

雑談対話におけるキーワードベース大規模対話ルール構築

光田 航 東中 竜一郎 大塚 淳史 片山 太一 齋藤 邦子 富田 準二

日本電信電話株式会社 NTT メディアインテリジェンス研究所

{mitsuda.ko, higashinaka.ryuichiro, otsuka.atsushi,
katayama.taichi, saito.kuniko, tomita.junji}@lab.ntt.co.jp

1 はじめに

雑談対話は人間の会話の約 60%を占めるという調査 [10] や、タスク指向対話の達成を助けるという報告 [5] があり、タスク指向対話とは異なる、非タスク指向対話（いわゆる雑談）を行う対話システムの研究が注目されている [4]。近年、seq2seq に代表される生成ベースの雑談対話システム [8] が数多く報告されているが、一般的な商用の対話システムでは、カスタマイズ性の高いルールベースの手法が利用されることが多い [14]。

図 1 にルールベースの雑談対話システムが利用する対話ルールの例を示す。図 1 に示したルールは、Artificial Intelligence Markup Language (AIML) [9] と呼ばれる形式で記述されており、対話システムは、ユーザ発話を pattern にマッチさせることで発話を理解し、マッチしたルール (category) の template を返すことでユーザに応答する。雑談対話で扱いたい対話の内容に合わせてあらかじめ人手でルールを用意しておくことで、ルールの範囲内で品質の高い応答を実現することができる。幅広い話題で高品質な雑談を行うためには、大規模な AIML 構築が必要となる。

日本語における大規模な AIML を作成し、雑談対話システムを構築した研究が Higashinaka らによって報告されている [2]。対話コーパスを元に、連続する適切な 2 発話のペアを pattern と template のペアとみなすことで AIML の初期ルールセットを作成し、評価セット中の発話に対して適切な応答が返るまで AIML を修正することで、AIML の構築を行っている。また、250 個のキーワードに対して作成された 25,000 発話を元に、キーワードと発話のペアを pattern と template のペアとみなして AIML の拡充を行うことで、最終的に 149,300 個の category で構成される AIML を構築している。

さらに、Higashinaka らが構築した AIML の改善を行う試みとして、不適切な発話の除去 [14] や、ルールの追加・修正 [3] を行った研究が報告されている。ルー

```
<aiml>
  <category>
    <pattern>お酒 * 飲め * か</pattern>
    <template>飲めますよ</template>
  </category>
  <category>
    <pattern>タバコ * 吸い * か</pattern>
    <template>全然吸いません</template>
  </category>
  ...
</aiml>
```

図 1: 雑談対話における対話ルール (AIML) の例。pattern は形態素単位で区切られており、'*' が任意の形態素列と対応することでユーザ発話とマッチする。

ルの追加・修正では、複数の作業者が作成したユーザ発話、および、複数ターンの対話を通じてシステムの評価を行い、その評価値が閾値以上になるまで繰り返しルールの追加や修正を行うことで、AIML の改善を行っている。最終的に、ルールの修正と評価を 8 回繰り返すことで作成された AIML の category 数は 333,295 個となったが、修正されたルールの有効性を確認するためのチャット形式の評価において、修正前の AIML と比較して対話の品質の向上がほとんど見られず、大規模な AIML の改善は容易ではないことが報告されている。その要因として、対応可能な話題が少ないためにシステムが同じ発話を繰り返してしまう点や、文脈理解が必要な対話に対応できていない点が挙げられている。

先行研究 [3] で述べられている問題のうち、文脈理解の問題については、文脈ごとに AIML を作成する必要があり、効率的に AIML を改善する方法は明らかではない。一方で、対応可能な話題が少ない問題については、その話題に対応可能な AIML を追加することで解決できると考えられる。先行研究で構築された AIML がカバーする話題の規模を調査するために、AIML の pattern に含まれる名詞の数を調査したところ、異なりで 11,360 個となり、例えば地名 (例. 大阪) や食べ物 (例. たこ焼き) など、典型的な話題を

