The Common Information Approach to Decentralized Stochastic Control

D. Teneketzis

Department of EECS
University of Michigan

(Joint work with Ashutosh Nayyar and Aditya Mahajan)
Decentralized Systems

- Communication Networks
Decentralized Systems

- Communication Networks
- Sensor Networks
Decentralized Systems

- Communication Networks
- Sensor Networks
- Transportation Networks
Decentralized Systems

- Communication Networks
- Sensor Networks
- Transportation Networks
- Supply Chain Systems
Decentralized Systems

- Communication Networks
- Sensor Networks
- Transportation Networks
- Supply Chain Systems
- Energy Systems
Decentralized Systems

- Communication Networks
- Sensor Networks
- Transportation Networks
- Supply Chain Systems
- Energy Systems
- Networked Control Systems
Decentralized Systems

- Communication Networks
- Sensor Networks
- Transportation Networks
- Supply Chain Systems
- Energy Systems
- Networked Control Systems
- Social Networks
Decentralized Systems

- Communication Networks
- Sensor Networks
- Transportation Networks
- Supply Chain Systems
- Energy Systems
- Networked Control Systems
- Social Networks
- Organizations
Key Features of Decentralized Systems

- Decisions are made by multiple decision-makers that have different information.
Key Features of Decentralized Systems

- Decisions are made by multiple decision-makers that have different information.

- Information of one decision-maker may depend on decisions made by other decision-makers.
Comparison with Centralized Stochastic Control

- In centralized stochastic control all decisions are made by a centralized decision maker who has access to all the information and has perfect recall.
Comparison with Centralized Stochastic Control

- In centralized stochastic control all decisions are made by a centralized decision maker who has access to all the information and has perfect recall.

- Key assumption of centralized stochastic control is violated in decentralized systems.
Comparison with Centralized Stochastic Control

- In centralized stochastic control all decisions are made by a centralized decision maker who has access to all the information and has perfect recall.

- Key assumption of centralized stochastic control is violated in decentralized systems.

- Techniques from centralized stochastic control cannot be directly applied to decentralized stochastic control.
Approaches to Decentralized Stochastic Control

- Person by person approach
Approaches to Decentralized Stochastic Control

- Person by person approach
- The designer’s approach
Approaches to Decentralized Stochastic Control

- Person by person approach
- The designer’s approach
- Combination of person by person and designer’s approach.
Approaches to Decentralized Stochastic Control

- Person by person approach
- The designer’s approach
- Combination of person by person and designer’s approach.
- Methods exploiting the system’s information structure.
Fix the control strategies of all decision-makers except one (say decision-maker $i$).
Person-by-person Approach

- Fix the control strategies of all decision-makers except one (say decision-maker $i$).

- Optimize with respect to the control strategy of decision maker $i$. 

Qualitative properties of globally optimal control strategies:
- Decentralized detection
- Real-time communication
- Decentralized control

Iterative process for determining person-by-person-optimal strategies (not globally optimal).
Person-by-person Approach

- Fix the control strategies of all decision-makers except one (say decision-maker $i$).

- Optimize with respect to the control strategy of decision maker $i$.

- Qualitative properties of globally optimal control strategies:
Person-by-person Approach

- Fix the control strategies of all decision-makers except one (say decision-maker $i$).

- Optimize with respect to the control strategy of decision maker $i$.

- Qualitative properties of globally optimal control strategies:
  - Decentralized detection
Person-by-person Approach

- Fix the control strategies of all decision-makers except one (say decision-maker $i$).

- Optimize with respect to the control strategy of decision maker $i$.

- Qualitative properties of globally optimal control strategies:
  - Decentralized detection
  - Real-time communication
Person-by-person Approach

- Fix the control strategies of all decision-makers except one (say decision-maker $i$).

- Optimize with respect to the control strategy of decision maker $i$.

- Qualitative properties of globally optimal control strategies:
  - Decentralized detection
  - Real-time communication
  - Decentralized control
Person-by-person Approach

- Fix the control strategies of all decision-makers except one (say decision-maker $i$).

- Optimize with respect to the control strategy of decision maker $i$.

- Qualitative properties of globally optimal control strategies:
  - Decentralized detection
  - Real-time communication
  - Decentralized control

- Iterative process for determining person-by-person-optimal strategies (not globally optimal).
Approaches to Decentralized Stochastic Control

- Person-by-person approach
Approaches to Decentralized Stochastic Control

- Person-by-person approach
- The designer’s approach
Approaches to Decentralized Stochastic Control

- Person-by-person approach
- The designer’s approach
- Combination of person-by-person and designer’s approach
Approaches to Decentralized Stochastic Control

- Person-by-person approach
- The designer’s approach
- Combination of person-by-person and designer’s approach
- Methods exploiting the system’s information structure
Person-by-person Approach

- Fix the control strategies of all decision-makers except one (say decision-maker $i$).

Qualitative properties of globally optimal control strategies:
- Decentralized detection
- Real-time communication
- Decentralized control

Iterative process for determining person-by-person-optimal strategies (not globally optimal).
Person-by-person Approach

- Fix the control strategies of all decision-makers except one (say decision-maker $i$).
- Optimize with respect to the control strategy of decision-maker $i$. 
Person-by-person Approach

- Fix the control strategies of all decision-makers except one (say decision-maker $i$).

- Optimize with respect to the control strategy of decision-maker $i$.

- Qualitative properties of globally optimal control strategies:
Person-by-person Approach

- Fix the control strategies of all decision-makers except one (say decision-maker $i$).

- Optimize with respect to the control strategy of decision-maker $i$.

- Qualitative properties of globally optimal control strategies:
  - Decentralized detection
Person-by-person Approach

- Fix the control strategies of all decision-makers except one (say decision-maker $i$).

- Optimize with respect to the control strategy of decision-maker $i$.

- Qualitative properties of globally optimal control strategies:
  - Decentralized detection
  - Real-time communication
Person-by-person Approach

- Fix the control strategies of all decision-makers except one (say decision-maker $i$).

- Optimize with respect to the control strategy of decision-maker $i$.

