

学生が作成した問題の質向上のための相互評価手法の提案

Proposal on a Mutual-Evaluation Method for Improving Quality of Quizzes Created by Students

佐々木 匠^{*1}, 井上 裕之^{*2}, 高木 正則^{*1}

Takumi SASAKI^{*1}, Hiroyuki INOUE^{*2}, Masanori TAKAGI^{*1}

^{*1}岩手県立大学ソフトウェア情報学部

^{*1}Faculty of Software and Information Science, Iwate Prefectural University

^{*2}岩手県立大学大学院ソフトウェア情報学研究科

^{*2}Graduate school of Software and Information Science, Iwate Prefectural University

Email: g031k068@s.iwate-pu.ac.jp

あらまし：著者らはこれまで学習者が問題を作成し、その問題をグループ内の学習者と相互に評価する学習活動を実践してきた。これにより、学生間のコミュニケーション機会の増加や理解力の向上が期待できる。しかし、作成された問題には誤りや不備が残っているため、問題の再利用性が低かった。そこで、学生が作成した問題の質向上のための相互評価手法を提案する。本稿では、大学の授業で実践した作問と相互評価の結果から本手法の有効性を評価する。

キーワード：作問学習，グループ学習，相互評価

1. はじめに

著者らは、学習者自身が問題を作成し、その問題を共有することで学習を進める WBT システム「CollabTest」を開発し、教育現場での作問学習・相互評価を実践してきた⁽¹⁾。しかし、作成した問題を学生同士で相互評価した場合、誤りに気づくことができている学生もおり、問題に誤りや不備が残ったまま教員に提出されることが多々あった。そのため、問題の再利用性が低くなっているのが現状である。これは、評価する学習者の理解度の低さと、評価する際の手順の違いが要因として考えられる。そこで、本稿では学生が作成した問題の質向上を目的とした相互評価手法を提案する。また、評価者の理解度と評価手順などを把握することを目的として実施したアンケート結果について述べる。

2. システムの提案

2.1 システム概要

問題を評価するには、その内容に対して理解している必要がある。さらに、問題の質を向上させるには、誤りを発見するための観点や方法などの評価者としての能力が必要となる。そこで、問題の誤りを全て発見できるようになるまで繰り返し問題を評価し、評価者としての能力を高めたあと、グループメンバーが作成した問題を相互評価できるシステムを提案する。

2.2 システム利用の流れ

(1)問題の評価

教師があらかじめ講義で学んだ単元に関する問題に対して意図的に誤りを含めた問題を登録しておく。学習者は誤りを含んだ問題を評価する。1つの問題には種類の異なる誤り（問題文の誤字・脱字、答えの誤り、解説の計算ミスなど）を複数含める。

(2)評価不備の判定

評価能力向上システムが学習者からの評価結果を受け取り、誤りの発見数と正しい指摘数を自動判定する。全ての誤りを発見した場合は(4)へ、発見できなかった場合(3)を行なう。

(3)アドバイス生成機能

誤りを発見できなかった箇所と、その誤りを発見するためのアドバイスを（フィードバック機能）誤りの種類は分類しておき、分類グループに対応するアドバイスを提供し(1)を再度行なう。

(4)評価者認定

正しい指摘と全ての誤りを発見することができた場合、評価者として必要な考え方や見方、理解度があると判断し、評価者として認定する。

(5)学生が作成した問題の評価

評価者としての認定を得ることができた学習者は、学生が作成した問題を評価することができることにする。

以上に示したように、1つの問題に含まれている誤りをすべて指摘できるようになるまで繰り返し問題を評価させることによって、問題を評価する能力を高めるとともに理解度を高めてもらう。

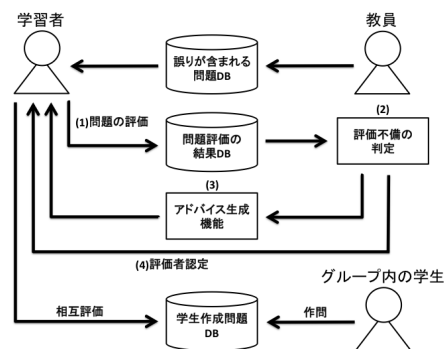


図1 システムの流れ

3. 問題評価の結果と分析

3.1 調査概要

評価者の理解力、評価手順と評価能力との関連を分析するため岩手県立大学ソフトウェア情報学部 1 年生の「情報基礎数学 C」(受講者 86 名)を対象として、調査を行った。この講義では数学の問題の作成と相互評価を「CollabTest」を利用して行っている。各回の授業は予習を必須としており、授業前に eラーニング教材を使って自学自習(予習)し、授業では事前・事後テストやグループ学習、作問演習を行っている。今回の調査では、ベクトルを対象とした授業を対象とした。

