Optimal decentralized stochastic control: The designer and common information approaches

Aditya Mahajan

McGill University

Joint work: Ashutosh Nayyar (UIUC/UC Berkeley) and Demosthenis Teneketzis (Univ of Michigan)

CEA-EDF-INRIA Summer School on Stochastic Optimization, Cadarache, June 25 – July 6, 2012

Common theme:

multi-stage multi-agent decision making under uncertainty

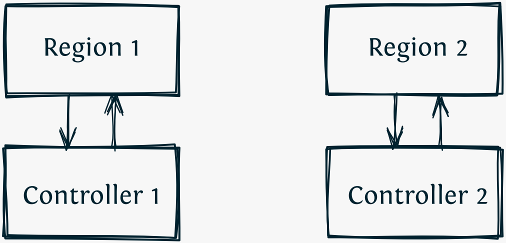
Interconnected Power Systems

Interconnected Power Systems

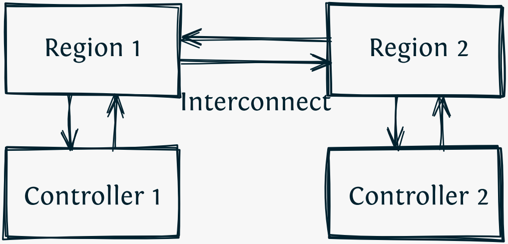
Region 1

Region 2

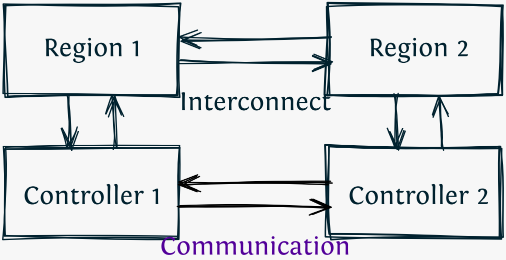
Interconnected Power Systems



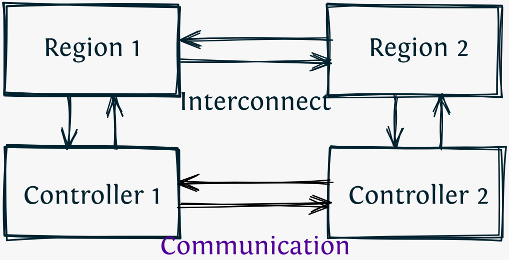
Interconnected Power Systems



Interconnected Power Systems



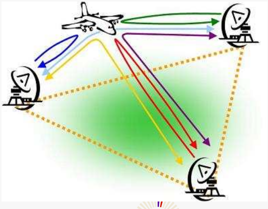
Interconnected Power Systems



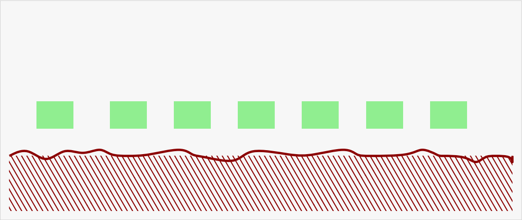
Challenges

- How to **coordinate**?
- When, what, and how to **communicate**?

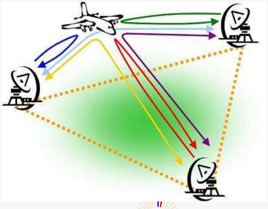
Sensor and Surveillance Networks



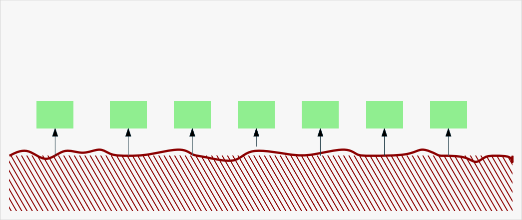
Sensor and Surveillance Networks



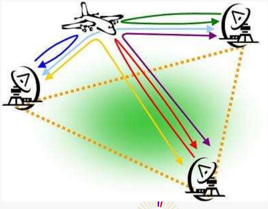
Limited resources



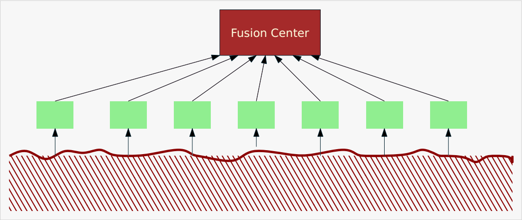
Sensor and Surveillance Networks

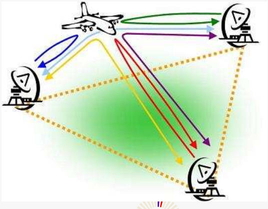


Limited resources   Noisy observations



Sensor and Surveillance Networks

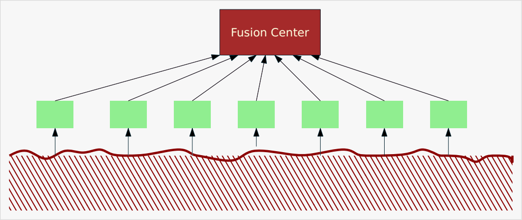


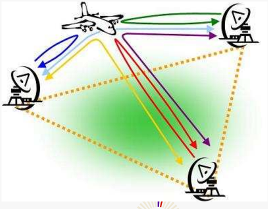


Limited resources Noisy observations

Communication

Sensor and Surveillance Networks



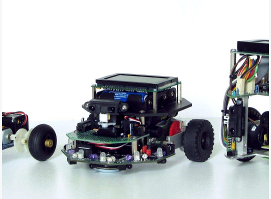


Limited resources Noisy observations
Communication

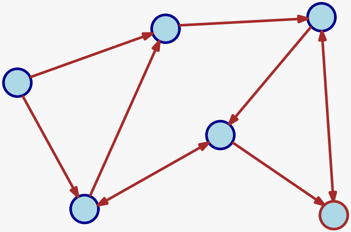
Challenges

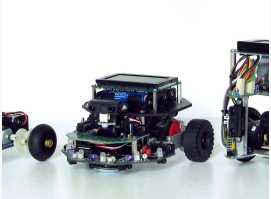
- Real-time communication
- Scheduling measurements and communication
- Detect node failures

Networked Control Systems

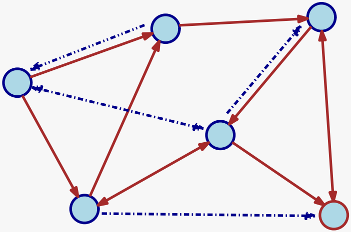


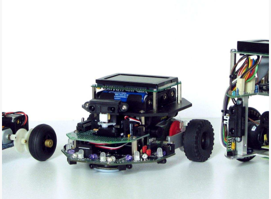
Networked Control Systems



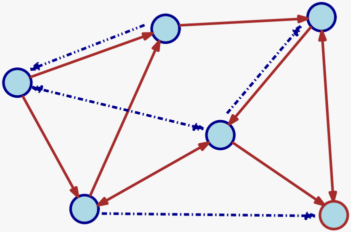


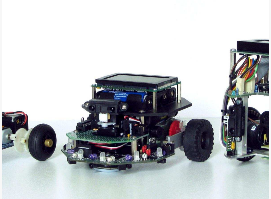
Networked Control Systems





Networked Control Systems

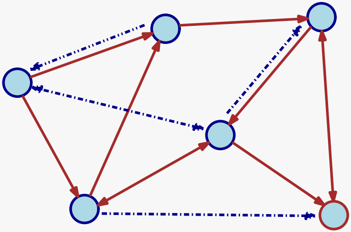


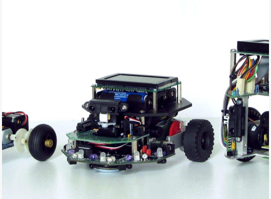


Challenges

- Control and communication over networks
  (internet $\Rightarrow$ delay, wireless $\Longrightarrow$ losses)

Networked Control Systems

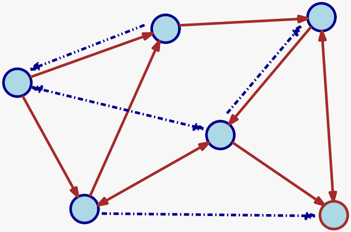


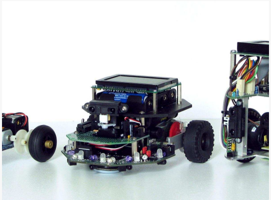


Challenges

- Control and communication over networks (internet $\Rightarrow$ delay, wireless $\Longrightarrow$ losses)
- Distributed estimation

Networked Control Systems





Challenges

- Control and communication over networks
  (internet $\Rightarrow$ delay, wireless $\Longrightarrow$ losses)
- Distributed estimation
- Distributed learning

Salient features in decentralized decision making

Multiple decision makers

Decisions made by multiple controllers in a stochastic environment

Salient features in decentralized decision making

Multiple decision makers

Decisions made by multiple controllers in a stochastic environment

Coordination issues

All controllers must coordinate to achieve a system-wide objective

Salient features in decentralized decision making

Multiple decision makers

Decisions made by multiple controllers in a stochastic environment

Coordination issues

All controllers must coordinate to achieve a system-wide objective

Communication issues

Controllers can communicate either directly or indirectly

Salient features in decentralized decision making

Multiple decision makers

Decisions made by multiple controllers in a stochastic environment

Coordination issues

All controllers must coordinate to achieve a system-wide objective

Communication issues

Controllers can communicate either directly or indirectly

Robustness

System model may not be completely known

Classification of decentralized systems

Controllers/agents are coupled in two ways:

1. Coupling due to cost/utility
2. Coupling due to dynamics

Classification of decentralized systems

Controllers/agents are coupled in two ways:

1. Coupling due to cost/utility
2. Coupling due to dynamics

Decentralized systems may be classified according to:

Classification of decentralized systems

Controllers/agents are coupled in two ways:

1. Coupling due to cost/utility
2. Coupling due to dynamics

Decentralized systems may be classified according to:

1. Objective

   Team vs Games

Classification of decentralized systems

Controllers/agents are coupled in two ways:

1. Coupling due to cost/utility
2. Coupling due to dynamics

Decentralized systems may be classified according to:

1. Objective
   Team vs Games
2. Dynamics
   Static vs Dynamic

Classification of decentralized systems

Controllers/agents are coupled in two ways:

1. Coupling due to cost/utility
2. Coupling due to dynamics

Decentralized systems may be classified according to:

1. Objective

   Team vs Games

2. Dynamics

   Static vs Dynamic

This talk will focus on Dynamic Teams

Classification of decentralized systems

Controllers/agents are coupled in two ways:

1. Coupling due to cost/utility
2. Coupling due to dynamics

Decentralized systems may be classified according to:

1. Objective

   Team vs Games

2. Dynamics

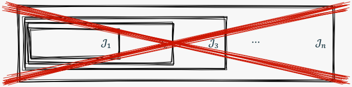
   Static vs Dynamic

This talk will focus on Dynamic Teams

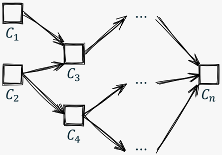
- Studied in economics and systems and control since the mid 50s.
- Unlike games, agents have no incentive to cheat.
- Instead of equilibrium, we seek globally optimal strategies.