- Qualitative properties of globally optimal control strategies:
  - Decentralized detection
  - Real-time communication
  - Decentralized control
Person-by-person Approach

- Fix the control strategies of all decision-makers except one (say decision-maker $i$).

- Optimize with respect to the control strategy of decision-maker $i$.

- Qualitative properties of globally optimal control strategies:
  - Decentralized detection
  - Real-time communication
  - Decentralized control

- Iterative process for determining person-by-person-optimal strategies (not globally optimal).
The Designer’s Approach

- System designer knows the system model and the statistics of the primitive random variable and chooses control strategies for all decision-makers.

Centralized planning problem whose solution results in globally optimal control strategies. Globally optimal strategies determined by a dynamic program where each step is a functional optimization problem. Computationally formidable problem.
System designer knows the system model and the statistics of the primitive random variable and chooses control strategies for all decision-makers.

Centralized planning problem whose solution results in globally optimal control strategies.
The Designer’s Approach

- System designer knows the system model and the statistics of the primitive random variable and chooses control strategies for all decision-makers.

- Centralized planning problem whose solution results in globally optimal control strategies.

- Globally optimal strategies determined by a dynamic program where each step is a functional optimization problem.
The Designer’s Approach

- System designer knows the system model and the statistics of the primitive random variable and chooses control strategies for all decision-makers.

- Centralized planning problem whose solution results in globally optimal control strategies.

- Globally optimal strategies determined by a dynamic program where each step is a functional optimization problem.
  - Computationally formidable problem.
Methods Exploiting System’s Info Structure

- Partially nested information structure
Methods Exploiting System’s Info Structure

- Partially nested information structure
- Stochastically nested information structure
Methods Exploiting System’s Info Structure

- Partially nested information structure
- Stochastically nested information structure
- Information sharing structures
Methods Exploiting System’s Info Structure

- Partially nested information structure
- Stochastically nested information structure
- Information sharing structures
  - Delayed sharing
Methods Exploiting System’s Info Structure

- Partially nested information structure

- Stochastically nested information structure

- Information sharing structures
  - Delayed sharing
  - Periodic sharing
Methods Exploiting System’s Info Structure

- Partially nested information structure

- Stochastically nested information structure

- Information sharing structures
  - Delayed sharing
  - Periodic sharing
  - Control sharing
Methods Exploiting System’s Info Structure

- Partially nested information structure
- Stochastically nested information structure
- Information sharing structures
  - Delayed sharing
  - Periodic sharing
  - Control sharing
  - Broadcast information sharing
- Partially nested information structure

- Stochastically nested information structure

- Information sharing structures
  - Delayed sharing
  - Periodic sharing
  - Control sharing
  - Broadcast information sharing
  - Common and private observations.
Partial history sharing information structure introduced in [NMT]
Partial history sharing information structure introduced in [NMT]

Key features
Partial history sharing information structure introduced in [NMT]

Key features
- Controllers sequentially share part of their past data (observations and control actions) with one another by means of a shared memory.
Partial history sharing information structure introduced in [NMT]

Key features
- Controllers sequentially share part of their past data (observations and control actions) with one another by means of a shared memory.
- All controllers have perfect recall of the commonly shared data (common information).
Model subsumes a large class of decentralized control models where information is shared among controllers.
Model subsumes a large class of decentralized control models where information is shared among controllers.

A general solution methodology was presented in [NMT].
The Common Information Approach

- Model subsumes a large class of decentralized control models where information is shared among controllers.

- A general solution methodology was presented in [NMT].

- Solution methodology based on the common information approach developed in Nayyar [Ph.D. thesis].
The Common Information Approach

- Model subsumes a large class of decentralized control models where information is shared among controllers.

- A general solution methodology was presented in [NMT].

- Solution methodology based on the common information approach developed in Nayyar [Ph.D. thesis]

- Common information approach applicable to all sequential decision-making problems.
The Common Information Approach

Key idea
The Common Information Approach

Key idea

- Reformulate original decentralized stochastic control problem into an equivalent centralized stochastic control problem from the point of view of a fictitious coordinator who has access only to common information and selects prescriptions that map each controller’s private information into actions.
The Common Information Approach

Key idea

- Reformulate original decentralized stochastic control problem into an equivalent centralized stochastic control problem from the point of view of a fictitious coordinator who has access only to common information and selects prescriptions that map each controller’s private information into actions.

- Solve the coordinator’s problem using ideas from Markov decision theory.
The Common Information Approach

- **Key idea**
  - Reformulate original decentralized stochastic control problem into an equivalent centralized stochastic control problem from the point of view of a fictitious coordinator who has access only to common information and selects prescriptions that map each controller’s private information into actions.
  
  - Solve the coordinator's problem using ideas from Markov decision theory
  
  - Translate the results back to the original problem.
The Common Information Approach

Features

- Structural results (qualitative properties) for globally optimal strategies.
The Common Information Approach

Features

- Structural results (qualitative properties) for globally optimal strategies.

- A dynamic program for determining globally optimal strategies for all controllers.
The Common Information Approach

Features

- Structural results (qualitative properties) for globally optimal strategies.
- A dynamic program for determining globally optimal strategies for all controllers.
- Structural results obtain by the common information approach can not be obtained by the person-by-person approach.
The Common Information Approach

Features

- Structural results (qualitative properties) for globally optimal strategies.

- A dynamic program for determining globally optimal strategies for all controllers.

- Structural results obtain by the common information approach can not be obtained by the person-by-person approach.

- Dynamic program obtained by the common information approach is simpler than that obtained by the designer’s approach.
The Common Information Approach

Features

- Provides a unified view of stochastic control
The Common Information Approach

Features

- Provides a unified view of stochastic control
  - If no controller has private information (all information is common) it reduces to a POMDP
The Common Information Approach

Features

- Provides a unified view of stochastic control

  - If no controller has private information (all information is common) it reduces to a POMDP

  - If there is no common information among all controllers it reduces to the designer’s approach.
The Common Information Approach

Features

- Provides a unified view of stochastic control
  
  - If no controller has private information (all information is common) it reduces to a POMDP
  
  - If there is no common information among all controllers it reduces to the designer’s approach.

