Team Decision Theory: Characterization of Information Structures, Basic Concepts and Solution Methods

Aditya Mahajan[1] Serdar Yüksel[2]

[1]McGill University

[2]Queen's University

CDC tutorial on
Information Structures in Optimal Control
10 Dec, 2012

Outline of the talk

Dynamic programming for centralized control

Dynamic programming for decentralized control

Salient features of common information approach

- Combining with person-by-person approach
- Contructing common information

Using decentralized dynamic programming to specific applications

Conclusion

Outline of the talk

Dynamic programming for centralized control

Dynamic programming for decentralized control

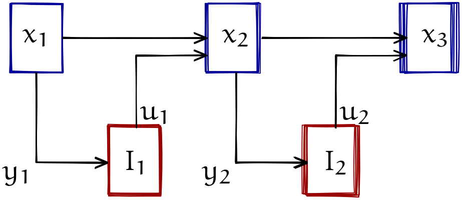
Salient features of common information approach

- Combining with person-by-person approach
- Contructing common information

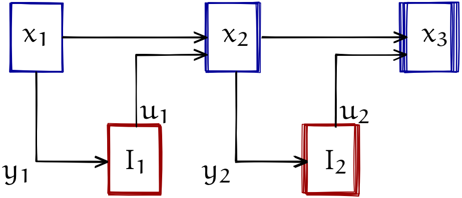
Using decentralized dynamic programming to specific applications

Conclusion

In centralized stochastic control, one DM with perfect recall takes multiple decision over time

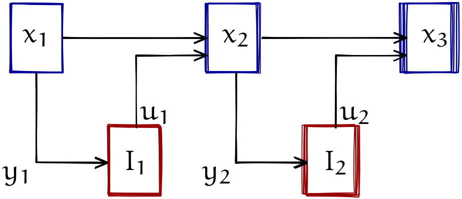


In centralized stochastic control, one DM with perfect recall takes multiple decision over time



- $I_t = \{y_{[1:t]}, u_{[1:t-1]}\}$ (perfect recall)
- $I_t \subseteq I_{t+1}$ (classical info-structure)

In centralized stochastic control, one DM with perfect recall takes multiple decision over time



- $I_t = \{y_{[1:t]}, u_{[1:t-1]}\}$ (perfect recall)
- $I_t \subseteq I_{t+1}$ (classical info-structure)

Conceptual difficulties

- $\gamma_t : I_t \mapsto u_t$
- $\min J(\gamma_1, \ldots, \gamma_T)$

Identifying an information state is a key step in centralized dynamic programming

Properties of information state

- π_t is a function of the information I_t
- π_{t+1} is a function of π_t and new information (u_t, y_{t+1})
- π_t is a sufficient statistic for predicting future observations

$$\mathcal{P}(y_{t+1} \mid I_t) = \mathcal{P}(y_{t+1} \mid \pi_t)$$

- π_t is sufficient for performance evaluation

$$\mathcal{E}[c(x_t, u_t) \mid I_t, u_t] = \mathcal{E}[c(x_t, u_t) \mid \pi_t, u_t]$$

Identifying an information state is a key step in centralized dynamic programming

Properties of information state

- π_t is a function of the information I_t
- π_{t+1} is a function of π_t and new information (u_t, y_{t+1})
- π_t is a sufficient statistic for predicting future observations

$$\mathcal{P}(y_{t+1} \mid I_t) = \mathcal{P}(y_{t+1} \mid \pi_t)$$

- π_t is sufficient for performance evaluation

$$\mathcal{E}[c(x_t, u_t) \mid I_t, u_t] = \mathcal{E}[c(x_t, u_t) \mid \pi_t, u_t]$$

Examples of information states

- In partially observable Markov decision problems (POMDPs), the belief state $\mathcal{P}(x_t \mid I_t)$ is an information state.
- In linear quadratic and Gaussain (LQG) problems, the state estimate $\hat{x}_{t|t}$ is an information state.

DP uses the information state to sequentially decompose the optimization problem

Structural result

Restricting attention to control laws of the form $u_t = \gamma_t(\pi_t)$ does not entail any loss of optimality.

- In some cases, also derive qualitative properties of the optimal policy (monotonicity, convexity, threshold, etc.)

DP uses the information state to sequentially decompose the optimization problem

Structural result

Restricting attention to control laws of the form $u_t = \gamma_t(\pi_t)$ does not entail any loss of optimality.

- In some cases, also derive qualitative properties of the optimal policy (monotonicity, convexity, threshold, etc.)

Dynamic programming decomposition

Recursively define

$$V_t(\pi_t) = \inf_{u_t \in \mathbb{U}_t} \mathcal{E}\left[c(x_t, u_t) + V_{t+1}(\pi_{t+1}) \mid \pi_t, u_t \right]$$

If the infimum is achieved, then the arg min at time t gives the optimal control action at information state π_t.

- Functional vs parameteric optimization

Outline of the talk

Dynamic programming for centralized control

Dynamic programming for decentralized control

Salient features of common information approach
- Combining with person-by-person approach
- Contructing common information

Using decentralized dynamic programming to specific applications

Conclusion

There is no obvious way to generalize DP to decentralized control

DP relies on classical information structure:

$$V_t(I_t) = \inf_{u_t \in \mathbb{U}_t} \mathcal{E}\left[c(\cdot) + \inf_{u_{t+1} \in \mathbb{U}_{t+1}} \mathcal{E}[c(\cdot) + \cdots \mid I_{t+1}] \;\middle|\; I_t\right]$$

There is no obvious way to generalize DP to decentralized control

DP relies on classical information structure:

$$V_t(I_t) = \inf_{u_t \in \mathbb{U}_t} \mathcal{E}\left[c(\cdot) + \inf_{u_{t+1} \in \mathbb{U}_{t+1}} \mathcal{E}[c(\cdot) + \cdots \mid I_{t+1}] \;\middle|\; I_t\right]$$

For consistency, we need $I_t \subseteq I_{t+1}$

There is no obvious way to generalize DP to decentralized control

DP relies on classical information structure:

$$V_t(I_t) = \inf_{u_t \in \mathbb{U}_t} \mathcal{E}\left[c(\cdot) + \inf_{u_{t+1} \in \mathbb{U}_{t+1}} \mathcal{E}[c(\cdot) + \cdots \mid I_{t+1}] \;\middle|\; I_t\right]$$

For consistency, we need $I_t \subseteq I_{t+1}$ which is not the case for decentralized control

There is no obvious way to generalize DP to decentralized control

DP relies on classical information structure:

$$V_t(I_t) = \inf_{u_t \in \mathbb{U}_t} \mathcal{E}\left[c(\cdot) + \inf_{u_{t+1} \in \mathbb{U}_{t+1}} \mathcal{E}[c(\cdot) + \cdots \mid I_{t+1}] \;\middle|\; I_t\right]$$

For consistency, we need $I_t \subseteq I_{t+1}$ which is not the case for decentralized control

Techniques from centralized control not directly applicable to decentralized control.

