

# 情報抽出と述語項の類似度を利用した音声対話システム

吉野 幸一郎

森 信介

河原 達也

京都大学 情報学研究科

## 概要

本稿では、日々更新される Web テキストに対して、述語項構造に着目した情報抽出を行い、その内容を扱う音声対話システムについて述べる。提案する対話システムは、述語項構造解析により抽出された情報に基づいて、対話の履歴を用いながら、ユーザの質問に対応する情報や、ユーザの質問に近い内容を含む文を応答として提示する。ドメインごとに有用な述語項構造のパターンを抽出するために Naive Bayes に基づく重要度を用いる。さらに、述語項構造における要素・用言における関連度を定義することによって、柔軟な情報検索を可能とする。評価実験の結果、音声情報検索における従来手法である Bag of Words (BOW) モデルと、BOW モデルに対して系列を考慮した Sequence of Words (SOW) モデルと比べて、的確な応答が実現できることが示された。

## 1 はじめに

近年、Web 上に集積する情報は爆発的に増加しており、それらを活用する機会が増大しているが、こうした情報へのアクセスは、現在はキーワード型検索が主である。これに対して、ユーザの漠然とした情報要求、自然文での検索に対する要求があるが、そうしたユーザの意図・嗜好を対話的に顕在化しながら情報を提示するシステム（情報コンシェルジュの研究 [11]）が行われている。

これまで研究・実用化されてきた音声対話システムはおおむね 2 種類に分類される [10]。フライト情報案内 [1, 9] やバスの運行案内 [6] などの明確なタスクを定義し、関係データベース (RDB) をバックエンドとした枠組みは、タスク達成に必要な意味表現の定義や対話のフローの記述が容易であった反面、Web などの大規模なテキスト情報に対して適用することが困難であった。それに対して、一般的な文書検索を用いた対話システムの研究 [12, 13] も行われてきたが、表層的なキーワードや係り受け関係、質問タイプなどのみに着目し、深い言語的解析や対話処理は扱われていない。その結果、対話の文脈やユーザの要求とは無関係な、不自然な応答が生成されることがあった。

これに対して本研究では、述語項構造に着目した情報抽出を行うことで、RDB のような構造を持たない

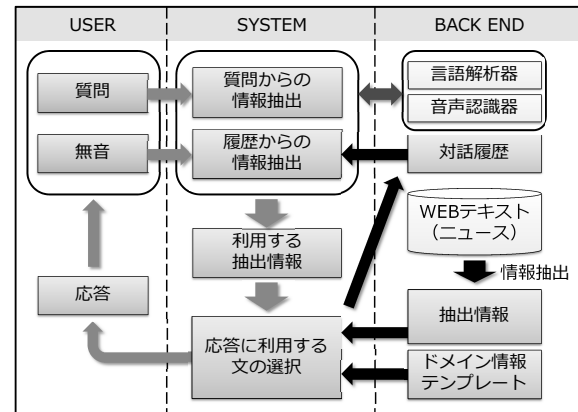


図 1: 対話システムの概要

Web 文書を扱いながら、その意味表現を扱えるシステムを構築する。文の係り受け関係を扱い、適切な文書を段階的に検索するシステムが先行研究として研究されている [15] が、本研究では、文中の述語項構造に着目した情報検索を行う。また、ユーザの質問に対して適切な応答を見つけれなかった場合、述語項同士の類似度を利用して、類似した情報の提示を行う。このような対話を行う上で有用な情報構造はドメインに依存しており、そのような情報構造のテンプレートを作成する必要性が指摘されている [8]。しかし、ドメインごとに人手でテンプレートを作成する手法は、Web に存在する様々なドメインに対して自動的に適用できない。そこで、本研究ではパーザを用いた述語項構造解析の結果から、自動で特定のドメインにおける重要な情報構造を抽出する [3, 4]。要素や用言に対して、コーパスから獲得した関連度を用いて、関連した情報の提示を行う。これによって、的確な検索及び推薦ができることを示す。

## 2 対話システムの概要

本システムは、Web に存在するニュースサイトなどのテキスト情報を用いて、ユーザの質問に答えながら対話を行う。日々更新される Web コンテンツを応答に利用することができる。異なるドメインに容易に拡張できるように設計を行っている。

