

論文審査の結果の要旨

論文提出者氏名 小松 孝徳

本研究の最終的な目的は、ユーザとの相互適応的な関係を通じた自然なコミュニケーションを成立させることのできる、適応的な音声インターフェイスを構築することである。ここで相互適応的な関係とは、ユーザがインターフェイスに対して適応学習し、その適応に基づいてインターフェイスもユーザに適応学習するというプロセスを繰り返す関係のことである。本論文では、そのための第一歩として、ユーザの適応を促しかつその適応を利用するという相互適応の一部分をユーザとの間に実現することで、ユーザの発話意図を理解する意味獲得モデルの構築を行っている。

第1章では、上記のような本論文の目的を説明している。

第2章では、本論文に関連した先行研究との比較を行っている。具体的には、(1) 適応する存在としての人間の認知的特性に注目し、それを具体的に観察・分析している点、(2) ユーザの適応を促しかつその適応を利用することでユーザの発話意図を理解する意味獲得モデルを構築している点、(3) ユーザの発話と行動とのマッピングを学習する意味獲得モデルを構築する際に、韻律情報に含まれる普遍的な情報を利用している点、に本論文の独創性があることを説明している。

第3章では、まず人間の適応能力を観察・分析する実験を行った上で、そこで観察された人間の適応能力を利用して、人間の発話意図を理解する意味獲得モデルの構築を行う、という本論文のアプローチが示されている。

第4章では、簡単なピンポンゲームを用いて人間の適応能力を観察するコミュニケーション実験について説明している。この実験では、ボールが隠されていて見えない操作者に別室の教示者が教示を行うが、その際、操作者は、教示者の教示発話中の音韻情報が利用できずに、発話中の韻律情報しか手がかりにできないような環境を設定し、教示者の発話意味を操作者がどのように理解していくのかを観察している。その結果、音韻情報としては未知の教示の意味を獲得することができた被験者ペアでは、お互いが相手のことを学習しながら相手に徐々に適応していくという相互適応的なプロセスが観察され、以下に示す点が、意味獲得の成功要因だと報告している。すなわち、教示者側にあつては、(1) 使用する教示の種類を減少させ、(2) 警告韻律を用いることで操作者の行動に注意を与えている点、操作者側にあつては、(1) 発話を与えられてから行動することで、現在の教示に対する理解度を教示者に示し、(2) 異なる種類の発話に対しては異なる行動をとり、(3) 与えられた警告韻律を利用して自分の行動を修正している点、である。

第5章では、コミュニケーション実験で観察された操作者側の学習を実現するような、発話意味を獲得する学習モデルの提案を行っている。このモデルは、ラケットにボールを当てられたときに与えられる正の報酬を受けた際の直前の教示音声・行動のデータ（音声－行動データ）を蓄積し、教示意味を表す複数のクラス（分布）にこのデータを分類することで発話の意味学習を実現している。この際、正規混合分布から音声－行動データが生成されると仮定したため、それぞれの正規分布のパラメータ（平均値・分散）を求めることが教示の意味学習に相当する。具体的には、EMアルゴリズムを用いて混合分布中の各分布のパラメータを推定しているが、従来のEMアルゴリズムでは負の報酬（警告韻律と呼ばれる、教示者が操作者に注意を与えた場合に現われる普遍的な韻律パターン）を学習に利用できないので、従来型のEMアルゴリズムのE stepを拡張した拡張型EMアルゴリズムを提案している。

第6章では、この提案された意味獲得モデルが、教示者とインタラクションしながら発話の意味を学習するモデルとして適しているのどうかを実験的に検討している。具体的には、第4章の

コミュニケーション実験で人間の操作者が操作していたピンポンゲームのラケットに第5章で提案した意味獲得モデルを実装し、実際に一般のユーザからリアルタイムで教示を受けた際のモデルの学習状況を観察している。その結果、実験者から事前に与えられるインストラクションの種類に依存するものの、このモデルが、一般のユーザの適応を促しかつそのユーザの適応を利用するという相互適応の一部分を構築することで、与えられた教示の意図通りに行動できたことを報告している。その際、韻律情報における普遍的な機能の一つとみなされる警告韻律を、意味学習において負の報酬として利用できたことも報告している。

第7章では、この意味獲得モデルによって実現できた、ユーザとインターフェイス間の部分的な相互適応を、人間同士で見られる真の意味での相互適応に発展させるための課題点について述べている。また、構築された意味獲得モデルが、インターフェイスの自動カスタマイゼーション技術や、人間とコミュニケーションが可能なペットロボットの基礎技術、などに応用可能なことを議論している。

第8章では、本論文の成果をまとめている。

従来のインターフェイス研究では、人間の適応性が考慮されておらず、本論文のように具体的な設定に基づく分析やモデルの構築はなされていなかった。これに対し、本論文では、人間の適応性という認知特性を実験的に分析した上で、実験者から与えられるインストラクションの種類には依存するものの、一般のユーザの適応を促しかつその適応を利用して、ユーザから与えられる発話の意図を理解できる意味獲得モデルを提案している点に、第一の独創性がある。さらに本論文では、従来の音声インターフェイス研究の多くが注目している音韻情報ではなく、韻律情報に焦点を当てている点にも独創性がある。

以上に記した通り、本論文は、ヒューマン・コンピュータ・インタラクション研究や適応的インタフェース研究において十分な独創性を有しており、その成果を将来的に工学システム（インターフェイスやロボット）へ応用できる可能性も高い。したがって本論文が、これらの分野に大きな貢献をもたらすものであると判断し、博士（学術）の学位を授与するにふさわしいものと審査委員会は認定する。