Секция «Математическая логика, алгебра и теория чисел»

Некоторые свойства антистохастических слов

Милованов Алексей Сергеевич

Аспирант

Московский государственный университет имени М.В.Ломоносова, Механико-математический факультет, Кафедра математической логики и теории алгоритмов, Москва, Россия

E-mail: almas239@gmail.com

Алгоритмическая статистика изучает объяснения для наблюдаемой информации, которые является хорошими в алгоритмическом смысле: объяснения должны с одной стороны быть простыми, с другой стороны - выявлять все алгоритмические закономерности, которые содержаться в наблюдаемой информации. Будем считать, что наблюдаемая информация закодирована как конечная бинарная строка x. В качестве объяснений рассматриваются конечные множества содержащие x. Пусть множество $A \ni x$. Колмогоров предложил[1] мерить качество объяснения A для x с помощью двух параметров - Колмогоровской сложности множества C(A) и его мощности $|A|$. Оба эти параметра не могут быть маленькими одновременно, т. к. $C(A) + \log |A| \gtrsim C(x)$. Колмогоров называл слово x *стохастическим*, если для него существует множество $A \ni x$, такое что $C(A) \approx 0$ и $\log |A| = C(x)$. Примером стохастической строки является случайная строка длины n, т. е. такая строка x длины n, сложность которой приблизительно равняется n. Сразу же возникает вопрос: существуют ли нестохастические строки? Положительный ответ был дан Шенем в [2]. Более подробно нестохастические строки были изучены Верещагиным и Витаньей в [4]. В частности ими было показано, что существуют в некотором смысле самые нестохастические строки, которые называются *антистохастическими*. В нашей работе антистохастические строки изучаются более подробно. Основным нашим результатом является следующая:

Теорема 1. *Пусть антистохастическая строка x длины n и сложности k содержится во множестве A размера 2^{n-k}. Тогда $C(x|A) = O(\log C(A) + \log n)$*

(Через $C(x|A)$ мы обозначили сложность x при условии A.) Неформальный смысл этой теоремы заключается в следующем: антистохастическую строку сложности k можно восстановить по любыми содержащимся в ней k битам информации (например, по первым или последним k битам). Можно показать, что строк длины n с похожим свойством восстановления довольно много, а именно 2^k (это следует из того, что сложность антистохастической строки равняется k). А из этого следует существование кода исправляющего стирания списком с оптимальными параметрами. Этот результат можно доказать и вероятностным методом[3].

Источники и литература

1) А. Н. Колмогоров, Сложность алгоритмов и объективное определение случайности. Представленно 16.04.1974 на заседании Московского Математического Общества // Успехи математических наук, 29 :4 (178), p. 155, 1974.

2) А. Х. Шень. Понятие (alpha,beta)-стохастичности по Колмогорову и его свойства // Доклады Академии наук СССР, 271(6), 1337-1340.

3) V. Guruswami, List decoding of error-correcting codes: winning thesis of the 2002 ACM doctoral dissertation competition, Springer, 2004.

4) N. Vereshchagin and P. Vit'anyi, Kolmogorov's Structure Functions with an Application to the Foundations of Model Selection // IEEE Transactions on Information Theory 50:12 (2004) 3265-3290. Preliminary version: Proc. 47th IEEE Symp. Found. Comput. Sci., 2002, 751–760.}