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

Aditya Mahajan and Serdar Yüksel
McGill University / Queen’s University

Teams and Information Structures

A decentralized control system is one where there are multiple controllers and these controllers have access to only local information variables.

Such a collection of decision makers who wish to achieve *a common goal* and who have *an agreement on the system variables* (that is, the policy and action spaces and the probability space on which the system is defined) is said to be a *team*.

This is in contrast to settings where the goals may not be aligned, as in *games*.

An example would be a large-scale power-grid with multiple coupled generators. Here, the decision center at each generator has access to only local measurements and based on local information, it must regulate variables that may impact the entire grid.

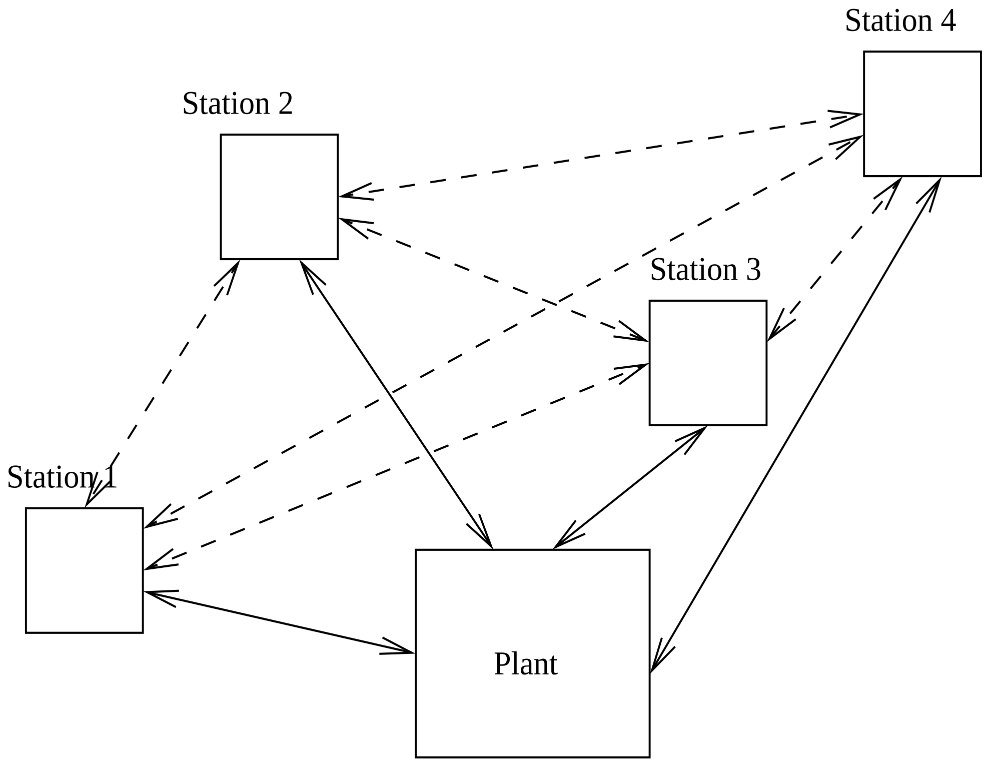


Figure 1: A decentralized control system with possible interactions.

Teams and Information Structures: Sequential Systems

A decentralized control system may either be *sequential* or *non-sequential*.

In a sequential system, the decision makers (DMs) act according to an order that is specified before the system starts running.

In a non-sequential system the DMs act in an order that depends on the realization of the system uncertainty and the control actions of other DMs.

It is more difficult to analyze and formulate a well-posed optimal control problem for non-sequential systems because we need to ensure that it is causal and deadlock-free [Andersland-Teneketzis'92][Teneketzis'96].

In this tutorial, we restrict attention to sequential systems.

Witsenhausen's Intrinsic Team Model

According to Witsenhausen's Intrinsic Model, any (finite horizon) team problem can be characterized by a tuple $\big((\Omega, \mathcal{F}), N, \{U^i, i = 1, \ldots, N\}, \{\mathcal{J}^i, i = 1, \ldots, N\}\big)$.

