

## Radiobots 型対話システムの提案

## Proposal of Radiobots based Spoken Dialogue System

木村泰知<sup>\*1</sup> ジェプカ ラファウ<sup>\*2</sup> 高丸圭一<sup>\*3</sup>  
Yasutomo Kimura Rafal Rzepka Keiichi Takamaru

<sup>\*1</sup>小樽商科大学 <sup>\*2</sup>北海道大学 <sup>\*3</sup>宇都宮共和大学  
Otaru University of Commerce Hokkaido University Utsunomiya Kyowa University

In our opinion "radio personalities and listeners interaction" perfectly refers to the relationship of "dialog system and its users". We introduce a new concept of a radio where its final goal is to become a dialog-supported interaction, enlist problems of existing systems and describe a prototype platform we have built to realize our proposed system.

## 1. はじめに

近年、雑談システム、あるいは、非タスク指向型対話処理と呼ばれる研究が盛んであり、人間らしい応答ができるシステムが登場しつつある [杉山 2015][東中 2014]. A.L.I.C.E.[Wallace 2009] は人間らしい応答ができる雑談システムとして有名であり、チューリングテストにより評価されるロープナー賞を受賞している。また、シンプルな対話規則を用いて実装された LINE アプリのパン田一郎<sup>\*1</sup> は 2015 年 3 月時点で 1200 万人のユーザによって登録され、2 億 6000 万回以上の入力がある。このほか、NTT ドコモの雑談対話 API<sup>\*2</sup> 等によって、雑談システムの構築が容易に行えるようになってきた。

しかしながら、これらの対話システムは基本的にユーザの 1 発話に対して、システムが 1 応答する仕組みである。人間同士の雑談でも特に話題がない場合に会話が進まないのと同様に、対話システムを起動した直後には、ユーザは発話を求められても何を話せばよいかかわからず、挨拶などしかできないことがある。対話システムからユーザに発話を求める場合、ある意味「強制的に」発話をさせられることから、「自ら発話したいという思い」が持ちにくく、自然な対話が成立しにくいという問題がある。このため、人間と見間違えほどの応答ができたとしても、ユーザがシステムを継続して利用してくれるとは限らない。また、対話システムの研究には、対話の分析や評価用の発話データが必要であり、大量のデータを準備する必要がある。しかしながら、長期的なスパンで、自然に発話されたデータを収集するなどの開発コストが問題となる。

このような問題を解決するために、本稿では、対話システムを継続して利用することを考慮し、ユーザの入力コストを抑えつつ、対話らしい娯楽性を保つ新たな対話形式を提案する。我々は「ラジオ番組のパーソナリティとリスナーのリアルタイムな反応」の関係性が「対話システムとユーザ」の最適な関係であると考え、対話システムの最終ゴールを示すとともに、対話システムのプラットフォーム構築に向けたプロトタイプを紹介する。ラジオ番組的なシステム発話とユーザからのリアルタイムな応答を模擬的に実現する環境として、対話システムのプラットフォームにニコニコ動画の生放送（ニコニコ生放送）を利用する。ニコニコ生放送は音声や動画のストリーミング配信に対する、リアルタイムの応答（コメント入力）の機能を備え

ている。また、コメント数や視聴者数の表示機能、ユーザへのアンケート機能などシステム評価に利用可能な機能を有している。さらに、知名度、放送枠の単位（1 放送 30 分間）などの点から本研究のプラットフォームとして適していると考えられる。本稿では、ニコニコ生放送を利用した Radiobots 型対話システムのプロトタイプについて説明する。

## 2. 提案

## 2.1 ねらい

我々は「ラジオ番組のパーソナリティとリスナーのリアルタイムな反応」の関係性が「対話システムとユーザ」の最適な関係と考えて、Radiobots 型対話システムの提案をする。本研究のねらいは、リアルタイムのラジオ番組を自動的に配信できる環境を整え、従来の対話システムの問題の解決を試みることである。

まず、従来の対話システムの問題点を述べる。

1. 入力コストに対するコスト対効果が低い（1 問 1 答の負担）
2. 1 システムを 1 ユーザだけが利用する
3. 発話しないと対話が展開しないため、不自然な発話になる

従来の対話システムにおける上記の問題をラジオ番組のパーソナリティとリスナーの関係を参考にして以下のように解決する。

1. 強制的な入力（自然文やキーワード）のコストを削減する
2. 複数キャラクターに対して複数ユーザが参加可能とする
3. 喋りたい気持ちになったときに発話をする

