

自動構築した日本語格フレームと英語 FrameNet の対応付け

笹野 遼平

名古屋大学 情報学研究科

sasano@i.nagoya-u.ac.jp

1 はじめに

言語理解のための基本的な知識の1つにフレーム知識がある。英語に着目すると、FrameNet や PropBank などの大規模なフレーム知識が人手で整備されており、多く意味役割付与に関する研究が行われている [1]。このうち、FrameNet は、図 1 の左に示すようなフレームが定義された辞書と、フレーム情報が付与された注釈付き文の集合から構成される言語資源である。FrameNet に基づき (1) に示す文の意味役割付与を行う場合、述語 broke が喚起するフレームとして図 1 に示した Cause_to_fragment フレームが選択され、He, the window, a stone が担う意味役割として、それぞれ Agent, Whole_patient, Instrument が付与される。

(1) He broke the window with a stone.

一方、日本語を対象とした述語項構造解析では意味役割ではなく表層格の解析が主流であり、フレームに基づく解析を行う場合も表層格フレームに基づく解析が行われることが多い。たとえば、述語項構造解析器 KNP¹に、(2) に示す文を与えた場合、述語「割った」の項「彼」、「石」、「窓」に Agent, Whole_patient, Instrument などといった意味役割が付与されるのではなく、図 1 の右に示す『割る:動18』という格フレームが選択され、「彼」、「石」、「窓」はそれぞれガ格、デ格、ヲ格に対応付けられる。

(2) 彼は石で窓を割った。

このような表層格フレームは、大規模コーパスから自動構築することが可能であるため、大規模でカバレッジが高く、多くの用例情報を含んでいるという特長がある。また、単に用例を収集するだけでなく、それらをクラスタリングすることで、意味・用法ごとに異なるフレームを構築することも可能となっている [2]。しかし、フレームの粒度が人間の直感に合ったものであるかは疑問であり、また、意味役割は付与されていないことから各項が担う意味内容は自明ではない。このため、より人間に理解しやすい文解析の実現を目指し、英語以外の言語に対しても人手によるフ

¹<http://nlp.ist.i.kyoto-u.ac.jp/?KNP>

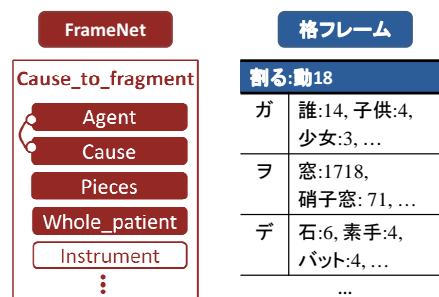


図 1: FrameNet におけるフレームと日本語格フレーム

レーム知識の整備が進められている。たとえば、日本語 FrameNet [4] は英語 FrameNet と互換性を持つ言語資源であり、言語横断的な意味理解に有用な言語リソースであると言える。しかし、注釈付き文の規模は英語 FrameNet の 5%程度であるなどカバレッジはまだ限定的である。また、日本語 FrameNet と自動構築した日本語格フレームを対応付けを目指した研究 [5] も存在するが、日本語 FrameNet の存在が前提となっており、人手により整備されたフレーム資源が存在しない言語に応用しようとした場合、基盤となるフレーム知識の人手による整備が必要となる。

そこで本研究では、自動構築された日本語格フレームを、日本語 FrameNet を介さず、直接、英語 FrameNet に自動的に対応付けることを考える。具体的には、近年急速に発展している言語横断的な単語意味ベクトル獲得技術 [6] を利用し、FrameNet に含まれる注釈付き文から得られる用例情報と、自動構築された日本語格フレームに含まれる用例情報の関連性に注目することで、人手を用いずに日本語格フレームと FrameNet の対応付けを行うことを目指す。

2 提案手法

2.1 対応付けの対象

本研究では、約 70 億文の Web テキストから自動構築された日本語格フレームである京都大学格フレーム²と、FrameNet version 1.7 を対象にフレームの対応

