A Theory of Semantic Communication

Qi Cao

Joint work with Yulin Shao and Daniz Gündüs

INC, 2025

西安电子科技大学
XIDIAN UNIVERSITY

广州研究院
Guangzhou Institute of Technology

What is Semantic Communication?

- A widely discussed but ambiguously defined concept

What is Semantic Communication?

- A widely discussed but ambiguously defined concept

Semantic communication is like a box of chocolates - you never know what definition you're gonna get.

What is Semantic Communication?

- A widely discussed but ambiguously defined concept
- Most existing works focus on:
  1. Replacing the encoder/decoder with a neural network
     - Often framed as "end-to-end learning"
     - But ultimately still falls under joint source-channel coding

What is Semantic Communication?

- A widely discussed but ambiguously defined concept
- Most existing works focus on:
  1. Replacing the encoder/decoder with a neural network
     - Often framed as "end-to-end learning"
     - But ultimately still falls under joint source-channel coding
  2. Using a knowledge base to reduce transmission
     - This essentially reduces conditional entropy: $H(X|Y) \leq H(X)$

What is Semantic Communication?

- A widely discussed but ambiguously defined concept
- Most existing works focus on:
  1. Replacing the encoder/decoder with a neural network
     - Often framed as "end-to-end learning"
     - But ultimately still falls under joint source-channel coding
  2. Using a knowledge base to reduce transmission
     - This essentially reduces conditional entropy: $H(X|Y) \leq H(X)$

So... what is really new?

Semantic & Technical

- Technical communication ✓ Semantic communication ✗


Semantic & Technical

- Technical communication ✓ Semantic communication ✗


- Technical communication ✗ Semantic communication ✓
  - It deosn't mttaer in waht oredr the ltteers in a wrod are.
  - The olny iprmoetnt tihng is taht the frist and lsat ltteer be at the rghit pclae.
  - The rset can be a toatl mses and you can sitll raed it wouthit porbelm.

Codebook

$$H = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Codebook

$$H = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Codebook

$$H = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Language

Codebook

$$H = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Language

We never consider the cost of transmitting the codebook.

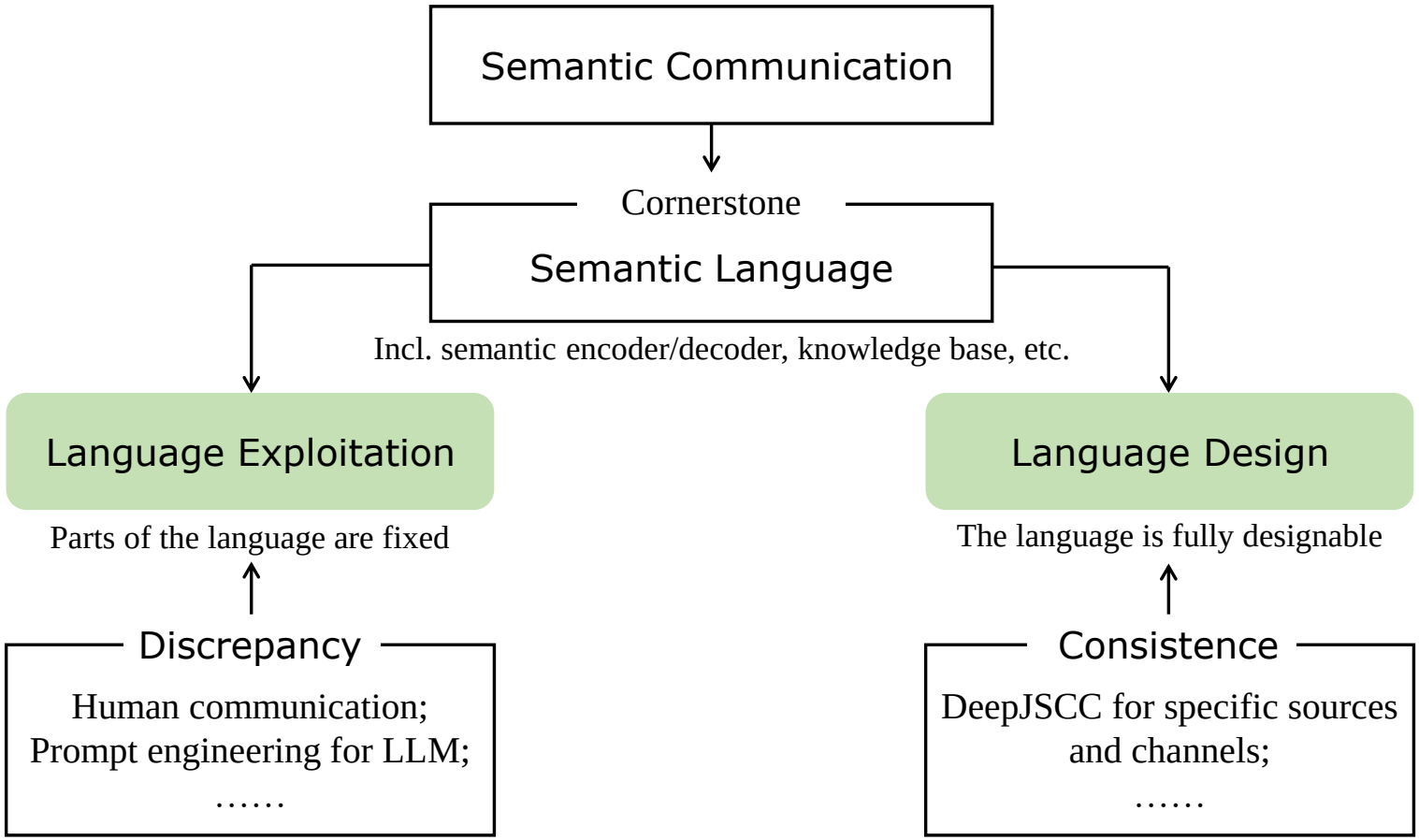
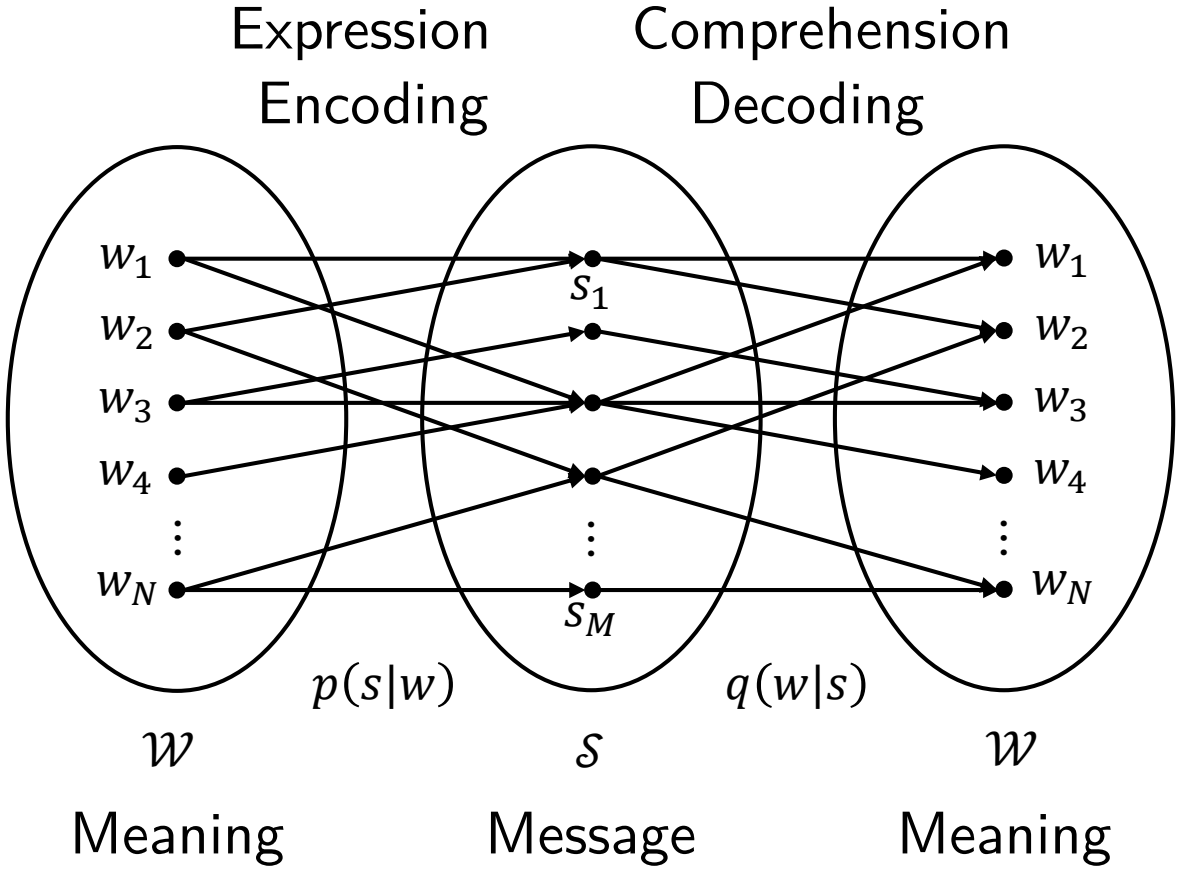


