

分类号: TN92

单位代码: 10335

密 级: 无

学 号: 22060598

浙 江 大 学

硕士专业学位论文



中文论文题目: 基于强化学习和语义表征的
语义通信及网络研究

英文论文题目: Research on Semantic Communication
and Network based on Reinforcement
Learning and Semantic Representation

申请人姓名: 卢 锟

指导教师: 李荣鹏

专业学位类别: 工程硕士

专业学位领域: 电子与通信工程

所在学院: 信息与电子工程学院

论文递交日期 2023 年 1 月 4 日

基于强化学习和语义表征的

语义通信及网络研究



论文作者签名:

李锐

指导教师签名:

李荣明

论文评阅人 1:

匿名

评阅人 2:

匿名

评阅人 3:

匿名

评阅人 4:

评阅人 5:

答辩委员会主席: 朱永东 研究员 之江实验室

委员 1: 刘安 研究员 浙江大学

委员 2: 杨倩倩 研究员 浙江大学

委员 3: 张宏纲 研究员 之江实验室

委员 4:

委员 5:

答辩日期 2023 年 3 月 10 日

**Research on Semantic Communication
and Network based on Reinforcement
Learning and Semantic Representation**



Author's signature: _____

Supervisor's signature: _____

External reviewers: _____ Anonymity
_____ Anonymity
_____ Anonymity

Examining Committee Chairperson:
_____ Prof. Yongdong Zhu

Examining Committee Members:
_____ Prof. An Liu
_____ Prof. Qianqian Yang
_____ Prof. Honggang Zhang

Date of oral defence: 10th March, 2023

浙江大学研究生学位论文独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得 **浙江大学** 或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名:

李银

签字日期: 2023 年 3 月 13 日

学位论文版权使用授权书

本学位论文作者完全了解 **浙江大学** 有权保留并向国家有关部门或机构送交本论文的复印件和磁盘，允许论文被查阅和借阅。本人授权 **浙江大学** 可以将学位论文的全部或部分内容编入有关数据库进行检索和传播，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

(保密的学位论文在解密后适用本授权书)

学位论文作者签名:

李银

导师签名:

李荣明

签字日期: 2023 年 3 月 13 日

签字日期: 2023 年 3 月 13 日

致谢

动笔致谢之时，忽然间已是疫情消退、生机勃勃的景象。好一个春回大地！从 2019 年底在实验室做毕设起，与实验室的同学和老师走到现在，已是三年有余。回想起一起努力的同学和老师，心中有许多不舍。这一路来，衷心地感谢所有在生活、科研上帮助过、包容过、以及支持我的朋友们。

在研究生阶段的学习中，实验室的老师是对我影响最深、也付出最多心血的。两位老师在学术追求上的崇高理想和踏实作风深刻影响了我的科研路程，也是激励我和整个实验室追求完美、不惧困难的榜样。“做没人做过的事”、“**talk is cheap, show me the code**”这些信念也成为了我开展工作的座右铭。想来能遇到两位以真理和纯粹科研为追求的指导老师真是万分荣幸。和而不同，我想这也是我从两位导师身上获得的宝贵财富——倘若如果没有两位导师对我的包容和支持，本文所呈现的工作也将不再具有可能。

独木不成林，来自信电楼 210、220 的同学们组成了一个快乐、自由的大家庭，这种良好的团队合作和温馨的生活环境对我的成长同样起到了不可或缺的作用。实验室的同学们非常幽默，总有很多新奇的想法，在学术问题和生活上相互帮助。在科研之余看到这样温馨的大家庭想来也是一种乐事。漆琛师姐稳重认真，帮助实验室处理了许多事务，为大家能够无忧无虑开展科研工作提供了良好的土壤。华郁秀师兄、徐行师兄、吴萌师兄、张莎莎师姐、倪霏师姐、陈颖师姐、苏冰滢师姐在我入学时给予了许多建议和转变角色上的帮助。申世博师兄、李晨晗师兄、陈明昊师兄、邵燕师姐还有陈焜师兄总能解答我许多科研上的疑惑。王楚捷师姐、代安娜师姐、以及刘再生师兄在我的日常生活和求职期间提供了许多细致的建议。能够和同届的刘云涛同学、周清炀同学、朱建行同学、陈浪飞同学一同成长，互相交流也是我难忘的回忆。也要感谢鲁芝琳、温若琪、肖柏狄、汪丙炎、易幸夫、项羽铭、童思予、陈宇轩、李斯照、于小雪几位同学的陪伴，看到大家活泼可爱、热爱钻研的样子，有时也会想起当时的自己，加油吧。

另外，感谢研究生期间朝夕相处的室友：金峡钊、卢云飞和蒋德琛同学，我们有各自的爱好和方向，但在宿舍这个小家里却显得格外融洽。王珏女士在我攻读研究生期间始终陪伴在我的身边，帮我打理着许多生活琐事、排解了我的许多烦恼，是激励我努力钻研的牵挂和支撑。我的父母勤勤恳恳，陪伴我度过了许多困难，也是我能够全心投入

事业无可替代的前提。最后，感谢为本文提出宝贵意见的匿名评审专家和答辩组老师们，帮助本文在可读性和科学性上得到近一步的提升。

所有的相遇都是缘分吧。祝愿大家健康平安，生活如意！

摘要

语义通信在近些年引起了研究人员的广泛关注，并被认为是一种可能会在下一代通信中起重要作用的传输方式。作为语义通信前中期的工作之一，本研究为点对点语义通信的实施架构和异构语义网络设计提供了一定的理论以及实验经验。

在语义通信的方案设计上，现有的许多方法以交叉熵和均方误差为设计准则，对语义提取模型和传输方式展开了较多的研究。不同于以上思路，本文指出一种（由学习目标语义缺陷导致的）在文献中通常被忽略的语义-符号准确率差异。为了解决这一问题，本文提出一种以广义语义相似度为优化目标的语义通信实现方法。这一思想回答了如何进行语义保真的问题，并实现了通信系统从比特准确度向语义测度下优化的转变。技术方法层面，为了实现在任意不可微语义测度上的通用优化，并同时处理通常被忽略的未知（特别是不可微）信道问题，本文提出使用强化学习中的决策过程建模语义通信系统，并使用基于语义相似度的策略梯度实现端到端的优化。实验表明本文所提方案更好地保留了消息中的高阶语义关系，并具有良好的鲁棒性和通用性。

进一步地，本文将语义通信的思想从点对点系统扩展到多对多的异构网络，并提出了一种“对语义负责（以语义为传输对象和目的），对（具有异构语义知识的）用户特异”的语义网络架构。特别地，在本文设想的语义网络框架下，来自异构语义体系下的不同符号消息被映射到统一的联合语义表征空间；网络对不同形式消息所含有的共同本原语义进行传输，并使用异构用户各自的语义知识在接收端进行特异化的符号表达。这一思想与以往“对用户负责，对语义特异”的多用户语义网络设计有着显著的区别，也是对现有方案的补充。此外，本文首次从表征学习的角度重新思考了语义通信系统设计中语义可解释性、通用性、安全性等问题，并提供了相应的解决方案。实验验证了本文所提语义网络在异构语义表征准确性与可解释性、对抗语义攻击的安全性、系统可拓展性等方面相较以往语义通信方案的显著优势。

本文所提使用语义相似度构建语义通信系统的方案是对早期点对点语义通信如何实现探索，回答了如何实现语义测度优化的基础问题（语义保真问题），而所提本原语义网络是对未来异构多用户语义网络架构的探索，进一步回答了点对点语义通信如何拓展、语义如何实现统一表征的发展性问题（通用语义表征问题）。所研究的两大问题

在目的上各有侧重，整体以发展的脉络承接，在语义通信历史进程的不同时期为学术和工业界提供了相应的解决方案。

关键词：语义通信，语义网络，语义相似度，语义表征，强化学习

Abstract

Semantic communication has raised extensive interest in recent years, and is believed to potentially play a key role in next-generation communication schemes. As one of the early and mid-term works, this paper provides theoretical and experimental experience for the architecture design of both point-to-point semantic communication and heterogeneous semantic networks.

On the schematic design of semantic communication, cross-entropy and mean square error are the most widely adopted criteria, upon which extensive research on semantic extraction methods and transmission schemes are investigated in the literature. Unlike these approaches, we point out the discrepancy between symbolic accuracy and semantic understandings (caused by flaws in objective function design) that is usually overlooked by previous efforts. To address this problem, we introduce a new semantic communication paradigm that learns from general semantic similarity metrics. The proposed approach provides answers to semantics preserving problem during the communicating process, and enables a schematic shift that learns from bit accuracy to semantic similarity. Technically, to facilitate learning on general non-differentiable objective functions, and also to tackle the impact of an unknown (and typically) non-differentiable transmission channel, we put forward a new formulation that models the transmission task as a decision-making process, and seek to optimize the decoupled semantic transceiver with semantic similarity-measured policy gradient in an end-to-end manner. Extensive experiments verified the benefits of the introduced framework in revealing high-order semantic relations, along with its robustness and generalization ability.

Further, this paper extends the idea of semantic communication from point-to-point system to heterogeneous networks and introduces a semantic network system that is “responsible for semantics (ie., treated as the ultimate goal and transmission media) and specific to users (who have heterogeneous semantic knowledge)”. Specifically under the proposed semantic network, heterogeneous symbolic messages are jointly mapped into a unified semantic representation space; wherein a semantic network then transmits the common primitive semantics and expresses them at the receiver side with heterogeneous semantic knowledge. The introduced idea

is significantly different from previous multi-user communication network designs which are “responsible for each user and specific to semantics”, and serves as a supplement to existing architectures. In addition, this paper provides a very first insight, from the perspective of representation learning, into several key problems in semantic system design, such as semantic interpretability and semantic security, and provides the corresponding solutions. Experiments have verified the notable advantages of the proposed semantic network compared with previous arts in terms of the accuracy and interpretability of heterogeneous semantic representation, semantic security against adversarial attacks, and scalability etc.

Building a semantic communication system from semantic similarity contributes to an early exploration on how to enable point-to-point semantic communication, and answers the basic question on how to realize semantic level optimization (we call it the semantic fidelity problem). The proposed primitive semantic network exemplifies one possible architecture for heterogeneous multi-user semantic network in the future, which further answers the developmental question on how to generalize point-to-point semantic communication as well as how to obtain unified semantic representations (we call it the general semantic representation problem). These two problems differ in their own emphasis, yet are connected in the context of development and both provide for academia and industry possible solutions in different periods of time.

Keywords: semantic communication, semantic network, semantic similarity, semantic representation, reinforcement learning

缩略表及符号表

缩略词表

英文缩写	中文全称	英文全称
5G	第五代移动通信技术	5th Generation Mobile Communication Technology
AI	人工智能	Artificial Intelligence
B5G	后 5G	Beyond 5G
BPE	字节对编码	Byte Pair Encoding
CE	交叉熵	Cross Entropy
CNN	卷积神经网络	Convolutional Neural Network
CV	计算机视觉	Computer Vision
DL	深度学习	Deep Learning
DNN	深度神经网络	Deep Neural Network
GAN	生成对抗网络	Generative Adversarial Network
HARQ	混合自动重复请求	Hybrid Automatic Repeat reQuest
IMT2030	“6G”的正式名称	International Mobile Telecommunications 2030
IoU	交并比	Intersection over Union
JSCC	联合信源信道编码	Joint Source-Channel Coding
MDP	马尔科夫决策过程	Markov Decision Process
ML	机器学习	Machine Learning
MMSE	最小均方误差	Minimum Mean Squared Error
MSE	均方误差	Mean Squared Error
NLP	自然语言处理	Natural Language Processing
OFDM	正交频分复用	Orthogonal Frequency Division Multiplexing
PSK	相移键控	Phase-Shift Keying
RL	强化学习	Reinforcement Learning
ROI	感兴趣区域	Region of Interest
SGD	随机梯度下降	Stochastic Gradient Descent
TF-IDF	词频逆文档频率	Term Frequency-Inverse Document Frequency

符号表

符号	含义	符号	含义
x	标量 x	\mathbf{x}	向量 \mathbf{x}
$\mathbf{m}, \hat{\mathbf{m}}$	输入和解码的消息	$\Theta(\cdot, \cdot)$	语义相似度
$\mathcal{F}_{\text{TX}}, \mathcal{F}_{\text{RX}}$	参数化的编码器和解码器	ϕ, θ	编解码器参数
$\mathcal{H}(\cdot)$	随机信道	h^{-1}	信道响应的逆
(\sim)	对变量的采样	Sample	采样操作
T, N	输入及解码消息的长度	$w^{(t)}, \hat{w}^{(t)}$	输入和解码的第 t 个词汇
\mathcal{W}	符号字典	V	符号字典的维度
d	消息嵌入的维度	$g(\cdot)$	词频逆文档频率权重
$p(\cdot)$	具体取值的概率	$P(\cdot)$	泛指的发生概率
$\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	均值 $\boldsymbol{\mu}$ 协方差 $\boldsymbol{\Sigma}$ 的高斯分布	$\mathcal{MD}([p_1, p_2 \dots])$	多项分布
M	并行采样次数	$r^{(t)}$	t 时刻获得的奖励
$\mathbf{s}^{(t)}$	智能体 t 时刻的状态	$\mathbf{a}^{(t)}$	智能体 t 时刻的动作
G_t	t 时刻的期望累积奖励	γ	强化学习折扣因子
σ	探索的标准差	$\pi(\cdot)$	智能体的策略
α	学习率	J	目标函数
\mathbf{s}	语义向量	\mathbf{k}	语义知识
Speak	从语义到符号的表达函数	Summary	从符号到语义的表征函数
$\det(\cdot)$	行列式	exp	自然常数为底的指数函数
$:=$	定义为	\Rightarrow	推导符号
$\langle \cdot, \cdot \rangle$	元组对	\leftarrow	赋值操作
\sim	服从于	$\neg x$	非 x , 除 x 以外的

目录

致谢	I
摘要	III
Abstract	V
缩略表及符号表.....	VII
目录	IX
图目录.....	XII
表目录.....	XIV
1 绪论	1
1.1 论文研究背景及意义	1
1.1.1 通信中的语义.....	1
1.1.2 语义通信的可能性	2
1.1.3 论文研究意义.....	4
1.2 国内外研究现状	6
1.2.1 语义信息理论.....	6
1.2.2 物理层深度学习	7
1.2.3 语义通信架构探索	9
1.2.4 点对点语义通信的优化和实用语义网络.....	11
1.3 论文研究内容与贡献	12
1.4 论文章节安排.....	13
2 语义模型和相关原理	16
2.1 语义通信基本模型	16
2.1.1 语义通信的概念和语义流	16
2.1.2 语义通信和考虑语义的通信系统	17
2.1.3 基于深度学习的通信系统	18
2.2 强化学习	19
2.2.1 强化学习概念简述	19
2.2.2 策略梯度简介.....	22

2.3	表征学习	23
2.3.1	语义通信与表征学习	23
2.3.2	表征学习简介	24
3	基于强化学习构建的语义通信系统	26
3.1	本章概述	26
3.2	优化语义相似度的通信系统模型	26
3.3	在广义语义相似度下的通信问题构造	28
3.4	基于强化学习的语义通信方法构建	30
3.4.1	对不可微语义相似度的优化	30
3.4.2	对不可微信道的优化	34
3.5	实验设计和结果评估	37
3.5.1	实验设置	37
3.5.2	模型效果	40
4	基于语义表征的异构语义网络	49
4.1	本章概述	49
4.2	传输本原语义的异构网络建模	50
4.3	统一语义表征与特异用户表达	52
4.3.1	统一的语义表征模型	52
4.3.2	特异的用户表达模型	54
4.3.3	语义网络安全性	55
4.4	实验设计和结果评估	57
4.4.1	实验设置	57
4.4.2	模型效果	59
5	总结与展望	67
5.1	研究工作总结	67
5.2	研究工作展望	69
	参考文献	71
	附录	79
A	定理 1 的证明	79

B 定理 2 的证明	80
硕士期间的成果.....	83

图目录

图 1.1	本文的叙述思路和整体结构安排。框图中实线表示了实体上的逻辑关系，虚线展示了所研究两个问题之间的内在联系	15
图 2.1	语义通信的语义流模型	17
图 2.2	模块化通信系统设计和联合信源信道编码对比.....	19
图 2.3	强化学习中的马尔科夫决策过程示意图.....	20
图 3.1	本文所提的使用广义语义相似度构建语义通信系统的整体结构.....	27
图 3.2	所提 SemanticRL 方法的具体网络结构（子图 (a)）以及具体的强化学习解码方案示意（子图 (b)）	33
图 3.3	SCSIU 变体训练过程与联合信源信道编码的差异。从左至右，本文依次展示了联合信源信道编码训练的梯度更新情况，以及所提 SCSIU 方法中编码器、解码器的梯度回传情况。图中 \mathcal{F}_{TX} 、 \mathcal{H} 、 \mathcal{F}_{RX} 分别表示编码器、信道，以及解码器.....	36
图 3.4	在 BLEU 指标上所提系统与基线模型的性能表现。在 BLEU-1 到 BLEU-4 指标上的结果分别呈现在子图 (a) 到 (d) 中	41
图 3.5	所提 SemanticRL-JSCC 方法在词准确率、BERT 相似度以及 CIDEr 相似度上的定量指标	42
图 3.6	使用不同语义相似度量作为奖励函数对所提系统性能的影响	42
图 3.7	所提系统的噪声鲁棒性研究。子图 (a) 展示了在 0-20dB 随机噪声训练下模型在不同信噪比环境下的适应性。子图 (b) 展示了在固定 10dB 信噪比环境训练后系统于变信噪比环境下推理的性能。为更清晰展示高低信噪比情况下的性能情况，本文将 10 到 20dB 区间单独展示在所有子图的右下角.....	43
图 3.8	SemanticRL-SCSIU 的收敛情况。子图 (a) 到 (c) 分别展示了训练时编码器、解码器和整体的 CIDEr 奖励值得分	45
图 3.9	在 CIFAR-10 数据集上对图像使用时序循环渐进解码恢复的训练曲线和视觉效果。所展示场景以 10 分贝信噪比高斯白噪声训练并以图像恢复时相邻时间步之间的均方误差增益作为语义相似度量	46

图 3.10	在 CIFAR-10 数据集上将图像传输建模为连续动作空间决策后，所提 SemanticRL 方法的适用性及不同训练方式的差异.....	48
图 4.1	本文所提面向多用户异构语义通信的本原语义网络架构图	50
图 4.2	仅强调自编码准确性的系统在语义符号表征上的缺陷。从左至右，依次展示了所提本原语义网络各组件的消融实验、仅使用自编码约束下的语义符号空间，以及增加掩膜推理约束后的符号空间对比	59
图 4.3	所提转述任务和对齐任务对语义符号统一表征的影响	61
图 4.4	所提语义对齐约束对于消息本原语义表征的影响	62
图 4.5	语义攻击鲁棒性对比。子图 (a) 和 (b) 分别展示了语义攻击者在困惑度（越低表示攻击越成功）以及语义攻击准确性（越高表示攻击越成功）上的得分情况.....	64

表目录

表 3.1	所提 SemanticRL 算法不同变体的比较	45
表 4.1	消息本原语义表征空间的可视化举例和消融性对比。进行了语义对齐的模型（前四个例子所示）在局部和整体上都拥有较好的语义分布和异构配对规律；而未使用语义对齐的模型（后两个例子）存在着异构语义表征不匹配的问题	63

1 绪论

1.1 论文研究背景及意义

1.1.1 通信中的语义

通信技术的进步一直是信息社会发展的重要推动力。纵观从 1G 的模拟调制到当下 5G 万物智联时代的发展历程，消息的传输已经不再成为通信的壁垒；在此之上，现代通信更多地关注于如何为信息社会提供更高速率、大容量、低延迟，以及智能化的业务方案。与之相对应的，学术和工业界也不再仅仅满足于比特误码率（Bit Error Rate, BER）的基础需求，而更多地致力于提升用户的服务质量（Quality of Service, QoS）。考虑到日益增长的业务需求，以及相对有限的通信资源和基础设施，信息社会也正期待着更智能和高效的传输技术。

尽管现代通信系统¹在建立以低误码率为准则的可靠通信上已经取得了极为重要而成功的突破（如波束赋形、大规模多输入多输出技术、高频通信等），比特级别的误码率和用户期望的服务质量并不完全一致。考虑以下的一个例子，当发送端期望发送“我刚刚买了一串黄色的香蕉”时，由于随机信道的存在，接收端不可避免地可能会收到一些错误的信息（例如，“我买了一串香蕉”，漏掉了“刚刚”和“黄色的”）。尽管发生了误码的情况，接收方依然可以大致理解原意。但并非所有的错误都有着这种理想的性质。例如，“我刚刚卖了一串黄色的香蕉”和“我刚刚买了一串黄色的手链”虽然有着更低的误码率，我们却很难从中推测出正确的语义（Semantics）信息。事实上，从现实生活中很多的例子上，我们都可以观察到误码率和用户体验的差异。比如考虑另一个例子：“研表究明，汉字的序顺并不定一能影阅响读”——由于认知主体对语义世界的知识先验存在，以上的错误并不会给理解或任务的执行带来显著的影响，而这一特性也基本上有别于追求每一比特都精准复现的通信系统。本文特别地定义一个以语义为传输目标的通信系统为“语义通信”系统，以和基于误码率的通信系统区分开来。

¹“系统”在本文特别指代一种由物理实体（收发信机、无线信道等）构成，运用信号处理方法，支持收发双方通过无线信道实现信息传输基本要求的功能整体。与许多同领域文献 [1-4] 一致，本文所提系统主要针对信号处理方法而言，不含对硬件实现部分的讨论。

对比于通信而言,语义是相对抽象的概念。一般来说,语义指的是消息(自然语言、符号、图像、音频等)在特定的任务背景下所蕴含的意义。特别地,同一种语义可能具有不同的表现形式(例如描述物体的多种符号和语言),而相同的符号在不同的知识体系中也可能拥有不同的含义(例如同一个单词在不同上下文和语种中意义不同)。也因此,能够建立语义通信的一大前提是收发双方所建立的“共识”^[5-6]。在人类社会,共识一般可以通过学习和社会交往建立;对机器而言,共识的建立则可通过施加人类的认知规则学习得到^[7-8]。不失一般地,本文将收发双方对某一语义的先验理解称为“语义知识”,其中具有相同因果关系的部分称为“共识”。本文所指的语义主要体现在人类认知中的语义,所研究的语义载体主要以自然语言的形式体现。需要单独指出的是,本文所研究的具体语义形式和研究角度并不局限于自然语言,可以拓展到其他语义场景或领域。

1.1.2 语义通信的可能性

语义概念和通信传输的碰撞为语义通信的产生和发展带来了可能。在历史上,可靠通信(即,追求符号复现的低误码率)的思想可以追溯到 Shannon 和 Weaver^[9]。1948年,Shannon 揭示了信息不确定性的度量以及基本的信息理论^[10]。随后,Weaver 在 1949 年将 Shannon 的经典信息理论拓展到额外的两个层次^[9]:语义层面(对应 Weaver 理论中的 Level B, semantic problem)以及效用层面(对应 Weaver 理论中的 Level C, effectiveness problem)。其中语义层研究如何提取和传输信息的语义;而效用层关注收到的消息如何影响知识主体(接收方)的理解和执行。由于当时社会对于向香农可靠传输理论的迫切需求和技术条件限制,Weaver 的观点在很长一段时间内并没有引起太多关注。

语义通信发展的新热潮出现在近两三年(2019, 2020年)。一方面,有限的通信资源对海量数据的传输提出了前所未有的要求,通信业务也愈发朝着智能化的方向演进。另一方面,伴随着机器学习(Machine Learning, ML)、人工智能(Artificial Intelligence, AI)技术的飞速发展,使用通用模型提取和加工语义的方法逐渐成熟。在内外变革的驱动下,研究人员对语义传输的探索热情再次兴起,语义通信系统也在当下再次成为一个研究热点^[3,5-6,8,11-13]。在目前语义通信的愿景下,经典通信理论和人工智能技术已经成为支撑语义通信发展的两大技术基石。前者为语义通信的理论构建和发展规划提供了宝贵的分析经验;后者为语义的表征和加工提供了基础工具。一些前沿的人工智能方法,如自然语言处理(Natural Language Processing, NLP)和计算机视觉(Computer Vision, CV)

技术利用表征学习 (Representation Learning) 提供了通用的语义计算及表征手段, 突破了语义信息理论在复杂语义建模上的困难。

从目前已有的研究进展来看, 语义通信的研究是对 Shannon 可靠通信的拓展。语义通信的研究体系, 例如语义信息理论、语义模型、语义度量等借鉴了经典信息理论中的研究方法, 并在此基础上对语义的提取和表征展开了挖掘。Lu 等人^[13]进一步分析了可靠通信和语义通信的包含关系: 语义通信可视为一种更为广泛而本原的通信手段, 而可靠通信则是语义通信使用符号 (如词汇) 进行表征的一种实例。特别地, 语义通信系统可以使用非自然语言作为传输媒介而只传输本原 (Primitive) 的语义信息; 反之当我们视所有消息符号同等重要时, 语义通信则退化为要求所有符号复现的一种可靠通信手段。语义通信通过利用通信双方共享的先验知识, 能够更好地实现“达意”的目的, 但也正因其对某一具体特定场景的特异性, 在全场景的通用性也受到一定限制。目前语义通信的研究多针对于某一特定的任务场景, 以提升语义传输的准确性为主要研究方向。而人工智能技术, 特别是基于深度学习 (Deep Learning, DL) 的联合信源信道编码 (Joint Source-Channel Coding, JSCC) 凭借强大的拟合能力、噪声鲁棒性和技术通用性成为主流的语义通信研究方法。在这一主流技术体系中, 经典的收发信机被可学习的参数替代, 在损失函数的约束下得到联合优化, 最终以自编码的形式实现语义传输。

然而, 现有的语义通信方案仍然存在一些难以妥善处理的潜在的问题。首先, 主流的语义通信方案集中在模型层面的语义提取 (例如使用更高级的神经网络结构或传输方案), 它们的比特级 (Bit-level) 目标函数, 例如均方误差 (Mean Squared Error, MSE) 和交叉熵 (Cross Entropy, CE) 损失不可避免地会引入语义差异; 也因此它们或多或少地被限制在可靠传输的范畴之内^[13]。其二, 现有大部分工作都以可微的目标函数为研究目标, 但无法为可微和不可微场景提供一种通用的优化框架; 第三, 目前几乎所有的语义编解码研究工作建立在通信信道可微的强假设下, 或是被设计为传输少量的语义符号, 很难在大规模、现实的无线通信场景中得到拓展和运用。最后, 多数语义通信系统的研究局限于点对点的简单同构化 (拥有相似的语义形式和收发共识) 通信体系, 所研究的场景具有较大的特异性, 很难拓展到未来的多对多、异构语义网络部署中。

本文所研究技术的动机主要来自于以上一些现有方法的不足, 并提出一系列方法实现本文所设想的一些语义通信关键特性。本文的主要内容将具体针对以上的四个问题进行逐一讨论, 并对语义通信的安全性^[13]做相应拓展。

1.1.3 论文研究意义

语义通信本身是近些年受到广泛关注的话题。在该领域的研究初期，语义传输在提升通信效率，节约通信负载方面的诸多潜在优势得到了国内外的许多学者的热烈讨论。早期国内关于语义通信的讨论主要由本学科具有影响力的学者推动。具有代表性地，石光明和张平教授团队^[12,14-16]指出语义通信是 5G 通信之后 (Beyond 5G, B5G) 突破香农限的一种强大方式，提出了对于语义通信可能架构²的展望并验证了考虑语义的通信系统相比传统通信系统的巨大研究潜力。国际上，Strinati 等人^[6,17]较早阐述了语义通信在语义压缩和传输准确性方面的良好性能，并对其在未来 6G 场景的可能运用做了总结。随后的一些工作^[7,11,18-20]也进一步完善了语义通信的体系以及其在通信质量、效率和噪声鲁棒性等方面的价值。在语义理论上，文献 [13,21-23] 另外从信息负载和语义知识信息的角度对语义通信如何降低通信开销进行了阐述。

从国际上对语义通信的研究趋势来看，语义通信的研究正经历着从“多点开花”到“百花齐放”的过程。不论是从参与的广度和研究课题的深度来说，语义通信的热度正经历着前所未有的增长趋势。目前，在 IMT2030 (6G) 推进组发布的《无线 AI 技术研究报告》白皮书中，语义通信作为单独的一章内容也得到了学术和工业界的广泛关注。相关的一些研究，诸如语义通信及语义认知网络架构等也正在推进之中。在这样的研究背景下，本文作为国内首批以语义通信为主要研究内容的学位论文，也旨在为学术和工业界提供宽口径的介绍和较为深入的研究报告。

目前，尽管语义通信的概念在文献中已有较为充分的讨论，但如何实现语义通信依然没有一个确定的答案。近两年来，不同的学者分别提出不同的见解，也涌现了一些可能的点对点语义通信架构，但现有的方案依然存在着语义保真、通用性、可解释性（部分内容在上小节已作简要讨论）等许多问题。本文基于语义传输的基本定义，着眼于构建一种以语义为传输目标、具有良好通用性和可解释性的语义通信架构，对现有方案起到了补充和拓展的作用，也提供了一种新的研究思路。

具体地，本文所设想的语义通信系统具有如下的期望特点：

- 语义保真性。语义通信的最基础要求，以传输语义为目标而非追求比特、单词、或像素上的精确复现。

² “通信架构”在本文特指为实现通信所需的组件以及联系各组件的具体运行方法。

- 业务和技术的通用性及可拓展性。语义通信方法或框架需要适用于不同的业务场景，支持用户自定义的语义相似测度，并能为高维空间下不可微信道的优化提供实践途径。特别地，语义通信系统应向下兼容，并在顶层业务上提供良好的拓展能力。
- 异构语义普适性。不同于经典的符号通信或点对点语义传输，多对多语义网络³存在语义异构的问题（各个用户处在不同的语义体系中，通常具有不同的文化和符号表达）。通用语义通信系统需要支持异构的语义表达形式，并以一种通用而普适的语义语言或语义空间加以描述。
- 语义可解释性。经典通信系统具有良好的因果关系和决策依据。相比之下许多基于深度学习的语义通信系统多关注于语义通信性能的优化，然而就语义通信是如何表征语义、如何对神经网络的特征和决策做出解释来说，目前并没有单独而较为系统的研究。可解释性是研究人员对基于人工智能的语义通信系统建立信任的重要依据，也是指导更深层次语义表征的重要手段。
- 语义安全性。除了一般通信系统中的安全考量，语义通信系统需要对语义传输过程中的语义攻击，例如窃听、破译等有着良好的鲁棒性。语义通信，特别是多用户语义网络应相应地提供基于语义的新加密或攻击对抗方法。