- $(\Omega, \mathcal{F})$: The realization $\omega \in \Omega$ is called the *primitive variable* of the system. Ω denotes realizations and $\mathcal{F}$ is a set of events in Ω.
- N denotes the number of decision makers (DMs) in the system. Each DM takes only one action.
- $\{U^i, i = 1, \ldots, N\}$ is a collection of action spaces for each DM.
- $\{\mathcal{J}^i, i = 1, \ldots, N\}$ is a collection of sets in $\mathcal{F}$ and represents the *information* available to a DM to take an action. We can also show this with a measurement function with range space I^i. The collection $\{\mathcal{J}^i, i = 1, \ldots, N\}$ is called the *information structure* of the system.
- A *control strategy (policy)*: $\{\gamma^i, i = 1, \ldots, N\}$ where $\gamma^i : \mathcal{J}^i \to U^i$.

Objectives and Policy Spaces

We focus on minimizing a loss function. Other performance measures include minimizing risk, ensuring safety, and ensuring stability.

We will assume that we are given a probability measure P on $(\Omega, \mathcal{F})$ and a real-valued loss function ℓ on $(\Omega \times U^1 \times \cdots \times U^N) =: H$.

Any choice $\gamma = (\gamma^1, \ldots, \gamma^n)$ of the control strategy induces a probability measure P^γ on H. We define the performance $J(\gamma)$ of a strategy as the expected loss, i.e.,

$$J(\gamma) = E^\gamma[\ell(\omega, u^1, \ldots, u^N)]$$

where ω is the primitive variable and u^i is the control action of DM i.

In the intrinsic model, if the system has a control station that takes multiple actions over time, it is modeled as a collection of DMs, one for each time instant.

Example

Consider the following model of a system with two decision makers.

$$\Omega = \{\omega_1, \omega_2, \omega_3\},$$

and action spaces

$$U^1 = \{U(\text{up}), D(\text{down})\},$$
$$U^2 = \{L(\text{left}), R(\text{right})\},$$

with information fields:

$$\mathcal{J}^1 = \{\emptyset, \{\omega_1\}, \{\omega_2, \omega_3\}, \Omega\}$$
$$\mathcal{J}^2 = \{\emptyset, \{\omega_1, \omega_2\}, \{\omega_3\}, \Omega\}.$$

Hence, DM^2 can't distinguish ω_2 from ω_3; and DM^2 can't distinguish ω_1 from ω_2.

Example

Suppose the probability measure P is given by

$$P(\omega_1) = 0.3, \quad P(\omega_2) = 0.3, \quad P(\omega_3) = 0.4,$$

and the loss function $\ell(\omega, u^1, u^2)$ is given by

$\omega_1 \leftrightarrow 0.3$ (rows: u^1, columns: u^2)

	L	R
U	1	0
D	3	1

$\omega_2 \leftrightarrow 0.3$

	L	R
U	2	3
D	2	1

$\omega_3 \leftrightarrow 0.4$

	L	R
U	1	2
D	0	2

For the above model, the unique optimal control strategy is given by

$$\gamma^{1,*}(y^1) = \begin{cases} U, & y^1 \in \{\omega_1\} \\ D, & \text{else} \end{cases}, \quad \gamma^{2,*}(y^2) = \begin{cases} R, & y^2 \in \{\omega_1, \omega_2\} \\ L, & \text{else} \end{cases}$$

Solutions to Teams

A solution to the generalized sequential decentralized stochastic control problem is in general difficult.

Most of the work in the literature has concentrated on identifying solution techniques for specific subclasses.

Typically, these subclasses are characterized on the basis of the information structure of the system.

In this tutorial, we provide an overview of some of these techniques and highlight the sources of difficulties.

Static and dynamic information structures

The simplest, and at first glance, the most critical, distinction is between static and dynamic information structures.

