А. Е. Ромащенко, Пары слов с нематериализуемой взаимной информацией, *Пробл. передачи информ.*, 2000, том 36, выпуск 1, 3–20

ПРОБЛЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Том 36 2000 Вып. 1

УДК 621.391.1:519.722:510.5

© 2000 г. **А.Е. Ромащенко**

ПАРЫ СЛОВ С НЕМАТЕРИАЛИЗУЕМОЙ ВЗАИМНОЙ ИНФОРМАЦИЕЙ

Пусть имеется пара слов $\langle a, b\rangle$ с достаточно большой взаимной информацией. Всегда ли можно «материализовать» эту информацию, т.е. указать такое слово c, которое легко вычисляется по a и b, и колмогоровская сложность которого равна взаимной информации a и b? В статье улучшается оценка на величину общей информации, которую можно материализовать для слов из конструкции Гача и Кёрнера, а также приводится новый метод построения пар слов с нематериализуемой взаимной информацией.

§ 1. Введение

Пусть имеются два файла: x и y. Предположим, что между записанными в них данными есть некоторая корреляция. Если нужно хранить (или передавать) информацию об x и y наиболее экономным способом, разумно потребовать, чтобы «общая» информация x и y хранилась в виде отдельного набора данных. Возникает вопрос: можно ли закодировать x в виде пары файлов $\langle u, v\rangle$, а y в виде пары $\langle u, w\rangle$ так, чтобы длина u соответствовала величине взаимной информации двух исходных файлов. Тогда можно было бы вместо пары $\langle x, y\rangle$ хранить тройку $\langle u, v, w\rangle$. При этом u должно легко вычисляться как по x, так и по y.

Сформулируем тот же вопрос более формально. Если даны слова x и y, можно рассмотреть колмогоровские сложности каждого из них и их взаимную информацию. Взаимная информация $I(x : y)$ показывает, насколько знание одного слова упрощает задачу порождения другого:

$$I(x : y) := K(y) - K(y \mid x).$$

Иногда этой величине удается дать наглядную интерпретацию. Рассмотрим простейший пример. Пусть слово x является конкатенацией слов u и v, а слово y – конкатенацией слов u и w.

$$x = uv, \quad y = uw. \tag{1}$$

Предположим, что все три слова u, v и w выбраны случайными и независимыми, и их длины равны n. Тогда сложности слов x и y равны $2n$, а их взаимная информация равна n. Это соответствует интуиции: у слов x и y есть общая часть u, и взаимная информация равна сложности этой общей части. Слово u является «материализацией» взаимной информации x и y.

В [1] был поставлен вопрос: всегда ли для двух слов можно найти третье, которое материализовало бы их взаимную информацию? Слово z, материализующее взаимную информацию x и y, должно легко вычисляться по каждому из них. Таким образом, мы интересуемся, всегда ли есть такое слово, которое имеет малую сложность относительно каждого из двух данных, и сложность которого равна их взаимной информации.

Как показали Гач и Кёрнер [1], ответ на этот вопрос – отрицательный. Однако, чтобы сформулировать точное утверждение, необходимо пояснить, что значит, что некоторое слово z легко получить по слову x и по слову y (т.е. условные колмогоровские сложности $K(z|x)$ и $K(z|y)$ малы). Кажется невозможным определить отношение «условной простоты» для индивидуальных слов (нельзя провести границу между «простыми» и «сложными» словами). Поэтому перейдем от индивидуальных слов к бесконечным последовательностям слов и будем рассматривать асимптотические свойства их сложностей. В таких терминах можно сформулировать результат Гача и Кёрнера.

Теорема 1. *Существуют такие последовательности слов x_n, y_n, что*

$$K(x_n) = n + o(n), \quad K(y_n) = n + o(n), \quad I(x_n : y_n) = an + o(n),$$

(a – положительная константа), и для любой последовательности слов z_n, удовлетворяющей условию

$$K(z_n|x_n) = o(n), \quad K(z_n|y_n) = o(n),$$

выполнено $K(z_n) = o(n)$.

Неформально говоря, данная теорема утверждает, что существуют такие x_n и y_n, что если сложность z_n мала относительно x_n и относительно y_n, то и сложность самого z_n мала. При этом взаимная информация слов x_n и y_n растет линейно по n. Таким образом, существуют последовательности слов, у которых нельзя материализовать взаимную информацию. Более того, теорема 1 утверждает, что нельзя материализовать даже часть взаимной информации x_n и y_n. Точнее, величина материализуемой взаимной информации бесконечно мала по сравнению с n.

В работе [1] описывается некоторый класс примеров пар $\langle x_n, y_n \rangle$, обладающих сформулированным выше свойством. При этом конструкция позволяет строить такие последовательности x_n, y_n для любых значений параметра a, $0 < a < 1$, т.е. можно указать такие x_n и y_n, взаимная информация которых очень велика (a близко к единице), но даже ее малая часть не может быть материализована.

В [1] не проводилась точная оценка остаточных членов. Но, анализируя доказательство, можно проверить, что теорема 1 останется верной, если в формулировке заменить члены $o(n)$ на $O(\sqrt{n})$ (или на $O(f(n))$, где $f(n)$ – любая функция, растущая быстрее $\sqrt{n}$, но медленнее n: $f(n) = o(n)$, $f(n) \geq \sqrt{n}$). Однако в утверждениях о колмогоровской сложности естественно формулировать равенства с точностью до логарифмического члена. Действительно, с точностью до логарифма длин слов выполнены такие свойства, как симметричность взаимной информации

$$I(x : y) = I(y : x) + O(\log(|x| + |y|))$$

или связь условной сложности и сложности пары слов

$$K(\langle x, y \rangle) = K(x) + K(y|x) + O(\log(|x| + |y|))$$

(см. [2, 3]). Поэтому кажется интересным рассмотреть усиление теоремы 1, а именно, доказать ее, заменив в формулировке члены $o(n)$ на $O(\log n)$. Более формально, естественным усилением теоремы Гача и Кёрнера является

Теорема 2. *Для любой функции $f(n)$ такой, что $f(n) = o(n)$ и $f(n) \geq \log n$, существуют такие последовательности слов x_n, y_n, что*

$$K(x_n) = n + O(f(n)), \quad K(y_n) = n + O(f(n)), \quad I(x_n : y_n) = an + O(f(n))$$

(a – некоторая положительная константа), и для любой последовательности слов z_n, удовлетворяющей условию

$$K(z_n|x_n) = O(f(n)), \quad K(z_n|y_n) = O(f(n)),$$

выполнено $K(z_n) = O(f(n))$.

В работах [4, 5] было получено доказательство теоремы 2 для произвольного значения параметра a, $0 < a < 1$. Другие примеры пар последовательностей слов с

нематериализуемой общей информацией (и, следовательно, другие доказательства теоремы 2 для некоторых специальных значений a) приводились также в [6, 7].

Таким образом, для любого $a < 1$ можно найти такие слова x_n и y_n, что сложности слов примерно равны n, их взаимная информация примерно равна an, и их взаимную информацию нельзя материализовать. Ан.А. Мучником был поставлен вопрос: при каких значениях параметра a для каждого x_n сложности n можно подобрать y_n сложности n такое, что взаимная информация $I(x_n : y_n)$ примерно равна an, но ее нельзя материализовать? Более точно, для каких значений параметра a имеет место следующее усиление теоремы 2.

Теорема 3. *Пусть $f(n)$ – такая функция, что $f(n) = o(n)$ и $f(n) \geq \log n$, и x_n – такая последовательность, что $K(x_n) = n + O(f(n))$. Тогда существует последовательность y_n такая, что*

$$K(y_n) = n + O(f(n)), \quad I(x_n : y_n) = an + O(f(n)),$$

и для любой последовательности слов z_n, удовлетворяющей условию

$$K(z_n|x_n) = O(f(n)), \quad K(z_n|y_n) = O(f(n)),$$

выполнено $K(z_n) = O(f(n))$.

Для $a = 1/2$ данная теорема была доказана в [5].

В данной работе рассматриваются два рассуждения, позволяющие доказывать теорему 3 для любых значений параметра a, $0 < a < 1$. В первом из них используется пример пар слов, предложенный в [1]; новый метод доказательства позволяет улучшить оценку на величину материализуемой общей информации (и, тем самым, доказать теоремы 2 и 3). Второе рассуждение основано на новой алгебраической конструкции, обобщающей метод, использованный в [5].

Замечание 1. Для простоты будем доказывать теорему 3 только для $f(n) = \log n$. На случай произвольной $f(n)$ все рассуждения переносятся почти дословно.

