

音声対話エージェントのための Web ブラウザを用いたシナリオエディタの開発

西村 良太^{1,2} 山本 大介^{1,2} 打矢 隆弘^{1,2} 内匠 逸^{1,2}

概要: 本論文では, 音声対話エージェントのシナリオ構築環境を改善するべく, シナリオエディタの開発を行う. このシナリオエディタを用いることで, 対話シナリオを見やすくし, 編集しやすくして, より簡単に扱うことができるようにするのが目的である. 開発したシナリオエディタには, 大きく 3 つの特徴がある. 1 つ目の特徴は, 様々なプラットフォームでの動作が可能ということである. MMDAgent 自体も, PC (windows, mac, Linux) や, スマートフォン (android) など様々なプラットフォームに移植されており, シナリオエディタについても, 同様の環境での需要が考えられる. これらの環境で広く動作させる為, web ブラウザベースでの開発を行なっている.

Development of Dialog Scenario Editor aimed at popularization of Spoken Dialog System

RYOTA NISHIMURA^{1,2} DAISUKE YAMAMOTO^{1,2} TAKAHIRO UCHIYA^{1,2} ICHI TAKUMI^{1,2}

1. はじめに

近年, 音声扱う諸技術が発展し, 音声認識などは, 実際にサービス化され広く普及してきている. そこから, 音声対話システムを構築する基盤技術が整いつつあり, Apple の Siri[1] や NTT ドコモのしゃべってコンシェル [2] のように, 商用化された音声対話エージェントも登場し, 音声対話エージェントの普及に火が付き始めている. しかし, まだ音声対話エージェントが広く一般に普及しているとはいえない.

簡単に誰もが音声対話エージェントを利用できるようにするため, 音声対話エージェント構築ツールキット (MMDAgent[3]) が開発され, オープンソース・ソフトウェアとして公開されている. そして実際の応用例として, 名古屋工業大学の正門前にデジタルサイネージとして設置され [4], 等身大 3D キャラクタ「メイちゃん」との音声対話が可能になっている. このソフトウェアは, PC 用に開発されているが, スマートフォン用に Android に移植したバージョン

を, 我々が開発している [5]. また, メイちゃんと Skype による VoIP のビデオ通信機能を用いて音声対話が可能システムであるモバイルメイちゃん [6] も開発し, オープンキャンパスなどでも活用されている.

このツールキットには, 音声対話システムに必要な, 音声認識, 音声合成, 3DCG・キャラクタ表示, 対話処理などのソフトウェアやモデルなどが含まれており, それらを導入するだけで, 音声対話エージェントを構築する環境が簡単に入手可能となっている. しかし, 音声対話エージェントの構築には, 音声対話エージェントに関わる専門的な知識 (音声認識, 音声合成, 対話処理など) が必要であるため, 初心者ユーザが, それらの知識を持たずに, 対話エージェントを構築する事は困難である.

また, 音声対話システム構築では, 対話のやり取りを決定するための, 対話シナリオの編集が必要不可欠である. しかし, 現状の編集環境では, その複雑さや見づらさで, 初心者ユーザだけではなく, 上級者ユーザでも, 編集作業に要する労力が大きくなっている. そこで, 本論文では, そのような環境を改善するべく, 対話シナリオ編集エディタを開発する.

¹ 名古屋工業大学
Nagoya Institute of Technology, Gokiso Showa, Nagoya,
Aichi 466-8555, Japan

