

# ゼロショットエンティティリンキングに寄与するエンティティ記述の分析

瀬谷 浩樹† 加藤 誠††

† 筑波大学大学院 情報学学位プログラム 〒 305-8550 茨城県つくば市春日 1-2

†† 筑波大学 図書館情報メディア系 〒 305-8550 茨城県つくば市春日 1-2

E-mail: †s2221653@u.tsukuba.ac.jp, ††mpkato@acm.org

**あらまし** 我々は、ゼロショットエンティティリンキング (ZSEL) に有効なエンティティ記述を作成する際の指針を得ることを目的として、エンティティ記述内のどの部分が ZSEL の正解率向上に寄与しているかを明らかにするための分析を行った。具体的な手法としては、メンションが言及しているエンティティ (つまり正解エンティティ) のエンティティ記述内の一部をマスクあるいは置換して、その結果 ZSEL の正解率がどのように変化するかを確認した。分析の結果、エンティティ記述内に出現するエンティティ間の関係、エンティティの出現回数、エイリアスの利用頻度が ZSEL の正解率に影響していることが示唆された。

**キーワード** エンティティリンキング, ゼロショットエンティティリンキング, XAI

## 1 はじめに

質問応答 [1] や情報検索 [2] 等の自然言語処理において、ユーザが入力した文章の内容を把握するための手法としてエンティティリンキング (以降 EL) がある。EL は、エンティティ (一意に識別可能な事物) の集合と、文章中のメンション (エンティティを言及している部分) が与えられたときに、メンションがどのエンティティを言及しているかを推定するタスクである。近年は EL の中でも、専門的なエンティティ等の訓練データが少ないエンティティに対しても有効なゼロショットエンティティリンキング (以降 ZSEL) が注目されている。ZSEL は、Logeswaran らによって提案された、エンティティは文章のみで表現される、訓練セット内のエンティティ集合とテストセット内のエンティティ集合が交わりを持たないという 2 つの制約の下で行われる EL である [3]。最初の制約を満たすために、ZSEL モデルは基本的に、エンティティとそれを表現する文章 (エンティティ記述) の集合と、メンションとその周囲の文脈のみを入力として受け取り、入力されたメンションがどのエンティティを言及しているかを推定する。そして、このモデルを 2 つ目の制約下で訓練することにより、エンティティに関する構造化された情報・訓練データが得られない場合でも EL を行える ZSEL モデルを作成できる。ZSEL の重要性について、例えば Oberhauser らは、訓練データを作るためのアノテーションに高いコストがかかる医療ドメインでは、訓練なしでも EL 可能な ZSEL が有効であることを指摘している [4]。

以上のように従来の EL では対処できない場面にも有効な ZSEL であるが、Wu らによる BLINK など [5]、ZSEL モデルは既に複数提案されている。そして、中には ZSEL 結果の分析を行なっている研究もある。しかし、これらの研究は ZSEL において有効なエンティティ記述の分析は行なっておらず、現状ではどういった記述が ZSEL に有効かが不明となっている。また、XAI と呼ばれる分野において、機械学習モデルの出力に対

して、入力の中のどの部分がある出力に寄与したかを可視化する研究が行われている。しかし、ZSEL という個別のタスクに注目して、モデルの出力に対して入力されたエンティティ記述のどの部分が寄与したかを分析する研究は行われていない。

そこで本研究では、ZSEL に有効な (ZSEL の正解率向上につながる) エンティティ記述を作成する際の指針を得ることを目的として、どのようなエンティティ記述が ZSEL に有効かを分析する。分析手法としては、メンションとエンティティ記述集合が与えられたときに、そのメンションが言及しているエンティティのエンティティ記述内の一部をマスクあるいは置換し、その結果 ZSEL の正解率がどのように変化するかを確認する。

本研究の実験では、ZSEL モデルとしては BLINK [5] を、エンティティ (記述) 集合としては BLINK と共に公開されているエンティティカタログを、EL を行うメンション集合としては WikilinksNED Unseen-Mentions [6] を ZSEL の制約を満たすように改変したデータセットをそれぞれ使用して分析を行った。

上の実験から、ZSEL に有効なエンティティ記述を作成する際の指針として、以下のようなことが示唆された。

- (1) エンティティ記述に、記述対象のエンティティと特定の関係にあるエンティティを含めると、代わりに他の多くの関係にあるエンティティを含めた時よりも ZSEL の正解率が向上する。また、特定のタイプに属するエンティティ記述に関しては、他にも重要な場合が多い関係が存在する。
- (2) エンティティ記述に、記述対象のエンティティと同タイプのエンティティを説明するときによく使われるようなエンティティを含めると、使われることが少ないエンティティを含めた時よりも ZSEL の正解率が向上する。
- (3) エンティティ記述内で、エンティティを参照するとき利用頻度が低いエイリアスを使うと、それ以外の利用頻度のエイリアスを使った時よりも ZSEL の正解率が低くなる場合が多い。なお、特定のタイプに属するエンティティ記述に関しては、上とは異なる傾向が見られる。

本論文の構成は以下の通りである。第 2 節では、本研究に関

連する研究である、エンティティリンクングや寄与部分の可視化に関する研究を紹介する。第3節では、本研究において分析の対象とする EL と ZSEL を定義した後に、どのようなエンティティ記述が ZSEL に有効かを分析する手法について説明する。第4節では、第3節で説明した手法を用いて行った分析の実験について述べる。第5節では、本研究の結論を述べる。

## 2 関連研究

本節では、関連研究について説明する。具体的には、EL に関する研究と、機械学習モデルの出力に対して、入力の中の部分が寄与したかを可視化する研究について説明する。

### 2.1 エンティティリンクング

EL は典型的に、大量のエンティティの中から文章中のメンションが言及している可能性のあるエンティティを探索する「候補生成」と、候補エンティティを言及されている可能性が高い順に並び替える「ランキング」の2段階で行われる。候補生成には、Wikipedia 内のハイパーリンクを元に計算したメンションがエンティティを言及している確率や [7,8]、文書検索手法である BM25 などが用いられる [3]。ランキングには、ニューラルネットワークを用いて行なっている例が多く見られる [9–11]。

近年は、通常の EL よりも厳しい制約の下で行われる ZSEL が提案され [3]、ZSEL モデルも複数研究されている [5,12–14]。

これらの中には EL 結果の分析を行なっている研究もある。Wu らは、提案した ZSEL モデルの EL 結果を分析し、文脈に誤解を招く単語が含まれている場合は正しいエンティティを推定できないことがあると述べている [5]。また、Sun らは、訓練時とテスト時のドメインの差異に対処できる ZSEL モデルの訓練手法を提案している [12]。そして、両ドメインが同じメタドメインに属しており密接に関連している場合、提案手法を通してその関連性を利用し、より良い結果が得られると述べている。