Key features of decentralized teams

- Non-Classical information structure



- Fixed **partial** order in which agents act



Why is decentralized stochastic control difficult?

An example of centralized static optimization

$$P = [\ \bullet \quad \bullet \quad \bullet \quad \bullet\ ]$$

ω_1	ω_2	ω_3	ω_4

An example of centralized static optimization

$$P = [\ \bullet \quad \bullet \quad \bullet \quad \bullet\ ]$$

	ω_1	ω_2	ω_3	ω_4
$x =$	1	1	2	2

An example of centralized static optimization

$$P = [\ \bullet \quad \bullet \quad \bullet \quad \bullet\ ]$$

	ω_1	ω_2	ω_3	ω_4
$x =$	1	1	2	2

$$u = g(x) \in \{1, 2, 3\}$$

An example of centralized static optimization

$$P = [\ \bullet \quad \bullet \quad \bullet \quad \bullet\ ]$$

	ω_1	ω_2	ω_3	ω_4
$x =$	1	1	2	2

$$u = g(x) \in \{1, 2, 3\}$$

$c(\omega, u)$

	ω_1	ω_2	ω_3	ω_4
$u = 1$	•	•	•	•
$u = 2$	•	•	•	•
$u = 3$	•	•	•	•

$$\boxed{J(g) = \mathbb{E}^g[c(\omega, u)]}$$

An example of centralized static optimization

$$P = [\ \bullet\quad \bullet\quad \bullet\quad \bullet\ ]$$

	ω_1	ω_2	ω_3	ω_4
$x =$	1	1	2	2

$$u = g(x) \in \{1, 2, 3\}$$

$c(\omega, u)$

	ω_1	ω_2	ω_3	ω_4
$u = 1$	•	•	•	•
$u = 2$	•	•	•	•
$u = 3$	•	•	•	•

$$J(g) = \mathbb{E}^g[c(\omega, u)]$$

Brute force search $\quad \min_g J(g), \quad |g| = |\mathcal{U}|^{|\mathcal{X}|} = 9$ possibilities.

An example of centralized static optimization

$$P = [\ \bullet \quad \bullet \quad \bullet \quad \bullet\ ]$$

	ω_1	ω_2	ω_3	ω_4
$x =$	1	1	2	2

$$u = g(x) \in \{1, 2, 3\}$$

$c(\omega, u)$

	ω_1	ω_2	ω_3	ω_4
$u = 1$	•	•	•	•
$u = 2$	•	•	•	•
$u = 3$	•	•	•	•

$$J(g) = \mathbb{E}^g[c(\omega, u)]$$

Brute force search $\min_g J(g), \quad |g| = |\mathcal{U}|^{|\mathcal{X}|} = 9$ possibilities.

Systematic search $3 + 3 = 6$ possibilities

$$u_1 = g(1) \qquad\qquad u_2 = g(2)$$

$$\min_{u_1} \mathbb{E}[c(\omega, u_1) \mid x = 1] \qquad\qquad \min_{u_2} \mathbb{E}[c(\omega, u_2) \mid x = 2]$$

An example of centralized static optimization

$P = [\ \bullet\ \ \bullet\ \ \bullet\ \ \bullet\ ]$

	ω_1	ω_2	ω_3	ω_4
$x =$	1	1	2	2

$c(\omega, u)$

	ω_1	ω_2	ω_3	ω_4
$u = 1$	•	•	•	•
$u = 2$	•	•	•	•
$u = 3$	•	•	•	•

$u = g(x) \in \{1, 2, 3\}$

$J(g) = \mathbb{E}^g[c(\omega, u)]$

(functional opt.)

Brute force search $\min_g J(g), \quad |g| = |\mathcal{U}|^{|\mathcal{X}|} = 9$ possibilities.

Systematic search $3 + 3 = 6$ possibilities **(parametric opt.)**

$u_1 = g(1)$ $\qquad$ $u_2 = g(2)$

$\min_{u_1} \mathbb{E}[c(\omega, u_1) \mid x = 1]$ $\qquad$ $\min_{u_2} \mathbb{E}[c(\omega, u_2) \mid x = 2]$

An example of decentralized static optimization

$P = [\ \bullet \quad \bullet \quad \bullet \quad \bullet\ ]$

ω_1	ω_2	ω_3	ω_4

An example of decentralized static optimization

$$P = [\ \bullet \quad \bullet \quad \bullet \quad \bullet\ ]$$

	ω_1	ω_2	ω_3	ω_4
$x =$	1	1	2	2
$y =$	2	1	1	2

An example of decentralized static optimization

$P = [\ \bullet \quad \bullet \quad \bullet \quad \bullet\ ]$

	ω_1	ω_2	ω_3	ω_4
$x =$	1	1	2	2
$y =$	2	1	1	2

$u = g(x) \in \{1, 2, 3\} \quad v = h(y) \in \{1, 2\}$

An example of decentralized static optimization

$P = [\ \bullet \quad \bullet \quad \bullet \quad \bullet\ ]$

	ω_1	ω_2	ω_3	ω_4
$x =$	1	1	2	2
$y =$	2	1	1	2

$u = g(x) \in \{1,2,3\} \quad v = h(y) \in \{1,2\}$

$c(\omega, u, v)$

	ω_1		ω_2		ω_3		ω_4	
$u = 1$	•	•	•	•	•	•	•	•
$u = 2$	•	•	•	•	•	•	•	•
$u = 3$	•	•	•	•	•	•	•	•
$v =$	1	2	1	2	1	2	1	2

$$\boxed{J(g,h) = \mathbb{E}^{g,h}[c(\omega,u,v)]}$$

An example of decentralized static optimization

$P = [\ \bullet \quad \bullet \quad \bullet \quad \bullet\ ]$

	ω_1	ω_2	ω_3	ω_4
$x =$	1	1	2	2
$y =$	2	1	1	2

$u = g(x) \in \{1, 2, 3\} \quad v = h(y) \in \{1, 2\}$

$c(\omega, u, v)$

$$\begin{array}{c|cc|cc|cc|cc|}
 & \omega_1 & & \omega_2 & & \omega_3 & & \omega_4 & \\
u = 1 & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
u = 2 & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
u = 3 & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
v = & 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2
\end{array}$$

$J(g, h) = \mathbb{E}^{g,h}[c(\omega, u, v)]$

Brute force search $\quad \min_{g,h} J(g, h), \quad |g| = |\mathcal{U}|^{|\mathcal{X}|}, |h| = |\mathcal{V}|^{|\mathcal{Y}|},$

$9 \times 4 = 36$ possibilities.

An example of decentralized static optimization

$P = [\ \bullet \quad \bullet \quad \bullet \quad \bullet\ ]$

	ω_1	ω_2	ω_3	ω_4
$x =$	1	1	2	2
$y =$	2	1	1	2

$c(\omega, u, v)$

	ω_1		ω_2		ω_3		ω_4	
$u = 1$	•	•	•	•	•	•	•	•
$u = 2$	•	•	•	•	•	•	•	•
$u = 3$	•	•	•	•	•	•	•	•
$v =$	1	2	1	2	1	2	1	2

$u = g(x) \in \{1, 2, 3\} \quad v = h(y) \in \{1, 2\}$

$$J(g, h) = \mathbb{E}^{g,h}[c(\omega, u, v)]$$

Brute force search $\quad \min_{g,h} J(g), \quad |g| = |\mathcal{U}|^{|\mathcal{X}|},\ |h| = |\mathcal{V}|^{|\mathcal{Y}|},$

$9 \times 4 = 36$ possibilities.

> For one controller/agent to choose an optimal action, it must second guess the other controller's/agent's policy

An example of decentralized static optimization

$P = [\ \bullet \quad \bullet \quad \bullet \quad \bullet\ ]$

	ω_1	ω_2	ω_3	ω_4
$x =$	1	1	2	2
$y =$	2	1	1	2

$c(\omega, u, v)$

	ω_1		ω_2		ω_3		ω_4	
$u = 1$	•	•	•	•	•	•	•	•
$u = 2$	•	•	•	•	•	•	•	•
$u = 3$	•	•	•	•	•	•	•	•
$v =$	1	2	1	2	1	2	1	2

$u = g(x) \in \{1, 2, 3\} \quad v = h(y) \in \{1, 2\}$

$$J(g, h) = \mathbb{E}^{g,h}[c(\omega, u, v)]$$

Orthogonal search

1. Suppose h is fixed: $\min\limits_{u_i} \mathbb{E}^{h}[c(\omega, u_i, v) \mid x = i], \quad i = 1, 2, 3.$
2. Suppose g is fixed: $\min\limits_{v_j} \mathbb{E}^{g}[c(\omega, u, v_j) \mid y = j], \quad j = 1, 2.$

An example of decentralized static optimization

$P = [\ \bullet \quad \bullet \quad \bullet \quad \bullet\ ]$

	ω_1	ω_2	ω_3	ω_4
$x =$	1	1	2	2
$y =$	2	1	1	2

$c(\omega, u, v)$

$$\begin{array}{c|cc|cc|cc|cc|}
 & \multicolumn{2}{c}{\omega_1} & \multicolumn{2}{c}{\omega_2} & \multicolumn{2}{c}{\omega_3} & \multicolumn{2}{c}{\omega_4} \\
u = 1 & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
u = 2 & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
u = 3 & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
v = & 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2
\end{array}$$

$u = g(x) \in \{1, 2, 3\} \quad v = h(y) \in \{1, 2\}$

$J(g, h) = \mathbb{E}^{g,h}[c(\omega, u, v)]$

Orthogonal search yields person-by-person opt strategy

1. Suppose h is fixed: $\min_{u_i} \mathbb{E}^{h}[c(\omega, u_i, v) \mid x = i], \quad i = 1, 2, 3.$
2. Suppose g is fixed: $\min_{v_j} \mathbb{E}^{g}[c(\omega, u, v_j) \mid y = j], \quad j = 1, 2.$

To find globally optimal strategies, in general, we cannot do better than brute force search

An example of centralized multi-stage optimization

ω_1	ω_2	ω_3	ω_4
ω_5	ω_6	ω_7	ω_8

An example of centralized multi-stage optimization

ω_1	ω_2	ω_3	ω_4	$y_1 = 1$
ω_5	ω_6	ω_7	ω_8	$y_1 = 2$

An example of centralized multi-stage optimization

ω_1	ω_2	ω_3	ω_4	$y_1 = 1$
ω_5	ω_6	ω_7	ω_8	$y_1 = 2$

$u_1 = g_1(y_1) \in \{1, 2\}$

An example of centralized multi-stage optimization

ω_1	ω_2	ω_3	ω_4	$y_1 = 1$
ω_5	ω_6	ω_7	ω_8	$y_1 = 2$

$u_1 = g_1(y_1) \in \{1, 2\}$

$u_1 = 1 \implies y_2 =$	1	1	2	2
$u_1 = 1 \implies y_2 =$	1	1	2	2

$u_1 = 2 \implies y_2 =$	1	2	2	1
$u_1 = 2 \implies y_2 =$	1	2	2	1

An example of centralized multi-stage optimization

ω_1	ω_2	ω_3	ω_4	$y_1 = 1$
ω_5	ω_6	ω_7	ω_8	$y_1 = 2$

$u_1 = g_1(y_1) \in \{1, 2\}$

$u_1{=}1 \implies y_2{=}$	1	1	2	2
$u_1{=}1 \implies y_2{=}$	1	1	2	2
$u_1{=}2 \implies y_2{=}$	1	2	2	1
$u_1{=}2 \implies y_2{=}$	1	2	2	1