Используемые обозначения

x, y, z, ... – двоичные слова (конечные последовательности нулей и единиц); длину слова x будем обозначать $|x|$;

$\boldsymbol{x} = \{x_n\}$, $\boldsymbol{y} = \{y_n\}$, $\boldsymbol{z} = \{z_n\}$, ... – бесконечные последовательности слов; будем предполагать, что во всех рассматриваемых последовательностях длина слов растет не более чем линейно: $|x_n| = O(n)$, $|y_n| = O(n)$, $|z_n| = O(n)$, ...; будем говорить, что последовательность $\boldsymbol{x} = \{x_n\}$ проста относительно последовательности $\boldsymbol{y} = \{y_n\}$, если $K(x_n|y_n) = O(\log n)$;

$\langle x_1, x_2, \ldots, x_n\rangle$ – кортеж двоичных слов; считаем фиксированной некоторую вычислимую нумерацию всех конечных кортежей слов;

α, β, γ, ... – дискретные случайные величины;

$K(x)$ – колмогоровская сложность слова x;

$K(x_1, x_2, \ldots, x_n)$ – колмогоровская сложность номера кортежа слов $\langle x_1, x_2, \ldots, x_n\rangle$ в выбранной нумерации;

$K(x|y)$ – колмогоровская сложность слова x относительно слова y;

$K(x_1, x_2, \ldots, x_n|y_1, y_2, \ldots, y_m)$ – колмогоровская сложность номера кортежа $\langle x_1, x_2, \ldots, x_n\rangle$ относительно номера кортежа $\langle y_1, y_2, \ldots, y_m\rangle$;

$I(x : y) := K(y) - K(y|x)$ – взаимная информация слов x и y;

$I(x : y|z) := K(y|z) - K(y|x, z)$ – взаимная информация слов x и y относительно z; аналогично определяется взаимная информация кортежей слов (относительно кортежей слов);

$H(\alpha)$ – энтропия Шеннона случайной величины α;

$\mathcal{I}(\alpha : \beta) := H(\alpha) + H(\beta) - H(\alpha, \beta)$ – взаимная информация случайных величин α и β;

· пусть $\langle x_1, x_2, \ldots, x_k\rangle$ – кортеж слов, $V = \{i_1, i_2, \ldots, i_r\} \subseteq \{1, 2, \ldots, k\}$ – набор индексов; тогда будем обозначать x^V кортеж $\langle x_{i_1}, x_{i_2}, \ldots, x_{i_r}\rangle$; аналогичное обозначение будем использовать для кортежей случайных величин;

все логарифмы в статье берутся по основанию 2.

§ 2. Стохастические пары

Мы рассмотрим частный случай из семейства примеров Гача и Кёрнера и приведем для него новое более простое доказательство нематериализуемости общей информации. Новое рассуждение позволит улучшить оценку на величину выделяемой информации и получить доказательства теорем 2 и 3.

Перед началом доказательства дадим несколько определений.

Определение 1. *Пусть случайные величины $\varphi^1, \varphi^2, \ldots, \varphi^k$ принимают значения в конечных алфавитах $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_k$ и имеют совместное распределение*

$$P(a^1, a^2, \ldots, a^k) = \mathrm{Prob}[\varphi^1 = a^1, \varphi^2 = a^2, \ldots, \varphi^k = a^k].$$

Тогда кортеж бесконечных последовательностей слов $\langle \boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^k\rangle$ назовем P-типичным, если для любого набора значений $\langle a^1, a^2, \ldots, a^k\rangle$ число таких позиций i, что в каждом слове x_n^j на i-м месте стоит буква a^j, равно $nP(a^1, a^2, \ldots, a^k) + +O(1)$.

Определение 2. *Будем называть P-типичный кортеж последовательностей слов $\langle \boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^k\rangle$ P-случайным, если для любого непустого набора индексов $V \subseteq \{1, 2 \ldots, k\}$ и любого (быть может пустого) набора индексов $W \subseteq \{1, 2 \ldots, k\}$ выполнено равенство*

$$K(x_n^V | x_n^W) = nH(\varphi^V | \varphi^W) + O(\log n). \tag{2}$$

(В случае, когда набор индексов W пуст, условная колмогоровская сложность и условная шенноновская энтропия в равенстве (2) обращаются в безусловные.)

Замечание 2. Пусть пара последовательностей $\langle \boldsymbol{x}, \boldsymbol{y}\rangle$ случайна относительно распределения P пары случайных величин $\langle \varphi, \psi\rangle$. Тогда величины $K(x_n)$, $K(y_n)$ и $K(x_n, y_n)$ равны (с точностью до логарифмического слагаемого) $nH(\varphi)$, $nH(\psi)$ и $nH(\varphi, \psi)$ соответственно. Следовательно, аналогичное равенство выполняется и для взаимной информации. В самом деле,

$$\mathcal{I}(\varphi : \psi) = H(\varphi) + H(\psi) - H(\varphi, \psi)$$

и

$$I(x_n : y_n) = K(x_n) + K(y_n) - K(x_n, y_n) + O(\log n).$$

Следовательно,

$$I(x_n : y_n) = n\mathcal{I}(\varphi : \psi) + O(\log n).$$

Утверждение 1. 1) *Если случайные величины $\varphi^1, \varphi^2, \ldots, \varphi^k$ имеют некоторое совместное распределение P, то для любых P-типичных последовательностей $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^k$, любого непустого набора индексов $V \subseteq \{1, 2, \ldots, k\}$ и любого (быть может пустого) набора индексов $W \subseteq \{1, 2, \ldots, k\}$ выполнено неравенство*

$$K(x_n^V | x_n^W) \leq nH(\varphi^V | \varphi^W) + O(\log n).$$

2) *Для любого распределения P существуют P-случайные кортежи.*

Доказательство. Для $k = 1$ доказательство утверждения приведено в [2], а также в [3]. В общем случае доказательство аналогично. ▲

Замечание 3. Пусть пара последовательностей $\langle \boldsymbol{x}, \boldsymbol{y}\rangle$ типична относительно распределения P пары случайных величин $\langle \varphi, \psi\rangle$, и $K(x_n, y_n) = nH(\varphi, \psi) + O(\log n)$. Тогда данная пара P-случайна. Докажем, например, что $K(x_n) = nH(\varphi) + O(\log n)$.

В самом деле, согласно утверждению 1 выполнено неравенство, $K(x_n) \leq nH(\varphi) + +O(\log n)$. Чтобы доказать обратное неравенство, достаточно сложить равенства

$$K(x_n, y_n) = K(x_n) + K(y_n|x_n) + O(\log n),$$
$$nH(\varphi, \psi) = K(x_n, y_n) + O(\log n)$$

и неравенство

$$K(y_n|x_n) \leq nH(\psi|\varphi) + O(\log n).$$

Аналогично можно доказать, что выполнены равенства

$$K(y_n) = nH(\psi) + O(\log n),$$
$$K(x_n|y_n) = nH(\varphi|\psi) + O(\log n),$$
$$K(y_n|x_n) = nH(\psi|\varphi) + O(\log n).$$

В дальнейшем нам потребуется конструкция, позволяющая расширять P-случайные кортежи. А именно, пусть k, l – натуральные числа, и $l < k$. Будем считать, что случайные величины $\varphi^1, \varphi^2, \ldots, \varphi^k$ имеют совместное распределение P. Обозначим через P' проекцию распределения P на первые l координат, т.е. совместное распределение величин $\varphi^1, \varphi^2, \ldots, \varphi^l$. Тогда любой P'-случайный кортеж можно дополнить до P-случайного. Более точно, выполнена

Лемма 1. *Пусть P – совместное распределение k случайных величин, каждая из которых принимает два значения 0 или 1, P' – проекция распределения P на первые l координат, и кортеж $\langle \boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^l \rangle$ является P'-случайным. Тогда существуют такие последовательности $\boldsymbol{x}^{l+1}, \boldsymbol{x}^{l+2}, \ldots, \boldsymbol{x}^k$, что кортеж $\langle \boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^k \rangle$ является P-случайным.*

Доказательство. Проведем доказательство для случая $k = 2, l = 1$ (доказательство в общем случае аналогично).

Пусть двоичные случайные величины φ, ψ имеют совместное распределение P:

$$\mathrm{Prob}[\varphi = i, \psi = j] = p_{ij},$$

где $i, j = 0, 1$. В данном случае P' – это распределение φ:

$$\mathrm{Prob}[\varphi = i] = p_{i0} + p_{i1}.$$

Пусть дана P'-случайная последовательность слов $\boldsymbol{x}$. Требуется подобрать такую последовательность $\boldsymbol{y}$, чтобы пара $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ была P-случайной.

