This is a preprint. Please cite as:
Dave Towey, Sen Yang, Zhihao Ying, Zhi Quan Zhou and Tsong Yueh 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.

Learning by doing: Developing the next generation of software quality assurance professionals

[1] Dave Towey, [2] Sen Yang, [3] Zhihao Ying,
[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] Institute of Cybersecurity and Cryptology
School of Computing and Information Technology
University of Wollongong, Wollongong, NSW 2522, Australia
[5] Department of Computer Science and Software Engineering
Swinburne University of Technology
Hawthorn VIC 3122, Australia
[1] dave.towey@nottingham.edu.cn, [2] zy18733@nottingham.edu.cn, [3] zy19067@nottingham.edu.cn,
[4] zhiquan@uow.edu.au, [5] tychen@swin.edu.au

Abstract

Purpose – The discipline of computer science (CS) has seen a surge in popularity in education, with student demand the highest it has been for many years. A reason for this enthusiasm may be linked to the recent growth of the industry, boosted by the increasingly ubiquitous ABC of Artificial Intelligence, Big Data, and Cloud Computing. However, in addition to the well-known challenges already facing CS education professionals, a newly identified dimension of difficulty relates to the nature of ABC systems, which may be considered "untestable systems": identifying the correctness of ABC system outputs or behaviour may not be feasible, a situation known as the "oracle problem." How can we prepare CS graduates to address the software quality assurance of such systems?

A software testing approach called metamorphic testing (MT) has a proven track record of alleviating the oracle problem, and has shown great potential as a candidate testing methodology for ABC systems. The literature has reported on many instances of teaching and training users to become proficient in MT, but reports continue to identify challenges. Metamorphic exploration (ME) is a new addition to the MT literature, and involves assisting the user to become more familiar with the software under study.

This paper reports on experiences using ME as a step towards becoming proficient in MT.

Design/methodology/approach – The study is based on reflective practice, and draws on the tradition of autoethnography. Literature reviews combined with the hands-on experience of the authors inform the action research underlying this report.

Findings – We have observed very positive results from using ME as a step towards full MT. Our experience teaching and learning about MT through ME indicates that this scaffolding of the learning process, and learning through doing, may indeed be an effective pedagogical approach to prepare student testers of ABC systems.

Originality/value/implications – This is, to the best of our knowledge, the first report of using ME as a teaching or training tool. Given MT's position as one of

very few approaches suitable for testing ABC systems (the next generation of CS systems), its inclusion in CS curricula and training is essential, and is already expanding. Any approaches that assist learners' mastery of MT, including ME, will see a positive impact in CS education. This report will help educators more easily teach MT.

Keywords: computer science education; metamorphic exploration; metamorphic testing; ABC; software quality assurance; oracle problem.

1 Introduction

Higher education, in the People's Republic of China (PRC) and elsewhere, has been evolving in many ways (Fitzgerald, Gunter and White, 2012; Scott, 2015; Wang and Liu, 2011). Many PRC HE changes have been linked to China's very successful change from a manufacturing-based economy to a services-oriented one. The modern PRC IT industry is a clear example of such a successful services-oriented industry. The history of the PRC HE changes includes the emergence of multiple Sino-foreign projects (QAA, 2013), including Sino-foreign HEIs, the first of which was University of Nottingham Ningbo China (UNNC), established in 2004 (Feng, 2013; Li et al., 2012). UNNC is the affiliation for three authors of this paper, and is the site for the teaching and learning activities described herein.

Computer science (CS) has recently been experiencing a dramatic surge in popularity, in industry and in education. The modern economy seems to be hungry for CS in ways and volume that we have never seen before. Some of this enthusiasm for CS may be linked to the recent growth and increasing ubiquity of artificial intelligence (AI), big data, and cloud computing—the so-called ABC. Companies are willing to pay a premium for workers with ABC development skills, and increasing numbers of potential workers are seeking ways to acquire these skills: higher education (HE), among other options (Towey, 2014), has seen demand for CS-related courses rise dramatically, putting increasing pressure on many HE institutions (HEIs) (Towey, 2016) and CS educators.