## 2.2 ラジオ番組と人工知能研究

ラジオパーソナリティの振る舞いを模した対話システムに必要な人工知能の研究タスク（言語処理技術や音声処理技術など）について、図 1 に示すラジオ番組における幾つかのコンテンツを例に述べる。まず、番組を構成する各コーナーにおいて必要となる研究タスクについて説明する。

ラジオの冒頭における複数パーソナリティの自己紹介では、自己紹介のための「文生成」やキャラクターごとの「役割語」やパーソナリティ間の会話における「あいづち」が必要となる [金水 2011][吉田 2009]。リクエスト募集・選択では、リスナーの投稿からの要望や意図を理解するための「要望抽出」「発話理

連絡先: 木村泰知, kimura@res.otaru-uc.ac.jp

\*1 <http://line.froma.com/>

\*2 <https://dev.smt.docomo.ne.jp/?p=docs.api.page>

番組内容 (コーナーの例)	概要	特に必要となる技術
ジングル・挨拶	複数エージェントの自己紹介	文生成・役割語・あいづち
コーナー告知 リクエスト募集	ユーザとの交流方法の告知	要望抽出・発話理解・意図理解
ニュース	2キャラクター以上による議論	情報抽出・自動要約
トリビア	情報収集を活かした時事ネタ	知識獲得・対話データの分析
ゲスト	実際の人間とのコラボ	常識 ジョーク・感情処理
CM 広告	ユーザの要望へ適合	マッチング
音楽リクエスト	利用者の要望	歌声音声
〇〇コーナー	利用者との対応	対話管理・自動要約
本紹介・映画紹介	評判・推薦の応用	意見抽出・推薦
天気予報 (時間調整)	時間を調整する対話制御	対話管理・発話制御

全体に必要な技術

音声合成  
対話管理  
発話制御  
発話理解  
意図理解  
マルチモーダル応答生成  
発話プランニング  
コンテンツ構築支援  
発話プランニング

図 1: 人工知能ラジオ番組に必要な研究分野の例

解・意図理解が必要となる [葦原 2012]. ニュースは, 2 キャラクター以上による議論や補足説明などがあるため, 前述のものに加えて, 重要箇所を抽出するための「情報抽出・自動要約」が必要である [平尾 2003]. トリビアコーナーは, ニュース記事・Wikipedia・対話履歴から驚くような事実を抽出するための「知識獲得」「対話データの分析」が必要となる. ゲストは, 2 キャラクターに加えて, 実際の間人が絡み, インタラクティブな対応をすることであり, 「常識」「ジョーク」「感情処理」「対話管理」が必要となる [徳久 2008]. CM 広告は, リスナーに適した広告を流すことが可能になれば, マッチング広告が必要となる. 音楽リクエストは, 著作権の問題もあることから, 歌声音声技術を利用した音楽の提供が考えられる. 任意のコーナーは, リスナーからのコメントなどの投稿を要約することが含まれるため「自動要約」が必要となる [浅原 2015]. 本・映画の紹介は, 本や映画の推薦文や意見に基づき, お薦めを紹介することを考えおり「意見抽出・推薦」が必要となる [中山 2015]. 天気予報 (時間調整) は, 30 分の番組を天気情報で調整することを考えており「自動要約」や「対話管理」が必要となる.

ラジオは複数のチャンネルで番組が放送されており, それぞれ異なった情報が提供される. ユーザはこの中から好みのチャンネルを選択し情報を受け取ることができる. ユーザが行うべき行動はチャンネルの選択のみである. 特定の具体的な情報を入手したいという情報検索的な用途には向かないものの, 漠然とした情報 (知識) を幅広く得たいという目的においては, 情報検索と比べて, ユーザの入力コストは極めて低く済む. また, ラジオの特徴として「複数人のリスナー (ユーザ) が番組を同時に視聴し, 場を共有している」「リスナー (ユーザ) は必ずしも発言しなくてよい」「パーソナリティはすべてのリスナー (ユーザ) のリクエスト (発言) に応答しなくてもよい」などが挙げられる. 他にも, ラジオではリスナー (ユーザ) を楽しませるために, 発話だけでなく, 効果音, ジングル, バックグラウンドの音を利用している.

我々は, ユーザからの応答をリアルタイムに受け付ける人工知能によるラジオを「Radiobots 型対話システム」と呼ぶ.

