清華大學 Tsinghua University

论文分享

刘柱

2023.10.17

大纲

- 语义组合性相关的研究
- A Systematic Search for **Compound Semantics** in Pretrained BERT Architectures （EACL 2023）
- From chocolate bunny to chocolate crocodile: Do Language Models Understand **Noun Compounds**? (ACL 2023)

与组合性相关的文章

1. 组合性成分的定义和识别
2. 组合性泛化（Compositional Generalization）
3. 组合性的评估和测量

4. 英语名词性的组合性短语（合成词）研究

1. 组合性成分的定义和识别

(1) 蒙塔古语义学：整体的语义是其各**部分语义**和它们语义组合**规则**的函数。
往往用来解释短语或者句子的语义

(1) 更加宽泛的定义：整体可以由部分以及一套程序解释/推断出来
不限于自然语言，只要具备一定有规则的结构即可

组合性泛化：从已知的组合性构造中推断由未知的词汇或者规则组合的新成分

1.1 语义组合性（蒙塔古语义学）

- Dataset - COGS (语义解析任务)

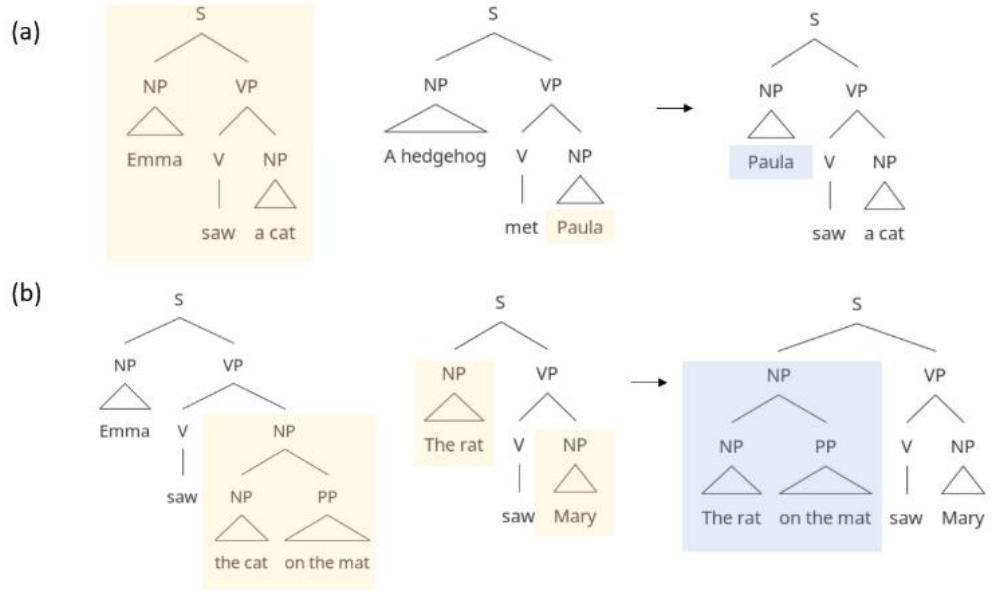


Figure 3: (a) Lexical generalization: a novel combination of a familiar primitive and a familiar structure. (b) Structural generalization: a novel combination of two familiar structures.

- Dataset - COGS (语义解析任务)

Case	Training	Generalization
S.3.1. Novel Combination of Familiar Primitives and Grammatical Roles		
Subject → Object (common noun)	A **hedgehog** ate the cake.	The baby liked the **hedgehog**.
Subject → Object (proper noun)	**Lina** gave the cake to Olivia.	A hero shortened **Lina**.
Object → Subject (common noun)	Henry liked a **cockroach**.	The **cockroach** ate the bat.
Object → Subject (proper noun)	The creature grew **Charlie**.	**Charlie** worshipped the cake.
Primitive noun → Subject (common noun)	**shark**	A **shark** examined the child.
Primitive noun → Subject (proper noun)	**Paula**	**Paula** sketched William.
Primitive noun → Object (common noun)	**shark**	A chief heard the **shark**.
Primitive noun → Object (proper noun)	**Paula**	The child helped **Paula**.
Primitive verb → Infinitival argument	**crawl**	A baby planned to **crawl**.
S.3.2. Novel Combination Modified Phrases and Grammatical Roles		
Object modification → Subject modification	Noah ate **the cake on the plate.**	**The cake on the table** burned.
S.3.3. Deeper Recursion		
Depth generalization: Sentential complements	Emma said **that** Noah knew **that** the cat danced.	Emma said **that** Noah knew **that** Lucas saw **that** the cat danced.
Depth generalization: PP modifiers	Ava saw the ball **in the bottle on the table**.	Ava saw the ball **in the bottle on the table on the floor**.

- Dataset - COGS (语义解析任务)

S.3.4. Verb Argument Structure Alternation		
Active → Passive	The crocodile **blessed** William.	A muffin **was blessed**.
Passive → Active	The book **was squeezed**.	The girl **squeezed** the strawberry.
Object-omitted transitive → Transitive	Emily **baked**.	The giraffe **baked a cake**.
Unaccusative → Transitive	The glass **shattered**.	Liam **shatterd** the jigsaw.
Double object dative → PP dative	The girl **teleported** Liam the cookie.	Benjamin **teleported** the cake **to** Isabella.
PP dative → Double Object Dative	Jane shipped the cake to John.	Jane shipped John the cake.
S.3.5. Verb Class		
Agent NP → Unaccusative subject	The **cobra** helped a dog.	The cobra **froze**.
Theme NP → Object-omitted transitive subject	The hippo **decomposed**.	The hippo **painted**.
Theme NP → Unergative subject	The hippo **decomposed**.	The hippo **giggled**.

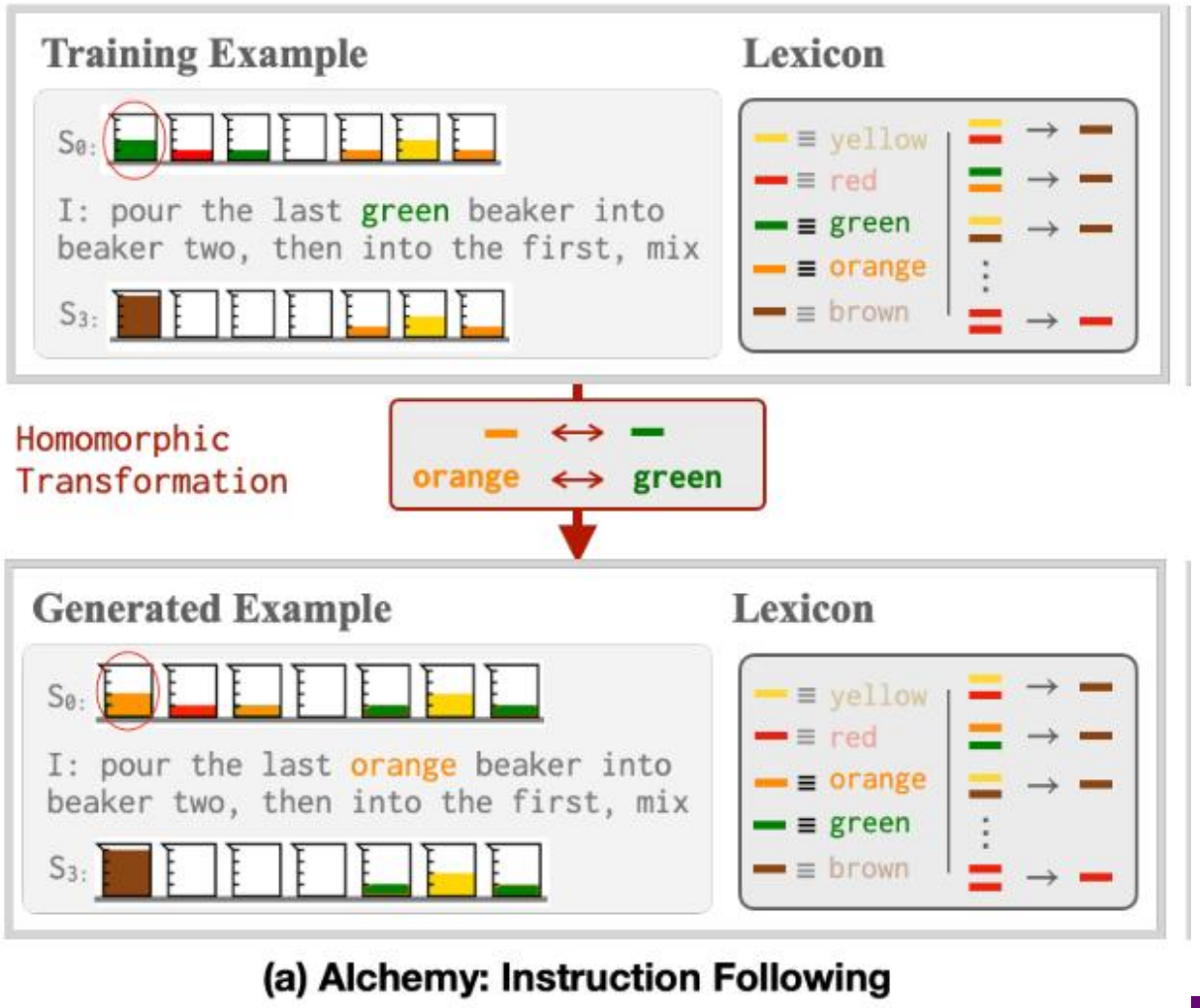
Dataset - SCAN (指令理解)