A critical element in the education and training of CS professionals is the development of software quality assurance (SQA) skills. Traditionally, SQA may refer to activities to help ensure that we "build the product right" (verification) and "build the right product" (validation) (Boehm, 1984). Software testing has long-been considered a core element of any SQA training, and basically involves execution of the software to see if its behaviour (or produced output) are correct. Clearly, for testing to be conducted successfully, some mechanism to determine the correctness of the behaviour/output is necessary: such a mechanism is called a (testing) oracle. Systems where there is no usable oracle are often referred to as untestable systems (Segura et al., 2018b), and are considered very challenging to perform testing and other SQA on. ABC systems, by their nature, are often considered untestable.

Metamorphic Testing (MT) is an approach to alleviating the oracle problem (Chen et al. 2018; Segura et al., 2016). MT has been very successfully applied in industry, and is growing in popularity in education. A recent addition to the MT literature is the approach of metamorphic exploration (ME) (Zhou et al., 2018), which instead of

exclusively focusing on testing, draws on MT techniques to assist a user in becoming more familiar with the system under study.

This paper examines some experiences, from the perspectives of both teachers and students, of becoming proficient at using MT through a process of applying ME.

The paper is laid out as follows: Section 2 presents the background, including a brief introduction to the institution where the activities described in this paper took place (UNNC). Section 2 also introduces SQA education at UNNC, and the basic concepts of both MT and ME. Section 3 discusses the two student projects related to MT that led to the experience of using ME as a step towards MT for students. Finally, Section 4 summarises the current situation, identifies some future work, and concludes the paper.

2 Background

2.1 University of Nottingham Ningbo China

The current PRC HE landscape is very different from that of just a few decades ago, and includes a number of Sino-foreign HEIs, the first of which, UNNC (Feng, 2013; Li et al., 2012; Towey, 2016) was established in 2004. UNNC basically delivers a British education in mainland China, and has been called both an educational innovation and a centre for innovation in education (Towey, 2014). It was created in response to the PRC's growing need to provide additional HE provision, and has seen many interesting developments and projects throughout its 15-year history. It has grown from a student body of just a few hundred undergraduates to the current 7,000+ bachelors, masters, and doctoral students.

2.2 SQA Education at UNNC

The School of Computer Science at UNNC currently offers two degree courses, Computer Science (CS) and Computer Science with AI (CSAI). Both degrees have multiple modules related to software engineering and to SQA, including the final year module Software Quality Assurance (SQA) convened by one author, and the final year projects (FYPs), which can focus on various aspects of the CS or CSAI curricula. Both student authors of this paper attended the SQA module and are currently completing their FYPs related to SQA, specifically in topics involving the application of metamorphic testing.

2.3 Metamorphic Testing & Exploration

Metamorphic Testing (MT) is an approach to alleviating the oracle problem that has been growing in popularity over the 20 years since it first appeared (Chen, Cheung, & Yiu, 1998; Chen et al. 2018; Segura et al., 2016). MT has been successfully applied in industry, and its potential for application to untestable systems is making it increasingly sought-after in SQA education programmes (Towey et al., 2016). MT

uses relationships among multiple executions of the system, called metamorphic relations (MRs): Without needing to identify any individual output as correct or incorrect, identification of a violation of the MR is sufficient to know that a fault exists.

Metamorphic exploration (ME) draws on the concepts from MT, but instead of rigorously and exclusively focusing on testing, is used to help (software) system users become more familiar with the system under study (Zhou et al., 2018). We recently have had opportunity to reflect and draw on some teaching and learning experiences related to MT. In particular, through the engagement with the UNNC FYPs, when attempting to conduct MT, we discovered that initially applying ME (even unintentionally) can lead to a faster ability to be able to apply MT.

3 Two MT-related Final Year Projects

UNNC CS/CSAI final year students may choose to take a 40-credit optional project module, the final year project (FYP). Over the 2018-2019 academic year, two authors choose to undertake FYPs related to MT: Addressing the Oracle Problem in Machine Learning; and Addressing the Oracle Problem in Big Data.

