Phase Transition in 3SAT

Yi Zhou

Phase Transition in 3SAT

Fine Grained Complexity Analysis

Outline

Phase Transition

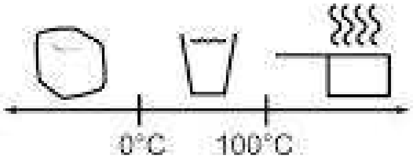

Figure: Phase Transition of H_2O

Phase Transition

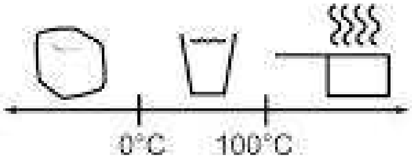

Figure: Phase Transition of H_2O

Sudden sharp transformation from one state to another at a certain point.

SAT & 3SAT

The satisfiability problem of propositional formulas, i.e., to determine whether there exists an interpretation satisfying a given propositional formula.

SAT & 3SAT

The satisfiability problem of propositional formulas, i.e., to determine whether there exists an interpretation satisfying a given propositional formula.

SAT can be linearly transformed to its subform 3SAT, where the propositional formula is a conjunction of clauses with no more than 3 literals.

SAT & 3SAT

The satisfiability problem of propositional formulas, i.e., to determine whether there exists an interpretation satisfying a given propositional formula.

SAT can be linearly transformed to its subform 3SAT, where the propositional formula is a conjunction of clauses with no more than 3 literals.

Example

The following formula is in 3SAT

$$(a \vee b \vee c) \wedge (a \vee b \vee \neg c) \wedge (a \vee \neg b \vee \neg c)$$

3SAT: an Important Problem

SAT/3SAT is (one of) the most important

- NP-complete problem
- constraint satisfaction problem
- combinatorial problem
- logic solving problem
- knowledge representation formalism

SAT/3SAT: Many Applications

SAT/3SAT has many applications in

- computational complexity
- computational learning theory
- hardware/software verification
- automatic test pattern generation
- AI planning
- theorem proving
- logic-based problem solving
- combinatorial search
- bioinformatics
- ......

Random 3SAT

Each clause is randomly generalized by the uniform distribution according to the clause/variable ratio r.

Random 3SAT

Each clause is randomly generalized by the uniform distribution according to the clause/variable ratio r.

Random 3SAT is important to understand SAT solving both in theory and in practice. In fact, it is one of the SAT competition category.

When Phase Transition Meets SAT/3SAT: the Observation

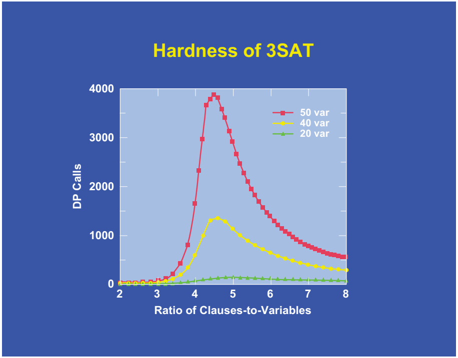

Figure: Hardness to solve 3SAT problems

When Phase Transition Meets SAT/3SAT: the Observation

The ”Easy-Hard-Easy” phenomenon

- Formulas with a low clause/variable ratio can easily be solved. Most likely satisfiable
- Formulas with a high clause/variable ratio can easily be solved. It varies.
- Formulas with a middle clause/variable ratio are hard to solve. Most likely satisfiable

When Phase Transition Meets SAT/3SAT: the Conjecture

Random 3SAT does embrace a phase transition phenomenon!

When Phase Transition Meets SAT/3SAT: the Conjecture

Random 3SAT does embrace a phase transition phenomenon!

There exists a real number r such that

- Almost all big 3SAT instances with a clause variable ratio less than r are satisfiable.
- Almost all big 3SAT instances with a clause variable ratio greater than r are unsatisfiable.

When Phase Transition Meets SAT/3SAT: the State-of-the-art from the Empirical Side

Empirical study supports the claim.

When Phase Transition Meets SAT/3SAT: the State-of-the-art from the Empirical Side

Empirical study supports the claim.

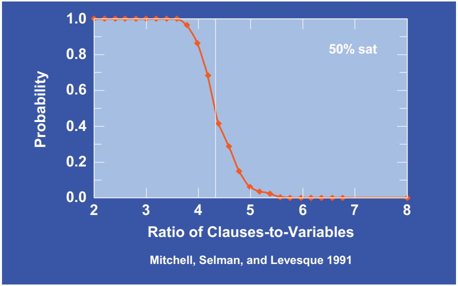

Figure: The probability of satisfying random 3SAT instances

When Phase Transition Meets SAT/3SAT: the State-of-the-art from the Empirical Side

Empirical study supports the claim.

It is concluded from the empirical studies that the claim is true. And the phase transition point is believed to be around 4.27 according to in statistical physics, more precisely, replica methods.

