

プログラムの人に近い採点のための考察及び実装

阪 春輝*1・北 英彦*1・高瀬 治彦*1

Email: 418m234@m.mie-u.ac.jp

*1: 三重大学大学院工学研究科

◎Key Words eラーニング、プログラム作成、自動採点

1. はじめに

情報化社会の発展に伴って、プログラミング教育の必要性が高まっている。学生にプログラミング技能を習得させるためには、様々なプログラムを数多く作成させることが重要である。しかし、講義では演習時間が足りないため、様々なプログラムを作成できていないという現状がある。講師やTAの数に対して学生が多い場合が多く、学生が書いたプログラムひとつひとつに対して講師がコメントを含めたフィードバックを返すとなると講師への負担が大きく現実的ではない。

プログラミング演習の方法として多くの大学が利用する、eラーニングシステムの小テスト機能を用いる。小テストを取り入れることにより演習以外でも学生に多くのプログラムを作成させることができる。しかし、ソースコードのような表現の自由度が高い記述式問題における学生の解答を自動採点することは困難であり、講師による手動採点への負担が増えるといった課題があった。

著者の所属する研究室の四方は、ソースコードを解答とする問題に対して軽度な動作エラーを自動で修正・補完する自動採点システムを考案¹⁾し、eラーニングシステムプラットフォームのMoodleに取り入れた。初学者が行いがちな軽微な間違いに対して修正を行い、部分点を与えることで初学者が学習しやすいシステムとなった。

しかし、解決すべき課題として、宣言した変数を使用していない場合や不必要なものが出力された場合でも、必要な出力が存在すれば満点になることがあげられる。

本研究では、解答に必須ワード、必罰ワードを規定することで学生が最適な記述を行っているかの判断を行い、自動採点の結果を人が採点したときの結果に近づけることを目的とする。これにより、講師は学生に対して多くのプログラム作成の問題を課すことができ、学生は多くのプログラムを作成することが可能になる。また、電子教科書を用いた自動採点の新たな活用方法を提案する。

2. 関連研究

本システムと同じようにプログラミング教育を行う上で自動採点に焦点を当てた研究はいくつか行われている。柳田らは、プログラムとトレース表を用い様々な難度の1行の組み合わせ穴埋め問題を作成し、解答後は自動採点を行い学習者に採点結果を示すシステム pgtracer を開発した²⁾。pgtracer は学生の解答と、正解のプログラムのコンパイルと実行結果の比較と、入力された文字列に対応するノードの値との比較という2通りの比較で正誤判定を行うことで自動採点を行っている。

しかし、コンパイルがエラーになった場合には結果が

表示されないため0点となってしまう、人による採点結果と比べると採点が厳しい。また、プログラミングのスタイルについて見ていないといった問題点がある。

3. 四方の自動採点システム

著者の所属する研究室の四方が作成した自動採点システムについて説明する。三重大学電気電子工学科ではJavaを用いたプログラミングを学んでいるので、対象はJavaで書かれたソースコードとする。ソースコードを解答とする小テストの目的のひとつは学習者に数多くのコードを書かせることにあるので、ソースコード全体を解答させるのではなく、小テストの問題の出題意図であるコアの部分のみを穴埋め問題形式で解答させる。また、採点基準は望月³⁾の採点基準を用いて採点を行っている。

3.1 採点項目

講師が採点の時に確認していると思われる以下の4つの項目に関して自動採点を行う。括弧内は配点の標準の割合であり変更できる。

- 構文エラーの有無 (30%)
- 動作の正しさ (60%)
- 課題で指定された機能の使用(上記60%の内10%)
- プログラミングスタイルの適切さ (10%)

3.2 四方の自動採点システム運用結果

四方の自動採点システムを用いた小テストを2018年度の三重大学電気電子工学科プログラミング演習の受講者89名を対象に実施した。ある課題における得点分布を図1に示す。結果は図のように大きく3群に分かれ、群に属さないものは少数という結果となった。これらの群はそれぞれ、コンパイルができプログラムが正常に動作した群、コンパイルはできたが動作が正しくなかった群、コンパイルができなかった群の3つと考えられる。

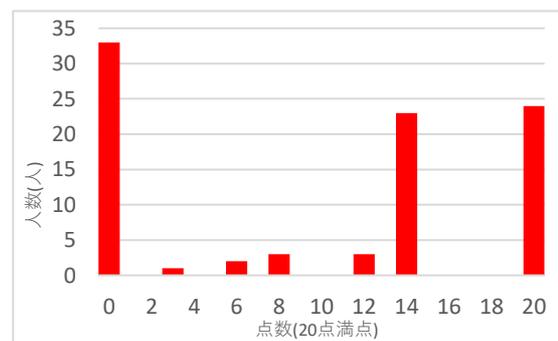


図1 四方の自動採点システムの得点分布

3.3 望月の採点基準の問題点

四方の自動採点で用いられている望月の採点基準の問題点を見つけるために次のことを行った。2016年度のプログラミング演習の受講者89名の解答のうち、満点でも0点でもない解答10個を無作為に抽出し、プログラムを理解している講師4人の採点と自動採点の結果を比較したところ、表1のような結果になった。望月の採点基準は、入力動作がないプログラムでは、解答の出力文に必須文字列が存在しているかで、入力動作があるプログラムではテストケースを複数用意し、学生の書いたプログラムの出力部とテストケースが一致するかで動作を確認している。しかし、この方法では出力文に不必要な値が入っている場合や、定義されている変数を使用しないといった場合でも満点になるという欠陥が存在する。例えば、2数の和を求めるプログラムにおいて、問題文中で計算用の変数 `sum` を定義してあっても、使用せずとも要求を満たすプログラムを作ることはできるが、講師が採点を行った場合、その解答に満点をつけることはないだろう。従って、望月基準は人間の基準と比べて甘いといえる。