All DP approaches exploit the common or shared information between DMs

Examples of information structures with sharing

- Periodic sharing information structure
  (Chong and Athans, 1976; Ooi et at, 1997)
- Delayed sharing information structure
  (Yoshikawa 1975, Aicardi et al 1987, Nayyar Mahajan Teneketzis 2011)
- Belief sharing (Yüksel 2009)
- Control sharing information structure
  (Bismut 1972, Sandell Athans 1974, Mahajan 2011)

All DP approaches exploit the common or shared information between DMs

Examples of information structures with sharing

- Periodic sharing information structure
  (Chong and Athans, 1976; Ooi et at, 1997)
- Delayed sharing information structure
  (Yoshikawa 1975, Aicardi et al 1987, Nayyar Mahajan Teneketzis 2011)
- Belief sharing (Yüksel 2009)
- Control sharing information structure
  (Bismut 1972, Sandell Athans 1974, Mahajan 2011)
- Specific information structures
  (Walrand Varaiya 1982, Veeravalli Başar Poor 1993)

All DP approaches exploit the common or shared information between DMs

Examples of information structures with sharing

- Periodic sharing information structure
  (Chong and Athans, 1976; Ooi et at, 1997)
- Delayed sharing information structure
  (Yoshikawa 1975, Aicardi et al 1987, Nayyar Mahajan Teneketzis 2011)
- Belief sharing (Yüksel 2009)
- Control sharing information structure
  (Bismut 1972, Sandell Athans 1974, Mahajan 2011)
- Specific information structures
  (Walrand Varaiya 1982, Veeravalli Başar Poor 1993)
- Partial history sharing
  (Nayyar 2011, Nayyar Mahajan Teneketzis (accepted 2013))

A unified framework for dynamic programming*

- $\gamma^i_t : I^i_t \mapsto u^i_t$

- $\min J(\gamma_1, \dots, \gamma_T)$

*Nayyar, Mahajan, Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," TAC (accepted) 2013

A unified framework for dynamic programming*

- $\gamma_t^i : I_t^i \mapsto u_t^i$

- $\min J(\gamma_1, \dots, \gamma_T)$

Identify a coordinator based on common information

- has access to the common information $C_t = \bigcap_{i=1}^{n} I_t^i$
- chooses prescriptions $\varphi = (\varphi_t^1, \dots, \varphi_t^n)$ where $\varphi_t^i : I_t^i \setminus C_t \mapsto u_t^i$.

*Nayyar, Mahajan, Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," TAC (accepted) 2013

A unified framework for dynamic programming*

- $\gamma_t^i : I_t^i \mapsto u_t^i$

Identify sufficient statistic/information state

- $\gamma_t^i : S_t^i \mapsto u_t^i$
- Use a person-by-person approach

- $\min J(\gamma_1, \dots, \gamma_T)$

Identify a coordinator based on common information

- has access to the common information $C_t = \bigcap_{i=1}^{n} I_t^i$
- chooses prescriptions $\varphi = (\varphi_t^1, \dots, \varphi_t^n)$ where $\varphi_t^i : I_t^i \setminus C_t \mapsto u_t^i$.

*Nayyar, Mahajan, Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," TAC (accepted) 2013

The common information approach to reformulate the system from the perspective of a coordinator

1. Construct a coordinated system in which the coordinator:

   - has access to the common information $C_t = \bigcap_{i=1}^{n} I_t^i$
   - chooses prescriptions $\varphi = (\varphi_t^1, \dots, \varphi_t^n)$ where $\varphi_t^i : I_t^i \setminus C_t \mapsto u_t^i$.

The common information approach to reformulate the system from the perspective of a coordinator

1. Construct a coordinated system in which the coordinator:
   - has access to the common information $C_t = \bigcap_{i=1}^{n} I_t^i$
   - chooses prescriptions $\varphi = (\varphi_t^1, \ldots, \varphi_t^n)$ where $\varphi_t^i : I_t^i \setminus C_t \mapsto u_t^i$.
2. Formulate the coordinated system as a centralized stochastic control system (POMDP)

The common information approach to reformulate the system from the perspective of a coordinator

1. Construct a coordinated system in which the coordinator:
   - has access to the common information $C_t = \bigcap_{i=1}^{n} I_t^i$
   - chooses prescriptions $\varphi = (\varphi_t^1, \dots, \varphi_t^n)$ where $\varphi_t^i : I_t^i \setminus C_t \mapsto u_t^i$.
2. Formulate the coordinated system as a centralized stochastic control system (POMDP)
3. Solve the resultant centralized system (POMDP)

The common information approach to reformulate the system from the perspective of a coordinator

1. Construct a coordinated system in which the coordinator:
   - has access to the common information $C_t = \bigcap_{i=1}^{n} I_t^i$
   - chooses prescriptions $\varphi = (\varphi_t^1, \ldots, \varphi_t^n)$ where $\varphi_t^i : I_t^i \setminus C_t \mapsto u_t^i$.
2. Formulate the coordinated system as a centralized stochastic control system (POMDP)
3. Solve the resultant centralized system (POMDP)
4. Show the equivalence between the two models

The common information approach to reformulate the system from the perspective of a coordinator

1. Construct a coordinated system in which the coordinator:
   - has access to the common information $C_t = \bigcap_{i=1}^{n} I_t^i$
   - chooses prescriptions $\varphi = (\varphi_t^1, \ldots, \varphi_t^n)$ where $\varphi_t^i : I_t^i \setminus C_t \mapsto u_t^i$.
2. Formulate the coordinated system as a centralized stochastic control system (POMDP)
3. Solve the resultant centralized system (POMDP)
4. Show the equivalence between the two models
5. Translate the results back to the original decentralized system

Example 1: Common information approach to a static team*

$\omega \mapsto (C, M^1, M^2),$

*Nayyar, Mahajan, Teneketzis, "The common-information approach to decentralized stochastic control," LCCC information control in networks workshop, 2012.

Example 1:
Common information approach to a static team*

$\omega \mapsto (C, M^1, M^2)$,

$$I^1 = (C, M^1), \qquad I^2 = (C, M^2)$$

*Nayyar, Mahajan, Teneketzis, "The common-information approach to decentralized stochastic control," LCCC information control in networks workshop, 2012.

Example 1:
Common information approach to a static team*

$\omega \mapsto (C, M^1, M^2)$,

$$I^1 = (C, M^1), \qquad I^2 = (C, M^2)$$
$$u^1 = \gamma^1(C, M^1), \qquad u^2 = \gamma^2(C, M^2)$$

*Nayyar, Mahajan, Teneketzis, "The common-information approach to decentralized stochastic control," LCCC information control in networks workshop, 2012.

Example 1: Common information approach to a static team*

$\omega \mapsto (C, M^1, M^2)$,

$$\left.\begin{aligned} I^1 &= (C, M^1), & I^2 &= (C, M^2) \\ u^1 &= \gamma^1(C, M^1), & u^2 &= \gamma^2(C, M^2) \end{aligned}\right\} \qquad \ell(\omega, u^1, u^2)$$

*Nayyar, Mahajan, Teneketzis, "The common-information approach to decentralized stochastic control," LCCC information control in networks workshop, 2012.

Example 1: Common information approach to a static team*

$\omega \mapsto (C, M^1, M^2)$,

$$\left.\begin{aligned} I^1 &= (C, M^1), & I^2 &= (C, M^2) \\ u^1 &= \gamma^1(C, M^1), & u^2 &= \gamma^2(C, M^2) \end{aligned}\right\} \qquad \ell(\omega, u^1, u^2)$$

- Brute force search: $\prod_{i=1}^{2} |\mathbb{U}^i|^{|\mathbb{C}||\mathbb{M}^i|}$.

---

*Nayyar, Mahajan, Teneketzis, "The common-information approach to decentralized stochastic control," LCCC information control in networks workshop, 2012.

Example 1: Common information approach to a static team*

$\omega \mapsto (C, M^1, M^2)$,

$$\left.\begin{array}{ll} I^1 = (C, M^1), & I^2 = (C, M^2) \\ u^1 = \gamma^1(C, M^1), & u^2 = \gamma^2(C, M^2) \end{array}\right\} \qquad \ell(\omega, u^1, u^2)$$

- Brute force search: $\displaystyle\prod_{i=1}^{2} |\mathbb{U}^i|^{|\mathbb{C}||\mathbb{M}^i|}$.

All binary alphabets $\implies$ 256 possibilities

*Nayyar, Mahajan, Teneketzis, "The common-information approach to decentralized stochastic control," LCCC information control in networks workshop, 2012.

Example 1:
Common information approach to a static team*

$\omega \mapsto (C, M^1, M^2)$,

$$\left.\begin{aligned} I^1 &= (C, M^1), & I^2 &= (C, M^2) \\ u^1 &= \gamma^1(C, M^1), & u^2 &= \gamma^2(C, M^2) \end{aligned}\right\} \qquad \ell(\omega, u^1, u^2)$$

Common information approach

- Contruct a coordinator: observes $C = I^1 \cap I^2$

*Nayyar, Mahajan, Teneketzis, "The common-information approach to decentralized stochastic control," LCCC information control in networks workshop, 2012.

