

プログラミング教育における IRT を用いたチェックシートの最適化と評価 Optimization and Evaluation of Check Sheet by IRT for Programming Class

古作創[†]
Hajime Kosaku

大枝真一[‡]
Shinichi Oeda

1. まえがき

絵画や書道、音楽などの芸術分野では、自身の技術向上を目的として他の作品を鑑賞することがある。他の作品を鑑賞することにより熟練者の作品の良い点を学び、自身の作品に取り組むことで技術向上、作品の品質向上に繋がる。先行研究において、プログラミングにおける他人のソースコードを読むことがスキルの修得に効果的であることが示されている [1]。そこで、プログラミング教育における学習の一環として、ソースコードレビューを導入する。

本研究では、プログラミング教育にコードレビューを取り入れる際、評価基準を統一するためにチェックシートを用いることを提案する。チェックシートを用いることで、初学者が熟練者の視点に立って評価を行うことが可能となる。また、チェックシートの項目は洗練されるべきであるため、各項目の品質調査を行う必要がある。そこで、通常試験項目に用いられる IRT (項目反応理論) をチェックシートの項目に適用し、項目の評価と最適化を行う。

2. 他分野でのチェックシートの活用

スポーツの分野では、練習の中にチェックシートを導入することで技術の向上を図ることがある。具体的な例として、バレーボール [2, 3]、卓球 [4]、バスケットボール [5] がある。バレーボールでは、準備姿勢や移動、頭の位置など基本的な動作を行うために必要な技術をチェック項目とし、評価者用 (教師) および実践者用 (学生) のスキルアップチェックシートの形式をとったものがある [2]。卓球では、よりレシーブが困難になるサービスの習得のためのチェックシートを作成し、練習時に使用している [4]。バスケットボールでは、シューティング時に動画を撮影してチェックシートを使いながら自身のフォームについて自己評価を行った [5]。

教育分野においては、授業後の振り返りシートとして自身がその日の講義の内容を習得できたかどうかをチェックするものとして使用している例 [6] や、ネットワークリテラシー教育における学習教材としてチェックシートを導入している例 [7] もある。先行研究 [7] では、Web ページを評価するための学習教材としてチェックシートが導入されていた。学習教材として用いることで、チェックシートを評価を行うためだけのものではなく、今後自身が Web ページを作成する際の道標として使うことを目的としている。

医療の分野では、養成校の学生が体外循環操作訓練

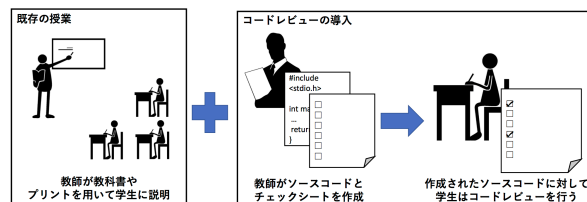


図 1: 既存の授業にコードレビューを導入した授業

の際に一連の流れを示すものとして使用し、指導者の存在に依存することなく訓練者相互でトレーニングし評価を行うためのツールとして用いていた [8]。

他分野においては、具体的な内容の項目を使用していることが特徴として見られた。しかし、スポーツにおいてはスキルの修得に対して明確な答えが存在することが多いが、プログラミングのソースコードには絵画や書道と同じように明確な答えが存在しない。そのため、本研究では明確な答えが存在しない分野でも評価を行うことができるような項目に設定した。

3. コードレビューを用いたプログラミング教育

既存のプログラミング教育は、教師が教科書やプリントを用いて説明を行い、その後演習に取り組むといった形態である。先行研究 [1] によりコードレビューの有用性が示されたことより、本研究では既存の授業に他のソースコードに対してのコードレビューを導入した新しい学習方法を提案する。熟練者のソースコードを読むことによって、学生は熟練者のソースコードの良いと思う点を自主的に取り入れる。これにより、学生のソースコードは熟練者と同じようなソースコードに近づき質の高いソースコードになるため、学生のスキル修得、技術向上に繋がると考える。

コードレビューの際には、教師と学生が評価基準を統一するためにチェックシートを導入する。チェックシートを導入することによって、学生はコードレビューのための評価基準としてだけでなく、今後自身がプログラミングを読んだり作成したりするための指針にすることができる。教師はコードレビューを行うソースコードと対応したチェックシートを作成し、学生がコードレビューを行ったあと教師と学生のチェックシートを比較し一致率を調べる。学生と教師の回答が一致していた場合、学生が教師と同じスキルを所持していると判断し正答とする。ここでの採点結果は学生の成績に反映させるためのものではなく、コードレビューを行うことによる学生の技術向上を目的とする。チェックシートを導入した際の授業の概要を図 1 に示す。

