

プログラミング教育における学習者の特徴的状況の定量化

Quantification of Learner's Characteristic Situation in Programming Exercise

藤原 理也†
Masaya Fujiwara田口 浩†
Hiroshi Taguchi島川 博光†
Hiromitsu Shimakawa

1. はじめに

プログラミング教育において指導を必要とする学習者を把握することは、教員にとって多大な負担である。そこで、指導を必要とする学習者を自動で把握する手法が必要となる。自動把握のためには、指導を必要としている学習者の状況に現れる特徴や傾向を定量化しなければならない。

本論文では、学習者の特徴的状況を定量化する手法を提案する。本手法は、学習者から複数のデータを時系列データとしてとらえ、グラフ上で同期表示することにより、学習者の学習状況を示す特徴や傾向を発見する。また、表現可能なグラフの特徴や傾向を定性表現としてモデル化しておくことにより、視覚的な操作で特徴や傾向を抽出する。抽出した特徴や傾向は、定性表現モデルに合わせた数式により定量化する。

2. 学習者データを用いた教員支援

2.1 プログラミング教育における教員支援の必要性

大学などにおけるプログラミング演習では、1人の教員が大勢のプログラミング学習者に指導を行う。教員が学習者ひとりひとりの学習状況を把握することは困難である。したがって、教員は理解の遅れなどの指導を必要とする学習者を見落とす。指導を必要とする学習者の学習状況を把握するためには、指導を必要とする学習者を自動で把握するシステムが必要である。図1に示すように、このようなシステムは、適切な指導を行うために、指導が必要な学習者の理解状況と取り組み状況を演習中に教員に伝える。その実現のためには、まず指導を必要とする学習者が理解困難な状況時に現れる特徴や傾向を自動で定量化し、把握しなければならない。ところが、学習者が理解困難な状況時に現れる特徴や傾向を定量化する適切な手法は未だ見当たらない。

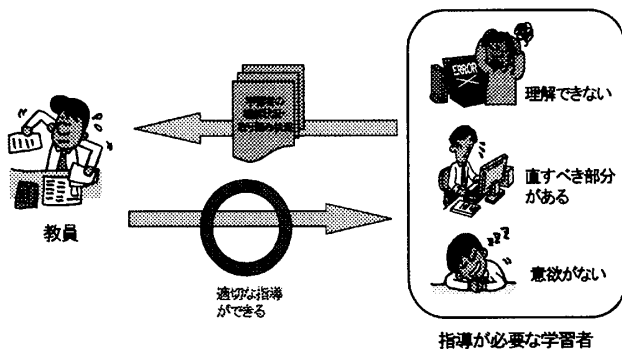


図1: 教員支援システム

2.2 学習者データ

学習者がプログラミング演習中に作成しているソースコードやコンパイルなどの操作回数を定期的に収集し、これらのデータからプログラミングの進捗を示す変量を算出する。本論文では、これらの変量を学習者データとして扱う。これらの変量にはソースコードのバイト数、コメントの数、予約語の数、関数の数、コンパイルの回数、コンパイル時のエラー数などが含まれる。

3. 学習者データからの特徴的状況の抽出

3.1 学習者データに現れる特徴の抽出手法

本論文では、学習者データから学習者の状況を示す特徴や傾向を抽出し、定量化する手法を提案する。学習者の特徴や傾向を抽出し定量化するまでの流れを図2に示す。まず、教員が指導を必要としている学生の特徴や傾向を示していると考えられる学習者データを表示する。次に、特徴や傾向を視覚的操作により抽出する。特徴や傾向を、定性モデルとしてグラフから抽出する。最後に、学習者データから見つけ出した特徴や傾向をデータ値の上昇率やデータ値の範囲として数式化することにより、抽出した定性モデルを定量モデルに変換する。

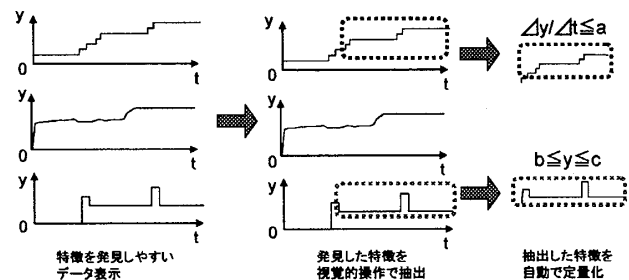


図2: 特徴の抽出

3.2 特徴検知の容易なデータ表示

学習者の状況を示す特徴や傾向は、複数の学習者データの関連情報が重要となる。たとえば、学習者が実行時のエラーを引き起こすバグがとれずに苦しんでいるときには、コンパイルを繰り返し、コンパイル時にエラーはないが、ソースコードのバイト数や予約語の数がほとんど変化していないことが多いと予測される。この傾向を把握するためには、コンパイルの回数やコンパイル時のエラー数、そしてソースコードのバイト数と予約語の数が学習者データとして必要である。これらの、時系列データとしてとらえられた学習者データを同期させて表示すると、学習者の学習状況を示す特徴や傾向を発見することが容易になる。

