

文献グラフにおける多項関係の埋め込み

井田 龍希 三輪 誠 佐々木 裕
豊田工業大学

{sd22401, makoto-miwa, yutaka.sasaki}@toyota-ti.ac.jp

概要

知識グラフ埋め込み (Knowledge Graph Embeddings; KGE) では、二つの節点間の関係、すなわち、二項関係を表現する埋め込みを学習している。二項関係は、複数のエンティティと関連するエンティティの検索など、より複雑な関係を推定する上では不十分であり、複数の節点間の関係である多項関係を考慮する必要がある。本研究では、多項関係を考慮する KGE の獲得を目指し、多項関係をモデル化する KGE モデルを提案する。実験では、文献グラフ内の多項関係の予測性能を評価する。

1 はじめに

知識グラフ (Knowledge Graph; KG) では、エンティティを節点、その間の関係を辺とするグラフ構造によって知識を表現する。KG は既存の情報を基にした新たな関係の推定などに利用されており、具体的にはソーシャルプラットフォームにおけるコンテンツのユーザーへの推薦 [1] や文献検索システム [2] に応用されている。新たな関係の推定など、KG の利用に向けて、KG の節点と辺をベクトル空間に埋め込む知識グラフ埋め込み (Knowledge Graph Embeddings; KGE) が頻繁に利用される。現在、多くの KGE モデルが提案されており [3, 4, 5], これらのモデルでは始点節点と終点節点間の関係 (以降、**二項関係**と呼ぶ) を表現する埋め込みを学習する。二項関係の例を図 1 (a) に示す。

一方で、既存モデルが対象とする二項関係は、複数のエンティティと同時に関連するエンティティの検索など、より複雑な関係を推定する上では不十分であり、図 1 (b) のような複数の節点間の関係 (以降、**多項関係**と呼ぶ) を考慮する必要がある。多項関係は、複数の始点節点 $H = \{h_1, \dots, h_n\}$ と一つの終点節点 t の間のそれぞれの関係 $R = \{r_1, \dots, r_n\}$ という複数の二項関係により構成される多対 1 の関係である。このため、多項関係を表現可能な KGE が

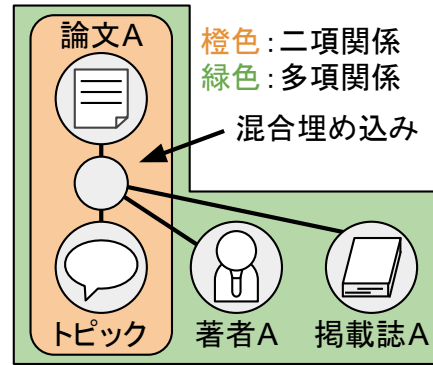


図 1 文献グラフにおける二項関係と多項関係の例

求められる。

そこで、本研究では、二項関係を対象とする既存モデルを拡張し、多項関係を考慮する KGE の獲得を目指し、多項関係をモデル化する KGE モデルを提案する。具体的には、多項関係を学習事例として、**複数の**始点節点 $H = \{h_1, \dots, h_n\}$ とそれぞれの関係 $R = \{r_1, \dots, r_n\}$ に対応する埋め込みを混合して得られる埋め込みと終点節点 t に対応する埋め込み間の距離を近づける学習を行う。この際、 H の部分集合である $H_{sub} \subseteq H$ とそれに対応する $R_{sub} \subseteq R$ と t から成る (H_{sub}, R_{sub}, t) も多項関係であるため¹⁾、多項関係の学習事例数は膨大になる。この問題を軽減するため、学習事例のサンプリングを行う。

評価では、文献グラフ内の多項関係の予測性能を評価する。提案手法の有効性を確認するために、多項関係を構成する各二項関係における既存の KGE モデルの予測を組み合わせ、多項関係を予測するベースラインモデルを作成し、性能を比較する。

