Mechanisms for Dynamic Environments

David C. Parkes
Division of Engineering and Applied Sciences
Harvard University
http://www.eecs.harvard.edu/econcs

Carnegie Mellon University Nov 6, 2005

Outline

- Prior-Free Online Auction Design:
  - Non-reusable Goods, Finite time horizon.
- General characterization for truthful online auctions
- Prior-Free Online Auction Design:
  - Reusable Goods, infinite time horizon.
- Model-based Online Mechanisms.
- Future Directions.

Related Papers

- Virtual Worlds: Fast and Strategyproof Auctions for Dynamic Resource Allocation. Chaki Ng, David C. Parkes, and Margo Seltzer. In *Proc. 4th ACM Conf. on Electronic Commerce* (EC'03), pp. 238-239, 2003 (Short paper).
- Adaptive Limited-Supply Online Auctions, Mohammad T. Hajiaghayi, Robert Kleinberg, and David C. Parkes. In *Proc. ACM Conf. on Electronic Commerce*, pp. 71-80, 2004.
- Online Auctions with Re-usable Goods, Mohammad Hajiaghayi, Robert Kleinberg, Mohammad Mahdian, and David C. Parkes. In *Proc. 6th ACM Conf. on Electronic Commerce* (EC'05), pp. 165-174, 2005.
- Models for Truthful Online Double Auction. Jonathan Bredin and David C. Parkes. In *Proc. 21st Conference on Uncertainty in Artificial Intelligence* (UAI'2005), pp. 50-59, 2005.
- Pricing WiFi at Starbucks– Issues in Online Mechanism Design, Eric Friedman and David C. Parkes. In *Proc. 4th ACM Conf. on Electronic Commerce* (EC'03), pp. 240-241, 2003. (Short paper).
- An MDP-Based Approach to Online Mechanism Design, David C. Parkes and Satinder Singh. In *Proc. 17th Annual Conf. on Neural Information Processing Systems* (NIPS'03), 2003.
- Approximately Efficient Online Mechanism Design, David C. Parkes, Satinder Singh, and Dimah Yanovsky. In *Proc. 18th Annual Conf. on Neural Information Processing Systems* (NIPS'04), 2004.

Example 1: Last-Minute Tickets

Value	\$100	\$80	\$60
Arrival:	11am	11am	12pm
Patience:	2hrs	2hrs	1hr

"Please bid your value and your patience. A decision will be made by the end of your stated patience."

How should you bid?

Value $100 $80 $60
Arrival: 11am 11am 12pm
Patience: 2hrs 2hrs 1hr

Auction: sell one ticket in each hour (given demand), to the highest bidder at second-highest bid price.

If truthful, then:
{ <1, $80>, <2, $60>}
However, bidder 1 could
a) reduce bid price to $65
{<2, $65>, <1, $60>}
b) delay bid until 12pm
{<2, $0>, <1, $60>}

Dynamic allocation problems

…are everywhere in computer science

- MoteLab (Berkeley)
  - distributed sensor network testbed
  - researchers compete for the right to sense, aggregate and propagate readings
- PlanetLab (Princeton)
  - global overlay network on the Internet
  - supports network research, long-running services
- Grid computing
  - much of science research is now intensively computational
  - globally-distributed computational infrastructure
- Network resource allocation
  - e.g. dynamic negotiation for WiFi bandwidth

Many systems are simultaneously both computational and economic systems.

…are can be found in e-commerce, elsewhere

- Sequential auctions on eBay
  - e.g. auctions for LCDs, each bidder wants one
- Expiring goods
  - e.g. auctions for last-minute tickets

Basic Model for Online Auctions

- Valuation $\theta_i = (a_i, d_i, w_i)$. Discrete time periods.
- Arrival time: a_i. Departure time: d_i .Value, w_i
- Allocation schedule $x \in X$.
- $v_i(x) = w_i$, if $x_i(t)=1$ for some $t \in [a_i,d_i]$
  $= 0$, otherwise
- Quasi-linear utility: $u_i(x,\text{price}) = v_i(x) - \text{price}$
- Auction: $A=\langle f, p \rangle$,
  - allocation rule, $f : \Theta^n \rightarrow X$
  - payment rule, $p : \Theta^n \rightarrow R^n$
- Truthful auction: reporting value $\langle a_i, d_i, w_i \rangle$ immediately upon arrival is a dominant strategy equilibrium.
- Assume: cannot under-report a_i.

Prior-Free: Key Variations

- Limited-supply ($k \geq 1$) of goods, sell in any period before time horizon, T.
  - single-unit demand
  - multi-unit demand
- Reusable goods, can sell up to k units in each time period. Finite time horizon, T.
  - single-period demand
  - multi-period demand

Prior-Free Auction Design

(c.f. Goldberg, Hartline et al.01)

$v^{(m)}$ is m-th highest value

$EFF(v) = \sum_{i \leq k} v^{(i)}$ "efficiency"

$\mathcal{F}^{(2)}(v) = \max_{2 \leq l \leq k} \{ l \cdot v^{(l)} \}$ "omniscient revenue"

(or Vickrey price if 1 item)

Value	\$500	\$80	\$60
Arrival:	11am	11am	12pm
Patience:	2hrs	2hrs	1hr

EFF:	\$580
OPT:	\$160

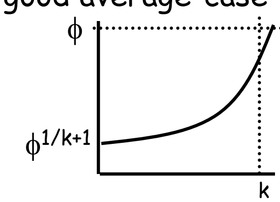
- c-competitive for efficiency if $E[Val(Auc_v)] \geq 1/c \; EFF(v)$, for all v
- c-competitive for revenue if $E[Rev(Auc_v)] \geq 1/c \; \mathcal{F}^{(2)}(v)$, for all v

Limited-Supply Auction

(Lavi & Nisan'00)

- Assume values in [L,U]. k-unit supply. Let $\phi = (U/L)$.
- <u>Adversarial model</u>: choose values and timing.
- Define a "price schedule": $p(j) = L \cdot \phi^{j/k+1}$, for j^{th} unit.
- Sell units while bid value $\geq$ price.

