Энтропия Шеннона. Коды. (5-й курс)

12 марта 2017 г.

1. Докажите, что энтропия Шеннона не меньше минимальной энтропии, определяемой как $H_{min} = \min_i(-\log p_i)$.

2. Пусть вероятности исходов случайной величины есть $1/2, 1/4, 1/8, ..., 1/2^n, 1/2^n$. К чему стремится ее энтропия, при $n \to \infty$? Тот же вопрос для случайной величины с вероятностями исходов $1/3, 1/3, 1/9, 1/9, ..., 1/3^n, 1/3^n, 1/3^n$.

3. Докажите формулу $h(p_1, \dots p_n, q_1, \dots q_m) = h(p_1 + \dots + p_n, q_1 + \dots + q_m) + (p_1 + \dots + p_n)h(p'_1, \dots p'_n) + (q_1 + \dots + q_m)h(q'_1, \dots q'_m)$ где $p'_i = \frac{p_i}{p_1 + \dots p_n}$ и $q'_i = \frac{q_i}{q_1 + \dots q_n}$.

4. Покажите, что величина $H(\zeta|A)$ может быть и больше и меньше величины $H(\zeta)$.

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

6. 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

$$H(\alpha, \beta, \gamma) = H(\alpha) + H(\beta) + H(\gamma).$$

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

8. Докажите, что

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

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

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

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

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

13. Существуют однозначно декодируемые коды, не являюющиеся префиксными, привидите пример.

14. Укажите явно взаимно-однозначное соответствие между множеством бесконечных последовательностей цифр 0,1,2 и множеством бескнечных последовательностей нулей и единиц.

15. Пусть слова $c_1, c_2, \ldots c_k$ и $d_1, d_2, \ldots d_k$ образуют префиксный код(по отдельности). Покажите, что kl слов $c_i d_j$ (приписываем одно слово к другому без разделителя) также образуют префиксный код.

16. Пусть целое число x выбирается случайным образом в интервале от 1 до 1000(все возможные значения x равновероятны). Докажите, что любой алгоритм, который с помощью вопросов с ответами ДА/НЕТ находит x, задает в среднем не меньше $\log 1000$ вопросов.

17. Докажите, что любое инъективное кодирование можно преобразовать в префиксное ценой небольшого увеличения средней длины кода: если у исходного кода средняя длина была ℓ, то у нового она будет не больше $\ell + 2\log \ell + 2$.

18. Пусть $\{a_1, a_2, \ldots a_n\}$ — произвольный алфавит и $p_1, p_2 \ldots p_n$ — 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 $p_1, p_2 \ldots p_n$.

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

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

21. (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

22. Докажите, что кодирование Хаффмана не является сбалансированным.

Минимальное количество задач для зачета по домашнему заданию: энтропия 8, коды 7.