Social Learning In Sensor Networks

Vikram Krishnamurthy
University of British Columbia,
Vancouver, Canada.

Social Learning In Sensor Networks

Vikram Krishnamurthy
University of British Columbia,
Vancouver, Canada.

... from a statistical signal processing/stochastic control perspective...

Social Learning In Sensor Networks

Vikram Krishnamurthy
University of British Columbia,
Vancouver, Canada.

... from a statistical signal processing/stochastic control perspective...

We consider autonomous decision making systems:
How do local and global decisions interact?

Social Learning In Sensor Networks

Vikram Krishnamurthy
University of British Columbia,
Vancouver, Canada.

... from a statistical signal processing/stochastic control perspective...

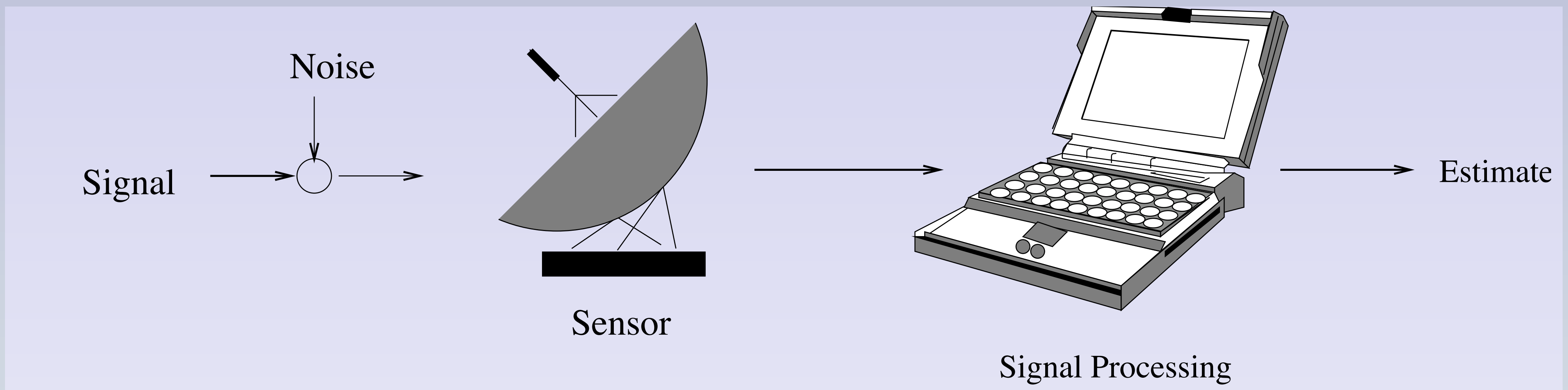
We consider autonomous decision making systems:
How do local and global decisions interact?

Game theory and social learning will be used as analysis/synthesis tools.

Sensor–Adaptive Signal Processing

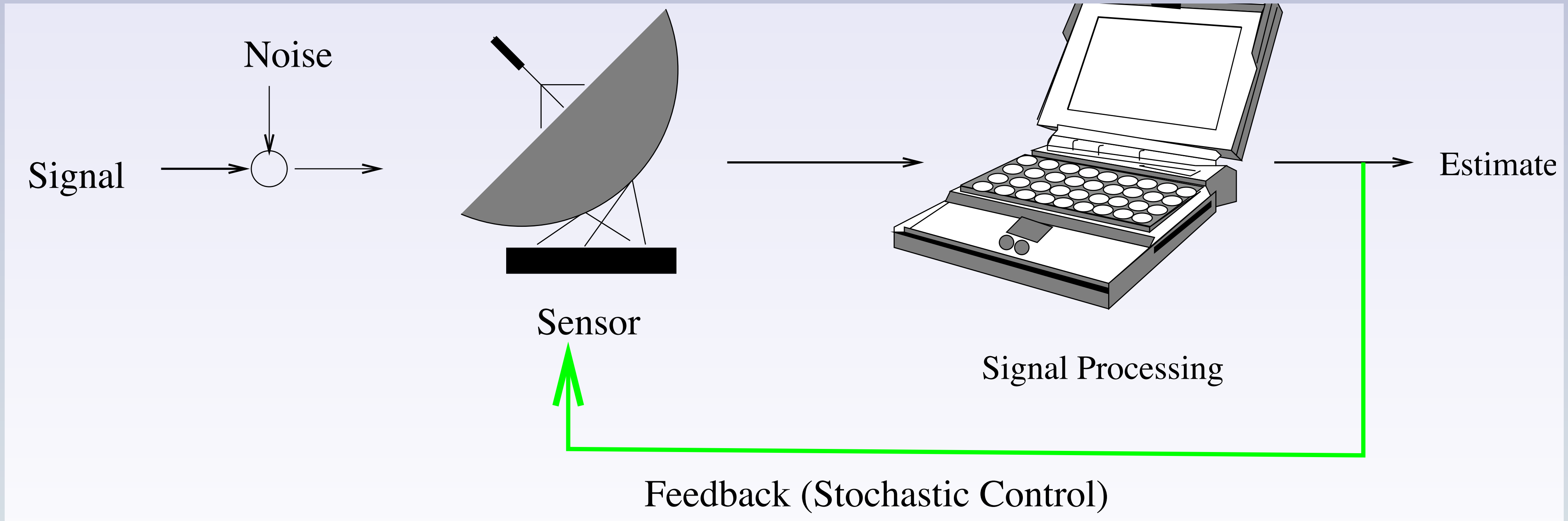
Sensor–Adaptive Signal Processing

Statistical signal processing: Extract signal from noise



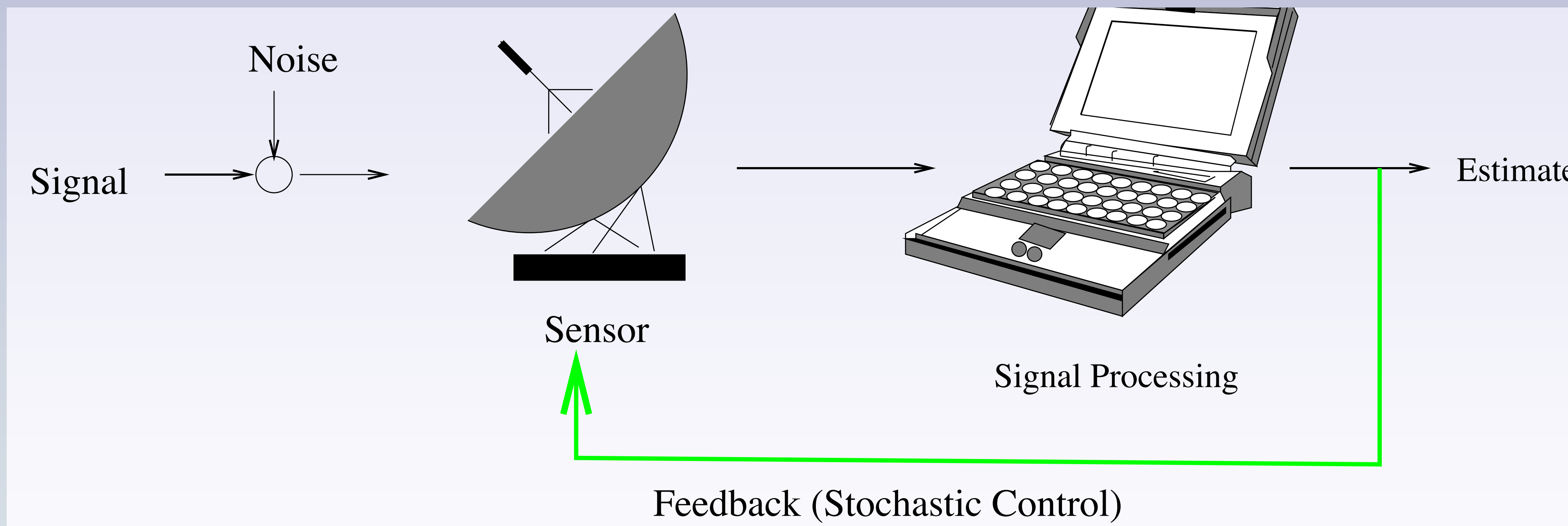
SENSOR–ADAPTIVE SIGNAL PROCESSING

Sensor-Adaptive signal processing: Dynamically adapt sensor behavior.

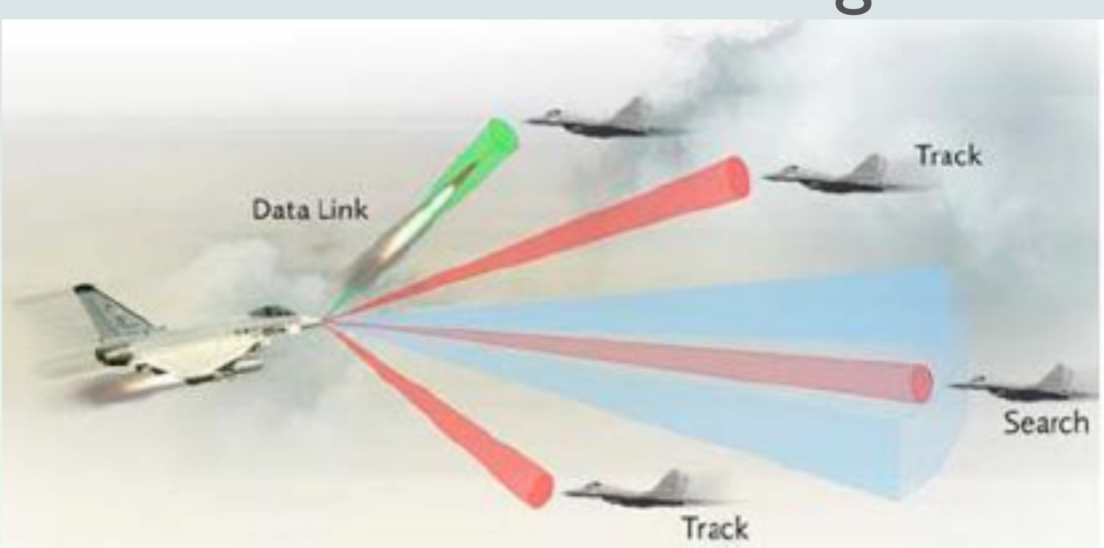


SENSOR–ADAPTIVE SIGNAL PROCESSING

Sensor-Adaptive signal processing: Dynamically adapt sensor behavior.



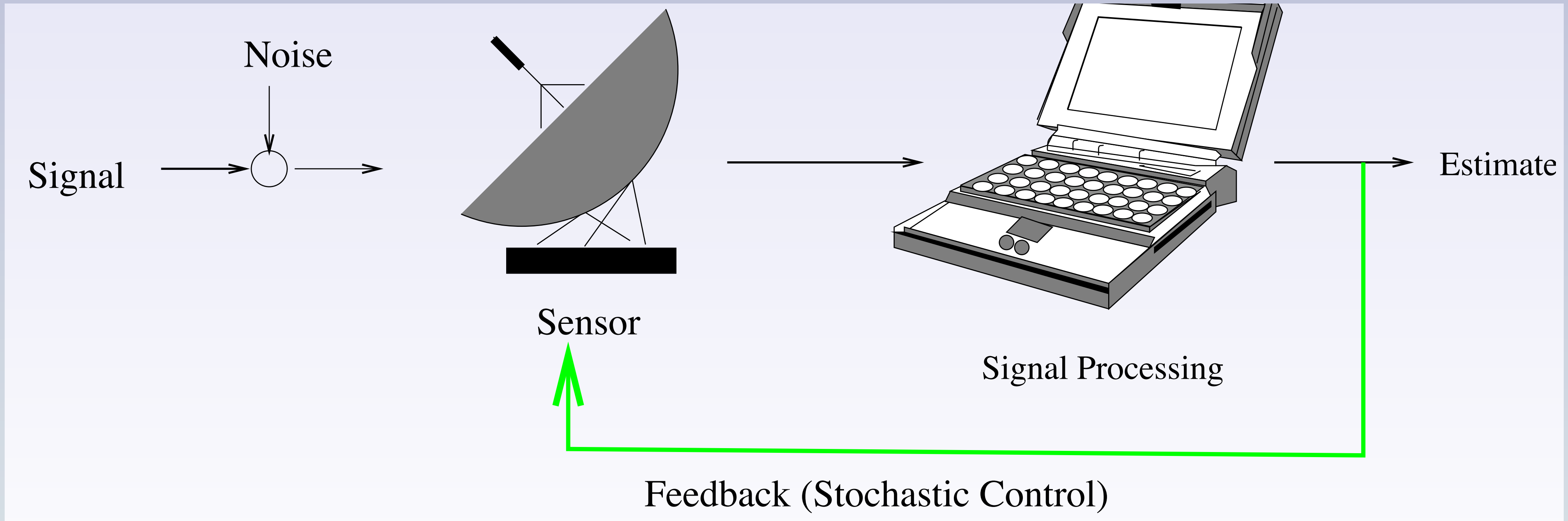
Centralized Sensor Management



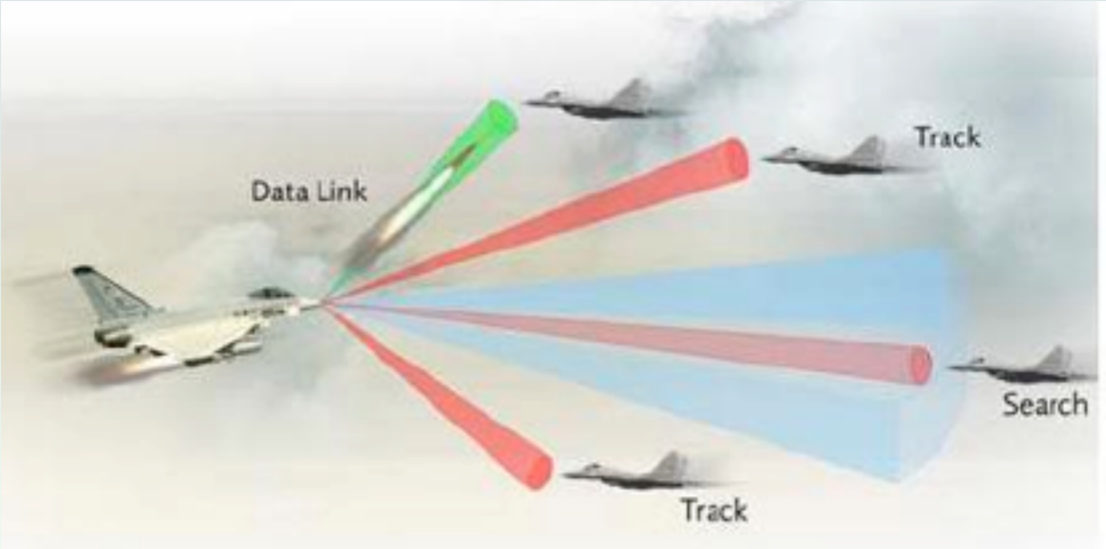
Stochastic control of multi-function radar

Sensor–Adaptive Signal Processing

Sensor-Adaptive signal processing: Dynamically adapt sensor behavior.

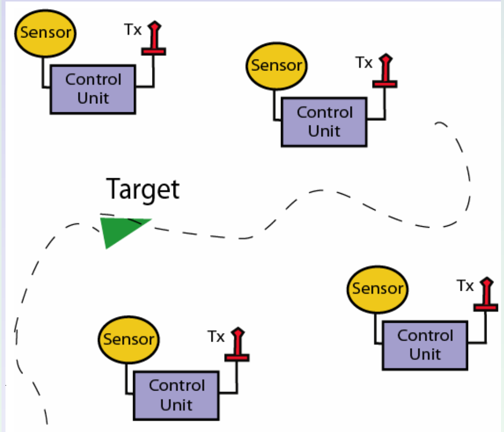


Centralized Sensor Management



Stochastic control of multi-function radar

Today's talk: Decentralized sensor management

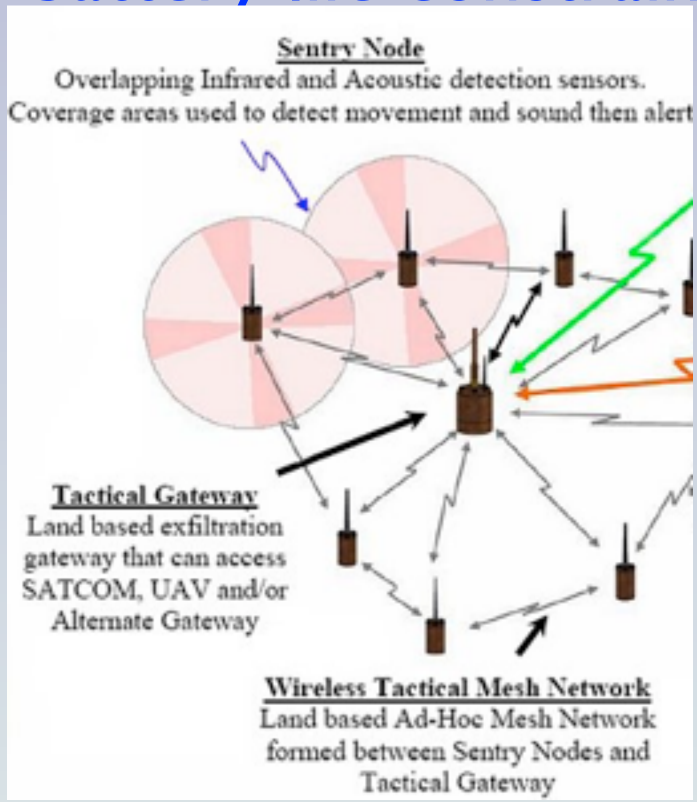


Key tool: game theory

DECENTRALIZED AUTONOMOUS DECISION MAKING

Example 1. Unattended Ground Sensor Network: Mass produced sensors; battery life constraints

Sentry sensors

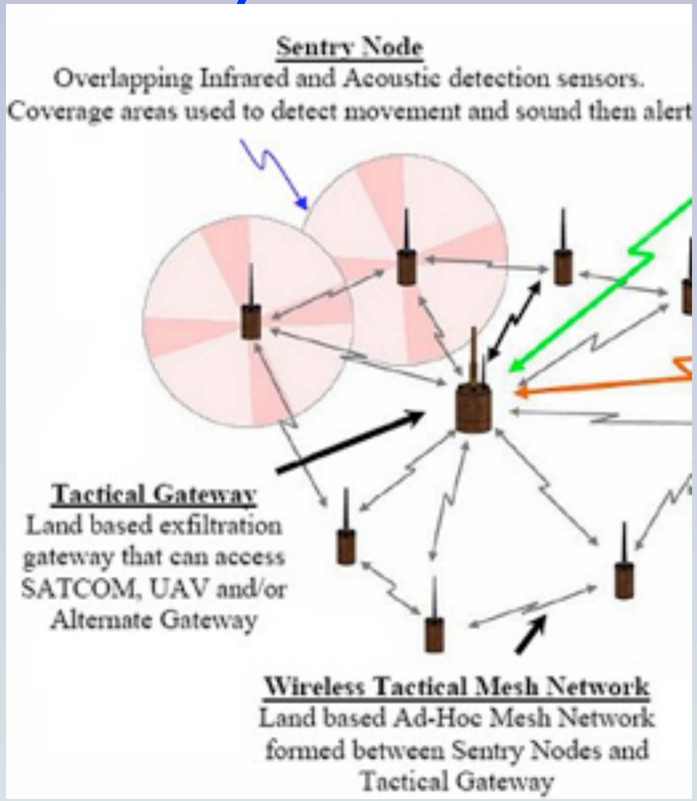


Weight	0.6 Kg
Onboard Sensors	IR Motion, GPS Low-res Imagery IR motion acoustic, seismic magnetic
RF Range	$\geq$ 300 m
Power	AA battery

DECENTRALIZED AUTONOMOUS DECISION MAKING

Example 1. Unattended Ground Sensor Network: Mass produced sensors; battery life constraints

Sentry sensors



Weight	0.6 Kg
Onboard Sensors	IR Motion, GPS Low-res Imagery IR motion acoustic, seismic magnetic
RF Range	≥ 300 m
Power	AA battery

1. Autonomous Sensor Activation: Can simple local behavior (mass-produced) yield useful global behavior? game theoretic analysis
2. How can sensors learn from other sensors to make local decisions? How do local decisions affect global decisions? social learning
3. If each sensor deploys a simple adaptive filter, is the global behavior rational? game theoretic learning

Decentralized Autonomous Decision Making

Example 2. Economic Systems: Speculative currency attacks Crashes, Bubbles and Booms, Information Delays in Financial markets

- George-Marios Angeletos and Ivan Werning (2006), "Crises and Prices: Information Aggregation, Multiplicity, and Volatility," *American Economic Review*, 96 (5): 1720–36.
- Andrew G. Atkeson, (2001), "Rethinking Multiple Equilibria in Macroeconomic Modeling: Comment." In NBER Macroeconomics Annual 2000, ed. Ben S. Bernanke and Kenneth Rogoff, 162–71. Cambridge, MA: MIT Press.
- Christian Hellwig, Arijit Mukherji and Aleh Tsyvinski (2006), "Self-Fulfilling Currency Crises: The Role of Interest Rates," *American Economic Review*, 96 (5): 1769-1787.