$u_2 = g_2(y_1, y_2, u_1) \in \{1, 2\}$

An example of centralized multi-stage optimization

ω_1	ω_2	ω_3	ω_4	$y_1 = 1$
ω_5	ω_6	ω_7	ω_8	$y_1 = 2$

$u_1 = g_1(y_1) \in \{1, 2\}$

$u_1{=}1 \implies y_2{=}$	1	1	2	2
$u_1{=}1 \implies y_2{=}$	1	1	2	2
$u_1{=}2 \implies y_2{=}$	1	2	2	1
$u_1{=}2 \implies y_2{=}$	1	2	2	1

$u_2 = g_2(y_1, y_2, u_1) \in \{1, 2\}$

$c_1(\omega, u_1) + c_2(\omega, u_2)$

$$J(g_1, g_2) = \mathbb{E}^{g_1, g_2}[c_1(\omega, u_1) + c_2(\omega, u_2)]$$

An example of centralized multi-stage optimization

	ω_1	ω_2	ω_3	ω_4	$y_1 = 1$	$u_1 = g_1(y_1) \in \{1, 2\}$
	ω_5	ω_6	ω_7	ω_8	$y_1 = 2$	$\boxed{d_1 = \{y_1\}}$
$u_1{=}1 \implies y_2{=}$	1	1	2	2		$u_2 = g_2(y_1, y_2, u_1) \in \{1, 2\}$
$u_1{=}1 \implies y_2{=}$	1	1	2	2		$\boxed{d_2 = \{y_1, y_2, u_1\}}$
$u_1{=}2 \implies y_2{=}$	1	2	2	1		
$u_1{=}2 \implies y_2{=}$	1	2	2	1		$c_1(\omega, u_1) + c_2(\omega, u_2)$

$$J(g_1, g_2) = \mathbb{E}^{g_1, g_2}[c_1(\omega, u_1) + c_2(\omega, u_2)]$$

An example of centralized multi-stage optimization

ω_1	ω_2	ω_3	ω_4	$y_1 = 1$
ω_5	ω_6	ω_7	ω_8	$y_1 = 2$

$u_1{=}1 \implies y_2{=}$	1	1	2	2
$u_1{=}1 \implies y_2{=}$	1	1	2	2

$u_1{=}2 \implies y_2{=}$	1	2	2	1
$u_1{=}2 \implies y_2{=}$	1	2	2	1

$u_1 = g_1(y_1) \in \{1, 2\}$

$\boxed{d_1 = \{y_1\}}$

$u_2 = g_2(y_1, y_2, u_1) \in \{1, 2\}$

$\boxed{d_2 = \{y_1, y_2, u_1\}}$

$c_1(\omega, u_1) + c_2(\omega, u_2)$

$$J(g_1, g_2) = \mathbb{E}^{g_1, g_2}[c_1(\omega, u_1) + c_2(\omega, u_2)]$$

Critical Assumption: Centralized information $\boxed{d_1 \subseteq d_2}$

Solution approach for centralized multi-stage optimization

Brute force search $\displaystyle \min_{g_1, g_2} J(g_1, g_2).$

$$|g_1| = |\mathcal{U}_1|^{|\mathcal{Y}_1|}, \quad |g_2| = |\mathcal{U}_2|^{|\mathcal{Y}_1| \times |\mathcal{Y}_2| \times |\mathcal{U}_1|}. \quad 2^2 \times 2^8 = 1024 \text{ possiblities.}$$

Solution approach for centralized multi-stage optimization

Brute force search

$$\min_{g_1, g_2} J(g_1, g_2).$$

$$|g_1| = |\mathcal{U}_1|^{|\mathcal{Y}_1|}, \quad |g_2| = |\mathcal{U}_2|^{|\mathcal{Y}_1| \times |\mathcal{Y}_2| \times |\mathcal{U}_1|}. \quad 2^2 \times 2^8 = 1024 \text{ possiblities.}$$

Dynamic programming decomposition

$$V_2(d_2) = \min_{u_2} \mathbb{E}[c_2(\omega, u_2) \mid d_2, u_2]$$

$$V_1(d_1) = \min_{u_1} \mathbb{E}[c_1(\omega, u_1) + V_2(d_2) \mid d_1, u_1]$$

Solution approach for centralized multi-stage optimization

Brute force search $\min_{g_1, g_2} J(g_1, g_2).$ **(functional opt.)**

$$|g_1| = |\mathcal{U}_1|^{|\mathcal{Y}_1|}, \quad |g_2| = |\mathcal{U}_2|^{|\mathcal{Y}_1| \times |\mathcal{Y}_2| \times |\mathcal{U}_1|}. \quad 2^2 \times 2^8 = 1024 \text{ possiblities.}$$

Dynamic programming decomposition **(parametric opt.)**

$$V_2(d_2) = \min_{u_2} \mathbb{E}[c_2(\omega, u_2) \mid d_2, u_2]$$

$$V_1(d_1) = \min_{u_1} \mathbb{E}[c_1(\omega, u_1) + V_2(d_2) \mid d_1, u_1]$$

Solution approach for centralized multi-stage optimization

Brute force search $\min_{g_1, g_2} J(g_1, g_2).$ **(functional opt.)**

$$|g_1| = |\mathcal{U}_1|^{|\mathcal{Y}_1|}, \quad |g_2| = |\mathcal{U}_2|^{|\mathcal{Y}_1| \times |\mathcal{Y}_2| \times |\mathcal{U}_1|}. \quad 2^2 \times 2^8 = 1024 \text{ possiblities.}$$

Dynamic programming decomposition **(parametric opt.)**

$$V_2(d_2) = \min_{u_2} \mathbb{E}[c_2(\omega, u_2) \mid d_2, u_2]$$

$$V_1(d_1) = \min_{u_1} \mathbb{E}[c_1(\omega, u_1) + V_2(d_2) \mid d_1, u_1]$$

- Step 1 works because $\mathbb{P}(\omega \mid d_2)$ **does not** depend on g_1.
- Step 2 works because $\mathbb{P}(d_2 \mid d_1, u_1)$ **does not** depend on g_1.

Solution approach for centralized multi-stage optimization

Brute force search $\min_{g_1, g_2} J(g_1, g_2).$ **(functional opt.)**

$$|g_1| = |\mathcal{U}_1|^{|\mathcal{Y}_1|}, \quad |g_2| = |\mathcal{U}_2|^{|\mathcal{Y}_1| \times |\mathcal{Y}_2| \times |\mathcal{U}_1|}. \quad 2^2 \times 2^8 = 1024 \text{ possiblities.}$$

Dynamic programming decomposition **(parametric opt.)**

$$V_2(d_2) = \min_{u_2} \mathbb{E}[c_2(\omega, u_2) \mid d_2, u_2]$$

$$V_1(d_1) = \min_{u_1} \mathbb{E}[c_1(\omega, u_1) + V_2(d_2) \mid d_1, u_1]$$

- Step 1 works because $\mathbb{P}(\omega \mid d_2)$ **does not** depend on g_1.
- Step 2 works because $\mathbb{P}(d_2 \mid d_1, u_1)$ **does not** depend on g_1.
- $\boxed{\text{Both steps work because } d_1 \subseteq d_2}$

An example of decentralized multi-stage optimization

ω_1	ω_2	ω_3	ω_4	$y_1 = 1$
ω_5	ω_6	ω_7	ω_8	$y_1 = 2$

$$u_1 = g_1(y_1) \in \{1, 2\}$$

$$\boxed{d_1 = \{y_1\}}$$

$u_1{=}1 \implies y_2{=}$	1	1	2	2
$u_1{=}1 \implies y_2{=}$	1	1	2	2

$$u_2 = g_2(y_2) \in \{1, 2\}$$

$$\boxed{d_2 = \{y_2\}}$$

$u_1{=}2 \implies y_2{=}$	1	2	2	1
$u_1{=}2 \implies y_2{=}$	1	2	2	1

$$c_1(\omega, u_1) + c_2(\omega, u_2)$$

$$J(g_1, g_2) = \mathbb{E}^{g_1, g_2}[c_1(\omega, u_1) + c_2(\omega, u_2)]$$

An example of decentralized multi-stage optimization

ω_1	ω_2	ω_3	ω_4	$y_1 = 1$	$u_1 = g_1(y_1) \in \{1, 2\}$
ω_5	ω_6	ω_7	ω_8	$y_1 = 2$	$\boxed{d_1 = \{y_1\}}$

$u_1{=}1 \implies y_2{=}$	1	1	2	2	$u_2 = g_2(y_2) \in \{1, 2\}$
$u_1{=}1 \implies y_2{=}$	1	1	2	2	$\boxed{d_2 = \{y_2\}}$
$u_1{=}2 \implies y_2{=}$	1	2	2	1	
$u_1{=}2 \implies y_2{=}$	1	2	2	1	$c_1(\omega, u_1) + c_2(\omega, u_2)$

$$J(g_1, g_2) = \mathbb{E}^{g_1, g_2}[c_1(\omega, u_1) + c_2(\omega, u_2)]$$

Critical Assumption: Decentralized information $\boxed{d_1 \not\subseteq d_2}$

Can we do better than brute force search?

Usual Dynamic programming does not work?

$$V_2(d_2) \stackrel{?}{=} \min_{u_2} \mathbb{E}^{g_1}[c_2(\omega, u_2) \mid d_2, u_2]$$

$$V_1(d_1) \stackrel{?}{=} \min_{u_1} \mathbb{E}^{g_1}[c_1(\omega, u_1) + V_2(d_2) \mid d_1, u_1]$$

Usual Dynamic programming does not work?

$$V_2(d_2) \stackrel{?}{=} \min_{u_2} \mathbb{E}^{g_1}[c_2(\omega, u_2) \mid d_2, u_2]$$

$$V_1(d_1) \stackrel{?}{=} \min_{u_1} \mathbb{E}^{g_1}[c_1(\omega, u_1) + V_2(d_2) \mid d_1, u_1]$$

A sequential decomposition is possible (Witsenhausen, 1973)

Define $\pi_t = \mathbb{P}(\omega \mid g_{1:t-1})$.

$$V_t(\pi_t) = \min_{g_t} \; \mathbb{E}^{g_t}[c_t(\omega, u_t) + V_{t+1}(\pi_{t+1}) \mid \pi_t]$$

But, the worst case complexity remains the same.

Finding optimal strategies

Can we obtain a systematic approach to find optimal strategies that does better than brute force search?

