

平成21年 5月 1日現在

研究種目： 基盤研究 (C)
 研究期間： 2006～2008
 課題番号： 18500119
 研究課題名 (和文)
 コンテキスト情報を活用する事例ベース推論によるプログラミング自動評価システム
 研究課題名 (英文)
 Case-based Evaluation System of Novice Programs Using Context Information
 研究代表者
 渡辺 博芳 (WATANABE HIROYOSHI)
 帝京大学・理工学部・准教授
 研究者番号：40240519

研究成果の概要：

プログラミング演習課題の過去の評価事例を用いて自動的に評価結果を導くシステムの性能向上のため、広く評価事例を共有できるようにすることが目的である。そのために、従来のシステムでは評価結果の導出に直接利用しない情報を「コンテキスト」として扱うことのできる「事例に基づく評価処理」を提案した。また、アセンブリ言語によるプログラミングを対象にシステムを実装することで、提案した手法が機能することを確認した。

交付額

(金額単位：円)

	直接経費	間接経費	合計
2006年度	1,800,000	0	1,800,000
2007年度	900,000	270,000	1,170,000
2008年度	900,000	270,000	1,170,000
年度			
年度			
総計	3,600,000	540,000	4,140,000

研究分野：情報学，教育工学

科研費の分科・細目：情報学・知能情報学

キーワード：事例ベース推論，プログラミング，自動評価，コンテキスト，知識ベース，知識システム

1. 研究開始当初の背景

報告者らは、以前、初等アセンブラプログラミング評価支援システムとアドバイスシステムを開発し、授業等で利用してきた。本評価システムは、「テストデータによるプログラムの動作評価」と「事例ベース推論によるプログラムの作成方法の評価」を行うものである。現状では、本評価システムは、報告者の所属する大学における授業での活用に留まっているが、事例ベースを用いたシステムであるので、組織の枠を超えてグローバルな環境で利用できれば、短期間で豊富な事例

ベースを構築できるため、システムの性能が飛躍的に向上すると考えられる。

このようなシステムをグローバルな環境で活用するためにはいくつかの課題をクリアしなければならないが、その一つが「同じプログラムに対する評価結果が問題の題意によって異なる場合があるので、事例がそのまま利用できない可能性があること」である。例えば、あるプログラミング問題を学習の最初の頃に提示した場合、「正しく動作すること」だけが合格の基準となる場合がある。しかし、同じプログラミング問題をもう少し学

習が進んだ時点で提示した場合は、「正しく動作すること」の他に「シンプルで可読性が高いこと」も合格の条件に加わることも考えられる。すると、そのプログラミング問題が学習の最初の頃に提示された時には合格と判定されたプログラムと同一のプログラムが、学習が少し進んだ時点で提示された時には不合格と判定されることになる。したがって、ある同一のプログラミング問題をA大学の授業では学習の最初に提示し、別のB大学では学習が少し進んだ時点で提示したとすれば、A大学で蓄積された評価事例はB大学ではそのまま使用できない。また、提出されたプログラムを評価する教員によって、評価基準の細かい部分が若干異なることもありえる。

事例ベースによるプログラミング評価支援システムをグローバルな環境で活用するためには、このような問題を解決する必要がある。

2. 研究の目的

本研究では、先に述べた問題を解決する手法として、事例ベース推論において、コンテキスト情報をうまく扱うことに着目した。そこで、本研究では、以下を目的とした。

(1) コンテキスト情報を活用した事例ベース推論の枠組み(事例表現方法, 事例ベース構築方法, 事例検索方法, 事例ベース管理方法)を検討し、一般的なモデル化を行う。

(2) モデルに基づいてプログラミング評価支援システムを実装し、その有効性を評価する。

(3) プログラミング評価分野におけるコンテキストの特徴を整理し、コンテキストの自動推定法の可能性を検証する。

プログラミング評価支援システムの具体的な対象領域は、COMET II / CASL IIを用いた初等アセンブラプログラミング演習授業である。COMET II / CASL IIは、情報処理技術者試験で採用されている仮想計算機(CPU)とそこで動作するアセンブリ言語である。