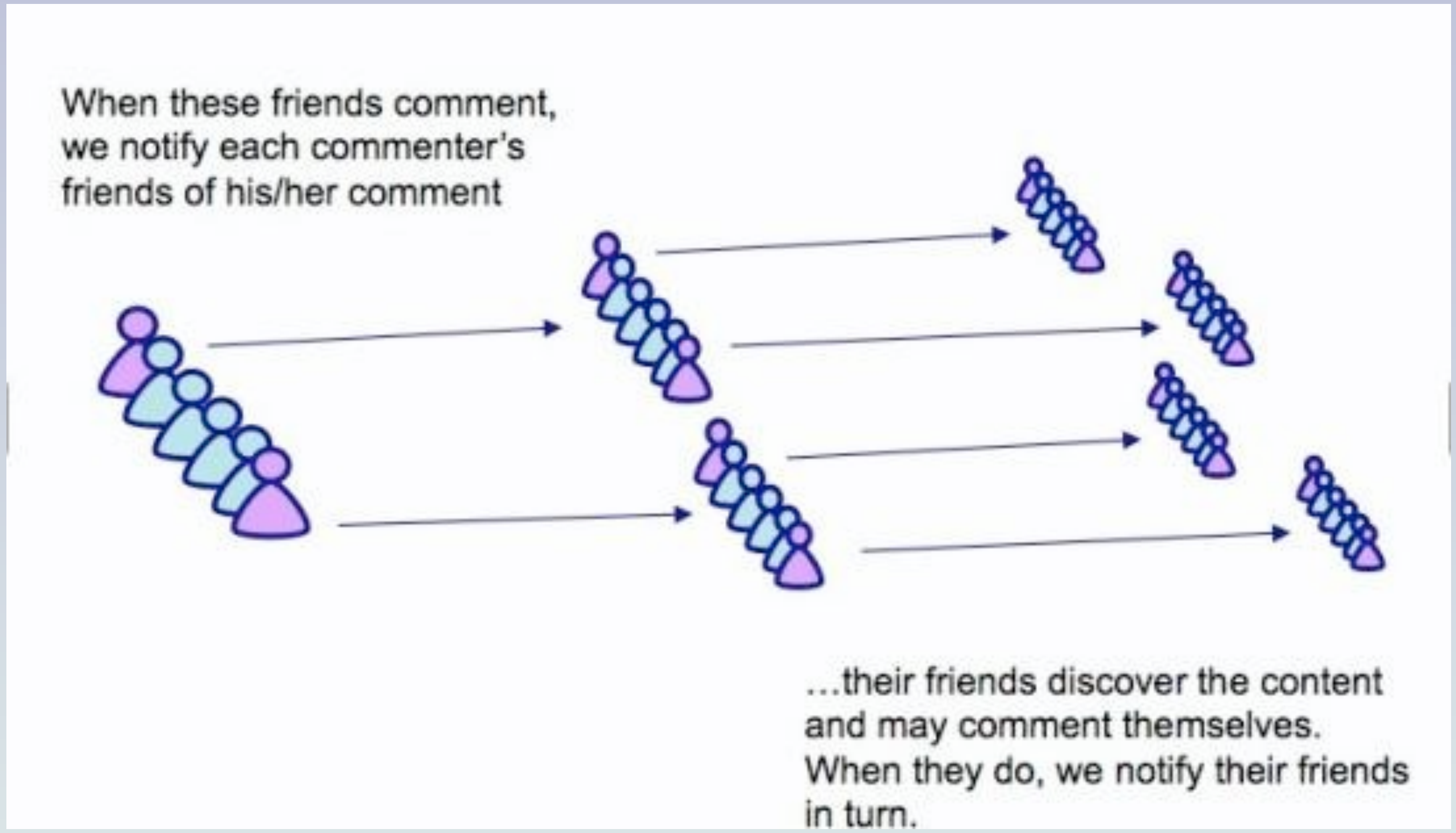
Adaptive heuristics are boundedly rational strategies (in fact, highly "bounded away" from full rationality). The main question of interest is whether such simple strategies may in the long run yield behavior that is nevertheless highly sophisticated and rational.

Econometrica, Vol. 73, No. 5 (September, 2005)

1. Autonomous Sensor Activation: Can simple local behavior (mass-produced) yield useful global behavior? game theoretic analysis
2. How can sensors learn from other sensors to make local decisions? How do local decisions affect global decisions? social learning
3. If each sensor deploys a simple adaptive filter, is the global behavior rational? game theoretic learning

Decentralized Autonomous Decision Making

Example 3. Social Networks: How to achieve consensus in decision making? Blogs...



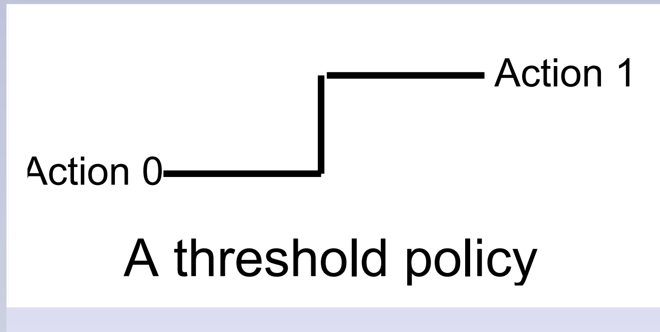
- Group behavior may not be as wise as we think.
- Crowds reduce diversity of responses
- Crowd opinion can be misleading.

1. Autonomous Sensor Activation: Can simple local behavior (mass-produced) yield useful global behavior? game theoretic analysis
2. How can sensors learn from other sensors to make local decisions? How do local decisions affect global decisions? social learning
3. If each sensor deploys a simple adaptive filter, is the global behavior rational? game theoretic learning

OUTLINE

1. **Global Games and Dynamic Sensor Activation**:

How can agents decide autonomously when to turn ON by predicting the actions of other agents?



Game Theory as an analysis tool

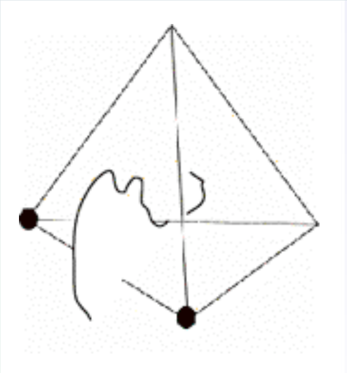
2. **Social Learning and Rational Herds:**

- How can agents learn from actions of other agents?
- How do local decisions affect global decisions?

Game Theory as a synthesis tool

3. **Adaptive Filtering Games:**

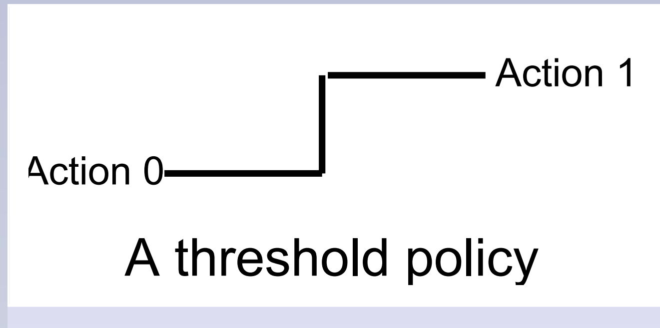
Each agent deploys an adaptive filter to optimize its decision.
Can the global system achieve consensus in policy space?
Game theory in adaptive learning.



OUTLINE

1. **Global Games and Dynamic Sensor Activation**:

How can agents decide autonomously when to turn ON by predicting the actions of other agents?



Game Theory as an analysis tool

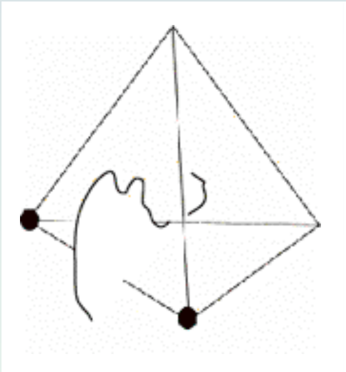
2. **Social Learning and Rational Herds:**

- How can agents learn from actions of other agents?
- How do local decisions affect global decisions?

Game Theory as a synthesis tool

3. **Adaptive Filtering Games:**

Each agent deploys an adaptive filter to optimize its decision.
Can the global system achieve consensus in policy space?
Game theory in adaptive learning.



Unifying theme: In autonomous decision making, how does local behavior affect global behavior?

PART 1: GLOBAL GAME FORMULATION

Example from Karp, Lee and Mason, *A global game with strategic substitutes and complements*, Games and Economic Behavior, 2007.
Univariate uniform distributed noise

1. Large number (continuum of agents):

$$y^{(i)} = x + w^{(i)}, \quad x \sim \pi_0, \quad w^{(i)} \sim p_W$$

2. Agent chooses action $u = \mu(y^{(i)})$

$u =$ sleep reward $= 0$

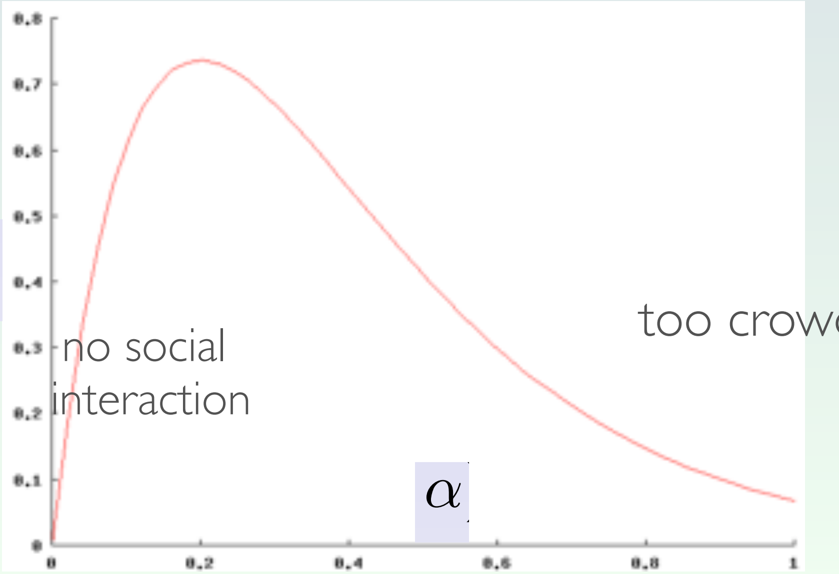
$u =$ active reward $h(x, \alpha) = cx + f(\alpha)$

$\alpha(\mu)$: frac of sensors choosing $u =$ active.

3. Agents are rational. (I know that you know). They predict and act.
What is optimal policy?

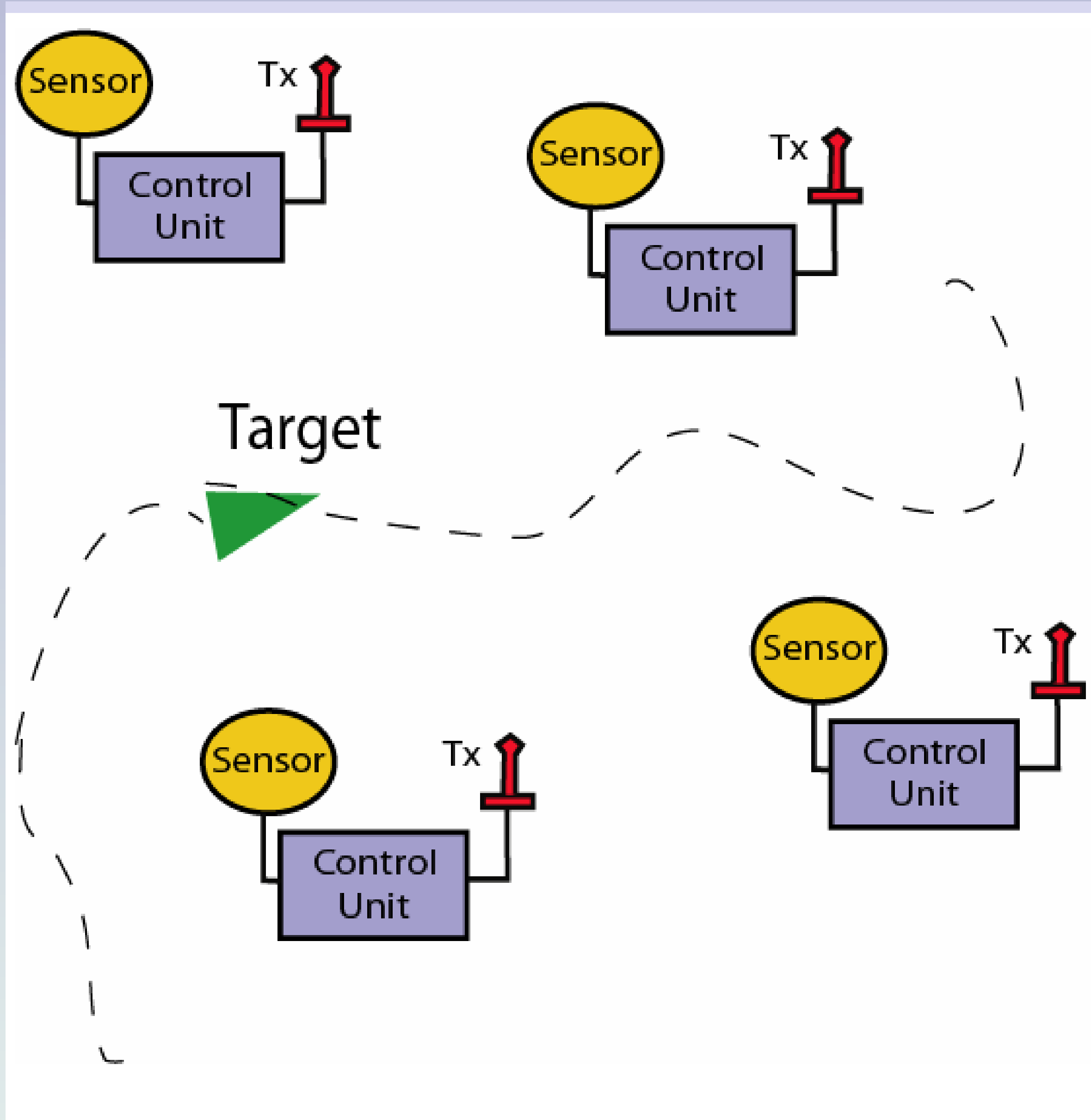
$\mu : y^{(i)} \rightarrow \{$ sleep, active$\}$ to max $\mathbf{E}\{$reward$\}$

$f(\alpha)$



Thursday, July 14, 2011

PART 1: GLOBAL GAME FORMULATION



Ex: x: measurement quality (dB, pH)
y: estimate at sensor

$$f(\alpha) = \frac{\text{Network throughput}}{\text{Global MSE}}$$

1. Large number (continuum of agents):

$$y^{(i)} = x + w^{(i)}, \quad x \sim \pi_0, \quad w^{(i)} \sim p_W$$

2. Agent chooses action $u = \mu(y^{(i)})$

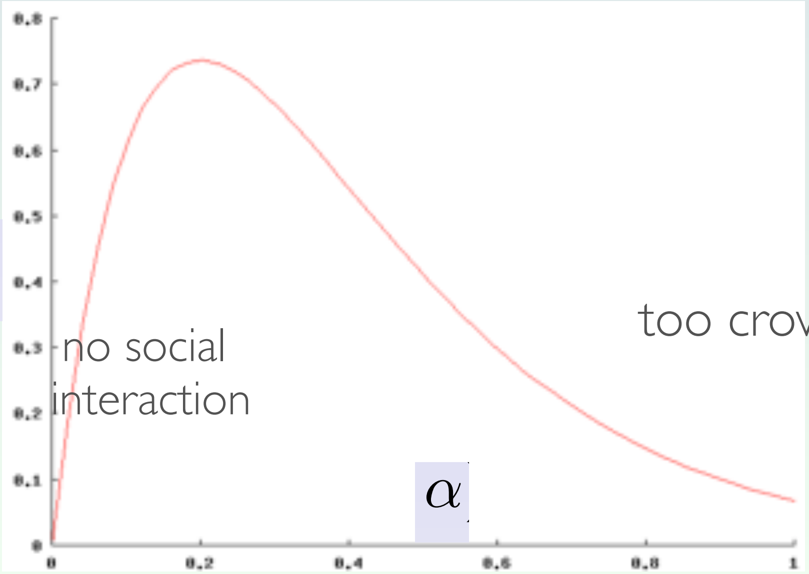
$u = \text{sleep}$ reward $= 0$
$u = \text{active}$ reward $h(x, \alpha) = cx + f(\alpha)$

$\alpha(\mu)$: frac of sensors choosing $u = \text{active}$.

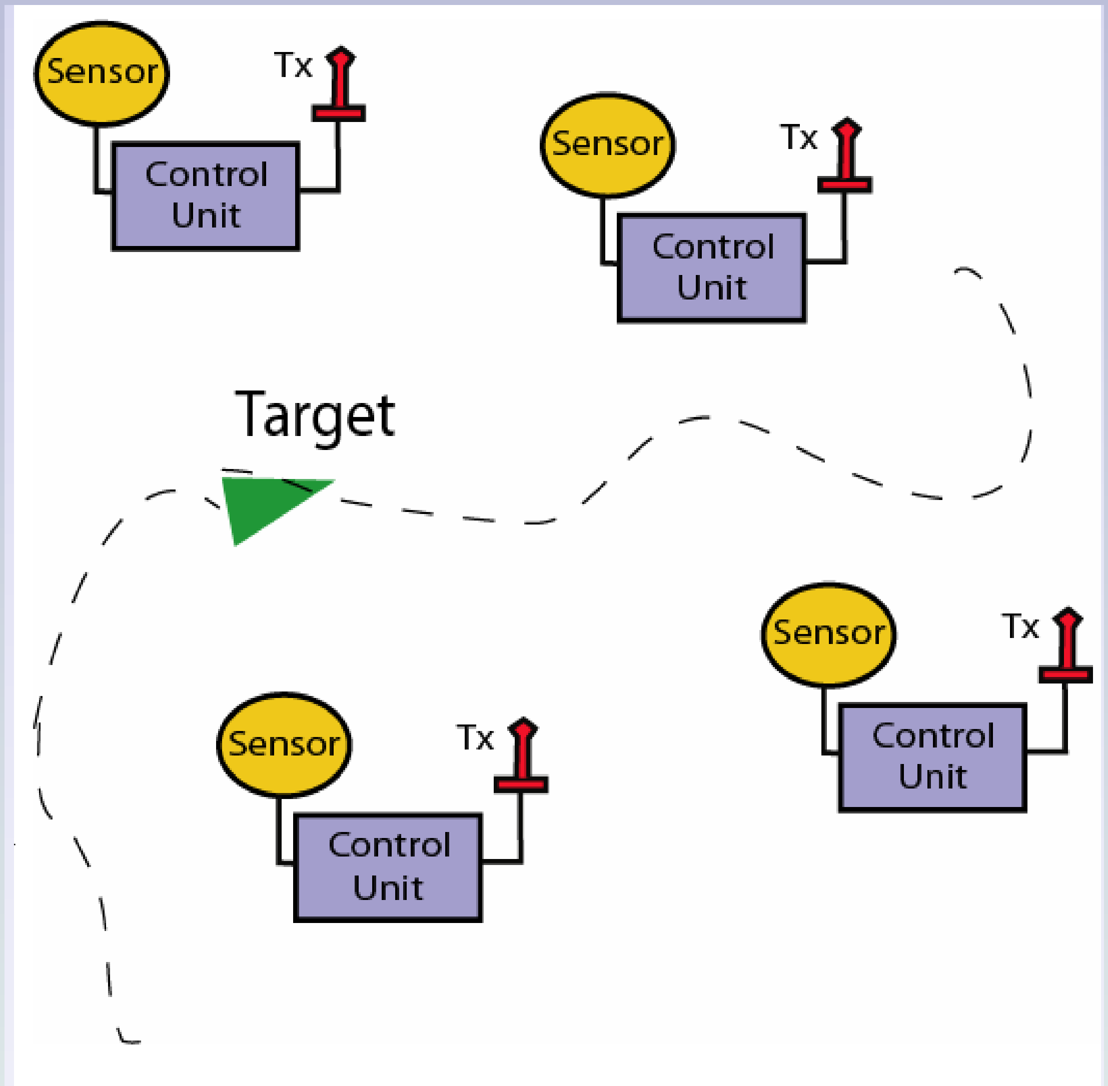
3. Agents are rational. (I know that you know). They predict and act.
What is optimal policy?

$$\mu : y^{(i)} \rightarrow \{\text{ sleep, active}\} \text{ to max } \mathbf{E}\{\text{reward}\}$$

$f(\alpha)$



Part 1: Global Game Formulation



1. Large number (continuum of agents):

$$y^{(i)} = x + w^{(i)}, \quad x \sim \pi_0, \quad w^{(i)} \sim p_W$$

2. Agent chooses action $u = \mu(y^{(i)})$

$u =$ sleep reward $= 0$

$u =$ active reward $h(x, \alpha) = cx + f(\alpha)$

$\alpha(\mu)$: frac of sensors choosing $u =$ active.

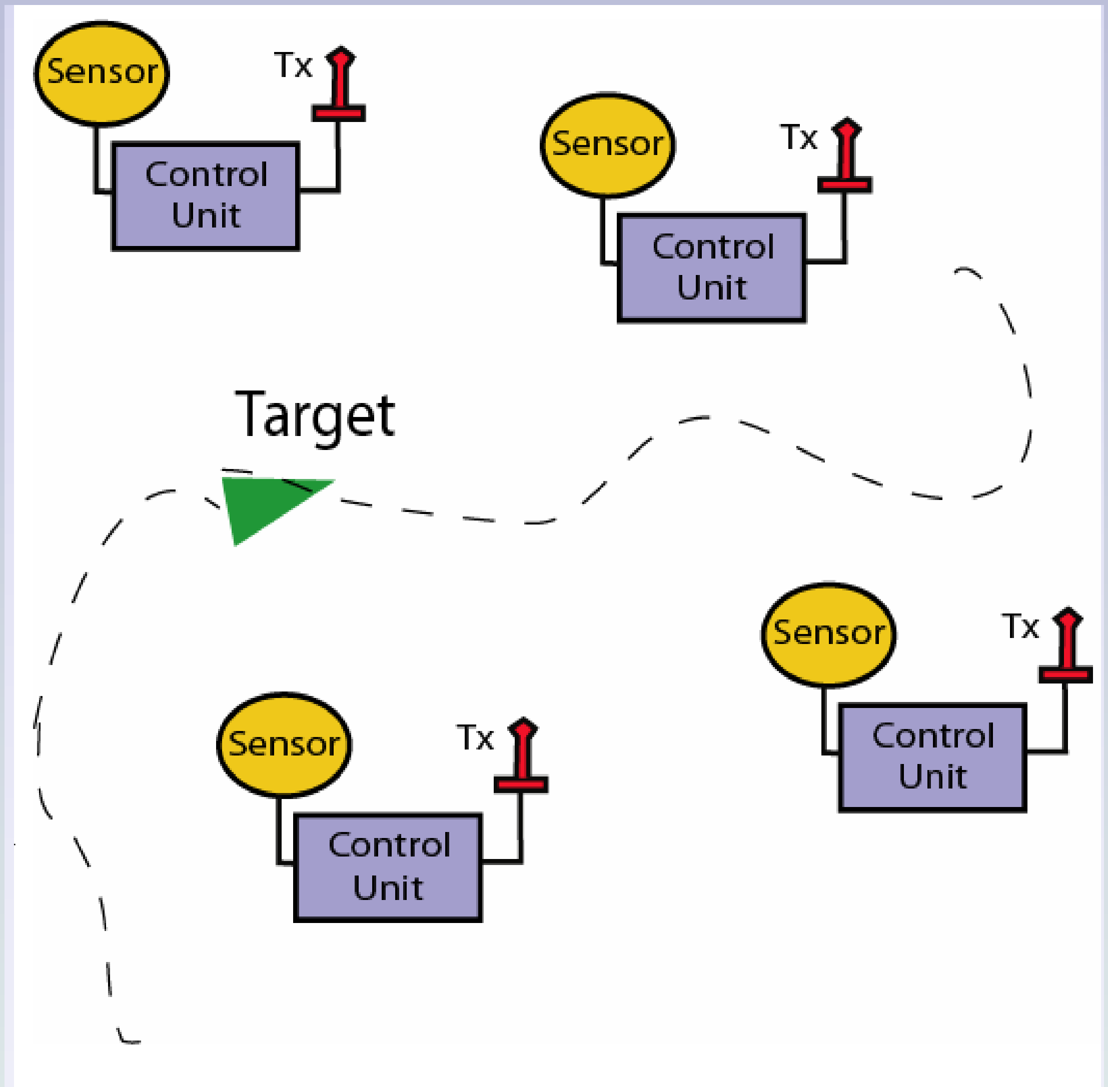
3. Agents are rational. (I know that you know). They predict and act.
What is optimal policy?

