Learning and Multiagent Reasoning for Autonomous Agents

Prof. Peter Stone

Director, Learning Agents Research Group
Department of Computer Sciences
The University of Texas at Austin

IJCAI 2007 Computers and Thought Lecture

A Goal of AI

Robust, **fully autonomous** agents in the real world

How?

- Build complete solutions to relevant challenge tasks
  Complete agents: sense, decide, and act — closed loop
  Challenge tasks: specific, concrete objectives
- Drives research on component algorithms, theory
  - Improve from experience (Machine learning)
  - Interact with other agents (Multiagent systems)
- A top-down, empirical approach

Bottom-Up Metaphors

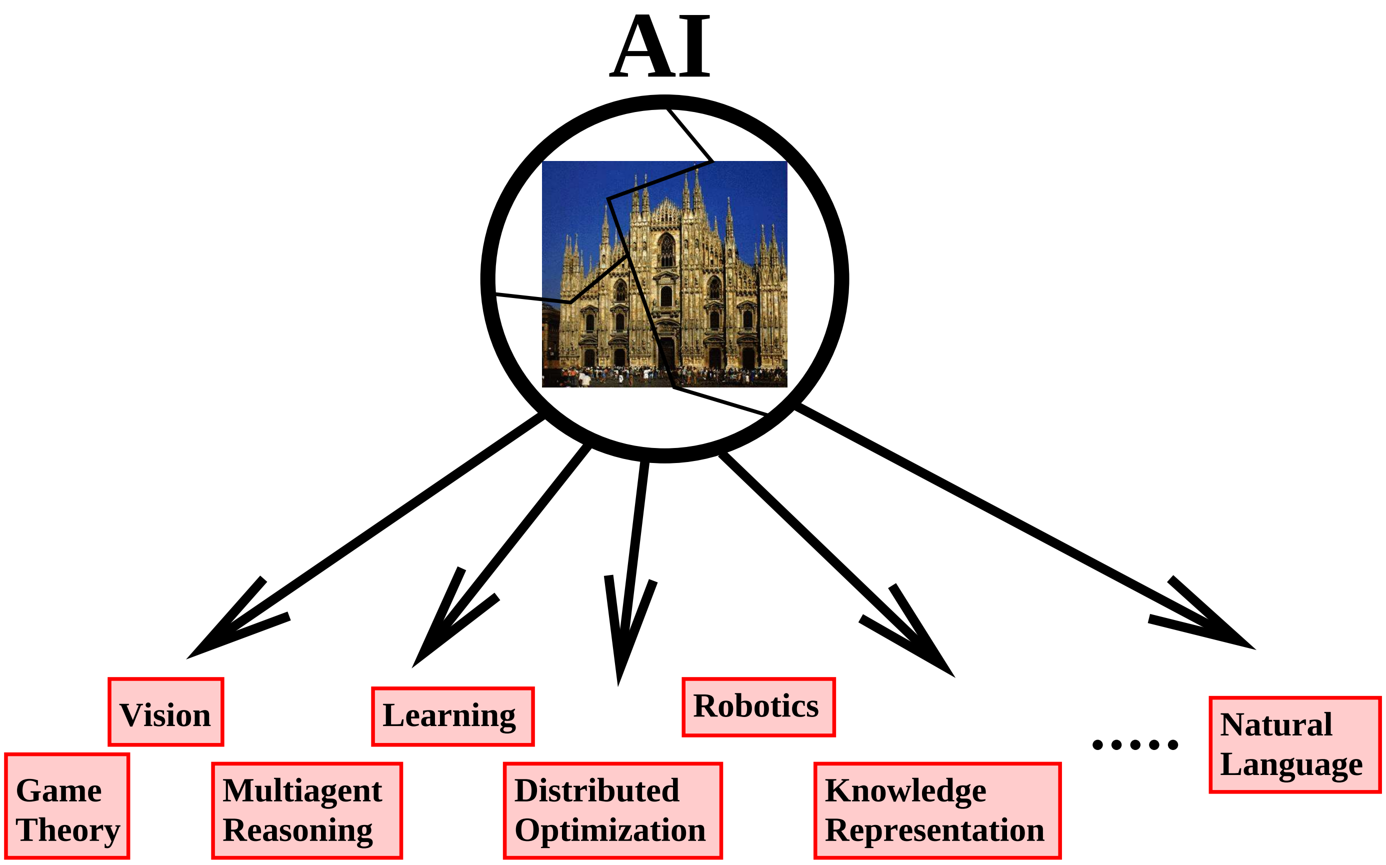
Russell, '95

"Theoreticians can produce the AI equivalent of bricks, beams, and mortar with which AI architects can build the equivalent of cathedrals."

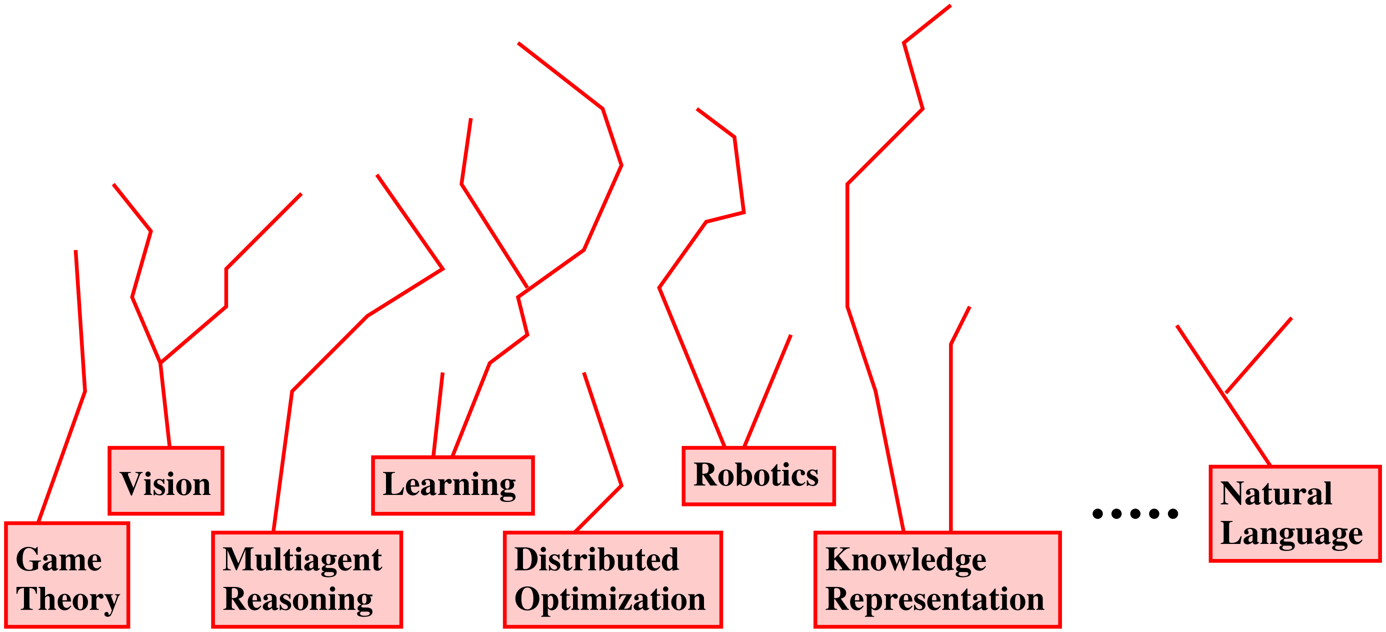
Koller, '01

"In AI . . . we have the tendency to divide a problem into well-defined pieces, and make progress on each one. . . . Part of our solution to the AI problem must involve building bridges between the pieces."

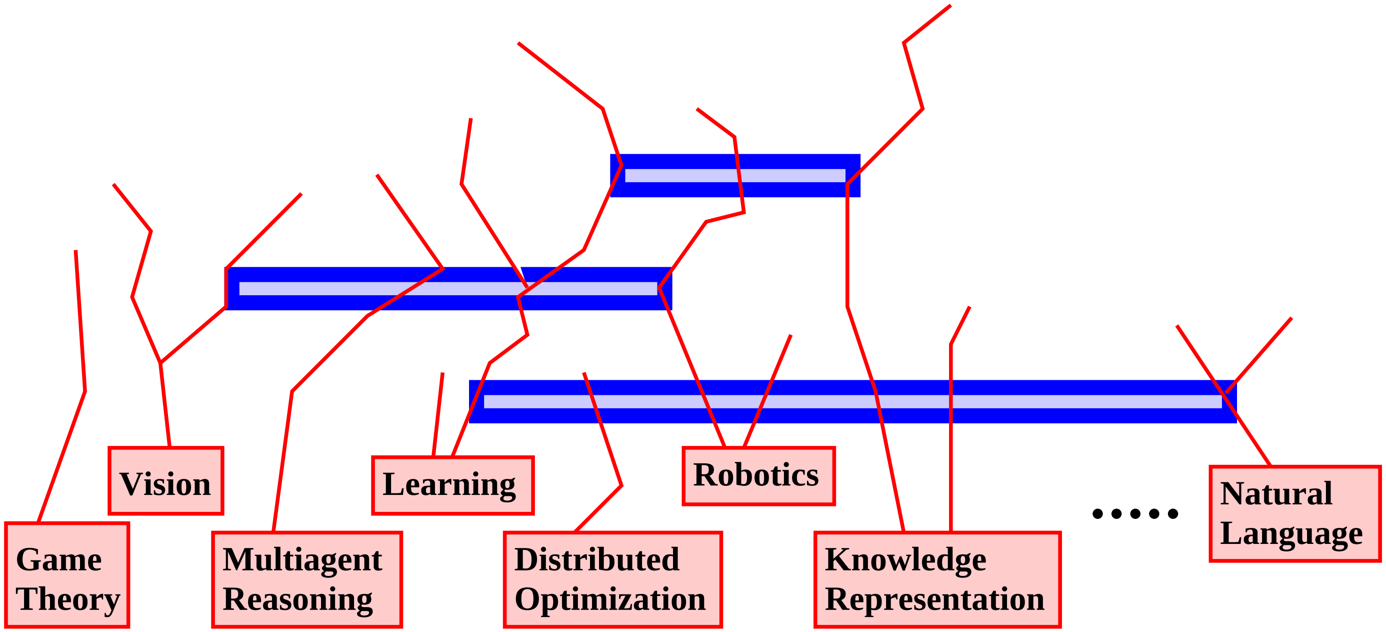
Dividing the Problem



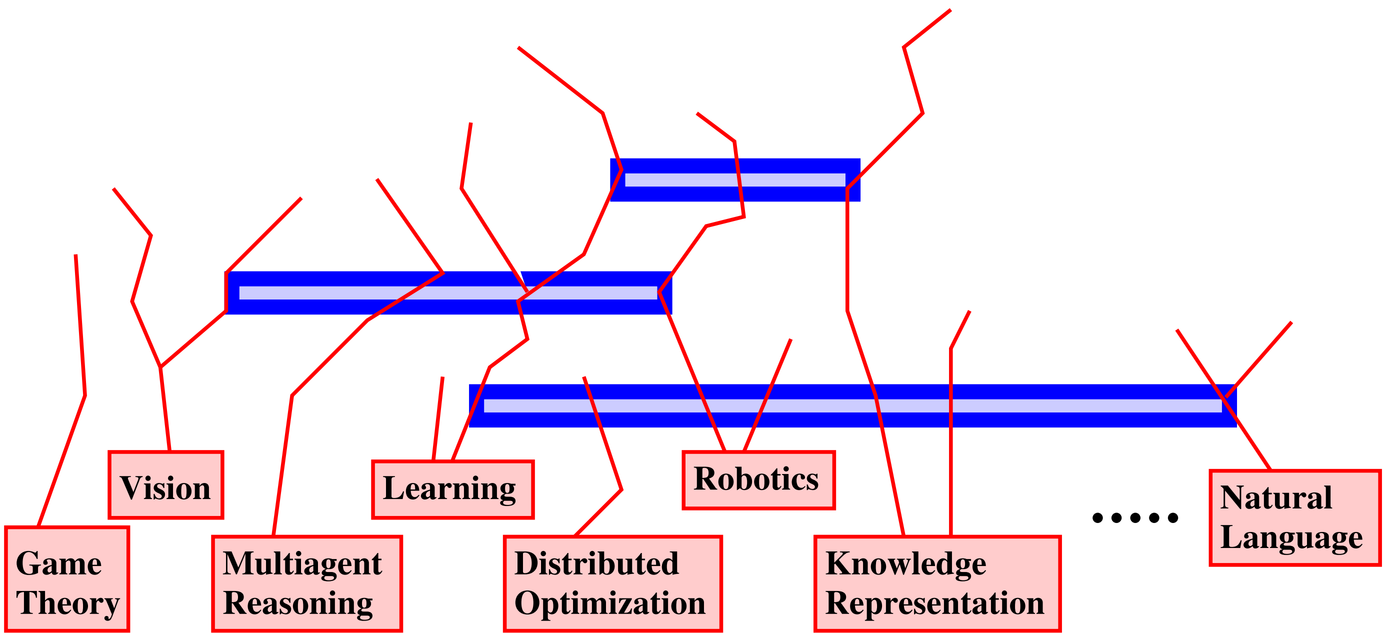
The Bricks


Vision
Learning
Robotics
Game Theory
Multiagent Reasoning
Distributed Optimization
Knowledge Representation
.....
Natural Language

