Laboratory Notebooks for Computer Science

Craig A. Struble, Ph.D.
Department of Mathematics, Statistics, and Computer Science
Marquette University

August 21, 2003

Abstract

Although *science* is included in the discipline named *computer science*, there appears to be little precedent for maintaining laboratory notebooks as part of a computer science education. Certainly, computer science combines the philosophies of mathematics, science, and engineering; and all three of these philosophies are seen throughout the field. In our experience, when learning to understand the behavior of software, an algorithm, or hardware, using the scientific method seems to prevail.

We describe the motivation and use of laboratory notebooks in the computer science classes that we teach. Included in this document are the required elements and good practices for maintaining a laboratory notebook.

1 Introduction

A scientific approach is one of the best ways to understand complex systems and behavior. While much of computer science is based around mathematics, a sound mathematical theory does not always provide the intuitive understanding of how an algorithm or approach behaves when provided with a variety of input. A scientific approach, generating hypotheses and carrying out experiments to test those hypotheses, can help to bridge the gaps between computer science theory and computer science practice.

Computer science students often approach learning with an undisciplined "trial and error" approach. This manifests itself when writing programs: "Let me add this one line of code and see if it works." It also manifests itself with understanding system behavior: "I'll change these five parameters at once and see if it works." These kinds of experiments prevent learning the true behavior, because the student does not formulate hypotheses for behavior, nor does the student focus on how a single change impacts the overall system behavior.

Taking a more scientific approach to computer science education, using controlled experiments and encouraging hypothesis generation, will improve a student's understanding. To accomplish this, the student needs to be familiar with the scientific method and carefully maintain a lab notebook, which records student observations. These two topics are the focus of the remainder of this paper.

2 The Scientific Method

The scientific method is generally considered to be the best approach for developing theories about complex systems, either man-made or natural. Typically the scientific method involves the following

steps:

1. Observe some aspect of the universe.
2. Formulate a hypothesis that is consistent with what you have observed and with your experience.
3. Make *testable* [1] predictions based on your hypothesis.
4. Test those predictions by experiments or further observations.
5. Modify your hypothesis in the light of your results.
6. Go to step 3.

The specifics of each step are beyond the scope of this paper, but a Google search for "scientific method" yields several pages going into detail about the scientific method and a few pages that refute the effectiveness of the scientific method. Whether or not the method itself is correct, it does provide a useful framework for learning.

In the context of computer science, building theories is often done through mathematical proof, since the computer and software is man-made and its basic workings are precisely understood. Although these theories are precise, their analysis is typically limited to best and worst case behavior, because analyzing the typical (or average case) behavior is impractical with existing mathematical techniques.

The scientific method plays its most important roles in understanding the typical behavior of a system. Furthermore it bridges the gap between theory and practice by carefully investigating the relationships between input, parameters, and environment. Through repeated application of the scientific method, a student learns to predict the behavior of a parameter change, writing a line of code, or the output generated from a given input. It is this intuitive understanding of a system's behavior that makes building new and more complex software an easier task.

One limiting factor for all humans is memory. It is impossible to remember every detail of every observation, every experiment, and every result. It is possible to record these observations as they occur, and then to recall those observations when formulating theories and hypotheses to predict behavior. This is the motivation for maintaining a laboratory notebook.

3 The Lab Notebook

In the context of a computer science class, the lab notebook can be viewed as a history of the work accomplished during the semester. Each student records the work they perform for lab assignments, carefully recording what they did and learned along the way.

For classes, the lab notebook is organized around lab assignments. For each lab assignment, a student records the following information: an introduction to the lab assignment; a diary of the steps performed and observations made during the assignment; an analysis of the observations; and a summary of the assignment.

The following subsections provide more detailed information about the content of the lab notebook.

[1] Predictions should be precisely stated and capable of being evaluated given available technology and knowledge.

3.1 Introduction

The introduction for each lab assignment contains a brief discussion of the important points of the assignment. The lab assignment number and title should be at the top of the sheet where the introduction is written. Any theory, formulas, definitions, designs, and hypotheses related to the assignment should be included in the introduction.

Pre-laboratory questions asked in the assignment should be answered in the introduction. The questions themselves should not be repeated, but it should be clear from your writing which question you are answering.

