

話し言葉解析に基づく話者知識の自動獲得

Speaker's Knowledge Acquisition based on Spoken Dialogue Analysis

小野貴博*1 Takahiro Ono 大野誠寛*1 Tomohiro Ohno 松原茂樹*2 Shigeki Matsubara 山口由紀子*2 Yukiko Yamaguchi 河口信夫*2 Nobuo Kawaguchi 吉川正俊*2 Masatoshi Yoshikawa

*1 名古屋大学大学院情報科学研究科
Graduate School of Information Science, Nagoya University

*2 名古屋大学情報連携基盤センター
Information Technology Center, Nagoya University

Conversation is one of the most effective ways for exchanging information, and it is very worth to share and reuse the speaker's knowledge. This paper proposes a method of information extraction from spoken conversations. The method extracts task-oriented information from spoken utterances and acquires speaker's knowledge by integrating them. The utterance intention tags and dependency structures are used for selecting the utterances to be processed and integrating the information linguistically, respectively. An extraction experiment using CIAIR in-car speech corpus achieves a precision of 85.7% and a recall of 64.2%, and has shown the method to be effective.

1. はじめに

自然言語における情報抽出とは、言語データから特定の情報を取り出す言語処理技術であり、大量に存在するデータを効果的かつ効果的に活用するための基盤技術として注目されている[4]。情報抽出に関する従来の研究は、新聞記事や科学技術論文、Web 文書などのテキストを対象としているものがほとんどであるが[5]、講義、講演、会議、問い合わせなどの音声もまた、参照価値のある情報を含み、大量に生産される言語データである。

そこで本稿では、音声対話データを対象とした情報抽出について述べる。対話は、人間にとって最も手軽な情報伝達手段であり、そこには話者が持つ情報や知識が現れる。本研究では、タスク指向対話から、タスク達成のために伝えられる情報を抽出する。抽出した情報を整備することにより、話者知識の共有や活用が可能になると考えられる。

本手法は、対話を構成する各発話からの情報抽出と対話全体からの情報抽出の二段階で実行される。すなわち、発話文の係り受け構造から抽出規則により情報を抽出し、各発話から得られた情報を統合することにより対話全体からの情報を得る。

名古屋大学 CIAIR 車内音声対話コーパスに収録されている飲食店案内をタスクとする 20 対話を用いて、情報抽出実験を実施し、本手法の有効性を確認した。

2. 対話データからの話者知識獲得

本研究では、対話のタスクに関連する情報を、対話参加者の知識として抽出する。本稿の以下では、飲食店案内をタスクとする対話データからの情報抽出について述べる。

飲食店案内対話の例を図 1 左に示す。この対話は、情報を提供される話者 (A) と情報を提供する話者 (B) の間で遂行され、付近にある飲食店の情報が話者 B から話者 A に音声で伝達される。これらの情報は話者 B の持つ知識であり、書き起こしテキストからの情報抽出により獲得できる。

飲食店に関する情報 (以下、店情報) を、

- 店名: 飲食店の名称
- ジャンル: 料理名, ジャンル名
- 予算: 金額, 金額の範囲
- 駐車場: 駐車場の有無

連絡先: 小野貴博, 名古屋大学大学院情報科学研究科社会システム情報学専攻吉川研究室, 〒464-8601 名古屋市千種区不老町, TEL: (052)789-1532, ono@dl.itc.nagoya-u.ac.jp

A: ラーメンがちょっと食べたいんだけどどこかないかしら	店名	金龍ラーメン
B: はいこの近くですと該当するお店が三軒ございます	ジャンル	ラーメン, ギョーザ
B: ラーメンギョーザの金龍ラーメンさっぽろ亭春帆亭で	駐車場	有
ございます		
A: さっぽろ亭お願いします	店名	さっぽろ亭
B: はいさっぽろ亭ですと駐車場がございませんが	ジャンル	ラーメン
よろしかったでしょうか	駐車場	無
A: じゃ駐車場があるところをお願いします		
B: はい金龍ラーメンと春帆亭ですと駐車場がございますが	店名	春帆亭
A: じゃ金龍ラーメンをお願いします	ジャンル	ラーメン
B: はいそれでは金龍ラーメンへご案内いたします	駐車場	有
A: はい		

