HOT RESEARCH

Debugging: From the Inside Out

Software is dynamic and complex. It is a rich artifact that reflects years of design, maintenance, creativity, utility, and changing requirements. Moreover, once executed, the software comes alive: The intentions of the developers are realized (or not) only when the software is put in action. It is this dynamic nature of software that intrigues ISR Prof. James A. Jones.

Jones conducts research with an aim to improve the quality of software and the efficiency with which it is developed and maintained. He specifically is interested in the research areas of software analysis, testing, and visualization. But it is the complex inner workings of the system — workings so complex that often no single developer can fully fathom all of the component mechanisms and their inter-relationships — that fascinate Jones and inform his research efforts.

“It’s like a clock. The face, with its minute and hour hands, performs a function, but it’s not as interesting as opening it up and seeing the cogs move and mesh — this is where the magic happens,” remarks Jones. “There are so many software artifacts that we can study: the static structure of the code, the dynamic behavior of the running system, the paths executed through the instructions, the values of the variables, the evolution of the system in a version control repository, the bug reports, and so on,” Jones observes. “It’s like peeking under the hood and figuring out what’s happening.”

As an example of this dynamic information, the software testing process of many software organizations produces “coverage information.” This coverage information is produced as a part of white-box testing to assess the extent to which the software under test has been exercised. The original program is “instrumented” to cause it to keep track of which instructions of the program were exercised by each test case. The goal is to ensure that all parts of the system are being tested — otherwise, the developers should have little confidence in the correctness of the parts of the software that have never been tested or executed.

Jones noticed that this coverage information can be a wealth of residual evidence that can be used to reveal how the software executes beyond its original intent of determining testing adequacy. The presence of this information caused Jones to question how the coverage information could be used to automatically suggest where bugs might lie in the program in the event of test case failures. The idea is to use statistical inferencing techniques to identify locations in the code that are most suspicious of causing failures. Jones likes to explain it as such: “Each instruction in the program likely influences the result of the test case. When the program fails, one of those instructions that were executed is likely to be the cause of the failure. Thus, if we can look at the aggregate of all of the test case passes and failures, we can start to infer which instructions are more suspicious of causing the failures. Those instructions more often executed in a failing context and least often in the passing context are suspicious of being the bug that caused those failures.”

RESEARCH BRIEFS

Director Richard N. Taylor has been awarded $500,000 by the NSF for research on “Making and Tracing: Architecture-centric Information Integration.” This work will be conducted with ISR alumna and post-doctoral researcher Hazel Asuncion.

Prof. Gloria Mark has been awarded $479,270 by the NSF for “Collaborative Research: Widescale Computer-Mediated Communication in Crisis Response: Roles, Trust & Accuracy in the Social Distribution of Information.” This project will examine the use of computer-mediated communication in citizen response to crisis and focuses on reliability, trust, and accuracy in information.

Profs. Walt Scacchi and Tony Wasserman have been named as Founding Members of Open Source for America, a coalition of industry leaders, non-government groups, and academic/research institutions organized to serve as a centralized advocate for encouraging broader U.S. Federal Government support of and participation in free and open source software. Wasserman has also been named to the Board of Advisors.

Prof. Crista Lopes, one of the main architects of the open source OpenSim virtual world development platform, served as a guest participant at Intel CEO Justin Rattner’s opening keynote address at Super Computing 2009.

IN THIS ISSUE:

2-Message from the Director
3-Meet Prof. Jim Whitehead
4-5-Student News
5-Research Briefs
6-Northrop Grumman/ISR Collaborate
6-Tech Reports
6-7-Student News
7-ISR Sponsorship
8-Does Culture Affect Software Use?
8-ISR Events
Jones created a technique and a tool called Tarantula that automatically finds parts of a program’s structure that are suspicious of causing failures. The output of the tool is presented to the developer with a high-level view of the internal components of the program, with the suspiciousness of the components encoded with color: red denotes suspiciousness and green denotes safety. For the programs and bugs studied, Jones found that up to 90% of the bugs were identified by the tool.

Another use of this execution information from the testing process allows Jones to automatically classify failures according to the bugs that caused them. Each execution produces a “fingerprint” of the execution behavior — the paths and profiles exercised in the system. These fingerprints are used by Jones and colleagues to cluster similar executions in an attempt to identify grouping of similar failures that were caused by similar bugs.

