Decentralized stochastic control

The person-by-person and the common information approaches

Aditya Mahajan

McGill University

REPARTI Annual Workshop, May 23, 2014

Decision making by a single agent

Static optimization

$$\min_{u \in \mathcal{C}} c(u)$$

- Linear programming
- Convex optimization
- Non-convex optimization

Bayesian optimization

$$\min_{g} \mathbb{E}[c(\omega, g(Y(\omega)))]$$

- Stochastic programming
- Stochastic approximation
- Markov Chain Monte Carlo

Dynamic optimization/ Stochastic control

$$\min_{(g_1,\dots,g_T)} \mathbb{E}\left[\sum_{t=1}^{T} c_t(x_t, u_t)\right]$$

where

$$x_t = f_t(x_t, u_t, W_t),$$

$$y_t = h_t(x_t, N_t),$$

$$u_t = g_t(y_{1:t}, u_{1:t})$$

- Dynamic programming
- Pontryagin maximum principle
- Multi-stage stochastic programming

Decision making by multiple agents

Game theory
Each agent has an individual objective. Agents compete to minimize individual costs.

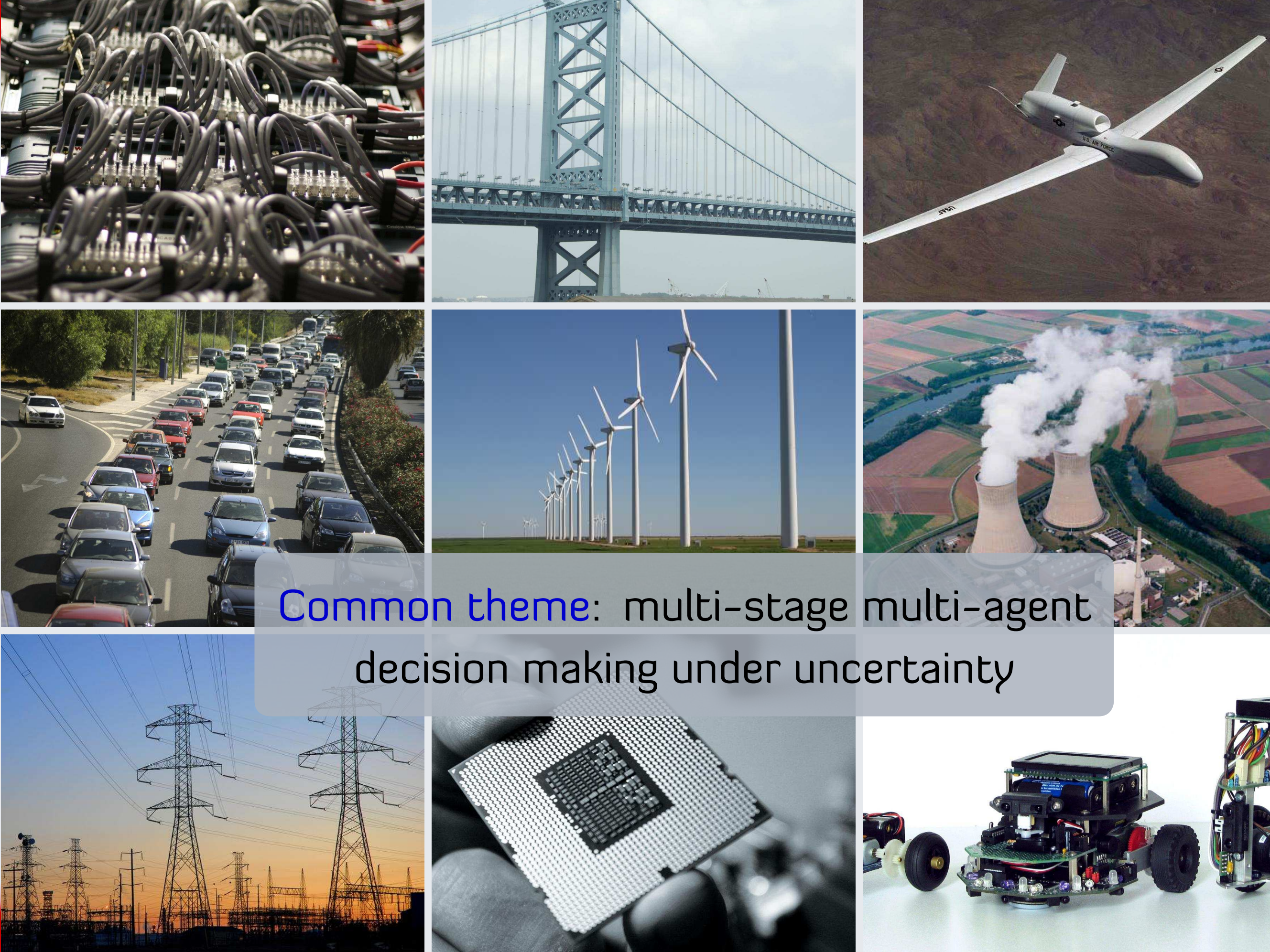
- Static games
- Bayesian games
- Dynamic games or multi-stage games with imperfect information

