

Block Hole: 関係行列の同時対角化に基づく知識グラフ埋め込みの問題点とその解決

林 克彦^{1,a)} 真鍋 陽俊^{2,b)} 石原 敬大^{2,c)} 新保 仁^{2,d)}

概要: 近年、知識グラフ埋め込みに関する研究が盛んに行われている。知識グラフ埋め込みは、未知の事実に対しても、その真偽値をスコア化することができるため、質問応答や対話などのタスクに応用されている。埋め込みモデルは、双線形、平行移動、ニューラルネットに基づく変換を用いた3種類のモデルに大別されるが、近年、平行移動、及び、ニューラルネットに基づくモデルはいくつかの問題点が指摘されており、現在では双線形モデルが主流となっている。しかし、DistMult、HolE、Complex等の双線形モデルは関係を表す行列を同時対角化して導出するため、関係行列同士が可換となる。本稿では関係行列における可換性の仮定がもたらす実用上の致命的な問題を指摘し、それを解決する新しい知識グラフ埋め込みモデルを提案する。

キーワード: 知識グラフ埋め込み、質問応答、ブロック巡回行列

1. はじめに

YAGO [18] や Freebase [1] などの大規模な知識グラフは質問応答や対話システムの構築に必要不可欠である。知識グラフには、三つ組 (triplet) (s, r, o) の形式で表現された事実 (fact) が格納されており、事実は二つのエンティティ (entity) s, o (それぞれ subject, object とよばれる) の間に関係 (relation) r が成り立つことを表している。例えば、知識グラフには (Leonard Nimoy, starredIn, Star Trek) のような三つ組が登録されており、その存在を知識グラフに問い合わせる操作は「Leonard Nimoy は Star Trek に出演したか?」を問う質問応答と見なせる。知識グラフが完全なものであれば、このような質問応答や複雑な推論もデータベース上での (離散的な) 演算手続きとして実現することができる。しかし、一般に、知識グラフでは多くの事実が欠損しており、不完全である [13]。

このような背景から、近年、知識グラフ埋め込みに関する研究が活発化している。埋め込みモデルでは、各エンティティと関係を埋め込み空間で表現し、その空間上で定

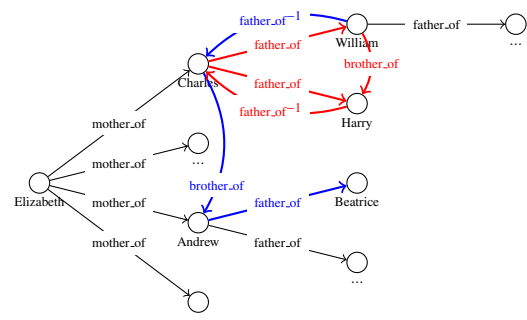


図1 知識グラフの例

義される演算を用いて、与えられた三つ組 (s, r, o) の真偽値を近似するスコアリング関数 $\phi(s, r, o)$ を定義する。スコアリング関数は学習によって汎化されるため、未知の事実に対してもその真偽値を近似的に推定することができる。埋め込みモデルは、双線形 (bilinear)、平行移動 (translation)、ニューラルネットに基づく変換を用いる3種類のモデルに大別され、表1のようにまとめられる。しかし、近年の研究から、平行移動モデルは表現力不足 [8], [20]、ニューラルネットモデルは過剰なパラメータ数や後述する関係パス質問応答タスクに不向きであること [6]、などが指摘されている。そのため、本稿では双線形モデルに議題を絞る。

代表的な双線形モデルに RESCAL [15] がある。RESCAL ではエンティティを n 次元ベクトル、関係を $n \times n$ 行列として n 次元ユークリッド空間に埋め込む。エンティティ s, o のベクトル埋め込みをそれぞれ $\mathbf{e}_s, \mathbf{e}_o$ 、関係 r の埋め込みを行列

¹ 大阪大学 産業科学研究所
The Institute of Scientific and Industrial Research, Osaka University

² 奈良先端科学技術大学院大学
Nara Institute of Science and Technology

a) katsuhiko-h@sanken.osaka-u.ac.jp

b) manabe.hitoshi.me0@is.naist.jp

c) isihara.takahiro.in0@is.naist.jp

d) shimbo@is.naist.jp

モデル	スコアリング関数 $\phi(s, r, o)$	パラメータ数	表現力	非可換性
RESCAL [15]	$\mathbf{e}_s^T \mathbf{W}_r \mathbf{e}_o$	$n \mathcal{E} + n^2 \mathcal{R} $	✓	✓
DistMult [22]	$\langle \mathbf{w}_r, \mathbf{e}_s, \mathbf{e}_o \rangle$	$n \mathcal{E} + n \mathcal{R} $	✗	✗
HolE [14]	$\mathbf{w}_r^T (\mathbf{e}_s \star \mathbf{e}_o)$	$n \mathcal{E} + n \mathcal{R} $	✓	✗
ComplEx [19]	$\text{Re}(\langle \mathbf{w}_r, \mathbf{e}_s, \bar{\mathbf{e}}_o \rangle)$	$2n \mathcal{E} + 2n \mathcal{R} $	✓	✗
Analogy [10]	$\langle \mathbf{w}_r, \mathbf{e}_s, \mathbf{e}_o \rangle + \text{Re}(\langle \mathbf{w}_r', \mathbf{e}_s', \bar{\mathbf{e}}_o' \rangle)$	$(d + 2n) \mathcal{E} + (d + 2n) \mathcal{R} $	✓	✗
Simple [8]	$\langle \mathbf{w}_r, \mathbf{e}_s, \mathbf{e}_o' \rangle$	$2n \mathcal{E} + n \mathcal{R} $	✓	✗
Block HolE	$\sum_{i=1}^b \sum_{j=1}^b \text{Re}(\langle \mathbf{w}_r^{(ij)}, \mathbf{e}_s^{(i)}, \bar{\mathbf{e}}_o^{(j)} \rangle)$	$2bn \mathcal{E} + 2b^2n \mathcal{R} $	✓	✓
TransE [2]	$-\ \mathbf{e}_s + \mathbf{w}_r - \mathbf{e}_o\ _2^2$	$n \mathcal{E} + n \mathcal{R} $	✗	✗
TransH [21]	$-\ (\mathbf{e}_s - \mathbf{w}_r^T \mathbf{e}_s \mathbf{w}_r) + \mathbf{z}_r - (\mathbf{e}_o - \mathbf{w}_r^T \mathbf{e}_o \mathbf{w}_r)\ _2^2$	$n \mathcal{E} + 2n \mathcal{R} $	✗	✓
TransR [9]	$-\ \mathbf{W}_r \mathbf{y}_s + \mathbf{z}_r - \mathbf{W}_r \mathbf{y}_o\ _2^2$	$d \mathcal{E} + (d + 1)n \mathcal{R} $	✗	✓
STransR [12]	$-\ \mathbf{W}_{r,1} \mathbf{y}_s + \mathbf{w}_r - \mathbf{W}_{r,2} \mathbf{y}_o\ _2^2$	$n \mathcal{E} + (2n + 1)n \mathcal{R} $	✗	✓
E-MRP [17]	$\mathbf{w}_r^T \tanh(\mathbf{Z}_r^T \begin{bmatrix} \mathbf{e}_s \\ \mathbf{e}_o \end{bmatrix})$	$d \mathcal{E} + (2d + 1)n \mathcal{R} $	✓	-
ER-MRP [4]	$\mathbf{v}^T \tanh(\mathbf{Z}^T \begin{bmatrix} \mathbf{e}_s \\ \mathbf{e}_o \\ \mathbf{w}_r \end{bmatrix})$	$d \mathcal{E} + d \mathcal{R} + (3d + 1)n$	✓	-
NTN [17]	$\mathbf{v}_r^T \tanh(\mathbf{e}_s^T \mathbf{W}_r^{[1..n]} \mathbf{e}_o + \mathbf{Z}_r^T \begin{bmatrix} \mathbf{e}_s \\ \mathbf{e}_o \end{bmatrix} + \mathbf{b}_r)$	$d \mathcal{E} + (d^2 + 2d + 2)n \mathcal{R} $	✓	-
ConvE [3]	$\text{ReLU}(\text{vec}(\text{ReLU}([M(\mathbf{e}_s); M(\mathbf{w}_r)] *_2 \omega)) \mathbf{W}) \mathbf{e}_o$	$n \mathcal{E} + d \mathcal{R} $	✓	-