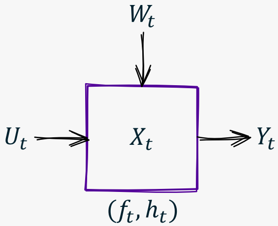
Finding optimal strategies

Can we obtain a systematic approach to find optimal strategies that does better than brute force search?

- Designer approach
- Common information approach

Designer approach

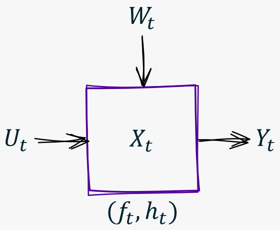
Block notation



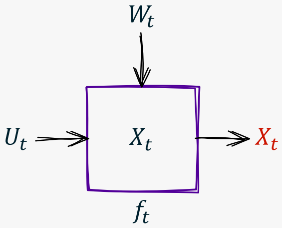
$$Y_t = h_t(X_t, U_t, W_t)$$

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

Block notation

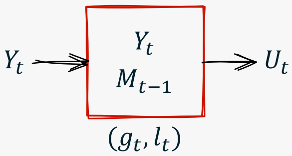


$$Y_t = h_t(X_t, U_t, W_t)$$
$$X_{t+1} = f_t(X_t, U_t, W_t)$$



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

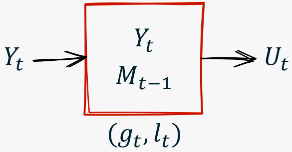
Block notation



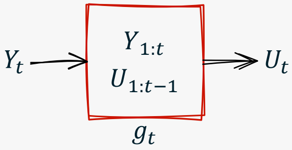
$$U_t = g_t(Y_t, M_{t-1})$$

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

Block notation

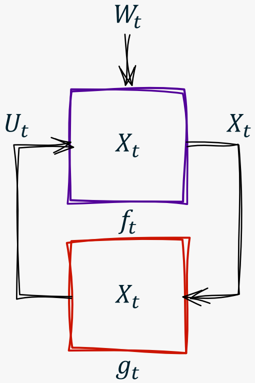


$$U_t = g_t(Y_t, M_{t-1})$$
$$M_t = l_t(Y_t, M_{t-1})$$

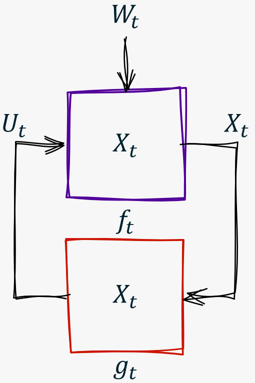


$$U_t = g_t(Y_{1:t}, U_{1:t-1})$$

Another look at dynamic consistency



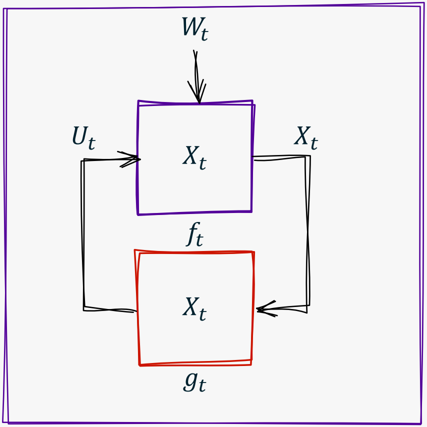
Another look at dynamic consistency



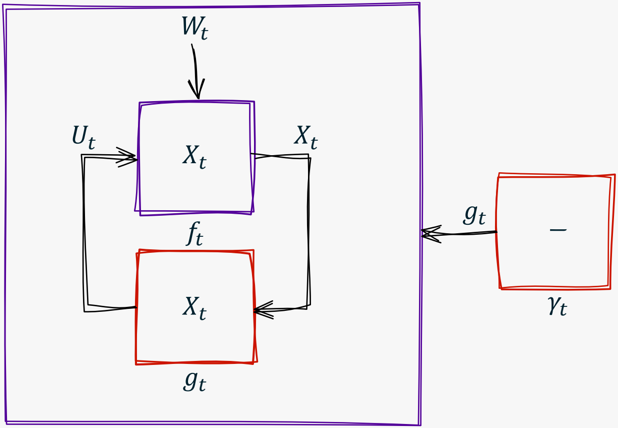
Define: $\pi_t = \mathbb{P}(X_t | g_{1:t-1})$. Then,

$$V_t(\pi_t) = \min_{g_t} \mathbb{E}[c(X_t, U_t) + V_{t+1}(L^g \pi_t)]$$

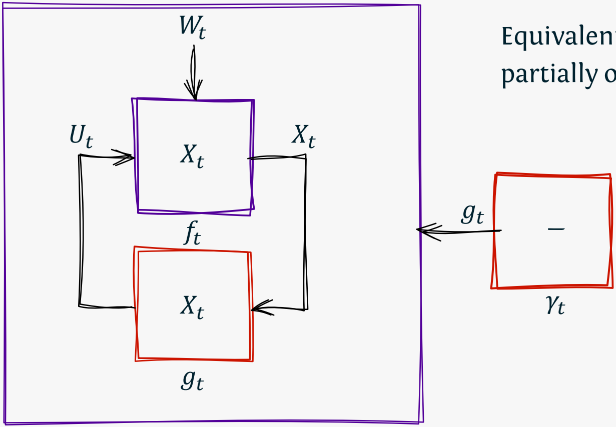
An alternative derivation



An alternative derivation

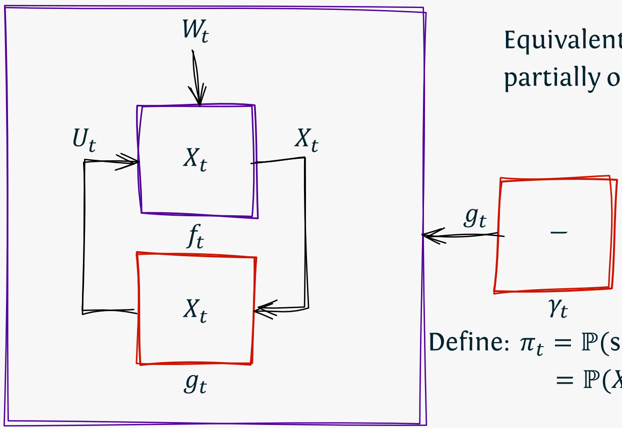


An alternative derivation



Equivalent to a **centralized** partially observed system

An alternative derivation

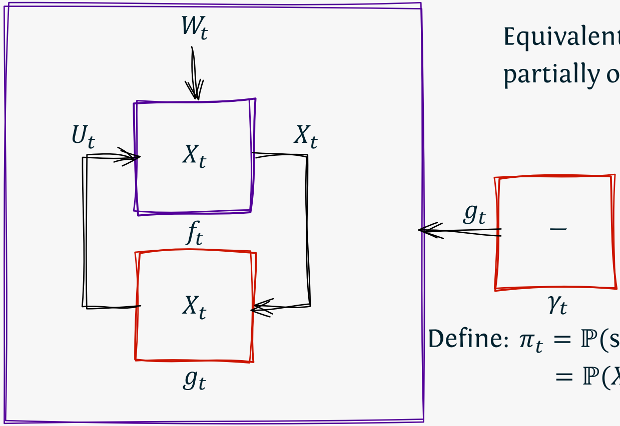


Equivalent to a **centralized** partially observed system

Define: $\pi_t = \mathbb{P}(\text{state} \mid \text{history of data})$

$= \mathbb{P}(X_t \mid g_{1:t-1})$

An alternative derivation

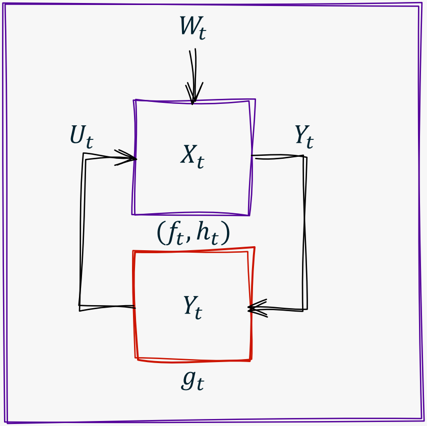


Equivalent to a **centralized** partially observed system

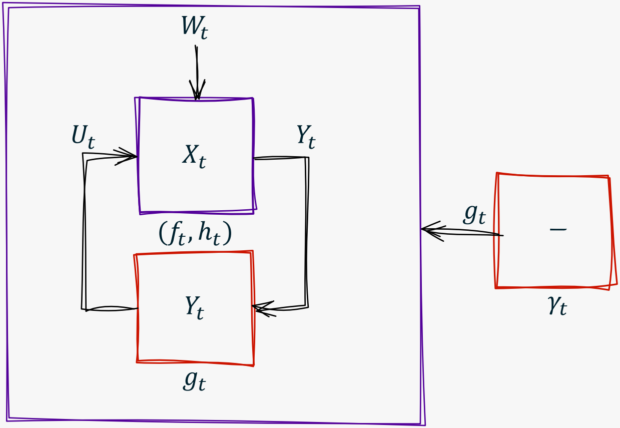
Define: $\pi_t = \mathbb{P}(\text{state} \mid \text{history of data})$
$= \mathbb{P}(X_t \mid g_{1:t-1})$

$$V_t(\pi_t) = \min_{g_t} \mathbb{E}[c(X_t, U_t) + V_{t+1}(L^g \pi_t)]$$

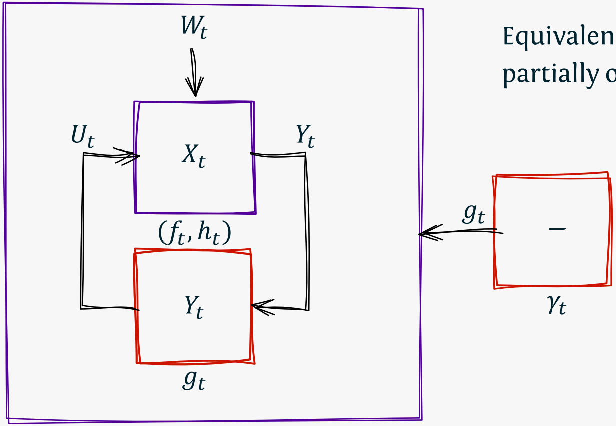
Generalization to decentralized system: finite memory controller



