Université Gustave Eiffel
M2 – Semester 1 – 2025/2026

**Complexity - Exercise Sheet 4**

**CHAU Dang Minh**

**Exercise 4.7.** What is wrong with the following proof of P $\neq$ NP.

Assume that P = NP. Then there exists an algorithm $A$ and a polynomial $p(n)$ such that SAT is decided by $A$ in time $O(p(n))$. Assume that $p(n) = O(n^{37})$. By the Time Hierarchy Theorem, there exists a problem $P \in \mathrm{DTIME}(n^{38})$ such that $P \notin \mathrm{DTIME}(n^{37} \log n^{37}) = \mathrm{DTIME}(n^{37} \log n)$. Since SAT is NP-complete, we can reduce $P$ to SAT and decided it in time $O(n^{37})$. But we have just shown that $P$ requires time $\omega(n^{37} \log n)$. This leads to a contradiction, hence the assumption P = NP must be false.

*Solution.* The assumption that $p(n) = O(n^{37})$ is not necessarily valid. But even if we make a weaker assumption that $p(n) = O(n^k)$ for some $k$, the proof is still flawed. The proof uses the Time Hierarchy Theorem to find a problem $P \in \mathrm{DTIME}(n^{k+1})$ such that $P \notin \mathrm{DTIME}(n^k)$, and arrive at a contradiction by reducing $P$ to SAT $O(n^k)$. However, suppose that $P$ is reduced to SAT in time $O(n^c)$ for some $c \geq 1$. Let $f$ be the reduction function. Then there is some $d \leq c$ such that if for every $x \in \{0,1\}^*$, we have $|f(x)| \in O(n^d)$, because the length of the output cannot exceed the time of the reduction. Therefore, the time complexity to decide $P$ is in $O(n^c + n^{dk})$. For the contradiction to hold i.e. $O(n^c + n^{dk}) = O^k$, we must have $dk \leq k$, or equivalently $d \leq 1$, which is not provided by the proof.

**Exercise 4.12.** Consider the problem of determining whether a DNF formula $\phi$ has an equivalent formula having less than $k$ literals.

$$\textsc{Min-DNF} = \{\langle \phi, k\rangle \mid \exists \text{ DNF } \psi \text{ s.t. } \phi \equiv \psi \text{ and } \psi \text{ has } \leq k \text{ occurrences of literals}\}.$$

Show that $\textsc{Min-DNF} \in \Sigma_2^p$.

*Solution.* We measure the size of a DNF formula $\phi$, denoted by $|\phi|$, in terms of the number of literals it contains. Let $x = (x_1, \ldots, x_k), k \leq |\phi|$ be the vector of variables in the formula $\phi$. We have $\langle \phi, k\rangle \in \textsc{Min-DNF}$ if and only if the following QBF is true.

$$\exists \psi \ \forall x \ R(\phi, x),$$

where $R(\phi, x) = (\phi(x) = \psi(x))$. The domain of discourse for $\psi$ is the set of all DNF formulas with at most $k$ literals and the domain of discourse for $x$ is $\{0,1\}^k$. It is clear that the length of each $x$ is linear in $k$, and hence linear in $|\phi|$. Every formula $\psi$ in the former domain has at most $k$ variables, hence it has at most $2k$ different literals. We will also count $\wedge$ and $\vee$. Hence we need at most $\log(2k+2)$ bits to encode each literal and operator. The encoding first converts $\phi$ to the prefix notation, then encodes each literal/operator using $\log(2k+2)$ bits. Therefore, the size of the encoding of $\psi$ is at most $(2k-1)\log(2k+2)$ ($k$ literals and $k-1$ operators), which is polynomial in $|\phi|$. Thus, $P \in \Sigma_2^p$.

**Exercise 4.13.** The complexity class DP is defined as those decision problems that can be written as an intersection of an NP problem and a co-NP problem.

