

対話意欲を喚起する価値観肯定・否定割合に基づく 自律対話ロボットの対話戦略

A Balance of Agreement/Disagreement Ratio to Build Users' Motivation for Conversation with Autonomous Robots

内田 貴久 *1*2
Takahisa Uchida

港 隆史 *2
Takashi Minato

石黒 浩 *1*2
Hiroshi Ishiguro

*1 大阪大学大学院基礎工学研究科
Osaka University

*2 株式会社 国際電気通信基礎技術研究所
Advanced Telecommunications Research Institute International

Existing spoken dialogue systems are insufficient for continuous use because they fail to inspire the user's motivation to talk with them. This is because a user is unable to feel that a robot possesses its own intentions, and thus the robot must acquire its own values to convey its sense of intentionality to users through its spoken communications. This paper focuses on a dialogue strategy aimed at building people's motivation under the assumption that the robot has a values-based dialogue system. People's motivation can be influenced by the intentionality as well as by the affinity of the robot. We hypothesized that there is a good disagreement/agreement ratio in a conversation to efficiently balance people's feelings of intentionality and affinity. The result of a psychological experiment using an android robot partially supported our hypothesis.

1. はじめに

昨今、音声対話システムの開発が進み、我々の社会に浸透しつつある。しかしながら、実用化されているシステムでは、ユーザの継続的な対話を行う意欲を引き出すには不十分である。例えば、Appleの音声認識アプリケーションSiri*1はユーザの様々な問いかけに返答することができるが、基本的に質問に答えるだけの一問一答のやり取りに留まっている。一方、複数ターン続くものとしては、バスの交通案内[Adachi 03]などタスク指向の対話の実現されているが、Siriのように様々な問いかけに対応できる訳ではなく、対話内容は非常に限られた範囲内に閉じている。ELIZA [Weizenbaum 66]に代表されるchatbotシステムや大量の対話データベースに基づいて複数ターン続く対話を生成することができるシステム [Higashinaka 14]も存在するが、現状ではユーザの状態に合わせた対話コンテキストに基づいて発話選択することができていない。

上述のように、既存のシステムでは自然な対話を実現できているとは言えず、結果としてユーザの対話意欲が喪失し、長期的な使用がなされないなどの状況に陥っているのが現状である。とりわけ、高齢者の対話支援や子供の道徳・コミュニケーション教育など対話そのものを目的とする場面において、対話システムとして解決されるべき事項である。

対話意欲の喪失を招く原因として、ユーザがシステムの発話に対する意図を帰属できないことが考えられる。人間の意思決定は価値観に基づくものである [Rokeach 73]、つまり、発話意図や理由は固有の価値観から生じるものであることを踏まえると、システムが相互の価値観を考慮して対話を行っていることをユーザが主観的に認識することができれば、システムの発話の意図や理由が帰属され、対話意欲が向上するものと考えられる。そのため、価値観に関するモデルを実装し、発話選択を行うシステムを構築することが必要であるが、本稿ではそれに先立ち、システムが価値観を持っているとユーザが感じるメカニズムについて調査することを目的とする。そこで本研



図 1: Appearance of Android ERICA

究では、「価値観」を「好みや意見の集合」と定義し、システムが価値観に従って発話しているとユーザが感じることが対話意欲の向上につながると考え、そのような対話戦略を構築する。対話においては、相手の意見への同意や非同意を繰り返しながら互いの価値観を認識している [井原 05]と考えられることから、この同意・非同意の割合に関する仮説を立て、心理実験で検証する。また、本研究ではアンドロイド ERICA (図 1)を用いて対話システムを実装する。

2. 仮説

2.1 対話における価値観

ここで対話について様々な分野でどのような捉え方がなされているのかを見てみると、教育学や社会学の分野では、対話において価値観というものを重要視している ([平田 12][永里 10])。さらに工学においても、価値観に依拠したインタラクションデザインを行っているものが存在する [井原 05]。また、価値観はユーザの嗜好や行動決定に強く影響を及ぼすと考えられており、マーケティングや Web インテリジェンスの分野で広く活用されている。Hattoriら [Hattori 12]はユーザの価値観に基づいた情報推薦システムを考案している。価値観を用いる理由として、価値観に基づいてユーザをモデル化することにより、こだわりの強い属性は少ない情報からモデリング可能であることが述べられている [Hattori 13]。すなわち、ユーザの価値観を把握することが情報処理の観点から効率的であることを示している。これらのことから、対話システムへの価値観の導入は対話意欲の向上だけでなく、相互理解や情報伝達の効率化にも寄与するということと言える。