An information structure is called *static* if the observation of all DMs depends only on the primitive random variable (and not on the control actions of others).

Systems that don't have static information structure are said to have *dynamic* information structure.

In such systems, some DMs influence the observations of others through their actions.

Classical, quasiclassical and nonclassical information structures

Centralized control systems are a special case of decentralized control systems.

The characterizing feature is centralization of information, i.e., any DM knows the information available to *all* the DMs that acted before it, or formally, $\mathcal{I}^i \subseteq \mathcal{I}^{i+1}$ for all i.

Such information structures are called *classical*.

Single-person stochastic control problems are typically classical problems.

Classical, *quasiclassical* and nonclassical information structures

A decentralized system is called *quasiclassical* or *partially nested* if the following condition holds:

Whenever DM i can influence DM j, then DM j must know the observations of DM i, or more formally, $\mathcal{J}^i \subseteq \mathcal{J}^j$.

This is an important sub-class: Such problems can be reduced to static problems. We will discuss this explicitly in the context of Linear Quadratic Gaussian teams.

Such settings will also be shown to admit convex formulations for a class of problems.

Classical, quasiclassical and *nonclassical* information structures

Any information structure that is not classical or quasiclassical is called *nonclassical*.

These problems are difficult to analyze and they lack a systematic approach which applies to all such problems.

In this tutorial, sources of this difficulty will be addressed together with some solution techniques.

Example: Decentralized systems admitting a state space model

In a state space model, we assume that system has a *state* variable x_t.

The initial state x_1 is a random variable and the state evolves as

$$\begin{aligned} x_{t+1} &= f_t(x_t, u_t^1, \ldots, u_t^N; w_t^0)\,, \quad t \in \mathcal{T}, \\ y_t^i &= g_t^i(x_t, w_t^i), \end{aligned}$$

The information at station i at time t is

$$I_t^i = \eta_t^i(x_{[1,t]}, \mathbf{y}_{[1,t]}, \mathbf{u}_{[1,t-1]}, \mathbf{w}_{[1,t]}), \tag{1}$$

where $\mathbf{u} = \{u^1, \ldots, u^N\}$ and $x_{[1,t]} = \{x_1, \ldots, x_t\}$.

Common Information Sharing Patterns

Some important information structures are

1. *Complete information sharing*: In complete information sharing, each DM has access to present and past measurements and past actions of all DMs. Such a system is equivalent to a centralized system.

$$I_t^i = \{\mathbf{y}_{[1,t]}, \mathbf{u}_{[1,t-1]}\}.$$

2. *Complete measurement sharing*: In complete measurement sharing, each DM has access to the present and past measurements of all DMs. Note that past control actions are not shared.

$$I_t^i = \{y_{[1,t]}\}.$$

Examples

3. *Delayed information sharing*: In delayed information sharing, each DM has access to n-step delayed measurements and control actions of all DMs.

$$I_t^i = \begin{cases} \{y^i_{[t-n+1,t]}, u^i_{[t-n+1,t-1]}\mathbf{y}_{[1,t-n]}, \mathbf{u}_{[1,t-n]}\}, & t > n \\ \{y^i_{[1,t]}, u^i_{[1,t-1]}\}, & t \leq n \end{cases} \tag{2}$$

4. *Delayed measurement sharing*: In delayed measurement sharing, each DM has access to n-step delayed measurements of all DMs. Note that control actions are not shared.

$$I_t^i = \begin{cases} \{y^i_{[t-n+1,t]}, u^i_{[1,t-1]}, \mathbf{y}_{[1,t-n]}\}, & t > n \\ \{y^i_{[1,t]}, u^i_{[1,t-1]}\}, & t \leq n \end{cases}$$

5. *Delayed control sharing*: In delayed control sharing, each DM has access to n-step delayed control actions of all DMs. Note that measurements are not shared.

$$I_t^i = \begin{cases} \{y^i_{[1,t]}, u^i_{[t-n+1,t-1]}, \mathbf{u}_{[1,t-n]}\}, & t > n \\ \{y^i_{[1,t]}, u^i_{[1,t-1]}\}, & t \leq n \end{cases}$$

Examples

6. *Completely decentralized information*: In a completely decentralized system, no data is shared between the DMs.

$$I_t^i = \{y_{[1,t]}^i, u_{[1,t-1]}^i\}.$$

In all the structures above, each DM has perfect recall (PR).