Example 1: Common information approach to a static team*

$\omega \mapsto (C, M^1, M^2)$,

$$\left.\begin{array}{ll} I^1 = (C, M^1), & I^2 = (C, M^2) \\ u^1 = \gamma^1(C, M^1), & u^2 = \gamma^2(C, M^2) \end{array}\right\} \qquad \ell(\omega, u^1, u^2)$$

Common information approach

- Contruct a coordinator: observes $C = I^1 \cap I^2$
- Coordinator chooses $(\varphi^1, \varphi^2) = \psi(C)$ where

$$u^1 = \varphi^1(M^1), \quad u^2 = \varphi^2(M^2)$$

*Nayyar, Mahajan, Teneketzis, "The common-information approach to decentralized stochastic control," LCCC information control in networks workshop, 2012.

Example 1: Common information approach to a static team*

$\omega \mapsto (C, M^1, M^2)$,

$$\left.\begin{aligned} I^1 &= (C, M^1), & I^2 &= (C, M^2) \\ u^1 &= \gamma^1(C, M^1), & u^2 &= \gamma^2(C, M^2) \end{aligned}\right\} \qquad \ell(\omega, u^1, u^2)$$

Common information approach

- Contruct a coordinator: observes $C = I^1 \cap I^2$
- Coordinator chooses $(\varphi^1, \varphi^2) = \psi(C)$ where

$$u^1 = \varphi^1(M^1), \quad u^2 = \varphi^2(M^2)$$

- Same cost $\ell(\omega, u^1, u^2)$

*Nayyar, Mahajan, Teneketzis, "The common-information approach to decentralized stochastic control," LCCC information control in networks workshop, 2012.

Example 1:
Common information approach to a static team*

$\omega \mapsto (C, M^1, M^2)$,

$$\left.\begin{aligned} I^1 &= (C, M^1), & I^2 &= (C, M^2) \\ u^1 &= \gamma^1(C, M^1), & u^2 &= \gamma^2(C, M^2) \end{aligned}\right\} \qquad \ell(\omega, u^1, u^2)$$

- Brute force search: $\prod_{i=1}^{2} |\mathbb{U}^i|^{|\mathbb{C}||\mathbb{M}^i|}$.

All binary alphabets $\implies$ 256 possibilities

*Nayyar, Mahajan, Teneketzis, "The common-information approach to decentralized stochastic control," LCCC information control in networks workshop, 2012.

Example 1: Common information approach to a static team*

$\omega \mapsto (C, M^1, M^2)$,

$$\left.\begin{array}{ll} I^1 = (C, M^1), & I^2 = (C, M^2) \\ u^1 = \gamma^1(C, M^1), & u^2 = \gamma^2(C, M^2) \end{array}\right\} \qquad \ell(\omega, u^1, u^2)$$

- Brute force search: $\prod_{i=1}^{2} |\mathbb{U}^i|^{|\mathbb{C}||\mathbb{M}^i|}$.

  All binary alphabets $\implies$ 256 possibilities
- Common info approach: $|\mathbb{C}| \prod_{i=1}^{2} |\mathbb{U}^i|^{|\mathbb{M}^i|}$

*Nayyar, Mahajan, Teneketzis, "The common-information approach to decentralized stochastic control," LCCC information control in networks workshop, 2012.

Example 1:
Common information approach to a static team*

$\omega \mapsto (C, M^1, M^2)$,

$$\left.\begin{array}{ll} I^1 = (C, M^1), & I^2 = (C, M^2) \\ u^1 = \gamma^1(C, M^1), & u^2 = \gamma^2(C, M^2) \end{array}\right\} \qquad \ell(\omega, u^1, u^2)$$

- Brute force search: $\prod_{i=1}^{2} |\mathbb{U}^i|^{|\mathbb{C}||\mathbb{M}^i|}$.
  All binary alphabets $\implies$ 256 possibilities
- Common info approach: $|\mathbb{C}| \prod_{i=1}^{2} |\mathbb{U}^i|^{|\mathbb{M}^i|}$
  All binary alphabets $\implies$ 32 possibilities

*Nayyar, Mahajan, Teneketzis, "The common-information approach to decentralized stochastic control," LCCC information control in networks workshop, 2012.

Example 2: Delayed sharing information structure*

$\omega \mapsto (x_1, w^0_{[1:T]}, w^1_{[1:T]}, w^2_{[1:T]}),$

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

Example 2: Delayed sharing information structure*

$\omega \mapsto (x_1, w^0_{[1:T]}, w^1_{[1:T]}, w^2_{[1:T]}),$

$$x_{t+1} = f_t(x_t, u^1_t, u^2_t, w^0_t), \qquad y^i_t = h^i_t(x_t, w^i_t)$$

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

Example 2: Delayed sharing information structure[*]

$$\omega \mapsto (x_1, w^0_{[1:T]}, w^1_{[1:T]}, w^2_{[1:T]}),$$

$$x_{t+1} = f_t(x_t, u^1_t, u^2_t, w^0_t), \qquad y^i_t = h^i_t(x_t, w^i_t)$$

$$I^1_t = (y^1_{[1:t]}, u^1_{[1:t-1]}, y^2_{[1:t-d]}, u^2_{[1:t-d]}),$$

$$I^2_t = (y^2_{[1:t]}, u^2_{[1:t-1]}, y^1_{[1:t-d]}, u^1_{[1:t-d]}),$$

[*]Nayyar, Mahajan, Teneketzis, "Optimal control strategies in delayed sharing information structures," IEEE TAC 2011

Example 2: Delayed sharing information structure*

$$\omega \mapsto (x_1, w^0_{[1:T]}, w^1_{[1:T]}, w^2_{[1:T]}),$$

$$\begin{aligned} x_{t+1} &= f_t(x_t, u^1_t, u^2_t, w^0_t), & y^i_t &= h^i_t(x_t, w^i_t) \\ I^1_t &= (y^1_{[1:t]}, u^1_{[1:t-1]}, y^2_{[1:t-d]}, u^2_{[1:t-d]}), & u^1_t &= \gamma^1_t(I^1_t) \\ I^2_t &= (y^2_{[1:t]}, u^2_{[1:t-1]}, y^1_{[1:t-d]}, u^1_{[1:t-d]}), & u^2_t &= \gamma^2_t(I^2_t) \end{aligned}$$

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

Example 2: Delayed sharing information structure[*]

$$\omega \mapsto (x_1, w^0_{[1:T]}, w^1_{[1:T]}, w^2_{[1:T]}), \qquad \ell(x_t, u^1_t, u^2_t)$$

$$\begin{aligned}
x_{t+1} &= f_t(x_t, u^1_t, u^2_t, w^0_t), & y^i_t &= h^i_t(x_t, w^i_t) \\
I^1_t &= (y^1_{[1:t]}, u^1_{[1:t-1]}, y^2_{[1:t-d]}, u^2_{[1:t-d]}), & u^1_t &= \gamma^1_t(I^1_t) \\
I^2_t &= (y^2_{[1:t]}, u^2_{[1:t-1]}, y^1_{[1:t-d]}, u^1_{[1:t-d]}), & u^2_t &= \gamma^2_t(I^2_t)
\end{aligned}$$

[*]Nayyar, Mahajan, Teneketzis, "Optimal control strategies in delayed sharing information structures," IEEE TAC 2011

Example 2: Delayed sharing information structure*

$$\omega \mapsto (x_1, w^0_{[1:T]}, w^1_{[1:T]}, w^2_{[1:T]}), \qquad \ell(x_t, u^1_t, u^2_t)$$

$$\begin{aligned} x_{t+1} &= f_t(x_t, u^1_t, u^2_t, w^0_t), & y^i_t &= h^i_t(x_t, w^i_t) \\ I^1_t &= (y^1_{[1:t]}, u^1_{[1:t-1]}, y^2_{[1:t-d]}, u^2_{[1:t-d]}), & u^1_t &= \gamma^1_t(I^1_t) \\ I^2_t &= (y^2_{[1:t]}, u^2_{[1:t-1]}, y^1_{[1:t-d]}, u^1_{[1:t-d]}), & u^2_t &= \gamma^2_t(I^2_t) \end{aligned}$$

