

対話エージェントを用いた認知行動療法の有効性検証

Evaluation of Cognitive Behavior Therapy with Embodied Conversational Agents

設楽 一碩*¹ 田中 宏季*¹ 足立 浩祥*² 金山 大祐*² 阪上 由香子*² 工藤 喬*²
Kazuhiro Shidara Hiroki Tanaka Hiroyoshi Adachi Daisuke Kanayama Yukako Sakagami Takashi Kudo

中村 哲*¹
Satoshi Nakamura

*¹奈良先端科学技術大学院大学
Nara Institute of Science and Technology

*²大阪大学 キャンパスライフ健康支援センター
Health and Counseling Center, Osaka University

Cognitive Behavior Therapy (CBT) is a training method that has been established to correct human cognitive distortions. We aim to achieve low cost and highly effective automation by an embodied conversational agent to recreate the human therapists' CBT. As a first step, we implemented dialogue systems with fixed questions based on cognitive restructuring. We prepared two types of dialogues: questions presented by text or spoken dialogue agents. The depression scales and changes in emotions during the session (effects of training) were measured from 28 graduate students and undergraduate students. As a result, users' emotions significantly changed between pre- and post- sessions were confirmed (25 out of 28 participants improved). We also analyzed some characteristics extracted from facial expressions. As a result, a facial expression characteristic correlated with the negative emotional intensity rate and the depression scale score.

1. はじめに

認知行動療法 (CBT) は、非適応的な思考や行動を修正し、抱えている問題の解決に取り組むトレーニングである [Beck 11]. CBT は精神科医療の中でも精神療法の一つとして確立された治療法であり、抑うつ障害や不安障害をはじめとする多くの精神疾患において有効性が認められている。精神療法において治療者とクライアントは言葉のみでなく、表情や韻律、ジェスチャーなどの非言語コミュニケーションをとることが知られており [Koole 16], これまで、医療的対話の自動化を目指し、非言語的行動に対応可能な音声対話エージェント (ECA) が提案されてきている [DeVault 14, Tanaka 17]. 我々は ECA との音声対話により、人間の治療者を再現し、低コストかつ高い効力を有した CBT の自動化を目指す。

CBT に基づいた対話エージェントの先行研究として、Woebot [Fitzpatrick 17], iCBT-AI [Mirai 16], WYSA [Inkster 18] が挙げられる。これらはクライアントの非適応的な思考を修正する機能を有しており、また反応によって応答を変化させている。抑うつ尺度を用いた有効性検証の結果、2 週間の Woebot の継続使用により、事前事後で有意に抑うつ傾向が減少したことが報告されている。一方で、iCBT-AI は応答を変化させる対話を 7 週間継続したが、非介入の待機グループとの比較で抑うつ傾向の有意な減少は見られなかった。しかし、応答変化がなく定型的な対話を行う対話エージェントを用いたところ、抑うつ傾向の有意な減少が見られた [Mirai 16]. これまでの研究は、いずれもテキストベースの対話であり、収集できるインタラクションのモダリティに限りがある。また、各セッションの詳細な効果検証、インタラクションの分析は行われていない。このため、CBT が効率よく進むための要因およびインタラクション中の行動特徴を知ることが難しいという問題がある。また、動的に変化するクライアントの気分を表情

や韻律から推定することも試みられていない。

本研究では、クライアントを ECA グループとテキスト対話グループに分け、定型の対話様式による影響、表情分析からの気分の変化の推定に関して調査を行った。

本論文では、まずはじめに本研究で採用する CBT について概説し、次に定型対話システムでの実装法、実験方法を述べる。実験で収集した各種データの分析結果を示し、最後に対話様式の違いによる CBT の有効性への影響と、気分の変化の推定に有効な特徴について考察を行う。

2. CBT の対話モデル

2.1 認知再構成法

本研究では、CBT の技法の一つである認知再構成法を適用した。認知再構成法は、状況に対する認知情報処理の偏りを検証し、思考を修正することで、反応の改善を行うトレーニングである。図 1 に、気分の改善の流れを示す [Beck 11]. 思考に否定的な偏りがあると、状況に対する認知情報処理が狭小化され、中立的あるいは肯定的な状況までも否定的に解釈し、非適応的な反応が現れる。治療者は、クライアントの置かれている状況の定式化、誘導的な問いかけ、を行い思考の修正を促す。

