Chaitin, G.J.
Osaka J. Math.
14 (1977), 139–149

PROGRAM SIZE, ORACLES, AND THE JUMP OPERATION

GREGORY J. CHAITIN

(Received January 17, 1976)

There are a number of questions regarding the size of programs for calculating natural numbers, sequences, sets, and functions, which are best answered by considering computations in which one is allowed to consult an oracle for the halting problem. Questions of this kind suggested by work of T. Kamae and D. W. Loveland are treated.

1. Computer programs, oracles, information measures, and codings

In this paper we use as much as possible Rogers' terminology and notation [1, pp. xv–xix]. Thus $N=\{0, 1, 2, \cdots\}$ is the set of (natural) numbers; $i, j, k, n, v, w, x, y, z$ are elements of N; A, B, X are subsets of N; f, g, h are functions from N into N; φ, ψ are partial functions from N into N; $\langle x_1, \cdots, x_k\rangle$ denotes the ordered k-tuple consisting of the numbers $x_1, \cdots, x_k$; the lambda notation $\lambda x[\cdots x\cdots]$ is used to denote the partial function of x whose value is $\cdots x\cdots$; and the *mu* notation $\mu x[\cdots x\cdots]$ is used to denote the least x such that $\cdots x\cdots$ is true.

The size of the number x, denoted $lg(x)$, is defined to be the number of bits in the xth binary string. The binary strings are: $\wedge$, 0, 1, 00, 01, 10, 11, 000,$\cdots$ Thus $lg(x)$ is the integer part of $\log_2(x+1)$. Note that there are 2^n numbers x of size n, and 2^n-1 numbers x of size less than n.

We are interested in the size of programs for a certain class of computers. The zth computer in this class is defined in terms of $\varphi_z^{(2)X}$ [1, pp. 128–134], which is the two-variable partial X-recursive function with Gödel number z. These computers use an oracle for deciding membership in the set X, and the zth computer produces the output $\varphi_z^{(2)X}(x, y)$ when given the program x and the data y. Thus the output depends on the set X as well as the numbers x and y.

We now choose the standard universal computer U that can simulate any other computer. U is defined as follows:

$$U^X((2x+1)2^z-1, y) = \varphi_z^{(2)X}(x, y)\,.$$

Thus for each computer C there is a constant c such that any program of size

n for C can be simulated by a program of size $\leq n+c$ for U.

Having picked the standard computer U, we can now define the program size measures that will be used throughout this paper.

The fundamental concept we shall deal with is $I(\psi/X)$, which is the number of bits of information needed to specify an algorithm relative to X for the partial function ψ, or, more briefly, the information in ψ relative to X. This is defined to be the size of the smallest program for ψ:

$$I(\psi/X) = \min lg(x)\ (\psi = \lambda y[U^X(x, y)])\,.$$

Here it is understood that $I(\psi/X)=\infty$ if ψ is not partial X-recursive.

$I(x\rightarrow y/X)$, which is the information relative to X to go from the number x to the number y, is defined as follows:

$$I(x\rightarrow y/X) = \min I(\psi/X)\ (\psi(x) = y)\,.$$

And $I(x/X)$, which is the information in the number x relative to the set X, is defined as follows:

$$I(x/X) = I(0\rightarrow x/X)\,.$$

Finally $I(\psi/X)$ is used to define three versions $I(A/X)$, $I_r(A/X)$, and $I_f(A/X)$ of the information relative to X of a set A. These correspond to the three ways of naming a set [1, pp. 69–71]: by r.e. indices, by characteristic indices, and by canonical indices. The first definition is as follows:

$I(A/X)=I(\lambda x[\text{if } x\in A \text{ then } 1 \text{ else undefined}]/X)$.

Thus $I(A/X)<\infty$ iff A is r.e. in X. The second definition is as follows:

$I_r(A/X)=I(\lambda x[\text{if } x\in A \text{ then } 1 \text{ else } 0]/X)$.

Thus $I_r(A/X)<\infty$ iff A is recursive in X. And the third definition, which applies only to finite sets, is as follows:

$$I_f(A/X) = I(\sum_{x\in A} 2^x/X)\,.$$

The following notational convention is used: $I(\psi)$, $I(x\rightarrow y)$, $I(x)$, $I(A)$, $I_r(A)$, and $I_f(A)$ are abbreviations for $I(\psi/\phi)$, $I(x\rightarrow y/\phi)$, $I(x/\phi)$, $I(A/\phi)$, $I_r(A/\phi)$, and $I_f(A/\phi)$, respectively.

