Value of Common Information in Static Teams

Mohammad Afshari and Aditya Mahajan

Electrical and Computer Engineering Department, McGill University

October 27th, 2016

Decision Making...

Decision Making:

✓ A single agent

✓ Multiple agents

Necessary where centralized is not available/ practical

Decision Making by Multiple Agents:

✓ *Game Theory*: Individual objectives

✓ *Team Theory*: Team objectives

Decision Making...

Decision Making:

✓ A single agent

✓Multiple agents

Necessary where centralized is not available/ practical

Decision Making by Multiple Agents:

✓ *Game Theory*: Individual objectives

✓*Team Theory*: Team objectives

Decision Making...

Decision Making:

✓ A single agent

✓Multiple agents

Necessary where centralized is not available/ practical

Decision Making by Multiple Agents:

✓ *Game Theory*: Individual objectives

✓ *Team Theory*: Team objectives

Decision Making...

Decision Making:

✓ A single agent

✓Multiple agents

Necessary where centralized is not available/ practical

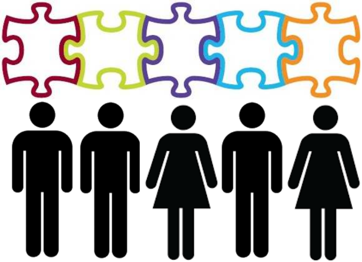
Decision Making by Multiple Agents:

✓ *Game Theory*: Individual objectives

✓ *Team Theory*: Team objectives

What is a Team?

A collection of multiple decision makers(agents) that have access to different information but aim to coordinate their actions to minimize(maximize) a common cost(reward) function.

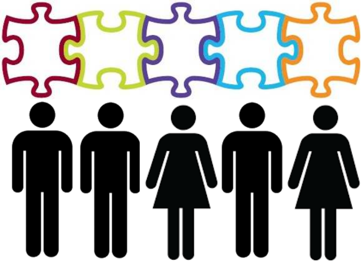


Applications:

Economics, Wireless Sensor Networks, Robotics, Traffic Management, and Smart Grids

What is a Team?

A collection of multiple decision makers(agents) that have access to different information but aim to coordinate their actions to minimize(maximize) a common cost(reward) function.



Applications:

Economics, Wireless Sensor Networks, Robotics, Traffic Management, and Smart Grids

Team Problems

Main challenges:

- ✓ Globally optimal solutions
- ✓ Communication
- ✓ Computational complexity

Types:

- ✓ *Static*: Agents do not affect observations of others
- ✓ *Dynamic*: Agents affect observations of others

Static Team Applications

Dynamic Teams Solution, Decentralized Estimation, Sensor Networks, Multi Robot Task Assignments, etc.

Team Problems

Main challenges:

- ✓ Globally optimal solutions
- ✓ Communication
- ✓ Computational complexity

Types:

- ✓ *Static*: Agents do not affect observations of others
- ✓ *Dynamic*: Agents affect observations of others

Static Team Applications

Dynamic Teams Solution, Decentralized Estimation, Sensor Networks, Multi Robot Task Assignments, etc.

Team Problems

Main challenges:

- ✓ Globally optimal solutions
- ✓ Communication
- ✓ Computational complexity

Types:

- ✓ *Static*: Agents do not affect observations of others
- ✓ *Dynamic*: Agents affect observations of others

Static Team Applications

Dynamic Teams Solution, Decentralized Estimation, Sensor Networks, Multi Robot Task Assignments, etc.

Team Problems

Main challenges:

- ✓ Globally optimal solutions
- ✓ Communication
- ✓ Computational complexity

Types:

- ✓ *Static*: Agents do not affect observations of others
- ✓ *Dynamic*: Agents affect observations of others

Static Team Applications

Dynamic Teams Solution, Decentralized Estimation, Sensor Networks, Multi Robot Task Assignments, etc.

Task Assignment

Example

3 robots with 8 tasks

Which tasks should be assigned to each robot?

Note: Centralized approach is not practical due to communication limits, robustness issues, and scalability.

