

Shrinking Embeddings for Hyper-Relational Knowledge Graphs

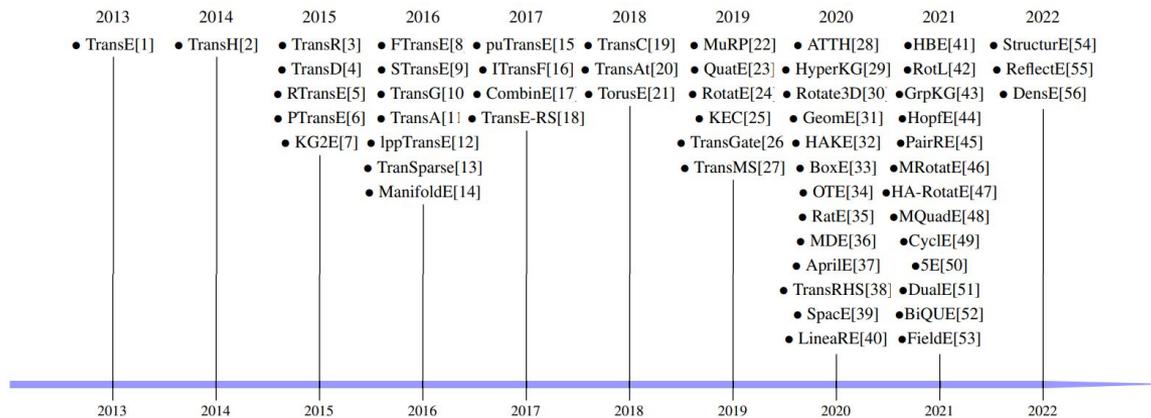
Bo Xiong¹, Mojtaba Nayyeri¹, Shirui Pan², Steffen Staab^{1,3}
¹University of Stuttgart, ²Griffith University, ³University of Southampton

@ACL2023

読み手：井田龍希（豊田工業大学 知能数理研究室 M2）

本論文を選んだ理由

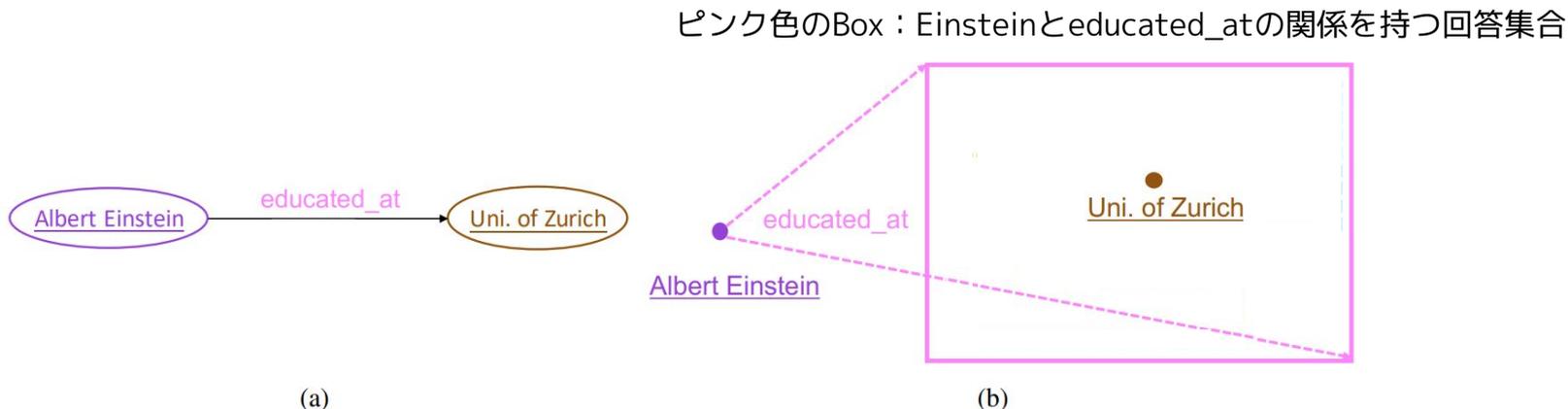
- 知識グラフ (Knowledge Graph) の埋め込み手法は数多く存在
- 最近の興味：KGの複雑な構造を簡略化せずに埋め込むには？
- 本論文の概要：関係に対するより詳細な事実をモデル化



