

ACL読み会

Improving Entity Linking by Modeling Latent Relations between Mentions

Phong Le¹ and Ivan Titov^{1,2}

¹University of Edinburgh ²University of Amsterdam

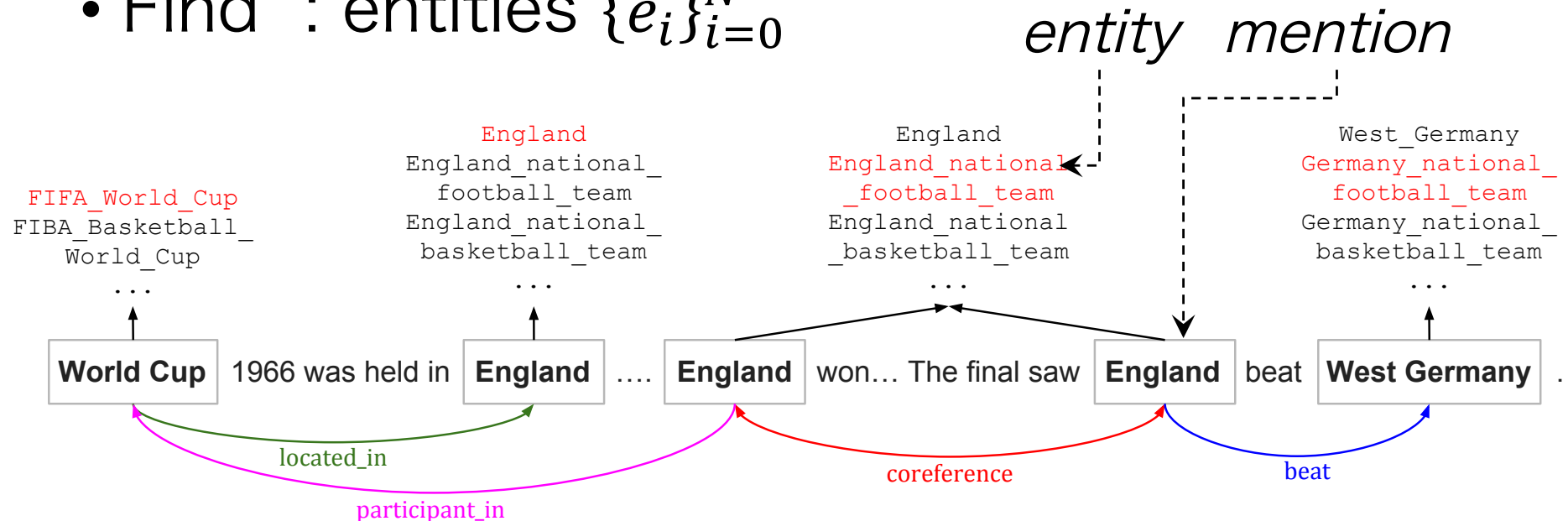
紹介者：駒水 孝裕（名古屋大学）

概要

- 問題：Entity Linking
- 従来：ヒューリスティクス，教師あり学習
- 課題：教師なしでEntity Linkingの実現
- 手法：
 1. 関係を隠れ変数として導入
 2. 複数の関係を考慮することで多様な関係を表現
- 実験：AIDA-CoNLLデータセットで最新研究と比較
- 結果：
 1. 最新研究よりも精度向上
 2. 複数の関係の導入が精度に貢献

Entity Linking

- Given: 文書 D , D 内の mentions $\{m_i\}_{i=0}^N$
- Find : entities $\{e_i\}_{i=0}^N$



Local model

$$e_i^* = \arg \max_{e_i \in C_i} \Psi(e_i, c_i)$$

Global model

$$E^* = \arg \max_{E \in C_1 \times \dots \times C_n} \sum_{i=1}^n \Psi(e_i, c_i) + \Phi(E, D)$$

