

決定木を用いた未知語属性推定による効率的対話方法

高橋 康博 堂坂 浩二 相川 清明

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

1 はじめに

音声対話システムとは、人と計算機が音声対話を通して決められた仕事(タスク)を遂行するシステムである。音声対話システムが保有する知識には制限があるが、ユーザはシステムの知識の範囲を知っているとは限らない。従って、ユーザがシステムにとって未知の語を含む発話をする事は避けられない。ユーザが未知語を含む発話をした場合、未知語の処理を行わないシステムでは、未知語を既知語に誤認識する。また、未知語を検出し、棄却するシステムでは、タスクの遂行に必要な情報を得るためにユーザに質問を繰り返す。このため、未知語の処理を行わないシステムや未知語を検出し、棄却するシステムでは、対話が非効率になる場合やタスクを完了できない場合がある。

本稿では、ユーザが未知語を含む発話をした場合でも、効率的な対話を実現することを狙いとして、ユーザ発話中の未知語の属性を推定し、推定した属性を利用して効率的に対話を行う方法を提案する。この方法では、決定木を用いて未知語の属性を推定する。そして、推定した属性が未知語の属性かどうかを必要に応じてユーザに確認し、確認された属性についてはユーザに質問を繰り返さず、確認された属性以外の属性についての情報により対話を進める。本稿で対象とする未知語は、システムにとって既知の属性をもつ単語(名詞)であり、未知語の音響処理により、ユーザ発話中の未知語を検出し、未知語ラベルを付けて出力できるとする。

ユーザ発話中の未知語の属性推定方法は、[1, 2]で提案されている。[1]の方法は、システムとユーザの過去の対話(実例)の距離に基づく方法である。この方法は、実例を蓄積し、未知語を含む現在の対話との距離が最小となる実例を基にする方法である。実例は、ある時点のシステム発話と直後のユーザ発話の組であるが、対話の書き起こしを実例として使うために、実例の作成に手間がかかる。[2]の方法は、未知語を含むユーザ発話の理解結果に属性推定規則を適用する方法である。しかし、属性推定規則を手動で作成するために手間がかかる。

そこで、実例を作成する手間や属性推定規則を作成する手間を必要としない属性推定方法を提案してきた[6]。この方法は、実例の距離に基づく方法である。しか

し、予備実験の結果、実例数を増加させても属性推定正解率が向上しないという問題があった。

本稿では、属性推定正解率が[6]の方法よりも高く、実例を作成する手間や属性推定規則を作成する手間を必要としない属性推定方法として、決定木を用いた方法を提案する。そして、[1, 6]の方法を改良した方法と比較して属性推定正解率が高いことを検証する。

効率的な対話方法は、[3, 4]で提案されている。[3]は、強化学習を利用する方法であり、[4]は、音声認識率を考慮する方法であるが、いずれの研究も、ユーザが未知語を含む発話をした場合を考慮していない。

本稿では、推定した属性を利用して効率的に対話を行う方法を提案する。そして、未知語の処理を行わないシステムや未知語を検出し、棄却するシステムと比較して短い対話でタスクを遂行できることをシミュレーション実験により検証する。

2 決定木を用いた未知語属性推定

2.1 提案方法

[6]で提案した実例の距離に基づく方法では、実例数を増加させても属性推定正解率が向上しないという問題があった。本稿では、属性推定正解率を向上させるために、[6]のようにヒューリスティックな距離関数を使うのではなく、[6]の実例を学習データとして決定木を作成し、その決定木を用いて未知語の属性を推定する方法を提案する。

未知語の属性を推定するための決定木の作成に使う学習データについて説明する。システム¹とユーザの対話において、システムは、ユーザ発話の理解状態に応じて発話し、その直後のユーザ発話により理解状態を変化させ、変化した理解状態に応じてまた発話をする。このときの

理解状態 → システム発話 → 理解状態変化

という一連の遷移(理解状態遷移)を学習データとする。学習データの例を以下に示す。(例1)

¹ システムの理解状態は属性と値の組の列で表現する。

理解状態 : (日付: 今日)
 システム発話 : 「番組名は何ですか」
 理解状態変化 : (番組名: 街角クイズ)

例1は、ビデオ録画予約システムが、ユーザの要求は今日についてであると理解している状態で、ユーザに番組名について質問し、直後のユーザ発話を理解した結果、ユーザの要求は街角クイズについてであると理解した対話から得られた学習データである。

