1 The Claimant, the Readers, and the Crowd

Vinay Deolalikar could be the featured subject of this chapter. Or it could be the small group of readers of his 102-page paper in August 2010 claiming to prove $\mathsf{P} \neq \mathsf{NP}$, and thus resolve in the negative the central open problem in our field of complexity theory. Initially this was a private group to whom the draft paper had been circulated, before its existence was leaked and made public on the blog of Greg and Kat Baker on August 7, 2010. However, the larger interest of the story told here by Ken is the way the review was *crowdsourced* and found what still seem to be all major issues in under two weeks.

Our story of the people and our field's greatest problem centers on a proof, or rather the claim of a proof and the social process this entails. A proof really ceases to be a proof when it is found to have mistakes, though we can still call its written-up form a "proof." The two-week adventure of the process that played out is best told as we experienced it personally, rather than reproducing the posts on our blog. The posts were not the real story anyway. Ken picks up the story.

1.1 The News

After a tumultuous summer attending the birth of a new niece in June just as my wife suddenly needed eye surgery, and a trip to Maine and New Jersey beginning in late July, I expected to settle into a few weeks of August summer routine at home before the Fall 2010 term began. Our first full day back in Buffalo was Sunday, August 8. After church service I forgot to switch my cellphone back on, and unusually for me, did not go online either. I had helped with a post late that Friday night from New Jersey, which Dick had posted Saturday afternoon, so I put the blog out of mind until Monday. Debbie cleaned while I mowed the lawn, and then we discussed what the family routine should be heading into the next school year, before taking a break with the newspapers. As dinner approached I did laundry in the basement, still oblivious to several progressively more excited e-mails and cell calls from Dick about the story which had broken the previous night. Then I switched on my laptop down there and remembered my cellphone too. If you recall a scene in a sci-fi movie where a starship's console suddenly lights up, that's what it felt like.

R.J. Lipton, K.W. Regan, *People, Problems, and Proofs*,
DOI 10.1007/978-3-642-41422-0_1,

Dick had in fact already drafted a post. He had glitches with the Georgia Tech e-mail system and expressed worry whether I had received any of the mails. I sent a quick acknowledgment at 5:39pm to Dick and Subruk, saying I would catch news online before reading the draft. I found some discussion about the paper already, including a quotation from Steve Cook that it looked like a "serious attempt," and the potential dimension of this started widening in my mind. With minimal delay of our 6:30 dinner, I modified a paragraph where Dick had raised a potential objection on whether the proof avoided the so-called *relativization barrier*, and added a paragraph on how it might avoid the *natural proofs* barrier. I also extended a third paragraph to try to remedy the paper's immediately obvious lack of an actual super-polynomial lower bound on deterministic time for a particular NP-complete problem such as SAT, which all but the most mysteriously indirect proofs of $\mathsf{P} \neq \mathsf{NP}$ should be expected to do.

1.2 The Paper and First Post

After dinner, still doing laundry in our basement, I noticed that Dick had obtained and included a link to the paper itself. I sat down at my laptop and examined the new wonder, whose first version had 63 pages. Though it ranged from physics to first-order logic to conditional probability spaces, I had some background in all of its areas, and could readily follow the structure of the argument. Most in particular, I had covered Gibbs distributions—as applied to neural networks—in a seminar in the early 1990's while supervising my first graduate student, Dr. Arun K. Jagota. I had first picked up relevant parts of mathematical physics during my graduate study at Oxford under Dominic J.A. Welsh, even fielding a question about Boltzmann distribution in my oral qualifying exam and refuting an assertion someone had made about percolation in a physics paper.

About forty minutes into reading, I perceived a potential flaw in the paper, one that harked back to a speculative idea I'd had in 1990. I sent Dick a query on something related to it which I didn't understand, and while waiting for reply, I started writing a potential section with my own objection. When we talked a half-hour later at 8:30, I explained my line of thought, and Dick concurred.

We both decided, however, that my point was better as a comment right below the main post, as we felt it was too early to put any judgments in the main body. Dick had sole control then of the blog's starship console, so he undertook the sometimes-wonky conversion from our LaTeX source to the WordPress HTML format. After one more typo-catch exchange he sent me a note at 9:28pm saying he had posted and to go with my comment—the times show an hour earlier since the blog's clock stays on Standard Time.

Dick titled it, "A Proof That P Is Not Equal To NP?" He noted that Deolalikar had sent the paper link with permission, and also linked the post by Baker which had broken the news the previous night. We were not fully aware that the news had been leaked to Baker against Deolalikar's own wishes to keep the paper in private circulation, a policy that we ourselves have observed and recommended in other cases. Dick posed the question, "Is the paper correct?" Several times starting in the next sentence, he—meaning we—referred to Deolalikar's paper as a "proof," "his proof." This caused consternation among many who would say we should refer to it only as a "claimed proof" or the like, but I had approved using "proof" to mean a piece of text representing a coherent mathematical argument, even though I myself already had a concrete doubt on its correctness. After a quick read, and before letting departmental colleagues know of this by e-mail, I started to put in my comment.