飲食店案内対話の例

抽出される店情報

図 1: 飲食店案内対話からの情報抽出

● 予約: 予約の必要性, 及び, 予約の可能性
の 5 種類とすると, 図 1 左の対話から抽出できる情報は図 1 右に示す通りとなる。このような情報は, 共有し再利用することによりその有用性が高まる。

3. 対話データからの情報抽出

本手法では, 各発話に係り受け解析及び意図推定が行われた対話データを入力とする。発話文の係り受け構造に対して, 発話意図の種類ごとに定めた抽出方法により情報抽出を行う。

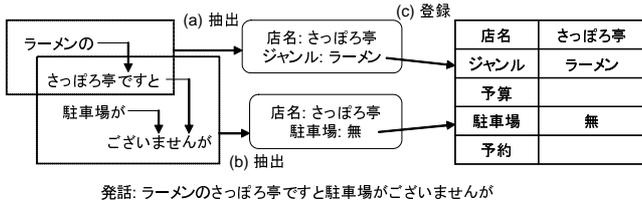
3.1 意図に基づく発話分類

対話データには, 抽出すべき情報を含む発話と含まない発話が存在する。発話が抽出すべき情報を含むかの判断は, 発話の意図を考慮することにより可能となる。たとえば, 飲食店案内対話では, 店情報を提示する発話, 及び, 検索条件を提示する発話において, 抽出すべき店情報が現れる。

本研究では, そのような発話の選択に名古屋大学 CIAIR で設計された階層的意図タグ (layered intention tag) を用いる。タグ付けは, コーパスに基づく自動推定により高い精度で実現できる [2]。この階層的意図タグは, 談話行為, 動作, 対象, 詳細の 4 階層で構成されており, 本研究では, 詳細を除く 3 階層を用いる。飲食店案内対話では 33 種類の意図タグが出現し, このうち, 抽出対象の発話に付与されるタグは以下の 5 種類である。

- 「店情報を提示する発話」に関する意図タグ
陳述+提示+店情報, 陳述+提示+検索結果, 提案+予約+店
- 「検索条件を提示する発話」に関する意図タグ
依頼+検索+店, 依頼+再検索+店

抽出規則は, これらのタグをもつ発話に対してのみ適用する。



発話: ラーメンのさっぽろ亭ですと駐車場がございませんが

図 2: 発話からの情報抽出



発話: ラーメンの春帆亭が ございます
発話: 春帆亭は予約が可能と なっております

図 3: 店情報データの統合

3.2 発話文からの情報抽出

発話文の係り受け構造に抽出規則を適用し情報を抽出する。図 2 に発話文からの情報抽出の例を示す。店内対話からの店情報の抽出のために、以下の 4 つの規則を作成した。

- 店名、ジャンル表現があれば、店名情報、ジャンル情報としてそれぞれ抽出する。ただし、ジャンル表現を含む文節と店名を含む文節との間に係り受け関係があるとき、それらを併せて抽出する (図 2(a))。
- 値段表現 (「数字+円」など) があれば、直後の範囲表現と併せて予算情報として抽出する。
- 駐車場表現を含む文節の受け文節から駐車場情報を抽出する。ただし、その文節と店名を含む文節との間に係り受け関係があるとき、それらを併せて抽出する (図 2(b))。
- 予約表現を含む文節の受け文節から予約情報を抽出する。ただし、その文節と店名を含む文節との間に係り受け関係があるとき、それらを併せて抽出する。

抽出した情報をもとに店情報データを作成する (図 2(c))。一つの発話から複数の店に関する情報が抽出されたときは、店ごとに店情報データを作成する。

3.3 抽出した情報の統合

対話データからの店情報の抽出は、発話文から生成された店情報データを統合することにより実現する。データの統合は、第 1 発話からの店情報データから順に、発話の種類に応じて以下の処理を実行する。

- 「店情報を提示する発話」から抽出した店情報データ 店名が同一の店情報データがすでにあればそれと統合する。
- 「検索条件を提示する発話」から抽出した店情報データ 直後に現れる「店情報を提示する発話」から抽出した店情報データと統合する。

