

Empirical Evaluation of SAT Provers

White Group

Integrated Logic Systems
Lab #2 - Spring 2015

Background

- Complexity theory tells us that SAT is a “hard” problem. But what if we asked about the average case complexity of SAT instead of the worst-case complexity?
- In early 80’s some researchers claimed that on average SAT is “easy” based on the fact that their solvers performed well on randomly generated formulae.
- In 1992 Mitchell, Selman and Levesque showed that the method used to generate random formulae affected the results of these experiments.
- Here we replicate their results.

In particular, they showed that results differ drastically for:

- *fixed clause-length model (FCM)*: fix the number of clauses and the number of variables, then for each clause randomly generate literals so that each literal is chosen with equal probability;
- *constant-probability model (CPM)*: clause length is not fixed, clauses are generated by including a variable in a clause with some constant probability.

Their claim: using FCM one can generate difficult cases, whereas with CPM one leaves them out.

Our experiment

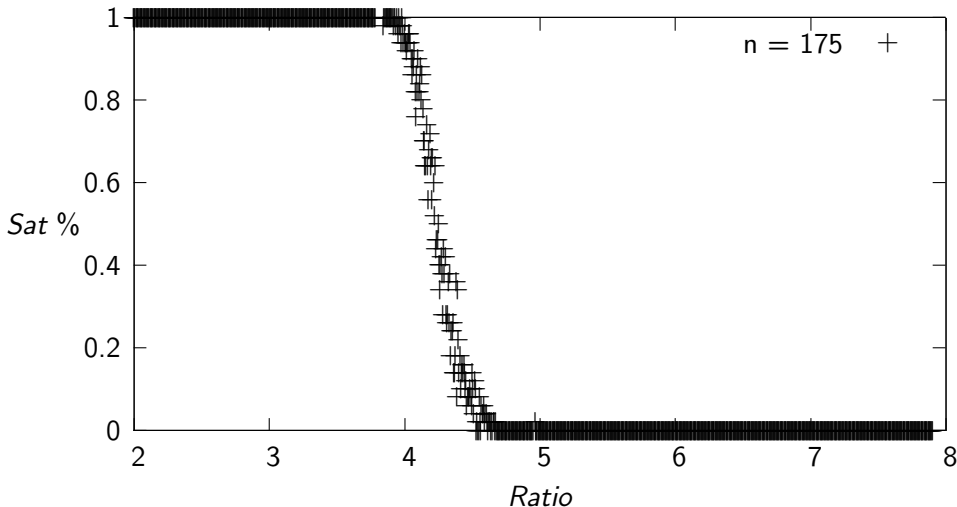
- We randomly generated FCM formulae of clause length 3 with 75, 100 and 175 variables and CPM formulae with 75 and 175 variables and average clause length 3.
- For each clauses-to-variables ratio we ran Z3 SAT solver 50 times and collected the following statistics:
 - Average time in seconds
 - Percentage of satisfiable formulae
- Our results confirm the claim made in the original paper.
- Worry we have: in order to compare FCM and CPM one has to consider formulae of average clause length 3. Implementation of the averaging might influence the results.

Result: Sharp phase transition

First, we observed the sharp phase transition phenomenon: a fast change from satisfiable to unsatisfiable 3-SAT problems happens around the ratio 4.25.

Result: Sharp phase transition, 1

Sharp phase transition, 175 variables

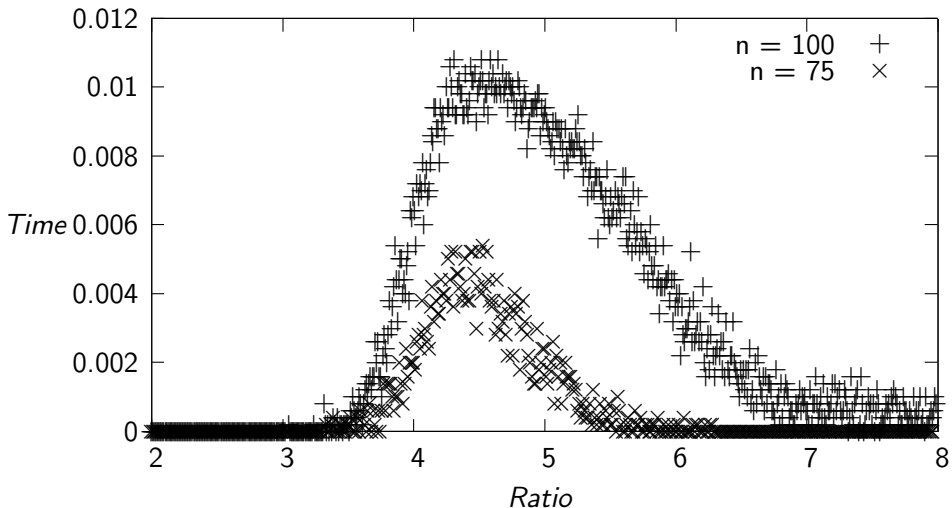


Result: Sharp phase transition, 2

Second: it takes more time to solve instances in the transition region.

