

## 対話のふるまい制御によるキャラクタ表現モデルと対話コーパスによる検証

山本 賢太 井上 昂治 中村 静 高梨 克也 河原 達也

京都大学 大学院情報学研究科 知能情報学専攻

## 1. はじめに

我々は人と自然なインタラクションを行う自律型アンドロイドの研究開発を進めている [1]. 対象とする対話は特定の社会的役割が与えられた状況を想定している. 社会的役割に沿った自然な対話を実現するためには, その役割に適したアンドロイドのキャラクタを表現することが重要である. これまでのキャラクタ表現方法として, 発話文のパターンを変更する手法が検討されてきた [2-4]. 一方で, 音声対話でみられる対話的ふるまいも制御する必要がある. そこで, 我々は発話量, 相槌, フィラー, 交替潜時に着目し, キャラクタが与えられたときに, それを表現するようにこれらのふるまいを制御するモデルを提案した [5]. 本稿では, このモデルで表現されるキャラクタの妥当性について, 実際の対話コーパスに出現する上記のふるまいとキャラクタとの関係性から検証する.

## 2. キャラクタ表現モデル

提案モデルでは, キャラクタ特性として, 外向性, 情緒不安定性, 丁寧さの3つを用いた. 外向性と情緒不安定性は, 心理学における性格分類のビッグ・ファイブ [6] を参考にした. 丁寧さは, アンドロイドを実社会で利用する際に重要になると考えられる [7]. キャラクタ表現に用いる特徴量は, 発話量, 相槌の頻度, フィラーの頻度, 交替潜時の長さの4種類である. これらの特徴量は, 先行研究 [8,9] を参考にして, 対話の印象に影響すると考えられるものを選んだ.

これらの特徴量の妥当性を検証するために被験者実験を実施し, 各特徴量と各キャラクタ特性の印象との関係性を調査した. 実験では, 各特徴量を調整した対話音声をもとに46人の被験者が聴取し, 話者の各キャラクタ特性について印象評定を行った. 結果から, 各特徴量と各キャラクタ特性のいくつかの組合せにおいて相関することが確認された. 例えば, 発話量と相槌が多く, 交替潜時が短い場合には, 外向性が高い傾向にあった.

上記の結果に基づき, キャラクタ特性を与えたときに, それに対応する各ふるまい特徴量の制御量を出力するモデルを構築した (図1). 具体的には, 3つのキャラクタ特性をそれぞれ7段階 (1~7) で入力し, 特徴量毎にロジスティック回帰モデルを用いて, キャラクタを表現するための制御量 (0~1) を出力する. 評価実験では, このモデルを用いてキャラクタを表現した音声対話に対して, キャラクタの印象評定を被験者に行ってもらった. その結果, 外向性と丁寧さについては, 与えたキャラクタと評定結果の間に相関が確認された [5].

## 3. 対話コーパスを用いた分析

前節で述べた実験では, 対話のふるまいは人工的なものであったため, 提案モデルの有効性を対話コーパスを用いて検証する. 使用する対話コーパスは複数の対話タスクを含み, タスク毎にそれに適したキャラクタが異な

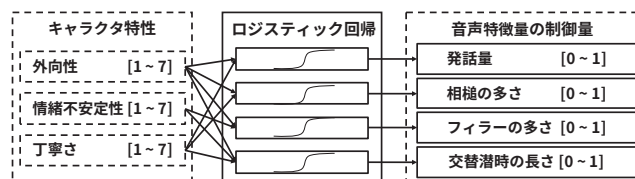


図1: キャラクタ特性に基づくふるまい制御モデル

る. そこで, 提案モデルを逆向きに用いることで, 対話コーパスで実際に観察されたふるまいから, それに対応するキャラクタを計算する. そして, 計算されるキャラクタの傾向を対話タスク毎に調べることで, そのタスクに適したキャラクタを提案モデルが表現できているかどうかを確認する.

## 3.1 対話コーパス

アンドロイド ERICA を用いた対話コーパス [1] を用いた. この対話は遠隔操作された ERICA と被験者との一対一対話である. ERICA の操作は別室のオペレータによって行われており, オペレータが話した音声をそのまま ERICA のスピーカから再生している. コーパス全体を通して, オペレータは4人である. 各対話は約10分であり, 被験者はアンドロイドが遠隔操作されていることは知らされていない. 対話タスクは, お見合い, 就職面接, 傾聴の3種類である. ERICA の役割は, お見合いでは練習相手, 就職面接では面接官, 傾聴では聞き手である. 対話コーパスには, 相槌, フィラー, ターンなどのアノテーションが付与されている. 分析に用いたセッション数は, お見合いが33, 就職面接が31, 傾聴が19である.

