

# 自律型アンドロイド ERICA による就職面接対話

## Job Interview Dialogue Performed by Autonomous Android ERICA

井上 昂治<sup>1\*</sup> Divesh Lala<sup>1</sup> 原 康平<sup>1</sup> 相原 邦光<sup>2</sup> 中村 静<sup>1</sup> 高梨 克也<sup>1</sup> 河原 達也<sup>1,2</sup>  
Koji Inoue<sup>1</sup> Divesh Lala<sup>1</sup> Kohei Hara<sup>1</sup> Kunimitsu Aihara<sup>2</sup> Shizuka Nakamura<sup>1</sup>  
Katsuya Takanashi<sup>1</sup> Tatsuya Kawahara<sup>1,2</sup>

<sup>1</sup> 京都大学 大学院情報学研究科

<sup>1</sup>Graduate School of Informatics, Kyoto University

<sup>2</sup> 京都大学 工学部情報学科

<sup>2</sup>Undergraduate School of Informatics and Mathematical Science, Kyoto University

**Abstract:** Conventional job interview systems ask only pre-defined questions towards interviewees. In the case of human-human job interviews, interviewers ask follow-up questions after interviewees' responses in order to elicit more information. The job interview system of autonomous android ERICA also generates follow-up questions based on the user responses on the fly. The follow-up questions are based on two types of generation methods. The first method is to select a proper question from a set of pre-defined questions, which is robust against any interviewee's responses. The other method is to use a keyword in an interviewee's response, which is reflective of understanding of the interviewee's response.

## 1 はじめに

スマートスピーカや会話ロボットが普及した今日では、人間とシステムが音声で対話をする光景が珍しくなくなってきている。ただし、ここでのやりとりは、基本的には一問一答型あるいは定型的なやりとりで構成されるものである。そのため、扱われるタスクは、その場で完結するものがほとんどである。一方、人間どうしの対話では、多くのやりとりを要する社会的な対話が多くなされている。例えば、就職の面接では、対面で様々な質問を面接官が投げかけることで、志願者の能力や志望動機などを判断している。そのため、志願者にとって、様々な質問に対する回答を予め練習しておくことが重要といえる。しかし、面接官役の相手を確認することは、人的および時間的なコストから難しい。そこで、面接の練習相手を音声対話システムで実現することが考えられるが、本番同様のリアルな面接を実現しなければその効果は期待できない。

我々はこれまでに自律型アンドロイド ERICA による音声対話の研究を進めてきた [1, 2]。アンドロイド ERICA の見かけは人間に酷似しており、非言語も含めた多様なふるまいを生成することができる。本システムでは、就職の面接について、その練習相手を ERICA で実現することを目指す。面接の自動評価に関しては、これまでに多くの研究がなされている [3, 4, 5]。一方、

面接対話自体をどのように実現するのかについては、予め用意した質問を投げかけるのみに留まっている。これでは志願者は事前に用意した回答を述べるだけで対応可能である。そこで、本システムでは、志願者の回答に応じて掘り下げ質問を動的に生成する。掘り下げ質問の生成に関する研究はまだ少数である [6]。ここでは、志願者の回答の充実度やキーワードに基づいた掘り下げ質問の自動生成を試みる。これにより、より多くの情報を志願者から自然な形で引き出すことを狙う。

## 2 想定する就職面接対話

本システムでは、ERICA が就職面接の練習相手として、面接官役を行う。ここで扱う就職面接は、特定の業種および企業に依存しないようにしている。具体的には、面接官からなされる質問は志願者自身に関する事柄に焦点をあてている。また、志願者の回答に対する掘り下げ質問も特定の業種および企業に依存しない。したがって、就職採用の一次面接に近い内容を想定している。このようにすることで、どのような条件の志願者に対しても、面接官が行う質問のリストを変更することなく、練習相手として対応することができる。ただし、志願者は、自身が志望する業種および企業を想定して練習に臨む必要がある。