立足于以上特点，本文首先设想并验证了一种以语义保真为目的新型通信系统，在语义通信系统架构方面做出了率先的探索。进一步地，本文为探索高维语义空间、不可微语义相似度和语义信道、多业务场景及多实现方法等实际场景下语义通信的通用性提供了基于强化学习的一种可行途径。此外，本文研究了在多对多异构语义网络中的通信问题，提出了一种对异构符号消息使用统一语义空间表征并利用异构语义知识实现差异性表达的方法。同时在语义安全性问题上，本文也提供了一种使用语义知识进行加密的新途径。

³在本文中“语义网络”指在多用户场景下，一种以用户为节点，由信息交互为边的网络状通信系统。本文讨论的语义网络是语义通信系统的一种表现形式，区别于人工智能和计算机网络领域的网络概念。

1.2 国内外研究现状

作为近些年兴起的一个研究方向，语义通信的理论和实验方法逐渐发展，但仍然处于初期到中期的探索阶段。目前，语义通信的研究主要可以分为语义信息理论、语义通信架构、语义性能优化、以及点对点通信到语义网络的拓展几个方面。本文的主要研究内容侧重于语义通信架构的探索和语义网络的构建。

1.2.1 语义信息理论

和 Shannon 经典信息理论类似，历史上关于语义通信的信息理论探索早于实验研究。在语义信息理论诸多的分支中，语义信息和语义不确定性是较早，也最为广泛研究的领域。Carnap 等人^[24]和 Bar 等人^[25]类比经典信息论，提出可以使用语义的逻辑概率计算某一语义概念发生的可能性，并使用对数概率模型来测量语义信息。Floridi^[26]和 D'Alfonso^[27]在 Carnap 工作的基础上进一步完善了 Carnap 等人理论中关于语义真值悖论的缺陷（文献通常称之为 Bar-Hillel-Carnap 悖论），并提出了使用语义空虚（Vacuity）、不准确性（Inaccuracy）和真实性（Truthlikeness）等语义度量。在文献 [5] 中，Bao 等人^[5]研究了一般意义上的语义信息，并在理论分析了一些诸如语义编码和语义通信的基本理论，如语义信道容量、语义信源编码和语义信道编码等。值得一提的是，不同于经典信息理论，语义概念随知识主体的认知差异而异，也有着不同的表现形式（符号消息），也因此给语义的度量带来了更加困难的挑战。除了早期在语义度量理论中的一些探索，最近一些相关工作，例如语义推理理论^[21]、语义共识的不确定性^[23]以及语义率失真理论^[22,28]等也得到了较为广泛的关注。

尽管基于信息理论的语义测度为衡量语义信息提供了可行的途径，目前大部分的语义信息度量建立在简单概率逻辑（如逻辑真和逻辑假）基础上。在实际的场景中，信息载体往往具有复杂的结构和表现形式。完整地将这些消息进行语义分解，并使用联合语义概率进行计算往往十分困难，难以在真实的复杂语义环境下得到拓展，也在一定程度上限制了语义信息理论对实际语义通信架构体系的指导和演进过程。在实际绝大多数的文献中，语义通信系统的构建往往并不直接采用基于信息论的语义信息度量，而是通过缩小收发双方的上下文语义距离^[21]或增大消息之间的语义相似度^[29]来实现。在文献 [29] 中，研究人员使用“广义语义相似测度”表示这一语义测度，在语义通信的学习目标上

进行了统一。

作为一种间接表征语义信息的方法，语义相似测度接受两个符号消息作为输入，并给出它们之间的语义相似性。一些早期测量语义相似度的方法主要基于结构化的规则来定义，例如使用语义符号的树状或图状编码^[30]等。在这种结构化表示下，不同语义信息的相似度可以用他们在结构化表征中的路径长度、深度、或编辑距离来表示^[31]。除了使用显式的结构表征，基于上下文和词性的统计分析方法通过统计语义之间的共现频率来计算语义间的相似性，例如词频逆文档频率^[32]（Term Frequency-inverse Document Frequency, TF-IDF）、语法分析^[33]、机器学习方法^[34-35]等。

近些年来，人工智能技术的进展也带来了更多复杂且特定于任务的指标，例如上下文词向量嵌入^[36-37]、预训练语言模型^[38-39]、跨模态相似性评价^[40]和非参数指标等^[41-43]。值得一提的是，一般意义上的语义相似度不仅限于自然语言，而可以是任何表达语义的信息媒介，甚至可以是执行某一任务的语义信息^[44-46]。

本文在第三章将介绍一种以语义相似度为优化目标的语义通信系统。在这一系统中，本文不再假设语义相似度量指标的可微性，从而可以接受任何用户定义的指标作为优化目标。得益于所提方法的良好通用性，本文为理论上的语义相似测度和实际的语义通信系统搭建了桥梁，并进一步验证了使用广义语义相似度构建实用语义通信系统的可行性和优越性，也期望所提方法可以推动更多在语义信息理论上的研究。

1.2.2 物理层深度学习

自从 2012 年开始，以 AlexNet^[47]为代表的深度学习模型依靠其优秀的拟合能力开始在诸多优化任务中超越经典的专家系统，并在计算机视觉、自然语言处理等领域根本性地变革了主流的技术方向。在 2016 至 2017 年，O'shea 等人^[48]首次将深度学习结合到物理层通信领域。在这一工作中，O'shea 等人提出了一种完全使用深度神经网络（Deep Neural Network, DNN）建模经典通信理论中编解码器的方法。特别地在这一方法中，编解码器使用可学习的神经网络来替代原本按照人工经验设计的模块，而信道被近似视为已知的随机过程。通过将每个符号的传输视为分类问题，该方法使用交叉熵（Cross Entropy, CE）损失函数对于符号的传输进行端到端的学习。尽管该系统建立在可微信道的假设下，O'shea 等人展示了神经网络系统在通信场景下的出色拟合能力，达到了与人类专家经典编解码策略，如汉明编码（Hamming Code）、相移键控（Phase-

Shift Keying, PSK) 等经典信源信道编码可比拟的性能。该论文也首次展示了以深度学习 (Deep Learning, DL) 为代表的人工智能技术在物理层通信运用的可行性。随后, 使用神经网络对编解码器进行联合信源信道编码 (Joint Source-Channel Coding, JSCC) 的工作得到了广泛的关注。Farsad 等人^[49]和 Bourtsoulatze 等人^[50]将 O'shea 的工作从符号传输拓展分别到符号流 (句子文本) 和结构化高维变量 (图像) 层面, 以交叉熵和均方误差为损失函数验证了使用联合信源信道编码在复杂信息载体下的可行性。需要指出的是, 得益于深度神经网络的复杂拟合能力和鲁棒性, 联合信源信道编码在低信噪比情况下展示出了比经典编码更优的性能; 后者虽在理论上存在任意误码率下的最优编码, 但由于实际通信场景下有限的码长设计和复杂环境限制往往无法达到理论性能。目前, 联合信源信道编码已经成为基于深度学习的物理层通信系统一项重要的使能技术。主流的语义通信架构多采用基于深度学习的联合信源信道编码来实现, 借助深度神经网络的上下文语义提取能力对语义进行表征。

直接对收发信机进行建模的联合信源信道编码技术可以视为一种特殊的自编码器。除此以外, 使用深度学习技术对未知且不可微信道的估计也是物理层通信的一大研究方向。为实现信道估计, 最直接的一类方法尝试直接对信道的逆变换进行建模。例如, Ye 等人^[51]将退化信号的恢复视为一种数值回归过程, 直接在深度网络的输出端重建原始信号, 并展示出该方案在正交频分复用 (Orthogonal Frequency Division Multiplexing, OFDM) 系统中相较于经典最小均方误差 (Minimum Mean Squared Error, MMSE) 方案更优的误码率。这类方法设计简单, 适用于数据丰富、拥有较大量的匹配样本的场景下, 属于判别模型。相比而言, 另一类方法不再直接恢复退化信号, 而是建模随机信道的退化机理。这种方法通常采用生成对抗网络 (Generative Adversarial Network, GAN) 学习从输入信号到退化信号的生成方式, 不需要严格匹配的监督样本, 属于生成模型。在部署时, 可使用预训练的网络参数提供代理梯度^[1]。

在常用的生成和判别模型之外, 强化学习作为一种按规则和累计奖励学习的方法在决策控制任务中已经展现出了最先进的效果^[52]。由于 (强化学习所要求的) 奖励函数并不需要保证可微性, 强化学习也被用于不可微信道状态估计并显示出了与已知信道情况下接近的性能^[53]。此外, 强化学习在通信中多被主要用于网络和应用层任务 (如切片、资源和路由调度、效率优化^[54-57]), 但与物理层任务的结合依然相对较少。比较成功的一些运用包括解码一些简单的符号^[53], 或传输简单的下游控制信号^[44,58]等。值得一提的

是，虽然强化学习具有良好的理论优势，该方法的训练相较监督学习面临着收敛性、网络规模、训练速度、以及可拓展性上的一些挑战^[57]。

本文提出了一种使用强化学习来优化语义相似度的语义通信系统，并首次将强化学习引入语义通信。所提方法同时松弛了对于语义相似度和高维空间下语义信道的可微性强假设，使用策略梯度方法为实现物理层语义通信在高维空间下的端到端优化提供了一种通用而实际的训练方法。

1.2.3 语义通信架构探索

语义通信系统是近年来最流行，也有着丰富研究成果的方向之一，同时也是早期语义通信研究的前沿阵地。虽然目前国内外学者所提实现方案各有不同，Seo 等人^[21]指出语义通信可以被视为拉进收发双方语义联合概率分布的过程。从这一角度出发，本文按照语义上下文概率建模的技术差异，将语义通信系统统一地分为两大类：显式语义通信和隐式语义通信^[8]。需要明确说明的是，本文所指的显式与隐式的定义差异在于：在收发双方进行语义通信的过程中，通信机制中是否存在按照数学规则建模推理的过程或存在某一概率推理实体。由这一定义，使用概率模型建模的方法，以及使用知识图谱等体系建立的语义通信机制存在显式的语义联想实体逻辑，在学习方法上属于生成式语义通信模型；而不含有逻辑实体的模型，如主流的自编码模型则是在学习过程中将特定的语义联想规则“嵌入”到参数表征中，在传达语义时表现类似一种判别模型。下面，本文将对这两类模型做更为具体的阐述。

- 显式语义通信模型

显式语义通信模型拥有语义联想和推理的实体逻辑。语义通信理论早期的一些工作，例如文献 [5,24-25,27,59] 均是建立在语义概率的基础上，这种方法通常使用语义概念之间的联合概率分布模型和贝叶斯估计等方法对语义选择的可能性进行描述。另外也有一些工作探索了收发双方在协作或对抗中进行语义优化^[5,60]的过程。最近，文献 [21] 和文献 [23] 另外分别从收发双方的上下文共识概率和本地语义知识的角度对语义推理问题展开了研究。在语义推理这种显式语义通信模型中，接收方不再直接将收到的消息作为最终解码的语义答案，而是利用这一语义线索在本地上下文知识中进行搜索和推断。基于概率的语义推断模型由于其良好的准确

性，不需要精确传输每一条语义关系，即可大致推断出发送方所期望表达的含义，在降低通信开销上展示出比经典通信系统更大的优势。此外，Choi 等人^[23]等人的研究工作另外展示了语义知识库更新在语义的推理过程所起的重要作用。

除了以上的一些以数学概率直接建模的语义通信系统，对概率的显示建模可以通过图网络、知识图谱等方法加以表征。例如，Wang 等人^[61]和 Liang 等人^[62]采用以三元组为表现形式的知识图谱将语义实体和语义关系建模为图网络中的节点和边，通过邻居节点的聚合推断出语义表达在本地知识体系中可能表达的含义。基于图网络的方法在实验中显示出知识补全、纠错、语义推断等方面的潜在价值。

值得一提的是，由于显式语义通信模型具有明确的概率推理实体或规则，将实体推理关系加以变化后这类方法易于实时更新和拓展，但学习所需的数据量较大。从学习系统的角度开来，显式语义通信系统将语义的传输的因果关系进行了显式建模，属于生成式模型。

- 隐式语义通信模型

隐式语义通信模型没有语义联想和推理的实体逻辑。目前较为主流的方法主要是基于自编码器的一些建模方法。与显式语义通信模型相比，这类方法的语义传输过程由预先学习到的参数所表征，其语义信息的解码过程并不具有明确的概率意义。隐式语义通信模型能够完成语义通信的关键在于能够利用深度神经网络对输入消息的内在语义联系（如共现频率、近义词等）进行建模，并能通过人类预先定义的损失函数将正确的语义匹配关系注入到可学习参数的更新、训练过程中。由于这类方法实现简单、准确性高，是目前语义通信架构，特别是基于深度学习实现的语义通信系统中的主要组成部分。

在使用可学习参数对语义的传输建模上，朴素深度学习方法（这里本文特指简单自编码器）是研究最早、成果也较为丰硕的方向。受到 O'shea 等人^[48]物理层深度学习工作的启发，Farsad 等人^[49]从实验上验证了使用联合源信道编码传输复杂文本的可行性。随后，Xie 等人^[2]正式以“语义通信”一词对联合语义编码的方法和概念展开了研究，并使用注意力模型来建模文本的上下文关系，也引起了学术界的广泛关注。随后的一些基于简单编码器的拓展性工作^[62-67]大多建立在 Xie 等人的基础上，其主要思想都是利用深度神经网络的拟合能力自动挖掘消息中内含

的语义关系，从而更好地实现语义编解码过程。但需要指出的是，基于自编码器的的工作基本都采用交叉熵和均方误差作为损失函数，但这些要求比特级别符号复现的工作在优化目标上依然与语义传达存在一定的差距。

为了解决比特级别距离和人类语义理解的差异，Lu 等人^[13,29]设想了一种以通用语义相似度为优化目标的语义通信架构。该方法通过将通信系统的传输过程建模为智能体的决策过程，引入强化学习的思想允许在任意不可微语义相似测度下的优化。随后，Wang 等人^[61]将以语义相似度为优化目标的强化学习系统进一步与知识图谱和注意力机制相结合，展示出在语义表征和语义准确性上更优的效果。除了上述点对点的强化学习方案，另外也有一些工作在多智能体联合语义通信上也做出了探索，例如多智能协作下的决策规划和无人机协作等^[44,58]。

本文在语义通信的历史发展进程中，首次指出了语义相似度量在构建语义通信系统中的重要作用，并提出了一种以优化广义语义相似度为目标的强化学习方法。该方法作为语义通信系统现有架构的补充，为研究人员对何为语义通信以及如何构建语义通信系统的探索提供了一定的帮助。

1.2.4 点对点语义通信的优化和实用语义网络

除了在语义通信系统架构上的研究，语义通信的优化和部署也在学术界和工业界受到了广泛的关注。目前语义通信的优化主要集中在提升准确率、降低通信开销、提升系统鲁棒性、大规模部署等方面。

在文献中，点对点语义通信的优化是受到最多关注的方向，例如轻量和分布式部署^[67]，混合自动重复请求（Hybrid Automatic Repeat reQuest, HARQ）方案^[63]、自适应比特率控制^[66]与自适应语义编码^[68]、码本辅助传输和语义噪声鲁棒性提升^[64]、信息瓶颈及压缩^[28]等。除了在通用任务上的一些优化，在基于特定任务场景的通信系统中，研究人员也对例如信息新鲜度^[69]、多模态传输^[70]、资源分配^[71]、分布式协作^[44]等具体任务提出了定制化的一些优化方案。

目前绝大多数语义通信系统专注于点对点语义通信方案的构建，学术界对于多对多语义网络的研究依然较少。Xie 等人^[72]设想了一种多用户多任务的语义通信系统，该方案为不同的用户分配各自不同的语义编码器，并在接收端使用多个对应的解码器完成对

应的任务。这一工作最近在 Zhang 等人^[4]的工作中被优化为多编码器-单解码器的方案，其中解码端使用多任务训练并以不同的子分类器对接各自的下游任务。然而，以上的设计依然存在着许多问题：首先，虽然以上的架构可以处理多模态异构语义消息的问题，由于其多编码器的设定，编码端的网络参数量随用户的数目线性增长，而网络整体的训练则面临着更高的复杂度，并不适合实际部署。其次，对多模态和异构任务使用分立的语义表征并没有抓住语义传输的本质；本文提出对不同的语义载体使用统一的语义空间加以描述，从而构建出一种通用的语义表征方法。此外，以上方法并没有特别考虑多用户场景中特有的窃听和攻击等安全性问题。

本文注意到现有的语义通信工作基本都以准确率优化为唯一考量，而忽略了使用深度学习进行语义通信的可解释性和拓展性。基于这一问题，本文首次尝试以表征学习的视角将通信问题从消息有效表征的角度加以解读，提出了在准确率之外的语义加工和表征优化新问题，也提供了相应的处理方法。

本文的第四章将介绍一种使用联合语义空间对异构的不同语义消息进行统一描述的方法，这一方法解决了多用户和异构语义下的联合语义表征问题，使得单编码器-单解码器的设计变得可能，也大大提高了语义通信的可解释性和拓展性；此外本文进一步拓展了 Lu 等人^[13]所设想的语义安全性问题，为物理层安全设计提供了一种使用语义进行加密的方法。

1.3 论文研究内容与贡献

本文的研究内容主要分为两大部分。本文的第一部分将介绍语义通信兴起之时本文在点对点语义通信系统架构上的研究成果。具体地，本文针对简单自编码方案在语义保真和通用性方面的缺陷，设计了一种以广义语义相似度为优化目标的通信系统，并提出了一种基于强化学习方法实现对任意不可微语义相似测度和不可微未知信道模型的优化。在第二部分，本文将语义通信的点对点架构设计拓展到多对多的语义网络中，并借鉴表征学习的思想设计了一种在多用户、异构语义场景下使用统一语义表征的语义网络架构，该方法大大减少了语义通信在学习和部署上的资源开销（单编码器-单解码器设计）并为统一异构语义的传输提供了可能。本文对语义表征空间的探索大大提升了现有语义通信系统的语义表征能力、可解释性和可拓展性，同时为物理层加密提供了基于语

义的新方案。

本文的贡献可以总结为以下几点：

- 本文设计了一种以传输语义为目的的通信系统。与经典通信系统和以往在语义通信上的一些尝试不同，本文所研究的系统不仅需要具有良好的语义建模能力，更要求在接收端实现语义的正确恢复。为此，本文重新思考了目前以主流的交叉熵和均方误差为优化目标的训练范式在语义保真上的潜在问题，并提出了一种基于广义语义相似度优化的通信系统，对以往的学习方法进行了统一。
- 本文提出了一种使用强化学习优化语义相似度以构建语义通信系统的方案。该方案使用广义语义相似度作为奖励函数，将通信过程转化为决策问题，使用策略梯度缩放了以往语义通信实践上对语义相似度量度和通信信道可微性的强假设，并实现了在高维语义空间上的端到端优化。实验表明本文所提方案相较于以往的基线模型更好地捕捉到了消息的高阶语义关系，而非专注于比特级别的符号复现。
- 本文将语义通信的点对点的设计拓展到多对多的异构语义网络中。具体地，本文提出使用联合语义空间对多用户、异构语义场景进行统一的语义嵌入。在这一框架下，单编码器-单解码器的设计相较于同期方法带来了训练和部署开销上的显著减少，并首次实现对异构语义载体的统一传输。此外本文首次以表征学习的视角对语义在异构场景下的提取和表征展开优化，拓展了以往以自编码准确性为唯一目标的语义通信设计准则，大大提升了语义表征的准确性、可解释和可拓展性。进一步地，本文对多用户语义网络中的语义窃听和攻击问题展开分析，为语义网络的实用化部署提供了一定的方法指导。

1.4 论文章节安排

本论文的整体组织结构如图1.1所示，主要章节安排如下：

第一章是本文的绪论部分，主要向读者介绍本研究领域的一些研究背景和国内外最近进展情况。此外，本文在本章也将对现有方法的一些研究划分和优缺点进行讨论，一方面是为使读者更好地了解该领域的最新动向，另一方面也凸显了本文研究内容与技术方法和现有工作的差异。在这一基础上，第一章作为文章的总起部分，向读者介绍了文章的主体内容和贡献。

在第二章中，本文将从语义通信基本模型、强化学习以及表征学习简介三个角度向读者介绍本文所涉及的一些基本知识。特别地，语义通信基本模型是本文后续介绍第三章的基本背景框架，而深度学习，特别是强化学习部分是第三章的核心内容；一些强化学习的基本术语和方法将在这一部分做简要介绍。表征学习的相关部分是本文第四章语义网络和语义表征总体设计思想的重要灵感来源，本文将表征学习也作为一个指导性内容做简要介绍。

第三章将介绍本文在使用强化学习构建语义通信系统方面的研究方法。该章节首先介绍了问题的由来和动机，并引出本文在广义语义相似度上优化的基本框架。随后，本文将对如何从强化学习角度将通信过程建模为决策问题，以及具体的概念定义和技术方法展开介绍。在第三章的最后，本文将对所提方法的语义保真度、鲁棒性以及可拓展性进行实验分析。

本文在第四章将点对点通信拓展到多对多的异构语义网络，并主要研究语义通信中的异构语义表征问题。在这一章节中，本文主要遵循和第三章类似的结构安排，依次介绍问题的由来、技术方法，以及最后的实验验证过程。

最后，本文在第五章将对所做工作进行总结，并对未来语义通信领域相关的发展方向作适当展望。

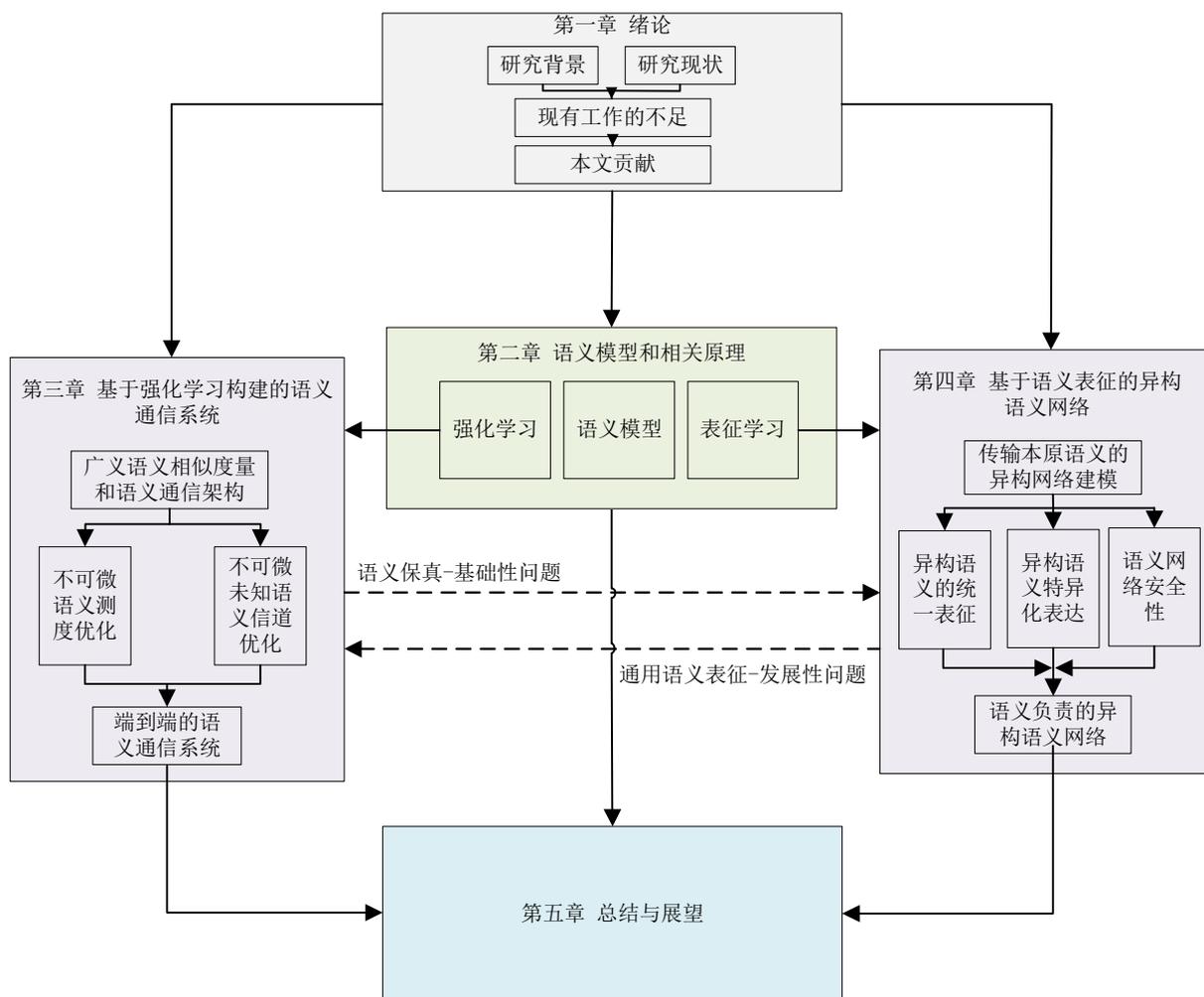


图 1.1 本文的叙述思路和整体结构安排。框图中实线表示了实体上的逻辑关系，虚线展示了所研究两个问题之间的内在联系

2 语义模型和相关原理

2.1 语义通信基本模型

为方便后续的建模，本节对语义通信中语义的加工过程以及主流的一些概念和分析方法做简要介绍。关于语义通信的一些概念性设想曾是早期研究中的重要内容，本文在后续章节介绍的技术方法也大体建立在这样的领域框架上。

2.1.1 语义通信的概念和语义流

首先，本文对语义，以及后文所涉及的一些名词做简要介绍：

- 语义 (Semantics)。一般而言，语义可以指向任何物理或虚拟世界中由实体和其关系构成的含义，在文献中通常认为语义可以从世界模型 (World Model) 观测得到^[5,26]，是某一事件的最本原刻画。
- 知识主体 (Participants with Semantic Knowledge)。知识主体是具有语义知识的智能体，能够对语义使用本地知识进行加工，并以消息的形式加以表述。不同的语义主体可能具有不同或部分不同的语义知识，其中语义知识可以通过训练或学习得到。在人类社会中，语义知识通常与文化和社会有关。未经社会化的个体除先天基因中含有的部分共性，通常与社会化个体存在较大的差异；来自不同文化和群体的个体对于同一语义也通常具有不同的理解。
- 符号消息 (Symbolic Message)。符号消息是知识主体利用自身语义知识对某一语义的具象描述。同一语义可能具有形式不同但语义一致的若干语义描述，例如描述同一事物的不同的语言和语法。语义符号消息的描述与语义本身，以及认知主体的知识结构有关。
- 推理 (Inference)。推理可以发生在语义和符号消息上，是利用已知信息对某一内容的推测。在语义通信的文献 [21,23] 中，推理一词多用于描述接收端从接收信号重构符号消息的过程。本文在第4章，对于从语义到符号的推理使用 **Speak** 函数表示；而从符号到语义的推理用 **Summary** 表示。

语义是消息的含义，消息是语义的具体表现。从语义的观测和通信过程中来看，完整的语义通信过程经历了从“语义观测”，到“符号表述”，再到“通信传输”和“语义还原”的过程。一个典型的语义通信场景如图2.1所示，本文特别地将语义信息在通信系统各个环节（表征、传输、表达等）下的存在形式和加工过程称为语义流。其中本文以 \mathcal{S} 表示语义的世界模型（集合）， s 表示某一具体的语义， m 表示在语义知识 k 下对语义的描述。

从通信的目的上来说，语义通信的最终目标是帮助接收方正确地将发送方传达的符号消息 m 恢复到语义层面 \hat{s} 。相比之下，以经典信息理论为依据的通信系统通常不考虑语义 s 和语义知识 k 的作用，而旨在尽可能地在接收端将发送的符号消息 m 复现。

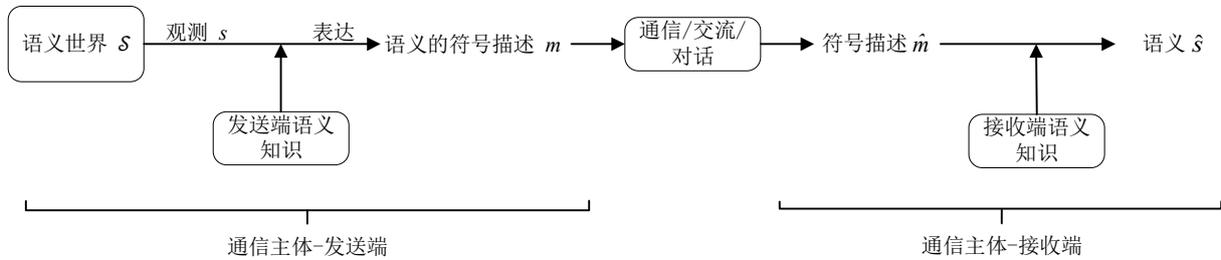


图 2.1 语义通信的语义流模型

2.1.2 语义通信和考虑语义的通信系统

以符号复现为目的的通信系统，其特征在于追求比特和符号层面上的准确性。我们可以很容易地使用误码率、均方误差、交叉熵等方法对（发送的符号消息 m ，恢复的符号消息 \hat{m} ）施加逐单词、逐像素的误码惩罚。类似地，一个语义通信系统，在理想情况下应当以语义 s 的恢复为传输目标。然而，语义本身是一个抽象的概念而非消息实体，通常情况下无法获得对语义空间或某一具体语义的具象表示。也因此，对语义准确率的优化是较为困难的。

本文第3章提出了使用广义语义相似度 $\Theta(m, \hat{m})$ 作为语义通信的优化目标的设想。通过使用 Θ 将符号空间映射到语义空间并计算符号消息在语义空间上的相似性，本文为使用语义相似度构建语义通信系统提供了示例。这一做法本质上是直接在语义空间 \mathcal{S} 上计算相似性的一种替代，利用知识主体从符号空间到语义空间转换的先验（语义相似度）从符号消息 m 间接地计算语义相似性，从而实现语义测度下的优化。

语义通信的目的在于恢复语义，而非恢复符号。只有以语义恢复为目的，施加了语义测度惩罚的系统是传输语义的系统。相对地，利用语义提取模型提高通信系统的符号准确率本质上依然是可靠通信（即，目标为追求低误码率）。需要特别指出，在本文中，语义通信的定义是严格的，区分语义通信和可靠通信的根本手段是其在通信目的上的差异；本文将“运用语义模型，但本质依然为可靠传输”的系统统称为“考虑语义的通信系统”^[13]。在文献中，另一种较为宽泛的语义通信定义包含了所有使用语义提取和加工技术的通信系统，而不是以通信目的为区分手段。

