Efficient Online mechanisms for persistent, periodically inaccessible self-interested agents

Ruggiero Cavallo (Harvard)

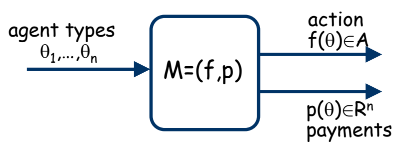
Satinder Singh (Michigan)

David C. Parkes
Harvard University
http://www.eecs.harvard.edu/econcs

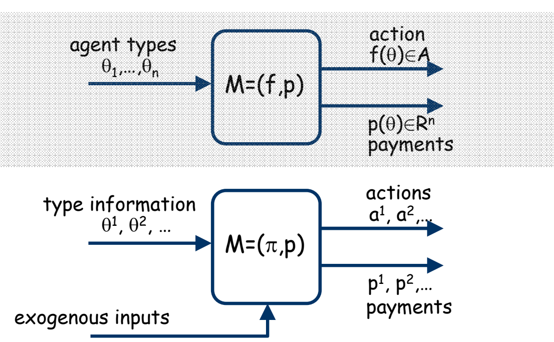
Motivating question

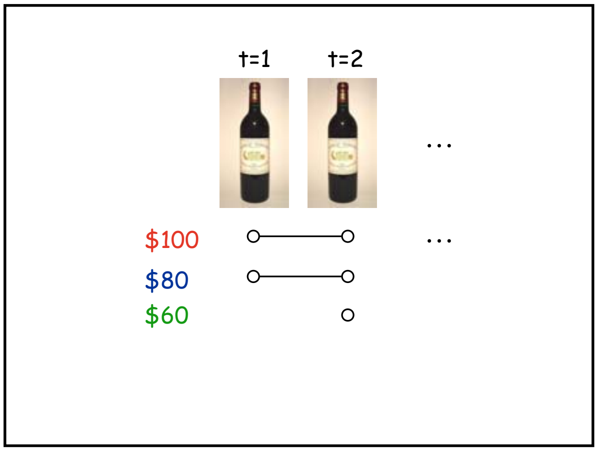
- MD in dynamic envionments!

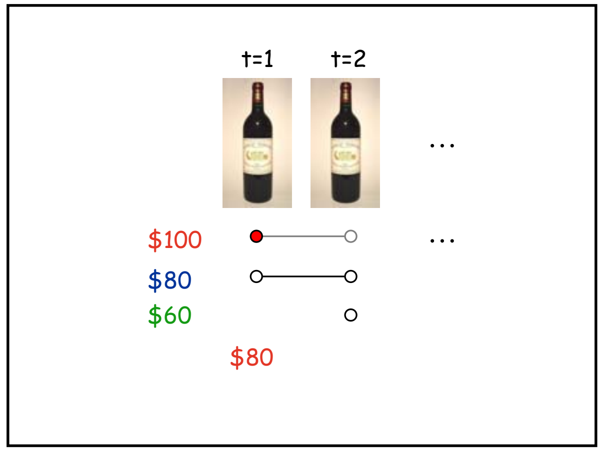
	dynamics	non-episodic
internet ads	demand & supply	contracts
social learning	information	persistent
last minute tix	demand	expressiveness
peer production	demand & supply	long tasks
pref. elicitation	information	bounded supply
task allocation	tasks	long tasks
…	…	…

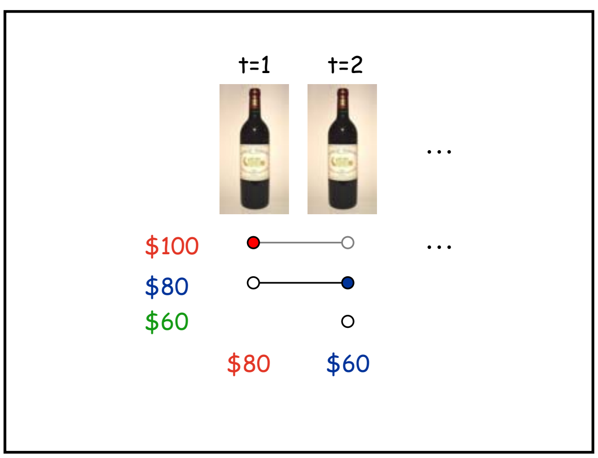


Dynamic Incentive Mechanisms









One view from CS

- Prior free:

$$\min_{z\in\mathcal{Z}} \mathbb{E}\left\{\frac{\mathrm{Val}(\pi(\theta_z))}{V^*(\theta_z)}\right\} \geq \frac{1}{c},$$

- Typical setting: (e_i, d_i, v_i, q_i)

One view from CS

- Prior free:

$$\min_{z\in\mathcal{Z}} \mathbb{E}\left\{\frac{\mathrm{Val}(\pi(\theta_z))}{V^*(\theta_z)}\right\} \geq \frac{1}{c},$$

- Typical setting: (e_i, d_i, v_i, q_i)
- Body of work
  - Lavi & Nisan '00
  - Awerbuch et al. '03
  - Porter '04
  - Hajiaghayi, Kleinberg, Parkes'04
  - Blum & Hartline '05
  - Hajiaghayi, Kleinberg, Mahdian, Parkes'05
  - Lavi & Nisan '05
  - …
- DSIC, monotonicity-based characterizations.
- Limited misreports: no-early arrival, no-late dep., etc.

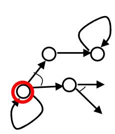
A second view

- Center (and/or agents) have a probabilistic model of dynamics of environment
- $\pi^*(s) \in \arg\max_a [r(s,a) + \gamma \sum_{s' \in S^{t+1}} Pr(s',a,s)V^*(s')]$
- Agents can misreport local model, local state

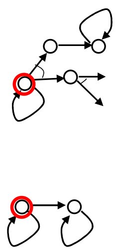
A second view

- Center (and/or agents) have a probabilistic model of dynamics of environment
- $\pi^*(s) \in \arg\max_a [r(s,a) + \gamma \sum_{s' \in S^{t+1}} Pr(s',a,s)V^*(s')]$
- Agents can misreport local model, local state
- Body of work:
  - Parkes & Singh '03, '04
  - Cavallo, Parkes & Singh '06
  - Bergemann and Valimaki '06
  - Cavallo, Parkes & Singh '07
  - Athey & Segal '07
- typically interim IC, sometimes DSIC

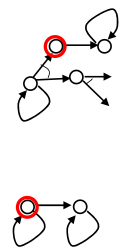
type = static component (local model)



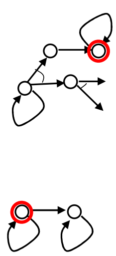
type = static component (local model)
+ dynamic component (local state)



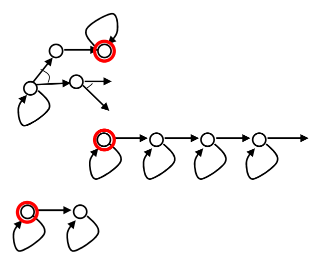
type = static component (local model)
+ dynamic component (local state)



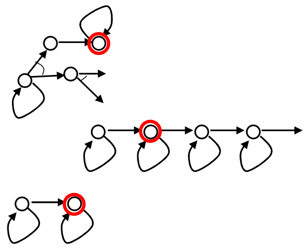
type = static component (local model)
+ dynamic component (local state)



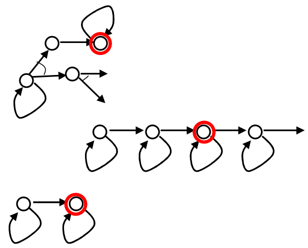
type = static component (local model)
+ dynamic component (local state)



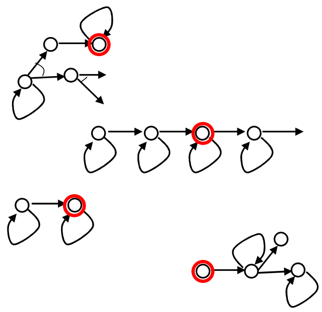
type = static component (local model)
+ dynamic component (local state)