²<http://nlp.ist.i.kyoto-u.ac.jp/index.php?> 京都大学格フレーム

付けを行う。また、対応付けの対象とする格は、京大格フレームのガ格、ヲ格、ニ格、および、FrameNetにおいて core クラスに分類された意味役割に限定する。

京大格フレームでは、フレームは各動詞の出現形ごと、用法ごとに、それぞれ複数、構築されていることから、FrameNet で定義されたフレーム数 1223 と比べ、圧倒的に多くのフレームが構築されている。このため本研究では、京大格フレームを入力とし、各格フレームと関連性をもっとも強いと考えられる FrameNet のフレームへ対応付ける。すなわち、FrameNet のフレーム 1 つに、複数の京大格フレームのフレームが対応付くという、1 対多の対応付けを行うことになる。以下、本稿では簡単のため、格フレームと記載した場合は京大格フレームのフレームを、FN フレームと記載した場合は FrameNet におけるフレームを指すものとする。

2.2 FrameNet からの用例収集

本研究で使用する FrameNet 1.7 では、1223 フレームに対し、約 20 万の注釈付き文が与えられている。フレームの対応付けに先立ち、これらの注釈付き文から、各 FN フレームのフレーム喚起語、および、各意味役割を担う用例の収集を行う。この際、フレーム喚起語は原形に戻さず出現形で収集し、また、用例が複数の語で構成されている場合は Stanford Parser³を用いて句構造解析を行い、主辞のみを収集対象とする。

2.3 フレームの対応付け

提案手法において、フレームの対応付けは大きく、**1. FN フレーム候補の絞り込み**と、**2. 最適な対応付けの決定**、の 2 ステップに分けられる。まず FN フレーム候補の絞りこみを行うのは、不必要な対応付けの評価を省き、計算コストを削減することが主な目的である。以下では各ステップの詳細について説明する。

1. FN フレーム候補の絞り込み 入力された日本語格フレームの動詞と、各 FN フレームのフレーム喚起語一覧との類似度のみを手掛りとして、対応付け候補となる FN フレーム候補を N_f 個に絞り込む。以降、本稿ではこの類似度を喚起語スコアと呼ぶ。喚起語スコアの具体的な算出手順は以下のとおりである。

まず、言語横断単語意味ベクトル空間において、対象の日本語動詞との余弦類似度が高い英語単語を 1000 個を類似度とともに列挙する。続いて、各 FN フレームに対し、列挙された英語単語集合に含まれるフレーム喚起語を、類似度が大きいものから順に最大 N_e 個

抽出し、それらの類似度の合計をその FN フレーム f の喚起語スコア $\text{score}_{\text{喚}}(f)$ とし、スコア $\text{score}_{\text{喚}}(f)$ の大きい N_f 個に FN フレーム候補を絞り込む。

2. 最適な対応付けの決定 絞り込まれた FN フレーム候補を対象に、格フレームの用例情報、および、FN フレームの各項の用例情報を用いて、フレームの対応付けスコアを算出し、喚起語スコアとの積がもっとも大きくなったものを最終的な対応付け結果とする。

対応付けスコアは、対象となる格フレームのガ格、ヲ格、ニ格と、候補となる FN フレームの core 意味役割の全組み合わせを考え、各格の対応付けスコアを、その格の用例数の平方根で重み付けした上で足し合わせることで算出する。すなわち、格フレーム f_{CF} の格 c に対し、その格の用例数を $\text{freq}(c)$ 、その対応付け先となる FN フレーム f_{FN} の意味役割を $r(c)$ 、対応付けスコアを $\text{score}_{\text{格}}(c, r(c))$ と表す場合、フレームの全体の対応付けスコア $\text{score}_F(f_{CF}, f_{FN})$ は以下により算出される。

$$\text{score}_F(f_{CF}, f_{FN}) = \sum_{c \in (\text{ガ}, \text{ヲ}, \text{ニ})} \text{score}_{\text{格}}(c, r(c)) * \sqrt{\text{freq}(c)}$$

