Устойчивость колмогоровских свойств при релятивизации

Ан. А. Мучник*

18 августа 2007 г.

Аннотация

Предположим, что у набора слов $x_1, \ldots, x_n$ пренебрежимо малая взаимная информация с некоторым z. Естественно предположить, что те свойства набора слов x_i, которые можно выразить на языке колмогоровской сложности, не должны значительно меняться при релятивизации относительно z. В данной статье мы попытаемся формализовать это предположение и докажем его для некотороых важных частных случаев.

1 Введение

В данной работе мы интересуемся устойчивостью *колмогоровских свойств* кортежей слов при релятивизации. Прежде всего следует пояснить, какие свойства мы называем колмогоровскими. Говоря кратко и нефомально, мы интересуемся свойствами, выразимыми в терминах колмогоровских сложностей заданных слов. При этом мы как правило будем интересоваться свойствами, выполняющимися с 'логарифмической точностью'. Прежде чем давать точное определение, приведём несколько примеров.

*Работа выполнена при частичной поддержке Российского Фонда Фундаментальных Исследований, проекты №04-01-00427, №02-01-10904, Совета поддержки научных школ при президенте РФ, и гранта 047.017.014 Голландско-Российской программы сотрудничества NWO/РФФИ

Пример 1. *Для любых слов x_1, x_2 выполнены неравенства*

$$K(x_1, x_2) \leq K(x_1) + K(x_2) + \mathcal{O}(\log K(x_1, x_2)),$$

и

$$K(x_1, x_2) \geq K(x_1) - \mathcal{O}(1).$$

Кроме того, выполнено равенство

$$K(x_1, x_2) = K(x_1) + K(x_2|x_1) + \mathcal{O}(\log K(x_1, x_2)).$$

Это наиболее простые свойства колмогоровских сложностей пары слов, выразимые с помощью *линейных равенств и неравенств*. В теории колмогоровской сложности часто оказывается, что интересные свойства выполнены с логарифмической точностью до аддитивного логарфмического слагаемого, как в рассмотренном примере. Кроме того, различные виды колмогоровской сложности (префиксная, монотонная сложности, сложность разрешения, априорная сложность) отличаются от простой сложности на величины не более чем логарифмического порядка. Поэтому свойства, выполненные с 'логарифмической точностью', одинаковы для всех видов колмогоровской сложности. Всё это даёт основания интересоваться свойствами сложности с логарифмической погрешностью.

Как описать такого рода свойств в наиболее общем виде? Для набора слов $x_1, \ldots, x_n$ мы рассмотрим колмогоровские сложности всех кортежей $x_{i_1}, \ldots, x_{i_k}$, где $1 \leq i_1 < \ldots < i_k \leq n$. Таким образом, набору из n слов соответствует $(2^n - 1)$ сложностей. Упорядочив все поднаборы из данного набора слов некоторым каноническим образом (скажем, лексикографически), мы получаем *сложностной профиль* – вектор в $\mathbb{Z}_+^{2^n-1}$

$$\vec{K}(x_1, \ldots, x_n) = (K(x_1), K(x_1, x_2), \ldots, K(x_2), K(x_2, x_3), \ldots).$$

Замечание. Не имеет смысла рассматривать кортежи, отличающиеся только порядком слов, я также кортежи, в которых некоторые слова встречаются несколько раз. При перестановке или дублировании членов кортежа, его колмогоровская сложность меняется на величину, зависящую только от числа слов в кортеже, но не от их сложностей. Поскольку мы собираемся пренебрегать логарифмической погрешностью, нет нужды различать между собой величины, отличающиеся друг от друга на $\mathcal{O}(1)$.

По тем же причинам нет необходимости принимать во внимание относительные колмогоровские сложности, поскольку с помощью теоремы Колмогорова – Левина их можно выразить как комбинацию безусловных сложностей:

$$K(x_1,\dots,x_n|y_1,\dots,y_m) = K(x_1,\dots,x_n,y_1,\dots,y_m) - K(y_1,\dots,y_m) + \mathcal{O}(\log N),$$

где $N = K(x_1,\dots,x_n,y_1,\dots,y_m)$.

Таким образом, мы можем говорить о свойствах сложностного профиля $\vec{K}(x_1,\dots,x_n)$ как о простейших колмогоровских свойствах набора слов $x_1,\dots,x_n$. Например, неравенство $K(x_1,x_2) \le K(x_1) + K(x_2) + \mathcal{O}(\log K(x_1,x_2))$ соответствует утверждению о профиле пары x_1, x_2: существует такая константа C, что для любых слов x_1, x_2 сложностной профиль $\vec{K}(x_1,x_2) = (K(x_1), K(x_2), K(x_1,x_2))$ принадлежит множеству

$$A = \{(u,v,w) \ : \ w \le u + v + C\log w\}.$$

В общем случае мы можем говорить, что простые утверждения о колмогоровских свойствах формулируются в виде

$$\vec{K}(x_1,\dots,x_n) \in A$$

для всевозможных множеств A. В частности, в таком виде можно выразить утверждения об истинности линейных информационных неравенств [11, 10, 5].

Описанные простые утверждения, выражающие колмогоровские свойства, можно рассмотреть как бесквантворные формулы. Разумеется, к такого вида формуле можно приписать кванторы всеобщности по всем переменным, получив (истинное или ложное) универсальное утверждение о колмогоровской сложности.

Однако предсталяют интерес и более сложные свойства колмогоровской сложности. Чтобы выразить их могут потребоваться формулы с переменами кванторов. Простейший пример – свойство выделяемости взаимной информации пары слов:

Пример 2. *Для заданных* x_1, x_2 *мы можем интересоваться, верно ли*

$$\exists y \ : \ K(y|x_1) = \mathcal{O}(\log N) \wedge K(y|x_2) = \mathcal{O}(\log N) \wedge K(y) \ge I(x_1 : x_2) - \mathcal{O}(\log N),$$

где $N = K(x_1, x_2)$. Конкретизируя константу в члене $\mathcal{O}(\log N)$, мы получаем вопрос следующего вида: верно ли, что

$$\exists y \vec{K}(x_1, x_2, y) \in A,$$

где A – соответствующее множество в $\mathbb{Z}^7$. Для сколь угодно большой константы в аддитивном члене $\mathcal{O}(\log N)$ ответ на данный вопрос можем быть положительным или отрицательным в зависимости от выбора пары $\langle x_1, x_2 \rangle$. Это следует из теоремы Гача и Кёрнера [2], см. также более сильное утверждение в [7].

Другие, более сложные примеры колмогоровских свойств, требующих перемен кванторов в формулировке, можно найти в [8].

Наиболее общий вид колмогоровского свойства для слов $x_1, \ldots, x_n$ может быть выражен формулой вида

$$\forall y_1 \exists y_2 \forall y_3 \ldots \vec{K}(x_1, \ldots, x_n, y_1, \ldots, y_m) \in A, \tag{1}$$

где $A \subset \mathbb{Z}^{2^{n+m}-1}$.

Пусть для $\bar{x} = (x_1, \ldots, x_n)$ выполнено свойство (1), а для кортежа $\bar{x}' = (x'_1, \ldots, x'_n)$ выполнено аналогичное свойство

$$\forall y_1 \exists y_2 \forall y_3 \ldots \vec{K}(x'_1, \ldots, x'_n, y_1, \ldots, y_m) \in A'$$

для некоторого A'. Можно говорить, что свойство $\bar{x}$ и $\bar{x}'$ близки, если близки множества множеств A и A'. В частности, если A лежит в $\mathcal{O}(logK(\bar{x}))$-окрестности A' и наоборот, мы говорим, что свойства кортежей $\bar{x}$ и $\bar{x}'$ близки с логарифмической точностью.

Итак, мы уточнили, какие свойства мы называем колмогоровскими. Теперь перейдём к вопросу об устойчивости этих свойств при релятивизации. Пусть задан некоторый оракул O (конечная или бесконечная двоичная последовательность). Можно рассмотреть колмогоровскую сложность, релятивизованную относительно данного оракула $K(x_1, \ldots, x_n|O)$. В случае конечного O релятивизованная сложность совпадает с обычной относительной сложностью.

Для произвольной n-ки слов мы можем рассмотреть релятивизованный сложностной профиль $\vec{K}(x_1, \ldots, x_n|O)$, содержащий все релятивизованные сложности $K(x_{i_1}, \ldots, x_{i_k}|O)$. При каких условиях релятивизация относительно O не изменяет колмогоровских свойств кортежа $\bar{x} =$

$(x_1, \ldots, x_n)$, или, во всяком случае, меняет их незначительно? Необходимое условие очевидно: оракул O не должен содержать информации о $\bar{x}$, то есть разность

$$K(\bar{x}) - K(\bar{x}|O)$$

должна быть пренебрежимо малой. Мы предполагаем, что данное условие является (при соответствующем уточнении 'пренебрежимой малости') также и достаточным.

Основная гипотеза: *Колмогоровские свойства кортежа* $\bar{x} = (x_1, \ldots, x_n)$ *мало меняются при релятивизации относительно оракула* O*, если и только если взаимная информация* $I(O : \bar{x}) = K(\bar{x}) - K(\bar{x}|O)$ *мала.*

Однако доказать это в сколько-нибудь общей ситуации оказывается трудно. В данной работе мы изучим некоторые частные случаи общей гипотезы. Далее мы сформулируем основные результаты статьи. Для простоты формулировок мы ограничимся формулировками для конечного оракула O (слово, задающее оракул, мы будем обозначать z).

1.1 Формулировки основных результатов

В дальнейшем мы будем часто использовать асимптотическое обозначение $\mathcal{O}(f(x_1, \ldots, x_n))$ для различных выражений $f()$, в которые входит колмогоровская сложность. Мы имеем в виду, что мультипликативная константа в 'О-большом' зависит только от конретизации функции колмогоровской сложности и от длины рассматриваемых кортежей. Более того, зависимость константы от длины кортежей эффективна.

1. Эквивалентность для бесквантворных формул.

Нетрудно проверить, что для свойств, выражающихся бесквантворными формулами, малость взаимной информации между $\bar{x}$ и оракулом O является необходимым и достаточным условием для устойчивости колмогоровского свойства при релятивизации. Для полноты изложения мы докажем этот факт:

Теорема 1 *Предположим, что для некоторых слов* $\bar{x} = (x_1, \ldots, x_n)$ *и слова* z

$$I(\bar{x} : z) = K(x) - K(x|z) \leq \delta.$$

Тогда соответствующие компоненты сложностных векторов $\vec{K}(\bar{x})$ *и* $\vec{K}(\bar{x}|z)$ *отличаются не более чем на* $\delta + \mathcal{O}(\log K(\bar{x}))$.

2. $\exists$-формулы.

Далее, перейдём к рассмотрению свойств, выражающихся $\exists$-формулами (с параметрами). В этом случае наша общая гипотеза может быть сформулирована в пары взаимно обратных утверждений. Прямое утверждение (теорема 2) тривиально; обратное утверждение (гипотеза 1) является главным открытым вопросом данной статьи:

Теорема 2 *Предположим, что для некоторых слов $\bar{x}$, z*

$$I(\bar{x} : z) \leq \delta.$$

Тогда для любого $\bar{y} = (y_1, \ldots, y_m)$ найдётся такой $\bar{y}' = (y'_1, \ldots, y'_m)$, что соответствующие компоненты сложностных профилей $\vec{K}(\bar{x}, \bar{y})$ и $\vec{K}(\bar{x}, \bar{y}'|z)$ отличаются друг от друга не более чем на $\delta + \mathcal{O}(\log K(\bar{x}, \bar{y}, z))$.

Гипотеза 1 *Предположим, для некоторых слов $\bar{x} = (x_1, \ldots, x_n)$ и z*

$$I(\bar{x} : z) \leq \delta.$$

Тогда для любого $\bar{y} = (y_1, \ldots, y_m)$ найдётся такой $\bar{y}' = (y'_1, \ldots, y'_m)$, что соответствующие компоненты сложностных профилей $\vec{K}(\bar{x}, \bar{y}|z)$ и $\vec{K}(\bar{x}, \bar{y}')$ отличаются не более чем на $\delta + \mathcal{O}(\log K(\bar{x}, \bar{y}, z))$.

Нам удаётся доказать гипотезу 1 только для стохастических слов.

Определение 1 *Слово x называется (α, β)-стохастическим, если существует множество A, содержащее x, такое что*

- *сложность кортежа $\hat{A}$, состоящего из всех элементов A в лексикографическом порядке, не превосходит α,*
- $K(x|\hat{A}) \geq \log |A| - \beta$

(здесь и далее в нашем тексте все логарифмы берутся по основанию 2). В частности, всякое несжимаемое слово длины N является $(\log N + \mathcal{O}(1), \mathcal{O}(1))$-стохастическим.