Literature overview

- Witsenhausen, Proc IEEE, 1971
- Sandell Athans, TAC 1974
- Yoshikawa, TAC 1975
- Varaiya Walrand, TAC 1978
- Nayyar Mahajan Teneketzis, TAC 2011
- Lamperski Doyle, arxiv 2012

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

Example 2: Delayed sharing information structure*

$$\omega \mapsto (x_1, w^0_{[1:T]}, w^1_{[1:T]}, w^2_{[1:T]}), \qquad \ell(x_t, u^1_t, u^2_t)$$

$$x_{t+1} = f_t(x_t, u^1_t, u^2_t, w^0_t), \qquad y^i_t = h^i_t(x_t, w^i_t)$$

$$I^1_t = (y^1_{[1:t]}, u^1_{[1:t-1]}, y^2_{[1:t-d]}, u^2_{[1:t-d]}), \quad u^1_t = \gamma^1_t(I^1_t)$$

$$I^2_t = (y^2_{[1:t]}, u^2_{[1:t-1]}, y^1_{[1:t-d]}, u^1_{[1:t-d]}), \quad u^2_t = \gamma^2_t(I^2_t)$$

Common information approach

- Contruct a coordinator: $C_t = I^1_t \cap I^2_t = (y^{1,2}_{[1:t-d]}, u^{1,2}_{[1:t-d]})$

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

Example 2: Delayed sharing information structure*

$$\omega \mapsto (x_1, w^0_{[1:T]}, w^1_{[1:T]}, w^2_{[1:T]}), \qquad \ell(x_t, u^1_t, u^2_t)$$

$$\begin{aligned} x_{t+1} &= f_t(x_t, u^1_t, u^2_t, w^0_t), & y^i_t &= h^i_t(x_t, w^i_t) \\ I^1_t &= (y^1_{[1:t]}, u^1_{[1:t-1]}, y^2_{[1:t-d]}, u^2_{[1:t-d]}), & u^1_t &= \gamma^1_t(I^1_t) \\ I^2_t &= (y^2_{[1:t]}, u^2_{[1:t-1]}, y^1_{[1:t-d]}, u^1_{[1:t-d]}), & u^2_t &= \gamma^2_t(I^2_t) \end{aligned}$$

Common information approach

- Contruct a coordinator: $C_t = I^1_t \cap I^2_t = (y^{1,2}_{[1:t-d]}, u^{1,2}_{[1:t-d]})$
- Coordinator chooses $(\varphi^1_t, \varphi^2_t) = \psi_t(C_t)$ where

$$u^1_t = \varphi^1_t(L^1_t), \quad u^2_t = \varphi^2_t(L^2_t)$$

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

Example 2: Delayed sharing information structure*

$$\omega \mapsto (x_1, w^0_{[1:T]}, w^1_{[1:T]}, w^2_{[1:T]}), \qquad \ell(x_t, u^1_t, u^2_t)$$

$$x_{t+1} = f_t(x_t, u^1_t, u^2_t, w^0_t), \qquad y^i_t = h^i_t(x_t, w^i_t)$$

$$I^1_t = (y^1_{[1:t]}, u^1_{[1:t-1]}, y^2_{[1:t-d]}, u^2_{[1:t-d]}), \quad u^1_t = \gamma^1_t(I^1_t)$$

$$I^2_t = (y^2_{[1:t]}, u^2_{[1:t-1]}, y^1_{[1:t-d]}, u^1_{[1:t-d]}), \quad u^2_t = \gamma^2_t(I^2_t)$$

Common information approach

- Contruct a coordinator: $C_t = I^1_t \cap I^2_t = (y^{1,2}_{[1:t-d]}, u^{1,2}_{[1:t-d]})$
- Coordinator chooses $(\varphi^1_t, \varphi^2_t) = \psi_t(C_t)$ where

$$u^1_t = \varphi^1_t(L^1_t), \quad u^2_t = \varphi^2_t(L^2_t)$$

- Same cost $\ell(x_t, u^1_t, u^2_t)$

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

Example 2: Delayed sharing information structure*

$$\omega \mapsto (x_1, w^0_{[1:T]}, w^1_{[1:T]}, w^2_{[1:T]}), \qquad \ell(x_t, u^1_t, u^2_t)$$

$$x_{t+1} = f_t(x_t, u^1_t, u^2_t, w^0_t), \qquad y^i_t = h^i_t(x_t, w^i_t)$$

$$I^1_t = (y^1_{[1:t]}, u^1_{[1:t-1]}, y^2_{[1:t-d]}, u^2_{[1:t-d]}), \quad u^1_t = \gamma^1_t(I^1_t)$$

$$I^2_t = (y^2_{[1:t]}, u^2_{[1:t-1]}, y^1_{[1:t-d]}, u^1_{[1:t-d]}), \quad u^2_t = \gamma^2_t(I^2_t)$$

Common information approach

- Contruct a coordinator: $C_t = I^1_t \cap I^2_t = (y^{1,2}_{[1:t-d]}, u^{1,2}_{[1:t-d]})$
- Coordinator chooses $(\varphi^1_t, \varphi^2_t) = \psi_t(C_t)$ where
$$u^1_t = \varphi^1_t(L^1_t), \quad u^2_t = \varphi^2_t(L^2_t)$$
- Same cost $\ell(x_t, u^1_t, u^2_t)$
- Centralized POMDP: state $= (x_t, L^1_t, L^2_t)$, action $= (\varphi^1_t, \varphi^2_t)$

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

Example 2: Delayed sharing information structure*

$$\omega \mapsto (x_1, w^0_{[1:T]}, w^1_{[1:T]}, w^2_{[1:T]}), \qquad \ell(x_t, u^1_t, u^2_t)$$

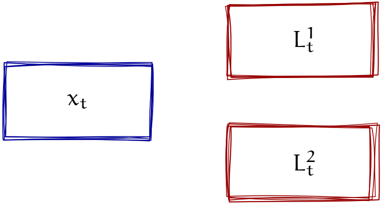
$$x_{t+1} = f_t(x_t, u^1_t, u^2_t, w^0_t), \qquad y^i_t = h^i_t(x_t, w^i_t)$$
$$I^1_t = (y^1_{[1:t]}, u^1_{[1:t-1]}, y^2_{[1:t-d]}, u^2_{[1:t-d]}), \quad u^1_t = \gamma^1_t(I^1_t)$$
$$I^2_t = (y^2_{[1:t]}, u^2_{[1:t-1]}, y^1_{[1:t-d]}, u^1_{[1:t-d]}), \quad u^2_t = \gamma^2_t(I^2_t)$$

Common information approach

- Contruct a coordinator: $C_t = I^1_t \cap I^2_t = (y^{1,2}_{[1:t-d]}, u^{1,2}_{[1:t-d]})$
- Coordinator chooses $(\varphi^1_t, \varphi^2_t) = \psi_t(C_t)$ where
$$u^1_t = \varphi^1_t(L^1_t), \quad u^2_t = \varphi^2_t(L^2_t)$$
- Same cost $\ell(x_t, u^1_t, u^2_t)$
- Centralized POMDP: state $= (x_t, L^1_t, L^2_t)$, action $= (\varphi^1_t, \varphi^2_t)$

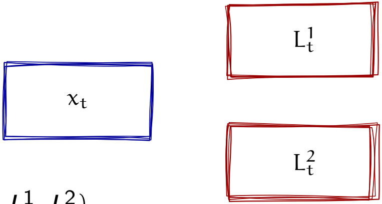
*Nayyar, Mahajan, Teneketzis, "Optimal control strategies in delayed sharing information structures," IEEE TAC 2011

Example 2: Delayed sharing information structure*



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

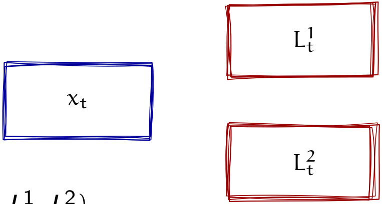
Example 2: Delayed sharing information structure*



- state: (x_t, L_t^1, L_t^2)
- observations: $y^{1.2}_{[1:t-d]}, u^{1,2}_{[1:t-d]}$
- actions: $(\varphi_t^1, \varphi_t^2)$, where $\varphi_t^i \colon L_t^i \mapsto u_t^i$.

