Reduction 3SAT to IND

We define an instance of 3SAT to be a Boolean expression E in conjunctive normal form where each clause has three terms. That is, $E = C_1 * C_2 * \cdots * C_k$ where $C_i = t_{i,1} + t_{i,2} + t_{i,3}$ and where each term $t_{i,j}$ is either of the form x or $!x$, where x is an identifier, an alphanumeric string which begins with a letter.. E is *satisfiable* if there exists a a *satifying assignment* of E, that is, an assignment of a Boolean value to each identifier in the expression such that the value of E, treating each identifer as a Boolean variable and assigning to each term the value given by that assignment, the value of E is true.

We define an instance of IND, the independent set problem, to be a string of the form $\langle G\rangle\langle k\rangle$ where G is a graph and k a positive integer. Let ISI (independent set instance) be the language consisting of all such strings. Let IND be the set of all $\langle G\rangle\langle k\rangle \in$ ISI such that there is some set I of vertices of G such that $|I| = k$ and no two members of I form an edge of G. Thus IND $\subseteq$ ISI. Similarly, 3SAT $\subseteq$ 3CNF, the language of all Boolean expressions in conjunctive normal where each clause has three terms.

We now define a $\mathcal{P}$–TIME reduction R of 3SAT to IND, a $\mathcal{P}$–TIME function R: 3CNF $\to$ ISI such that $R(E) \in$ IND if and only if $E \in$ 3SAT.

Let $E = (t_{1,1} + t_{1,2} + t_{1,3}) * (t_{2,1} + t_{2,2} + t_{2,3}) * \cdots * (t_{k,1} + t_{k,2} + t_{k,3}) \in$ 3CNF. We say that two terms $t_{i,j}$ and $t_{i',j'}$ *contradict* if $t_{i,j} = x$ and $t_{i',j'} = !x$ for some identifier x.

We define a graph $G = (V, E)$ as follows.

- $V = \{v_{i,j}\}_{1\le i\le k, 1\le j\le 3}$
- $E_1 = \left\{\left\{v_{i,j}, v_{i,j'}\right\} : j \neq j'\right\}$
- $E_2 = \left\{\left\{v_{i,j}, v_{i',j'}\right\} : i \neq i' \text{ and } t_{i,j} \text{ and } t_{i',j'} \text{ contradict }\right\}$
- $E = E_1 \cup E_2$ We refer to the complete subgraph containing $\{v_{i,1}, v_{i,2}, v_{i,3}\}$ as a *triangle*

Theorem 1 *If* $E \in$ 3SAT *then* $\langle G\rangle\langle k\rangle \in$ IND.

Proof: Suppose $E \in$ 3SAT. Fix a satisfying assignment of E. For each $1 \le i \le k$, there is at least one term of the i^{th} clause of E which is assigned the value true. Pick $J(i)$ such that $t_{i,J(i)}$ is assigned true. Let $I = \left\{v_{i,J(i)}\right\}$, a set of k vertices of G. We need to prove no two members of I form an edge. Pick two vertices of I, $v = v_{i,J(i)}$ and $v' = v_{i',J(i')}$. Since J is 1-1, $i \neq i'$, hence v and v' are in different triangles, and thus no edge of E_1 connects them. Suppose v and v' are connected by an edge in E_2. Thus, their corresponding terms contradict, *i.e.*,, without loss of generality, $t_{i,J(i)} = x$ and $t_{i',J(i')} = !x$ for some variable x. However, by definition of J, both terms are assigned true, contradiction. ∎

Theorem 2 *If* $\langle G\rangle\langle k\rangle \in$ IND *then* $E \in$ 3SAT.

Proof: Let I be an independent set of k vertices of G. Since each triangle is a complete subgraph, and since there are k triangles, there must be exactly one member of I in each triangle. Let $v_{i,J(i)}$ be the member of I in the ith triangle. We define an assignment of the variables of E as follows.

- If $t_{i,J(i)} = x$ for some variable x, then assign the value true to x.

- If $t_{i,J(i)} = !y$ for some variable y, then assign the value false to y.
- Assign an arbitrary truth value to each of the remaining variables.

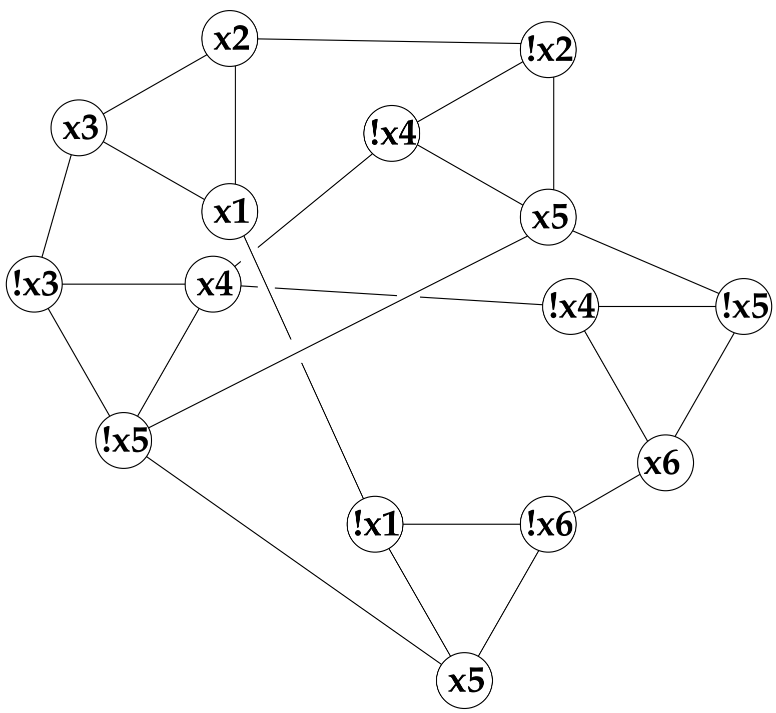
No variable can be assigned both false and true, because otherwise the corresponding vertices would be connected by an edge in E_2. For each i, $t_{i,J(i}$ is assigned the value true, and hence the i^{th} clause has the value true, and hence E has the value true. Thus, E has a satisfying assignment, hence is in 3SAT. ▮

Theorem 3 *If* 3SAT *is NP–*COMPLETE *then* IND *is NP–*COMPLETE.

Proof: R is a polynomial time reduction of 3SAT to IND, 3SAT is $\mathcal{NP}$ complete, and IND is clearly $\mathcal{NP}$. ▮

Example

Let $E = (x1 + x2 + x3) * (!x3 + x4 + !x5) * (!x1 + !x6 + x5) * (!x4 + !x5 + x6) * (!x2 + !x4 + x5)$



Then $E \in$ 3SAT and $< G > 5 \in$ IND, where G is the following graph. In this example, there are many correct choices of certificate.