

プログラミング学習における意欲向上のための 生成 AI 利用の検討

市村衣未^{†1} 神場知成^{†1}

概要: プログラミング学習において生成 AI (大規模言語モデル) の利用は有効であるものの、使い方によっては生成 AI のコード生成能力に学習者が依存して自主的なプログラミング力が育たない可能性があるという課題も知られている。本研究では、コード作成支援ではなく学習意欲向上のために生成 AI を利用する試みとして、学習サイトで与えたプログラミング課題に対する学習者の回答をふまえて「論理フィードバック」と「情緒フィードバック」の2種類を生成 AI にそれぞれ異なるプロンプトで生成させるシステムを開発した。被験者となる大学生 12 名が未学習の言語 Ruby のプログラミングを対象として 2 群に分けた簡易的な実験を実施した。その結果、情緒フィードバックを含めて提示した場合には、論理フィードバックだけの場合と比較してこの言語に対する学習意欲の向上が示唆される傾向が見られた。

1. はじめに

大規模言語モデルを活用した生成 AI は幅広い問題の解決能力に優れており、教育分野での活用が注目されている [1]。学習したデータをもとに細かなカスタマイズを行うことで教材の作成や演習の補助など、従来よりも高度にパーソナライズされた学習体験を提供することが可能になってきた。

プログラミング領域においても様々なコード生成ツールが開発されプログラミングの手段が大きく変化した [2]。GitHub Copilot は一般的な統合開発環境のプラグインであり、コードの意図を汲み取りより最適な提案を行うことができる。また ChatGPT はコードの疑問点などを自然言語で質問することができるようになり、個人学習の方法にも変化が起きている。

プログラミング学習において学習意欲の継続は非常に重要である [3]。コード自体は生成 AI が記述できるようになったが、学習者が AI によって生成されたコードに依存し学習を放棄する懸念もある。学習者が高い学習意欲を持ち、生成 AI に依存せずにコードの生成および理解力が高まる状態にすることが望ましい。本研究では初学者を対象に、プログラミングを生成 AI を用いて学習する際に、学習意欲を高める方法の検討を行う。

2. 関連研究

2.1 プログラミングにおける学習意欲

学習意欲の構成要素は複雑であり、プログラミング学習においても様々な定義が用いられている。Marwan 等はプログラミング学習の中で一定時間以上コードを書いていないアイドルタイムを学習意欲の高低と関連づけた [4]。Lee

等は、時間無制限という条件下で、学習意欲が向上すればより多くの問題に取り組むと仮定し、解いた問題数を学習意欲の基準とした [5]。

以上のような研究において一般に、学習者の回答に対し即座に正誤や解説を行う即時フィードバックが学習意欲の向上につながる事が指摘されており、特に初学者に有効とされている。これはエラーを早期に修正することで課題の進行がスムーズになり、より多くの成功体験を積むことができるため高い動機づけに繋がるためとされている [6]。

2.2 生成 AI を用いた学習

生成 AI によってコード生成能力を支援するための複数の調査において、特に初学者においては生成 AI の利用が有効であるだけでなく学習弊害も報告されている。有効な面としては、たとえば Leinonen 等は生成 AI がエラーメッセージの説明と修正方法の提案をすることで、コードの可読性が高まり初学者の理解を促進したと述べている [7]。Phung 等は初学者に多い Syntax Error の指摘とコード概要の説明に生成 AI を用いることで、より質の高いコード生成の支援を行った [8]。一方、これらの研究において学習者はコード学習に容易に取り組めるという利点だけでなく、生成 AI のコード生成能力に依存しがちであるという問題も発生した。

生成 AI によって生成されたコードへの学習者の依存を防ぐための様々な工夫も行われている。直接的にコードを使用せずに解説を行う方法や、ヒントの提示に留め学習者の深い理解を促すなどの対策が取られている [9] [10]。プログラミングを学ぶ上で初学者の間は基本的な概念を理解することが以降の開発体験において重要であり、どの程度生成 AI を用いるかは慎重に検討する必要がある。

