

框架语义自动解析研究



常宝宝

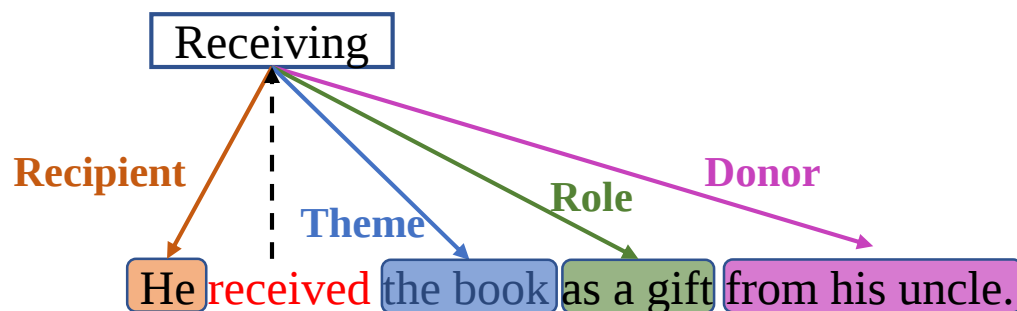
北京大学计算语言学研究

主要内容

- 框架语义解析简介
- 主要研究内容
 - 框架语义解析的建模研究
 - 知识驱动的框架语义解析
 - 大模型和小模型差异分析及级联
- 总结

框架语义解析简介

- 建立在框架语义学基础上的句子意义解析任务。

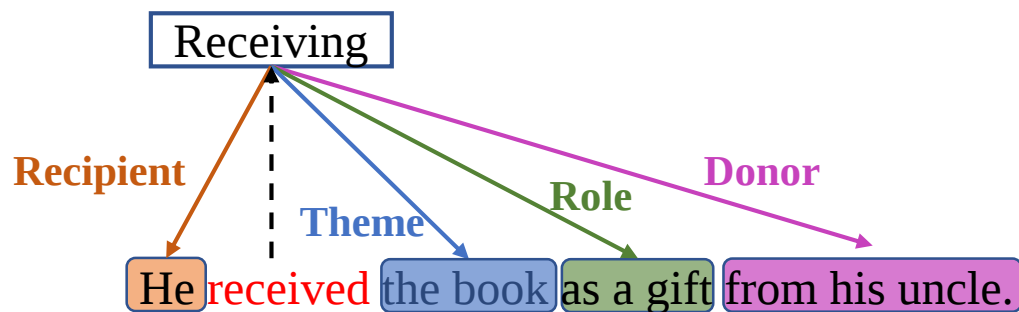


- 狭义上，框架作为认知结构，是人们理解句子意义的认知基础
- 广义上，框架作为表示机制，也是知识存储和表示的手段

框架语义解析简介

- 框架语义解析是一个复杂的任务，有许多挑战！
 - 目标词识别 (target Identification) → received
 - 框架识别 (frame Identification) → Receiving
 - 论元识别 (Argument Identification) → He/the book/as a gift/from his uncle
 - 角色分类 (Role Classification) → Recipient/Theme/Role/Donor

FrameNet中的语义角色即框架元素



框架语义解析简介

- （基于FrameNet的）自动框架语义解析工作开始于大约2000年
- 2000年，统计方法正在复苏过程中
框架语义解析工作几乎都基于统计方法（或者机器学习方法）
- 许多方法都被应用于框架语义解析任务，性能也持续得到改进
 - 最大熵方法、条件随机场、深度学习方法等
 - 序列标注法、基于片段的方法、基于转移的方法等

D. Gildea and Daniel Jurafsky. (2000). "Automatic labeling of semantic roles." ACL-2000.

Collin F. Baker, Charles J. Fillmore, and John B. Lowe. (1998). "The Berkeley FrameNet Project. ACL-1998.

主要内容

- 框架语义解析简介
- 主要研究内容
 - 框架语义解析的建模研究
 - 知识驱动的框架语义解析
 - 大模型和小模型差异分析及级联
- 总结

框架语义分析建模研究

- 主要关注点
 - 主要考虑多子任务联合分析，减少错误积累
 - 考虑框架元素间的选择限制关系
 - 一个框架元素的出现会影响另一个框架元素的概率分布
 - 表示和利用框架语义本体知识
- 近年来，自然语言处理领域取得很大进步，建模范式不断变化
 - 架构工程范式
 - 微调预训练语言模型范式 (面向特定任务的知识迁移)
 - 大模型提示范式 (面向所有任务的知识迁移)
- 基于不同建模范式的框架语义分析