本システムの構成を図 1 に示す。まず、システムは事前に Web から得たテキストに対して、述語項構造

解析を用いた情報抽出を行う。対話中のユーザ発話に対しても同様の解析を行い、抽出した情報間のマッチングを行うことで、ユーザの要求に最も関連の深いテキストを検索し、応答を生成する。

### 3 述語項構造解析に基づく情報抽出

テキストから情報抽出を行い、その意味表現を捉えた応答を生成するためには、意味表現の単位と、それを表現する形式を定義しなければならない。本研究では、述語項構造を用いる。述語項構造は、「要素 - 格用言」の関係性に基づいて意味表現を表すものである。述語項構造は、古くから自然言語処理で利用されており、古典的な対話システムにおいては、必要な述語項構造を手で定義したものが一般的に利用されていた。近年統計的手法による大規模で一般的な述語項構造解析の研究 [2] が進んでおり、本研究では、述語項構造解析のパーザとして JUMAN/KNP<sup>1</sup>を用いる。

述語項構造解析によって、要素とその係り先である用言、その関係性として格を抜き出す。この際、一つの用言に対して複数の要素と格が存在するが、その組み合わせもあわせて保持する。日本語の場合は自明な場合の主格が省略される傾向にあるが、そうしたゼロ代名詞推定の問題は今回は扱わず、対話生成の際には直前の対話履歴から主格を補う。

大規模テキストに述語項構造解析を行うと、非常に多くのパターンが抽出されるが、情報検索・推薦の対話で有用なものはドメインに依存して限定される。例えば、野球ドメインでは「A 選手 - が 打つ」や、「B チーム - が 連勝」など、経済ドメインでは「A 社株 - が ストップ高」や、「B 社と C 社 - が 提携する」などの表現が典型例となる。パーザが出力する大量のパターンをすべて利用することは効率が悪いだけでなく、不要なパターンの増加は音声認識誤りや解析誤りなどに対する頑健性の点でも好ましくない。そこで、自動で抽出された述語項構造の中から、当該ドメイン内で対話を行う上で重要なパターンを、NaiveBayes 法による重要度  $NBScore$  に基づいて抽出する [3]。

### 4 抽出情報を用いた情報の検索

まずユーザ発話の述語項構造に完全に一致するパターンを検索する。その後、部分的に一致するパターンを検索するため、その制約条件を徐々に緩和する。具体的には、重要度 ( $NBScore$ ) の低い要素、用言から順に捨象 (ワイルドカード扱い) する。その上で、複数の候補がマッチングすれば、単語間関連度、用言間関連度の高いものを選択する。図 2 にその例を示す。例では、ユーザ発話の中に「金本 (格:ガ) 打つ」「ホームラン (格:ヲ) 打つ」という述語項構造があり、その中で重要度が低い「ホームラン」の情報を捨象して、「金本 (格:ガ) 打つ」「\* (格:ヲ) 打つ」という形の情報検索を行っている。最終的には、典型的な従来手法

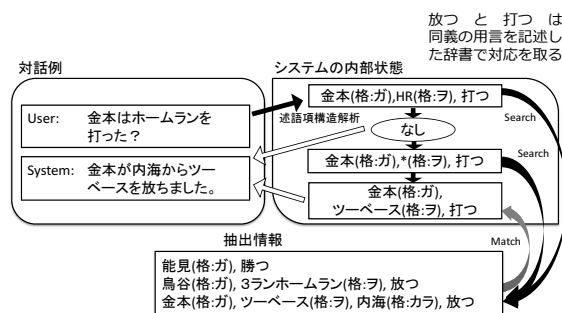


図 2: 対話の生成例

である Bag of Words (BOW) モデルを用いた類似度による検索までバックオフを行う。

本章ではまず、要素と用言の関連度の定義について説明し、その後それらを用いた上記の検索の詳細について説明する。

#### 4.1 要素の関連度を用いた検索の拡張

部分的な述語項構造パターンの一致を利用して検索を行う場合、一致する部分パターン以外の情報を利用して、応答に利用する文を決定する必要がある。図 2 のように「格 + 用言」の一致を利用する場合、どの要素を含む述語項を応答に利用するか選択する指標が必要である。そこで単語間関連度  $S(w_i, w_j)$  を、以下のように文書中で共起する割合で定義する。

$$S(w_i, w_j) = \frac{C^2(w_i, w_j)}{C(w_i) \times C(w_j)}.$$

$C(w_i, w_j)$  は文書における  $w_i$  と  $w_j$  の共起頻度、 $C(w_i)$  は  $w_i$  の頻度である。「格 + 用言」の部分パターンを検索に利用する場合、より関連性の高いものを優先して応答に利用するのが適当と考えられる。