A DM may only have access to its current observation, in which case the information structure is

$$I_t^i = \{y_t^i\}.$$

The Cost (Loss) Function

To complete the description of the team problem, we have to specify the loss function. We will assume that the loss function is of an additive form:

$$\ell(x_{[1,T]}, \mathbf{u}_{[1,T]}) = \sum_{t \in \mathcal{T}} c(x_t, \mathbf{u}_t) \tag{3}$$

where each term in the summation is known as the *incremental* (or *stagewise*) *loss*.

The objective is to choose control laws γ_t^i such that $u_t^i = \gamma_t^i(I_t^i)$ so as to minimize the expected loss.

In the sequel, we will denote the set of all measurable control laws γ_t^i under the given information structure by Γ_t^i.

Solutions to Static Teams: Convex Static Team Problems

Let

$$J(\gamma) := E[c(\omega, \gamma^1(\eta^1(\omega)), \ldots, \gamma^N(\eta^N(\omega)))].$$

A policy γ^* is *person-by-person-optimal* if for all k

$$J(\gamma^*) \leq J(\gamma^{*1}, \ldots, \gamma^{*(k-1)}, \beta, \gamma^{*(k+1)}, \ldots), \qquad \beta \in \Gamma^k.$$

A policy γ^* is optimal if

$$J(\gamma^*) \leq J(\gamma), \quad \gamma \in \Gamma = \Gamma^1 \times \Gamma^2 \times \cdots \times \Gamma^N$$

Solutions to Static Teams: Convex Static Team Problems

It has been observed by [Radner'62] and [Krainak-Speyer-Marcus'82] that a static team problem with a loss (or cost) function $c(\omega, u^1, \cdots, u^N)$ which is

(i) continuously differentiable in the actions
(ii) convex in the actions

a person-by-person-optimal strategy is globally optimal.

Sketch: Convex Static Team Problems [Radner'62] - [Krainak-Speyer-Marcus'82]

Let γ^* be person-by-person-optimal and γ be another finite-cost policy:

$$
\begin{aligned}
& J(\gamma) - J(\gamma^*) \\
& \geq \lim_{h \to 0} E[\frac{c(\omega, \gamma^* + h(\gamma - \gamma^*)) - c(\omega, \gamma^* + h(\gamma - \gamma^*))}{h}] \\
& = E[\lim_{h \to 0} \frac{c(\omega, \gamma^* + h(\gamma - \gamma^*)) - c(\omega, \gamma^* + h(\gamma - \gamma^*))}{h}] \\
& = E[\sum_i \nabla_{u^i} c(\omega, \gamma^*)|_{u^i = \gamma^*(y^i)}] \\
& = 0
\end{aligned}
$$

Hence, γ^* is optimal.

Application: Static LQG Team Problems

Consider a two-controller system evolving in $\mathbb{R}^n$ with the following description: Let x_1 be Gaussian and

$$x_2 = Ax_1 + B^1u_1^1 + B^2u_1^2 + w_1$$
$$y_1^1 = C^1x_1 + v_1^1, \qquad y_1^2 = C^2x_1 + v_1^2,$$

with w, v^1, v^2 zero-mean, i.i.d. disturbances. For $\rho_1, \rho_2 > 0$, let the goal be the minimization of

$$J(\gamma^1, \gamma^2) = E\left[||x_1||_2^2 + \rho_1||u_1^1||_2^2 + \rho_2||u_1^2||_2^2 + ||x_2||_2^2\right] \tag{4}$$

over the control policies of the form:

$$u_t^i = \mu_t^i(y_1^i), \quad i = 1, 2$$

Solutions to Static Teams: Convex Static Team Problems

For this problem, the cost function is convex in the actions of the decision makers, and it is continuously differentiable.

