

令和 3 年 6 月 16 日現在

機関番号：13801

研究種目：挑戦的研究（萌芽）

研究期間：2018～2020

課題番号：18K18504

研究課題名（和文）心理学的に妥当な自然言語処理システムによる会話の自然な「間」の自動生成

研究課題名（英文）Natural Pause Generation in Conversations by Psychologically Plausible Natural Language Processing

研究代表者

狩野 芳伸（Kano, Yoshinobu）

静岡大学・情報学部・准教授

研究者番号：20506729

交付決定額（研究期間全体）：（直接経費） 4,800,000円

研究成果の概要（和文）：より人間に近いモデルを採用することで、人間のように自然な「間」を自動的に生成し、相槌および会話中の割込タイミングを、構築した解析器を用いて自動的に推測、さらに対話システムによる応用的取り組みの端緒をつかむことが目的である。音声コーパスに相槌・割込みタイミングをはじめ各種の音声・言語学的アノテーションを付与し、音声処理とテキスト処理を統合したアプローチにより実際に推測システムを構築した。また話し言葉処理を応用した研究として対話システムの構築を行い、会話ゲーム「人狼」の自動プレイを行うコンテストとして毎年主催している人狼知能大会自然言語部門に参加するエージェントを実装し、有効性を確認した。

研究成果の学術的意義や社会的意義

統計的手法の隆盛により、既存の自然言語処理研究のほとんどは大規模データを取集ししやすい書き言葉を前提に研究がおこなわれている。そのために話し言葉の処理性能は未だ不十分なのが現状である。また、人間の言語処理は話し言葉を基盤としている可能性が高く、書き言葉の処理においても話し言葉の処理が有用な可能性があり、「間」をはじめとする音声的な要素も取り込んだシステム構築は、自然言語処理全般に資すると共に、人間の言語処理メカニズムの解明にもつながるものである。

研究成果の概要（英文）：By adopting a model closer to humans, we aim to generate a natural pause like a humans, and the interrupt/backchannel timing during conversations. We further aim to apply such a spoken language processing to dialogue systems. We recorded a spoken language corpus, then transcribed and assigned annotations of backchannel/interrupt timing, etc. We constructed an inference system using these annotations based on speech and linguistic features. We built a dialogue system and implemented an agent which participates in the natural language division of the annual AI Werewolf Contest, where agents automatically play the conversation game "Mafia".

研究分野：自然言語処理

キーワード：自然言語処理 間 生成

1. 研究開始当初の背景

相槌や会話への割り込みを含めた、発話のタイミングの研究は古くからおこなわれてきた。しかし、既存研究は基本的に記録済みのデータに対してタイミング判定を行う分析的なもので、計算機では未だ人間のようにリアルタイムでタイミングを制御できない。これは計算機能力の不足で計算が間に合わないだけでなく、人間同様のタイミングでの発話には相手のその先の発話予測が本質的に不可欠だからである。

既存の解析器の問題として、書き言葉を主眼にしている点、統計的手法の隆盛により、書き言葉を前提に大規模データと潤沢な計算資源を用いた処理が行われてきた点があげられる。膨大な数の候補を同時に保持し、その中から確率的に最も妥当と思われる候補を最終結果として出力しており、候補数を減らすと性能が大幅に低下してしまう。一方で人間のほうが高精度であるだけでなく、計算量も少ないであろうことは、古くから心理学的実験により知られている。例えば埋め込み構文の埋め込み回数の限界、袋小路文での処理の失敗とバックトラックから、解析候補はほとんど一つしか持たないであろうと考えられる。であれば、処理モデルや処理過程もより人間に近づけるアプローチが有効ではないかと考えた。

2. 研究の目的

本研究の第一の目的は、より人間に近いモデルを採用することで、人間のように自然な「間」を自動的に生成することである。たとえば人間同士であれば、「お互いに発話しかけて譲りあいどちらも話し始められない」といったシーンがみられるが、既存システム同士ではこれを本質的に再現できていない。また近年の多くの研究では、書き言葉を主眼にしている点、統計的手法の隆盛により、書き言葉を前提に大規模データと潤沢な計算資源を用いた処理が行われてきた。一方で人間のほうが高精度であるだけでなく、計算量も少ないであろうことは、古くから心理学的実験により知られている。こうした点を踏まえ、心理学的に妥当なモデルにより解析器を構築する。

相槌および会話中の割込タイミングを、この解析器を用いて自動的に推測することを目標とし、その判断基準となる手がかり（特徴量）の発見と分析を行うことが第二の目的である。