表1 知識グラフ埋め込みモデルの比較: Block HolE は $b = 1$ のとき、ComplEx、 $n = 1$ のとき、RESCAL と等価である。 $M(\mathbf{e})$ はベクトル \mathbf{e} の行列化、 $*_2$ は二次元の畳み込み操作とする。

\mathbf{R}_r と書くと、RESCAL のスコア関数は $\phi(s, r, o) = \mathbf{e}_s^T \mathbf{R}_r \mathbf{e}_o$ と定義される。しかし、関係を表す行列 (関係行列) \mathbf{R}_r は n^2 個のパラメータを持つため、計算効率の悪化や過学習の原因となることが知られている [8]。そのため、近年では、関係行列を同時対角化することでパラメータをベクトル化するモデルがいくつか提案されている。その中でも、ComplEx [19] は RESCAL の表現力を落とすことなく、その問題点を解決しており、知識グラフ埋め込みの最も有力なモデルの一つである。

一方で、文献 [6] では、三つ組に対する質問応答を拡張して、複数の関係を連ねた系列に基づく質問応答 (関係パス質問応答) を埋め込みモデルで定式化する方法が提案されている。例えば、「William の父親の兄弟の子供に Beatrice がいるか?」という質問応答は、(William, father_of⁻¹/brother_of/father_of, Beatrice) という三つ組の真偽値を確かめることに一致する。ここで father_of⁻¹/brother_of/father_of は三つの関係を連ねた新しい関係である*1。いま、一般に、 $(s, r_1 / \dots / r_k, o)$ が与えられ、RESCAL を用いてこの三つ組のスコアを計算することを考える。この場合、系列を構成する個々の関係 r_1, \dots, r_k に対応する関係行列 $\mathbf{R}_{r_1}, \dots, \mathbf{R}_{r_k}$ の積をその系列の埋め込みとみなすのが自然であり、したがって与えられた三つ組のスコアは $\mathbf{e}_s^T \mathbf{R}_{r_1} \dots \mathbf{R}_{r_k} \mathbf{e}_o$ としてモデル化できる。また、関係のパラメータをベクトル化 (対角行列化) したモデルでも同様にして、行列積によって関係の系列に対す

*1 いくつかの関係が一つのエンティティを共有するが、系列とならない場合、例えば、 $r_1(a, b) \wedge r_2(a, c)$ のような場合、逆関係を導入することで $r_1^{-1}(b, a) \wedge r_2(a, c)$ として連結できる。

(a)



(b)



図2 (a) 正しい関係パス質問応答、(b) 間違った関係パス質問応答

る計算が行える。ただし、RESCAL とは異なり、ベクトル化された双線形モデルでは、 $\mathbf{R}_r \mathbf{R}_{r'} = \mathbf{R}_{r'} \mathbf{R}_r$ となり、関係を表すパラメータ同士が可換となる。従来研究において、この事実は特に問題視されてきていないが、関係パス質問応答をモデル化する場合、この性質には致命的な欠陥が存在する。例えば、father_of⁻¹/brother_of/father_of を入れ換えて、brother_of/father_of⁻¹/father_of という関係を作った場合、図2の (b) は「William の兄弟の父親の子供に Beatrice がいるか?」という意味になり、(a) とは異なる質問応答となる。しかし、可換な関係パラメータを持つ双線形モデルでは、これら異なる二つの質問応答は全く同じスコアとなり、スコアの大小によってそれらの真偽値を区別することができない。

以上の議論から、本稿では関係のパラメータにブロック巡回行列を仮定した新しい知識グラフ埋め込みモデルを提案する。このモデルは、HolE、または、ComplEx をブロック化によって一般化したモデルと見なせるため、RESCAL の表現力を残しつつ、パラメータ削減を実現することがで

きる。本稿では、このモデルを **Block HolE** と呼ぶ。また、**Block HolE** では関係行列が非可換となるため、上記した **CompLEx** などにおける関係行列の可換性の問題も解決することができる。この有効性を検証するため、**WordNet** と **Freebase** を用いたいくつかの質問応答タスクで実験を行ったので、その結果について報告する。