Linear policies are person by person optimal since linear policies adopted by the other decision makers reduce the problem to a standard Linear Quadratic Gaussian cost optimization problem with partial, Gaussian observations.

Hence, the solution to this problem is affine.

This remarkable observation allows one to show that optimal team policies are affine.

Reduction of Quasi-Classical Problems to Static Problems

For partially observed LQG problem, [Ho-Chu'72] showed that affine control laws are optimal.

The result relies on showing that an invertible linear transformation can convert the partially nested LQG system into a static LQG system:

As an example, consider the following dynamic team with measurements:

$$y^k = C^k \xi + \sum_{i:i \to k} D_{ik} u^i,$$

where ξ is an exogenous random variable picked by nature, and $i \to k$ denotes the precedence relation that the action of DM i affects the information of DM k and u^i is the action of DM i.

Static Reduction of Quasi-Classical LQG Teams

If the information structure is quasi-classical, then

$$\mathcal{I}^k = \{y^k, \{\mathcal{I}^i, i \to k\}\}.$$

That is, DM k has access to the information available to all the signaling DMs. The above original information structure *IS* can be converted to the following static information structure

$$\tilde{y}^k = \Big\{C^k\xi, \{C^i\xi, i \to k\}\Big\},$$

provided that the policies adopted by the agents are deterministic .

The restriction of using only deterministic policies is **without any loss of optimality**: With policies of all other agents fixed (possibly randomized) no agent can benefit from randomized decisions for such team problems [Blackwell'64].

Static Reduction of Sequential Dynamic Teams

[Witsenhausen'88] showed that *any* sequential dynamic team can be converted to a static decentralized control system by an appropriate change of measures and cost functions, provided that a mild conditional density condition holds.

This is a conceptually useful construction, as we discuss later.

Non-Classical Settings

The above analysis is not applicable to stochastic dynamic team problems with nonclassical information as we will see shortly.

As discussed above, nonclassical information structure (*IS*) arises if a Decision Maker (DM) i's action affects the information available to another DM j, who however does not have access to the information available to DM i based on which her action was constructed.

A more precise way of stating this is that the information of DM j is dependent explicitly on the policy (decision rule, or control law) of DM i.

Signaling

What makes a large number of problems possessing the nonclassical information structure difficult is the fact that signaling is present:

Signaling is the policy of communication through control actions.

In this case, the control policies induce a probabilistic map (hence, a channel) from the primitive random variable space to the observation space of the signaled decision makers.

For the nonclassical case, the problem thus also features an information transmission aspect, and the signaling decision maker's objective also includes the design of an optimal measurement channel.

Introducing Probability into Information Structures

The characterizations due to Witsenhausen are independent of probability measures on the system. Incorporating probability lets us expand the class of systems which are effectively partially nested.

Stochastic nestedness [Yüksel'09] captures this: Let x, w^1, w^2 be independent Gaussian variables.

$$y^1 = x + w^1 + w^2, \qquad y^2 = [x + w^1, u^1]$$

$$J(\gamma) = E[|x + u^1 + u^2|^2 + |u^1|^2]$$

DM^1 agent does not have an incentive to signal, since its private information is *not informative*: Optimal policies are affine, even though the problem is not partially nested.

This was extended in [Mahajan-Yüksel'10] to the Intrinsic Model.

Expansion of Information Structures

Solution through Expansion of Information Structures consists of three steps:

1 Provide additional information to make the problem solvable. More information does not degrade performance in a team.

2- Obtain a solution.

3- And show that the solution can be realized with the original information structure.

In the example provided, if we let

$$\tilde{y}^2 = [x + w^1, u^1, y^1]$$

The problem would be partially nested. The optimal solution would be realizable with the original y^2.

