MSR—CMU Center for Computational Thinking Project Proposal

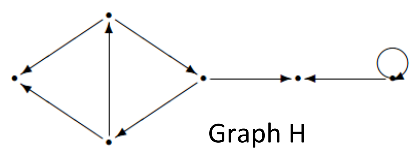
The Dichotomy Conjecture

Ryan O'Donnell

Project Background

Which algorithmic problems are solvable efficiently? One of the greatest exports of computer science to the rest of the sciences is the notion of NP-completeness versus polynomial-time (P) solvability. This dichotomy has been tremendously helpful in understanding the difficulty of tasks in areas as diverse as statistical mechanics, combinatorial chemistry, and molecular biology. Although it often taken as gospel in computer science that almost every natural problem (except Factoring and Graph-Isomorphism) is known to be in P or NP-hard, this is far from correct. Here are three examples of algorithmic tasks which are not known to be in P and are not known to be NP-hard:

1. Given is a graph. Find a cut in the graph with size at least 90% of the Max-Cut.
2. There are n Boolean variables. Given is a list of *constraints*, each operating on at most 3 variables. It is promised that there is an assignment to the variables satisfying all the constraints. Find an assignment that satisfies at least 2/3 of the constraints.
3. Given is a directed graph G. Decide if it is possible to place each of its vertices onto one of the 6 vertices of graph H (to the right) so that edges of G only go onto edges of H.



These problems are examples of *constraint satisfaction problems* (CSPs). CSPs model a vast number of theoretically and practically important algorithmic tasks. In general, a CSP is defined by two things: The *domain* of the variables, and the *types* of constraints allowed. For example, in Problem 2, the domain is {0, 1} and any constraints on 3 variables are allowed.

Given a CSP, there are several objectives one might consider. The simplest is called the *decision problem*: the objective is to decide if there is an assignment to the variables which satisfies all the constraints. Problem 3 above is such an example. Also important is the *optimization problem*: here the goal is to satisfy as many of the constraints as possible. Problems 1 and 2 above are examples. A third objective is the *counting problem*: the task here is to *count* the number of assignments which satisfy all the constraints. (The counting problem is of course always harder than the decision problem.)

One of the great open problems in computer science is the **Dichotomy Conjecture** of Feder and Vardi. It states quite simply that for every CSP, the associated decision problem is either in P or is NP-complete. Since its introduction in the '90s, The Dichotomy Conjecture has fueled an enormous body of research in complexity theory, programming languages, combinatorics, and especially algebra. The *counting version* of the Dichotomy Conjecture is also closely related to problems in statistical physics.

Another grand open problem in computer science is the **Unique Games Conjecture** made by Khot in 2002. The Unique Games Conjecture has led to major progress in the understanding the complexity of the optimization question for CSPs. Much of this progress came from the development of new mathematical tools in areas of probability, geometry, and harmonic analysis of Boolean functions.

Project Description. Researchers have been making slow but steady progress on the Dichotomy Conjecture over the years, and the problem, though difficult, is viewed as "within reach". The main approach has been very heavily algebraic; the current state of affairs is that certain algebraic conditions are known to make a CSP decision problem in P, but these conditions need to be weakened to prove the conjecture. Peculiarly, research on the complexity of optimization for CSPs has proceeded along a completely separate track. Assuming the Unique Games Conjecture, my work from 2004–2007 shows that one can use *analytic* tools (probability, harmonic analysis, etc.) to obtain Dichotomy-type results for basic CSP optimization. For example, it was known that approximating the Max-Cut to within 87.856% is in P, and I showed that anything better is NP-hard. Recently, Raghavendra has extended these tools and used them to show a fantastic result: essentially, a dichotomy theorem for CSP optimization (still assuming the Unique Games Conjecture). Unfortunately, the analytic tools cannot distinguish between CSPs where 100% of the constraints are satisfiable and where only 100%–ε are satisfiable. For optimization of CSPs, this is not important, but it makes all of the difference to the decision problem for CSPs.

Although the two tracks have operated completely separately, I and a few others have recently noticed that the objects being studied analytically by optimization researchers are *identical* to the objects being studied algebraically by Dichotomy Conjecture researchers. A very recent paper by Kun and Szegedy has made preliminary steps in identifying the correspondence. **My proposal is the following: with Yuan Zhou of CMU and Pinyan Lu of MSRA, we will unify the analytic tools and the algebraic tools used by the two groups to produce a solution to the Dichotomy Conjecture.**

I believe that Yuan, Pinyan, and I are the right people to attack this problem. I am an expert on the analytic tools used to study CSP optimization, and Yuan is on his way to becoming an expert too. (He has four strong papers in the area now.) The two of us also have some of the only work in the field on optimization hardness in the "100% satisfiable regime", which is precisely what is relevant for the Dichotomy Conjecture. Finally, as discussed briefly below, Pinyan is a world expert on the counting version of the Dichotomy Conjecture, and the tools used previously to attack this version are an interesting mix of the algebraic and the analytic.

Project Context. Our proposed research project will undoubtedly lead to advances in another branch of science: namely, mathematics. As mentioned earlier, much of the recent work on the optimization question for CSPs has led to new developments in pure mathematics. For example, my work on the Max-Cut optimization problem led to new Invariance Principles in probability theory, the resolution of important conjectures in the mathematics of social choice, and the solution to geometric problems concerning bubbles and foams. I will further foster interdisciplinary work between mathematicians and computer scientists when I give my invited lectures on CSP optimization at the 2011 Summer School on CSPs and the Dichotomy Conjecture, taking place June through August at the Fields Institute for Mathematics in Toronto. The summer school's series of workshops and lectures will mark the first significant meeting between the algebraists working on the Dichotomy Conjecture and the computer scientists and operations researchers with experience on the analytical side of CSP optimization.

Collaboration with MSR. Our main point of collaboration will be Pinyan Lu at Microsoft Research Asia and his interns. Pinyan is recognized as a world leader in the area of Dichotomy Conjecture research – especially counting dichotomies – with over a dozen papers on the topic. In addition, Pinyan is a close collaborator with Yuan. They have published research together on algorithmic game theory. Further, Yuan and Pinyan have already begun preliminary work on the Dichotomy Conjecture when Yuan was an intern at MSRA in the summer of 2010 with Pinyan as his mentor.

Budget. For years 1 and 2 of the project, we are requesting the following support:

Yuan Zhou academic year graduate student stipend:	$21,405.
Yuan Zhou graduate student tuition remission:	$37,188.
Overhead:	$11,848.
Ryan O'Donnell one month summer salary:	$10,595.
Funds for travel between CMU and MSRA:	$ 5000.
Budget Total per year:	$86,036.