

授業支援システム Moodle における小テスト結果からの 理解度不足の検出機能の提案

加藤利康

日本工業大学

A Proposal of Detection Function of Degree of Understanding from Quiz Responses in Course Management Systems

Toshiyasu KATO

Nippon Institute of Technology

katoto@nit.ac.jp

概要: 本論文は、授業支援システムを用いた小テストの結果から理解度不足の学習内容を検出し教員へ提示する理解度不足の検出機能を提案する。理解度不足の検出機能は、小テストの設問を学習項目で分類し、その正答率の差を分割表によって検定して有意に正答率が低い学習内容を検出する。評価実験の結果、実際の授業で小テストの実施後に教員が理解度不足と判断し説明を行った学習内容のうち、提案機能が理解度不足と判定した割合が86%あった。この結果から提案機能は、理解度不足の学習内容を教員が的確に把握することに有用である。

Abstract: This paper proposes a function that analyses the quiz results and determines the content that a student has not sufficiently understood. This information is subsequently presented to the teacher. The proposed function statistically tests for differences in systematically classified correct answer rates and detects which of the studied content has significantly low correct answer rates. The results of an evaluation experiment showed that 86% of the content that was determined to have been understood insufficiently had been explained by the teacher. This result demonstrates that the proposed function is useful in enabling teachers to accurately ascertain which parts of the studied content have not been sufficiently understood.

キーワード: 授業支援システム、小テスト、対面授業、学習履歴、学習分析

Keywords: Course Management System, Quiz, Face-to-Face lesson, Learning History, Learning Analytics

1. はじめに

Moodle (Moodle) などの授業支援システムの小テスト機能が対面授業中における学習内容の理解度の把握に利用されている(江本, 2010)(清水, 2007)。小テスト機能は、教員が問題と模範解答を用意するだけでオンラインでの小テストを実現し、自動で採点を行う。そのため、小テスト機能には、小テストの実施後すぐに正答率を学生の理解状況の代理指標として確認できるメリットがある(植野, 2007)(William, 2008)。

小テスト機能が提供する小テストの結果は、設問ごとの正答率を提示するだけで、クラスで相対的に理解が不足している学習内容を指摘していない。教員は小テストを理解度確認に用いる場合、小テストの結果をその場で分析し、

理解度不足の学習内容を再説明している(駒林, 1986)。小テストの結果には、クラスに再説明が必要な理解度不足の学習内容を判別しなければならない問題がある。そのため、理解度不足の検出を目的とした研究が行われている(Romero, 2013)(Romero, 2008)(生田目, 2006)(和田, 2012)。しかしながら、第2節で述べるように、対面授業中に理解度不足の検出が行われていないため、教員は理解度不足の学習内容を随時に的確に把握できていない。また、しきい値による検出では、小テストに合わせてしきい値を設定する必要があるため、とくに学習指導が必要な学習内容を教員が精査する必要がある。

そこで本研究は、小テスト機能の不足を補うことを目的として、理解度不足の検出機能を提案する。理解度不足の検出機能は、授業支援システム Moodle における小テストの結果からクラスで相対的に理解度不足の学習内容を検

出し教員へ提示する。本研究における理解度不足の定義は、学習内容として単元と学習項目と設問がそれぞれ 1 対 n で系統的に分類されている分類項目間の正答数の差が統計的に有意に低い場合とする。教員が対面授業という限られた時間の中でクラスの理解度不足の学習内容を的確に把握できることは、対面授業における授業支援システムの小テスト機能の有用性をさらに高めることが期待される。本研究が想定している小テストを用いた授業は、小テストを学習内容の復習として授業の始めか途中に実施するものである。また、本論文における学習内容は、単元や学習項目で系統化されているものとする。

2. 関連研究

小テストの結果から理解度不足の学習内容を検出することを目的とした関連研究の提案手法には、データマイニングの応用、複数回の誤答分析、テスト理論に基づく分析がある。

データマイニングの応用として、相関ルール分析の適用 (Romero, 2013) は、小テストの得点と解答時間を教員が用意し、データマイニング手法 (福田, 2001) の相関ルール分析 (Agrawal, 1993) によって難しい問題と易しい問題、および解答時間の長い問題を検出する。その提案手法は、問題 A と問題 B が誤答のとき問題 C も誤答するという、2 つの誤答から誤る問題のルールにより難しい問題を検出することである。問題 A, B, C が正答の場合には、易しい問題を検出する。問題が誤答で小テストの得点が低いとき解答時間が長いというルールからは解答時間の長い問題を検出する。この有効性は、問題に対する学生の傾向を提示することで、問題の関連性を教員が把握できることである。また、決定木の適用 (Romero, 2008) は、小テストの得点と解答時間、および解答数を教員が用意し、データマイニング手法の決定木 (Quinlan, 1993) によって学生を FAIL, PASS, GOOD, EXCELLENT の 4 つに分類する。その提案手法は、FAIL や PASS に分類された学生が誤答した共通の問題を検出することである。この有効性は、レベル分けされた理解度不足の問題を提示することで、学習指導の検討材料を教員が得られることである。