本研究の貢献は以下の 2 点である。

- 多項関係をモデル化する新たな KGE モデルを提案した。
- 多項関係の予測において、既存の KGE モデルを用いたベースラインと比較して、提案手法が

1) H_{sub} と R_{sub} のサイズが 1 の時、 (H_{sub}, R_{sub}, t) は一つの始点節点と一つの終点節点間の関係、すなわち二項関係となる。

より高い性能を示すことを示し、提案手法の多項関係の予測における有効性を示した。

- 学習事例に含まれる項数以上の多項関係において、Transformer が有効であることを示した。

2 関連研究

KGE は、推薦や知識グラフ補完などのタスクでの利用に向けて、知識グラフの節点と辺を密な低次元空間へ埋め込む手法である。どのような制約を満たす埋め込みを目的とするかは手法によって異なるが、一般的にグラフ構造を始点節点、終点節点、およびそれらの関係から構成される二項関係の集合とみなして、それらの二項関係を表現できる埋め込みを学習する。TransE [3] では、始点節点の埋め込みと関係の埋め込みの和が終点節点の埋め込みに近づくように、埋め込みを更新する。TransE では、表現できない関係があるため、関係固有の重み行列を利用して、関係固有の空間に節点表現を射影する TransR [4] や始点節点と終点節点間の関係を回転で表現する RotatE [5] などの後続の改善手法が多く提案されている。

また、二項関係を超越る関係をモデル化する KGE も提案されており、二項関係にその成立する期間を付与した時間知識グラフ [6, 7] やより詳細な説明を付与した超関係知識グラフ (Hyper-Relational Knowledge Graph) [8, 9] などを対象とした手法がある。これらの手法では、二項関係に付与された追加情報のモデル化を対象としており、複数の二項関係により構成される多項関係は対象としていない。

3 提案手法

提案手法である多項関係をモデル化する KGE モデルについて説明する。まず、3.1 節では本研究で対象とする文献グラフの構築方法について説明し、次に 3.2 節で多項関係を学習事例として、多項関係をモデル化する KGE モデルの説明をする。

3.1 文献グラフの構築

各論文から書誌情報として、題目・要旨・著者・出版年・掲載誌というタイプの異なる情報を取り出す。また、論文間の引用関係も取り出す。さらに、題目と要旨から自動的に抽出したキーワードとトピックを書誌情報に追加する。キーワード抽出には、文と n-gram の埋め込みの類似度に基づいてキーワードを抽出する KeyBERT [10] を用い、トピック

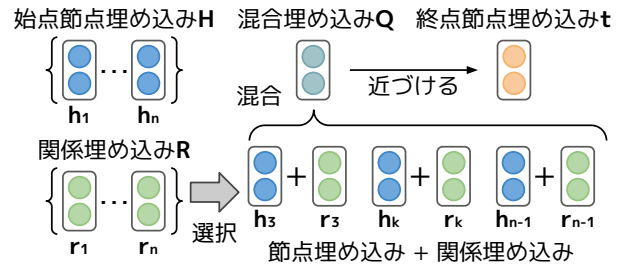


図2 多項関係をモデル化するグラフ埋め込み。

ク抽出には、Sentence-BERT [11] と c-TF-IDF を活用して密なクラスタを作成する BERTopic [12] を用いる。この際、トピックがその他に分類された場合、そのトピックは論文に割り当てない。論文そのものを表現する論文節点は、その論文に関する書誌情報節点に接続される。各論文節点間には、引用関係がある場合に接続される。さらに、各論文の書誌情報節点間も接続する。また、題目、要旨節点では、対応するテキスト情報が利用可能な場合には、テキスト情報を付与する。

