УДК 517.11

А.Х. ШЕНЬ

АЛГОРИТМИЧЕСКИЕ ВАРИАНТЫ ПОНЯТИЯ ЭНТРОПИИ

(Представлено академиком А.Н. Колмогоровым 1 VI 1983)

В настоящей работе предлагается общая схема получения различных алгоритмических вариантов понятия энтропии. Эта схема использует понятие f_0-пространства в смысле Ю.Л. Ершова [1] и интерпретацию логических операций как операций над задачами в смысле А.Н. Колмогорова [2]. Частными случаями этой схемы оказываются простая и условная колмогоровские энтропии [3, 4], энтропия разрешения, монотонная и префиксная энтропии [4–6], а также энтропия вычислимых функций, равная логарифму минимального номера при оптимальной в смысле К. Шнорра [4], с. 151, нумерации. С точки зрения этой схемы рассматривается также понятие априорной вероятности [5, 6].

1. Понятие f_0-пространства. Это понятие введено Ю.Л. Ершовым. Приведем его определение в удобной для нас форме. Тройка $\langle X, X_0, \leqslant\rangle$, где $\langle X, \leqslant\rangle$ – упорядоченное множество, а $X_0 \subset X$, называется f_0-пространством, если: 1) X содержит наименьший элемент $\perp$, который принадлежит X_0; 2) любые два элемента X_0, имеющие общую мажоранту в X, имеют в X точную верхнюю грань, которая принадлежит X_0; 3) если $x, y \in X$ и $x \nleqslant y$, то существует $x_0 \in X_0$, для которого $x_0 \leqslant x$ и $x_0 \nleqslant y$. Объектами f_0-пространства $\langle X, X_0, \leqslant\rangle$ будем называть элементы X. Объект $\perp$ будем называть неопределенностью, элементы X_0 – конечными объектами, или f-объектами. Объекты x и y, имеющие общую мажоранту, будем называть совместными.

Назовем множество $I \subset X_0$ идеалом, если оно непусто, вместе с каждым f-объектом z содержит и все f-объекты, меньшие z, и любые два объекта, $y,z \in I$ совместны, причем $\sup(y, z) \in I$. Назовем f_0-пространство полным, если всякий идеал I равен множеству $I_x = \{x_0 \in X_0 \mid x_0 \leqslant x\}$ для некоторого объекта x. В дальнейшем мы рассматриваем только полные f_0-пространства.

Опишем несколько операций над f_0-пространствами. Произведением f_0-пространств $\langle X, X_0, \leqslant_1\rangle$ и $\langle Y, Y_0, \leqslant_2\rangle$ называется пространство $\langle X \times Y, X_0 \times Y_0, \leqslant_1 \times \leqslant_2\rangle$ (произведение порядков определяется покомпонентно). Суммой f_0-пространств $\langle X, X_0, \leqslant_1\rangle$ и $\langle Y, Y_0, \leqslant_2\rangle$ при $X \cap Y = \phi$ мы считаем пространство $\langle X \cup Y \cup \{\perp\}, X_0 \cup Y_0 \cup \{\perp\}, \leqslant\rangle$, где $\perp$ – элемент, не входящий ни в X, ни в Y, а порядок определяется так: $\perp \leqslant x$ и $\perp \leqslant y$ для всех $x \in X$, $y \in Y$, порядок внутри X и Y сохраняется и элементы X не сравнимы с элементами Y. Пространство непрерывных функций из $\langle X, X_0, \leqslant\rangle$ в $\langle Y, Y_0, \leqslant\rangle$ состоит из всюду определенных функций из X в Y, непрерывных относительно естественной топологии f_0-пространств, в которой базовыми открытыми множествами считаются множества, состоящие из всех объектов, больших данного f-объекта. Порядок на функциях поточечный: $f \leqslant g \iff (\forall x \in X)\,(f(x) \leqslant g(x))$. Конечными объектами в пространстве функций являются функции вида

$$f_{x_0, y_0}(x) = \text{if } x_0 \leqslant x \text{ then } y_0 \text{ else } \perp$$

для всех f-объектов $x_0 \in X_0$, $y_0 \in Y_0$, а также точные верхние грани совместных конечных семейств таких функций. Описанные операции, примененные к полным пространствам, дают полные пространства.

