

音声対話コンテンツ循環環境を実現するための対話シナリオのプロファイル自動収集およびフィードバック

Autonomous Collection and Feedback of Dialogue Profile for Dialogue Contents Circulation

森 雅希 李 晃伸
Masaki Mori Akinobu Lee

名古屋工業大学大学院工学研究科
Nagoya Institute of Technology, Japan

In this research, we regard a spoken dialogue systems as “dialogue media content” from user’s view, and tries to realize dialogue media content circulation between creators and users. Focusing on the feedback from users to the creator of the content, an autonomous collection of dialogue profile and its effective feedback scheme is proposed. Experimental results showed that among the automatically collected feedback information, displaying recognition word list and occupation time at each state helps creators for content refinement and updates. Further interview with a professional creator after a long free-creation run also indicates that this scheme can motivate ones to create the content.

1. はじめに

近年、音声認識や音声合成、自然言語処理技術の発展に伴い音声対話システムが注目されており、様々な音声対話システムや音声言語インタフェースが実用化されつつある。

個々に異なる音声対話システムが社会に溢れるとき、ユーザにとって音声対話システムは対話という体験を提供する一種のコンテンツとして捉えられと考えられる。映像メディアや画像メディアのように、音声対話システムもその構成要素（辞書・ボイス・対話シナリオ・エージェントモデル・モーション等）からなるコンテンツとして考えることができる。音声対話がコンテンツとして認知され、対話が消費・模倣・創造される環境が実現されれば、多様な対話システムの創造や、データドリブンで実践的な対話システムの知見の蓄積が期待される。筆者らは音声インタラクション構築ツールキット MMDAgent を用いた音声対話コンテンツの研究を行っている [1][2]。

本研究では、音声対話コンテンツがメディアとして循環するための要件の一つとして、ユーザからシステム構築者（クリエイター）へのフィードバックに着目し、利用したユーザの体験や動向をクリエイターへ積極的にかつ有効にフィードバックする仕組みを提案する。以下、対話シナリオ動作時のユーザごとのプロファイルを自動収集し、それをクリエイターに有益な形でフィードバックするための仕組みを提案し、実験を行った結果を示す。

2. 音声対話コンテンツの循環系

コンテンツが流通する背景にはコンテンツ循環が存在する [3]。音声対話においては、図 1 のようなコンテンツ、ユーザ、クリエイターが主体となった循環系を成すと考えられる。コンテンツがこれら主体の間を循環していくことで、多彩かつ魅力のあるコンテンツがユーザから生み出される。

音声対話のコンテンツ循環系については、これまでに、ユーザがキャラクターとの対話を登録・編集できるコンテンツ作成システム [1] や、ユーザ層ごとのインセンティブ設計と誘引の分析 [2] などが試みられているが、実際に音声対話コンテンツのポジティブループを回すには至っていない。



図 1: 音声対話コンテンツの循環系

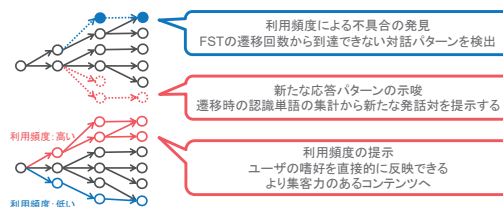


図 2: 対話シナリオ利用統計からのフィードバック

音声対話は、動画や画像と異なり、体験を提供するインタラクティブなメディアである。このため、音声対話コンテンツを一方に提供するだけでなく、ユーザの話した内容や受け答えへの反応などの動向をクリエイターへ強くフィードバックする仕組みが重要であると考えられる。フィードバックによって、想定と実際に噛み合ったより質の高いコンテンツへと発展することが容易になることが期待される。また、ユーザの反応が積極的にクリエイターへフィードバックされることは、クリエイターへ満足度等のインセンティブ付与や、新たなコンテンツへのアイデアを供する効果も期待される。

3. 対話プロファイルの自動収集およびフィードバックの実現

コンテンツをユーザが利用した際の利用情報（対話プロファイル）を自動収集し、コンテンツ作成者へフィードバックするシステムを提案する。

連絡先: 李 晃伸, 名古屋工業大学, 愛知県名古屋市昭和区御器所町, 052-735-7550, ri@nitech.ac.jp

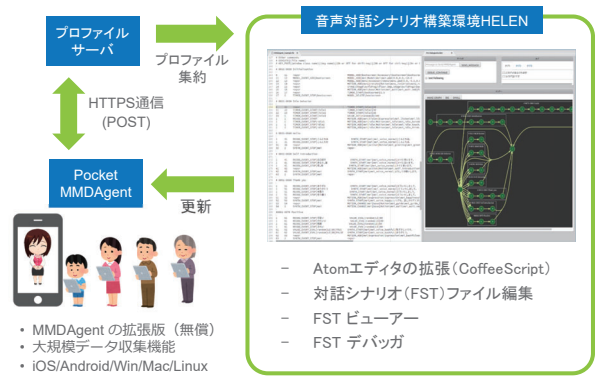


図 3: 提案システムの全体像

3.1 対話プロフィールからのフィードバック

図 2 に、対話プロフィールから得られるフィードバックの例を示す。図の左側は対話コンテンツにおける対話シナリオ（状態遷移モデル）である。ユーザが実際に辿った対話の状態遷移と各状態における発話内容、滞在時間を記録し、その情報を収集することで、状態ごとの利用率やユーザ離脱率、応答できなかった発話リストが得られる。これにより、シナリオ上で定義したが実際には利用されない対話パターンの検出や、ユーザがある状況で頻繁に発話するが対話シナリオで対応できていないパターンの抽出、ユーザの嗜好などが把握できる。