1.3 My Snap-Doubt and Comment

When I clicked to open a comment box there were as yet no comments, but two would sneak in ahead of mine before I finished fixing up the symbols and changing some things I had written. I hadn't reckoned with a basic fact of "Blog Geometry" that sundry replies to those two comments would soon push mine far down the page. Here is what I wrote, formatted back into LaTeX:

> Having seen this only since 5:30 this (Sunday) afternoon, here's my attempt at a somewhat more fleshed-out description of the proof strategy, still at a very high (and vague) level: Deolalikar has constructed a vocabulary V such that:
> (1) Satisfiability of a k-CNF formula can be expressed by NP-queries over V—in particular, by an NP-query Q over V that ties in to algorithmic properties.
> (2) All P-queries over V can be expressed by FO+LFP formulas over V.
> (3) $\mathsf{NP} = \mathsf{P}$ implies Q is expressible by an LFP+FO formula over V.
> (4) If Q is expressible by an LFP formula over V, then by the algorithmic tie-in, we get a certain kind of polynomial-time LFP-based algorithm.
> (5) Such an algorithm, however, contradicts known statistical properties of randomized k-SAT when $k \geq 9$.
> If there is a conceptual weakness, it might be in step 3. This step may seem obvious given step 2, and maybe the paper has been set up so that it does follow immediately. However, it does need comment, because general statements of the P–LFP characterization are careful to say that every polynomial-time decidable query over a vocabulary V can be encoded and represented by an LFP+FO formula over $\mathsf{Strings}$.

I went on to make an analogy with that old idea from 1990, in which I defined a class Q based on polynomial-time machines that could learn their input only via certain logical queries. Then NQ equaled NP, because the machines could guess queries to make that would reveal the entire input, but Q itself is only a small subclass of P. Thus I had $\mathsf{NQ} \neq \mathsf{Q}$ with $\mathsf{NQ} = \mathsf{NP}$, but this did not entail $\mathsf{Q} = \mathsf{P}$. I finished my comment by speculating that the paper might also be defining a "Q" that likewise falls short of being equal to P. This hunch turned out to be right in several respects.

No one replied to the comment, but I took that as meaning no one found fault with my outline of the paper's strategy. After sending e-mail telling colleagues about this, I went up from the basement to rejoin my family. But I snuck back down after making sure Debbie was comfortable and saying good-whatever to our more-nocturnal kids. I noticed a comment with a query from Cris Moore, and answered it tying back to my long comment. Then I looked for more information around the Net, and noted that certain other computational theory blogs had not yet picked up the story. And so to bed, still with little idea what I'd be waking up to the next morning—especially given that we had posted in time for morning not just in Europe but even the Far East.

1.4 Comments and Comments and Comments

When I arrived at my office in mid-morning, there were 13 new top-level comments, some with replies already, beginning with one Moore had entered right after midnight as a followup to the query I'd answered. The last was a comment from David Mix Barrington passing on a potential flaw noted by Lance Fortnow. It shows as "9:30am" because the blog times are Eastern Standard, i.e., GMT-5. Of course I could not make 13 replies, so I chose one detailed comment by Arthur Milchior, a student from France who also does standup comedy and edits Wikipedia pages in logic and computer science theory. I answered him but also addressed Barrington's comment and some other points. Then after printing a longer version of the paper which Deolalikar had released in the meantime, and starting to read it more closely, I noticed Barrington himself had replied to *my* reply only 30 minutes later. This was my first cue about the degree to which the comment threads would become *interactive*.

Dick sent me e-mail noting that WordPress projected the post would have over 30,000 readers in 24 hours, fifteen times our typical rate, and could become one of the top WordPress pages for the day. My colleague Atri Rudra noted that fellow complexity theory blogger Scott Aaronson had picked up the story that morning, and had offered $200,000 of his own money as a supplement if the proof turned out to be correct and full enough to win the $1,000,000 Clay Millennium Prize. Lance Fortnow and Bill Gasarch, the original complexity bloggers, had links to that and to us, and gave some specific reasons for doubting that the ingredients of Deolalikar's paper would measure up. I could have spent more time troving for evaluations on other blogs besides those.

However, I did what one might expect a researcher rather than an online maven to do. I went over with the paper to the University at Buffalo Commons to buy lunch and stake out a table outside in the shade for an afternoon of reading on a beautiful August Monday. I looked over the first-order logic aspects that they all had been querying, but was most interested in the latter part of the paper where the actual lower bound was argued. The writing in terms of "phase transitions" was hazy, and as we had already noted in the post, did not give a concrete time lower bound on a concrete problem.

Most in particular, I thought it must be possible to extract from the argument a particular *hardness predicate* that was being employed. A hardness predicate $R(H, n)$ is one that is false for all sufficiently large n when H is an easy problem, and true for all sufficiently large n (or at least infinitely many n) for *some* other problems H. The issue is whether you can define R so that it applies to a problem H that you are interested in, such as SAT, and then whether you can *prove* $R(H, n)$

for the n's you need. Then you have a proof that H is hard—and if H is SAT and your "easy" class really does include all of P, then you have a proof that $\mathsf{NP} \neq \mathsf{P}$.

The celebrated 1994 "Natural Proofs" paper of Alexander Razborov and Stephen Rudich placed notable restrictions on predicates R for which this strategy could succeed, when "easy" was taken to refer to the class P/poly of problems having polynomial-sized *circuits*, which includes P. They proved that unless *all* one-way functions—including the ones underlying the RSA cryptosystem as used for Internet commerce—are substantially less secure than currently propounded, then a predicate R for which this strategy could possibly succeed must either be *strangely thin* or *extraordinarily complex*. The former means that $R(H, n)$ would hold only for a negligible fraction of problems H that are actually hard. The latter means that merely evaluating $R(H, n)$—given any representation of the problem H on inputs of length n such as a 2^n-sized table of values—would take time more than singly exponential in n, such as doubly exponential in n. After demonstrating that all hardness predicates heretofore employed in successful lower bounds were neither thin nor complex, and noting that basically all properties commonly used in relevant mathematical fields are decidable in time singly exponential in n, Razborov and Rudich generated a consensus that theirs was a substantial barrier any hardness proof such as Deolalikar's must overcome. There was a potential "out" in the difference between P/poly and P, but it was not clear that Deolalikar's strategy was really driving at it.

