

5M-9 CPUとアセンブラ授業のための事例に基づくプログラム 評価支援システム – 授業での実用に基づく評価 –

渡辺 博芳

荒井 正之

武井 恵雄

帝京大学理工学部

1. はじめに

我々は、CASLとCOMETを対象として、提示した課題に対して学生が作成したプログラムの評価を支援するシステムを実現した[1,2]。本システムが対象とする評価作業は、学生のプログラムが問題の題意をみたしているかどうかの判定と学生のプログラムに対するアドバイスの作成である。本稿では、実現したシステムの概要を述べた後、実際の授業での使用に基づく評価結果を報告する。

2. 評価支援システムの概要

2.1 プログラム評価基準

プログラムが問題の題意を満たしているとは、(a)プログラムの動作と(b)プログラムの実現方法が問題の題意を満たしていることとする。(a)は問題を解くという意味で必須条件である。(b)は教員の教育的な意図に基づいて評価される。本手法では、教員が行った評価の事例に基づいて(b)を評価するので、その判定のための知識等をシステムに埋め込む必要はない。

2.2 プログラム評価処理の流れ

前節で述べた評価基準をテストするために、(1)プログラム動作の評価、(2)冗長命令の検出、(3)事例に基づくプログラム実現方法の評価の3つの処理を行う。最初に(1)プログラム動作の評価では、あらかじめ用意したテストデータについて動作評価を行い、正しく動作しない場合は、不合格とする。次に、正しく動作したプログラムに対して、(2)冗長命令の検出処理を行う。冗長命令とは、その命令がなくともプログラムの動作に影響を与えないような命令(削除可能な命令)である。冗長命令の検出処理は合否判定には直接関係ないが、冗長命令情報を、事例照合処理と教員が判定を行ったり、アドバイスを記述する際に使用するために行う。最後に、(3)事例に基づくプログラム実現方

法の評価を行う。ここで、照合する事例が存在すれば、その事例を適用し、合否の結果を得る。照合する事例が存在しない場合は、システムによる合否判定は不可となり、教員の判定が必要になる。

2.3 事例に基づく評価処理の概要

評価事例を、プログラムリスト、判定結果、アドバイスとその他の情報から構成する。評価対象のプログラムを事例ベース内の評価事例と比較を行い、実現方法が同じと判断される事例の中で最もプログラムリストが類似している事例を選択する。実現方法が同じと判断される場合を「評価対象と事例は照合する」と言い、特にプログラムリストがほとんど同一と判断される場合を「完全照合」と呼ぶ。照合する事例が存在する場合、事例を適用する。事例の適用処理は、事例の判定結果を評価対象の判定結果とすること、事例のアドバイスに簡単な修正を施して評価対象に対するアドバイスとすることである。

2.4 実現したシステムの概要

本システムは、サーバ、教員用クライアント、学生用クライアントから構成される。教員用クライアントはhttpサーバ上のCGIとして開発し、Webベースの使い易いインターフェースを実現した。また、システムの評価結果の利用法に関して、以下の評価モードを定義し、問題ごとに評価モードを設定可能とした。

(a) 自動モード：プログラムが提出された時点で、事例との照合を行い、事例と完全に照合した場合、事例から生成された判定結果とアドバイスを自動的に学生に通知するモード。

(b) 手動モード：システムの判定結果を基に、最後は必ず、教員が判定を行うモード。

(c) 動作のみ評価モード。

3. 授業での使用に基づくシステムの評価

3.1 システム評価の目的と授業での使用の概要

実現したシステムに関して、(a)システムの判定精度と(b)教員の評価作業軽減の効果を評価することを目的として、本学2年生を対象とした演習授業の2つ

A Case-Based Evaluating System for the Students' Programs of the CPU and Assembler Course – System Evaluation Based on Practical Use

Hiro Yoshi Watanabe, Masayuki Arai and Shigeo Takei
School of Science and Engineering, Teikyo University

表 1: 問題の概要

問題名	問題の説明	対象
P1	2つの値の大きい方を求める	B組
P2	N個のデータの加算	B組
P3	N個のデータの最大値を求める	B組
P4	Nビットの循環右シフト	B組
P5	Nビットの循環左シフト	A組
P6	スタックを使って()の対応検査	B組
P7	スタックを使って文字列反転	A組

のクラス(仮にA組, B組と呼ぶ)で, 本システムを実際に使用した. 授業を履修した学生の人数は, A組が79人, B組が73人である. 本システムを使用した際に出題した問題を表1に示す. 全員が合格となるまで, プログラムの再提出を行うので, プログラムの提出数は, 学生数よりも多くなる.