基于架构工程的框架语义解析—MultiDecoder Joint Model

Author Text Form
I can write my name on the deposit slip .
Text_creation

- 基本思路：将框架语义分析视作生成任务

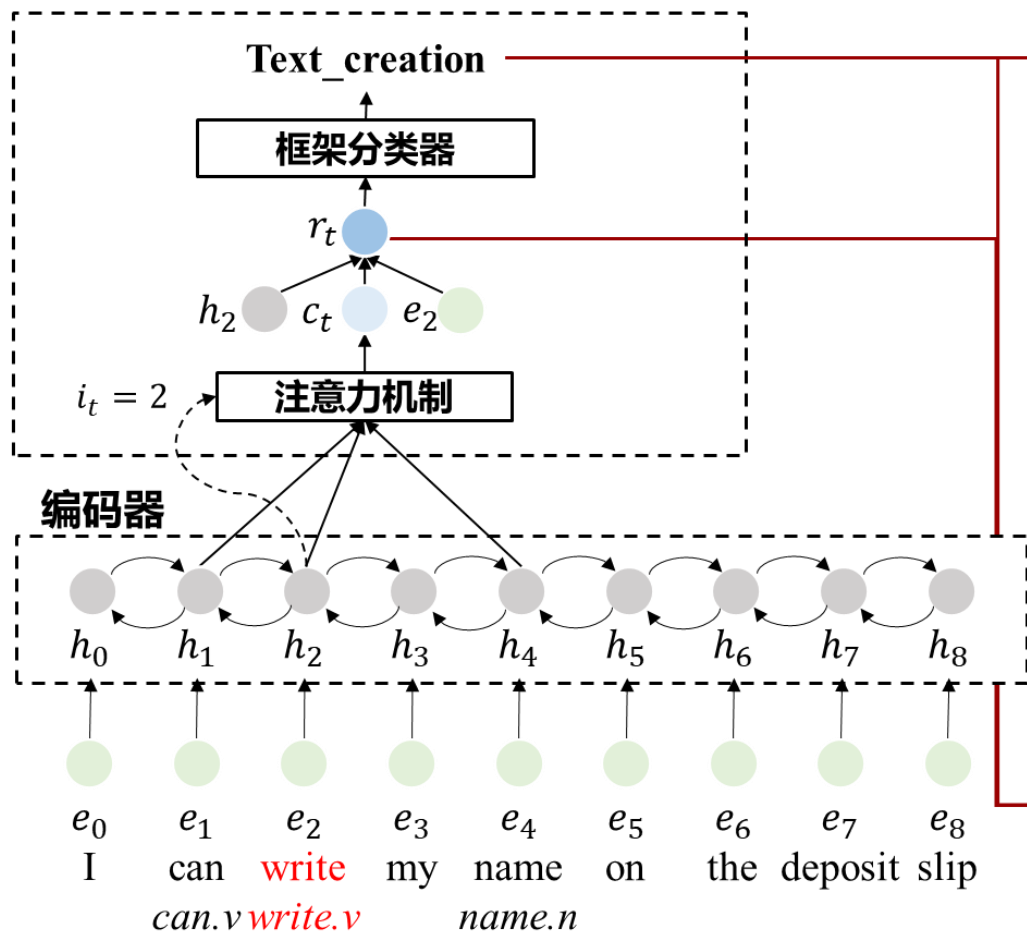
write → Text_Creation

→ I → Author → my name → Text → on this deposit slip → Form

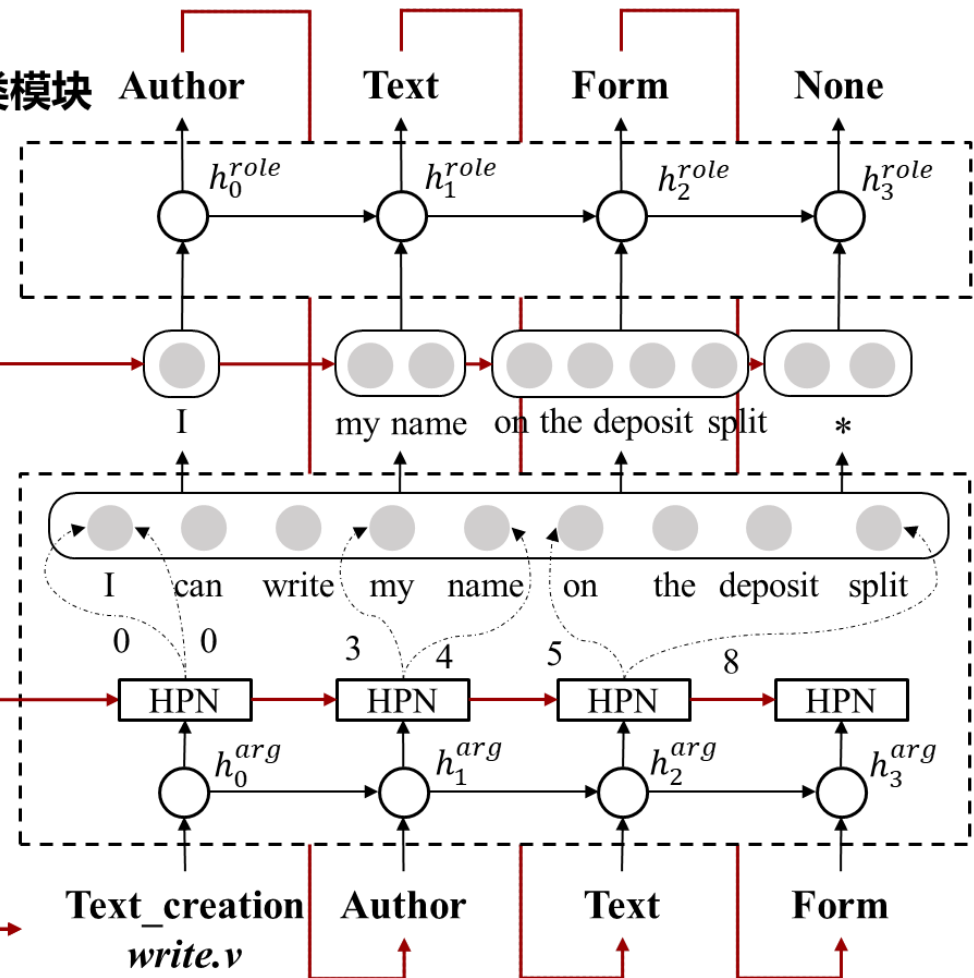
- 建模动机
 - 框架识别、论元识别、角色分类联合建模、联合训练和解码，缓解错误积累
 - 有完整的解析历史建模，后面的决策基于所有历史决策，建模框架元素之间的制约关系

Xudong Chen, Ce Zheng, Baobao Chang. Joint Multi-Decoder Framework with Hierarchical Pointer Network for Frame Semantic Parsing. ACL 2021 findings.

