IX Зимняя школа «Комбинаторика и алгоритмы»
Даниил Мусатов, «Вычислительные задачи поиска»
Список упражнений и задач №1: **NP**-полнота

Через $\{0,1\}^*$ обозначается множество всех конечных слов из 0 и 1. *Языком* называется любое подмножество $\{0,1\}^*$. Язык называется *разрешимым*, если существует некоторый алгоритм M, такой что $M(x) = 1$ при $x \in L$ и $M(x) = 0$ при $x \notin L$.

1. Докажите, что не все языки разрешимы.

Классом $\mathbf{TIME}(f(n))$ называется множество языков L, для которых существует некоторая константа c и некоторый алгоритм M, такой что $M(x) = 1$ при $x \in L$, $M(x) = 0$ при $x \notin L$ и при всех x вычисление $M(x)$ длится не более $cf(|x|)$ шагов. Классом $\mathbf{P}$ называется объединение $\bigcup_{k=1}^{\infty} \mathbf{TIME}(n^k)$.

2. Докажите, что если A и B лежат в $\mathbf{P}$, то $A \cup B$, $A \cap B$, $\overline{A}$ также лежат в $\mathbf{P}$.

Классом $\mathbf{NP}$ называется множество языков L, для которых существуют полином $p(n)$ и алгоритм $V(x,s)$, такой что при всех $x \in L$ для некоторого s верно $V(x,s) = 1$, при всех $x \notin L$ для всех s верно $V(x,s) = 0$ и при всех x и s время работы $V(x,s)$ ограничено $p(|x|)$. (В частности, $V(x,s)$ прочтёт не больше $p(|x|)$ символов s).

3. Докажите, что если A и B лежат в $\mathbf{NP}$, то $A \cup B$ и $A \cap B$ также лежат в $\mathbf{NP}$.

4. Докажите, что $\mathbf{P} \subset \mathbf{NP}$.

Классом $\mathbf{coNP}$ называется множество $\{L \mid \overline{L} \in \mathbf{NP}\}$.

5. Докажите, что если $\mathbf{P} = \mathbf{NP}$, то $\mathbf{NP} = \mathbf{coNP}$.

Говорят, что язык A полиномиально сводится (по Карпу) к языку B, если существует полиномиально вычислимая функция $f\colon \{0,1\}^* \to \{0,1\}^*$, такая что $x \in A \Leftrightarrow f(x) \in B$. Обозначение: $A \leqslant_p B$.

6. Докажите, что:

а) $A \leqslant_p A$;
б) Если $A \leqslant_p B$ и $B \leqslant_p C$, то $A \leqslant_p C$;
в) Если $A \in \mathbf{P}$, $B \neq \varnothing$ и $B \neq \{0,1\}^*$, то $A \leqslant_p B$;
г) Если $B \in \mathbf{P}$ и $A \leqslant_p B$, то $A \in \mathbf{P}$;
д) Если $B \in \mathbf{NP}$ и $A \leqslant_p B$, то $A \in \mathbf{NP}$.

Язык B называется $\mathbf{NP}$-*трудным*, если для любого $A \in \mathbf{NP}$ выполнено $A \leqslant_p B$, и $\mathbf{NP}$-*полным*, если, помимо этого, $B \in \mathbf{NP}$.

7. Докажите, что:

а) Если A — $\mathbf{NP}$-трудный и $A \leqslant_p B$, то B — $\mathbf{NP}$-трудный;
б) Если $\mathbf{P} = \mathbf{NP}$, то любой язык из $\mathbf{NP}$ (кроме $\varnothing$ и $\{0,1\}^*$) является $\mathbf{NP}$-полным;
в) Если A — $\mathbf{NP}$-полный и $A \in \mathbf{P}$, то $\mathbf{P} = \mathbf{NP}$.

Теорема Кука–Левина утверждает, что язык $\mathsf{SAT} = \{\varphi \mid \varphi$ — выполнимая логическая формула$\}$ является $\mathbf{NP}$-полным.

8. Сведите к SAT без использования теоремы Кука–Левина следующие языки:

а) $\mathsf{3COL} = \{G \mid$ вершины графа G можно раскрасить в 3 цвета, так чтобы не было одноцветных рёбер$\}$;

б) $\mathsf{dHAMPATH} = \{(G.s,t) \mid$ в ориентированном графе G есть ориентированный гамильтонов путь из s в $t\}$ (путь называется гамильтоновым, если проходит ровно один раз через каждую вершину);

в) EXACT-SET-COVER $= \{(S_1, \ldots, S_m) \mid$ из семейства множеств $S_1, \ldots, S_m$ можно выбрать подсемейство, покрывающее каждый элемент их объединения ровно по одному разу$\}$;

г) SUBSET-SUM $= \{(n_1, \ldots, n_k, N) \mid \exists m\ \exists i_1 < \cdots < i_m \colon n_{i_1} + \cdots + n_{i_m} = N\}$;

д) INDSET $= \{(G, k) \mid$ в графе G есть независимое множество из k вершин$\}$ (множество вершин называется независимым, или антикликой, если между этими вершинами нет рёбер);

е) VERTEXCOVER $= \{(G, k) \mid$ в графе G есть вершинное покрытие из k вершин$\}$ (вершинное покрытие — такое множество вершин, в котором у каждого ребра лежит хотя бы один конец).

9. Докажите, что язык 3SAT $= \{\varphi \mid \varphi$ — выполнимая КНФ с 3 литералами в каждом дизъюнкте$\}$ является **NP**-полным. (Указание: сведите SAT к 3SAT, введя дополнительные переменные для всех промежуточных результатов).

10. Докажите, что язык INDSET является **NP**-полным. (Указание: сведите 3SAT к INDSET. Каждой скобке сопоставьте треугольник, помеченный литералами. Также соедините рёбрами противоположные литералы).

11. Докажите, что язык VERTEXCOVER является **NP**-полным. (Указание: что будет дополнением к вершинному покрытию?)

12. Докажите, что языки EXACT-SET-COVER и SUBSET-SUM являются **NP**-полными. (Указание: первый ко второму сводится довольно легко. К первому сводится 3SAT, нужно покрывать множество из переменных, вхождений переменных в тройки и самих троек).

Задачей поиска для **NP**-языка, заданного верификатором V, называется задача нахождения по входу x любого сертификата s, такого что $V(x, s) = 1$. Говорят, что задача поиска сводится (по Куку) к задаче распознавания, если существует полиномиальный алгоритм, использующий решатель для задачи распознавания как «чёрный ящик». (Т.е. можно полимномиальное число раз запускать чёрный ящик и производить полиномиальные вычисления с результатами).

13. Сведите задачу поиска к задаче распознавания для SAT, INDSET, EXACT-SET-COVER, SUBSET-SUM, 3COL, dHAMPATH, VERTEXCOVER.

14*. Язык называется *унарным*, если все слова в нём имеют вид 1^k (т.е. состоят из одних единиц). Докажите, что если существует унарный **NP**-полный язык, то **P** = **NP**. (Указание: воспользуйтесь тем, что этот язык и SAT) сводятся друг к другу, и запустите динамическое программирование.