Quantum Computing since Democritus

Written by noted quantum computing theorist Scott Aaronson, this book takes readers on a tour through some of the deepest ideas of math, computer science, and physics.

Full of insights, arguments, and philosophical perspectives, the book covers an amazing array of topics. Beginning in antiquity with Democritus, it progresses through logic and set theory, computability and complexity theory, quantum computing, cryptography, the information content of quantum states, and the interpretation of quantum mechanics. There are also extended discussions about time travel, Newcomb's Paradox, the Anthropic Principle, and the views of Roger Penrose. Aaronson's informal style makes this fascinating book accessible to readers with scientific backgrounds, as well as students and researchers working in physics, computer science, mathematics, and philosophy.

SCOTT AARONSON is an Associate Professor of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology. Considered one of the top quantum complexity theorists in the world, he is well known both for his research in quantum computing and computational complexity theory, and for his widely read blog *Shtetl-Optimized*. Professor Aaronson also created Complexity Zoo, an online encyclopedia of computational complexity theory, and has written popular articles for *Scientific American* and *The New York Times*. His research and popular writing have earned him numerous awards, including the United States Presidential Early Career Award for Scientists and Engineers and the Alan T. Waterman Award.

Quantum Computing since Democritus

SCOTT AARONSON

Massachusetts Institute of Technology

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9780521199568

First published 2013
3rd printing 2015

Printed in the United Kingdom by Clays, St Ives plc.

A catalog record for this publication is available from the British Library

Library of Congress Cataloging in Publication data
Aaronson, Scott.
Quantum computing since Democritus / Scott Aaronson.
pages cm
Includes bibliographical references and index.
ISBN 978-0-521-19956-8 (pbk.)
1. Quantum theory – Mathematics. 2. Quantum computers. I. Title.
QC174.17.M35A27 2013
621.39′1 – dc23 2012036798

ISBN 978-0-521-19956-8 Paperback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

To my parents

Contents

Preface *page* ix
Acknowledgments xxix

1. Atoms and the void 1

2. Sets 8

3. Gödel, Turing, and friends 18

4. Minds and machines 29

5. Paleocomplexity 44

6. P, NP, and friends 54

7. Randomness 71

8. Crypto 93

9. Quantum 109

10. Quantum computing 132

11. Penrose 150

12. Decoherence and hidden variables 160

13. Proofs 186

14. How big are quantum states? 200

15. Skepticism of quantum computing 217

16. Learning 228

17. Interactive proofs, circuit lower bounds, and more 243

18. Fun with the Anthropic Principle 266

19. Free will 290

20. Time travel 307

21. Cosmology and complexity 325

22. Ask me anything 343

Index 363

Preface

A CRITICAL REVIEW OF SCOTT AARONSON'S *QUANTUM COMPUTING SINCE DEMOCRITUS*
by Scott Aaronson.

Quantum Computing since Democritus is a candidate for the weirdest book ever to be published by Cambridge University Press. The strangeness starts with the title, which conspicuously fails to explain what this book is *about*. Is this another textbook on quantum computing – the fashionable field at the intersection of physics, math, and computer science that's been promising the world a new kind of computer for two decades, but has yet to build an actual device that can do anything more impressive than factor 21 into 3 × 7 (with high probability)? If so, then what does *this* book add to the dozens of others that have already mapped out the fundamentals of quantum computing theory? Is the book, instead, a quixotic attempt to connect quantum computing to ancient history? But what does Democritus, the Greek atomist philosopher, really have to do with the book's content, at least half of which would have been new to scientists of the 1970s, let alone of 300 BC?

Having now read the book, I confess that I've had my mind blown, my worldview reshaped, by the author's truly brilliant, original perspectives on everything from quantum computing (as promised in the title) to Gödel's and Turing's theorems to the **P** versus **NP** question to the interpretation of quantum mechanics to artificial intelligence to Newcomb's Paradox to the black-hole information loss problem. So, if anyone were perusing this book at a bookstore, or with Amazon's "Look Inside" feature, I would *certainly* tell that person to buy a copy immediately. I'd also add that the author is extremely handsome.

Yet it's hard to avoid the suspicion that *Quantum Computing since Democritus* is basically a "brain dump": a collection of thoughts about theoretical computer science, physics, math, and philosophy that were on the author's mind around the fall of 2006, when he gave a series of lectures at the University of Waterloo that eventually turned into this book. The material is tied together by the author's nerdy humor, his "Socratic" approach to every question, and his obsession with the theory of computation and how it relates to the physical world. But if there's some overarching "thesis" that I'm supposed to take away, I can't for the life of me articulate what it is.

More pointedly, one wonders who the *audience* for this book is supposed to be. On the one hand, it has *way* too much depth for a popular book. Like Roger Penrose's *The Road to Reality* – whose preface promises an accessible adventure even for readers who struggled with fractions in elementary school, but whose first few chapters then delve into holomorphic functions and fiber bundles – *Quantum Computing since Democritus* is not for math-phobes. A curious layperson could *certainly* learn a lot from this book, but he or she would have to be willing to skip over some dense passages, possibly to return to them later. So if you're someone who can stomach "science writing" only after it's been carefully cleansed of the science, look elsewhere.

