Problems in the intersection of Information theory and Control

Achilleas Anastasopoulos
anastas@umich.edu

EECS Department
University of Michigan

Dec 5, 2013
Former PhD student: Junghuyn Bae

Current PhD students: Deepanshu Vasal, Jui Wu

Thanks to Demos Teneketzis and his former students Aditya Mahajan and Ashutosh Nayyar
What this talk is NOT about

Controlling of a plant from a distance through a communication link
Instead, we are interested in

A) Viewing point-to-point communication as a Control problem

The act of transmitting a signal (partially) controls the overall communication system, with the hope of bringing it to a “desirable” state.
B) Viewing multi-agent communications as a Control problem

Multiple agents (partially) control a communication network to bring it to a state beneficial for all (cooperatively/competitively)
C) More subtle: Viewing off-line optimization problems relevant to Information theory as control problems, e.g., Shannon capacity

\[
C = \sup_{\{P_{X_t|X_{t-1},Y_{t-1}}(\cdot|\cdot)\}_{t}} \frac{1}{T} \sum_{t=1}^{T} I(X_t \wedge Y_t | Y_{t-1})
\]

Only connection to Communications is the problem origin.

No clear connection to Control either.
Where is the controller? where is the plant? what is the observation/control action?

Achilleas Anastasopoulos
anastas@umich.edu
(U of Michigan)

Problems in the intersection of Information
Whenever there is feedback there is an intimate relation between Communication and Control

One possible classification of problems

1. Use Control techniques to design transmission schemes that achieve general performance measures (e.g., real-time coding\(^1\) \rightarrow \text{examples A,B})

2. Use Control techniques to design transmission schemes that achieve certain information-theoretic-inspired measures (e.g., capacity, error exponents\(^2\) \rightarrow \text{examples A,B})

3. Use Control techniques to evaluate information-theoretic quantities (e.g., capacity, error exponents\(^3\) \rightarrow \text{example C})

---

\(^1\) [Walrand and Varaiya, 1983, Mahajan and Teneketzis, 2009]
\(^3\) [Chen and Berger, 2005, Tatikonda and Mitter, 2009, Kavcic et al., 2009, Bae and Anastasopoulos, 2010]
Whenever there is feedback there is an intimate relation between Communication and Control

One possible classification of problems

1. Use Control techniques to design transmission schemes that achieve general performance measures (e.g., real-time coding\(^1\) → examples A,B)
2. Use Control techniques to design transmission schemes that achieve certain information-theoretic-inspired measures (e.g., capacity, error exponents\(^2\) → examples A,B)
3. Use Control techniques to evaluate information-theoretic quantities (e.g., capacity, error exponents\(^3\) → example C)

\(^1\)[Walrand and Varaiya, 1983, Mahajan and Teneketzis, 2009]
\(^3\)[Chen and Berger, 2005, Tatikonda and Mitter, 2009, Kavcic et al., 2009, Bae and Anastasopoulos, 2010]
Another possible classification of problems

1. Problems involving a single controller (e.g., point-to-point transmission)
2. Problems involving multiple controllers (e.g., multi-user transmission)
   a) Agents act as members of a team trying to achieve a common goal
   b) Agents act strategically having individual goals (games)

Generally, dynamic problems with multiple agents and/or strategic interaction are more difficult: no standard solution methodology

In this presentation we will discuss centralized/decentralized sequential team problems, and a sequential problem with strategic interaction (game)
Another possible classification of problems

1. Problems involving a single controller (e.g., point-to-point transmission)
2. Problems involving multiple controllers (e.g., multi-user transmission)
   a) Agents act as members of a team trying to achieve a common goal
   b) Agents act strategically having individual goals (games)

Generally, dynamic problems with multiple agents and/or strategic interaction are more difficult: no standard solution methodology

In this presentation we will discuss centralized/decentralized sequential team problems, and a sequential problem with strategic interaction (game)
Overview

1. Point-to-point channels with memory and noiseless feedback
2. Multiple access channel with noiseless feedback
3. Cooperative communications in relay networks
Overview

1. Point-to-point channels with memory and noiseless feedback
2. Multiple access channel with noiseless feedback
3. Cooperative communications in relay networks
Background: DMC with noiseless feedback

- Information message $W \in \{1, 2, \ldots, 2^{nR}\}$
- Transmitted symbols $X_t \in \mathcal{X}, \ t = 1, 2, \ldots, n$
- Received symbols $Y_t \in \mathcal{Y}, \ t = 1, 2, \ldots, n$
- Discrete-memoryless channel (DMC) defined by $Q(y_t|x_t)$
- Message estimate $\hat{W} \in \mathcal{W}$
- Encoding functions $X_t = f_t(W, Y^{t-1}), \ t = 1, 2, \ldots, n$
- Decoding function $\hat{W} = d(Y^n)$
Fact [Shannon]: feedback does not increase capacity!

Capacity given by an **off-line, static, single-letter** optimization problem over distributions on $\mathcal{X}$

$$C = \sup_{P_X(\cdot)} I(X_t \land Y_t),$$

with mutual information evaluated as

$$I(X_t \land Y_t) \overset{\text{def}}{=} \sum_x \sum_y P_X(x) Q(y|x) \log \frac{Q(y|x)}{\sum_{x'} P_X(x') Q(y|x')}$$
Channel with memory and noiseless feedback