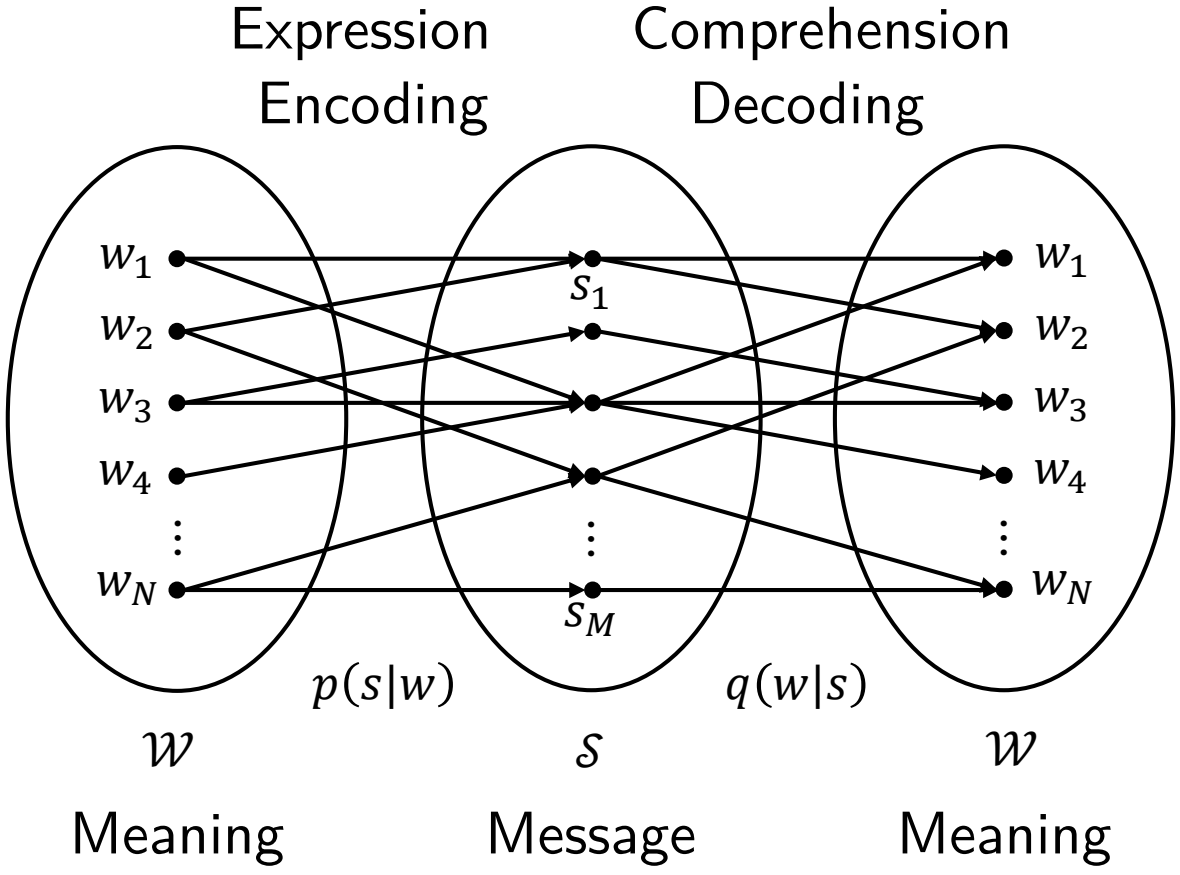
Figure: The framework of semantic communications.[1]

[1] Yulin Shao, Qi Cao, and Deniz Gündüz. "A Theory of Semantic Communication". In: *IEEE Transactions on Mobile Computing* 23.12 (2024), pp. 12211–12228.

A semantic language $(\mathcal{W}, \mathcal{S}, \boldsymbol{P}, \boldsymbol{Q})$.



A semantic language $(\mathcal{W}, \mathcal{S}, \boldsymbol{P}, \boldsymbol{Q})$.



Semantic Distortion

$$D_{\boldsymbol{U},\boldsymbol{Q}} = \sum_{w,s,\hat{s},\hat{w}} p(w)u(s|w)c(\hat{s}|s)q(\hat{w}|\hat{s})d(w,\hat{w}).$$