I had studied hardness predicates arising in algebraic geometry and the theory of ideals that were hard for exponential space—hence ostensibly outside single exponential time—and yet were mathematically tractable to work with. I had also written a survey article explaining how the predicates employed in a far-reaching program by Ketan Mulmuley escaped a single exponential bound. Hence I expected a similarly complex predicate to lie at the heart of the phase-transition argument, one whose novelty would be valuable even if the paper itself failed. I tried to tease it out, but falling short of success as the afternoon wore on, I started getting frustrated with the paper's lack of explicitness. I returned to my office after 3pm. Little did I know that from then on my time would be mostly occupied in systematizing and summarizing the way others were reading the paper, not so much from other blogs but from myriad comments in ours.

I came back to find all the comment threads expanding, while some other researchers had provided detailed notes and queries on points in the paper. Indeed I noticed that the overnight comment by Moore, which my morning reply had passed over, actually struck at more-immediate aspects of the part of the argument I had just been examining. When Dick and I spoke that afternoon, we realized the need to organize and summarize what people were saying.

The comments clustered around four major issues. We started drafting a post titled "Issues in the Proof That $\mathsf{P} \neq \mathsf{NP}$". We hyperlinked certain comments in the original post for each of the four issues. We finished and uploaded it Monday evening. Right after it went up, I added a comment noting a question raised by Ryan Williams on his Twitter account, a comment related to one of the issues in Slashdot's item from the previous day, and several matters raised by a commenter named "vloodin" whose real identity I still do not know. The next morning a commenter named "harrison" replied by querying another issue raised by Ryan about isolation of k-SAT formulas in the phase-transition argument, and Ryan himself replied twice in return. I had an urge to reply with matters from my read of the paper that we had left out of the post, but realized doing so would take me too long past midnight and so went to bed.

1.5 Back to the Future

When I awoke Tuesday morning, I found a flurry of e-mails in my box from between 1am and 3:30am, four from Terence Tao, two from "Geomblog" creator Suresh Venkatasubramanian, one from Gil Kalai, and all copying Dick and Timothy Gowers. They proposed organizing the investigation of the paper under the "PolyMath" wiki-based structure proposed in a paper by Gowers with Michael Nielsen, "Massively collaborative mathematics," in *Nature* vol. 466, Oct. 2009, pages 879–881. After Dick signaled our co-operation, I replied:

> I also agree, and will be happy to take advantage of the framework. I've been interested in the co-ordination of research discussion by Internet ever since the astounding Queen endgame of the Kasparov–World match in 1999.

This brought a flashback of pain and promise. In summer 1999 I was at the flood of my own concerted attempt to prove $\mathsf{NP} \neq \mathsf{P}$, or rather an algebraic analogue of it made famous by Leslie Valiant. My main graduate student in this work told me of a chess match sponsored by Microsoft to advertise their online "MSN Gaming Zone" in which Garry Kasparov was to take on the entire world voting on moves at the MSN site, at the rate of one move per 24 hours. I've never taken up playing chess online, and I ignored his repeated suggestions until well into July.

I found that MSN had set up a system where three young teenage players of about my playing strength (one division above Master but short of Grandmaster) would each suggest a move, and the World would vote for one of those moves or for a write-in. They were Elisabeth Paehtz of Germany, Etienne Bacrot of France, and Irina Krush of Brooklyn. I picked up the game at Move 27 in a position that was *wild*, and found that the wildness had come from a fantastic new sacrificial move suggested by Krush for Move 10 of her side playing Black. In fact, early this year, 2013, world champion Viswanathan Anand won a brilliant game using Krush's move when his opponent dared him to repeat it, so it's a durably good move. The World players in 1999 were so appreciative that they voted for every move suggested by Krush subsequently, until the critical final phase in which the vote was apparently first hacked by a personage calling himself "Joe One-Two" in Spanish, and then a communications breakdown occurred between Krush and MSN on the final relevant move.

Two things about the match disquieted me. First, players of full Grandmaster strength were being organized to analyze and provide input on the forum provided by MSN, one from St. Petersburg, Russia, calling themselves the "GM School." I felt there should have been a stated policy that no one above the level of the three junior panelists should take part. Second, there were a number of people posting "flames" on the board, directed largely at Krush when she recommended a move other than what they favored, especially when she proposed trading Queens to enter a desperate-looking endgame. Krush was being managed by a childhood chess friend of mine, and suddenly this seemed personal. I looked at the position after the trade enough to tell it was still fully dynamic. So I waded in on the forum myself to say hey guys, let's appreciate how amazing the game has been and still is, with every single one of the seven remaining pawns—two for Kasparov as White and five for the World as Black—being a so-called "passed pawn" readily capable of becoming a new Queen.

Vacation more than research had made me a less-involved onlooker in August 1999, but as the battle raged into September, a long looming unavoidable sequence emerged in which both sides would make a Queen. In this I saw an aspect that *is* research. The sequence would leave just the King, new Queen, and another pawn for Kasparov, versus King, new Queen, and two pawns for the World. But Kasparov's other pawn would be a greater menace than our two combined, and I recognized that the World would still have to play with utmost care to earn the golden ticket of a draw. The GM-schoolers had peeled away, but amateur-level players on the forum started looking ahead at the sheaves of variations burgeoning from the two female coronations at the end of the forced sequence. There were to be only seven pieces, but the complexity would ramp up to levels unseen even in this game.

