



Masked Language Modeling を用いた Knowledge Graph 補完手法の検討

創発ソフトウェア研究室

M1 堀本 隆誠

発表の流れ

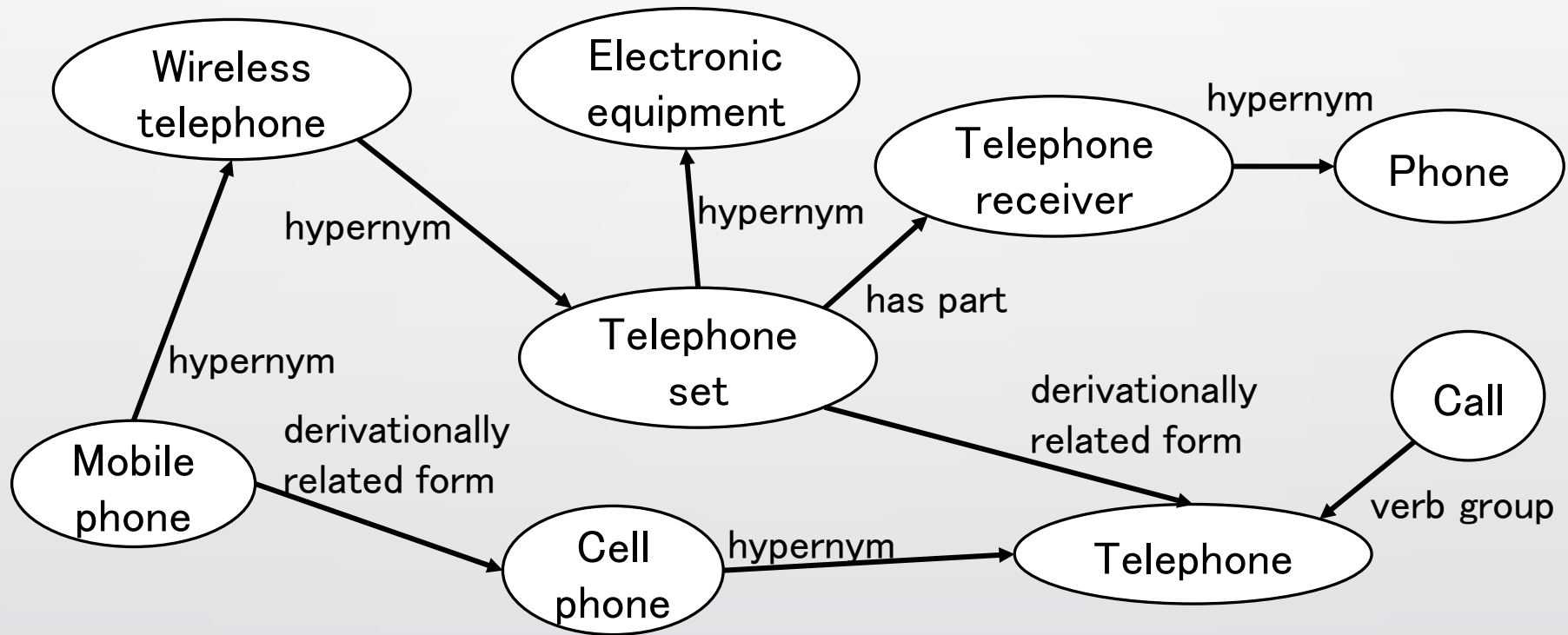
1. はじめに
2. 要素技術
3. 提案手法
4. 実験
5. 実験結果
6. まとめと今後の課題

発表の流れ

1. はじめに
2. 要素技術
3. 提案手法
4. 実験
5. 実験結果
6. まとめと今後の課題

はじめに

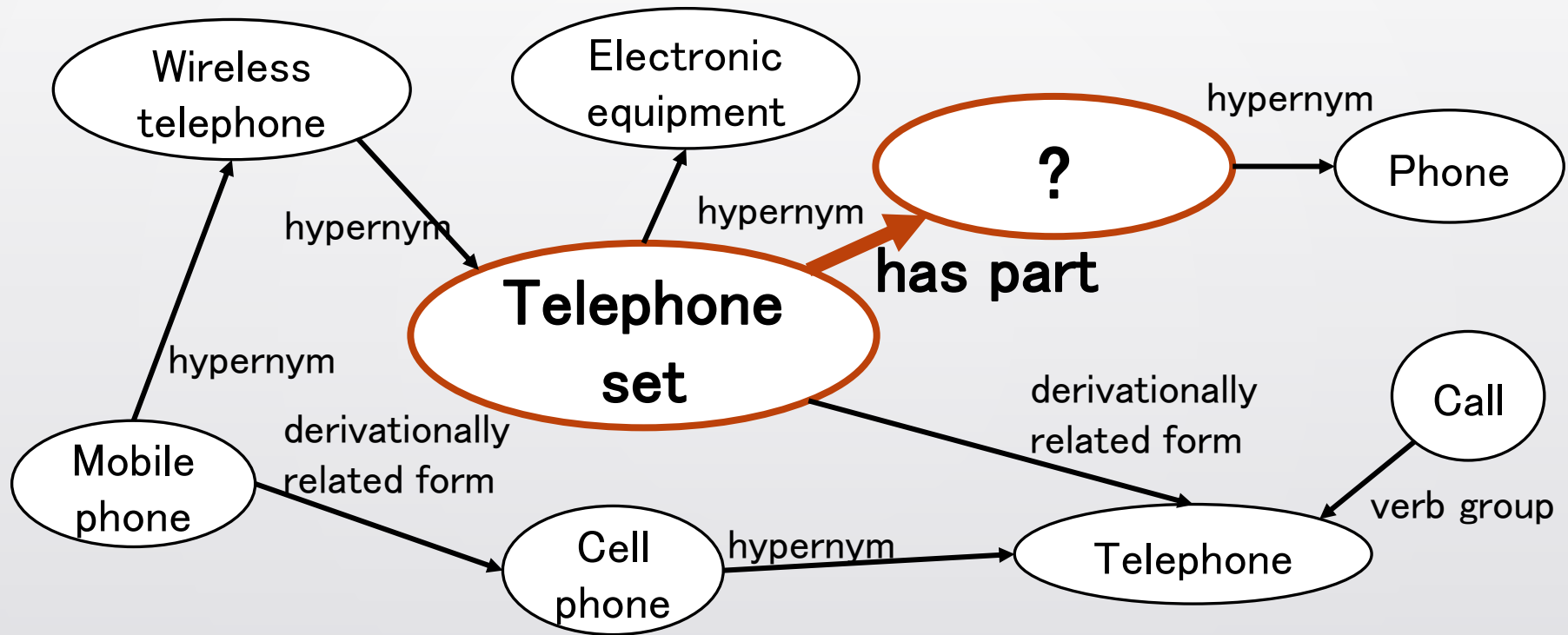
● 人工知能の基盤技術：Knowledge Graph



⇒ 知識間の関係を自動的に補完する手法が必要

はじめに

● 人工知能の基盤技術：Knowledge Graph



⇒ 知識間の関係を自動的に補完する手法が必要

はじめに

● Knowledge Graph 補完手法

代表的な手法 : TransE, ComplEx etc.