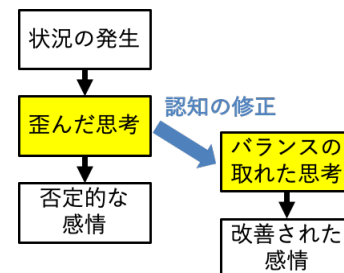


図 1: 認知再構成法のモデル

連絡先: 設楽 一碩, 奈良先端科学技術大学院大学, 奈良県生駒市高山町 8916-5 情報棟 B712, 電話: 0743-72-5264, 電子メール: shidara.kazuhiro.sc5@is.naist.jp

2.2 CBT の流れ

実地臨床の CBT では、トレーニングを的確に行うため治療の流れが構造化されている、特に、本研究で扱う認知再構成法では、問いかげの項目とその順番が決まっている。このことから、対話中のエージェントからの問いかげが完全に定型でも、一定の有効性があると考えられる。本実験では、同一の問いかげを異なる対話様式で実装し、CBT の対話様式による比較検証とセッション中のインタラクションの分析を正確に行うことを目的とする。本実験の定型対話には認知再構成法で用いられる 7 項目の思考記録表 [Greenberger 95] の問いかげに加え、自己紹介、共感を示す言葉、ソクラテス的質問法 [Beck 11] を含めた。ソクラテス的質問法は、クライアントが自身の思考を自覚、評価することを目的とした質問法である。セッションは表 1 の構成で行う。対話エージェントの問いかげの文章は精神科医師による監修のもと作成した。

3. 定型対話エージェントを用いた CBT

3.1 対話様式のグループ分類

本研究では音声対話とテキスト対話それぞれを行うため、2 種類のインターフェースを用意した。条件を揃えるため、問いかげの文章は同一にした。対話プラットフォームはテキスト対話に Slack bot を、ECA に MMDAgent を用いた。

3.2 研究協力者

研究協力者として、計 28 名を募集した。研究協力者の条件は精神疾患を診断されていない大学生である。各研究協力者は、事前に十分な説明を受けた上で本実験の参加に同意した。このうち、13 名はテキスト対話グループとして収録に参加し、15 名は ECA グループとして収録に参加した。研究協力者は事前に抑うつ尺度である K6 [Kessler 02] と QIDS [Rush 03] への記入を行った。K6 (0-24 点) の平均値 (SD) はテキスト対話で 6.23 (4.40)、ECA で 8.07 (4.28) であった。QIDS (0-27 点) の平均値 (SD) はテキスト対話で 6.92 (4.25)、ECA で 7.93 (3.41) であった。各抑うつ尺度に対し、両グループの比較のためウェルチの t 検定を行った結果、K6、QIDS ともに有意差は見られなかった。

3.3 実験手続き

研究協力者はまず抑うつ尺度に回答し、CBT の概要が解説されているリーフレット *1 を読んだ後、定型対話エージェントと対話する。実験の所要時間は全体で 1 時間程度である。対話エージェントを使用するために、HP Probook または MacBook Air を使用し、PC の内部カメラから動画を収録した。

4. 自動化 CBT に向けた対話様式の比較

単独セッションの有効性を、実際の CBT で用いられている評価指標から検証する。また、継続使用への適性を検証するため、かかった時間や発話量 (文章量) についても分析する。

4.1 セッション中の気分の変化

研究協力者が自ら述べた [気分の強さ] は図 1 の思考の歪みの基準として用いられている。思考記録表における気分の変化により、認知の修正が適切に行われたかを評価する。研究協力者が始めに述べた否定的な [気分の強さ] をセッションの開始時点の気分の強さ、セッションの最後に述べる [気分の変化] を終了時点の気分の強さとする。気分の強さの変化と対話様式

の交互作用を調査するため二元配置分散分析を行い、単独セッションの有効性を評価した。

4.2 ECA とテキスト対話の有効性

研究協力者が自ら述べた気分の種類は、[不安、憂鬱、落ち込み、悲しみ、苦痛、劣等感、疲労感、残念] など、いずれも苦痛な気分だった。本研究では、これら全てを一律に [否定的な気分] として扱い、気分の強さにより評価した。図 2 に、縦軸に気分の強さの平均値、横軸に事前事後をとる交互作用図を示す。分析の結果、ECA グループで 14 名、テキスト対話グループで 11 名の否定的な気分の強さが減少した。対話様式における主効果、交互作用は見られなかったが、事前事後における主効果が見られた (対話様式, $F(1, 26) = 0.00, p = .97, \eta_p^2 = 0.00$; 事前事後, $F(1, 26) = 48.3, p < .001, \eta_p^2 = 0.65$; 交互作用, $F(1, 26) = 0.76, p = .39, \eta_p^2 = 0.03$)。