We use the coding τ^* of finite sequences of numbers into individual numbers [1, p. 71]; τ^* is an effective one-one mapping from $\cup_{k=0}^{\infty} N^k$ onto N. And we also use the notation $\bar{f}(x)$ for τ^* of the sequence $\langle f(0), f(1),\cdots, f(x-1)\rangle$ [1, p. 377]; for any function f, $\bar{f}(x)$ is the code number for the finite initial segment of f of length x.

The following theorems, whose straight-forward proofs are omitted, give some basic properties of these concepts.

Theorem 1.

(a) $I(x/X)\leq lg(x)+c$

(b) *There are less than 2^n numbers x with $I(x/X)<n$.*
(c) $|I(x/X)-I(y/X)|\leq 2\ lg(|x-y|)+c$
(d) *The set of all true propositions of the form "$I(x\to y/X)\leq z$" is r.e. in X.*
(e) $I(x\to y/X)\leq I(y/X)+c$

Recall that there are 2^n numbers x of size n, that is, there are 2^n numbers x with $lg(x)=n$. In view of (a) and (b) most x of size n have $I(x/X)\approx n$. Such x are said to be X-random. In other words, x is said to be X-random if $I(x/X)$ is approximately equal to $lg(x)$; most x have this property.

Theorem 2.
(a) $I(\tau^*(\langle x, y\rangle))\leq I(\tau^*(\langle y, x\rangle))+c$
(b) $I(\tau^*(\langle x, y\rangle)\to\tau^*(\langle y, x\rangle))\leq c$
(c) $I(x)\leq I(\tau^*(\langle x, y\rangle))+c$
(d) $I(\tau^*(\langle x, y\rangle)\to x)\leq c$

Theorem 3.
(a) $I(x\to\psi(x)/X)\leq I(\psi/X)$
(b) *For each ψ that is partial X-recursive there is a c such that $I(\psi(x)/X)\leq I(x/X)+c$.*
(c) $I(x\to\bar{f}(x)/X)\leq I(f/X)+c$
(d) $I(\bar{f}(x)\to x/X)\leq c$ *and* $I(x/X)\leq I(\bar{f}(x)/X)+c$

Theorem 4.
(a) $I(x/X)\leq I(\lambda y[x]/X)$ *and* $I(\lambda y[x]/X)\leq I(x/X)+c$
(b) $I(x/X)\leq I_f(\{x\}/X)+c$ *and* $I_f(\{x\}/X)\leq I(x/X)+c$
(c) $I(x/X)\leq I_r(\{x\}/X)+c$ *and* $I_r(\{x\}/X)\leq I(x/X)+c$
(d) $I(x/X)\leq I(\{x\}/X)+c$ *and* $I(\{x\}/X)\leq I(x/X)+c$
(e) $I_r(A/X)\leq I_f(A/X)+c$ *and* $I(A/X)\leq I_r(A/X)+c$

See [2] for a different approach to defining program size measures for functions, numbers, and sets.

2. The jump and limit operations

The jump X' of a set X is defined in such a manner that having an oracle for deciding membership in X' is equivalent to being able to solve the halting problem for algorithms relative to X [1, pp. 254–265].

In this paper we study a number of questions regarding the information $I(\psi)$ in ψ relative to the empty set, that are best answered by considering $I(\psi/\phi')$ and $I(\psi/\phi'')$, which are the information in ψ relative to the halting problem and relative to the jump of the halting problem. The thesis of this paper is that in order to understand $I(\psi/X)$ with $X=\phi$, which is the case of practical significance, it is sometimes necessary to jump higher in the arithmetical hierarchy to $X=\phi'$ or $X=\phi''$.

The following theorem, whose straight-forward proof is omitted, gives some facts about how the jump operation affects program size measures.

Theorem 5.