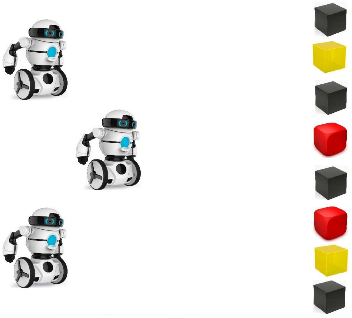
Task Assignment

Example

3 robots with 8 tasks

Which tasks should be assigned to each robot?

Note: Centralized approach is not practical due to communication limits, robustness issues, and scalability.



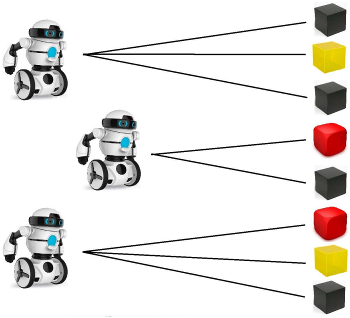
Task Assignment

Example

3 robots with 8 tasks

Which tasks should be assigned to each robot?

Note: Centralized approach is not practical due to communication limits, robustness issues, and scalability.



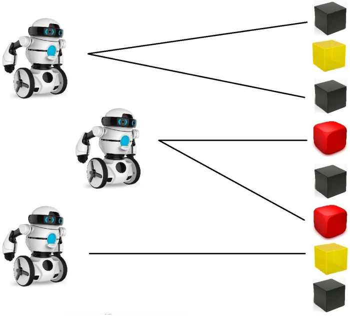
Task Assignment

Example

3 robots with 8 tasks

Which tasks should be assigned to each robot?

Note: Centralized approach is not practical due to communication limits, robustness issues, and scalability.



Static Team Problem Formulation

Problem (Formulation)

- ✓ *n agents try to estimate x and $(x, y_1, \dots, y_n)$ jointly Gaussian*
- ✓ $\mathbb{E}[x] = \bar{x}$, $\mathbb{E}[y_i] = \bar{y}_i$, $\mathbb{E}[xy_i^\intercal] = \Theta_i$, *and* $\mathbb{E}[y_i y_j^\intercal] = \Sigma_{ij}$.
- ✓ *Agent i observes y_i and chooses $u_i = g_i(y_i)$.*

The performance is measured by

$$c(x, u_1, \dots, u_n) = \sum_{i \in N} \sum_{j \in N} u_i^\intercal R_{ij} u_j + 2 \sum_{i \in N} u_i^\intercal P_i x = u^\intercal R u + 2 u^\intercal P x$$

$$u = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}, P = \begin{bmatrix} P_1 \\ \vdots \\ P_n \end{bmatrix}, R = \begin{bmatrix} R_{11} & \cdots & R_{1n} \\ \vdots & \ddots & \vdots \\ R_{n1} & \cdots & R_{nn} \end{bmatrix}$$

Static team problem formulation

Static Team Problem

Assume:

A(1) : $R \geqslant 0$ and $R = R^\intercal$.

A(2) : P, R, Θ, and Σ are common knowledge to all agents.

then choose decision rules $g = (g_1, \dots, g_n)$ for all agents to minimize $J(g) = \mathbb{E}^g[c(x, u)]$.

Theorem (Radner 1962)

The optimal decision rules for static team problem are given by

$$u_i = K_i(y_i - \bar{y}_i) + H_i\bar{x}, \quad \forall i \in N;$$

$$u = \begin{bmatrix} H_1 \\ \vdots \\ H_n \end{bmatrix} = -R^{-1}P, K := \begin{bmatrix} K_1 \\ \vdots \\ K_n \end{bmatrix} = -\Gamma^{-1}\eta$$

where $\eta := \operatorname{vec}(P_1\Theta_1, \dots, P_n\Theta_n)$ *and* $\Gamma := [(\Sigma_{ij} \otimes R_{ij})_{ij}]$
Moreover, the optimal cost J^* *is given by*

$$J^* = -\eta^\intercal\Gamma^{-1}\eta - \bar{x}^\intercal P^\intercal R^{-1}P\bar{x}.$$

Motivation for common information

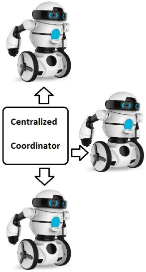
Example

Before the system starts running, suppose it is possible to build an observation channel and broadcast its measurements to all agents. **What is the value of such common information?**.