The endgame has always been my strength as a player, and I recognized general strategy principles to inform and organize the analysis efforts. I spent several days—that is nights—writing up what became a manifesto of *thirty* numbered points, titled "World Team Endgame Strategy Explained." It had the same dozen-page length as a typical conference research paper submission. It was hailed by those doing the hard work on the forum. Someday I hope to tell the story fully, including how my strategy caused us to play a move now known to be losing, and an ingenious trap discovered by a player one division below Master that Kasparov's own notes showed no inkling of. Alas the aforementioned communications breakdown prevented a crucial preparatory move for the trap from being posted as Irina's choice, and the World voted for a superficial centering move that lost in short order.

Of Kasparov I greatly appreciate that even though the unexpected effort and duration of the game had derailed his preparations for his forthcoming title defense, he still cared enough to release thirteen pages of notes to claim that he would have won even without the mishap. I did, however, refute his claim within two days. Then the "postgame" entered a new interface of computer and chess spheres, in which exhaustive enumeration of possible positions took over from the brave human analysts. The resulting tabulation of perfect play (modulo ignorance of underpromotion which was later rectified without major change) upheld my refutation 100 %, but also revealed just before Thanksgiving 2010 that White has a winning line no human had thought of. Kasparov later published it without comment in his book on the match.

By this time I was well distracted from my student and our research drive, which was soon blunted anyway by a refutation of its workable hypothesis by a mathematician in Italy. I could say much more about my own chess analysis, but when I had written my strategy article in September I wasn't promoting my own chess—I was helping a host of others, and that was the story.

Come a decade later, and the larger story was recognized by Nielsen himself. While writing his 2011 book *Reinventing Discovery: The New Era of Networked Science*, he made Kasparov-World one of his main running examples, and drew upon my materials. Now I was graced with the opportunity, indeed the obligation, to do the same action regarding the most central problem in my professional field—and with some of the same people involved. I felt amazed and excited.

I also felt bulldozed—as I stared at the reasons why scrolling to the ends of our two posts still left the scrollbar a tiny little square down only an inch from the top of the window:

205 Comments 127 Comments

—numbers that would go over 250 and 300 in succeeding posts. Indeed when Dick started drafting a post announcing the PolyMath wiki page, he did so partly to enable a fresh comment thread.

1.6 The Group

Our Tuesday post on "The Group" also addressed three suggestions to Dr. Deolalikar himself, though in third person, on easier partial results that should be entailed by his methodology that would already be significant advances, and whose demonstration would do a lot to satisfy the community. Dick did the posting—I was not yet on the masthead—while I prepared my own contribution to the wiki page based on what I had gleaned from the paper the day before. I sent it as an e-mail internally to the five other group members.

I expected some internal discussion. That would be an *in*volution—the way things had been discussed for centuries. What was happening, however, was *e*volution with heavy emphasis on the root stem *e-* or *ex-* meaning *out*—evolution meaning *turned outwards*, so onto the wiki it went, to see what response my writeup might get from outside. Outsiders, however, had plenty to investigate before getting to my hypotheticals. Nobody I know (including myself) has taken time since then to delve into them further, although they are still the only way I see to extract something solid and publishable from Deolalikar's work.

I read the initial forms of the wiki page, and then went over to the burgeoning comment threads, and read not just each for content but also to see that they were in sync with each other. From then on I did not contribute directly to the wiki page, but kept this indirect role, and I also started posting some comments to clarify and answer queries myself. Especially since Dick was occupied enough already by communications with ACM officials and other principals, I saw keeping abreast of the comment threads as my best purpose.

1.7 Internet Non-resistance

Overnight Tuesday there began some criticism of the whole event. Russell Impagliazzo, whom I might nominate as the most quietly powerful researcher pushing the field for a quarter century, commented at midnight Tuesday ET that he was "distressed to see so many people spending a lot of time on this," and that the paper was not a serious attempt and had no original ideas. By the time I woke up Wednesday morning, there were a couple dozen replies agreeing and disagreeing, very respectfully, including a followup by Russell himself. But there was also a long anonymous comment by someone labeled "Concerned Citizen," who began by quoting Russell's words and went on:

> Thank you Russell for saying something non-anonymously that the rest of us are thinking. Look at the paper. It is not a serious attempt. It is not a rigorously written mathematical argument. If any serious mathematician really believed they had an approach to P vs. NP, they would definitely seek to write it in a straightforward, easily verifiable way. And Deolalikar has a PhD in mathematics; he should know the standards of mathematical writing.
>
> Read the paper. He is more interested in buzzwords than clarity. It is quite clear that there is something very wrong in the model theory part of the paper, which is where all the interesting action has to be happening. After having a very weak (and incorrect) characterization of P, the proof could end in a variety of ways. The author has chosen to reach a contradiction using random k-SAT. And this, apparently, is leaving some people (who are unfamiliar with the work on random CSPs) hailing it as a "new approach."
>
> I think that many bloggers have become quite enamoured with being part of the publicity storm, and are thus complicit in its continuation. The situation is very simple: If he believes he has a proof, then he should write a 10-page version without all the unnecessary background material, stream of consciousness rambling, and overly verbose ruminations. Just the math, please.
>
> What you will find is not that he has discovered a shocking new technique. This is not very good work that happens to have a subtle flaw. This will not solve some special case like NL vs. NP. Instead, it is simply a misunderstanding caused by a lack of rigor and a lack of professional mathematical colleagues against whom he can check his ideas.
>
> Let's try to be responsible here, as a community. Let's not say things like "Vinay's use of conditional independence seems strikingly new." How can you say this without understanding the proof, even at a very basic level? This reeks of bombastic pseudoscience.