$\mu : y^{(i)} \to \{$ sleep, active$\}$ to max $\mathbf{E}\{$reward$\}$

Morris & Shin, 2000: speculative currency attacks.

monotone $f(\alpha)$

Part 1: Global Game Formulation



1. Large number (continuum of agents):

$$y^{(i)} = x + w^{(i)}, \quad x \sim \pi_0, \quad w^{(i)} \sim p_W$$

2. Agent chooses action $u = \mu(y^{(i)})$

$u =$ sleep reward $= 0$

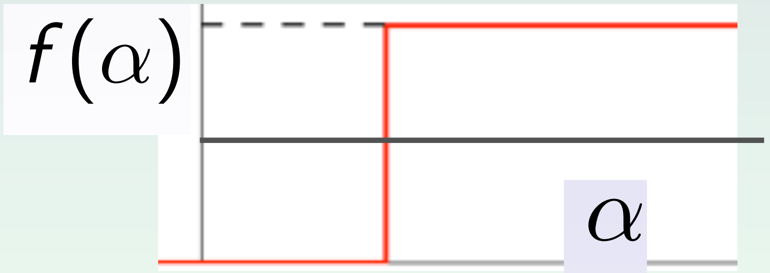
$u =$ active reward $h(x, \alpha) = cx + f(\alpha)$

$\alpha(\mu)$: frac of sensors choosing $u =$ active.

3. Agents are rational. (I know that you know). They predict and act.
What is optimal policy?

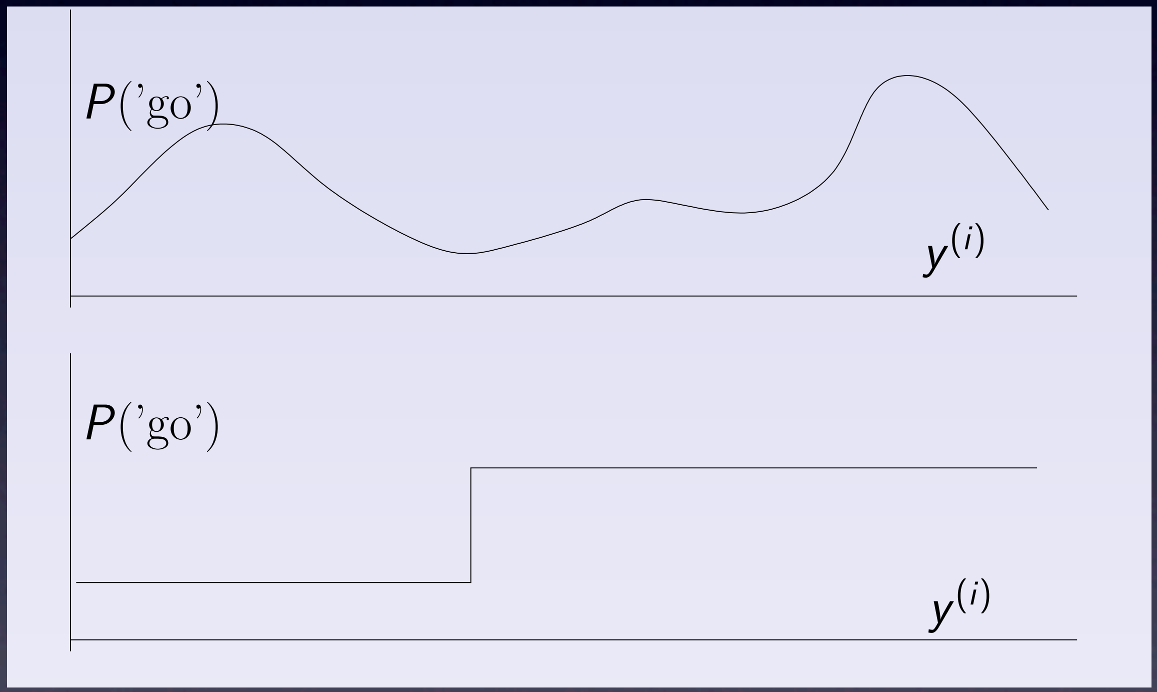
$\mu : y^{(i)} \to \{$ sleep, active$\}$ to max $\mathbf{E}\{$reward$\}$

Angeletos, Hellwig, Econometrica, 2007:
Dynamic global games of Regime change



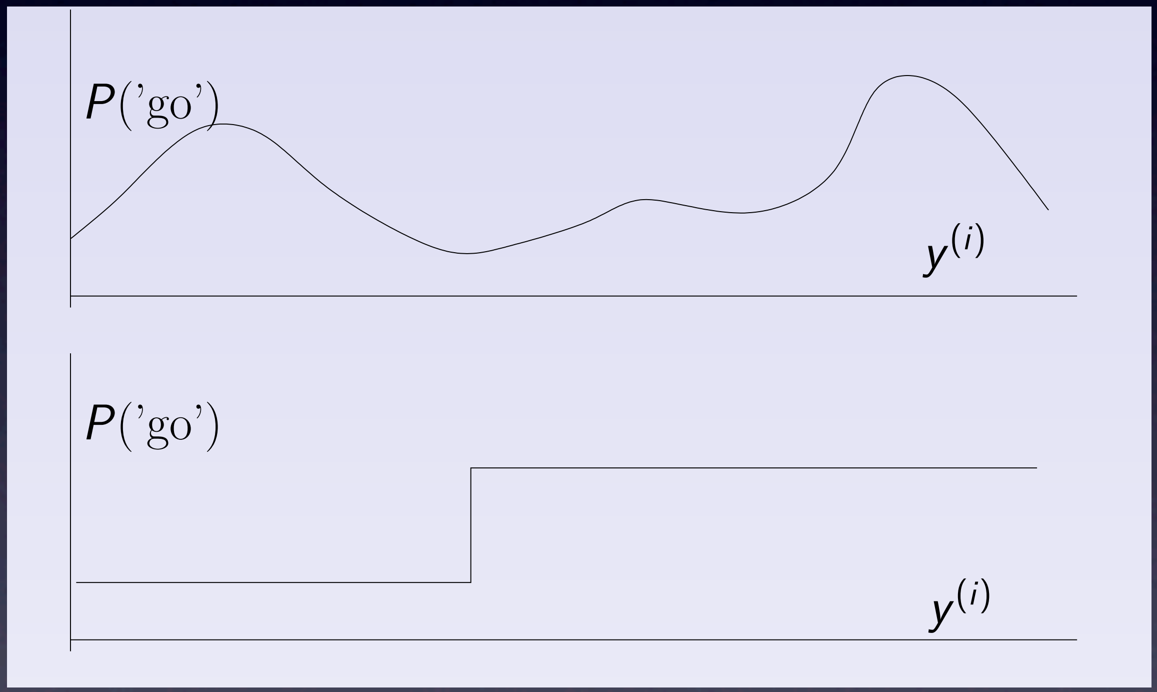
Under what conditions is optimal policy (Bayesian Nash equilibrium) a threshold?

BNE $\mathbb{E}_{Y^{-i}}[u_i(\pi_i^*(Y^i), \pi_{-i}^*(Y^{-i})|Y^i] \geq \mathbb{E}_{Y^{-i}}[u_i(\pi_i(Y^i), \pi_{-i}^*(Y^{-i})|Y^i]$



Under what conditions is optimal policy (Bayesian Nash equilibrium) a threshold?

BNE $\mathbb{E}_{Y^{-i}}[u_i(\pi_i^*(Y^i), \pi_{-i}^*(Y^{-i})|Y^i] \geq \mathbb{E}_{Y^{-i}}[u_i(\pi_i(Y^i), \pi_{-i}^*(Y^{-i})|Y^i]$

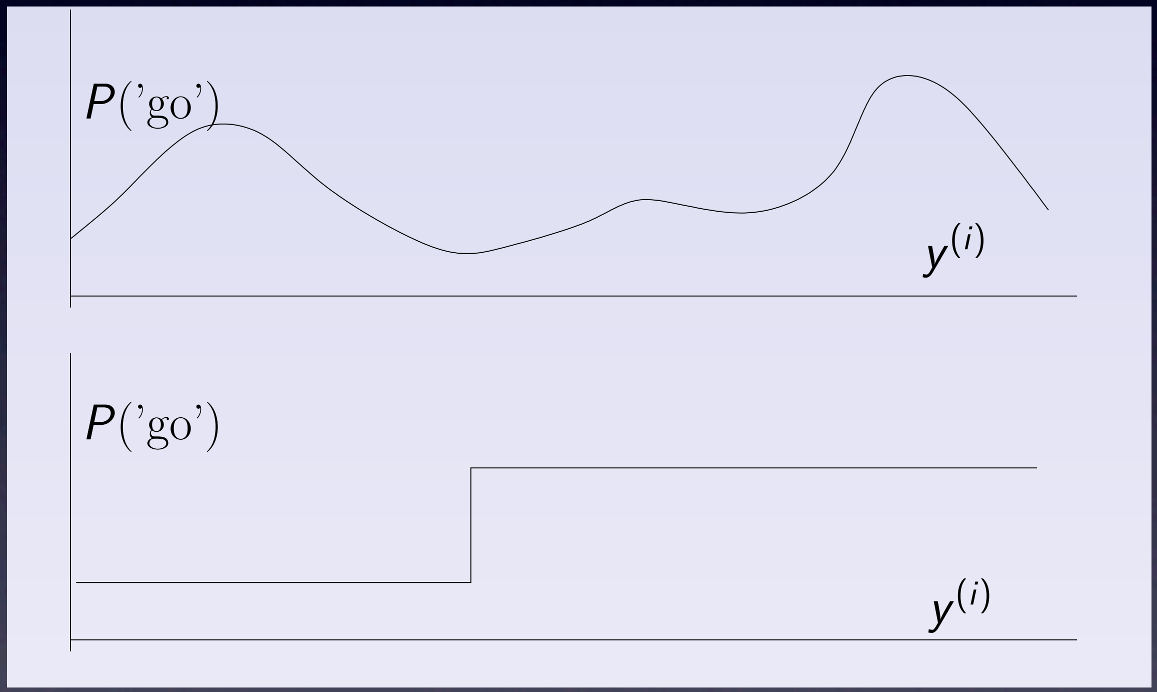


Why?

- Simple autonomous decision making by each agent
- Threshold can be estimated easily.

Under what conditions is optimal policy (Bayesian Nash equilibrium) a threshold?

BNE $\mathbb{E}_{Y^{-i}}[u_i(\pi_i^*(Y^i), \pi_{-i}^*(Y^{-i})|Y^i] \geq \mathbb{E}_{Y^{-i}}[u_i(\pi_i(Y^i), \pi_{-i}^*(Y^{-i})|Y^i]$



•IEEE Trans SP 2008 & 2009: Multiple classes of sensors; multiple night clubs;

•Gaussian noise; Proof via lattice programming and stochastic orders.

MAIN RESULT

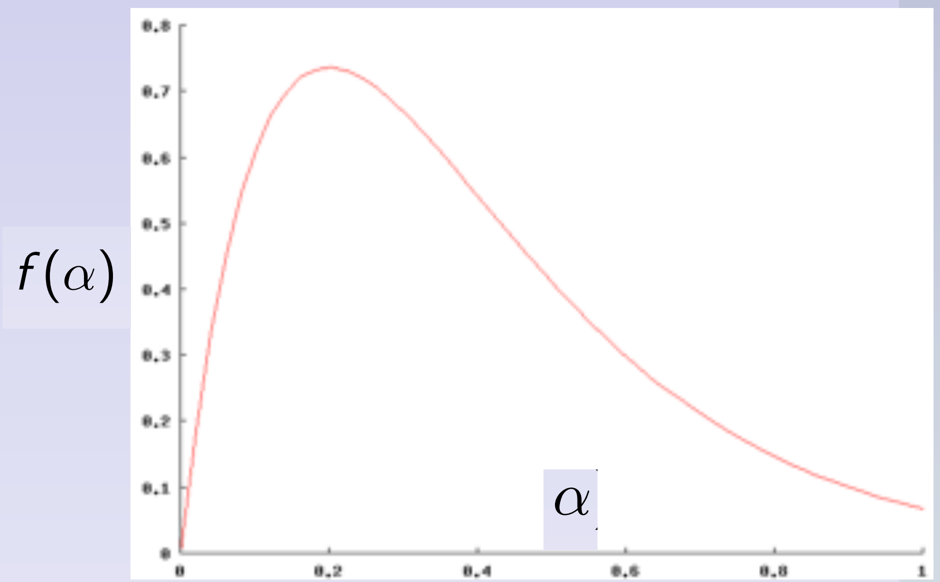
Result: If network congestion $\frac{df}{d\alpha} \geq -M$, there is unique threshold y^* such that Bayesian Nash Equilibrium of each sensor is a threshold policy:

$$\mu(Y) = \begin{cases} \text{active} & \text{if } Y > y^* \\ \text{sleep} & \text{if } Y \leq y^* \end{cases}$$

y^* satisfies $E[cX + f(\alpha(X))|Y = y^*] = 0$.

Rationality: $\alpha(x) = P(u = \text{active}|x)$

$$= \int_{y^*}^{\infty} p_{Y|X}(y|x)dy = 1 - \Phi_W(y^* - x).$$



Main result

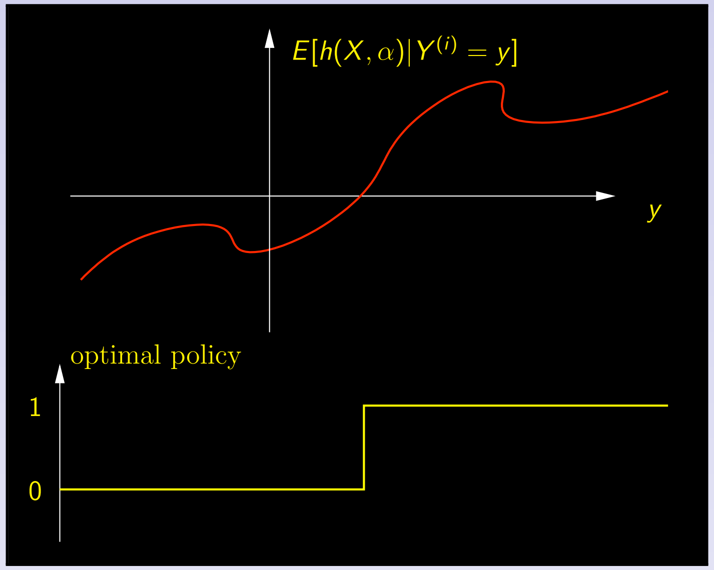
Result: If network congestion $\frac{df}{d\alpha} \geq -M$, there is unique threshold y^* such that Bayesian Nash Equilibrium of each sensor is a threshold policy:

$$\mu(Y) = \begin{cases} \text{active} & \text{if } Y > y^* \\ \text{sleep} & \text{if } Y \leq y^* \end{cases}$$

y^* satisfies $E[cX + f(\alpha(X))|Y = y^*] = 0$.

Rationality: $\alpha(x) = P(u = \text{active}|x)$

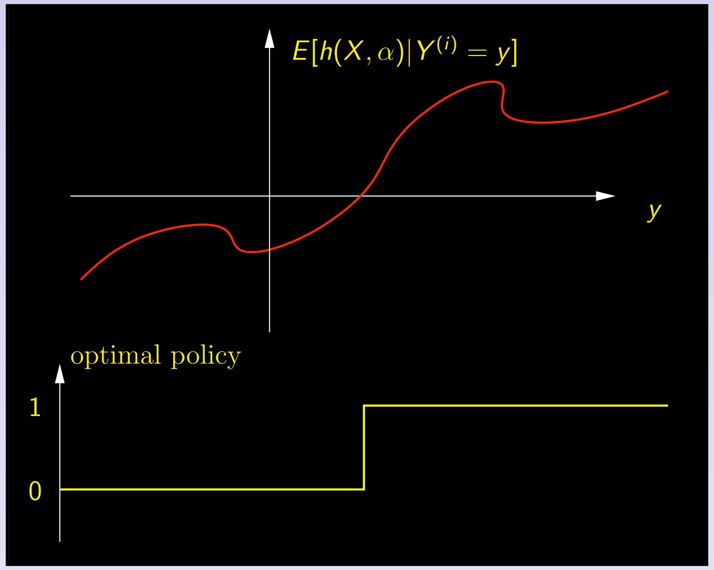
$$= \int_{y^*}^{\infty} p_{Y|X}(y|x)dy = 1 - \Phi_W(y^* - x).$$



MAIN RESULT

Result: If network congestion $\frac{df}{d\alpha} \geq -M$, there is unique threshold y^* such that Bayesian Nash Equilibrium of each sensor is a threshold policy:

$$\mu(Y) = \begin{cases} \text{active} & \text{if } Y > y^* \\ \text{sleep} & \text{if } Y \leq y^* \end{cases}$$



y^* satisfies $E[cX + f(\alpha(X))|Y = y^*] = 0$.

Rationality: $\alpha(x) = P(u = \text{active}|x)$

$$= \int_{y^*}^{\infty} p_{Y|X}(y|x)dy = 1 - \Phi_W(y^* - x).$$

Existence: Glicksberg fixed point theorem

MAIN RESULT

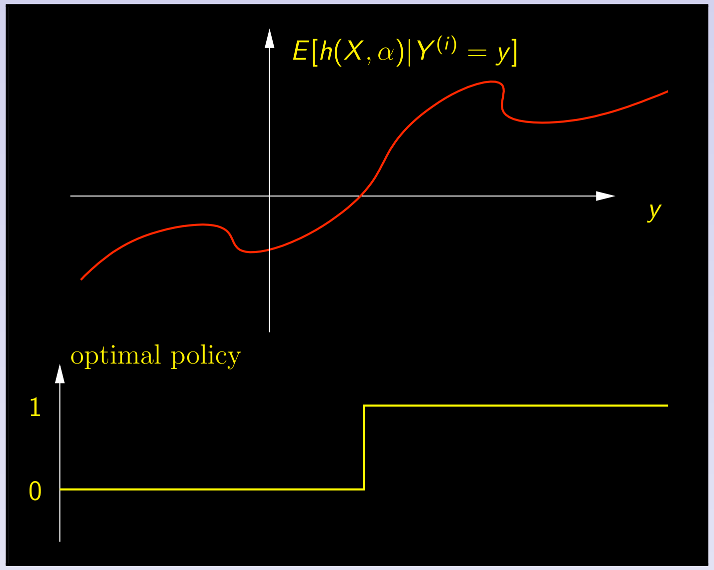
Result: If network congestion $\frac{df}{d\alpha} \geq -M$, there is unique threshold y^* such that Bayesian Nash Equilibrium of each sensor is a threshold policy:

$$\mu(Y) = \begin{cases} \text{active} & \text{if } Y > y^* \\ \text{sleep} & \text{if } Y \leq y^* \end{cases}$$

y^* satisfies $E[cX + f(\alpha(X))|Y = y^*] = 0$.

Rationality: $\alpha(x) = P(u = \text{active}|x)$

$$= \int_{y^*}^{\infty} p_{Y|X}(y|x)dy = 1 - \Phi_W(y^* - x).$$

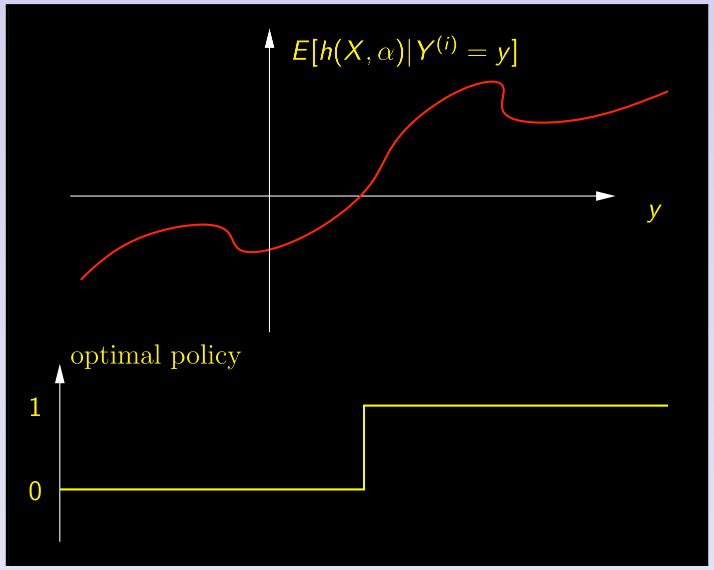


When is Nash $\mu^{(i)}(Y) = \arg\max\{\underbrace{0}_{(\text{sleep})\ u=1}, \underbrace{E[cX + f(\alpha(X)|Y]}_{(\text{active})\ u=2}\} \uparrow$ in Y?

MAIN RESULT

Result: If network congestion $\frac{df}{d\alpha} \geq -M$, there is unique threshold y^* such that Bayesian Nash Equilibrium of each sensor is a threshold policy:

$$\mu(Y) = \begin{cases} \text{active} & \text{if } Y > y^* \\ \text{sleep} & \text{if } Y \leq y^* \end{cases}$$



y^* satisfies $E[cX + f(\alpha(X))|Y = y^*] = 0$.

Rationality: $\alpha(x) = P(u = \text{active}|x)$

$$= \int_{y^*}^{\infty} p_{Y|X}(y|x)dy = 1 - \Phi_W(y^* - x).$$

When is Nash $\mu^{(i)}(Y) = \arg\max\{\underbrace{0}_{(\text{sleep})\ u=1}, \underbrace{E[cX + f(\alpha(X)|Y]}_{(\text{active})\ u=2}\} \uparrow$ in Y?

Monotone comparative statics:

Under what conditions does $\mu(y) = \arg\max_u \{H(y,u)\} \uparrow y$?

