

多言語述語項構造ベクトル表現の学習

宇野 真矢^{1,a)} 柴田 知秀^{1,2,b)} 河原 大輔^{1,c)} 黒橋 禎夫^{1,2,d)}

概要：

近年、単語の意味をベクトルで表現する取り組みが広く行われており、多言語で単語ベクトルを学習する手法も提案されている。これらの手法においては、単語、特に述語の意味に曖昧性が大きいことが問題となるが、述語項構造のような単語より大きな単位で学習すれば曖昧性を解消することができる。本論文では、述語項構造のベクトル表現を言語横断的に学習する手法を提案する。この手法では、対訳コーパスを用いて多言語の単語ベクトルを学習する手法を拡張し、多言語の述語項構造ベクトルを学習する。評価実験では、学習された述語項構造ベクトルを用いて、日英中3言語で述語項構造の対応付けを行い、単語ベクトルに基づくベースラインと比べて高い精度を達成した。

1. はじめに

近年、単語の意味をコーパスから学習し、ベクトルとして表現する試みが広く行われている [1], [2], [3]。学習された単語ベクトルは構文解析、感情分析、文書分類などの様々な自然言語処理タスクに用いられ、一定の成果を挙げている。

これらの手法においては、単語、特に述語の意味に曖昧性が大きいことが問題となるが、句や述語項構造のような単語より大きな単位で学習すれば曖昧性を解消することができる。そこで、句や述語項構造、さらには文に対してベクトル表現を学習する研究が多く行われている [4], [5], [6], [7], [8], [9], [10], [11], [12]。これらの研究では基本的に、語のベクトルを合成することによって、句や文のベクトルを表現している。

上記は単言語についての研究であるが、複数言語で共通のベクトル表現を学習することができれば、言語横断情報検索、言語横断文書分類、機械翻訳などの多言語自然言語処理タスクにおいて有用である [13], [14], [15]。以下に、対訳関係にある日本語と英語の表現の例を示す。

- (1) a. 取引先と契約を 結ぶ
b. make a contract with a client
- (2) a. 鉄道が主要な都市を 結ぶ

b. the railway connects major cities

これらの日本語文には、どちらも動詞「結ぶ」が含まれている。しかし英語文中でこの動詞「結ぶ」に対応する単語は、例文 (1-b) では “make”, (2-b) では “connect” と異なる。複数の語からなる句または述語項構造に対して、多言語で共通のベクトル表現を学習し、上記の例文における「契約を結ぶ = make a contract」「都市を結ぶ = connect cities」といった対応関係を認識することができるようになれば、多言語自然言語処理タスクの精度向上に貢献できると考えられる。これまで、複数言語で共通の単語ベクトル表現を学習する手法が提案されているが、単語単位のベクトルでは「契約を結ぶ = connect a contract」のような誤った対応関係を認識してしまう可能性がある。句や述語項構造のような単位で複数言語で共通のベクトルを学習している手法は、これまでに存在しない。

本論文では、多言語述語項構造ベクトル表現を学習するモデルを提案する。本モデルは、文単位で対応付けされた対訳コーパスを用いて多言語の単語ベクトルを学習する手法 [16] を拡張し、多言語の述語項構造ベクトル表現を学習する。単言語内で、類似した述語項構造に類似したベクトル表現が割り当てられるようにし、それと同時に、言語を横断して、対訳関係にある述語項構造のベクトルが類似するようにベクトル表現を学習する。

学習された述語項構造ベクトルを用いて、日英中3言語で述語項構造の対訳を推定する実験を行い、単語ベクトルや述語項ペアベクトルに基づくベースライン手法よりも高い精度を達成した。

¹ 京都大学 大学院情報学研究所

² 科学技術振興機構 CREST

a) uno@nlp.ist.i.kyoto-u.ac.jp

b) shibata@i.kyoto-u.ac.jp

c) dk@i.kyoto-u.ac.jp

d) kuro@i.kyoto-u.ac.jp

2. 関連研究

単言語単語ベクトル表現の学習手法を応用し、多言語単語ベクトル表現を同一空間上で学習する手法が提案されている。各手法は、学習に用いる言語リソースによって以下のように分類できる。