Пусть в слове x_n символы 0 и 1 встречаются s_0 и s_1 раз соответственно (величины s_0 и s_1 зависят от n). Поскольку последовательность $\boldsymbol{x}$ является P'-случайной,

$$s_i = n(p_{i0} + p_{i1}) + O(1)$$

при $i = 0, 1$.

Следовательно, можно представить числа s_0 и s_1 в виде сумм $s_i = s_{i0} + s_{i1}$ так, что

$$s_{ij} = np_{ij} + O(1)$$

(числа s_{ij} также будут зависеть от n).

Будем называть двоичное слово $\hat{y}$ длины n допустимым, если его можно получить из слова x_n следующим преобразованием: в слове x_n нужно заменить $(s_{01} + s_{10})$ битов на противоположные, а именно, заменить s_{01} нулей на единицы, и s_{10} единиц на нули.

Ясно, что если $\hat{y}$ допустимо, то частоты нулей и единиц в этом слове соответствуют распределению вероятностей ψ, и пара $\langle x_n, \hat{y} \rangle$ является P-типичной. Остается выбрать такое допустимое $\hat{y}$, чтобы полученная пара была P-случайной. Но для этого достаточно взять допустимое слово, имеющее самую большую сложность относительно x_n.

Для фиксированного x_n число всех допустимых слов $\hat{y}$ равно $(C_{s_0}^{s_{01}} \cdot C_{s_1}^{s_{10}})$. Если в качестве y_n взять допустимое слово, имеющее максимальную сложность относительно x_n, то

$$K(\hat{y}_n|x_n) = \log(C_{s_0}^{s_{01}} \cdot C_{s_1}^{s_{10}}) + O(\log n) = nH(\psi|\varphi) + O(\log n).$$

(Последнее равенство легко доказать, оценивая величины биномиальных коэффициентов с помощью формулы Стирлинга.) Далее, поскольку последовательность $\boldsymbol{x}$ является P-случайной, $K(x_n) = nH(\varphi) + O(\log n)$. Учитывая равенство

$$K(x_n, y_n) = K(x_n) + K(y_n|x_n) + O(\log n),$$

получаем

$$K(x_n, y_n) = nH(\varphi) + nH(\psi|\varphi) + O(\log n) = nH(\varphi, \psi) + O(\log n). \tag{3}$$

Согласно замечанию 3 из равенства (3) вытекает P-случайность построенной пары. ▲

Лемма 1 верна не только для двоичных, но и для произвольных совместно распределенных случайных величин. Однако мы ее доказали (и будем применять) только для совместных распределений двоичных случайных величин.

Введем еще одно понятие. Рассмотрим совместное распределение пары случайных величин $\langle \varphi, \psi \rangle$ со следующими свойствами: обе величины φ, ψ принимают значения 0 и 1 с вероятностью $1/2$ (т.е. являются двоичными равномерно распределенными); при этом φ и ψ принимают разные значения с вероятностью α (и, соответственно, совпадают с вероятностью $(1-\alpha)$). Таким образом,

$$\mathrm{Prob}[\varphi = 0, \psi = 0] = \mathrm{Prob}[\varphi = 1, \psi = 1] = \frac{1-\alpha}{2},$$

$$\mathrm{Prob}[\varphi = 1, \psi = 0] = \mathrm{Prob}[\varphi = 0, \psi = 1] = \frac{\alpha}{2}.$$

Данное распределение задается табл. 1.

Таблица 1

Распределение P пары случайных величин φ, ψ

φ \ ψ	0	1
0	$\frac{1-\alpha}{2}$	$\frac{\alpha}{2}$
1	$\frac{\alpha}{2}$	$\frac{1-\alpha}{2}$

Определение 3. *Назовем последовательности слов $\boldsymbol{x}$, $\boldsymbol{y}$ α-парой, если они являются P-случайной парой относительно распределения P, заданного табл.* 1.

Пусть случайные величины φ и ψ имеют совместное распределение P, указанное в табл. 1, а последовательности $\boldsymbol{x}$ и $\boldsymbol{y}$ образуют α-пару. Тогда x_n, y_n являются случайными словами длины n и отличаются друг от друга в $\alpha n + O(1)$ позициях.

Далее, нетрудно вычислить шенноновские энтропии случайных величин φ и ψ, а также энтропию пары $\langle \varphi, \psi \rangle$

$$H(\varphi) = H(\psi) = 1,$$

$$H(\varphi, \psi) = 1 - (\alpha \log \alpha + (1-\alpha)\log(1-\alpha)).$$

Получаем взаимную информацию данных случайных величин

$$\mathcal{I}(\varphi : \psi) = 1 + (\alpha \log \alpha + (1-\alpha)\log(1-\alpha)).$$

Обозначим величину взаимной информации

$$c(\alpha) = 1 + (\alpha \log \alpha + (1-\alpha)\log(1-\alpha)). \qquad (4)$$

Очевидно, $c(\alpha) > 0$ при $\alpha \neq 1/2$.

Если $\boldsymbol{x}$ и $\boldsymbol{y}$ образуют α-пару, то

$$K(x_n) = n + O(\log n),$$
$$K(y_n) = n + O(\log n),$$
$$I(x_n, y_n) = c(\alpha)n + O(\log n).$$

Таким образом, при $\alpha \neq 1/2$ взаимная информация x_n и y_n растет линейно по n.

Следует выделить случай $\alpha = 1/2$. Поскольку $c(1/2) = 0$, взаимная информация слов x_n и y_n равна $O(\log n)$. Это соответствует интуиции: если два слова выбраны случайно и независимо, то они отличаются примерно в половине битов. Но согласно определению 1/2-пары слова x_n и y_n как раз и должны быть парой случайных слов, отличающихся примерно в половине битов.

Именно α-пары оказываются примерами слов, удовлетворяющих теореме 2. Мы будем доказывать, что для любого $\alpha \in (0;1)$ взаимная информация случайных α-пар не материализуется. Для этого нам понадобятся следующие технические леммы.

Лемма 2. *Пусть последовательность $\boldsymbol{z}$ проста относительно $\boldsymbol{x}$ и относительно $\boldsymbol{y}$, т.е. $K(z_n|x_n) = O(\log n)$ и $K(z_n|y_n) = O(\log n)$. Тогда $K(z_n) \leq I(x_n : y_n) + O(\log n)$.*

Доказательство. Для любых x_n, y_n, z_n имеют место соотношения

$$K(x_n, z_n) = K(x_n) + K(z_n|x_n) + O(\log n),$$
$$K(x_n) - K(x_n|y_n) = I(x_n : y_n) + O(\log n),$$
$$K(x_n|y_n) \leq K(x_n|z_n) + K(z_n|y_n) + O(\log n),$$
$$K(x_n|z_n) + K(z_n) = K(x_n, z_n) + O(\log n).$$

Складывая их, получаем

$$K(z_n) \leq K(z_n|x_n) + K(z_n|y_n) + I(x_n : y_n) + O(\log n). \qquad (5)$$

Учитывая $K(z_n|x_n) = O(\log n)$ и $K(z_n|y_n) = O(\log n)$, получаем $K(z_n) \leq I(x_n : y_n) + O(\log n)$. ▲

Лемма 3. *Пусть даны четыре последовательности $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{x}'$, $\boldsymbol{y}'$ такие, что x_n и y_n независимы относительно x'_n и относительно y'_n:*

$$I(x_n : y_n|x'_n) = O(\log n), \quad I(x_n : y_n|y'_n) = O(\log n).$$

Тогда всякая последовательность $\boldsymbol{z}$, простая относительно $\boldsymbol{x}$ и $\boldsymbol{y}$ ($K(z_n|x_n) = O(\log n)$, $K(z_n|y_n) = O(\log n)$), проста также и относительно $\boldsymbol{x}'$ и $\boldsymbol{y}'$ ($K(z_n|x'_n) = O(\log n)$, $K(z_n|y'_n) = O(\log n)$).

