А.Н. КОЛМОГОРОВ

совместно с Л.А.Левиным и Н.В.Петри будет вести семинар

"СЛОЖНОСТЬ АЛГОРИТМОВ И АЛГОРИТМИЧЕСКИЙ ПОДХОД К ТЕОРИИ ИНФОРМАЦИИ И ТЕОРИИ ВЕРОЯТНОСТЕЙ"

Семинар посвящен новому подходу к понятиям информации и случайности, который излагался в ряде последних выступлений А.Н.Колмогорова. Большое внимание будет уделяться понятиям сложности алгоритмов, времени их работы и т.п.

Семинар будет понятен студентам, которые имеют (или сумеют быстро получить) начальные представления теории алгоритмов, теории информации и теории вероятностей.

Те, кто желает быстрее войти в курс дела могут обратиться к литературе:

I) А.Н.Колмогоров "Три подхода к понятию количества информации " журнал "Проблемы передачи информации" I965, № I, стр. 3-7.

2) Обзор А.К.Звонкина и Л.А.Левина "Алгоритмические сложности и обоснование понятий информации и случайности с помощью теории алгоритмов". Успехи математических наук, I970, № 6, стр. 85-I27.

Собираться семинар будет по четвергам с 16[0] до 18[0] в ауд. 12-13 .

На первом занятии 7 октября предполагается доклад А.Н.Колмогорова.

Вот результат, который я объявил на семинаре.

Будем называть простой функцией такую, которая может быть вычислен в зоне, не превышающей её результата (самого результата, а не его длины).

Очевидно, что такие и только такие функции могут задавать зону работы какого-нибудь алгоритма (машины Тюринга с любым алфавитом)

Будем говорить, что $g(x)$ является точной границей зоны вычисления функции $F(x)$ если для всякой простой функции $f(x)$ верно:

$\exists c \quad c \cdot f(x) \geq g(x) \iff$ ($F(x)$ вычислима в зоне $f(x)$)

Теорема

а) Для любой вычислимой (не обязательно простой) функции g найдется F , для которой g будет точной границей.

б) Для любой вычислимой F найдется g , являющаяся её точной границей.

Дорогой Андрей Николаевич! Я на днях получил результат, который мне очень нравится. Может быть он пригодится и Вам, если Вы на корабле будете заниматься этими вещами.

Результат состоит в том, что я придумал иную, чем у Мартин-Лёфа формулировку обоснования теории вероятностей. Она, как мне кажется, ближе к Вашей первоначальной идее о связи сложности и случайности и (как показалось, например, Медведеву) намного прозрачнее философски.

Мартин-Лёф рассматривал для произвольной вычислимой меры P алгоритм, который изучает последовательность и находит в ней всё больше и больше отклонений от гипотезы о том, что она P-случайна. Такой алгоритм должен быть P-корректен, т.е. констатировать отклонение величины m лишь на множестве меры $\leq 2^{-m}$. Очевидно, что результаты m такого алгоритма на слове x находятся в пределах $0 \leq m \leq -\log_2 P(x)$. Будем рассматривать дополнительную величину $(-\log_2 P(x)) - m$ и называть её доп.тестом (легко переформулировать для неё условия P-корректности). [redacted]

Теорема Логорифм априорной вероятности $(-\log_2 R(x))$ является P-корректным доп.тестом при любой мере P и алгоритмические свойства его обычны.
Можно показать, что $(-\log_2 R(x))$ является оптимальным (с точностью до адд.константы) среди всех других доп. тестов, обладающих свойствами, указанными в теореме.

Напомню Вам, что априорной вероятностью я называю универсальную полувычислимую меру введенную в нашей со Звонкиным статье. Она, как там доказывается, численно близка к сложности.

Если мы рассмотрим конкретную вычислимую меру P, то по сравнению с универсальным тестом f Мартин-Лёфа, корректным

только относительно одной меры P, наш тест не будет оптимальным с точностью до аддитивной константы, но будет оптимален ассимптотически, т.е. если универсальный Мартин-Лёфовский тест найдет отклонение величины m, то наш тест найдет отклонение величины не меньше чем $m - 2\log_2 m - c$. Таким образом класс бесконечных двоичных случайных последовательностей не изменится.