複数回の誤答分析 (生田目, 2006) は、小テストの得点とアクセス履歴を教員が用意し、複数回の誤答内容によって

当て推量で解答した問題を検出する。その提案手法は、複数回の小テストの解答において類似問題に解答が一貫していない、同一問題を 2 回以上誤答する、以前正解した問題を誤答する、といった規準が全て当てはまる場合に当て推量と判断し理解度不足の問題を検出する。この有効性は、理解度不足の学習内容を学習項目の単位で提示することで、学習指導の検討材料を教員が得られることである。

テスト理論に基づく分析 (和田, 2012) は、小テストの問題と得点を教員が用意し、項目反応理論 (竹内, 2006) と S-P 表理論 (佐藤, 1975) によって問題の妥当性と理解度不足の学習内容を検出する。その提案手法は、学生の解答パターンから項目反応理論による項目難易度と S-P 表理論による受験者注意係数を示すことによって、理解度不足の問題と学習上に問題のある学生を検出する。この有効性は、個別学生の学習上の問題点を提示することで、個別指導の検討材料を教員が得られることである。

これらの関連研究に共通する問題点は、対面授業におけるクラスに対して学習指導に必要な分析が行われていないことである。教員は、小テストの結果から学習定着度を確認し、再度説明すべき学習内容を把握していると言われている (生田目, 2006) ため、クラスの理解度不足の学習内容を的確に把握する必要がある。この問題点に対して本研究は、クラスに対する学習指導に必要な学習内容を指摘する機能を提案している。

相関ルール分析の適用 (Romero, 2013) は、各分類を比較する点で本研究と共通している。Romero らの提案手法が解答時間と得点を分析しているのに対して、本研究では系統的に分類した得点を分析している点が異なる。本研究の手法は、単元や学習項目ごとに理解度不足を把握できる。また、決定木の適用 (Romero, 2008) は、理解度不足の学習内容をレベル分けする点で本研究と共通している。Romero らの提案手法が必ず理解度不足を検出するのに対して、本研究では統計的に分析しているため、有意差のある理解度不足を検出する点が異なる。本研究の手法は、対面授業の限られた時間の中でクラスの理解度不足の有無と、その学習内容を把握できる。

複数回の誤答分析 (生田目, 2006) は、理解度不足の学習内容を学習項目などの単位で検出している点で本研究と共通している。生田目の提案手法が複数回分の小テスト

の結果を用意する必要があることに対して、本研究では 1 回分の小テストの結果から分析できる点が異なる。本研究の手法は、1 回目の授業から理解度不足のある学習内容を検出できる。

テスト理論に基づく分析(和田, 2012)は、理解度不足の問題を検出している点で本研究と共通している。和田の提案手法が小テストの問題ごとに妥当性を分析するのに対して、本研究では小テストの問題を単元や学習項目で系統的に分類している点が異なる。本研究の手法は、単元や学習項目の単位で理解度不足のある学習内容を検出できる。

本研究の手法は、分析にカイ二乗独立性検定を応用しているため、クラスでとくに学習指導が必要な学習内容を教員が把握できるようになる。また、分析対象を単元、学習項目、設問に分けて分析するため、1 回の授業だけでなく過去の授業で実施した小テストを含めた理解度不足の学習内容を検出できる。教員は、正答率が低い問題を全て確認しなくても単元別、学習項目別、設問別に理解度不足の学習内容を把握できる。

3. 対面授業中の小テストによる理解度把握

教員が授業支援システムの小テスト機能を利用し、対面授業中に学生の理解度を把握している。しかしながら、国内で最も利用されている(放送大学学園, 2011)Moodle の小テスト機能が提供する小テストの結果からでは、クラスでとくに学習指導が必要な学習内容を把握できない。そのため、教員は、その場でクラスの理解度不足の学習内容を判別し、再説明を行っている。

3.1 授業支援システム Moodle の小テスト機能

小テスト機能による設問は、一般的につぎのような形式の問題が作成可能である。

- 作文:解答を文章で入力する。
- 記述:解答を語句で入力する。
- ○×:問題に対する正誤を選択する。
- 多肢選択:選択肢の中から 1 つ以上の解答を選択する。
- 組み合わせ:問題群と解答群から、各問題に対応す

る解答を選択する。

Moodle の小テスト機能による小テストの結果の例としては、学生ごとの正答数とヒストグラム、およびアイテム分析などがある。アイテム分析は、設問に対するクラスの正答と誤答の解答率や標準偏差などを提示する。これらの小テストの結果は、学習内容に対する分類が行われていないことや理解度不足の統計的な検出には活用されていないため、理解度不足の学習内容を的確に把握することが困難である。

3.2 対面授業における小テスト機能の利用シナリオ

教員は、小テスト実施後に小テストの結果からクラスで相対的に理解が不足している学習内容を判別し、学習指導を行っている。学習内容が複数の設問で構成されていることがあるため、正答率の低い設問から、その学習項目を把握する必要がある。対面授業における現状の小テスト機能の利用シナリオは下記のとおりである。

