Autonomous Multiagent Systems

Peter Stone

Week 12, Thursday (11/14)

Any Questions?

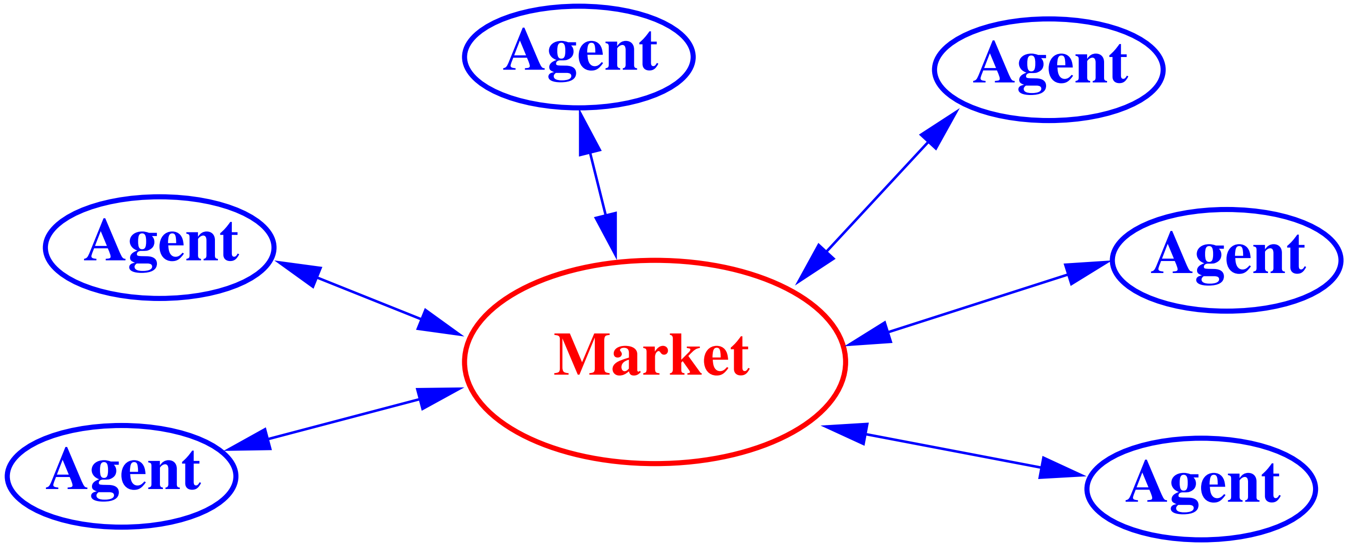
TAC

- The reading was an overview
- Here, some specific strategies
- 2 years: Details differed in first year
- Current, active research
- Lots of slides — may skip some
- (Should the class be run with this as the testbed?)

Autonomous Bidding Agents

eBay bidding agents

- Input max price, single good
- agent tops bids up to max price
- snipers bid at last minute: substitutable goods



Objective

Create **autonomous agents** to bid in auctions so as to maximize **profit**, given an **auction mechanism** and a set of **competing bidders**

Profit: Utility($winnings$) $-$ $expenditures$

Mechanism: Fixed and known

- (Simultaneous auctions)

Competing Bidders: Not necessarily fixed or known

- Beyond agent's control
- Possibly manipulable
- Off equilibrium

Toy Example

	utility
camera alone	\$50
flash alone	10
both	100
neither	0

- What's the value of the flash?
  - Auctions are simultaneous
  - Auctions are independent (no combinatorial bids)

Toy Example

	utility
camera alone	$50
flash alone	10
both	100
neither	0

- What's the value of the flash?
  - Auctions are simultaneous
  - Auctions are independent (no combinatorial bids)

- $\in [10, 50]$ — **Depends on the price of the camera**

Trading Agent Competition

- Put forth as a **benchmark problem** for e-marketplaces, simultaneous auctions [Wellman, Wurman, et al., 2000]

- Autonomous agents act as **travel agents**
  - **Game**: 8 *agents*, 15 min.
  - **Agent**: simulated travel agent with 8 *clients*
  - **Client**: TACtown $\leftrightarrow$ Boston within 5-day period

