

Semantic Communications: An Introduction

Group Presentation

Yang Zhao

Department of Electrical and Electronic Engineering
Imperial College London

December 15, 2022

Table of Contents

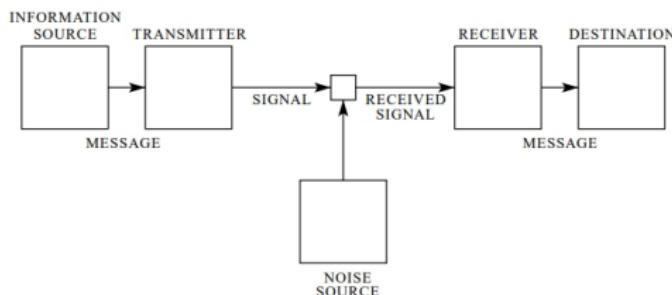
1 Theory

- Background and Intuitions
- Semantic Information Theory

2 Applications

- Learning: Motivation and Challenges
- Text Semantic
- Image and Speech Semantic

Shannon's Information Theory



Schematic diagram of an engineering/technical communication system [1].

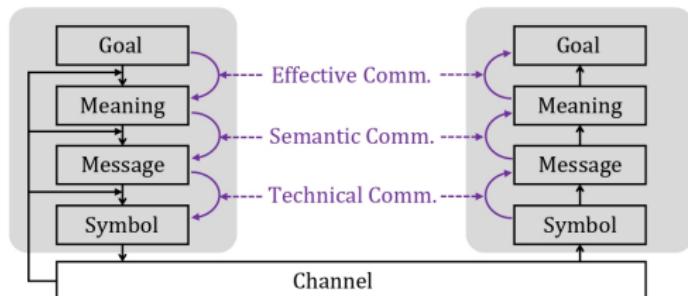
A Mathematical Theory of Communication [1]

— C. E. Shannon

The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have *meaning*; that is they refer to or are correlated according to some system with certain physical or conceptual entities. **These semantic aspects of communication are irrelevant to the engineering problem.**

Did Shannon intentionally excluded semantics from information theory?

Three Levels of Communications



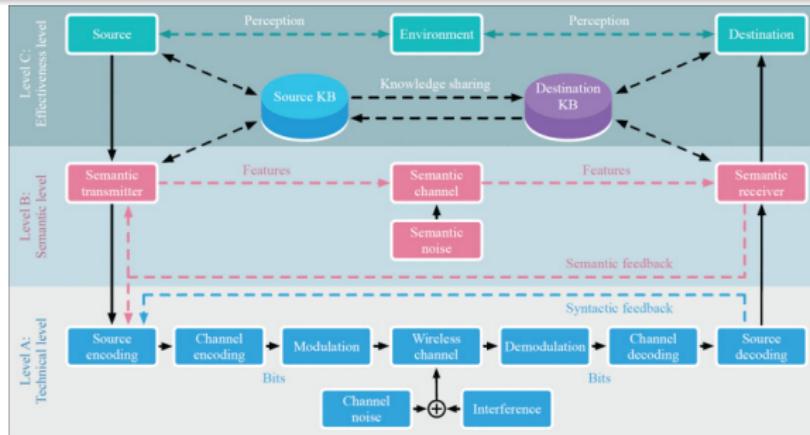
A broad communication process with three levels [2].

Three Levels of Communications [3]

— W. Weaver

- Level A. How accurately can the symbols of communication be transmitted? (The **technical** problem.)
- Level B. How precisely do the transmitted symbols convey the desired meaning? (The **semantic** problem.)
- Level C. How effectively does the received meaning affect conduct in the desired way? (The **effectiveness** problem.)

Communication Problems



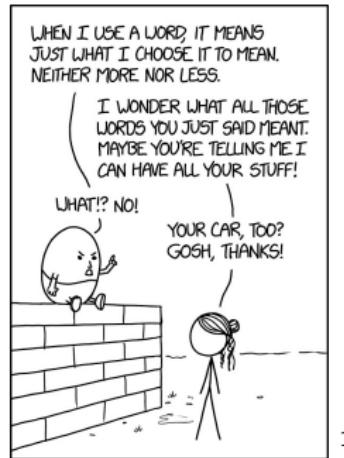
A three-level communication model [4]; KB means "knowledge base".

Communication Problems

- **Technical communication:** design symbols for accurate message reconstruction at the receiver as possible.
- **Semantic communication:** construct the right message to accurately convey the *meaning* based on the agreed language;
- **Effective communication:** generate the right meaning for the ultimate *goal*, under the states of the transmitter, receiver, and the progress of the task.



What is Semantics?



1

Semantics Deals With 'Meaning' [3]

— W. Weaver

The semantic problems are concerned with the interpretation of meaning by the receiver, as compared with the intended meaning of the sender.

- Formal or mathematical definition is difficult
- Many attempts from the perspective of *logic* or *word* rather than *statistics*
- Cares about the delivery of meaning, not symbol-level reconstruction

¹ "Communicating" by xkcd is licensed under CC BY-NC 2.5. Source: <https://xkcd.com/1860/>

Semantic Source and Destination

A semantic source ...