Further discussions in [Yüksel'09] [Mahajan-Yüksel'10].

More on Signaling: Comparison of Information Structures

How do we compare information structures?

Let x be a variable and y be observed through $y = g(x, w)$ for some independent variable w.

This measurement model induces a channel: $Q(y \in S|x) = P(g(x, w) \in S|x)$.

Consider the following optimization problem where the controller has access to y.

$$J(P, Q) = \inf_{\gamma} \int_{\mathbb{X}\times\mathbb{Y}} c(x, \gamma(y))Q(dy|x)P(dx),$$

A Concavity Property on the Space of Information Structures

We first present the following concavity result [Yüksel-Linder'10].

The function

$$J(P,Q) = \inf_{\gamma \in \Gamma} \int c(x,\gamma(y)) PQ(dx,dy)$$

is concave in Q.

Hence, the *information structure design* or *channel design* is a *non-convex* problem.

Signaling leads to a non-convex problem.

Comparison of Information Structures

Using the above, we can estbalish a *comparison of experiments* result due to D. Blackwell ('55) and Le Cam ('64).

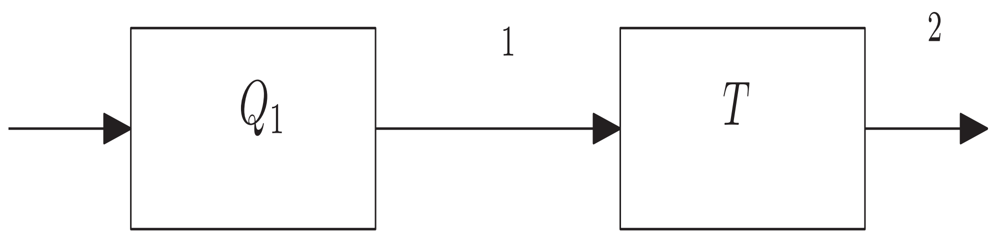


Figure 2: Q_1 defines a better information structure than the composite channel $Q_2 := Q_1 T$.

Comparison of Information Structures: Remarks

When $\mathbb{X}$ is finite, Blackwell showed that the above condition also has a converse: For an information structure to be more informative, weak stochastic degradedness is a necessary condition.

For more general $\mathbb{X}$ and $\mathbb{Y}$, the converse result holds under technical conditions on the stochastic kernels further technical conditions on the set of kernels [Le Cam'96] [Boll'55].

These generalize *garbling* discussions of [Marshak-Radner'72].

The comparison argument applies also for multi-stage problems. [Yüksel-Başar'13].

The comparison result also applies for multi DM problems [Lehrer et al'10].

Further properties of information structures are available in [Yüksel-Linder'10].

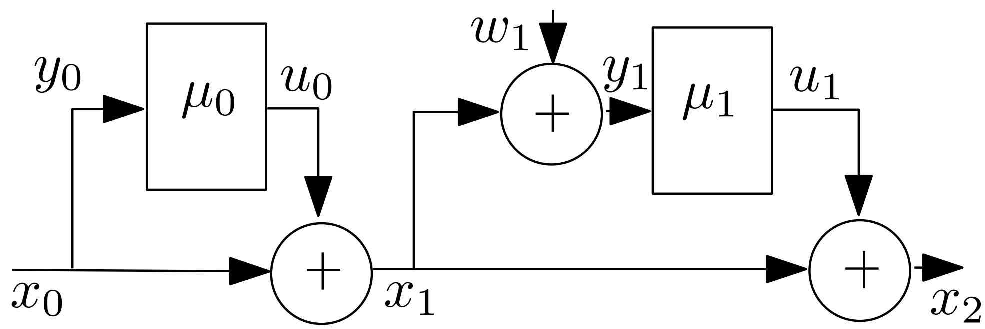
Adding Cost Dependency: Witsenhausen's Counterexample and the Generalized Gaussian test channel

It is known that classical LQG team problems admit solutions which are linear.