- **Auctions** for flights, hotels, entertainment tickets
  - **Server** maintains markets, sends prices to agents
  - Agent sends bids to server **over network**

28 Simultaneous Auctions

Flights: Inflight days 1-4, Outflight days 2-5 (8)

- Unlimited supply; prices random walk; immediate clear; no resale

28 Simultaneous Auctions

Flights: Inflight days 1-4, Outflight days 2-5 (8)

- Unlimited supply; prices random walk; immediate clear; no resale

Hotels: Grand Hotel/Le Fleabag Inn 1-4 (8)

- 16 rooms per auction; 16th-price ascending auction; quote is ask price; no resale
- Auctions can close early

28 Simultaneous Auctions

Flights: Inflight days 1-4, Outflight days 2-5 (8)

- Unlimited supply; prices random walk; immediate clear; no resale

Hotels: Grand Hotel/Le Fleabag Inn 1-4 (8)

- 16 rooms per auction; 16th-price ascending auction; quote is ask price; no resale
- Auctions can close early

Entertainment: Baseball/Symph/Theater days 1-4 (12)

- Continuous double auction; initial endowments; quote is bid-ask spread; resale allowed

Client Preferences and Utility

Preferences: randomly generated per client

- Ideal arrival, departure days
- Good Hotel Value
- Entertainment Values

Utility: 1000 (if valid) − travel penalty + hotel bonus
+ entertainment bonus

Score: Sum of client utilities − expenditures

Example Game

Client Preferences

Client	*IAD*	*IDD*	*GHV*	*BEV*	*SEV*	*TEV*
1	Day 2	Day 5	73	175	34	24
2	Day 1	Day 3	125	113	124	57
⋮	⋮	⋮	⋮	⋮	⋮	⋮
8	Day 1	Day 3	50	79	92	136

Allocation

Client	*AD*	*DD*	Hotel	Ent'ment	Utility
1	Day 2	Day 5	Bad	B4	1175
2	Day 1	Day 2	Good	B1	1138
⋮	⋮	⋮	⋮	⋮	⋮
8	Day 1	Day 2	Good	T1	1086

Outline

- **TAC 2000**
  - The Domain
  - **The Agent: ATTac-2000**
  - The Competition
  - Controlled Experiments
- TAC 2001

Department of Computer Sciences
The University of Texas at Austin

High-Level Strategy

1. **While auctions open:**
   - Obtain updated market prices
   - Compute G^*: most profitable allocation of goods given holdings and prices
   - Bid (2 different modes)

     Passive: bid to keep options open

     Active: at end, bid aggressively

2. **Allocate:**
   - Compute G^* with closed auctions

Allocation

Given holdings, prices, determine G^*:
Optimal complete itinerary assignments

- Mixed-integer LP with 3 constraints:
  1. Purchase enough to meet demand
  2. Entertainment tickets must be used legally
  3. All variables integral
- Enforce 1 and 2 (admissible);
  "Branch and bound" over adjustments for 3
- Globally optimal solution; *usually* $< .01$ sec

Department of Computer Sciences
The University of Texas at Austin

Flights

- Expected change in flight auctions is \$0
- Flights are expensive
- More information about hotel markets over time

$\boxed{\textbf{Delay flight commitments}}$

(Provably optimal in isolation)

Passive: No bidding on flights

Active: Bid high for all flights needed by G^*

Hotels

Passive: Keep options open

- *Low* prices: bid low for several rooms (win)
- *Middle* prices: bid low for 1 room (keep open)
- *High* prices: don't bid

Active: Bid marginal utility in G^*

- Since flights bought, usually $> \$1000$

Department of Computer Sciences
The University of Texas at Austin

Toy Example

	utility
camera alone	\$50
flash alone	10
both	100
neither	0

- Let current camera price = \$80
  - $score(G_{\mathrm{f}}^{*}) = \max\{100 - 80, 10 - 0\} = 20$

