

技術・研究報告

テーブルトップ型コミュニケーションロボットを利用したクイズシステムの開発

原口和貴, 林拓実, 瀬澤樹, 梅谷智弘, 北村達也

甲南大学 知能情報学部 知能情報学科
神戸市東灘区岡本 8 - 9 - 1, 658-8501

(受理日 2017 年 5 月 9 日)

概要

ヴイストン社のコミュニケーションロボット Sota は、頭部、胸部、両腕から成る高さ 28 cm のロボットである。このロボットは、専用の GUI ソフトウェアもしくは Java プログラムによりその動作を制御することができる。本研究では、クラウドサービスを利用した音声認識・合成、Sota 頭部のカメラを利用した画像認識を活用したクイズシステムを開発した。このシステムは、まず、顔検出によりユーザーがロボットの前に立ったことを認識する。そして、ユーザーと挨拶を交わした後、クイズを出す。ユーザーは、音声による返答、もしくは QR コードが印刷されたカードの提示によって Sota に回答を示す。オープンキャンパスにて本システムの実演を行い、ロボットとコミュニケーションする楽しさを来場者に提供した。

キーワード: ロボット, Q&A システム, 音声認識・合成, 画像認識, QR コード

1 はじめに

コミュニケーションロボットとは、ユーザーの顔、表情、音声などを認識し、それに応答することができるロボットである。日本においては、2015 年の Pepper (ソフトバンクロボティクス社) 発売により、コミュニケーションロボットが一般に認知される存在となった。その後のコミュニケーションロボットの普及スピードは目覚ましく、商業施設、公共機関など様々な場所で Pepper をはじめとする様々なロボットと人々がコミュニケーションをとる場面を目にするようになった。

コミュニケーションロボットの要素技術は、ロボット工学、メディア情報処理、パターン認識、プログラミング、人工知能など本学部 (甲南大学知能情報学部) の教育・研究分野と密接な関連がある。したがって、コミュニケーションロボットは、本学部の教育・研究内容を対外的にわかりやすく提示するのに適した素材の 1 つである。そこで、本研究ではヴイストン社のテーブルトップ型コミュニケーションロボット Sota [1] を用いてクイズシステムを開発し、オープンキャン

ンパスや科学体験イベントなどの一般向けイベントにて幅広い年齢層の来場者, 参加者にロボットとのコミュニケーション体験を提供することを目的とする。

本稿ではまず Sota の仕様や機能について概説する。次に, 音声対話に基づくクイズシステムについて説明する。このシステムでは音声認識を利用してユーザーの回答の正誤を判定する。さらに, QR コードを利用したクイズシステムについて説明する。このシステムは, 周囲雑音が大きく音声認識の利用が難しい状況でも正確にユーザーの回答を得ることができる。

2 Sotaの仕様・機能

ヴイストーン社Sotaは図1に示すように, 頭部, 胴部, 両腕から成る。頭部にはLEDによって色が変わる目や口があり, そのサイズ(高さ28 cm)もあいまって, 親しみやすいデザインとなっている。頭部3軸, 胴部1軸, 腕部各2軸の計8自由度を持っており, 多彩な動作が可能である。メディア入出力機能としては, 頭部のカメラ, マイク, 胸部のスピーカーがある。胴部には制御用のマイコンボードが格納されており, Raspberry Piを利用した機種とIntel社のEdisonを利用した機種が存在する。本研究では前者を利用している。Raspberry PiにはWiFi子機アダプターが接続されており, インターネット接続が可能である。そして, これを利用した音声認識・合成のクラウドサービスが提供されている。



図1: ヴイストーン社Sota

Sotaの制御を行うプログラムの開発は, ヴイストーン社が提供するソフトウェアVstoneMagicもしくはJavaにより行うことができる[2]。前者はWindows上で動作するソフトウェアであり, 様々な機能を有するブロックをGUIによる操作にて連結することによってプログラミングを行う。ブロックは, Sotaの動作やLEDの色の指定, 音声認識・合成, 顔画像処理, ループ・分岐, 通信などの機能を持つものが提供されている。作成されたプログラムは, Javaプログラムに変換され, Sota本体に送信された後に実行される。したがって, SotaはJavaにより直接プログラミングすることも可能である。

