

# 集合知を用いた大規模意味的フレーム知識の構築

小原京子  
慶大/理研 AIP

河原大輔  
早大/理研 AIP

笹野遼平  
名大/理研 AIP

関根聡  
理研 AIP

ohara@hc.st.keio.ac.jp dkw@waseda.jp

sasano@i.nagoya-u.ac.jp satoshi.sekine@riken.jp

## 1 はじめに

本研究の目的は、高度な言語理解の実現に向けて、自然言語処理の手法で構築された「京都大学格フレーム」(KCF)の格フレームと人手で構築中の「日本語フレームネット」(JFN)の意味フレームとを集合知を用いて対応付け、フレーム知識を統合することである。具体的な目標は2つある。究極的目標は、KCF格フレームをJFN意味フレームに対応付けし、KCF格フレームを構成するすべての文がそのJFN意味フレームにあてはまるとみなすことによって、JFN意味フレームアノテーション付きコーパスを大幅に増強することである。もう1つはクラウドソーシングの結果からJFN上の既存のフレーム定義が不完全であることがわかる可能性があるため、これらについて人手でチェックすることによりJFNを改良することである。本論文は、1つ目の究極的目標に向けた中間報告であるとともに、2つ目の目標の指針を決定するためのケーススタディである。

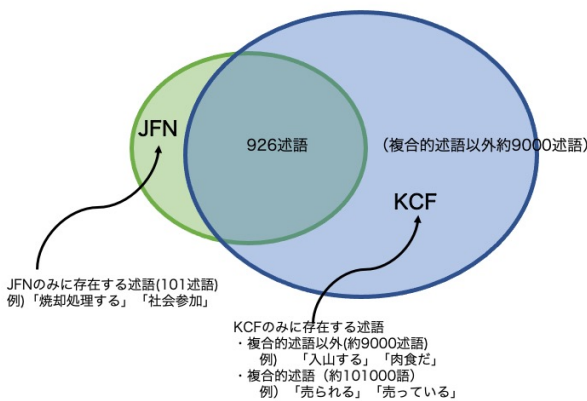


図1 KCF格フレームとJFN意味フレーム

本研究では、ある述語に対するそれぞれのKCF格フレームについて、その述語に対して定義されたJFN意味フレームのいずれかに対応付ける(図1)。この対応付けを大規模かつ高速に行うためにクラウド

ソーシングを用いる。具体的には、一つのKCF格フレームを代表する例文を提示文、JFN意味フレームの例文2文を選択肢1,2として、提示文と用法が似ている選択肢を選ぶタスクを実行する(図2)。選択肢1,2に加えて、「いずれにも似ていない、もしくは判断できない」という選択肢(以下では“OTHER”と呼ぶ)を別途設け、JFNのいずれの例文にも似ていない場合、もしくは提示文からは判断が難しい場合に選択してもらう。この選択肢が多量のクラウドワーカーに選ばれた場合には、提示文、選択肢、あるいはそれ以外の何かを見直した方が良い可能性がある。

提示文中の[...]の語の意味を考え、意味がもっとも似ている選択肢を選んでください。意味が似ている選択肢がない場合、もしくは判断できない場合は「※どれにも似ていない、もしくは判断できない」を選んでください。

**ファイルサイズが指定サイズ以上のファイルを【収集】します**

提示文

選択肢1

我が国のそれぞれの地域・企業において実際に行われている取組みの事例を【収集】し、分析することとした。

選択肢2

廃棄物等は品目別毎に【収集】され、廃棄物発電などのサーマルリサイクルも活発化します。

※どれにも似ていない、もしくは判断できない

OTHER

図2 クラウドソーシングタスク例

KCFとJFNの両方に存在する述語は図1に示した通り926述語あり、そのうちJFNにアノテーション例文があるのは713述語である。これらのうち185述語については既にクラウドソーシングを行い、言語分析者が人手でその結果の評価を行った。その内訳は、表1の通りである。

