Learning Anti-Coordination

Boi Faltings, Ludek Cigler

AI Laboratory, EPFL
boi.faltings@epfl.ch
http://liawww.epfl.ch/

2016

Resource Sharing

N agents $A_1, .., A_N$ share C resources $R_1, ..R_C$:

- C wireless channels and N transmitters.
- C keywords, N advertisers.
- C roadways, N cars.

Characteristics

- Every agent wants to use the capacity of 1 resource.
- Multiple agents accessing at the same time = collision:
  - channel interference ⇒ loss of data.
  - many bidders for a keyword ⇒ high price.
  - many cars on a road ⇒ slow traffic.
- Need to *anti-coordinate*.

Why is anti-coordination difficult?

- Coordination (everyone does the same) is *symmetric*
- Anti-coordination is *asymmetric*
- Creating asymmetry among selfish agents has a cost: price of anonymity.

Resource Allocation Game

		B	
		Y	A
	Y	(0,0)	(0,1)
A	A	(1,0)	(γ, γ)

- $\gamma < 0 =$ payoff of collision
- 2 efficient, but unfair pure-strategy Nash equilibria (PSNE) with payoff 1
- 1 fair mixed-strategy Nash equilibrium with payoff 0

Mixed Strategy Payoff

- Playing a mixed strategy between access and yield means agent must be indifferent between access and yield.
- But yield has utility 0 $\Rightarrow$ expected utility of accessing must be the same.
- $p(access) = \frac{1}{1-\gamma}$ (recall $\gamma < 0$)
- definitely not a good equilibrium!

Symmetric Strategies with Highest Payoff

- Agents could use mixed strategies that maximize their combined payoff:

$$p(access) = \frac{1}{2(1-\gamma)}$$

$\Rightarrow$ combined payoff $= \frac{1}{4(1-\gamma)}$

- but not a Nash equilibrium!
- could be made an equilibrium in a repeated game if agents are punished for deviating (Folk Theorem)

Price of Anonymity

- Anonymous agents: all treated the same.

$\Rightarrow$ Equilibrium σ must be symmetric - otherwise agents would fight for their favorite one.

- Price of Anonymity:

$$R(\sigma) = \frac{maxE(\tau)}{E(\sigma)}$$

where τ is any equilibrium and $E(x) =$ social welfare of equilibrium x.

- Price of Anonymity in simple resource allocation game: $1/0 = \infty$.

Correlated Equilibria (Aumann, 1974)

- correlation device recommends action for each agent.
- equilibrium if no agent gains by deviating from recommendation.
- correlation device should ensure *fairness* and *efficiency*.
- agents need no intelligence: just follow the recommendation.

Correlated Equilibria in Resource Allocation

- Correlation device flips coin between efficient equilibria.

$\Rightarrow$ ex-ante both agents have same expected utility.

- Price of Anonymity $= 1$
- Many agents and resources $\Rightarrow$ many efficient equilibria, fairness becomes more complex.

Smart Agents, Stupid Device

Settings have no room for central authority $\Rightarrow$ replace coordination device by a combination of:

- correlation signal $X \in \{0, 1, .., K - 1\}$, and
- for agent i, an agent-specific mapping μ_i from $X \rightarrow A_i$ mapping each signal value to a different action of the agent.
- agents start out identical, and must learn the mapping.
- mappings should be fair and distribute access evenly.

Correlation Signals

Must be observable by all agents.
Should fluctuate at a reasonable rate.
Examples:

