Detecting Defects

SWE 795, Fall 2019
Software Engineering Environments
Today

• Part 1 (Lecture)(~80 mins)

• Break!

• Part 2 (Discussion)(~60 mins)
  • Discussion of readings
Detecting Defects

• Where do defects come from?

• How can defects be prevented?

• How should potential defects be communicated to developers?
Where do defects come from?

1. Omitted logic
   Code is lacking which should be present. Variable A is assigned a new value in logic path X but is not reset to the value required prior to entering path Y.

2. Failure to reset data
   Reassignment of needed value to a variable omitted. See example for "omitted logic."

3. Regression error
   Attempt to correct one error causes another.

4. Documentation in error
   Software and documentation conflict; software is correct. User manual says to input a value in inches, but program consistently assumes the value is in centimeters.

5. Requirements inadequate
   Specification of the problem insufficient to define the desired solution. See Figure 4. If the requirements failed to note the interrelationship of the validity check and the disk schedule index, then this would also be a requirements error.

6. Patch in error
   Temporary machine code change contains an error. Source code is correct, but "jump to 14000" should have been "jump to 14004."

7. Commentary in error
   Source code comment is incorrect. Program says DO i=1,5 while comment says "loop 4 times."

8. IF statement too simple
   Not all conditions necessary for an IF statement are present.
   IF A<B should be IF A<B AND B<C.

9. Referenced wrong data variable
   Self-explanatory
   See Figure 3. The wrong queues were referenced.

10. Data alignment error
    Data accessed is not the same as data desired due to using wrong set of bits. Leftmost instead of rightmost substring of bits used from a data structure.

11. Timing error causes data loss
    Shared data changed by a process at an unexpected time. Parallel task B changes XYZ just before task A used it.

12. Failure to initialize data
    Non-preset data is referenced before a value is assigned.

[Glass TSE81]
Where do defects come from?

<table>
<thead>
<tr>
<th>Gould [14]</th>
<th>Assignment bug</th>
<th>Software errors in assigning variables’ values</th>
<th>Requires understanding of behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novice Fortran</td>
<td>Iteration bug</td>
<td>Software errors in iteration algorithms</td>
<td>Requires understanding of language</td>
</tr>
<tr>
<td></td>
<td>Array bug</td>
<td>Software errors in array index expressions</td>
<td>Requires understanding of language</td>
</tr>
<tr>
<td>Novice APL</td>
<td>Naive bug</td>
<td>Iteration instead of parallel processing</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Logical bug</td>
<td>Omitting or misusing logical connectives</td>
<td>‘…seem to be syntax oversights’</td>
</tr>
<tr>
<td></td>
<td>Dummy bug</td>
<td>Experience with other languages interfering</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inventive bug</td>
<td>Inventing syntax</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Illiteracy bug</td>
<td>Difficulties with order of operations</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gestalt bug</td>
<td>Unforeseen side effects of commands</td>
<td>‘…failure to see the whole picture’</td>
</tr>
</tbody>
</table>

Adapted from Ko & Myers, JVLC05
Where do defects come from?

- Knuth [18] While writing TeX in SAIL and Pascal
- Algorithm awry
- Improperly implemented algorithms
- ‘proved…incorrect or inadequate’
- ‘not…enough brainpower’
- Blunder or botch
- Accidentally writing code not to specifications
- ‘did not preserve…invariants’
- Data structure debacle
- Software errors in using data structures
- ‘I did not remember everything’
- Forgotten function
- Missing implementation
- Language liability
- Misunderstanding language/environment
- ‘I forgot the conventions I had built’
- Module mismatch
- Imperfectly knowing specification
- ‘tried to make the code bullet-proof’
- Robustness
- Not handling erroneous input
- ‘forced me to change my ideas’
- Surprise scenario
- Unforeseen interactions in program elements
- Trivial typos
- Incorrect syntax, reference, etc.
- ‘my original pencil draft was correct’

Adapted from Ko & Myers, JVLC05
Where do defects come from?