ただし、同一の意味役割に対応付ける格フレームの格は 1 つまでとし、ガ格は必ずいずれかの意味役割に対応付けるという制約を加える。また、格の対応付けスコア $\text{score}_{\text{格}}(c, r)$ は、格フレームの格 c の用例集合 E_c に含まれる語の意味ベクトルの平均ベクトルと、意味役割 r の用例集合 E_r に含まれる語の意味ベクトルの平均ベクトルの余弦類似度により定義する。また、格フレームのある格 c がいずれの意味役割とも対応付けられなかった場合もスコアは 0 とはせず、定数 $c_{NIL} > 0$ を与えるものとする。これは、意味ベクトル空間における語ベクトルは偏在しており [3]、ほぼ無関係な語の集合同士であっても各集合の平均ベクトルの余弦類似度を計算すると正の値となる傾向があるためであり、 c_{NIL} を導入することで不適切な格と意味役割の対応付けが抑制されることが期待できる。

3 日本語 FrameNet に基づく評価

フレーム対応付けの評価は、図 2 に示すような、対応情報を用いた日本語意味役割付与を通して行う。この際、意味役割の正解情報には、日本語 FrameNet に含まれる注釈付き文⁴を利用する。評価の具体的な手順は次の通りである。

⁴日本語と英語の FrameNet には互換性があることから、基本的に共通するフレームが付与されている。

³<https://nlp.stanford.edu/software/lex-parser.shtml>

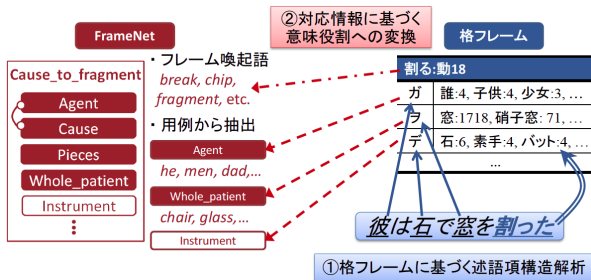


図 2: 対応付け情報を用いた日本語意味役割付与

手順0 日本語 FrameNet に含まれる注釈文から、フレーム喚起語とその項の主辞だけから構成される単文を生成する。この際、注釈文に付与された FN フレームが FrameNet 1.7 に存在しないもの、述語が単独の動詞で構成されていないもの、項が名詞句以外であるものは解析対象から除く。

手順1 手順0 で生成した単文を、格フレームに基づく日本語述語構造解析器である KNP を用いて解析し、各項を選択された格フレームの各格と対応付ける。格フレームの格と対応付けられた項が存在しなかった文は解析対象から除く。

手順2 手順1 で選択された格フレームと FrameNet の対応付け情報を用い、手順1 で対応付けられた格を FrameNet の意味役割に変換し、注釈で与えられた意味役割と比較、評価する。

手順0, 1 において、フレーム喚起語とその項だけから構成される単文を生成したり、一部の文を除外したりしているのは、フレーム対応付け結果の評価が目的であることから、構文解析誤りなどフレームの対応付けと関係のない解析誤りの影響をできる限り少なくするためである。また、述語が単独の動詞でないものとは「食べ始める」や「言われる」などのように補助動詞や助動詞が付属しているものが相当する。京大格フレームでは、これらの述語に相当する格フレームが構築されているが、本研究では単語意味ベクトルを基本的に形態素単位で構築しているため、これらの述語に相当する意味ベクトルが存在しないため解析対象から除いた。

実際の解析例として、(3) に日本語 FrameNet に含まれる注釈文の例を示す。この例では、述語「信じれ」に対し、Religious.belief フレームが付与され、「正しい神」に Element という意味役割が割り当てられている。この文に対し意味役割付与を行う場合、手順0 では「正しい神」の主辞である「神」と述語「信じれ」だけからなる単文「神を信じる」を生成し、手順1 でこれを KNP で解析することで動詞「信ずる」が格フ

