

# オントロジーを用いた音声入力質問文の解析手法の提案

寺村 真加寿

荒木 雅弘

京都工芸繊維大学情報工学専攻

## 概要

柔軟な音声対話システムを構築するためにはユーザーにできるだけ自由な発話を許すことが重要である。しかし音声対話システムにおいて、ユーザーの発話の意図を理解する言語理解部は、そのシステムのドメインに依存するところが大きく、ユーザーの発話を大きく制限するものが多い。本研究では、推論が可能なベイジアンネットワークとオントロジーを組み合わせることで、多様な入力に対して柔軟な言語理解手法を提案する。

## 1 はじめに

自由な発話における表現は多種多様であり、それらすべてのルールを記述するのは困難である。しかし、あるひとつのタスクにおいて、発話の意図を理解するためには認識結果中のすべての単語を理解する必要はなく、必要最低限の単語だけで十分な場合が多い。また一方では音声認識誤りにより生じる不要単語を取り除く必要がある。このように、自由発話の解析には柔軟な処理が重要となる。

そこで本研究では、柔軟な発話理解を推論ネットワークとオントロジーを用いて実現する方法を提案する。まず、タスクの進行に必要な単語の概念だけでオントロジーを構成しその関係を記述しておく。さらに、複数の表現方法が存在する事柄もひとつの概念としてオントロジーでまとめておくことで、多様な言い回しに対応できるようにする。そしてユーザーの意図と行為をあらわしたベイジアンネットワークと関係付け、意味解析ネットワークを構成する。このネットワークを利用することで、様々な表現の発話から必要な単語を抜き出し、意図を推測する。

オントロジーを利用する利点はいくつかある。ひとつは、今後のセマンティック WEB の進展によってオントロジーの整備が期待できることである。これによってすべてを手で書く必要がなくなり作業量が格段に減る。また、オントロジー記述言語である OWL/RDF の知識表現は、ドメインにおける推

論や音声対話システムの応答生成部にも利用できるかもしれないということである。

## 2 意図理解

今回提案する手法では、音声認識と言語理解は別々に行う。大語彙音声認識器から送られてきた単語ごとに分割された認識結果を言語処理部で処理する。発話から言語理解までの大まかな流れは図 1 に示すとおり。

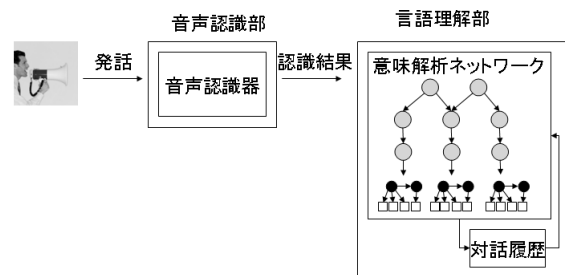


図 1: 言語理解までの流れ

### 2.1 意味解析ネットワーク

意味解析ネットワークは大語彙音声認識器によって認識された音声入力を解析し、意味表現を作成するためのものである。意味解析ネットワークはベイ

ジアンネットワーク部とオントロジー部からなる。オントロジー部の一部の要素はベイジアンネットワーク部に含まれており、オントロジーは RDF トリプルの形であらわされる。入力された単語はオントロジー部のインスタンスにヒットするものがあるかどうかを調べられる。ヒットするとオントロジーからその単語の概念を取得し、概念と関連付けられたベイジアンネットワークのノードからユーザーの意図の推論をおこなう。

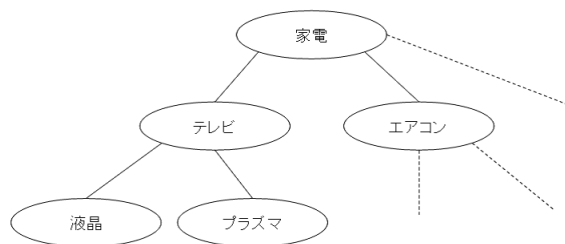


図 3: オントロジーの例

### 2.1.1 オントロジー部

オントロジー部ではシステムで使用される単語とその概念、単語間の関係が記述されている。ベイジアンネットワーク部の末端のノードはこのオントロジー部で記述されている概念とその関係である。

このオントロジーは図 2 のような RDF トリプルをベースにした形で記述される。RDF トリプルは主語と目的語、それらをつなぐ述語で構成される。すべての質問文はこの形に簡略化して書くことができる。「テレビ」というオブジェクトに対しては、「電源を入れる」「チャンネルを変える」などのアクションがあり、そしてこのオブジェクトとアクションの間に「操作」という関係を持たせる。これをベースにすべての概念を体系的に分類し、関係を記述したものがオントロジーであり、その記述言語が OWL である。これらはすべてオントロジーに記述される。ほとんどの質問文において、それが明示的に発話されるかどうかは別として、これらの要素は含まれる。ベイジアンネットワーク部の末端ノードに割り当てられるオントロジーにはこの RDF トリプルのセットが少なくとも 1 つあり、これらが埋められるとそのノードが観測されたと判断する。