連絡先: 内田 貴久, 大阪大学大学院基礎工学研究科, 大阪府豊中市待兼山町 1-3, 06-6850-6362, uchida.takahisa@irl.sys.es.osaka-u.ac.jp

*1 <http://www.apple.com/ios/siri/>

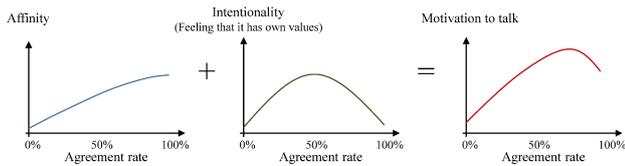


図 2: Concept of Hypothesis

2.2 価値観を感じる対話

対話を大まかにみると、互いの価値観に対する肯定と否定を繰り返しており、その中で、互いの価値観を認識している [井原 05] と考えられる。システムの発言における肯定と否定のバランスは、価値観を持っているかどうかの印象に影響すると考えられる。例えば、相手の発言に対してすべて肯定するようでは、本当に価値観を持っているのかと感じてしまうであろう。この考えに類似する研究が Yamaoka ら [Yamaoka 07] によってなされている。この研究では、ロボットの相手動作模倣率が高くなるほど、好感度は高まるが、自律性は失われるという結果になっている。このように、模倣率によってロボットの印象が変わるという研究結果も存在することから、相手への同調度合いと考えられる相手の発言への肯定率が対話に影響するということは十分に考えられる。

2.3 対話意欲を喚起するための仮説

システムが価値観に従って発話しているという印象、すなわち、システムが自身の価値観を持っており、それに従って発話しているという印象をユーザに与えるために、システムがユーザにどの程度同調（ユーザの発話を肯定）すれば良いかを考える。これに関して、システムの発言が最も予測しにくい場合に、システムが固有の価値観を持っているとユーザを感じる度合いが最も高まると考えられる。情報理論の観点から、最も予測しにくいのは 50% の場合であることを考慮すると、システムが価値観を持って話しているかどうかの印象と肯定割合の関係は図 2 中央のようになると考えられる。一方、システムが自身の主張ばかりしている（ユーザの発言を否定ばかりしている）ようでは、ユーザの対話意欲を損ないかねない。そこで、ユーザの意見に対して理解を示す、すなわち、肯定する度合いが高くなるほど、ユーザに親近感を与えることができると考えられる。そこで、親近感・親和性とユーザの発言に対する肯定割合の関係は、図 2 左のようになると考えられる。対話意欲は、両者を同時に備えている場合に感じられるものであると考えられるため、対話意欲は図 2 右のように、肯定側に偏った肯定割合によって高まるという仮説が考えられる。これらの仮説をまとめると、以下のようになる。

- 対話意欲を引き出す最適な肯定否定のバランスが存在する
- システムが固有の価値観を有しており、かつ、ユーザに対する親和性を兼ね備えていると感じる肯定割合、対話意欲を喚起する肯定割合は、50% でも 100% でもない、その間に存在する

対話はその内容に依存して印象が変化するという考え方が一般的であるが、肯定・否定割合を変化させるだけで対話エージェントの印象を変化させ、対話意欲を向上させることができれば、対話そのものがサービスとして求められる場面での設計パラメータとして有益であると考えられる。



図 3: Video Stimulus

3. 仮説の検証

3.1 実験目的

実験者とアンドロイドが対話を行っている映像を被験者に見せた後、その映像をアンケートで評価させることで、前章で導いた仮説、すなわち、対話相手の発言に対する肯定・否定割合によって親和性（理解してくれる）と意思性（自分の価値観を持っている）に対する評価が変化し、両者を備えている肯定割合において、最も対話意欲が喚起されることを検証することが目的である。