本論文では、表1の85述語と同様にJFN上で現在1つの意味フレームに割り当てられており、複数のアノテーション例文が存在する81述語のクラウドソーシング結果について報告と考察を行う。考察においては、JFNで2つの意味フレームに割り当て

られた 37 述語のクラウドソーシング結果に関する先行研究[1]、[2]の知見と比較する。

表 1 実施済みクラウドソーシング対象

JFN上で割り当てられていた意味フレーム数 (語義数に相当)	JFN上に例文の存在する述語数
2	37
3以上	63
1	85 <sup>i</sup>

分析の結果、JFN に一義語として登録されている述語についても、正解率の高低やクラウドワーカーからの回答には多義語の場合と同様の傾向が見られることがわかった。加えて、クラウドソーシング結果の正解率の向上と JFN 改良のための指標が得られた。

以下では、まず第 2 節で先行研究の知見を概観した後に、第 3 節で今回の実験について論じる。第 4 節では今回の実験結果を先行研究のそれと比較する。最後の第 5 節で今後の展望について述べる。

## 2 先行研究

先行研究[1]、[2]では、JFN において 2 つの意味フレームに割り当てられており（つまり 2 つの意味を持つ多義語）、それぞれの JFN 意味フレームで例文が 1 文以上存在する 37 述語について実験を行った。その結果、述語や格フレームによって最多票を得た回答の正解率にばらつきが見られるが、2 つの語義が明らかに隔たっていて JFN 上の意味フレーム定義に不備がない述語は正解率が高いこと、最多票の回答が OTHER で得票数が多い格フレーム例文には複単語表現 (multi-word expression、以下 MWE) が出現しているものが多いことが明らかとなった。

10 人のクラウドワーカーからの回答について、最多票を獲得した JFN 意味フレームとその例文が正解だったかどうかを言語分析者が人手で評価・分析した。各述語の 2 つの語義間の類似度と、それらの意味が現行の JFN に正確に意味フレームとして登録されているかどうかにより、37 述語を 3 つに分類した：

**カテゴリ I:** 2 つの語義が明らかに隔たっている (9 述語、例：「送る」)。

**カテゴリ II:** 2 つの語義が似通っている (11 述語、

例：「書く」)。

**カテゴリ III:** JFN 上に定義済みの JFN 意味フレームが、当該述語の KCF 格フレームに見られる全ての語義を正確に表現していない (17 述語)。以下の 2 種類のケースがある：1) JFN 上で誤った意味フレームが当該述語に割り当てられていた (例：「適切だ」)；2) KCF 格フレームから当該述語が 3 つ以上の語義を持つことが判明したが、JFN 上の当該述語に対する意味フレーム割り当てが不完全であった (例：「与える」)。

述語カテゴリ別の、述語の平均正解率は以下の通りであった。

カテゴリ I: 83.9%

カテゴリ II: 55.0%

カテゴリ III: 39.9%

この結果から、クラウドワーカーらはカテゴリ I の述語、つまり 2 つの語義が明らかに隔たっている述語に対しては、概ね正しく KCF 格フレームを JFN 意味フレームに対応づけられたと言える。

また、最多票が OTHER でその得票数が多い KCF 格フレーム例文には複単語表現が関与するものが多かった。(1)から(3)はその例である (文末のカッコ内の数字は最多票を得た OTHER の得票数)。

- (1) 自分が青春時代を【送った】 (10)
- (2) その日に私が話題に【上った】 そうだ (9)
- (3) ここは私が話を【通そう】 (8)

KCF の述語には MWE は含まれず、JFN でも今のところ MWE の見出し語は少ないので、KCF 格フレームと JFN 意味フレームの対応付けの際には MWE の扱いに留意する必要がある。

本研究の究極的目標である KCF 上のそれぞれの格フレームに対して JFN 意味フレームを効率的に対応づけるためにも、当面の目標である JFN コンテンツの改良のためにも、クラウドソーシング結果を言語分析者が人手で詳細に評価する前にあらかじめ正解率の高い述語や留意すべき格フレームを推定できることが望ましい。従って、この先行研究で得られた、JFN 上の意味フレーム定義に不備がない述語の正解率が高いという傾向と、OTHER が最多票でなおかつ得票数が多い場合に MWE を含むケースが多

