Role of information structures in decentralized decision making

Aditya Mahajan

February 8, 2010, UCLA

Acknowledgments

Collaborators

- Demos Teneketzis, Univ of Michigan,
- Sekhar Tatikonda, Yale Univ
- Ashutosh Nayyar, Univ of Michigan,
- Serdar Yüksel, Queen's Univ

Funding Agencies

- NSF
- NASA

Decentralized systems are everywhere . . .

Communication Networks

Sensor Networks

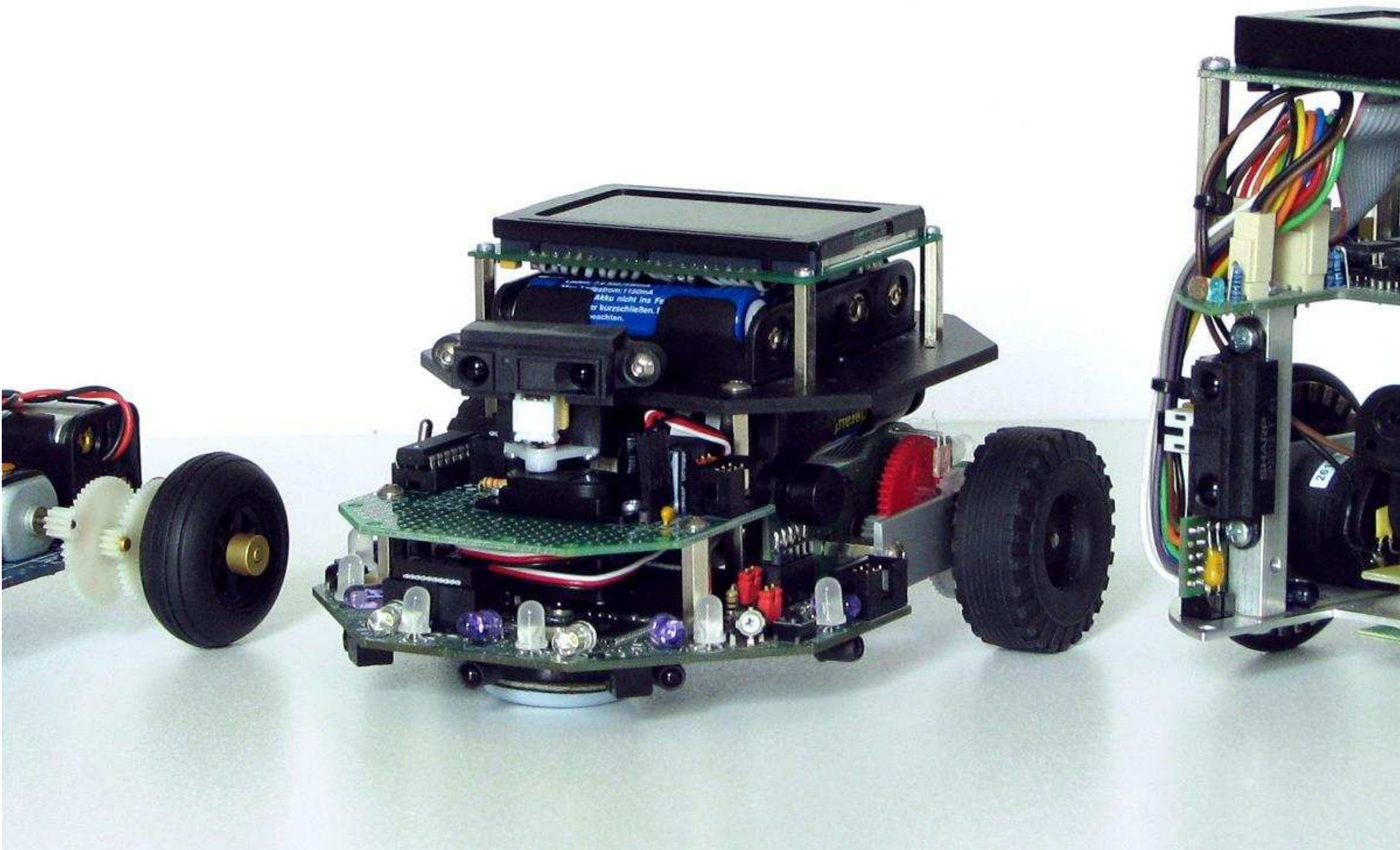
Surveillance Networks

Transportation Networks

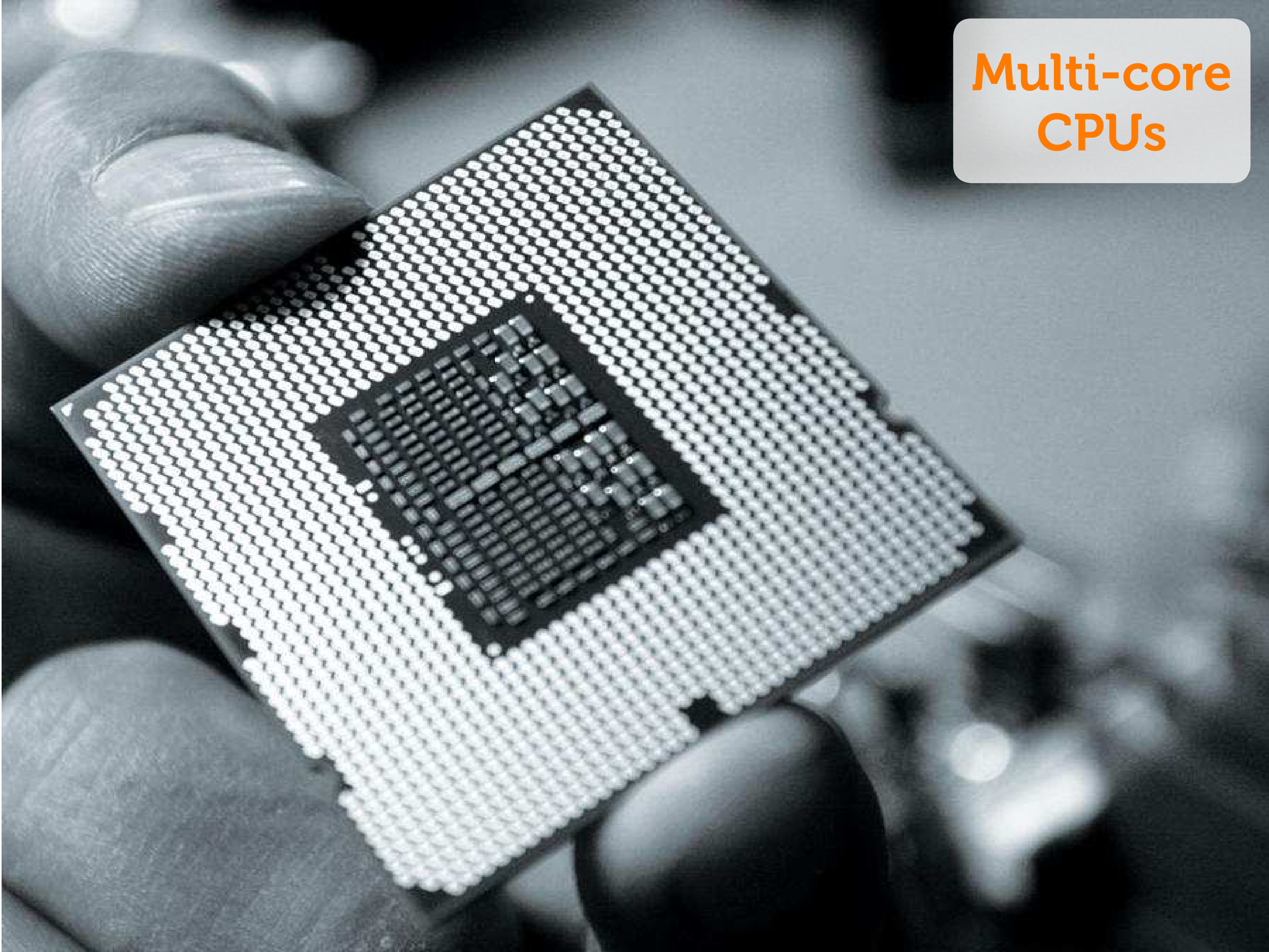
Control Systems

Monitoring and Diagnostic Systems

Robotics



Multi-core CPUs



Power Distribution

. . . and many others . . .

Basic research premise

- The various applications where decentralized systems arise are **independent areas of research** with dedicated communities.
- However, most applications share **common features** and **common design principles**.
-

> Develop a **systematic methodology**
> that addresses these commonalities.

- Such a methodology will provide **design guidelines** for all applications.