On the other hand, the book is *also* too wide-ranging, breezy, and idiosyncratic to be used much as a textbook or reference work. Sure, it has theorems, proofs, and exercises, and it covers the basics of an astonishing number of fields: logic, set theory, computability, complexity, cryptography, quantum information, and computational learning theory, among others. It seems likely that students in any of those fields, from the undergraduate level on up, could gain valuable insights from this book, or could use it as an entertaining self-study or refresher course. Besides these basics, the book also has significant material on quantum complexity theory – for example, on the power of quantum proofs and advice – that (to this reviewer's knowledge) hasn't appeared anywhere else in book form. But still, the book flits from topic to topic too hastily to be a definitive text on anything.

So, is the book aimed at non-scientists who won't *actually* make it past the first chapter, but want something to put on their coffee table to impress party guests? The only other possibility I can think of is that there's an underserved audience for science books that are neither "popular" nor "professional": books that describe a piece of the intellectual landscape from one researcher's heavily biased vantage point, using the same sort of language you might hear in a hallway conversation with a colleague from a different field. Maybe, besides those colleagues, this hypothetical "underserved audience" would include precocious high-school students, or programmers and engineers who enjoyed their theoretical courses back in college and want to find out what's new. Maybe this is the same audience that frequents these "science blogs" I've heard about: online venues where anyone in the world can apparently watch real scientists, people at the forefront of human knowledge, engage in petty spats, name-calling, and every other juvenile behavior, and can even egg the scientists on to embarrass themselves further. (The book's author, it should be noted, writes a particularly crass and infamous such blog.) *If* such an audience actually exists, then perhaps the author knew exactly what he was doing in aiming at it. My sense, though, is that he was having too much fun to be guided by any such conscious plan.

NOW FOR THE ACTUAL PREFACE

While I appreciate the reviewer's kind words about my book (and even my appearance!) in the preceding pages, I also take issue, in the strongest possible terms, with his ignorant claim that *Quantum Computing since Democritus* has no overarching thesis. It *does* have a thesis – even though, strangely, I wasn't the one who figured out what it was. For identifying the central message of this book, I need to thank Love Communications, an advertising agency based in

Sydney, Australia, which put the message into the mouths of fashion models for the purpose of selling printers.

Let me explain – the story is worth it.

In 2006, I taught a course entitled "Quantum Computing since Democritus" at the University of Waterloo. Over the next year, I posted rough notes from the course on my blog, *Shtetl-Optimized*[1] – notes that were eventually to become this book. I was heartened by the enthusiastic response from readers of my blog; indeed, that response is what convinced me to publish this book in the first place. But there was one response neither I nor anyone else could have predicted.

On October 1, 2007, I received an email from one Warren Smith in Australia, who said he had seen a television commercial for Ricoh printers. The commercial, he went on, featured two female fashion models in a makeup room, having the following conversation:

> Model 1: But if quantum mechanics isn't physics in the usual sense – if it's not about matter, or energy, or waves – then what *is* it about?
>
> Model 2: Well, from my perspective, it's about information, probabilities, and observables, and how they relate to each other.
>
> Model 1: That's interesting!

The commercial then flashed the tagline "A more intelligent model," followed by a picture of a Ricoh printer.

Smith said he was curious where the unusual text had come from, so he googled it. Doing so brought him to Chapter 9 of my "Quantum Computing since Democritus" notes (p. 110), where he found the following passage:

> But if quantum mechanics isn't physics in the usual sense – if it's not about matter, or energy, or waves, or particles – then what *is* it

[1] www.scottaaronson.com/blog

> about? From my perspective, it's about information and probabilities and observables, and how they relate to each other.

So, it seemed, there was exactly one bit of dialogue in the commercial that I *didn't* write ("That's interesting!"). Smith found a link[2] where I could see the commercial for myself on YouTube, and his story checked out.

Far more amused than annoyed, I wrote a post for my blog, entitled "Australian Actresses Are Plagiarizing My Quantum Mechanics Lecture to Sell Printers."[3] After relating what had happened and linking to the video, the post ended

> For almost the first time in my life, I'm at a loss for words. I don't know how to respond. I don't know which of 500 000 possible jokes to make. Help me, readers. Should I be flattered? Should I be calling a lawyer?

This would become the most notorious blog post I ever wrote. By the next morning, the story had made the *Sydney Morning Herald* ("Ad agency cribbed my lecture notes: professor"[4]), Slashdot ("Scott Aaronson, Printer Shill"[5]), and several other news sites. I happened to be in Latvia at the time, visiting my colleague Andris Ambainis, but somehow journalists tracked me down to my hotel room in Riga, waking me up around 5 a.m. to ask for interviews.