2.1.3 基于深度学习的通信系统

一个简单的通信系统模型由编解码器和信道构成。香农分离定理^[10]（Separation Theorem）指出，信源编解码和信道编解码可以分离设计而得到最优的通信系统。这一思想也是现代通信系统设计的基石之一。通常情况下，通信系统遵循着模块化的设计，编解码器可进一步由信源编解码器和信道编解码器构成，一个典型的点对点通信系统如图2.2上部分所示。然而，受到实际场景中带宽、码长、信噪比等条件的限制和动态变化，信源信道分离设计往往无法取得理论上的最优性能，而联合信源信道编码方式（如图2.2下部分所示）由于其一体化和联合优化的性质展示出了更优的传输性能^[49-50]。

特别地，近些年来得益于深度学习的发展，使用深度神经网络对编解码器进行建模成为了学术界研究的一大热点问题^[48]。在这一方法中，随机信道被建模为加噪的（部分参数已知的）随机过程，编解码器在损失函数（通常为交叉熵或均方误差）的约束下进行联合信源信道编码。其中，为了便利神经网络的梯度反向传播过程，学术界往往对信道的扰动有一个较强的可知假设。从通信目的来说，基于深度学习的联合信源信道编码以符号消息复现为目的，在传输方式上使用神经网络进行了改进。

在现有的一些设想和实验性工作中，语义通信的框架基本建立在图2.2所示的分离或联合信源信道编码基础上。主流的语义通信实现方案在概念上额外地引入语义编码和解码器，在实现上使用多层神经网络对发送方和接收方进行联合信源信道编码优化。整个通信系统可以视为一个自编码器，一般采取交叉熵或均方误差为损失函数，利用以随机梯度下降法（Stochastic Gradient Descent, SGD）为代表的深度学习方法进行训练。注意到，相比于以往使用深度学习进行联合信源信道编码的方法，一些早期的语义通信探索尽管在概念上提出了语义编解码器，但在实现中并没有真正从语义传输准确性的角度

对语义通信系统进行深入的分析，在优化目标、整体设计、优化方法的考量上存在着诸多不尽完善的方面。也因此早期的一些工作仍然属于“考虑语义的通信系统”而非严格意义上的“以语义保真为目的语义通信系统”。尽管如此，语义通信初期的一些探索为本文提出使用语义相似度构建语义通信系统提供了良好的借鉴和启发。

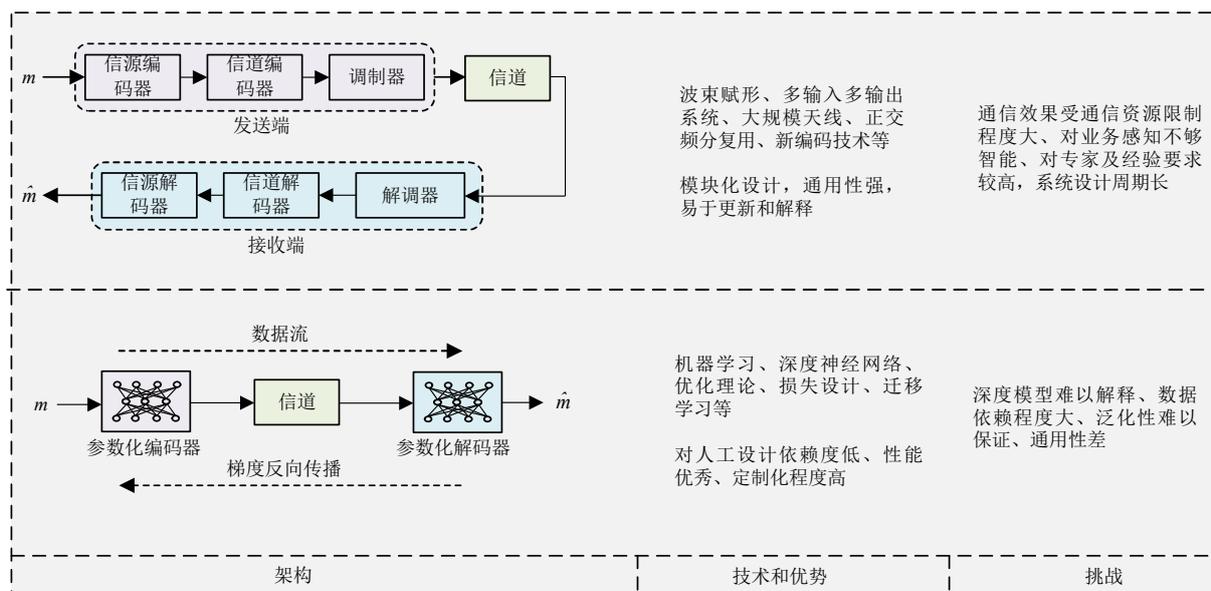


图 2.2 模块化通信系统设计和联合信源信道编码对比

2.2 强化学习

2.2.1 强化学习概念简述

本文第3章提出了一种使用语义相似度构建语义通信系统的方法。在这一方法中，本文进一步利用强化学习克服了不可微语义测度的优化问题。在这一小节本文将对强化学习的有关内容作简要介绍。

强化学习 (Reinforcement Learning, RL) 在机器学习和深度学习中通常作为与监督学习和无监督学习之外并列的一类方法。不同于一般的机器学习方法，强化学习并不直接利用损失函数对神经网络进行直接更新，而是将学习过程建模为时序决策问题，在不断试错的过程中获得不同的奖励 (进行抑制或增强)，并以获得累计奖励最大化作为优化的目标。

智能体 (Agent) 是强化学习决策的主体。在每个时间步中, 智能体需要根据其对环境 (Environment) 的观察 (Observation) \mathbf{o} , 利用某一策略 (Policy) π 做出决策动作 (Action) \mathbf{a} , 并以转移概率 P 更新自己的状态 (State) \mathbf{s} , 同时获得来自环境的奖励值 (Reward) r 。在强化学习中, 智能体与环境的交互可以被建模为马尔科夫决策过程 (Markov Decision Process, MDP), 如图2.3所示。

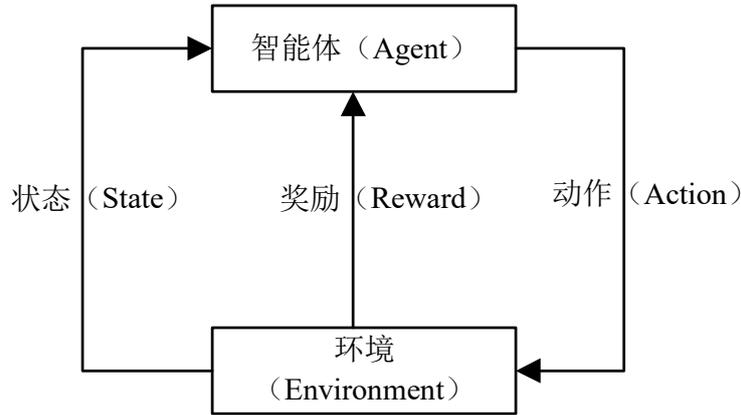


图 2.3 强化学习中的马尔科夫决策过程示意图

在强化学习中, 智能体的目的在于通过与环境的交互, 获得最大的累计奖励值。为了描述这一累计奖励的大小, 强化学习引入了状态值函数 $V(\mathbf{s})$ 和状态动作函数 $Q(\mathbf{s}, \mathbf{a})$ 来量化当前状态 (和动作) 在未来可能获得长期奖励的大小:

$$V(\mathbf{s}^{(t)}) = \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^k r^{(k+t+1)} | \mathbf{s} = \mathbf{s}^{(t)} \right] \quad (2-1)$$

$$Q(\mathbf{s}^{(t)}, \mathbf{a}^{(t)}) = \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^k r^{(k+t+1)} | \mathbf{s} = \mathbf{s}^{(t)}, \mathbf{a} = \mathbf{a}^{(t)} \right] \quad (2-2)$$

其中 $\gamma \in (0, 1]$ 表示折扣因子, 控制着未来时刻累计奖励对现在时刻的影响程度。 $V(\mathbf{s})$ 在数学上等同于 $Q(\mathbf{s}, \mathbf{a})$ 对动作 \mathbf{a} 的期望。状态值函数 $V(\mathbf{s})$ 和状态动作函数 $Q(\mathbf{s}, \mathbf{a})$ 统称为价值函数, 可依据贝尔曼方程 (Bellman Equation) 展开为递推形式:

$$V(\mathbf{s}^{(t)}) = \sum_{\mathbf{a}^{(t)}} \pi(\mathbf{a}^{(t)} | \mathbf{s}^{(t)}) \left(r^{(t)} + \sum_{\mathbf{s}^{(t+1)}} \gamma \cdot P(\mathbf{s}^{(t+1)} | \mathbf{s}^{(t)}, \mathbf{a}^{(t)}) V(\mathbf{s}^{(t+1)}) \right)$$

$$Q(\mathbf{s}^{(t)}, \mathbf{a}^{(t)}) = r^{(t)} + \gamma \cdot \sum_{\mathbf{s}^{(t+1)}} P(\mathbf{s}^{(t+1)} | \mathbf{s}^{(t)}, \mathbf{a}^{(t)}) \sum_{\mathbf{a}^{(t+1)}} \pi(\mathbf{a}^{(t+1)} | \mathbf{s}^{(t+1)}) Q(\mathbf{s}^{(t+1)}, \mathbf{a}^{(t+1)}) \quad (2-3)$$

式(2-3)显示了状态值函数和状态动作函数的前后的递推关系。根据这一关系,当环境中的奖励值、状态转移函数等参数完全已知时,可以使用动态规划方法对以上的强化学习决策过程进行高效求解。使用动态规划对 $V(\mathbf{s})$ 及 $Q(\mathbf{s}, \mathbf{a})$ 的求解有值迭代和策略迭代两种常用的方式^[73]。由于在动态规划方法中,环境模型被认为完全已知,这一方法属于有模型学习 (Model-based Learning)。

在很多问题中,环境并非完全已知或难以全部感知,此时的学习方法可以被粗略统称为无模型学习 (Model-free Learning)。在无模型学习中,智能体不对状态转移概率进行直接建模,而是通过蒙特卡洛模拟 (Monte Carlo Rollout) 或时序差分方法 (Temporal-Difference Learning) 的方式对价值函数进行估计:

- 蒙特卡洛方法。该方法直接对于当前时刻以后的动作轨迹进行采样模拟,直到执行结束动作或达到规定时间步。根据大数定理,当采样次数越多时,使用采样轨迹对于当前状态累计奖励值的估计将变得更为准确(均值不变,方差降低)。该方法对累计奖励的估计是无偏差的,但要求获得连续且完整的采样,通常实施的复杂度较大。
- 时序差分方法。该方法利用后续若干(通常为后续一步)步的(折扣)价值作为当前期望累计奖励的估计,无需等到整个决策轨迹结束即可完成价值函数更新。与蒙特卡洛方法不同,时序差分方法利用后继状态进行价值函数更新,属于有偏估计;但因为通常时序差分仅考虑后续若干步骤,轨迹采样的方差较小。

在完成价值函数更新(策略评估)后,通常可遵循与策略迭代或值迭代类似的思路,按照 ϵ -greedy 方法^[73]按照状态动作值函数 $Q(\mathbf{s}, \mathbf{a})$ 进行策略选择,并更新价值函数。常用的一些更新方法包括 Q-学习^[74] (Q-learning)、SARSA^[75]等。

直接使用价值函数 $V(\mathbf{s})$ 和 $Q(\mathbf{s}, \mathbf{a})$ 的方式虽然非常直观,在复杂高维问题上,建立状态表格将面临着巨大的动作和状态空间,并难以提供准确的价值函数估计。考虑到深度学习良好的函数拟合能力,对 $V(\mathbf{s})$ 和 $Q(\mathbf{s}, \mathbf{a})$ 函数使用非线性神经网络进行数值近似的方法提供了良好的解决思路。在 DQN^[76]中,研究人员首次展示了这样一种完全利用深度神经网络对强化学习中价值函数和决策网络进行建模,并实现大规模端到端学习的方法。这一工作也启发了后续许多深度 Q 网络的变种,开启了深度强化学习领域的研究热潮。

2.2.2 策略梯度简介

以上提到的一些方法主要基于对价值网络 $V(\mathbf{s})$ 和 $Q(\mathbf{s}, \mathbf{a})$ 的评估和更新进行建模，从而训练智能体。除了这些常用的方法之外，另一大类强化学习方法使用策略梯度进行学习。

在策略梯度中，智能体的目标函数 J 依然是最大化累计奖励 $\sum_{k=0}^{\infty} \gamma^k r^{(k+1)}$ 或平均奖励 $\bar{r} = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T r^{(k+1)}$ 。以平均奖励为例，策略梯度方法的目标函数可以表达为：

$$\begin{aligned} J &= \mathbb{E}_{\pi(\mathbf{s}, \mathbf{a})} [\bar{r}] \\ &= \sum_{\mathbf{s}} P(\mathbf{s}) \sum_{\mathbf{a}} \pi(\mathbf{s}, \mathbf{a}) \bar{r} \end{aligned} \quad (2-4)$$

使用对数梯度技巧 $\nabla \pi(\mathbf{s}, \mathbf{a}) = \pi(\mathbf{s}, \mathbf{a}) \nabla \log \pi(\mathbf{s}, \mathbf{a})$ ，式(2-4)的梯度为：

$$\begin{aligned} \nabla J &= \sum_{\mathbf{s}} P(\mathbf{s}) \sum_{\mathbf{a}} \nabla \pi(\mathbf{s}, \mathbf{a}) \bar{r} \\ &= \sum_{\mathbf{s}} P(\mathbf{s}) \sum_{\mathbf{a}} \pi(\mathbf{s}, \mathbf{a}) \nabla \log \pi(\mathbf{s}, \mathbf{a}) \bar{r} \\ &= \mathbb{E}_{\pi(\mathbf{s}, \mathbf{a})} [\nabla \log \pi(\mathbf{s}, \mathbf{a}) \bar{r}] \end{aligned} \quad (2-5)$$

不同于上小节对价值网络的建模，策略梯度^[77]直接使用奖励值（或累计奖励函数）对模型参数的更新方向提供指导，而无需单独维护价值函数 $Q(\mathbf{s}, \mathbf{a})$ 与 $V(\mathbf{s})$ 。这一策略可以理解为对带来高价值动作的奖励和低价值动作的抑制。

策略梯度的目标函数定义和累计奖励估计方法可以是多样的，除了上面提到的平均奖励外，Actor-critic 算法^[78]将决策部分建模为演员（Actor）网络，使用累计期望奖励 $Q(\mathbf{s}, \mathbf{a})$ 提供优化方向，并使用批评家（Critic）网络对 Q 函数进行估计。而在 REINFORCE^[77]算法中，研究人员则是利用蒙特卡洛方法直接估计当前状态下的累计期望奖励。

基于策略梯度的方法不需要对价值函数进行迭代维护，而是每次都朝着带来策略提升的方向更新，因此通常有着较好的稳定性；此外由于该方法高度的简洁性和假设的一般性，在连续动作空间以及高维动作空间中不需要进行大量的迭代计算，在高维空间下往往更加有效。

值得一提的是，由于直接使用式(2-5)进行估计的方式具有较大的方差（特别是在蒙

特卡洛方法中), 在对策略梯度使用目标函数进行加权时通常会额外引入与决策动作无关的偏置项 b 来降低策略梯度估计的方差, 而不改变策略梯度估计的期望值 (无偏):

$$\nabla J = \mathbb{E}_{\pi(\mathbf{s}, \mathbf{a})} [\nabla \log \pi(\mathbf{s}, \mathbf{a}) (\bar{r} - b)] \quad (2-6)$$

其中累计 (或平均) 奖励与偏置项的差通常被称为优势函数 (Advantage Function), 表示进一步对高于基准奖励的动作进行奖励, 对低于基准阈值的动作进行惩罚。在使用优势函数的 actor-critic 方法中, 研究人员使用 $V(\mathbf{s})$ 作为偏置项。如何设计合理的偏置也是策略梯度中的一个研究方向^[79-80]。

自 Williams 等人在 1992 年提出 REINFORCE 算法后, 策略梯度方法目前已经得到了长足的发展, 成为强化学习领域一个重要的学习方法, 在连续控制和离散控制上都得到了广泛的运用。目前, 策略梯度的一些代表性方法有深度确定性策略梯度 DDPG^[81]、信任区域策略优化 TRPO^[82]、近端策略优化 PPO^[83]等。

2.3 表征学习

2.3.1 语义通信与表征学习

语义通信系统的目的在于语义恢复, 从学习系统的角度来说这是一个自监督自编码任务。纵观语义通信中系统设计、架构设计、模型设计等一系列研究问题, 在很大程度上都可以归结到研究如何对语义进行合适的加工和表征。一个好的语义通信系统, 不仅应当能够对语义进行恢复, 更需要具有语义的理解和正确的表征能力、可解释性, 便利后续顶层语义任务实现 (如语义网络中的聚类、推荐、调度等功能)。从符号消息到语义挖掘, 语义通信在文献中常被视为自编码任务, 以准确性为度量, 却往往忽视了语义本身的表征情况和语义通信的解释性问题。

本文首次在语义通信中引入表征学习的思想, 主要的动机来自观察到于语义通信对语义进行自编码的本质。语义表征的好坏直接决定了语义能否被正确表达, 以及能否具有良好的适用性和可拓展性。语义通信, 包括基于深度学习的通信系统, 从星座图到语义提取器, 本质都是在符号或语义空间中为消息 (符号或语义) 寻找合适的 (例如正交的、稀疏的、簇状的等) 隐空间嵌入。

在本文所呈现技术贡献的第二部分 (第4章), 本文将探索语义和符号消息之间的映

射关系，从语义表征的角度对异构语义进行统一表征，实现一种以本原语义为传输流的异构语义网络。本文也期为语义通信系统，包括以往一些使用深度学习构建语义编解码的方案，提供一些在表征学习上的解释；也期待后续在语义通信上的工作能够从语义表征的角度对语义传输展开更深入的挖掘。

2.3.2 表征学习简介

表征学习 (Representation Learning, 中文也译为表示学习) 是研究数据如何得到有效表示的研究领域, 通常是将高维输入数据以低维流形进行呈现。深度学习先驱 Bengio 在 2013 年指出^[84], 有效的表征需要寻找到低维空间上符合人类先验、而又有利于下游任务的数据特征。

对数据进行高效表征曾是机器学习领域中困扰研究人员的一大难题。经典的一些机器学习算法, 例如支持向量机、决策树、聚类方法等, 都期望输入数据具有较为分明而易于决策的特征。如何设计这些特征也是特征工程的主要研究内容。早期一些代表性的特征提取算法通常经人工选择并需要大量专家先验知识; 为不同的输入信号找到各自领域有效的表征方法往往需要耗费大量的人力与时间, 且并不通用。

深度学习是现代表征学习的重要方法。与使用人工特征不同, 深度神经网络在损失函数的约束下进行参数调整以改变特征提取的方式 (例如卷积操作本身即为一种特征提取的滤波过程), 从而获得对输入信号具有整体把握的重要特征。深度表征学习的出现使得研究人员将精力从特征工程转移到更高层次的任务分析和模块设计上, 也大大加速了机器学习领域的进展。在 2016 年前后, 研究人员发现通用数据集训练得到的特征表征 (往往是深度网络提取的高阶语义特征) 不仅仅可以用于所研究的特定问题, 通常也具有良好的泛化性。利用这一特性, 深度表征学习在辅助迁移学习上也得到了广泛的关注。

一般而言, 深度表征学习可以借助有监督、半监督或无监督的方式实现:

- 基于监督学习的方法。监督学习是早期深度学习的主要方式, 常用的监督学习任务, 例如分类、目标检测都是常用的表征学习任务。特别地, 在目标分类任务的许多著名网络, 例如 VGG^[85]、ResNet^[86]等到目前为止依然是许多任务上常用的特征提取器。

- 基于自监督学习的方法。自监督学习通常使用样本本身提供训练所需的标签，一些常用的自监督学习有自编码 (Autoencoder) 任务、噪声自编码 (Noise Autoencoder)、掩膜预测 (Mask Predicting)、对比学习 (Contrastive Learning) 等。自监督学习，特别是对比学习方法往往会对表征空间进行进一步地约束，以获得更为平滑和高效的特征表达。
- 基于无监督学习的方法。无监督学习不需要事先获得训练的配对标签，一般可使用生成对抗 (Generative Adversarial)、伪标签 (Pseudo Labeling) 等方式实现。自监督学习也可以归类为无监督学习的一种。

在许多时候，上述的三种学习方式往往是互补的。监督学习目标函数提供了基于具体任务的引导；而自监督和无监督方法擅长对数据的内在分布特征和语义推理关系进行建模。考虑到配对数据的获取难度，使用自监督和无监督训练方法是表征学习中采用较多的方式^[38,87]。在一些需要配对数据的场景中，使用半监督和伪标签进行训练的方式也是常常被采取的方法^[88-89]。

作为数据内在特征的代表，表征学习所构建的参数空间不仅仅在便利下游任务上具有良好的性质，这些隐变量表征本身也在数据表征方面，特别是语义计算方面获得了成功的运用。BERT 语义表征^[38]、多模态语义检索模型 CLIP^[40]、文本引导图像生成 GLIDE^[90]等都是近些年较为流行并成功商用的一些例子。

3 基于强化学习构建的语义通信系统

3.1 本章概述

本章对于所研究的第一个关键问题，也即使用语义相似度构建语义通信系统展开叙述。在本章的3.2节中，本文将介绍该方法提出的动机并介绍本文所研究模型的基础架构。3.3节对所述方法进行了正式的数学建模。3.4节将详细介绍使用强化学习来建模通信系统的具体方法和方案设计。最后，3.5节对所提方法的性能进行了评估。实验结果表明，本文所提方法 **SemanticRL** 在语义保真和鲁棒性上都有着更好的性能。

3.2 优化语义相似度的通信系统模型

已有的语义通信系统多以交叉熵或均方误差作为系统的目标函数^[2,49-50]。这些损失函数要求发送消息和接收消息以逐字或逐像素的方式复现。虽然以上一些方式在经验上可以达到传输目的，它们的本质依然是追求比特层面的复现，而非在语义意义上的传输。一些后续的很多工作^[63-64,66]在语义提取的主干网络和传输方法上提出了改进，但受到目标函数的限制，它们或多或少地依然受限在可靠通信的范畴中。需要指出的是，仅使用更复杂和高性能的语义提取单元确实提供了更优的语义建模能力；但从学习系统的角度来说，这与我们传输语义的最终目的依然存在显著的差异。

要实现以语义为传输目标的通信架构，一个自然的途径是重新定义学习系统的目标函数，将优化目标从比特级相似度转移到语义相似测度上。基于这样的考虑，本文定义一种广义的优化目标-语义相似度 Θ 。相比于使用逐比特的监督，语义相似度更接近认知主体对消息的理解和感知过程，也因而更适合构建语义通信系统。从语义信息理论上来说，语义相似度在物理层比特传输之外，额外引入了认知主体对“语义”概念的先验，建模的是主体之间对于语义符号的理解不确定性。而单纯使用比特层面的监督没有考虑到主体的认知差异，其学习的过程可以视为一种“考虑上下文语义的编解码策略”。

需要说明的是，本文探讨的重点并非如何去定义这样一个相似度量。事实上，使用何种语义度量在很大程度上取决于通信主体对于消息和语义符号的认知偏好，并因具体的任务而异。举例来说，在最常研究的自然语言语义中，我们可以借助于自然语言处理

领域已有的成熟指标来作为语义的相似度量；在图像传输中，我们也可以定制不同的度量方式，例如当我们期望传输图像的显著部分时，感兴趣区域（Region of Interest, ROI）和收到消息的交并比（Intersection over Union, IoU）可以被考虑是一种语义度量。本文不具体定义某种新的度量方式，也不再假设这种度量的可微性。但可以预见的是，一种好的语义度量方式应该能够准确反映认知主体的语义理解偏好，并具有良好的优化性质（例如，输入输出较为平滑、对不同表述具有良好的灵敏度、具有良好的语义表征一致性等）。事实上，在本文使用广义语义相似度和强化学习构建语义通信系统的构想提出后，后续也有学者对如何确定具体场景的语义相似测度开展了一些讨论^[61,91]。此外，本文也认为来自信息论角度的语义相似测度是一个值得期待也值得继续研究的重要方向。

在广义语义相似测度的基础上，本文所提出的语义通信系统，其基本模型如图3.1所示。其中，子图3.1(a)展示了本文所研究语义通信模型的具体技术结构，子图(b)展示了“广义语义相似度量”的可能表现形式。从结构上来看，本文沿用了与现有工作相似的编解码结构，但不再假设通信信道的可知性和可微性；从模型上看，本文放缩了对于何种具体编解码器（如长短时记忆网络、卷积神经网络等）的限制，而是着眼于构建一种通用的优化和学习方法；从学习目的上看，本文定义了一个广义的目标函数 Θ ，不再局限于常用的交叉熵或均方误差，而是允许用户定义任何的，甚至是不可微的一些语义度量。

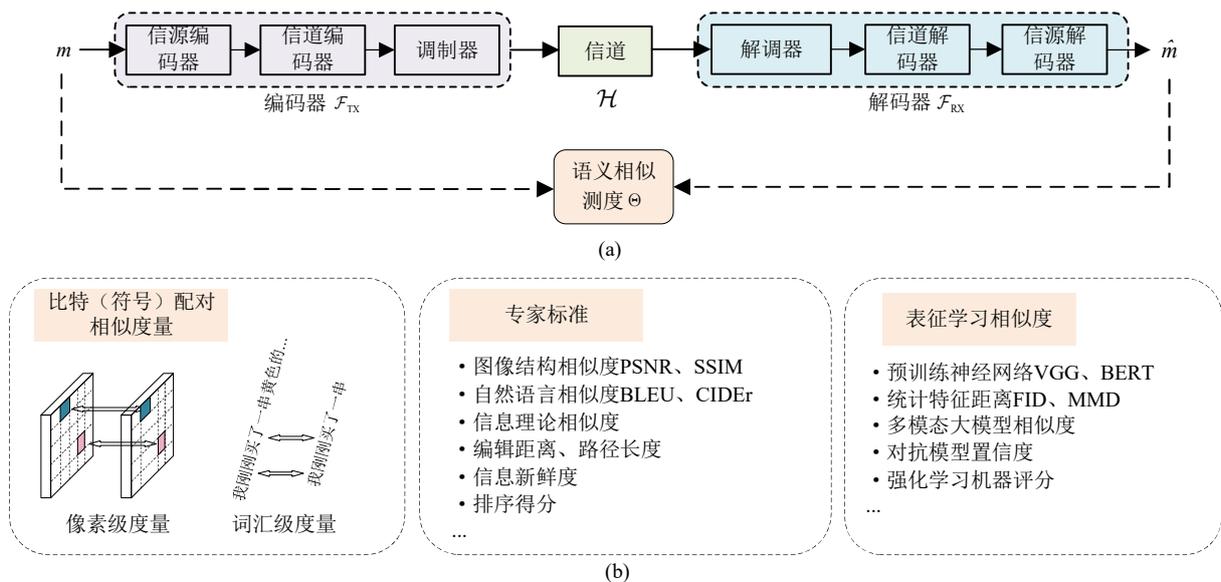


图 3.1 本文所提的使用广义语义相似度构建语义通信系统的整体结构

3.3 在广义语义相似度下的通信问题构造

本文考虑如图3.1(a)所示的语义通信系统，其中编码器和解码器分别用可学习模型 $\mathcal{F}_{\text{TX}}(\cdot)$ 和 $\mathcal{F}_{\text{RX}}(\cdot)$ 表示（本文以花体符号 \mathcal{F} 表示可学习的模型，下标 TX 和 RX 分别表示该参数模型位于发送端和接收端）。信道则用 $\mathcal{H}(\cdot)$ 表示。需要注意的是，本文所指的编解码器在实际系统中可能由若干子模块构成，但这里为了讨论的简洁性本文认为所有子模块可视为一个整体；更进一步地本文将以 ϕ 和 θ 分别表示编解码器的参数。作为语义通信系统，对于某一条消息 \mathbf{m} ，接收方的基本任务是，需要从其观测到的经过信道退化的 d 维向量 $\mathcal{H}(\mathcal{F}_{\text{TX}}(\mathbf{m})) \in \mathbb{R}^d$ 中解码得到与 \mathbf{m} 语义接近的表达⁴。

除了特别指出外，本文遵循以往在语义通信上的工作^[2,49]，主要以自然语言为研究的语义载体。在这一问题下，本文记某一句子为 $\mathbf{m} = \{w^{(1)}, w^{(2)}, w^{(t)}, \dots, w^{(T)}\}$ ，其中 T 表示句长， t 表示在 $[0, T]$ 之间的某个时间步；句子中的每一词汇 $w^{(t)}$ 都来自词典 \mathcal{W} 。在解码端，接收方在每一个时间步内需要按照参数模型 $\mathcal{F}_{\text{RX}}(\cdot)$ 的输出概率分布 $[p(\hat{w}_1^{(t)}), p(\hat{w}_2^{(t)}), \dots, p(\hat{w}_V^{(t)})]^T$ 从 V 维字典空间中选择最合适的单词作为输出。不失一般性地，本文假设解码后的句子用 $\hat{\mathbf{m}} = \{\hat{w}^{(1)}, \hat{w}^{(2)}, \dots, \hat{w}^{(N)}\}$ 来表示，其中考虑到解码的句子可能与输入存在一定的差异，本文用 N 表示解码句长。

对于以上的通信系统，其最优编解码器由下式给出：

$$\langle \phi^*, \theta^* \rangle = \underset{\mathcal{F}_{\text{TX}}, \mathcal{F}_{\text{RX}}}{\operatorname{argmax}} \Theta(\mathbf{m}, \hat{\mathbf{m}}) \quad (3-1)$$

注意在这里和以往的语义通信工作不同，我们不再假设 \mathcal{H} 和 Θ 的可微性。

为实现对式(3-1)的优化，广义语义相似度 Θ 可以有多种表征的形式。由于本文不要求 Θ 的可微性，在1.2.1节和图3.1中介绍过的一些度量方法都可以作为具体的优化指标。特别地，考虑到使用预训练模型提取的相似性指标给训练带来的额外开销和使用深度学习来指导另一套黑盒的欠解释性，本文选择了解释性更强，更易于拓展，也较为成熟的一些例如 BLEU^[41]和 CIDEr^[42]等的自然语言指标。BLEU 最初用于评价机器翻译的准确性，是一种广泛接受的自然语言相似度指标，也是目前语义通信领域最主流的评价指标之一^[2,13,49]。BLEU 通过计算词汇上下文的 n 元语法 (n -gram, 该方法考虑由连续 n

⁴本文在第二章指出语义通信的最终目的是对语义而非符号消息进行恢复。考虑到语义是一个较为抽象并难以表征的概念，文献中关于语义通信的研究通常以具体的符号消息为解码的目标，要求收发消息（作为语义载体）之间语义一致。本文使用广义语义相似度 Θ 作为符号消息之间的语义相似性度量。