3.1 Addressing the Oracle Problem in Machine Learning

Mitchell has defined machine learning (ML) as: "A computer program is said to learn from experience *E* with respect to some task *T* and some performance measure *P*, if its performance on *T*, as measured by *P*, improves with experience *E*" (Mitchell, 1997, p. 2). ML enables computers to make decisions and behave based on their learned experience. The computers can also improve through further encounters with new data. Supervised ML algorithms aim to construct a mathematical model that enables prediction of expected outputs based on class labels and features of the training data (Kotsiantis, Zaharakis, & Pintelas, 2007). Unsupervised ML algorithms usually take data without labels or classifications, and attempt to find structure in the data (Alpaydin, 2014).

Yang (2019) has spent the last year working on this project. Specifically, he aimed to conduct MT on unsupervised clustering software, including Weka[1]. He described his journey to being able to use MT as essentially having three stages. At first, he only had a basic understanding of MT concepts, and was able to implement simple instances of MT, such as for the *sine* or *shortest path* functions (Chen et al., 2004). This, he says, structured a basic comprehension and framework of MT for him.

Yang's second phase was about doing ME—although at the time, he was not aware of ME, and thought his activities were MT. As he explains, ME is a way of exploring a software system to enhance understanding and better use it. Unlike standard MT, ME does not require strictly necessary properties for construction of metamorphic relations (MRs) (Zhou et al., 2018): The MRs can be defined by developers or users to enhance familiarity with the system. In this context, they can initially be

[1] https://www.cs.waikato.ac.nz/ml/weka/

hypothesized MRs (HMRs), and may later be discarded as not truly valid. Because the testing frameworks for both ME and MT are the same, the difficulty finding correct MRs, and doing ME, is minimised. In this phase, based on his then understanding, and his reading of the MT literature, seven HMRs were identified for Weka's *k-means* clustering algorithm (Cheung, 2003). As he later remarked, the HMRs were not, at that time, validated. Attempting to follow standard MT protocol, experiments and test plans using a single data set were designed. At this point, Yang was still unaware that his activities were ME, not MT. The challenges and experiences surrounding his MT (ME) experiences further enhanced his understanding and proficiency of both Weka, and MT.

The final stage in Yang's development has been the transition to being able to use full MT. This transition has been gradual, with no clear discrete moment of ME morphing into MT. Nevertheless, there are clear differences between the ME phase and this MT phase. As with all novice uses, there are still times when mistakes are made: however, these mistakes are indeed those of a novice MT practitioner.

A further element of note relates to the desire to develop reflective practice, and to help students to become professionals with an ability to self-evaluate, reflect, and grow (Schön, 1987; Towey, 2015). Yang has been able to demonstrate this, and has been able to share his ME experiences in a major international MT forum (Yang, Towey, & Zhou, 2019).

3.2 Addressing the Oracle Problem in Big Data

Big Data means that the amount of data is so large that they cannot be captured, curated, analysed or processed by common or mainstream methods within a tolerable cost and time (Snijders, Matzat, & Reips, 2012). It has three main properties of big data, volume, variety and velocity, also known as 3Vs model, which mean large size, multiples types and high speed of generation of data (Laney, 2001). Big data has permeated our everyday lives, Facebook and WeChat being good examples, changing many aspects of how we live, work, and think (Mayer-Schönberger & Cukier, 2013). Estimates of the growth in big data include an expectation of 44 zettabytes (2^70 bytes) of global data by 2020, and 163 zettabytes by 2020 (Hajirahimova & Aliyeva, 2017; Reinsel, Gantz, & Rydning, 2017).

Ying (2019) has spent the last year working on this project, examining the oracle problem in systems including eCommerce websites, and map applications. As with Yang (Section 3.1), unknowingly, while considering his activities to be MT, Ying also spent time carrying out ME. In addition to his journey from ME to MT, Ying has also been attempting to expand the MT theory with some fundamental concept contributions relating to potential identification of MR patterns (Segura et al., 2018a; Zhou et al., 2018).

