

ユーザの感情の種類と原因を考慮した対話エージェントの応答モデルの開発

沼田崇志^{†1} 朝康博^{†1} 北垣友博^{†1†2} 橋本剛明^{†3} 唐沢かおり^{†3}

概要：近年の人工知能（AI）に関する技術の発展に伴い、新たな価値を持つサービスやシステムの提案が期待されている。AI を活用し快適で豊かな生活を送るためには、AI のインタフェースとして、ユーザの感情に合わせて適切に応答し、ポジティブな気分を誘起することで、ユーザと信頼関係を構築できる対話エージェントが有用となる。本研究では、特定の感情とその原因を示すエピソードを呈示し、その際の対話エージェントの応答による主観的気分を評価することで、感情の種類と原因による最適な応答を抽出した。その結果、ユーザが喜び、悲しみまたは驚きを感じた場合は、原因に関係なく感情表現の模倣が最適であった。一方、ユーザ自身や対話エージェントに向けて怒りを感じた場合は、感情表現の模倣ではなく、悲しみの感情表現が最適であった。これらの結果から、感情の種類と原因を考慮して応答する対話エージェントを開発することで、よりユーザのポジティブな気分を誘起し得ることが分かった。

1. はじめに

1.1 背景

近年、人工知能（AI; Artificial Intelligence）に関する様々な技術が発展しており、AI を用いて新たな価値をもつサービスやシステムが提供されることで、より快適で便利な生活となることが期待されている[1]. AI を用いたサービスやシステムは、インタフェースを介して利用されるため、AI のインタフェースが重要な役割を担う。そのインタフェースの一つとして、ヒトと対話できる機能を持つエージェント（対話エージェント）が注目されており、様々な対話エージェントが開発されている[2]-[4]. 対話エージェントは、AI による新たな価値をスムーズに提供するだけでなく、対話を通じてヒトと信頼関係を構築し、ヒトがポジティブに暮らせるように支援し得るため、有用性が高い。

信頼関係を構築する方法として、ヒト同士では相手の行動の模倣により、共感を誘起するアプローチが有用である[2]. ただ、先行研究では、対話エージェントがヒトの行動を模倣しても、必ずしもヒトの共感を誘起できていなかった[3][4]. そこで我々は、単純な顔や身体の動きではなく、感情表現の模倣が共感誘起に有用な要因と考えて、感情表現を模倣する対話エージェントを開発してきた[5][6]. そして、対話エージェントがポジティブな感情表現を模倣することで、ポジティブな気分へ導けることを示している[7].

しかし、ネガティブな感情についても、感情表現の模倣により共感を誘起し、ポジティブな気分へと導くアプローチが最適であるかは明らかでない。ヒト同士の対話において、模倣そのものにポジティブな気分を誘起する効果があること[2][8]や、ネガティブな表情も模倣する傾向があること[9]は知られている。一方で、ネガティブな感情の発露は、

相手のネガティブな感情を誘起する可能性があること[10][11]も報告されている。対話エージェントは、ヒトの様々な感情に対してポジティブな気分を誘起することが期待されるため、様々な種類の感情に対する対話エージェントの適切な応答モデルを開発することは有用である。

また、同じ感情が生起されても、感情を引き起こす原因となった存在（感情の原因）が異なると注意状態が異なる[12]など、感情の原因により外部刺激の影響が異なることが知られている。そのため、対話エージェントの応答も、感情の原因により異なる影響を及ぼす可能性があり、感情の原因を考慮することで、より適切に対話エージェントが応答できると考えた。感情の原因については、対話を想定し、自分、相手（対話エージェント）、第三者の三種類に区別して評価する必要がある。

1.2 目的

本研究では、ユーザの感情の種類と原因による対話エージェントの最適な応答を抽出し、ユーザの感情の種類と原因を考慮した対話エージェントの応答モデルを開発することを目的とした。

具体的には、ユーザの感情として喜び、悲しみ、怒り、驚きの4種類、その感情の原因として自分、相手、第三者の3種類、対話エージェントの応答として喜び、悲しみ、怒り、驚きの4種類を設定した上で、その中の特定の感情とその原因を示すエピソードを呈示し、そのエピソードにおける対話エージェントの応答を見た際の主観的気分を回答する課題を課した。そして、応答による主観的気分を比較することで、対話エージェントの最適な応答を抽出した。

