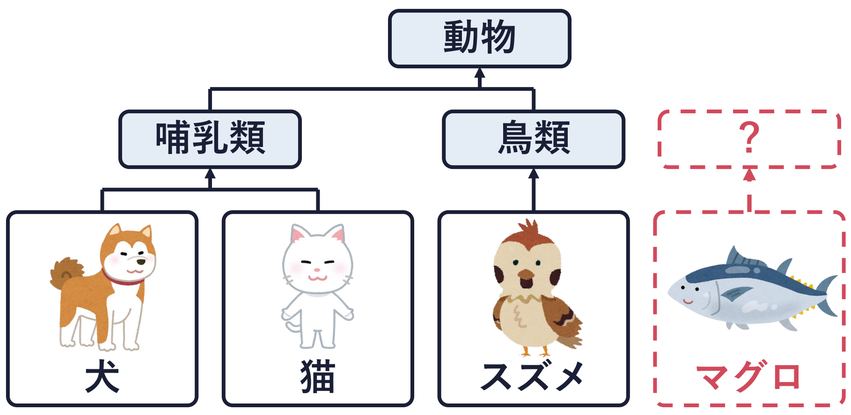
異種オントロジー埋め込みモデルの予測統合による知識グラフ補完

柴田 陽向　兼岩研究室
電気通信大学 情報理工学域 I 類 CS プログラム

① 背景と目的

(1) 何を補完するのか

オントロジー：概念と関係を体系的に表した知識構造



箱：概念
例）動物，犬

線：親子関係
例）犬 → 哺乳類

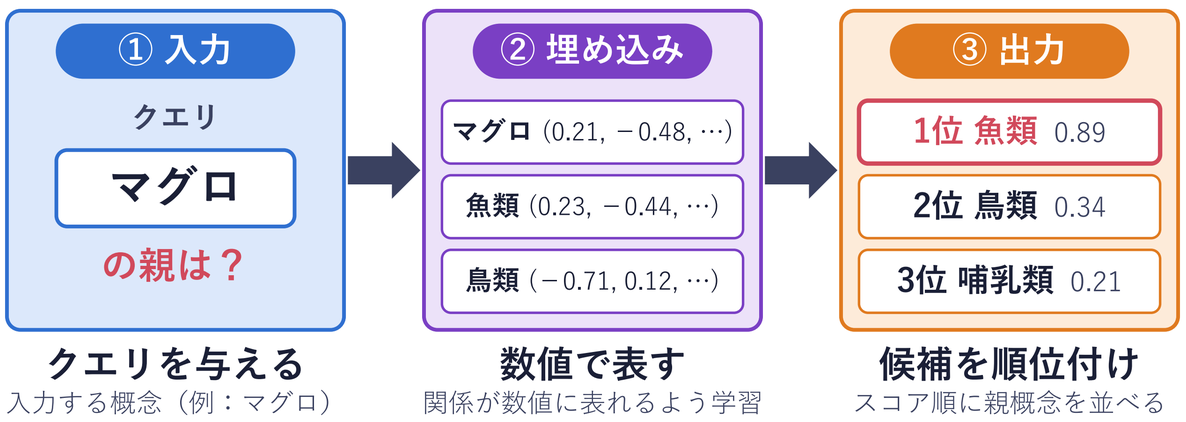
？：未記述の関係
例）マグロの親



「？」を既知の関係から予測して埋める → **補完**

(2) どうやって補完するのか

埋め込みモデル：概念を数値で表し，親概念の候補を順位付け



クエリを与える
入力する概念（例：マグロ）

数値で表す
関係が数値に表れるよう学習

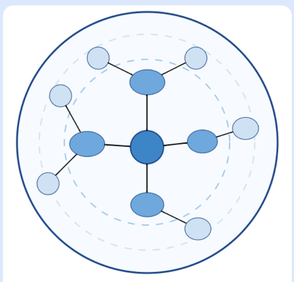
候補を順位付け
スコア順に親概念を並べる

着眼点：モデルごとに使う情報・表現が違う
→ 得意なクエリも違うのでは？

目的：複数モデルの予測を統合し，補完性能が上がるかを検証する

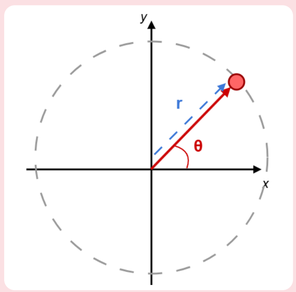
② 使用モデルと評価指標

OnT [1]



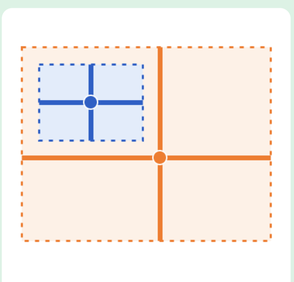
双曲空間で表現

RotatE-HAKE [2]



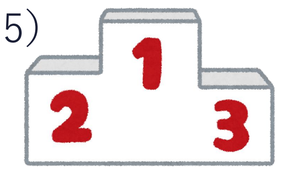
角度と半径で表現

ELAE [3]