^{†1} 東洋大学 情報連携学部 (INIAD)

2.3 プログラミング学習におけるフィードバック

プログラミングでは学習者は一般的にエラーといった失敗のメッセージを多く受け取るため、ポジティブなメッセージを受け取る機会に乏しい。既存のコード演習ツールで提供される自動フィードバックシステムの多くが間違いを指摘することに重点をおいている [11]。これは初学者がプログラミングを行う際に混乱やフラストレーションを感じる要因になり、意欲の低下につながる可能性がある [12]。

学習意欲の向上につながるフィードバック提示方法の検討として、Lee 等は擬人化したキャラクターを用いて失敗を共有し、学習者がキャラクターを助けるというインタラクション設計により、初学者の学習意欲が向上したと示した [5]。また Marwan 等は進捗状況を褒める短文をポップアップメッセージとして表示することで、学習者は励まされているように感じ学習の継続に繋がると報告した [4]。いずれもの研究も学習者が自信を高めるフィードバックを返すことが、プログラミング学習意欲の向上に貢献することを示している。

本研究ではこれらの成果も踏まえ、生成 AI を用いることにより、学習者の状況にあわせて学習意欲を向上するようなフィードバックを与える手法の検討を行う。

3. 開発システム

学習者が簡単なプログラミング課題に回答しながらプログラミング学習を行うサイトを開発した。このシステムは、課題を表示して学習者が提出したコードに対して、正誤を伝える論理的なフィードバック（以降、論理 FB）と、学習への意欲を高める情緒的なフィードバック（以降、情緒 FB）の 2 種類を提供する。従来、生成 AI を活用したプログラミング学習では、ChatGPT や Gemini といったチャットボットが多く利用されていたが、その場合に生成 AI のレスポンスにはコードそのものに関するものや、アドバイスに相当するものがランダムに含まれていた。筆者らは生成 AI を用いた学習において、学習者が提出したコードの正誤に関するコメントと、学習プロセスに関するポジティブなコメントを区別することが重要と考えた、具体的には、コードの正誤に関しては、学習者が生成 AI に依存せずに自ら考えられるような独立した解説が必要である。一方、学習プロセスに関しては、従来は教師が行っていたような努力を評価し励ます役割が有効であろう。それにより、学習者が一人で学習するオンライン学習サイトにおいても、モチベーションの維持・向上につながる可能性がある。

3.1 システム構成

図 1 に本システムのアーキテクチャを示す。学習者がコード作成に用いるエディタは JavaScript で埋め込み可能な Ace エディタ (<https://ace.c9.io/>) とした。課題に対して学習者がコードを提出すると、システムは Web 上のコード実行

機能を持つ WebAssembly (<https://webassembly.org/>) で実行する（以降、Wasm）。それをもとにした学習者へのフィードバック生成のためには OpenAI 社の API (GPT 4o-mini) を使用する。GPT 4o-mini には、課題の文章、学習者が提出したコード、それを Wasm で実行した結果。さらに論理 FB と情緒 FB それぞれを生成するために必要な情報を含むプロンプトを渡す。また、一連のやりとりで、過去に学習者が提出したコードやそれに対するシステムのフィードバックは、学習者ごとに保管している。

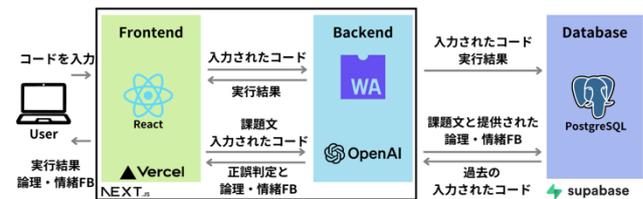


図 1 システム構成図

3.2 論理 FB の設計

生成 AI が作成したコードに初学者が依存することを防ぐために、生成 AI は課題に対する正解コードを生成せずに解説を行う。具体的には、各課題ごとに GPT 4o-mini には、プロンプトに対する付属情報として次の 4 種類を与える。