Motivation for common information

Example

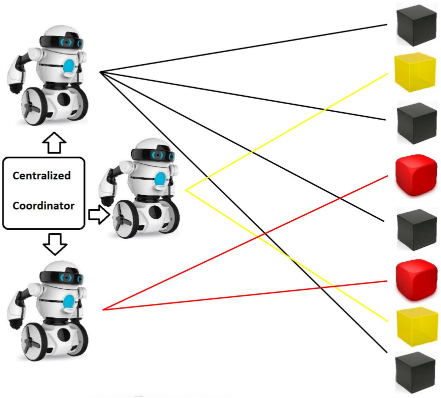
Before the system starts running, suppose it is possible to build an observation channel and broadcast its measurements to all agents. **What is the value of such common information?**.



Motivation for common information

Example

Before the system starts running, suppose it is possible to build an observation channel and broadcast its measurements to all agents. **What is the value of such common information?**.



Static team with common information formulation

- ✓ In addition to x and y_i there is z available to every agent.
- ✓ $(x, y_1, \dots, y_n, z)$ are jointly Gaussian.
- ✓ $\mathbb{E}[z] = \bar{z}$, $\mathbb{E}[xz^\intercal] = \Theta_c$, and $\mathbb{E}[zz^\intercal] = \Sigma_{cc}$.
- ✓ Agent i observes (y_i, z) and choose $\tilde{u}_i = \tilde{g}_i(y_i, z)$.
- ✓ The performance is measured by the same cost.

Problem Statement

Assume (A1) and (A2) are held. Then, choose decision rules $\tilde{g} = (\tilde{g}_1, \dots, \tilde{g}_n)$ for all agents to minimize $\tilde{J}(\tilde{g}) = \mathbb{E}^{\tilde{g}}[c(x, \tilde{u})]$.

Results

It can be solved the same way as Radner solved it.

Theorem (An extension of Radner's Solution)

The optimal decision rules for this problem are given by

$$\tilde{u}_i = \bar{K}_i(y_i - \bar{y}_i) + \bar{L}_i(z - \bar{z}) + H_i\bar{x}, \quad \forall i \in N;$$

$$\begin{bmatrix} \bar{K} \\ \bar{L} \end{bmatrix} := \begin{bmatrix} \bar{K}_1 \\ \vdots \\ \bar{K}_n \\ \bar{L}_1 \\ \vdots \\ \bar{L}_n \end{bmatrix} = -\bar{\Gamma}^{-1}\bar{\eta}$$

Theorem (An extension of Radner's Solution)

$\bar{\eta} \coloneqq \text{vec}(\eta, P_1\Theta_c, \ldots, P_n\Theta_c)$ *and*

$$\bar{\Gamma} \coloneqq \begin{bmatrix} \Gamma & \bar{\Gamma}_{12}^{ij} \\ \bar{\Gamma}_{21}^{ij} & \bar{\Gamma}_{22}^{ij} \end{bmatrix}$$

where, $\bar{\Gamma}_{12}^{ij} \coloneqq [\Sigma_{ic} \otimes R_{ij}]$, $\bar{\Gamma}_{21}^{ij} \coloneqq [\Sigma_{cj} \otimes R_{ij}]$, *and* $\bar{\Gamma}_{11}^{ij} \coloneqq [\Sigma_{cc} \otimes R_{ij}]$. *Moreover, the optimal cost* $\bar{J}^*$ *is given by*

$$\bar{J}^* = -\bar{\eta}^\intercal \bar{\Gamma}^{-1} \bar{\eta} - \bar{x}^\intercal P^\intercal R^{-1} P \bar{x}.$$

Then the effect of the link can be found from the following equation.

$$J^* - \bar{J}^* = \bar{\eta}^T \bar{\Gamma}^{-1} \bar{\eta} - \eta^T \Gamma^{-1} \eta = \bar{\eta}^T \bar{\Gamma}_d \bar{\eta}$$

Positive definiteness of $\bar{\Gamma}_d$ *can be obtained with the same method as Radner(1962).*

Evaluation

In above theorem broadcasting a common information to every agent decreases the optimal cost but, the linear equation which has to be solved is of a higher order.