\*連絡先: k.inoue@i.kyoto-u.ac.jp

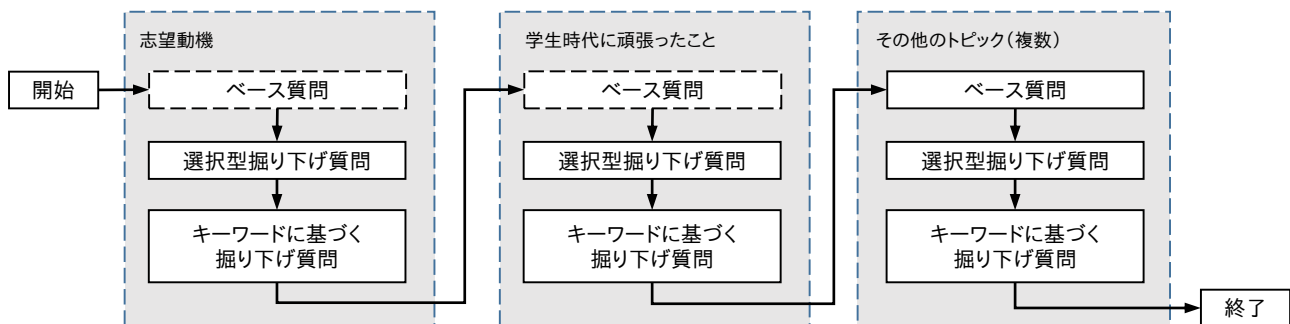


図 1: 面接対話の流れ

### 3 面接対話システム

上記の想定に基づいた面接対話システムの実現方法について述べる。対話制御は状態制御に基づくが、志願者の回答に応じて、掘り下げ質問を生成することでリアル面接の実現を目指す。

#### 3.1 対話の流れ

面接対話の流れを図1に示す。この対話は状態遷移に基づいて制御される。この面接対話は、複数のトピックで構成される。トピックは、就職面接において一般的に行われる典型的な内容を指す。例えば、「志望動機」や「学生時代に頑張ったこと」などがこれに相当する。各トピックは、ベース質問とこれに続く掘り下げ質問から成る。ベース質問とは、そのトピックに関してオープンに尋ねるものである。例えば、「志望動機」のトピックの場合は、「あなたの志望動機について聞かせてください」などとなる。トピックおよび対応するベース質問は予め用意しておく。次に、ベース質問に対する志願者の回答に基づいて、選択型掘り下げ質問がなされる。その後、選択型掘り下げ質問への志願者の回答に基づいて、キーワードに基づく掘り下げ質問がなされる。この掘り下げ質問に対する回答が終わると、次のトピックへと遷移する。各掘り下げ質問の生成方法や具体例については次節で述べる。

志願者が回答を述べる際の、試験官であるERICAのふるまいについて述べる。まず、志願者が回答を述べているときには、回答に合わせてERICAはうなずきを表出する。うなずきのタイミングは、相槌予測のモデル[7]を用いて、韻律情報をもとに決定する。発話権の制御(ターンテイキング)は、志願者が回答を確実に言い終えるために、4秒程度の沈黙を検出した場合に回答が終了したとみなす。ただし、この間の沈黙が不自然にならないように、沈黙が続く間はERICAの

うなずきを表出する。また、志願者の回答の内容語(名詞)数が少ない場合には、「えっ」や「もう一度お願いします」などをERICAが言うことで、志願者が十分な回答を述べるまで待つ。

#### 3.2 掘り下げ質問生成

掘り下げ質問の生成方法について述べる。はじめに、ベース質問への回答に対して、選択型の掘り下げ質問を生成する。次に、この掘り下げ質問への回答について、キーワードに基づく掘り下げ質問を生成する。

##### 3.2.1 選択型掘り下げ質問

ベース質問に関連する掘り下げ質問を予め複数用意しておき、志願者の回答に応じて選択する。ここでは、志願者の回答の充実度に応じた以下の3つのカテゴリのそれぞれについて、掘り下げ質問を複数用意する。

- **カテゴリ 1 (優先度高)**  
必ず答えて欲しかったことに関する掘り下げ質問
- **カテゴリ 2 (優先度中)**  
答えるとよかったことに関する掘り下げ質問
- **カテゴリ 3 (優先度低)**  
その他に関する掘り下げ質問

カテゴリ3は、カテゴリ1と2のどの掘り下げ質問も適さない場合のバックアップである。例えば、「志望動機」に関するベース質問の場合、各カテゴリに対応する掘り下げ質問の例は以下となる。