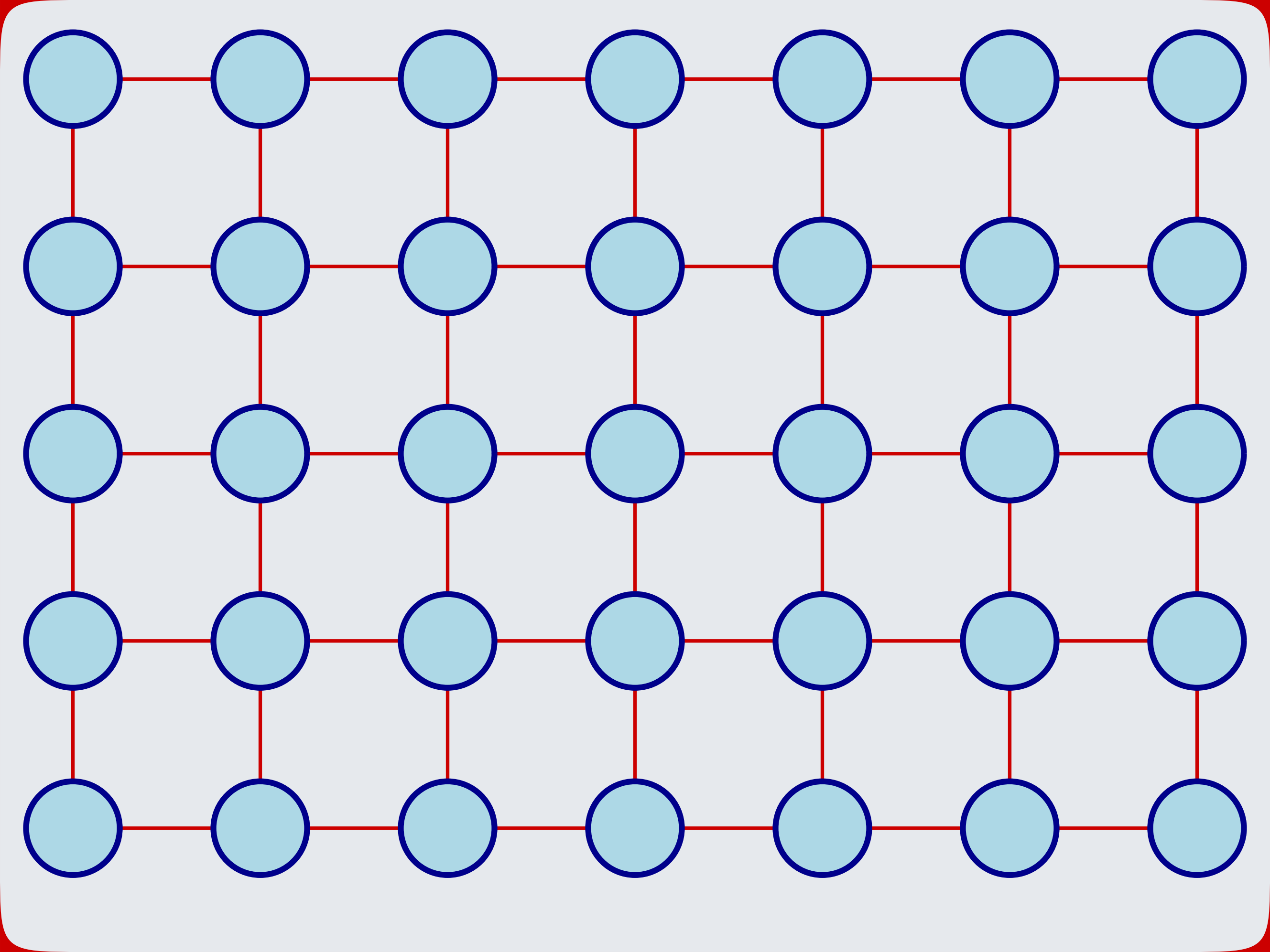
Team theory/ Decentralized stochastic control
All agents have a common objective. Agents cooperative to minimize team costs.

- Static (Bayesian) teams
- Dynamic teams or decentralized stochastic control

Research in team theory started in Economics in mid 50's in the context of organizational behaviour. It has been studied in Systems and Control since the late 60's and in Artificial Intelligence since late 90's.

The motivation of decentralized control is not that it is more powerful than centralized control; rather it is necessary in systems where centralized information is not available or is not practical.