既存手法

$$E^* = \arg \max_{E \in C_1 \times \dots \times C_n} \sum_{i=1}^n \Psi(e_i, c_i) + \Phi(E, D)$$

- mention間の関係

- coreference : 2つのmentionが同じentityを指示
- それ以外の関係
 - Cheng and Roth (2013), Ren et al. (2017), など
 - 人手で作成したルールに基づいてentity間の関係を定義
 - 関係抽出自体が難しいタスクなので, ほとんどない

- 表現学習

- コンテキストの表現とentityの表現の関連性
- feature engineeringから発展
 - entityに対応するWikipediaのページタイトルの単語ベクトルのコサイン類似度
- 各表現の学習とその間の“単一の”関連性の学習

$$\Phi(e_i, e_j, D) = \frac{1}{n-1} \mathbf{e}_i^T \mathbf{R} \mathbf{e}_j$$

提案手法：multi-relational model

- エンティティ間の関連性

$$E^* = \arg \max_{E \in C_1 \times \dots \times C_n} \sum_{i=1}^n \Psi(e_i, c_i) + \Phi(E, D)$$

- K 個の潜在的な関係の和で表現

$$\Phi(e_i, e_j, D) = \sum_{k=1}^K \alpha_{ijk} \Phi_k(e_i, e_j, D)$$

確信度

k 番目の関連性に
基づくスコア

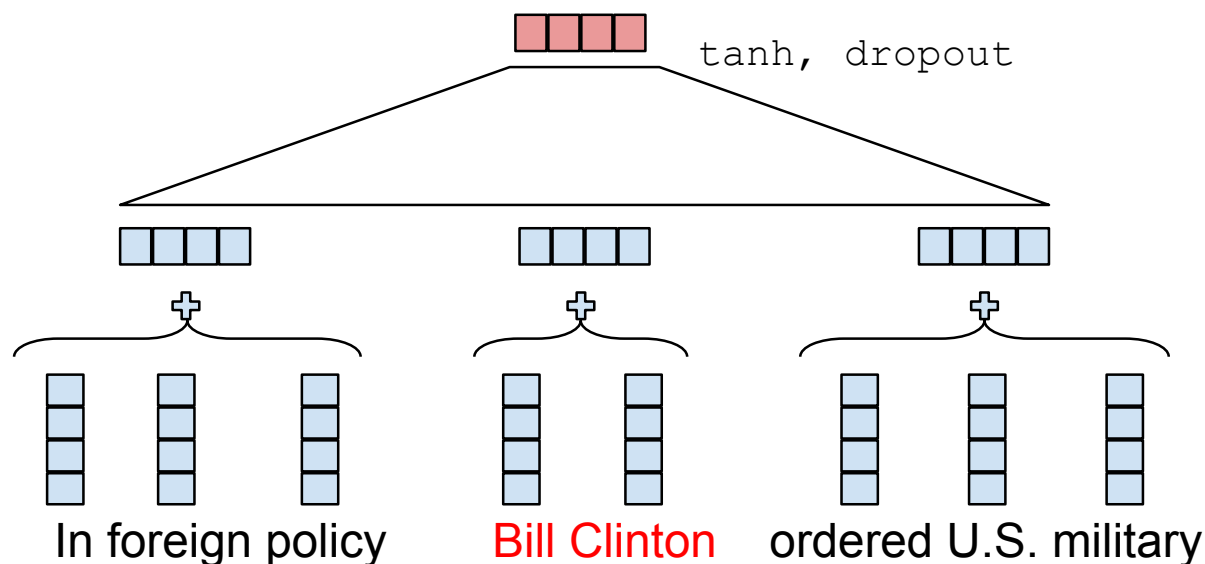
確信度

$$\alpha_{ijk} = \frac{1}{Z_{ijk}} \exp \left\{ \frac{f^T(m_i, c_i) \mathbf{D}_k f(m_j, c_j)}{\sqrt{d}} \right\}$$