Meanwhile, reactions on my blog and in other online forums were mixed. Some readers said I'd be foolish if I didn't sue the ad agency for all it was worth. What if they had played a few beats of a *Rolling Stones* song, without first getting permission? Cases like that, I was assured, are sometimes settled for millions of dollars. Others said that even *asking the question* made me a stereotypical litigious American, a personification of everything wrong with the world. I should be flattered, they continued, that the ad writers had

[2] www.youtube.com/watch?v=saWCyZupO4U
[3] www.scottaaronson.com/blog/?p=277
[4] www.smh.com.au/news/technology/professor-claims-ad-agency-cribs-lecture-notes/2007/10/03/1191091161163.html
[5] idle.slashdot.org/story/07/10/02/1310222/scott-aaronson-printer-shill

seen fit to give *my* take on quantum mechanics all this free publicity. Dozens of commenters offered variations on the same insipid joke, that I should ask for a date with the "models" as my compensation. (I replied that I'd rather have a free printer, if it came down to it.) One commenter simply wrote, "This really could be the funniest thing that has ever happened."

For its part, Love Communications admitted that it had appropriated material from my lecture, but said it had consulted a lawyer and thought it was perfectly within its fair-use rights to do so. Meanwhile, I *did* get in touch with an Australian intellectual property lawyer, who said that I might have a case – but it would take time and energy to pursue it. I felt torn: on the one hand, plagiarism is one of the academic world's few unforgivable sins, and I was miffed by the agency's completely unapologetic response, after they'd been caught so red handed. On the other hand, if they had just *asked* me, I probably would have gladly given them permission to use my words, for either a token sum or no money at all.

In the end, we found a solution that everyone liked. Love Communications apologized (without admitting wrongdoing), and donated $5000 to two science outreach organizations of my choice in Australia.[6] In return, I didn't pursue any further action – and indeed, I mostly forgot about the affair, except when colleagues would rib me (as they continue to do) about Australian models.

But there's a final irony to the tale, and that's why I'm recounting it here (well, besides just that it's a hilarious true story involving this book). If I had to choose one passage from the entire book to be broadcast on TV, I think I would have chosen the exact same one that the commercial writers chose – even though they were presumably just trawling for some sciencey-sounding gobbledygook, and I hadn't highlighted the passage in any way, as its centrality hadn't occurred to me.

[6] See www.scottaaronson.com/blog/?p=297

The idea that quantum mechanics is "about" information, probabilities, and observables, rather than waves and particles, certainly isn't an original one. The physicist John Archibald Wheeler said similar things in the 1970s; and today an entire field, that of quantum computing and information, is built around the idea. Indeed, in the discussion on my blog that followed the Australian models episode, one the commonest (and to me, funniest) arguments was that I had no right to complain, because the appropriated passage *wasn't special in any way*: it was an obvious thought that could be found in any physics book!

How I wish it were so. Even in 2013, the view of quantum mechanics as a theory of information and probabilities remains very much a minority one. Pick up almost any physics book – whether popular or technical – and you'll learn that (a) modern physics says all sorts of paradoxical-seeming things, like that waves are particles and particles are waves, (b) at a deep level, no one really understands these things, (c) even translating them into math requires years of intensive study, but (d) they make the atomic spectra come out right, and that's what matters in the end.

One eloquent statement of this "conventional view" was provided by Carl Sagan, in *The Demon-Haunted World*:

> Imagine you seriously want to understand what quantum mechanics is about. There is a mathematical underpinning that you must first acquire, mastery of each mathematical subdiscipline leading you to the threshold of the next. In turn you must learn arithmetic, Euclidean geometry, high school algebra, differential and integral calculus, ordinary and partial differential equations, vector calculus, certain special functions of mathematical physics, matrix algebra, and group theory . . . The job of the popularizer of science, trying to get across some idea of quantum mechanics to a general audience that has not gone through these initiation rites, is daunting. Indeed, there are no successful popularizations of quantum mechanics in my opinion – partly for this reason. These

> mathematical complexities are compounded by the fact that quantum theory is so resolutely counterintuitive. Common sense is almost useless in approaching it. It's no good, Richard Feynman once said, asking why it *is* that way. No one knows why it is that way. That's just the way it is (p. 249).

It's understandable why physicists talk this way: because physics is an experimental science. In physics you're *allowed* to say, "these are the rules, not because they make sense, but because we ran the experiment and got such-and-such a result." You can even say it proudly, gleefully – *defying* the skeptics to put their preconceived notions up against Nature's verdict.

Personally, I simply *believe* the experimentalists, when they say the world works in a completely different way than I thought it did. It's not a matter of convincing me. Nor do I presume to predict what the experimentalists will discover next. All I want to know is: *What went wrong with my intuition? How should I fix it, to put it more in line with what the experiments found? How could I have reasoned, such that the actual behavior of the world* ***wouldn't*** *have surprised me so much?*

With several previous scientific revolutions – Newtonian physics, Darwinian evolution, special relativity – I feel like I more-or-less know the answers to the above questions. If my intuition isn't yet fully adjusted even to those theories, then at least I know how it *needs* to be adjusted. And thus, for example, if I were creating a new universe, I might or might not decide to make it Lorentz invariant, but I'd certainly *consider* the option, and I'd understand why Lorentz-invariance was the inevitable consequence of a couple of other properties I might want.