jump	⇒	JUMP
jump left	⇒	LTURN JUMP
jump around right	⇒	RTURN JUMP RTURN JUMP RTURN JUMP RTURN JUMP
turn left twice	⇒	LTURN LTURN
jump thrice	⇒	JUMP JUMP JUMP
jump opposite left and walk thrice	⇒	LTURN LTURN JUMP WALK WALK WALK
jump opposite left after walk around left	⇒	LTURN WALK LTURN WALK LTURN WALK LTURN WALK LTURN LTURN JUMP

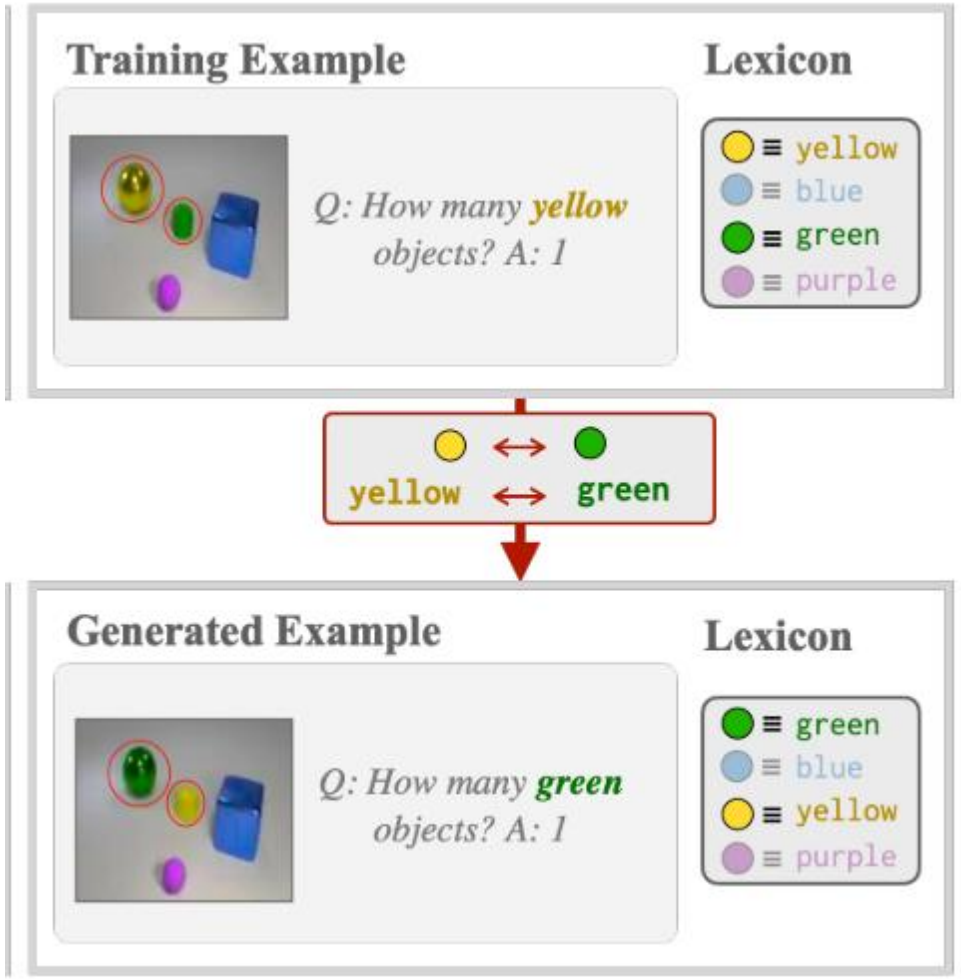
Figure 1. Examples of SCAN commands (left) and the corresponding action sequences (right).