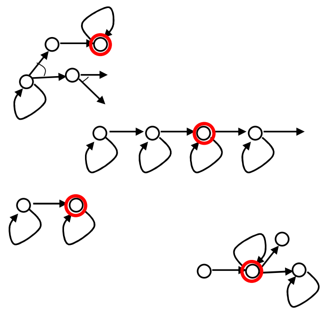
type = static component (local model)
+ dynamic component (local state)



type = static component (local model)
+ dynamic component (local state)



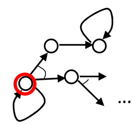
type = static component (local model)
+ dynamic component (local state)



- Dynamic types, persistent agents
- Dynamic types, arrival + departures
- Dynamic types, accessible/inaccessible

MDP: (S_i, A, r_i, τ_i)
$\tau_i(s_i,a)$ $r_i(s_i,a)$
initial state s_i^0

policy π_i: $S_i \rightarrow A$
$V^{\pi}_i(s)=E[\sum_{k=t} \gamma^{k-t} r^k_i(s^k_i,\pi(s^k))]$
$V^*_i(s)$: $\pi^*_i \in \arg\max_\pi V^{\pi}_i(s)$

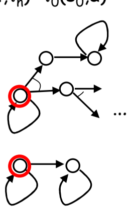


MDP: (S_i, A, r_i, τ_i)
$\tau_i(s_i,a)$ $r_i(s_i,a)$
initial state s_i^0

$S=S_0 \times S_1 \times \ldots \times S_n$
$r(s,a)=\sum_i r_i(s_i,a)$
$\tau=(\tau_0,\tau_1,\ldots,\tau_n)$ $\tau_0(s_0,a)$
$a \in A(s)$

policy π_i: $S_i \rightarrow A$
$V^{\pi}_i(s)=E[\sum_{k=t} \gamma^{k-t} r^k_i(s^k_i,\pi(s^k))]$
$V^*_i(s)$: $\pi^*_i \in \arg\max_\pi V^{\pi}_i(s)$

policy π: $S \rightarrow A$
$V^{\pi}(s)=E[\sum_{k=t} \gamma^{k-t} r^k(s^k,\pi(s^k))]$
$V^*(s)$: $\pi^* \in \arg\max_\pi V^{\pi}(s)$

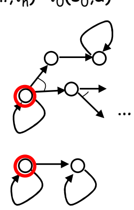


MDP: (S_i, A, r_i, τ_i)
$\tau_i(s_i,a)$ $r_i(s_i,a)$
initial state s_i^0

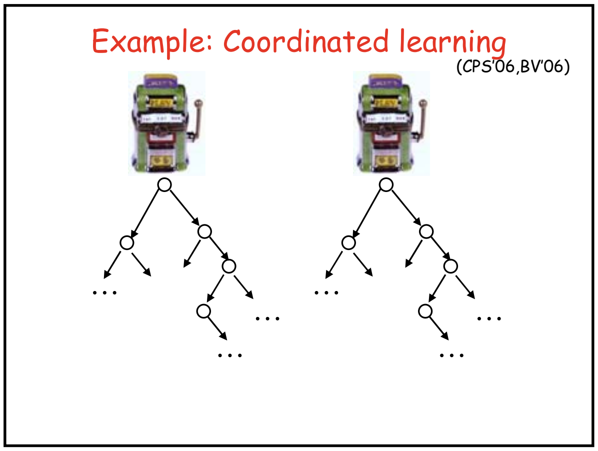
$S=S_0 \times S_1 \times \ldots \times S_n$
$r(s,a)=\sum_i r_i(s_i,a)$
$\tau=(\tau_0,\tau_1,\ldots,\tau_n)$ $\tau_0(s_0,a)$
$a \in A(s)$