表 1: 正解フレームが含まれる割合 () 内の数字は 1407 事例中、正解フレームが含まれた数を表す)

N_f	1	3	10	30	100	300
割合	.379 (533)	.593 (835)	.722 (1016)	.808 (1137)	.907 (1277)	.950 (1337)

レーム『信ずる:動1』に、「神」はそのヲ格に対応付けられ、手順2 でフレーム対応付け情報を用い FrameNet の意味役割に変換される。

(3) 正しい神を信じれば、正しくない神は自然に消滅してしまいます。

日本語 FrameNet に含まれる注釈文 7899 文のうち、最終的に評価の対象となった注釈文は 1407 文であった。さらに、core クラスに分類された意味役割が付与された項で、KNP でガ格に対応付けられたものは 691 個、ヲ格に対応付けられたものは 401 個、ニ格に対応付けられたものは 174 個で、合計 1266 個⁵であった。本研究では、これら 1407 文に含まれる 1266 個の項を評価に利用する。

4 実験

4.1 実験設定

日本語と英語を対象とした言語横断単語意味ベクトルは、MUSE⁶の教師ありモデルを用いて構築した。元となる日本語単語意味ベクトルは、CommonCrawl の日本語テキスト⁷から抽出し JUMAN を用いて形態素区切りを行った日本語 1 億文に対し、skipgram モデル⁸を適用することで構築した。英語単語意味ベクトル⁹、および、日本語英語対訳辞書¹⁰は Web 上に公開されているものを利用した。

また、フレーム対応付けに使用する各種パラメータのうち、FN フレーム候補の列挙において考慮するフレーム喚起語の最大数 N_e は 3、対応付けられなかった格フレームの格に与えるスコア c_{NIL} は 0.5 とした。

4.2 FN フレーム候補の絞り込みの評価

まず、日本語動詞とフレーム喚起語の類似度のみから、どのくらい正解となる FN フレームを絞り込めるかを確認するため、絞り込むフレーム数 N_f と絞り込まれた FN フレーム中に正解 FN フレームが含まれてい

⁵文数 1407 より小さいのは、ガ格、ヲ格、ニ格のいずれかと対応付けられた core 意味役割が 1 つもない文が存在するためである。

⁶<https://github.com/facebookresearch/MUSE>

⁷<http://statmt.org/ngrams/>

⁸<https://github.com/facebookresearch/fastText>

⁹<https://s3-us-west-1.amazonaws.com/fasttext-vectors/crawl-300d-2M.vec.zip>

¹⁰<https://s3.amazonaws.com/arrival/dictionaries/ja-en.txt>

表 2: 意味役割付与の精度

フレームを 与えたか	フレーム 一致率	フレームが一致した場合の格対応正解率			
		ガ格	ヲ格	ニ格	合計
自動推定	.392 (552/1407)	.764 (243/318)	.551 (81/147)	.603 (41/68)	.685 (365/533)
正解付与	1.000 (1407/1407)	.740 (511/691)	.453 (182/401)	.477 (83/174)	.613 (776/1266)

る割合の関係を調査した。結果を表 1 に示す。約 38% の事例で最上位に、約 72% の事例で上位 10 フレーム中に、正解フレームが含まれていることが確認できる。FrameNet 1.7 で定義されているフレーム数は 1223 であることを考慮すると、喚起語スコアにより効率的にフレーム候補を絞り込んでいると考えられる。

4.3 フレーム対応付けの評価

フレーム対応付けの精度を確認するため、意味役割付与結果の評価を行った。本実験では $N_f = 10$ とした。また、正しいフレームが与えられた場合の、格の対応付けの精度を確認するため、正しいフレームを与えた実験も行った。結果を表 2 に示す。フレームの自動推定を行う場合は、格対応の正解率は、正しいフレームが推定できた場合のみを対象に行っている。