Generalization to decentralized system: finite memory controller

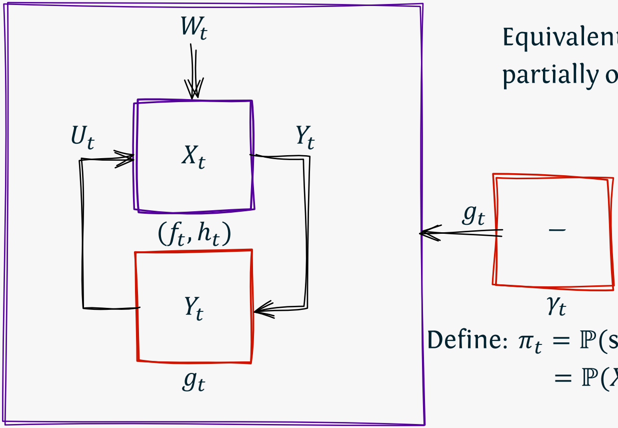


Generalization to decentralized system: finite memory controller



Equivalent to a **centralized** partially observed system

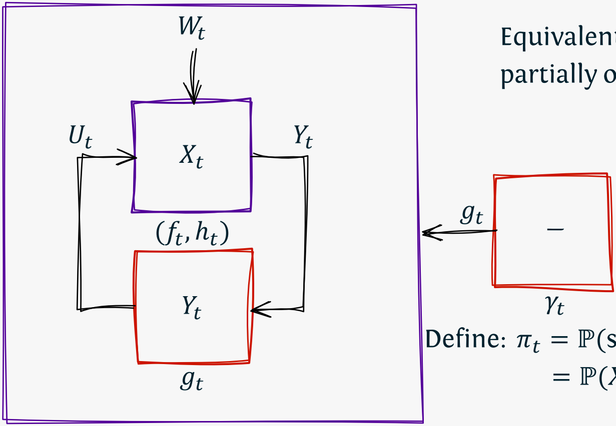
Generalization to decentralized system: finite memory controller



Equivalent to a **centralized** partially observed system

Define: $\pi_t = \mathbb{P}(\text{state} \mid \text{history of data})$
$= \mathbb{P}(X_t, Y_t \mid g_{1:t-1})$

Generalization to decentralized system: finite memory controller



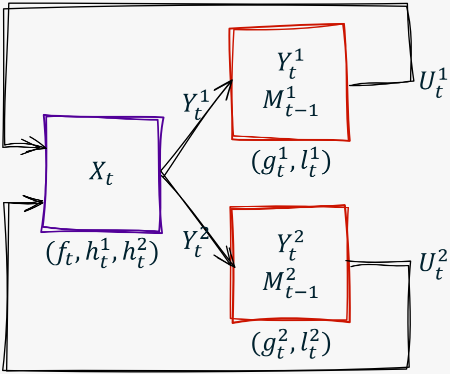
Equivalent to a **centralized** partially observed system

Define: $\pi_t = \mathbb{P}(\text{state} \mid \text{history of data})$

$$= \mathbb{P}(X_t, Y_t \mid g_{1:t-1})$$

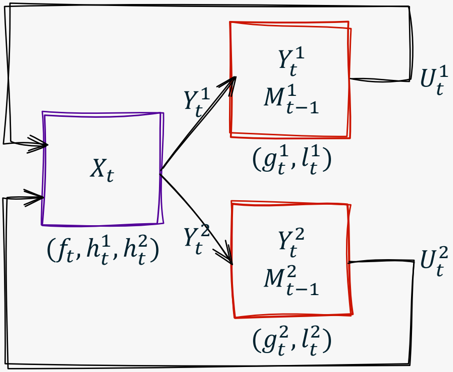
$$V_t(\pi_t) = \min_{g_t} \mathbb{E}[c(X_t, U_t) + V_{t+1}(L^g \pi_t)]$$

Same idea works for arbitrary systems with finite memory controllers



$$\min_{g^{1,2}_{1:T}, l^{1,2}_{1:T}} \mathbb{E}\Big[\sum_{t=1}^{T} c(X_t, U^1_t, U^2_t)\Big]$$

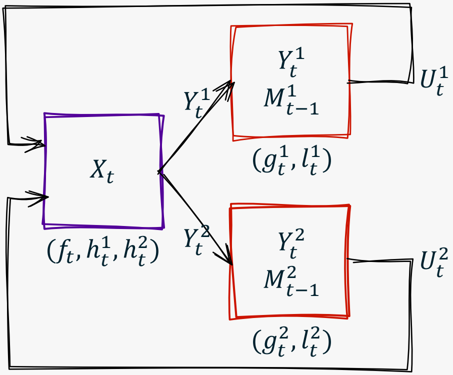
Same idea works for arbitrary systems with finite memory controllers



$$\min_{g^{1,2}_{1:T}, l^{1,2}_{1:T}} \mathbb{E}\Big[\sum_{t=1}^{T} c(X_t, U^1_t, U^2_t)\Big]$$

$$\pi_t = \mathbb{P}(X_t, Y^{1,2}_t, M^{1,2}_{t-1} \mid g^{1,2}_{1:t-1}, l^{1,2}_{1:t-1})$$

Same idea works for arbitrary systems with finite memory controllers

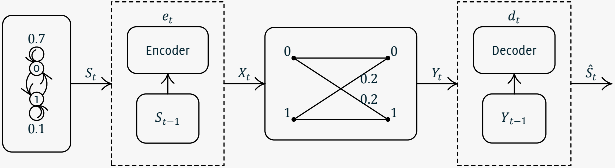


$$\min_{g_{1:T}^{1,2}, l_{1:T}^{1,2}} \mathbb{E}\Big[\sum_{t=1}^{T} c(X_t, U_t^1, U_t^2)\Big]$$

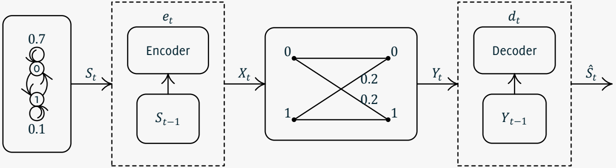
$$\pi_t = \mathbb{P}(X_t, Y_t^{1,2}, M_{t-1}^{1,2} \mid g_{1:t-1}^{1,2}, l_{1:t-1}^{1,2})$$

$$V_t(\pi_t) = \min_{g_t^{1,2},\ l_t^{1,2}} \mathbb{E}[c(X_t, U_t^1, U_t^2) + V_{t+1}(L^{g_t^{1,2},\ l_t^{1,2}} \pi_t)]$$

Example: Real-time communication



Example: Real-time communication



Finite Memory

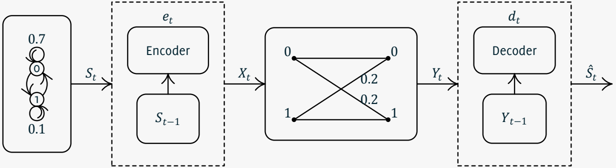
Encoder

$x_t = e_t(s_t, s_{t-1})$

Decoder

$\hat{s}_t = d_t(y_t, y_{t-1})$

Example: Real-time communication



Finite Memory

Encoder

$$x_t = e_t(s_t, s_{t-1})$$

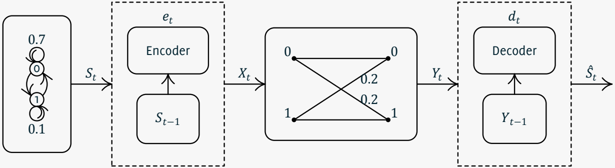
Decoder

$$\hat{s}_t = d_t(y_t, y_{t-1})$$

- Communication Strategy

$$E = (e_1, e_2, \ldots, e_T), \quad D = (d_1, d_2, \ldots, d_T)$$

Example: Real-time communication



Finite Memory

Encoder

$$x_t = e_t(s_t, s_{t-1})$$

Decoder

$$\hat{s}_t = d_t(y_t, y_{t-1})$$

- Communication Strategy

$$E = (e_1, e_2, \dots, e_T), \quad D = (d_1, d_2, \dots, d_T)$$

- Performance

$$\mathcal{J}(E, D) = \lim_{T \to \infty} \mathbb{E}\left\{ \sum_{t=2}^{T} \beta^{t-1}\, \mathbb{P}(\hat{s}_t \neq s_{t-1}) \right\}$$

Gaarder and Slepian's (1982) approach

Brute force search of an optimal policy

- Pick a **time invariant** strategy $E = (e, e, ..., e)$, $D = (d, d, ..., d)$.
- Find the steady-state distribution of the MC $\{S_{t-1}, S_t, Y_{t-1}\}$
- Find the steady-state probability of error

$$\lim_{t\to\infty} \mathbb{E}\left\{\mathbb{P}(\hat{s}_t \neq \hat{s}_{t-1})\right\}$$

- **Repeat** for all time invariant strategies.

Gaarder and Slepian's (1982) approach

Brute force search of an optimal policy

- Pick a **time invariant** strategy $E = (e, e, \ldots, e)$, $D = (d, d, \ldots, d)$.
- Find the steady-state distribution of the MC $\{S_{t-1}, S_t, Y_{t-1}\}$
- Find the steady-state probability of error

$$\lim_{t \to \infty} \mathbb{E}\left\{ \mathbb{P}(\hat{s}_t \neq \hat{s}_{t-1}) \right\}$$

- **Repeat** for all time invariant strategies.

Difficulty with the approach

- Steady-state distribution of a Markov chain is discontinuous in its transition matrix
- For some (E, D), the Markov chain may not have a unique steady-state distribution

Designer approach based solution

Define $\pi_t = \mathbb{P}(s_{t-1}, s_t, y_{t-1} | e_{1:t-1}, d_{1:t-1})$.

Designer approach based solution

Define $\pi_t = \mathbb{P}(s_{t-1}, s_t, y_{t-1} | e_{1:t-1}, d_{1:t-1})$.

Finite horizon: An optimal communication strategy can be determined by the solution of the following nested optimality equations

$$V_T(\pi_T) = \min_{e_T, d_T} \mathbb{E}\Big[\, \mathbb{P}(\hat{s}_T \neq s_{T-1}) \Big| \pi_T, e_T, d_T \Big]$$

$$V_t(\pi_t) = \min_{e_t, d_t} \mathbb{E}\Big[\, \mathbb{P}(\hat{s}_t \neq s_{t-1}) + V_{t+1}(\pi_{t+1}) \Big| \pi_t, e_t, d_t \Big]$$

Designer approach based solution

Define $\pi_t = \mathbb{P}(s_{t-1}, s_t, y_{t-1} | e_{1:t-1}, d_{1:t-1})$.

Finite horizon: An optimal communication strategy can be determined by the solution of the following nested optimality equations

$$V_T(\pi_T) = \min_{e_T, d_T} \mathbb{E}\Big[\, \mathbb{P}(\hat{s}_T \neq s_{T-1}) \Big| \pi_T, e_T, d_T \Big]$$

$$V_t(\pi_t) = \min_{e_t, d_t} \mathbb{E}\Big[\, \mathbb{P}(\hat{s}_t \neq s_{t-1}) + V_{t+1}(\pi_{t+1}) \Big| \pi_t, e_t, d_t \Big]$$

Infinite horizon: . . . fixed point equation

$$V(\pi) = \min_{e,d} \mathbb{E}\Big[\, \mathbb{P}(\hat{s}_t \neq s_{t-1}) + \beta V(\pi_+) \Big| \pi, e, d \Big]$$

Designer approach based solution

Define $\pi_t = \mathbb{P}(s_{t-1}, s_t, y_{t-1} | e_{1:t-1}, d_{1:t-1})$.