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

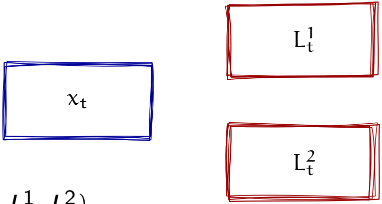
Example 2: Delayed sharing information structure*



- state: (x_t, L_t^1, L_t^2)
- observations: $y^{1,2}_{[1:t-d]}, u^{1,2}_{[1:t-d]}$
- actions: $(\varphi_t^1, \varphi_t^2)$, where $\varphi_t^i \colon L_t^i \mapsto u_t^i$.
- information state:
  $\pi_t = \mathcal{P}(\text{state} \mid \text{history}) = \mathcal{P}(x_t, L_t^1, L_t^2 \mid C_t)$

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

Example 2: Delayed sharing information structure*

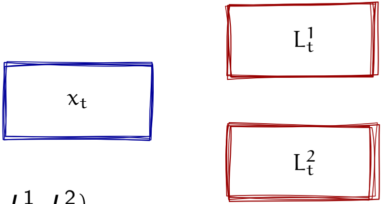


- state: (x_t, L_t^1, L_t^2)
- observations: $y_{[1:t-d]}^{1.2}, u_{[1:t-d]}^{1,2}$
- actions: $(\varphi_t^1, \varphi_t^2)$, where $\varphi_t^i \colon L_t^i \mapsto u_t^i$.
- information state:
  $\pi_t = \mathcal{P}(\text{state} \mid \text{history}) = \mathcal{P}(x_t, L_t^1, L_t^2 \mid C_t)$
- Structure of optimal controller: Without loss of optimality:

$$\varphi_t = \psi_t(\pi_t),$$

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

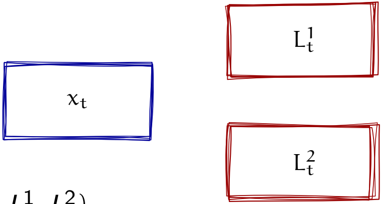
Example 2: Delayed sharing information structure*



- state: (x_t, L_t^1, L_t^2)
- observations: $y^{1.2}_{[1:t-d]}, u^{1,2}_{[1:t-d]}$
- actions: $(\varphi_t^1, \varphi_t^2)$, where $\varphi_t^i \colon L_t^i \mapsto u_t^i$.
- information state:
  $\pi_t = \mathcal{P}(\text{state} \mid \text{history}) = \mathcal{P}(x_t, L_t^1, L_t^2 \mid C_t)$
- Structure of optimal controller: Without loss of optimality:
  $$\varphi_t = \psi_t(\pi_t), \qquad u_t^i = \varphi_t^i(L_t^i) = \gamma_t^i(\pi_t, L_t^i).$$

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

Example 2: Delayed sharing information structure*



- state: (x_t, L_t^1, L_t^2)
- observations: $y^{1.2}_{[1:t-d]}, u^{1,2}_{[1:t-d]}$
- actions: $(\varphi_t^1, \varphi_t^2)$, where $\varphi_t^i \colon L_t^i \mapsto u_t^i$.
- information state:
  $\pi_t = \mathcal{P}(\text{state} \mid \text{history}) = \mathcal{P}(x_t, L_t^1, L_t^2 \mid C_t)$
- Structure of optimal controller: Without loss of optimality:
  $$\varphi_t = \psi_t(\pi_t), \qquad u_t^i = \varphi_t^i(L_t^i) = \gamma_t^i(\pi_t, L_t^i).$$
- Also obtain a dynamic programming decomposition

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

Example 3: For centralized system, common information approach collapses to POMDPs

$\omega \mapsto (x_1, w_{[1:T]})$

Example 3: For centralized system, common information approach collapses to POMDPs

$\omega \mapsto (x_1, w_{[1:T]})$

$$I_t = (y_{[1:t]}, u_{[1:t-1]}),$$

Example 3: For centralized system, common information approach collapses to POMDPs

$\omega \mapsto (x_1, w_{[1:T]})$

$$I_t = (y_{[1:t]}, u_{[1:t-1]}), \quad u_t = \gamma_t(I_t)$$

Example 3: For centralized system, common information approach collapses to POMDPs

$\omega \mapsto (x_1, w_{[1:T]}) \qquad \ell(x_t, u_t)$

$$I_t = (y_{[1:t]}, u_{[1:t-1]}), \quad u_t = \gamma_t(I_t)$$

Example 3: For centralized system, common information approach collapses to POMDPs

$\omega \mapsto (x_1, w_{[1:T]}) \qquad \ell(x_t, u_t)$

$$I_t = (y_{[1:t]}, u_{[1:t-1]}), \quad u_t = \gamma_t(I_t)$$

Common information approach

- Contruct a coordinator: observes $C_t = I_t$.

Example 3: For centralized system, common information approach collapses to POMDPs

$\omega \mapsto (x_1, w_{[1:T]}) \qquad \ell(x_t, u_t)$

$$I_t = (y_{[1:t]}, u_{[1:t-1]}), \quad u_t = \gamma_t(I_t)$$

Common information approach

- Contruct a coordinator: observes $C_t = I_t$.
- Local information $L_t = I_t \setminus C_t = \emptyset$

Example 3: For centralized system, common information approach collapses to POMDPs

$\omega \mapsto (x_1, w_{[1:T]}) \qquad \ell(x_t, u_t)$

$$I_t = (y_{[1:t]}, u_{[1:t-1]}), \quad u_t = \gamma_t(I_t)$$

Common information approach

- Contruct a coordinator: observes $C_t = I_t$.
- Local information $L_t = I_t \setminus C_t = \emptyset$
- Coordinator chooses $\varphi_t = \psi_t(C_t)$ where

$$\varphi_t : L_t \mapsto U_t \quad \implies \varphi_t \equiv u_t$$

Example 3: For centralized system, common information approach collapses to POMDPs

$\omega \mapsto (x_1, w_{[1:T]}) \qquad \ell(x_t, u_t)$

$$I_t = (y_{[1:t]}, u_{[1:t-1]}), \quad u_t = \gamma_t(I_t)$$

Common information approach

- Contruct a coordinator: observes $C_t = I_t$.
- Local information $L_t = I_t \setminus C_t = \emptyset$
- Coordinator chooses $\varphi_t = \psi_t(C_t)$ where

$$\varphi_t : L_t \mapsto U_t \quad \implies \varphi_t \equiv u_t$$

- Same cost $\ell(x_t, u_t^1)$

Example 3: For centralized system, common information approach collapses to POMDPs

$\omega \mapsto (x_1, w_{[1:T]}) \qquad \ell(x_t, u_t)$

$$I_t = (y_{[1:t]}, u_{[1:t-1]}), \quad u_t = \gamma_t(I_t)$$

Common information approach

- Contruct a coordinator: observes $C_t = I_t$.
- Local information $L_t = I_t \setminus C_t = \emptyset$
- Coordinator chooses $\varphi_t = \psi_t(C_t)$ where

$$\varphi_t : L_t \mapsto U_t \qquad \Longrightarrow \varphi_t \equiv u_t$$

- Same cost $\ell(x_t, u_t^1)$
- Coordinator is equivalent to single DM

Outline of the talk

Dynamic programming for centralized control

Dynamic programming for decentralized control

Salient features of common information approach

- Combining with person-by-person approach
- Contructing common information

Using decentralized dynamic programming to specific applications

Conclusion

When does common information approach work?

- Common information is increasing with time: $C_t \subseteq C_{t+1}$.

When does common information approach work?

- Common information is increasing with time: $C_t \subseteq C_{t+1}$.
- Local information ($L_t^i = I_t^i \setminus C_t$) has fixed size/dimension $|L_t^i| \leqslant k$.
- Extension of results to infinite horizon and optimality of time-invariant strategies.

When does common information approach work?

- Common information is increasing with time: $C_t \subseteq C_{t+1}$.
- Local information ($L_t^i = I_t^i \setminus C_t$) has fixed size/dimension $|L_t^i| \leqslant k$.
- Extension of results to infinite horizon and optimality of time-invariant strategies.
- Equivalent to POMDP in which the action space is a function space.

