

オントロジを利用した会話理解手法の提案

杉山 聡 赤埴 淳一 佐藤 哲司

日本電信電話株式会社 NTT コミュニケーション科学基礎研究所
〒619-0237 京都府相楽郡精華町光台 2-4
{sugiyama, akahani}@cslab.kecl.ntt.co.jp, satoh.tetsuji@lab.ntt.co.jp

本報告では、オントロジを利用して会話を理解する手法を提案する。インターネットの発達により、音声通信に高度なサービスを統合することが現実的になってきた。会話内容を理解することが出来れば、会話内容に応じた情報提供などのサービスが可能となる。対話理解規則や統計的手法を利用した従来の対話理解の手法では、さまざまな話題が話される自由な会話に対応するのは難しい。さまざまな概念を体系化した知識であるオントロジを利用して会話を理解する方法を示し、会話内容に応じた情報提供サービスへの適用を提案する。

A Method for Understanding Natural Conversation by Using Ontology

Akira Sugiyama Jun-ichi Akahani Tetsuji Satoh

NTT Communication Science Laboratories, NTT Corporation
2-4, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0237 Japan

This paper proposes a method for understanding natural conversation by using ontology. The evolution of the Internet makes it possible to combine useful services into speech communication. If a system understand human conversation, it is possible to provide appropriate information to speakers. However, it is difficult to interpret natural conversation by using task knowledge or statistical processing. We therefore propose a method for understanding natural conversation by using ontology and an application that provides related information to conversation.

1 はじめに

インターネットの発達や高速な加入者線の普及により、音声通信の環境が大きく変わってきている。特に、IP 電話に代表されるようなインターネットを介した音声通信を可能とする技術を使うことで、さまざま

な付加価値をつけた音声通信サービスが考えられるようになった。その一つとして、インターネット上に多数存在する Web ページなどを情報源とする情報検索の技術と組み合わせ、会話内容に応じた情報を提供するサービスが考えられる。

会話内容に応じた情報を提供するために

は、会話内容を理解する方法が必要である。会話内容を理解する従来の方法には、音声対話システムのように対話理解規則を用意して理解する方法や、対話コーパスを構築して統計的手法を用いる方法などがある。これらの方法では、ドメインに依存した知識やドメインごとの対話コーパスが必要となる。従来の方法を電話のようにさまざまな話題が話されるサービスに適用するには、ドメイン依存の知識やデータの収集のコストが大きな問題となる。さまざまな話題の会話を適切に理解する方法が必要である。

一方、セマンティック Web の研究分野では、さまざまなドメインの知識をオントロジとして記述する活動がおこなわれており、Web で利用可能となることが期待されている。これらのオントロジは、会話内容を理解するために構築されているものではないが、話題が特定できない会話を理解するための知識源として有用であると考えられる。本報告では、オントロジを利用した会話理解の方法を提案する。

2 関連研究

対話の談話構造理論の研究では、Grosz, Sidner [1]が、対話の談話構造を発話列の構造と意図、注視状態の 3 つの構成でモデル化している。また、対話システムの研究では、プランを用いて談話理解を行う方法が提案されている[2,3]。計算機と人間との対話で適切な応答を生成するためには、このような方法は有効である。しかし、人同士の自由な会話の場合、発話が断片的であることや音声認識の誤認識の問題を考慮すると、実現は難しい。

音声対話システムでは、プランを用いて推論する[4]、属性と値の構造でドメイン知識を用意して管理する[5]などの方法で対話を理解する。いずれの方法でも、会話を理解するドメイン知識が必要である。目的達成型の対話で限定された話題ならばドメイン知識を用意して運用することも可能だが、さまざまな話題がかわされる会話に適用す

るには、会話を理解する知識を構築するのが困難である。

知識を構築するコストを削減するために、語彙と言語モデルをドメインのデータベースから自動的に構築する方法が提案されている[6]。この方法は、ユーザがシステムに質問することを想定しており、人同士がシステムを意識せずに会話をする本研究の想定とは異なる。