<table>
<thead>
<tr>
<th>Source</th>
<th>Type of Defect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eisenstadt [19]</td>
<td>Clobbered memory: Overwriting memory, subscript out of bounds</td>
</tr>
<tr>
<td>Industry experts</td>
<td>Also identified why software errors were difficult to find: cause/effect chasm; tools inapplicable; failure did not actually happen; faulty knowledge of specs; “spaghetti” code.</td>
</tr>
<tr>
<td>COBOL, Pascal, Fortran, C</td>
<td></td>
</tr>
</tbody>
</table>

- Vendor problems: Buggy compilers, faulty hardware
- Design logic: Unanticipated case, wrong algorithm
- Initialization: Erroneous type or initialization of variables
- Variable: Wrong variable or operator used
- Lexical bugs: Bad parse or ambiguous syntax
- Language: Misunderstandings of language semantics

Adapted from Ko & Myers, JVLC05
Where do defects come from?

- Ko & Myers proposed a model for understanding the *cognitive* causes of defects.
- Latent errors become active errors when they breach defenses of the system.

Adapted from Ko & Myers, JVLC05
Skill / Rule / Knowledge

• James Reason proposed a taxonomy of cognitive breakdowns based on differences in type of cognition being used
  • Skill-based activity: routine, proceduralized activity
    • e.g., typing a string, opening a source file, compiling a program
  • Rule-based activity: use of rules for acting in certain contexts
    • e.g., starting to type a for loop in order to perform an action on each element of a list
  • Knowledge-based activity: forming plans & making high-level decisions based on knowledge of program
    • e.g., forming a hypothesis about cause of runtime failure

Adapted from Ko & Myers, JVLC05
# Types of skill breakdowns

<table>
<thead>
<tr>
<th>Inattention</th>
<th>Type</th>
<th>Events resulting in breakdown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Failure to attend to a routine action at a critical time causes forgotten</td>
<td>Strong habit</td>
<td>In the middle of a sequence of actions → no attentional check → contextually frequent action is taken instead of intended action.</td>
</tr>
<tr>
<td>actions, forgotten goals, or inappropriate actions.</td>
<td>intrusion</td>
<td></td>
</tr>
<tr>
<td>Interruptions</td>
<td></td>
<td>External event → no attentional check → action skipped or goal forgotten</td>
</tr>
<tr>
<td>Delayed action</td>
<td></td>
<td>Intention to depart from routine activity → no attentional check between intention and action → forgotten goal</td>
</tr>
<tr>
<td>Exceptional stimuli</td>
<td></td>
<td>Unusual or unexpected stimuli → stimuli overlooked → appropriate action not taken</td>
</tr>
<tr>
<td>Interleaving</td>
<td></td>
<td>Concurrent, similar action sequences → no attentional check → actions interleaved</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Overattention</th>
<th>Type</th>
<th>Events resulting in breakdown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attending to routine action causes false assumption about progress of</td>
<td>Omission</td>
<td>Attentional check in the middle of routine actions → assumption that actions are already completed → action skipped</td>
</tr>
<tr>
<td>action.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repetition</td>
<td></td>
<td>Attentional check in the middle of routine actions → assumption that actions are not completed → action repeated</td>
</tr>
</tbody>
</table>

