

人形との相互作用によるプログラミング時の思考整理支援 Support for organizing thoughts during programming through interaction with dolls

倉田 寛大 †* 佐々木 耕平 †† 辻 愛里 ††† 藤波 香織 †††
Kandai KURATA Kouhei SASAKI †† Airi TSUJI ††† Kaori FUJINAMI †††

† 東京農工大学 工学部 知能情報システム工学科

† Department of Electrical Engineering and Computer Science, Tokyo University of Agriculture and Technology

†† 東京農工大学 生物システム応用科学府 生物機能システム科学専攻

†† Department of Bio-Functions and Systems Science, Tokyo University of Agriculture and Technology

††† 東京農工大学 大学院 工学研究院 先端情報科学部門

††† Division of Advanced Information Technology and Computer Science, Tokyo University of Agriculture and Technology

Abstract

In recent years, when the shortage of IT human resources has become an issue, it is significant to provide multifaceted support to programmers of various skill levels through systems that do not require a lot of manpower. Until now, there have been many attempts to develop systems to train users' tracing and algorithmic thinking abilities, and to improve their skills by providing hints from the outside. In this study, we focused on rubber duck debugging with the aim of helping users organize information in their brains so that they can maximize their abilities. This is a debugging method in which users find clues to solving problems as they explain program details and errors to a duck doll, and is considered to be an effective debugging method because it allows users to infer and organize causal and conceptual relationships among objects in the process of explaining them. We consider that the doll could be more effective in organizing thoughts if it were interactive, so a doll equipped with an interactive function was implemented and investigated its effect on user support.

1 はじめに

IT人材の不足が問題となっている昨今、人手を割かずシステムによってさまざまな熟練度のプログラマを多角的に支援することには意義がある。今日までユーザのトレース力やアルゴリズム的思考力を鍛錬するシステムを開発して、外部からヒントを提示するような形でユーザのスキルアップを図る試みが盛んに行われてきた [1][2]。また HRI (Human Robot Interaction) に関する研究として、角ら [3] は、ヒト型のエージェントとその人相を顔面に投影したロボットを用意し、ユーザへの励ましによるモチベーション維持の効果を比較し、結果として、エージェント条件の方がロボット条件よりも課題遂行回数が多かったことから、エージェントの方がロボットよりも説得力が高いことがわかった。Kiesler ら [4] は、同じ人に似せたロボットの外観でも、2D と 3D、対面と遠隔、のような条件の違いによって起こる作用を調査し、信頼度、社交的、反応のよさ、尊敬できる等あらゆる点でロボット条件がエージェント条件を上回る高評価を得たことから、ロボット条件のほうがエージェント条件と比べて人間のように感じることを示している。

本研究では、ユーザの脳内情報の整理を支援することで持てる力を最大限発揮させることを目的とし、ラバーダックデバッグに着目した。これは、アヒル人形に向かってプログラムの詳細およびエラーについて説明するうちに問題解決の糸口が見つかるといったデバッグ方法であり、事柄について説明を行う過程で対象の事柄の因果関係や概念的な関係を推論して整理するため効果的であると考えられている。我々はこの人形をインタラクティブにすることで、さらなる思考整理効果が得られると考え、相槌を打つ人形を実装し、ユー

ザの支援に及ぼす影響について調査した。

2 提案システムの設計と実装

2.1 設計

図 1 にシステム概要図を、図 2 にシステムの外観を示す。ラバーダック・デバッグに倣いゴム製のアヒル人形にデバイスを組み込む形で提案システム (uh-huh duck) を開発した。使用例は以下の通りである。

1. プログラミング作業中に行き詰るとトリガーであるスイッチを押下する。
2. システムからユーザに一声かけて説明を促す。
例「どんな構造なの？」
3. ユーザは説明を開始する。(その間システムは会話の切れ目で相槌を打つ)
4. 説明が終わり次第プッシュスイッチを押下してシステムを停止する。
5. 再度悩みが生じたら 1 に戻る。(2 の質問内容は前回の説明を踏まえた内容)

