Debugging

Mohammad Mousavi

Department of Informatics, King's College London

Software Testing and Measurement

(Automated) Debugging: A Sorting Program

```
int main(int argc, char * argv[])
{
int *a;
int i;
a = (int *) malloc( (argc - 1) * sizeof(int) );
for (i = 0; i < argc - 1; i ++)
    a[i] = atoi(argv[i + 1]);
shell_sort(a, argc);
printf("Output: ");
for (i = 0; i < argc - 1; i++)
    printf("%d ", a[i]);
free(a);
return 0;
}
```

```
void shell_sort(int a[], int size)
{ int i, j; int h = 1;
do {
     h = h * 3 + 1;
} while (h <= size);
do {
    h /= 3;
    for (i = h; i < size; i++)
    {
        int v = a[i];
        for (j = i; j >= h && a[j - h] > v; j -= h)
              a[j] = a[j - h];
        if (i != j)    a[j] = v;
    }
} while (h != 1);
}
```

(Automated) Debugging: A Sorting Program

Once upon a time, a tester found the following bug:

```
$ ./simple 5 4 3 2 1 666666
Output:  0 1 2 3 4 5
```

How do we find the fault?

Find and Focus

- Scientific method:
  1. assume,
  2. organize an experiment,
  3. if refuted, refine your assumption and repeat.

  possible formalization: invariants and assertions
- Observing: logging the value of infected variables

  e.g., `print` command in gdb
- Watching: keeping an eye on infected variables

  e.g., `break` and `watch` commands in gdb
- Slicing: find the slice responsible for infection

  see the lecture on slicing

Getting Our Hands Dirty...

We use gdb (any other debugger will do)

- Reproduce the test:
  `run 5 4 3 2 1 666666` Damn, the tester was right!
  (Not always that easy, try `55 4`.)
- Simplify the test-case
  `run 5 4 3 2`
- Find the possible the origins,
  focus on a problem area,
  e.g., `a[0]` and `shell_sort` (slicing...)
- Isolate the causes
  what makes `a[0]` wrong?
  compare it with the sane situation, what is different?
- Correct the problem

TRAFFIC

1. Track the problem
2. Reproduce the failure
3. Automate and simplify the test-case: minimal test-case $\Leftarrow$
4. Find possible origins: where it first went wrong
5. Focus on the most likely origins: what part of state is infected
6. Isolate the chain: what causes the state to be infected $\Leftarrow$
7. Correct the defect

Automated Debugging is about Perfection

Perfection

> *Perfection is achieved not when you have nothing more to add, but when there is nothing more left to take away.*
>
> Antoine de Saint-Exupéry

Automated Debugging is about Perfection

Perfection

Perfection is achieved not when you have nothing more to add, but when there is nothing more left to take away.

Antoine de Saint-Exupéry

Automated Debugging

Take out all that has nothing to do with the failure...

Debugging: An Example

- My slides for today (in LaTeX) did not compile
- some part of it did work before (older slides)

Debugging: An Example

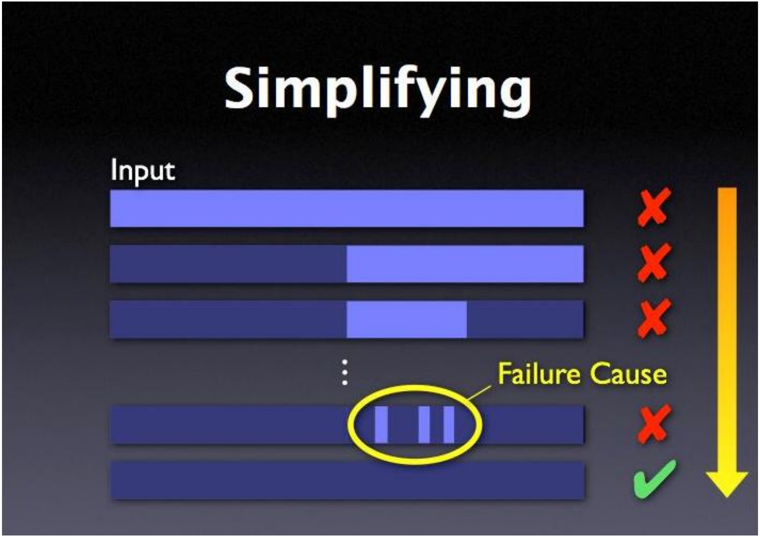
- My slides for today (in LaTeX) did not compile
- some part of it did work before (older slides)
- divide the new parts into two:
  1. remove first half part
  2. if the problem is there, repeat until one (new) slide is left
  3. if not, put back the second half and and remove the first, repeat
- apply the same technique to the content of the remaining slide

Debugging: An Example

- My slides for today (in LaTeX) did not compile
- some part of it did work before (older slides)
- divide the new parts into two:
    1. remove first half part
    2. if the problem is there, repeat until one (new) slide is left
    3. if not, put back the second half and and remove the first, repeat
- apply the same technique to the content of the remaining slide

This is called delta debugging:
our order of business for today.