(dog, hypernym (上位語), poodle dog)

↓
 v_{dog}

↓ ベクトル化
 v_{hypernym}

↓
 v_{poodle}

↓ 関係式を満たすように更新

$$v'_{\text{dog}} + v'_{\text{hypernym}} \approx v'_{\text{poodle}}$$

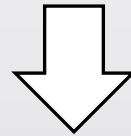
Bordes, Antoine, et al. “Translating embeddings for modeling multi-relational data.”
Advances in neural information processing systems. 2013.

Théo Trouillon, et al. “Complex embeddings for simple link prediction.” Vol. 48 of Proceedings of Machine Learning Research, pp. 2071–2080, New York, New York, USA, 20–22 Jun 2016. PMLR.

はじめに

● 事前学習済み言語モデル BERT

- ◆ 知識（単語）の情報を保持
- ◆ 知識自体の意味情報を効果的に活用
- ◆ Masked Language Modeling でマスクされた単語を予測



BERT の Masked Language Modeling を用いた
tail 予測による Knowledge Graph 補完手法の検討

発表の流れ

1. はじめに

2. 要素技術

3. 提案手法

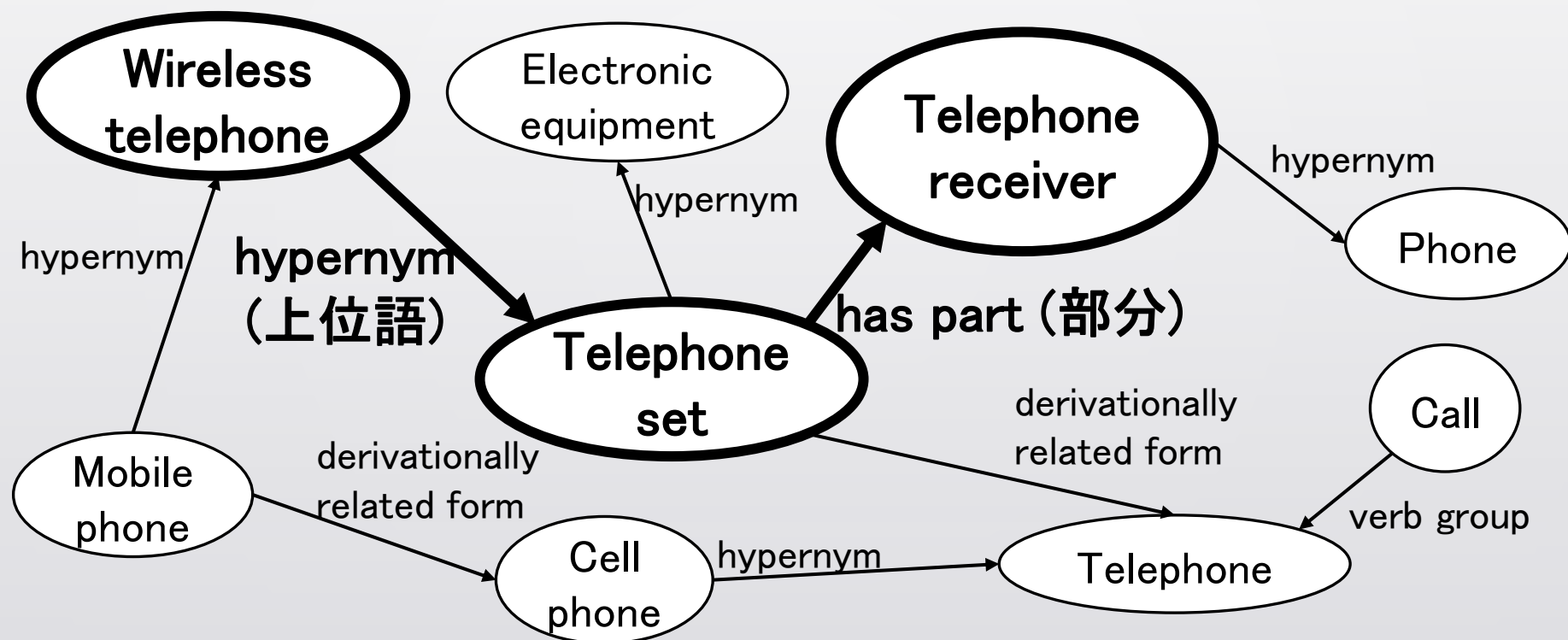
4. 実験

5. 実験結果

6. まとめと今後の課題

➤ Knowledge Graph

知識をグラフ構造として表した知識ネットワーク



要素技術

➤ Knowledge Graph



head	relation	tail
Wireless telephone	hypernym	Telephone set
Telephone set	has part	Telephone receiver
Telephone set	derivationally related form	Telephone
Call	verb group	Telephone
⋮	⋮	⋮

要素技術

➤ Bidirectional Encoder Representations from Transformers (BERT)

◆ Transformer をベースとして双方向エンコーダで構成された言語モデル

□ Next Sentence Prediction (NSP)