But quantum mechanics is different. Here, the physicists assure us, *no one knows* how we should adjust our intuition so that the behavior of subatomic particles would no longer seem so crazy. Indeed, maybe there *is* no way; maybe subatomic behavior will always remain an arbitrary brute fact, with nothing to say about it beyond "such-and-such formulas give you the right answer." My response is radical: if that's true, then *I don't much care* how subatomic particles behave. No doubt other people *need* to know – the people designing lasers or transistors, for example – so let them learn. As for me, I'll simply study another subject that makes more sense to me – like, say, theoretical computer science. Telling me that my physical intuition was wrong, without giving me any path to *correct* that intuition, is like flunking me on an exam without providing any hint about how I could've done better. As soon as I'm free to do so, I'll simply gravitate to other courses where I get As, where my intuition *does* work.

Fortunately, I think that, as the result of decades of work in quantum computation and quantum foundations, we *can* do a lot better today than simply calling quantum mechanics a mysterious brute fact. To spill the beans, here's the perspective of this book:

> *Quantum mechanics is a beautiful generalization of the laws of probability: a generalization based on the 2-norm rather than the 1-norm, and on complex numbers rather than nonnegative real numbers. It can be studied completely separately from its applications to physics (and indeed, doing so provides a good starting point for learning the physical applications later). This generalized probability theory leads naturally to a new model of computation – the quantum computing model – that challenges ideas about computation once considered a priori, and that theoretical computer scientists might have been driven to invent for their own purposes, even if there were no relation to physics. In short, while quantum mechanics was invented a century ago to solve technical problems in physics, today it can be fruitfully explained from an extremely different perspective: as part of the history of ideas, in math, logic, computation, and philosophy, about the limits of the knowable.*

In this book I try to make good on the above claims, taking a leisurely and winding route to do so. I start, in Chapter 1, as near to the "beginning" as I possibly can: with Democritus, the ancient Greek philosopher. Democritus's surviving fragments – which speculate, among other things, that all natural phenomena arise from complicated interactions between a few kinds of tiny "atoms," whizzing around in mostly empty space – get closer to a modern scientific worldview than anything else in antiquity (and certainly closer than any of Plato's or Aristotle's ideas). Yet no sooner had Democritus formulated the atomist hypothesis, than he noticed uneasily its tendency to "swallow whole" the very sense-experiences that he was presumably trying to explain in the first place. How could *those* be reduced to the motions of atoms? Democritus expressed the dilemma in the form of a dialogue between the Intellect and the Senses:

> Intellect: By convention there is sweetness, by convention bitterness, by convention color, in reality only atoms and the void.
>
> Senses: Foolish intellect! Do you seek to overthrow us, while it is from us that you take your evidence?

This two-line dialogue will serve as a sort of touchstone for the entire book. One of my themes will be how quantum mechanics seems to give both the Intellect *and* the Senses unexpected new weapons in their 2300-year-old argument – while still (I think) not producing a clear victory for either.

In Chapters 2 and 3, I move on to discuss the deepest knowledge we have that intentionally *doesn't* depend on "brute facts" about the physical world: namely, mathematics. Even there, something inside me (and, I suspect, inside many other computer scientists!) is suspicious of those *parts* of mathematics that bear the obvious imprint of physics, such as partial differential equations, differential geometry, Lie groups, or anything else that's "too continuous." So instead, I start with some of the most "physics-free" parts of math yet discovered: set theory, logic, and computability. I discuss the great discoveries of Cantor, Frege, Gödel, Turing, Church, and Cohen, which helped to map the contours of mathematical reasoning itself – and which, in the course of showing why all of mathematics can't be reduced to a fixed "mechanical process," also demonstrated just how much of it *could* be, and clarified what we mean by "mechanical process" in the first place. Since I can't resist, in Chapter 4 I then wade into the hoary debate about whether the human mind, too, is governed by "fixed mechanical processes." I set out the various positions as fairly as I can (but no doubt reveal my biases).

Chapter 5 introduces computability theory's modern cousin, *computational complexity theory*, which plays a central role in the rest of the book. I try to illustrate, in particular, how computational complexity lets us systematically take "deep philosophical mysteries" about the limits of knowledge, and convert them into "merely" insanely difficult unsolved mathematical problems, which arguably capture most of what we want to know! There's no better example of such a conversion than the **P** versus **NP** problem, which I discuss in Chapter 6. Then, as warmups to quantum computing, Chapter 7 examines the many uses of *classical* randomness, both in computational complexity and in other parts of life; and Chapter 8 explains how computational complexity ideas were applied to revolutionize the theory and practice of *cryptography* beginning in the 1970s.

