В знаменитой статье 1944 года Э. Пост ввел понятия простого и гиперпростого множества для построения перечислимых множеств, которые были бы $btt$-неполными или даже $tt$-неполными.

Оказывается, простые и гиперпростые множества естественно возникают при рассмотрении вычислимых верхних оценок на функцию колмогоровской (простой) и префиксной энтропии. Простая энтропия была введена А. Колмогоровым в 1965 году, префиксная энтропия была введена Л. Левиным в 1970 году.

Как известно, функция префиксной энтропии может рассматриваться как $-\log_2 \mu(x)$, где $\mu(x)$ — некоторый сходящийся ряд (так называемая априорная полумера). Это позволяет нам большую часть изложения построить на языке положительных рядов.

Основной вопрос, связанный с рядами, — это вопрос об их сходимости. Обозначим через $\log^{[i]} x$ $i$-ую итерацию двоичного логарифма. Классическими примерами сходящихся рядов являются $a^0_m = \frac{1}{m^2}$, $a^1_m = \frac{1}{m(\log m)^2}$, $a^2_m = \frac{1}{m \log m (\log^{[2]} m)^2}$ и так далее. Каждый следующий ряд из этой серии существенно больше предыдущего, то есть $\forall i$ ($a^{i+1}_m / a^i_m \to \infty$ при $m \to \infty$). Классическими примерами расходящихся рядов являются $b^0_m = \frac{1}{m}$, $b^1_m = \frac{1}{m \log m}$, $b^2_m = \frac{1}{m \log m \log^{[2]} m}$ и так далее. Каждый следующий ряд из этой серии существенно меньше предыдущего, то есть $\forall i$ ($b^{i+1}_m / b^i_m \to 0$ при $m \to \infty$). Ряды $a^i_m$ и $b^i_m$ очень близки, их отношение $\log^{[i]} m$ растёт, но очень медленно. Возникает естественный вопрос: нет ли наибольшего сходящегося или наименьшего расходящегося ряда? Ответ на этот вопрос отрицателен. Впрочем, нас будут интересовать вычислимые ряды. Без ограничения общности будем считать, что члены рядов принимают только значения вида $2^{-n}$. Каждое положительное число можно заменить на ближайшее к нему снизу число указанного вида. При этом оно уменьшится не более чем вдвое, и сохранится сходимость или расходимость ряда.

Каждый вычислимый расходящийся ряд можно существенно уменьшить, оставив его расходящимся.

**Теорема.** *По каждому вычислимому расходящемуся ряду $\alpha_m$ можно построить вычислимый расходящийся ряд $\beta_m$, такой что $\beta_m / \alpha_m \to 0$.*

*Доказательство.* Построим последовательность $\{m_i\}$, так что $\sum\limits_{m=m_i}^{m_{i+1}} \alpha_m > 2^i$, $m_1 = 1$. Положим $\beta_m = \alpha_m 2^{-i}$ для $m_i \leqslant m < m_{i+1}$. □

Каждый вычислимый *эффективно* сходящийся ряд можно существенно увеличить, оставив его эффективно сходящимся.

**Теорема.** *По каждому вычислимому эффективно сходящемуся ряду $\alpha_m$ можно построить вычислимый эффективно сходящийся ряд $\beta_m$, такой что $\beta_m/\alpha_m \to \infty$.*

*Доказательство.* Построим последовательность $\{m_i\}$, так что $\sum\limits_{m=m_i}^{\infty} \alpha_m < 2^{-2i}$. Положим $\beta_m = \alpha_m 2^i$ для $m_i \leqslant m < m_{i+1}$. □

Неожиданно оказывается, что существует вычислимый сходящийся ряд, который нельзя существенно увеличить, оставив его сходящимся.

**Теорема 1.** *Существует вычислимый сходящийся ряд $\alpha_m$, для которого нет вычислимого сходящегося ряда $\beta_m$, такого что $\beta_m/\alpha_m \to \infty$.*

Доказательство этой теоремы мы приведём ниже.