注釈付き文に付与されたフレームを正しく推定できた割合は 39.2%(552/1407) であり、喚起語スコアが最大となるフレームを出力した場合の精度 37.9%(533/1407) より大きな値となった。これは、喚起語スコアのみでフレームをランキングした場合は 2 番目以降に順位付けられたフレームが、対応付けも考慮した結果、最終的に選択されたケースが存在するためである。たとえば、(3) に示した文の解析では、喚起語スコアのみを用いてフレームを評価した場合、スコア上位 2 つは順に Awareness フレームと Certainty フレームであり、Religious_belief フレームは 3 番目であるが、フレーム対応付けスコアまで考慮することで、最終的に正しいフレームと意味役割付与が出力される。また、フレームと意味役割のペアを正しく推定できたものの割合は 28.8%(365/1266) であった。

フレームが一致した場合の格対応の正解率に目を移すと、まず、人手で正解を与えた場合の全体の精度よりも、自動推定で正しいフレームを選択できた場合の精度の方が高いことが確認できる。これは、正しいと考えられる格の対応付けが見つかった場合の方が、選択されたフレームの正解率も高いと考えられることから、妥当な結果であると言える。自動推定で正しいフレームを選択できた場合の格対応正解率は全体で 68.5%、ガ格に限定すると 76.4% と比較的高い精度が達成されていることが確認できる。このため、今後、意味役割

付与の精度を向上させるためには、フレームの選択精度を向上させることが重要であると考えられる。

5 まとめと今後の展望

本研究では、自動構築された日本語格フレームに含まれる用例情報と FrameNet の注釈付き文を用いることで、日本語格フレームと FrameNet の言語横断的な対応付けに取り組んだ。また、日本語を対象とした意味役割付与に応用し、日本語の教師データを使わない設定において、39.2% の精度でフレーム推定を、正しいフレームが推定できた場合は 61.3% の精度で意味役割の推定を実現できることを示した。

本研究は自動構築された日本語格フレームと英語 FrameNet の自動的な対応付けの第一ステップであるとの考えから、本稿で提案した手法では基本的な情報しか利用しなかった。しかし、ガ格、ヲ格、ニ格以外の格や、core クラス以外に分類された意味役割、FrameNet において同一のフレームの意味役割間に付与されている CoreSet や Excludes relation 等の情報も、フレーム対応付けの手掛りとなると考えられる。したがって、これらの情報を活用することで、さらに高度な対応付けが実現可能であると考えられる。また、本稿で提案した手法は、人手による日本語 FrameNet のアノテーションの補助にも活用可能であると考えられる。

参考文献

- [1] Collin Baker, Michael Ellsworth, and Katrin Erk. SemEval'07 task 19: Frame semantic structure extraction. In *Proc. of SemEval'07*, pp. 99–104, 2007.
- [2] Daisuke Kawahara, Daniel Peterson, Octavian Popescu, and Martha Palmer. Inducing example-based semantic frames from a massive amount of verb uses. In *Proc. of EACL'14*, pp. 58–67, 2014.
- [3] David Mimno and Laure Thompson. The strange geometry of skip-gram with negative sampling. In *EMNLP'17*, pp. 2873–2878, 2017.
- [4] Kyoko Ohara. Relating frames and constructions in Japanese FrameNet. In *Proc. of LREC'14*, pp. 2474–2477, 2014.
- [5] Kyoko Ohara, Daisuke Kawahara, Satoshi Sekine, and Kentaro Inui. Linking Japanese FrameNet with Kyoto University Case Frames using crowdsourcing. In *Proc. of the International FrameNet Workshop 2018: Multilingual Framenets and Constructicons*, pp. 56–61, 2018.
- [6] Sebastian Ruder, Ivan Vulic, and Anders Sogaard. A survey of cross-lingual word embedding models. In *In Arxiv preprint https://arxiv.org/abs/1706.04902*, 2017.