- 多言語対訳辞書を用いる手法 [17], [18]
言語ごとに単語ベクトルを学習した後、対訳辞書を用いて、一方の言語の単語ベクトルを他方の言語の単語ベクトル空間へ写像する行列を学習する。述語項構造に関する情報はほぼ対訳辞書に記載されていないため、これらの手法は述語項構造ベクトルの学習に適用することができない。
- 単語アライメント情報が自動付与された対訳コーパスを用いる手法 [13], [19], [20], [21], [22], [23]
言語ごとに単語ベクトルを学習した後、対訳コーパスの自動単語アライメント結果を用いて各言語の単語ベクトル空間を近づける。自動アライメントの結果にはノイズが含まれており、ベクトル表現学習に悪影響を及ぼす可能性がある。
- 対訳コーパスを用いる手法 [16], [24], [25], [26], [27]
Auto Encoder を用いた方法 [24], [26]、対訳文間の自動単語アライメントの学習と単語ベクトルの学習を並行して行う方法 [25]、各言語の単語ベクトルを学習しながら、多言語間で単語ベクトルを近づける学習を同時に行う方法 [16], [27] がある。
- コンパラブルコーパスを用いる手法 [28]
文書単位で対応付けられたコンパラブルコーパスを用いて、多言語間で共通の単語ベクトルを同時に学習する。

本論文で提案する述語項構造ベクトル表現を学習するモデルは、対訳コーパスを用いる手法の一つである Gouws ら [16] の BilBOWA (Fast Bilingual Distributed Representations without Word Alignments)*¹ を基にした。BilBOWA は、2 言語の対訳コーパスおよびそれぞれの言語の単言語コーパスを用いて、多言語の単語ベクトル表現を同一のベクトル空間上で学習する。Giza++ などによる単語アライメントはコストがかかり、またノイズが含まれることから用いない。学習は各言語の単言語学習および多言語学習に分かれており、それぞれの学習を同時に行う。単言語学習は word2vec [2] と同様に行われる。BilBOWA では単語ベクトル学習モデルとして skip-gram モデル、高速化手法として Negative Sampling が用いられている。多言語学習は、単語アライメントを用いないことから単語の対応関係がわからないが、文中の単語のベクトル表現を平均し、これが対訳文同士で類似するように学習する。

*¹ <https://github.com/gouwsmeister/bilbowa>

BilBOWA を含む上記の手法は、多言語単語ベクトル表現を学習するが、述語項構造のような複数単語からなる単位でベクトル表現を学習するものではない。機械翻訳の分野では、フレーズテーブルを用いて多言語に渡る句のベクトル表現を学習する手法がある [29], [30] が、言語的構造をもつ句を対象としていないという点で本研究とは異なる。

3. 提案手法

本論文では対訳コーパスを利用して述語項構造のベクトル表現を言語横断的に学習する手法を提案する。本モデルでは、単言語内で、類似した述語項構造に類似したベクトル表現が割り当てられるようにし、それと同時に、言語を横断して、対訳関係にある述語項構造のベクトルが類似するようにベクトル表現を学習する。述語項構造のベクトル表現は述語と項のベクトルおよび述語と項の係り受け関係を基に計算する。

単言語内については、Mikolov ら [2] の提案した word2vec と同様に周辺単語との共起に基づく学習を行う。多言語については、BilBOWA にならい、文に含まれる述語項構造ベクトルの平均が対訳文の間で近づくように学習する。日英を例として、以下の対訳文に含まれる述語項構造に対する学習の概略を図 1 に示す。

- (3) a. I went to Oxford University by train and had a meeting with the professor.
b. 私はオックスフォード大学に電車で行き、教授とミーティングをした。

両言語ともに構文解析を行うと、図 1 に示すような構文木が得られる。単言語での学習では英語、日本語それぞれにおいて、述語項構造から周辺の単語が予測できるようにベクトル表現を学習する。例えば英語では述語項構造 “have meeting with professor” から周辺の単語 “go”, “Oxford”, “University”, “train” などが予測できるように、また同様に、述語項構造 “I go to University by train” から周辺の単語 “have”, “meeting” などが予測できるようにベクトル表現を学習する。日本語でも同様に、述語項構造 “教授とミーティングをする” から周辺の単語 “行く”, “大学”, “電車” などが予測できるようにベクトル表現を学習する。そして、多言語では、英語における述語項構造 “I go to University by train” のベクトルと “have meeting with professor” のベクトルの平均と、日本語における述語項構造 “私が大学に電車で行く” のベクトルと “教授とミーティングをする” のベクトルの平均が近くなるようにベクトルを学習する。