3. 研究の方法

以下のような手順で研究を進めることとした。

(1) 普及が著しい「コース管理システム」との比較や連携においてプログラミング評価支援システムの位置づけを明確にするとともに、グローバルな環境で事例を共有可能なシステムの全体像を明確にする。

(2) コンテキスト情報を扱う事例ベース推論

システムの一般的枠組みについて、事例表現、事例ベース推論処理の流れ等を検討する。

(3) 対象領域において、コンテキスト情報を扱う事例ベース推論システムとしてのプログラミング評価支援システムを実際に設計し、実装する。システム設計においてはグローバルな環境での活用を考慮しながら、その全体像を設計し、評価支援システムとしての機能を中心に部分的に実装する。

(4) 実装したシステムの授業等での活用により、プログラミング評価分野でのコンテキスト情報を整理し、コンテキスト推定の手法について検討する。

しかし、実際には研究対象とすることができる「コンテキストの異なる実授業」が存在しなかったために、(4)については十分に進めることができなかった。

4. 研究成果

(1) 大学の基盤システムとしてのコース管理システムとの連携

研究開始時点で、高等教育の情報化や学習支援を目的とした、大学の基盤システムとしてコース管理システムが普及し始めていたので、最初に、このようなコース管理システムとプログラミング評価支援システムのような特定の分野を対象としたシステムの位置づけを明確にした。コース管理システムを中心に据え、個別の学習支援システムからの成績データ等をコース管理システムに書き込むような形での連携方法を提案した。

(2) グローバルな環境で事例を共有するシステムの全体像

本研究で対象とするシステムの全体像を検討した結果が図1である。

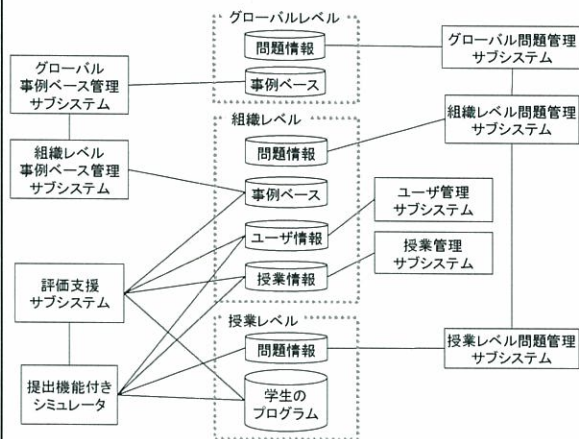


図1 システムの全体像

昨今のセキュリティ事情や基盤となる学習システムとの連携を考慮すると、大学等の組織レベルでの管理の層を介して、組織間での演習問題や評価事例ベースの共有を行うのが望ましいと考え、グローバルレベル、組織レベル、授業レベルの3層から成るシステム構成とした。

システムは、全体で7つのサブシステムで構成されている。提出機能付シミュレータで学生がプログラムを提出すると、評価支援サブシステムからプログラムと動作結果がデータベースに保存される。教員が評価を開始すると、組織レベルの事例ベース管理サブシステムが類似事例を検索・比較し、仮判定を行う。最終的には教員が評価する。比較の基準となるのは、類似事例の他、コンテキスト情報である。事例は、グローバルレベルと組織レベル間でそれぞれ、ダウンロード・アップロードが可能である。今回、開発したサブシステムは、組織レベル問題情報管理サブシステム、授業レベル問題情報管理サブシステム、授業情報管理サブシステムである。

(3) コンテキスト情報を扱う事例ベース推論の一般的な枠組み

事例ベース推論における事例は、問題記述、解記述、その他の属性から構成される。問題記述は過去に生じた問題の特徴や詳細な記述、解記述はその問題に対する解である。その他の属性は作成日時等である。また、事例ベース推論処理は、一般に、問題解析、事例検索、事例評価、事例修正・適用により、解を得るとともに、事例格納・事例ベース修復により事例ベースに対して「学習」を行う。