3.3 Learning by Doing

Towey (2015; 2016) has noted a common complaint among many CS students to be a perception that the materials they are required to engage with, in class and in

independent study, may appear "too theoretical." A repeated request from some students is to make things more practical, more meaningful to them. This mirrors the tensions seen in discussions of fundamental education approaches, including traditional HE and vocational and professional education and training (VPET) (Towey, Walker, & Ng, 2018; 2019).

The experience reported in this paper, although in the context of an FYP, illustrates the potential for students to learn how to do something (MT) while actually doing it (ME). The parallels with VPET are clear: VPET has its roots in learning through apprenticeships, which is perhaps the original learning through doing. Although this FYP case of ME used to develop as an MT practitioner is small, with only two students, it is something that the advisor/instructor (Towey) is planning to use more in future, both in FYPs and in more traditional module delivery.

4 Discussion & Conclusion

The FYPs are now almost completed, and the final report submissions, presentations, and conclusions will all take place before the end of May 2019. In addition to the actual project outcomes, an important take-away from this year's experience (for Towey, at least) has been the reminder of the potential for students to surprise teachers in terms of innovating beyond the planned activities (Towey, 2015). The use of ME to enable their MT was not planned by Towey—arguably, since the students were not aware of ME, it was not planned by them either: nevertheless, it was the students use of ME as a step to using MT that led to the reflection and observation of its potential to truly be an approach that can be deployed in the SQA classroom.

Our future work will obviously include completion of the FYP projects, and further exploration (including verification and validation) of their findings and implications. It is planned that our experiences of ME and MT will be incorporated into a revision of the UNNC SQA module to potentially enable such an approach to better scaffold the students' engagement with MT. Future MT-related FYP projects (and probably more senior postgraduate projects, including MSc and PhD level) will also potentially evolve to directly include ME strategies.

Finally, given the potential for MT to support the SQA of so many current and future systems (including ABC), we anticipate that the positive impact of ME's inclusion in MT training will further expand the popularity and impact of metamorphic testing.

Acknowledgements

This research was partly supported by a University of Nottingham Ningbo China (UNNC) T&L Development Grant and a UNNC Faculty of Science and Engineering T&L Conference Grant.

References

Alpaydin, E. (2014). *Introduction to Machine Learning*, 3rd ed., The MIT Press.

Boehm, B.W. (1984). "Verifying and validating software requirements and design specifications", *IEEE Software*, 1(1), pp. 75–88.

Chen, T.Y., Cheung, S.C., & Yiu, S.M. (1998). "Metamorphic testing: a new approach for generating next test cases", *Technical Report HKUST-CS98-01*, Department of Computer Science, Hong Kong University of Science and Technology, Hong Kong.

Chen, T.Y., Kuo, F.-C., Liu, H., Poon, P.L., Towey, D., Tse, T.H., & Zhou, Z.Q. (2018). "Metamorphic Testing: A Review of Challenges and Opportunities", *ACM Computing Surveys*, 51(1), pp. 4:1–4:27. http://doi.acm.org/10.1145/3143561

Chen, T.Y., Huang, D.H., Tse, T.H., & Zhou, Z.Q. (2004). "Case studies on the selection of useful relations in metamorphic testing," in *Proceedings of the 4th Ibero-American Symposium on Software Engineering and Knowledge Engineering (JIISIC'04)*, pp. 569–583.

Cheung, Y.-M. (2003). "k*-means: A new generalized k-means clustering algorithm," *Pattern Recognition Letters*, 24(15), pp. 2883–2893.

Feng, Y. (2013). "University of Nottingham Ningbo China and Xi'an Jiaotong-Liverpool University: Globalization of Higher Education in China", *Higher Education: The International Journal of Higher Education and Educational Planning*, 65(4), pp. 471–485.

Hajirahimova, M.S. & Aliyeva, A.S. (2017). "About big data measurement methodologies and indicators", *International Journal of Modern Education and Computer Science*, 9(10), pp. 1–9.