- (Nayyar, Ph.D. Thesis)
Illustrate the approach.
The Common Information Approach

- Illustrate the approach.

- Present solution to 40-year old conjecture on delayed sharing information structures (Witsenhausen 1971).
The Common Information Approach

- Book chapter contains
The Common Information Approach

- Book chapter contains
  - Point-to-point real-time communication with feedback.
Book chapter contains

- Point-to-point real-time communication with feedback.
- Networked control systems.
Book chapter contains

- Point-to-point real-time communication with feedback.
- Networked control systems.
- Paging and registration in cellular networks.
The Common Information Approach

- Book chapter contains
  - Point-to-point real-time communication with feedback.
  - Networked control systems.
  - Paging and registration in cellular networks.
  - Multiaccess broadcast systems.

*Static team example*

*Dynamic team example*
The Common Information Approach

Book chapter contains

- Point-to-point real-time communication with feedback.
- Networked control systems.
- Paging and registration in cellular networks.
- Multiaccess broadcast systems.
- Static team example* 

*(illustrate how to simplify search of globally optimal strategies)
The Common Information Approach

- Book chapter contains
  - Point-to-point real-time communication with feedback.
  - Networked control systems.
  - Paging and registration in cellular networks.
  - Multiaccess broadcast systems.
  - Static team example*
  - Dynamic team example*
The Common Information Approach

- Book chapter contains
  - Point-to-point real-time communication with feedback.
  - Networked control systems.
  - Paging and registration in cellular networks.
  - Multiaccess broadcast systems.
  - Static team example*
  - Dynamic team example*
  - *(illustrate how to simplify search of globally optimal strategies)
Optimal Control Strategies in Delayed Sharing Information Structures
The Model

Dynamic System

\[ X_{t+1} = f_t(X_t, U_t, W_{0,t}), \quad t = 1, 2, \ldots \]

\[ U_t := (U_{1,t}, U_{2,t}), \text{ two decision makers} \]

\[ \{W_{0,t}, t = 1, 2, 3, \ldots \} \text{ independent of } X_1. \]

Observations

\[ Y_{jt} = h_{jt}(X_t, W_{jt}), \quad t = 1, 2, \ldots, j = 1, 2. \]

\[ Y_{jt} \Rightarrow \text{DM}_j's \text{ observations at } t \]

\[ \{W_{jt}, t = 1, 2, 3, \ldots \}, j = 0, 1, 2 \text{ mutually independent.} \]

Note: All r.v.'s take values in finite spaces.
The Model

Dynamic System

\[ X_{t+1} = f_t (X_t, U_t, W_t^0), \ t = 1, 2, \ldots \]
**The Model**

**Dynamic System**

- $X_{t+1} = f_t \left( X_t, U_t, W_t^0 \right)$, $t = 1, 2, \ldots$
- $U_t := \left( U_t^1, U_t^2 \right)$, two decision makers (DMs)
The Model

Dynamic System

- \( X_{t+1} = f_t(X_t, U_t, W_t^0), \ t = 1, 2, \ldots \)
- \( U_t := (U_1^t, U_2^t), \) two decision makers (DMs)
- \( \{W_t^0, \ t = 1, 2, 3, \ldots \} \) i.i.d. noise process independent of \( X_1. \)
The Model

Dynamic System

- \( X_{t+1} = f_t(X_t, U_t, W_t^0), \ t = 1, 2, \ldots \)
- \( U_t := (U_t^1, U_t^2), \) two decision makers (DMs)
- \( \{W_t^0, \ t = 1, 2, 3, \ldots \} \) i.i.d. noise process independent of \( X_1. \)

Observations
The Model

Dynamic System

- \( X_{t+1} = f_t \left( X_t, U_t, W_t^0 \right), \ t = 1, 2, \ldots \)
- \( U_t := (U_1^t, U_2^t) \), two decision makers (DMs)
- \( \{W_t^0, \ t = 1, 2, 3, \ldots\} \) i.i.d. noise process independent of \( X_1 \).

Observations

- \( Y_t^j = h_t^j \left( X_t, W_t^j \right), \ t = 1, 2, \ldots, j = 1, 2. \)
The Model

Dynamic System

- \( X_{t+1} = f_t \left( X_t, U_t, W_t^0 \right), \ t = 1, 2, \ldots \)
- \( U_t := (U_t^1, U_t^2) \), two decision makers (DMs)
- \( \{W_t^0, \ t = 1, 2, 3, \ldots \} \) i.i.d. noise process independent of \( X_1 \).

Observations

- \( Y_t^j = h_t^j \left( X_t, W_t^j \right), \ t = 1, 2, \ldots, j = 1, 2. \)
- \( Y_t^j \Rightarrow \) DM \( j \)'s observations at \( t \)
The Model

**Dynamic System**

- \( X_{t+1} = f_t (X_{t}, U_{t}, W^0_{t}), \ t = 1, 2, \ldots \)
- \( U_t := (U^1_t, U^2_t) \), two decision makers (DMs)
- \( \{ W^0_t, \ t = 1, 2, 3, \ldots \} \) i.i.d. noise process independent of \( X_1 \).

**Observations**

- \( Y^j_t = h^j_t (X_t, W^j_t), \ t = 1, 2, \ldots, j = 1, 2. \)
- \( Y^j_t \Rightarrow \) DM \( j \)'s observations at \( t \)
- \( \{ W^j_t, \ t = 1, 2, 3, \ldots \} \) i.i.d. noise process
The Model

Dynamic System

- \( X_{t+1} = f_t \left( X_t, U_t, W_t^0 \right) \), \( t = 1, 2, \ldots \)
- \( U_t := (U_t^1, U_t^2) \), two decision makers (DMs)
- \( \{ W_t^0, t = 1, 2, 3, \ldots \} \) i.i.d. noise process independent of \( X_1 \).

