From *SIAM News*, Volume 38, Number 4, May 2005

Math/Physics Collaboration Sheds New Light on Computational Hardness

By Sara Robinson

All of the familiar jokes about mathematicians and physicists highlight their very different approaches to research: Mathematicians tend to focus on simplified models about which they can prove theorems, while physicists, guided by intuition, come up with solutions that are often not rigorous. Despite these cultural differences, points of contact between the disciplines have been increasing as each side has begun to see the advantages of viewing problems from a different angle.

During the past year, a new collaborative effort has been gaining steam. A group of mathematicians and theoretical computer scientists have joined forces with statistical physicists in an attempt to better understand a model of long-standing interest to both sides. Thus far, the collaboration has been spectacularly successful: Researchers in each discipline have borrowed tools from the other to make significant progress on open questions, and a completely new picture of the model is developing. Though it is still a work in progress, researchers on both ends are hopeful that there will be further leaps of insight.

"It used to be that we weren't interested in the same issues and paid little attention to what others were doing," says Marc Mézard, a physicist at the Université de Paris Sud in Orsay, France, who just wound up a three-month visit to the Mathematical Sciences Research Institute in Berkeley. "Now, we really start to understand each other."

The model at the heart of this unusual interaction lies at the nexus of computer science and statistical physics. Known as "random k-SAT," it captures an apparent connection between computational hardness, the domain of mathematicians and computer scientists, and phase transitions, a focus for statistical physicists.

The Random k-SAT Model

Scientists in many fields have long been interested in solutions to what are known as "constraint satisfaction problems." Such a problem involves a large set of variables, each taking values in a small domain, and a collection of constraints, each forbidding some of the possible joint values for the variables. For a given instance, the question is whether there are values for the variables that simultaneously satisfy all the constraints.

Constraint satisfaction problems seem to crop up everywhere: In information theory, they are part of the theory of error-correcting codes. In statistical physics, they take the form of spin glasses—physical systems, realized as crystals with impurities, where each particle's spin is constrained by the spins of particles around it. Mathematicians have long studied graph colorings, yet another version of constraint satisfaction. For computer scientists, the canonical constraint satisfaction problem is k-SAT.

A k-SAT problem involves a fixed set of n Boolean variables. Variables or their negations are strung together with "or" symbols into "clauses" of length k, which are then joined by "and" symbols into "formulas." Each clause thus precludes one of the 2^k possible joint values of the variables it contains. The question, then, is whether there are assignments for the variables that "satisfy" every clause, making the formula a logically true statement.

Long known to be NP-complete for $k > 2$, k-SAT is in fact the archetypal NP-complete problem: It was shown to be so in the paper that defined the notion of NP-completeness. But even though worst-case instances of k-SAT are thought to be intractable (the P $\neq$ NP conjecture), it could be that most instances are easy. In the 1970s, hoping to understand the complexity of typical problems, some computer scientists turned their attention to randomly generated instances of k-SAT.

Random formulas sampled from most distributions turned out to be easy to solve. Yet in the early 1990s, when considering formulas with a fixed ratio of clauses to variables, computer scientists noticed a curious phenomenon. When the ratio is small, formulas have many variables and few constraints; there are many satisfying assignments and it's easy for simple algorithms to find one of them. When the ratio is large, the variables tend to be overly constrained, and formulas almost certainly have no satisfying assignments. Remarkably, as the ratio, r, of clauses to variables grows, the transition from probably satisfiable to probably unsatisfiable is not gradual, but abrupt.

Researchers conjectured that there is a single threshold value of r, depending only on k, such that as the number of variables goes to infinity, random formulas with $r < r_k$ are almost surely satisfiable, while random formulas with $r > r_k$ are almost surely unsatisfiable. This dramatic shift, viewed in the right way, is a type of phase transition, akin to ice melting into water at a fixed temperature.

Computer scientists also observed that random k-SAT problems are hardest right around this transition point. There, the time needed for algorithms to decide whether a formula is satisfiable is maximal. This seeming connection between computational hardness and the "satisfiability threshold" prompted a race to prove the existence of the threshold and determine its exact location.