これらの研究は ZSEL 結果を分析しているという点で本研究に類似しているが、ZSEL に有効なエンティティ記述の分析は行なっていない。本研究では、どのようなエンティティ記述が ZSEL に有効かを分析することで、ZSEL に有効なエンティティ記述の作成という、現在主流なモデルの改善とは異なる方法で ZSEL の正解率を上げられるようになることを期待する。

### 2.2 寄与部分の可視化

XAI という分野において、機械学習モデルから出力が得られたときに、入力の中の部分がその出力に寄与したかを人間が解釈できるように可視化するという研究が行われている。

文章も対象とした代表的な可視化手法として、LIME [15] や SHAP [16] が挙げられる。エンティティ記述を作成する際の指針となるような、エンティティ記述内のどの部分が ZSEL の正解率向上に寄与しているかに関する知見を得るという本研究の目的を踏まえると、上の手法には2つの問題点がある。(1) あくまで元の複雑なモデルを局所的に近似した線形モデルを用いており、元のモデルにおける入力内の寄与部分が正しく求められない可能性がある。(2) 事前に定義された観点で入力内の寄

与部分进行分析する訳ではないので、得られた寄与部分に指針となり得る規則性を見つけれられるとは限らない。よって本研究では、実際の ZSEL モデルに編集したエンティティ記述を入力し、その結果 ZSEL の正解率がどのように変化するかを確認する。

上のような広く利用可能な可視化手法の提案を行なった研究とは異なり、より狭いタスクに注目した研究も行われている。例えば、Li らは自然言語処理タスクにおいて、ニューラルネットワークに入力された文章内の単語や単語ベクトルの一部を削除し、入力のどの部分が出力に寄与したかを分析する手法を提案している [17]。当該研究の手法は本研究で用いる分析手法と類似しているが、当該研究ではあくまで個別の単語それぞれの重要性を求めており、同じ実験を ZSEL を対象に行ったとしても、得られた分析結果からエンティティ記述を作成する際の指針となるような規則性を見つけることができるとは限らない。

## 3 分析手法

本節では、どのようなエンティティ記述が ZSEL に有効かを分析するための手法について説明する。まず、分析の対象とする EL と ZSEL を定義する。そして、本研究では複数の観点から分析を行うため、観点間で共通している分析手法の概要を説明した後、個々の観点ごとの分析手法の詳細を説明していく。

### 3.1 エンティティリンクングの定義

EL は、エンティティ集合  $E$  と、文章中においてエンティティ  $e \in E$  を言及している部分（これをメンションと呼び、 $m$  で表す）が与えられたときに、 $m$  が  $E$  内のどのエンティティを言及しているかを推定するタスクである。

ZSEL は、Logeswaran らによって提案された、2つの制約の下で行われる EL である [3]。まず、エンティティに関するリソースとして与えられるのは、エンティティ  $e$  とその記述  $d$  のペアの集合であるエンティティ辞書  $\mathcal{E} = \{(e_i, d_i)\}_{i=1, \dots, K}$  のみである ( $K$  はエンティティ数)。つまり、エンティティに関する構造化された情報が保存されている知識グラフ（以降 KG）等があったとしても、それは ZSEL では使用できない。2つ目として、訓練セット内のエンティティ集合  $\mathcal{E}_{\text{train}}$  とテストセット内のエンティティ集合  $\mathcal{E}_{\text{test}}$  は交わりを持たない ( $\mathcal{E}_{\text{train}} \cap \mathcal{E}_{\text{test}} = \emptyset$ )。

### 3.2 各観点間で共通な分析手法

本研究では、どのようなエンティティ記述が ZSEL に有効かを分析するということを、エンティティ記述内のどの部分が ZSEL の正解率向上に寄与しているかを明らかにすることと捉える。この目的に対して本研究では、メンションとエンティティ記述集合が与えられたときに、そのメンションが言及しているエンティティ（つまり正解エンティティ）のエンティティ記述の一部をマスクあるいは置換して、ZSEL の正解率の変化を確認するという手法を用いる。なお、本研究においてマスクするとは、特殊なトークンで置換することを指す。

以降、上述の手法をより詳細に説明する。マスク・置換対象  $T$  を数式で表すと、次式のようになる：

$$T = \{e \in E_{d_c} \mid s_e \geq \theta\} \quad (1)$$

$T$  は、正解エンティティ  $c$  のエンティティ記述  $d_c$  内に出現しているエンティティの集合  $E_{d_c}$  の各要素の内、対応するスコア  $s_e$  が閾値  $\theta$  以上のエンティティである。具体的な  $s_e$  と  $\theta$  は観点ごとに異なるため、次節以降で説明する。 $T$  のマスク・置換が終わったら、メンション集合  $M$  とエンティティ集合  $\mathcal{E}$ 、マスク・置換済みの正解エンティティ記述を含むエンティティ記述集合を用い、各メンション  $m \in M$  に対して ZSEL を行って正解率を求める。これを各観点ごとに定めた複数のパターンで行い、正解率が最も低かった（置換の場合は最も高かった）パターンのエンティティ記述が ZSEL の正解率向上に寄与していると考えられる。

### 3.3 観点 1：エンティティ間の関係

本節では、どのようなエンティティ記述が EL に有効かを、記述対象のエンティティと特定の関係にあるエンティティという観点で分析する手法について説明していく。

本観点で分析を行う理由は、次の通りである。最も有名なエンティティ記述集合の例の 1 つである Wikipedia がそうであるように、エンティティ記述は基本的に、記述対象であるエンティティと他のエンティティの間に存在する関係を記述する。よって、既述対象のエンティティとどのような関係にあるエンティティが ZSEL の正解率向上に寄与しているかを明らかにできれば、ZSEL に有効なエンティティ記述を作成する際の有用な指針となると考えられる。

これを明らかにするために、3.2 節の手法を用いて分析を行う。後述の各パターン間でマスクするトークン量が同一になるように正解エンティティと特定の関係にあるエンティティをマスクするために、まず式 1 内の  $s_e$  を次のようにする：

$$s_e =: g(c, e, r) \quad (2)$$

上式の  $g$  は 2 つのエンティティ  $c$  と  $e$ 、関係  $r$  を入力すると、両エンティティ間にその関係が存在する度合い（スコア）を出力する関数である。続いて、式 1 内の  $\theta$  は、正解エンティティ集合  $C$  内の各要素とそのエンティティ記述内に出現しているエンティティ間に  $r$  が存在するスコアの集合  $S_r = \{g(c', e', r) \mid c' \in C, e' \in E_{d_{c'}}\}$  において、そのスコア以上のスコアを持つエンティティの総トークン数が正解エンティティ記述集合  $D_C = \{d_c \mid c \in C\}$  の総トークン数の  $p\%$  以上となるようなスコアの内、最大のスコアとする。そして、本観点では、 $r$  として関係集合  $R$  内の各関係を用いたパターンの ZSEL の正解率を比較する。