I had written the line in the post causing the offense of this paragraph, based on the context of his Gibbs argument and my reading that Monday. I was gratified that my Sunday evening inkling about the characterization of "P" being too weak was seen as such. After again noting the unclarity in Deolalikar's paper, the comment concluded:

> The biggest mistake here is the response of the community. We should definitely NOT be supportive of such efforts. Think about it: Do you want to have such a reaction every week, to every new proof claiming to resolve a famous open question? We should reinforce the responsible action: To produce a clear, concise, carefully written proof and send it quietly to a few experts who can give it a first reading. After it has been revised and rewritten to their standards, humbly circulate it more widely; say that you are hopeful that your proof is correct, and you would be exceedingly grateful to address the comments of interested readers.
>
> As currently written and announced, Deolalikar's manuscript is disrespectful to the community. We are given a bad name so he can have 15 minutes of fame. Please stop propagating this madness. It's irresponsible.

I woke up on Wednesday to find a few responses to this comment also: an anonymous "Amen," and then "Hear, hear" from Hugo van den Berg of Warwick, who had also chimed in with support of Russell and scorn of the paper, and some longer defenses of Deolalikar and the situation. I imagined that people would be interested to see what the reaction of Dick and me would be. And I quickly reached a conclusion on exactly what that reaction should be:

> Nothing.

My reason was based on what was happening just below this section of the thread. Kalai had posed a technical question on what another person examining the paper had commented and got a reply from him. Alberto Atserias stated his agreement with Russell but with his very next sentence helped us by saying exactly where he thought the flaw in the model-theory part of the paper was, and referenced two comments in previous posts (one in Aaronson's item). Timothy Gowers sharpened the call for a synopsis of how the proof worked. A stream of further substantive notes leaning on previous ones was starting, including especially quickly perceptive ones by "vloodin."

I decided it was of utmost importance not to disturb this discussion by doing anything to give oxygen to the "meta-discussion." I called Dick first-thing about this and he agreed. Hence even when some others defended us, even about what we had opined was new, Dick and I just stayed silent. That afternoon, Tao gave what we still regard as the central view:

> [Russell,] I understand your concern about the snowball effect, but actually I think this time around, the fact that we concentrated all the discussion into a single location made it more efficient than if everybody just heard rumours that everybody else was looking at it, and dozens of experts each independently and privately wasted hours of time doing the same basic groundwork of reading the entire paper before reaching their conclusions. I think the net time cost has actually been reduced by this method, though of course the cost is far more visible as a consequence.
> I think there are a number of useful secondary benefits too. For instance, I think this experience has greatly improved our $\mathsf{P} \neq \mathsf{NP}$ Bayesian spam filter that is already quite good (but not perfect) at filtering out most attempts at this problem; and I think we are better prepared to handle the next proposed solution to a major problem that makes it past our existing spam filters. Also, I think the authors of such proposed solutions may perhaps also have a better idea of what to expect, and how to avoid unnecessary trouble (well, the best advice here is not even to try solving such problems directly), but somehow I think this advice will not be heeded [this came with a smiley emoticon]. Such authors often complain that their work gets ignored, but actually they may be lucky in some ways, when compared against those authors whose work gets *noticed*...

It is important to be reminded that Deolalikar himself had originally not wanted to be noticed. By Wednesday we understood fully that he had been "outed" by a leak from someone in the original distribution. Once that happened he had no choice but to go all the way public.

Now we hoped not only that he would respond to the three suggestions in our post, but also that he too would become a part of the public discussion. We did in fact get a private response from him, and he made a short statement that accompanied a revision of his paper.

1.8 Company...

Vinay Deolalikar's response, however, answered only a relatively easy question: how and why his paper distinguished between values of k in arguing structural properties of a space associated to the k-SAT problem that imply non-membership in P, so as to avoid the appearance queried by some that it would imply the false statement 2SAT $\notin$ P. It did not engage with suggestions and queries in our posts or the many comments which we and the wiki page had summarized. We were disappointed, but Dick went ahead and framed a post for Wednesday around his answer, our fourth post in four days.

Meanwhile, we had communications that brought into focus what is now regarded as the first of three main flaws in the paper, about the use of logic. Neil Immerman, whom we each regard as a personal friend as well as one of the foremost experts on finite model theory, addressed a two-page letter to Dr. Deolalikar in full detail, and gave us leave to release it in a fifth post, which we did on Thursday, August 12. We highlighted the following sentences from the letter:

> Thus, your restriction to only have successor and to restrict to monadic fixed points is fundamental. In this domain—only monadic fixed points and successor—FO(LFP) does not express all of P!

This post, which was current for three days and remained on the WordPress "Top Posts" list even through the following Monday, attracted 327 comments. Robert Solovay piped up to say he had placed Immerman's letter on the wiki page. Anuj Dawar, perhaps Britain's leading finite model theory expert, concurred. Janos Simon and Mauricio Karchmer commented about possible satellite ideas. There was also a followup comment from "Concerned Citizen," but this time the 'meta-discussion' was shorter with fewer siding with him. My favorite reply was from someone styled "Quantum Tennis Referee":

> As if the community needs protecting! ha! What a wild notion!

