InT 1. Докажите, что:

а) код Шеннона–Фано является префиксным;

б) если центральный отрезок относить туда, куда попала его большая часть, то кодирование Шеннона–Фано не является сбалансированным (то есть не существует константы d, для которой выполнено $\ell(c_i) < -\log p_i + d$ для любых k и любых исходных вероятностей $p_1, \ldots, p_k$);

в) если центральный отрезок всегда относить к правой половине, то кодирование Шеннона–Фано также не является сбалансированным.

InT 2. Докажите, что арифметическое кодирование сбалансировано с константой 2.

InT 3. Приведите пример такого распределения вероятностей, что код Шеннона–Фано не является оптимальным.

InT 4. Рассмотрим функцию $\mathsf{IP}\colon \{0,1\}^n \to \{0,1\}$, где $\mathsf{IP}(x,y) := \sum\limits_i x_i y_i \mod 2$. Докажите, что коммуникационная сложность IP равна $n - \mathcal{O}(1)$.

InT 5. 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 $\mathcal{O}(\log n)$, который позволяет Бобу узнать x.

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

InT 6. Докажите, что если $I(x\colon y) = I(x\colon y \mid a) = 0$, то $I(a\colon b) \le I(a\colon b \mid x) + I(a\colon b \mid y)$.

InT 7. Пусть X — случайная величина, распределённая на $\{0,1\}^n$. Докажите, что для любого распределения S на подмножествах $[n]$, при котором $\Pr[i \in S] \ge \mu$, выполняется $\mathop{\mathbb{E}}\limits_S[\mathrm{H}(X_S)] \ge \mu \cdot \mathrm{H}(X)$, где X_S — проекция X на координаты из множества S.

Подсказка: попробуйте применить chain rule.