Truthful.
$\ln(\phi)$-competitive w.r.t. efficiency and Vickrey revenue,
Matching lower-bound, and good average-case performance in simulation.



Our model: Fixed, Unknown Distribution

(Hajiaghayi, Kleinberg, P., ACM'EC04)

- More realistic adversarial model: Lavi & Nisan allowed arbitrary sequencing of arbitrary values
- Instead, we model values as i.i.d. from some unknown distribution.
- Want good performance <u>whatever</u> the distribution.
- Should lead to an auction with better performance in practice.

Aside: The Online Selection Problem

- Remove incentives, and specialize to the case of disjoint arrival-departure intervals.

5 | 2 | 7 | 1,000 | 3 | . . .

Aside: The Online Selection Problem

- Remove incentives, and specialize to the case of disjoint arrival-departure intervals.
- Reduces to the secretary problem:
  - interview n job applicants in random order, want to max prob of selecting best applicant (told n)
  - told *relative ordering* w.r.t. applicants already interviewed, must hire or pass

5 | 2 | 7 | 1,000 | 3 | . . .

Aside: The Online Selection Problem

- Remove incentives, and specialize to the case of disjoint arrival-departure intervals.
- Reduces to the secretary problem:
  - interview n job applicants in random order, want to max prob of selecting best applicant (told n)
  - told *relative ordering* w.r.t. applicants already interviewed, must hire or pass

useful info: 1 | 2 | 1 | 1 | 4

5 | 2 | 7 | 1,000 | 3 | . . .

- Samples 1…n
- Candidate: a sample that is max across seen so far
- Want to accept a candidate when Prob(winner | candidate) $\geq$ Prob(find winner in future with optimal policy)

Prob(winner | candidate): increases; Prob(find winner in future with optimal policy): decreases

So, unique round in which start accepting.

E.g., n=1, s*=1, Pr(succ)=1
n=5, s*=3, Pr(succ)=0.433
n=10, s*=4, Pr(succ)=0.399
n=20, s*=8, Pr(succ)=0.384
n=100, s*=38, Pr(succ)=0.371
n=1000, s*=369, Pr(succ)=0.368 $\approx$ 1/e

The Secretary Algorithm

- **Theorem** (Dynkin, 1962): The following stopping rule picks the maximum element with probability approaching $1/e$ as $n\to\infty$.
  - Observe the first $\lfloor n/e \rfloor$ elements. Set a threshold equal to the maximum quality seen so far.
  - Stop the next time this threshold is exceeded.
- Asymptotic success probability of $1/e$ is best possible, *even if the numerical values of elements are revealed.*
  - i.e. optimal competitive ratio in the large n limit

Straw model for an Auction

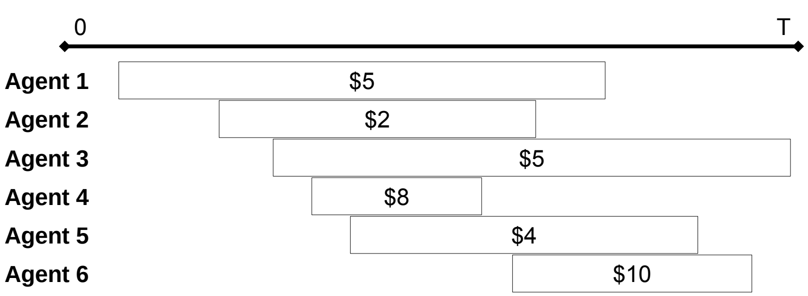
- **Auction**: $p(t)=\infty$, then set $p(t\geq\tau)=\max_{i\leq j} w_i$ after $j=\lfloor n/e \rfloor$ bids received. Sell to first subsequent bid with $w_i \geq p(t)$, then set $p(t)=\infty$.
- **Not truthful**: Bidders that span transition, and with high enough values, should delay arrival.

Truthful Auction:

- At time τ (for n/e arrival) let $p\geq q$ be the top two bids yet received.
- If any agent bidding p has not yet departed, sell to that agent (breaking ties randomly) at price q.
- Else, sell to the next agent whose bid is at least p (breaking ties randomly)

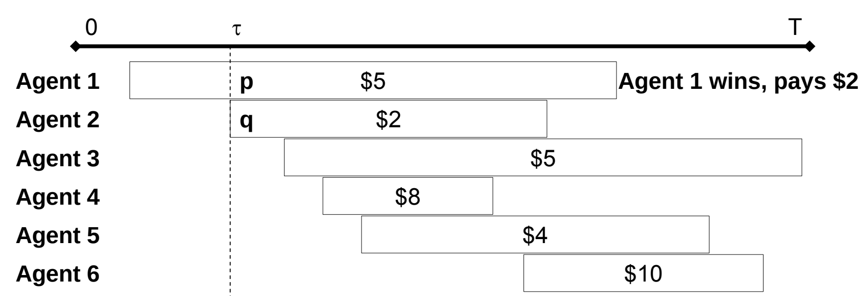
Adaptive Limited-Supply Auction

- At time τ, denoting arrival $j=\lfloor n/e \rfloor$, let $p\geq q$ be the top two bids yet received.
- If any agent bidding p has not yet departed, sell to that agent (breaking ties randomly) at price q.
- Else, sell to the next agent whose bid is at least p.