个词汇构成的短语，其中 n 多被称为阶数）匹配程度（从 1 阶到 4 阶）进行打分：

$$\Theta_{\text{BLEU}} = \exp \left(1 - \frac{T_r}{\min(T_c, T_r)} + \sum_{n=1}^4 \xi_n \log q_n \right) \quad (3-2)$$

其中， q_n 是解码句子 $\hat{\mathbf{m}}$ 的规范化的 n 元语法匹配得分， ξ_n 为加权因子； T_r 和 T_c 分别代表参考句子和解码句子的长度。CIDEr 指标则利用了词频逆文档频率^[32] (TF-IDF) $g(\cdot)$ 构建语义符号的向量表征，并使用余弦相似度对 n 元短语进行相似性评估：

$$\Theta_{\text{CIDEr}} = \sum_{n=1}^4 \xi_n \frac{g_n(\mathbf{m}) \cdot g_n(\hat{\mathbf{m}})}{\|g_n(\mathbf{m})\| \|g_n(\hat{\mathbf{m}})\|} \quad (3-3)$$

其中 $g_n(\cdot)$ 是利用词频逆文档频率对消息中第 n 元语法的向量表征； ξ_n 为加权因子。式(3-3)中分子为向量内积的形式，分母为向量 2 范数的相乘。

一般情况下，BLEU 和 CIDEr 被用于衡量 1 到 4 阶的语义相似度。更高阶的 n 元语法相似度在语义上下文给予了更多的侧重，而 1 阶相似度可以视为完全忽略上下文语义的一种退化（此时只关心某个语义概念是否出现而忽略其语义联系）。一个好的语义通信系统通常被期望拥有着较高的高阶语义相似度得分。

作为一种广义的语义相似测度，特别地，当 Θ 使用交叉熵（或图像传输中常用的均方误差）来表示时（例如 Θ_{CE} 和 Θ_{MSE} ），本文所提方法退化为以往的符号自编码器模型。交叉熵损失视符号的传输过程为时序分类任务，在逐单词层面上对符号传输的准确性进行约束（ ϕ 、 θ 分别是编解码器参数）：

$$\Theta_{\text{CE}} = \log \prod_{t=1}^T p(w^{(t)} | \hat{w}^{(0)}, \dots, \hat{w}^{(t-1)}; \phi, \theta) \quad (3-4)$$

需要说明的是，以交叉熵和均方误差为代表的一些符号级相似测度通常是可微的，在这种情况下可以直接使用随机梯度下降法进行优化，几乎所有以往的语义通信方法均采用此类方式^[2,48-49,63,67-68]；但在例如 BLEU 和 CIDEr 等不可微指标下，以往的深度学习将不再适用。为进一步解决这一问题，本章后续小节将介绍一种使用强化学习提供代理梯度（Surrogate Gradient）的学习策略。此外，尽管本文没有采用基于深度学习预训练模型的一些度量方法，例如 BERT 模型^[38]，利用这些预训练模型计算语义相似度（如使用余弦相似度构造 Θ_{BERT} ）也是完全可行的。

3.4 基于强化学习的语义通信方法构建

3.4.1 对不可微语义相似度的优化

3.4.1.1 以马尔科夫决策过程定义的通信系统

为了叙述的简洁性，同时尽量保持和现有的语义通信架构的可比性，在这一节中，本文考虑句子的传输过程，并暂时假设信道是可微的，也即允许使用联合信源信道编码（JSCC）的方式对收发信机完成梯度优化，也因此本文将本小节所提算法命名为 SemanticRL-JSCC。本文的3.4.2节将对信道的可微问题做单独讨论。

以往以交叉熵和均方误差为损失函数的语义通信框架建立在梯度反向传播的基础之上对模型做迭代更新，这一方法的基础假设是模型处处可微。考虑到通用语义通信架构不再假设语义相似度的可微性，已有的许多语义通信方法便无法继续运行。为了实现对于不可微优化目标的优化，一种可行的方法是使用可微算子来近似代替导致不可微的运算部分。然而这种设计也存在着极大的问题：1) 为近似不同语义相似度量，我们需要根据不同的指标定制不同的可微算子，并且难以保证在环境（例如信道和基础设施条件）变化时依然能做到良好的适应；2) 为不可微函数设计可微算子本身也涉及到大量的技巧，这对于提高语义通信方案的通用程度也造成了一定的影响^[92]。为了获得更好的普适性以及鲁棒性，本文借鉴于强化学习方法，提出一种使用奖励函数来提供优化目标的学习方案。

事实上，符号的时序解码过程与经典决策问题，例如围棋游戏有着十分相似的特性。一方面，时序解码的最终结果只有当句子完整解码后才具有参考价值，这一点与棋类博弈中未到死局不定胜负的决策过程极为类似；另一方面，两者决策过程的优劣情况都无法轻易使用可微的指标进行描述，而是采用“奖励函数”的形式对决策过程进行指导。这一特点进一步启发本文将通信过程（在本小节主要考虑解码过程）转化为一个时序决策过程加以描述。

在这样的框架下，本文首先定义强化学习的几大要素：

- 决策主体：联合收发信机（本小节主要考虑解码器）。
- 状态：解码器的隐状态 $\mathbf{s}_{\text{dc}}^{(t)}$ ，以及历史生成的符号序列。具体地，本文使用 $\mathbf{s}^{(t)} = \{\mathbf{s}_{\text{dc}}^{(t)}, \hat{w}^{(0)}, \hat{w}^{(1)}, \dots, \hat{w}^{(t)}\}$ 进行表示。在符号传输的过程中，一旦下一时刻的输出符号

被确定，状态之间的转移关系是确定的。

- 策略：在当前时间步和状态下，选择何种语义符号作为输出。
- 动作：本文定义动作为从 V 维字典 \mathcal{W} 中被选择出的符号 $\hat{w}^{(t)}$ 。与以往在通信中的强化学习问题不同，本文所研究的时序符号解码问题在动作空间上比前者高若干数量级，这也使得构造合适的策略变得更为困难。
- 奖励：解码消息和发送消息之间的语义相似度。

本文对以上定义的语义通信方法做简要分析。该方法与以往基于自编码的深度学习方法的差别在于功能上的转变。在解码流程上，以往的符号解码建立在分类（传输符号的情况）或回归（传输图像的情况）的强监督学习框架下，可被视为一种判别式学习；而强化学习的思路将每个语义符号的解码视为一种决策过程，并不严格专注于当下解码时刻的优劣与否，而更关心句子解码完毕后与输入文本的长期语义差异。在学习目标上，通过将严格可微的强监督转换为基于奖励函数的动态指引，强化学习为不可微语义相似度的优化提供了可能，也更接近于语义传输的本质。

3.4.1.2 基于强化学习的不可微目标优化

不失一般地，本文首先给出以上问题的累积回报 $G^{(t)}$ ：

$$G^{(t)} = r^{(t+1)} + \gamma r^{(t+2)} + \gamma^2 r^{(t+3)} + \dots = \sum_{k=0}^{N-t-1} \gamma^k r^{(t+k+1)} \quad (3-5)$$

其中 γ 是强化学习中的折扣因子，代表了未来预期奖励对当前决策的折算程度。一般而言在句子传输中，可以简单设置 $\gamma = 1$ 来简化问题的讨论^[93-95]。在文献中， $G^{(t)}$ 也可以被条件地定义在智能体的状态或状态-动作对上，相对应地称为状态值函数 $V(\mathbf{s}^{(t)}) = \mathbb{E}[G^{(t)} | \mathbf{s} = \mathbf{s}^{(t)}]$ 和状态动作函数 $Q(\mathbf{s}^{(t)}, \mathbf{a}^{(t)}) = \mathbb{E}[G^{(t)} | \mathbf{s} = \mathbf{s}^{(t)}, \mathbf{a} = \mathbf{a}^{(t)}]$ 。

于是，对于使用强化学习优化语义相似度的通信系统（在3.3节中，本文以 ϕ 和 θ 分别表示了编码器与解码器的参数），其目标函数 J 定义如下：

$$\begin{aligned} J(\phi; \theta) &= \mathbb{E}_{\hat{w}^{(1)}, \dots, \hat{w}^{(N)}} [G^{(0)} | \mathbf{s}^{(0)}] \\ &= \mathbb{E}_{\hat{w}^{(1)}, \dots, \hat{w}^{(N)}} \left[\sum_{t=1}^N r^{(t)} \right] \end{aligned} \quad (3-6)$$

其中, $\hat{w}^{(1)}, \dots, \hat{w}^{(N)}$ 是解码的一条完整轨迹 (由历史决策动作构成), 也即解码器从初始状态 $\mathbf{s}^{(0)}$ 开始解码出的一个完整句子。目标函数 $J(\boldsymbol{\theta})$ (对解码器而言) 可被视为从初始状态开始, 对未来决策的总期望奖励, 更大的目标函数对应了更高的语义相似度。需要指出的是, 当我们暂时假设信道的可微性时, 本文以联合信源信道编码的方式对编解码器整体进行优化, 其最优参数由 $\langle \boldsymbol{\phi}^*, \boldsymbol{\theta}^* \rangle = \operatorname{argmax} J(\boldsymbol{\phi}; \boldsymbol{\theta})$ 给出。

定理 1 (语义策略梯度) 设 $\hat{w}^{(1)}, \dots, \hat{w}^{(N)}$ 是在策略 $\pi_{\boldsymbol{\phi}, \boldsymbol{\theta}}$ 下决策得到的时序解码结果, 策略 π 在 t 时间步对动作 $\hat{w}^{(t)}$ 的输出概率为 $\pi_{\boldsymbol{\phi}, \boldsymbol{\theta}}(\hat{w}^{(t)} | \mathbf{s}^{(t)})$, $J(\boldsymbol{\phi}; \boldsymbol{\theta})$ 的梯度可以使用如下的蒙特卡洛估计 (Monte-Carlo Estimation) 近似:

$$\nabla J(\boldsymbol{\phi}; \boldsymbol{\theta}) \approx \sum_{t=1}^N \nabla_{\boldsymbol{\phi}, \boldsymbol{\theta}} \log \pi_{\boldsymbol{\phi}, \boldsymbol{\theta}}(\hat{w}^{(t)} | \mathbf{s}^{(t)}) \Theta(\mathbf{m}, \hat{\mathbf{m}}) \quad (3-7)$$

证明过程见附录A。

推论 1 对于连续控制任务, 智能体可以采用连续动作策略, 定理 1 依然成立。

本文对以上的定理和推论展开分析。与直接使用损失函数对语义通信系统 $\{\mathcal{F}_{\text{TX}}, \mathcal{F}_{\text{RX}}\}$ 进行更新不同, 强化学习对于正确的决策加以奖励, 从而激励系统做出更有利的决策, 也即获得更大的语义相似度。在广义相似度指标 Θ 不可微的情况下, 式(3-7)依然能提供合理梯度的关键在于, 它并不试图直接对 Θ 求导, 而是利用广义语义相似度指标的标量值为通信系统决策的对数置信度 ($\log \pi_{\boldsymbol{\phi}, \boldsymbol{\theta}}$) 提供指引, 以奖励得到更高语义相似度的动作。也正因此, 本文所提系统在学习方法的角度上根本区别于以往的深度学习方法, 为优化不可微目标函数提供了可行的方法, 也直接建立起了连接广义语义相似度和实际语义通信系统的桥梁。

3.4.1.3 高维语义空间下的强化学习实现

本小节将具体介绍为实现解码端不可微语义相似度优化而采取的解码策略和优化方案。

在以往的工作中, $\hat{w}^{(t)}$ 通常通过在解码器决策分布中利用贪婪方式 (Argmax) 选择得到。然而在强化学习决策问题的框架下, 解码输出轨迹 $\hat{w}^{(1)}, \dots, \hat{w}^{(N)}$ 的产生将不再是一个确定的解码过程。对于给定的某一策略分布 $\left[p(\hat{w}_1^{(t)}), p(\hat{w}_2^{(t)}) \dots p(\hat{w}_V^{(t)}) \right]^T$, 本文采用 Softmax 策略对该分布进行采样, 并以此更新解码端的状态变量:

$$\pi_{\phi, \theta}^{(t)} := \text{Sample} \left(\mathcal{MD} \left(\left[p(\hat{w}_1^{(t)}), p(\hat{w}_2^{(t)}) \dots p(\hat{w}_V^{(t)}) \right]^T \right) \right) \quad (3-8)$$

其中， \mathcal{MD} 表示服从多项分布模型， Sample 表示采样操作。使用多项分布对归一化的采样概率建模确保了原本具有较大可能性采取的动作依然具有较大的采样概率，而一些可能性较低的动作（语义词汇）依然有可能被探索到。

经过 **Softmax** 操作和多项采样，理论上式(3-7)中的所有参数都已经具备了完备的计算条件。然而，按照式(3-8)进行的蒙特卡洛采样虽然为语义策略梯度提供了无偏的估计量，这种从高维空间的长时间采样不可避免地给代理梯度的估计带来了较大的方差。为了降低式(3-7)估计的方差，一种常用的技巧^[77,79,96]是在累计奖励项中加入额外的偏置项 b 。在 **Actor-critic** 算法中，偏置项使用估计的状态值函数 $V(\mathbf{s}^{(t)})$ 来表示，然而为了计算这一额外的偏置项，就不可避免地引入另一个状态值估计（神经）网络，也将极大地增大训练的计算开销。类似地，也有一些工作将偏置项 b 的估计视为额外的估计任务^[96]，但无论如何以上的一些做法都不可避免地引入额外的参数并使得整体训练对于偏置估计的准确性更为敏感，特别是在高维语义空间下可能面临着训练发散的风险。

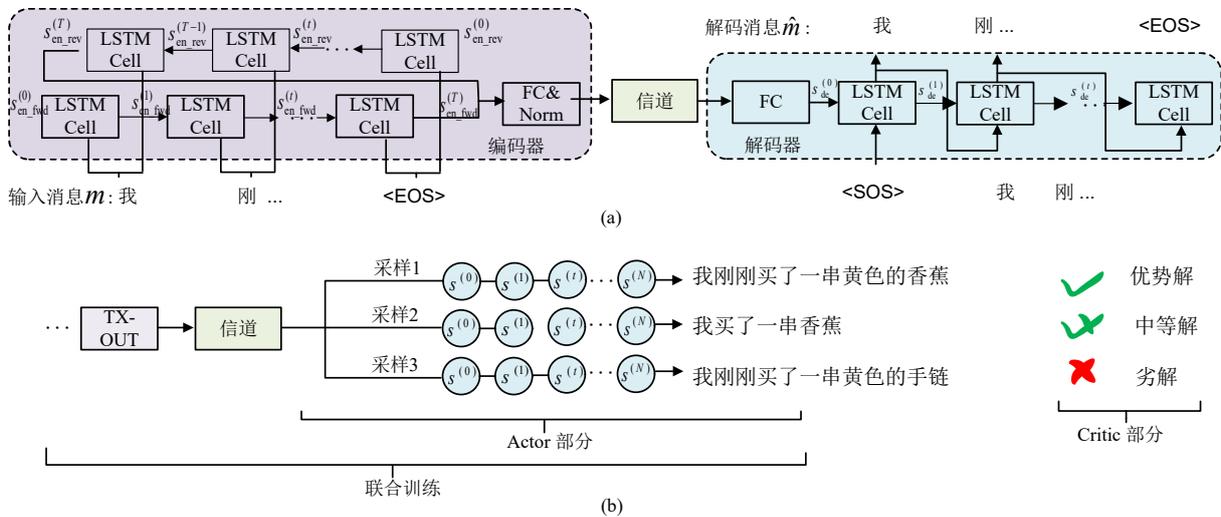


图 3.2 所提 **SemanticRL** 方法的具体网络结构（子图 (a)）以及具体的强化学习解码方案示意（子图 (b)）

针对在高维语义空间下的优化问题，本文采取了一种称为“**self-critic**”的技术^[79]，该方法通过计算同一输入语义向量经并行采样解码的平均累计奖励替代了额外的状态值估计网络。需要指出的是，对于如何准确提供偏置项 b 的探索也是学术界乐于探索的一个重要问题^[80]，属于强化学习本身的一些技巧。文本并不着眼于探索强化学习领域的

前沿技术，而是关注于如何利用强化学习工具实现以语义传输为最终目标的通信系统。

具体而言，self-critic 方法对初始状态 $\mathbf{s}^{(0)}$ 进行多次并行采样，得到 M 条采样轨迹 $\{\hat{\mathbf{m}}_1, \hat{\mathbf{m}}_2, \dots, \hat{\mathbf{m}}_M\}$ 。对于其中某条采样轨迹，该方法利用其余 $M - 1$ 条并行轨迹的回报均值作为其本身估计的偏置项：

$$\nabla J(\phi; \theta) \approx \frac{1}{M} \sum_{i=1}^M \left[\sum_{t=1}^N \nabla_{\phi, \theta} \log \pi_{\phi, \theta, i}(\hat{w}^{(t)} | \mathbf{s}^{(t)}) \cdot \left(\Theta_i - \underset{k \sim M; k \neq i}{\text{avg}} (\Theta_k) \right) \right] \quad (3-9)$$

其中 $\text{avg}(\cdot)$ 代表均值函数， $\pi_{\phi, \theta, i}$ 中的下标 i 表示第 i 条采样轨迹。此处 Θ_i 表示采取第 i 条轨迹所带来的回报。在实际意义上， M 起到平衡并行采样复杂度和奖励估计准确性的作用。

如图3.2所示，式(3-9)可以解释为一种基于优势函数（Advantage Function）的策略梯度方法——对于并行决策中回报高于平均值的动作加以奖励，并对于劣势决策进行惩罚。在实际的通信系统中，式(3-9)在不引入其他参数的条件下提供了当前策略的准确代理梯度，既缓解了朴素策略梯度的高方差问题，也避免了训练额外的（尤其是高维空间下不稳定的）状态值估计网络。相对于以往使用交叉熵进行优化的方法，并行采样的方式允许系统在计算资源富足的条件下（例如中心服务器）以近乎相同的时间复杂度进行训练。

最后，本文将使用强化学习优化不可微语义相似度的整体算法流程归纳于**算法 1**。

3.4.2 对不可微信道的优化

3.4.2.1 信道不可微下的语义通信模型

在实际场景下，无线随机信道通常是随机且不可知的。将其假设为可微且已知的模型仅仅便利了基于深度学习的仿真过程，却很难运用到实际的通信场景中。在这一小节，本文进一步放宽这一强假设，考虑信道的不可微优化问题。此时由于我们不再假设随机信道模型已知，联合信源信道编码的梯度仅能回传到解码器，而无法优化到编码器，从而导致现有方法无法运用到此类场景。

在上一小节中，本文针对损失函数不可微的问题给出了语义策略梯度的公式(3-7)。现考虑信道不可微的场景，其与上一节的不同之处在于：1) 编码器的输出不再是一个离散空间的概率分布（用于分类），而是用于编码的浮点数（连续动作空间）；2) 编码

算法 1 基于强化学习的语义相似度优化算法 SemaniticRL-JSCC 算法

输入: 批次大小, 初始学习率 α , self-critic 并行采样次数 M , 输入消息 m , 预训练轮次 E_p , 训练结束轮次 E_e , 广义语义相似度准则 Θ

输出: 编码器参数 ϕ , 解码器参数 θ

// 预训练阶段 (可选)

- 1: **for** 训练轮次 = 1: E_p **do**
- 2: 采样一批次数据, 按3.5.1节所述更新学习率 α
- 3: 编码器将消息 m 编码为 $\mathcal{F}_{\text{TX}}(m)$ 。对编码信号做功率归一化, 并将其发送至噪声信道中
- 4: 接收端进行解码, 并按照式(3-4) (作损失函数时取相反数) 计算交叉熵损失 \mathcal{L}_{CE}
- 5: 按照随机梯度下降法更新编解码器参数: $\langle \phi; \theta \rangle \leftarrow \langle \phi; \theta \rangle - \alpha \cdot \nabla \mathcal{L}_{\text{CE}}$
- 6: **end for**

// 基于强化学习的语义相似度优化阶段

- 7: **for** 训练轮次 = $E_p + 1$: E_e **do**
- 8: 采样一批次数据, 按3.5.1节所述更新学习率 α
- 9: 编码器将消息 m 编码为 $\mathcal{F}_{\text{TX}}(m)$ 。对编码信号做功率归一化, 并将其发送至噪声信道中
- 10: **for** $i = 1:M$ **do**
- 11: 按照式(3-8)所述采样得到解码序列
- 12: **end for**
- 13: 为每次采样得到的轨迹计算与输入消息之前的语义相似度 $\Theta(m, \hat{m})$
- 14: 按照式(3-9)计算目标函数 J 的梯度
- 15: 更新编解码器参数: $\langle \phi; \theta \rangle \leftarrow \langle \phi; \theta \rangle + \alpha \cdot \nabla J(\phi; \theta)$
- 16: **end for**

// 训练结束

- 17: 返回编解码器参数 $\langle \phi; \theta \rangle$

器的输出不再是一个循环的时序过程。虽然这一问题与前述的不可微语义相似度优化有所区别, 我们注意到根据**推论 1**, 非循环决策过程可以视为循环解码中时间步为 1、动作空间连续的特例。因此在编码过程中, 本文将上节的语义策略梯度思想进一步拓展到编码器中, 使得原通信系统从单解码器决策优化转化为编解码器的协作优化。具体地本文给出该场景下的系统模型:

$$\phi^* = \operatorname{argmax}_{\mathcal{F}_{\text{TX}}} \Theta(m, \underbrace{\mathcal{F}_{\text{RX}}(\mathcal{H}(\mathcal{F}_{\text{TX}}(m)))}_{\text{参数固定}}) \quad (3-10a)$$

$$\theta^* = \operatorname{argmax}_{\mathcal{F}_{\text{RX}}} \Theta(m, \underbrace{\mathcal{F}_{\text{RX}}(\mathcal{H}(\mathcal{F}_{\text{TX}}(m)))}_{\text{参数固定}}) \quad (3-10b)$$

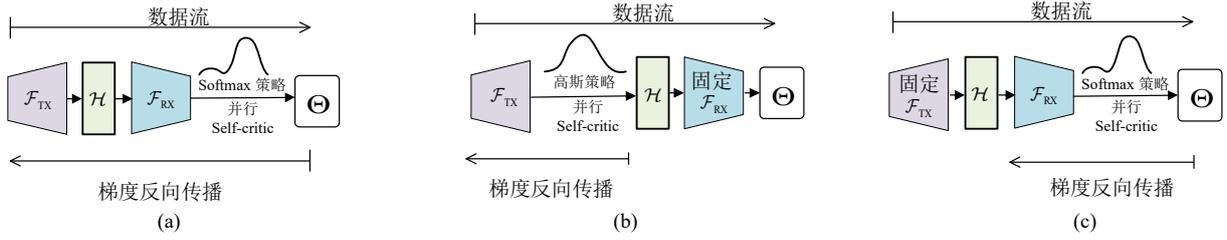


图 3.3 SCSIU 变体训练过程与联合信源信道编码的差异。从左至右，本文依次展示了联合信源信道编码训练的梯度更新情况，以及所提 SCSIU 方法中编码器、解码器的梯度回传情况。图中 \mathcal{F}_{TX} 、 \mathcal{H} 、 \mathcal{F}_{RX} 分别表示编码器、信道，以及解码器

如式(3-10)所示，当收发端被单独地建模为两个决策智能体时，其参数相互独立，但共享着同一优化目标 $\Theta(\mathbf{m}, \hat{\mathbf{m}})$ 。与在3.4.1节中使用的联合优化系统不同，本文在图3.3中比较了两者系统建模之间的差异。其中，联合信源信道编码模型（图3.3(a)）假设通信信道的可微性，并对编解码器采取联合优化的方式；而在通信信道不可微的情况下，需要对编解码器分别进行优化（图3.3(b)及(c)）。

3.4.2.2 基于随机迭代优化实现的语义通信

为实现式(3-10)描述的语义通信模型，本节将分别对编解码器的决策方案和优化策略展开介绍。特别地，本节将提出一种基于 **self-critic** 方法的随机迭代优化方案，以实现在信道和广义语义相似度同时不可微情况下的系统优化。本文将这一算法称为 **SemanticRL-SCSIU** (Self-Critic Stochastic Iterative Updating)。

首先，本文将对编解码器的决策策略做简要介绍。在编码器端，由于信道的不可微性质阻碍了梯度从解码器的反向传播过程，编码器也被视为一个智能体进行单独优化。具体地，本文对编码器的输出 $\mathcal{F}_{\text{TX}}(\mathbf{m})$ 采用高斯策略 (Gaussian Policy) 进行建模，此时原本在联合信源信道编码中编码器的输出不再被视为确定的浮点向量，而是以均值 $\boldsymbol{\mu}$ 和协方差 $\boldsymbol{\Sigma}$ 的形式加以描述：

$$\pi_{\phi} := \text{Sample}(\mathcal{N}(\boldsymbol{\mu} = \mathcal{F}_{\text{TX}}(\mathbf{m}), \boldsymbol{\Sigma} = (\sigma \mathbf{I})^2)) \quad (3-11)$$

其中 \mathbf{I} 表示单位矩阵。在式(3-11)中，本文将原先编码器的输出值设置为高斯分布的均值 $\boldsymbol{\mu}^{d \times 1}$ ，并使用固定大小的协方差 $\boldsymbol{\Sigma}^{d \times d}$ 对采样过程中的随机性加以表征。在不特别指出的情况下，本文默认以对角形式的协方差矩阵 $(\sigma \mathbf{I})^2$ 来建模，并设置每个维度上的标准差 σ 为 0.1 平衡决策过程中的探索与利用。

在解码器端，本文沿用在3.4.2节介绍的 Softmax 策略进行采样：

$$\pi_{\theta}^{(t)} := \text{Sample} \left(\mathcal{MD} \left(\left[p(\hat{w}_1^{(t)}), p(\hat{w}_2^{(t)}) \dots p(\hat{w}_V^{(t)}) \right]^T \right) \right) \quad (3-12)$$

相比式(3-8)而言，式(3-12)对解码器进行了单独建模，在解码器参数的更新中，梯度传播仅对解码器有效。

结合式(3-7)，本文在**定理 2**中给出上述系统梯度计算的具体形式：

定理 2（语义策略梯度传播）设 $\mathbb{1}^{V \times 1}$ 为指示函数，仅对采样符号所在维度为设为 1，其余位置为 0。以 $\mathcal{F}_{\text{RX}}^{(t)}(\cdot) \in \mathbb{R}^V$ 表示解码器在 t 时刻的输出，并以 $\widetilde{\mathcal{F}}_{\text{TX}}(\mathbf{m})$ 表示编码器对消息 \mathbf{m} 采样后的结果。目标函数传递至式(3-11)及式(3-12)所定义编解码器输出端的梯度由下式给出：

$$\nabla_{\mathcal{F}_{\text{TX}}(\mathbf{m})} \log(\pi_{\phi}) = \Sigma^{-1} \left[\widetilde{\mathcal{F}}_{\text{TX}}(\mathbf{m}) - \mathcal{F}_{\text{TX}}(\mathbf{m}) \right] \quad (3-13a)$$

$$\nabla_{\mathcal{F}_{\text{RX}}^{(t)}(\cdot)} \log(\pi_{\theta}(\hat{w}^{(t)}|s^{(t)})) = \mathbb{1}^{V \times 1} - \left[p(\hat{w}_1^{(t)}), p(\hat{w}_2^{(t)}) \dots p(\hat{w}_V^{(t)}) \right]^T \quad (3-13b)$$

证明过程见附录B。由于神经网络通常由非线性的复杂模块构成，本文暂仅给出目标函数至网络输出端的梯度结果；关于特定情形下目标函数对 ϕ 及 θ 的梯度解析形式置于附录B。

在实际的优化过程中，本文对编解码器进行交替优化。具体地，在训练编码器时，本文将其余部分视为参数固定而不可训练的部分，并将解码器的解码策略使用贪婪解码替代（使用 **argmax** 方式从 V 维分布 $\left[p(\hat{w}_1^{(t)}), p(\hat{w}_2^{(t)}) \dots p(\hat{w}_V^{(t)}) \right]^T$ 中选择概率最大的语义符号输出）；在训练解码器时，本文类似地将编码器的决策过程视为确定而不可训练的部分，并设置其采样过程中的方差为 0（也即，令式(3-11)中的参数 $\sigma = 0$ ）。此外，本文对于编解码器分别使用式(3-7)中描述的语义策略梯度进行优化，并使用 **self-critic** 方法提供策略梯度中的偏置项估计。整体的训练算法 **SemanticRL-SCSIU** 如**算法 2**所示。

3.5 实验设计和结果评估

3.5.1 实验设置

3.5.1.1 神经网络和参数设定

本文参考了 Farsad 等人的神经网络设计，使用双向长短时记忆网络（**Bidirectional LSTM**）作为编码器，并以单向长短时记忆网络（**LSTM**）作为解码器。在解码端，本文

利用解码器收到的信号 $\mathcal{H}(\mathcal{F}_{\text{TX}}(\mathbf{m}))$ 初始化解码 LSTM 的隐藏状态。在解码的初始时刻 ($t = 0$)，解码器得到固定的输入 “<SOS>” 并开启解码过程；在随后的时间步，用于解码的 LSTM（神经）网络使用上一步输出的语义符号作为当前时刻的输入开展循环解

算法 2 在不可微信道下基于强化学习的语义相似度优化 SemanticRL-SCSIU 算法

输入： 批次大小，初始学习率 α ，self-critic 并行采样次数 M ，输入消息 \mathbf{m} ，预训练轮次 E_p ，训练结束轮次 E_e ，广义语义相似度准则 Θ ，更新标记 $\text{flag} = 0$ ，编码器采样标准差 σ