- Information message $W \in \{1, 2, \ldots, 2^{nR}\}$
- Transmitted symbols $X_t \in \mathcal{X}$, $t = 1, 2, \ldots, n$
- Channel state $S_t \in \mathcal{S}$, $t = 1, 2, \ldots, n$
- Received symbols $Y_t \in \mathcal{Y}$, $t = 1, 2, \ldots, n$ and channel state (perfect Rx CSI)
- Finite state channel (FSC) defined by $Q_y(y_t|x_t, s_t)$, $Q_s(s_{t+1}|s_t, x_t)$
- Message estimate $\hat{W} \in \mathcal{W}$
- Encoding functions $X_t = f_t(W, Y^{t-1}, S^{t-1})$, $t = 1, 2, \ldots, n$ (delayed Tx CSI)
- Decoding function $\hat{W} = d(Y^n, S^n)$
Capacity of this channel is the result of the following **off-line** optimization problem\(^4\) over infinitely many conditional distributions on \(\mathcal{X}\)

\[
C = \sup_{\{P_{X_t|X_{t-1},S^{t-1},Y^{t-1}}\}_{t=1}^{\infty}} \liminf_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} I(X_t, X_{t-1} \land S_t, Y_t|S^{t-1}, Y^{t-1}).
\]

- **Observe:** \(P_{X_t|X_{t-1},S^{t-1},Y^{t-1}} \in \mathcal{X} \times \mathcal{S}^{t-1} \times \mathcal{Y}^{t-1} \to \mathcal{P}(\mathcal{X})\), so its domain increases with \(t\)

- **Q:** How can we utilize Control theory to solve this problem?

---

\(^4\)[Tatikonda and Mitter, 2009, Bae and Anastasopoulos, 2010]
A Markov Decision Process (MDP) is a random process with:

- **State** \( S_t \in \mathcal{S} \),
- **Control action** \( U_t \in \mathcal{U} \),
- **Instantaneous reward** \( R_t \in \mathcal{R} \),

defined by the following dynamics:

\[
P(S_{t+1}|S_t, U_t, R_t) = Q_s(S_{t+1}|S_t, U_t)
\]
\[
P(R_t|S_t, U_t, R_{t-1}) = Q_r(R_t|S_t, U_t)
\]

\[
U_t = g_t(S^t)
\]

**Problem:** Design the sequence of mappings \( g = \{g_t\}_t \) to maximize the average reward

\[
J(g) \overset{\text{def}}{=} \mathbb{E}^g \left\{ \sum_t R_t \right\}
\]
A single controller observes perfectly the state and takes an action (centralized control with perfect state observation)

Solution: Optimal control policy is Markov, i.e.,

\[ U_t = g^*_t(S^t) = g^*_t(S_t) \]

Interpretation: If state is perfectly observed by single controller, then it perfectly summarizes the entire history of observations.
A Partially Observed MDP (POMDP) is a random process with:

- **State** \( S_t \in \mathcal{S} \),
- **Observation** \( Y_t \in \mathcal{Y} \),
- **Control action** \( U_t \in \mathcal{U} \),
- **Instantaneous reward** \( R_t \in \mathcal{R} \),

defined by the following dynamics

\[
P(S_{t+1}|S^t, U^t, R^t, Y^t) = Q_S(S_{t+1}|S_t, U_t)
\]
\[
P(Y_t|S^t, U^{t-1}, R^{t-1}, Y^{t-1}) = Q_Y(Y_t|S_t)
\]
\[
P(R_t|S^t, U^t, R^{t-1}, Y^t) = Q_R(R_t|S_t, U_t)
\]

\[
U_t = g_t(Y^t)
\]

**Problem**: Design the sequence of mappings \( g = \{g_t\}_t \) to maximize the average reward

\[
J^g \overset{\text{def}}{=} \mathbb{E}^g \{ \sum_t R_t \}
\]
Parenthesis: MDPs and POMDPs in 5 mins

Solution: Optimal control policy has the structure,

\[ U_t = g_t^*(Y^t) = g_t^*(\Pi_t) \]

where \( \Pi_t \in \mathcal{P}(\mathcal{S}) \) and \( \Pi_t(s) \overset{\text{def}}{=} Pr(S_t = s|U^{t-1}, Y^t) \quad \forall s \in \mathcal{S} \)

Interpretation: If state is imperfectly observed by controller, then the posterior belief of the state, \( \Pi_t \), perfectly summarizes the entire history of observations

Takeaway: MDPs/POMDPs are useful tools when we want to summarize the dependence of previous observations in our present decisions
Back to our problem: FSC capacity

\[ C = \sup_{\{P_{X|X_{t-1},S^{t-1},Y^{t-1}}\}_{t=1}^{\infty}} \liminf_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} I(X_t, X_{t-1} \land S_t, Y_t | S^{t-1}, Y^{t-1}). \]

- Think of $C$ as the total average reward of some “fictitious” MDP with appropriate states, actions, instantaneous rewards, etc.
- Hint:

\[
I(X_t, X_{t-1} \land S_t, Y_t | S^{t-1}, Y^{t-1}) = \]

\[
= \mathbb{E}\{ \log \frac{P(S_t, Y_t | X_t, X_{t-1}, S^{t-1}, Y^{t-1})}{P(S_t, Y_t | S^{t-1}, Y^{t-1})} \}
\]

\[
= \mathbb{E}\{ \log \frac{\sum_{x_t, x_{t-1}} Q_y(Y_t | S_t, x_t) Q_s(S_t | S_{t-1}, x_{t-1}) P(x_t | x_{t-1}, S^{t-1}, Y^{t-1}) P(x_{t-1} | S^{t-1}, Y^{t-1})}{P(X_t | X_{t-1}, S^{t-1}, Y^{t-1})} \} \]
Back to our problem: FSC capacity

Instantaneous reward depends on

a) some current variables, e.g., $S_{t-1}, S_t, X_{t-1}, X_t, Y_t$

b) the input distribution $P(x_t|x_{t-1}, S^{t-1}, Y^{t-1})$

c) the quantity $P(x_{t-1}|S^{t-1}, Y^{t-1})$

- Define the Control action $U_t \in \mathcal{X} \rightarrow \mathcal{P}(\mathcal{X})$ (conditional distribution of $X_t$ given $X_{t-1}$)

- Allow $U_t$ to be a deterministic function of $S^{t-1}, Y^{t-1}$, $U_t = g_t(S^{t-1}, Y^{t-1})$