Определение стохастичности очевидным образом переносится на *кортежи* слов. Для наших целей наибольший интерес представляют $(\mathcal{O}(\log N), \mathcal{O}(\log N))$-стохастические кортежи $\bar{x}$, где $N = K(\bar{x})$. Такие кортежи мы для краткости называем просто *стохастическими*. Отметим, что сам по себе тот

факт, что не все слова являются нестохастическими, является весьма нетривиальным [13].

В приложениях комогоровской сложности как правило используются именно стохастические слова или кортежи слов. Поэтому изучение свойств стохастических наборов слов заслуживет особого внимания.

Теорема 3 *Гипотеза 1 выполнена для стохастических* $\bar{x}$.

3. Существование кортежей, неэквивалентных стохастическим

Будем говорить, что слова a, b C-эквивалентны ($a \sim_C b$), если

$$K(a|b) \le C \log K(a,b) \wedge K(b|a) \le C \log K(a,b).$$

Далее, будем называть кортежи $\bar{a} = (a_1, \ldots, a_n)$ и $\bar{b} = (b_1, \ldots, b_n)$ C-эквивалентными (обозначая это $\bar{a} \sim_C \bar{b}$), если для каждого $i = 1, \ldots, n$ слово a_i C-эквивалентно b_i. Поскольку мы изучаем колмогоровские свойства с логарифмической точностью, эквивалентные кортежи с нашей точки зрения неотличимы друг от друга.

Возникает искушение доказывать гипотезу 1, сводя произвольный кортеж $\bar{x}$ к эквивалентному ему стохастическому $\bar{x}'$ и применяя к $\bar{x}'$ теорему 3. Но можно ли для произвольного $\bar{x}$ найти эквивалентный ему стохастический $\bar{x}'$?

Прежде всего отметим, что для индивидуального слова x (для кортежа длины 1) можно найти C-эквивалентное ему $(C \log K(x), C \log K(x))$-стохастическое слово x'. Разумеется, мы требуем, чтобы константа C была достаточно большой, но независящей от x. В самом деле, для любого слова x найдётся кратчайшее описание p для данного слова x, такое что

$$K(p|x) = \mathcal{O}(\log K(x)).$$

Это p и можно взять в качестве x'. Очевидно, данное слово $\mathcal{O}(\log K(x))$-эквивалентно слову x. В то же время

$$K(x') \ge |x'| - \mathcal{O}(1),$$

так как оно является кратчайшим описанием x. Следовательно, x' стохастическое.

Верно ли аналогичное утверждение для пары $\langle x_1, x_2 \rangle$? Если x_1, x_2 независимы, то мы можем отдельно заменить каждое из x_i на эквивалентное стохастическое x_i'. Нетрудно видеть, что пара $\langle x_1', x_2' \rangle$ является

стохастической и эквивалентна $\langle x_1, x_2\rangle$. Другой простой пример: предположим, что $K(x_1|x_2) \le \mathcal{O}(\log K(x_2))$ (неформально это значит, что x_1 является частью x_2). Тогда найдутся такие несжимаемые слова p и q, что x_1 эквивалентно p, а x_2 эквивалентно $\langle p, q\rangle$. При этом пара $(p, \langle p, q\rangle)$ является стохастической.

Однако в общем случае для пары слов нельзя найти эквивалентную ей стохастическую. Сформулируем это утверждение более точно.

Теорема 4 *Пусть даны рациональные числа $\alpha, \beta, \gamma, C > 0$ такие, что $\alpha+\beta > \gamma$ и $\alpha, \beta < \gamma$. Тогда для достаточно больших n существует пара слов x_1, x_2, для которой*

- $K(x_1) = \alpha n + \mathcal{O}(\log n)$,
- $K(x_2) = \beta n + \mathcal{O}(\log n)$,
- $K(x_1, x_2) = \gamma n + \mathcal{O}(\log n)$,

и не существует $(C \log n, C \log n)$-стохастической пары x_1', x_2', которая была бы C-эквивалентна паре (x_1, x_2).

Замечание: условие $\alpha + \beta > \gamma$ гарантирует, что x_1 и x_2 зависимы, а условия $\alpha, \beta < \gamma$ показывают, что ни одно из слов x_i не может быть слишком просто относительно другого.

Отметим, что поскольку уже для кортежа длины 2 в общем случае нельзя подобрать эквивалентный ему стохастический кортеж, методы теоремы 3 не позволяют доказать гипотезу 1 в общем случае.

4. Ослабленные варианты гипотезы 1

Нетривиалным оказывается даже частный случай гипотезы 1 для $m = 1$:

Гипотеза 2 *Предположим, для некоторых слов $\bar{x} = (x_1, \ldots, x_n)$ и z*

$$I(\bar{x} : z) \le \delta.$$

Тогда для любого y найдётся y' такое, что сложностные профили $\vec{K}(\bar{x}, y|z)$ и $\vec{K}(\bar{x}, y')$ отличаются не более чем на $\delta + \mathcal{O}(\log K(\bar{x}, y, z))$.

Мы покажем, что гипотеза 2 эквивалентна следующему техническому утверждению (вопрос о его истинности был поставлен в [12]):

Гипотеза 3 *Для любой n-ки $\bar{x} = (x_1, \ldots, x_n)$ и любого w найдётся такое слово w', что*

- $K(x_{i_1}, \ldots, x_{i_k}|w') \leq K(x_{i_1}, \ldots, x_{i_k}|w) + \mathcal{O}(\log K(\bar{x}, w))$ *для любых* $1 \leq i_1 < \ldots < i_k \leq n$,
- $K(w') \leq I(\bar{x} : w) + \mathcal{O}(\log K(\bar{x}, w))$.

Кроме того, мы докажем ослабленный вариант этой гипотезы.

Теорема 5 *Для любой n-ки $\bar{x} = (x_1, \ldots, x_n)$ ($n > 1$) и любого w найдётся такое слово w', что*

- $K(x_{i_1}, \ldots, x_{i_k}|w') \leq K(x_{i_1}, \ldots, x_{i_k}|w) + \delta + \mathcal{O}(\log N)$ *для любых* $1 \leq i_1 < \ldots < i_k \leq n$,
- $K(w') \leq I(\bar{x} : w) + \delta + \mathcal{O}(\log N)$,

где $N = K(\bar{x}, w)$, и

$$\delta = I(\bar{x} : w) - \frac{1}{n-1}(\sum_{i=1}^{n} K(x_i) - K(\bar{x})) + \frac{2}{n-1}(\sum_{i=1}^{n} K(x_i|w) - K(\bar{x}|w)).$$

Отличие от гипотезы 3 состоит в добавочном члене δ, который в общем случае имеет порядок $\Omega(N)$.

Отметим частный случай теоремы 5 для $n = 2$.

Следствие 1 *Для любых x_1, x_2, w найдётся такое слово w', что*

- $K(x_1|w') \leq K(x_1|w) + \delta + \mathcal{O}(\log N)$,
- $K(x_2|w') \leq K(x_2|w) + \delta + \mathcal{O}(\log N)$,
- $K(x_1, x_2|w') \leq K(x_1, x_2|w) + \delta + \mathcal{O}(\log N)$,
- $K(w') \leq I(\bar{x} : w) + \delta + \mathcal{O}(\log N)$,

где $\delta = I(x_1 : x_2|w) + I(x_1 : w|x_2) + I(x_2 : w|x_1)$, а $N = K(x_1, x_2, w)$.

Данное следствие является усилением леммы 4 из [12], где был доказан аналогичный результат для в двое большего δ. Следствие 1 наиболее интересно в случае $\delta = \mathcal{O}(\log N)$, когда из трёх слов x_1, x_2, w каждые два независимы относительно третьего (см. применения этого утверждения в [12]).

Техника, используемая при доказательстве теоремы 5, основана на свойствах *гроздей*, вперве введённых в [12].

5. Выделение взаимной информации.

В [12] было показано, что в случае $I(x_1, x_2 : z) \approx 0$ взаимная информация пары $\langle x_1, x_2 \rangle$ может быть выделена при релятивизации относительно z, если и только если взаимная информация пары выделяется без релятивизации. Более точно, имеет место следующее утверждение.

Утверждение 1 *[12] Пусть для слов x_1, x_2, z и некоторой константы C*

$$I(x_1, x_2 : z) \leq C \log N,$$

(пара $\langle x_1, x_2 \rangle$ независима с z), и у слов x_1, x_2 выделяется (с точностью $C \log N$) взаимная информация при релятивизации относительно z, то есть

$$\exists w \ : \ K(w|x_i, z) \leq C \log N, \ i = 1, 2 \ и \ K(w|z) \geq I(x_1 : x_2) - C \log N.$$

Тогда для некоторой константы D (зависящей только от C) взаимная информация x_1 и x_2 выделяется с точностью $D \log N$ без релятивизации, то есть

$$\exists w' \ : \ K(w'|x_i) \leq D \log N, \ i = 1, 2 \ и \ K(w') \geq I(x_1 : x_2) - D \log N.$$

(Мы используем обозначение $N = K(x_1, x_2, z)$.)

Мы предполагаем, что верно более сильное утверждение: если $I(x_1, x_2 : z) \approx 0$ и у слов x_1, x_2 при релятивизации относительно z выделяется некоторая *часть* взаимной информации, то аналогичное свойство выполнено и без релятивизации. Нам не удаётся доказать это утверждение с логарифмической точностью. Мы покажем лишь, что оно верно с погрешностью $o(N)$. Аналогичное утверждение верно и для произвольного $n \geq 2$:

Теорема 6 *Для любой функции $f(N)$, $f(N) = o(N)$ найдётся функция $g(N)$ ($g(N) = o(N)$) такая, что для любых слов $z, \bar{x} = (x_1, \ldots, x_n)$ если $I(z : \bar{x}) \leq f(N)$ и*

$$\exists w : \ K(w|z) \geq \alpha, \ K(w|x_i, z) \leq f(N)(i = 1, \ldots, n),$$

где $N = K(\bar{x}, z)$, (то есть у слов $x_1, \ldots, x_n$ можно выделить α битов взаимной информации с погрешностью $f(N)$, имея z в качестве оракула), то

$$\exists y : \ K(y) \geq \alpha - g(N), \ K(y|x_i) \leq g(N)(i = 1, \ldots, n),$$

то есть те же α битов взаимной информации выделяются без релятивизации (с погрешностью $g(N)$).

Доказательство этого результата также использует метод гроздей.

1.2 Как организована статья

В разделе 2 мы приводим основные определения и формулируем несколько технических лемм. В разделе 3 мы доказываем основные результаты статьи – теоремы 1, 2, 3. В разделе 4 мы доказываем теорему 4. В разделе 5 мы показываем эквивалентность гипотез 2 и 3, доказываем теорему 5. В разделе 6 доказывается теорема 6. В приложении мы приводим доказательства некоторых технических лемм.

2 Определения и обозначения

2.1 Сложностные профили

Обозначение 1 *Пусть фиксирована n-ка слов $\bar{x} = x_1, \ldots, x_n$. Для любого множества $V = \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n\}$ ($1 \leq i_1 < \ldots < i_k \leq n$) будем обозначать $\bar{x}_V$ кортеж слов x_j с индексами $j \in V$:*

$$\bar{x}_V = \langle x_{i_1}, \ldots, x_{i_k} \rangle.$$

Мы будем использовать данное соглашение для обозначения колмогоровской сложности соответствующего кортежа:

$$K(\bar{x}_V) := K(x_{i_1}, \ldots, x_{i_k}).$$

Если $V = \emptyset$, то полагаем $K(\bar{x}_V) := K(\lambda)$ (где λ – пустое слово).

Аналогичное обозначение будем использовать для условных сложностей: для любых подмножеств $V = \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n\}$ и $W = \{j_1, \ldots, j_l\} \subseteq \{1, \ldots, n\}$ полагаем

$$K(\bar{x}_V|\bar{x}_W) := K(x_{i_1}, \ldots, x_{i_k}|x_{j_1}, \ldots, x_{j_l}).$$

При этом, если W пусто, то считаем $K(\bar{x}_V|\bar{x}_W) := K(\bar{x}_V|\lambda)$.