Finite horizon: An optimal communication strategy can be determined by the solution of the following nested optimality equations

$$V_T(\pi_T) = \min_{e_T, d_T} \mathbb{E}\Big[\, \mathbb{P}(\hat{s}_T \neq s_{T-1}) \Big| \pi_T, e_T, d_T \Big]$$

$$V_t(\pi_t) = \min_{e_t, d_t} \mathbb{E}\Big[\, \mathbb{P}(\hat{s}_t \neq s_{t-1}) + V_{t+1}(\pi_{t+1}) \Big| \pi_t, e_t, d_t \Big]$$

Infinite horizon: . . . fixed point equation

$$V(\pi) = \min_{e,d} \mathbb{E}\Big[\, \mathbb{P}(\hat{s}_t \neq s_{t-1}) + \beta V(\pi_+) \Big| \pi, e, d \Big]$$

The designer strategy $\gamma_t : \pi_t \to (e_t, d_t)$ is time-invariant. The choice of (e_t, d_t) is **not time invariant**.

Optimal communication scheme

- Example with $\beta = 0.9$

$$(e_t, d_t) = \begin{pmatrix} s_t & , & 0 \\ s_{t-1} \oplus s_t & , & y_{t-1} \oplus y_t \\ s_{t-1} & , & y_{t-1} \oplus y_t \end{pmatrix}_{t \pmod 3}$$

Optimal communication scheme

- Example with $\beta = 0.9$

$$(e_t, d_t) = \begin{pmatrix} s_t & , & 0 \\ s_{t-1} \oplus s_t & , & y_{t-1} \oplus y_t \\ s_{t-1} & , & y_{t-1} \oplus y_t \end{pmatrix}_{t \pmod 3}$$

Source	s_1	s_2	s_3	s_4	s_5	s_6	s_7
Encoder							
Decoder							
Estimate	–	s_1	s_2	s_3	s_4	s_5	s_6

Optimal communication scheme

- Example with $\beta = 0.9$

$$(e_t, d_t) = \begin{pmatrix} s_t & , & 0 \\ s_{t-1} \oplus s_t & , & y_{t-1} \oplus y_t \\ s_{t-1} & , & y_{t-1} \oplus y_t \end{pmatrix}_{t \pmod 3}$$

Source	s_1	s_2	s_3	s_4	s_5	s_6	s_7
Encoder	s_1						
Decoder	0						
Estimate	–	s_1	s_2	s_3	s_4	s_5	s_6

Optimal communication scheme

- Example with $\beta = 0.9$

$$(e_t, d_t) = \begin{pmatrix} s_t & , & 0 \\ s_{t-1} \oplus s_t & , & y_{t-1} \oplus y_t \\ s_{t-1} & , & y_{t-1} \oplus y_t \end{pmatrix}_{t \pmod 3}$$

Source	s_1	s_2	s_3	s_4	s_5	s_6	s_7
Encoder	s_1	$s_1 \oplus s_2$					
Decoder	0	$y_1 \oplus y_2$					
Estimate	–	s_1	s_2	s_3	s_4	s_5	s_6

Optimal communication scheme

- Example with $\beta = 0.9$

$$(e_t, d_t) = \begin{pmatrix} s_t & , & 0 \\ s_{t-1} \oplus s_t & , & y_{t-1} \oplus y_t \\ s_{t-1} & , & y_{t-1} \oplus y_t \end{pmatrix}_{t \pmod 3}$$

Source	s_1	s_2	s_3	s_4	s_5	s_6	s_7
Encoder	s_1	$s_1 \oplus s_2$	s_2				
Decoder	0	$y_1 \oplus y_2$	$y_2 \oplus y_3$				
Estimate	–	s_1	s_2	s_3	s_4	s_5	s_6

Optimal communication scheme

- Example with $\beta = 0.9$

$$(e_t, d_t) = \begin{pmatrix} s_t & , & \color{red}{0} \\ s_{t-1} \oplus s_t & , & y_{t-1} \oplus y_t \\ s_{t-1} & , & y_{t-1} \oplus y_t \end{pmatrix}_{t \pmod 3}$$

Source	s_1	s_2	s_3	s_4	s_5	s_6	s_7
Encoder	s_1	$s_1 \oplus s_2$	s_2	s_4			
Decoder	0	$y_1 \oplus y_2$	$y_2 \oplus y_3$	0			
Estimate	–	s_1	s_2	s_3	s_4	s_5	s_6

Optimal communication scheme

- Example with $\beta = 0.9$

$$(e_t, d_t) = \begin{pmatrix} s_t & , & 0 \\ s_{t-1} \oplus s_t & , & y_{t-1} \oplus y_t \\ s_{t-1} & , & y_{t-1} \oplus y_t \end{pmatrix}_{t \pmod 3}$$

Source	s_1	s_2	s_3	s_4	s_5	s_6	s_7
Encoder	s_1	$s_1 \oplus s_2$	s_2	s_4	$s_4 \oplus s_5$		
Decoder	0	$y_1 \oplus y_2$	$y_2 \oplus y_3$	0	$y_4 \oplus y_5$		
Estimate	–	s_1	s_2	s_3	s_4	s_5	s_6

Optimal communication scheme

- Example with $\beta = 0.9$

$$(e_t, d_t) = \begin{pmatrix} s_t & , & 0 \\ s_{t-1} \oplus s_t & , & y_{t-1} \oplus y_t \\ s_{t-1} & , & y_{t-1} \oplus y_t \end{pmatrix}_{t \pmod 3}$$

Source	s_1	s_2	s_3	s_4	s_5	s_6	s_7
Encoder	s_1	$s_1 \oplus s_2$	s_2	s_4	$s_4 \oplus s_5$	s_5	
Decoder	0	$y_1 \oplus y_2$	$y_2 \oplus y_3$	0	$y_4 \oplus y_5$	$y_5 \oplus y_6$	
Estimate	–	s_1	s_2	s_3	s_4	s_5	s_6

Optimal communication scheme

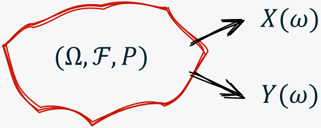
- Example with $\beta = 0.9$

$$(e_t, d_t) = \begin{pmatrix} s_t & , & 0 \\ s_{t-1} \oplus s_t & , & y_{t-1} \oplus y_t \\ s_{t-1} & , & y_{t-1} \oplus y_t \end{pmatrix}_{t \pmod 3}$$

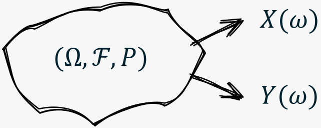
Source	s_1	s_2	s_3	s_4	s_5	s_6	s_7
Encoder	s_1	$s_1 \oplus s_2$	s_2	s_4	$s_4 \oplus s_5$	s_5	s_7
Decoder	0	$y_1 \oplus y_2$	$y_2 \oplus y_3$	0	$y_4 \oplus y_5$	$y_5 \oplus y_6$	0
Estimate	–	s_1	s_2	s_3	s_4	s_5	s_6

Coordinator approach

Common Knowledge (Aumann, 1976)

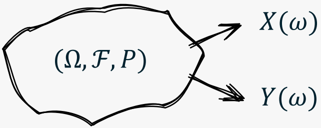


Common Knowledge (Aumann, 1976)



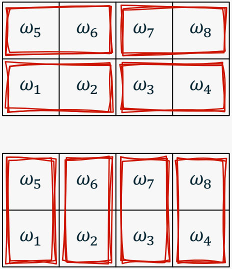
$\sigma(X) \cap \sigma(Y)$

Common Knowledge (Aumann, 1976)

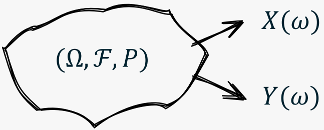


$\sigma(X) \cap \sigma(Y)$

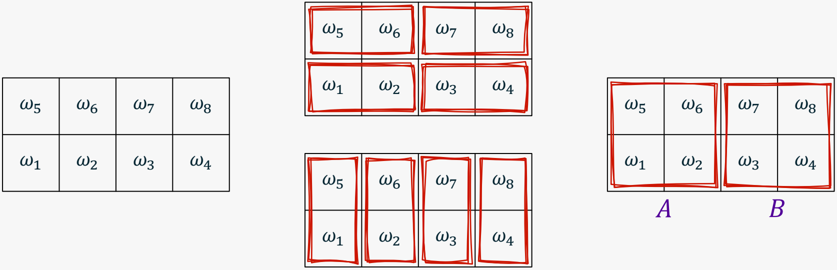
ω_5	ω_6	ω_7	ω_8
ω_1	ω_2	ω_3	ω_4



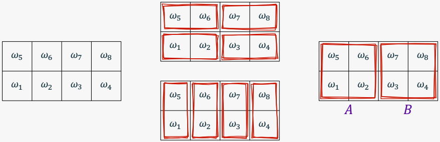
Common Knowledge (Aumann, 1976)



$\sigma(X) \cap \sigma(Y)$



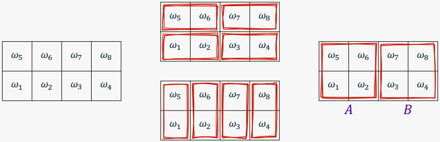
Exploiting common knowledge to simplify decentralized static optimization



$$u = g(x), \quad v = h(y)$$

$$J(g,h) = \mathbb{E}^{g,h}[c(\omega, u, v)]$$

Exploiting common knowledge to simplify decentralized static optimization



$$u = g(x), \quad v = h(y)$$

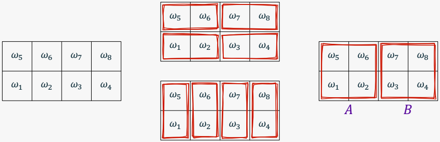
$$J(g,h) = \mathbb{E}^{g,h}[c(\omega,u,v)]$$

Let k denote the common knowledge between x and y. Write:

$$x \equiv (k,p), \quad y \equiv (k,q),$$

$$u = \tilde{g}(k,p). \quad v = \tilde{h}(k,q).$$

Exploiting common knowledge to simplify decentralized static optimization



$$u = g(x), \quad v = h(y)$$

$$J(g,h) = \mathbb{E}^{g,h}[c(\omega,u,v)]$$

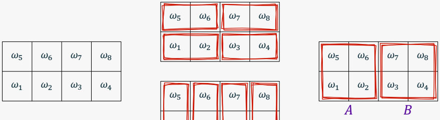
Let k denote the common knowledge between x and y. Write:

$$x \equiv (k,p), \quad y \equiv (k,q),$$

$$u = \tilde{g}(k,p). \quad v = \tilde{h}(k,q).$$

$$\tilde{g} : (k,p) \mapsto u, \quad \tilde{g} : k \mapsto \underbrace{(p \mapsto u)}_{\gamma}$$

Exploiting common knowledge to simplify decentralized static optimization



$$u = g(x), \quad v = h(y)$$

$$J(g,h) = \mathbb{E}^{g,h}[c(\omega,u,v)]$$

Let k denote the common knowledge between x and y. Write:

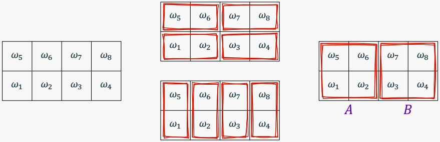
$$x \equiv (k,p), \quad y \equiv (k,q),$$

$$u = \tilde{g}(k,p). \quad v = \tilde{h}(k,q).$$

$$\tilde{g} : (k,p) \mapsto u, \quad \tilde{g} : k \mapsto \underbrace{(p \mapsto u)}_{\gamma}$$

Let $\gamma(\cdot) = \tilde{g}(k,\cdot)$ and $\eta(\cdot) = \tilde{h}(k,\cdot)$

Exploiting common knowledge to simplify decentralized static optimization



$$u = g(x), \quad v = h(y)$$

$$J(g, h) = \mathbb{E}^{g,h}[c(\omega, u, v)]$$

Let k denote the common knowledge between x and y. Write:

$$x \equiv (k, p), \quad y \equiv (k, q),$$

$$u = \tilde{g}(k, p). \quad v = \tilde{h}(k, q).$$

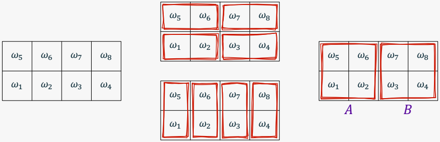
$$\tilde{g} : (k, p) \mapsto u, \quad \tilde{g} : k \mapsto \underbrace{(p \mapsto u)}_{\gamma}$$

Let $\gamma(\cdot) = \tilde{g}(k, \cdot)$ and $\eta(\cdot) = \tilde{h}(k, \cdot)$

A common knowledge based solution

$$\min_{\gamma, \eta} \mathbb{E}^{\gamma,\eta}[c(\omega, u, v) | k]$$

Exploiting common knowledge to simplify decentralized static optimization



$$u = g(x), \quad v = h(y)$$

$$J(g,h) = \mathbb{E}^{g,h}[c(\omega, u, v)]$$

Let k denote the common knowledge between x and y. Write:

$$x \equiv (k, p), \quad y \equiv (k, q),$$

$$u = \tilde{g}(k, p). \quad v = \tilde{h}(k, q).$$

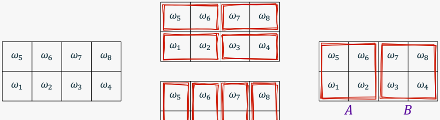
$$\tilde{g} : (k, p) \mapsto u, \quad \tilde{g} : k \mapsto \underbrace{(p \mapsto u)}_{\gamma}$$

Let $\gamma(\cdot) = \tilde{g}(k, \cdot)$ and $\eta(\cdot) = \tilde{h}(k, \cdot)$

A common knowledge based solution (functional opt. over smaller space)

$$\min_{\gamma, \eta} \mathbb{E}^{\gamma, \eta}[c(\omega, u, v) | k]$$

Exploiting common knowledge to simplify decentralized static optimization



$$u = g(x), \quad v = h(y)$$

$$J(g,h) = \mathbb{E}^{g,h}[c(\omega,u,v)]$$

Let k denote the common knowledge between x and y. Write:

$$x \equiv (k,p), \quad y \equiv (k,q),$$

$$u = \tilde{g}(k,p). \quad v = \tilde{h}(k,q).$$

$$\tilde{g} : (k,p) \mapsto u, \quad \tilde{g} : k \mapsto \underbrace{(p \mapsto u)}_{\gamma}$$

Let $\gamma(\cdot) = \tilde{g}(k,\cdot)$ and $\eta(\cdot) = \tilde{h}(k,\cdot)$

A common knowledge based solution (functional opt. over smaller space)

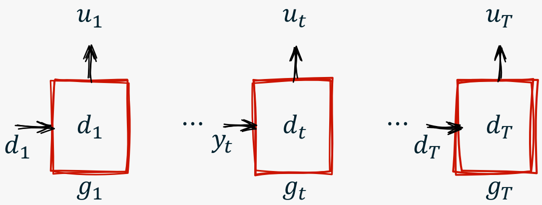
$$\min_{\gamma,\eta} \mathbb{E}^{\gamma,\eta}[c(\omega,u,v)|k]$$

Brute force: $2^4 \times 2^4$ possiblities. CI-based soln: $2 \cdot (2^2 \times 2^2)$ possibilities.

Main idea: Extend CI-based approach to decentralized multi-stage systems.

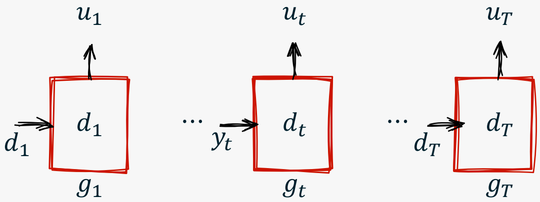
A common information based approach for decentralized multi-stage systems

(Nayyar, 2010; Nayyar, Mahajan, Teneketzis, 2011)



A common information based approach for decentralized multi-stage systems

(Nayyar, 2010; Nayyar, Mahajan, Teneketzis, 2011)



A common information based approach for decentralized multi-stage systems

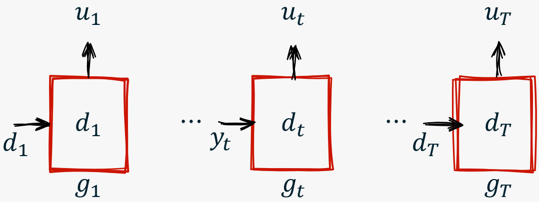
(Nayyar, 2010; Nayyar, Mahajan, Teneketzis, 2011)

- **Common information**:

  $$k_t = \bigcap_{s \ge t} d_s$$

- **Private information**:

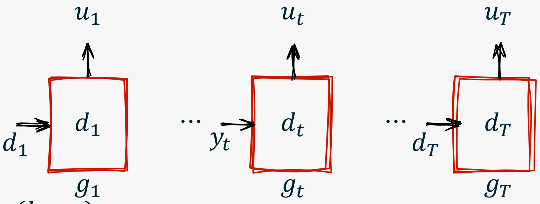
  $$p_t = d_t \setminus k_t$$



A common information based approach for decentralized multi-stage systems

(Nayyar, 2010; Nayyar, Mahajan, Teneketzis, 2011)

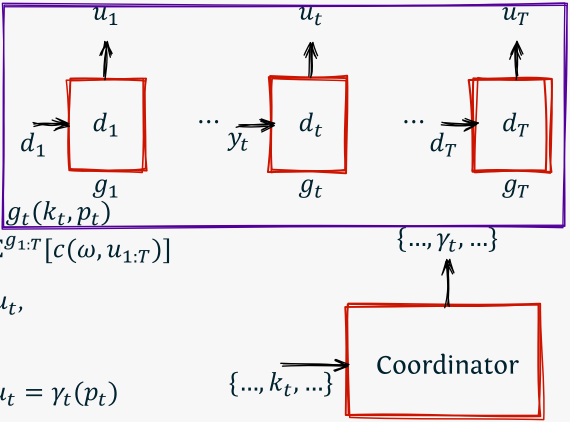
- **Common information**:
  $$k_t = \bigcap_{s \ge t} d_s$$
- **Private information**:
  $$p_t = d_t \setminus k_t$$
- **Objective**: Choose $u_t = g_t(k_t, p_t)$
  to minimize $J(g_{1:T}) = \mathbb{E}^{g_{1:T}}[c(\omega, u_{1:T})]$



A common information based approach for decentralized multi-stage systems

(Nayyar, 2010; Nayyar, Mahajan, Teneketzis, 2011)

- **Common information**:
  $$k_t = \bigcap_{s \ge t} d_s$$
- **Private information**:
  $p_t = d_t \setminus k_t$
- **Objective**: Choose $u_t = g_t(k_t, p_t)$ to minimize $J(g_{1:T}) = \mathbb{E}^{g_{1:T}}[c(\omega, u_{1:T})]$
- **Prescription**: $\gamma_t : p_t \mapsto u_t$, chosen according to
  $$\gamma_t = \psi_t(k_t, \gamma_{1:t-1}), \qquad u_t = \gamma_t(p_t)$$

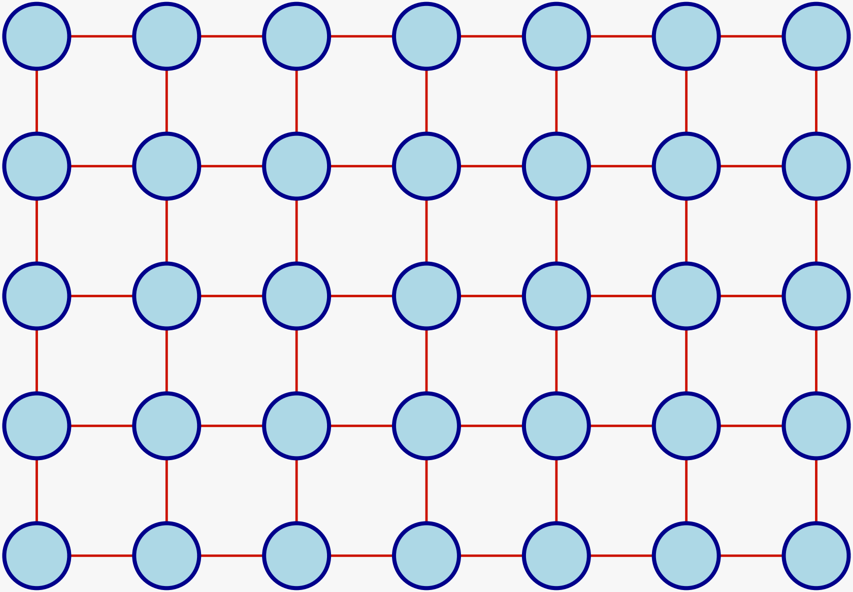


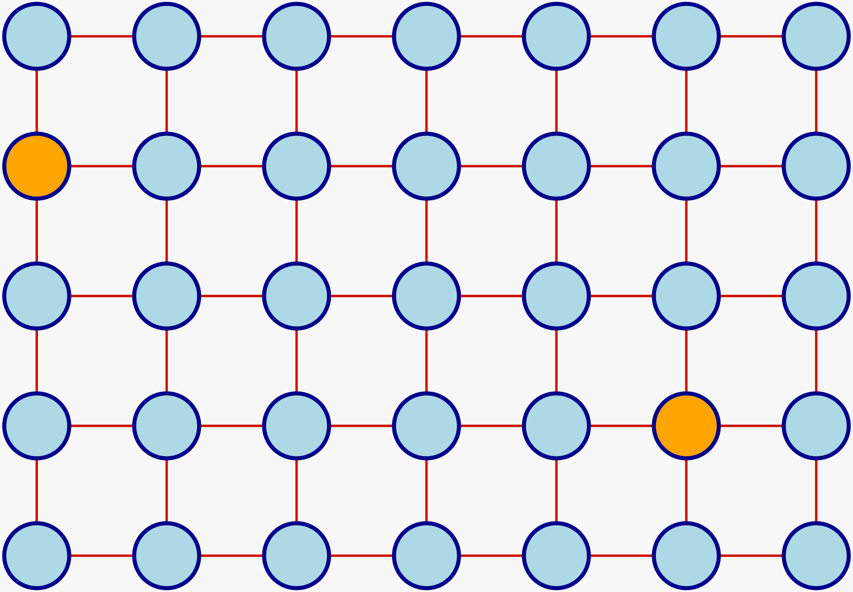
A common information based approach for decentralized multi-stage systems