**Теорема.** *Не существует наибольшего с точностью до мультипликативной константы вычислимого сходящегося ряда.*

*Доказательство.* Пусть $\alpha_m$ — вычислимый сходящийся ряд. Построим последовательность $\{m_i\}$, так что $\alpha_{m_i} < 2^{-2i}$. Положим $\beta_{m_i} = \alpha_{m_i} 2^i$, и $\beta_m = \alpha_m$ для остальных $m$. □

Замечательным открытием в своё время явилось то, что в некотором естественном расширении класса вычислимых рядов, наоборот, существует наибольший с точностью до мультипликативной константы сходящийся ряд. Этим расширением является класс вычислимо аппроксимируемых снизу рядов. (В каждый момент времени нижняя аппроксимация ряда отлична от нуля на конечном множестве. Подчеркнём, что не предполагается *равномерность* аппроксимации.) Для удобства обозначений в дальнейшем мы будем допускать члены ряда, равные нулю; номер члена ряда будем писать не в индексе, а в скобках.

**Теорема 2 (Левин).** *Существует вычислимо аппроксимируемый снизу ряд $\mu(x)$, сумма которого $\leqslant 1$, и по каждому вычислимо аппроксимируемому снизу ряду $\nu(x)$, сумма которого $\leqslant 1$, можно указать константу $C$, для которой $\forall x\ \mu(x) > \nu(x)2^{-C}$.*

*Доказательство.* По аналогии с перечислимыми множествами существует такая универсальная вычислимая функция $U\colon n \mapsto \nu_n$, нумерующая все вычислимо аппроксимируемые снизу ряды, что каждая другая вычислимая функция $V$, нумерующая вычислимо аппроксимируемые снизу ряды, $\mathsf{m}$-сводится к $U$.

Каждый вычислимо аппроксимируемый снизу ряд $\nu$ можно эффективно заменить на вычислимо аппроксимируемый снизу ряд $\nu'$ так, что $\sum \nu'(x) \leqslant 1$ и, если $\sum \nu(x) \leqslant 1$, то $\forall x\ \nu'(x) = \nu(x)$. Программа, аппроксимирующая снизу $\nu'$, действует так же, как программа, аппроксимирующая снизу $\nu$, пока сумма значений текущей аппроксимации $\nu$ не превышает 1; если же такое превышение случится, то программа для $\nu'$ останавливается.

Искомый ряд $\mu$ можно определить как $\sum_n 2^{-n}\nu'_n$. □

Очевидно, что ряд $\mu$, существование которого мы доказали, единствен с точностью до мультипликативной константы.

Левин доказал, что ряд $\mu$ невычислим. Займёмся вопросом о его вычислимых верхних и нижних оценках.

Ряд $\mu$ не имеет нетривиальных частично вычислимых верхних оценок (аналог теоремы Маранджана для простой энтропии).

**Теорема 3.** *Для каждой частично вычислимой функции $\gamma$ можно указать такую константу $C$, что*

$$\forall x \in Dom(\gamma)\ \ \mu(x) \leqslant \gamma(x) \quad \Rightarrow \quad \forall x \in Dom(\gamma)\ \ \gamma(x) \geqslant 2^{-C}\,.$$

*Доказательство.* Рассмотрим следующий вычислимо аппроксимируемый снизу ряд $\delta = \lim \delta_n$. Аппроксимация $\delta_0$ тождественно равна нулю. Аппроксимация $\delta_{n+1}$ совпадает с аппроксимацией $\delta_n$ всюду, кроме, возможно, одного аргумента $x$, который ищется так. Параллельно вычисляя $\gamma$ на всех аргументах, ищем такое $x$, что $\gamma(x) < 2^{-2n}$. На первом найденном $x$ (если оно вообще нашлось) положим $\delta_{n+1}(x) = \max\{\delta_n(x), 2^{-n-1}\}$. Понятно, что $\sum_x \delta(x) \leqslant 1$. Следовательно, можно указать (эффективно по номеру $\gamma$) такое $c$, что $\forall x\ \mu(x) > \delta(x)2^{-c}$. Соединяя нижнюю и верхнюю оценки на $\mu(x)$, получаем, что для $n \geqslant c$ число $x$ в определении $\delta_n$ найдено не будет, и $\forall x\ \gamma(x) \geqslant 2^{-2c}$. □

Возникает предположение, что у $\mu$ нет и хороших вычислимых нижних оценок, то есть отношение $\mu$ к любой вычислимой нижней оценке $\mu$ стремится к бесконечности. Следующая теорема опровергает это предположение и заодно доказывает теорему 1.