†1 (株)日立製作所 研究開発グループ 基礎研究センタ

†2 東京大学 文学部

†3 東京大学 人文社会系研究科

2. 対話エージェント

本研究では、対話エージェントとして、ヒヨコ型の対話エージェント「ピョタ」を用いた[5]-[7] (図 1)。ピョタは、デフォルメを用いて、大きな目と小さな鼻と口により好ましいと感じやすい外見[13]をもち、Plutchik の感情モデル[14]に基づいて開発した表現 24 種 (8 相の感情と各 3 段階の感情レベルの組み合わせ) などの様々な表現を有する。



図 1 対話エージェント「ピョタ」

3. 対話エージェントの表現と感情の関係評価

3.1 評価方法

対話エージェントの表現は、特定の感情を示すように意図して開発したが、実際にユーザが読み取る感情には個人差が存在することが考えられる。個人差が大きい表現を用いると、対話エージェントの応答に対する解釈が複雑になり、汎用性の高い対話エージェントの応答モデルを開発することが困難となる。そこで本研究では、対話エージェントの表現のうち、特定の感情を示していることが明らかな表現を抽出し、それらの感情表現を用いて評価を行った。まず、対話エージェントの表現のうち、特定の感情を示す表現を抽出した方法について述べる。

被験者は、健常な男女 124 名 (平均年齢 28.3±11.8 歳) とした。対話エージェントの表現に対して、それぞれ感情 7 種類 (喜び、悲しみ、怒り、嫌悪、驚き、恐れ、同情) を用意し、対話エージェントの表現に対して最も近いと思う感情を選択するように指示した。対話エージェントの表現の呈示および感情の選択は、質問紙により実施した (図 2)。

本実験のデータは、東京大学文学部の実験倫理委員会の承認を得るとともに、(株) 日立製作所研究開発グループで定める倫理審査基準に則り審査し、取得した。

記入例) キャラクタの表情を見て、キャラクタは「悲しみ」の感情になっていると思ったら、下のように該当する場所に○をつけてください。

表情	喜び	悲しみ	怒り	嫌悪	驚き	恐れ	同情
		○					

図 2 対話エージェントの表現と感情の関係評価における質問紙の一例

3.2 評価結果および考察

対話エージェントの表現のうち、最も特定の感情が選択された割合が高かった表現を感情別に抽出した結果、喜び、悲しみ、怒り、驚きの 4 種類において、95%以上の被験者が特定の感情を表現していると感じられる表現を抽出できた (図 3)。一方、嫌悪、恐れ、同情については、多くの被験者が特定の感情を表現していると感じられる表現を抽出できず、対話エージェントの表現から推定される感情の個人差が大きいと判断した。そのため、本研究では喜び、悲しみ、怒り、驚きの 4 種類の感情に限定して、対話エージェントの応答について評価した。

感情の種類	喜び	悲しみ	怒り	嫌悪
対話エージェントの表現				
一致率	96.0%	98.4%	95.2%	83.1%
感情の種類	驚き	恐れ	同情	
対話エージェントの表現				
一致率	95.2%	71.8%	46.0%	

図 3 各感情における最も特定の感情が選択された割合が高かった表現とその割合

4. 対話エージェントの応答評価

4.1 評価方法

ユーザの感情の種類と原因による対話エージェントの適切な応答を抽出するため、特定の感情とその原因を示すエピソード (感情エピソード) を呈示し、その感情エピソードにおける対話エージェントの応答を見た際の主観的気分を回答する課題を課した。

被験者は、健常な男女 145 名 (平均年齢 19.8±1.1 歳) とした。課題では、ユーザの 4 種類の感情 (喜び、悲しみ、怒り、驚き) と 3 種類の原因 (自分、相手、第三者) を組み合わせた 12 種類の感情エピソードについて、対話エージェントの 4 種類の応答 (喜び、悲しみ、怒り、驚き) を見た際の主観的気分を取得した。感情エピソードの呈示および主観的気分の回答は、質問紙により実施した (図 4)。