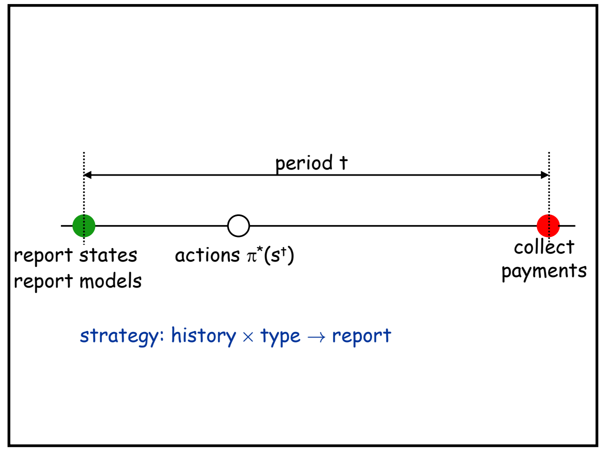
policy π_i: $S_i \rightarrow A$
$V^{\pi}_i(s)=E[\sum_{k=t} \gamma^{k-t} r^k_i(s^k_i,\pi(s^k))]$
$V^*_i(s)$: $\pi^*_i \in \arg\max_\pi V^{\pi}_i(s)$

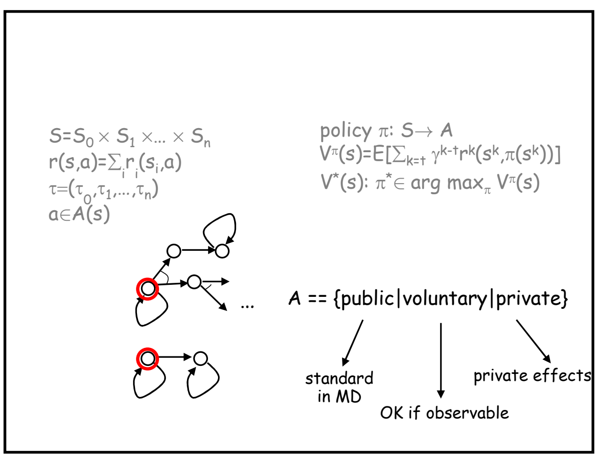
policy π: $S \rightarrow A$
$V^{\pi}(s)=E[\sum_{k=t} \gamma^{k-t} r^k(s^k,\pi(s^k))]$
$V^*(s)$: $\pi^* \in \arg\max_\pi V^{\pi}(s)$

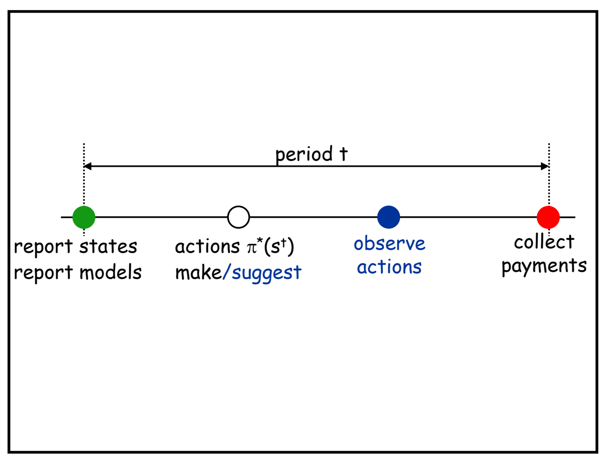


Assumption: *CIA*
(**C**onditional **I**ndependence given **A**ctions)
$r_i((s_i,s_{-i}),a)=r_i((s_i,s_{-i}'),a)$
$\tau_i((s_i,s_{-i}),a)=\tau_i((s_i,s_{-i}'),a)$


Example: Coordinated learning
(CPS'06,BV'06)

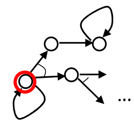

period t
report states
report models
actions $\pi^*(s^t)$
collect
payments
strategy: history × type → report