**Теорема 4.** *Существует вычислимый ряд $\alpha$, являющийся нижней оценкой, и на бесконечном множестве равный $\mu$.*

*Доказательство.* Построим вспомогательный вычислимый ряд $\beta$.

| | | | | | |
|---|---|---|---|---|---|
| 1 | | | | | |
| 2 | | | | | |
| $\vdots$ | | | | | |
| $i$ | $\beta(x_i^1) = 2^{-i}$<br>$\mu(x_i^1) > 2^{-i}$ | $\beta(x_i^2) = 2^{-i}$<br>$\mu(x_i^2) > 2^{-i}$ | $\dots$ | $\beta(x_i^j) = 2^{-i}$ | |
| $\vdots$ | | | | | |

Ряд $\beta$ получается в результате вычислимого процесса заполнения указанной таблицы. К каждой строке мы возвращаемся бесконечно много раз, и параллельно аппроксимируем снизу ряд $\mu$. При первом обращении к $i$-ой строке мы берём первый ещё не определённый член ряда $\beta$ (обозначим его номер через $x_i^1$), и полагаем $\beta(x_i^1) = 2^{-i}$. При следующем обращении к $i$-ой строке мы сравниваем текущую аппроксимацию $\mu(x_i^1)$ с числом $2^{-i}$, и если $\mu(x_i^1) > 2^{-i}$, то снова берём первый ещё не определённый член ряда (обозначим его номер через $x_i^2$), и полагаем $\beta(x_i^2) = 2^{-i}$, и так далее. Ясно, что каждый номер $x$ ровно один раз встречается в таблице.

Поскольку $\sum \mu(x) \leqslant 1$, то длина $i$-ой строки заведомо меньше $2^i$. Рассмотрим множество $D$ номеров $x_i^j$, для которых $\mu(x_i^j) \leqslant 2^{-i}$. Сумма ряда $\beta(x)$ по этому множеству равна $\sum_i 2^{-i} = 1$. Сумма ряда $\beta(x)$ по дополнению до $D$ меньше $\sum \mu(x) \leqslant 1$. Итак, ряд $\beta(x)$ сходится и $\exists c \forall x\, \mu(x) > \beta(x)2^{-c}$.

С другой стороны, $\mu(x) \leqslant \beta(x)$ на бесконечном множестве $D$. Для каждого $C$ рассмотрим множество $M_C = \{x : \mu(x) \leqslant \beta(x)2^{-C}\}$. Для $C = 0$ это множество бесконечно, для $C = c$ это множество пусто. С ростом $C$ множество $M_C$ уменьшается. Рассмотрим наибольшее $d$, для которого $M_d$ бесконечно. Так как $\mu(x) > \beta(x)2^{-(d+1)} \Leftrightarrow \mu(x) \geqslant \beta(x)2^{-d}$, то на дополнении до конечного множества $M_{d+1}$ будет $\mu(x) \geqslant \beta(x)2^{-d}$ и на бесконечном множестве $M_d \setminus M_{d+1}$ будет $\mu(x) = \beta(x)2^{-d}$.

Искомый ряд $\alpha$ определяется формулой $\alpha(x) = \min\{\mu(x), \beta(x)2^{-d}\}$. □

**Теорема 5.** *Для произвольного вычислимого ряда $\alpha$, оценивающего снизу $\mu$ и на бесконечном множестве равного $\mu$, множество $\{x : \mu(x) > \alpha(x)\}$ гиперпростое.*

(Напомним, что перечислимое множество с бесконечным дополнением называется гиперпростым, если в любой вычислимой бесконечной последовательности непересекающихся отрезков натуральных чисел существует отрезок, целиком вложенный в это множество.)

*Доказательство.* Пусть $T_j$ — вычислимая последовательность непересекающихся отрезков натуральных чисел.

Построим вспомогательный вычислимый ряд $\beta$. Для каждого $m$ найдём $j_m$, для которого $\sum_{x\in T_{j_m}} \alpha(x) < 2^{-2m}$. Это возможно, так как ряд $\mu(x)$ сходится, $\alpha(x) \leqslant \mu(x)$ и, следовательно, вычислимый ряд $\alpha(x)$ сходится. На отрезках $T_{j_m}$ положим $\beta(x) = \alpha(x)2^m$, в остальных местах $\beta(x) = \alpha(x)$. Поскольку

$$\sum_{x\in\bigcup\limits_m T_{j_m}} \beta(x) < \sum_m 2^m \cdot 2^{-2m} = 1\,,$$

то ряд $\beta(x)$ сходится. Поэтому $\exists c \forall x\, \mu(x) > \beta(x)2^{-c}$. Из определения $\beta$ получается, что на отрезке $T_{j_c}$ не может быть $x$, в котором $\alpha(x) = \mu(x)$. □

Перейдём к языку сложности описаний.

