This is a preprint. Please cite as:
Zhihao Ying, Dave Towey, Anthony Bellotti, Zhi Quan Zhou, & Tsong Yueh Chen, (2021). "Preparing SQA Professionals: Metamorphic Relation Patterns, Exploration, and Testing for Big Data", in Proceedings of the International Conference on Open and Innovation Education (ICOIE 2021), pp. 22-30.

Preparing SQA Professionals: Metamorphic Relation Patterns, Exploration, and Testing for Big Data

[1]Zhihao Ying, [2]Dave Towey, [3]Anthony Bellotti,
[4]Zhi Quan Zhou and [5]Tsong Yueh Chen

[1,2,3] School of Computer Science
University of Nottingham Ningbo China
Ningbo, Zhejiang 315100, People's Republic of China
[4] School of Computing and Information Technology,
University of Wollongong, Australia
[5] Department of Computer Science and Software Engineering,
Swinburne University of Technology, Australia

[1,2,3]{zhihao.ying2,Dave.Towey,Anthony-Graham.Bellott}@nottingham.edu.cn,
[4]zhiquan@uow.edu.au, [5]tychen@swin.edu.au

Abstract
Purpose – Computer Science (CS) has been rapidly growing in popularity, partly due to the great growth of big data (BD) and other new technologies, and the services that support them. BD involves data sets that are too large to be captured, curated, analyzed, or processed (for an acceptable time and cost) using traditional CS techniques. Although BD has changed our way of living and working, it may still frequently encounter software quality assurance (SQA) problems: when testing BD systems, it may be too difficult to check whether the system behaves correctly or produces the correct output. This situation is known as the Oracle Problem, and is a major challenge for current and future (student) SQA professionals.
Metamorphic Testing (MT) has been identified as an effective approach to alleviate the Oracle Problem. MT makes use of Metamorphic Relations (MRs) across multiple test cases executions to help identify program faults. Metamorphic Exploration (ME) is a related approach that can enable a better understanding, and thus testing, of a system. Both MT and ME rely on MRs for their successful implementation. However, identification of MRs is often a manual task requiring creative thinking, and a good understanding of the system. Metamorphic Relation Patterns (MRPs) are abstractions, or templates, for multiple actual MRs.
This paper reports on an experience using MRPs to guide the identification of MRs for the implementation of ME/MT as a step towards training the next generation of SQA professionals.

Design/methodology/approach – The ME/MT experience is examined through reflection and comparison with existing MRP studies reported in the literature. Our case study involved using MRPs to identify MRs for ME/MT of BD systems. In addition to MRP-derived MRs, other MRs were derived directly from the user perspective, or from previous studies in the literature. The future use of MRP as a pedagogical tool to help train CS students and SQA professionals is examined.

Findings –We found that MRPs are useful in guiding, teaching and training the identification of MRs for MT/ME. Testers and students are able to implement MT/ME easily since the identification of MRs is a vital step of conducting MT/ME.

Originality/value/implications – This is, we believe, the first report on using MRPs to support ME/MT as a step towards training and teaching SQA professionals. We highlight the importance of BD SQA, and how MT/ME can support this. We show the usefulness of ME to prepare for MT, and MRPs to prepare good MRs to support MT/ME. We also outline several directions for follow-up action, relating to both MRP-based research and training.

Keywords: software quality assurance; big data; metamorphic testing; metamorphic exploration; metamorphic relation; metamorphic relation pattern;