Пусть $\langle X, X_0, \leqslant\rangle$ – f_0-пространство, ν – натуральная нумерация множества X_0, множества $\{\langle m, n\rangle \mid \nu(m) \leqslant \nu(n)\}$ и $\{\langle m, n\rangle \mid \nu(m)$ совместно с $\nu(n)\}$ разрешимы и существует вычислимая функция $f: \mathbf{N}^2 \to \mathbf{N}$, для которой $\nu(f(m, n)) = \sup(\nu(m),$

$\nu(n)$), если $\nu(m)$ и $\nu(n)$ совместны. В этом случае четверку $\langle X, X_0, \leqslant, \nu\rangle$ назовем эффективным f_0-пространством. Если X и Y – эффективные f_0-пространства, то на их произведении $X \times Y$, сумме $X + Y$ и на пространстве непрерывных функций $C(X, Y)$ естественно вводятся структуры эффективных f_0-пространств.

Приведем несколько примеров f_0-пространств, используемых в дальнейшем. Через $N_\perp$ обозначим пространство, объектами которого являются натуральные числа и символ $\perp$, все объекты конечны, объект $\perp$ меньше всех других, а натуральные числа не сравнимы друг с другом. Через Ω обозначим пространство, элементами которого являются конечные и бесконечные последовательности цифр 0 и 1, f-объектами являются конечные последовательности, а $x \leqslant y$ означает, что x является началом y. Через Ξ обозначим пространство частичных функций из N в $\{0, 1\}$, f-объектами которого являются функции с конечной областью определения, а $x \leqslant y$ означает, что y продолжает x. Заменив $\{0, 1\}$ на N, получим пространство, которое мы будем обозначать F. Во всех перечисленных пространствах естественно вводится структура эффективного f_0-пространства. Все они полны. В дальнейшем эффективные полные f_0-пространства будем для краткости называть просто пространствами.

Объект x пространства $\langle X, X_0, \leqslant, \nu\rangle$ вычислим, если множество $\{n | \nu(n) \leqslant x\}$ перечислимо. Для всякого пространства X существует вычислимый объект из $C(N_\perp, X)$, образом которого является множество всех вычислимых объектов пространства X.

Функцию l, сопоставляющую с f-объектами пространства X натуральные числа, мы будем называть объемом, если из $x_1 \leqslant x_2$ следует $l(x_1) \leqslant l(x_2)$ и функция $n \mapsto l(\nu(n))$ вычислима. Основные для нас примеры объемов: в $N_\perp$ определим объем так, что $l(\perp) = 0, l(n) =$ (целая часть $\log_2(1+n)$)+1, на Ω объем совпадает с длиной, на Ξ объем x равен числу элементов в области определения функции x.

2. Задачи и их энтропия. Пусть X – пространство, A – некоторое множество объектов X. Всякую пару $\langle X, A\rangle$ мы будем называть задачей, X – пространством задачи $\langle X, A\rangle$, а объекты из A – решениями задачи $\langle X, A\rangle$. Мы будем интерпретировать ее как задачу отыскания среди объектов, принадлежащих пространству X, объекта, входящего в множество A. Назовем задачу монотонной, если из $x \in A$, $x \leqslant y$ следует $y \in A$, и разрешимой, если в A имеется вычислимый объект. Пусть X и Y – пространства, l – объем на пространстве X. Способом описания объектов Y с помощью объектов X будем называть любой вычислимый объект пространства $C(X, Y)$. Пусть дан способ описания $f \in C(X, Y)$ и задача $\alpha = \langle Y, A\rangle$ в пространстве Y. Число

$$K_f(\alpha) = K_f(\langle Y, A\rangle) = \inf\{ l(x_0) \mid x_0 \text{ – конечный объект } X, f(x_0) \in A\}$$

называется сложностью задачи α при способе описания f. Будем говорить, что способ описания $f \in C(X, Y)$ эффективнее способа описания $g \in C(X, Y)$, если существует такое C, что для любой задачи $\alpha = \langle Y, A\rangle$ в пространстве Y выполнено неравенство $K_f(\alpha) \leqslant K_g(\alpha) + C$. Способ описания $f \in C(X, Y)$ назовем оптимальным, если он эффективнее любого способа описания $g \in C(X, Y)$. Назовем пространство X с объемом l регулярным, если для любого пространства Y в $C(X, Y)$ существует оптимальный способ описания.