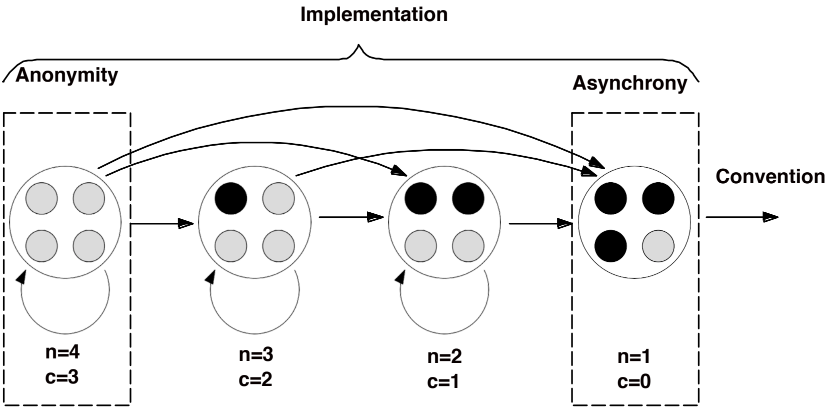
- explicit coordinator.
- time (time-division schemes).
- weather, sunspots, foreign exchange rates, etc.
- anything that fluctuates in an ergodic fashion.

$\Rightarrow$ over time, different equilibria will be played so all agents can get access to the resource.

Conventions (Bhaskar, 2000)

- Agents need to learn consistent mappings signal $\Rightarrow$ action (Y/A)
- Highest welfare: for each signal value, at most C of N agents access, the others yield.
- Once mapping is learned, agents keep following it as a *convention* (Bhaskar, 2000).
- Learning phase = *implementation*
- Theorem (Cigler & Faltings, 2012): For any convention, there exists an equilibrium implementation.

Convention Learning



Learning Algorithm

- Bhaskar: random play until a consistent mapping is reached, then this becomes the convention $\Rightarrow$ too inefficient.
- Use a distributed no-regret learning algorithm instead.
- Agents observe a fluctuating *coordination signal* $s \in \{1..k\}$
- Every change in s initiates a new stage game.
- Choice of actions: access or yield; if yield can monitor another resource.
- Feedback: success or failure if access, empty or used if yield.

Algorithm (Cigler & Faltings, 2011/2013)

Agent i learns strategy $f_i : \{1..k\} \to \{0, 1..C\}$; initialized randomly.
$f_i(j) = l$: for $s = j$, access resource l or yield if $l = 0$.

- Observe coordination signal s
- If $f_i(s) > 0$:
    - access resource $f_i(s)$
    - if failure, with probability p set $f_i(s) \leftarrow 0$
- else
    - monitor random resource c
    - if c was free, set $f_i(s) \leftarrow c$

Converged when $(\forall i, j) f_i(j) > 0 \Rightarrow (\forall k) f_i(j) \neq f_k(j)$

Efficiency

Agents learn an *efficient* set of strategies:

- all collisions get resolved.
- all resources are used.

Theorem: Expected number of steps until convergence is bounded by

$$O\left(k^2 C \frac{1}{1-p}\left[C + \frac{1}{p}\log N\right]\right)$$

Quadratic in k and C, but can tolerate large N.

Fairness

- Anonymous: all players have equal chance to win access.
- Jain index $J[X] = \frac{E[x]^2}{E[x^2]}$ measures fairness, J[X]=1 means perfectly fair.
- Fairness depends on value space of coordination signal:
  - if $k < \frac{N}{C}$, some agents can never access the resource.
  - if $k = \omega(\frac{N}{C})$, Jain Index goes to 1 as $N \to \infty$.
  - if $k > \frac{1-\epsilon}{\epsilon}\left(\frac{N}{C} - 1\right)$, then $J > 1 - \epsilon$.
- Choosing backoff probabilities so that agents that already have many resources back off more easily improves convergence and fairness (Cigler & Faltings, 2013).

Rationality

- Why would rational agents play along with the algorithm?
- Why not grab all slots and force others to yield?
- Rational agents want to grab all channels....

Equilibrium Payoff

Claim: equilibrium payoff is equal to 0.

- No agent will yield (in a collision) unless it's indifferent between access and yield.

$\Rightarrow$ expected payoff (any play) $=$ expected payoff (all yield)

- but expected payoff (all yield) $= 0$ - because other agents will grab all access to the resource.

Escaping the dilemma

- Ensure that supply satisfies all demand.

