博士（工 学） 木 村 泰 知

学 位 論 文 題 名

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における性能向上に関する研究

学位論文内容の要旨

人間の優れた能力の一つに言語の使用があげられる．人間のような言語能力を有するコンピュータ，すなわち，人間と自然に対話できるコンピュータの研究が1950年代から行われており，コンピュータの発明当初，人間と同等の対話処理能力を有するシステムを実現することは時間の問題と考えられていた．しかしながら，現在，実現されたことは電話による天気予報案内，観光案内，スケジュール管理などの非常に限定された状況においてのみである．この大きな理由として，自然言語には曖昧な表現及び多様な表現が数多く存在することがあげられる．さらに，予め与える知識には限界があり，未知語処理などが必要になることがあげられる．対話システムの本質は，「どれだけ多様な表現を使って対話が行えるか」，「どれだけ自然にやりとりが行えるか」にあるといわれている．これは，音声対話には同一の内容を表現する場合にも，多様な表現が存在し，ユーザに適応した利用方法を考える必要があることを示している．

音声対話では，情報は文字言語のように符号というディジタル表現ではなく，連続した音声波形というアナログ表現で与えられる．コンピュータでは，アナログ表現からディジタル表現に変換する際に誤りが生じる．他にも音声対話の問題としては，発話自体の誤りである言い誤りから始まり，「話し言葉」特有の問題，すなわち，フィラー，言い直しが含まれる．これらの問題点を考慮しながら，頑健な処理を行う場合，予め規則を与えることでは，ユーザに依存している状況，あるいは未知の状況に適応することが困難である．すなわち，予め与える知識に加えて，対象へ適応する能力を付与することが望まれる．

本研究の目的は，音声対話処理の頑健性を向上させるために，実対話例から動的な規則を獲得することにより，ユーザあるいは，対象に依存した多様な表現を扱うことである．音声対話における表現は，会話の状況及びタスクの違いに加えて，ユーザの口癖についても考慮しなければならず，それら全て，コンピュータにとって未知の状況に遭遇したこととなる．コンピュータの場合，未知の状況での対処は困難であるが，人間は未知の状況に対して高い適応能力を有している．人間の言語獲得は連続した未知語列から言語を獲得しているため，未知語処理の問題を扱っていると考えられる．人間は，日常の言語使用を通じて新しい表現や語彙を獲得することで，言語知識を向上させている．音声対話処理においても，日常の会話を通して，言語知識を向上させる仕組みを考えるべきである．コンピュータ上で言語獲得過程を実装することを考える場合，獲得されている知識や能力に違

いがあるため，予め与える能力と獲得すべき能力に分ける必要がある．帰納的学習では，予め与える能力を「2つの事物に対して共通部分とそれ以外の差異部分を識別する能力」と仮定する．提案手法では，予め与える言語知識を少なくすることで，汎用性が高くなり，ユーザに依存した表現を獲得することが可能となる．実対話例を用いて，少量のデータから有効な応答を行うことが可能となる．まず，学習能力として，表層的な一致による共通部分とそれ以外の差異部分を利用してユーザ発話とシステム応答の対応関係を獲得する．日常会話では特に明確なタスクが設定されていないコミュニケーションが多いことから，ここでは雑談を対象とする．対話例から応答のためのルールを獲得し，交叉・突然変異・淘汰・帰納的学習という5つの処理を用いて行う応答をGA-IL（Inductive Learning with Genetic Algorithm）応答と呼ぶ．GA-ILのルール辞書に適したルールが存在しない場合，本システムは，キーワードを用いて応答を行うELIZAのアルゴリズムも用いる．このような応答をELIZA型応答と呼ぶ．ELIZA型応答はユーザとシステムの対話を継続し，対話例を獲得するために行う．雑談を対象に評価実験を行い，実対話例から学習することで実際のユーザが使用している表現や季節に依存した応答など行えることを確認した．評価実験の結果，本手法を用いた音声対話システムは，正応答と準応答の合計が76.1%となり，ELIZA型システムに対して9.6ポイントの向上が確認された．