Теорема 1. *Пространство X с объемом l регулярно тогда и только тогда, когда существует способ описания $f \in C(X, X \times N_\perp)$, для которого*

$$(\forall n \in N)(\exists C)(\forall x_0 \in X_0)(K_f(\langle X \times N_\perp, \{\langle x_0, n\rangle\}\rangle) \leqslant l(x_0) + C).$$

Из этой теоремы следует, что пространства N, Ω и Ξ регулярны.

Пусть пространство X с объемом l регулярно. Для любого пространства Y выберем оптимальный способ описания $f \in C(X, Y)$ и назовем энтропией $K_X(\alpha)$ задачи α в пространстве Y относительно пространства X ее сложность при способе описания f. Таким образом, для данного пространства Y энтропия задач в этом пространстве определяется с точностью до ограниченного слагаемого.

Т е о р е м а 2. *Пусть X – регулярное пространство с объемом, Y – произвольное пространство, α – задача в Y. Тогда энтропия $K_X(\alpha)$ конечна тогда и только тогда, когда задача α разрешима.*

Пусть (X_1, l_1) и (X_2, l_2) – регулярные пространства с объемом, f – монотонно возрастающая функция, удовлетворяющая условию Липшица.

Т е о р е м а 3. *Следующие свойства равносильны:*

1) *для любого пространства Y существует такое C, что для любой задачи $\alpha = \langle Y, A \rangle$ имеет место неравенство*

$$K_{X_1, l_1}(\alpha) \leqslant f(K_{X_2, l_2}(\alpha)) + C;$$

2) *существует такое C, что для всякого конечного объекта $x_2 \in X_2$ выполнено неравенство*

$$K_{X_1, l_1}(\langle X_2, \{x_2\} \rangle) \leqslant f(l_2(x_2)) + C.$$

Теорема останется справедливой, если заменить в условии 1) "для любой задачи" на "для любой монотонной задачи", а в условии (2) заменить $\langle X, \{x_2\} \rangle$ на $\langle X, \Gamma_{x_2} \rangle$, где $\Gamma_{x_2} = \{x \in X_2 \mid x \geqslant x_2\}$. Полученные условия назовем 1') и 2'). Если условия 1') и 2') выполнены для функции $f(n) = n$, будем говорить, что X_1 н е х у ж е X_2; если они выполнены для функции $f(n) = n + C \log_2 n$ при некотором C, то будем говорить, что X_1 п о ч т и н е х у ж е X_2.

Т е о р е м а 4. *Имеют место соотношения $\mathrm{N}_\perp \rightleftarrows \Omega \leftarrow \Xi$, где $X \to Y$ означает, что X не хуже Y, а $X \rightleftarrows Y$ означает, что X почти не хуже Y; иных соотношений (за исключением соотношения $\Xi \dashrightarrow \mathrm{N}_\perp$, вытекающего из указанных) нет.*

Определим логические операции над задачами. Пусть $\alpha = \langle X, A \rangle$ и $\beta = \langle Y, B \rangle$ – две задачи. Определим задачи $\alpha \wedge \beta = \langle X \times Y, A \times B \rangle$, $\alpha \vee \beta = \langle X + Y, A \cup B \rangle$ (мы считаем, что X и Y не пересекаются) и $\alpha \supset \beta = \langle C(X, Y), \{f \mid f(A) \subset B\} \rangle$. Ложью назовем задачу $F = \langle P, \phi \rangle$, где P – пространство, содержащее единственный конечный объект.

Энтропию $K_X(\alpha \supset \beta)$ задачи $\alpha \supset \beta$ мы будем называть у с л о в н о й э н т р о п и е й β п р и и з в е с т н о й α и обозначать $K(\beta \mid \alpha)$.