- Single crossing condition: Milgrom & Shannon, Econometrica, 1994
- Supermodularity: Topkis (1978,1998). $H(y,2) - H(y,1) \uparrow y$

(ii) Stochastic dominance: Whitt [1984]: $p_{X|Y}(x|y_1) \underset{r}{\geq} p_{X|Y}(x|y_2)$ iff $p_W(y-x) \underset{r}{\geq} p_W(y-x') \implies$ conditions on noise density $p_W(w)$.

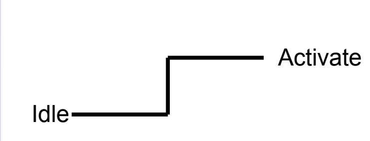
Gaussian, uniform,..

Thursday, July 14, 2011

Summary for Part 1

Sensors predicting behavior of other sensors in a global game yields:

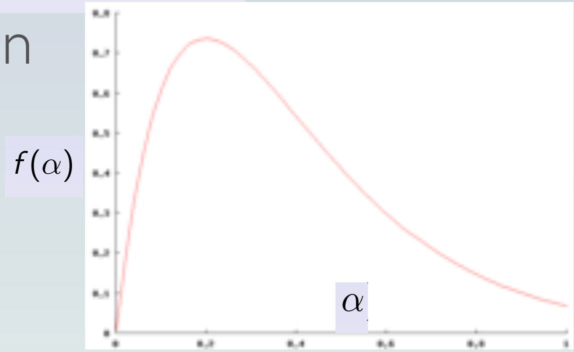
1. Simple intuitive optimal activation policy:
   $y < y^* \implies$ sleep, otherwise awaken.



2. **Each sensor can estimate threshold y^* indept of others.**
   $E[cX + f(\alpha(X)) \,|\, Y = y^*] = 0$ via simulation based stochastic approximation algorithm

3. **The network exhibits complex behavior.** e.g. high congestion

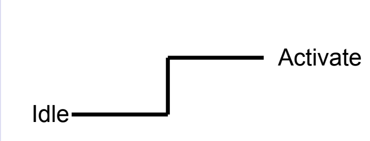
   No one goes there anymore ... it is always too crowded (Yogi Berra)



SUMMARY FOR PART 1

Sensors predicting behavior of other sensors in a global game yields:

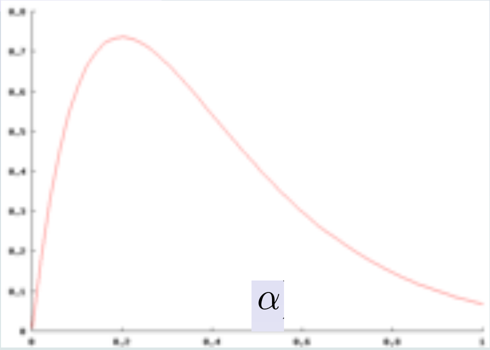
1. Simple intuitive optimal activation policy:
   $y < y^* \implies$ sleep, otherwise awaken.



2. **Each sensor can estimate threshold y^* indept of others.**
   $E[cX + f(\alpha(X)) \,|\, Y = y^*] = 0$ via simulation based stochastic approximation algorithm
3. **The network exhibits complex behavior.**

   No one goes there anymore ... it is always too crowded (Yogi Berra)

e.g. high congestion

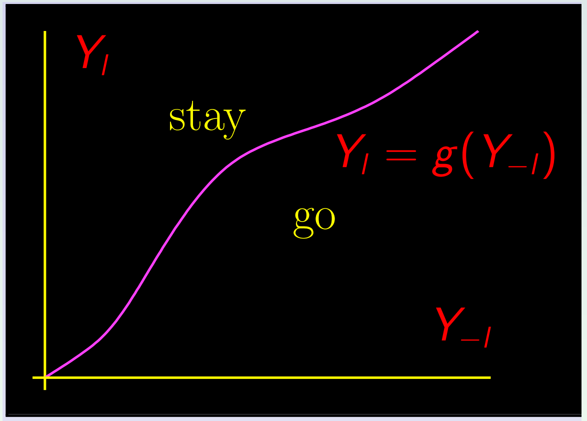


Multivariate Generalizations and Cognitive Radios

How to autonomously choose spectrum gap?

- Too many users in spectrum gap = excess interference.
- Too few users in spectrum gap = under-utilization.

[IEEE TSP 2009]: multi-variate opportunistic scheduling



OUTLINE

1. Global Games and Sensor Activation: Act simultaneously by predicting actions of other agents.
2. **Social Learning and Rational Herds:** Agents act sequentially.

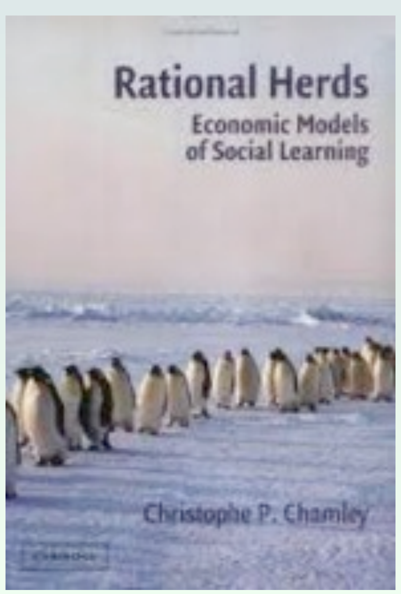
OUTLINE

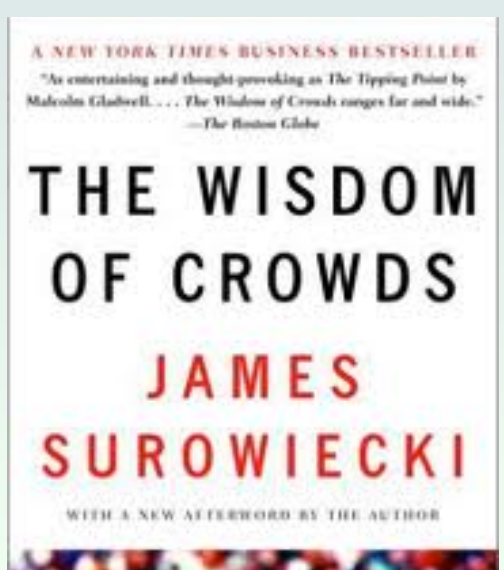
1. Global Games and Sensor Activation: Act simultaneously by predicting actions of other agents.

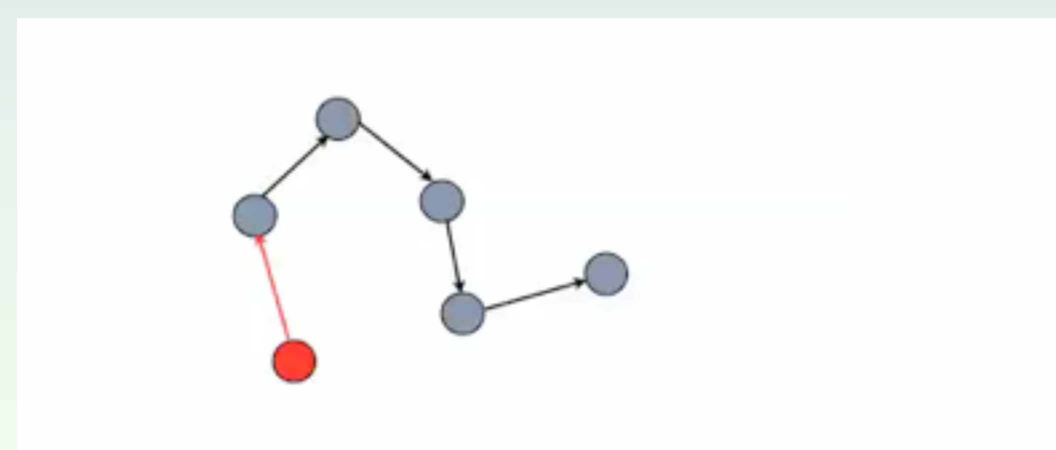
2. **Social Learning and Rational Herds:** Agents act sequentially.

Q1: How can agents learn by observing actions of other agents?

Herding - rational agents end up blindly following previous agents.







OUTLINE

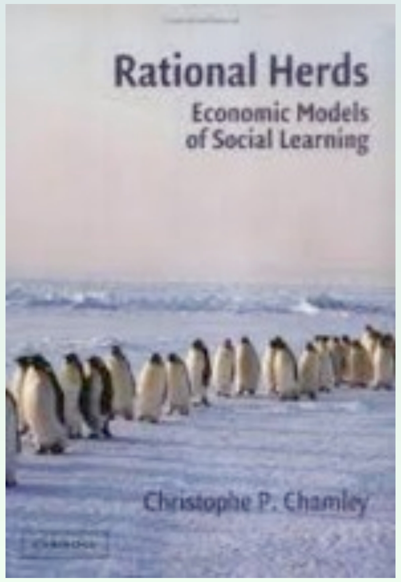
1. Global Games and Sensor Activation: Act simultaneously by predicting actions of other agents.

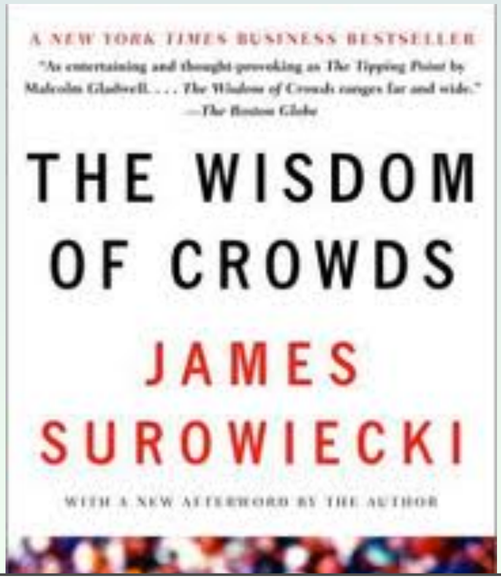
2. **Social Learning and Rational Herds:** Agents act sequentially.

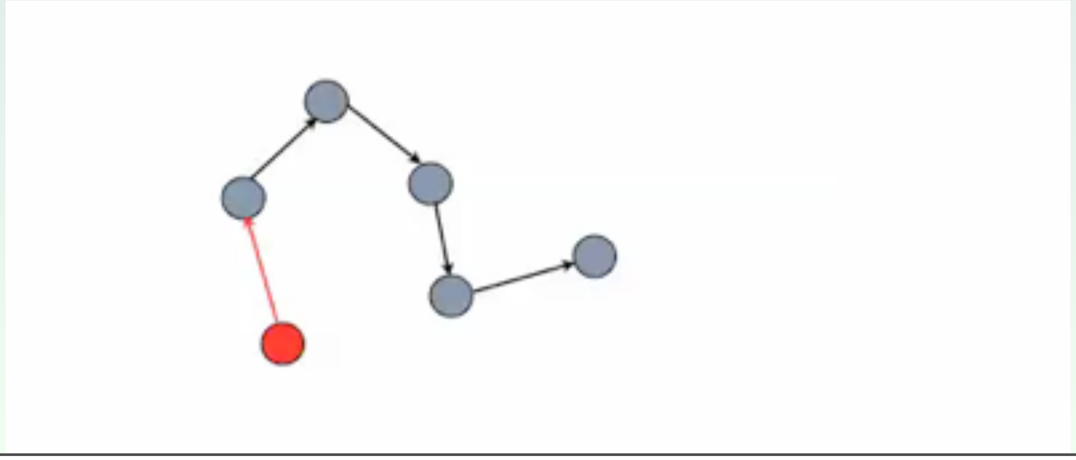
Q1: How can agents learn by observing actions of other agents?

Herding - rational agents end up blindly following previous agents.

In 1995, management gurus Treacy & Wiersema secretly bought 50,000 copies of their own book. Made NY times best seller list. How to cope with malicious agents?







OUTLINE

1. Global Games and Sensor Activation: Act simultaneously by predicting actions of other agents.

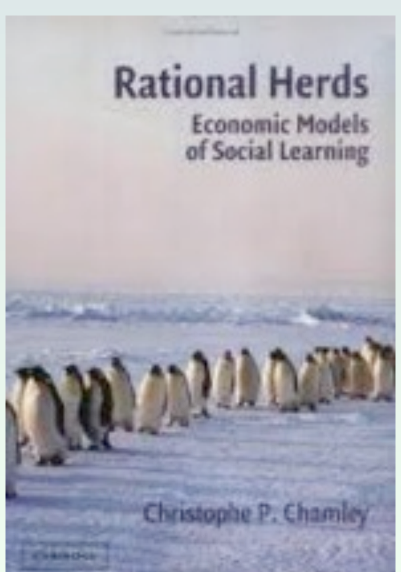
2. **Social Learning and Rational Herds:** Agents act sequentially.

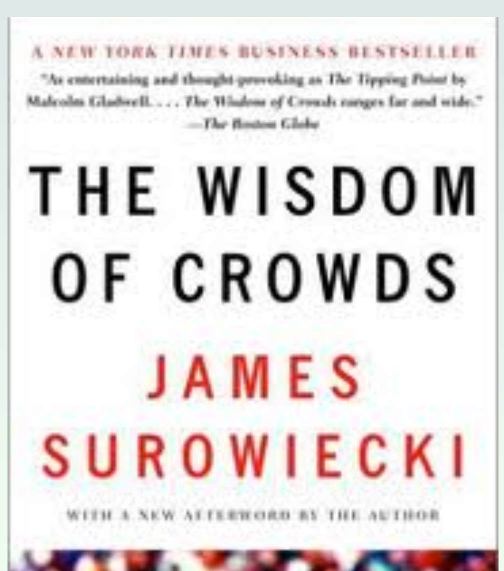
Q1: How can agents learn by observing actions of other agents?

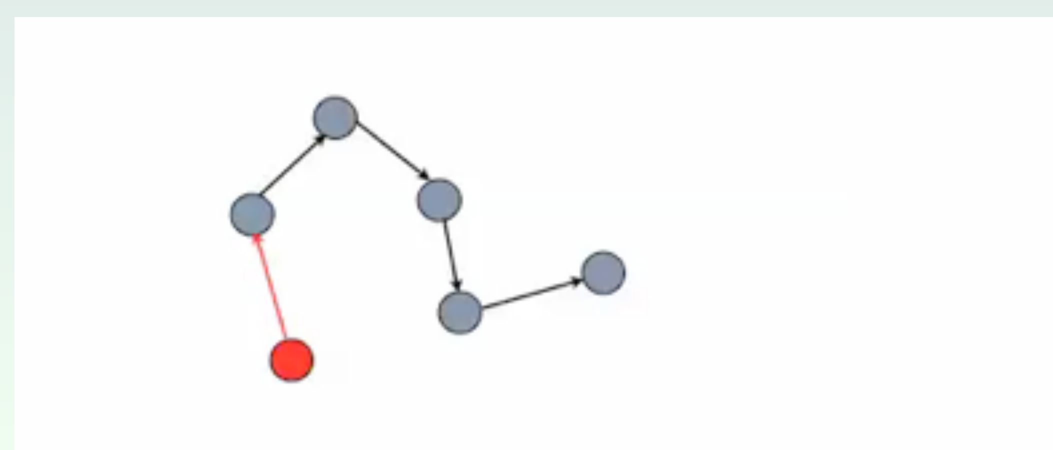
Herding - rational agents end up blindly following previous agents.

- When I see others taking umbrellas, I take an umbrella without checking the weather forecast. I assume their private info is accurate.

Chamley, 2004. Rational Herds







OUTLINE

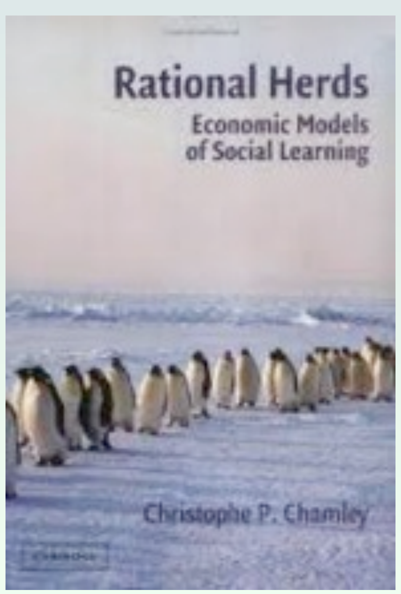
1. Global Games and Sensor Activation: Act simultaneously by predicting actions of other agents.

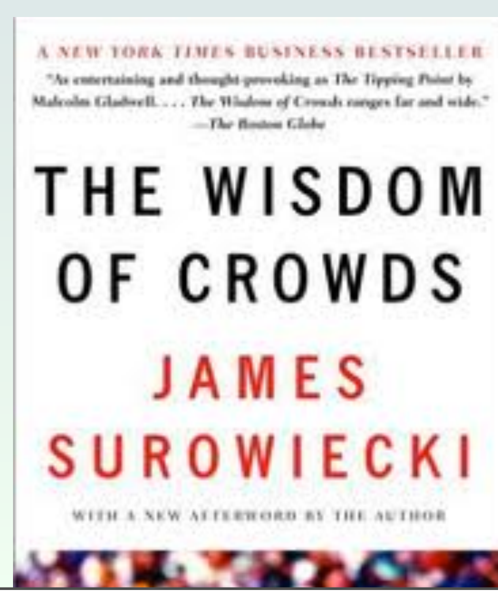
2. **Social Learning and Rational Herds:** Agents act sequentially.

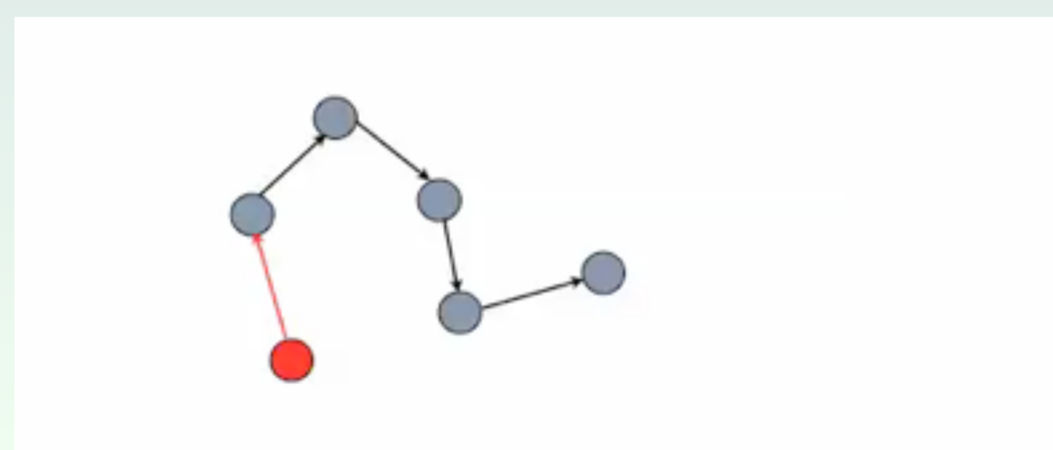
Q1: How can agents learn by observing actions of other agents?

Herding - rational agents end up blindly following previous agents.

How to build a sophisticated protocol that delays herding?







OUTLINE

1. Global Games and Sensor Activation: Act simultaneously by predicting actions of other agents.

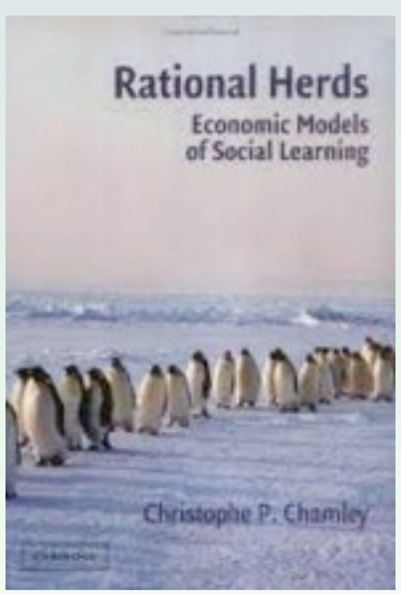
2. **Social Learning and Rational Herds:** Agents act sequentially.

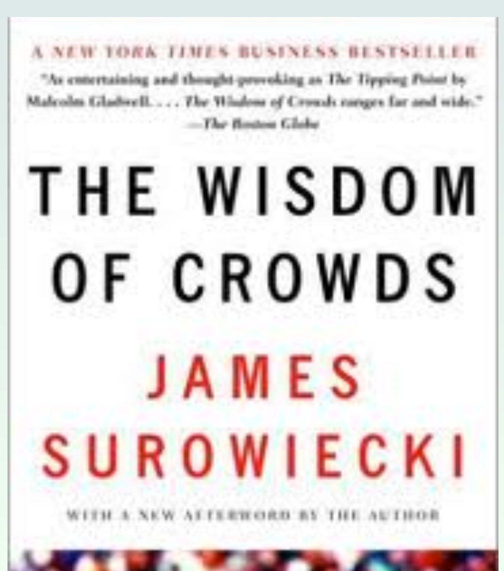
Q1: How can agents learn by observing actions of other agents?

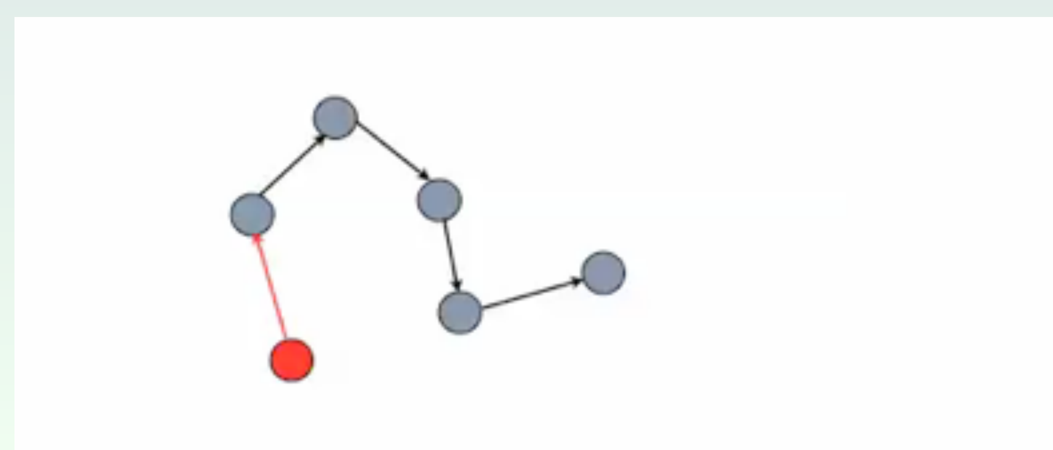
Herding - rational agents end up blindly following previous agents.

How to build a sophisticated protocol that delays herding?