1. 小テストを実施する。
2. 設問ごとに正答率を確認する。
3. 正答率の低い学習項目を確認する。
4. とくに学習指導が必要な学習項目を判別する。
5. 理解度不足の学習内容の再説明を行う。

4. 理解度不足の検出機能

現状の小テスト機能の不足を補うため、本論文は理解度不足の検出機能を提案する。従来の小テスト機能に理解度不足の検出機能を加えた利用シナリオと利用図をそれぞれ図 1 と図 2 に示す。

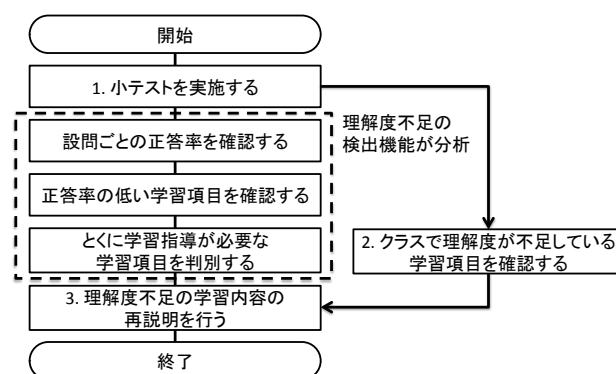


図 1 理解度不足の検出機能の利用シナリオ

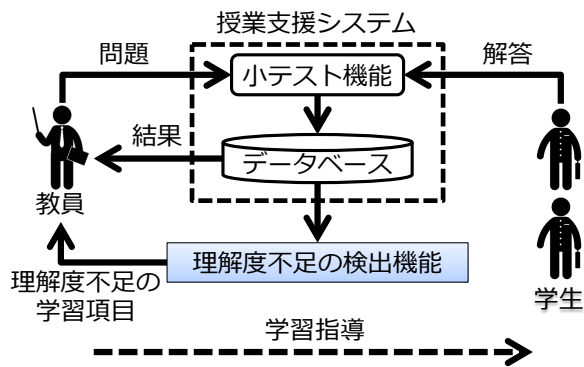


図2 理解度不足の検出機能を加えた小テスト機能の利用

理解度不足の検出機能の目的は、小テストの結果からクラスの理解度不足の学習内容を的確に検出することである。理解度不足の検出機能は、小テストの設問を学習内容に対応付けるため、作問時に設問ごとに授業の単元と学習項目を対応させる。

4.1 小テストで理解度を把握するための作問

理解度不足の提案機能が有効に機能するためには、学習項目に対応した設問を作成する必要がある。本研究は、特定の学習項目の理解を設問に反映させるために、授業で使用しているテキストの章毎にある「まとめ」を利用する。テキストを用いた理由は、評価実験対象の授業で利用されているためである。テキストのまとめの例は下記のとおりである。

- GUI プログラムを作成するために、Swing を利用することができます。
- Swing のウィンドウ部品を、コンポーネントと呼びます。
- コンポーネントを含むことができるコンポーネントを、コンテナと呼びます。
- ユーザーが行う操作などは、イベントとして処理されます。
- イベント処理はソース・イベント・リスナによって行われます。
- リスナをあらかじめソースに登録しておく、イベントが発生したときに、ソースからリスナにイベントが渡されて処理されます。

作問は設問集合間で難易度や識別力を揃えるために、作問の原則(橋本, 1972)に基づいた。作問の原則は、理解や思考などをテストの目標としたときに教員が自作でテ

ストを作るための規則である。文献(橋本, 1972)によると、テストの目標が理解の場合には、多肢選択による選択完成法が適当である。選択完成法による作問の原則は、つぎのとおりである。

- 空所の数を増やしすぎて、全体の意味や関連の把握が不可能にならないようにする。
- 各選択肢間の差異を縮めて、どれもある程度正答であるようにする。
- 選択肢の数を5, 6個程度にする。

本研究は、テキストの内容の理解を偏りなく妥当な問題とするために、上記の原則に基づいて作問手順を下記のように定義する。

- (1) 小テストの様式は選択完成法とし、問題形式は多肢選択とする。
- (2) 問題はテキストの章単位とする。章単位としたのは、評価実験対象の授業がテキストの章ごとに行われているためである。
- (3) 設問は章ごとに5個から9個ある「まとめ」とする。
- (4) 設問で空欄にする内容は章ごとに4個から8個ある「Check Point」の項目とする。空欄にする内容をCheck Point としたのは、新しい学習内容が他の学習内容と区別できているかの理解(佐伯, 2007)を確認するためである。Check Point の例は、「GUI、コンテナ、イベント、リスナ、ソース、コンポーネント」である。
- (5) 選択肢はCheck Point の項目をランダムにすべて列挙する。