Доказательство. Пусть $\boldsymbol{z}$ проста относительно $\boldsymbol{x}$ и $\boldsymbol{y}$. Покажем, что $\boldsymbol{z}$ проста также относительно $\boldsymbol{x}'$ и $\boldsymbol{y}'$. Рассмотрим релятивизованный вариант неравенства (5)

$$K(z_n|x'_n) \leq K(z_n|x_n, x'_n) + K(z_n|y_n, x'_n) + I(x_n : y_n|x'_n) + O(\log n).$$

Ослабляя его, получаем

$$K(z_n|x'_n) \leq K(z_n|x_n) + K(z_n|y_n) + I(x_n : y_n|x'_n) + O(\log n).$$

Поскольку $\boldsymbol{z}$ проста относительно $\boldsymbol{x}$ и $\boldsymbol{y}$ и последовательности $\boldsymbol{x}$ и $\boldsymbol{y}$ независимы относительно $\boldsymbol{x}'$, получаем $K(z_n|x'_n) = O(\log n)$. Аналогично $K(z_n|y'_n) = O(\log n)$. ▲

Следующая лемма позволит сводить задачу о нематериализуемости взаимной информации α-пары к задаче о нематериализуемости взаимной информации некоторой β-пары, причем $\beta > \alpha$.

Лемма 4. *Пусть $\alpha < 1/2$, последовательности $\boldsymbol{x}$ и $\boldsymbol{y}$ являются α-парой. Тогда для любого β, удовлетворяющего неравенству*

$$\alpha < \beta \leq \min\{1 - \sqrt{1-2\alpha}, 1/2\},$$

существует такая β-пара $\boldsymbol{x}'$, $\boldsymbol{y}'$, что любая последовательность $\boldsymbol{z}$, простая относительно $\boldsymbol{x}$ и $\boldsymbol{y}$, является также простой относительно $\boldsymbol{x}'$ и $\boldsymbol{y}'$.

Доказательство. Построим последовательности $\boldsymbol{x}'$, $\boldsymbol{y}'$, которые являются случайной β-парой, и

$$I(x_n : y_n | x'_n) = O(\log n), \quad I(x_n : y_n | y'_n) = O(\log n).$$

Тогда согласно лемме 3 всякая последовательность, простая относительно $\boldsymbol{x}$ и $\boldsymbol{y}$, будет также простой и относительно $\boldsymbol{x}'$ и $\boldsymbol{y}'$. Для построения $\boldsymbol{x}'$ и $\boldsymbol{y}'$ воспользуемся свойствами четверок последовательностей специального вида.

Рассмотрим четверку двоичных случайных величин φ_1, φ_2, φ_3 и φ_4, имеющих следующее совместное распределение P'. Во-первых, потребуем, чтобы совместные распределения пар случайных величин $\langle\varphi_1, \varphi_2\rangle$ и $\langle\varphi_3, \varphi_4\rangle$ были такими, как показано в табл. 2.

Таблица 2

Проекции распределения P'

Распределение φ_1 и φ_2			Распределение φ_3 и φ_4		
φ_1 \ φ_2	0	1	φ_3 \ φ_4	0	1
0	$\frac{1-\alpha}{2}$	$\frac{\alpha}{2}$	0	$\frac{1-\beta}{2}$	$\frac{\beta}{2}$
1	$\frac{\alpha}{2}$	$\frac{1-\alpha}{2}$	1	$\frac{\beta}{2}$	$\frac{1-\beta}{2}$

Во-вторых, условные распределения вероятностей пары $\langle\varphi_1, \varphi_2\rangle$ при известных значениях φ_3 и φ_4 должны быть такими, как в табл. 3.

Таблица 3

Распределения φ_1, φ_2 при фиксированных значениях φ_3, φ_4

$\varphi_3 = 0$, $\varphi_4 = 0$			$\varphi_3 = 0$, $\varphi_4 = 1$		
φ_1 \ φ_2	0	1	φ_1 \ φ_2	0	1
0	$\frac{1-\beta-t}{1-\beta}$	0	0	$\frac{\beta-\alpha}{2\beta}$	$\frac{\alpha}{2\beta}$
1	0	$\frac{t}{1-\beta}$	1	$\frac{\alpha}{2\beta}$	$\frac{\beta-\alpha}{2\beta}$
$\varphi_3 = 1$, $\varphi_4 = 0$			$\varphi_3 = 1$, $\varphi_4 = 1$		
φ_1 \ φ_2	0	1	φ_1 \ φ_2	0	1
0	$\frac{\beta-\alpha}{2\beta}$	$\frac{\alpha}{2\beta}$	0	$\frac{t}{1-\beta}$	0
1	$\frac{\alpha}{2\beta}$	$\frac{\beta-\alpha}{2\beta}$	1	0	$\frac{1-\beta-t}{1-\beta}$

При этом в качестве значения параметра t возьмем $\dfrac{1-\beta-\sqrt{1-2\alpha}}{2}$. Выражение для величины t имеет смысл (подкоренное выражение неотрицательно), так как по условию $\alpha \le 1/2$.

Данное определение корректно, т.е. табл. 2 и 3 действительно задают совместное распределение четверки случайных величин, если все числа, стоящие в таблицах, неотрицательны. Это условие выполнено, если $t \ge 0$ и $1-\beta-t \ge 0$. Легко проверить, что оба неравенства выполнены для выбранного значения параметра. В самом деле, первое неравенство вытекает из ограничения $\beta \le 1-\sqrt{1-2\alpha}$ в условии леммы, а второе неравенство выполнено для любых α и β из интервала $(0, 1/2)$.

Докажем, что при выбранном значении t случайные величины φ_1 и φ_2 независимы относительно φ_3 и φ_4, а значит,

$$\mathcal{I}(\varphi_1 : \varphi_2|\varphi_3) = 0, \quad \mathcal{I}(\varphi_1 : \varphi_2|\varphi_4) = 0.$$

Доказательства независимости φ_1 и φ_2 при условиях $\varphi_3 = 0$, $\varphi_4 = 0$, $\varphi_3 = 1$ и $\varphi_4 = 1$ совершенно аналогичны, и мы рассмотрим только случай $\varphi_3 = 0$.

Итак, докажем, что при указанном выборе параметра t случайные величины φ_1 и φ_2 независимы при условии $\varphi_3 = 0$. Легко видеть, что совместное распределение φ_1 и φ_2 при условии $\varphi_3 = 0$ будет таким, как в табл. 4.

Таблица 4

Распределение φ_1 и φ_2 при условии $\varphi_3 = 0$

φ_1 \ φ_2	0	1
0	$1-\beta-t+\dfrac{\beta-\alpha}{2}$	$\dfrac{\alpha}{2}$
1	$\dfrac{\alpha}{2}$	$t+\dfrac{\beta-\alpha}{2}$

Независимость пары двоичных случайных величин означает, что задающая распределение матрица из четырех чисел имеет ранг один. Остается найти такое значение t, при котором определитель матрицы, заданной табл. 4, равен нулю. Мы получаем квадратное уравнение

$$(1-\beta-t+\frac{\beta-\alpha}{2})(t+\frac{\beta-\alpha}{2}) - \frac{\alpha^2}{4} = 0,$$

одним из корней которого и будет число $\dfrac{1-\beta-\sqrt{1-2\alpha}}{2}$.

Мы доказали, что случайные величины φ_1 и φ_2 независимы относительно φ_3 и φ_4. Следовательно, если последовательности слов $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{x}', \boldsymbol{y}'$ образуют P-случайную четверку, то

$$I(x_n : y_n|x'_n) = O(\log n), \quad I(x_n : y_n|y'_n) = O(\log n).$$

Кроме того, $\boldsymbol{x}$ и $\boldsymbol{y}$ образуют α-пару, а $\boldsymbol{x}'$ и $\boldsymbol{y}'$ – β-пару.

Для доказательства леммы остается заметить, что если дана произвольная α-пара $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$, то по лемме 1 ее можно достроить до P-случайной четверки $\langle \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{x}', \boldsymbol{y}' \rangle$. ▲

Следствие. *Пусть* $\dfrac{3}{8} < \alpha < \dfrac{1}{2}$, *и* $\boldsymbol{x}$, $\boldsymbol{y}$ – *α-пара. Тогда всякая последовательность* $\boldsymbol{z}$, *простая относительно* $\boldsymbol{x}$ *и* $\boldsymbol{y}$, *имеет логарифмическую сложность* $(K(z_n) = O(\log n))$.

Доказательство. Пусть $\boldsymbol{z}$ – последовательность, простая относительно $\boldsymbol{x}$ и $\boldsymbol{y}$. Нетрудно проверить, что $\dfrac{1}{2} < 1-\sqrt{1-2\alpha}$, т.е. $\alpha \in (3/8, 1/2)$ и $\beta = 1/2$ удовлетворяют условиям леммы 4. Следовательно, $\boldsymbol{z}$ также является простой относительно

некоторых x' и y', образующих случайную 1/2-пару. Но взаимная информация x'_n и y'_n не превосходит $O(\log n)$. Из леммы 2 получаем $K(z_n) = O(\log n)$. ▲

Пусть теперь α – произвольное число из интервала $(0,1)$. Для того чтобы доказать, что для всякой α-пары взаимная информация не материализуется, достаточно несколько раз повторить прием из доказательства следствия. Проведем это рассуждение формально.