输出： 编码器参数 ϕ ，解码器参数 θ

// 预训练阶段（可选）

1: 本步骤与 SemanticRL-JSCC 一致，不再赘述

// 基于强化学习的随机迭代优化阶段

2: **for** 训练轮次 = $E_p+1:E_e$ **do**

3: 采样一批次数据

4: **if** $\text{flag}=0$ **then**

5: // 训练编码器

6: 固定解码器的参数（解码端使用贪婪解码策略替代 Softmax 采样策略以获得确定性输出）。编码器对消息 \mathbf{m} 进行编码

7: 对编码信号做功率归一化

8: **for** $i=1:M$ **do**

9: 按照式(3-11)采样得到编码器的一次并行采样输出，再次进行功率归一化，并将其发送至噪声信道中

10: **end for**

11: 解码器使用贪婪解码策略解码，并计算解码消息与原编码器发送消息之间的语义相似度

12: （可选）累积梯度以进一步降低方差

13: 按照式 (3-10a) 和式 (3-13a) 对使用高斯策略的编码器参数 ϕ 进行更新

14: **else**

15: // 训练解码器

16: 固定编码器参数（设置编码器采样标准差 $\sigma = 0$ 以获得确定性输出）。对编码信号做功率归一化，并将其发送至噪声信道中

17: **for** $i=1:M$ **do**

18: 解码器按式(3-12)采样得到一条解码轨迹（输出消息），并计算与编码器发送消息之间的语义相似度

19: **end for**

20: （可选）累积梯度以进一步降低方差

21: 按照式 (3-10b) 和式 (3-13b) 对使用 Softmax 策略的解码器参数 θ 进行更新

22: **end if**

23: 翻转更新标记 flag

24: **end for**

// 训练结束

25: 返回编解码器参数 $\langle \phi; \theta \rangle$

码。该解码过程持续到规定的最大解码长度 N 或输出终止符号 “<EOS>”。

如同在**算法 1** 和**算法 2** 中描述的，在实际的实验中，本文将编码器的输出视为复数向量，并在其进入随机信道前施加平均功率限制。在信道的仿真上，本文参考文献 [97] 使用了加性高斯白噪声（Additive White Gaussian Noise, AWGN）和相位不变衰落（Phase-Invariant Fading）两种模型。

本文将交叉熵作为语义相似度的方法作为基线模型。由于本文所提两种方法（SemanticRL-JSCC 和 SemanticRL-SCSIU）均采用 self-critic 和策略梯度方法训练，所提方法不会引入任何额外参数。也正因此，相较于基线模型本文所提方法有着完全一致的模型结构和参数量，仅仅有着学习方法上的不同。在后续的实验中，本文旨在验证以强化学习构建语义通信系统在语义保真上的可行性和优越性。在语义相似度量指标上，除非特别指出，本文使用自然语言处理中的 CIDEr 指标（式(3-3)）提供语义相似度 Θ_{CIDEr} 。更多的参数讨论可参考后续的3.5.2.1节。

本文注意到后续的一些语义通信方法多使用 Transformer^[98]作为其编解码器。需要指出的是，本文在3.2节设计以广义语义相似度为优化目标的通信架构时，并不对某种具体的模型结构、使用场景以及语义相似度量做任何限制；相反地本文关注的重点在于设计一种通用的架构和具体的学习方法。有关所提方法的泛化性和通用性将在3.5.2.4节做单独讨论。

3.5.1.2 训练细节

以往在通信场景使用强化学习的工作，例如文献 [44,99]，多研究了在低维动作空间下的决策问题。然而，考虑到在高维空间下的探索难度（例如在本文的时序传输场景，每个时间步上的动作空间在 10^4 级别），从零开始训练一个强化学习方法不仅极大地浪费了计算资源，也给模型的训练带来了极大的挑战。除了使用式(3-9)和式(3-13)处理高维空间的梯度估计问题，本文也借鉴了一种常用的强化学习预训练方法^[79,100]，使用监督学习将模型参数引导到合理（但未收敛）的初始值后，再继续使用强化学习方法完成后续的训练过程。特别地，本文在前 87 个迭代轮次中使用交叉熵进行预训练（ $E_p = 87$ ），并在后续的学习过程（至第 200 轮，也即 $E_e = 200$ ）中使用所提 SemanticRL 算法在广义语义相似度下优化。作为对照实验，基线模型则是全程使用交叉熵完成学习。在 SemanticRL-SCSIU 变体中，由于该算法每轮仅更新一半的参数（编码器或解码器），本

文对其使用了接近两倍的强化学习训练轮次（从第 88 轮至第 300 轮）。当然，本文特别指出，预训练仅仅是可选的步骤而不是本文所提算法的核心，其目的在于提高强化学习的速度和学习效率；在低维决策空间下，可以不使用预训练而直接从零开始完成本文所提算法的训练（见3.5.2.4节）。

在具体的超参数设定上，除非特别说明，本文选择训练批次大小为 64，并行采样次数 $M = 5$ ；初始学习率 α 为 10^{-4} ，并在第 160 轮训练后降低为原先的一半。

3.5.1.3 数据集及指标

本章遵循以往语义通信架构研究^[2,49]中最常用的欧洲议会数据集^[101]。该数据集包含了 200 余万条有关国际关系的 21 种语言语料，在语义通信文献中多以英文为研究对象。在预处理阶段，本文将所有的英文字母转为小写并移除了标点符号。在移除超长句和超短句后，本文收集了 88 万余语料，其中 80% 的句子用于训练本文所提语义通信系统，而余下的 20% 用于评估方法的性能。

与以往的研究工作一致，本文使用在语义通信中常用的指标 BLEU、BERT 相似度和单词准确率来衡量。其中 BLEU 和 BERT 相似度量了句子传输的语义相似度，单词准确率和误码率相对。特别地，本文使用学术社区中流行的开源 BERT 模型来计算句子嵌入的相似度。此外，本文额外提供了在自然语言处理中常用的 CIDEr 指标上的性能表现。另外需要指出的是，由于现有语义通信方法几乎都建立在语义信道可微的假设下（对通信问题做了简化），为了对比的公平性，除非特别指出本文将默认报告 SemanticRL-JSCC 的性能。

3.5.2 模型效果

3.5.2.1 语义保真性能

作为本文首要考虑的指标，本文首先对于所提方法的语义保真性能进行了分析。图3.4展示了本文所提 SemanticRL-JSCC 算法和使用交叉熵训练的基线模型在 BLEU 指标下的性能表现。BLEU- n 指标考量了连续 n 个词汇共现的匹配情况，在3.3节中已有所介绍。从 BLEU-1 到 BLEU-4，该指标逐渐考虑到更多的语义上下文，反映了更为高阶的语义相似度。从图3.4中可以首先观察到，所提模型（黄色曲线）在强化学习的训练的框架下，整体上有着稳定的学习曲线并逐渐收敛。尽管在训练的初期由于交叉熵预训练

和强化学习的差异，所提模型存在一定的性能下降，该模型能够迅速按照强化学习所提供的优化方向展开调整，并逐渐朝着更高语义相似度的方向演进。这些实验现象验证了本文使用强化学习优化语义相似度这一做法的可行性。

其次，从 BLEU-1 到 BLEU-4，本文所提方法相较基线模型的优势逐渐明显。这一有趣的现象强力佐证了使用广义语义相似度而非比特级目标函数在保留语义信息，特别是高阶语义关系上的明显优势。特别地，本文观察到所提方法从 BLEU-2 开始相比于直接使用交叉熵的学习模型取得了更好的语义保真效果。虽然所提方法在 BLEU-1 指标上劣于基线模型，我们观察到 BLEU-1 的得分曲线与图3.5(a) 中单词准确率的指标非常接近，也印证了 BLEU-1 更接近于不考虑语义上下文情况下的指标。相反地，由于所提方法以语义相似度而非词汇准确度为优化目标，SemanticRL-JSCC 将更多地侧重于高阶语义的表征，而允许一定的符号错误，与语义传输而非符号传输的目标更为接近。

除了 BLEU 指标，图3.5(b-c) 分别展示了所提方法在 BERT 语义相似度和 CIDEr 上的性能表现。其中，本文所提方法在 CIDEr 指标上的优势最为明显。考虑到本实验中正是采用 CIDEr 指标作为优化目标，所训练得到的模型在同一评价指标上的显著优势并不意外。有趣的是，除了 BLEU 和 CIDEr 这些基于 n 元词组共现频率的经验指标，基于表征学习的语义度量，例如 BERT 相似度，如图3.5(b) 所示，同样遵循着类似的变化趋势。尽管本文并没有使用 BERT 相似度 Θ_{BERT} 作为优化目标，所提方法在 CIDEr 为代表的专家语义相似测度优化下，更好地捕捉到了所传输消息的语义信息。这是因为以往的一些研究方法虽然引入了更为有效的语义提取和表征手段，其比特和符号级别的优化目标与真正的语义传输依然存在差异，最终限制了模型的语义传输能力。总的来说，以

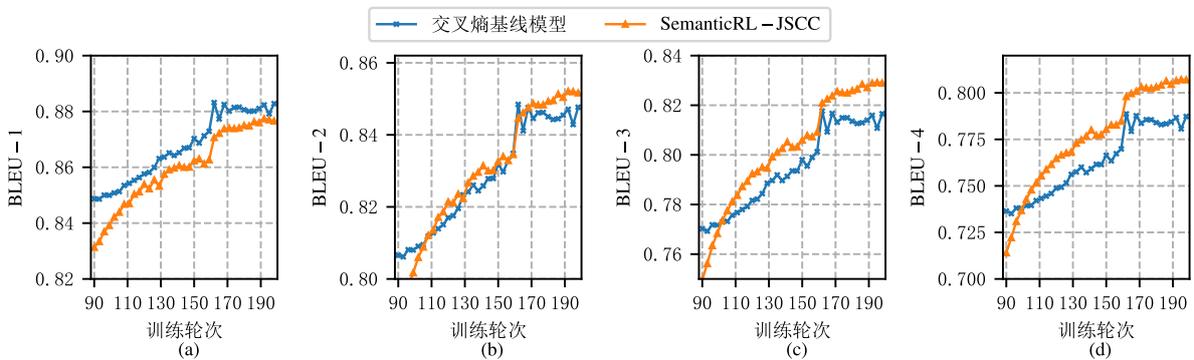


图 3.4 在 BLEU 指标上所提系统与基线模型的性能表现。在 BLEU-1 到 BLEU-4 指标上的结果分别呈现在子图 (a) 到 (d) 中

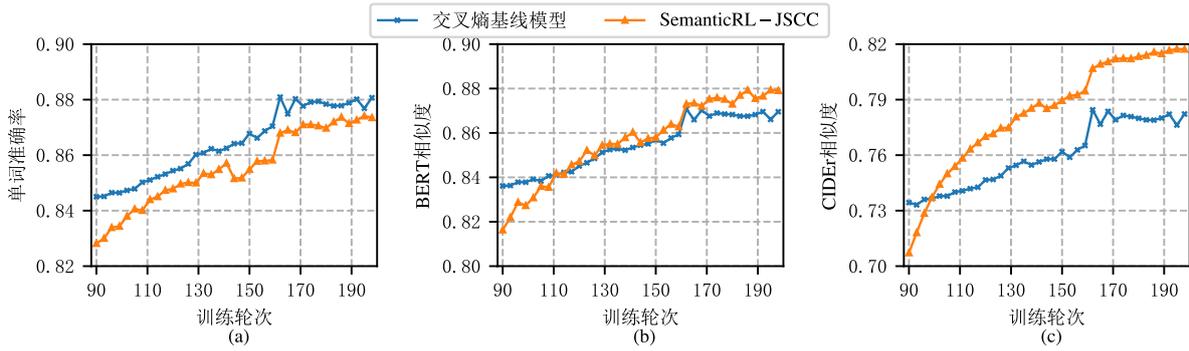


图 3.5 所提 SemanticRL-JSCC 方法在词准确率、BERT 相似度以及 CIDEr 相似度上的定量指标

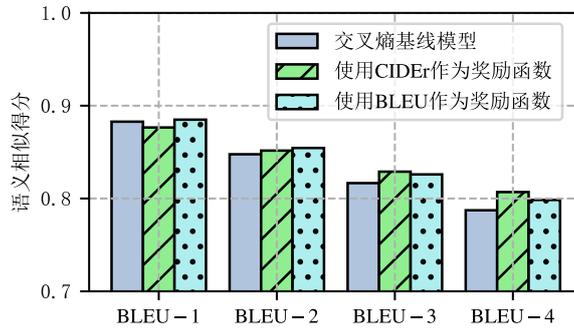


图 3.6 使用不同语义相似度量作为奖励函数对所提系统性能的影响

上的一些定量分析为本文使用语义相似度构建语义通信系统的方法提供了强力支撑，也验证了所提方法相较于比特级别目标函数在语义保真上的优势。

受到在 CIDEr 指标上的启发，本文进一步地对语义相似度的选择开展了一定的研究。在图3.6中，本文展示了使用 BLEU 指标（这里使用 BLEU-1 和 BLEU-4 的均值）、CIDEr 指标和基线模型的对比情况。可以看到，整体上模型会在所优化的特定方面表现更好；另外，无论使用何种广义语义相似度量，所提方法均有着较好的稳定性和适用性。值得一提的是，后续文献 [61] 进一步对于本文所提的语义相似度量提出了具体的实例化形式；亦有文献 [69,91] 从信息新鲜度、错误率等方面对语义相似度的选择展开了讨论。由于本文不再假设相似度量 Θ 的可微性，任何合理的、基于具体任务的相似度量都可以用于指导特定的语义通信系统。

3.5.2.2 噪声鲁棒性

除了语义保真性能之外，本文对于所提方法和基线模型的噪声鲁棒性开展了定量分析。首先，本文考虑了在没有训练-推理差异下的噪声适应性。具体地在这一实验中，本

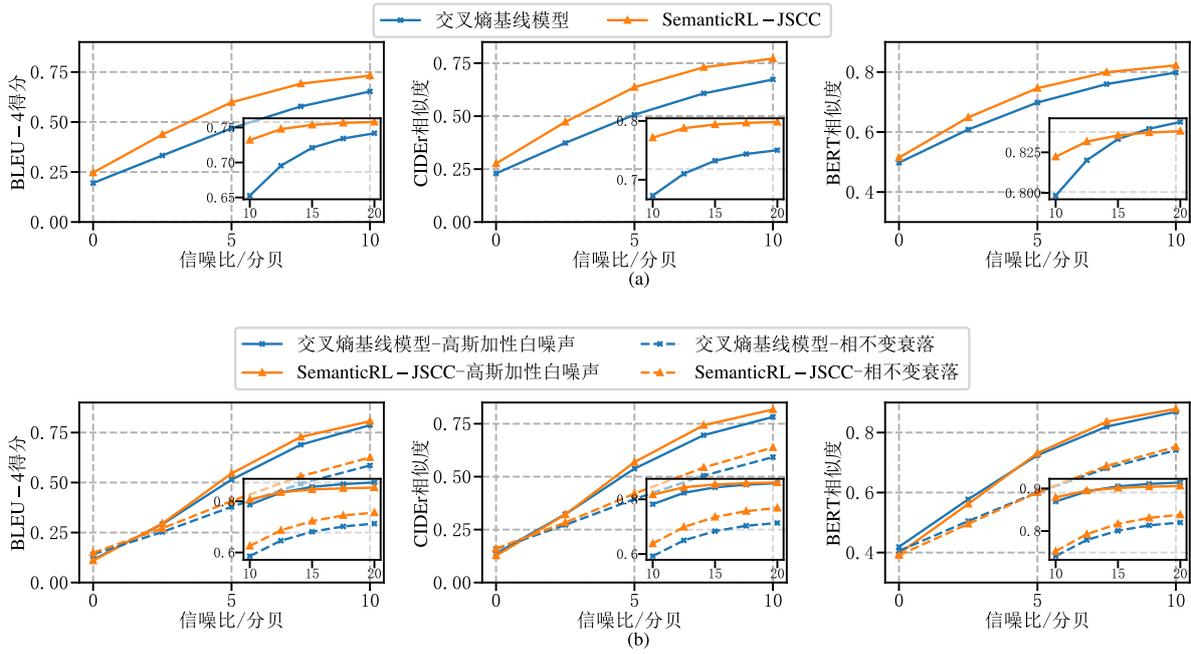


图 3.7 所提系统的噪声鲁棒性研究。子图 (a) 展示了在 0-20dB 随机噪声训练下模型在不同信噪比环境下的适应性。子图 (b) 展示了在固定 10dB 信噪比环境训练后系统于变信噪比环境下推理的性能。为更清晰展示高低信噪比情况下的性能情况，本文将 10 到 20dB 区间单独展示在所有子图的右下角

文研究了 0-20dB 随机噪声训练下的模型在不同测试信噪比下的表现情况，如图3.7(a)所示。在这一实验中，本文观察到：1) 所提方法在几乎所有测试情况下相比基线模型均有着更好的语义相似得分；2) 在边界信噪比情况下，所提方法和基线模型的差距均小于 10dB 附近，这也预示着所提方法在不同信噪比下均有着较好的平衡性和普适性，而非在训练过程中特别专注于某一特定的信噪比环境；3) 在较低的信噪比环境下（例如在 2.5dB 和 7.5dB 附近），所提方法相较于基线模型有着更为明显的优势，也预示着对语义相似度的优化具备在低信噪比环境更好保留语义的潜在可能。

为了更全面地评价所提方法，并进一步评估所提方法的噪声鲁棒性，本文另外研究了两个更为严苛的实验场景：固定信噪比训练-变信噪比推理、以及在此基础上额外测试系统在更复杂噪声信道下的鲁棒性。如图3.7(b)所示，基线模型与所提方法均因训练和推理的差异存在一定的性能退化，但整体而言，以语义相似度为优化目标的模型在低信噪比环境下有着更好的鲁棒性。特别地，在信噪比环境和衰落信道环境（图中虚线）下，所提方法相较基线模型的优势更为显著。对于这一现象，本文推测以语义相似度为优化目标的系统更接近于语义传输的本质，也因此更有可能学习到语义不变的鲁棒表征方式；相反地，以符号复现为目的的通信系统并不具有良好的语义表达能力，受到环境

退化的影响更大。

3.5.2.3 所提两种 SemanticRL 算法的对比

在以上的分析中，本文均使用 **SemanticRL-JSCC** 与基线模型进行对比（两者均假设信道可微，在联合信源信道编码的框架下训练）。但在实际的通信场景中，联合信源信道编码仅仅是相当理想的假设。为了研究当信道不可微时，使用 **SemanticRL-SCSIU** 变体和 **JSCC** 变体的差异，本文以加性高斯白噪声信道为例对于两种模型的训练情况开展了进一步的分析。

在 **SemanticRL-SCSIU** 中，本文将编解码器分别建模为强化学习的智能体，并使用整个系统获得的语义相似度奖励来更新参数。对比于 **SemanticRL-JSCC**，两者共享着相同的解码器学习过程，其主要差异来自于编码器训练的不同。在 **SCSIU** 变体中，编码器面临的主要挑战来自于随机信道对奖励值估计的干扰；而 **JSCC** 变体则是面临着梯度估计的干扰，本文将这两者的差异用式(3-14)表示：

$$\begin{aligned} \text{grad } \mathcal{H}_{\text{TX}} &= h_{\text{JSCC}}^{-1} \cdot (\text{grad } \mathcal{H}_{\text{RX}}) \\ \Rightarrow \text{梯度估计方差} & \end{aligned} \quad (3-14a)$$

$$\begin{aligned} \Theta(\mathbf{m}, \mathcal{F}_{\text{RX}} \mathcal{F}_{\text{TX}}(\mathbf{m})) &= h_{\text{SCSIU}}^{-1} \cdot \Theta(\mathbf{m}, \mathcal{F}_{\text{RX}} \mathcal{H} \mathcal{F}_{\text{TX}}(\mathbf{m})) \\ \Rightarrow \text{奖励估计方差} & \end{aligned} \quad (3-14b)$$

其中 $\text{grad } \mathcal{H}_{\text{TX}}$ 表示编码器端靠近信道部分的梯度， $\text{grad } \mathcal{H}_{\text{RX}}$ 表示在接收端靠近信道部分的梯度； $h_{(\cdot)}^{-1}$ 表示信道响应函数的逆。

从式(3-14)可以看到，无论是 **SCSIU** 还是 **JSCC** 变体，它们对于信道干扰的估计都是无偏的，但存在一定的方差干扰。也因此理论上 **SCSIU** 和 **JSCC** 变体一样都可以收敛，但存在收敛速度和收敛上界上的差异。可以预见的是，当不再假设随机信道 \mathcal{H} 的可微性后，对信道的估计和适应将变得更为困难。

本文在图3.8展示了 **SCSIU** 训练的收敛情况。在这一实验中，本文研究了与 **SemanticRL-JSCC** 模型使用相同的预训练（**Pre-training**）初始化情况下，**SCSIU** 算法编解码器的语义相似度随训练轮次的变化曲线。可以观察到，**SCSIU** 变体即便是在信道和优化目标均不可微的情况下，编解码器依然能够稳定学习并趋于收敛。另外，如表3.1中 **SCSIU-PT**

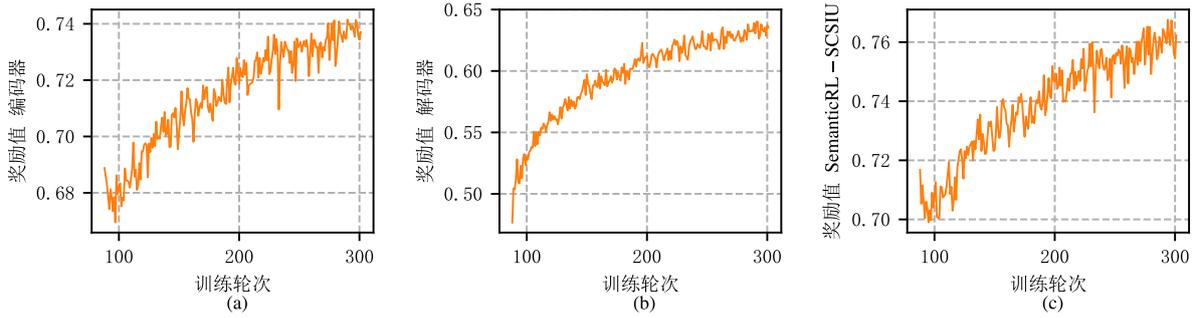


图 3.8 SemanticRL-SCSIU 的收敛情况。子图 (a) 到 (c) 分别展示了训练时编码器、解码器和整体的 CIDEr 奖励值得分

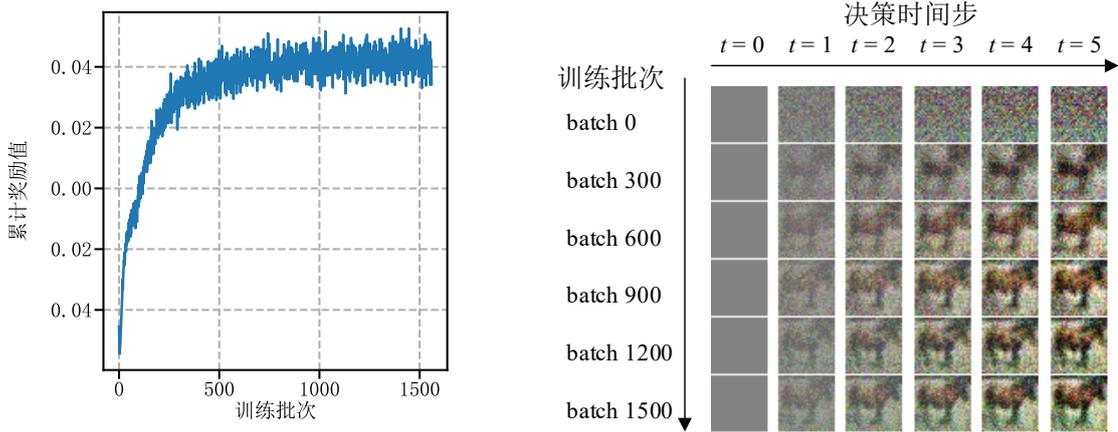
(PT 表示预训练) 所示, 由于信道不可微的设定, 网络探索和适应信道的难度加大, 在与 SemanticRL-JSCC 相近的编解码器总更新次数下表现略逊色于 JSCC 变种。另一方面本文推测也归因于在编解码交替优化的情况下, 编解码器的更新方向有着一定的偏差, 从而需要更多的迭代周期; 此外不可微信道的优化本身给学习过程带来了更大的挑战, 也是导致 SCSIU 算法收敛速度慢的一大原因。为了更进一步地研究 SCSIU 变体的可能上限, 本文对于 SemanticRL-JSCC 变体使用 SCSIU 算法进行了微调 (Fine-Tune, FT), 如表 3.1 所示, 该微调模型 SCSIU-FT (FT 表示在 SemanticRL-JSCC 最终训练所得模型上进行微调) 相较于 JSCC 本身有着一定程度的提升。但值得说明的是, 受限于不可微的假设以及强化学习本身的一些训练技巧, 直接使用 SCSIU 算法在收敛速度和收敛上限上未必能达到和信道可微假设下相同的性能。对于 SemanticRL 算法中提升性能的一些技巧将在 3.5.2.4 节做讨论, 本节展示的是最朴素的 SemanticRL-SCSIU 算法, 主要目的在于验证该算法的可行性。

3.5.2.4 方法的泛化性敏感性

在本节中, 本文将对所提 SemanticRL 算法的泛化性和通用性做更为广泛的讨论, 其中本文展示了所提算法在不同模态 (文本、图像)、不同网络及传输方案、不同语义度量下的通用性。并将对所提算法中的一些关键设定, 例如是否需要进行预训练, 关键参

表 3.1 所提 SemanticRL 算法不同变体的比较

所提 SemanticRL 算法变体	BLEU-1	BLEU-2	BLEU-3	BLEU-4	CIDEr	BERT 相似度
SemanticRL-JSCC	0.877	0.852	0.829	0.807	0.817	0.879
SemanticRL-SCSIU-PT	0.849	0.812	0.780	0.749	0.747	0.840
SemanticRL-SCSIU-FT	0.882	0.858	0.836	0.815	0.824	0.884



(a) CIFAR-10 数据集上时序恢复方法的训练曲线 (b) 随训练进行的图像恢复情况视觉效果

图 3.9 在 CIFAR-10 数据集上对图像使用时序循环渐进解码恢复的训练曲线和视觉效果。所展示场景以 10 分贝信噪比高斯白噪声训练并以图像恢复时相邻时间步之间的均方误差增益作为语义相似度度量

数的动态更新做简要分析。

作为最主要的考量，本文首先测试了所提算法在句子传输之外的适用性。在这一问题中，本文考虑了联合信源信道编码中常用的图像传输场景，在语义通信领域较为常用的 CIFAR-10 数据集^[102]上对如何使用强化学习建模图像传输这一问题做简要讨论。该数据集包含 50000 张 32×32 大小的彩色训练图像，以及 10000 张测试图像。就本文所提 SemanticRL 框架而言，本文展示了对于图像传输的两种可行建模方法：

- 将图像的恢复建模为时序的循环决策解码过程。该方法允许图像被渐进修复，适用于算力和带宽受限而时间要求不严格的场景。
- 将图像的恢复建模为时间步为 1 的连续空间决策过程，该方法与 SemanticRL-SCSIU 中编码器的优化类似，适用于时间要求严格的场景。

以上两种建模方法都可以实现对不可微语义相似度的优化。为简单起见本文在本小节假设信道可微，研究所提 SemanticRL 算法在多模态场景的适用性。

对于第一种建模，本文参考文献 PixelRL^[103]的思路，将图像进行量化后，定义每个时间步上的决策空间为 {像素增，像素减，像素不变}，并逐时间步地对图像的像素数值进行逼近。其决策过程的奖励函数可以定义为相邻时间步上均方误差或图像结构相似度上的改善程度。例如，在对图像做 11 种像素等级量化并归一化到 $[0, 1]$ 之间（也即，像素取值在 $\{0, 0.1, 0.2, \dots, 1.0\}$ ）后，本文从初始时刻所有像素为 0.5 出发，在 5 个时间

步内可以实现对原输入图像的渐进恢复，如图3.9所示。可以观察到，所提 **SemanticRL** 算法是一种通用的使用强化学习优化语义相似度的方法，不受限于具体的任务场景、某一主干模型或具体的语义相似度量，在图像空间同样可以正确恢复出消息的语义信息。特别地，本文指出（相对于对图像的一步恢复）这种渐进传输的方式虽然需要经历多次解码才可获得较为细腻的结果，在解码的初期已隐约可按照解码轮廓辨别所传输语义的可能结果；并且在每个时间步 **SemanticRL** 的决策子任务是简单的。对符号消息的渐进传输是学术界研究较少的内容，本文所提方法也为扩充渐进传输的通信方法提供了一定的参考。

对于第二种建模，通过将动作规定在连续空间中，方法可以实现在一个时间步范围内对图像的恢复，此时的奖励函数可直接定义为解码图像与输入图像的语义相似度。特别地，本文使用图像的结构相似度 **SSIM** 作为语义相似指标，并使用式(3-11)所述策略进行解码。所提方法在 **CIFAR-10** 数据集上的训练曲线如图3.10所示。

需要指出的是，本文所提方法并不局限于何种任务、使用何种具体的语义提取模型以及何种语义相似度量。将特定的场景建模为强化学习的决策过程（本文对于离散空间和连续空间均进行了举例），并使用合适的广义语义相似度量提供奖励函数后，一般的语义通信问题均可使用本文所提方法进行优化。

本文方法并不严格要求如3.5.1节所述的预训练过程。如图3.10所示，在低维空间（这一实验中的低维体现在时间步为1）和合适的模型设定下，所提方法可以从随机初始化的模型开始逐渐学习并收敛到最优值。相反，当动作空间和决策时序较大时，使用监督学习方法预先缩小模型空间上的搜索域，将大大提升模型在解空间搜索的收敛速度，并提供较为鲁棒的起始解。

此外，本文在 **CIFAR-10** 场景下，对于影响强化学习探索过程的参数 σ 进行了敏感性讨论。图3.10展示了使用固定的标准差 0.1、手动降低标准差 σ （退火）、以及将 σ 作为可学习参数一并加以学习三组实验进行了对照。可以看到，在 **SemanticRL** 中，并行采样策略的探索与利用是影响最终收敛上界的一大因素；当探索的程度（与标准差 σ 成正比）较大时，网络倾向于对不同的动作选择进行探索从而迅速跳出劣势解；相比之下较大的探索程度并不利于网络的收敛过程，而呈现出在较优解附近震荡的模式。然而需要指出的是，尽管可以人为地对于 **SemanticRL** 中有关强化学习部分的诸多参数进行调整，这些调参的优化或改进并非本文关注的重点。一般而言，使用**算法 1**和**算法 2**中的

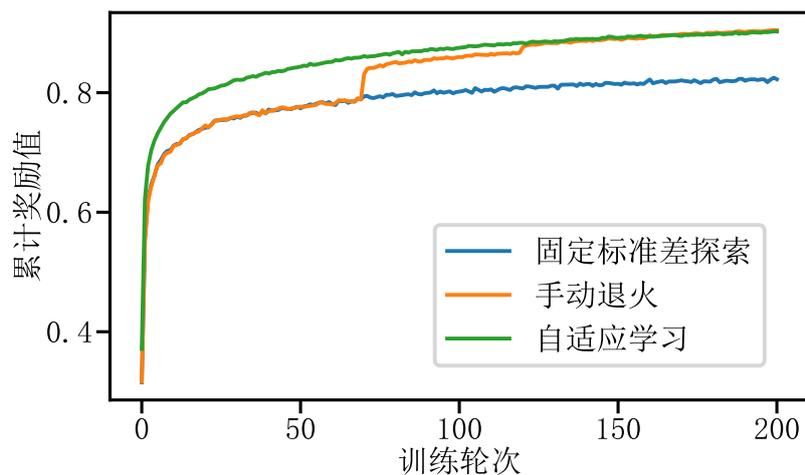


图 3.10 在 CIFAR-10 数据集上将图像传输建模为连续动作空间决策后，所提 SemanticRL 方法的适用性及不同训练方式的差异

策略已可实现基本的语义通信系统，一些关键的参数，例如语义相似度指标 Θ 的选择以及具体的强化学习学习策略属于优化的范畴或特定领域的研究，本文对于该部分参数进行敏感性分析，但仍然需要指出具体的设定应由任务而异。