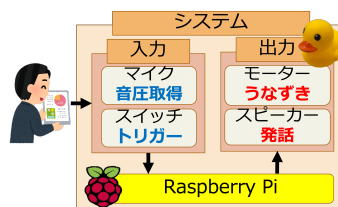


図 1: システム概要図

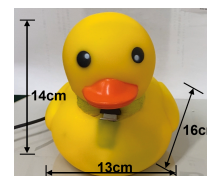


図 2: システム外観

また、システムの要件として以下の 2 点を定義し、実装に着手した。

- 使用者の会話の快適性を上げる
- アヒル人形内にデバイスがすべて収まる

* 連絡先: 東京農工大学工学部知能情報システム工学科
〒 184-8588 東京都小金井市中町 2-24-16
E-mail: s217337z@st.go.tuat.ac.jp

相槌によってユーザの発話の快適性を上げること、アヒル人形一つでデバッグ作業が行える手軽さを損なわないことに注意し、インタラクティブ化を行った詳細を次節に記す。

2.2 実装

うなずき

人形の頭部を胴部と切り離し、Raspberry Piにて制御可能なパンチルト機構付きサーボモータを頭部内側に組み込むことで、うなずき機能を実装した。また、モータ駆動音による思考への弊害を鑑みて、防音シートを人形の内側に張りモータを覆い、スピーカーからの発話によって掻き消すことで、消音化した。なお、首の接着にも防音シートを使用した。

発話

テキスト読み上げソフト・Open JTalk を Python 上で動作させることで発話機能を実装した。人形の外見との組み合わせを考慮し、主観的に適切と判断した女声を採用した。相槌として予め用意した三種の単語（表1）からランダムに一つを、説明開始時には実験者の入力したテキストを読み上げるよう設計した。これらテキストは Raspberry Pi に接続された USB スピーカーを用いて出力した。

表 1: 発話内容

| タイミング | 発話内容 |
|-------|------------|
| 相槌 | 「うんうん」 |
| | 「へえー」 |
| | 「なるほど」 |
| 説明開始時 | 実験者がその都度決定 |
| 説明終了時 | 「説明ありがとう」 |

相槌のタイミング

小林ら [5] は、対話システムが会話の切れ目で相槌をうつと、会話中のユーザの快適性が高まることを示した。これに倣い本システムは、Raspberry Pi に接続した USB マイクによって、0.3 秒おきに周囲の音圧を測定し、閾値 45dB を 2 秒以上上回り続け、次に下回ったタイミングで相槌をうつよう設計した（図 3）。

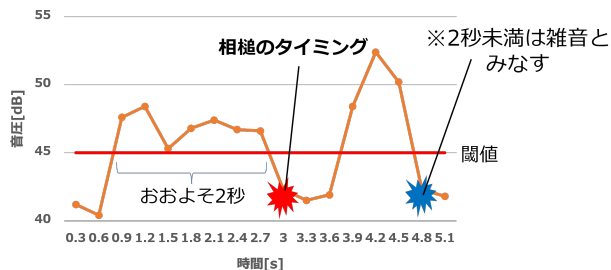


図 3: 会話の切れ目検出の概要

デバイスの収容

要件を満たすため、入力用のスイッチとマイク、出力用のモータとスピーカー、制御用の Raspberry Pi をアヒル人形内に収容した。なおマイクに関しては周囲の音を拾う必要があるため、首の付け根の前面から外部に露

出させる形で配置した。その他例外として、Raspberry Pi の給電用ケーブルと、トリガーとなるマイクロスイッチを外部へ露出している。なお首の接着にもモータ駆動音の消音化にも用いた防音シートを使用した。最終的なデバイスの配置を図 4 に示す。

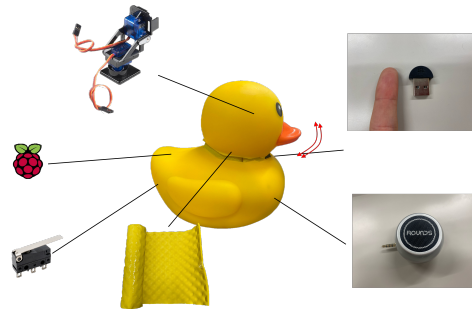


図 4: デバイス配置図

3 評価実験と考察

3.1 実験方法

インタラクティブ化による効果と、インタラクションの有無や適切さによる影響を調査するため、以下の 3 条件を設けて比較実験を行った。

1. アヒル人形のみ（対照群）
2. 提案システム（以下、正システム）
3. 返答内容と相槌のタイミングが無作為なシステム（以下、誤システム）

正システムでは実験者が Wizard of Oz 法にてその都度適切に返答内容を決定した。誤システムでは 10 秒を 1 区間とし、1 区間でランダムなタイミングで相槌を行い、あらかじめ設定したプログラミングに関する定型文 7 種（表 2）のいずれかを読み上げるよう設計した。