That the threshold is sharp was known for 2-SAT, instances of which are easily solvable in linear time, but for $k > 2$ the exact nature of the threshold is still unknown. In 1998, Israeli mathematician Ehud Friedgut proved a broad theorem, a consequence of which is a near resolution of the satisfiability threshold conjecture. As n grows, Friedgut showed, the transition from satisfiable to unsatisfiable indeed becomes a sharp jump. His proof, however, leaves open the possibility that the location of the threshold does not converge to a limit, but rather oscillates between several different values. Nevertheless, most computer scientists and mathematicians regard this as a technicality and believe that random k-SAT does undergo a sharp transition at a critical ratio r_k.

Many computer scientists were skeptical of attempts to link the threshold to computational complexity, noting that the hardness of randomly generated formulas is not necessarily related to the hardness of the k-SAT problems that come up in practice. A series of recent results, however, have changed the face of the problem, and are changing attitudes as well.

Two years ago, a group of physicists gave evidence that the hardness of random k-SAT is connected not to the satisfiability threshold phase transition, but to another phenomenon altogether. For a given formula, the geometry of the space of solutions undergoes a dramatic transformation—also thought to be a phase transition—at a ratio well below the satisfiability threshold. It is this latter transition, researchers now believe, that is the key to understanding what makes some instances of k-SAT hard. The description of this second transition was initially based only on intuition gained from studying physical phenomena. Only now is a clear and mathematically rigorous picture beginning to unfold.

Separately, a computer scientist and a mathematician dramatically improved the lower bound for the satisfiable–unsatisfiable transition, confining the threshold within a small interval. Remarkably, the proof techniques used to help pinpoint the location of the threshold are now being adopted by physicists, who are using them to make their insights rigorous.

Zeroing In on the Threshold

For the past two decades, computer scientists have been zeroing in on the value r_k that defines the transition from almost surely satisfiable to almost surely unsatisfiable. An upper bound of $2^k \ln 2$ on r_k follows from the fact that the probability of at least one satisfying assignment is bounded by $2^n(1 - 2^{-k})^{rn}$, which tends to 0 for $r > 2^k \ln 2$. Until recently, lower bounds were given by producing algorithms that can quickly find a satisfying assignment up to that value of r. These algorithms fail far below the upper bound, however, giving lower bounds of order $c2^k/k$ that differ only in the leading constant. This left a chasm that grows rapidly with k between the largest value of r for which solutions provably exist and the smallest value for which solutions are known not to exist.

To significantly improve the lower bound, researchers had to turn to arguments that were not algorithmic. In 2003, Dimitris Achlioptas of Microsoft Research and Yuval Peres, a probabilist at the University of California at Berkeley, proved that satisfying assignments almost surely exist for r as large as $2^k \ln 2 - k$, far higher than what the best algorithms could handle and within a linear term of the upper bound. Using a probabilistic technique, the researchers proved that solutions exist without giving a way to find them. The proof technique—a surprising application of the "second moment" method—has since been used by physicists to prove that their intuitions about random k-SAT are accurate.

The second moment method relies on a simple probabilistic inequality: The probability that a non-negative random variable is strictly positive is bounded below by the ratio of its expected value squared to its second moment. In the case of k-SAT, when the random variable is simply the number of satisfying assignments for the given formula, the inequality yields no useful information: The ratio converges rapidly to zero as n grows. Still, for a given formula, it is equally valid to consider any random variable X that is positive only if the formula is satisfiable. For example, X could be a weighted sum over the satisfying assignments that weights some assignments more than others.

The first bound along these lines came in a 2002 paper by Achlioptas and Cristopher Moore, a computer scientist at the University of New Mexico. They defined a random variable that assigns non-zero weight only to those satisfying assignments whose complements are also satisfying. For this random variable, the ratio is bounded away from zero for all $r < 2^{k-1} \ln 2 - O(1)$, yielding a lower bound for the critical threshold that is within a factor of 2 of the upper bound. In 2003, Achlioptas joined forces with Peres, who has specialized in weighted second moment arguments. Together, they came up with a way of doing a systematic search for weights that asymptotically maximize the ratio, yielding the lower bound $2^k \ln 2 - k$. The optimal weighting, it turns out, promotes those satisfying assignments that minimally satisfy every clause.

Survey propagation brought disparate research communities together. Although, to date, the algorithm's run-time and probability of success are unproven, its performance in experiments is so spectacular "that people in other communities had to look at it, even if it seemed to come from strange ideas."