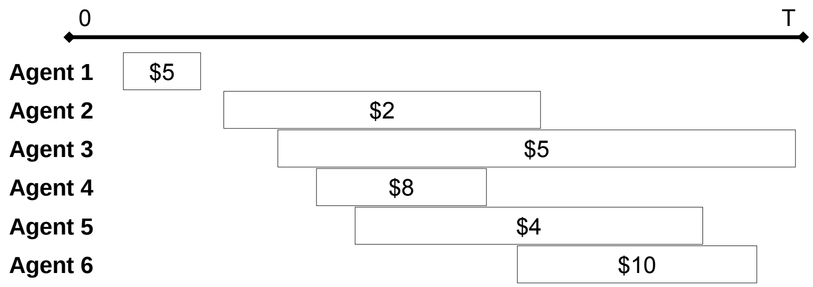
Adaptive Limited-Supply Auction

- At time τ, denoting arrival $j=\lfloor n/e \rfloor$, let $p\geq q$ be the top two bids yet received.
- If any agent bidding p has not yet departed, sell to that agent (breaking ties randomly) at price q.
- Else, sell to the next agent whose bid is at least p.



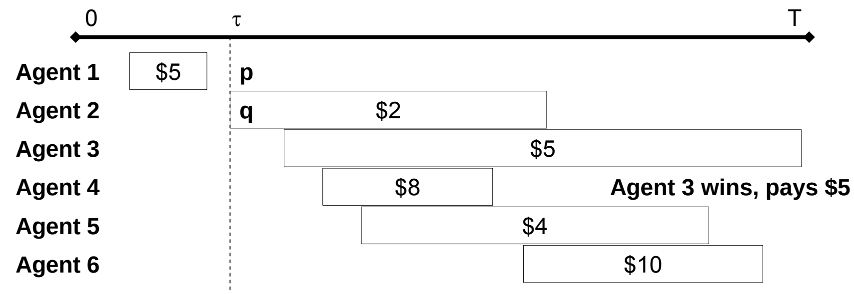
Adaptive Limited-Supply Auction

- At time τ, denoting arrival $j=\lfloor n/e \rfloor$, let $p \geq q$ be the top two bids yet received.
- If any agent bidding p has not yet departed, sell to that agent (breaking ties randomly) at price q.
- Else, sell to the next agent whose bid is at least p.



Adaptive Limited-Supply Auction

- At time τ, denoting arrival $j=\lfloor n/e \rfloor$, let $p \geq q$ be the top two bids yet received.
- If any agent bidding p has not yet departed, sell to that agent (breaking ties randomly) at price q.
- Else, sell to the next agent whose bid is at least p.



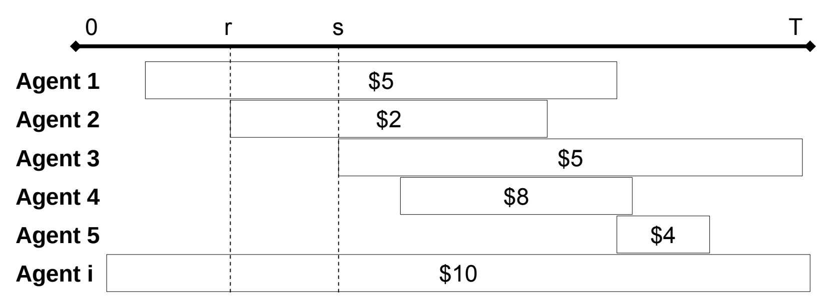
Analysis: Truthfulness

- If agent i wins, the price charged to her does not depend on her reported valuation.
- Pr(agent i wins) is (weakly) increasing in w_i, hence no incentive to understate w_i.
- Reporting $w'_i > w_i$ cannot increase the probability that agent i wins at a price $\leq w_i$, hence no incentive to overstate w'_i.
- Price facing agent i is never influenced by d_i, so no incentive to misstate d_i

… just need to check effect of arrival time.

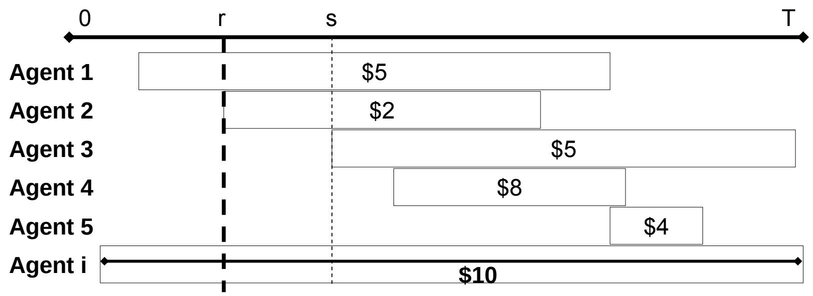
Analysis: Truthfulness

- **Claim**: Given two arrival times $a_i < a'_i$, it's always better to report a_i if possible.
- Let r,s be the $(\lfloor n/e \rfloor-1)$-th and $\lfloor n/e \rfloor$-th arrival times excluding agent i.