以上が本観点での分析の詳細であるが、最後に、何を  $g$  として用いるかを述べる。1 つ目は KG である。なお、本手法における KG の使用は直接的に ZSEL の正解率を上げるためではなく、あくまでエンティティ記述内において記述対象のエンティティとどのような関係にあるエンティティが ZSEL の正解率向上に寄与しているかを明らかにするためであるので、ZSEL の制約に反することにはならない。さらに、KG には含まれていないが実際には 2 つのエンティティ間に存在する関係を推定するために、知識グラフ埋め込みモデル (KGE モデル) を用いる。

### 3.4 観点 2：エンティティの出現回数

本節では、どのようなエンティティ記述が EL に有効かを、エンティティ記述内に出現しているエンティティの、記述対象のエンティティと同タイプのエンティティ記述内での出現回数という観点で分析する手法について説明していく。

本観点で分析を行う理由は、次の通りである。3.3 節で述べたように、エンティティ記述は基本的に、記述対象であるエンティティと他のエンティティとの間に存在する関係を記述する。そして、エンティティによっては、複数のエンティティとの間に、ある同一の関係が存在する場合がある。このようなエンティティのエンティティ記述を作成する際に、その複数のエンティティの中からどれをエンティティ記述に含めるべきかという情報は、エンティティ記述を作成する際の有用な指針となると考えられる。そして、エンティティ記述に含めるべきエンティティを選ぶ際の基準として、本手法では記述対象のエンティティと同タイプのエンティティ記述内での出現頻度に着目する。その理由は、次の通りである。ZSEL モデルにおいて多く採用されるバイエンコーダは、一般的には類似エンティティの距離が近くなるようにエンティティをベクトル空間に埋め込む [12, 18, 19]。よって、同タイプのエンティティは ZSEL において正解エンティティと間違われる可能性が高く、それらを弁別するためにエンティティ記述内に出現するエンティティの差異が使われていると考えられるためである。

これを明らかにするために、3.2 節の手法を用いて分析を行う。後述の各パターン間でマスクするトークン量が同一になるように正解エンティティと同タイプのエンティティ記述内での出現回数に応じてエンティティをマスクするために、まず式 1 内の  $s_e$  を次のようにする：

$$s_e =: \text{frq}(e, D_{t_c}) \quad (3)$$

上式の  $\text{frq}$  は、エンティティ  $e$  と、 $c$  が属するタイプ  $t_c$  とタイプが同じであるエンティティのエンティティ記述集合  $D_{t_c} = \{d_{e'} \mid e' \in \mathcal{E}, t_{e'} = t_c\}$  を入力すると、 $e$  の  $D_{t_c}$  内での出現回数を出力する関数である。続いて、式 1 内の  $\theta$  は、正解エンティティのエンティティ記述内に出現している各エンティティの  $D_{t_c}$  内の出現回数の集合  $F_c = \{\text{frq}(e', D_{t_c}) \mid e' \in E_{d_c}\}$  において、その出現回数以上の出現回数を持つエンティティの総トークン数がエンティティ記述  $d_c$  の総トークン数の  $p\%$  以上となるような出現回数の内、最大の出現回数とする。なお、人間のエンティティ記述には父親が出現する可能性があるが、映画の場合はその可能性がないなど、正解エンティティと特定の関係にあるエンティティが出現するかは正解エンティティ記述ごとにまちまちである。よって、3.3 節の手法においては正解エンティティと特定の関係にあるエンティティをより多くマスクするために全正解エンティティ記述を横断してマスク対象のエンティティを決めたが、本手法ではその必要はないため、各正解エンティティ記述ごとにマスク対象のエンティティを決める。そして、本観点では、出現回数が多いエンティティと少ないエンティティという 2 パターンのエンティティをマスクした際の ZSEL の正解率を比較する。これまで説明してきたのは出

現回数が多いエンティティをマスクする際の手法であり、少ないエンティティをマスクする際は反対の手法を用いる。

### 3.5 観点3：エイリアスの利用頻度

本節では、どのようなエンティティ記述が EL に有効かを、エンティティ記述内に出現しているエンティティを表すエイリアス（別名）の利用頻度という観点で分析する手法について説明していく。

本観点で分析を行う理由は、次の通りである。これまでに述べたように、エンティティ記述はエンティティを含んでいる。そして、エンティティを含めるためには、当然のことながらその名前をエンティティ記述内に書く必要があるが、複数の名前、すなわちエイリアスを持っているエンティティが存在する。このようなエンティティは珍しくないため、そのエイリアスの中のどれをエンティティ記述内で使用すれば ZSEL の正解率が向上するかを明らかにできれば、ZSEL に有効なエンティティ記述を作成する際の有用な指針となると考えられる。そして、どのエイリアスを使うべきかの基準として、本手法ではエイリアスの利用頻度に着目する。これは、エイリアスが使われ始めた時期など他の基準もあり得るが、それらも結局のところ利用頻度と相関がある場合が多いと考えられるためである。

これを明らかにするために、3.2 節の手法を用いて分析を行う。なお、前の 2 つの観点ではエンティティ記述内のエンティティを置換して分析を行うことが難しくマスクで代替したが、本観点ではそれが可能であるため、より実際の状況に近くなるようにエンティティを置換して分析を行う。後述するパターンでの分析を行うために、利用頻度が異なる複数のエイリアスを持っているエンティティを置換対象  $T$  とする。具体的には、式 1 内の  $s_e$  と  $\theta$  を次のようにする：

$$s_e =: \#\{\text{utl}(a_e) \mid a_e \in A_e\}, \quad \theta =: 2 \quad (4)$$

上式の  $A_e$  はエンティティ  $e$  のエイリアス集合を指しており、 $\text{utl}$  は引数として渡されたエイリアスの利用頻度を出力する関数である。そして、本観点では、 $T$  内の各エンティティを以下の 3 パターンの異なる利用頻度のエイリアスで置換した際の ZSEL の正解率を比較する。(1)  $A_e$  内で最も利用頻度が高いエイリアス、(2)  $A_e$  内で最も利用頻度が低いエイリアス、(3)  $A_e$  内で利用頻度が中程度 ( $\text{utl}(a_e)$  が中央値) のエイリアス。

## 4 実 験

本節では、エンティティ記述内のどの部分が ZSEL の正解率向上に寄与しているかを明らかにするために行った分析の実験について説明する。まず、本実験で共通して使用したデータセットと EL モデルについて説明した後、3 節で述べた 3 つの観点による分析の実験についてそれぞれ述べる。

