Decentralized decision making under uncertainty

Aditya Mahajan
Yale University

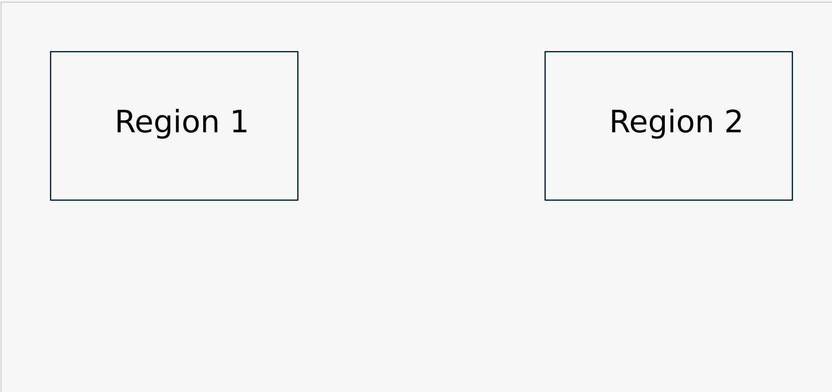
Acknowledgments

- Mentors
  - Prof. Demos Teneketzis (Univ of Michigan)
  - Prof. Sekhar Tatikonda (Yale Univ)
- Collaborators
  - Prof. Serdar Yüksel (Queen's Univ),
  - Prof. Mingyan Liu (Univ of Michigan),
  - Prof. Mahta Moghaddam (Univ of Michigan),
  - Prof. Dara Entekhabi (MIT),
  - Ashutosh Nayyar, David Shuman (Univ of Michigan)
- Funding Agencies
  - EECS departmental and Rackham graduate school fellowships
  - NSF, ONR, NASA

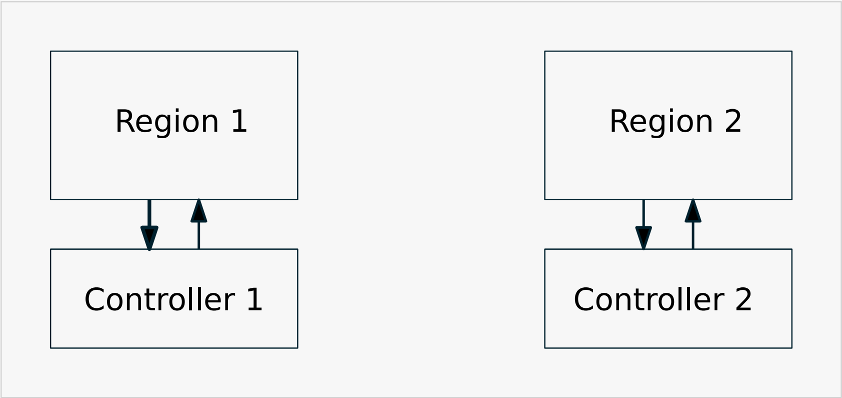

Decentralized systems
are ubiquitous

Interconnected power systems

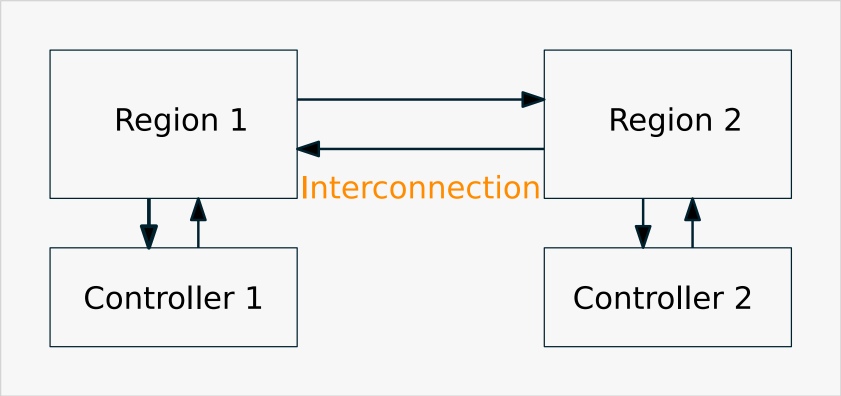
Interconnected power systems



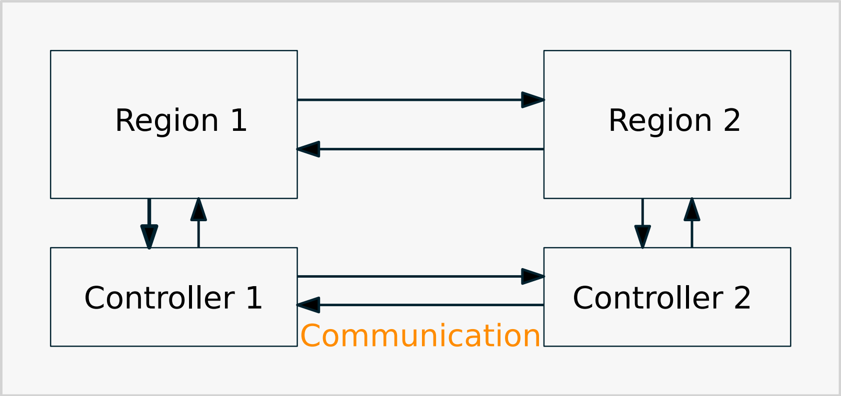
Interconnected power systems



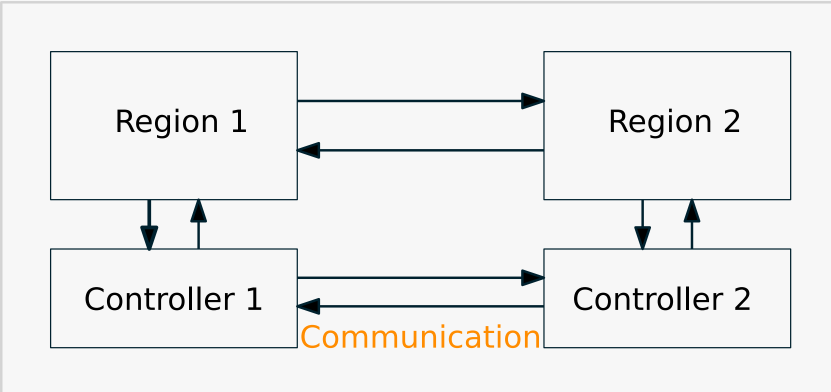
Interconnected power systems



Interconnected power systems

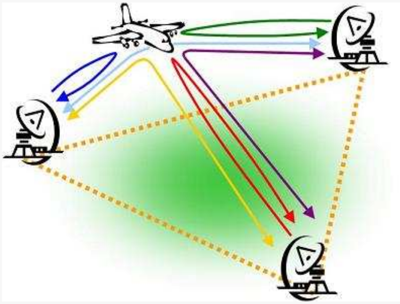


Interconnected power systems

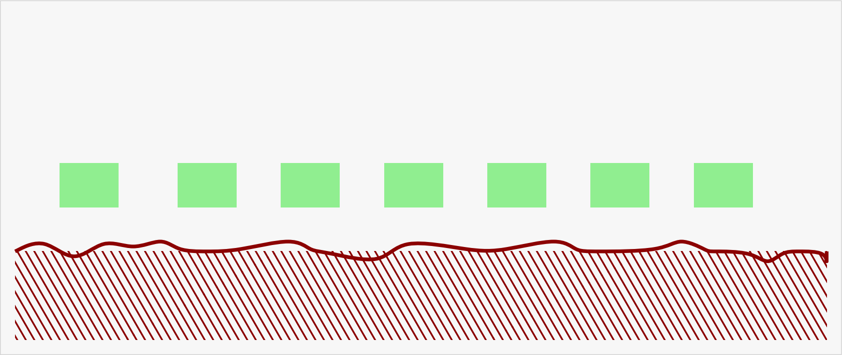


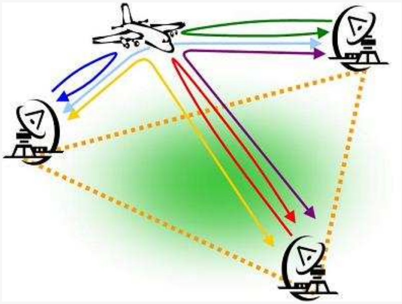
- Challenges
  - How to coordinate
  - What to communicate
  - How to communicate
  - When to communicate

Sensor and Surveillance networks



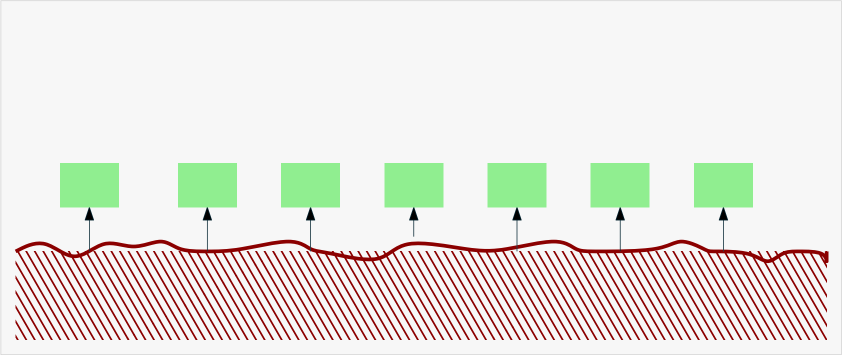
Sensor and Surveillance networks

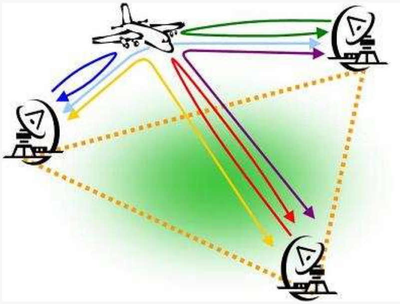




- Limited resources

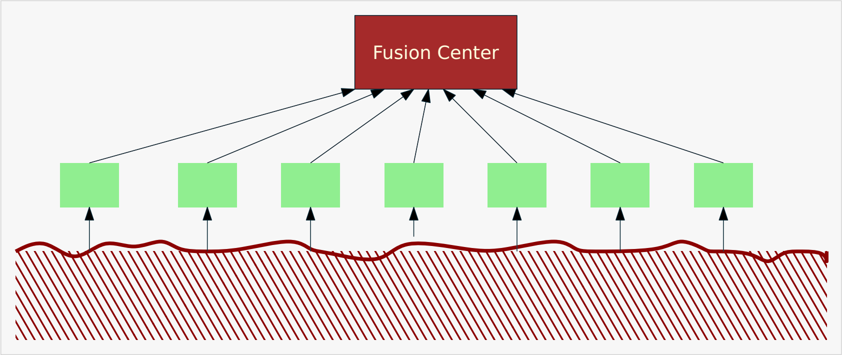
Sensor and Surveillance networks



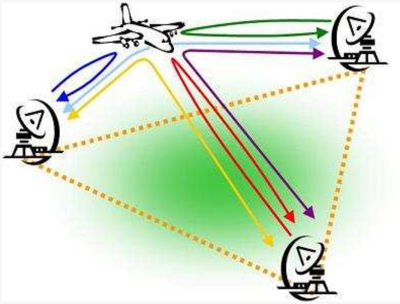


- Limited resources
- Noisy observations

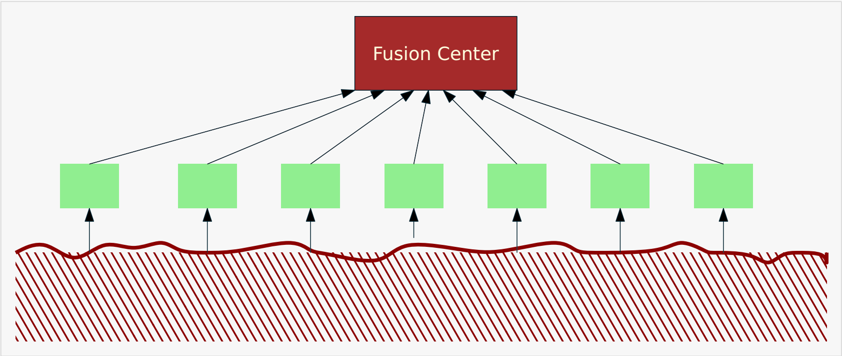
Sensor and Surveillance networks

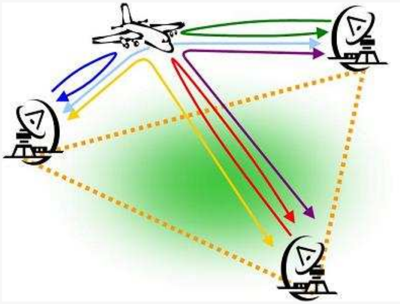


- Limited resources
- Noisy observations
- Communication



Sensor and Surveillance networks



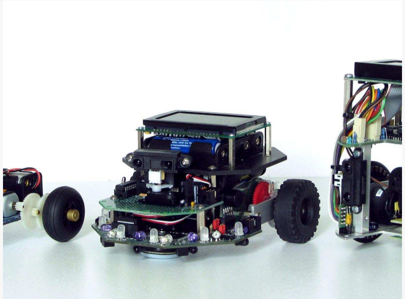


- Limited resources
- Noisy observations
- Communication

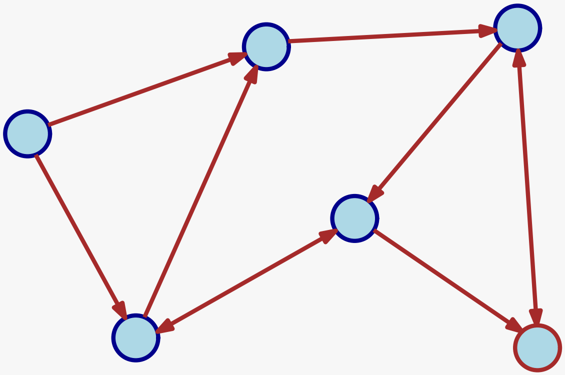
Challenges

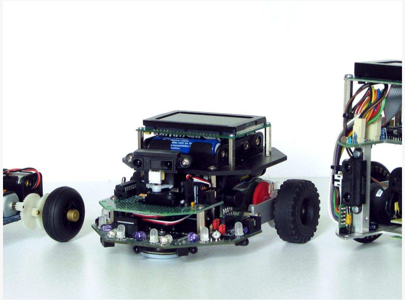
- Real-time communication
- Scheduling measurements and communication
- Detect node failures