MESSAGE FROM THE DIRECTOR

Industry-university partnerships take many forms, and December 2009 saw a new partnership and a new form – at least for me. Starting December 7th, I taught a class on software architecture at the Mysore, India campus of Infosys Technologies, Ltd. The class was 3 weeks long, and team-taught with USC Professor and ISR faculty affiliate Neno Medvidovic, and assisted by ISR’s Dr. Hazel Asuncion and USC’s George Edwards. The class was based upon the software architecture textbook that Neno, Dr. Eric Dashofy from The Aerospace Corporation, and I completed a year ago.

For me, this was an opportunity to understand at a deeper level the strategies and approaches taken by a major worldwide player in the software industry. India is an indisputable force in the global software world and insights for shaping successful corporate strategies arise from understanding what approaches are effective in a global software engineering context.

The experience was truly eye-opening. The sheer scale of the software industry and its commitment to software engineering education is staggering. The Infosys Mysore campus, for instance, had, at the time I was teaching, 9000 students in residence, with room for another 5000 shortly to come on line. These are all Infosys employees – resident here for 3 or 4 months. The campus architecture was stunning, with the newest building bigger than the national capital, with 100 very well equipped large classrooms, plus additional theatres. Each morning’s edition of “The Times of India” included a story on some aspect of the software industry. One article pointed out that the industry is logging annual revenues of $60 billion U.S., and the industry is expected to grow 16% this fiscal year – despite the global economic slowdown.

But what impressed me at least as much as the scale is the depth of knowledge and the commitment to excel in software engineering. My class was composed of managers, teachers, and senior software architects. Their objective was to learn the software architecture material in depth, so that they could then take that knowledge and teach others. Given the enormous investment that Infosys has made in training and re-training engineers, I have no doubt that they will succeed.

One evidence of the Indian commitment to developing currency with the latest technical results appeared in concert with our class, and concerns our textbook. Since the class was focused on software architecture I wanted each member of the class to have a copy available. Not only did that happen, but as a consequence Wiley India Pvt. Ltd, the Indian affiliate of our publisher John Wiley & Sons, agreed to license the text for an Indian marketplace reprint edition. A small item in one sense, but suggestive.

While India is still looking to the West for technology innovation, the time is past when their presence on the software engineering world stage can be ignored. They have made, and are making, a tremendous effort to have a highly skilled cadre of software engineers. My question: is the American software industry willing to make a similar investment? Historically the answer has been no – we didn’t need to. In my opinion, that time is now long gone. Unless American industry renews its partnerships with the research community, ultimately it will be swept aside by global competition.

ISR Director Richard N. Taylor can be reached at taylor@uci.edu.
This is work that Jones started with Profs. Mary Jean Harrold and James Bowring at Georgia Tech and continues today with undergraduate researchers Frank Morales and Jordaniel Wolk at UCI. They have shown that automatically grouping failures enables programmers to debug multiple bugs independently, resulting in a 50% savings in debugging costs.

Another debugging front Jones is exploring is the process through which programmers build their understanding and context of the behavior of the program when debugging with traditional debuggers. Together with undergraduate researchers Carter Jones and Donald Stern, Jones is building systems that will enable programmers to “steer” their exploration of the code through related and suspicious parts of the program to identify and understand the bugs in order to fix them.

The prevention of bugs in the first place is the ideal. With this goal in mind, a system is being developed that analyzes the structure of the program and the live changes that are being made by multiple developers on a distributed development project. This work is being conducted with Jones’s Ph.D. student Francisco Servant and collaborator, ISR Prof. André van der Hoek. The system enables developers to be immediately aware of each other’s code changes and their impact at the moment they make them. The goal is to prevent bugs that can be introduced by a lack of awareness of this impact. Their prototype implementation displays “Spheres of Influence” on a visualization of the program for each change to the code. Overlapping spheres should prompt developers to contact one another to ensure that new bugs or incompatible expectations of the software are not introduced to the code base.

Jones, with his collaborators, is currently pursuing these projects and others in an effort to better understand software systems, produce programs with fewer bugs, and debug the systems more efficiently. And as a fortunate side effect, as Jones puts it, he will continue to be “fascinated to see the cogs whir.”

