

## SUMMARY: Analysis of Boolean Functions for Algorithms and Complexity

**Researcher:** Dr. Ryan O'Donnell

Associate Professor, Dept. of Computer Science, Carnegie Mellon Univ., Pittsburgh USA

<http://www.cs.cmu.edu/~odonnell>

**Project Coordinator and Host Institution:**

Dr. A. C. Cem Say

Professor, Dept. of Computer Engineering, Boğaziçi Univ., Istanbul Turkey

<http://www.cmpe.boun.edu.tr/~say>

The scientific focus of the project is in the field Theoretical Computer Science. This is the mathematical branch of Computer Science/Engineering research dealing with the theoretical capabilities and limitations of computation, not specifically tied to any particular technological model. Researcher Ryan O'Donnell's particular area of speciality involves bringing tools from probability and discrete analysis to the study of computation; Project Coordinator Cem Say's particular area of speciality is in theoretical models for computational complexity; especially, models based on quantum mechanics. Thus much of the research focus of the project has involved the study of quantum computational complexity and probabilistic algorithms. We discuss both the project research and the transfer of knowledge outcomes below.

### RESEARCH OBJECTIVES AND OUTCOMES:

**Quantum complexity research.** The first component of the quantum research was a collaboration between O'Donnell and Say concerning the computational complexity of *time travel*. This may sound fanciful. However: a) so-called "closed timelike curves (CTCs)" are apparently consistent with the laws of physics; b) the topic has previously been studied by renowned physicists (e.g., David Deutsch) and computer scientists (e.g., Scott Aaronson); c) research on the model provides insights into the relative powers of quantum and classical computation; d) even with the "time travel" motivation removed, the resulting questions concerning the computational complexity of Markov chains and quantum channels have impact for basic classical and quantum complexity theory. Building on previous work by the Say and his former Ph.D. student Yakaryilmaz, O'Donnell and Say were able to show that the algorithmic benefit conferred by any fixed number  $w$  of "time-traveling quantum bits" is no greater than that conferred by a single time-traveling classical bit.

A second component of the quantum research was joint between O'Donnell's and his Ph.D. student John Wright. It concerned the following fundamental problem in quantum theory: Suppose a physical experiment produces a particle in some unknown  $d$ -dimensional quantum state. The experimenter wishes to know some information about this state; perhaps complete information (this is called the quantum tomography problem) or perhaps only some partial information (say, just about the probabilities  $p_1, \dots, p_d$ ; this is called the *spectrum testing* problem). We assume the experimenter can prepare  $n$  copies of the particle and then measure the result. The fundamental question is: How large does  $n$  need to be, as a function of  $d$ , to reliably answer various questions? O'Donnell and Wright completed an extensive investigation into the spectrum testing problem, establishing several new sharp upper and lower bounds. It is to be hoped that the upper bounds provided have impact on the practical problem of quantum tomography.

**Algorithmic statistics and information theory.** The project pursued several objectives in this area; we highlight here only one. We consider arguably the most basic problems in statistics: Suppose we are given access to from an unknown probability distribution  $p_1, \dots, p_d$  on  $\{1, 2, \dots, d\}$ . How many samples  $n$  are required to reliably determine various properties of this distribution? (In fact, this is precisely the classical analogue of the quantum spectrum testing problem discussed above.) For many problems of this nature, the answer is — unfortunately — that  $n$  must be quite large compared to  $d$ .

For example, to estimate the *entropy* of the probability distribution, roughly  $n \approx d$  samples are required. However, recently algorithmic research has shown that this sample complexity can be dramatically lowered, to  $n \approx \log^2 d$ , if the experimenter is also allowed to make "distribution queries"; i.e., request the value of  $p_i$  for any  $i$ . These algorithms may have impact for many "Big Data" applications in computer science. During the project, O'Donnell, Say, and Boğaziçi students Cafer Caferov and Baris Kaya collaborated to show that these algorithms are best possible; i.e. that roughly  $n \approx \log^2 d$  samples/queries are always required.

Besides the aforementioned works, several other research projects — on topics in information theory, mathematical biology, algorithms for optimization tasks, among others — are ongoing or completed. O'Donnell has completed (or nearly completed) eight separate research works during the course of the project; three with the Project Coordinator and his students, two others with European visitors to Boğaziçi, and the other three with his Ph.D. students. Several have already been published (in the FSTTCS 2014, ITW 2015, and STOC 2015 conferences), and the rest are being submitted for publication.

### **TRANSFER OF KNOWLEDGE OBJECTIVES AND OUTCOMES :**

Another major component of the project is to achieve transfer of knowledge, both from the Project Coordinator and other Boğaziçi researchers to the Researcher, and from the Researcher to students at Bogaziçi in Turkey, and in Europe. In the first direction, O'Donnell has benefited from the quantum computation expertise of Dr. Say and from the mathematical and probabilistic expertise of Dr. Atilla Yilmaz from the Bogazici University Mathematics Department. We also summarize transfer of knowledge outcomes in the other direction.

**Teaching and student mentoring.** In the fall 2014 semester, O'Donnell taught *CmpE 587: Introduction to Research in Theoretical Computing*, variant of O'Donnell's 2013 Carnegie Mellon graduate course *A theorists' toolkit* (<http://www.cs.cmu.edu/~odonnell/toolkit13/>). In the spring 2014 semester, O'Donnell and Say organized a reading seminar for Dr. Say's group, to study quantum algorithms in general and the recent Ph.D. thesis of Aleksandr Belovs in particular. These two activities had impact on approximately 15 students, bringing them up to speed on the forefront of research in theoretical computer science. In addition, O'Donnell spent time mentoring and advising several Bogazici students; indeed, this led to a joint research paper with one master's student (Cafer Caferov) and one undergraduate (Baris Kaya).