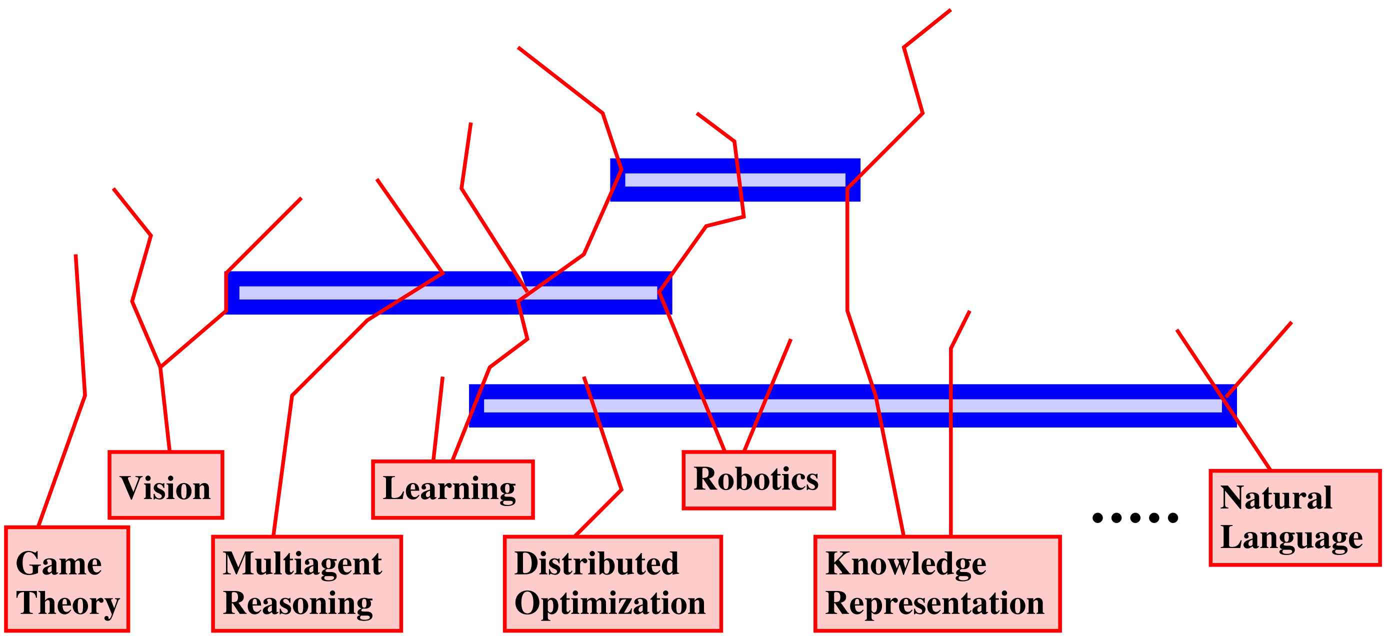
The Beams and Mortar



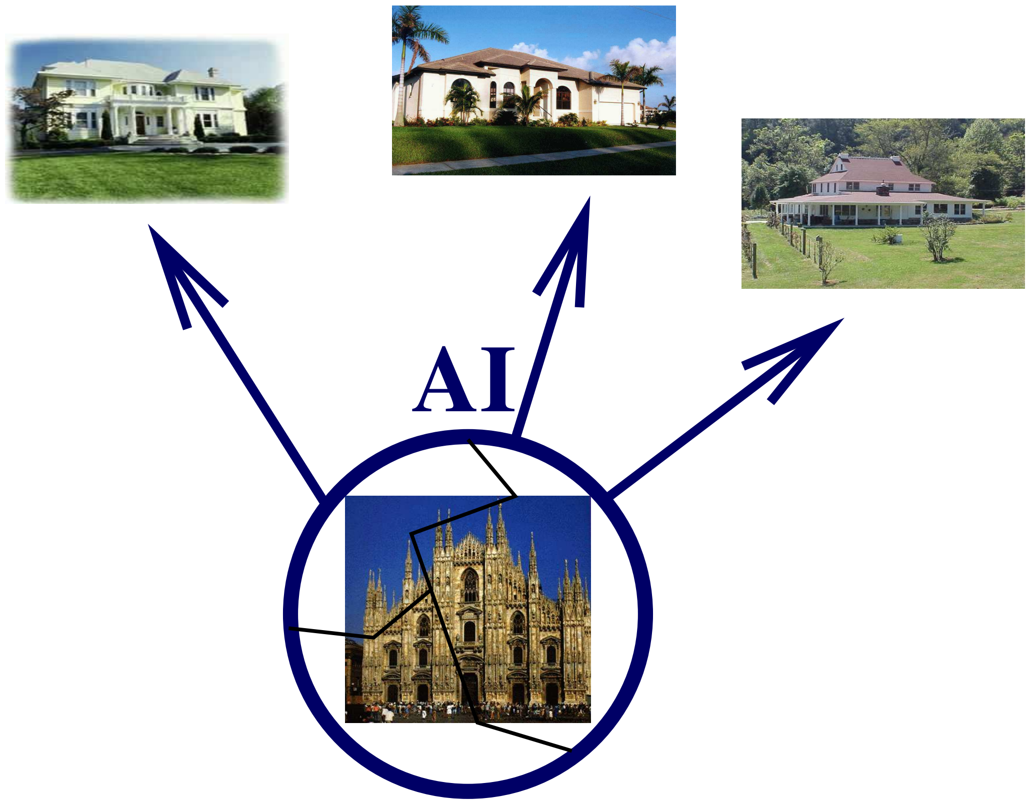
Towards a Cathedral?



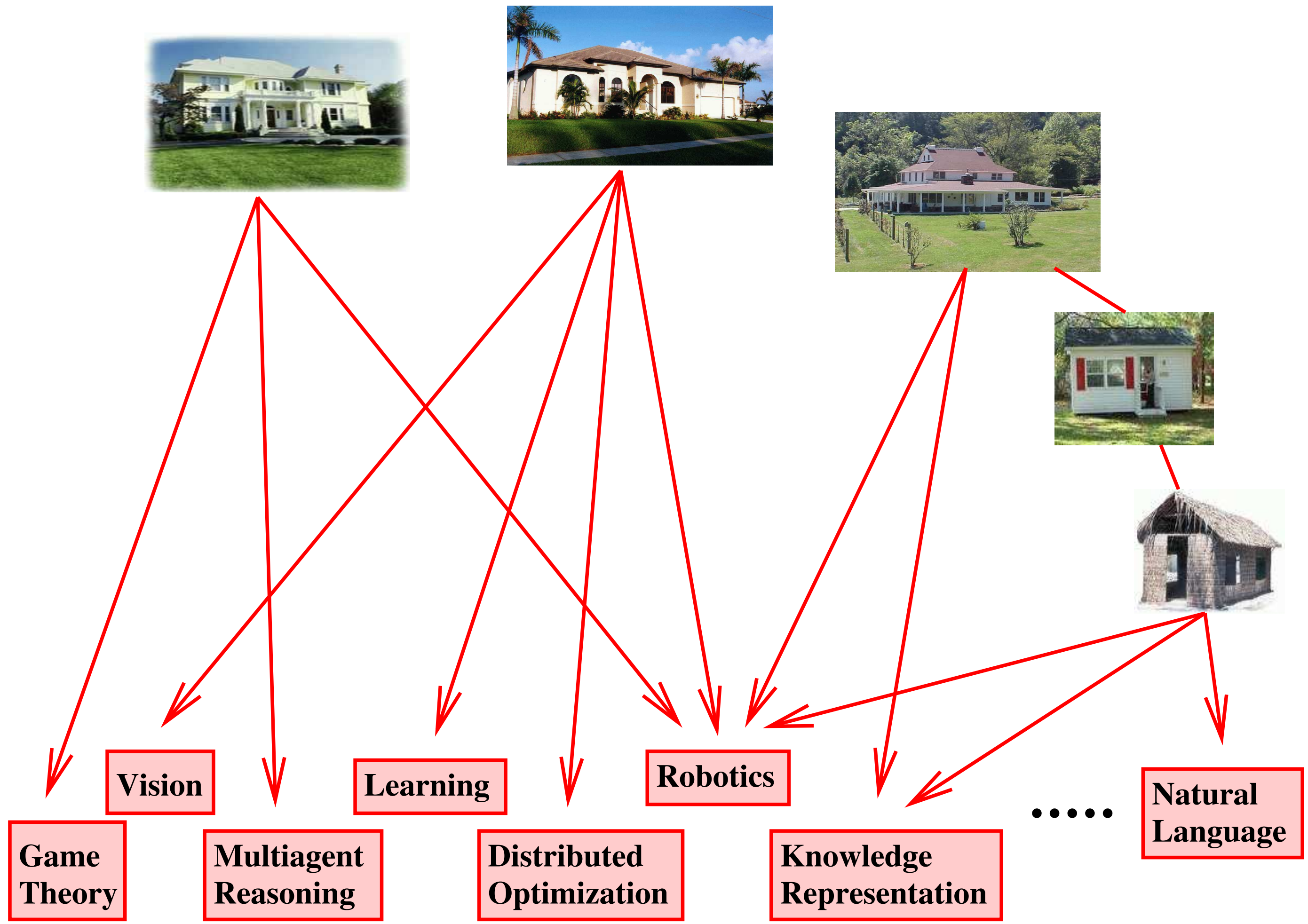
Or Something Else?



