

# 初心者を対象とするプログラミング能力評価ツール作成のための 評価基準の提案

小宅 直樹<sup>†</sup> 山内 大輔<sup>††</sup> 鈴木 裕利<sup>†</sup> 石井 成郎<sup>‡</sup>

<sup>†</sup>中部大学大学院工学研究科 <sup>††</sup>中部大学工学部 <sup>‡</sup>愛知きわみ看護短期大学

## 1. はじめに

中部大学工学部情報工学科では、プログラミング教育の一環として、PAD(Problem Analysis Diagram)を導入した授業を進めている。PADは、アルゴリズムの理解の促進と、プログラムの理解度の判断に有効と考えるからである。これは、先行研究において、PADの理解度とプログラムの理解度の相関が観察された報告に基づく[1]。一方、学習者が、PADを作成する際に、表記法の誤り等、理解の不十分な点も多く確認されている[1]。そこで我々は、これらの知見を踏まえて、アルゴリズムの理解を促進するためのPAD作成支援ツールの開発を進めている。また、学習者が作成するPADは紙媒体で、評価方法も手作業であり暫定的な基準によって評価を行っている。そこで、本研究ではPAD評価ツールに実装可能であり、実験によるデータに基づいた評価基準の提案を進めている。本稿では、評価基準のために実施した評価実験について報告する。

## 2. 評価基準の提案

現在の暫定的な評価基準を表1に、実験のために作成した新評価基準を表2に示す。

PAD作成支援ツールで作成したPADは、現評価基準の分類である「線」が不必要になるため、新評価基準では「配置」に変更している。また、評価者が異なる場合に評価結果が異なることを避けるために、新評価基準の内容を現評価基準よりも詳細になっている。

表1：現評価基準

分類	内容
正解	正解
線	線の書く場所が間違えている
要素	中身の書き方が間違えている
形	形を間違えている
比較	”if”の部分を”while”で書いている
T/F	”if”の部分を”while”で書いている
複数	1つのPADで書かれていない
その他	上記以外の間違い
未記入	未記入

表2：新評価基準

分類	内容
正解	問題文から最低限のキーワードや計算式を判断できる
配置	PAD部品の位置やPAD部品同士 のつなぎ方の間違い
要素	PAD部品の中身の定義の間違い
形	PAD部品の使うものの間違い
比較	比較の必要な所で使えていない
T/F	比較の部品のT/Fの条件の間違い
不等号	不等号が逆などの間違い
複数	PADを複数書いている
その他	上記以外の間違い

## 3. 評価実験

本章では、提案した新評価基準で複数の評価者が評価した場合の一致率を観るために行った実験について報告する。

本実験では、PAD評価者2名がPADデータA(115件)、PADデータB(138件)を分類の例を見ながら評価を行う。評価結果の例を図1に示す。

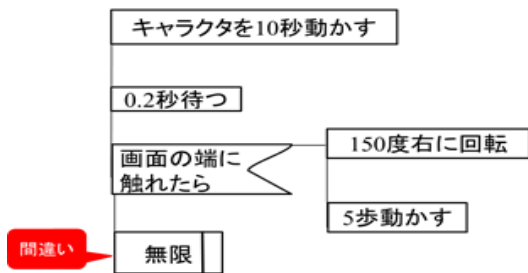


図1：評価結果の例（配置）

4. 分析

評価者 2 名の評価結果の一致度を判断するため、分類ごとに k 係数を算出する。また、k 係数が 0.60~1.00 であった場合、高い一致度であると判断する。

PAD データ A の評価の一致度を表 3 に、PAD データ B の評価の一致度を表 4 に示す。

表 3：評価の一致度（PAD データ A）

分類 \ 一致度	見かけの一致率	偶然による一致率	k 係数
正解	0.96	0.53	0.91
配置	0.96	0.86	0.68
要素	0.97	0.67	0.90
形	0.99	0.86	0.94
比較	1.00	0.89	1.00
T/F	1.00	0.98	1.00
不等号	1.00	1.00	0.00
複数	0.99	0.96	0.80
その他	1.00	1.00	0.00

表 4：評価の一致度（PAD データ B）

分類 \ 一致度	見かけの一致率	偶然による一致率	k 係数
正解	0.99	0.51	0.97
配置	0.99	0.64	0.98
要素	0.99	0.64	0.98
形	0.96	0.85	0.76
比較	0.99	0.95	0.85
T/F	0.99	0.99	0.66
不等号	1.00	1.00	0.00
複数	1.00	1.00	0.00
その他	1.00	0.99	1.00

表 3 では「不等号」と「その他」、表 4 では「不等号」と「複数」が見かけの一致率と偶然による一致率が 1.00 で k 係数が 0.00 となっていることがわかる。また、それ以外の k 係数が 0.60 以上であることが確認された。

5. まとめ

本研究では、PAD 作成支援ツールに合わせた PAD 評価ツールを作成するにあたって、新評価基準を提案し、その評価基準が複数の評価者でどの程度一致しているか確認するために実験を行った。

前述したように、表 3 では「不等号」と「その他」、表 4 では「不等号」と「複数」の見かけの一致率と偶然による一致率が 1.00 で k 係数が 0.00、それ以外の分類の k 係数が 0.60 以上であることが確認された。

見かけの一致率と偶然による一致率が 1.00 で k 係数が 0.00 である「不等号」の分類は、本実験で評価した PAD データが「不等号」を扱うデータではなかったためである。また、表 3 の「その他」、表 4 の「複数」は評価者 2 名が 1 度も評価しなかったケースである。このケースを除いて、0.60 以上となっている。

結果からは、新評価基準は複数の評価者の評価結果の一致が一定の割合で確認できた。

今後は、PAD 作成支援ツールと PAD 評価ツールの開発を進め、完成次第、新評価基準を実装して実験を行う予定である。

謝辞

本研究は、JSPS 科研費基盤研究 (C) 26350204 の助成を受けている。

文献

[1] 鈴木裕利, 小宅直樹, 山内康一郎, 奥居哲, 石井成郎, “初心者向けプログラミング教育における PAD の導入とその効果,” 電子情報通信学会技術研究報告. 教育工学研究会 (名古屋工業大学), 2016.