さらに、こうした話し言葉の解析は、対話システムから SNS 投稿や掲示板の書き込みまで幅広く応用できる可能性がある。そうした応用的取り組みの端緒をつかむことが、第三の目的である。

3. 研究の方法

- (A) 心理学的な制約を取り入れかつ解析可能（ホワイトボックス）な設計に基づき、解析候補数を限定しつつ逐次処理を行うモデルを採用した統合解析器を構築する。
- (B) 音声処理とテキスト処理を統合したアプローチにより実際にリアルタイムで動作するシステムを構築する。音声コーパスに相槌・割込みタイミングのアノテーションを付与し、構築した解析器による推測と特徴量の発見を行い、必要に応じてタイミング以外のアノテーション修正を行う。
- (C) 話し言葉処理を応用した研究として対話システムの構築を行い、応用の道筋をつける。

4. 研究成果

間や相槌が発生しやすい状況として、話し手・聞き手に役割が分担されるような傾聴対話の状況を設定し、30 分程度の音声収録を左右分離のステレオ音声で記録した。この音声に対し、書き起こしおよび発話区間、各種アノテーションを付与した。アノテーションの基準は、千葉大学 3 人対話コーパスをもとに、相槌の種類など独自の拡張を行ったもので、アノテーションの種類は表のとおりである。

記号	種類	説明
(.)	休止	注釈内での休止
:	引き伸ばし	標準的な発音に含まれない語中、語末の音の引き伸ばし

%	詰まり	非語彙的な音の詰まり。休止との違いは非語彙的であるかどうか。ちょうど文節で区切れている場合は休止記号を付与する。
-	語の中断	言いかけた語を途中で止めている場合その中断箇所で使用。言い直す場合が多いので言いさし記号と併用することが多い。ただし名詞の中で言いかけは言いさし記号と併用するが、連用形と動詞の言いかけは言いさし記号とは併用しない。
(F_word)	フィラー	「えっと」などのフィラー
?	上昇調	質問など上昇調に使用
(B_word)	応答形感動詞	「うん」「はい」などの応答系感動詞
(E_word)	感情表出系感動詞	「あっ」「へえ」などの感情表出系感動詞
(L_word)	語彙的応答	「なるほど」「確かに」などの慣習化された短い表現
(A_word)	評価応答	「すごい」「こわ」などの短い表現で評価する語
(P_word)	繰り返し	直前の相手発話を繰り返したもの
(C_word)	共同補完	相手発話に後続するであろう言葉を補ったもの
(T_phon word)	言いさし (意図された語が同定可能)	言いさしたもののうち同定可能なもの。記号中の phon は実際の音列、word は意図された語を表す。
(D_phon)	言いさし (意図された語が不明)	言いさしたもののうち不明なもの
(W_phon word)	いい誤り	いい誤りや発音の怠けなど
(K_kana kanji)	漢字表記できなくなった文字	語中での引き伸ばしやつまりによって漢字表記できなくなった文字
(R_*)	仮名	個人情報保護のために別の名称に置き換える
(歌_*)	歌	歌いながらの発話
<声>	聞き取れない言語音	全く聞き取れないか、言語音と見なせない音声
<笑>	笑い声	発話を伴わない笑い。どれだけ長く笑っていても単に<笑> と書かれる
<息>	呼吸	ため息などの呼吸。
(Y_TEXT)	読み仮名	注釈終了時にその注釈の読み仮名をカタカナで追加
[重複開始位置	複数の話者間で発話が重複している場合は、重複開始位置を「[」で記す。
(笑_)	笑い	笑いながら話している音声
<続>	続き	1 つ前の発話区間にまたがった短単位の後半部分
(迷:)		書き起こせない何かしらの理由

表 アノテーションの種類

構築したアノテーション付き傾聴対話コーパスと、既存の千葉大学3人対話コーパスを用いて、相槌のタイミング予測を行った。予測には音声学的特徴量と言語学的特徴量を用い、深層学習手法により分割された音声区間ごとの二値分類として学習したものをを用いた。

表に記載したように、相槌には異なる種類があり、異なる役割をもつことが知られている。この分類を同様に音声・言語学的特徴量を用いた深層学習手法により実装した。

話し言葉処理の応用として、対話システムの構築に取り組んだ。雑談的な非タスク指向の対話システムにおいては、深層学習手法による統計的なものが主流であるが、「うん」「はい」などのよくある応答のみに終始しがちであるという問題があった。そこで、学習時の損失関数に単語並びの頻度を取り込み、自然な文生成を実現しつつもより多様な応答ができる対話システムを構築した。