测试时候会出现新指令的组合

1.1 更宽泛的组合性



(a) Alchemy: Instruction Following



(c) CLEVR-CoGenT: VQA

2. 组合性泛化的方法

一、基于模型的

学习策略：对比学习

二、基于数据的

数据增强

示例设计

对比学习策略

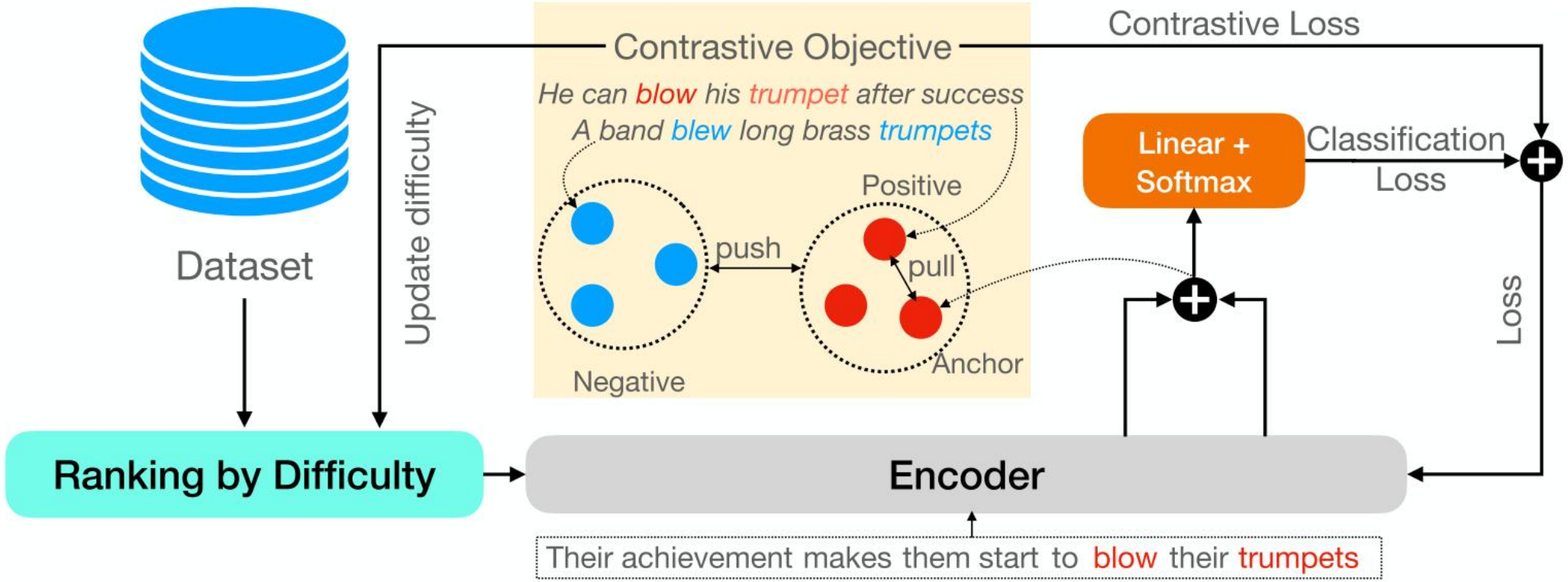
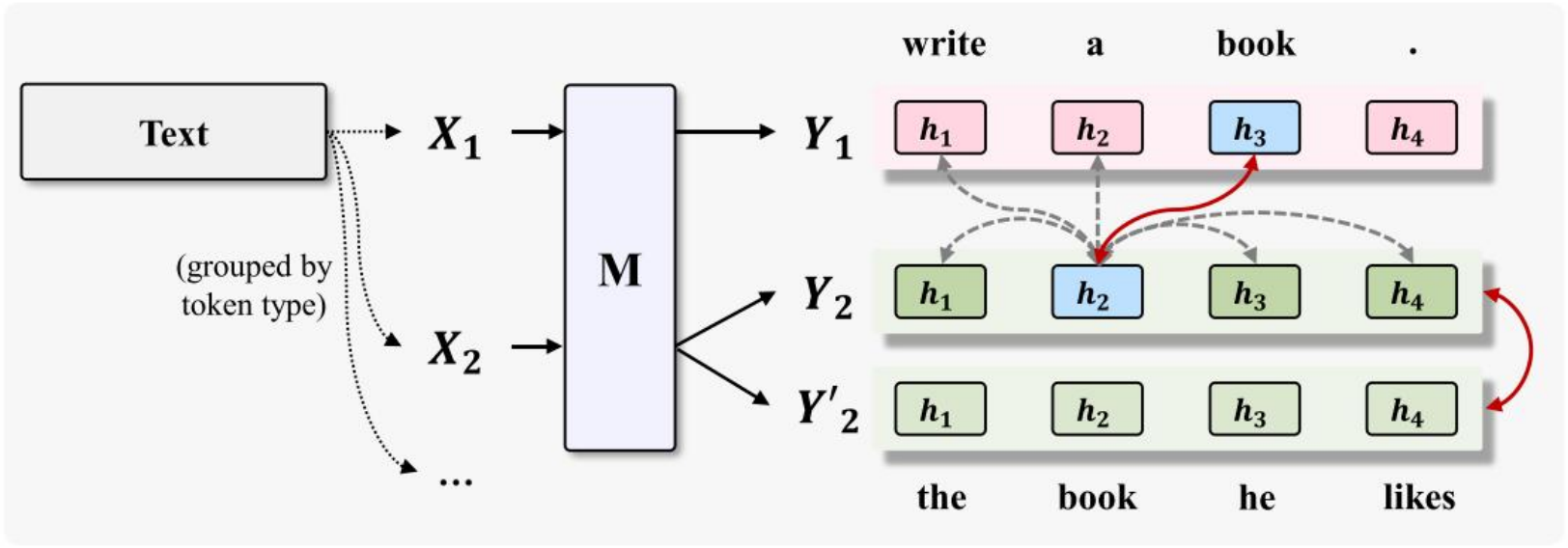


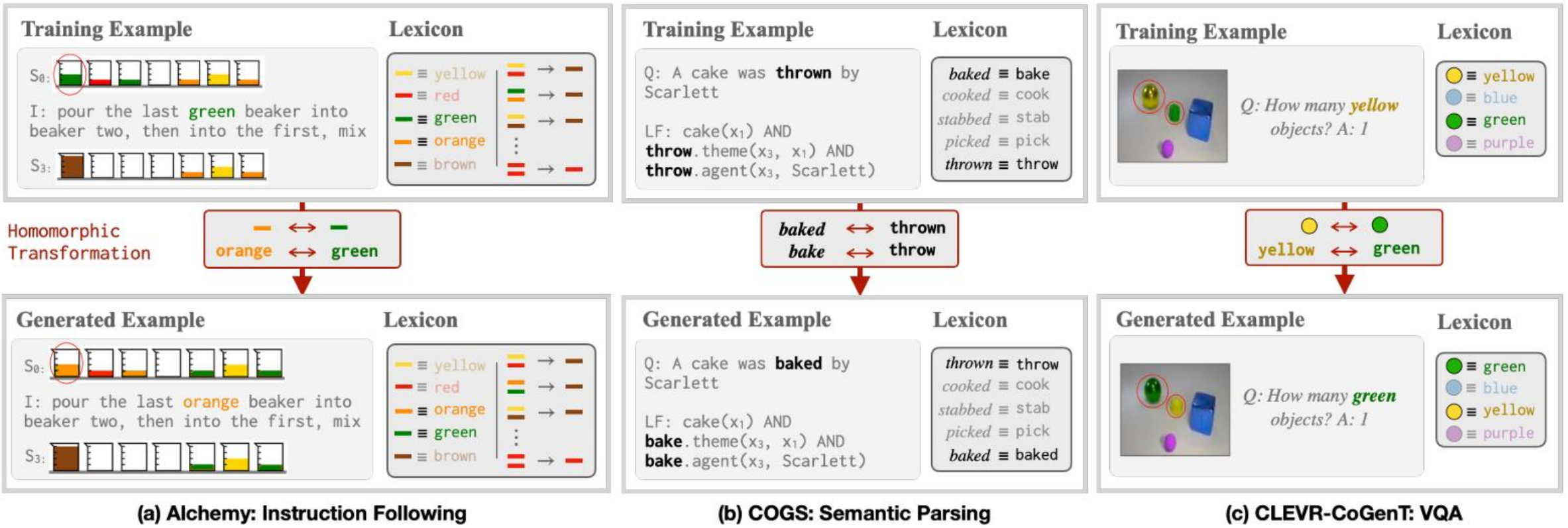
Figure 1: The diagram illustrates the CLCL framework.

CLCL: Non-compositional Expression Detection with Contrastive Learning and Curriculum Learning （ACL'23）

对比学习策略


Text
(grouped by token type)
X_1
X_2
...
M
Y_1
Y_2
Y'_2
write a book .
h_1 h_2 h_3 h_4
the book he likes
write a book
the book on wildlife
book
got a book
checked a book out
Model
Consistency Regularization
Contextualized Representation
Context
Token Boundary

数据增强


Training Example
Lexicon
I: pour the last green beaker into beaker two, then into the first, mix
yellow
red
green
orange
brown
Homomorphic Transformation
orange ↔ green
Generated Example
Lexicon
I: pour the last orange beaker into beaker two, then into the first, mix
(a) Alchemy: Instruction Following
Training Example
Lexicon
Q: A cake was thrown by Scarlett
LF: cake(x1) AND throw.theme(x3, x1) AND throw.agent(x3, Scarlett)
baked ≡ bake
cooked ≡ cook
stabbed ≡ stab
picked ≡ pick
thrown ≡ throw
baked ↔ thrown
bake ↔ throw
Generated Example
Lexicon
Q: A cake was baked by Scarlett
LF: cake(x1) AND bake.theme(x3, x1) AND bake.agent(x3, Scarlett)
thrown ≡ throw
cooked ≡ cook
stabbed ≡ stab
picked ≡ pick
baked ≡ baked
(b) COGS: Semantic Parsing
Training Example
Lexicon
Q: How many yellow objects? A: 1
yellow
blue
green
purple
yellow ↔ green
Generated Example
Lexicon
Q: How many green objects? A: 1
green
blue
yellow
purple
(c) CLEVR-CoGenT: VQA