1 Introduction

Computer science (CS) has been experiencing a rapid expansion in popularity, which motivates the demand for CS-related education and training. This may be partly caused by the development of big data (BD) together with the services that support it. BD involves data sets that are too large to be captured, curated, analyzed, or processed (for an acceptable time and cost) using traditional CS techniques. According to Laney (2001), BD has been described as having three main properties of big volume, variety and velocity, also known as the 3Vs model. BD has been changing people's way of living, working and thinking (Armstrong, 2014). Towey (2014) pointed out that due to the growing ubiquity of BD systems, companies are seeking workers with BD development skills, which motivates the increasing numbers of potential workers looking for ways to develop the related skills. Software testing is a major element of any software quality assurance (SQA) education and training. Basically, software testing consists of execution of a system under test (SUT) to check whether its behavior or output is correct. The mechanism used to determine correctness is called a (test) oracle. However, sometimes it is impossible or too expensive to implement an oracle, which is a situation referred to as the (test) oracle problem: systems facing the oracle problem have been called "untestable systems" (Segura et al., 2018). BD systems are typical examples of untestable systems because there is no precise functional specification for checking whether the behavior or outputs are as expected. Also, it is typically too expensive to check the behavior or outputs of BD systems due to the huge size of the data involved. Therefore, testers skilled in addressing the oracle problem for BD systems are in urgent demand (Davoudian & Liu, 2020).

Metamorphic testing (MT) is a well-developed property-based software testing technique that uses metamorphic relations (MRs) to identify possible faults in the SUT. Many studies have reported on MT being an efficient approach to alleviate the oracle problem (Chen et al., 2018; Segura et al., 2016; Zhou et al., 2018). MRs are an essential part of MT, and are used to illustrate the relation between at least two sets of test cases, the *source* and *follow-up* test cases. To conduct MT, firstly, source test cases are generated and then a pre-defined MR can be used to generate the follow-up test case(s) based on the source test case(s). Secondly, these test cases are run against the SUT, and the source and follow-up outputs are compared with the MR to check whether or not there has been a violation, which would indicate a software fault. Recently, an addition to the MT literature is metamorphic exploration (ME) by Zhou et al. (2018), which constructs MRs from the perspective of users. It enables testers and users to have a better understanding, and thus testing and use, of a system. More recently, ME has been identified as a means to scaffold the teaching and training of MT (Zhou et al., 2018; Towey et al., 2019; Yang et al., 2019).

The performance of MT is strongly influenced by the quality of MRs used (Chen et al., 2018; Segura et al., 2016). However, identification of appropriate MRs is still a major challenge, and requires creative thinking, and an adequate understanding of the SUT, especially for good MRs (Chen et al., 2018; Segura et al., 2016). Metamorphic relation patterns (MRPs) are abstractions, or templates, for various real MRs, and have been proven effective at helping to identify appropriate concrete MRs.

This paper reports on an experience using MRPs to guide the identification of likely MRs for the implementation of ME and MT as a step towards training the next generation of SQA professionals.

2 Background: MRs and MRPs

Generally, in MT, MRs are constructed based on the specifications of the SUT. However, as reported by Zhou et al. (2018), in ME, MRs need not be necessary properties of the SUT, but are properties hypothesized by users, which are so-called hypothesized MRs (HMRs). These HMRs are used for exploring the SUT, for the aim of better understanding its performance. However, the identification of likely MRs is still a challenge for testers especially beginners since it is often a manual task requiring creative thinking, and a good understanding of the system (Chen et al., 2018; Segura et al., 2016).

Recently, Segura et al. (2017) introduced the term *metamorphic relation output pattern* (MROP), which they defined as an abstract relation among the source and follow-up outputs from which multiple concrete metamorphic relations can be derived. Their work opened a new MT research direction on "metamorphic relation patterns," in a broad sense, as foreseen by Segura in his keynote at the third International Workshop on Metamorphic Testing (ICSE MET '18) (Segura, 2018).

More recently, Zhou et al. (2018) further investigated the notion of "patterns" and formally defined the general concept of a *metamorphic relation pattern* (MRP) as "an abstraction that characterizes a set of (possibly infinitely many) metamorphic relations." Zhou et al. also defined a concept of a *metamorphic relation input pattern* (MRIP) as "an abstraction that characterizes the relations among the source and follow-up inputs of a set of (possibly infinitely many) metamorphic relations."

