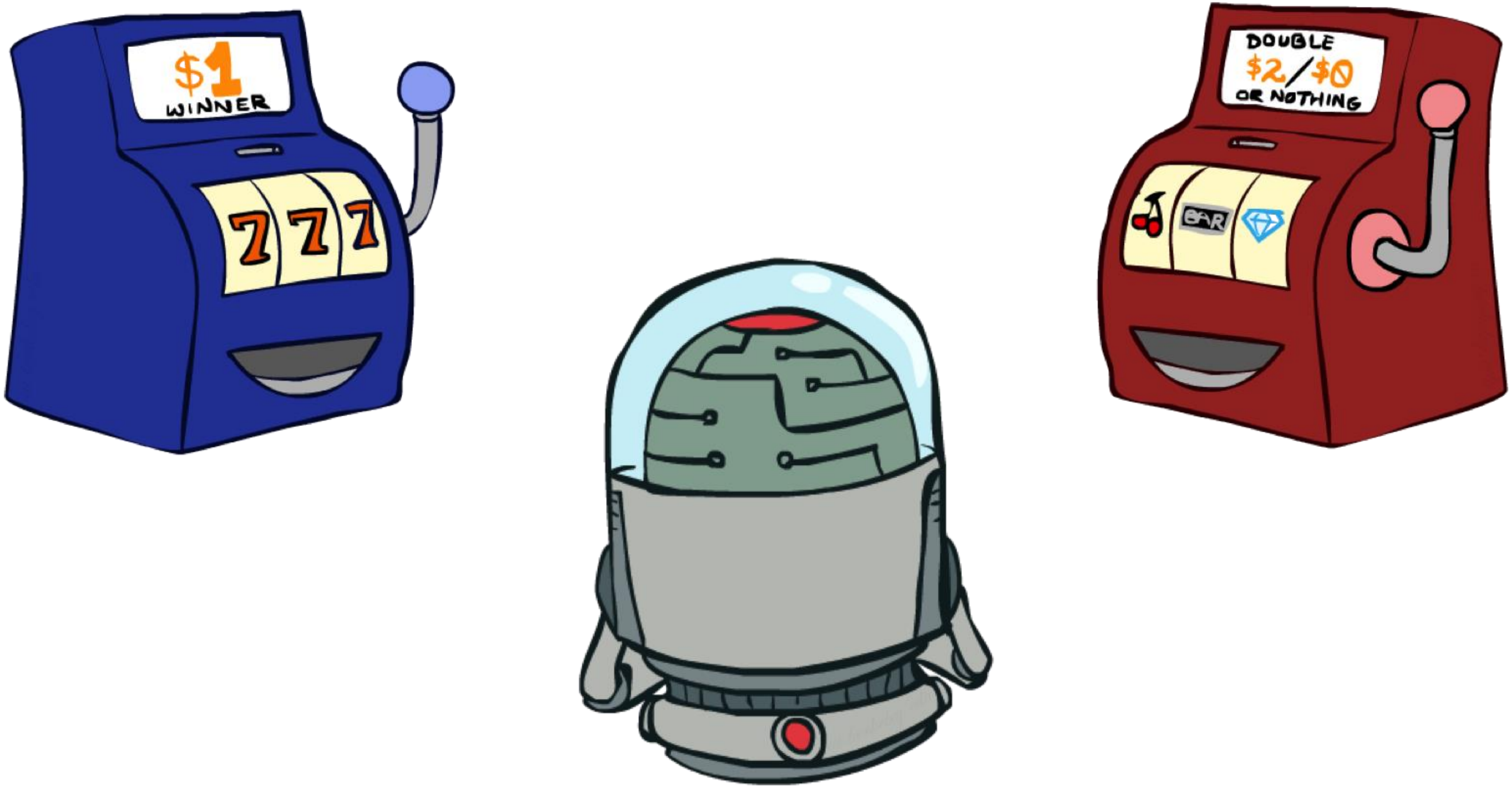
Co-financed by the Connecting Europe Facility of the European Union