自然言語処理の分野では、情報検索の研究が盛んに行われている。ユーザの質問文を意味解析することで検索式を生成し検索する方法[7]や、質問文や文書を多次元ベクトルで表現[8]し、機械学習的な手法を用いる方法などの研究がある。意味解析をする方法は、ドメイン知識が必要であり、領域を限定しない会話では扱いが難しい。また、機械学習的な手法は、さまざまな話題の会話コーパスの収集や、実時間での処理などの問題がある。

3 オントロジとセマンティック Web

オントロジとは、概念間の関係を体系化した知識である。例えば、店舗のオントロジでは、店舗の特徴をあらわす概念として店名や、所在地、種類などの概念があることが記述される。オントロジは特定の概念ごとに定義される。例えば、店舗のオントロジやビールのオントロジ、車のオントロジなど、多くのオントロジが定義される。

このようなオントロジを利用することで、Web ドキュメントを機械処理可能にするセマンティック Web の研究が盛んになってきている[9]。セマンティック Web では、Web ドキュメントの内容をオントロジを使って意味付けし、メタ情報として記述する。オントロジ記述言語の標準化[10]が進められており、さまざまなオントロジが Web で利用可能となることが期待される。

図 1 は、セマンティック Web におけるオントロジの例である。この例では、「飲食店」のオントロジの例を示している。「飲食店」には「居酒屋」や「和食店」のサブクラス

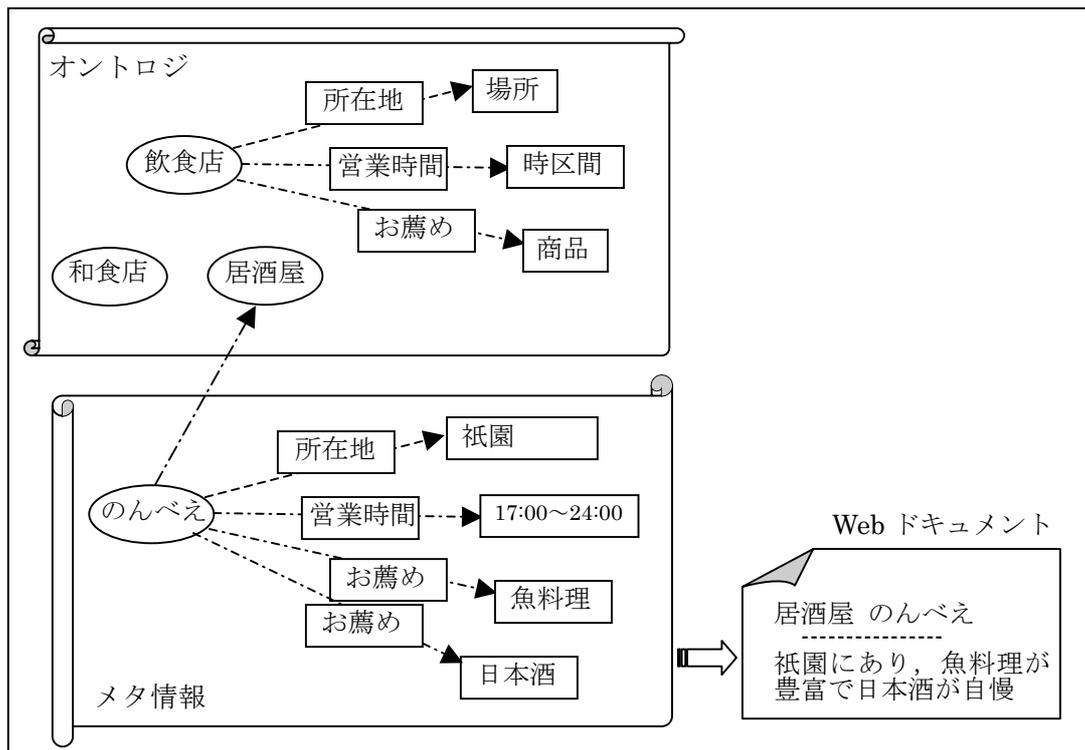


図1 オントロジとセマンティック Web の例

があることや、「所在地」や「営業時間」などの属性があるなどの知識構造がオントロジとして記述されている。Web ドキュメントの内容はオントロジに基づいて意味付けされ、メタ情報として公開される。この例では、「のんべえ」という店の Web ドキュメントとメタ情報が示されている。