Q2: How do local decisions affect global decisions? Social learning with quickest time change detection







Q1: How do agents learn from decisions of other agents?

Restaurant problem.

SOCIAL LEARNING PROTOCOL

SOCIAL LEARNING PROTOCOL

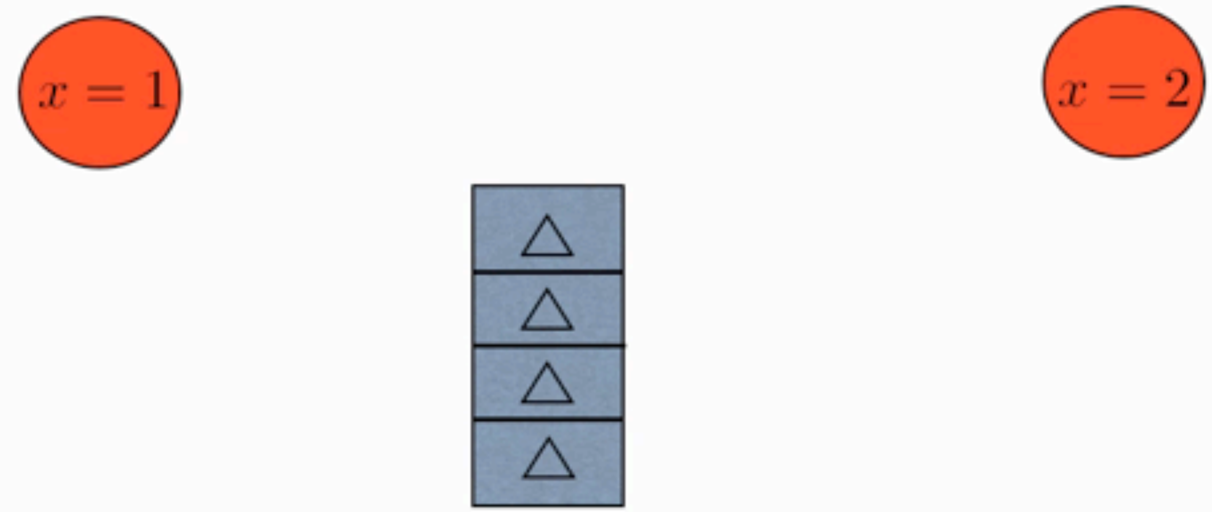
The Tale of Two Restaurants

SOCIAL LEARNING PROTOCOL

Aim: Agents $k = 1, 2, \ldots$ act sequentially to estimate state $x \sim \pi_0$.

Protocol: Given public belief $\pi_{k-1} = P(x|a_1, \ldots, a_{k-1})$

- Agent k observes $y_k \sim p(y|x)$
- Takes action to greedily minimize cost $a_k = \operatorname{argmin}_a \mathbf{E}_{\pi_{k-1}, y_k}\{c(x, a)\}$
- Using action a_k, other agents update belief $\pi_k = P(x|a_1, \ldots, a_k)$



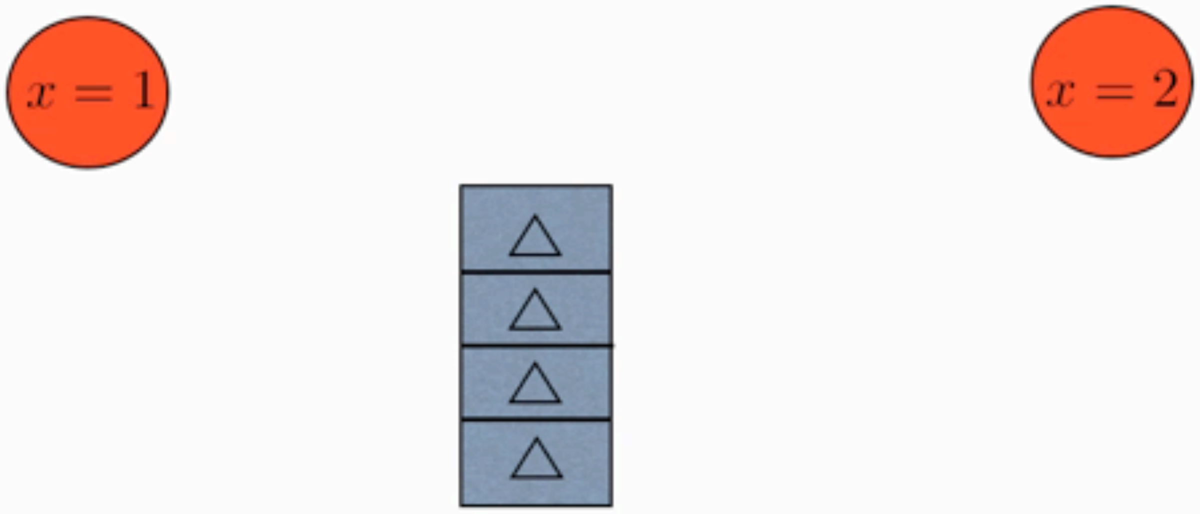
The Tale of Two Restaurants

SOCIAL LEARNING PROTOCOL

Aim: Agents $k = 1, 2, \ldots$ act sequentially to estimate state $x \sim \pi_0$.

Protocol: Given public belief $\pi_{k-1} = P(x|a_1, \ldots, a_{k-1})$

- Agent k observes $y_k \sim p(y|x)$
- Takes action to greedily minimize cost $a_k = \text{argmin}_a \mathbf{E}_{\pi_{k-1}, y_k}\{c(x, a)\}$
- Using action a_k, other agents update belief $\pi_k = P(x|a_1, \ldots, a_k)$



The Tale of Two Restaurants

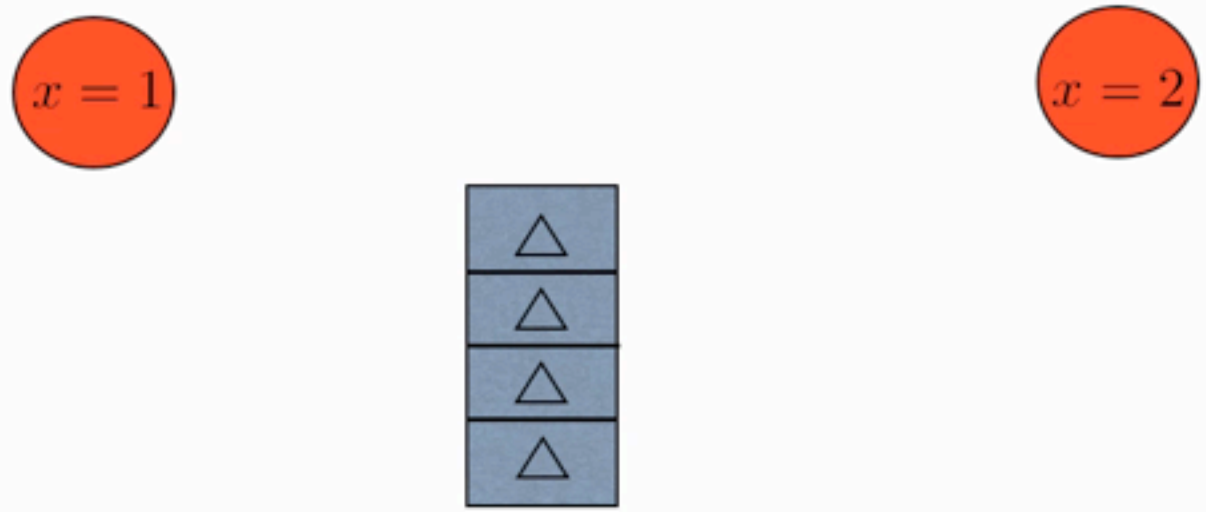
Key point: Agents learn from the actions of previous agents.

Social learning Protocol

Aim: Agents $k = 1, 2, \ldots$ act sequentially to estimate state $x \sim \pi_0$.

Protocol: Given public belief $\pi_{k-1} = P(x|a_1, \ldots, a_{k-1})$

- Agent k observes $y_k \sim p(y|x)$
- Takes action to greedily minimize cost $a_k = \text{argmin}_a \mathbf{E}_{\pi_{k-1}, y_k}\{c(x, a)\}$
- Using action a_k, other agents update belief $\pi_k = P(x|a_1, \ldots, a_k)$



The Tale of Two Restaurants

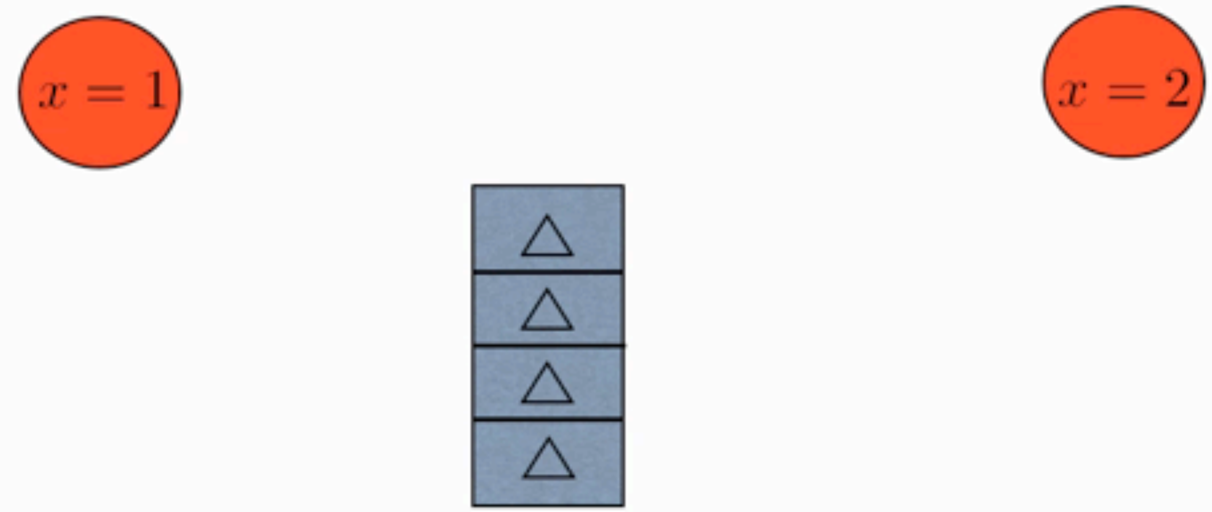
Theorem: [Bikchandani, J. Political Economy, 1992] Agents eventually herd, i.e. take the same action. (Social learning stops).

Social learning Protocol

Aim: Agents $k = 1, 2, \ldots$ act sequentially to estimate state $x \sim \pi_0$.

Protocol: Given public belief $\pi_{k-1} = P(x|a_1, \ldots, a_{k-1})$

- Agent k observes $y_k \sim p(y|x)$
- Takes action to greedily minimize cost $a_k = \text{argmin}_a \mathbf{E}_{\pi_{k-1}, y_k}\{c(x, a)\}$
- Using action a_k, other agents update belief $\pi_k = P(x|a_1, \ldots, a_k)$



The Tale of Two Restaurants

Theorem: [Bikchandani, J. Political Economy, 1992] Agents eventually herd, i.e. take the same action. (Social learning stops).

This is bad news for sensor networks where agents make local greedy decisions. We need a more sophisticated protocol.

E.g. pay agents to sit in restaurant 2

- Herding is caused by agents making greedy (capitalistic) local decisions.
- How to delay herding?

Q2: How to optimize Social learning?

Social Learning: Choose local decision greedily: $a_k = \min_a \mathbb{E}_{\pi_{k-1},y_k}\{c(x,a)\}$.
Results in herding. Posterior $\pi_k = P(x|a_1,\ldots,a_k)$ freezes.

Q2: How to optimize Social learning?

Social Learning: Choose local decision greedily: $a_k = \min_a \mathbb{E}_{\pi_{k-1},y_k}\{c(x,a)\}$.
Results in herding. Posterior $\pi_k = P(x|a_1,\ldots,a_k)$ freezes.

Socialistic Learning: [More Sophisticated Protocol] To estimate $x \sim \pi_0$

I choose my local decision to sacrifice my local ultility so that my action provides useful information to subsequent agents

Benevolent agents choose local decision by minimizing *social welfare* cost:

Q2: How to optimize Social learning?

Social Learning: Choose local decision greedily: $a_k = \min_a \mathbb{E}_{\pi_{k-1},y_k}\{c(x,a)\}$.
Results in herding. Posterior $\pi_k = P(x|a_1,\ldots,a_k)$ freezes.

Socialistic Learning: [More Sophisticated Protocol] To estimate $x \sim \pi_0$

> *I choose my local decision to sacrifice my local ultility so that my action provides useful information to subsequent agents*

Benevolent agents choose local decision by minimizing *social welfare* cost:

$$a_k = \mu^*(\pi_{k-1}, y_k) \in \{y_k(\text{reveal}), a_{k-1}(\text{herd})\}, \quad \mu^* = \arg\min_\mu \mathbb{E}^\mu_{\pi_0}\{\sum_{k=1}^N c(x,a_k)\}$$

global decision

Q2: How to optimize Social learning?

Social Learning: Choose local decision greedily: $a_k = \min_a \mathbb{E}_{\pi_{k-1},y_k}\{c(x,a)\}$.
Results in herding. Posterior $\pi_k = P(x|a_1,\ldots,a_k)$ freezes.

Socialistic Learning: [More Sophisticated Protocol] To estimate $x \sim \pi_0$

> *I choose my local decision to sacrifice my local ultility so that my action provides useful information to subsequent agents*

Benevolent agents choose local decision by minimizing *social welfare* cost:

$$a_k = \mu^*(\pi_{k-1}, y_k) \in \{y_k(\text{reveal}), a_{k-1}(\text{herd})\}, \quad \mu^* = \arg\min_{\mu} \mathbb{E}^{\mu}_{\pi_0}\{\sum_{k=1}^{N} c(x,a_k)\}$$

Partially Observed Stochastic Control Problem.

Q2: How to optimize Social learning?

Social Learning: Choose local decision greedily: $a_k = \min_a \mathbb{E}_{\pi_{k-1},y_k}\{c(x,a)\}$.
Results in herding. Posterior $\pi_k = P(x|a_1,\ldots,a_k)$ freezes.

Socialistic Learning: [More Sophisticated Protocol] To estimate $x \sim \pi_0$

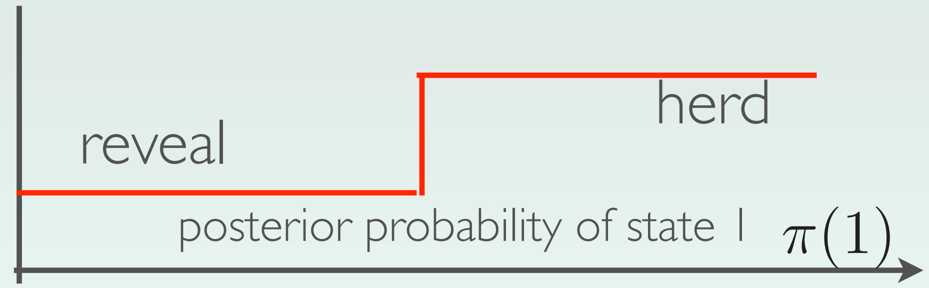
I choose my local decision to sacrifice my local ultility so that my action provides useful information to subsequent agents

Benevolent agents choose local decision by minimizing *social welfare* cost:

$$a_k = \mu^*(\pi_{k-1}, y_k) \in \{y_k(\text{reveal}), a_{k-1}(\text{herd})\}, \quad \mu^* = \arg\min_{\mu} \mathbb{E}^{\mu}_{\pi_0}\{\sum_{k=1}^{N} c(x, a_k)\}$$

Can show: [IEEE Trans. Info. Theory, 2011]

- Under supermodular assumptions global decision policy is threshold.



- Global decision policy: Initially socialistic then capitalistic.

Q3. How do local decisions in social learning affect global decisions in Quickest Time Change Detection?

Q3: How do local decisions affect Global Decisions?

Example: Multi-agent Quickest Time Change Detection

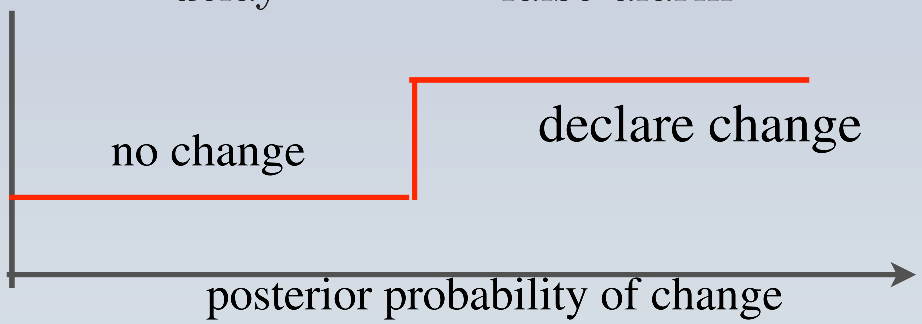
Q3: How do local decisions affect Global Decisions?

Example: Multi-agent Quickest Time Change Detection

Observations $y_k \sim \begin{cases} B_1(\cdot) & k \leq \tau^0 \\ B_2(\cdot) & k > \tau^0 \end{cases}$, where τ^0 = change time (usually geometric)

Aim: Compute time τ to annouce change: Minimize $\mathbb{E}^{\mu}_{\pi_0}\{\underbrace{d|\tau - \tau^0|^+}_{\text{delay}} + \underbrace{f\, I(\tau < \tau^0)}_{\text{false-alarm}}\}$

Classical: Given posterior $\pi_k = P(\text{change}|y_1, \ldots, y_k)$:
Optimal decision policy is threshold.



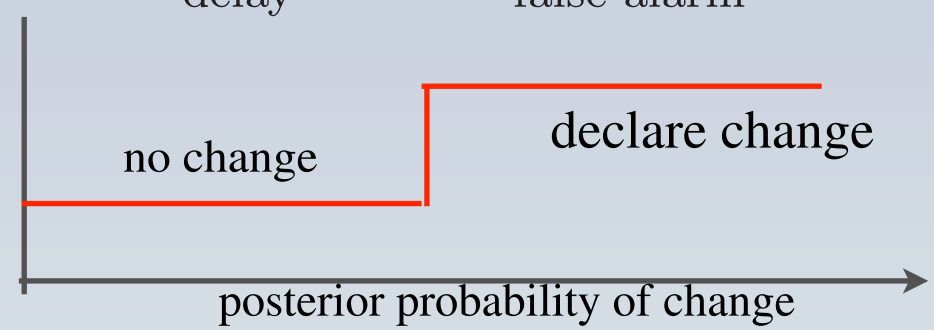
Q3: How do local decisions affect global decisions?

Example: Multi-agent Quickest Time Change Detection

Observations $y_k \sim \begin{cases} B_1(\cdot) & k \le \tau^0 \\ B_2(\cdot) & k > \tau^0 \end{cases}$, where τ^0 = change time (usually geometric)

Aim: Compute time τ to annouce change: Minimize $\mathbb{E}^{\mu}_{\pi_0}\{\underbrace{d|\tau - \tau^0|^+}_{\text{delay}} + \underbrace{f\, I(\tau < \tau^0)}_{\text{false-alarm}}\}$

Classical: Given posterior $\pi_k = P(\text{change}|y_1, \dots, y_k)$:
Optimal decision policy is threshold.



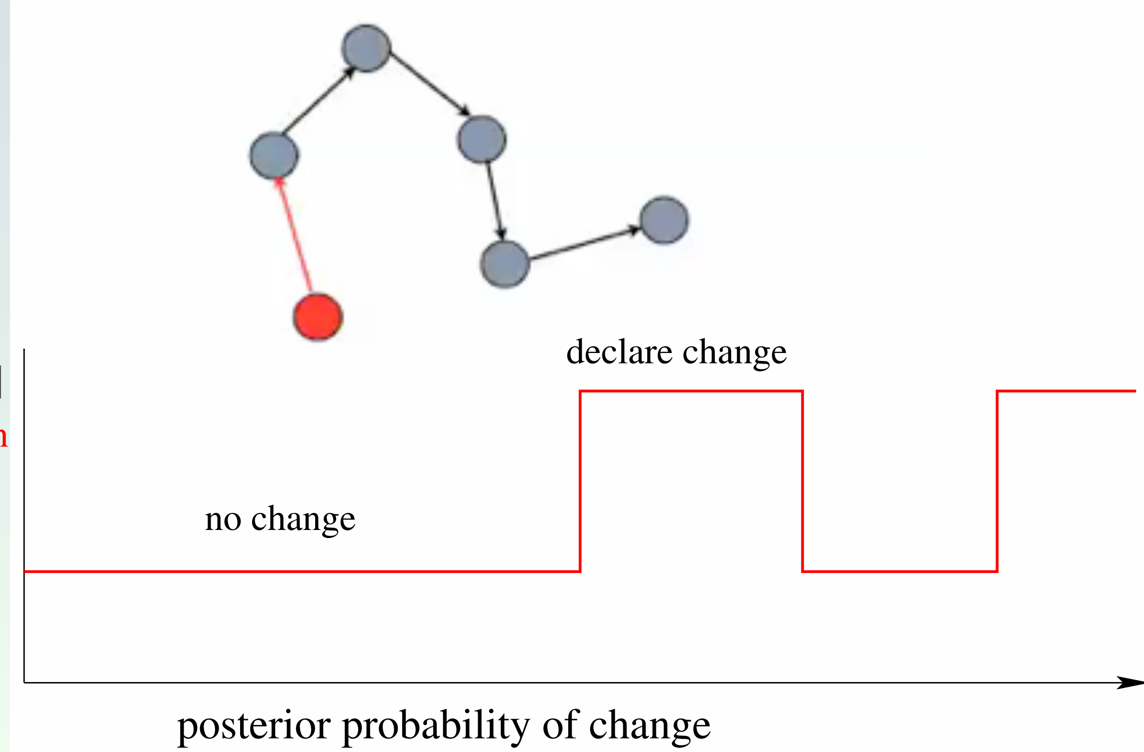
Social Learning: $\pi_{k-1} = P(\text{change}|a_1, \dots, a_{k-1})$

- Agent k receives observation y_k.
- Chooses local decision greedily

$$a_k = \min_a \mathbb{E}\{c(\text{state}, a)|a_1, \dots, a_{k-1}, y_k\}$$

- Broadcasts action a_k.

global
decision
policy



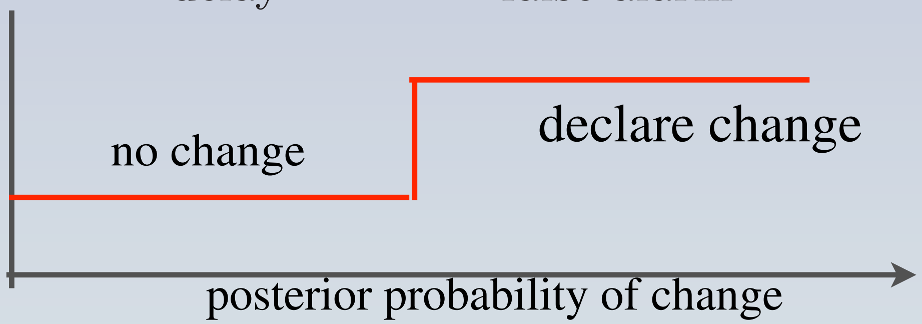
Q3: How do local decisions affect Global Decisions?