$S=S_0 \times S_1 \times ... \times S_n$
$r(s,a)=\sum_i r_i(s_i,a)$
$\tau=(\tau_0,\tau_1,...,\tau_n)$
$a \in A(s)$
policy π: $S \rightarrow A$
$V^\pi(s)=E[\sum_{k=t} \gamma^{k-t} r^k(s^k,\pi(s^k))]$
$V^*(s)$: $\pi^* \in \arg\max_\pi V^\pi(s)$
A == {public|voluntary|private}
standard
in MD
OK if observable
private effects

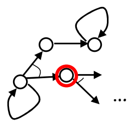

period t
report states
report models
actions $\pi^*(s^t)$
make/suggest
observe
actions
collect
payments

- Dynamic types, persistent agents
- Dynamic types, arrival + departures
- Dynamic types, accessible/inaccessible

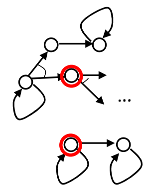
arrival/departure process



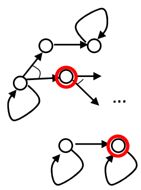
arrival/departure process



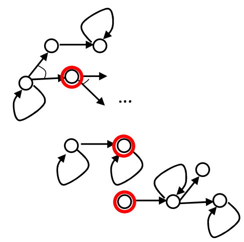
arrival/departure process



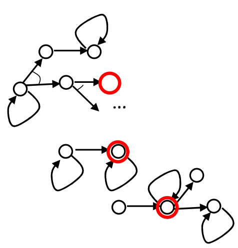
arrival/departure process



arrival/departure process



arrival/departure process



static type == (local model,initial state)

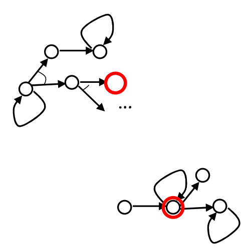
$H(s_0)$: set of static types present

$s=(s_0, \{s_i\}^{i \in H(s_0)}) \in S$

τ_0: $S_0 \times A \rightarrow S_0$ ⟵ arrival process

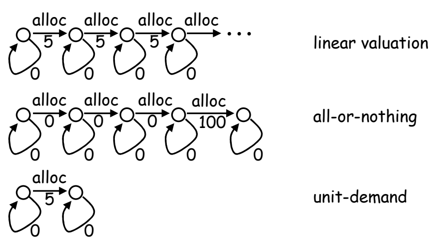
$r(s,a)=\sum_{i \in H(s_0)} r_i(s_i,a)$

departure: local absorbing state



Special case: Deterministic Local Models

- Only dynamics are arrival/departure of static types
- Agents arrive and declare reward for all future sequences of actions via deterministic local model



static type == (local model,initial state)

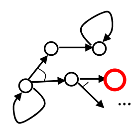
$H(s_0)$: set of static types present

$s=(s_0, \{s_i\}^{i \in H(s_0)}) \in S$

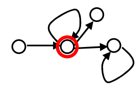
τ_0: $S_0 \times A \rightarrow S_0$ ⟵ arrival process

$r(s,a)=\sum_{i \in H(s_0)} r_i(s_i,a)$

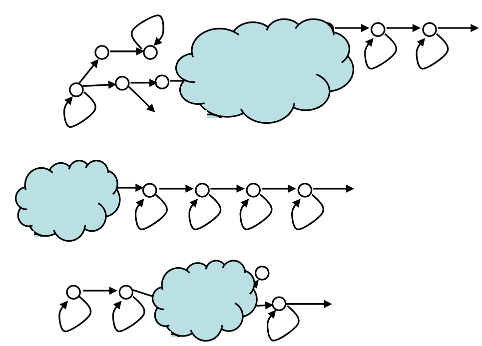
departure: local absorbing state



Assumption: CIA
Conditional Independence given Actions
$\Pr(\theta^t \mid \theta^{1..t-1}, a^{1..t-1}) = \Pr(\theta^t \mid a^{1..t-1})$