3 An MT/ME Experience

3.1 System under test (SUT)

Electronic commerce website: Our target SUT category for ME and MT was websites or services used for electronic commerce (e-commerce). Typically, these services may provide access to very large amounts of data, including product information for potential purchases. The servers examined belong to Amazon in English and Chinese (Deloitte, 2017).

3.2 SUT Metamorphic Relation Patterns

In this study, we used MRPs and from Zhou et al. (2018) and from Segura et al. (2019) to construct new MRs.

Zhou et al. (2018) proposed a *symmetry* MRP together with a *change direction* MRIP that can be applied to different kinds of applications. These two patterns are defined based on the notion of symmetries in various areas, such as real-life nature and mathematics. The symmetry MRP leads to new MRs by using different viewpoints of the SUT from which the SUT should look the same. For instance, a query "sweaters for women" on online commerce websites should return the same number of outputs, regardless of whether the outputs are sorted by price from low to high or from high to low. The change direction MRIP constructs MRs by modifying the direction element in the source test cases to generate follow-up test cases. In the previous example, the direction in the source test case is price from *low to high*, while the direction in the follow-up test case is changed to *high to low*, thus enabling the MR to be defined.

A query represents how to retrieve and display data. Segura et al. (2019) refer to software systems that support the use of queries as query-based systems. In such systems, users can enter queries for searching information, and sorting or filtering results by specific criteria. For example, a query "sweaters for women" sorted by price from low to high on an e-commerce service like Amazon. Segura et al. (2019) presented seven MRPs for query-based systems, as follows:

- **<u>Input equivalence</u>**: The source and follow-up test cases should return the same outputs for inputs that accept fully equivalent values expressed in different forms: (1 dollar = 100 cents) should return same outputs.
- **<u>Shuffling:</u>** If the ordering criterion in the source and follow-up inputs changes, then the source and follow–up outputs should contain the same items. For instance, a query "sweaters for women" in Amazon ordered by "price from low to high" should contain the same items compared with the same query ordered by "price from high to low".
- **<u>Conjunctive conditions:</u>** If a series of new conjunctive conditions are added into source inputs to construct follow-up inputs, then the source outputs should include the follow-up outputs. For instance, a query "sweater" in Amazon should contain the items that searched through the same query filtered by a specific brand like "ZESICA".
- **<u>Disjunctive conditions:</u>** If a series of new disjunctive conditions are added into source inputs to construct follow-up inputs, then the source outputs should be included in the follow-up outputs. For instance, the outputs of query "software" in Google should be included in the outputs of query "'software' OR 'application'".
- **<u>Disjoint partitions:</u>** Suppose the input domain of at least one parameter of the source and follow-up inputs can be divided into different partitions, this MRP represents the relation that the source and follow-up outputs should be pairwise disjoint. For example, a query "sweater" filtered by a brand such as "LEANI" in Amazon should contain totally different items compared with the results of the

same query but filtered by different brand such as “ZESICA”.

- **<u>Complete partitions:</u>** Suppose the input domain of at least one parameter of the source and follow-up inputs can be divided into different partitions, this MRP represents the relations that the source outputs should contain all the follow-up outputs. For example, a query “sweater” in Amazon should contain all the items searched through the queries “sweater” filtered by “Women’s Fashion” and “Men’s Fashion”.
- **<u>Partition difference:</u>** Suppose the input domain of at least one parameter of the source and follow-up inputs can be divided into different partitions, this MRP represents the relations that the follow-up outputs are pairwise disjoint to each other and their union contains the same items as the source outputs. For example, a query “sweater” filtered by a brand such as “LEANI” in Amazon should contain totally different items compared with the results of the same query but filtered by a different brand such as “ZESICA”, and these two outputs should be same as the results searched using the same query filtered by “LEANI” and “ZESICA”.

3.3 SUT Metamorphic Relations

A total of five HMRs were used in the ME/MT study for the e-commerce sites.

