Monotone complexity of a pair

Pavel Karpovich (Moscow State University)

CSR2010

Kolmogorov complexity

- Decompressor D - a program that reads input binary word and prints output binary word
- Binary word p is a description of q with respect to decompressor D if $D(p) = q$
- Complexity of a word q with respect to D is

$$K_D(q) = min\{|p||D(p) = q\}$$

Kolmogorov complexity

- Decompressor D - a program that reads input binary word and prints output binary word
- Binary word p is a description of q with respect to decompressor D if $D(p) = q$
- Complexity of a word q with respect to D is

$$K_D(q) = min\{|p||D(p) = q\}$$

Theorem. There is a universal decompressor U such that for every other decompressor D there is a constant c_D such that

$$K_U(q) \leq K_D(q) + c_D.$$

for every q.

Different complexities

- $KS(q)$ - Kolmogorov complexity. Decompressor reads finite binary word (with explicit delimiter), prints finite binary word and stops.

Different complexities

- $KS(q)$ - Kolmogorov complexity. Decompressor reads finite binary word (with explicit delimiter), prints finite binary word and stops.
- $KP(q)$ - Prefix complexity. Decompressor reads infinite binary sequence on the tape (deciding where to stop by itself), prints finite binary word and stops.

Different complexities

- $KS(q)$ - Kolmogorov complexity. Decompressor reads finite binary word (with explicit delimiter), prints finite binary word and stops.

- $KP(q)$ - Prefix complexity. Decompressor reads infinite binary sequence on the tape (deciding where to stop by itself), prints finite binary word and stops.

- $KM(q)$ - Monotone complexity. Decompressor reads infinite sequence on the tape, print bits to infinite output tape sequentially; we measure how many input bits were read before the desired word q appears on the output tape.

Different complexities

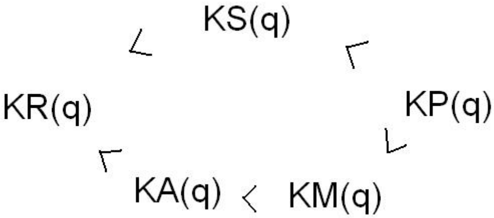
- $KS(q)$ - Kolmogorov complexity. Decompressor reads finite binary word (with explicit delimiter), prints finite binary word and stops.
- $KP(q)$ - Prefix complexity. Decompressor reads infinite binary sequence on the tape (deciding where to stop by itself), prints finite binary word and stops.
- $KM(q)$ - Monotone complexity. Decompressor reads infinite sequence on the tape, print bits to infinite output tape sequentially; we measure how many input bits were read before the desired word q appears on the output tape.
- $KR(q)$ - Decision complexity. Decompressor read finite binary word with explicit delimiter, prints bits to infinite output tape sequentially (the output should start with the desired word q).

Different complexities

$KA(q)$ - A priori complexity. Decompressor is applied to random input (or, equivalently, has not input but has access to random bit generator $($rand()). Prints bits to the infinite output tape sequentially. We measure that probability of the event "q is a prefix of the output".

A priori complexity of word q with respect to D is $-log(P_q)$, where P_q is the probability to get word q as a prefix of the output sequence
Note that the probability of this event is at least $2^{-|p|}$ where p is the shortest preimage of q.

Inequalities for complexities



$KR(q) < KA(q) \Leftrightarrow KR(q) \leq KA(q) + c, KA(q) - KR(q)$ is unbounded

Difference $KM(q) - KA(q)$

It is known that the difference $KM(q) - KA(q)$ can be as large as $log(log(|q|))$. First, Peter Gács proved (1983) that the difference between $KM(q)$ and $KA(q)$ is not bounded. The lower bound was improved recently to $log(log(|q|))$ by Adam Day (2009). Upper bound:

$$KM(q) < KA(q) + (1 + \epsilon) log(KA(q)) + c$$

Complexity of pairs

Our goal is to investigate the difference between KA and KM for pairs; it turns out that this case is much simpler and we get (almost) tight bounds.
It's possible to generalize definitions of complexities to pairs.

Decompressor still has one input tape, but has two output tapes instead of one. For plain and prefix complexity (where the exact answer is needed on the output tape) it does not make much sense (we can replace two tapes by one using encoding), but for monotone, a priori and decision complexities it gives a new notion.

A priori compexity of pairs

- $$KA(q) \leq |q| + O(1)$$
- $$KA(q,p) \leq |q| + |p| + O(1)$$

Indeed, the decompressor may just send random bits to output tapes (independently for different tapes)

Upper bounds for monotone complexity

- $$KM(q) \leq |q| + c$$

- $$KM(q,p) \leq |q| + |p| + log(|q| + |p|) + 2log(log(|q| + |p|)) + c$$

Question: can the logarithmic term be avoided?

Main result

Theorem. For each $\alpha < 1$ and constant c there is a pair $\langle q, p \rangle$, such that the following inequality holds :

$$KM(q,p) > |q| + |p| + \alpha \cdot log(|q| + |p|) + c$$

Corollary. The difference $KM(q,p) - KA(q,p)$ can be as large as $(1 - \epsilon) log(|q| + |p|)$.

For upper bound the following inequality holds:

$$KM(q,p) < KA(q,p) + (1 + \epsilon) log(|q| + |p|) + c$$

Upper bounds for decision complexity

Observation.$KR(q) \leq |q| + c_1$; $KR(q,p) \leq |q| + |p| + c$.

Proof.

Let D will be decompressor which reads a binary word s from input tape, print s on first tape and print s^R (reverted word) on second tape. Description of a pair $< q, p >$ will be a concatenation of q and p^R. For this decompressor we have

$$KR(q,p) \leq KR_D(q,p) + c = |q| + |p| + c$$

Upper bounds for decision complexity

Observation.$KR(q) \leq |q| + c_1$; $KR(q,p) \leq |q| + |p| + c$.

Proof.

Let D will be decompressor which reads a binary word s from input tape, print s on first tape and print s^R (reverted word) on second tape. Description of a pair $< q, p >$ will be a concatenation of q and p^R. For this decompressor we have

$$KR(q,p) \leq KR_D(q,p) + c = |q| + |p| + c$$

Theorem. $KR(q,p,r) \leq |q| + |p| + |r| + c$.

Combinatorial lemma

Lemma. There is a set $A = \{a_1, .., a_n, b_1, .., b_n, c_1, .., c_n\}$ of $3n$ binary vectors in linear space F_2^n such that every subset $B = \{a_1, .., a_p, b_1, .., b_q, c_1, .., c_r\}$ of A with $p + q + r = n$ $(p \geq 0, q \geq 0, r \geq 0)$ is a linearly independent set.

$$\begin{array}{ccc}
a_n & b_n & c_n \\
... & ... & ... \\
... & b_q & ... \\
... & * & ... \\
a_p & * & ... \\
* & * & ... \\
* & * & c_r \\
* & * & * \\
a_1 & b_1 & c_1
\end{array}$$

Combinatorial lemma

Construction. (This simplified construction was suggested by Ilya Razenshteyn.) Let $e_1, ..., e_n$ is some basis of F_2^n. Let

$$a_i = e_i, b_i = e_{n-i}, c_i = Pe_i,$$

where P is a matrix of Pascal triangle ($p_{ij} = C_j^i \; mod \, 2$).

Lemma doesn't generalize to 4 columns. Open question: is the inequality true for quadruples ?

$$KR(p, q, r, s) \leq |p| + |q| + |r| + |s| + c$$

Thank you.