Утверждение 2. *Пусть $0 < \alpha < 1$, а x и y – α-пара. Тогда для всякой последовательности z, простой относительно x и относительно y,*

$K(z_n) = O(\log n)$.

Доказательство. Итак, пусть последовательность z проста относительно x и относительно y. Сначала отметим, что если x и y являются α-парой, то, заменив все биты слов последовательности y на противоположные, мы получим $(\frac{1}{2} - \alpha)$-пару с такими же свойствами материализуемости взаимной информацией. Поэтому достаточно рассмотреть $\alpha \le 1/2$.

Случай $\alpha = 1/2$ тривиален. Слова случайной 1/2-пары независимы, т.е. $I(x_n : y_n) = O(\log n)$. Согласно лемме 2 сложность слова z_n не больше, чем взаимная информация x_n и y_n. Таким образом, остается рассмотреть $\alpha < 1/2$.

Пусть теперь $0 < \alpha < 1/2$. Выберем параметр α^1:

$$\alpha^1 = \min\{1 - \sqrt{1-2\alpha}, 1/2\}.$$

Согласно лемме 3 существует такая α^1-пара x^1, y^1, что всякая последовательность z, простая относительно x и y, является также простой и относительно x^1, y^1. Если $\alpha^1 = 1/2$, то все доказано. В противном случае снова применим лемму 3, согласно которой существует такая α^2-пара x^2, y^2,

$$\alpha^2 = \min\{1 - \sqrt{1-2\alpha^1}, 1/2\},$$

что всякая z, простая относительно x^1 и y^1, является простой и относительно x^2, y^2. Повторяя применение леммы 3, получим последовательность пар

$$\langle x^1, y^1\rangle, \langle x^2, y^2\rangle, \ldots, \langle x^n, y^n\rangle, \ldots,$$

где для каждого n последовательности x^n и y^n образуют α^n-пару,

$$\alpha^{n+1} = \min\{1 - \sqrt{1-2\alpha^n}, 1/2\}, \quad n = 1, 2, \ldots \tag{6}$$

При этом всякая последовательность z, простая относительно x и y, является также простой относительно каждой из последовательностей x^n и y^n. Остается доказать, что на некотором шаге будет получена 1/2-пара:

$$\exists N \ \alpha^N = 1/2.$$

Предположим противное. Тогда $\{\alpha^n\}$ является бесконечной строго возрастающей последовательностью, все члены которой меньше 1/2. Значит, последовательность имеет некоторый предел α_∞. Подставляя α_∞ в рекуррентное соотношение (6), получаем

$$\alpha_\infty = 1 - \sqrt{1-2\alpha_\infty},$$

откуда $\alpha_\infty = 0$. Но это противоречит возрастанию α^n. ▲

Утверждение 2 позволяет доказывать теорему 2 для любого значения параметра a. Действительно, согласно (4) для любого a из интервала $(0,1)$ существует такое α, что $c(\alpha) = a$; тогда для α-пары x, y

$$K(x_n) = n + O(\log n), \ K(y_n) = n + O(\log n), \ I(x_n : y_n) = an + O(\log n).$$

Но согласно утверждению 2 для всякой α-пары общая информация не материализуется. Для доказательства теоремы 3 остается заметить, что если $\alpha \in (0,1)$ и x такая последовательность, что $|x_n| = n$ и $K(x_n) = n + O(\log n)$, то согласно лемме 1 можно подобрать такую последовательность y, что x и y образуют α-пару.

§ 3. Ортогональные линейные подпространства

В этом параграфе мы рассмотрим вторую конструкцию, позволяющую получить последовательности слов $\boldsymbol{x}$, $\boldsymbol{y}$ с нематериализуемой взаимной информацией. Зафиксируем два параметра – натуральные числа m и k такие, что $2k < m$. Для каждого $n \in \mathbb{N}$ выберем некоторое конечное поле F_n (поле F_n будет содержать $2^{\Theta(n)}$ элементов). Обозначим теперь через V_n линейное m-мерное пространство над F_n. Считаем, что в V_n выбран некоторый базис. Будем говорить, что векторы v и w из V_n *ортогональны*, если в фиксированной системе координат

$$v = (v^1, v^2, \ldots, v^m), \quad w = (w^1, w^2, \ldots, w^m)$$

и

$$v^1 w^1 + v^2 w^2 + \ldots + v^m w^m = 0.$$

Соответственно, будем называть линейные подпространства $A, B \subseteq V_n$ *ортогональными*, если любой вектор из A ортогонален любому вектору из B.

В качестве x_n и y_n будем брать пары ортогональных k-мерных подпространств из V_n. Если P_n – число всех пар ортогональных k-мерных подпространств в V_n, то, очевидно, $K(x_n, y_n) \leq \log P_n + O(\log n)$. Будем интересоваться случайными парами $\langle x_n, y_n \rangle$, т.е. такими парами, что $K(x_n, y_n) = \log P_n + O(\log n)$. Отметим, что в случае $k = 1$, $m = 3$ мы получаем конструкцию из [5].

Параметрами данной конструкции являются числа m и k, а также размеры полей F_n. Далее мы выберем такие значения параметров, что сложности x_n и y_n будут близки к n; величина $I(x_n : y_n)$ будет зависеть от отношения k и m. Наиболее интересен случай, когда k выбирается близким к $m/2$, так как при этом взаимная информация x_n и y_n оказывается близкой к n.

При фиксированных значениях параметров всякое слово x_n сложности n можно рассматривать как код случайного k-мерного подпространства из V_n. Для любого k-мерного подпространства x_n найдется ортогональное ему k-мерное пространство y_n такое, что условная сложность $K(y_n|x_n)$ имеет максимальную возможную величину (точнее, $K(y_n|x_n)$ есть логарифм числа k-мерных подпространств в V_n, ортогональных подпространству x_n). Ясно, что тогда пара $\langle x_n, y_n \rangle$ будет случайной, т.е. будет иметь сложность $\log P_n + O(\log n)$. Чтобы получить доказательство теоремы 3, остается проверить, что для случайных пар ортогональных подпространств выполняется теорема 2 (нельзя материализовать взаимную информацию).

Доказательство основано на следующем свойстве ортогональных подпространств. Рассмотрим граф G_n, вершинами которого являются все k-мерные подпространства из V_n. Ребрами в графе будут соединены ортогональные подпространства. Зафиксируем некоторую вершину графа v_0. Рассмотрим случайное блуждание на графе, начинающееся в v_0. Пусть v_i – вершина графа, полученная после i шагов случайного блуждания. Для каждого i случайная величина v_i распределена на множестве вершин G_n (т.е. на множестве k-мерных подпространств V_n). Мы покажем, что для некоторого s распределение v_s близко к равномерному. При этом s зависит от m и k, но не зависит от n.

Итак, пусть для каждого n слова x_n и y_n кодируют случайную пару ортогональных k-мерных подпространств из V_n. Перейдем к формальному доказательству.

Утверждение 3. *Колмогоровские сложности и взаимная информация $\boldsymbol{x}$, $\boldsymbol{y}$ растут линейно по n:*

$$K(x_n) = (mk - k^2)|F_n| + O(\log n), \tag{7}$$

$$K(y_n) = (mk - k^2)|F_n| + O(\log n), \tag{8}$$

$$I(x_n : y_n) = k^2|F_n| + O(\log n). \tag{9}$$

Доказательство. Пусть W – линейное пространство над F_n. Найдем число последовательностей $e_1, e_2, \ldots, e_k$, состоящих из k линейно независимых векторов пространства W.

Обозначим $s = \dim(W)$ и $N = |F_n|$. В качестве e_1 можно взять любой ненулевой вектор W. Таким образом, для выбора первого вектора в последовательности имеется $N^s - 1$ вариант. Пусть вектор e_1 уже выбран. Тогда для выбора второго вектора последовательности имеется $N^s - N$ вариантов, поскольку e_2 не должен линейно зависеть от e_1. Далее, если выбраны первые i векторов последовательности, то в качестве вектора e_{i+1} может быть взят любой вектор, не лежащий в линейной оболочке $e_1, \ldots, e_i$, т.е. для выбора e_{i+1} имеется $N^s - N^i$ вариантов. Следовательно, в пространстве W имеется

$$(N^s - 1)(N^s - N)\ldots(N^s - N^{k-1}) = N^{ks}(1 + O(1/N))$$

последовательностей из k линейно независимых векторов. Подставляя вместо s число m, находим количество последовательностей из k линейно независимых векторов во всем пространстве V_n. Далее, подставляя вместо s число k, находим количество последовательностей из k линейно независимых векторов в каждом k-мерном подпространстве V_n. Отношение этих величин

$$Q_n = N^{mk-k^2}(1 + O(1/N))$$

дает количество k-мерных подпространств в V_n. Поскольку x_n и y_n выбираются случайными,

$$K(x_n) = \log Q_n + O(\log n), \quad K(y_n) = \log Q_n + O(\log n).$$

Остается найти взаимную информацию x_n и y_n.