3.1 述語項構造のベクトル表現

本モデルでは Socher ら [10] や橋本ら [11] と同様に、述語項構造のベクトル表現を、述語のベクトルおよび述語に

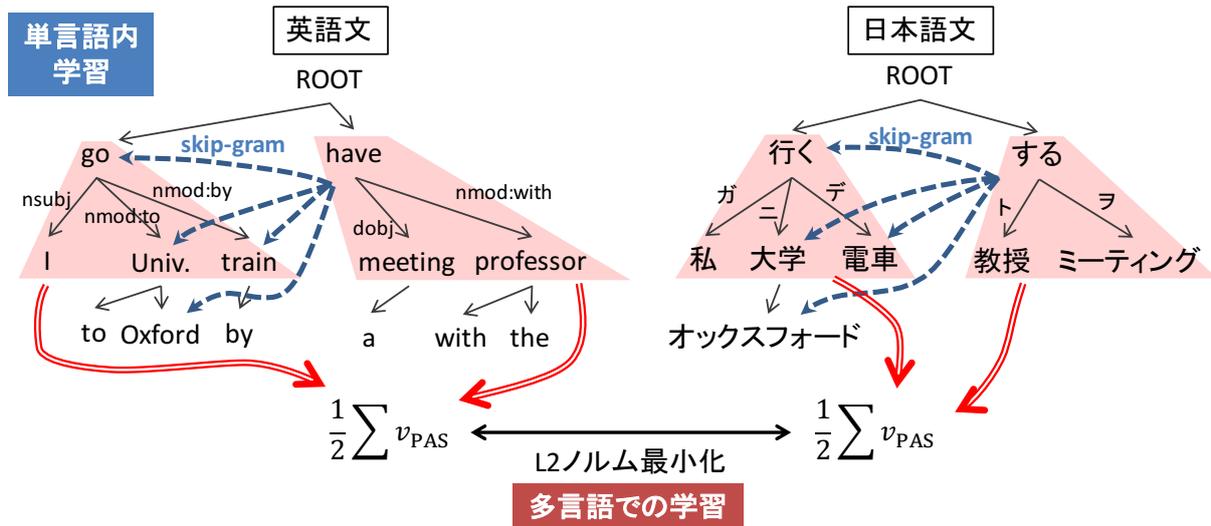


図1 提案モデルの概略. 英語文 “I went to Oxford University by train and had a meeting with the professor.” と日本語文 “私はオックスフォード大学に電車で行き、教授とミーティングをした.” における学習の過程を示している.

係る項のベクトルに述語との係り受け関係に応じた行列を掛けたものの和で表されると仮定する.

入力文中のある述語 p について、 p に係る項の集合を A とする. 述語 p の単語ベクトルを $v_p \in \mathbb{R}^n$ とし、項 $a \in A$ について、単語ベクトルを $v_a \in \mathbb{R}^n$ 、述語 p との間の係り受け関係を $R(p, a)$ で表し、述語 p および項集合 A からなる述語項構造 x のベクトル $v_{PAS}(x)$ を次式で計算する.

$$v_{PAS}(x) = f \left(W_{\text{pred}} v_p + b_{\text{pred}} + \sum_{a \in A} [W_{R(p,a)} v_a + b_{R(p,a)}] \right) \quad (1)$$

ここで $W_{\text{pred}} \in \mathbb{R}^{n \times n}$, $b_{\text{pred}} \in \mathbb{R}^n$ は述語ベクトルを写像する重み行列およびバイアス項で、 $W_{R(p,a)} \in \mathbb{R}^{n \times n}$, $b_{R(p,a)} \in \mathbb{R}^n$ は述語と項の係り受け関係における役割に応じて項ベクトルを写像する重み行列およびバイアス項である. $R(p, a)$ は日本語の場合、ガ、ヲ、ニなど、英語の場合、nsubj, dobjなどに相当する. また f は非線形の活性化関数であり、本研究では \tanh を用いた.

3.2 単言語内での述語項構造ベクトルの学習

単言語の単語ベクトル学習においては word2vec[2] が成果を挙げており、特に skip-gram モデルとその高速化手法である Negative Sampling の組み合わせが広く用いられている. skip-gram モデルは、コーパス内の語 (対象語と呼ぶ) とその周辺語を与え、対象語から周辺語が予測できるように単語のベクトル表現が学習される.