An interesting sequence near the top of that thread comes I believe closest to the sociological heart, though still from the cynical side. Amir Michail, the creator of DropZap for iPhone and other games, first imputed (by reference to "cynical people") that the attention had been motivated "to promote the field, attract graduate students, and increase theory funding." I hope this chapter makes clear that this opinion is wrong. But following up to a responder he said something closer:

> Do you think the publicity from this will have an impact on complexity theory? If so, what sort of impact?
> I find it hard to believe that those who commented on the proof gave no thought whatsoever to the fact that the world is watching. This has become a spectator sport.

To which an anonymous responder affirmed, "Amir: People are commenting here *precisely* because they know the world is watching them."

Now here is my take: Most people who combine the terms "Internet" and "global village" attach an emotion of hope to their statements, and expectation that this is for good. There is a worldwide community at the juncture of computer science and mathematics. When we come together for conferences or workshops there is a large component of *conviviality*, meant according to its roots of "together" and "live" more than drink and merriment. We cannot otherwise get the *con-* part, even by e-mails, which are largely individual. But the comment threads took a non-negligible step toward a virtual gathering. Various personal opinions amounted to statements of community values, and even when some of those values were opposed they jointly mapped community feelings.

In brief, the threads became a source of *company*. I myself noticed some people I'd been closer to in the past century, and sent short greeting e-mails. There is not space here to mention them, or many others I've known or known-of who made substantive contributions.

1.9 ... And a Crowd

There was more in that comment sequence. A commenter named "Random" whom I'd seen on the first day wrote:

> Paradoxically, I think a lot of people who are still paying attention are non-specialists (like me). For people who know enough about barriers, it must have been easy to take a cursory glance, convince oneself that the strategy was flawed, and go back to one's own research.

Alexander Razborov—he of the "Natural Proofs" barrier mentioned above—responded "yes" but added:

> P.S. But I have to admit I am a little bit harassed by the Russian media these days (I am sure many colleagues have similar problems), and it is *extremely* handy to have this blog around as a pointer. Dick, the Group, et al.—thank you *very* much for investing effort into this.

Some media had indeed reached us. Lee Gomes of Forbes Online contacted me on the first Tuesday and kept pace with our posts and the comments:

(1) "The Non-Flaming of an HP Mathematician" (August 10).
(2) "A Beautiful Sausage" (August 11).
(3) "How to Think Like a Mathematician" (August 12).
(4) "Now It's the Slow Roasting of an HP Mathematician" (August 17).

As the titles already make clear, the main subject was not the paper itself but rather the process of reviewing research collaboratively online, and of doing research (that way) to begin with. His first column remarked on the decorum of the review and the comment threads. In an update to it he quoted me defending Deolalikar's actions and the autonomy of research at places like HP. In the second he quoted a comment by Tao:

> Regardless of how significant or useful Deolalikar's proof is (and how much of a "waste of time" it would be for so many people to be looking at it), it has provided a rare opportunity to gather a surprisingly large number of experts to have an extremely high quality conversation around $\mathsf{P} \neq \mathsf{NP}$. One may have strong feelings about the specific circumstances that created this opportunity, but I do not think that should be cause to squander the opportunity altogether.

This article went on to tell how aspects of the discussion were being made "accessible to anyone" in both the posts and the comments, even joking that mutual-fund managers could thereby "learn a lot about how to think clearly." The last reflected the reality that our fifth post had the words "fatal flaws" in its title, as consensus grew that the jig was up with the paper. But the paper was less and less the story anyway, and the article by John Markoff in the Tuesday August 17 "Science Times" section of The New York Times led with the main subject.

> The potential of Internet-based collaboration was vividly demonstrated this month when complexity theorists used blogs and wikis to pounce on a claimed proof for one of the most profound and difficult problems facing mathematicians and computer scientists.

Dick and I started another post to mark a week since the release of the paper, and to provide a fresh space for comments. Only rarely does the photo atop our posts have more than one person, but this time Dick chose to feature a *crowd*. We were grateful that the intent expressed at the beginning by Suresh on his own blog of "crowdsourcing" the review had worked out as well as it had. For the body of the post we worked on summarizing what had been established about the paper and its issues thus far.

1.10 The Issues Crystallize

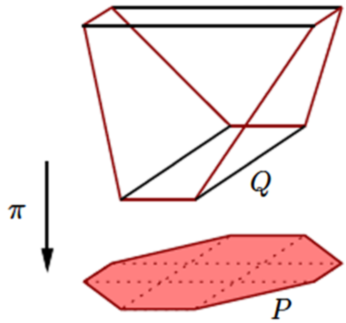
I in particular was occupied that Saturday with the current long comment thread, besides also helping edit an item by Dick for the *Communications of the ACM* blog. Saturday afternoon I came again over comments regretting that someone like Tao—meaning a generic person of his talent—should "waste his time" on something like this. At the same time I noticed that the actual Tao was contributing some new comments, until by about 3pm he had 8 of the 15 most recent comments appearing in a sidebar to the blog.

I was still catching up with observations Tao had made in the same thread the previous day, including pointing to a deep assessment by Gowers on his own blog, so I did not immediately pick up what Tao was doing. I helped Dick draft the body around what we called "The Two Main Issues". Immerman's point reappeared as the first issue, and then I wrote two sections giving larger shrift to the solution space issues which we had first noticed in the comments from Ryan Williams. I actually tweaked and shortcutted Ryan's up-to-date argument a little, and wondered if I'd be corrected—there not being time to run my writeup by him first. I also referenced the beginning of the string of comments by Tao that I'd marveled at that afternoon, but by the time we hit send on the post early on Sunday the 15th, I hadn't appreciated where they led.