**Outreach to undergraduates students in Turkey and in Europe.** Over the summer of 2014, O'Donnell gave two week-long, 15-hour "summer school" courses on the topic of Analysis of Boolean Functions — a mathematical tool heavily used in the theory of computation. These took place at Turkey's Nesin Mathematics Village <https://matematikkoyu.org/eng/> and in Stockholm at KTH's Swedish Summer School in Computer Science <http://s3cs.csc.kth.se/>. Approximately 50 undergraduates and graduate students from Turkey and Europe were impacted by these lectures.

**International Theoretical Computer Science seminar series.** O'Donnell organized a series of lectures at Boğaziçi University and at the Istanbul Center for Mathematical Sciences, with internationally-renowned European researchers: Sangxia Huang (KTH Stockholm); Dr. Guy Kindler (Hebrew University); Dr. Ilias Diakonikolas (Edinburgh University); Dr. Rahul Santhanam (Edinburgh University). In addition, O'Donnell facilitated an invited lecture by his Carnegie Mellon University colleague, the noted Human Computation researcher/entrepreneur Luis von Ahn. On Sept. 29, 2014, von Ahn inaugurated the 2014-2015 Bogazici Lecture Series by speaking to people from all departments of the university on the topic of *Massive Education for the Future*. <http://takvim.boun.edu.tr/bu/?q=node/195>

## A: USE AND DISSEMINATION OF FOREGROUND

The Researcher habitually publishes all of his research:

- freely, on his website, <http://www.cs.cmu.edu/~odonnell/>;
- in conferences devoted to theoretical computer science;
- in peer-reviewed academic journals (preferably ones with open access).

The research produced during this project is no exemption. As mentioned, eight research papers were completed (or nearly completed) during the project. We list the dissemination methods here:

1. *Improved NP-inapproximability for 2-variable linear equations*, joint with Johan Hastad (KTH Stockholm), Sangxia Huang (KTH Stockholm), Rajsekar Manokar (KTH Stockholm), John Wright (Carnegie Mellon).

This paper appears on the Researcher's web page and is under submission to the APPROX 2015 conference (Princeton). The journal to which this paper will be submitted is not decided, but the open-access *Theory of Computing* is a candidate.

2. *Conditioning and covariance on caterpillars*, joint with Sarah R. Allen (Carnegie Mellon).

This paper appears on the Researcher's web page, and will appear in ITW (Information Theory Workshop) 2015, Jerusalem. The journal to which it will be submitted is not yet decided.

3. *One time-traveling bit is as good as logarithmically many*, joint with A. C. Cem Say (Bogazici).

This paper (also available on the Researcher's web page) appeared at FSTTCS (Foundations of Software Technology and Theoretical Computer Science) 2014, New Delhi. The journal to which it will be submitted is not yet decided.

4. *Quantum spectrum testing*, joint with John Wright (Carnegie Mellon).

This paper appears on the Researcher's web page, and will appear in STOC (Symposium on the Theory of Computation) 2015, Portland USA. It is also under submission to Communications in Mathematical Physics.

5. *The weakness of CTC qubits and the power of approximate counting*, joint with A. C. Cem Say (Bogazici).

A version of this paper is completed, but since the research is still ongoing it is not yet available on the Researcher's web page. The conference and journal to which it will be submitted is not yet decided.

6. *"The most informative Gaussian function"*. joint with Guy Kindler and David Witmer.

7. *Optimal lower bounds for estimating entropy with PMF queries*, joint with Cafer Caferov, Baris Kaya, A.C. Cem Say (all of Bogazici).

This paper is essentially complete and will shortly appear on the Researcher's web page. The conference to which it will be submitted is not yet decided; candidates include COCOON (Computing and Combinatorics Conference) 2015 and RANDOM 2015. The journal to which it will be submitted is not yet decided.

8. *Refuting random CSP: stronger, sparser, and more general*, joint with Sarah R. Allen and David Witmer (both of Carnegie Mellon).

This paper is nearing completion; it will appear on the Researcher's web page later this semester. It will be submitted to FOCS (Foundations of Computer Science) 2015. The journal to which it will be submitted is not yet decided.

In addition to publishing research papers, O'Donnell has given a number of international talks during the reporting period. These have promoted the project work, as well as fostered collaboration with other European researchers. Complete slides for all of these talks are available on the Researcher's website. They include:

- *Testing surface area*, a talk given in the May 2014 workshop on Sublinear Algorithms at the Bertinoro International Center for Informatics, Bertinoro Italy.
- *Social choice, computational complexity, Gaussian geometry, and Boolean functions*, a talk on the main project topic, given in August 2014 at the International Congress of Mathematicians, Seoul.
- *One time-traveling bit is as good as logarithmically many*, a talk given in December 2014 on the aforementioned paper, at FSTTCS 2014 in New Delhi.

Besides this, the Researcher has disseminated his expertise in Analysis of Boolean Functions (the main project topic) in three ways: By giving a week-long lecture series to undergraduates at the Nesin Mathematics Village; by giving a week-long lecture series to graduate students at KTH's Swedish Summer School in Computer Science; and, by the free publishing in October 2014 of his textbook *Analysis of Boolean Functions* at <http://get.analysisofbooleanfunctions.org>.