МФТИ, бакалавриат ФИВТ. 2014.
Упражнения к первой части курса
***Введение в теорию информации.* А.Е. Ромащенко**

1. Докажите, что для любых совместно распределённых случайных величин выполнены неравенства

a) $2H(a,b,c) \leq H(a,b) + H(a,c) + H(b,c \mid a)$,
b) $H(a,b,c) + H(a) \leq H(a,b) + H(a,c)$,
c) $H(c \mid d) \leq H(c \mid a) + H(c \mid b) + I(a : b \mid d)$.

2. Докажите, что неравенство

$$2H(a,b,c) \leq H(a,b) + H(a,c \mid b) + H(b,c \mid a)$$

выполнено не для всех троек случайных величин.

3. (a) Существует ли такое распределение вероятностей для тройки (a,b,c), что $I(a;b) = 0$ и $I(a;b \mid c) > 0$? (b) Существует ли такое распределение вероятностей для тройки (a,b,c), что $I(a;b) > 0$ и $I(a;b \mid c) = 0$?

4. Пусть последовательность случайных величин $a - b - c$ образует цепь Маркова. Докажите, что $I(a;c) \leq I(a;b)$ и $I(a;c) \leq I(b;c)$.

5. Пусть последовательность случайных величин $a - b - c - d$ образует цепь Маркова. Докажите, что $I(a;d) \leq I(b;c)$.

6. Пусть энтропия случайной величины a равна n, а взаимная информация пар a и b, а также a и c больше $3n/4$. Докажите, что взаимная информация b и c больше $n/2$.

7. Случайные функции a и b принимают значения в 3-элементном множестве, и $a = b$ с вероятностью $2/3$. Докажите, что $H(a \mid b) < 4/3$.

8. Сколько нужно взвешиваний на чашечных весах, чтобы найти среди 40 монет одну фальшивую, отличающуюся от настоящих по весу (узнавать относительный вес фальшивой монеты не требуется).

9. Имеется набор из n камней. С помощью чашечных весов без гирь можно сравнить по весу любые два камня. Сколько необходимо взвешиваний, чтобы найти одновременно самый тяжелый и самый легкий камень?

10. Приведите пример распределения вероятностей, для которого код Шеннона–Фано не является оптимальным.

11. Рассмотрим задачу разделение секрета для следующей структуры доступа с 4 участниками: минимальными группами участников, знающих секрет, являются четыре пары

$$\{1,2\}, \{1,3\}, \{2,3\}, \{3,4\}.$$

Докажите, что для данной структуры доступа невозможно идеальное разделение секрета.

12. Докажите, что энтропийный профиль

$$\begin{aligned} &H(a \mid b,c) = H(b \mid a,c) = H(c \mid a,b) = 0, \\ &I(a;b \mid c) = I(a;c \mid b) = I(b;c \mid a) = k, \\ &I(a;b;c) = -k, \end{aligned}$$

может быть реализован некоторым распределением (a,b,c), если и только если $k = \log N$ для некоторого целого N.

13*. Пусть для некоторого распределения (a,b,c)

$$I(a;b \mid c) = I(a;c \mid b) = I(b;c \mid a) = 0.$$

Докажите, что найдётся такая случайная величина d, что $H(d) = I(a;b;c)$, и a,b,c независимы относительно d (т.е., у тройки a,b,c можно “выделить” взаимную информацию).

14*. Имеется некоторое совместное распределение случайных величин (a, b, c, x, y). Известно, что

$$I(a; b \mid c) = I(a; c \mid b) = I(b; c \mid a) = 0.$$

Докажите, что в таком случае выполнено неравенство

$$I(a; b) \leq I(a; b \mid x) + I(a; b \mid y) + I(x; y).$$

МФТИ, бакалавриат ФИВТ. 2014.
Упражнения ко второй части курса
***Введение в теорию информации.* А.Е. Ромащенко**
(upd. 11.06.2014)

1. Рассмотрим двоичный канал без памяти, а котором ноль всегда передаётся без искажений, а единица с вероятностью ε превращается в ноль. Вычислите пропускную способность этого канала.

2. Рассмотрим канал без памяти с входным алфавитом $\{0, 1\}$ и выходным алфавитом $\{0, 1, *\}$. Символы 0 и 1 с вероятностью $(1-\varepsilon)$ передаются без искажения, а с вероятностью ε превращаются в $*$. Вычислите пропускную способность этого канала.

3. В некотором канале с шумом (без памяти) входной алфавит состоит из 3 символов, а выходной из 4. (а) Может ли пропускная способность такого канала быть равна 1? (б) Может ли пропускная способность такого канала быть равна 2?

4. (а) Существует ли такой оптимальный алгоритмический способ описания U, что величина колмогоровской сложности $KS(x)$ для любого слова x чётна? (б) Такой, что колмогоровская сложность любого слова x является степенью двойки?

5. Пусть вычислимая функция U является оптимальным способом описания для колмогоровской сложности. Докажите, что отображение V, определяемое для всякого слова p как $V(p) = U(U(p))$, тоже является оптимальным способом описания.

6. Докажите, что множество пар $\{(x, n) \mid KS(x) < n\}$ является перечислимым, но не является разрешимым.

7. Докажите, что существует такая константа C, что для любого палиндрома x длины n выполнено неравенство

$$KS(x) \leq n/2 + c.$$

8. Докажите, что $KS(x, KS(x)) = KS(x) + O(1)$.

9. Докажите неравенства для колмогоровской сложности:
(а) $2KS(x, y, z) \leq KS(x, y) + KS(x, z) + KS(y, z|x) + O(\log(|x| + |y| + |x|))$,
(б) $KS(x, y, z) + KS(z) \leq KS(x, z) + KS(y, z) + O(\log(|x| + |y| + |x|))$,
(в) $KS(z) \leq KS(z \mid x) + KS(z \mid x) + I(x : y) + O(\log(|x| + |y| + |x|))$.