Common theme: multi-stage multi-agent
decision making under uncertainty



Conceptual difficulties in decentralized control

Witsenhausen Counterexample

- A two step dynamical system with two controllers
- Linear dynamics, quadratic cost, and Gaussian disturbance
- Non-linear controllers outperform linear control strategies . . .

  . . . cannot use Kalman filtering + Riccati equations

Whittle and Rudge Example

- Infinite horizon dynamical system with two symmetric controllers
- Linear dynamics, quadratic cost, and Gaussian disturbance
- A priori restrict attention to linear controllers
- Best linear controllers not representable by recursions of finite order

Complexity analysis

- All random variables are finite valued
- Finite horizon setup
- The problem of finding the best control strategy is in NEXP

- Witsenhausen, "A counterexample in stochastic optimum control," SICON 1969.
- Whittle and Rudge, "The optimal linear solution of a symmetric team control problem," App. Prob. 1974.
- Bernstein, et al, "The complexity of decentralized control of Markov decision processes," MOR 2002.
- Goldmand and Zilberstein, "Decentralized control of cooperative systems: categorization and complexity," JAIR 2004.

Overview of my research in decentralized stochastic control

Research theme

- Identify specific information structures that capture key features of applications but, at the same time, are amenable to analysis.
- Develop analytic and computation approaches to optimally design controllers for these information structures.

Two main approaches

- The person-by-person approach
  - Static teams with infinite players
- The common-information approach
  - Delayed sharing information structure

The person-by-person approach

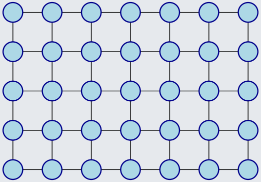
Static teams with finite number of agents (Marschak 1955)

Observations $(Y^1, \dots, Y^n)$ defined on a common probability space

Control Action $U^i = g^i(Y^i)$ is the control action of agent i

Objective $\min_{(g^1,\dots,g^n)} \mathbb{E}\big[\text{any function of } (Y, U)\big]$

Example



- Neighbors have correlated observations

$$\Sigma_{ii} = \sigma^2 \quad \text{and} \quad \Sigma_{ij} = \alpha\sigma^2 \text{ for } j \in N_i$$

- Objective: Choose $U^i = g^i(Y^i)$ to minimize

$$\mathbb{E}\Big[\sum_i (U^i)^2 + \sum_i \sum_{j \in N_i} \big((Y^i - U^i) - (Y^j - U^j)\big)^2\Big]$$

Salient features

- Correlated observations and coupled costs.
- Static optimization problem.
- Seeking an optimal off-line design, not an iterated solution with communication between neighbors.

- Marschak, "Elements for a theory of teams," Management Science, 1955

Solution to static LQG teams (Radnar 1962)

Solution approach

1. Identify sufficient conditions for optimality
   - **(GO)**: Sufficient conditions for global optimality

     $$\forall(\tilde{g}^1, \dots, \tilde{g}^n) : \quad J(g^1, \dots, g^n) \leqslant J(\tilde{g}^1, \dots, \tilde{g}^n)$$

   - **(PBPO)**: Sufficient conditions for person-by-person optimality

     $$\forall(\tilde{g}^1, \dots, \tilde{g}^n) \text{ and } \forall i : \quad J(g^1, \dots, g^n) \leqslant J(\tilde{g}^i, g^{-i})$$

2. Show that when Y are jointly Gaussian and cost is quadratic in (Y, U)

   $$(\mathbf{PBPO}) \implies (\mathbf{GO})$$

3. Assume all controllers are linear, i.e., $U^i = H^i Y^i$

   $$(\mathbf{PBPO}) \equiv \text{set of } n \text{ linear equations} : A_{n\times n} h_{n\times 1} = b_{n\times 1}$$

   where $h_{n\times 1} = \text{vec}[H_1 \mid \cdots \mid H_n]$

Therefore, globally optimal solution obtained by solving n linear equations

- Radner, "Team decision problems," Ann. Math. Statist., 1962
- Marshak and Radner, Economic Theory of Teams, Yale University Press, 1972.

Static LQG teams with infinite agents (MMY 2013)

Objective $$\min_{(g^1, g^2, \dots)} \limsup_{n \to \infty} \frac{1}{n} \mathbb{E}\left[U^\top Q U + Y^\top P U\right]$$

Motivation
- Intermediate step for extending some results in dynamic teams to infinite horizon.
- Proxy for large scale systems.

Key difficulty Radnar's approach breaks down because $(\mathbf{PBPO}) \not\Rightarrow (\mathbf{GO})$.

