

C言語のエラーメッセージの出現頻度に基づいた学習支援の検討

A study of learning support based on the frequency of error messages in C language

森重 嘉優^{*1}, 井上 大地^{*2}, 松本 慎平^{*1}

Kayu MORISHIGE^{*1}, Daichi INOUE^{*2}, Shimpei MATSUMOTO^{*1}

^{*1} 広島工業大学 情報学部

^{*2} 広島工業大学大学院 工学系研究科

^{*1} Faculty of Applied Information Science, Hiroshima Institute of Technology

Email: {bm20124, s.matsumoto.gk}@cc.it-hiroshima.ac.jp

^{*2} Graduate School of Engineering, Hiroshima Institute of Technology

Email: md22001@cc.it-hiroshima.ac.jp

あらまし: 本稿では, エラーメッセージの教育支援への実用性を検証するために, 実際に使われている C 言語学習システムのデータを収集し, エラーメッセージの傾向を調査した. 今回, 実際の授業の中で用いられた演習問題を分析対象とした. 取り組んだ人数やソースコードの長さは問題ごとに異なるため, これらの要因の影響を除いた「ミス度合い」を問題ごとに算出した. その結果, ミス度合いが極端に高い問題は初歩的なエラーメッセージが多いことを確認した. また, 1つの問題においては1種類のエラーメッセージが多いことを確認した.

キーワード: プログラミング, エラーメッセージ, データ, アルゴリズム

1. はじめに

高等教育機関のプログラミング講義では, それを特に苦手と感じている学習者が多く存在している. この原因として, プログラミングは様々な能力や活動を必要とするために, 初学者にとっては学習の負担が大きいということが考えられている⁽¹⁾⁽²⁾. また, 初学者はタイプミス等による文法エラーが多発すると指摘されており⁽³⁾, 学習者によっては文法等の学習の負荷が大きくなる. その結果として, アルゴリズム等の学習が不十分となり, プログラミングに苦手意識を感じるようになると考えられている.

上述した点について, 仮に問題ごとのエラー度合い及びエラー傾向を明らかにすることが可能になれば, 教授者は, 受講者全体のエラー傾向を把握し, エラー傾向を基に学習者にアドバイスを行うことが可能となる. それにより, プログラミングを苦手とする学習者を減らせるのではないかと考えられる.

そこで本研究では, 先行研究⁽⁴⁾に基づいて, エラーメッセージによる教員の的確な指導の実現を目指し, 初学者のエラー傾向を可視化する. なお, 本研究の対象者は, 初歩的なプログラミング教育を受けている大学生とする.

2. 先行研究

Paul ら⁽⁴⁾は, プログラミングを学習している学生は, 数種類のエラーに頻繁に遭遇し, そのエラーの解決には多くの時間が必要であると考えているため, 的を絞った教育が学生の生産性の向上という点で大きな見返りをもたらす可能性があるとして述べられている. しかし, Paul らは, Java についての研究であり, C 言語とは言語が異なる. C 言語と Java で同様のエ

ラーの種類が多いことは明らかになっていない. また, Paul らは, エラーが出た理由について明示していない. そこで, C 言語という, より初学者が多い言語で, エラーが多い理由まで調査することで, 初学者の教育支援になると仮説が立てられる.

3. 本研究の準備

本研究を行うにしたがって, 「使用するツール」「分析するデータ」「可視化の方法」を事前に決めることが必要になる.

3.1 使用するツール

今回は, Hello C と呼ばれるシステムを採用する⁽⁵⁾. Hello C とは, C 言語初学者のための C 言語学習システムである. サーバアンドクライアントシステムで, プログラミング初学者が容易に学習を行うことが可能である. また, 学習者のログを収集するロギング機能がある. 学習者は, 学習中にコンパイルを行うごとにログデータを保管し, 解答提出時にデータベースにデータを送信する. 得られるデータとして, コード, エラー内容, 日時, 正誤などがある.

3.2 分析するデータ

今回は, インフォテックサーブ社の「はじめてのアルゴリズム」¹を基に作られた問題の演習ログを分析対象とした. 教科書は全6章で構成され, 章が進むほど難易度が上がる. 今回, 234人の学習者から得た204問の演習ログを分析対象とした. なお, 学習者は, 情報学を学ぶ大学1・2年生である.