When Phase Transition Meets SAT/3SAT: the State-of-the-art from the Theoretical Side

2SAT does embrace the phase transition phenomenon with the phase transition point to be 1, by using implication graph and branching process in random graph theory, originally developed Erdos and Renyi.

When Phase Transition Meets SAT/3SAT: the State-of-the-art from the Theoretical Side

2SAT does embrace the phase transition phenomenon with the phase transition point to be 1, by using implication graph and branching process in random graph theory, originally developed Erdos and Renyi.

The 3SAT phase transition problem remains open .

When Phase Transition Meets SAT/3SAT: the State-of-the-art from the Theoretical Side

2SAT does embrace the phase transition phenomenon with the phase transition point to be 1, by using implication graph and branching process in random graph theory, originally developed Erdos and Renyi.

The 3SAT phase transition problem remains open for a long time.

When Phase Transition Meets SAT/3SAT: the State-of-the-art from the Theoretical Side

2SAT does embrace the phase transition phenomenon with the phase transition point to be 1, by using implication graph and branching process in random graph theory, originally developed Erdos and Renyi.

The 3SAT phase transition problem remains open for a long time.

Researchers are trying to find lower bound and upper bound instead, and the gap gradually thins.

Phase Transition in 3SAT: Upper Bound

For upper bound

- $r = 5.1909$ (1983) Franco, Paull (and others)
- $r = 5.19 - 10^{-7}$ (1992) Frieze and Suen
- $r = 4.758$ (1994) Kamath, Motwani, Palem, Spirakis
- $r = 4.667$ (1996) Kirousis, Kranakis, Krizanc
- $r = 4.642$ (1996) Dubois, Boufkhad
- $r = 4.602$ (1998) Kirousis, Kranakis, Krizac, Stamatiou
- $r = 4.596$ (1999) Janson, Stamatiou, Vamvakari (1999)
- $r = 4.571$ (2007) Kaporis, Kirousis, Stamatiou, Vamvakari
- $r = 4.506$ (1999) Dubois, Boukhand, Mandler
- $r = 4.49$ (2008) Diaz, Kirousis, Mitsche, Perez
- $r = 4.453$ (2008) Maneva, Sinclair

Phase Transition in 3SAT: Lower Bound

For lower bound

- $r = 2.66$ (1986) Chao, Franco
- $r = 2.99$ (1986) Chao, Franco
- $r = 3.003$ (1992) Frieze, Suen
- $r = 3.145$ (2000) Achlioptas
- $r = 3.26$ (2001) Achlioptas and Sorkin
- $r = 3.42$ (2002) Kaporis, Kirousis, Lalas
- $r = 3.52$ Kaporis, Kirousis, Lalas (2003)
- $r = 3.52$ Hajiaghayi, Sorkin (2003)

Phase Transition in 3SAT: an Important Open Problem

It is believed that the approaches for showing the upper and lower bounds cannot prove the ultimate claim. However, any slight improvement is highly technical, tedious and important.

Phase Transition in 3SAT: an Important Open Problem

It is believed that the approaches for showing the upper and lower bounds cannot prove the ultimate claim. However, any slight improvement is highly technical, tedious and important.

The phase transition problem in 3SAT still remains open.

Phase Transition for other SAT Classes

The phase transition phenomenon also exists for other subclasses of SAT.

Phase Transition for other SAT Classes

The phase transition phenomenon also exists for other subclasses of SAT.

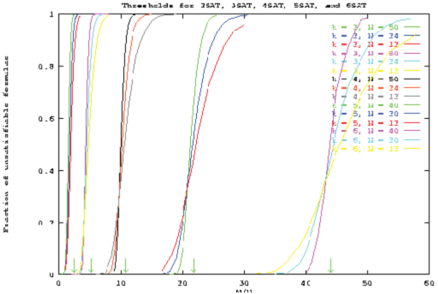

Figure: The probability of satisfying random *k*-SAT instances

Phase Transition for other SAT Classes

The phase transition phenomenon also exists for other subclasses of SAT.

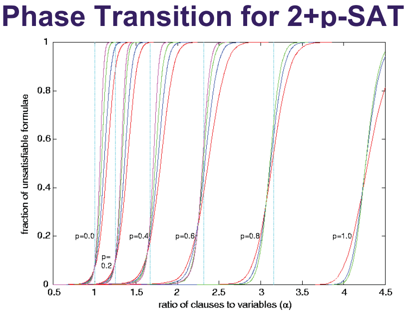

Figure: The probability of satisfying random $2 + p$-SAT instances

Phase Transition 3SAT: Something More

Papers published in *Nature* and *Science*.

Phase Transition 3SAT: Something More

Papers published in *Nature* and *Science*.

From empirical studies, many NP problems embrace the phase transition phenomenon.

Phase Transition 3SAT: Something More

Papers published in *Nature* and *Science*.

From empirical studies, many NP problems embrace the phase transition phenomenon.