Kotsiantis, S.B., Zaharakis, I., & Pintelas, P. (2007). "Supervised machine learning: A review of classification techniques," *Emerging artificial intelligence applications in computer engineering*, 160, pp. 3–24.

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

Li, Y., Whalley, J., Zhang, S. & Zhao, X. (2012). "The higher educational transformation of China and its global implications", in Ennew, C. & Greenaway, D. (eds), *The Globalization of Higher Education*, Palgrave Macmillan, Chapter 10, pp. 135–162.

Mayer-Schönberger, V. & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.

Mitchell, T. M. (1997). *Machine Learning*, 1st ed., New York, NY, USA: McGraw-Hill, Inc.

QAA (2013). *Review of UK transnational education in China 2012: Overview*, The Quality Assurance Agency for Higher Education. Available from https://www.qaa.ac.uk/docs/qaa/international/tne-china-overview-(1).pdf?sfvrsn=e43ff481_2 [Accessed: 25 April 2019]

Reinsel, D., Gantz, J., & Rydning, J. (2017). "Data age 2025: The evolution of data to life-critical", *IDC White Paper*, pp. 1-25.

Schön, D.A. (1987) *Educating the reflective practitioner*. San Francisco: Jossey-Bass.

Scott, R.W. (2015). “Higher Education in America: Multiple Field Perspectives” in Stevens, M, & Kirst, MW (eds), *Remaking College: The Changing Ecology of Higher Education*, Stanford University Press, Palo Alto. Available from: ProQuest Ebook Central. [Accessed: 10 April 2019]

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

Segura, S., Parejo, J. A., Troya, J., & Ruiz-Cortés, A. (2018a). “Metamorphic testing of RESTful web APIs”, *IEEE Transactions on Software Engineering*, 44(11), pp. 1083-1099.

Segura, S., Towey, D., Zhou, Z.Q., & Chen, T.Y. (2018b) “Metamorphic testing: Testing the Untestable”, *IEEE Software*. https://doi.org/10.1109/MS.2018.2875968

Snijders, C., Matzat, U., & Reips, U.-D. (2012). “‘Big data’: big gaps of knowledge in the field of internet science”, *International Journal of Internet Science*, 7(1), pp. 1–5.

Towey, D. (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.

Towey, D. (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.

Towey, D. (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.

Towey, D., Chen, T.Y., Kuo, F.-C., Liu, H., & Zhou, Z.Q. (2016). “Metamorphic Testing: A New Student Engagement Approach for a New Software Testing Paradigm”, in *Proceedings of the IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE 2016)*, pp. 218–225.

Towey, D., Walker, J., & Ng, R.Y.-k. (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.

Towey, D., Walker, J., & Ng, R.Y.-k. (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

Wang, X. & Liu, J. (2011). “China’s Higher Education Expansion and the Task of Economic Revitalization”, *Higher Education: The International Journal of Higher Education and Educational Planning,* 62(2), pp. 213–229.

Yang, S., Towey, D., & Zhou, Z.Q. (2019). “Metamorphic Exploration of an Unsupervised Clustering Program”, in *Proceedings of the 4th International Workshop on Metamorphic Testing (ICSE MET 2019), in conjunction with the 41st International Conference on Software Engineering (ICSE 2019)*, ACM Press. [Accepted to Appear]

Yang, S. (2019). “Addressing the Oracle Problem in Machine Learning”, *2019 Final Year Project Report, School of Computer Science, University of Nottingham Ningbo China*. [Unpublished manuscript]

Ying, Z. (2019). “Addressing the Oracle Problem in Big Data”, *2019 Final Year Project Report, School of Computer Science, University of Nottingham Ningbo China*. [Unpublished manuscript]

Zhou, Z.Q., Sun, L., Chen, T.Y., & Towey, D. (2018). “Metamorphic relations for enhancing system understanding and use”, *IEEE Transactions on Software Engineering*, DOI: 10.1109/TSE.2018.2876433