Теперь посмотрите, как красиво выглядит философия. Мы говорим, что гипотеза о возникновении слова x случайно по мере P опровергается с уверенностью m, если мера P согласуется с появлением x намного хуже, чем априорная вероятность (точнее просто если $P(x) < \frac{R(x)}{2^m}$). Это закон теории вероятностей, т.к. он нарушается только с вероятностью $\leqslant 2^{-m}$ Его нарушение можно установить эффективно, т.к. R — полувычислима. Но уже если он выполняется, то тогда выполняются и все другие законы теории вероятностей (Мартин-Лёфовские тесты). Правда значение нарушений этот закон может дать немного меньше (вместо m только $m - 2\log_2 m - c$), но это плата за универсальность (в смысле произвольности распределения вероятностей).

P.S. Связь со сложностью обусловлена тем, что $-\log_2 R(x)$ это почти сложность x. Теперь эта связь не зависит от меры.

Интересно заметить, что универсальная полувычислимая мера имеет интересные приложения во многих вопросах. Кроме вышеприведенного, Вы знаете её приложения к изучению вероятностных алгоритмов. Ещё её часто удобно применять в доказательствах (например, в доказательстве гипотезы Дж.Т.Шварца о сложности п.в. траекторий произвольных динамических систем). Когда-то я применил эту меру к построению одного определения интуиционистской истенности. Всё это показывает, что она является довольно естественной величиной.

Л. Левин

Дорогой Андрей Николаевич !

Я хочу показать Вам, что обычная сложность не годится для точного определения случайности даже в конечном случае. В случае равномерного распределения вероятностей на словах одинаковой длины дефект сложностей определяется как длина минус сложность. В случае неравномерного распределения дефект сложности определяется как логарифм вероятности минус сложность.
Оказывается, что даже для распределения вероятностей на конечном числе слов дефект сложности может быть большим на множестве большой меры.

ПРИМЕР. Пусть $P(x) = \begin{cases} 2^{-(l(x)+100)} & \text{при } l(x) \le 2^{100} \\ 0 & \text{при } l(x) > 2^{100} \end{cases}$

Тогда $|\log_2 P(x)| - K(x)$ будет больше 100 на всех словах

Аналогичный пример можно построить и на словах одинаковой длины (добавив спереди нулей). Такие нарушения могут иметь порядок логарифма сложности.
Я покажу Вам как уточнить определение сложности, чтобы всё получалось точно (и для конечных и для бесконечных последовательностей).

ОПРЕДЕЛЕНИЯ

Пусть A монотонный алгоритм (т.е. $\forall x, y \subset x$ если $A(x)$ – определено, то $A(y)$ – определено и $A(y) \subset A(x)$)

$$Km_A(x) \stackrel{def}{=} \begin{cases} \min l(p) : A(p) \supset x \\ \infty, \text{ если такого } p \text{ нет} \end{cases}$$

сложность по оптимальному алгоритму обозначим просто $Km(x)$
Пусть $P(x)$ – вычислимое распределение вероятностей на Ω
($P(x)$ – вероятность Γ_x)

Теорема I $Km(x) \lesssim |\log_2 P(x)|$

Теорема 2.

Для P-почти всех ω $Km(\omega_n) \gtrsim |\log_2(\omega_n)|$

Причем вероятность того что дефект сложности хотя бы на одном фрагменте превзойдет m меньше или равна 2^{-m}

Теорема 3.

Те ω для которых $Km(\omega_n) \gtrsim |\log_2 P(\omega_n)|$ выдерживают любой закон теории вероятностей (Мартин-Лёфовский тест)

Изложу заодно результаты, которые я докладывал в лаборатории о том, почему можно не рассматривать невычислимых (не задаваемых в сильном языке) тестов.

Для этого нужно еще раз усовершенствовать определение сложности. Обычная сложность $K(x)$ обладает следующим важным свойством.

Замечание. Пусть $\mathcal{A}_i$ - эффективная последовательность алгоритмов таких , что

$$\forall i, x \qquad K_{\mathcal{A}_{i+1}}(x) \leq K_{\mathcal{A}_i}(x)$$

Тогда $\exists$ алгоритм $\mathcal{A}_0$ такой, что

$$K_{\mathcal{A}_0}(x) = 1 + \min_i K_{\mathcal{A}_i}(x)$$

К сожалению $Km(x)$ этим свойством, по-видимому, не обладает. Положение легко исправить . Пусть $\mathcal{A}_i$ - эффективная последовательность монотонных алгоритмов с конечной областью определения (заданных в виде таблицы) такая, что

$$\forall i, x \quad Km_{\mathcal{A}_{i+1}}(x) \leq Km_{\mathcal{A}_i}(x)$$

Тогда определим $\overline{Km}_{\{\mathcal{A}_i\}}(x) \stackrel{def}{=} \min_i Km_{\mathcal{A}_i}(x)$

Среди таких последовательностей существует оптимальная; сложность, по ней мы обозначим через $\overline{Km}(x)$. Она будет совпадать с логарифмом универсальной полувычислимой меры.