### 4.1 データセット

本研究では、WikilinksNED Unseen-Mentions [6] を、3.1 節で述べた ZSEL の制約を満たすように改変して得られたデータセット（以降、WikilinksNED UM ZS と呼ぶ）を使用した。

改変元として WikilinksNED Unseen-Mentions を選んだ理由は、以下の通りである。当該データセットは、950 万件の Web ページを対象に、Wikipedia ページをリンクしているアンカーテキスト（メンション）、リンク先の Wikipedia ページ（エンティティ）、アンカーテキストの周りの文章（メンションの文脈）を収集して構築された Wikilinks コーパス [20] を基にしている。よって、当該データセットのメンションとその周囲の文脈は、個人によるブログやテレビドラマのセリフなど多様な文章であり、より実際の例に近いと考えられる。

続いて、WikilinksNED Unseen-Mentions をどのように改変して WikilinksNED UM ZS を作成したのかを説明する。Wu らも述べているように [5]、WikilinksNED Unseen-Mentions はテストセットと検証セットの両方に、訓練セット内のメンションと言及しているエンティティが同一であるメンションを含んでいる。よって、そのままでは ZSEL の制約を満たすことができない。そこで、テストセットと検証セットから、上述の訓練セット内のメンションと言及しているエンティティが同一であるメンションを取り除くことにより、ZSEL の制約を満たすことができるデータセットである WikilinksNED UM ZS を作成した。WikilinksNED UM ZS のテストセットと検証セットのメンション数は、それぞれ 7,493 件と 6,872 件である。

### 4.2 エンティティリンキングモデル

本研究では、ZSEL モデル「BLINK」[5] を使用した。その理由は、現在の EL モデルでは一般的に採用されているバイエンコーダアーキテクチャ [12, 18, 19] を最初に採用したモデルの 1 つであり、かつ多くの EL 研究においてベースラインとして用いられているためである [13, 21, 22]。なお、前述の通りバイエンコーダを採用している EL モデルは多いが、それと比較してクロスエンコーダを採用している EL モデルは少ないため、本研究では BLINK のバイエンコーダのみを使用して分析を行った。そして、ZSEL 設定における EL モデルの性能を分析するために、上述の WikilinksNED UM ZS を用いて訓練した BLINK を実験に使用した。訓練時のハイパーパラメータを表 1 で示し、訓練後のモデルの WikilinksNED Unseen-Mentions における性能を表 2 で示す。なお、エンティティ（記述）集合として、4.3.1 節で述べるエンティティカタログを使用したため、10,000 個のテストデータの内、メンションが言及しているエンティティが当該カタログに含まれていない 12 個のデータは取り除いた。また、クロスエンコーダ使用の有無などの条件が異なるため正確な比較はできないが、参考のため Wu らによる BLINK の性能も同表に記載する [5]。

### 4.3 関係に応じてエンティティをマスク

本節では、エンティティ記述内において、記述対象のエンティティとどのような関係にあるエンティティが ZSEL の正解率向上に寄与しているかを明らかにするために、3.3 節で述べた手法に沿って行った分析の実験について説明する。順番としては、実験設定、実験結果の順で述べる。

表 1 BLINK 訓練時のハイパーパラメータ.

パラメータ	値
BERT モデル	bert-large-uncased
損失関数	Cross Entropy Loss
オプティマイザ	Adam
学習率	$2 \times 10^{-6}$
エポック数	1
バッチサイズ	128
エンティティ記述の上限トークン数	128
メンションの文脈の上限トークン数	32

表 2 WikilinksNED Unseen-Mentions における BLINK の性能.

BLINK	Accuracy	Recall@10
本実験のために訓練	0.475	0.868
Wu ら (2020)	0.747	-

### 4.3.1 実験設定

ZSEL モデルとしては, 4.2 節で述べた BLINK を用いた.

エンティティ (記述) 集合としては, BLINK の公式 GitHub リポジトリにおいて公開されているエンティティカタログ<sup>1</sup>を使用した. このカタログは 2019 年 8 月 1 日の Wikipedia ダンプから作られており 5,903,527 個のエンティティを含んでいるが, 本実験では 3.1 節で述べた ZSEL の制約を満たすために, BLINK の訓練に用いた WikilinksNED UM ZS の訓練セット内のメンションが言及しているエンティティを除いた 5,817,551 個のエンティティのみを使用した.

ZSEL を行う対象のメンション集合としては, 4.1 節で述べた WikilinksNED UM ZS のテストセットの一部を用いた. 具体的には, テストセットの中から, 言及しているエンティティが上述の Wikipedia ダンプに含まれていないメンションと後述の KGE モデルの訓練に用いた Wikidata5m データセット内に含まれていないメンションを除いた 6,778 個を使用した.

評価指標としては, EL において最も重要だと考えられる正解率 (Accuracy) を主に用いるが, 同時に報告されることが多い Recall@10 も報告する.

関係集合としては, 後述の KGE モデルに合わせて Wikidata5m データセットに含まれている 822 個の関係の一部を用いた. 具体的には, 822 個の関係の中から, Wikidata ダンプを用いて取得した, エンティティ記述内に出現しているエンティティと記述対象のエンティティ間の関係として出現頻度が高い上位 100 個の関係のみを使用した. すべての関係を使わなかった理由は, 出現頻度が低い関係は訓練データが少なく KGE モデルが適切な学習を行えないことが考えられ, その結果モデルが誤った推論をしてしまい誤ったエンティティがマスクされてしまう可能性を低くするためである.

KGE モデルとしては, 「NodePiece」[23] を, Wikipedia から作られた KG である Wikidata が基となっている Wikidata5m データセット [24] で訓練して使用した. NodePiece を使用した理由としては次の 2 つがある. 1 つ目は, Wikidata5m のような数百万個のエンティティを含む巨大な KG を埋め込むのに適

表 3 NodePiece 訓練時のハイパーパラメータ.

パラメータ	値
スコア付け関数	RotatE
損失関数	Margin Ranking Loss
オプティマイザ	Adam
学習率	$1 \times 10^{-3}$
エポック数	5
バッチサイズ	256
関係文脈数	12
埋め込み次元数	200
負例数	1

していることである. 2 つ目は, Wikidata5m もそうであるが, エンティティ間の関係の種類が豊富である KG に対する関係推定タスクで比較的高い性能を発揮したと報告されていることである. 訓練時のハイパーパラメータを表 3 に示す.

閾値  $\theta$  を決める際のマスク割合  $p$  は 15% とした. これは, マスクされる文章量が少ないと ZSEL の正解率が下がらず, エンティティ記述内において ZSEL の正解率向上に最も寄与しているエンティティと記述対象のエンティティとの間にある関係を確認するのが難しいためである.

