

Institution: King's College London
Unit of Assessment: B11: Computer Science and Informatics
Title of case study: Search Based Software Testing as a Practical Tool [PEX]
<p>1. Summary of the impact (indicative maximum 100 words)</p> <p>KCL researchers, led by Professor Mark Harman, developed algorithms and contributed code to the Microsoft Pex automatic software testing tool, which is now part of the Fakes integral component on Microsoft Visual Studio 2012/2013 development environment used by tens of thousands of professional software developers. Harman's work on Search Based Software Testing (SBST) has been pioneering in the development of automatic test case generation. The Pex/Fakes tool supports software engineers in automatic test-case generation, and is in very wide use in the commercial and industrial software development environment (tens of thousands of software developers have downloaded the Pex plug-in or acquired it as the Microsoft Visual Studio Fakes component). Further impact of the SBST work carried out by the KCL research team can be seen in the adoption of SBST techniques by Google software developers and in work on the Vada software testing system carried out with DaimlerChrysler.</p>
<p>2. Underpinning research (indicative maximum 500 words)</p> <p>The Search-based Software Engineering (SBSE) research programme is based on the insight that software engineering problems can be formulated as search problems. Harman and his colleagues have pioneered the development of this agenda in software engineering over the past decade. The research work has demonstrated feasibility of SBSE across a surprisingly wide range of software engineering activities. It is fair to say that this work has changed the way that many software engineering researchers think about software engineering and, that as a result, more recently, SBSE is starting to change the practice of software engineering, through uptake in several large organisations, particularly with respect to their testing programmes.</p> <p>Professor Harman is well known for many contributions to the field of Search based Software Engineering (SBSE), a field he founded in 2001 and now with active research groups in 25 countries. At KCL, Harman founded and led the Centre for Research on Evolution, Software and Testing (CREST), which included researchers Yuanyuan Zhang, Yue Jia, Shin Yoo and Kiran Lakhotia and other KCL academic staff, including Dr Kathleen Steinhofel. Their joint work, during the REF period, has underpinned impact in search based software testing [1-6].</p> <p>Search-based Software Engineering exploits adaptive optimisation algorithms, including genetic algorithms, particle swarms, simulated annealing and other stochastic search and meta-search techniques, to optimise software engineering artefacts such as requirements, plans, test cases, architectures and the software products created by Software Engineering. Systematic search for optimisation is also exploited, including tools such as the Z3 SAT Modulo Theories solver which can solve constraints generated from code analysis in order to construct test cases that cover the behaviour of a program or program fragment. Harman and his team have developed the framework and specific instantiations of the search-based approach to provide tools that perform automatic test case generation, greatly easing the process of software development in an industrial context, where it is essential to create appropriate test cases to cover a space of possible software failures. The use of meta-heuristic search techniques allows the generation and careful selection of test</p>

Impact case study (REF3b)

cases (and, in other contexts, of requirements, plans etc) to be subjected to automated evaluation and refinement. This process, iterated many times, leads to the identification of a suite of high-value artefacts (eg test cases) fully automatically.

Key researchers

Prof Mark Harman. Professor and director of the CREST centre (at KCL from 2005 to 2010, at UCL from 2010).

Kiran Lakhotia, Research Associate in the CREST centre (at KCL from 2004, at UCL from 2010).

Yue Jia, PhD Student in the CREST centre (at KCL from 2006, at UCL from 2010).

Shin Yoo, Research Associate in the CREST centre (at KCL from 2005, at UCL from 2010 to 2012), Lecturer in Software Engineering at UCL from 2012.

Yuanyuan Zhang, Research Associate in the CREST centre (at KCL from 2006; at UCL from 2010).

Kathleen Steinhöfel, KCL academic staff 2003–

Drs. Jia, Yoo and Lakhotia were all KCL MSc students then KCL PhD students and then KCL RAs under Harman's supervision while at KCL. Zhang was funded by the KCL KC Wong scholarship and subsequently became an RA at KCL.

3. References to the research (indicative maximum of six references)

*[1] Mark Harman, Lin Hu, Rob Hierons, Joachim Wegener, Harmen Sthamer, Andre Baresel and Marc Roper. "*Testability Transformation*". IEEE Trans. on Software Engineering. 30(1): 3-16, 2004. DOI: 10.1109/TSE.2004.1265732.

*[2] Paul Baker, Mark Harman, Kathleen Steinhöfel and Alexandros Skaliotis. "*Search Based Approaches to Component Selection and Prioritization for the Next Release Problem*". 22nd Int. Conf. on Software Maintenance (ICSM 06). pp. 176-185. 2006. DOI: 10.1109/ICSM.2006.56

*[3] Shin Yoo, Mark Harman "*Using hybrid algorithm for Pareto efficient multi-objective test suite minimization*". Journal of Systems and Software 83(4): 689-701 (2010). DOI: 10.1016/j.jss.2009.11.706

[4] K. Lakhotia, M. Harman, P. McMinn. "*Handling Dynamic Data Structures in Search Based Testing*". Proc. of the Genetic and Evolutionary Computation Conference (GECCO 2008). pp. 1759-1766. 2008. DOI: 10.1145/1389095.1389435