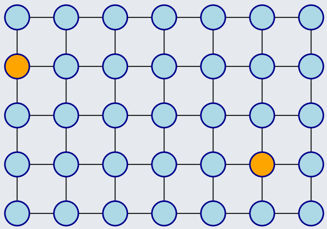
Our approach Use spectral properties of infinite dimensional Toeplitz matrices to identify sufficient conditions under which $(\mathbf{PBPO}) \implies (\mathbf{GO})$.

Main Theorem Under appropriate symmetry and regularity conditions, the optimal strategy for infinite agents is periodic and obtained by solving a finite dimensional system of linear equations.

- Mahajan, Martins, Yüksel, "Static LQG Teams with Countably Infinite Players", CDC 2013.

The common-information approach

Delayed sharing information structure



Dynamics $X_{t+1} = f_t(X_t, (U^1_t, \dots, U^n_t), W^0_t)$

Observations $Y^i_t = h^i_t(X_t, W^i_t)$

Delayed sharing Agent i observes k-step delayed observations and control of all other agents.

Objective $$\min_{(g^1_{1:T}, \dots, g^n_{1:T})} \mathbb{E}\Big[\sum_{t=1}^{T} c_t(X_t, (U^1_t, \dots, U^n_t))\Big]$$

Literature overview

- Witsenhausen 1971
  - Proposed as a bridge between centralized and decentralized systems.
  - Asserted structure of optimal control strategies.
- Varaiya and Walrand, 1978
  - Proved Witsenhausen's assertion for $k = 1$ (one-step delay).
  - Counterexample to disprove Witsenhausen's assertion for $k \geqslant 2$.

- Witsenhausen, "Separation of estimation and control," Proc IEEE, 1971.
- Varaiya and Walrand, "On delayed sharing patterns," IEEE TAC 1978.

Common-info approach for delayed sharing (NMT 2011)

Solution approach

1. Split available information into two parts
   - Common information: $C_t = \cap_{s \geqslant t} \cap_{i=1}^{n} I_s^i = \{Y_{1:t-k}, U_{1:t-k}\}$
   - Local information: $L_t^i = C_t \setminus L_t^i = \{Y_{t-k+1:t}^i, U_{t-k+1:t-1}^i\}$

2. Construct an equivalent centralized coordinated system where
   - Observation history: C_t
   - Control action: $(\gamma_t^1, \dots, \gamma_t^n)$ where $\gamma_t^i \colon L_t^i \mapsto U_t^i$.
   - Coordination law: $\psi_t(C_t) = (\gamma_t^1, \dots, \gamma_t^n)$

3. Solve the centralized coordinated system
   - Information state: $\pi_t = \mathbb{P}(\text{state for I/O mapping} \mid \text{data at controller})$
     $= \mathbb{P}(X_t, L_t^{1:n} \mid C_t)$.
   - Structure of optimal controller: $(\gamma_t^1, \dots, \gamma_t^n) = \psi_t(\pi_t)$
     Equivalently, $U_t^i = g_t^i(\pi_t, L_t^i)$.
   - Appropriate dynamic program to find $(\gamma_t^1, \dots, \gamma_t^n)$.

- Nayyar, Mahajan, Teneketzis, "Optimal control strategies in delayed sharing information structures," IEEE TAC 2011.

An easy solution to a

long-standing open problem!

Generalization and refinements

Partial history sharing

- Most general system solvable by common-information approach.
- Many existing results in decentralized control are special cases
- In the worst case, solution scales double exponentially with n.

- Nayyar, Mahajan, Teneketzis, “Decentralized stochastic control with partial history sharing: A common information approach,” IEEE TAC 2013.

Control sharing

- Motivated by communication networks where control actions are observed by all agents.
- Show that under an appropriate conditional independence assumption, the solution scales exponentially with n.

- Mahajan, “Optimal decentralized control of coupled subsystems with control sharing,” IEEE TAC 2013.

Mean-field sharing

- Motivated by smart grids where agents are weakly coupled through the mean-field.
- Show that under an appropriate symmetry assumption, the solution scales polynomially with n.

- Arabneydi and Mahajan, “Team optimal control of coupled subsystems with mean field sharing,” CDC 2014 (submitted).

Applications Real-time communication, sensor networks, smart grids.

Thank you

Contact details
- email: aditya.mahajan@mcgill.ca
- homepage: http://www.ece.mcgill.ca/~amahaj1/index.html

Collaborators
- Nuno Martins, Univ of Maryland
- Ashutosh Nayyar, Univ of Southern California
- Sekhar Tatikonda, Yale University
- Demosthenis Teneketzis, Univ of Michigan
- Serdar Yüksel, Queen's University

PhD students
- Jalal Arabneydi
- Jhelum Chakravorty

Master's students
- Mehnaz Mannan
- Prokopis Prokopiou
- Can Cui