チェックシートの中には、評価項目として有用な項目とそうでない項目がある。チェックシートの項目は

[†]木更津工業高等専門学校 制御・情報システム工学専攻, Advanced Course of Control and Information Engineering, National Institute of Technology, Kisarazu College

[‡]木更津工業高等専門学校 情報工学科, Department of Information and Computer Engineering, National Institute of Technology, Kisarazu College

洗練されるべきであるため、各項目の品質調査をして項目の精査を行う。

4. IRT(Item Response Theory)

IRTは日本語では項目反応理論と呼ばれ、試験項目に対する反応データから試験項目の難易度や識別力、被験者の特性を測定するための試験理論である。IRTを用いることで、試験項目を定量的に評価を行うことができる。試験項目や被験者のパラメータを推定するときに使用する反応データは、2値(正答なら1、誤答なら0)の正誤データを用いる。欧米では幅広く使用されており、日本でもTOEICやTOEFL、情報処理技術者試験、日本語能力試験などに使用されている[9]。

本研究では本来試験項目に使用するIRTを、コードレビューの際に用いるチェックシートの質問項目に対して使用し評価を行う。

4.1. 2母数ロジスティックモデル

被験者の項目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使用する。本研究では式(1)に示す2母数ロジスティックモデルを使用した。

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

θ は被験者の能力値、 b_j は質問項目 j の難易度、 a_j は質問項目 j の識別力、 $p_j(\theta)$ 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。 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。

被験者の能力と被験者の項目に対する正答確率の関係を表したものととして項目特性曲線がある。項目特性曲線は項目における難易度と識別力によって曲線が変化する。難易度が変化すると曲線が横方向に平行移動し、識別力が変化すると曲線の傾きが変化する。難易度と識別力が変化したときの曲線の変化を図2,3に示す。

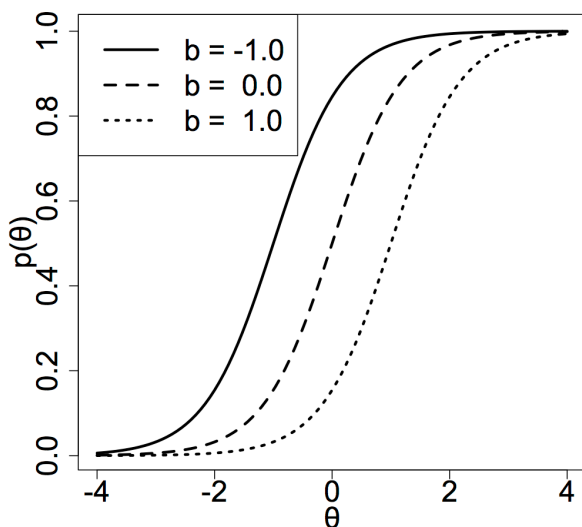


図2: 難易度 b_j を変化させたときの項目特性曲線

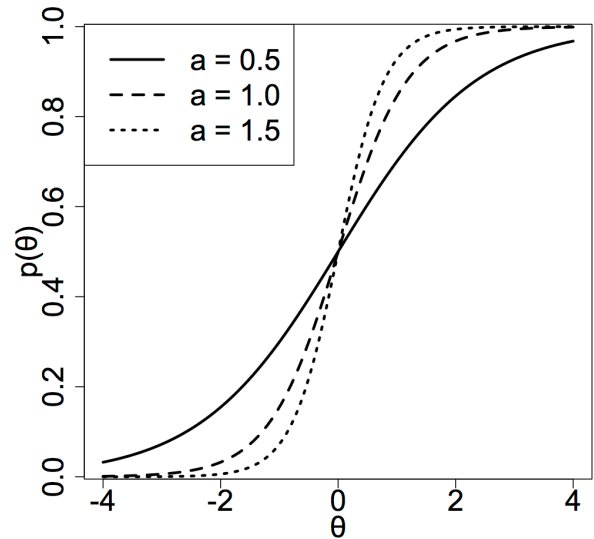


図3: 識別力 a_j を変化させたときの項目特性曲線

4.2. 情報関数

テストでの測定の精度を測り、適切なテストであるかどうかを判断するために情報量を用いて項目の評価を行う。IRTでは評価を行うための情報量としてフィッシャー情報量を用いる。フィッシャー情報量は式(2)で定義され、 $q_j = 1 - p_j$ とすると2母数ロジスティックモデルでは式(3)に変形される。