For more on Prof. Jones’s research, see: http://www.isr.uci.edu/~jajones/

Jones can be reached at jajones@ics.uci.edu.

FOCUS ON FACULTY

Meet Jim Whitehead, Ludologist and Software Entomologist

ISR Prof. Jim Whitehead is intrigued by software bugs. “Bugs seem a bit magical,” say Whitehead. “There you are, writing code that seems perfectly reasonable, in a normal state of mind and—zip—a bug slips in. How does this happen?”

This fascination with bugs has led to recent efforts to characterize the kinds of bugs that occur most frequently. “We software engineers still don’t have a general consensus on which kinds of software bugs are most common,” he says. A recently published paper with former students Sunghun Kim (now an Assistant Professor at Hong Kong University of Science and Technology, HKUST) and Kai Pan in the journal Empirical Software Engineering provides some insight. Fine-grained bug fix patterns were mined from the history of seven open source software projects. The underlying idea is that commonality in the behavior of fixing bugs can tell you something about the underlying type of bug that was fixed.

Interestingly, this work found that the most common category is bug fixes that change the parameters of method calls. “This is an interesting result,” says Whitehead. “Several researchers, including ISR Professor David Redmiles, have identified the challenges in software inter-
Bortis

Their collaboration has continued after the internship was completed. In November, Mangano and his advisor, ISR Prof. André van der Hoek, presented a position paper at the “Flexibility in Modeling Tools” workshop at IBM’s Centers for Advanced Studies annual conference, CASCON 2009, in Toronto. They demonstrated their free hand sketch-based GUI to interface with a business user.

Gerald Bortis (A. van der Hoek, advisor) spent his second consecutive summer as an intern at the IBM T.J. Watson Research Center, in the Services Department. The research group he worked with was developing a new method for modeling business process models; Mangano contributed by developing a sketch-based GUI to interface with a business user.

A Design Knowledge Collaboration Approach,” was presented at the Third Int’l Workshop on Knowledge Collaboration in Software Development in November. Both papers were co-authored by his advisor, ISR Prof. André van der Hoek.

For more information on students:
http://www.isr.uci.edu/people.html

faces. However, there is no awareness that interface problems are a leading category of bugs.” Changes to “if” statement conditionals are the second most common type of bug fix pattern. “This makes me wonder whether it’s possible to create a better ‘if,’” muses Whitehead.

Predicting where bugs are going to occur has captured much of Prof. Whitehead’s recent attention. Today, software developers typically fix bugs in a reactive way. “It’s like the game whack-a-mole,” he says. “A bug appears, is logged, then gets whacked.” Research has shown that bugs are not spread evenly across a software project. Instead, in another example of the 80/20 rule, a small number of files contain a large proportion of the bugs. If bug fixing effort could be focused proactively on that small number of buggy files, it should save money, and result in higher quality software. According to Whitehead, “there is the potential to shift software quality efforts away from reactive bug fixing and towards a proactive approach, using statistical quality assurance.”

Whitehead is currently working on creating a software imp. “Imagine you had this little imp sitting on your shoulder watching all the code you write and able to tell you whether the change you’re making is buggy or clean,” he says. “Imps are mischievous by nature, and will be wrong some of the time. But how correct does the imp need to be for you to believe what he says, and take action?” A machine learning classifier, trained on the history of changes made in a single software project, is able to take a new software change, and classify it, i.e. “predict” whether it is buggy or clean. Recent research, performed with Sunghun Kim, UC Santa Cruz Professor Ram Akella, and current student Shiv Shivaji, published at the 2009 Automated Software Engineering conference, shows it is possible to achieve an average precision of .96, meaning that 96% of the time the predictor claims a change is buggy, it really is buggy. Average recall is .67, meaning that 67% of all bugs are detected. “I think we’re at the point now where practitioners can have faith in these predictions,” says Whitehead.