#### 4.2 用言の関連度を用いた検索の拡張

用言についても、完全に一致しない場合に対応する必要がある。特に、「打った」「放った」のような同義表現を扱える必要があるが、これはドメインに依存する。以前の研究 [4] では、同義的用言を手で定義したドメイン類義用言辞書を用いていたが、コーパスから自動で獲得した用言間関連度を利用する方法を導入する。先行研究において、コーパスから類義用言の自動獲得を行う手法 [14, 7] が研究されており、そうした手法を応用する。具体的には、述語項構造における用言に係る要素の分布類似度計算を用いる。

述語項構造における要素、格、用言の 3 つ組のうち、要素と格の組み合わせである格要素  $w_{role_p, arg_q}$  を用言  $w_{pre_p}$  に対する分布として捉え、これらの条件付き確率を考える。一般的な分布類似度の計算では、これらの組をベクトルとして、ベクトル同士の類似度を、コ

<sup>1</sup><http://nlp.kuee.kyoto-u.ac.jp/nl-resource/>

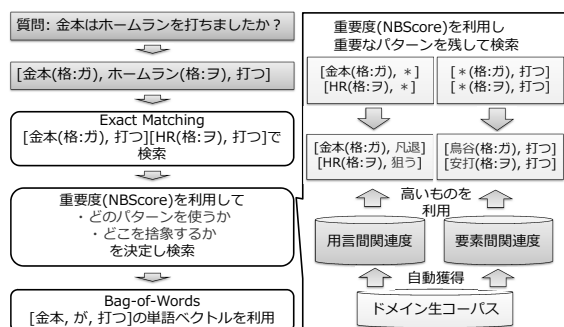


図 3: 述語項構造の段階的緩和によるマッチング

サイン距離を用いたベクトル空間モデルによって表す。

$$\text{sim}_{pre}(w_{pre_i}, w_{pre_j}) = \cos_{pre}(\vec{u}_i, \vec{u}_j) = \frac{\vec{u}_i \cdot \vec{u}_j}{\|\vec{u}_i\| \|\vec{u}_j\|}$$

ここで、 $\vec{u}_i = (C(w_{role_1, arg_1}), \dots, C(w_{role_l, arg_m}))$ 。

しかし、述語項構造のような、コーパスにおいてスパースな素性を用いる場合には、単純なベクトル空間モデルはコーパスの偏りに大きく依存するという問題が生じる。

そこで、要素同士の関連度を利用することによって、スムージングを行う。ここで利用する関連度は、4.1節で定義した要素間関連度と異なり、コーパスにおいて内容語の直前、直後に現れる内容語の分布類似度をベクトルとする。ただし、格要素同士の類似度を全て用いると、爆発的に増加してしまうので、用言の頻度が10以上のものについて計算を行う。さらに、類似度が高い格要素の組み合わせのみを利用した。

このような同義用言獲得の先行研究[14]では、同義用言の他に、対義、時間経過の用言が獲得されるという問題が指摘されているが、本研究の応用である情報検索に基づく対話システムでは問題ないとする。

### 4.3 関連度を用いた情報の検索

関連度を用いた情報検索の具体的な例を図3に示す。例えば、「金本(格:ガ), ホームラン(格:ヲ), 打つ」というクエリがユーザから与えられた場合、まずシステムは「金本(格:ガ), 打つ」「ホームラン(格:ヲ), 打つ」というパターンでのマッチング(Exact Matching)を行う。その後、捨象をした場合は、要素の場合は「鳥谷(格:ガ), 打つ」「ヒット(格:ヲ), 打つ」のように、用言の場合は「金本(格:ガ), 凡退」「ホームラン(格:ヲ), 狙う」のように、部分マッチング(Partial Matching)により、関連度の高い情報の検索を行う。最終的には、「金本, ホームラン, 打つ」というベクトルを用いたBOWモデルまで遡って検索を行う。