提案モデルでは、対象語を述語項構造に変更し、述語項構造から周辺語を予測する学習を行う. 周辺語として、付属語などのストップワードを除いた前後 c 単語を用いる. 例えば、例文 (3-a) の中の述語項構造 “have meeting with

professor” の場合、周辺語として “Oxford”, “University”, “train” が用いられる. このようにして、周辺語が類似する述語項構造のベクトルが類似するように、述語、項のベクトルならびに各種行列を学習する.

3.3 多言語での述語項構造ベクトルの学習

BilBOWA では、対訳文間の単語アライメントを用いずに多言語で単語ベクトルを学習している. 本モデルでは BilBOWA にならい、対訳文間の単語アライメントを用いず、各文に含まれる述語項構造ベクトルの平均を求め、対訳文間でこの平均ベクトル間の L_2 ノルムを最小化することで多言語で述語項構造ベクトルを学習する. 2 言語 e, f の各対訳文中の述語項構造を $E = \{e_1, e_2, \dots, e_m\}$, $F = \{f_1, f_2, \dots, f_n\}$ (m, n は各文に含まれる述語項構造の数) とし、この対訳文に対するコスト $\Omega(E, F)$ を次式で計算する.

$$\Omega(E, F) = \left\| \frac{1}{m} \sum_{i=1}^m v_{PAS}(e_i) - \frac{1}{n} \sum_{j=1}^n v_{PAS}(f_j) \right\|^2 \quad (2)$$

3.2 節で述べた単言語での学習と本節で述べた多言語での学習を行うことにより、多言語での単語ベクトルならびに行列を学習する. 具体的には対訳コーパスを与え、 k 対訳文を読み、単言語、多言語それぞれでのコストを計算し、誤差逆伝播法で単語ベクトルならびに行列を更新する. この操作を対訳コーパス全体に対して行い、これを繰り返す. 本モデルは BilBOWA と同様に、学習リソースとして対訳コーパスを必要とするが、さらに単言語コーパスを加えて大規模な学習を行うこともできる.

4. 実験

提案した多言語述語項構造ベクトル表現の妥当性を確認するため、述語項構造の対訳を選択する実験を行った。本研究では、日英、英中の2つの言語ペアについて実験を行った。

4.1 ベースラインモデル

ベースラインモデルとして、BilBOWAを用いたモデルを2つ考える。

単語モデル: BilBOWAを用いて単語ベクトルを学習し、述語項構造ベクトルを、述語項構造を構成する単語のベクトルを平均したものとして定義する。

述語項ペアモデル: 文中の項に対して述語と格の情報を付与し、述語、項およびその間の格の情報を含む**述語項ペア**のベクトル表現を学習する。まず学習用コーパスに対して、述語に直接係る項について「述語-項 [格]」という形へ置換する前処理を行う。この前処理を行ったコーパスを用いてBilBOWAで学習を行うことで、述語、項、格の情報を含む述語項ペアを単語とみなし、このベクトル表現が学習される。前処理の例として、例文(4-a)の前処理結果を(4-b)に示す。

- (4) a. 私は京都大学に電車で行き、昼を北部食堂で食べた。
b. 行く-私 [ガ] は 京都 行く-大学 [ニ] に 行く-電車 [デ] で行く, 食べる-昼 [ヲ] を 北部 食べる-食堂 [デ] で食べる.

4.2 モデルの学習

日英モデルの学習用コーパスとして、日英新聞記事対応付けデータ (JENAAD)[31]を使用した。このデータには新聞記事対訳文108万文が含まれるが、本研究ではこのうち対訳スコアの高い対訳文25万文をモデルの学習に用いた。また英中モデルの学習には、Chinese English News Magazine Parallel Text (LDC2005T10) [32]を用いた。このデータは新聞記事対訳文28万文を含む。

日本語については、JUMAN^{*2}・KNP^{*3}による構文・格解析結果から述語項構造を抽出した。述語の表記はKNPの用言代表表記を用い、項と係り受け関係はKNPの格解析結果から得た。英語と中国語については、Stanford CoreNLP[33]による構文解析結果から述語項構造を抽出した。ただし英語については、助動詞および子を持たないbe動詞は除外した。また、項は述語に直接係る語を抽出し、項が固有名詞である場合は“<name>”, 項が節の主辞である場合は“<s>”という表記に汎化した。

提案モデルはchainer^{*4}で実装した。単語ベクトルの次元数は200次元とし、単語ベクトルの初期値として、BilBOWAによる上記対訳コーパスの単語ベクトル学習結果を使用した。BilBOWAによる事前学習結果がないベクトルと各種行列は乱数で初期化した。3.2節で述べた周辺語のウィンドウ幅は $c=5$ とした。また3.3節で述べたミニバッチサイズは $k=100$ とし、学習を5エポック行った。