感情エピソードは、各種類の感情につき 4 つ作成した。そして、それぞれの感情エピソードに原因を示す情報を追加し、特定の感情の種類と原因を示すエピソードにした。これにより、計 48 つ (ヒトの 4 種類の感情×4 エピソード×3 種類の感情の原因) の感情エピソードを作成した。尚、原因間で感情エピソードは共通しているため、各被験者に

対して特定の原因の感情エピソードのみ呈示することにして、被験者ごとに異なる原因の感情エピソードを呈示した。

主観的気分の計測は、VAS (Visual Analogue Scale) [15]を用いて行った。本研究における VAS は、70mm の水平線について、左端を「とても悪い気分」、右端を「とても良い気分」として、自身の気分に相当する位置に垂線を描くように指示した。

感情エピソードにおける対話エージェントの応答を見た際の主観的気分については、統計解析により評価を行った。具体的には、それぞれのユーザの感情について、3種類の感情の原因(自分、相手、第三者)と4種類の対話エージェントの応答(喜び、悲しみ、怒り、驚き)の2要因12条件について比較した。まず、取得した主観的気分のVASデータに対して最大3.5点～最小-3.5点の点数を付与した。次に、各被験者の各感情の種類の感情エピソード間の主観的気分の平均値を算出した。そして、2要因分散分析(two-way ANOVA; two-way measures analysis of variance)を用いて、感情の原因および対話エージェントの応答による主観的気分を比較した。その後、Bonferroni補正により多重比較補正を行った上で、感情の原因間については対応のないt検定、対話エージェントの応答間については対応のあるt検定により比較した。交互作用が有意であれば原因間および応答間の組み合わせ、有意でなければ原因間と応答間それぞれで比較を行った。尚、統計解析の有意水準(p値)は0.05とした。

対話エージェントの表現と感情の関係評価と同様に、本実験のデータは、東京大学文学部の実験倫理委員会の承認を得るとともに、(株)日立製作所研究開発グループで定める倫理審査基準に則り審査した上で、取得した。

あなたがピヨタに表出した言葉・感情：あなたはピヨタに次の言葉を言いながら「怒り」感情を表出しました。

あなた「明日は早起きだから早めに寝たいのに、目が冴えて全然眠れない！すっごくイライラする！」

ピヨタが返した表情	設問
	(1)ピヨタはあなたの感情を理解してくれましたか？ 全くそう思わない どちらでもない 強くそう思う
	(2)ピヨタの返した表情はあなたにとって適切でしたか？ 全くそう思わない どちらでもない 強くそう思う
	(3)ピヨタの返した表情を見てどんな気分になりましたか？ とても悪い気分 どちらでもない とても良い気分

図4 対話エージェントの応答評価における質問紙の一例

4.2 評価結果および考察

ユーザの感情の種類に対する感情の原因と対話エージェントの応答による主観的気分を比較した結果(図5)、被験者が喜び、悲しみ、驚きの感情をもった場合は、感情の原因に拘わらず、対話エージェントが被験者の感情を模倣す

