

音によるヒューマン・エージェント・インタラクションのための プロトタイプシステムの試作

棚橋 徹^{a)} 小林 一樹^{b)} 北原 鉄朗^{c)}

^{a)} 日本大学 大学院 総合基礎科学研究科

^{b)} 信州大学 学術研究院

^{c)} 日本大学 文理学部

1. はじめに

近年, Papper¹⁾ や Sota²⁾ を筆頭に多くの対話型ロボットの研究, 開発が進められており, 人とロボットがコミュニケーションを行う場面は数多く存在するようになりつつある。また, 一対一の対話だけでなく複数ユーザや複数エージェントによるコミュニケーション^{3),4)} も研究がなされている。このようなロボットやエージェントとのコミュニケーションで最も一般的なものは, 日本語などの自然言語による会話である。しかし, 音声認識や言語理解の点で難しい課題が少なくない。

一方, 犬や猫などのペットは自然言語によるコミュニケーションが全くできないにも関わらず, 鳴き声の高さや長さなどから人間は親しみを感じたり, 敵意を感じることができると考える。このような, 自然言語を用いない音によるコミュニケーションによって, ペットに対して人間側が心を通わせていると考え, 家族の一員として扱われることもある。

そこで本研究では, 家庭内で言語情報を持たない音を用いて人間とコミュニケーションを行うエージェントの実現を目指す。また本稿では, その実現のために, 試作したプロトタイプシステムについて述べる。

2. システム構成

コミュニケーションとは, 感情や言語情報を伝達をしようことである。特に言語情報を持たない音では, 感情の伝達によってコミュニケーションを図っていると考え。そこで, 本研究の目標は次の3点の事柄を検証することとする。

- 人間は, 言語情報を持たず, 周波数や振幅の単純な変化のみを持つ音しか発しないエージェントに対して, 感情を見出すことができるか。
- 感情を見出す上で エージェント同士のやりとりを観察することが有効に働くかどうか。
- 人間がエージェントに対して感情を見出すことで, 人間とエージェントの間に人間とペットのような関係性を構築することができるか。

これらの事柄を確かめるために, 次のような仕様のプロトタイプシステムを試作する。

- (1) 画面上に2体のエージェントが表示され, 各エージェントが音を発する(ただし, 片方が音を発している最中は他方が音を発しないものとする)。
- (2) 発する音はできるだけ単純なものとし, 言語情報は一切持たない。

- (3) 発する音の特徴(周波数, 振幅, 反復回数など)は, エージェント内部の感情に基づく。感情はVAモデル⁵⁾を離散化した平静, 喜び, 楽しみ, 怒り, 哀しみの5種類とする。

- (4) ユーザの発する声は各エージェントが認識する。ただし, 韻律のみを認識するものとし, 言語情報は一切認識しない。

- (5) ユーザの発した声に呼応して各エージェントの感情が遷移し, その感情に基づき音を発する。これによってインタラクションを行う。

- (6) ユーザの声に対する呼応の仕方はエージェントによって違いを持たせる。

本プロトタイプシステムを用いた対話の一例を図1に示す。ただし, 現状では(1), (2)と(5)の一部を実装したのみである。(3)の各感情に基づく音と(4)~(6)のユーザとのインタラクションについては今後の課題とし, 実装されていない。そのため現状のプロトタイプシステムでは, 2体のエージェントが発する音をユーザが聴くことや, ユーザ, もしくは一方のエージェントが発話した際に他方のエージェントが相槌を行うのみである。

2.1 外見の設計

プロトタイプシステムのエージェントの外見を図2に示す。本研究では音によるコミュニケーションが目的のため, 外見はできるだけシンプルになるように設計を行った。しかし, どちらが音を発しているかはユーザが観察する際にわかるようにするため, 発話しているエージェントの目を赤くすることとした。

2.2 発する音の設計

一切言語情報を持たない音を発するために, 原波形(正弦波, 矩形波)に対して, 振幅と周波数のみを変化させて音を作成する。振幅と周波数は, 効果音ラボ⁶⁾の犬の鳴き声から抽出し, 更にそれを加工し決定する。加工は, 全体の振幅や周波数を上げ下げする他, 部分的に上げ下げ, 話速度, 発話時間の変化を行う。ただし, 加工の際のパラメータは本来エージェント内部の感情に基づき決定すべきだが, 現状のプロトタイプシステムでは犬の鳴き声から抽出した振幅と周波数を原波形に加工したのみである。

2.3 発話タイミングの設計

本来は, ユーザもしくは他のエージェントが発した音に呼応して発話タイミングを考慮すべきだが, 現状のプロトタイプシステムでは, 特定のタイミングでのみ発話, もしくは相槌を行う。具体的には, 発話は誰も音を発しない時間が長時間継続しないよう, 4秒間無音区間が連続した場合にエージェントが音を発することとした。また相槌としての発話は, 文献⁷⁾を参考に, 200ms以上のユーザ, もしくはエージェント

A Prototype System for Sound-based Human-Agent Interaction

by Tetsu Tanahashi (Nihon University), Kazuki Kobatashi (Shinshu University) and Tetsuro Kitahara (Nihon University)