3.2 対話プロフィールの自動収集

提案システムの全体像を図 3 に示す。音声対話コンテンツの構築環境として、編集時の対話シナリオの対話状態遷移をグラフィカルに表示可能な MMDAgent 用の対話シナリオ編集環境 HELEN^{*1} を用いる。コンテンツ再生アプリは、ログ記録・収集機能を持つ Pocket MMDAgent[4] を用いる。Pocket MMDAgent は、コンテンツ利用時のログをサーバへ自動アップロードする。アップロードされたログからサーバ上でプロフィール情報を集計・抽出し、コンテンツ開発元のクリエイターへ提供する。HELEN では提供されたプロフィール情報から対話シナリオの状態遷移図（図 3 右部）上に利用率を色分け表示するとともに、利用率、離脱率、平均滞在時間および認識単語リストを状態ごとにポップアップ表示する。これによって、フィードバックを確認しながら編集を行える。

4. 評価実験

まずフィードバックがコンテンツ構築に与える効果について検証した。10 名の学生を対象に、既存の作成済み音声対話コンテンツ（自己紹介・10 対話）を与え、エディタを用いて自由に対話内容を追加してもらった実験を行った。まずプロフィールなしの状態課題を行い、次に事前実験で収集したプロフィールに基づいたフィードバック情報を表示した状態で課題を行った。利用後にアンケートによる主観評価を行った結果を表 1 に示す。フィードバック情報は音声対話コンテンツの制作・更新の際に有用な情報となり、認識単語リストおよび利用率表示が有用であることが示唆された。一方、離脱率表示はあまり有用でなく、複雑なコンテンツの制作にも大きくは寄与しなかった。コンテンツの修正や簡易な追加といった小規模の修正には有用であるが「新たな対話内容を発想する」という創作に近い行為への敷居を下げる効果は小さかったといえる。

*1 近日中に GitHub にて公開予定

表 1: 実験結果

質問内容	評価（5 段階）
コンテンツ制作に役立ったか	4.0
編集が楽になったか	4.0
コンテンツ制作の敷居を下げると思うか	3.2
思い通りにコンテンツを制作できたか	4.0
今後もツールを利用してみたいか	3.7
認識単語リストは有用だったか	4.5
離脱率表示は有用だったか	2.7
利用率表示は有用だったか	3.9
コンテンツ制作を手助けするか	3.8
複雑なコンテンツを制作できたか	3.1
簡単にコンテンツを制作できたか	4.0
コンテンツ制作の励みになったか	4.1

プロフィール情報の創作行為への効果をさらに調べるため、実際にクリエイターとして活発に活動しているアマチュア制作者（対話システムは構築未経験）に本システムを利用してもらった実験を行った。システムを口頭説明し、3D モデルなど素材を提供したのち、3 週間の期限で自由にコンテンツを創作してもらい、実験後にインタビューを行った。作成は期間を前半と後半に分け、前半で作成したシステムを数人のユーザに利用してもらい、そのフィードバックを与え、後半の作成を行うようにした。結果として、クリエイターが構築したコンテンツは、早口言葉を認識するゲーム（150 状態・200 遷移）であった。インタビューの結果、利用率や認識単語リストがユーザ動向の把握に重要な手がかりとなり、コンテンツ制作における有効な指標となり、コンテンツ制作のモチベーションとなることが示唆された。一方で、開発当初に想定した内容以上のものがフィードバックによって新たに生まれることはなかった。モチベーションの維持には、今回のようなパッシブな情報だけでなく、ユーザからのコメントや要望といったアクティブなフィードバックも必要であると考えられる。

5. おわりに

音声対話コンテンツにおいて対話プロフィールの自動収集とフィードバックがコンテンツ循環の一助となることを示した。今後の展開としては、音声対話コンテンツ配信の仕組みやユーザ間の体験の共有方法、アクティブなフィードバック、大規模実験等が挙げられる。

参考文献

[1] 宮木京介, 飯塚遼, 李晃伸, “利用者による履歴付き対話の共同構築・拡張が可能なユーザ生成型音声対話システム”, 日本音響学会秋季研究発表会講演論文集, 3-Q-22, 2015.

[2] 飯塚遼, 李晃伸, “ユーザ生成型音声対話システムにおけるクリエイターとユーザの相互刺激によるインセンティブ向上の検討”, 研究報告 自然言語処理 (NL) 2014-NL-216(6), 1-6, 2014.

[3] 関亜紀子, 亀山渉, “コンテンツ循環における権利詠唱の自動化”, 情報処理学会論文誌, Vol.48, No.5, p.1952-1964, 2007.

[4] 李晃伸, “音声対話コンテンツのネットワーク配信および大規模ログ収集を可能にするスマートフォン版 MMDAgent の開発”, 日本音響学会秋季研究発表会講演論文集, 2-2-8, 2019.