[Ge et al., Knowledge Graph Embedding: An Overview]

まとめ

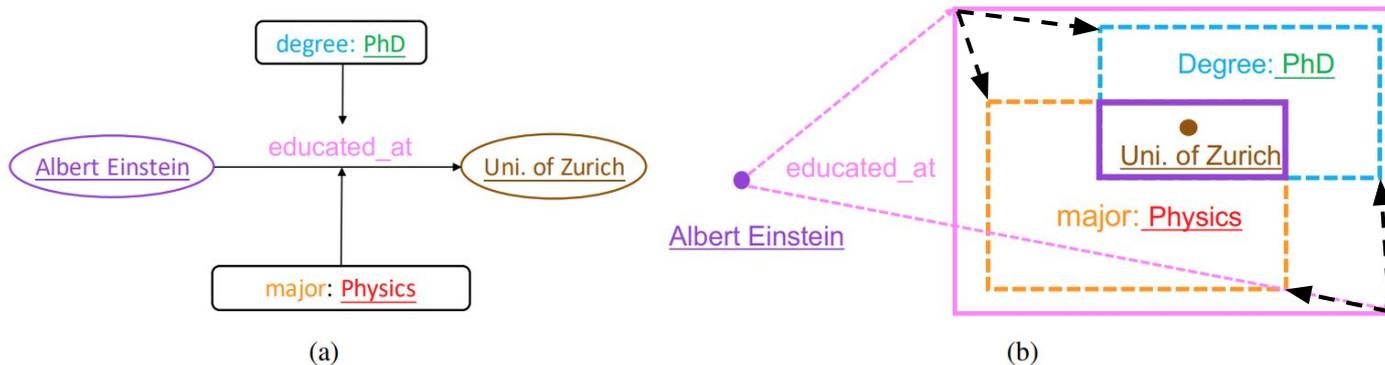
- 用語を点, その用語と特定の関係を持つ回答候補をBoxで表現
- 関係に対するより詳細な事実の追加をBoxの縮小で表現
 - 追加情報により, 回答の候補が絞られるイメージ



まとめ

- 用語を点, その用語と特定の関係を持つ回答候補をBoxで表現
- 関係に対するより詳細な事実の追加をBoxの縮小で表現
 - 追加情報により, 回答の候補が絞られるイメージ

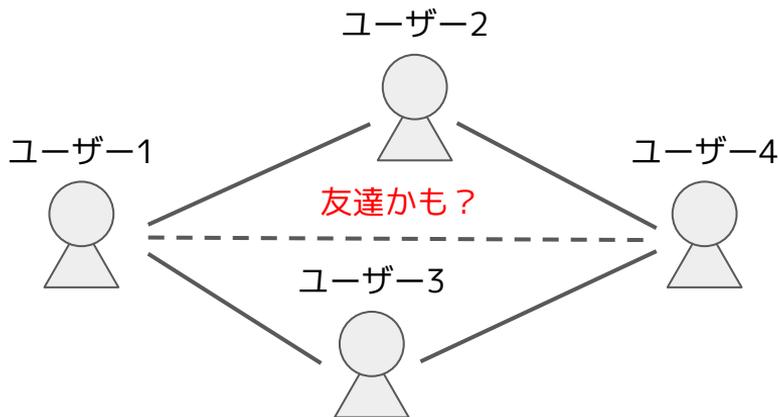
水色のBox : 取得した学位PhDという追加情報によりピンク色のBoxを縮小
オレンジ色のBox : 物理専攻という追加情報によりピンク色のBoxを縮小