4 オントロジを利用した会話理解

会話理解の課題として、ドメインごとに会話を理解する規則を用意しなければならないことや、ドメインごとに会話コーパスを用意しなければならないという課題がある。前節で説明したように、Webの機械処理を可能とするために、共通のオントロジ記述言語を決め、さまざまなドメインのオントロジをWebで利用可能にする活動が行われている。この活動で構築されるオントロジは、会話理解のための規則ではないが、会話理解のための知識として有用であると考えられる。さまざまなドメインのオントロジ

が利用できるようになれば、話題を限定しない会話の理解に利用できるという利点がある。本節では、オントロジを利用して会話を理解し、会話内容に応じた情報を提供する仕組みを提案する。

会話に応じた情報を提供する処理の流れを図2に示す。インターネット上には、さまざまなドメインのオントロジと、オントロジに基づいて意味付けされたメタ情報を持った Web ドキュメントがあるとする。これらを知識源として会話を理解し、メタ情報の付いた Web ドキュメントから会話に応じた情報を検索する。

まず、前処理としてオントロジと Web ドキュメントのメタ情報からオントロジ毎の語彙と頻度を生成する。メタ情報から語彙を収集することで、Web に存在する語彙をオントロジの概念ごとに整理することが出来る。検索対象である Web ドキュメントから、検索対象に出現する語彙と頻度をオントロジ毎に集計する。