3.2 多項関係をモデル化する KGE

多項関係は複数の二項関係によって構成され、始点節点の集合 $H = \{h_1, \dots, h_n\}$ と一つの終点節点 t の間のそれぞれの関係 $R = \{r_1, \dots, r_n\}$ から成る (H, R, t) の組で表現できる。それぞれの節点、関係に対応する埋め込みを $h_n, r_n \in \mathbb{R}^k$ と表現する²⁾。多項関係のモデル化では、 H と R を混合した埋め込み $Q(H, R)$ と t の距離を近づけるように学習する。提案手法の概要を図2に示す。 $Q(H, R)$ は、式(1)に示す通り、 H と R の和を混合して作成される。混合 (Mix) の方法として、単純な平均と Transformer [13] に入力し、その出力を利用する手法の二つを比較する。

$$\begin{aligned} \mathbf{X} &= (\mathbf{h}_1 + \mathbf{r}_1, \mathbf{h}_2 + \mathbf{r}_2, \dots, \mathbf{h}_n + \mathbf{r}_n) \\ \mathbf{Q}(\mathbf{H}, \mathbf{R}) &= \text{Mix}(\mathbf{X}) \end{aligned} \quad (1)$$

具体的に、Transformer の出力を利用する手法では、式(2)のように、Transformer を L 層通して獲得した \mathbf{X}^L の平均ベクトルを線形変換して $\mathbf{Q}(\mathbf{H}, \mathbf{R})$ を作成する。

$$\begin{aligned} \mathbf{X}^L &= \{\mathbf{x}_1^L, \mathbf{x}_2^L, \dots, \mathbf{x}_n^L\} \\ \mathbf{X}_{\text{Avg}} &= \text{Average}(\mathbf{x}_1^L, \mathbf{x}_2^L, \dots, \mathbf{x}_n^L) \\ \mathbf{Q}(\mathbf{H}, \mathbf{R}) &= \text{FC}(\mathbf{X}_{\text{Avg}}) \end{aligned} \quad (2)$$

$\mathbf{Q}(\mathbf{H}, \mathbf{R})$ と t の距離は、 $d(\mathbf{Q}(\mathbf{H}, \mathbf{R}), t) = \|\mathbf{Q}(\mathbf{H}, \mathbf{R}) - t\|$ と計算される。負例サンプリングでは、 n 個の節

2) 本稿では埋め込みを太字で表記する。

点 t' を t と同じ節点タイプの負例からサンプリングする [3]. また, 予測が困難な負例に注目するために, モデルの予測に基づく重み α_i を使用する自己敵対的負例サンプリングを採用して, 損失はマージン Γ を用いて式 (3) のように計算する [5].

$$\begin{aligned} \text{Loss} = & -\log \sigma(\Gamma - d(\mathbf{Q}(\mathbf{H}, \mathbf{R}), \mathbf{t})) \\ & - \sum_{i=1}^n \alpha_i \log \sigma(-\Gamma + d(\mathbf{Q}(\mathbf{H}, \mathbf{R}), \mathbf{t}'_i)) \end{aligned} \quad (3)$$

また, 複数の節点の組み合わせで構成される多項関係の事例数が膨大になるという問題を避けるため, 事例のサンプリングを行い, 計算コストを削減する. 具体的に, (H, R, t) について, 部分始点集合 $H_{sub} \subseteq H$ と一つの終点節点 t の間の関係 $R_{sub} \subseteq R$ から成る (H_{sub}, R_{sub}, t) は, $\sum_{i=1}^{|H|} C_i$ 個と膨大である. そこで, 各エポックにおいて, 各 (H, R, t) に対して, 1 個から $\min(\text{max_num}, |H|)$ 個の節点をランダムに選択して学習事例とする. max_num は, 選択する節点数の最大値を表すハイパーパラメタであり, max_num 個の始点節点と t の間に成り立つ $\text{max_num} + 1$ 項関係までを学習対象にする. この際, 学習の初期には予測しやすい項数の多い多項関係から学習して, 学習が進むにつれて予測が難しい項数が少ない多項関係を学習するようにして, 最終的には二項関係の学習をするというカリキュラム学習を採用する.