A Different Problem Division

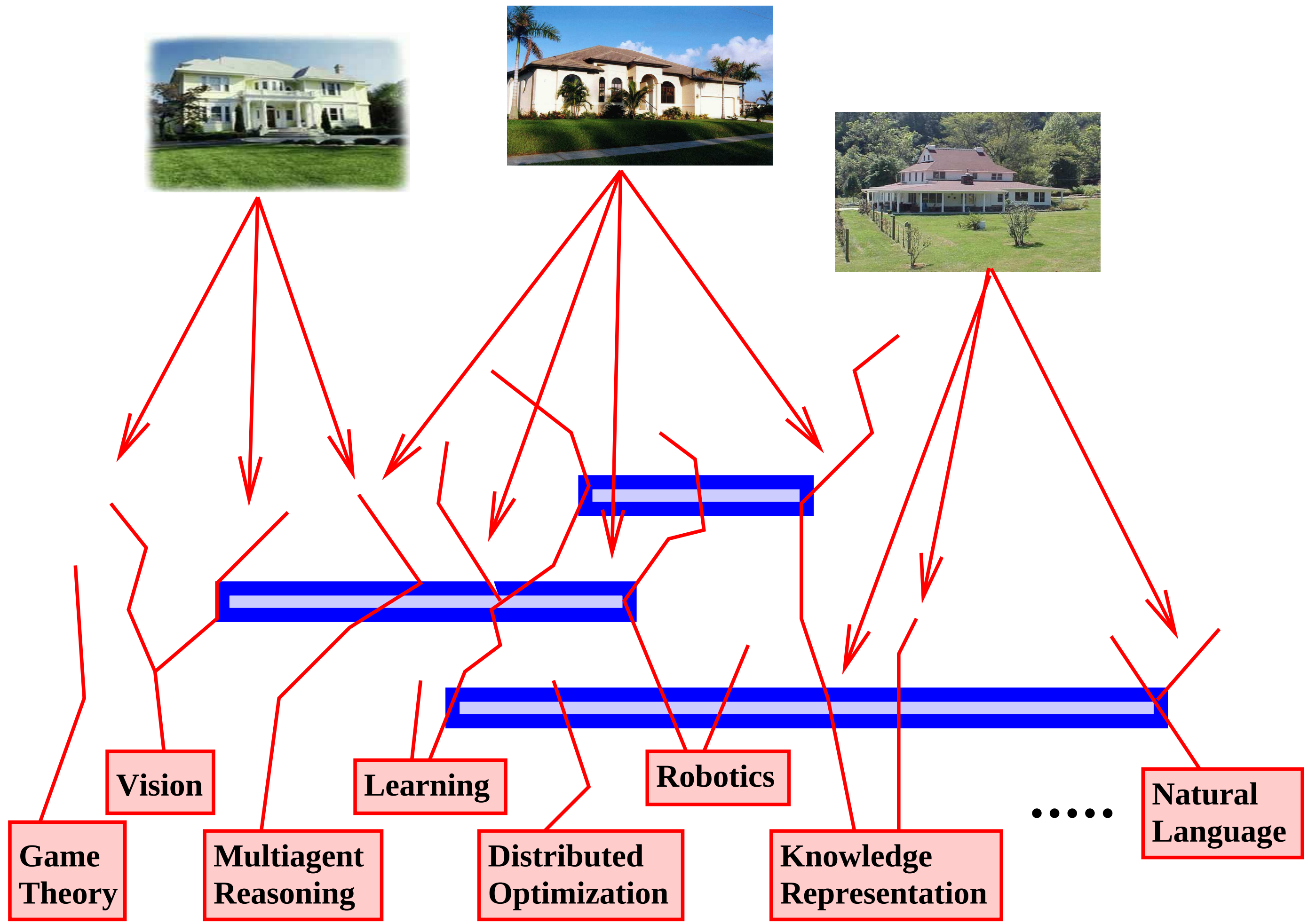


Top-Down Approach



"Good problems . . . produce good science" [Cohen, '04]

Meeting in the Middle



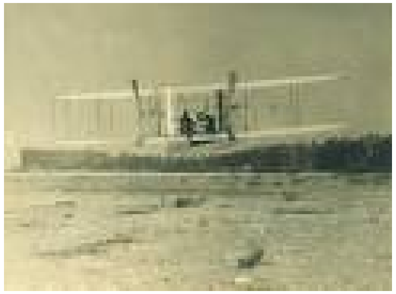
Top-down approaches underrepresented: (IJCAI '05: 35/237)

Choosing the Challenge

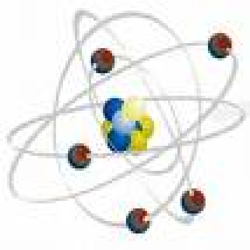
- Features of good challenges: [Cohen, '04]
  - Frequent tests; Graduated series of challenges
  - Accept poor performance; Complete agents
- Closed loop + specific goal (beyond [Brooks, '91])
- 50-year technical, scientific goals
  - Beyond commercial applications — not possible now
  - Moore's law not enough
- There are many — choose one that inspires you
  - Leverage "bricks and mortar" from past
  - Hybrid symbolic/probabilistic methods

[Richardson & Domingos, '06]

Good Problems Produce Good Science

Manned flight 

Apollo mission 

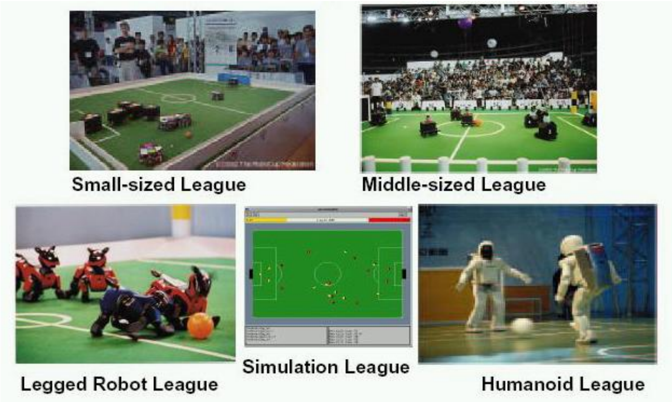
Manhattan project 

RoboCup soccer 

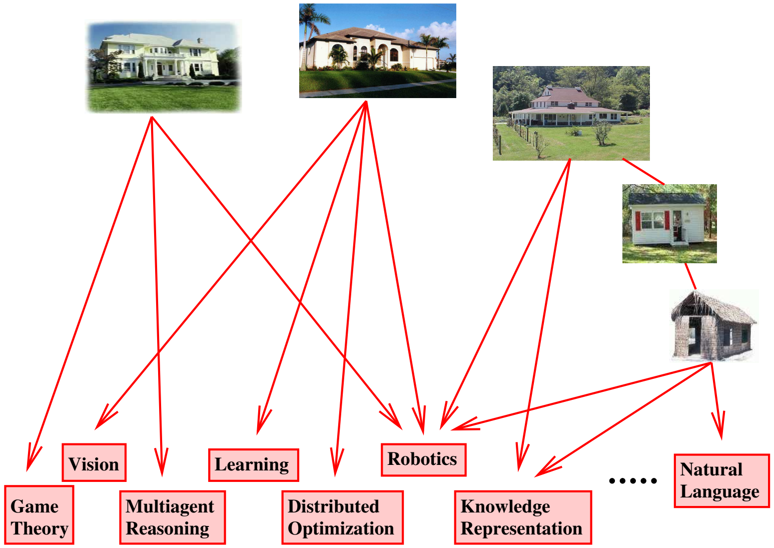
Goal: By the year 2050, a team of humanoid robots that can beat the human World Cup champion team.
[Kitano, '97]

RoboCup Soccer

- Still in the early stages (small houses)

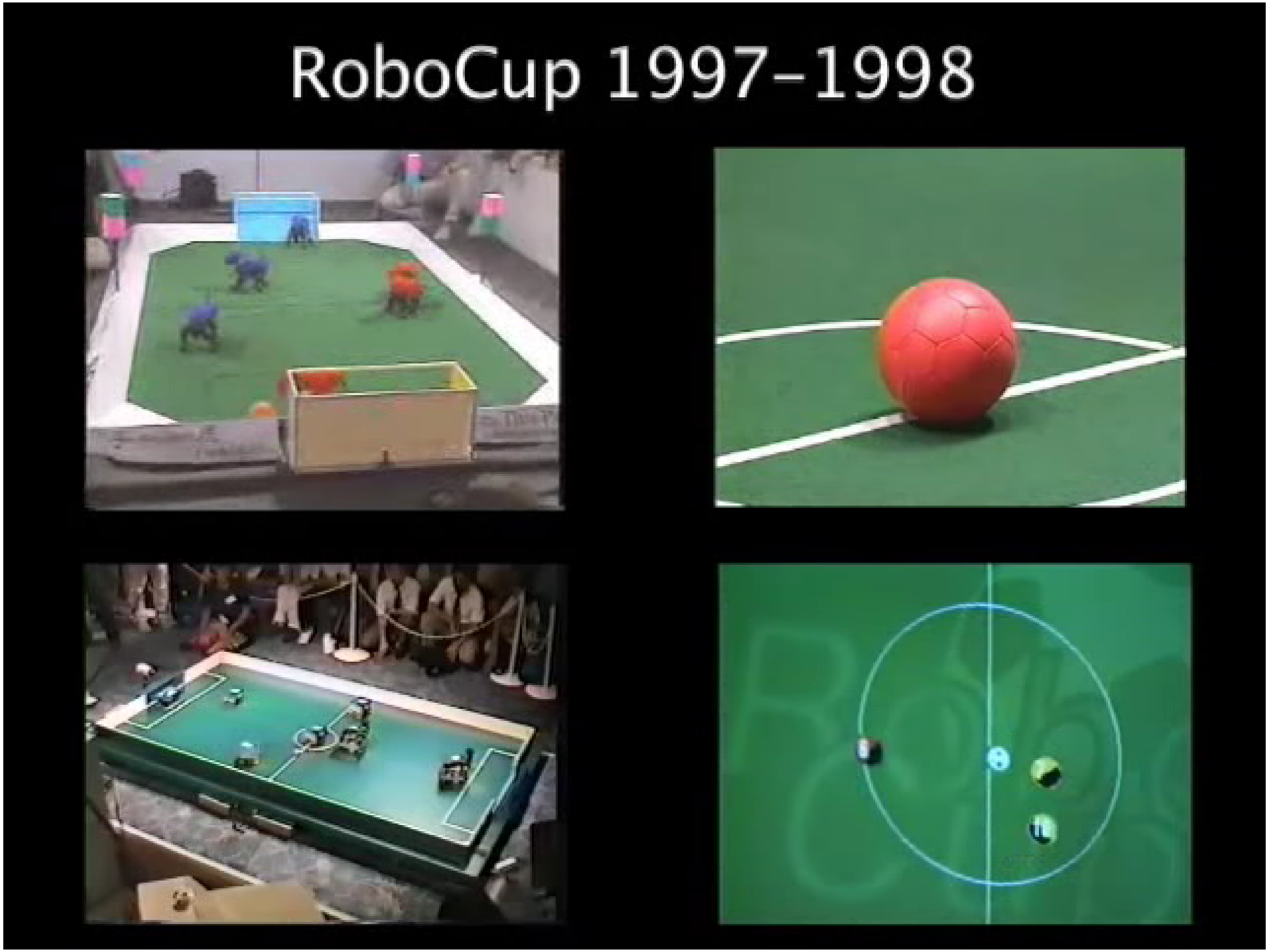


- Many virtues:
  - Incremental challenges, closed loop at each stage
  - Relatively easy entry
  - Multiple robots possible
  - Inspiring to many