- Meaning: $U_t(x_{t-1})(\cdot) = g_t(S^{t-1}, Y^{t-1})(x_{t-1})(\cdot) = P(X_t = \cdot|X_{t-1} = x_{t-1}, S^{t-1}, Y^{t-1})$
Further define the r.v. (information state) $\Theta_t \in \mathcal{P}(\mathcal{X})$ with

$$\Theta_t(x) \overset{\text{def}}{=} \Pr(X_t = x|S^t, Y^t), \quad \forall x \in \mathcal{X}$$

Average instantaneous reward becomes

$$I(X_t, X_{t-1} \wedge S_t, Y_t|S^{t-1}, Y^{t-1}) =$$

$$= \mathbb{E}\left\{ \log \frac{Q_y(Y_t|S_t, X_t)Q_s(S_t|S_{t-1}, X_{t-1})}{\sum_{x_t, x_{t-1}} Q_y(Y_t|S_t, x_t)Q_s(S_t|S_{t-1}, x_{t-1})U_t(x_{t-1})(x_t)\Theta_{t-1}(x_{t-1})} \right\}$$

Observe: dependence on $S^{t-1}, Y^{t-1}$ is "hidden" in the generation of $U_t = g_t(S^{t-1}, Y^{t-1})$ and the evolution of $\Theta_{t-1}$
Theorem ([Bae and Anastasopoulos, 2010])

The original optimization problem is equivalent to an MDP with

- **State** \( \{(S_{t-1}, \Theta_{t-1})\}_t \)
- **Control action** \( U_t \)

Markov policies are optimal, i.e., optimal actions can be of the form

\[
U_t = g^*_t(S_{t-1}, \Theta_{t-1}) \Leftrightarrow P^*_X|X_{t-1}, S'_{t-1}, Y'_{t-1} = P^*_X|X_{t-1}, S_{t-1}, \Theta_{t-1}
\]

Capacity is now simplified to a **single-letter** expression

\[
C = \sup_{P_X|X_{t-1}, S_{t-1}, \Theta_{t-1}} I(X_t, X_{t-1} \land S_t, Y_t|S_{t-1}, \Theta_{t-1})
\]
Lessons learned

- Complex optimization problems in Information theory can be translated to simple centralized control problems.
- Starting point: some multi-letter capacity expression.
- Methodology: appropriately define a control system to unveil an MDP/POMDP.
- These ideas can also help in designing actual on-line capacity-achieving transmission schemes (not in this talk).
Overview

1. Point-to-point channels with memory and noiseless feedback

2. Multiple access channel with noiseless feedback

3. Cooperative communications in relay networks
System model: the information theoretic setup

- Messages $W^i \in \{1, 2, \ldots, 2^{nR^i}\}$, $i = 1, 2$
- Transmitted symbols $X^i_t \in \mathcal{X}^i$, $i = 1, 2$, $t = 1, 2, \ldots, n$
- Received symbols $Y_t \in \mathcal{Y}$, $t = 1, 2, \ldots, n$
- Discrete-memoryless MAC (DM-MAC) $Q(y_t|x^1_t, x^2_t)$
- Message estimates $(\hat{W}^1, \hat{W}^2) \in \mathcal{W}^1 \times \mathcal{W}^2$
- Encoding functions $X^i_t = f^i_t(W^i, Y^{t-1})$, $i = 1, 2$, $t = 1, 2, \ldots, n$
- Decoding function $(\hat{W}^1, \hat{W}^2) = d(Y^n)$
NLF-MAC capacity


- For the general NLF-MAC the capacity is not known as a single-letter expression!

- Multi-letter expression was developed by [Kramer, 2003]
Fact ([Kramer, 2003], [Salehi, 1978])

The problem of evaluating the NLF-MAC capacity region is equivalent to solving the following optimization problem for every $\lambda = (\lambda_1, \lambda_2, \lambda_3) \geq 0$

$$J_\lambda = \sup \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \left\{ \lambda_1 I(X_1^t \land Y_t | X^{2,t}, Y^{t-1}) + \lambda_2 I(X_2^t \land Y_t | X^{1,t}, Y^{t-1}) + \lambda_3 I(X_1^t, X_2^t \land Y_t | Y^{t-1}) \right\},$$

and the supremum is over all input distributions of the form

$$\{P(X_1^t | X^{1,t-1}, Y^{t-1}), P(X_2^t | X^{2,t-1}, Y^{t-1})\}_t$$
Multiple access channel with noiseless feedback

Why is the NLF-MAC capacity still an open problem?

Three main difficulties

1. The optimal input distributions on $X_i^t$ depend on entire history $X_i^{t-1}$ and $Y^{t-1}$
2. The optimization problem involves two controllers with different observations (decentralized control!)
3. The per-stage rewards (mutual info expressions) are complicated functions of the involved random variables

Claim: we can address the first two of the above three difficulties

Solve a slightly different problem: real-time communication over the NLF-MAC
Why is the NLF-MAC capacity still an open problem?

- Three main difficulties
  1. The optimal input distributions on $X^i_t$ depend on entire history $X^{i,t-1}$ and $Y^{t-1}$
  2. The optimization problem involves two controllers with different observations (decentralized control!)
  3. The per-stage rewards (mutual info expressions) are complicated functions of the involved random variables

- Claim: we can address the first two of the above three difficulties

- Solve a slightly different problem: real-time communication over the NLF-MAC
System model: real-time communication

Same model as before, except

- Message estimates for each time $t$, $(\hat{W}_t^1, \hat{W}_t^2) \in \mathcal{W}^1 \times \mathcal{W}^2$
- Decoding functions $(\hat{W}_t^1, \hat{W}_t^2) = d_t(Y^t)$, $t = 1, 2, \ldots, n$
- Instantaneous reward function $\rho_t(W^1, W^2, \hat{W}_t^1, \hat{W}_t^2)$
- Find a set of encoding/decoding functions $g \overset{\text{def}}{=} \{f^1_t, f^2_t, d_t\}_t$ that maximize

$$J(g) = \mathbb{E}_g \left\{ \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \rho_t(W^1, W^2, \hat{W}_t^1, \hat{W}_t^2) \right\},$$
Problem statement: discussion