Определение 2 *Будем называть* **сложностным профилем** *n-ки слов $\bar{x} = \langle x_1, \ldots, x_n\rangle$ вектор, состоящий из $(2^n - 1)$ сложностей $K(\bar{x}_W)$ для всех непустых подмножеств $W \subseteq \{1, \ldots, n\}$ (считаем, что подмножества W располагаются в лексикографическом порядке):*

$$\vec{K}(x_1, \ldots, x_n) = (K(x_1), K(x_1, x_2), \ldots, K(x_2), K(x_2, x_3), \ldots).$$

Будем называть **относительным сложностным профилем** *n-ки $\bar{x}$ при условии y вектор, состоящий из $(2^n - 1)$ сложностей $K(\bar{x}_W|y)$ для всех непустых подмножеств $W \subseteq \{1, \ldots, n\}$ (снова полагаем, что все подмножества W располагаются в лексикографическом порядке):*

$$\vec{K}(x_1, \ldots, x_n|y) = (K(x_1|y), K(x_1, x_2|y), \ldots, K(x_2|y), K(x_2, x_3|y), \ldots).$$

Определение 3 *Будем называть* **расширенным сложностным профилем** *n-ки слов $x_1, \ldots, x_n$ вектор, состоящий из всех условных сложностей $K(\bar{x}_V|\bar{x}_W)$, где $V, W \subseteq \{1, \ldots, n\}$, причём $V \cap W = \emptyset$ и $V \neq \emptyset$. Заметим, что в случае $W = \emptyset$ мы получаем значение безусловной сложности: $K(\bar{x}_V|\bar{x}_\emptyset) = K(\bar{x}_V) + O(1)$. Считаем, что все пары подмножеств (V, W) располагаются в лексикографическом порядке:*

$$\vec{K}'(x_1, \ldots, x_n) = (K(x_1), K(x_1|x_2), \ldots, K(x_2|x_1), K(x_2|x_3), \ldots).$$

Аналогично определим **расширенный относительный сложностной профиль** *$x_1, \ldots, x_n$ при условии y. Для этого рассмотрим вектор из всех сложностей вида $K(\bar{x}_V|\bar{x}_W, y)$:*

$$\vec{K}'(x_1, \ldots, x_n|y) = (K(x_1|y), K(x_1|x_2, y), \ldots, K(x_2|x_1, y), K(x_2|x_3, y), \ldots).$$

Нам потребуется сравнивать сложностные профили для разных наборов слов. Для этого мы введём обозначения для сравнения векторов в $\mathbb{R}^k$.

Обозначение 2 *Будем говорить что сложностной вектор $\bar{\alpha} \in \mathbb{R}^n$ не больше вектора $\bar{\beta} \in \mathbb{R}^n$ (обозначение: $\bar{\alpha} \leq \bar{\beta}$), если каждая компонента первого вектора не превосходит соответствующей компоненты второго вектора, т.е. $\alpha_i \leq \beta_i$ для $i = 1, \ldots, n$.*

Кроме того, будем использовать l_∞-норму для измерения расстояния между векторами:

$$\rho(\bar{\alpha}, \bar{\beta}) := \max_i \{|\alpha_i - \beta_i|\}.$$

В частности, будем говорить что сложностной профиль для $\bar{x} = (x_1, \ldots, x_n)$ не больше сложностного профиля для $\bar{y} = (y_1, \ldots, y_n)$, если каждая компонента первого профиля не превосходит соответствующей компоненты второго профиля, т.е. для каждого $V \subseteq \{1, \ldots, n\}$ выполнено $K(\bar{x}_V) \leq K(\bar{y}_V)$. Аналогично, мы будем говорить, что расстояние между сложностными профилями кортежей $\langle x_1, \ldots, x_n \rangle$ и $\langle y_1, \ldots, y_n \rangle$ не превосходит ε, если для каждого набора индексов V выполнено $|K(\bar{x}_V) - K(\bar{y}_V)| \leq \varepsilon$.

2.2 Типизация

В дальнейшем мы будем применять прием 'типизации' (этот приём использовался в [11, 9, 5])

Определение 4 *Пусть даны наборы слов $\bar{x} = \langle x_1, \ldots, x_n \rangle$ и $\bar{y} = \langle y_1, \ldots, y_m \rangle$. Будем называть типизацией $\bar{x}$ относительно $\bar{y}$ следующее множество n-ок слов:*

$$T(\bar{x}|\bar{y}) := \{\bar{x}' = \langle x'_1, \ldots, x'_n \rangle \mid \vec{K}'(\bar{x}', \bar{y}) \leq \vec{K}'(\bar{x}, \bar{y})\}.$$

Далее, будем называть k-строгой типизацией $\bar{x}$ относительно $\bar{y}$ следующее множество:

$$ST_k(\bar{x}|\bar{y}) := \{\bar{x}' = \langle x'_1, \ldots, x'_n \rangle \mid \vec{K}'(\bar{x}', \bar{y}) \leq \vec{K}'(\bar{x}, \bar{y}) \text{ и } \rho(\vec{K}'(\bar{x}', \bar{y}), \vec{K}'(\bar{x}, \bar{y})) \leq k\}.$$

Отметим, что множество $T(\bar{x}|\bar{y})$ можно перечислять, зная $\bar{y}$ и все числа из расширенного профиля $\vec{K}'(\bar{x}, \bar{y})$. Множество $ST(\bar{x}|\bar{y})$ таким свойством не обладает.

Имеют место следующие леммы, доказанные в [11, 5].

Лемма 1 *Для любых наборов слов* $\bar{x} = (x_1, \ldots, x_n)$ *и* $\bar{y} = (y_1, \ldots, y_m)$

$$\log |T(\bar{x}|\bar{y})| = K(\bar{x}|\bar{y}) + O(\log N),$$

где $N = K(\bar{x}, \bar{y})$.

Лемма 2 *Для любых натуральных* n, m *существует* $C = C(n, m)$ *такое, что для всякой* n*-ки* $\bar{x} = (x_1, \ldots, x_n)$ *и всякой* m*-ки* $\bar{y} = (y_1, \ldots, y_m)$

$$|ST_{C \log N}(\bar{x}|\bar{y})| > \frac{1}{2}|T(\bar{x}|\bar{y})|,$$

где $N = K(\bar{x}, \bar{y})$. *При этом можно считать, что* C *вычислимо зависит от* m *и* n.

Обозначение 3 *Для краткости мы будем обозначать*

$$ST(\bar{x}|\bar{y}) = ST_{C \log N}(\bar{x}|\bar{y}),$$

где константа C *такая же, как в лемме 2.*

Также нам потребуетя следующий несложный технический факт:

Лемма 3 *Пусть даны наборы слов* $\bar{x} = \langle x_1, \ldots, x_n \rangle$, $\bar{y} = \langle y_1, \ldots, y_m \rangle$. *Имеют место следующие утверждения:*

(1) Для любого $\bar{x}' = (x'_1, \ldots, x'_n)$, *если* $\vec{K}'(\bar{x}', \bar{y}) \leq \vec{K}'(\bar{x}, \bar{y}) + \delta_1 \cdot \vec{e}$ *и* $K(\bar{x}', \bar{y}) \geq K(\bar{x}, \bar{y}) - \delta_2$, *то*

$$\rho(\vec{K}'(\bar{x}', \bar{y}), \vec{K}'(\bar{x}, \bar{y})) \leq (2\delta_1 + \delta_2) + O(\log N)$$

(здесь $N = K(\bar{x}, \bar{y})$, $\vec{e} = (1, \ldots, 1)$*).*

(2) Для любого z *и для любого* $\bar{x}' = (x'_1, \ldots, x'_n)$ *такого, что* $\vec{K}'(\bar{x}', \bar{y}) \leq \vec{K}'(\bar{x}, \bar{y}) + \delta_1 \cdot \vec{e}$ *и* $K(\bar{x}', \bar{y}|z) \geq K(\bar{x}, \bar{y}) - \delta_2$, *имеем*

$$\rho(\vec{K}'(\bar{x}', \bar{y}|z), \vec{K}'(\bar{x}, \bar{y})) \leq (2\delta_1 + \delta_2) + O(\log N),$$

где $\vec{e} = (1, \ldots, 1)$ *и* $N = K(\bar{x}, \bar{y})$.

Доказательство леммы приведено в Приложении.

2.3 Комбинаторная энтропия

Обозначение 4 *Пусть $A \subset X_1 \times \ldots \times X_n$ – некоторое множество n-ок (в качестве множеств X_i мы как правило будем брать конечные множества двоичных слов). Для произвольного набора индексов $I = \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n\}$ будем обозначать $\pi_I(A)$ проекцию A на соответствующие оси координат:*

$$\pi_I(A) := \{\bar{x}_I \mid \bar{x} \in A\}.$$

В частности, для $I = \{1, \ldots, n\}$ получаем $\pi_I(A) = A$.

Далее, для всякого кортежа $\bar{x} = \langle x_1, \ldots, x_k \rangle$ будем обозначать $\sigma_I(A|\bar{x})$ сечение A, соответствующее значению $\bar{x}$ проеции на оси I:

$$\sigma_I(A|\bar{x}) := \{\bar{y} \mid \bar{y} \in A \text{ и } \bar{y}_I = \bar{x}\}.$$

Обозначение 5 *Пусть $X_1, \ldots, X_n$ – некоторые конечные множества, и $A \subset X_1 \times \ldots \times X_n$ – произвольное множество n-ок. Будем использовать следующие обозначения:*

- $n_I(A)$ *– число элементов в $\pi_I(A)$.*
- $n_{I|J}(A|\bar{x})$ *– число элементов в $\pi_I\sigma_J(A|\bar{x})$.*
- $n_{I|J}(A) = \max\limits_{\bar{x} \in \pi_J(A)} n_{I|J}(A|\bar{x})$. *В частности, если $J = \emptyset$, то $n_{I|J}(A) = n_I(A)$.*

Определение 5 *Пусть $X_1, \ldots, X_n$ – некоторые конечные множества, и $A \subset X_1 \times \ldots \times X_n$ – произвольное множество n-ок. Для любых наборов индексов $I, J \subseteq \{1, \ldots, n\}$ полагаем*

- $\mathrm{ent}_I(A) := \lceil \log n_I(A) \rceil$.
- $\mathrm{ent}_{I|J}(A) := \lceil \log n_{I|J}(A) \rceil$ *(если $J = \emptyset$, то $\mathrm{ent}_{I|\emptyset}(A) = \mathrm{ent}_I(A)$).*

Лемма 4 *Пусть $A \subset \Sigma^n$ – некоторое конечное множество n-ок слов. Обозначим $list(A)$ список всех элементов A (в некоторой вычислимой кодировке). Тогда для любого $\bar{x} \in A$, для любых $V, W \subset \{1, \ldots, n\}$*

$$K(\bar{x}_V | \bar{x}_W, list(A)) \leq \mathrm{ent}_{V|W}(A) + O(1).$$

Доказательство очевидно. Имея список всех элементов A и кортеж $\bar{x}_W$ можно найти список всех элементов в множестве

$$B = \pi_V \sigma_W(A|\bar{x}_W).$$

При этом $\bar{x}_V \in B$. Чтобы указать $\bar{x}_V$, остаётся указать номер кортежа $\bar{x}_V$ в списке элементов B, что требует не более $\lceil \log n_{V|W}(A) \rceil$ битов.

3 Эквивалентность для бесквантерных и экзистенциальных формул

Доказательство теоремы 1 тривиально. С одной стороны, для любого набора индексов $V \subseteq \{1, \ldots, n\}$

$$K(\bar{x}_V|z) \leq K(\bar{x}_V) + O(1).$$

С другой стороны,

$$K(\bar{x}_V) - K(\bar{x}_V|z) = I(\bar{x}_V : z) \leq I(\bar{x} : z) + O(\log N) \leq \delta + O(\log N).$$

□

Доказательство теоремы 2. Требуется доказать, что расстояние между $\vec{K}(\bar{x}, \bar{y})$ и $\vec{K}(\bar{x}, \bar{y}'|z)$ не превосходит $\delta + O(\log N)$. Мы докажем формально более сильный факт: расстояние между соответствующими *расширенными* профилями не превосходит $\delta + O(\log N)$.

Рассмотрим множество $T(\bar{y}|\bar{x})$. Согласно лемме 1 в данном множестве содержится $2^{K(\bar{y}|\bar{x})+O(\log N)}$ m-ок слов. Следовательно, мы можем выбрать $\bar{y}' \in T(\bar{y}|\bar{x})$ такое, что

$$K(\bar{y}'|\bar{x}, z) \geq K(\bar{y}|\bar{x}) - O(\log N).$$

Для выбранного набора слов $\bar{y}'$

$$K(\bar{x}, \bar{y}'|z) = K(\bar{x}|z) + K(\bar{y}'|\bar{x}, z) \geq K(\bar{x}) - \delta + K(\bar{y}|\bar{x}) - O(\log N) = K(\bar{x}, \bar{y}) - \delta - O(\log N).$$

Следовательно, можно применить лемму 3, т.е.

$$\rho(\vec{K}'(\bar{x}, \bar{y}), \vec{K}'(\bar{x}, \bar{y}'|z)) \leq \delta + O(\log N).$$

□

Мы предполагаем, что верна гипотеза 1, являющаяся обращением теоремы 2. Нам не известно, верна ли эта гипотеза в общем случае. Далее мы докажем её для *стохастических* $\langle x_1, \ldots, x_n \rangle$.

Доказательство теоремы 3.