4.2 理解度不足の検出方法

本研究における理解度不足の定義は、単元と学習項目、および解答について、対象の正答率と対象の各分類の平均正答率を相対的に比較して統計的に有意に低い場合とする。本研究では、個々の測定(正答あるいは誤答)は独立と見なして検定を行うものとする。統計的に有意に低い場合としたのは、理解度不足が予想の範囲外であることを把握する必要があり(駒林, 1986)、判別するためである。また、正答率を理解度の代理指標と位置付けるために、小テストごとに理解度不足を判別できるようにする。

ある学習内容が理解度不足であることの判定は、学習

内容間の相違点を発見するためにカイ二乗独立性検定によって決定する。カイ二乗独立性検定は、2 つの変数が独立しているかどうかを検定する統計的手法である(倉田, 2009)。検定式における観測度数は、対象とする単元、学習項目、設問のいずれか(以下、項目)の正答数および誤答数である。

理解度不足を判定するアルゴリズムの手順は、あらかじめ作問時に対応させた単元と学習項目に基づいて分析を行うため、下記とする。

- (1) 単元、学習項目、設問のいずれか(以下、項目)と有意水準値を入力する。有意水準値は、統計的仮説検定を行う場合に、仮説を棄却するかどうかを判定する基準である。
- (2) 項目に対応した受験者全員の解答を取得する。
- (3) 項目ごとの正答率と各分類(単元、学習項目、設問)の平均正答率を算出する。
- (4) 項目の正答率が各分類の平均正答率より下回っているならば、カイ二乗独立性検定を用いて項目の正答数・誤答数と分類におけるその他の正答数・誤答数からカイ二乗値を算出する。下回っていない場合は、この手順を終了する。
- (5) カイ二乗値から、カイ二乗分布の片側確率の関数値を算出する。
- (6) 関数値が有意水準値を下回っているならば、項目を理解度不足として出力する。

例えば、特定の設問についての判定を行う場合、その設問に対する正答数と誤答数、それ以外の設問に対する正答数と誤答数を用意する。これらを用いて、2*2 の分割表を構成し、独立性検定によって理解度不足を判定する。

この理解度不足の検出方法は、アルゴリズムの対象をクラスから個別学生にし、検定方法に超幾何分布(宮川, 2002)を用いることで、個別学生の理解度不足の学習内容の検出に応用できる(Kato, 2013)。超幾何分布は、母集団 N 個の要素からなり、そのうち属性 A をもつものが M 個含まれている離散型の確率分布である。教員は個別指導の際の検討材料として、学生ごとの理解度不足の学習内容を把握できるようになる。また、上記のアルゴリズムにおいて、項目を設問の解答に変え、正答数を解答数に変えたと誤答分析が可能である。この誤答分析により教員は、統計的に学生がどの解答と誤っているのかを把握できるように

なる。

4.3 検出方法の有効性の評価

理解度不足の判定の有効性は、判定が妥当であるかを確認するため理解度不足の判定数、および有意水準の変化による理解度不足の検出数を検証する。理解度不足の判定割合は、単元や学習項目の対象間の正答数にどの程度の差があれば検出できるかを検証する。検証の方法は、基準と対象の 2 つの正答数において、基準の正答数を 100 から 20 刻みで 20 までとし、有意水準の変化に応じて理解度不足と判定される対象の正答数を算出する。ここの正答数と誤答数の合計は 100 とする。理解度不足の判定割合の結果を図 3 に示す。図 3 の一番左側の点が有意水準 0.1%であり、一番右側が有意水準 50%である。基準と対象の正答数の差が有意水準 50%の場合は、100 に対して 97 から 20 に対して 13 の差で理解度不足が検出された。また、有意水準 0.1%の場合は、100 に対して 56 から 20 に対して 0 の差で理解度不足が検出された。これらの結果は、有意水準に応じて、正答数の差が 3 から理解度不足を検出し、基準の正答数が 20 でも理解度不足の判定が可能であることを示している。

有意水準による理解度不足の検出数は、単元と学習項目、および設問において、理解度不足の検出された項目数が有意水準に応じて変化することを検証する。実際の小テストにおいては、単元数と学習項目数、および設問数が比例していない。そこで、単元、学習項目、設問のどの水準が対象でも有意水準を調整することで検出率を段階的に変化させることができることを確かめる必要がある。評価に用いる小テストの結果は、本学の 2 年秋学期「ソフトウェア技法・演習」における小テスト 11 回分である。小テストにおける単元数は、11、学習項目数は 54、設問数は 68 である。なお、受験者数は 49 名である。有意水準の変化によるクラスの理解度不足の検出数を図 4 に示す。

理解度不足の検出数は、有意水準に応じて段階的に変化した。この結果から、理解度不足の判定は、有意水準の変更により、教員の対応能力に応じて検出率を制御することが可能である。