- Many reasonable choices for reward functions $\rho_t(\cdot)$, e.g.,

  $$\rho_t(W^1, W^2, \hat{W}^1_t, \hat{W}^2_t) = 1_{W^1=\hat{W}^1_t \text{ and } W^2=\hat{W}^2_t} \Rightarrow$$

  $$\mathbb{E}\rho_t(W^1, W^2, \hat{W}^1_t, \hat{W}^2_t) = Pr(W^1 = \hat{W}^1_t \text{ and } W^2 \neq \hat{W}^2_t)$$

- Focus on structural properties of the communication system that are common regardless of these choices.

Salient features of the problem:

1. Domain of encoding functions $X^i_t = f^i_t(W^i, Y^{t-1})$ increases with time.
2. Existence of common information at encoders ($Y^{t-1}$ at time $t$) and private information ($W^i$)
3. Decentralized, non-classical information structure (this is not a MDP/POMDP-like problem!)
Problem statement: discussion

- Many reasonable choices for reward functions $\rho_t(\cdot)$, e.g.,
  \[
  \rho_t(W^1, W^2, \hat{W}_t^1, \hat{W}_t^2) = 1_{W^1 = \hat{W}_t^1 \text{ and } W^2 = \hat{W}_t^2} \Rightarrow \\
  \mathbb{E}\rho_t(W^1, W^2, \hat{W}_t^1, \hat{W}_t^2) = Pr(W^1 = \hat{W}_t^1 \text{ and } W^2 \neq \hat{W}_t^2)
  \]

- Focus on **structural properties** of the communication system that are common regardless of these choices.

- **Salient features of the problem:**
  1. Domain of encoding functions $X_t^i = f_t^i(W^i, Y^{t-1})$ increases with time.
  2. Existence of common information at encoders ($Y^{t-1}$ at time $t$) and private information ($W^i$)
  3. Decentralized, non-classical information structure (this is **not** a MDP/POMDP-like problem!)
Introduction of pre-encoder

- Equivalent encoder description:
  Each user’s transmission $X^i_t = f^i_t(W^i, Y^{t-1})$ can be thought of as a two-stage process

  1. Based on available feedback $Y^{t-1}$ select encoding functions
     
     $$E^i_t : W^i \rightarrow X^i, \quad i = 1, 2,$$

     through a pre-encoder mapping
     
     $$(E^1_t, E^2_t) = h_t(Y^{t-1}).$$

  2. Generate transmitted signals by evaluating the encoding functions at $W^i$, i.e.,

     $$X^i_t = E^i_t(W^i), \quad i = 1, 2.$$
Introduction of pre-encoder

- Equivalent encoder description:
  Each user’s transmission $X^i_t = f^i_t(W^i, Y^{t-1})$ can be thought of as a two-stage process

  1. Based on available feedback $Y^{t-1}$ select encoding functions
     
     $$E^i_t : \mathcal{W}^i \rightarrow \mathcal{X}^i, \quad i = 1, 2,$$

     through a pre-encoder mapping

     $$(E^1_t, E^2_t) = h_t(Y^{t-1}).$$

  2. Generate transmitted signals by evaluating the encoding functions at $W^i$, i.e.,

     $$X^i_t = E^i_t(W^i), \quad i = 1, 2.$$
Introduction of pre-encoder

- Equivalent encoder description: Each user’s transmission $X_t^i = f_t^i(W^i, Y^{t-1})$ can be thought of as a two-stage process
  1. Based on available feedback $Y^{t-1}$ select encoding functions $E_t^i: W^i \rightarrow X^i$, $i = 1, 2$, through a pre-encoder mapping $(E_t^1, E_t^2) = h_t(Y^{t-1})$.
  2. Generate transmitted signals by evaluating the encoding functions at $W^i$, i.e., $X_t^i = E_t^i(W^i)$, $i = 1, 2$.

---

[Walrand and Varaiya, 1983, Nayyar and Teneketzis, 2008]
Decentralization of information is imposed by design ($h_t$ only uses the common information $Y^{t-1}$ available to both encoders)

Both encoders can evaluate each-other’s encoding functions through $(E^1_t, E^2_t) = h_t(Y^{t-1})$ (can be thought of as a fictitious coordinator)
Transforming to a centralized control problem

The control problem boils down to selecting encoding functions \((E^1_t, E^2_t) = h_t(Y^{t-1})\). Generation of \(X^i_t\) is a “dumb” function evaluation \(X^i_t = E^i_t(W^i)\).

New equivalent design \(g \overset{\text{def}}{=} \{f^1_t, f^2_t, d_t\}_t \Rightarrow \tilde{g} \overset{\text{def}}{=} \{h_t, d_t\}_t\).