- **カテゴリ 1**  
当社でないといけない理由というのがあるのでしょか

- カテゴリ 2  
実際に入ってからやりたい仕事や実現したいことなどはありますか
- カテゴリ 3  
当社の他にこの業界のどのような企業を受けていますか

カテゴリ 1 の例は、その会社を選んだ理由が述べられていなければ、この掘り下げ質問を行うことは妥当と考えられる。また、カテゴリ 2 の例は、その会社を選んだ理由は述べられているが、その後具体的にどうしたいのかについて述べられていなければ、妥当な掘り下げ質問といえる。

志願者の回答の内容に応じて、上記のように用意された複数の掘り下げ質問から 1 つを選択する。まず、名詞の数が一定数以上あれば、カテゴリ 1 に属する掘り下げ質問を候補から除外する。カテゴリ 1 の掘り下げ質問の内容は、ある程度の発話がなされている場合、すでに述べられている可能性が高いためである。次に、すべてのカテゴリの各掘り下げ質問の内容について、ベース質問に対する回答ですでに述べられているか否かを推定する。はじめに、各掘り下げ質問を表す単語集合を定義する。ここでは、掘り下げ質問の文中に出現する名詞およびそれに関連する名詞を列挙しておく。例えば、「実際に入ってからやりたい仕事や実現したいことなどはありますか」という掘り下げ質問の場合、これを表す単語集合は「仕事」、「実現」、「入社」などとなる。そして、ベース質問に対する回答中の名詞とこの単語集合中の名詞との意味的な距離が 1 つの組合せでも近い場合には、その掘り下げ質問を候補から除外する。前述の例の場合、ベース質問に対する回答で志願者が「入社」あるいは、これに近い「就職」という単語を述べていた場合には、この掘り下げ質問文は候補から除外される。意味的な近さを測るために、単語のベクトル表現 (word2vec) を用いて、単語間のコサイン距離を閾値判定する。以上の処理を経て、候補に残った掘り下げ質問のうち、できるだけ高い優先度のカテゴリから 1 つを選択する。同じカテゴリ内に候補が複数存在する場合には、それらからランダムに選択する。

### 3.2.2 キーワードに基づく掘り下げ質問

選択型の掘り下げ質問は確実に生成できるが、掘り下げ質問の内容自体は動的に生成しているわけではないため、志願者にとって自分の回答が「聞かれている」という感覚が得られにくい可能性がある。そこで、志願者の回答の内容に基づき、掘り下げ質問文の内容を動的に生成することも行う。質問文全体を生成するのは難しいため、ここでは志願者の回答からキーワード



図 2: アンドロイド ERICA による面接対話の様子

を抽出し、これを用いたテンプレートにより掘り下げ質問を生成する。はじめに、志願者の回答中のすべての名詞に対して、その単語の珍しさを表す idf 値とその名詞のモーラ数について、それぞれでランキングを行う。idf 値については、日本語 Wikipedia のテキストデータから算出した。この方針は、珍しくかつ長い単語がキーワードになりやすいという仮説に基づく。ただし、名詞が連続している場合は複合名詞とみなし、1 単語として扱う。さらに、「人工知能の研究」といった「(名詞 1) の (名詞 2)」などの場合も、これら全体をキーワードの候補として加え、複合名詞と同様に扱う。これらのように複数の名詞がキーワードに含まれる場合には、各単語の idf 値の中で最大のものを、そのキーワード候補の idf 値とする。各キーワード候補について、idf 値とモーラ数の 2 種類のランキングの順位を足し合わせ、その数値が最小のものを最終的なキーワードとする。その後、予め用意したスロットにキーワードをあてはめることで、掘り下げ質問を生成する。例えば、キーワードが「音声対話システム」の場合、「今おっしゃった音声対話システムについて、もう少し詳しく説明してください」という掘り下げ質問が生成される。

## 4 対話例

上記の面接対話システムによる対話例を以下に示す<sup>1</sup>。以下では、U はユーザ、E は ERICA の発話をそれぞれ表す。また、ERICA が面接対話を行う様子を図 2 に示す。

はじめに、ベース質問が志望動機の場合について示す。

**E1** それでは、あなたの志望動機について聞かせてください。(ベース質問)

<sup>1</sup>対話の様子を動画としても公開している。  
(日本語) <https://youtu.be/JFc90m9TJ6I>  
(英語) <https://youtu.be/JpT1SVp2zx8>