3.3 節で述べた手法ではエンティティ記述内でマスクするのはエンティティのみであったが, 本実験ではエンティティに加えて, エンティティと係り受け関係にある単語<sup>2</sup>もマスクした. これは, マスクを行った後のエンティティ記述をより自然な文章にするためである.

### 4.3.2 実験結果

正解エンティティ記述内に含まれている, 正解エンティティと特定の関係にあるエンティティをマスクした状態で行った ZSEL の結果を表 4 に示す. 表の各行は ZSEL の結果を示しており, 左から, (1) Wikidata 内での関係の ID, (2) 正解エンティティとどの関係にあるエンティティをマスクしたか, あるいはマスクしていないか, (3) 正解率, (4) recall@10 を表している. なお, ここでは 4.3.1 節で述べた 100 個の関係の内, 対応する正解率の昇順で上から 10 個の関係における結果のみ示す. また, 「†」は, Holm 補正を用いて Wilcoxon signed-rank test を行ったところ, 対応する正解率の昇順に並べたときに 100 個の関係中 50 番目に位置する関係 「discoverer or inventor (P61)」の正解率よりも  $\alpha = 0.05$  で有意に低かった正解率に付している. そして, 「†」が付された正解率に対応する関係は, 半数の 50 個以上の関係よりも正解率が有意に低かった.

上の結果から, 正解エンティティ記述内に含まれている, 正解エンティティと 「instance of (P31)」関係にあるエンティティを, 正解エンティティ記述集合全体でのマスク割合が 15% になるように各正解エンティティ記述からマスクした時の ZSEL の正解率は, 他の関係にあるエンティティをマスクした時よりも低い場合が多いということが読み取れる. よって, 出現頻度で上位 100 位以内に入らない使用頻度が低い関係は ZSEL におい

1: <http://dl.fbaipublicfiles.com/BLINK/entity.jsonl>

2: より厳密には, dependency parse tree においてエンティティが依存している単語.

表4 正解エンティティ記述内の正解エンティティと特定の関係にあるエンティティをマスクした ZSEL 結果.

Wikidata ID	関係名	Accuracy	Recall@10
P31	instance of	0.3200†	0.8138
P279	subclass of	0.3265	0.8020
P527	has part(s)	0.3272	0.8163
P101	field of work	0.3277	0.8150
P641	sport	0.3280	0.8135
P407	language of work or name	0.3287	0.8157
P460	said to be the same as	0.3292	0.8169
P171	parent taxon	0.3294	0.8176
P136	genre	0.3299	0.8187
P355	has subsidiary	0.3302	0.8203
	マスクなし	0.5003	0.8870

でも重要性が低いと考えられることを踏まえて、上の結果より、エンティティ記述を作成する際に、記述対象のエンティティと「instance of」関係にあるようなエンティティを含むことは、他の関係にあるエンティティを代わりに含めるよりも ZSEL の正解率を向上させる場合が多いと考えられる。

また、上と同様の実験を、言及しているエンティティのタイプによって分割したメンション集合に対しても行った。これは、ZSEL に有効なエンティティ記述を作成する際のより有用な指針を得ることを目的として、あるタイプのエンティティ記述内において、記述対象のエンティティとどのような関係にあるエンティティが ZSEL の正解率向上に寄与しているかを明らかにするためである。本研究では「関根の拡張固有表現階層」<sup>3</sup>を参考に、表5に示す11個のタイプを用いた。なお、1つのエンティティに複数のタイプが割り当てられることもあったため、メンション数の合計は6,778にはならない。エンティティにタイプを割り当てる方法としては、Wikidata をグラフとして見て、あるエンティティ(ノード)から出発し、「instance of」または「subclass of」エッジ上を移動した(すなわち上位概念であるようなエンティティを収集した)ときに、最も距離が近いタイプ集合をそのエンティティが属するタイプであると見做した。これは、関根の拡張固有表現階層は階層構造になっており、その中でより具体性の高いタイプを割り当てるためである。

以降、実験結果を示す。なお、ここでは比較的多くの ZSEL の結果間に有意差が見られたタイプにおける結果のみ紹介する。

human タイプにおける結果を表6に示す。表4と同様に各行は ZSEL の結果を示しており、「†」は100個の関係中50番目に位置する関係「filming location (P915)」の正解率よりも有意に低かった正解率に付している。そして、「†」が付された正解率に対応する関係は、半数の50個以上の関係よりも正解率が有意に低かった。また、「\*」は、表4で結果を示した実験において、他関係よりも有意に正解率が低い場合が最も多かった関係「instance of」の正解率よりも有意に低かった正解率に付している。

上の結果から、正解エンティティ記述内に含まれている、正

表5 使用したエンティティのタイプ.

タイプ名	Wikidata ID	拡張固有表現	メンション数
human	Q5	人名	531
deity	Q178885	神名	83
organization	Q43229	組織名	619
geographical object	Q618123	地名	928
construction	Q811430	施設名	103
goods	Q28877	プロダクト名	1,180
occurrence	Q1190554	イベント名	401
natural physical object	Q16686022	自然物名	141
disease	Q12136	病名	29
color	Q1075	色名	2
concept	Q151885	概念	2,904

解エンティティと「member of sports team (P54)」関係にあるエンティティと「employer (P108)」関係にあるエンティティそれぞれを、各正解エンティティ記述からマスクした時の ZSEL の正解率は、他の関係にあるエンティティをマスクした時よりも低い場合が多いということが読み取れる。前者について、スポーツチームに加入している人は全体のごく一部であるはずにも関わらずこのような結果になった理由を探るために、どのようなエンティティがマスクされたかを確認した。その結果、スポーツチームに加えて、(1) 記述対象の人物が所属している組織、(2) 受賞した賞や発明品、リリースした楽曲などの記述対象の人物による功績がマスクされている例が比較的多く見られた。これらのエンティティがマスクされた理由について、(1)には記述対象の人物の所属先、(2)には記述対象の人物の貢献対象あるいは功績により得たものと、それぞれスポーツチームとの共通点があるため、記述対象のエンティティとこれらのエンティティ間にはスポーツチームとの間にあるのと同様の関係があると KGE モデルが判断した可能性があり得る。

