



**INT 1.** Пусть последовательность  $\alpha \rightarrow \beta \rightarrow \gamma$  образует цепь Маркова. Докажите, что:

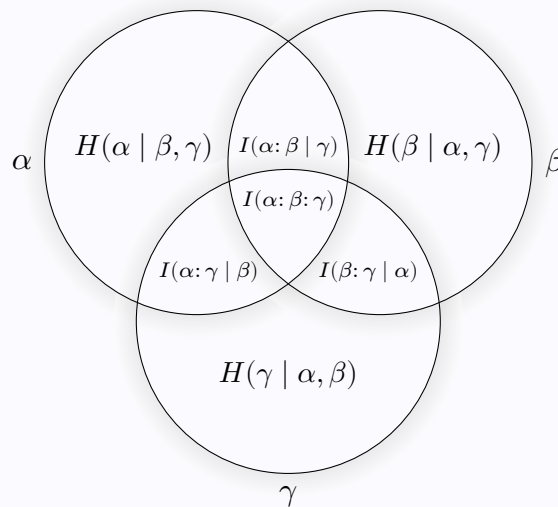
- а)  $I(\alpha: \gamma) \leq I(\alpha: \beta)$ ;
- б)  $I(\alpha: \gamma) \leq I(\beta: \gamma)$ .

### Определение

Определим **общую информацию** трёх случайных величин:

$$I(\alpha: \beta: \gamma) := I(\alpha: \beta) - I(\alpha: \beta | \gamma).$$

Соотношения на информационные величины имеют удобную геометрическую интерпретацию. При помощи диаграмм Эйлера можно сопоставить площади каждой из получившихся замкнутых области некоторую информационную величину. В частности, площадь каждого круга соответствует энтропии указанной случайной величины.



**INT 2.** Постройте три таких случайных величины  $\alpha, \beta, \gamma$ , что  $I(\alpha: \beta: \gamma) < 0$ .

**INT 3.** Пусть энтропия случайной величины  $a$  равна  $n$ , а взаимная информация пар  $a$  и  $b$ , а также  $a$  и  $c$  больше  $3n/4$ . Докажите, что  $I(b: c) > n/2$ .

### Определение

**Коммуникационный протокол** для функции  $f: X \times Y \rightarrow Z$  — это корневое двоичное дерево, которое описывает совместное вычисление Алисой и Бобом функции  $f$ . В этом дереве каждая внутренняя вершина  $v$  помечена меткой  $a$  или  $b$ , означающей очередь хода Алисы или Боба соответственно. Для каждой вершины, помеченной  $a$ , определена функция  $g_v: X \rightarrow \{0, 1\}$ , которая говорит Алисе, какой бит нужно послать, если вычисление находится в этой вершине. Аналогично, для каждой вершины  $v$  с меткой  $b$  определена функция  $h_v: Y \rightarrow \{0, 1\}$ , которая определяет бит, который Боб должен отослать в этой вершине. Каждая внутренняя вершина имеет двух потомков, ребро к первому потомку помечено 0, а ребро ко второму потомку помечено 1. Каждый лист помечен значением из множества  $Z$ .

Каждая пара входов  $(x, y)$  определяет путь от корня до листа в описанном двоичном дереве естественным образом. Будем говорить, что коммуникационный протокол вычисляет функцию  $f$ , если для всех пар  $(x, y) \in X \times Y$ , этот путь заканчивается в листе с пометкой  $f(x, y)$ .

**Коммуникационной сложностью** функции  $f$  назовем наименьшую глубину протокола, вычисляющего функцию  $f$ , и будем ее обозначать  $D(f)$ .



**INT 4.** Пусть  $f: X \times Y \rightarrow \{0, 1\}$ . Рассмотрим матрицу  $M^f$ , строки которой соответствуют элементам множества  $X$ , столбцы элементам множества  $Y$ , а в ячейке  $M_{x,y}$  написано значение  $f(x, y)$ . Покажите, что:

- а)  $\text{rk}_{\mathbb{R}}(M_{x,y})$  не превосходит числа листьев в коммуникационном протоколе для  $f$  (*подсказка*: для вершины протокола опишите множество входов, которое приводит в нее);  
б)  $D(\text{EQ}_n) \geq n$ , где  $\text{EQ}_n: \{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\}$  и  $\text{EQ}_n(x, y) = 1 \Leftrightarrow x = y$ .

**INT 5.** Рассмотрим функцию  $\text{SUM}: \{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\}^{n+1}$ , возвращающую сумму чисел в двоичной системе счисления. Докажите, что  $D(\text{SUM}) = 2n$ .

**INT 6.** Покажите, что  $D(\text{Med}) = \mathcal{O}(\log^2 n)$ , где  $x$  и  $y$  это характеристические функции подмножеств  $[n]$ , а  $\text{Med}(x, y)$  — медиана мультимножества  $x \cup y$  (если элемент встречается и в  $x$  и в  $y$ , то считаем его дважды).

*Комментарий:* на самом деле  $D(\text{Med}) = \Theta(\log n)$ .

**INT 7.** Пусть  $\{a_1, a_2, \dots, a_n\}$  — произвольный алфавит и  $p_1, p_2, \dots, p_n$  — вероятности букв этого алфавита. Докажите, что для любого инъективного кодирования букв этого алфавита средняя длина кода не меньше  $h - 2 \log(h + 1) - 2$ , где  $h$  — энтропия распределения с вероятностями  $p_1, p_2, \dots, p_n$ .