

チェックシートを用いた相互採点法における質問項目の評価

古作 創[†]

木更津工業高等専門学校 情報工学科[†]

大枝 真一[‡]

木更津工業高等専門学校 情報工学科[‡]

1. 研究概要

先行研究により、プログラミングの学習において他人のソースコードを読むことは自身のスキル向上に繋がることが示されている [1]。これにより、相互採点では単に採点手法としてだけでなく、採点者に対して学習効果も期待できるといえる。

そこで、学習の一環として相互採点を導入した授業形態の提案を行う。相互採点の際には、初学者も熟練者も同じ観点で採点を行うことができるようにチェックシートを用いる。また、項目反応理論を使用して、チェックシートの質問項目の難易度の差を定量的に評価する。

2. 提案する授業形態

現在、情報工学のプログラミングの授業でよく利用されている授業形態は、実施されたテストを教師が採点し、採点結果をもとに学生の評価を行うものである。提案する授業形態では、テストの実施後、回収した学生のソースコードをランダムに配布し、学生同士で相互採点を行う。学生が正しく採点できているかを、相互採点結果をもとに教師は評価を行う。

現在の授業形態を図 1、提案する授業形態を図 2 に示す。

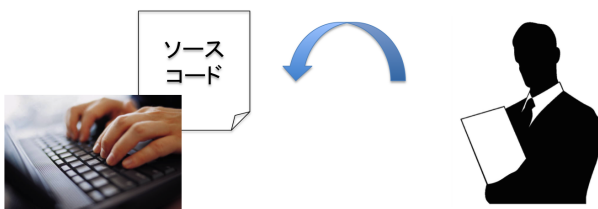


図 1 現在の授業形態

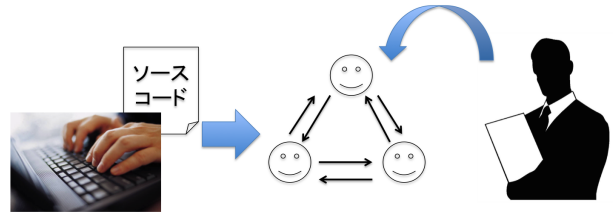


図 2 提案する授業形態

3. チェックシートを用いた相互採点

絵画や書道、音楽などの芸術分野の作品は、専門的なスキルを所持していないとどのような観点で評価を行えばよいのかわからないため、作品に対しての評価を行うことは難しい。これはプログラミングのソースコードレビューにも当てはめることができ、初学者ではソースコードの良し悪しを評価することができない。提案する授業形態では、ソースコードに対する相互採点の際にチェックシートを用いる。チェックシートを用いて採点を行うことで、教師の採点基準が明確化されるため、学生は教師と同じ視点で採点を行うことができる。共通のソースコードに対して、学生と教師がチェックシートを用いて採点を行う。学生と教師の採点結果を比較し、項目が一致すれば正しく採点できたとみなして加点をする。チェックシートを用いた相互採点を図 3 に示す。図 3 では、1 問 1 点で評価を行っている。

しかし、チェックシートの質問項目の中には難易度に差があり、高いスキルを必要とする項目と、そうでない項目がある。そこで、質問項目の難易度を定量的に評価するために、項目反応理論を用いる。

4. 項目反応理論

項目反応理論 [2] とは、試験に対する反応データから試験項目の難易度・識別力、被験者の特性・能力を測定するための試験理論であり、試験項目の難易度や識別力から試験の設問を定量的に評価することができる。

Evaluation of Question Difficulty in Check Sheet for Peer Review
[†]Hajime Kosaku · National Institute of Technology, Kisarazu College
[‡]Shinichi Oeda · National Institute of Technology, Kisarazu College

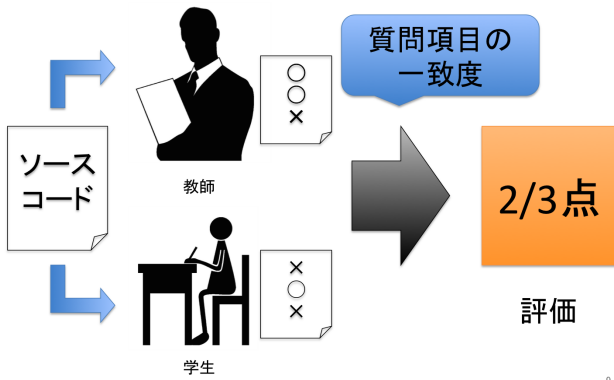


図3 チェックシートを用いた相互採点

本研究では、試験項目でなくチェックシートの質問項目に対して項目反応理論を適用した。被験者の能力に対する正答確率を表現するために、正規累積モデルの近似式であるロジスティックモデルを用いた。本研究で使用した、1母数ロジスティックモデルは式(1)で表される。

$$p_j(\theta) = \frac{1}{1 + \exp(-D \times (\theta - b_j))} \quad (1)$$

θ は被験者の能力値、 b_j は質問項目 j の難易度、 $p_j(\theta)$ は質問項目 j に対する能力値 θ を持つ被験者の正答確率を表す。項目の難易度の b_j の数値が高いと難しい項目であることを表す。 D は尺度因子であり、 θ の全ての値に対しての正規累積モデルとロジスティックモデルの誤差が、0.01 以下になることから 1.7 に定められている。

5. 評価実験

授業教材としての相互採点の有効性を示すために評価実験を行う。木更津工業高等専門学校情報工学科2年生41人を対象に質問項目10項目のチェックシートを用いて実験を行った。チェックシートの質問項目を表1に示す。学生を模して作成された共通のプログラムのソースコードを読んでもらい、チェックシートをもとに相互採点を行った。学生の中の採点結果に対して項目反応理論を用いて、質問項目の評価を行う。

実験結果を表2に示す。項目1,2,4は難易度が低い項目であることがわかる。これはソースコードの中身にはあまり関係のない項目であったため、難易度が低くなったと考える。逆に項目9は全体を通しての質問であること、抽象的な質問であったこ

表1 チェックシートの質問項目

1	インデントがずれている。
2	間違いは三箇所である。
3	set_string関数は正しく動作する。
4	main関数は読みにくい。
5	count_alphabet関数はセンスを感じる。
6	count_alphabet関数のように記述しないほうが良い。
7	初期化は適切である。
8	変数名、関数名は適切である。
9	このプログラムは良く出来ている。
10	disp_count関数は正しく表示されない。

表2 実験結果

項目	項目難易度 b
1	-3.12
2	-3.70
3	-1.09
4	-3.66
5	-0.94
6	-0.55
7	-0.07
8	-1.08
9	0.16
10	-1.55

とから難易度が高くなったと考える。

6. まとめ

本研究では、チェックシートを用いた相互採点を導入した授業形態の提案とチェックシートの質問項目の評価実験を行った。今後は、項目反応理論を用いて実験を繰り返し、教材としてよりよいチェックシートの作成を行う。

謝辞 本研究はJSPS 科研費 25750095 の助成を受けたものです。

参考文献

- [1] 菊池祥平, 大枝真一, “相互採点を用いた能動的評価法の開発”, 第77回全国大会(情報処理学会), 2ZF-06, 2015.
- [2] 豊田秀樹, “項目反応理論 [入門編]-テストと測定の科学-”, 朝倉書店, pp.i-29, 2002.