2. 準備

本稿では、実数の集合を \mathbb{R} 、複素数の集合を \mathbb{C} と書く。また、 $[\mathbf{v}]_j$ はベクトル \mathbf{v} の j 番目の要素、 $[\mathbf{M}]_{jk}$ は行列 \mathbf{M} の (j, k) 要素を表す。ある複素数 z 、複素ベクトル \mathbf{z} 、複素行列 \mathbf{Z} に対して、 \bar{z} 、 $\bar{\mathbf{z}}$ 、 $\bar{\mathbf{Z}}$ はその複素共役とする。ちなみに、実数領域での複素共役は入力をそのまま返す。共役転置は $\mathbf{Z}^* = \bar{\mathbf{Z}}^T$ と書く。また、 $\text{Re}(z)$ 、 $\text{Re}(\mathbf{z})$ 、 $\text{Re}(\mathbf{Z})$ と $\text{Im}(z)$ 、 $\text{Im}(\mathbf{z})$ 、 $\text{Im}(\mathbf{Z})$ はそれぞれ複素数、複素ベクトル、複素行列の実部と虚部を表すと約束する。

n 次元ベクトル $\mathbf{x}, \mathbf{y}, \mathbf{z}$ に対して、 $\mathbf{x} \odot \mathbf{y}$ は \mathbf{x}, \mathbf{y} の要素積を表す。すなわち、 $[\mathbf{x} \odot \mathbf{y}]_i = [\mathbf{x}]_i [\mathbf{y}]_i$ ($i = 1, \dots, n$)。さらに、三重内積 $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle = [\mathbf{x}]_i [\mathbf{y}]_i [\mathbf{z}]_i$ と定義する。また、 n 次元の実数ベクトル $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ に対して^{*2}、 $\mathbf{x} * \mathbf{y}$ 、 $\mathbf{x} \star \mathbf{y}$ は巡回畳み込み、相互相関をそれぞれ表し、 $[\mathbf{x} * \mathbf{y}]_i = \sum_{j=1}^n [\mathbf{x}]_{(i-j+1)} [\mathbf{y}]_j$ 、 $[\mathbf{x} \star \mathbf{y}]_i = \sum_{j=1}^n [\mathbf{x}]_{(j-i+1)} [\mathbf{y}]_j$ とする ($i = 1, \dots, n$)。ただし、上の二つの式において、ベクトル要素の添字が $1, \dots, n$ の範囲にない場合には、 $0 = n, -1 = n-1, \dots, -n+1 = 1$ と解釈するものとする。

n 次元の実数ベクトル $\mathbf{v} \in \mathbb{R}^n$ に対して、

$$\text{circ}(\mathbf{v}) = \begin{bmatrix} [\mathbf{v}]_1 & [\mathbf{v}]_n & \dots & [\mathbf{v}]_3 & [\mathbf{v}]_2 \\ [\mathbf{v}]_2 & [\mathbf{v}]_1 & [\mathbf{v}]_n & & [\mathbf{v}]_3 \\ \vdots & [\mathbf{v}]_2 & [\mathbf{v}]_1 & \ddots & \vdots \\ [\mathbf{v}]_{n-1} & & \ddots & \ddots & [\mathbf{v}]_n \\ [\mathbf{v}]_n & [\mathbf{v}]_{n-1} & \dots & [\mathbf{v}]_2 & [\mathbf{v}]_1 \end{bmatrix} \quad (1)$$

は上のような巡回行列を作る操作として定義する。巡回行列 $\text{circ}(\mathbf{v}) \in \mathbb{R}^{n \times n}$ は n 次離散フーリエ変換行列 $\mathbf{F}_n \in \mathbb{C}^{n \times n}$ を用いて、 $\mathbf{F}_n^{-1} \text{diag}(\mathbf{F}_n \mathbf{v}) \mathbf{F}_n$ として対角化できることが知られている。なお、 $\text{diag}(\mathbf{v})$ はベクトル \mathbf{v} の要素を対角成分に持つ対角行列を表す。また、 $\mathbf{x} * \mathbf{y} = \text{circ}(\mathbf{x}) \mathbf{y}$ 、 $\mathbf{x} \star \mathbf{y} = \text{circ}(\mathbf{x})^T \mathbf{y}$ となる。ここで $\mathbf{x} * \mathbf{y} = \mathbf{F}_n^{-1} (\mathbf{F}_n \mathbf{x} \odot \mathbf{F}_n \mathbf{y})$ 、 $\mathbf{x} \star \mathbf{y} = \mathbf{F}_n^{-1} (\bar{\mathbf{F}}_n \mathbf{x} \odot \mathbf{F}_n \mathbf{y})$ とできることから、これらの計算は高速フーリエ変換を用いることで $O(n \log n)$ で計算可能となる。

3. 双線形変換に基づく知識グラフ埋め込み

知識グラフ \mathcal{G} は $(\mathcal{E}, \mathcal{R}, \mathcal{F})$ として定義される。ここで、 \mathcal{E} はエンティティの集合、 \mathcal{R} は関係の集合であり、集合 $\mathcal{F} \subseteq \mathcal{E} \times \mathcal{R} \times \mathcal{E}$ に含まれる三つ組 (s, r, o) は「事実」と呼ばれる。(ただし、 $s, o \in \mathcal{E}, r \in \mathcal{R}$)。文献 [16] などと言及され

^{*2} 巡回畳み込み、相互相関、及び、巡回行列は複素領域でも定義可能であるが、本稿では不要なため、実数空間でのみ定義する。

ているように、 \mathcal{F} に含まれる既知の三つ組 (すなわち事実) であれば、エンティティの one-hot encoding $\mathbf{x}_s, \mathbf{x}_o \in \{0, 1\}^n$ 、および、関係の真偽値行列 $\mathbf{A}_r \in \{+1, -1\}^{n \times n}$ を用いて、その真偽値を $\mathbf{x}_s^T \mathbf{A}_r \mathbf{x}_o$ で判定することができる。

しかし、知識グラフ補完では、観測された事実集合 \mathcal{F} が不完全である状況を想定している。すなわち、観測できない真の事実集合 $\mathcal{F}_{\text{complete}}$ を、観測された $\mathcal{F} \subset \mathcal{F}_{\text{complete}}$ から予測することが目的である。この目的のためには、未知の三つ組 $(s, r, o) \notin \mathcal{F}$ の真偽値を予測することが可能な、汎化能力を持ったモデルが必要となる。**RESCAL** [15] では、 $\mathbf{x}_s, \mathbf{x}_o, \mathbf{A}_r$ を低次元の線形空間へと埋め込み、その表現を学習することで汎化を達成する。以下では、まず **RESCAL** について説明し、その後 **RESCAL** の改善版とみなせる **DistMult**、**HolE**、**CompLEx** について説明を行う。これらモデルはいずれも、関係 r に関する真偽値行列 \mathbf{A}_r を、スコアリング関数 $\phi(s, r, o)$ を用いて、

$$[\mathbf{A}_r]_{so} \approx \text{sign}(\phi(s, r, o))$$

と近似するモデルになっている。ここで、 $\text{sign}(a)$ は $a > 0$ の場合に 1、そうでない場合に -1 を返す符号関数である。

3.1 RESCAL モデル

文献 [15] において提案された **RESCAL** はもっとも一般製の高い双線形モデルである。**RESCAL** のスコアリング関数は、

$$\phi(s, r, o) = \mathbf{e}_s^T \mathbf{R}_r \mathbf{e}_o.$$

と書ける。ここで $\mathbf{e}_s, \mathbf{e}_o \in \mathbb{R}^n$ はエンティティ s, o を埋め込んだ n 次元の実数ベクトル、 $\mathbf{R}_r \in \mathbb{R}^{n \times n}$ は関係 r を埋め込んだ $n \times n$ 次元の実数行列である。**RESCAL** のパラメータ数は $n|\mathcal{E}| + n^2|\mathcal{R}|$ となり、関係行列が持つ n^2 個のパラメータのために計算効率が悪い。また、その過剰なパラメータのため、過学習の問題を引き起こすことも知られている [8]。

