

Class 6: SAT Solving

Schedule

Problem Set 2 is due **Friday at 6:29pm**. A few of the problems on PS2 are quite challenging! The problems are organized by topic, not by difficulty (the hardest problems are #4 and #7). I hope everyone makes a good attempt to solve all of the problems, but don't be overly stressed if you cannot solve these ones and move on to the other problems if you are stuck on them (to hopefully return to these problems after you've had more time to think about them). Good solutions to these problems will impress us, but are not necessary to get a good score on PS2.

I will be out-of-town Monday, and not able to hold my usual office hours on Monday. I will have "make-up" office hours on Wednesday, 3:30-4:30pm (in addition to my usual Wednesday 1-2pm office hours).

Notes and Questions

Definition. A formula is in 3SAT if it is in 3CNF form and it is satisfiable.

$$(x_1 \vee x_2 \vee x_3) \wedge (x_1 \vee \bar{x}_2 \vee x_3) \wedge (\bar{x}_1 \vee x_2 \vee \bar{x}_3)$$

$$\begin{aligned} & (x_{48} \vee x_4 \vee \bar{x}_9) \wedge (\bar{x}_{44} \vee x_{50} \vee \bar{x}_{37}) \wedge (\bar{x}_8 \vee \bar{x}_1 \vee x_{28}) \wedge (x_{21} \vee x_{27} \vee \bar{x}_{32}) \wedge (x_{17} \vee x_{29} \vee \bar{x}_{30}) \wedge (x_{30} \vee x_{24} \vee x_{37}) \wedge (\bar{x}_{22} \vee \bar{x}_{27} \vee \bar{x}_{44}) \wedge (x_8 \vee \bar{x}_{25} \vee \bar{x}_{24}) \wedge (\bar{x}_{44} \vee x_{50} \vee x_{14}) \wedge (x_{45} \vee x_{15} \vee x_{37}) \wedge (\bar{x}_{16} \vee \\ & x_{14} \vee \bar{x}_{36}) \wedge (\bar{x}_{33} \vee x_5 \vee x_{26}) \wedge (x_{18} \vee \bar{x}_7 \vee \bar{x}_{24}) \wedge (x_{31} \vee x_{38} \vee x_{28}) \wedge (x_{31} \vee \bar{x}_{33} \vee \bar{x}_8) \wedge (x_{49} \vee x_7 \vee \bar{x}_6) \wedge (x_{34} \vee \bar{x}_8 \vee x_{46}) \wedge (x_4 \vee \bar{x}_5 \vee \bar{x}_{35}) \wedge (x_{43} \vee x_{27} \vee x_{39}) \wedge (\bar{x}_{46} \vee \bar{x}_{40} \vee \bar{x}_{27}) \wedge (\bar{x}_{25} \vee x_{14} \vee \\ & \bar{x}_{49}) \wedge (x_{38} \vee x_5 \vee x_{15}) \wedge (x_9 \vee x_{14} \vee \bar{x}_{19}) \wedge (x_{45} \vee \bar{x}_{42} \vee \bar{x}_{39}) \wedge (x_{34} \vee \bar{x}_{22} \vee \bar{x}_{28}) \wedge (\bar{x}_{20} \vee x_{15} \vee \bar{x}_8) \wedge (\bar{x}_{44} \vee \bar{x}_{10} \vee \bar{x}_9) \wedge (x_{22} \vee \bar{x}_{31} \vee x_{14}) \wedge (\bar{x}_9 \vee \bar{x}_{42} \vee \bar{x}_{15}) \wedge (\bar{x}_{40} \vee x_{12} \vee \bar{x}_{32}) \wedge (\bar{x}_{20} \vee \bar{x}_6 \vee \bar{x}_{15}) \wedge \\ & (\bar{x}_{37} \vee x_{39} \vee \bar{x}_{23}) \wedge (\bar{x}_3 \vee \bar{x}_{40} \vee \bar{x}_{32}) \wedge (\bar{x}_4 \vee \bar{x}_{25} \vee x_7) \wedge (\bar{x}_{20} \vee \bar{x}_{36} \vee \bar{x}_{37}) \wedge (\bar{x}_{40} \vee \bar{x}_{35} \vee x_{39}) \wedge (\bar{x}_{43} \vee \bar{x}_{40} \vee \bar{x}_7) \wedge (x_{34} \vee x_{44} \vee x_{26}) \wedge (x_{13} \vee x_{27} \vee x_{28}) \wedge (x_{12} \vee \bar{x}_{36} \vee