数据增强

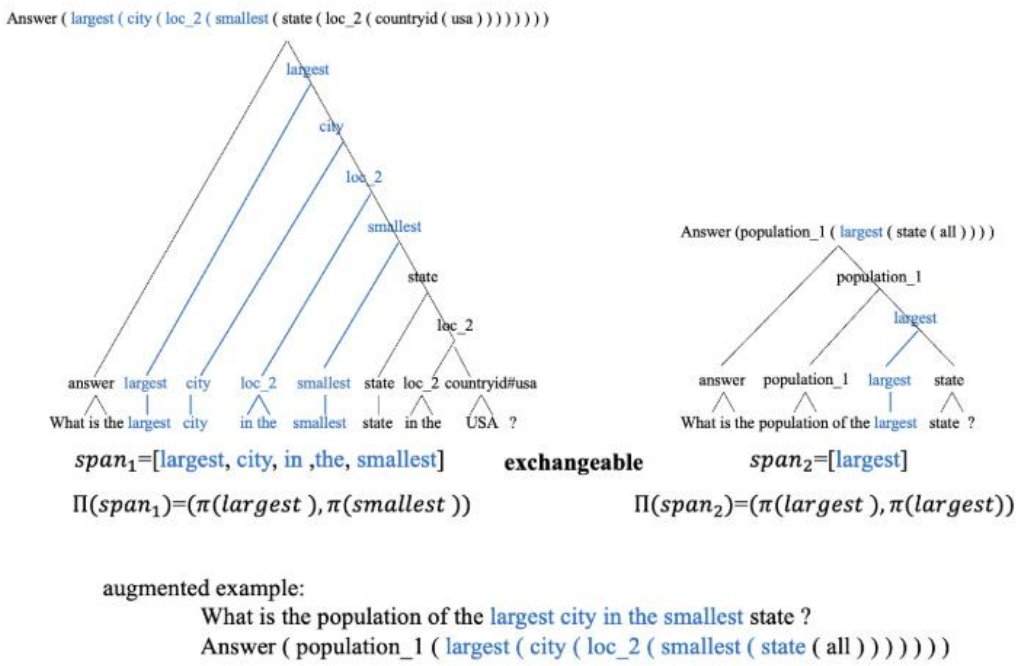


Figure 2: An augmentation example by SpanSub. SpanSub substitutes a span "largest" with another span "largest city in the smallest", and augments a new question "What is the population of the largest city in the smallest state?".

示例学习

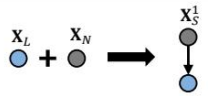
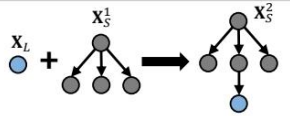
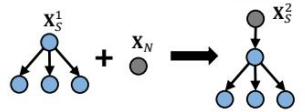
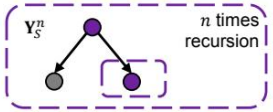
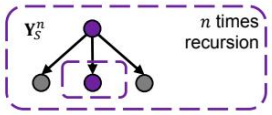
Category	In-Context Examples	Test Case	Illustration of Combination
Primitive Substitution	input: **shark** output: SHARK input: A girl drew the boy . output: DRAW (GIRL , BOY , NONE)	input: **The shark** drew a boy . output: DRAW (SHARK , BOY , NONE)	
Primitive Structural Alternation	input: The goose **baked** . output: BAKE (GOOSE , NONE , NONE) input: A teacher noticed a chicken . output: NOTICE (TEACHER , CHICKEN , NONE)	input: A teacher **baked** the chicken . output: BAKE (TEACHER , CHICKEN , NONE)	
Phrase Recombination	input: Logan mailed Stella **the cake in the pile** . output: MAIL (LOGAN , IN (CAKE , PILE) , STELLA) input: The goose rolled **a baby in a room** . output: ROLL (GOOSE , IN (BABY , ROOM) , NONE)	input: **A visitor in the pile** rolled a resident . output: ROLL (IN (VISITOR , PILE) , RESIDENT , NONE)	
Longer Chain	input: The boy admired that Noah confessed that \ Emma was given a cookie . output: ADMIRE (BOY , NONE , NONE) \ **CCOMP** CONFESS (NOAH , NONE , NONE) \ **CCOMP** GIVE (NONE , COOKIE , EMMA)	input: The girl wished that a crocodile declared that \ the boy admired that Emma liked that \ Evelyn was passed a drink . output: WISH (GIRL , NONE , NONE) \ **CCOMP** DECLARE (CROCODILE , NONE , NONE) \ **CCOMP** ADMIRE (BOY , NONE , NONE) \ **CCOMP** LIKE (EMMA , NONE , NONE) \ **CCOMP** PASS (NONE , DRINK , EVELYN)	
Deeper Nesting	input: Noah appreciated a girl in a house \ beside the chair . output: APPRECIATE (NOAH , \ **IN (** GIRL , \ **BESIDE (** HOUSE , CHAIR\ **))** , NONE)	input: A dog painted the girl beside the chair \ in a house beside a road on a dish . output: PAINT (DOG , \ **BESIDE (** GIRL , \ **IN (** CHAIR , \ **BESIDE (** HOUSE , \ **ON (** ROAD , DISH\ **))))** , NONE)	

Figure 2: Five categories of aiming combinations. The key parts in combinations are marked with underlines and colors (blue in NL-side and purple in code-side). The last column follows the notations defined in Section 3.2.

3. 组合性的评估和测量

1）下游任务的准确性，往往在新的组合性

任务：语义解析任务、机器翻译任务、问答任务等等

2）从表征的角度

TRE: Tree Reconstruction Error

“...to treat a set of primitive meaning representations as hidden, and optimize over them to find an explicitly compositional model that **approximates** the true model as well as possible…”

Measuring compositionality in representation learning. (ICLR, 2019)

4. 英语名词性的合成词研究

1. English Noun-Noun Compound
2. 主要的相关任务：

1）合成词解释 （Compound Interpretation）解释合成词部分之间的关系

2）合成词推断（Compound Conceptualization）解释新词部分之间的关系

1. 主要工作

1）BERT模型的表征是否学习到了与人类一致的组合性 (3篇文章提到0.58-0.7之间)

2）合成词解释与推断 (free paraphrase or template-based)

现有工作不足

（1）缺乏名名复合外的类型探究

（2）把组合性当作一个预测或者相似性任务，没有更加细粒度地探究组合性的特征。例如对合成词内部进行一些插入、重复、删除等操作如何改变表征

（3）缺乏对汉语这类以复合构词为主的语言的研究

A Systematic Search for Compound Semantics in Pretrained BERT Architectures

Filip Miletic and **Sabine Schulte im Walde**
Institute for Natural Language Processing, University of Stuttgart
{filip.miletic, schulte}@ims.uni-stuttgart.de

EACL 2023 main; Cited by 3

背景

1. 多词表达（合成词）组合性是一个连续统
2. 静态词向量模型 （word2vec等） vs. 动态词表示模型（BERT等）
3. 已有研究(Garcia et al., 2021a).表明静态词向量更可以捕捉到组合性
4. 本文认为动态词表示的能力不足很有可能是由于未充分利用模型编码的信息
5. 本文贡献：

（1）更加全面地探究了预训练BERT模型对于合成词组合性的捕捉。

（2）探讨了影响模型组合性的一些经验性因素

（3）一些指导性的发现：一定程度编码了多词表达的语义；初始层

相关研究

- 心理语言学：语义透明度
- 探讨一系列合成词特性（词义预测、部分间的语义关系、组合性）
- 不同的语言，包括英语、德语、法语、葡萄牙语
- 影响组合性的因素探究（频率、能产性、歧义性）
- 在与人类评分的相关性指标上，有研究指出静态词向量的相关性更好

但前人的工作没有考虑更多的参数配置（一般是在token水平对后四层的表征进行平均）

数据

- 280个英语名词性合成词
- 组合性评分：0-5（从非字面义到字面义）；合成词的组合性以及各部分的贡献程度

Compound	Compositionality rating		
	Modifier	Head	Phrase
guinea pig	0.47 ± 0.72	0.47 ± 0.72	0.24 ± 0.56
flea market	0.38 ± 0.81	4.71 ± 0.84	1.52 ± 1.13
biological clock	4.71 ± 0.47	1.76 ± 1.35	2.29 ± 1.21
health insurance	4.53 ± 0.88	4.83 ± 0.58	4.40 ± 1.17

Table 1: Sample gold standard compounds with compositionality ratings (mean and standard deviation).