Observations

- \( Y_t^j = h_t^j \left( X_t, W_t^j \right) \), \( t = 1, 2, \ldots, j = 1, 2 \).
- \( Y_t^j \Rightarrow \) DM \( j \)'s observations at \( t \)
- \( \{ W_t^j, t = 1, 2, 3, \ldots \} \) i.i.d. noise process
- \( X_1, \left\{ W_t^j, t = 1, 2, 3, \ldots \right\}, j = 0, 1, 2 \) mutually independent.
The Model

Dynamic System

- \( X_{t+1} = f_t \left( X_t, U_t, W^0_t \right), \ t = 1, 2, \ldots \)
- \( U_t := (U^1_t, U^2_t) \), two decision makers (DMs)
- \( \{ W^0_t, \ t = 1, 2, 3, \ldots \} \) i.i.d. noise process independent of \( X_1 \).

Observations

- \( Y^j_t = h^j_t \left( X_t, W^j_t \right), \ t = 1, 2, \ldots, j = 1, 2. \)
- \( Y^j_t \Rightarrow \) DM \( j \)'s observations at \( t \)
- \( \{ W^j_t, \ t = 1, 2, 3, \ldots \} \) i.i.d. noise process
- \( X_1, \ \{ W^j_t, \ t = 1, 2, 3, \ldots \}, j = 0, 1, 2 \) mutually independent.
- Note: All r.v.'s take values in finite spaces.
DMs’ Information

- DMs share their observations and actions with one another with $k$-step delay.
DMs’ Information

- DMs share their observations and actions with one another with $k$-step delay.
- DM $j$’s information at $t$, $j = 1, 2$
The Model

DMs’ Information

- DMs share their observations and actions with one another with $k$-step delay.
- DM $j$’s information at $t$, $j = 1, 2$

$$I^j_t := (Y_{1:t-k}, U_{1:t-k}, Y^j_{t-k+1:t}, U^j_{t-k+1:t-1})$$

$$Y_t := (Y^1_t, Y^2_t)$$

$$U_t := (U^1_t, U^2_t)$$

$$Y_{1:t-k} := (Y_1, Y_2, \ldots, Y_{t-k})$$

$$U_{1:t-k} := (U_1, U_2, \ldots, U_{t-k})$$
DMs’ Information

- DMs share their observations and actions with one another with $k$-step delay.
- DM $j$’s information at $t$, $j = 1, 2$

$$\mathcal{I}_t^j := \left( Y_{1:t-k}^1, U_{1:t-k}^1, Y_{t-k+1:t}^j, U_{t-k+1:t-1}^j \right)$$

$$Y_t := (Y_t^1, Y_t^2)$$

$$U_t := (U_t^1, U_t^2)$$

$$Y_{1:t-k} := (Y_1, Y_2, \ldots, Y_{t-k})$$

$$U_{1:t-k} := (U_1, U_2, \ldots, U_{t-k})$$

$$\mathcal{I}_t^j := \left( C_t, P_t^j \right)$$

$$C_t := (Y_{1:t-k}^1, U_{1:t-k}^1) \quad \text{(Common Info)}$$

$$P_t^j := \left( Y_{t-k+1:t}^j, U_{t-k+1:t-1}^j \right) \quad \text{(Private Info)}$$
The Model

DMs’ Strategies

\[ U_t^j = g_t^j \left( Y_{1:t-k}, U_{1:t-k}, Y_{t-k+1:t}, U_{t-k+1:t-1} \right) \]
DMs’ Strategies

\[ U_t^j = g_t^j \left( Y_{1:t-k}, U_{1:t-k}, Y_{t-k+1:t}^j, U_{t-k+1:t-1}^j \right) \]

\[ g_t^j \Rightarrow \text{DM } j\text{'s control/decision strategy at } t \]
DMs’ Strategies

- $U_t^j = g_t^j \left( Y_{1:t-k}, U_{1:t-k}, Y_{t-k+1:t}, U_{t-k+1:t-1} \right)$

- $g_t^j \Rightarrow$ DM $j$’s control/decision strategy at $t$

- $g^j := \left( g_1^j, g_2^j, \ldots, g_T^j \right) \Rightarrow$ DM $j$’s control/decision strategy
The Model

Instantaneous Cost

- \( l_t(X_t, U_t) = l_t(X_t, U^1_t, U^2_t), \quad t = 1, 2, \ldots, T \)
The Model

Instantaneous Cost

\[ l_t(X_t, U_t) = l_t(X_t, U^1_t, U^2_t), \quad t = 1, 2, \ldots, T \]

Objective
The Model

Instantaneous Cost

- \( l_t (X_t, U_t) = l_t (X_t, U_t^1, U_t^2), \ t = 1, 2, \ldots, T \)

Objective

- Determine a control/decision strategy

\[
g^{1:2} := (g^1, g^2) \\
g^i := (g^i_1, g^i_2, \ldots, g^i_T), \ i = 1, 2
\]

To minimize expected total cost

\[
J (g^{1:2}) = \mathbb{E} \left\{ \sum_{t=1}^{T} l_t (X_t, U_t^1, U_t^2) \right\}
\]
History of the Problem

- Structure of an optimal control strategy conjectured by Witsenhausen in 1971

\[ U_t^i = g_t^j \left( P_t^i, Pr(X_{t-k+1} | C_t) \right) \]  \textit{(*)}

Varaiya-Walrand, 1978  
\textit{(*)} is correct when \( k = 1 \)  
\textit{(*)} is incorrect when \( k > 1 \) (counterexample)  
Structure of an optimal control strategy has remained unknown when \( k > 1 \) until 2011 [NMT, IEEE TAC, July 2011].
History of the Problem

- Structure of an optimal control strategy conjectured by Witsenhausen in 1971
  \[ U_t^j = g_t^j \left( P_t^j, Pr(X_{t-k+1}|C_t) \right) \]  
  \[ (*) \]

- Varaiya-Walrand, 1978
  \[ (*) \text{ is correct when } k = 1 \]
  \[ (*) \text{ is incorrect when } k > 1 \text{ (counterexample)} \]
History of the Problem