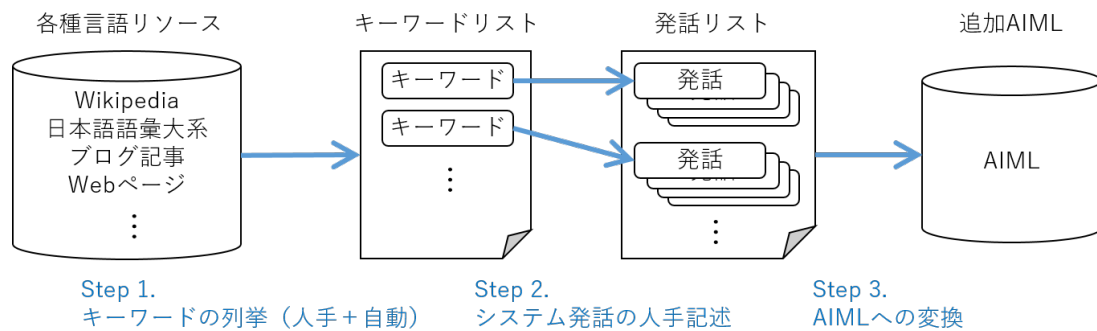


図 2: キーワードベースの対話ルール (AIML) の構築手順

カバーできていない場合があることが判明した。これは、先行研究では主に対話コーパスに基づいて AIML を構築・整備しており、雑談対話で扱われるさまざまな話題をカバーするような AIML を直接的に作成してはいないためである。

本研究では、先行研究 [3] で構築された AIML を改善するために、雑談で頻出すると考えられる話題 (キーワード) を各種言語リソースを用いて列挙し、それらに対応するシステム発話をルールとして人手で大規模に作成することで、AIML が対応可能な話題の拡充を行った。作成したキーワードベースの AIML に基づく応答発話を人手で評価したところ、評価値 (発話の有益さ) の改善が確認できた。本稿では、AIML の構築方法、および、その評価について述べる。

2 キーワードに基づくルールの構築

図 2 にキーワードベースの AIML の構築手順を示す。まず、Step 1 として、Wikipedia や日本語語彙大系などの言語リソース、および、人手による作業に基づき、雑談対話で頻出すると考えられるキーワードを列挙する。次に、Step 2 として、列挙されたキーワードを用いて、各キーワードに対応するシステム発話を人手で記述する。最後に、Step 3 として、キーワードと発話のペアを pattern と category とみなして AIML に変換することで、キーワードベースの AIML を構築する。

2.1 キーワードの列挙

雑談対話で扱う必要があるさまざまな話題をキーワードとしてカバーするためには、人手で列挙する単語だけでなく、言語リソースから自動的に抽出した単語も合わせて利用することが効果的である。異なる言語リソースには異なる特徴の単語が含まれているため、それらを適切に組み合わせることで、網羅性の高いキーワードリストを作成できると考えられる。本稿

では、既存の言語リソースから抽出できるキーワードをカテゴリとして整理し、適宜人手も用いることで、2 万語のキーワードリストを作成した。

表 1 に列挙したキーワードをカテゴリごとにまとめたものを示す。表 1 にはキーワードのカテゴリごとにキーワード数、概要、列挙方法と使用リソース、キーワード例が示されており、全部で 16 種類のカテゴリが掲載されている。ID1 から ID3 までのカテゴリは新語に対応するためのカテゴリであり、ID4 以降のカテゴリはそれ以外のキーワードが属するカテゴリである。言語リソースから自動的にキーワードを列挙するカテゴリ (ID4, ID5, ID14) には、雑談対話においてキーワードとして出現しにくいと考えられる単語が多く含まれていたため、クロールしたブログ記事コーパス (1ヶ月分) でフィルタを行い、10 回以上出現しているもののみをキーワードとして利用した。人手が必要なカテゴリ (ID1 から ID3, ID6 など) はカテゴリごとに 1 名の作業者が作業を実施することでキーワードを列挙した。キーワードを列挙する際には、商用サービスなどで問題となる不適切なキーワード [14] や商品名に該当するキーワードへのラベル付与を行った。

2.2 システム発話の人手記述

作成したキーワードリストに含まれる各キーワードに対して、そのキーワードを含むユーザ発話に対する応答となるシステム発話を人手で記述した。キーワードごとに作成する発話は、先行研究 [2] におけるキーワードベースの AIML 構築の方法にならって 4 発話とし、以下の 4 発話を作成した。

- (1) キーワードを含む平叙文
 - (2) キーワードを含む質問文
 - (3) キーワードを含まず、関連する内容の平叙文
 - (4) キーワードを含まず、関連する内容の質問文
- 質問文や、キーワードを含まない発話も合わせて作成