よって、human タイプのエンティティ記述を作成する際に、記述対象の人物が所属している組織や功績を含むことは、他の関係にあるエンティティを代わりに含めるよりも、エンティティ集合の要素が human タイプのエンティティに限定された ZSEL の正解率を向上させる場合が多いと考えられる。なお、表6内の「\*」で示したように、関係「instance of」よりも有意に正解率が低い関係が存在するため、上述の限定された ZSEL の正解率を向上させるエンティティ記述を作成する際には、記述対象のエンティティと「instance of」関係にあるようなエンティティを含むだけでは不十分だと考えられる。

goods タイプにおける結果を表7に示す。表4と同様に各行は ZSEL の結果を示しており、「†」は100個の関係中50番目に位置する関係「occupation (P106)」の正解率よりも有意に低かった正解率に付している。そして、「†」が付された正解率に対応する関係は、半数の50個以上の関係よりも正解率が有意に低かった。また、「\*」は表6と同様である。

上の結果から、正解エンティティ記述内に含まれている、正解エンティティと表7内で「†」を付した7つの関係にあるエンティティを、各正解エンティティ記述からマスクした時の ZSEL の正解率は、他の関係にあるエンティティをマスクした時より

3: [http://liat-aip.sakura.ne.jp/ene/ene9/definition\\_jp/html/enetree.html](http://liat-aip.sakura.ne.jp/ene/ene9/definition_jp/html/enetree.html)(参照 2023-12-15).

表 6 正解エンティティ記述内の正解エンティティと特定の関係にあるエンティティをマスクした ZSEL 結果 (human タイプ).

Wikidata ID	関係名	Accuracy	Recall@10
P54	member of sports team	0.3239*†	0.7552
P108	employer	0.3239*†	0.7533
P69	educated at	0.3333*	0.7571
P22	father	0.3371	0.7439
P102	member of political party	0.3371	0.7589
P172	ethnic group	0.3371*	0.7439
P19	place of birth	0.3390	0.7627
P1344	participant in	0.3390	0.7495
P27	country of citizenship	0.3409	0.7665
P1830	owner of	0.3409	0.7533
	マスクなし	0.4689	0.8098

表 7 正解エンティティ記述内の正解エンティティと特定の関係にあるエンティティをマスクした ZSEL 結果 (goods タイプ).

Wikidata ID	関係名	Accuracy	Recall@10
P161	cast member	0.2983*†	0.8042
P1877	after a work by	0.3068*†	0.8458
P58	screenwriter	0.3076*	0.8390
P162	producer	0.3076*†	0.8364
P921	main subject	0.3085*†	0.8449
P86	composer	0.3127	0.8483
P674	characters	0.3127*†	0.8542
P449	original broadcaster	0.3136*†	0.8508
P840	narrative location	0.3153*†	0.8508
P725	voice actor	0.3153	0.8441
	マスクなし	0.5153	0.9339

も低い場合が多いということが読み取れる。よって, goods タイプのエンティティ記述を作成する際に, 記述対象のエンティティと上の 7 つの関係にあるようなエンティティを含むことは, 他の関係にあるエンティティを代わりに含めるよりも, エンティティ集合の要素が goods タイプのエンティティに限定された ZSEL の正解率を向上させる場合が多いと考えられる。なお, 表 7 内の「\*」で示したように, 関係「instance of」よりも有意に正解率が低い関係が存在するため, 上述の限定された ZSEL の正解率を向上させるエンティティ記述を作成する際には, 記述対象のエンティティと「instance of」関係にあるようなエンティティを含むだけでは不十分だと考えられる。

#### 4.4 出現回数に応じてエンティティをマスク

本節では, エンティティ記述内において, 記述対象のエンティティと同タイプのエンティティ記述内で出現回数が多いエンティティと少ないエンティティのどちらが ZSEL の正解率向上に寄与しているかを明らかにするために, 3.4 節で述べた手法に沿って行った分析の実験について説明する。順番としては, 実験設定, 実験結果の順で述べる。

##### 4.4.1 実験設定

使用した ZSEL モデル, エンティティ (記述) 集合, ZSEL 対象のメンション集合, 評価指標は 4.3.1 節と同様である。

正解エンティティが属するタイプ  $t_c$  としては, 4.3.2 節で述

表 8 正解エンティティ記述内のエンティティを記述対象のエンティティと同タイプのエンティティ記述内における出現回数に応じてマスクした ZSEL 結果.

出現回数	マスク割合	Accuracy	Recall@10
多い	19.88%	0.2992	0.8008
少ない	20.06%	0.3126	0.8103
マスクなし	0%	0.5003	0.8870

べた 11 個のタイプを用いた。また, 記述対象のエンティティ (正解エンティティ) と同タイプのエンティティ記述内におけるエンティティの出現回数  $\text{frq}(e, D_{t_c})$  は, Wikipedia ダンプ内のエンティティ記述として使用されている箇所に含まれているハイパーリンクのリンク先を確認することによって数えた。

4.3.1 節と同様に分析が行いやすくなるように, 閾値  $\theta$  を決める際のマスク割合  $p$  は 20% とした。なお, 3.4 節での手法の説明でも触れたように, 本実験で用いる手法では全正解エンティティ記述を横断してマスク対象のエンティティを決めるのではなく, 各正解エンティティ記述ごとにマスク対象のエンティティを決めるため, マスク割合の平均は完全に 20% にはならない。また, 4.3.1 節と同様に, エンティティ記述内においてエンティティに加えてエンティティと係り受け関係にある単語もマスクした。

##### 4.4.2 実験結果

正解エンティティ記述内に含まれている, 正解エンティティと同タイプのエンティティ記述内で出現回数が (1) 多いエンティティと (2) 少ないエンティティの 2 種類のエンティティをそれぞれマスクした状態で行った ZSEL の結果を表 8 に示す。表の各行は ZSEL の結果を示しており, 左から, (1) 正解エンティティと同タイプのエンティティ記述内で多少どちらの出現回数のエンティティをマスクしたか, あるいはマスクしていないか, (2) 正解エンティティ記述内のマスクされたトークン割合の平均, (3) 正解率, (4) recall@10 を表している。なお, ZSEL の結果ペアに対して, 4.3.2 節と同様の検定を行ったところ, どの結果ペアについても  $\alpha = 0.05$  で有意差が見られた。

上の結果から, 正解エンティティ記述内に含まれている, 正解エンティティと同タイプのエンティティ記述内で出現回数が多いエンティティを, 平均マスク割合が約 20% になるように各正解エンティティ記述からマスクした際の ZSEL の正解率は, 出現回数が少ないエンティティをマスクした時よりも低いということが読み取れる。よって, エンティティ記述を作成する際に, 現状のエンティティ記述の文章量がある一定以上まで増加するように, 記述対象のエンティティと同タイプのエンティティを説明するときによく使われるようなエンティティを含めると, 使われることが少ないエンティティを含めたときよりも ZSEL の正解率を向上させることができると考えられる。

なお, 上の結果に関連して, Scialvolino らは, BERT [25] ベースのエンコーダを用いた情報検索手法である DPR は Wikipedia 内での出現頻度が低い稀なエンティティを上手く検索できないことを指摘している [26]。本実験で用いた BLINK も BERT ベースのエンコーダを採用しているため, これと同様の現象