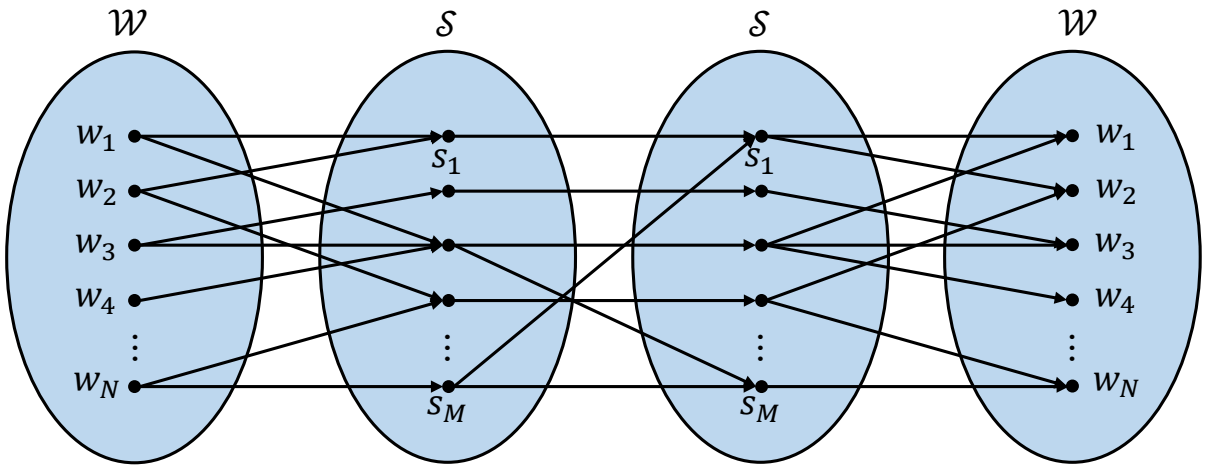
Semantic Cost

$$L_{\boldsymbol{U}} = \sum p(w)u(s|w)\ell(s).$$

Three Problems

- Semantic encoding: encoding intended meanings at the transmitter
- Semantic decoding: reconstructing meaning from received messages
- Combined semantic encoding and decoding (CSED)

Three Problems



(a)	$p(w)$	$u(s\|w)$	$c(\hat{s}\|s)$	$q(\hat{w}\|\hat{s})$	Semantic Encoding
(b)	$q(w)$	$p(s\|w)$	$c(\hat{s}\|s)$	$v(\hat{w}\|\hat{s})$	Semantic Decoding
(c)	$q(w)$	$u(s\|w)$	$c(\hat{s}\|s)$	$v(\hat{w}\|\hat{s})$	CSED

Semantic Encoding (Decoding Scheme is Fixed)

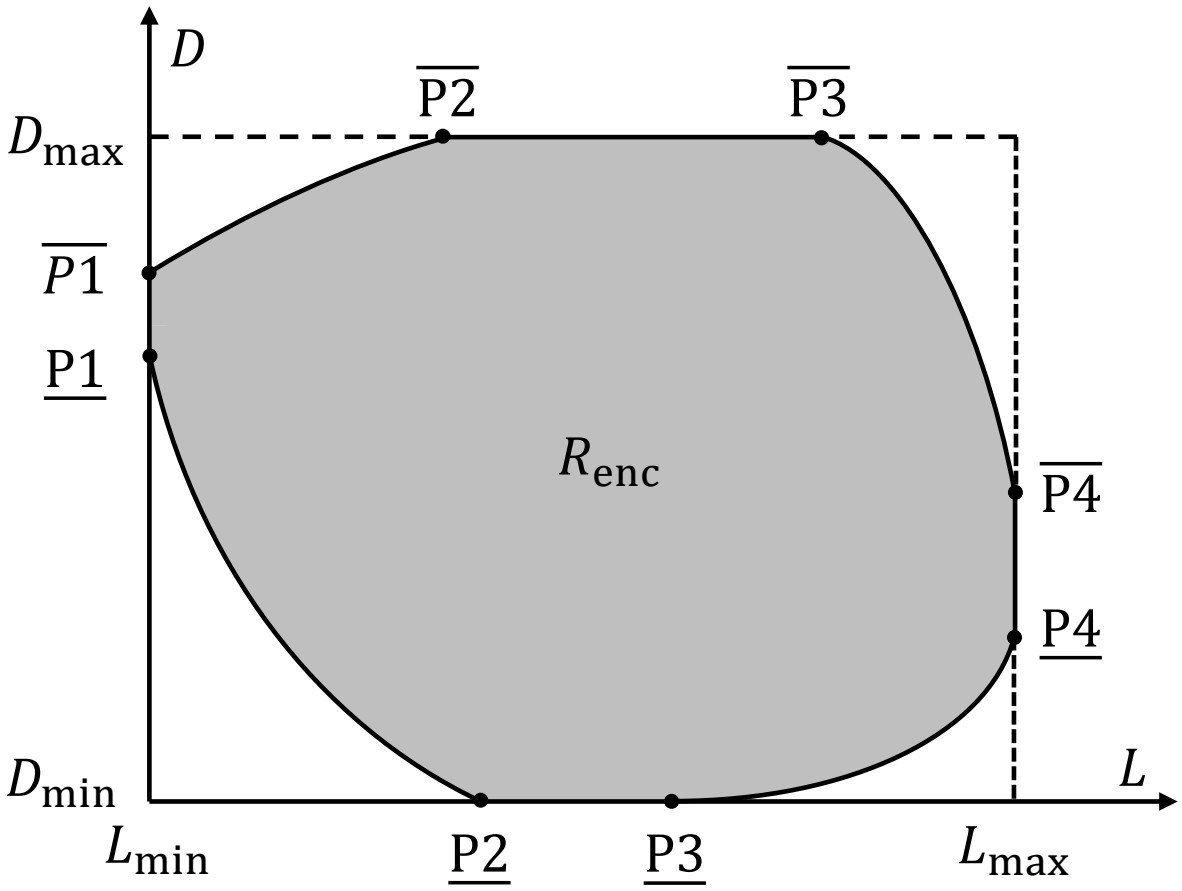


Figure: Semantic rate-distortion region (encoding).

Semantic Encoding (Decoding Scheme is Fixed)