**U1** はい。私はこれまで人工知能に関する技術について大学で学んできました。また、御社では人工知能技術の開発に力を入れておられます。そのため、御社では、私のスキルを活かしつつ、また新しい技術を学ぶことができると思い志望いたしました。

**E2** そうなんです。わかりました。では、あなたは自分のやりたいことを実現するためにどんな努力や準備をされていますか。  
(選択型掘り下げ質問)

**U2** はい。この業界は変化が激しいので、常に新しい技術を学んでいます。現在ですと、機械学習の研究に力を入れています。その中でも特に音声対話システムについて学んでいます。

**E3** そうなんです。今おっしゃった音声対話システムについて、もう少し詳しく説明してもらえますか。  
(キーワードに基づく掘り下げ質問)

**U3** はい。音声対話システムは、音声で人間と会話をするシステムのことです。音声認識、言語処理、対話制御など様々な技術が必要となります。そのため、多くのことを学ぶ必要があります。最終的には人間のように自然な会話ができるシステムの実現を目指しています。

**E4** わかりました。ありがとうございます。

次に、別のベース質問として、学生時代に一番頑張ったことについての例を示す。

**E1** では、学生時代に一番頑張ったことを教えてください。(ベース質問)

**U1** 学生時代に一番頑張ったことは部活動における仕事です。具体的には、道具の管理を行う役職のリーダーをしていました。全てを管理するのは大変でしたが、このような管理能力や経験は御社に入っても役に立てると思います。

**E2** ありがとうございます。その取り組みにおいて一番難しかった課題や問題点は何か。  
(選択型掘り下げ質問)

**U2** 練習場所の立替え工事があり、その際の引っ越し作業が難しかったです。リーダーとしてきちんとしたスケジュール管理を行い、他の人に指示を出すことで、効率よく作業を進めることができました。

**E3** そうなんです。今おっしゃったスケジュール管理について、より具体的に教えてください。  
(キーワードに基づく掘り下げ質問)

**U3** 具体的には、部員の授業や練習日程を細かく管理し、その中で空いた時間に仕事を割り当てるようにしていました。そうすることで、時間を無駄に

することなく、時間内に仕事を終わることができました。

**E4** わかりました。ありがとうございます。

## 5 まとめ

本稿では、自律型アンドロイド ERICA による面接対話のための音声対話システムについて述べた。このシステムは、予め用意した質問を投げかけるだけでなく、志願者の回答に応じて掘り下げ質問を生成する。掘り下げ質問は、選択型の掘り下げ質問とキーワードに基づく掘り下げ質問の2種類である。選択型は予め用意されたものであるが、キーワードに基づくものはユーザの回答内容を動的に反映することができる。これらの掘り下げ質問により、リアルな面接を実現する。

## 謝辞

本研究は、JST ERATO 石黒共生ヒューマンロボットインタラクションプロジェクト JPMJER1401 の支援を受けて実施した。

## 参考文献

- [1] Koji Inoue, Pierrick Milhorat, Divesh Lala, Tianyu Zhao, and Tatsuya Kawahara. Talking with ERICA, an autonomous android. In *SIGDIAL*, pp. 212–215, 2016.
- [2] Tatsuya Kawahara. Spoken dialogue system for a human-like conversational robot ERICA. In *IWSDS*, 2018.
- [3] Mohammed E. Hoque, Matthieu Courgeon, Jean-Claude Martin, Bilge Mutlu, and Rosalind W. Picard. MACH: My automated conversation coach. In *UbiComp*, pp. 697–706, 2013.
- [4] Mohammad R. Ali, Dev Crasta, Li Jin, Agustin Baretto, Joshua Pachter, Ronald D. Rogge, and Mohammed E. Hoque. LISSA-Live interactive social skill assistance. In *ACII*, pp. 173–179, 2015.
- [5] Pooja Rao S. B., Sowmya Rasipuram, Rahul Das, and Dinesh B. Jayagopi. Automatic assessment of communication skill in non-conventional interview settings: A comparative study. In *ICMI*, pp. 221–229, 2017.
- [6] Ming-Hsiang Su, Chung-Hsien Wu, Kun-Yi Huang, Qian-Bei Hong, and Huai-Hung Huang. Follow-up question generation using pattern-based seq2seq with a small corpus for interview coaching. *INTERSPEECH*, pp. 1006–1010, 2018.
- [7] Tatsuya Kawahara, Takashi Yamaguchi, Koji Inoue, Katsuya Takanashi, and Nigel G Ward. Prediction and generation of backchannel form for attentive listening systems. In *INTERSPEECH*, pp. 2890–2894, 2016.