Toy Example

	utility
camera alone	\$50
flash alone	10
both	100
neither	0

- Let current camera price = \$80
  - $score(G^*_{\text{f}}) = \mathsf{max}\{100 - 80, 10 - 0\} = 20$
  - $score(G^*_{\text{no-f}}) = \mathsf{max}\{50 - 80, 0 - 0\} = 0$
  - So $\text{value}(\text{flash}) = 20 - 0 = \$20$

Toy Example

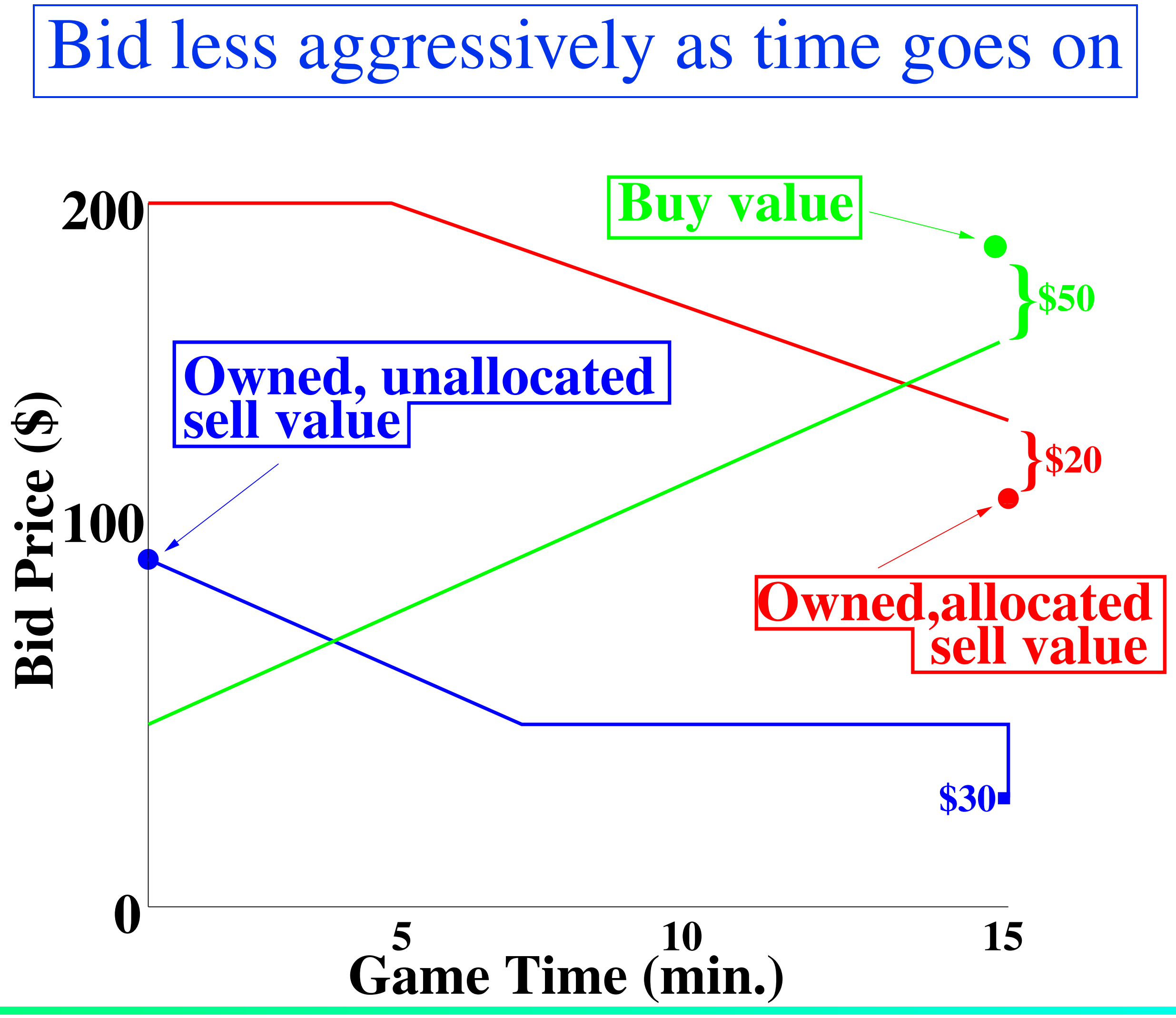
	utility
camera alone	\$50
flash alone	10
both	100
neither	0