All of that is just to set the stage for the most notorious part of the book: Chapter 9, which presents my view of quantum mechanics as a "generalized probability theory." Then Chapter 10 explains the basics of my own field, the *quantum theory of computation*, which can be briefly defined as the merger of quantum mechanics with computational complexity theory. As a "reward" for persevering through all this technical material, Chapter 11 offers a critical examination of the ideas of Sir Roger Penrose, who famously holds that the brain is not merely a quantum computer but quantum *gravitational* computer, able to solve Turing-uncomputable problems – and that this, or something like it, can be shown by an appeal to Gödel's Incompleteness Theorem. It's child's play to point out the problems with these ideas, and I do so, but what I find more interesting is to ask whether there *might* be nuggets of truth in Penrose's speculations. Then Chapter 12 confronts what I see as the central conceptual problem of quantum mechanics: not that the future is indeterminate (who cares?), but that the past is *also* indeterminate! I examine two very different responses to that problem: first, the appeal, popular among physicists, to *decoherence*, and to the "effective arrow of time" supplied by the Second Law of Thermodynamics; and second, "hidden-variable theories" such as Bohmian mechanics. Even if hidden-variable theories are rejected, I find that they lead to some extremely interesting mathematical questions.

The rest of the book consists of applications of the perspective developed earlier, to various big, exciting, or controversial questions in math, computer science, philosophy, and physics. Much more than the earlier chapters, the later ones discuss *recent research* – mostly in quantum information and computational complexity, but also a bit in quantum gravity and cosmology – that strikes me as having some hope of shedding light on these "big questions." As such, I expect that the last chapters will be the first to become outdated! While there are minor dependencies, to a first approximation the later chapters can be read in any order.

- Chapter 13 discusses new notions of mathematical proof (including probabilistic and zero-knowledge proofs), then applies those notions to understanding the computational complexity of hidden-variable theories.
- Chapter 14 takes up the question of the "size" of quantum states – do they encode an exponential amount of classical information, or not? – and relates this question to the quantum interpretation debate on the one hand, and to recent complexity-theoretic research on quantum proofs and advice on the other.
- Chapter 15 examines the arguments of quantum computing *skeptics*: the people who hold, not merely that building a practical quantum

computer is hard (which everyone agrees about!), but that it can *never be done* for some fundamental reason.

- Chapter 16 examines Hume's Problem of Induction, using it as a jumping-off point for discussing *computational learning theory*, as well as recent work on the learnability of quantum states.
- Chapter 17 discusses some breakthroughs in our understanding of classical and quantum *interactive proof systems* (e.g., the **IP** = **PSPACE** and **QIP** = **PSPACE** theorems), but is mostly interested in those breakthroughs insofar as they've led to *non-relativizing circuit lower bounds* – and, therefore, might illuminate something about the **P** versus **NP** question.
- Chapter 18 examines the famous Anthropic Principle and "Doomsday Argument"; the discussion starts out highly philosophical (of course), but eventually winds its way to a discussion of *postselected quantum computing* and the **PostBQP** = **PP** theorem.
- Chapter 19 discusses Newcomb's Paradox and free will, leading into an account of the Conway–Kochen "free will theorem," and the use of Bell's Inequality to generate "Einstein-certified random numbers."
- Chapter 20 takes up time travel: in a now-familiar pattern, starting with a wide-ranging philosophical discussion, and ending with a proof that classical *or* quantum computers with closed timelike curves yield exactly the computational power of **PSPACE** (under assumptions that are open to interesting objections, which I discuss at length).
- Chapter 21 discusses cosmology, dark energy, the Bekenstein bound, and the holographic principle – but, not surprisingly, with an eye toward what all these things mean for *the limits of computation*. For example, how many bits can one store or search through, and how many operations can one perform on those bits, without using so much energy that one instead creates a black hole?
- Chapter 22 is "dessert": it's based off the final lecture of the Quantum Computing Since Democritus class, in which the students could ask me anything whatsoever, and watch me struggle to respond. Topics addressed include the following: the possible breakdown of quantum mechanics; black holes and "fuzzballs"; the relevance of oracle results in computational complexity; **NP**-complete problems and creativity;

"super-quantum" correlations; derandomization of randomized algorithms; science, religion, and the nature of rationality; and why computer science is not a branch of physics departments.

A final remark. One thing you *won't* find in this book is much discussion of the "practicalities" of quantum computing: either physical implementation, or error correction, or the details of Shor's, Grover's, or other basic quantum algorithms. One reason for this neglect is incidental: the book is based on lectures I gave at the University of Waterloo's Institute for Quantum Computing, and the students were already learning all about those aspects in their other classes. A second reason is that those aspects are covered in *dozens* of other books[7] and online lecture notes (including some of my own), and I saw no need to reinvent the wheel. But a third reason is, frankly, that the technological prospect of building a new kind of computer, exciting as it is, is not why I went into quantum computing in the first place. (*Shhh*, please don't tell any funding agency directors I said that.)