Characteristics of decentralized systems

Multiple agents that have **different information** need to **cooperate and coordinate**

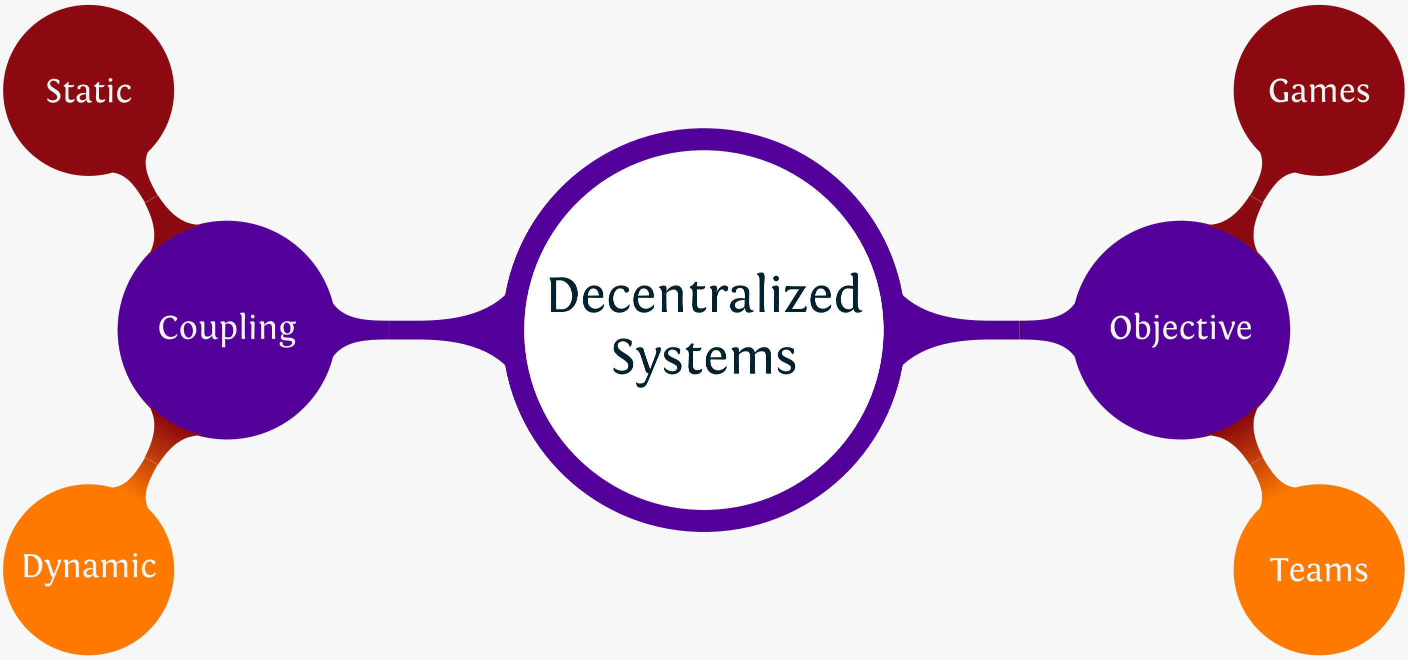
Required:

a theory for decentralized decision making

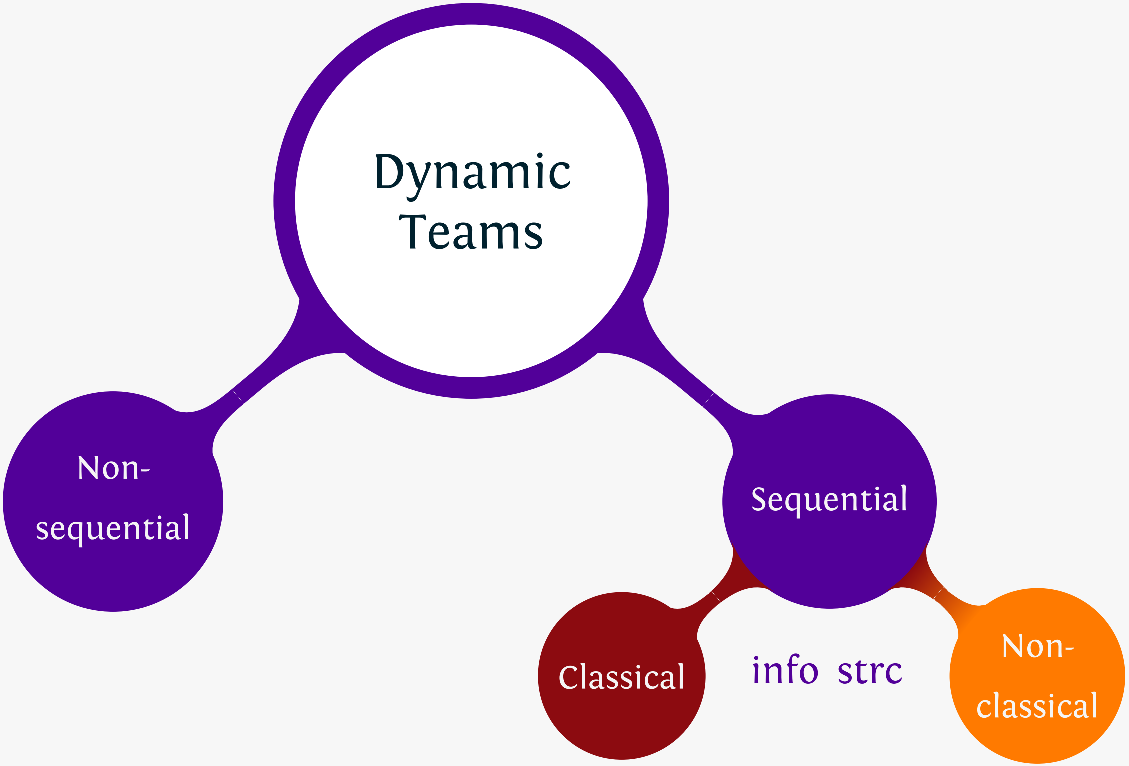
Outline

1. Overview of decentralized systems
   - **Classification**: games vs. teams; single- vs. multi-stage; etc.
   - **Objective**: structural results and sequential decomposition
2. Why can't we directly use Markov decision theory
3. n-step delayed sharing structure
   - Information states
   - Summarizing the affect of past on future performance
4. Conclusion

Classification of decentralized systems



Classification of decentralized systems



Sequential dynamic teams
with
non-classical info strc

Salient Features

- Sequential Team

Order in which the agents act can be fixed in advance: $A_1, A_2, \dots, A_n$.

- Non-classical information structures

Let $\mathcal{J}_i$ represent the "information" known to A_i.
$\exists i$ such that $\mathcal{J}_i \not\subseteq \mathcal{J}_{i+1}$

Literature Overview

Economics

- R. Radner, *Team decision problems*, Ann. Math. Statistics, 1962.
- J. Marschak and R. Radner, *Economic Theory of Teams*, Yale Univ Press, 1972.
- C.B. McGuire, *Comparison of Information Structures*, Cowles Foundation, 1959.

Controls

- H.S. Witsenhausen, *On information structures, Feedback and Causality*, SIAM J. Control, 1971
- H.S. Witsenhausen, *Separation of estimation and control for discrete time systems*, Proc of IEEE, 1971.

Literature Review

Difficulty of the problem

- H.S. Witsenhausen, *counterexample in stochastic control*, SICON 1968

  Linear policies are not optimal for linear quadratic Gaussian systems under non-classical information structure

- D.S. Bernstein, S. Zilberstein, and N. Immerman, 2000

  In general, the problem is NEXP-complete: **no polynomial time** solution can exist.

Literature Review

- Results for general setup
  - **Standard form**: Witsenhausen 1973
  - **Non-classical LQG problems**: Sandell and Athans, 1974
  - **Multi-criteria problems**: Basar, 1978
  - **Equivalence of static and dynamic teams**: Witsenhausen 1988
  - **Non-sequential systems**: Andersland and Teneketzis, 1992 and 1994.

Literature Review

- Results for specific information structures
  - **Partially nested info structures**, Ho and Chu, 1972, Ho, Kastner, and Wong, 1978, Ho, 1980
  - **Delayed sharing info structures**, Witsenhausen 1971, Varaiya and Walrand, 1978, Mahajan, Nayyar, and Teneketzis 2010.
  - **Common past**, Aicardi *et al* 1987
  - **Partially observed and partially nested**, Casalino *et al* 1984
  - **Periodic sharing info structure**, Ooi *et al* 1997
  - **Tower info structures**, Swigart and Lall 2008
  - **Stochastic nested and belief sharing**, Yüksel 2009
  - **P-classical and P-quasiclassical**, Mahajan and Yüksel, 2010

This talk

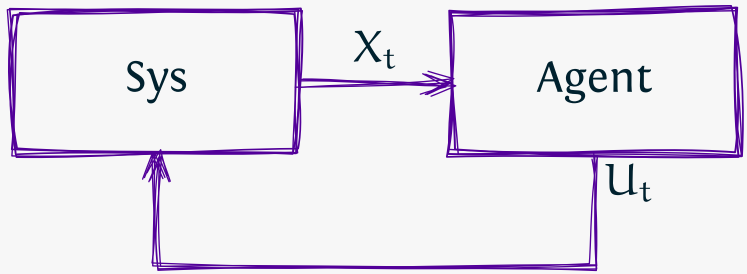
- A. Mahajan, A. Nayyar, and D. Teneketzis, *Identifying tractable decentralized control problems on the basis of information structure,* Allerton 2008.
- A. Mahajan and S. Tatikonda, *Sequential team form and its simplification using graphical models,* Allerton 2009.
- A. Nayyar, A. Mahajan, and D. Teneketzis, *Optimal control strategies in delayed sharing information structures,* ACC 2010.