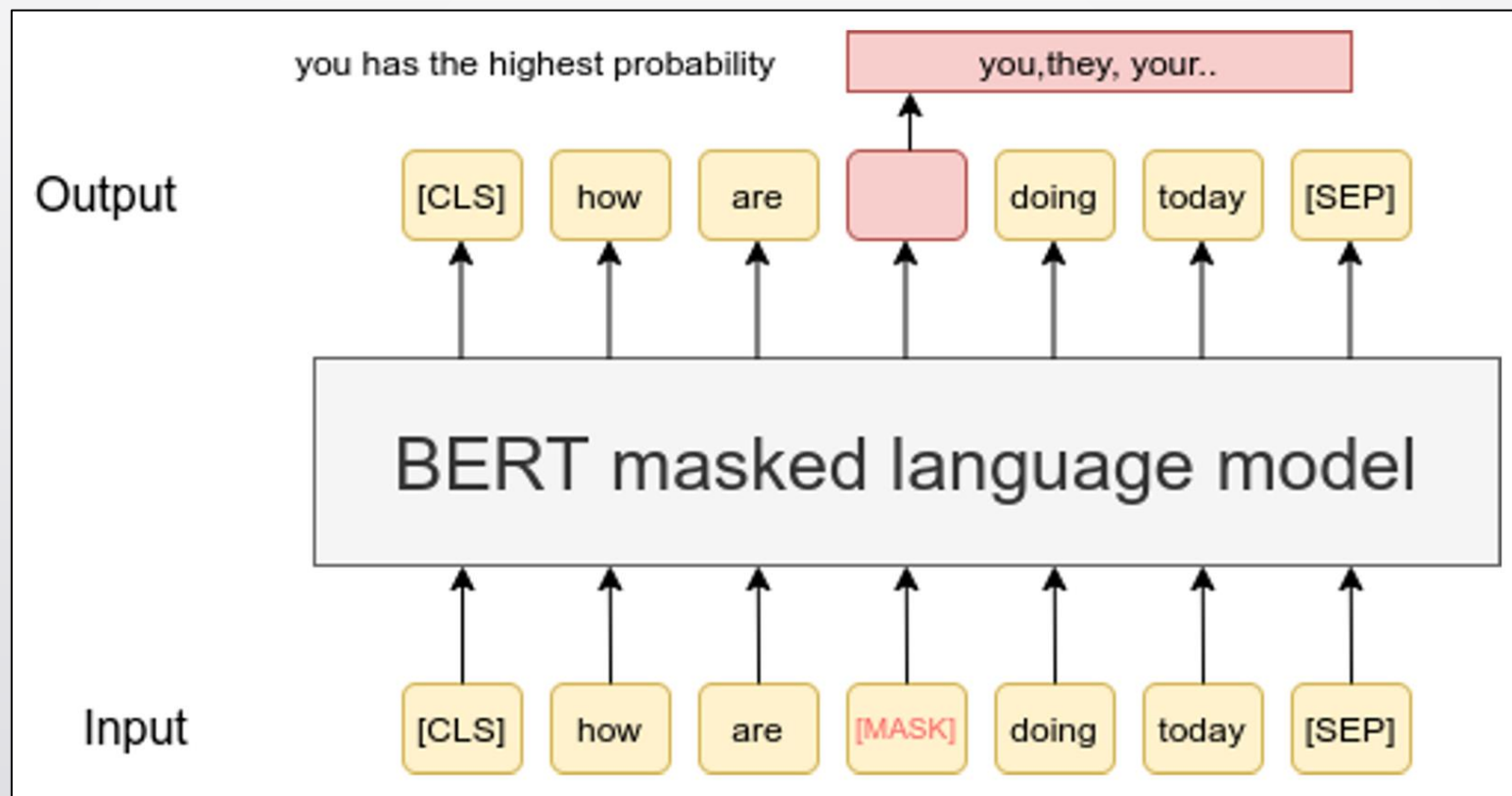
□ Masked Language Modeling (MLM)

◆ マスクされた単語の予測

◆ 転移学習および fine-tuning によりさまざまなタスクに対応

➤ BERT

□ Masked Language Modeling (MLM)



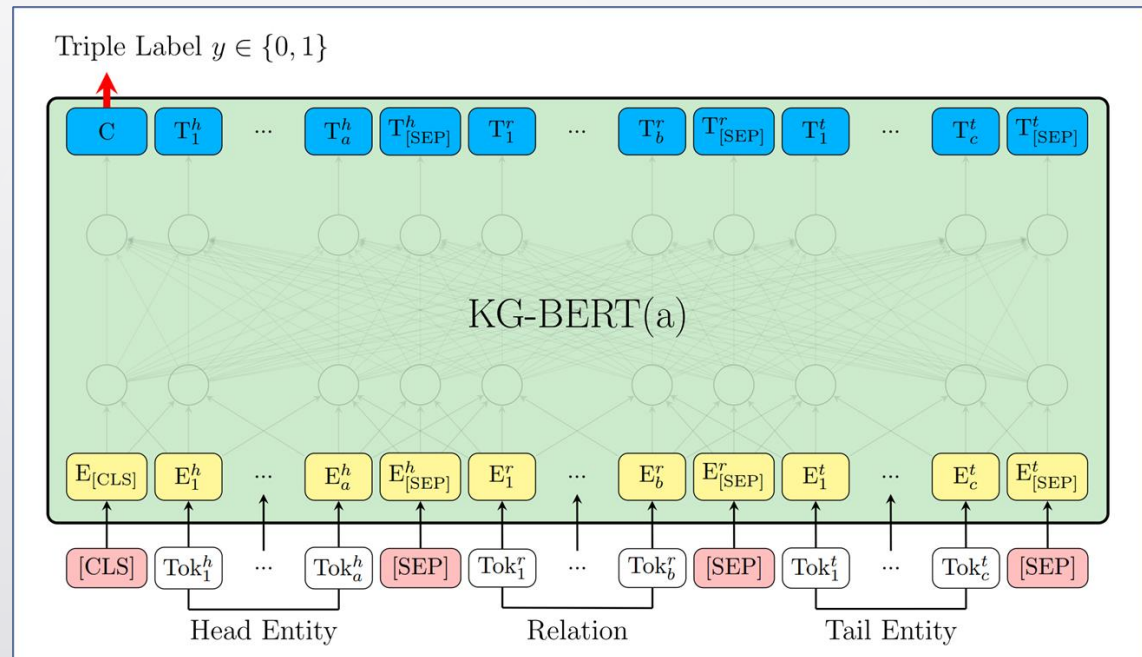
(参考文献 図 1 参照)

要素技術

➤ Knowledge Graph BERT (KG-BERT)

◆ BERT を用いた Knowledge Graph 補完手法の 1 つ

- head, relation, tail を [SEP] トークンで区切った
トークン列を入力
- triple が存在するか否かの 2 値分類



(参考文献 図 1 参照)