¹

https://www.infotech-s.co.jp/teaching_material/45_firs_t_algo.html

3.3 可視化の方法

本研究では、問題ごとのエラーの傾向を可視化し、その中から特徴がある箇所のエラーメッセージの種類、出現回数を調べていくことを考えている。可視化を行う際、エラー数をそのまま度数として用いた場合、利用者が多い場合や難易度が難しい場合、度数が高くなるのは当然であるため、可視化として相応しくない。そこで、人数や難易度の関係性をなくした度数 D を式 1 のように表す。

$$D = \frac{E}{SC} \dots \dots \dots (1)$$

ここで、 E はエラー数、 S は使用した人数、 C は空白以外のソースコードの文字数である。今回、ソースコードの長さを難易度と考えることができる。以下、この計算で出した度数をミス度合いとする。

4. 結果

全ての問題のミス度合いを可視化したものを図 1 に示す。x 軸は、3.2 で示した各問題を、y 軸は、問題のミス度合いを示している。1 つの問題のミス度合いが著しく高いことが示唆された。図 1 だけでは、他の問題の傾向が掴めないため、ミス度合いが著しく高い問題を除いたミス度合いの可視化を図 2 に示す。見方は、図 1 と同様である。

図 1、図 2 の中から、ミス度合いが高い問題ごとのエラーメッセージの種類を調査した。その結果、図 1 では、「initialization makes integer」が、図 2 では、「stray」と「expected」のどちらかが著しく高い問題が多いということを確認した。「initialization makes integer」は、記憶域が大きいものを記憶域が小さいものに代入すると値が消えるという警告である。「stray」は、文字列やコメント以外の箇所に全角文字を使用した際に出てくるエラーである。「expected」は、シンタックスエラー(セミコロンやカッコを付け忘れるなど)の際に出てくるエラーである。そこから、ミス度合いが高い問題は、初歩的なエラーが多いことを確認した。また、図 2 のミス度合いが高い箇所が広範囲に散らばっていることも確認した。問題の難易度が上がると、アルゴリズムに関するエラーすると一般的に思われているが、初歩的なミスが多いことが明らかとなった。その原因としては、初歩的なエラーメッセージは、一か所のミスで多く出現することが考えられる。例として、ある 1 つの変数を日本語にした場合、その変数を使うごとに、「stray」のエラーメッセージが出現してしまうことが挙げられる。このように、エラー数が難易度と比例していないことが可視化できた。

併せて、「stray」が著しく高い問題の出題方法の調査を行った。問題の出題方法の 1 つとして、問題文と共にアルゴリズムも記載されており、そのアルゴリズムの変数が日本語で記入されていた。そこから、プログラムとは別の要因によりエラーを出してしまう初学者が存在することを確認した。その原因として、初学者の中には、何かヒントになるものに縋り

たくなる人が存在し、その結果、アルゴリズムをそのまま鵜呑みにし、エラーを出してしまうことが考えられる。このように、教授者が良かれと思った行動に初学者のエラーを促進させる可能性があることが可視化された。

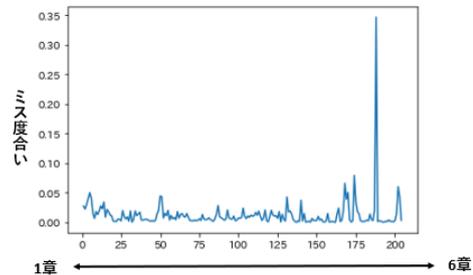


図 1 全体のミス度合いの結果

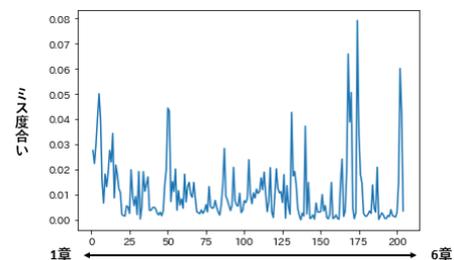


図 2 ミス度合いが高い問題を除いたミス度合いの結果

5. おわりに

本研究では、実際に使われている C 言語学習システムのデータを収集し、エラーメッセージの傾向を調査した。その結果、エラーメッセージは教育支援に利用可能であることを確認した。

謝辞

本研究は、独立行政法人日本学術振興会科学研究費助成事業(基盤研究(C)20K0319, 22K02815)の助成を受けて実施した成果の一部である。ここに記して謝意を表します。

参考文献

- (1) Lisack S. K.: "Helping Students Succeed in a First Programming Course: A Way to Correct Background Deficiencies." International Association for Computer Information Systems Conference, Mexico (1998)
- (2) Garner S.: "A Tool to Support the Use of Part-Complete Solutions in the Learning of Programming." Proceeding de conference, pp.222-228 (2001)
- (3) 岡本雅子, 喜田一: "プログラミングの"写経型学習"における初学者のつまずきの類型化とその考察." 実践センター紀要, Vol.22, pp.49-53 (2014)
- (4) Paul D, Andrew L. R., Ewan T: All Syntax Errors Are Not Equal (2012)
- (5) 松本慎平, 大下昌紀, 買田康介, C 言語初学者及びその教授者のためのサーバ・クライアントに基づくプログラミング学習支援システムの開発, 電気学会論文誌 C, Vol.140, No.9, pp.1096-1109 (2020)