表 1: 雑談対話における対話ルール (AIML) 構築のために列挙されたキーワードの一覧

ID	カテゴリ	キーワード数	概要	列挙方法	使用リソース	キーワード例
1	gendaiyugo	508 (2.5%)	現代用語の基礎知識に掲載されている単語	人手でキーワード化可能な名詞を抽出	現代用語の基礎知識 2018 [15]	ノーコン, パシュー ト, ハッカー
2	neologism	345 (1.7%)	Web に掲載されている新語・流行語	人手でキーワード化可能な名詞を抽出	新語・流行語大賞 ¹ , 今年の新語 ²	インスタ映え, 歩き スマホ
3	koujien	146 (0.7%)	広辞苑第 7 版に新しく掲載された名詞 (一部)	Web ページを参照し, 人手で新語を列挙	広辞苑第 7 版 [12]	朝ドラ, 小腹, 健康 寿命
4	qa	7,120 (35.6%)	質問ランキングサイトにおける頻出の回答	別リソースで閾値回数以上出現する名詞を抽出	質問ランキングサイトの回答データ [13], ブログ記事	コミュニケーション 能力, ビラティス
5	ene	4,289 (21.4%)	拡張固有表現 (ENE) でフィルタした Wikipedia	人手列挙の ENE が自動付与されたページ名を列挙	Wikipedia, ENE [6], ブログ記事	インフルエンザ, リ チウム, 真鯛
6	trip	1,834 (9.2%)	国内の観光スポットやグルメ	人手で県ごとに観光スポットやグルメを列挙	観光案内サイトなど観光関連情報の Web ページ	牛タン, ロイズ, な めろう
7	famous	1,271 (6.4%)	有名な人物, 団体, 作品の名称	Web ページを参照し人手でキーワードを列挙	ランキングなどをまとめた各種 Web ページ	バイマックス, ス ターウォーズ
8	city	984 (4.9%)	国内の主要都市	国内の主要都市をまとめたリストから抽出	基準地・重要地・主要地一覧 ³	帯広, 瀬田, さつま
9	pdb	975 (4.9%)	Person DB (PDB) 中の質問の典型回答	PDB の各質問で典型的と考えられる応答を人手列挙	PDB [7]	マニキュア, 遺跡巡 り, DS
10	work	464 (2.3%)	職業名	Wikipedia の「職業一覧」ページのリストから抽出	Wikipedia の職業一覧 ⁴	大統領, 外交官, 塾 講師
11	genre	374 (1.9%)	雑談対話における各話題で特徴的な名詞	決められたジャンルごとに名詞を人手で列挙	Higashinaka らが定義したジャンル一覧 [2]	つけまつげ, プロバ イダ, 従兄弟
12	prefecture	264 (1.3%)	都道府県や県庁所在地, および, 主な地名	Web ページを参照し人手でキーワードを列挙	都道府県, 県庁所在地をまとめた Web ページ	北海道, 札幌
13	transport	185 (0.9%)	主要駅などの公共交通機関関連名称	Web ページのリストから抽出	利用者数の順位や主要駅一覧をまとめた Web ページ	渋谷駅, 東名, ゆり かもめ
14	goitaikai	155 (0.8%)	日本語語彙大系の一部の名詞意味カテゴリの名詞	人手で選択したカテゴリに属する名詞を抽出	日本語語彙大系 [11], ブログ記事	博士, 朝風呂, 食い しん坊
15	date	96 (0.5%)	日付関連の用語やイベントの名称	Web ページを参照し人手でキーワードを列挙	行事, 記念日, イベントなどをまとめた Web ページ	1 月, 春分の日, 2019
16	misc	990 (5.0%)	その他 (海外の観光地や大学名など)	Web ページを参照し人手でキーワードを列挙	ランキングなどをまとめた各種 Web ページ	オスロ, ピサの斜 塔, 東大
-	合計	20,000 (100%)	-	-	-	-

しておくことで, 話題を変えるなど幅広い発話が可能になる. 各キーワードにつき作業員 1 名が発話を作成した. 発話の作成者にはインストラクションを与え, 作成する発話は, キーワードを含むユーザ発話に幅広く応答可能で, 文脈に依存せず, 有益な情報をなるべく多く含み, 対話システムとして問題のない (例えば, 政治的思想を含まない) 発話であると指示した.