- Let current camera price = \$80
  - $score(G^*_{\mathrm{f}}) = \max\{100 - 80, 10 - 0\} = 20$
  - $score(G^*_{\text{no-f}}) = \max\{50 - 80, 0 - 0\} = 0$
  - So $\text{value}(\text{flash}) = 20 - 0 = \$20$
- Already bought camera $\Rightarrow$ price = \$0 $\Rightarrow$ $\text{value}(\text{flash}) = 100 - 50 = \$50$

Department of Computer Sciences
The University of Texas at Austin

Entertainment

- Offer to buy *and* sell each ticket type ($0–$200)



Adaptability

Allocation: LP sometimes too slow

- Fallback: heuristic allocation (greedy)

Bidding mode: Want 2 iterations in active mode

- Unpredictable network/server lag
- Active too late $\Rightarrow$ miss flights and/or hotels
- Active too soon $\Rightarrow$ early commitment
- Solution: monitor loop times

Hotel bids: Avoid skyrocketing rooms

Department of Computer Sciences
The University of Texas at Austin

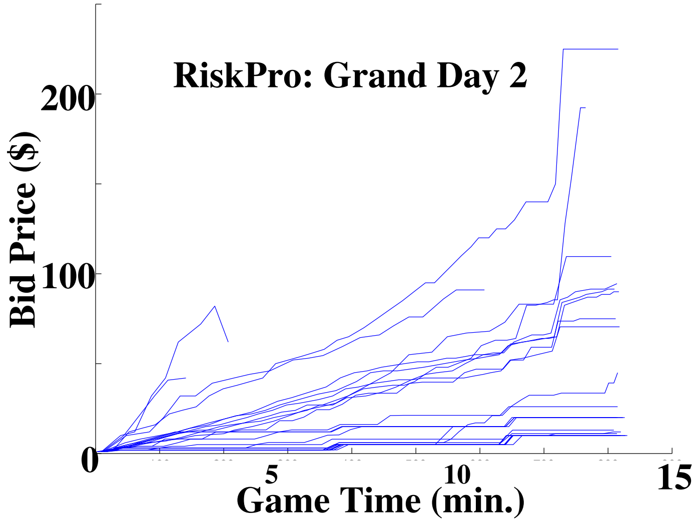
Outline

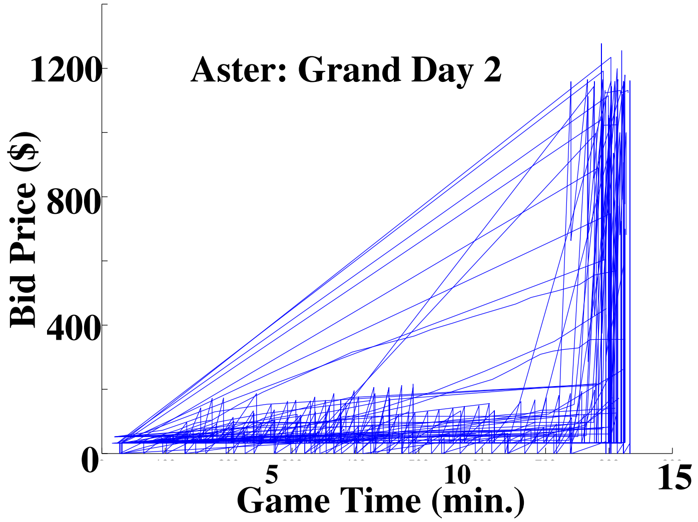
- **TAC 2000**
  - The Domain
  - The Agent: ATTac-2000
  - **The Competition**
  - Controlled Experiments
- TAC 2001

Department of Computer Sciences
The University of Texas at Austin

Hotel Escalation

- Preliminaries: 3 of 22 (14%) agents bid high
- After: 7 of 12 (58%) start bidding high





- Games with several high-bidders **disastrous**
  - Hotel prices skyrocket to $> \$1000$
  - **No-win situation**

Department of Computer Sciences
The University of Texas at Austin

Adaptability (cont.)