The Bellman Equations



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

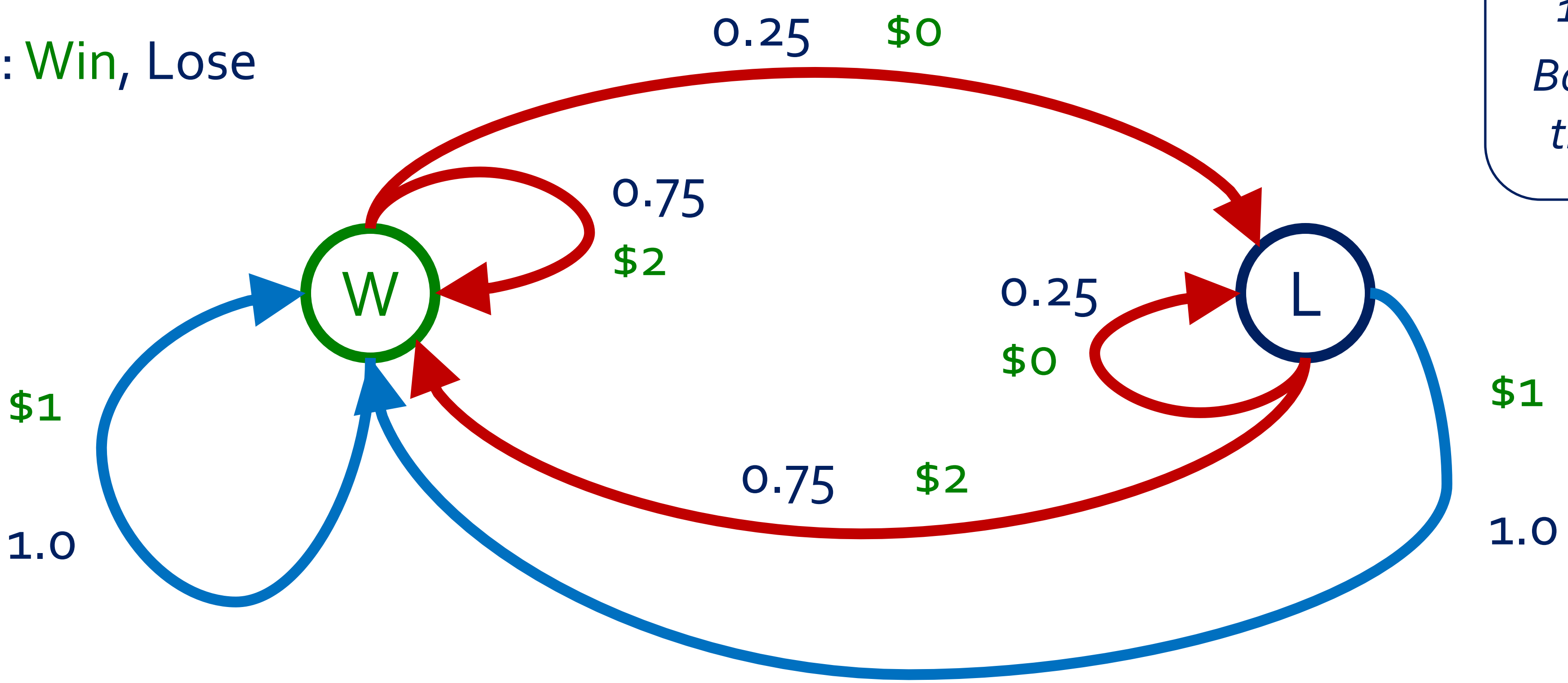
Double Bandits



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

Double-Bandit MDP

- Actions: *Blue*, *Red*
- States: Win, Lose



No discount
100 time steps
Both states have the same value

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

Offline Planning