Example: Multi-agent Quickest Time Change Detection

Observations $y_k \sim \begin{cases} B_1(\cdot) & k \leq \tau^0 \\ B_2(\cdot) & k > \tau^0 \end{cases}$, where τ^0 = change time (usually geometric)

Aim: Compute time τ to annouce change: Minimize $\mathbb{E}^{\mu}_{\pi_0}\{\underbrace{d|\tau - \tau^0|^+}_{\text{delay}} + \underbrace{f\, I(\tau < \tau^0)}_{\text{false-alarm}}\}$

Classical: Given posterior $\pi_k = P(\text{change}|y_1, \ldots, y_k)$:
Optimal decision policy is threshold.



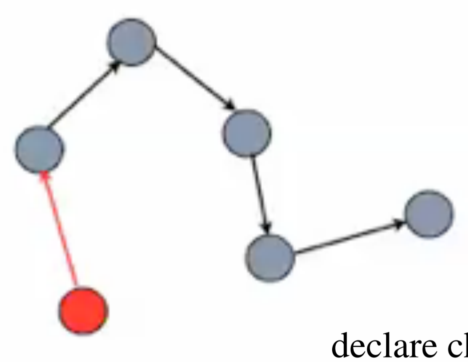
Social Learning: $\pi_{k-1} = P(\text{change}|a_1, \ldots, a_{k-1})$

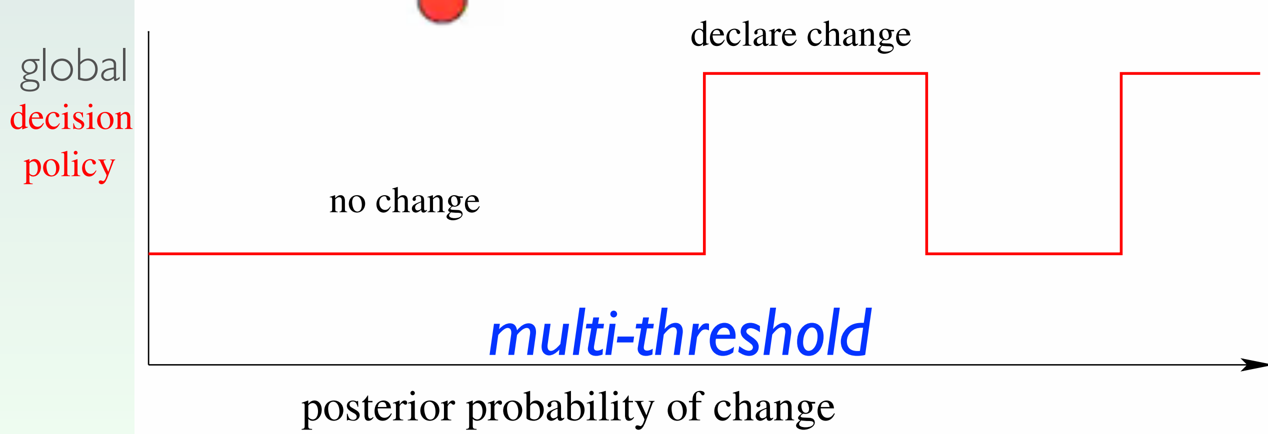
- Agent k receives observation y_k.
- Chooses local decision greedily

$$a_k = \min_a \mathbb{E}\{c(\text{state}, a)|a_1, \ldots, a_{k-1}, y_k\}$$

- Broadcasts action a_k.

When should global decision-maker declare change?





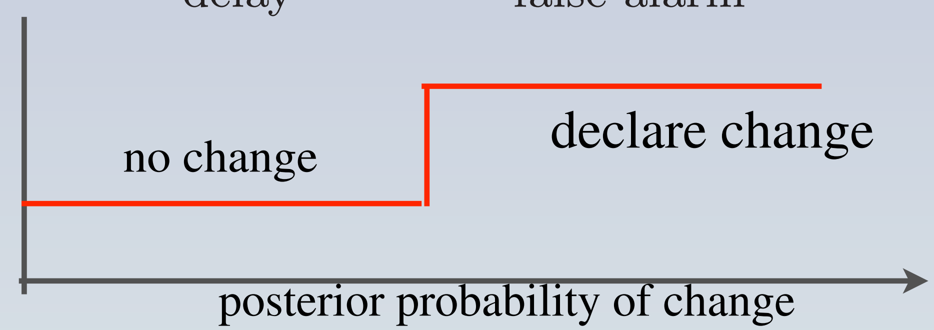
Q3: How do local decisions affect global decisions?

Example: Multi-agent Quickest Time Change Detection

Observations $y_k \sim \begin{cases} B_1(\cdot) & k \leq \tau^0 \\ B_2(\cdot) & k > \tau^0 \end{cases}$, where τ^0 = change time (usually geometric)

Aim: Compute time τ to annouce change: Minimize $\mathbb{E}^{\mu}_{\pi_0}\{\underbrace{d|\tau - \tau^0|^+}_{\text{delay}} + \underbrace{f\, I(\tau < \tau^0)}_{\text{false-alarm}}\}$

Classical: Given posterior $\pi_k = P(\text{change}|y_1, \ldots, y_k)$:
Optimal decision policy is threshold.

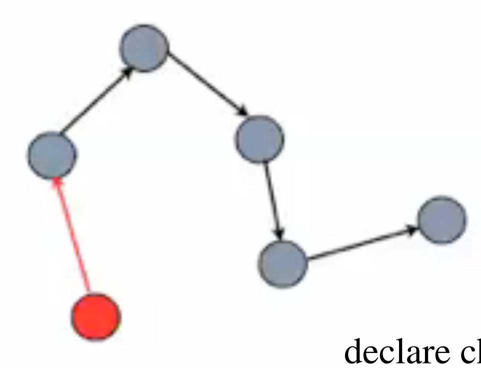


Social Learning: $\pi_{k-1} = P(\text{change}|a_1, \ldots, a_{k-1})$

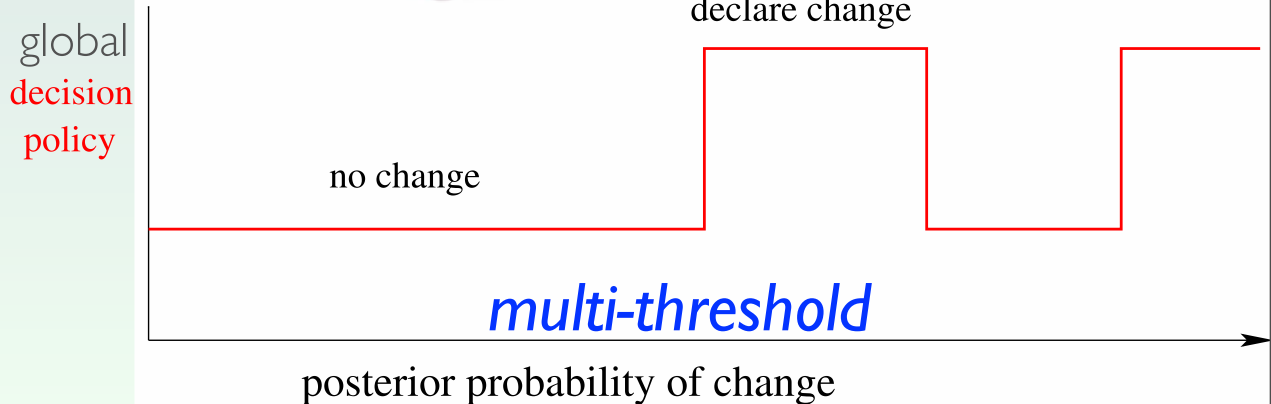
- Agent k receives observation y_k.
- Chooses local decision greedily

$$a_k = \min_a \mathbb{E}\{c(\text{state}, a)|a_1, \ldots, a_{k-1}, y_k\}$$

- Broadcasts action a_k.



When should global decision-maker declare change?



Non-standard Partially Observed Markov Decision Process

Likelihood: $P(a_k|\pi_{k-1}, x_k)$

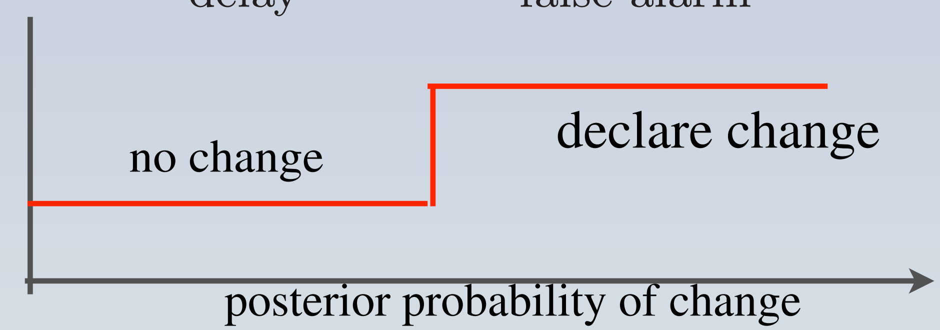
Q3: How do local decisions affect Global Decisions?

Example: Multi-agent Quickest Time Change Detection

Observations $y_k \sim \begin{cases} B_1(\cdot) & k \leq \tau^0 \\ B_2(\cdot) & k > \tau^0 \end{cases}$, where τ^0 = change time (usually geometric)

Aim: Compute time τ to annouce change: Minimize $\mathbb{E}^{\mu}_{\pi_0}\{\underbrace{d|\tau - \tau^0|^+}_{\text{delay}} + \underbrace{f\, I(\tau < \tau^0)}_{\text{false-alarm}}\}$

Classical: Given posterior $\pi_k = P(\text{change}|y_1, \ldots, y_k)$:
Optimal decision policy is threshold.



Social Learning: $\pi_{k-1} = P(\text{change}|a_1, \ldots, a_{k-1})$

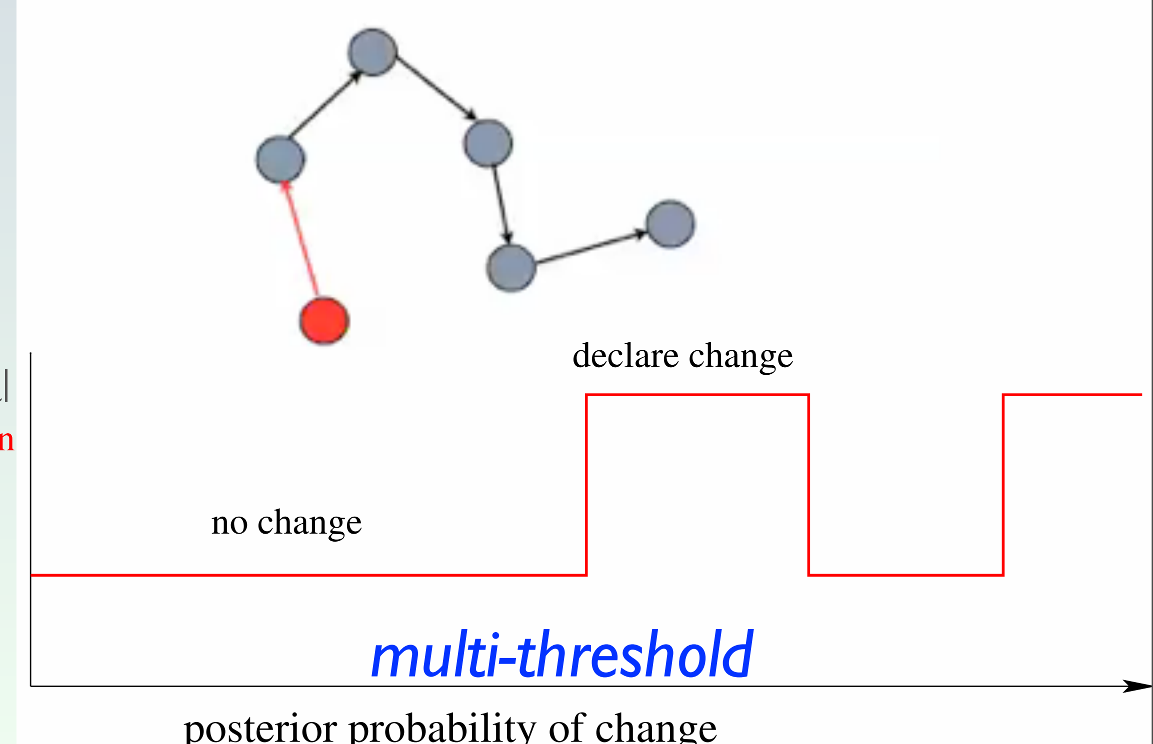
- Agent k receives observation y_k.
- Chooses local decision greedily

$$a_k = \min_a \mathbb{E}\{c(\text{state}, a)|a_1, \ldots, a_{k-1}, y_k\}$$

- Broadcasts action a_k.

When should global decision-maker declare change?

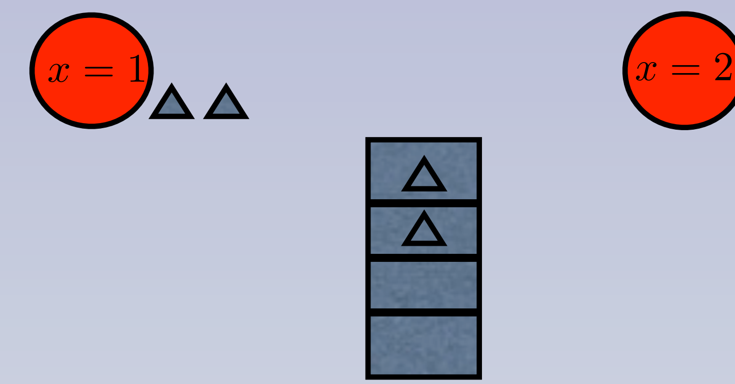
global decision policy



Summary: Global Decision making using local decisions is complex!

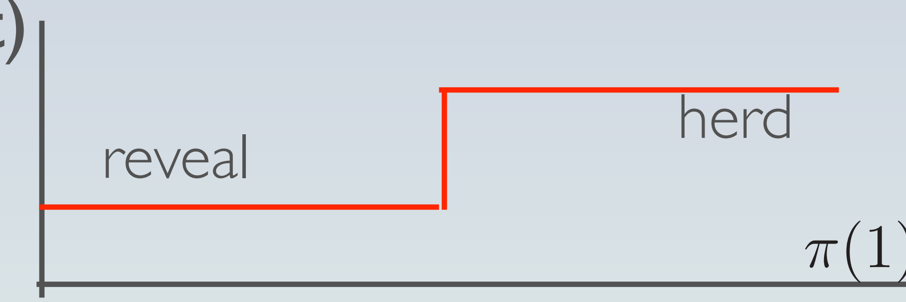
Summary for Part 2

Sequential Social Learning and Herding

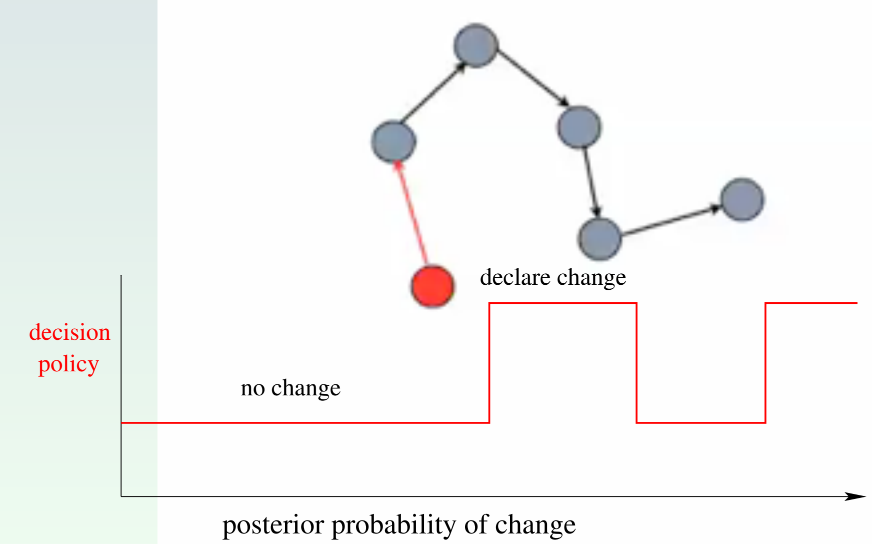


- **Model 1**. Greedy social learning from local decisions causes herding - information cascade. Individuals imitate actions of others.

- **Model 2**. Social learning with benevolent agents. Global decision specifies local decision (micromanagement) Stochastic control problem
  - Threshold policy is optimal (supermodularity)
  - Simulation-based stochastic optimization to estimate threshold.

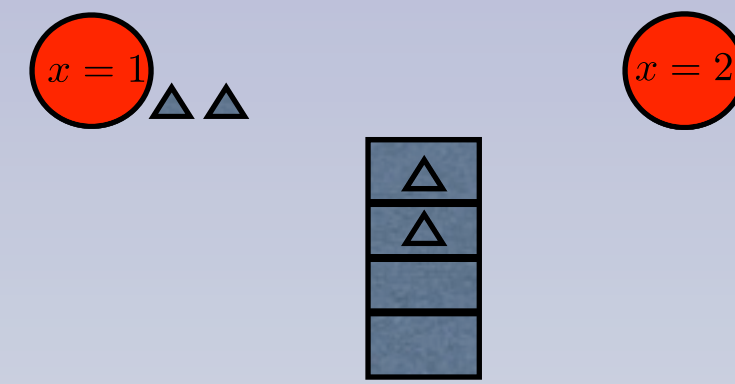


- **Model 3**. Making global decision (quickest time change detection) using local decisions has multi-threshold behavior.



Summary for Part 2

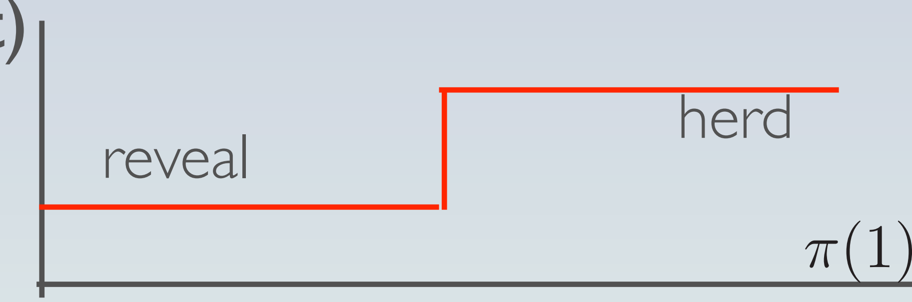
Sequential Social Learning and Herding



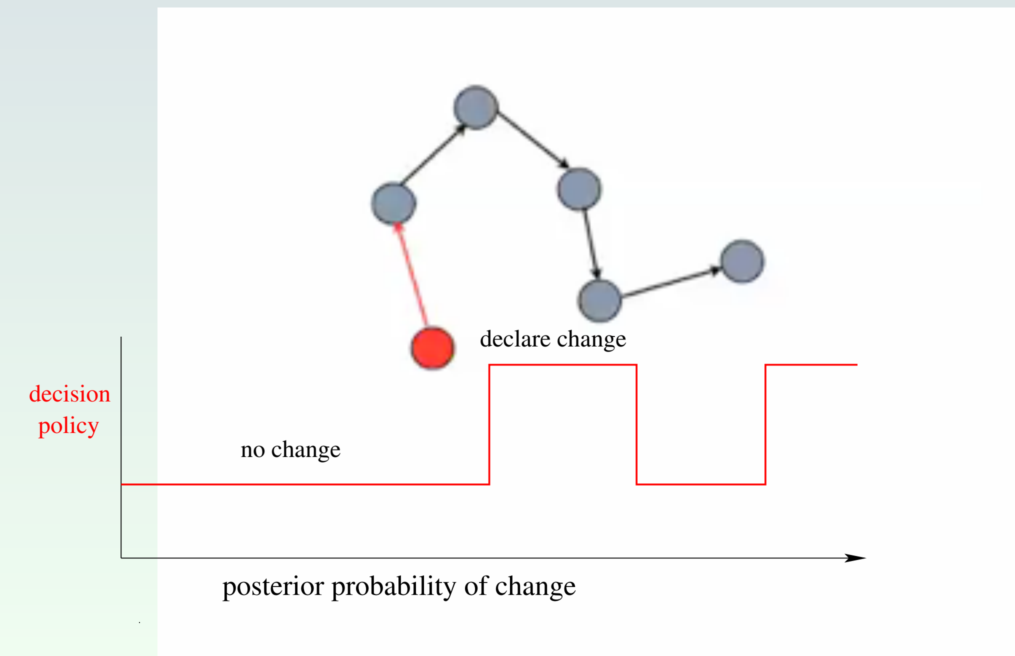
•**Model 1**. Greedy social learning from local decisions causes herding - information cascade. Individuals imitate actions of others.

•**Model 2**. Social learning with benevolent agents. Global decision specifies local decision (micromanagement) Stochastic control problem

- Threshold policy is optimal (supermodularity)
- Simulation-based stochastic optimization to estimate threshold.



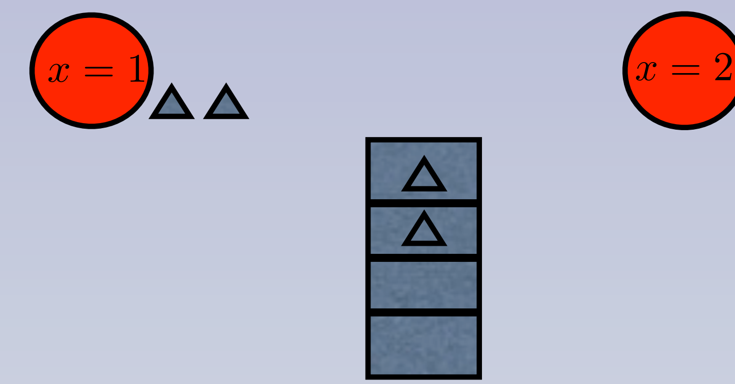
•**Model 3**. Making global decision (quickest time change detection) using local decisions has multi-threshold behavior.