² CREST, JST

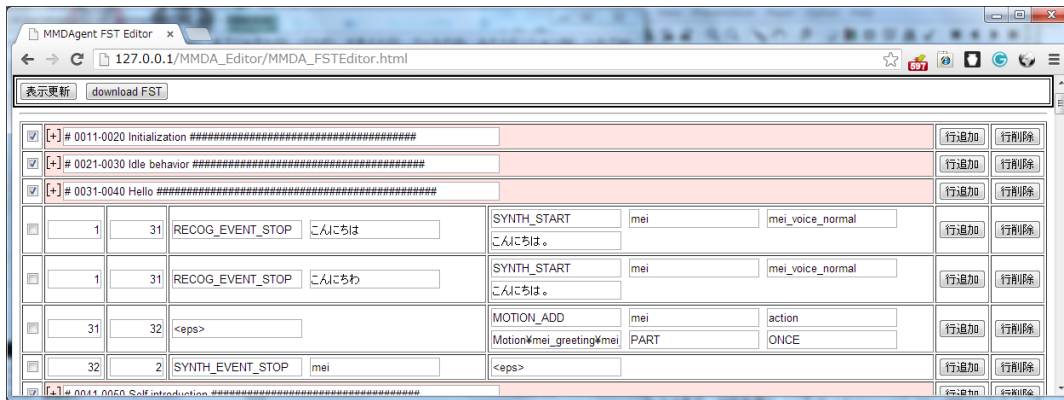


図 1 対話シナリオエディタ

```

1 10 RECOG_STOP| こんにちは <eps>
1 10 RECOG_STOP| おはよう <eps>
10 11 <eps> MOTION_ADD|mei|greet|greet.vmd
11 12 <eps> SYNTH_START|mei|normal| おはよう。
12 13 SYNTH_STOP|mei MOTION_ADD|mei|happy.vmd
13 14 <eps> SYNTH_START|mei|happy| 素敵な朝ですね。
14 1 SYNTH_STOP|mei <eps>

```

図 2 対話シナリオ FST の例 (一部表記を簡略化)

2. 対話シナリオ

MMDAgent の対話シナリオは、システム内で有限状態トランスデューサ (FST : Finite State Transducer) 形式で管理されており、ユーザがシナリオを作成する場合には、FST の状態遷移をテキストファイルで記述して、用意する必要がある (図 2)。FST は、状態番号、遷移先状態番号、受入条件、コマンドの 4 つ組 (スペース区切り) のリストで構成されている。図 2 の例では、状態番号が 1 のとき、“RECOG_STOP| こんにちは” という内部イベントを受信すると、 ϵ 遷移により状態番号 10 に遷移し、“MOTION_ADD|mei|greet|greet.vmd” という内部コマンドメッセージを出力して、状態番号 11 に遷移する。このように、状態遷移と内部メッセージのやり取りを繰り返すことによって、音声対話の制御を可能にしている。なお、<eps> は ϵ 遷移を表し、無条件で遷移することを表す。

シナリオの内容が記述されているテキストファイルには、図 2 のような行がたくさん並んでおり、それぞれの項目でインデントが揃っていないかったり、命令コマンドの意味が理解できなかつたりして、可読性が悪く、テキストエディタでの編集は、上級者ユーザでも不便で難しい。

3. 対話シナリオエディタ (MMDAE)

本論文では、上述したような問題を受けて、音声対話

エージェントのシナリオ構築環境を改善するべく、シナリオエディタ : MMDAE (MMDAgent scenario Editor) を開発した (図 1)。MMDAE を用いることで、対話シナリオを見やすくし、編集しやすくして、より簡単に扱えるようにすることが目的である。MMDAE には、大きく 3 つの特徴がある。1 つ目の特徴は、入力の補助を行う機能である。シナリオの記述には、図 2 のように、4 つの項目の入力が必要である。この中の、条件やコマンド部分を全て手入力するのは大変である。そこで、入力途中で候補を補完するようにした。また、命令を選択すると、引数を入力するテキストエリアが、それぞれの命令に応じた数だけ表示される。

2 つ目の特徴は、様々なプラットフォームでの動作が可能ということである。MMDAgent 自体も、PC (windows, mac, Linux) や、スマートフォン (android) など様々なプラットフォームに移植されており、シナリオエディタについても、同様の環境での需要が考えられる。これらの環境で広く動作させる為、web ブラウザベースでの開発を行っている。3 つ目の特徴は、環境に応じて編集モードを変更できることである。ユーザの音声対話エージェントに対する知識・経験の差に応じて、それぞれに合ったレベルの対話シナリオ編集ができるように、編集モードの変更が可能である。初心者ユーザに対しては、限定された項目だけを表示し、シナリオ記述に対する細かな知識がなくても、対話シナリオを記述できるようにする。上級者ユーザに対しては、細かい設定項目も編集できるように表示を行い、MMDAgent の全機能を用いて柔軟にシナリオを作り込める環境を提供する。また、シナリオ編集を行うハードウェア的な環境の違い (PC, スマートフォン) に対して、それぞれに適した表示を行うことで、シナリオ編集がしやすくする。例えば、スマートフォンは画面領域が小さい為、表示をコンパクトにするなどである。

4. 被験者実験

MMDAE がどれほど効果的であるのかを調査するため

```

●対話 1
ユーザ   : おはよう
メイちゃん : おはよう

# 0031-0040 Hello ##
1  31  RECOG_STOP| こんにちは ...
<●ここに追記●>
31  32  <eps> ...
32  2   SYNTH_STOP|mei ...