Шаг 1. Пусть $N = K(\bar{x}, \bar{y})$. Заменим каждое из слов $x_1, \ldots, x_m, y_1, \ldots, y_m$ на соответствующую ему кратчайшую программу (в оптимальном языке программирования). При этом величины сложностей в расширенном профиле $\vec{K}'(\bar{x}, \bar{y}|z)$ изменятся не более чем на $O(\log N)$, а также сохранится стохастичность набора $x_1, \ldots, x_m$. Далее без ограничения общности мы будем предполагать, что $\bar{x} \in (\mathbb{B}^N)^n$, $\bar{y} \in (\mathbb{B}^N)^m$. По условию n-ка $\bar{x}$ является стохастической, то есть лежит в некотором простом множестве $S \subset (\mathbb{B}^N)^n$:

$$K(S) = O(\log N),$$

$$\log |S| = K(\bar{x}) + O(\log N).$$

Таким образом, $\langle \bar{x}, \bar{y} \rangle \in S \times (\mathbb{B}^N)^m \subseteq (\mathbb{B}^N)^{n+m}$.

Шаг 2. Рассмотрим множество $A_0 = T(\bar{x}, \bar{y}|z) \cap S \times (\mathbb{B}^N)^n$. Размеры проекций и сечений A_0 не превосходят экспоненты от соответствующих значений расширенного сложностного профиля $\vec{K}'(\bar{x}, \bar{y}|z)$. Будем называть множество $A \subset S \times (\mathbb{B}^N)^m$ *правильным*, если все его комбинаторные энтропии $ent_{I|J}(A)$ не превосходят соответствующих комбинаторных энтропий A_0. Другими словами, множество A называется правильным, если для любых $I, J \subseteq \{1, \ldots, (n+m)\}$

$$\log n_{I|J}(A) \leq ent_{I|J}(A_0).$$

В частности, само множество A_0 является правильным. Отметим, что из правильности A следует $n_{I|J}(A) < 2n_{I|J}(A_0)$.

Зная значения всех координат расширенного сложностного профиля $\vec{K}'(\bar{x}, \bar{y}|z)$ и все элементы множества S, можно алгоритмически найти список *всех* правильных множеств (этот список чрезвычайно велик, но конечен!). Поскольку список всех элементов множества S имеет лишь логарифмическую сложность, список всех правильных множеств также может быть получен со сложностью $O(\log N)$. Обозначим $A_1, A_2, \ldots$ лексикографически упорядоченный список всех правильных множеств.

Согласно определению, размер каждого сечения правильного множества не более чем в двое превосходит размер соответствующего сечения A_0. При этом, однако, некоторые сечения могут быть значительно меньше указанной границы.

Назовём *сильной проекцией* множества A_i на первые n координат (т.е. на S) множество B_i, состоящее из точек проекции множества A_i на S, которым соответствуют достаточно большие сечения:

$$B_i = \{\bar{x}' \in \pi_{1,\ldots,n}(A_i) \mid \log|\sigma_{1,\ldots,n}(A_i|\bar{x}')| \geq ent_{n+1,\ldots,n+m|1,\ldots,n}(A_0) - C_1 \log N\}$$

(константу C_1 мы выберем ниже). В частности, сильную проекцию множества A_0 будем называть B_0. Для дальнейшего фиксируем алгоритмы, первый из которых по слову z и набору чисел $\vec{K}'(\bar{x}, \bar{y})$ перечисляет A_0, а второй по слову z, числам $\vec{K}'(\bar{x}, \bar{y})$ и C_1 перечисляет B_0.

Заметим, что зная z и расширенный профиль $\vec{K}'(\bar{x}, \bar{y}|z)$, мы можем перечислять элементы сечения A_0, соответствующего значению $\bar{x}$ (в наших обозначениях это $\sigma_{1,\ldots,n}(A_0|\bar{x})$). Следовательно,

$$ent_{n+1,\ldots,n+m|1,\ldots,n}(A_0) \leq K'(\bar{y}|\bar{x}, z) \leq \log|\sigma_{1,\ldots,n}(A_0|\bar{x})| + O(\log N).$$

Таким образом, мы можем выбрать такое значение C_1, чтобы $\bar{x}$ принадлежало B_0.

Шаг 3. Теперь выберем из последовательности правильных множеств специальную подпоследовательность по следующем правилу. Пусть правильные множества $A_1, \ldots, A_{s-1}$ уже рассмотрены, причем $A_{i_1}, \ldots, A_{i_k}$ включены в подпоследовательность. Очередное по списку правильное множество A_s включается в подпоследовательность, если разность

$$B_s \setminus \left(\bigcup_{r \leq k} B_{i_r} \right)$$

имеет мощность не меньше $2^{K(\bar{x}|z) - C_2 \log N}$ (константа C_2 будет выбрана позже).

Отметим, что в данную подпоследовательность будет включено не более $\frac{|S|}{2^{K(\bar{x}|z) - C_2 \log N}} = 2^{I(\bar{x}:z) + C_2 \log N + O(\log N)}$ правильных множеств. Обозначим $\hat{A}$ объединение всех правильных множеств из выбранной подпоследовательности $A_{i_1}, A_{i_2}, \ldots$. Проекцию $\hat{A}$ на первые n координат обозначим $\hat{B}$.

Очевидно, $K(\hat{A}) = O(\log N + \log C_2)$ и $K(\hat{B}) = O(\log N + \log C_2)$, т.к. список элементов этих множеств может быть найден алгоритмически, если известны расширенный сложностной профиль $\vec{K}'(\bar{x}, \bar{y}|z)$, множество S и константа C_2.

Замечание. Очевидно, что

$$\log n_{I|J}(\hat{A}) \le \mathrm{ent}_{I|J}(A_0) + I(\bar{x} : z) + C_2 \log N + O(\log N)$$

для любых $I, J \subset \{1, \ldots, n\}$. *При этом список всех элементов* $\hat{A}$ *может быть получен со сложностью* $O(\log N + \log C_2)$. *Применяя лемму 4, находим, что для любого* $\bar{u} \in \hat{A}$

$$K(\bar{u}_I|\bar{u}_J) \le \mathrm{ent}_{I|J}(A_0) + I(\bar{x} : z) + C_2 \log N + O(\log N + \log C_2)$$

для любых $I, J \subset \{1, \ldots, n\}$.

Лемма 5 $\bar{x} \in \hat{B}$ *(при подходящем выборе* $C_2 = C_2(n, m)$*).*

Доказательство леммы. Предположим противное: $\bar{x}$ не принадлежит сильной проекции построенного множества $\hat{A}$. В этом случае множество A_0 (которое входит в список всех правильных множеств) не было включено в выбранную подпоследовательность. Но это значит, что мощность разности $B_0 \setminus \hat{B}$ заведомо меньше, чем $2^{K(\bar{x}|z) - C_2 \log N}$. Следовательно, чтобы задать n-ку $\bar{x}$ при известных z и $\vec{K}'(\bar{x}, \bar{y})$ достаточно указать список всех элементов $\hat{B}$, расположенных в порядке перечисления B_0 и номер кортежа $\bar{x}$ в списке всех элементов множества $B_0 \setminus \hat{B}$ в порядке перечисления. Таким образом,

$$K(\bar{x}|z) \le \log |(B_0 \setminus \hat{B})| + O(\log N + \log C_2) \le K(\bar{x}|z) - C_2 \log N + O(\log N + \log C_2).$$

Выбирая достаточно большую константу C_2, получаем противоречие. Лемма доказана.

Шаг 4. Таким образом, $\bar{x} \in \hat{B}$. Обозначим Q сечение множества $\hat{A}$, соответствующее $\bar{x}$:

$$Q = \{\bar{y} | \langle \bar{x}, \bar{y} \rangle \in \hat{A}\}.$$

Из построения $\hat{A}$ следует, что число элементов в Q не может быть слишком мало. Точнее,

$$\log |Q| \ge K(\bar{y}|\bar{x}, z) - \mathcal{O}(\log N).$$

Остается выбрать из Q m-ку $\bar{y}'$, имеющую максимальную возможную сложность относительно $\bar{x}$. А именно, существует $\bar{y}' \in \hat{S}$, для которой

$$K(\bar{y}'|\bar{x}) \ge K(\bar{y}|\bar{x}, z) - \mathcal{O}(\log N).$$

Таким образом, поскольку

$$K(\bar{x}) \geq K(\bar{x}|z) + I(\bar{x} : z) - \mathcal{O}(\log N),$$

мы получаем

$$K(\bar{x}, \bar{y}') \geq K(\bar{x}, \bar{y}|z) + I(\bar{x} : z) - \mathcal{O}(\log N).$$

С другой стороны, из построения $\hat{A}$ следует (см. замечание выше), что

$$\vec{K}'(\bar{y}'|\bar{x}) \leq \vec{K}'(\bar{y}|\bar{x}) + (I(\bar{x} : z) - \mathcal{O}(\log N)) \cdot \vec{e}.$$

Полагая $\delta_1 = -\delta_2 = I(\bar{x} : x)$ в лемме 3 имеем

$$\rho(\vec{K}'(\bar{x}, \bar{y}'), \vec{K}'(\bar{x}, \bar{y}|z)) \leq I(\bar{x} : z) + \mathcal{O}(\log N),$$

что заканчивает доказательство теоремы. □

4 Не для всякой пары существует эквивалентная ей стохастическая

Доказательство теоремы 4.

Для краткости мы будем называть $(C \log n, C \log n)$-стохастические пары просто стохастическими. Зафиксируем достаточно большое n. Обозначим S_1 множество слов длины αn и S_2 множество слов длины βn. Мы построим некоторое перечислимое множество $A \subset S_1 \times S_2$ размера $2^{\gamma n - \mathcal{O}(\log n)}$. При этом мы покажем, что некоторый элемент A является искомым – он имеет требуемый сложностной профиль и для него нет C-эквивалентной стохастической пары.

Нам будет удобно представлять A как множество рёбер в двудольном графе с левой долей S_1 и правой долей S_2.

Всякое слово x_1' C-эквивалентное некоторому слову из S_1 имеет сложность не более $\alpha n + 2C \log n$. Обозначим множество всех слов с такими сложностями L_1. Аналогично, всякое слово x_2', эквивалентное некоторому слову из S_2, имеет сложность не более $\beta n + 2C \log n$. Множество слов с таким сложностями мы обозначим L_2.

Нас будут интересовать множества $R \subset L_1 \times L_2$ такие, что

- $|R| \leq 2^{\gamma n + 3C \log n}$,

- $K(R) \leq C \log N$, то есть список элементов R можно получить с помощью алгоритма сложности не более $C \log N$.

Очевидно, если пара слов (x_1, x_2) из множества A C-эквивалентна некоторой стохастической паре (x_1', x_2'), то данная пара (x_1', x_2') должна принадлежать одному из множеств R указанного вида. Число всех таких множеств R не превосходит $C \log n$. Зная их точное число (чтобы сообщить это число, нужно логарифмическое число битов), мы можем найти все множества R. Обозначим $\hat{R}$ их объединение. Множество $\hat{R}$ также удобно рассматривать как двудольный граф; правую и левую доли вершин составляют L_1 и L_2 соответственно.

Таким образом, с помощью программы длины $\mathcal{O}(\log n)$ мы можем получить двудольный граф $\hat{R}$. Остаётся построить граф A, у которого будет достаточно много рёбер неэквивалентных ни одному из рёбер $\hat{A}$.

Некоторая трудность состоит в том, что отношение C-эквивалентности не является разрешимым. Однако мы можем описать относительно небольшой (и вычислимый) класс отношений, который будет заведомо содержать интересующее нас отношение C-эквивалентности. А именно, назовём *отношением сходства* всякое

$$D \subset S_1 \times L_1 \cup S_2 \times L_2,$$

удовлетворяющее следующим условиям:

- для любого $x \in S_i$ имеется не более $2^{C \log n + 1}$ элементов $y \in L_i$, для которых $(x, y) \in D$;
- для любого $y \in L_i$ имеется не более $2^{C \log n + 1}$ элементов $x \in S_i$, для которых $(x, y) \in D$,

$i = 1, 2$. Очевидно, отношение

$$D_0 = \{(x, y) \in S_1 \times L_1 \cup S_2 \times L_2 \ : \ x \sim_C y\}$$

удовлетворяет данным условиям. Заметим, что общее число отношений сходства не превосходит

$$(|L_1|^{poly(n)})^{|S_1|} \cdot (|L_2|^{poly(n)})^{|S_2|}.$$

Без ограничения общности можно считать, что $\alpha \geq \beta$. Тогда число различных отношений сходства равно

$$2^{2^{\alpha n + \mathcal{O}(\log n)}}.$$

Будем говорить, что ребро $(x_1, x_2) \in A$ D-сходно с ребром $(x_1', x_2') \in \hat{R}$, если $(x_i, x_i') \in D$ для $i = 1, 2$.