赤字: 対話エージェントの応答による比較結果

青字: 感情の原因による比較結果

*: p<0.05, **: p<0.01, NS: 有意差なし

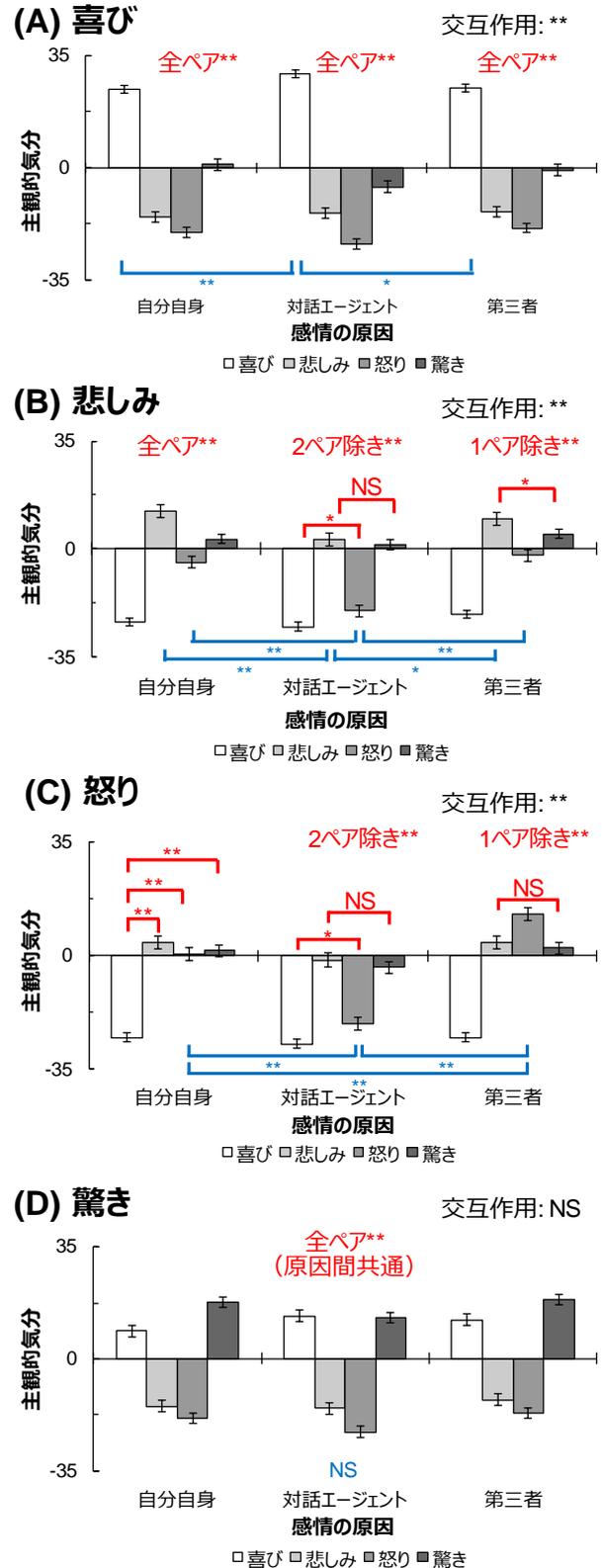


図5 ユーザの感情の種類に対する感情の原因と対話エージェントの応答による主観的気分の比較結果

るように感情表現（感情表現を模倣）することで、最もポジティブな気分を示した。一方、被験者が怒りの感情をもった場合のうち、怒りの感情が第三者に因る場合は感情表現を模倣することで最もポジティブな気分を示したが、怒りの感情が自分または相手に因る場合は、悲しみの感情表現により最もポジティブな気分を示した。これらを踏まえると、感情の種類および原因によっては、必ずしも感情表現の模倣が最適ではないことが示唆された。ネガティブな感情においても、感情の種類と原因により感情表現の模倣が最適かどうかは異なっていたため、被験者の感情の種類と原因に応じて適切な感情を表現することで、対話エージェントの応答によりポジティブな気分を誘起できることが分かった。

これらの結果から、対話エージェントの応答について、ユーザが喜び、悲しみ、驚きを感じた場合全てと、ユーザが怒りを感じた場合のうち、その原因が第三者であった場合は感情表現の模倣が最適であり、ユーザが怒りを感じた場合のうち、その原因が自分や対話エージェントであった場合は、悲しみの感情表現で応答することが最適であることが分かった。これらをまとめて、ユーザの感情の種類と原因を考慮した対話エージェントの応答モデルを開発できた（表 1）。

表 1 ユーザの感情の種類と原因を考慮した対話エージェントの応答モデル

対話エージェントの最適な応答		ユーザの感情の種類			
		喜び	悲しみ	怒り	驚き
ユーザの感情の原因	ユーザ自身	模倣 (喜び)	模倣 (悲しみ)	悲しみ	模倣 (驚き)
	対話エージェント			模倣 (怒り)	
	第三者				

本研究の結果において、感情表現の模倣が最適でなかった怒りの感情についても、ヒト同士では表情を模倣することが報告されている[9][16]。そのため、ヒト同士の対話で見られる現象をそのまま対話エージェントの機能として実装することは、必ずしも最適ではないことが示唆された。また、相手の気分をポジティブに導くという観点では、自然に表情を模倣してしまうヒトよりも、感情表現を任意に選択できる対話エージェントの方が適切に応答し、ポジティブな気分を誘起し得る可能性を示したと考えられる。

