Algorithms: SAT and 3-SAT

Model 1: SAT

Each variable x_i represents a Boolean value (T or F). An *assignment* consists in specifying a T or F value for each variable. Recall that $\wedge$ denotes logical AND, and $\vee$ denotes logical OR. $\overline{x}$ denotes logical negation. For example, $x_1 \wedge \overline{x_3}$ means "x_1 and not x_3".

	Examples	Non-examples
Term	x_1 x_3 $\overline{x_3}$	$x_1 \wedge x_2$ $x_2 \vee x_4$
Clause	x_1 $x_1 \vee x_2$ $\overline{x_1} \vee x_3 \vee x_2$ $\overline{x_3}$ $\overline{x_5} \vee x_3 \vee \overline{x_7} \vee x_9$	$x_1 \wedge x_2$ $x_2 \vee \overline{x_3} \wedge x_4$ $x_1 \Rightarrow x_5$
CNF formula	x_1 $x_2 \vee x_5 \vee \overline{x_1}$ $(x_1 \vee x_2) \wedge (\overline{x_3} \vee x_5 \vee x_9 \vee x_8 \vee x_2)$ $x_4 \wedge (x_1 \vee x_2) \wedge \overline{x_5} \wedge (x_1 \vee \overline{x_3} \vee x_5)$ $x_1 \wedge (\overline{x_1} \vee x_3) \wedge (\overline{x_1} \vee \overline{x_3})$ $x_1 \wedge \overline{x_3}$	$(x_1 \wedge x_3) \vee (x_2 \wedge x_5)$ $x_1 \wedge (x_2 \vee (x_3 \wedge (x_4 \vee (x_5 \wedge x_6))))$
3-CNF formula	$x_2 \vee x_5 \vee \overline{x_1}$ $(x_1 \vee x_2 \vee x_3) \wedge (\overline{x_2} \vee x_3 \vee \overline{x_1})$ $(x_4 \vee \overline{x_2} \vee x_3) \wedge (x_2 \vee \overline{x_3} \vee x_1) \wedge (x_2 \vee x_2 \vee x_3)$	x_1 $x_1 \vee x_2$ $(x_4 \vee \overline{x_2} \vee x_3) \wedge (x_2 \vee \overline{x_3}) \wedge (x_2 \vee x_2 \vee x_3)$

(Page intentionally left blank.)

1 Based on the examples and non-examples of *terms* in the first row of the chart, write down a definition of a *term*.

2 Based on the examples and non-examples, write a definition of a *clause*. Be sure to use the word *term* in your definition.

3 Again based on the model, write a definition of a *CNF formula*. Be sure to use the word *clause*.

CNF stands for *conjunctive normal form*.

4 Consider the assignment setting each x_i to T when i is even, and F when i is odd. For each CNF formula in the left-hand column, say whether it evaluates to T or F (*i.e.* whether it is *satisfied*) under this assignment using the usual rules of Boolean logic.

5 Find an assignment that makes $x_1 \land (\overline{x_2} \lor \overline{x_3} \lor \overline{x_1}) \land (x_2 \lor \overline{x_1})$ true. (You only need to specify values for x_1, x_2, and x_3.) Such an assignment is called a *satisfying assignment*.

6 Does every *clause* have some assignment which makes it true (*i.e.* a

satisfying assignment)? If so, explain why; if not, give a counterexample.

7 Does every *CNF formula* have a satisfying assignment? If so, explain why; if not, give a counterexample.

8 Based on your previous answer, state an interesting decision problem about CNF formulas.

This is a famous decision problem called SAT. If we restrict every *clause* to have *exactly three* terms—as in the *3-CNF formulas* shown in the model—the corresponding decision problem is known as 3-SAT.

9 Explain why 3-SAT $\leq_P$ SAT.

It turns out that SAT $\leq_P$ 3-SAT as well, although this is extremely nonobvious! In fact, 3 is the smallest k for which SAT $\leq_P$ k-SAT. 1-SAT is trivial and 2-SAT can be solved in linear time by a very clever application of DFS.

It turns out that we can reduce 3-SAT to another problem we have studied before:

Theorem 1. 3-SAT $\leq_P$ INDEPENDENT-SET.

Let's prove it!

Recall that the INDEPENDENT-SET problem takes as input a graph G and a natural number k, and outputs whether there is an independent set in G of size k or greater.