$$I(\theta_i) = E \left[\frac{\partial}{\partial \theta} \log L(\mathbf{u}_i | \theta) \Big|_{\theta = \theta_i} \right]^2 \quad (2)$$

$$I_j(\theta_i) = D^2 a_j^2 p_j(\theta_i) q_j(\theta_i) \quad (3)$$

項目 j に関しての情報量を表したものを項目情報量と呼ぶ。被験者の能力 θ を母数とした項目情報量を項目情報関数(Item Information Function: IIF)と呼ぶ。項目情報量が大きいほど測定精度が高いことを表し、質が高い項目であることを表す。難易度、識別力が変化したときの項目情報関数を図4に示す。

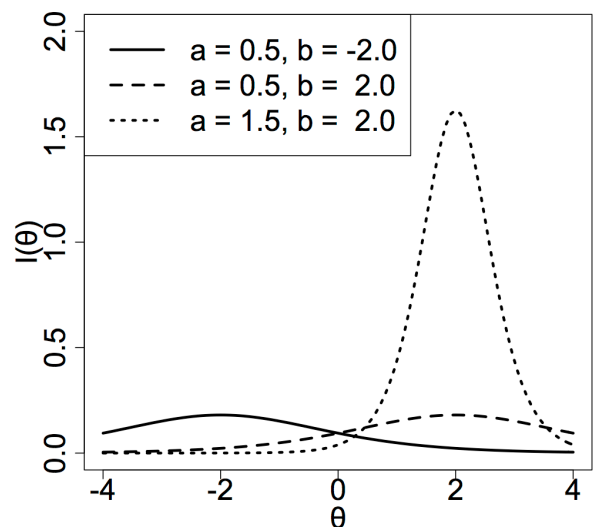


図4: 難易度、識別力を変化させたときの項目情報関数

被験者にとって 50%の確率で問題に正答するとき、項目情報関数の頂点は高くなる。項目情報関数の頂点は、難易度 b_j が低いとき左側に、高いとき右側に平行移動する。項目識別力 a_j が変化するとグラフの頂点が高くなり、値が大きいほど頂点は高くなる。項目情報関数は1つの項目に対して1つ算出され、頂点の高い複数の項目情報関数が θ 全域に対して配備されていると理想的な問題構成となる。

5. 評価実験

プログラミングの授業にコードレビューを導入し、チェックシートの項目が効果的なものであるかを調べるために評価実験を行う。実験は木更津工業高等専門学校情報工学科3年生38名、4年生34名、5年生39名の計111名に対して行った。評価実験の流れを図5に示す。

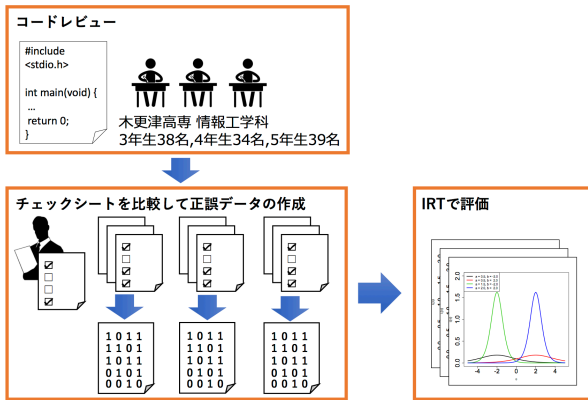


図 5: 評価実験の流れ

コードレビューの際に使用する問題およびソースコード、チェックシートは教師により作成されたものである。学生には0から9までの数値をランダムに並び替えるという問題に対して作成された2つのソースコードを読んで、チェックシートの項目を解答してもらった。実験に用いたチェックシートの項目を表1に示す。

表 1: 実験に用いたチェックシートの項目

1	No1 と No2 はどちらもランダムに並び終わる。
2	計算量の観点では No2 の方が良い。
3	No1 の方がプログラマーのスキルレベルが高い。
4	No2 に改良の余地はない。
5	No1 のコーディングは無駄がない。
6	自分なら No2 の考え方でコーディングする。
7	一緒に仕事するなら No2 のプログラマーだ。