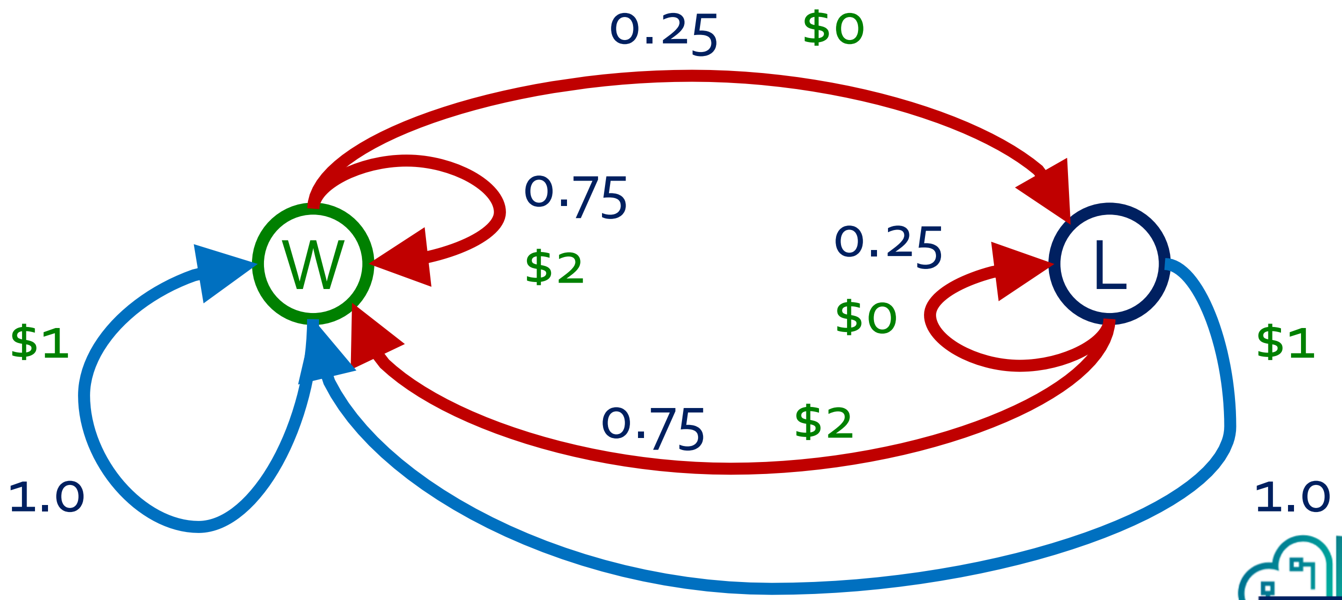
- Solving MDPs is offline planning
  - You determine all quantities through computation
  - You need to know the details of the MDP
  - You do not actually play the game!

No discount

100 time steps

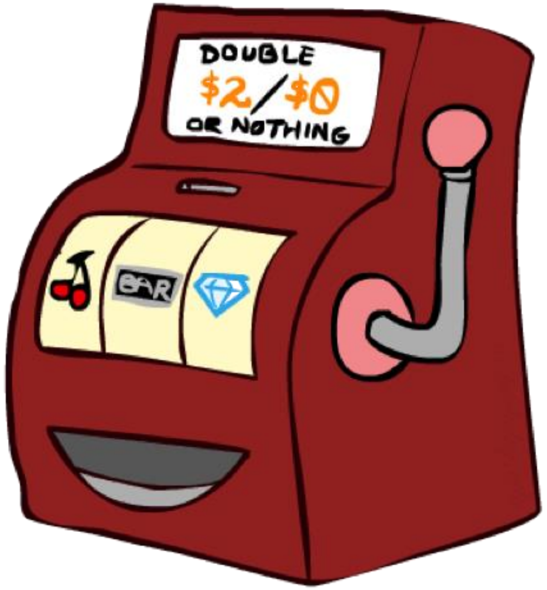
Both states have the same value

	Value
Play Red	150
Play Blue	100



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

Let's Play!