### 2.3 Radiobots 型対話システム

我々は, 既存のラジオ番組の構成を参考に, Radiobots 型対話システムを提案する. 図 1 において対話プラットフォームの

イメージを伝えるとともに, 必要となる技術の例を示した.

本節では, Radiobots 型対話システム構成及びシナリオの作成及び更新方法について説明する. 図 3 は, ニコニコ動画の生放送をする準備時間と放送時間の時間の流れ, および, データの流れを示す. 本研究では, Windows 上で「Nikoniko Live Encoder<sup>\*3</sup>」を用いて放送しており, シナリオの生成, コメントの自動取得などの処理については Linux 上で行い, ファイル共有のソフトウェアを用いてシナリオを受け渡ししている.

下記に Radiobots 型対話システムの処理概要を示す.

1. シナリオの作成  
放送前にウェブ上の情報に基づいてシナリオ作成を行う. シナリオはトピックあるいはコーナーで話される最小単位のシナリオを複数もつことで管理する. この最小単位のシナリオを「対話セット」と呼ぶ. 複数エージェントの対話制御は, シナリオを用いて行っており, キャラクターごとの発言が書き込まれている.
2. ラジオ放送の開始 (ニコニコ生放送+放送 ID の取得)  
ニコニコ生方法は予め作成したシナリオに基づいて放送する. 開始時点でコメントの取得のために放送 ID を取得する.
3. コメントの取得とシナリオの更新  
ラジオの方法中の更新は, シナリオを更新することで行う. コメントを反映した「対話セット」を挿入することでコメントを反映した対話を実現する.

#### 2.3.1 シナリオソース

シナリオは, ニュース・本・映画などの情報や Wikipedia のような百科事典を用いて作成する. 例えば, シナリオの情報源となるニュースなどは, 「News-日付-時間」フォルダにトピックごとに収集し, 「収集時間」「タイトル」「本文」を含めてファイルに保存する.

#### 2.3.2 対話規則

現時点では, 対話規則は ELIZA の規則を参考に 3 つの規則を用いている.

\*3 <http://live.nicovideo.jp/s/encoder>

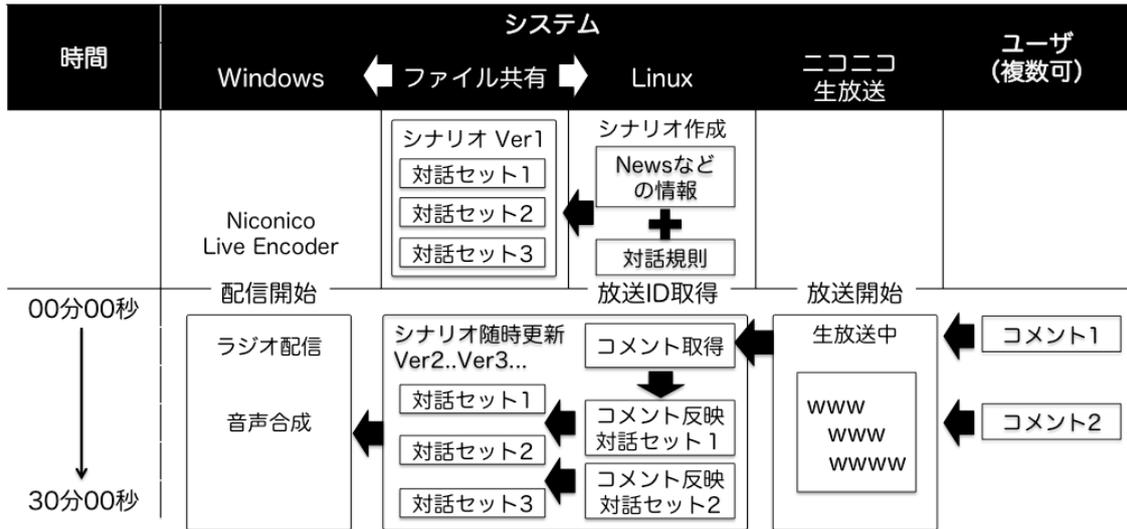


図 2: Radiobots 型対話システムの構成とデータの流れ

1. いつでも利用できる応答表現  
例. 「そうだね」「おもしろい」「いいね」など
2. 置換規則によるオウム返し  
例. 変更前:「あなた」→変更後:「わたし」
3. キーワードの応答  
例. 「天気」→「晴れるといいね」