4 基于语义表征的异构语义网络

4.1 本章概述

在第3章中，本文从现有语义通信系统在语义保真方面的缺陷出发介绍了一种使用广义语义相似度构建语义通信系统的方法，并进一步提出了一种使用强化学习对不可微语义相似度进行优化的策略。然而在语义通信领域，几乎所有的工作（包括在第3章中介绍的方法）均是建立在某一具体的点对点语义通信基础上，考虑单一、特定收发对象通信中的优化问题。

本文所设想的语义通信系统以语义为通信媒介，不仅要求在点对点传输上尽量做到语义保真，也意图构建一种考虑在多用户、异构语义知识场景下的语义网络架构。在文献中，多用户或多模态的语义通信场景直到最近一段时间才有所讨论，例如文献 [72] 为不同的用户和任务设计独立的编解码器，从而实现多编码器-多解码器的通信方案；在文献 [4] 中研究人员提出了多编码器-单解码器的设想。尽管这些方法提出了各自的多用户或多模态通信方法，以上的方法依然存在着许多问题：

- 一对一通信系统可以通过直接训练学习得到，但考虑到多用户和异构语义知识的场景，为每一用户都训练其自己的编解码器将变得极为昂贵并难以部署。
- 在多用户场景下，存在着窃听、语义攻击等特有的问题，这一点常常在现有的研究中被忽略。

本文设想的语义网络并不局限在实现多用户通信这一层面。本文认为语义网络是阐述语义的网络，这也就要求语义网络中传输的所有信息，应以统一的、与具体任务和本地知识无关的语义形式存在；在此基础上，网络应能按照不同用户的需求对达成特异性响应。此外，在多对多场景中，语义网络能利用语义本身的理解差异提供全新而又向下兼容的加密方案。

针对于现有问题的不足和本文对语义网络的设想，本文提出了在多用户和异构语义场景下的统一的语义网络架构。本文使用的“统一”一词旨在说明本文以统一的语义表征作为语义网络中的信息流（不论用户为何，任务为何）的特征，而非为不同用户定

制其各自的语义表征。特别地，在本文设想的语义网络中，语义流仅对语义负责；不同形式的同类型语义共享相同的语义表征，而将用户作为语义表征的实例。所提语义网络的整体架构如图4.1所示。其中语义网络将异构的符号表达映射到统一的语义表征空间，使用归纳（Summary）函数提取其本原语义并进行传输（图4.1上半部分）；在语义的表达上，本文使用表达（Speak）函数结合异构用户的语义知识对本原语义做特异化表达（图4.1下半部分）。

本文所提语义网络“对语义负责，对用户特异”的性质从根本上区别于以往“对用户负责，对语义特异”的设计思路。由于所提方法以统一而通用的语义本原作为语义网络中的数据流，本文将命名为本原语义网络（Primitive Semantic Network, PSN）。

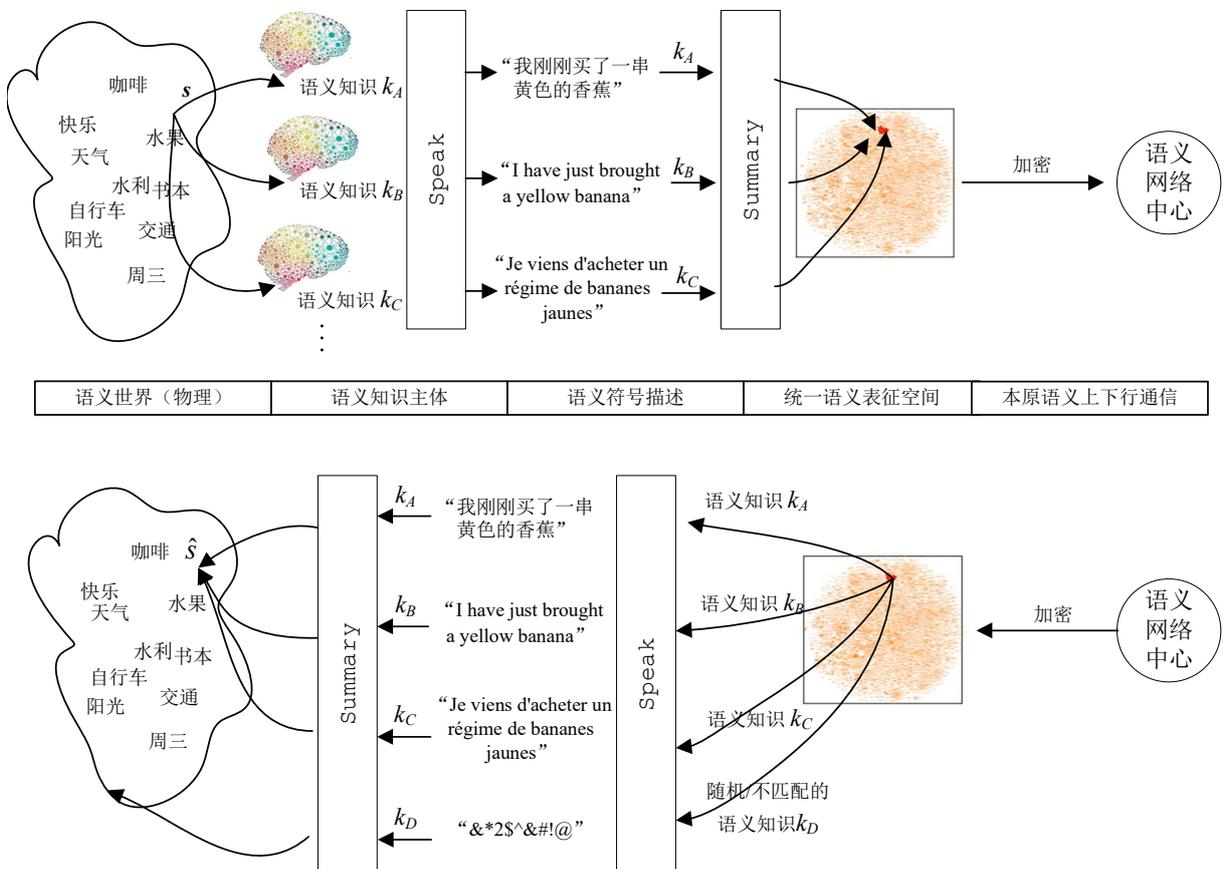


图 4.1 本文所提面向多用户异构语义通信的本原语义网络架构图

4.2 传输本原语义的异构网络建模

本文认为语义是唯一需要传输的信息媒介。基于这一设想，同一语义信息允许出现不同的符号描述（例如在异构的中文和英文语义知识中；或是在同一种语义知识中不同

的符号描述)。为后续讨论的方便,本文不失一般性地认为不同的用户具有不同的本地语义知识 $k_{(\cdot)}$, 并可使用表达函数 Speak 将该语义信息进行描述。沿用本文在上一章节中对于消息 m 的表示, 对于某一语义 s , 其在不同语义知识中的符号描述可以表达为 $m = \text{Speak}(\langle s, k \rangle)$ 。尽管本原语义 s 在不同的用户中存在不同的符号表达方式 m , 这些符号消息本质上都是同一语义意向的描述。

基于以上的特性, 本文设想一种能够从不同描述 m 中还原语义表述 s 的方法 Summary , 从而为在不同知识体系下使用语义空间 \mathcal{S} 中的相同表征提供可能。在这一框架下语义异构的用户只需要将所期望表述的本原语义在网络中传输, 而忽略掉与语义无关的信息。对于接收方来说, 可以使用 Speak 函数将收到的语义使用其特异的知識表达进行转述, 如图4.1所示。本文将以上的一次语义符号传输用公式表达为:

$$s = \text{Summary}(\langle m, k_A \rangle) \quad (4-1a)$$

$$\hat{s} = D[\mathcal{H}(E(s))] \quad (4-1b)$$

$$\hat{m} = \text{Speak}(\langle \hat{s}, k_B \rangle) \quad (4-1c)$$

其中下标 A 和 B 用以区分表示语义异构的不同信源(用户); \mathcal{H} 、 D 以及 E 与第3章中类似地分别表示通信信道、编码器和解码器(本节仅要求编解码器能做到准确传输即可, 因此未沿用第3章中语义编解码器 \mathcal{F}_{TX} 与 \mathcal{F}_{RX} 的符号表述)。值得注意的是, 由于本章主要考虑语义网络的构建和语义的表征, 后文将不再对物理层传输过程(式(4-1b))做特别讨论。

式(4-1)以两个异构语义信源为例介绍了本文所设想的语义网络。相比较于在第3章介绍的点对点通信模型, 本章更关注于语义的统一表征(式(4-1a))和特异性表达问题(式(4-1c)), 而非物理层的点对点传输(式(4-1b))。从式(4-1b)中, 可以清楚地观察到在本文所设想的语义网络中, 一切符号消息都以“语义流”的形式进行传达。特别地, 当式(4-1)中收发双方的语义知识 k_A 与 k_B 相同时, 该式所表达的语义转述问题退化为第3章中介绍的自编码问题。

于是, 在上述的语义网络框架中, 本文所需要解决的问题成为了如何准确而高效地建模 Summary 和 Speak 函数。

4.3 统一语义表征与特异用户表达

4.3.1 统一的语义表征模型

与已有的一些多用户语义通信方法不同, 本文对不同的信息载体使用统一的语义表征模型 **Summary** 加以表述。为实现这一目标, **Summary** 函数需要具备如下的两个特点: 1) 对于语义接近的表达, 其在语义空间中应具有较近的距离; 2) 对于语义不同的表达, 在语义空间中应尽可能远离。本文以 **Sim** 表示在投影语义空间上的某一统计相似性度量, **Summary** 函数的优化目标可用下式表述:

$$\text{Summary}^* = \begin{cases} \operatorname{argmax} & \text{Sim}(\text{Summary}(\langle \mathbf{m}, \mathbf{k} \rangle), \text{Summary}(\langle \mathbf{m}, \mathbf{k} \rangle^+)) \\ & \langle \mathbf{m}, \mathbf{k} \rangle^+ \text{表示具有相近语义的符号-语义知识组合} \\ \operatorname{argmin} & \text{Sim}(\text{Summary}(\langle \mathbf{m}, \mathbf{k} \rangle), \text{Summary}(\langle \mathbf{m}, \mathbf{k} \rangle^-)) \\ & \langle \mathbf{m}, \mathbf{k} \rangle^- \text{表示具有不相似语义的符号-语义知识组合} \end{cases} \quad (4-2)$$

其中, 具有相近语义的符号-语义知识是较为严格的要求。语义不匹配则是较为常见和易于发生的情况, 例如位于同一语义知识体系中无关的符号消息, 或相同的符号消息在其他语义知识体系中被转述时产生了语义错误和偏差 (这种表达偏差在文献 [5,9] 中被称为一种语义噪声)。

式(4-2)涉及到两个优化目标。尽管最小化语义接近样本之间距离的优化策略具有唯一解, 最大化不匹配样本之间的距离却具有无穷多解, 以同样的力度施加这一惩罚可能导致优化目标偏离正确方向 (对系统来说, 最大化距离带来的增益可能远大于最小化同语义间的距离)。对于以上的优化目标, 对比学习^[104-105]指出可以通过构造正负匹配样本的批次采样方法对其中的正样本加以奖励, 并对负样本 (不同语义的样本) 的匹配程度进行抑制。

具体地, 对于本原语义对应于 \mathbf{s} 的某一 “⟨符号消息 \mathbf{m} - 语义知识 \mathbf{k} ⟩” 对, 本文从其余消息-语义知识对中采样与其语义相符 (通常通过事先标定得到, 例如后续重点研究的中-英双语翻译问题, 或使用伪标签获得) 的一对样本构造为正样本对, 并随机采样 $B - 1$ 个非匹配样本构造负样本对。可以预见的是, 在理想情况下, 语义匹配样本之间的相似度应显著高于非语义匹配情况下的相似度得分, 于是本文在这 B 个样本上对

其匹配的可能性使用 Softmax 概率进行描述:

$$p_{\text{match}}(\langle \mathbf{m}, \mathbf{k} \rangle, \langle \mathbf{m}, \mathbf{k} \rangle_i) = \frac{\exp[\text{Sim}(\text{Summary}(\langle \mathbf{m}, \mathbf{k} \rangle), \text{Summary}(\langle \mathbf{m}, \mathbf{k} \rangle_i))]}{\sum_{j=1}^B \exp[\text{Sim}(\text{Summary}(\langle \mathbf{m}, \mathbf{k} \rangle), \text{Summary}(\langle \mathbf{m}, \mathbf{k} \rangle_j))]} \quad (4-3)$$

其中, 本文借鉴文献文本-图像语义匹配任务中常用的乘性注意力策略^[40], 式(4-3)中的 Sim 函数由单层全连接神经网络 f 以及向量点乘得到:

$$\text{Sim}(\text{Summary}(\langle \mathbf{m}, \mathbf{k} \rangle), \text{Summary}(\langle \mathbf{m}, \mathbf{k} \rangle_i)) = f(\text{Summary}(\langle \mathbf{m}, \mathbf{k} \rangle)) \cdot f(\text{Summary}(\langle \mathbf{m}, \mathbf{k} \rangle_i)) \quad (4-4)$$

在实际的场景中, 由于正例样本通常远小于负例样本, 所得到的 p_{match} 通常具有较为尖锐的分布。在文献 [104] 和文献 [105] 中, 研究人员指出了构造这种在对比学习中使用单正例样本和多负例样本与最大似然估计方法在大批次 B 条件下的一致性。约束语义联合表征的配对损失 $\mathcal{L}_{\text{Match}}$ 由式(4-5)给出

$$\mathcal{L}_{\text{Match}} = \mathbb{E}_{\langle \mathbf{m}, \mathbf{k} \rangle} \left[\sum_{i=1}^B -\mathbb{1}(i) \log p_{\text{match}}(\langle \mathbf{m}, \mathbf{k} \rangle, \langle \mathbf{m}, \mathbf{k} \rangle_i) \right] \quad (4-5)$$

其中 $\mathbb{1}(\cdot)$ 与第3章一致表示指示函数, 只有当某 (消息 - 知识) 对的内在语义与给定的 $\langle \mathbf{m}, \mathbf{k} \rangle$ 一致时才为真。本文主要研究跨知识体系的语义匹配问题, 并主要以异构语义下的文本作为联合语义表征的研究对象, 但同时也指出使用对比学习方法在其他领域, 如文本-图像多模态场景下的通用性^[40]。

不同于现有的一些以自编码准确率为优化目标的点对点语义通信系统, 语义网络需要具有较高的拓展性和良好的语义表征能力, 也就是“知其然而知其所以然”。现有的一些工作多强调了准确性却忽视了语义表征的可解释性和拓展性。本文进一步借鉴在表征学习中常用的自监督学习方法, 将消息中的一部分内容使用掩膜 (Mask) 进行覆盖, 并使用其余部分预测所缺失的语义内容, 从而对语义网络的训练施加额外的语义推理监督, 更进一步完善语义网络对消息中各部分语义实体及关系的理解和表达能力:

$$\mathcal{L}_{\text{MLM}} = -\log p(w_{\text{mask}} | \text{mask}(\mathbf{m})) \quad (4-6)$$

其中 mask 表示掩膜操作, “MLM” 取自英文 “Masked Language Modeling” 的开头字母缩写。这里与以往一些语义通信方法^[2,50]的最大不同在于, 所述方法不仅仅需要完整对消息进行传输, 还需要对于消息中的每一部分语义内容进行理解和推断。本文所提表征

语义学习方法也是在语义通信领域除自编码准确性约束之外使用自监督学习优化语义表征的首次尝试。

4.3.2 特异的用户表达模型

在4.3.1节中, 本文将同一语义的不同表达映射到同一语义空间, 从而为不同语义载体的统一表征提供了可能。进一步地, 考虑解码端多用户的场景, 需要对于语义网络中的联合语义表征还原到对应的异构符号空间。本节主要研究 **Speak** 函数 (最终将以神经网络建模) 的设计。

为实现不同语义知识下同一语义 \hat{s} 的表达问题, 本文将不同的语义知识 $\mathbf{k}_{(\cdot)}$ 使用可学习的参数加以建模, 并将其作为语义网络输入的一部分构造条件生成模型 $\text{Speak}((\hat{s}, \mathbf{k}))$ 。通过改变 \mathbf{k} , 表达模型将在不同的语义知识体系中输出对应的符号消息描述。

在本文的表达模型中, 本文认为语义在所有用户和任务是共享的, 而造成特异表达的关键在于不同的语义知识 \mathbf{k} 。于是, 本文从收发双方语义知识的匹配性出发, 对于 **Speak** 网络的功能设计做了以下的两点讨论:

- 同构语义知识下的语义表达问题。收发双方具有相同的语义知识, 此时语义编解码等同于普通自编码任务。
- 异构语义知识下的语义表达问题。收发双方具有不同的语义知识, 此时语义解码等同于语义转述任务。

注意到, 普通语义自编码任务是点对点语义通信场景中的核心 (主要) 内容, 而异构的语义表达属于语义网络中独有的任务, 是对于点对点语义通信在通用性上的拓展, 也是本文相较于点对点通信, 以及现有多对多语义通信方法的贡献之一。

同构自编码任务假设收发双方有着相同的语义知识体系。在解码的每个时间步, 自编码旨在对输入消息本身进行在某一语义相似测度 Θ 下的重建:

$$\mathcal{L}_{\text{AE}} = -\Theta(m, \hat{m}) \quad (4-7)$$

其中下标 AE 是自编码器 (Autoencoder) 的缩写。

异构转述任务假设收发双方有着不同的知识结构和语言 (符号描述) 体系。考虑 A, B 两种不同的语义体系, 当发送端将消息 m_A 以语义 s 进行传输后, 解码端可使用不

同的知识体系 k_B 对收到的本原语义 \hat{s} 进行转述（或翻译）成为 $\hat{m}_B = \text{Speak}(\langle \hat{s}, k_B \rangle)$ 。转述任务期望 m_A 与 \hat{m}_B 具有一致的符号语义和语义空间表征：

$$\mathcal{L}_{\text{Tr}} = - \left[\underbrace{\Theta(m_A, \hat{m}_B)}_{\text{符号语义一致性}} + \underbrace{\text{Sim}(\text{Summary}(\langle m_A, k_A \rangle), \text{Summary}(\langle \hat{m}_B, k_B \rangle))}_{\text{语义空间一致性}} \right] \quad (4-8)$$

其中损失函数下标 Tr 是英文单词“翻译”（Translation）的开头字母。

4.3.3 语义网络安全性

安全性通常不是一对一通信场景下被主要考虑的因素。然而考虑多对多的语义通信问题，窃听和对抗攻击成为了语义网络实用性的实际挑战。尽管前文所述以语义为传输对象的语义网络架构极大地便利了语义的统一表征，一旦本原语义 \hat{s} 或对应不同语义知识的 k 遭到泄露，将存在着语义不安全的潜在问题。

基于语义安全性考虑，本文另外设想了一种语义安全的语义网络架构，并具体提出了语义知识加密和本原语义加密两种互补方案加以处理。其中，所提出的语义知识加密方法展示了一种不同于经典加密技术的新加密方案；而本原语义加密展示了所提系统对于以往加密方案的向下兼容性。基于所提方法，文本展示了所提语义网络抵抗语义系统外攻击以及语义系统内部其他用户窃听的良好能力。

4.3.3.1 语义知识加密

考虑一个在多用户场景下使用异构语义知识 k 对接收端本原语义 \hat{s} 进行表达的解码方案 $\hat{m} = \text{Speak}(\langle \hat{s}, k \rangle)$ 。对于同一本原语义 \hat{s} ，在不同的语义知识下应具有差异性的表达。特别地，对于具有语义知识为 k_A 的解码主体 A ，当且仅当 Speak 函数中的语义知识正确匹配 k_A 时才可正确解码出在语义主体为 A 下所对应相似度量 Θ_A 的符号表达。

$$\begin{aligned} & \max \Theta_A(m_A, \text{Speak}(\langle \hat{s}, k_A \rangle)) \\ & \min \Theta_A(m_A, \text{Speak}(\langle \hat{s}, k_{\neg A} \rangle)) \end{aligned} \quad (4-9)$$

其中 $k_{\neg A}$ 表示除了 A 以外的语义知识体系。以中文传输为例，只有当对应中文体系的语义知识 $k_{\text{中文}}$ 正确输入后才可得到正确的中文描述，使用其他语义知识（如 $k_{\text{英文}}$ ）进行解

码虽然可能得到对应的符号转述，但将无法得到满足 $\Theta_{\text{中文}}$ 的中文消息描述。式(4-9)给予了 \mathbf{k} 在语义安全场景下的解释：对于异构知识结构的用户，其本地语义知识相当于正确解码出对应符号表达的语义密钥。

进一步地，考虑来自系统外的随机语义攻击情况。首先需要指出的是，由于语义知识并非人为指定的参数向量，该参数可随着网络进行实时学习和更新，也并非需要遵循任何已知的概率分布，通过随机生成语义密钥进行破译的方式几乎不可能实现。在更为宽松的条件下，倘若某一窃听者已经获取到部分解码消息 $\hat{\mathbf{m}}$ 以及输入语义 $\hat{\mathbf{s}}$ 的对应关系，窃听者可通过强化学习或监督学习的方式逆向破译出对应用户的语义密钥 $\mathbf{k}^* = \operatorname{argmax}_{\mathbf{k}} \Theta(\hat{\mathbf{m}}, \operatorname{Speak}(\langle \hat{\mathbf{s}}, \mathbf{k} \rangle))$ 。针对这一可能，一种可行的设计思路是使得 \mathbf{k} 的参数空间尽可能崎岖，不利于普通梯度下降破译方法的更新过程。具体的，本文给出一种使用随机噪声构建非平滑语义密钥空间的方法：

$$\begin{aligned} \max \quad & \Theta(\mathbf{m}, \operatorname{Speak}(\langle \hat{\mathbf{s}}, \mathbf{k} \rangle)) \\ \min \quad & \Theta(\mathbf{m}, \operatorname{Speak}(\langle \hat{\mathbf{s}}, \operatorname{Dist}(\mathbf{k}) \rangle)) \end{aligned} \quad (4-10)$$

其中 $\operatorname{Dist}(\cdot)$ 表示对某一参数的随机扰乱。该式通过将探索空间位置 \mathbf{k} 的邻域构造为高峰和低谷的形式将参数空间逆平滑处理，以增强对抗攻击的难度。

对于式(4-9)及(4-10)所示的对比学习思想，本文使用文献 [106] 中提出的连续空间对比损失加以约束：

$$\mathcal{L}_{\text{Sec}} = \|\max(\epsilon - \Delta, 0)\|^2 \quad (4-11)$$

其中 Δ 表示式(4-9)和(4-10)中需要最大化目标与最小化目标的数值差异（相似度 Θ 之差）， ϵ 为设定的正常数，称为边界（Margin）；损失函数下标 Sec 为安全（Security）的开头字母。式(4-11)旨在鼓励正样本和负样本拉开至少 ϵ 的得分（此处为语义相似度）差距。

4.3.3.2 本原语义加密

上一节中本文对于语义知识 \mathbf{k} 的设计实现进行了讨论，提供了一种在语义层面进行加密的方法。此外，所提语义通信系统同样可以向下兼容已有的一些加密方法。

特别地,为防止来自语义系统内的语义泄露(例如,用户 B 窃听了 A 收到的本原语义 \hat{s} ,并使用 k_B 进行解码实现窃听),往往还需要对发送到特定接收方的语义本原在传输过程中进行加密。本文在这一小节简单设计了一种在上下行链路均使用经典加密方法对本原语义使用符号加密,并在多用户系统中使用语义知识进行语义加密的整体框架。

如图4.1所示,本文在这一系统中设想了一种中心网络(以此为例,亦可构造为去中心网络),不同用户和中心服务器之间可使用各类经典加密方法(如 MD5、SHA、RSA 等)对本原语义进行加密;语义表征下发到其他用户时,亦可利用接收端与中心服务器的协议进行再次加密。在非中心网络中,经典加密可由收发双方协商确定。特别地,即便接收端窃听者获取了正确的语义本原表征 \hat{s} ,还需要提供正确的语义知识 k 方可完成正确的符号解码。考虑到所提框架对于经典加密方法的兼容性,不论是传输在语义网络中的语义信息流,还是用于解码的语义知识,均可实现密钥的动态更新。

注意到,本原语义加密属于所提语义网络对经典加密方法的一种兼容性拓展,在实际系统的学习(训练)过程中并不需要特别考虑。本文在异构语义网络安全性上的主要贡献体现于语义加密概念和语义网络安全性概念的提出,以及运用语义手段提供新加密方法的可行性验证上。

最后,本文给出所提语义网络 PSN 的整体训练损失函数:

$$\mathcal{L}_{\text{PSN}} = \mathcal{L}_{\text{AE}} + \mathcal{L}_{\text{MLM}} + \mathcal{L}_{\text{Tr}} + \mathcal{L}_{\text{Match}} + \mathcal{L}_{\text{Sec}} \quad (4-12)$$

其中,从左至右,整体损失分别由自编码损失、掩膜推理损失、转述损失、语义匹配损失以及语义安全损失组成;其中各部分的消融实验将在4.4.2节介绍。值得注意的是,绝大多数点对点语义通信网络仅仅考虑到自编码损失^[63,65-67];本文后续将进一步分析式(4-12)在语义网络中起到的作用。本文将所提“统一表征,特异表达”的异构语义网络训练算法总结在**算法 3**。

4.4 实验设计和结果评估

4.4.1 实验设置

本文考虑一种典型的异构语义场景:多语言语义通信的实现。在这一系统中,本文特别选择了中英两种语言为研究对象;两种异构语义具有不同的语法结构和词汇表达,

算法 3 面向异构语义的本原语义网络 PSN 训练算法

输入: 异构语义符号消息 m ., 学习率 α , 本原语义提取模型 Summary, 语义到符号的特异表达模型 Speak, 广义语义相似度准则 Θ , 随机初始化的异构语义知识 $k_{(\cdot)}$

输出: 参数化表征的模型 Summary、Speak 以及 $k_{(\cdot)}$

```

1: while 未达最大迭代次数或未收敛 do
2:   for 对每一语义体系遍历, 以优化语义体系  $A$  为例 do
3:     采样得到来自某一语义体系  $A$  下的符号消息  $m_A$ 
       // 自编码和语义推理
4:     使用语义相似度  $\Theta$  对自编码任务以式(4-7)优化
5:     按照式(4-6)对掩膜推理损失优化
       // 异构语义统一表征和转述
6:     从其余异构语义体系采样与  $\langle m_A, k_A \rangle$  语义一致或接近的消息-语义知识对  $\langle m_{-A}, k_{-A} \rangle$ 。不限语义体系语义, 采样得到与  $\langle m_A, k_A \rangle$  语义不同的负样本  $\langle m, k \rangle^-$ , 以及正样本  $\langle m, k \rangle^+$  (如条件允许)
7:     使用语义相似度  $\Theta$  对转译任务以损失函数(4-8)优化
8:     按照式(4-5)对异构语义中的本原语义表征进行优化
       // 异构语义安全性
9:     采样异构语义知识体系中语义知识密钥  $k_{-A}$  以及邻域扰动噪声 Dist, 对不同语义体系下的语义知识使用对比损失按照式(4-11)对自身邻域以及语义密钥空间进行优化
10:   end for
11: end while
       // 训练结束

```

并能为大多数读者所接受。

本文使用了来自 OpenSubtitles^[107]的中英文数据, 该数据集提供了 65 种语言的配对电影字幕, 是多语言翻译中常用的数据集。其中本文对所研究的中英文数据采用了 50 万条样本用于训练。为了构建统一的多语言字典, 本文使用了常用的字节对编码^[108] (Byte Pair Encoding, BPE) 方法进行分词, 对中英两种语言构建了总维度为 40015 的字典 (该字典同时包含中英两种单词)。

在模型选择上, 本章使用了与 BERT 模型^[38]类似的结构, 其中 Summary 与 Speak 函数分别由堆叠的 Transformer 模型^[98]加以表征。在编解码阶段, 用户的本地语义知识 k 被视为可学习的参数, 与词嵌入函数 WE 以及位置编码函数 PE 一同提供语义符号的向量嵌入 (Embedding):

$$\text{Embedding}(w^{(t)}) = \mathbf{k} + \text{PE}(t) + \text{WE}(w^{(t)}) \quad (4-13)$$

其中，词嵌入函数和位置编码函数同样使用可学习的参数进行建模。本文不对 **Summary** 及 **Speak** 中语义知识向量、位置编码函数以及词嵌入函数的共享情况做要求，也即，这两个函数可共享一套词嵌入方法，亦可使用各自独立的参数进行词嵌入建模。

在语义相似度量指标的选择上，由于本章节所述语义网络侧重于异构语义的提取和统一表征，对物理层传输要求不高，不失一般地本章将使用计算量最小的交叉熵作为语义相似度量。

4.4.2 模型效果

本章着重研究了语义，特别是异构语义、被提取并加以统一表征和特异化表达的过程；在后续部分中，本文将对所提出的语义网络训练方式(4-12)（特别是其中每一种组件的必要性）的实际效果做逐一分析，回答语义如何被加工和表征的问题。本小节整体按照消融实验的逻辑关系进行组织。

4.4.2.1 语义通信与语义理解

在文献中，大多数的点对点语义通信系统专注于自编码损失 \mathcal{L}_{AE} 约束下的通信问题，以 BLEU 等语义相似度准则作为评价指标；并在此基础上对语义传输准确性和效率问题开展了一定的研究。上述的系统在一定程度上回答了如何更准确、高效传输语义的问题，却鲜有对语义的加工过程展开挖掘。本文意识到仅仅以“准确率”为研究重点而忽略语义加工的可解释性可能会在语义通信研究中带来一些潜在问题，并期望本节及后