1. プログラミング課題本文
2. 課題に対して学習者が回答したコード
3. 一連の学習過程で学習者が提出したコード
4. 学習者が回答したコードの Wasm 上の実行結果

上記に加えて GPT 4o-mini には最初に、学習者の学習歴を与える。本研究では初学者を対象にしているため、プロンプトには「初学者を対象に解説してください」と指示した。これによりコード解説をシンプルかつ理解が容易な表現で生成することが期待できる。

3.3 情緒 FB の設計

学習者の学習意欲を高めるため、取り組みに対して肯定的なメッセージを生成する。具体的には GPT 4o-mini にプロンプトとして与える情報には次の 4 種類がある。

1. プログラミング課題本文
2. 課題に対して学習者が回答したコード
3. 一連の学習過程で学習者が提出したコード
4. システムから学習者に対する情緒 FB の例

1~3 は論理 FB と同じであり、4 だけが異なる。情緒 FB の例としては、次の 3 種類を与える。1) できたことを端的に褒める例 2) 過去と比較しできるようになったことを褒める例 3) 以前できていたことが解けなかった場合、できるはずなので再度取り組みを促しながらほめる例。なお、GPT 4o-mini に対するプロンプトにおいて明示的に「褒めてください」というプロンプトを与えた場合は、「天才です」

「とても優秀です」といった過度な表現の使用が増えたため、そのような言葉は用いずに褒める例を示すようにした。

また生成 AI によるコメントは、論理 FB、情緒 FB いずれも学習者の回答提出時に自動的に表示されるのではなく、学習者が「表示」のボタンを押したときだけ表示されるようにした。

3.4 学習者インタフェース

図 2 は論理 FB と情緒 FB という 2 種類の即時フィードバックが表示された画面の例である。ヘッダーには学習者 ID とログアウトボタンが表示され (a)、フッターにはページ遷移のボタンがまとまっている (e)。画面は左右で構造が分割されており、画面の左側に文法などの説明とともに課題文が表示され、右側で学習者は回答のコードを入力したり、生成 AI からのフィードバックを閲覧する。課題提示の方法をより詳細に述べると、まず学習すべき文法知識が紹介され (b)、これに応じた例題 (c) と課題 (d) が表示される。学習者が対応する回答を入力して実行ボタンを押すと (f, g)、システムは Wasm を用いたコードの実行結果を返す (h)。さらに学習者がフィードバックの生成ボタンを押すと (i)、システムは論理 FB と情緒 FB を返す (j, k)。



図 2 フィードバック返答時の画面

3.5 動作例

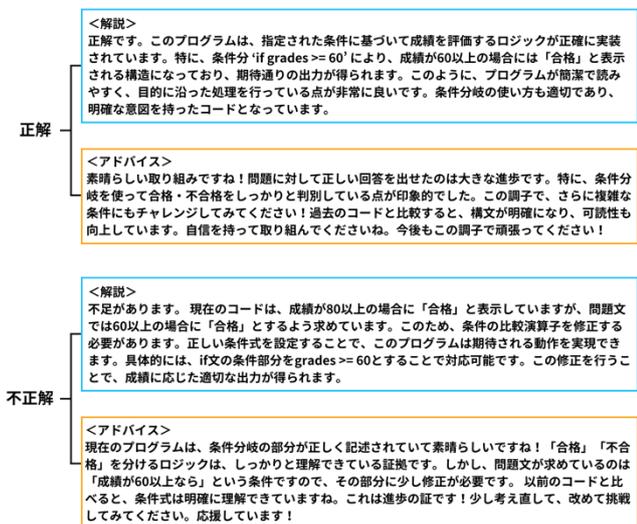


図 3 問題の正誤に応じたフィードバック内容

図 3 は問題の正誤に応じてフィードバックが生成されている画面である。青枠は論理 FB、オレンジ色の枠は情緒 FB を示す。

まず、回答が正解の場合、論理 FB については正誤のみ、もしくは具体的に正解に導いたロジックを再度表示する。情緒 FB については回答したコードの中でも具体的に良かった部分を表示し肯定的なメッセージを生成する。