Вычислим величину условной сложности $K(y_n|x_n)$. Если подпространство x_n уже задано, то y_n лежит в подпространстве векторов V_n, ортогональных x_n. Размерность этого подпространства равна $m - k$. Но во всяком $(m-k)$-мерном пространстве имеется

$$T_n = N^{(m-k)k-k^2}(1 + O(1/N))$$

k-мерных подпространств. Сложность $K(y_n|x_n)$ с точностью до $O(\log n)$ равна логарифму величины T_n. Следовательно,

$$I(x_n : y_n) = K(y_n) - K(y_n|x_n) = \log(\frac{Q_n}{T_n}) + O(\log n).$$

Непосредственные вычисления логарифмов Q_n и T_n доказывают требуемое утверждение. ▲

Будем считать m, k и F_n ($|F_n| = 2^{\Theta(n)}$) выбранными такими, что

$$K(x_n) = n + O(\log n), \tag{10}$$

$$K(y_n) = n + O(\log n), \tag{11}$$

$$I(x_n : y_n) = an + O(\log n), \tag{12}$$

где a – некоторая положительная константа. Таким образом, взаимная информация x_n и y_n растет линейно по n. Нетрудно заметить, что когда отношение k/m стремится к $1/2$, соответствующее значение a стремится к единице. Следовательно, выбирая значения параметров m и k, можно сделать величину a сколь угодно близкой к единице.

Покажем, что у построенных последовательностей x_n, y_n нельзя материализовать взаимную информацию, т.е. для них выполняется утверждение теоремы 2.

Утверждение 4. *Для построенных последовательностей* $\boldsymbol{x}$, $\boldsymbol{y}$ *и для всякой последовательности* $\boldsymbol{z}$, *простой относительно* $\boldsymbol{x}$ *и относительно* $\boldsymbol{y}$, *выполнено равенство* $K(z_n) = O(\log n)$.

Доказательство. Пусть z_n – последовательность слов, простых относительно x_n и y_n (т.е. $K(z_n|x_n) = O(\log n)$ и $K(z_n|y_n) = O(\log n)$). Докажем, что $K(z_n) =$

$= O(\log n)$. Зафиксируем натуральное n. Далее для простоты обозначений будем опускать нижний индекс n всюду, где это не приведет к путанице.

Поскольку $\boldsymbol{z}$ просто относительно $\boldsymbol{x}$ и $\boldsymbol{y}$,

$$K(x|z) = K(x,z) - K(z) + O(\log n) =$$
$$= K(x) + K(z|x) - K(z) + O(\log n) = K(x) - K(z) + O(\log n). \qquad (13)$$

Аналогичные вычисления можно провести для y. Положим

$$D = \max\{K(x|z), K(y|z)\}.$$

(Отметим, что $|K(x|z) - K(y|z)| = O(\log n)$.) В новых обозначениях

$$K(x|z) \le D,\ K(y|z) \le D.$$

Далее мы докажем, что существует достаточно много слов, сложность которых относительно z не превосходит D (а значит, число D достаточно велико). Более точно, покажем, что $D = K(x) - O(\log n)$. Это будет значить, что условная сложность $K(x|z)$ очень мало отличается от безусловной сложности $K(x)$. Далее, используя (13), мы получим логарифмическую оценку на сложность z.

В доказательстве будут рассматриваться цепочки подпространств – конечные последовательности

$$x^0 - y^1 - x^1 - y^2 - \ldots - y^r - x^r, \qquad (14)$$

где x^i, y^i – k-мерные подпространства V_n, причем любые два соседних в цепочке подпространства ортогональны. Подпространство x^0 будем называть левым концом, а подпространство x^r – правым концом цепочки. Число r назовем длиной цепочки (r не зависит от n). Будем интересоваться только такими цепочками, в которых $x^0 = x$. Такая цепочка является траекторией случайного блуждания на графе G_n. Отметим, что число шагов блуждания четно, нечетные шаги обозначены x^i, $i = 0, 1, \ldots, r$, а четные шаги, соответственно, y^i, $i = 1, \ldots, r$.

Случайное блуждание на графе соответствует равномерному распределению на множестве цепочек с фиксированным левым концом. Равномерное распределение на цепочках индуцирует некоторое распределение на множестве их правых концов. Мы подберем такое значение параметра r, что получаемое распределение на множестве правых концов цепочек окажется близким к равномерному. В то же время мы покажем, что с достаточно большой вероятностью правый конец случайно выбранной цепочки имеет сложность не более D относительно z. Используя это, мы докажем, что число правых концов цепочек, имеющих сложность не более D относительно z, велико.

Прежде всего покажем, что для полиномиальной доли всех цепочек вида (14) сложность правого конца x^r мала относительно z.

Лемма 5. *Пусть X^r – множество всех цепочек вида* (14). *Тогда число таких цепочек, для правых концов которых выполнено неравенство*

$$K(x^r|z) \le D,$$

не меньше $\dfrac{|X^r|}{\mathrm{poly}(n)}$ (*где* $\mathrm{poly}(n)$ – *некоторый многочлен*).

Доказательство. Докажем более сильное утверждение. А именно, покажем, что не менее чем полиномиальную долю составляют такие цепочки подпространств, для которых выполняются следующие два условия:

а) *все* элементы цепочки x^i, y^i имеют сложность не более D относительно z;

б) каждая пара соседних (в цепочке) подпространств $\langle y^j, x^j\rangle$ или $\langle x^j, y^{j+1}\rangle$ случайна, т.е. имеет сложность не меньше $\log P_n - O(\log n)$ (здесь и далее мы пользуемся обозначениями из доказательства утверждения 3).

Отметим, что пара $\langle y^j, x^j\rangle$ имеет сложность, близкую к P_n, тогда и только тогда, когда сложность $K(x^j)$ близка к $\log Q_n$, а условная сложность $K(y^j|x^j)$ близка

к $\log T_n$. Точнее, случайность пары $\langle y^j, x^j\rangle$ эквивалентна тому, что для некоторой константы C выполнены неравенства

$$K(x^j) \geq \log Q_n - C\log n,$$
$$K(y^j|x^j) \geq \log T_n - C\log n.$$

Доказательство леммы проведем по индукции по длине цепочки. Пусть имеется не менее $\dfrac{|X^i|}{n^c}$ цепочек длины i, удовлетворяющих условию леммы. Выберем любую из них и рассмотрим все возможные ее продолжения $\ldots - y^{i+1} - x^{i+1}$. При этом подпространство y^{i+1} должно быть ортогонально x^i, а x^{i+1} – ортогонально y^{i+1}. Всего имеется T_n^2 таких продолжений. Достаточно доказать, что среди этих продолжений по крайней мере полиномиальная доля удовлетворяет условиям а) и б).

По предположению $K(x^i|z)$ и $K(y^i|z)$ не превосходят D, и пара $\langle x^i, y^i\rangle$ случайна. (При $i = 0$ будем считать $y^0 = y$.) Учитывая определение D и соотношение (13), получаем

$$K(z|x^i) = K(x^i, z) - K(x^i) + O(\log n) =$$
$$= K(x^i|z) + K(z) - K(x^i) + O(\log n) \leq$$
$$\leq D + K(z) - K(x) + O(\log n) = O(\log n).$$

Рассмотрим множество L всех k-мерных подпространств $\hat{y}$, которые ортогональны x^i, и

$$K(\hat{y}|z) \leq D.$$

Заметим, что, зная слово x^i, можно с логарифмической сложностью получить z, а затем запустить процесс перечисления множества L. Но подпространство y^i лежит в L. Поэтому для нахождения y^i достаточно иметь программу, перечисляющую L, и знать номер y^i в этом перечислении. Таким образом,

$$K(y^i|x^i) \leq \log|L| + O(\log n).$$

По предположению индукции пара $\langle x^i, y^i\rangle$ случайна, и сложность y^i относительно x^i не меньше $\log T_n - \mathrm{C}\log n$. Следовательно, $|L| \geq T_n/\mathrm{poly}(n)$.