When does common information approach work?

- Common information is increasing with time: $C_t \subseteq C_{t+1}$.
- Local information ($L_t^i = I_t^i \setminus C_t$) has fixed size/dimension $|L_t^i| \leqslant k$.
- Extension of results to infinite horizon and optimality of time-invariant strategies.

- Equivalent to POMDP in which the action space is a function space.
- When all system variables are finite valued, numerical methods for POMDPs (following Smallwood and Sondik) may be used.

When does common information approach work?

- Common information is increasing with time: $C_t \subseteq C_{t+1}$.
- Local information ($L_t^i = I_t^i \setminus C_t$) has fixed size/dimension
  $|L_t^i| \leqslant k$.
- Extension of results to infinite horizon and optimality of time-invariant strategies.
- Equivalent to POMDP in which the action space is a function space.
- When all system variables are finite valued, numerical methods for POMDPs (following Smallwood and Sondik) may be used.
- In the LQG setup, solving possibly non-convex functional optimization problem is difficult.

When does common information approach work?

- Common information is increasing with time: $C_t \subseteq C_{t+1}$.
- Local information ($L_t^i = I_t^i \setminus C_t$) has fixed size/dimension $|L_t^i| \leqslant k$.
- Extension of results to infinite horizon and optimality of time-invariant strategies.

- Equivalent to POMDP in which the action space is a function space.
- When all system variables are finite valued, numerical methods for POMDPs (following Smallwood and Sondik) may be used.

- In the LQG setup, solving possibly non-convex functional optimization problem is difficult.
- DP may be solvable when:
  - π_t is parameteric (e.g., Gaussian)
  - φ_t^i is parameteric (e.g., linear)

Outline of the talk

Structural results allow using the common information approach to a larger class of systems

- For the common information approach to work, local information must have a fixed size.

Structural results allow using the common information approach to a larger class of systems

- For the common information approach to work, local information must have a fixed size.
- Even when the local information does not have a fixed size, a structural result may identify redundant information or an information state that has fixed size.

Structural results allow using the common information approach to a larger class of systems

- For the common information approach to work, local information must have a fixed size.
- Even when the local information does not have a fixed size, a structural result may identify redundant information or an information state that has fixed size.

Person-by-person approach to identify structural result

- Arbitarily fix the strategy of all but one DM, say DM i.
- Look at the centralized subproblem of finding the best response strategy of DM i.

Structural results allow using the common information approach to a larger class of systems

- For the common information approach to work, local information must have a fixed size.
- Even when the local information does not have a fixed size, a structural result may identify redundant information or an information state that has fixed size.

Person-by-person approach to identify structural result

- Arbitarily fix the strategy of all but one DM, say DM i.
- Look at the centralized subproblem of finding the best response strategy of DM i.
- If we identify a structure for the best response strategy of DM i that does not depend on the exact choice of the strategy of other DMs, then that structure is globally optimal.

Example 4: Coupled subsystems with control sharing*

$$x_t = (x_t^1, x_t^2)$$

$$x_t^1 = f_t^1(x_t^1, u_t^1, u_t^2, w_t^1) \qquad x_t^2 = f_t^2(x_t^2, u_t^1, u_t^2, w_t^2)$$

*Mahajan "Optimal decentralized control of coupled subsystems with control sharing," CDC 2011

Example 4: Coupled subsystems with control sharing*

$$x_t = (x_t^1, x_t^2)$$

$$x_t^1 = f_t^1(x_t^1, u_t^1, u_t^2, w_t^1) \qquad x_t^2 = f_t^2(x_t^2, u_t^1, u_t^2, w_t^2)$$

$$\omega \mapsto (x_1, w^0_{[1:T]}, w^1_{[1:T]}, w^2_{[1:T]}),$$

*Mahajan "Optimal decentralized control of coupled subsystems with control sharing," CDC 2011

Example 4: Coupled subsystems with control sharing*

$$x_t = (x_t^1, x_t^2)$$

$$x_t^1 = f_t^1(x_t^1, u_t^1, u_t^2, w_t^1) \qquad x_t^2 = f_t^2(x_t^2, u_t^1, u_t^2, w_t^2)$$

$$\omega \mapsto (x_1, w^0_{[1:T]}, w^1_{[1:T]}, w^2_{[1:T]}),$$

$$I_t^1 = (x^1_{[1:t]}, u^1_{[1:t-1]}, u^2_{[1:t-1]}),$$
$$I_t^2 = (x^2_{[1:t]}, u^2_{[1:t-1]}, u^1_{[1:t-1]}),$$

*Mahajan "Optimal decentralized control of coupled subsystems with control sharing," CDC 2011

Example 4: Coupled subsystems with control sharing*

$$x_t = (x_t^1, x_t^2)$$
$$x_t^1 = f_t^1(x_t^1, u_t^1, u_t^2, w_t^1) \qquad x_t^2 = f_t^2(x_t^2, u_t^1, u_t^2, w_t^2)$$

$$\omega \mapsto (x_1, w^0_{[1:T]}, w^1_{[1:T]}, w^2_{[1:T]}),$$

$$I_t^1 = (x^1_{[1:t]}, u^1_{[1:t-1]}, u^2_{[1:t-1]}), \quad u_t^1 = \gamma_t^1(I_t^1)$$
$$I_t^2 = (x^2_{[1:t]}, u^2_{[1:t-1]}, u^1_{[1:t-1]}), \quad u_t^2 = \gamma_t^2(I_t^2)$$

*Mahajan "Optimal decentralized control of coupled subsystems with control sharing," CDC 2011

Example 4: Coupled subsystems with control sharing*

$$x_t = (x_t^1, x_t^2)$$

$$x_t^1 = f_t^1(x_t^1, u_t^1, u_t^2, w_t^1) \qquad x_t^2 = f_t^2(x_t^2, u_t^1, u_t^2, w_t^2)$$

$$\omega \mapsto (x_1, w^0_{[1:T]}, w^1_{[1:T]}, w^2_{[1:T]}), \qquad \ell(x_t, u_t^1, u_t^2)$$

$$I_t^1 = (x^1_{[1:t]}, u^1_{[1:t-1]}, u^2_{[1:t-1]}), \quad u_t^1 = \gamma_t^1(I_t^1)$$

$$I_t^2 = (x^2_{[1:t]}, u^2_{[1:t-1]}, u^1_{[1:t-1]}), \quad u_t^2 = \gamma_t^2(I_t^2)$$

*Mahajan "Optimal decentralized control of coupled subsystems with control sharing," CDC 2011

Example 4: Coupled subsystems with control sharing*

- Common information: $C_t = I_t^1 \cap I_t^2 = u^{1,2}_{[1:t-1]}$
- Local information: $L_t^i = I_t^i \setminus C_t = x^i_{[1:t]}$.

*Mahajan "Optimal decentralized control of coupled subsystems with control sharing," CDC 2011

Example 4: Coupled subsystems with control sharing*

- Common information: $C_t = I_t^1 \cap I_t^2 = u^{1,2}_{[1:t-1]}$
- Local information: $L_t^i = I_t^i \setminus C_t = x^i_{[1:t]}$.

Directly using the common information approach not tractable

*Mahajan "Optimal decentralized control of coupled subsystems with control sharing," CDC 2011

Example 4:
Coupled subsystems with control sharing*

- Common information: $C_t = I_t^1 \cap I_t^2 = u_{[1:t-1]}^{1,2}$
- Local information: $L_t^i = I_t^i \setminus C_t = x_{[1:t]}^i$.

- Use a person-by-person approach to show that, without loss of optimality

$$u_t^i = \gamma_t^i(x_t^i, u_{[1:t-1]}^1, u_{[1:t-1]}^2)$$

*Mahajan "Optimal decentralized control of coupled subsystems with control sharing," CDC 2011

Example 4: Coupled subsystems with control sharing*

- Common information: $C_t = I_t^1 \cap I_t^2 = u_{[1:t-1]}^{1,2}$
- Local information: $L_t^i = I_t^i \setminus C_t = x_{[1:t]}^i$.