背景 | リンク予測

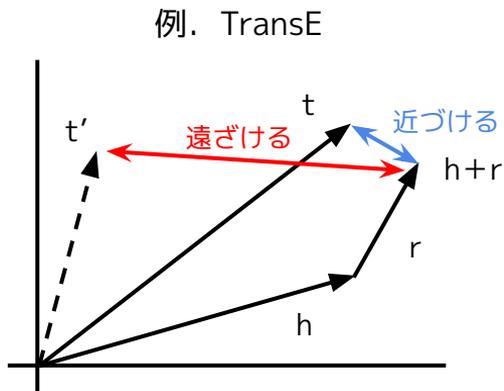
- グラフ：人や概念などを表す **ノード間の関係をエッジで表現したデータ構造**
- リンク予測：**ノード間の新たな繋がりを予測するタスク**
 - 推薦タスク, 知識グラフの補完などで利用

例. SNSのユーザー間の関係を表したグラフにおけるリンク予測



背景 | グラフ埋め込み手法

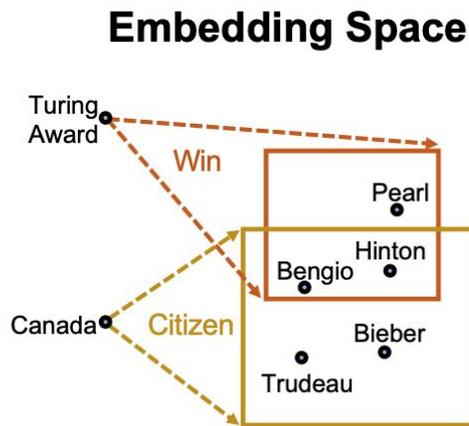
- リンク予測に対するアプローチの一つ
 - ノード, エッジをグラフ構造を考慮したベクトルに埋め込む
- **グラフをトリプルの集合**とみなし, そのトリプルをモデル化する手法が多数
 - トリプル: 始点ノード h と終点ノード t の間の関係 r を用いた (h, r, t)



t' : 始点ノード h と関係 r を持たないノード
目標: $h+r$ と t の距離が近く, かつ,
 $h+r$ と t' の距離が遠い埋め込みの獲得

背景 | Box埋め込み

- ノードを点, 関係をBoxで表現
- 各次元についてBoxの最小の座標と最大の座標を埋め込み



オレンジ色のBox : Turing AwardとWinの関係を持つ回答集合
黄色のBox : CanadaとCitizenの関係を持つ回答集合

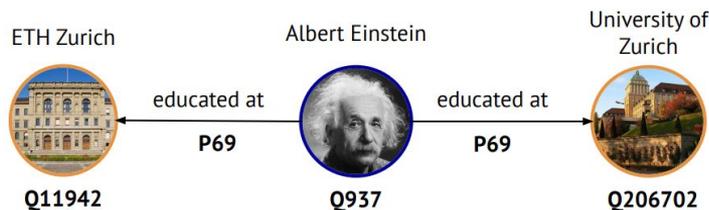
[Ren et al., Query2box: Reasoning over Knowledge Graphs in Vector Space using Box Embeddings, ICLR2020]

背景 | Hyper-relational facts

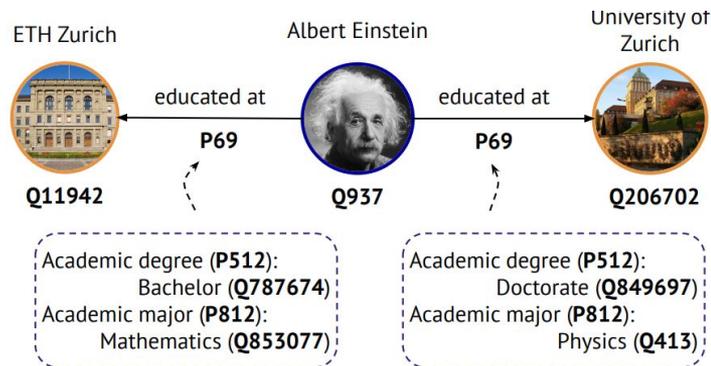
- トリプルの問題点：関係に対する詳細な知識を表現不可
- Hyper-relational facts：主要トリプルと複数の修飾子で表現
 - 主要トリプルが修飾子によりコンテキスト化され，曖昧性解消