(a) $\lambda xy[I(x\rightarrow y/X)]$ *and* $\lambda x[I(x/X)]$ *are* X'*-recursive.*

(b) $I(\psi/X')\leq I(\psi/X)+c$

(c) *For each* n *consider the least* x *such that* $lg(x)\geq n$ *and* $I(x/X)\geq n$. *This* x *has the property that* $lg(x)=n$, $I(x/X)\leq n+c_1$, *and* $I(x/X')\leq I(n/X')+c_2$ $\leq lg(n)+c_3$.

(d) $I(\bar{A}/X')\leq I(A/X)+c$

(e) $I_r(A/X')\leq I(A/X)+c$

(f) *If* A *is finite* $I_f(A/X')\leq I(A/X)+c$.

(g) $I(X'/X)\leq c$ *and* $I_r(X'/X)=\infty$

It follows from (b) that X'-randomness implies X-randomness. However (c) shows that the converse is false: there are X-random numbers that are not at all X'-random.

Having examined the jump operation, we now introduce the limit operation. The following theorem shows that the limit operation is in a certain sense analogous to the jump operation. This theorem is the tool we shall use to study work of Kamae and Loveland in the following sections.

DEFINITION.

Consider a function f. $\lim_x f(x)$ denotes a number z having the property that there is an x_0 such that $f(x)=z$ if $x\geq x_0$. If no such z exists, $\lim_x f(x)$ is undefined. In other words $\lim_x f(x)$ is the value that $f(x)$ assumes for almost all x (if there is such a value.)

Theorem 6.

(a) *If* $I(z/X')<n$, *then there is a function* f *such that* $z=\lim_x f(x)$ *and* $I(f/X)<n+c$.

(b) *If* $I(f/X)<n$ *and* $\lim_x f(x)=z$, *then* $I(z/X')<n+c$.

Proof.

(a) By hypothesis there is a program w of size less than n such that $U^{X'}(w, 0)=z$. Given w and an arbitrary number x, one calculates $f(x)$ using the oracle for membership in X as follows. Choose a fixed algorithm relative to X for enumerating X'.

One performs x steps of the computation of $U^{X'}(w, 0)$. This is done using a fake oracle for X' that answers that v is in X' iff v is obtained during the first x steps of the algorithm relative to X for enumerating X'. If a result is obtained by performing x steps of $U^{X'}(w, 0)$ in this manner, that is the value of $f(x)$. If not, $f(x)$ is 0.

It is easy to see that $\lim_x f(x)=U^{X'}(w, 0)=z$ and $I(f/X)\leq lg(w)+c<n+c$.

(b) By hypothesis there is a program w of size less than n such that $\lim_x U^X(w,x) =z$. Given w one can use the oracle for X' to calculate z. At stage i one asks the oracle whether there is a $j>i$ such that $U^X(w,j)\neq U^X(w,i)$. If so, one goes to stage $i+1$ and tries again. If not, one is finished because $U^X(w,i)=z$. This shows that $I(z/X')\leq lg(w)+c<n+c$. Q.E.D.

See [3] for applications of oracles and the jump operation in the context of self-delimiting programs for sets and probability constructs; in this paper we are only interested in programs with endmarkers.

3. The Kamae information measure

In this section we study an information measure $K(x)$ suggested by work of Kamae [4] (see also [5]).

$I(y\rightarrow x)$ is less than or equal to $I(x)+c$, and it is natural to call $I(x)-I(y\rightarrow x)$ the degree to which y is helpful to x. Let us look at some examples. By definition $I(x)=I(0\rightarrow x)$, and so 0 is no help at all. On the other hand some y are very helpful: $I(y\rightarrow x)<c$ for all those y whose prime factorization has 2 raised to the power x. Thus every x has infinitely many y that are extremely helpful to it.

Kamae proves in [4] that for each c there is an x such that $I(y\rightarrow x)<I(x)-c$ holds for almost all y. In other words, for each c there is an x such that almost all y are helpful to x more than c. This is surprising; one would have expected there to be a c with the property that every x has infinitely many y that are helpful to it less than c, that is, infinitely many y with $I(y\rightarrow x)>I(x)-c$.

We shall now study how $I(y\rightarrow x/X)$ varies when x is held fixed and y goes to infinity. Note that $I(y\rightarrow x/X)$ is bounded (in fact, by $I(x/X)+c$). This suggests the following definition: $K(x/X)$ is the greatest w such that $I(y\rightarrow x/X)=w$ holds for infinitely many y. In other words, $K(x/X)$ is the least v such that $I(y\rightarrow x/X) \leq v$ holds for almost all y.