CollabTest を活用した 2 回目の作問演習の前に、意図的に誤りを含ませた問題の評価と、評価に関するアンケート調査を実施した。調査では、問題文と答え、選択肢、解説に誤りを合計 5 つ含ませたベクトルの問題を利用した。学生には、問題文と図、4 択の選択肢、答えを導く計算過程を示した解説を閲覧してもらい、問題を評価してもらった。誤りの箇所、誤りの修正案、解説の分かりやすさなどについて記述してもらった。

3.2 アンケート内容

評価項目については⁽²⁾を参考にして作成した。また、評価手順については①問題文を読む、②答えを見る、③選択肢を見る、④解説を読む、⑤答えを求める、⑥その他の 6 つの中からどの順序で評価していくのかを選択してもらい、さらに、特に着目して見る観点を「誤字・脱字」、「表現の適切さ」、「内容の正確さ」、「その他」の中から選んでもらった。

3.3 問題評価結果の分析

問題の評価結果のデータは 86 名から収集した。また、評価手順に関するアンケートは 85 名が回答した。評価結果のデータを分析した結果、38 種類の評価手順が得られた。誤り発見数と正しい指摘数ごとに分類したグループの人数と、ベクトルの授業で実施した事前・事後テストの平均点を表 1~3 に示す。事前・事後テストの平均点は 10 点満点のテスト 10 回分、合計 100 点満点の得点を表している。

表 1 誤りの発見数毎のテストの平均点と所属人数

誤り発見数	0 個	1 個	2 個	3 個	4 個	5 個
平均点	34.7	50.9	46.0	54.3	53.8	51.3
人数	3 人	17 人	33 人	18 人	12 人	3 人

表 2 正しい指摘数毎のテストの平均点と所属人数

正しい指摘数	0 個	1 個	2 個	3 個	4 個	5 個
平均点	57.3	42.6	47.8	54.9	52.2	66.0
人数	8 人	21 人	27 人	16 人	13 人	1 人

表 3 評価時に注目して見ている観点の回答人数

注目している観点	誤字・脱字	表現の適切さ	内容の正確さ	その他	なし

人数	18 人	5 人	20 人	2 人	40 人

3.4 考察

表 1~3 の結果からテスト得点と誤り発見数・正しい指摘数の相関係数を算出した。表 4 に求めた相関係数を示す。表 4 よりテスト得点と誤り発見数・正しい指摘数の間には相関がないことが分かった。つまり、テストの得点が高くなって誤りの発見や正しい指摘ができていたことが示唆される。今回調査は、ベクトルの授業を 6 回分受講したあとに実施したため、問題を評価した学生はベクトルに関する知識をある程度持っていたため、このような結果になったと考えられる。

また、解説の誤りを発見している人のうち問題文の誤りを発見できなかった人の割合は 22%、問題文の誤りを発見している人のうち解説の誤りを発見できていなかった人の割合は 24%であった。このことから、全体の約半分の人が誤りを発見している箇所に偏りがあることが分かった。これらの偏りを減らすことで、より多くの誤りを発見できると考える。

評価手順では表 5 より、問題文を見たあとに解説部分を見る、最後に選択肢を見るということが、誤りを多く発見できる順序だと考えられる。表 5 は 3.2 章で示した評価手順①~⑥に対応している。

表 4 テスト得点と誤り発見数・正しい指摘数との相関係数

テスト合計点と誤り発見数	0.168
テスト合計点と正しい指摘数	0.132

表 5 誤り発見数 0 個と 5 個の人の評価手順

0 個	①→⑤→③→②→④, ①→③→②→④→⑤, ①→②→③→④→⑤
5 個	①→④→⑤→②→③(2 名), ①→⑤→②→④→③

4. まとめ

問題を評価する際、誤りや不備を発見できていないことで、問題の質が低くなり、再利用性が低くなっている。そこで、アンケートを実施し、現状としてどのようにして問題の評価をしているのか調査し分析した結果、理解度が高くなって誤りの発見や正しい指摘ができることと分かった。これらの結果を考慮し、システムを利用してもらうことで問題の質が高くなり、再利用性が高くなるようにしたいと考えている。

参考文献

- (1) 高木正則, 田中充, 勅使河原可海: “学生による問題作成およびその相互評価を可能とする協調学習型 WBT システム”, 情報処理学会論文誌 48(3), pp.1532-1545 (2007)
- (2) 高木正則, 星野大輔, 望月雅光, 勅使河原可海: “学生が作成した問題の改善を促すピアレビューナビゲーション機能の開発と評価”, 教育システム情報学会誌 vol.27, No.1 pp.87-99 (2010)