## 3.2 分析方法

分析の単位として各対話を2分程度に分割したものをを用いた. 対話全体を1つの単位として用いた場合には, サンプル数が少なくなってしまう. また, 2分程度であれば, 複数のターン交替が含まれ, 相槌やフィラーなどのふるまいも対話内容に依存せず十分に観測され得ると考えられる. 各データに対して, 被験者とオペレータそれぞれで発話時間, 相槌の数, フィラーの数, 交替潜時の長さの特徴量を算出した. 発話量については多い少ないで中央値を用いて二値化した. 相槌の数は対話相手の間休止単位 (IPU) 数, フィラーの数は発話者の IPU 数でそれぞれ正規化した. 前節で述べたキャラクタ表現モデルにキャラクタ特性のすべての組合せ ( $7^3 = 343$ 通り) を入力し, 出力された特徴量と, 各データで計算した上記の制御量とを比較し, 二乗距離が最小のキャラクタ特性をそのデータに割り当てた.

## 3.3 分析結果

算出したキャラクタ特性の分布についてタスク毎に分析を行った. 後述の傾聴対話を除き, オペレータと被験者で同様の傾向がみられたため, ここでは両者を区別せずに扱った.

A character expression model controlling dialogue behaviors and model validation with a dialogue corpus : Kenta Yamamoto, Koji Inoue, Shizuka Nakamura, Katsuya Takanashi, Tatsuya Kawahara (Kyoto Univ.)

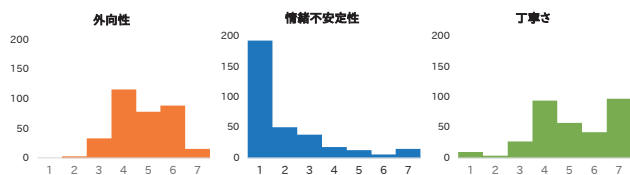


図 2: お見合い対話のキャラクタ分布

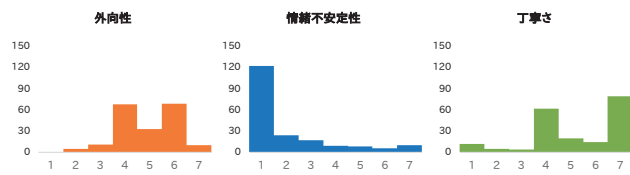


図 4: 傾聴のキャラクタ分布

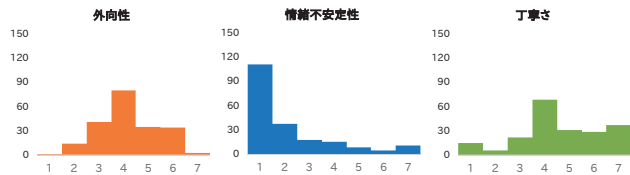


図 3: 就職面接のキャラクタ分布

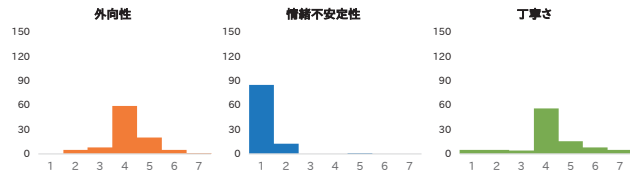
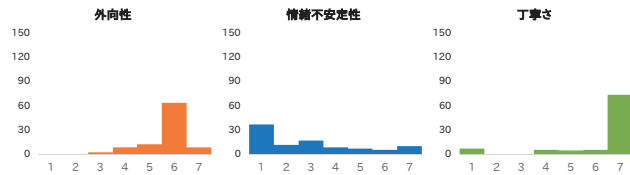


図 5: 傾聴のキャラクタ分布 (上: オペレータ=聞き手, 下: 被験者=話し手)