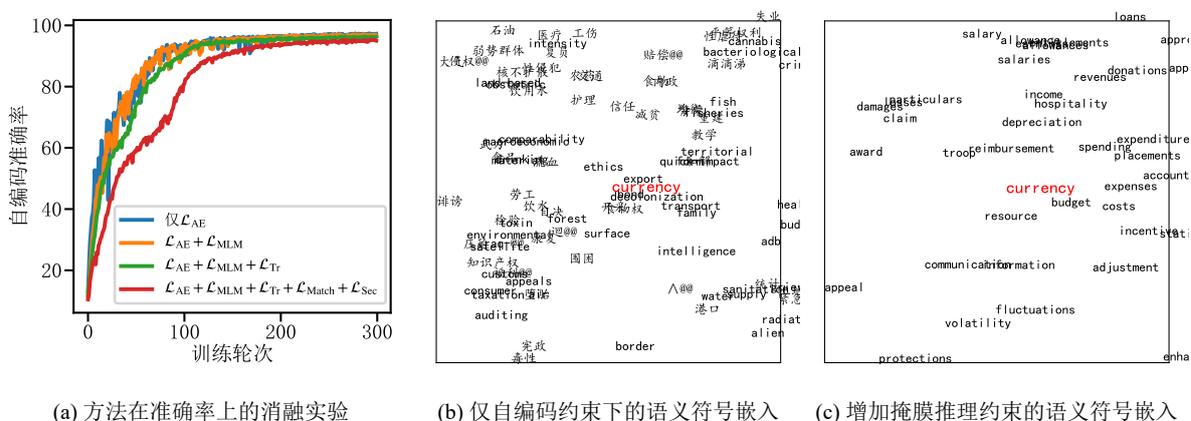


图 4.2 仅强调自编码准确性的系统在语义符号表征上的缺陷。从左至右，依次展示了所提本原语义网络各组件的消融实验、仅使用自编码约束下的语义符号空间，以及增加掩膜推理约束后的符号空间对比

续部分能够对未来语义通信的研究起到一定的帮助作用。

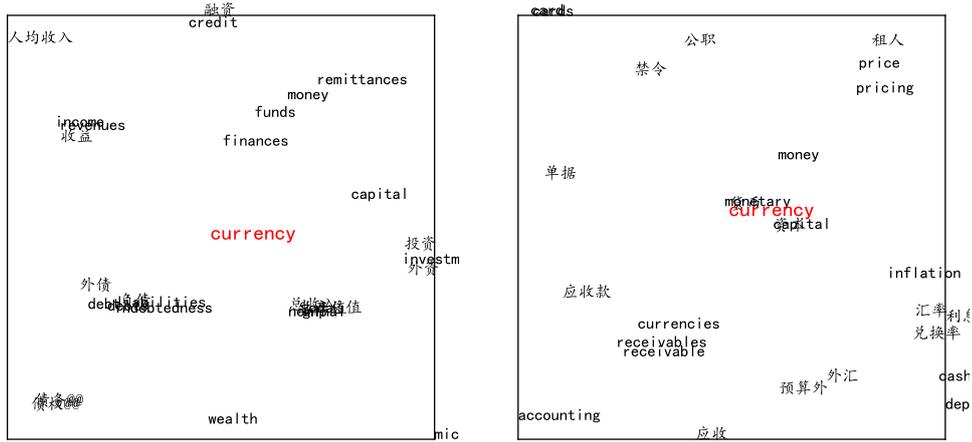
首先,本小节重新思考了仅以准确率为目标的学习范式在语义表征上的不足,并进一步介绍了语义表征在通用语义通信架构设计中的重要作用。作为对比实验,本文首先在暂时不引入异构语义知识的情况下与仅使用自编码器损失 \mathcal{L}_{AE} 的基线模型进行对比。特别地,图4.2的三幅子图展示了仅使用自编码损失约束与同时施加自编码损失和掩膜推理损失 \mathcal{L}_{MLM} 在自编码性能以及词汇语义空间嵌入上的差异。

从图4.2(a)中可以观察到,加入掩膜推理损失后,原自编码的准确性有着一定程度,但极为微小的下降。考虑到新增损失必将改变原自编码问题的优化轨迹,性能上的下降并不意外。事实上,自编码损失可视为一种利用某一位置之前的语义符号(在这里为词汇)对当前符号的预测任务,其目的在于单向的未来语义预测;而掩膜推理可视为使用其他所有位置的语义符号预测当前位置语义的双向推理任务。也因此,额外的掩膜推理损失在学习过程中和自编码任务有着一定的相似性,并未给自编码性能带来较大的退化。同时,在图4.2(c)中,加入双向语义推理任务后语义符号的表征空间更为紧凑,出现了更多同义词聚集的语义团簇。这一证据表明仅仅对自编码的优化虽然可以达到略好的准确性,语义编解码器在学习语义符号之间深层次关系时存在着较大的困难。相反地,使用自监督的掩膜推理任务帮助语义模型从上下文中提取出语义实体和关系间的一些复杂、双向的连接关系,更有利于统一语义表征空间的建立。

值得注意的是,虽然引入额外的损失函数(语义任务)有着给原本特定指标的评估带来性能下降的风险,本文注意到所提语义网络在其他语义特性(例如此处为可解释性和表征能力)上得到显著增强。另一方面需要指出的是,仅强调准确率在点对点通信中具有较大的价值,但并不绝对适用于异构语义网络,特别是异构网络中其他更为复杂的性能考虑(例如均衡性、公平性、安全性等)。在后续小节,本文将对所提语义网络的其他特性做进一步分析。

4.4.2.2 异构语义的统一表征

上一小节中,本文展示了一种使用语义推理任务优化语义表征的方法。该方法不同于以往仅仅以性能为考量的通信系统设计准则,为同构符号语义的联合表征以及可解释语义通信展示了一条基于表征学习的可行的途径。延续这一思路,本文更进一步对于所提方法在统一异构语义上的研究结果进行展示。



(a) 在自编码和掩膜约束上加入转述任务后的语义符号空间 (b) 使用语义对齐约束后的语义符号表征情况

图 4.3 所提转述任务和对齐任务对语义符号统一表征的影响

考虑异构的语义知识（例如中文和英文），为每一种异构语义设计独立的语义编码器和语义表征空间是极为昂贵且并不优雅的途径。本文从 **Speak** 和 **Summary** 函数出发，分别对异构的语义符号空间和异构的语义消息空间进行了统一表征：

- 异构语义符号的统一表征** 在多语义知识 $k(\cdot)$ 的情况下，本文在4.3.2节中将语义自编码任务拓展到异构语义下的转述任务。为研究转述任务对语义通信的影响，本文在 \mathcal{L}_{AE} 和 \mathcal{L}_{MLM} 基础上增加了异构语义的转述约束 \mathcal{L}_{Tr} ，并在中文编码-英文解码和英文编码-中文解码两个转述任务上同时进行了训练。如图4.3(a)所示，本文进一步观察到在多语言语义上的符号对齐现象。这一实验结果进一步论证了使用 $\text{Summary}(\langle m, k \rangle)$ 为异构语义建立统一符号表征的可行性，以及使用 $\text{Speak}(\langle \hat{s}, k \rangle)$ 进行语义转述对于提升语义联合表征可解释性的重要作用。事实上，从上一小节语义推理的角度来看，语义转述是一种同时利用双向异构语义，以及单向同构语义进行语义预测的任务，对于建立起跨语义知识的符号级对应关系起到了极大的帮助。此外，从整体的词汇嵌入表示中，本文也另外观察到常用连词处于词汇聚类中心，特定场景词汇分块聚类的现象。
- 异构本原语义的统一表征** 通过将自编码任务进一步拓展到异构语义的转述任务，语义网络能够实现高效的符号级语义对齐。然而，在消息语义的整体把握上，如图4.4(a)所示，由于异构语义体系在语法和表达上的差异，直接使用消息符号（例如对句中所有词汇嵌入取均值等常用的语义表征方式）作为本原语义并不能获

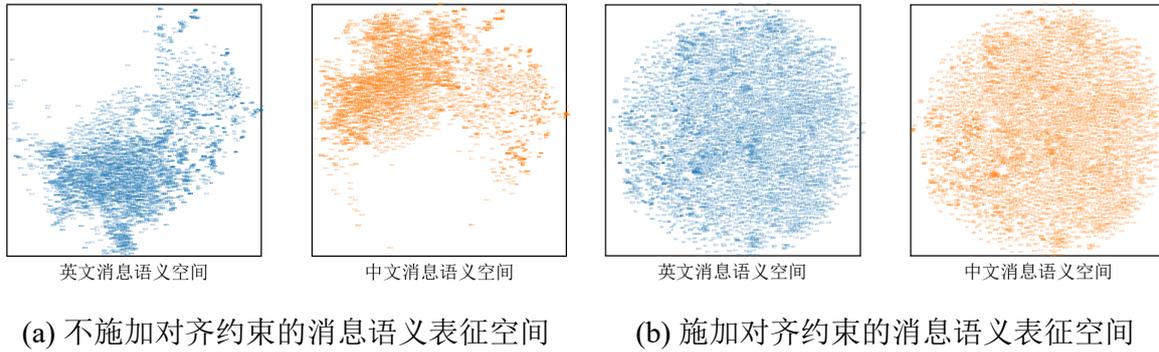


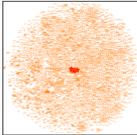
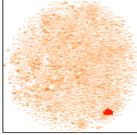
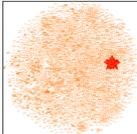
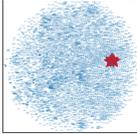
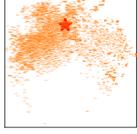
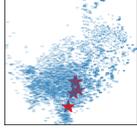
图 4.4 所提语义对齐约束对于消息本原语义表征的影响

得较好的效果。特别地，在本原语义表征空间的边缘，一些在不同语义知识中作用类似的简单表达能够呈现出对齐的模式；而更多的本原语义则是呈现出明显的语义隔阂（由语义的异构性质决定）。此外，如表4.1所示，在这种简单的本原语义表征下，投影空间中距离较近的样本并不一定严格拥有着接近的语义。

为处理这一问题，本文在4.3.1节提出了一种使用对比学习方法对本原语义进行对齐的方式。在具体的实现过程中，本文使用单层全连接神经网络将句向量投影到联合语义空间，并使用式(4-5)加以优化。该系统中的语义空间可视化如图4.4(b)所示。与未使用 \mathcal{L}_{Match} 的系统4.4(a)相比，语义对齐任务将语义接近的不同表达映射到统一的语义空间，实现对不同异构语义消息 m 中同一（或接近）语义本原 s 的高效提取。本文同时注意到，使用消息级语义对齐 \mathcal{L}_{Match} 对所提本原语义网络的提升主要来自于消息整体语义的表征，其对符号级语义对齐的帮助（图4.3(b)）相比掩膜推理任务（图4.2(c)）和转述任务（图4.3(a)）并不十分明显。

除了异构语义的对齐，本文特别地在所构成的统一语义表征空间中，观察到同构语义（如不同的中文表达或英文表达）话题团簇聚集的性质。在表4.1中，本文首先展示了使用语义对齐后对应于中文语义空间两处不同位置的符号消息簇。其中可以明显观察到在同构语义表征空间中，话题一致的表达倾向于拥有接近的语义表征，并使得整体语义空间呈现出宏观对称而微观特异的模式。这一消息级别语义聚类的特性（区别于使用 \mathcal{L}_{MLM} 和 \mathcal{L}_{Tr} 获得的符号级别语义聚类）为使用联合语义表征拓展高级语义任务（如分类、推荐、调度等）提供了可能。进一步地，本文以环境话题为例，将与中文样本共享相同本原语义的英文表达在对应的英文语义空间中进行了索引（见图4.1中红色五角星标注部分）。在这一实验中，语义对齐约

表 4.1 消息本原语义表征空间的可视化举例和消融性对比。进行了语义对齐的模型（前四个例子所示）在局部和整体上都拥有较好的语义分布和异构配对规律；而未使用语义对齐的模型（后两个例子）存在着异构语义表征不匹配的问题

语义空间嵌入	对应的符号表达
	(语义话题：资金) 截至 2007 年 12 月 31 日的 2006-2007 两年期收入和支出报表及储备金和资金余额变动财务报表是按照“儿童基金会财务条例和细则”以及联合国会计标准编列的 应收账款包括应向各全国委员会及其他合作伙伴收取的账款，包括... 在... 法庭任命一名财务调查员 该信托基金原先由联合国在 1977 年设立，名为...
	(语义话题：外交) 寻求加强同致力于谈判和外交并反对恐怖主义的伙伴们不断对话 ... 愿意竭尽全力重振裁军谈判会议 我们务必不要忘记，它也是唯一一个可真正做出决策和采取行动的环球论坛 我国对促进国际和平与安全的承诺及始终不渝的执着，是我国外交政策的... 我热切希望，他们能够按照计划于 8 月 20 日开始进行对话
	(语义话题：环境) 一些国家不存在有效的干旱和荒漠化评估制度，而另外一些国家却... 提供一套容易辨别荒漠化状况并且可加以衡量的指标，为加强... 评估易受气候变化影响程度的资金援助需求 其他可选办法涉及支助天然气管道铺设项目和支助利用目前被燃烧掉的... 工作队一直在处理城市环境日益脆弱，如何更好地支持...
	(上一段符号消息的异构转述) some countries have no effective systems for evaluating drought and desertification... disposing of a set of desertification state indicators that are easily... needs for financial assistance to assess vulnerability to climate change other options relate to supporting projects to construct gas pipelines... it has been addressing the increasing vulnerability of urban environments, how to...
	(语义话题：协商) 希望... 能够不经表决通过整个决议草案 希望与会者就议定成果达成共识，这有助于.. 决议草案周全地陈述相关的事实，他促请委员会协商一致通过决议草案 ... 以这样的理解赞同协商一致的意见 毫无疑问，... 将以协商一致方式通过该行动纲领
	(上一段符号消息的异构转述) hope..., it would be possible to adopt the draft resolution as a whole without a vote ...favoured a negotiated outcome that was fully owned and shared by all participants... ..., and he urged the committee to adopt it by consensus ...will join consensus with this understanding ...confident that the general assembly would endorse the programme of action by consensus

束成功地将异构语义消息所含有的同一本原语义，在联合语义表征空间进行了对齐。各个语义接近的同构及异构符号消息都能够在统一的语义空间得到接近的参数表征。对比而言，未进行语义空间对齐的模型（后两个例子）虽然在某种特定语义体系中也存在着类似的聚类性质，却无法对同一语义在异构知识上的多种表达形式实现统一；这一错位的语义表征也带来了，使用不同语义知识衡量下冲突的相似性度量。

值得一提的是，本文对语义符号和本原语义空间对齐展开的探索，大多以证明其可行性为目的，以期为统一的语义网络提供理论和实践方案。可以预见的是，将多对多语

义网络使用联合语义空间加以描述后，语义网络在语义表征上的统一性（相比于为不同用户设计独立编码器的方案^[4,72]）、可解释性、业务拓展性（例如中心节点可利用对本原语义统一表征的性质实现快速业务分类和统一调度）将极大地便利语义网络的实际运用和大规模部署。所提架构的理论优点不是本文企图展示的重点，但考虑到未来的一些可能拓展，本文对其做简要展望。

4.4.2.3 语义网络安全

在通常情况下，由于语义知识 \mathbf{k} 和本原语义 \mathbf{s} 均可兼容经典的加密算法，企图对语义网络进行窃听和破译是极为困难的。为了验证本文所提语义加密方案（将 \mathbf{k} 作为语义解码密钥）的语义攻击鲁棒性，本文对一种极端理想的破译情况展开实验：假设已知完整的 **Speak** 模型（视为可微部分而非黑箱）以及配对的本原语义 $\hat{\mathbf{s}}$ 和符号消息 $\hat{\mathbf{m}}$ ，攻击者需要逆向破译出该语义体系中的密钥 \mathbf{k} 。

从学习系统的角度看，这样一个语义攻击问题可以使用经典的监督学习方法、强化学习方法甚至生成对抗方法加以建模，其中强化学习和生成对抗模型对配对样本量有着较大的要求。不失一般地，本文使用训练数据中的 10 万条样本作为受窃听数据，并使用监督学习方法对原语义密钥进行回归： $\mathbf{k}^* = \operatorname{argmax}_{\mathbf{k}} \Theta(\hat{\mathbf{m}}, \operatorname{Speak}(\langle \hat{\mathbf{s}}, \mathbf{k} \rangle))$ 。

本文在图4.5中分别展示了仅去除语义安全约束 \mathcal{L}_{Sec} 的基线模型（记为 **baseline**）、在式(4-10)使用随机正态噪声建模负样本 $\operatorname{Dist}(\mathbf{k}) = \mathcal{N}(0, \sigma^2) \quad \sigma \in [0, 1]$ （记为 **randn**），

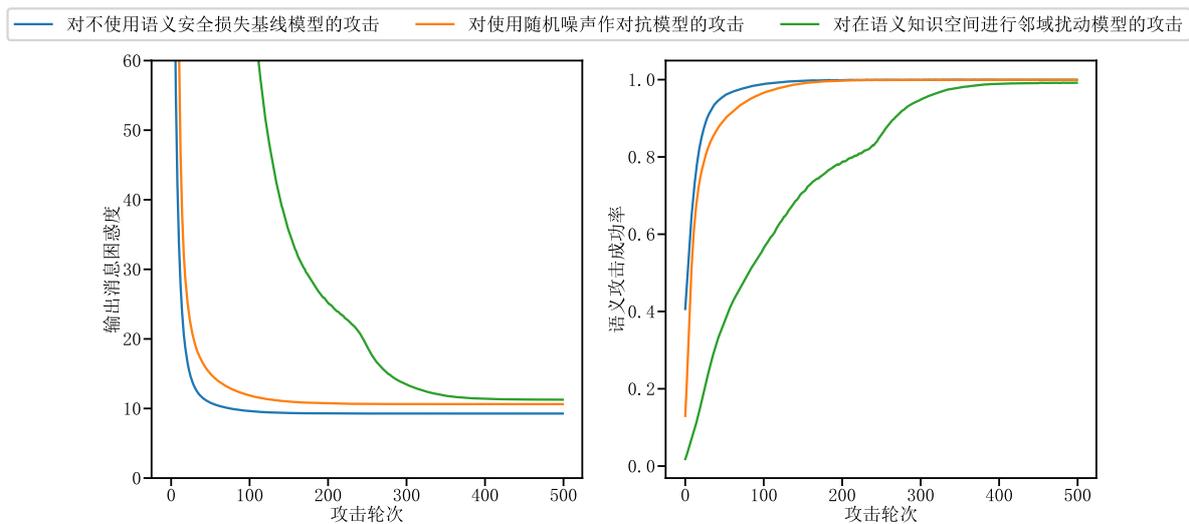


图 4.5 语义攻击鲁棒性对比。子图 (a) 和 (b) 分别展示了语义攻击者在困惑度（越低表示攻击越成功）以及语义攻击准确性（越高表示攻击越成功）上的得分情况

以及向原语义知识 \mathbf{k} 添加零均值随机高斯噪声作为扰动 $\text{Dist}(\mathbf{k}) = \mathbf{k} + 0.4\|\mathbf{k}\| \cdot \mathcal{N}(0, 1)$ (记为 **disturb**) 三种负样本构造方法在语义攻击鲁棒性上的实验结果。子图 (a) 和 (b) 分别展示了语义攻击者在困惑度 (**Perplexity**, 计算公式为 $1/\sqrt[N]{\prod_{t=1}^N p(\hat{w}^{(t)}|\hat{w}^{(0)}\dots\hat{w}^{(t-1)})}$, 越低攻击越成功) 和成功率 (计算方式为语义攻击的词准确率归一化到原模型准确率的比值, 越高攻击越成功) 上的得分情况。其中可以较为明显地看到, 在不对语义知识空间进行安全性约束的基线模型中, 语义攻击者可以迅速经由梯度下降法对原始解码密钥 \mathbf{k} 进行破译。在使用较弱的随机对抗样本 (**randn**) 对语义知识流形进行约束后, 窃听者对语义攻击的难度有所加大。特别地, 在第 0 时刻的迭代情况 (相当于使用随机初始化的语义密钥进行猜测), 基线模型相当一部分样本被破译, 而考虑到语义攻击的简单语义安全模型 (**randn**) 已可以较好地对语义网络中的语义解码进行保护。考虑到现实语义攻击可能并不能获取到足够的窃听样本, 语义网络对随机攻击的鲁棒性有着较大的实际意义, 也表现了本文语义安全考量的一大优势。更进一步地, 本文对使用更精细的语义知识约束 (**disturb** 变体) 进行了实验分析。其中可以明显观察到, 使用更高细粒度的对抗训练后, 窃听者语义攻击的难度大大增加, 需要耗费比基线模型 10 余倍的迭代次数方可达到较为理想的性能。对比使用普通随机噪声进行对抗的方式, **disturb** 变体将语义知识 \mathbf{k} 的邻域塑造得极为崎岖, 大大增加了窃听攻击破译真实密钥的难度。

在构建更鲁棒的语义系统之外, 本文也注意到在图4.5中, 几种考虑语义安全的解码模型 (**randn** 及 **disturb** 变体) 的一些局限性。从攻击的结果来说, 即便对语义知识流形进行约束的方法可以增大语义密钥破译的难度, 但 (在上述 **Speak** 模型已知的情况下) 语义攻击者最终依然可以获得较高的破译得分 (图4.5(b))。这是因为当网络中除语义密钥未知外所有参数已知时, 所优化的语义密钥本身维度较低, 解空间较为低维, 使用监督学习和梯度下降方法, 加上大量的训练样本即可逐渐破译。然而, 考虑到实际的语义网络并非会将敏感的编解码网络完全暴露 (例如仅开放 API), 攻击者只能通过黑箱模型对语义密钥进行猜测, 可以预想到所提知识流形约束给语义系统带来的实际价值。另外, 从系统训练的准确率来说, 额外的语义知识流形约束难免会给所提多用户语义系统的解码准确性带来一定程度上的下降 (如图4.5(a) 所示困惑度绝对数值), 语义安全性增益和解码准确性的权衡也应当是实际部署的考虑因素之一。

值得一提的是, 上述语义攻击情景属于在理想的攻击环境下对所提算法的可行性验证。在实际的系统部署角度, 由于所提语义网络对经典加密方案的兼容性, 不同的异构

语义用户可对其发送的本原语义和用于解码的密钥进行加密并实现密钥的动态更新；语义加密亦可和经典物理层、网络层加密互补地构建语义网络系统。作为语义网络的初步探索性工作，本论文对未来语义网络的安全性设想了一种可能的架构，并展示了使用语义密钥进行语义加密的可行性。

5 总结与展望

5.1 研究工作总结

伴随着人工智能和下一代通信技术的发展,语义通信研究的趋势方兴未艾,并已成为一种得到广泛关注的新研究方向。也正是在这一背景下,语义通信架构的研究经历了初步探索、自我批评到逐步发展的阶段,并依然处在迅速的更新迭代之中。本文对语义通信的研究起始于点对点语义通信的早期探索时期,针对当时语义通信体系在语义保真上的问题提出了一种在语义相似测度上的优化方法,为语义通信摆脱符号层级优化的方向提供了一定的理论和方法论指导,在语义通信发展历史中也获得了学术界的一些关注。

与许多在具体的语义提取方法和场景部署等方向的研究不同,本文所设想的语义通信系统以消息中最本原的语义为传输和加工对象,一切的设计原则均以语义的保真和表征为基础,也因此更多地关注于语义系统架构设计本身。站在现在的历史时刻回望过去点对点语义通信研究的发展历程,本文也意识到当前学术界在点对点语义研究泛化性和鲁棒性上的诸多遗憾,并进一步对广义的、统一的语义网络展开了架构设想和实验设计。在本文的第二个研究内容中,本文设想了一种统一的语义表征和表达模型,对异构语义知识、消息无关的本原语义表征,以及多用户语义通信下的安全性展开了研究。相比于现有的多对多语义通信方法,所提方案更好地统一了异构语义和知识表达方法,具有良好的效率、拓展性、可解释性和安全性。

具体地,本文所提出的方法和贡献如下:

1. 本文对于近几年来较为流行的语义通信从发展历史、使能技术、研究方法分类和技术细节等角度展开了分析,首先是对当前语义通信发展历程较为及时而完整的综述性质参考资料。特别地,本文立足于语义传输的基本要求,对现有的一些语义通信方法展开了讨论,指出潜在的缺陷并从点对点语义通信和广义语义网络两个角度设想了一种面向未来的语义通信体系。所提系统是对语义通信领域研究方法的补充,对于后续的研究工作起到了一定的引导作用,也为该领域未来的发展提供了一定的参考价值。

2. 特别地，对于点对点语义通信的问题，本文提出了一种以广义语义相似度为优化目标的统一的语义通信系统。该系统为语义通信从符号层优化到语义层优化提供了现实的理论和方法借鉴。相比于以复现为目标的自编码通信系统，所提方法更好地捕捉到语义消息的高阶语义相似度，为语义传输而非符号传输提供了可行性方面的证据。为了实现不可微语义相似度的优化，本文借鉴了强化学习中使用奖励函数建模决策问题的方式，将以往常用的回归和分类问题转化为使用策略梯度优化的决策问题，并在连续和离散空间上展示了方法的可行性。在技术方法上，所提方法首次将强化学习引入语义通信系统，并开辟了使用语义相似度优化语义通信系统的先例，对后续的一些研究起到了一定的引导作用。此外，所提方法首次将语义信息理论中的语义相似度与实验性语义通信研究中的语义通信系统通过强化学习加以桥接，是利用语义理论指导实际语义通信系统的良好示例，也帮助语义理论和实际语义系统这两个领域的研究联系地更为紧密。
3. 进一步地，本文意识到点对点语义通信在实际运用上的困难，并设想了一种多用户、异构语义知识体系下的本原语义网络模型。在这一架构中，本文提出使用统一的联合语义空间对不同的语义实体（符号消息）进行一致性映射的方法，首次实现了使用单编码器-单解码器传输通用语义本原的目标。从学习系统的角度，本文所提语义网络是对于点对点语义通信在优化目标、实现方法、理论体系上的扩充，使得使用统一的语义分析方法研究语义网络和点对点通信变得可能。本文所提方法也是表征学习在语义通信领域的首次实践，论证了语义表征在语义通信中的重要价值，是对以往简单语义自编码体系在可解释性、语义表征和推理能力、可拓展性等方面的扩充和重塑，在丰富语义通信研究方向和研究深度上起到了一定的作用。实验表明本文所提方法相较于简单自编码优化，在异构语义符号联合表征和语义本原统一表征上都有着优良的性质。此外，针对于异构语义网络中的安全性问题，本文提出了一种使用异构语义知识本身作为语义密钥的语义加密方案，并设计了一种兼容经典加密方法、支持动态密令更新的语义网络体系，在对抗语义攻击的鲁棒性方面带来了明显的提升。

5.2 研究工作展望

尽管语义通信上的研究已经取得了长足的进展，受到异构语义难以量化和表征的影响，语义信息理论的发展整体滞后于语义通信的试验性实践。就目前的研究情况来看，在如何更高效地使用语义信息理论指导语义通信系统设计方面，学术界的一些成果，如语义率失真和语义压缩^[22,28]、语义相似度和语义通信系统^[29]、语义信道容量^[5]依然处于起步和探索阶段。现有的一些实验性探索多以专家经验和仿真验证为主，并没有与语义信息理论构成良好的补充。本文第三章展示了一种使用语义相似度指导语义通信系统设计的思路，并在第四章展示了对异构语义进行统一表征的方法，克服了异构语义信息在理论度量的一些困难，为语义信息理论提供了一定的基础工具和实验方法。期待本文所做的一些工作可以在未来推动语义信息理论的发展，使得语义理论研究更多指导实验设计，并和语义通信实验研究相互促进。

使用语义相似度量而非比特级目标函数为通信系统朝语义传输发展提供了良好的经验指导。然而，在使用何种语义相似度量上，却很难找到良好而通用的准则，这也是通用语义信息表征的一大困难。在本文所提方案中，为构建通用的语义通信学习方法，强化学习的思想被首次引入语义通信。然而，强化学习本身也有着难以训练、收敛速度慢的缺陷，这些不足之处也是后续通用语义通信系统研究中值得深入挖掘的方向。特别地，在撰写本文之际，使用强化学习进行自监督和人工干预的方式取得了长足的进展，例如在 ChatGPT⁵中，研究人员使用人类反馈的强化学习（Reinforcement Learning from Human Feedback, RLHF）在文本的语义理解和生成上取得了史无前例的成功。该方法与本文所提思想有着较高的相似性，其中语义相似度量使用人工标注排序得到。另一方面，利用人工智能技术对智能体进行自监督训练，实现自我提升（例如 AlphaZero^[109]）的方法，也是通用语义信息度量和语义通信系统上可能的未来方向。

语义网络是语义通信中暂时受到较少关注的内容，却对于未来大规模的语义通信部署有着至关重要的作用。本文设想的语义网络克服了异构语义表征的困难，为统一而高效率的语义网络提供了一种可能的架构。但目前而言，绝大多数语义通信研究尚处于点对点、特殊场景下的设计，并不能很好拓展到通用语义网络情景。本文期待所提语义网络架构能够在未来语义网络发展上起到一定的推动和指引作用。

⁵该模型是 2022 年 12 月 OpenAI 公司在互联网发布的对话模型。

坦率地说，近些年语义通信的研究是崎岖的。目前的很多工作往往关注到符号层面的准确性，却忽视了语义保真和语义通信可解释性及通用性的问题。良好的语义通信生态也需要来自语义理论、语义实验、语义可信（可解释性、安全性、鲁棒性、公平性等）、语义运用等一系列研究和参与者的推动。本文期待未来语义通信的研究能够更多地对语义信息理论和语义表征问题展开进一步的探索，也期待未来语义通信可以成功地到通信业务中，为高质量语义传输提供新的解决方案。

整体上，语义通信的发展是前进的。能够在这样的背景下为学术和工业界添砖加瓦、带来新方法和新思路，笔者感到十分荣幸。期待所呈现的文章能够对语义通信的研究人员和感兴趣的同仁有所帮助，推动在该领域百家争鸣的壮观景象。

参考文献

- [1] Ye H, Liang L, Li G Y, et al. Deep learning-based end-to-end wireless communication systems with conditional GANs as unknown channels[J]. IEEE Transactions on Wireless Communications, 2020, 19(5): 3133-3143.
- [2] Xie H, Qin Z, Li G Y, et al. Deep learning enabled semantic communication systems[J]. IEEE Transactions on Signal Processing, 2021, 69: 2663-2675.
- [3] Uysal E, Kaya O, Ephremides A, et al. Semantic communications in networked systems: A data significance perspective[J]. IEEE Network, 2022, 36(4): 233-240.
- [4] Zhang G, Hu Q, Qin Z, et al. A unified multi-task semantic communication system for multimodal data[J]. arXiv preprint arXiv:2209.07689, 2022.
- [5] Bao J, Basu P, Dean M, et al. Towards a theory of semantic communication[C]//IEEE Network Science Workshop. Washington, USA, 2011: 110-117.
- [6] Strinati E C, Barbarossa S. 6G networks: Beyond Shannon towards semantic and goal-oriented communications[J]. Computer Networks, 2021, 190: 107930.
- [7] Lan Q, Wen D, Zhang Z, et al. What is semantic communication? A view on conveying meaning in the era of machine intelligence[J]. Journal of Communications and Information Networks, 2021, 6(4): 336-371.
- [8] Lu Z, Li R, Lu K, et al. Semantics-empowered communication: A tutorial-cum-survey[J]. arXiv preprint arXiv:2212.08487, 2022.
- [9] Shannon C E, Weaver W. The mathematical theory of information[M]. Urbana, IL: University of Illinois Press, 1949.
- [10] Shannon C E. A mathematical theory of communication[J]. The Bell System Technical Journal, 1948, 27(3): 379-423.
- [11] Qin Z, Tao X, Lu J, et al. Semantic communications: Principles and challenges[J]. arXiv preprint arXiv:2201.01389, 2021.
- [12] Shi G, Gao D, Song X, et al. A new communication paradigm: From bit accuracy to semantic fidelity [J]. arXiv preprint arXiv:2101.12649, 2021.
- [13] Lu K, Zhou Q, Li R, et al. Rethinking modern communication from semantic coding to semantic communication[J]. IEEE Wireless Communications, 2023, 30(1): 158-164.
- [14] Shi G, Xiao Y, Li Y, et al. From semantic communication to semantic-aware networking: Model, architecture, and open problems[J]. IEEE Communications Magazine, 2021, 59(8): 44-50.

