CSC490 Project Proposal by Tong Zou, Feb 7, 2011

Assessing the influence of PEY on graduating UofT CS students' knowledge and application

Purpose / Abstract: One of the main forms of CS education is not just rooted in classroom theory, but practical application of concepts, which manifests itself in the form of doing internships. The purpose of this project is to study if and how the PEY program affects the knowledge and technical abilities of graduating CS students at the University of Toronto. Is there a significant difference in the methodology and knowledge between students who have gone on an internship versus those who haven't? Does PEY adequately prepare students for a full time job after graduation? These are the questions this study seeks to answer. At first, one might say that the influence of an internship will always result in more practical knowledge of CS concepts, but this isn't necessarily always the case. There could exist students who have not taken PEY, yet has a more solid understanding of the fundamental areas of CS; if that was indeed the case for students, then the PEY program will have a poor showing in its relation to CS education. If the outcome does support the consensus that a 12 or 16 month internship will *significantly* affect the knowledge and skills of a graduating student, and better prepare them for obtaining a job, then the PEY program will have merited some measure of success. Moreover, this study allows us to determine if graduating students who have taken PEY, have more expertise on areas that are *not* immediately relevant to their internship. This does bring up a causality issue: how do we know that knowing more is a result of the internship and not a result of the student inherently being bright? That is not the primary focus of this study, and for our purposes, we will assume that being bright, and going on an internship, are independent of each other.

Method: Ideally, I want to be able to do a mock interview with students. That would give me the richest set of data to work with. So I am going to ask students for a mock interview first, but knowing that some students will be busy, I also have a questionnaire. The questionnaire is not intrusive, and easy to distribute. However, I will have to rely on the student's honesty to do these questionnaires in good faith, and not look up answers on the internet; because the questionnaire is anonymous and there is no marking/penalty, there is very little incentive to 'cheat'. I will obtain a fair sample size from asking various students, split them into two groups of PEY + graduating, and no PEY + graduating, and I will have two sets of data from both the questionnaire and the mock interview if I have the chance to do those. They will contain questions that test the student's knowledge of fundamental areas of Computer Science, and are taken from actual interview questions by Google, Amazon, Microsoft, etc. I will also ask them to elaborate on some questions to assess their way of thinking and get a better sense of context. Then, I will compare their solutions and answers with the expected solutions and answers (I will be taking these questions from a coding book), and graph the results and record their implications. If the sample size happens to be small (which is likely), then this study will very much resemble a set of case studies. If I get to do an interview with some students, and if that data is very different, then I will take that into account.

Conclusions: The conclusions from the set of data, both quantitative and qualitative, will give us a better understanding of how PEY affects a student's application of computer science concepts, if at all. This can lead use to make several more observations on how useful the PEY program is, and whether or not the value of the internship experience is supplementary or complementary to the classroom experience. Without a large sample size, it will still give us a limited measure of 'quantifying' an internship experience, and what implications it may hold for future computer science students who may be considering the significance of doing an internship vs. the standard classroom CS education.

Deliverables: The written report will include an analysis and implications of the data that I have gathered from the distributed questionnaires and the presentation will focus on the distribution and results of the data. The questionnaire is available immediately for feedback. The interview questions will look similar to those on the questionnaire. Here is a rough timeline between now and Apr 5.

Feb 7 – Feb 28: Distributing questionnaire / mock interviews and gathering data.

Feb 28 – Mar 14: Assessing and analyzing data into information and conclusions about the study.

Mar 14 – Mar 31 – Finalizing report + presentation.

Mar 31 – presentation due

Apr 5 – written report due