其他

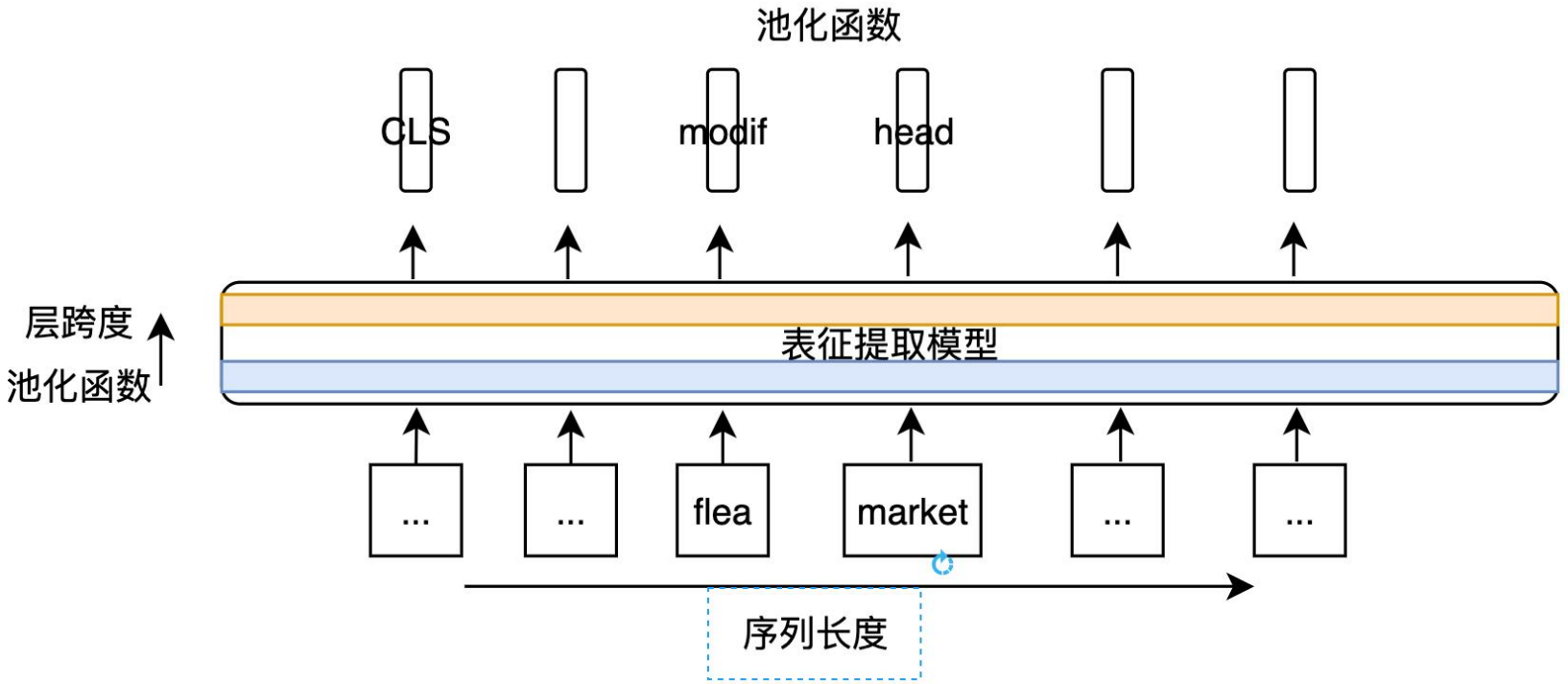
- 语料库 ENCOW：用来提供合成词所在的上下文
- 经验的影响组合性的因素：

  （1）频率

  （2）能产性

  （3）歧义性。词网 WordNet

实验设定



- BERT-base-uncased, 768d, 12层
- 表示类型：modif, head, comp, cont, cls
- 池化函数：1）词级别；2）层级别；3）类符级别
- 层跨度组合：
- 序列长度
- 序列数量

组合性估计

- 直接估计：从<head, modif, comp, cont, cls>两两计算余弦相似性
- 组合性函数：<comp, cont, cls>中任一个与<head, modif>做如下函数

$$\begin{aligned} \text{ADD} &= \cos(\text{comp}, \text{modif}) + \cos(\text{comp}, \text{head}) \\ \text{MULT} &= \cos(\text{comp}, \text{modif}) \cdot \cos(\text{comp}, \text{head}) \\ \text{COMB} &= \text{ADD} + \text{MULT} \end{aligned}$$

- 相同形符实例汇总：

  类符：先池化表征，再进行计算

  形符：先计算得分，再平均得分

结果1:

	ρ	layers	pool	len	seqs	estimate		agg	ρ	layers	pool	len	seqs	estimate		agg
COMP	0.706	1-1	sum	3	1k	COMB	cont	token	-0.642	2-3	avg	3	1k	cont	cls	type
	0.706	1-1	avg	3	1k	COMB	cont	token	-0.644	3-3	avg	3	1k	cont	cls	type
	0.706	1-1	sum	20	1k	MULT	cont	token	-0.645	1-5	avg	3	1k	cont	cls	type
	0.706	1-1	avg	20	1k	MULT	cont	token	-0.646	2-4	avg	3	1k	cont	cls	type
	0.706	1-1	sum	3	1k	MULT	cont	token	-0.649	1-4	avg	3	1k	cont	cls	type
HEAD	0.645	1-1	sum	3	1k	head	cont	token	-0.598	0-7	avg	3	1k	cont	cls	type
	0.645	1-1	avg	3	1k	head	cont	token	-0.599	1-4	avg	3	1k	cont	cls	type
	0.638	1-1	sum	3	1k	COMB	cont	token	-0.600	0-6	avg	3	1k	cont	cls	type
	0.638	1-1	avg	3	1k	COMB	cont	token	-0.604	1-5	avg	3	1k	cont	cls	type
	0.638	1-1	sum	3	1k	ADD	cont	token	-0.606	1-6	avg	3	1k	cont	cls	type
MODIF	0.553	1-1	avg	20	1k	modif	cont	token	-0.464	2-4	avg	3	1k	cont	cls	type
	0.553	1-1	sum	20	1k	modif	cont	token	-0.465	1-5	avg	3	1k	cont	cls	type
	0.548	1-1	sum	3	1k	modif	cont	token	-0.471	1-3	avg	3	1k	cont	cls	type
	0.548	1-1	avg	3	1k	modif	cont	token	-0.474	1-4	avg	3	1k	cont	cls	type
	0.546	1-1	avg	20	1k	modif	cont	type	-0.476	1-2	avg	3	1k	cont	cls	type

Table 2: Best (left) and worst (right) evaluated implementations. Abbreviations: *pool* = pooling function; *len* = minimum tokens per sequence; *seqs* = minimum number of modeled sequences; *agg* = aggregation of occurrences (type vs. token-level).