Outline



(Ack. figures are due to Andreas Zeller.)

Minimizing Delta Debugging: Basic Idea

Try to find the minimal environment causing the failure by:

- Divide the circumstances C in n parts C_i,
- remove a part C_i such that $C \setminus C_i$ causes failure, repeat the algorithm with $C \setminus C_i$,
- if no such part exists, choose a bigger $n <\mid C \mid$ and repeat.

Minimizing Delta Debugging: Formalization

- Circumstances: C (input but could be: program, environment, etc.)
- Test: $\mathit{test} : 2^C \rightarrow \{\times, \checkmark, ?\}$
- Starting state: $C_x \subseteq C$, such that $\mathit{test}(C_\times) = \times$
- Goal: find a minimal subset $C'_\times \subseteq C_x$ such that $\mathit{test}(C'_\times) = \times$

Minimizing Delta Debugging: Algorithm

$ddmin(C_\times, 2)$, where

$ddmin(C'_\times, n) =$

$$\begin{cases} C'_\times, & \text{if } \mid C'_\times \mid = 1, \\ ddmin(C'_\times \setminus C_i, max(n-1, 2)) & \text{else if } \exists_{i \leq n} test(C'_\times \setminus C_i) = \times \\ ddmin(C'_\times, max(2n, \mid C'_\times \mid)) & \text{else if } n < \mid C'_\times \mid \\ C'_\times & \text{otherwise} \end{cases}$$

where C_i's are partitions of $C'_\times$ of (almost) equal size.

Application in Random Testing

Idea

- feed huge inputs to the system
  (guaranteed crash on huge input)
- simplify input
- present the simplified result as a test-case

Application in Random Testing

Examples

- applied to command UNIX tools
- FLEX (lexical analyzer): crashed on a test-case of 2121 characters
- NROFF (document formatter): crashed on a single control character
- CRTPLOT (plotter output): crashed on single characters ‘t’ or ‘f’

Improvements

- caching: save the test outcomes,
  use the saved data
- stop early: define a criterion to stop the algorithm, e.g.,
  1. no progress
  2. reaching a certain granularity
  3. upper bound on time
- use structures, e.g., blocks instead of characters
- differences vs. circumstances
  (compare sane with insane)

What is a Cause?

- Effect: the failure
- Cause: an event preceding effect,
  without which effect would not have happened

Isolating the cause

- Cause: the minimal difference between the worlds with and without the failure

Isolating the cause

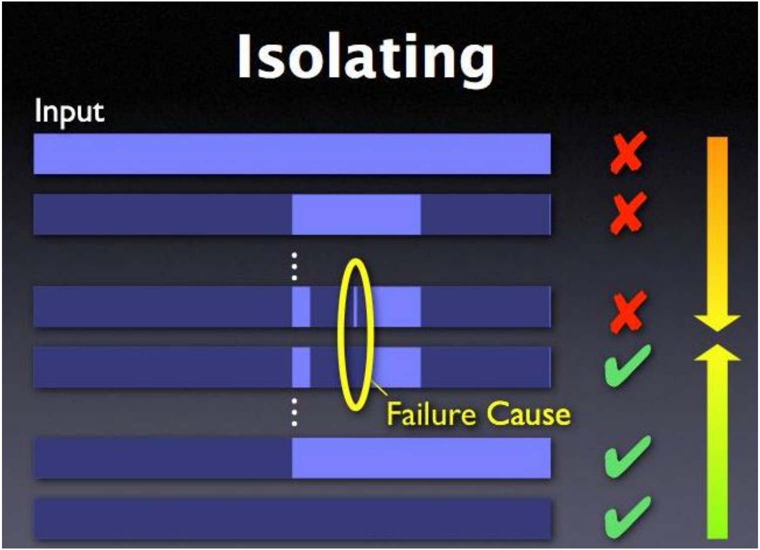
- Cause: the minimal difference between
  the worlds with and without the failure
- Challenge: the world without failure: the goal of debugging

Isolating the cause

- Cause: the minimal difference between the worlds with and without the failure
- Challenge: the world without failure: the goal of debugging
- Two solutions:
  1. manipulate the world by a debugger: turn infected to sane
  2. use another test-case in which no fault appears