- Structure of an optimal control strategy conjectured by Witsenhausen in 1971

\[ U_t^j = g_t^j \left( P_t^j, Pr( X_{t-k+1} | C_t) \right) \]  

- Varaiya-Walrand, 1978

\( (*) \) is correct when \( k = 1 \)

\( (*) \) is incorrect when \( k > 1 \) (counterexample)

- Structure of an optimal control strategy has remained unknown when \( k > 1 \) until 2011 [NMT, IEEE TAC, July 2011].
Step 1 - Show equivalence between the original system and coordinated system

- Identify common information among all controllers at each time $t$
Step 1 - Show equivalence between the original system and coordinated system

- Identify common information among all controllers at each time $t$
- Identify the private information of each controller at each $t$
The Common Information Approach - Key Steps

Step 1 - Show equivalence between the original system and coordinated system

- Identify common information among all controllers at each time $t$
- Identify the private information of each controller at each $t$
- If common information is non-empty, construct a coordinated system in which at each $t$ the coordinator has access to the common information at $t$. 

Based on the common information the controller selects prescriptions (according to a coordination strategy) for each controller. Prescription maps the controller's private information to control action. System dynamics, instantaneous cost function, performance metric same as in original system.

Objective: Select a coordination strategy to minimize the expected total loss.
Step 1 - Show equivalence between the original system and coordinated system

- Identify common information among all controllers at each time $t$
- Identify the private information of each controller at each $t$
- If common information is non-empty, construct a coordinated system in which at each $t$ the coordinator has access to the common information at $t$.
- Based on the common information the controller selects prescriptions (according to a coordination strategy) for each controller.
Step 1 - Show equivalence between the original system and coordinated system

- Identify common information among all controllers at each time $t$
- Identify the private information of each controller at each $t$
- If common information is non-empty, construct a coordinated system in which at each $t$ the coordinator has access to the common information at $t$.
- Based on the common information the controller selects prescriptions (according to a coordination strategy) for each controller.
- Prescription maps the controller’s private information to control action
Step 1 - Show equivalence between the original system and coordinated system

- Identify common information among all controllers at each time $t$
- Identify the private information of each controller at each $t$
- If common information is non-empty, construct a coordinated system in which at each $t$ the coordinator has access to the common information at $t$.
- Based on the common information the controller selects prescriptions (according to a coordination strategy) for each controller.
- Prescription maps the controller’s private information to control action
- System dynamics, instantaneous cost function, performance metric same as in original system
The Common Information Approach - Key Steps

Step 1 - Show equivalence between the original system and coordinated system

- Identify common information among all controllers at each time $t$
- Identify the private information of each controller at each $t$
- If common information is non-empty, construct a coordinated system in which at each $t$ the coordinator has access to the common information at $t$.
- Based on the common information the controller selects prescriptions (according to a coordination strategy) for each controller.
- Prescription maps the controller’s private information to control action
- System dynamics, instantaneous cost function, performance metric same as in original system
- Objective: Select a coordination strategy to minimize the expected total loss.
Step 1 - Construct a coordinated system

The coordinator selects prescriptions $\gamma^1_t, \gamma^2_t$

- $\gamma^1_t$: Space of DM1’s private info $\rightarrow$ Space of DM1’s actions
- $\gamma^2_t$: Space of DM1’s private info $\rightarrow$ Space of DM2’s actions
Step 1 - Construct a coordinated system

- The coordinator selects prescriptions $\gamma_t^1, \gamma_t^2$
  
  $\gamma_t^1$ : Space of DM1’s private info $\rightarrow$ Space of DM1’s actions
  
  $\gamma_t^2$ : Space of DM1’s private info $\rightarrow$ Space of DM2’s actions

- $U_t^1 = \gamma_t^1(P_t^1), \ U_t^2 = \gamma_t^2(P_t^2)$
The Common Information Approach - Key Steps

Step 1 - Construct a coordinated system

- The coordinator selects prescriptions $\gamma_t^1, \gamma_t^2$
  - $\gamma_t^1$ : Space of DM1’s private info $\rightarrow$ Space of DM1’s actions
  - $\gamma_t^2$ : Space of DM1’s private info $\rightarrow$ Space of DM2’s actions
- $U_t^1 = \gamma_t^1(P_t^1), U_t^2 = \gamma_t^2(P_t^2)$
- $\Gamma_t^1 = \psi_t^1(C_t), \Gamma_t^2 = \psi_t^1(C_t)$
Step 1 - Construct a coordinated system

The coordinator selects prescriptions $\gamma_1^t, \gamma_2^t$

- $\gamma_1^t$ : Space of DM1’s private info $\rightarrow$ Space of DM1’s actions
- $\gamma_2^t$ : Space of DM1’s private info $\rightarrow$ Space of DM2’s actions

- $U_1^t = \gamma_1^t (P_1^t)$, $U_2^t = \gamma_2^t (P_2^t)$
- $\Gamma_1^t = \psi_1^t (C_t)$, $\Gamma_2^t = \psi_1^t (C_t)$
- $\{\psi_1^t, \psi_2^t, t = 0, 1, 2, \ldots, T - 1\}$ coordination law
Step 1 - Construct a Coordinated System

- Plant dynamics, instantaneous cost function the same as in the original system
The Common Information Approach - Key Steps

Step 1 - Construct a Coordinated System

- Plant dynamics, instantaneous cost function the same as in the original system

- Objective: Determine a coordination law

\[ \psi := (\psi_0, \psi_1, \ldots, \psi_{T-1}), \psi_t := (\psi^1_t, \psi^2_t) \]

so as to minimize the expected total loss

\[ J(\psi) = \mathbb{E}^{\psi} \left\{ \sum_{t=1}^{T} l_t (X_t, \gamma^1_t, \gamma^2_t) \right\} \]
Step 2 - Formulate coordinated system as POMDP

- Coordinator’s decision problem is a centralized stochastic control problem when common information increases with time.
Step 2 - Formulate coordinated system as POMDP

- Coordinator’s decision problem is a centralized stochastic control problem when common information increases with time.

