

## 報告書

### 1 実験結果

#### 1.1 パラメータ

最大 epoch 数を設定しているが, 検証データを用いて最も結果が良かった model に対してテストデータを使用しているため, model の訓練は最大 epoch 数より少ない epoch 数のものが存在する.

表 1: 実験パラメータ

パラメータ	値
最大 epoch 数	200
埋め込み次元	128
batch size	100
Regularization weight	0.0
学習率	0.001

#### 1.2 データセット

##### 1.2.1 KG Challenge (A Case of Identity)

ナレッジグラフ推論チャレンジにおいて公開されているデータセットの 1 つ. シャーロック・ホームズの小説『A Case of Identity』の内容を用いている.

##### 1.2.2 WN18RR

WordNet のサブセットである WN18 から作成されたリンク予測のデータセット.

##### 1.2.3 FB15k-237

FB15k から作成されたリンク予測のデータセット.

表 2: データセット

データセット	エンティティ数	リレーション数	トリプル数
KG Challenge (A Case of Identity)	4,414	34	4,434
WN18RR	40,943	11	93,003
FB15k-237	14,505	237	310,079

#### 1.3 結果

ホームズデータセットの 1 つ, 『A Case of Identity (花婿失踪事件)』について, 11 の model を用いて実験をする.

### 1.3.1 MRR (Mean Reciprocal Rank)

情報検索システムの評価指標の 1 つ。ある検索結果において最初に見つかった適合文書のランクの逆数をスコアとしており，こうして得たスコアの平均をとったもの。

$$MRR = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{1}{rank_i}$$

$|Q|$  は文書数， $rank_i$  は適合文書のランクを表している。

### 1.3.2 H@k

Hit k のこと。予測したエンティティを順位付けしたときに，上位 k 個以内に正解が含まれている割合を指す。例えば「H@3」において，推定したエンティティのうち上位 3 個以内に正解が含まれている割合を表している。

### 1.3.3 結果 1 (学習率 0.001)

表 3: 結果 (データセット : KG Challenge (A Case of Identity))

model	最適化関数	epoch	MRR	H@1	H@3	H@10
TransE	Adam	200	0.764	0.694	0.802	0.897
CP	Adam	180	0.909	0.851	0.962	0.985
MurE	Adam	185	0.787	0.723	0.828	0.896
RotE	Adam	195	0.793	0.716	0.852	0.931
RefE	Adam	200	0.790	0.713	0.843	0.921
AttE	Adam	200	0.781	0.708	0.835	0.906
RotatE	Adagrad	200	0.266	0.184	0.305	0.421
ComplEx	Adagrad	85	0.275	0.192	0.321	0.407
RotH	Adagrad	200	0.301	0.234	0.332	0.418
RefH	Adagrad	200	0.265	0.201	0.288	0.393
AttH	Adagrad	200	0.254	0.187	0.267	0.384

### 1.3.4 結果 2 (学習率 0.1, 最適化関数 Adagrad)

表 4: 結果 (データセット : KG Challenge (A Case of Identity))

model	Regularization weight	epoch	MRR	H@1	H@3	H@10
RotatE	0.0	105	1.000	1.000	1.000	1.000
ComplEx	0.1	200	0.996	0.995	0.997	0.998
RotH	0.0	95	0.936	0.910	0.955	0.980
RefH						
AttH						

## 2 金田さんの実験コード

金田さんからもらいましたが，まだ実行できていません．今週実行できるように試してみます．

## 参考文献