次に回答が不正解の場合である。論理 FB では答えのコードを示さず、再考が必要な箇所のロジックの説明に留めている。情緒 FB については誤りを伝えた上で、中でも回答のコードの中から良かった部分、また過去のコードと比較して改善が見られた部分を表示し、ポジティブなメッセージを生成する。

4. 実験と結果

4.1 実験方法

前章に述べたシステムを用いて、論理 FB だけを表示する場合と、論理 FB と情緒 FB の両方を表示する場合との比較実験を行った。被験者はコンピュータ関連専攻の大学生 12 名である。実験は各自の PC で指定サイトにアクセスし、回答コードと、そこで返されるフィードバックを記録した。学生は全員 Python の基本を学習済みのため、実験における学習対象の言語は Ruby とし、被験者の募集にあたっては、「Ruby 言語の学習経験がないこと」を条件とした。

評価実験は一人あたり約 20 分で、12 人の被験者を無作為に 6 人ずつ、論理 FB のみ返される論理 FB 群と、論理 FB と情緒 FB の両方を返される論理・情緒 FB 群に分けた。論理 FB 群の場合は該当エリアの表示が空欄になるが、その他の違いはない。

被験者には Ruby の基本的な文法の説明に続き、それに関連した問題が表示された。各問題の学習内容は Ruby の公式ドキュメントを参考に基本的な制御構造を学ぶ問題を選定した。選定された学習内容は if (条件分岐)、for (繰り返し)、raise-begin (例外処理) である。被験者は本システムと Ruby の公式ドキュメントのみを参照可能な状態で実験に参加した。

約 20 分間の学習の事前と事後に、プログラミング学習に対するアンケートを行った。

4.2 結果

表 1 はプログラミングの意欲に関するアンケート項目、表 2 はその結果である。問 6 を除き、事前と事後に同一の質問を実施した。アンケートは 6 項目から構成され、10 段階のリッカート尺度で評価した (数値が大きいほど評価が高い)。初学者のプログラミング学習時に感情に関与する楽しさ、満足感、興味をたずね、対応のある t 検定およびウェルチの t 検定を用いた [13]。

アンケート結果では、論理 FB 群および論理・情緒 FB 群

の学習前数値を比較したところ、問 5 に統計的有意差が確認された (p 値 \approx 0.021, 有意水準=5%)。学習前・後の数値差分を比較した結果、全問で両群の差分に統計的有意差は見られなかった。一方、両群の学習後数値を比較した結果、新しいプログラミング言語学習に関する問 4 および Ruby を学ぶ意欲に関する問 5 で統計的有意差が確認された (問 4: p 値 \approx 0.035, 問 5: p 値 \approx 0.008)。ただし、問 5 は学習前数値の違いが影響している可能性が高く、学習条件による効果とは断定できない。また、同一群内の学習前後比較では、論理 FB 群は問 2, 論理・情緒 FB 群は問 2 および問 5 で有意な向上が確認された。問 4 については、学習条件の違いが論理・情緒 FB 群の数値向上に寄与している可能性が示唆される。しかし、両群の学習効果差に統計的有意差がないことから、学習条件の効果は限定的と考えられる。

被験者からは情緒 FB に関する肯定的意見が得られた。例えば「やる気を鼓舞するフィードバックを得られ、プログラミングが嫌いな自分でも学んでみたいという意欲が高まった」や、「間違えば他人からフィードバックを得るのは気を遣うため、生成 AI から取得できると気軽に学べる」といった意見があり、学習時の心理的障害の軽減が示唆される。

