Review 13

3/6/12

CS6V81.502

Seungtack Baek

**Scalable and Systematic Detection of Buggy Inconsistencies**

**in Source Code**

In this Paper, Gabel et al. developed a system called “DejaVu” that detects inconsistencies in the code clones. As opposed to the other similar studies, the author claims that this system focuses on detecting maximal clones (yet without any duplicate reports) and scalable design. The design of DejaVu is as following.

1. Enumerate similar code groups (code clones) using abstract syntax tree (AST), which can increase the maximality of the bugs found.
2. Analyze each clone then determine each clone to be buggy/benign
3. Generate bug report from the analysis in the previous step.

They used this system to a large (75+ MLOC) commercial code base, which resulted in 8104 bug reports, of those 2,070-2760 (why is this in tens?) are likely to be bugs.

I like how they started the motivation of this study with real world example. From figure1, we can clearly see that inconsistencies in the code clones have potential to cause unforeseen bugs. Thus we know what they have done is something useful, given they made sufficient contribution to the topic.

Yet, the author suggests possible performance upgrade through using more “cores,” without considering coordination overhead. If you add more cores, it might cause some overhead on the master node to adapt to the newly added cores (which can depend on the protocol used and the hardware architecture in integration with operating system). Also, they chose “retroactive detection” approach, as opposed to proactive solution, without fully explaining why they did so.

**Question:**

1. First, just a grammar question: the first bullet point on page 2, it says “changes need not necessarily be contiguous.” Aren’t “need” and “necessary” the same thing? So in a way it is repeating itself (like “for awhile”, as opposed to “for a while”)?
2. I know when the researchers use technical term, they usually italicizes the word. Yet, how about the word “completeness” from last sentence from section “Large-scale Inconsistent Clone Analysis” on page 2? Does complete means general completeness? Or has some other connotations?
3. Could you explain what you mean by “incremental clone detection,” which is from last sentence of the section “Clone Detection ‘Front End’” on page 4?
4. Who came up with the name “DeajVu?”