<sup>i</sup> これらの述語には各 JFN 意味フレームにおいて複数のアンテーション例文が存在する。

いという傾向が、一義語についても見られるかどうかを調査することにした。

### 3 実験

#### 3.1 実験設定

今回の実験では、JFN 上で1つの JFN 意味フレームに登録されている(すなわち一義語とされている)81 述語、合計 1319KCF 格フレーム(1 述語あたり平均 16.7 格フレーム)を対象とした。

#### 3.2 結果

先行研究と同様に、最多票を獲得した選択肢が正解かどうかを評価した。OTHER が最多票を獲得した際にそれが正解かどうかを得票数別に調べた。その結果を表 2 の OTHER 得票数 6-10 に示す。

表 2 格フレームごとの得票数別最多票 OTHER の正解率

OTHER 得票数	左に相当する格フレーム数	左中 OTHER 正解格フレーム数	OTHER の正解率%
10	4	4	100
9	4	4	100
8	8	7	87.5
7	19	9	47.4
6	39	9	23.1
5	77	16	20.8
4	128	20	15.6
3	218	41	18.8
2	309	52	16.8
1	329	41	12.5
0	184	23	12.5
Total	1319	226	17.1

まず、OTHER が最多票を取得し、なおかつ正解だった格フレームについて分析した。OTHER が正解となった理由としては次の 4 つがある：1) 述語が MWE の一部となっており MWE 全体で別のフレームを喚起する、2) JFN 上には 1 つの意味フレームのみ割り当てられていたが、実際には他にも語義があり KCF 格フレーム例文の用法はその別の語義に関するものであった、3) JFN に割り当てられた意味フ

レームかアノテーション例文に誤りがあった、4)提示文である KCF 格フレーム例文の意味がわかりづらかった、である(表 3)。それぞれの例を(4)から(7)に示す(文末のカッコ内の数字は最多票だった OTHER の得票数)。

(4) MWE

多くの人が会場で門前払いを【食べた】(10)

(5) 多義

有効の方がメモリを【食い】ません(10)

[Ingestion フレームに加え、Expend\_resource フレームへ割り当て必要]

(6) JFN の意味フレーム割り当て不正確

いつでも気持ちは「日に【新た】」です(6)

[Familiarity フレームではなく正しくは Age フレーム]

(7) KCF 例文意味不明

今日も一日が【大幅に】しないとね(7)

表 3 OTHER が正解だった理由の内訳

OTHER の得票数	OTHER 正解の格フレーム数	OTHER 正解の理由の内訳 (格フレーム数)			
		1)MWE	2)多義	3)JFN 不完全	4)KCF 例文意味不明
10	4	3	1		
9	4	2	1		1
8	7	3	4		
7	9	2	4	2	1
6	9			5	4
5	16		4	6	6
4	20		3	17	
3	41		7	34	
2	52		11	41	
1	41		14	27	
0	23		12	11	
Total	226	10	61	143	12

表 3 から分かる通り、MWE の関わる KCF 格フレーム例文は、最多票 OTHER が正解でなおかつ得票

数が7から10までの格フレームに限られた。これは先行研究で見られた傾向と同じである。これに対して、JFN上の意味フレームかアノテーション例文の誤りに関わるKCF格フレームは、OTHER得票数が少ない格フレームに多く見られた。

次に、OTHERが最多票を獲得した(つまり得票数が6から10までだった)が、正解ではなかった41格フレームについて調べた(表2参照)。これらの格フレームにおいてOTHERが不正解となった理由として考えられる顕著なものにKCF例文(提示文)とJFN例文(選択肢)の文構造の違いがある(14述語、計25格フレーム)。以下の例(8)-(10)に示すように、同一語義・用法であるにもかかわらず両者の文構造が違うためにクラウドワーカーらが同義と認識できず、OTHERを選択したと考えられる。

#### (8) 文末 vs. 名詞修飾

男のほうが顔を【重視】するのだろうか(KCF) ;  
【重視】すべき機能に応じた森林整備の推進(JFN)