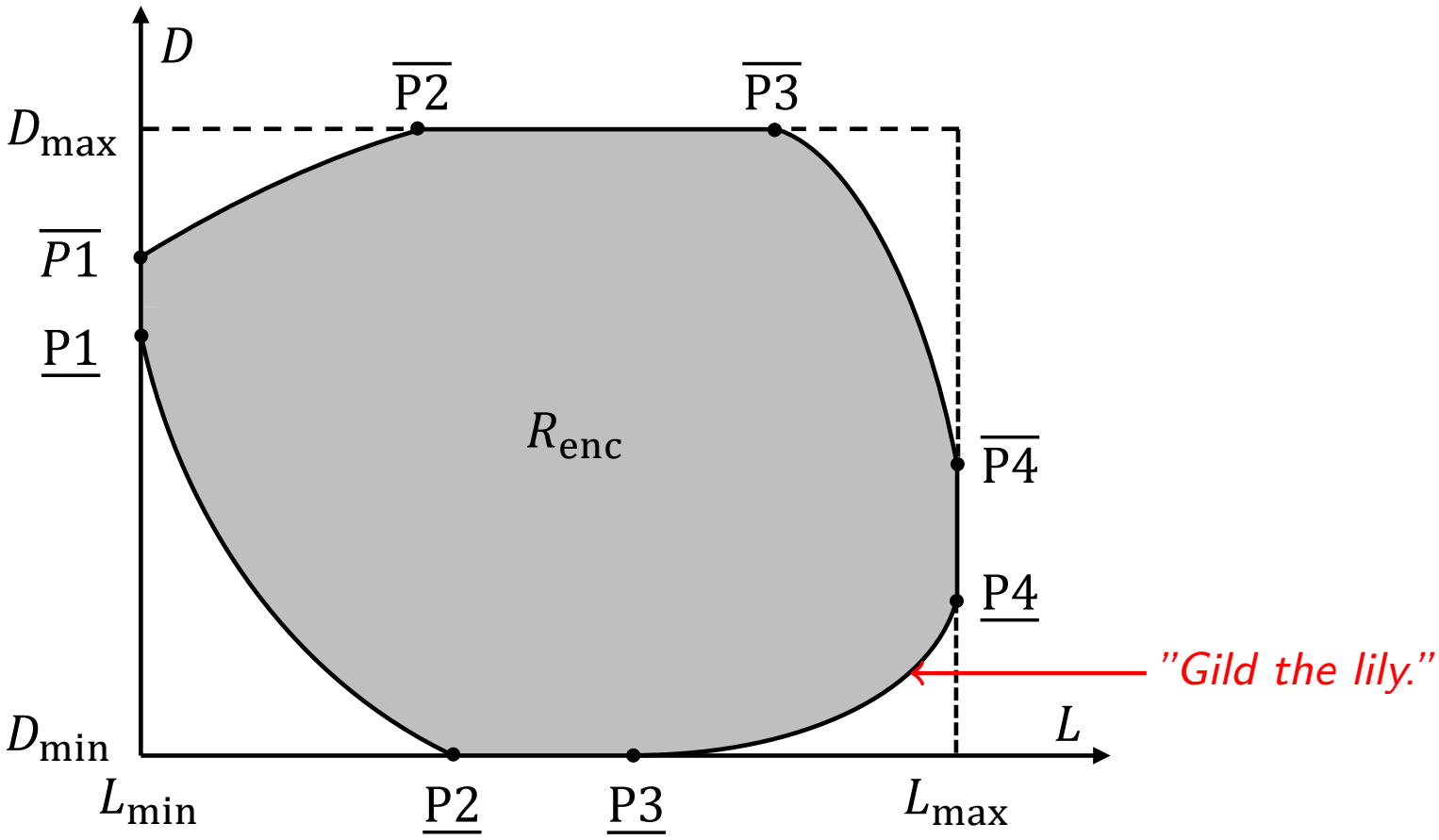


Figure: Semantic rate-distortion region (encoding).

Semantic Encoding (Decoding Scheme is Fixed)

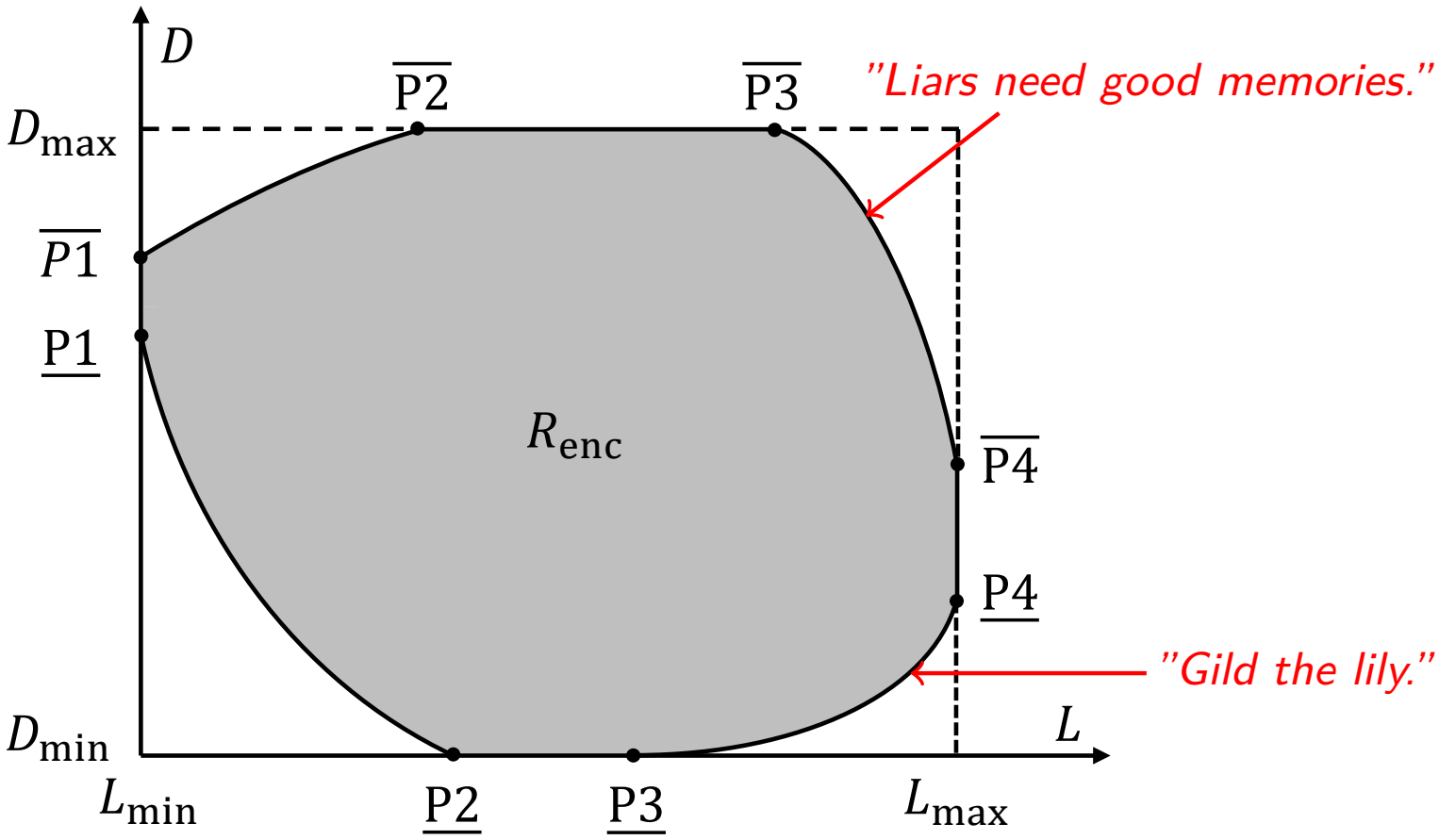


Figure: Semantic rate-distortion region (encoding).

Semantic Decoding (Encoding Scheme is Fixed)

The semantic rate-distortion region (decoding) R_{dec} is the vertical line between

$$\left(L_{\boldsymbol{P}}, D_{\boldsymbol{P},\widetilde{\boldsymbol{\Delta}}_{n_1',n_2',\ldots,n_M'}}\right) \text{ and } \left(L_{\boldsymbol{P}}, D_{\boldsymbol{P},\widetilde{\boldsymbol{\Delta}}_{n_1'',n_2'',\ldots,n_M''}}\right).$$

Semantic Decoding (Encoding Scheme is Fixed)

The semantic rate-distortion region (decoding) R_{dec} is the vertical line between

$$\left(L_{\boldsymbol{P}}, D_{\boldsymbol{P},\widetilde{\boldsymbol{\Delta}}_{n_1',n_2',\ldots,n_M'}}\right) \text{ and } \left(L_{\boldsymbol{P}}, D_{\boldsymbol{P},\widetilde{\boldsymbol{\Delta}}_{n_1'',n_2'',\ldots,n_M''}}\right).$$

- The receiver may not have accurate prior information on the distribution of the meanings $p(w)$.