Networked control systems

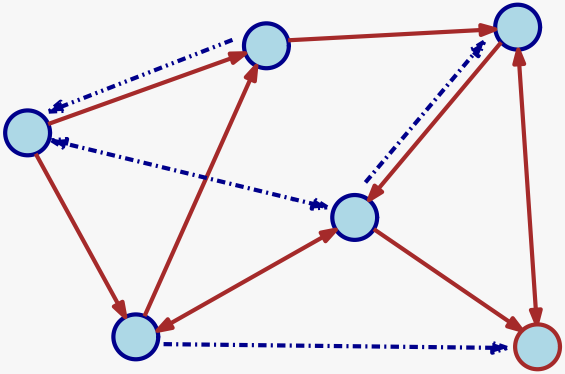


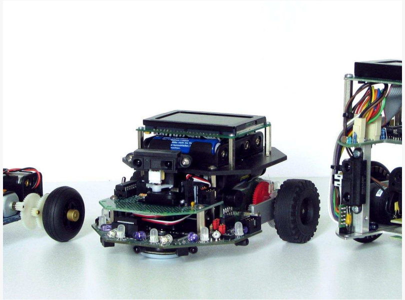
Networked control systems



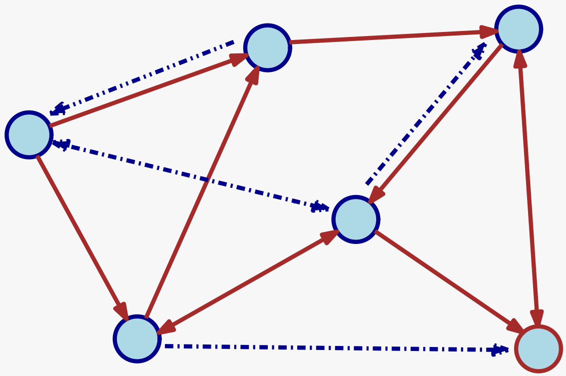


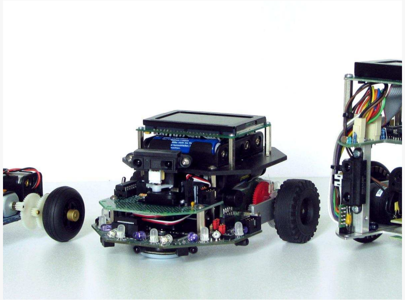
Networked control systems





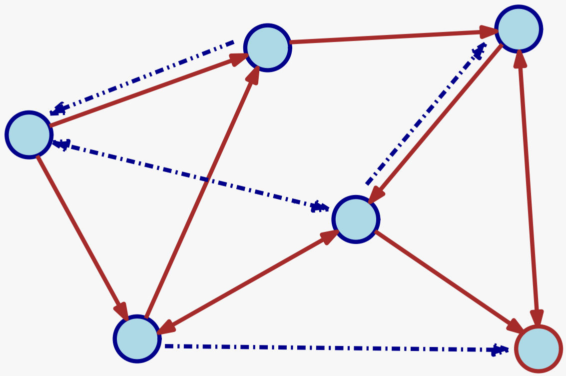
Networked control systems

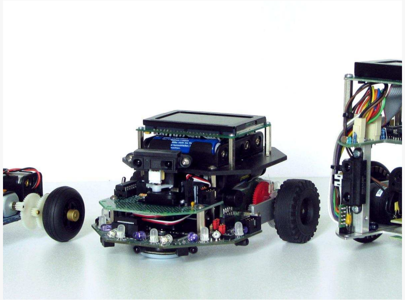




- Challenges
  - Control and communication over networks (internet $\Rightarrow$ delay, wireless $\Longrightarrow$ losses)

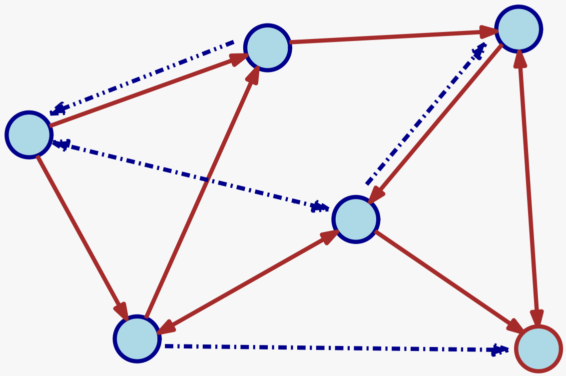
Networked control systems

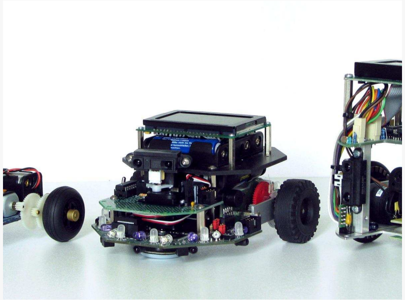




- Challenges
  - Control and communication over networks (internet ⇒ delay, wireless ⟹ losses)
  - Distributed estimation

Networked control systems





- Challenges
  - Control and communication over networks (internet $\Rightarrow$ delay, wireless $\Longrightarrow$ losses)
  - Distributed estimation
  - Distributed learning

Salient Features

Salient Features

- Multiple agents

  Decision making by multiple agents in stochastic dynamic environment

Salient Features

- Multiple agents

  Decision making by multiple agents in stochastic dynamic environment

- Coordination issues

  All agents must coordinate to achieve a system-wide objective

Salient Features

- Multiple agents

  Decision making by multiple agents in stochastic dynamic environment

- Coordination issues

  All agents must coordinate to achieve a system-wide objective

- Communication constraints

  Data must be communicated within fixed finite delay

Salient Features

- Multiple agents

  Decision making by multiple agents in stochastic dynamic environment

- Coordination issues

  All agents must coordinate to achieve a system-wide objective

- Communication constraints

  Data must be communicated within fixed finite delay

- Robustness

  System model may not be known completely

Salient Features

- Multiple agents

  Decision making by multiple agents in stochastic dynamic environment

- Coordination issues

  All agents must coordinate to achieve a system-wide objective

- Communication constraints

  Data must be communicated within fixed finite delay

- Robustness

  System model may not be known completely

- Exploiting domain knowledge

  Application specific modeling assumptions

Research Directions

Research Directions

- Real-time communication

Delay sensitive communication

Research Directions

- Real-time communication

  Delay sensitive communication

- Optimal control over noisy channels

  Communication and coordination

Research Directions

- Real-time communication

  Delay sensitive communication

- Optimal control over noisy channels

  Communication and coordination

- Delay sharing pattterns

  Coordination

Research Directions

- Real-time communication

  Delay sensitive communication

- Optimal control over noisy channels

  Communication and coordination

- Delay sharing pattterns

  Coordination

- Communication over unknown channels

  Robustness

Research Directions

- Real-time communication

  Delay sensitive communication

- Optimal control over noisy channels

  Communication and coordination

- Delay sharing pattterns

  Coordination

- Communication over unknown channels

  Robustness

- Calibration and validation of remote sensing observations

  Exploiting domain knowledge

Research Directions

- **Real-time communication**

  Communication constraint

- **Optimal control over noisy channels**

  Coordination

- **Delayed sharing patterns**

  Coordination

- Communication over unknown channels

  Robust communication

- Calibration and validation of remote sensing observations

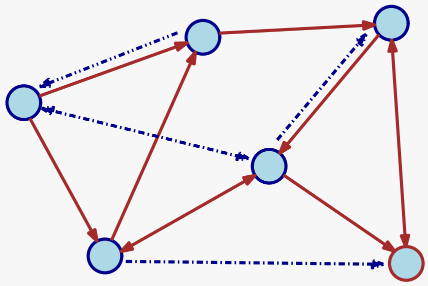
  Exploiting domain knowledge

Real-time communication

Real-time communication

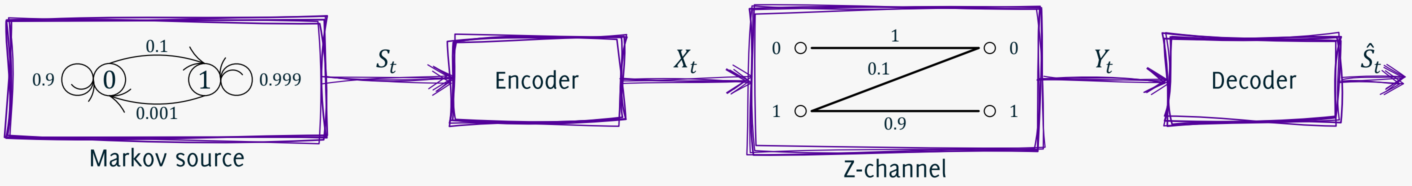
Simplest setup: A node observes a stream of data and has to communicate it to another node (over possibly noisy channels) within a fixed finite delay

- Integral component of many applications
  - Sensor and surveillance networks
  - Transportation networks
  - Fault diagnosis in power systems
  - Networked control systems

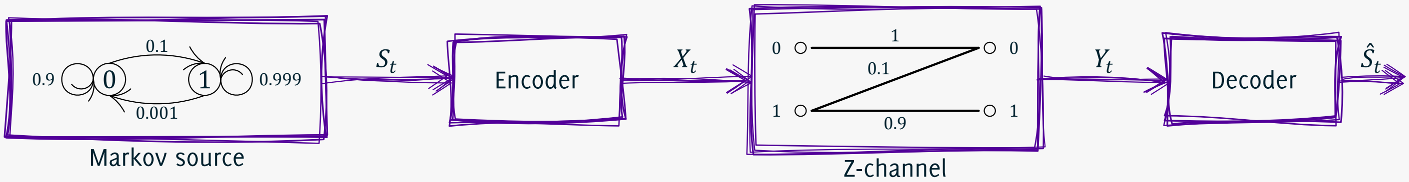


Importance and Challenges: An example

Importance and Challenges: An example



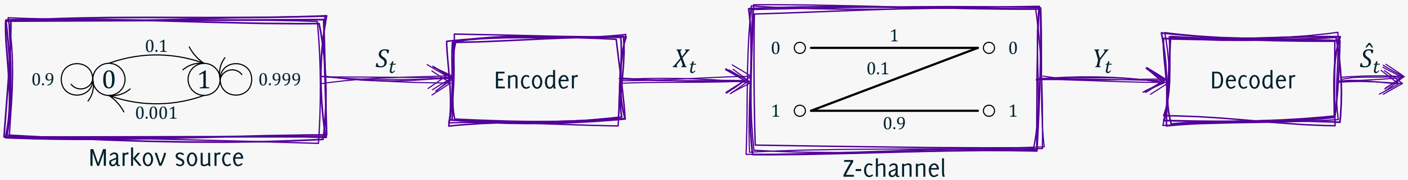
Importance and Challenges: An example



- Probability of error

$$\mathbb{P}(s_1 \neq \hat{s}_1) + \mathbb{P}(s_2 \neq \hat{s}_2) + \mathbb{P}(s_3 \neq \hat{s}_3)$$

Importance and Challenges: An example



- Probability of error

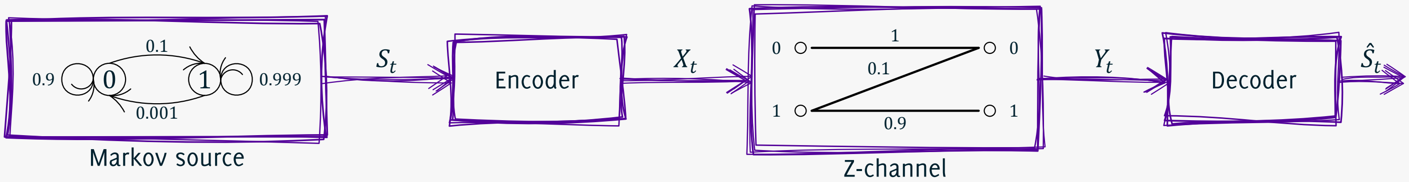
$$\mathbb{P}(s_1 \neq \hat{s}_1) + \mathbb{P}(s_2 \neq \hat{s}_2) + \mathbb{P}(s_3 \neq \hat{s}_3)$$

- Memoryless coding strategies

$$x_1 = e_1(s_1), \quad x_2 = e_2(s_2), \quad x_3 = e_3(s_3)$$

$$\hat{s}_1 = e_1(y_1), \quad \hat{s}_2 = e_2(y_2), \quad \hat{s}_3 = e_3(y_3)$$

Importance and Challenges: An example



- Probability of error

$$\mathbb{P}(s_1 \neq \hat{s}_1) + \mathbb{P}(s_2 \neq \hat{s}_2) + \mathbb{P}(s_3 \neq \hat{s}_3)$$

