

イベント履歴収集システムの試作と年齢推定への活用

大竹 栄一[†] 遠藤 正隆[‡] 中嶋 裕一[‡] 松井 瑠偉人[‡] 菱田 隆彰[†]

愛知工業大学[†] 株式会社リオ[‡]

1. はじめに

対話システムは、主に特定の情報の授受のために対話を行うタスク指向型と、会話の継続を目的とする非タスク指向型に分けられる。最近では、人とより自然に会話することのできる対話システムの研究が盛んに行われている。定型文の質問を行うだけでなく、前後に雑談を挟むことで、相手とのより自然に近い対話をすることで少ないストレスで目的を達成することが可能となることが示されている[1]。相手を楽しませながら自然な会話を行うためには、相手の興味を持つ話題を選んで提供する必要がある。そのためには、相手の個人的な情報を会話から入手する必要がある。小林らの研究[2]では、会話の中に現れる特徴的なフレーズを個性として紐付けることで、相手の嗜好や状態を推測できることが示されている。しかし、実際には個人を特徴付けるような情報を平易な会話の中から取得するのは困難な作業である。

そこで我々は、よりユーザ個人にあった話題を提供するために、雑談の最中に意図的に特定の情報が出しやすくなるような対話の誘導を行い、平易な内容に付帯する個性的な情報を入手することを検討し始めた。一つの例として、相手が体験した特定の出来事の話から、その時期とその時の相手の年齢がわかれば自ずと現在の年齢が推定できることになる。この時、問題点は様々な出来事からその実施時期を得ることは容易ではないという点である。

我々は、この問題を解決するために大小様々な出来事について発生した時期を一元的に管理し、必要に応じて相互に変換することのできるイベント履歴収集システムを提案し、出来事情報の収集方法について検討を行った[3]。提案システムは Historical information Acquisition System と名付け、以降 HAS と呼ぶ。本稿では、HAS の試作及び、HAS が利用する出来事情報の収集を行う。その上で、HAS を利用した年齢推定可能な対話システムの構築を目指し、年齢推定のための LINE のチャットボットを作成し、HAS とそのテストデータの有効性を検証する。

```
{
  "year":2005,
  "year_alt":["平成 17 年", "21 世紀"],
  "date":"3/25-9/25",
  "event":"2005 年日本国際博覧会",
  "event_alt":"愛・地球博",
  "category":["イベント","愛知県"],
  "location":"35.1718934,137.0898625"
},
```

図 1 : 愛・地球博のデータ例 (json 形式)

表 1 : データの取得コマンド

Has get X	X の取得
-----------	-------

表 2 : オプションコマンド (一部)

オプション	説明
at A	A の年代
from A to B	A から B の間の年
by C	カテゴリ C を含む

A, B : データの属性値, C : カテゴリ

2. イベント履歴収集システム

イベント履歴収集システムである HAS は過去の様々な分野のあらゆる出来事の時期を提供するためのシステムである。基本的な機能は、出来事の名前や属性とその発生時期(年月日)を相互に変換することである。ドメイン名と IP アドレスの変換を行う Domain Name Service のように、HAS によって世界中の出来事の発生時期を一元的に集積・提供が可能となることで、過去の出来事を一意で定量的な情報として様々なシステムで活用することができる。

HAS は出来事情報を蓄積するデータベースとそのデータを抽出するためのコマンドラインインタフェース(CLI)から成る。例えば、2005 年に開催された「愛・地球博」の場合、図 1 のようにデータが登録される。登録されたデータを元に必要な出来事情報を得るためには表 1 に示す「get」コマンドを用いて行う。X に取得した属性を指定することで、該当する情報のみの抽出が可能となる。また、表 2 に示すようなオプションコマンドを付加することで、取得する情報に条件をつけることが可能となる。

A prototype of historical information acquisition systems and the application to age estimation method

[†] Eiichi Otake, [‡] Masataka Endo, [‡] Yuichi Nakashima,

[‡] Ruito Matsui, [†] Takaaki Hishida

[†] Aichi Institute of Technology, [‡] RIO CORPORATION

3. 年齢推定への活用

3.1 テストデータの収集

対話の中から得られるユーザーの特徴は様々に考えられるが、本稿ではユーザーの年齢に注目する。直接年齢を聞くことなく対話しているユーザーの年齢推定を行うための一つの方法として、ユーザーが体験した特定の出来事と体験した当時の年齢から算出する方法が考えられる。この場合、体験した出来事からその発生時期を得る必要が生じるため、対話システムに HAS を連携させることで、対話の中からより少ない手間ですべて年齢推定ができるようになる。利用の際には HAS に様々な分野の情報を多く持っていることが望まれる。我々は、年齢推定に利用できそうな分野として「事件、災害、映画、アニメ映画、ゲーム、音楽、漫画」の分野を対象とし、Wikipedia 上のデータから HAS の出来事情報のテストデータの生成を試みる。