[Witsenhausen'68] showed that when there are measurability and information constraints leading to a non-classical information structure, this result is no longer true.

[Witsenhausen'68]: Even LQG problems may admits solutions which are non-linear.

Witsenhausen's Counterexample



$$y_0 = x_0, \qquad u_0 = \mu_0(y_0), \qquad x_1 = x_0 + u_0,$$
$$y_1 = x_1 + w_1, \qquad u_1 = \mu_1(y_1), \qquad x_2 = x_1 + u_1.$$

The goal is to minimize the expected performance index for some $k > 0$

$$Q_W(x, u_0, u_1) = k(u_0)^2 + x_2^2$$

Witsenhausen’s Counterexample

This is the celebrated Witsenhausen’s counterexample.

It is described by a linear system; all primitive variables are Gaussian.

Yet optimal team policy is non-linear [Witsenhausen’68].

Witsenhausen established that a solution exists ([Wu-Verdú’11] provided an alternative proof using Transport Theory), and established that an optimal policy is non-linear.

Properties of optimal solutions have been discussed in [Wu-Verdú’11].

Witsenhausen's Counterexample: Equivalent Representation

Suppose x and w_1 are two independent, zero-mean Gaussian random variables with variance σ^2 and 1. An equivalent representation is:

$$u_0 = \gamma_0(x), \qquad u_1 = \gamma_1(u_0 + w).$$

$$Q_W(x, u_0, u_1) = k(u_0 - x)^2 + (u_1 - u_0)^2\,, \tag{5}$$

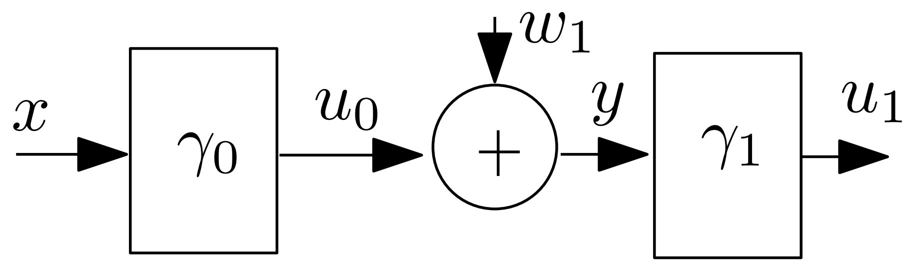


Figure 3: Flow of information in Witsenhausen's counterexample.

Witsenhausen's Counterexample and the Generalized Gaussian test channel

Now consider a different choice for Q:

$$Q_{\mathsf{TC}}(x, u_0, u_1) = k(u_0)^2 + (u_1 - x)^2 \,, \tag{6}$$

where again $k > 0$.

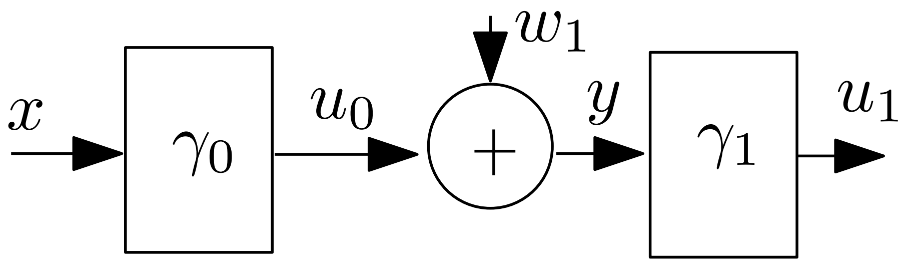


Figure 4: Flow of information in Witsenhausen's counterexample.

Witsenhausen's Counterexample and the Generalized Gaussian test channel

The version of this problem where the soft constraint is replaced by a hard power constraint, $E[(u_0)^2] \leq k$, is known as the *Gaussian Test Channel* (GTC).