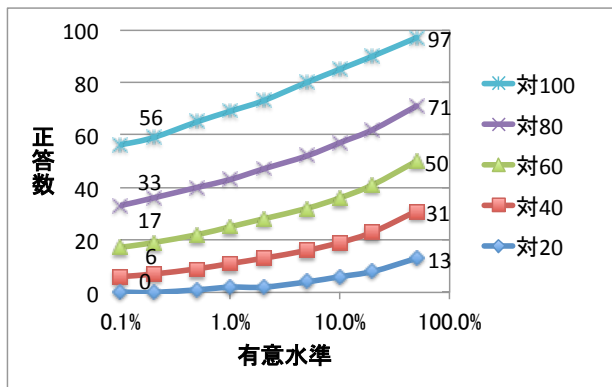


図3 理解度不足の判定数

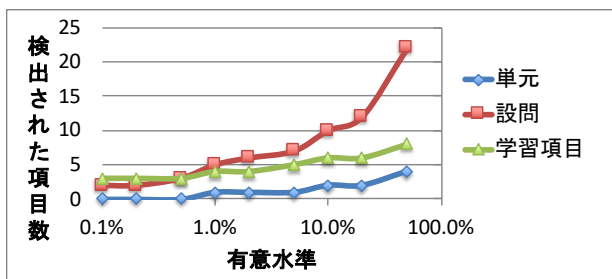


図4 クラスの理解度不足の検出数

5. 理解度不足の検出機能における有用性の評価実験

5.1 実験の目的

実験の目的は、提案する理解度不足の検出機能(以下、提案機能)が実際の授業で有用となるのかを評価することである。ここでの有用とは、教員が実際の授業における学習指導の際に検討材料となるかどうかである。

5.2 実験の方法

本研究は、提案機能が学習指導に有用であるかの評価を行った。評価の前に提案機能の有意水準を決めるため、教員の認識アンケート調査を行った。調査結果に基づき、提案機能により理解度不足と判定された学習項目と小テスト実施後に教員による再説明のあった学習項目を比較する。

認識アンケート調査の方法は、まず提案機能による理解度不足の判定が、教員の判断と合っているかを確認するため、小テストの実施ごとに提案機能による小テストの結果を確認してもらう。その後、全授業終了後にアンケートを回

答してもらう。この提案機能による小テストの結果は、学習内容で分類された単元や学習項目ごとの正答率に加えて、可変な有意水準に基づいた理解度不足が示される(図5参照)。図5における下線は、理解度不足を示す。教員は有意水準を自由に設定できる。対象とした授業は、小テストを実施している授業担当教員2名2クラスの本年2年春学期「プログラミング技術・演習」である。実験条件として、分析対象の小テストの結果は授業で実施されたすべての9回分とし、提案機能には、対象の授業で Moodle が使われていたため、その小テストの結果を分析して Web ブラウザで表示できるように PHP 言語と R 言語(The R Project)を用いた。また、アンケートの内容はつぎのとおりである。

問 1. クラスの理解度不足の判定結果は、教員の判断と合っていますか？

(1. ほぼ合っている 2. まあまあ合っている 3. やや合っていない 4. あまり合っていない)

問 2. 教員の判断と最も合っている検出判定の有意水準は何%ですか？

(1. 5% 2. 10% 3. 15% 4. 20% 5. 25%)

問 3. 最も合っている検出判定の有意水準を選んだ根拠は何ですか？(複数回答可)

(1. 理解度不足のあった学習項目が検出されている 2. 理解度不足のない学習項目が検出されていない 3. 検出される理解度不足の数が指導するのに適当だから 4. その他:具体的に書いてください)