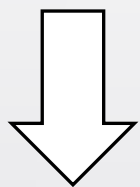
◆ 提案手法では入力方法が同じ, 2 値分類ではなく MLM を使用

要素技術

➤ WN18RR

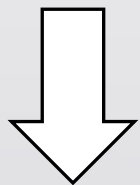
WordNet

… 英語の大規模な語彙データベース
単語間の関係が階層構造として表現



WN18

… 自動抽出された triple データセット

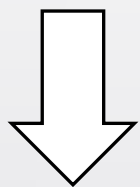


WN18RR

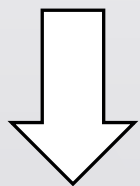
… 逆関係の triple を除去したデータセット

➤ WN18RR

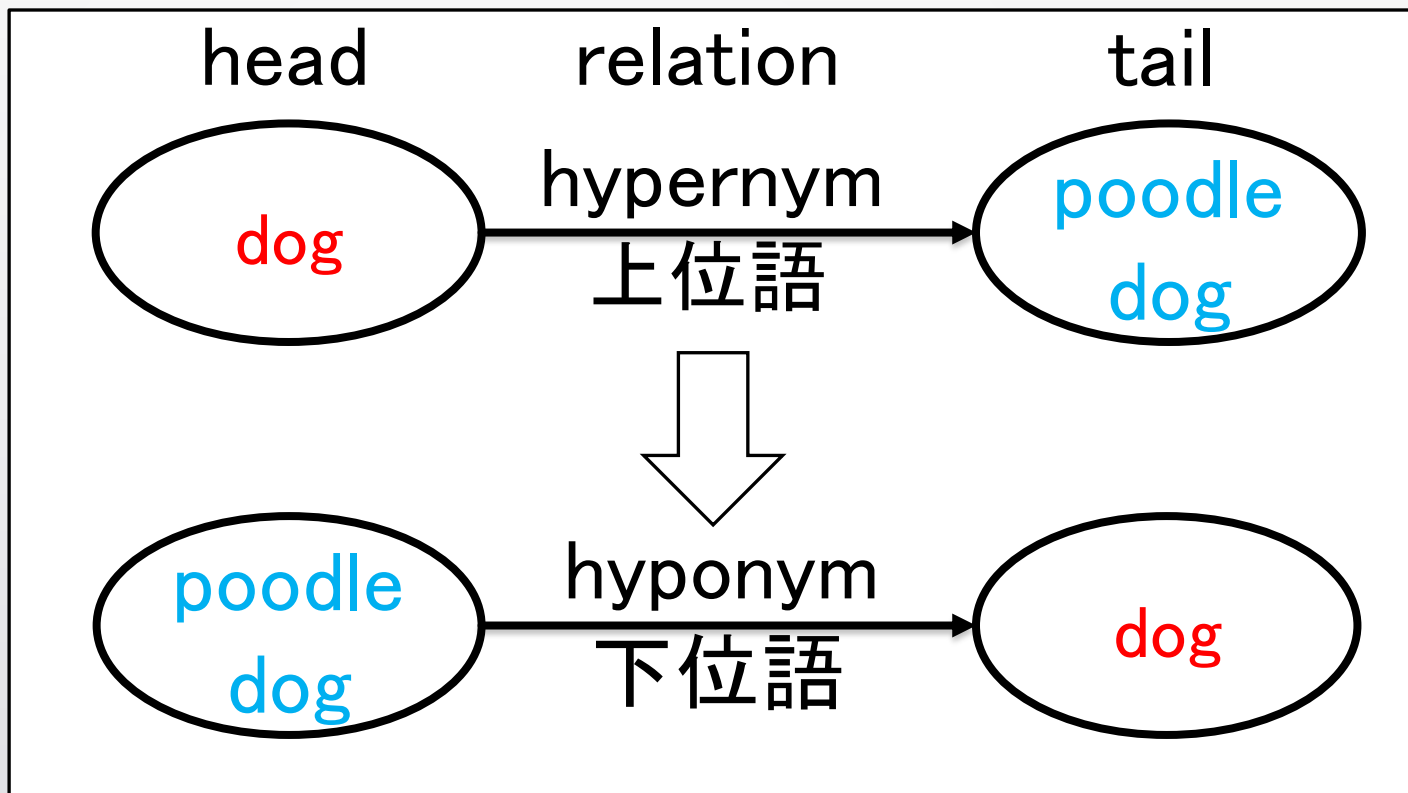
WordNet



WN18

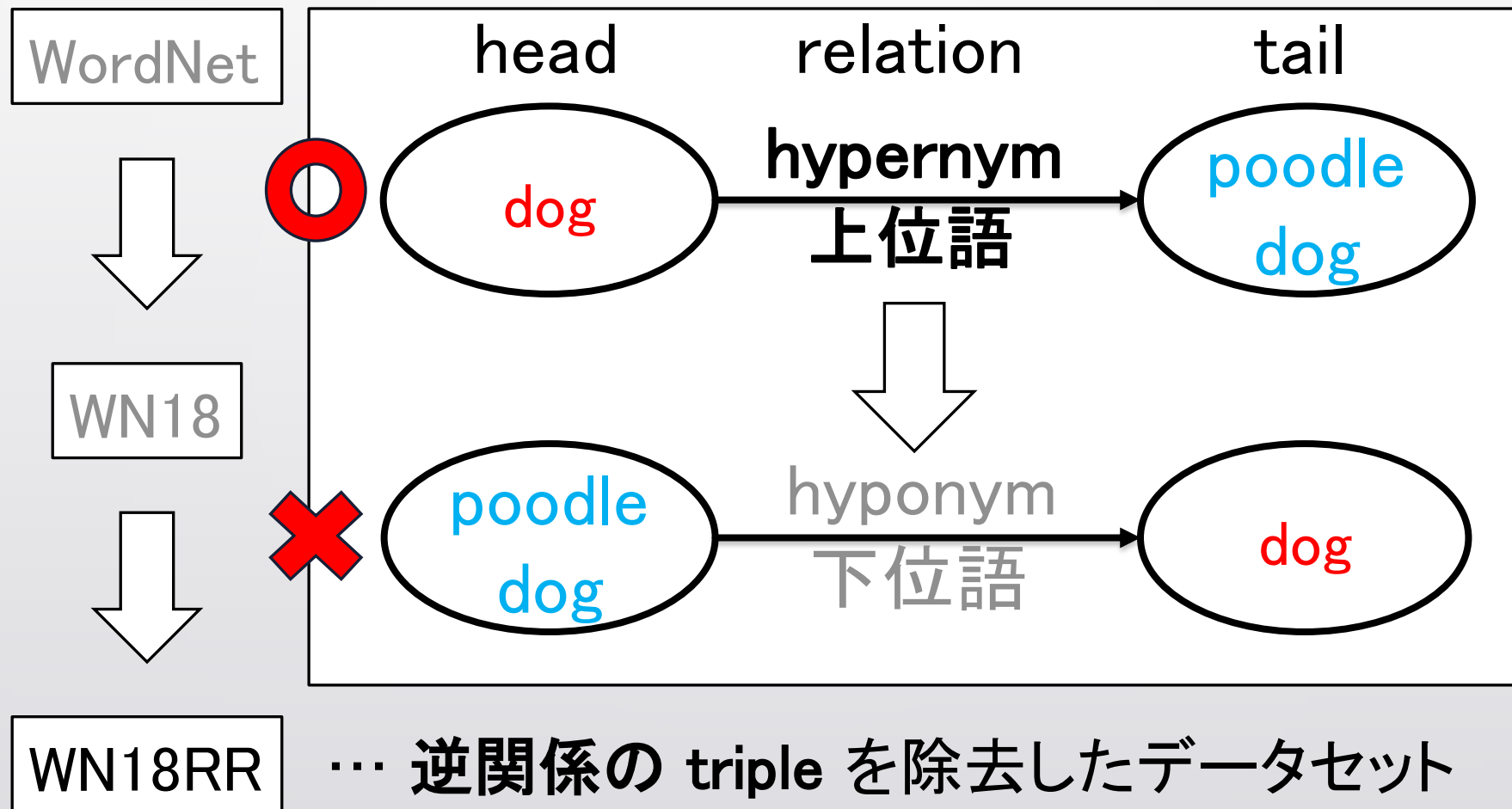


WN18RR



… 逆関係の triple を除去したデータセット

➤ WN18RR



要素技術

➤ WN18RR

entity (head or tail)	relation	triples
40,943	11	93,003

entity (見出し語, その説明文)