属性推定のための決定木は、未知語を含む理解状態遷移に対し、理解状態に含まれる属性、システム発話内容、理解状態変化に含まれる属性の情報を使って、未知語がどの属性をもつかを推定する。簡単のため、理解状態変化の第一番目だけに未知語が含まれる場合について、未知語の属性を推定する決定木を作成する方法を述べる。学習データの理解状態から得る情報は、各属性が値をもっているかどうかという情報である。システム発話から得る情報は、システム発話がどの属性についての質問またはどの属性についての確認であるかという情報である。理解状態変化から得る情報は、理解状態変化の第一番目に含まれる属性は何かという情報と理解状態変化の第二番目以降に各属性が値をもっているかどうかという情報である。これらの情報を学習データから取り出し、C4.5のアルゴリズムを適用することにより、決定木を作成する。

例えば、ビデオ録画予約システムにおいて、システムのもつ属性が番組名、日付、ジャンル名であるとき、例1の学習データから得る情報について述べる。理解状態から得る情報は、日付が値をもち、番組名とジャンル名は値をもっていないという情報である。システム発話から得る情報は、番組名についての質問であるという情報である。理解状態変化から得る情報は、第一番目に含まれる属性が番組名であり、第二番目以降には、番組名、日付、ジャンル名は値をもたないという情報である。

理解状態変化に未知語が複数含まれる場合について、未知語の属性を推定するためには、理解状態変化の各位置ごとに、上と同様にして決定木を作成する。

2.2 比較する方法

提案方法と [1, 6] の方法を改良した方法を比較する。比較する方法のアルゴリズムは、[6] と同様である。ユーザが未知語を含む発話をした場合、システムは、ユーザ発話に含まれる未知語を検出し、未知語を含む理解状態遷移を構成する。未知語を含む理解状態遷移と実例データベースの各実例との距離を計算し、距離が最小となる実例を選択する。選択された実例において、未知語

	f_1	f_2	f_3	o
	理解状態	システム発話	理解状態変化	属性列
実例1	(日付: 今日)	番組名は?	(番組名: 街角クイズ)	(番組名)
実例2	(番組名: ザックデー)	いつの番組?	(日付: 明日)	(日付)
⋮	⋮	⋮	⋮	⋮

図1: 実例データベース

に対応する単語の属性を未知語の属性とする。

実例としては、決定木を作成する際に使った学習データを使う。この実例は、[1] で使われている実例と異なり、自動的に収集することが可能である。また、距離関数は、[6] で使われている距離関数ではなく、Memory-Based Reasoning における Value Difference Metric を使う。この距離は、実例に基づく機械翻訳 [5] で使われている距離である。

実例データベースは、図1のように、入力フィールド f_1, f_2, f_3 と出力フィールド o からなる表とする。 f_1, f_2, f_3 は、それぞれ、理解状態、システム発話、理解状態変化を値としても、 o は、理解状態変化から得られる属性の列を値としてもつ。ただし、例えば、理解状態変化の第一番目に検出された未知語の属性を推定する場合、 o の値は属性の列の第一番目の属性とする。

未知語を含む理解状態遷移 τ と実例 ρ の距離 Δ を次のように定義する。

$$\Delta(\tau, \rho) = \sum_{k=1}^3 \delta_k^2(D, \tau.f_k, \rho.f_k)$$

$$\delta_k(D, \tau.f_k, \rho.f_k) = \sum_{v \in V_o} \left| \frac{|D[f_k = \tau.f_k][o = v]|}{|D[f_k = \tau.f_k]|} - \frac{|D[f_k = \rho.f_k][o = v]|}{|D[f_k = \rho.f_k]|} \right|$$

ただし、 V_o は属性の集合、 $\tau.f_k$ は τ のフィールド f_k の値、 $|D[f_k = \tau.f_k]|$ は実例データベース D の中で、 f_k の値が $\tau.f_k$ と等しい実例の数、 $|D[f_k = \tau.f_k][o = v]|$ は D の中で、 f_k の値が $\tau.f_k$ と等しく、 o の値が v である実例の数を表す。

2.3 評価実験

学習データは、我々の研究グループで開発された天気情報案内システム「飛遊夢」と一般のユーザとの対話データから 600 データを収集した。属性の種類は、時間、場所、警報種別、情報種別の 4 つである。飛遊夢には、ユーザ発話中の未知語を検出する機能はないので、