(a) Show that $\textsc{3Sat-3Unsat} = \{\langle \phi, \psi\rangle \mid \phi \in \textsc{3Sat}, \psi \notin \textsc{3Sat}\}$ is DP-complete.

(b) Let $\alpha(G)$ be the independence number of a graph $G$. Let $\textsc{ExactIndSet} = \{\langle G, k\rangle \mid \alpha(G) = k\}$. Show that $\textsc{ExactIndSet} \in \mathrm{DP}$.

(c) For two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, the lexicographic product of $G_1$ with $G_2$ is

$$G = (V_1 \times V_2, \{((u_1, u_2), (v_1, v_2)) \mid (u_1, v_1) \in E_1 \text{ or } u_1 = v_1 \text{ and } (u_2, v_2) \in E_2\}).$$

Show that $\alpha(G) = \alpha(G_1) \cdot \alpha(G_2)$.

(d) Show that ExactIndSet is DP-complete.

(e) Show that $\mathrm{NP} \cup \mathrm{coNP} \subseteq \mathrm{DP} \subseteq \Sigma_2^p \cap \Pi_2^p$.

*Solution.*

(a) We first show that 3Sat-3Unsat $\in$ DP. Let

$$L_1 = \{\langle \phi, \psi \rangle \mid \phi \in \text{3Sat}\} \text{ and } L_2 = \{\langle \phi, \psi \rangle \mid \psi \notin \text{3Sat}\}.$$

The problem $L_1$ is in NP, because we use a verifier for 3Sat on the first component $\phi$ of every instance $\langle \phi, \psi \rangle \in L_1$. Similarly, $L_2^c = \{\langle \phi, \psi \rangle \mid \psi \in \text{3Sat}\} \in \mathrm{NP}$, hence $L_2 \in \mathrm{coNP}$. We have 3Sat-3Unsat $= L_1 \cap L_2$, hence 3Sat-3Unsat $\in$ DP.

Next, we show that 3Sat-3Unsat is DP-hard. In fact, $L_1$ is NP-complete because we can reduce every instance $\phi \in$ 3Sat to $\langle \phi, \psi_0 \rangle \in L_1$, where $\psi_0$ is a fixed formula. Similarly, $L_2^c$ is NP-complete. Therefore, for every co-NP problem $M$, we have

$$x \in M \iff x \notin M^c \xLeftrightarrow{L_2^c \in \mathrm{NP}} f(x) \notin L_2^c \iff f(x) \in L_2,$$

where $f$ is a polynomial-time reduction from $M^c$ to $L_2^c$, or $L_2$ is co-NP-complete. Therefore, for every $M = M_1 \cap M_2 \in \mathrm{DP}$, where $M_1 \in \mathrm{NP}$ and $M_2 \in$ co-NP, there exist polynomial-time reductions $u$ from $M_1$ to $L_1$ and $v$ from $M_2$ to $L_2$. We define the reduction $h$ from $M$ to 3Sat-3Unsat as $h(x) = \langle u(x), v(x) \rangle$. It is clear that $h$ is computable in polynomial time. Therefore, 3Sat-3Unsat is DP-complete.

(b) Let $L_1 = \{\langle G, k \rangle \mid \alpha(G) \geq k\}$ and $L_2 = \{\langle G, k \rangle \mid \alpha(G) \leq k\}$. The problem $L_1$ is exactly our well-known IndSet problem, because $\langle G, k \rangle \in L_1$ if an only if $G$ has an independent set of size at least $k$. Hence $L_1$ is NP-complete. Using similar argument as in question (a), we derive that $L_2$ is co-NP-complete, because its complement is $\{\langle G, k \rangle \mid \alpha(G) \geq k+1\}$, a slightly modification of $L_1$, which is NP-complete. Therefore, ExactIndSet $= L_1 \cap L_2 \in$ DP.