クラス全体の小テスト結果						
* 赤字の回答率は、その項目が他の項目と比較して統計的に有意に低いことを表します。 有意水準: 20 %						
実施日	単元	設問番号	問題文	学習項目	選択肢	回答率
2014-10-03	はじめての	1-1	【 】は、単体で動作するプログラムです。	アプリケーション	アプリケーション アプレット Java ソースコード コンパイル クラスファイル インタプリタ	48.9 % 13.3 % 17.8 % 2.2 % 2.2 % 6.7 % 8.9 %
		1-2	【 】は、Webブラウザなどの上で動作するプログラムです。	アプレット	アプレット アプリケーション Java ソースコード コンパイル クラスファイル インタプリタ	46.7 % 26.7 % 20.0 % 2.2 % 0.0 % 0.0 % 4.4 %
		1-3	【 A 】を使って作成したテキスト形式のプログラムを、【 B 】と呼びます。空欄Aの解答	Java	Java アプリケーション アプレット ソースコード コンパイル クラスファイル インタプリタ	40.0 % 2.2 % 2.2 % 22.2 % 0.0 % 13.3 % 20.0 %
	二歩	2-1	【 】は、単体で動作するプログラムです。	アプリケーション	アプリケーション アプレット Java ソースコード コンパイル クラスファイル インタプリタ	48.9 % 13.3 % 17.8 % 2.2 % 2.2 % 6.7 % 8.9 %
		2-2	【 】は、Webブラウザなどの上で動作するプログラムです。	アプレット	アプレット アプリケーション Java ソースコード コンパイル クラスファイル インタプリタ	46.7 % 26.7 % 20.0 % 2.2 % 0.0 % 0.0 % 4.4 %
		2-3	【 A 】を使って作成したテキスト形式のプログラムを、【 B 】と呼びます。空欄Aの解答	Java	Java アプリケーション アプレット ソースコード コンパイル クラスファイル インタプリタ	40.0 % 2.2 % 2.2 % 22.2 % 0.0 % 13.3 % 20.0 %
		2-4	【 】は、単体で動作するプログラムです。	アプリケーション	アプリケーション アプレット Java ソースコード コンパイル クラスファイル インタプリタ	48.9 % 13.3 % 17.8 % 2.2 % 2.2 % 6.7 % 8.9 %
		2-5	【 】は、Webブラウザなどの上で動作するプログラムです。	アプレット	アプレット アプリケーション Java ソースコード コンパイル クラスファイル インタプリタ	46.7 % 26.7 % 20.0 % 2.2 % 0.0 % 0.0 % 4.4 %
		2-6	【 A 】を使って作成したテキスト形式のプログラムを、【 B 】と呼びます。空欄Aの解答	Java	Java アプリケーション アプレット ソースコード コンパイル クラスファイル インタプリタ	40.0 % 2.2 % 2.2 % 22.2 % 0.0 % 13.3 % 20.0 %
		2-7	【 】は、単体で動作するプログラムです。	アプリケーション	アプリケーション アプレット Java ソースコード コンパイル クラスファイル インタプリタ	48.9 % 13.3 % 17.8 % 2.2 % 2.2 % 6.7 % 8.9 %

図5 提案機能による小テストの結果

評価方法は、提案機能による理解度不足の学習内容が学習指導に反映されるのかを確認するために、理解度不足と判定された学習項目と小テスト実施後に教員による再

説明のあった学習項目の割合を示す。対象とした授業は、小テスト実施後に教員が Moodle のアイテム分析を見ながら再説明をしている本学 2 年秋学期「ソフトウェア技法・演習」である。評価の教員は 1 名 1 クラスであり、アンケート調査の教員とは異なる。実験条件として、分析対象の小テストの結果は授業で実施されたすべての 11 回分とし、提案機能の理解度不足の検出における有意水準は、アンケート結果で妥当とされた 20%とする(後述 5.3 節参照)。再説明の基準は、著者の判断で口頭の説明だけでなく、関連する資料の提示やホワイトボードへの書き込みによる説明が行われた場合とする。

5.3 実験の結果

教員による理解度不足の学習内容に対する認識アンケート結果を表 1 に示す。教員 2 名は、有意水準 20%のときに、理解度不足のあった学習項目が検出されていると回答した。

理解度不足と判定された学習項目と小テスト実施後に教員による再説明のあった学習項目を表 2 に示す。理解度不足と判定された 14 項目 A に対して再説明があった学習項目 B (表 2 の○で囲んだ 12 項目)の割合を示す適合率 $P(A|B)$ は、0.86 であった。また、再説明のあった 28 項目 C に対して理解度不足と判定された学習項目 B の割合を示す再現率 $P(B|C)$ は、0.43 であった。

表 1 理解度不足の学習内容に対する認識アンケート結果

	問1 理解度不足の判定 と教員の判断	問2 妥当な 有意水準	問3 問2の根拠
教員 A	やや合っていない	20%	理解度不足の あった学習項目 が検出されている
教員 B	まあまあ 合っている	20%	理解度不足の あった学習項目 が検出されている

表 2 理解度不足と判定された学習項目と再説明のあった

学習項目

単元	学習項目	理解度不足と判定された学習項目 (14 項目)	再説明のあった学習項目 (28 項目)
はじめの 一歩	a.アプリケーション, b.アプレット, c.java, d.ソースコード, e.コンパイル, f.クラスファイル, g.インタプリタ	c, g	a, ㉔, ㉕
クラスライブラリ	a.文法, b.クラスライブラリ, c.実行, d.クラス, e.パッケージ	なし	a, b
Swing の 基本	a.GUI, b.コンテナ, c.イベント, d.リスナ, e.ソース, f.コンポーネント	e	b, ㉔, f
Swing の 応用	a.レイアウトマネージャ, b.パネル, c.ラベル, d.アイコン, e.ボーダー, f.ボタン, g.テキストフィールド	a	㉕, b, g
Swing の 活用	a.リスト, b.コンボボックス, c.テーブル, d.モデルとビュー, e.メニューバー, f.ツールバー, g.ダイアログ, h.フレーム	c, d	㉔, ㉕, g, h
サーブレット	a.Web, b.HTTP, c.サーブレット, d.フォーム, e.セッション管理, f. リクエストの転送	c	b, ㉔, e
JSP	a.JSP, b.JSP の書式, c.暗黙の オブジェクト, d.コンポーネント, e.javaBeans	なし	なし
JDBC	a.JDBC, b.データベース, c.SQL, d.Web	b, d	a, ㉕, ㉖
ファイル 操作	a.ファイルチューザ, b.ファイル 情報, c.テキストファイル, d.バイナリ ファイル, e.シーケンシャル アクセス, f.ランダムアクセス, g. 正規表現	g	なし
ネット ワーク	a.DOM, b.XSL, c.スタイル シート, d.インターネットアドレ ス, e.ネットワーク, f.スレッド, g.ソケット	c	a, ㉔, g
大きなプ ログラム の作成	a.仕様, b.外観, c.データ, d.機能, e.クラス, f.処理, g.変更, h.拡張	b, d, e	c, ㉕, ㉖, f