Above transformation still suffers from increasing domain \(Y^{t-1}\) of the pre-encoder \(h_t\), i.e., \((E^1_t, E^2_t) = h_t(Y^{t-1})\).
Introduction of information state

- We would like to summarize $Y^{t-1}$ in a quantity (state) with time invariant domain.
- Related attempts:
  1. Introduction of auxiliary variables in information theory (e.g., [Cover and Leung, 1981, Bross and Lapidoth, 2005])
  2. Form a graph describing the correlation structure of the messages after receiving $Y^{t-1}$ [Venkataramanan and Pradhan, 2009]
- A more direct approach: define the random quantities

$$\Pi_t \in P(W^1 \times W^2), \quad t = 0, 1, 2, \ldots$$

as

$$\Pi_t(w^1, w^2) \overset{\text{def}}{=} Pr(W^1 = w^1, W^2 = w^2 | Y^t),$$

i.e., the posterior distribution of the message pair given the observation.
Introduction of information state

- We would like to summarize $Y^{t-1}$ in a quantity (state) with time invariant domain
- Related attempts:
  1. Introduction of auxiliary variables in information theory (e.g., [Cover and Leung, 1981, Bross and Lapidoth, 2005])
  2. Form a graph describing the correlation structure of the messages after receiving $Y^{t-1}$ [Venkataramanan and Pradhan, 2009]
- A more direct approach: define the random quantities

\[ \Pi_t \in \mathcal{P}(W^1 \times W^2), \quad t = 0, 1, 2, \ldots \]

as

\[ \Pi_t(w^1, w^2) \overset{\text{def}}{=} Pr(W^1 = w^1, W^2 = w^2 | Y^t), \]

i.e., the posterior distribution of the message pair given the observation.
Introduction of information state

- We would like to summarize $Y_{t-1}$ in a quantity (state) with time invariant domain.
- Related attempts:
  1. Introduction of auxiliary variables in information theory (e.g., [Cover and Leung, 1981, Bross and Lapidoth, 2005])
  2. Form a graph describing the correlation structure of the messages after receiving $Y_{t-1}$ [Venkataramananan and Pradhan, 2009]
- A more direct approach: define the random quantities

\[ \Pi_t \in \mathcal{P}(\mathcal{W}_1 \times \mathcal{W}_2), \quad t = 0, 1, 2, \ldots \]

as

\[ \Pi_t(w^1, w^2) \overset{\text{def}}{=} Pr(W^1 = w^1, W^2 = w^2 | Y^t), \]

i.e., the posterior distribution of the message pair given the observation.
Lemma

1. *The quantity* \( \Pi_t \) *can be recursively updated as*

\[
\Pi_t = \Phi(\Pi_{t-1}, E_1^t, E_2^t, Y_t), \quad t = 1, 2, \ldots
\]

2. \((\Pi_t)_t\) *is a controlled Markov process with control action* \((E_1^t, E_2^t)\)

3. *The optimal decoder function at time* \( t \) *is only a function of* \( \Pi_t \)

\[
(\hat{W}_t^1, \hat{W}_t^2) = d_t^*(Y_t) = d_t^*(\Pi_t)
\]

4. *The average instantaneous costs are functions of* \( \Pi_{t-1}, E_1^t, E_2^t \), *i.e.,*

\[
E\{\rho_t(W_1^1, W_2^2, \hat{W}_t^1, \hat{W}_t^2)\} = E\{\Psi_t(\Pi_{t-1}, E_1^t, E_2^t)\}.
\]

*where* \( \Psi_t \) *are known functions.*
Main structural result

Theorem

The optimal communication system for the NLF-MAC consists of

1. Encoders of the form \( X_t^i = E_t^i(W^i) \), \( i = 1, 2 \), where

\[
(E_t^1, E_t^2) = h_t(\Pi_{t-1})
\]

2. A receiver that generates message estimates as

\[
(\hat{W}_t^1, \hat{W}_t^2) = d_t(\Pi_t),
\]

where \( d_t \) is a known function.

3. The optimal \( h_t \) can be determined as the solution of a fix-point equation (dynamic program)
Multiple access channel with noiseless feedback

Equivalent optimal communication system

\[ \Phi(\cdot) \]

\[ h_t(\Pi_{t-1}) \]

\[ E_t^1(\cdot) \]

\[ E_t^2(\cdot) \]

\[ Y_{t-1} \]

\[ \Phi(\cdot) \]

\[ D \]

\[ \Pi_t \]

\[ d_t(\Pi_t) \]

\[ (\hat{W}_t^1, \hat{W}_t^2) \]

\[ W^1 \]

\[ W^2 \]
Overview

1. Point-to-point channels with memory and noiseless feedback

2. Multiple access channel with noiseless feedback

3. Cooperative communications in relay networks
Bernoulli arrivals at Source (w.p. $p$) and at Relay (w.p. $q$)

Packets waiting at Source’s and Relay’s queues $X_t = (X^1_t, X^2_t) \in \mathbb{N} \times \mathbb{N}$

Actions $U^1_t \in \mathcal{U}^1 \overset{\text{def}}{=} \{0, E_{13}, E_{12}\}$, $U^2_t \in \mathcal{U}^2 \overset{\text{def}}{=} \{0, E_{23}\}$

Simple collision model. Feedback $Y_t \in \{\emptyset, \text{ACK}, \text{NACK}\}$
Instantaneous costs are functions of energy and “delay”

\[ C_t^i = \rho^i(X_t^i, U_t^i) \] (e.g., \( = X_t^i + U_t^i \)), \( i = 1, 2 \)

Reasonable assumptions: \( E_{12} + E_{23} < E_{13}, p + q < 1 \), units either receive or transmit
Cooperative communications in relay networks

Three scenaria of interest: Scenario A

- **Centralized** control of queues with **perfect** observation

\[(U_t^1, U_t^2) = f_t(X_t^{1,t}, X_t^{2,t}, Y_t^{t-1})\]

- Find a set of policies \(f \overset{\text{def}}{=} \{f_t\}_t\) that minimize

\[J(f) = \mathbb{E}^f \left\{ \sum_t \rho^1(X_t^1, U_t^1) + \rho^2(X_t^2, U_t^2) \right\}\]

- **Solution**: Centralized stochastic control problem. Can be formulated as an MDP. \((U_t^1, U_t^2) = f^*_t(X_t^1, X_t^2)\)
Three scenaria of interest: Scenario A

- **Centralized** control of queues with **perfect** observation

\[(U^1_t, U^2_t) = f_t(X^{1,t}, X^{2,t}, Y^{t-1})\]

- Find a set of policies \(f = \{f_t\}_t\) that minimize

\[J(f) = \mathbb{E}^f \left\{ \sum_t \rho^1(X^1_t, U^1_t) + \rho^2(X^2_t, U^2_t) \right\}\]