3.2 DistMult モデル

RESCAL における関係行列 \mathbf{R}_r のパラメータ数を削減するため、文献 [22] では **DistMult** と呼ばれるモデルが提案された。このモデルは対称行列に対する同時対角化の観点から導出することができる。行列 $\mathbf{R}_r \in \mathbb{R}^{n \times n}$ が対称行列の場合 ($\mathbf{R}_r = \mathbf{R}_r^T$)、

$$\mathbf{R}_r = \mathbf{O}_r \text{diag}(\mathbf{w}_r) \mathbf{O}_r^T \quad (2)$$

のように対角化できることが知られている。ここで $\mathbf{O}_r \in \mathbb{R}^{d \times d}$ は直交行列であり ($\mathbf{O}_r^T \mathbf{O}_r = \mathbf{O}_r \mathbf{O}_r^T = \mathbf{I}$)、 $\mathbf{w}_r \in \mathbb{R}^d$ は \mathbf{R}_r の固有値を並べたベクトルである。また、 $\forall r, r' \in \mathcal{R}$ に対して、 $\mathbf{R}_r \mathbf{R}_{r'} = \mathbf{R}_{r'} \mathbf{R}_r$ (可換) であるとき、 \mathcal{R} における全ての関係行列 \mathbf{R}_r はその固有ベクトルを共有することが知

られている。すなわち、式 (2) の \mathbf{O}_r を共通の直交行列 \mathbf{O} に置き換えることができる。これは同時対角化と呼ばれている。

したがって、RESCAL における全ての関係行列が対称、かつ、互いに可換であると仮定すると、以下のようなモデルを導出できる。

$$\begin{aligned}\phi(s, r, o) &= \mathbf{e}_s^T \mathbf{R}_r \mathbf{e}_o = \mathbf{e}_s^T \mathbf{O} \text{diag}(\mathbf{w}_r) \mathbf{O}^T \mathbf{e}_s \\ &= \mathbf{e}'_s{}^T \text{diag}(\mathbf{w}_r) \mathbf{e}'_o = \langle \mathbf{w}_r, \mathbf{e}'_s, \mathbf{e}'_o \rangle.\end{aligned}\quad (3)$$

ここで $\mathbf{e}' = \mathbf{e}\mathbf{O}$ とする。このモデルは DistMult と呼ばれ、パラメータ数は $n|\mathcal{E}| + n|\mathcal{R}|$ となり、線形時間で計算できる。しかし、対称なスコアリング関数となるため ($\phi(s, r, o) = \phi(o, r, s)$)、非対称な関係をモデル化するには不向きである。例えば、father_of のような関係は非可換であり、一般に、知識グラフにはこのような関係が多数存在する。

3.3 HolE モデル

文献 [14] では、DistMult モデルが対称関係しか扱えない弱点を克服するため、相互相関を用いたモデルを提案している。このモデルは HolE と呼ばれ、以下のように定式化される。

$$\phi(s, r, o) = \mathbf{w}_r^T (\mathbf{e}_s \star \mathbf{e}_o).\quad (4)$$

ここで $\mathbf{w}_r, \mathbf{e}_s, \mathbf{e}_o \in \mathbb{R}^n$ は関係、及び、エンティティを埋め込んだ n 次元の実数ベクトルである。HolE のパラメータ数は DistMult と同様に $n|\mathcal{E}| + n|\mathcal{R}|$ となるが、相互相関の計算には高速フーリエ変換を用いた場合でも $O(n \log n)$ の計算量が必要となる。相互相関の性質から、一般に、 $\phi(s, r, o) \neq \phi(o, r, s)$ となる。

3.4 ComplEx モデル

文献 [19] では、HolE モデルとは異なる視点から、より直接的に DistMult モデルの弱点を克服する新たなモデルを提案している。任意の実正方形行列 $\mathbf{R}_r \in \mathbb{R}^{n \times n}$ に対して、 $\mathbf{R}_r = \text{Re}(\mathbf{W}_r)$ となる正規行列 $\mathbf{W}_r \in \mathbb{C}^{n \times n}$ が必ず存在することが知られている [19]。ここで正規行列とは $\mathbf{A}^* \mathbf{A} = \mathbf{A} \mathbf{A}^*$ を満たす行列である。また、正規行列は

$$\mathbf{W}_r = \mathbf{U}_r \text{diag}(\mathbf{w}_r) \mathbf{U}_r^* \quad (5)$$

として対角化できる。ここで、 $\mathbf{U}_r \in \mathbb{C}^{n \times n}$ はユニタリ行列 ($\mathbf{U}_r^* \mathbf{U}_r = \mathbf{U}_r \mathbf{U}_r^* = \mathbf{I}$)、 $\mathbf{w}_r \in \mathbb{C}^n$ は \mathbf{W}_r の固有値を並べたベクトルである。対称行列における同時対角化の議論と同様に、 $\forall r, r' \in \mathcal{R}$ に対して、 $\mathbf{W}_r \mathbf{W}_{r'} = \mathbf{W}_{r'} \mathbf{W}_r$ (可換) となるとき、 \mathbf{U}_r は共有のユニタリ行列 \mathbf{U} とできる。上記の議論から、RESCAL における全ての関係行列が可換であると仮定すると、

$$\begin{aligned}\phi(s, r, o) &= \mathbf{e}_s^T \mathbf{R}_r \mathbf{e}_o = \mathbf{e}_s^T \text{Re}(\mathbf{U} \text{diag}(\mathbf{w}_r) \mathbf{U}^*) \mathbf{e}_o \\ &= \text{Re}(\mathbf{e}'_s{}^T \text{diag}(\mathbf{w}_r) \overline{\mathbf{e}'_o}) = \text{Re}(\langle \mathbf{w}_r, \mathbf{e}'_s, \overline{\mathbf{e}'_o} \rangle).\end{aligned}\quad (6)$$

