Энтропия Шеннона. Коды. (3-й курс)

12 марта 2017 г.

1. Докажите, что $I(f(\alpha) : \beta) \leq I(\alpha : \beta)$ для любой функции f.

2. Докажите, что величины α, β, γ независимы в совокупности(вероятность события $(\alpha = \alpha_i, \beta = \beta_j, \gamma = \gamma_k)$ равна произведению трех отдельных вероятностей) тогда и только тогда, когда
$$H(\alpha, \beta, \gamma) = H(\alpha) + H(\beta) + H(\gamma).$$

3. Докажите, что $I((\alpha, \beta) : \gamma) \geq I(\alpha : \gamma)$.

4. Докажите, что
$$I((\alpha, \beta) : \gamma) = I(\alpha : \gamma) + I(\beta : \gamma|\alpha).$$

5. Докажите, что если $I(\alpha : \gamma|\beta) = 0$, то $I(\alpha : \gamma) \leq I(\alpha : \beta)$, а значит и $I(\alpha : \gamma) \leq H(\beta)$.

6. Докажите, что $I((\alpha, \beta) : \gamma) \geq I(\alpha : \gamma)$ и что разность между левой и правой частями равна $I(\beta : \gamma|\alpha)$.

7. Докажите неравенство $2H(\alpha, \beta, \gamma) \leq H(\alpha, \beta) + H(\alpha, \gamma) + H(\beta, \gamma)$

8. Докажите следующее обобщение предыдущего неравенства. Пусть $T_1, ..., T_k$ — произвольные кортежи, составленные из переменных $\alpha_1, \dots \alpha_n$, причем каждая переменная входит ровно в r кортежей. Тогда $rH(\alpha_1, \dots, \alpha_n) \leq H(T_1) + H(T_2) + \dots + H(T_k)$ (неравенство Шерера(Shearer)).

9. Пусть $\{a_1, a_2, \dots a_n\}$ — произвольный алфавит и $p_1, p_2 \dots p_n$ — внроятности букв этого алфавита. Докажите, что для любого инъективного кодирования букв этого алфавита средняя длина кода не меньше $H - 2\log H - 2$. H — это энтропия распределения с вероятностями $p_1, p_2 \dots p_n$.

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

11. Докажите, что константу 2 в предыдущей задаче нельзя понизить, даже в предположении, что $p_1, \dots p_n$ упорядочены по величине.

12. (а) Докажите, что код Шеннона–Фано является префиксным.

    (б) Докажите, что 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 $l(c_i) < -\log p_i + d$ для любых k и любых исходных вероятностей $p_1, ..., p_k$).

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

13. Докажите, что кодирование Хаффмана не является сбалансированным.

Минимальное количество задач для зачета по домашнему заданию: энтропия 4, коды 2.