したがって、コンテキストを事例ベース推論で扱う方法として、コンテキスト情報を問題記述に含めることが考えられる。しかし、コンテキストと本質的な問題そのものをある程度明確に分離できるドメインでは、問題記述とコンテキストを分けて扱うことで、より高度な推論処理や事例管理を行える可能性がある。そこで、コンテキスト情報を問題記述とは分離するような推論処理を検討し以下のように定義した。

事例は、コンテキスト、問題記述、解記述から構成されるものとする。事例に基づく推論処理を以下のように行う。

- ①問題解析：入力された問題記述から検索のキーとなる特徴を抽出するとともに、問題が与えられた状況からコンテキスト情報を収集する。
- ②事例検索：まずコンテキストが同じと判定される範囲で事例を検索する。候補となる事例が得られない場合や次の事例評価の処理で適用可能な事例が得られなかった場合は、コンテキストが異なる事例を検索する。
- ③事例評価：検索の結果、候補として得られ

た事例を詳細に現問題と比較して、現問題への適用可能性を評価する。適用可能な事例が存在しない場合は、事例ベース推論で解を得ることはできない。

④事例修正・適用：コンテキストが同じ事例の場合は、事例の問題記述と現問題の差異に基づいて事例の解を修正する。コンテキストが異なる事例の場合は、問題記述の差異に加えて、コンテキストの差異に基づいた修正も必要となる。解の修正が成功したら、その解を現問題に適用する。

概ね以上のようになるが、現実の問題領域では実装が難しい部分もあるのも事実である。どんなケースにおいても問題やコンテキストの差異に基づく事例修正が可能であれば、そもそも事例を用いずに解を生成可能であることが多い。事例修正機能の実装は可能な範囲で行い、事例を豊富にすることで対応できるのが、事例ベース推論の特長である。

(4) コンテキスト情報を扱う事例ベース推論によるプログラミング評価

図1の評価支援サブシステムは、学生が提出したプログラムの動作を評価し、正しく動作しない場合は再提出を通知する。正しく動作した場合には、事例ベース推論による評価を行う。この事例ベース推論による評価処理を、先に述べた一般的な枠組みに基づいて検討した。

まず、事例の表現について、事例は、コンテキスト、問題記述、解記述、その他の属性から構成する。問題記述は、「演習問題」と学生が提出した「プログラムリスト」であり、解記述は「判定(合否)」と「アドバイス文」である。コンテキストは「学校名」「年度」「対象学年」「授業名」「教員名」「出題した授業の時限」「出題意図」「合格基準」とした。出題意図、合格基準は教員が入力した文章のまま保持する。この他に事例ベースにはインデックス情報として「各命令の数」「命令の順序情報」が記録されており、各事例と関連付ける。

次に事例に基づく評価の処理は以下のように行う。

- ①問題解析：学生が提出したプログラムリストを分析し、インデックス情報となる各命令の数をカウントする。また、問題情報にアクセスして現演習問題の出題意図と合格基準を取得する。
- ②事例検索：まず、コンテキスト情報のうち年度以外の属性が一致している範囲で評価事例を検索する。ただし、出題意図と合格基準は文章に含まれるキーワードが一致すれば一致すると判定する。複数の事例が得られた場合は、年度が近いものを優先する。候補となる事例が得られない場合や次の事例評価の処理で適用可能な事例が得られなかつ

た場合は、コンテキストが異なる事例を検索する。この場合、コンテキスト情報で一致しない属性の数が少ないものを優先する。事例検索には命令数と順序情報を用いる。

③事例評価：候補事例について、学生が提出したプログラムリストとの対応を詳細に比較して、適用可能性を判定する。この方法は以前の研究で考案した。

④事例修正・適用：問題記述(プログラムリスト)の差異による修正を行う。つまり、事例評価で得た比較結果を基にアドバイス文の語を置き換える修正を施す。一方、コンテキストの差異による修正は行わず、解の提案時に警告メッセージとしてコンテキストが異なることを提示する。

