Optimal Real-time communication

Aditya Mahajan

DEPT. OF EECS,
UNIVERSITY OF MICHIGAN,
ANN ARBOR, MI. USA.

Joint work with *Demosthenis Teneketzis*

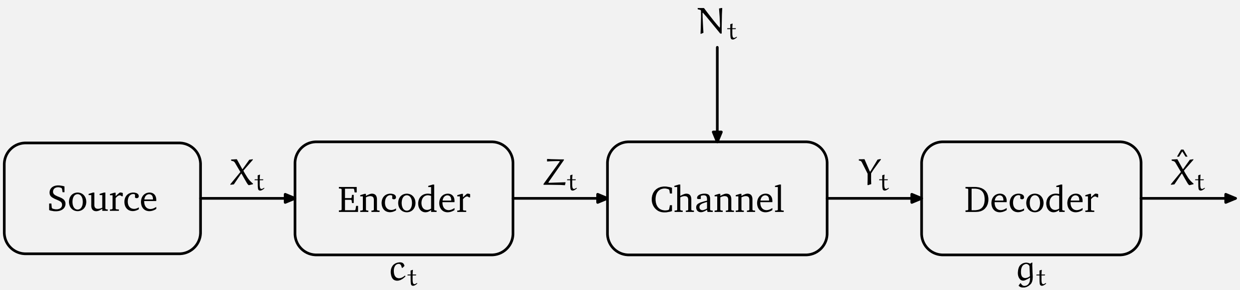
Presented at Information Theory and Applications Workshop, January 30, 2008.

Motivation

Real-Time Communication Communication systems in which information should be transmitted and decoded within a fixed delay constraint.

- *Applications*
  - ▷ Sensor networks
  - ▷ QoS over communication networks
  - ▷ Vehicular traffic control
  - ▷ Surveillance networks
  - ▷ Networked controlled systems
- *Features*
  - ▷ Informationally decentralized systems
  - ▷ Communication is delay sensitive
  - ▷ Channels may be noisy

Model



Encoder $Z_t = c_t(X_1, \dots, X_t)$

Decoder $\hat{X}_t = g_t(Y_1, \dots, Y_t)$

Delay δ

Distortion $\rho(X_{t-\delta}, \hat{X}_t)$

Total Cost $\mathbf{E}\left\{ \sum_{t=\delta+1}^{T} \rho(X_{t-\delta}, \hat{X}_t) \,\middle|\, c_1, \dots, c_T, g_1, \dots, g_T \right\}$

Model

Salient Features

- *Sequential operation*

$$\cdots \to \text{encoder at } t \to \text{decoder at } t \to \text{encoder at } t+1 \to \text{decoder at } t+1 \to \cdots$$

- *Decentralized information*

$$\underset{\text{info. at encoder}}{\sigma(X_1, X_2, \dots, X_t)} \quad \begin{matrix} \not\subseteq \\ \not\supseteq \end{matrix} \quad \underset{\text{info. at decoder}}{\sigma(Y_1, Y_2, \dots, Y_t)}$$

Non-classical information structure

Comparison with Info Thy

	Shannon Formulation		Real-time communication
Encoder	$Z_t = c_t(X_1, \ldots X_T)$		$Z_t = c_t(X_1, \ldots X_t)$
Decoder	$\hat{X}_t = g_t(Y_1, \ldots Y_T)$		$\hat{X}_t = g_t(Y_1, \ldots Y_t)$
Distortion	$\rho(X^T, \hat{X}^T) = \sum_{t=1}^{T} \rho(X_t, \hat{X}_t)$		$\rho(X^T, \hat{X}^T) = \sum_{t=\delta+1}^{T} \rho(X_{t-\delta}, \hat{X}_t)$

- *Information theoretic results are not applicable*
  - Cannot use asymptotic equipartition theorem.
  - No concentration of measure on typical sequences.
  - Separate source channel coding is not optimal
- *Asymptotic concepts not appropriate*
  - Source entropy
  - Transmission rate
  - Channel capacity

Our Approach

Formulate the real-time communication problem as a stochastic optimization problem