Hotel Bidding: Avoid skyrocketing rooms

- Will they? Which ones?
- Determine opponents: # high bidders
- Group hotels by symmetry (day 1=4, 2=3)
- Keep track of average closing prices
- When *price* $<$ *average*, use average
- Empirical tests: averts catastrophe, little cost

Department of Computer Sciences
The University of Texas at Austin

Toy Example

	utility
camera alone	\$50
flash alone	10
both	100
neither	0

- Let current camera price = \$20, flash = \$10
  - value(flash) would be $80 - 30 = \$50$
  - value(camera) would be $90 - 0 = \$90$
- But what if prices jump at the end?

Toy Example

	utility
camera alone	\$50
flash alone	10
both	100
neither	0

- Let current camera price = \$20, flash = \$10
  - value(flash) would be $80 - 30 = \$50$
  - value(camera) would be $90 - 0 = \$90$
- But what if prices jump at the end?
  - Let average past camera price = \$80, flash = \$30
  - value(flash) $= \$20$
  - value(camera) $= \$70$

Semi-Finals and Finals

Rank	Team	Avg. Score	Institution
1	ATTac	3398	AT&T
2	RoxyBot	3283	NASA/Brown
3	aster	3068	Intertrust
4	umbctac1	3051	UMBC
5	ALTA	2198	A-Life (Russia)
6	m_rajatish	1873	U. of Tulsa
7	RiskPro	1570	Stockholm (Sweden)
8	T1	1167	Industilogik (Sweden)

- We're consistent – no bad games

Controlled Experiments

ATTac vs. non-adaptive high and low bidders

#high	agent 2	agent 3	agent 4	agent 5	agent 6	agent 7	agent 8
7	⟵	9526	⟶	⟶	⟶	⟶	⟶
6	⟵	10679	⟶	⟶	⟶	⟶	1389
5	⟵	10310	⟶	⟶	⟶	⟵	2650
4	⟵	10005	⟶	⟶	⟵	⟵	4015
3	⟵	5067	⟶	⟵	⟵	⟵	3639
2	⟵	*209*	⟵	⟵	⟵	⟵	2710

- Shows ATTac's average score difference
- ATTac adapts over successive runs
- All numbers positive, most are significant

Lessons Learned

$\boxed{\textbf{The competing agents matter}}$

- Flexibility to different scenarios important
- ATTac not stable against itself – "parasite"
- Allocation problem solvable

$\boxed{\text{To come: }\textbf{The mechanism matters}}$

Outline

- TAC 2000
- **TAC 2001**
  - **Rule changes**
  - Agent changes
  - The competition
  - Controlled Experiments

Department of Computer Sciences
The University of Texas at Austin

Slight Rule Changes

Flights: prices tend to increase

- Reason to bid early

Hotels: last 8 minutes: random one closes, new quotes

- Must bid early
- Full minute to make decision

Entertainment: initial endowments skewed

- 4 of 2; 2 of 2; 0 of other 8

New Approach

- Start with optimal allocator
  - No longer an advantage
  - Use as subroutine over many possible futures
- Learn model of expected price **distributions**
  - New boosting-based conditional density estimation algorithm for price prediction [ICML, 2002]
- Decision theoretic reasoning over predictions [AMEC, 2002]

High-Level Strategy

- For each auction:
  - Repeatedly sample price vector from distributions
  - Bid avg marginal expected utility: $v(G_w^*) - v(G_l^*)$
- Bid for all goods — not just those in G^*

$\boxed{\textbf{Goal: analytically calculate optimal bids}}$

Department of Computer Sciences
The University of Texas at Austin

Toy Example

	utility
camera alone	\$50
flash alone	10
both	100
neither	0

- What's the value of the flash?
  - Camera price = \$70 $\Rightarrow$ value(flash) = \$30
  - Camera price = \$20 $\Rightarrow$ value(flash) = \$50
  - Camera price = \$40 $\Rightarrow$ value(flash) = \$50

- Expected value: resample camera price, take avg.