To be clear, I think it's entirely possible that I'll see practical quantum computers in my lifetime (and also possible, of course, that I *won't* see them). And if we *do* get scalable, universal quantum computers, then they'll almost certainly find real applications (not even counting codebreaking): mostly, I think, for specialized tasks like quantum simulation, but to a lesser extent for solving combinatorial optimization problems. If that ever happens, I expect I'll be as excited about it as anyone on earth – and, of course, tickled if any of the work I've done finds applications in that new world. On the other hand, if someone gave me a practical quantum computer tomorrow, then I confess that I can't think of anything that I, personally, would want to use it for: only things that *other people* could use it for!

[7] The "standard reference" for the field remains *Quantum Computation and Quantum Information*, by Michael Nielsen and Isaac Chuang.

Cambridge University Press
978-0-521-19956-8 - Quantum Computing since Democritus
Scott Aaronson
Frontmatter
More information

Partly for that reason, if scalable quantum computing were proved to be *im*possible, that would excite me a thousand times more than if it were proved to be possible. For such a failure would imply something wrong or incomplete with our understanding of quantum mechanics itself: a revolution in physics! As a congenital pessimist, though, my *guess* is that Nature won't be so kind to us, and that scalable quantum computing will turn out to be possible after all.

In summary, you could say that I'm in this field less because of what you could do with a quantum computer, than because of what the *possibility* of quantum computers *already* does to our conception of the world. *Either* practical quantum computers can be built, and the limits of the knowable are not what we thought they are; *or* they can't be built, and the principles of quantum mechanics themselves need revision; *or* there's a yet-undreamt method to simulate quantum mechanics efficiently using a conventional computer. All three of these possibilities sound like crackpot speculations, but at least one of them is right! So whichever the outcome, what can one say but – to reverse-plagiarize a certain TV commercial – "that's interesting?"

WHAT'S NEW

In revising this manuscript for publication, the biggest surprise for me was how much *happened* in the fields discussed by the book between when I originally gave the lectures (2006) and "now" (2013). This book is supposed to be about deep questions that are as old as science and philosophy, or at the least, as old as the birth of quantum mechanics and of computer science almost a century ago. And at least on a day-to-day basis, it can *feel* like nothing ever changes in the discussion of these questions. And thus, having to update my lectures extensively, after the passage of a mere six years, was an indescribably pleasant burden for me.

Just to show you how things are evolving, let me give a partial list of the developments that are covered in this book, but that *couldn't* have been covered in my original 2006 lectures, for the simple reason that they hadn't happened yet. IBM's Watson computer defeated the *Jeopardy!* world champion Ken Jennings, forcing me to update my discussion of AI with a new example (see p. 37), very different in character from previous examples like ELIZA and Deep Blue. Virginia Vassilevska Williams, building on work of Andrew Stothers, discovered how to multiply two $n \times n$ matrices using only $O(n^{2.373})$ steps, *slightly* beating Coppersmith and Winograd's previous record of $O(n^{2.376})$, which had held for so long that "2.376" had come to feel like a constant of nature (see p. 49).

There were major advances in the area of *lattice-based cryptography*, which provides the leading candidates for public-key encryption systems secure even against quantum computers (see pp. 105–107). Most notably, solving a 30-year-old open problem, Craig Gentry used lattices to propose the first *fully homomorphic cryptosystems*. These systems let a client delegate an arbitrary computation to an untrusted server – feeding the server encrypted inputs and getting back an encrypted output – in such a way that only the client can decrypt (and verify) the output; the server never has any clue what computation it was hired to perform.

In the foundations of quantum mechanics, Chiribella *et al.* (see p. 131) gave a novel argument for "why" quantum mechanics should involve the specific rules it does. Namely, they proved that those rules are the only ones compatible with certain general axioms of probability theory, *together with* the slightly mysterious axiom that "all mixed states can be purified": that is, whenever you don't know everything there is to know about a physical system A, your ignorance must be fully explainable by positing correlations between A and some faraway system B, such that you *would* know everything there is to know about the combined system AB.

In quantum computing theory, Bernstein and Vazirani's "Recursive Fourier Sampling" (RFS) problem – on which I spent a fair bit of time in my 2006 lectures – has been superseded by my "Fourier Checking" problem (see p. 145). RFS retains its place in history, as the first black-box problem ever proposed that a quantum computer can provably solve superpolynomially faster than a classical probabilistic computer – and, as such, an important forerunner to Simon's and Shor's breakthroughs. Today, though, if we want a candidate for a problem in **BQP**\PH – in other words, something that a quantum computer can easily do, but which is not even in the classical "polynomial-time hierarchy" – then Fourier Checking seems superior to RFS in every way.