(c) Let $I_1 \subseteq V_1$ and $I_2 \subseteq V_2$ be two independent sets of $G_1$ and $G_2$ respectively. We show that $I = I_1 \times I_2$ is an independent set of $G$. For every $(u_1, u_2), (v_1, v_2) \in I$, we have $u_1, v_1 \in I_1$ and $u_2, v_2 \in I_2$. Since $I_1$ and $I_2$ are independent sets, we have $(u_1, v_1) \notin E_1$ and $(u_2, v_2) \notin E_2$. Therefore, by the definition of the lexicographic product, $((u_1, u_2), (v_1, v_2)) \notin E$. Hence, $I$ is an independent set of $G$. Therefore, if $I_1$ and $I_2$ are maximum independent sets of $G_1$ and $G_2$ respectively, then $I$ is an independent set of $G$ with size $|I| = |I_1| \cdot |I_2|$. This shows that $\alpha(G) \geq \alpha(G_1) \cdot \alpha(G_2)$.

On the other hand, let $I \subseteq V$ be an independent set of $G$. Let $I_1 = \{u \mid (u, v) \in I\}$. We claim that $I_1$ is an independent set of $G_1$. If $|I_1| = 1$, we are done. Otherwise, for every $u, v \in I_1$, there exist $(u, u'), (v, v') \in I$ for some $u', v' \in V_2$. Since $I$ is an independent set of $G$, we have $((u, u'), (v, v')) \notin E$. By the definition of the lexicographic product, this implies that $(u, v) \notin E_1$. Hence, $I_1$ is an independent set of $G_1$. Similarly, $I_2 = \{v \mid (u, v) \in I\}$ is an independent set of $G_2$. Since $I \subseteq I_1 \times I_2$, we have $|I| \leq |I_1| \cdot |I_2|$. Therefore, if $I$ is a maximum independent set of $G$, then $\alpha(G) = |I| \leq |I_1| \cdot |I_2| \leq \alpha(G_1) \cdot \alpha(G_2)$. Combining this with the previous result, we have $\alpha(G) = \alpha(G_1) \cdot \alpha(G_2)$.

(d) Let $M \in$ DP. Since 3Sat-3Unsat is DP-complete, we can reduce $M$ to 3Sat-3Unsat in polynomial time. The remaining is to reduce 3Sat-3Unsat to ExactIndSet. Let $\langle \phi, \psi \rangle$

be an instance of 3SAT-3UNSAT. Without loss of generality, assume that both $\phi$ and $\psi$ have $n$ clauses (by adding true clause $(x \vee \neg x \vee y)$ to the formula having fewer clauses). For every formula $\phi$, we construct the graph $H(\phi)$ as in our previous proof for that INDSET is NP-complete.

1. For each clause $C_i = (l_{i1} \vee l_{i2} \vee l_{i3})$ in $\phi$, we create three vertices $v_{i1}, v_{i2}, v_{i3}$ corresponding to the three literals $l_{i1}, l_{i2}, l_{i3}$ and add edges between every pair of them.
2. For every pair of vertices $v_{ij}$ and $v_{kl}$, where $i \neq k$, we add an edge $(v_{ij}, v_{kl})$ if and only if the literals $l_{ij}$ and $l_{kl}$ are complementary.

If $\phi$ is satisfiable, we can select one true literal from each clause to form an independent set of size $n$. Suppose that there is an independent set of size $n+1$. Then there exist two vertices from the same set $\{l_{i1}, l_{i2}, l_{i3}\}$ for some $i$. But this contradicts the fact that they are all connected by edges. Hence, $\alpha(H(\phi)) = n$.

If $\phi$ is not satisfiable, then for any selection of $n$ vertices, either there are two vertices corresponding to complementary literals, or there exists at least one clause $C_i$ such that none of its literals is selected. The first case contradicts the independence of the set, while the second case brings us back to the satisfiable case with $n-1$ clauses, which also raises a contradiction. Therefore, $\alpha(H(\phi)) \leq n-1$.