EinsteinがETH Zurichで与えられた学位は？専攻は？

⇒ Einsteinと複数の大学間の関係は**区別不可**



educated_atの関係に対して学位・専攻の修飾子が付与
⇒ Einsteinと複数の大学間の関係は**区別可**



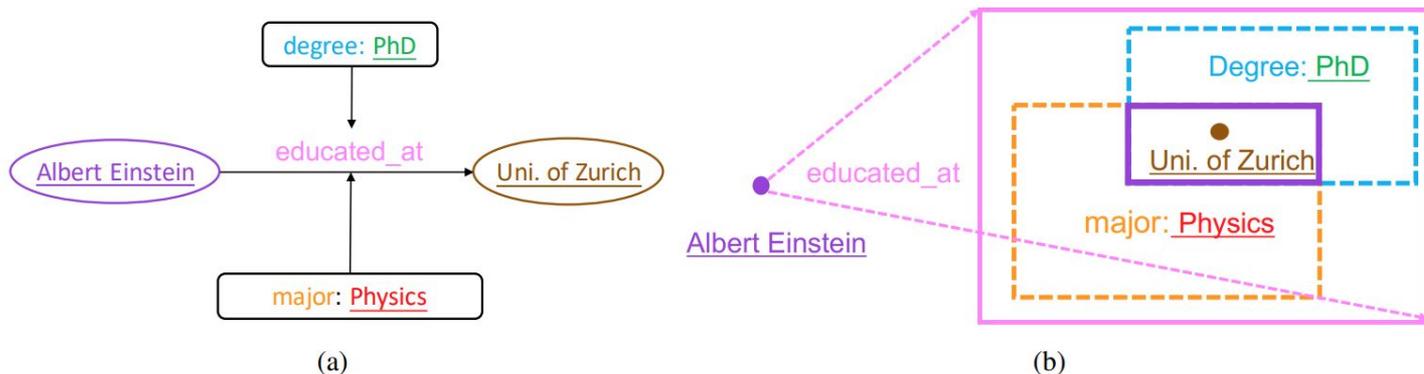
[Galkin et al., Message Passing for Hyper-Relational Knowledge Graphs, EMNLP2020]

背景 | Hyper-relational KGsの埋め込み

- Hyper-relational factsを単純な複数トリプルの組み合わせで表現
- 修飾子の論理的な特性を表現不可
 - 修飾子の単調性：修飾子を追加した時、候補となる回答セットは増加しないなど
- 目的：修飾子の論理的な特性を明示的にモデル化した埋め込みの実現

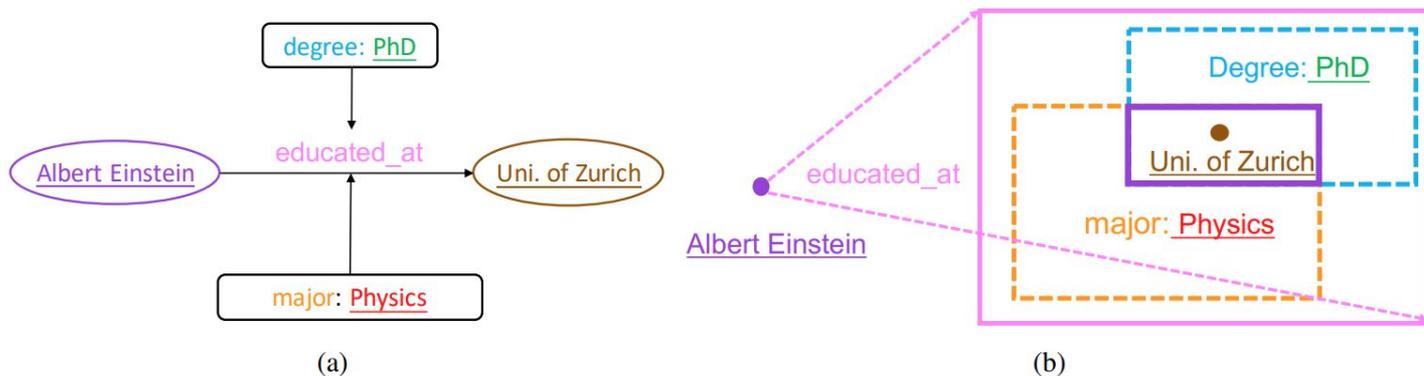
ShrinkE | 概要

- Hyper-relational factsの修飾子による論理的な特性を明示的にモデル化
 - 関係をBoxで表現し, そのBox間の空間的な関係によって特性をモデル化
 - 主要トリプル: 始点ノードから関係固有のBoxへの変換により表現
 - 修飾子: 主要トリプルのBoxの縮小により表現



