

Doing a PhD in Computer Science in Four Years (after Bachelors)

Pankaj Jalote, Professor, Computer Science, IIT Kanpur

PhD is a unique degree in that it is not focused on acquiring more knowledge, but is oriented around doing research. As research requires many abilities as well as creation of new knowledge, a PhD program is necessarily a long one.

Though many students who join the PhD program are not unduly worried about the time the program takes, there is also a section of students for whom the possibility of a long program is clearly a deterrent. So, even though the careers for PhDs in Computer Science have increased and continue to increase¹, in many cases, the perceived duration is discouraging these students from undertaking the PhD program.

In this note we will discuss how a PhD program can be comfortably finished in four years, after the B.Tech./BE, if the student is motivated and does not loosen up during the program. It is hoped that this note will help students who are inclined towards PhD take a more informed decision, and not just give up on PhD due to its large duration. (This is not just a hypothetical approach – I have successfully applied it on one candidate. Though one candidate is not enough data, as it was applied consciously and right from the start, it does provide a limited validation of its feasibility.)

There are essentially three phases in any PhD program. (1) Course work to gain breadth, (2) Identifying the area and problem to work on, and (3) doing the research work and writing results. All these three phases are roughly of equal duration – about 1.5 to 2 years each. They are often done linearly, leading to the program length of about 5 to 6 years.

The most difficulty part of a PhD program is the phase two – identifying the area and defining the problem to be worked on in the thesis (which includes some a sense of how to approach the problem). It is often this phase in which a lot of time is spent – there is no tight course schedule to help, no deadlines, and no pressure of time as not too much time has passed. With some planning and initiative, however, it is possible to shorten the impact of this phase by doing parts of it in parallel with other phases.

To put this strategy in action, a rough schedule for a 4-year PhD is suggested, specifying what should be achieved in each semester.

Semester 1: Focus is on courses. The candidate should meet with different faculty, find out their areas of interest, and what type of problems they are working on. If the candidate has an idea of the areas he may be interested in, this search can be more focused. **Sem goal:** Identify a small set of faculty members whom he may want to work with.

Semester 2: Do some advanced courses, and/or some “independent study” with the professor(s). Use these courses and extra study/independent study to engage with the professors to identify possible problem-areas for PhD. **Sem goal:** Identify the problem-area, and the advisor. Here, it will help a lot if the potential advisor has some pre-defined problem which the student likes.

Summer, Semester 3: Finish the comprehensive (or qualifier or whatever it is called). Do further independent study to focus more on the problem-area, and generate (maybe small) ideas which can be developed further. **Goal:** The comprehensive is finished, and active research has started.

¹ See a related note: <http://www.cse.iitk.ac.in/users/jalote/GenArticles/ITresearch-TOI-7-05.pdf>

Semester 4: Focus on developing the ideas and doing related reading. The developed ideas should be converted into research papers for conferences/journals. **Sem goal:** Have some paper(s) submitted to some conference(s).

Semester 5, 6: Refine and/or expand the problem. Or work on some related ideas that got generated while working/reading. I.e. if the problem you are working on is not expandable, you can work on some related problems. **Goal:** Write at least two papers in this year. (It is better to start writing papers early, as they may not get accepted in the first attempt, and might need revision. As it is best to have some papers to your credit when you graduate, it is imperative that by the end of third year, a few papers have been written and submitted.)

Semester 7: Further development of ideas. Papers accepted in conferences which can be expanded, should be expanded as journal papers. Any pending or new ideas should be developed and papers written and submitted.

Semester 8: This is the closing semester. With papers written earlier, the thesis will be largely these papers plus a few chapters. The thesis writing effort reduces, but it is still a lot. **Sem goal:** Tie the loose threads and close pending issues. Complete and submit the thesis.

The basic strategy in this approach is to start research work early, and to break the large problem of PhD into smaller problems of writing papers in closely related problems (or different aspects of some problem). In the process, it is ensuring that phase 2 of the PhD process is done in parallel with the other two phases of doing course work and doing research. Besides giving the satisfaction of visible progress in terms of publications, this approach has another advantage – with publications, the case becomes stronger and evident in the eyes of the thesis examiners.

In this schedule, the advisor has to tightly manage the program (assuming that completing the PhD soon is a shared objective of him and the student). This approach implicitly assumes that the advisor has a relatively well defined problem to offer to the candidate. With the problem defined, in the second year, the advisor has to really guide the student in doing the work – i.e. the work is to a large extent decided and directed by the advisor, but executed by the student. The paper writing exercise will also be the same – the advisor will guide the student in the art of paper writing, and through many rounds of the first paper, help the student develop the ability to properly write scientific papers. In the third year the role of the advisor matures to being a collaborator – actively engaged in developing and generating ideas together with the student and writing papers. In the final year, the role should actively become more passive so as to let the student take the lead. The advisor should become more of a bouncing board, giving suggestions and feedback.

This approach is primarily geared towards completing the PhD quickly and developing research doing capability, on which a research career can be built. There are, however, some disadvantages of this approach. Due to its focus on time, it discourages looking at hard problems, which are risky and may require much more time. Also, the approach of having papers in related problems can provide breadth, but it can discourage working at a larger problem for a longer periods where results may come only towards the end. And it can also reduce the development of the problem-finding ability of the PhD candidate, as it relies a lot on the advisor proposing the problem. Overall, this approach is geared towards someone who treats PhD as a starting stage of a research career where he is learning the skills of doing research, which he will further develop in the rest of his career. For those who want to seriously attempt making a mark on the state of knowledge and go for breakthroughs, the schedule may be mildly useful as some sort of guide.