- Memoryless coding strategies

$$x_1 = e_1(s_1), \quad x_2 = e_2(s_2), \quad x_3 = e_3(s_3)$$

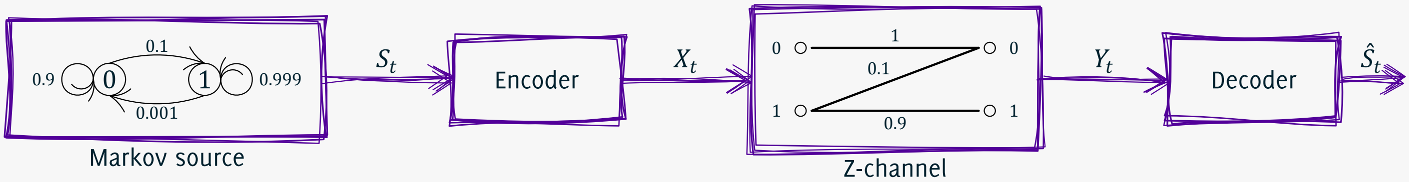
$$\hat{s}_1 = e_1(y_1), \quad \hat{s}_2 = e_2(y_2), \quad \hat{s}_3 = e_3(y_3)$$

- Causal real-time coding strategies

$$x_1 = e_1(s_1), \quad x_2 = e_2(s_1, s_2), \quad x_3 = e_3(s_1, s_2, s_3)$$

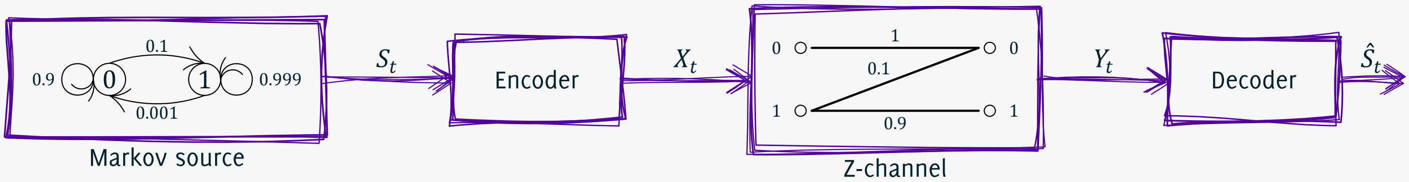
$$\hat{s}_1 = e_1(y_1), \quad \hat{s}_2 = e_2(y_1, y_2), \quad \hat{s}_3 = e_3(y_1, y_2, y_3)$$

Importance and Challenges: An example



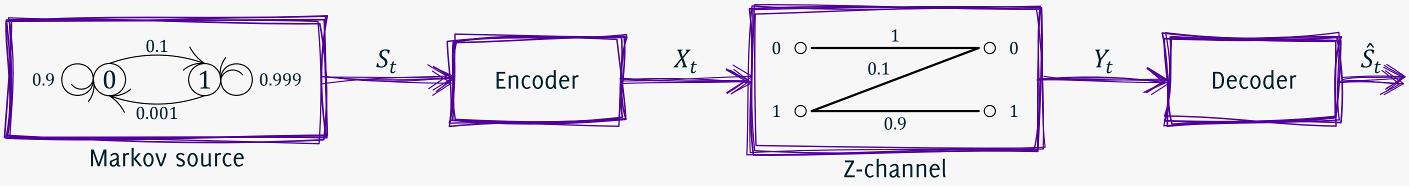
	Memoryless	Causal real-time	
Prob of error	0.1346	0.0564	240% better

Importance and Challenges: An example



	Memoryless	Causal real-time	
Prob of error	0.1346	0.0564	240% better
# of strategies	$(2^2 \times 2^2 \times 2^2)^2$	$(2^2 \times 2^4 \times 2^8)^2$	simple example
	$O(10^3)$	$O(10^8)$	

Importance and Challenges: An example



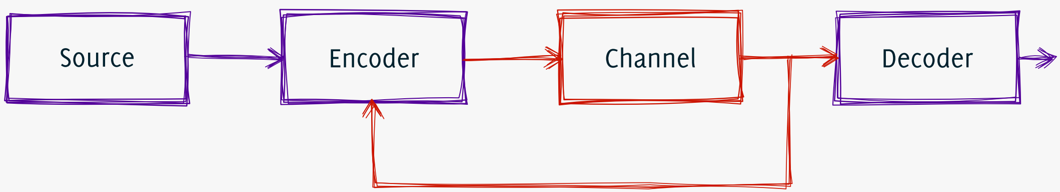
	Memoryless	Causal real-time	
Prob of error	0.1346	0.0564	240% better
# of strategies	$(2^2 \times 2^2 \times 2^2)^2$	$(2^2 \times 2^4 \times 2^8)^2$	simple example
	$O(10^3)$	$O(10^8)$	

Can we search for optimal real-time strategies efficiently?

Literature Overview

- Encoder knows what decoder knows

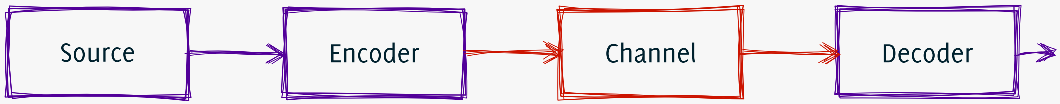
Noiseless channel: Lloyd, 1977; Witsenhausen, 1979; Neuhoff and Gilbert, 1982; Linder and Lugosi, 2001; Weissman and Merhav, 2002; Linder and Zamir, 2006.



Noisy channel with noiseless feedback:
Walrand and Varaiya, 1982

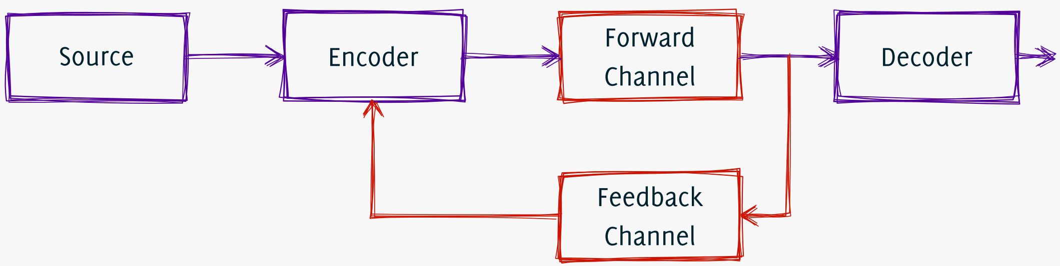
Literature Overview

- Encoder and decoder have different information



Finite memory: Gaarder and Slepian, 1982;
Mahajan and Teneketzis, 2006

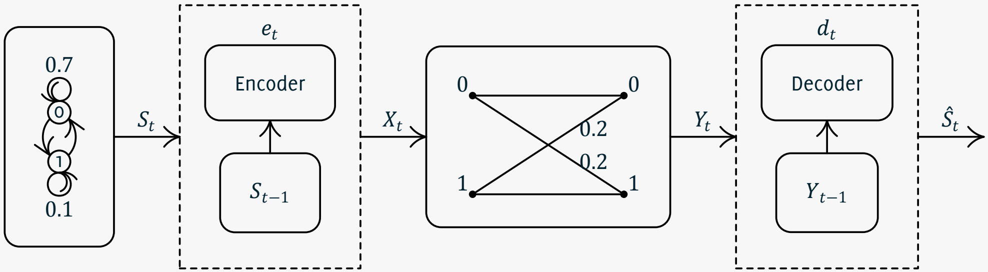
Noisy channel: Teneketzis, 2006,
Mahajan and Teneketzis, 2009b



Noisy channel with noisy feedback:
Mahajan and Teneketzis, 2008

Memory and delay consideration

Memory and delay consideration



Finite Memory

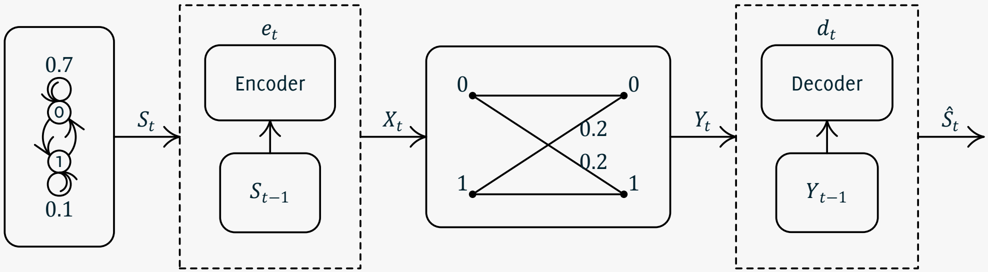
Encoder

$$x_t = e_t(s_t, s_{t-1})$$

Decoder

$$\hat{s}_t = d_t(y_t, y_{t-1})$$

Memory and delay consideration



Finite Memory

Encoder

$$x_t = e_t(s_t, s_{t-1})$$

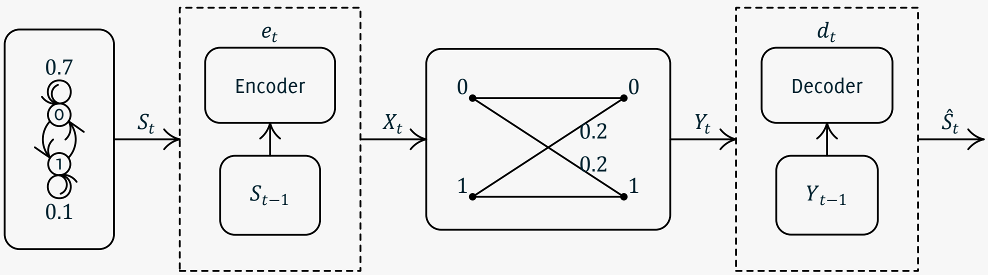
Decoder

$$\hat{s}_t = d_t(y_t, y_{t-1})$$

- Communication Strategy

$$E = (e_1, e_2, \dots, e_T), \quad D = (d_1, d_2, \dots, d_T)$$

Memory and delay consideration



Finite Memory

Encoder

$$x_t = e_t(s_t, s_{t-1})$$

Decoder

$$\hat{s}_t = d_t(y_t, y_{t-1})$$

- Communication Strategy

$$E = (e_1, e_2, \dots, e_T), \quad D = (d_1, d_2, \dots, d_T)$$

- Performance

$$\mathcal{J}(E, D) = \lim_{T \to \infty} \mathbb{E}\left\{ \sum_{t=2}^{T} \beta^{t-1} \mathbb{P}(\hat{s}_t \neq s_{t-1}) \right\}$$

Gaarder and Slepian's approach

[Gaarder and Slepian, 1982]

- Choose a time invariant strategy $E = (e, e, \dots, e)$, $D = (d, d, \dots, d)$.

Gaarder and Slepian's approach

[Gaarder and Slepian, 1982]

- Choose a time invariant strategy $E = (e, e, \ldots, e)$, $D = (d, d, \ldots, d)$.
- $\{S_{t-1}, S_t, Y_{t-1}\}$ forms a Markov chain

Gaarder and Slepian's approach

[Gaarder and Slepian, 1982]

- Choose a time invariant strategy $E = (e, e, \ldots, e)$, $D = (d, d, \ldots, d)$.
- $\{S_{t-1}, S_t, Y_{t-1}\}$ forms a Markov chain
- Find its steady-state distribution (if it is unique)

Gaarder and Slepian's approach

[Gaarder and Slepian, 1982]

- Choose a time invariant strategy $E = (e, e, \ldots, e)$, $D = (d, d, \ldots, d)$.
- $\{S_{t-1}, S_t, Y_{t-1}\}$ forms a Markov chain
- Find its steady-state distribution (if it is unique)
- Find the steady-state probability of error

$$\lim_{t \to \infty} \mathbb{E}\left\{ \mathbb{P}(\hat{s}_t \neq \hat{s}_{t-1}) \right\}$$

Gaarder and Slepian's approach

[Gaarder and Slepian, 1982]

- Choose a time invariant strategy $E = (e, e, \ldots, e)$, $D = (d, d, \ldots, d)$.
- $\{S_{t-1}, S_t, Y_{t-1}\}$ forms a Markov chain
- Find its steady-state distribution (if it is unique)
- Find the steady-state probability of error

$$\lim_{t\to\infty} \mathbb{E}\left\{ \mathbb{P}(\hat{s}_t \neq \hat{s}_{t-1}) \right\}$$

- Repeat for all time invariant strategies.

so we deem it best to publish now the results we do have; albeit, incomplete and unsatisfactory as they are. Perhaps others will pick up the fallen torch and run more deftly!