- 其他研究中静态向量 0.726 vs. BERT 0.37
- 搜索空间大（41,496），变化也比较大
- 最佳配置vs. 最差配置

各项因素

	min. 3 tokens	min. 20 tokens
COMP	0.146 (-0.649, 0.706)	0.134 (-0.587, 0.706)
HEAD	0.102 (-0.606, 0.645)	0.087 (-0.561, 0.637)
MODIF	0.099 (-0.476, 0.548)	0.093 (-0.460, 0.553)

Table 3: Spearman's ρ (mean, min, max) for minimum sequence length.

	avg	sum
COMP	0.139 (-0.649, 0.706)	0.141 (-0.587, 0.706)
HEAD	0.094 (-0.606, 0.645)	0.095 (-0.563, 0.645)
MODIF	0.095 (-0.476, 0.553)	0.097 (-0.460, 0.553)

Table 5: Spearman's ρ (mean, min, max) for pooling functions.

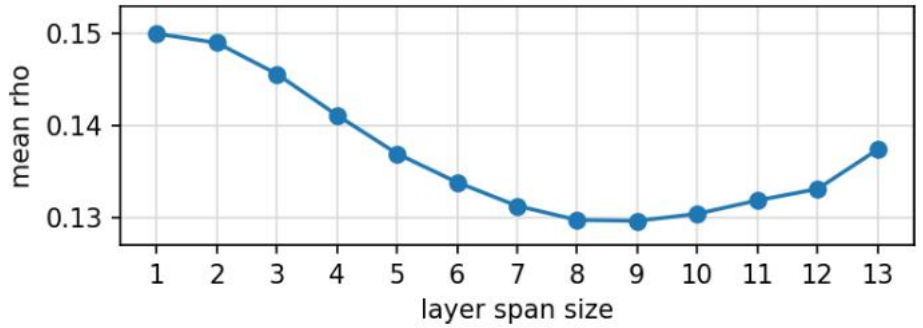
	10 sequences	100 sequences	1,000 sequences
C	0.135 (-0.394, 0.622)	0.142 (-0.607, 0.689)	0.143 (-0.649, 0.706)
H	0.093 (-0.384, 0.565)	0.094 (-0.551, 0.621)	0.096 (-0.606, 0.645)
M	0.089 (-0.367, 0.495)	0.101 (-0.459, 0.544)	0.098 (-0.476, 0.553)

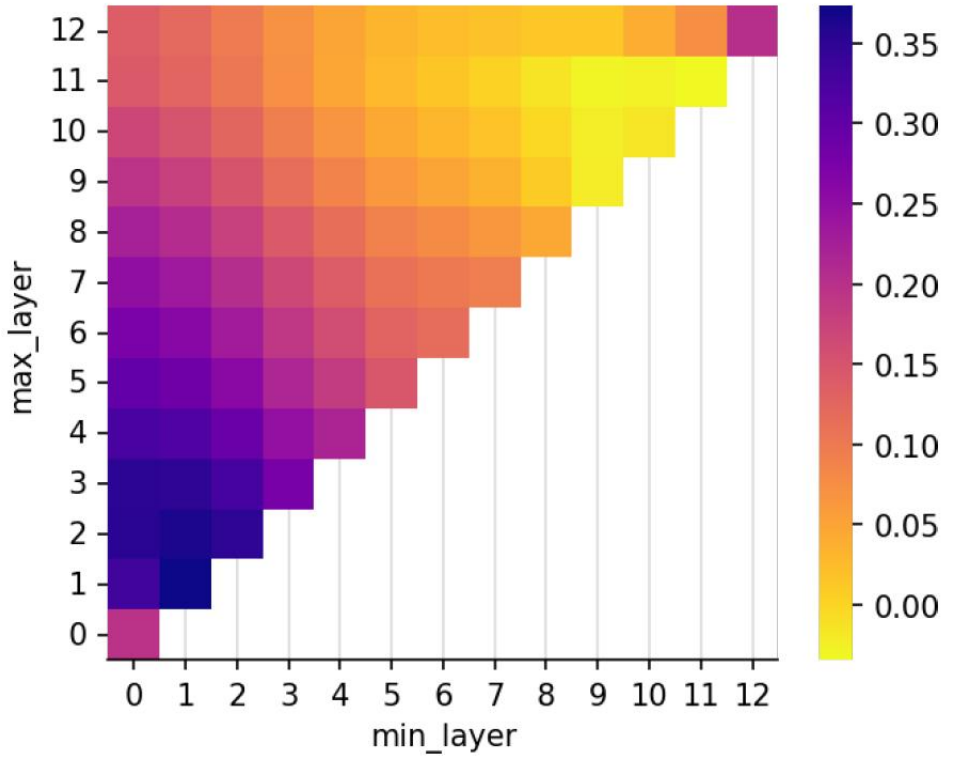
Table 4: Spearman's ρ (mean, min, max) for number of sequences. C, H, M = compound, head, modifier.

	token-level	type-level
COMP	0.150 (-0.584, 0.706)	0.130 (-0.649, 0.699)
HEAD	0.103 (-0.556, 0.645)	0.085 (-0.606, 0.628)
MODIF	0.100 (-0.460, 0.553)	0.092 (-0.476, 0.546)

Table 6: Spearman's ρ (mean, min, max) for token- vs. type-level processing.

各项因素


mean rho
0.15
0.14
0.13
1 2 3 4 5 6 7 8 9 10 11 12 13
layer span size


max_layer
12 11 10 9 8 7 6 5 4 3 2 1 0
0 1 2 3 4 5 6 7 8 9 10 11 12
min_layer
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00

各项因素

- 组合性估计：

Head更加重要

之后会有针对head做的实验

	modif	head	comp	cont	cls
COMP	0.135	0.274	0.245	0.172	-0.128
	-0.383	-0.133	-0.324	-0.649	-0.649
	0.615	0.630	0.666	0.706	0.611
HEAD	0.071	0.242	0.194	0.130	-0.161
	-0.384	-0.130	-0.327	-0.606	-0.606
	0.464	0.645	0.598	0.645	0.558
MODIF	0.106	0.167	0.164	0.133	-0.094
	-0.274	-0.130	-0.229	-0.476	-0.476
	0.553	0.415	0.517	0.553	0.477

Table 7: Spearman's ρ (mean, min, max) for embedding types, across all direct and composite estimates if used.

