

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 16 日現在

機関番号：14603

研究種目：研究活動スタート支援

研究期間：2015～2016

課題番号：15H06408

研究課題名(和文) 実用的な話し言葉の解析技術に関する研究

研究課題名(英文) A Study of Practical Spoken Language Processing

研究代表者

吉野 幸一郎 (Koichiro, Yoshino)

奈良先端科学技術大学院大学・情報科学研究科・助教

研究者番号：70760148

交付決定額(研究期間全体)：(直接経費) 2,300,000円

研究成果の概要(和文)：本研究課題においては話し言葉における意味構造、特に述語項構造の解析技術について、後方参照などの話し言葉特有の現象を考慮した解析器の研究開発を行った。まず、話し言葉で明示的に解析することが難しい係り受け情報を無くした場合、どの程度の影響があるのかについての調査を行った。構築したシステムでは点予測を用い、後方参照については特定のラベルを導入することで、入力文が途中であっても述語項構造解析を行い、ユーザの話した内容だけでなく、これから話す内容が存在するかどうかを予測しつつ解析を行うことができる。また、実際の対話コーパスにおいてこうした現象がどの程度発生するのかを検証するため、対話コーパスを構築した。

研究成果の概要(英文)：We developed a method of semantic analysis, especially for predicate argument structure on spoken language. The proposed method considers backward reference, a characteristic of spoken language to robustly process spoken languages. We developed a predicate argument structure analyzer that does not use dependency information of words to verify the influence of lack of dependency information, which is difficult to extract from spoken sentences. We used pointwise method to build the system to analyze sentences, even if it is unfinished sentence, by introducing backward reference label. This architecture make it possible to predict not only structure of spoken contents by the user but also existences of unseen future content words. We also developed a dialogue corpus to confirm the phenomena of backward reference.

研究分野：音声対話

キーワード：述語項構造 対話システム

1. 研究開始当初の背景

対話ロボットなどの研究において、実際に人間が発話を行ってシステム・ロボットを操作するために高精度な話し言葉音声の認識と解析が必要とされている。特に述語を中心とした意味を解析する述語項構造解析などの意味解析は、対話・検索などの言語を用いたアプリケーションにおける基盤的な技術であり、これら精度はそのままアプリケーションの精度に直結する。しかし、これまでこうした解析技術は書き言葉を対象としたものが中心であり、対話システムなどのアプリケーションで用いられるような話し言葉に対しての研究は少ない。また、近年人間の言語発生にあわせてリアルタイムに処理を行うアプリケーションが提案されつつあり、こうしたアプリケーションに対応する意味解析器が必要であった。

2. 研究の目的

言語の意味理解は言語処理における究極の目的であり、対話システム、翻訳システムなど多様なシステムがこうした言語処理による意味解析をベースとしている。しかし、これまで言語処理が扱ってきた書き言葉とこれらのアプリケーションが利用する話し言葉では、構造・現象などにギャップがあり既存のシステムをそのままアプリケーションの入力とすることは困難であった。本研究では、今まで特に書き言葉に対して行われてきた自然言語処理技術、特に述語と項の関係を解析する述語項構造解析などの意味解析の技術を、話し言葉に最適な解析モデルとして再構築し、音声対話などの実アプリケーションにおいて話し言葉の意味を解析・利用できるシステムの構築を目指す。

3. 研究の方法

本研究では、書き言葉と話し言葉、具体的には既存の新聞記事と対話コーパスを比較し、どのような現象の違いがあるかを考察した上で話し言葉のための解析モデルを構築する。本研究が取り扱う範囲としては、対話などの応用研究で必要とされる述語項構造と、音声認識結果から得られる形態素区切りについてである。単語同士の依存関係を解析する係り受け解析は、既存の研究では述語項構造解析の前段階として考えられてきたが、これが本当に必要かどうかについては議論の余地がある。そこでまず、係り受け関係を明示的に用いて述語項構造解析を行う場合と、係り受け関係を用いずに点予測によって項構造解析を行う場合の精度差について評価した。実際の話し言葉で綺麗な構造を得ることが難しい係り受け関係を利用せず、実際の話し言葉に対してより直接的に述語項構造を解析することを考える。

また、実際に人間が行っている意味処理を考慮した場合、どのように話し言葉を処理しているかについても考慮しモデルを構築する。

人間が対話している間に係り受け構造を考えるかということではなく、むしろ係り受け構造は書き言葉を書く際の内省の中で構成されるのではないかと考えられる。こうした仮説のもと、より人間に近い述語項構造解析のモデルを構築することで、実際に人間が意味の処理をどう行っているのかという疑問に対する答えも考察する。

係り受け構造の項構造解析に対する影響

本調査では、係り受け構造が実際にどの程度述語項構造解析の精度に対して影響を及ぼすかについて検証する。点予測に基づく述語項構造解析器を構築し、ここに係り受け構造から得られる素性を明示的に加味するかどうかによってどの程度精度が変化するかどうかの検証を行った。

実際の対話中の発話に出現する述語項構造の調査

本調査では、実際に人同士の対話にどのような述語項構造の現象が生じるかを調査する。具体的には人同士の対話を収録・書き起こしてコーパスを整備し分析する。この際、実際に人間の意味理解がどう行われているかをモデル化するため、実際に話し言葉を聞いて処理するプロセスにおいてどのような処理を行っているのかを、被験者に実際の事例の音声を聞きながら内省してもらい、アンケートを取る。この結果から認定基準を決定し、コーパスを構築していくという手法を取る。