Gaarder and Slepian, 1982

Difficulty with Gaarder and Slepian's approach

- Steady-state distribution of a Markov chain is discontinuous in its transition matrix

Difficulty with Gaarder and Slepian's approach

- Steady-state distribution of a Markov chain is discontinuous in its transition matrix
- For some (E, D), the Markov chain may not have a unique steady-state distribution
  - Multiple recurrence classes $\Rightarrow$ uncountable steady-state distributions

Determining optimal encoders and decoders

None of the existing approaches work

Difficulty with other approaches

Difficulty with other approaches

- Brute force search is computationally challenging

  Number of communication strategies: $\left(|\mathcal{X}|^{|\mathcal{S}|^2} |\hat{\mathcal{S}}|^{|\mathcal{Y}|^2} \right)^T$

Difficulty with other approaches

- Brute force search is computationally challenging

  Number of communication strategies: $\left(|\mathcal{X}|^{|\mathcal{S}|^2} |\hat{\mathcal{S}}|^{|\mathcal{Y}|^2} \right)^T$

- Information theory is not appropriate
  - Based on law of large numbers for long sequences
  - Long sequences introduce delay and require big memory

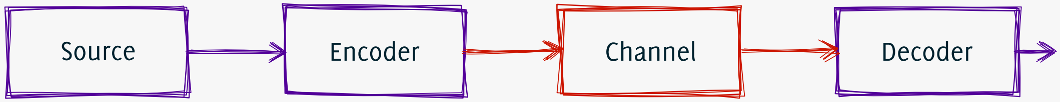
Difficulty with other approaches

- Brute force search is computationally challenging

  Number of communication strategies: $\left(|\mathcal{X}|^{|\mathcal{S}|^2} |\hat{\mathcal{S}}|^{|\mathcal{Y}|^2} \right)^T$

- Information theory is not appropriate
  - Based on law of large numbers for long sequences
  - Long sequences introduce delay and require big memory

- Markov decision theory does not apply

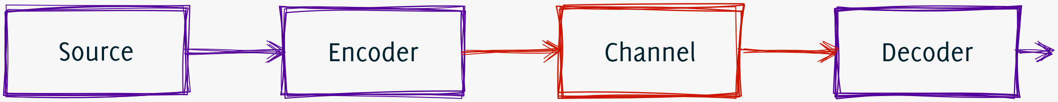


Difficulty with other approaches

- Brute force search is computationally challenging

  Number of communication strategies: $\left(|\mathcal{X}|^{|\mathcal{S}|^2} |\hat{\mathcal{S}}|^{|\mathcal{Y}|^2} \right)^T$

- Information theory is not appropriate
  - Based on law of large numbers for long sequences
  - Long sequences introduce delay and require big memory
- Markov decision theory does not apply



- Orthogonal search (Coordinate descent)
  - May not converge
  - Gives only local optima

Global Optimization

Global Optimization

Research Contributions

- Sequential decomposition
  - Sequential search algorithm
  - Exponentially reduces the search complexity

Research Contributions

- Sequential decomposition
  - Sequential search algorithm
  - Exponentially reduces the search complexity
- Information state
  - No existing methodology
  - Provided an axiomatic approach

Research Contributions

- Sequential decomposition
  - Sequential search algorithm
  - Exponentially reduces the search complexity
- Information state
  - No existing methodology
  - Provided an axiomatic approach
- Common knowledge
  - What can two agents with different information agree upon?
  - Key notion in finding information states

Research Contributions

- Sequential decomposition
  - Sequential search algorithm
  - Exponentially reduces the search complexity
- Information state
  - No existing methodology
  - Provided an axiomatic approach
- Common knowledge
  - What can two agents with different information agree upon?
  - Key notion in finding information states
- Finite or infinite time-steps
  - No priori approach for infinite time-steps
  - Proposed approach works for both

Sequential decomposition

- Divide and Conquer

  Algorithm: One step optimization → sequence of nested optimizations

Sequential decomposition

- Divide and Conquer

  Algorithm: One step optimization → sequence of nested optimizations

- Information state

  - **For each time t, an information state π_t**

    Compression: All the past relevant to the future

Sequential decomposition

- Divide and Conquer

  Algorithm: One step optimization → sequence of nested optimizations

- Information state

  - **For each time t, an information state π_t**

    Compression: All the past relevant to the future

  - **For each info state π_t, a value function $V_t(\pi_t)$**

    Represents: minimum probability of error in the future

Sequential decomposition

- Divide and Conquer

  Algorithm: One step optimization → sequence of nested optimizations

- Information state

  - **For each time t, an information state π_t**

    Compression: All the past relevant to the future

  - **For each info state π_t, a value function $V_t(\pi_t)$**

    Represents: minimum probability of error in the future

- Recursion

  Express: $V_t(\pi_t)$ in terms of $V_{t+1}(\pi_{t+1})$

Sequential decomposition

- **Divide and Conquer**

  Algorithm: One step optimization → sequence of nested optimizations

- **Information state**

  - **For each time t, an information state π_t**

    Compression: All the past relevant to the future

  - **For each info state π_t, a value function $V_t(\pi_t)$**

    Represents: minimum probability of error in the future

- **Recursion**

  Express: $V_t(\pi_t)$ in terms of $V_{t+1}(\pi_{t+1})$

- **Backward Induction**

  Evaluate: $V_t(\pi_t)$ for each π_t moving backward in time

Sequential decomposition

- Divide and Conquer

Exponential reduction in
the search complexity

$$O\left((2^A)^T \right) \to O(T \cdot K \cdot 2^A)$$

- Recursion

  Express: $V_t(\pi_t)$ in terms of $V_{t+1}(\pi_{t+1})$

- Backward Induction

  Evaluate: $V_t(\pi_t)$ for each π_t moving backward in time

Sequential Decomposition

First example of sequential decomposition for
optimal solution of
general non-linear decentralized systems

Sequential Decomposition

First example of sequential decomposition for
optimal solution of
general non-linear decentralized systems

- How do we choose information state π_t

No previous known technique for finding
information states for decentralized systems

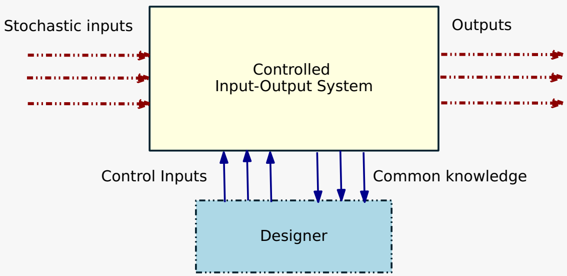
An axiomatic approach to choosing information state

Choosing information state

[Mahajan, 2008, Mahajan and Teneketzis 2008, 2009b]

Do not think in terms of encoders or decoders;

think in terms of a system designer

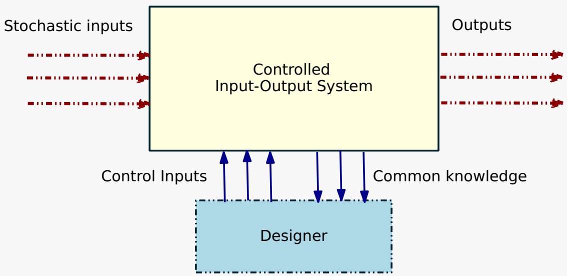


Choosing information state

[Mahajan, 2008, Mahajan and Teneketzis 2008, 2009b]

Do not think in terms of encoders or decoders;
think in terms of a system designer

- Controlled input-output system
  - Control inputs: encoding and decoding functions
  - Stochastic inputs: source output and channel noise
  - Output: source estimates and prob of error

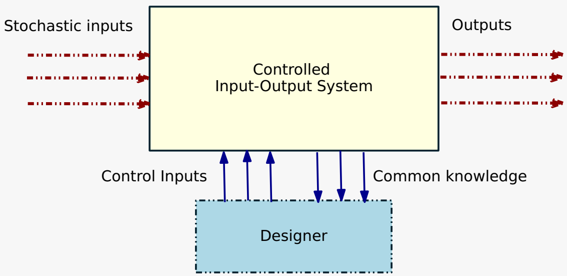


Choosing information state

[Mahajan, 2008, Mahajan and Teneketzis 2008, 2009b]

Do not think in terms of encoders or decoders;
think in terms of a system designer

- Controlled input-output system
  - Control inputs:
    encoding and decoding functions
  - Stochastic inputs:
    source output and channel noise
  - Output: source estimates and prob of error



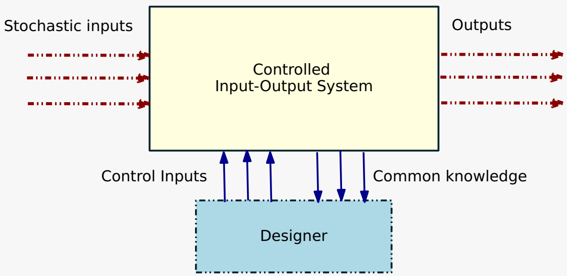
- From the point-of-view of the designer, find

Choosing information state

[Mahajan, 2008, Mahajan and Teneketzis 2008, 2009b]

Do not think in terms of encoders or decoders;
think in terms of a system designer

- Controlled input-output system
  - Control inputs: encoding and decoding functions
  - Stochastic inputs: source output and channel noise
  - Output: source estimates and prob of error



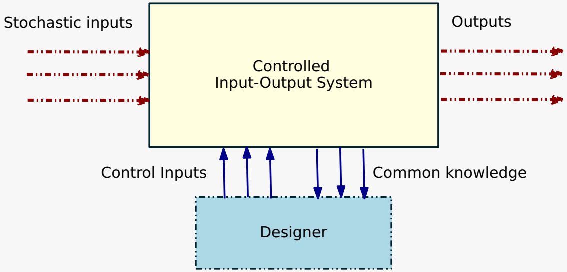
- From the point-of-view of the designer, find
  - **common knowledge between encoder and decoder**

Choosing information state

[Mahajan, 2008, Mahajan and Teneketzis 2008, 2009b]

Do not think in terms of encoders or decoders;
think in terms of a system designer

- Controlled input-output system
  - Control inputs:
    encoding and decoding functions
  - Stochastic inputs:
    source output and channel noise
  - Output: source estimates and prob of error



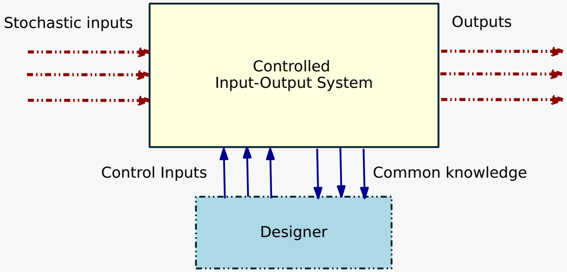
- From the point-of-view of the designer, find
  - **common knowledge between encoder and decoder** $(e_{1:t-1}, d_{1:t-1})$

Choosing information state

[Mahajan, 2008, Mahajan and Teneketzis 2008, 2009b]

Do not think in terms of encoders or decoders;
think in terms of a system designer

- Controlled input-output system



  - Control inputs:
    encoding and decoding functions
  - Stochastic inputs:
    source output and channel noise
  - Output: source estimates and prob of error

- From the point-of-view of the designer, find

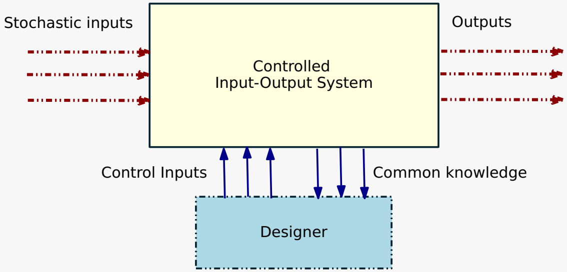
  - **common knowledge between encoder and decoder** $(e_{1:t-1}, d_{1:t-1})$
  - **state sufficient for input-output mapping**

Choosing information state

[Mahajan, 2008, Mahajan and Teneketzis 2008, 2009b]

Do not think in terms of encoders or decoders;
think in terms of a system designer

- Controlled input-output system



  - Control inputs:
    encoding and decoding functions
  - Stochastic inputs:
    source output and channel noise
  - Output: source estimates and prob of error

- From the point-of-view of the designer, find

  - **common knowledge between encoder and decoder** $(e_{1:t-1}, d_{1:t-1})$
  - **state sufficient for input-output mapping** (s_t, s_{t-1}, y_{t-1})

Choosing information state

[Mahajan, 2008, Mahajan and Teneketzis 2008, 2009b]

Information state

$\mathbb{P}$(State for i/o mapping | common knowledge)

source output and channel noise

- Output: source estimates and prob of error

- From the point-of-view of the designer, find
  - **common knowledge between encoder and decoder** $(e_{1:t-1}, d_{1:t-1})$
  - **state sufficient for input-output mapping** (s_t, s_{t-1}, y_{t-1})

Choosing information state

[Mahajan, 2008, Mahajan and Teneketzis 2008, 2009b]

Information state

$\mathbb{P}$(State for i/o mapping | common knowledge)

$$\pi_t = \mathbb{P}(s_t, s_{t-1}, y_{t-1} \mid e_{1:t-1}, d_{1:t-1})$$

- From the point-of-view of the designer, find
  - common knowledge between encoder and decoder $(e_{1:t-1}, d_{1:t-1})$
  - state sufficient for input-output mapping (s_t, s_{t-1}, y_{t-1})

Information state Seq decomposition

[Mahajan and Teneketzis, 2009b]

Finite horizon: An optimal communication strategy can be determined by the solution of the following nested optimality equations

$$V_T(\pi_T) = \min_{e_T, d_T} \mathbb{E}\Big[\, \mathbb{P}(\hat{s}_T \neq s_{T-1}) \,\Big|\, \pi_T, e_T, d_T \,\Big]$$

$$V_t(\pi_t) = \min_{e_t, d_t} \mathbb{E}\Big[\, \mathbb{P}(\hat{s}_t \neq s_{t-1}) + \beta V_{t+1}(\pi_{t+1}) \,\Big|\, \pi_t, e_t, d_t \,\Big]$$

where $\pi_t = \mathbb{P}(s_{t-1}, s_t, y_{t-1} | e_{1:t-1}, d_{1:t-1})$

Information state Seq decomposition

[Mahajan and Teneketzis, 2009b]