Bugs aren’t the only things on Whitehead’s mind. He has recently been putting effort into computer games, and was a key participant in developing the undergraduate Computer Game Design degree program at his home university, the University of California, Santa Cruz. “I knew from talking to our incoming freshmen that there was substantial interest in learning about how to create computer games,” says Whitehead about the times before the degree was in place. “I wondered what would happen if we provided a focused degree in this area.” Launched in Fall 2007, the new degree has been attracting 100 students a year to study computer science, with an end goal of creating computer games. Involvement with the game degree naturally led to research interests in computer games. “I am fascinated by computer game level design, the way the architecture and enemy placement in a game combine to create patterns of play. Software engineers have traditionally focused on the design of code, not of content. This is primarily a content design problem, including some interplay with enemy AI.” This led to an interest in trying to create computer algorithms that could automatically create computer game levels and work in tandem with human game designers. Whitehead is one of the organizers of a workshop on procedural content generation for computer games, being held the day before the Foundations of Digital Games conference (Monterey, CA, June, 2010). Does Prof. Whitehead ever get a chance to actually play games? “Shhh. Yes! Don’t tell my advisor (ISR Director Richard Taylor), but my Ph.D. dissertation was finished two weeks later than it should have been due to the game Civilization.”
RESEARCH BRIEFS

Prof. André van der Hoek has been awarded $500,000 by the NSF for his research on “Calico: Improving Software Design Education with Tool Support, Design Exercises, and Course Modules for Sketch-Based, Continuous Design Practice and Reflection.”

Prof. Paul Dourish, together with Prof. Irina Shklovski of the University of Copenhagen, Denmark, has received a $500,000 grant from the NSF to study social networking in Kazakhstan and Russia. The project is titled “From Local Ties to Transnational Connections: The Role of Computer-mediated Communication in Relational Maintenance.”

Prof. Walt Scacchi has been awarded $120,000 by the Naval Postgraduate School for his research on “Investigating Advances in the Acquisition of Systems based on Open Architecture and Open Source Software.”

Director Richard N. Taylor and alumnus Roy T. Fielding were featured on KOCI’s “Inside OC” in an episode on the 40th Anniversary of the Internet. The episode first aired on October 7; it can be downloaded from the Inside OC website: http://www.rickreiff.com/insideoc_shows.html

Prof. David Redmiles and Donald Bren School of ICS Dean Debra J. Richardson have been named Fellows of Automated Software Engineering. ASE Fellows are bestowed in perpetuity on those who are deemed by the Steering Committee of the IEEE/ACM International Conference on Automated Software Engineering to have rendered significant and sustained contributions.

Prof. Alfred Kobsa has been appointed as Associate Editor of the new ACM Transactions on Intelligent Interactive Systems (TIIS).

ISR alumna Emily Navarro and Prof. André van der Hoek have been recognized with the 2009 Premier Award for SimSE, a game-based educational software engineering simulation environment that allows students to practice “virtual” software engineering processes in a graphical, interactive, and fun setting.

ISR STUDENT NEWSBRIEFS

Jose Romero-Mariona (D. Richardson, advisor) spent part of Summer and Fall as an intern at Panasonic Avionics Corporation in Lake Forest, CA. He focused on research and development of new technologies to secure communications between aircraft operating Panasonic equipment and ground control. In October, Romero-Mariona was awarded the “Best Paper/Presentation” award for the engineering/computer science category at the 2009 SACNAS (Society for the Advancement of Chicanos and Native Americans in Science) for his paper “Engineering Secure Software Through Requirements.” The paper was co-authored by ISR Prof. Hadar Ziv and his advisor, ISR Prof. and Donald Bren School of ICS Dean Debra J. Richardson.

Eric Kabisch (P. Dourish, advisor) has been awarded the competitive Newkirk Center for Science and Society, Newkirk Fellowship award to support his project Datascape – a periscope device that allows its operator to view invisible data about the surrounding city. Kabish also presented his paper “Mobile after-media, cultural narratives and the data Imaginary” at the Digital Arts and Culture 2009 conference (DAC 09) in December in Irvine, CA.

For more on Whitehead’s research, see: http://users.soe.ucsc.edu/~ejw/

Contact Whitehead at ejw@cs.ucsc.edu.

For more on the Workshop on Procedural Content Generation in Games, see:
http://pcgames.fdg2010.org/

A portion of a generated platform game level, showing final artwork.
Northrop Grumman and ISR join hands in Virtual World and Game Technology Research Collaboration

Working closely with sponsors results in lasting relationships between ISR and our industry partners. The relationship between ISR and long-time supporter Northrop Grumman was spear-headed by Distinguished Technical Fellow Maria “Lolo” Penedo in the Information Systems Sector, Defense Systems Division in Carson, CA years ago, and is currently also championed by Ed Tse, Chief Technology Director in the Ground Systems Business Unit in the Intelligence Systems Division of the Information Systems Sector in Redondo Beach, CA.

