Rectangles, communication, Kolmogorov complexity and a combinatorial game

following Andrei Romashchenko and Alexander Kozachinskiy

Abstract

A well known result in communication complexity says that the triple information of two inputs and a transcript is non-negative. Romashchenko and Zimand (2019) generalized this result for the case of arbitrary rectangle partitions (both for Shannon entropy and Kolmogorov complexity). We observe that essentially one rectangle is needed for the Kolmogorov complexity case, and extend this modified version to multidimensional case. Then we translate these inequalities into a combinatorial language and prove the existence of winning strategies in some simple combinatorial games. This gives new examples of combinatorial results that can be proved using Kolmogorov complexity. (We do not know a direct combinatorial proof.)

1 Deterministic protocol and random inputs

Consider a communication protocol for Alice and Bob on inputs x and y where x and y form a pair of (jointly distributed) random variables. The transcript (a sequence of messages) is then a function of x and y and so it becomes a random variable $\pi = \pi(x, y)$ on the same space.

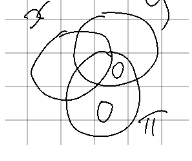
Proposition 1 (Reference???).

$$I(x:y:\pi) \geqslant 0.$$

We assume that protocol is a binary tree where each vertex is a leaf (that determines the answer, but this does not matter for now), or specifies who should send the next bit and has two children (corresponding to bits 0/1 sent). It is possible that after a bit is sent by some participant, the next bit (according to the tree) should be sent by the same participant, so each message is not a single bit but a bit string, and the transcript is some path from a root to a leaf split into alternating messages.

For the definition of a triple information and corresponding Venn-style information diagrams see, e.g., [5, chapter 6].

Proof. We can use induction. The protocol determines who sends the first message; say it is Alice. Then π is a function of x, the information diagram has two zeros, and $I(x:y:\pi) = I(y:\pi) \geqslant 0$.



Induction step. Assume that we know that $I(x:y:\pi_1 \ldots \pi_k) \geqslant 0$ for first k messages, and we want to prove that $I(x:y:\pi_1 \ldots \pi_{k+1}) \geqslant 0$. Note that we have the chain rule $H(ab) = H(a) + H(b|a)$ for entropies, therefore we have chain rule $I(ab:c) = I(a:c) + I(b:c|a)$ for mutual information, and then in the same way we get

$$I(ab:c:d) = I(a:c:d) + I(b:c:d|a).$$

We can apply this to our case:

$$I(\pi_1 \ldots \pi_k \pi_{k+1}:x:y) = I(\pi_1 \ldots \pi_k:x:y) + I(\pi_{k+1}:x:y|\pi_1 \ldots \pi_k).$$

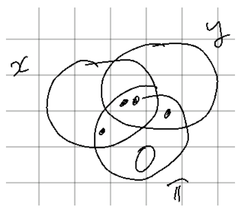
This first term is non-negative by the induction assumption, and the second term is non-negative for the same reasons as before (when $\pi_1 \ldots \pi_k$ is known, π_{k+1} is a function of one of the variables x, y, depending on the sender of $(k+1)$th message).

It remains to note that we may assume that the number of messages in the protocol is the same for all x and y by adding dummy constant messages after the end of protocol: when Alice and Bob see that the path in the tree reaches a leaf, they send special messages $*$ until the required total number of messages is reached. This does not change the distribution of π (just trailing $*$ are added).

An alternative argument [Reference???] goes as follows. Let X and Y be the ranges of random variables x and y. The communication protocol splits $X \times Y$ into combinatorial rectangles that correspond to leaves of the tree. The variable $\pi(x, y)$ is the rectangle that contains (x, y).

The inequality $I(x:y:\pi) \geqslant 0$ can be rewritten as

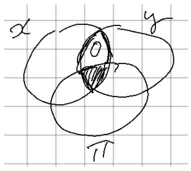
$$H(\pi) \geqslant H(\pi|x) + H(\pi|y). \tag{1}$$



Consider new triple of random variables π, x', y': first we choose π with the same distribution as before, and then choose x' using the conditional distribution of x for the chosen value of π, and y' in a similar way, *independently from choosing* x'. (This is known — in some equivalent version — as the *copy lemma*.) Obviously the distribution of (π, x') is the same as for (π, x), so all information quantities (entropies, conditional entropies, mutual information) for (π, x') are the same as for (π, x), and the same for (π, y') and (π, y). Also by construction x' and y' are independent given π.

What is not so obvious (and very important) is that π is not only a function of (x, y), but also a function of (x', y'). Indeed, if the value of π is some rectangle, then the conditional distribution of x is concentrated in the first side (projection) of the rectangle, so the same is true for x'. Applying the same reasoning to y', we see that (x', y') is in the same rectangle, so π can be reconstructed from (x', y') in the same way as for (x, y).