Теперь выберем некоторую константу $C' > C$ и отбросим те подпространства из L, сложность которых относительно x^i меньше $T_n - C'\log n$. Более точно, пусть $L' \subset L$ состоит из всех таких подпространств $\hat{y}$, что

$$K(\hat{y}|x^i) \geq T_n - C'\log n.$$

Если константа C' достаточно велика, то $|L'| \geq T_n/\mathrm{poly}(n)$. Любое подпространство из L' можно взять в качестве y^{i+1}. Действительно, для любого $\hat{y} \in L'$ пара $\langle x^i, \hat{y}\rangle$ случайна, и $K(\hat{y}|z) \leq D$.

Аналогично можно доказать, что если $y^{i+1} \in L'$ выбрано, то имеется не менее $T_n/\mathrm{poly}(n)$ подпространств $\hat{x}$, каждое из которых можно взять в качестве x^{i+1}.

Таким образом, среди T_n^2 продолжений выбранной цепочки длины i имеется не менее $\dfrac{T_n^2}{\mathrm{poly}(n)}$, удовлетворяющих условиям а) и б). (Отметим, что степень получаемого многочлена $\mathrm{poly}(n)$ зависит от r.) ▲

Определим последовательность чисел l_i следующим рекуррентным соотношением:

$$l_0 = k, \tag{15}$$
$$l_{i+1} = \max\{l_i + 2k - m; 0\}. \tag{16}$$

Заметим, что, начиная с некоторого номера, все числа l_i равны нулю.

Назовем цепочку подпространств *правильной*, если для $i = 1, 2, \ldots, r$

$$\dim(x^0 \cap x^i) = l_i. \tag{17}$$

Далее покажем, что случайно выбранная цепочка с экспоненциально близкой к единице вероятностью является правильной. Более точно, имеет место

Лемма 6. *Среди цепочек подпространств вида* (14) *доля правильных не меньше* $1 - 2^{-cn}$ *для некоторого* $c > 0$.

Доказательство. Прежде всего докажем две простые комбинаторные сублеммы.

Сублемма 1. *Выберем случайно (относительно равномерного распределения) систему из* q *уравнений с* s *переменными над полем* F_n. *Тогда с вероятностью не меньше* $1 - 2^{-cn}$ *(для некоторой константы* $c > 0$*) ранг данной системы будет равен* $\min\{s; q\}$.

Доказательство. Пусть $q \leq s$. Покажем, что с экспоненциально близкой к единице вероятностью все уравнения системы линейно независимы. Действительно, если уравнения линейно зависимы, то одно из них является линейной комбинацией остальных. Всего имеется $|F_n|^{q-1}$ линейных комбинаций $(q-1)$ уравнений. А одно уравнение можно выбрать $|F_n|^s$ способами. Таким образом, для каждого i вероятность того, что i-е уравнение системы есть линейная комбинация остальных, не превосходит $|F_n|^{(q-1)-s}$. Остается просуммировать данные вероятности по всем i от 1 до q. Так как $q \leq s$, доля линейно зависимых систем уравнений не превосходит $q|F_n|^{-1}$. Но $|F_n| = 2^{\Theta(n)}$, и утверждение доказано.

Если $q > s$, то с вероятностью, экспоненциально близкой к единице, первые s уравнений системы линейно независимы, и ранг системы равен s. ▲

Замечание 4. Отметим очевидное следствие доказанной сублеммы. Пусть случайно выбрана система из q линейных уравнений с s переменными над конечным полем F. Предположим, о некоторых совокупностях уравнений полученной системы известно, что они оказались линейно независимыми. Ясно, что при данном условии вероятность события "ранг всей выбранной системы равен $\min\{s; q\}$" тем более экспоненциально близка к единице.

Сублемма 2. 1) *Пусть* W *– линейное пространство над конечным полем* F, *и пусть* a *–* s_1*-мерное подпространство в* W. *Выберем случайно (относительно равномерного распределения)* s_2*-мерное подпространство* b *в* W. *Тогда вероятность события* $\dim(a \cap b) = r$ *зависит только от величин* $r, s_1, s_2, \dim(W), |F|$ *(но не зависит от выбора* s_1*-мерного пространства* a*).*

2) *Пусть* W *– линейное пространство над конечным полем* F, a *и* b *–* s*-мерные линейные подпространства* W. *Выберем случайно* s*-мерное подпространство* a_1 *из ортогонального дополнения* a *и* s*-мерное подпространство* b_1 *из ортогонального дополнения* b. *Тогда для любого* l *вероятности событий* $\dim(a_1 \cap b) = l$ *и* $\dim(a \cap b_1) = l$ *равны.*

Доказательство. 1) Пусть a' – произвольное линейное подпространство W размерности s_1. Очевидно, существует автоморфизм φ пространства W такой, что $\varphi a = a'$. Тогда для любого линейного подпространства b из пространства W имеем $\dim(a \cap b) = \dim(a' \cap \varphi b)$. Следовательно,

$$\mathrm{Prob}_b[\dim(a \cap b) = l] = \mathrm{Prob}_b[\dim(a' \cap \varphi b) = l] = \mathrm{Prob}_b[\dim(a' \cap b) = l].$$

2) Заметим, что $\dim(a^\perp \cap b) = \dim(a \cap b^\perp)$. В самом деле,

$$\dim(W) - \dim(a \cap b^\perp) = \dim((a \cap b^\perp)^\perp) = \dim(a^\perp \oplus b) =$$

$$= \dim(a^\perp) + \dim(b) - \dim(a^\perp \cap b) = \dim(W) - \dim(a^\perp \cap b)$$

(здесь $(a^\perp \oplus b)$ обозначает сумму подпространств $a^\perp$ и b). Остается применить утверждение 1) сублеммы 2 к пространству $a^\perp$, в котором лежат $b \cap a^\perp$ и случайно выбранное a_1, и к пространству $b^\perp$, в котором лежат $a \cap b^\perp$ и случайно выбранное b_1. ▲

Будем доказывать лемму индукцией по длине цепочки. База индукции очевидна. Для выполнения шага индукции достаточно показать, что если цепочка

$$x^0 - y^1 - x^1 - y^2 - \ldots - y^i - x^i$$

правильна, то все ее продолжения

$$x^0 - y^1 - x^1 - y^2 - \ldots - y^i - x^i - y^{i+1} - x^{i+1},$$

кроме экспоненциально малой доли, также правильны. Рассмотрим заключительный фрагмент данной цепочки

$$x^i - y^{i+1} - x^{i+1}.$$

Для выбора y^{i+1} и x^{i+1} имеется T_n^2 различных возможностей (в обозначениях из доказательства утверждения 3). Подсчитаем, с какой вероятностью $\dim(x^0 \cap x^{i+1}) = l_{i+1}$.

Предположим, что y^{i+1} уже выбрано. Далее можно рассматривать только пару подпространств x^0 и y^{i+1}: требуется узнать, какова вероятность того, что случайно выбранное третье подпространство (мы назовем его x^{i+1}) имеет l_{i+1}-мерное пересечение с x^0 при условии, что y^{i+1} и x^{i+1} оказались перпендикулярными. Согласно утверждению 2) сублеммы 2 последняя задача эквивалентна другой: какова вероятность того, что случайно выбранное подпространство y^0 будет иметь l_{i+1}-мерное пересечение с y^{i+1} при условии, что x^0 и y^0 перпендикулярны.

Таким образом, для того чтобы вычислить вероятность, с которой выполняется равенство $\dim(x^0 \cap x^{i+1}) = l_{i+1}$, можно решить другую задачу: пусть y^0 – случайно выбранное k-мерное подпространство в V_n, перпендикулярное x^0, а y^{i+1} – случайно выбранное k-мерное подпространство в V_n, перпендикулярное x^i; какова вероятность того, что $\dim(y^0 \cap y^{i+1}) = l_{i+1}$?

Подпространства y^0 и y^{i+1} можно задать системами из $m-k$ линейных уравнений. Без ограничения общности можно считать, что первые k уравнений в первой системе соответствуют ортогональности y^0 пространству x^0, а первые k уравнений во второй системе – ортогональности y^{i+1} пространству x^i. Можно также считать, что первые l_i уравнений в обеих системах одинаковы и соответствуют требованию ортогональности обоих подпространств y^0 и y^{i+1} подпространству $x^0 \cap x^i$.