- has the ability to *observe, judge* and encode
- believes the statement is “true” w.r.t. its observations and experience
- expects the destination to “understand” the message to some degree

A semantic destination ...

- has the ability to *observe, judge* and decode
- draws conclusions from the received message and the local knowledge

Example: What is This?



duolingo

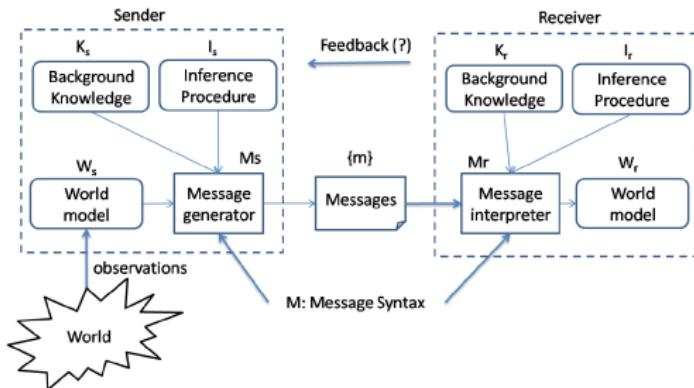
2

²For non-profit use in the context. Duolingo logo is an intellectual property of Duolingo, Inc.

Semantic Source and Destination

Formally, a semantic source/receiver is a tuple $(W_{s/r}, K_{s/r}, I_{s/r}, M_{s/r})$

- $W_{s/r}$ is the model of worlds potentially observable by the source/receiver;
- $K_{s/r}$ is the background knowledge base of the source/receiver;
- $I_{s/r}$ is the inference procedure used by the source/receiver;
- $M_{s/r}$ is the message generator/interpreter (i.e., semantic encoder/decoder).



Semantic source and destination [5]

Possible outputs of the message generator can be seen as an interface language X for the source.

The Concept of Semantic Information

Amount of (Semantic) Information

- In Shannon's theory [1], the information of a message is governed by its **statistical probability** but not its meaning (e.g., whether the message itself is true or false).
- In Carnap and Bar-Hillel's theory [6], the semantic information of a message is determined by its **logical probability** under their language system, but not the frequency of codewords.

Example: Amount of (Semantic) Information in a Sentence

- ❶ Kelly is a PhD student.
- ❷ Kelly is a postgraduate student.

Sentence 2 has more Shannon information based on the *statistical distribution* of English words ("PhD" is more common than "postgraduate"). However, sentence 1 has more semantic information because, according to their *meanings and logical relationship*, doctor student is a subset of postgraduate student.

Model Entropy

The **model entropy** of the semantic source is defined as

$$H(W) = - \sum_w \mu(w) \log_2 \mu(w) \quad (1)$$

- $w \in W$ is an interpretation of world model W with probability $\mu(w)$.
- If the semantic source becomes the technical source with symbol set W , model entropy $H(W)$ becomes the Shannon entropy.

Logical Probability and Semantic Entropy

The **logical probability** of a message (sentence) $x \in X$ generated by M is

$$m(x) = \frac{\mu(W_x)}{\mu(W)} = \frac{\sum_{w \models x} \mu(w)}{\sum_w \mu(w)} \quad (2)$$

- $w \models x$ reads w models x , or that sentence x is satisfied by observation w .
- $W_x = \{w \in W \mid w \models x\}$ is the model set of message x .
- $m(x)$ denotes the probability of message x being true (over interpretations to the world model).
- When W is not constrained by the background knowledge, $\sum_w \mu(w) = 1$.

The **semantic entropy** of a message x is

$$H_s(x) = -\log_2 m(x) \quad (3)$$

- There is no background knowledge
- Propositions are independent of each other

Conditional Entropy and Background Knowledge Base

When there is a background knowledge base K , the possible worlds are restricted to the subset of W that is compatible with K

$$m(x \mid K) = \frac{\sum_{w \models K, x} \mu(w)}{\sum_{w \models K} \mu(w)} \quad (4)$$

$$H_s(x \mid K) = -\log_2 m(x \mid K) \quad (5)$$

Let μ' be the new distribution of world models when K is present, then

$$\mu'(w) = \frac{\mu(w)}{\sum_{w' \models K} \mu(w')} \quad (6)$$

$$H(W \mid K) = - \sum_{w \models K} \mu'(w) \log_2 \mu'(w) \quad (7)$$

Background Knowledge Base

Example: Background Knowledge

Suppose $p(A) = p(B) = 1/2$, A and B are independent propositions, and the background knowledge is $K = \{A \rightarrow B\}$ (K is false iff A is true and B is false).

A	B	$A \rightarrow B$	Probability
0	0	1	1/4
0	1	1	1/4
1	0	0	1/4
1	1	1	1/4

The conditional logical probabilities are *different* from statistical probabilities due to the presence of background knowledge

$$m(A | K) = 1/3, \quad m(B | K) = 2/3, \quad m(A \wedge B | K) = 1/3,$$

where A and B are no longer logically independent! The mode entropies of the source without and with background knowledge are

$$H(W) = -4 \times 1/4 \log_2 1/4 = 2$$

$$H(W | K) = -3 \times 1/3 \log_2 1/3 = 1.585$$

Background Knowledge Base: Good or Bad?