Isolating: The Sorting Program Case

1. `./sample` produces a failure on `5 4 3 666666`
2. works fine on `5 4 3`
3. find combinations of
   1. states of 1 with 2 such that the program passes
   2. states of 2 with 1 such that the program fails
4. the difference between the two leads to a cause


Isolating
Input
Failure Cause

Delta Debugging: The Algorithm

Start from:

- $C_{\checkmark} = \emptyset$: passing circumstances and
- $C_{\times}$: failing circumstances

1. compute the difference Δ between the failing and the passing circ., divide into n parts: Δ_i,

Delta Debugging: The Algorithm

Start from:

- $C_{\checkmark} = \emptyset$: passing circumstances and
- $C_{\times}$: failing circumstances

1. compute the difference Δ between the failing and the passing circ., divide into n parts: Δ_i,
2. remove Δ_i from the failing circ.; it is the new passing circ., if it passes

Delta Debugging: The Algorithm

Start from:

- $C_{\checkmark} = \emptyset$: passing circumstances and
- $C_{\times}$: failing circumstances

1. compute the difference Δ between the failing and the passing circ., divide into n parts: Δ_i,
2. remove Δ_i from the failing circ.; it is the new passing circ., if it passes
3. add Δ_i to the passing circ.; it is the new failing circ., if it fails
4. add Δ_i to the passing circ.; it is the new passing circ., if it passes
5. remove Δ_i from the failing circ.; it is the new failing circ., if it fails

Delta Debugging: The Algorithm

Start from:

- $C_{\checkmark} = \emptyset$: passing circumstances and
- $C_{\times}$: failing circumstances

1. compute the difference Δ between the failing and the passing circ., divide into n parts: Δ_i,
2. remove Δ_i from the failing circ.; it is the new passing circ., if it passes
3. add Δ_i to the passing circ.; it is the new failing circ., if it fails
4. add Δ_i to the passing circ.; it is the new passing circ., if it passes
5. remove Δ_i from the failing circ.; it is the new failing circ., if it fails
6. increase n if none of the above holds

Delta Debugging: The Algorithm

Start from:

- $C_{\checkmark} = \emptyset$: passing circumstances and
- $C_{\times}$: failing circumstances

1. compute the difference Δ between the failing and the passing circ., divide into n parts: Δ_i,
2. remove Δ_i from the failing circ.; it is the new passing circ., if it passes
3. add Δ_i to the passing circ.; it is the new failing circ., if it fails
4. add Δ_i to the passing circ.; it is the new passing circ., if it passes
5. remove Δ_i from the failing circ.; it is the new failing circ., if it fails
6. increase n if none of the above holds
7. repeat until the difference is a singleton

Delta Debugging: Algorithm

$dd(C_{\checkmark}, C_{\times}, 2)$,
where $ddmin(C'_{\checkmark}, C'_{\times}, n)$ is defined recursively as:

$$\begin{cases} (C'_{\checkmark}, C'_{\times}) & \text{if } \mid \Delta \mid = 1, \\ dd(C'_{\times} \setminus \Delta_i, C'_{\times}, 2) & \text{else if } \exists_{i \leq n} test(C'_{\times} \setminus \Delta_i) = \checkmark \\ dd(C'_{\checkmark}, C'_{\checkmark} \cup \Delta_i, 2) & \text{else if } \exists_{i \leq n} test(C'_{\checkmark} \cup \Delta_i) = \times \\ dd(C'_{\checkmark} \cup \Delta_i, C'_{\times}, max(n-1, 2)) & \text{else if } \exists_{i \leq n} test(C'_{\checkmark} \cup \Delta_i) = \checkmark \\ dd(C'_{\checkmark}, C'_{\times} \setminus \Delta_i, max(n-1, 2)) & \text{else if } \exists_{i \leq n} test(C'_{\times} \setminus \Delta_i) = \times \\ dd(C'_{\checkmark}, C'_{\times}, min(2n, \mid \Delta \mid)) & \text{else if } n < \mid \Delta \mid \\ (C'_{\checkmark}, C'_{\times}) & \text{otherwise} \end{cases}$$

where $\Delta = C'_{\times} \setminus C'_{\checkmark}$ and Δ_i's are n partitions of Δ of (almost) equal size.