poodle dog, an intelligent dog with a heavy curly ~

dog, informal term for a man; “you lucky dog”

position, cause to be in an appropriate place, state, or relation

point, a brief version of the essential meaning of something; ~

drill, a tool with a sharp point and cutting edges for making ~

⋮

要素技術

➤ WN18RR の例 (entity の説明文は省略)

head	relation	tail
position	hypernym	point
drill	has part	chuck
rest	derivationally related form	breath
sophisticated	also see	worldly
brasov	instance hypernym	city
⋮	⋮	⋮

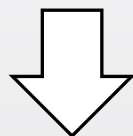
発表の流れ

1. はじめに
2. 要素技術
- 3. 提案手法**
4. 実験
5. 実験結果
6. まとめと今後の課題

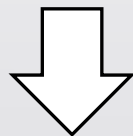
提案手法

◆Knowledge Graph 補完

triple (head, relation, ?)



“?” に入る entity を回答 (tail 予測)



entity 間の関係性を予測

既存手法

➤ KG-BERT

◆ 学習

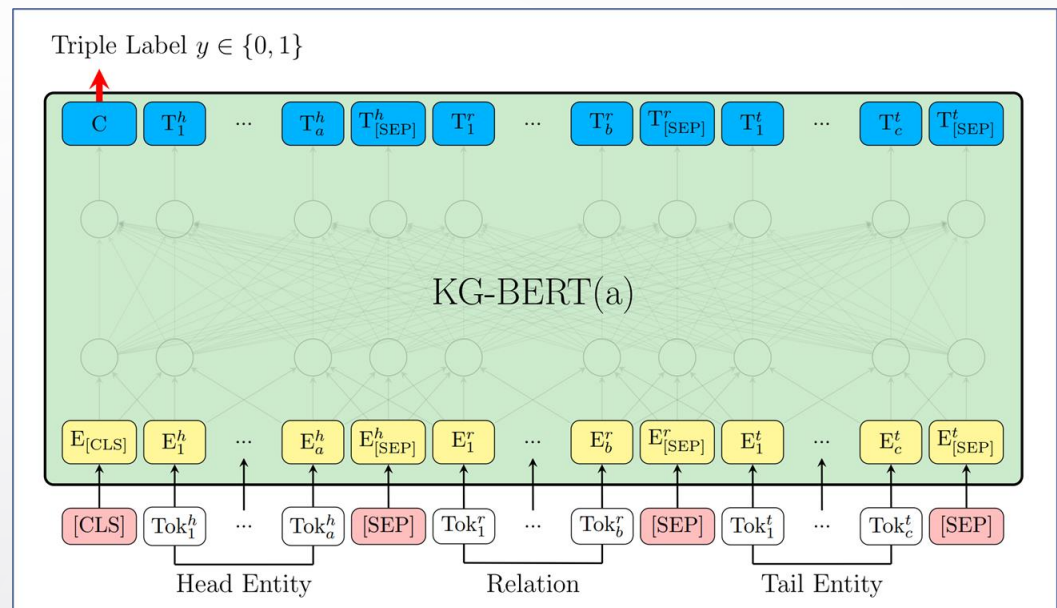
- 1 つの triple に対して

その tail を他の entity と置換した 5 つの triple を生成

- それらが存在しない triple だと判定するように学習

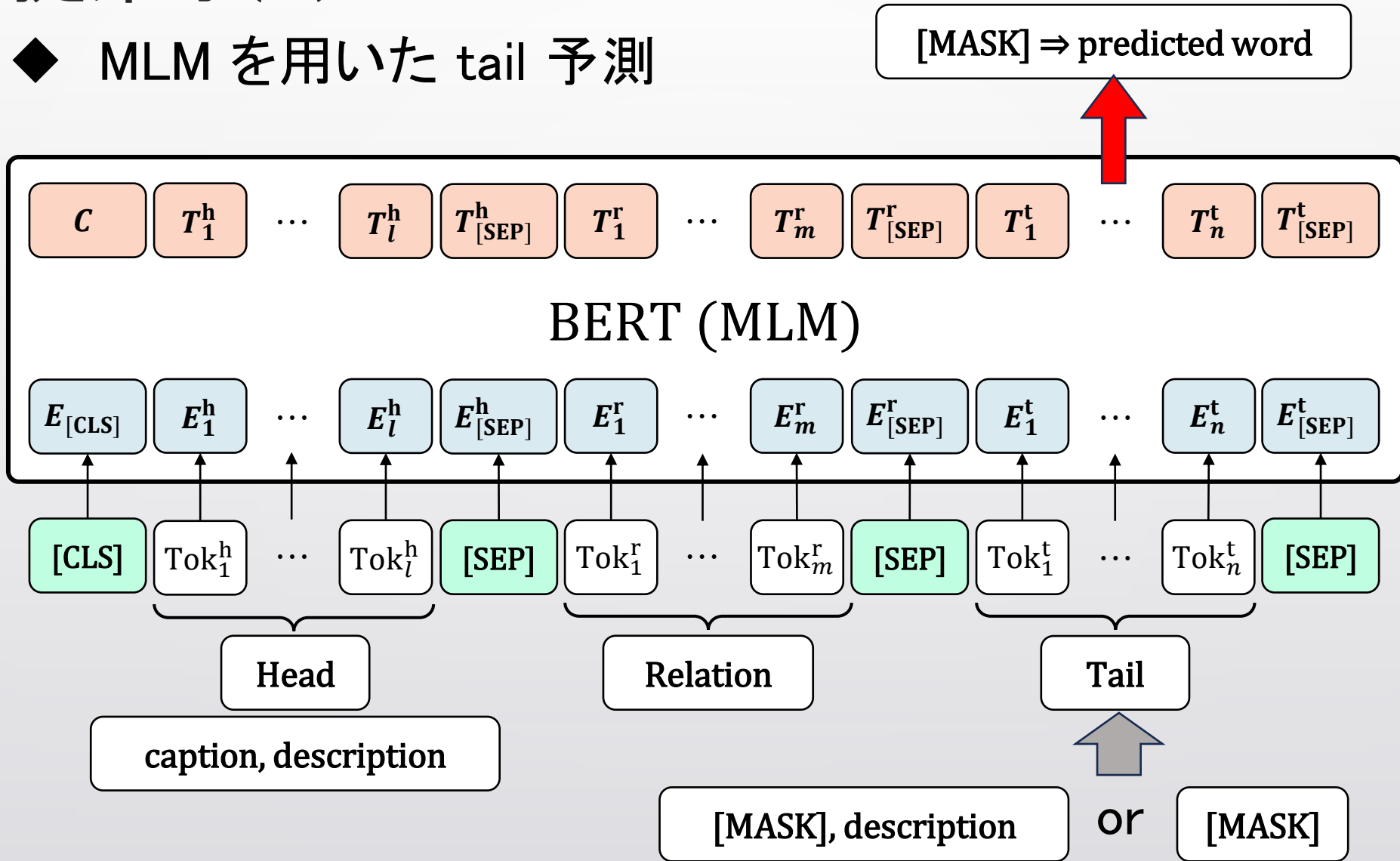
◆ テスト

- 1 つの triple に対してその tail を他の entity と置換した triple を entity の数だけ生成
- それらの triple の分類スコアをランク付けして評価