- **<u>HMR1</u>**: This HMR was inspired by the symmetry MRP, and states that an online commercial server should return equivalent outputs for the same input and sorting rule, regardless of the time it is running. We consider that this HMR can be easily identified on the basis of the input equivalence pattern.
- **<u>HMR2</u>**: This HMR was also inspired by the symmetry MRP, and states that an e-commerce server should return equivalent amounts of outputs for the same input regardless of what sorting rule is used. In this study, the default and the “sort by price” sorting rules were used. We consider that this HMR can also be easily identified on the basis of the shuffling pattern.
- **<u>HMR3</u>**: This HMR was inspired by the complete partitions pattern, and states that if a user uses search functions and obtains a number of corresponding outputs, *S1*; then if the outputs are further filtered by a specific criterion, obtaining a number of corresponding outputs, *S2*; then *S1* should be greater than or equal to *S2*.

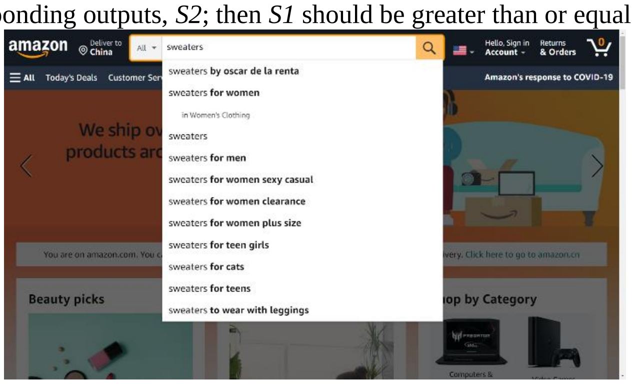

Fig. 1. Recommendation list from English Amazon web server.

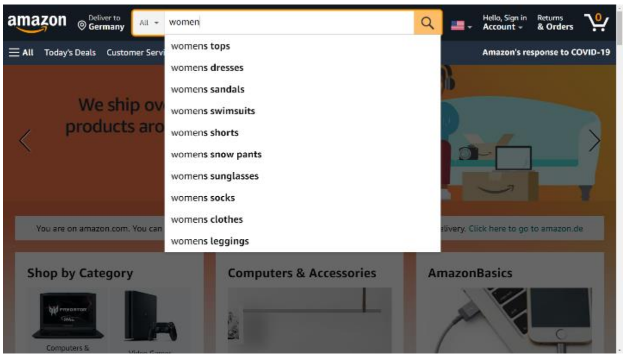

Fig. 2. Recommendation list from English Amazon web server.

- **<u>HMR4</u>**: This HMR was constructed by exploring the ways and habits of users on the e-commerce system. Related to different input languages (e.g., Chinese compared with English), an input query may be changed simply by the word order, or through some additional grammatical changes. English input queries, for example, may involve a change to the structure: “sweaters for women” compared with “womens sweaters”. Figs 1 and 2 illustrate this with recommendation lists from English Amazon. HMR4 states that if a query is slightly changed to construct a new query without changing its original meaning, then the new result should contain the same or similar items compared with the original one. We consider that this HMR can also be easily identified on the basis of the input equivalence pattern.
- **<u>HMR5</u>**: This HMR was also constructed by exploring the ways and habits of users on the e-commerce system. E-commerce servers typically offer several ways for users to search for goods, including (1) searching for items through a query only, and (2) searching for items with a query and a filter function. The filter offers many types of the items to users. Each type has several possible values. Users can select some of these values for further details of goods. HMR5 is: If a user searches for the same item using the search and filter functions with the same inputs, the results should be similar to each other. We consider that this HMR can also be easily identified on the basis of the symmetry MRP and the input equivalence pattern.

4 Evaluation and Discussion

In our exploration of both the English and Chinese versions of Amazon, we found that all hypothesized MRs (HMR1-5) were violated.

A typical example of violation on the English version of Amazon is shown in Figs. 3 and 4. The source output shown in Fig. 3 is over 300,000 items, but when the same input query was repeated at a different time, the new result (shown in Fig. 4) was over 400,000. Note that this experiment has been repeated many times on different days. Since these two queries are totally the same and the difference is not a small number, this is an apparent violation.