The phase transition problem in k-SAT is the key points in Vinay Deolalikar's wrong proof of P$\neq$NP.

Phase Transition in 3SAT: the New Conjecture

Phase transition does exist for random 3SAT and the transition point is $\sqrt{10} + 1 \approx 4.16$.

Phase Transition in 3SAT: the New Conjecture

Phase transition does exist for random 3SAT and the transition point is $\sqrt{10} + 1 \approx 4.16$.

This is less than the conjectured point 4.27 obtained by applying the replica method for empirical studies. But it does make sense as proving unsatisfiability is always much harder than finding a satisfiable interpretation. The former needs to explore the whole search tree while the latter only needs find one solution, possibly by chance.

Outline

Rethink the Hardness for Random 3SAT Instances

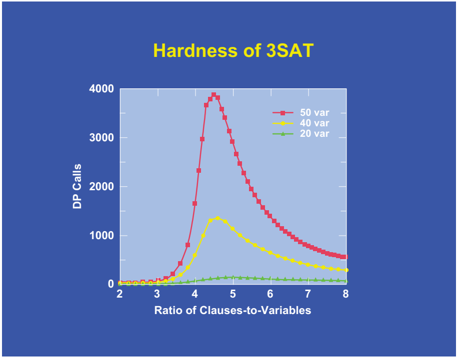

Figure: Hardness to solve 3SAT problems

Rethink the Hardness for Random 3SAT Instances

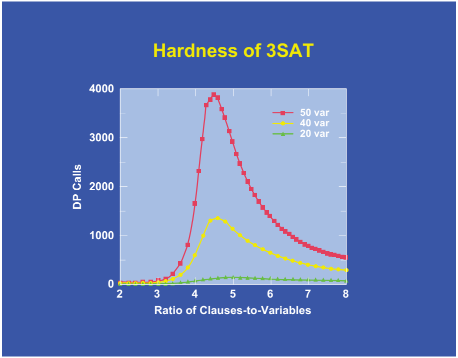

Figure: Hardness to solve 3SAT problems

The sizes of instances are increasing!!!

Rethink the Hardness for Random 3SAT Instances

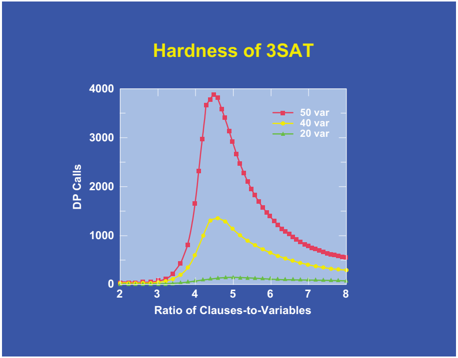

Figure: Hardness to solve 3SAT problems

So what?

Worst-Case Complexity Analysis

In computer science, we traditionally use worst-case complexity analysis. And we use the big O notation to analyze the efficiency of an algorithm, which is represented as a function of the input size.

Worst-Case Complexity Analysis

In computer science, we traditionally use worst-case complexity analysis. And we use the big O notation to analyze the efficiency of an algorithm, which is represented as a function of the input size.

Bubble sort has worst-case time complexity to be $O(n^2)$, while quick sort has worst-case time complexity to be $O(n \log n)$.

Worst-Case Complexity Analysis

In computer science, we traditionally use worst-case complexity analysis. And we use the big O notation to analyze the efficiency of an algorithm, which is represented as a function of the input size.

Bubble sort has worst-case time complexity to be $O(n^2)$, while quick sort has worst-case time complexity to be $O(n \log n)$.

Average-case analysis and smooth analysis are also introduced. However, both of them also take the input size as the main factor.

The Problem

Fixing the number of variables, when the clause/variable ratio is becoming bigger, so is the formula size.

The Problem

Fixing the number of variables, when the clause/variable ratio is becoming bigger, so is the formula size.

However, it is NOT necessary that bigger formulas are always harder to solve. Also, sometimes an SAT instance with 1,000,000 variables can be solved quickly but the same solver will stuck with some instances with 100 variables.

Observation and Motivation

For random 3 SAT instances, the size of the input instance is not the major factor, the inherit complicatedness of the instance is.

Observation and Motivation

For random 3 SAT instances, the size of the input instance is not the major factor, the inherit complicatedness of the instance is.

This motivates us to consider fine-grained complexity analysis for algorithms with respect to particular instance.

How?

How to analyze

How?

How to analyze

Kolmogorov complexity may shed some insights

Kolmogorov Complexity

To characterize the complexity of a string by some language, which is defined as the minimal number of objects in the language to describe it.

Kolmogorov Complexity

To characterize the complexity of a string by some language, which is defined as the minimal number of objects in the language to describe it.

Example

For the string *ababababababababababababababab*, although it has 30 characters. In English, we can describe it by 15 "ab"s. But the string $4c1j5b2p0cv4w1x8rx2y39umgw5q85s$ seems to have no short description.