The areas of mutual interest between ISR and Northrop Grumman are diverse, including open source software, software architecture, software testing, knowledge management, and virtual worlds and computer game technology.

As part of our relationship, Northrop Grumman periodically hosts ISR faculty and senior researchers on site. In November, Prof. Crista Lopes spoke on the topic of “The Matrix: How Virtual Worlds Affect the Real World and Vice-Versa” in the Defense Systems Division’s Technonet speaker series. Lively discussion with Northrop Grumman attendees resulted.

Recent interactions on virtual worlds and game technology have been particularly fruitful, laying the foundation for a joint effort seeded by a Discover grant from Northrop Grumman Intelligence Systems Division, Ground Systems Business Unit. At ISR, the project is led by Prof. Walt Scacchi who brings extensive experience with computer game development to the project. He is joined by Prof. Crista Lopes, a core developer of the open source

ISR Technical Reports Available Online

ISR technical reports present information resulting from student and faculty research carried out under the auspices of the Institute. They showcase early results not available in print elsewhere. ISR technical reports are available in PDF on the ISR website. Recent reports include:

“CREST: A new model for Decentralized, Internet-Scale Applications”
Justin R. Erenkrantz, Michael M. Gorlick, Richard N. Taylor
UCI-ISR-09-4, September 2009

“An Analysis of Publish/Subscribe Middleware Versatility”
Roberto S. Silva Filho, David F. Redmiles
UCI-ISR-09-3, August 2009

“Software Transactional Distributed Shared Memory”
Alokika Dash, Brian Demsky
UCI-ISR-09-2, February 2009

“Disjointness Analysis for Java-Like Languages”
James C. Jenista, Brian Demsky
UCI-ISR-09-1, February 2009

All ISR technical reports are available at:
http://www.isr.uci.edu/tech-reports.html
OpenSim virtual world development platform, and Prof. Gloria Mark, who brings expertise in collaborative work practices performed by high tech professionals.

“Virtual worlds and computer game technologies offer many exciting new opportunities for innovation and research advances in application domains such as aerospace, advanced manufacturing, and distributed collaboration among high tech professionals,” says Scacchi. “This collaborative research partnership with Northrop Grumman can help realize new breakthroughs.”

Tse and Penedo believe that “collaborative game technology and advanced visualization techniques from virtual worlds can potentially offer game changing paradigms for how traditional military systems will operate in the future.”

Tse and Penedo believe that “collaborative game technology and advanced visualization techniques from virtual worlds can potentially offer game changing paradigms for how traditional military systems will operate in the future.”

We are in the early stages of this virtual worlds and game technology research collaboration and look forward to reporting on a sustained and rewarding joint effort in future issues of the ISR Connector.

More information on virtual worlds and game technology research is available in the Fall/Winter 2008 issue of the ISR Connector, and at:

http://www.isr.uci.edu/research-games.html

Profs. Scacchi, Lopes, and Mark can be reached at wscacchi@ics.uci.edu, lopes@ics.uci.edu, and gmark@ics.uci.edu.

Want to Get Involved?

Sponsoring ISR has many benefits. It enables your company to form closer ties with our faculty and students, puts you on the fast track to our leading edge research, and gives you first crack at our experimental software tools. Choose from five levels of sponsorship:

<table>
<thead>
<tr>
<th>Support Level</th>
<th>Annual Contribution</th>
<th>Contribution goes to</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friend</td>
<td>$10,000</td>
<td>ISR’s general research fund.</td>
</tr>
<tr>
<td>Affiliate—Research</td>
<td>$30,000</td>
<td>A designated ISR research area.</td>
</tr>
<tr>
<td>Affiliate—Visiting</td>
<td>$40,000</td>
<td>Faculty host’s research area.</td>
</tr>
<tr>
<td>Affiliate—Grad Student</td>
<td>$60,000</td>
<td>Graduate student fellowship.</td>
</tr>
<tr>
<td>Partner</td>
<td>$100,000 or more</td>
<td>Large-scale research project.</td>
</tr>
</tbody>
</table>

For more information about ISR Sponsorship, please contact:

Debra A. Brodbeck
brodbeck@uci.edu
(949) 824-2260

Become Part of the ISR Family

Rubbing elbows with ISR faculty, staff and students gives you a valuable window into the technology landscape of the future. But a relationship with ISR can be much more: Think of us as an extension of your company—a think tank, an R&D department, a research library, a consulting firm, a training department, and an employment agency, all rolled into one. More importantly, when you sponsor ISR you become part of a friendly group of folks who speak the same language and are eager to work with you to solve your current technical problems in the most cost-effective way possible.