### 3.3.1 お見合い

お見合いでの各キャラクタ特性の分布を図 2 に示す。情緒不安定性については、情緒安定なキャラクタが大部分を占めた。外向性は 4 から 6、丁寧さは 4 から 7 にかけて分布しており、どちらも高い傾向にあった。お見合いの場面では、お互いに初対面であるため丁寧になり、また積極的に話そうとするため、外向的になると考えられる。したがって、お見合いのタスクに適したキャラクタが表現されているといえる。

### 3.3.2 就職面接

就職面接での各キャラクタ特性の分布を図 3 に示す。情緒不安定性と丁寧さの傾向はお見合いの場合と同様であった。外向性はお見合いの場合よりも中央に集中していた。お見合いに比べて面接はフォーマルな対話であるため、外向性が抑えられたと考えられる。ふるまいの傾向を分析すると、面接では相槌の頻度が低下し、逆にフィラーが増加する傾向にあり、これが外向性の分布に影響したとみられる。

### 3.3.3 傾聴

傾聴での各キャラクタ特性の分布を図 4 に示す。情緒不安定性は他のタスクと同様に情緒安定に集中していた。外向性と丁寧さにはそれぞれ二つのピークがみられた。傾聴に関しては、被験者とオペレータで傾向が異なっていたため、これを区別した分布を図 5 に示す。オペレータは聞き手であり、被験者の話を聞き出す必要があることから、外向的かつ丁寧になったと考えられる。一方、被験者は話し手であり、自分のペースで話すことができるため、ニュートラルなキャラクタに集中したと考えられる。

### 3.3.4 全体の傾向

外向性はタスクに限らず、低いものはほとんど見られなかった。また、情緒不安定性は全てのタスクで低い箇所に集中しており、情緒安定な場合が多いことがわかった。丁寧さについては、外向性ほどではないが丁寧さが低いキャラクタは少なめであった。これらは、対話参加者が初対面であることの影響と考えられる。

## 4. おわりに

本稿では、自律型アンドロイドのキャラクタ表現のためのふるまい制御モデルについて述べた。自然な対話におけるモデルの有効性について、対話コーパスを用いた検証を行なった。分析では、外向性と丁寧さについては、対話タスク毎に分布の違いが現れた。また、傾聴で

は話し手と聞き手の役割毎に異なる傾向が見られた。したがって、提案モデルが出力するキャラクタ表現のためのふるまい制御は実際の対話を反映していることが示唆された。今後の課題として、提案モデルをアンドロイド ERICA の対話システム上に実装することが挙げられる。これにより、実際の対話において、指定したキャラクタ特性と対話相手が知覚した ERICA のキャラクタ特性がどの程度一致するかを確認することができる。また、表現するキャラクタ特性を外向性、情緒不安定性、丁寧さだけでなく、各対話タスクに即したものを追加することで、より実用的にすることも検討している。

謝辞 本研究は JST ERATO JPMJER1401 の支援を受けた。

## 参考文献

- [1] Tatsuya Kawahara. Spoken dialogue system for a human-like conversational robot ERICA. *IWSDS*, 2018.
- [2] Francois Mairesse *et al.* Controlling user perceptions of linguistic style: Trainable generation of personality traits. *Computational Linguistics*, 37(3):455-488, 2011.
- [3] 沈睿 ほか. 音声生成を前提としたテキストレベルでのキャラクタ付与. *情報処理学会論文誌*, 53(4):1269-1276, 2012.
- [4] 宮崎千明 ほか. 文節機能部の確率的書き換えによるキャラクタ性変換. *言語処理学会 第 21 回年次大会 発表論文集*, 277-280, 2015.
- [5] 山本賢太 ほか. 人間型ロボットのキャラクタ表現のための対話の振る舞い制御モデル. *人工知能学会論文誌*, 33(5):C-137-1-9, 2018.
- [6] 和田さゆり. 性格特性用語を用いた Big Five 尺度の作成. *心理学研究*, 67(1):61-67, 1996.
- [7] Swati Gupta *et al.* How Rude Are You?: Evaluating Politeness and Affect in Interaction. *ACII*, 203-217, 2007.
- [8] Etienne D. Sevin *et al.* Influence of personality traits on backchannel selection. *IVA 2010 LNAI 6356*, 187-193, 2010.
- [9] 長岡千賀 ほか. 音声対話における交替潜在時が対人認知に及ぼす影響. *ヒューマンインタフェースシンポジウム 2002 論文集*, 171-174, 2002.