- ✓ Common information approach
- ✓ Hierarchical approach

Evaluation

In above theorem broadcasting a common information to every agent decreases the optimal cost but, the linear equation which has to be solved is of a higher order.

- ✓ Common information approach
- ✓ Hierarchical approach

Results

Theorem (Common information approach)

The optimal decision rules for this problem are given by

$$\tilde{u}_i = \tilde{K}_i(y_i - \tilde{y}_i) + H_i\tilde{x}, \quad \forall i \in N;$$

$$\tilde{K} \coloneqq \begin{bmatrix} \tilde{K}_1 \\ \vdots \\ \tilde{K}_n \end{bmatrix} = -\tilde{\Gamma}^{-1}\tilde{\eta}$$

where $\tilde{\eta} \coloneqq \mathrm{vec}(P_1\tilde{\Theta}_1, \dots, P_n\tilde{\Theta}_n)$, $\tilde{\Theta}_i \coloneqq \mathbb{E}[xy_i^\intercal | z]$, $\tilde{\Gamma} = [\tilde{\Sigma}_{ij} \otimes R_{ij}]$,
$\tilde{\Sigma}_{ij} = \mathbb{E}[y_i y_j^\intercal | z]$
Moreover, the optimal cost $\tilde{J}^*$ *is given by*

$$J^* = -\tilde{\eta}^\intercal\tilde{\Gamma}^{-1}\tilde{\eta} - \bar{x}P^\intercal R^{-1}P\bar{x} - \mathrm{Tr}(\Theta_c\Sigma_{cc}^{-1}\Theta_c^\intercal P^\intercal R^{-1}P)$$

Results

Theorem (Hierarchical approach)

The optimal decision rules for common information problem are given by

$$\tilde{u}_i = \tilde{K}_i(y_i - \bar{y}_i) + L_i(z - \bar{z}) + H_i\bar{x}, \quad \forall i \in N;$$

where

$$L = -\Gamma_L^{-1}(\eta_c - \Gamma_c\tilde{\Gamma}^{-1}\tilde{\eta})$$

where $\Gamma_L = [(\Sigma_{cc} \otimes R_{ij})_{ij}]$, $\Gamma_c = [(\Sigma_{cj} \otimes R_{ij})_{ij}]$,
$\eta_c \coloneqq \text{vec}(P_1\Theta_c, \ldots, P_n\Theta_c)$.
Moreover, the optimal cost $\tilde{J}^*$ *is given by*

$$J^* = -\tilde{\eta}^\intercal\tilde{\Gamma}^{-1}\tilde{\eta} - \bar{x}P^\intercal R^{-1}P\bar{x} - \text{Tr}(\Theta_c\Sigma_{cc}^{-1}\Theta_c^\intercal P^\intercal R^{-1}P).$$

example

A decentralized estimation problem

- ✓ n sensors. x and y_i are zero mean. Let $w_0, w_1, \cdots, w_n$ be independent Gaussian random variables with distribution $\mathcal{N}(0, \Sigma_{w_i})$.
- ✓ Assume local information of each agent is obtained by $y_i = C_i x + w_i$.
- ✓ Minimize the following performance with cross terms