Solution Concept

- Structural results
  - **Discard irrelevant** information
  - **Compress relevant** information to a compact statistic
  - **Restrict attention to a sub-class** of decision rules
- Sequential decomposition
  - **Divide and conquer**: Exploit sequential and multi-stage nature of the problem
  - **Convert** a one-shot optimal design problem into a **sequence of nested** optimization problems.

Same solution concepts as centralized systems

Markov decision process (MDP)

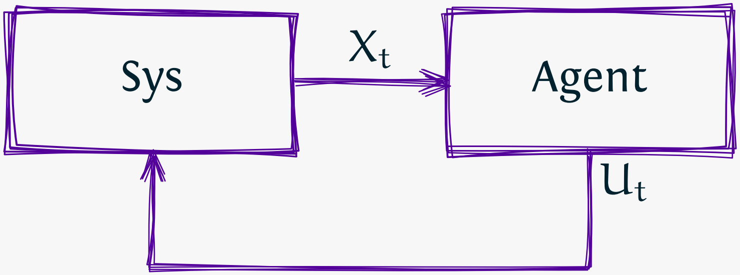


$$X_{t+1} = f_t(X_t, U_t, W_t)$$

$$U_t = g_t(X_{1:t}, U_{1:t-1})$$

$$\min \mathbb{E} \left\{ \sum_{t=1}^{T} c_t(X_t, U_t) \right\}$$

Markov decision process (MDP)



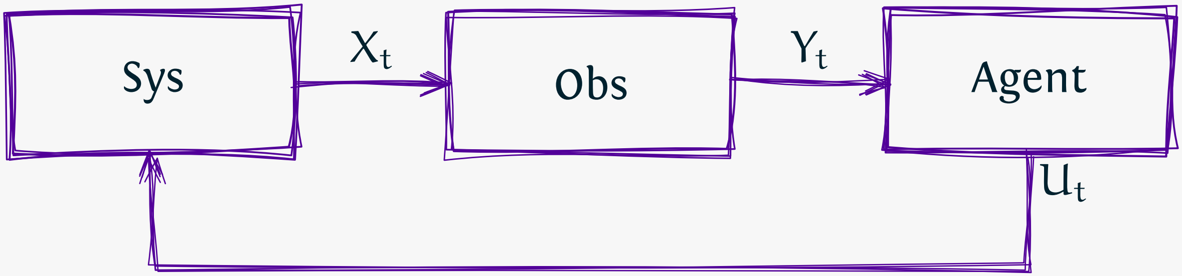
- Structural Results: **Discard past observations and actions**

  Choose current action based on **current state X_t**

- Sequential decomposition: **Dynamic programming**

  Recursively **compute the next action U_t** for each realization of the current state X_t

Partially observable MDP (POMDP)



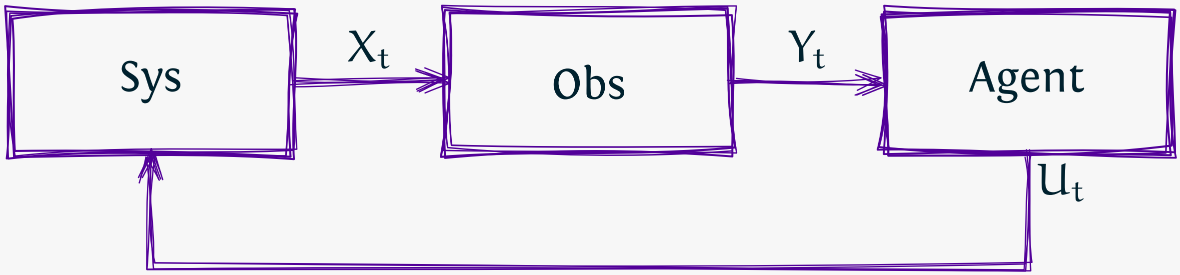
$$X_{t+1} = f_t(X_t, U_t, W_t)$$

$$Y_t = h_t(X_t, Q_t)$$

$$U_t = g_t(Y_{1:t}, U_{1:t-1})$$

$$\min \mathbb{E}\left\{ \sum_{t=1}^{T} c_t(X_t, U_t) \right\}$$

Partially observable MDP (POMDP)



- Structural Results: **Compress past observations and actions**

  Choose current action based on **current info state**
  $\pi = \Pr(\text{state of system} \mid \text{all data at agent})$

- Sequential decomposition: **Dynamic programming**

  Recursively **compute the next action U_t** for each realization of the current **information** state π_t

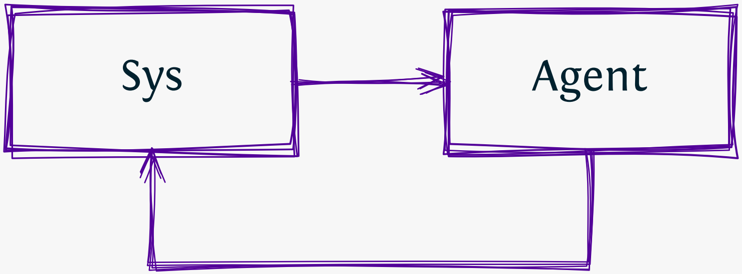
Objective

Find a systematic methodology to determine structural results and sequential decomposition for sequential teams

Why can't we use what we know?

Markov decision theory makes an implicit assumption: **information is centralized**