となる。ここで $\mathbf{e}' = \mathbf{e}\mathbf{U}$ とした。これが ComplEx と呼ばれるモデルである。ComplEx のパラメータ数は $2n|\mathcal{E}| + 2n|\mathcal{R}|$ となり、線形時間で計算できるため、HolE よりも処理効率が良い。また、DistMult とは異なり、ComplEx では一般に $\phi(s, r, o) \neq \phi(o, r, s)$ となるため、非対称な関係もモデル化できる。

3.5 HolE と ComplEx の等価性

本稿の著者である林と新保は文献 [7] において、HolE と ComplEx が等価なモデルであることを示した。単純には、以下のように HolE を書き換えることができるため、HolE は ComplEx でもモデル化可能だと言える。

$$\begin{aligned}\phi(s, r, o) &= \mathbf{w}_r^T (\mathbf{e}_s \star \mathbf{e}_o) = \mathbf{w}_r^T \mathbf{F}_n^{-1} (\overline{\mathbf{F}_n \mathbf{e}_s} \odot \mathbf{F}_n \mathbf{e}_o) \\ &= \frac{1}{n} (\overline{\mathbf{F}_n \mathbf{w}_r})^T (\overline{\mathbf{F}_n \mathbf{e}_s} \odot \mathbf{F}_n \mathbf{e}_o) \\ &= \frac{1}{n} \text{Re}(\langle \mathbf{w}'_r, \mathbf{e}'_s, \overline{\mathbf{e}'_o} \rangle).\end{aligned}\quad (7)$$

ここで $\mathbf{w}'_r = \overline{\mathbf{F}_n \mathbf{w}_r}$ 、 $\mathbf{e}'_s = \overline{\mathbf{F}_n \mathbf{e}_s}$ 、 $\mathbf{e}'_o = \mathbf{F}_n \mathbf{e}_o$ である。 $\frac{1}{n}$ は定数のため、無視できる。逆に ComplEx を HolE でモデル化する方法については、文献 [7] を参考にしてもらいたい。

4. 従来法の問題点

節 1 で言及したように、関係パス質問応答とは、 $(s, r_1 / \dots / r_k, o)$ の真偽値を確かめる操作である。双線形モデルでは、 $\mathbf{e}_s^T \mathbf{R}_r \approx \mathbf{e}_o$ となる性質から、 $\phi(s, r_1 / \dots / r_k, o) = \mathbf{e}_s^T \mathbf{R}_{r_1} \dots \mathbf{R}_{r_k} \mathbf{e}_o$ として関係パス質問応答をモデル化できる [6]。DistMult や ComplEx でも同様にして、 $\mathbf{e}_s^T \text{diag}(\mathbf{w}_{r_1}) \dots \text{diag}(\mathbf{w}_{r_k}) \mathbf{e}_o$ 、 $\text{Re}(\mathbf{e}_s^T \text{diag}(\mathbf{w}_{r_1}) \dots \text{diag}(\mathbf{w}_{r_k}) \overline{\mathbf{e}_o})$ のようにモデル化することができる。しかし、一般に対角行列は可換であるため、例えば、 $(s, r_k / r_{k-1} / \dots / r_1, o)$ のような事実を作っても、そのスコアは $(s, r_1 / \dots / r_k, o)$ に対するスコアと全く同じ値になる。すなわち、 $\phi(s, r_1 / \dots / r_k, o) = \phi(s, r_k / r_{k-1} / \dots / r_1, o)$ になってしまうため、これらの真偽値をスコアで区別することができない。

また、ニューラルネットに基づくモデルは、双線形モデルにおける $\mathbf{e}_s^T \mathbf{R}_r \approx \mathbf{e}_o$ や平行移動モデルにおける $\mathbf{e}_s + \mathbf{w}_r \approx \mathbf{e}_o$ のような性質を持たないため、関係パス質問応答をそのままの形式でモデル化することは難しい [6]。さらに、TransE のような平行移動モデルでは、 $-\|\mathbf{e}_s + \mathbf{w}_{r_1} + \dots + \mathbf{w}_{r_k} - \mathbf{e}_o\|_2^2$ としてモデル化するが、これも可換なスコアリングとなっている。また、TransE を始めとした平行移動モデルの多くはそもそも表現力が足りないことが証明されている [8]。双線形モデルでは、Analogy [10] や SimpleE [8] といった新しいモデルも提案されているが、これらも関係のパラメー

タがベクトル (対角行列) となるため、可換なスコアリングとなる。SimpleE については、subject と object のエンティティを別々のベクトル空間で表現するため、関係パス質問応答をどのようにモデル化するか自体が自明ではない。

5. Block Hole: ブロック巡回行列によるモデル化

本稿では RESCAL モデルにおける関係行列にブロック巡回行列を仮定することで従来法の問題点を解決する。ブロック巡回行列は

$$\begin{bmatrix} \mathbf{W}^{(11)} & \dots & \mathbf{W}^{(1b)} \\ \vdots & \ddots & \vdots \\ \mathbf{W}^{(b1)} & \dots & \mathbf{W}^{(bb)} \end{bmatrix} \quad (8)$$

のような行列として定義される。ここで $\mathbf{W}^{(ij)}$ は n 次元の実数ベクトル $\mathbf{w}^{(ij)}$ から作られる巡回行列 $\text{circ}(\mathbf{w}^{(ij)})$ とする。ブロック巡回行列 $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{bn \times bn}$ ($\mathbf{A} \neq \mathbf{B}$) は一般に $\mathbf{AB} \neq \mathbf{BA}$ となるので非可換である。

RESCAL モデルにおける関係 r に対する埋め込み行列 \mathbf{R}_r をブロック巡回行列に置き換えると、

$$\phi(s, r, o) = [\mathbf{e}_s^{(1)\top} \dots \mathbf{e}_s^{(b)\top}] \begin{bmatrix} \mathbf{W}_r^{(11)} & \dots & \mathbf{W}_r^{(1b)} \\ \vdots & \ddots & \vdots \\ \mathbf{W}_r^{(b1)} & \dots & \mathbf{W}_r^{(bb)} \end{bmatrix} \begin{bmatrix} \mathbf{e}_o^{(1)} \\ \vdots \\ \mathbf{e}_o^{(b)} \end{bmatrix} \quad (9)$$