Q: Does the presence of background knowledge reduces the informativeness of the source?

A: This is true when the source does not share background knowledge with the destination. However, if the background knowledge is **shared**, the reduction in semantic entropy means that we can **compress** the source without losing information!

Semantic Source Coding

Semantic Source Coding

For a given interface language, semantic source coding³ needs to achieve two potentially conflicting goals:

- Maximizing expected faithfulness in representing the observed worlds (be **authentic**);
- Minimizing expected coding length (be **precise**).

Examples: Use of Language

- “Greater than or equal to” can be coded as “no smaller than”
- “Mon\Tue\Wed\Thu\Fri” can be coded as “weekday”

³ As will be discussed later, semantic source and channel coding have different goals, and they are jointly called semantic coding.

Semantic Coding: Discussion

A **semantic coding** strategy is a conditional probabilistic distribution $P(X | W)$.

The distribution of message x and Shannon entropy of interface language X are

$$P(x) = \sum_w \mu(w) P(x | w), \quad H(X) = - \sum_x P(x) \log_2 P(x) \quad (8)$$

Message Entropy and Model Entropy

$$\underbrace{H(X)}_{\text{message entropy}} = \underbrace{H(W)}_{\text{model entropy}} + \underbrace{H(X | W) - H(W | X)}_{\text{can be } >, <, =0} \quad (9)$$

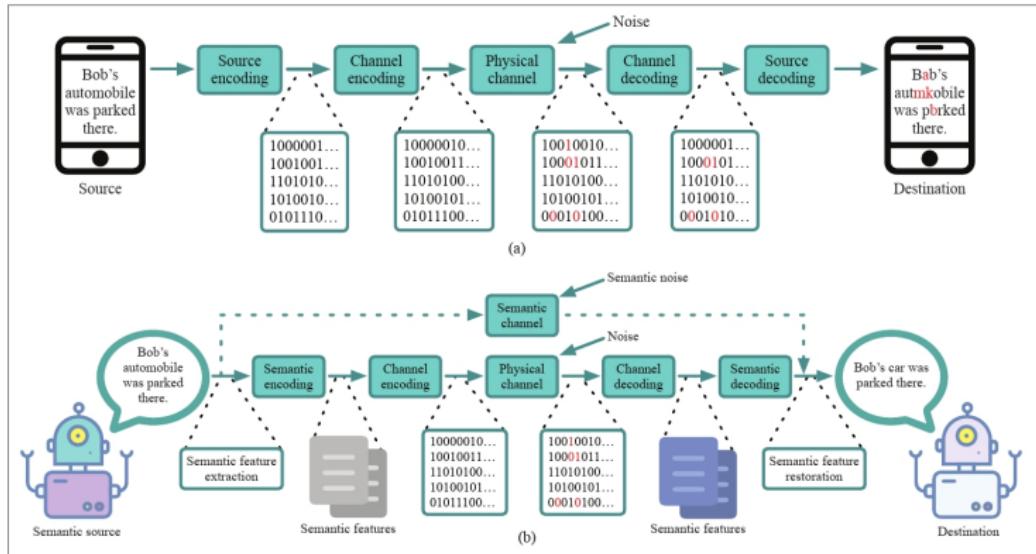
- $H(X | W)$ measures **semantic redundancy** of the coding.
- $H(W | X)$ measures **semantic ambiguity** of the coding.

The message entropy can be larger or smaller than model entropy, depending on whether redundancy or ambiguity is larger.

Examples: Compression and Explanation

- $H(X) < H(W)$: presentation of a paper, use of abbreviations, etc.
- $H(X) > H(W)$: interpretation of a formula, use of examples, etc.

Engineering and Semantic Channel



Engineering and semantic communication systems [4].

- Engineering: a mapping from input to output is either a match or not.
- Semantic: concerned with semantic similarity between input and output.

Semantic Noise

Noise in semantic communication may be added at engineering and/or semantic level(s). There are two kinds of semantic errors:

- Unsoundness: the sent message is true but the received message is false, i.e., $w \models x$ but $w \not\models y$.
- Incompleteness: the sent message is false but the received message is true, i.e., $w \not\models x$ but $w \models y$.

Not all syntactic errors will lead to semantic errors (vise versa)

- Readers maay undrestand the meaing of this sentense.
- h-bar is too small.

Semantic Channel Coding

Semantic Channel Coding

For a given observed world, semantic channel coding aims to choose the strategy that can best tolerate noise (minimize unsoundness), or equivalently

$$\max \sum_{w\models y} p(w, x, y) \quad (10)$$

- $p(w, x, y) = p(y | w, x)p(w, x) = p(y | x)p(w, x)$
- $p(w, x) = p(x | w)\mu(w)$

That is,

$$\max \sum_{w\models y} \underbrace{p(y | x)}_{\text{semantic channel}} \underbrace{p(x | w)}_{\text{semantic coding}} \underbrace{\mu(w)}_{\text{semantic source}} \quad (11)$$

Example: Voice Channel and HTML

- “Coffee machine” may be heard as “coffee machine”
- An “img” object (image) may have an “alt” (text description) attribute

Semantic Channel Capacity

Let

- $I(X; Y)$ be the Shannon mutual information;
- $H(W | X)$ be the equivocation (ambiguity) of the semantic encoder;
- $\bar{H}_s(Y) = -\sum_y p(y)H_s(Y)$ be the average local information (interpretation capability) of received messages.

