ELSEVIER

Theoretical Computer Science 207 (1998) 383–386

Theoretical Computer Science

On a complexity of the formula $(A \vee B) \Rightarrow C$

K.Yu. Gorbunov *,[1]

Institute of New Technologies, Nizhnaya Radishevskaya 10, Moscow, 109004, Russia

Abstract

By the complexity $KF(\Phi)$ of the formula Φ: $(A \vee B) \Rightarrow C$ we mean the minimal length of a program which on input $\langle 0,A\rangle$ outputs C and on input $\langle 1,B\rangle$ outputs C. We prove that there exist words A, B, C such that $KF(\Phi)$ is close to $K(C|A)+K(C|B)$.

Keywords: Algorithmic information theory; Complexity of formulae; Game approach

According to ideas of A.N. Kolmogorov, the problem of estimation of a complexity of a formula of propositional calculus become sensible if we consider each variable as a task of indication of some word. Then each formula also corresponds to a task. For example, the formula $A \Rightarrow B$ corresponds to the following task: "given word A construct word B". It is natural to define complexity of this formula as minimal length of a program which on input A outputs B. Clearly, it is conditional Kolmogorov complexity $K(B|A)$.

In the paper [1] five authors estimated the complexity of the formula $A \Leftrightarrow B$, that is minimal length of a program which on input $\langle 0,A\rangle$ outputs B and on input $\langle 1,B\rangle$ outputs A. They proved that

$$KF(A \Leftrightarrow B) = \max(K(A|B), K(B|A)) + O(\log(|A|+|B|))$$

(we denote the complexity of a formula φ by $KF(\varphi)$, the length of a word X by $|X|$; all logarithms in our paper have base 2).

We consider the formula Φ: $(A \vee B) \Rightarrow C$. Its complexity is minimal length of a program which on input $\langle 0,A\rangle$ outputs C and on input $\langle 1,B\rangle$ outputs C. It is natural to put the question about correlation between the complexity of Φ and conditional complexities $K(C|A)$, $K(C|B)$. In the present paper we prove that unlike the case of the formula $A \Leftrightarrow B$ in our case $KF(\Phi)$ can be close to $K(C|A)+K(C|B)$. Evidently,

* Fax: 7095 915 69 63; e-mail: kgorbunov@int.glasnet.ru.

[1] This paper was supported in part by the INTAS project No 93-0893.

PII S0304-3975(98)00074-7

the estimate $KF(\Phi) \leqslant K(C|A)+K(C|B)+2(\log K(C|A)+\log K(C|B))+const$ is true, so this estimate is sharp enough.

Let us consider an optimal programming system S to be fixed.

Theorem. *$\exists d\, \forall m,n\, \exists A,B,C$ $[K(C|A)\leqslant m+2(\log m+\log n)+d$, $K(C|B)\leqslant n+2(\log m+\log n)+d$, $KF(\Phi)\geqslant n+m-1]$, where the number d depends only on programming system (here m, n are natural numbers; A, B, C are words).*

Proof. We will construct some special programming system S' with corresponding complexity K' such that for some words A,B,C $K'(C|A)\leqslant m+2(\log m+\log n)+d$, $K'(C|B)\leqslant n+2(\log m+\log n)+d$ and there is no program from S which has length $\leqslant n+m-1$ and outputs C on inputs $\langle 0,A\rangle$ and $\langle 1,B\rangle$. To construct S' we must define an enumerable set of triples: $\langle n,x,y\rangle$ where n is number of a program (in lexicographic numeration), x is an input, y is the output of the program on the input x (all three words are in the alphabet $\{0,1\}$). Let us denote by M the set of triples which corresponds to the fixed system S.

We will use the game approach developed by An.A. Muchnik in [2]. Let us describe a game with two players. The players know the numbers m and n. The first player (which is denoted by P_1) enumerates triples of S' as follows. On every move P_1 can ascribe some word y to some word x. The word y differs from all words which are already ascribed to x. If there are exactly $i-1$ words already ascribed to x then $\langle i,x,y\rangle \in S'$, and hence $K'(y|x,m,n)\leqslant \log i+\text{const}$, $K'(y|x)\leqslant \log i+2(\log m+\log n)+\text{const}$. Assume without loss of generality that $n\geqslant m$. Note that there is not a reason for P_1 to ascribe to a word more than 2^n words because the fact that for some x,y $K'(y|x) > n$ does not help to find required triple $\langle A,B,C\rangle$. So, we consider that P_1 can ascribe not more than 2^n words to every word. To accelerate the game we permit him to make finite number of moves in succession.

The second player (which is denoted by P_2) can construct step-by-step $j=2^{m+n-1}$ ordered pairs of functions from words to words: $\langle f_1^1,f_1^2\rangle$, $\langle f_2^1,f_2^2\rangle,\dots,\langle f_j^1,f_j^2\rangle$. On every move P_2 can define finite number of functions on finite number of words.

The game can last infinitely long. The first player wins if after some move of P_2 there exists a triple $\langle A,B,C\rangle$ such that for some $i_1\leqslant 2^m$ and $i_2\leqslant 2^n$ we have: $\langle i_1,A,C\rangle\in S'$, $\langle i_2,B,C\rangle \in S'$ and there is no i such that $f_i^1(A)=C$, $f_i^2(B)=C$. Otherwise, the second player wins.