表 1 プログラミング意欲の調査項目

	設問
問 1	プログラミングの課題に挑戦する楽しさ(略称: 楽しさ)
問 2	プログラミング全般に関する自信(満足感 1)
問 3	プログラミング課題を終えたときの達成感(満足感 2)
問 4	新しいプログラミング言語学習への意欲(興味 1)
問 5	Ruby を学ぶことへの学習意欲(興味 2)
問 6	フィードバックによる Ruby の学習意欲の変化(効果)

表 2 プログラミング意欲の自己評価

(表の値は平均値, カッコ内は標準偏差を示す)

	論理 FB		論理・情緒 FB	
	学習前	学習後	学習前	学習後
問 1 (楽しさ)	4 (1.90)	7 (1.41)	5.5 (2.59)	8.17 (1.17)
問 2 (満足感 1)	4 (2.83)	4 (2.53)	4.16 (1.83)	4.5 (2.59)
問 3 (満足感 2)	6.33 (2.58)	7.17 (1.83)	7.33 (1.51)	8 (1.10)
問 4 (興味 1)	5.17 (2.48)	6.33 (1.03)	6.66 (0.82)	7.83 (1.47)
問 5 (興味 2)	4 (1.79)	5.33 (1.03)	6.33 (1.37)	7.5 (1.52)
問 6 (効果)	-	7.5 (1.04)	-	8.16 (1.17)

5. 考察

5.1 論理 FB 生成における Wasm 利用の必要性

前章の実験を行うにあたり、事前に Wasm を用いずに、学習者が提出したコードの正誤を生成 AI に直接判定してもらう試みも行ったが、一部で問題が発生したために、Wasm を利用するようにした。図 4 では、論理 FB の生成にあたり Wasm の実行結果を GPT 4o-mini に与えた場合、Wasm を利用せず直接 GPT 4o-mini に問題文と学習者の回答を与えてフィードバックを作成した一例を比較している。学習者が回答したコードは問題に正しい出力を行えるものであったにも関わらず、直接 GPT 4o-mini に判定を行わせたところ不正解と判定された。これは、プログラムの実行結果は正しいが、生成 AI が想定したアルゴリズムと異なるなどの理由で不正解判定になったと考えられる。また、例えば条件分岐を学ぶ問いで、プログラム中の `unless` という単語を GPT 4o-mini が自然言語として解釈し、誤った正誤判定や解説が提供される場合があった。Wasm の実行結果もプロンプトに含めた場合はこのようなことが起きなかった。実行結果もプロンプトに渡すことで、生成 AI が特定の解法のみを正解としたり問題文の誤認をしたりするリスクが軽減できると考えられる。

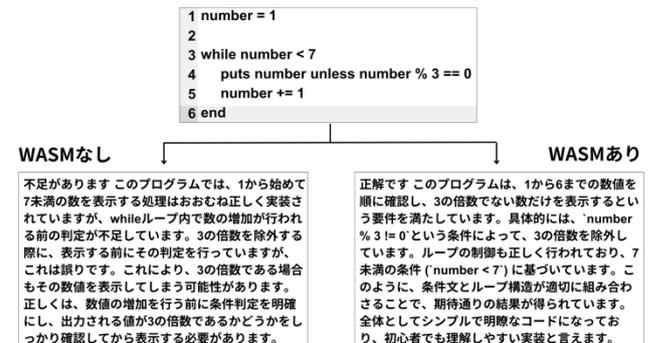


図 4 Wasm 使用有無による論理 FB の内容

5.2 情緒 FB と学習意欲への貢献

上記の実験では、生成 AI を用いて学習者にポジティブなメッセージを提供することが、学習意欲の向上に寄与する可能性が示唆された。フィードバックにより学習意欲が向上したかどうかを明示的に問うアンケート (問 6) の結果は、統計的に有意とは言えなかったものの、情緒 FB による意欲向上の傾向は見られた。しかし、初めて学習するプログラミング言語への意欲や、学んだプログラミング言語の学習を継続する意欲が向上した点は注目に値する。これらの結果は、生成 AI を活用したポジティブなフィードバックが、学習者の動機付けに一定の効果をもたらす可能性を示唆している。ただし、本実験では Ruby 言語の学習経験はないものの、プログラミングの一定の経験