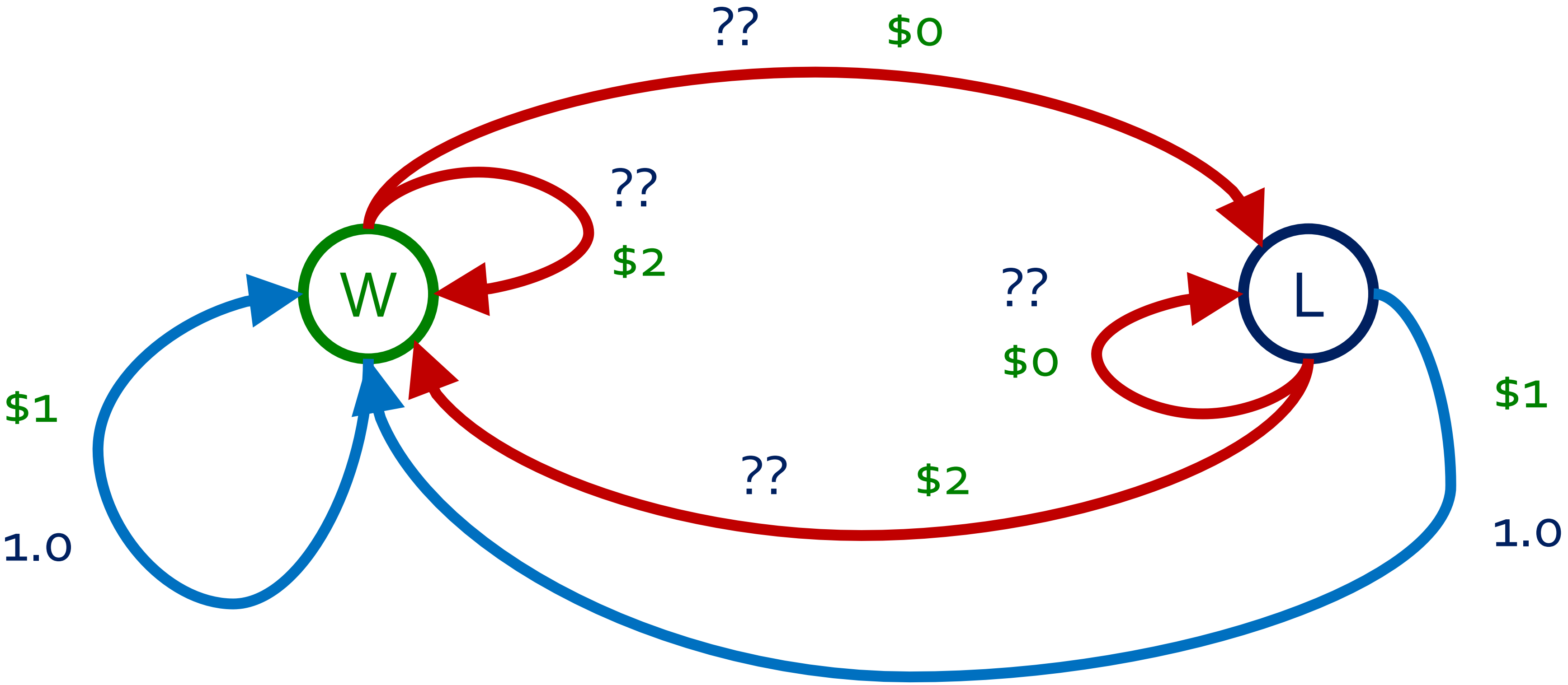
$2 $2 $0 $2 $2

$2 $2 $0 $0 $0

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

Online Planning

- Rules changed! Red's win chance is different.



Let's Play!

$0 $0 $0 $2 $0

$2 $0 $0 $0 $0

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

What Just Happened?

- That wasn't planning, it was learning!
  - Specifically, reinforcement learning
  - There was an MDP, but you couldn't solve it with just computation
  - You needed to actually act to figure it out

- Important ideas in reinforcement learning that came up
  - Exploration: you have to try unknown actions to get information
  - Exploitation: eventually, you have to use what you know
  - Regret: even if you learn intelligently, you make mistakes
  - Sampling: because of chance, you have to try things repeatedly
  - Difficulty: learning can be much harder than solving a known MDP

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

Advanced Topics in AI

Next: 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