Теперь мы готовы перейти к построению графа A. Будем называть ребро $(x_1, x_2) \in A$ *правильным*, если степени вершин x_1 и x_2 не превосходили $2^{(\gamma-\alpha)n+C_1 \log n}$ и $2^{(\gamma-\beta)n+C_1 \log n}$ соответственно. Будем требовать, чтобы

$$|A| = 2^{\gamma n - C_1 \log n},$$

а также чтобы выполнялось следующее условие:

$$\begin{gathered} \textit{Для любого отношения сходства } D \textit{ найдется не менее} \\ 2^{\gamma n - C_2 \log n} \textit{правильных рёбер } (x_1, x_2) \in A, \\ \textit{не являющихся } D\textit{-сходными ни с одним ребром из } \hat{R} \end{gathered} \tag{2}$$

(константы C_1, C_2 будут выбраны позднее).

Мы построим такое множество A эффективно, то есть его сложность будет равна $\mathcal{O}(\log n)$. При этом не менее $2^{\gamma n - C_2 \log n}$ рёбер из A не будут иметь эквивалентных стохастических пар. Остаётся лишь выбрать среди данных рёбер одно, имеющее сложность не менее $\gamma n - C_2 \log n$. Для выбранной пары (x_1, x_2) будут выполнены условия

$$K(x_1) \leq \alpha n, K(x_2) \leq \beta n,$$

а также

$$K(x_2|x_1) \leq (\gamma - \alpha)n + C_1 \log n$$

и

$$K(x_1|x_2) \leq (\gamma - \beta)n + C_1 \log n.$$

Кроме того, при достаточно большой константе C_1

$$K(x_1, x_2) < \gamma n,$$

и тем самым теорема будет доказана.

Отметим, что свойство (2) можно проверить алгоритмически. Мы покажем, что для множества, состоящего из $2^{\gamma n - C_1 \log n}$ случайно выбранных рёбер из $S_1 \times S_2$, с положительной вероятностью выполнено условие (2). Тем самым будет доказано, что множества с нужными нам свойствами существуют, и мы сможем найти одно из них (скажем, лексикографически первое) перебором.

Зафиксируем одно из отношений сходства D. Назовём ребро $(x_1, x_2) \in S_1 \times S_2$ *плохим*, если для него найдется D-сходное ребро в $\hat{R}$. Подсчитаем вероятность оказаться плохим для ребра случайно выбранного среди всех пар $S_1 \times S_2$. Граф $\hat{R}$ содержит $2^{\gamma n + \mathcal{O}(\log n)}$ рёбер. Для каждого из них имеется не более $poly(n)$ D-сходных пар в $S_1 \times S_2$. Следовательно,

$$Prob[(x_1, x_2) \text{ плохое }] \leq \frac{2^{\gamma n + \mathcal{O}(\log n)}}{2^{(\alpha+\beta)n}} \ll 1/2.$$

Здесь мы использовали условие $\alpha + \beta > \gamma$.

Пусть $k = 2^{\gamma n - C_1 \log n}$ и $l \leq 2^{\gamma n - C_2 \log n + 1}$. Тогда вероятность того, что среди k случайно выбранных рёбер $(k - l)$ оказались плохими, не превосходит

$$C_k^l (1/2)^{k-l} \leq k^l (1/2)^{k-l} \leq \frac{1}{2^{2^{\gamma n - \mathcal{O}(\log n)}}}$$

при достаточно большой разнице между C_1 и C_2. Теперь мы должны просуммировать данную вероятность по всем $l \leq 2^{\gamma n - C_2 \log n + 1}$. Ясно, что умножение полученной вероятности на число различных l (то есть на $2^{\gamma n}$) не изменит существенно асимптотики убывания.

Итак, мы оценили вероятность того, что случайно выбранное A содержит много плохих рёбер для фиксированного отношения сходства D. Остаётся просуммировать эту вероятность по всем отношениям сходства. Это даст оцеку вероятности того, что в случайно выбранном множестве A для любого отношения сходства D (в том числе и для собственно отношения C-экаивалентности) не менее $2^{\gamma n - C_2 \log n + 1}$ рёбер не эквивалентны ни одному из рёбер $\hat{R}$. А именно, данная вероятность не меньше

$$\frac{2^{2^{\alpha + \mathcal{O}(\log n)}}}{2^{2^{\gamma n - \mathcal{O}(\log n)}}} < 1.$$

Здесь мы использовали условие $\alpha < \gamma$.

Мы доказали, что с положительной вероятностью $(1 - p_0)$ случайно выбранное множество A содержит не менее $2^{\gamma n - C_2 \log n + 1}$ рёбер, не эквивалентных ни одному из пар в $\hat{R}$. Однако некоторые из этих рёбер могут оказаться *неправильными* (одна из вершин имеет слишком большую степень). Покажем, что с большой вероятностью таких рёбер не более половины. Точнее, вероятность того, что в случайном графе окажется больше $2^{\gamma n - C_2 \log n}$ неправильных рёбер, не преосходит некоторого $p_1 < (1 - p_0)$. Таким образом, мы докажем, что с положительной вероятностью множество A содержит не менее $2^{\gamma n - C_2 \log n}$ рёбер, которые

одновременно являются и *хорошими*, и *правильными*, то есть удовлетворяют условию (2).

Рассмотрим произвольную вершину $x \in S_1$. Среднее число рёбер, инцедентных x в случайно выбранним A, равно

$$\frac{2^{\gamma n - C_1 \log n}}{2^{\alpha n}} = 2^{(\gamma-\alpha)n - C_1 \log n}.$$

Из неравенства Чебышёва следует, что верятность того, что вершина x инцедентна более чем $2^{(\gamma-\alpha)n+C_1 \log n}$ рёбрам, не превосходит $1/n^{2C}$. Следовательно, математическое ожидание числа вершин из S_1, имеющих аномально большую степень (более $2^{(\gamma-\alpha)n+C_1 \log n}$) не превосходит $2^{\alpha n}/n^{2C_1}$. Ещё раз воспользуемся неравенством Чебышёва: вероятность того, что число вершин в S_1 с аномально большой степенью оказалось больше $2^{(\gamma-\alpha)n-C_1 \log n}$, не может быть больше $1/n^C$. Аналогично, верятность того, что число вершин в S_2 с аномально большой (более $2^{(\gamma-\beta)n+C_1 \log n}$) степенью превысит $2^{\beta n - C_1 \log n}$, также не может быть больше $1/n^C$. При больших n

$$p_1 \geq 1 - 2/n^c \gg 1 - p_0.$$

Таким образом, сушествование множества A, удовлетворяющего (2), доказано. □

5 Доказательство ослабленной гипотезы

В этом разделе мы докажем эквивалентность гипотез 2 и 3 и теорему 5.

Замечание. Для слова w' в гипотезе 3 мы имеем $K(w') \leq I(\bar{x} : w) + \mathcal{O}(\log N)$ и $K(\bar{x}|w') \leq K(\bar{x}|w) = K(\bar{x}) - I(\bar{x} : w)$. Следовательно, $K(w'|\bar{x}) = \mathcal{O}(\log N)$.

Утверждение 2 *Гипотеза 2 и гипотеза 3 эквивалентны.*

Доказательство. (1) **Гипотеза 2** $\Rightarrow$ **Гипотеза 3**. Пусть даны n-ка $\bar{x}$ и слово w. Согласно [6] найдется слово v такое, что

1. $K(v) = K(w|\bar{x}) + \mathcal{O}(\log N)$,
2. $K(v|w) = \mathcal{O}(\log N)$,
3. $K(w|\bar{x}, v) = \mathcal{O}(\log N)$.

Говоря неформально, v есть кратчайшая (с точностью до логарифмического слагаемого) программа, перерабатывающая $\bar{x}$ в слово w, причём v просто относительно w. Легко видеть, что

$$I(\bar{x} : v) = \mathcal{O}(\log N).$$

Теперь применим гипотезу 2 к заданному $\bar{x}$, $z = v$ и $y = w$. Получаем слово y' такое, что сложностные профили $\vec{K}'(\bar{x}, y')$ и $\vec{K}'(\bar{x}, w|v)$ отличаются не более чем на $\mathcal{O}(\log N)$. Отсюда следуют два факта:

1. $K(y') = K(w|v) + \mathcal{O}(\log N) = I(\bar{x} : w) + \mathcal{O}(\log N)$;
2. $\rho(\vec{K}'(\bar{x}|w, v), \vec{K}'(\bar{x}|y')) = \mathcal{O}(\log N)$, а значит $\vec{K}'(\bar{x}|y') \leq \vec{K}'(\bar{x}|w, v) + \mathcal{O}(\log N) \cdot \vec{e}_n \leq \vec{K}'(\bar{x}|w) + \mathcal{O}(\log N) \cdot \vec{e}_n$.

Таким образом, мы можем взять слово y' в качестве w'.

(2) **Гипотеза 3 $\Rightarrow$ Гипотеза 2**.

Прежде всего применим [6] и получим слово z' такое, что

1. $K(z') = K(z|\bar{x}) + \mathcal{O}(\log N) \geq K(z) - \delta - \mathcal{O}(\log n)$,
2. $K(z'|z) = \mathcal{O}(\log N)$,
3. $K(z|\bar{x}, z') = \mathcal{O}(\log N)$.

Нетрудно видеть, что для данного z'

$$I(z' : \bar{x}) = \mathcal{O}(\log N)$$

и

$$\rho(\vec{K}'(\bar{x}, y|z), \vec{K}'(\bar{x}, y|z')) \leq \delta + \mathcal{O}(\log N).$$

Таким образом, мы можем заменить слово z на z', изменив относительный сложностной профиль $\langle \bar{x}, y \rangle$ не более чем на $\delta + \mathcal{O}(\log N)$.

Далее, назовём $w = \langle y, z' \rangle$. Согласно гипотезе 3 найдется w' сложности $I(\bar{x} : w) = I(\bar{x} : y|z')$ такое, что

$$\vec{K}'(\bar{x}|w') \leq \vec{K}'(\bar{x}|y, z') + \mathcal{O}(\log N) \cdot \vec{e}_n.$$

Как мы замечали выше, для такого w'

$$K(w'|\bar{x}) = \mathcal{O}(\log N).$$

Далее, для любого набора индексов $I \subseteq \{1,\dots,n\}$ и $\bar{I} = \{1,\dots,n\} \setminus I$

$$\begin{aligned} K(x_I|yz') + K(x_{\bar{I}}|x_I, y, z') &= K(\bar{x}) - I(\bar{x} : y, z') = \\ &= K(\bar{x}) - I(\bar{x} : y|z') = \\ &= K(\bar{x}) - K(w') = \\ &= K(\bar{x}) - I(w' : \bar{x}) \\ &= K(\bar{x}|w') \le K(x_I|w') + K(x_{\bar{I}}|x_I, w') \\ &= \le K(x_I|w') + K(x_{\bar{I}}|x_I, y, z'). \end{aligned}$$

Таким образом, $K(x_I|yz) \le K(x_i|w') + \mathcal{O}(\log N)$. Следовательно, $\vec{K}(\bar{x}|y,z') \le \vec{K}(\bar{x}|w') + \mathcal{O}(\log N)\cdot e_n$. Таким образом,

$$\rho(\vec{K}(\bar{x}|w'), \vec{K}(\bar{x}|y,z')) = \mathcal{O}(\log N).$$

Остаётся выбрать произвольное слово w'' сложности $K(y|\bar{x},z)$ независимое от $\bar{x}, w'$:

$$K(w''|x,w') = K(y|\bar{x},z),\ \ I(w'' : x, w') = \mathcal{O}(\log N).$$

Положим $y' = \langle w', w''\rangle$. Нетрудно проверить, что $\rho(\vec{K}(\bar{x},y'), \vec{K}(\bar{x},y|z')) = \mathcal{O}(\log N)$, а значит $\rho(\vec{K}(\bar{x},y'), \vec{K}(\bar{x},y|z)) \le \delta + \mathcal{O}(\log N)$. $\square$

Далее мы докажем теорему 5, являющуюся ослабленным вариантом гипотезы 3.

Доказательство теоремы 5

Шаг 1. Рассмотрим строгую типизацию w относительно $\bar{x}$:

$$ST(w|\bar{x}) = \{\hat{w} \mid \vec{K}'(\bar{x},\hat{w}) \le \vec{K}'(\bar{x},w))\}.$$

Напомним,

$$\log|ST(\hat{w}|\bar{x})| \ge K(w|\bar{x}) - k_1 \log N$$

Фиксируем произвольное $w_1 \in ST(w|\bar{x})$. Заметим, что для более чем половины $w_2 \in T'(w|\bar{x})$

$$I(w_1 : w_2|\bar{x}) \le k_2 \log N$$

(k_2 зависит только от n).