ShrinkE | 概要

- Hyper-relational factsの修飾子による論理的な特性を明示的にモデル化
 - 関係をBoxで表現し、そのBox間の空間的な関係によって特性をモデル化
 - 主要トリプル：始点ノードから関係固有のBoxへの変換により表現
 - 修飾子：主要トリプルのBoxの縮小により表現



ShrinkE | 主要トリプルの埋め込み

- 始点ノードから関係固有のBoxへの変換により表現
 - Box内の全ての点が回答の候補に対応

具体的には, 関係固有の点に変換 → 点をBoxに拡張

始点ノードの表現

$$\mathcal{B}_r : \mathbb{R}^d \rightarrow \text{Box}(d)$$

各次元についてBoxの最小の座標と最大の座標

$$\text{Box}^d(\mathbf{m}, \mathbf{M}) = \{\mathbf{x} \in \mathbb{R}^d \mid \mathbf{m}_i \leq \mathbf{x}_i \leq \mathbf{M}_i, i = 1, \dots, d\}. \quad (3)$$

Point-to-point

関係固有の平行移動 + 回転

$$\mathcal{H}_r(\mathbf{e}_h) = \Theta_r \mathbf{e}_h + \mathbf{b}_r \quad (5)$$

Point-to-Box

関係固有のオフセットベクトルにより範囲を決定

$$\mathcal{B}_r(\mathbf{e}_h) = \text{Box}(\mathcal{H}_r(\mathbf{e}_h) - \tau(\boldsymbol{\delta}_r), \mathcal{H}_r(\mathbf{e}_h) + \tau(\boldsymbol{\delta}_r)), \quad (4)$$

τ はソフトプラス関数

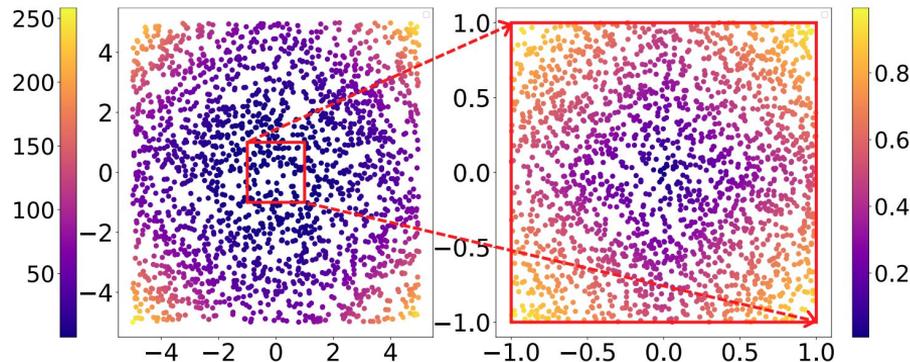
ShrinkE | 主要トリプルの埋め込み

- 主要トリプルの埋め込みの妥当性の指標として, 回答とBoxの距離関数
 - Boxの内側: 距離は比較的遅く増加、Boxのサイズと逆相関
 - Boxの外側: 距離は急速に増加

e: 回答のベクトル (点)