10. (а) Докажите, что для любых чисел C и D найдутся такие слова x, y, z, что

$$2KS(x, y, z) \not\leq KS(x, y) + KS(x, z|y) + KS(y, z|x) + C \log(KS(x, y, z)) + D.$$

(б) Докажите, что для любых чисел C и D найдутся такие слова x, y, z, что

$$I(x : y) \not\leq I(x : y|z) + C \log(KS(x, y, z)) + D.$$

(в) Докажите, что для любых чисел C и D найдутся такие слова x, y, v, w, что

$$I(x : y) + I(v : w) \not\leq I(x : y|v) + I(x : y|w) + C \log(KS(x, y, z)) + D$$

11. (а) Докажите, что ряд $\sum\limits_{x\in\{0,1\}^*} 2^{-KS(x)}$ расходится.

(б) Докажите, что ряд $\sum\limits_{x\in\{0,1\}^*} 2^{-KP(x)}$ сходится.

12. Докажите, что для любого m, для всех достаточно больших n найдется не менее m слов x длины n таких, что $KS(x) \geq n$.

13. Докажите, что для префиксной колмогоровской сложности выполняется неравенство

$$KP(x, y) \leq KP(x) + KP(y) + O(1).$$

14. Докажите, что $KP(x) \leq KS(x) + \log KS(x) + 2 \log\log KS(x) + O(1)$.

15. Докажите, что для любой константы c, для бесконечно многих x

$$KP(x) > KS(x) + c.$$

16. Известно, что для колмогоровской сложности взаимная информация симметрична с логарифмической погрешностью:

$$|I(x : y) - I(y : x)| = O(\log KS(x, y)).$$

(а) Покажите, что $O(\log KS(x, y))$ в этом равенстве нельзя заменить на $O(1)$.

(б) Покажите, что $O(\log KS(x, y))$ в этом равенстве нельзя заменить на $O(\sqrt{\log KS(x, y)})$.

17. (а) Докажите, что если последовательность $x_1x_2x_3\ldots$ случайна по Мартин-Лёфу, то она не может быть вычислимой. (б) Докажите, что если последовательность $x_1x_2x_3\ldots$ случайна по Мартин-Лёфу, то подпоследовательность $x_2x_4x_6\ldots$ не может быть вычислимой. (в) Докажите, что никакая последовательность вида $x_1x_1x_2x_2x_3x_3\ldots$ не может быть случайной по Мартин-Лёфу

18. Докажите, что в любой случайной по Мартин-Лёфу последовательности $x_1x_2x_3\ldots$ встречаются сколь угодно длинные последовательности нулей.

19. В последовательности битов $x_1x_2x_4\ldots x_n\ldots$ в каждом префиксе длине n не встречается больше $\frac{\log n}{2}$ нулей подряд. Докажите, что такая последовательность не может быть случайной по Мартин-Лёфу.

20. Рассмотрим предикат $\mathsf{GT}_n : \{1, \ldots, 2^n\} \times \{1, \ldots, 2^n\} \to \{0, 1\}$

$$\mathsf{GT}_n(x, y) = \begin{cases} 1, & \text{если } x > y, \\ 0, & \text{иначе.} \end{cases}$$

Докажите, что детерминированная коммуникационная сложность предиката GT_n равна $n + 1$.

21. Рассмотрим предикат *дизъюнктности* DISJ_n: для подмножеств $x, y \subset \{1, \ldots, n\}$

$$\mathsf{DISJ}_n(x, y) = \begin{cases} 1, & \text{если } x \cap y = \emptyset, \\ 0, & \text{иначе.} \end{cases}$$

Докажите, что детерминированная коммуникационная сложность предиката DISJ_n равна $n+1$.

22. (а) Докажите, что предикат GT_n из задачи 20 имеет вероятностную коммуникационную сложность $O(\log^2 n)$ (для вероятности ошибки $\varepsilon = 1/3$). (б)* Докажите, что этот предикат имеет вероятностную коммуникационную сложность $O(\log n)$ (также для вероятности ошибки $\varepsilon = 1/3$).

23. Рассмотрим функцию $\mathsf{max}_n : \{0, \ldots, 2^n - 1\} \times \{0, \ldots, 2^n - 1\} \to \{0, 1\}$

$$\mathsf{max}_n(x, y) = \begin{cases} x, & \text{если } x \geq y, \\ y, & \text{иначе.} \end{cases}$$

(а) Докажите, что детерминированная коммуникационная сложность функции max_n не превосходит $2n$. (б) Докажите, что детерминированная коммуникационная сложность функции max_n не превосходит $\frac{3}{2}n + O(1)$. (в)* Докажите, что детерминированная коммуникационная сложность этой функции не превосходит $n + O(\sqrt{n})$.

24. Докажите, что для любого $\varepsilon < 1/2$ вероятностная коммуникационная сложность предиката GT_n из задачи 20 не меньше $\Omega(\log n)$ (константа в $\Omega(\cdot)$ может зависеть от ε).

25. У Алисы имеется n-битная строка x, а у Боба n-битная строка y. Известно, что y получен из x инвертированием одного бита.

(а) Придумайте детерминированный коммуникационный протокол сложности $O(\log n)$, который позволяет Бобу узнать x.

(б) Придумайте *однораундовый* детерминированный коммуникационный протокол сложности $O(\log n)$, который позволяет Бобу узнать x. (В *однораундовом* протоколе Алиса посылает некоторое сообщение Бобу, после чего Боб вычисляет результат.)