(5) システムの実装と結果

先に述べた事例の他、事例インデックス、問題情報、授業情報、ユーザ情報、提出課題データの内容をクラス図レベルで検討し、SQL のデータベーススキーマとして定義した。その後、図1におけるユーザ管理サブシステム、授業管理サブシステム、問題管理サブシステムを perl で作成した。問題管理サブシステムは、グローバル、組織レベル、授業レベルとも同機能で、保存するデータベースだけが異なるので、設定ファイルでの調整だけでプログラム自体は流用可能である。また、提出機能付きシミュレータは、JavaApplet として作成した。以上のサブシステムを統合して動作させ、帝京大学理工学部情報科学科(通信教育課程)の情報科学演習1において試用して実用性を確認した。さらに、評価支援サブシステムの評価処理の部分を Java で作成し、サンプルデータを用いて(4)で述べた評価処理が機能することを確認した。ただし、現状ではコンテキストの異なる実授業がなく、実データに対する評価はできなかった。

以上を通して、コンテキスト情報を扱う事例ベース推論の枠組みを用いたプログラミング評価は、実装可能であり、適切に機能することは確認できた。コンテキストが異なる場合でも、システムは信頼度に関する警告とともに評価結果を提示できる。これは、適用可能ではあるが、類似度の小さい事例を用いて評価結果を導出する場合と同じで、教員による最終判断は必要だが、教員に評価のための素材を提供できるケースである。したがって、本研究で提案した手法でコンテキスト情報を扱い、グローバルに事例を共有できれば、システムが評価結果を提示できるケースが広がることが示された。

また、本研究で定義したプログラミング評価におけるコンテキストには、「対象学年」「授業名」「教員名」等、容易に収集できるものと、「出題意図」「合格基準」のように教員の労力が伴うものもある。サンプルデータ

による動作実験からは、コンテキストが異なることの警告だけであれば、前者を用いるだけで十分であることが示唆された。一方、後者を用いることでコンテキストの差異に基づく事例修正の可能性があるが、後述するように事例に基づく評価処理と手続き的評価処理と融合の観点で検討すべきであると考えられる。

(6) 今後の展望

本研究では、コンテキスト情報を扱う事例ベース推論の枠組みをプログラミング評価支援システムに適用することで、システムが評価結果を提示できるケースを大きく増やせる可能性があることを示した。また、グローバルに演習問題や評価事例ベースを共有する環境での評価処理を確立した。今後、大規模化した事例ベースの保守の課題をクリアすれば、実際にグローバルに演習問題や評価事例ベースを共有する環境を有効に機能させることができると考えられる。

また、プログラミング評価分野で、コンテキストの差異に基づく事例修正の可能性を検討したとき、教員が記述する合格基準を利用することが有効である。そのためには、例えば「正しく動作すること」「シフト命令が使われていること」「可読性を高めない、冗長な命令が存在しないこと」のようなレベルで合格基準を整理し、それらを標準化した表現で表す必要がある。ただし、そこまで整理が進めば、その合格条件を確認するための処理手続きを作成することも困難ではない。そこで、合格条件を検査するための手続きを定義し、手続き的処理による評価結果と事例に基づく評価結果を総合して評価処理を行うことで評価性能を高めるアプローチを検討したい。

5. 主な発表論文等

〔雑誌論文〕(計0件)

〔学会発表〕(計1件)

- ① 渡辺博芳, 高井久美子, 武井恵雄, 古川文人, 及川芳恵: 大学の教育基盤としての CMS と個別の学習支援システムをどう連携するか?, 情報処理学会第2回 CMS 研究会予稿集, pp. 17-22 (2006).

6. 研究組織

(1) 研究代表者

渡辺 博芳 (WATANABE HIROYOSHI)

帝京大学・理工学部・准教授

研究者番号: 40240519

(2) 研究分担者

なし