4.3 多言語述語項構造ベクトルの評価

4.3.1 評価手法

提案モデルが正しく述語項構造のベクトル表現を学習できているかを評価する。4.2節で用いた対訳コーパスとは別に評価用の対訳コーパスを用意し、述語項構造対を抽出したものを正解データとして、学習されたモデルを用いて述語項構造の対訳を選択する実験を行った。

評価用のコーパスとして、日英中基本文データ^{*5}から日英中対訳文5,304文を使用した。このコーパスは、格フレームをベースに抽出された日本語の基本的な文およびこれを人手で英語と中国語に翻訳した文からなり、述語項構造を1つのみ持つ文が多く含まれている。コーパス中の対訳文の例を表1に示す。評価に用いる正解述語項構造対応を得るために、評価する2言語について、各文に述語項構造1つのみを含む対訳文対を抽出した。このような対訳文として、日英では3,210文、英中では2,251文を得た。この対訳文から述語項構造を抽出し、述語項構造対応の正解データとした。実験は、日英、英日、英中、中英の4方向について行った。

日英方向の実験では、ある日本語述語項構造のベクトルを計算し、それと全英語述語項構造ベクトルとのcos類似度を計算し、英語述語項構造をランキングした。評価は、正解の英語述語項構造がランキングの上位K位以内($K=1, 5, 20$)に含まれる割合を計算することで行った。同様の手法により、英日、英中、中英方向の実験も行った。

4.3.2 結果と考察

各モデルの評価結果を表2に示す。表の数字はTopKに正解が含まれる精度を表す。提案モデルより得られた述語項構造対応付けの例を表3に示す。

提案モデルは、ベースラインモデル2種に比べて高い精度を達成した。表3の例(1)(2)(5)(6)は、提案モデルによって正しく対応付けられた述語項構造対の例である。例(1)(2)は日本語では同じ動詞「入る」からなる述語項構造だが、対応する英語述語項構造はそれぞれ“enter”, “reach”と異なる動詞からなるものが得られた。また例(5)については、日本語動詞「受ける」に対して英語動詞“receive”か

^{*4} <http://chainer.org/>

^{*5} <http://nlp.ist.i.kyoto-u.ac.jp/index.php?%E6%97%A5%E8%8B%B1%E4%B8%AD%E5%9F%BA%E6%9C%AC%E6%96%87%E3%83%87%E3%83%BC%E3%82%BF>

^{*2} <http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN>

^{*3} <http://nlp.ist.i.kyoto-u.ac.jp/index.php?KNP>

日本語	英語	中国語
今日の記事を読む	I read today's diary.	看今天的日记。
彼がみなさんのコメントを読みました	He read everyone's comments.	他看了大家的留言。
彼が会社で仕事をします	He works at a company.	他在公司里工作。
誰もが経験をする	Everyone experiences it.	谁都会经历。

表 1 日英中基本文データの例

言語	モデル	Top1	Top5	Top20
英 → 日	単語	0.007	0.022	0.045
	述語項ペア	0.011	0.021	0.041
	述語項構造 (提案手法)	0.160	0.283	0.390
日 → 英	単語	0.064	0.135	0.219
	述語項ペア	0.014	0.029	0.048
	述語項構造 (提案手法)	0.101	0.208	0.316
英 → 中	単語	0.019	0.051	0.083
	述語項ペア	0.004	0.014	0.028
	述語項構造 (提案手法)	0.108	0.236	0.368
中 → 英	単語	0.043	0.095	0.167
	述語項ペア	0.006	0.015	0.032
	述語項構造 (提案手法)	0.133	0.268	0.403

表 2 実験結果

らなる述語項構造と対応付けられた一方で、例 (6) では、日本語で「虐待を受ける」と表層的には能動態で表される意味が、英語で“is abused”と受動態で表されるという対応関係が得られた。このことから、提案モデルでは項による述語の意味の変化をうまく学習できていると言える。