3 音声対話に基づくクイズシステム

3.1 クイズシステムの概要

本研究で開発したクイズシステムのフローチャートを図2に示す。システムが起動すると、顔検出機能によりユーザーの顔を探査する。このとき、Sotaの頭部が動いて周囲を探査する。ユーザーの顔を検出すると、目のLEDの色を変化させ、合成音声で「こんにちは」と挨拶する。ユーザーが「こんにちは」と返事をすると、その音声を認識し、次の処理へ進む。10秒以内に「こんにちは」と認識できない場合には、再び顔検出の処理に戻る。なお、上述の通り、音声合成と音声認識はクラウドサービスにて提供されている。

ユーザーの挨拶を認識すると、Sotaは合成音声でクイズを出題する。クイズはあらかじめ複数用意されており、それらの中からランダムに選択される。クイズはユーザーが「はい」または「いいえ」の二択で答えるものになっている。当初は、「はい」、「いいえ」以外の回答も想定していたが、音声認識の精度が十分ではなかったため、認識対象を2語に限定した。ユーザーの回答を認識できれば正誤の判定を行い、「当たり」、「外れ」という音声を返す。その後、回答の正誤に関わらず問題の解説を行い、再び顔検出プロセスに戻る。

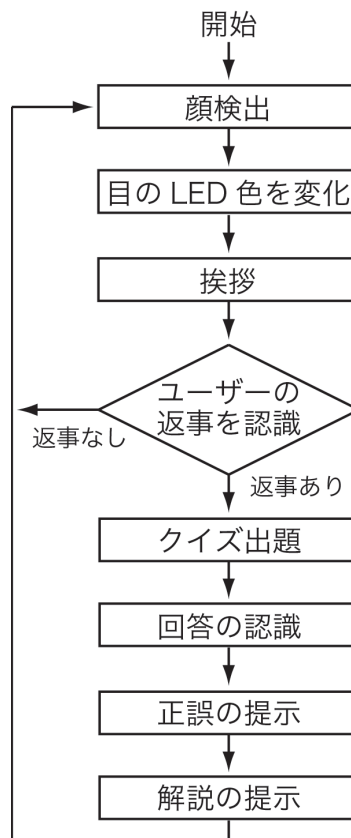


図2：音声対話に基づくクイズシステムのフローチャート

3.2 VstoneMagicによるクイズシステムの実装

VstoneMagic を用いて上述のクイズシステムを実装した。ロボットに求められる機能として、顔認識によるユーザーの検出、合成音声による挨拶や出題、ユーザーの音声の認識、答え合わせ等が挙げられる。VstoneMagic の制御ブロックを用い、これらの機能を実装し、クイズシステムを構築した。

ユーザーの検出には「顔追従ブロック」(図3)を用いた。これは、Sota の頭部を動かしながらユーザーの顔を検出し、その方向へと顔を向けるブロックである。

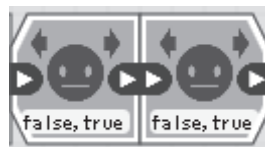


図3：顔追従ブロック

顔の検出後、ユーザーに対して合成音声で「こんにちは」と挨拶をする。これは「発話ブロック」を用いて実現した(図4)。「発話ブロック」は初期設定で「こんにちは」と発話する。