Stochastic Opt --- MDP

- *Markov decision theory*
  - Classical methodology for solving stochastic optimization problems
- *Assumption: Centralized system*
  - One controller
  - Perfect recall at the controller
- *Real-time communication*
  - Has two “controllers”
  - Decentralized information

Markov decision theory is not applicable to real-time communication problem

Real-time communication

Conceptual difficulties with real-time comm

- Information theory does not apply
- Markov decision theory does not apply
- Brute force search is computationally very difficult

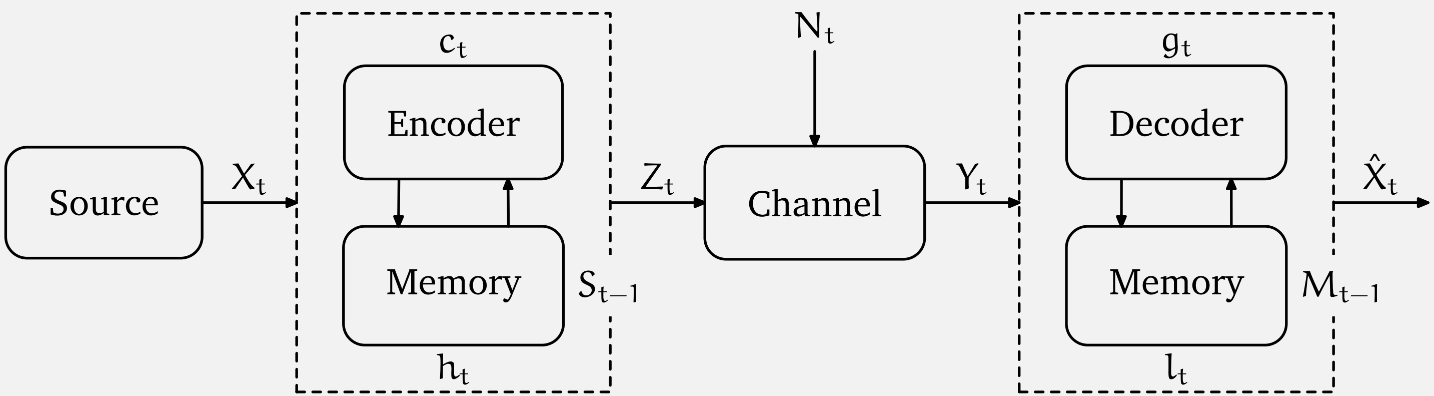
Our Contribution

- *Provide sequential decomposition*
  - Break one shot optimization problem into sequence of nested optimization problems
  - Exponential reduction in the search complexity