4 実験と考察

4.1 実験設定

データセットには ACL Anthology [14] に掲載されている ACL 2023 までに発表された 82,414 件の論文とその書誌情報を利用した. ACL Anthology には引用情報が含まれないため, Semantic Scholar Academic Graph API [15] を用いて取得した. 実験では, 他の節点との接続が少ない (次数が 5 未満) 著者節点は削除した. KeyBERT と BERTopic では, sentence-transformers/all-mpnet-base-v2 [11] を使用した. また, キーフレーズ抽出では, 3-gram までのフレーズを対象とした. 共有されないキーフレーズや一般的すぎるキーフレーズを除去して, 論文間でのキーフレーズの共有性を向上させるため, 各論文からキーフレーズを抽出した後, 2 件から 1,000 件の論文が持つキーフレーズのみを含む辞書を作成し, 文字列照合によって辞書に含まれるキーフレーズを論文に付与した. この処理により, キーフレーズの種類は 216,143 から 37,678 に減少したが, 各論文に

付与されたキーフレーズの数の平均は 5.7 から 11.9 に増加した. 題目節点と要旨節点は SciBERT [16] で初期化し, それに合わせて文献グラフの各節点の表現は 768 次元とした. 本研究では簡単のため過去 5 年の論文とその文献情報を使用して文献グラフを作成した. 文献グラフの統計は付録 A に示す.

開発データと評価データについては, 書誌情報から論文を検索するという設定を想定して, 終点節点が論文節点であり, かつ始点節点と終点節点と共に, 学習データ内に 5 回以上含まれる二項関係を各関係タイプごとに 1 割ずつ選択した. この開発データ, 評価データのそれぞれで, 終点節点と同じである二項関係をマージして, 多項関係のデータに変換した. 二項関係と多項関係の両方で評価を行うために, Word2vec [17] に着想を得た 3/4 次確率分布から対象節点と同じタイプの負例 127 個を負例サンプリングした. なお, 多項関係においては, 全ての始点節点と関連する論文節点が正例であるため, 一部の節点と関連する節点も負例に含まれる. 評価指標には Hit@1 と MRR を用いた.

ベースラインモデルとして, 二項関係をモデル化する KGE の予測を組み合わせた手法を用意した. 具体的には, 多項関係を構成する二項関係ごとに予測を行い, それぞれの **ランキング・スコア・埋め込みの平均** を利用して, 多項関係の予測をする. 二項関係をモデル化する KGE には, TransE と TransE のスコア関数を $\|\text{Transformer}(\mathbf{h}) + \mathbf{r} - \mathbf{t}\|$ に変更したモデルを用意した.

また, カリキュラム学習のスケジューリングについて, 3 種類の方法 (Replace, Add, Add Weighted) を検討した. **Replace** では, 各ステップにおいて $n-1$ 個の始点節点をランダムに選択して, n 項関係のみを対象にする. n 項関係が作れない (H, R, t) に対しては, 可能な最大の項数を対象とする. **Add** では, 前ステップまでの対象である事例に加えて, 項数を一つ減らした事例を対象とする. **Add Weighted** では, Add とほぼ同様であるが, 後で加えた事例がサンプリングされやすいように $1/(n-1)$ の重みづけサンプリングを行う. 全てのスケジューリングで MRR が 5 回更新されなければカリキュラムの次のステップに進むようにした.