x_7) \wedge (\bar{x}_{16} \vee x_9 \vee \bar{x}_{24}) \wedge (\bar{x}_{48} \vee \\ & x_{14} \vee x_{28}) \wedge (x_{16} \vee x_4 \vee x_{40}) \wedge (\bar{x}_{25} \vee x_{15} \vee x_{37}) \wedge (x_{47} \vee \bar{x}_{26} \vee \bar{x}_{23}) \wedge (x_4 \vee \bar{x}_{13} \vee x_{36}) \wedge (x_{48} \vee \bar{x}_{13} \vee \bar{x}_{37}) \wedge (x_4 \vee x_{35} \vee \bar{x}_{27}) \wedge (\bar{x}_{22} \vee x_{47} \vee x_{26}) \wedge (\bar{x}_{22} \vee \bar{x}_{46} \vee x_{27}) \wedge (\bar{x}_{20} \vee x_{49} \vee x_{11}) \wedge (x_{42} \vee \bar{x}_{10} \vee \\ & x_{28}) \wedge (\bar{x}_{45} \vee x_{28} \vee \bar{x}_{37}) \wedge (x_{14} \vee \bar{x}_{32} \vee \bar{x}_{23}) \wedge (x_{22} \vee x_{14} \vee x_{23}) \wedge (\bar{x}_{17} \vee \bar{x}_{46} \vee \bar{x}_7) \wedge (\bar{x}_{31} \vee x_{46} \vee \bar{x}_{50}) \wedge (x_{34} \vee \bar{x}_{41} \vee x_{43}) \wedge (x_{17} \vee \bar{x}_9 \vee x_{15}) \wedge (x_{46} \vee x_{14} \vee \bar{x}_{12}) \wedge (\bar{x}_{20} \vee x_{12} \vee x_{14}) \wedge (x_{41} \vee x_{42} \vee \\ & \bar{x}_{15}) \wedge (x_{48} \vee x_{46} \vee \bar{x}_{36}) \wedge (\bar{x}_{22} \vee \bar{x}_4 \vee \bar{x}_{49}) \wedge (x_{22} \vee x_{12} \vee \bar{x}_{42}) \wedge (x_{13} \vee x_{38} \vee x_{39}) \wedge (x_{48} \vee \bar{x}_{16} \vee \bar{x}_{27}) \wedge (x_{17} \vee \bar{x}_{18} \vee \bar{x}_{26}) \wedge (x_{48} \vee \bar{x}_{40} \vee \bar{x}_{35}) \wedge (\bar{x}_{43} \vee \bar{x}_{40} \vee \bar{x}_{49}) \wedge (x_{29} \vee x_{11} \vee \bar{x}_{32}) \wedge (x_{33} \vee \bar{x}_{17} \vee \\ & x_{39}) \wedge (\bar{x}_{25} \vee \bar{x}_9 \vee \bar{x}_6) \wedge (x_{40} \vee \bar{x}_{50} \vee x_{19}) \wedge (x_8 \vee x_{10} \vee \bar{x}_{27}) \wedge (x_5 \vee x_9 \vee \bar{x}_{26}) \wedge (x_{45} \vee \bar{x}_{38} \vee \bar{x}_{27}) \wedge (\bar{x}_4 \vee \bar{x}_{40} \vee \bar{x}_{42}) \wedge (x_{21} \vee x_{50} \vee x_{12}) \wedge (\bar{x}_8 \vee \bar{x}_{14} \vee \bar{x}_{42}) \wedge (\bar{x}_{17} \vee x_{47} \vee \bar{x}_{27}) \wedge (x_{49} \vee \bar{x}_{12} \vee \bar{x}_6) \wedge \\ & (x_{27} \vee x_{49} \vee \bar{x}_{32}) \wedge (\bar{x}_{29} \vee \bar{x}_{12} \vee \bar{x}_{26}) \wedge (x_{48} \vee \bar{x}_2 \vee x_6) \wedge (x_{16} \vee x_{36} \vee x_{49}) \wedge (x_{33} \vee \bar{x}_{12} \vee \bar{x}_{26}) \wedge (\bar{x}_{33} \vee x_{29} \vee x_{49}) \wedge (\bar{x}_{48} \vee x_2 \vee x_{19}) \wedge (x_{25} \vee x_{36} \vee x_{49}) \wedge (x_{21} \vee x_{40} \vee \bar{x}_{14}) \wedge (\bar{x}_{34} \vee \bar{x}_{44} \vee \bar{x}_6) \wedge (x_{48} \vee \\ & \bar{x}_{50} \vee \bar{x}_1) \wedge (x_5 \vee \bar{x}_{12} \vee x_7) \wedge (x_{21} \vee \bar{x}_{35} \vee \bar{x}_{27}) \wedge (\bar{x}_{22} \vee \bar{x}_{16} \vee \bar{x}_{14}) \wedge (\bar{x}_{13} \vee \bar{x}_{35} \vee \bar{x}_{12}) \wedge (\bar{x}_4 \vee \bar{x}_{35} \vee \bar{x}_{42}) \wedge (\bar{x}_{50} \vee \bar{x}_{40} \vee x_7) \wedge (x_{25} \vee x_{47} \vee \bar{x}_{12}) \end{aligned}$$