Adapted from Ko & Myers, JVLC05
# Types of rule breakdowns

<table>
<thead>
<tr>
<th>Wrong rule</th>
<th>Type</th>
<th>Events resulting in breakdown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of a rule that is successful in most contexts, but not all.</td>
<td>Problematic signs</td>
<td>Ambiguous or hidden signs $\rightarrow$ conditions evaluated with insufficient info $\rightarrow$ wrong rule chosen $\rightarrow$ inappropriate action</td>
</tr>
<tr>
<td></td>
<td>Information overload</td>
<td>Too many signs $\rightarrow$ important signs missed $\rightarrow$ wrong rule chosen $\rightarrow$ inappropriate action</td>
</tr>
<tr>
<td></td>
<td>Favored rules</td>
<td>Previously successful rules are favored $\rightarrow$ wrong rule chosen $\rightarrow$ inappropriate action</td>
</tr>
<tr>
<td></td>
<td>Favored signs</td>
<td>Previously useful signs are favored $\rightarrow$ exceptional signs not given enough weight $\rightarrow$ wrong rule chosen $\rightarrow$ inappropriate action</td>
</tr>
<tr>
<td></td>
<td>Rigidity</td>
<td>Familiar, situationally inappropriate rules preferred over unfamiliar, situationally appropriate rules $\rightarrow$ wrong rule chosen $\rightarrow$ inappropriate action</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bad rule</th>
<th>Type</th>
<th>Events resulting in breakdown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of a rule with problematic conditions or actions.</td>
<td>Incomplete encoding</td>
<td>Some properties of problem space are not encoded $\rightarrow$ rule conditions are immature $\rightarrow$ inappropriate action</td>
</tr>
<tr>
<td></td>
<td>Inaccurate encoding</td>
<td>Properties of problem space encoded inaccurately $\rightarrow$ rule conditions are inaccurate $\rightarrow$ inappropriate action</td>
</tr>
<tr>
<td></td>
<td>Exception proves rule</td>
<td>Inexperience $\rightarrow$ exceptional rule often inappropriate $\rightarrow$ inappropriate action</td>
</tr>
<tr>
<td></td>
<td>Wrong action</td>
<td>Condition is right but action is wrong $\rightarrow$ inappropriate action</td>
</tr>
</tbody>
</table>
### Types of knowledge breakdowns

<table>
<thead>
<tr>
<th>Bounded rationality</th>
<th>Type</th>
<th>Events resulting in breakdown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem space is too large to explore because working memory is limited and costly.</td>
<td>Selectivity</td>
<td>Psychologically salient, rather than logically important task information is attended to → biased knowledge</td>
</tr>
<tr>
<td></td>
<td>Biased reviewing</td>
<td>Tendency to believe that all possible courses of action have been considered, when in fact very few have been considered → suboptimal strategy</td>
</tr>
<tr>
<td></td>
<td>Availability</td>
<td>Undue weight is given to facts that come readily to mind → facts that are not present are easily ignored → biased knowledge</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Faulty models of problem space</th>
<th>Type</th>
<th>Events resulting in breakdown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formation and evaluation of knowledge leads to incomplete or inaccurate models of problem space.</td>
<td>Simplified causality</td>
<td>Judged by perceived similarity between cause and effect → knowledge of outcome increases perceived likelihood → invalid knowledge of causation</td>
</tr>
<tr>
<td></td>
<td>Illusory correlation</td>
<td>Tendency to assume events are correlated and develop rationalizations to support the belief → invalid model of causality</td>
</tr>
<tr>
<td></td>
<td>Overconfidence</td>
<td>False belief in correctness and completeness of knowledge, especially after completion of elaborate, difficult tasks → invalid, inadequate knowledge</td>
</tr>
<tr>
<td></td>
<td>Confirmation bias</td>
<td>Preliminary hypotheses based on impoverished data interfere with later interpretation of more abundant data → invalid, inadequate hypotheses</td>
</tr>
</tbody>
</table>
Breakdown chain example (Part 1)

P2 has difficulty creating the specifications for the Boolean logic to check if all of the dots are eaten, as evidenced by verbal utterances; also, part of the expression was obscured, and she though the "BigDot" reference was off-screen.

Rule breakdown creating specifications for Boolean logic for seeing if all dots are eaten *(wrong action)*

Skill breakdown implementing Boolean logic *(problematic sign)*

Conditional becomes true after one dot is eaten

Missing reference to BigDot.isShowing

Conditional doesn’t check BigDot.isShowing

Pac bounces before all the dots are eaten

Knowledge breakdown understanding runtime failure *(biased reviewing)*

Rule breakdown modifying Boolean logic *(wrong action)*

She only forms one hypothesis about the cause of the failure, which is incorrect.

This causes a breakdown in modifying the Boolean logic.

Adapted from Ko & Myers, JVL05
Breakdown chain example (Part 1)

Because camera was pointing down at Pac, she was unaware that Pac was bouncing.

The fact that Pac doesn't seem to be bouncing leads her to believe he is not.

After 20 minutes, P2 reorients the camera and notices that Pac is bouncing, but assumes it was due to more recent changes and not the earlier error.