Finite horizon: An optimal communication strategy can be determined by the solution of the following nested optimality equations

$$V_T(\pi_T) = \min_{e_T, d_T} \mathbb{E}\Big[\, \mathbb{P}(\hat{s}_T \neq s_{T-1}) \,\Big|\, \pi_T, e_T, d_T \,\Big]$$

$$V_t(\pi_t) = \min_{e_t, d_t} \mathbb{E}\Big[\, \mathbb{P}(\hat{s}_t \neq s_{t-1}) + \beta V_{t+1}(\pi_{t+1}) \,\Big|\, \pi_t, e_t, d_t \,\Big]$$

where $\pi_t = \mathbb{P}(s_{t-1}, s_t, y_{t-1} | e_{1:t-1}, d_{1:t-1})$

Infinite horizon: . . . fixed point equation

$$V(\pi) = \min_{e,d} \mathbb{E}\Big[\, \mathbb{P}(\hat{s}_t \neq s_{t-1}) + \beta V(\pi_+) \,\Big|\, \pi, e, d \,\Big]$$

Optimal communication scheme

- Example with $\beta = 0.9$

$$(e_t, d_t) = \begin{pmatrix} s_t & , & 0 \\ s_{t-1} \oplus s_t & , & y_{t-1} \oplus y_t \\ s_{t-1} & , & y_{t-1} \oplus y_t \end{pmatrix}_{t \pmod 3}$$

Optimal communication scheme

- Example with $\beta = 0.9$

$$(e_t, d_t) = \begin{pmatrix} s_t & , & 0 \\ s_{t-1} \oplus s_t & , & y_{t-1} \oplus y_t \\ s_{t-1} & , & y_{t-1} \oplus y_t \end{pmatrix}_{t \pmod 3}$$

Source	s_1	s_2	s_3	s_4	s_5	s_6	s_7
Encoder							
Decoder							
Estimate	$-$	s_1	s_2	s_3	s_4	s_5	s_6

Optimal communication scheme

- Example with $\beta = 0.9$

$$(e_t, d_t) = \begin{pmatrix} s_t & , & 0 \\ s_{t-1} \oplus s_t & , & y_{t-1} \oplus y_t \\ s_{t-1} & , & y_{t-1} \oplus y_t \end{pmatrix}_{t \pmod 3}$$

Source	s_1	s_2	s_3	s_4	s_5	s_6	s_7
Encoder	s_1						
Decoder	0						
Estimate	$-$	s_1	s_2	s_3	s_4	s_5	s_6

Optimal communication scheme

- Example with $\beta = 0.9$

$$(e_t, d_t) = \begin{pmatrix} s_t & , & 0 \\ s_{t-1} \oplus s_t & , & y_{t-1} \oplus y_t \\ s_{t-1} & , & y_{t-1} \oplus y_t \end{pmatrix}_{t \pmod 3}$$

Source	s_1	s_2	s_3	s_4	s_5	s_6	s_7
Encoder	s_1	$s_1 \oplus s_2$					
Decoder	0	$y_1 \oplus y_2$					
Estimate	$-$	s_1	s_2	s_3	s_4	s_5	s_6

Optimal communication scheme

- Example with $\beta = 0.9$

$$(e_t, d_t) = \begin{pmatrix} s_t & , & \color{red}{0} \\ s_{t-1} \oplus s_t & , & y_{t-1} \oplus y_t \\ s_{t-1} & , & y_{t-1} \oplus y_t \end{pmatrix}_{t \pmod 3}$$

Source	s_1	s_2	s_3	s_4	s_5	s_6	s_7
Encoder	s_1	$s_1 \oplus s_2$	s_2				
Decoder	0	$y_1 \oplus y_2$	$y_2 \oplus y_3$				
Estimate	$-$	s_1	s_2	s_3	s_4	s_5	s_6

Optimal communication scheme

- Example with $\beta = 0.9$

$$(e_t, d_t) = \begin{pmatrix} s_t & , & {\color{red}0} \\ s_{t-1} \oplus s_t & , & y_{t-1} \oplus y_t \\ s_{t-1} & , & y_{t-1} \oplus y_t \end{pmatrix}_{t \pmod 3}$$

Source	s_1	s_2	s_3	s_4	s_5	s_6	s_7
Encoder	s_1	$s_1 \oplus s_2$	s_2	s_4			
Decoder	0	$y_1 \oplus y_2$	$y_2 \oplus y_3$	0			
Estimate	$-$	s_1	s_2	s_3	s_4	s_5	s_6

Optimal communication scheme

- Example with $\beta = 0.9$

$$(e_t, d_t) = \begin{pmatrix} s_t & , & \color{red}{0} \\ s_{t-1} \oplus s_t & , & y_{t-1} \oplus y_t \\ s_{t-1} & , & y_{t-1} \oplus y_t \end{pmatrix}_{t \pmod 3}$$

Source	s_1	s_2	s_3	s_4	s_5	s_6	s_7
Encoder	s_1	$s_1 \oplus s_2$	s_2	s_4	$s_4 \oplus s_5$		
Decoder	0	$y_1 \oplus y_2$	$y_2 \oplus y_3$	0	$y_4 \oplus y_5$		
Estimate	$-$	s_1	s_2	s_3	s_4	s_5	s_6

Optimal communication scheme

- Example with $\beta = 0.9$

$$(e_t, d_t) = \begin{pmatrix} s_t & , & \color{red}{0} \\ s_{t-1} \oplus s_t & , & y_{t-1} \oplus y_t \\ s_{t-1} & , & y_{t-1} \oplus y_t \end{pmatrix}_{t \pmod 3}$$

Source	s_1	s_2	s_3	s_4	s_5	s_6	s_7
Encoder	s_1	$s_1 \oplus s_2$	s_2	s_4	$s_4 \oplus s_5$	s_5	
Decoder	0	$y_1 \oplus y_2$	$y_2 \oplus y_3$	0	$y_4 \oplus y_5$	$y_5 \oplus y_6$	
Estimate	$-$	s_1	s_2	s_3	s_4	s_5	s_6

Optimal communication scheme

- Example with $\beta = 0.9$

$$(e_t, d_t) = \begin{pmatrix} s_t & , & 0 \\ s_{t-1} \oplus s_t & , & y_{t-1} \oplus y_t \\ s_{t-1} & , & y_{t-1} \oplus y_t \end{pmatrix}_{t \pmod 3}$$

Source	s_1	s_2	s_3	s_4	s_5	s_6	s_7
Encoder	s_1	$s_1 \oplus s_2$	s_2	s_4	$s_4 \oplus s_5$	s_5	s_7
Decoder	0	$y_1 \oplus y_2$	$y_2 \oplus y_3$	0	$y_4 \oplus y_5$	$y_5 \oplus y_6$	0
Estimate	–	s_1	s_2	s_3	s_4	s_5	s_6

Optimal strategies may be time-varying

Optimal strategies may be time-varying

- Although the system is time homogeneous, time-invariant strategies need not be optimal

Time homogeneous meta-strategy $(e_t, d_t) = \Delta(\pi_t)$

Optimal strategies may be time-varying

- Although the system is time homogeneous, time-invariant strategies need not be optimal

  Time homogeneous meta-strategy $(e_t, d_t) = \Delta(\pi_t)$

- Gaarder and Slepian assumed time-invariant strategies

This restriction made their setup extremely difficult to solve.

Optimal strategies may be time-varying

- Although the system is time homogeneous, time-invariant strategies need not be optimal

Time homogeneous meta-strategy $(e_t, d_t) = \Delta(\pi_t)$

- Gaarder and Slepian assumed time-invariant strategies

This restriction made their setup extremely difficult to solve.

Design of decentralized systems requires a paradigm shift

Summary

Summary

- Highly non-trivial problem

  No existing technique was directly applicable

Summary

- Highly non-trivial problem

  No existing technique was directly applicable

- Novel way to approach system design

  - System designer
  - Common knowledge
  - Information information

Summary

- Highly non-trivial problem

  No existing technique was directly applicable

- Novel way to approach system design

  - System designer
  - Common knowledge
  - Information information

- Qualitative difference in results

Even for infinite horizon problems,
time-invariant strategies may not be optimal

Generality of the approach

Generality of the approach

- General model for real-time communication
  - Full-memory encoder or decoder (Mahajan and Teneketzis, 2009b)
  - Channels with memory (Mahajan and Teneketzis, 2009b)
  - Higher order Markov sources (Mahajan and Teneketzis, 2009b)
  - Presence of noisy feedback (Mahajan and Teneketzis, 2008)

Generality of the approach

- General model for real-time communication
  - Full-memory encoder or decoder (Mahajan and Teneketzis, 2009b)
  - Channels with memory (Mahajan and Teneketzis, 2009b)
  - Higher order Markov sources (Mahajan and Teneketzis, 2009b)
  - Presence of noisy feedback (Mahajan and Teneketzis, 2008)
- Towards a theory of real-time network communication

Generality of the approach

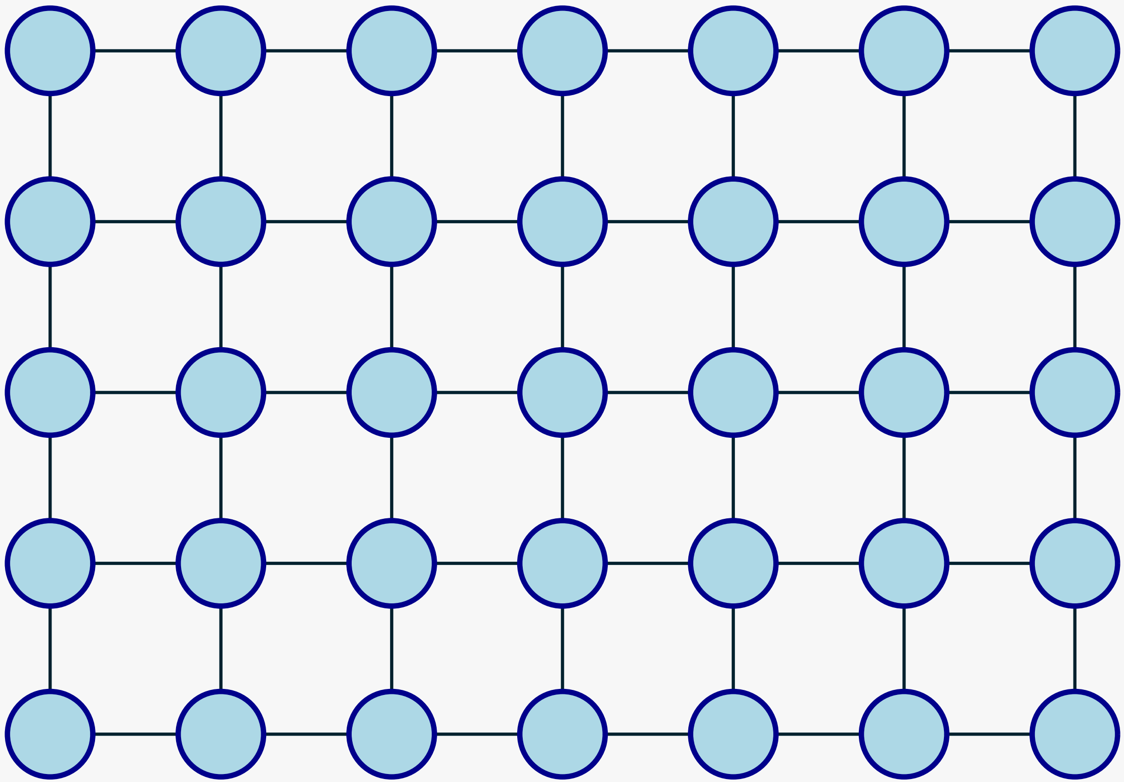
- General model for real-time communication
  - Full-memory encoder or decoder (Mahajan and Teneketzis, 2009b)
  - Channels with memory (Mahajan and Teneketzis, 2009b)
  - Higher order Markov sources (Mahajan and Teneketzis, 2009b)
  - Presence of noisy feedback (Mahajan and Teneketzis, 2008)
- Towards a theory of real-time network communication
- Optimal control over noisy channels