Assumptions

1. Closing prices somewhat predictable
2. Own bids don't affect final prices
3. Can change bids in open auctions
4. Future decisions made with complete information

Hotel Price Prediction

- Need hotel price predictions at various points in game
- Use machine learning methods applied to data from previously played games
- Want to estimate:
  - expected price
  - distribution of prices given current conditions (permits sampling of prices)

Setting up the Learning Problem

- **Features:**
  - Current hotel and flight prices
  - Current time in game
  - Hotel closing times
  - Agents in the game (when known)
  - Variations of the above

- Built **8 separate predictors**:
  - Good Hotel vs. Bad (2)
  - Day 2,3 vs. day 1,4 — exploit symmetry (2)
  - 1st min. (pre-hotel-quotes) vs. later (2)

- **Data**: seeding round games
  - Features $\mapsto$ actual clear prices

Department of Computer Sciences
The University of Texas at Austin

The Learning Algorithm

- X = feature vector
- Given X, estimate distribution of Y = (hotel closing price) − (current bid)
- Break Y into $k \approx 50$ cut points $b_1 \leq \cdots \leq b_k$
- For each b_i, estimate probability $Y \geq b_i$, given X
  - Can convert to estimated distribution of $Y|X$
- Say X belongs to class C_i if $Y \geq b_i$
  - k-class problem: each example in many classes
- Use **BoosTexter** (boosting "decision stumps")
  - Can convert output to class prob. estimates

Department of Computer Sciences
The University of Texas at Austin

The Learning Algorithm (cont.)

- Use logistic regression form of boosting
- Given X, get estimates p_j of $\Pr[Y \geq b_j]$
- So can estimate $\Pr[b_j \leq Y < b_{j+1}]$ by $Q(j) = p_j - p_{j+1}$
- Gives distribution Q over bins $[b_j, b_{j+1})$
- To sample:
  - Choose one bin at random according to Q
  - Choose point within bin uniformly at random
- Piecewise constant density function

Department of Computer Sciences
The University of Texas at Austin

Hotel Expected Values

- Repeat until time bound, for each hotel:
  1. Assume this hotel closes next
  2. Sample prices from predicted price distributions
  3. Given these prices compute $V_0, V_1, \ldots V_8$
     - $V_i = v(G^*)$ if own **exactly** i of the hotel
     - $V_0 \leq V_1 \leq \ldots \leq V_8$
- Value of ith copy is avg($V_i - V_{i-1}$)

Other Uses of Sampling

Flights: Cost/benefit analysis for postponing commitment

Cost: Price expected to rise over next n minutes

Benefit: More price info becomes known

- Compute expected marginal value of buying some different flight

Entertainment: Bid more (ask less) than expected value of having one more (fewer) ticket

Outline

- TAC 2000
- **TAC 2001**
  - Rule changes
  - Agent changes
  - **The competition**
  - Controlled Experiments

Department of Computer Sciences
The University of Texas at Austin

Finals

Team	Avg.	Adj.	Institution
ATTac	3622	4154	AT&T
livingagents	3670	4094	Living Systems (Germ.)
whitebear	3513	3931	Cornell
Urlaub01	3421	3909	Penn State
Retsina	3352	3812	CMU
CaiserSose	3074	3766	Essex (UK)
Southampton	3253*	3679	Southampton (UK)
TacsMan	2859	3338	Stanford

- ATTac improves over time
- livingagents is an open-loop strategy

Controlled Experiments

- *ATTac*$_s$: "“full-strength” agent based on boosting
- *SimpleMean*$_s$: sample from empirical distribution
  (previously played games)
- *ConditionalMean*$_s$: condition on closing time
- *ATTac*$_{ns}$, *ConditionalMean*$_{ns}$, *SimpleMean*$_{ns}$:
  predict expected value of the distribution
- *CurrentPrice*: predict no change
- *EarlyBidder*: motivated by TAC-01 entry livingagents
  - Immediately bids high for G^* (with *SimpleMean*$_{ns}$)
  - Goes to sleep