Fig. 3. HMR1 source test case for English Amazon web server.

Fig. 4. HMR1 follow-up test case for English Amazon web server.

Towey (2015; 2016) reported that many CS students complain that the materials they use in class and independent learning may appear too theoretical, and they prefer more practical and useful things. This reflects the tensions existing in education and training methods, such as traditional higher education and vocational and professional education and training (Towey, Walker, & Ng, 2018; 2019). In this study, we combined the education and training of SQA with software testing. We used the MRP to guide and train the identification of MRs for MT/ME. Through the implementation of MRPs, the ability to identify MRs was improved. Testers, especially beginners, should thus more easily be able to implement MT/ME since the identification of MRs is a vital step of conducting MT/ME and MRPs can support identifying a specific type of MRs.

MRPs, we argue, can not only guide and train the identification of MRs for MT/ME, but can also offer students a way to better understand and use the SUTs, especially those MRPs that can only be applied to a specific area. Because an MRP is an abstraction among a series of MRs, and in MT, the MRs generally are constructed based on the specifications of the SUT, MRPs can be regarded as a high-level abstraction among these specifications and are able to help the students better understand and use these specifications, especially the students who are new to the SUT.

5 Conclusion and Future Work

Given the growing popularity of BD systems, the need of testers to be able to apply appropriate SQA techniques is increasingly urgent. The performance of MT for testing untestable systems has been recognized, and MT has been rapidly growing in popularity. However, there is still a need to provide appropriate education and training to ensure good understanding and implementation. ME has been shown to facilitate this adoption. In this study, we have reported on an experience using MRPs to guide the identification of MRs for the implementation of ME/MT as a step towards training the next generation of SQA professionals.

This is, we believe, the first report on using MRPs to support ME or MT and train SQA professionals. In this study, a case study involved use existing MRP studies reported in the literature to identify MRs for ME or MT of BD systems. Except the MRP-derived MRs, other MRs used in this study were derived directly from the user perspective or previous MT related studies. A popular BD system has been explored and tested.

In this paper, we have highlighted that the BD SQA is an indispensable part in CS, together with how to use MT to support this. We have shown the use of ME to prepare for MT, and MRPs to construct good MRs for both ME and MT. Since the identification of MRs is a vital step in the implementation of MT and ME, we believe that based on the using of MRP, students should be better able to conduct the MT and ME.

During reflection, we found that the relations used in our study, that were derived from previous literature or users' perspectives, could be easily identified based on existing MRPs. This further supports the idea that MRPs can offer testers a direction and guideline to identify a certain type of MRs, especially those MRPs that can only be applied to a specific area. Although their application scope is limited, on the other hand, they may be able to offer both beginners and experts a more detailed and in-depth guideline for a certain type of systems.

Our future work will include the exploration and use of more MRPs to guide the identification of MRs for MT and ME. Furthermore, more MRs are also required to further validate our findings. Based on the potential for MT and ME to teach and train SQA of so many current and future systems such as BD systems, it is predicted that the positive impact of MRP's inclusion in MT and ME education and training should be able to broaden the popularity and influence of MT and ME.

Acknowledgements

The work in this paper was partly supported by a Natural Science Foundation of China (project code 61872167), and a Western River entrepreneurship grant. The authors also acknowledge the support of the University of Nottingham Ningbo China (UNNC) Artificial Intelligence and Optimisation (AIOP) Research Group, and the Faculty of Science and Engineering. We thank Morphick Solutions Pty Ltd for their discussions.

References

A. Davoudian, and M. Liu (2020). "Big Data Systems: A Software Engineering Perspective." *ACM Computing Surveys*, 53, 5, pp.110:1-110:39.

Deloitte (2017). "Global powers of retailing 2017: The art and science of customers. Deloitte Touché Tohmatsu Limited". Available: https://www2.deloitte.com/global/en/pages/consumerbusiness/articles/global-powers-of-retailing.html.

