



言語に基づく推論のための 事象間関係知識データベースの構築

奈良先端科学技術大学院大学 情報科学研究科
松吉 俊、村上 浩司、乾 健太郎、松本 裕治

背景

- 高度な言語情報処理能力を実現するためには、辞書や文法などの言語知識の他に、**大量の世界知識が必要**である

太郎が読書に熱中する。

<同義>

太郎が読書にふける。

国民は新しい制度に賛成している。

<反義>

国民は新しい制度に反対している。

三郎が四郎に真実を告げた。

<結果>

四郎が真実を知った。

自分のスペースが広がる。

<結果>

自分のスペースが広い。

背景

- 高度な言語情報処理能力を実現するためには、辞書や文法などの言語知識の他に、大量の世界知識が必要である

太郎

三郎、四郎: 固有名詞-人名-名
真実: 一般名詞

国
賛

(類義語: 事実、実態、現実、真理)
告げる: 動詞-一段動詞 (他動詞)

読書にふける。

新しい制度に
ている。

三郎が四郎に真実を告げた。

<結果>

四郎が真実を知った。

自分のスペースが広がる。

<結果>

自分のスペースが広い。

背景

- 高度な言語情報処理能力を実現するためには、辞書や文法などの言語知識の他に、大量の世界知識が必要である

太郎が読書に熱中する。

<同義>

太郎が読書にふける。

国民は新しい制度に賛成し

X ガ Y ヲ Z ニ告げる

<結果>

国民は新しい制度に

Z ガ Y ヲ知る

三郎が四郎に真実を告げた。

<結果>

四郎が真実を知った。

自分のスペースが広がる。

<結果>

自分のスペースが広い。

本研究の目的

- 以下のような関係知識を収集し、
自然言語に基づく推論を行なう基盤となる
事象間関係知識データベースを構築する

太郎が読書に熱中する。

太郎が読書にふける。

- X が Y に熱中する ↔ X が Y にふける
<同義>

国民
賛成

- X が Y に賛成する ↔ X が Y に反対する
<反義>

三良

- X が Y を Z に告げる → Z が Y を知る
<結果>

- X が広がる → X が広い
<結果>

自分のスペースが広がる。

<結果>

自分のスペースが広い。



事象(コト、event)

- 事象を表す言語表現の例:
 - 雨(= 雨が降る)
 - 彼が友人に手紙を送る (述語項構造)
 - 酒を飲みすぎて酔っぱらう (述語項構造の組み合わせ)
- 本研究では、事象を表す表現の単位として、**述語項構造**を用いる
 - 利点1: 新しい概念に対する頑健性
 - 利点2: 多様な言語表現の標準化
 - 利点3: 述語項構造解析技術の利用



関連研究

■ 英語

- WordNet (Fellbaum 1998)、
FrameNet (Baker et al. 1998)、
VerbOcean (Chklovski et al. 2004)

■ 日本語

- 日本語WordNet (Bond et al. 2008)
- 動詞語釈文構造化データ(大西ら 2008)
- 動詞項構造シソーラス(竹内ら 2008)
- コーパスからの知識獲得手法(Inui et al. 2005;
Torisawa 2006; 阿部ら 2008)

事象間関係知識データベース

	項構造		関係知識		
	基本	下位事象を含む	シソーラス (類義・反義)	二項関係	その他
動詞	約1万4千語 [29,555語義] (大西ら 2008; 竹内ら 2008)	約600語 [728語義] (2008)	9,135語義 のシソーラス (竹内ら 2008)	45,905件 (大西ら 2008)	
形容詞	約700語 [1,035語義] (2008)			2,071件 (2008)	
形容動詞	約2千語 [2,474語義] (2008)			4,746件 (2008)	
複合用言 (慣用句等)	約2千語			約2千8百件	
事象名詞					

事象間関係知識データベース

■ 対象品詞:

- 動詞、形容詞、形容動詞、
- 複合用言(「役に立つ」、「二進も三進も行かない」など)、
- 事象名詞(「雨」、「(からの)電話」など)