- Dynamic types, persistent agents
- Dynamic types, arrival + departures
- Dynamic types, accessible/inaccessible



arrival/depature = inaccess → access → inaccess

- Augment local state with accessible/inaccessible
- Inaccessible == no messages, no payments
- Can pretend to be inaccessible when not accessible (but not vice versa)
- Actions can make an agent become inaccessible
- Can have reward for actions while inaccessible

- Augment local state with accessible/inaccessible
- Inaccessible == no messages, no payments
- Can pretend to be inaccessible when not accessible (but not vice versa)
- Actions can make an agent become inaccessible
- Can have reward for actions while inaccessible

- Assumption: run but can't hide // pay the piper

A Belief-state MDP model

- **P**artially Observable **M**arkov **D**ecision **P**rocess
- Model as a belief-state MDP (Kaelbling et al. 96)
  - $BS = S_0 \times BS_1 \times \ldots \times BS_n$, $BS_i = \Delta(S_i)$
  - when accessible, $bs_i \in BS_i$, reduces to point mass
  - $r_i(bs_i^t a)$: in expectation on underlying states
  - policy π^*: $BS \to A$

Persistent & Dynamic type

CIA: $r(s_i,a)$ $\tau(s_i,a)$

BV'06

$V^*_{-i}(s) - V^*_{-i}(s|\pi^*(s))$

charge each period

Arr/dep & Static type

$r(s_i,a)$ $\tau(s_i,a)$

$Pr(\theta^t|\theta^{1..t-1}, a^{1..t-1}) = Pr(\theta^t|a^{1..t-1})$

PS'03

$v_i - (V^*(s^e) - V^*(s^e_{-i}))$

charge @ departure

Arr/dep & Dynamic type

$r(s_i,a)$ $\tau(s_i,a)$

$Pr(\theta^t|\theta^{1..t-1}, a^{1..t-1}) = Pr(\theta^t|a^{1..t-1})$

CPS'07

$V^*_{-i}(s) - V^*_{-i}(s|\pi^*(s))$

charge each period

Persistent, Access/inaccess, dynamic type

$r(s_i,a)$ $\tau(s_i,a)$ RBCH

CPS'07

charge $\hat{T}_i^t(bs^t) = \sum_{k=t-\delta(t)}^{t} \frac{T_i^k(bs^k)}{\gamma^{t-k}}$

where δ^t is # periods been inaccessible

- Expected equilibrium payoff to agent i is $V^*(s^t|f_i)-C_i(s^t)$, where $C_i(s^t)$ is independent of strategy f_i

- persistent: in every state
- arrival/departure: in arrival period
- access/inaccess: in every accessible state

- *Groves intuition: agent i's payoff is aligned with expected true value of policy, maximizes this by reporting true type given that policy π is optimal*

- Expected equilibrium payoff to agent i is $V^*(s^t|f_i)-C_i(s^t)$, where $C_i(s^t)$ is independent of strategy f_i

- persistent: in every state
- arrival/departure: in arrival period
- access/inaccess: in every accessible state

- *Groves intuition: agent i's payoff is aligned with expected true value of policy, maximizes this by reporting true type given that policy π is optimal*

- Subtle: $V^*(s^t) - V^*(s^t_{-i})$ does not work!

Relating to "Offline VCG"

offline VCG	dynamic VCG
• optimal x^* // DSIC	• optimal π^* // interim IC
• ex post IR	• interim IR ($V^*(s^t) \geq V^*_{-i}(s^t_{-i})$)
• ex post ND	• interim ND ($V^*_{-i}(s^t_{-i}) \geq V^*_{-i}(s^t)$)
need private values	

ε–correctness of model
ε–optimality of policy
critical for incentive properties

Computational issues

- Special cases will be feasible to solve offline, e.g. via value or policy iteration. Payments come from Q-values.