と書ける。ここで $\mathbf{e}_s^{(i)}, \mathbf{e}_o^{(i)} \in \mathbb{R}^n$, $\mathbf{W}_r^{(ij)} \in \mathbb{R}^{n \times n}$ であり (ただし、 $i = 1, \dots, b, j = 1, \dots, b$), $\mathbf{W}_r^{(ij)} = \text{circ}(\mathbf{w}_r^{(ij)})$ とする。上式は、巡回行列の性質を用いて、

$$\begin{aligned} \phi(s, r, o) &= [\mathbf{e}_s^{(1)\top} \dots \mathbf{e}_s^{(b)\top}] \begin{bmatrix} \sum_{j=1}^b \mathbf{w}_r^{(1j)} * \mathbf{e}_o^{(j)} \\ \vdots \\ \sum_{j=1}^b \mathbf{w}_r^{(bj)} * \mathbf{e}_o^{(j)} \end{bmatrix} \\ &= \sum_{i=1}^b \mathbf{e}_s^{(i)\top} \left(\sum_{j=1}^b \mathbf{w}_r^{(ij)} * \mathbf{e}_o^{(j)} \right) \\ &= \sum_{i=1}^b \sum_{j=1}^b \mathbf{w}_r^{(ij)\top} (\mathbf{e}_o^{(j)} \star \mathbf{e}_s^{(i)}) \end{aligned} \quad (10)$$

と書き換えることができる。本稿ではこのモデルを **Block HoleE** と呼ぶ。なぜなら、 $b = 1$ のとき、Block HoleE は HoleE と本質的に等価であり、それをブロック化によって一般化したモデルと見なせるからである。また、 $n = 1$ のとき、次元数 b の RESCAL モデルと等価である。Block HoleE のパラメータ数は $bn|\mathcal{E}| + b^2n|\mathcal{R}|$ 、計算量は $O(b^2n \log n)$ となる。

5.1 複素計算を用いた高速化

節 3.5 のように相互相関の計算を全て複素空間で行うことにより、高速フーリエ変換を排除し、Block HoleE のスコア計算を高速化する [7], [23]。まず、式 (10) を以下のように書き換える。

$$\begin{aligned} \phi(s, r, o) &= \sum_{i=1}^b \sum_{j=1}^b \mathbf{w}_r^{(ij)\top} (\mathbf{e}_o^{(j)} \star \mathbf{e}_s^{(i)}) \\ &= \sum_{i=1}^b \sum_{j=1}^b \mathbf{w}_r^{(ij)\top} \mathbf{F}_n^{-1} (\mathbf{F}_n \mathbf{e}_o^{(j)} \odot \mathbf{F}_n \mathbf{e}_s^{(i)}) \\ &= \frac{1}{n} \sum_{i=1}^b \sum_{j=1}^b (\mathbf{F}_n \mathbf{w}_r^{(ij)})^\top (\mathbf{F}_n \mathbf{e}_o^{(j)} \odot \mathbf{F}_n \mathbf{e}_s^{(i)}) \\ &= \frac{1}{n} \sum_{i=1}^b \sum_{j=1}^b \text{Re}(\langle \mathbf{F}_n \mathbf{w}_r^{(ij)}, \mathbf{F}_n \mathbf{e}_s^{(i)} \overline{\mathbf{F}_n \mathbf{e}_o^{(j)}} \rangle). \end{aligned} \quad (11)$$

ここで $\mathbf{w}_r^{(ij)'} = \frac{1}{n} \mathbf{F}_n \mathbf{w}_r^{(ij)}$, $\mathbf{e}_s^{(i)'} = \overline{\mathbf{F}_n \mathbf{e}_s^{(i)}}$, $\mathbf{e}_o^{(j)'} = \overline{\mathbf{F}_n \mathbf{e}_o^{(j)}}$ と置くと、

$$\phi(s, r, o) = \sum_{i=1}^b \sum_{j=1}^b \text{Re}(\langle \mathbf{w}_r^{(ij)'}, \mathbf{e}_s^{(i)'} \overline{\mathbf{e}_o^{(j)'}} \rangle) \quad (12)$$

が得られる。このモデルのパラメータ数は $2bn|\mathcal{E}| + 2b^2n|\mathcal{R}|$ となり、計算量は $O(b^2n)$ となる。このモデルでは通常 $b^2 \ll n$ として設定するため、実用上、 b は定数項として扱える。

5.2 関係パス質問応答のモデル化

ここでは Block HoleE で関係パス質問応答をモデル化する方法について述べる。式 (12) を行列表記に戻すと、

$$\text{Re} \left([\mathbf{e}_s^{(1)\top} \dots \mathbf{e}_s^{(b)\top}] \begin{bmatrix} \text{diag}(\mathbf{w}_r^{(11)}) & \dots & \text{diag}(\mathbf{w}_r^{(1b)}) \\ \vdots & \ddots & \vdots \\ \text{diag}(\mathbf{w}_r^{(b1)}) & \dots & \text{diag}(\mathbf{w}_r^{(bb)}) \end{bmatrix} \begin{bmatrix} \mathbf{e}_o^{(1)} \\ \vdots \\ \mathbf{e}_o^{(b)} \end{bmatrix} \right) \quad (13)$$

と書ける。ここで式 (13) における関係行列を \mathbf{W}'_r と置くと、従来手法と同様にして、 $\phi(s, r_1 / \dots / r_k, o) = \text{Re}(\mathbf{e}_s^\top \mathbf{W}'_{r_1} \dots \mathbf{W}'_{r_k} \overline{\mathbf{e}_o})$ として関係パス質問応答をモデル化することができる。 $b \geq 2$ の時、一般に、 $\mathbf{W}'_{r_i} \mathbf{W}'_{r_j} \neq \mathbf{W}'_{r_j} \mathbf{W}'_{r_i}$ となるので、Block HoleE では非可換性を持ったモデル化を実現することができる。

6. 実験

6.1 分類データセット

文献 [6] では、WN11、及び、FB13 データの知識グラフ上でランダムウォークを行って、 $(s, r_1 / \dots / r_k, o)$ をサンプリングしている。WN11 と FB13 は文献 [17] において WordNet [5] と Freebase から作られたデータセットである。表 2 にデータサイズを示す。Base は元々の WN11 と FB13 を表し、Path が文献 [6] によって構築された関係パス質問応答タスクの評価セットである。Base の test データは分類設定となっており、実際は正解を崩して作成した負例も含むため、表記の倍の三つ組が存在する。Path の設定では $(s, r_1 / \dots / r_k, ?)$ のように Object エンティティを隠して、そのランキング精度を評価するタスクとなっている。

		WN11	FB13
	$ \mathcal{R} $	11	13
	$ \mathcal{E} $	38,696	75,043
Base	#train	112,581	316,232
	#valid	2,609	5,908
	#test	10,544	23,733
Path	#train	2,129,539	6,266,058
	#valid	11,277	27,163
	#test	46,577	109,557
CPath	#train	1,203,554	5,276,240
	#valid	5,315	19,048
	#test	22,533	77,565
CPathR	#train	2,019,178	5,949,826
	#valid	8,671	21,255
	#test	36,115	85,824