框架识别模块



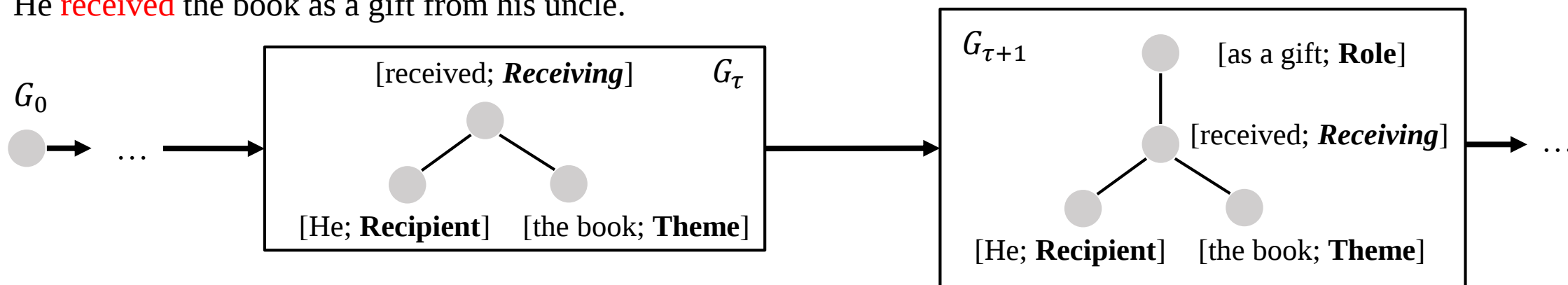
角色分类模块



角色识别模块

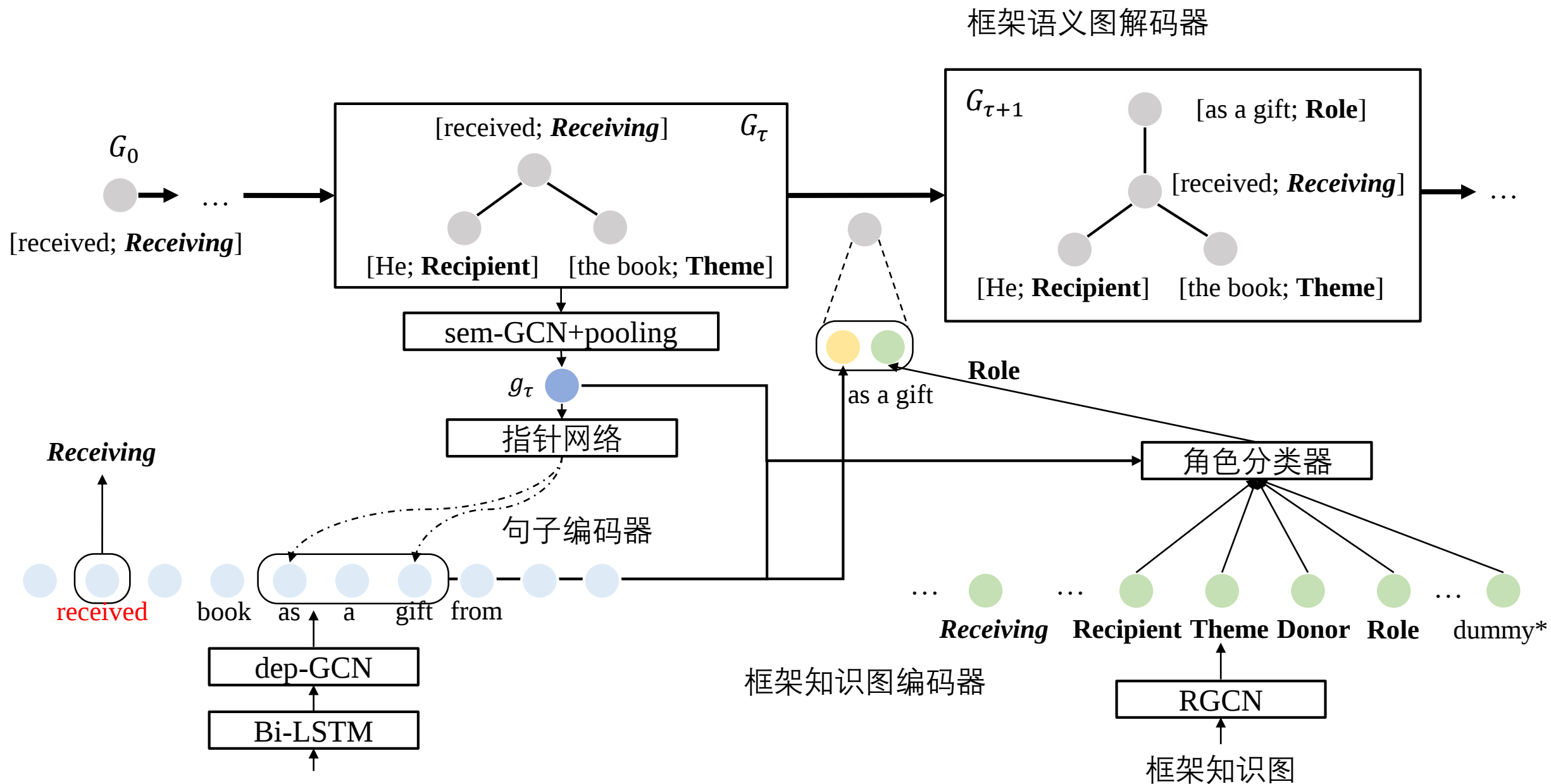
基于架构工程的框架语义解析 --- KID 模型

He **received** the book as a gift from his uncle.



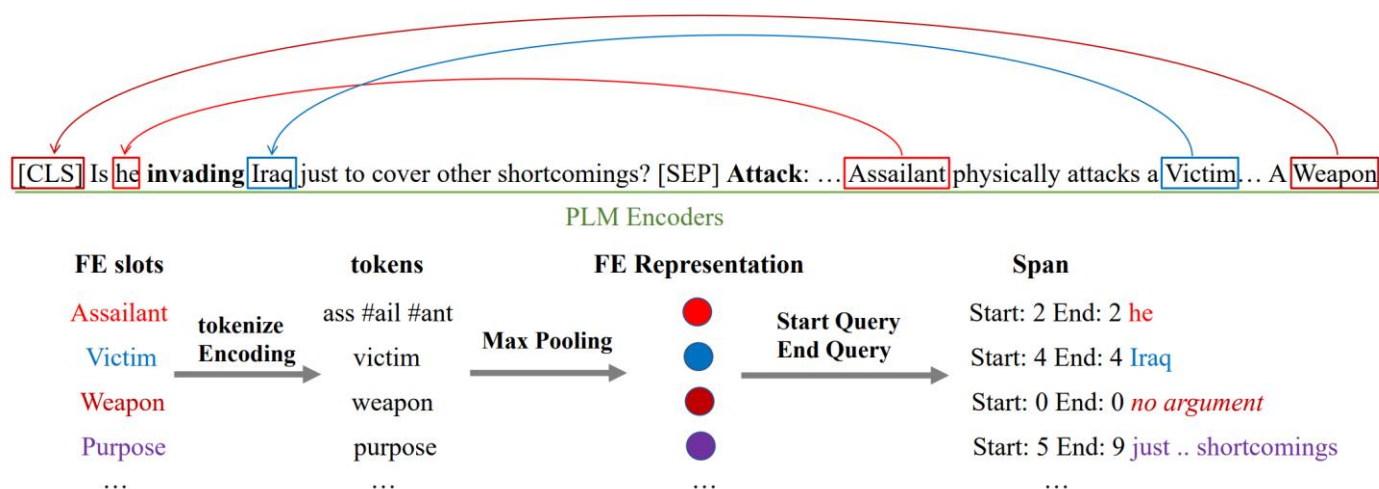
- 基本思路：框架语义解析过程视作增量式构图过程(框架语义图FSG)
- 建模动机：
 - 框架识别、论元识别、角色分类联合建模
 - 用框架语义图(FSG)建模框架语义解析历史
 - 用框架语义图建模句子框架语义的结构信息

Ce Zheng, Xudong Chen, Runxin Xu, Baobao Chang, A Double-Graph Based Framework for Frame Semantic Parsing. NAACL 2022.



基于微调范式的框架语义角色标注---AGED

- 基本思路：
 - 将待解析的句子与框架定义拼接，交由预训练语言模型编码表示
 - 利用框架定义中的框架元素提及(mention)的表示作为查询(query)识别句子中的框架元素
 - 采用微调预训练语言模型范式



Ce Zheng, Yiming Wang, Baobao Chang, Query Your Model with Definitions in FrameNet: an Effective Method for Frame Semantic Role Labeling, AAAI 2023.

基于微调范式的框架识别--- COFFTEA

- 基本思路

- 基于对比学习，引入两种负例产生方式

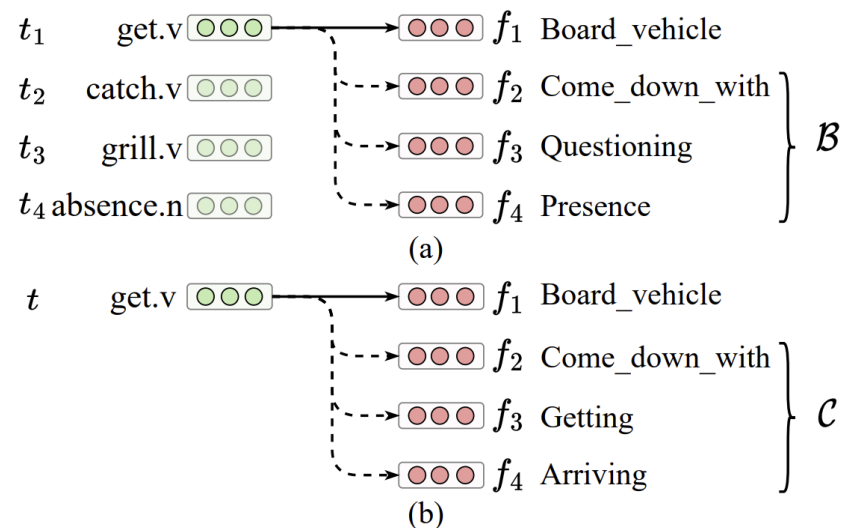
- (1) In-Batch Learning (IBL)

- (2) In-Candidate Learning (ICL)

