

The University of W.A.

## The Software Process - CITS 5502

Terry Woodings & Du Huynh    Second Semester, 2013

### Assignment Three – Modelling the Production of Code

This assignment is worth 20% of the final mark for this unit. The report is due on Monday the 7<sup>th</sup> October (office closes at 4.30pm) with a (ten pages or thereabouts) discussion covering the results and conclusions. Extra marks are available for early submission.

The work will be based upon an analysis of several sets of data on the effort and size of software projects. In previous years, each student wrote a new copy of a program every three days (four copies in all). Three sets of programs were produced: Problem 1 in language A followed by a similar type of Problem 2 in language A and then Problem 1 again but this time in language B. This is similar to work done in 2013 when you were asked to write a program to calculate the line which best discriminates between two clusters of points (in two dimensions). Ideally, each student should do this at three-day intervals (without reference to previous code listings). Provide a table of times (to the nearest minute) and program lengths (in LOC).

Describe the issues being studied and discuss the measurement process and the data. There are a number of questions requiring your attention in the report.

1. To what extent do we carry our learning from previous projects over to the next project of a similar type but different specification? This is a question of the relevance of the problem domain in predicting future effort.
2. To what extent do we carry our learning from previous projects over to the next project which requires different resources? This is a question concerning the resources domain.
3. To what extent does practice enable us to produce a more predictable estimate of effort?
4. Is the learning pattern similar between the three sets of data?
5. What appears to be the limiting factor (the minimum time) on the production of the code in each of the three sets of tasks?
6. How similar (consistent) are the solutions as we learn the details of the problem?

Consider the models for reduction in effort over four ( $t = 0, 1, 2, 3$ ) attempts ( $b > 0, a > c > 0$ ):

$$\text{Effort} = (a + bct) / (bt + 1)$$

$$\text{Effort} = (a-c)(t+1)^{-b} + c$$

$$\text{Effort} = (a-c) e^{-bt} + c$$

$$\text{Effort} = a + bt + ct^2$$

and discuss their suitability for modelling the learning process. One of these models is quite inappropriate. Which and why? Choose two models and use them to fit curves to the averages of each of the three sets of data from previous years (these sets will be discussed at the lab on 19<sup>th</sup> Sep). Which fits the best? Use it to fit a curve to your personal data for the discrimination problem. Comment on how well the model suits your own data and personal process.

[Note: A comparison of the **a** parameters (the initial effort) should assist with questions (1) and (2). A comparison of the **b** parameters (the amount of curvature or rate of improvement) will provide an answer to question (4) and similarly the **c** parameters (the amount of practiced effort) should assist with question (5). The coefficient of variation of the LOC answers question (6).]