2.3 AIML への変換

図 3 に最終的に作成された AIML の例を示す. 列挙したキーワードと各キーワードに対して記述された発話をペアとみなし, 機械的に変換することで AIML を作成した. キーワードは, 形態素解析 [1] を行って分かち書きした後, 任意の形態素列にマッチする形態素 '*' をキーワードの前後に挿入することで, $2 \times 2 = 4$ 通りの pattern を作成し, 作成されたシステム発話 (template) と組み合わせることで AIML を作成した.

```
<aiml>
<category>
  <pattern>たこ焼き</pattern>
  <template>たこ焼きは熱くてやけどします</...>
</category>
...
<category>
  <pattern>* たこ焼き *</pattern>
  <template>粉物がお好きなんですか?</...>
</category>
</aiml>
```

図 3: キーワード「たこ焼き」に対して作成された対話ルール (AIML) の例

3 構築した対話ルールの評価

キーワードに基づいて構築した新規の AIML の有効性を確認するために, 先行研究 [3] で作成された AIML との比較を行った.

比較を行うために, 100 発話からなる評価セット 1, および 2 を用意した. 評価セット 1 は, キーワードに対して作成された AIML の有効性を確認するために, 本研究で列挙したキーワードからランダムに選択した 100 単語を含む入力発話を人手で作成したものであり,

¹<https://ja.wikipedia.org/wiki/新語・流行語大賞>

²<https://ja.wikipedia.org/wiki/今年の新語>

³<http://www.mlit.go.jp/road/sign/sign/annai/6-hyou-timei.htm>

⁴<https://ja.wikipedia.org/wiki/職業一覧>

表 2: 構築された対話ルールの人手評価

ルールセット	評価セット 1		評価セット 2	
	自然さ	有益さ	自然さ	有益さ
OLD	2.76	2.45	2.33	1.93
OLD+NEW_1/4	2.66	2.60	2.30	1.99
OLD+NEW_1/2	2.49	2.61	2.45	2.21
OLD+NEW_ALL	2.51	2.82	2.45	2.28

評価セット 2 は、一般の雑談対話における AIML の有効性を確認するために、質問応答や感情理解など 10 個の観点で入力発話を列挙したものである。

評価値として、入力発話に対する AIML の応答のつながりの良さを表す「自然さ」、および、応答に有益な情報や話題が含まれるかを表す「有益さ」を用い、5 段階の Likert 尺度で評価を実施した。評価は AIML 構築とは無関係な 2 名の作業者が行った。

比較する対話ルールは 4 種類とし、先行研究 [3] で作成された AIML (“OLD”), 本研究で作成した AIML を追加した AIML (“OLD+NEW_ALL”), 本研究で作成した AIML を部分的 (1/4 または 1/2) に追加した AIML (“OLD+NEW_1/4” と “OLD+NEW_1/2”) とした。

表 2 に構築した対話ルールを人手評価した結果を示す。評価セット 1, および, 2 において, “OLD” と “OLD+NEW_ALL” の有益さには統計的に有意な差が見られ (ANOVA および t 検定, 有意水準: .05), キーワードに基づいて新規に作成した AIML が有効であることが確認できた。

図 4 にキーワードベースの対話ルールを追加する前 (“OLD”) と後 (“OLD+NEW_ALL”) の対話ルールに基づく応答の例を示す。1 つ目の例については, 追加前の AIML では「(地名) はどこにあるのですか」という応答を返しているのに対し, 追加後の AIML では地名 (大阪) に関連する応答ができています。2 つ目の例については, 追加前の AIML では「たこ焼き *」にマッチするルールが存在せず, デフォルトの応答が帰っているのに対し, 追加後の AIML では, 作成した「たこ焼き」のルールを用いて, たこ焼きに関連する応答ができており, より広い話題をカバーできていることがわかる。

4 おわりに

本稿では, 先行研究 [3] で構築された対話ルール (AIML) の性能を改善するために, 言語リソースに基づいて列挙したキーワードから作成した AIML を大規模に追加する手法を提案した。雑談対話で扱われ