人間の発話モデルに従った解析モデルの設計・構築

点予測に基づく漸進的な項構造解析器を構築する。後方参照などの話し言葉特有の現象を考慮した解析器の研究開発を行った。

4. 研究成果

まず、係り受けを利用しない述語項構造解析器として、点予測によるものを構築した。具体的には、図1に示すような点予測に基づく項構造解析器を構築した。この解析器におい

ID	S-ID	word	(translation)	Tag from classifiers (arg_id:tag:prob.)
1	1	東京	(Tokyo)	
2	1	上野	(Ueno)	
3	1	の	(function word)	
...				
21	2	ユリカモメ	(pewit gull)	
22	2	。	(.)	
23	3	釣り糸	(fishline)	
24	3	を	(function word)	
25	3	引っ掛け	(tangle)	1:ga_inter:0.648 , 21:ga_inter:0.866, 23:wo_intra:0.993
...				
Decides one candidate for one label				↓
				21:ga, 23:wo

Result: pew it gull (subject), fishline (object), tangle

図 1. 構築した述語項構造解析器の枠組み

では、文内・文外とガ・ヲ・ニ格それぞれの組み合わせ6種類に対して、述語と格要素候補が項構造の関係にあるかどうかの二値分類問題として解析を行った。各分類器は回帰として解き、そこから得られる事後確率の値によって、述語のそれぞれの項にあたる格要素がどれであるかの推定を行った。

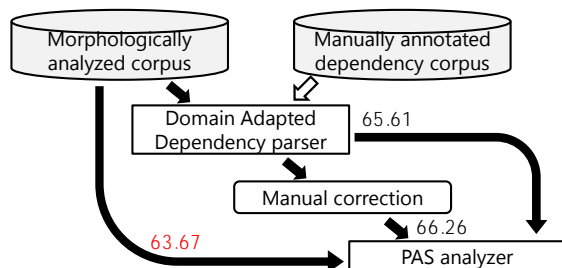


図 2. 係り受け解析結果の影響

この解析器に対して様々な素性を入力として分析を行った。コーパスとしてはNAISTテキストコーパスを用い、先行研究[Iida 2007]に従って学習・開発・テストにコーパスを分割した上で、テストセット上の解析精度によって評価を行った。この結果、図2に示すように、係り受け解析の自動解析結果を用いた場合は65.61%、理想的な正解の係り受け解析結果を用いた場合は66.26%の精度になるのに対し、係り受け解析結果を用いない場合の精度は63.67となり、係り受け解析器をドメインに対して構築するコストを考慮すると十分な解析精度が得られることが確認された。

構築したシステムでは点予測を用いているが、これに対して後方参照については特定のラベルを導入することで、入力文が途中であっても述語項構造解析を行い、ユーザの話した内容だけでなく、これから話す内容が存在するかどうかを予測しつつ解析を行うことができるようにデザインを変更した。特に後方参照は、書き言葉を対象とするNAISTテキストコーパスにおいても297,070件中6,337件、より話し言葉に近いBCCWJにおいても12,029件中3,148件が存在することがわかり、一定の考慮すべき量であることがわかる。これらの結果を踏まえ、実際に収録・書き起こしを行った対話データに対して述語項構造のアノテーションを行っている。このデータ整備は、本研究課題で得られた知見をベースとしてJST さきがけ・研究領域：新しい社会システムデザインに向けた情報基盤技術の創出において新たに研究課題として採択され、継続的にアノテーション・調査・解析器の精度向上を続けている。

参考文献

[Iida 2007] Ryu Iida, Kentaro Inui, and Yuji Matsumoto. 2007. Zero-anaphora resolution by learning rich syntactic pattern features. *ACM*

Transactions on Asian Language

Information Processing (TALIP), 6(4):12:1-

12:22.

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

{ 雑誌論文 } (計 0 件)

{ 学会発表 } (計 4 件)

1. Seokhwan Kim, Luis Fernando D'Haro, Rafael Enrique Banchs, Jason D. Williams, Matthew Henderson, Koichiro Yoshino.

The Fifth Dialog State Tracking Challenge.

2016 IEEE Workshop on Spoken Language Technology (SLT 2016), San Diego, California, USA

2. Koichiro Yoshino, Naoki Hirayama, Shinsuke Mori, Fumihiko Takahashi, Katsutoshi Itoyama and Hiroshi G. Okuno
Parallel Speech Corpora of Japanese Dialects

10th edition of the Language Resources and Evaluation Conference (LREC). pp.4652--4657, Portoroz, Slovenia. May 2016

3. Koichiro Yoshino, Takuya Hiraoka, Graham Neubig, Satoshi Nakamura.
Dialogue State Tracking using Long Short Term Memory Neural Networks
7th International Workshop on Spoken Dialog Systems (IWSDS). Riekkonlinna, Finland. January 2016

4. 吉野 幸一郎, 平山 直樹, 森 信介, 高橋 文彦, 糸山 克寿, 奥乃 博.
日本語方言における音声対訳コーパスの構築

言語処理学会第22回年次大会 (NLP2016), 東北大学(宮城県仙台市), 2016年3月

{ 図書 } (計 0 件)

{ 産業財産権 }

出願状況 (計 0 件)

取得状況 (計 0 件)

{ その他 }

なし

6. 研究組織

(1)研究代表者

吉野 幸一郎 (YOSHINO, Koichiro)

奈良先端科学技術大学院大学, 情報科学研究科, 助教

研究者番号 : 70760148