RGC Ref. No.:
UGC/FDS11(14)/E02/15
_____
(please insert ref. above)

RESEARCH GRANTS COUNCIL
COMPETITIVE RESEARCH FUNDING SCHEMES FOR
THE LOCAL SELF-FINANCING DEGREE SECTOR

FACULTY DEVELOPMENT SCHEME (FDS)

Completion Report
(for completed projects only)

Submission Deadlines:	1. *Auditor's report with unspent balance, if any: within **six** months of the approved project completion date.*
	2. *Completion report: within **12** months of the approved project completion date.*

Part A: The Project and Investigator(s)

1.   Project Title

An Automated Student Program Assessment Framework with Tailorable and Automated

Test Oracle for Computer Science Education

2.   Investigator(s) and Academic Department(s) / Unit(s) Involved

Research Team	Name / Post	Unit / Department / Institution
Principal Investigator	POON, Chung-keung / Professor	Department of Computing / The Hang Seng University of Hong Kong
Co-Investigator(s)	Dr. WONG Tak-lam / Faculty Member	Department of Computing Studies and Information Systems / Douglas College
	Dr. YU Yuen-tak / Associate Professor	Department of Computer Science / City University of Hong Kong
	Dr. LEE Victor Chung-sing / Assistant Professor	Department of Computer Science / City University of Hong Kong
Others		

3.   Project Duration

	Original	Revised	Date of RGC / Institution Approval *(must be quoted)*

Project Start Date	1/1/2016		
Project Completion Date	31/12/2018	30/6/2019	11/10/2018
Duration *(in month)*	36	42	11/10/2018
Deadline for Submission of Completion Report	31/12/2019	30/6/2020	11/10/2018

Part B: The Final Report

5. Project Objectives

5.1 Objectives as per original application

1. *To develop an automatic program assessment framework which can recognize a variety of admissible output variants and be applicable to a broader range of programming exercises*
2. *To develop an enhanced model of program outputs with higher expressiveness which can capture their essential characteristics for improving the effectiveness of output comparisons*
3. *To automate the matching rule generation, discovery and tailoring process by learning instructors' assessment requirements in determining the correctness of program outputs*
4. *To evaluate the effectiveness of our enhanced automatic program assessment framework by implementing and deploying it in an online platform in our institutions and collecting feedback from instructors and students via case studies and experiments*

5.2   Revised objectives

Date of approval from the RGC: N/A

Reasons for the change:

_____

_____

1.

2.

3. ....

5.3   Realisation of the objectives
(Maximum 1 page; please state how and to what extent the project objectives have been achieved; give reasons for under-achievements and outline attempts to overcome problems, if any)

For objectives 1 and 2, we designed a tree-like structure, which we called the Hierarchical Program Output Structure (HiPOS), to model the program outputs and a new automatic program assessment framework based on HiPOS. The model is more expressive than previous models and the framework is applicable to a wider varieties of program outputs.

For objective 3, we designed new algorithms for discovering frequent patterns in the presence of noise and for discovering knowledge/skill relationships using machine learning techniques. It is not straightforward to directly apply these algorithms to matching rules due to various reasons, such as the lack of labelled dataset to generalize the patterns and the highly diverse programs submitted by students. We worked around the difficulty by incorporating some heuristic techniques.

For objective 4, we have designed and implemented a platform for automated program assessment systems (APAS) and evaluated our approach for program output matching. Using the platform, we have also evaluated the effectiveness of an automatic debugging aid.

Thus, all four objectives have been accomplished.

5.4   Summary of objectives addressed to date

Objectives *(as per 5.1/5.2 above)*	Addressed *(please tick)*	Percentage Achieved *(please estimate)*
1. To develop an automatic program assessment framework which can recognize a variety of admissible output variants and be applicable to a broader range of programming exercises	✓	100%
2.To develop an enhanced model of program outputs with higher expressiveness which can capture their essential characteristics for improving the effectiveness of output	✓	100%

comparisons		
3. To automate the matching rule generation, discovery and tailoring process by learning instructors' assessment requirements in determining the correctness of program outputs	✓	100%
4. To evaluate the effectiveness of our enhanced automatic program assessment framework by implementing and deploying it in an online platform in our institutions and collecting feedback from instructors and students via case studies and experiments	✓	100%

6. Research Outcome

6.1 Major findings and research outcome
(Maximum 1 page; please make reference to Part C where necessary)

In this project, we studied the problem of testing the correctness of programs based on their outputs. In general, this is a very challenging problem in software testing. Here, we focused on students' programs in typical introductory programming courses. Such programs have much more tractable output patterns. We designed a new model, called HiPOS, to capture the structure of the program outputs and developed a framework for program output matching based on HiPOS. Our approach can handle a wider range of program outputs compared with existing state-of-the-art approaches. Our initial design makes use of a natural language parser to analyze the program output and construct the HiPOS (1st publication in Section 9). Subsequently, we enhanced our design by introducing more powerful matching rules in the framework and performed more testing (3rd publication in Section 9).

We also designed an algorithm for automatic discovery of frequent patterns in the presence of noise (1st publication in Section 8) and an algorithm for discovery of knowledge/skill relationships using machine learning technique (2nd publication in Section 9). In addition, we applied heuristics to enhance the matching rule generation (4th publication in Section 9).

Finally, we designed and implemented a platform for automated program assessment systems (APASs) and tested our approach of matching program outputs. The evaluation showed that our approach can substantially improve the matching quality and students' satisfaction (3rd paper in Section 8). We also integrated into the platform an automatic debugging aid. Evaluation results showed that the debugging aid is effective in helping students to debug their programs (2nd paper in Section 8).