- $f(m, c)$: mention m とコンテキスト c を埋め込む関数
- \mathbf{D}_k : ベクトル成分の重みを表現する対角行列
- d : 埋め込むベクトル空間の次元数
- Z_{ijk} : 正規化項

図は論文からの引用

埋め込み関数 $f(m, c)$



単層ニューラルネットワーク

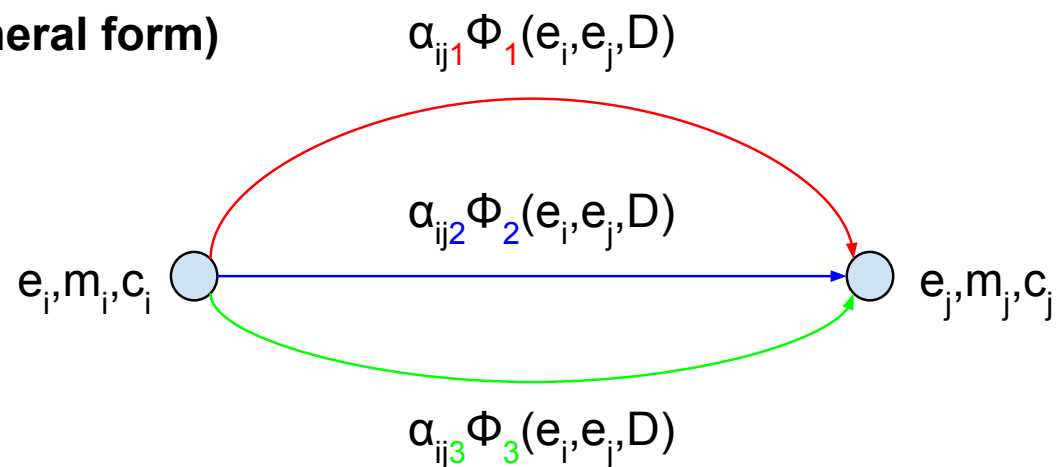
- 活性化関数：tanh
- ドロップアウト

正規化項

図は論文からの引用

$$\alpha_{ijk} = \frac{1}{Z_{ijk}} \exp \left\{ \frac{f^T(m_i, c_i) \mathbf{D}_k f(m_j, c_j)}{\sqrt{d}} \right\}$$

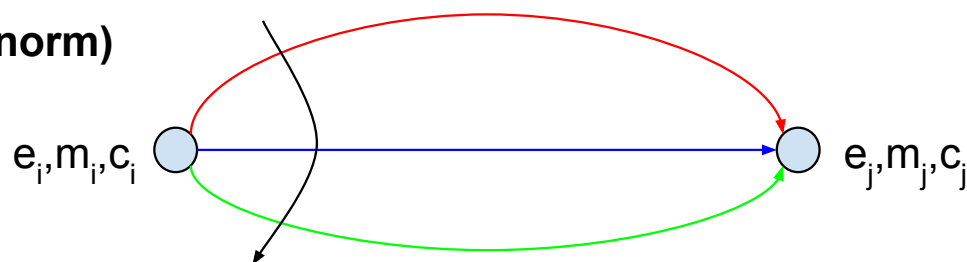
(general form)



正規化方式の選択

関係について正規化

(rel-norm)