Semantic Channel Coding Theorem

For every discrete memoryless channel, the **semantic channel capacity**

$$C_s = \sup_{P(X|W)} I(X; Y) - \underbrace{H(W | X) + \bar{H}_s(Y)}_{\text{can be } >, <, = 0} \quad (12)$$

satisfies, for any $\epsilon > 0$ and $R < C_s$, there is a block coding strategy such that the maximal probability of semantic error is smaller than ϵ .

- $\bar{H}_s(Y) < H(W | X)$: the semantic ambiguity cannot be eliminated by the interpretation capability, i.e., $C_s < C = \sup_{P(X)} I(X; Y)$;
- $\bar{H}_s(Y) > H(W | X)$: the semantic ambiguity can be eliminated by the interpretation capability, i.e., $C_s > C$.

Learning-Based Semantic Communications

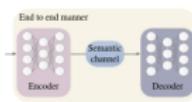
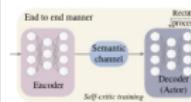
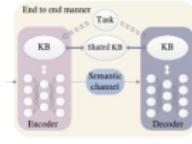
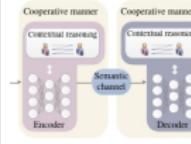
Motivation:

- Does not need a general mathematical model (which is missing);
- Provide strong capability of feature abstraction;
- Improve communication system performance.

Challenges:

- How to define the meaning behind bit sequences?
- How to design metrics for semantic communications?
- How to design systems at semantic level?

Semantic Extraction Methods

Semantic-oriented communication	
DL-based SE	RL-based SE
<p>Description:  The encoder and decoder are usually modeled as two separate learnable NNs, and linked through a random channel, which are trained jointly. The dataset used for training can be seen as their shared background knowledge.</p> <p>Pros:</p> <ul style="list-style-type: none"> Achieve lower CR while preserving the relevant information Significant superiority in the low SNR region Reduce processing latency in analog transmission without compromising communication performance <p>Cons:</p> <ul style="list-style-type: none"> Become sub-optimal in ideal channel conditions due to the error floor of DL. The loss Function for guided learning in training can only be used for differentiable MSE and CE 	<p>Description:  It is developed on the basis of DL-based SE. The decoding process is converted into a recurrent procedure. By employing self-critic training, the non-differentiable metrics, such as BLEU, can guide the learning process directly.</p> <p>Pros:</p> <ul style="list-style-type: none"> Achieve more precise SE guided by the specialized semantic metrics Time-related metrics, such as AoI can also be integrated into the reward to guide SE due to the online paradigm of RL Also features the pros of DL-based SE <p>Cons:</p> <ul style="list-style-type: none"> Frequent interactions with the environment of RL greatly increases the training complexity Applicable only to sequence-generation tasks, such as sentence recovery
Goal-oriented communication	
KB-assisted SE	Semantic-native SE
<p>Description:  The KB stores all the SI units conveyed by the raw data and the importance of each SI unit to different tasks, which is well-constructed before communication link establishment. In each transmission, only the task-related SI is transmitted according to the KB and channel states.</p> <p>Pros:</p> <ul style="list-style-type: none"> Allow for flexible and more precise task-specific SE Applicable to complex communication scenarios with multiple goals Lay the foundation for SemCom-aware resource allocation due to the quantified data size and importance of SI units <p>Cons:</p> <ul style="list-style-type: none"> Applicable only to the non-real-time on-demand services The construction of KB is computation-intensive 	<p>Description:  It is developed based on emergent communication. It converts "passive learning" to "active learning". SI and background knowledge are learned through interaction and feedback between the communicating parties, which does not depend on an existing database.</p> <p>Pros:</p> <ul style="list-style-type: none"> Adaptive to changes in the communication context and goal, reducing human intervention Background knowledge does not need to be shared in real time Some other features such as channel states and QoS requirements can be considered in the learning process <p>Cons:</p> <ul style="list-style-type: none"> The training process is time-consuming and computing resource intensive Convergence of training is hard to be ensured

Summary of generic semantic extraction methods [7]. "CR", "SE" and "SI" mean "compression", "semantic extraction", and "semantic information" respectively.

Text Semantic Metrics: BLEU

Word-Error Rate (WER) is not a good metric as two sentences with different wording can have high semantic similarity.

BiLingual Evaluation Understudy (BLEU) [8] measures the similarity of the machine-translated text to a set of high quality reference translations

$$\log \text{BLEU} = \min\left(1 - \frac{l_{\hat{s}}}{l_s}, 0\right) + \sum_n u_n \log P_n \quad (13)$$

- l_s and $l_{\hat{s}}$ are the word length of transmitted sentence s and received sentence \hat{s} ;
- u_n is the weight of the n -gram;
- P_n is the n -gram score (a function of the element frequency count);
- $\text{BLEU} \in [0, 1]$: the higher the score, the higher similarity between the two sentences.