6.2 Potential for further development of the research and the proposed course of action
(Maximum half a page)

Recent research in software debugging has shown increasingly better results in fault localization and program repair. A potentially interesting direction to extend our framework is to incorporate these new ideas to improve our automatic debugging aids. This will greatly enhance the usefulness of APASs and facilitate the teaching of computer programming in an online mode, which is becoming more and more important nowadays.

7. Layman's Summary

(Describe <u>in layman's language</u> the nature, significance and value of the research project, in no more than 200 words)

Automated Program Assessment Systems (APASs) have been used in many universities as a tool to aid the teaching and learning of computer programming. A well-known challenge in such systems is how to assess the program correctness based on its outputs in a flexible way so that minor and unimportant variations in the program output can be accepted as correct. Existing approaches are rather rigid, causing many programs deemed correct by a human instructor to be rejected by an APAS as incorrect. In this project, we developed a new approach for program output matching that is more flexible than the state-of-the-art approaches. Thus, we have made a step forward in making those APASs more useful and in improving the teaching and learning environment for programming classes.

Part C: Research Output

8. Peer-Reviewed Journal Publication(s) Arising <u>Directly</u> From This Research Project
(Please attach a copy of the publication and/or the letter of acceptance if not yet submitted in the previous progress report(s). All listed publications must acknowledge RGC's funding support by quoting the specific grant reference.)

The Latest Status of Publications				Author(s) *(denote the corresponding author with an asterisk[*])*	Title and Journal / Book *(with the volume, pages and other necessary publishing details specified)*	Submitted to RGC *(indicate the year ending of the relevant progress report)*	Attached to this Report *(Yes or No)*	Acknowledged the Support of RGC *(Yes or No)*	Accessible from the Institutional Repository *(Yes or No)*
Year of Publication	Year of Acceptance *(For paper accepted but not yet published)*	Under Review	Under Preparation *(optional)*						
2018				Shengxin Liu and Chung Keung Poon*	"On mining approximate and exact fault-tolerant frequent itemsets", Knowledge and Information Systems, 55(2): 361-391	Yes (2017)	Yes	Yes	Yes
2018				V.C.S. Lee, Y.T. Yu, C.M. Tang, T.L. Wong and C.K. Poon*	"ViDA: A virtual debugging advisor for supporting learning in computer programming courses", Journal of Computer Assisted Learning, 34(3): 243-258	No	Yes	Yes	Yes
2020		✓		C.M. Tang, Y.T. Yu and C.K. Poon*	Improving the Quality of Automated Program Assessment via a Versatile Test Oracle	No	Yes	Yes	N/A

9. Recognized International Conference(s) In Which Paper(s) Related To This Research Project Was / Were Delivered
(Please attach a copy of each conference abstract)

Month / Year / Place	Title	Conference Name	Submitted to RGC *(indicate the year ending of the relevant progress report)*	Attached to this Report *(Yes or No)*	Acknowledged the Support of RGC *(Yes or No)*	Accessible from the Institutional Repository *(Yes or No)*
June 2016 Atlanta, USA	Toward more robust automatic analysis of student program outputs for assessment and learning	40th IEEE Annual Computer Software and Applications Conference (COMPSAC 2016)	Yes (2017)	Yes	Yes	Yes
December 2017 Christchurch, New Zealand	An artificial intelligence approach to identifying skill relationship	25th International Conference on Computers in education (ICCE 2017)	No	Yes	Yes	Yes
July / August 2018 Osaka, Japan	Automatic assessment via intelligent analysis of students' program output patterns	International Conference on Blended Learning (ICBL 2018)	No	Yes	Yes	Yes
August 2020 Macao, China	Automatic generation of matching rules for programming exercise assessment	**Under review** for 5th International Conference on Technology in Education (ICTE 2020)	No	Yes	Yes	N/A

10. Whether Research Experience And New Knowledge Has Been Transferred / Has Contributed To Teaching And Learning
(Please elaborate)

Two undergraduate students have been involved in this research project as student helpers. Through this participation, they have learned a number of research skills as well as some advanced knowledge in natural language processing and software testing. Their undergraduate experience has been enriched.

With this experience of supervision of undergraduate students in research projects, we can improve our routine teaching and learning by incorporating some parts of the research work as part of undergraduate final year projects.

11. Student(s) Trained
(Please attach a copy of the title page of the thesis)

Name	Degree Registered for	Date of Registration	Date of Thesis Submission / Graduation
N/A			

12. Other Impact
(e.g. award of patents or prizes, collaboration with other research institutions, technology transfer, teaching enhancement, etc.)

The project strengthens the collaboration between The Hang Seng University of Hong Kong and other universities, including City University of Hong Kong, Open University of Hong Kong and Douglas College (one of the largest degree-granting colleges located in British Columbia, Canada).   The PI and co-Is are discussing ideas for future proposals.

13. Statistics on Research Outputs

	Peer-reviewed Journal Publications	Conference Papers	Scholarly Books, Monographs and Chapters	Patents Awarded	Other Research Outputs (please specify)	
					Type	No.
No. of outputs arising directly from this research project	2 plus 1 more under review	3 plus 1 more under review				

14. Public Access Of Completion Report
(Please specify the information, if any, that cannot be provided for public access and give the reasons.)

Information that Cannot Be Provided for Public Access	Reasons
N/A	