- Use a person-by-person approach to show that, without loss of optimality

$$u_t^i = \gamma_t^i(x_t^i, u_{[1:t-1]}^1, u_{[1:t-1]}^2)$$

- Common information: $C_t = I_t^1 \cap I_t^2 = u_{[1:t-1]}^{1,2}$
- Local information: $L_t^i = I_t^i \setminus C_t = x_t^i$.

*Mahajan "Optimal decentralized control of coupled subsystems with control sharing," CDC 2011

Example 4: Coupled subsystems with control sharing*

- Common information: $C_t = I_t^1 \cap I_t^2 = u_{[1:t-1]}^{1,2}$
- Local information: $L_t^i = I_t^i \setminus C_t = x_{[1:t]}^i$.

- Use a person-by-person approach to show that, without loss of optimality

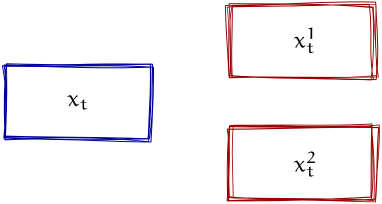
$$u_t^i = \gamma_t^i(x_t^i, u_{[1:t-1]}^1, u_{[1:t-1]}^2)$$

- Common information: $C_t = I_t^1 \cap I_t^2 = u_{[1:t-1]}^{1,2}$
- Local information: $L_t^i = I_t^i \setminus C_t = x_t^i$.

With structural results the common information approach is tractable

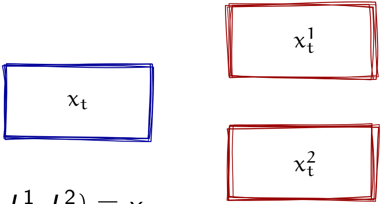
*Mahajan "Optimal decentralized control of coupled subsystems with control sharing," CDC 2011

Example 4: Coupled subsystems with control sharing*



*Mahajan "Optimal decentralized control of coupled subsystems with control sharing," CDC 2011

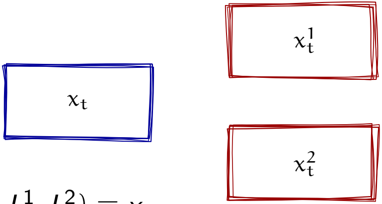
Example 4: Coupled subsystems with control sharing*



- state: $(x_t, L_t^1, L_t^2) \equiv x_t$
- observations: $u_{[1:t-1]}^{1,2}$
- actions: $(\varphi_t^1, \varphi_t^2)$, where $\varphi_t^i \colon x_t^i \mapsto u_t^i$.

*Mahajan "Optimal decentralized control of coupled subsystems with control sharing," CDC 2011

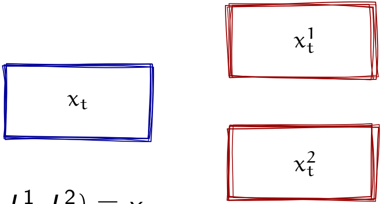
Example 4: Coupled subsystems with control sharing*



- state: $(x_t, L_t^1, L_t^2) \equiv x_t$
- observations: $u_{[1:t-1]}^{1,2}$
- actions: $(\varphi_t^1, \varphi_t^2)$, where $\varphi_t^i \colon x_t^i \mapsto u_t^i$.
- info state: $\pi_t = \mathcal{P}(\text{state} \mid \text{history}) = \mathcal{P}(x_t \mid C_t)$

*Mahajan "Optimal decentralized control of coupled subsystems with control sharing," CDC 2011

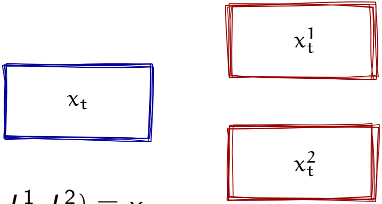
Example 4:
Coupled subsystems with control sharing*



- state: $(x_t, L_t^1, L_t^2) \equiv x_t$
- observations: $u_{[1:t-1]}^{1,2}$
- actions: $(\varphi_t^1, \varphi_t^2)$, where $\varphi_t^i \colon x_t^i \mapsto u_t^i$.
- info state: $\pi_t = \mathcal{P}(\text{state} \mid \text{history}) = \mathcal{P}(x_t \mid C_t)$
$$\equiv \big(\mathcal{P}(x_t^1 \mid C_t), \mathcal{P}(x_t^2 \mid C_t) \big)$$

*Mahajan "Optimal decentralized control of coupled subsystems with control sharing," CDC 2011

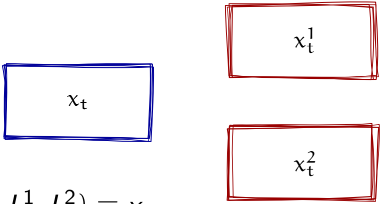
Example 4: Coupled subsystems with control sharing*



- state: $(x_t, L_t^1, L_t^2) \equiv x_t$
- observations: $u_{[1:t-1]}^{1,2}$
- actions: $(\varphi_t^1, \varphi_t^2)$, where $\varphi_t^i \colon x_t^i \mapsto u_t^i$.
- info state: $\pi_t = \mathcal{P}(\text{state} \mid \text{history}) = \mathcal{P}(x_t \mid C_t)$
  $$\equiv \big(\mathcal{P}(x_t^1 \mid C_t), \mathcal{P}(x_t^2 \mid C_t) \big)$$
- Structure of optimal controller: Without loss of optimality:
  $$\varphi_t = \psi_t(\pi_t),$$

*Mahajan "Optimal decentralized control of coupled subsystems with control sharing," CDC 2011

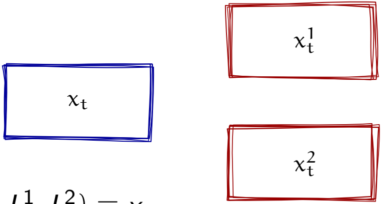
Example 4: Coupled subsystems with control sharing*



- state: $(x_t, L_t^1, L_t^2) \equiv x_t$
- observations: $u_{[1:t-1]}^{1,2}$
- actions: $(\varphi_t^1, \varphi_t^2)$, where $\varphi_t^i \colon x_t^i \mapsto u_t^i$.
- info state: $\pi_t = \mathcal{P}(\text{state} \mid \text{history}) = \mathcal{P}(x_t \mid C_t)$
  $$\equiv \big(\mathcal{P}(x_t^1 \mid C_t), \mathcal{P}(x_t^2 \mid C_t)\big)$$
- Structure of optimal controller: Without loss of optimality:
  $$\varphi_t = \psi_t(\pi_t), \qquad u_t^i = \varphi_t^i(x_t^i) = \gamma_t^i(\pi_t, x_t^i).$$

*Mahajan "Optimal decentralized control of coupled subsystems with control sharing," CDC 2011

Example 4: Coupled subsystems with control sharing*



- state: $(x_t, L_t^1, L_t^2) \equiv x_t$
- observations: $u^{1,2}_{[1:t-1]}$
- actions: $(\varphi_t^1, \varphi_t^2)$, where $\varphi_t^i \colon x_t^i \mapsto u_t^i$.
- info state: $\pi_t = \mathcal{P}(\text{state} \mid \text{history}) = \mathcal{P}(x_t \mid C_t)$

$$\equiv \big(\mathcal{P}(x_t^1 \mid C_t), \mathcal{P}(x_t^2 \mid C_t) \big)$$

- Structure of optimal controller: Without loss of optimality:

$$\varphi_t = \psi_t(\pi_t), \qquad u_t^i = \varphi_t^i(x_t^i) = \gamma_t^i(\pi_t, x_t^i).$$

- Also obtain a dynamic programming decomposition

*Mahajan "Optimal decentralized control of coupled subsystems with control sharing," CDC 2011

Outline of the talk

How to construct common information that is nested with time*

What if $I^1_t \cap I^2_t \not\subseteq I^1_{t+1} \cap I^2_{t+1}$?

*Nayyar "Sequential decision making in decentralized systems," PhD thesis 2011

How to construct common information that is nested with time*

What if $I_t^1 \cap I_t^2 \not\subseteq I_{t+1}^1 \cap I_{t+1}^2$?