Note that there are less than 2^n numbers x with $K(x/X)<n$, so that $K(x/X)$ clearly measures bits of information in some sense. The trivial inequality $K(x/X)\leq I(x/X)+c$ relates $K(x/X)$ and $I(x/X)$, but the following theorem shows that $K(x/X)$ is actually much more intimately related to the information measures $I(x/X')$ and $I(x/X'')$ than to $I(x/X)$.

Theorem 7.

(a) $K(x/X)\leq I(x/X')+c$

(b) $I(x/X'')\leq K(x/X)+c$

Proof.

(a) Consider a number x_0. By Theorem 6a there is a function f such that

$\lim_y f(y)=x_0$ and $I(f/X)\leq I(x_0/X')+c$. Hence $I(y\to f(y)/X)\leq I(f/X)\leq I(x_0/X')+c$. In as much as $f(y)=x_0$ for almost all y, it follows that $I(y\to x_0/X)\leq I(x_0/X')+c$ for almost all y. Hence $K(x_0/X)\leq I(x_0/X')+c$.

(b) By using an oracle for membership in X' one can decide whether or not $I(y\to x/X)<n$. Thus by using an oracle for membership in X'' one can decide whether or not y_0 has the property that $I(y\to x/X)<n$ for all $y\geq y_0$. It follows that the set $A_n=\{x\mid K(x/X)<n\}$ is r.e. in X'' uniformly in n.

Suppose that $x_0\in A_n$. Consider $j=2^n+k$, where $k=$(the position of x_0 in a fixed X''-recursive enumeration of A_n uniform in n). Since there are less than 2^n numbers in A_n, $k<2^n$ and $2^n\leq j<2^{n+1}$. Thus one can recover from j the values of n and k. And if one is given j one can calculate x_0 using an oracle for membership in X''. Thus if $K(x_0/X)<n$, then $I(x_0/X'')\leq lg(j)+c_1\leq n+c_2$.

Q.E.D.

What is the significance of this theorem? First of all, note that most x are ϕ''-random and thus have $lg(x)\approx I(x/\phi'')\approx I(x/\phi')\approx I(x)\approx K(x)$. In other words, there is a c such that every n has the property that at least 99% of the x of size n have all four quantities $I(x/\phi'')$, $I(x/\phi')$, $I(x)$, and $K(x)$ inside the interval between $n-c$ and $n+c$. These x are "normal" because there are infinitely many y that do not help x at all, that is, there are infinitely many y with $I(y\to x)>I(x)-c$.

Now let us look at the other extreme, at the "abnormal" x discovered by Kamae.

Consider the first ϕ-random number of size n, where n itself is ϕ''-random. More precisely, let x_n be the first x such that $lg(x)=n$ and $I(x)\geq n$. (There is such an x_n, because there are 2^n numbers x of size n, and at most 2^n-1 of these x have $I(x)<n$.) Moreover, we stipulate that n itself be ϕ''-random, so that $I(n/\phi'')\approx lg(n)$.

It is easy to see that these x_n have the property that $lg(n)\approx I(x_n/\phi'')\approx I(x_n/\phi')\approx K(x_n)$ and $I(x_n)\approx lg(x_n)=n$. Thus most y help these x_n a great deal, because $I(x_n)\approx n$ and for almost all y, $I(y\to x_n)\lesssim \log_2 n$.

Theorem 7 enables us to make very precise statements about $K(x)$ when $I(x/\phi'')\approx I(x/\phi')$. But where is $K(x)$ in the interval between $I(x/\phi'')$ and $I(x/\phi')$ when this interval is wide? The following theorem shows that if $I(x/\phi'')$ and $I(x/\phi')$ are many orders of magnitude apart, then $K(x)$ will be of the same order of magnitude as $I(x/\phi')$. To be more precise, Theorems 7a and 8 show that $\frac{1}{2}I(x/\phi')-c\leq K(x)\leq I(x/\phi')+c$.

Theorem 8.

If $K(x_0/X)<n$, then $I(x_0/X')<2n+c$.

Proof.