情報の抽出・収集は Wikipedia の記事のインフォボックスに記載されている基本情報を DBpedia Japanese を経由し、用意されている SPARQL のエンドポイントから抽出を行い、HAS のフォーマットへの整形を行う。DBpedia では分野毎に対象データを取得する方法が異なるため、異なるクエリとデータ整形処理を用意する必要があった。対象とした7つの分野で1980年から2016年までのデータを対象に収集を行った結果を表3に示す。1980年から2016年までの各分野総合計で41,348件のデータを収集することができ、HASでの利用が可能となった。

3.2 年齢推定を行うチャットボットの試作

対話システムのプロトタイプとして収集したテストデータを搭載した HAS を利用して年齢推定を行う LINE のチャットボットを作成する。

チャットボットの基本的な対話の流れは

STEP1：過去の体験に関して質問する

STEP2：体験した当時の年齢を質問する

STEP3：推定年齢の結果を返す

STEP1 では過去の体験について事件、災害、映画、アニメ映画、ゲーム、音楽、漫画のいずれかの体験についてランダムに問いかける。ユーザーの回答を受けて、HAS に問い合わせを行い、出来事の年代情報を取得することができたら、STEP2 の発話を行いユーザーの返答を待つ。ユーザーが適切な年齢に関する返答をした時、STEP3 に移行し推定年齢の算出を行う。今回は、HAS を用いたチャットボットの動作確認とともに、収集したテストデータが年齢推定をする上でどの程度機能するかを検証する。

検証は1995年度生まれ(23又は24歳)の男性10人に対して、チャットボットと最大20回の対話を行ってもらい、STEP1の問いかけの回答がテストデータと適合した回数と年齢推定の結果を集計する。

表4は被験者の体験内容がテストデータにヒットし

表3：HASで提供可能なデータ件数

分野	件数
事件	1,023
災害	305
映画	12,355
アニメ映画	1,420
ゲーム	8,649
音楽	16,343
漫画	1,343

表4：テストデータのヒット率

	問合せ回数	ヒット回数	ヒット率
映画	28	17	60.7%
曲	41	22	53.7%
漫画	40	19	47.5%
ゲーム	52	30	57.7%
事件	13	3	23.1%
災害	11	5	45.5%
総件数	185	96	51.9%

た回数とその比率を示している。問いかけの半数は年齢推定に利用できることがわかった。

4. まとめ

本研究では、過去の出来事とその発生時期を相互に変換可能なイベント履歴収集システム HAS の設計と試作を行い、Wikipedia からテストデータを収集して LINE のチャットボットに HAS を組み込み対話形式で情報の取得、年齢推定を行うチャットボットの試作を行なった。チャットボットを介してユーザーが体験した過去の出来事情報とその時期を得ることで、LINE 上でユーザーの推定年齢の算出が可能となった。

今後の課題としては、検証結果からより多くの出来事情報が必要であることがわかった。情報の拡充方法の検討を行い、雑談を行いながら自然な流れで年齢推定が行えるような対話システムの構築を行いたいと考えている。

謝辞

本研究は、JSPS 科研費 JP19K12073 の助成を受けたものです。

参考文献

- [1]大村祐司, 川端豪, 「雑談可能な目的達成型音声対話システム」, 情報処理学会研究報告, Vol2012-SLP-94, No. 9, pp.1-5, (2012)
- [2]小林峻也, 萩原将文, 「ユーザーの嗜好や人間関係を考慮する非タスク指向型対話システム」, 人工知能学会論文誌, Vol31, No. 2, pp.1-10, (2016)
- [3]大竹栄一, 遠藤正隆, 中嶋裕一, 三浦哲郎, 菱田隆彰, 「Historical information Acquisition System(HAS)の実装」, マルチメディア, 分散, 協調とモバイル (DICOM2019)シンポジウム, pp.1099-1102, (2019)