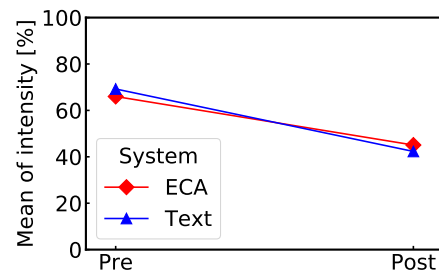


図 2: 否定的な気分の強さの交互作用図

4.3 セッションにかかる時間、発話量 (文章量)

対話エージェントの継続的な使用を考慮する際、セッションにかかる時間は使用者の積極的な継続使用と効率的な反映に影響することが考えられる。そこで対話様式とセッションにかかる時間の関係を分析した。さらに、かかった時間と対話エージェントが引き出した発話量 (文章量) の関係を調査するため、延べ語彙数についても分析した。発話量に関しては対話様式の違いによる使用語彙への影響を調査するため、延べ語彙数中の異なり語数の割合である Type Token Ratio (TTR) を算出した。分析結果を表 2 に示す。各項目に両グループでウェルチの t 検定を行った結果、時間にも有意な差は見られなかった。このことから、ECA グループでは比較的短時間で CBT を行うことができ、かつ引き出せる発話量に差はないと判断できる。また、抑うつ尺度 (K6、QIDS) と、時間、延べ語彙数、TTR に対してピアソンの積率相関係数の算出と無相関検定を行った結果、有意な相関関係は見られなかった (K6 と時間, $r = 0.34, p = .21$; K6 と延べ語彙数, $r = -0.054, p = .84$; K6 と TTR, $r = 0.038, p = .89$; QIDS と時間, $r = 0.081, p = .77$; QIDS と延べ語彙数, $r = -0.068, p = .81$; QIDS と TTR, $r = 0.0042, p = .99$)。よって、発話量と抑うつ傾向に相関関係はないと判断できる。

5. 気分の変化の推定に向けた特徴分析

5.1 否定的な気分と抑うつ傾向の関係

認知再構成法は思考の修正のスキルを身に付けることが目的のため、一度のみのセッションで抑うつ傾向を改善することを前提としておらず、事前事後の否定的な気分の強さはそのセッションでの思考の修正が適切に行われたかを評価するために用いられる。そのため、抑うつ傾向による単独のセッションの評

*1 <https://www.cbtjp.net/downloads/skillup/pdf/>

表 1: 各ターンのシステムの問いかけ

項目	問いかけの文章
状況	こんにちは、セラピストのマイと申します。 これから私と一緒に、悩みと向き合うトレーニングをしましょう。今何か悩みを抱えていませんか。 ここ最近で、気がかりなことや、向き合うのが辛いことはありますか。 もし良ければ、あなたが辛いと思っていることを私に教えてください。
気分	そのとき、どんな気分になりましたか。
気分の強さ	その気分の強さを 0 から 100 で表すとどれくらいですか。
自動思考	それは大変でしたね…今からその出来事をテーマにトレーニングを始めたいと思います。 なぜその状況だとそのような気分になるのかを知りたいです。 そのできごとに直面した際、どのような考えが頭に浮かびましたか。
根拠	なるほど…そのような、ある出来事に直面した時に頭に思い浮かぶ考えのことを、自動思考と呼びます。 自動思考は無意識に浮かんでくるものなので、ついあたかも真実であるかのように感じてしまうことがあります。 あなたのその自動思考が正しいとしたら、その根拠はなんだと思いますか。教えてください。
反証	なるほど、では、その根拠に対して別の見方はできるでしょうか。 例えば、その自動思考に反する事実はありませんか。
反証の補足	その調子です。他にも思い浮かびますか。なければこれで全部と言ってください。
ソクラテス的質問法 1	これまで考えた自動思考の根拠と、それに反する事実からバランスの取れた考え方を検討しましょう。 そのための質問をいくつか行います。今回の状況で、考える最悪の結末はどのようなものですか。
ソクラテス的質問法 2	逆に、思い通りに行った場合の結末はどのようなものですか。
ソクラテス式質問法 3	では、今の二つの予想から、一番現実的なシナリオを考えて、教えてください。
適応的思考	その答えから、新しい思考を作り出してみよう。 今回の出来事に対して、どのように考え直すことができますか。
気遣いの言葉 1	いいですね。今回話してくれた話題について、身の回りに頼れる人はいますか。
気遣いの言葉 2	なるほど、なんとかやっつけていけそうですね。
気分の変化	そうですね、誰かと相談すると気持ちも軽くなることもあります。私でよければいつでも話してください。 さて、はじめに感じていた気分の強さはどのくらいに変わりましたか。 0 から 100 の数値で表してください。
セッション終了	もし始めと変わっていたら、うまく考えを整理できた証です。 今日はこれで終了です。お疲れ様でした。またいつでも声をかけてくださいね。