[5] Yuanyuan Zhang, Enrique Alba, Juan J. Durillo, Sigrid Eldh and Mark Harman. "*Today/future importance analysis*". ACM Genetic and Evolutionary Computation Conference (GECCO 2010). pp. 1357-1364. 2010. DOI: 10.1145/1830483.1830733

[6] Kiran Lakhotia, Phil McMinn and Mark Harman, "*Automated Test Data Generation for Coverage: Haven't We Solved This Problem Yet?*", Proceedings of the 2009 Testing: Academic and Industrial Conference - Practice and Research Techniques, p.95-104, September 04-06, 2009 DOI: 10.1109/TAICPART.2009.15

* Publications indicating quality of underpinning research.

Harman was supported by EPSRC Platform grant EP/G060525/1 "CREST: Centre for Research on Evolution, Software and Testing", worth £1,091,081 2009-2010 (held at KCL)

4. Details of the impact (indicative maximum 750 words)

Although the use of search-based methods, including constraint reasoning, has had an impact across many areas with software engineering, the SBSE approach has had particularly strong impact on Software Testing. All software organisations depend crucially on their testing to ensure that software performs within acceptable engineering tolerances for both the functional properties and the behavioural (non-functional) properties of the systems. Research carried out by Professor Harman and his colleagues is used by many commercial and industrial software developers across the world, embodied in the Microsoft Pex tool and the Microsoft Visual Studio integral component, Fakes. The head of the Microsoft Pex development team reports that *“development of Pex...has been very significantly influenced by the work of Professor Harman and his group.”*

Harman and his post-doctoral researcher Kiran Lakhotia worked with Microsoft to incorporate their work on search based test optimisation into the Microsoft Pex automated software testing tool. These techniques are available online as plug-ins for the .Net Visual Studio development tool and have, together, been downloaded more than 50,000 times. The work that underlies this tool was published in 2008 [LHM08]. The work on Pex was cited in the recent Microsoft report on the impact of research on Microsoft's software testing programme that appeared at the top-tier Software Engineering conference, ICSE, in 2011.

Lakhotia was seconded to Microsoft in 2009, while employed as a Research Assistant at KCL in Harman's CREST team. During his secondment, he implemented the floating point test case analysis parts of Pex. This work is described in detail in the paper on FloPSy [A] coauthored by Lakhotia, Harman and the head of the Microsoft Pex development team, in 2010. In addition to this direct contribution to the Pex code-base, Harman and his team influenced and informed the development of Pex throughout the project and *“have continued to influence and contribute to the development of Pex”* [C].

Microsoft's Pex development team leader reports [C]: *“Without Harman's work, there is little doubt that the Pex tool would not exist in its current form and his contributions have had a significant impact on the relevance and value of the Visual Studio plug-in downloaded by more than 50,000 users worldwide. The overall success of the plug-in eventually caused the Visual Studio product group to ship Fakes, one core component of Pex, as an integral part of Visual Studio 2012/2013.”*

Harman and another post-doctoral research, Shin Yoo worked with Google to implement a regression test optimisation system. This system was deployed for a period of two weeks (before it became overloaded with internal demand and had to be disabled) as an integrated part of Google's entire code base to collect data for analysis and optimisation. Google made a research award of £25k to support this work. The ideas that underpin this work were published in a 2007 paper at the ACM International Symposium on Software Testing and Analysis (ISSTA), subsequently presented in a more detailed journal version [3], and a paper describing the collaboration and its effects at Google [B] was published more recently.

DaimlerChrysler funded research on development of the Vada system by Harman and his team at KCL, in a series of research projects that accrued cash funding of approximately £60k, much of which was received by KCL.

SBST is changing the way software engineers build unit tests for code and is affecting the software engineering process. The Pex/Code Digger tool is widely discussed in the software development

Impact case study (REF3b)

community [E] demonstrating its impact on some of the most influential figures in the software development world.

5. Sources to corroborate the impact (indicative maximum of 10 references)

The following papers record the evidence of the collaborations between KCL researchers and industrial users:

[A] K. Lakhotia, N. Tillmann, M. Harman, J. de Halleux, “*FloPSy - Search-Based Floating Point Constraint Solving for Symbolic Execution.*” 22nd IFIP International Conference on Testing Software and Systems, Natal, Brazil, November 8-12, 2010, pp. 142-157.

[B] Shin Yoo, R. Nilsson, M. Harman. “*Faster Fault Finding at Google using Multi Objective Regression Test Optimisation*” ACM SIGSOFT Symposium on the Foundations of Software Engineering (FSE 2011), Industry track. Szeged, Hungary, 5th-9th September 2011.

The following people can corroborate the contributions of KCL researchers to the work, as described above and confirm the quoted material in the above.

[C] Head of Microsoft Pex Development Team (Letter supplied)

[D] Test Software Engineering Team, Google, Zurich

[E] Articles corroborating range of users of Pex: <http://research.microsoft.com/en-us/projects/pex/>