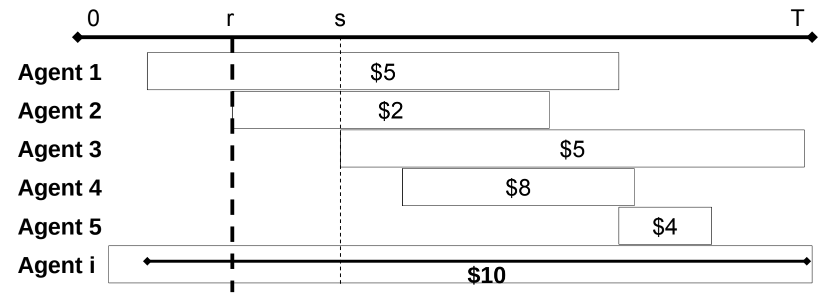
Analysis: Truthfulness

- Stating true arrival, agent 2 defines transition. Offered price \$5 on transition.



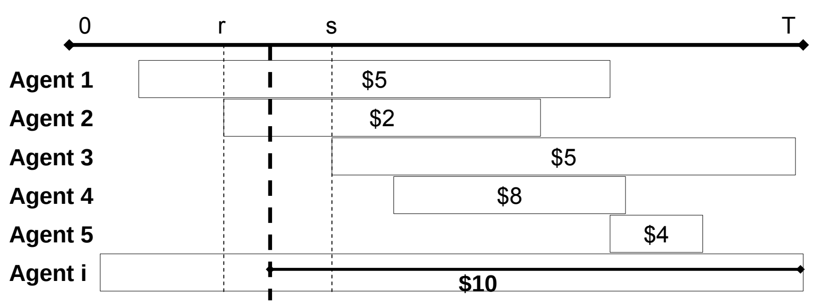
Analysis: Truthfulness

- Stating arrival time in $(a_i, r]$ changes nothing. Offered price \$5 on transition.



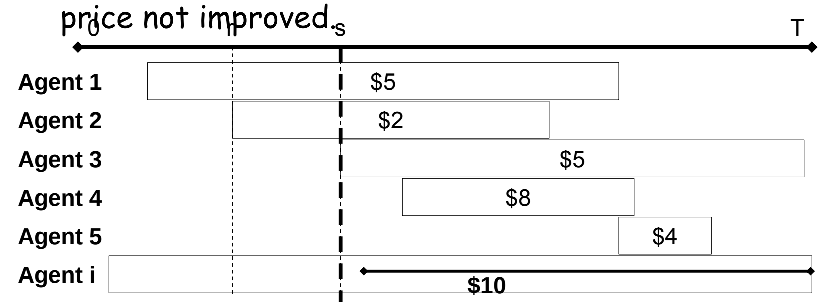
Analysis: Truthfulness

- Stating arrival time in $(a_i, r]$ changes nothing.
- Stating arrival time in (r, s) influences the transition time τ but not the pricing. Still offered price \$5.



Analysis: Truthfulness

- Stating arrival time in $(a_i, r]$ changes nothing.
- Stating arrival time in (r, s) influences the transition time τ but not the pricing.
- Stating arrival time $\geq s$ influences the transition, but price not improved.



Analysis: Competitive Ratio

- **Claim**: Competitive ratio for efficiency is *e+o(1)*, assuming all valuations are distinct.
- Case 1: Item sells at time t. Winner is highest bidder among first $\lfloor n/e \rfloor$. With probability ~1/e, this is also the highest bidder among all n agents.
- Case 2: Otherwise, the auction picks the same outcome as the secretary algorithm, whose success probability is ~1/e.

Analysis: Competitive Ratio

- **Claim**: Competitive ratio for revenue (wrt Vickrey) is e^2+o(1), assuming all valuations are distinct.
- Estimate probability of selling to highest bidder at second-highest price. Use same two cases as before.
- Case 1: Probability ~$1/e^2$.
  - (prob 1/e that second highest also in first half)
- Case 2: Probability ~(1/e)(1/e).
  - (prob. that highest in first-half is the second-highest overall is 1/e conditioned on highest in second-half, prob. that choose highest in case 2 is 1/e)

- 4+o(1)-competitive for revenue (and also efficiency), by setting transition time at n/2.
- Lower-bounds of 2-competitive for efficiency, 1.5-competitive for revenue (in our model).

General approach -- Two phase

- "Learning phase"
  - use a sequence of bids to set price for rest of auction

Transition:
  - be sure that remains truthful for agents on transition
- "Accepting phase"
  - exploit information, retain truthfulness

Multi-Item Online Auction (k>1)

- Adopt a variation on the Dual-Price Sampling Optimal Threshold (DSOT) auction (Goldberg, Hartline et al'01; also Segal'03).
- (Learning) Choose pivotal bidder, $j \sim \text{Binom}(n, \frac{1}{2})$.
- (Transition) Sell up to $s = \lceil k/2 \rceil$ items at time τ, to agents present and bidding above (s+1)-st bid price so far.
- (Accepting) After τ, set price to be the revenue-optimizing fixed price, p^{opt} for bids in first half. Sell item to bid$\geq p^{opt}$ while supply.