3.2 実験刺激

本実験では図 3 のように画面上にアンドロイド ERICA が現れ、画面には登場しない対話相手（実験者）と対話している様子をビデオに録画し、被験者はこれを見て評価を行う。ビデオ作製にあたり、音声合成として HOYA 株式会社の VOICE TEXT ERICA*2、動作生成として Ishi ら [Ishi 12] が開発した口の動きを音声から生成するシステムと境ら [境 16] が考案した音声から頭部動作を生成するシステムを使用した。対話相手はビデオ上では見えないが、対話相手の声は聞こえるようになっている。対話内容は、ERICA が相手の価値観（好みや意見）を聞き出し、それに対して自分の価値観を述べるというやり取りである。トピックは 8 項目（好きなお菓子、好きな飲み物、都会と田舎、電車と車、映画とドラマ、小説とマンガ、好きな科目、国内旅行と海外旅行）用意し、これらに関して ERICA が相手の好みや意見に対する肯定・否定を繰り返す。以下に、あるトピックにおける対話の例を示す。また、表情生成に関して、質問時は無表情、相手の好みや意見を肯定する場合は微笑み、否定する場合は戸惑う表情を生成した。

ERICA : 「ところで、住むとしたら都会と田舎のどっちが良いですか。」

実験者 : 「都会が良いです。」

ERICA (肯定する場合) : 「確かに。都会って何でも揃って便利ですよ。」

ERICA (否定する場合) : 「うーん。都会はなんだか落ち着かないから嫌だなあ。」

ビデオは 3 種類あり、対話におけるトピックと順番はどれも同じであるが、相手の発言内容に対する肯定率が 50%、75%、100% と異なる。50% では、好きな飲み物、電車と車、小説とマンガ、国内旅行と海外旅行のトピック、75% では、電車と車、

*2 <http://voicetext.jp/>

表 1: 主効果の p 値 * : $p < 0.05$, ** : $p < 0.01$

項目	対話意欲	親和性	意思性
p 値	3.1×10^{-1}	$6.3 \times 10^{-3**}$	$5.3 \times 10^{-6**}$

国内旅行と海外旅行のトピックに関してアンドロイドが対話相手の発言内容を否定するものとした。否定するトピックは被験者間で同一のものとした。肯定・否定によって受ける印象の強さは、被験者の各トピックに対する関心度合いに影響されると考えられるが、関心度合いは個人依存であるため、否定するトピックを固定することで、個人依存の影響を平均化することとした。

3.3 実験手順

まず、各肯定割合の ERICA に名前を付す（例えば、50% : A, 75% : B, 100% : C）。名前と肯定割合との対応はカウンターバランスをとった。そして、被験者は A と B の対話ビデオを見て、それに関するアンケートに答える。これと同様の手順を A と C, B と C の組み合わせに関しても行った。ここで、順序効果を打ち消すため、ある組み合わせにおいて2つのビデオのうちどちらから先に見るかは被験者の自由とし、ビデオの再生回数に制限は設けなかった。

3.4 評価方法

今回の検証実験では相手の発言に対する肯定・否定割合によって印象がどのように変化するかを調べる、すなわち、心理的尺度を構成することが目的であるので、評価方法として一対比較法を採用した。仮説を検証するため、「対話意欲」、「親和性」、「意思性」に関する評価を目的としてアンケートを作成した。実施した質問項目を以下に示す。

1. 自分が対話するなら、
2. 友達になるなら、
3. 自分の意思をもって話しているのは、

上記の質問に対して、1. 確かに A, 2. A, 3. どちらかといえば A, 4. どちらとも言えない, 5. どちらかといえば B, 6. B, 7. 確かに B の中から一つ選択させた。これと同様の質問を A と C, B と C のそれぞれの組み合わせで実施した。そして、最後にインタビューを行い、評価基準や感想を述べるよう求めた。

3.5 実験結果

評価実験に関して、被験者は 19 歳から 26 歳までの 14 名（男性 8 名、女性 6 名）、平均年齢は 22.1 歳（標準偏差 2.0 歳）である。

アンケート結果を解析するにあたり、比較順序を考慮せず、かつ 1 人の被験者がすべての組み合わせを 1 回ずつ比較することから、シェッフェの一対比較法の中屋の変法 [三浦 73] を用いた。この手法により得た主効果の p 値を表 1 に示す。

表 1 より、親和性、意思性の項目において有意水準 1% で主効果が有意であった。有意な主効果が見られた項目についてヤードスティックを用いて多重比較を行うと、親和性に関して 50 - 75% 間で有意水準 1%, 50% - 100% 間で有意水準 5%, 意思性に関して 50% - 75% 間で有意水準 1%, 75% - 100% 間で有意水準 1% で有意差が現れた。得られた主効果の尺度値と、これらの検定結果を合わせてグラフ化したものを図 4~図 6 に示す。