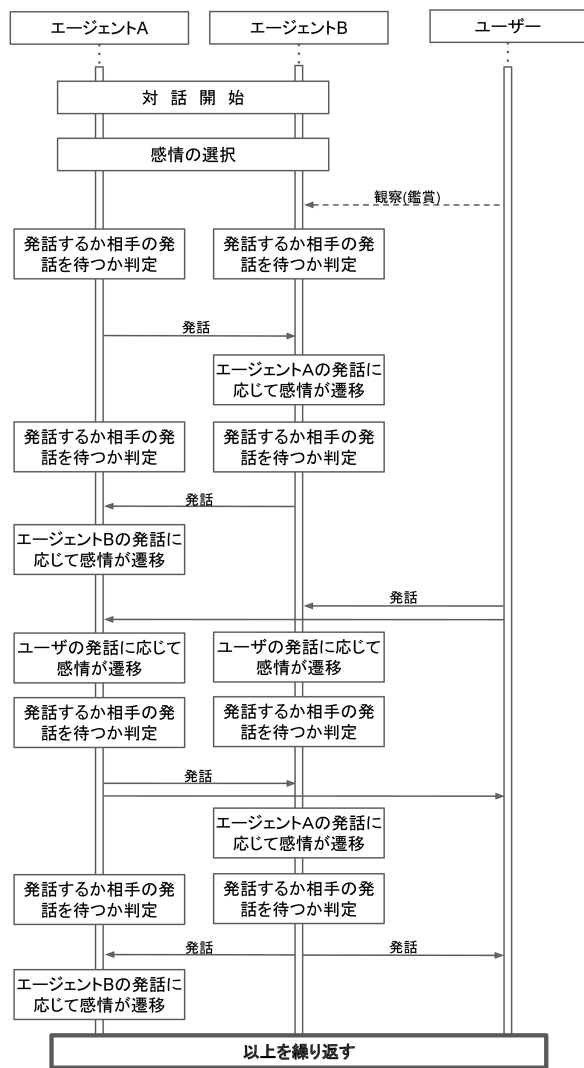


図1 本プロトタイプシステムを用いた対話の一例

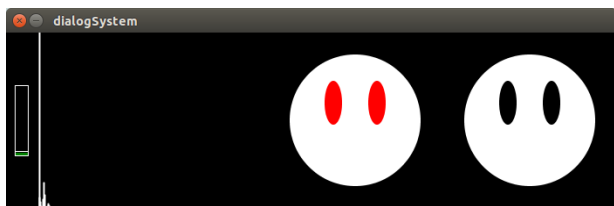


図2 左側のエージェントが発話している際の外見

の発話の後に 200ms 以上の無音区間を検出した場合に音を発することとした。

3. 今後の方針

本節では、今後の方針として大きく 3 点について述べる。

3.1 エージェント同士のインタラクション

現在のプロトタイプシステムでは、一方のエージェントの発話に対して、他方のエージェントがランダムに感情を遷移し発話するシステムになっている。本来、エージェントの感情遷移は一方のエージェントの発話に呼応して感情が遷移す

るべきである。そこで、一方のエージェントの音から振幅や F0 などの韻律の特徴量を抽出し、それに応じて他方のエージェントの感情が遷移し、その感情に従ってエージェントが発話する処理をシステムに加えていきたい。また、人間でも性格が異なるように、呼応の仕方も各エージェントで異なるべきだと考える。例えば、楽しい方向に感情が遷移しやすいエージェントや哀しい方向へ感情が遷移しやすいエージェントなど、それぞれのエージェントに感情の遷移の傾向を持たせることも検討していきたい。

3.2 エージェント同士のインタラクションの観察

エージェント同士のインタラクションが可能となった場合、人間がそれを観察した際にそれぞれのエージェントの発する音から感情が見出だせるのか実験によって確かめたいと考えている。これは、本研究の目標で掲げた、人間は言語情報を持たない音しか発しないエージェントに対して感情を見出だせるのか、また、感情を見出す上でエージェント同士のインタラクションは有効に働くかを検討するためである。もし実験を行った結果、人間が感情が見いだせない場合、エージェントが発する言語情報を持たない音の改善や感情の遷移の方法の検討を行っていきたい。

3.3 人間を含めたインタラクション

3.1, 3.2 節の実装、実験の後に人間を含めた 3 者によるインタラクション (鼎談) を検討している。これは、人間がエージェントから感情を受け取ることによってどのように感情が遷移するのか、また、韻律などはエージェントに影響されるのか、持続的なインタラクションを構築できるかなど実験によって確かめたいと考えている。

4. おわりに

本稿では、音によるヒューマン・エージェント・インタラクションのためのプロトタイプシステムの試作について述べた。今後は、ユーザとのインタラクションを実装し、感情に基づいた音の生成や、ユーザやエージェントの発話に呼応する音の生成、外観などに関して改良を行い、ユーザに使用してもらうことによって 2 章で述べた 3 点について検証していきたい。

謝辞 本研究は、日本学術振興会の科学研究費補助事業 26280089, 26118005 から支援を受けたものである。

参考文献

- 1) Pepper, Softbank, URL:<http://www.softbank.jp/robot/consumer/products/>
- 2) Sota, ヴインストーン株式会社, URL:<https://www.vstone.co.jp/products/sota/>
- 3) 藤堂祐樹, 西村良太, 山本一公, 中川聖一, 単一対話エージェントと複数対話エージェントを用いた音声対話システムの分析と評価, 情報処理学会研究報告 (SLP), 19, 1-8, 2012.
- 4) 乙木翔地, 堀田怜, 黄宏軒, 馬場直哉, 中野有紀子, 川越恭二, 複数人ユーザ会話におけるエージェントの割り込みタイミングの推定手法の検討, HAI シンポジウム, 2012.
- 5) James A. Russell: A Ciucumplex Model of Affect, Journal of Personality and Social Psychology, 39(6), 1161-1178, 1980.
- 6) 効果音ラボ, URL:<http://soundeffect-lab.info/sound/animal/>
- 7) 小林一樹, 船越孝太郎, 小松孝徳, 山田誠二, 中野幹生, ASE に基づく相槌によるロボットとの対話体験の向上, 人工知能学会論文誌, 30(4), 604-612, 2015.