MDP revisited



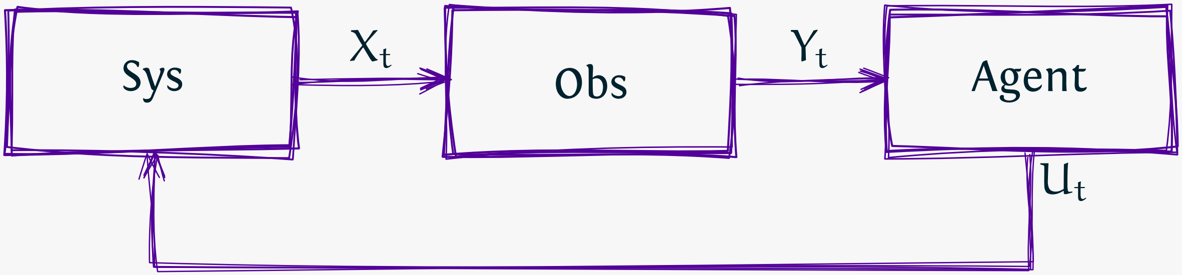
- Information State

current **state X_t** of the system

- Implicit Assumption

one agent with **perfect recall**

POMDP revisited



- Information State

$$\pi_t = \Pr(\text{state of system} \mid \text{all data at agent})$$

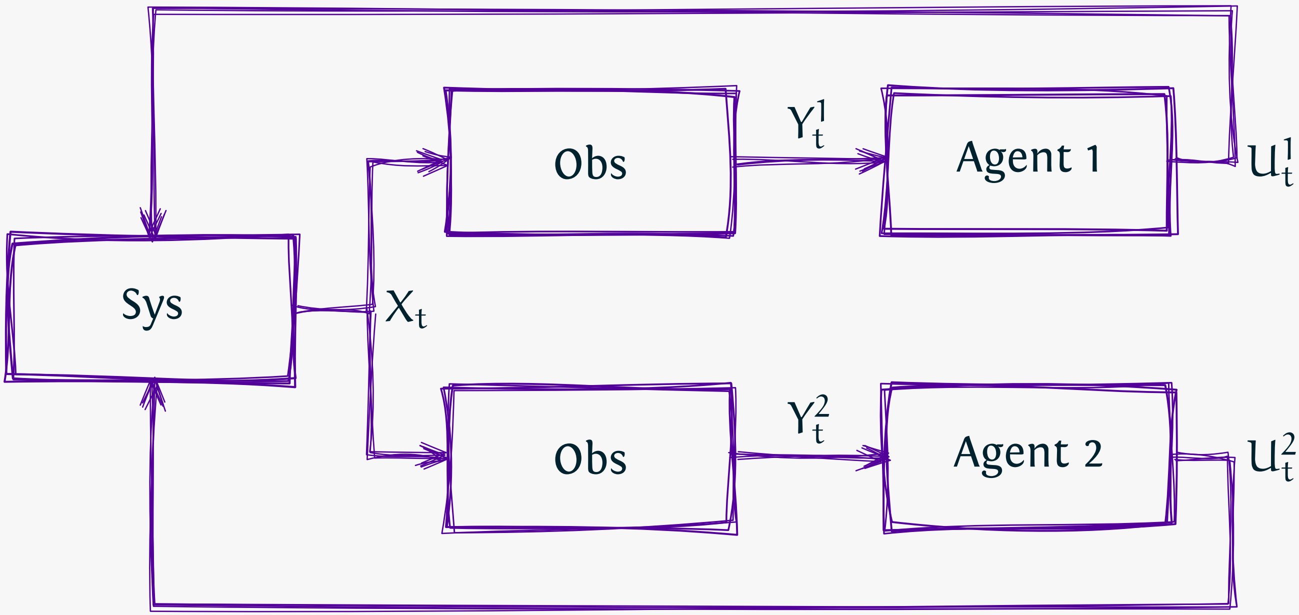
- Implicit Assumption

$$\text{data at time } t \subseteq \text{ data at time } t+1$$

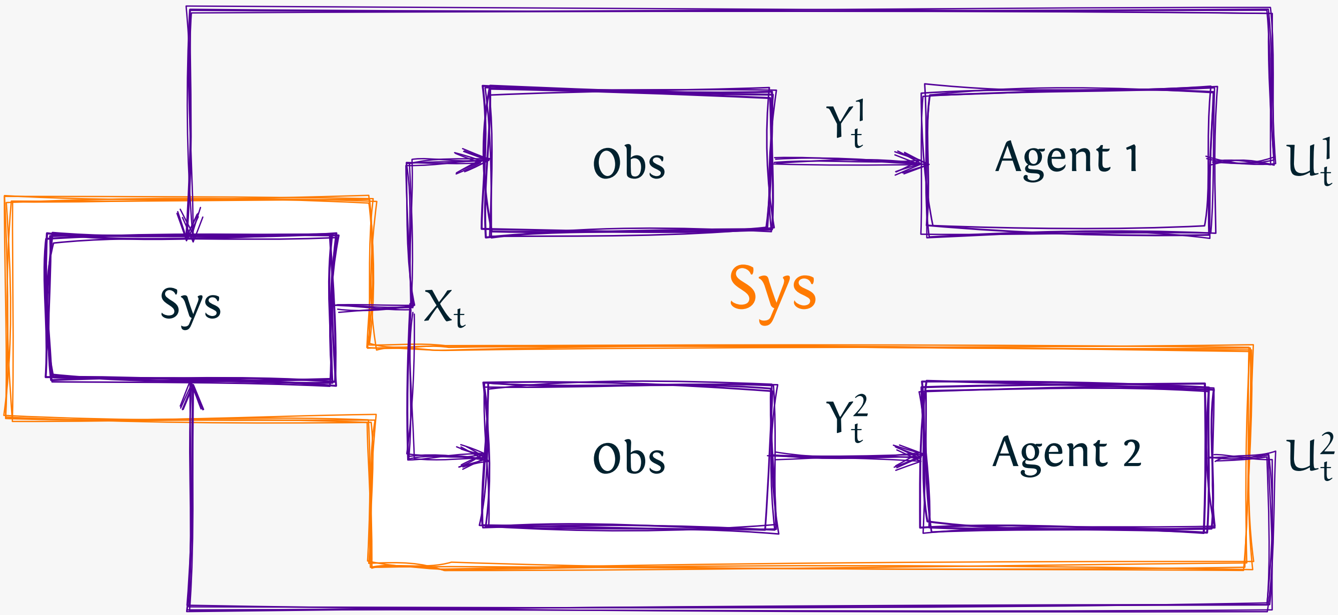
one agent with **perfect recall**

What happens if this assumption is not satisfied?

An example with two agents

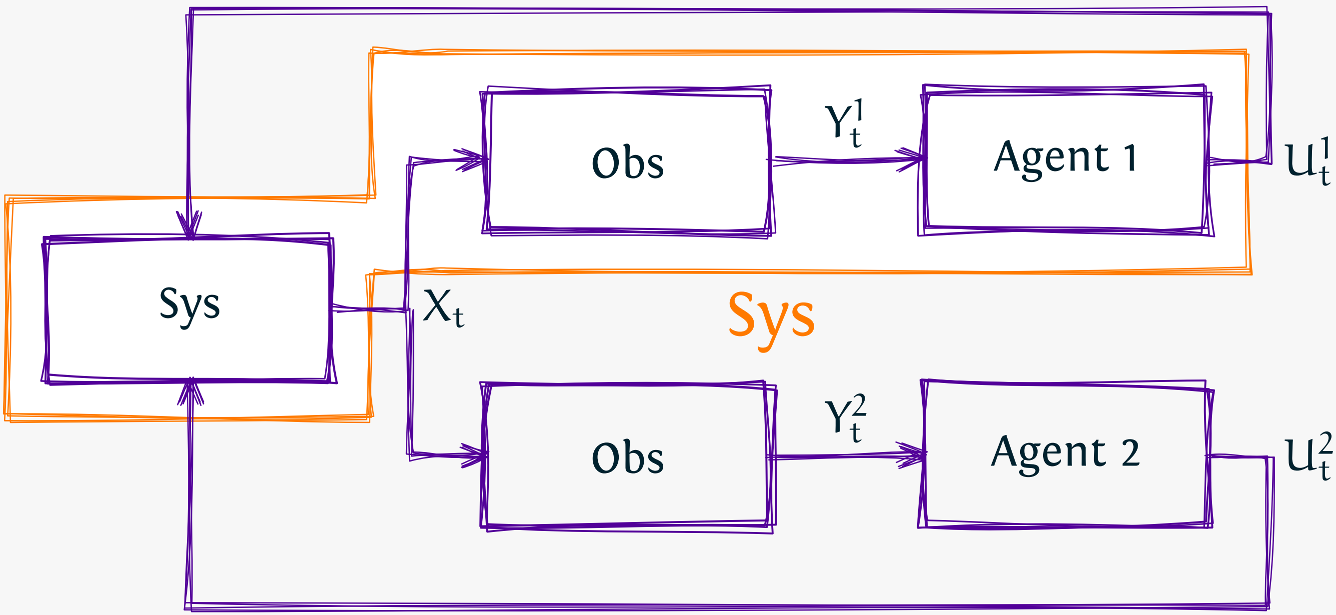


From the p.o.v. of agent 1



$$\pi_t^1 = \Pr(X_t, Y_{1:t}^2, U_{1:t-1}^2 \mid Y_{1:t}^1, U_{1:t-1}^1)$$

From the p.o.v. of agent 2



$$\pi_t^2 = \Pr(X_t, Y_{1:t}^1, U_{1:t-1}^1 \mid Y_{1:t}^2, U_{1:t-1}^2)$$

What happens when we try to combine the two p.o.v.