4.2 多項関係における予測性能の比較

多項関係における予測性能について表 1 にまとめる. Transformer の層数は, 開発データセットを用い

表 1 多項関係の予測における性能比較

Method	多項関係		
	Transformer	Hit@1	MRR
二項関係 KGE			
ランキング平均		0.2862	0.4262
	✓	0.3227	0.4603
スコア平均		0.3564	0.5026
	✓	0.4141	0.5543
埋め込み平均		0.3290	0.4790
	✓	0.4143	0.5536
多項関係 KGE			
カリキュラム学習なし		0.5770	0.6777
	✓	0.5599	0.6635
Replace		0.5241	0.6324
	✓	0.5579	0.6606
Add		0.5626	0.6660
	✓	0.5618	0.6651
Add Weighted		0.5500	0.6562
	✓	0.5618	0.6644

表 2 二項関係の予測における性能比較

Method	Transformer	二項関係	
		Hit@1	MRR
二項関係 KGE		0.3817	0.5098
	✓	0.4367	0.5530
多項関係 KGE		0.4420	0.5543
	✓	0.4330	0.5463

てチューニングし、最も性能が良いモデルでの結果を報告する。使用したハイパーパラメタやチューニングの結果は付録 B に示す。

提案手法はベースラインと比較して高い性能を達成した。この結果から、多項関係のモデル化が有効であることが分かった。Transformer を使用しない場合に、特に高い性能となった。これは、今回使用した文献グラフは、単純な平均で多項関係を表現できたためだと考えられる。また、カリキュラム学習によって性能は改善せず、Transformer を使用しない場合の **Replace** では性能が大幅に低下した。これは、**Replace** では初期に学習した多項関係の混合方法を忘却してしまうためだと考えられる。ベースラインにおいて、ランキング平均の性能はスコア平均、埋め込み平均の性能よりも低く、スコア平均と埋め込み平均の性能はほぼ同じであった。

次に、提案手法が二項関係の予測にも適用可能かを検証するために行った実験の結果を表 2 に示す。提案手法はベースラインと同等の性能を達成した。この結果から、提案手法は既存の KGE と同等の二項関係に対する表現力を維持しつつ、多項関係に対する表現力を向上させたと言える。

続いて、提案手法の多項関係に対する表現力をさ

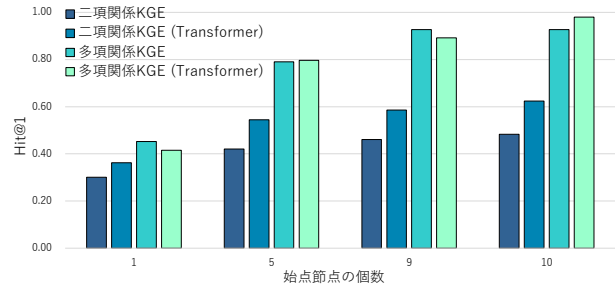


図 3 始点節点の数ごとの Hit@1

表 3 学習事例を超える項数での性能評価 (Hit@1)

Method	Transformer	項数 1~5	項数 6~10
二項関係 KGE		0.3329	0.4661
二項関係 KGE	✓	0.4168	0.6066
多項関係 KGE		0.5854	0.8783
多項関係 KGE	✓	0.5708	0.8828

らに詳細に確かめるために、多項関係のデータにおける項数ごとの評価を行う。始点節点の個数が 1 の時、また学習対象の最大数である 5 の時、さらに学習対象以上の項数で予測できるかを確認するために、9,10 の時の性能比較を図 3 に示す³⁾。この結果から、項数が多いときに、特にベースラインよりも高い性能を示す。また、始点節点の数が 10 個の時には、表 1 の結果と反し、Transformer を使用するモデルの方が性能が高い。さらに、表 3 に示す通り、始点節点が 6~10 個の時の性能は Transformer を使用するモデルのほうが高いことから、Transformer を使用するモデルのほうが項数がより多くなった際に有効であると言える。

5 おわりに

本研究では、多項関係を考慮できる KGE の獲得を目的として、多項関係をモデル化する KGE モデルを提案した。その結果、既存の KGE モデルの予測を用いたベースラインと比べて、高い性能を達成した。実験では 2~6 項関係を学習事例としてモデルを学習した。その範囲では Transformer を用いないモデルが高い性能となり、それ以上の項数では Transformer を用いたモデルが高い性能となった。この結果から、Transformer によって、より多くの項を考慮できる可能性がある。また、カリキュラム学習は、性能の改善に寄与しなかった。今後は、より有効な学習事例をサンプリングする方法を模索する。