Happily, several things discussed as "open problems" in my 2006 lectures have since lost that status. For example, Andrew Drucker and I showed that **BQP/qpoly** is contained in **QMA/poly** (and, moreover, the proof relativizes), falsifying my conjecture that there should be an oracle separation between those classes (see p. 214). Also, in a justly celebrated breakthrough in quantum computing theory, Jain *et al.* proved that **QIP** = **PSPACE** (see p. 263), meaning that quantum interactive proof systems are no more powerful than classical ones. In that case, at least, I conjectured the right answer! (There was actually *another* breakthrough in the study of quantum interactive proof systems, which I *don't* discuss in the book. My postdoc Thomas Vidick, together with Tsuyoshi Ito,[8] recently showed that **NEXP** $\subseteq$ **MIP***, which means that any *multiple*-prover interactive proof system can be "immunized" against the possibility that the provers secretly coordinate their responses using quantum entanglement.)

Chapter 20 of this book discusses David Deutsch's model for quantum mechanics in the presence of closed timelike curves, as well as my (then-)new result, with John Watrous, that Deutsch's model provides exactly the computational power of **PSPACE**. (So that, in particular, quantum time-travel computers would be no more powerful than *classical* time-travel computers, in case you

[8] T. Ito and T. Vidick, A Multi-prover Interactive Proof for NEXP Sound against Entangled Provers. In *Proceedings of IEEE Symposium on Foundations of Computer Science* (2012), pp. 243–252.

were wondering.) Since 2006, however, there have been important papers questioning the assumptions behind Deutsch's model, and proposing alternative models, which generally lead to computational power *less* than **PSPACE**. For example, one model, proposed by Lloyd *et al.*, would "merely" let the time traveler solve all problems in **PP**! I discuss these developments on pp. 319–322.

What about circuit lower bounds – which is theoretical computer scientists' codeword for "trying to prove **P** ≠ **NP**," in much the same way that "closed timelike curves" is the physicists' codeword for "time travel?" I'm pleased to report that there have been interesting developments since 2006, certainly more than I would have expected back then. As one example, Rahul Santhanam used interactive proof techniques to prove the non-relativizing result that the class **PromiseMA** doesn't have circuits of any fixed polynomial size (see p. 257). Santhanam's result was part of what spurred Avi Wigderson and myself, in 2007, to formulate the *algebrization barrier* (see p. 258), a generalization of Baker, Gill, and Solovay's relativization barrier from the 1970s (see pp. 245–246). Algebrization explained why the interactive proof techniques can take us only so far and no further in our quest to prove **P** ≠ **NP**: as one example, why those techniques led to superlinear circuit lower bounds for **PromiseMA**, but not for the class **NP** just "slightly below it." The challenge we raised was to find new circuit lower bound techniques that convincingly *evade* the algebrization barrier. That challenge was met in 2010, by Ryan Williams' breakthrough proof that **NEXP** $\not\subset$ **ACC**$^{\mathbf{0}}$ (discussed on pp. 260–261).

Of course, even Williams' result, exciting as it was, is a helluva long way from a proof of **P** ≠ **NP**. But the past six years have also witnessed a flowering of interest in, and development of, Ketan Mulmuley's Geometric Complexity Theory (GCT) program (see pp. 261–262), which is to proving **P** ≠ **NP** almost exactly as string theory is to the goal of a unified theory of physics. That is, in terms of concrete results, the GCT program hasn't yet come anywhere close to fulfilling its initial hopes, and even the program's most ardent proponents predict a slog of many decades, while its mathematical complexities frighten everyone else. What GCT has going for it is two things: firstly, that it's forged mathematical connections "too profound and striking to be mere coincidence," and secondly, that it's perceived (by no means universally!) as "the only game in town," the only hunter currently in the forest who's even carrying a sharp stick.

Let me mention just three other post-2006 developments relevant to this book. In 2011, Alex Arkhipov and I proposed "BosonSampling" (see pp. 287–288): a rudimentary, almost certainly *non*-universal quantum computing model involving non-interacting photons, which was just recently demonstrated on a small scale. Interestingly, the evidence that BosonSampling is hard to simulate on a classical computer seems *stronger* than the evidence that (say) Shor's factoring algorithm is hard to simulate. In 2012, Umesh Vazirani and Thomas Vidick, building on earlier work of Pironio *et al.*, showed how to use violations of the Bell inequality to achieve *exponential randomness expansion* (see p. 305): that is, converting n random bits into 2^n bits that are guaranteed to be almost-perfectly random, *unless* Nature resorted to faster-than-light communication to bias the bits. Meanwhile, the debate about the "black hole information paradox" – i.e., the apparent conflict between the principles of quantum mechanics and the locality of spacetime, when bits or qubits are dropped into a black hole – has evolved in new directions since 2006. Possibly the two most important developments have been the increasing popularity and sophistication of Samir Mathur's "fuzzball" picture of black holes, and the controversial argument of Almheiri *et al.* that an observer falling into a black hole would never even get near the singularity, but would instead encounter a "firewall" and burn up at the event horizon. I cover these developments as best I can on pp. 346–349.