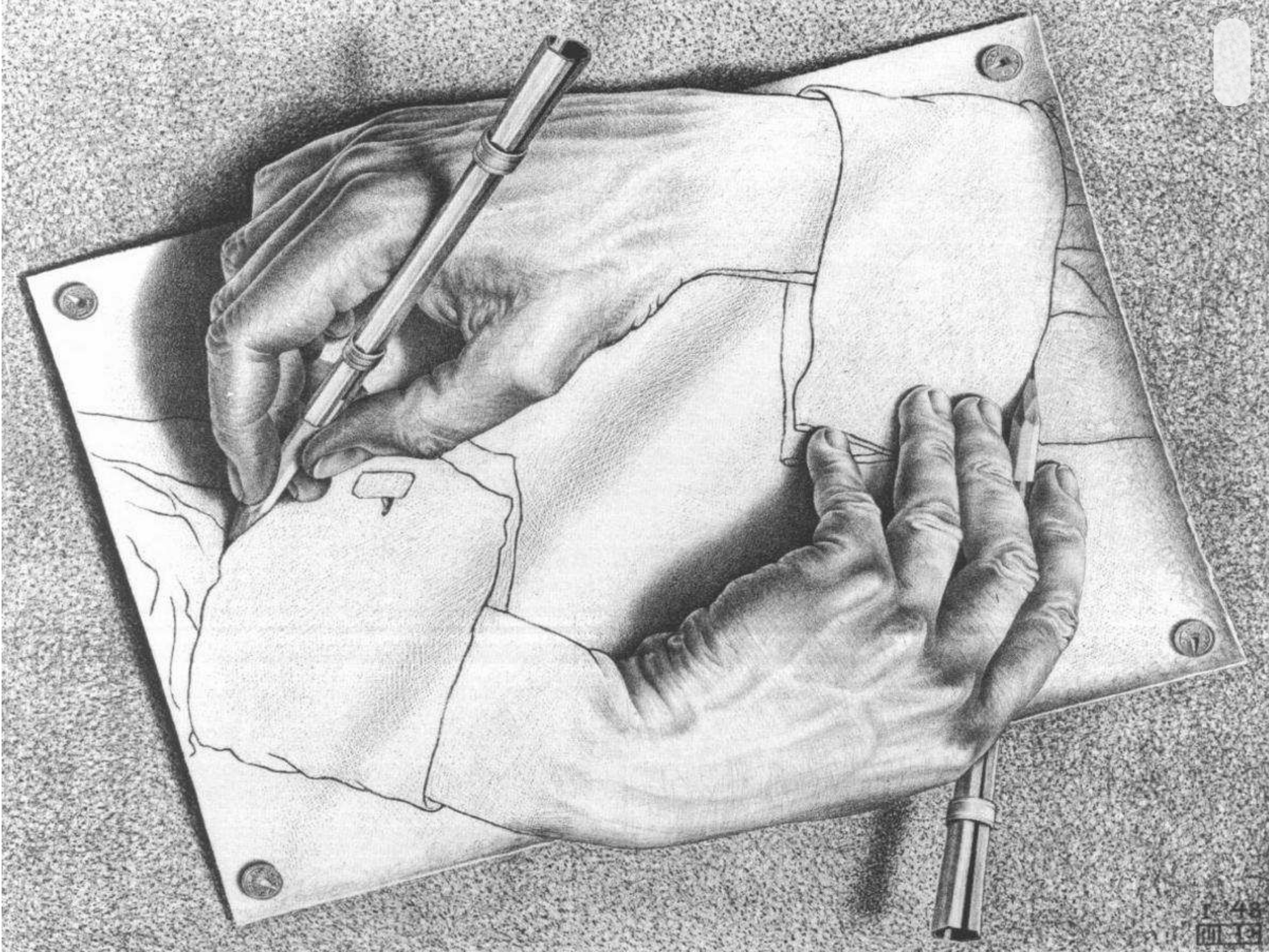
- Each agent's belief on the other agent's obs

$$\pi_t^1 = \Pr(X_t, Y_{1:t}^2, U_{1:t-1}^2 \mid Y_{1:t}^1, U_{1:t-1}^1)$$
$$\pi_t^2 = \Pr(X_t, Y_{1:t}^1, U_{1:t-1}^1 \mid Y_{1:t}^2, U_{1:t-1}^2)$$

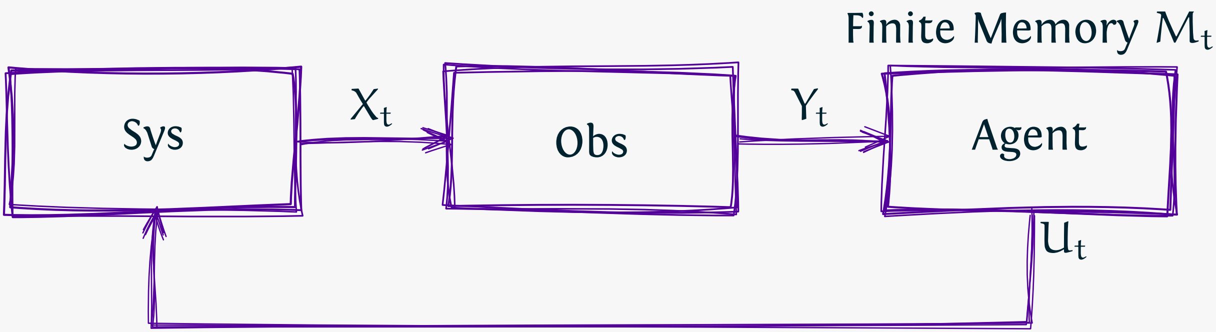
- Each agent's belief on the other agent's belief on the first agent's obs

$$\hat{\pi}_t^1 = \Pr(X_t, \pi_t^2 \mid \pi_t^1)$$
$$\hat{\pi}_t^2 = \Pr(X_t, \pi_t^1 \mid \pi_t^2)$$

Each agent is **second-guessing** the other



An example with one agent without perfect recall



$$X_{t+1} = f_t(X_t, U_t, W_t)$$

$$Y_t = h_t(X_t, Q_t)$$

$$U_t = g_t(Y_t, M_t)$$

$$M_{t+1} = l_t(Y_t, M_t)$$

What happens if we use the same approach as POMDPs

$$\pi_t = \Pr(X_t \mid Y_t, M_t)$$

$$\sigma(Y_t, Mt) \not\subseteq \sigma(Y_{t+1}, M_{t+1})$$

π_t cannot be updated recursively

What is the correct

notion of state

Difficulties in decentralized control

- The notion of state

How do we choose information states

- The second guessing argument

How does an agent know what other agent think about what it knows

Triple-aspect of control — estimation, control, and **communication/signaling**

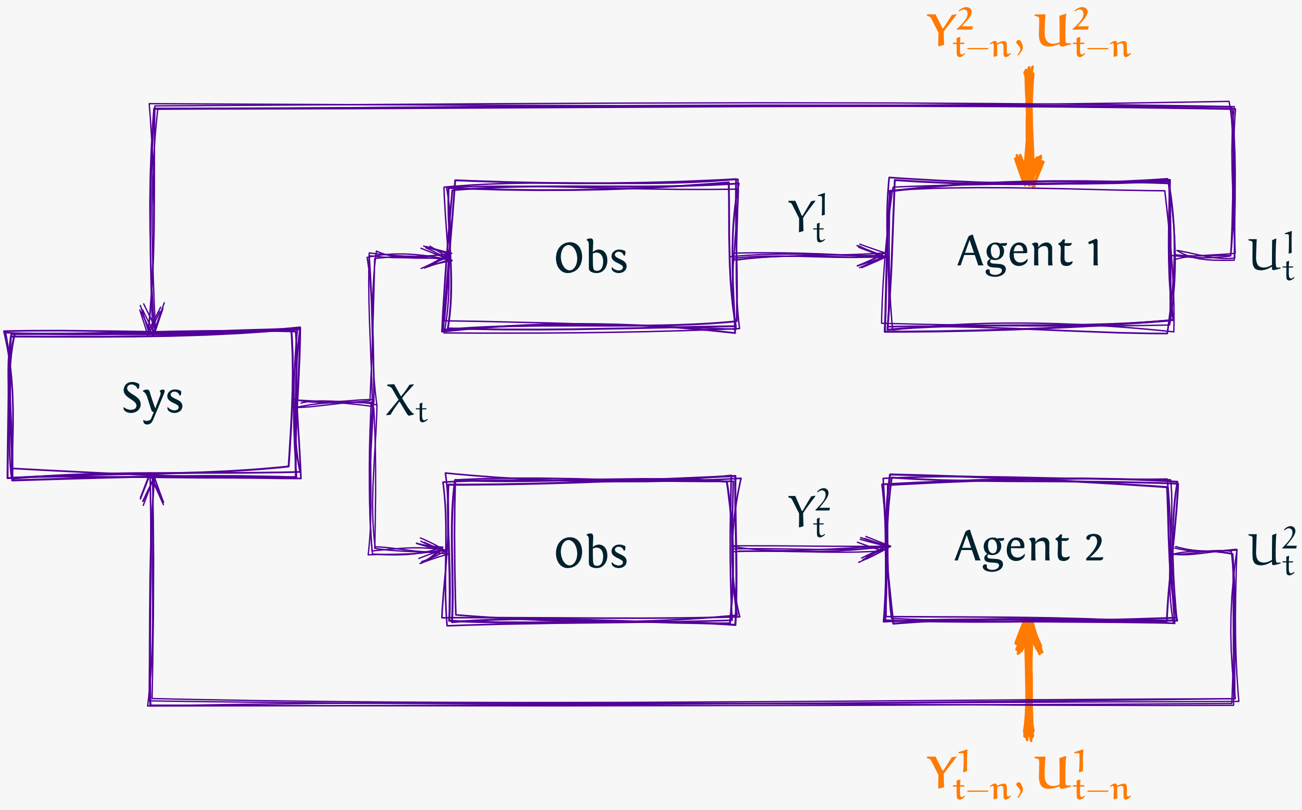
Our approach

- The notion of state
  - Start from first principles
  - State for **what purpose**? State for **whom**?
- The second guessing argument
  - Exploit common knowledge
- Triple-aspect of control
  - Each step of the dynamic program is a **functional** optimization problem