Шаг 2. Покажем, что для любого $w_1 \in T'(w|\bar{x})$ для более чем половины $w_2 \in T'(w|\bar{x})$

$$I(w_1 : w_2) \ge \frac{1}{n-1}(\sum K(x_i) - K(\bar{x})) - \frac{2}{n-1}(\sum K(x_i|w) - K(\bar{x}|w)).$$

В самом деле, для любых $\bar{x} = \langle x_1 \ldots, x_n \rangle, w_1, w_2$ выполнено неравенство

$$\begin{aligned} \sum K(x_i) + K(\bar{x}, w_1, w_2) &+ (n-1)K(w_1, w_2) \\ &\leq \sum (K(x_i, w_1) + K(x_i, w_2)) \end{aligned} \tag{3}$$

(доказательство см. в приложении). Напомним, что кортеж $\bar{x}$ нам задан в условии теоремы, а слово w_1 мы зафиксировали выше. Далее, заметим, что для фиксированных $\bar{x}, w_1$ и для более чем половины $w_2 \in T'(w|\bar{x})$ мы имеем $I(w_1 : w_2|\bar{x}) = \mathcal{O}(\log N)$. Следовательно,

$$\begin{aligned} K(x_i, w_1, w_2) &= K(x_i) + K(w_1|x) + K(w_2|x) + \mathcal{O}(\log N) \\ &= K(x_i) + 2K(w|x) + \mathcal{O}(\log N). \end{aligned}$$

Для таких w_2 неравенство (3) принимает вид

$$\begin{aligned} (n-1)I(w_1 : w_2) &= (n-1)(K(w_1) + K(w_2) - K(w_1, w_2)) \\ &\geq (n-1)(K(w_1) + K(w_2)) + K(\bar{x}, w_1, w_2) - \mathcal{O}(\log N) \\ &\quad + \sum (K(x_i) - K(x_i, w_1) - K(x_i, w_2)) - \mathcal{O}(\log N) \\ &= K(\bar{x}) + 2K(w|\bar{x}) + 2(n-1)K(w) + \sum (K(x_i) - 2K(x_i, w)) - \mathcal{O}(\log N) \\ &= \sum K(x_i) - K(\bar{x}) - 2(\sum K(x_i|w) - K(\bar{x}|w)) - \mathcal{O}(\log N). \end{aligned}$$

Таким образом,

$$\begin{aligned} K(w_1|w_2) &\leq K(w) - \tfrac{1}{n-1}(\sum K(x_i) - K(\bar{x})) + \tfrac{2}{n-1}(\sum K(x_i|w) - K(\bar{x}|w)) \\ &= K(w|\bar{x}) + \delta(\bar{x}, w), \end{aligned}$$

где $\delta(\bar{x}, w) = I(w : \bar{x}) - \frac{1}{n-1}(\sum K(x_i) - K(\bar{x})) + \frac{2}{n-1}(\sum K(x_i|w) - K(\bar{x}|w))$.

Шаг 3. В [12] было сформулировано следующее определение *грозди слов*:

Определение 6 *Назовём (α, β, γ)-гроздью множество слов X такое, что*

1. $|X| = 2^\alpha$,
2. $K(x_1|x_2) < \beta$ *для всех* $x_1, x_2 \in X$,
3. $K(x) < \gamma$ *для всех* $x \in X$.

Лемма 6 ([12]) *Для любых натуральных чисел α, β, γ можно построить алгоритм сложности $\mathcal{O}(\log \gamma)$, который перечисляет некоторый список (α, β, γ)-гроздей $U_0, \ldots, U_q$ со следующими свойствами:*

- *для любой* (α, β, γ)*-грозди* U *найдется номер* $i \leq q$ *такой, что* $|U \cap U_i| \geq 2^{\beta - \epsilon}$*, где* $\epsilon = 2(\beta - \alpha) + \mathcal{O}(1)$*,*
- $q < 2^{\beta + \gamma - 2\alpha + \mathcal{O}(1)}$.

Нам потребуется несколько модифицировать определение грозди слов.

Определение 7 *Назовём* (α, β, γ)*-полугроздью множество слов* X *такое, что*

1. $|X| = 2^{\alpha}$*,*
2. *для любого* $x_1 \in X$ *для более чем половины* $x_2 \in X$ *выполнено* $K(x_1|x_2) < \beta$
3. $K(x) < \gamma$ *для всех* $x \in X$*.*

Следующая лемма аналогична лемме 6 о свойствах гроздей.

Лемма 7 *Для любых натуральных чисел* α, β, γ *можно построить алгоритм сложности* $\mathcal{O}(\log \gamma)$*, который перечисляет некоторый список* (α, β, γ)*-полугроздей* $U_0, \ldots, U_q$ *со следующими свойствами:*

- *для любой* (α, β, γ)*-полугрозди* U *найдется номер* $i \leq q$ *такой, что* $|U \cap U_i| \geq 2^{\beta - \epsilon}$*, где* $\epsilon = 2(\alpha - \beta) + \mathcal{O}(1)$*,*
- $q < 2^{\beta + \gamma - 2\alpha + \mathcal{O}(1)}$.

Мы не приводим Доказательство леммы 7, поскольку оно совершенно аналогично доказательству леммы 6 в [12]. Полугрозди $U_0, \ldots, U_q$ из леммы 7 мы будем называть *стандартными* (для заданных α, β, γ). Очевидно, для любых α, β, γ и любого $i \leq q$ сложность списка элементов стандартной полугрозди U_i при известном i равна $\mathcal{O}(\log \gamma)$.

Теперь заметим, что по доказанному выше (шаг 2) множество $T'(w|\bar{x})$ является полугроздью с параметрами

$$(K(w|\bar{x}) - \mathcal{O}(\log N), K(w|\bar{x}) + \delta + \mathcal{O}(\log N), K(w)).$$

Таким образом, по лемме 7 найдется некоторая стандартная полугроздь U_i, для которой

$$|T(w|\bar{x}) \cap U_i| \geq 2^{K(w|\bar{x}) - \delta - \mathcal{O}(\log N)}.$$

В качестве w' мы возьмём двоичную запись числа i. Поскольку $i \leq 2^{I(w:\bar{x})+\delta+\mathcal{O}(\log N)}$, получаем

$$K(v) \leq I(w:\bar{x}) + \delta + \mathcal{O}(\log N).$$

Шаг 4. Нам остаётся доказать, что

$$\vec{K}(\bar{x}|w') \leq \vec{K}(\bar{x}|w) + (\delta + \mathcal{O}(\log N))\vec{e}_n.$$

Требуется проверить, что для любого $I \subseteq \{1,\ldots,n\}$ сложность $K(x_I|v)$ ненамного больше $K(x_I|w)$. Обозначим $\hat{x} = x_I$. Заметим, что для любого

$$\hat{w} \in T(w|\bar{x}) \cap U_i$$

выполнены условия

1. $K(\hat{w}|w') \leq \log|U_i| + \mathcal{O}(\log n) \leq K(w|\bar{x}) + \mathcal{O}(\log N)$ (поскольку зная w' и $\mathcal{O}(\log N)$ битов дополнительной информации можно перечислять множество U_i),
2. $K(\hat{x}|\hat{w}) \leq K(\hat{x}|w)$ (из определения $T(w|\bar{x})$).

Таким образом, $\hat{x}$ принадлежит множеству

$$X = \{\hat{x}' \quad | \quad \text{существует не менее } 2^{K(w|\bar{x})-\mathcal{O}(\log N)} \text{ слов } \hat{w}, \text{ для которых } K(\hat{w}|w') \leq K(w|\bar{x}) + \mathcal{O}(\log N) \text{ и } K(\hat{x}|\hat{w}) \leq K(\hat{x}|w)\}.$$

Очевидно,

$$|X| \leq \frac{2^{K(w|\bar{x})+\mathcal{O}(\log N)} \cdot 2^{K(\hat{x}|w)}}{2^{K(w|\bar{x})-\delta-\mathcal{O}(\log N)}} = 2^{K(\bar{x}|w)+\delta+\mathcal{O}(\log N)}.$$

При этом алгоритм перечисления списка элементов X имеет сложность $\mathcal{O}(\log N)$. Следовательно, $K(x_I|w') \leq K(x_I|w) + \delta + \mathcal{O}(\log N)$. $\square$.

6 Выделение взаимной информации с релятивизацией и без неё.

Доказательство теоремы 6. Прежде, чем приступить к доказательству, примем несколько соглашений. Прежде всего, без ограничения общности

можно считать, что $f(N) > \log N$, причём $f(N)$ неубывает. Обозначим $m = K(w)$. Далее, введём обозначение $\delta(N) = N/\sqrt{\log \frac{N}{f(N)}}$. Мы будем писать для краткости просто δ, если значение N ясно из контекста. Наконец, мы будем считать, что

$$g(n) = C(3^{\sqrt{\log \frac{N}{f(N)}}} \cdot f(N) + \delta(N))$$

для достаточно большой константы $C > 0$. Данные функции подобраны так, чтобы выполнялись следующие условия:

1. $\delta(N) = o(N)$,
2. $3^{\sqrt{\frac{N}{f(N)}}} f(N) = o(N)$,
3. $\delta(N) \cdot \sqrt{\log \frac{N}{f(N)}} \geq N$,
4. $f(N) = o(\delta(N))$,
5. $g(N) = o(\delta(N))$.

В доказательстве мы будем пользоваться этими соотношениями. Разумеется, фунции $g(N)$ и $\delta(N)$, удовлетворяющие указанным соотношениям, можно было выбрать многими другими способами (это не повлияло бы на дальнейшее доказательство).

Рассмотрим пару $\langle y, w\rangle$ и её строгую типизацию относительно кортежа x: $A = ST(y, w|\bar{x})$. Согласно лемме 1 имеем $|A| \geq 2^{K(y,w|\bar{x}) - \mathcal{O}(f(N))} = 2^{K(w) - \mathcal{O}(f(N))}$. Отметим, что для любой пары $\langle y', w'\rangle \in A$

$$K(y', w') = m + \alpha + \mathcal{O}(f(N)).$$

Возможны два случая:

Случай 1^0. Для любого $\langle y', w'\rangle \in A$ для большинства $\langle y'', w''\rangle \in A$

$$I(y'w' : y''w'') \geq \alpha - \delta.$$

Это значит, что

$$K(y'w'|y''w'') \leq m + \delta + \mathcal{O}(f(N)).$$

В этом случае A является полугроздью с параметрами

$$(m - \mathcal{O}(f(N)), m + \delta + \mathcal{O}(f(N)), m + \alpha + \mathcal{O}(f(N)).$$

Следовательно, мы можем воспользваться леммой 7: существует стандартная полугроздь U_j с теми же параметрами, для которой

$$|A \cap U_j| \geq 2^{m-\delta+\mathcal{O}(f(N))},$$

причём если z – двоичная запись числа j, то

$$K(z) = \alpha + \delta + \mathcal{O}(f(N)).$$

Далее, для каждого $i = 1, \ldots, n$ для слова x_i выполнены следующие условия:

- для любой пары $\bar{v} \in A \cap U_j$ выполнено неравенство $K(x_i|\bar{v}) \leq K(x_i|y, w)$ (по определению множества $A = ST(y, w|\bar{x})$);
- для любой пары $\bar{v} \in A \cap U_j$ выполнено неравенство $K(\bar{v}|z) \leq \log |U_j| + \mathcal{O}(\log N) \leq m$ (поскольку список элементов стандартной грозди U_j можно перечислять, зная число j и параметры грозди).

Это значит, что слово x_i принадлежит множеству

$$X(i) = \{\hat{x} \quad | \quad \text{существует не менее } 2^{m-\delta+\mathcal{O}(f(N))} \text{ таких } \bar{v}, \text{ что } K(\hat{x}|\bar{v}) \leq K(x_i|y, w) \leq K(x_i) - \alpha + f(N) \text{ и } K(\bar{v}|z) \leq m - \mathcal{O}(f(N))\}.$$

Очевидно, множество $X(i)$ перечислимо при известном z программой сложности $\mathcal{O}(\log N)$, и $\log |X(i)| \leq K(x_i) - \alpha + \delta + \mathcal{O}(f(N))$. Следовательно,

$$K(x_i|z) \leq K(x_i) - \alpha + \delta + \mathcal{O}(f(N)).$$

Поскольку

$$K(z|x_i) = K(x_i|z) + K(z) - K(x_i) + \mathcal{O}(\log N),$$

мы получаем $K(z|x_i) \leq 2\delta + \mathcal{O}(f(N))$. Вспоминая определение функции $g(n)$, получаем $K(z|x_i) \leq g(N)$, и теорема доказана.

Случай 2^0. Для некоторого $\langle y', w' \rangle \in A$ и большинства $\langle y'', w'' \rangle \in A$

$$I(y'w' : y''w'') \leq \alpha - \delta.$$

Это значит, что

$$K(y''y''w'w'') \geq 2m + \alpha + \delta.$$

Поскольку данное свойство верно для большинства пар $\langle y'', w'' \rangle \in A$, мы можем выбрать такие $\langle y'', w'' \rangle$, что $\langle y', w' \rangle$ и $\langle y'', w'' \rangle$ независимы относительно $\bar{x}$. При этом w' и w'' будут также независимы относительно $\bar{x}$, а значит

$$I(w'w'' : \bar{x}) \leq I(w' : \bar{x}) + I(w'' : \bar{x}) + I(w' : w''|\bar{x}) + \mathcal{O}(\log N) \approx 0.$$

Более точно, будем полагать, что $I(w'w'' : \bar{x}) \leq 3f(n)$. Отметим, что мы огрубляем оценку – можно было бы потребовать, чтобы $I(w'w'' : \bar{x}) \leq 2f(N) + \mathcal{O}(\log N)$. Однако в данном случае мы не заботимся о максимальной точности оценки.