○印は提案機能が理解度不足と判定した学習項目を示す(12 項目)

5.4 結果の考察

適合率が 0.86 あったことから、提案機能による理解度不足の判定はクラスに対する学習指導の際に検討材料となる。教員は、小テスト実施後すぐに理解度不足の学習内容を再説明できるようになる。

再現率が 0.43 と低い原因は、理解度不足の学習内容が他の学習内容と関連していることが挙げられる。理解度不足と判定された学習項目よりも再説明のあった学習項目

のほうが多いのは、学習内容の理解が既存の知識構造へと新しい学習内容をマッピングするプロセスであると言われていたためである(佐伯, 2007)。再現率が低くなるのは、新しい学習内容は既存の知識構造に関連づけられる必要があり、再説明のときに既存の知識にも言及する必要があるためと考えられる。学習項目に関連があることを示した結果として、小テスト 11 回分のうち再説明のあった学習項目が 1 つだけの回はなかったことが挙げられる。有意水準を上げることにより他の理解度不足の学習内容を検出できるが、その学習内容が関連しているかどうかは分析が必要である。

教員の妥当有意水準が 20%のときに再現率が 0.43 であったことについて、理解度不足の学習項目と実際に学習指導する学習項目は異なることが考えられる。有意水準を変化させたときの理解度不足の判定を図 6 に示す。有意水準を上げることによって、再現率が上がるため、教員は学習指導が必要な学習項目を把握できるようになる。しかしながら、適合率が下がるため、学習項目間に関連しているかどうかは分析が必要である。

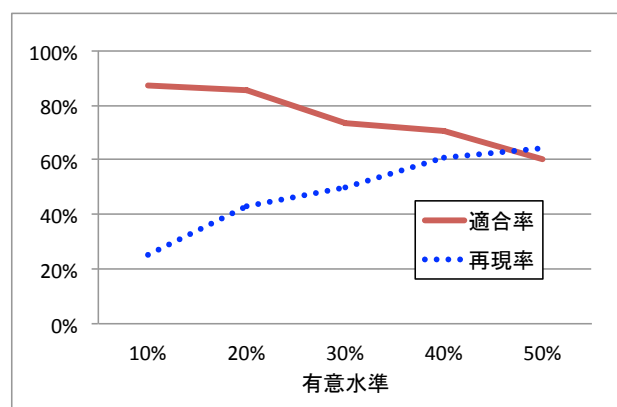


図6 理解度不足の判定評価

5.5 小テストの基準関連妥当性の考察

評価実験における小テストの作問は、設問集合間で難易度や識別力を揃えるために、作問の原則(4.1 節に詳述)に基づいた。この原則による作問の妥当性について本研究は、本学 2 年春学期「プログラミング技術・演習」の期末試験の得点と 2 年秋学期「ソフトウェア技法・演習」の小テスト 11 回分において、両方の科目を受講した学生の結果を比較することで評価する。これらの科目を対象としたのは、

毎回の授業開始時に前回の授業における学習内容の理解度確認小テストを行っていること、および春学期から秋学期へ続いている授業内容のためである。評価方法は、授業担当教員が作成した春学期の期末試験の得点と本研究の作問による秋学期の小テストの総合点を正答率にして比べ、成績の差異と相関係数を確認する。成績の差異は、春学期と秋学期の科目を受講した学生 46 名の平均値の比較である。相関係数は、学生 46 名の期末試験の正答率と小テストの正答率の関係指標である。

評価の結果を図 7 に示す。図 7 は、46 名の学生を横軸にならべ、それぞれの学生について、期末試験の正答率と小テストの正答率を縦に並べたプロットである。46 名の平均正答率の差異は、春学期から秋学期を引いた差で 7.2%であった。対応のある平均の差の検定の結果は、有意差がなかった($p>.5$)。その標準偏差は、期末試験が 18.2%、小テストが 16.5%であった。また、相関係数は 0.5 ($p<.001$)であった。この結果から、期末試験と小テストには相関があり、成績の差異には有意差がないことから、本研究における小テストの作問は妥当である。