Semantic Decoding (Encoding Scheme is Fixed)

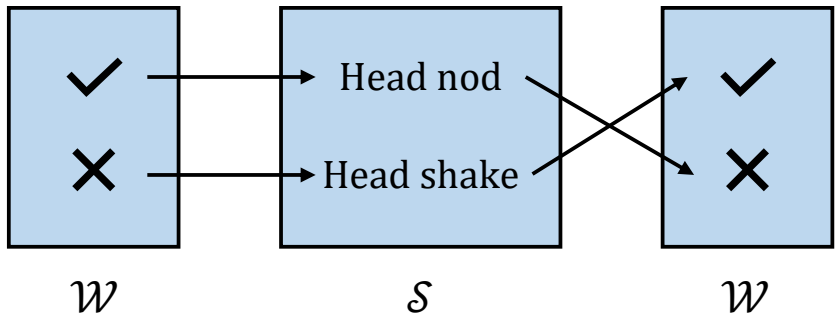
The semantic rate-distortion region (decoding) R_{dec} is the vertical line between

$$\left(L_{\boldsymbol{P}}, D_{\boldsymbol{P},\widetilde{\boldsymbol{\Delta}}_{n_1',n_2',\ldots,n_M'}}\right) \text{ and } \left(L_{\boldsymbol{P}}, D_{\boldsymbol{P},\widetilde{\boldsymbol{\Delta}}_{n_1'',n_2'',\ldots,n_M''}}\right).$$

- The receiver may not have accurate prior information on the distribution of the meanings $p(w)$.
- Semantic decoding achieves the optimal distortion if and only if for any message s

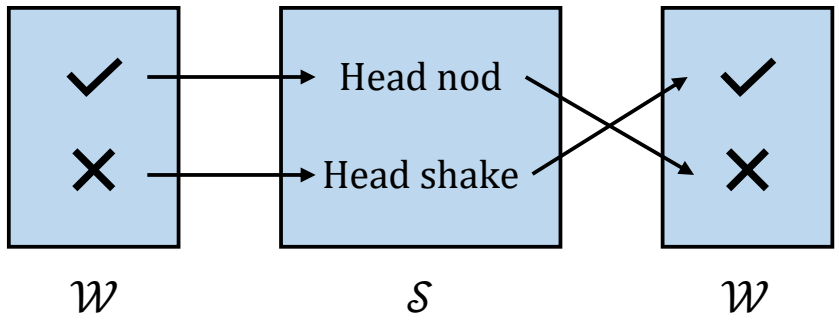
$$\arg\max_w q(w)p(\hat{s}|w) \subseteq \arg\max_w p(w)p(\hat{s}|w).$$

Combined semantic encoding and decoding



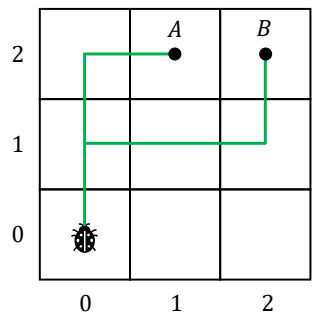
- Maggie's Gift: both semantic encoding and semantic decoding is better than CSED.

Combined semantic encoding and decoding

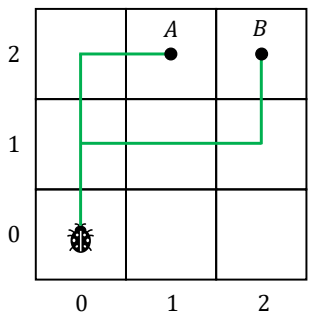


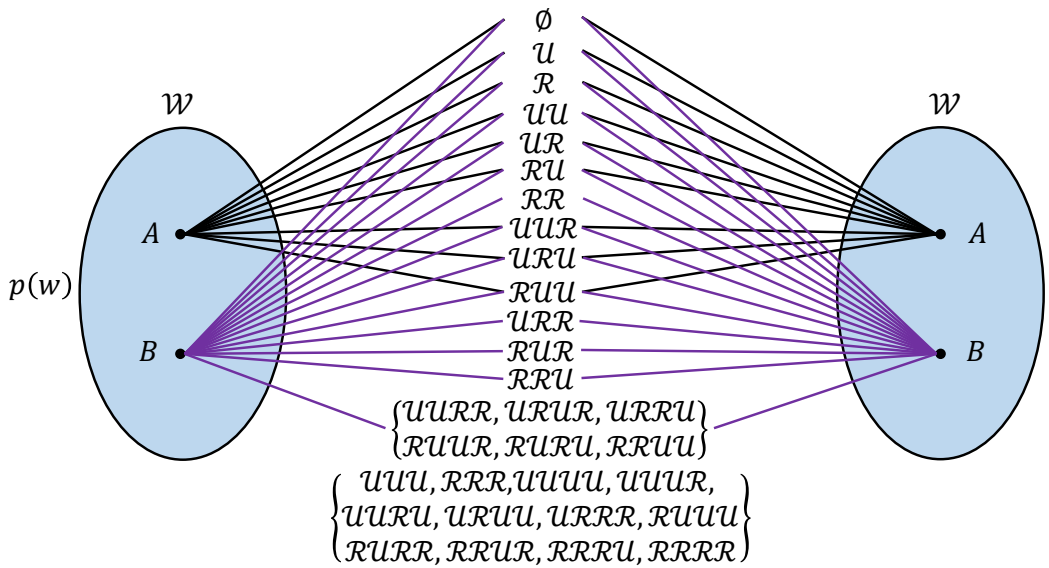
- Maggie's Gift: both semantic encoding and semantic decoding is better than CSED.
- What if the transmitter is clever enough and predicts that the receiver would improve its interpretation?

A bug walking in the grid world.



A bug walking in the grid world.





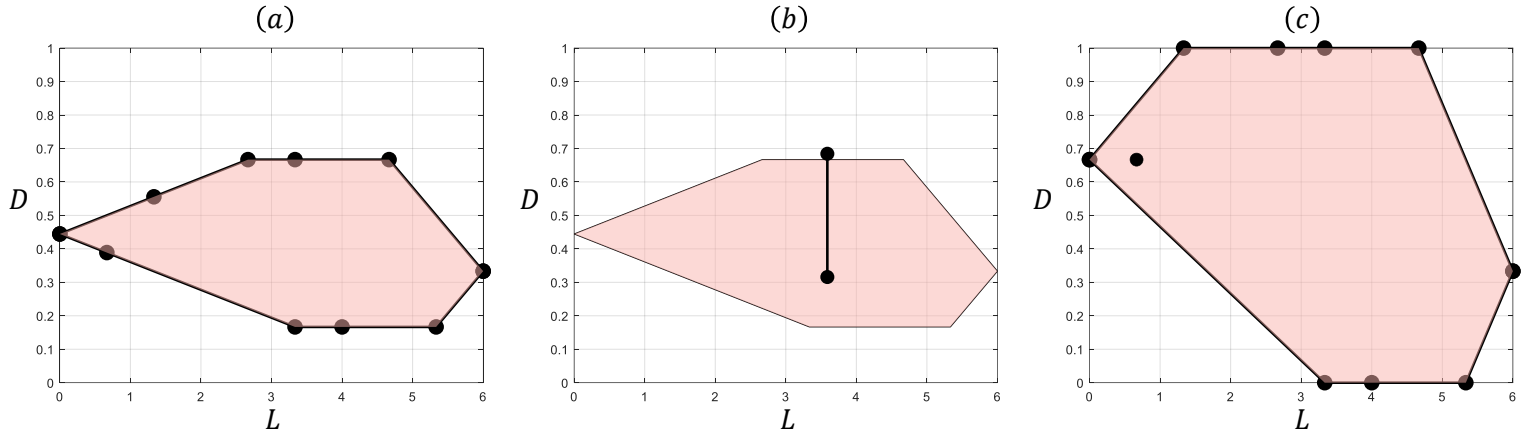
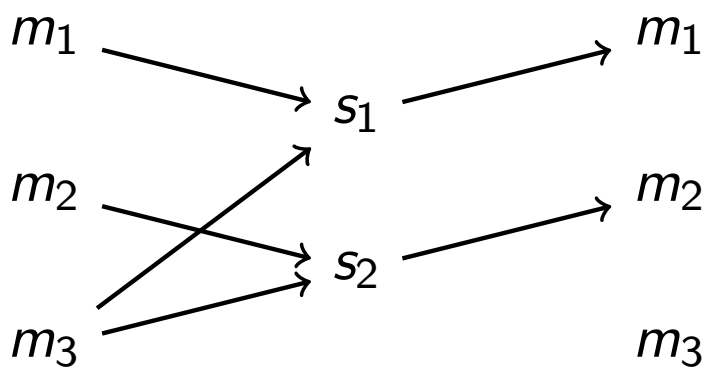


