

## 2019 年度 卒業論文

# NAO を活用した音声認識による対話システムの作成

龍谷大学 理工学部 数理情報学科

T150003 池田 翠

指導教員 佐野 彰

## 概要

人工知能 (AI) の発達により、私たちの生活はこれまでよりも便利に、快適に変わってきた。そんな AI の中でも近年注目されているのが音声認識の分野である。音声認識が注目されている理由の 1 つとしてロボット分野での活躍が挙げられる。ロボットが人の代わりに店頭などで接客をしたり、観光地で外国人観光客向けの案内ロボットとして活躍したりする光景を見ることが多くなった。ロボットと人のより自然なコミュニケーションを支えているのが音声認識の技術である。音声認識の現状を知るとともに音声認識のシステムを実装したいと考えたことが本研究を始めるきっかけとなった。

本研究では、理工学部で実施されたオープンラボの研究室公開の出展一覧のデータをもとに、オープンラボでの案内を想定した対話システムの作成に取り組んだ。

対話システムのインターフェイスとして、アルデバランロボティクス社 (現 ソフトバンクロボティクス社) が開発した「NAO」を用いて行った。システム作成には NAO でアプリケーションの開発を行うための公式の開発ツールである Choregraphe と外部の音声認識サービスを活用したものと 2 つの手法を用いて作成し、それぞれのシステムについての動作検証を行う。

対話システムの実装には、Choregraphe を用いて Dialog ボックスと呼ばれる NAO の発話を決定するソースファイルの作成が必要である。Dialog ボックスは、QiChat スクリプトと呼ばれる Choregraphe 上でプログラミングを行う際に使用する専用の言語で記述する。しかし、あらかじめ定義された言葉のみを認識するため、オープンラボで聞かれるであろう質問と回答を全パターン考え、記述しなければならない。そこで、QiChat の機能を活用し重複した会話を concept と呼ばれる変数リストに入れておくことで会話パターンを宣言し、文中から concet を呼び出すことのできるため、効率的にプログラムを作成した。また、同じ内容でも人によって質問の仕方が変わることを考慮して、QiChat の機能で任意の言葉や文章を認識できる Wildcard(\*) という記号を用いて NAO が応答できるようにした。

Choregraphe での対話システムは実際にオープンラボで動作検証ができ、音声の聞き取りと応答に成功した。しかし、言葉により認識できず応答できない場合が多く見られたことから NAO 本体の音声認識では精度が低いと判断した。

外部の音声認識サービスを活用したシステムの作成では、Python を中心とした各機能ごとに作成したシステムを組み合わせたものを最終的に NAO と実際に接続して対話システムとして音声認識の精度を検証することを目標とし本研究を進めた。外部の音声認識サービスには、Google Speech Recognition を選択し、人とコンピュータ間で音声ファイルを用いての音声認識とリアルタイムでの音声認識を行った。また、データの読み込みを行い、質問と回答の記述をシステム化させた。

本研究では、Choregraphe では対話システムを完成させることができた。Speech Recognition では、NAO をインターフェイスとしてリアルタイムで質問と回答を返すシステムの作成と動作検証が未完成である。今後の課題は、未完成である Speech Recognition の対話システムを完成させ、NAO で実際に動作検証し、音声認識の精度を上げることである。