## 5 生成された応答の評価

重要度(NBScore)と単語間関連度を10年分(2000-2009年)の毎日新聞記事データベース(CD-

毎日新聞データ集)のプロ野球関連記事から学習し、対話システムを構築した。システムの評価のために、毎日新聞社のWebサイトに存在する2010年9月19-26日の記事の内容について設問を201問用意した。システムの応答のうち、適切な述語項構造パターンを1つ提示できた場合を「的確」とした。それに対して、適切な述語項構造パターンを含んでいるものの、複数の述語項構造パターンや文を提示し、ユーザの質問に対する適切な回答を1つに絞りきれなかったものを「曖昧」とし、誤った述語項構造パターンのみを出力した場合を「誤り」とした。回答が全く生成されない場合は、「回答なし」とした。これらの評価指標を用いて、テキスト入力、音声入力について評価する。この結果を表1に示す。

表におけるExactとは、4.3で示したExact Matchingのみを利用した応答の評価である。Exact + PartialとはExact Matchingで一致する情報が見つからなかった場合に、Partial Matchingを利用するモデルであり、Exact + Partial + BOWとは、最終的にBOWモデルまで遡って検索を行うモデルである。また、SOWモデルとはBOWモデルに語系列を考慮したものである。

提案手法ではBOWモデルと比較して、ユーザの質問に対して回答を的確に提示できている割合が大幅に増加しており、簡潔で直接的な応答ができていることがわかる。

提案手法について詳細に見ると、Exact Matchingによって応答を行ったものは、ユーザの質問に対して適切な応答を選択できている反面、回答が得られない場合も多い。Partial Matchingによって、「回答なし」が減って正解率は向上したが、一方で誤ったり、冗長な回答も生じた。この中には、単語間関連度や重要度(NBScore)を利用する際に、スコアの低いものまで遡って用いた結果誤ったパターンを提示してしまったものがいくつかあった。BOWモデルまで遡った場合は少ないが、BOWモデルを利用することで「回答なし」はほぼなくなり、的確な応答を選択できる場合も増えている。これによって、BOWモデル単体よりも、同義用言辞書と重要度を用いた場合、F値で15.8%、SOWモデル単体よりも10.5%の精度向上が見られた。また、用言間関連度を用いた場合F値で17.6%、SOWモデル単体よりもF値で12.3%の精度向上が見られた。

音声入力を用いた場合は、精度が音声の認識誤りの分だけ低くなっている。今回用いた音声認識は、Julius<sup>2</sup>に対してドメイン適応した言語モデルを用い、質問文の単語認識率は76%であった。特に、提案手法のPartial Matchingが、Exact Matchingで検索失敗した場合に対して効果的に作用しており、認識誤りがあった場合でも、ドメインにおいて重要な述語項構造パターンのテンプレートを利用することによって、適切な情報

<sup>2</sup><http://julius.sourceforge.jp>

表 1: システムの応答の評価 (用言間関連度を用いる場合)

入力	モデル	正解	曖昧	誤り	回答なし	適合率	再現率	F 値
text input	Exact	29.9%	0.5%	1.5%	68.1%	93.8%	30.3%	45.8%
	Exact+Partial	66.2%	5.0%	20.3%	8.5%	72.5%	71.1%	71.8%
	Exact+Partial+BOW	69.7%	5.0%	25.3%	0.0%	70.1%	74.6%	72.3%
	(cf) Bag-of-words (BOW)	46.8%	13.9%	39.3%	0.0%	49.8%	60.7%	54.7%
	(cf) Sequence-of-words (SOW)	54.2%	11.4%	34.3%	0.0%	55.2%	65.6%	60.0%
speech input (ASR)	Exact	19.4%	1.0%	0.5%	79.1%	89.1%	20.4%	33.2%
	Exact+Partial	57.2%	6.0%	18.9%	17.9%	65.8%	63.2%	64.5%
	Exact+Partial+BOW	64.1%	6.5%	28.9%	0.5%	61.7%	70.6%	65.9%
	(cf) Bag-of-words (BOW)	39.8%	9.4%	48.8%	0.0%	42.9%	49.3%	45.9%
	(cf) Sequence-of-words (SOW)	46.3%	10.4%	43.3%	0.0%	48.3%	56.7%	52.2%

