

情報量に基づく帰納的学習による音声対話処理手法の提案

E-26 Proposition of Spoken Dialogue Processing Method Using Inductive Learning Based on Amount of Information

木村 泰知* 荒木 健治* 栃内 香次†
 Yasutomo KIMURA Kenji ARAKI Koji TOCHINAI

1 はじめに

近年、音声認識技術の進展に伴い、音声対話システムの研究がさかに行われている。音声対話システムの構築には一般に対話システムを応用することが考えられる。しかし、人間の音声対話を書き起こしてみると、間投詞、言い誤り、言い直し、(人間には理解できる) 非文が多く存在するため、一般的な対話処理のように知識を予め与え、形態素解析、構文解析、意味解析を用いる方式では、話し言葉特有の現象に対処することは困難である。人間は上に述べた状況でも相手の意図を理解し、対話が成立している。これらの状況に対応するために、従来より我々は汎用的な学習能力のみを与え、実際の対話例から応答規則を獲得する帰納的学習を提案してきた [1]。ここで帰納的学習とは「2つの事物に対して共通部分と差異部分を見分ける能力」を用いて実際の例から規則を獲得することである。差異部分は共通部分の確定後に決まり、共通部分の定義が学習の根幹にかかわるといえる。さらに、学習能力を向上させるためにも、共通部分の拡張が必要であると考えられる。従来手法における共通部分の定義は、「字面(あるいは単語単位)による比較」により一致した部分を共通部分としていた。本稿では字面の比較により共通部分を決定するのではなく、情報量、共起関係、相互情報量に基づいて共通部分、差異部分を決定することにより応答規則を獲得する手法を提案する。本手法を「情報量に基づく帰納的学習」と呼ぶ。

2 共通部分と差異部分

最初に従来の共通部分と差異部分の獲得方法を説明する。我々は予め言語に依存した知識を与えず、実例から応答規則を獲得することを考えているため、このような方法ではシステムにとっては全ての言語は未知語となる。そのため、表1のようにギリシャ文字を使用して説明する。AとBが対話しているとする場合、A1に対するB2やA3に対するB4は何らかの因果関係を有している。そのため、Aの入力に対してBを出力する規則とする。そのような規則の中で共通部分を決定する。共通部分は下線部のように字面(または単語単位)の一致部分が現れた部分であり、共通部分以外の部分を差異部分としている。差異部分は強い因果関係が含まれていると考えられる。さらに、差異部分を変数

表1: 従来の対話規則獲得方法

A 1		$\alpha\delta\nu\beta\epsilon$
B 2		$\kappa\epsilon\zeta\vartheta\eta$
A 3		$\alpha\delta\Gamma\pi\epsilon$
B 4		$\xi\sigma\zeta\vartheta\eta$
左辺(Aの発話内容) → 右辺(Bの発話内容)		
$\alpha\delta\nu\beta\epsilon$	→	$\kappa\epsilon\zeta\vartheta\eta$
$\alpha\delta\Gamma\pi\epsilon$	→	$\xi\sigma\zeta\vartheta\eta$
対話規則		
$\underline{\alpha\delta}$	$\nu\beta$	ϵ
$\underline{\alpha\delta}$	$\Gamma\pi$	ϵ
差異部分(対応関係)		
	$\nu\beta$	
	$\Gamma\pi$	
共通部分+変数(文生成規則)		
$\alpha\delta$	@diff	ϵ
	@diff	$\zeta\vartheta\eta$

@diffは変数

化し、代入可能な文生成の規則を獲得する。差異部分の意味とは、「特徴ある対応関係」であり、共通部分の意味とは「文生成規則(テンプレート)」と考えることができる。

さらに、差異部分を獲得する条件として、各辺で+が1つとする。差異部分を複数獲得することも考えられるが、差異部分の対応関係が複数存在することとなり、再現する能力(「差異部分」と「共通部分+変数」を結合する能力)も低くなるため、複数の対応関係を獲得することは文生成の能力を低下させることとなる。

情報量に基づいて考えた場合、差異部分とは「情報量の高い部分(特徴ある対応関係)」であり、共通部分とは「情報量の低い部分(文生成規則)」と捉えることができる。

図1に従い、共通部分、差異部分の獲得方法を説明する。分割数は従来の帰納的学習に従い、3分割までとする。分割するために、隣接単語の共起頻度を利用する。Step1のように共起頻度の最も低い部分を分割点とする。3分割とするために、分割点は共起頻度の最も低い点の上位2点とする。3分