- 两阶段策略，先易后难，
递进增加区分难度

- ICL 负例

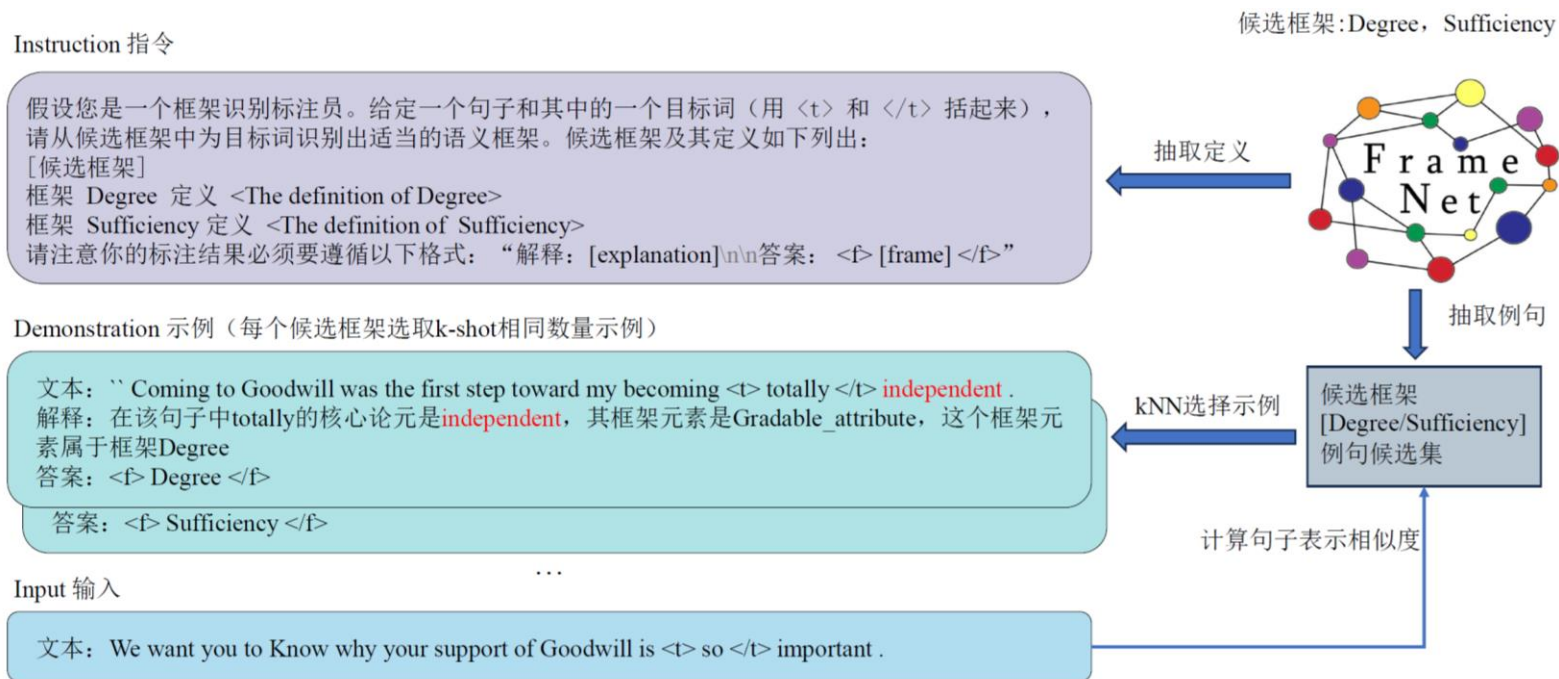
- 目标词可能对应的其他框架实例，与
目标框架有继承或其他关系的框架实例



Kaikai An, Ce Zheng, Bofei Gao, Haozhe Zhao, Baobao Chang, Coarse-to-Fine Dual Encoders are Better Frame Identification Learners, EMNLP 2023 Findings

基于提示范式的框架识别

- 基于ChatGPT
- 参数冻结
- 撰写指令
- 提供示例



- MultiDecoder joint Model结果

Model	All	Ambiguous
SEMAFOR	83.6	69.2
Open-SESAME	87.0	-
Hartmann et al.	87.6	73.8
Yang and Mitchell	88.2	75.7
Hermann et al.	88.4	73.1
Peng et al.(BASIC)	89.2	76.3
Our Model	89.4	76.7
Our Model+Bert	90.5	79.1

Table 3: Frame identification accuracy result.

Model	P	R	F1
SEMAFOR	69.2	65.1	67.1
Framat	71.1	63.7	67.2
Framat+context	71.1	64.8	67.8
Open-SESAME	71.0	67.8	69.4
FitzGerald et al.	74.8	65.5	69.9
Yang and Mitchell (SEQ)	69.6	70.9	70.2
Yang and Mitchell(REL)	77.1	68.7	72.7
Peng et al.(BASIC)	79.2	71.7	75.3
Our Model	75.1	76.9	76.0
Our Model+Bert	78.2	82.4	80.2

Table 4: Full structure extraction result on the FN test set.

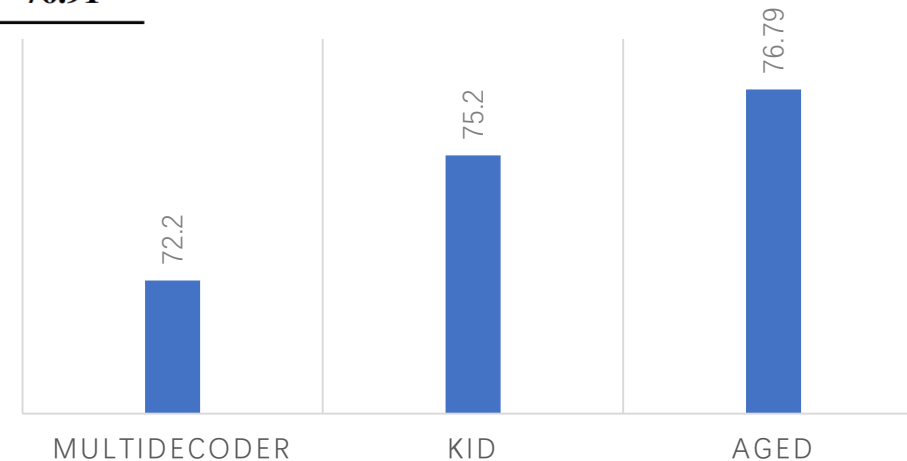
- KID模型结果

Model	Frame Id	Arg Id (gold frame)			Full structure		
	Accuracy	Precision	Recall	F1-score	Precision	Recall	F1-score
SEMAFOR (2014)	83.6	65.6	53.8	59.1	-	-	-
Hermann et al. (2014)	88.4	-	-	-	74.3	66.0	69.9
SEMAFOR (HI) (2015)	-	67.2	54.8	60.4	-	-	-
Täckström et al. (2015)	-	-	-	-	75.4	65.8	70.3
FitzGerald et al. (2015)	-	-	-	-	74.8	65.5	69.9
open-SESAME (2017)	86.9	69.4	60.5	64.6	71.0	67.8	69.4
KID (GloVe)	89.5	64.6	68.2	66.4	73.8	76.8	75.3
SEMAFOR (HI + exemplar) (2015)	-	66.0	60.4	63.1	-	-	-
Yang and Mitchell (2017)	88.2	70.2	60.2	65.5	77.3	71.2	74.1
Swayamdipta et al. (2018)	-	67.8	66.2	67.0	-	-	-
Peng et al. (2018)	89.2	-	-	-	79.2	71.7	75.3
Marcheggiani and Titov (2020)	-	69.8	68.8	69.3	-	-	-
Chen et al. (2021)	89.4	-	-	-	75.1	76.9	76.0
KID (GloVe + exemplar)	90.0	66.8	73.7	70.1	75.5	80.1	77.7
Bastianelli et al. (2020) (JL)	90.1	74.6	74.4	74.5	-	-	-
Chen et al. (2021) (BERT)	90.5	-	-	-	78.2	82.4	80.2
Jiang and Riloff (2021)	91.3	-	-	-	-	-	-
Su et al. (2021)	92.1	-	-	-	-	-	-
KID (BERT)	91.7	71.7	79.0	75.2	79.3	84.2	81.7