表 2: 用言の検索拡張手法の比較

入力	用言の検索手法	F 値
text input	辞書+重要度	70.5%
	用言間関連度	72.3%
speech input	辞書+重要度	62.6%
	用言間関連度	65.9%

選択ができています。更に、最後には BOW モデルを用いることによって、音声認識誤りが生じてでも頑健な情報のマッチングができています。

表 2 においては、コーパスから求めた用言間関連度によって検索を行った場合と、人手で定義した同義用言辞書と重要度によって用言の選択を行った場合を比較しており、前者の有効性が確認できた。

## 6 まとめ

述語項構造解析による情報抽出を利用した音声対話システムを実装した。文の意味構造を利用しながら頑健なマッチングを行うために、ドメインにおいて重要な述語項構造パターンと、述語項における各要素の類似度を利用した検索を行う。従来手法である BOW モデル、SOW モデルによる検索より高い応答精度を示すことができた。これらの知識はいずれも、コーパスからの教師なし学習によって獲得することができ、様々なドメインに対して適応した対話システムを作成することが容易である。

今後は、今回提案した手法を用いて、他のドメインへの適用を行うことを検討している。

## 参考文献

- [1] D.A.Dahl. Expanding the scope of the ATIS task: The ATIS-3 corpus. In *Proc. ARPA Human Language Technology Workshop*, pp. 43–48, 1994.
- [2] 河原大輔, 黒橋禎夫. 自動構築した大規模格フレームに基づく構文・格解析の統合的確率モデル. *自然言語処理*, Vol. 14, No. 4, pp. 67–81, 2007.
- [3] 吉野幸一郎, 河原達也. Web からの情報抽出を用いた音声対話システム. *情処研報*, SLP-82-20, 2010.
- [4] 吉野幸一郎, 河原達也. Web からの情報抽出を用いた対話システムの評価. *人工知能研資*, SIG-SLUD-B002-04, 2010.
- [5] L.Ramshaw and R.M.Weischedel. Information extraction. In *IEEE-ICASSP*, Vol. 5, pp. 969–972, 2005.
- [6] 安達史博, 河原達也, 奥乃博, 岡本隆志, 中嶋宏. Voicexml の動的生成に基づく自然言語音声対話システム. *情処研報*, SLP-40-23, pp. 133–138, 2002.
- [7] P.Patrick, C.Eric, B.Arkady, P.Ana-Maria, and V.Vishnu. Web-scale distributional similarity and entity set expansion. In *Proc. EMNLP*, pp. 938–947, 2009.
- [8] R.Grishman. Discovery methods for information extraction. In *Proc. ISCA & IEEE Workshop on Spontaneous Speech Processing and Recognition*, pp. 243–247, 2003.
- [9] R.Pieraccini, E.Tzoukermann, Z.Gorelov, J-L.Gauvain, E.Levin, C.-H Lee, and J.G.Wilpon. A speech understanding system based on statistical representation of semantics. In *Proc. IEEE-ICASSP*, Vol. 1, pp. 193–196, 1992.
- [10] T.Kawahara. New perspectives on spoken language understanding: Does machine need to fully understand speech? In *Proc. IEEE-ASRU*, pp. 46–50, 2009.
- [11] 河原達也, 川島宏彰, 平山高嗣, 松山隆司. 対話を通じてユーザの意図・興味を探り情報検索・提示する情報コンシェルジェ. *情報処理*, Vol. 49, No. 8, pp. 912–918, 2008.
- [12] T.Misu and T.Kawahara. Dialogue strategy to clarify user's queries for document retrieval system with speech interface. *Speech Communication*, Vol. 48, No. 9, pp. 1137–1150, 2006.
- [13] T.Misu and T.Kawahara. Bayes risk-based dialogue management for document retrieval system with speech interface. *Speech Communication*, Vol. 52, No. 1, pp. 61–71, 2010.
- [14] 柴田知秀, 黒橋禎夫. 文脈に依存した述語の同義関係獲得. *情処研報*, NL-199-13, 2010.
- [15] 清田陽司, 黒橋禎夫, 木戸冬子. 大規模テキスト知識ベースに基づく自動質問応答 ユーザインタフェース. *自然言語処理*, Vol. 10, No. 4, pp. 145–175, 2003.