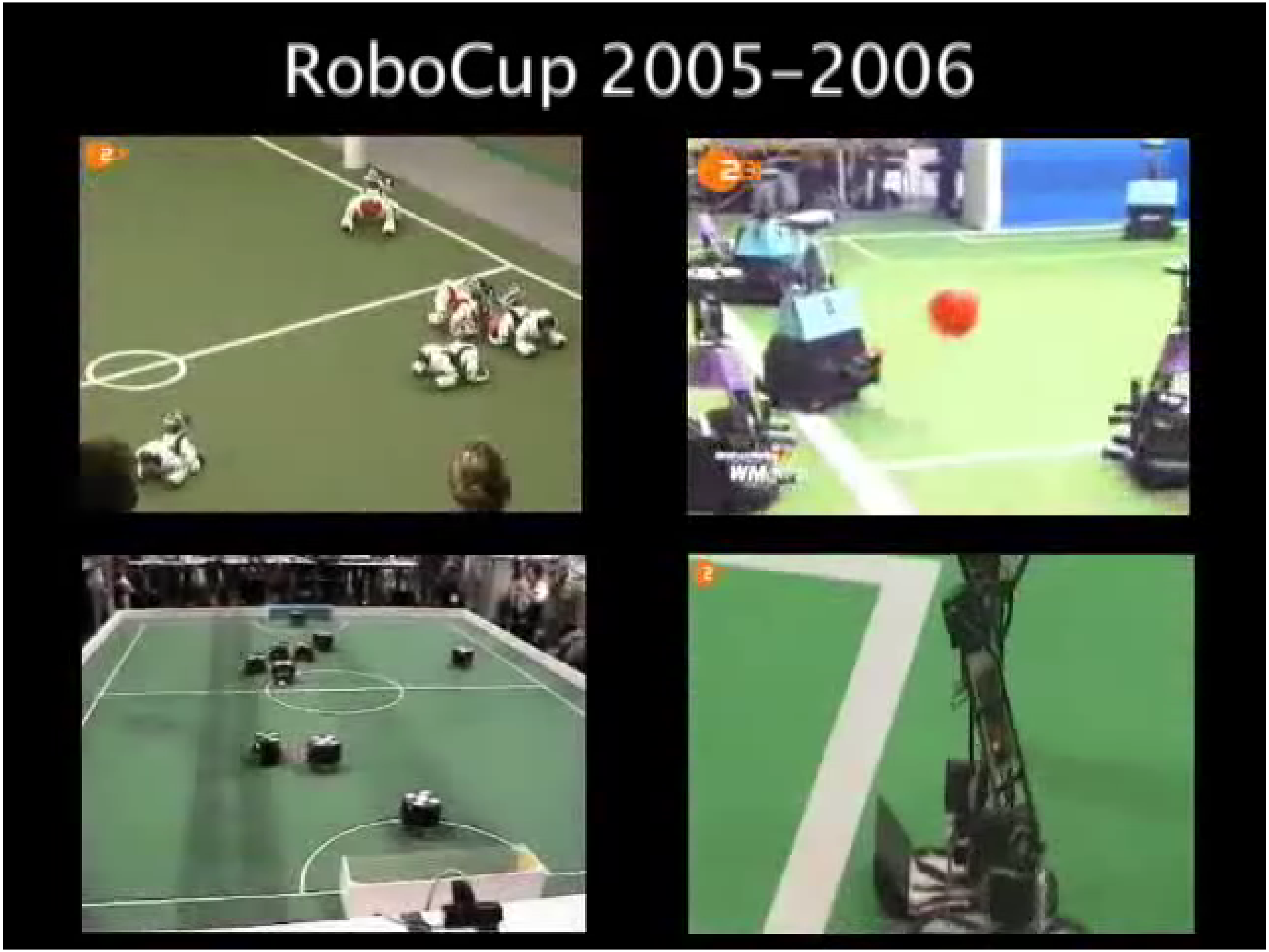


- Visible progress

The Early Years

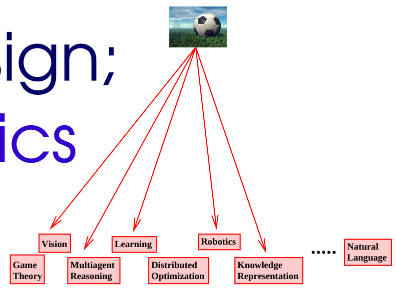


A Decade Later



Advances due to RoboCup

- Drives **research** in many areas:
  - Control algorithms; computer vision, sensing; localization;
  - Distributed computing; real-time systems;
  - Knowledge representation; mechanical design;
  - Multiagent systems; machine learning; robotics



- **200+ publications** from simulation league alone
- **200+** from 4-legged league
- **15+** Ph.D. theses

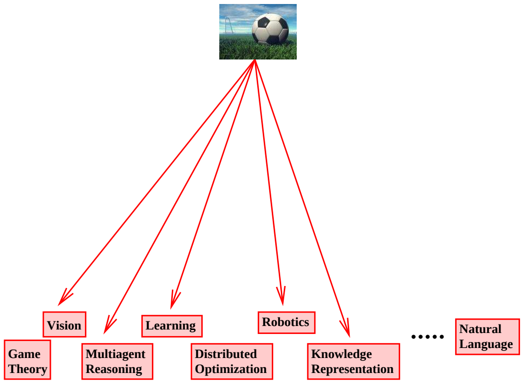
Layered Learning

- For domains too complex for tractably mapping state features $S \longmapsto$ outputs O
- Hierarchical subtask decomposition given: $\{L_1, L_2, \ldots, L_n\}$
- Machine learning: exploit data to train, adapt
- Learning in one layer feeds into next layer

Layered Learning in Practice

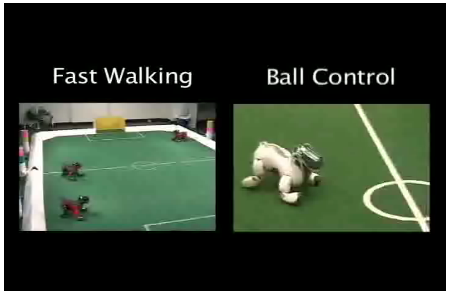
First applied in simulated robot soccer [Stone & Veloso, '97]

	Strategic Level	Example
L_1	individual	ball interception
L_2	multiagent	pass evaluation
L_3	team	pass selection

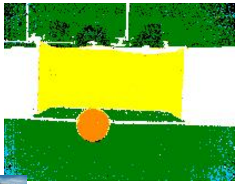


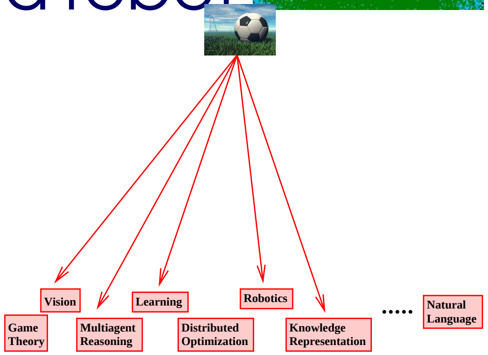
Recently applied on real robots [Stone, Kohl, & Fidelman, '06]

	Strategic Level	Example
L_1	individual	fast walking
L_2	individual	ball control



Robot Vision

- Great progress in computer vision
  - Shape modeling, object recognition, face detection...
- Robot vision offers new challenges 
  - Mobile camera, limited computation, color features
- **Autonomous color learning** [Sridharan & Stone, '05]
  - Learns color map based on known object locations
  - Recognizes and reacts to illumination changes 
  - Object detection in real-time, on-board a robot