Generality of the approach

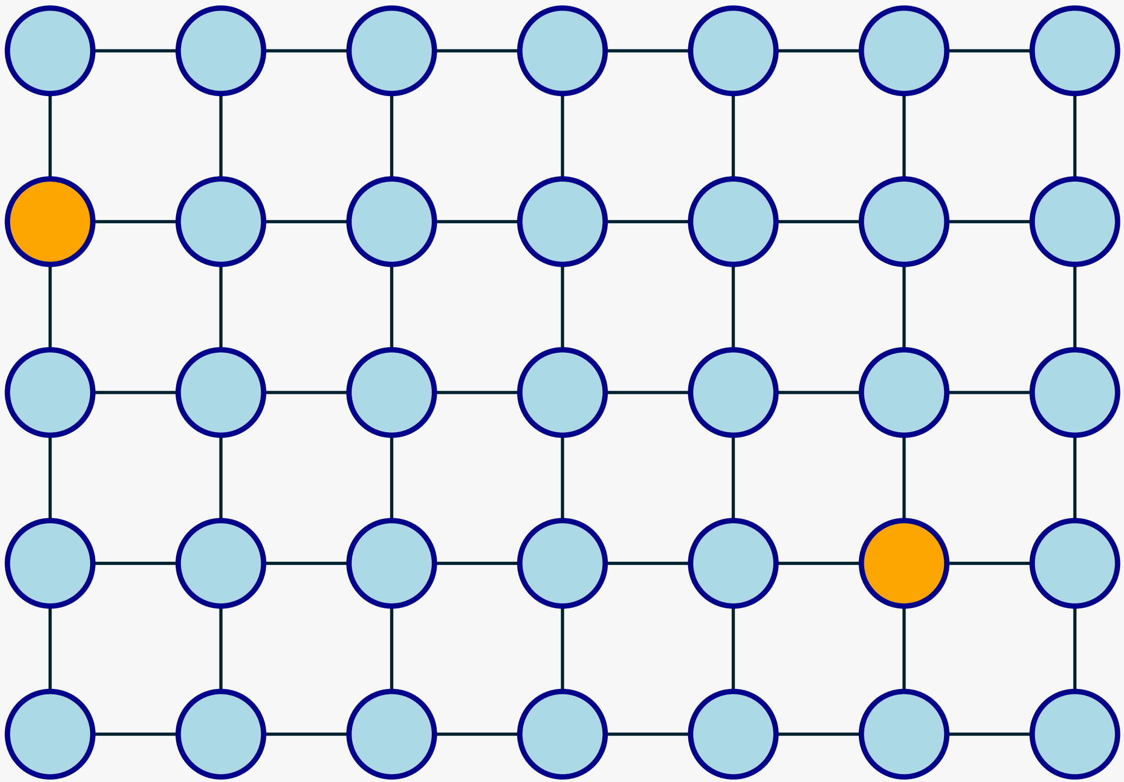
- General model for real-time communication
  - Full-memory encoder or decoder (Mahajan and Teneketzis, 2009b)
  - Channels with memory (Mahajan and Teneketzis, 2009b)
  - Higher order Markov sources (Mahajan and Teneketzis, 2009b)
  - Presence of noisy feedback (Mahajan and Teneketzis, 2008)
- Towards a theory of real-time network communication
- Optimal control over noisy channels
- Delayed sharing patterns

  A 40 year old open problem

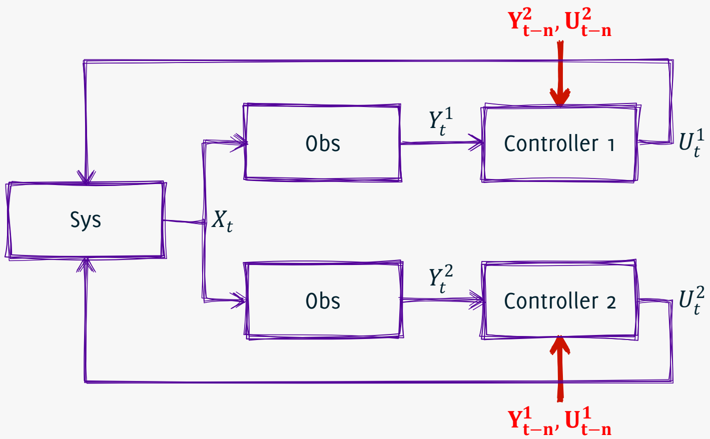
Delayed Sharing Patterns



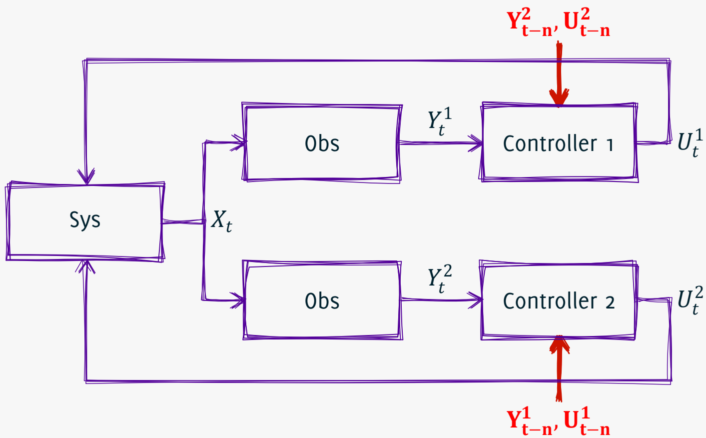
Delayed Sharing Patterns



Delayed Sharing Patterns

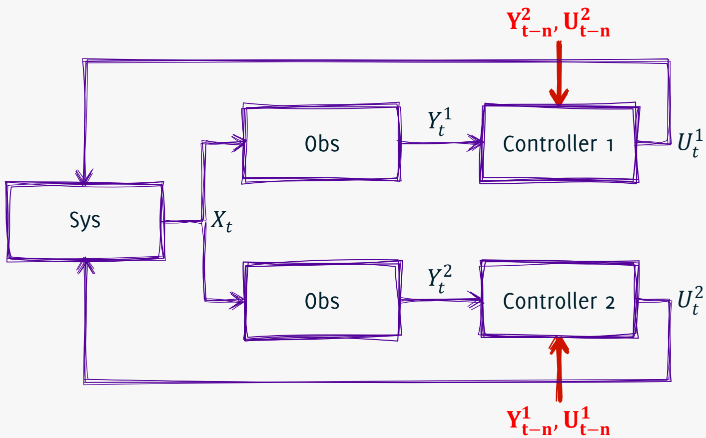


Delayed Sharing Patterns



$$U^i_t = g^i_t(C_t, L^i_t)$$

Delayed Sharing Patterns



$$U_t^i = g_t^i(C_t, L_t^i)$$

where

$$C_t = \left(\begin{bmatrix} Y_1^1 \\ \vdots \\ Y_{t-n}^1 \end{bmatrix}, \begin{bmatrix} Y_1^2 \\ \vdots \\ Y_{t-n}^2 \end{bmatrix}, \begin{bmatrix} U_1^1 \\ \vdots \\ U_{t-n}^1 \end{bmatrix}, \begin{bmatrix} U_1^2 \\ \vdots \\ U_{t-n}^2 \end{bmatrix} \right) \qquad L_t^i = \left(\begin{bmatrix} Y_{t-n+1}^1 \\ \vdots \\ Y_t^1 \end{bmatrix}, \begin{bmatrix} U_{t-n+1}^1 \\ \vdots \\ U_{t-1}^1 \end{bmatrix} \right)$$

Common Information

Local Information

Design Difficulty

Common information C_t is increasing with time

Design Difficulty

Common information C_t is increasing with time

- Conjecture (Witsenhausen, 1971)

Without loss of optimality, each controller can replace C_t by $\mathbb{P}(X_{t-n} \mid C_t)$

Design Difficulty

Common information C_t is increasing with time

- Conjecture (Witsenhausen, 1971)

  Without loss of optimality, each controller can replace C_t by $\mathbb{P}(X_{t-n} \mid C_t)$

- Varaiya and Walrand (1979)
  - True for $n = 1$
  - False for $n > 1$

Open problem for 40 years

Does a information state for C_t exist?

How do we find such a information state?

Importance of the problem

- Applications (of one step delay sharing)
  - Power systems: Altman *et. al*, 2009
  - Queueing theory: Kuri and Kumar, 1995
  - Communication networks: Grizzle *et. al*, 1982
  - Stochastic games: Papavassilopoulos, 1982; Chang and Cruz, 1983
  - Economics: Li and Wu, 1991

Importance of the problem

- Applications (of one step delay sharing)
  - Power systems: Altman *et. al*, 2009
  - Queueing theory: Kuri and Kumar, 1995
  - Communication networks: Grizzle *et. al*, 1982
  - Stochastic games: Papavassilopoulos, 1982; Chang and Cruz, 1983
  - Economics: Li and Wu, 1991
- Conceptual significance
  - Understanding the design of networked control systems
  - Bridge between centralized and decentralized systems
  - Insights for the design of general decentralized systems

Solution approach

[Nayyar, Mahajan, and Teneketzis, 2010]

- Common knowledge between all agents

Solution approach

[Nayyar, Mahajan, and Teneketzis, 2010]

- Common knowledge between all agents

$$(C_t, g^1_{1:t-1}, g^2_{1:t-1})$$

Solution approach

[Nayyar, Mahajan, and Teneketzis, 2010]

- Common knowledge between all agents

$$(C_t, g^1_{1:t-1}, g^2_{1:t-1})$$

- State sufficient for i/o mapping

Solution approach

[Nayyar, Mahajan, and Teneketzis, 2010]

- Common knowledge between all agents

$$(C_t, g^1_{1:t-1}, g^2_{1:t-1})$$

- State sufficient for i/o mapping

$$(X_t, L^1_t, L^2_t)$$

Solution approach

[Nayyar, Mahajan, and Teneketzis, 2010]

- Common knowledge between all agents

$$(C_t, g^1_{1:t-1}, g^2_{1:t-1})$$

- State sufficient for i/o mapping

$$(X_t, L^1_t, L^2_t)$$

- Information state

Solution approach

[Nayyar, Mahajan, and Teneketzis, 2010]

- Common knowledge between all agents

$$(C_t, g^1_{1:t-1}, g^2_{1:t-1})$$

- State sufficient for i/o mapping

$$(X_t, L^1_t, L^2_t)$$

- Information state

$$\pi_t = \mathbb{P}(\text{state for i/o mapping} \mid \text{common knowledge})$$

Solution approach

[Nayyar, Mahajan, and Teneketzis, 2010]

- Common knowledge between all agents

$$(C_t, g^1_{1:t-1}, g^2_{1:t-1})$$

- State sufficient for i/o mapping

$$(X_t, L^1_t, L^2_t)$$

- Information state

$$\begin{aligned} \pi_t &= \mathbb{P}(\text{state for i/o mapping} \mid \text{common knowledge}) \\ &= \mathbb{P}(X_t, L^1_t, L^2_t \mid C_t, g^1_{1:t-1}, g^2_{1:t-1}) \end{aligned}$$

Solution approach

[Nayyar, Mahajan, and Teneketzis, 2010]

- Common knowledge between all agents

Structure of optimal control law

Without loss of optimality, each controller can replace C_t by $\mathbb{P}(X_t, L^1_t, L^2_t \mid C_t, g^1_{1:t-1}, g^2_{1:t-1})$

Can also write a sequential decomposition based on this information state

A systematic approach can easily resolve long-standing conceptual difficulties in decentralized systems

Conclusions
Optimal design of decentralized systems

Decentralized system: Salient features

- Multiple agents

  Decision making by multiple agents in stochastic dynamic environment

- Coordination issues

  All agents must coordinate to achieve a system-wide objective

- Communication constraints

  Data must be communicated within fixed finite delay

- Robustness

  System model may not be known completely

- Exploiting domain knowledge

  Application specific modeling assumptions

Decentralized systems: Research directions

- Real-time communication

  Delay sensitive communication

- Optimal control over noisy channels

  Communication and coordination

- Delay sharing pattterns

  Coordination

- Communication over unknown channels

  Robustness

- Calibration and validation of remote sensing observations

  Exploiting domain knowledge

Summary

- Systematic approach to design decentralized systems

Summary

- Systematic approach to design decentralized systems
- Algorithm to sequentially synthesize optimal controllers

  Based on information states

Summary

- Systematic approach to design decentralized systems
- Algorithm to sequentially synthesize optimal controllers

  Based on information states

- Axiomatic approach to find information states
  - Find common knowledge
  - Find state for i/o mapping
  - $\mathbb{P}$(state for i/o mapping | common knowledge)

Summary

- Systematic approach to design decentralized systems
- Algorithm to sequentially synthesize optimal controllers

  Based on information states

- Axiomatic approach to find information states
  - Find common knowledge
  - Find state for i/o mapping
  - $\mathbb{P}$(state for i/o mapping | common knowledge)
- Delayed sharing pattern

  Able to resolve a long standing open conjecture

Future Directions

- Control of power systems
  - Renewable energy: unpredictable generation
  - Energy markets: Game theoretic considerations


Future Directions

- Control of power systems


  - Renewable energy: unpredictable generation
  - Energy markets: Game theoretic considerations

- Environmental sensor networks

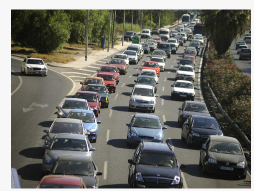

  - Climate change: cheap yet reliable monitoring
  - Calibration validation of remote sensing observations Time varying sampling

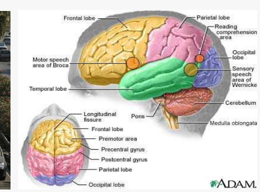
Future Directions

- **Control of power systems**
  - Renewable energy: unpredictable generation
  - Energy markets: Game theoretic considerations

- **Environmental sensor networks**
  - Climate change: cheap yet reliable monitoring
  - Calibration validation of remote sensing observations Time varying sampling

- **Control and coordination**
  - Transportation networks
  - Bioscience and medicine
  - . . .





References: real-time communication

- A. Mahajan and D. Teneketzis, Optimal design of sequential real-time communication systems, *IEEE Trans Information Theory*, Nov 2009.
- A. Mahajan and D. Teneketzis, On the design of globally optimal communication strategies for noisy real-time communication systems with noisy feedback, *IEEE Selected Areas in Comm*, May 2008
- A. Mahajan, Structure of optimal block Markov superposition coding for multiple access channel with feedback, *Information Theory and Applications (ITA) Workshop*, 2010
- A. Mahajan, Optimal sequential transmission over broadcast channels with nested feedback, *47th Allerton Conference*, 2009
- A. Mahajan and D. Teneketzis, Fixed delay optimal joint source-channel coding for finite-memory systems, *IEEE Int. Symp. of Information Theory (ISIT)*, 2006
- A. Mahajan and D. Teneketzis, “Real-time communication systems with noisy feedback,” *IEEE Information Theory Workshop (ITW)*, 2007.