This argument shows that we can assume without loss of generality that x and y are independent given π when proving (1). But in this case the statement is obvious, since $I(x:y:\pi) = I(x:y) \geqslant 0$.



□

2 Kolmogorov complexity version

The Kolmogorov complexity version of the Proposition 1 can be formulated as follows. Let x and y be two strings, and let $\pi_1, \dots, \pi_k$ be a sequence of strings. Assume that π_1 is simple given x, then π_2 is simple given y and π_1, then π_3 is simple given x, π_1, π_2 and so on. (This corresponds to the setting when the first message is sent by Alice who knows x.) Then $I(x:y:\pi_1, \dots, \pi_k) \geqslant 0$ up to a small error. Here the triple information is understood in the complexity sense (using Kolmogorov complexities instead of entropies). More precisely, the following proposition holds:

Proposition 2 (Romashchenko, Zimand, see section 8.5 in [2]). *Let x, y and $\pi_1, \dots, \pi_k$ be strings of length at most n. Then*

$$I(x:y:\pi_1, \dots, \pi_k) \geqslant -O(\mathrm{C}(\pi_1|x) + \mathrm{C}(\pi_2|\pi_1, y) + \mathrm{C}(\pi_3|\pi_1, \pi_2, x) + \dots) - O(k \log n).$$

Proof. We may repeat the induction proof of Proposition 1 replacing entropies by complexities; all the inequalities that we used for entropies remain true for complexities (they are positive linear combinations of Shannon's inequalities). □

This proposition makes sense only if the number of messages is small, while in communication protocols the number of messages could be of order n, and then the statement becomes vacuous. Also it would be interesting to have a version of the second proof of Proposition 1 adapted to the complexity case. This can be done as follows.

Let x, y be binary strings of length n and let $\Pi \subset \mathbb{B}^n \times \mathbb{B}^n$ be a combinatorial rectangle (a product of two subsets of $\mathbb{B}^n$) containing (x, y). Assume that Π has the following *regularity* property: *Π is simple conditional to each of its points (x', y')*. Then $I(x:y:\Pi)$ is (almost) non-negative.

In general, for a finite set A of finite objects we may introduce the "irregularity parameter"[1]

$$i(A) = \max\{\mathrm{C}(A|x) : \ x \in A\};$$

the smaller $i(A)$ is, the more "regular" is the set A.

Proposition 3 (Romashchenko, Zimand, lemma 4.6 in [2] adapted to one rectangle).

$$I(x:y:\Pi) \geqslant -O(i(\Pi)) - O(\log n)$$

for every combinatorial rectangle $\Pi \subset \mathbb{B}^n \times \mathbb{B}^n$ that contains (x, y).

[1]Strangely, this quantity appeared in quite a different context when considering everywhere complex configurations, see Rumyantsev's paper [3] or its exposition in [4, Theorem 162, p. 251]. The property of having small $i(A)$ can be reformulated as follows: A appears in a simple enumeration of finite sets that covers every point few times. More precisely, $i(A) = O(\log n)$ means that A appears in an enumeration of complexity $O(\log n)$ with multiplicity poly(n).

In the communication complexity setting x and y could be inputs of Alice and Bob (encoded by n-bit strings) and Π could be the set of all input pairs that produce the same transcript as (x, y). The $i(\Pi)$ is bounded by the complexity of the communication protocol (knowing the protocol, we may run it on arbitrary point (x', y') in Π to get the output and then Π).

Proof sketch. We use the same "enforced independence" tool as before, but in a Kolmogorov complexity setting. For that we use the "typization" trick switching from some object s to a set of objects "similar to s" (that includes s itself). For example, a string s of some complexity $m = C(s)$ is an element of the set of all strings that have complexity at most s. (We say "at most s" and not "exactly s" since we will need to enumerate those strings.)

For a pair of strings (u, v) we might consider all pairs (u', v') such that

$$C(u') \leqslant C(u),\ C(v') \leqslant C(v),$$
$$C(u', v') \leqslant C(u, v),$$
$$C(u'|v') \leqslant C(u|v),\ C(v'|u') \leqslant C(v|u).$$

In our case we consider objects similar to x (and separately do the same for y) in the context of a given combinatorial rectangle $\Pi = U \times V$ (which is fixed: we do not consider other combinatorial rectangles). Namely, we consider the set X of all $x' \in U$ such that all the quantities

$$C(x), C(x|\Pi), C(x, \Pi), C(x|\Pi), C(\Pi|x) \tag{2}$$

do not increase when x is replaced by x'. Note that we consider only $x' \in U$ (belonging to the first projection of the rectangle) but use the entire rectangle (and not only U) in these expressions.