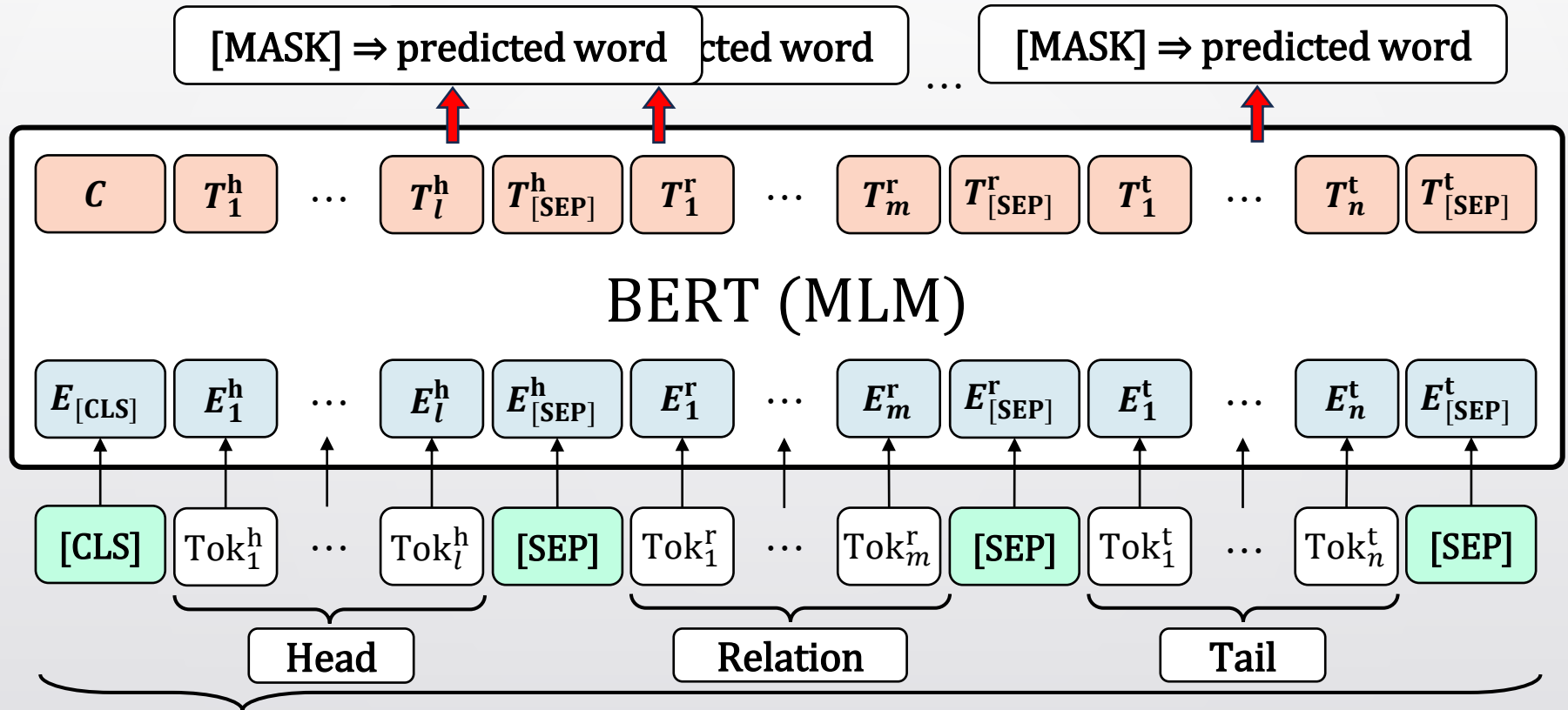
提案手法

◆ MLM を用いた tail 予測



提案手法

◆ BERT の MLM による fine-tuning



15% のうち, 80% : [MASK], 10% : ランダム, 10% : そのまま

└→ 15% 内に必ず tail の見出し語

発表の流れ

1. はじめに
2. 要素技術
3. 提案手法
- 4. 実験**
5. 実験結果
6. まとめと今後の課題

実験

◆ 実験 1

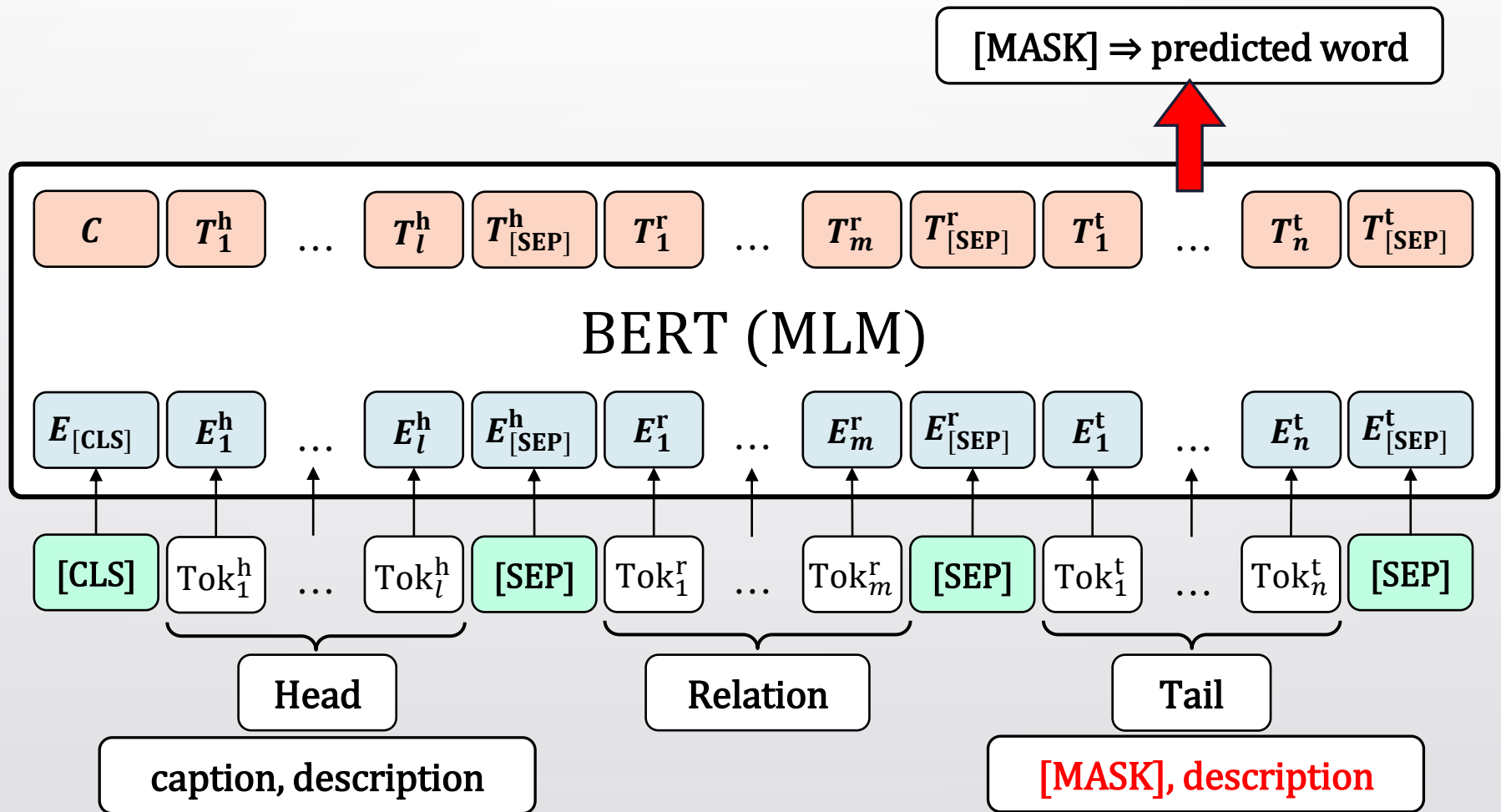
- tail の見出し語を説明文と一緒に入力

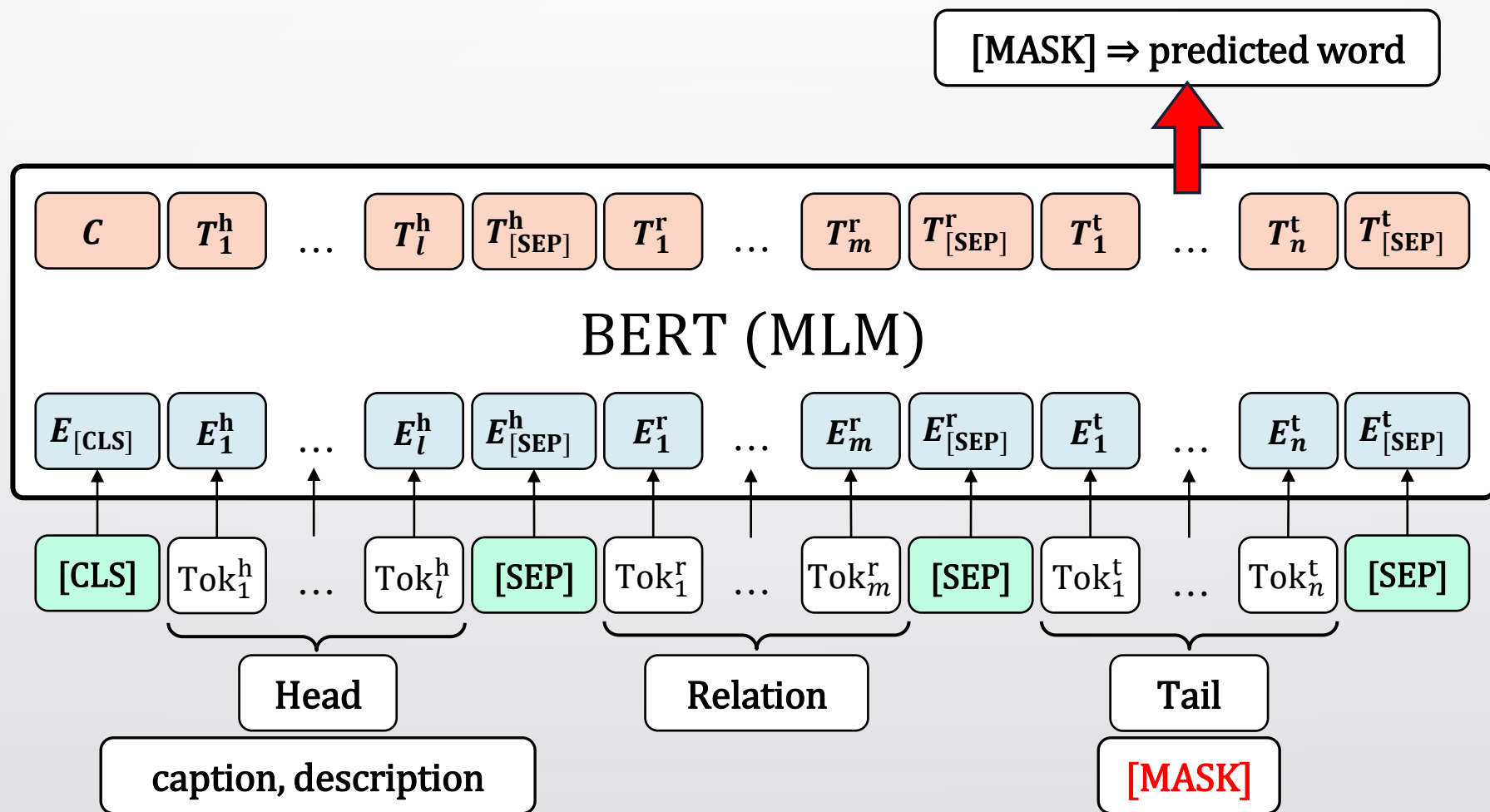
ex) 入力 tail = [MASK], an intelligent dog with a heavy curly ~

◆ 実験 2

- tail の見出し語のみを入力（説明文なし）

ex) 入力 tail = [MASK]



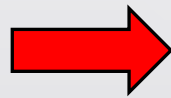


実験（評価指標）

➤ ランク

予測結果のうち、
正解 tail を予測した
順位

正解 tail



ランク = r

スコア順	予測結果
1	predict 1
2	predict 2
3	predict 3
4	predict 4
⋮	⋮
r	predict r (正解)
⋮	⋮
n	predict n