Delta Debugging: Applied to Test-Case Simplification

Start from:

- $C_{\checkmark} = \emptyset$: the empty test-case
- $C_{\times}$: the test-case leading to failure
- Much more efficient than minimizing delta debugging

Delta Debugging: Applied to Regression Testing

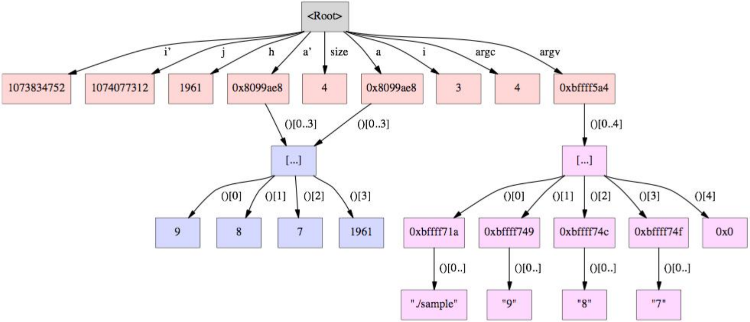
Start from:

- Goal: find out what went wrong in the new development (the old version worked well)
- $C_{\checkmark} = \emptyset$: basis is the old program, no changes needed
- $C_{\times}$: difference between the old and the new i.e., changes needed to obtain the new program from the old one

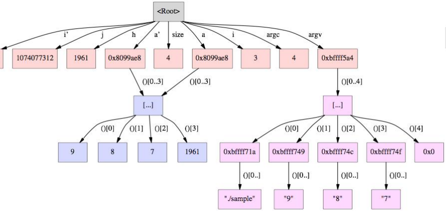
Isolating the Cause: Idea

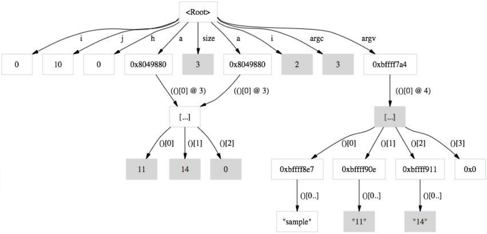
- Capture the state of the program
- Compare the states of a passes and a failed run
- The smallest difference Δ is the variable causing the problem
- Find out what influences this variable

Program State: Memory Graphs



Comparing the Differences





Implementable as debugger commands,
e.g., `set variable size = 2`.

Isolating the Cause: Implementation

- Compute the common subgraph of the passing and failing memory graphs. Let the difference be $C_\times$.
- Implement $C_\times$ as debugger commands.
- Apply delta-debugging to $C_\checkmark = \emptyset$ and $C_\times$
  1. Apply differences to the memory graphs and test.
  2. At each step of *dd* if the changed state is not a valid state (program does not run), return ?, if it is a valid state, return the result of the test,
- The result Δ leads to a cause.

Isolating the Cause: Sorting Case

Run the algorithm before calling `shell_sort` with the state of `./sample 7 8 9` as passing and `./sample 11 14` as failing.
If 0 at the state: test fails $\times$, passes $\checkmark$ otherwise.

1. $C_\times = \{$ `a[]`, `i`, `size`, `argc`, `argv[]` $\}$, $C_\checkmark = \emptyset$.
2. new failing state: `a[]`, `argv[1]` $\times$
3. new passing state: `argv[1]` $\checkmark$
4. new passing state: `a[0]` $\checkmark$
5. new passing state: `a[0]` and `a[1]` $\checkmark$
6. $\Delta = \{$ `a[2]` $\}$

Isolating the Cause: Illustrated Case

■ = δ is applied, □ = δ is *not* applied

#	$a'[0]$	$a[0]$	$a'[1]$	$a[1]$	$a'[2]$	$a[2]$	$argc$	$argv[1]$	$argv[2]$	$argv[3]$	i	$size$	Output	Test
1	□	□	□	□	□	□	□	□	□	□	□	□	7 8 9	✔
2	■	■	■	■	■	■	■	■	■	■	■	■	0 11	✘
3	■	■	■	■	■	■	□	□	□	□	□	□	0 11 14	✘
4	■	■	■	□	□	□	□	□	□	□	□	□	7 11 14	?
5	□	□	□	■	■	■	□	□	□	□	□	□	0 9 14	✘
6	□	□	□	■	□	□	□	□	□	□	□	□	7 9 14	?
7	□	□	□	□	■	■	□	□	□	□	□	□	0 8 9	✘
8	□	□	□	□	■	□	□	□	□	□	□	□	0 8 9	✘
Result					■									

Isolating the Chain of Causes

- Apply delta-debugging at the start, determine the minimal passing and running state
- Choose a common point (e.g., a function call) in the middle
- Apply delta-debugging on the states of the minimal passing and failing run
- Repeat the algorithm with the rest of the program and the new passing and failing states

Finding the Culprits

- The previous algorithm gives different Δ's (causes at different points)
- Track the change of causes
- A smelling point: a ceases to be a cause and b becomes a cause

Automated Debugging

- A natural mechanization of simple debugging principles
- Provides (partial) solutions to
  1. testing,
  2. simplifying the test-cases,
  3. isolating the causes and
  4. isolating the cause-effect chain.