Common information is information commonly known to all future DMs

$$C_t = \bigcap_{s \geqslant t} \bigcap_{i=1}^{n} I_s^i$$

(See [Nayyar PhD] for an equivalent representation in terms of observation maps)

*Nayyar "Sequential decision making in decentralized systems," PhD thesis 2011

How to construct common information that is nested with time*

What if $I_t^1 \cap I_t^2 \not\subseteq I_{t+1}^1 \cap I_{t+1}^2$?

Common information is information commonly known to all future DMs

$$C_t = \bigcap_{s \geqslant t} \bigcap_{i=1}^{n} I_s^i$$

(See [Nayyar PhD] for an equivalent representation in terms of observation maps) By construction, $C_t \subseteq C_{t+1}$.

If $C_t = \emptyset$,

*Nayyar "Sequential decision making in decentralized systems," PhD thesis 2011

How to construct common information that is nested with time*

What if $I_t^1 \cap I_t^2 \not\subseteq I_{t+1}^1 \cap I_{t+1}^2$?

Common information is information commonly known to all future DMs

$$C_t = \bigcap_{s \geqslant t} \bigcap_{i=1}^{n} I_s^i$$

(See [Nayyar PhD] for an equivalent representation in terms of observation maps) By construction, $C_t \subseteq C_{t+1}$.

If $C_t = \emptyset$, $L_t^i = I_t^i$,

*Nayyar "Sequential decision making in decentralized systems," PhD thesis 2011

How to construct common information that is nested with time*

What if $I_t^1 \cap I_t^2 \not\subseteq I_{t+1}^1 \cap I_{t+1}^2$?

Common information is information commonly known to all future DMs

$$C_t = \bigcap_{s \geqslant t} \bigcap_{i=1}^{n} I_s^i$$

(See [Nayyar PhD] for an equivalent representation in terms of observation maps) By construction, $C_t \subseteq C_{t+1}$.

If $C_t = \emptyset$, $L_t^i = I_t^i$,

$$\varphi_t^i \colon L_t^i \mapsto u_t^i$$

*Nayyar "Sequential decision making in decentralized systems," PhD thesis 2011

How to construct common information that is nested with time*

What if $I_t^1 \cap I_t^2 \not\subseteq I_{t+1}^1 \cap I_{t+1}^2$?

Common information is information commonly known to all future DMs

$$C_t = \bigcap_{s \geqslant t} \bigcap_{i=1}^{n} I_s^i$$

(See [Nayyar PhD] for an equivalent representation in terms of observation maps) By construction, $C_t \subseteq C_{t+1}$.

If $C_t = \emptyset$, $L_t^i = I_t^i$,

$$\varphi_t^i \colon L_t^i \mapsto u_t^i \quad \therefore \varphi_t^i = \gamma_t^i.$$

Coordinator's approach is same as Witsenhausen's standard form (or the designer's approach) (Witsenhausen, 1973).

*Nayyar "Sequential decision making in decentralized systems," PhD thesis 2011

Outline of the talk

Dynamic programming for centralized control

Dynamic programming for decentralized control

Salient features of common information approach
- Combining with person-by-person approach
- Contructing common information

Using decentralized dynamic programming to specific applications

Conclusion

Using the dynamic programming approach to specific applications

- Real time communication
  - Witsenhausen 1978, Walrand Varaiya 1982, Teneketzis 2006, Mahajan Teneketzis 2010, Yüksel 2013 (to appear), Kaspi Merhav (arixv 2011)
- Networked control systems
  - Walrand Varaiya 1983, Imer Yüksel Başar 2006, Schenato, Sinopoli, Franceschetti, Poola, Sastry 2007, Mahajan Teneketzis 2009,
- Decentralized sequentail hypothesis testing
  - Teneketzis Ho 1987, LaVigna Makowski Baras 1986, Veeravalli Başar Poor 1993, Nayyar Teneketzis 2011
- Mobile cellular networks
  - Hajek Mitzel Yang 2008
- Sensor Networks
  - Imer Başar 2005, Lipsa Martins 2011, Nayyar Başar Teneketzis Veeravalli (arxiv 2012)

Outline of the talk

Dynamic programming for centralized control

Dynamic programming for decentralized control

Salient features of common information approach
- Combining with person-by-person approach
- Contructing common information

Using decentralized dynamic programming to specific applications

Conclusion

Summary of DP approach for decentralized control

The common information approach provides a unified way to obtain DP decomposition for decentralized control.

- Difficulty with decentralized control

Summary of DP approach for decentralized control

The common information approach provides a unified way to obtain DP decomposition for decentralized control.

- Difficulty with decentralized control
  - Since future cost depends on actions of all DMs, each DM needs to predict the actions of other DMs.

Summary of DP approach for decentralized control

The common information approach provides a unified way to obtain DP decomposition for decentralized control.

- Difficulty with decentralized control
  - Since future cost depends on actions of all DMs, each DM needs to predict the actions of other DMs.
  - DMs take actions based on their belief of the state of the world.

Summary of DP approach for decentralized control

The common information approach provides a unified way to obtain DP decomposition for decentralized control.

- Difficulty with decentralized control
  - Since future cost depends on actions of all DMs, each DM needs to predict the actions of other DMs.
  - DMs take actions based on their belief of the state of the world.
  - DMs have different information $\implies$ different beliefs $\implies$ cannot predict actions of others (c.f. signalling).

Summary of DP approach for decentralized control

The common information approach provides a unified way to obtain DP decomposition for decentralized control.

- Difficulty with decentralized control
  - Since future cost depends on actions of all DMs, each DM needs to predict the actions of other DMs.
  - DMs take actions based on their belief of the state of the world.
  - DMs have different information $\implies$ different beliefs $\implies$ cannot predict actions of others (c.f. signalling).
- Conceptual simplification using common information

Summary of DP approach for decentralized control

The common information approach provides a unified way to obtain DP decomposition for decentralized control.

- Difficulty with decentralized control
  - Since future cost depends on actions of all DMs, each DM needs to predict the actions of other DMs.
  - DMs take actions based on their belief of the state of the world.
  - DMs have different information $\implies$ different beliefs $\implies$ cannot predict actions of others (c.f. signalling).
- Conceptual simplification using common information
  - Beliefs based on common information are constent.

Summary of DP approach for decentralized control

The common information approach provides a unified way to obtain DP decomposition for decentralized control.

- Difficulty with decentralized control
  - Since future cost depends on actions of all DMs, each DM needs to predict the actions of other DMs.
  - DMs take actions based on their belief of the state of the world.
  - DMs have different information $\implies$ different beliefs $\implies$ cannot predict actions of others (c.f. signalling).
- Conceptual simplification using common information
  - Beliefs based on common information are constent.
  - We haven't circumvented non-convexity! Each step of the DP is a (possibly non-convex) functional optimization problem

Summary of DP approach for decentralized control

The common information approach provides a unified way to obtain DP decomposition for decentralized control.

- Difficulty with decentralized control
  - Since future cost depends on actions of all DMs, each DM needs to predict the actions of other DMs.
  - DMs take actions based on their belief of the state of the world.
  - DMs have different information $\implies$ different beliefs $\implies$ cannot predict actions of others (c.f. signalling).
- Conceptual simplification using common information
  - Beliefs based on common information are constent.
  - We haven't circumvented non-convexity! Each step of the DP is a (possibly non-convex) functional optimization problem
  - The coordinator is fictitious. All DMs know the common information and may simulate the actions of the coordinator in a decentralized manner.

Summary of DP approach for decentralized control

The common information approach provides a unified way to obtain DP decomposition for decentralized control.

- Difficulty with decentralized control
  - Since future cost depends on actions of all DMs, each DM needs to predict the actions of other DMs.
  - DMs take actions based on their belief of the state of the world.
  - DMs have different information $\implies$ different beliefs $\implies$ cannot predict actions of others (c.f. signalling).
- Conceptual simplification using common information
  - Beliefs based on common information are constent.
  - We haven't circumvented non-convexity! Each step of the DP is a (possibly non-convex) functional optimization problem
  - The coordinator is fictitious. All DMs know the common information and may simulate the actions of the coordinator in a decentralized manner.
  - Obtain structural results and DP decomposition.

Thank you