Пусть $\Phi(p_1, p_2, \ldots, p_n)$ – пропозициональная формула, содержащая знаки $\wedge$, $\vee$, $\supset$ и F (ложь). Если вместо $p_1, p_2, \ldots, p_n$ подставить задачи $\alpha_1 = \langle X_1, A_1 \rangle$, $\alpha_2 = \langle X_2, A_2 \rangle, \ldots, \alpha_n = \langle X_n, A_n \rangle$, то возникнет некоторая задача $\Phi(\alpha_1, \alpha_2, \ldots, \alpha_n)$. Пространство этой задачи определяется пространствами $X_1, X_2, \ldots, X_n$ и не зависит от A_i; обозначим его $\Phi(X_1, X_2, \ldots, X_n)$.

Т е о р е м а 5. *Пусть $\Phi(p_1, p_2, \ldots, p_n)$ выводима в интуиционистском исчислении высказываний, $X_1, X_2, \ldots, X_n$ – пространства.*

Тогда существует вычислимый объект в пространстве $\Phi(X_1, X_2, \ldots, X_n)$, являющийся решением задачи $\Phi(\langle X_1, A_1 \rangle, \langle X_2, A_2 \rangle, \ldots, \langle X_n, A_n \rangle)$ при любых $A_i \subset X_i$.

Т е о р е м а 6. *Пусть $\Phi(p_1, p_2, \ldots, p_n) \supset \Psi(p_1, p_2, \ldots, p_n)$ – выводимая формула интуиционистского исчисления высказываний, $X_1, X_2, \ldots, X_n$ – пространства, X – регулярное пространство с объемом.*

Тогда существует такое C, что для любых задач $\alpha_1, \alpha_2, \ldots, \alpha_n$ в пространствах $X_1, X_2, \ldots, X_n$ выполнено неравенство $K_X(\Psi(\alpha_1, \alpha_2, \ldots, \alpha_n)) \leqslant K_X(\Phi(\alpha_1, \alpha_2, \ldots, \alpha_n)) + C$.

Отсюда вытекают неравенства $K_X(\alpha) \leqslant K_X(\alpha \wedge \beta) + O(1)$, $K_X(\alpha \mid \beta) \leqslant K_X(\alpha) + O(1)$, $K_X(\beta) \leqslant K_X(\alpha \wedge (\alpha \supset \beta)) + O(1)$ и многие другие.

Рассмотрим множество Q формул, для которых выполнено утверждение теоремы 5. Оно является суперинтуиционистской логикой.

Т е о р е м а 7. *Логика Q не совпадает ни с интуиционистской, ни с классической; она также отличается от логики финитных задач Ю.Т. Медведева* [7].

Т е о р е м а 8. а) $K_{\mathbf{N}_\perp}(\langle \mathbf{N}_\perp, \{n\}\rangle)$ = (*сложность n в смысле* [3]) + $O(1)$;

б) $K_{\mathbf{N}}(\langle \Omega, \Gamma_x\rangle)$ = (*сложность разрешения последовательности x в смысле* [5]) + $O(1)$;

в) $K_\Omega(\langle \mathbf{N}_\perp, \{n\}\rangle)$ = (*префиксная энтропия n в смысле* [6]) + $O(1)$;

г) $K_\Omega(\langle \Omega, \Gamma_x\rangle)$ = (*монотонная энтропия последовательности x в смысле* [6]) + $O(1)$;

д) $K(\langle \mathbf{N}_\perp \{n\}\rangle \supset \langle \mathbf{N}_\perp \{m\}\rangle)$ = (*условная сложность m относительно n в смысле* [3]) + $O(1)$;

е) $K_{\mathbf{N}}(\langle F, \{f\}\rangle)$ = (*логарифм номера вычислимой функции f при оптимальной в смысле* [4], *с.* 151, *нумерации*) + $O(1)$.

Напомним, что через Γ_x обозначается $\{y \mid x \leqslant y\}$.

