Exercise 1: Prove that `DIRECTED DOMINATING SET` is NP-Hard through a series of Karp reductions starting at `3SAT`.

Solution: I will first introduce some notation. Let F be a boolean formula with N variables and M CNF clauses with 3 literals each. I will denote the variables as $x_1, x_2, \ldots, x_N$ and the clauses as $c_1, c_2, \ldots, c_M$. For convenience, I will enumerate all possible literals as $l_j = x_j, l_{N+j} = \overline{x_j}$ (there are $2N$ of them). I will denote $c_i = \left(l_{j_{i,1}}, l_{j_{i,2}}, l_{j_{i,3}}\right) = \left(l_1^i, l_2^i, l_3^i\right) \in \{l_1, \ldots, l_{2N}\}^3$.

I will now construct a directed graph $G = (V, E)$ and $k \in \mathbb{N}$ such that G contains a dominating set of size at most k if and only if F is satisfiable. Furthermore, the construction of the graph will clearly be polynomial in time, thus providing the Karp reduction we need directly.

First, I define the vertices V of G as:

- A vertex L_j for each literal l_j ($2N$ in total, which can be created in linear time by scanning F). For convenience, I will denote $L_u^i \coloneqq L_{j_{i,u}}$.
- A vertex C_i for each clause c_j (M in total, which similarly can be created in linear time).

Next, I define the edges E of G as:

- $\left(L_u^i, C_i\right)$ for $1 \leq u \leq 3$, $1 \leq i \leq M$ (each literal points to the clauses it appears in, which we can construct in linear time).
- (L_s, L_{s+N}) and (L_{s+N}, L_s) for $1 \leq s \leq N$ (each literal points to its negation, which we can construct in linear time).

Finally, I define $k = N$. It remains to be proven that G has a dominating set of size $k \iff F$ is satisfiable:

$\Longleftarrow$) Suppose we have an assignment $x_s = B_s \in \{\text{True}, \text{False}\}$ that satisfies F. I will show that the set $S \coloneqq \{l_s | B_s = \text{True}\} \cup \{l_{s+N} | B_s = \text{False}\}$, which has size $N = k$, is dominating:

- All L_j are either in S or pointed to by $\overline{L_j} \coloneqq L_{j \pm N} \in S$.
- All C_i are pointed to by their literals, at least one of which is in S.

$\Longrightarrow$) Suppose there is a dominating set S of size at most k. for each variable x_s, L_s must either be in S or pointed to by an element of S (that is, one of L_s, L_{s+N} is in S). In fact, because there are $N = k$ variables, *exactly* one of them is in S (otherwise S would have more than k elements). Furthermore, S only contains vertices of the form L_j (and not C_i).

This means that a variable assignment $x_s = B_s$ where $B_s = \text{True}$ if $L_s \in S$ and $B_s = \text{False}$ if $L_{s+N} \in S$ can be defined. To show that this assignment satisfies F, note that for each clause c_i, there is a literal l_j such that L_j points to C_i and $L_j \in S$. If $j \leq N$, this means we have assigned $x_j = \text{True}$ and $x_j = l_u^i$ for $1 \leq i \leq 3$, satisfying the clause. Otherwise, we have assigned $x_{j-N} = \text{False}$ and $\overline{x_{j-N}} = l_u^i$ for $1 \leq i \leq 3$, satisfying the clause as well.

Remark. This construction works just as well for arbitrary `SAT` instances, not just `3SAT`.

Exercise 2:

References

[1] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.