表 2: 時間、延べ語彙数、TTR の平均値 (SD) と t 値

	テキスト対話	ECA	t (df)
時間 (sec)	681.5 (276.1)	401.1 (142.6)	3.30 (17.4)**
延べ語彙数	195.5 (150.2)	349.5 (247.6)	-2.02 (23.5)
TTR	0.53 (0.11)	0.45 (0.12)	1.81 (25.9)

** $p < .01$

価は適切ではない。一方で、抑うつ傾向の高低は本人の気分の変化を推定する上で重要な指標である。クライアントが自ら述べた否定的な気分を抑うつ傾向の判断基準とするべきかを調査するため、K6、QIDS の評点とセッションの開始時点の気分の強さの関係を、ピアソンの積率相関係数の算出と無相関検定により分析した。ECA グループを対象に調査を行った結果、K6 と気分の強さの間には有意な相関関係が見られなかった ($r = 0.36, p = .19$)。また、QIDS と気分の強さの間にも有意な相関関係は見られなかった ($r = 0.36, p = .19$)。結果として、本実験の収録内容からは、抑うつ傾向の否定的な気分の強さへの影響は見られなかった。

5.2 抑うつ傾向、否定的な気分と表情の関係

先行研究 [Cohn 09] において、表情の特徴である Action Units (AUs) [Ekman 97] の AU14 (えくぼを作る) が抑うつ傾向の識別に有効であることが報告されている。本研究では OpenFace*2 を用いて収録した動画から AUs の抽出を行っ

た。動画のフレームレートは 30fps であり、1 フレームごとに特徴の抽出を行った。実際の CBT のコミュニケーションに近づけるため、ECA グループを対象とした。前処理として、表情認識に失敗したフレームは取り除き、CBT による思考の修正の影響がないと考えられる、最初の 3 ターンの間映像を使用した。OpenFace で抽出可能な特徴量 AU (01, 02, 04, 05, 06, 07, 09, 10, 12, 14, 15, 17, 20, 23, 25, 26, 28, 45) の出現率と否定的な気分の強さでピアソンの積率相関係数算出と無相関検定を行った結果、AU14 の出現率との間に有意な相関関係が見られた ($r = -0.57, p < .05$)。さらに、各 AU の出現率と抑うつ尺度 (K6) について同様に分析した結果でも、AU14 の出現率との間に有意な相関関係が見られた ($r = -0.54, p < .05$)。AU14 の出現率と否定的な気分の散布図を図 3 に、K6 の評点との散布図を図 4 に示す。

5.3 セッション前後の表情変化

次に、セッションの開始時点と終了時点の表情の映像から、各 AU で対応のある t 検定を行った結果を述べる。ここで、開始時点は思考の修正前であるセッション開始直後の 3 ターンの間映像を、終了時点は思考の修正後であるセッション終了直前の 3 ターンの間映像を使用した。結果、AU12 の出現率に有意な変化が起き、出現率が有意に減少していた (平均値の差分 $= -15.7\%$, $t(14) = 2.85, p < .05$)。AU12 の出現は口角を上げることが示している。表情の中で、AU12 の出現率のみが顕著に上がる状況としては、フラストレーションの高ま

*2 <https://github.com/TadasBaltrusaitis/OpenFace>

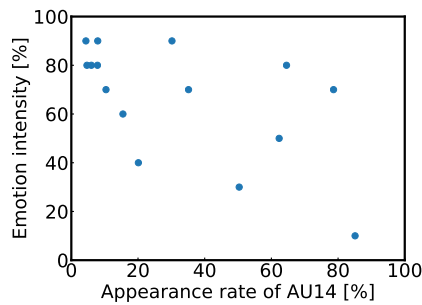


図 3: AU14 の出現率と否定的な気分の強さの散布図

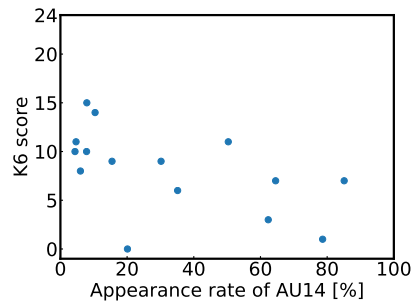


図 4: AU14 の出現率と K6 の評点の散布図

りが報告されている [Hoque 12]. これより、研究協力者は対話の開始時点ではフラストレーションを感じていたが、対話の進行につれ軽減したと捉えられる。フラストレーションの要因としては、自身が抱える問題や、ECA に対して抱く印象が考えられる。AU12 の出現率の低下は、これらのような要因に前向きな変化が起き、表情として現れた可能性を示唆している。