$$O\big((2^{A})^{T}\big) \to O(T \cdot K \cdot 2^{A})$$

Solution Methodology: An Example

Example



Encoder

$$Z_t = c_t(X_t, S_{t-1})$$

$$S_t = h_t(X_t, S_{t-1})$$

Decoder

$$\hat{X}_t = g_t(Y_t, M_{t-1})$$

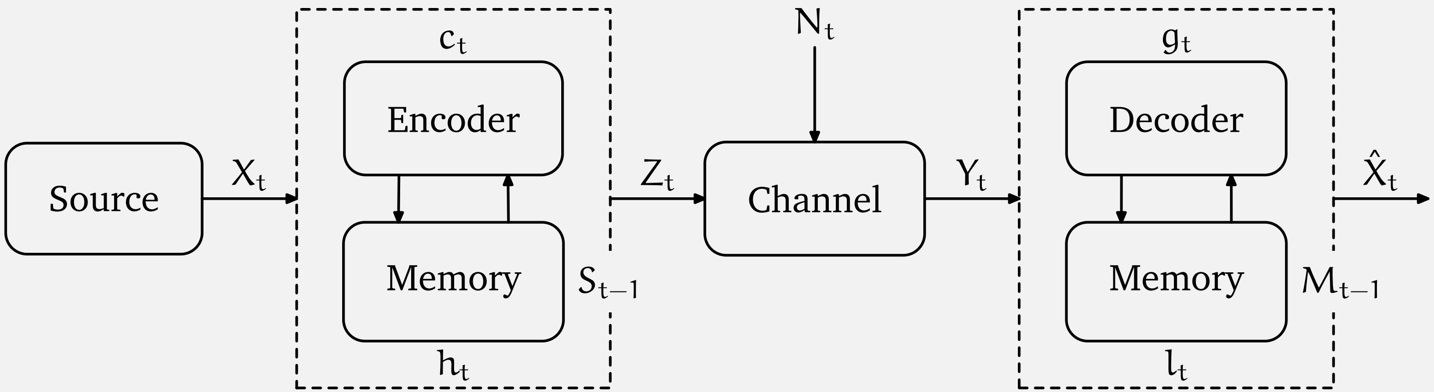
$$M_t = l_t(Y_t, M_{t-1})$$

$$T = 10$$

$$\delta = 1$$

Distortion $\rho(X_{t-1}, \hat{X}_t) =$ hamming distortion

Example



Communication Scheme

$$C := (c_1, \dots, c_{10}) \qquad H := (h_1, \dots, h_{10})$$

$$G := (g_1, \dots, g_{10}) \qquad L := (l_1, \dots, l_{10})$$

Performance

$$\mathcal{J}(C, H, G, L) := \mathbf{E}\left\{ \sum_{t=2}^{10} \rho(X_{t-1}, \hat{X}_t) \,\middle|\, C, H, G, L \right\}$$

Brute Force Approach

- Fix a communication scheme (C, H, G, L)
- Evaluate $\Pr\left(X_1, \dots, X_{10}, \hat{X}_1, \dots, \hat{X}_{10} \,\middle|\, C, H, G, L\right)$
- Evaluate $\mathbf{E}\left\{\sum_{t=2}^{10} \rho(X_{t-1}, \hat{X}_t) \,\middle|\, C, H, G, L\right\}$
- Repeat for all choices of (C, H, G, L)
- Pick the scheme with best performance

Complexity

- Possible choices for $c_t = 2^{2\times 2} = 16.$
- Possible choices for $(c_t, h_t, g_t, l_t) = 16^4 = 65,536$
- Possible choices for $(C, H, G, L) = (16^4)^{10} \approx 1.5 \times 10^{48}$

Recall, this is for a "simple" example.

Solution Approach: Sequential Decomposition

Key Idea: Information state

Information State

Requirements on information state (π_t)

- *π_t should be a "state"*
  - $\pi_t = \text{fn}(c_{t-1}^{T}, h_{t-1}^{T}, g_{t-1}^{T}, l_{t-1}^{T})$
  - $\pi_{t+1} = \text{fn}(\pi_t, c_t, h_t, g_t, l_t)$
- *π_t should absorb the effect of past functions on future performance*
  - $\mathbf{E}\left\{\rho(X_{t-1}, \hat{X}_t) \,\middle|\, c_1^t, h_1^t, g_1^t, l_1^t\right\} = \mathbf{E}\left\{\rho(X_{t-1}, \hat{X}_t) \,\middle|\, \pi_t, c_t, h_t, g_t, l_t\right\}$

Sequential Decomposition

$$V_t(\pi_t) = \min_{\gamma_t}\left\{\hat{\rho}(\pi_t, \gamma_t) + V_{t+1}\big(\pi_{t+1}(\pi_t, \gamma_t)\big)\right\}$$

$$\mathcal{J}^* = V_1(\pi_1)$$

where $\gamma_t = (c_t, h_t, g_t, l_t)$.

Identifying appropriate information states is highly non-trivial

Seq. Decomposition

Information State

$$\pi_t = \Pr\left(X_t, S_{t-1}, Y_{t-1}, M_{t-1}\right)$$

$$\pi_t \in \Delta(\mathcal{X} \times \mathcal{S} \times \mathcal{Y} \times \mathcal{M})$$

$$\begin{tikzcd}
{} \arrow[r] & \pi_t \arrow[r, "{(c_t, h_t, l_t)}"] \arrow[d, "{(c_t, g_t)}"] & \pi_{t+1} \\
& \mathbf{E}\left\{\rho(X_{t-1}, \hat{X}_t) \,\middle|\, \pi_t, c_t, g_t\right\} &
\end{tikzcd}$$