Next, for any two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, let $G_1 \vee G_2$ by the graph obtained by collecting all vertices and edges of $G_1$ and $G_2$ and adding edges between every pair of vertices $u \in V_1$ and $v \in V_2$. It is clear that $\alpha(G_1 \vee G_2) = \max(\alpha(G_1), \alpha(G_2))$, because any independent set of $G_1 \vee G_2$ can only contain vertices from either $G_1$ or $G_2$. Also denote by $G_1 \circ G_2$ the lexicographic product of $G_1$ and $G_2$.

Let $E_n$ be the graph with $n$ isolated vertices. For every formula $\phi$, consider the graph

$$G(\phi) = (H(\phi) \circ E_{n+1}) \vee E_{(n-1)(n+1)}.$$

If $\phi \in$ 3SAT, then $\alpha(H(\phi)) = n$. Hence $\alpha(G(\phi)) = \max\{n(n+1), (n-1)(n+1)\} = n(n+1)$. If $\phi \notin$ 3SAT, then $\alpha(H(\phi)) \leq n-1$. Hence $\alpha(G(\phi)) = (n-1)(n+1)$.

Now for the instance $\langle \phi, \psi \rangle$ of 3SAT-3UNSAT, we construct the graph

$$G = G(\phi) \circ G(\phi) \circ G(\psi).$$

Consider four cases.

1. If $\phi \in$ 3SAT and $\psi \in$ 3SAT, then $\alpha(G) = n^3(n+1)^3$.
2. If $\phi \in$ 3SAT and $\psi \notin$ 3SAT, then $\alpha(G) = (n-1)n^2(n+1)^3$.
3. If $\phi \notin$ 3SAT and $\psi \notin$ 3SAT, then $\alpha(G) = (n-1)^2 n(n+1)^3$.
4. If $\phi \notin$ 3SAT and $\psi \in$ 3SAT, then $\alpha(G) = (n-1)^3(n+1)^3$.

It is clear that the value in the second case is not equal to other values. Let $k = (n-1)n^2(n+1)^3$, we have $\langle \phi, \psi \rangle \in$ 3SAT-3UNSAT $\iff \alpha(G) = k$. Our constructs are all computable in polynomial time. Therefore, we have reduced 3SAT-3UNSAT to EXACTINDSET in polynomial time, and hence EXACTINDSET is DP-complete.

(e) Note that $\{0,1\}^*$ and $\varnothing$ are in NP, since we can use the verifier that always accepts and rejects, respectively. Hence $\{0,1\}^*$ is in co-NP. Therefore for every $L \in$ NP, we have $L = L \cap \{0,1\}^* \in$ DP. For every $L \in$ co-NP, we have $L = \{0,1\}^* \cap L \in$ DP. This shows that NP $\cup$ co-NP $\subseteq$ DP.

Next, let $L \in \mathrm{DP}$. Then there exist $L_1 \in \mathrm{NP}$ and $L_2 \in \text{co-NP}$ such that $L = L_1 \cap L_2$. The corresponding quantified boolean formula is $\exists y_1\, R_1(x, y_1)$ and $\forall y_2\, R_2(x, y_2)$. Therefore, we can express $x \in L$ if and only if

$$T = \exists y_1\, R_1(x, y_1) \wedge \forall y_2\, R_2(x, y_2)$$

is true. Since $y_1$ does not appear in $R_2$ and $y_2$ does not appear in $R_1$, we can rewrite $T$ in two equivalent forms.

$$T = \exists y_1\, \forall y_2\, R_1(x, y_1) \wedge R_2(x, y_2) = \forall y_2\, \exists y_1\, R_1(x, y_1) \wedge R_2(x, y_2).$$

Therefore, $x \in L$ if and only if the above QBFs are true. This shows that $L \in \Sigma_2^p \cap \Pi_2^p$. Thus, $\mathrm{DP} \subseteq \Sigma_2^p \cap \Pi_2^p$.