D. Laney. 3D data management: Controlling data volume, velocity and variety. META group research note, 6(70):1, 2001.

D. Towey (2014). "Open and flexible learning as an alternative in mainland Chinese higher education." In *Li, K.C. & Yuen K.S. (eds.), Emerging modes and approaches in open and flexible education*, The Open University of Hong Kong Press, Chapter 2, pp. 12–26.

D. Towey (2015). "Lessons from a Failed Flipped Classroom: The Hacked Computer Science Teacher." In *Proceedings of the IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE) 2015*, pp. 11–15.

D. Towey (2016). "Preparing the next Generation of China's Computer Scientists: A Snapshot of Challenges for Sino-foreign Computer Science Education." In *Proceedings of the Third International Conference on Open and Flexible Education (ICOFE 2016)*, Hong Kong, China, pp. 224–231.

D. Towey, J. Walker, & R.Y.-k. Ng (2018). "Traditional Higher Education Engineering versus Vocational and Professional Education and Training: What can we Learn from Each Other?" In *Proceedings of the 2018 International Conference on Open and Innovative Education (ICOIE 2018)*, pp. 474-486.

D. Towey, J. Walker, & R.Y.-k. Ng (2019). "Embracing Ambiguity: Agile Insights for Sustainability in Engineering in Traditional Higher Education and in Technical and Vocational Education and Training." *Interactive Technology and Smart Education*. https://doi.org/10.1108/ITSE-10-2018-0088.

D. Towey, S. Yang, Z. Ying, Z. Q. Zhou & T. Y. Chen (2019). "Learning by doing: Developing the next generation of software quality assurance professionals." In *Proceedings of the 2019 International Conference on Open and Innovative Education (ICOIE 2019)*, Hong Kong, July 10-12, pp. 347-355.

K. Armstrong. Big data: a revolution that will transform how we live, work, and think. *Information, Communication & Society*, 17(10):1300–1302, 2014.

S. Segura, G. Fraser, A. B. Sanchez, & A. Ruiz-Cortés (2016), "A survey on metamorphic testing." *IEEE Transactions on Software Engineering*, 42, 9, pp. 805-824.

S. Segura, J. A. Parejo, J. Troya, & A. Ruiz-Cortés (2017). "Metamorphic testing of restful web APIs." *IEEE Transactions on Software Engineering*, 44, 11, pp. 1083-1099.

S. Segura (2018). "Metamorphic testing: Challenges ahead (keynote speech)," in *Proceedings of the IEEE/ACM 3nd International Workshop on Metamorphic Testing (MET '18)*, in conjunction with the *40th International Conference on Software Engineering (ICSE '18)*, May 27, p. 1.

S. Segura, D. Towey, Z. Q. Zhou, & T. Y. Chen (2018). "Metamorphic testing: Testing the Untestable." in *IEEE Software*, vol. 37, no. 03, pp. 46-53, 2020.

S. Segura, A. Duran, J. Troya, & A. Ruiz-Cortes (2019). "Metamorphic relation patterns for query-based systems." In *2019 IEEE/ACM 4th International Workshop on Metamorphic Testing (MET)*, IEEE, pp. 24-31.

S. Yang, D. Towey, Z. Q. Zhou & T. Y. Chen (2019). "Preparing Software Quality Assurance Professionals: Metamorphic Exploration for Machine Learning." In *Proceedings of the IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE 2019)*, pp. 874-880.

T. Y. Chen, F.-C. Kuo, H. Liu, P. L. Poon, D. Towey, T. H. Tse, & Z. Q. Zhou (2018). "Metamorphic Testing: A Review of Challenges and Opportunities." *ACM Computing Surveys*, 51, 1, pp. 4:1-4:27.

Z. Q. Zhou, L. Sun, T. Y. Chen, & D. Towey (2018). "Metamorphic relations for enhancing system understanding and use." *IEEE Transactions on Software Engineering*, 46, 10, pp. 1120-1154.