図4：発話ブロック

ユーザーが Sota に挨拶を返したことを確認した後、クイズを出題する。ユーザーの返事の有無の判定には「音声認識 (ベストスコア)」(図5)を用いた。

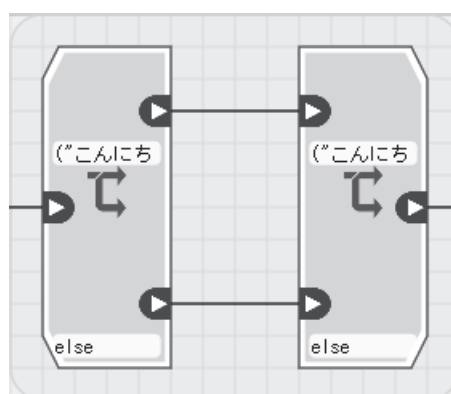


図5：音声認識 (ベストスコア) ブロック

クイズの出題には挨拶と同様に「発話ブロック」(図4)を用いた。あらかじめ複数の問題を用意し、「ランダムブロック」(図6)を用いてランダムに選択した。本システムにおけるクイズは、甲南大学の各学部に関するもので、すべて「はい」または「いいえ」で答えることができるものとした。

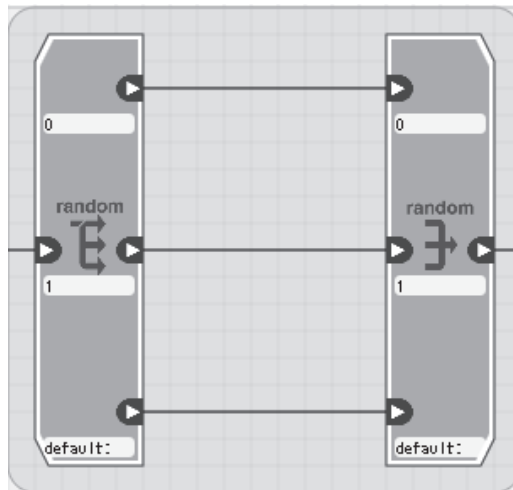


図6：ランダムブロック

VstoneMagicでは「はい」または「いいえ」を認識するための「はいいいえの取得」ブロック(図7)が提供されている。本システムではこれを用いてユーザーの回答を認識し、正解の場合には「当たり」と合成音声で答え、不正解の場合には「外れ」と答える。音声認識が失敗した場合には「残念」と答えるようにした。それに続いてクイズについての解説を提示した後、顔検出の処理に戻る。

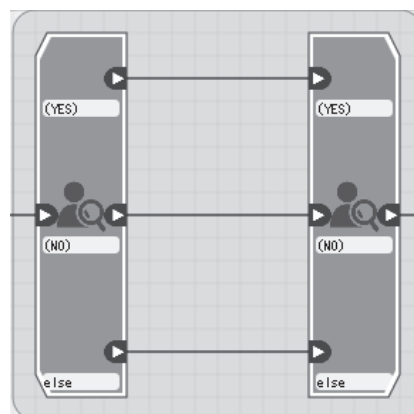


図7：はい/いいえの取得ブロック

図8にシステム全体のブロック構成(プログラム)を示す。この図では、1つの問題にのみ対応したプログラムを示している。複数の問題に対応させるためには、このブロックを複製し、その内容を変更する必要がある。