A few updates were occasioned not by any new discovery or argument, but simply by me (gasp) *changing my mind* about something. One example is my attitude toward the arguments of John Searle and Roger Penrose against "strong artificial intelligence." As you'll see in Chapters 4 and 11, I still think Searle and Penrose are *wrong* on crucial points, Searle more so than Penrose. But on rereading my 2006 arguments for *why* they were wrong, I found myself wincing at the semi-flippant tone, at my eagerness to *laugh* at these celebrated scholars tying themselves into logical pretzels in quixotic, obviously doomed attempts to defend human specialness. In effect, I was lazily relying on the fact that everyone in the room already agreed with me – that to these (mostly) physics and computer science graduate students, it was simply self-evident that the human brain is nothing other than a "hot, wet Turing machine," and weird that I would even waste the class's time with such a settled question. Since then, I *think* I've come to a better appreciation of the immense difficulty of these issues – and in particular, of the need to offer arguments that engage people with different philosophical starting-points than one's own.

Here's hoping that, in 2020, this book will be as badly in need of revision as the 2006 lecture notes were in 2013.

Scott Aaronson
Cambridge, MA
January 2013

Acknowledgments

As my summer student in 2008, Chris Granade enthusiastically took charge of converting the scattered notes and audio recordings from my course into coherent drafts that I could post on my website, the first step on their long journey into book form. More recently, Alex Arkhipov, my phenomenal PhD student at MIT, went through the drafts with a fine-tooth comb, flagging passages that were wrong, unclear, or no longer relevant. I'm deeply grateful to both of them: this book is also *their* book; it wouldn't exist without their help.

It also wouldn't exist without Simon Capelin, my editor at Cambridge University Press, who approached me with the idea. Simon understood what I needed: he prodded me every few months to see if I'd made progress, but never in an accusatory way, always relying on my own internal guilt to see the project through. (And I *did* see it through – eventually.) Simon also assured me that, even though *Quantum Computing since Democritus* was . . . a bit *different* from CUP's normal fare, he would make every effort to preserve what he called the book's "quirky charm." I also thank all the others at CUP and Aptara Corp. who helped to make the book a reality: Sarah Hamilton, Emma Walker, and Disha Malhotra.

I thank the students and faculty who sat in on my "Quantum Computing since Democritus" course at the University of Waterloo in Fall 2006. Their questions and arguments made the course what it was (as you can still see in this book, especially in the last chapters). On top of that, the students also took care of the audio recordings and preliminary written transcripts. More broadly, I remember my two years as a postdoc at Waterloo's Institute for Quantum Computing as one of the happiest times of my life. I thank everyone there, and especially IQC's director Ray Laflamme, for not only *letting* me teach such a nutty course but *encouraging* it, and even (in Ray's and several other cases) sitting in on the course themselves and contributing many insights.

I thank MIT's Computer Science and Artificial Intelligence Laboratory and its Electrical Engineering and Computer Science Department, as well as the US National Science Foundation, the Defense Advanced Research Projects Agency, the Sloan Foundation, and TIBCO, Inc., for all the support they've given me over the last six years.

I thank the readers of my blog, *Shtetl-Optimized* (http://www.scottaaronson.com/blog), for their many comments on the draft chapters that I posted there, and for catching numerous errors. I especially thank those readers who encouraged me to turn *Quantum Computing since Democritus* into a book – some even promised they'd buy it when it came out.

I thank the people who advised me from my high school to my postdoc years: Chris Lynch, Bart Selman, Lov Grover, Umesh Vazirani, and Avi Wigderson. John Preskill was never "formally" an advisor, but I still think of him as one. I owe all of them more than I can say. I also thank everyone else in (and beyond) the quantum information and theoretical computer science communities whose discussions and arguments with me over the years left their imprints on this book. I can't possibly produce a full list of such people, so here's a partial one: Dorit Aharonov, Andris Ambainis, Dave Bacon, Michael Ben-Or, Raphael Bousso, Harry Buhrman, Sean Carroll, Greg Chaitin, Richard Cleve, David Deutsch, Andy Drucker, Ed Farhi, Chris Fuchs, Daniel Gottesman, Alex Halderman, Robin Hanson, Richard Karp, Elham Kashefi, Julia Kempe, Greg Kuperberg, Seth Lloyd, Michele Mosca, Michael Nielsen, Christos Papadimitriou, Len Schulman, Lenny Susskind, Oded Regev, Barbara Terhal, Michael Vassar, John Watrous, Ronald de Wolf. I apologize for the inevitable omissions (or to those who don't want their names in this book, you're welcome!).

I thank the following alert readers for catching errors and omissions in the first printing of this book: Boaz Barak, Evan Berkowitz, Ernest Davis, Bob Galesloot, Vijay Ganesh, Yuri Gurevich, John Kadvany, Andrew Marks, Cris Moore, Reviel Netz, and Tyler Singer-Clark.

Lastly, I thank my mom and dad, my brother David, and of course my wife Dana, who will now finally be able to know me while I'm *not* putting off finishing the damn book.