3.2 評価項目

表1の問題のうちのP3に関して, 提出されたプログラムに対するシステムと教員の判定結果を表2に示す. 本システムによる提出では学生が“気軽に”提出するため, 教員に直接提出する場合は動作評価により不合格になるケースはもっと少ないと予想される. これを考慮して, 1人の学生が提出した動作評価で不合格になるプログラムを多くて2個と限定すると, 動作で不合格のケース, 提出数の合計はそれぞれ[]内の値になる. 以下の計算ではこれらの値を用いる.

(1) 判定精度: システムが判定を行ったプログラムのうち, システムと教員の判定結果が等しいケースの割合を判定精度とする. 表2から, システムが判定を行ったプログラム数は140から判定不可の19を引いた121であり, そのうちの2個を除いて教員の判定と一致している. このことから, 判定精度は約98%となる. 事例が完全照合した場合に限ると, 100%である.

(2) 教員の評価が必要なプログラム数の割合(減少率1): 事例が完全照合したケースは教員は評価を行う必要がないとすると, 表2のうちで, 教員の評価が必要なケースは判定不可の19個と, 完全照合でない4個である. これらの提出総数(140)に対する割合は約17%である. これを減少率1とする.

(3) 教員が行う必要のある評価作業の割合(減少率2): プログラムの評価作業が(a)動作の評価, (b)実現方法の評価, (c)アドバイスの作成の3つのタスクから成ると仮定し, 教員が行う必要のあるタスク数で評価する. 全体のタスク数は, (a)が140, (b)が96, (c)が140で合計376である. 教員が行った評価タスクは(b)が23, (c)が28(判定不可のケースとシステムが生成し

表 2: P3に関する判定結果の比較

システムの判定	教員の判定		合計
	合格	不合格	
合格(完全照合)	60(59)	0(0)	60(59)
不合格(完全照合)	2(0)	2(0)	4(0)
判定不可	14	5	19
小計	76	20	96
動作評価で不合格	0	116 [44]	116 [44]
合計	76	136 [64]	212 [140]

表 3: 各問題に対する評価結果

	P1	P2	P3	P4	P5	P6	P7
初期事例数	2	2	5	2	2	2	2
最終事例数	15	10	29	34	38	34	40
命令数	9	10	15	19	20	26	20
精度(%)	100	100	98	100	100	100	100
減少率1(%)	11	10	17	22	23	26	27
減少率2(%)	7.3	11	14	22	20	21	26

たアドバイスがそのまま使えなかったケースの合計)で合計51であった. 従って, 教員の評価タスク数は約14%に減少した. これを減少率2とする.

3.3 評価結果

表1の7つの問題に関して, 前節で述べた評価値をまとめたのが表3である.

(1) システムの判定精度: システムの精度に関しては, P3で98%の他は100%であり, 完全照合のケースに限ると, 全ての問題で100%であった.

(2) 教員の評価作業の軽減の効果: 表3の結果から, 教員の評価作業は, 10%~30%程度に軽減されており, 本システムによる教員の評価作業軽減の効果は顕著である. また, 最終事例数や命令数(事例の命令数の平均)が解答のバリエーションの多さを表すと見れば, 解答のバリエーションが多いほど, 教員が行う必要のある評価作業は多くなる傾向があることが分かる.

4. おわりに

実現したプログラム評価支援システムを実際の授業で用いた結果に基づき, その有効性を示した. 今後は, 問題情報エディタの開発や事例ベース管理法の検討を行う予定である.

参考文献

- [1] Watanabe, H., Arai, M. and Takei, S.: Automated Evaluation of Novice Programs Written in Assembly Language, Proc. of ICCE99, Vol.2, pp.165-168, 1999.
- [2] 渡辺博芳, 荒井正之, 武井恵雄: CPUとアセンブラ授業のための事例に基づくプログラム評価支援システム, 情報処理, Vol.99-CE-54, pp.33-40, 1999.