もう一つの応用として、会話ゲーム「人狼」を自動プレイするエージェント構築を行った。これは人狼知能プロジェクトの一環である人狼知能大会として毎年実施しているもので、人工的な言語で対戦するプロトコル部門と、日本語・英語で対戦する自然言語部門があり、我々は自然言語部門を主催側としてホストするとともに、参加者としてシステム構築を行った。将棋やチェスなど盤面上にすべての情報が開示される完全情報ゲームとは異なり、人狼は不完全情報ゲームである。この特徴と、人狼が基本的に会話を通じてのみ行われるということが、プレイヤーの自動化という視点でとらえたときに特有の興味深い研究テーマを生み出している。まず、状況判断がプレイヤーの言動に依存するため、多くの局面で本質的に「正解」を知りえず、「推理」が必要である。推理する際にも、他プレイヤーの意図をモデル化するという高度な作業が必要となる。こうした推理やモデルの上に、人狼陣営はいやおうなく「嘘をついて騙す」ことが求められる一方、村人陣営はそれを「見破る」必要がある。これは見方を変えると、いかに他者を「説得」し「信頼を得る」ということでもある。一般に対話システムの評価は「対話が成立しているように見える」という表層的なものになりがちで、結果ある種の逃げに走るほうが良い評価を得られることも多い。人狼知能においては、会話に制約がないと同時にゲームを成立させ勝利に導くという目標があるため、表層的な発話に終始しては内容の一貫したふるまいにならず不自然に映るだろう。そのため適切な項目を設けた主観評価によって、より本質的な要素に高得点をつける評価が可能ではないかと考えている。さまざまな参加チームとの対戦を通じて、より多様な文脈での応答が必要となり、話し言葉処理の性能向上につながる取り組みである。

人狼知能大会自然言語部門の 2018 年は国内最大のゲームカンファレンスである CEDEC にて、2019 年には国際会議 INLG (International Conference of Language Generation) 併設の国際ワークショップ AIWolfDial 2019 (The 1st International Workshop of AI Werewolf and Dialog System) として自然言語部門のみで独立して実施、2020 年は国際会議 IJCAI 2020 (29th International Joint Conference on Artificial Intelligence) 併設の ANAC 2020 (11th Automated Negotiating Agents Competition) のコンテストの一つとして実施した。評価は審査員による主観評価で、

- A 発話表現は自然か
 - B 文脈を踏まえた対話は自然か
 - C 発話内容は一貫しており矛盾がないか
 - D ゲーム行動は対話内容を踏まえているか
 - E 発話表現は豊かか・エージェントごとに一貫して豊かなキャラクター性が出ているか
- の 5 軸を各 5 段階評価するものとした。

参加者としては、人狼ゲームの対戦掲示板の大規模データを用いて深層学習モデルを事前学習し、プレイヤーの発話からの役職推定を行ったものを対話システムに組み込み、対戦エージェントとして実装した。

今後は、作成した傾聴対話音声コーパスのデータ増加やアノテーション拡充、人狼知能大会のリアルタイム対戦実施などを通じて、より人間に近い自然な対話システムの構築を目指したい。

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件／うち国際共著 0件／うちオープンアクセス 0件）

1. 著者名 Minowa Takashi, Kano Yoshinobu	4. 巻 35
2. 論文標題 Middle Expression and Its Converter from Natural Language for Conversation Game “Mafia”	5. 発行年 2020年
3. 雑誌名 Transactions of the Japanese Society for Artificial Intelligence	6. 最初と最後の頁 DSI ~ F_1-13
掲載論文のDOI（デジタルオブジェクト識別子） 10.1527/tjsai.DSI-F	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Sakishita Masahito, Kishimoto Taishiro, Takinami Akiho, Eguchi Yoko, Kano Yoshinobu	4. 巻 843
2. 論文標題 Large-Scale Dialog Corpus Towards Automatic Mental Disease Diagnosis	5. 発行年 2019年
3. 雑誌名 Studies in Computational Intelligence	6. 最初と最後の頁 111 ~ 118
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/978-3-030-24409-5_10	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Sakishita Masahito, Ogawa Chihiro, Tsuchiya Kenji J., Iwabuchi Toshiki, Kishimoto Taishiro, Kano Yoshinobu	4. 巻 843
2. 論文標題 Autism Spectrum Disorder 's Severity Prediction Model Using Utterance Features for Automatic Diagnosis Support	5. 発行年 2019年
3. 雑誌名 Studies in Computational Intelligence	6. 最初と最後の頁 83 ~ 95
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/978-3-030-24409-5_8	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計15件（うち招待講演 2件／うち国際学会 6件）