表 3 の例 (3)(4) では、正解の英語述語項構造が上位 3 位以内に現れなかった。例 (3) の 1 位と 2 位は、日本語述語項構造と反義の関係にある英語述語項構造となっている。他にも例 (1) の 3 位、(5) の 2 位、(7) の 2 位に反義の英語述語項構造が現れた。これは提案モデルの多言語での述語項構造ベクトル学習の方法に起因すると思われる。多言語での述語項構造ベクトルの学習では、対訳文それぞれの中の述語項構造ベクトルを平均し、これを近づけるよう学習した。しかしこの時、上記の例のような意味は異なるが同じ文脈でよく用いられるような述語項構造の組については、複数の対訳文に同時に出現する可能性が高く、それゆえ誤った述語項構造が対応づけられて学習される可能性が高くなると考えられる。その結果、よく同じ文脈で用いられる述語項構造については、意味が異なっても類似度が高くなる場合があると思われる。

例 (4) の 2 位と 3 位には、日本語述語項構造と類似した英語述語項構造が得られた。正解の英語述語項構造は 214 位となったが、これは態が異なることからデータスパース性の影響を受けたためと思われる。

例 (7) の 1 位と 3 位に現れた英語述語項構造には、共に“attention”という項が含まれているが、これらは述語項

正解	順位	類似度	英語述語項構造
(1) 彼が大学に入る			
✓	1	0.609	he entered university
	2	0.606	he unexpectedly came across this book
	3	0.590	I will graduate from a school
(2) プロジェクトがよいよ段階に入る			
✓	1	0.520	project finally reached the final stages
	2	0.504	the city will promote development
	3	0.488	promenade will be built along wetlands
(3) 学校が休みに入る			
	1	0.590	child will attend classes
	2	0.567	children attended school for three years
	3	0.561	students have jobs

✓	8	0.548	schools have started holidays
(4) 情報がすぐに手に入る			
	1	0.581	he unexpectedly came across this book
	2	0.578	I know information
	3	0.575	I also gathered information

✓	214	0.479	information will be obtained quickly
(5) 彼が給付を受ける			
✓	1	0.634	he receives benefit
	2	0.588	he pays premiums himself
	3	0.551	he takes Route (6) towards Kyoto
(6) 彼が虐待を受ける			
✓	1	0.578	he is abused
	2	0.571	he unexpectedly came across this book
	3	0.545	he respects individuality
(7) 患者が治療を受ける			
	1	0.671	patients are given attention by doctors
	2	0.629	he had a disease
✓	3	0.618	patients receive attention

表 3 提案モデルによる日 → 英述語項構造対応付け例

上位 3 位以内に正解の述語項構造が含まれなかったものについては、正解の順位及び cos 類似度も合わせて示している。

構造を抽出した元の文では“medical attention”という複合語であり、述語項構造を抽出する際に“medical”の意味が脱落している。これに対応する日本語の項は「治療」となっており、項の意味に相違が生じている。このように、提案モデルでは述語に直接係る語のみを項としており、複合語を考慮していないため、述語項構造を抽出する際に項の修飾語の意味が脱落する問題がある。このような修飾語の考慮は今後の検討課題である。

5. おわりに

本論文では、多言語述語項構造ベクトルの学習手法を提案した。提案モデルは、多言語述語項構造の対応付けにおいて、単語ベクトルや述語項ペアベクトルに基づくベースラインモデルと比べて高い優位性を示した。

本研究ではモデルの学習に対訳コーパスのみを用いたが、本研究で提案したモデルはこれに単言語コーパスを加えて学習することもできる。大規模な単言語コーパスを加えた学習を行うことで、より広い範囲の述語項構造について学習できることが期待される。また、2言語に限らず、3言語以上について多言語述語項構造ベクトルの学習を同時に行うことが今後の課題に挙げられる。