3.3 視覚的手法による特徴の抽出と定量化

本手法では、学習状況を示す特徴や傾向を、定性表現としてモデル化する。本手法で用いる定性表現モデルには、値の下限、値の上限、変化量、値の増減、微分値の増減の5つである。これらの5つのモデルを用いることで、グラフの特徴や傾向が表現できる。たとえば、ある学習者のコンパイルエラーが解消できない状況を考える。コンパイルエラーを解消できない状況において、コンパイル回数が増加しているが、エラー数は減少していないという特徴を発見したとする。教員はコンパイル回数の増加傾向を示す部分に、微分値の増減を表すモデルを当てはめる。また、教員はエラー数が減少していない特徴を示す部分に、値の下限値を表すモデルを当てはめる。これらの定性表現モデルは、数式で表現可能である。その結果、コンパイル回数 y が増加傾向にある特徴が $\Delta y / \Delta t \geq 4$ のように数式化される。また、エラー数 e が減少していない特徴が $e \geq 1$ のように数式化される。

4. 特徴的状況抽出システム

4.1 システム構成

システム構成図を図3に示す。本システムは、主にグラフ表示機能、詳細データ表示機能、時間範囲指定機能、特徴指定機能、特徴定量化機能の5つから構成される。

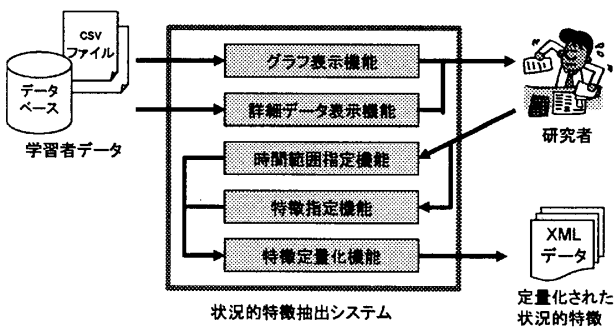


図3: システム構成図

4.2 データの表示方法

システム画面を図4に示す。本システムは、データベースやcsv形式のデータをグラフ表示する。各データは横軸に時間を取り、常に同期表示する。

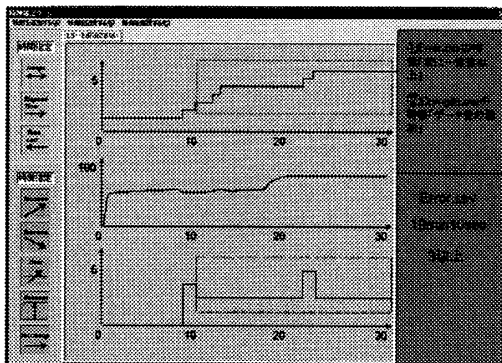


図4: システム画面

4.3 特徴の抽出方法

グラフの特徴や傾向を示すために、指定した時間範囲における上下限、変化量、値の増減、微分値の増減の5つを示す。本システムでは、これらのグラフの特徴を画面左のアイコンを用いて指定する。アイコンには時間範囲指定アイコンと、特徴指定アイコンの2種類がある。特徴を抽出する場合、最初に時間範囲指定アイコンを選択し、グラフ上の特徴が存在する範囲を指定する。続いて特徴に応じた特徴指定アイコンを選択し、特徴が現れる範囲を選択する。指定した時間範囲において複数の特徴を指定することも可能である。時間と特徴を指定すると、指定した複数の特徴や傾向を自動で数式化して記録する。

5. 評価

これまでの手法では、学習者データから学習者の状況を示す特徴や傾向の抽出が困難である。特徴の定量化に必要な機能を、特徴発見の容易性、複数データにおける特徴指定、特徴定量化の容易性、定量化された特徴の管理、汎用性の5項目について比較検証を行う。文献[1]は為替変動の予測などを行うために、データの時間的変動の特徴を分析する研究である。主に単一データの分析を目的としているため、他のデータとの比較や複数のデータにおける特徴の抽出を行うことができない。Excelは、Microsoft社の表計算ソフトである。データを様々な形式の表にして表示できることや、グラフの一部を数式化できることなどの豊富な機能から汎用性は高い。しかし、特徴の管理はユーザがしなければならないため、ユーザにかかる負担は大きい。

表1: 学習者データの分析における評価

| | 本研究 | 文献[1] | Excel |
|---------------|-----|-------|-------|
| 特徴発見の容易性 | ○ | △ | ○ |
| 複数データにおける特徴指定 | ○ | × | × |
| 特徴定量化の容易性 | ○ | ○ | △ |
| 定量化された特徴の管理 | ○ | ○ | × |
| 汎用性 | × | × | ○ |

6. おわりに

本論文では、複数の学習者データから学習者の特徴的状況を定量化する手法を提案した。本手法を用いると、複数のデータから指導が必要な学習者に現れる特徴や傾向を発見できる。また、視覚的な操作で抽出した特徴や傾向を定量化できる。今後は本システムの実装を進め、提案手法の有効性を検証する予定である。

参考文献

- [1] H. Shimakawa, H. Yamahara, Y. Imayama, M. Ushijima, S. Azuma, Pattern Refinement with Model Data Fusion to Predict Exchange Rate Movement, in A. Abraham, et al., ed., Design and Application of Hybrid Intelligent Systems, pp.662-671, IOS press, Dec., 2003