We claim that if there exists computable winning strategy for the first player then the triple $\langle A,B,C\rangle$ is required. Indeed, let the second player plays according to the system S in the following way. He enumerates ordered quadruples $\langle k,x,y,z\rangle$ such that the program with number $k\leqslant 2^{m+n-1}$ (hence, its length $\leqslant m+n-1$) from S outputs z on inputs $\langle 0,x\rangle$ and $\langle 1,y\rangle$. For each such quadruple P_2 defines: $f_k^1(x)=z$, $f_k^2(y)=z$. P_1 has a winning way of constructing of S'. Since $\forall x K(x)\leqslant K'(x)+const$, this implies existence of the required words A, B, C. So it remains to prove the following lemma.

Lemma. *There exists a computable winning strategy for P_1.*

Proof. Let us represent the game in a more convenient form. Consider horizontal lexicographic sequence of all words where over each word there are 2^n places located on a vertical line. These places are for those words which the first player ascribes to the corresponding word. But it is easy to see that result of the game depends only on facts of *equality* of the words (that is a result does not change if we replace all equal words used in the game by another equal words). This implies the following rules.

At any moment of the game there is a finite number of finite mutually disjoint sets of places, and each set contains on any vertical line not more than one element. The first player can add some place (not belonging to any set) to some set or declare some such place as a new set (the meaning is: words corresponding to places from the same set are equal). The second player can mark elements of the sets by numbers: 1_1, 1_2, 2_1, $2_2, \ldots, 2_1^{m+n-1}$, 2_2^{m+n-1} (the meaning is: the mark k_i for a place y over a word x means that $f_k^i(x) = y$). The requirement is: there are no equal marks on every vertical line. P_1 wins if at some moment after a move of P_2 there exist places x, y from one set for which there is no k such that x is marked by k_1 and y is marked by k_2, and there are no more than 2^m sets on the vertical containing x.

Let us describe a winning strategy for P_1. P_1 puts on the first horizontal a large (we will see from the construction how large) number of two-element sets ordered from left to right. After the next move of P_2 for each set there exists k (otherwise, clearly, P_1 wins) such that the left element is marked by k_1 and the right element is marked by k_2. Among the sets we take many one-type sets (that is with the same k). Let us divide these sets into neigboring pairs: ⟨left set, right set⟩. P_1 adds to each left set the place on the second horizontal over the right element of corresponding right set. These extended sets will be called by step-sets. Applying to the step-sets the above argument (that is consideration of pairs: ⟨the most left element of the step set, the most right its element⟩) we get for each step-set a second number which must differ from k (because by the rules of the game P_2 cannot put k_2 over k_2). Then we take many one-type pairs of the original sets. These pairs we again divide into pairs of the pairs. Then P_1 adds to each step-set the place on the third horizontal over the right element of the right set of correspondind quadruple. Applying to the (extended) step-sets the same argument, we find a third number differing from the first two numbers. Then, in the same way, we consider octuples of the sets, 16-tuples,...,2^{2^n-1}-tuples. Finally, we have on the first horizontal *many* elements (the most left from the 2^{2^n-1}-tuples) marked by the same 2^n different numbers (with the index 1), and there are no other sets on the corresponding vertical lines. The next (second) stage of the game will run only on these verticals. It runs in the same way as the first stage, but the lowest horizontal is the second now. After the second stage we have on the second horizontal many elements marked by the same $2^n - 1$ numbers differing from the 2^n numbers on the first horizontal. Then, in the same way, we run the third stage (on verticals defined by the elements of the second stage), then stages with numbers $4, 5, \ldots, 2^m$ (after each stage height of the game decreases by 1 and the set of used verticals becomes more sparse).

It is easy to see that if the second player marks the elements of all considered pairs (according to the rules of the game), then he has to use not less than

$2^n + (2^n - 1) + (2^n - 2) + \cdots + (2^n - (2^m - 1)) = 2^n 2^m - (1 + 2 + \cdots + (2^m - 1)) = 2^{n+m} - (2^m - 1)2^{m-1} > 2^{n+m-1}$ numbers (since $n \geqslant m$) that is more than he has. Note that we considered only the pairs of places such that the first place has height not more than 2^m (height of the second, of course, $\leqslant 2^n$). Therefore, whenever P_2 refuses to mark the elements of such a pair, P_1 wins. The lemma and the theorem are proved. $\square$

It would be interesting to determine a sharp lower estimate on horizontal zone of the game, that is on number of words used by the first player. In particular, the author does not know an answer to the following question. Is it true that $\exists \alpha > 1 \, \forall \varepsilon > 0 \, \exists n_0 \, \forall n \geqslant n_0$ [the first player has a winning strategy (for $m = n$) on horizontal zone $\exp(\exp(\varepsilon n))$ (here $\exp(x) = 2^x$) in the game where the second player can mark places by $2^{\alpha n}$ pairs of numbers]? (vertical zone, as earlier, is equal to 2^n). It is not difficult to prove that if the answer is positive then the following is true:

$$\forall d \, \exists \beta > 1 \, \exists n_0 \, \forall n \geqslant n_0 \, \exists A, B, C \; [K(C|A) \leqslant n, \; K(C|B) \leqslant n$$

and

$$KF(\Phi) \geqslant \beta n + d(\log(K(A)) + \log(K(B)))].$$

Acknowledgements

The author is thankful to An.A. Muchnik, who drew author's attention to the considered problem and told the author about the game approach.

References

[1] C.H. Bennett, P. Gács, M. Li, P.M.B. Vitányi, W.H. Zurik. Information distance and thermodynamics of computation, Proc. 25-th ACM symp. theory of computers. 1993, pp. 21–30.

[2] An.A. Muchnik. On the main structures of descriptive theory of algorithms, Dokl. Akad. Nauk SSSR 285 (2), (1985) pp. 280–281 (in Russian).