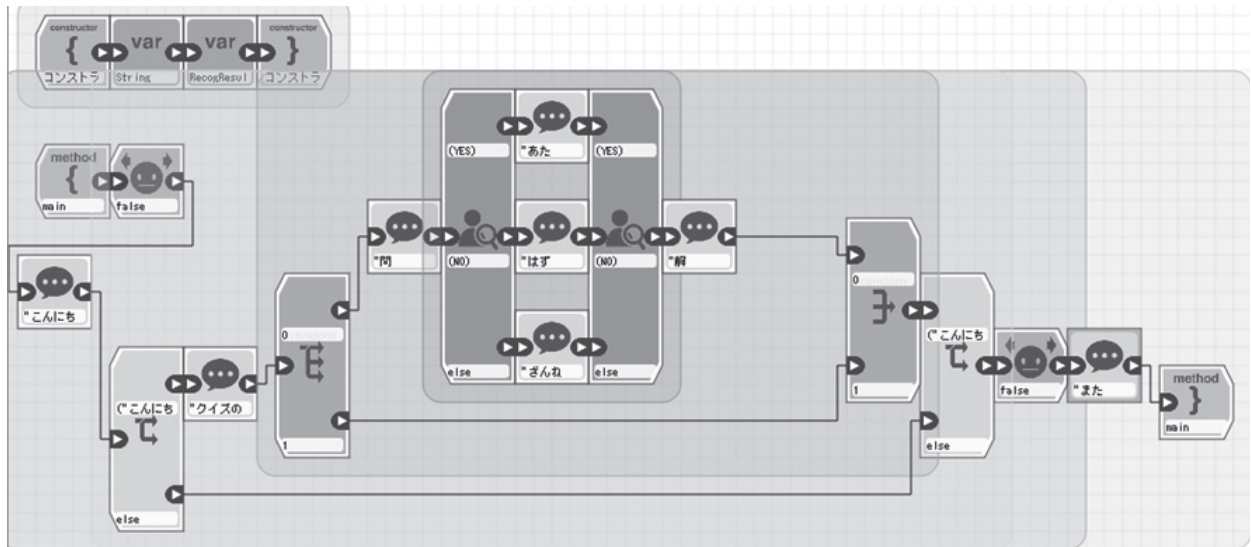


図8 : VstoneMagicで作成したクイズシステムのプログラム

3.3 システムの実演と課題

2016年6月12日に開催された甲南大進学説明会の学部相談会会場にて本システムを実演した。本学への進学を考える高校生やその保護者が参加することを考慮して、クイズは本学の各学部に関するものを用意した。例えば、「甲南大学のマスコットキャラクターになっている動物はイノシシである」、「甲南大学の学部の数は7つである」である。

来場者の反応はおおむね好意的であったが、周囲の雑音が大きくなると音声認識を正確にできないケースがあった。また、合成音声特有の聞き取りにくさにより、クイズの問題が来場者に正しく伝わらないケースもあった。さらに、上述のように、本システムにおいてクイズを追加するためにはクイズの出題、回答認識、正誤判定、解説のブロックを複製し、その内容を変更する必要がある手間がかかる。

そこで、著者らは、これらの問題を解消するシステムを開発することにした。VstoneMagicでは、図8のようなプログラムが、実行直前にJavaのプログラムに変換される。つまり、SotaはJavaのプログラムにより制御される。このとき自動生成されたプログラムに手を加えることによって、あらかじめ用意されたブロックだけでは実現が難しい複雑な機能も実現できる。

4 QRコードを利用したクイズシステム

4.1 システムの改良

このシステムでは以下の3点について改良を施した。第1の改良点は、クイズの変更や追加に関するものである。3.3節のシステムでは、クイズの変更や追加のたびにブロックをコピーしてその内容を変更する必要があった。これは手間がかかる上に、バグも生じやすい。そこで、本シス

テムでは、問題文、答え、解説文から成るCSVファイルを用意し、システム起動時にこれを読み込むようにした。これによって、クイズの変更や追加が格段に容易になった。

第2の改良点は、合成音声の聞き取りにくさに関するものである。3.3節のシステムでは、音声合成時の読み、アクセント、形態素解析の誤りなどにより、問題文が理解しにくい場面があった。そこで、本システムでは、合成音声による提示に加え、クイズの問題文をPCのディスプレイにテキストで表示して、ユーザーがそれを読むことができるように改良した。SotaをTCP/IPのサーバ、PCをクライアントとして接続し、Sota側でクイズをランダムに選択した後、そのテキストをPCに送信する。それをPCのディスプレイに表示させるようにした。

第3の改良点は、音声認識の問題に関するものである。少なくとも2016年6月段階では音声認識の精度に問題があったため、本システムでは音声認識を利用せず、ユーザーがQRコードを用いて回答を示すことにした。QRコードはSotaの頭部にあるカメラにて認識した。QRコードの認識にはZXing (Zebra Crossing) を利用した。