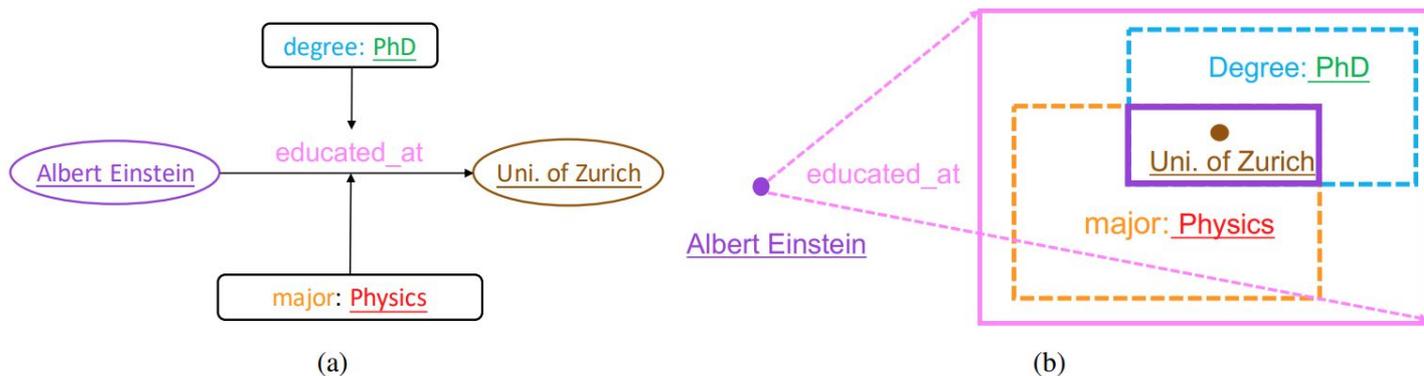
c: Boxの中心点

$$D(\mathbf{e}, \text{Box}(\mathbf{m}, \mathbf{M})) = \frac{|\mathbf{e} - \mathbf{c}|_1}{|\max(\mathbf{0}, \mathbf{M} - \mathbf{m})|_1} + (|\mathbf{e} - \mathbf{m}|_1 + |\mathbf{e} - \mathbf{M}|_1 - |\max(\mathbf{0}, \mathbf{M} - \mathbf{m})|_1)^2.$$



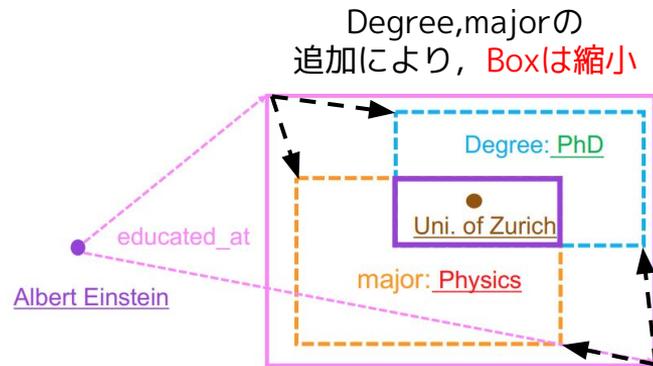
ShrinkE | 概要

- Hyper-relational factsの修飾子による論理的な特性を明示的にモデル化
 - 関係をBoxで表現し、そのBox間の空間的な関係によって特性をモデル化
 - 主要トリプル：始点ノードから関係固有のBoxへの変換により表現
 - 修飾子：主要トリプルのBoxの縮小により表現



ShrinkE | 修飾子の埋め込み

- 修飾子の追加によりBoxは必ず縮小
 - 修飾子の単調性を実現
- Boxの縮小度合いをモデル化
 - 主要トリプル, 修飾子の関係が深い → 小さな縮小, トリプルに弱い制約
 - 主要トリプル, 修飾子の関係が浅い → 大きな縮小



$$\mathbf{s}_{r,k,v}, \mathbf{S}_{r,k,v} = \text{MLP}(\text{concat}(r_{\theta}, k_{\theta}, v_{\theta}))$$

それぞれ
 r_{θ} : 主要トリプルの関係
 k_{θ} : 修飾子のkey
 v_{θ} : 修飾子のvalueの埋め込み

Boxの縮小

$$\mathcal{S}_{r,k,v}(\text{Box}(\mathbf{m}, \mathbf{M})) = \text{Box}(\mathbf{m} + \sigma(\mathbf{s}_{r,k,v}) \odot \mathbf{L}, \mathbf{M} - \sigma(\mathbf{S}_{r,k,v}) \odot \mathbf{L}), \quad (\mathbf{L} = \mathbf{M} - \mathbf{m})$$

(8)