5. おわりに

本研究では、ユーザの特定の感情とその原因を示すエピソードを呈示し、その感情の原因および対話エージェントの応答による主観的気分を比較した。その結果、対話エージェントによる感情表現の模倣は必ずしも最適な応答では

なく、ユーザの感情の種類と原因に応じて適切な感情を表現することで、対話エージェントの応答によりポジティブな気分を誘起できることが示唆された。

今後は、本研究で提案した、ユーザの感情の種類と原因を考慮した応答モデルを対話エージェントに実装することで、適切に感情を表現する対話エージェントの開発をめざす。

参考文献

- [1] 内閣府. 第 5 期科学技術基本計画. 2016
- [2] Chartrand, T. L. and Bargh J. A.. The Chameleon Effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, 1999, vol. 76, no.6, pp.893-910
- [3] Gratch, J. Wang, N. Gerten, J. Fast E. and Duffy, R.. Creating Rapport with Virtual Agents. *International Conference on Intelligent Virtual Agents*, 2007, pp.125-138
- [4] Wang, N. and Gratch, J. Can virtual human build rapport and promote learning? *The 14th International Conference on Artificial Intelligence in Education*, 2009, pp.737-744
- [5] “株式会社日立製作所 人に言寄り添う AI の実現に向け、感情表現を模倣する技術を開発”. <http://www.hitachi.co.jp/rd/news/topics/2018/1106.html>, (参照 2018-12-18)
- [6] 沼田崇志, 朝康博, 牧敦. 感情表現を模倣する対話エージェントの開発とその脳活動計測による擬人化効果の実証. *日本バーチャルリアリティ学会研究報告 サイバースペースと仮想都市研究会*, 2018, CSVC2018-25
- [7] Sato, H. Numata, T. Asa, Y. Koike, T. Miyata, K. Nakagawa, E. Sumiya, M. Sadato, N. An fMRI investigation on the positive consequences of being imitated by a virtual non-human agent. *Neuroscience*, 2018, 110
- [8] Catmur, C. and Heyes, C. Is it what you do, or when you do it? The roles of contingency and similarity in pro-social effects of imitation. *Cognitive Science*, 2013, vol. 37, pp.1541-1552
- [9] Lundqvist, L-O. Facial EMG reactions to facial expressions: A case of facial emotional contagion? *Scandinavian Journal of Psychology*, 1995, vol. 36, pp.130-141
- [10] Friedman, R. Anderson, C. Brett, J. Olekalns, M. Goates, N. and Lisco, C. C. The positive and negative effects of anger on dispute resolution: evidence from electronically mediated disputes. *Journal of Applied Psychology*, 2004, vol. 89, no. 2, pp.369-376
- [11] Kulesza, W. M. Cistak, A. Vallacher, R. R. Nowak, A. Czekiel, M. and Bedynska, S. The face of the Chameleon: The experience of facial mimicry for the mimicker and the mimickee. *The Journal of Social Psychology*, 2015, vol. 155, no. 6, pp.590-604
- [12] Imbir, K. Origin and source of emotion as factors that modulate the scope of attention. *Roczniki Psychologiczne / Annals of Psychology*, 2013, vol. 14, no. 2, pp.287-310
- [13] 平家雅之, 川崎裕達, 田中貴紘, 藤田欣也. 類似度と好感度を両立する個人化アバタのデフォルメ法の検討. *ヒューマンインタフェース学会論文誌*, 2011, vol.13, no.3, pp.243-254
- [14] Plutchik, R. The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American Scientist*, 2001, vol.89, no.4, pp.344-350
- [15] Torrance, G. W. Feeny, D. and Furlong, W. Visual Analog Scale: Do they have a role in the measurement of preferences for health states? *Medical Decision Making*, 2001, vol. 21, pp.329-334.
- [16] 市川寛子, 牧野順四郎. 刺激表情に対する観察者の同調的表情. *心理学研究*, 2004, vol. 75, no. 2, pp.142-147