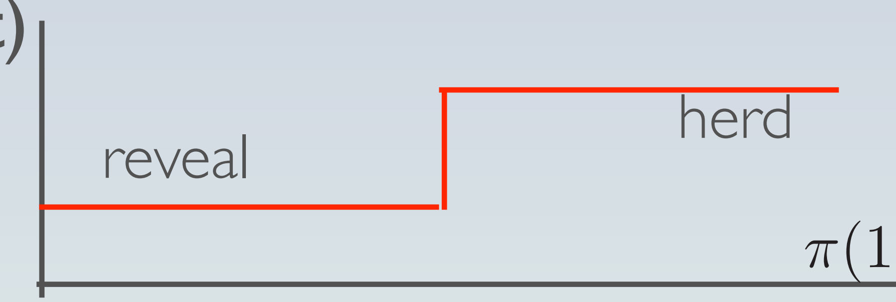
How to extend to more general communication graphs?

Summary for Part 2

Sequential Social Learning and Herding

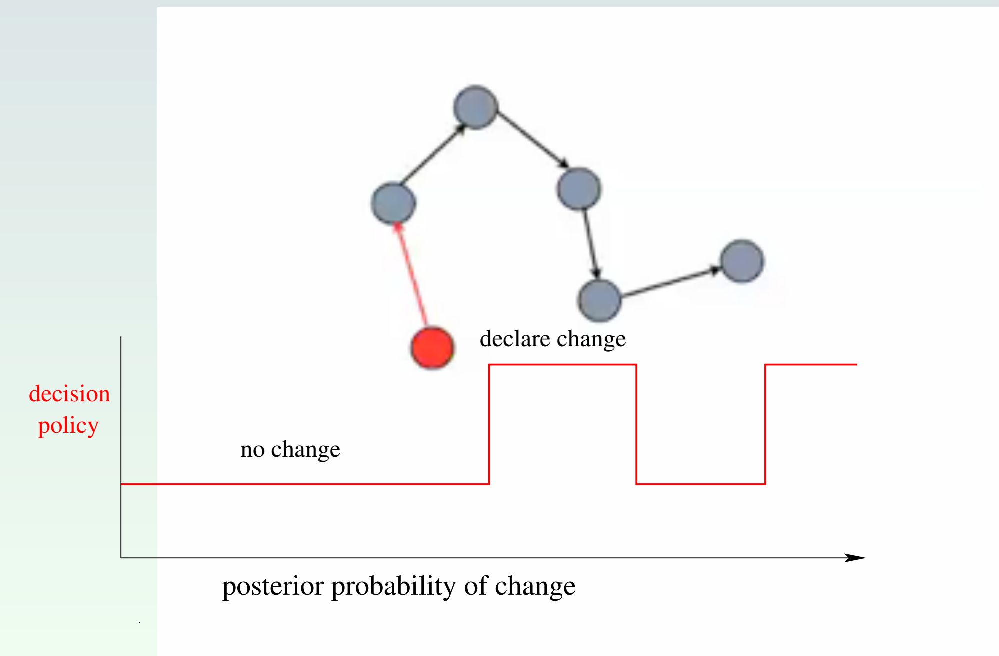


- **Model 1**. Greedy social learning from local decisions causes herding - information cascade. Individuals imitate actions of others.



- **Model 2**. Social learning with benevolent agents. Global decision specifies local decision (micromanagement) Stochastic control problem
  - Threshold policy is optimal (supermodularity)
  - Simulation-based stochastic optimization to estimate threshold.

- **Model 3**. Making global decision (quickest time change detection) using local decisions has multi-threshold behavior.



In which order should agents act?

Extension: Panel of Experts

High reputation

Good reputation

In which order to poll agents?
If senior agents talk first, they unduly affect junior agents.

No reputation

Low reputation

Extension: Panel of Experts

High reputation

Seniority Rule?

Good reputation

In which order to poll agents?
If senior agents talk first, they unduly affect junior agents.

No reputation

Low reputation

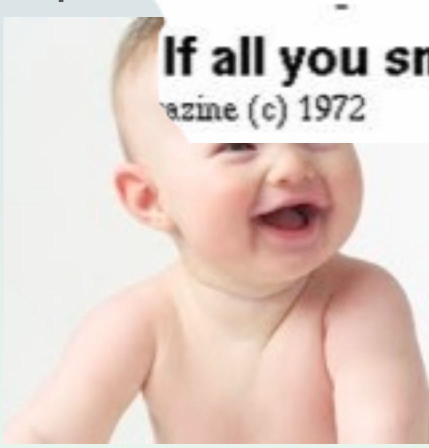
Extension: Panel of Experts

High reputation

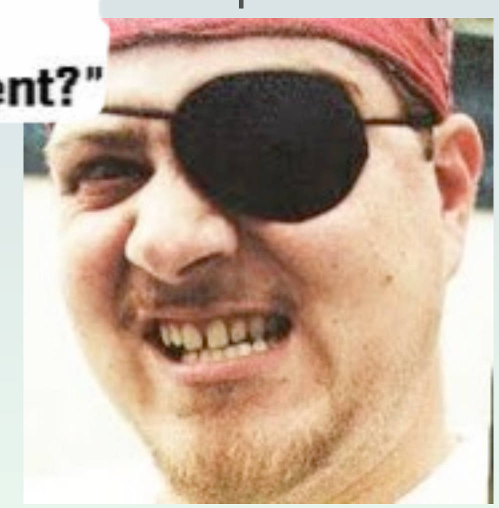
Seniority Rule?

Good reputation

No repu



reputation



Extension: Panel of Experts

High reputation

Seniority Rule?

Good reputation

For 94% of problems, the group's final answer was the first answer suggested, and people with dominant personalities tend to speak first and most forcefully...

Anderson & Kilduff, Berkeley Hass School, 2009.

No reputation

Low reputation

Extension: Panel of Experts

High reputation

Seniority Rule?

Good reputation

For 94% of problems, the group's final answer was the first answer suggested, and people with dominant personalities tend to speak first and most forcefully...

Anderson & Kilduff, Berkeley Hass School, 2009.

No reputation

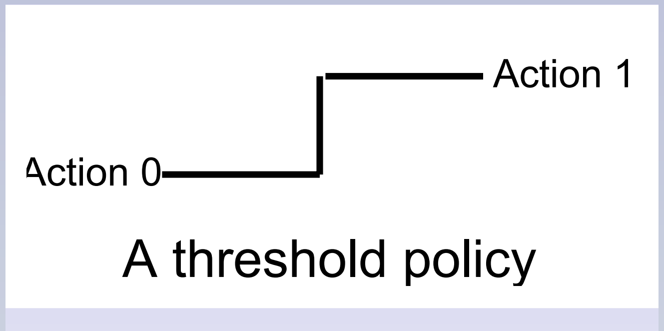
Low reputation

Ottaviani, Sorensen, 2001. *Information Aggregation in Debate: Who should speak first?*

OUTLINE

1. Global Games and Dynamic Sensor Activation:

Suppose each sensor deploys a *simple* algorithm.
What can you say about global performance?
Game Theory as an analysis tool



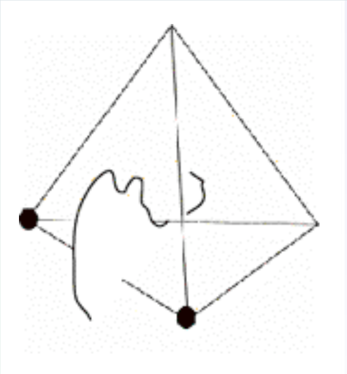
2. Social Learning and Rational Herds:

- How can agents learn from actions of other agents?
- How do local decisions affect global decisions?

Game Theory as a synthesis tool

3. Adaptive Filtering Games:

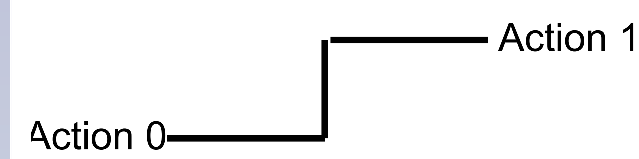
Each node deploys an adaptive filter to optimize its decision.
Can the global system achieve consensus in policy space?
Game theory in adaptive learning.



OUTLINE

1. **Global Games and Dynamic Sensor Activation**:

Suppose each sensor deploys a *simple* algorithm.
What can you say about global performance?
Game Theory as an analysis tool



A threshold policy

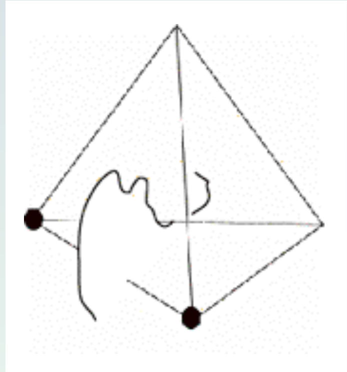
2. **Social Learning and Rational Herds:**

- How can agents learn from actions of other agents?
- How do local decisions affect global decisions?

Game Theory as a synthesis tool

3. **Adaptive Filtering Games:**

Each node deploys an adaptive filter to optimize its decision.
Can the global system achieve consensus in policy space?
Game theory in adaptive learning.



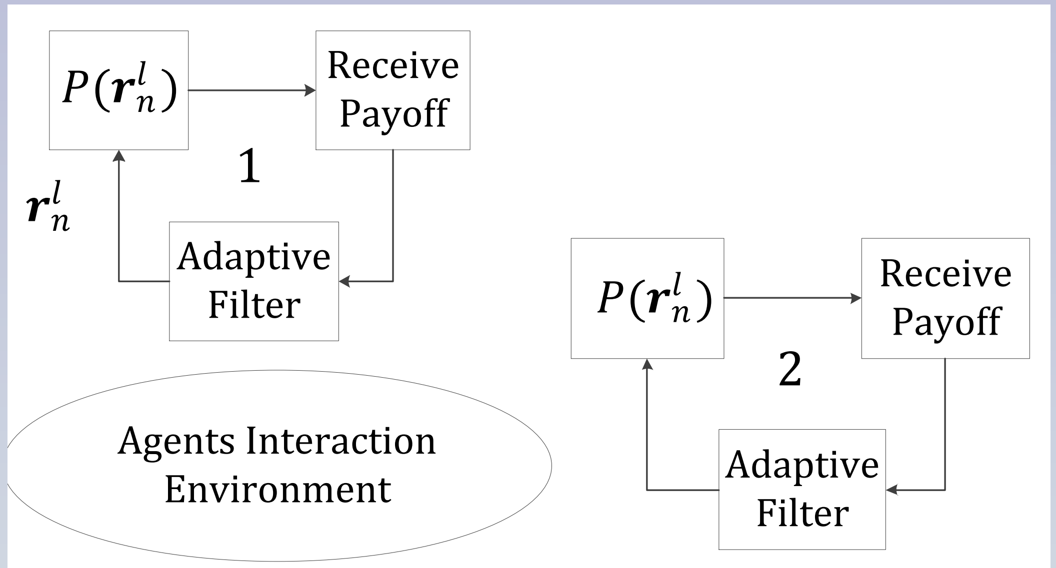
Unifying theme: local to global behavior for autonomous decision making

PART 3: ADAPTIVE FILTERING GAMES

Non-Bayesian - game theory in adaptive learning

- Each agent l:

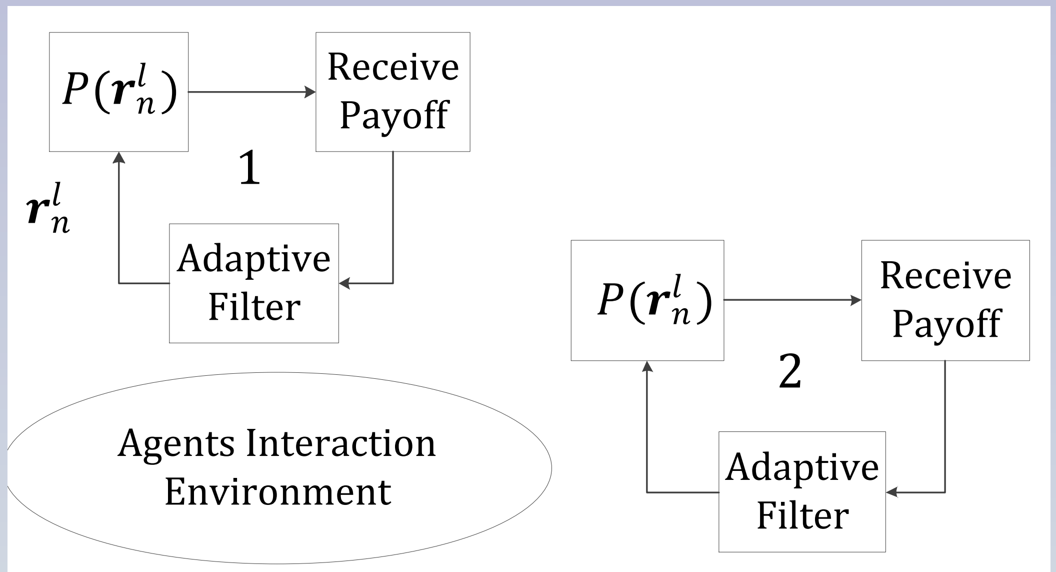
  1. Chooses action randomly $\{x^l_{n+1} = i\} \sim P(\mathbf{r}^l_n)$
  2. Receives stage utility $u^l(i, \mathbf{x}^{-l}_{n+1})$
  3. $\text{Regret}_{n+1}(i,j) = u^l(j, \mathbf{x}^{-l}_{n+1}) - u^l(i, \mathbf{x}^{-l}_{n+1})$
  4. Adaptive filter $\mathbf{r}^l_{n+1} = \mathbf{r}^l_n + \epsilon \left[\text{Regret}_{n+1} - \mathbf{r}^l_n\right]$



Part 3: Adaptive Filtering Games

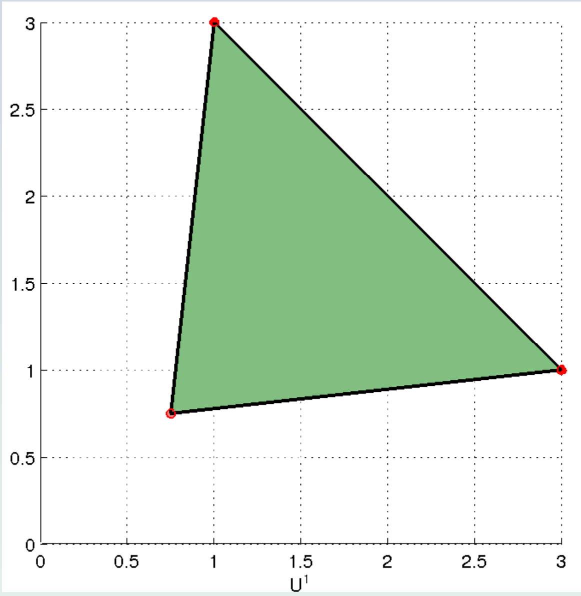
Non-Bayesian - game theory in adaptive learning

- Each agent l:
  1. Chooses action randomly $\{x^l_{n+1} = i\} \sim P(\mathbf{r}^l_n)$
  2. Receives stage utility $u^l(i, \mathbf{x}^{-l}_{n+1})$
  3. $\text{Regret}_{n+1}(i,j) = u^l(j, \mathbf{x}^{-l}_{n+1}) - u^l(i, \mathbf{x}^{-l}_{n+1})$
  4. Adaptive filter $\mathbf{r}^l_{n+1} = \mathbf{r}^l_n + \epsilon\left[\text{Regret}_{n+1} - \mathbf{r}^l_n\right]$



Main Result: **Consensus in decision space**

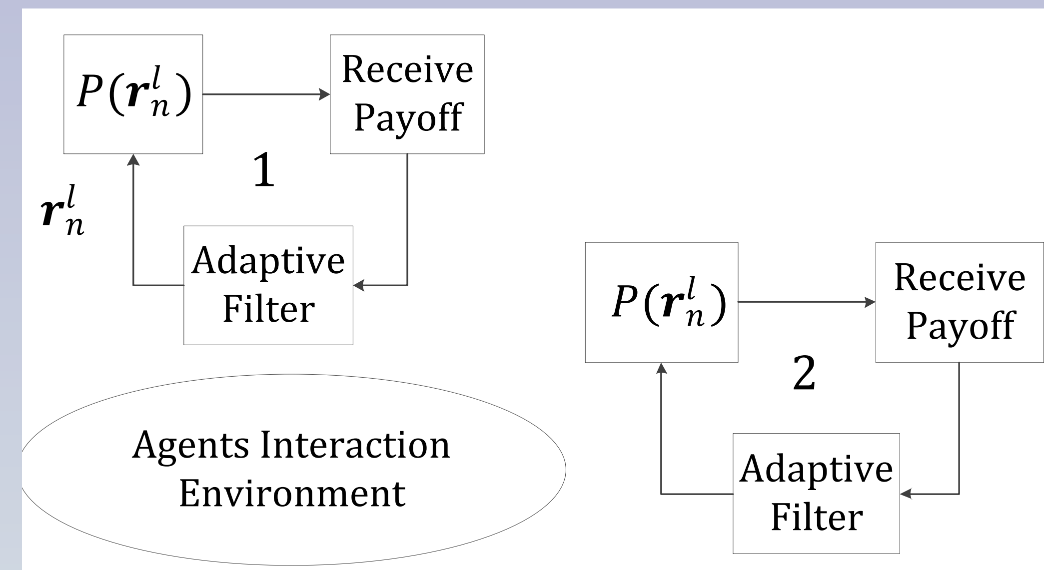
Can this simple local behavior lead to rational global behavior?
Global behavior converges to correlated equilibrium set (weakly or wp1).



PART 3: ADAPTIVE FILTERING GAMES

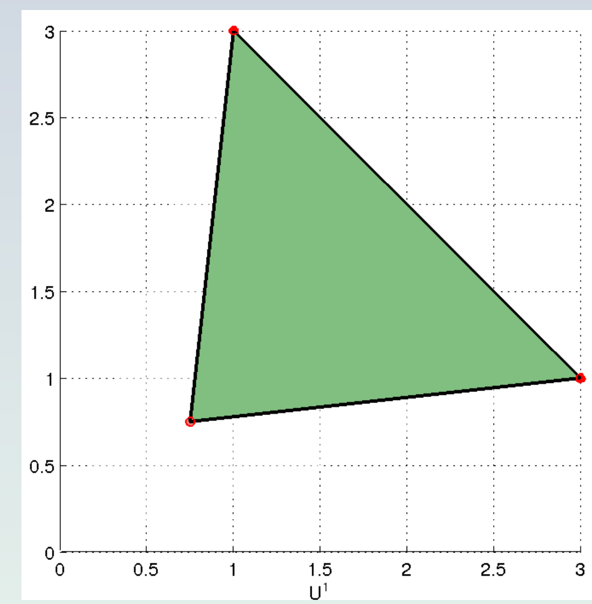
Non-Bayesian - game theory in adaptive learning

- Each agent l:
  1. Chooses action randomly $\{x^l_{n+1} = i\} \sim P(\mathbf{r}^l_n)$
  2. Receives stage utility $u^l(i, \mathbf{x}^{-l}_{n+1})$
  3. $\text{Regret}_{n+1}(i,j) = u^l(j, \mathbf{x}^{-l}_{n+1}) - u^l(i, \mathbf{x}^{-l}_{n+1})$
  4. Adaptive filter $\mathbf{r}^l_{n+1} = \mathbf{r}^l_n + \epsilon \left[\text{Regret}_{n+1} - \mathbf{r}^l_n\right]$



Main Result: **Consensus in decision space**

Can this simple local behavior lead to rational global behavior?
Global behavior converges to correlated equilibrium set (weakly or wp1).

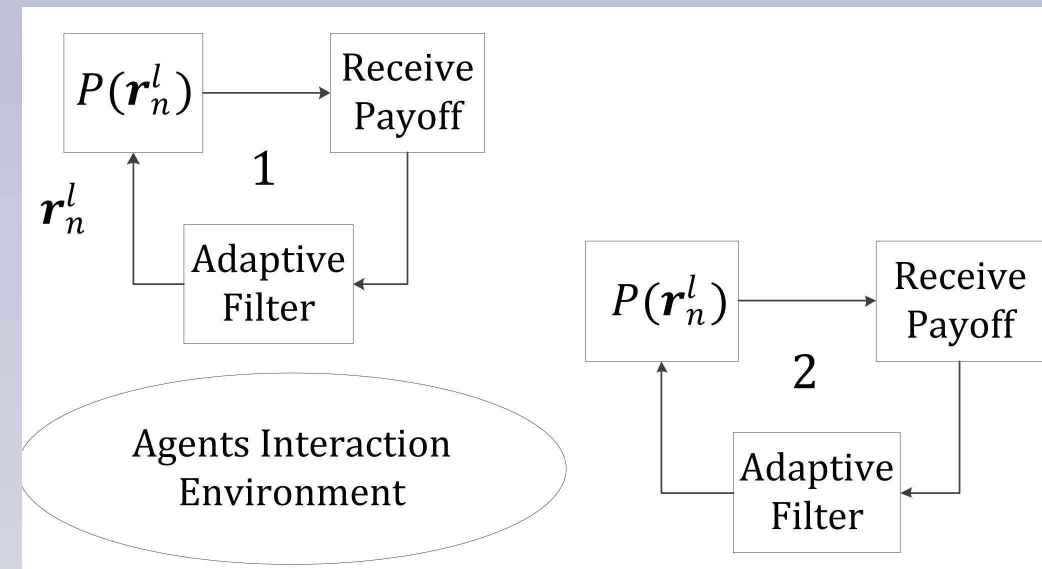


[1] Hart, S. and Mas-Collel, A (2000) A simple procedure leading to correlated equilibrium, *Econometrica*.
[2] Hart, S (2005) Adaptive Heuristics, Econometrica.