- Formulate the centralized stochastic control problem as a POMDP. Identify:
Step 2 - Formulate coordinated system as POMDP

- Coordinator’s decision problem is a centralized stochastic control problem when common information increases with time.

- Formulate the centralized stochastic control problem as a POMDP. Identify:
  - (unobserved) state
Step 2 - Formulate coordinated system as POMDP

- Coordinator’s decision problem is a centralized stochastic control problem when common information increases with time.

- Formulate the centralized stochastic control problem as a POMDP. Identify:
  - (unobserved) state
  - observations
Step 2 - Formulate coordinated system as POMDP

- Coordinator’s decision problem is a centralized stochastic control problem when common information increases with time.

- Formulate the centralized stochastic control problem as a POMDP. Identify:
  - (unobserved) state
  - observations
  - control actions
Step 2 - Formulate coordinated system as POMDP

- Coordinator’s decision problem is a centralized stochastic control problem when common information increases with time.

- Formulate the centralized stochastic control problem as a POMDP. Identify:
  - (unobserved) state
  - observations
  - control actions
  - instantaneous cost
Step 2 - Formulate coordinated system as a POMDP

- State Process $S_t$, $t = 1, 2, \ldots, T$

  \[ S_t = (X_t, P_t^1, P_t^2) \]

  \[ P^i_t = (Y_{t-k+1:t}^i, U_{t-k+1:t-1}^i), \quad i = 1, 2 \]
Step 2 - Formulate coordinated system as a POMDP

- State Process $S_t$, $t = 1, 2, \cdots, T$
  \[ S_t = (X_t, P^1_t, P^2_t) \]
  \[ P^i_t = (Y^i_{t-k+1:t}, U^i_{t-k+1:t-1}), \ i = 1, 2 \]

- Observation Process $O_t$, $t = 1, 2, \cdots, T$
  \[ O_t = (Y^1_{t-k}, U^1_{t-k}, Y^2_{t-k}, U^2_{t-k}) \]
Step 2 - Formulate coordinated system as a POMDP

- **State Process** $S_t, t = 1, 2, \cdots, T$
  
  $S_t = (X_t, P^1_t, P^2_t)$
  
  $P^i_t = (Y^i_{t-k+1:t}, U^i_{t-k+1:t-1}), \ i = 1, 2$

- **Observation Process** $O_t, t = 1, 2, \cdots, T$
  
  $O_t = (Y^1_{t-k}, U^1_{t-k}, Y^2_{t-k}, U^2_{t-k})$

- **Action Process**
  
  $A_t = (\Gamma^1_t, \Gamma^2_t)$
Step 2 - Formulate coordinated system as a POMDP

- **State Process** $S_t$, $t = 1, 2, \ldots, T$
  \[ S_t = (X_t, P^1_t, P^2_t) \]
  \[ P^i_t = (Y^i_{t-k+1:t}, U^i_{t-k+1:t-1}) , \ i = 1, 2 \]

- **Observation Process** $O_t$, $t = 1, 2, \ldots, T$
  \[ O_t = (Y^1_{t-k}, U^1_{t-k}, Y^2_{t-k}, U^2_{t-k}) \]

- **Action Process**
  \[ A_t = (\Gamma^1_t, \Gamma^2_t) \]

- We can verify that
  \[ Pr (S_{t+1}, O_{t+1} | S_{1:t}, A_{1:t}) = Pr (S_{t+1}, O_{t+1} | S_t, A_t) \]
  \[ l_t (X_t, U_t) = \tilde{l}_t (S_t, A_t) \]
Step 2 - Formulate coordinated system as a POMDP

- **State Process** $S_t$, $t = 1, 2, \cdots, T$
  
  $$S_t = (X_t, P^1_t, P^2_t)$$
  
  $$P^i_t = (Y^i_{t-k+1:t}, U^i_{t-k+1:t-1}), \ i = 1, 2$$

- **Observation Process** $O_t$, $t = 1, 2, \cdots, T$
  
  $$O_t = (Y^1_{t-k}, U^1_{t-k}, Y^2_{t-k}, U^2_{t-k})$$

- **Action Process**
  
  $$A_t = (\Gamma^1_t, \Gamma^2_t)$$

- We can verify that
  
  $$Pr (S_{t+1}, O_{t+1} | S_{1:t}, A_{1:t}) = Pr (S_{t+1}, O_{t+1} | S_t, A_t)$$
  
  $$l_t (X_t, U_t) = \tilde{l}_t (S_t, A_t)$$

- Decision problem at coordinator is POMDP.
Step 3 - Solve the resultant POMDP

- Use Markov decision theory to identify:
Step 3 - Solve the resultant POMDP

- Use Markov decision theory to identify:
  - the structure of optimal coordination strategies in the coordinated system.
Step 3 - Solve the resultant POMDP

- Use Markov decision theory to identify:
  1. the structure of optimal coordination strategies in the coordinated system.
  2. a dynamic program to obtain an optimal coordinated strategy with such structure.
Step 3 - Solve the resultant POMDP

Let

$$\Theta_t(s) = Pr \left( S_t \mid C_t, \Gamma^1_{1:t-1}, \Gamma^2_{1:t-1} \right)$$
The Common Information Approach - Key Steps

Step 3 - Solve the resultant POMDP

- Let

  \[ \Theta_t(s) = Pr \left( S_t \mid C_t, \Gamma_{1:t-1}^1, \Gamma_{1:t-1}^2 \right) \]

- The realization \( \theta_t \) of \( \Theta_t \) updates according to

  \[ \theta_{t+1} = \eta_t \left( \theta_t, y_{t-k+1}^1, y_{t-k+1}^2, u_{t-k+1}^1, u_{t-k+1}^2, l_t^1, l_t^2 \right) \]

  (non-linear filtering equation)
Step 3 - Solve the resultant POMDP

- Let

\[ \Theta_t(s) = Pr(S_t \mid C_t, \Gamma^1_{1:t-1}, \Gamma^2_{1:t-1}) \]