Adapted from Ko & Myers, JVLC05
Causes of defects: API misuse

• Components expose APIs which have rules about how they should be used

• What types of rules do components impose?
Causes of defects: API misuse

• Based on survey of APIs, categorized directives APIs impose on clients
• Restrictions on when to call
  • Do not call from UI thread, for debugging use only
• Protocols specifying ordering constraints
  • Method must only be called once, method must be called prior to other method
• Locking describing thread synchronization
• Restrictions on possible parameter values
  • String.replaceAll() should not include $ or \ characters in replacement string

Causes of defects: Object protocol misuse

• Examined Java code for presence of protocols, found 7.2% of types defined protocols & 13% of classes used protocols

• Most frequent causes:
  • Initialization (28.1%): calls to an instance method \( m \) without first calling initializing method \( i \)
  • Deactivation (25.8%): calls to an instance method \( m \) after calling a deactivation method \( d \)
  • Type Qualifier (16.4%): object enters a state during which method \( m \) will always fail

Causes of defects in JavaScript

- Examined 502 bug reports from 19 repos, categorizing the cause of each error
- Most common types of errors:
  - Erroneous input validation (16%): inputs passed into JS code are not validated or sanitized
  - Error in writing a string literal (13%): incorrect CSS selectors, regular expressions, forgetting prefixes, etc.
  - Forgetting null / undefined check (10%)
  - Neglecting differences in browser behavior (9%): differences in behavior of browser API across browsers
  - Errors in syntax (7%)

Ocariza et al, A Study of Causes and Consequences of Client-Side JavaScript Bugs, TSE 2016
**GRAPHICAL IDIOMS**

**GRAPHICAL SETTERS** CHANGING GRAPHICAL PROPERTIES OF THE DOM VIA API METHODS OR CSS PROPERTIES

**GB5 Unidentified Setter:**
visual property change → code fragment to mutate property

**GB6 Unobservable Setter:**
setterA and visual property change → setterB to mutate property

**GB7 Indirect Setter:**
setterA → elements which inherit properties from setterA or occlude elements mutated by setter

**GB8 Conflicted Setter:**
setterA → setterB which overwrites setterA and code fragment with alternative setter or sequencing

**GRAPHICAL QUERIES** RETRIEVING DOM ELEMENTS OR SIMILAR REPRESENTATIONS VIA QUERY METHODS OR CSS SELECTORS

**GB1 Incomplete Queries:**
queryA and elements to be matched → queryB matching only elements

**GB2 Live Query Results:**
queryA → changes to query result set over time and alternative fragment

**GB3 Overwritten Query:**
queryA → queryB intersecting mutations made by queryA and queryB' which does not

**GRAPHICAL GETTERS** OBTAINING GRAPHICAL PROPERTIES OF THE DOM VIA API METHODS

**GB4 Unidentified Getter:**
visual property → getter code fragment to retrieve

**CALLBACK IDIOMS**

**BINDING TARGETS** IDENTIFYING OR CHOOSING AN EVENT, LIFECYCLE HOOK, OR TRIGGER TO REGISTER A CALLBACK

**51%**

**29%**
Some techniques for helping developers better work with defects

• Help developers engage in better information seeking to prevent defects from ever occurring

• Use tool to find defect, report error message to developer

• Use tests to find defect, report test failures to developers
Preventing defects by supporting better information seeking

1. Help programmers recover from interruptions or delays by reminding them of their previous actions
2. Highlight *exceptional* circumstances to help programmers adapt their routine strategies
3. Help programmers manage *multiple tasks* and detect interleaved actions
4. Design task-relevant information to be visible and unambiguous
5. Avoid *inundating* programmers with information
6. Help programmers consider all relevant *hypotheses*, to avoid the formation of invalid hypotheses
7. Help programmers identify and understand *causal relationships*, to avoid invalid knowledge
8. Help programmers identify *correlation* and recognize illusory correlation
9. Highlight *logically* important information to combat availability and selectivity heuristics
10. *Prevent* programmer’s *overconfidence* in their knowledge by testing their assumptions

Adapted from Ko & Myers, JVLC05
Tools for preventing defects

• Early work in program analysis and formal methods made possible analyzing code to find inconsistencies with a specification

• But...
  • Often required extensive work to write a specification of behavior
Early 2000s

