Monotone complexity of a pair

Pavel Karpovich (MSU, Moscow)

June 2009

Upper bounds for plain complexity

$$C(x) \leq |x| + c$$

where $C(x)$ is Kolmogorov complexity of x and $|x|$ is the length of x.

FACT.

- $$C(x,y) \leq C(x) + C(y) + log(C(x)) + 2log(log(C(x))) + c$$
  $$C(x,y) \leq |x| + |y| + log(|x|) + 2log(log(|x|)) + c$$
- $$\exists x, y : C(x,y) \geq |x| + |y| + log(|x|) - c$$

Monotone description modes

Let $x \preceq x'$ means that x is the prefix of x' (it's possible $x = x'$).

Monotone description mode is an enumerable set W of pairs of binary strings such that:

1. if $\langle p, x\rangle \in W$ and $x' \preceq x$, then $\langle p, x'\rangle \in W$.
2. if $\langle p, x\rangle \in W$ and $\langle p, x'\rangle \in W$, then $x \preceq x'$ or $x' \preceq x$.
3. if $\langle p, x\rangle \in W$ and $p \preceq p'$, then $\langle p', x\rangle \in W$.

If $\langle p, x\rangle \in W$, then p is called a description for the string x with respect to W.
The monotone complexity $K_m^W(x)$ of x is the length of the shortest description for x with respect to W.

Monotone complexity

There is an optimal monotone description mode S such that:

$$K_m^S(x) \leq K_m^W(x) + C^W$$

for every monotone description mode W and a binary string x.
$K_m^S(x)$ is called *monotone Kolmogorov complexity* of the string x.

Monotonicity: if a binary string x is a prefix of another string x',

$$K_m(x) \leq K_m(x')$$

Monotone Kolmogorov complexity of a pair

We can generalize the definition of monotone complexity to pairs.

A monotone description mode for pairs is a pair of enumerable sets W_1 and W_2; each of them is a monotone description mode (as defined earlier).

The complexity $K_m^{W_1,W_2}(x,y)$ of a pair of binary strings x and y is the length of the shortest string p, such that $\langle p,x\rangle \in W_1$ and $\langle p,y\rangle \in W_2$ (i.e., p describes x with respect to W_1 and p describes y with respect to W_2).

There is an optimal description mode for pairs and we can define monotone complexity of a pair, denoted by $K_m(x,y)$.

Upper bounds for monotone complexity

- $$K_m(x) \leq |x| + c$$

- $$K_m(x, y) \leq |x| + |y| + log(|x| + |y|) + 2log(log(|x| + |y|)) + C$$

Question: can the logarithmic term be avoided?

Main result

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

$$K_m(x, y) > |x| + |y| + \alpha \cdot log(|x| + |y|) + c$$

Decision complexity

Decision description mode is an enumerable set W of pairs of binary strings such that:

1. if $\langle p, x\rangle \in W$ and $x' \preceq x$, then $\langle p, x'\rangle \in W$.
2. if $\langle p, x\rangle \in W$ and $\langle p, x'\rangle \in W$, then $x \preceq x'$ or $x' \preceq x$.

The decision complexity $KR^W(x)$ of x is the length of the shortest description for x with respect to W.

There is an optimal decision description mode and we can define decision complexity, denoted by $KR(x)$.

Decision complexity of a pair

A decision description mode for pairs is a pair of enumerable sets W_1 and W_2(we need three modes for triples); each of them is a decision description mode.

The complexity $KR^{W_1,W_2}(x,y)$ of a pair of binary strings x and y is the length of the shortest string p, such that $\langle p,x\rangle \in W_1$ and $\langle p,y\rangle \in W_2$

There is an optimal description mode for pairs and we can define decision complexity of a pair, denoted by $KR(x,y)$(and decision complexity of a triple $KR(x,y,z)$).

Upper bounds for decision complexity

FACT. $KR(x, y) \leq |x| + |y| + c$.

Proof. A word p is a description of a pair $\langle p, p^R \rangle$, where p^R is the reversed string p. A description of a pair $\langle x, y \rangle$ will be a string t, it is a concatenation of the string x and the reversed string y $(|x| + |y| = |t|)$.

Theorem. $KR(x, y, z) \leq |x| + |y| + |z| + c$.

Combinatorial lemma

Lemma. There is a set $A = \{a_1, .., a_n, b_1, .., b_n, c_1, .., c_n\}$ from $3n$ 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 linear independent set.

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

Question. Is it true for quadruples ?

$$KR(x, y, z, w) \leq |x| + |y| + |z| + |w| + c$$

A priori complexity

A lower semicomputable semimeasure on a binary tree is a lower semicomputable function $a(x)$ (x is a binary string) with the following properties :

1. $a(x) \geq 0$
2. $a(\lambda) = 1$
3. $a(x) \geq a(x0) + a(x1)$

There is universal lower semicomputable semimeasure $\mu(x)$. It holds:

$$a(x) \leq c_a \mu(x)$$

for every lower semicomputable semimeasure a and a binary string x. The value of $\mu(x)$ is called 'a priori probability' of the string x. Minus logarithm of the 'a priori probability' is called 'a priori complexity' $KA(x)$.

A priori complexity for a pair

A lower semicomputable semimeasure on pairs of binary strings is a lower semicomputable function $a(x, y)$ with the following properties :

1. $a(x, y) \geq 0$
2. $a(\lambda, \lambda) = 1$
3. $a(x, y) \geq a(x0, y0) + a(x0, y1) + a(x1, y0) + a(x1, y1)$

There is universal lower semicomputable semimeasure on pairs $\mu(x, y)$. It holds:

$$a(x, y) \leq c_a \mu(x, y)$$

for every lower semicomputable semimeasure on pairs a and a pair $\langle x, y \rangle$. The value of $\mu(x, y)$ is called 'a priori probability' of the pair. Minus logarithm of the 'a priori probability' is called 'a priori complexity' $KA(x, y)$.

Difference $K_m(x) - KA(x)$

For strings (not pairs) it is known that the difference $KM(x) - KA(x)$ can be as large as $O(log(log(|x|)))$. First, Peter Gács proved (1983) that the difference between $KM(x)$ and $KA(x)$ is not bounded. The lower bound was improved recently to $O(log(log(|x|)))$ by Adam Day.

Upper bound:

$$K_m(x) < KA(x) + (1 + \epsilon)log(KA(x)) + c$$

Difference $K_m(x, y) - KA(x, y)$

Remark.

- $KA(x, y) < K_m(x, y) + c$
- $KA(x, y) \leq |x| + |y| + c$ (a function $2^{-|x|-|y|}$ is a lower semicomputable semimeasure)

Corollary. The difference $K_m(x, y) - KA(x, y)$ can be as large as $(1 - \epsilon)log(|x| + |y|)$.

For upper bound the following inequality holds:

$$K_m(x, y) < KA(x, y) + (1 + \epsilon)log(|x| + |y|) + c$$

Thank you.