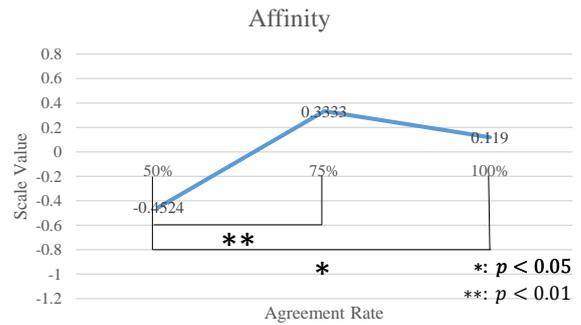


図 4: Scale Value for Affinity

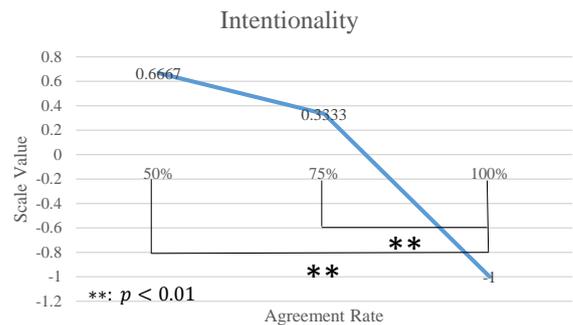


図 5: Scale Value for Intentionality

4. 考察

4.1 親和性・意思性について

まず、親和性と意思性について、親和性の項目に関して図 4 から、50% - 75% 間で有意水準 1%, 50% - 100% 間で有意水準 5% で有意差が現れる結果となった。すなわち、50% より高い割合で被験者が感じる親和性が高まると考えられる。一方、意思性の項目では図 5 より、50% - 75% 間で有意水準 1%, 75% - 100% 間で有意水準 1% で有意差が現れていることから、100% より低い割合で、自分の意思をもって話している、すなわち、固有の価値観を持っているという印象を高めることができると考えられる。以上をまとめると、肯定割合を 50% と 100% の間に設定することが、親和性と意思性の両者を兼ね備える対話戦略として有効であると考えられる。また、親和性と意思性のグラフ（図 4, 図 5）を比較すると、親和性を高める

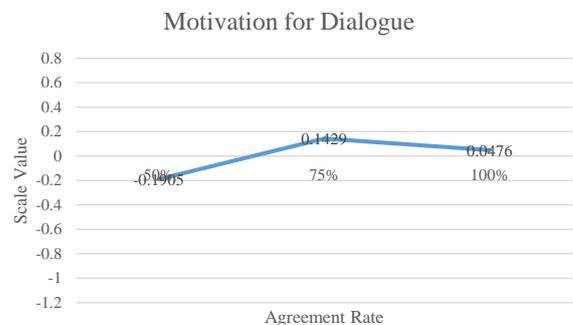


図 6: Scale Value for Motivation of Dialogue

と意思性が減少し、意思性を高めると親和性が減少する、すなわち、トレードオフの関係になっていることが観察される。

4.2 対話意欲について

親和性と意思性の両者を兼ね備えているとユーザが感じる肯定割合は、上記のように50%と100%の間に存在することが確認された。このことから、対話意欲が最も喚起される割合が75%であることが予想されるが、図6を見ると、尺度値としてこの割合が最も高い評価となっているものの、他の割合との有意差が現れるまでには至っていない。この原因として、対話意欲を喚起する要因が個人間で異なることが考えられる。アンケート回答後のインタビューにおいて、どのような基準で対話意欲の項目を評価したかを尋ねたところ、すべて肯定するものとは話したくないと言う被験者と、簡単に否定するようなものとは話したくないという、2つの考えのうちのどちらかに従って回答している傾向にあった。このことは、対話したいかという質問に対して、個人によってその動機が異なっていることを示している。また、今回の実験では、被験者はビデオを通してアンドロイドを評価しており、実際にアンドロイドと対面したときと異なる印象を持った可能性がある。そのため、今後はアンドロイドと対面しての印象評価実験を行う必要がある。また、質問として対話したいかという主観的な評価だけではなく、より客観的な評価指標を設定し、対話意欲のメカニズムを考察することも重要である。さらに、対話開始時では肯定割合を増やしてユーザからの親和性を獲得した後、肯定割合を減らして、意思性を感じさせるといったような肯定・否定割合の動的な変化が対話意欲に対してもたらす効果についても、今後検証する余地がある。