ShrinkE | スコア関数

- 修飾子集合Qによる**全ての縮小Boxが重なる部分と回答の距離**
 - hと関係rを持たないノードt'をサンプリングして負例を作成
 - h, r, QによるBoxとtが近く, かつ, t'と遠くなるように学習

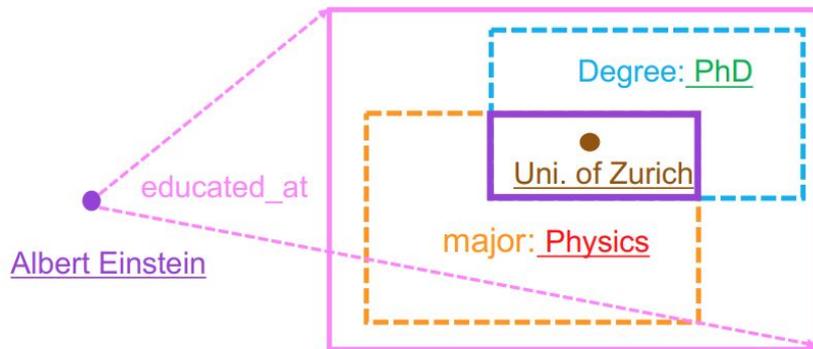
スコア関数

$$f(((h, r, t), Q)) = D(\mathbf{e}_t, \text{Box}_Q(\mathbf{m}, \mathbf{M})), \quad (9)$$

(再掲) 距離関数

$$D(\mathbf{e}, \text{Box}(\mathbf{m}, \mathbf{M})) = \frac{|\mathbf{e} - \mathbf{c}|_1}{|\max(\mathbf{0}, \mathbf{M} - \mathbf{m})|_1} + (|\mathbf{e} - \mathbf{m}|_1 + |\mathbf{e} - \mathbf{M}|_1 - |\max(\mathbf{0}, \mathbf{M} - \mathbf{m})|_1)^2.$$

水色のBoxとオレンジ色のBoxが重なった
紫色のBoxが回答集合



ShrinkE | 修飾子の特性

- 修飾子の単調性：修飾子を追加した時，回答の候補は増加しない
 - 修飾子の含意：既存の修飾子に暗黙的に含意される修飾子を追加しても回答の候補は同じ
 - 修飾子の相互排他性：相互に排他的な2つの修飾子を追加すると矛盾
- ⇒ 縮小したBox間の空間的な関係によりこれらの特性を表現可能

実験設定

	All facts	Higher-arity facts (%)	Entities	Relations	Train	Dev	Test
JF17K	100,947	46,320 (45.9%)	28,645	501	76,379	–	24,568
WikiPeople	382,229	44,315 (11.6%)	47,765	193	305,725	38,223	38,281
WD50k	236,507	32,167 (13.6%)	47,156	532	166,435	23,913	46,159
WD50K(33)	102,107	31,866 (31.2%)	38,124	475	73,406	10,568	18,133
WD50K(66)	49,167	31,696 (64.5%)	27,347	494	35,968	5,154	8,045
WD50K(100)	31,314	31,314 (100%)	18,792	279	22,738	3,279	5,297

- 評価：Hyper-relational factsの欠損した始点，終点ノードのリンク予測
 - 評価指標：Mean reciprocal rank (MRR), H@K (K=1,10)
- データセット：JF17K, WikiPeople, WD50k
 - WikiPeople, WD50Kは修飾子を含む関係の割合が低い
 - Hyper-relational factsの割合が高いWD50kの3つの分割
 - それぞれ33%, 66%, 100%の関係が修飾子を含む関係
- 比較手法
 - m-TransH, StarEなどのHyper-relational KGsを対象にした既存手法

実験結果

- すべてのデータセットで高性能
 - Hyper-relational factsの割合が高いJF17Kでは、最高性能

データセット名(Hyper-relational factsの割合)