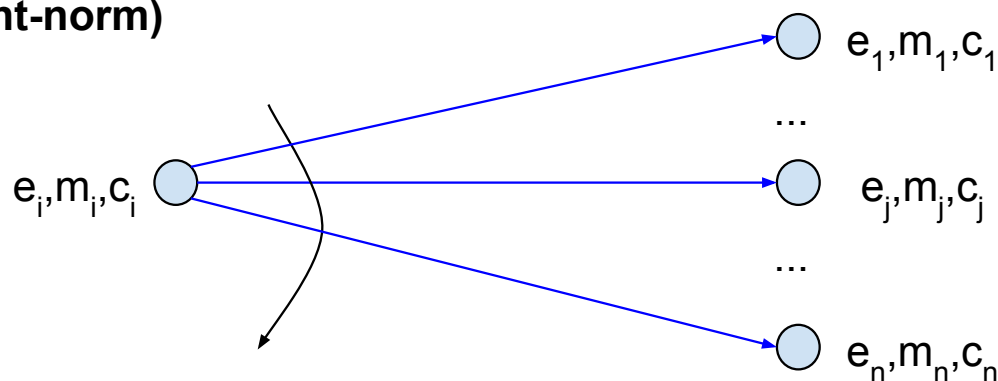
normalize over relations: $\alpha_{ij1} + \alpha_{ij2} + \alpha_{ij3} = 1$

α の役割

mention ペアがどの
関係に関連するかを
表す確率

mention について正規化

(ment-norm)



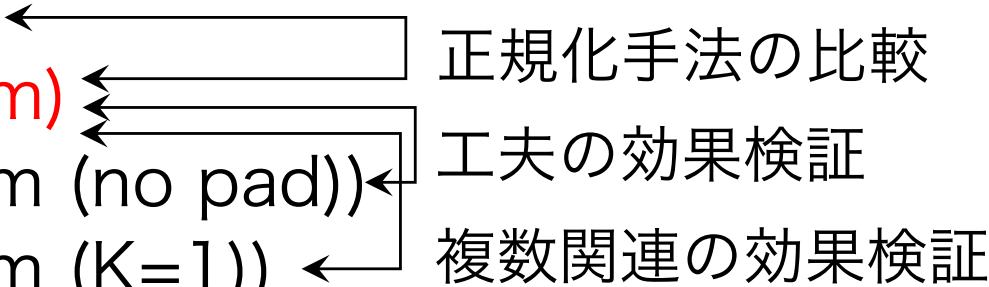
normalize over mentions:

$$\alpha_{i12} + \alpha_{i22} + \dots + \alpha_{ij2} + \dots + \alpha_{in2} = 1$$

α の役割

mention m_i に
対して関係 k を
通じて mention m_j が
関係しているかを
表す確率

実験

- Entity Linkingの精度を評価
 - 比較手法
 - 提案手法 (rel-norm) ←
 - 提案手法 (ment-norm) ←
 - 提案手法 (ment-norm (no pad)) ←
 - 提案手法 (ment-norm (K=1)) ←
 - 既存手法
 - 設定
 - データセット：AIDA-CoNLL
 - 評価指標：F1のマイクロ平均（5回, 95%信頼区間）
 - 潜在関連数：6 (rel-norm), 3 (ment-norm)
 - 実験的に決定
- 

実験結果

Methods	Aida-B	
Chisholm and Hachey (2015)	88.7	
Guo and Barbosa (2016)	89.0	
Globerson et al. (2016)	91.0	
Yamada et al. (2016)	91.5	
Ganea and Hofmann (2017)	92.22 ± 0.14	
rel-norm	92.41 ± 0.19	←
ment-norm	93.07 ± 0.27	←←
ment-norm ($K = 1$)	92.89 ± 0.21	←
ment-norm (no pad)	92.37 ± 0.26	←

ment-norm が
rel-norm より良い

複数の潜在関連を
取り入れることが良い

- 提案手法 (ment-norm) が最良
- 計算量的には既存手法より高いが収束が早いので問題にはならない

まとめ

- 問題：Entity Linking
- 従来：ヒューリスティクス，教師あり学習
- 課題：教師なしでEntity Linkingの実現
- 手法：
 1. 関係を隠れ変数として導入
 2. 複数の関係を考慮することで多様な関係を表現
- 実験：AIDA-CoNLLデータセットで最新研究と比較
- 結果：
 1. 最新研究よりも精度向上
 2. 複数の関係の導入が精度に貢献