Refine time

$$^1\pi_t = \Pr\left(X_t, S_{t-1}, Y_{t-1}, M_{t-1}\right)$$

$$^2\pi_t = \Pr\left(X_t, S_{t-1}, Y_t, M_{t-1}\right)$$

$$^3\pi_t = \Pr\left(X_t, S_t, Y_t, M_{t-1}\right)$$

$$\begin{tikzcd}
^1\pi_t \arrow[r, "c_t"] & ^2\pi_t \arrow[r, "h_t"] & ^3\pi_t \arrow[r, "l_t"] \arrow[d, "g_t"] & ^1\pi_{t+1} \\
& & \mathbf{E}\left\{\rho(X_{t-1}, \hat{X}_t) \,\middle|\, {}^3\pi_t, g_t\right\} &
\end{tikzcd}$$

Seq. Decomposition

(Initialization) $\quad {}^3V_{T+1}({}^3\pi) = 0$

(Recursion)

$$\begin{cases} {}^1V_t({}^1\pi_t) = \min_{c_t} {}^2V_t\big({}^2\pi_t({}^1\pi_t, c_t)\big) \\ {}^2V_t({}^2\pi_t) = \min_{h_t} {}^3V_t\big({}^3\pi_t({}^2\pi_t, h_t)\big) \\ {}^3V_t({}^3\pi_t) = \min_{g_t} \hat{\rho}({}^3\pi_t, g_t) + \min_{l_t} {}^1V_{t+1}\big({}^4\pi_{t+1}({}^3\pi_t, l_t)\big) \end{cases}$$

- *Functional optimization problem*
  - Different from Markov decision theory
- *Two step solution*
  - Step One: Computations — The backward step (off-line)
  - Step Two: Implementation — The forward step (off-line or on-line)

Seq. Decomposition

- *Computations — The backward step*
  - For each time instant t and each ${}^i\pi_t \in \Pi$
    - evaluate the cost to go ${}^iV_t({}^i\pi_t)$
    - and store the corresponding arg minimum ${}^i\Phi_t({}^i\pi_t)$
  - $\mathcal{J}^* = V_1(\pi_1^\circ)$
- *Implementation — The forward step*
  - Start at time 1. We know ${}^1\pi_1^\circ$. Look-up $c_1^\circ = {}^1\Phi_t({}^1\pi_1^\circ)$.
  - ${}^1\pi_1^\circ$ and c_1° determine ${}^2\pi_1^\circ$. Look-up $h_1^\circ = {}^2\Phi_t({}^2\pi_2^\circ)$.
  - And so on . . .

Determine optimal $(c_1^\circ, h_1^\circ, \dots, g_T^\circ, l_T^\circ)$ *off line*

Complexity

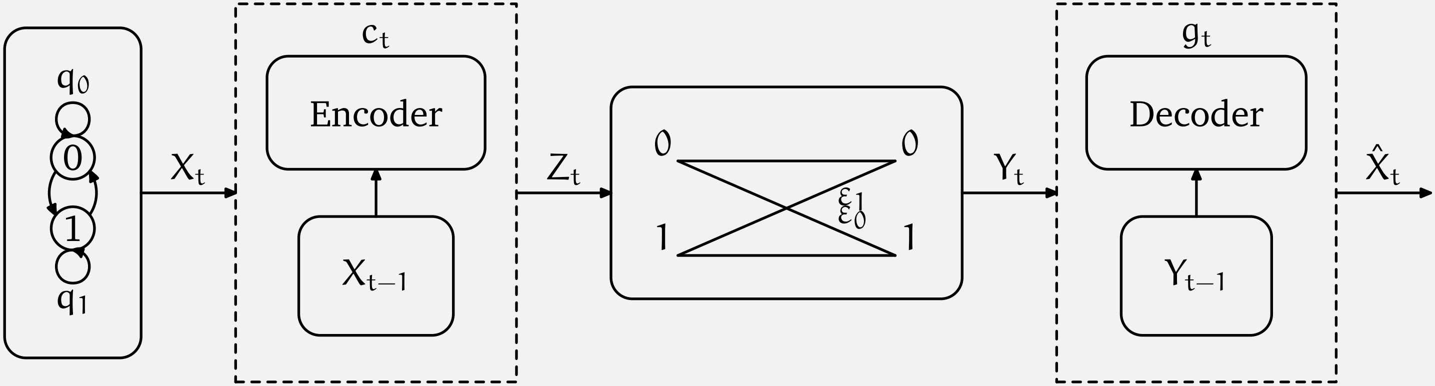
- Each equation is non-convex in function space.
- For a fixed t and π_t, there are 2^4 alternatives.
- There are $3 \times T = 30$ nested optimality equations. (linear in T)
- However, π takes value in a continuous space.
- Suppose we partition Π into 10^6 points.