• Static analysis tools becoming sufficiently scalable to be used on real-world programs
• More emphasis on finding real-world defects rather than simply focusing on improvements in underlying analysis technology
• Several tools adopted in industry, often to address specific and important problems
Slam

Rules governing lock

Iteratively refines boolean abstraction of program to determine if there exists path that violates rules

Static Driver Verifier

06/13/2019 • 2 minutes to read • 📚 🔴 ⏰

Static Driver Verifier (also known as “StaticDV” or “SDV”) is a static verification tool that systematically analyzes the source code of Windows kernel-mode drivers. SDV is a compile time tool that is capable of discovering defects and design issues in a driver. Based on a set of interface rules and a model of the operating system, SDV determines whether the driver correctly interacts with the Windows operating system kernel.

Installing Static Driver Verifier

Static Driver Verifier is available as part of the Windows Driver Kit (WDK) in both the full WDK experience and in the standalone Enterprise WDK. In addition, the Visual C++ Redistributable Packages for Visual Studio are required for SDV to run. See the following:

- Visual Studio 2019 Redistributable
- Visual C++ Redistributable Packages for Visual Studio 2017
- Visual C++ Redistributable Packages for Visual Studio 2013

For versions of SDV available in the WDK for Windows 10, Version 1809 or earlier, the Visual C++ Redistributable Packages for Visual Studio 2012 should be installed instead of the 2017 packages.

Visual Studio Integration

Static Driver Verifier is integrated into Visual Studio. You can run static analysis on your Visual Studio driver project. You can launch, configure, and control Static Driver Verifier from the Driver menu in Visual Studio.

Static Driver Verifier Documentation

- Static Driver Verifier Known Issues: Lists latest known issues for Static Driver Verifier
- Using Static Driver Verifier to Find Defects in Drivers: Tells you what you need to get started analyzing your driver code in the Visual Studio environment.
- Static Driver Verifier commands (MSBuild): Lists the MSBuild commands to use to run SDV in a Visual Studio Command Prompt window.
- Introducing Static Driver Verifier: Provides an overview of the static analysis tool.
- Using Static Driver Verifier: Provides the details about using and configuring the static analysis tool.
- Static Driver Verifier Report: Describes the viewer that displays the detailed trace of the static code analysis.
- Static Driver Verifier Rules: The rules define the requirements for proper interaction between a driver model and the kernel interface of the operating system.
- Static Driver Verifier Reference: Provides reference information about the function role types, SDV configuration files, error, and warning messages.
The DDI compliance rules for audio (PortClis) miniport drivers verify the DDI interface between PortClis.sys and its miniport drivers.

### In this section

<table>
<thead>
<tr>
<th>Topic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PcAddAdapterDevice</strong></td>
<td>The PcAddAdapterDevice rule specifies that a PortClis miniport driver correctly uses the PcAddAdapterDevice function, specifically that the DeviceExtensionSize should be either zero (0) or no less than PORT_CLASS DEVICE EXTENSION SIZE.</td>
</tr>
</tbody>
</table>
| **PcAllocateAndMapPages** | The PcAllocateAndMapPages rule specifies that a PortClis miniport driver calls the following interfaces, using the correct parameters:  
  - IPORTWaveRTStream:AllocatePagesForMdl  
  - IPORTWaveRTStream:AllocateContiguousPagesForMdl  
  - IPORTWaveRTStream:MapAllocatedPages |
| **PcAllocatedPages**  | The PcAllocatedPages rule specifies that a PortClis miniport driver frees previous allocated pages by calling AllocatePagesForMdl or AllocateContiguousPagesForMdl methods. |
| **PcIrqDDIs**         | The PcIrqDDIs rule specifies that a PortClis miniport driver must call PortClis DDIs at the correct IRQL level.                                |
| **PcIrqIport**        | The PcIrqIport rule specifies that a PortClis miniport driver must call PortClis IPort interfaces at the correct IRQL level.                 |
FindBugs

Null pointer deref

```java
// Eclipse 3.0,
// org.eclipse.jdt.internal.ui.compare,
// JavaStructureDiffViewer.java, line 131

Control c = getControl();
if (c == null && c.isDisposed())
    return;
```