本方式では会話音声を音声認識し、形態

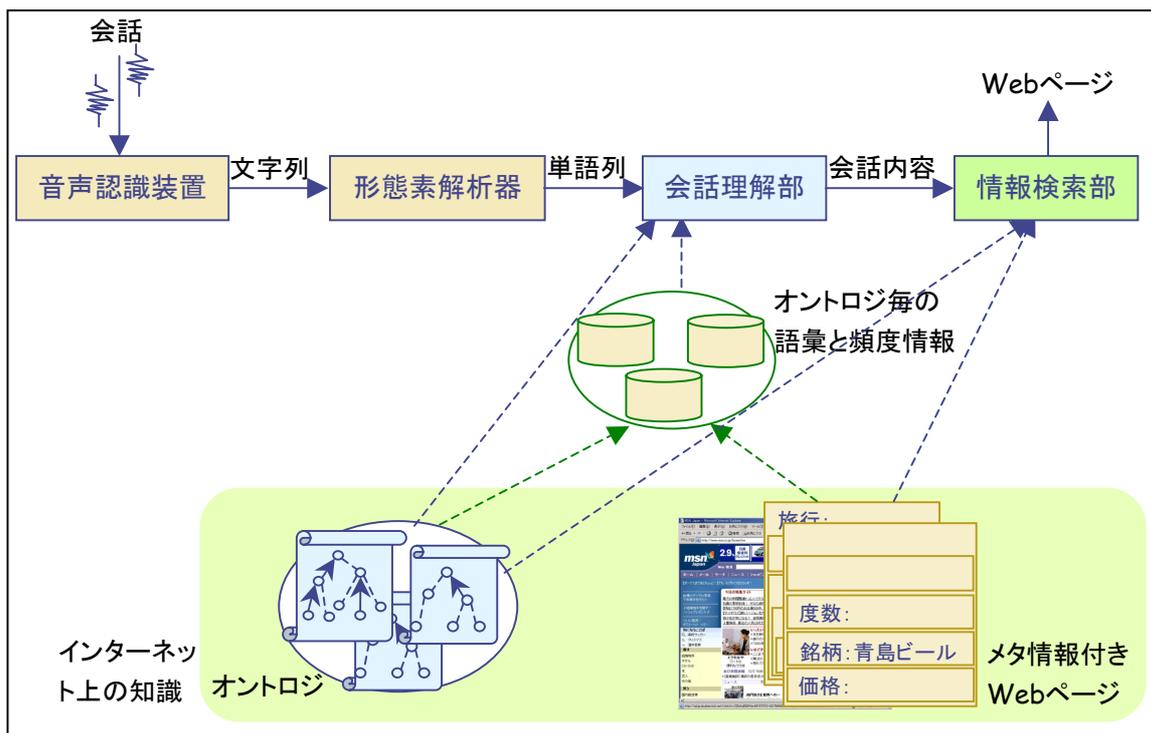


図2 会話内容に応じた情報提示処理

- A1: そうそう ところで石川
 B2: なに
 A3: 演奏会の打ち上げ どこでやる
 B4: うーん なにがいいかな 魚料理が
 食べたいなあ
 A5: 魚ね
 B6: うん 魚
 A7: 僕は日本酒が飲みたいな
 A8: 場所は祇園のあたりで飲むんだよね

図3 会話例

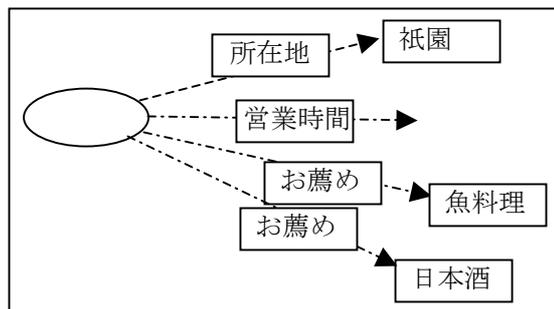


図4 オントロジを利用した会話理解の例

素解析して単語列に変換する。会話理解部では、前処理で作成したオントロジ毎の語彙と頻度の情報を利用して、会話に現れる単語をオントロジのインスタンスを構築する。この際、一つの単語が複数のオントロジに割り当てられることがあり、あいまい性が生じる。一定区間の会話の単語列をすべてオントロジに当てはめ、多くの単語が当てはめられたオントロジを選択することにより、このあいまい性を解消する。本報告では、会話に現れた単語を当てはめたオ

ントロジが会話内容を表しているものと考えられる。情報検索部では、会話から得られたオントロジの値と同じ値をメタ情報に持つ Web ドキュメントを検索し、関連情報として提示する。

図3の会話例を用いて具体的に説明する。会話例をオントロジに割り当てた例が図4である。あらかじめ構築したオントロジ毎の語彙と頻度を利用して、会話に現れた語彙を同じ語彙をもつオントロジの概念に割り当てる。例えば、前処理で構築したオントロジ毎の語彙と頻度には、「飲食店」オントロジの中の、「飲食店」概念と「お薦め」の関係にある「品物」概念に「魚料理」と

「日本酒」という語彙があることが、「のんべえ」の情報のメタ情報から得られている。会話理解部では、発話 A7 に現れる「日本酒」を、オントロジ毎の語彙と頻度で「日本酒」を語彙にもつ概念を調べることで、「飲食店」オントロジの「品物」概念の値であることがわかる。発話 A8 の「祇園」も同様に、「飲食店」オントロジの「場所」概念のインスタンスであることがわかり、図 4 のような会話理解が得られる。情報検索部で、図 4 の概念と値の組をメタ情報に持つ Web ドキュメントを検索することで、会話内容に適した情報を検索する。

5 考察

本報告では、会話に現れる語をオントロジに割り当てた結果を会話内容とした。従来の対話理解の研究のように、発話意図の認識や談話構造を構築しておらず、会話から得られるはずの情報をかなり捨てている。しかし、前節で述べたような、会話に関連する情報を提示する程度なら、十分に役に立つ情報だと考える。

会話内容に応じた情報を提供する簡単な方法として、会話に現れるキーワードを検索エンジンに入力する方法が考えられる。キーワードで検索する方法では、会話に現れるどのキーワードをまとめて検索すればよいのか明らかではない。また、音声認識の誤認識の影響も大きい。本報告の方法では、オントロジを使って会話に現れる語をまとめている。会話に現れる語は複数のオントロジに対応すると考えられるが、もっとも多く値が埋まったオントロジのインスタンスを検索条件とすることで、キーワードが選択できると考えている。例えば、前節の例では、「石川」や「演奏会」などの語が現れている。これらの語は、「場所」オントロジや「音楽」オントロジなどにも当てはまるが、「飲食店」オントロジを選択することで、これらの語を検索条件から除外できている。単純にキーワードを使って検索をするよりも、適切な検索条件の選択が