3) 始点節点の個数が 1 の時は、二項関係である。多項関係のデータ作成では、終点節点と同じ二項関係をマージする。この際、開発データセット内に一度しか登場しない終点節点は二項関係となる。表 2 の結果は文献グラフの全ての二項関係を対象としているため、評価対象が異なる。

謝辞

本研究は JSPS 科研費 JP20K11962 の助成を受けたものです。

参考文献

- [1] Rex Ying, Ruining He, Kaifeng Chen, Pong Eksombatchai, William L Hamilton, and Jure Leskovec. Graph convolutional neural networks for web-scale recommender systems. In **Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining**, pp. 974–983, 2018.
- [2] Xiao Yu, Quanquan Gu, Mianwei Zhou, and Jiawei Han. Citation prediction in heterogeneous bibliographic networks. In **SDM**, 2012.
- [3] Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko. Translating embeddings for modeling multi-relational data. In C.J. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K.Q. Weinberger, editors, **Advances in Neural Information Processing Systems**, Vol. 26. Curran Associates, Inc., 2013.
- [4] Yankai Lin, Zhiyuan Liu, Maosong Sun, Yang Liu, and Xuan Zhu. Learning entity and relation embeddings for knowledge graph completion. In **Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence**, AAAI'15, p. 2181–2187. AAAI Press, 2015.
- [5] Zhiqing Sun, Zhi-Hong Deng, Jian-Yun Nie, and Jian Tang. Rotate: Knowledge graph embedding by relational rotation in complex space. In **International Conference on Learning Representations**, 2019.
- [6] Julien Leblay and Melisachew Wudage Chekol. Deriving validity time in knowledge graph. In **Companion Proceedings of the The Web Conference 2018**, WWW '18, p. 1771–1776, Republic and Canton of Geneva, CHE, 2018. International World Wide Web Conferences Steering Committee.
- [7] Pengpeng Shao, Dawei Zhang, Guohua Yang, Jianhua Tao, Feihu Che, and Tong Liu. Tucker decomposition-based temporal knowledge graph completion. **Know.-Based Syst.**, Vol. 238, No. C, feb 2022.
- [8] Jianfeng Wen, Jianxin Li, Yongyi Mao, Shini Chen, and Richong Zhang. On the representation and embedding of knowledge bases beyond binary relations. In **Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence**, IJCAI'16, p. 1300–1307. AAAI Press, 2016.
- [9] Bo Xiong, Mojtaba Nayyeri, Shirui Pan, and Steffen Staab. Shrinking embeddings for hyper-relational knowledge graphs. In Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki, editors, **Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)**, pp. 13306–13320, Toronto, Canada, July 2023. Association for Computational Linguistics.
- [10] Maarten Grootendorst. KeyBERT: Minimal keyword extraction with BERT., 2020.
- [11] Nils Reimers and Iryna Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. In **Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing**. Association for Computational Linguistics, 11 2019.
- [12] Maarten Grootendorst. BERTopic: Neural topic modeling with a class-based tf-idf procedure. **arXiv preprint arXiv:2203.05794**, 2022.
- [13] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, L ukasz Kaiser, and Illia Polosukhin. Attention is all you need. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, **Advances in Neural Information Processing Systems**, Vol. 30. Curran Associates, Inc., 2017.
- [14] ACL Anthology. ACL anthology. (Accessed on 2023-07-23).
- [15] Semantic Scholar. Semantic scholar academic graph api. (Accessed on 2023-07-23).
- [16] Iz Beltagy, Kyle Lo, and Arman Cohan. SciBERT: A pre-trained language model for scientific text. In **Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)**, pp. 3615–3620, Hong Kong, China, November 2019. Association for Computational Linguistics.
- [17] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. In C.J. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K.Q. Weinberger, editors, **Advances in Neural Information Processing Systems**, Vol. 26. Curran Associates, Inc., 2013.