- Truthfulness: have $p(s+1) \leq p^{opt}$
- Constant-competitive with $\mathcal{F}^{(2)}$ for revenue.
- Constant-competitive for efficiency (and also revenue), by setting $s = \lceil k/3 \rceil$, and adopting p(t)=(s+1)-st bid in accepting phase. (i.e. a lower price.)

Characterization of Truthful auctions

(Hajiaghayi, Kleinberg, Mahdian, and P., ACM-EC05)

- **Definition**. Allocation rule $f: \Theta^n \rightarrow \{0,1\}^n$ is **monotonic** if for every agent i and every $(\theta,\theta')\in\Theta^n$ with $[a'_i,d'_i]\subseteq[a_i,d_i]$, and $w_i \geq w'_i$, we have $f_i(\theta)\geq f_i(\theta')$.
- **Definition**. The "critical value" price is:
  $$ps_i(a_i,d_i,\theta_{-i}) = \begin{cases} \min w'_i \text{ s.t. } f_i(\langle a_i,d_i,w'_i\rangle, \theta_{-i})=1 \\ \infty, \quad \text{if no such } w'_i \text{ exists} \end{cases}$$
- **Definition**. The "critical period" is the first $t\in[a_i,d_i]$ with minimal $ps_i(a_i,t,\theta_{-i})$.

Theorem. An online auction is truthful if and only if the allocation rule, f, is monotonic, sets payment equal to critical value, and assigns item after the critical period.

Via an Agent-independent Price Schedule

- Define an agent-independent price schedule, $ps_i(t,\theta_{-i})$ for allocation in period t
- Allocate good to agent if and only if $ps_i(t',\theta_{-i}) \leq w_i$ for some $t'\in[a_i,d_i]$, at price $ps_i(a_i,d_i,\theta_{-i})=\min_{t'\in[a_i,d_i]} ps_i(t',\theta_{-i})$.
- Allocate no earlier than period t' for which $ps_i(t',v_{-i})$ is minimal in $[a_i,d_i]$.

Via an Agent-independent Price Schedule

- Define an agent-independent price schedule, $ps_i(t,\theta_{-i})$ for allocation in period t
- Allocate good to agent if and only if $ps_i(t',\theta_{-i}) \leq w_i$ for some $t'\in[a_i,d_i]$, at price $ps_i(a_i,d_i,\theta_{-i})=\min_{t'\in[a_i,d_i]} ps_i(t',\theta_{-i})$.
- Allocate no earlier than period t' for which $ps_i(t',v_{-i})$ is minimal in $[a_i,d_i]$.
- **Example**: single-unit auction. Let $j=\lfloor n/e \rfloor$, and use "outside bid" refer to a bid from an agent ≠i.

$$ps_i(t,\theta_{-i}) = \begin{cases} \infty & \text{for} < j-1 \text{ outside bids} \\ b^{(1)}_{<j\backslash i} & \text{for } j-1 \text{ outside bids} \\ b^{(1)}_{\leq j\backslash i} & \text{for} \geq j-1 \text{ outside bids, before item sells} \\ \infty & \text{otherwise} \end{cases}$$

Prior-Free: Key Variations

- Limited-supply ($k\geq1$) of goods, sell in any period before time horizon, T.
  - single-unit demand
  - multi-unit demand
- Reusable goods, can sell up to k units in each time period. Finite time horizon, T.
  - single-period demand
  - multi-period demand

Formal Model: Re-usable Goods

(Hajiaghayi, Kleinberg, Mahdian, and P., ACM-EC05)

- One good in each time slot (can extend to k>1).
- Agent value $\langle a_i, d_i, w_i \rangle$. Value for one time slot in $[a_i, d_i]$.
- No-late departures (i.e. $[a'_i, d'_i] \subseteq [a_i, d_i]$)
  - (WiFi) suppose can verify presence, and fine an agent that reports $d'_i > d_i$ but leaves at d_i.
  - (Grid) reasonable to hold result until time d' with some small probability.
- Necessary to assume NLD to achieve a bounded competitive ratio on efficiency (Lavi & Nisan'05)

Theorem. Online auction is truthful if and only if the allocation rule, f, is monotonic, sets payment equal to critical value. Can assign at any time in interval w/ NLD.

Example: Grid scheduling

Value	\$100	\$80	\$60
Arrival:	11am	11am	12pm
Patience:	2hrs	2hrs	1hr
Duration:	1hr	1hr	1hr

Allocation rule: In each period, t, allocate the good to the highest unassigned bid.
Payment rule: Pay smallest amount could have bid and still received good (in some period).

monotone: smaller [a',d'], smaller w'_i cannot help.
reduces to seq. of Vickrey for impatient bidders.

Efficiency: Competitive Analysis

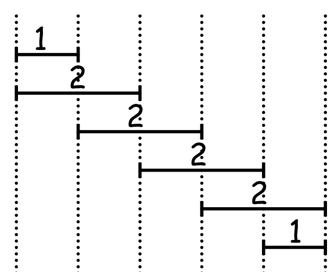
2-competitive wrt efficiency, (maximum-weighted matching in bipartite graph).

(**Tight**. But, 1.618 poss. without incentives!)