What Tao had actually done was find a third issue, more fundamental than the other two. Commenter "vloodin" piped up about our omission right away, but Tao—who among all his talents has mastered arts of community—quickly put all three into proper perspective, in a most genteel way. And this issue turns on a principle whose statement can be appreciated by all:

A shadow can sometimes be more complex than the object it is projected from.

Here is a visual example reproduced from a paper by Sebastian Pokutta and others in the case of polytopes:



Here "complexity" refers to counting the facets of the polytopes, which is the same as counting the number of half-spaces used to define them. The 3D object has 6 facets, but its 2D shadow has 8.

In formal mathematics, a shadow is defined by a *projection* operation. This notion applies formally even in complexity theory. By that notion, NP is a projection of P, exactly the one that Deolalikar himself had purposed to prove is more complex.

What Tao had ferreted out of the part of the paper dealing with physics and conditional probability distributions—the part least familiar to us complexity-theory people—was an implicit assumption that a certain complexity property would not increase under projections. In that context it was an easy intuition to have, but a false one. Tao phrased this in terms of what he called "ppp" not being equal to Deolalikar's "pp", with the extra 'p' standing for "projection." Once Dick and I understood this Sunday evening into Monday, it was our greatest "aha" moment of the whole episode.

1.11 The Falling Action

Our Sunday post also had a three-page response by Deolalikar to "3 issues," but while the third was on the subject of how his conditional-probability argument was parameterized, it did not yet see the projection-closure issue in its clear form. He promised that these issues would be addressed more fully in a "version being prepared, which will be sent for peer and journal review." Thus he signalled his desire to revert to how he had originally intended his work to be appraised, not under the eyes of the whole world. Alas in a play one cannot go back from Act IV to Act II. We quoted a paragraph that he had addressed to the first-order objection, but followed by writing:

> The last sentence aims to answer Neil's objection, but does it? Recall Neil's contention is that the class X whose power is quantified by the finite-model theory section of the paper is not P, but rather a class already known to be properly contained in P. No one has disputed Immerman's counter-claim. We are, as we stated earlier, disappointed that this essence of Neil's communication has not been engaged.

This post attracted 264 comments and links over three days. Over that time Dick tried to draw more out of Deolalikar himself, but he instead removed the paper drafts from public access on his webpage, leaving only the three-page synopsis. Almost everyone I've mentioned and more contributed to the technical discussion, which crystallized exactly how the projection argument connected to the other components of the paper. The question of whether the space called SAT0 in the post could be the support of a 'ppp' distribution was hammered out in the affirmative, but this did not patch the paper. The thread then shifted over to Deolalikar's shutting his doors, while Dick and I answered 'no' about whether we or anyone else in the "Group" circle had received a revised manuscript.

Dick drafted a new post for the 18th. We felt it important to do a "rewind" and explain for the general public why a *proof* of $\mathsf{P} \neq \mathsf{NP}$ would be important. What is a proof itself good for? Why desire a proof of this inequality when most practitioners have latched on to the reality that most NP-complete problems have very many instances that fast programs *are* able to solve well? He then attached my roundup of the developments with Deolalikar and perspective of the status, the latter linking an opinion by Williams.

Meanwhile, I occupied myself with explaining the projection issue for the public. This became only my second full post for the blog—recall I was not yet on the masthead. For some time I had been intending to present a famous lower bound argument by Volker Strassen that my own effort toward $\mathsf{NP} \neq \mathsf{P}$ had focused on trying to extend. I was excited to realize that I could do both at once, because I could cast Strassen's argument as an instance where closure under projections *does* succeed. Thus it would present the intellectual impasse in Deolalikar's paper as a compare-contrast, while having value apart from the previous ten days' episode.

Dick looked at my incomplete first draft on the 18th, and drew from his knowledge of mathematical history an event that made a perfect lead-in. The great Henri Lebesgue had once made the error of assuming that the collection of sets defined by Émile Borel was closed under projection. That this was false came as a double surprise to me. First, I had forgotten it from my undergraduate topology and analysis courses. Second and more notably, I recalled Mike Sipser's analogy between Borel sets and the P, NP, co-NP, ... hierarchy, under which projections *do* stay within the hierarchy.

After getting Dick's preamble that evening, I worked past 3am to finish the post. I did this and still managed to fulfill my fatherly duty of taking the kids to a 9:30am haircut appointment. This was in someone's house, and while playing with the family dogs and stepping over young kids' toys and sitting outside in the yard, I appreciated the material concreteness away from a glowing screen, while letting the swirling projections of the eleven days subside into a dazed satisfaction. I also appreciated a supportive family in circumstances that made unusual demands at a difficult time. This post went out on the 19th and is reproduced here as Chap. 32.

Dick followed four days later with a short post looking back and referencing his article for *Communications of the ACM*, but there was no more news and only a little more discussion of the paper, as comments went under 50 in both posts. After a weekend preparing for the beginning of classes at Georgia Tech and the University at Buffalo, we reverted to an "easy" and "uncontroversial" topic: quantum computation.