店情報データの統合例を図 3 に示す。

4. 実験と評価

4.1 実験の概要

本手法の実現可能性を評価するために、抽出実験を行った。実験には、名古屋大学 CIAIR 車内音声対話コーパス [1] に収録された飲食店案内をタスクとする 20 対話データ (210 発話) を使用した。意図タグは人手で付与し、係り受け解析は話し言葉用解析器 [3] を用いた*1。本手法を正解データに対する精度及び再現率により評価した。すなわち、

$$\text{精度} = \frac{\text{正解した店情報数}}{\text{抽出した店情報数}}, \text{再現率} = \frac{\text{正解した店情報数}}{\text{正解データの店情報数}}$$

*1 解析器の本コーパスに対する解析精度は 87.0% である [3]。

表 1: 正解データの内訳

抽出項目	店名	ジャンル	予算	駐車場	予約
個数	58	72	6	10	13

表 2: 実験結果

	本手法	手法 1	手法 2	手法 3
精度	85.7% (102/119)	84.7% (100/118)	78.4% (91/116)	77.6% (90/116)
再現率	64.2% (102/159)	62.9% (100/159)	57.2% (91/159)	56.6% (90/159)

である。ここで、抽出された店情報が正解であるとは、正解データと同一の値が同一の店情報データにあることとする。正解データの店情報は 159 個であった。その内訳を表 1 に示す。

4.2 実験結果

本手法による抽出性能を、係り受け構造を用いない方法 (手法 1)、発話意図を用いない方法 (手法 2)、及び、いずれも用いない方法 (手法 3) と比較評価した。結果を表 2 に示す。精度、再現率とも本手法による結果が最も良好であり、その効果を確認した。

係り受け構造を使用することにより、例えば、発話文「イタリア料理のキャナリイロウ、ハンバーグの気楽家がございます」から、店名情報とジャンル情報を正しく組み組み合わせることができた。また、発話意図については、抽出すべき発話文の選択だけでなく、店情報データの統合に関しても効果的に機能することが明らかになった。

本手法では、57 個の店情報については正しく抽出できなかった。その第一の理由は、形態素解析や文節まとめ上げの誤りであり、全体の 64.9% (37/57) を占めている。その他の問題として、別の話題に関する情報と誤って統合したり、代名詞の指示対象が特定できないことが挙げられる。

5. まとめ

本稿では、対話データからの情報抽出により話者の知識を獲得する手法を提案した。本手法では、発話文の係り受け構造に抽出規則を適用することにより、発話に含まれる情報を抽出し、それらを統合することにより、会話全体からの情報としてまとめ上げる。名古屋大学 CIAIR 車内対話コーパスを用いた店情報抽出実験により、本手法の有効性を確認した。

今後は、対話構造を考慮した情報の統合手法を検討するとともに、大規模な対話データを用いて実験を行う予定である。

参考文献

- [1] Kawaguchi, N., Matsubara, S., Takeda, K., Itakura, F.: CIAIR In-Car Speech Corpus -Influence of Driving States-, IEICE Trans. on Inf. and Sys., E88-D(3), pp.578-582 (2005).
- [2] Irie, Y., Matsubara, S., Kawaguchi, N., Yamaguchi, Y., and Inagaki, Y.: Speech Intention Understanding based on Decision Tree Learning, Proc. of ICSLP-2004 (2004)
- [3] Ohno, T., Matsubara, S., Kawaguchi, N., and Inagaki, Y.: Robust Dependency Parsing of Spontaneous Japanese Spoken Language, IEICE Trans. on Inf. and Sys., E88-D(3), pp.545-552 (2005).
- [4] 関根聡: 情報抽出 -情報を整理して提示する-, 情報処理, Vol.45, No.6, pp.563-568 (2004)
- [5] S. Miller, M. Crystal, H. Fox, L. Ramshaw, R. Schwartz, R. Stone, R. Weischedel, and the Annotation Group: Algorithms that learn to extract information; BBN: Description of the SIFT system as used for MUC-7, Proc. of the MUC-7 (1998)