6. まとめ

本研究では、CBT を行う対話エージェントをテキスト対話と、ECA の対話様式で構築し、セッションの収録を行った。計 28 名のテキスト対話と音声対話により、CBT を基にした定型対話の評価を行った。結果として、対話様式によらず否定的な気分の強さが減少した。さらにセッション中のインタラクションの分析を行った。発話量とかかった時間の分析の結果、両グループの延べ語彙数に差はなかったがセッションの時間は短縮されていた。表情に関しては、特定の表情の特徴と抑うつ傾向。自ら述べた否定的な気分の強さ、に有意な相関関係が見られた。これらから、音声対話による非言語的行動の検出と、ECA との音声対話による単独セッションの有効性が確認できた。今後は音声特徴を含めたより詳細な分析を進め、言語的・非言語的特徴から認知の歪みの分類や否定的な気分の予測を行い、クライアントの気分の変化に適応可能な ECA の開発につなげていく。

謝辞

本研究は CREST (グラント番号: JPMJCR19A5) の支援によって行われた。

参考文献

[Beck 11] Beck, J. S. and Beck, A.: Cognitive behavior therapy, *New York: Basics and beyond. Guilford Publication* (2011)

- [Cohn 09] Cohn, J. F., Kruez, T. S., Matthews, I., Yang, Y., Nguyen, M. H., Padilla, M. T., Zhou, F., and Torre, De la F.: Detecting depression from facial actions and vocal prosody, in *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, pp. 1–7IEEE (2009)
- [DeVault 14] DeVault, D., Artstein, R., Benn, G., Dey, T., Fast, E., Gainer, A., Georgila, K., Gratch, J., Hartholt, A., Lhommet, M., et al.: SimSensei Kiosk: A virtual human interviewer for healthcare decision support, in *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, pp. 1061–1068 (2014)
- [Ekman 97] Ekman, R.: *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*, Oxford University Press, USA (1997)
- [Fitzpatrick 17] Fitzpatrick, K. K., Darcy, A., and Vierhile, M.: Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial, *JMIR mental health*, Vol. 4, No. 2, p. e19 (2017)
- [Greenberger 95] Greenberger, D. and Padesky, C. A.: *Mind over Mood: a cognitive therapy treatment manual for clients*. (1995)
- [Hoque 12] Hoque, M. E., McDuff, D. J., and Picard, R. W.: Exploring temporal patterns in classifying frustrated and delighted smiles, *IEEE Transactions on Affective Computing*, Vol. 3, No. 3, pp. 323–334 (2012)
- [Inkster 18] Inkster, B., Sarda, S., and Subramanian, V.: An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study, *JMIR mHealth and uHealth*, Vol. 6, No. 11, p. e12106 (2018)
- [Kessler 02] Kessler, R. C., Andrews, G., Colpe, L. J., Hiripi, E., Mroczek, D. K., Normand, S.-L., Walters, E. E., and Zaslavsky, A. M.: Short screening scales to monitor population prevalences and trends in non-specific psychological distress, *Psychological medicine*, Vol. 32, No. 6, pp. 959–976 (2002)
- [Koole 16] Koole, S. L. and Tschacher, W.: Synchrony in psychotherapy: A review and an integrative framework for the therapeutic alliance, *Frontiers in psychology*, Vol. 7, p. 862 (2016)
- [Mirai 16] Mirai, S., Yoichi, S., Yoshitake, T., et al.: Can Artificial Intelligence Heal Human Hearts? A randomized controlled trial on the effects of internet cognitive behavioral therapy with artificial intelligence on depression (Japanese), Technical report (2016)
- [Rush 03] Rush, A. J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., Markowitz, J. C., Nisan, P. T., Kornstein, S., Manber, R., et al.: The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression, *Biological psychiatry*, Vol. 54, No. 5, pp. 573–583 (2003)
- [Tanaka 17] Tanaka, H., Negoro, H., Iwasaka, H., and Nakamura, S.: Embodied conversational agents for multimodal automated social skills training in people with autism spectrum disorders, *PloS one*, Vol. 12, No. 8 (2017)