表2 分類タスク用のデータセット

本稿の目的は可換性が関係パスのモデル化にもたらす問題を調べることにある。そのため、モデルの Precision を重要視して、本稿では Path の設定に負例を混ぜた分類タスクを考える。具体的には、Path の train、dev、test データに含まれる長さ $(k) \geq 2$ 以上の $(s, r_1/\dots/r_k, o)$ から $(s, r_k/r_{k-1}/\dots/r_1, o)$ を作り、これが知識グラフ上に存在しない場合、 $(s, r_1/\dots/r_k, o)$ を正例、 $(s, r_k/r_{k-1}/\dots/r_1, o)$ を負例としてデータに登録する。この設定を CPath とする。また、負例をランダムに生成したデータセットを CPathR とする。CPath、CPathR において、表2では負例を除いた数を表記しているが、train、dev、test 共にその倍の数のサンプルが存在する。

6.2 分類タスクの実験設定

全てのモデル学習は L2 正則化付きロジスティック回帰

$$\sum_{((s,r,o),y) \in \mathcal{D}} \log\{1 + \exp(-y\phi(s,r,o))\} + \lambda \|\Theta\|_2^2$$

で行う。ここで $\mathcal{D} \subseteq (\mathcal{E} \times \mathcal{R} \times \mathcal{E}) \times \{-1, +1\}$ 、 Θ は関係、及び、エンティティのパラメータとする。まず、Base の train データに (s, r, o) がある場合、 (o, r^{-1}, s) という逆関係を追加し、通常の三つ組学習を行う。この埋め込みを用いて、Base 設定の分類タスクを評価する。正負の判定は $\text{sign}(\phi(s, r, o))$ で行う。

また、三つ組学習後の埋め込みを初期値として、CPath、CPathR の学習をそれぞれ行う。ただし、CPath 設定では、非可換性を持たないモデル (DistMult、HolE、ComplEx、Analogy、Simple) は正例か負例のどちらかしか当てられないため、正解率は常に 50% である。そのため、CPath 設定では RESCAL と Block HolE の学習のみ行った。

また、各モデルのハイパーパラメータは L2 正則化項の係数 $\lambda \in \{0.01, 0.001, 0.0001, 0\}$ 、学習率 $\eta \in \{0.15, 0.1, 0.05\}$ 、

	WN11			FB13		
	Base	CPath	CPathR	Base	CPath	CPathR
DistMult	70.2	50.0	60.1	68.9	50.0	58.5
ComplEx	80.7	50.0	64.3	77.1	50.0	62.5
RESCAL	79.8	66.1	67.1	75.8	62.5	63.8
Block HolE	80.2	72.1	68.5	76.8	66.9	65.2

表3 分類精度の比較

	$ \mathcal{E} $	$ \mathcal{R} $	#train	#valid	#test
WN18	40,943	18	141,442	5,000	5,000
FB15k	14,951	1,345	483,142	50,000	59,071

表4 WN18 と FB15k のデータセットの統計量

次元数 $n \in \{5, 10, 20, 50, 100, 200\}$ の範囲でグリッドサーチして決定した。Block HolE のブロック数 b は 2 で固定した。

6.3 実験結果: 分類タスク

表3に分類精度を示す。Base 設定では RESCAL、ComplEx、Block HolE に大きな精度差は見られなかった。一方で、CPath 設定では Block HolE が従来法を大きく上回る精度を達成した。この結果は本稿で指摘した非可換性のモデル化が有効に機能していることを明らかに示唆している。また、ランダムに生成した関係パス質問応答タスク (CPathR) に対しても、Block HolE がもっとも高い精度となっており、より一般的な関係パスのモデル化においてもその有効性が確認された。

6.4 ランキングデータセット

Block HolE の性能を最も標準的な実験設定で評価するため、WN18 と FB15k によるランキングタスクの実験を行う。各データセットの統計量を表4に示す。

6.5 ランキングタスクの実験設定

評価は先行研究にならい、各 test サンプル (s, r, o) に対して、object o (または subject s) を全てのエンティティ e で置き換えて、 (s, r, e) (または (e, r, o)) のスコアを計算する。そして、そのスコアに基づいて降順にソートし、平均逆順位 (MRR: mean reciprocal rank) とトップ N に正解が含まれる割合 (“Hits at N”) を求めて、“raw” または “filtered” の設定で評価する。“filtered” では train、dev、test データ上の $e \neq o$ (または $e \neq s$) となる (s, r, e) (または (e, r, o)) を全て削除して評価し、“raw” では削除を行わず評価する。

学習は L2 正則化付きロジスティック回帰で行った。各モデルのハイパーパラメータは L2 正則化項の係数 $\lambda \in \{0.01, 0.001, 0.0001, 0\}$ 、学習率 $\eta \in \{0.15, 0.1, 0.05, 0.025\}$ 、次元数 $n \in \{5, 10, 20, 50, 100, 200\}$ 、Block HolE のブロック数 $b \in \{2, 4, 8, 20, 40\}$ の範囲でグリッドサーチして決定した。

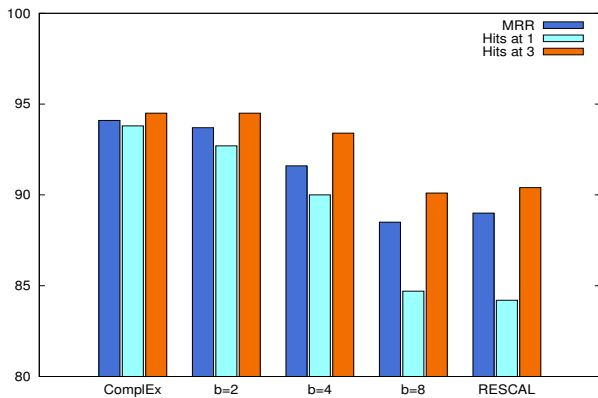


図3 $bn = 200$ に固定して、Block HoIE のブロック数 b を 2、4、8 と変化させた時の WN18 上でのランキング精度

6.6 実験結果: ランキングタスク

表5に各種モデルのランキング精度を示す。この結果から、Block HoIE は ComplEx と同程度の精度を達成できることがわかった。これはブロック数 $b = 2$ 、次元数 $n = 100$ に設定した時の結果であり、三つ組のモデル化には不要である関係行列の非対角要素があるにも関わらず [8]、RESCAL のような過学習は起こっていない。しかし、一方で、図3に示すように、 $bn = 200$ に固定して、ブロック数を 2、4、8 と増やしていくと、そのランキング精度が低下しているのがわかる。これは余剰な非対角成分が増加することによって引き起こされる過学習だと考えられ、RESCAL で起こっている問題と本質的に同じ現象だと考えられる。

7. まとめ

本稿では、知識グラフ埋め込みモデルにおける従来法の問題点を指摘し、それを解決するための新しいモデルを提案した。また、いくつかの質問応答タスクにおいて、その有効性を確認した。今後の課題として、ブロック化によって増えるパラメータを自動的に削減して過学習の問題を緩和することが考えられる。これには、真鍋らによって提案された ComplEx のための L1 正則化を Block HoIE へ拡張することが考えられる [11], [24]。

参考文献

[1] Kurt Bollacker, Colin Evans, Praveen Paritosh, Tim Sturge, and Jamie Taylor. Freebase: A collaboratively created graph database for structuring human knowledge. In *Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data (SIGMOD)*, pp. 1247–1250, 2008.