Solution approach

1. Construct a **coordinated system** (that has classical info-struct.)
2. Show that coordinated system $\equiv$ original system.
3. Find a solution to coordinated system using centralized stoc. control.
4. Translate the result back to original system

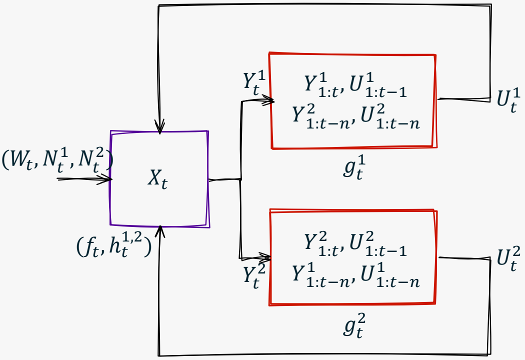
An example: delayed sharing information structure





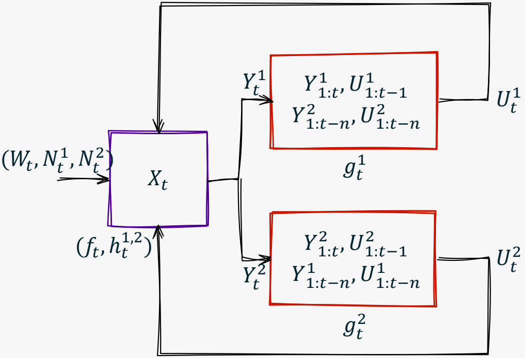
Delayed sharing information structure

(Nayyar, Mahajan, Teneketzis, 2011)



Delayed sharing information structure

(Nayyar, Mahajan, Teneketzis, 2011)



$$J(g^{1,2}_{1:T}) = \mathbb{E}^{g^{1,2}_{1:T}}[c(X_t, U^{1,2}_t)]$$

Literature Overview

How to compress data into a sufficient statistic?

- (Witsenhausen, 1971):
  - Proposed delayed-sharing information structure.
  - **Asserted** a structure of optimal control law (without proof).

Literature Overview

How to compress data into a sufficient statistic?

- (Witsenhausen, 1971):
  - Proposed delayed-sharing information structure.
  - **Asserted** a structure of optimal control law (without proof).
- (Varaiya and Walrand, 1978):
  - Proved Witsenhausen's assertion for $n = 1$.
  - Counter-example to **disproved** the assertion for delay $n > 2$.

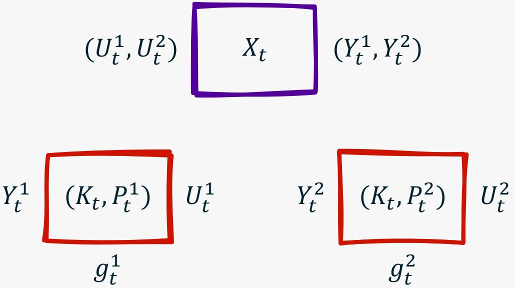
Literature Overview

How to compress data into a sufficient statistic?

- (Witsenhausen, 1971):
  - Proposed delayed-sharing information structure.
  - **Asserted** a structure of optimal control law (without proof).
- (Varaiya and Walrand, 1978):
  - Proved Witsenhausen's assertion for $n = 1$.
  - Counter-example to **disproved** the assertion for delay $n > 2$.

A good model for many applications The result of one-step delayed sharing used in various applications in queuing theory, communication networks, stochastic games, and economics. (Kuri Kumar, 1995; Altman *et al*, 2009; Grizzle *et al* 1982; Papavassilopoulos, 1982; Chang and Cruz, 1983; Li and Whu, 1991).

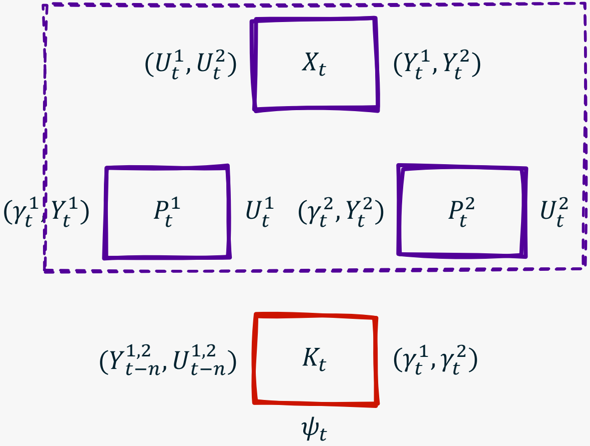
The coordinated system: state for I/O mapping



Common information $K_t = (Y^{1,2}_{1:t-n}, U^{1,2}_{1:t-n}).$

Private information $P_t^i = (Y^i_{t-n+1:t}, U^i_{t-n+1:t-1})$

The coordinated system: state for I/O mapping



Information state for coordinated system

The coordinated system is a centralized partially observed system.

$$\text{Info state} = \mathbb{P}(\text{state for I/O mapping} \mid \text{data at controller})$$

Information state for coordinated system

The coordinated system is a centralized partially observed system.

$$\text{Info state} = \mathbb{P}(\text{state for I/O mapping} \mid \text{data at controller})$$

$$\pi_t = \mathbb{P}(X_t, P_t^1, P_t^2 \mid K_t, \gamma_t^1, \gamma_t^2)$$

Structural Result There is no loss of optimality in restricting prescriptions of the form

$$\gamma_t = \psi_t(\pi_t) \quad \text{and hence,} \quad U_t^i = g_t^i(\pi_t, P_t^i)$$

Information state for coordinated system

The coordinated system is a centralized partially observed system.

$$\text{Info state} = \mathbb{P}(\text{state for I/O mapping} \mid \text{data at controller})$$

$$\pi_t = \mathbb{P}(X_t, P_t^1, P_t^2 \mid K_t, \gamma_t^1, \gamma_t^2)$$

Structural Result There is no loss of optimality in restricting prescriptions of the form

$$\gamma_t = \psi_t(\pi_t) \quad \text{and hence,} \quad U_t^i = g_t^i(\pi_t, P_t^i)$$

Dynamic Programming decomposition An optimal coordination strategy is given by the solution to the following dynamic program

$$V_t(\pi_t) = \min_{\gamma_t^1, \gamma_t^2} \mathbb{E}[c(X_t, \gamma_t^1(P_t^1), \gamma_t^2(P_t^2)) + V_{t+1}(\pi_{t+1} \mid \pi_t, \gamma_t^1, \gamma_t^2)]$$

Information state for coordinated system

The coordinated system is a centralized partially observed system.

$$\text{Info state} = \mathbb{P}(\text{state for I/O mapping} \mid \text{data at controller})$$

$$\pi_t = \mathbb{P}(X_t, P^1_t, P^2_t \mid K_t, \gamma^1_t, \gamma^2_t)$$

Structural Result There is no loss of optimality in restricting prescriptions of the form

$$\gamma_t = \psi_t(\pi_t) \quad \text{and hence,} \quad U^i_t = g^i_t(\pi_t, P^i_t)$$

Dynamic Programming decomposition An optimal coordination strategy is given by the solution to the following dynamic program

$$V_t(\pi_t) = \min_{\gamma^1_t, \gamma^2_t} \mathbb{E}[c(X_t, \gamma^1_t(P^1_t), \gamma^2_t(P^2_t)) + V_{t+1}(\pi_{t+1} \mid \pi_t, \gamma^1_t, \gamma^2_t)]$$

Setting $g^i_t(\pi_t, P^i_t) = \psi^i_t(\pi_t)(P^i_t)$ gives optimal control strategy.

An easy solution to long standing open problem

Connections

(Nayyar, Mahajan, Teneketzis, 2011)

Many existing results on decentralized control are special cases

- Delayed state sharing (Aicardi *et al*, 1987)
- Periodic sharing information structures (Ooi *et al*, 1997)
- Control sharing (Bismut, 1972; Sandell and Athans, 1974; Mahajan 2011)
- Finite state memory controllers (Sandell, 1974, Mahajan, 2008)

Connections

(Nayyar, Mahajan, Teneketzis, 2011)

Many existing results on decentralized control are special cases

- Delayed state sharing (Aicardi *et al*, 1987)
- Periodic sharing information structures (Ooi *et al*, 1997)
- Control sharing (Bismut, 1972; Sandell and Athans, 1974; Mahajan 2011)
- Finite state memory controllers (Sandell, 1974, Mahajan, 2008)

Generalization to other models

- **Infinite horizon** (discounted and average cost) models using standard results for POMDPs
- **Computation algorithms** based on algorithms for POMDPs
- Extend results to systems with **unknown models** based on Q-learning and adaptive control algorithms

Summary of the main idea

- Find **common information** at the controllers
- Look from the point of view of a **coordinator** that observes common information and chooses prescriptions to the controllers
- Find **information state** for the coordinated system and use it to set up a dynamic program
- When common information is nil, the approach reduces to **designer's approach**

Future Directions

Identify other tractable information structures

The common information approach (almost) all known results for Markov chain setup. Are there other structures that are tractable?

Computational algorithms

Develop computation algorithms that are tuned to the type of DP equations that arise in decentralized control.

Connections with stochastic optimization

Can techniques from stochastic optimization used to solve decentralized stochastic control problems?

Connections with sequential games

Does the common information approach help in identifying sequential equilibrium in sequential games?

Thank you

References

1. H. Witsenhausen, A Standard Form for Sequential Stochastic Control, *Mathematical System Theory,* vol 7, no 1, pp. 1-11, 1973.
2. A. Mahajan, Sequential decomposition of sequential teams: applications to real-time communication and networked control systems,*PhD thesis*, Univ of Michigan, Ann Arbor, MI. 2008.
3. A. Nayyar, Sequential decision making in decentralized systems, *PhD thesis*, Univ of Michigan, Ann Arbor, MI. 2011.
4. A. Nayyar, Sequential decision making in decentralized systems, *PhD thesis*, Univ of Michigan, 2011.
5. H. Witsenhausen, Separation of estimation and control for discrete time systems, *Proceedings of the IEEE*, vol 59, no 11, pp. 1557-1566, Nov 1971.
6. A. Nayyar, A. Mahajan, D. Teneketzis, Optimal control strategies for delayed sharing information structures, *IEEE Trans. on Automatic Control,* vol. 56, no. 7, pp. 1606-1620, July 2011.

References

7. A. Nayyar, A. Mahajan, D. Teneketzis, Dynamic programming for decentralized stochastic control with partial information sharing: a common information approach, submitted to *IEEE Trans. on Automatic Control*, Dec 2011.
8. A. Mahajan, Optimal decentralized control of coupled subsystems with control sharing,submitted to *IEEE Trans. on Automatic Control*, Dec 2011.