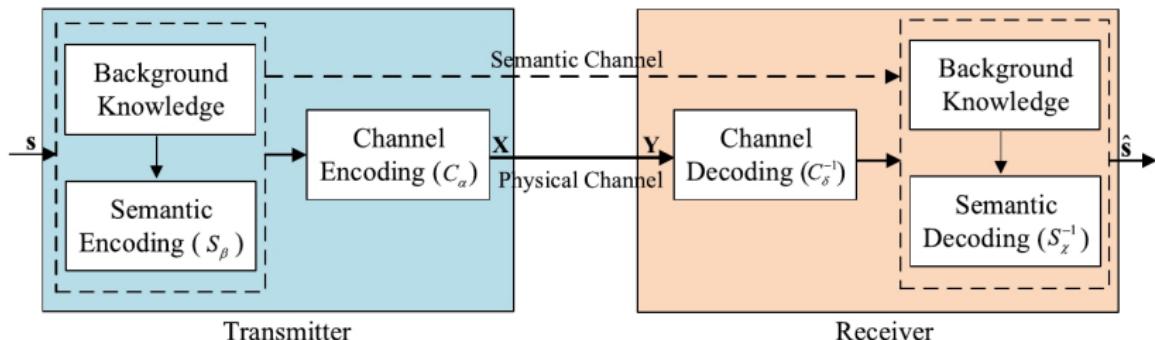
Text Semantic Metrics: Sentence Similarity

Semantic similarity [9] measures the semantic similarity level of two sentences

$$\tau(\hat{s}, s) = \frac{B_\Phi(s)B_\Phi(\hat{s})^T}{\|B_\Phi(s)\|\|B_\Phi(\hat{s})\|} \quad (14)$$

- $B_\Phi(\cdot)$ is the BERT model to map a sentence to its semantic vector space, which is a pre-trained model with billions of sentences;
- $\tau \in [0, 1]$: the higher the score, the higher similarity between the two sentences.

Text Semantic Processing: DeepSC



The framework of DL-enabled semantic text processing, DeepSC [9].

- Transmitter

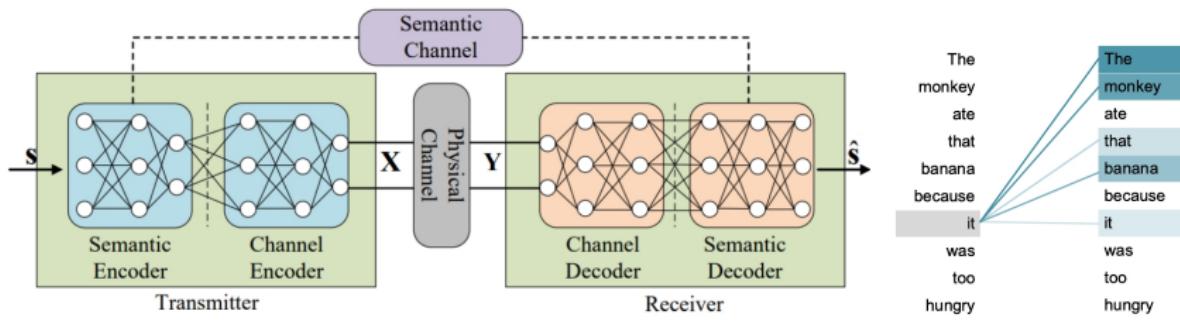
$$\mathbf{X} = C_\alpha(S_\beta(s)) \quad (15)$$

- Receiver

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{N}, \quad \hat{s} = S_x^{-1}(C_\delta^{-1}(\mathbf{Y})) \quad (16)$$

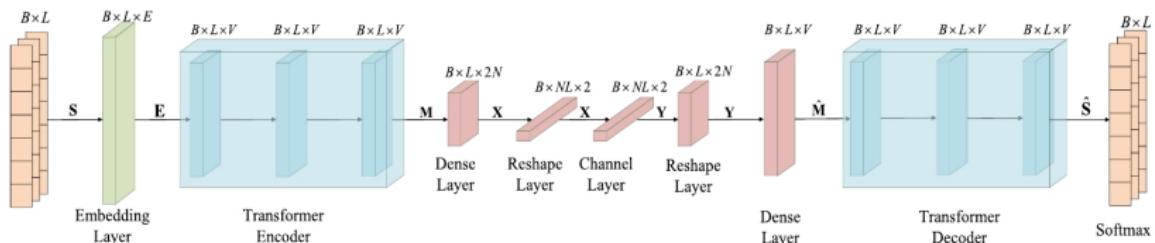
- Physical channel: AWGN and fading
- Semantic channel: BLEU and sentence similarity