Converting Truth Tables to CNF

| P | Q | $P \implies Q$ | $P \oplus Q$ |
|----------|----------|----------------|--------------|
| T | T | T | F |
| T | F | F | T |
| F | T | T | T |
| F | F | T | F |

The output of the operator is **T** if and only if the inputs do not match *any row* where the output is **F**. We can ensure the inputs do not match a row, but OR-ing the negation of each input: in the disjunction, at least one must be **T** to satisfy the clause.

Problems

Definition. A *problem* is a precise description of set of possible inputs and desired property of an output corresponding to each input.

Define the *ADDITION* problem (adding two integers):

Definition. A *decision problem* is a problem where the output is either T or F . Equivalently, we can view a decision problem as testing set membership: $x \in S$.

The *SUM* problem:

Input. Three integers, x , y , and z .

Output. T iff $z = x + y$.

How could we solve *ADDITION* using *SUM*?

Definition. A *procedure* is a precise description of an information process.

Definition. An *algorithm* for a particular problem is a procedure that *solves* that problem. To *solve* a problem, an algorithm must always (eventually) produce the correct output for any problem input.

Definition. A *program* is a description of a procedure that can be executed by a computer.

Definition. The *3SAT decision problem* takes as input a logical formula written in CNF, and outputs T if the input formula is satisfiable and outputs F otherwise.

How many uses of a solver for the *3SAT* decision problem are sufficient to always find a satisfying assignment for a satisfiable formula?

SAT Solving

Sahand Saba's *Understanding SAT by Implementing a Simple SAT Solver in Python* (Code with a few modifications: <https://github.com/evansuva/simple-sat>)

What is the *worst-case* number of guesses needed to solve a SAT instance with v variables and c clauses using the “brute-force guessing” strategy?

What is the *worst-case* number of guesses needed to solve a SAT instance with v variables and c clauses using the “watchlists” strategy?

Why is it possible to solve *most* SAT instances reasonably quickly?