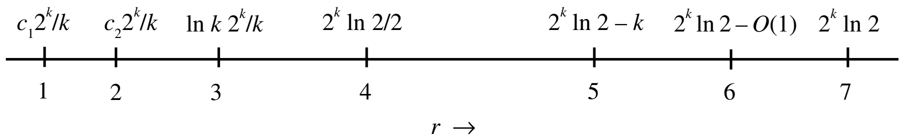
Enter the Physicists

Meanwhile, some physicists, too, were captivated by the random k-SAT model because it involves a phase transition. Their goal was to understand the transition in terms of a phase diagram, which requires knowing (for them, not the same as proving) exactly where the transition occurs. Toward this end, three European physicists—Mézard, Giorgio Parisi of the University of Rome, and Riccardo Zecchina of the International Centre for Theoretical Physics in Trieste—devised in 2002 a new algorithm for solving random constraint satisfaction problems that far outperformed existing methods.

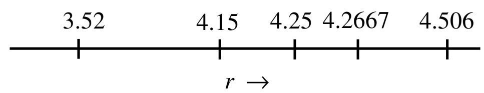
Inspired by two powerful ideas from statistical physics, the algorithm—called "survey propagation"—solves instances of random 3-SAT with millions of variables for $r = 4.25$ (a hairsbreadth from the conjectured threshold of around 4.27) in a matter of minutes. By comparison, standard approaches would take billions of years to solve problems of this size. "It blew us away," says Achlioptas.

The new algorithm was the catalyst that brought the disparate research communities together, Mézard says. Although, to date, the algorithm's run-time and probability of success are unproven, its performance in experiments is so spectacular, he says, "that people in other communities had to look at it, even if it seemed to come from strange ideas." Moore, whose background is in both computer science and physics, agrees: "It's been an evolution for the computer scientists. Rather than pooh-poohing the physicists for being non-rigorous, there was an effort to understand what they were saying."

Key to the physicists' approach was intuition about how the geometry of the solution space behaves as formulas approach the threshold. Consider the space of all possible assignments for a given random formula, expressed as bit strings, where two assignments are neighbors if they differ by a single bit flip. The physicists' conjecture, based on similar behavior in some systems in the natural world, was that for low values of r, the solutions to that formula are connected in essentially one big ball. As r gets closer to the threshold, the ball develops holes, which stretch into larger holes. Then, abruptly, in another phase transition, the ball shatters into exponentially many clusters of solutions spaced far apart. Inside each cluster, a large fraction of the variables are frozen; that is, their values are fixed.



Key values of r, the ratio of clauses to variables, in random k-SAT.
1. The largest r for which known algorithms, with high probability, find solutions quickly. 2. Certain practical algorithms are known to take exponential time beyond this point. 3. Physicists say this is about where fragmentation occurs. 4. At this point, the solution space is known to be fragmented into clusters with frozen variables. 5. The Achlioptas–Peres lower bound for the satisfiability threshold. 6. Physicists conjecture that this is the satisfiability threshold, r_k. 7. The (trivial) upper bound on the threshold.



The picture for k = 3.
3.52 is the highest value of r for which an algorithm for 3-SAT provably finds solutions. By r = 4.15, physicists believe, fragmentation has occurred. Survey propagation produces solutions for values of r as high as 4.25. By 4.2667, formulas are experimentally shown to be no longer satisfiable. 4.506 is the best proven upper bound for the random 3-SAT threshold, given by Dubois, Boufkhad, and Mandler in 2002.

As the value of r increases further, approaching the satisfiability threshold, the clusters shrink almost uniformly and, on reaching the threshold, disappear altogether. Physicists believe that the fragmentation occurs at about $\ln k\, 2^k/k$, which is just past the region where all rigorously analyzed algorithms for k-SAT break down, Achlioptas points out. It is this transition, some researchers now believe, that is central to understanding the complexity of the problem.

This complex picture, Mézard says, was developed from physical intuition, and checked for self-consistency by long computations that took many months to complete, but is still very far from proven. Nevertheless, it was critical to the design of the new algorithm. For $k = 3, 4$, and 5, survey propagation finds solutions far beyond the conjectured point of fragmentation and very near the conjectured threshold. (For larger values of k, such experiments are impossible because they would require too much computer memory.)

Survey Propagation

At its core, survey propagation is just a very complex variation of belief propagation, a method long used in information theory and machine learning, as well as in statistical physics. Belief propagation tries to satisfy global constraints by first finding solutions to local constraints and then attempting to patch them together. Applying belief propagation to an instance of k-SAT requires transforming the formula into a "factor graph" in which the nodes are variables and clauses. A variable is connected to a clause by an edge only if the variable or its negation is contained in the clause.

