

情報量に基づく帰納的学習による音声対話処理手法の有効性について

Effectivity 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的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、言い誤りが存在するため、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。このような状況に適応するために、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用いる手法を提案してきた [1]。我々が提案した帰納的学習で仮定されている生得的能力とは「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」である。この帰納的学習において我々は学習能力を向上させるために、共通部分と差異部分の定義が重要であると考えている。従来手法における共通部分の定義は「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。本稿で提案する情報量に基づく帰納的学習では、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。本稿では、従来の帰納的学習と本手法に対して、獲得規則数及び獲得規則による文生成の精度の比較実験を行った。実験により本手法の有効性が確認されたので、その結果について述べる。

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 [1] について説明する。話者 A と話者 B が対話している場合、話者 A の発話に対する話者 B の発話は何らかの因果関係を有している。応答規則は、話者 A の発話に対して話者 B に発話を出す規則とする。獲得された規則を比較することで共通部分を決定する。共通部分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。ここで獲得された差異部分は、強い因果関係が含まれていると考えられる。さらに、差異部分を変数化し、代入可能な文生成の規則を獲得する。

情報量に基づいて考えた場合、差異部分は特徴ある部分であるため「情報量の大きい部分」、共通部分は頻繁に出現するため「情報量の小さい部分」と捉えることができる。次節において本手法の獲得方法を述べる。

3 処理過程

本システムは、学習データから応答規則と生成規則の獲得を行う。応答規則とは伝達内容を選択する規則であり、生成規則とは応答規則を基に応答文を生成する規則である。この規則の獲得方法を述べる。

応答規則は対話例から特徴的な対応関係を獲得した規則である。差異部分は共通部分より特徴ある情報が含まれるため、応答規則は差異部分の対応関係を獲得する。3分割中で最も情報量の大きい単語を含む部分が「差異部分」とし

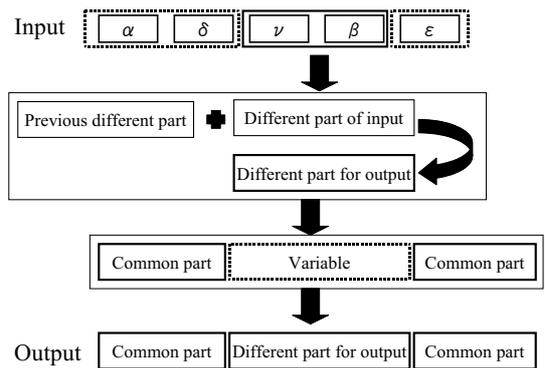


図 1: 応答生成処理

て抽出される。情報量の式を以下に示す。

$$\text{情報量 } I(x) = -\log P(x) \cdots (1)$$

$P(x) = \frac{F(x)}{N}$ であり、 $F(x)$ は単語 x の出現頻度、 N は総単語数である。式 (1) より、出現頻度の少ない単語は情報量が大きい単語となる。差異部分の対応関係は 1 対 1 ではなく、左辺が 2 つの要素に対して右辺が 1 つの要素から構成される対応関係である。左辺は前回と今回入力された文の差異部分であり、右辺は応答に使用する差異部分である。右辺が応答に使用されるキーワードとなり、生成規則と結合することで応答文となる。

生成規則は共通部分と変数化された差異部分から構成され、応答文生成のテンプレートとして利用される。変数化された差異部分に、応答規則の右辺が代入されることで応答文が生成される。

2 つの規則の使用方法を図 1 に示す。入力文が「 $\alpha\delta\nu\beta\epsilon$ 」の場合を考える。入力文は $\alpha\delta - \nu\beta - \epsilon$ と 3 分割され、差異部分が $\nu\beta$ のように決定されたとする。入力文の差異部分と応答規則の左辺のマッチングにより応答規則が選択される。応答規則の候補が複数存在した場合には、前回の差異部分のマッチングも行う。選択された応答規則の右辺を生成規則に代入することにより、応答文が生成される。

3.1 対話処理過程

対話処理の概要を図 2 に示す。本システムは音声認識ツールを用いて、ユーザ発話の認識を行う。認識結果から「差異部分の抽出」が行われ、抽出された差異部分を基に「応答規則の選択」を行う。選択された応答規則の右辺と整合性のある「生成規則を選択」して、結合する。結合された文が応答文となる。もし、差異部分の対応関係から応答生成が行えない場合には「情報量の小さい応答」により応答文が生成される。応答文は音声合成ツールを利用して、ユーザに伝えられる。本稿では、生成された文の比較実験を行う。