An Example

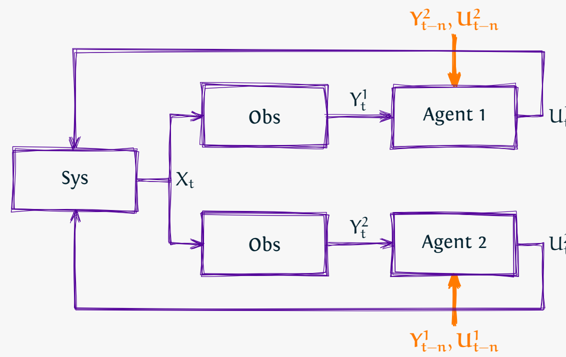
Delayed sharing information structure

Delayed sharing info structure (DSIS)



Delayed sharing info structure (DSIS)

- K controllers that **share information with a delay** of n time steps



- $n = 0 \Rightarrow$ classical info structure (centralized system)

- $n = \infty \Rightarrow$ non-classical info structure with no sharing (completely decentralized system)

Delayed sharing info structure (DSIS)

- **Common information**: $\Delta_t = (Y^1_{1:t-n}, Y^2_{1:t-n}, U^1_{1:t-n}, U^2_{1:t-n})$
- **Private information**: $\Lambda^k_t = (Y^k_{t-n+1:t}, U^k_{t-n+1:t-1}), \quad k = 1, 2.$
- Dynamics

$$X_t = f_t(X_{t-1}, U^1_t, U^2_t, W_t)$$
$$Y^k_t = h^k_t(X_t, N^k_t)$$
$$U^k_t = g^k_t(\Lambda^k_t, \Delta_t)$$

- Objective

$$\min \mathbb{E}\left\{ \sum_{t=1}^{T} c_t(X_t, U^1_t, U^2_t) \right\}$$

History of the problem

- **Witsenhausen, 1971** proposed the n-DSIS and **asserted** a structural result

$$U_t^k = g_t^k(\Lambda_t^k, \Theta_t)$$

where $\Theta_t = \Pr(X_{t-n} \mid \Delta_t)$.

> The domain of Δ_t increases with time, the domain of Θ_t does not.

- **Varaiya and Walrand, 1979** proved that Witsenhausen's assertion is **true for $n = 1$** but **false of $n > 1$**

What is the structure of optimal controllers for DSIS?

(Open problem for 39 years)

Our Results

Derive two structural results

- $U_t^k = g_t^k(\Lambda_t^k, \Pi_t)$, where $\Pi_t = \Pr(X_{t-1}, \Lambda_t^1, \Lambda_t^2 \mid \Delta_t)$
- $U_t^k = g_t^k(\Lambda_t^k, \Theta_t, r_t^1, r_t^2)$, where $\Theta_t = \Pr(X_{t-n} \mid \Delta_t)$ and r_t^k is a collection of partial functions of the previous $n-1$ control laws of each controller.

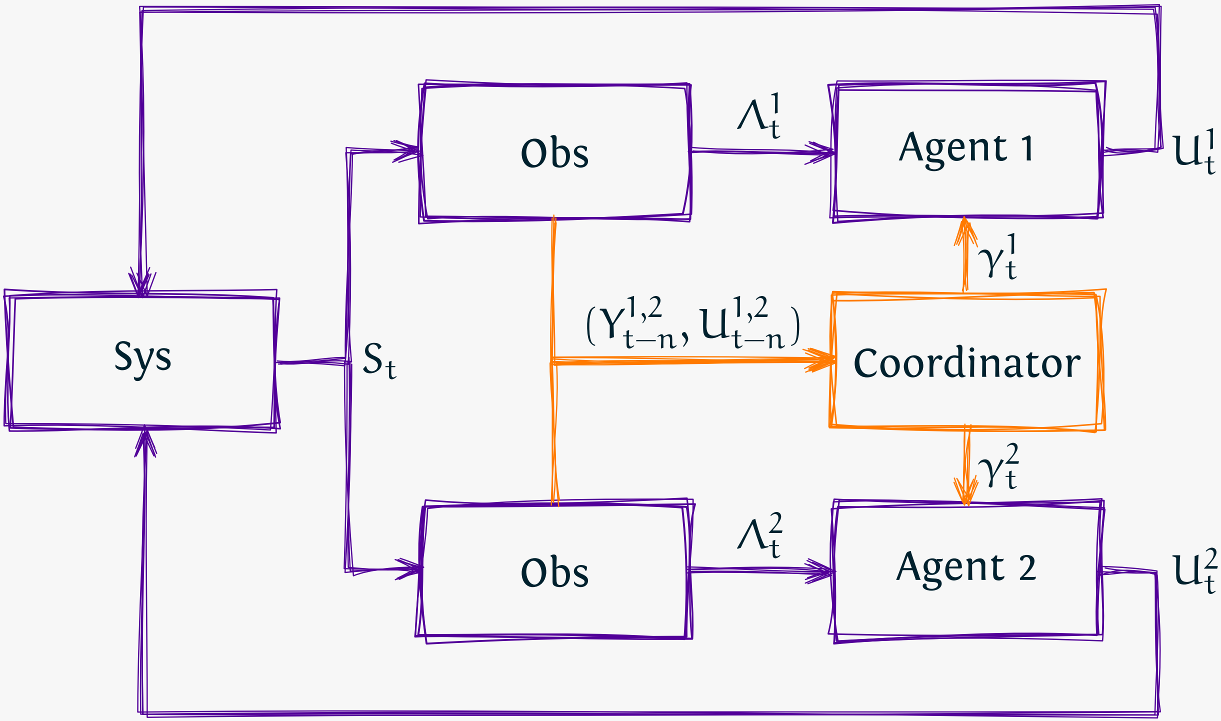
$$r_t^k = \{ \, g_{m+n}^k(\cdot, Y_{m+1:t-n}^k, U_{m+1:t-1}^k, \Delta_{m+n}), \; m = t-2n+1, \dots, t-n-1 \, \}$$

Both Π_t and (Θ_t, r_t^1, r_t^2) have time-invariant domains

First Solution Approach

- Consider a **coordinator** that observes common information Δ_t (but does not observe the private information $(\Lambda_t^1, \Lambda_t^2)$).
- Formulate a **centralized optimization problem** from the point of view of the coordinator
- Show that the coordinator's problem is equivalent to the original problem
- Find **states** sufficient for **input-output mapping** for the **coordinator**
- Find **information states** (state sufficient for dynamic programming) for the coordinator

A coordinator for system



$$(\gamma_t^1, \gamma_t^2) = \psi_t(\Delta_t)$$

$$U_t^1 = \gamma_t^1(\Lambda_t^1) \quad U_t^2 = \gamma_t^2(\Lambda_t^2)$$

State sufficient for input-output mapping

Define: $S_t = (X_{t-1}, \Lambda_t^1, \Lambda_t^2)$

- **Recursive update**:

$$S_{t+1} = \hat{f}_t(S_t, \gamma_t^1, \gamma_t^2, W_t, N_t^1, N_t^2)$$

- **Observation function**:

$$(Y_{t-n}^1, Y_{t-n}^2, U_{t-n}^1, U_{t-n}^2) = \hat{h}_t(S_t)$$

- **Cost can be written in terms of state**

$$c_t(X_t, U_t^1, U_t^2) = \hat{c}_t(S_t, S_{t+1}, \gamma_t^1, \gamma_t^2)$$

Information State

$$\Pi_t = \Pr(\text{state} \mid \text{past data})$$
$$= \Pr(S_t \mid \Delta_t, \gamma^1_{1:t}, \gamma^2_{1:t})$$

- **Recursive update**:

$$\pi_{t+1} = \tilde{f}_t(\pi_t, \gamma^1_t, \gamma^2_t, (Y^1_{t-n}, Y^2_{t-n}, U^1_{t-n}, U^2_{t-n}))$$

- **Controlled Markov process**:

$$\Pr(\Pi_{t+1} \mid \Delta_t, \Pi_{1:t}, \gamma^1_{1:t}, \gamma^2_{1:t}) = \Pr(\Pi_{t+1} \mid \Pi_t, \gamma^1_t, \gamma^2_t)$$

- **Expected cost**:

$$\mathbf{E}\{\hat{c}_t(S_t, S_{t+1}, \gamma^1_t, \gamma^2_t) \mid \Delta_t, \Pi_{1:t}, \gamma^1_{1:t}, \gamma^2_{1:t}\} = \tilde{c}_t(\Pi_t, \gamma^1_t, \gamma^2_t)$$

First structural result

- For the coordinator's problem

$$(\gamma^1_t, \gamma^2_t) = \psi_t(\Pi_t)$$

- For the original problem

$$U^k_t = g^k_t(\Lambda^k_t, \Pi_t) = \gamma^k_t(\Pi_t)(\Lambda^k_t)$$

- dynamic programming decomposition

$$V_T(\pi) = \min_{(\gamma^1, \gamma^2)} \tilde{c}_T(\pi, \gamma^1, \gamma^2)$$

$$V_t(\pi) = \min_{(\gamma^1, \gamma^2)} \Big[\tilde{c}_T(\pi, \gamma^1, \gamma^2) + \mathbf{E}\big\{ V_{t+1}(\Pi_{t+1}) \mid \pi, \gamma^1, \gamma^2 \big\} \Big]$$

Features of the solution

$$\Pi_t = \Pr(X_{t-1}, \Lambda^1_t, \Lambda^2_t \mid \Delta_t \gamma^1_{1:t}, \gamma^2_{1:t})$$

- Π_t has **time invariant domain**

- Π_t is **not independent of the agent's policies** (it is independent of the coordinator's policies)

- In each step of the dynamic program, we are choosing partial **functions** (γ^1_t, γ^2_t).