Computational issues

- Special cases will be feasible to solve offline, e.g. via value or policy iteration. Payments come from Q-values.
- Local Markov chains (Gittins'74)
  - discounting & infinite time horizon
  - activate agent with maximal index

Computational issues

- Special cases will be feasible to solve offline, e.g. via value or policy iteration. Payments come from Q-values.
- Local Markov chains (Gittins'74)
  - discounting & infinite time horizon
  - activate agent with maximal index
- In general, large (even continuous) state space. Need online sample-approximation methods.

Computational issues

- Special cases will be feasible to solve offline, e.g. via value or policy iteration. Payments come from Q-values.
- Local Markov chains (Gittins'74)
  - discounting & infinite time horizon
  - activate agent with maximal index
- In general, large (even continuous) state space. Need online sample-approximation methods.
- Tree-sampling (Kearns, Mansour & Ng'99, Ng & Jordan'00)
  - applied to online MD by Parkes & Singh'04
  - scale exponentially in number of actions and look-ahead, but independent of state space
  - ε-optimality and ε-equilibrium (poly in $1/\varepsilon$)

Computational issues

- Special cases will be feasible to solve offline, e.g. via value or policy iteration. Payments come from Q-values.
- Local Markov chains (Gittins'74)
  - discounting & infinite time horizon
  - activate agent with maximal index
- In general, large (even continuous) state space. Need online sample-approximation methods.
- Tree-sampling (Kearns, Mansour & Ng'99, Ng & Jordan'00)
  - applied to online MD by Parkes & Singh'04
  - scale exponentially in number of actions and look-ahead, but independent of state space
  - ε-optimality and ε-equilibrium (poly in $1/\varepsilon$)
- Trajectory-sampling (van Hentenryck et al.)
  - suitable in deterministic local model // independent arrival process setting
  - avoids exponential scaling, but hard to prove bounds (but see Mercier, Upfal & van Hentenryck '07)

Parallel literature: DSIC

- set-valued setting (e,d,r,L)
- limited misreports and/or delayed allocation and late departures
- makes monotonicity in sense of earlier arrival, later departure, smaller L sufficient

→ competitive analysis

→ online stochastic combinatorial optimization + "output ironing" (Parkes&Duong, Constantin&Parkes)

Parallel literature: DSIC

- set-valued setting (e,d,r,L)
- limited misreports and/or delayed allocation and late departures
- makes monotonicity in sense of earlier arrival, later departure, smaller L sufficient

→ competitive analysis

→ online stochastic combinatorial optimization + "output ironing" (Parkes&Duong, Constantin&Parkes)

Open problems

Indirect methods

- Direct revelation of type information costly, and counter to privacy concerns
- Prefer to allow agents to retain private model, perform local planning, report enough information to allow optimal joint actions
- Look for price-based methods

Cashless design

- What about dynamic coordination in systems without money?
- Dynamic voting protocols
- c.f. Jackson "linking mechanism"

Richer models, Economic Qs

- Two-sided models, e.g. dynamic (combinatorial) exchanges. Only have results for simple settings. E.g., Blum et al.'06, Parkes & Bredin'07
- Redistribution methods. Can the methods of Bailey, Cavallo, Guo & Conitzer, Moulin, Hartline etc. be leveraged?
- What about revenue?
- Learning by center?
- etc…

Approximate incentive compatibility

- Likely will need to relax IC requirements
- One approach
  - have local agents help with computation
  - suggest improved policies
  - fold new computational methods into the mechanism
  - c.f. Nisan & Ronen (second-chance)
  - c.f. Cavallo (belief-coordination mechanism)
- Another approach
  - local stability
  - first price vs. generalized second price vs. GSP, Threshold rule, etc.
  - tolerable manipulability

Applications/Comput. Directions

- Meta-deliberation auctions
  - use these techniques to coordinate the deliberation process of agents
  - extend MD to embrace more of the process of decision making
- Online CAs
  - obvious application area
  - lots of algorithmic challenges
- AI architectures
  - can this be used to coordinate computational processes?

Summary

- Rich agenda in embracing dynamics within mechanism design
- Relevant to many kinds of coordination problems, including computational processes
- Fun new challenges ☺

www.eecs.harvard.edu/econcs