#### (9) 文末 vs. て接続

アップグレードビジネスクラスを【比較】しよう！(KCF) ;  
商業借款と【比較】して一定の水準以上に優遇された貸付条件であることが必要とされている(JFN)

#### (10) 否定形 vs. 肯定形、能動態 vs. 受動態

パソコンがモデムを【認識】しなくなりました(KCF) ;  
大量破壊兵器やその運搬手段である弾道ミサイルの移転・拡散は、冷戦後の大きな脅威の一つとして【認識】され続けてきた(JFN)

以上、JFNに一義語として登録されている述語についても、最多票の回答がOTHERで得票数が多い格フレームにMWEが関与しているケースが多いことがわかった。さらに、MWE以外にもその述語の他の語義が見つかるケースがあることが判明した。また、クラウドワーカーらは提示文と選択肢の文構造を比較して回答しているらしいことが示唆された。

## 4 考察

今回の実験対象はJFN上で1つのJFNフレームに登録されている述語であったが、クラウドソーシング結果を評価・分析した結果、81述語のうちの9述語が多義語であることが判明した。そこで先行研

究に依い、述語の意味的特徴と、それらが現行のJFNに正確に反映されているかどうかに基づき、これら9つの多義語をカテゴリ-IからIIIに分類した。述語カテゴリ別に述語の平均正解率を算出したところ、以下の通りとなった。

カテゴリ-I: 2つの語義が明らかに隔たっている(2述語: 「食う」、「浮かぶ」): 100%

カテゴリ-II: 2つの語義が似通っている(2述語: 「教える」、「泳ぐ」): 50.95%

カテゴリ-III: JFN上に定義済みのJFN意味フレームが、当該述語のKCF格フレームに見られる全ての語義を正確に表現していない(5述語): 52.7%

1) JFNに誤り(2述語: 「記録」、「重ねる」): 28.1%

2) 3つ以上の語義(3述語: 「吸う」、「決定」、「飲む」): 69.1%

今回の実験においても2つの語義が明らかに隔たっておりJFNのコンテンツに不備のない述語(カテゴリ-I述語)は正解率が高いことが示された。ちなみに、一義語の72述語のうちJFN上の意味フレーム割り当てやアノテーション例文に不備のない59述語の正解率は63.2%以上100%以下に分布しており、不備のある残りの一義語(13述語)の正解率はそれより低い0%以上56.3%以下に分布していた。

## 5 まとめと今後の展望

JFN上で1つのJFN意味フレームに登録されている述語を対象に、KCF格フレーム例文にJFN意味フレーム例文を対応づけるクラウドソーシングを行い、その結果を正解率の向上とJFN改良の観点から考察した。得られた知見は以下の通りである。1) OTHERが正解となる理由は4つある。2) OTHERが最多票で得票数が多い場合は正解率が高く、MWEが関与している可能性が高い。3) 一義語・多義語ともにJFN上で不備のない述語の正解率が高い。4) クラウドワーカーらは回答の際にKCF格フレーム例文とJFNフレーム例文の構造を考慮していると推定できる。今後JFNに1例文のみ存在する一義語の実験の際にこれらの仮説をさらに検証していく。

また、多義語と分析した9述語の述語カテゴリへの分類では語義間の意味の隔たりを考慮したが、JFN意味フレームの粒度とJFN上のフレーム間関係に鑑み、多義語については集合知のみならず自動処理を用いて今後検討していきたい。

## 参考文献

1. クラウドソーシングによる日本語 FrameNet と自動構築した格フレームとの対応付け. 河原大輔, 小原京子, 関根聡, 乾健太郎. 言語処理学会第 24 回年次大会, 2018.
2. Linking Japanese FrameNet with Kyoto University Case Frames Using Crowdsourcing. Ohara, Kyoko Hirose, Daisuke Kawahara, Satoshi Sekine, and Kentaro Inui. Proceedings of LREC2018 International FrameNet Workshop 2018: Multilingual Framenets and Constructicons, 2018.