- [15] Niu K, Dai J, Yao S, et al. A paradigm shift towards semantic communications[J]. *IEEE Communications Magazine*, 2022, 60(11): 113-119.
- [16] Zhang P, Xu W, Gao H, et al. Toward wisdom-evolutionary and primitive-concise 6G: A new paradigm of semantic communication networks[J]. *Engineering*, 2022, 8: 60-73.
- [17] Sana M, Strinati E C. Learning semantics: An opportunity for effective 6G communications[C]// *IEEE Annual Consumer Communications & Networking Conference (CCNC)*. Virtual, 2022: 631-636.
- [18] Iyer S, Khanai R, Torse D, et al. A survey on semantic communications for intelligent wireless networks[J]. *Wireless Personal Communications*, 2022: 1-43.
- [19] Yang W, Du H, Liew Z, et al. Semantic communications for 6G future Internet: Fundamentals, applications, and challenges[J]. *arXiv preprint arXiv:2207.00427*, 2022.
- [20] Wheeler D, Natarajan B. Engineering semantic communication: A survey[J]. *arXiv preprint arXiv:2208.06314*, 2022.
- [21] Seo H, Park J, Bennis M, et al. Semantics-native communication with contextual reasoning[J]. *arXiv preprint arXiv:2108.05681*, 2021.
- [22] Xiao Y, Zhang X, Li Y, et al. Rate-distortion theory for strategic semantic communication[C]// *IEEE Information Theory Workshop (ITW)*. Mumbai, India, 2022: 279-284.
- [23] Choi J, Loke S W, Park J. A unified view on Semantic information and communication: A probabilistic logic approach[J]. *arXiv preprint arXiv:2201.05936*, 2022.
- [24] Carnap R, Bar-Hillel Y, et al. An outline of a theory of semantic information[R]. *RLE Technical Report No. 247*. 1952.
- [25] Bar-Hillel Y, Carnap R. Semantic information[J]. *The British Journal for the Philosophy of Science*, 1953, 4(14): 147-157.
- [26] Floridi L. Outline of a theory of strongly semantic information[J]. *Minds and machines*, 2004, 14(2): 197-221.
- [27] D'Alfonso S. On quantifying semantic information[J]. *Information*, 2011, 2(1): 61-101.
- [28] Liu F, Tong W, Sun Z, et al. Task-oriented semantic communication systems based on extended rate-distortion theory[J]. *arXiv preprint arXiv:2201.10929*, 2022.
- [29] Lu K, Li R, Chen X, et al. Reinforcement learning-powered semantic communication via semantic similarity[J]. *arXiv preprint arXiv:2108.12121*, 2021.
- [30] Miller G A. WordNet: a lexical database for English[J]. *Communications of the ACM*, 1995, 38(11): 39-41.
- [31] Rada R, Mili H, Bicknell E, et al. Development and application of a metric on semantic nets[J]. *IEEE*

- Transactions on Systems, Man, and Cybernetics, 1989, 19(1): 17-30.
- [32] Jones K S. A statistical interpretation of term specificity and its application in retrieval[J]. Journal of Documentation, 1972, 28(1): 11-21.
- [33] Le Y, Wang Z J, Quan Z, et al. ACV-tree: A new method for sentence similarity modeling[C]// International Joint Conference on Artificial Intelligence (IJCAI). Stockholm, Sweden, 2018: 4137-4143.
- [34] Landauer T K, Foltz P W, Laham D. An introduction to latent semantic analysis[J]. Discourse Processes, 1998, 25(2-3): 259-284.
- [35] Blei D M, Ng A Y, Jordan M I. Latent dirichlet allocation[J]. Journal of machine Learning Research, 2003, 3(Jan): 993-1022.
- [36] Mikolov T, Chen K, Corrado G, et al. Efficient estimation of word representations in vector space [J]. arXiv preprint arXiv:1301.3781, 2013.
- [37] Pennington J, Socher R, Manning C D. Glove: Global vectors for word representation[C]// Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha, Qatar, 2014: 1532-1543.
- [38] Devlin J, Chang M W, Lee K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding[C]//Conference of the North American Chapter of the Association for Computational Linguistics (NAACL). Minneapolis, USA, 2019: 4171-4186.
- [39] Brown T B, Mann B, Ryder N, et al. Language models are few-shot learners[C]//Conference on Neural Information Processing Systems (NeurIPS). Virtual, 2020: 1877-1901.
- [40] Radford A, Kim J W, Hallacy C, et al. Learning transferable visual models from natural language supervision[C]//International Conference on Machine Learning (ICML). Virtual, 2021: 8748-8763.
- [41] Papineni K, Roukos S, Ward T, et al. BLEU: a method for automatic evaluation of machine translation [C]//Annual Meeting of the Association for Computational Linguistics (ACL). Philadelphia, USA, 2002: 311-318.
- [42] Vedantam R, Lawrence Zitnick C, Parikh D. CIDEr: Consensus-based image description evaluation [C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA, 2015: 4566-4575.
- [43] Zhu G, Iglesias C A. Computing semantic similarity of concepts in knowledge graphs[J]. IEEE Transactions on Knowledge and Data Engineering, 2017, 29(1): 72-85.
- [44] Tung T Y, Kobus S, Roig J P, et al. Effective communications: A joint learning and communication framework for multi-agent reinforcement learning over noisy channels[J]. IEEE Journal on Selected Areas in Communications, 2021, 39(8): 2590-2603.

- [45] Maatouk A, Kriouile S, Assaad M, et al. The age of incorrect information: A new performance metric for status updates[J]. *IEEE/ACM Transactions on Networking*, 2020, 28(5): 2215-2228.
- [46] Ceran E T, Gündüz D, György A. A reinforcement learning approach to age of information in multi-user networks with HARQ[J]. *IEEE Journal on Selected Areas in Communications*, 2021, 39(5): 1412-1426.
- [47] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [48] O'shea T, Hoydis J. An introduction to deep learning for the physical layer[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2017, 3(4): 563-575.
- [49] Farsad N, Rao M, Goldsmith A. Deep learning for joint source-channel coding of text[C]//*IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Alberta, Canada, 2018: 2326-2330.
- [50] Bourtsoulatze E, Kurka D B, Gündüz D. Deep joint source-channel coding for wireless image transmission[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2019, 5(3): 567-579.
- [51] Ye H, Li G Y, Juang B H. Power of deep learning for channel estimation and signal detection in OFDM systems[J]. *IEEE Wireless Communications Letters*, 2017, 7(1): 114-117.
- [52] Silver D, Huang A, Maddison C J, et al. Mastering the game of Go with deep neural networks and tree search[J]. *Nature*, 2016, 529(7587): 484-489.
- [53] Carpi F, Häger C, Martalò M, et al. Reinforcement learning for channel coding: Learned bit-flipping decoding[C]//*Annual Allerton Conference on Communication, Control, and Computing*. Monticello, USA, 2019: 922-929.
- [54] Li R, Zhao Z, Sun Q, et al. Deep reinforcement learning for resource management in network slicing [J]. *IEEE Access*, 2018, 6: 74429-74441.
- [55] Xu Z, Wang Y, Tang J, et al. A deep reinforcement learning based framework for power-efficient resource allocation in cloud RANs[C]//*IEEE International Conference on Communications (ICC)*. Paris, France, 2017: 1-6.
- [56] Gadaleta M, Chiariotti F, Rossi M, et al. D-DASH: A deep Q-learning framework for DASH video streaming[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2017, 3(4): 703-718.
- [57] Luong N C, Hoang D T, Gong S, et al. Applications of deep reinforcement learning in communications and networking: A survey[J]. *IEEE Communications Surveys & Tutorials*, 2019, 21(4): 3133-3174.
- [58] Yun W J, Lim B, Jung S, et al. Attention-based reinforcement learning for real-time UAV seman-

- tic communication[C]//International Symposium on Wireless Communication Systems (ISWCS). Virtual, 2021: 1-6.
- [59] Weaver W. Recent contributions to the mathematical theory of communication[J]. *ETC: A Review of General Semantics*, 1953, 10(4): 261-281.
- [60] Güler B, Yener A, Swami A. The semantic communication game[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2018, 4(4): 787-802.
- [61] Wang Y, Chen M, Saad W, et al. Performance optimization for semantic communications: An attention-based learning approach[C]//IEEE Global Communications Conference (GLOBECOM). Madrid, Spain, 2021: 1-6.
- [62] Liang J, Xiao Y, Li Y, et al. Life-long learning for reasoning-based semantic communication[J]. arXiv preprint arXiv:2202.01952, 2022.
- [63] Jiang P, Wen C K, Jin S, et al. Deep source-channel coding for sentence semantic transmission with HARQ[J]. *IEEE Transactions on Communications*, 2022.
- [64] Hu Q, Zhang G, Qin Z, et al. Robust semantic communications against semantic noise[J]. arXiv preprint arXiv:2202.03338, 2022.
- [65] Weng Z, Qin Z, Li G Y. Semantic communications for speech signals[C]//IEEE International Conference on Communications (ICC). Virtual, 2021: 1-6.
- [66] Zhou Q, Li R, Zhao Z, et al. Adaptive bit rate control in semantic communication with incremental knowledge-based HARQ[J/OL]. *IEEE Open Journal of the Communications Society*, 2022, 3: 1076-1089. <https://ieeexplore.ieee.org/abstract/document/9877924>.
- [67] Xie H, Qin Z. A lite distributed semantic communication system for Internet of Things[J]. *IEEE Journal on Selected Areas in Communications*, 2020, 39(1): 142-153.
- [68] Zhou Q, Li R, Zhao Z, et al. Semantic communication with adaptive universal transformer[J]. *IEEE Wireless Communications Letters*, 2021, 11(3): 453-457.
- [69] Maatouk A, Assaad M, Ephremides A. The age of incorrect information: An enabler of semantics-empowered communication (In Press)[J]. *IEEE Transactions on Wireless Communications*, 2022.
- [70] Li A, Wei X, Wu D, et al. Cross-modal semantic communications[J]. *IEEE Wireless Communications*, 2022, 29(6): 144-151.
- [71] Yan L, Qin Z, Zhang R, et al. Resource allocation for text semantic communications[J]. *IEEE Wireless Communications Letters*, 2022, 11(7): 1394-1398.
- [72] Xie H, Qin Z, Tao X, et al. Task-oriented multi-user semantic communications[J]. *IEEE Journal on Selected Areas in Communications*, 2022, 40(9): 2584-2597.
- [73] Sutton R S, Barto A G. Reinforcement learning: An introduction[M]. MIT Press, Cambridge, USA,

- 2018.
- [74] Watkins C J C H. Learning from delayed rewards[D]. King's College, Cambridge, United Kingdom, 1989.
- [75] Sutton R S. Generalization in reinforcement learning: Successful examples using sparse coarse coding[C]//Conference on Neural Information Processing Systems (NeurIPS). Denver, USA, 1995: 1038-1044.
- [76] Mnih V, Kavukcuoglu K, Silver D, et al. Playing Atari with deep reinforcement learning[J]. arXiv preprint arXiv:1312.5602, 2013.
- [77] Williams R J. Simple statistical gradient-following algorithms for connectionist reinforcement learning[J]. Machine Learning, 1992, 8(3): 229-256.
- [78] Konda V R, Tsitsiklis J N. Actor-critic algorithms[C]//Conference on Neural Information Processing Systems (NeurIPS). Denver, USA, 1999: 1008-1014.
- [79] Rennie S J, Marcheret E, Mroueh Y, et al. Self-critical sequence training for image captioning[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA, 2017: 7008-7024.
- [80] Luo R. A better variant of self-critical sequence training[J]. arXiv preprint arXiv:2003.09971, 2020.
- [81] Lillicrap T P, Hunt J J, Pritzel A, et al. Continuous control with deep reinforcement learning[C/OL]//International Conference on Learning Representations (ICLR). San Juan, Puerto Rico, 2016. <https://arxiv.org/abs/1509.02971>.
- [82] Schulman J, Levine S, Abbeel P, et al. Trust region policy optimization[C]//International Conference on Machine Learning (ICML). Lille, France, 2015: 1889-1897.
- [83] Schulman J, Wolski F, Dhariwal P, et al. Proximal policy optimization algorithms[J]. arXiv preprint arXiv:1707.06347, 2017.
- [84] Bengio Y, Courville A, Vincent P. Representation learning: A review and new perspectives[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(8): 1798-1828.
- [85] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [C/OL]//International Conference on Learning Representations (ICLR). San Diego, USA, 2015. <https://arxiv.org/abs/1409.1556>.
- [86] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA, 2016: 770-778.
- [87] Chen T, Kornblith S, Norouzi M, et al. A simple framework for contrastive learning of visual representations[C]//International Conference on Machine Learning (ICML). Virtual, 2020: 1597-1607.
- [88] Feng Y, Ma L, Liu W, et al. Unsupervised image captioning[C]//IEEE Conference on Computer

- Vision and Pattern Recognition (CVPR). Long Beach, USA, 2019: 4125-4134.
- [89] Cascante-Bonilla P, Tan F, Qi Y, et al. Curriculum labeling: Revisiting pseudo-labeling for semi-supervised learning[C]//AAAI Conference on Artificial Intelligence: vol. 35: 8. Vancouver, Canada, 2021: 6912-6920.
- [90] Nichol A, Dhariwal P, Ramesh A, et al. GLIDE: Towards photorealistic image generation and editing with text-guided diffusion models[C]//International Conference on Machine Learning (ICML). Baltimore, USA, 2022: 16784-16804.
- [91] Yang W, Liew Z Q, Lim W Y B, et al. Semantic communication meets edge intelligence[J]. arXiv preprint arXiv:2202.06471, 2022.
- [92] Li H, Tao C, Zhu X, et al. Auto Seg-loss: Searching metric surrogates for semantic segmentation [C/OL]//International Conference on Learning Representations (ICLR). Vienna, Austria, 2021. <https://arxiv.org/abs/2010.07930>.
- [93] Pitis S. Rethinking the discount factor in reinforcement learning: A decision theoretic approach[C] //AAAI Conference on Artificial Intelligence: vol. 33: 01. Honolulu, USA, 2019: 7949-7956.
- [94] Bahdanau D, Brakel P, Xu K, et al. An actor-critic algorithm for sequence prediction[C/OL]// International Conference on Learning Representations (ICLR). Toulon, France, 2017. <https://arxiv.org/abs/1607.07086>.
- [95] Ren Z, Wang X, Zhang N, et al. Deep reinforcement learning-based image captioning with embedding reward[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA, 2017: 290-298.
- [96] Ranzato M, Chopra S, Auli M, et al. Sequence level training with recurrent neural networks[C/OL] //International Conference on Learning Representations (ICLR). 2016. <https://arxiv.org/abs/1511.06732>.
- [97] Kurka D B, Gündüz D. Deep joint source-channel coding of images with feedback[C]//IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Barcelona, Spain, 2020: 5235-5239.
- [98] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//Conference on Neural Information Processing Systems (NeurIPS). Long Beach, USA, 2017: 5998-6008.
- [99] Aoudia F A, Hoydis J. End-to-end learning of communications systems without a channel model[C] //Asilomar Conference on Signals, Systems, and Computers (ACSSC). Pacific Grove, USA, 2018: 298-303.
- [100] Yun S, Choi J, Yoo Y, et al. Action-decision networks for visual tracking with deep reinforcement learning[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu,

- USA, 2017: 2711-2720.
- [101] Koehn P, et al. Europarl: A parallel corpus for statistical machine translation[C]//Proceedings of Machine Translation Summit X. Phuket, Thailand, 2005: 79-86.
- [102] Krizhevsky A. Learning multiple layers of features from tiny images[R]. Toronto, ON, Canada, 2009.
- [103] Furuta R, Inoue N, Yamasaki T. PixelRL: fully convolutional network with reinforcement learning for image processing[J]. IEEE Transactions on Multimedia, 2019, 22(7): 1704-1719.
- [104] Gutmann M, Hyvärinen A. Noise-contrastive estimation: A new estimation principle for unnormalized statistical models[C]//International Conference on Artificial Intelligence and Statistics. Sardinia, Italy, 2010: 297-304.
- [105] Oord A v d, Li Y, Vinyals O. Representation learning with contrastive predictive coding[J]. arXiv preprint arXiv:1807.03748, 2018.
- [106] Lu K, Li R, Zhang H. Contrastive monotonic pixel-level modulation[C]//European Conference on Computer Vision (ECCV). Tel Aviv, Israel, 2022: 493-510.
- [107] Tiedemann J. Parallel data, tools and interfaces in OPUS[C]//International Conference on Language Resources and Evaluation. Istanbul, Turkey, 2012: 2214-2218.
- [108] Sennrich R, Haddow B, Birch A. Neural machine translation of rare words with subword units[C]//Annual Meeting of the Association for Computational Linguistics. Berlin, Germany, 2016: 1715-1725.
- [109] Silver D, Hubert T, Schrittwieser J, et al. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play[J]. Science, 2018, 362(6419): 1140-1144.

附录

A 定理 1 的证明

以 $P(\hat{\mathbf{m}}|\phi; \theta)$ 表示采取某一完整轨迹（这里为时序符号） $\hat{\mathbf{m}}$ 的概率，式 (3-6) 中的期望可以被展开为：

$$\begin{aligned} J(\phi; \theta) &= \mathbb{E}_{\hat{w}^{(1)}, \dots, \hat{w}^{(N)}} \left[\sum_{t=1}^N r^{(t)} \right] \\ &= \sum_{\hat{\mathbf{m}}} P(\hat{\mathbf{m}}|\phi; \theta) \left(\sum_{t=1}^N r^{(t)} \right) \end{aligned} \quad (\text{A-1})$$

对等式两边求微分，可得

$$\begin{aligned} \nabla J(\phi; \theta) &= \nabla_{\phi, \theta} \left[\sum_{\hat{\mathbf{m}}} P(\hat{\mathbf{m}}|\phi; \theta) \left(\sum_{t=1}^N r^{(t)} \right) \right] \\ &= \sum_{\hat{\mathbf{m}}} \left[\nabla_{\phi, \theta} P(\hat{\mathbf{m}}|\phi; \theta) \left(\sum_{t=1}^N r^{(t)} \right) \right] \\ &= \sum_{\hat{\mathbf{m}}} \left[P(\hat{\mathbf{m}}|\phi; \theta) \nabla_{\phi, \theta} \log(P(\hat{\mathbf{m}}|\phi; \theta)) \left(\sum_{t=1}^N r^{(t)} \right) \right] \\ &= \mathbb{E}_{\hat{\mathbf{m}}} \left[\nabla_{\phi, \theta} \log(P(\hat{\mathbf{m}}|\phi; \theta)) \left(\sum_{t=1}^N r^{(t)} \right) \right] \end{aligned} \quad (\text{A-2})$$

其中第二个等式的成立是考虑到 1) 系统的总期望奖励并不依赖于参数 θ （见参考文献^[77]中的附录 A1）；以及 2) 交换求和以及微分操作在此处并不影响结果。第三个等式利用到了对数函数的微分公式 $\nabla \log(x) = \nabla x / x$ （其中 x 指代某一随机变量）。

在实际中，可以使用一次蒙特卡洛展开（Monte-Carlo Rollout）对 $\hat{\mathbf{m}}$ 的数学期望进行估计，也就是

$$\nabla J(\phi; \theta) \approx \nabla_{\phi, \theta} \log(P(\hat{\mathbf{m}}|\phi; \theta)) \left(\sum_{t=1}^N r^{(t)} \right) \quad (\text{A-3})$$

考虑到解码的完整采样由 N 次连续决策构成，解码轨迹 $\hat{\mathbf{m}}$ 的生成概率 $P(\hat{\mathbf{m}}; \theta)$ 可以使用条件概率展开为 $P(\hat{\mathbf{m}}; \theta) = P(\mathbf{s}^{(0)}) \prod_{t=1}^N [\pi_{\phi, \theta}(\hat{w}^{(t)}|\mathbf{s}^{(t)}) P(\mathbf{s}^{(t)}|\mathbf{s}^{(t-1)}, \mathbf{a}^{(t-1)})]$ 。于

是，对于式 (A-3)，有

$$\begin{aligned}\nabla J(\phi; \theta) &\approx \nabla_{\phi, \theta} \log \left(P(\mathbf{s}^{(0)}) \prod_{t=1}^N \pi_{\phi, \theta}(\hat{w}^{(t)} | \mathbf{s}^{(t)}) \right) \left(\sum_{t=1}^N r^{(t)} \right) \\ &= \sum_{t=1}^N \nabla_{\phi, \theta} \log \pi_{\phi, \theta}(\hat{w}^{(t)} | \mathbf{s}^{(t)}) \left(\sum_{t=1}^N r^{(t)} \right)\end{aligned}\quad (\text{A-4})$$

考虑到在句子传输过程中，中间过程的奖励通常为 0（如同围棋一样，只有句子完整生成才可获得奖励 $\Theta(\mathbf{m}, \hat{\mathbf{m}})$ ）。也有一些蒙特卡洛方法将中间过程的累积期望奖励视为与最后一步一致，并以平均奖励为优化目标，这种考虑的角度最后得到的结果是一致的），将语义通信问题的优化目标 $\Theta(\mathbf{m}, \hat{\mathbf{m}})$ 代入式 (A-4)，得到：

$$\nabla J(\phi; \theta) \approx \sum_{t=1}^N \nabla_{\phi, \theta} \log \pi_{\phi, \theta}(\hat{w}^{(t)} | \mathbf{s}^{(t)}) \Theta(\mathbf{m}, \hat{\mathbf{m}}) \quad (\text{A-5})$$

定理 1 证毕。 ■

B 定理 2 的证明

1) 式 (3-13a) 的证明

均值为 $\boldsymbol{\mu}^{d \times 1}$ 以及协方差矩阵为 $\boldsymbol{\Sigma}^{d \times d}$ 的高斯分布具有如下的概率密度函数：

$$\pi(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{d/2} \sqrt{\det \boldsymbol{\Sigma}}} \exp \left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right) \quad (\text{B-1})$$

上式的对数似然为：

$$\begin{aligned}\log \pi(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) &= \\ &-\frac{d}{2} \log 2\pi - \frac{1}{2} \log(\det \boldsymbol{\Sigma}) - \frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\end{aligned}\quad (\text{B-2})$$

对式 (B-2) 求梯度，得到

$$\nabla_{\boldsymbol{\mu}} \log \pi(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \boldsymbol{\Sigma}^{-1} [\mathbf{x} - \boldsymbol{\mu}] \quad (\text{B-3})$$

当 $\mathbf{x} = \widetilde{\mathcal{F}}_{\text{TX}}(\mathbf{m})$ ，以及 $\boldsymbol{\mu} = \mathcal{F}_{\text{TX}}(\mathbf{m})$ 时，代入即可得到式 (3-13b)。

式 (3-13a) 证毕。 ■

2) 式 (3-13b) 的证明

為表示的簡單起見，本文將解碼器在第 t 個時間步的輸出表示為 $\mathbf{z}^{(t)} \in \mathbb{R}^V$ ，其中 V 表示字典維度（候選詞匯的範圍）。Softmax 策略對 $\mathbf{z}^{(t)}$ 計算每個候選動作的概率為：

$$a_k^{(t)} = \frac{e^{z_k^{(t)}}}{\sum_{i=1}^V e^{z_i^{(t)}}} \quad (\text{B-4})$$

於是 $\mathbf{a}^{(t)} = [p(\hat{w}_1^{(t)}), p(\hat{w}_2^{(t)}) \dots p(\hat{w}_V^{(t)})]^T$ 。其中的每一元素（表示每個動作選擇的概率）被限制在 $[0, 1]$ 範圍內，與本文在式 (3-8) 中的表示一致。在不引起歧義的情況下，後續將省略上標 t 。

策略梯度方法對動作的概率計算其對數，本文記這一過程為 $L = \log(\pi_{\theta}(\hat{w}^{(t)} | \mathbf{s}^{(t)}))$ 。設 \mathbf{a} 中第 j 個元素在式 (3-12) 中被採樣到，根據鏈式法則， $\frac{\partial L}{\partial z_k} = \frac{\partial L}{\partial a_j} \cdot \frac{\partial a_j}{\partial z_k}$ 。對於等號右邊的第一項，有 $\frac{\partial L}{\partial a_j} = \frac{1}{a_j}$ （對數求導性質）；對於等號右邊第二項，有：

$$\begin{aligned} \frac{\partial a_j}{\partial z_k} &= e^{z_j} \cdot \frac{-e^{z_k}}{(\sum_{i=1}^V e^{z_i})^2} \quad \text{if } k \neq j \\ &= -a_j \cdot a_k \\ \frac{\partial a_j}{\partial z_k} &= \frac{e^{z_j} \cdot (\sum_{i=1}^V e^{z_i}) - e^{z_j} \cdot e^{z_j}}{(\sum_{i=1}^V e^{z_i})^2} \quad \text{if } k = j \\ &= a_j - (a_j)^2 \end{aligned} \quad (\text{B-5})$$

考慮式 (B-5) 中的兩種情況， L 對 z_k 的偏導可以寫為：

$$\frac{\partial L}{\partial z_k} = \begin{cases} -a_k & \text{if } a_k \text{ 未被採樣到} \\ 1 - a_k & \text{if } a_k \text{ 被採樣到} \end{cases} \quad (\text{B-6})$$

於是，對數概率對解碼器輸出向量 $\mathbf{z}^{(t)}$ 的求導可以表示為：

$$\frac{\partial L}{\partial \mathbf{z}^{(t)}} = \mathbb{1}^{V \times 1} - [p(\hat{w}_1^{(t)}), p(\hat{w}_2^{(t)}) \dots p(\hat{w}_V^{(t)})]^T \quad (\text{B-7})$$

式 (3-13b) 證畢。 ■

需要指出的是，在本文中編解碼器分別由 ϕ 及 θ 參數表征，最終目標函數對於神經網絡參諸元的導數可以在式 (3-13) 基礎上進一步由鏈式法則求得。然而，在以人工神經網絡為代表的非線性複雜映射關係中，求得目標函數對任意結構參數的解析導數是較為複雜的。就一個較為簡單的情形，例如考慮線性映射而言 ($\mathcal{F}_{\text{TX}}^{d \times 1}(\mathbf{m}) = \phi^{d \times T} \mathbf{m}^{T \times 1}$ 、

$\mathcal{F}_{\text{RX}}^{V \times 1}(\mathcal{H}(\mathcal{F}_{\text{TX}}(\mathbf{m}))) = \boldsymbol{\theta}^{V \times d} [\mathcal{H}(\mathcal{F}_{\text{TX}}(\mathbf{m}))]^{d \times 1}$, 目标函数对 $\boldsymbol{\phi}$ 及 $\boldsymbol{\theta}$ 的导数具有解析形式:

$$\nabla_{\boldsymbol{\phi}} \log(\pi_{\boldsymbol{\phi}}) = \boldsymbol{\Sigma}^{-1} \left[\widetilde{\mathcal{F}_{\text{TX}}}(\mathbf{m}) - \mathcal{F}_{\text{TX}}(\mathbf{m}) \right] \mathbf{m}^T \quad (\text{B-8a})$$

$$\begin{aligned} \nabla_{\boldsymbol{\theta}} \log(\pi_{\boldsymbol{\theta}}(\hat{w}^{(t)} | s^{(t)})) = \\ \sum_{k=1}^V \left[\mathbb{1}^{V \times 1} - \left[p(\hat{w}_1^{(t)}), p(\hat{w}_2^{(t)}) \dots p(\hat{w}_V^{(t)}) \right]^T \right]_k \cdot \nabla_{\boldsymbol{\theta}} \left[\mathcal{F}_{\text{RX}}^{(t)} \right]_k \end{aligned} \quad (\text{B-8b})$$

其中, 式 (B-8b) 中下标 k 表示矩阵或向量中的第 k 元。

在更一般的情况下, 考虑对神经网络参数的导数, 需要根据具体的结构进行讨论。

硕士期间的成果

获得奖励和荣誉：

2021 年 10 月，浙江大学优秀研究生

2022 年 7 月，ECCV 2022 Oral

2022 年 10 月，研究生国家奖学金

硕士阶段发表的学术论文：

1. Rethinking modern communication from semantic coding to semantic communication[J]. IEEE Wireless Communications, 2023, 30(1): 158-164. 第一作者，对应本文第二章
2. Contrastive monotonic pixel-level modulation[C]//European Conference on Computer Vision (ECCV). Tel Aviv, Israel, 2022: 493-510. 第一作者，对应本文第四章

硕士阶段的其他论文：

1. Reinforcement learning-powered semantic communication via semantic similarity[J]. arXiv preprint arXiv:2108.12121, 2021. 第一作者，对应本文第三章
2. Semantics-empowered communication: A tutorial-cum-survey[J]. arXiv preprint arXiv: 2212.08487, 2022. 第三作者，对应本文第二章

参研项目：

1. 浙江省杰出青年基金项目，LR23F010005，面向移动分布式智能业务的任务中心网络，2023/1-2025/12。
2. 国家自然科学基金面上项目，62071425，基于业务感知的内生智能通信网络研究，2021/1-2024/12。
3. 国家自然科学基金重点项目，61731002，基于用户行为预测的无线边缘网络资源利用理论研究，2018/1-2022/12。

4. 浙江省“领雁”项目，2022C01093，面向多网融合低功耗传感设备的数能同传一体化集成技术研发，2022/1-2024/12。
5. 华为技术有限公司，6G无线智能网络架构中智能内生理论研究，2020/7-2023/7。