4 比較実験

ここでは本手法と従来の帰納的学習との比較実験を行う。比較実験の内容は、規則の獲得数と生成された応答文の精

*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

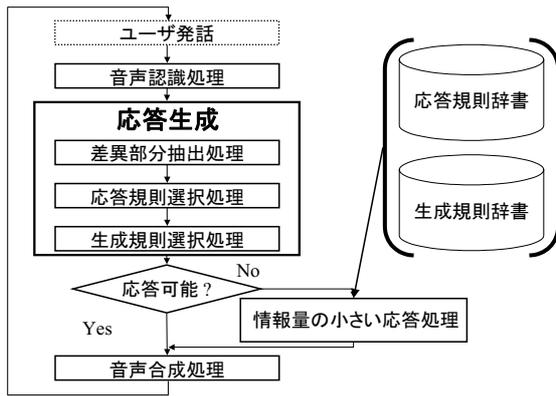


図 2: 本手法の処理過程

表 1: 獲得された規則数

	帰納的学習	本手法
差異部分	29	364
生成規則	18	367
生成文数	452	3,592
正生成率	43%	62%
誤生成率	57%	38%

度に関しての比較である。実験データは自然発話音声・言語データベース (SLDB) の 400 文 (200 ターン) とする。その対話から獲得された規則数と獲得規則による生成可能な応答文の精度について評価を行う。

帰納的学習により獲得された規則と、本手法により獲得された規則を表 1 に示す。帰納的学習により獲得された差異部分は 29 個、共通部分と変数は 18 個となり、生成可能な組合せは $29 \times 18 = 522$ 文、学習データと重複した文を削除した場合、452 文となった。本手法の獲得された差異部分は 364 個、共通部分と変数は 367 個となり、生成可能な組合せは $364 \times 367 = 133,588$ 文である。しかし、本手法では共起頻度が "0" の場合、相互情報量の計算が行えないため、生成数は 3,952 文であった。さらに、学習データと同一文の生成を削除した場合、3,574 文となった。

ここで、各手法で生成された文の精度を評価する。各文に対して、「正生成文」と「誤生成文」のいずれかで評価する。「正生成文」は文法的な誤りが含まれない文であり、「誤生成文」は文法的な誤りが含まれる文である。評価データとして、各手法により生成された文から 100 文ずつ選択する。帰納的学習の評価データは、文生成の基準がさまざまであるため、452 文からランダムに 100 文を選択した。本手法の評価データは、3,574 文から相互情報量の大きい上位 100 文を選択した。この 100 文における正生成数と誤生成数を表 1 に示す。本手法の正生成率 62% であり、帰納的学習の正生成率は 43% であったため、本手法が 19 ポイント上回った。

本手法の正生成率が上回った理由は、規則の獲得方法にあると考えられる。規則の獲得における分割箇所は、応答文生成処理に影響を与える。本手法では共起頻度の低い部分が共通部分と差異部分の分割箇所となり、帰納的学習では表層的にマッチングしている部分が共通部分であり、その共通部分の端がそのまま分割箇所となる。本手法は共起頻度の低い部分であるため、応答規則の右辺と生成規則から生成された文は比較的、自然な文となる。表 2 に同じ対話例から獲得された規則を示す。帰納的学習において獲得された差異部分は「そうです」「フロントでございます」である。本手法で獲得された差異部分は「そう」「フロント」である。本手法で獲得された差異部分は、帰納的学習で獲

表 2: 獲得規則の比較

獲得元の対話例	
受付	はい、十日にえーこちらにご到着ということでございますね。
お客	はい、そうです。
受付	はい、失礼いたします。
お客	はい、フロントでございます。
帰納的学習により獲得された規則	
差異部分	十日にえーこちらにご到着ということでございますね
差異部分	そうです
差異部分	失礼いたします
差異部分	フロントでございます
生成規則	はい、@diff。
本手法により獲得された規則	
差異部分	はい、十日にえーこちら
差異部分	そう
差異部分	失礼
差異部分	フロント
生成規則	@diff にご到着ということでございますね。
生成規則	はい、@diff です。
生成規則	はい、@diff いたします。
生成規則	はい、@diff でございます。

得された差異部分より短い単位で獲得されている。帰納的学習は表層的なマッチングであるため、共通部分は短く、差異部分は長くなる傾向にある。本手法のように短く獲得された差異部分は、生成規則として利用しやすい。本手法で獲得された差異部分は、生成規則の「はい、@diff です。」や「はい、@diff でございます。」に代入が可能である。これらの結果から、本手法が効率的に規則の獲得が行えることを確認した。

5 おわりに

本稿では、従来の帰納的学習と本手法に対して、獲得規則数及び獲得規則による文生成の精度の比較実験を行った。比較実験では、本手法の生成可能な文数が帰納的学習の約 8 倍であることを確認した。さらに、本手法は文の正生成率において、帰納的学習を 19 ポイント上回った。この結果より、本手法が共通部分と差異部分の識別を効率的に行えることを示した。

今後、本手法が固有の言語に依存しない点を確認するために英語などの他言語に対して、音声対話の実験を行う予定である。

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.