Obviously, this set contains x, so it is not empty. On the other hand, its log-size is bounded by $C(x|\Pi)$, since its elements have bounded complexity given Π. The crucial observation is that *this bound is $O(\log n)$-tight*. Indeed, one can enumerate X knowing Π and numerical parameters (complexities), so if there were much less elements in X, the complexity of all its elements (including x) would be much smaller than $C(x|\Pi)$. Now, knowing that X contains about $2^{C(x|\Pi)}$ elements, we conclude that $C(x'|\Pi) = C(x|\Pi)$ with $O(\log n)$-precision for most elements of U, and the fraction can be made arbitrarily close to 1 by adjusting the constant in $O(\log n)$.

This implies that for most $x' \in X$ we have $C(x', \Pi) = C(x, \Pi)$ with $O(\log n)$ precision, since $C(x', \Pi) = C(\Pi) + C(x'|\Pi)$ with the same precision (and we do not change Π). We can also recall that $C(x', \Pi) = C(x') + C(\Pi|x')$, and both summands do not exceed the corresponding quantities for x by construction, so if the sum is (almost) the same, the summands are (almost) unchanged, too.

We conclude that *for most $x' \in X$ all the quantities from* (2) *remain the same with $O(\log n)$ precision when x is replaced by x'*.

Now we can "make x and y independent relative to Π", so to say. We consider a random element $(x', y') \in X \times Y$. With probability close to 1 we have

$$C(x', y'|\Pi) \geqslant \log \#(X \times Y) = \log \#X + \log \#Y = C(x'|\Pi) + C(y'|\Pi)$$

as well as $C(x'|\Pi) = C(x|\Pi)$ and $C(y'|\Pi) = C(y|\Pi)$ with $O(\log n)$-precision, so $I(x' : y'|\Pi) = O(\log n)$ for most $(x', y') \in X \times Y$.

We finish the proof using the same argument as in the second proof of Proposition 1. The expression of $I(x:y:\Pi)$ can be rewritten as

$$C(\Pi) - C(\Pi|x) - C(\Pi|y) + O(C(\Pi|x,y))$$

(we have seen this on the first picture in Section 1 for the case when $C(\Pi|x,y)) = 0$, but in general we just have $O(i(\Pi))$-term, which is absorbed by our error term). All the quantities remain the same when we replace x, y by x', y', and it remains to use the second picture of that Section to prove the required statement. □

Remarks. 1. The initial motivation for this statement was communication complexity, but it is a general statement about points in combinatorial rectangles that just involves the quantity $i(\Pi)$. It is interesting to find some other inequalities that involve this quantity and/or use them in some other contexts.

2. The construction of X and Y is robust in the sense that we can add other enumerable conditions for x' and y', assuming that these conditions have logarithmic complexity (and are true for x, y). For example, we may use the condition saying that the common information in x and y is extractable (to some extent) — does it give something interesting?

3 Combinatorial version

Let us recall our initial goal in Section 2: we wanted to prove that $I(x:y:\pi) \geqslant 0$ for every two inputs x and y of some communication protocol (of small complexity) and the transcript π that corresponds to these two inputs. Transcripts are in one-to-one correspondence with the partition rectangles, and we want to prove for the corresponding rectangle Π that $I(x:y:\Pi) \geqslant 0$. Using that Π is a (simple) function of x and y, we wrote this inequality as

$$C(\Pi) \geqslant C(\Pi|x) + C(\Pi|y).$$

How can we now try to prove this inequality combinatorially? The standard way is to replace statements of the form $C(u) \leqslant k$ (for some u and k) by statements like "u is an element of a simple set of cardinality at most 2^k". The usual problem here is the interpretation of the statement $C(u) + C(v) \leqslant m$ where we have two complexities in the left hand side. It can be first converted into an equivalent statement "for every k, l such that $k + l = m$, either $C(u) \leqslant k$, or $C(v) \leqslant m$" (and then these two inequalities are interpreted combinatorially). In our specific case these considerations motivate the following combinatorial game.