Figure: (a) semantic encoding R_{enc}, (b) semantic decoding R_{dec}, (c) CSED R_{csed}.

Future Work

Multiple Transmission

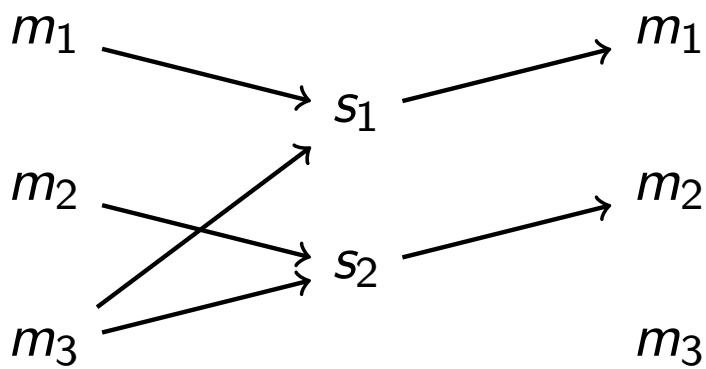
- Transmit a single meaning
  - semantic encoding: repetition coding or different encoding schemes across multiple channel uses?
  - semantic decoding: how to jointly decode the intended meaning (e.g., the majority vote, maximum ratio combing, etc.)?



Future Work

Multiple Transmission

- Transmit a single meaning
  - semantic encoding: repetition coding or different encoding schemes across multiple channel uses?
  - semantic decoding: how to jointly decode the intended meaning (e.g., the majority vote, maximum ratio combing, etc.)?

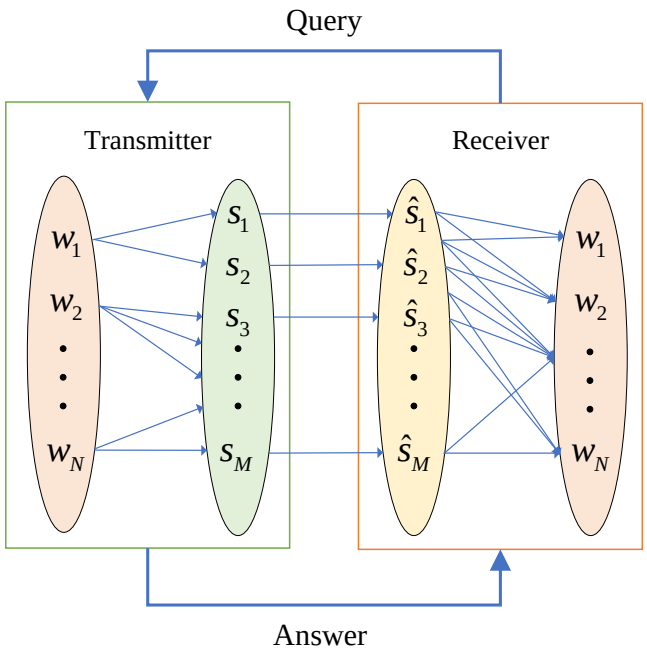


- After multiple times of transmission, the receiver can estimate $p(\hat{s})$ of the received symbols, and hence $p(s)$ and $p(w)$.

Future Work

Interaction-based transmission

- Feedback-aided transmission.
- Semantic communication with no language agreement.
- Effective and goal-oriented communication.



Future Work

- Network semantic communication
  - broadcast channels
  - multiple-access channels
  - wiretap channels
  - relay channels
  - mesh networks
- Beyond human language
  - Interacting with generative models like ChatGPT.
  - simulating an individual's painting style.
  - Predicting earthquakes based on seismic wave properties.
  - Creating anthropomorphic robots.

One More Thing...

Is a channel with memory considered a semantic channel?

My Thesis: Parallel Network

The output of any channel in the network does not connect to any channel.