しかしながら，音声対話は多様な表現が存在するため，表層的な比較を用いた規則獲得方法では，獲得が困難な場面も存在する．そこで，学習能力を向上させるために，情報量に基づいた重要部分の抽出に基づく規則獲得手法を提案する．学習能力を向上させるために，共通部分としていた表層的な一致とそれ以外である差異部分の考え方を変更する．情報量に基づき共通部分と差異部分を決定することにより，効率的に規則を獲得する．間投詞は個人ごとに表現や回数に違いがあるため予め設定するのは困難であるが，比較的出現回数が多いため，情報量が小さくなることを利用する．本手法は情報量の大きい部分に基づいて応答文を生成することで，間投詞，雑音による影響を削減する．ここでは，ホテルクラークとお客の対話を書き起こしたATRのSLDBコーパスを用いた評価実験に関して説明する．表層的な共通部分による帰納的学習との文生成率を比較した結果，19ポイントの向上が確認された．

また，音声対話の多様性を解決するために，自立語及び情報量の大きな単語のみを扱うのではなく，全ての単語に重み付けを行うことで，発話文と学習データ中の文の類似性を測定するアプローチを試みる．これは，自然言語処理には同じ意味を持つ表現が数多く存在するため，ユーザの発話に注目し，発話の同定を行う．しかし，字面が完全に一致していない場合，一致率の高い文が必ずしも同じような内容をあらわしているとは限らないため，それぞれの対象に適した計算方法が必要となる．提案手法では，各発話文を決定された重みを付与したベクトル表現に変換し，ユークリッド距離の近い文を同じ内容を持つ文とする．ここでは，各単語に対して情報量による重みと単語の出現頻度を掛け合わせる．これを$tf \cdot AoI(termfrequency \times AmountofInformation)$と呼ぶ．この提案手法を評価するために，類似度を利用した発話文の同定を行う．提案手法に基づいて構築した対話処理による評価実験の結果，字面上の一致率による選択方法に対して13ポイント，“$tf \cdot idf$の重み付けを用いたユークリッド距離に基づく類似度”に対して6.5ポイントの正応答率の向上が確認された．

以上の結果から，情報量により重要部分の抽出及び$tf \cdot idf$の重み付けによる類似度評価を用いることにより，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における性能向上が確認された．

学位論文審査の要旨

主　査　教　授　荒　木　健　治
副　査　教　授　青　木　由　直
副　査　教　授　北　島　秀　夫
副　査　助教授　伊　藤　敏　彦

学位論文題名

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における性能向上に関する研究

著者は，音声対話システムにおける多様な表現に関する問題を解決するために，遺伝的アルゴリズムを用いた帰納的学習による手法の提案を行い，性能向上に関する研究を行った．著者は，近年の音声認識技術の向上に伴う音声対話処理の需要を考慮し，言語的なアプローチによる音声対話処理に着目した．音声対話には，発話自体の誤りである言い誤りから始まり，「話し言葉」特有の問題であるフィラー，言い直しが含まれる．そこで，これらの問題を解決するために，著者は実対話例を用いることで各対象に内在する対話規則の獲得を試みた．対話規則を獲得する能力を高めることに着目し，予め与えた言語知識を少量にすることで，汎用性を高め対象へ適応することを可能とした．

音声対話における表現は，会話の状況及びタスクの違いに加えて，ユーザの口癖についても考慮しなければならず，それら全て，コンピュータにとって未知の状況に遭遇したこととなる．コンピュータとは違い，人間の場合は未知の状況に対して高い適応能力を有している．人間の言語獲得は連続した未知語列から言語を獲得しているため，未知語処理の問題を扱っていると考えられる．人間は，日常の会話を通じて新しい表現や語彙を獲得することで，言語知識を向上させているため，著者のシステムは，日常の会話を通して言語知識を向上させる仕組みを実現させている．著者が提案する帰納的学習では，予め与える能力を「2つの事物に対して共通部分とそれ以外の差異部分を識別する能力」と仮定する．ここで共通部分は字面の比較により一致した部分とする．予め与える言語知識を少なくすることで，汎用性が高くなり，ユーザに依存した表現を獲得するという理由からである．帰納的学習だけでは効率の良い規則獲得が困難であるため，遺伝的アルゴリズムの交叉・突然変異・淘汰の考え方を導入することで少量の対話例における問題を解決する．