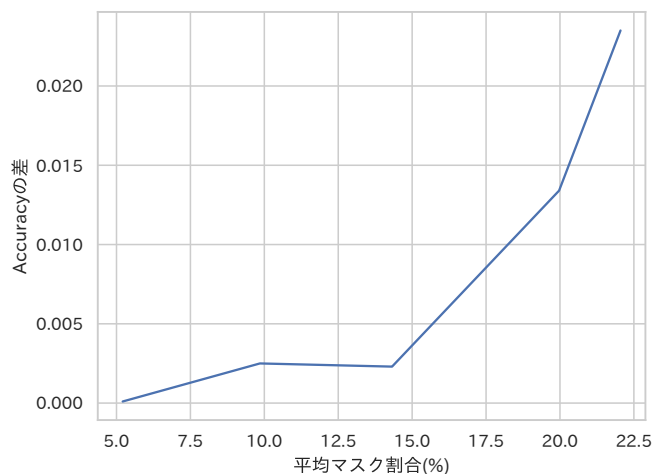


図1 出現回数の多少間の Accuracy の差.

が起こり、上のような結果が得られたと思われる。この結果の理由として、BERTの事前学習に使われたコーパスの1つがWikipediaであるため、その中で出現頻度が低い単語（エンティティ）は上手く埋め込むことができなかった可能性がある。

また、平均マスク割合を変えたときの、正解エンティティと同タイプのエンティティ記述内で出現回数が多いエンティティと少ないエンティティの2種類のエンティティをそれぞれマスクした状態で行ったZSELの正解率の差を図1に示す。

#### 4.5 異なる利用頻度のエイリアスで置換

本節では、エンティティ記述内に出現するエンティティについて、どの利用頻度のエイリアスがZSELの正解率向上に寄与しているかを明らかにするために、3.5節で述べた手法に沿って行った分析の実験について説明する。順番としては、実験設定、実験結果の順で述べる。

##### 4.5.1 実験設定

使用したZSELモデル、エンティティ（記述）集合、ZSEL対象のメンション集合、評価指標は4.3.1節と同様である。

式4内のエイリアス集合  $A_e$  は、2019年1月28日のWiki-data ダンプを用いて収集した。具体的には、ダンプ内の各エンティティについて、“labels”と“aliases”をエイリアスとして使用した。なお、本実験で用いたZSELモデル内のBERTは英語用なので、それに合わせて得られたエイリアスの中で英語でないと思われるものは除外した。具体的には、Unicodeの「基本ラテン文字」～「ラテン文字拡張B」ブロック外の文字が含まれている、またはBERTのトークナイザに入力して得られたトークン列に[UNK]トークンが含まれるエイリアスは除外した。また、式4内のエイリアスの利用頻度  $utl$  としては、実際の社会における利用頻度をある程度反映していると考えられるため、2019年8月1日のWikipediaダンプ内におけるエイリアスの出現回数を用いた。

##### 4.5.2 実験結果

正解エンティティ記述内のエンティティを、そのエンティティの(1)利用頻度が高いエイリアス、(2)利用頻度が中程度のエイリアス、(3)利用頻度が低いエイリアスでそれぞれ置換した

表9 正解エンティティ記述内のエンティティを異なる利用頻度のエイリアスで置換したZSEL結果.

利用頻度	Accuracy	Recall@10
高	0.4951	0.8752
中	0.5025	0.8793
低	0.4879	0.8690
置換なし	0.5003	0.8870

状態で行ったZSELの結果を表9に示す。表の各行はZSELの結果を示しており、左から、(1)正解エンティティ記述内のエンティティをどの利用頻度のエイリアスで置換したか、あるいは置換していないか、(2)正解率、(3)recall@10を表している。なお、ZSELの結果ペアに対して、4.3.2節と同様の検定を行ったところ、(1)「利用頻度：中」と「利用頻度：低」、(2)「利用頻度：低」と「置換なし」の2組のペアについては $\alpha = 0.05$ で有意差が見られた。

上の結果から、正解エンティティ記述内のエンティティを、そのエンティティの利用頻度が低いエイリアスで置換した際のZSELの正解率は、それ以外の利用頻度のエイリアスで置換した（あるいは置換しなかった）際の正解率よりも低い場合が多いということが読み取れる。よって、ZSELの正解率を向上させるエンティティ記述を作成する際には、その中に含めるエンティティのエイリアスとして利用頻度が低いエイリアスは使うべきでない場合が多いと考えられる。

また、上と同様の実験を4.3.2節と同じく、言及しているエンティティのタイプ(11個)によって分割したメンション集合に対しても行った。ここではタイプで分割しないメンション集合に対して行った実験と異なる傾向の結果が見られたgeographical objectタイプにおける結果のみ紹介する。

geographical objectタイプにおける結果を表10に示す。本表内のZSELの結果ペアに関して、(1)「利用頻度：高」と「利用頻度：中」、(2)「利用頻度：高」と「利用頻度：低」、(3)「利用頻度：中」と「置換なし」、(4)「利用頻度：低」と「置換なし」の4組のペアでは $\alpha = 0.05$ で有意差が見られた。

この結果からは、先ほどとは反対に、正解エンティティ記述内のエンティティを、そのエンティティの利用頻度が低いエイリアスで置換した際のZSELの正解率は、それ以外の利用頻度のエイリアスで置換した（あるいは置換しなかった）際の正解率よりも高い場合が多いということが読み取れる。

上のような結果が得られた理由を探るために、どのようなエイリアスに置換されたかを確認した。その結果、「London」が「London, United Kingdom」に置換されるなど、正解エンティティ記述内に出現する地名エンティティが、その地名が属する国名や州名を含むエイリアスに置換されている例が多く見られた。このことから、geographical objectタイプに属するエンティティのエンティティ記述内にはそのエンティティと同じ国や州などに位置する地名エンティティが多く含まれており、それらがエイリアスへ置換されることによって位置する国や州などに関する情報が増え、ZSELが上手く行えるようになったということが考えられる。



表 10 正解エンティティ記述内のエンティティを異なる利用頻度のエリアスで置換した ZSEL 結果 (geographical object タイプ).