```

図 3 編集作業課題の例

に、被験者実験を行った。被験者は、20代の大学生・大学院生の男性13名である。

実験は、以下の手順で行った。

- (1) 実験説明
- (2) 対話シナリオ (FST) の説明
- (3) アンケート項目確認
- (4) 編集実験
 - テキストエディタ (メモ帳) による編集
 - MMDAE の使い方の説明, MMDAE による編集
- (5) アンケート

MMDAE を用いることで、対話シナリオの編集がどれほど改善されたのかを比較するため、テキストエディタであるメモ帳での編集と MMDAE での編集をしてもらい、その間の比較を行った。実験では、予め用意された対話シナリオファイルに、目的の対話内容を追記する形で行ったが、何もない状態から対話シナリオを作る事も想定し、メモ帳では、コピー&ペーストの使用は禁止とした。

編集ツールを使う順番によって、編集への慣れが実験結果に影響することが考えられるため、被験者を2つのグループに分け、各グループで編集ツールを使う順番を変更した (メモ帳→MMDAE のグループと、MMDAE→メモ帳のグループ)。実験中には、各被験者が編集にどれだけ時間がかかったかも記録した。実験は、研究室の各人の席で、被験者が普段使用している PC を用いて行った。

編集作業については、図3に示すような形で被験者にタスクを提示した。タスクは、「対話1」から「対話4」までの4つが用意され、全く同じ課題をメモ帳、MMDAE の両方で実行する。被験者は、MMDAgent の対話シナリオ編集経験が無く、難しい課題は困難である為、図に示すような、簡単な対話で、FST 命令を1文追加すれば良いような、簡単な課題にしてある。

実験後にアンケートを行い、以下の項目について、5段階評価をつけてもらった。Q1 の場合は、編集しにくい:1点～編集しやすい:5点となっている。

- 編集しやすさ (Q1:メモ帳, Q2:MMDAE)
- 表示の見やすさ (Q3:メモ帳, Q4:MMDAE)

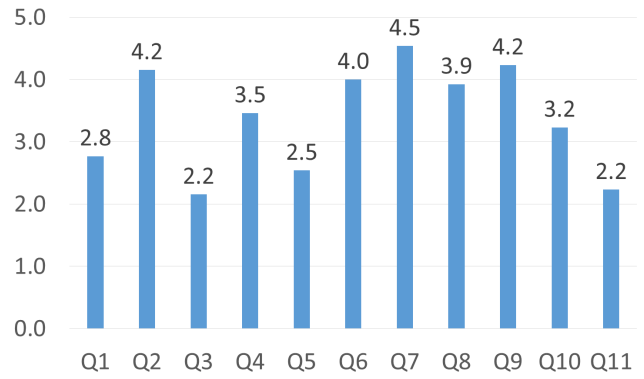


図 4 被験者実験の評価結果 (5段階評価)

表 1 編集時間

	メモ帳	MMDAE
全体平均	4:50	3:58

- 理解のしやすさ (Q5:メモ帳, Q6:MMDAE)
 - 機能の便利さ
 - (Q7:入力補完, Q8:TAB 移動, Q9:テキストエリア分割)
 - MMDAgent について
 - (Q10:知っているか, Q11:使ったことがあるか)
- また、以下の項目について、自由筆記によるアンケートを行った。
- Q12: どのような機能があれば便利だと思いますか?
 - Q13: その他 システムの良かった点, 悪かった点, 気になる点など

5. 実験結果

5段階評価による主観評価と、編集時間による客観評価を示す。

5.1 主観評価

図4に、実験結果を示す。質問内容は、4節に示したとおりである。図を見ると、メモ帳とMMDAE とを比べると (Q1 対 Q2, Q3 対 Q4, Q5 対 Q6), すべての項目でMMDAEの方が高評価を得ており、1点以上の差が開いている。特に、編集しやすさ、理解のしやすさにおいては、MMDAE が平均4点以上の評価を得ていた。また、MMDAE の特徴的な機能に対する評価 (Q7:入力補完, Q8:TAB 移動, Q9:テキストエリア分割) についても、全体に非常に高い点数を得ている。これらのことから、MMDAE がメモ帳と比較して、対話編集に用いる効果が得られることがわかる。