References: control over noisy channels

- A. Mahajan and D. Teneketzis, Optimal performance of networked control systems with non-classical information structures, *SIAM Journal of Control and Optimization*, 2009
- A. Mahajan and D. Teneketzis, Optimal performance of feedback control systems with limited communication over noisy channels, *45th IEEE Conference on Decision and Control (CDC)*, 2006

References: General decentralized control problem

- A. Nayyar, A. Mahajan and D. Teneketzis, Optimal control strategies om delayed sharing information structures, *IEEE Transactions on Automatic Control*, Feb 2010 (submitted)
- A. Nayyar, A. Mahajan, and D. Teneketzis, A separation result for delayed sharing information structures, *American Control Conference (ACC)*, 2010
- A. Mahajan and S. Yüksel, Measure and cost dependent properties of information structures, *American Control Conference (ACC)*, 2010
- A. Mahajan and S. Tatikonda, Sequential team form and its simplification using graphical models, *47th Allerton Conference*, 2009
- A. Mahajan, Sequential decomposition of systems with non-classical information structures: Some examples, *Information Theory and Applications (ITA) Workshop*, 2009.

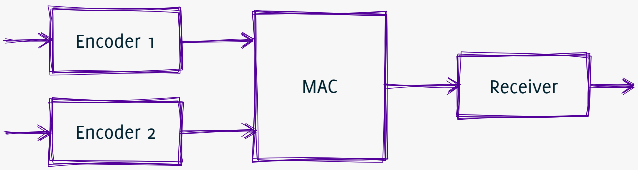
Thank you

Backup Slides

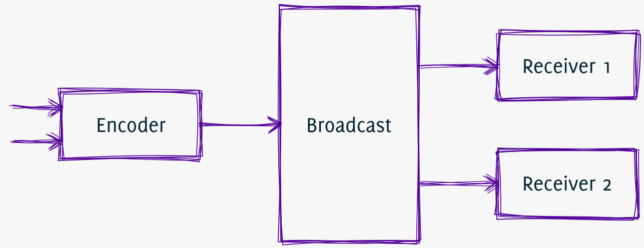
Contents

Towards a theory of real-time network communication

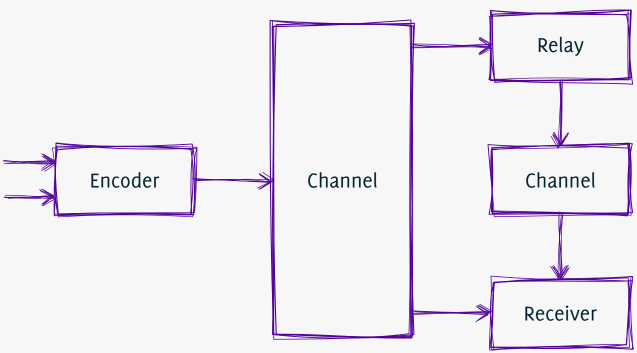
Towards a theory of real-time network communication



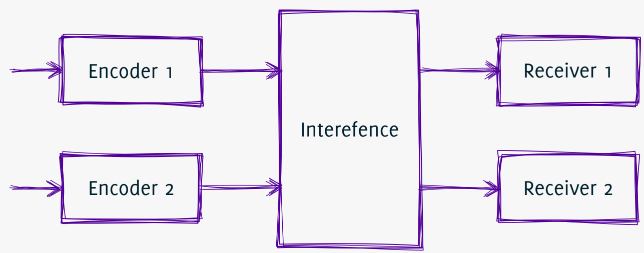
Multiple access channel (M, 2009a)



Broadcast channel (M, 2009b)



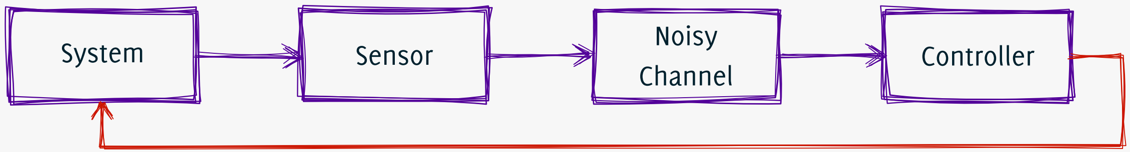
Relay channel



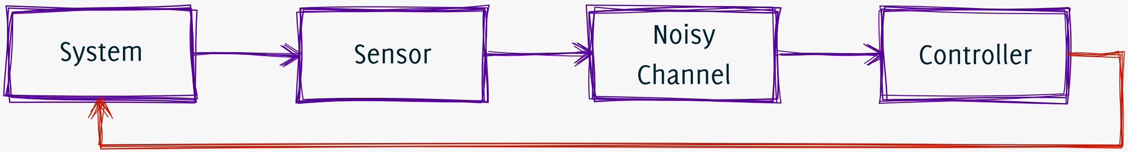
Interference channel

A surprisingly related problem

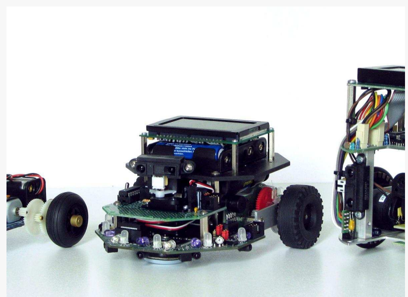
Optimal control over noisy channels

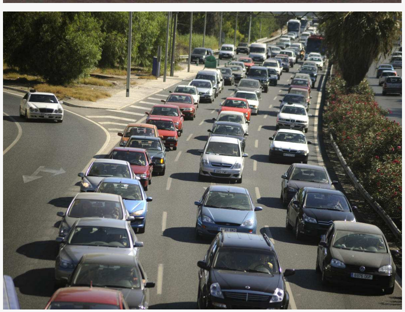
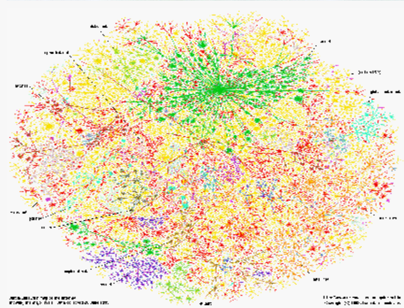
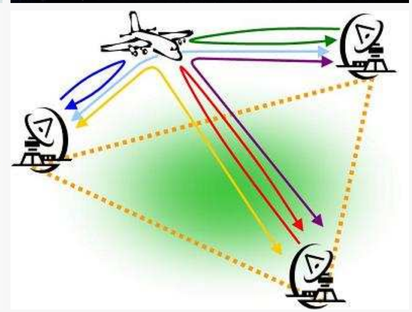


Optimal control over noisy channels



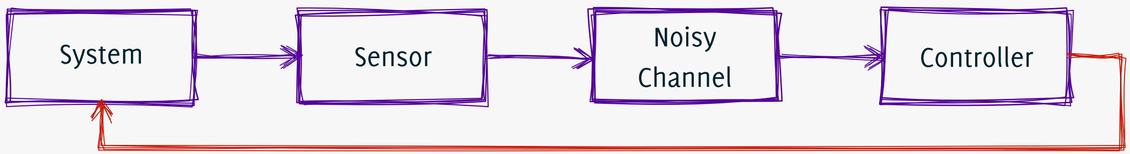
- Simplest example of optimal control of networked systems

Optimal control over noisy channels

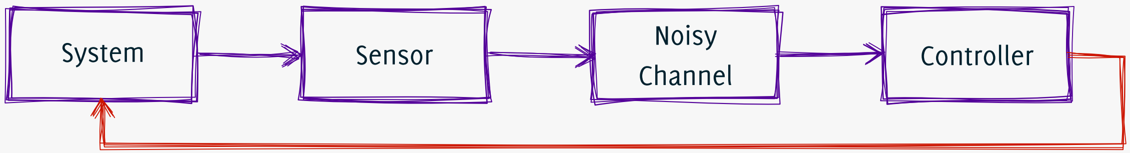
[M and Teneketzis, 2009a]



- Salient features
  - Noisy channel
  - Performance optimization

Optimal control over noisy channels

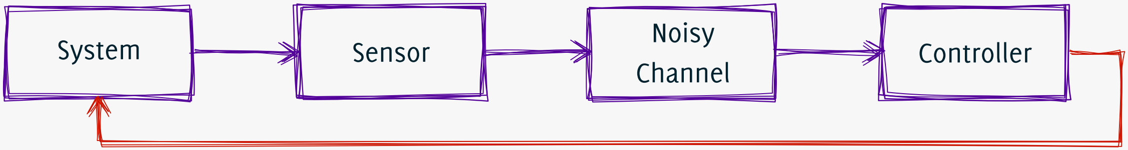
[M and Teneketzis, 2009a]



- Salient features
  - Noisy channel
  - Performance optimization
- Control vs (real-time) communication

Optimal control over noisy channels

[M and Teneketzis, 2009a]



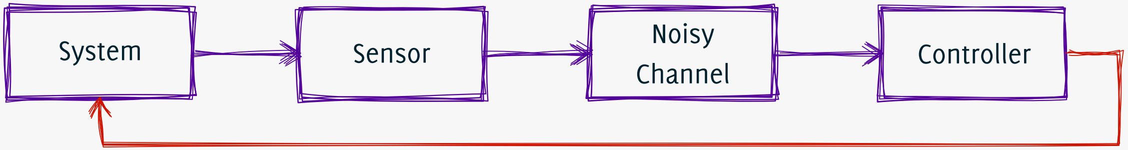
- Salient features
  - Noisy channel
  - Performance optimization
- Control vs (real-time) communication
  - **Similar design difficulties**

    non-comparable information

    - No brute force
    - No Markov decision theory
    - No orthogonal search

Optimal control over noisy channels

[M and Teneketzis, 2009a]



- Salient features
  - Noisy channel
  - Performance optimization
- Control vs (real-time) communication
  - **Similar design difficulties**

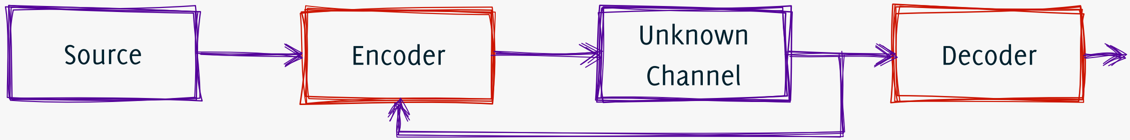
    non-comparable information

    - No brute force
    - No Markov decision theory
    - No orthogonal search
  - **Similar solution approach works**

    Algorithm to sequentially search for the optimal encoding and control strategies

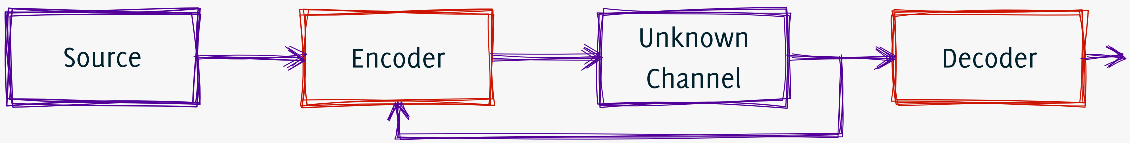
Feedback communication over unknown channels

Feedback communication over unknown channels



- Trade off between learning and communication

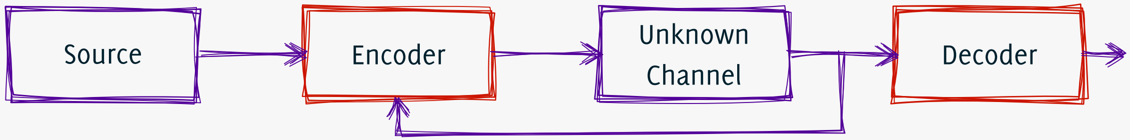
Feedback communication over unknown channels



- Trade off between learning and communication
- Performance criterion:

Error exponent $E = \lim_{n\to\infty} \mathbb{P}^{(n)}_{\text{error}}$

Feedback communication over unknown channels



- Trade off between learning and communication
- Performance criterion:

  Error exponent $E = \lim_{n\to\infty} \mathbb{P}^{(n)}_{\text{error}}$

- Training based scheme
  - Send training sequence
  - Estimate channel and communicate

Feedback communication over unknown channels

- Proposed coding scheme

$$E_{\text{proposed}} = \alpha E_{\text{known}}$$

Feedback communication over unknown channels

- Proposed coding scheme

$$E_{\text{proposed}} = \alpha E_{\text{known}}$$

- Main insights
  - Need to send multiple training sequences
  - Channel estimation for each training sequence should be done independently

Feedback communication over unknown channels

- Proposed coding scheme

$$E_{\text{proposed}} = \alpha E_{\text{known}}$$

- Main insights
  - Need to send multiple training sequences
  - Channel estimation for each training sequence should be done independently
- Reference

