

対話中に現れる未知インスタンスの オントロジーを用いたクラス同定

中野 領祐[†]

武田 龍[‡]

駒谷 和範[‡]

[†] 大阪大学 工学部電子情報工学科

[‡] 大阪大学 産業科学研究所

1. はじめに

音声対話システムにおいて、個々の対象ドメインの知識を完全に記述するのは困難である。システムが正しく応答するには、一般的な属性だけでなく、対象ドメインにおける知識が不可欠である。例えば、図1左の例のように、「中野さん」が人名であると理解できたとしても、他の概念との関連はわからず、正しく応答できない。

本研究では、対話中に現れた未知語に関する知識を獲得するシステムの構築を目指す。未知語は、システムのオントロジー内に位置づけられていない語とし、かつ、実際にはオントロジー内の最下位クラスのインスタンスであるとする。これが所属する最下位クラスを同定することで、その語に関する知識を獲得できたとする。図2のようなオントロジーを対象とした場合、例えば「中野さん」が「B4」という最下位クラスに属することを同定する。これにより、例えば「学生の予定を教えてください」「学部生の予定を教えてください」という質問に対して、「中野さん」の予定を含めて応答できるようになる(図1右)。未知語に関する属性値を推定する先行研究もあるが、このような知識の構造には着目していない[1]。

本稿では、未知語が所属するクラスを同定するために、オントロジーの木構造を利用した質問をユーザに行う手法について述べる。本手法では、図3のように、知識の木構造に基づき最適な質問を行う手法[2]の適用を試み、所属クラスを特定するまでの質問回数効率化を狙う。この際に、ユーザの発話文とオントロジー内の各クラスとの類似度を計算し、質問を効率化するための指標を得る。スケジュール登録タスクを対象とし、所属クラス特定までの質問回数によって本手法を評価する。

2. 未知インスタンス同定問題とアプローチ

2.1 問題設定

クラス同定の問題は、ユーザ発話文に未知インスタンスがあると判断された時点から、対話の開始 $t = 0$ とすると、次のように定式化される。まず、 t 回目の発話における N_t 個の単語 $w_{t,n}$ ($n = 1, \dots, N_t$) を含むユーザ発話文を $\mathbf{w}_t = [w_{t,1}, w_{t,2}, \dots, w_{t,N_t}]$ 、最下位クラス $c \in C_t$ に未知インスタンスが所属する確率を $b_{t,c}$ (ただし、 $\sum_c b_{t,c} = 1$) と定義する。また、これらの情報に基づいてシステム質問 a_t の生成を行う。この時、効率的なクラス同定は、真の所属クラス c_{opt} に対して $b_{t,c_{opt}} = 1$ であり、かつ、 t を最小化する最適なシステム質問列 $a_t(t = 0, 1, \dots)$ を選択する問題とみなせる。

本研究では、クラス同定に効果的な質問 a_t をオントロジーを用いて選択することがポイントとなる。例えば、 a_0 として、ある最下位クラスについて Yes/No 質問を行い、かつ、これが c_{opt} を確認する質問である場合、一回で同定が完了する。逆に、この質問が誤りである場合は同定までの質問回数が増加するリスクもある。一方、

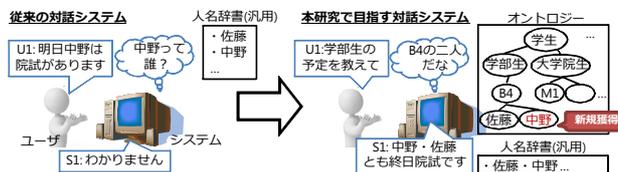


図1: 本研究の目標システム

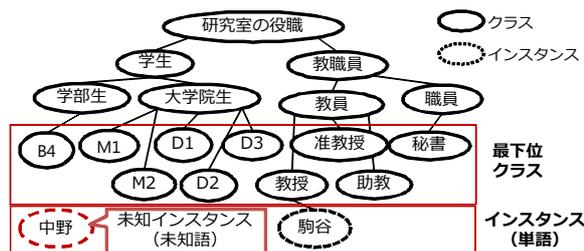


図2: 本研究室における役職オントロジー



図3: オントロジーに基づいた質問とクラス同定

オントロジーの最上位クラスから順番に質問を繰り返すことでも所属クラスを同定可能である。本稿ではこれをベースラインに定め、オントロジーのおおよその深さよりも小さい t でクラス同定ができれば、効率的に質問を設計できたとする。