他にも「Wikipedia を用いた解説」を行なっており、ニュース記事に出現した最も長い単語を Wikipedia で調べ、1 行目の定義文を利用する。

### 2.3.3 対話セット

対話セットは、新聞記事などのシナリオソースから対話規則を用いて生成する。例として、テレビ取材のニュースから作成した対話セットを表 1 に示す。

表 1: 対話セットの例

開始時間	発話時間	名前	内容
0:00	4 秒	オトハ	取材をビデオ撮影するのやめてについて話しますよ
0:04	1 秒	タエ	えええー
0:05	5 秒	オトハ	テレビ取材の撮影やめるよう市長に求める決議案…

### 2.3.4 シナリオの更新

シナリオの更新は、既存の対話セットを動的に組み合わせることで実現する。シナリオの更新は、ユーザからのコメントが入力されると、既存の対話セット間にコメントに回答するための対話セットを挿入する。

## 3. プロトタイプ

Radiobots 型対話システムは、ニコニコ動画の生放送を用いて実装している。ニコニコ生放送は、生放送を行うための敷

居が低いことに加えて、音声合成の利用が進んでおり、ユーザからのコメント収集も容易である。また、ニコニコ生放送は、放送予約も可能であることから、全ての処理を自動化することも可能である。

プロトタイプでは、シナリオを生成し、ユーザのコメントやニュースなどの状況を考慮して、シナリオを更新する処理を作成した。2 キャラクターの声は、音声合成ソフトを利用しており、VOICEROID の結月ゆかり<sup>\*4</sup> と棒読みちちゃん<sup>\*5</sup> を使い分けることにより、異なるキャラクターの音声として出力している。

### 3.1 研究課題

現時点の課題について述べる。

- 発言の統一化 (意見・立場・役割語など)
- キャラクターの設定
- 娯楽性の向上 (ジョーク, 感情処理, 俳句の自動生成など)
- 話しかけやすさ

Radiobots 型対話システムには、発言・意見・役割語などの統一が必要である。発言の統一については、過去の発言を何度も繰り返さず、参照するときには過去に話したことを明確にすることである。例えば、過去の内容は「〇月〇日に話したように..」などの対応が考えられる。意見については、過去の意見と異なるように統一することである。例えば、一般のラジオでは、数多くのニュース情報を取り上げることから、過去のニュースの意見との整合性も考慮する必要がある。役割語の統一については、文末表現などをキャラクターに合わせることである [宮崎 2015]。例えば、女性のキャラクターに適した文末表現として「ですね」「わね」などに統一する必要がある。

また、キャラクターの設定を明確にする必要があり、上記の役割語にも関連している。プロトタイプでは、複数のキャラクターがシナリオを発言しているだけであったため、今後

\*4 <http://www.ah-soft.com/voiceroid/yukari/>

\*5 <http://chi.usamimi.info/Program/Application/BouyomiChan/>

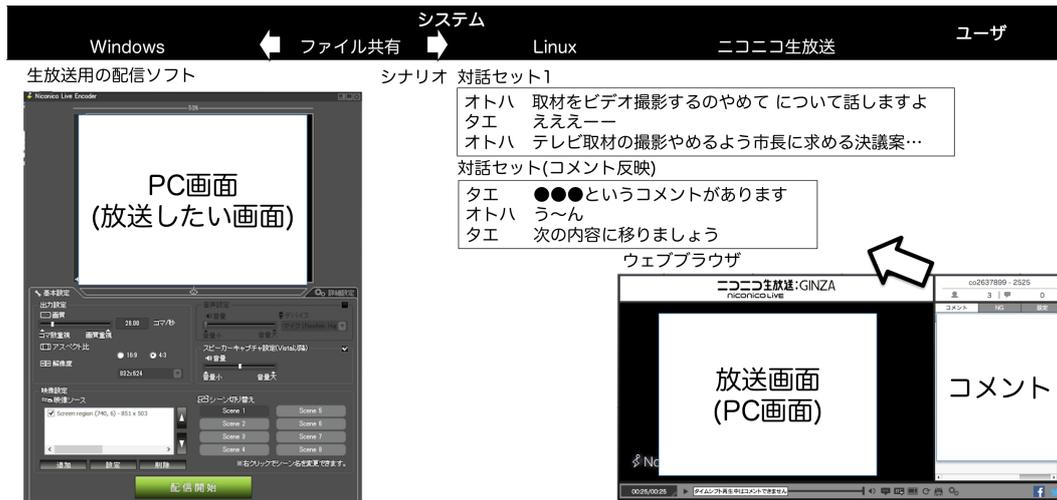


図 3: プロトタイプ例

は、マルチエージェントとして MMD agent の利用も検討する [李 2011]. 例えば、エージェントが多い研究のひとつに「人狼プロジェクト」がある [片上 2015]. 人狼プロジェクトでは、MMD Agent を用いて、5 人の初音ミクがゲームに参加しており、複数のエージェントを利用した研究である。