Consider a fixed number n and a fixed set X. Let $\sharp_x$ be the cardinality of the set $B_x=\{z \mid I(x\to z/X)<n\}$. Note that $\#_x$ is bounded (in fact, by 2^n-1). Let i be the greatest w such that $\#_x=w$ holds for infinitely many x, which is also the least v such that $\#_x\leq v$ holds for almost all x. Let $j=\mu z[\#_x\leq i$ if $x\geq z]$, and let A be the infinite set of x greater than or equal to j such that $\#_x=i$. Thus B_x has exactly i elements if $x\in A$.

It is not difficult to see that if one knows n and i, then one can calculate j by using an oracle for membership in X'. And if one knows n, i, and j, by using an oracle for membership in X one can enumerate A and simultaneously calculate for each $x\in A$ the canonical index $\sum 2^z(z\in B_x)$ of the i-element set B_x.

Define $J(x)$ as follows: $J(x)=$(the greatest w such that $I(y\to x/X)=w$ holds for infinitely many $y\in A$)$=$(the least v such that $I(y\to x/X)\leq v$ holds for almost all $y\in A$). It is not difficult to see from the previous paragraph that if one is given n and i and uses an oracle for membership in X', one can enumerate the set of all x such that $J(x)<n$.

Note that there are less than 2^n numbers x with $J(x)<n$, and that if $K(x/X)<n$, then $J(x)<n$. Suppose that x_0 has the property that $K(x_0/X)<n$. Consider the number $k=(2^n+i)2^n+i_2$, where $i_2=$(the position of x_0 in the above-mentioned X'-recursive enumeration of $\{x \mid J(x)<n\}$). Since $i<2^n$ and $i_2<2^n$, one can recover n, i, and i_2 from k.

It is not difficult to see that if one is given k, then one can calculate x_0 using an oracle for membership in X'. Thus if $K(x_0/X)<n$, then $I(x_0/X')\leq lg(k)+c_1<2n+c_2$. Q.E.D.

4. The Loveland information measure

Define $L(f/X)$ to be $\max_x I(x\to \bar{f}(x)/X)$, and to be ∞ if $I(x\to \bar{f}(x)/X)$ is unbounded. This concept is suggested by work of Loveland [6]. Since there are less than 2^n functions f with $L(f/X)<n$, it is clear that in some sense $L(f/X)$ measures bits of information. $I(x\to \bar{f}(x)/X)$ is bounded if f is X-recursive, and conversely A. R. Meyer [6, pp. 525–526] has shown that if $I(x\to \bar{f}(x)/X)$ is bounded then f is X-recursive. Thus $L(f/X)<\infty$ iff $I(f/X)<\infty$.

But can something more precise be said about the relationship between $L(f)$ and $I(f)$? $L(f)\leq I(f)+c$, but as is pointed out in [6, p. 515], the proof that $I(f)<\infty$ if $L(f)<\infty$ is nonconstructive and does not give an upper bound on $I(f)$ in terms of $L(f)$. We shall show that in fact $I(f)$ can be enormous for reasonable values of $L(f)$. The proof that $I(f)<\infty$ if $L(f)<\infty$ may therefore be said to be extremely nonconstructive.

In [7] it is shown that $I(f)<\infty$ iff there is a c such that $I(\bar{f}(x))-I(x)\leq c$ for all x. This result is now also seen to be extremely nonconstructive, because $I(f)$ may be enormous for reasonable c.

Furthermore, R. M. Solovay has studied in [8] what is the situation if the endmarker program size measure I used here is replaced by the self-delimiting program size measure H of [9]. He shows that there is a nonrecursive function f such that $H(\bar{f}(x))-H(x)$ is bounded. This result previously seemed to contrast sharply with the fact that f is recursive if $I(x\rightarrow\bar{f}(x))$ is bounded [6] or if $I(\bar{f}(x))-I(x)$ is bounded [7]. But now a harmonious whole is perceived since the sufficient conditions for f to be recursive just barely manage to keep $I(f)$ from being ∞.

Theorem 9.

If $I(k/X')\leq n$, then there is a function f such that $L(f/X)\leq n+c$ and $I(f/X)\geq k-c$.

Proof.