[2] Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko. Translating embeddings for modeling multi-relational data. In *Advances in Neural Information Processing Systems 26 (NIPS)*, pp. 2787–2795, 2013.

[3] Tim Dettmers, Pasquale Minervini, Pontus Stenetorp, and Sebastian Riedel. Convolutional 2d knowledge graph embeddings. *arXiv preprint arXiv:1707.01476*, 2017.

[4] Xin Dong, Evgeniy Gabrilovich, Jeremy Heitz, Wilko Horn,

Ni Lao, Kevin Murphy, Thomas Strohmann, Shaohua Sun, and Wei Zhang. Knowledge vault: A web-scale approach to probabilistic knowledge fusion. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining (SIGKDD)*, pp. 601–610. ACM, 2014.

[5] Christiane Fellbaum. *WordNet: An Electronic Lexical Database*. MIT Press, 1998.

[6] Kelvin Guu, John Miller, and Percy Liang. Traversing knowledge graphs in vector space. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 318–327. Association for Computational Linguistics, September 2015.

[7] Katsuhiko Hayashi and Masashi Shimbo. On the equivalence of holographic and complex embeddings for link prediction. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 554–559, 2017.

[8] Seyed Mehran Kazemi and David Poole. Simple embedding for link prediction in knowledge graphs. *CoRR*, Vol. abs/1802.04868, , 2018.

[9] Yankai Lin, Zhiyuan Liu, Maosong Sun, Yang Liu, and Xuan Zhu. Learning entity and relation embeddings for knowledge graph completion. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI)*, Vol. 15, pp. 2181–2187, 2015.

[10] Hanxiao Liu, Yuexin Wu, and Yiming Yang. Analogical inference for multi-relational embeddings. In *Proceedings of the 34th International Conference on Machine Learning (ICML)*, pp. 2168–2178, 2017.

[11] Hitoshi Manabe, Katsuhiko Hayashi, and Masashi Shimbo. Data-dependent learning of symmetric/anti-symmetric relations for knowledge base completion. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*, 2018.

[12] Dat Quoc Nguyen, Kairit Sirts, Lizhen Qu, and Mark Johnson. Stranse: a novel embedding model of entities and relationships in knowledge bases. In *Proceedings of the 15th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pp. 460–466, 2016.

[13] Maximilian Nickel, Kevin Murphy, Volker Tresp, and Evgeniy Gabrilovich. A review of relational machine learning for knowledge graphs. *Proceedings of the IEEE*, Vol. 104, No. 1, pp. 11–33, 2016.

[14] Maximilian Nickel, Lorenzo Rosasco, and Tomaso Poggio. Holographic embeddings of knowledge graphs. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI)*, pp. 1955–1961, 2016.

[15] Maximilian Nickel, Volker Tresp, and Hans-Peter Kriegel. A three-way model for collective learning on multi-relational data. In *Proceedings of the 28th International Conference on Machine Learning (ICML)*, pp. 809–816, 2011.

[16] Taisuke Sato. A linear algebraic approach to datalog evaluation. *Theory and Practice of Logic Programming*, Vol. 17, No. 3, pp. 244–265, 2017.

[17] Richard Socher, Danqi Chen, Christopher D Manning, and Andrew Ng. Reasoning with neural tensor networks for knowledge base completion. In *Advances in Neural Information Processing Systems 26 (NIPS)*, pp. 926–934, 2013.

[18] Fabian M. Suchanek, Gjergji Kasneci, and Gerhard Weikum. Yago: A core of semantic knowledge. In *Proceedings of the 16th International Conference on World Wide Web (WWW)*, pp. 697–706, 2007.

[19] Théo Trouillon, Johannes Welbl, Sebastian Riedel, Éric Gaussier, and Guillaume Bouchard. Complex embeddings for

Models	WN18					FB15k				
	MRR		Hits at			MRR		Hits at		
	Filter	Raw	1	3	10	Filter	Raw	1	3	10
TransE*	45.4	33.5	8.9	82.3	93.4	38.0	22.1	23.1	47.2	64.1
RESCAL**	89.0	60.3	84.2	90.4	92.8	35.4	18.9	23.5	40.9	58.7
DistMult*	82.2	53.2	72.8	91.4	93.6	65.4	24.2	54.6	73.3	82.4
HoIE*	93.8	61.6	93.0	94.5	94.9	52.4	23.2	40.2	61.3	73.9
ComplEx*	94.1	58.7	93.6	94.5	94.7	69.2	24.2	59.9	75.9	84.0
ComplEx	94.1	59.5	93.8	94.5	94.7	69.0	23.7	59.5	75.6	83.8
Block HoIE	93.7	60.3	92.7	94.5	94.7	68.2	24.5	59.0	75.5	83.5

表 5 ランキング精度: (Filtered / raw) MRR と filtered Hits at {1, 3, 10} (%). *は [19]、**は [14] から引用した結果を示す。

simple link prediction. In *Proceedings of the 33rd International Conference on Machine Learning (ICML)*, pp. 2071–2080, 2016.

- [20] Yanjie Wang, Rainer Gemulla, and Hui Li. On multi-relational link prediction with bilinear models. *arXiv preprint arXiv:1709.04808*, 2017.
- [21] Zhen Wang, Jianwen Zhang, Jianlin Feng, and Zheng Chen. Knowledge graph embedding by translating on hyperplanes. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI)*, Vol. 14, pp. 1112–1119, 2014.
- [22] Bishan Yang, Wen-tau Yih, Xiaodong He, Jianfeng Gao, and Li Deng. Embedding entities and relations for learning and inference in knowledge bases. *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, 2015.
- [23] 林克彦, 新保仁, 永田昌明. フーリエ領域上でのホログラフィック埋め込み. 言語処理学会第 23 回年次大会, pp. 314–317, 2017.
- [24] 真鍋陽俊, 林克彦, 新保仁. 知識ベース埋め込みのためのペアワイズ積 11 正則化. 言語処理学会第 24 回年次大会, pp. 1003–1006, 2018.