A copy of the lab assignment should be placed just after the introduction.

3.2 Diary

The foundation of the lab notebook is maintaining a diary. The goal of the diary is to provide enough information for you or someone else to reproduce your work in the future. Details are important, so write down everything you did, including the machine you used, operating system you used, the software version you used, parameters you set and selected, etc.

Diary entries should be recorded in chronological order. Each entry should be dated in pen and written on the right hand side of the notebook. The facing left hand side should contain computer plots, analysis, or other comments related to your diary entry.

3.3 Data

Data is typically output generated from the computer system, but may also include summary information, the length of time it took to execute a program, or the fact that the program crashed.

Data should be recorded neatly in tables. Each table should be labeled and contain a caption (describing the data recorded) for later reference in your diary, analysis, or summary. Each column should have a header and include the units for the data in the corresponding column.

As you record data, carefully consider whether the data is consistent. Note any outliers, or other abnormalities. If a mistake is made recording data, do not erase the data value. Strike through it with a single line and record the correct value.

3.4 Figures

Figures in a lab notebook include plots or graphs (visual representations of data), screenshots, sketches (e.g., for designs, to note trends), etc.

Each figure should be labeled and given a caption for later reference in diary entries or analysis.

Plots or graphs should contain labels for each axis, with appropriate units included. If the plot or graph contains lines (e.g. such as drawing lines through data points), a legend (a label for each line) should be provided.

3.5 Source Code

Because we assume students are keeping a lab notebook where programming will take place, it is important to include the source code for any programs written and used to answer questions during the course of the experiment. The source code should be well structured and documented, carefully

explaining how the program works. Recording source code is important, as it is a necessary for others to reproduce your work.

3.6 Analysis

After the raw data is collected and the experiment or lab assignment completed, an analysis of the results should be written. In an analysis section, several references to texts, papers, etc. should be included. Remember, everything you say is assumed to be false unless supported by other references.

Discuss the meaning of the data collected. Compare the results to relevant theoretical background. Note points of agreement and disagreement with the theory. Note any additional trends obtained from plots, and generate new hypotheses.

The analysis section also includes your answers to post-laboratory questions posed in the assignment. Clearly identify your answers as you write. Generally, the best way to accomplish this is with an introductory sentence to a paragraph answering the question. Using bullets or a number list is not good practice, unless in your answer you are highlighting important points of your solution.

3.7 Summary

Each lab writeup should include a brief (up to 1 page) summary of the point under study, what was observed, and what was concluded during the laboratory assignment.

3.8 Maintaining a Lab Notebook

To ease the tedium of maintaining a laboratory notebook, we recommend creating a template or boilerplate that contains an outline of the essential laboratory notebook sections. It is acceptable to refer to earlier experiments; for example, saying that the experimental conditions, such as computer platform and software, are the same as those in *Experiment 1: Using Telepathy to Move Mountains*. Abbreviations can be used, as long as they are clearly defined elsewhere in your notebook.

Remember, a main purpose of a laboratory notebook is to allow someone else to reproduce your work. All that may be available of your work is what you write. Most people are incapable of reading minds should you not be around to explain what you wrote.

4 Evaluation

The purpose of maintaining a lab notebook is to provide enough information for others to reproduce your results. Your notebooks will be evaluated on this principle. If the instructors feel they could reproduce your results given your notebook, then you will receive a better grade.

Several other points are considered when your notebook is evaluated:

- Is the notebook neatly maintained? (i.e., Can someone else follow what you did?)
- Did the student answer pre-lab and post-lab questions posed?
- Are the tables and figures properly labeled?
- Are correct and well founded conclusions drawn from the results?

- Are the introduction, analysis, and summary sections in well written English?
- Were proper procedures followed during the course of the experiment?

Grades for lab notebooks will be assigned based on the instructor's overall impression of the lab. Feedback on portions of the notebook may be made, but individual point deductions will not be assigned.

5 Acknowledgements

Norie Dela Cruz, Ph.D. graciously read and edited earlier versions of this paper. Several ideas came from the document titled "Keeping A Lab Notebook" maintained by the Harvey Mudd Department of Physics. Other ideas were taken from several web pages about laboratory notebook procedures in chemistry, biology, and physics departments worldwide. Thanks to Google for helping me to track down those resources.