6. 実験結果

評価実験の結果を以下に示す。3年生の結果の項目情報関数を図6、4年生の結果の項目情報関数を図7、5年生の結果の項目情報関数を図8に示す。

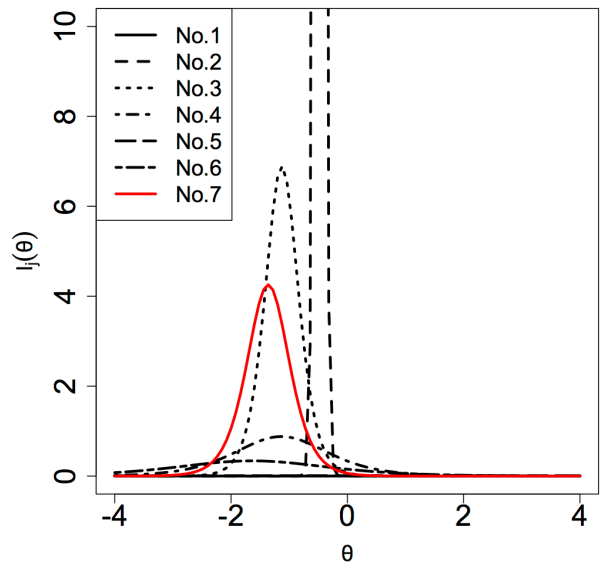


図 6: 項目情報関数 (3 年生)

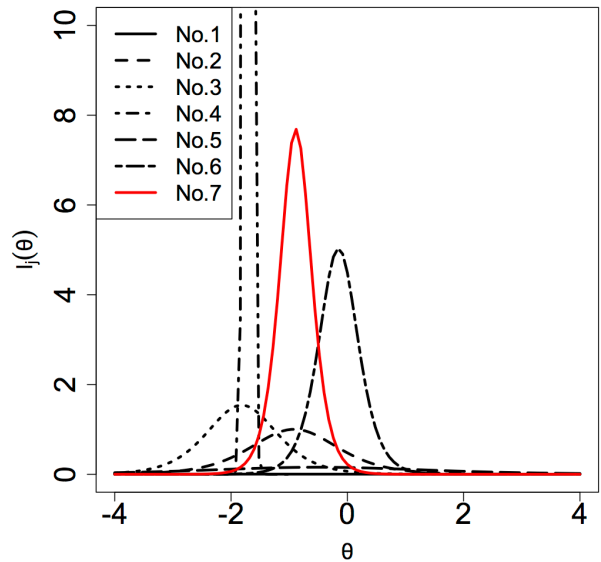


図 7: 項目情報関数 (4 年生)

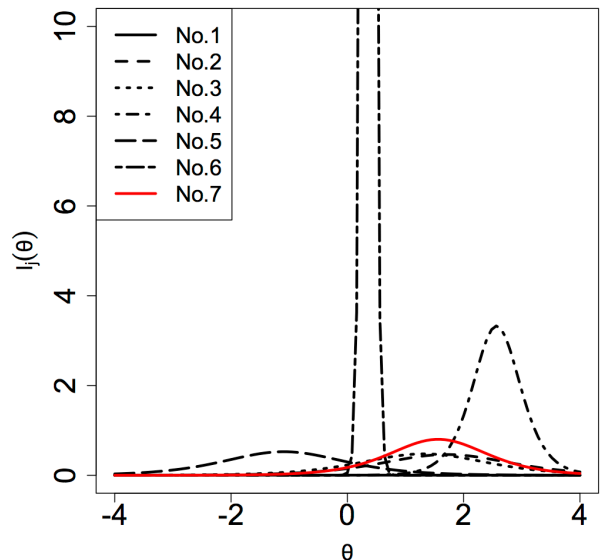


図 8: 項目情報関数 (5 年生)

7. 考察

今回の評価実験では、学年が上がるに連れてプログラミングスキルが高いと仮定して実験を行った。同じ項目に対しての項目情報関数が、学年が上がるに連れて左側に位置されることを予測していた。これは、左側に平行移動することによって被験者に対して易しい項目になっていることを意味する。つまり、プログラミングスキルが高いため同じ項目でも5年生にとっては易しい項目になることを表すためである。1つの項目に対しての項目情報関数の各学年の期待していた項目情報関数を図9に示す。