Теорема 4. $\overline{Km}(x)$ -минимальная полувычислимая величина, для которой выполняется теорема 2.

Таким образом дальше усовершенствовать $\overline{Km}$ невозможно.

Рассмотрим язык функций вычислимых относительно фиксированной невычислимой последовательности α . Пусть α достаточно сложна, чтобы этот язык содержал характеристическую функцию универсального перечислимого множества.

Аналогично алгоритмической сложности $\overline{Km}(x)$ легко определить "языковую сложность" $\overline{Km}_{\sim}(x)$, заменив понятие "алгоритма" на

понятие "функция из языка".

Определение

Назовем последовательность (w_n) нормальной, если

$$Km(w_n) \gtrsim Km_{\mathcal{L}}(w_n)$$

(конечно это это определение зависит от $\mathcal{L}$, задающей язык). Конечная последовательность может иметь "дефект нормальности":

$$Km(w_n) - Km_{\mathcal{L}}(w_n)$$

Теорема 5. Последовательность, которая получается алгоритмом из нормальной, сама тоже будет нормальной.

Теорема 6. Если P -распределение вероятностей, задаваемое (в естественной кодировке) нормальной последовательностью, то P-почти любая последовательность будет нормальной.

Эта теорема выявляет закон теории вероятностей о том, что в случайном процессе не может возникнуть ненормальной последовательности (кроме случая, когда само распределение вероятностей уже было ненормальным). Этот закон намного более общий, чем обычные законы теории вероятностей, так как не зависит от распределения. Кроме того теорема 5 показывает, что это закон не только теории вероятностей. Можно этот закон возвести в ранг универсального закона природы:

Тезис. Любая встречающаяся в природе последовательность (конечная или бесконечная) имеет дефект нормальности не больше сложности описания физического способа ее получения (на русском языке) - или местонахождения и т.п.

Оказывается, что из закона нормальности (который можно считать выходящим за рамки теории вероятностей) и закона "выдерживания универсального вычислимого теста" вытекает любой другой закон теории вероятностей (не обязательно вычислимый, но задаваемый в языке). А именно:

Теорема 7. Пусть P -вычислимое распределение вероятностей. Если последовательность нормальна и выдерживает универсальный вычислимый P— тест, то она выдерживает любой закон теории вероятностей, задаваемый в нашем языке (т.е. тест* вычислимый относительно α).

Приведем еще одну любопытную теорему. Она показывает, что все нормальные последовательности устроены сходным образом.

Теорема 8. Всякая нормальная последовательность может быть получена алгоритмом из последовательности, случайной по равномерной мере.

---

* Заметим, что для любого множества меры 0 найдется тест (необязательно вычислимый), который отбрасывает все его элементы.

Нам понадобится последовательность d, задающая „плотную" кодировку универсального перечислимого множества. В качестве такой d выберем, например, двоичную запись числа $\sum_{p \in A} \frac{1}{p \cdot \log^2 p}$, где A – это область определения оптимального алгоритма.

Будем говорить, что двоичное слово p является „хорошим" кодом для X, если оптимальный алгоритм (пару $(p, K(x))$) перерабатывает p в список слов, содержащий X, причём логарифм мощности списка не превосходит $K(x) + 2\log K(x) - \ell(p)$. (Существование такого кода означает „случайность" X при $n \geq \ell(p)$)

Будем говорить, что двоичное слово p является каноническим кодом X, если любой начальный фрагмент p либо является „хорошим" кодом X, либо совпадает с соответствующим (начальным) фрагментом d, причём $\ell(p) = K(x) + 2\log K(x)$

Теорема 1. Для любого X, кроме конечного числа существует канонический код p, причем p и X восстанавливаются друг по другу при известном $K(x)$.

Таким образом, „Неслучайность" X может возникать только за счёт очень специальной информации (начального фрагмента d), содержащейся в X. Я не могу себе представить, как в природе можно получить и выделить такое X (вне d – невообразимо). И к тому же задача изучения свойств начальных фрагментов конкретной последовательности d является очень специальной.

2008 Note: struck out text reads:
"суммарной априорной вероятности всех натуральных чисел!"