実験（評価指標）

➤ Mean Reciprocal Rank (MRR)

$$\text{MRR} = \frac{1}{|T|} \sum_{i=1}^{|T|} \frac{1}{r_i}$$

$|T|$: triple 数

r_i : triple _{i} における
正解 tail のランク

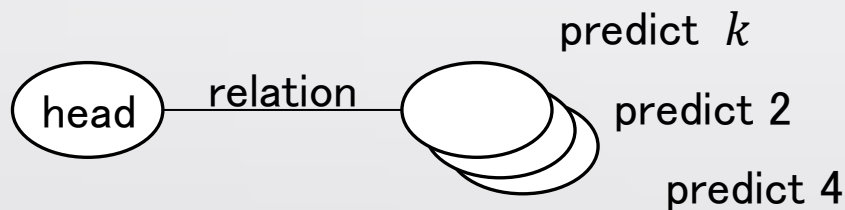
➤ Hits@ k

上位 k 個以内に正解が出力される割合

- $k = 1, 3, 10$ で評価

実験（評価指標）

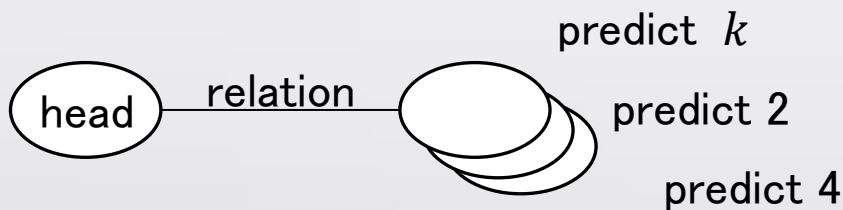
- テスト triple の tail は predict k
- テスト triple と同じ head, relation をもち, predict 2, predict 4 を tail とする triple が存在



スコア順	予測結果
1	predict 1
2	predict 2
3	predict 3
4	predict 4
⋮	⋮
k	predict k (正解)
⋮	⋮
n	predict n

実験（評価指標）

- テスト triple の tail は predict k
- テスト triple と同じ head, relation をもち, predict 2, predict 4 を tail とする triple が存在



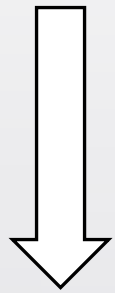
↓ 正確なランクではない

テスト triple 以外の存在する triple を予測結果から除去

スコア順	予測結果
1	predict 1
2	predict 2
3	predict 3
4	predict 4
⋮	⋮
k	predict k (正解)
⋮	⋮
n	predict n

実験（評価指標）

□ 正解した予測結果の
順位を引き上げ



補正したランクを使用

➤ Filtered MRR

➤ Filtered Hits@ k

スコア順	予測結果
1	predict 1
2	predict 2
2	predict 3
4	predict 4
⋮	⋮
$k - 2$	predict k (正解)
⋮	⋮
$n - 2$	predict n

データセット	train	validation	test
WN18RR	86,835	3,034	3,134

パラメータ	KG-BERT (default)	提案手法 (実験 1, 2)
learning rate	5.0×10^{-5}	5.0×10^{-5}
mask probability	–	0.15
batch size	32	32
eval batch size	128 (5000)	–
max seq length	32 (50)	128
epoch	5	20

発表の流れ

1. はじめに
2. 要素技術
3. 提案手法
4. 実験
- 5. 実験結果**
6. まとめと今後の課題

実験結果

34

model		MRR	Hits@1	Hits@3	Hits@10
KG-BERT	文献値	—	—	—	52.4
	再現実験	0.25	12.41	29.44	51.85
提案手法	実験 1	0.546	52.55	56.16	57.79
	実験 1 (Filtered)	0.550	52.81	56.38	57.82
	実験 2	0.168	10.94	19.37	27.95
	実験 2 (Filtered)	0.169	11.04	19.66	28.05

実験結果

35

model		MRR	Hits@1	Hits@3	Hits@10
KG-BERT	文献値	—	—	—	52.4
	再現実験	0.25	12.41	29.44	51.85
提案手法	実験 1	0.546	52.55	56.16	57.79
	実験 1 (Filtered)	0.550	52.81	56.38	57.82
	実験 2	0.168	10.94	19.37	27.95
	実験 2 (Filtered)	0.169	11.04	19.66	28.05

model		MRR	Hits@1	Hits@3	Hits@10
KG-BERT	文献値	—	—	—	52.4
	再現実験	0.25	12.41	29.44	51.85
提案手法	実験 1	0.546	52.55	56.16	57.79
	実験 1 (Filtered)	0.550	52.81	56.38	57.82
	実験 2	0.168	10.94	19.37	27.95
	実験 2 (Filtered)	0.169	11.04	19.66	28.05

実験結果

□ 実験 1

- ◆ 提案手法が KG-BERT の結果を上回ることを確認
 - ⇒ tail 予測に対して MLM は有効
- ◆ Filtered なしとありの変化は微細
 - ⇒ tail の説明文の情報が結果に大きく影響

□ 実験 2

- ◆ 提案手法が Hits@1 において KG-BERT と同等の結果
- ◆ Filtered なしとありはほとんど変化なし
 - ⇒ head, relation のみの入力では tail 予測は困難

考察 (実験 1)

➤ tail 予測

	head	relation	tail
test triple	position	hypernym	?

正解 tail

= “**point**”

	予測結果
KG-BERT	situation
提案手法	point

考察 (実験 1)

➤ tail 予測

	予測結果
KG-BERT	situation
提案手法	point

	test triple (見出し語は太字)
head	position , the particular portion of space occupied ~
relation	hypernym
tail	point , the precise location of something; a spatially limited location; “she walked to a point where she could survey the whole street”

考察 (実験 2)

➤ tail 予測

	head	relation	tail
test triple	evidence	hypernym	?

正解 tail

= “**indication**”

	予測結果
KG-BERT	avermment
提案手法	indication

考察 (実験 2)

➤ tail 予測

	予測結果
KG-BERT	avermment
提案手法	indication

	test triple (見出し語は太字)
head	evidence , an indication that makes something evident; ~
relation	hypernym
tail	indication , something that serves to indicate or suggest; ~

発表の流れ

1. はじめに
2. 要素技術
3. 提案手法
4. 実験
5. 実験結果
- 6. まとめと今後の課題**

まとめと今後の課題

➤ まとめ

- ◆ MLM を用いた Knowledge Graph 補完手法の提案
- ◆ 実験 1 において既存手法 KG-BERT を上回る結果
- ◆ 実験 2 において一部の triple に対して有効
- ◆ 提案手法の tail 予測における有効性の確認

➤ 今後の課題

- ◆ MLM の出力候補を entity に限定
- ◆ head 予測や relation 予測への応用

