

## Planning, Execution & Learning: Multi-Agent Planning I

Reid Simmons  
Manuela Veloso

Planning, Execution & Learning: Multi-Agent Planning I 1

Simmons & Veloso : Fall 2008

## Multi-Agent Domains

- **Cooperative**
  - Agents interact to maximize (expected) global reward
  - Action of agent *independent* of state of others
  - Reward for actions are independent of what others are doing:  
 $R(a_1, a_2) = R(a_1) + R(a_2)$
- **Coordinated**
  - Agents interact to maximize (expected) global reward
  - Action of agent *dependent* on state of others
  - Reward may depend on *joint* action:  $R(a_1, a_2) \neq R(a_1) + R(a_2)$ 
    - Correct coordination may lead to large reward
    - Miscoordination may lead to large penalty
- **Adversarial**
  - Agents interact to maximize their own reward
  - May try to *minimize* reward of the other agents
  - Rewards *interfere* with each other:  $R(a_1) \propto -R(a_2)$

Planning, Execution & Learning: Multi-Agent Planning I 2

Simmons & Veloso : Fall 2008

## Multi-Agent Planning

- **Cooperative Domains**
  - Focus on efficient and effective task allocation and dynamic re-allocation
  - *Behavior-based* approaches
  - *Market-based* approaches
- **Coordinated Domains**
  - Focus on performing joint actions
  - *MDP-based* approaches (both centralized and decentralized)
- **Adversarial Domains**
  - Focus on modeling and reasoning about opponent
  - *Mini-max* (game) search
  - *Game theoretic* approaches

Planning, Execution & Learning: Multi-Agent Planning I 3

Simmons & Veloso : Fall 2008

## Behavior-Based Approaches to Cooperation

- Agents compute actions independently
- Little, or no, communication between agents
- Dependent upon sensing other agents
- Cooperative behavior is *emergent property* of agents' actions
- + Simple tactics, minimal computation
- Overall behavior may be very suboptimal

Planning, Execution & Learning: Multi-Agent Planning I 4

Simmons & Veloso : Fall 2008

## Social Potentials (Balch)

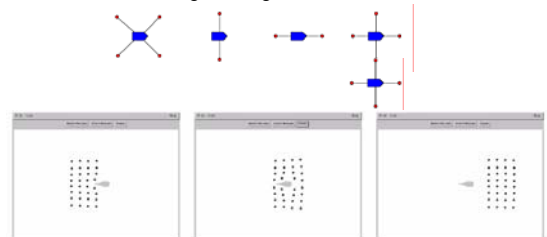
- Potential-Field Based Approach
  - Each behavior computes attractive or repulsive force
    - **avoid\_obstacle, move\_to\_goal**
  - Overall action vector sum of each behavior
    - Modulate with gain factors
- **Social Potentials**
  - Some behaviors based on knowledge of other agents
    - **avoid\_robots, maintain\_formation, move\_to\_unit\_center**

Planning, Execution & Learning: Multi-Agent Planning I 5

Simmons & Veloso : Fall 2008

## Social Potentials & Formation Keeping

- Each Agent has Multiple “Attachment Sites”
  - Inspired by crystal formation
  - Different configurations give rise to different formations



Planning, Execution & Learning: Multi-Agent Planning I 6

Simmons & Veloso : Fall 2008

### Market-Based Methods

- Addresses the problem of *Task Allocation*
- **Basic Idea:**
  - Agents bid on tasks (based on cost and rewards)
  - Auctioneer allocates tasks (based on increased utility)
- **Decentralized, Satisficing, Cooperation (mainly)**
  - Deals well with heterogeneous systems
- **Some recent work in coordination**

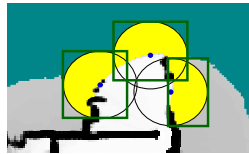
### Bidding for Information Gain

- Decentralized approach to exploration
- **Basic Idea:**
  - Individual robots construct “bids”
  - “Auctioneer” collects bids and assigns tasks
  - Robots periodically (and asynchronously) continue bidding
  - Tasks may be reassigned based on new bids



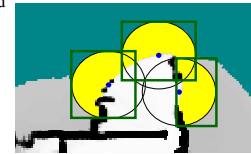
### Bidding for Information Gain II

- **Individual robots construct “bids”**
  - Bid is estimate of cost *and* information gain of visiting a location
  - Determine cost of optimal path to frontier cell
  - Estimate number of “unknown” cells within the “nominal” sensor range of frontier cell
  - Approximate region of information gain by bounding rectangle