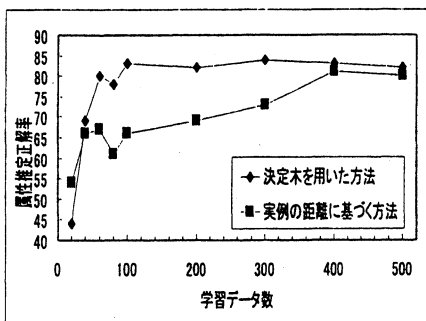


図 2: 属性推定実験結果

収集した学習データの中から 100 データを取り出し、取り出したデータの中の理解状態変化の一部を未知語と仮定することにより未知語を含む理解状態遷移を作成した。理解状態変化の第一番目だけに未知語が含まれる場合について実験を行った結果、図 2 が得られた。

決定木を用いた方法は、学習データ数が 100 の場合、属性推定正解率が 83 % であり、学習データ数を増やした場合と同等の属性推定正解率に到達する。一方、実例の距離に基づく方法は、実例数が 100 の場合、属性推定正解率が 67 % であり、実例数を増やすに従って、属性推定正解率が 81 % になるが、決定木を用いた方法より属性推定正解率が低い。

3 推定した属性を利用した効率的な対話方法

3.1 提案方法

未知語の処理を行わないシステムや未知語を検出し、棄却するシステムでは、未知語の属性についての質問をユーザに繰り返し、対話が非効率になる場合やタスクを完了できない場合がある。そこで、提案方法では、まず、未知語の属性を推定する。属性推定方法は、前節で属性推定正解率が高いことが検証された決定木を用いた方法である。次に、推定した属性がシステムの理解状態に含まれているかどうかを確認する。これは、既に理解状態に含まれている属性をもつ未知語は棄却するためである。推定した属性が理解状態に含まれていない場合、システムは推定した属性が未知語の属性と一致するかどうかをユーザに確認する。未知語の属性が確認された後は、システムは未知語の属性についての質問を行わない。これにより、未知語の属性についての質問をユーザに繰り返し行うことを防ぐ。

以上をまとめると、ユーザ発話に未知語が含まれる場合、提案方法は次のような処理を行う。

1. 未知語の属性推定

理解状態変化の第一番目だけに未知語が含まれる場合、決定木を用いた方法により属性推定を行う。未知語が複数含まれる場合や理解状態変化の第二番目以降に含まれる場合には、未知語は棄却する。

2. 属性決定の必要性の判断

推定された未知語の属性が、システムの理解状態に含まれていないことを確認する。理解状態に含まれている場合には、未知語は棄却する。

3. 属性決定の対話

未知語の属性である可能性が高いと推定された属性が未知語の属性かどうかをユーザに確認する。

上の処理により、未知語の属性が確認された後は、確認された属性についてユーザに質問しない。

3.2 実験条件

提案方法の効率性をシミュレーション実験により検証する。シミュレーション実験は、模擬ユーザとシステムの対話による実験である。やりとりされる情報は、音声ではなく、属性と値の組の列である。提案方法を評価するためには、タスク完了までの対話の効率性を評価するため、1 つの要求を実現するために、システムと対話する模擬ユーザを導入した。実ユーザではなく模擬ユーザを導入したのは、新たな方式を短期間で評価し、評価の基準を統一できるからである。

タスクはビデオ録画予約であり、属性の種類は、日付、時間、ジャンル名、番組名、人名、情報種別の 6 つである。各属性に対し、それぞれ、7 種類、12 種類、10 種類、50 種類、50 種類、4 種類の値を用意し、ランダムに組み合わせさせて 84 個の番組からなる番組表を作った。ユーザは、番組表から 1 つ番組を選択し、その番組を予約するかその番組と同じ条件の番組を検索する。

模擬ユーザ (以下では、ユーザと呼ぶ) は、システムの実現できる要求をランダムに選択し、その要求を達成するために必要な属性の値をシステムに伝える。このとき、ユーザは未知語率に従って要求を表現する。例えば、「今日のドラマ」という要求を表現する場合、(日付: 今日, ジャンル名: ドラマ) のように属性と値の組の列で表現する。未知語率が 0.6 であれば、各属性の値を 0.6 の確率で未知語で表現する。ユーザは未知語率に従って要求を表現するので、未知語率が 1 でない限り、システムがある属性について質問を繰り返すと、ユーザはその属性の値を既知語で表現する。ユーザのこのような設定は、システム 1 やシステム 2 にとって不利でない実験条件にするためである。なぜなら、ユー