複数の表現方法がある単語については、図 4 のようにその要素に複数の表現方法を関連付け、それと入力発話の単語を比較するようにする。



図 2: RDF トリプル

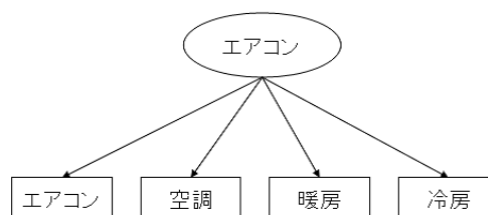


図 4: 複数の表現の割り当て

### 2.1.2 ベイジアンネットワーク部

ベイジアンネットワーク部では図 5 に示すように概念のレベルごとに階層が分けられている。階層は「意図」「行為」「手段」「概念」の 3 つからなる。

「意図」階層 「意図」階層はベイジアンネットワークの最上位に位置し、各ノードはユーザーの意図を表す。

「行為」階層 「行為」階層は、ユーザーが意図を満足するための行為を表すノードが集まった階層である。「ニュースがみたい」という意図に対して、「テレビを見る」「ラジオを聴く」「インターネットをする」などの行為がこの階層にあたる。取得すべき意味表現はこの階層に記述されている。

「手段」階層 「手段」階層では、各ノードは「行為」が実際に実行されるための手段、手法を現している。「テレビを見る」という行為をするために「テレビの電源を入れる」という手段が必要となる。この階層のノードはベイジアンネットワークの末端であり、オントロジー部の概念部

分がこれらのノードになる。この階層のノードの内容は、ユーザーの発話にほぼ等しい。

「概念」階層 「概念」階層は、オントロジー部の概念部分のことであり、手段階層の内容を概念と関係だけで表したものである。手段階層から各概念の要素に対して、タグつきコーパスから得られた確率を与える。ベイジアンネットワーク部においては主語と目的語の関係を表す述語もひとつのノードとして扱われる。

これらの各ノードの確率は、ユーザーとの対話履歴から更新される。

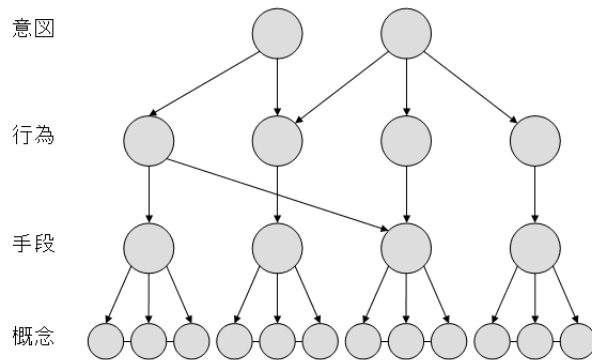


図 5: ベイジアンネットワーク部

## 2.2 対話

ユーザーからの発話が入力されると、認識器は上位数個の認識結果を出力する。得られた認識結果の上位の各単語から調べていき、ネットワークの末端に割り当てられたオントロジーにマッチするものがあるかどうかを調べる。マッチするものがあれば、そのノードから推論を行い、意味表現を抽出する。ある質問文について、異なるいくつかの表現の仕方が存在する場合がある。情報家電の操作システムを例にとると、

U:「今テレビでどんな番組が放送されていますか？」

と

U:「今テレビで放送されてる番組表を見せてください」

は表現や使っている単語に違いがあるが、意味は同じである。この質問で必要な概念は「テレビ」の「番組表」を「表示」する、ということであり、各概念に設定された単語を参照することで表現の差を吸収し、同じ意味表現を取得することができる。また、

U:「今どんな番組がテレビで放送されていますか？」

のように単語の順番が入れ替わっても、同様に意味表現を取得することができる。さらに対話が続くときは、発話中に主語が省略されたり、代名詞を使用することがある。先の対話に続いて

U:「6ch を映してください」

と発話すれば、これは明示的にはテレビとは言っていないもののテレビのことをさしているのは明らかである。このときは対話履歴を参照することで解決する。履歴は最近のものほど重みを重くする。

もし、認識誤りが起こり、どのノードも観測できなかった場合は次の候補文を利用する。

## 3 まとめ

ベイジアンネットワークとオントロジーを組み合わせた質問文解析手法を提案した。この手法ではユーザーの自由な発話に対して、オントロジーを利用することで同じ意味の異なる表現の差を吸収し概念として取り扱うことができるので推論に利用しやすくなる。さらにベイジアンネットワークを用いることで、ユーザーとの発話履歴などを用いて各ノードの確率を変えることができ、自動的にユーザーに適合したシステムにすることもできると考えられる。

今後はオントロジーの詳細な記述ルールの決定や、効率的なベイジアンネットワークの構成方法を考えていく。オントロジーの記述は現在は手動で行うしかないが、将来セマンティック WEB が普及すれば、比較的楽にオントロジーを入手できるようになると考えられる。また、ベイジアンネットワークもオントロジーの形で記述することができればひとまとめに書くことができるようになる。[1] や [2] のようなベイジアンネットワークをオントロジー記述言語で記述する研究もおこなわれている。また情報検索にオントロジーをもちいたり [3]、音声対話システムにオントロジーを積極的に利用する研究も行われている。

る [4] ので、これから OWL の利用価値は上昇すると思われる。

## 参考文献

- [1] Z.Ding and Y.Peng. A Probabilistic Extension to Ontology Language OWL. Proc. 37th Hawaii Int'l Conf. System Sciences (HICSS 04), IEEE CS Press, (2004).
- [2] Z.Ding, Y.Peng and R.PAN. A Bayesian Approach to Uncertainty Modelling in OWL Ontology. In Proceedings of AISTA-04, Luxembourg, (2004).
- [3] M.Holi and E.Hyvonen. Probabilistic Information Retrieval Based on Conceptual Overlap in Semantic Web Ontologies. In Proceedings of the 11th Finnish AI Conference, Web Intelligence, vol. 2, Finnish AI Society, Finland, (2004).
- [4] J.Kopisa, Z.Mikovec and P.Slavik. Ontology driven voice-based interaction in mobile environment. Human Language Technology Conference, Proceedings of the first international conference on Human language technology research, (2001).