- The realization \( \theta_t \) of \( \Theta_t \) updates according to

\[ \theta_{t+1} = \eta_t \left( \theta_t, y^1_{t-k+1}, y^2_{t-k+1}, u^1_{t-k+1}, u^2_{t-k+1}, l^1_t, l^2_t \right) \]

(non-linear filtering equation)

- There exists an optimal coordination strategy of the form

\[ \Gamma^1_t = \psi^1_t(\Theta_t), \quad \Gamma^2_t = \psi^2_t(\Theta_t) \]
Step 3 - Solve the resultant POMDP

An optimal coordination strategy is determined by the following DP. Define:

\[
V_t(\theta) = \min_{\gamma_1^T, \gamma_2^T} \mathbb{E} \{ l_T(X_T, \gamma_1^T, \gamma_2^T) | \Theta_T = \theta, \Gamma_T^1 = \gamma_1^T, \Gamma_T^2 = \gamma_2^T \} 
\]

(1)

and for \( t = 1, 2, \ldots, T - 1 \)

\[
V_t(\theta) = \min_{\gamma_1^T, \gamma_2^T} \mathbb{E} \{ l_t(X_t, \gamma_1^t, \gamma_2^t) \\
+ V_{t+1} \left( \eta_t(\theta, Y_{t-K+1}^1, Y_{t-K+1}^2, U_{t-K+1}^1, U_{t-K+1}^2, \gamma_1^t, \gamma_2^t) \right) | \Theta_t = \theta, \Gamma_t^1 = \gamma_1^t, \Gamma_t^2 = \gamma_2^t \} 
\]

(2)
Step 3 - Solve the resultant POMDP

- An optimal coordination strategy is determined by the following DP. Define:

\[
V_t(\theta) = \min_{\gamma_1^t, \gamma_2^t} \mathbb{E}\left\{ l_T(X_T, \gamma_1^T, \gamma_2^T) \mid \Theta_T = \theta, \Gamma_1^T = \gamma_1^T, \Gamma_2^T = \gamma_2^T \right\}
\]

(1)

and for \( t = 1, 2, \cdots, T - 1 \)

\[
V_t(\theta) = \min_{\gamma_1^t, \gamma_2^t} \mathbb{E}\left\{ l_t(X_t, \gamma_1^t, \gamma_2^t) + V_{t+1}(\eta_t(\theta, Y_{t-K+1}^1, Y_{t-K+1}^2, U_{t-K+1}^1, U_{t-K+1}^2, \gamma_1^t, \gamma_2^t)) \mid \Theta_t = \theta, \Gamma_1^t = \gamma_1^t, \Gamma_2^t = \gamma_2^t \right\}
\]

(2)

- For each \( t \) and for each \( \theta \in \Theta_t \), the optimal prescription is the minimizer of \( V_t(\theta) \).
Step 4 - Show equivalence between original system and coordinated system

For any coordination strategy in the coordinated system, there exists a control strategy in the original system that results in the same expected total cost and vice versa.
Step 2: Equivalence of Problems 1 and 2

For going from Problem 2 to Problem 1, consider any choice of coordinator's policy $\psi$. Define $g^1, g^2$ as

$$g_t^k(\cdot, C_t) := \psi_t^k (C_t)$$

for $k = 1, 2$, $t = 1, 2, \ldots, T$. Then,

$$\mathcal{J}(g^1, g^2) = \mathcal{J}(\psi)$$

**Figure:** Going from coordinated system to original system.

**Lemma 1**

Consider any choice of coordinator’s policy $\psi = (\psi_1^1, \psi_1^2, \ldots, \psi_T^1, \psi_T^2)$ in the coordinated system. Define $g^1, g^2$ as

$$g_t^k(\cdot, C_t) := \psi_t^k (C_t)$$

for $k = 1, 2$, $t = 1, 2, \ldots, T$. Then,

$$\mathcal{J}(g^1, g^2) = \mathcal{J}(\psi)$$
Step 2: Equivalence of Problems 1 and 2 for original system as \( g = , t = \).

Consider any choice of coordinator's policy. Then, the decision makers' policies in original and coordinated systems are equivalent.

**Lemma 2**

Consider any choice of decision makers' policies \( g^1, g^2 \) in original system. Define \( \psi \) as

\[
\psi^k_t (C_t) := g^k_t (\cdot, C_t)
\]

for \( k = 1, 2, t = 1, 2, \ldots, T \). Then,

\[
\mathcal{J} (\psi) = \mathcal{J} \left( g^1, g^2 \right)
\]
Step 5 - Translate the solution of the coordinated system to the original system

Use the equivalence of the fourth step to translate the structural results and the dynamic program obtained in the third step for the coordinated system to structural results and dynamic program for the original system.
Step 5 - Translate the solution of the coordinated system to the original system

For the $k$-th step delayed sharing information structure, there exist optimal control/decision strategies of the form

$$U^j_t = g^j \left( P_t^j, \Pr(X_t, P_t^1, P_t^2|C_t) \right), \ j = 1, 2$$
Step 5 - Translate the solution of the coordinated system to the original system

- For the \( k \)-th step delayed sharing information structure, there exist optimal control/decision strategies of the form

\[
U_t^j = g^j \left( P_t^j, Pr \left( X_t, P_t^1, P_t^2 | C_t \right) \right), \ j = 1, 2
\]

- If \( \psi^*_1: T = (\psi^*_1: T, \psi^*_1: T) \) is an optimal coordination strategy (i.e. the solution of DP (1)-(2)), then an optimal control strategy \( g^*_1: T := (g^*_1: T, g^*_1: T) \) is

\[
g^*_t (\cdot, \theta_t) = \psi^*_t (\theta_t), \ j = 1, 2, \ t = 1, 2, \ldots, T.
\]

\[
\theta_t = Pr \left( X_t, P_t^1, P_t^2 | c_t \right)
\]
Witsenhausen’s Conjecture, 1971:
There are optimal policies of the form

\[ U^1_t = g^1_t (P^1_t, Pr(X_{t-k+1} | C_t)) \]
\[ U^2_t = g^2_t (P^2_t, Pr(X_{t-k+1} | C_t)) \]
Comparison with Witsenhausen’s Conjecture