PART 3: ADAPTIVE FILTERING GAMES

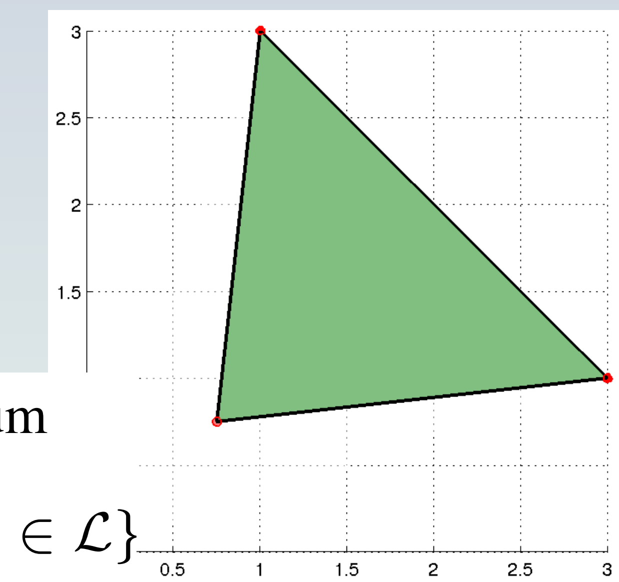
Non-Bayesian - game theory in adaptive learning

- Each agent l:
  1. Chooses action randomly $\{x^l_{n+1} = i\} \sim P(\mathbf{r}^l_n)$
  2. Receives stage utility $u^l(i, \mathbf{x}^{-l}_{n+1})$
  3. $\text{Regret}_{n+1}(i,j) = u^l(j, \mathbf{x}^{-l}_{n+1}) - u^l(i, \mathbf{x}^{-l}_{n+1})$
  4. Adaptive filter $\mathbf{r}^l_{n+1} = \mathbf{r}^l_n + \epsilon\left[\text{Regret}_{n+1} - \mathbf{r}^l_n\right]$



Main Result: **Consensus in decision space**

Can this simple local behavior lead to rational global behavior?
Global behavior converges to correlated equilibrium set (weakly or wp1).



Blackwell approachability

- CORRELATED EQUILIBRIUM (CE) [Aumann, 1987]: a generalization of Nash equilibrium

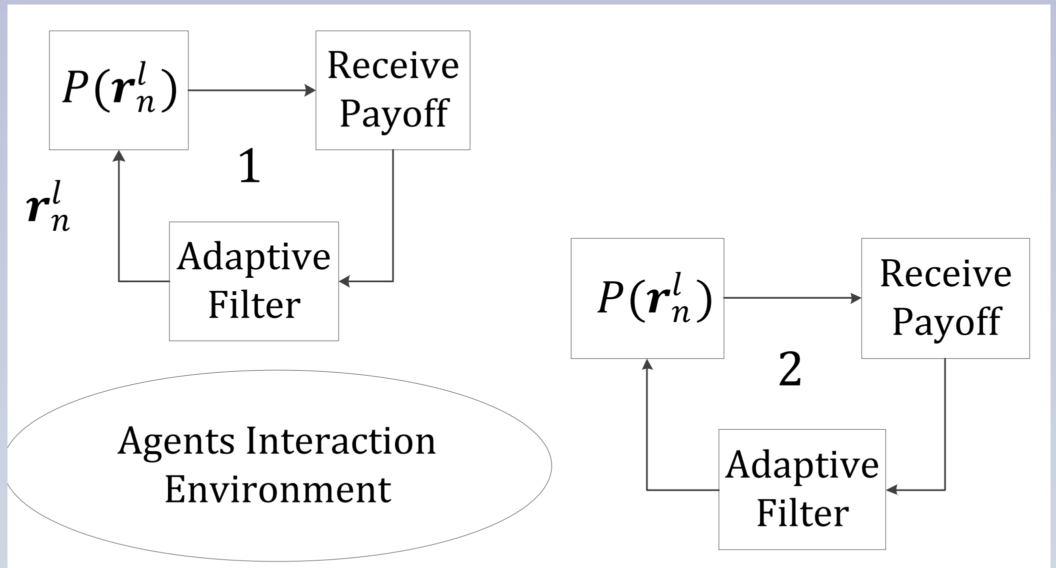
$$\mathcal{C}_e = \{\pi : \sum_{\mathbf{x}^{-l} \in A^{-l}} \pi^l(i, \mathbf{x}^{-l})[u^l(j, \mathbf{x}^{-l}) - u^l(i, \mathbf{x}^{-l})] \leq 0, \quad \forall i, j \in A^l, l \in \mathcal{L}\}$$

- *Why ?* 1) Correlation device: common history of actions, 2) Structural Simplicity: convex polytope, 3) Provably convergent learning algorithms

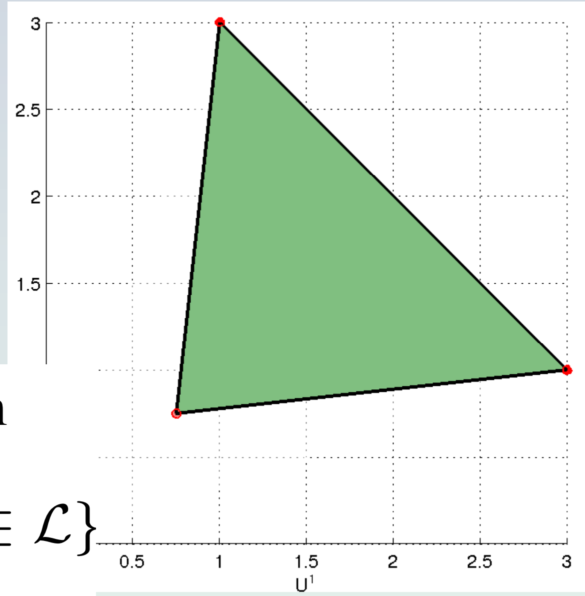
PART 3: ADAPTIVE FILTERING GAMES

Non-Bayesian - game theory in adaptive learning

- Each agent l:
  1. Chooses action randomly $\{x^l_{n+1} = i\} \sim P(\mathbf{r}^l_n)$
  2. Receives stage utility $u^l(i, \mathbf{x}^{-l}_{n+1})$
  3. $\text{Regret}_{n+1}(i,j) = u^l(j, \mathbf{x}^{-l}_{n+1}) - u^l(i, \mathbf{x}^{-l}_{n+1})$
  4. Adaptive filter $\mathbf{r}^l_{n+1} = \mathbf{r}^l_n + \epsilon \left[\text{Regret}_{n+1} - \mathbf{r}^l_n\right]$



Application: Any problem where multiple agents need to establish consensus in decision making:
- autonomous sensor activation
- spectrum allocation in cognitive radio



Blackwell approachability

- CORRELATED EQUILIBRIUM (CE) [Aumann, 1987]: a generalization of Nash equilibrium

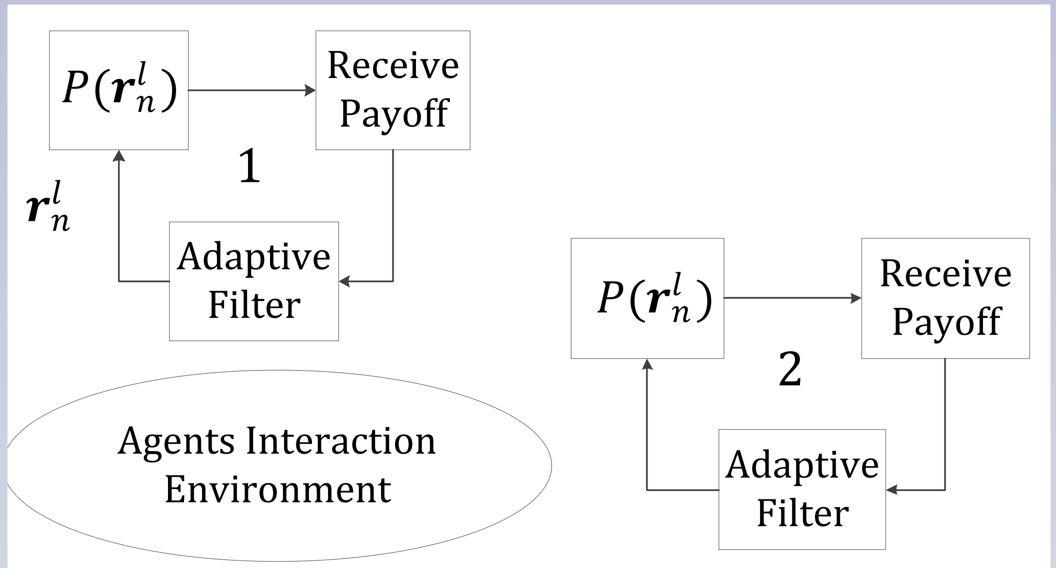
$$\mathcal{C}_e = \{\pi : \sum_{\mathbf{x}^{-l} \in A^{-l}} \pi^l(i, \mathbf{x}^{-l})[u^l(j, \mathbf{x}^{-l}) - u^l(i, \mathbf{x}^{-l})] \leq 0, \quad \forall i,j \in A^l, l \in \mathcal{L}\}$$

- *Why ?* 1) Correlation device: common history of actions, 2) Structural Simplicity: convex polytope, 3) Provably convergent learning algorithms

PART 3: ADAPTIVE FILTERING GAMES

Non-Bayesian - game theory in adaptive learning

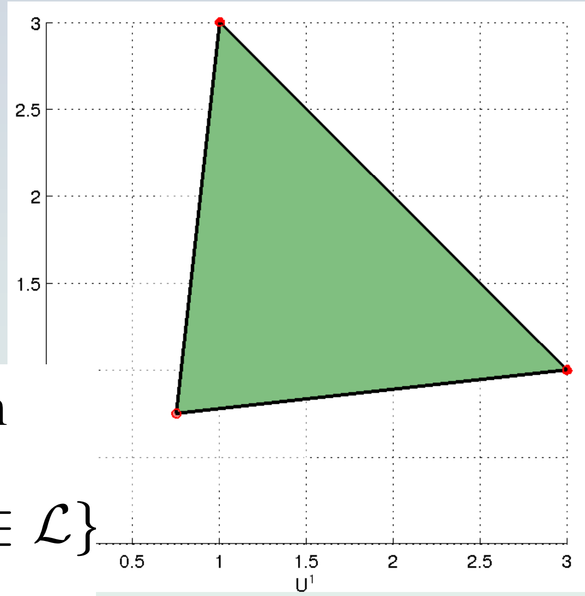
- Each agent l:
  1. Chooses action randomly $\{x^l_{n+1} = i\} \sim P(\mathbf{r}^l_n)$
  2. Receives stage utility $u^l(i, \mathbf{x}^{-l}_{n+1})$
  3. $\text{Regret}_{n+1}(i,j) = u^l(j, \mathbf{x}^{-l}_{n+1}) - u^l(i, \mathbf{x}^{-l}_{n+1})$
  4. Adaptive filter $\mathbf{r}^l_{n+1} = \mathbf{r}^l_n + \epsilon\left[\text{Regret}_{n+1} - \mathbf{r}^l_n\right]$



Key tool: Stochastic averaging theory (Kushner & Yin)

Benaim, Math OR, 2006
Differential inclusion

$$\frac{d\mathbf{r}(i,j)}{dt} \in \sum_{\mathbf{x}^{-l}} \pi^l(i, \mathbf{x}^{-l})[u^l(j, \mathbf{x}^{-l}_{n+1}) - u^l(i, \mathbf{x}^{-l}_{n+1})]$$



- CORRELATED EQUILIBRIUM (CE) [Aumann, 1987]: a generalization of Nash equilibrium

$$\mathcal{C}_e = \{\pi : \sum_{\mathbf{x}^{-l} \in A^{-l}} \pi^l(i, \mathbf{x}^{-l})[u^l(j, \mathbf{x}^{-l}) - u^l(i, \mathbf{x}^{-l})] \le 0, \quad \forall i, j \in A^l, l \in \mathcal{L}\}$$

- *Why ?* 1) Correlation device: common history of actions, 2) Structural Simplicity: convex polytope, 3) Provably convergent learning algorithms

Blackwell approachability

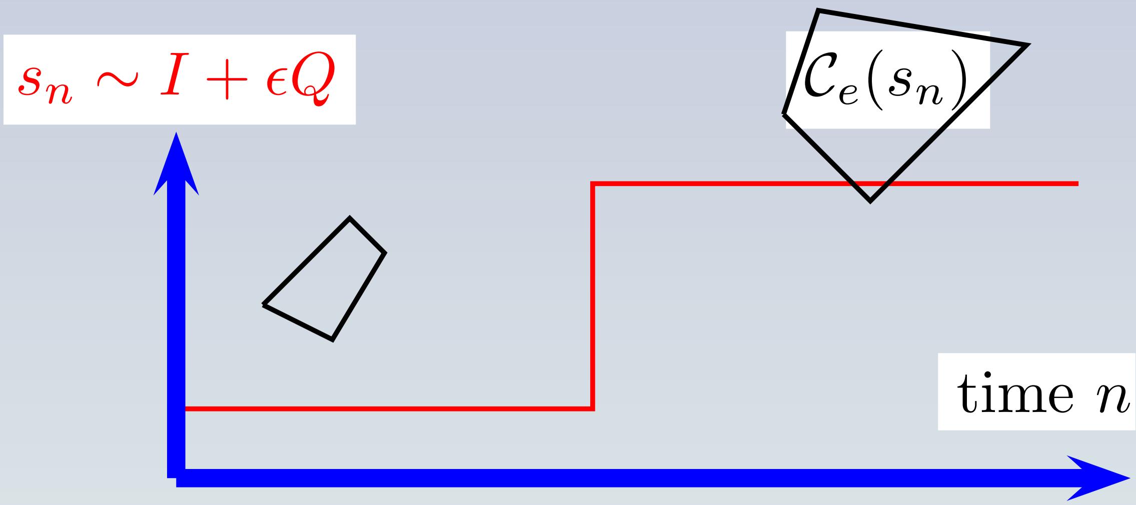
Tracking Time-varying Equilibria

Tracking Time–varying Equilibria

Each node deploys simple algorithm - global performance achieves consensus in decision space

Tracking Time–Varying Equilibria

Each node deploys simple algorithm - global performance achieves consensus in decision space

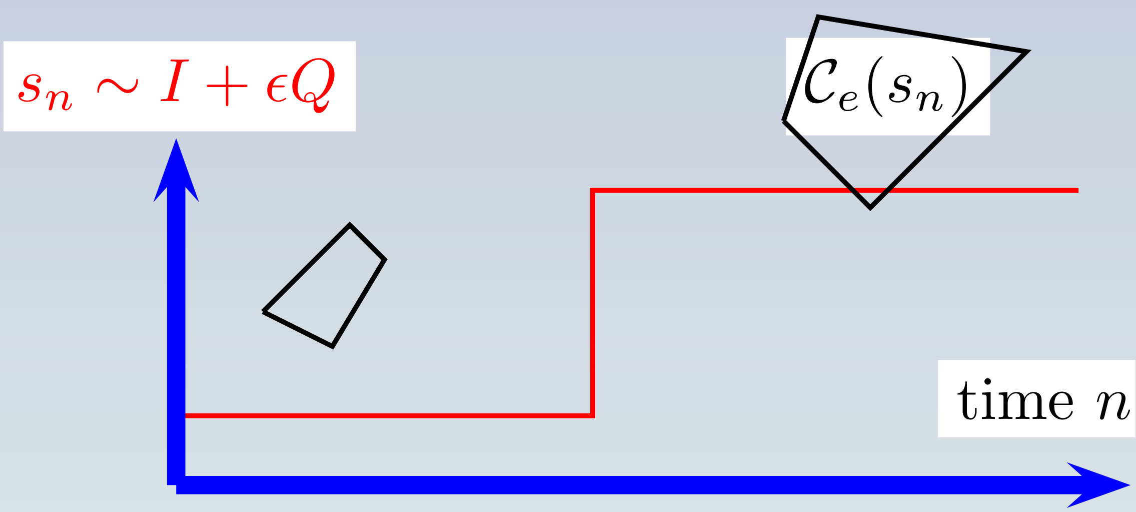


- Track time varying correlated equilibria -- dynamic spectrum allocation in cognitive radio.
- Asynchronous updates

Krishnamurthy, Yin, SIAM J.Opt [2004],
Kushner & Yin [2003],
Krishnamurthy, Yin, SIAM J. Multiscale [2009].

TRACKING TIME–VARYING EQUILIBRIA

Each node deploys simple algorithm - global performance achieves consensus in decision space



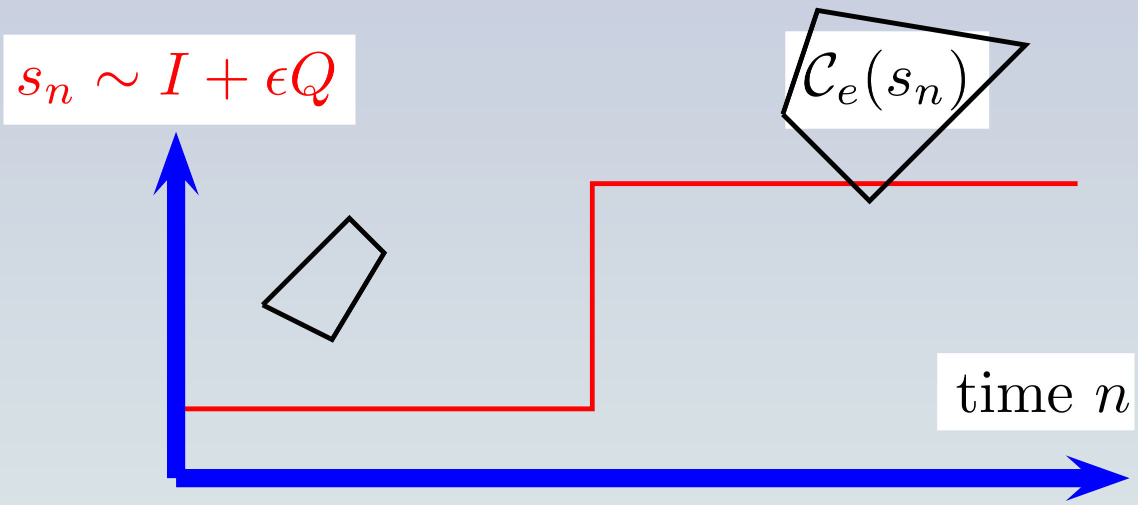
- Track time varying correlated equilibria -- dynamic spectrum allocation in cognitive radio.
- Asynchronous updates

Krishnamurthy, Yin, SIAM J.Opt [2004],
Kushner & Yin [2003],
Krishnamurthy, Yin, SIAM J. Multiscale [2009].

Non-standard stochastic averaging – requires use of "Martingale problem" of Strook and Varadhan. [Kushner & Yin 2003], [Ethier & Kurtz]

TRACKING TIME–VARYING EQUILIBRIA

Each node deploys simple algorithm - global performance achieves consensus in decision space



- Track time varying correlated equilibria -- dynamic spectrum allocation in cognitive radio.
- Asynchronous updates

Krishnamurthy, Yin, SIAM J.Opt [2004],
Kushner & Yin [2003],
Krishnamurthy, Yin, SIAM J. Multiscale [2009].

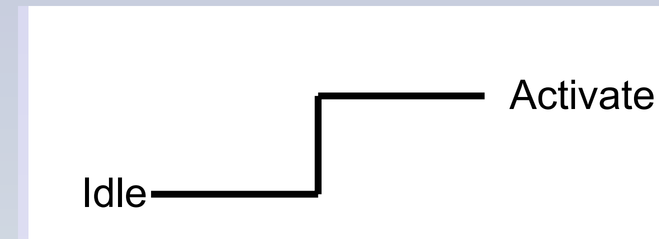
Adaptive heuristics are boundedly rational strategies (in fact, highly "bounded away" from full rationality). The main question of interest is whether such simple strategies may in the long run yield behavior that is nevertheless highly sophisticated and rational.

Hart, Econometrica, 2005

DISCUSSION

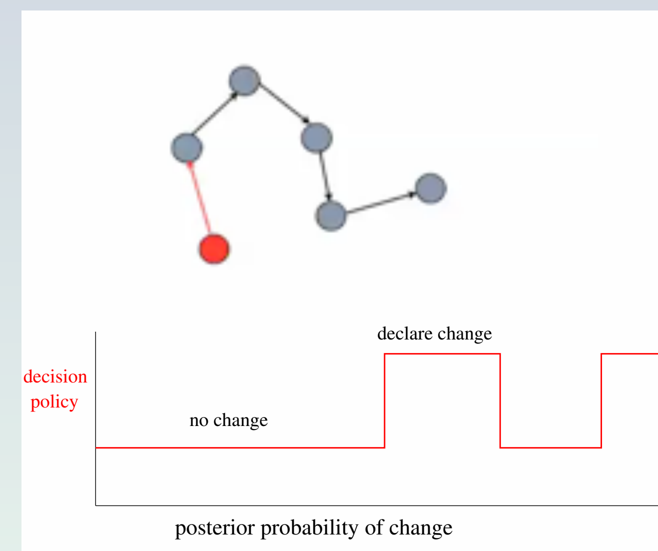
Vision: Analysis/design of interacting stochastic dynamical systems. Given simple local behavior, game theory/social learning is useful for anlaysis/synthesis of global behavior.

Part 1: Bayesian Global Game: Agents act *simultaneously*. Learn from data and predict other agents actions.

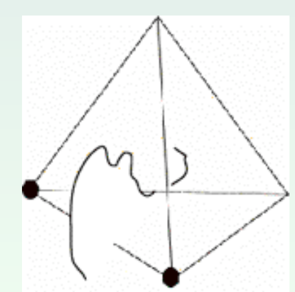


Part 2: Social Learning. Agents learn *sequentially* from actions of previous agents.

- Pure capitalism causes herding. Socialism delays herding.
- Global decisions from local decisions: multi-threshold



Part 3: Adaptive filtering games: Global behavior achieves consensus in action space - correlated equilibrium.



REFERENCES

Part 3: Learning Correlated Equilibria (adaptive filtering games)

- Hart & Mas-Collell, Econometrica, 2000.
- Benaim, Hofbauer & Sorin, SIAM J Control Opt, 2005 and Math OR 2006.
- Krishnamurthy, Topley, Yin, *Consensus Formation in two-time-scale Markovian system*, SIAM Journal Multiscale Modelling, 2009.
- Krishnamurthy, Maskery, Yin, *Decentralized Adaptive Filtering Algorithms for Sensor Activation*, IEEE Trans SP, 2007.

Part 2: Social Learning (Herding, Optimal Social Learning, Quickest Detection)

- Chamley, *Rational Herds,* 2004. Cambridge Univ Press.
- Smith & Sorenson, *Informational Herding and Experimentation*,
- Krishnamurthy, *Sequential Detection with Phase-Distributed Change Times and Nonlinear Penalty*, IEEE Transactions Information Theory (65 pages), 2011.

Part 1: Global Games (Structural results of Bayesian Nash equilibrium)

- Karp, Lee and Mason, *Global Game with strategic substitutes & complements*, Games and Economic Behavior, 2007.
- Krishnamurthy, *Decentralized Activation in Dense Sensor Networks via Global games*, IEEE Trans SP, Oct.2008...
- Krishnamurthy, Decentralized Spectrum Access amongst Cognitive Radios -- Multivariate Interacting Global Games Approach, IEEE Trans SP, 2009.