Stability

- 7 *EarlyBidder*'s with 1 *ATTac*

Agent	Score	Utility
ATTac	2431 ± 464	8909 ± 264
EarlyBidder	-4880 ± 337	9870 ± 34

- 7 *ATTac*'s with 1 *EarlyBidder*

Agent	Score	Utility
ATTac	2578 ± 25	9650 ± 21
EarlyBidder	2869 ± 69	10079 ± 55

EarlyBidder gets more utility; *ATTac* pays less

Results

- *Phase I*: Training from TAC-01 (seeding round, finals)
- *Phase II*: Training from TAC-01, phases I, II
- *Phase III*: Training from phases I – III

Agent	*Relative Score*	
	Phase I	*Phase III*
$ATTac_{ns}$	105.2 ± 49.5 (2)	166.2 ± 20.8 (1)
$ATTac_s$	27.8 ± 42.1 (3)	122.3 ± 19.4 (2)
EarlyBidder	140.3 ± 38.6 (1)	117.0 ± 18.0 (3)
$SimpleMean_{ns}$	-28.8 ± 45.1 (5)	-11.5 ± 21.7 (4)
$SimpleMean_s$	-72.0 ± 47.5 (7)	-44.1 ± 18.2 (5)
$ConditionalMean_{ns}$	8.6 ± 41.2 (4)	-60.1 ± 19.7 (6)
$ConditionalMean_s$	-147.5 ± 35.6 (8)	-91.1 ± 17.6 (7)
CurrentPrice	-33.7 ± 52.4 (6)	-198.8 ± 26.0 (8)

TAC-02

- Repeat of the TAC-01 rules
- Several new and returning agents
- Seeding round: 440 games
  - ATTac ran with the learned models from TAC-01
  - **ATTac won the seeding round** (barely)
- Semifinals (14 games)... not so good.

Offshoots

Practical application: FCC spectrum auction 35

- 422 licenses; 80+ bidders; ≈$8 billion spent; ran Dec 12 – Jan 26, 2001
- Randomized SDR [Reitsma et al., 2001]

Theoretical abstraction: iterated bimatrix games

- Issuing, responding to threats [Littman, Stone, 2001]

Future Work — Generalize Approaches

TAC-03:

- Artificial supply chain scenario:
  - Bid to acquire raw materials
  - Manage assemply process
  - Respond to customer orders

Penn Automated Trading Project:

- Simulation based on real-time, real market data
- Uses the full order book from ECNs
- No distinction between real orders and agent orders

FCC Auction Simulator (FAucS):

- Learning into the FCC agents
- Represent, reason about realistic human preferences

Last-minute bidding [Roth,Ockenfels, 2001]

– eBay: first-price, ascending auction
– Amazon: auction extended if bid in last 10 minutes
– eBay: bots exist to incrementally raise your bid to a maximum

- Still people *snipe*. Why?
  – There's a risk that the bid might not make it
  – However, common-value $\Longrightarrow$ bid conveys info
  – Late-bidding can be seen as implicit collusion
  – Or . . . , lazy, unaware, etc. (Amazon and eBay)

- Finding: more late-bidding on eBay,
  – even more on antiques rather than computers

Small design-difference matters

Department of Computer Sciences
The University of Texas at Austin

Late Bidding as Best Response

- Good vs. incremental bidders
  - They start bidding low, plan to respond
  - Doesn't give them time to respond

- Good vs. other snipers
  - Implicit collusion
  - Both bid low, chance that one bid doesn't get in

- Good in common-value case
  - protects information

Overall, the analysis of multiple bids supports the hypothesis that last-minute bidding arises at least in part as a response by sophisticated bidders to unsophisticated incremental bidding.

Next Week's Readings – Agent Modeling

- Recursive Modeling
- Team modeling in air combat
- Strategy assessment in robotic soccer