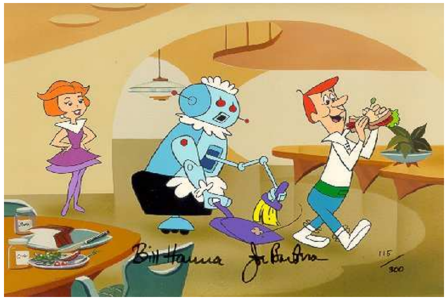


Other Good AI Challenges

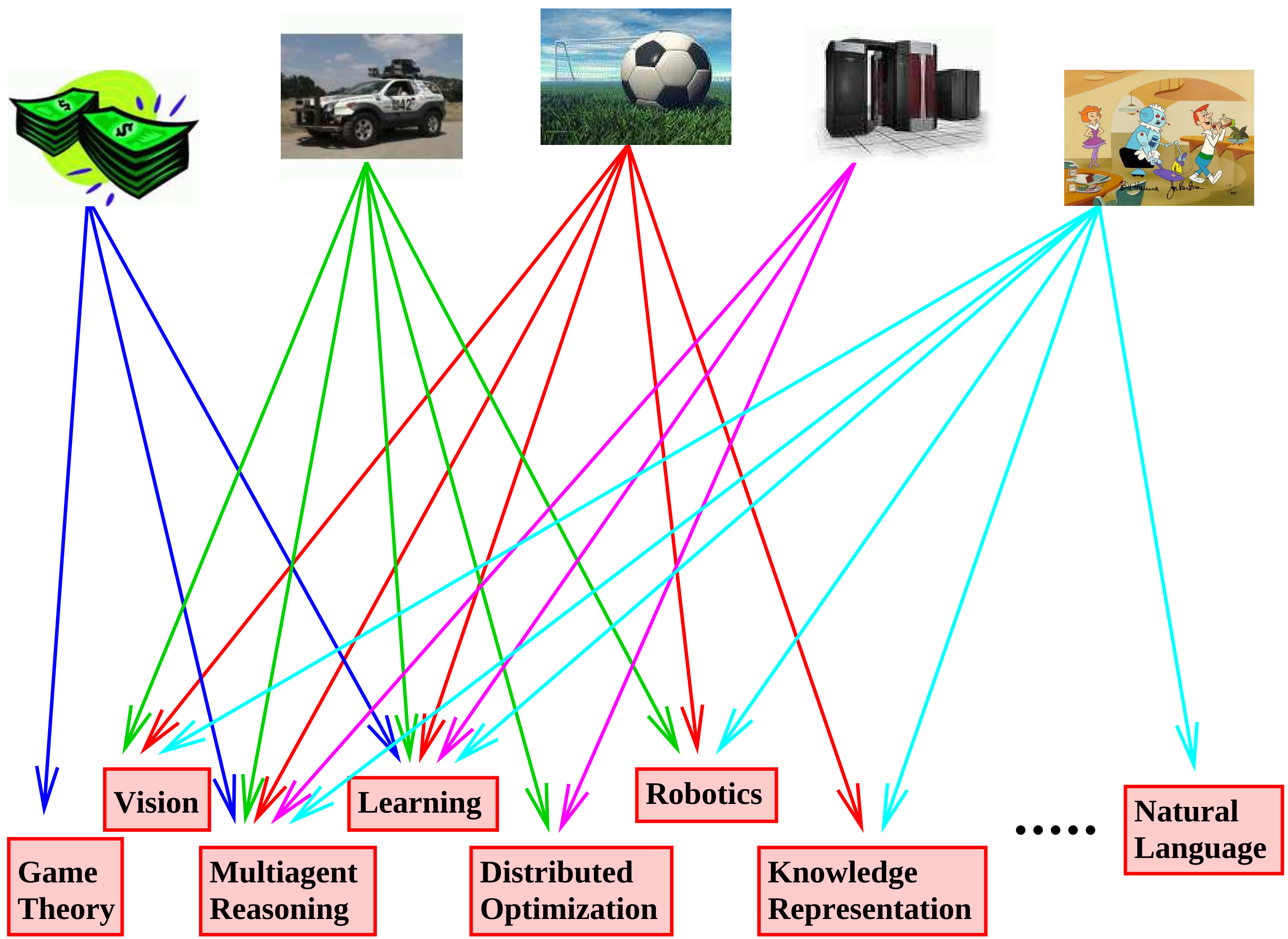
Trading agents 

Autonomous vehicles 

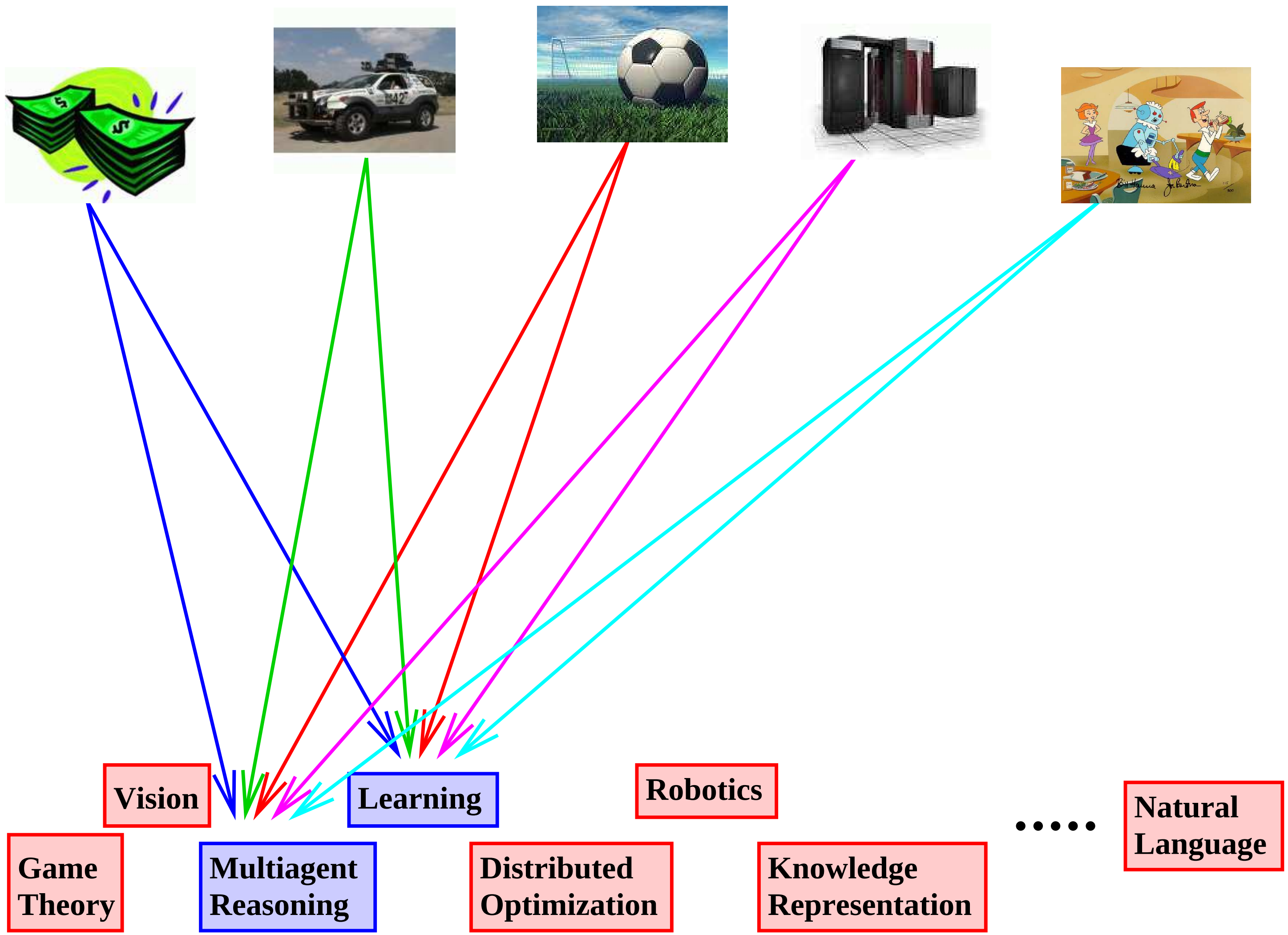
Autonomic computing 

Socially assistive robots 

Challenge Problems Drive Research

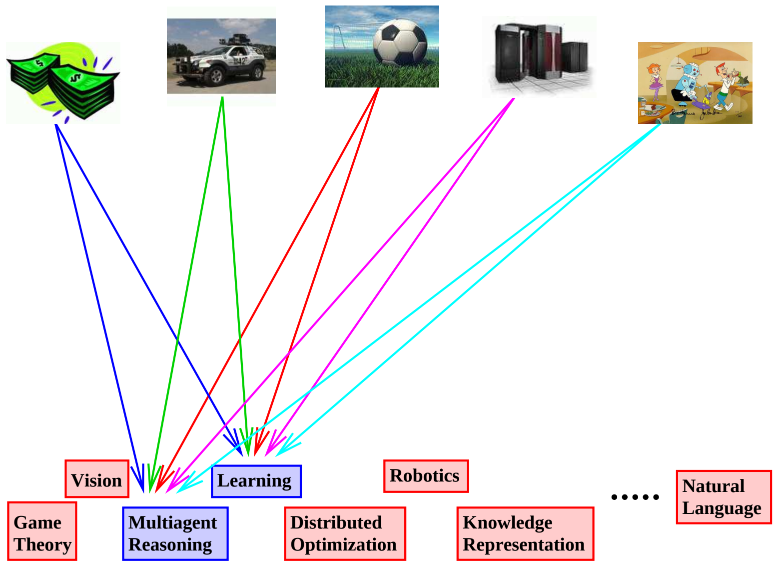


Learning and Multiagent Reasoning



Outline

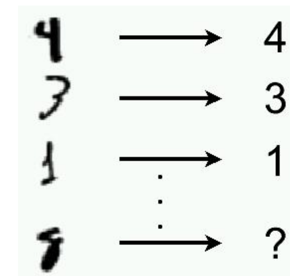
- Build complete solutions to relevant challenge tasks
  - Drives research on component algorithms, theory
- **Learning Agents**
  - Scaling up Reinforcement Learning
  - Adaptive representations
- Multiagent reasoning
  - Prepare for the unexpected
  - Adaptive interaction protocols



Machine Learning

"... resurgence of interest in machine learning" [Mitchell, '83]

Supervised learning mature [Kaelbling, '97]

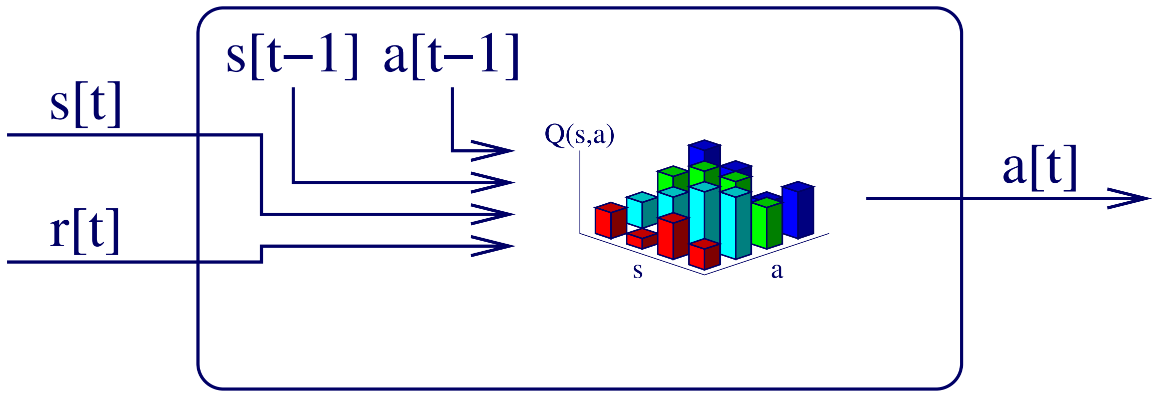


For agents, reinforcement learning most appropriate

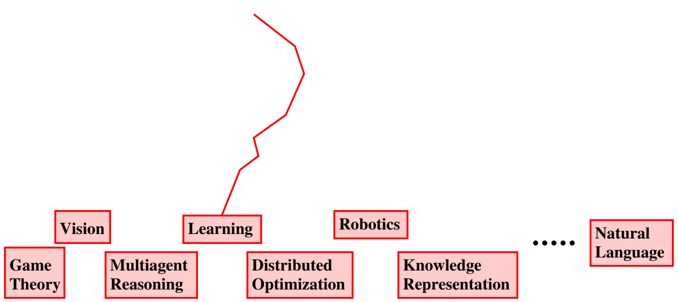
– Foundational theoretical results
– Challenge problems require innovations to scale up

RL Theory

Success story: Q-learning converges to π^* [Watkins, 89]



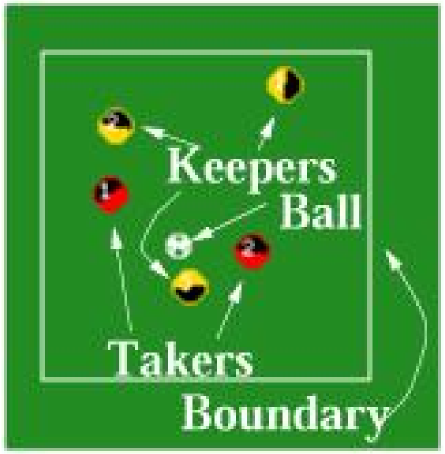
– Table-based representation
– Visit every state infinitely often

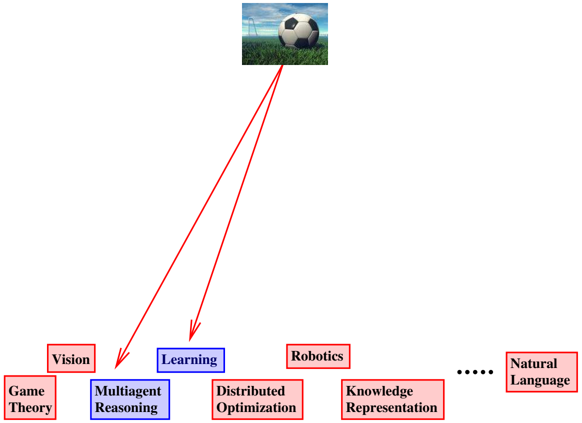


Scaling Up

- Backgammon [Tesauro, '94]
- Helicopter control [Ng et al., '03]

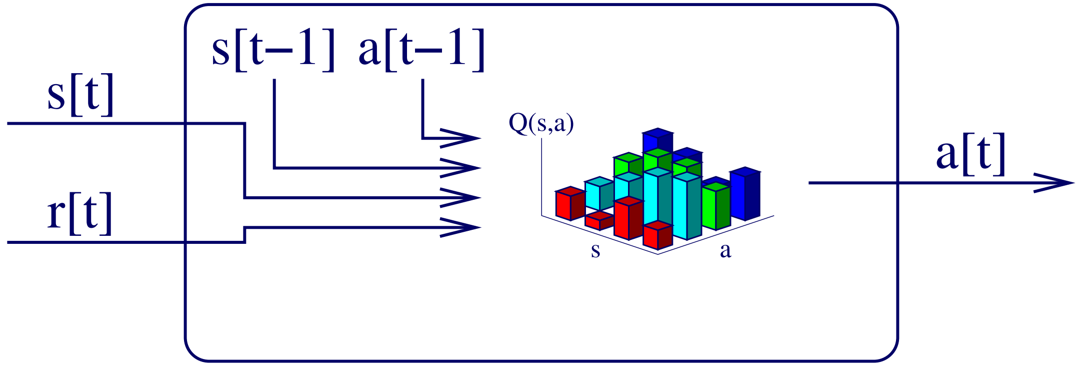
- **RoboCup Soccer Keepaway** [Stone & Sutton, '01]
  - Play in a small area (20m $\times$ 20m)
  - Keepers try to keep the ball
  - Takers try to get the ball
  - Performance measure: average possession duration



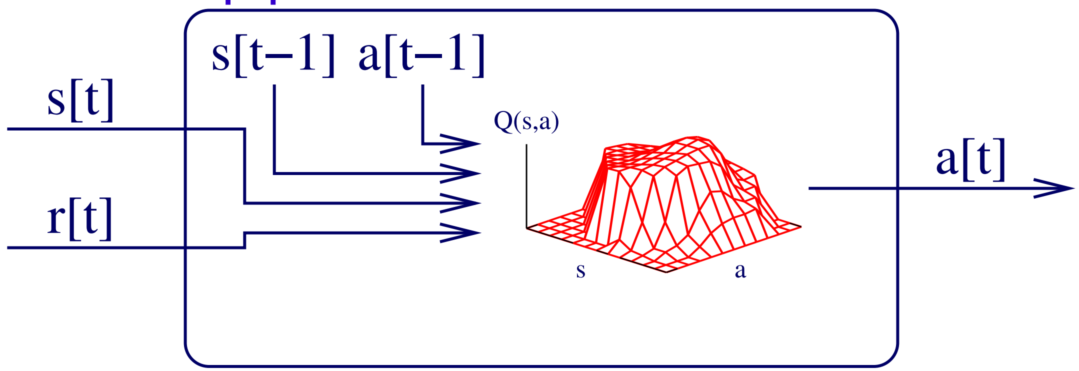


Function Approximation

In practice, visiting every state impossible

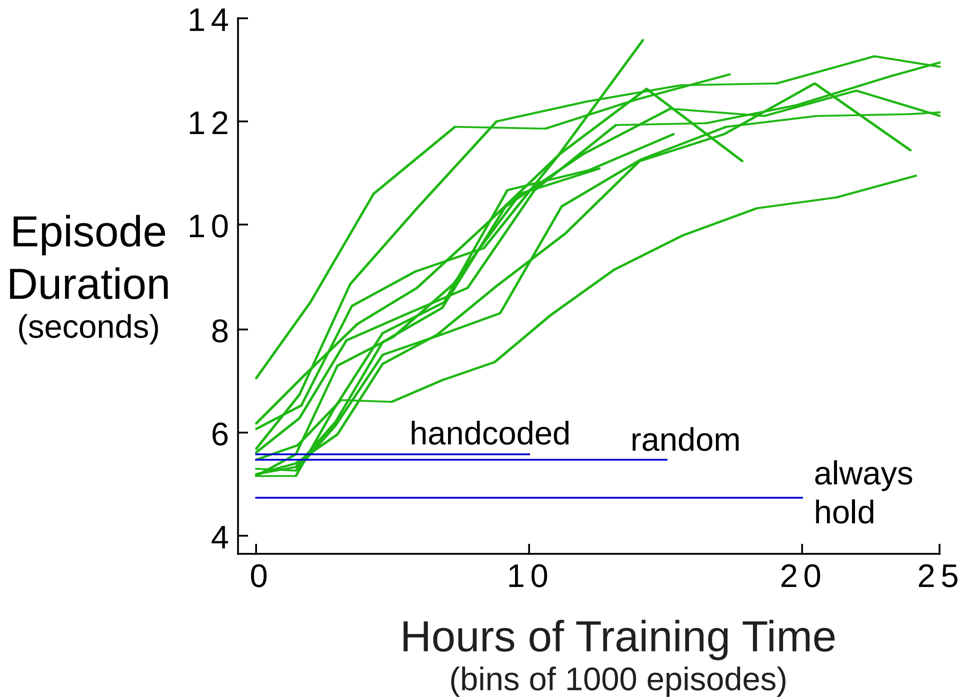


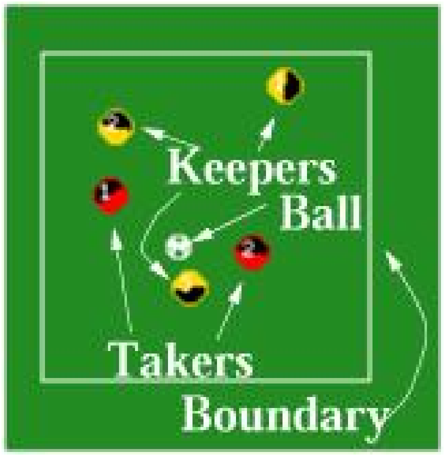
Function approximation of value function