包含関係を箱で表現

MRR：正解順位の逆数の平均（1位→1，2位→0.5）
H@k：正解が上位k件以内に入る割合
→ どちらも**高いほど良い**



③ モデル間の予測差

データ	OnT最良	RotatE最良	ELAE最良	単体最高MRR	クエリ別最良 MRR*
GO	59.1%	29.0%	8.5%	0.1716	**0.2839**
GALEN	61.7%	17.5%	19.9%	0.1680	**0.2207**

*クエリごとに事後的に3モデル中の最良順位を選んだ参考値（実際には使えない）

最良のモデルはクエリごとに異なる
クエリ別最良は単体最高を上回る → **相補性がある可能性**

④ 予測の統合手法

例：クエリ「マグロ」の親概念候補の順位（イメージ）　**正解：魚類**

OnT	RotatE-HAKE	ELAE	統合結果
1位　魚類	1位　鳥類	1位　哺乳類	**1位　魚類**
2位　鳥類	**2位　魚類**	**2位　魚類**	2位　鳥類
3位　哺乳類	3位　哺乳類	3位　鳥類	3位　哺乳類

順位を使う
重み付きRRF
各モデルの順位を合成

スコアを使う
重み付きスコア統合
各モデルの点数を足し合わせ

信頼できるモデルほど重く扱う
（重みは検証用データで決定）



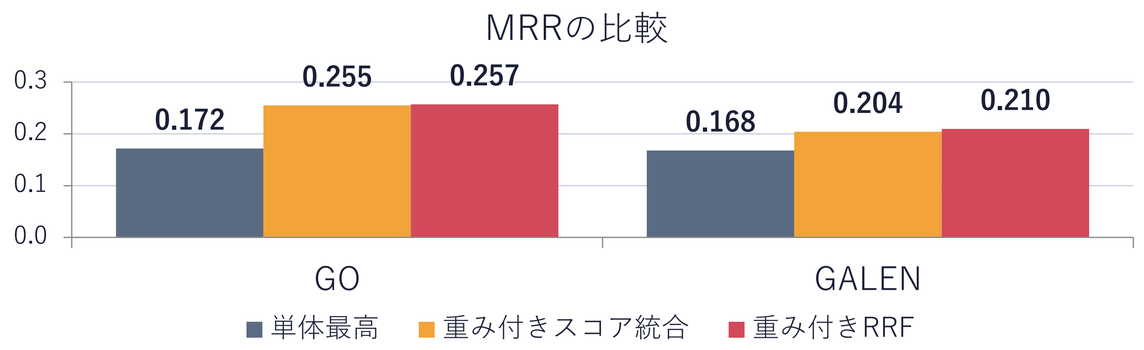
⑤ 実験設定

データ：GO（遺伝子機能），GALEN（医療・解剖学）
3モデルの予測は固定し，**統合部分のみを評価**
重みは検証用データで選び，別の評価用データで測定

クエリ数／候補概念数　GO：6,882／44,244　GALEN：5,390／23,141

⑥ 実験結果

両データセットで統合手法が単体最高MRRを上回った
GO 0.1716 → 0.2571　／　GALEN 0.1680 → 0.2096



手法	GO				GALEN			
	MRR	H@1	H@10	H@100	MRR	H@1	H@10	H@100
OnT	0.1335	0.0705	0.2553	0.6622	0.1680	0.1022	0.2917	0.5388
RotatE-HAKE	0.1716	0.1177	0.2870	0.4054	0.0561	0.0330	0.0972	0.2508
ELAE	0.0564	0.0257	0.1186	0.3344	0.0246	0.0056	0.0506	0.2926
重み付きスコア統合	0.2549	**0.1880**	0.3797	0.5974	0.2039	0.1276	0.3503	0.5868
重み付きRRF	**0.2571**	0.1699	**0.4350**	**0.7380**	**0.2096**	**0.1314**	**0.3638**	**0.6239**

太字・赤字は最高値．GOのH@100のみ，重み付きスコア統合はOnT単体を下回る

⑦ まとめと今後の課題

【まとめ】
- モデルごとに最良のクエリが異なる（相補性を示唆）
- 統合手法は両データセットで単体最高MRRを上回った

【今後の課題】
- クエリの構造的特徴から，統合の重みを学習して動的に決める
- 統合で絞った上位候補を，学習モデルで並べ替える（リランキング）
- 統合とリランキングを組み合わせる

参考文献

[1] H. Yang et al., "Language Models as Ontology Encoders," ISWC 2025, LNCS 16140, pp. 443–461, 2025.
[2] 飯島, 兼岩, 「オントロジーの論理構造と語彙情報を考慮した概念包含のための埋め込み学習」, SWO-062, 2024.
[3] 山中, 白石, 兼岩, 「概念包含の論理的推論に対するロール埋め込み手法」, 知識ベースシステム研究会 第127回, 2022.
[4] G. V. Cormack et al., "Reciprocal Rank Fusion Outperforms Condorcet and Individual Rank Learning Methods," SIGIR 2009.
[5] S. Varma and R. Simon, "Bias in Error Estimation When Using Cross-Validation for Model Selection," BMC Bioinformatics 7:91, 2006.
[6] A. Nandi et al., "DynaSemble: Dynamic Ensembling of Textual and Structure-Based Models for Knowledge Graph Completion," ACL 2024, pp. 205–216.