- AGED模型结果

Model	FN 1.5			FN 1.7		
	Precision	Recall	F1-score	Precision	Recall	F1-score
semi-CRF (2017)	-	-	73.56	-	-	72.22
Lin, Sun, and Zhang (2021)	-	-	73.28	-	-	72.06
Kalyanpur et al. (2020)	-	-	-	71	73	72
AGED (ours) w/o exemplar	71.93	76.78	74.28	74.02	75.46	74.73
AGED (ours) + FE definition w/o exemplar	73.43	76.31	74.84	75.39	75.36	75.37
Chen, Zheng, and Chang (2021)	69.27	75.39	72.20	-	-	-
Bastianelli, Vanzo, and Lemon (2020)	74.23	76.94	75.56	-	-	-
KID (2022)	71.7	79.0	75.2	74.1	77.3	75.6
AGED (ours)	73.06	79.84	76.30	75.84	77.87	76.84
AGED (ours) + FE definition	74.04	79.75	76.79	75.80	78.05	76.91

F1 SCORE ON FN1.5



• COFFTEA模型结果

Models	FN 1.5					FN 1.7				
	Acc	R@1	R@3	R@5	Overall	Acc	R@1	R@3	R@5	Overall
Hartmann et al. (2017)†	87.63	77.49	-	-	82.25	83.00	76.10	-	-	79.40
KGFI (2021)†	92.13	85.63	-	-	88.76	92.40	85.81	91.59	92.88	88.98
FIDO (2021)†	91.30	-	-	-	-	92.10	-	-	-	-
Tamburini (2022)	92.57	-	-	-	-	92.33	-	-	-	-
Lookup Table	88.60	78.48	86.74	88.16	83.23	89.02	78.75	87.43	88.71	83.57
+ <i>init.</i> with Def	90.40	83.56	88.25	89.32	86.85	90.30	85.17	88.93	89.87	87.66
+ $\mathcal{L}_B$	91.06	86.52	91.73	92.68	88.73	90.94	87.13	92.30	93.21	89.00
$\mathcal{L}_B$ only†	91.82	84.73	94.47	96.18	88.13	91.58	84.88	93.88	95.90	88.10
$\mathcal{L}_C$ only†	92.50	76.51	87.35	91.01	83.75	92.70	78.33	88.93	91.64	84.91
$\mathcal{L}_B$ only	90.94	86.77	95.14	96.79	88.81	91.36	87.38	95.29	96.77	89.33
COFFTEA	92.55	87.69	92.64	94.71	90.05	92.64	87.34	92.91	94.29	89.91

• ChatGPT结果

模型	准确率	困难样本准确率
FIDO+AGED	92.07	54.01
FIDO+AGED (4-shot)	81.40	-
ChatGPT (0-shot)	83.06	48.66
ChatGPT (4-shot)	85.21	59.36
CAP-Fi (不含思维链)	92.43	62.03
CAP-Fi	92.70	67.91

微调模型效果显著优于ChatGPT效果

主要内容

- 框架语义解析简介
- 主要研究内容
 - 框架语义解析的建模研究
 - **知识驱动的框架语义解析**
 - 大模型和小模型差异分析及级联
- 总结

知识驱动的框架语义解析

- FrameNet定义了框架语义解析任务
- FrameNet也促进了框架语义解析技术的进步
- 作为一个语义知识资源，FrameNet中信息未能得到充分利用
 - 多数研究只用到了FrameNet中标注例句
- 作为框架本体知识库，FrameNet中的语义信息
 - 结构化信息
 - 框架-框架关系、框架-元素关系、跨框架元素关系
 - 描述性信息
 - 框架定义、框架元素定义
- 如何利用框架语义本体知识，实现知识驱动的框架语义解析？

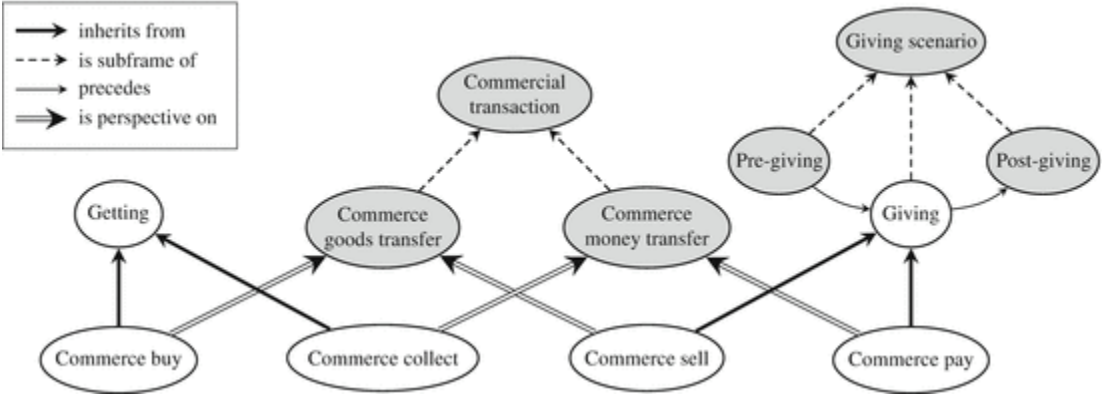
Ce Zheng, Xudong Chen, Runxin Xu, Baobao Chang, A Double-Graph Based Framework for Frame Semantic Parsing. NAACL 2022.

知识驱动的框架语义解析

- FrameNet作为一个词典
- FrameNet作为本体知识库
- FrameNet作为标注例句集

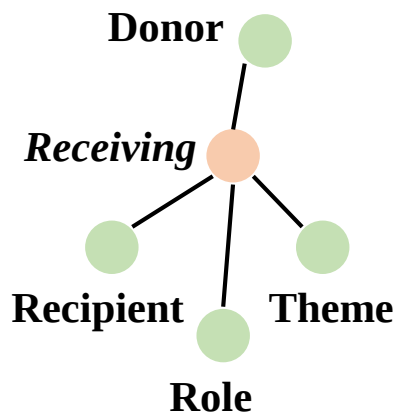
表 2.1 FrameNet 框架语义关系及其示例

语义关系	示例	
	头框架	尾框架
Inheritance (继承)	Receiving	Getting
Perspective_on (视角)	Commerce_buy, Commerce_sell	Commerce_transfer-goods
Subframe (成分)	Arrest, Trial	Criminal_process
Precedence (后继)	Trial	Arrest
Causative (致使)	Cause_change_position_on_a_scale	Change_position_on_a_scale
Inchoative (表始)	Change_position_on_a_scale	Position_on_a_scale
Using (使用)	Judgement_communication	Judgement, Statement
See also (参照)	Apply_heat	Cook_creation



知识驱动的框架语义解析

- 结点: 框架 (frame) 和 框架元素 (FE)
- 关系类型:
 - Frame-FE relations: 表示框架识别和其他子任务的交互

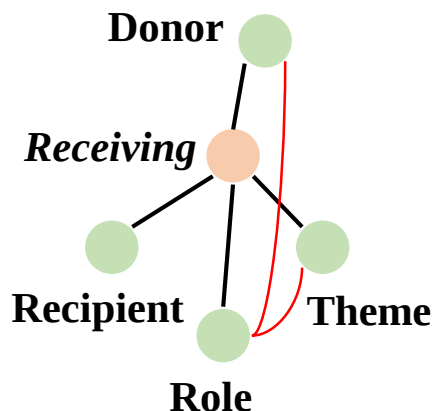


知识驱动的框架语义解析

- 结点: 框架 (frame) 和 框架元素 (FE)
- 关系类型:
 - Frame-FE relations: 表示框架识别和其他子任务的交互
 - Intra-Frame FE-FE relation: 框架元素的定义, 框架内推理