First we define the function g as follows: $g(x)$ is the first non-zero y such that $I(y/X)\geq x$. Note that g is X'-recursive.

By hypothesis $I(k/X')\leq n$. Hence $I(g(k)/X')\leq I(k/X')+c_1\leq n+c_1$. By Theorem 6a, there is a function h such that $I(h/X)\leq n+c_2$ and $\lim_x h(x)=g(k)$. Let $x_0=\mu z[h(x)=g(k)$ if $x\geq z]$. Thus $h(x)=g(k)$ if $x\geq x_0$.

The function f whose existence is claimed is defined as follows:

$$f(x)=\begin{cases}0 \text{ if } x<x_0, \text{ and}\\ h(x) \text{ if } x\geq x_0\,.\end{cases}$$

Thus $f(x)=g(k)$ if $x\geq x_0$.

First we obtain a lower bound for $I(f/X)$. The following holds for any function f:

$$I(f(\mu x[f(x)\neq 0])/X)\leq I(f/X)+c_3\,.$$

Hence for this particular f we see that $I(f/X)+c_3\geq I(f(x_0)/X)=I(g(k)/X)$. Thus, by the definition of g, $I(f/X)+c_3\geq I(g(k)/X)\geq k$.

Next we obtain an upper bound for $I(x\rightarrow\bar{f}(x)/X)$. There are two cases: $x\leq x_0$ and $x>x_0$. If $x\leq x_0$, then $\bar{f}(x)$ is the code number for a sequence of x 0's and thus $I(x\rightarrow\bar{f}(x)/X)\leq I(\lambda x[\tau^*(\langle 0\rangle^x)]/X)=c_4$, where $\langle 0\rangle^x$ denotes a sequence of x 0's. If $x>x_0$, then

$$\begin{aligned}I(x\rightarrow\bar{f}(x)/X)\leq I(x\rightarrow\tau^*(\langle h(x),\, x_0\rangle)/X)+c_5 =&\\ I(x\rightarrow\tau^*(\langle h(x),\, \mu z[h(x)=h(y) \text{ if } x\geq y\geq z]\rangle)/X)+c_5&\\ \leq I(h/X)+c_6\leq n+c_2+c_6\,.&\end{aligned}$$

Thus $I(x\rightarrow\bar{f}(x)/X)$ is either bounded by c_4 or by $n+c_2+c_6$. Hence $I(x\rightarrow\bar{f}(x)/X)\leq n+c_7$ and $L(f/X)\leq n+c_7$.

To recapitulate, we have shown that this f has the property that $I(f/X)\geq k-c_3$ and $L(f/X)\leq n+c_7$. Taking $c=\max c_3, c_7$, we see that $I(f/X)\geq k-c$ and $L(f/X)\leq n+c$. Q.E.D.

Why does Theorem 9 show that $I(f)$ can be enormous even though $L(f)$ has a reasonable value? Consider the function $g(x)$ defined to be $(\cdots((x!)!)!\cdots!)$ in which there are x!'s. $g(x)$ quickly becomes astronomical as x increases. However, $I(g(n)/\phi') \leq I(g(n)) + c_1 \leq I(n) + c_2 \leq lg(n) + c_3$, and $lg(n)+c_3$ is less than n for almost all n. Hence almost all n have the property that there is a function f with $L(f) \leq n+c$ and $I(f) \geq g(n)-c$.

In fact the situation is much worse. It is easy to define a function h that is ϕ'-recursive and grows more quickly than any recursive function. In other words, h is recursive in the halting problem and for any recursive function g, $h(x) > g(x)$ for almost all x. As before we see that $I(h(n)/\phi') < n$ for almost all n. Hence almost all n have the property that there is a function f with $L(f) \leq n+c$ and $I(f) \geq h(n)-c$.

5. Other applications

In this section some other applications of oracles and the jump operation are presented without proof.

First of all, we would like to examine a question raised by C. P. Schnorr [10, p. 189] concerning the relationship between $I(x)$ and the limiting relative frequency of programs for x. However, it is more appropriate to ask what is the relationship between the self-delimiting program size measure $H(x)$ [9] and the limiting relative frequency of programs for x (with endmarkers). Define $F(x, n)$ to be $-\log_2$ of (the number of programs w less than or equal to n such that $U^{\phi}(w, 0)=x)/(n+1)$. Then Theorem 12 of [10] is analogous to the following:

Theorem 10.
There is a c such that every x satisfies $F(x, n) \geq H(x)-c$ for almost all n.