In this context γ_0 is the *encoder* and γ_1 the *decoder*, where the latter's optimal choice is clearly the conditional mean of x given y, that is $E[x|y]$.

The best encoder for the GTC can be shown to be linear (a scaled version of the source output, x), which in turn leads to a linear optimal decoder.

The approach here is through information theoretic arguments [Goblick'65][Berger'71].

Witsenhausen's Counterexample and the Generalized Gaussian test channel

Now, consider the more general version of (6):

$$Q_{\text{GTC}}(x, u_0, u_1) = k(u_0)^2 + (u_1 - x)^2 + b_0 u_0 x \;, \tag{7}$$

where b_0 is a scalar. In this case, an optimal solution is linear [Bansal-Başar'87]. The difference between (7) and Witsenhausen's problem is that Q in the former has a product term between the decision rules of the two agents while here it does not.

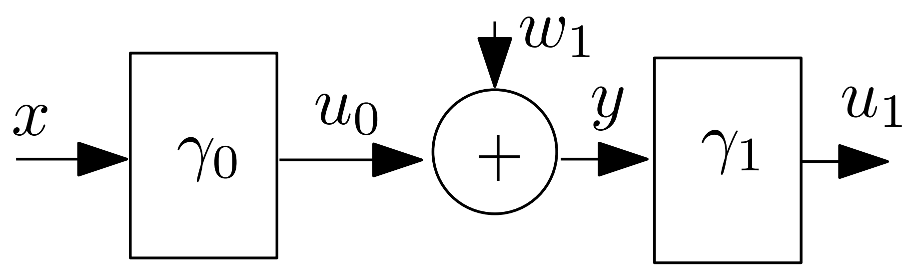


Figure 5: Flow of information in Witsenhausen's counterexample.

Witsenhausen's Counterexample and the Generalized Gaussian test channel

Hence, it is not only the nonclassical nature of the information structure but also the structure of the performance index that determines whether linear policies are optimal in these quadratic dynamic decision problems with Gaussian statistics and nonclassical information.

More on this aspect has been discussed in [Yüksel-Başar'13].

Static Reduction of Witsenhausen's Counterexample

[Witsenhausen'88] showed that it is generally possible to transform any sequential team problem into a static team problem.

The static reduction of the Witsenhausen's counterexample is a two controller static team where the observations y^1 and y^2 of the two controllers are *independent* zero-mean Gaussian random variables with variance σ^2 and 1, respectively.

The control laws γ^1 and γ^2 are to be chosen to minimize

$$J(\gamma^1, \gamma^2) = E\big[(y^1 + u^1 - u^2)^2 + (ku^1)^2 e^{(y^1+u^1)(2y^2-y^1-u^1)/2}\big]$$

The above static reduction clearly illustrates that the cost is not convex.

Linear Policies may be optimal in min-max settings

In a class of quadratic cost minimization problems, one poses the problem not as an expectation minimization but as a min-max optimization where nature acts as the maximizer and the controllers act as the minimizers for cost functions of the form

$$\inf_{\{\gamma^i\}} \sup_{\omega} J(\gamma, \omega),$$

with J being a quadratic function, γ^i denoting controller policies and ω a disturbance with norm constraints.

Linear policies are optimal for a large class of such settings in both encoder-decoder design as well as controller design (see different contexts in [Başar'71] [Başar'85] [Rotkowitz'06] [Gattami-Bernhardsson-Rantzer'12]).

The proof of such results typically use the fact that such min-max problems can be converted to a quadratic optimization problem by a re-parametrization of the cost function.

Concluding Remarks

In the first part of this tutorial, we defined information structures and provided various characterizations.

In particular, we reviewed Witsenhausen's Intrinsic Model.

Solutions to Static and Dynamic Teams have been investigated.

Probability measure and cost dependency of information structures have been discussed and relaxations to partially nested information structure have been obtained.

LQG problems have been discussed as a case study.

In the following talks, further solution approaches and various optimization problems will be discussed.