NP-Completeness of 3SAT, 1-IN-3SAT and MAX 2SAT

3SAT

The 3SAT problem is the following.

INSTANCE : Given a boolean expression E in conjunctive normal form (CNF) that is the conjunction of clauses, each of which is the disjunction of three distinct literals.
QUESTION : Is E satisfiable ?

The CSAT problem is the following.

INSTANCE : Given a boolean expression E in CNF.
QUESTION : Is E satisfiable; that is, is there a truth assignment to the variables of E so that each clause in E becomes true ?

The Cook-Levin theorem asserts that SATISFIABILITY is NP-complete. Although 3-CNF expressions are a subset of the CNF expressions, they are complex enough in the sense that testing for satisfiability turns out to be NP-complete.

Theorem : 3SAT is NP-complete.

Proof **:** Evidently 3SAT is in NP, since SAT is in NP.
To determine whether a boolean expression E in CNF is satisfiable, nondeterministically guess values for all the variables and then evaluate the expression. If E turns out to be true, then accept. This can be carried out in nondeterministic polynomial time. Thus 3SAT is in NP.

To prove NP-completeness, we shall reduce CSAT to 3SAT. Let a given CNF expression be $E = e_1 \wedge e_2 \wedge ... \wedge e_k$ where each e_i is a disjunction of literals. We replace each clause e_i (as shown below) to create a new expression F such that

i) F is a clause in 3SAT form.

ii) time taken to construct F is linear in the length of E.

iii) a truth assignment satisfies E if and only if it can be extended to a satisfying truth assignment for F.

The construction of F is as follows.

1. If an e_i is a single literal, say (x) or $(\overline{x})$ we introduce two new variables u and v. We replace (x) by the conjunction of four clauses as $(x+u+v)(x+u+\overline{v})(x+\overline{u}+v)(x+\overline{u}+\overline{v})$. Since u and v appear in all combinations, the only way to satisfy all four clauses is to make x true.

2. Suppose an e_i is the disjunction of two literals, $(x + y)$. We introduce a new variable z and replace e_i by the conjunction of two clauses $(x + y + z)(x + y + \overline{z})$. As in case 1, the only way to satisfy both clauses is to satisfy $(x + y)$.

3. If an e_i is the disjunction of three literals it is already in the form required for 3-CNF, so we take e_i as such to construct F.

4. Suppose $e_i = (x_1+x_2+...+x_m)$ for some $m \geq 4$. We introduce new variables $y_1, y_2, ..., y_{m-3}$ and replace e_i by the conjunction of clauses

$$(x_1+x_2+y_1)(x_3+\overline{y_1}+y_2)(x_4+\overline{y_2}+y_3)\dots(x_{m-2}+\overline{y_{m-4}}+y_{m-3})(x_{m-1}+x_m+\overline{y_{m-3}}). \quad (1)$$

If there is a truth assignment τ that satisfies e_i then at least one literal (one of the x's) in $x_1 + x_2 + ... + x_m$ should be true; say τ makes x_j true; that is in

$$\dots(x_{j-1} + \overline{y_{j-3}} + y_{j-2})(x_j + \overline{y_{j-2}} + y_{j-1})(x_{j+1} + \overline{y_{j-1}} + y_j)\dots$$

Then in (1) above, if we make $y_1, y_2, ..., y_{j-2}$ true and make $y_{j-1}, y_j, ..., y_{m-3}$ false, we satisfy all the clauses of (1). Thus τ can be extended to satisfy these clauses. On the other hand, if τ makes all the x's false we can reason (as follows) that it is not possible to extend τ to make (1) true. The reason is that, since x_1 and x_2 are false, y_1 must be true (otherwise this clause will become false and will collapse e_i to false) in the first clause; since x_3 is false and $\overline{y_1}$ is false, y_2 must be true to keep the situation alive. By continuing the argument, we will reason that y_{m-3} is true. But, alas, x_{m-1} is false, x_m is false and $\overline{y_{m-3}}$ is false and we see that e_i cannot be satisfied.

The above argument shows how to reduce an instance E of CSAT to an instance F of 3SAT, such that F is satisfiable if and only if E is satisfiable. The construction evidently requires time that is linear in the length of E, because none of the four cases above expand a clause by more than a factor 32/3 (that is the ratio of symbol counts in case 1) and it is possible to built F in polynomial time.

Since CSAT is NP-complete, it follows that 3SAT is also NP-complete. □

1-IN-3SAT

The 1-IN-3SAT problem is the following.

INSTANCE : A collection of clauses $C_1, ..., C_m$, $m > 1$; each C_i is a disjunction of exactly three literals.
QUESTION : Is there a truth assignment to the variables occurring so that exactly one literal is true in each C_i ?

Let $X = \{x_1, ..., x_5\}$ be the set of variables. Let the clause set $C = \{C_1, C_2, C_3\}$ be the following : $C_1 = \{\overline{x_1}, \overline{x_2}, x_3\}$, $C_2 = \{\overline{x_1}, x_4, x_5\}$, $C_3 = \{\overline{x_2}, x_4, x_5\}$. With this (X, C) we consider the 1-IN-3SAT problem. We say a clause is correctly satisfied if and only if the clause is satisfied due to exactly one literal in it. To correctly satisfy C_1 we try setting only $\overline{x_1} = \top$ in C_1. This when applied to C_2 requires $x_4 = \bot$ and $x_5 = \bot$ for correct satisfaction of C_2. But $x_4 = \bot$ and $x_5 = \bot$ in C_3 requires $\overline{x_2} = \top$ for correct satisfaction of C_3—this violates the correct satisfaction of C_1. Similarly we get a contradiction if we try setting only $\overline{x_2} = \top$ in C_1. The only solution is $\overline{x_1} = \bot$, $\overline{x_2} = \bot$ and $x_3 = \top$ and exactly one of x_4 or x_5 is $\top$. Note that as a 3SAT instance (X, C) admits other solutions as well.