1. 発表者名 Issei Tsunoda and Yoshinobu Kano
2. 発表標題 AI Werewolf Agent with Reasoning Using Role Patterns and Heuristics
3. 学会等名 The 1st International Workshop of AI Werewolf and Dialog System (AIWolfDial2019), collocated with the INLG 2019 conference (国際学会)
4. 発表年 2019年

1. 発表者名 Yoshinobu Kano, Claus Aranha, Michimasa Inaba, Fujio Toriumi, Hirotaka Osawa, Daisuke Katagami, Takashi Otsuki, Issei Tsunoda, Shoji Nagayama, Dolca Tellois, Yu Sugawara and Yohei Nakata
2. 発表標題 Overview of AIWolfDial 2019 Shared Task: Contest of Automatic Dialog Agents to Play the Werewolf Game through Conversations
3. 学会等名 The 1st International Workshop of AI Werewolf and Dialog System (AIWolfDial2019), INLG 2019 conference (国際学会)
4. 発表年 2019年

1. 発表者名 上山彩夏, 狩野芳伸
2. 発表標題 生成文の文脈に応じ動的に変化する損失関数による多様な対話応答生成
3. 学会等名 言語処理学会第26回年次大会 (2020)
4. 発表年 2020年

1. 発表者名 岩間寛悟, 狩野芳伸
2. 発表標題 再帰的にエンコードを行う階層型Transformerによるマルチターン雑談対話の応答生成
3. 学会等名 言語処理学会第26回年次大会 (2020)
4. 発表年 2020年

1. 発表者名 伊藤毅志, 松原仁, 山本雅人, 狩野芳伸, 大澤博隆
2. 発表標題 パネル討論OS09ゲーム研究の新展開と認知科学
3. 学会等名 日本認知科学会第36回大会 (2019)
4. 発表年 2020年

1. 発表者名 Yoshinobu Kano
2. 発表標題 Japanese Medical NLP: Automatic Diagnosis of Mental Disorders from Dialogs and Processing Electronic Health Records
3. 学会等名 6th Biomedical Linked Annotation Hackathon (BLAH6) (招待講演) (国際学会)
4. 発表年 2020年

1. 発表者名 Yoshinobu Kano
2. 発表標題 Beyond end-to-end learning: Dialog system, sentence generation, and conversation analysis for automatic mental disorder diagnosis
3. 学会等名 31st Conference on Computational Linguistics and Speech Processing (ROCLING 2019) (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 島田渉平, 狩野芳伸
2. 発表標題 N-bestのスコアと文字列類似度を素性に用いた音声認識結果における誤り検出の検討
3. 学会等名 言語処理学会第25回年次大会
4. 発表年 2019年

1. 発表者名 崎下雅仁, 小川ちひろ, 土屋賢治, 岩淵俊樹, 岸本泰士郎, 狩野芳伸
2. 発表標題 自閉スペクトラム症の発話者コーパス作成と自動診断支援システム構築
3. 学会等名 言語処理学会第26回年次大会
4. 発表年 2019年

1. 発表者名 梅村奏子, 狩野芳伸
2. 発表標題 類似単語を元にしたメタファー生成の検討
3. 学会等名 メタファー研究会
4. 発表年 2018年

1. 発表者名 服部充, 狩野芳伸
2. 発表標題 対話音声合成のためのCRFによる日本語東京方言のアクセント結合推定
3. 学会等名 第125回音声言語情報処理研究会 (IPSJ-SLP)
4. 発表年 2018年

1. 発表者名 福野将人, 狩野芳伸
2. 発表標題 音響的特徴を用いた応答の使い分け・挿入を行う傾聴対話システムの試作
3. 学会等名 第9回対話システムシンポジウム(言語・音声理解と対話処理研究会)
4. 発表年 2018年

1. 発表者名 黒田和矢, 狩野芳伸
2. 発表標題 単語の分散表現を用いた相槌生成タイミングの予測
3. 学会等名 第10回対話システムシンポジウム(言語・音声理解と対話処理研究会)
4. 発表年 2018年

1. 発表者名 M. Sakishita, T. Kishimoto, A. Takinami, Y. Eguchi, Y. Kano.
2. 発表標題 Large-scale Dialog Corpus Towards Automatic Mental Disease Diagnosis
3. 学会等名 International Workshop on Health Intelligence (W3PHAI 2019), AAI-19 (国際学会)
4. 発表年 2019年

1. 発表者名 M. Sakishita, C. Ogawa, K. J. Tsuchiya, T. Iwabuchi, T. Kishimoto, Y. Kano.
2. 発表標題 Autism Spectrum Disorder 's Severity Prediction Model Using Utterance Features for Automatic Diagnosis Support
3. 学会等名 International Workshop on Health Intelligence (W3PHAI 2019), AAI-20 (国際学会)
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関