Extends to k>1 case (still 2-competitive).

Revenue Analysis: Consider VCG

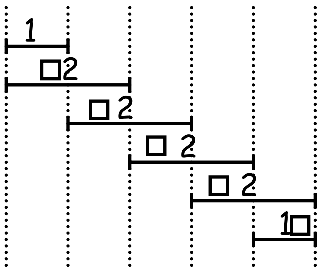
n slots, n+1 bids



n-1 high bidders

Revenue Analysis: VCG

n slots, n+1 bids

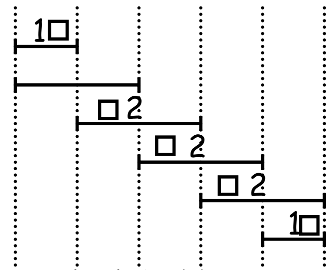


n-1 high bidders

VCG: $V^*=2(n-1)+1$

Revenue Analysis: VCG

n slots, n+1 bids



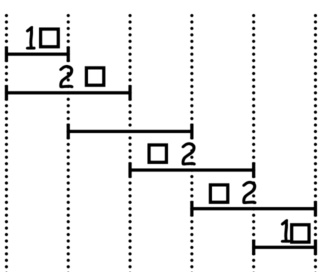
n-1 high bidders

VCG: $V^*=2(n-1)+1$

$V_{-2}=2(n-2)+2$

Revenue Analysis: VCG

n slots, n+1 bids



n-1 high bidders

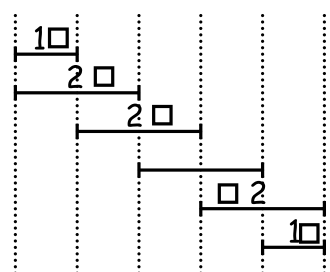
VCG: $V^*=2(n-1)+1$

$V_{-2}=2(n-2)+2$

$V_{-3}=2(n-2)+2$

Revenue Analysis: VCG

n slots, n+1 bids



n-1 high bidders

VCG: $V^*=2(n-1)+1$

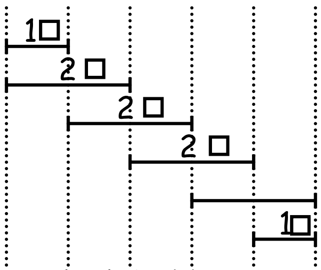
$V_{-2}=2(n-2)+2$

$V_{-3}=2(n-2)+2$

$V_{-4}=2(n-2)+2$

Revenue Analysis: VCG

n slots, n+1 bids



n-1 high bidders

VCG: $V^*=2(n-1)+1$

$V_{-2}=2(n-2)+2$

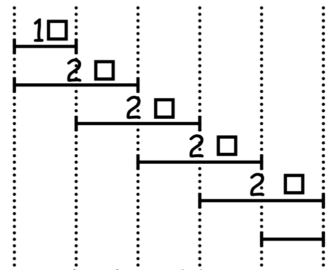
$V_{-3}=2(n-2)+2$

$V_{-4}=2(n-2)+2$

$V_{-5}=2(n-2)+2$

Revenue Analysis: VCG

n slots, n+1 bids



n-1 high bidders

VCG: $V^*=2(n-1)+1$

$V_{-2}=2(n-2)+2$

$V_{-3}=2(n-2)+2$

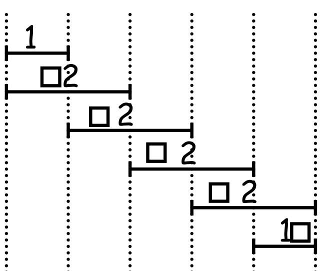
$V_{-4}=2(n-2)+2$

$V_{-5}=2(n-2)+2$

$V_{-6}=2(n-1)+1$

Revenue Analysis: VCG

n slots, n+1 bids



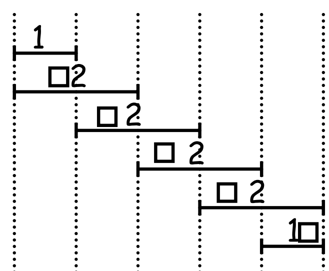
n-1 high bidders

VCG:		payments
	$V^*=2(n-1)+1$	
	$V_{-2}=2(n-2)+2$	1
	$V_{-3}=2(n-2)+2$	1
	$V_{-4}=2(n-2)+2$	1
	$V_{-5}=2(n-2)+2$	1
	$V_{-6}=2(n-1)+1$	1

Revenue(VCG) = 1(n-1)+1

Revenue: Competitive Analysis

n slots, n+1 bids



	payments
Bid_2	1
Bid_3	0
Bid_4	0
Bid_5	0
Bid_6	0

Revenue(VCG) = 1(n-1)+1

Revenue(Auc) = 1

⇒ competitive ratio can be arbitrarily bad!

- Actually, have a general negative result available for the revenue-competitiveness of a deterministic online auction for this problem.

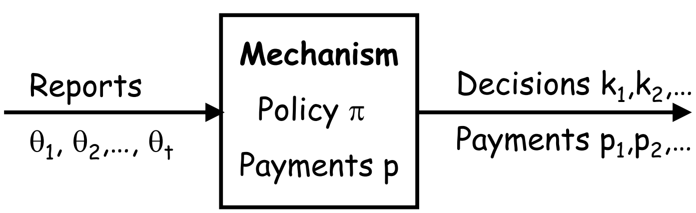
  Can achieve $O(\log_2(\phi))$ competitive with a randomized auction, for $\phi=(U/L)$, even with unknown bounds.