Простая энтропия (введённая Колмогоровым в 1965 году) — это минимальная длина кода относительно оптимального кодирования.

Кодирование называется префиксным, если ни один код не может быть продолжением другого кода. Требование префиксности позволяет делить последовательность закодированных сообщений на отдельные сообщения без использования вспомогательного символа (например, пробела). Префиксная энтропия (введённая Левиным в 1970 году) — это минимальная длина кода относительно оптимального префиксного кодирования.

Бескодовые определения.

Простая энтропия — это минимальная с точностью до аддитивной константы перечислимая сверху функция $KS$ (со значениями в $\mathbb{N}$), для которой при любом $n$ выполнено $|\{x : KS(x) < n\}| < 2^n$.

Префиксная энтропия — это минимальная с точностью до аддитивной константы перечислимая сверху функция KP (со значениями в $\mathbb{N}$), для которой выполнено $\sum_x 2^{-KP(x)} \leqslant 1$.

Из определений очевидно, что функции $KS$ и $KP$ стремятся к бесконечности, и $\forall x\, KS(x) \leqslant KP(x)+O(1)$. Значительно сложнее доказывается, что разность $(KP-KS)$ стремится к бесконечности. (Четыре месяца назад доказательство было рассказано нами на международном семинаре в Гейдельберге. Ли и Витаньи в своей книге приводят этот факт без доказательства со ссылкой на неопубликованную рукопись Соловэя.)

Функции $KS$ и $KP$ имеют редкие, но непредсказуемые падения.

*Здесь рисунок из файла graph.eps*

Г. Маранджан в 1969 году доказал, что функция $KS$ не имеет нетривиальных частично вычислимых нижних оценок (для функции $KP$ рассуждение аналогично).

Что касается вычислимых верхних оценок простой энтропии, то такой оценкой является, например, $\ell(x)+C$, где $\ell(x)$ — длина двоичной записи натурального числа $x$.

Из мощностных соображений следует, что

$$\forall n \exists x\, \ell(x) = n \,\&\, KS(x) \geqslant n - 1\,.$$

С другой стороны, по определению $KS$ имеем $\exists c \forall x\ KS(x) < \ell(x) + c$. Для каждого $C$ рассмотрим множество $M_C = \{x : \ell(x) + C \leqslant KS(x)\}$. Для $C = -1$ это множество бесконечно, для $C = c$ это множество пусто. С ростом $C$ множество $M_C$ уменьшается. Рассмотрим наибольшее $d$, для которого $M_d$ бесконечно. Ясно, что на дополнении до конечного множества $M_{d+1}$ будет $\ell(x) + d \geqslant KS(x)$ и на бесконечном множестве $M_d \setminus M_{d+1}$ будет $\ell(x) + d = KS(x)$. Пусть функция $f$ равна $KS$ на конечном множестве $M_{d+1}$ и равна $\ell(x) + d$ в остальных местах. Понятно, что $f$ — вычислимая функция, оценивающая сверху $KS$ и на бесконечном множестве равная $KS$.

**Теорема 6.** *Для* произвольной *вычислимой функции $f$, оценивающей сверху $KS$ и на бесконечном множестве равной $KS$, множество $\{x : KS(x) < f(x)\}$ простое.*

(Напомним, что перечислимое множество с бесконечным дополнением называется простым, если в его дополнение нельзя вложить бесконечное перечислимое подмножество.)

*Доказательство.* Предположим, что существует бесконечное перечислимое множество $R$, на котором $f(x) = KS(x)$. Пусть $f'$ — ограничение $f$ на $R$. Тогда $f'$ — частично вычислимая *нижняя* оценка на $KS$. Последнее невозможно, поскольку $KS(x) \to \infty$. □

Теоремы 4 и 5 на языке префиксной энтропии приведены без доказательства в книге Ли и Витаньи со ссылкой на неопубликованную рукопись Соловэя. Сформулируем их.

**Теорема 7.** *Существует всюду определённая вычислимая функция $f$, оценивающая сверху $KP$ и на бесконечном множестве равная $KP$.*

**Теорема 8.** *Для произвольной всюду определённой вычислимой функции $f$, оценивающей сверху $KP$ и на бесконечном множестве равной $KP$, множество $\{x : KP(x) < f(x)\}$ гиперпростое.*