- **Solution**: Centralized stochastic control problem. Can be formulated as an MDP. \((U^1_t, U^2_t) = f^*_t(X^1_t, X^2_t)\)
Three scenarios of interest: Scenario B

- Decentralized control of queues: each agent $i$ observes only his own queue length $X_i^t$ and both agents have a common goal (team problem)

  $$U_i^t = f_i^t(X_i^{i,t}, Y^{t-1}), \quad i = 1, 2$$

Find a set of policies $f \overset{\text{def}}{=} (f_1^t, f_2^t)_t$ that minimize

$$J(f) = \mathbb{E}^f \left\{ \sum_t \rho^1(X_1^t, U_1^t) + \rho^2(X_2^t, U_2^t) \right\},$$

- Salient features of the problem:
  1. Domain of control mappings $U_i^t = f_i^t(X_i^{i,t}, Y^{t-1})$ increases with time.
  2. Presence of common information ($Y^{t-1}$ at time $t$) and private information ($X_i^{i,t}$ at time $t$ for agent $i$)
  3. Decentralized, non-classical information structure (this is not a MDP/POMDP-like problem!)
Three scenarios of interest: Scenario B

- Decentralized control of queues: each agent $i$ observes only his own queue length $X_i^t$ and both agents have a common goal (team problem)

  $$U_i^t = f_i^t(X_i^{i,t}, Y^{t-1}), \quad i = 1, 2$$

  Find a set of policies $f \equiv (f^1_t, f^2_t)_t$ that minimize

  $$J(f) = \mathbb{E}^f \{ \sum_t \rho^1(X^1_t, U^1_t) + \rho^2(X^2_t, U^2_t) \}$$

- Salient features of the problem:
  1. Domain of control mappings $U_i^t = f_i^t(X_i^{i,t}, Y^{t-1})$ increases with time.
  2. Presence of common information ($Y^{t-1}$ at time $t$) and private information ($X_i^{i,t}$ at time $t$ for agent $i$)
  3. Decentralized, non-classical information structure (this is not a MDP/POMDP-like problem!)
Structural results for the team problem

Lemma ([Vasal and Anastasopoulos, 2012])

1. **Knowledge of** $Y^{t-1}$ **and** $U^{i,t-1}$ **reveals** $U^{t-1} = (U^{1,t-1}, U^{2,t-1})$, **so** $U^{t-1}$ **is common knowledge** ($Y^{t-1}$ **is not needed further)**

$$U^i_t = f^i_t(X^i_t, U^{t-1}), \quad i = 1, 2$$

2. **Optimal policy depends only on the current private state** $X^i_t$

$$U^i_t = f^i_t(X^i_t, U^{t-1}), \quad i = 1, 2$$

Still we have not addressed the decentralization issue and the expanding domain of $f^i_t$ issue.
Equivalent controller description:
Each agent’s decision \( U^i_t = f^i_t(X^i_t, U^{t-1}) \) can be thought of as a two-stage process

1. Based on common info \( U^{t-1} \) select “prescription” functions \( \Gamma^i_t : \mathbb{N} \rightarrow \mathcal{U}^i \), \( i = 1, 2 \) through the pre-encoder mapping

\[
(\Gamma^1_t, \Gamma^2_t) = h_t(U^{t-1})
\]

2. The actions \( U^i_t \) are determined by evaluating \( \Gamma^i_t \) at the private information \( X^i_t \), i.e.,

\[
U^i_t = \Gamma^i_t(X^i_t), \quad i = 1, 2
\]
Introduction of pre-encoder\(^6\)

Equivalent controller description:
Each agent’s decision \(U^i_t = f^i_t(X^i_t, U^{t-1}_t)\) can be thought of as a two-stage process

1. Based on common info \(U^{t-1}_t\) select “prescription” functions \(\Gamma^i_t : \mathbb{N} \rightarrow \mathcal{U}^i, \quad i = 1, 2\) through the pre-encoder mapping \((\Gamma^1_t, \Gamma^2_t) = h_t(U^{t-1}_t)\)

2. The actions \(U^i_t\) are determined by evaluating \(\Gamma^i_t\) at the private information \(X^i_t\), i.e.,

\[
U^i_t = \Gamma^i_t(X^i_t), \quad i = 1, 2
\]

\(^6\)[Walrand and Varaiya, 1983, Nayyar and Teneketzis, 2008]
Equation controller description:
Each agent’s decision $U_t^i = f_t^i(X_t^i, U_{t-1}^i)$ can be thought of as a two-stage process

1. Based on common info $U_{t-1}^i$ select “prescription” functions $\Gamma_t^i : \mathbb{N} \rightarrow \mathcal{U}^i$, $i = 1, 2$ through the pre-encoder mapping

$$(\Gamma_t^1, \Gamma_t^2) = h_t(U_{t-1}^i)$$

2. The actions $U_t^i$ are determined by evaluating $\Gamma_t^i$ at the private information $X_t^i$, i.e.,

$$U_t^i = \Gamma_t^i(X_t^i), \quad i = 1, 2$$

---

6 [Walrand and Varaiya, 1983, Nayyar and Teneketzis, 2008]
Transformation to a centralized problem

- Generation of $U^i_t$ is a “dumb” function evaluation $U^i_t = \Gamma^i_t(X^i_t)$
- The control problem boils down to selecting prescription functions $h \overset{\text{def}}{=} \{h_t\}_t$,
- Both agents can evaluate each-other’s prescription functions through $(\Gamma^1_t, \Gamma^2_t) = h_t(U^{t-1})$ (can be thought of as a fictitious controller)
- The decentralized control problem has been transformed to a centralized control problem
- Last issue to address: increasing domain $\mathcal{U}^{t-1}$ of the pre-encoder mappings $h_t$. 
Introduction of information state

- We would like to summarize $U^{t-1}$ in a quantity (state) with time invariant domain
- Define the random quantities $\Pi_t \in \mathcal{P}(\mathbb{N} \times \mathbb{N})$, $t = 0, 1, 2, \ldots$

$$\Pi_t(x^1_t, x^2_t) \overset{\text{def}}{=} P(X_t^1 = x_t^1, X_t^2 = x_t^2 | U^{t-1})$$

i.e., the posterior distribution of the queue lengths given the observation.
We would like to summarize $U^{t-1}$ in a quantity (state) with time invariant domain.