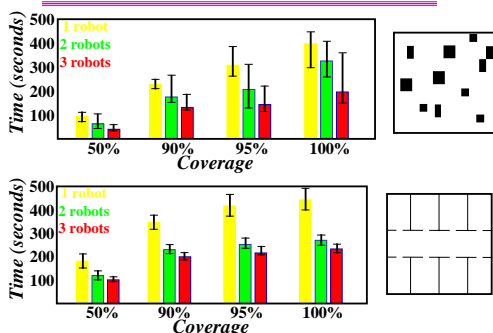


### Bidding for Information Gain III

- **“Auctioneer” collects bids and assigns tasks**
  - Try to maximize overall expected utility
  - Find maximum utility bid (info gain - cost)
  - Assign that task to robot
  - Discount other bids based on percent overlap with selected info gain region
  - Continue until no expected net gain



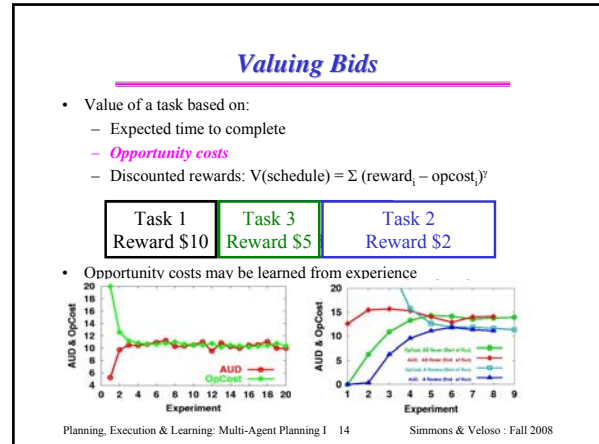
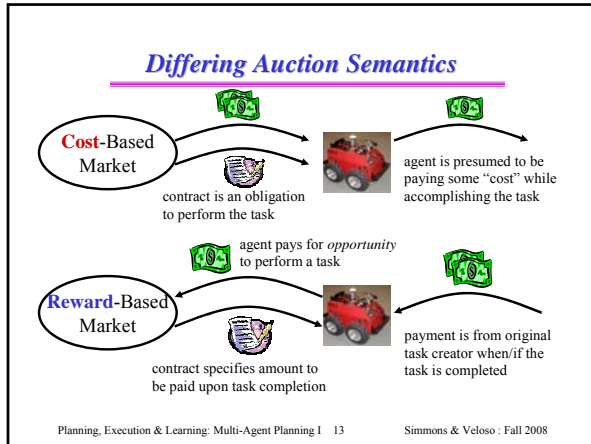
### Results



### Combinatorial Auctions for Exploration

- Auctioneer announces tasks for auction
- Agents (optionally) bid on tasks
  - May cluster tasks to decrease overall cost
- Auctioneer “clears” the auction
  - May use full combinatorial solution
  - May use heuristic / greedy solution
- Tasks are awarded if they increase the auctioneers utility
- Agents receiving tasks may achieve them, or rebid





- ### Learning When and What to Bid (Busquets)
- Observation:** There is a statistical likelihood that bids of a certain value will win, on average
  - Approach:** Learn the probability distribution by tracking which bids win and lose
    - For the **auctioneer**: Given the minimum bid price, determine whether to offer the task
    - For the **bidder**: Given the calculated bid, is it reasonable to send it
  - Result:** Decreases computation and communication nearly order of magnitude, with little or no affect on performance
    - Main win is with auctioneer's decision
- Planning, Execution & Learning: Multi-Agent Planning I 15      Simmons & Veloso : Fall 2008

- ### Learning When and What to Bid
- Bidder**
    - $n_{\text{win}}(b) = \# \text{ auctions where winning bid} \leq b$
    - $n_{\text{loss}}(b) = \# \text{ auctions where no winner and top bid} \geq b$
    - $p_{\text{win}}(b) = n_{\text{win}}(b) / (n_{\text{win}}(b) + n_{\text{loss}}(b))$
  - Auctioneer**
    - $n_{\text{award}}(b) = \# \text{ winning bids received} \geq b$
    - $n_{\text{no-award}}(b) = \# \text{ losing bids received} \geq b$
    - $n_{\text{auctions}} = \text{total} \# \text{ auctions}$
    - $p_{\text{award}}(b) = (n_{\text{award}}(b) + n_{\text{no-award}}(b)) / n_{\text{auctions}}$
  - Learn iteratively, since policy changing
- 
- Planning, Execution & Learning: Multi-Agent Planning I 16      Simmons & Veloso : Fall 2008

- ### Market-Based Coordination (Kalra)
- Key Idea:** Agents bid on *plans* rather than tasks
  - "Passive" Coordination**
    - Generate and evaluate based on plans of other agents
    - Broadcast best plan to other agents
    - Agents re-evaluate plans based on new information; iterate
  - "Active" Coordination**
    - Generate joint (team) plans
    - Request "price quotes" from other agents
    - Adopt if team profit minus costs better than individual profit
    - Obtain commitments from teammates
      - Pays "breach penalty" if commitment is broken
  - Switch to active coordination when constraint violations are likely
- Planning, Execution & Learning: Multi-Agent Planning I 17      Simmons & Veloso : Fall 2008

