Test Driven Development and the Scientific Method

Rick Mugridge

Department of Computer Science,

University of Auckland, New Zealand

Daniel Preisler Larsen

Abstract

The article makes comparisons between T.D.D. and the Scientific Method (S.M.) used in most sciences today. The purpose is to better the understanding of both (in our case mostly the understanding of T.D.D.)

S.M. main elements Theories evolve in large or small steps and must be consistent with experiment results. Experiments examining the predictions of the theory must be repeatable so that the results can be verified and the theory refined.

T.D.D. The design of a system corresponds to the theories of S.M, and the tests that are written are the experiments.

Tests, like experiments, are repeatable and can be run (and rerun) to test new versions of the design.

Many things in the world are refined this way : the XP process, Aeroplanes, Martial arts etc.

T.D.D. can be used at two levels in XP:

When the overall design of the system is made - Customer stories (E.G. use cases) where tests are made for each story

At the micro-iterations-level when completing sub stories.

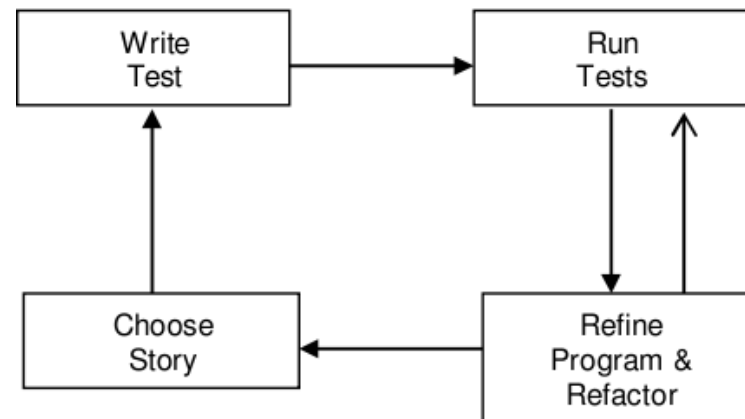In both cases the process can be illustrated :

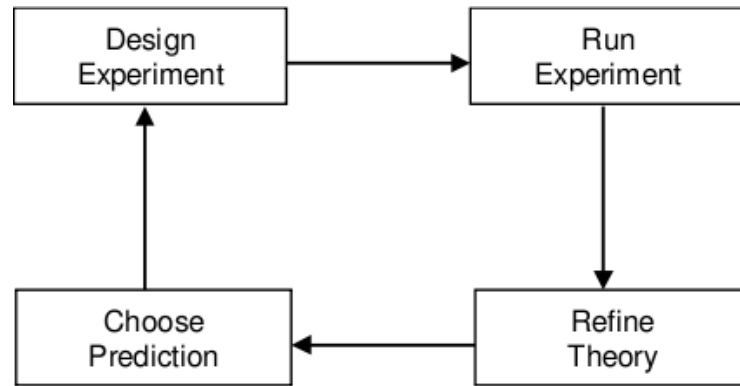Figure 1. Model of TDD

The same applies with the S.M

Figure 2. Model of the scientific method

Formulating a theory / design

The theories of the natural sciences are often stated with mathematical terms and laws of physics. This is often done in order to be precise and to build on the works of others.

In T.D.D. the design is specified with code, pseudo code, diagrams and drawings. Predictions can be tested by executing the tests.

Creating a theory / design

The scientist may not know the precise nature of the problem at the beginning − only that the existing theory (if any) doesn't fit well enough. The theory might therefore be based on vague ideas and intuitions in the beginning.

Analog to this, the customer might not know what he wants.

The first tests might be difficult to choose, but the process gets easier quickly.

Selecting experiments / tests

S.M.: The theory must be sufficiently well defined before experiments can be made.

Results of experiments that conforms with the theory usually gives little information. As a consequence the 'edges' and 'extremes' of the theory are tested.

Experiments usually have costs associated. Cost/benefit ratios should be considered when choosing experiments. This can be a hard − impacts of the experiments are not yet known.

The same applies to T.D.D.

Performing tests / experiments

In science, control experiments are made to heighten the possibility of a correct result. Uncertainties and lack of precision have to be accounted for.

In T.D.D. when a test fails, small steps are used to zero in on the error. The test is run again to make sure it fails, and to learn more. Uncertainties can also arise with concurrency. Statistical results can come from profiling tests.

Reproducibility

In science, experiments must be reproducible. Costs can be very high!

In T.D.D. testes are usually easy and cheap to rerun but can be expensive to develop.

Experiments and tests as examples

Experiments can be used in educational contexts, and can illustrate an aspect of nature concretely.

Tests can give a picture of the progress of the development. A history of the development process.

Simplicity

In science, the simplest of two or more theories, describing the same thing, is probably the most correct one. Occam's Razor

In T.D.D. the simplest design (e.g. fewest components) should be chosen, all things being equal. Choose the simplest thing that will work.

Evo- / Revo - lution

Theories are refined over time. Ideas are discarded and new added – they evolve. But sometimes they undergo revolutions.

Kuhn[1]: 'cumulative acquisition of novelty is not only rare in fact, but improbable in principle'

I recently read the opposite on slashdot – people disagree – even the smart ones.

In T.D.D. a design is made through evolutionary development. The design adapts to the needs of the customer (which might change over time). Changes have different impact (from very large to very small effects)

The Human Factor

Rosenthal[2]: '... trained, conscientious, intelligent experimenters unconsciously bias their tests, avoid running experiments that might cause trouble for their theories, misanalyze, misinterpret, and ignore test results that show their ideas are wrong.'

The same goes for testing in the conventional style. Writing test for finished code has many of the same problems.

In T.D.D. the tests are made FIRST thereby minimizing that risk.

Conclusion

T.D.D. and S.M has a lot in common. Theories and designs evolve over time and sometimes undergoes big changes. By using T.D.D. some of the problem with conventional development can be reduced.

Over Stene Til Stjerne.

References

[1] T. Kuhn, The Structure of scientific Revolutions, University of Chicago Press, 1962.

[2] R. Rosenthal. Experimental Effects in behavioral Research, Appleton-Century Crofts, 1966.