著者らは、このシステムを用いて都道府県に関するクイズを作成した。基本的な処理の流れは図2と同じであるが、上記3点の機能を利用している。47都道府県の名称をQRコードに変換し、ハガキ大のカードに印刷した(図9参照)。Sotaがユーザーの顔を検出し、あいさつを交わした後、「日光東照宮があるのはどこ？(答え：栃木県)」、「一番多くの都道府県と接しているのはどこ？(答え：長野県)」などの問題を合成音声にて提示する。同時にPCのディスプレイに問題文を表示する。ユーザーは47枚のカードの中から回答を探し出し、Sotaに見せるようにする(図9参照)。すると、SotaはQRコードを認識し、正誤判定し、不正解であれば合成音声で正解を示す。

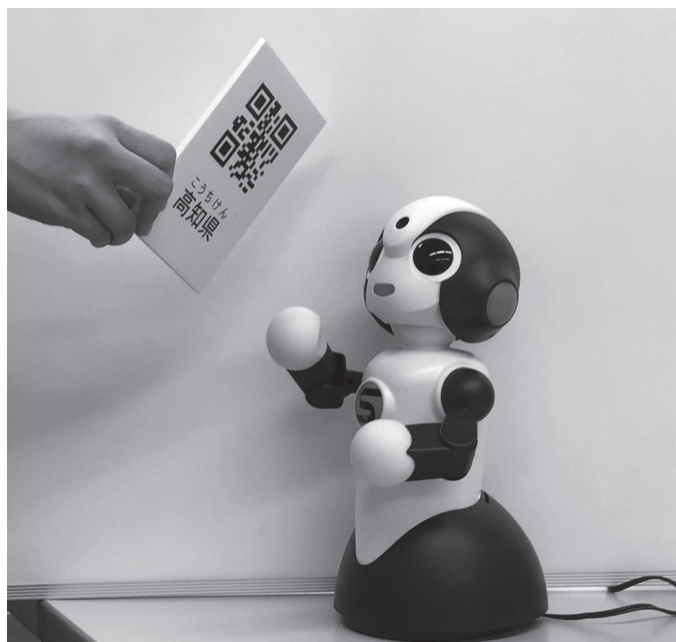


図9：SotaにQRコードを認識させている様子

4.2 システムの実演

2016年8月7日に開催された甲南大学オープンキャンパス・甲南大学知能情報学部「科学とあそぼ! 2016」にてこのシステムの実演を行った。ユーザーが回答用カードをSotaのカメラに近づけすぎると避けるため、テーブル上に回答用カードをかざす位置を示しておいた。本システムでは、回答の入力に音声認識を用いないため、多数の参加者が来場して騒がしい環境でも問題なく運用できた。また、回答のカードを探す部分に競争の要素が入ったので、複数の参加者で競いながら楽しむ場面も見られた。

なお、本システムを1時間程度運用した時点で、熱が原因と考えられる動作不良が発生した。そのため、以降はSotaを扇風機で冷やしながら展示した。

5 おわりに

本研究では、コミュニケーションロボットSotaを用いたクイズシステムを開発した。本システムは、音声認識・合成、画像認識、通信など本学部の教育・研究分野を体現したものだといえる。このシステムを本学の進学説明会やオープンキャンパスなどで展示し、ロボットとコミュニケーションする楽しさを来場者に体験してもらうことができた。本研究ではクイズという形のコミュニケーションを実現したが、今後また別のコミュニケーションの形を模索していきたい。

謝辞

本研究の一部は私立大学等経常費補助金特別補助「大学間連携等による共同研究」および甲南大学KONANプレミア・プロジェクトによる支援を受けたことを付記し深謝する。

参考文献

- [1] 塩見昌裕, 大和信夫, 前田武志, 横山智彰, 深津将生, 今川拓郎, 石黒浩, “テーブルトップ型対話ロボットプラットフォーム「Sota(ソータ): Social Talker」の開発,” 第33回日本ロボット学会学術講演会, RSJ2015AC3G3-06, 2 pages, 2015.
- [2] ヴイストーン株式会社, Sota Community, <https://sota.vstone.co.jp/sota/> (2017年5月1日閲覧).