Объединяя системы уравнений, задающие y^0 и y^{i+1}, мы получаем систему из $2(m-k) - l_i$ уравнений. Согласно сублемме 1 с вероятностью, экспоненциально близкой к единице, данная система уравнений имеет ранг $\min\{m, 2(m-k) - l_i\}$, а значит, выполнено равенство

$$\dim(y^0 \cap y^{i+1}) = \max\{0; 2k - (m - l_i)\} = l_{i+1}.$$

Следовательно, с такой же вероятностью и $\dim(x^0 \cap x^{i+1}) = l_{i+1}$. Таким образом, с экспоненциально близкой к единице вероятностью для цепочки выполняется условие (17). ▲

Пусть A^r – множество всех k-мерных подпространств в V_n, пересечение которых с x имеет размерность l_r. Согласно лемме 6 для большинства цепочек вида (14) x^r принадлежит A^r. Теперь докажем, что имеет место однородность: все подпространства из A^r являются правыми концами одинакового числа правильных цепочек длины r (по-прежнему рассматриваем только такие цепочки, левый конец которых совпадает с x).

Лемма 7. *Если $v, w \in A^r$, то число правильных цепочек, правый конец которых совпадает с v, равно числу правильных цепочек, правый конец которых совпадает с w.*

Доказательство. Доказательство проведем индукцией по r. Нужно вычислить для каждого подпространства $x^r \in A^r$ количество цепочек

$$x^{r-1} - y^r - x^r,$$

где $x^{r-1} \in A^{r-1}$. Точнее, достаточно показать, что количество таких цепочек не зависит от выбора $x^r \in A^r$, а зависит только от номера r. Зафиксируем подпространство $x^r \in A^r$ и будем выбирать случайное перпендикулярное ему y^r. Далее выберем случайное x^{r-1}, перпендикулярное y^r. Нас интересует, какова вероятность события $x^{r-1} \in A^{r-1}$ (т.е. $\dim(x^0 \cap x^{r-1}) = l_{r-1}$).

Пусть подпространство y^r уже выбрано. Требуется определить, с какой вероятностью случайно выбранное подпространство x^{r-1} имеет l_{r-1}-мерное пересечение с x^0 при условии, что x^{r-1} и y^r перпендикулярны. Согласно утверждению 2) сублеммы 2 данная вероятность равна вероятности того, что случайно выбранное подпространство y^0 имеет l_{r-1}-мерное пересечение с y^r при условии, что y^0 и x^0 перпендикулярны.

Таким образом, мы можем перейти к решению новой задачи: найти вероятность того, что пара случайно выбранных k-мерных подпространств y^0, y^r, первое из которых перпендикулярно x^0, а второе перпендикулярно x^r, имеют l_{r-1}-мерное пересечение.

Подпространства y^0 и y^r задаются системами $n-k$ линейно независимых уравнений. Будем считать, что первые k уравнений первой системы соответствуют ортогональности y^0 подпространству x^0, а первые k уравнений второй системы – ортогональности y^r подпространству x^r. Кроме того, можно считать, что первые l_r уравнений обеих систем совпадают (и соответствуют ортогональности обоих подпространств y^0, y^r подпространству $x^0 \cap x^r$). Объединим две данные системы уравнений. Пространство решений новой системы (из $2(n-k)-l_r$ уравнений) есть пересечение y^0 и y^r. Нас интересует вероятность того, что его размерность равна l_{r-1}. Очевидно, эта вероятность определяется значениями n, k и r, но не зависит от выбора конкретного x^r. Нам не требуется значение данной вероятности. Важно лишь, что оно однозначно определяется номером r (при данном n). ▲

Закончим доказательство утверждения 4. Рассмотрим множество X^r цепочек подпространств (14) длины r. Согласно лемме 6 все такие цепочки за исключением экспоненциально малой доли являются правильными, а значит, их правые концы лежат в A^r. Далее, по лемме 5 по крайней мере $\dfrac{|X^r|}{\mathrm{poly}(n)}$ цепочек имеют правые концы, простые относительно z (т.е. их сложность относительно z не превосходит D). Значит, не менее $\dfrac{|X^r|}{\mathrm{poly}(n)}$ цепочек одновременно являются правильными и имеют правый конец сложности не более D относительно z. Но в силу однородности (лемма 7) все подпространства из A^r являются правыми концами одинакового числа правильных цепочек. Следовательно, не менее $\dfrac{|A^r|}{\mathrm{poly}(n)}$ элементов множества A^r имеют сложность не больше D относительно z.

Выберем минимальное r, для которого $l_r = 0$. Тогда A^r состоит из всех k-мерных подпространств, имеющих нулевое пересечение с x. В этом случае все k-мерные подпространства из V_n, кроме экспоненциально малой доли, лежат в A^r. Действительно, подпространство x в V_n задается системой $(m-k)$ линейных уравнений. Выберем случайно еще $(m-k)$ линейно независимых уравнений (задающих некоторое k-мерное подпространство в V_n). Объединим две данные системы уравнений. Новая система содержит $2(m-k)$ уравнений. Поскольку $2k < m$, с экспоненциально близкой к единице вероятностью ранг данной системы равен m (сублемма 1), и

она не имеет нетривиальных решений. Это значит, что случайно выбранное подпространство в V_n с экспоненциально близкой к единице вероятностью имеет нулевое пересечение с x.

Итак, A^r содержит все k-мерные подпространства из V_n, кроме экспоненциально малой доли. Мы знаем, что не менее чем полиномиальная доля всех подпространств из A^r имеет сложность не более D_n относительно z. Следовательно, и среди всех k-мерных подпространств V_n по крайней мере полиномиальная доля имеет сложность не более D относительно z. Таким образом, если Q_n – число всех k-мерных подпространств в V_n, то

$$2^D \geq \frac{Q_n}{\mathrm{poly}(n)}.$$

Напомним, что $K(x) = \log Q_n + O(\log n)$. Далее, $D = K(x|z) + O(\log n) = K(x) - -K(z) + O(\log n)$. Следовательно,

$$K(x) - K(z) \geq K(x) - O(\log n).$$

Таким образом, $K(z) = O(\log n)$. ▲

Замечание 5. Рассмотренное построение можно обобщить. Пусть k_1, k_2, m – такие натуральные числа, что $k_1 + k_2 < m$. Рассмотрим последовательности слов $\{x_n\}$ и $\{y_n\}$, где x_n и y_n – случайные ортогональные линейные подпространства в F_n^m, размерности которых равны k_1 и k_2 соответственно.

Тогда для некоторых a, b, c (которые определяются параметрами k_1, k_2 и m)

$$K(x_n) = bn + O(\log n), \quad K(y_n) = cn + O(\log n), \quad I(x_n : y_n) = an + O(\log n).$$

Рассуждение, аналогичное доказательству утверждения 4, показывает, что взаимная информация $\{x_n\}$ и $\{y_n\}$ не может быть материализована, т.е. для любой последовательности $\{z_n\}$, простой относительно $\{x_n\}$ и $\{y_n\}$, выполнено $K(z_n) = O(\log n)$.

Автор благодарен Н.К. Верещагину за научное руководство и помощь в работе над статьей.

СПИСОК ЛИТЕРАТУРЫ

1. *Gács P., Körner J.* Common Information is Far Less Then Mutual Information // Probl. Control and Inform. Theory. 1973. V. 2. № 2. P. 149–162.
2. *Колмогоров А.Н.* Комбинаторные основания теории информации и исчисления вероятностей // УМН. 1983. Т. 38. № 4. С. 27–36.
3. *Звонкин А.К., Левин Л.А.* Сложность конечных объектов и обоснование понятий информации и случайности с помощью теории алгоритмов // УМН. 1970. Т. 25. № 6. С. 85–127.
4. *Мучник Ан.А.* О выделении общей информации двух слов // Первый Всемирный конгресс общества математической статистики и теории вероятностей им. Бернулли. Тезисы докл. М.: Наука, 1986. Т. 1. С. 453.
5. *Muchnik An.A.* On Common Information // Theoretical Computer Science. 1998. V. 207. P. 319-328.
6. *Muchnik An.A., Shen A., Romashchenko A., Vereshchagin N.K.* Upper semi-lattice of binary strings with the relation x is simple conditional to y: Preprint DIMACS TR 97-74. Rutgers University, 1997.
7. *Hammer D., Romashchenko A., Shen A., Vereshchagin N.* Inequalities for Shannon entropies and Kolmogorov complexities // Proc. Twelfth Annual IEEE Conference on Computational Complexity. Ulm, Germany. June, 1997. P. 13–23.

Поступила в редакцию
29.03.99
После переработки
30.11.99