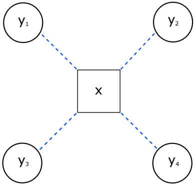
$$c(x, u_1, \ldots, u_n) = \sum_{i=1}^{n} (x - u_i)^\intercal R_{ii} (x - u_i) + \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} (u_i - u_j)^\intercal R_{ij} (u_i - u_j)$$

example

A decentralized estimation problem

- ✓ n sensors. x and y_i are zero mean. Let $w_0, w_1, \cdots, w_n$ be independent Gaussian random variables with distribution $\mathcal{N}(0, \Sigma_{w_i})$.
- ✓ Assume local information of each agent is obtained by $y_i = C_i x + w_i$.
- ✓ Minimize the following performance with cross terms

$$c(x, u_1, \ldots, u_n) = \sum_{i=1}^{n} (x - u_i)^\intercal R_{ii} (x - u_i) + \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} (u_i - u_j)^\intercal R_{ij} (u_i - u_j)$$

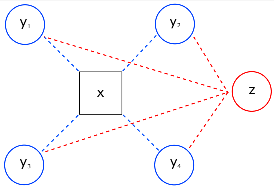


example

A decentralized estimation problem

- ✓ n sensors. x and y_i are zero mean. Let $w_0, w_1, \cdots, w_n$ be independent Gaussian random variables with distribution $\mathcal{N}(0, \Sigma_{w_i})$.
- ✓ Assume local information of each agent is obtained by $y_i = C_i x + w_i$.
- ✓ Minimize the following performance with cross terms

$$c(x, u_1, \ldots, u_n) = \sum_{i=1}^{n} (x - u_i)^\intercal R_{ii} (x - u_i) + \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} (u_i - u_j)^\intercal R_{ij} (u_i - u_j)$$



Solution

Following common information approach the optimal decision rules are given by

$$u_i = \tilde{K}_i(y_i - \tilde{y}_i) + H_i\tilde{x}$$

where

$$\tilde{K} = \tilde{\Gamma}_e^{-1}\tilde{\eta}_e$$

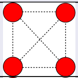
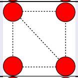
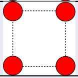
$$L = \begin{bmatrix} \sum_{j=1}^n R_{1j} & -R_{12} & \cdots & -R_{1n} \\ \vdots & \ddots & \vdots & \\ -R_{n1} & -R_{n2} & \cdots & \sum_{j=1}^n R_{nj} \end{bmatrix}^{-1} \begin{bmatrix} R_{11} \\ \vdots \\ R_{nn} \end{bmatrix} = R^{-1}R_d$$

$$\tilde{\Gamma}_e := \begin{bmatrix} \tilde{\Sigma}_{11} \otimes \sum_{j=1}^n R_{1j} & -\tilde{\Sigma}_{12} \otimes R_{12} & \cdots & -\tilde{\Sigma}_{1n} \otimes R_{1n} \\ \vdots & \ddots & \vdots & \\ -\tilde{\Sigma}_{n1} \otimes R_{n1} & -\tilde{\Sigma}_{n2} \otimes R_{n2} & \cdots & \tilde{\Sigma}_{nn} \otimes \sum_{j=1}^n R_{nj} \end{bmatrix}$$

$\tilde{\Gamma}_e$ is strictly diagonally dominant. So by circle theorem it is positive definite. The optimal cost J^* is given by

$$J^* = -\tilde{\eta}_e^\intercal \tilde{\Gamma}_e^{-1} \tilde{\eta}_e - \mathrm{Tr}(\Sigma_x C_0^\intercal (C_0 \Sigma_x C_0^\intercal)^{-1} C_0 \Sigma_x R_d^\intercal R^{-1} R_d) + \mathrm{Tr}(\sum_{i=1}^{n} R_{ii} \Sigma_x)$$

Table : Cost function

Cost matrix	With common info	Without common info
	18.655	22.244
	17.968	21.4807
	17.135	20.8493

Conclusion

- ✓ Value of common information in a LQG system.
- ✓ Two methods to compute the optimal strategy performance.
    - ✓ Common Information Approach: Means and covariance of the observations are the conditional means and covariance given the realization of the common information.
    - ✓ Hierarchical Structure Approach: For high-dimensional (e.g., a video) observations. Communicating corrective terms instead of the common observations.

Future Works:

- ✓ Which link is the best to broadcast?
- ✓ Dynamic Team Problem

Thank you!