5. 結論

本研究では、対話における相手の発言に対する肯定割合に着目し、システムの価値観を感じさせる意思性と親和性の両者を兼ね備える肯定・否定のバランスが存在し、その割合において対話意欲が最も喚起されるという仮説を立て、検証実験を行った。肯定割合50%以上の条件で調べた結果、親和性を高めるには肯定割合を50%より大きくし、意思性を高めるには100%より小さくするべきであるという知見を得た。さらに、親和性と意思性は価値観の肯定・否定割合に関してトレードオフの関係にあることを発見した。このことから、両者を同時に表現するというもののみならず、親和的印象を与えた後、意思性を獲得するといったように、価値観の肯定・否定割合をフェーズごとに動的に変化させるダイナミクスが対話意欲を高めるのに効果的であるのではないかという新たな仮説を得ることができた。一方、上記で得られた肯定割合において対話意欲が最も喚起されるという結果を予想したが、各個人の対話意欲に対する動機が異なることから、肯定・否定割合に依存した変化を確かめることはできなかった。これに関して、実際に被験者がアンドロイドと対面したインタラクションを行っていないことも、原因として考えられる。今後は、上記の問題を考慮した実験のデザイン、さらに、対話意欲に関してより客観的な評価指標を設定することが課題となる。

参考文献

[Adachi 03] Adachi, K., Fumihiko, K., Ueno, S., Kawahara, T., and Okuno, H. G.: Flexible spoken dialogue system based on user models and dynamic generation of

VoiceXML scripts, in *Proc. of the 4th SIGdial Workshop on Discourse and Dialogue*, pp. 87–96 (2003)

[Weizenbaum 66] Weizenbaum, J.: ELIZA — a computer program for the study of natural language communication between man and machine, *Communications of the ACM*, Vol. 9, No. 1, pp. 36–45 (1966)

[Higashinaka 14] Higashinaka, R., Kobayashi, N., Hirano, T., Miyazaki, C., Meguro, T., Makino, T., and Matsuo, Y.: Syntactic filtering and content-based retrieval of Twitter sentences for the generation of system utterances in dialogue systems, in *Proc. of the International Workshop Series on Spoken Dialogue Systems Technology*, pp. 113–123 (2014)

[Rokeach 73] Rokeach, M.: *The nature of human values*, Free press, New York (1973)

[井原 05] 井原雅行, 小林稔: 価値観を用いたコミュニケーション支援のための発言開始語句分析, 情報処理学会論文誌, Vol. 46, No. 1, pp. 138–146 (2005)

[平田 12] 平田オリザ: わかりあえないことから コミュニケーション能力とは何か, 講談社 (2012)

[永里 10] 永里智広, 假屋園昭彦: 道德授業の対話学習における自己内対話の推移分析: 対話プロセスの実相と完全応答型対話形式, 鹿児島大学教育学部教育実践研究紀要, Vol. 20, pp. 121–140 (2010)

[Hattori 12] Hattori, S. and Takama, Y.: Investigation About Applicability of Personal Values for Recommender System, *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol. 16, No. 3, pp. 404–411 (2012)

[Hattori 13] Hattori, S. and Takama, Y.: Proposal of user modeling method employing reputation analysis on user reviews based on personal values, in *Proc. of the 27th Annual Conference of Japanese Society for Artificial Intelligence*, pp. 1A3–IOS–3a–4 (2013)

[Yamaoka 07] Yamaoka, F., Kanda, T., Ishiguro, H., and Hagita, N.: How contingent should a lifelike robot be? The relationship between contingency and complexity, *Connection Science*, Vol. 19, No. 2, pp. 143–162 (2007)

[Ishi 12] Ishi, C. T., Liu, C., Ishiguro, H., and Hagita, N.: Evaluation of formant-based lip motion generation in tele-operated humanoid robots, in *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2377–2382 IEEE (2012)

[境 16] 境くりま, 港隆史, 石黒浩: 音声に対応する頭部動作のオンライン生成システムと遠隔操作における効果, 電子情報通信学会論文誌, Vol. 99, No. 1, pp. 14–24 (2016)

[三浦 73] 三浦新他: 官能検査ハンドブック, 日技科連, Vol. 1, pp. 249–252 (1973)