がある被験者を対象としたため、全くのプログラミング初心者を対象とした場合は、やや異なる結果となる可能性がある。

6. 結論

生成AIを用いてプログラミングの学習を行う際に、学習意欲の向上につながるのか不明瞭であるという課題に対し、コードの解説と、取り組みに対してポジティブなメッセージを生成するという目的の異なるフィードバックを出し分けるシステムを開発した。ポジティブなメッセージを含むフィードバックを提供した場合、一部の学習者において学習意欲の向上が示唆されたが、全体としては統計的に有意な差が確認されなかった。また、被験者からの意見により、ポジティブなフィードバックが心理的障害の軽減に寄与する可能性が示された。今後は学習者の状況や学習段階に応じた精度の高いフィードバックを生成し、効果的な学習支援を行う方法を検討する必要がある。

謝辞

本研究は、東洋大学重点研究推進プログラムにより助成を受けたものです。同助成に感謝いたします。

参考文献

- [1] P. Denny et al., "Computing Education in the Era of Generative AI," *Commun of the ACM.*, vol. 67, no. 2, pp. 56–67, Jan. 2024.
- [2] Brett A. Becker, et al., "Programming Is Hard - Or at Least It Used to Be: Educational Opportunities and Challenges of AI Code Generation". in *Proc. of the 54th ACM Technical Symposium on Computer Science Education (SIGCSE)*, 2023, pp. 500–506
- [3] Kris M.Y. Law et al., "Learning motivation in e-learning facilitated computer programming courses", *Computers & Education*, vol.55, issue 1, pp.218-228, Aug. 2010.
- [4] S. Marwan et al., "Adaptive Immediate Feedback Can Improve Novice Programming Engagement and Intention to Persist in Computer Science," in *Proc. of the 2020 ACM Conference on International Computing Education Research (ICER)*. 2020, pp. 194–203
- [5] Michael J. Lee, and Amy J. Ko, "Personifying programming tool feedback improves novice programmers' learning," in *Proc. of the seventh international workshop on Computing education research (ICER)*, 2011, pp. 109–116
- [6] C. Ott et al., "Translating Principles of Effective Feedback for Students into the CS1 Context," *ACM Transactions on Computing Education*, vol. 16, issue 1, Feb 2016.
- [7] J. Leinonen et al., "Using Large Language Models to Enhance Programming Error Messages," in *Proc. of the 54th ACM Technical Symposium on Computer Science Education (SIGCSE)*. 2023, pp. 563–569
- [8] T. Phung et al., "Generating High-Precision Feedback for Programming Syntax Errors using Large Language Models," *16th International Conference on Educational Data Mining*. 2023
- [9] A. Singh et al., "Bridging Learnersourcing and AI: Exploring the Dynamics of Student-AI Collaborative Feedback Generation," in *Proc. of the 14th Learning Analytics and Knowledge Conference (LAK)*. 2024, pp. 742–748
- [10] M. Kazemitabaar et al., "CodeAid: Evaluating a Classroom Deployment of an LLM-based Programming Assistant that Balances Student and Educator Needs," *In Proc. of the 2024 CHI Conference on Human Factors in Computing Systems (CHI)*. 2024, pp. 1–20
- [11] H. Keuning et al., "A Systematic Literature Review of Automated Feedback Generation for Programming Exercises," *ACM Transactions on Computing Education*. vol. 19, issue 1, no. 3, pp.1-43, Mar 2019.
- [12] N. Bosch, and S. D’Mello, "The Affective Experience of Novice Computer Programmers," *International Journal of Artificial Intelligence in Education*. vol. 27, pp.181-206, Mar 2017.
- [13] G. Kanaparan et al., "Effect of Self-efficacy and Emotional Engagement on Introductory Programming Students," *Australasian Journal of Information Systems (ACIS)*, vol.23, Jul 2019