结果2: 消融实验

- 在最佳设置中，变化其中一项参数类型，其他参数保持不变。
- 总趋势与之前的类似，波动主要出现在模型特征的选择上，而不是预处理上

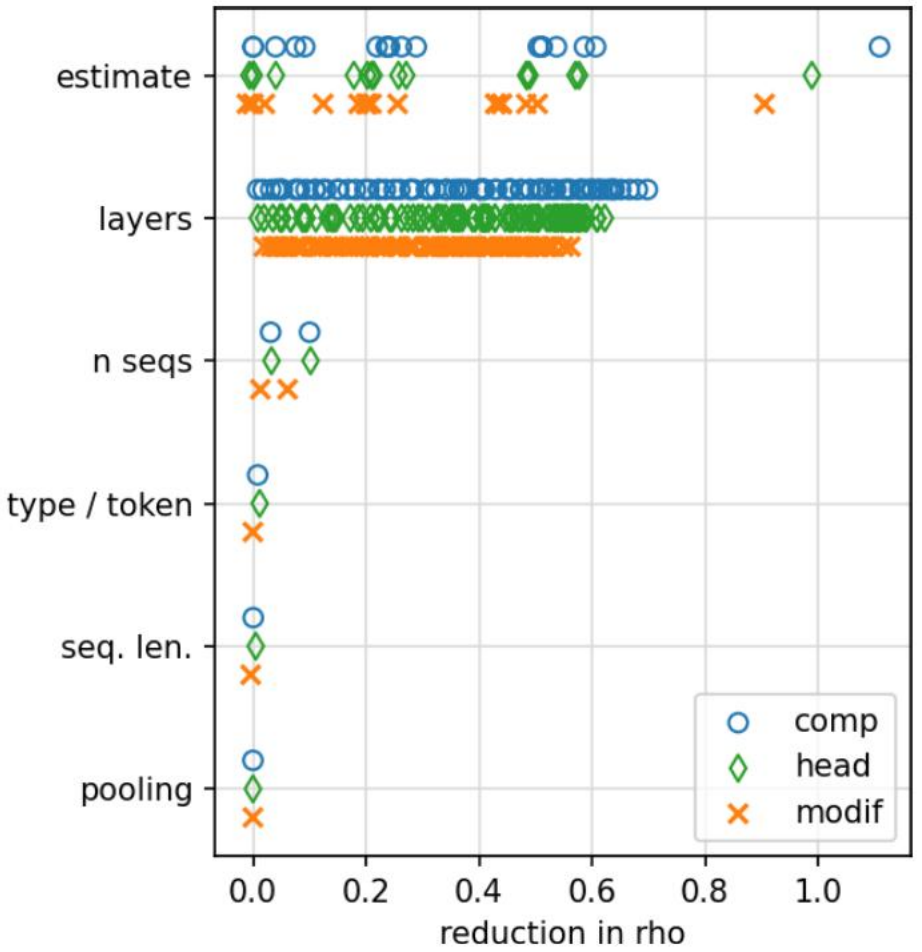


Figure 2: Effect of alternative parameter values compared to the top parameter constellation.

结果3:分析影响因素

- 由于head对于结果影响很大，仅分析head的经验性因素对于组合性评分的影响。
- 三类因素：频率、能产性和歧义性
- 根据值大小，分为五档，取一三五档

Feature	Mean	Std.	Example
Frequency (thousands)	42	± 30	*silver* ***spoon***
	452	± 108	*labor* ***union***
	3,614	± 2,438	*crash* ***course***
Productivity	7	± 5	*night* ***owl***
	75	± 19	*time* ***difference***
	448	± 208	*birth* ***rate***
Ambiguity	2	± 1	*research* ***project***
	5	± 1	*flea* ***market***
	13	± 4	*application* ***form***

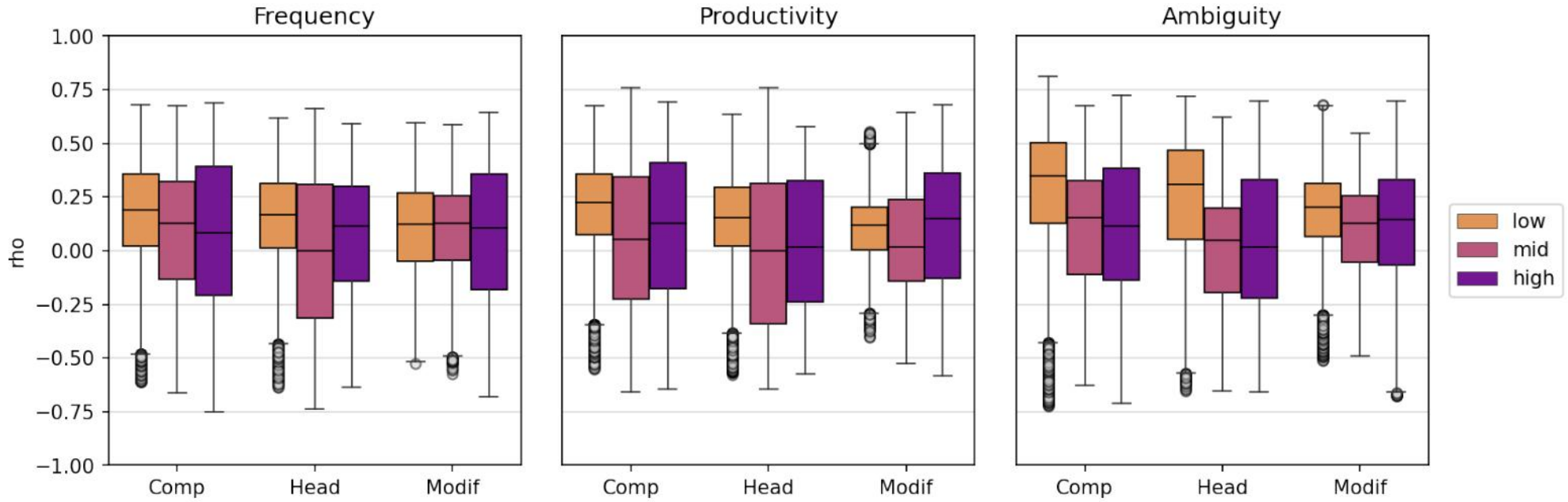


Figure 3: Effect of empirical properties of the head on model performance, observed across the evaluated implementations. Values on the x-axis indicate prediction targets (compound, head, and modifier scores).

- 更低的频率、更低的能产性、更低的歧义性 都会得到更好的组合性
- 三者有相关性

结论

- BERT可以得到较高的相关性（0.706）
- 重要的内部参数：相同形符实例汇总、层数范围、embedding选择
  不重要的内部参数：长度、池化函数
- 外部因素：频率、能产性、多义性

From *chocolate bunny* to *chocolate crocodile*: Do Language Models Understand Noun Compounds?

Jordan Coil[1] **and Vered Shwartz**[1,2]
[1] University of British Columbia [2] Vector Institute for AI
jcoil93@students.cs.ubc.ca, vshwartz@cs.ubc.ca

Findings of ACL 2023

背景

- NCI (Noun Compound Interpretation)发掘组成旧有合成词的部分之间的隐性的语义关系
- NCC（Noun Compound Conceptualization）推断新的合成词的语义关系
- 简约主义、民族主义…“大同主义”

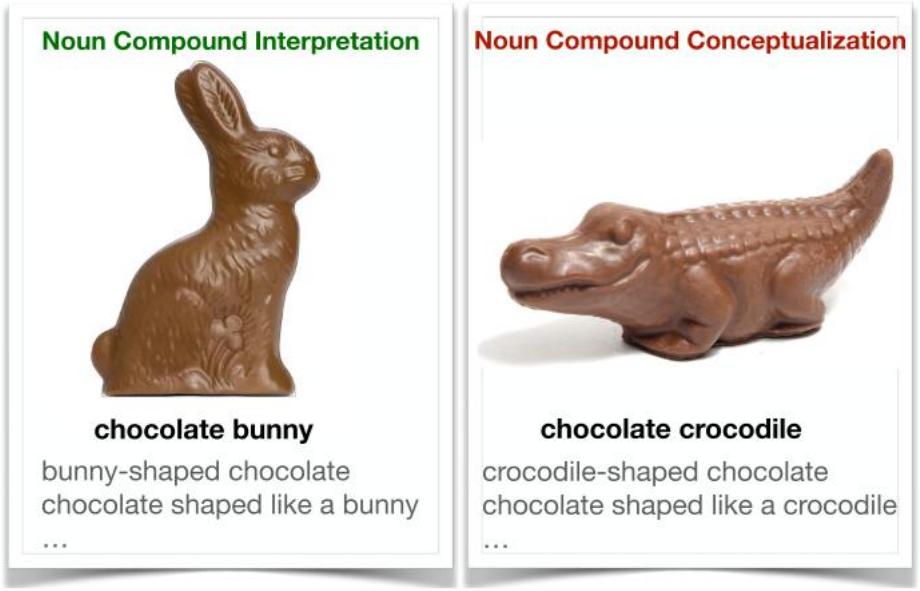


Figure 1: An example NC (input) and paraphrases (output) for each of the NCI and NCC tasks.