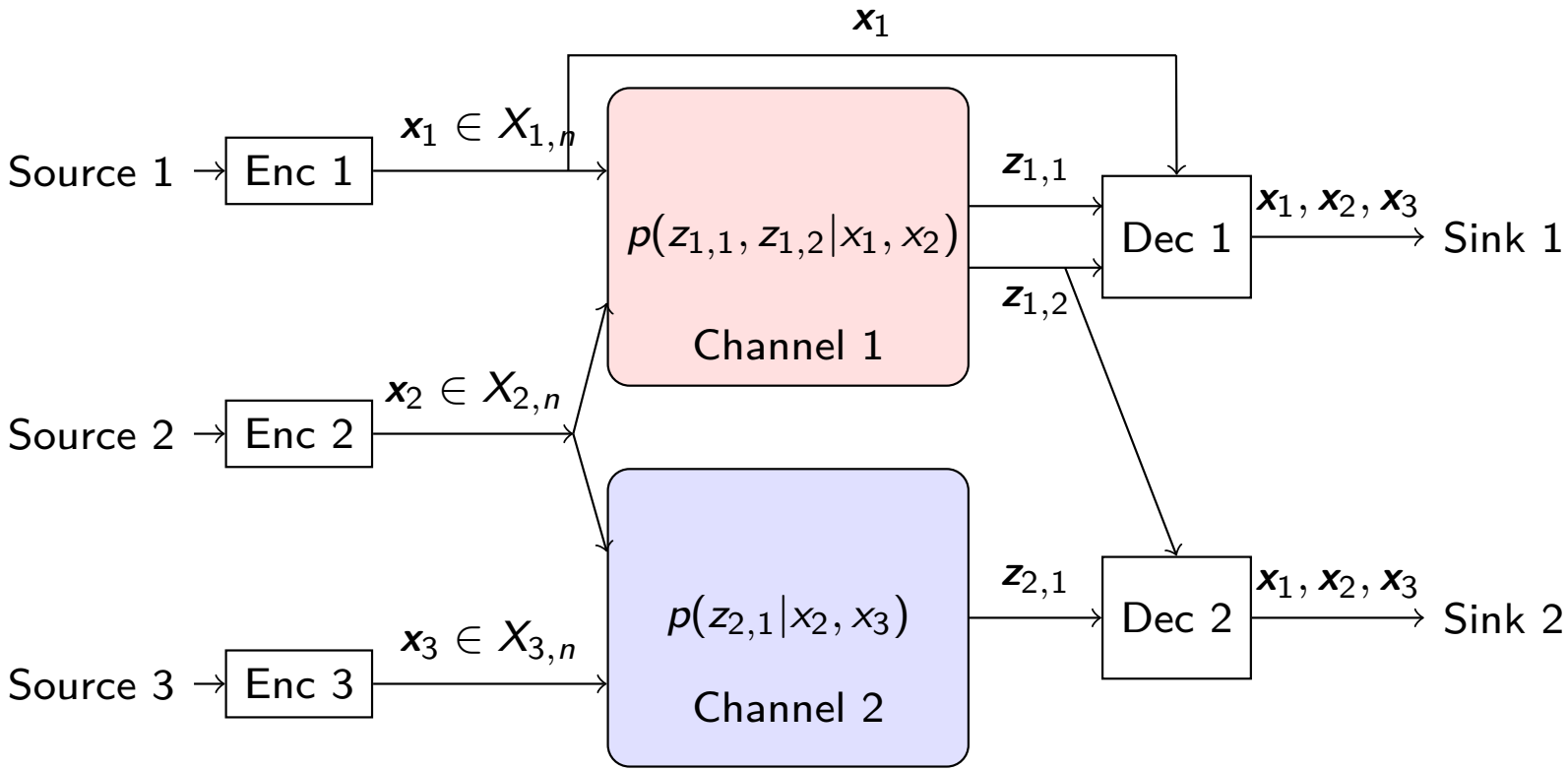


Figure: An Example of the Parallel Network.

$$\boldsymbol{G} = \{G^{DEC_1}, G^{DEC_2}\}$$

My Thesis: Parallel Network

The output of any channel in the network does not connect to any channel.

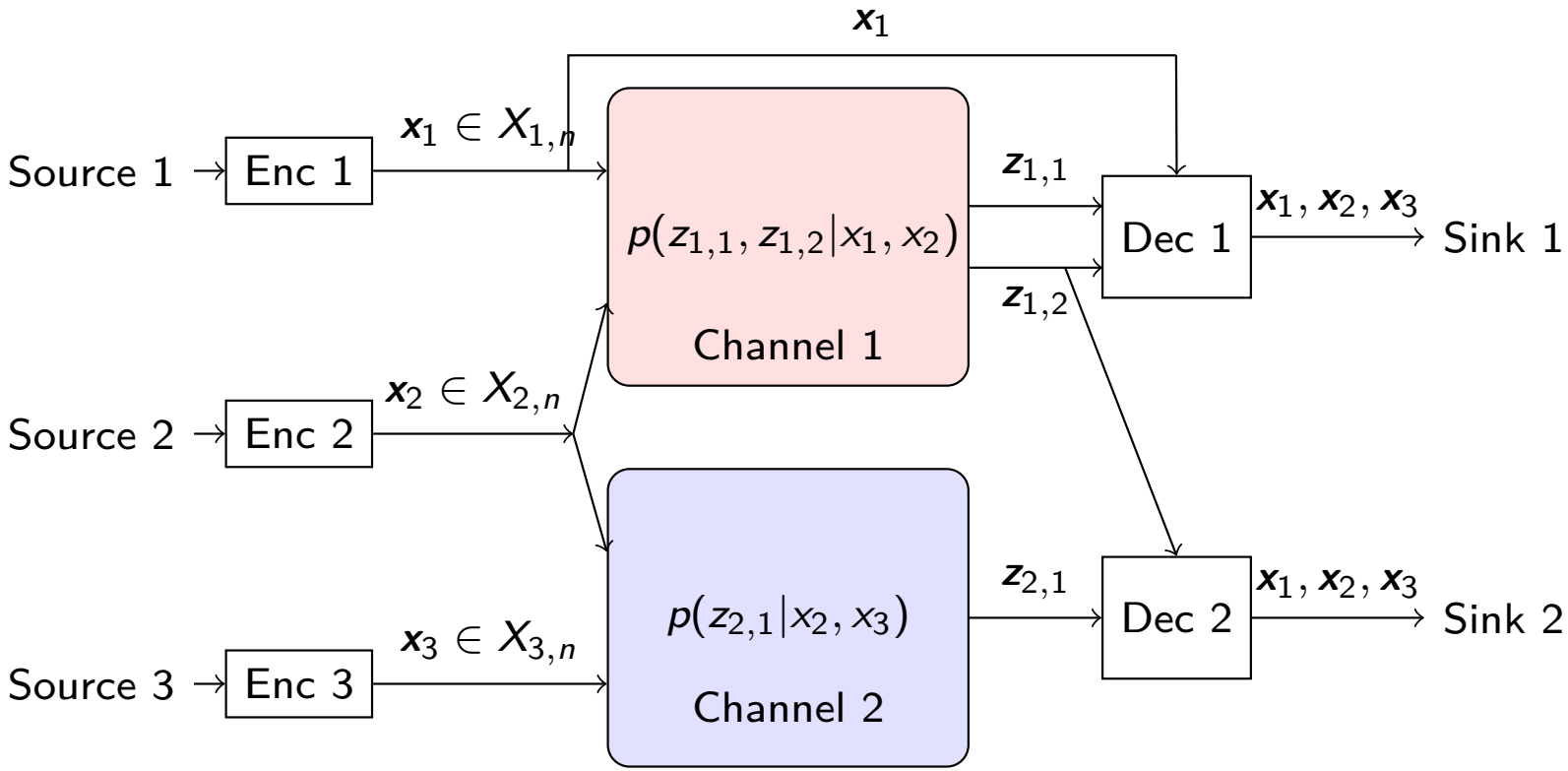


Figure: An Example of the Parallel Network.

$\boldsymbol{G} = \{G^{DEC_1}, G^{DEC_2}\}$

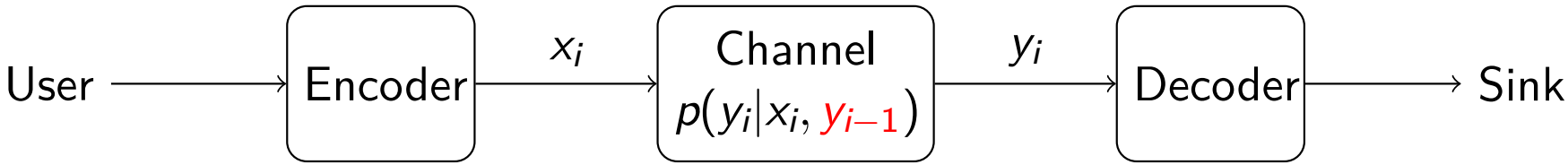
What if the output of a channel connects to its own input?