$$\text{Number of calculations} = 5 \times 10^8 \qquad (\text{cf. } 10^{48} \text{ for brute force})$$

Numerical Computations

- *Reachability Analysis*
  - Find all reachable π_t and solve the nested optimality equations for them
- *Smallwood & Sondik-like Algorithm*
  - ${}^{i}V_t(\cdot)$ is piecewise linear and convex
  - Can be represented as pointwise minimum of a finite family of affine functions
  - These affine functions can be computed by linear programming
- *Approximation Algorithms*
  - Grid based solutions
  - Rust's probabilistic algorithm
- *Specialized Algorithms ??*

Numerical Example



Source $(0.7, 0.1)$

Channel $(0.2, 0.1)$

Encoder

X_t	X_{t-1}	Z_t
0	0	1
0	1	0
1	0	1
1	1	0

Decoder

Y_t	Y_{t-1}	$\hat{X}_t$
0	0	1
0	1	1
1	0	0
1	1	0

Numerical Example

Source $(0.7, 0.1)$ Channel $(0.2, 0.2)$

Encoder 1

X_t	X_{t-1}	Z_t
0	0	1
0	1	0
1	0	0
1	1	0

Decoder 1

Y_t	Y_{t-1}	$\hat{X}_t$
0	0	1
0	1	0
1	0	0
1	1	0

Encoder 2

X_t	X_{t-1}	Z_t
0	0	1
0	1	1
1	0	0
1	1	0

Decoder 2

Y_t	Y_{t-1}	$\hat{X}_t$
0	0	1
0	1	0
1	0	0
1	1	0

Encoder 3

X_t	X_{t-1}	Z_t
0	0	1
0	1	0
1	0	1
1	1	0

Decoder 3

Y_t	Y_{t-1}	$\hat{X}_t$
0	0	0
0	1	0
1	0	0
1	1	0

$$\text{Encoder 1} \longrightarrow \text{Encoder 2} \longrightarrow \text{Encoder 3} \longrightarrow \text{Encoder 1} \longrightarrow \cdots$$

Summary

So far . . .

- Formulated real-time communication as a stochastic optimization problem.
- Obtained a methodology for sequential decomposition

Key Ideas

- Information state
- Information structures

The key ideas are fundamental and are also applicable to other problems

- *Three different explanations of how to choose information states*
  - Related to Aumann's notion of common knowledge
  - Resolve the second guessing argument

Other problems

- *Variations of real-time communication problem*
  - Arbitrary (but finite) delay
  - Channels with memory
  - Active noisy feedback
- *Control and communication*
  - Optimal feedback control over noisy communication channels.
- *Decentralized diagnosis with communication*
  - Fault diagnosis in discrete event systems with communication between diagnosers.

Future Directions

- *Connections with classical information theory*
  - Smooth transition from real-time to asymptotic
  - Sequential information theory problems as stochastic optimization problem
- *Connections with other approaches to real-time communication*

  (e.g. Linder & Lugosi, Matloub & Weissman)
- *Connections with mathematical economics*
  - Mechanism design
  - Games with communication
- *Decentralized systems with a communication component*
  - Networks: Communication, control, and detection

References

- Aditya Mahajan and Demosthenis Teneketzis, On jointly optimal encoding, decoding, and memory update for noisy real-time communication systems, under review in *IEEE Transactions on Information Theory*.
- ——, On the design of globally optimal communication strategies for noisy real-time communication systems with noisy feedback, accepted for publication in the *IEEE Journal on Selected Areas in Communication*, Decemeber 2007.
- ——, Globally optimal performance of feedback control systems with limited communication over noisy channels, under review in *SIAM Journal of Control and Optimization*.
- ——, Optimal design of communication strategies for decentralized diagnosis of discrete event systems with communicating diagnosers, in preparation.

Thank you