SATE V Ockham Sound Analysis Criteria

Paul E. Black

paul.black@nist.gov

http://samate.nist.gov/

The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.

How can I measure sound analysis?

… the tool is never wrong …

How can I measure sound analysis?

SATE V Ockham Sound Analysis Criteria:

1. **The tool is claimed to be sound.**
2. **… the tool produces findings for a minimum of 60% of buggy sites OR of non-buggy sites.**
3. **Even one incorrect finding disqualifies a tool for this SATE.**

http://samate.nist.gov/SATE5OckhamCriteria.html

Definitions

- **A *site* is a location in code where a weakness might occur.**
- **A *buggy site* is one that has an instance of the weakness. A *non-buggy site* does not.**

```
char data[100] = "";
size_t dataLen = strlen(data);
FILE * pFile = fopen(FILENAME, "r");
if (pFile != NULL) {
    if (fgets(data+dataLen, (int)(100-dataLen), pFile) == NULL) {
        printLine("fgets() failed");
        /* Restore NUL terminator if fgets() failed */
        data[dataLen] = '\0';
    }
    fclose(pFile);
}
/* No format allowing a possible format string vulnerability */
printf(data);
```

Number of sites

	CWE-121 Stack-based Buffer Overflow[1]	CWE-476 NULL Pointer Dereference	CWE-190 Integer Overflow[2]	CWE-369 Divide by Zero	CWE-457 Use of Uninitialized Variable
U all sites	86 612	77 945	124 081	3018	339 407
N notices	18 598	303	1 356	1399	769
F = U - N	70 107	77 642	122 725	1619	338 638
B buggy	3472	303	3 306	684	200

1. CWE-121: |U| ≠ |N| + |F| because some notices are not sites.
2. CWE-190: |B| > |N| because our sites included “short” numbers.

More Definitions

- **A *notice* is a tool report about a site.**
  - **A notice may be conservative, so we allow for …**
- **A *finding* is a judgment on a site.**
- ***Sound* means every finding is correct.**
  - **A tool need not produce a finding for every site; that is *completeness*.**

SATE V Ockham Criteria – Frama-C

- **Frama-C reports sites with bugs, *but the analysis is conservative. Some notices are wrong, that is, the sites are not buggy.***
- **If Frama-C reports nothing, the site is sure to be ok (not buggy).**
- **So in this case, a finding (of a good site) is a site with no notice for it.**

- **CEA ran Frama-C on C files in Juliet 1.2**

Procedure for Each Weakness

1. Decide what constitutes a site
2. Determine the sites **U=the set of all sites**
3. Determine the notices **N=the set of notices**
4. Check that $N \subseteq U$
   - If that is not true, reconcile definition of site and notice
5. Determine buggy sites **B=the set of buggy sites**
6. Determine the findings **$F = U - N$**
7. Check that $|F| \geq 0.6\times(|U| - |B|)$
   - If that is true, Criteria 2 is satisfied
8. Check that $F \cap B = \varnothing$
   - If that is true, Criteria 3 is satisfied

Results So Far for Frama-C

- CWE121 Stack-based Buffer Overflow
- CWE122 Heap-based Buffer Overflow
- CWE123 Write-what-where Condition
- CWE124 Buffer Underwrite ('Buffer Underflow')
- CWE126 Buffer Over-read and CWE127 Buffer Under-read
- ✓ CWE476 NULL Pointer Dereference
- ✓ CWE190 Integer Overflow or Wraparound
- CWE191 Integer Underflow (Wrap or Wraparound)
- ✓ CWE369 Divide by Zero
- ✓ CWE457 Use of Uninitialized Variable
- ✓ CWE562 Return of Stack Variable Address

Problems With SATE V Ockham

- **definition of CWE**
  - **uninitialized *variable***
  - **return of stack *variable* address. Also, what if returned but never used?**
- **definition of sites**
  - **Return of Stack Variable Address?** `return 1;`
- **align tool's notices with CWEs**
- **what is a "site" for path weaknesses, e.g., failure to filter input - SQL injection**

Next Steps

- Finish checking criteria for all weaknesses
- Crosscheck ‘buggy’ list with other SATE results
- Integrate with automated synthetic test case checking to develop master bug list for Juliet 1.2
- Finish final report by July