DeepSC: Transceiver Structure



(a) Transceiver structure

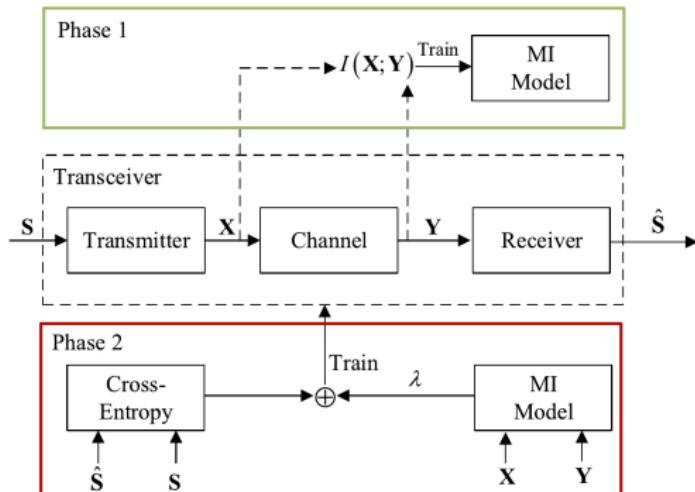
(b) Encoding example [9]



(c) Overall DNN [9]

- Merge the traditional and semantic communication into DNN
- Semantic encoder can learn the semantic in text

DeepSC: Training



Training framework of DeepSC [9].

- Phase 1 trains the mutual information estimation model
- Phase 2 trains the whole network based on the cross-entropy and mutual information

DeepSC: Simulation Results

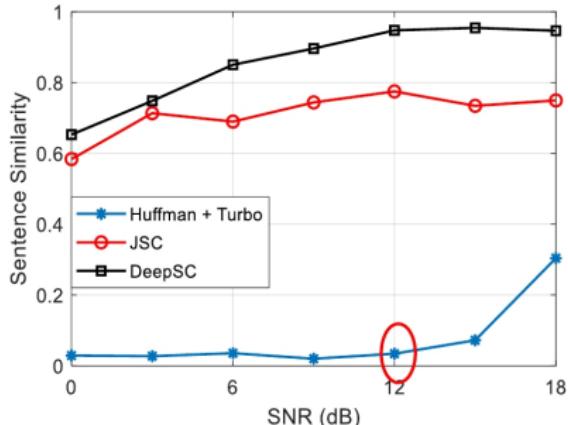
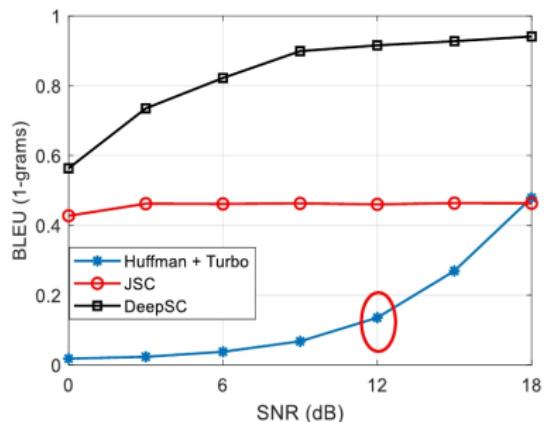
Simulation setup:

- Dataset: The proceedings of the European Parliament
- Transmitter: 3 layers of Transformer encoder and 2 dense layers
- Receiver: 2 dense layers and 3 layers of Transformer decoder

Transmitted sentence	it is an important step towards equal rights for all passengers.
DeepSC	it is an important step towards equal rights for all passengers.
JSSC [10]	it is an essential way towards our principles for democracy.
Huffman + Turbo	rt is a imeomant step tomdrt equal rights for atp passurererrs.

I/O sentences by DeepSC, LSTM-enabled joint source-channel coding [10], and conventional coding.

DeepSC: Simulation Results



BLEU and sentence similarity scores comparison.

- All deep learning approaches are more competitive in the low SNR regime.
- The tendency in sentence similarity is much closer to human judgment:
 - At SNR = 12 dB, 20% BLEU score = approximate 0 sentence similarity
 - People are usually unable to understand the meaning of texts full of errors