Unconditional wait

```java
// JBoss 4.0.0RC1
// org.jboss.deployment.scanner
// AbstractDeploymentScanner.java, line 185

// If we are not enabled, then wait
if (!enabled) {
    try {
        synchronized (lock) {
            lock.wait();
        }
    }
    ...
```

Some initial Findbugs bug patterns

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CN</td>
<td>Cloneable Not Implemented Correctly</td>
</tr>
<tr>
<td>DC</td>
<td>Double Checked Locking</td>
</tr>
<tr>
<td>DE</td>
<td>Dropped Exception</td>
</tr>
<tr>
<td>EC</td>
<td>Suspicious Equals Comparison</td>
</tr>
<tr>
<td>Eq</td>
<td>Bad Covariant Definition of Equals</td>
</tr>
<tr>
<td>HE</td>
<td>Equal Objects Must Have Equal Hashcodes</td>
</tr>
<tr>
<td>IS2</td>
<td>Inconsistent Synchronization</td>
</tr>
<tr>
<td>MS</td>
<td>Static Field Modifiable By Untrusted Code</td>
</tr>
<tr>
<td>NP</td>
<td>Null Pointer Dereference</td>
</tr>
<tr>
<td>NS</td>
<td>Non-Short-Circuit Boolean Operator</td>
</tr>
<tr>
<td>OS</td>
<td>Open Stream</td>
</tr>
<tr>
<td>RCN</td>
<td>Redundant Comparison to Null</td>
</tr>
<tr>
<td>RR</td>
<td>Read Return Should Be Checked</td>
</tr>
<tr>
<td>RV</td>
<td>Return Value Should Be Checked</td>
</tr>
<tr>
<td>Se</td>
<td>Non-serializable Serializable Class</td>
</tr>
<tr>
<td>UR</td>
<td>Uninitializable Read In Constructor</td>
</tr>
<tr>
<td>UW</td>
<td>Unconditional Wait</td>
</tr>
<tr>
<td>Wa</td>
<td>Wait Not In Loop</td>
</tr>
</tbody>
</table>
Current list of Findbugs bug patterns

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BC</td>
<td>Equals method should not assume anything about the type of its argument</td>
</tr>
<tr>
<td>BIT</td>
<td>Check for sign of bitwise operation</td>
</tr>
<tr>
<td>CN</td>
<td>Class implements Cloneable but does not define or use clone method</td>
</tr>
<tr>
<td>CN</td>
<td>clone method does not call super.clone()</td>
</tr>
<tr>
<td>CN</td>
<td>Class defines clone() but doesn't implement Cloneable</td>
</tr>
<tr>
<td>CNT</td>
<td>Rough value of known constant found</td>
</tr>
<tr>
<td>Co</td>
<td>Abstract class defines covariant compareTo() method</td>
</tr>
<tr>
<td>Co</td>
<td>compareTo()/compare() incorrectly handles float or double value</td>
</tr>
<tr>
<td>Co</td>
<td>compareTo()/compare() returns Integer.MIN_VALUE</td>
</tr>
<tr>
<td>Co</td>
<td>Covariant compareTo() method defined</td>
</tr>
<tr>
<td>DE</td>
<td>Method might drop exception</td>
</tr>
<tr>
<td>DE</td>
<td>Method might ignore exception</td>
</tr>
<tr>
<td>DMI</td>
<td>Adding elements of an entry set may fail due to reuse of Entry objects</td>
</tr>
<tr>
<td>DMI</td>
<td>Random object created and used only once</td>
</tr>
</tbody>
</table>

http://findbugs.sourceforge.net/bugDescriptions.html
Some challenges in preventing defects

• How do you know what is incorrect behavior?

• How do you explain to a developer the cause of the (potential) defect?