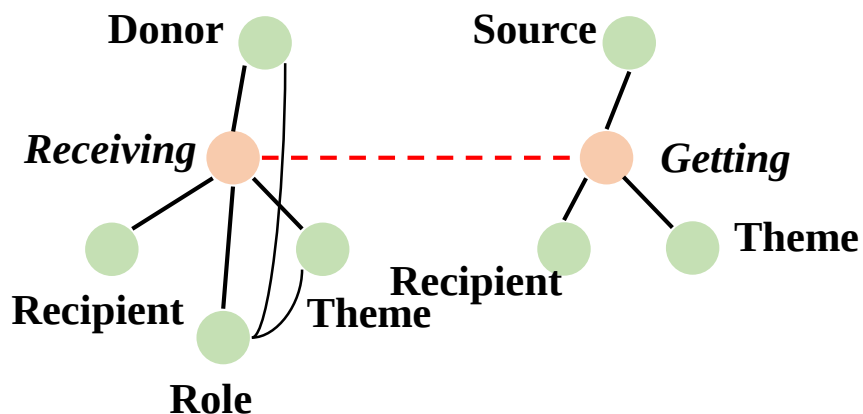
FE Definition

Role: A role that the **Theme** is intended to fill by the **Donor**



知识驱动的框架语义解析

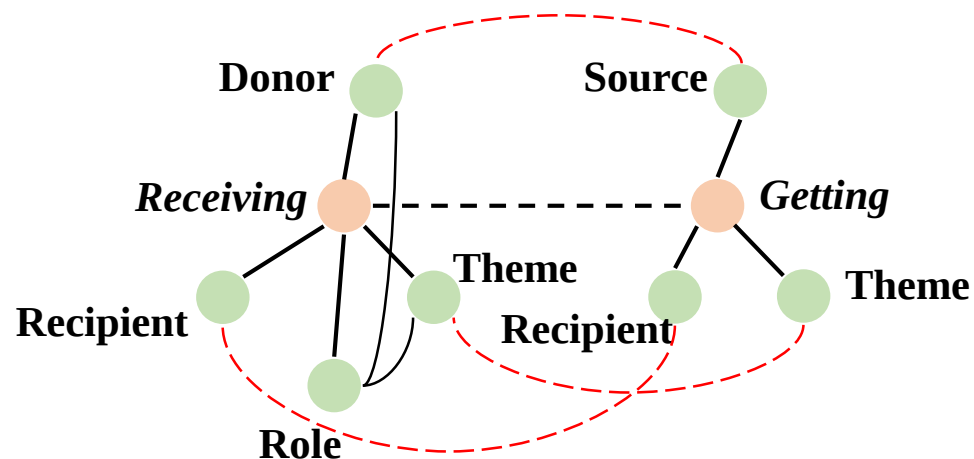
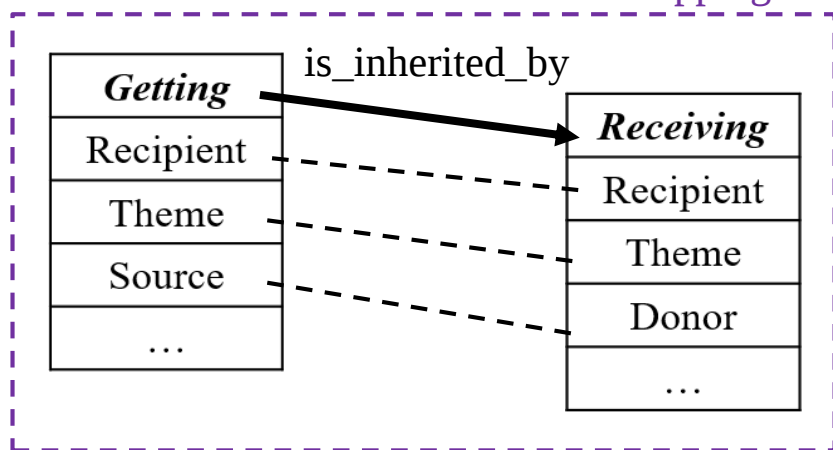
- 结点: 框架 (frame) 和 框架元素 (FE)
- 关系类型:
 - Frame-FE relations: 表示框架识别和其他子任务的交互
 - Intra-Frame FE-FE relation: 框架元素的定义, 框架内推理
 - Frame-Frame relation: *Receiving* 继承自 *Getting*, 框架间推理



知识驱动的框架语义解析

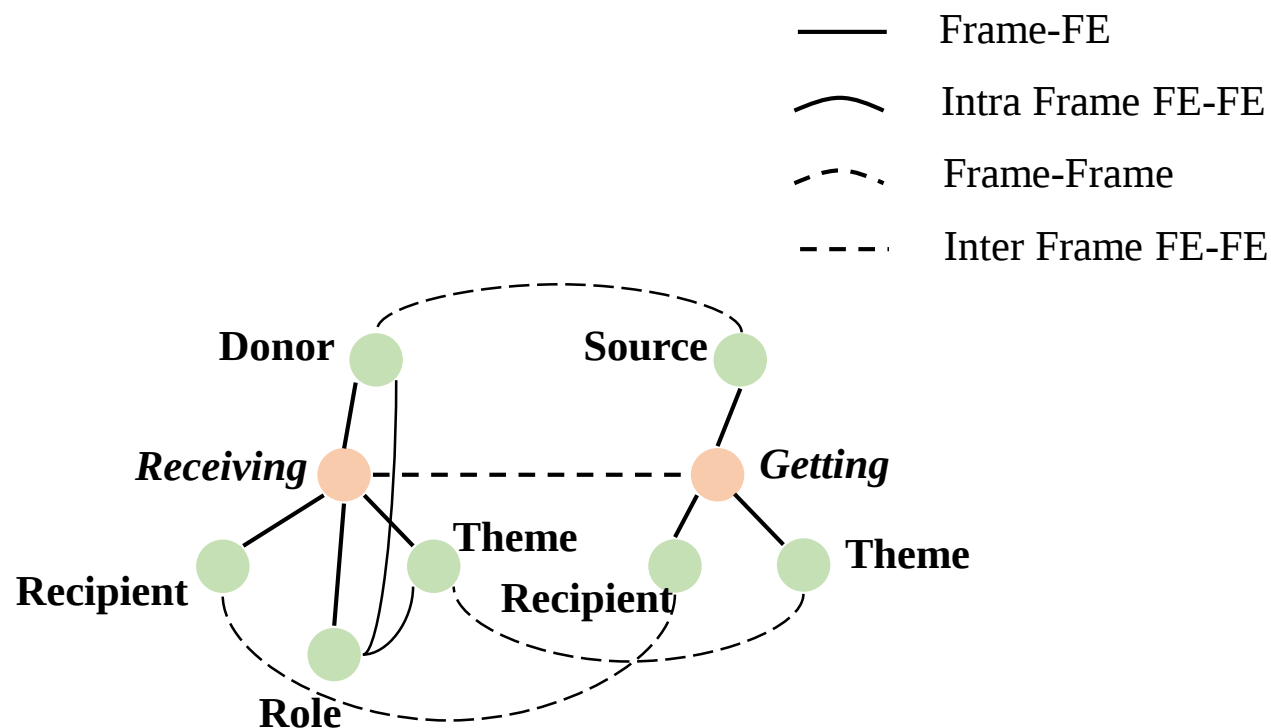
- 结点: 框架 (frame) 和 框架元素 (FE)
- 关系类型:
 - Frame-FE relations: 表示框架识别和其他子任务的交互
 - Intra-Frame FE-FE relation: 框架元素的定义, 框架内推理
 - Frame-Frame relation: *Receiving* 继承自 *Getting*, 框架间推理
 - Inter-frame FE-FE relations: 框架元素映射, 框架间推理

