

CSE 504 Assignment 1

Nan-Chen Chen

1. Debug runtime error related to memory

For languages like C/C++, where we can directly access and manipulate memory, it is usually hard to debug runtime errors. For example, there may be one line in a code that accesses an over-range array index, but it may not lead to problem when no other program declares that block of memory. When we move a program that is developed on a machine with plenty of memory to one with smaller memory size, such an error may occur and it may not be trivial for figuring out this is the bug.

Underlying cause

- (1) Memory operations are not intuitive and requires special attention to use it correctly.
- (2) Such errors are not always easy to reproduce and it may be challenging to surface the problems. Also, memory dump is not always easy to interpret. Therefore, tracing back to what leads to the problem requires great efforts.

What data/tools to help

- (1) It may be helpful if we can extract the usage of memory in a program and examine if there are potential memory traps.
- (2) If we can collect runtime errors and see if there are patterns for different types of errors, we may have a tool to infer the potential type of runtime errors and guide us to finding errors.

2. Fast development vs. Better maintenance

Sometimes, I write a program for one-time use purpose. In order to finish it shortly, I may not consider the design of the program structures carefully. However, sometimes it turns out that others want to use the program and request adding other features, which may not be easy under current design. On the other hand, sometimes I develop a program with very careful design, but it turns out that lots of the functionalities become not necessary, and it take too long to finish development.

Underlying cause

The goal of developing such software is uncertain and it changes after development starts.

What data/tools to help

- (1) It can be helpful if the requirements are predefined clearly. We can also user studies to understand what parts may be extended in the future and what parts are less important.
- (2) If we can have a tool to take current codes and refactor them to simpler/more sophisticated design, it can be helpful.

3. Too many scripts to remember parameters and correct workflow

When analyzing data, I usually write a set of scripts (in perl or python), and each script take different parameters. If I do not use the set of scripts for a while, it often takes some time to recall

Summary of Comments on chen.pdf

Page: 1

1 Number: 1 Author: mernst Subject: Highlight Date: 1/11/2016 8:12:11 PM

There exist many tools that warn about such problems, make them more likely (in order to easier reproduce them), or make them less likely. A popular open-source tool is Valgrind. What is wrong with these tools, or how do you want to improve on them?

1 Number: 2 Author: mernst Subject: Highlight Date: 1/11/2016 8:08:43 PM

Why is this relevant? Is it only because the memory manager may put data in different locations?

Also, different programs generally get their own heaps, and they are isolated by virtual memory, so there is no possibility for an out of balance array accessing one program to fall in memory declared by another program. Are you talking about threads? It doesn't seem like it, but that's the closest I can come to understanding your example.

1 Number: 3 Author: mernst Subject: Highlight Date: 1/11/2016 8:13:16 PM

Do you have ideas about reproduction? If this is not the problem you want to address, why bring it up?

I'm not sure how the different parts of point two here are related to one another.

1 Number: 4 Author: mernst Subject: Highlight Date: 1/11/2016 8:09:32 PM

What would be the representation of this?

This possible solution feels very vague to me, and would be helpful to be specific about how this would help a programmer. What that is represented, and what concrete action with the programmer do as a result of having access to this data?

1 Number: 5 Author: mernst Subject: Highlight Date: 1/11/2016 8:16:21 PM

To think this possible to do? That is, is the program the same after you write it as you had in mind before you started? I suspect that off that is not the case: the act of writing the program causes you to better understand the problem itself.

For all of these problems, I think would be helpful if you had in mind a specific anecdote: a concrete example that is happen to you and you remember well. Then, you can devise a solution that would've helped you then. I feel like your descriptions are overly vague; because they are lacking details, it is hard to understand what the real problem was and you are unlikely to come up with a solution that will actually be useful.

Doing this will help make your ideas much more concrete, and once you have something useful you can generalize it afterward.

what parameters to give and how to execute them in order.

Underlying cause

Sometime this is due to the scripts are not documented well, but other times it may due to that the scripts are overly complicated.

What data/tools to help

It can be convenient if there is a program that can record the usage of scripts and show **the steps** in an interface to help recall the steps or to help programmers figure out how to simplify the workflow

4. **Unexpected changes in dependency**

For languages like python, where lots of packages are available for direct use, sometimes the software will be broken after a dependent package is updated and the updates are not backward-compatible. In some cases, updates are done unintentionally (e.g., we move the codes to another machine and reinstall the development environment, and we may not have versions of packages specified before doing that). Since python programs usually have lots of dependencies, it may **not be easy to notice** **version difference** leads to problems.

Underlying cause

Things we do not have control get changed and it may take some time to diagnose the errors we get are due to those changes in dependent packages.

What data/tools to help

Keep track of dependencies and send alerts when **something we currently use** in the program **is changed** in the latest update.

5. **Team development: Hard to be consistent**

When developing software in a team, since members may have different programming background and habits, splitting tasks but at the same time remain consistency requires lots of communication. For example, everyone may have different naming conventions. Even though it is possible to combine components without consistent variable names/functions name, it may make the software less understandable for future development or sharing. Another example can be different way of using a framework: There are various ways of using Angular.js, and one may like to use curly braces for template (`{{ variable }}`) but one may like to use `ng-bind` (`<div ng-bind='variable'></div>`).

Underlying cause

Everyone has different programing experiences, habits, and styles. The resources of learning a library may also impact how a person uses it.

What data/tools to help

(1) If we can find inconsistency between codes/programmers, we can use it in code review and to improve consistency.

(2) When declaring a variable/function, we can have a tool **suggesting its name**. The suggestions come from existing codes and other people's common styles.

1 Number: 1 Author: mernst Subject: Highlight Date: 1/11/2016 8:19:23 PM
What steps?

Is there something about the initial implementation of the scripts that made it difficult to document them? Is it that the scripts were easy to remember at the time but you've lost that memory?

If the scripts have to be run in a particular order, what is the reason that you didn't create a script that itself calls all the other scripts? The good reasons not to do this, such as the need to change the order of execution or to examine intermediate results; other tools that can help with that?

1 Number: 2 Author: mernst Subject: Highlight Date: 1/11/2016 8:20:26 PM
I presume you don't have tests that you run when you install?

1 Number: 3 Author: mernst Subject: Highlight Date: 1/11/2016 8:21:58 PM
Many programming languages have a way to require a particular version (or a range of versions) for a dependency. Does Python?

Can you think of tools that would be helpful for the maintainers of the projects as well is for you? That is, what is the reason that the maintainers released a new incompatible version?

1 Number: 4 Author: mernst Subject: Highlight Date: 1/11/2016 8:22:34 PM
How would you define this?

1 Number: 5 Author: mernst Subject: Highlight Date: 1/11/2016 8:22:11 PM
How would you define this?

1 Number: 6 Author: mernst Subject: Highlight Date: 1/11/2016 8:25:07 PM
There are lint like tools that will warn about deviations from a coding style. I don't know what exists for Python and JavaScript, but one example of a tool for Java is checkstyle, and surely this exists for their languages as well.

1 Number: 7 Author: mernst Subject: Highlight Date: 1/11/2016 8:26:30 PM
What are its inputs? Is the routine already written? What is the programmer enter a name and then the tool suggests a better name?

This seems like a much more interesting problem than how templates are used.