名词合成词解释任务NCI

- 早期定义为一个分类任务，预先定义好一些关系类别（介词,动作等， of, from, contains, purpose）-> 类别有限、不正交
- 本文是自由生成关系内容，基于模式[n2]…[n1]；或者完全无限制
- 这一任务主要来源于 SemEval 2013 task 4

形成和解释新概念NCC

- 认知科学中研究人脑如何解释新的名词合成词，以及最终结果的倾向性。例如〝绣球（花）〞主要是参考了绣球的形状，而非颜色
- 认知科学中有工作发现人们对于NC的敏感性反应时间更短，但对于解释该合成词需要更长时间
- NCC的数据收集来自新的NC或者罕见出现的NC，可以通过语料库收集得到。

贡献

- 早期方法还没有在生成式预训练语言模型上评测。
- 本文改善了已有数据集，并在GPT-3上进行了NCI任务的评估。
- 本文进一步探讨了可能的原因，并通过NCC任务进行验证。

数据

- 主要来自SemEval-2023数据集，每个合成词包含一个释义：[n2]…[n1]。
- 本文去除掉了不正确或者不合理的合成词关系释义
- 去除掉了32个测试集中出现在训练集的合成词
- 数据增强：释义同义增强

	Original			Revised		
	train	**dev**	**test**	**train**	**dev**	**test**
#NCs	174	0	181	160	28	110
#paraphrases	4,256	0	8,190	5,441	1,469	4,820

Table 1: Statistics of the original SemEval 2013 dataset (Hendrickx et al., 2013) vs. our revised version (henceforth: the NCI dataset).

方法

- 监督模型：序列到序列的T5-large模型，重新训练
- 少样本学习模型：GPT-3，包含10个示例。
- 人类（MTurk）

```
Q: what is the meaning of <NC>?
A:<paraphrase>
```

评估

NC	GPT-3	T5
access road	road that provides access	road for access
reflex action	a sudden, involuntary response to a stimulus	action performed to perform reflexes
sport page	a page in a publication that is devoted to sports	page dedicated to sports
computer format	the way in which a computer organizes data	format used in computers
grief process	process of grieving or mourning	process that a grief sufferer experiences

Table 2: Example paraphrases generated using GPT-3 and T5 for NCs in the revised SemEval 2013 test set.

Method	METEOR	ROUGE-L	BERTScore	Human
T5	**69.81**	**65.96**	**95.31**	65.35
GPT-3	56.27	47.31	91.94	**95.64**

Table 3: Performance of the T5 and GPT-3 models on the revised SemEval 2013 test set.

- 自动化评估：METEOR、ROUGE-L、BERTScore；以及人类评估
- GPT-3生成的更加多样，人类更加偏向GPT-3

NCC任务评估

- 动机：GPT-3出色的表现是由于它记住了训练集中的样例，还是具备了一种解释能力?
- 本文定义了名词合成词概念化NCC这一任务，去看模型解释新的合成词的能力

数据

- 前人数据集Dhar and van der Plas (2019)，来自Google Ngram Corpus，它的训练集在2000年之前，测试集在2000年之后。
- 由于GPT-3的数据来自于近期，本文选取上述数据集中出现频率最低的500个合成词
- 通过过滤掉一些不恰当的名词合成词、增加社交媒体上出现的新的合成词、最终本文收集到了105个名词合成词。

- 最后使用人类评估的方式评估模型和人类的结果

结果

- GPT-3的结果比人类表现得还好。
- GPT-3可能受益于大量的数据，而神经认知方面的成果显示人类对于新概念也是比较费时的。

Test Set	NCI	NCC
Human Performance	-	73.33
GPT-3	95.64	83.81

Table 4: Human evaluation performance (percent of correct paraphrases) of paraphrases proposed by people or generated by GPT-3 for the NCI and NCC test sets.

探究实验

- GPT-3是否只是重复已有的释义？Parrot?
- 不同输出结果中的N-gram与C4语料库重合的频率在正确样本和错误样本中的占比情况

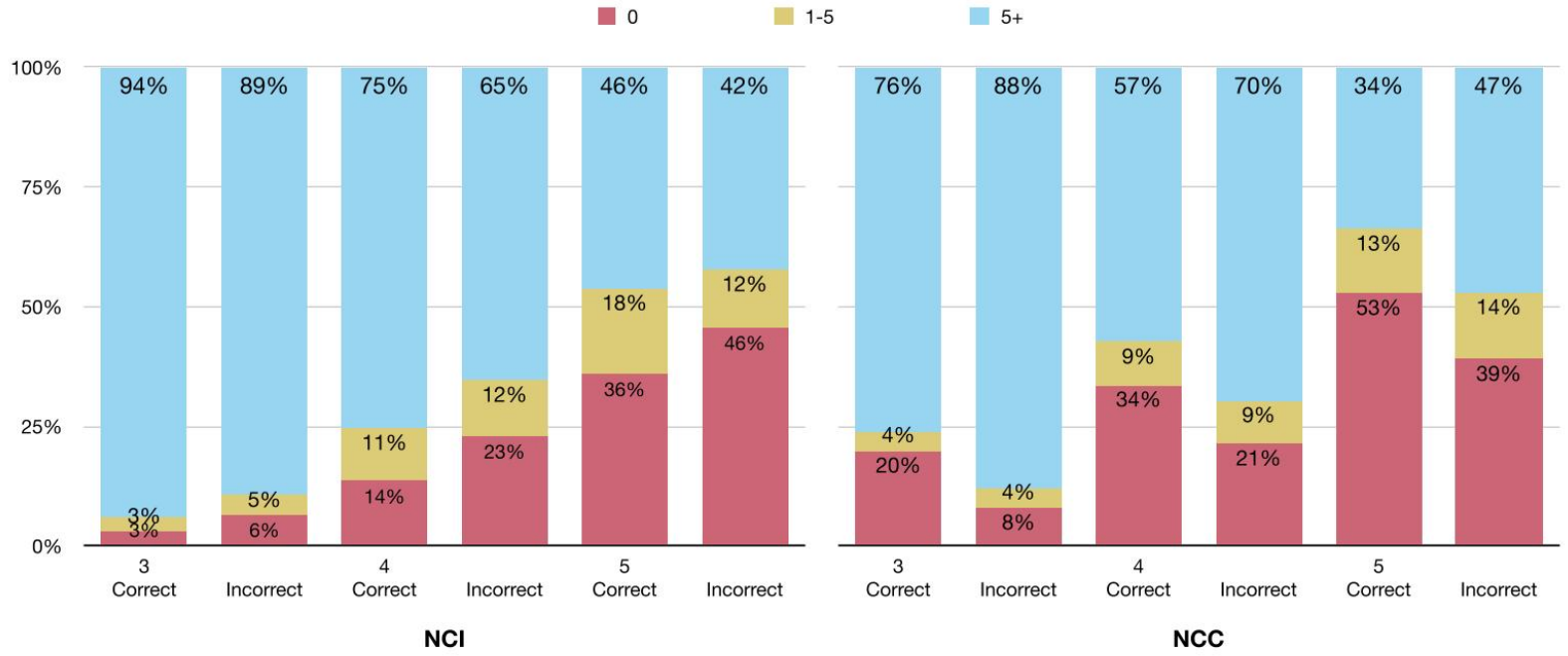


Figure 2: The percent of n-grams among the generated paraphrases (for $n = \{3, 4, 5\}$) that occur in the C4 corpus 0, 1-5, or 5+ times, for each of the NCI and NCC test sets, grouped by correct vs. incorrect generated paraphrases.

- 表现正确的情况绝大多数还是出现在训练集的。
- NCC和NCI在正确不正确中5+的相对大小不一致->NCC更加不需要Parrot

清華大學
Tsinghua University

Q & A

谢谢

Note