Frame Semantic Relations & FE mappings



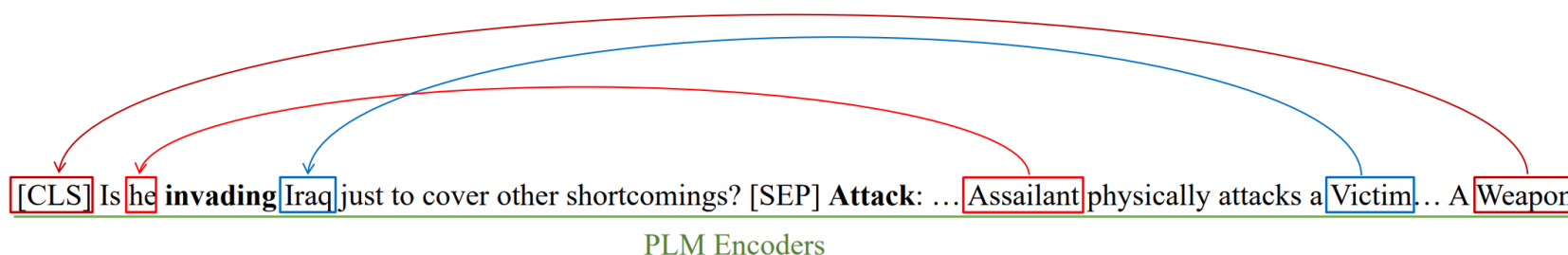
知识驱动的框架语义解析

- 结点: 框架 (frame) 和 框架元素 (FE)
- 关系类型:
 - Frame-FE relations
 - Intra-Frame FE-FE relations
 - Frame-Frame relations
 - Inter-frame FE-FE relations
- 使用RGCN对这种多关系图进行建模



知识驱动的框架语义解析

- 在基于微调范式的框架语义角色标注(AGED)中
 - 利用框架及元素定义实现了框架内元素关系建模
 - 基于框架及元素定义实现了元素的标签语义建模



Ce Zheng, Yiming Wang, Baobao Chang, Query Your Model with Definitions in FrameNet: an Effective Method for Frame Semantic Role Labeling, AAAI 2023.

跨框架知识迁移能力

- 对本体知识的建模赋予模型跨框架知识迁移能力

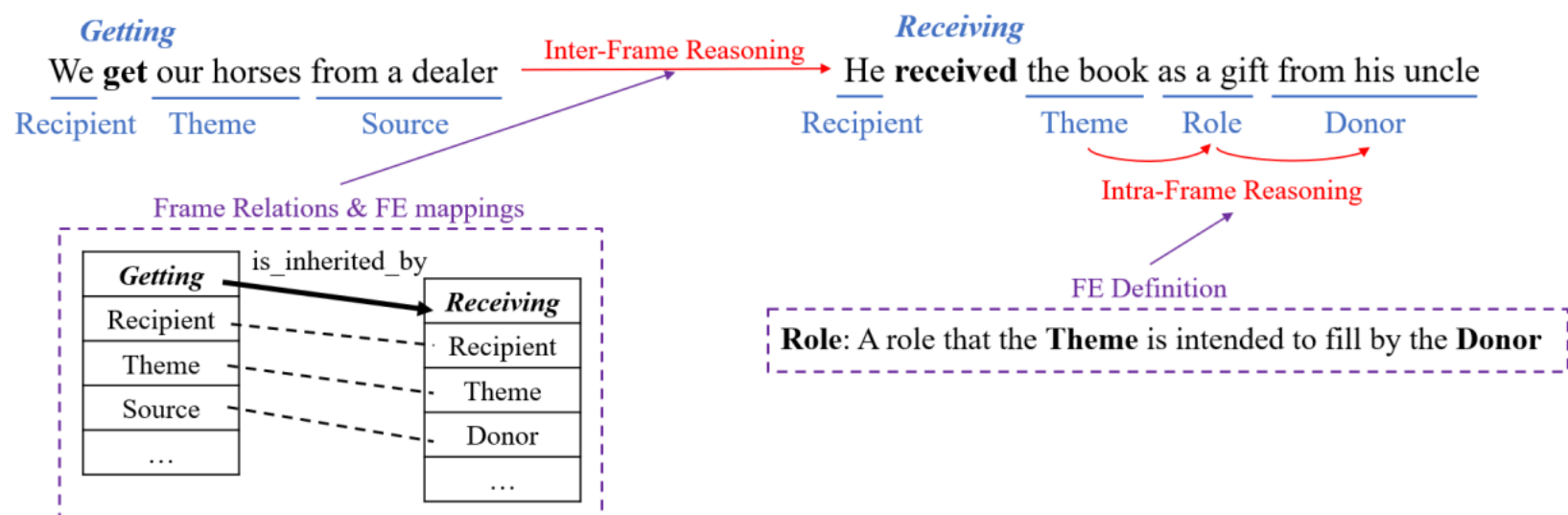


Figure 2: An example of how frame knowledge contributes to frame semantic parsing. The frame semantic relations and FE mappings guide inter-frame reasoning (from left sentence to right); and the FE definitions help with intra-frame reasoning (**Theme** to **Role** and **Role** to **Donor**).

跨框架知识迁移能力

对本体知识的建模赋予模型跨框架知识迁移能力

Model	K				
	0	4	16	32	full
KID (GloVe)	56.26	63.28	65.32	65.95	70.32
w/o FKG	0.00	50.70	56.40	57.59	63.94
Δ	56.26	12.58	8.92	8.36	6.38

Table 6: Experiments on confirming transfer learning ability of FKG. K denotes the number of instances of each frame in training set. Full means adding all instances of these frames except those including target *get* in train and development sets. Lack of labeled instances has much less impact on Arg F1 performance of KID with FKG, which confirms our assumption.