Features of the solution

$$\Pi_t = \Pr(X_{t-1}, \Lambda_t^1, \Lambda_t^2 \mid \Delta_t \gamma_{1:t}^1, \gamma_{1:t}^2)$$

- Π_t has **time invariant domain**
- Π_t is **not independent of the agent's policies** (it is independent of the coordinator's policies)
- In each step of the dynamic program, we are choosing partial **functions** (γ_t^1, γ_t^2).

Can we **exploit** the "partial function is control action" nature of the problem at the coordinator

Affect of functions on future can be compressed by **partially evaluating the function**

Partially evaluating a function

Consider

$$X_{t+1} = f_t(X_t, Y_t)$$

How do we compress the affect (f_t, X_t) affect X_{t+1}?

Partially evaluating a function

Consider

$$X_{t+1} = f_t(X_t, Y_t)$$

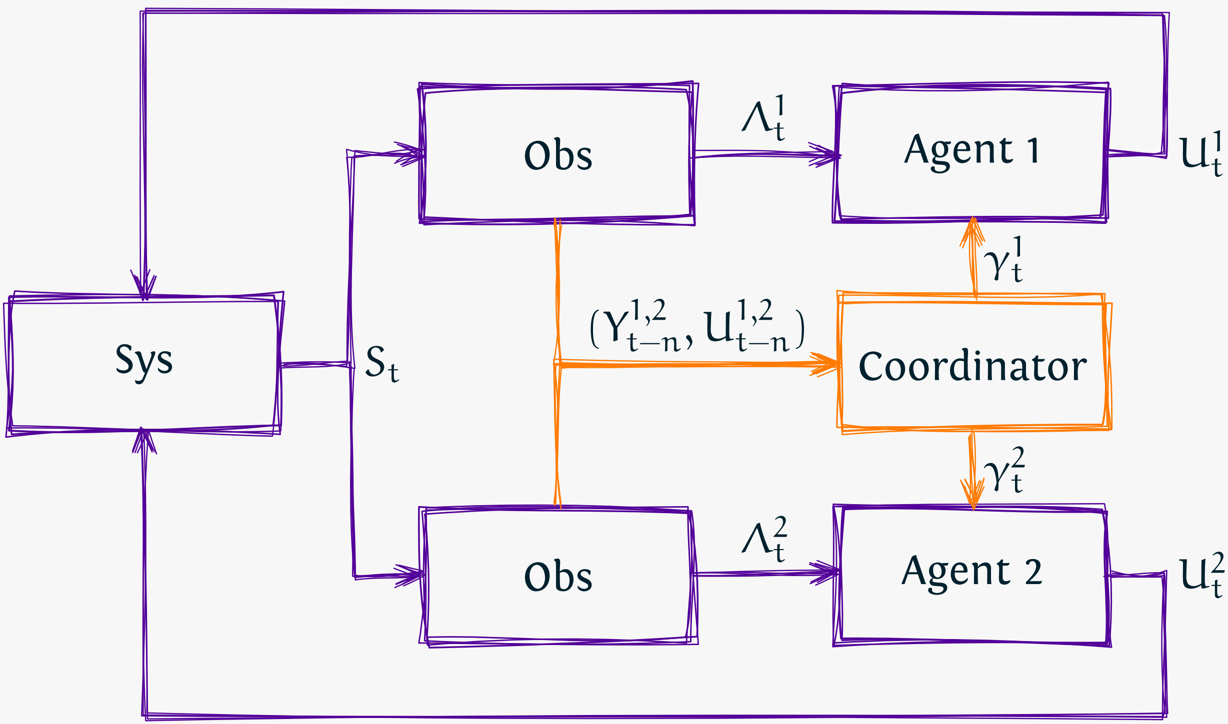
How do we compress the affect (f_t, X_t) affect X_{t+1}?

$$f_t(X_t, \cdot)$$

Second Solution Approach

- Consider a **coordinator** that observes common information Δ_t (but does not observe the private information $(\Lambda_t^1, \Lambda_t^2)$).
- Formulate a **centralized optimization problem** from the point of view of the coordinator
- Show that the coordinator's problem is equivalent to the original problem
- **Compress** the information at the coordinator into **control law independent** part and **partially evaluating** past control laws.

A coordinator for system



$$(\gamma_t^1, \gamma_t^2) = \psi_t(\Delta_t)$$

$$U_t^1 = \gamma_t^1(\Lambda_t^1) \quad U_t^2 = \gamma_t^2(\Lambda_t^2)$$

Information state for optimization

Define: $\Theta_t = \Pr(X_{t-n} \mid \Delta_t)$

$r^k_t = \{\, g^k_{m+n}(\cdot, Y^k_{m+1:t-n}, U^k_{m+1:t-1}, \Delta_{m+n}),\ m = t-2n+1, \dots, t-n-1 \,\}$

- **Recursive update**:

$$\Theta_t = Q_t(\Theta_t, Y^1_{t-n+1}, Y^2_{t-n+1}, U^1_{t-n+1}, U^2_{t-n+1})$$

$$r^k_{t+1} = Q^k_t(r^k_t, Y^1_{t-n+1}, Y^2_{t-n+1}, U^1_{t-n+1}, U^2_{t-n+1}, \gamma^k_t)$$

- **Controlled Markov process**:

$$\Pr(\Theta_{t+1} \mid \Delta_t, \Pi_{1:t}, r^1_{1:t}, r^2_{1:t}, \gamma^1_{1:t}, \gamma^2_{1:t}) = \Pr(\Theta_{t+1} \mid \Delta_t, r^1_t, r^2_t, \gamma^1_t, \gamma^2_t)$$

- **Expected cost**

$$\mathbb{E}\{c_t(X_t, U^1_t, U^2_t) | \Delta_t r^1_{1:t}, r^2_{1:t}, \gamma^1_{1:t}, \gamma^2_{1:t}\} = \hat{c}_t(\Theta_t, r^1_t, r^2_t, \gamma^1_t, \gamma^2_t)$$

Second structural result

- For the coordinator's problem

$$(\gamma^1_t, \gamma^2_t) = \psi_t(\Theta_t, r^1_t, r^2_t)$$

- For the original problem

$$U^k_t = g^k_t(\Lambda^k_t, \Theta_t, r^1_t, r^2_t) = \gamma^k_t(\Theta_t, r^1_t, r^2_t)(\Lambda^k_t)$$

- dynamic programming decomposition

$$V_T(\theta, r^1, r^2) = \min_{(\gamma^1,\gamma^2)} \hat{c}_T(\theta, r^1, r^2, \gamma^1, \gamma^2)$$

$$V_t(\theta, r^1, r^2) = \min_{(\gamma^1,\gamma^2)} \Big[\hat{c}_T(\theta, r^1, r^2, \gamma^1, \gamma^2) + \mathbf{E}\left\{ V_{t+1}(\Theta_{t+1}, r^1_{t+1}, r^2_{t+1}) \mid \theta, r^1, r^2, \gamma^1, \gamma^2 \right\} \Big]$$

Features of the solution

$$\Theta_t = \Pr(X_{t-n} \mid \Delta_t)$$

$$r_t^k = \{\, g_{m+n}^k(\cdot, Y_{m+1:t-n}^k, U_{m+1:t-1}^k, \Delta_{m+n}),\ \ m = t-2n+1, \ldots, t-n-1 \,\}$$

- (Θ_t, r_t^1, r_t^2) has **time invariant domain**
- Θ_t is **independent of the agent's policies**
- The past control laws affect the information state only through (r_t^1, r_t^2).
- In each step of the dynamic program, we are choosing partial **functions** (γ_t^1, γ_t^2).