表 2: 誤システムの発話内容

| |
|---------------------|
| 何を出力したいの？ |
| 入力された値をどうしたいの？ |
| ループはちゃんと抜け出せるの？ |
| 処理順は正しいの？ |
| 無駄な処理は本当はないの？ |
| コーディングしやすい構造になってるの？ |
| データの流れはどうなってるの？ |

評価実験は、情報系の学部及び大学院に所属する 20 代の 12 名（男性：6 名、女性：6 名）を対象に行った。タスク実施中の被験者と実験者のフローを図 5 に示す。順序効果を相殺するため、各条件の順序を変えた 6 通りの順序で実験を実施した。実験タスクは鈴木らの手法 [6] を参考に、被験者にはプログラミングの問題文を提示し、そこからフローチャートを作成するよう設計した。フローチャートの完成次第、3 条件のいずれかを使用して検討する時間を与え、修正を加えるか否かを被験者に確認した。なおデータ収集の利便性等を考慮して、タスク実施用のツールはゲームエンジン・Unity を用いて作成した（図 6）。このツールの操作の慣れによる影響を除外するため、本実験までに簡易的なタスクを与え被験者がツールに触れる機会を設けた。また適切な問題難易度の設定のため、被験者には事前

に Web サイト・paiza[7] を用いてコーディングスキルの測定を実施させ、実験タスク用の問題も同サイトから各被験者の適正レベルの問題を選定した。なお問題と条件の組合せによる影響を除外するため、同じ組合せを複数の被験者に与えることのないよう設計した。

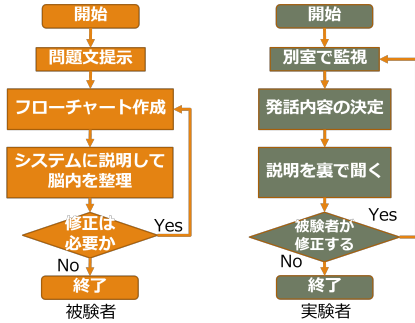


図 5: 被験者および実験者のフロー

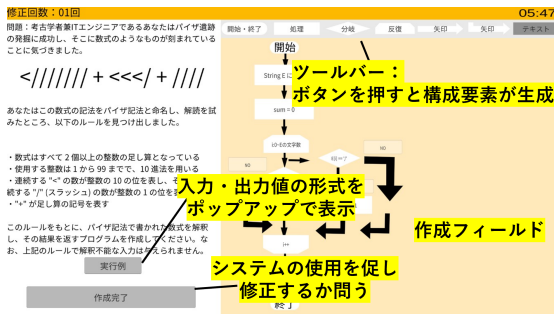


図 6: 実験用ツール

3.2 実験環境

図 7 に実験環境を示す。システムの仕様上、周囲が静かでない则会話の切れ目が判定できないため実験室の空調をオフにする等静音環境を保ち実験を実施した。また、タスク実施用ツールの取扱説明書をそばに配置しておき、いつでも確認できるようにした。



図 7: 実験環境

3.3 評価方法

定性評価のため、各条件間でシステムの印象や使用感に関するアンケートを実施した。その他にも、実験後にはシステムのユーザビリティを測定するため、SUS (System Usability Scale) を調査するアンケートを提案システムを対象に実施した。また、実験で得られたフローチャートの正しさや修正方針の分類などによる定量的な評価も実施した。

3.4 実験結果

タスク実施後アンケートの結果を図 8 に示す。各条件間で有意水準 5% でライアンの多重比較法により検定を行ったところ、システムの有用性と印象について、有意差が見られた。また、条件設定時には全ての項目において正システムが誤システムより高評価を得ることを想定していたが、部分的に棄却された。