• What happens if the tool approximates program behavior and comes to an incorrect conclusion?
Use of defect prevention tools in OSS projects (Dec 2014)

<table>
<thead>
<tr>
<th>Source</th>
<th>Projects</th>
<th>Use 1 ASAT</th>
<th>Use &gt; 1 ASATs</th>
</tr>
</thead>
<tbody>
<tr>
<td>GitHub</td>
<td>83</td>
<td>34%</td>
<td>30%</td>
</tr>
<tr>
<td>OpenHub</td>
<td>9</td>
<td>67%</td>
<td>22%</td>
</tr>
<tr>
<td>SourceForge</td>
<td>10</td>
<td>30%</td>
<td>0%</td>
</tr>
<tr>
<td>Gitorious</td>
<td>20</td>
<td>30%</td>
<td>5%</td>
</tr>
<tr>
<td>Total</td>
<td>122</td>
<td>36%</td>
<td>23%</td>
</tr>
</tbody>
</table>

TABLE III
DESCRIPTION OF THE ASATs FOR RQ 2 AND 3.

<table>
<thead>
<tr>
<th>Tool</th>
<th>Language</th>
<th>Format</th>
<th>Extendable</th>
<th>Released</th>
<th># of Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHECKSTYLE</td>
<td>Java</td>
<td>XML</td>
<td>Yes</td>
<td>2001</td>
<td>179</td>
</tr>
<tr>
<td>FINDBUGS</td>
<td>Java</td>
<td>Text</td>
<td>Yes</td>
<td>2003</td>
<td>160</td>
</tr>
<tr>
<td>PMD</td>
<td>Java</td>
<td>XML</td>
<td>Yes</td>
<td>2002</td>
<td>330</td>
</tr>
<tr>
<td>ESLINT</td>
<td>JavaScript</td>
<td>JSON</td>
<td>Yes</td>
<td>2013</td>
<td>157</td>
</tr>
<tr>
<td>JSCS</td>
<td>JavaScript</td>
<td>JSON</td>
<td>Yes</td>
<td>2013</td>
<td>116</td>
</tr>
<tr>
<td>JSHINT</td>
<td>JavaScript</td>
<td>JSON</td>
<td>No</td>
<td>2011</td>
<td>253</td>
</tr>
<tr>
<td>JSL</td>
<td>JavaScript</td>
<td>Text</td>
<td>No</td>
<td>2005</td>
<td>63</td>
</tr>
<tr>
<td>PYLINT</td>
<td>Python</td>
<td>Text</td>
<td>Yes</td>
<td>2006</td>
<td>390</td>
</tr>
<tr>
<td>RUBOCOP</td>
<td>Ruby</td>
<td>YAML</td>
<td>Yes</td>
<td>2012</td>
<td>221</td>
</tr>
</tbody>
</table>

Why developers don't use defect prevention tools

• Not integrated. The tool is not integrated into the developer’s workflow or takes too long to run
• Not actionable. The warnings are not actionable;
• Not trustworthy. Users do not trust the results due to, say, false positives
• Not manifest in practice. The reported bug is theoretically possible, but the problem does not actually manifest in practice
• Too expensive to fix. Fixing the detected bug is too expensive or risky
• Warnings not understood. Users do not understand the warnings.
Challenges with customizability

- Many tools have many false positives
- Want to have the ability to turn on and off useful and not useful rules
- Teams may customize settings, but then results in issues when different teams use different settings and find different issues with shared code
Working with developer intent

• How do you know what behavior is incorrect? (i.e., the oracle problem)
  • Have developers write specifications for a program for properties they care about
  • Build rules about how an API should be used, check that clients use it correctly
  • Look at lots of code, find atypical behaviors
Writing specifications

Model classes should have ‘private’ fields and getters.

```java
//CompilationUnit
//PackageDeclaration/Name[@Image="com.bankapplication.model"]
//ClassOrInterfaceDeclaration[
  count(
    ClassOrInterfaceBody/
    ClassOrInterfaceBodyDeclaration/FieldDeclaration[@Private="true"]=0
  )
  or
  count(ClassOrInterfaceBody/
    ClassOrInterfaceBodyDeclaration/MethodDeclaration/MethodDeclarator[starts-with(@Image,"get")]=0)
]
```

Natural language spec and corresponding implementation in PMD

- Specifying constraints on code often requires learning and using a new language defined by tool
- Often done by dedicated tool expertise with expertise in writing necessary specs
- May capture company-wide policies
How should potential defects be communicated to developers?