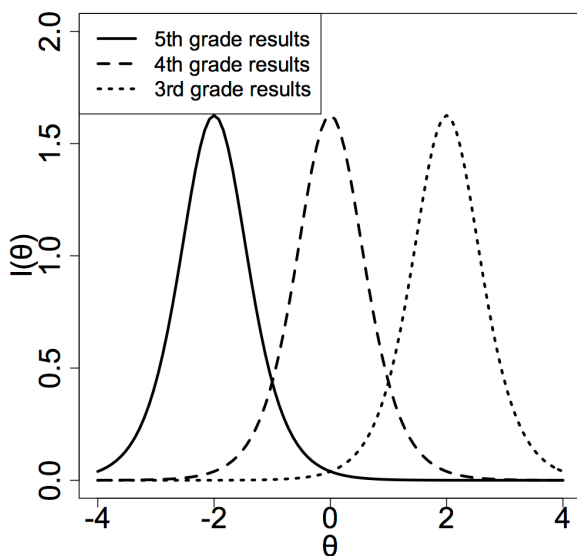


図9: 期待していた各学年の項目情報関数

前述した仮定のもと実験を行い期待通りの結果を得ることができれば、チェックシートの項目を定量的に評価できると考えていた。

しかし、5年生に対しての実験結果が全体的に右側に位置している。そのため、今回の実験に対しては5年生よりも3,4年生の方がプログラミングスキルが高いことを示している。今回の結果になった要因として、3,4年生の方が今回の問題に必要なスキルについての授業を行ってから日が浅いこと、高学年になるに連れてプログラミングの授業では基本的なことよりも応用的なものを作成する授業が多いことが要因としてあげられると考察する。他の要因として、チェックシートの項目が最適でなかったことや、評価実験として今回の実験方法が適していなかった可能性も考えられる。そのため、評価実験の方法を再考しつつ実験回数を重ねてチェックシートとして効果的な項目を探る必要がある。

また、3,4年生間で比較すると全体的な項目情報関数が同程度の θ の区間に位置している。このことから、今回の問題から得られるプログラミングスキルでは3年生と4年生ではあまり差がないことがわかった。

一方で、本実験により、項目情報関数を用いてチェックシートの項目難易度や項目識別力を可視化できた点については、新しい成果を得たと考えている。

8. まとめ

プログラミング教育の中にチェックシートを用いたコードレビューを導入し、チェックシート項目の評価実験を行った。期待通りの実験結果が得られなかったことから、実験方法やチェックシート項目について考え直す必要があると考える。しかし、これまで教員の感覚によって作成していた項目を定量的に評価し、修正を行うための指標を得ることができた。

今後は、プログラミングのソースコードに対してのチェックシートの有用性について示すことを課題とし、それが達成できれば、スポーツや絵画などの他分野にも応用したいと考える。

謝辞

本研究は JSPS 科研費 16K01095 の助成を受けたものである。

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] 澤井亨, “バレーボール授業におけるスキルチェックシートの作成”, 大阪産業大学人間環境論集 15, pp.167-177, 2016 - 03.
- [3] 荻原朋子, “仲間学習モデルの体育授業への適用過程とその成果”, 体育科教育学研究 21(2), pp.42-48, 2015.
- [4] 吉田和人, 山田耕司, 玉城将, 加賀勝, “卓球サービス動作に関する指導の観点: 日本ジュニア世代トップレベル選手を対象とした競技サポートから”, スポーツ教育学研究, Vo36, No.2, pp.49-59, 2016.
- [5] 長田則子, 斎藤秀平, 梅野圭史, 林修, 上原禎弘 “バスケットボールにおける Back Ring Target 投法の有効性に関する実験的研究-練習効果に着目して-”, 鳴門教育大学学校教育研究紀要, 第28号, pp.55-60, 2014.
- [6] 猫田和明, “教員養成における外国語活動指導者養成プログラム-チェックシートを用いた対話型の講義・演習を通して-”, 中国地区英語教育学会研究紀要, No.42, 2012.
- [7] 有賀妙子, 吉田智子, “ネットワークリテラシー教育の授業設計と教材開発”, 日本教育工学会論文誌 27(2), pp.181-190, 2003.
- [8] 小松真也, 二宮信治, “基本操作手技習熟度の定量評価とチェックシートを用いたシナリオ達成度評価の関係性の検討”, 体外循環技術 Vol.42 No.4, pp.393-399, 2015.
- [9] 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, 2002.