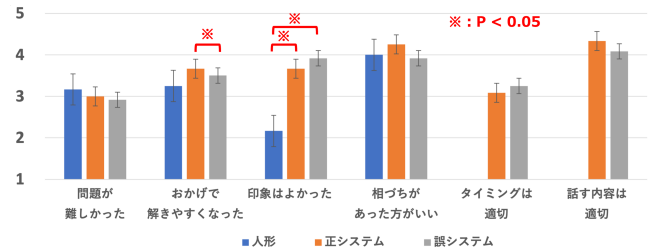


図 8: アンケート結果 (5 段階)

(1: 当てはまらない 3: どちらでもない 5: 当てはまる)

次に問題難易度が高いほど提案システムの効果が高まるという仮説のもと、「難しく感じた」と「解きやすくなった」の 2 項目間の相関係数 (表 3) を算出したところ、正システム条件において強い正の相関があることが確認できる。

表 3: 主観的難易度と有用性の相関係数

| 人形のみ | 正システム | 誤システム |
|-------|-------|-------|
| -0.19 | 0.77 | 0.46 |

最後に、実験後アンケートで実施した SUS のスコアは 73.8 となった。SUS は平均スコアは 68 とされており [8]、提案システムはこれを上回ったことから、高いユーザビリティを示すことを確認した。

また、定量評価として得られたフローチャートの正しさを独自に設定した基準 (表 4) に従い採点した結果を表 5 に示す。

表 4: フローチャート正しさの判定基準

| 点数 | 該当基準 |
|----|----------------|
| 5 | ミスがない、完璧な状態 |
| 4 | スペルミス等の初歩的なミス |
| 3 | 処理に抜けがある |
| 2 | 問題をあまり理解できていない |
| 1 | 見間違い |

表 5: フローチャート正しさの採点結果 平均値 (満点者数)

| 人形のみ | 正システム | 誤システム |
|----------|----------|----------|
| 3.58 (2) | 3.92 (4) | 4.25 (5) |

さらに条件間の差異を明らかにするために、実験タスク中のシステム使用による修正を、条件と方針の点から分類した。方針の分類項目として、行き詰まりを解消するためにシステムを用いた場合と、システムの使用によって初歩的ミスの解消を行った場合と、処理の大きな変更を行った場合の 3 つを設けた。

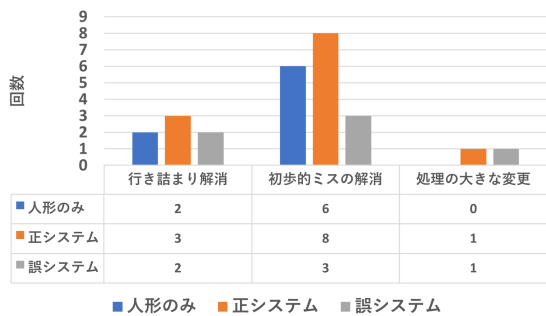


図 9: 修正方針の分類結果

3.5 考察

3.5.1 システムに対する印象

人形が正・誤システムより低い評価を得たことの原因としては、被験者からの「相槌が無いと不安になる」「反応のないモノに話しかけるのは抵抗が強い」といった意見が挙げられる。システムの好印象の要因としても、「うなずいてくれたから」「反応があるので聞こえていると感じて安心感を覚えた」といった被験者の意見があった。しかし、本研究における被験者は本研究の概要について知識を有していたことから、「アヒル人形型システムは相槌をするモノ」といったバイアスが加かっていた可能性が存在する。

3.5.2 相槌タイミングの適切さ

タスク実施後アンケートの結果を示した図 8 より、ランダムに相槌を打つ誤システムの方が会話の切れ目で相槌を打つ正システムよりも高い評価を得た。この項目では正システムの方が高い評価を得ることを想定して設計したが棄却された。要因としては、会話の切れ目抽出のための閾値の設定に誤りがあったことが考えられたが、正システムに対して「切れ目を抽出していることはわかったが、タイムラグで次の会話に相槌が被っていた」という意見が被験者から得られたことから、実装の精度に問題があったことがわかる。実際に、本システムは会話の切れ目を抽出してから動作まで 0.5 秒ほどのタイムラグが発生することから、話者の特性によっては今後の課題として解決を図るべき点である。

