File 20100128.0736: Notes from meeting with Dr Martin this morning:

I started off by noting that the 19th February talk is confirmed; I need a title and abstract to Ralph soon.

I complained that the thesis statement at the beginning of the outline has been simplified to such an extent that it looks trivial. Dr Martin came back to this later. He said thagt if it's too stark, it can seem to be trivial. Want something more general, general enough that it can be applied by other people to other problems in future, with predictive value.

Dr Martin concurred that a lot of data points might be needed to show statistical significance of such a stark statement. Make it less stark for the audience to key on.

We discussed the various metrics available to measure both level of effort and security improvement. The problem of how to measure and compare levels of security is a hard one, not yet solved, and it would be a contribution in itself (note: ANOTHER POSSIBLE CONTRIBUTION) to find one. I asked whether any of these metrics (e.g., calendar time, person-hours, total number of individuals involved, budget spent, number of test procedures, code coverage analysis) alone would be sufficient, or whether people like seeing a synthetic measure better. Dr Martin noted that they may be highly correlated anyway, and I can easily look at them in a spreadsheet, compare them, add them and subtract them, combine them and graph them in ways that produce a strong synthetic measure if I need to later.

For measuring security improvement, I like the second measure: fewer findings during subsequent testing. The nature of these products (cross domain solutions) is such that they are commonly installed in diverse security domains, each controlled by accreditors who do not necessarily trust one another, and hence the products tend to be tested and retested again using similar criteria but by different groups, e.g., FAT, IV&V, CT&E, ST&E (at multiple sites in the same security domain), followed by new rounds of CT&E and ST&E in other security domains and by other certifiers. Nevertheless, the criteria are often the same or similar, since NSA, NIST, JIS and GCHQ are the ultimate authorities on the subject of computer security. Different military standards tend to trace back to the same recommendations. This, I think is the best measure of security improvement that I can hope for, not terribly applicable outside the peculiar environment I am researching in, but laser focussed on this particular problem domain, highly applicable, and most importantly—available.

(Dr Martin cautioned that getting statistics on number of findings may be hard, but I do have access to them, and I think I can argue that aggregate counts of Cat I, Cat II, Cat III, Cat IV findings, not identified by product or version or application or site, would be unclassified—note: have to get those numbers through the trusted downgrade process—and could therefore be published if suitably sanitised. I know sanitisation, at least.)

Another issue brought up by Dr Martin is the problem of measurement perturbing the thing being measured. I argued back that I think it won't be a problem because of the deliberate and careful disconnect between stages of testing: *eg* FAT, IV&V, CT&E, ST&E. I think I can make a convincing case that Heisenberg uncertainty is not a problem here. I need to make it convincing, though, to be persuasive.

I would like to justify every step in my methodology as being both necessary and sufficient. The goal of this dissertation, something important to keep in mind, is to make a contribution that has predictive value, that others can use in future. If it has no predictive value then it was a waste of time.

Dr Martin came back to my concerns about the apparent triviality of the (simplified) thesis statement at the beginning of the outline. I think I can address this by writing prose to expand upon it.

Note that the triviality of the hypothesis at the beginning is due to the fact that it's my strawman, not my thesis. My thesis is more detailed and general. Make that clear. I am trying to anticipate questions that people will ask on the 19th.

I asked whether this level of introspection, of considering different alternatives, how much of this disucssion belongs in the methodology chapter. Dr Martin mentioned that if one of the main contributions of the dissertation is the methodology, then showing how it was developed and why does belong in the methodology chapter. I was intrigued by this comment, and I think I may run with it some more, because I am short of contributions and my methodology is getting bigger and bigger. I like that idea. I will explore it further in the next few day, think about it.

Plan for next few weeks: first, finish the methodology chapter and get it put to bed. Then prepare talk for the 19th. Next, do that first survey. It is slightly unfortunate that I didn't identify the need for it until this late date, but it needs to be done first, it is simple, and I can do it quickly. Plan to have that done in the next month. Next task after that is to update the schedule and get it to Dr Martin. The first case study (CC) is largely written already; the second case study is in progress, hopefully it will be successful, but it's ongoing. I have not done much planning or detail on the tool development at all yet.

We ran out of time so I didn't bring up the *nihil obstat* paper that I was thinking last night of dusting off and publishing (after I get it security and IP cleared, of course). I may try to talk about that on Monday, or I may wait until I have it in a form to show to Dr Martin directly.

This was a good meeting. I had progress to show, and we actually ran out of time because I talked about so much. We only had 0700 to 0730 scheduled; I deliberately did not run over time out of consideration for Dr Martin's schedule.

Work on the *nihil obstat* paper. Convert it to LaTeX.

Next meeting: Monday at 1715 GMT; Dr Martin is teaching next week so he will call me. It might be as late as 1730. Have Skype showing available.

References