5.2 客観評価

編集にかかった時間により、それぞれの効率の良さの比較を行った。その結果を、表1に示す。メモ帳による編集時間が、平均で4分50秒かかっていたのに対して、



図 5 MMDAE の簡易編集モード

MMDAE による編集時間は、平均で 3 分 58 秒であった。MMDAE を用いることで、平均で 52 秒 (メモ帳による編集時間の約 18%) も編集時間の効率が上がっていた。

5.3 自由筆記による意見

自由筆記によるアンケート結果から、被験者が実験中に感じた、MMDAE に対する意見で、代表的なものを以下に挙げる。

ポジティブな意見としては、以下のものがあつた。

- 入力補完によって作業効率はかなり良くなった
- テキストエディタと比べてとても見やすかつた
- FST の知識が無くても、会話文の定義が自動補完により容易にできた

このように、MMDAE の効率改善は、被験者にも感じられていた。また、見やすさや機能についても、ポジティブな意見が多く見られた。今回の被験者は、対話シナリオファイルをほとんど触ったことがなく、FST のコマンドなどについても知識がなかつたが、MMDAE によって命令の候補などがリストとして表示されることで、初心者が FST の勉強をしながら、編集出来る環境として使えることがわかつた。

一方、ネガティブな意見としては、以下のものがあつた。

- テキストボックスが多くて見づらい

見やすさに関しては、ポジティブな意見が多かつたものの、ネガティブな意見も多少挙がつており、人によって合う表示形式があることが予想される。今回の実験では、詳細編集を想定した編集画面のみ編集を行ったが (図 1)、図 5 に示すような編集画面モードもある。テキストボックスが多くて見づらいつと感じた被験者は、このような簡素な表示による編集画面を用いると、編集がしやすくなるものと考えられる。

また、要望としては、以下のものがあつた。

- ボタンなどの説明が、フォーカスをあてると出ると良い

- 行のコピー

被験者は、今回はじめて MMDAE を使用したため、使い方などが、それぞれの部分に表示されると、ツールの使い方を学びながら、対話シナリオの編集ができる。この機能は、初心者には重要な機能であるため、今後 MMDAE に追加していきたい。また、1 行毎、または複数行をコピーすることができれば、これまでに記述した対話シナリオと似た構造の対話シナリオを記述する際には、有用である。この機能についても、追加したい。

6. まとめ

本論文では、音声対話エージェントのシナリオ構築環境を改善するべく、シナリオエディタ (MMDAE) の開発を行った。被験者実験の結果、メモ帳などのテキストエディタで編集するよりも、MMDAE で編集するほうが、編集時間が約 18% も短縮された。また、使用感に関するアンケートを行った結果、こちらも MMDAE のほうが高い評価を得ていた。

今後は、被験者アンケートにより得られた結果をもとに、必要な機能を追加していく予定である。具体的には、5.2 節に挙げた他に、FST 番号の自動付与や、遷移先のチェックなど、誤入力に対処する機能。システム応答文のレビュー (音声合成で再生)

また、音声認識キーワードを、音声で入力することも考えている。これにより、シナリオ作成時点で、システムの挙動を確認しながら作業ができるため、非常に有用である。また、応答文も音声で入力することで、話す速さや、アクセントの付け方、音の高低などの情報も入力できるようにしたい。

参考文献

- [1] Apple Inc.: Siri, <http://www.apple.com/ios/siri/>.
- [2] NTT ドコモ: シャベってコンシェル, http://www.nttdocomo.co.jp/service/information/shabette_concier/.
- [3] 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, 情報処理学会研究報告, Vol. 2011-SLP-89, pp. 1-6 (2011).
- [4] 大浦圭一郎, 山本大介, 内匠 逸, 李 晃伸, 徳田恵一: キャンパスの公共空間におけるユーザ参加型双方向音声案内デジタルサイネージシステム, 特集「音声対話」, 人工知能学会誌, Vol. 28, No. 1, pp. 60-67 (2013).
- [5] 山本大介, 大浦圭一郎, 西村良太, 打矢隆弘, 内匠 逸, 李 晃伸, 徳田恵一: スマートフォン単体で動作する音声対話 3 D エージェント「スマートメイちゃん」の開発, インタラクション 2013, IPSJ Symposium, Vol. 2013, No. 1, pp. 675-680 (2013).
- [6] 打矢隆弘, 山本大介, 柴川元宏, 吉田真基, 西村良太, 内匠 逸: ビデオ通話型音声対話サービス「モバイルメイちゃん」の開発, JAWS2012 講演論文集 (2012).