3.5.3 問題難易度とシステム有用性の相関

表 3 の相関係数について、正システムがその他 2 条件と比べて強い正の相関があることが確認できる。正システムに対して被験者からは「的確な質問をしてくれたおかげでミスに気づけた」「自分の説明を理解しているかのような質問を返してくれて脳内が整理できた」といった意見が得られたことから、ミスが生じやすい高難易度の問題ほど、適切な内容の質問をする正システムの方が脳内整理効果が高かったことが明らかになった。すなわち、返答内容の違いがこの結果の要因であることが考えられる。なお図 8 に示す通り、被験者の主観によって測定した問題難易度は全ての条件間で有意差が見られなかったため、影響はないと考えられる。

3.5.4 提案システムのユーザビリティ

提案システムのユーザビリティを定量的に表した SUS のスコアについては、システムはスイッチの押下以外は特に複雑な操作の必要がないため、高い評価を得られたと考えられる。

3.5.5 インタラクティブ化による影響

表 5 の採点結果について、満点獲得者数の条件間での差に注目した。満点獲得の要因としては、システムを使用する必要がないほど問題を簡単に感じた場合（以下、可能性 1）と、システムを使用したことで被験者は些細なミスに気づいて修正できた場合（以下、可能性 2）が存在すると考えられる。具体的に、可能性 1 はシステム使用回数が 1 回かつアンケートの「問題を難しく感じた」に対する回答が 2 以下、可能性 2 はシステム使用回数が 2 回以上かつアンケートの「システムのおかげで解きやすくなった」に対する回答が 4 以上であるとし、条件ごとに内訳を確認した（図 10）。

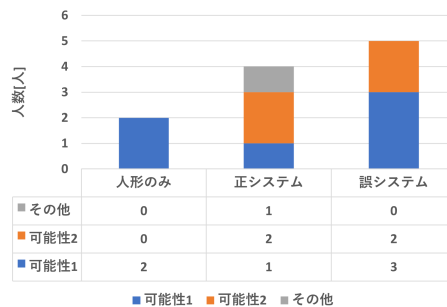


図 10: 満点獲得者の条件・要因内訳

可能性 2) に注目すると、人形のみ条件では該当者が 0 人なのに対して正・誤システム条件では 2 人ずつ該当した。このことから、インタラクティブ化したことによって思考整理効果が高まったことがわかる。また、修正方針の分類結果（図 9）のうち大きな処理の変更について注目すると、人形のみ条件では該当者が 0 人なのに対して正システム、誤システム条件では 1 人ずつ該当した。このことから、同様の知見が得られた。

4 おわりに

本研究では、インタラクティブなアヒル人形型システムを製作し、インタラクティブ化による影響とインタラクションの有無や適切さによる影響を検証した。実験の結果から、インタラクティブ化がシステムに対する印象と思考整理効果に良い影響を及ぼすことが示された。また、ユーザーが問題難易度を難しいと感じる際は特に提案システムが有効であることを確認した。今後は、タイムラグの解消を図りつつ、瞬きや首傾げなどの新たなインタラクションの実装に取り組む。

参考文献

- [1] 江木鶴子ほか. プログラミング初心者にとレースを指導するデバッグ支援システムの開発と評価. 日本教育工学会論文誌, 32(4), pp. 369–381, 2009.
- [2] 相知政司ほか. フローチャートを利用したアルゴリズム学習支援システムの開発. 電気学会論文誌 A, 126(12), pp. 1199–1204, 2006.
- [3] 角薫ほか. モチベーション維持におけるキャラクターエージェントとロボットの比較. Human-Agent Interaction Symposium 2011.
- [4] Sara Kiesler ほか. Anthropomorphic interactions with a robot and robot-like agent. *Social Cognition*, 26(2), pp. 169–181, 2008.
- [5] 小林一樹ほか. Ase に基づく相槌によるロボットとの対話体験の向上. 人工知能学会論文誌, 30(4), pp. 604–612, 2015.
- [6] 鈴木優実ほか. プログラミング初学者にむけたアルゴリズム的思考習得のための学習支援手法. Technical Report 6, 筑波大学情報学群知識情報・図書館学類, 2018.
- [7] paiza. IT エンジニア向け国内最大の転職・就職・学習プラットフォーム. <https://paiza.jp>. (最終アクセス 2023 年 1 月 7 日)
- [8] Jeff Sauro. Measuring usability with the system usability scale (SUS). <https://measuringu.com/sus/>. (最終アクセス 2023 年 2 月 7 日)