Кроме того, отметим, что $K(w'w'') \leq 2K(w) + 3f(N) \leq 3N$ (снова делаем очень грубую оценку).

Теперь мы можем заключить, что для слов $y^1 = \langle y', y'' \rangle$ и $w^1 = \langle w', w'' \rangle$ мы имеем

$$K(y^1|w^1) \geq \alpha + \delta - 3f(N) - \mathcal{O}(\log N) \geq \alpha + 1/2\delta.$$

Таким образом, для n-ки слов $\bar{x}$ существует слово w^1, такое что $I(w^1 : \bar{x}) \leq 3f(N)$, и

$$\exists y^1 : \ K(y^1|w^1) \geq \alpha + 1/2\delta, \ K(y^1|x_i, w^1) \leq 3f(N) (i = 1, \ldots, n).$$

Вместо пары $\langle y, w \rangle$ мы получили новую – $\langle y^1, w^1 \rangle$. При этом слово w^1 независимо с $\bar{x}$ (хотя и с большей погрешностью, чем w, а именно $I(w^1 : \bar{x}) \leq 3f(n)$). При релятивизации относительно w^1 слово y^1 попрежнему просто относительно каждого из x_i (также с погрешностью не более $3f(n)$). При этом сложность слова y^1 относительно w^1 не меньше $\alpha + 1/2\delta$. Таким образом, у слов $\bar{x}$ при релятивизации относительно w^1 можно выделить $\alpha + 1/2\delta$ битов общей информации при допустимом уровне погрешности $3f(n)$. Отметим, что сложности слов w^1, y^1 заведомо не превосходят $3N$.

Далее мы итерируем доказательство, повторяя с w^1, y^1 те же рассуждения, которые ранее проводились для пары слов w, y. Для удобства обозначений положим $\alpha^1 = \alpha + \delta/2$, $m_1 = K(w^1)$, и $f_1(N) = 3f(N)$.

Итак, рассмотрим типизацию пары $\langle y^1, w^1 \rangle$ относительно $\bar{x}$: $A^1 = ST(y^1, w^1|\bar{x})$. Снова возможны два варианта.

Случай 1^1. Для любого $\langle y', w' \rangle \in A^1$ для большинства $\langle y'', w'' \rangle \in A^1$

$$I(y'w' : y''w'') \geq \alpha_1 - \delta.$$

В этом случае A^1 является полугроздью с параметрами

$$(m^1 - \mathcal{O}(f_1(N)), m_1 + \delta + \mathcal{O}(f_1(N)), m_1 + \alpha_1 + \mathcal{O}(f_1(N)).$$

Рассуждаем как в случае 1^0. Применяя лемму 7, находим слово z такое, что

$$K(z) = \alpha_1 + \delta + \mathcal{O}(f_1(N)) > \alpha,$$

и для $i = 1, \ldots, n$

$$K(z|x_i) \leq 2\delta + \mathcal{O}(f_1(N)),$$

и теорема доказана.

Случай 2^1. Предположим, для некоторого $\langle y', w' \rangle \in A^1$ и большинства $\langle y'', w' \rangle \in A^1$

$$I(y'w' : y''w'') \leq \alpha_1 - \delta.$$

Рассуждаем как в случае 2^0: находим такую пару $\langle y^2, w^2 \rangle$, что

1. $K(w^2) = m_2 < 3m_1 < 9N$,
2. $I(w^2 : \bar{x}) \leq f_2(N) < 3f_1(N) < 9f(N)$,
3. $K(y^2|w^2, x_i) \leq f_2(N) < 9f(N)$,
4. $K(y^2|w^2) = \alpha_2 \geq \alpha_1 + \delta/2 \geq \alpha + \delta$.

Повторяя рассуждение снова и снова, на шаге j мы будем получать слова w^j, y^j, для которых

1. $K(w^j) = m_j < 3m_{j-1} < 3^j m < 3^j N$,
2. $I(w^j : \bar{x}) \leq f_j(N) < 3f_{j-1}(N) < 3^j f(N)$,
3. $K(y^j|w^j, x_i) \leq f_j(N) < 3^j f(N)$,
4. $K(y^j|w^j) = \alpha_j > \alpha_{j-1} + \delta/2 > \alpha + j\delta/2$.

Итерации рассуждений $2^1, 2^2, 2^3 \ldots, 2^j, \ldots$ будут повторяться, пока на некотором шаге j_{max} мы не получим случай $1^{j_{max}}$.

Заметим, что итерации не могут повторяться слишком долго. Если бы было сделано $j = D\sqrt{\log \frac{N}{f(N)}}$ шагов, мы получили бы противоречие с неравенством

$$K(y^j|w^j) \leq K(y^j|x_1, w^j) + K(y^j|x_2, w^j) + I(x_1 : x_2|w^j) + \mathcal{O}(\log N)$$

для достаточно большой константы D (нетрудно проверить, что данное неравенство выполнено для любых слов). В самом деле, в левой части неравенства мы бы получили величину не меньше DN, а в правой –

$$2f_j(N) + I(x_1 : x_2|w_j) + \mathcal{O}(\log N) \leq N + o(N).$$

Замечание: Во всех вычислениях мы игнорировали слагаемые порядка $\mathcal{O}(\log K(y^j, w^j))$, *полагая, что* $\log K(y^j, w^j) \ll f(N)$. *Это предположение законно, поскольку для* $j < \log N$ *выполняется* $K(y^j), K(w^j) < N^2$.

Таким образом, после нескольких итераций шагов 2^j для $j_{max} < D\sqrt{\log \frac{N}{f(N)}})$ мы придём к рассмотрению случая $1^{j_{max}}$. При этом мы получим некоторое слово z, для которого

$$K(z) \geq \alpha + j_{max}\delta/2 - \mathcal{O}(f_{j_{max}}(N)) > \alpha - g(N)$$

и

$$K(z|x_i) \leq 2\delta + f_{j_{max}} < 2\delta + 3\sqrt{\frac{N}{f(N)}} f(N) < g(N)(i = 1, \ldots, n).$$

Таким образом, у слов x_i выделяется не менее α битов общей информации с погрешностью не более $g(N)$. $\square$

7 Получение новых информационных неравенств

В этом разделе мы применим результаты из раздела 3 для доказательства нового класса линейных информационных неравенств. Мы получим неравенства, обобщающие результаты из [5, 4].

Прежде всего мы приведём интуитивно ясное доказательство частного утверждения, а затем получим общий результат.

Утверждение 3 *Пусть для слов* x_1, x_2, y, z_1, z_2 *выпонены следующие соотношения:*

$$I(x_1 : x_2|y),\ I(x_1 : y|x_2),\ I(x_2 : y|x_1),\ I(x_1 : x_2|z_1),\ I(x_1 : x_2|z_2) \leq C \log N$$

для некоторой константы C *(*$N = K(x_1, x_2, y, z_1, z_2)$*). Тогда существует* D, *зависящее только от* C, *такое что*

$$I(z_1 : z_2) \geq I(x_1, x_2 : y) - D \log N.$$

Доказательство. Для начала мы докажем, что данное утверждение выполнено в предположении гипотезы 3. Позднее мы получим доказательство, свободной от недоказанных предположений.

Прежде всего обозначим $a = I(x_1, x_2 : y)$. Поскольку $I(x_1 : y|x_2)$ и $I(x_2 : w|x_1)$ ограничены $C \log N$, имеем

$$I(x_i : w) = a + \mathcal{O}(\log N),\ i = 1, 2.$$

Мы предполагаем, что выполнена гипотеза 3. Применим её к словам x_1, x_2, y и получим такое w, что

1. $K(x_i|w) \leq K(x_i|y) + \mathcal{O}(\log N), i = 1, 2,$
2. $K(x_1, x_2) \leq K(x_1, x_2|y) + \mathcal{O}(\log N),$
3. $K(w) \leq I(x_1, x_2 : y) + \mathcal{O}(\log N).$

Прежде всего отметим, что из условия (2) следует $K(w) \geq I(x_1, x_2 : y) - \mathcal{O}(\log N) = a - \mathcal{O}(\log N)$, так что

$$K(w) = I(x_1, x_2 : y) + \mathcal{O}(\log N).$$

Далее,

$$K(w|x_i) = K(w) - I(x_i : w) = a - a + \mathcal{O}(\log N) = \mathcal{O}(\log N)\ i = 1, 2.$$

Таким образом, у слов x_1 и x_2 выделяется взаимная информация: слово w просто относительно каждого из x_i, и его сложность равна $I(x_1 : x_2)$.

По условию слова x_1, x_2 независимы относительно каждого из $z_j, j = 1, 2$. Отсюда следует, что слово w должно быть просто относительно каждого из z_j:

$$K(w|z_j) \leq K(w|x_1) + K(w|x_2) + I(x_1 : x_2|z_j) + \mathcal{O}(\log N) = \mathcal{O}(\log N)$$

(неравенство выполнено для любых слов, см. Приложение).

Но поскольку у пары слов z_1, z_2 можно выделить общую часть w, их взаимная информация не может быть меньше $K(w) \approx a$:

$$a - \mathcal{O}(\log N) \leq K(w) \leq K(w|z_1) + K(w|z_2) + I(z_1 : z_2) + \mathcal{O}(\log N),$$

то есть $I(z_1 : z_2) \geq a - \mathcal{O}(\log N)$, и утверждение доказано. □

Далее мы докажем общую теорему **??**. Отметим, что утверждение 3 является следствием теоремы **??**. Таким образом, доказав теорему, мы получим доказательство утверждения 3, не зависящее от гипотезы 3.

Нам потребуется следующая лемма:

Лемма 8 *Следующие условия на вещественные коэффициенты* $\{\lambda_W\}$ *(*$W \subseteq \{1,\dots,n\}$*) эквивалентны:*

1. *информационное неравенство с коэффициентами* λ_W *верно для колмогоровских сложностей* n*-ок слов, то есть*

$$\exists C > 0 \forall \bar{x} = (x_1,\dots,x_n) \sum_W \lambda_W K(\bar{x}_W) \geq -C \log K(\bar{x}).$$

2. *информационное неравенство с коэффициентами* λ_W *верно для колмогоровских сложностей стохастических* n*-ок слов, то есть* $\forall C > 0 \exists D > 0$ *такое, что для всякой* $C \log N$*-стохастической* n*-ки* $\bar{x} = (x_1,\dots,x_n)$ *сложности* N *выполнено*

$$\sum_W \lambda_W K(\bar{x}_W) \geq -D \log N.$$

По существу лемма 8 вытекает из доказательства эквивалентности неравенств для колмогоровской сложности и шенноновской энтропии в [11]. Более подробно доказательство леммы обсуждается в приложении.

Доказательство теоремы **??**. Согласно лемме 8 достаточно доказать неравенство для стохастических $\bar{x}$. Таким образом, мы можем воспользоваться теоремой 3 и считать, что для $\bar{x}, y$ верно заключение гипотезы 3. Это значит, что существует w такое, что

1. $K(\bar{x}_V|w) \leq K(\bar{x}_V|y) + \mathcal{O}(\log N)$ для любого набора индексов $V \subseteq \{1,\dots,n\}$,
2. $K(w) \leq I(\bar{x} : y) + \mathcal{O}(\log N)$.

Для данного w и для $i = 1,\dots,n$ имеем

$$K(w|x_i) \leq K(x_i|w) + K(w) - K(x_i) + \mathcal{O}(\log N) \leq K(x_i|y) + K(w) - K(x_i) + \mathcal{O}(\log N). \quad (4)$$

Далее воспользуемся неравенством, верным для любого набора из $(n+2)$ слов $x_1,\dots,x_n,w,z_j$

$$(n-1)K(w|z_j) \leq \sum_i K(w|x_i) + (\sum_i K(x_i|z_j) - K(\bar{x}) + \mathcal{O}(\log N).$$

Применяя (4), получаем

$$\begin{aligned} K(z_j|w) &= K(w|z_j) + K(z_j) - K(w) \leq \\ &\leq \tfrac{1}{n-1}[\textstyle\sum K(w|x_i) + \sum K(x_i|z_j) - K(\bar{x}|z_j] + K(z_j) - K(w) \\ &\leq \tfrac{1}{n-1}[\textstyle\sum (K(x_i|y) + K(w) - K(x_i)) + \sum K(x_i|z_j) - K(\bar{x}|z_j] + K(z_j) - K(w) \\ &\leq \tfrac{1}{n-1}[\textstyle\sum K(x_j|z_j) - K(\bar{x}|z_j) - \sum I(x_i : y)] + K(z_j) + \tfrac{1}{n-1}K(w) \end{aligned}$$

Следовательно,

$$\begin{aligned} K(\bar{z}) &\leq K(w) + \sum_j K(z_j|w) \leq \\ &\leq \tfrac{1}{n-1}[\sum_j \sum_i K(x_i|z_j) - K(\bar{x}|z_j)] - \tfrac{m}{n-1}\sum_i I(x_i : y) + \sum_j K(z_j) + \tfrac{m+n-1}{n-1} I(\bar{x} : y). \end{aligned}$$

Остается воспользоваться оценкой $K(w) \leq I(\bar{x} : y)$, и мы получаем требуемое неравенство.