- Prior-Free Online Auction Design:
  - Non-reusable Goods, Finite time horizon.
- General characterization for truthful online auctions
- Prior-Free Online Auction Design:
  - Reusable Goods, infinite time horizon.
- Model-based Online Mechanisms
- Future directions.

Model-Based Online Mechanisms

(P. & Singh'03, P., Singh & Yanovsky'04)

- Agents, and the auctioneer, have a common prior.
- θ iid from distribution $g(\theta)$.
- Mechanism makes a sequence of decisions $\{k_1,k_2,\ldots\}$
- Agents $\theta_i=[a_i,d_i,v_i]$. $v_i(k)\geq 0$.
- Goal: maximize the expected sequential value.



As a Markov Decision Process

- State: $h_t=(\theta_1,\ldots,\theta_t;\ k_1,\ldots,k_{t-1})$. Time horizon T.
- Model: $Pr(h_{t+1}|h_t,k_t)$; $R(h_t,k_t)=\sum_i[v_i(k_{\leq t})-v_i(k_{\leq t-1})]$
- Policy: $\pi=\{\pi_1,\ldots,\pi_T\}$, π_t: $H_t\rightarrow K_t$
- $V^\pi(h_t)=E_\pi\{R(h_t,\pi(h_t))+R(h_{t+1},\pi(h_{t+1}))+ \ldots+R(h_T,\pi(h_T))\}$

- Efficient policy, π^*, maximizes MDP value in all states; value $V^*(h_t)$.
- Solve via dynamic programming, policy iteration, linear programming, etc.

- "Stalling" == "Action space rich enough that cannot improve policy by delaying the arrival of an agent."

- How to handle self-interest?

An Online VCG Mechanism

- Receive reports. Implement $\pi^*(h'_t)$.
- Payment: $p_i = v'_i(k^*) - \{V^*(h_{ai'}) - V^*(h_{ai'}^{-i})\}$

An Online VCG Mechanism

- Receive reports. Implement $\pi^*(h'_t)$.
- Payment: $p_i = v'_i(k^*) - \{V^*(h_{ai'}) - V^*(h_{ai'}^{-i})\}$

- **Theorem**. Given a correct model, and a policy with stalling, the online VCG is **Bayes-Nash** IC and implements the efficient policy.

An Online VCG Mechanism

(P.&Singh'03)

- Receive reports. Implement $\pi^*(h'_t)$.
- Payment: $p_i = v'_i(k^*) - \{V^*(h_{ai'}) - V^*(h_{ai'}^{-i})\}$

- **Theorem**. Given a correct model, and a policy with stalling, the online VCG is **Bayes-Nash** IC and implements the efficient policy.

$EU(\theta_i') = v_i(\pi^*(h_{ai'})) + V^*(h_{ai'}) - v'_i(\pi^*(h_{ai'})) - V^*(h_{ai'}^{-i})$

expected value to all other agents given reported type of agent i

expected value to all other agents plus expected **true** value to agent i

Remarks.

- BNIC not DSIC. Correctness of π^* requires correct model $f(\theta)$, which requires other agents play equilibrium.
- c.f. offline VCG, where the center can make the value-maximizing choice (based on reports), whatever the reports.

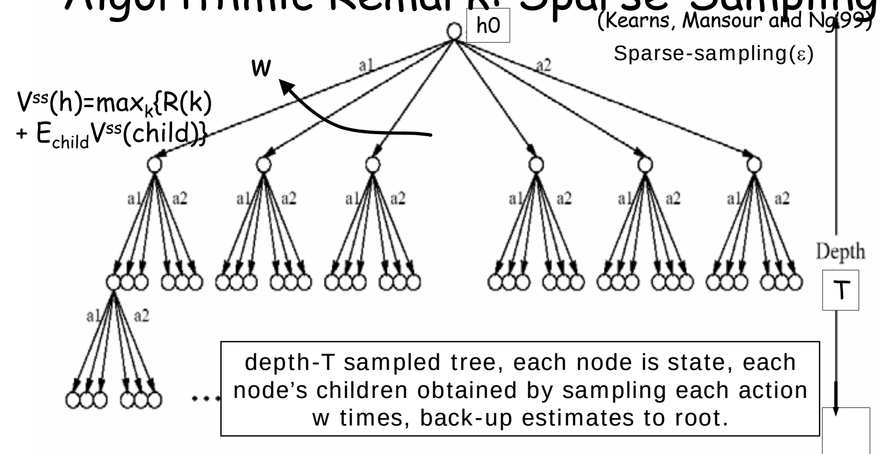
Remarks.

- BNIC not DSIC. Correctness of π^* requires correct model $f(\theta)$, which requires other agents play equilibrium.
- c.f. offline VCG, where the center can make the value-maximizing choice (based on reports), whatever the reports.

- *ex post individual-rational* given "value monotonicity", i.e. addition of an agent has a (weakly) +ve effect on total MDP value.
- *ex ante no-deficit* given "no positive externalities", i.e. addition of an agent has a (weakly) -ve effect on MDP value to others.