10 In order to show 3-SAT $\leq_P$ INDEPENDENT-SET, we need to assume that we have a black box to solve ____________, and show how we can use it to construct a solution to __________________.

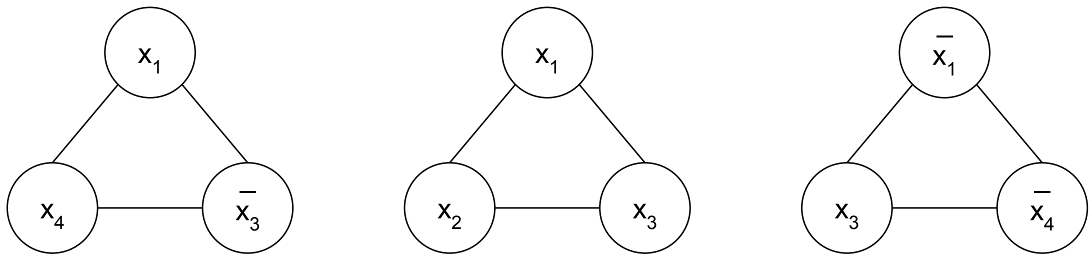
11 Draw a picture of the situation using nested boxes. What are the inputs and outputs?

12 Fill in this statement based on your picture: given a ____________, we have to construct a ____________________________________ and pick a ____________________________________ such that the newly constructed ____________________________________ has ____ if and only if the original input ____________________________.

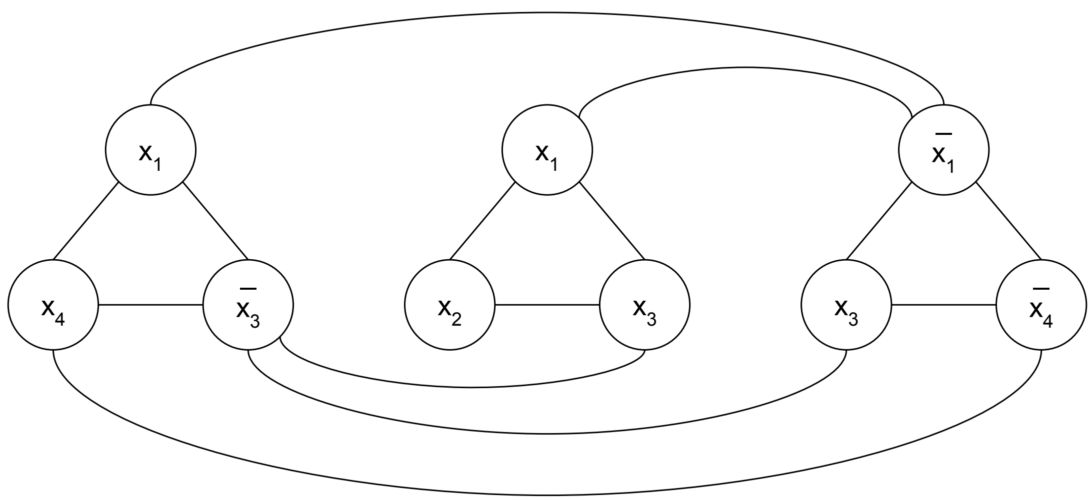
Model 2: 3-SAT $\leq_P$ INDEPENDENT-SET

$$F = (x_1 \lor \overline{x_3} \lor x_4) \land (x_2 \lor x_1 \lor x_3) \land (\overline{x_1} \lor x_3 \lor \overline{x_4})$$

Graph A:



Graph B:



Let's first consider formula F and graph A.

13 What is the relationship of formula F to graph A?

14 How many clauses does F have? How many vertices does A have?

15 What is the size of a maximum independent set in graph A?

16 In general, instead of formula F, suppose we started with a 3-CNF formula F' having k clauses. How would we make the corresponding graph A'?

17 How many vertices would A' have?

18 What would be the size of a maximum independent set in A'?

Now consider graph B.

19 How is graph B related to graph A?

20 In general, if we started with some 3-CNF formula F' and made it into a graph A' of triangles, what do you think we would add to turn it into a corresponding graph B'?

21 Find a maximal independent set in B.

22 Explain how you can use your independent set to find a satisfying assignment for F.

23 Explain why in a satisfying assignment, at least one term must be true in each clause.

24 Explain how you could use any satisfying assignment for F to find an independent set of size 3 in B.

25 Would the previous arguments still work if we used graph A instead of graph B? In other words, what is the importance of the edges added in graph B?