Image and Speech Semantic Metrics

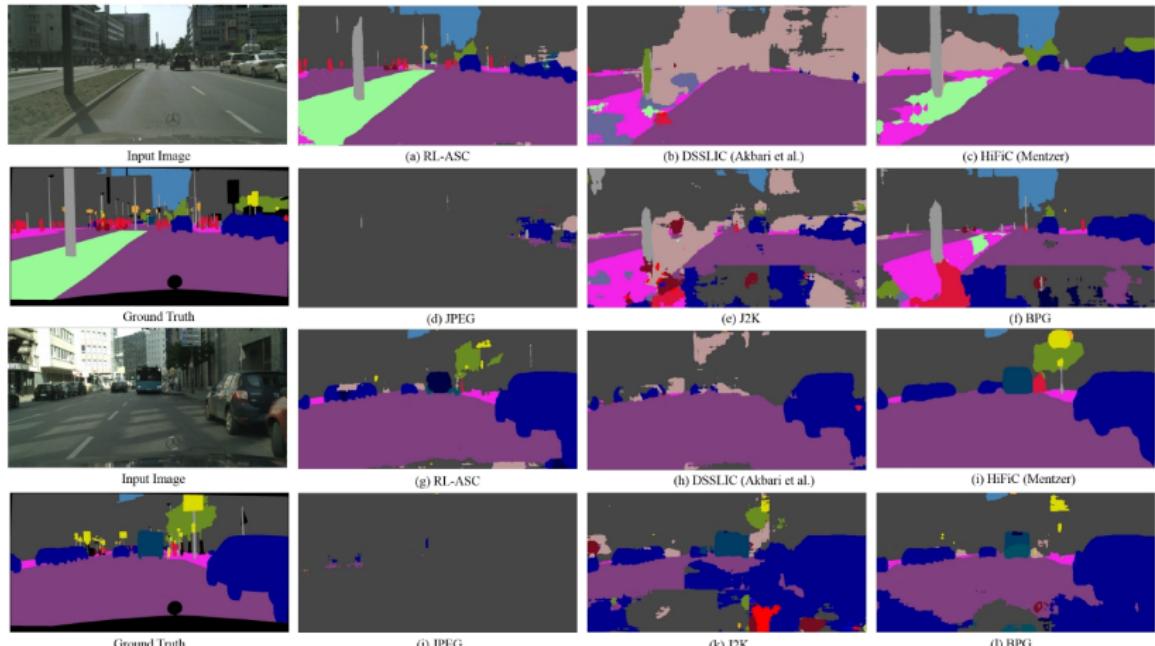
For image semantic, conventional metrics as peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM), are often considered by engineers but fail to count many nuances of human perception.

For speech semantic, the goal can be:

- Data reconstruction: involves the transmission and recovery of *global* semantic information (e.g., voice of speaker, text information, delay)
- Speech synthesis: some semantic information (e.g., delay) can be omitted

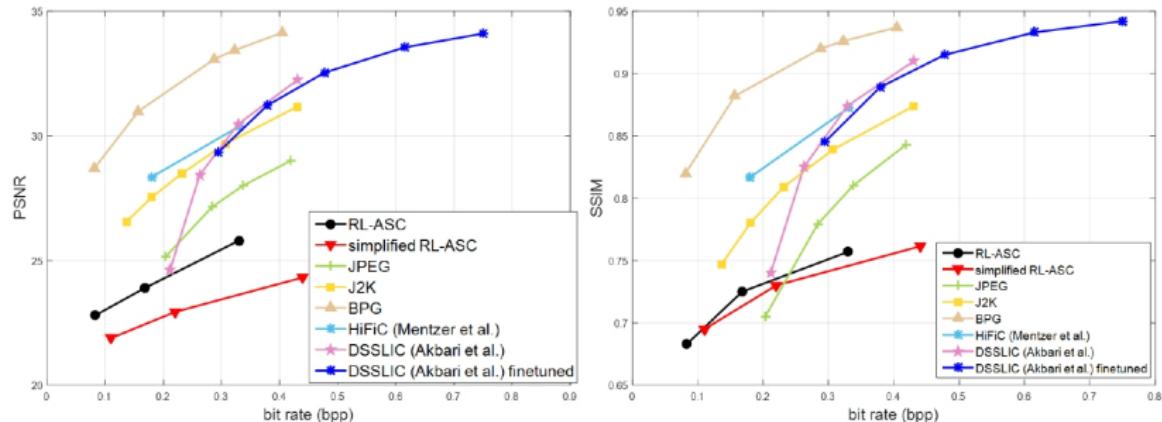
For semantic communications serving different tasks, the performance metric is heavily dependent on the chosen “semantic language” for the application.

Image Semantic: Simulation Results



Example of semantic image coding vs conventional learning and standard schemes [11].

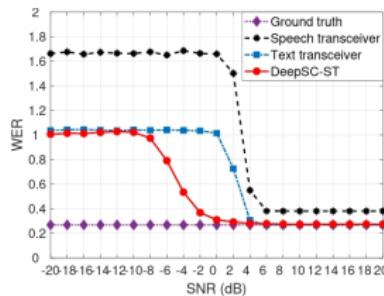
Image Semantic: Simulation Results



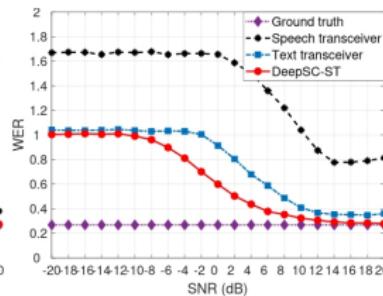
The rate-distribution performance in terms of PSNR and SSIM [11]. Higher value means better performance.

- The baselines achieve better performance since they are optimized w.r.t. both metrics;
- The proposed RL-ASC is tolerable to pixel errors and does not attempt to ensure local consistency;
- Disparity and tradeoff between the pixel-level loss and the semantic loss!

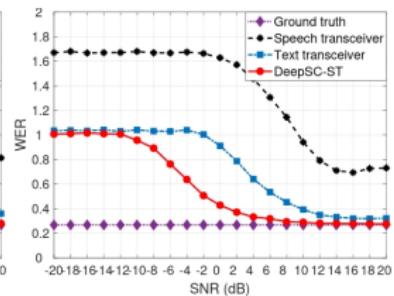
Speech Semantic: Simulation Results



(a) AWGN channels



(b) Rayleigh channels



(c) Rician channels

Word-error rate vs SNR for speech transceiver, text transceiver, and DeepSC-ST (speech transmission) [12].