1.12 What Did We Learn?

There have unfortunately not been any happy developments to report with Vinay Deolalikar himself in the three years since. His revised paper doubled in length to 200 pages, but we glean that it did not increase appreciably in content. He has not retracted his claim or otherwise responded about it in a public way. He continues to work at HP Labs but at press time his webpage has no research entries after 2010, and on this affair has only a statement that seems frozen in time:

> BASIC RESEARCH
> Vinay Deolalikar. $\mathsf{P} \neq \mathsf{NP}$. The preliminary version was meant to solicit feedback from a few researchers as is customarily done. It illustrated the interplay of principles from various areas, which was the major effort in constructing the proof. I have fixed all the issues that were raised about the preliminary version in a revised manuscript; clarified some concepts; and obtained simpler proofs of several claims. Once I hear back from the journal as part of due process, I will put up the final version on this website.

I stand by my belief that a publishable paper can be salvaged from the conditional-probability and first-order part, plus a scholium on hardness predicates, but with so much going on, including successful development of my model of human decision-making at chess and its employment in several cheating cases since then, I have not taken any step to further this. Nothing directly positive that we know has been traced back to this work.

As for what we learned about mass collaborative research, this is still awaiting the proverbial "proof of the pudding." We have not reported on the blog any mathematical theorem that was obtained since then by open discussion by a dozen or more researchers. It is easier to tear potential proofs down than to build them up.

At least we showed that constructive discussion can be maintained in an open online forum. The hopeful relevance for scientific practice was hailed by Stephen Landsburg, an economist at the nearby University of Rochester and popular writer, in a post on his "Big Questions" blog titled "O Brave New World"—on August 16th in the fullest swing. It was picked up in the New Scientist magazine the following month, and has been referenced since, including being the sole subject of a May 4, 2012 article by Srinath Perur for India's "Fountain Ink" e-magazine, titled "How the science works in science." Well, there are many ways science works in science, and in an arena of survival of the fittest, we still need to see how this fits.

We have, however, been able to apply the most close-in lessons to our own practice. We have been slowly and carefully and privately mediating a potential approach and claim to prove the opposite, that $\mathsf{P} = \mathsf{NP}$. We expect that this claim will be public before this book appears, but our private contacts both pro and con recognize the need to find the best presentation of both a lemma–theorem–proof structure and fine details of notation, while keeping the paper well under 100 pages.

Most important, in 2012 we repeated the experiment in a different subject with two willing parties in a controlled fashion from the start, and stimulated some positive research. Yes, the subject was quantum computation. We hosted a debate on whether quantum computers will ever be feasible to build on a large enough scale to be useful. Gil Kalai, whom I've mentioned, who also publishes the blog "Combinatorics and More," took the "con" side, while Aram Harrow, recently hired to MIT's physics faculty, defended the consensus of the quantum computation community. What had been envisioned as a three-part debate became eight posts spanning much of the year, and several of these posts topped 200 comments with much technical discussion that never veered from decorum. Out of this so far has come a paper by Harrow with Steve Flammia showing that one of Kalai's conjectures fails for ternary quantum primitives in place of binary quantum bits, and two papers by renowned physicist John Preskill who also took part in the discussions. Preskill's papers, one updating the other, clarified the argument over proper noise models for quantum computers in a way supporting Harrow's position.

Dick and I ourselves intend to further this debate, I with an argument that Strassen's lower-bound predicate—the same one in my projections post—also furnishes an algebraic measure of multi-way entanglement in quantum circuits. This is still for the future as I write, as is so much else.

1.13 Notes and Links

The eight posts:

http://rjlipton.wordpress.com/2010/08/08/a-proof-that-p-is-not-equal-to-np/
http://rjlipton.wordpress.com/2010/08/09/issues-in-the-proof-that-p≠np/
http://rjlipton.wordpress.com/2010/08/10/update-on-deolalikars-proof-that-p≠np/
http://rjlipton.wordpress.com/2010/08/11/deolalikar-responds-to-issues-about-his-p≠np-proof/
http://rjlipton.wordpress.com/2010/08/12/fatal-flaws-in-deolalikars-proof/
http://rjlipton.wordpress.com/2010/08/15/the-p≠np-proof-is-one-week-old/
http://rjlipton.wordpress.com/2010/08/18/proofs-proofs-who-needs-proofs/
http://rjlipton.wordpress.com/2010/08/19/projections-can-be-tricky/

Greg Baker post:

http://gregbaker.ca/blog/2010/08/07/p-n-np/

Vinay Deolalikar's webpage at HP Labs:

http://www.hpl.hp.com/personal/Vinay_Deolalikar/

Michael Nielsen, *Reinventing Discovery*:
http://press.princeton.edu/titles/9517.html
My September 1999 chess strategy guide:
http://www.cse.buffalo.edu/~regan/chess/K-W/wtstrategy.html
Lee Gomes' columns in Forbes Online:
http://www.forbes.com/sites/leegomes/2010/08/10/the-non-flaming-of-an-hp-mathematician/
http://www.forbes.com/sites/leegomes/2010/08/11/a-beautiful-sausage/
http://www.forbes.com/sites/leegomes/2010/08/12/how-to-think-like-a-mathematician/
http://www.forbes.com/sites/leegomes/2010/08/17/now-its-the-slow-roasting-of-an-hp-mathematician/
Dick's CACM post:
http://cacm.acm.org/blogs/blog-cacm/97587-a-tale-of-a-serious-attempt-at-p≠np/fulltext
Scott Aaronson's $200,000 offer:
http://www.scottaaronson.com/blog/?p=456
Steven Landsburg's blog post:
http://www.thebigquestions.com/2010/08/16/o-brave-new-world/
John Markoff's New York Times article:
http://www.nytimes.com/2010/08/17/science/17proof.html

Picture credits:
Polytope and projections: presentation slides by Sebastian Pokutta,
http://spokutta.wordpress.com/2012/01/05/1311/