Chemical Residual Channel[2]

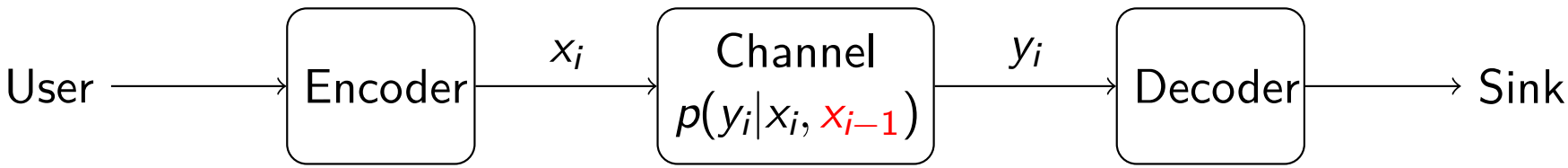
The previous experiments will affect what experiments are later run.

[2] Qi Cao and Qiaoqiao Zhou. "Zero-Error Capacity of the Chemical Residual Channel". In: *IEEE Transactions on Information Theory* 70.2 (2024), pp. 854–864.

Chemical Residual Channel & Channel with Memory



Chemical Residual Channel (Channel with Output Memory)



Channel with Input Memory[3][4][5][6]

[3] R. Ahlswede, N. Cai, and Z. Zhang. "Zero-error capacity for models with memory and the enlightened dictator channel". In: *IEEE Trans. Inf. Theory* 44.3 (1998), pp. 1250–1252. ISSN: 0018-9448.

[4] G. Cohen, E. Fachini, and J. Körner. "Zero-Error Capacity of Binary Channels With Memory". In: *IEEE Trans. Inf. Theory* 62.1 (2016), pp. 3–7. ISSN: 0018-9448.

[5] Qi Cao et al. "On Zero-Error Capacity of Binary Channels With One Memory". In: *IEEE Transactions on Information Theory* 64.10 (2018), pp. 6771–6778.

[6] Qi Cao, Qi Chen, and Baoming Bai. "On Zero-Error Capacity of Graphs With One Edge". In: *IEEE Transactions on Information Theory* 71.5 (2025), pp. 3350–3359.

Semantic Channel[7][8]

$$
\begin{aligned}
&\Pr(y_t|x_t, x_{t-1}, x_{t-2}, \ldots, x_1, y_{t-1}, y_{t-2}, \ldots, y_1) \\
=&\Pr(y_t|x_t, x_{t-1}, x_{t-2}, \ldots, x_{t-k_1+1}, y_{t-1}, y_{t-2}, \ldots, y_{t-k_2+1}) \\
=&\begin{cases}
\frac{1}{2}, & \text{if (a) } x_t \neq x_{t-1} = x_{t-2} = x_{t-k_1+1} \text{ for } t \geq k_1, \\
\frac{1}{2}, & \text{if (b) } x_t \neq y_{t-1} = y_{t-2} = \ldots = y_{t-k_2+1},\ t \geq k_2, \\
1, & \text{if } y_t = x_t \text{ and not (a), not (b),}
\end{cases}
\end{aligned}
$$

[7] R. Ahlswede, N. Cai, and Z. Zhang. "Zero-error capacity for models with memory and the enlightened dictator channel". In: *IEEE Trans. Inf. Theory* 44.3 (1998), pp. 1250–1252. ISSN: 0018-9448.

[8] Qi Cao, Yulin Shao, and Shangwei Ge. "On the Zero-Error Capacity of Semantic Channels With Input and Output Memories". In: *IEEE Wireless Communications Letters* 14.3 (2025), pp. 896–900.

Semantic Channel[7][8]

$$
\begin{aligned}
&\Pr(y_t|x_t, x_{t-1}, x_{t-2}, \ldots, x_1, y_{t-1}, y_{t-2}, \ldots, y_1) \\
=&\Pr(y_t|x_t, x_{t-1}, x_{t-2}, \ldots, x_{t-k_1+1}, y_{t-1}, y_{t-2}, \ldots, y_{t-k_2+1}) \\
=&\begin{cases}
\frac{1}{2}, & \text{if (a) } x_t \neq x_{t-1} = x_{t-2} = x_{t-k_1+1} \text{ for } t \geq k_1, \\
\frac{1}{2}, & \text{if (b) } x_t \neq y_{t-1} = y_{t-2} = \ldots = y_{t-k_2+1},\ t \geq k_2, \\
1, & \text{if } y_t = x_t \text{ and not (a), not (b),}
\end{cases}
\end{aligned}
$$

Zero-error capacity is

- If $k_2 \in \{2, 3\}$ or $k_1 \geq k_2 \geq 4$, then $C(M_{k_1,k_2}) = \log \omega_{k_2}$.
- If $k_1 = 2$ and $k_2 > 3$, then $0 \leq C(M_{k_1,k_2}) \leq 1/2$.
- If $k_2 > k_1 \geq 3$, then $\log \omega_{k_1} \leq C(M_{k_1,k_2}) \leq \log \lambda_{k_1}$.

---

[7] R. Ahlswede, N. Cai, and Z. Zhang. "Zero-error capacity for models with memory and the enlightened dictator channel". In: *IEEE Trans. Inf. Theory* 44.3 (1998), pp. 1250–1252. ISSN: 0018-9448.

[8] Qi Cao, Yulin Shao, and Shangwei Ge. "On the Zero-Error Capacity of Semantic Channels With Input and Output Memories". In: *IEEE Wireless Communications Letters* 14.3 (2025), pp. 896–900.

Reference

[1] Yulin Shao, Qi Cao, and Deniz Gündüz. "A Theory of Semantic Communication". In: *IEEE Transactions on Mobile Computing* 23.12 (2024), pp. 12211–12228.

[2] Qi Cao and Qiaoqiao Zhou. "Zero-Error Capacity of the Chemical Residual Channel". In: *IEEE Transactions on Information Theory* 70.2 (2024), pp. 854–864.

[3] R. Ahlswede, N. Cai, and Z. Zhang. "Zero-error capacity for models with memory and the enlightened dictator channel". In: *IEEE Trans. Inf. Theory* 44.3 (1998), pp. 1250–1252. ISSN: 0018-9448.

[4] G. Cohen, E. Fachini, and J. Körner. "Zero-Error Capacity of Binary Channels With Memory". In: *IEEE Trans. Inf. Theory* 62.1 (2016), pp. 3–7. ISSN: 0018-9448.

[5] Qi Cao et al. "On Zero-Error Capacity of Binary Channels With One Memory". In: *IEEE Transactions on Information Theory* 64.10 (2018), pp. 6771–6778.

[6] Qi Cao, Qi Chen, and Baoming Bai. "On Zero-Error Capacity of Graphs With One Edge". In: *IEEE Transactions on Information Theory* 71.5 (2025), pp. 3350–3359.

[7] R. Ahlswede, N. Cai, and Z. Zhang. "Zero-error capacity for models with memory and the enlightened dictator channel". In: *IEEE Trans. Inf. Theory* 44.3 (1998), pp. 1250–1252. ISSN: 0018-9448.

[8] Qi Cao, Yulin Shao, and Shangwei Ge. "On the Zero-Error Capacity of Semantic Channels With Input and Output Memories". In: *IEEE Wireless Communications Letters* 14.3 (2025), pp. 896–900.

Thank you!

"The most incomprehensible thing about the world is that it is at all comprehensible."

- Albert Einstein