Define the random quantities

$$\Pi_t \in \mathcal{P}(\mathbb{N} \times \mathbb{N}), \quad t = 0, 1, 2, \ldots$$

as

$$\Pi_t(x_t^1, x_t^2) \overset{\text{def}}{=} P(X_t^1 = x_t^1, X_t^2 = x_t^2 | U^{t-1})$$

i.e., the posterior distribution of the queue lengths given the observation.

Achilleas Anastasopoulos
anastas@umich.edu
(U of Michigan)

Problems in the intersection of Information theory and Control

Dec 5, 2013 46 / 53
Main structural result

**Theorem ([Vasal and Anastasopoulos, 2012])**

The original decentralized control problem is equivalent to an MDP with

- **State** $\Pi_t$
- **Control actions** $\Gamma_t \overset{\text{def}}{=} (\Gamma^1_t, \Gamma^2_t)$
- **Instantaneous costs** $\mathbb{E}\{\rho^1_t(X^1_t, U^1_t) + \rho^2_t(X^2_t, U^2_t)|\Pi_t, \Gamma_t\}$

**Markov policies are optimal, i.e., optimal actions can be of the form**

$$\Gamma_t = h^*_t(\Pi_t) \quad \Rightarrow \quad U^i_t = f^i_t(\Pi_t, X^i_t)$$

It turns out there is a further simplification: instead of joint posterior distributions, we can use the two marginals! A general version of this result in [Nayyar et al., 2011]
Main structural result

Theorem ([Vasal and Anastasopoulos, 2012])

The original decentralized control problem is equivalent to an MDP with

- **State** \( \Pi_t \)
- **Control actions** \( \Gamma_t \overset{\text{def}}{=} (\Gamma^1_t, \Gamma^2_t) \)
- **Instantaneous costs** \( \mathbb{E}\{\rho^1_t(X^1_t, U^1_t) + \rho^2_t(X^2_t, U^2_t) | \Pi_t, \Gamma_t\} \)

Markov policies are optimal, i.e., optimal actions can be of the form

\[
\Gamma_t = h^*_t(\Pi_t) \quad \Rightarrow \quad U^i_t = f^{i*}_t(\Pi_t, X^i_t)
\]

It turns out there is a further simplification: instead of joint posterior distributions, we can use the two marginals! A general version of this result in [Nayyar et al., 2011]
Three scenarios of interest: Scenario C

- The Source/Relay act strategically: they want to minimize their own average costs over the given time horizon

\[ J^i(f) = \mathbb{E}^f \left\{ \sum_t \rho^i(X^i_t, U^i_t) \right\}, \quad i = 1, 2 \]

- Enlarge action space for Relay (to allow acceptance/rejection of Source packet)

\[ U^2_t \in \mathcal{U}^2 = \{0a, 0r, E_{23}\} \]

- One can study the resulting dynamic game and find Nash/sub-game perfect equilibria

- Unfortunately the equilibria of this game do not coincide with the optimal centralized solution of scenario A! (a.k.a., price of anarchy)

Example: if optimal centralized action was \((E_{12}, 0a)\) this can never be a NE, because Relay is better off playing \(0r\) (reject packet from source)
Three scenarios of interest: Scenario C

- The Source/Relay act **strategically**: they want to minimize their own average costs over the given time horizon
  \[
  J^i(f) = \mathbb{E}^f \left\{ \sum_t \rho^i(X^i_t, U^i_t) \right\}, \quad i = 1, 2
  \]

- Enlarge action space for Relay (to allow acceptance/rejection of Source packet)
  \[
  U^2_t \in \mathcal{U}^2 = \{0a, 0r, E_{23}\}
  \]

- One can study the resulting dynamic game and find Nash/sub-game perfect equilibria

- Unfortunately the equilibria of this game do not coincide with the optimal centralized solution of scenario A! (a.k.a., price of anarchy)
  Example: if optimal centralized action was \((E_{12}, 0a)\) this can never be a NE, because Relay is better off playing \(0r\) (reject packet from source)
Three scenaria of interest: Scenario C

- The Source/Relay act strategically: they want to minimize their own average costs over the given time horizon
  \[ J^i(f) = \mathbb{E}^f \left\{ \sum_t \rho^i(X^i_t, U^i_t) \right\}, \quad i = 1, 2 \]

- Enlarge action space for Relay (to allow acceptance/rejection of Source packet)
  \[ U^2_t \in \mathcal{U}^2 = \{0a, 0r, E_{23}\} \]

- One can study the resulting dynamic game and find Nash/sub-game perfect equilibria

- Unfortunately the equilibria of this game do not coincide with the optimal centralized solution of scenario A! (a.k.a., price of anarchy)
  Example: if optimal centralized action was \((E_{12}, 0a)\) this can never be a NE, because Relay is better off playing 0r (reject packet from source)
Three scenarios of interest: Scenario C

- The Source/Relay act **strategically**: they want to minimize their own average costs over the given time horizon

  \[ J^i(f) = E^f \left\{ \sum_t \rho^i(X^i_t, U^i_t) \right\}, \quad i = 1, 2 \]

- Enlarge action space for Relay (to allow acceptance/rejection of Source packet)

  \[ U^2_t \in \mathcal{U}^2 = \{0a, 0r, E_{23}\} \]

- One can study the resulting dynamic game and find Nash/sub-game perfect equilibria