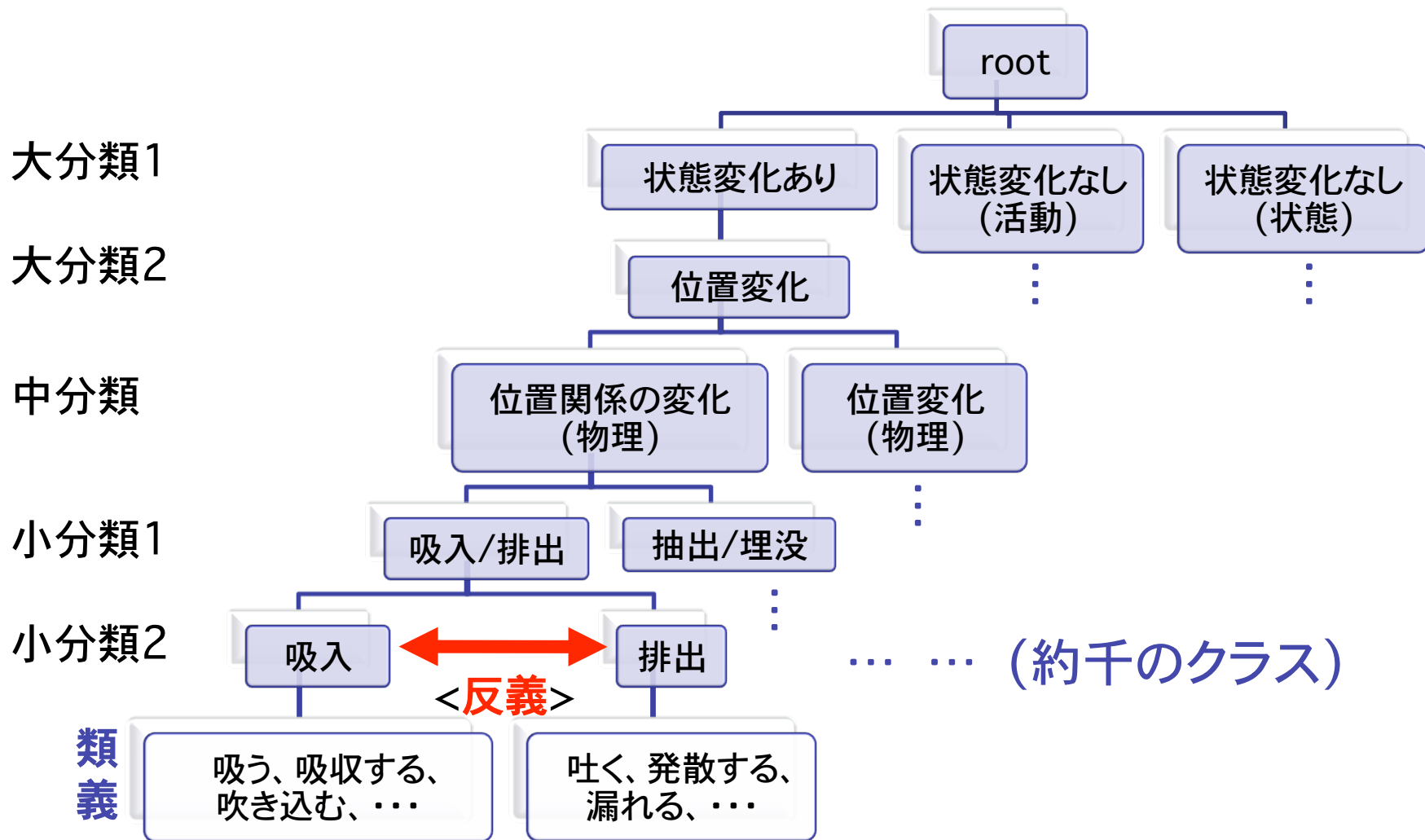
■ 項構造:

- 基本項構造(ガ、ヲ、ニ、デ、ト、カラ、…)
- 下位事象を含む項構造
(部長が部下にノルマを達成することを命令した)
- 格交替(「友人に会う」→「友人と会う」、
「することを」→「するように」、「しろと」)

■ 関係知識:

- シソーラス(類義関係、反義関係)、二項関係
- その他(順序関係、必要十分関係、部分全体関係など)

動詞項構造シソーラス(竹内ら 2008)



二項関係(大西ら 2008)

- 国語辞典の語釈文から人手で関係知識を抽出

見出し語: 倒す

立っている物に力を加えて傾け、横にする。

- 
- X ガ Y ヲ倒す → Y ガ立つ <前提条件>
 - X ガ Y ヲ倒す → X ガ Y ニ力ヲ加える <手段>
 - X ガ Y ヲ倒す → X ガ Y ヲ傾ける <手段>
 - X ガ Y ヲ倒す → X ガ Y ヲ横にする <同義・上位>

- 9種類の関係:

- 同義(言い換え)、同義・上位、反義語、前提条件、結果(状態)、付帯状況、不可分、目的、手段



事象間関係知識データベース

	項構造		関係知識		
	基本	下位事象を含む	シソーラス (類義・反義)	二項関係	その他
動詞	約1万4千語 [29,555語義] (大西ら 2008; 竹内ら 2008)	約600語 [728語義] (2008)	9,135語義 のシソーラス (竹内ら 2008)	45,905件 (大西ら 2008)	
形容詞	約700語 [1,035語義] (2008)			2,071件 (2008)	
形容動詞	約2千語 [2,474語義] (2008)			4,746件 (2008)	
複合用言 (慣用句等)	約2千語			約2千8百件	
事象名詞					

本データベースを用いた推論

三郎が真実を四郎に教える

X ガ Y ヲ Z ニ教える $\xrightarrow{\text{手段}}$ X ガ Y ヲ Z ニ告げる

X ガ Y ヲ Z ニ告げる $\xrightarrow{\text{結果}}$ Z ガ Y ヲ知る

四郎が真実を知る

類義

X ガ Y ヲ Z ニ言う

X ガ Y ヲ Z ニ言いふらす

X ガ Y ヲ Z ニ打ち明ける

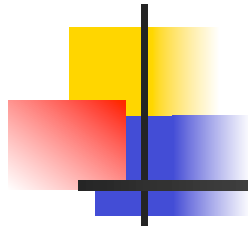
⋮

伝達



まとめ

- 自然言語に基づく推論のための事象間関係知識データベースを作成した
 - 語義単位で、動詞29,555エントリー
 - 9,135語義の動詞項構造シソーラス
 - 45,905件の事象間関係知識



今後の課題

	項構造		関係知識		
	基本	下位事象を含む	シソーラス (類義・反義)	二項関係	その他
動詞	約1万4千語 [29,555語義] (大西ら 2008; 竹内ら 2008)	約600語 [728語義] (2008)	9,135語義 のシソーラス (竹内ら 2008) <拡張予定>	45,905件 (大西ら 2008) <精錬中>	統計的 手法により獲得された大規模な関係知識の整理
形容詞	約700語 [1,035語義] (2008)		<上記の拡張の一環>	2,071件 (2008)	
形容動詞	約2千語 [2,474語義] (2008)		<上記の拡張の一環>	4,746件 (2008)	
複合用言 (慣用句等)	約2千語 <精錬予定>		<上記の拡張の一環>	約2千8百件 <精錬予定>	
事象名詞	<収集予定>				

実テキストに対する関係認識精度評価