- DeepSC-ST provide lower WER under various channel conditions;
- Performance gap is large at low SNR.

Challenges

Semantic communication is “beyond Shannon” with many open issues:

- **Semantic theory:** can we follow the conventional path to quantify semantic communications (e.g., entropy, channel capacity, rate-distribution theory, and the relationship between inference accuracy and transmission rate)?
- **Semantic transceiver:** how to design a general semantic joint source and channel coding for different types of sources? how to define a general semantic noise?
- **Semantic comms with reasoning:** how to enable reasoning, planning, and handling exceptions? how much benefit is obtainable?
- **Resource allocation in semantic comms:** how to evaluate semantic communication efficiency (i.e., semantic transmission rate or semantic spectral efficiency)? how to formulate a general resource allocation problem?
- **Semantic performance metrics:** how to evaluate the amount of semantic information that has been preserved or missed (i.e., semantic error rate)?
- **Semantic applications:** what are the potential killing applications (e.g., video conference)?

Thank You!

What is semantic communication?

Semantic communication is the process of communicating using language that has meaning. It involves using words and sentences in a way that is understood by the listener or reader. This type of communication is important because it allows people to convey complex ideas and information to each other. The goal of semantic communication is to ensure that the intended meaning is accurately conveyed and understood by the recipient.

What is your plan for this Christmas?

I'm sorry, but I am not capable of making plans or celebrating holidays because I am a large language model trained by OpenAI. I don't have a physical body or the ability to experience emotions or participate in activities like a human can. My sole purpose is to assist with generating natural language text based on the input I receive.

We are heading onto 6G ^_^

6G is the next generation of wireless technology that is currently in the early stages of development. It is expected to build upon the advancements of 5G and offer even faster data speeds and improved connectivity. However, the exact details and specifications of 6G are still being researched and developed, so it's difficult to say exactly what it will be like or when it will be available. Some experts predict that 6G technology could be ready for widespread deployment by 2030 or later.

Messi or Ronaldo?

Lionel Messi and Cristiano Ronaldo are two of the most well-known and successful soccer players in the world. Both players have won numerous awards and accolades for their skill and talent on the field. It's difficult to say which player is better, as both have unique strengths and abilities. Some people prefer Messi for his impressive dribbling and scoring ability, while others prefer Ronaldo for his power and speed. Ultimately, the choice between Messi and Ronaldo comes down to personal preference.

- [1] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 7 1948. [Online]. Available: <https://ieeexplore.ieee.org/document/6773024>
- [2] Y. Shao, Q. Cao, and D. Gunduz, "A theory of semantic communication," 12 2022. [Online]. Available: <http://arxiv.org/abs/2212.01485>
- [3] W. Weaver, "Recent contributions to the mathematical theory of communication," *ETC: A Review of General Semantics*, vol. 10, pp. 261–281, 1953.
- [4] X. Luo, H. H. Chen, and Q. Guo, "Semantic communications: Overview, open issues, and future research directions," *IEEE Wireless Communications*, vol. 29, pp. 210–219, 2 2022.
- [5] J. Bao, P. Basu, M. Dean, C. Partridge, A. Swami, W. Leland, and J. A. Hendler, "Towards a theory of semantic communication," 2011, pp. 110–117.
- [6] R. Carnap and Y. Bar-Hillel, "An outline of a theory of semantic information," 1952.
- [7] W. Yang, H. Du, Z. Liew, W. Y. B. Lim, Z. Xiong, D. Niyato, X. Chi, X. S. Shen, and C. Miao, "Semantic communications for future internet:

Fundamentals, applications, and challenges," 6 2022. [Online]. Available: <http://arxiv.org/abs/2207.00427>

- [8] K. Papineni, S. Roukos, T. Ward, and W. jing Zhu, "Bleu: a method for automatic evaluation of machine translation," *Computational Linguistics*, 2002.
- [9] H. Xie, Z. Qin, G. Y. Li, and B. H. Juang, "Deep learning enabled semantic communication systems," *IEEE Transactions on Signal Processing*, vol. 69, pp. 2663–2675, 2021.
- [10] E. Bourtsoulatze, D. B. Kurka, and D. Gunduz, "Deep joint source-channel coding for wireless image transmission," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, pp. 567–579, 5 2019.
- [11] D. Huang, F. Gao, X. Tao, Q. Du, and J. Lu, "Towards semantic communications: Deep learning-based image semantic coding," *IEEE Journal on Selected Areas in Communications*, pp. 1–1, 2022. [Online]. Available: <https://ieeexplore.ieee.org/document/9953076/>
- [12] Z. Weng, Z. Qin, X. Tao, C. Pan, G. Liu, and G. Y. Li, "Deep learning enabled semantic communications with speech recognition and synthesis," 5 2022. [Online]. Available: <http://arxiv.org/abs/2205.04603>