⇒ even agents who always yield will eventually get to use the resource.

⇒ payoff (all yield) no longer zero; rational backoff probability p exists.

- (Cigler & Faltings 2014) analyze a simplified algorithm.

Simplified Algorithm (Cigler & Faltings 2014)

- Agent i learns strategy $f_i : \{1..k\} \rightarrow \{0, 1..C\}$.
- initialize to all 0.
- observe coordination signal s.
- If $f_i(s) > 0$, access resource $f_i(s)$.
- Otherwise let l be one of c unclaimed resources (no traffic or collision in previous episode) and with probability p/c,
    - access l
    - if success, $f_i(s) \leftarrow l$.

Maintain list of unclaimed resources by monitoring (not entirely realistic).

Market Convention (Cigler & Faltings, 2014)

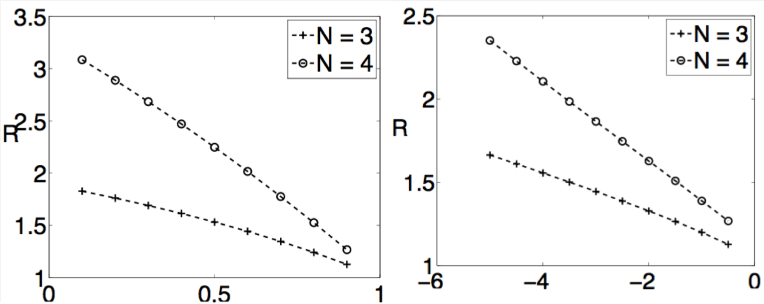
- Reduce demand by charging for successful use of resource to balance capacity and demand.
- Can be implemented by resource monitor (charge for use of bandwidth, roadway, keywords, etc.)
- Often already exists naturally: limited needs and budgets.

Results from Simplified Game

- "Bourgeois" convention (no limit on demand): probability of access $p = 1$ for every signal, expected utility is 0 (only collisions). (in fact, it would never converge).
- "Market" convention: demand limited to one resource/agent: supports $p \in (0..1)$ as equilibrium, with positive expected utility.

Price of Anonymity

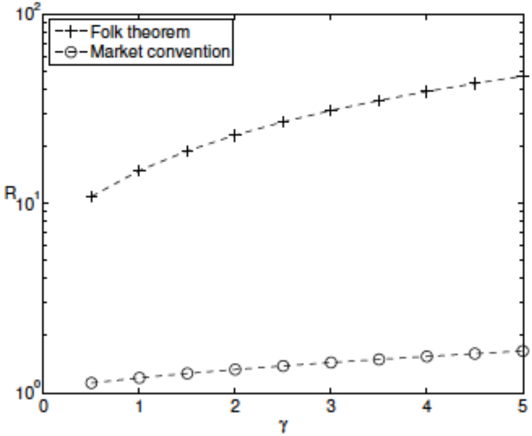
- Price of Anonymity depends strongly on discount factor δ, cost of collision γ



- (shown for single resource, k=N)

Comparison with Folk Equilibrium

Coordinated equilibrium gives much higher efficiency than uncoordinated folk equilibrium:



Price of Anonymity is almost =1!

Punishment

- Fix allowable resource use by each agent.
- Monitor actual use; if it exceeds allocation make resource unusable (wireless jamming, block road, etc.).

⇒ exceeding allowed usage not an equilibrium strategy for anyone.

Conclusions

- Sharing resources requires anti-coordination.
- Fairness requires symmetric equilibria: price of anonymity
- Anti-coordination using a common signal can be learned...
- ...but participating in the learning algorithm may not be rational.
- Managing supply/demand can solve this problem.

L. Cigler, B. Faltings. Decentralized Anti-coordination Through Multi-agent Learning. Journal of Artificial Intelligence Research, 47, 441-473, 2013.

L. Cigler, B. Faltings, Symmetric Subgame-Perfect Equilibria in Resource Allocation. Journal of Artificial Intelligence Research, 49, 323-361, 2014,