Пусть X – произвольное пространство, $f \in C(\Omega, X)$ – способ описания. С каждой задачей $\alpha = \langle X, A\rangle$, где A – борелевское подмножество пространства X (относительно рассмотренной топологии), сопоставим число $P_f(\alpha)$ = мера ($\{\omega$ – бесконечная последовательность цифр 0 и 1 $\mid f(\omega) \in A\}$), называемое в е р о я т н о с т ь ю р е ш е н и я з а д а ч и α п р и с п о с о б е f. Среди всех способов описания существует оптимальный, при котором $P_f(\alpha)$ максимальна с точностью до мультипликативной константы: для всякого другого способа g найдется такое $C > 0$, что $P_f(\alpha) \geqslant CP_g(\alpha)$ для всех задач α в пространстве X. Выбрав и зафиксировав оптимальный способ f, будем называть $P_f(\alpha)$ а п р и о р н о й в е р о я т н о с т ь ю задачи α и обозначать $P(\alpha)$. При $X = \mathbf{N}_\perp$ априорная вероятность задачи $\langle \mathbf{N}_\perp, \{n\}\rangle$ совпадает (с точностью до ограниченного отделенного от нуля множителя) с введенной в [6], с. 26; при $X = \Omega$ априорная вероятность задачи $\langle \Omega, \Gamma_x\rangle$ совпадает с введенной в [5], с. 49 (полумера M из теоремы 4.1).

Пусть P – мера, определенная на борелевских подмножествах пространства X. Будем называть ее п е р е ч и с л и м о й, если множество $\{\langle n, r\rangle \in \mathbf{N} \times \mathbf{Q} \mid r < P(\Gamma_{\nu(n)})\}$ перечислимо (здесь ν – нумерация, фигурирующая в определении эффективного пространства). Априорная вероятность является перечислимой мерой.

Т е о р е м а 9. *Если для любых совместных конечных объектов x и y пространства X выполнено $x \leqslant y$ или $y \leqslant x$, то априорная вероятность на X является максимальной с точностью до мультипликативной константы перечислимой мерой. Условие, накладываемое на пространство X, существенно: в пространстве Ξ априорная вероятность не является максимальной перечислимой мерой.*

Т е о р е м а 10. а) *Имеет место неравенство* $-\log_2 P(\alpha) \leqslant K_\Omega(\alpha) + O(1)$, *где величина* $O(1)$ *зависит только от пространства задачи* α.

б) *Обратное неравенство* $K_\Omega(\alpha) \leqslant -\log_2 P(\alpha) + O(1)$ *не имеет места уже для задач вида* $\langle \Xi, \Gamma_x\rangle$.

в) *Существует регулярное пространство M с объемом l, для которого* $K_{M,l}(\alpha) = -\log_2 P(\alpha) + O(1)$ *для всех задач вида* $\langle X, \Gamma_x\rangle$, *где X – произвольное пространство (от него зависит граница для* $O(1)$), *а x – любой конечный объект в X.*

г) *Не существует регулярного пространства, для которого соотношение п.* в) *выполнялось бы для всех монотонных задач во всяком пространстве X.*

Институт проблем передачи информации
Академии наук СССР, Москва

Поступило
1 VII 1983

ЛИТЕРАТУРА

1. *Ершов Ю.Л.* – Алгебра и логика, 1972, т. 11, № 4, с. 367–437. 2. *Kolmogoroff A.N.* – Math. Z., 1932, Bd. 35, H. 1, S. 58–65. 3. *Колмогоров А.Н.* – Проблемы передачи информации, 1965, т. 1, вып. 1, с. 3–11. 4. *Uspensky V.A., Semenov A.L.* In: Algoritms in modern mathematics and computer science. B.: Springer–Verlag, 1981. (Lecture notes in computer sci., vol. 122), p. 100–234. 5. *Звонкин А.К., Левин Л.А.* – УМН, 1970, т. 25, вып. 6 (156), с. 85–127. 6. *Вьюгин В.В.* – Семиотика и информатика, 1980, вып. 16, с. 14–43. 7. *Медведев Ю.Т.* – ДАН, 1962, т. 142, № 5, с. 1015–1018.