Algorithmic Remark: Sparse-Sampling:

(Kearns, Mansour and Ng'99)



Policy π^{ss}, estimate $V^{ss}(h)$:

$|V^*(h)-V^{ss}(h)| \leq \varepsilon$

$|V^*(h)-E\{\hat{V}^{ss}(h)\}| \leq \varepsilon$

in time $O((K\cdot w)^T)$, with $w=\text{poly}(K, 1/\varepsilon, R_{max}, T)$, for R_{max} bound on reward in a state.

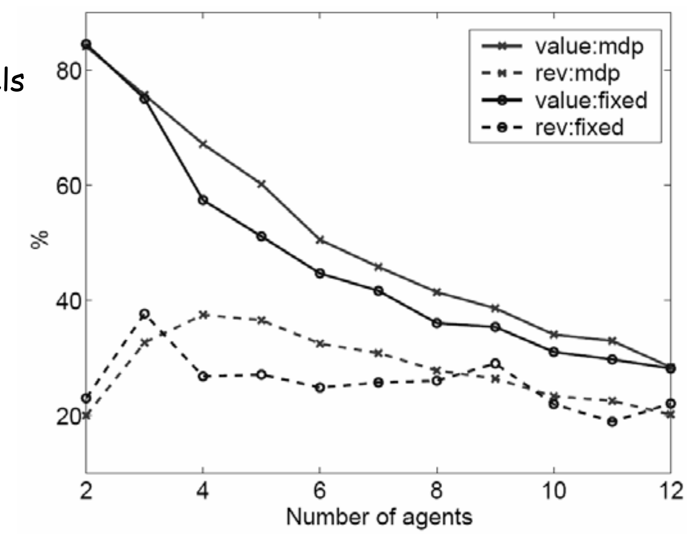
- ε-BNIC: no agent can improve its expected utility by more than ε, for any type, as long as other agents are bidding truthfully.

Theorem. For any ε, and a correct model, the sparse-sampling online VCG mechanism is ε-efficient, truthful reporting is a 4ε-BNE, and the run-time is independent in the size of state space.

Example: Eff, Rev in WiFi problem

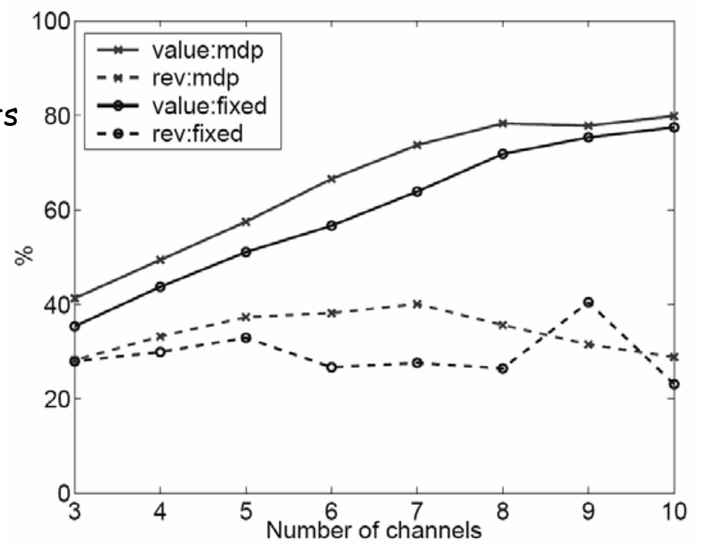
(P., Singh & Yanovsky'04)

5 channels



Revenue and value normalized by the unlimited supply value.

5 agents



Future Direction: Introduce Learning.

- What if center has only a distribution on priors, and a MLE of the model, denoted f′(θ)?
- Would like to converge to optimal π^* over time.

Future Direction: Introduce Learning.

- What if center has only a distribution on priors, and a MLE of the model, denoted f′(θ)?
- Would like to converge to optimal π^* over time.
- Main problems:
  - (A1) retaining incentive-compatibility with respect to time despite the adaptiveness of the policy.
  - (A2) retaining incentive-compatibility despite an approximate policy.

Future Direction: Introduce Learning.

- What if center has only a distribution on priors, and a MLE of the model, denoted f′(θ)?
- Would like to converge to optimal π^* over time.
- Main problems:
  - (A1) retaining incentive-compatibility with respect to time despite the adaptiveness of the policy.
  - (A2) retaining incentive-compatibility despite an approximate policy.
- Remark: the online VCG mechanism is not BNIC with an approximate model.

Future Direction: Introduce Learning.

- What if center has only a distribution on priors, and a MLE of the model, denoted $f'(\theta)$?
- Would like to converge to optimal π^* over time.
- Main problems:
  - (A1) retaining incentive-compatibility with respect to time despite the adaptiveness of the policy.
  - (A2) retaining incentive-compatibility despite an approximate policy.
- Remark: the online VCG mechanism is not BNIC with an approximate model.

- Current work: focus on a "single-minded domain". In that domain, optimal policies are monotonic, whatever the model $\Rightarrow$ can get a positive result.
- General problem of learning + MDPs is open.

Summary

- Many computational systems present dynamic resource allocation problems.
- Need to extend MD to handle dynamics.
- Two styles of analysis.
- Prior-free: DSIC mechanisms with online competitive results for non-reusable and reusable-good scenarios.
- Model-based: BNIC mechanisms to implement optimal MDP policies.

- Future direction: Allow for learning.