Колмогоровская сложность

Будем рассматривать функции над множеством *двоичных слов* (слов из нулей и единиц) $\{0,1\}^*$. Длину слова $x \in \{0,1\}^*$ обозначим $|x|$.

Определение. Пусть фиксирована (частичная) вычислимая функция $D\colon \{0,1\}^* \to \{0,1\}^*$, называемая *способом описания. Колмогоровской сложностью* слова $x \in \{0,1\}^*$ относительно способа описания D называется число

$$\mathrm{KS}_D(x) = \min\{|y| \mid D(y) = x\}.$$

Если же такого y, что $D(y) = x$, не существует, то $\mathrm{KS}_D(x) = +\infty$.

1. Докажите, что для любых двух способов описания D_1 и D_2 существует такой способ описания D_0, что $\mathrm{KS}_{D_0}(x) \leqslant \mathrm{KS}_{D_1}(x) + c$ и $\mathrm{KS}_{D_0}(x) \leqslant \mathrm{KS}_{D_2}(x) + c$, где c — константа, не зависящая от x.

Теорема Соломонова – Колмогорова (об оптимальном способе описания). Существует такой способ описания D, что для любого другого способа описания D' существует константа c, такая что для всех $x \in \{0,1\}^*$

$$\mathrm{KS}_D(x) \leqslant \mathrm{KS}_{D'}(x) + c.$$

В дальнейшем будем считать, что фиксирован некоторый оптимальный способ описания D, и колмогоровскую сложность будем обозначать просто $\mathrm{KS}(x)$ (эта величина определена с точностью до аддитивной константы, зависящей от выбора оптимального способа описания).

2. а) Докажите, что существует константа c, такая что для любого слова x верно $\mathrm{KS}(x) \leqslant |x| + c$.
б) Докажите, что $\mathrm{KS}(\underbrace{111\ldots111}_{n \text{ единиц}}) \leqslant \log_2 n + c$.
в) Докажите, что $\mathrm{KS}(\underbrace{111\ldots111}_{n \text{ единиц}}) \leqslant \mathrm{KS}(n) + c$ (при вычислении $\mathrm{KS}(n)$ считается, что число n представлено в двоичной записи).
г) Оцените сверху величину $\mathrm{KS}(\underbrace{111\ldots111}_{4^{2^n+3n} \text{ единиц}})$.

3. а) Докажите, что существует константа c, такая что для любых двух слов x и y верно

$$\mathrm{KS}(xy) \leqslant \mathrm{KS}(x) + \mathrm{KS}(y) + 2\log_2 \mathrm{KS}(x) + c,$$

где xy — слово, получаемое приписыванием слова y к слову x.
б) Докажите более сильное неравенство

$$\mathrm{KS}(xy) \leqslant \mathrm{KS}(x) + \mathrm{KS}(y) + \log_2 \mathrm{KS}(x) + 2\log_2\log_2 \mathrm{KS}(x) + c.$$

4. Докажите, что для любого n существует «несжимаемое» слово x длины n, такое что $\mathrm{KS}(x) \geqslant n$.

5. Чему асимптотически равно среднее арифметическое сложностей всех слов длины n?

6. Докажите, что если f вычислима, то для некоторой константы c неравенство $\mathrm{KS}(f(x)) \leqslant \mathrm{KS}(x) + c$ верно для всех x.

7. Существует ли вычислимая функция $f\colon \mathbb{N} \to \{0,1\}^*$, такая что $|f(n)| = n$ и $\mathrm{KS}(f(n)) \geqslant n$ для всех n?

8. а) Докажите, что функция $\mathrm{KS}\colon x \mapsto \mathrm{KS}(x)$ невычислима. **б)** Докажите, что у функции KS не существует нетривиальных вычислимых нижних оценок: если вычислимая функция $k\colon \{0,1\}^* \to \mathbb{N}$ такова, что для всех x из области определения k верно $k(x) \leqslant \mathrm{KS}(x)$, то k ограничена сверху константой.