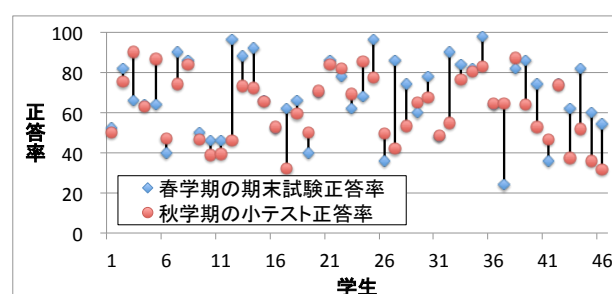


図7 春学期と秋学期の正答率の差異

6. おわりに

本論文は、授業支援システム Moodle における小テストの結果から理解度不足の学習内容を検出して提示する理解度不足の検出機能を提案した。理解度不足の検出機能は、小テストの設問を単元と学習項目で系統的に分類し、その正答率の差を分割表によって検定して有意に正答率が低い学習内容を検出する。この提案機能により教員は、従来の小テストの結果に加えて、統計的に検出される理解度不足の学習内容を随時に把握できるようになる。提案機能の有用性に対する評価実験の結果、実際の授業で復習に用いる小テストの実施後に授業担当教員が理解度不

足と判断し説明を行った学習内容のうち、提案機能が理解度不足と判定した学習内容の割合が 86%あった。この結果から提案機能は、小テスト機能の不足を補ってクラスの学習上の問題点を客観的に明らかにし、クラスの理解度不足の学習内容を教員が的確に把握することに有用である。

今後の課題は、考察で明らかになった理解度不足の要因となる学習内容を分析し、教員へ提示できるようにすることである。

謝辞 評価実験に協力していただいた日本工業大学情報工学科の石川孝教授、および高瀬浩史准教授と勝間田仁准教授に感謝いたします。

参考文献

植野真臣. (2007). e ラーニングにおけるデータマイニング, 日本教育工学会論文誌, Vol. 31, No. 3, pp. 271-283.

江本理恵. (2010). ICT を活用した教育支援システムの導入とファカルティ・ディベロップメント-岩手大学の事例から-, 国立教育政策研究所紀要 第 139 集, pp. 73-84.

倉田博史, 星野崇宏. (2009). 入門統計解析, 新世社.

駒林邦男. (1986). 学校のつまづきをどうするか, 明治図書.

佐伯胖. (2007). 理解とは何か (コレクション認知科学 2), 東京大学出版会.

佐藤隆博. (1975). S-P 表の作成と解釈 -授業分析・学習診断のために-, 明治書店.

清水康敬. (2007). ICT 活用による FD の現状と NIME の取り組み, メディア教育研究, 第 4 巻, 第 1 号, pp. 1-8.

竹内俊彦. (2006). 項目反応理論入門, 青山学院大学総合研究所.

生田目康子. (2006). WBT による形成的評価の改善, 日本教育工学会論文誌, Vol. 29, No. 4, pp. 483-490 .

橋本重治. (1972). 教師自作テストのつくり方 思考・創造・理解・技能などの作問法, 日本図書文化協会.

福田剛志, 徳山豪, 森本康彦. (2001). データマイニング, 共立出版.

放送大学学園. (2011). 平成 21 年度・22 年度 文部科学

省先導的大学改革推進委託事業「ICT 活用教育の推進に関する調査研究」, 放送大学学園.

宮川治, 当麻喜弘. (2002). サービス動作時におけるソフトウェアの信頼性評価, 電子情報通信学会論文誌, J85-D-I, No. 2, pp. 202-209.

和田武. (2012). CMS 小テスト問題分析による授業改善の試み, 学術情報処理研究, No. 16, pp. 167-173.

Agrawal, R., Imielinski, T., & Swami, A. (1993). Mining association rules between sets of items in large databases., In ACM SIGMOD Record, Vol. 22, No. 2, pp. 207-216.

Kato T., & Ishikawa T. (2013). Analyzing Online Quiz Responses to Support One-to-One Instruction in the Classroom, International Conference on Computers in Education 2013, C4-P-198.

Moodle. <http://moodle.org/>

Quinlan, R. (1993). C4.5: Programs for machine learning., Morgan Kaufman Publishers.

Romero, C., Zafra, A., Luna, J. M., & et al. (2013). Association rule mining using genetic programming to provide feedback to instructors from multiple-choice quiz data, Expert Systems, Vol. 30, No. 2, pp. 162-172.

Romero, C., Ventura, S., & Garcia, E. (2008). Data mining in course management systems: Moodle case study and tutorial., Computers & Education, Vol. 51, No. 1, pp. 368-384.

The R Project. The R project for statistical computing. <http://www.r-project.org/> (参照 2015. 2. 21)

William H. Rice IV. (2008). Moodle 1.9 E-Learning Course Development, Packt Publishing.

著者紹介

加藤利康



2003 年より民間企業において Web システムの開発に従事。2014 年日本工業大学大学院博士後期課程修了。博士（工学）。2018 年より日本工業大学 共通教育学群 講師。現在に至る。e ラーニングによる学習履歴に対する学習分析の研究

に従事。情報処理学会、電子情報通信学会、教育システム情報学会、学習分析学会各会員。