さらには、学習した多言語述語項構造ベクトルを言語横断情報検索や言語横断文書分類などに適用し、その効果を検証する予定である。

参考文献

- [1] Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K. and Kuksa, P.: Natural Language Processing (Almost) from Scratch, *Journal of Machine Learning Research*, Vol. 12, pp. 2493–2537 (2011).
- [2] Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S. and Dean, J.: Distributed Representations of Words and Phrases and their Compositionality, *Advances in Neural Information Processing Systems 26* (Burgess, C., Bottou, L., Welling, M., Ghahramani, Z. and Weinberger, K., eds.), Curran Associates, Inc., pp. 3111–3119 (online), available from <http://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf> (2013).
- [3] Pennington, J., Socher, R. and Manning, C. D.: GloVe: Global Vectors for Word Representation, *Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1532–1543 (online), available from <http://www.aclweb.org/anthology/D14-1162> (2014).
- [4] Socher, R., Lin, C. C., Ng, A. Y. and Manning, C. D.: Parsing Natural Scenes and Natural Language with Recursive Neural Networks, *Proceedings of the 26th International Conference on Machine Learning (ICML)* (2011).
- [5] Socher, R., Pennington, J., Huang, E. H., Ng, A. Y. and Manning, C. D.: Semi-Supervised Recursive Autoencoders for Predicting Sentiment Distributions, *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (2011).
- [6] Socher, R., Huval, B., Manning, C. D. and Ng, A. Y.: Semantic Compositionality through Recursive Matrix-Vector Spaces, *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, Jeju Island, Korea, Association for Computational Linguistics, pp. 1201–1211 (online), available from <http://www.aclweb.org/anthology/D12-1110> (2012).
- [7] Socher, R., Bauer, J., Manning, C. D. and Andrew Y., N.: Parsing with Compositional Vector Grammars, *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Sofia, Bulgaria, Association for Computational Linguistics, pp. 455–465 (online), available from <http://www.aclweb.org/anthology/P13-1045> (2013).
- [8] Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A. and Potts, C.: Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank, *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, Seattle, Washington, USA, Association for Computational Linguistics, pp. 1631–1642 (online), available from <http://www.aclweb.org/anthology/D13-1170> (2013).
- [9] Li, P., Liu, Y. and Sun, M.: Recursive Autoencoders for ITG-Based Translation, *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, Seattle, Washington, USA, Association for Computational Linguistics, pp. 567–577 (online), available from <http://www.aclweb.org/anthology/D13-1054> (2013).
- [10] Socher, R., Karpathy, A., Le, Q. V., Manning, C. D. and Ng, A. Y.: Grounded Compositional Semantics for Finding and Describing Images with Sentences, *Transactions of the Association for Computational Linguistics*, Vol. 2, pp. 207–218 (2014).
- [11] Hashimoto, K., Stenetorp, P., Miwa, M. and Tsuruoka, Y.: Jointly Learning Word Representations and Composition Functions Using Predicate-Argument Structures, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, Association for Computational Linguistics, pp. 1544–1555 (online), available from <http://www.aclweb.org/anthology/D14-1163> (2014).
- [12] Liu, Y., Wei, F., Li, S., Ji, H., Zhou, M. and Wang, H.: A Dependency-Based Neural Network for Relation Classification, *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, Beijing, China, Association for Computational Linguistics, pp. 285–290 (online), available from <http://www.aclweb.org/anthology/P15-2047> (2015).
- [13] Klementiev, A., Titov, I. and Bhattarai, B.: Inducing Crosslingual Distributed Representations of Words, *Proceedings of COLING 2012*, Mumbai, India, The COLING 2012 Organizing Committee, pp. 1459–1474 (online), available from <http://www.aclweb.org/anthology/C12-1089> (2012).
- [14] Hermann, K. M. and Blunsom, P.: Multilingual Models for Compositional Distributed Semantics, *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Baltimore, Maryland, Association for Computational Linguistics, pp. 58–68 (online), available from <http://www.aclweb.org/anthology/P14-1006> (2014).
- [15] 林 佑明, 酒井哲也: 言語の分散表現による文脈情報を利用した言語横断情報検索, *DEIM Forum 2015*, (オンライン), 入手先 <http://db-event.jp/deim2015/paper/350.pdf> (2015).
- [16] Gouws, S., Bengio, Y. and Corrado, G.: BiBOWA: Fast Bilingual Distributed Representations without Word Alignments, *Proceedings of The 32nd International Conference on Machine Learning*, pp. 748–756 (2015).
- [17] Mikolov, T., Le, Q. V. and Sutskever, I.: Exploiting Similarities among Languages for Machine Translation, *CoRR*, Vol. abs/1309.4168 (online), available from <http://arxiv.org/abs/1309.4168> (2013).
- [18] Ishiwatari, S., Kaji, N., Yoshinaga, N., Toyoda, M. and Kitsuregawa, M.: Accurate Cross-lingual Projec-