- Witsenhausen’s Conjecture, 1971:
  There are optimal policies of the form

  \[ U^1_t = g^1_t (P^1_t, Pr(X_{t-k+1} | C_t)) \]
  \[ U^2_t = g^2_t (P^2_t, Pr(X_{t-k+1} | C_t)) \]

- Nayyar, Mahajan & Teneketzis, 2011:
  There are optimal policies of the form

  \[ U^1_t = g^1_t (P^1_t, Pr(X_t, P^1_t, P^2_t | C_t)) \]
  \[ U^2_t = g^2_t (P^2_t, Pr(X_t, P^1_t, P^2_t | C_t)) \]
Discussion

- Centralized stochastic control
Centralized stochastic control

Controller’s belief fundamental for predicting future costs
Centralized stochastic control

- Controller's belief fundamental for predicting future costs

- If control strategy for the future is fixed as a function of future beliefs, current belief is a sufficient statistic for future cost under any current action.
Centralized stochastic control

- Controller’s belief fundamental for predicting future costs

- If control strategy for the future is fixed as a function of future beliefs, current belief is a sufficient statistic for future cost under any current action.

- Optimal action is only a function of current belief on state.
Decentralized stochastic control

- Two difficulties
Decentralized stochastic control

- Two difficulties

  (1) Any prediction of future costs must involve a belief on system state and some means of predicting other DMs actions (as cost depend on state and other DMs actions).
Decentralized stochastic control

Two difficulties

1. Any prediction of future costs must involve a belief on system state and some means of predicting other DMs actions (as cost depend on state and other DMs actions).

2. Different DMs have different information ⇒ beliefs formed by each controller and their prediction of future costs can not be consistent.
Discussion

Common Information

- Beliefs based on common info are consistent among all DMs and can serve as a consistent sufficient statistic.
Common Information

- Beliefs based on common info are consistent among all DMs and can serve as a consistent sufficient statistic.

- Based on (realization of) common info DMs can know how each DM will map its private info to control action.
Common Information

- Beliefs based on common info are consistent among all DMs and can serve as a consistent sufficient statistic.

- Based on (realization of) common info DMs can know how each DM will map its private info to control action.

- Common info beliefs and prescriptions must play a fundamental role in a general theory of decentralized stochastic control $\Rightarrow$ fictitious coordinator.
Discussion

- Common Information
Discussion

- Common Information
  - Fictitious coordinator's beliefs based on common info
Discussion

- Common Information
  - Fictitious coordinator's beliefs based on common info
  - Fictitious coordinator's decisions are prescriptions
Discussion

- Common Information
  - Fictitious coordinator's beliefs based on common info
  - Fictitious coordinator's decisions are prescriptions
  - Fictitious coordinator’s objective same as in original problem
Discussion

- Common Information
  - Fictitious coordinator’s beliefs based on common info
  - Fictitious coordinator’s decisions are prescriptions
  - Fictitious coordinator’s objective same as in original problem
  - If common info increases with time coordinator’s problem is a POMDP
Discussion

- Common Information
Common Information

- POMDP equivalent to original problem
Common Information

- POMDP equivalent to original problem

Equivalence $\Rightarrow$
Discussion

- Common Information
  - POMDP equivalent to original problem
  - Equivalence ⇒
    1. Qualitative properties of globally optimal strategies
Discussion

- Common Information
  - POMDP equivalent to original problem
  - Equivalence $\Rightarrow$
    1. Qualitative properties of globally optimal strategies
    2. Sequential decomposition (DP) of the original problem.
Discussion

- Common Information
Discussion

- Common Information
  - Fictitious coordinator invented for conceptual clarity
Discussion

- **Common Information**
  - Fictitious coordinator invented for conceptual clarity
  - Each DM in the original system can solve coordinator’s problem
Discussion

- Common Information
  - Fictitious coordinator invented for conceptual clarity
  - Each DM in the original system can solve coordinator’s problem
  - Presence of coordinator allows to look at problem from the point of view of a “higher level authority”
Discussion

- **Common Information**
  - Fictitious coordinator invented for conceptual clarity
  - Each DM in the original system can solve coordinator’s problem
  - Presence of coordinator allows to look at problem from the point of view of a “higher level authority”
  - Higher level authority simultaneously determines DMs’ prescriptions.
Discussion

- Common Information
Common Information

- Connection of original problem with POMDP can be used for computational purposes.
Common Information

- Connection of original problem with POMDP can be used for computational purposes.

- Value function of POMDP piecewise linear and concave function of coordinator’s belief
Common Information

- Connection of original problem with POMDP can be used for computational purposes.

- Value function of POMDP piecewise linear and concave function of coordinator’s belief

- Efficient algorithms for POMDPs, based on the above property, exist.
Witsenhausen’s Conjecture, 1971:
There are optimal policies of the form

\[ U_t^1 = g_t^1 (P_t^1, Pr(X_{t-n+1}|C_t)) \]
\[ U_t^2 = g_t^2 (P_t^2, Pr(X_{t-n+1}|C_t)) \]
Witsenhausen’s Conjecture, 1971:
There are optimal policies of the form

\[ U^1_t = g^1_t \left( P^1_t, Pr\left(X_{t-n+1} | C_t\right) \right) \]
\[ U^2_t = g^2_t \left( P^2_t, Pr\left(X_{t-n+1} | C_t\right) \right) \]

Nayyar, Mahajan & Teneketzis, 2011:
There are optimal policies of the form

\[ U^1_t = g^1_t \left( P^1_t, Pr\left(X_t, P^1_t, P^2_t | C_t\right) \right) \]
\[ U^2_t = g^2_t \left( P^2_t, Pr\left(X_t, P^1_t, P^2_t | C_t\right) \right) \]
Discussion

- Common Information in other areas
Discussion

- Common Information in other areas
  - Consensus (Washburn-Teneketzis 1984)
Discussion

Common Information in other areas

- Consensus (Washburn-Teneketzis 1984)

- Trading (Milgrom-Stokey 1978)


Thank you!