- Unfortunately the equilibria of this game **do not coincide** with the optimal centralized solution of scenario A! (a.k.a., price of anarchy)

  Example: if optimal centralized action was \((E_{12}, 0a)\) this can never be a NE, because Relay is better off playing 0r (reject packet from source)
Incentivizing cooperation

- Our approach: Devise a protocol that provides incentives to Source/Relay so that the resulting dynamic game has equilibria that coincide with the solutions of the optimal centralized problem (Scenario A)

- Introduce a state/action-dependent monetary transfer $c(X_t, U_t)$ between agents

\[
\hat{\rho}^1(X_t, U_t) = \rho^1(X^1_t, U^1_t) + c(X_t, U_t)
\]

\[
\hat{\rho}^2(X_t, U_t) = \rho^2(X^2_t, U^2_t) - c(X_t, U_t)
\]

- Observe: the total societal cost is the same as in the centralized problem

\[
\hat{\rho}^1(X_t, U_t) + \hat{\rho}^2(X_t, U_t) = \rho^1(X^1_t, U^1_t) + \rho^2(X^2_t, U^2_t)
\]
Incentivizing cooperation

- Our approach: Devise a protocol that provides incentives to Source/Relay so that the resulting dynamic game has equilibria that coincide with the solutions of the optimal centralized problem (Scenario A)
- Introduce a state/action-dependent monetary transfer $c(X_t, U_t)$ between agents

$$
\hat{\rho}^1(X_t, U_t) = \rho^1(X_t^1, U_t^1) + c(X_t, U_t)
$$
$$
\hat{\rho}^2(X_t, U_t) = \rho^2(X_t^2, U_t^2) - c(X_t, U_t)
$$

- Observe: the total societal cost is the same as in the centralized problem

$$
\hat{\rho}^1(X_t, U_t) + \hat{\rho}^2(X_t, U_t) = \rho^1(X_t^1, U_t^1) + \rho^2(X_t^2, U_t^2)
$$
Incentivizing cooperation: basic result

- Important assumption: users know each others cost functions $\hat{\rho}^i$ (strategic behaviour does not manifest itself in desire for privacy/untruthful revelation of cost structure)

Theorem ([Vasal and Anastasopoulos, 2013])

There exist monetary transfers $c(\cdot, \cdot)$ such that the unique Nash (sub-game perfect) equilibrium of the resulting dynamic game is exactly the optimal solution of the centralized control problem.

- Implication: Source and Relay are incentivized to behave in a way that coincides with the optimal centralized solution.
Incentivizing cooperation: basic result

- Important assumption: users know each others cost functions $\hat{\rho}^i$ (strategic behaviour does not manifest itself in desire for privacy/untruthful revelation of cost structure)

**Theorem ([Vasal and Anastasopoulos, 2013])**

*There exist monetary transfers $c(\cdot, \cdot)$ such that the unique Nash (sub-game perfect) equilibrium of the resulting dynamic game is exactly the optimal solution of the centralized control problem*

- Implication: Source and Relay are incentivized to behave in a way that coincides with the optimal centralized solution
Incentivizing cooperation: basic result

- Important assumption: users know each others cost functions $\hat{\rho}^i$ (strategic behaviour does not manifest itself in desire for privacy/untruthful revelation of cost structure)

Theorem ([Vasal and Anastasopoulos, 2013])

*There exist monetary transfers $c(\cdot, \cdot)$ such that the unique Nash (sub-game perfect) equilibrium of the resulting dynamic game is exactly the optimal solution of the centralized control problem*

- Implication: Source and Relay are incentivized to to behave in a way that coincides with the optimal centralized solution
Conclusions

- A number of communications problems can be viewed as centralized/decentralized control systems.
- Using ideas from Control we can derive structural results and simplify the solution of these problems.
- Can handle: dynamics; cooperation (team problems); and to some extent competition (games).
- Still a lot of open problems in this area.
  - Capacity-achieving / Error exponent-achieving actual communication systems (single/multi-user).
  - Single-letter capacity for MAC with feedback...
Open problem: Capacity of Multiple Access Channel with noiseless feedback

- Messages $W^i \in \{1, 2, \ldots, 2^{nR^i}\}$, $i = 1, 2$
- Transmitted symbols $X_t^i \in \mathcal{X}^i$, $i = 1, 2$, $t = 1, 2, \ldots, n$
- Received symbols $Y_t \in \mathcal{Y}$, $t = 1, 2, \ldots, n$
- Discrete-memoryless MAC (DM-MAC) $Q(y_t|X_t^1, X_t^2)$
- Message estimates $(\hat{W}^1, \hat{W}^2) \in \mathcal{W}^1 \times \mathcal{W}^2$
- Encoding functions $X_t^i = f_t^i(W^i, Y_t^{t-1})$, $i = 1, 2$, $t = 1, 2, \ldots, n$
- Decoding function $(\hat{W}^1, \hat{W}^2) = d(Y^n)$
Thank you!


An achievable rate region for the multiple-access channel with feedback.


---

The capacity region of a multiple-access discrete memoryless channel can increase with feedback (corresp.).


---

Sequential transmission using noiseless feedback.


---

Upper bounds on the capacities of non-controllable finite-state channels using dynamic programming methods.


---

Capacity results for the discrete memoryless network.


Optimal design of sequential real-time communication systems.

(accepted for publication).

Optimal control strategies in delayed sharing information structures.


On globally optimal real-time encoding and decoding strategies in multi-terminal communication systems.

In *Proc. IEEE Conf. on Decision and Control*, pages 1620–1627, Cancun, Mexico.

Cooperative communications in relay networks

The capacity of the white Gaussian multiple access channel with feedback.


Cardinality bounds on auxiliary variables in multiple-user theory via the method of ahlswede and korner.

A coding scheme for additive noise channels with feedback–I: No bandwidth constraint.

The posterior matching feedback scheme: Capacity achieving and error analysis.