娯楽性を向上するためには、ジョーク、感情処理、俳句の自動生成などの研究を考えている。また、「話しかけやすさ」を考慮することも必要であり、従来から研究されているが、[杉山 2014] らの研究はロボットとユーザーを想定しており、本研究に適した評価を検討する必要がある。

他には、ラジオの特徴を活かすことも考えられ、エコーや発話速度・間・なども効果を利用することも考えられる。

#### 4. おわりに

本稿では「ラジオ番組のパーソナリティとリスナーのリアルタイムな反応」の関係性が「対話システムとユーザー」の最適な関係であると考え、Radiobots 型対話システムを提案した。今後の課題については、発言・意見・役割語の統一化やキャラクターの設定などがある。

#### 参考文献

[東中 2014] 東中竜一郎, "チューリングテスト「合格」のシステム", 情報処理, Vol.55, No.9, pp.904-907, 2014.

[Wallace 2009] Richard S. Wallace, "The Anatomy of A.L.I.C.E.", Parsing the Turing Test, pp.181-210, 2009.

[片上 2015] 片上大輔, 鳥海不二夫, 大澤博隆, 稲葉通将, 篠田孝祐, 松原仁, "人狼知能プロジェクト", 人工知能学会誌, Vol. 30, No. 1, pp. 65-73, 2015.

[杉山 2015] 杉山 弘晃, 目黒 豊美, 東中竜一郎, 南泰浩, "任意の話題を持つユーザー発話に対する係り受けと用例を利用した応答文の生成", 人工知能学会論文誌, Vol.30, No.1, pp.183-194, 2015.

[宮崎 2015] 宮崎千明, 平野徹, 東中竜一郎, 牧野俊朗, 松尾義博, 佐藤理史, "文節機能部の確率的書き換えによるキャラクター性変換", 言語処理学会第 21 回年次大会, B1-4, pp. 277-280, 2015.

[徳久 2008] 徳久良子, 乾健太郎, 松本裕治, "Web から獲得した感情生起要因コーパスに基づく感情推定", 情報処理学会論文誌, pp.33-36, 2008.

[杉山 2014] 杉山貴昭, 駒谷和範, 佐藤理史, "ロボットへの話しかけやすさモデルの評価と個人差や教示による変動への対応", 人工知能学会論文誌, Vol.29, No.1 pp.32-40, 2014.

[吉田 2009] 吉田奈央, 高梨克也, 伝康晴, "対話におけるあいづち表現の認定とその問題点について", C3-4, p430-433, 2009.

[金水 2011] 金水敏, 役割語研究の展開, くろしお出版, 2011.

[中山 2015] 中山祐輝, 藤井敦, "意見文の対象読者を限定する条件の抽出", 言語処理学会第 21 回年次大会, pp.437-440, 2015.

[葦原 2012] 葦原史敏, 木村泰知, 荒木健治, "地方議会議録における要求・要望表現抽出の提案", 言語処理学会第 18 回年次大会論文集, P1-27, 2012.

[平尾 2003] 平尾努, 磯崎秀樹, 前田英作, 松本, 裕治, "Support Vector Machine を用いた重要文抽出法 (自然言語)", 情報処理学会論文誌, Vol. 44, No. 8, pp.2230-2243, 2003.

[浅原 2015] 浅原 正幸, 加藤 祥, 今田 水穂, "単一文書自動要約のための言語資源構築に向けて", 研究報告自然言語処理 (NL), 2015-NL-220(15), pp.1-27, 2015. (2015-01-12)

[李 2011] 李晃伸, 大浦圭一郎, 徳田恵一, "魅力ある音声インタラクションシステムを構築するためのオープンソースツールキット MMDAgent", 電子情報通信学会技術研究報告, pp.159-164, 2011.