8 Приложение

В этом разделе для полноты изложения мы приводим доказательства технических лемм, использовавшихся в основном тексте.

8.1 Технические утверждения о типизации

В этом разделе мы приведём доказательства технических лемм о свойствах сложностных профилей типичных последовательностей. Аналогичные свойства типичных последовательностей отмечались в работах [11, 9, 5].

Доказательство леммы 1. Прежде всего, для любого $\bar{x}' \in T(\bar{x}|\bar{y})$

$$K(\bar{x}'|\bar{y}) \leq K(\bar{x}|\bar{y}).$$

Следовательно, число таких $\bar{x}'$ не превосходит $2^{K(\bar{x}|\bar{y})+1}$. Далее, оценим размер $T(\bar{x}|\bar{y})$ снизу. Заметим, что зная $\bar{y}$ и все числа из профиля $\vec{K}'(\bar{x}|\bar{y})$, можно перечислять список всех элементов множества $T(\bar{x}|\bar{y})$ (разумеется, не имея большой дополнительной информации, нельзя определить, когда данный процесс перечисления закончится). Таким образом, чтобы получить $\bar{x}$ из $\bar{y}$ достаточно знать все компоненты расширенного сложностного профиля $\vec{K}(\bar{x}|\bar{y})$, а также номер кортежа $\bar{x}$ в указанном списке (в порядке перечисления). Следовательно,

$$K(\bar{x}|\bar{y}) \leq \log |T(\bar{x}, \bar{y})| + \mathcal{O}(\log N),$$

что и даёт требуемую оценку на размер $T(\bar{x}|\bar{y})$. $\square$

Доказательство леммы 2. Согласно лемме 1

$$|T(\bar{x}|\bar{y})| \geq 2^{K(\bar{x}|\bar{y})-C\log N}$$

для некоторой константы C. Следовательно, не меньше половины $\bar{x}' \in T(\bar{x}|\bar{y})$ имеют сложность относительно $\bar{y}$ большую или равную

$$K(\bar{x}|\bar{y}) - C\log N - 1.$$

Именно эти $\bar{x}' \in T(\bar{x}|\bar{y})$ мы и включим в $ST(\bar{x}|\bar{y})$.

Поскольку каждый элемент $\bar{x}' \in ST(\bar{x}|\bar{y})$ также принадлежит и $T(\bar{x}|\bar{y})$, мы имеем

$$\vec{K}'(\bar{x}', \bar{y}) \leq \vec{K}'(\bar{x}, \bar{y}).$$

Остаётся доказать, что с точностью до $\mathcal{O}(\log N)$ выполнено обратное неравенство. Иначе говоря, для любых $V_1, V_2 \subset \{1, \ldots, n\}$, $W_1, W_2 \subset \{1, \ldots, m\}$ мы должны показать, что

$$K(\bar{x}'_{V_1}, \bar{y}_{W_1}|\bar{x}'_{V_2}, \bar{y}_{W_2}) \geq K(\bar{x}_{V_1}, \bar{y}_{W_1}|\bar{x}_{V_2}, \bar{y}_{W_2}) - \mathcal{O}(\log N). \qquad (5)$$

Для этого заметим, что

$$K(\bar{x}', \bar{y}) = K(\bar{x}'_{V_2}, \bar{y}_{W_2}) + K(\bar{x}'_{V_1}, y_{W_1}|\bar{x}'_{V_2}, y_{W_2}) + K(\bar{x}', \bar{y}'|\bar{x}'_{V_1\cup V_2}, \bar{y}_{W_1\cup W_2}) + \mathcal{O}(\log N).$$

Правая часть неравенства не превосходит

$$K(\bar{x}_{V_2}, \bar{y}_{W_2}) + K(\bar{x}'_{V_1}, y_{W_1}|\bar{x}'_{V_2}, y_{W_2}) + K(\bar{x}, \bar{y}|\bar{x}_{V_1\cup V_2}, \bar{y}_{W_1\cup W_2}) + \mathcal{O}(\log N),$$

поскольку $\bar{x}' \in T(\bar{x}|\bar{y})$. Далее, учитывая равенство

$$\begin{aligned} K(\bar{x}', \bar{y}) + \mathcal{O}(\log N) &= K(\bar{x}, \bar{y}) = \\ &= K(\bar{x}_{V_2}, \bar{y}_{W_2}) + K(\bar{x}_{V_1}, y_{W_1}|\bar{x}_{V_2}, y_{W_2}) + \\ &\quad + K(\bar{x}, \bar{y}'|\bar{x}_{V_1\cup V_2}, \bar{y}_{W_1\cup W_2}) + \mathcal{O}(\log N), \end{aligned}$$

получаем (5). $\square$

Доказательство леммы 3. Доказательство аналогично доказательству предыдущей леммы. Сразу приступим к доказательству более общего утверждения пункта (2). Для любых V_1, V_2, W_1, W_2 имеем

$$K(\bar{x}'_{V_1}, \bar{y}_{W_1}|\bar{x}'_{V_2}, \bar{y}_{W_2}, z) \leq K(\bar{x}'_{V_1}, \bar{y}_{W_1}|\bar{x}'_{V_2}, \bar{y}_{W_2}) + \mathcal{O}(1) \leq K(\bar{x}_{V_1}, \bar{y}_{W_1}|\bar{x}_{V_2}, \bar{y}_{W_2}) + \delta.$$

С другой стороны, если сложность $K(\bar{x}'_{V_1}, \bar{y}_{W_1}|\bar{x}'_{V_2}, \bar{y}_{W_2}, z)$ меньше, чем $K(\bar{x}_{V_1}\bar{y}_{W_1}|\bar{x}_{V_2}, \bar{y}_{W_2}) - \delta - D\log N$, то (аналогично рассуждению в доказательстве леммы 2) получаем

$$K(\bar{x}', \bar{y}|z) < K(\bar{x}, \bar{y}) - \delta - D\log N + \mathcal{O}(\log N),$$

что даёт противоречие для достаточно большой константы D. $\square$

8.2 Доказательство неоторых информационных неравенств

В этом разделе мы доказываем несколько линейных неравенств для колмогоровской сложности. Все неравнства, доказываемые в данном разделе, принадлежат к классу неравенств шенноновского типа, то есть могут быть представлены в виде комбинациями *базисных* неравенств вида $I(\bar{x}_U : \bar{x}_V|\bar{x}_W) + \mathcal{O}(\log K(\bar{x})) \geq 0$.

А именно, мы докажем следующие неравенства.

1. $K(w) \leq K(w|a) + K(w|b) + I(a : b) + \mathcal{O}(\log K(a, b, w))$,
2. $K(w|z) \leq K(w|a) + K(w|b) + I(a : b|z) + \mathcal{O}(\log K(a, b, w, z))$,
3. $(n-1)K(w|z) \leq \sum_{i=1}^{n} K(w|x_i) + \sum_{i=1}^{n} K(x_i|z) - K(x_1, \ldots, x_n|z) + \mathcal{O}(\log K(x_1, \ldots, x_n, w, z))$.
4. для любых $\bar{x} = (x_1, \ldots, x_n), w_1, w_2$

$$\begin{aligned} \sum_{i=1} nK(x_i) + K(\bar{x}, w_1, w_2) \quad &+ \quad (n-1)K(w_1, w_2) \leq \\ &\leq \quad \sum_{i=1} n(K(x_i, w_1) + K(x_i, w_2)) + \mathcal{O}(\log K(\bar{x}, w_1, w_2)). \end{aligned}$$

Доказательство. Очевидно, первое неравенство вытекает из второго (при $z = \lambda$), а второе является частным случаем третьего (при $n = 2$). Доказажем третье неравенство. Перепишем его в виде

$$K(\bar{x}|z) + (n-1)K(w|z) \leq \sum K(w|x_i) + \sum K(x_i|z) + \mathcal{O}(\log N),$$

где $N = K(\bar{x}, w, z)$. Далее, правая часть только уменьшится, если добавить слово z в условие в величинах $K(w|x_i)$:

$$K(\bar{x}|z) + (n-1)K(w|z) \leq \sum K(w|x_i, z) + \sum K(x_i|z) + \mathcal{O}(\log N).$$

Вычтем $nK(w|z)$ из обеих частей равенства. Получим

$$K(\bar{x}|w, z) \leq \sum K(x_i|w, z) + \mathcal{O}(\log N).$$

Но данное неравенство уже совершенно очевидно: сложность кортежа $\bar{x}$ относительно $\langle w, z\rangle$ не превосходит суммы сложностей x_i относительно той же пары $\langle w, z\rangle$.

Перейдём к доказательству четвёртого неравенства. Будем обозначать $N = K(\bar{x}, w_1, w_2)$. Передставим требуемое неравенство в виде суммы неравенства

$$K(\bar{x}, w_1, w_2) + (n-1)K(w_1, w_2) \leq \sum K(x_i, w_1, w_2) + \mathcal{O}(\log N)$$

и неравенств

$$K(x_i) + K(x_i, w_1, w_2) \leq K(x_i, w_1) + K(x_i, w_2) + \mathcal{O}(\log N)$$

для $i = 1, \ldots, n$. Заметим, что первое из этих неравенство эквивалентно утверждению

$$K(\bar{x}|w_1, w_2) \leq \sum K(x_i|w_1, w_2) + \mathcal{O}(\log N),$$

а последующие выражают неотрицательность взаимной информации w_1 и w_2 относительно x_i:

$$K(w_1, w_2|x_i) \leq K(w_1|x_i) + K(w_2|x_i) + \mathcal{O}(\log N).$$

Таким образом, неравенство доказано. □

Список литературы

[1] M.Li and P.Vitányi, *An introduction to Kolmogorov complexity and its applications*. Second edition, Springer-Verlag, New York, 1997,

[2] P.Gács, J.Körner, *Common information is far less than mutual information*. Problems of Control and Information Theory, 2, 1973, p. 49-62.

[3] R.Ahlswede, J.Körner, *On common information and related characteristics of correlated information sources*. Presented at the 7th Prague Conf. on Inf. Th., Stat. Dec. Fct's and Rand. Proc., included in "Information Theory" by I.Csiszár and J.Körner, NY Acad. Press, 1981. See also
`http://www.mathematik.uni-bielefeld.de/ahlswede/pub/ahlswede/source.ps`

[4] Z.Zhang and R.W.Yeung, *On characterization of entropy functions via information inequalities*. IEEE Trans. on Information Theory, Vol. 44, pp.1440-1452, Jul 1998.

[5] K.Makarychev, Yu.Makarychev, A.Romashchenko, N.Vereshchagin. *A New class of non Shannon type inequalities for entropies.* Communications in Information and Systems, 2:2 (2002) 147-166.

[6] An.A.Muchnik, A.L. Semenov, *Multi-conditional Descriptions and Codes in Kolmogorov Complexity.* Electronic Colloquium on Computational Complexity (ECCC) 7(15) 2000.

[7] An.A.Muchnik, *On common information.* Theoretical Computer Science, 207, 1998, p.319–328.

[8] A.Chernov, An.A.Muchnik, A.Shen, A.Romashchenko, N.K.Vereshchagin, *Upper semi-lattice of binary strings with the relation "x is simple conditional to y".* Theoretical Computer Science. 271 (2002) pp. 69-95.

[9] A.Romashchenko, A.Shen, N.Vereshchagin, *Combinatorial Interpretation of Kolmogorov Complexity.* Theoretical Computer Science. 271 (2002) pp. 111-123.

[10] A.Romashchenko. *Pairs of Words with Nonmaterializable Mutual Information.* Problems of Information Transmission. 36 (2000) No. 1, pp. 1-18.

[11] D.Hammer, A.Romashchenko, A.Shen, N.Vereshchagin, *Inequalities for Shannon Entropy and Kolmogorov Complexity.* Journal of Computer and System Sciences. 60 (2000) pp. 442-464.

[12] A.Romashchenko. *Extracting the Mutual Information for a Triple of Binary Strings.* Proc. 18th Annual IEEE Conference on Computational Complexity (2003).

[13] A.Kh.Shen. *The concept of Kolmogorov* (α, β)*-stochastocity and its properties.* Soviet Math. Dokl. 28 (1983) pp. 295–299.