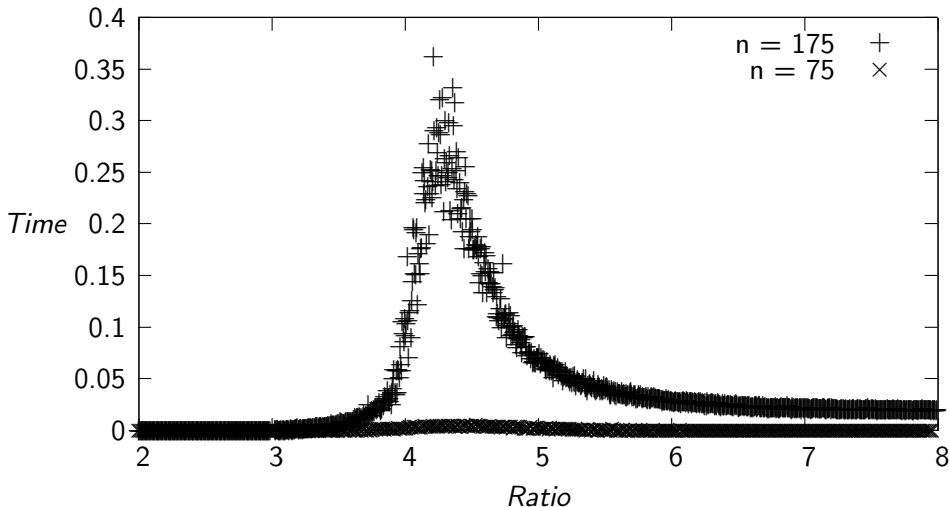
Result: Sharp phase transition,3

Sharp phase transition for formulae with 75 and 100 variables



Result: Sharp phase transition, 4

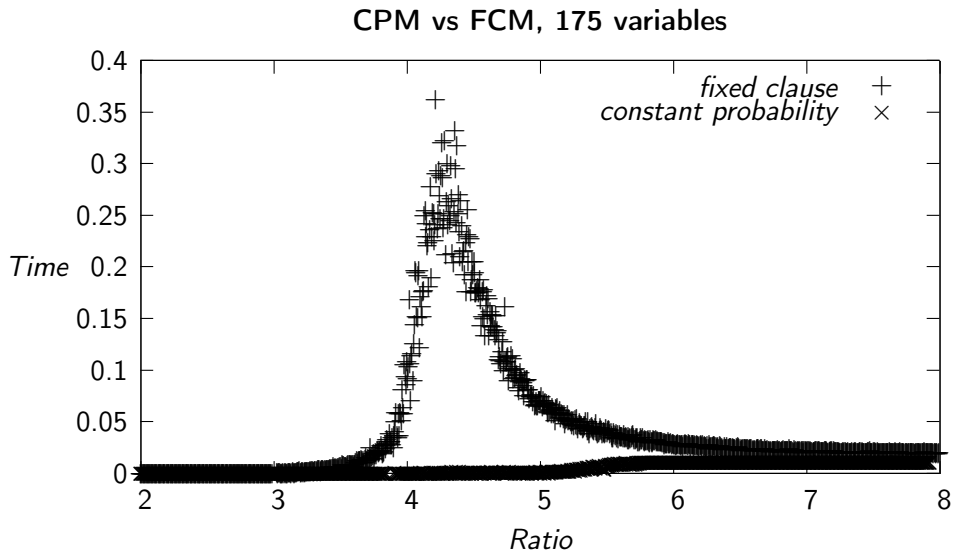
Sharp phase transition for formulae with 75 and 175 variables



Result: FCM vs CPM

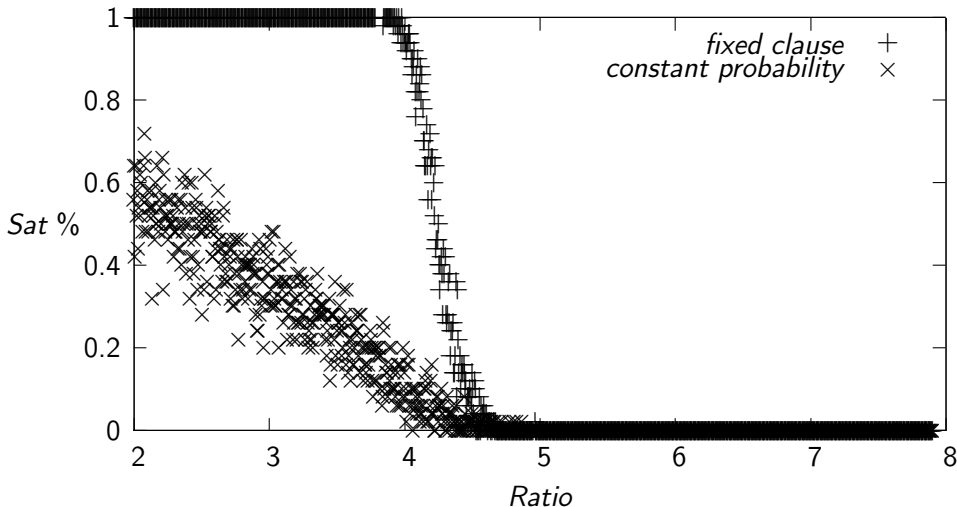
Moreover, we observed the difference between fixed clause and constant probability model.

Result: FCM vs CPM, 1



Result: FCM vs CPM, 2

Satisfiability probability for CPM and FCM, 175 variables



Take home message:

- Easy and hard SAT cases are not evenly distributed
- This should be taken into account when testing SAT solvers on random formulae in order to avoid incorrect results

Thank you!