This shows that if $H(x)$ is small, then x has many programs. Schnorr asks whether the converse is true. In fact it is not:

Theorem 11.
There is a c such that every x satisfies $F(x, n) \geq H(x/\phi')-c$ for almost all n.

Thus even though $H(x)$ is large, x will have many programs if $H(x/\phi')$ is small.

We would like to end by examining the maximum finite cardinality $\sharp A$ and co-cardinality $\sharp \bar{A}$ attainable by a set A of bounded program size. First we define the partial function G:

$$G(x/X) = \max z \; (I(z/X) \leq x) \, .$$

The following easily established results show how gigantic G is:

(a) If ψ is partial X-recursive and $x > I(\psi/X)+c$, then $\psi(x)$, if defined, is less than $G(x/X)$.

(b) If ψ is partial X-recursive, then there is a c such that $\psi(G(x/X))$, if defined, is less than $G(x+c/X)$.

Theorem 12.

(a) $G(x-c)<\max \#A\ (I_f(A)\leq x)<G(x+c)$

(b) $G(x-c/\phi')<\max \#A\ (I_r(A)\leq x)<G(x+c/\phi')$

(c) $G(x-c/\phi')<\max \#\bar{A}\ (I_r(A)\leq x)<G(x+c/\phi')$

(d) $G(x-c/\phi')<\max \#A\ (I(A)\leq x)<G(x+c/\phi')$

(e) $G(x-c/\phi'')<\max \#\bar{A}\ (I(A)\leq x)<G(x+c/\phi'')$.

Here it is understood that the maximizations are only taken over those cardinalities which are finite.

The proof of (e) is beyond the scope of the method used in this paper; (e) is closely related to the fact that $\{x \mid W_x \text{ is co-finite}\}$ is $\sum_3$-complete [1, p. 328] and to Theorem 16 of [3].

Appendix

Theorem 3b can be strengthened to the following:

$$I(\psi(x)/X)\leq I(x/X)+I(\psi/X)+lg(I(\psi/X)) \\ +lg(lg(I(\psi/X)))+2\ lg(lg(lg(I(\psi/X))))+c\,.$$

There are many other similar inequalities.

To formulate sharp results of this kind it is necessary to abandon the formalism of this paper, in which programs have endmarkers. Instead one must use the self-delimiting program formalism of [9] and [3] in which programs can be concatenated and merged. In that setting the following inequalities are immediate:

$$H(\psi(x)/X)\leq H(x/X)+H(\psi/X)+c,$$
$$H(\lambda x[\psi(\varphi(x))]/X)\leq H(\varphi/X)+H(\psi/X)+c\,.$$

IBM THOMAS J. WATSON RESEARCH CENTER

References

[1] H. Rogers, Jr.: Theory of recursive functions and effective computability, McGraw-Hill, New York, 1967.

[2] G.J. Chaitin: *Information-theoretic limitations of formal systems*, J. ACM. **21** (1974), 403–424.

[3] G.J. Chaitin: *Algorithmic entropy of sets*, Comput. Math. Appl. **2** (1976), 233–245.

[4] T. Kamae: *On Kolmogorov's complexity and information*, Osaka J. Math. **10** (1973), 305–307.

[5] R.P. Daley: *A note on a result of Kamae,* Osaka J. Math. **12** (1975), 283–284.
[6] D.W. Loveland: *A variant of the Kolmogorov concept of complexity,* Information and Control **15** (1969), 510–526.
[7] G.J. Chaitin: *Information-theoretic characterizations of recursive infinite strings,* Theoretical Comput. Sci. **2** (1976), 45–48.
[8] R.M. Solovay: unpublished manuscript on [9] dated May 1975.
[9] G.J. Chaitin: *A theory of program size formally identical to information theory,* J.ACM **22** (1975), 329–340.
[10] C.P. Schnorr: *Optimal enumerations and optimal Gödel numberings,* Math. Systems Theory **8** (1975), 182–191.
[11] G.J. Chaitin: *Algorithmic information theory,* IBM J. Res. Develop. **21** (1977), in press.