Theoretical guarantees harder to come by

Main Result





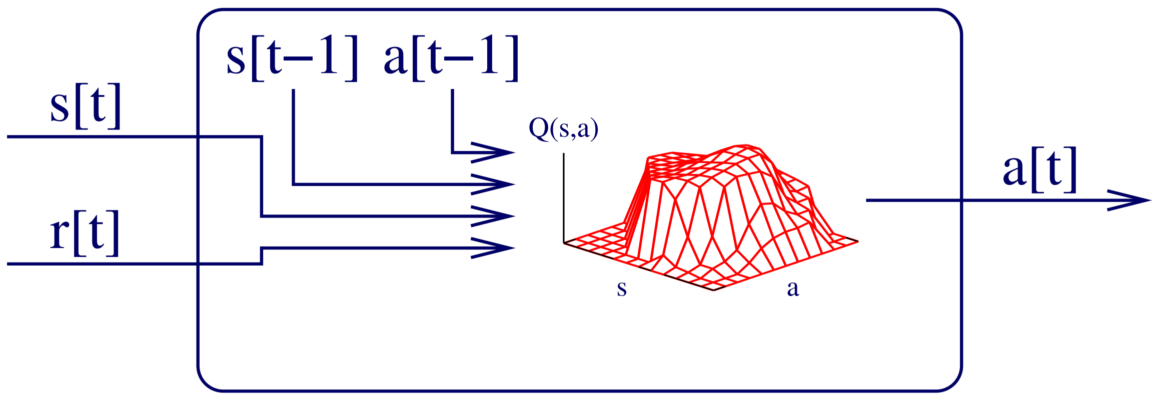
Learning: Distributed SMDP SARSA(λ) with CMACs

– Algorithm modified to enable distributed updates

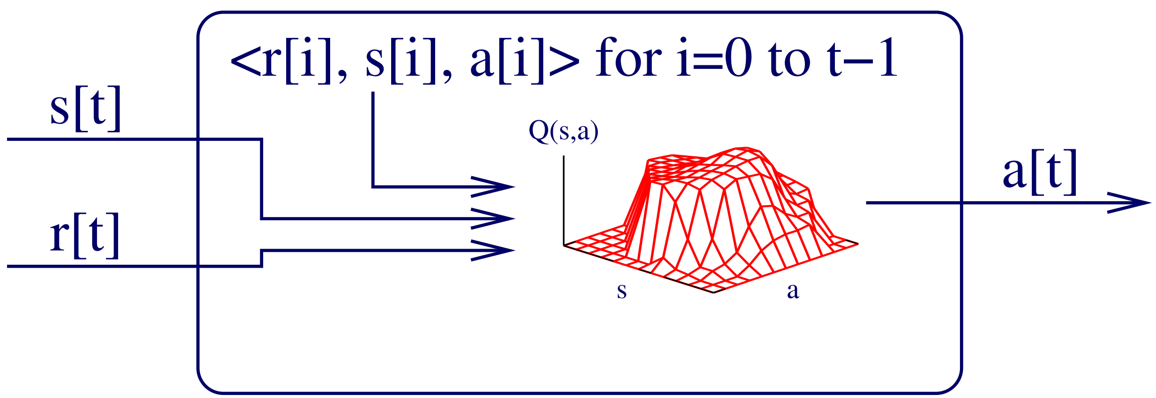
1 hour = 720 5-second episodes

Batch Methods

In practice, often experience is scarce

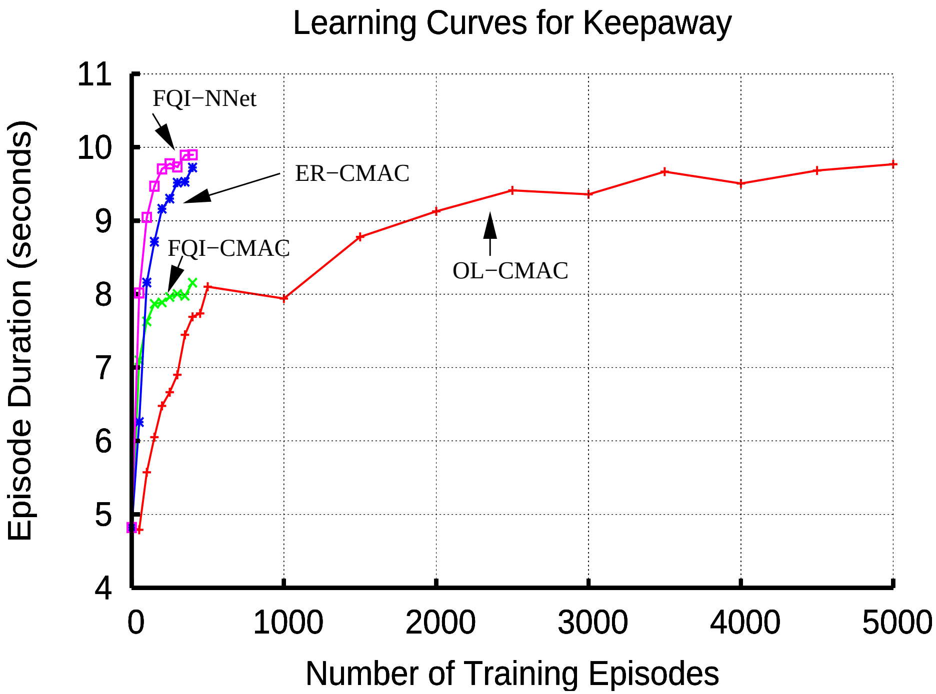


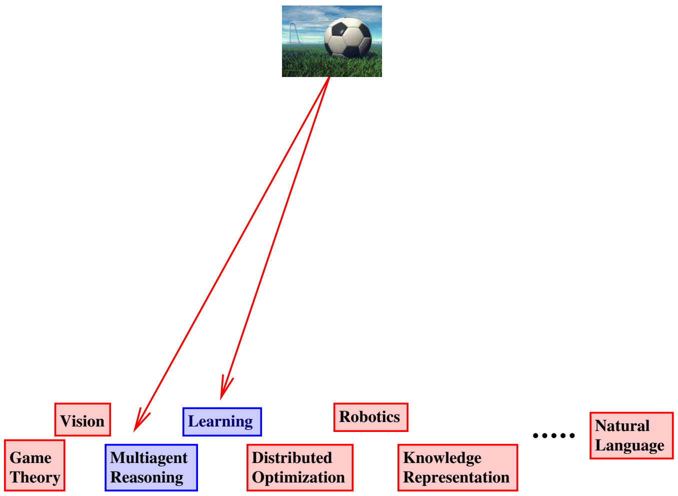
Save transitions:



“Few Zeroes” [Kalyanakrishnan & Stone, ’07]

Experience replay [Lin, ‘92], Fitted Q Iteration [Ernst et al., ‘05]



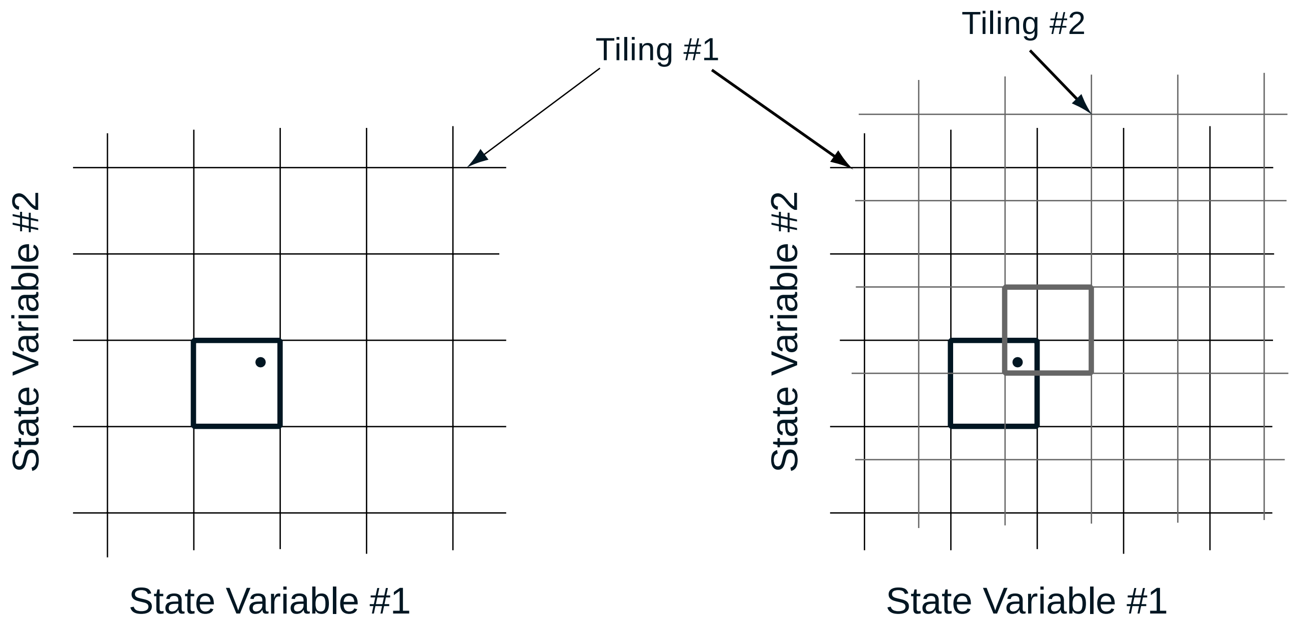


Other ways to scale up

- Advice/demonstration, state/temporal abstraction
- Hierarchical representations, transfer learning