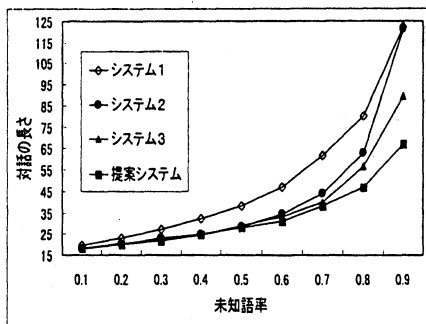


図 3: 実験結果 (置換誤り率 = 0.2)

ザが未知語で要求を表現し続けると、システム 1 とシステム 2 は、無限に質問をし続けるからである。

システムとユーザの対話は、2 つのフェーズに分割できるとする。1 つは、システムがユーザに質問または確認を行い、ユーザの要求を決定する確認フェーズである。もう 1 つは、ユーザの要求に合致する情報を提供する情報提供フェーズである。対話の効率性の評価基準は、タスク完了までの対話の長さとする。対話の長さとは、各フェーズの対話の長さの和とする。確認フェーズの対話の長さは、システムとユーザの間でやりとりされた属性の個数とする。情報提供フェーズの対話の長さは、システムがユーザに提示した情報の個数とする。

音声認識誤りをシミュレートするために、ユーザがシステムに情報を伝えるとき、0.2 の確率で属性の値が置換されることにする (置換誤り率 = 0.2)。

評価実験では、次のシステムを比較する。

- システム 1: 未知語を既知語に誤認識する
- システム 2: 未知語を検出し、棄却する
- システム 3: 出現頻度の高い属性を未知語の属性と推定し、提案システムと同じ未知語の処理を行う

3.3 評価実験

各システムについて、未知語率を変化させて 1000 対話の平均の対話の長さを計算した結果、図 3 が得られた。提案システムは属性推定に用いる決定木を作成するために、模擬ユーザとシステムの対話から 700 の学習データを収集し、決定木を作成した。提案システムの属性推定正解率は、71% であった。

未知語率が高い状況は、未知語を含む発話をユーザが繰り返す状況である。この状況では、提案システムの効率が良い。これは、提案システムが未知語の属性についての質問を繰り返さないのに対し、システム 1 とシステム 2 は、未知語が既知語に言い換えられるまで質

問を繰り返すからである。システム 3 の効率が悪いのは、属性推定正解率が低いために、未知語の属性を確認する対話が長くなるからである。

未知語率が低い状況は、未知語を含む発話をユーザが繰り返さない状況である。この状況では、システム 2 の効率が最も良いが、提案システムもシステム 2 と比較して効率性を損なわずにタスクを遂行している。システム 2 が提案システムと比較して効率が良いのは、未知語が既知語に言い換えられるまでの対話の長さ、未知語の属性を確認する対話の長さと比較して短くなるからである。しかし、未知語の属性を確認する対話の長さは、未知語が既知語に言い換えられるまでの対話の長さと比較して、非常に長いものではないために、提案システムは効率性を損なわない。

4 おわりに

本稿では、決定木を用いた未知語属性推定方法を提案し、従来の実例の距離に基づく方法を改良した方法と比較して属性推定正解率が高いことを検証した。また、推定した属性を利用した効率的な対話方法を提案し、未知語を含む発話をユーザが繰り返す状況では、提案システムが効率的に対話を行い、未知語を含む発話をユーザが繰り返さない状況では、提案システムが効率性を損なわずに対話を行うことを検証した。

謝辞 日頃より御指導頂くメディア情報研究部 村瀬洋部長、有益な示唆を頂くマルチモーダル対話研究グループの諸氏に感謝致します。

参考文献

- [1] 荒木雅弘, 堂下修司, 対話事例ベースによる発話内容の推定および未知語の解析, 情報処理学会第 49 回全国大会発表論文集 Vol. 3, pp. 155-156, 1994.
- [2] H. Meng, et al., ISIS : A Learning System with Combined Interaction and Delegation Dialogs, EUROSPEECH-2001, pp. 1551-1554, 2001.
- [3] D. J. Litman, M. S. Kearns, and M. A. Walker, Automatic Optimization of Dialogue Management, COLING-2000, 2000.
- [4] Y. Niimi and T. Nishimoto, Mathematical Analysis of Dialogue control strategies, EUROSPEECH-1999, pp. 1403-1406, 1999.
- [5] 佐藤理史, MBT1 : 実例に基づく訳語選択, 人工知能学会誌 Vol. 6 No. 4, pp. 129-136, 1991.
- [6] 高橋康博, 堂坂浩二, 相川清明, 音声対話における実例に基づく未知語属性推定, 電子情報通信学会技術研究報告, NLC2001-22~35, pp. 101-106, 2001.