• Static analysis tools increasingly part of the build process
• Builds compile code, run static analysis tools
• Individual teams may build their own static analysis rules

• How should these tools communicate analysis results to developers?
Tricorder

• Goals:
  • Low false positives—error reports should result in code changes
  • Empower users to contribute—let developers write their own checkers
  • Make data-driven usability improvements
  • Effective workflow integration
  • Quick fixes

Communicating errors to developers

• Study at Google based on **26.6 million** builds
• Developers frequently see error messages
  • ~30% of builds fail due to compiler error
• Median resolution time is ~12 minutes

• Dependency errors are the most common

<table>
<thead>
<tr>
<th>Count</th>
<th>Error</th>
<th>Fix</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Misspelled identifier</td>
<td>Fix spelling</td>
</tr>
<tr>
<td>5</td>
<td>Wrong number of args to constructor call</td>
<td>Add or remove arguments</td>
</tr>
<tr>
<td>4</td>
<td>Missing import</td>
<td>Add import</td>
</tr>
<tr>
<td>2</td>
<td>Missing dependency</td>
<td>Add dependency to BUILD file</td>
</tr>
<tr>
<td>2</td>
<td>Incorrect type parameter in arg to method</td>
<td>Fix type parameter</td>
</tr>
<tr>
<td>1</td>
<td>Called a non-existent method</td>
<td>Removed method call</td>
</tr>
<tr>
<td>1</td>
<td>Accessed a non-existent field</td>
<td>Added field</td>
</tr>
<tr>
<td>1</td>
<td>Removed a class but didn’t remove all uses</td>
<td>Removed remaining uses of class</td>
</tr>
</tbody>
</table>

Communicating error messages

```java
void m() {
    final int x;
    while (true) {
        x = read();
    }
}
```

F.java:5: error: variable x might be assigned in loop
```
x = read();
^ 1 error
```

VS.

F.java:5: error: The blank final variable "x" cannot be assigned within the body of a loop that may execute more than once.
```
x = read();
^
Communicating errors

F.java:5: error: The blank final variable "x" cannot be assigned within the body of a loop that may execute more than once.
    x = read();
    ^

Claim: there is a problem

Grounds: why is this a problem

The claim is the concluding assertion or judgment about a problem in the code.

Resolutions suggest concrete actions to the source code to remediate the problem.

Facts, rules, and evidence to support the claim.

Bridging statements that connect the grounds to the claim. Provides justification for using the grounds to support the claim.

Examples

OpenJDK  cannot find symbol
symbol: variable varnam
location: class Foo

Jikes  No field named "varnam" was found in type "Foo". However, there is an accessible field "varname" whose name closely matches the name "varnam".

• OpenJDK only presents a claim. Jikes presents a ground (there is an accessible field "varname"), which is qualified through a rebuttal (However).

How do developers themselves explain errors on StackOverflow?

(a) Claim-only (CEM = 191, SO = 0)

(b) Claim-resolution (CEM = 10, SO = 59)

(c) Simple argument layout (CEM = 8, SO = 49)

(d) Extended argument layout (CEM = 1, SO = 102)

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Simple Argument Components</strong></td>
<td></td>
</tr>
<tr>
<td>CLAIM (Section 5.3.1)</td>
<td>The claim is the concluding assertion or judgment about a problem in the code.</td>
</tr>
<tr>
<td>RESOLUTION (Section 5.3.2)</td>
<td>Resolutions suggest concrete actions to the source code to remediate the problem.</td>
</tr>
<tr>
<td>GROUNDS (Section 5.3.3)</td>
<td>Facts, rules, and evidence to support the claim.</td>
</tr>
<tr>
<td>WARRANT (Section 5.3.4)</td>
<td>Bridging statements that connect the grounds to the claim. Provides justification for using the grounds to support the claim.</td>
</tr>
<tr>
<td><strong>Extended Argument Components</strong></td>
<td></td>
</tr>
<tr>
<td>BACKING (Section 5.3.5)</td>
<td>Additional evidence to support the warrant, if the warrant is not accepted.</td>
</tr>
<tr>
<td>QUALIFIER (Section 5.3.6)</td>
<td>This is the degree of belief for a claim, often used to weaken a claim.</td>
</tr>
<tr>
<td>REBUTTAL (Section 5.3.7)</td>
<td>Exceptions to the claim or other components of the argument.</td>
</tr>
</tbody>
</table>