A satisfying assignment is found via an iterative message-passing system in which each variable node passes to each neighboring clause the probability that its other neighboring clauses will be satisfied when it takes the value 0 or the value 1. The clause nodes, in turn, tell the variables the probabilities, in each case, that they will be satisfied given the distributions of values of their other neighboring nodes. In updating their probabilities, nodes assume that their incoming messages are statistically independent.

Convergence of this process is guaranteed only for trees, i.e., factor graphs that have no cycles, but the algorithm often works on general graphs, if they are locally tree-like. Once the probabilities converge, the algorithm chooses the most biased variable and sets it to its favored value. Continuing in this way, the algorithm eventually solves the problem.

Belief propagation successfully solves instances of random k-SAT with ratios well below the threshold, where the solution space is still connected. In that region, partial solutions for a subset of the clauses can typically be extended into full solutions. Near the fragmentation point and past it, however, the method breaks down. The problem, physicists realized, is that the partial solutions to which the local computations converge may belong to different clusters (so that a variable is frozen to two different values) or to no solution cluster at all.

Working from their geometric picture, Mézard, Parisi, and Zecchina were able to devise a method that seems to get around this problem. They theorized that specifying which variables are frozen in each cluster provides a sufficiently accurate description of the cluster, enabling them to aggregate variable preferences over the cluster descriptions.

Specifically, survey propagation, like belief propagation, works on the factor graph of the input formula, but with variables now allowed to take values of 0, 1, or *, a joker state indicating that the variable is unconstrained. Variables send messages saying how much they are constrained by other clauses, and clauses send messages to variables telling them whether they must take a particular value to satisfy the clause. Once this process converges, the most constrained variable is set to its favored value and the problem is reduced to a smaller one. Eventually, only unconstrained variables are left, and a solution can be found via a simpler algorithm.

A Budding Collaboration

When the paper describing survey propagation appeared, mathematicians and computer scientists were simultaneously fascinated and flummoxed. The paper was filled with references to unfamiliar physical models and methods, and it contained no proofs. Indeed, over the past year, two groups of researchers wrote papers whose goal was simply to give a more mathematical description of the algorithm, though still without proofs of the limits of its performance.

During their initial interactions with computer scientists and mathematicians, Mézard, Parisi, and Zecchina learned of the new lower bound for the threshold and the probabilistic method used to derive it. Recently, Mézard and Zecchina, working with Thierry Mora, a student at Orsay, were able to use the weighted second moment method of Achlioptas and Peres to count pairs of satisfying assignments at each distance. In this way, they demonstrated that, nearing the threshold, pairs must be either close together (in the same cluster) or far apart.

With this geometric picture, says Achlioptas, may come a deeper understanding of why k-SAT is hard—the goal of computer scientists who have been studying the phase transition. "The interesting story is the changing geometry of the solution space and its effect on algorithms, not the satisfiability threshold, although that was the original focus," he says.

It was because the interaction with mathematicians was proving so fruitful, Mézard says, that he decided to spend three months at MSRI to take part in a semester-long program in statistical physics. "My purpose here is to explain my work and understand theirs," he says. In addition to making progress on understanding k-SAT, Mézard says that he learned about other interesting problems in computer science to which he thinks he could apply methods from physics.

For Achlioptas, too, the interaction has proved fruitful. At a recent MSRI workshop, he described work, with Federico Ricci-Tersenghi, a physicist at the University of Rome, proving that solution clusters indeed have frozen variables, and that the fraction of such variables tends to one as the ratio r approaches the satisfiability threshold. Their proof relies heavily on the approximate description of clusters introduced by Mézard, Parisi, and Zecchina in their paper on the survey propagation algorithm.

Despite the recent leaps in understanding, the exact behavior of the random k-SAT model is still a mystery. It remains to be seen, for example, whether the success of survey propagation persists for large k. "Proving that some algorithm works past the fragmentation point for all k would be big news," says Achlioptas.

Nonetheless, mathematicians and physicists believe that by trading ideas, they will continue to make progress. "The cultural differences have created a nice, creative tension," Moore says, pointing out that certain ways of thinking are deeply ingrained in the intellectual landscape of each discipline. "Computer scientists have traditionally thought in terms of worst-case complexity as defined by an adversary. In physics, there are no adversaries; problems come from nature, and they have simpler solutions."

Sara Robinson is a freelance writer based in Los Angeles, California.