Be part of the ISR Family—a Friend, Affiliate, or Partner.

For more information, visit:
http://www.isr.uci.edu/sponsorship.html
or contact:
Debra A. Brodbeck
Technical Relations Director
brodbeck@uci.edu, (949) 824-2260

ISR Student Newsbriefs

Nilmax Moura and Tiago Preneca (A. van der Hoek, advisor) both spent the summer as interns at Siemens Corporate Research in Princeton, New Jersey. They worked on visualization of models, efficient test case generation, and derivation of executable test cases for SCR’s Test Development Environment (TDE), a tool that automatically generates test cases from previously created UML 2.0 models. Their project manager was ISR alumnus Dr. Marlon Vieira. Moura also presented a paper, “On the Use of Emerging Design as a Basis for Knowledge Collaboration,” at the 3rd Int’l Workshop on Knowledge Collaboration in Software Development in Tokyo, Japan. The paper was co-authored by Preneca and their advisor André van der Hoek.

Derek Phister (R. Taylor, advisor) interned this summer at The Aerospace Corporation where he worked on the development of a Model Driven Development and Analysis Workbench, which provides functionality to model AADL (Architecture Analysis and Design Language) graphically using Eclipse.

For more information on students:
http://www.isr.uci.edu/people.html
Does Culture Affect Software Use?

Yong Ming Kow and his advisor Bonnie Nardi addressed the question of whether culture affects software use by comparing World of Warcraft modding in the U.S. and China. “Mods” are end-user modifications to video games. Although China has twice as many World of Warcraft players as the U.S., Chinese modders created less than a third as many mods.

U.S. modders focused on the creation of original mods. One U.S. modder alone developed over seventy mods! The most technically expert Chinese modders, however, each produced no more than 10 mods. Instead, they became teachers to their community. For the Chinese, World of Warcraft is their first encounter with a game supporting mods. A majority of Chinese modders request all the knowledge these teachers can impart. The teachers continue to create mods, but to a much lesser extent than the U.S. modders.

U.S. modders first encountered modding in 1993, with the game Wolfenstein 3D. By 2004, U.S. modders had a solid blueprint for supporting a community. When World of Warcraft was introduced in November 2004, U.S. modders had already created a knowledge sharing chatroom and wiki. Blizzard Entertainment, the maker of the game, provides an official forum in which modders and Blizzard developers interact. In April 2008, John Wiley and Sons published a book on programming in World of Warcraft written by three U.S. modders.

Chinese modders have no official forum from Blizzard. They do not have books on programming, and had much less experience organizing themselves on the Internet. The Chinese learned and experimented, and in October 2006, set up their first forum, chatroom, and wiki.

This research highlighted the importance of community culture in the developmental trajectory of software innovations. The Chinese, by prioritizing learning in their first experiences with modding, are paving their own unique community development path.

A paper based on this research, Culture and Creativity: World of Warcraft Modding in China and the U.S., will be published in Online Worlds: Convergence of the Real and the Virtual, edited by William Bainbridge (Springer 2009).

For more information on Kow and Nardi, see:
http://www.kowym.com/
http://www.isr.uci.edu/~nardi

Kow can be reached at mail@kowym.com.
Nardi can be reached at nardi@ics.uci.edu.

SPECIAL THANKS

The UCI Institute for Software Research is generously supported by:

IBM
Intel Corporation
Northrop Grumman
Panasonic Shikoku Electronics
The Aerospace Corporation

To receive the ISR Connector, send an email request to: isr@uci.edu

ISR news, including the ISR Connector, is available at the ISR website: http://www.isr.uci.edu/

For more information, contact:
Debra A. Brodbeck
Technical Relations Director
brodbeck@uci.edu, (949) 824-2260