2.2 アプローチと課題

本稿では問題の簡単化のため、システム質問 a_t はあるクラス c に対する Yes/No 質問に限定する。つまり、 a_t の決定は質問に適したクラス c_q を決定する問題と等価である。よって、どのクラスを質問すべきかを表す尺度 $q(c)$ を定義し、それに従い質問クラス c_q を決定する。

本研究では、知識の木構造に基づく最適な質問生成手法により尺度 $q(c)$ を定義する[2]。この手法を適用すると、所属クラス確率を表す $b_{t,c}$ に基づき、質問が誤りである場合のリスクを考慮して Yes/No 質問をすべきクラスを判断できる。以上の議論から、所属クラス確率 $b_{t,c}$ をどう推定するかが課題となる。さらに、ユーザの最初の発話文以降はすべて Yes/No で回答させるため、初期の発話文 \mathbf{w}_0 の情報を如何に活用するかが特に重要となる。

3. 単語類似度を利用したクラス同定

3.1 本手法の着目点

本稿では、初期のユーザ発話文中の単語とオントロジー上の各クラスとの類似度に着目し、質問応答に必要な未知インスタンスの所属クラスの疑似確率を計算する。

Dialogue-based Class Identification of Unknown Instances Using Ontology: Ryosuke Nakano, Ryu Takeda, and Kazunori Komatani (Osaka Univ.)

例えば、「中野は院試を受ける」という発話文において「中野」が未知インスタンスであるとする。この時、「院試」という単語から「中野」は「D3」「M1」ではなく、「B4」や「学部生」の可能性が高いと判断できる(図3)。

一方、このような類似度を用いるだけでは、オントロジー上の包含関係を考慮していないため、その情報を活用していないに等しい。この問題に対しては、発話文と各クラスとの類似度に基づいて算出する疑似確率 $\hat{b}_{t,c}$ で代用し解決する。

3.2 発話文とクラスの類似度の計算

本研究では、単語を実数ベクトルに変換する word2vec [3] を用いて、発話文とクラスの類似度計算に利用する。ここでは、発話文中の各名詞に対応するベクトルから発話文のベクトル表現を算出し、それをを用いてクラスとの類似度を計算する。

まず、word2vec により、単語 $w_{0,n}$ に対応する数値ベクトル $\mathbf{v}_{w_{0,n}}$ を得る。これらのベクトルを用いて、発話文 w_0 に対する数値ベクトルを

$$\mathbf{v}_{w_0} = \sum_n \mathbf{v}_{w_{0,n}} / \left\| \sum_n \mathbf{v}_{w_{0,n}} \right\|, \quad (1)$$

で定義する。クラス名と発話文の間の類似度は、それぞれのベクトル $\mathbf{v}_c, \mathbf{v}_{w_0}$ に基づくコサイン距離で計算する。

$$s_{c,w_0} = \cos(\mathbf{v}_c, \mathbf{v}_{w_0}) + d. \quad (2)$$

ここで、 d は類似度を正の値に保つためのバイアスパラメータである。この類似度は以降の疑似確率 $\hat{b}_{t,c}$ の計算で利用する。

3.3 類似度とオントロジーに基づく疑似確率計算

はじめに、質問の対象となるクラスの集合 C_q を導入する。この集合の $t=0$ での初期状態では、オントロジー上で質問対象となるすべてのクラス c を含む。ユーザからの Yes/No 回答によって、集合の要素は変化していく。

疑似確率 $\hat{b}_{t,c}$ は、集合 C_q 、および、初期発話文とクラスの間の類似度 s_{c,w_0} を一度最下位クラスに伝搬させたスコア $y(c)$ を用いることで計算する。オントロジーの最上位クラスから、そのスコアを子クラスへそれらの類似度の比に従って分配する。

$$y(c) = s_{c,w_0} + \frac{s_{c,w_0}}{\sum_{c' \in \text{child}(\text{parent}(c))} s_{c',w_0}} y(\text{parent}(c)), \quad (3)$$

$$\hat{b}_{t,c} = \begin{cases} y(c) / \sum_{c' \in C_q} y(c') & \text{if } c \in C_q \text{ and } c \in C_t, \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