提案手法の持つ学習能力の高さを実証するために，数多くの話題が混在する雑談を対象としたシステムの構築を行い，キーワード及び文生成テンプレートを利用したELIZAシステムとの比較実験を行う．雑談を対象とした1,000ターンの評価実験では，実対話例か

ら学習することで実際のユーザが使用している表現や季節に依存した応答等が可能であることを確認した．評価実験の結果，本手法を用いた音声対話システムは，意味が正しく表現が自然である正応答と意味は正しいが表現が不足している準応答の合計が76.1%となり，ELIZA型システムに対して9.6ポイントの向上が確認された．

しかしながら，音声対話は多様な表現が存在するため，字面の比較を用いた規則獲得方法では，表層的に類似した2例が存在しない場合には規則の獲得が困難となる．そこで，著者は学習能力を向上させるために，情報量に基づいた重要部分の抽出に基づく規則獲得手法を提案した．学習能力を向上させるために，共通部分としていた表層的な一致とそれ以外である差異部分の考え方を改善する．情報量に基づき共通部分と差異部分を決定することにより，効率的に規則を獲得する．間投詞及びユーザの口癖は個人ごとに表現や回数に違いがあるため予め設定するのは困難であるが，比較的出現回数が多いため，情報量が小さくなることを利用する．提案手法の持つ学習能力の高さを実証するために，ホテルクラークとお客の対話例であるATRのSLDB（Speech and Language DataBas）コーパスを用いた評価実験を行った．表層的な比較を用いた規則獲得を行う帰納的学習との文生成率を比較した結果，19ポイントの向上が確認された．

情報量を用いた手法では，音声対話の多様性を解決するために，重要部分を情報量により抽出したが，不要語として扱う単語が多くなり，利用可能な単語まで不要語とするという問題が存在した．そこで，類似性の考えに基づいて，全ての単語に重み付けを行うことで，不要語処理を行わずに発話文と学習データ中の文の類似性を測定した．これは，自然言語処理の多義性及び多様な表現の問題を解決することになる．著者は，各発話文を決定された重みを付与したベクトル表現に変換し，各単語に対して情報量による重みと単語の出現頻度を掛け合わせることにより，ユークリッド距離Dを求め，類似度を$Sim = \frac{1}{D+1}$とした．これを$tf \cdot AoI(termfrequency \times AmountofInformation)$の重み付けを用いたユークリッド距離に基づく類似度と呼ぶ．この提案手法を評価するために，類似度を利用した発話文の同定を行う．提案手法の持つ学習能力の高さを実証するために，ATRのSLDBに含まれる約12,000発話から，同じ内容で異なる2つの表現の発話を154組収集し，字面の一致率による選択方法と，他の文書に出現しないタームほど文章の特徴をあらわすタームとして重みを大きくする$tf \cdot idf(termfrequency \times inversedocumentfrequency)$の重み付けを用いたユークリッド距離に基づく類似度との比較実験を行った．提案手法に基づいて構築した対話処理による評価実験の結果，字面上の一致率による選択方法に対して13ポイント，$tf \cdot idf$の重み付けを用いたユークリッド距離に基づく類似度に対して6.5ポイントの正応答率の向上が確認された．

以上を要約すると，音声対話における表現の多様性に関する問題を解決するために，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の提案を行った．さらに，情報量による規則の獲得及び類似度による文の比較を行うことで，音声対話システムの性能を向上させることが可能であることを示した．本研究を通じての，情報メディア工学，自然言語処理工学の発展に貢献するところ大なるものがある．よって，著者は北海道大学博士（工学）の学位を授与される資格あるものと認める．