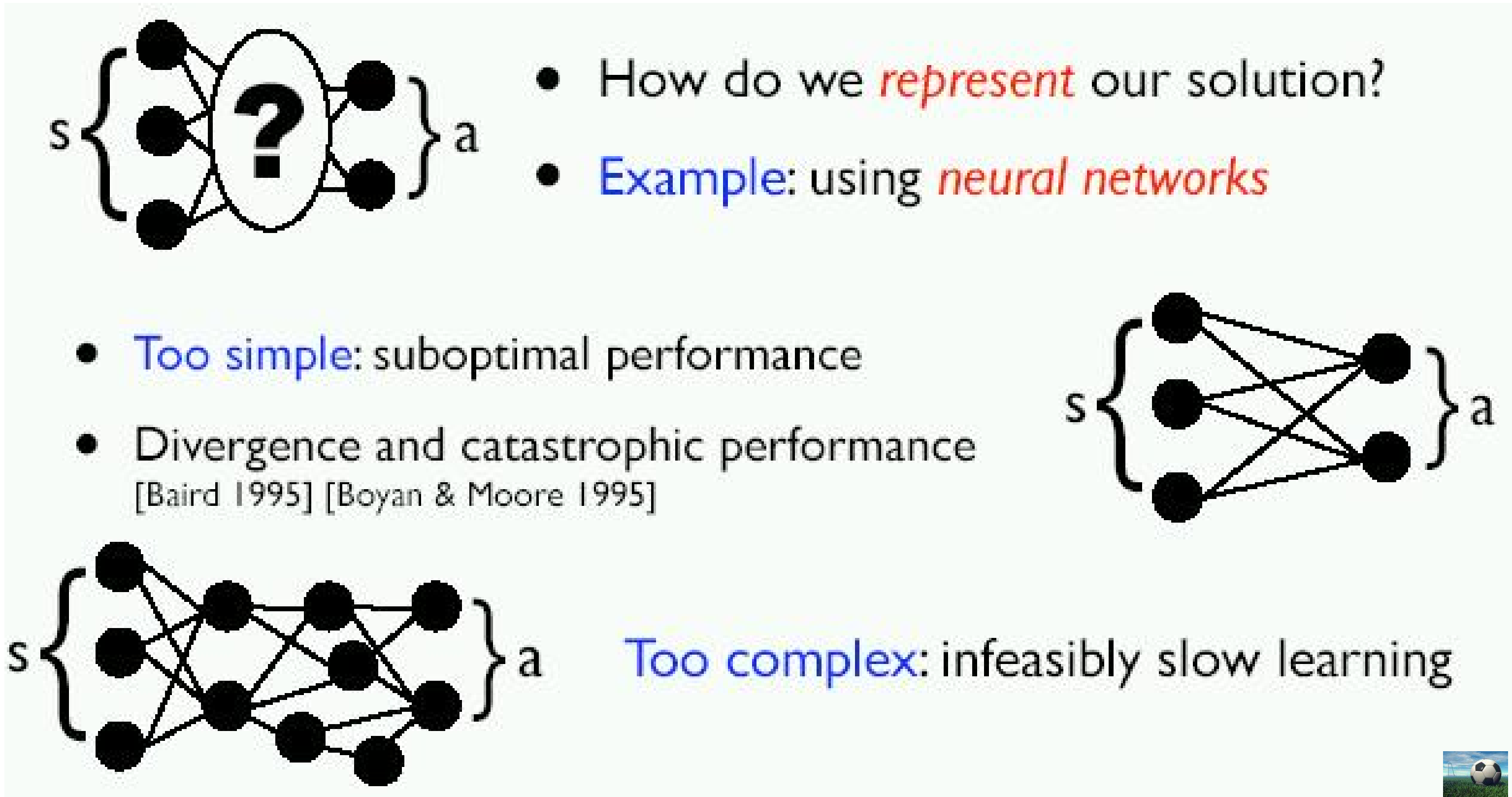
A Big Caveat

So far, **representations chosen manually**

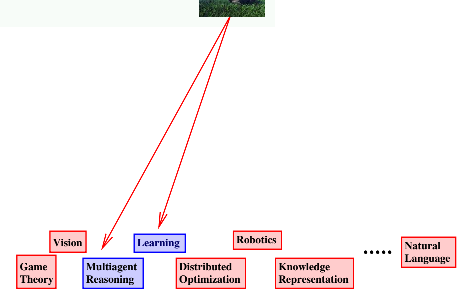


The crucial factor for a successful approximate algorithm is the choice of the parametric approximation architecture. . . ." [Lagoudakis & Parr,'03]

Representations for RL



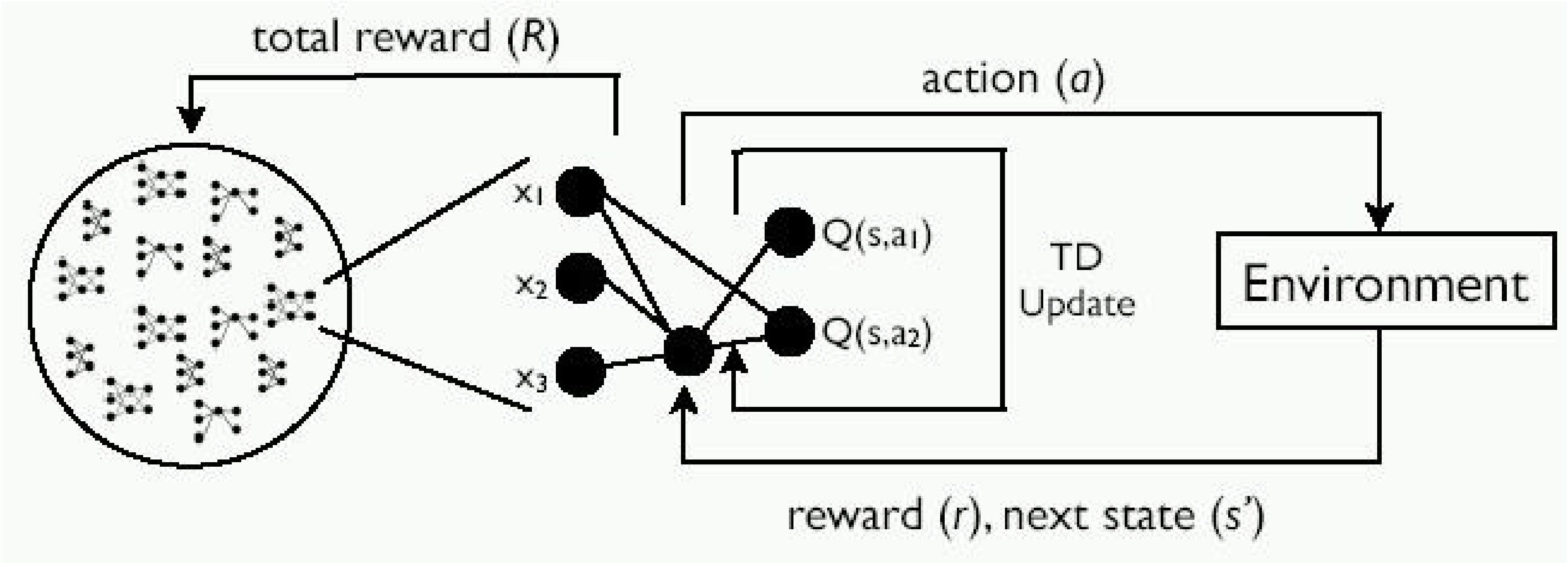
Can RL agents automatically **learn** effective **representations?**



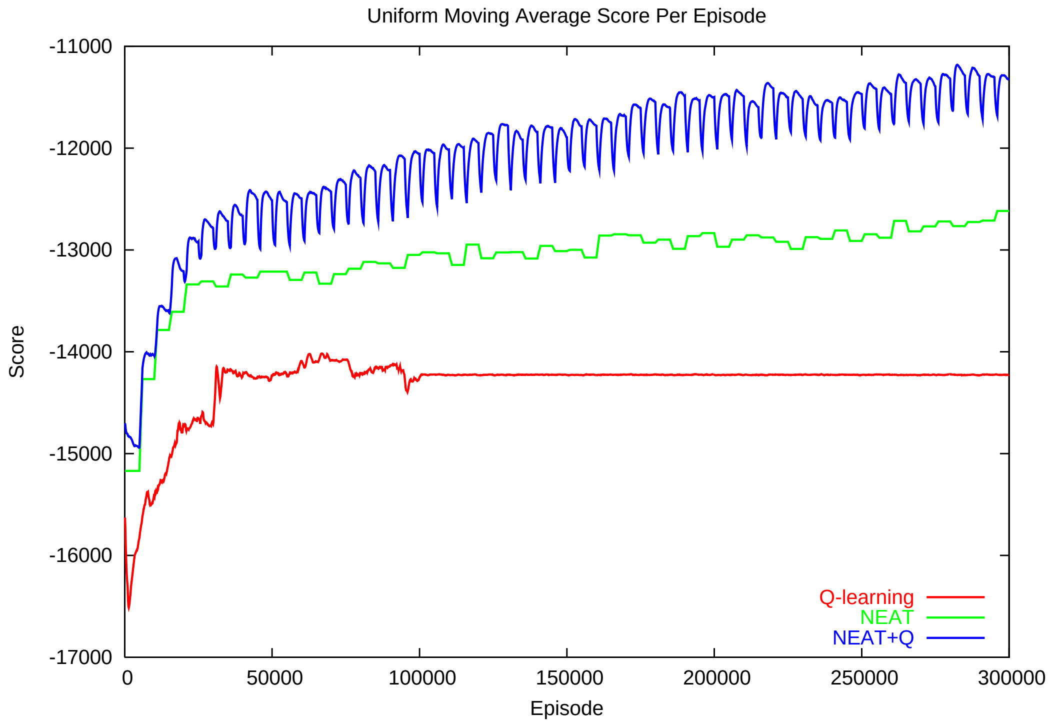
NEAT+Q [Whiteson & Stone, '06]

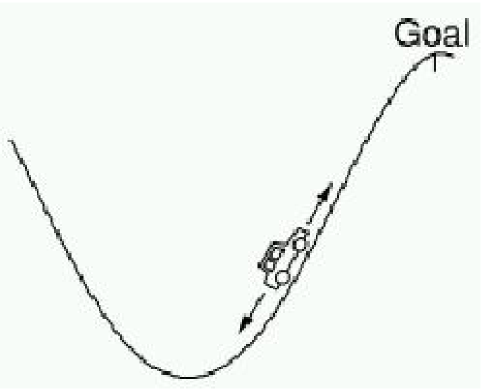
Evolve agents that are better able to **learn**

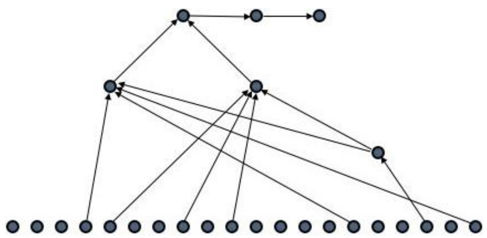
- Evolution chooses representation and initial weights
  - NEAT learns NN topologies [Stanley & Miikkulainen, '02]
- Q-learning learns weights that approximate value function



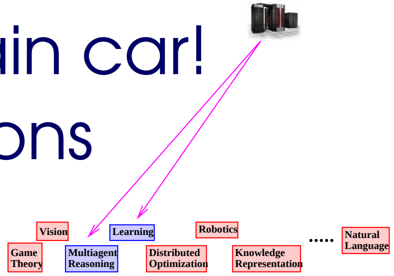
NEAT+Q Results





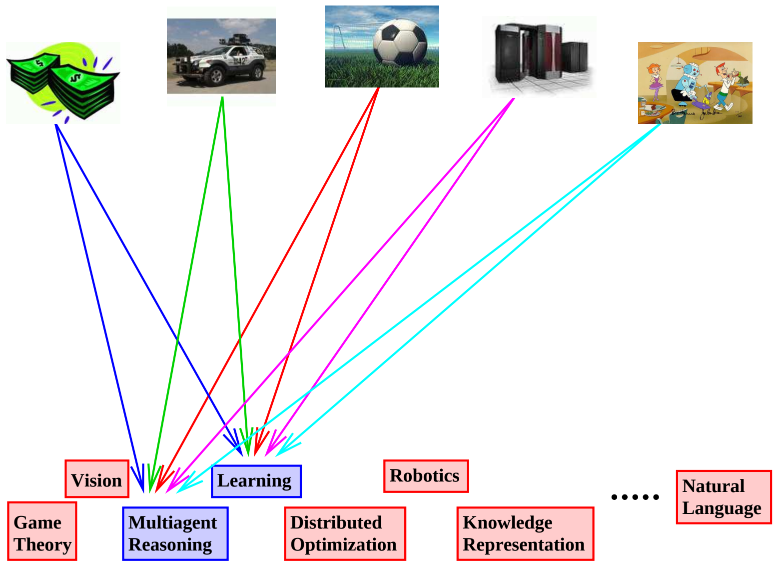


- Neural net function approx. works on mountain car!
- Tested Q-learning with 24 manual configurations



Outline

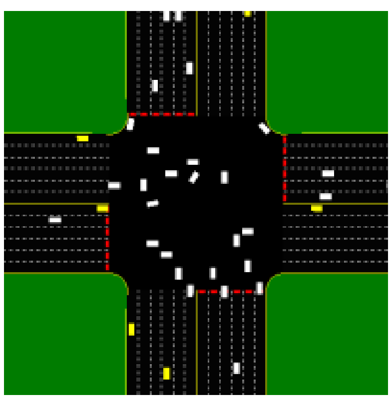
- Build complete solutions to relevant challenge tasks
  - Drives research on component algorithms, theory
- Learning Agents
  - Scaling up Reinforcement Learning
  - Adaptive representations
- **Multiagent reasoning**
  - Prepare for the unexpected
  - Adaptive interaction protocols