*北海道大学大学院工学研究科

†北海学園大学大学院経営学研究科

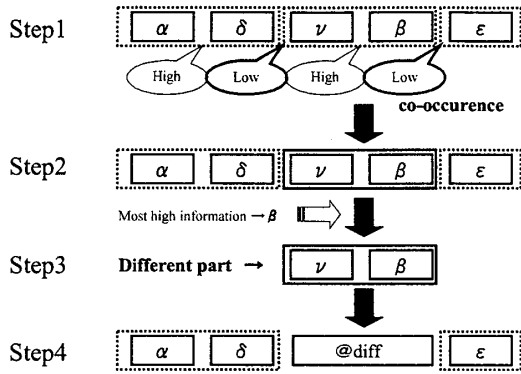


図 1: 本手法による共通部分と差異部分の獲得

表 2: コーパスの例

通訳者	1	かしこまりました。お客様のお名前のスペルをいただけますでしょうか。
申込者	2	はい。ティーエーエヌエーケーエーです。
通訳者	1	(かしこまり:178)[178](ました:313)[298](。:2243)[12](お客:47)[47](様:187)[43](の:640)[44](お:557)[40](名前:55)[2](の:640)[1](スペル:9)[4](を:263)[1](いただけ:28)[26](ます:783)[21](でしょう:88)[88](か:431)[386](。:2243)
通訳者	1	[スペルを]
申込者	2	[ティーエーエヌエーケーエー]

割された中で最も高い情報量の単語を含む部分を差異部分とする。

3 実験

実験の目的は、ATRの音声対話コーパスから情報量に基づき共通部分と差異部分に分割すること、獲得された規則から生成された応答文を評価することである。差異部分は発話の中で重要度の高い部分であり、共通部分は冗長語が多く含まれる。

実験の処理過程を説明する。対話コーパスは、予めJUMAN[2]を用いて各文の単語分割を行い、空白を区切り文字とする。実験データはATR自然発話音声・言語データベースの書き起こし、1574文(787ターン)である。ここでは、品詞情報、構文情報、意味情報は利用せずに生コーパスで行う。分割された対話を各単語ごとの頻度計算と、隣接頻度の低い部分を分割点として3分割する。情報量が高い単語を含む部分を差異部分として、差異部分の対応関係を獲得した規則を獲得する。これを対応規則と呼ぶ。さらに、差異部分のみを変数化し、文法生成規則と呼ぶ。獲得した規則を用いて、同一コーパスにより応答文を生成でき

表 3: 応答結果の分類

1	同一文を再現	25.5%(201/787)
2	意味が等しい	28.9%(228/787)
3	表現不足, 内容伝わる	7.1%(56/787)
4	誤応答	38.5%(303/787)

表 4: パターン 2 の例

応	あはい、確かにそうですね。
正	はい、そうです。
応	田中様、どういったご予約をご希望でしょうか。
正	はい、田中様。いつご予約をご希望ですか。

るか評価する。

入力文から単語頻度と隣接共起頻度を利用して、差異部分を抽出する。抽出された差異部分と一致する規則の応答を選択する。複数存在した場合には文脈情報として前回入力文の差異部分と比較する。選択された規則に対して、文法規則を利用して、自然な表現に近づける。文法規則はテンプレートになっているので、対応規則の右辺を代入することで、応答文が生成される。結合する場合、自己相互情報量の高い規則を適用する。

$$\text{自己相互情報量 } M(x, y) = \log \frac{P(x, y)}{P(x)P(y)}$$

$P(x, y)$ は単語 x と単語 y の共起確率, $P(x)$ は単語 x の出現確率, $P(y)$ は単語 y の出現確率である。

表 3 では応答生成結果を 4 パターンに分け評価した。意味が等しい文を生成した規則が 28.9% 存在した。例を表 4 に示す。パターン 1~3 を正しい応答と判断した場合、正応答生成率は 61.5% であり、情報量に基づく帰納的学習の可能性を示した。

4 おわりに

本稿では情報量に基づく帰納的学習を提案し、音声対話コーパスを用いて実験を行った。その結果、正応答率が 61.5% であった。今後は本手法を用いて実際に人間と対話を行う予定である。

謝辞 本研究の一部は文部科学省科学研究費補助金(第14658097号)によって行なわれている。

参考文献

- [1] 木村, 荒木, 桃内, 柘内: “遺伝的アルゴリズムを用いた帰納的学習による音声対話処理手法”, 信学論, Vol. J84-D-II, No.9, pp.2079-2091
- [2] 黒橋 禎夫, 長尾 真, “日本語形態素解析システム JUMAN” version3.61, 1999.