利用頻度	Accuracy	Recall@10
高	0.4375	0.9256
中	0.4731	0.9256
低	0.4806	0.9267
置換なし	0.4267	0.9224

よって、エンティティ集合の要素が geographical object タイプのエンティティに限定された ZSEL の正解率を向上させるエンティティ記述を作成する際に、その中に同じ地域に位置するエンティティを含める場合は、そのエンティティの名前に加えて位置する地域も記載するべきだと考えられる。

## 5 ま と め

本研究では、ZSEL に有効なエンティティ記述を作成する際の指針を得ることを目的として、エンティティ記述内のどの部分が ZSEL の正解率向上に寄与しているかを明らかにするための分析を行った。分析のために、メンションが言及しているエンティティのエンティティ記述内の一部をマスクあるいは置換して、その結果 ZSEL の正解率がどのように変化するかを確認するという手法を用いた。分析の結果、エンティティ間の関係、エンティティの出現回数、エリアスの利用頻度が ZSEL の正解率に影響していることが示唆された。

今後の研究課題としては、ZSEL に有効なエンティティ記述に関する更なる分析や、ZSEL に有効なエンティティ記述と人間が読みやすいエンティティ記述の両立が挙げられる。

**謝辞** 本研究は JSPS 科研費 JP21H03775, JP23H03400 の助成を受けたものです。ここに記して謝意を表します。

## 文 献

- [1] Fengbin Zhu, Wenqiang Lei, Chao Wang, Jianming Zheng, Soujanya Poria, and Tat-Seng Chua. Retrieving and reading: A comprehensive survey on open-domain question answering. *arXiv preprint arXiv:2101.00774*, 2021.
- [2] Mei Kobayashi and Koichi Takeda. Information retrieval on the web. *ACM computing surveys (CSUR)*, Vol. 32, No. 2, pp. 144–173, 2000.
- [3] Lajanugen Logeswaran, Ming-Wei Chang, Kenton Lee, Kristina Toutanova, Jacob Devlin, and Honglak Lee. Zero-shot entity linking by reading entity descriptions. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 3449–3460, 2019.
- [4] Tom Oberhauser, Tim Bischoff, Karl Brendel, Maluna Menke, Tobias Klatt, Amy Siu, Felix Alexander Gers, and Alexander Löser. Trainx-named entity linking with active sampling and bi-encoders. In *Proceedings of the 28th International Conference on Computational Linguistics: System Demonstrations*, pp. 64–69, 2020.
- [5] Ledell Wu, Fabio Petroni, Martin Josifoski, Sebastian Riedel, and Luke Zettlemoyer. Scalable zero-shot entity linking with dense entity retrieval. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 6397–6407, 2020.
- [6] Yasumasa Onoe and Greg Durrett. Fine-grained entity typing for domain independent entity linking. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34, pp. 8576–8583, 2020.
- [7] Ikuya Yamada, Hiroyuki Shindo, Hideaki Takeda, and Yoshiyasu Takefuji. Joint learning of the embedding of words and entities for named entity disambiguation. In *Proceedings of the 20th SIGNLL Conference on Computational Natural Language Learning*, pp. 250–259, 2016.
- [8] Octavian-Eugen Ganea and Thomas Hofmann. Deep joint entity disambiguation with local neural attention. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pp. 2619–2629, 2017.
- [9] Zhengyan He, Shujie Liu, Mu Li, Ming Zhou, Longkai Zhang, and Houfeng Wang. Learning entity representation for entity disambiguation. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pp. 30–34, 2013.
- [10] Yaming Sun, Lei Lin, Duyu Tang, Nan Yang, Zhenzhou Ji, and Xiaolong Wang. Modeling mention, context and entity with neural networks for entity disambiguation. In *Proceedings of the 24th International Conference on Artificial Intelligence*, pp. 1333–1339, 2015.
- [11] Nikolaos Kolitsas, Octavian-Eugen Ganea, and Thomas Hofmann. End-to-end neural entity linking. In *Proceedings of the 22nd Conference on Computational Natural Language Learning*, pp. 519–529, 2018.
- [12] Kai Sun, Richong Zhang, Samuel Mensah, Yongyi Mao, and Xudong Liu. A transformational biencoder with in-domain negative sampling for zero-shot entity linking. In *Findings of the Association for Computational Linguistics: ACL 2022*, pp. 1449–1458, 2022.
- [13] GP Shrivatsa Bhargav, Dinesh Khandelwal, Saswati Dana, Dinesh Garg, Pavan Kapanipathi, Salim Roukos, Alexander Gray, and L Venkata Subramaniam. Zero-shot entity linking with less data. In *Findings of the Association for Computational Linguistics: NAACL 2022*, pp. 1681–1697, 2022.
- [14] Xuhui Sui, Ying Zhang, Kehui Song, Baohang Zhou, Guoqing Zhao, Xin Wei, and Xiaojie Yuan. Improving zero-shot entity linking candidate generation with ultra-fine entity type information. In *Proceedings of the 29th International Conference on Computational Linguistics*, pp. 2429–2437, 2022.
- [15] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. "why should i trust you?" explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 1135–1144, 2016.
- [16] Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pp. 4768–4777, 2017.
- [17] Jiwei Li, Will Monroe, and Dan Jurafsky. Understanding neural networks through representation erasure. *arXiv preprint arXiv:1612.08220*, 2016.
- [18] Jan A Botha, Zifei Shan, and Dan Gillick. Entity linking in 100 languages. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 7833–7845, 2020.
- [19] Rajarshi Bhowmik, Karl Stratos, and Gerard de Melo. Fast and effective biomedical entity linking using a dual encoder. In *Proceedings of the 12th International Workshop on Health Text Mining and Information Analysis*, pp. 28–37, 2021.
- [20] Sameer Singh, Amarnag Subramanya, Fernando Pereira, and Andrew McCallum. Wikilinks: A large-scale cross-document coreference corpus labeled via links to wikipedia. 2012.
- [21] Xiaodong Yu, Wenpeng Yin, Nitish Gupta, and Dan Roth. Event linking: Grounding event mentions to wikipedia. In

*Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, pp. 2671–2680, 2023.

- [22] Belinda Z Li, Sewon Min, Srinivasan Iyer, Yashar Mehdad, and Wen-tau Yih. Efficient one-pass end-to-end entity linking for questions. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 6433–6441, 2020.
- [23] Mikhail Galkin, Etienne Denis, Jiapeng Wu, and William L Hamilton. Nodepiece: Compositional and parameter-efficient representations of large knowledge graphs. In *International Conference on Learning Representations*, 2021.
- [24] Xiaozhi Wang, Tianyu Gao, Zhaocheng Zhu, Zhengyan Zhang, Zhiyuan Liu, Juanzi Li, and Jian Tang. Kepler: A unified model for knowledge embedding and pre-trained language representation. *Transactions of the Association for Computational Linguistics*, Vol. 9, pp. 176–194, 2021.
- [25] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.
- [26] Christopher Scialolino, Zexuan Zhong, Jinhyuk Lee, and Danqi Chen. Simple entity-centric questions challenge dense retrievers. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pp. 6138–6148, 2021.