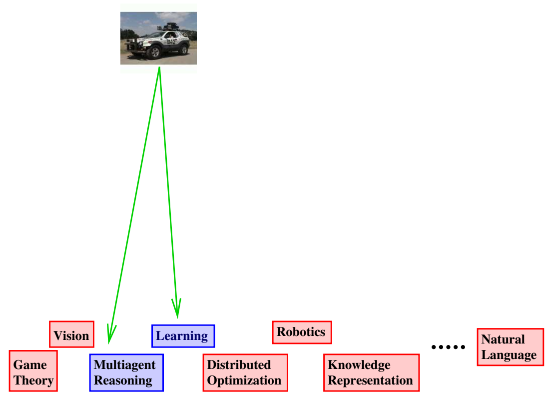


Multiagent Reasoning

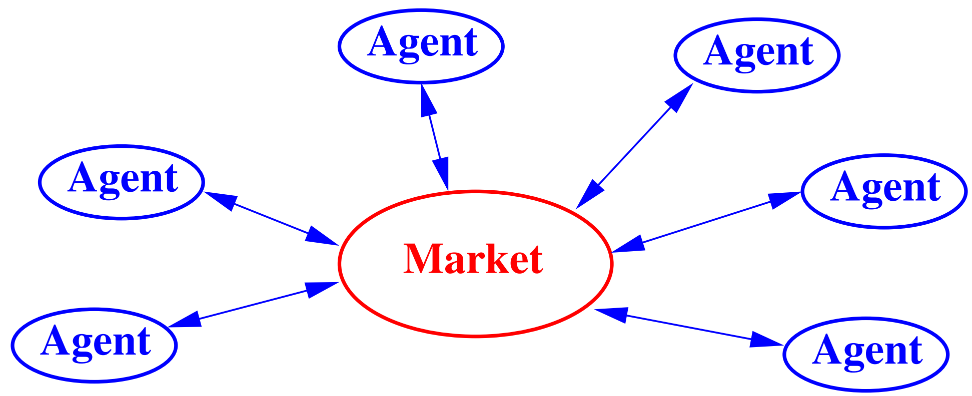
Robust, fully autonomous agents in the real world

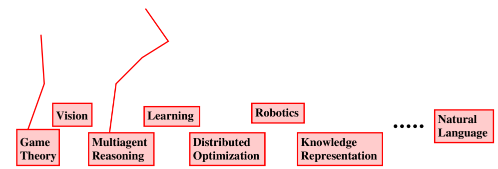
- Once there is one, there will soon be many
- To coexist, agents need to interact
- Example: **autonomous vehicles**
  - DARPA "Grand Challenge" was a great first step
  - Urban Challenge continues in the right direction
  - **Traffic lights and stop signs still best?** [Dresner & Stone, '04]





Autonomous Bidding Agents





- Usual assumption: rational agents

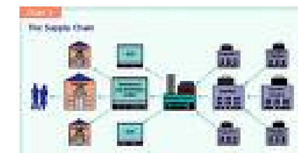
- In practice, must prepare for the unexpected
  - Other agents created by others
  - Teammate/opponent modeling
  - Especially in competition scenarios

Trading Agent Competitions

ATTac: champion travel agent [Stone et al., '02]

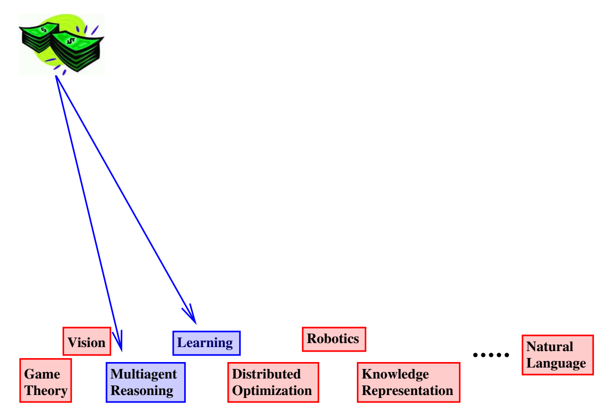
– Learns model of auction closing prices from past data
– Novel algorithm for conditional density estimation

TacTex: champion SCM agent [Pardoe & Stone, '06]



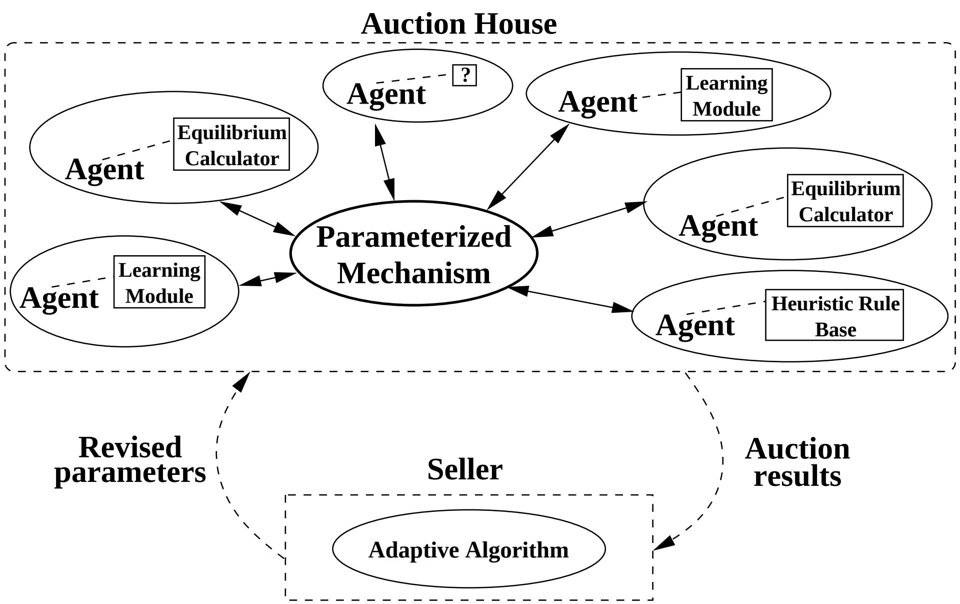
– Adapts procurement strategy based on recent data
– Predictive planning and scheduling algorithms

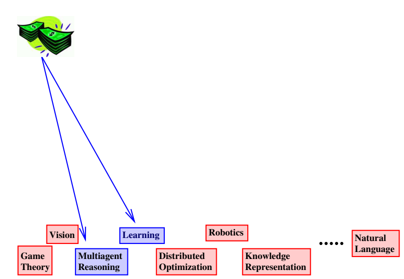
Common multiagent tradeoff:
learn detailed static model vs. adapt minimally on-line



Adaptive Mechanism Design

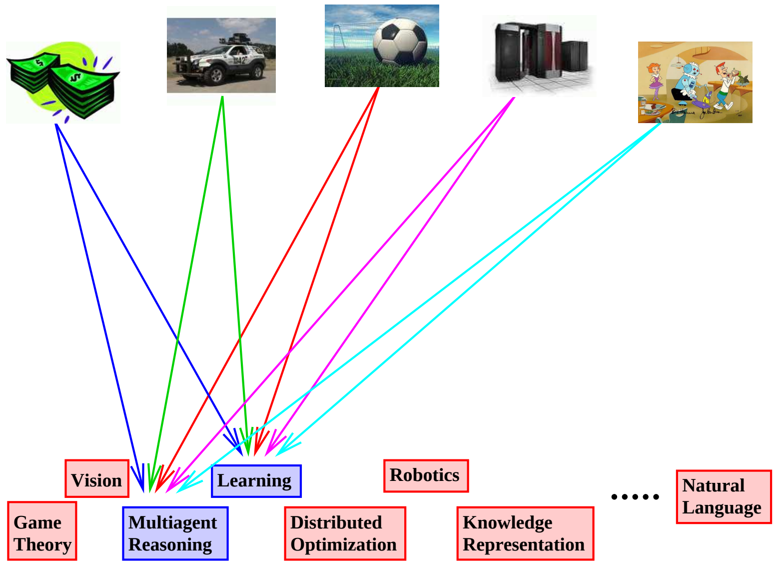
- Traditional mechanism design done manually
  - Protocols fixed and given
  - e.g. Telecom spectrum auctions
- Like RL representations, protocols can be the hard part!
  - Language learning [Steels '96; Jim & Giles, '00]
  - Automated mechanism design [Sandholm, '03]
  - Let the **mechanism adapt** itself: [Pardoe & Stone, '06]





Outline

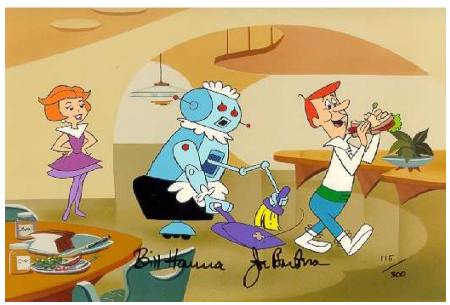
- Build complete solutions to relevant challenge tasks
  - Drives research on component algorithms, theory
- Learning Agents
  - Scaling up Reinforcement Learning
  - Adaptive representations
- Multiagent reasoning
  - Prepare for the unexpected
  - Adaptive interaction protocols
- **Implications**



A Goal of AI

Robust, fully autonomous agents in the real world

What happens **when** we achieve this goal

 ?

?

- Question: Would you rather live
  - 50 years ago? Or 50 years in the future?
- Not clear — world changing in many ways for the worse

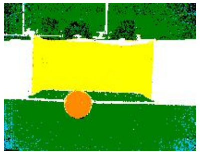
AI can be a part of the solution

Acknowledgments

- Manuela Veloso
- Collaborators on this research: János Csirik, Kurt Dresner, Peggy Fidelman, Shivaram Kalyanakrishnan, Nate Kohl, Gregory Kuhlmann, David McAllester, David Pardoe, Michael Littman, Patrick Riley, Richard Sutton, Robert Schapire, Mohan Sridharan, Manuela Veloso, Shimon Whiteson
- Help preparing the talk: Dana Ballard, Kristen Grauman, Nicholas Jong, Nate Kohl, Gregory Kuhlmann, Ben Kuipers, Vladimir Lifschitz, Risto Miikkulainen, Ray Mooney, Bruce Porter, Elaine Rich, Astro Teller, Shimon Whiteson
- Other contributors: Mazda Ahmadi, Bikramjit Banerjee, Daniel Stronger, Mathhew Taylor, Jonathan Wildstrom, Yaxin Liu
- Sponsors: NSF, DARPA, ONR, Sloan Foundation, IBM, NASA
- **AI researchers over the past 50 years**

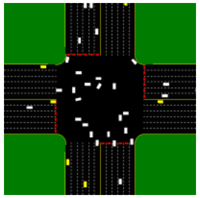
Additional Technical Details

- Learned robot vision [Sridharan & Stone] (Wed., 2:40pm)

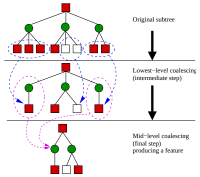


- Autonomic computing [Wildstrom & Stone] (Wed., 3:00pm)

- Intersection management [Dresner & Stone] (Thurs., 3:00pm)



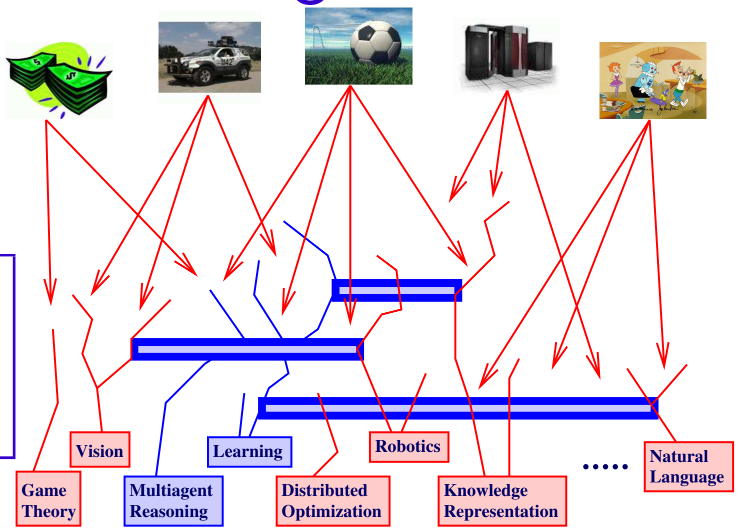
- Transfer learning [Banerjee & Stone] (Thurs., 5:00pm)



Summary

Robust, fully autonomous agents in the real world

- Build complete solutions to relevant challenge tasks
  - Good problems drive research



Combine algorithmic research, problem-oriented approaches

- Current challenges need learning, multiagent reasoning
  - Adaptive representations
  - Adaptive interaction protocols