A データセットの統計

実験で利用した ACL Anthology に掲載されている過去 5 年分の論文とその書誌情報を使って作成した文献グラフの統計を表 4, 表 5, 表 6 に示す.

表 4 文献グラフの節点の統計

節点タイプ	
Paper	48,995
Author	6,029
Title	30,673
Abstract	28,744
Journal	39
Year	5
Topic	569
Keyphrase	28,220
Total	143,274

表 5 文献グラフの関係の統計. 80:10:10 で分割

関係タイプ	全体	Train	Valid	Test
Cites	385,173	308,139	38,517	38,517
Author	68,761	55,009	6,876	6,876
Title	30,673	30,673	-	-
Abstract	28,744	28,744	-	-
Journal	30,622	24,498	3,062	3,062
Year	30,673	24,539	3,067	3,067
Topic	19,993	15,995	1,999	1,999
Keyphrase	537,339	429,873	53,733	53,733
Total	1,131,978	917,470	107,254	107,254

表 6 学習データにおける書誌情報間の関係の統計

関係タイプ	
Author-Author	73,060
Author-Title	55,009
Author-Abstract	52,742
Author-Journal	10,828
Author-Year	19,681
Author-Topic	22,099
Author-Keyphrase	679,639
Title-Abstract	28,744
Title-Journal	24,498
Title-Year	24,539
Title-Topic	15,995
Title-Keyphrase	429,873
Abstract-Journal	22,984
Abstract-Year	22,984
Abstract-Topic	15,117
Abstract-Keyphrase	423,044
Journal-Year	70
Journal-Topic	1,699
Journal-Keyphrase	62,436
Year-Topic	1,965
Year-Keyphrase	70,405
Topic-Keyphrase	125,787
Keyphrase-Keyphrase	4,816,760

B チューニング結果

実験で使用したハイパーパラメータを表 7 に示す. Transformer を使用する二項関係をモデル化する KGE モデルのチューニング結果を表 8, Transformer を使用しない多項関係をモデル化する KGE モデルのチューニング結果を表 9, Transformer を使用する多項関係をモデル化する KGE モデルのチューニング結果を表 10 に示す.

表 7 実験で使用したハイパーパラメータ

パラメータ	数値
学習対象の最大項数	6
学習時の負例サンプリング数	128
ベクトルの次元数	768
マージン	20
学習率	0.0001
Transformer のヘッド数	8
入力のドロップアウト率	0.2
Transformer のドロップアウト率	0.1

表 8 Transformer を使用する二項関係をモデル化する KGE モデルの開発データにおけるチューニング結果

Transformer 層数	二項関係	
	Hit@1	MRR
1	0.4169	0.5384
2	0.4302	0.5479
3	0.4361	0.5523

表 9 Transformer を使用しない多項関係をモデル化する KGE モデルの開発データにおけるチューニング結果

スケジューリング	多項関係	
	Hit@1	MRR
Random	0.5854	0.6833
Replace	0.5319	0.6380
Add	0.5691	0.6715
Add Weighted	0.5551	0.6607

表 10 Transformer を使用する多項関係をモデル化する KGE モデルの開発データにおけるチューニング結果

スケジューリング	層数	多項関係	
		Hit@1	MRR
Random	1	0.5610	0.6627
	2	0.5631	0.6651
	3	0.5375	0.6417
Replace	1	0.5520	0.6571
	2	0.5515	0.6553
	3	0.5553	0.6587
Add	1	0.5662	0.6693
	2	0.5578	0.6621
	3	0.5564	0.6598
Add Weighted	1	0.5708	0.6711
	2	0.5619	0.6646
	3	0.5601	0.6631