- tion between Count-based Word Vectors by Exploiting Translatable Context Pairs, *Proceedings of the Nineteenth Conference on Computational Natural Language Learning*, Beijing, China, Association for Computational Linguistics, pp. 300–304 (online), available from <http://www.aclweb.org/anthology/K15-1030> (2015).
- [19] Zou, W. Y., Socher, R., Cer, D. and Manning, C. D.: Bilingual Word Embeddings for Phrase-Based Machine Translation, *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, Seattle, Washington, USA, Association for Computational Linguistics, pp. 1393–1398 (online), available from <http://www.aclweb.org/anthology/D13-1141> (2013).
- [20] Faruqui, M. and Dyer, C.: Improving Vector Space Word Representations Using Multilingual Correlation, *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, Gothenburg, Sweden, Association for Computational Linguistics, pp. 462–471 (online), available from <http://www.aclweb.org/anthology/E14-1049> (2014).
- [21] Zhao, K., Hassan, H. and Auli, M.: Learning Translation Models from Monolingual Continuous Representations, *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Denver, Colorado, Association for Computational Linguistics, pp. 1527–1536 (online), available from <http://www.aclweb.org/anthology/N15-1176> (2015).
- [22] Lu, A., Wang, W., Bansal, M., Gimpel, K. and Livescu, K.: Deep Multilingual Correlation for Improved Word Embeddings, *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Denver, Colorado, Association for Computational Linguistics, pp. 250–256 (online), available from <http://www.aclweb.org/anthology/N15-1028> (2015).
- [23] Luong, T., Pham, H. and Manning, C. D.: Bilingual Word Representations with Monolingual Quality in Mind, *Proceedings of the 1st Workshop on Vector Space Modeling for Natural Language Processing*, Denver, Colorado, Association for Computational Linguistics, pp. 151–159 (online), available from <http://www.aclweb.org/anthology/W15-1521> (2015).
- [24] Lauly, S., Boulanger, A. and Larochelle, H.: Learning Multilingual Word Representations using a Bag-of-Words Autoencoder, *CoRR*, Vol. abs/1401.1803 (online), available from <http://arxiv.org/abs/1401.1803> (2014).
- [25] Kočiský, T., Hermann, K. M. and Blunsom, P.: Learning Bilingual Word Representations by Marginalizing Alignments, *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, Baltimore, Maryland, Association for Computational Linguistics, pp. 224–229 (online), available from <http://www.aclweb.org/anthology/P14-2037> (2014).
- [26] Chandar, A. P. S., Lauly, S., Larochelle, H., Khapra, M. M., Ravindran, B., Raykar, V. C. and Saha, A.: An Autoencoder Approach to Learning Bilingual Word Representations, *CoRR*, Vol. abs/1402.1454 (online), available from <http://arxiv.org/abs/1402.1454> (2014).
- [27] Coulmance, J., Marty, J.-M., Wenzek, G. and Benhaloum, A.: Trans-gram, Fast Cross-lingual Word-embeddings, *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, Lisbon, Portugal, Association for Computational Linguistics, pp. 1109–1113 (online), available from <http://aclweb.org/anthology/D15-1131> (2015).
- [28] Vulić, I. and Moens, M.-F.: Bilingual Word Embeddings from Non-Parallel Document-Aligned Data Applied to Bilingual Lexicon Induction, *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, Beijing, China, Association for Computational Linguistics, pp. 719–725 (online), available from <http://www.aclweb.org/anthology/P15-2118> (2015).
- [29] Zhang, J., Liu, S., Li, M., Zhou, M. and Zong, C.: Bilingually-constrained Phrase Embeddings for Machine Translation, *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Baltimore, Maryland, Association for Computational Linguistics, pp. 111–121 (online), available from <http://www.aclweb.org/anthology/P14-1011> (2014).
- [30] Su, J., Xiong, D., Zhang, B., Liu, Y., Yao, J. and Zhang, M.: Bilingual Correspondence Recursive Autoencoder for Statistical Machine Translation, *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, Lisbon, Portugal, Association for Computational Linguistics, pp. 1248–1258 (online), available from <http://aclweb.org/anthology/D15-1146> (2015).
- [31] Utiyama, M. and Isahara, H.: Reliable Measures for Aligning Japanese-English News Articles and Sentences, *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics*, Sapporo, Japan, Association for Computational Linguistics, pp. 72–79 (online), DOI: 10.3115/1075096.1075106 (2003).
- [32] Ma, X.: *Chinese English News Magazine Parallel Text LDC2005T10*, Linguistic Data Consortium (2005).
- [33] Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J. and McClosky, D.: The Stanford CoreNLP Natural Language Processing Toolkit, *Association for Computational Linguistics (ACL) System Demonstrations*, pp. 55–60 (online), available from <http://www.aclweb.org/anthology/P/P14/P14-5010> (2014).