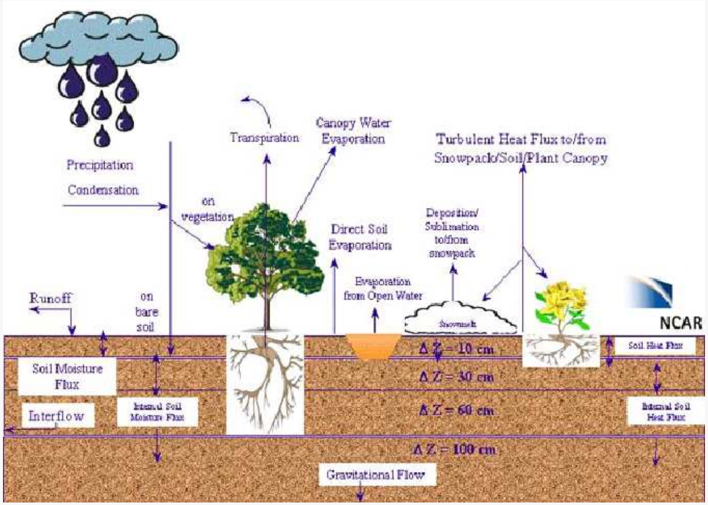
A. Mahajan and S. Tatikonda, Opportunistic capacity and error exponents of compound channel with feedback, *IEEE Trans of Info Theory*, 2010 (submitted)

Calibration and Validation of Remote sensing observations

Monitoring soil moisture

- Measurement need for earth science

  NASA Earth Science focus:
  climate, carbon, weather,
  water, surface, and atmosphere

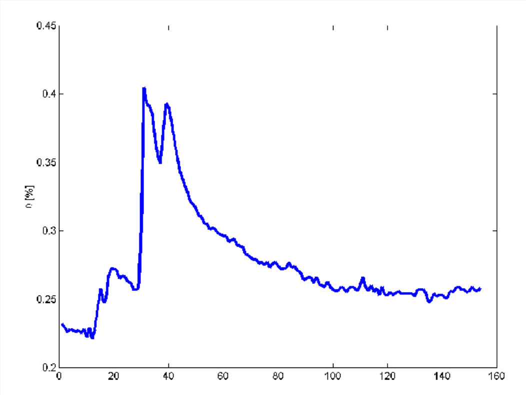


- Challenges

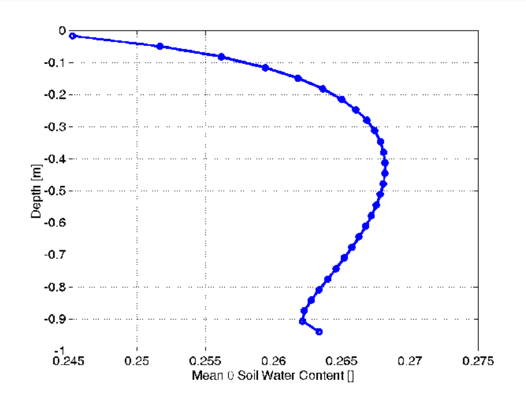
  - Complicated variation
    Depends on temperature, vegetation, precipitation, soil texture, topology, etc.

  - Remote sensing gives coarse estimates
    $O(1\text{km})$ to $O(10\text{km})$

Variation of Soil Moisture

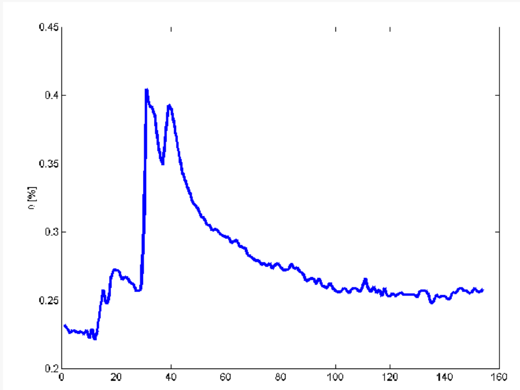


Variation with time

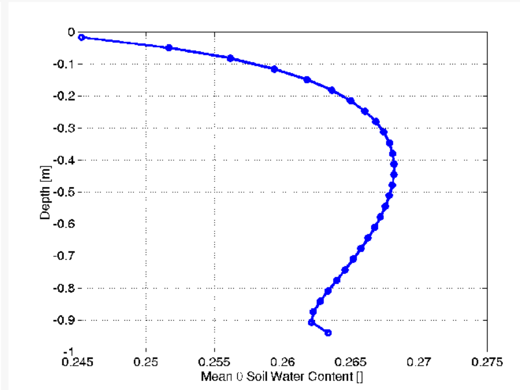


Variation with depth

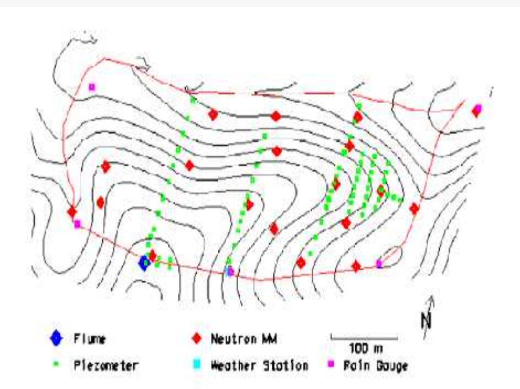
Variation of Soil Moisture



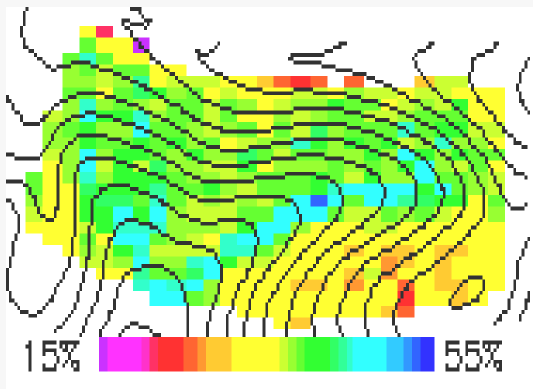
Variation with time



Variation with depth

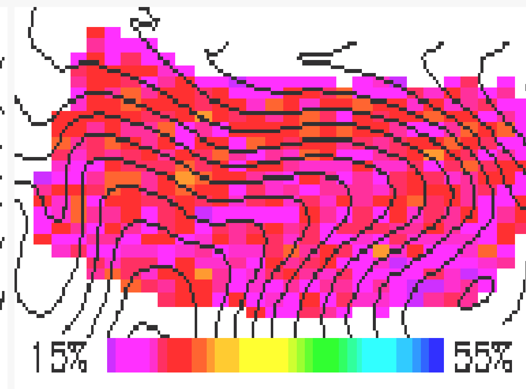


Topology



After rainfall

Follows topology



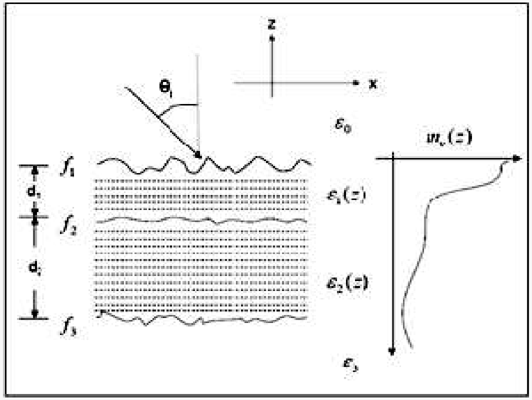
After dry run

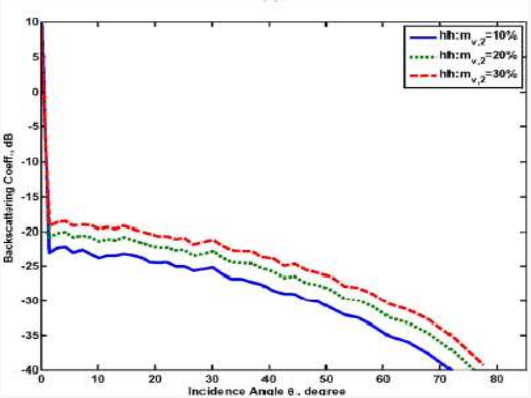
Follows vegetation

Radar measurements

- Backscattering coefficient

  Depends on incidence angle, soil depth, soil type, and vegetation cover





Sensor Scheduling

- Limited battery life

  Measurement and communication consumes energy

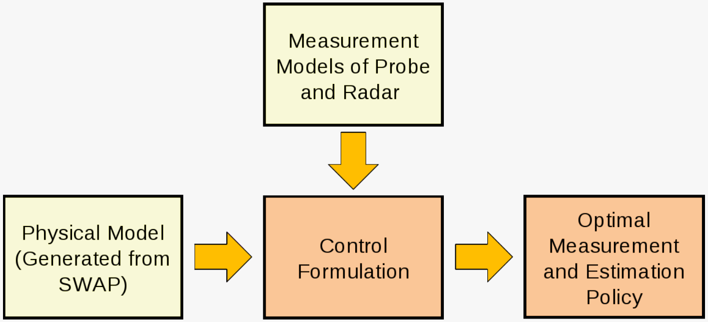
- Sleep scheduling

  Does the sensor need to switch on its radio to determine when to take a measurement?

- Sensor placement

  Need to cover a large area to match satellite footprint

Overview

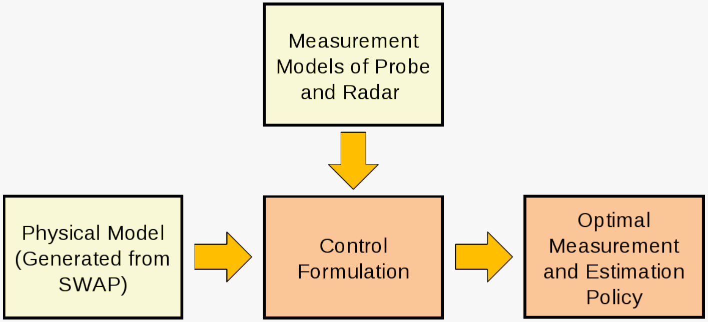


Measurement Models of Probe and Radar
Physical Model (Generated from SWAP)
Control Formulation
Optimal Measurement and Estimation Policy

Overview

- Physical model

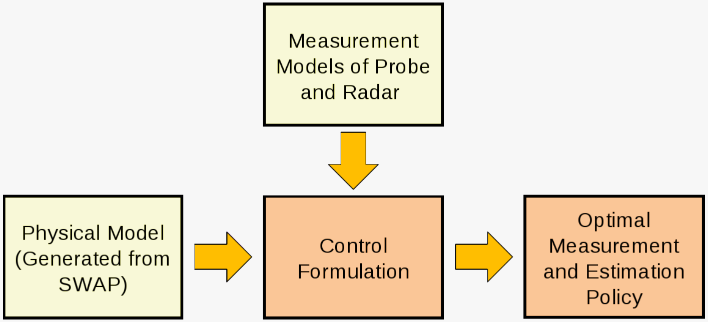
  Use a community standard numerical model developed over 35 years (SWAP)



Overview

Physical model

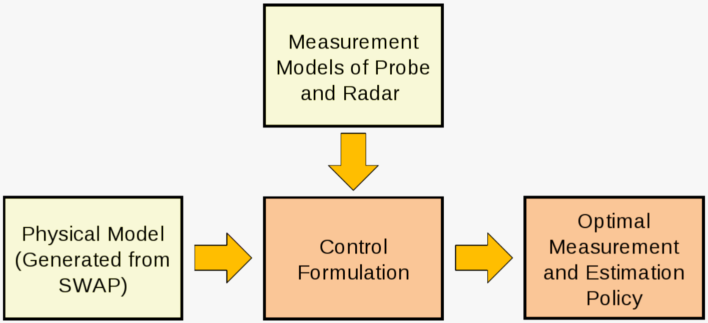
Use a community standard numerical model developed over 35 years (SWAP)



Measurement model

- Forward model for electromagnetic backscattering
- Sensor observation model based on calibration curves

Overview



- **Physical model**

  Use a community standard numerical model developed over 35 years (SWAP)

- **Measurement model**

  - Forward model for electromagnetic backscattering
  - Sensor observation model based on calibration curves

- **Optimal control formulation**

  Use physical and measurement model to guide sensor scheduling and measurement

Overview

- Physical model

Obtain a scalable algorithm
for sensor scheduling

  - Sensor observation model based on calibration curves

- Optimal control formulation

  Use physical and measurement model to guide sensor scheduling and measurement

Field Testing

- Location

  Matthaei Botanical Gardens, Ann Arbor, MI

- Multiple nodes at 3 depths

http://soilscape.eecs.umich.edu

References: Soil moisture measurement

- M. Moghaddam, D. Entekhabi, Y. Goykhman, K. Li, M. Liu, A. Mahajan, A. Nayyar, D. Shuman, and D. Teneketzis, A wireless soil moisture smart sensor web using physics-based optimal control: concept and initial demonstrations, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2010 (accepted)
- D. Shuman, A. Nayyar, A. Mahajan, Y. Goykhman, K. Li, M. Liu, D. Teneketzis, M. Moghaddam, D. Entekhabi, Measurement scheduling for soil moisture sensing: From physical models to optimal control, *Proceedings of the IEEE*, (in revision)
- M. Moghaddam, D. Entekhabi, Y. Goykhman, M. Liu, A. Mahajan, A. Nayyar, D. Shuman, and D. Teneketzis, Soil moisture smart sensor web using data assimilation and optimal control: formulation and first laboratory demonstration, *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2008.

Contents

- Main Slides
  - Real-time communication
  - Memory and delay considerations
  - Research contributions
  - Delayed sharing patterns
  - Conclusions
- Backup slides
  - Towards a theory of real-time network communication
  - Optimal control over noisy channels
  - Communication over unknown channels
  - Calibration and validation of remote sensing observations