表1 自動採点と人の採点の比較

	講師 A	講師 B	講師 C	講師 D	自動採点
(1)	10	2	0	10	8
(2)	15	10	10	15	14
(3)	13	15	10	5	12
(4)	5	15	0	5	8
(5)	10	10	10	1	14
(6)	0	0	0	10	6
(7)	13	15	10	5	14
(8)	10	13	0	15	0
(9)	10	10	0	10	9
(10)	8	10	5	6	14
平均	9.4	10	4.5	8.2	9.9

4. 自動採点の改良

本研究では自動採点で以下のような機能を実装することで、より人に近い採点を行うことを考えている。

- 解答のプログラムの必須ワードの指定
現在の自動採点は出力文の必須ワードを指定し、そのワードとの一致度で動作の採点を行っているが、プログラム内にも必須ワードを課すことで、宣言した変数の使用・未使用を判断し、採点を行うことができる。
- 解答のプログラムの必罰ワードの指定
現在の自動採点では指定された機能の有無は判別できるが、値の条件には対応していない。例えば、自然数を表示するプログラムで-1から表示されても、1以降がきちんと表示されれば正解となるが、これは出題者の要求を満たしていない。従って、必罰ワードを設けることで、出題者の意図に沿っていないプログラムに関しても減点を行うことができる。

また、それ以外の自動採点の活用方法として、以下のような使用方法を考案する。

- 複数穴埋め問題の設定

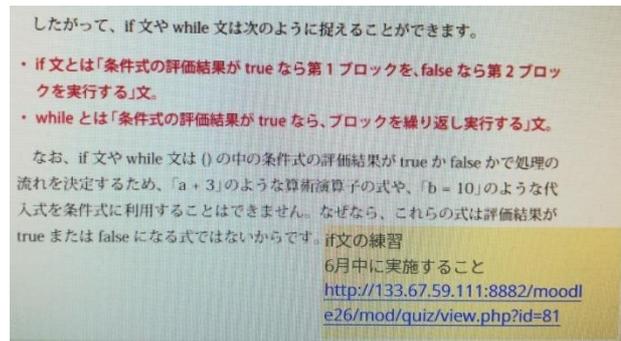


図2 電子教科書を用いた例

現在のシステムでは、プログラムに対して、1行の穴埋めを複数配置する形か、数行程度を1ヶ所に配置するといった形だが、実際にアプリなどのプログラムは大規模なプログラムが相互に作用しながら動作している。そのため、1つのプログラムに対して複数の穴埋め問題の実施は、プログラム全体の動きを見る視野の育成にもつながる。

- 電子教科書と連携した出題

教科書の練習問題は学習者の理解度を図るために使用できるが、プログラムの練習問題は自分で採点しても理解できていない箇所を見つけるのは難しい。そこで、図2のように電子教科書の付箋機能を用いて、学習者の学習状況に合わせて、小テストを受講させ、その解答に対して自動採点を行う。そこで間違っていた部分のフィードバックを行うことで、学習者が自分の学習状況を正しく把握することができる。

5. まとめ

学生がプログラミング技能を向上させるための演習時間が足りないという現状がある。そこで、多くの大学でも導入されているeラーニングシステムの小テスト機能をプログラミング科目に導入することで時間不足の問題を解決することを考えた。しかし、ソースコードのように表現の自由度が高い記述式問題の解答を自動採点することは困難であり採点における講師の負担は非常に大きなものとなってしまふ。

本研究は、学生が書くソースコードに問題ごとのルールを指定することで、自動採点の結果を人による採点に近づけることを目的とした。また、新たな自動採点の活用方法を提案することで、小テストの問題が意図している部分の理解度を測ることができるようになると思われる。

参考文献

- (1) 四方雅晴、北英彦：プログラムのソースコードを解答とする問題の自動採点、コンピュータ利用教育協議会、CIEC 夏季研究会 2017 (2017)
- (2) 柳田峻、太田康介、大月美佳、掛下哲郎：穴埋め問題を用いたプログラミング教育支援ツール `pgtracer` の運用実験、情報処理学会、情報処理学会論文誌教育とコンピュータ (TCE) 2(2), 20-36, 2016
- (3) 望月将行、森田直樹、高瀬治彦、北英彦、林照峯：自動テスト機能を備えたプログラミング演習支援システム、日本教育工学会、日本教育工学会大会講演論文集 20, 677-678, 2004