期待できる。

6 まとめと今後の課題

本報告では、オントロジを利用して会話を理解する手法を提案した。インターネットの発達により、音声通信に高度なサービスを統合することが現実的になってきた。会話内容を理解することが出来れば、会話内容に応じた情報提供などのサービスが可能となる。会話理解規則や統計的手法を利用した従来の会話理解の手法では、さまざまな話題が話される電話サービスに対応できない。さまざまな概念を体系化した知識であるオントロジを利用することで、人同士の自由な会話をある程度理解することが可能となる。また、会話に応じた情報を提供するサービスの枠組みを示した。本手法では、オントロジを用いて会話内容を表現している。オントロジの構造を利用することで、統計的手法とは異なる検索が可能となる。

本報告では、オントロジを利用して会話を理解し情報を提供する方法のファーストステップを述べた。本方式に、ドメインに依存しない会話理解規則を導入することにより、文脈を考慮した会話理解が可能と考えられる。また、情報を提示するタイミングなどの情報提示戦略を導入することで、より会話内容に即した情報提供が可能となる。

今後の課題として、まず音声認識の認識誤りへの対処方法があげられる。音声認識を意識しない人同士の会話を精度よく認識するのは、現在の技術では難しいため、音声認識誤りに頑健な方法が必要である。また、本手法では会話内容をオントロジの形式で表現している。オントロジの構造を使って、会話内容から適切な Web ページを検索する方法が必要である。さらに、電話会話では、さまざまな話題が次々に話される。話題の転換に迅速に対応できる方法が必要である。

謝辞

本研究において、有益なアドバイスをいただいた社会インタラクション研究グループの皆様には感謝します。

参考文献

- [1] B. J. Grosz, C. L. Sidner. Attention, Intentions, and the Structure of Discourse. Computational Linguistics, Vol.12, Num.3, pp.175-204, 1986.
- [2] S. Carberry. Plan Recognition in Natural Language Dialogue. MIT Press, Cambridge, 1990.
- [3] J. F. Allen. Natural Language Understanding (second edition). Benjamin / Cummings, 1995.
- [4] J. F. Allen, G. Ferguson, A. Stent. An Architecture For More Realistic Conversational Systems. In Proc. Intelligent User Interfaces 2001, pp.1-8, 2001.
- [5] M. Nakano, K. Dohsaka, N. Miyazaki, J. Hirasawa, M. Tamoto, M. Kawamori, A. Sugiyama, T. Kawabata. Handling Rich Turn-Taking in Spoken Dialogue Systems. In Proc. 6th Eurospeech, pp.1167-1170, 1999.
- [6] K. Komatani, K. Tanaka, H. Kashima, T. Kawahara. Domain-Independent Spoken Dialogue Platform using Key-Phrase Spotting based on Combined Language Model. In Proc. 7th Eurospeech, pp. 1319-1322, 2001.
- [7] 佐々木裕, 磯崎秀樹, 平博順, 平尾努, 賀沢秀人, 鈴木潤, 前田英作. SAIQA: 大量文書に基づく日本語質問応答システム. 情処研究報告 NL145/FI64, pp.77-82, 2001
- [8] G. Salton, A. Wong, C. Yang. A Vector Space Model for Automatic Indexing. Communication of the ACM, Vol.11, No.18, pp.613-620, 1975.
- [9] T. Berners-Lee, J. Hendler, and O. Lassila. The semantic web. Scientific American, Vol.284 No.5, pp.34-43, 2001.

[10] D. L. McGuinness and F. van Harmelen (eds.). OWL Web Ontology Language Overview. W3C Working Draft, <http://www.w3.org/TR/2003/WD-owl-features-20030331/>, 2003.