入力発話: 「大阪出身です」
 OLD: 「大阪はどこにあるのですか?」
 OLD+NEW_ALL: 「大阪は商人の町です」

入力発話: 「たこ焼きが好きです」
 OLD: 「そうなんですね」
 OLD+NEW_ALL: 「たこ焼きは熱くてやけどします」

図 4: キーワードに基づく対話ルールを追加する前と後の対話ルールに基づく応答発話の例

るさまざまな話題をカバーするために, 複数の言語リソース, および, 人手の作業を組み合わせることで, 2 万語のキーワードを列挙し, 各キーワードに対して 4 発話ずつシステム発話を人手で作成することで, キーワードベースの AIML を作成した。作成した AIML を評価したところ, 先行研究の AIML と比較して, 応答発話の有益さに改善が見られた。

今後は作成した AIML を既存の雑談対話システムや生成ベースの手法と組み合わせ, より自然かつ有益な応答が可能な対話システムの構築を目指す。

参考文献

- [1] Takeshi Fuchi and Shinichiro Takagi. Japanese morphological analyzer using word co-occurrence -JTAG-. In *Proc. COLING*, 1998.
- [2] Ryuichiro Higashinaka, Kenji Imamura, Toyomi Meguro, Chiaki Miyazaki, Nozomi Kobayashi, Hiroaki Sugiyama, Toru Hirano, Toshiro Makino, and Yoshihiro Matsuo. Towards an open domain conversational system fully based on natural language processing. In *Proc. COLING*, pp. 928–939, 2014.
- [3] Ryuichiro Higashinaka, Toyomi Meguro, Hiroaki Sugiyama, Toshiro Makino, and Yoshihiro Matsuo. On the difficulty of improving hand-crafted rules in chat-oriented dialogue systems. In *Proc. APSIPA*, pp. 1014–1018, 2015.
- [4] Alan Ritter, Colin Cherry, and William B. Dolan. Data-driven response generation in social media. In *Proc. EMNLP*, pp. 583–593, 2011.
- [5] Kazuki Sakai, Ryuichiro Higashinaka, Yuichiro Yoshikawa, Hiroshi Ishiguro, and Junji Tomita. Introduction method for argumentative dialogue using paired question-answering interchange about personality. In *Proc. SIGDIAL*, 2018.
- [6] Satoshi Sekine. Extended named entity ontology with attribute information. In *Proc. LREC*, 2008.
- [7] Hiroaki Sugiyama, Toyomi Meguro, Ryuichiro Higashinaka, and Yasuhiro Minami. Large-scale collection and analysis of personal question-answer pairs for conversational agents. In *Proc. IVA*, 2014.
- [8] Oriol Vinyals and Quoc Le. A neural conversational model. In *Proc. ICML Deep Learning Workshop*, 2015.
- [9] Richard S. Wallace. *The Anatomy of A.L.I.C.E.* A.L.I.C.E. Artificial Intelligence Foundation, Inc., 2004.
- [10] 小磯花絵, 石本祐一, 菊池英明, 坊農真弓, 坂井田瑠衣, 渡部涼子, 田中弥生, 伝康晴. 大規模日常会話コーパスの構築に向けた取り組み—会話収録法を中心に—. 第 74 回言語・音声理解と対話処理研究会, 第 5 巻, pp. 37–42, 2015.
- [11] 池原悟, 白井論, 横尾照男, 中岩浩巳, 小倉健太郎, 大山芳史, 林良彦. 日本語語彙大系 CD-ROM 版. 岩波書店, 1997.
- [12] 新村出. 広辞苑 第七版 (普通版). 岩波書店, 2018.
- [13] 大塚淳史, 片山太一, 杉山弘晃, 東中竜一郎, 松尾義博. 対話シナリオ構築におけるユーザ回答候補の推定. 人工知能学会第 31 回全国大会, 2017.
- [14] 角森唯子, 大西可奈子, 藤本拓, 角野公亮, 吉村健, 磯田佳徳. カスタマイズ可能なオープンドメイン雑談対話エンジンの開発. 第 84 回言語・音声理解と対話処理研究会, 2018.
- [15] 佐藤優, 谷本道哉, 島園進, 五野井郁夫, 伊藤真, 金子勝, 武田砂鉄, 吉村誠, 清家篤, 湯川れい子, 塚田穂高, 伊勢崎賢治. 現代用語の基礎知識 2018 年版. 自由国民社, 2017.