ここで、 $\text{parent}(c)$ はクラス c の親クラスを、 $\text{child}(c)$ はクラス c の子クラスの集合を意味する。 $\hat{b}_{t,c}$ は正の値、かつ、質問の対象となるクラスに関して合計が1であるため、所属クラスの疑似的な確率値とみなせる。

3.4 曖昧性に基づく Yes/No 質問生成とユーザ応答解釈

質問対象とするクラスは、文献 [2] に従い決定する。まず、あるクラス c の質問すべき度合い $q(c)$ を、 c より下位に属する最下位クラスの疑似確率 $\hat{b}_{t,c}$ の和と 0.5 との距離により定義する。この値は、未知インスタンスが関わるクラスに対する曖昧性を表していると考えられる。ここで、 $q(c)$ が最も小さな値を取るクラスを質問に用いるクラス c_q とする。

ユーザの回答に従い、質問対象のクラス集合 C_q の要素を修正する。回答が肯定であれば、 c_q の子クラス以外

表 1: 平均質問回数

類似度での判断	発話文の数	平均質問回数	
		上位から質問	提案手法
可	13	3.59	2.92
不可	7	3.02	3.17

のクラスを C_q からすべて除外する。回答が否定であれば、 c_q の子クラスを C_q からすべて削除する。疑似確率計算と質問生成・応答解釈を繰り返し、 C_q の要素が1つになれば、クラス同定は完了する。そのクラスへ未知インスタンスを追加することで知識の獲得を達成する。

4. 評価実験

4.1 実験条件

スケジュール登録システムを実装し、具体的な発話文に対してクラス同定に必要な質問回数を評価した。評価用の発話文は、未知インスタンスであると仮定した人名と予定の内容を表す単語で構成されており、人手で20文作成した。また、20文の内、未知インスタンスの所属クラスを、単語の類似度で同定可能と思われる数を主観で判断したところ、13文となった。

word2vec の学習データは、Wikipedia 日本語版の全文章(約 2.8G)を IPADIC を用いて mecab で形態素解析したものを用いている。同様の設定で入力発話文も形態素解析を行って、単語に分割している。この時、助詞など類似度推定に不要だと判断した単語は除外し、 w_0 とした。その他のアルゴリズム・パラメータは経験的に設定している。

比較対象としては、最上位のクラスから順に Yes/No 質問でクラスを下位へ辿る方法を取り上げる。最下位クラスの1つ上の分岐では、各ノードを1ずつランダムに選択し・質問すると仮定し、その平均質問回数を入力発話文に対する質問回数とカウントした。

4.2 実験結果と考察

表1に結果を示す。類似度で判断可能と思われるものについては、上位から質問する場合よりも、提案法により平均質問回数が少ないことがわかる。本研究で着目した、発話文とクラスの類似度の利用は有効である可能性が高い。一方、それ以外については、提案手法での質問回数は逆にやや多い。これは、例えば、発話中の「出張」という単語から「教授」と「准教授」を区別できない状況などが該当する。そのため、人名につく敬称、システムの利用履歴や複数のユーザからの発話文を利用するなどして解決する必要がある。

5. おわりに

本稿では、オントロジーにない未知インスタンスの所属クラスを対話から同定する手法を提案した。ユーザの発話文とクラスとの類似度の利用とオントロジーの木構造に基づいた質問を行うことで、単純な同定方法よりも平均質問回数を約 0.6 回削減した。今後は、発話文以外の情報を利用することでより効果的な同定を目指す。

参考文献

- [1] 大塚嗣巳, 駒谷和範, 佐藤理史, 中野幹生: データベース検索音声対話システムにおける店舗属性値取得のための質問生成. 人工知能学会第 27 回全国大会, 2013.
- [2] 伊藤亮介, 駒谷和範, 河原達也: 機器操作マニュアルの知識と構造を利用した音声対話ヘルプシステム. 情報処理学会論文誌, Vol.43, No.7, pp.2147-2154, 2002.
- [3] word2vec - Tool for computing continuous distributed representations of words. - Google Project Hosting <https://code.google.com/p/word2vec/>