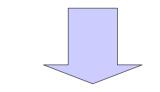
新聞記事の意味解析における 意味フレームの有効性についての考察

橋本祐樹 鈴木基之 任福継 (徳島大学)

本研究は、近年格フレームに代わる新たな知識表現方法として注目されている意味フレームの有効性を 検証する.格フレームと意味フレームの知識表現方法としての性能を比較し,両者の特性の違いを明ら かにするため,実際に新聞記事を用いて意味解析の実装を行い,これを検証した.

格フレームを用いた意味解析

入力文の深層格と対応する格フレームを決定



意味素の類似度等から意味的曖昧性を解消



意味素を決定

深層格の解析に失敗する可能性がある

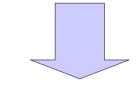
みんなが 母の退院を喜んだ。 (主語) (目的語) (動詞) ・母の退院・ 喜んだ。 みんな で (主語) (目的語) (動詞)

格助詞が省略されていたり異なっている

2. 意味フレームの利点

意味フレームを用いた意味解析

入力文の語句と意味フレームの コーパスの語句の類似度を計算



入力文に対応する意味フレームを決定



フレーム要素が決定

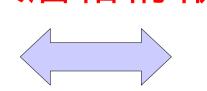
格助詞等の変化に頑健な解析が可能

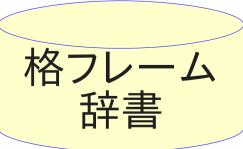


解析過程の違い

表層格情報

入力文

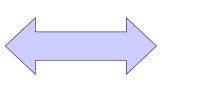




格フレームを用いた解析では 表層格情報を根拠にフレーム同定する

背景場面

入力文

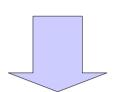




意味フレームを用いた解析では 背景場面を根拠にフレーム同定する

表層格情報を考える必要が無いので助詞や 語順等モダリティの問題に対して頑健

意味フレームが格フレームより入力文の モダリティに対して頑健な解析ができる事を検証する



実際の新聞記事の1文から動詞の主語を特定し その解析精度を比較する

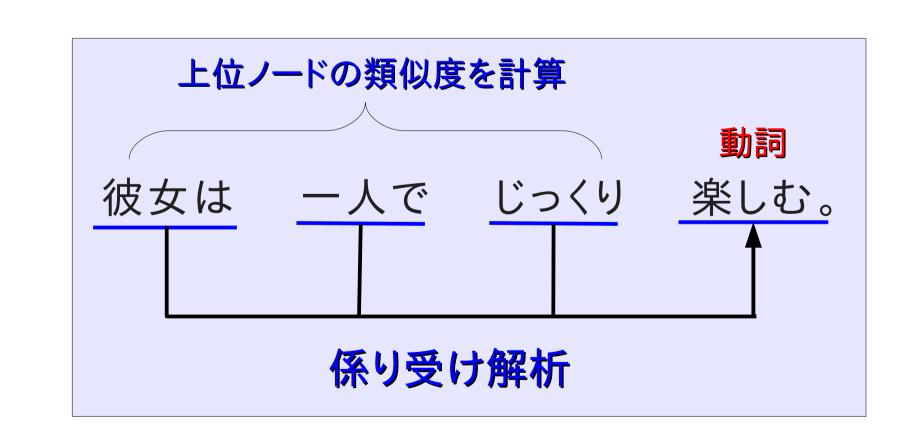
格フレームを用いた意味解析

- 入力文を係り受け解析
- 解析対象の動詞を検索 (喜ぶ,怒る,悲しむ,楽しむ)
- 動詞に係るノードから入力文中の格助詞を特定
- 「ハ格」「ガ格」に相当する語句を、動作主格とする



意味フレームを用いた意味解析

- 日本語フレームネット(小原ら2005)を使用
- 入力文を係り受け解析し、上位ノードの語句を 意味フレームのコーパスの語句と比較
- 最も類似度の高いものを Experiencer フレーム要素とする
- 意味フレームによる解析では, Experiencer sbj フレームの Experiencer フレーム要素に該当する 要素を主語とする



実験結果と考察

全入力文数···115文

格フレームによる解析結果

意味フレームによる解析結果

	成功	失敗
文の数	24	91
割合	20.8%	79.2%

	成功	失敗
文の数	34	81
割合	29.5%	70.5%

解析結果の成功例

夫婦そろってニセコでスキーを楽しむ。

結果) 格フレーム: そろっ 意味フレーム: 夫婦

原因)格助詞が明記されていないため 格フレームの同定に失敗している 〇〇社は, 航空機に乗らない人も 楽しめるようにする考えだ。

結果) 格フレーム: 空港 意味フレーム: 人

原因)動詞「楽しむ」の主語は「乗らない人」だが 係助詞「は」が「〇〇社」に係っているため

解析結果の失敗例

主な購入者は家庭で楽器を楽しむ。

結果) 格フレーム: 購入者 意味フレーム: 家庭

原因)意味フレームの類似度計算でコーパスに 収録されている語と最も似ているものを 指している. 家庭 - みんな(コーパス収録語)にマッチ

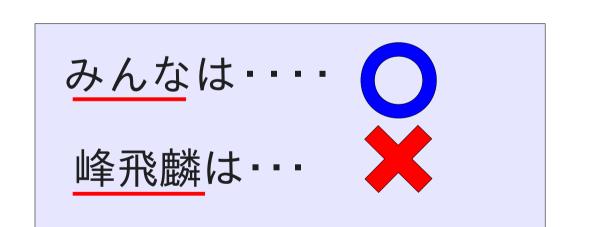
○○する××

食事を楽しむ夫婦・・・・ 主語 無事を喜ぶ人々・・・・

動詞 主語

- 全体の約39%を占める失敗例
- 係り受け解析の上位ノードのみを解析対象としたのが原因
- 格助詞は記述されないため、格フレームによる解析は不可

フレーム辞書コーパスとの類似度計算



珍しい言葉では 入力文の語との類似度が高くならない

- 類似度計算の際には比較対象の性質が重要
- 特殊な用法や稀な用法のコーパスは 類似度計算の精度を著しく低下させる
- "よくある"言葉の使われ方が望ましい

まとめ

- 本研究は意味フレームによる解析のモダリティの 問題に対する頑健性に着目しそれを検証した。
- 実験は、日本語フレームネットを用いて新聞記事 の解析を行い、動詞の主語を特定して、その精度を 意味フレームと格フレームの場合で比較した.
- その結果, 従来行われていた表層格情報から格フ レームを同定するよりも、意味的類似度計算を行って 意味フレームを特定する方が高精度だった。
- 以上より、意味フレームによる解析が格フレーム を用いた従来手法よりもモダリティの問題に対して 頑健であることが確認できた。