Summary of approach

Solution Methodology

- Find **common information** at each time
- Look at the problem for the point of view of a **coordinator** that observes this common information and choose **partial functions**
- Find an **information state** for the problem at the coordinator
  - Pr(state for input-output mapping | common information)
  - (Pr(past state | common information),
    past partial control laws)

Salient Features

- Information state has time invariant domain

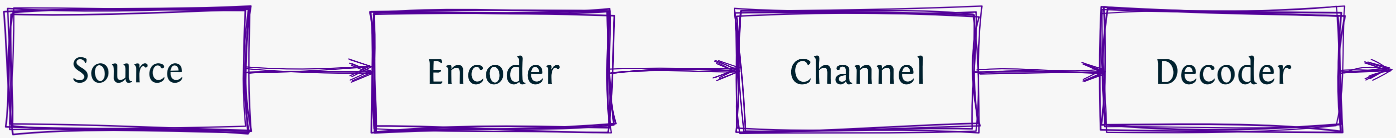
The methodology is also applicable
to infinite horizon problems

- Each step of DP is a functional optimization problem
  - Form of the DP is similar to that of POMDP

Methodology applicable to general problems

- General two-agent teams (M, *Sequential decomposition of sequential teams,* 2008)
- Sufficient conditions for sequential decomposition of dynamic teams (M, Nayyar, and Teneketzis, *Identifying tractable decentralized problems on the basis of information structures,* 2008) (1st set of general conditions in the last 35 years)
- Automated tools to derive structural results for sequential teams (M and Tatikonda, *Sequential team form and its simplification using graphical models,* 2009)
- **Applications**:
  - Real-time communication (M and Teneketzis, 2008, 2009)
  - Control over noisy channels (M and Teneketzis 2009)
  - Decentralized sequential detection (Nayyar and Teneketzis 2009)
  - Multi-terminal communication (M 2009, Nayyar and Teneketzis 2009

Real-time communication

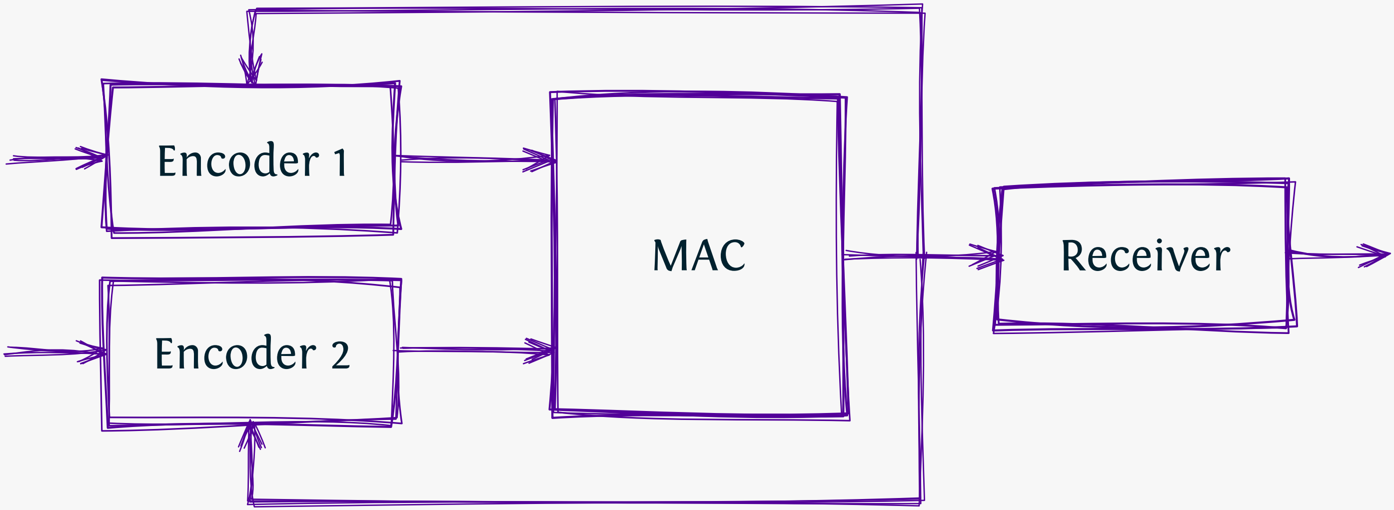


- Communication with zero-delay or fixed finite delay.
- noisy communication channels

- Structure of optimal encoding and decoding strategies
- Sequential decomposition to find optimal strategies

M and Teneketzis, *Optimal design of real-time communication*, IT-2009.

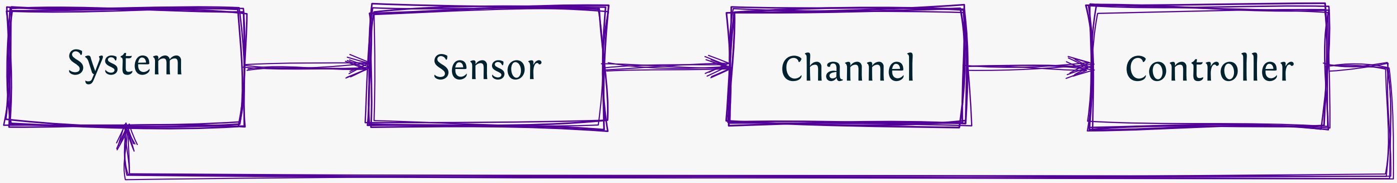
Block Markov superposition codes for multiple access channel



- Use block Markov coding scheme that decode with a finite delay
- Structure of optimal sequential transmission systems

M, *Block Markov superposition coding schemes for MAC with feedback*, ITA-2010.

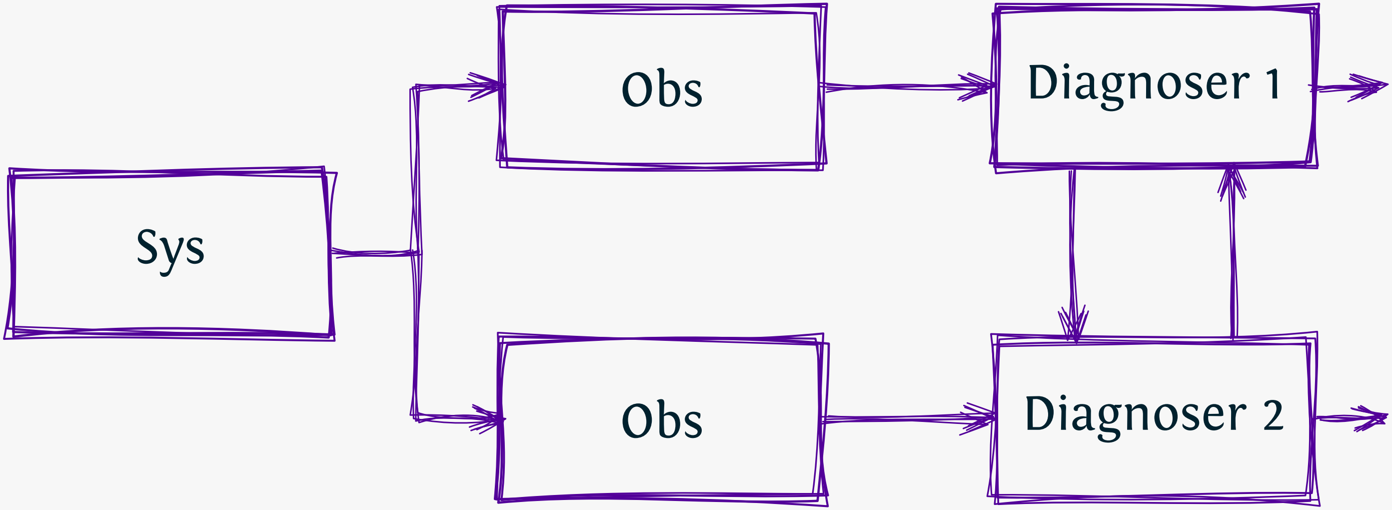
Optimal control over noisy channels



- Sensor and controller are connected over **noisy** communication channel
- Optimize performance (minimize total cost)
- Structure of optimal sensor and controller strategies
- Sequential decomposition to find optimal strategies

M and Teneketzis, *Optimal performance of networked control systems with non-classical information structures*, SICON, 2009

Decentralized diagnosis with communication



- decentralized diagnosers that can communicate information.
- Modeled as discrete event systems: non-sequential and non-probabilistic

Conclusion

Conclusion

- Important concepts
  - coordinator
  - information state
  - partial functions
- Axiomatic approach
  - Insights can be generalized
  - Solution can be **automated**
  - Developing a software to algorithmically identify structural results

    http://pantheon.yale.edu/~am894/code/teams/

Reflections

- Non-sequential information structures
  - **Conceptual difficulties**
  - Computational difficulties
- Provides high-level design guidelines
  - The optimal solution needs to computed **numerically**
  - Provides some design insights: structural properties, which modeling assumption makes the problem easier, etc.
- Actual solution requires "domain knowledge"

Domain knowledge tells us how to approximate a model.

Stochastic control tells us which simplification of the model **makes the overall design easier**

Thank you