Method	WikiPeople (2.6)			JF17K (45.9)			WD50K (13.6)		
	MRR	H@1	H@10	MRR	H@1	H@10	MRR	H@1	H@10
m-TransH	0.063	0.063	0.300	0.206	0.206	0.463	—	—	—
RAE	0.059	0.059	0.306	0.215	0.215	0.469	—	—	—
NaLP-Fix	0.420	0.343	0.556	0.245	0.185	0.358	0.177	0.131	0.264
NeuInfer	0.350	0.282	0.467	0.451	0.373	0.604	—	—	—
HINGE	0.476	0.415	0.585	0.449	0.361	0.624	0.243	0.176	0.377
Transformer	0.469	0.403	0.586	0.512	0.434	0.665	0.264	0.194	0.401
BoxE	0.395	0.293	0.503	0.560	0.472	0.722	—	—	—
StarE	0.491	0.398	0.648	0.574	0.496	0.725	0.349	<u>0.271</u>	0.496
ShrinkE	<u>0.485</u>	0.431	<u>0.601</u>	0.589	0.506	0.749	<u>0.345</u>	0.275	<u>0.482</u>

実験結果

- Hyper-relational factsの割合が高いデータセットでは性能向上が顕著
 - すべての関係が修飾子を持つWD50k(100)では性能向上の幅が大きい

Method	WD50K (33)			WD50K (66)			WD50K (100)		
	MRR	H@1	H@10	MRR	H@ 1	H@ 10	MRR	H@ 1	H@ 10
NaLP-Fix	0.204	0.164	0.277	0.334	0.284	0.423	0.458	0.398	0.563
HINGE	0.253	0.190	0.372	0.378	0.307	0.512	0.492	0.417	0.636
Transformer	0.276	0.227	0.371	0.404	0.352	0.502	0.562	0.499	0.677
StarE	<u>0.331</u>	<u>0.268</u>	0.451	<u>0.481</u>	<u>0.420</u>	<u>0.594</u>	<u>0.654</u>	<u>0.588</u>	<u>0.777</u>
ShrinkE	0.336	0.272	<u>0.449</u>	0.511	0.422	0.611	0.695	0.629	0.814

Ablation Study

- 平行移動, 回転, Boxの縮小のすべてが性能向上に寄与
 - Boxの縮小を削除した時の性能低下が大きい
- ⇒ 縮小による修飾子の特性の明示的なモデル化は高い有用性

Method	MRR	H@ 1	H@ 10
ShrinkE (w/o translation)	0.583	0.495	0.729
ShrinkE (w/o rotation)	0.581	0.497	0.724
ShrinkE (w/o shrinking)	0.571	0.490	0.711
ShrinkE	0.589	0.506	0.749

Ablation Study

- 平行移動, 回転, Boxの縮小のすべてが性能向上に寄与
 - Boxの縮小を削除した時の性能低下が大きい
- ⇒ 縮小による修飾子の特性の明示的なモデル化は高い有用性

Method	MRR	H@ 1	H@ 10
ShrinkE (w/o translation)	0.583	0.495	0.729
ShrinkE (w/o rotation)	0.581	0.497	0.724
ShrinkE (w/o shrinking)	0.571	0.490	0.711
ShrinkE	0.589	0.506	0.749

Case Study

- bodyが満たされる時, headも満たされる

body	head
(residence: Monte Carlo)	(country, Monaco)
(residence: Belgrade)	(country, Serbia)
(owned_by: X)	(of, voting interest)
(emergency phone number: Y)	(has_use, police)
(emergency phone number: Z)	(has_use, fire department)
(used_by: software)	(via, operating_system)

Table 5: Example pairs of qualifiers with implication relations (body \rightarrow head). $X \in [Eric\ Schmidt, Mark\ Zuckerberg, Dustin\ Moskovitz, Larry\ Page]$ denotes a CEO name of a company. $Y \in [112, 115, 113, \dots]$ and $Z \in [912, 18, 192, \dots]$ are emergency numbers involving *police* and *fire department*, respectively. Qualifier exclusion pairs are ubiquitous and are hence omitted.

まとめ

- トリプルよりも複雑な関係を表現可能なHyper-relational factsを対象
- 修飾子の単調性などの特性を明示的にモデル化した埋め込みでリンク予測
 - 縮小による修飾子の特性の明示的なモデル化の有効性を確認