跨框架知识迁移能力

- 在微调范式中，对标签语义的建模也展现出知识迁移能力

Model	<i>K</i>		
	0	32	full
KID (GloVe)	56.26	65.95	70.32
AGED	82.45	84.06	86.26

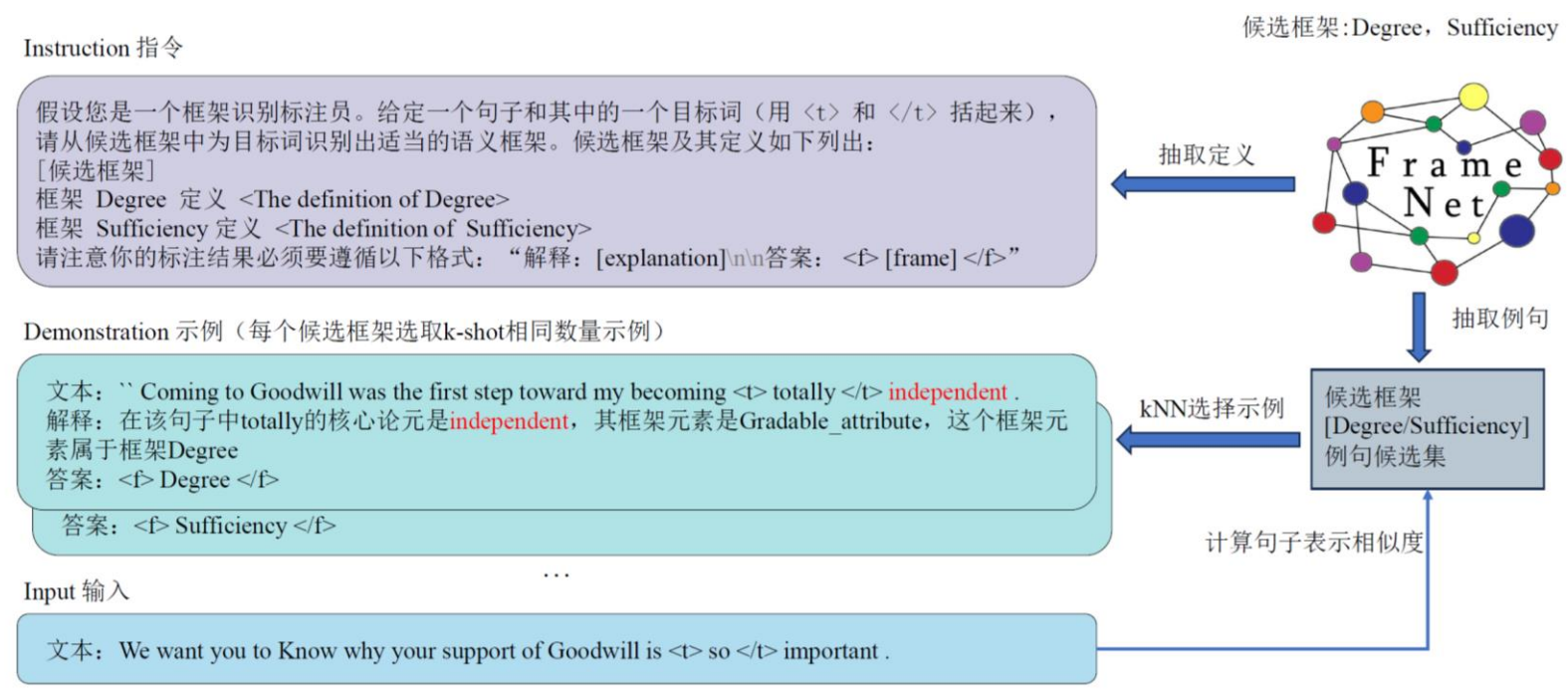
Model	Precision	Recall	F1-score	Δ
15 $\rightarrow$ 17	73.28	74.17	73.72	-1.01
15 $\rightarrow$ 17 (new)	71.85	73.90	72.86	-1.87
17 $\rightarrow$ 17	74.02	75.46	74.73	-

主要内容

- 框架语义解析简介
- 主要研究内容
 - 框架语义解析的建模研究
 - 知识驱动的框架语义解析
 - **大模型和小模型差异分析及级联**
- 总结

基于提示范式的框架识别（大模型）

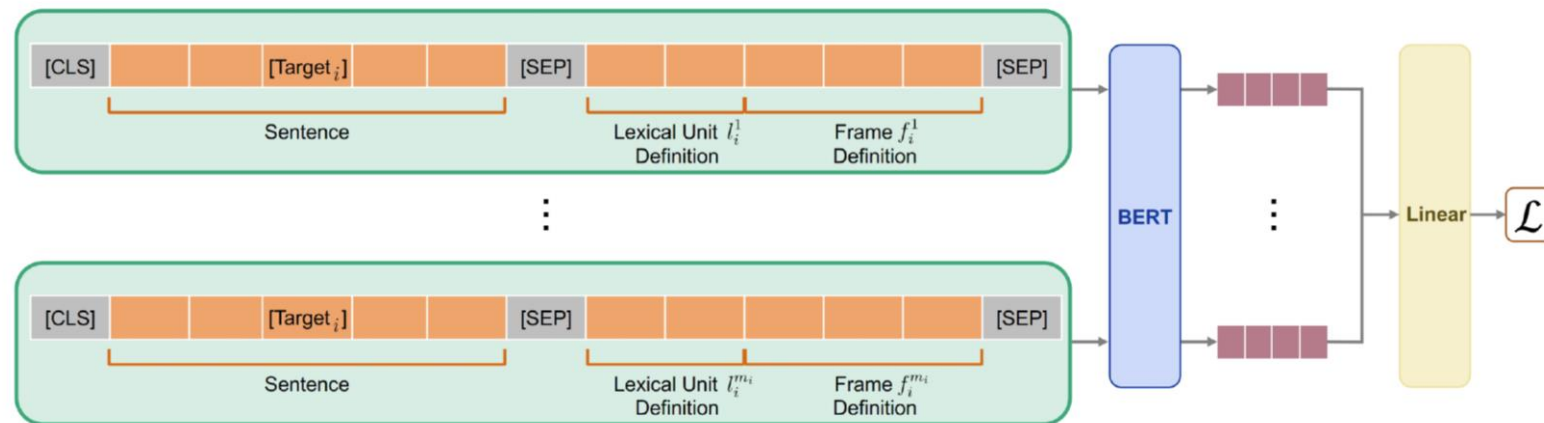
- 基于ChatGPT
- 撰写指令
- 选择示例(KNN)
- 构造思维链



基于框架元素信息构造思维链，在识别核心框架元素的基础上识别框架

基于微调模式的框架识别（小模型）

- FIDO
- FIDO+
 - 统一了候选框架数量
- FIDO+AGED
 - 与角色标注联合训练



小模型结果

- | | |
|-------------|------|
| • FIDO | 91.3 |
| • FIDO+ | 91.6 |
| • FIDO+AGED | 92.1 |

大模型结果

- | | |
|-------------------|-------|
| • ChatGPT(0-shot) | 83.06 |
| • ChatGPT(4-shot) | 85.21 |

Tianyu Jiang, Ellen Riloff. Exploiting Definitions for Frame Identification, EACL 2021.

基于大模型和小模型的框架语义解析差异

- 小模型
 - 经过框架语义标注数据训练，具有框架语义的专门知识
- 大模型
 - 未经微调，但参数规模巨大且经过大规模预训练、指令微调和对齐微调，具有广泛的通用知识
- 猜测
 - 对于标注例子贫乏的框架而言，小模型应该有局限性
 - 对于小模型处理不好的句子，大模型应该可以处理的更好
 - 单纯从提高解析精度的角度出发，可以把小模型处理不好的句子交给大模型处理

- 困难样本
 - 小模型识别置信度较低的样本
- 级联策略CAP-Fi
- 在指令中提供更多示例，有助于提升大模型处理困难样本的能力

模型	准确率	困难样本准确率
FIDO+AGED	92.07	54.01
FIDO+AGED (4-shot)	81.40	-
ChatGPT (0-shot)	83.06	48.66
ChatGPT (4-shot)	85.21	59.36
CAP-Fi (不含思维链)	92.43	62.03
CAP-Fi	92.70	67.91

主要内容

- 框架语义解析简介
- 主要研究内容
 - 框架语义解析的建模研究
 - 知识驱动的框架语义解析
 - 大模型和小模型差异分析及级联
- 总结

- 探讨了基于不同范式的框架语义解析模型，得益于预训练语言模型，框架语义分析效果提升的同时，模型结构越来越简单
- 在框架语义分析过程中引入框架语义本体知识有助于提升模型的跨框架知识迁移能力，提升模型性能
- 大模型在未见过训练例子或仅见过少量例子的前提下，可以取得可观的处理效果，在处理困难样本方面具有一定优势

谢谢大家!