Game: classifying rectangles

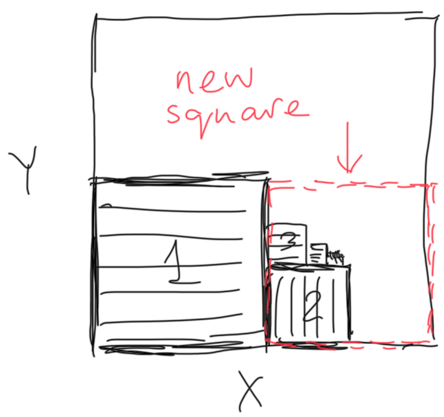
There are two players that alternate. The first player **N** ("Nature") provides disjoint combinatorial rectangles in $\mathbb{Z} \times \mathbb{Z}$. At each move it produces a new rectangle (that does not intersect the previous ones). The second player **M** ("Mathematician") classifies this rectangle: it should declare it as "horizontal" or "vertical". Then **N** provide the next rectangle, etc.

The game has two parameters K and L. The total number of moves (and rectangles) is KL. After KL moves the winner is declared: **M** wins the game if *every horizontal line crosses at most K horizontal rectangles, and every vertical line crosses at most L vertical rectangles.* (Horizontal lines may cross vertical rectangles without problems, and vice versa.)

Of course, we can take any infinite set instead of $\mathbb{Z}$; we use $\mathbb{Z}$ so we can speak about horizontal and vertical lines on the $\mathbb{Z}^2$ -grid.

It is easy to see that if the classification can be postponed to the moment when all the rectangles are known, **M** wins the game easily. Indeed, if there is no horizontal line that crosses more than K rectangles, we may declare all rectangles as horizontal. If some horizontal line α intersects K or more rectangles, we declare all these rectangles as vertical, delete them and reduce the situation to a smaller game: we have at most $KL - K = K(L-1)$ rectangles and use (by induction) the winning strategy for $K, (L-1)$ game: the rectangles that are taken away on the first step, are horizontally disjoint (since some horizontal line crosses them and they are disjoint), so each vertical line can cross at most one of them (and $(L-1)+1 = L$).

However, as noted by Ilya Kondakov and Mikhail Yaroshevich, in the online setting **N** has a winning strategy. Given K and L, **N**ature choose a large enough square and uses one quarter of this square for the first move.



This rectangle is declared to be horizontal or vertical by **M**; for example, let it be horizontal. Then **N** continues the game in the new square (shown on the right of the first square), and uses its quarter for the second move. Note that in this way **N** adds 1 for the number of horizontal intersections in the original game (compared to the game in the new square). Then **N** repeats the trick and uses one quarter of the new square for the second move, etc. If the original square is large enough, **N** guarantees to get more than K horizontal intersections or more than L vertical intersections after $K + L + 1$ moves (much less than KL for large K and L).

So we modify the game assuming that it is played on some finite field $\mathbb{B}^n \times \mathbb{B}^n$ and allowing a polynomial in n factor for the allowed number of horizontal and vertical intersections making the task of **M** easier. For this game, as we will see later, **M** has a winning strategy. But first let us explain how this game is related to the inequality for transcripts.

Proposition 4. *Assume that **M** has a winning strategy in all games of this type (for some polynomial factor). Then for every communication protocol C with n-bit inputs and for every $x, y \in \mathbb{B}^n$ we have*

$$I(x:y:\pi) \geqslant -O(\mathrm{C}(C)) - O(\log n)$$

if π is the transcript of C on inputs x, y.

Proof sketch. As before, with required $O(\log n + \mathrm{C}(C))$ precision we can rewrite the inequality in the form

$$\mathrm{C}(\Pi) \geqslant \mathrm{C}(\Pi|x) + \mathrm{C}(\Pi|y)$$

where Π is the rectangle of C containing (x, y). Assume that it is not the case and choose some k and l such that

$$\mathrm{C}(\Pi) \leqslant k + l, \quad \mathrm{C}(\Pi|x) \geqslant k, \quad \mathrm{C}(\Pi|y) \geqslant l,$$

and all three inequalities are true with some margin of order $O(\log n + \mathrm{C}(C))$. Let us then play the game with parameters $K = 2^k$ and $L = 2^l$: **N** enumerates all combinatorial rectangles of complexity at most $k+l$ that appear in C-partitioning; there are at most $2^{k+l} = KL$ of them (we may omit constants and even polynomial in n factors, since we are interested in $O(\log n)$ precision). We apply the winning strategy of **M** against this strategy of **N**, and get the rectangles classified as horizontal or vertical. In this way we get enumerations of horizontal and vertical rectangles. The complexity of this enumeration is $O(\mathrm{C}(C)) + O(\log n)$, since we need to know C and some numerical parameters, such as n, k, and l. (The complexity of the strategy may be omitted, since we may take the first winning strategy for the game in some natural ordering.)

The rectangle Π should be horizontal or vertical. Assume, for example, that Π is horizontal. Then the horizontal line with height y (as well as any other horizontal line) crosses Π in at most $\mathrm{poly}(n) \cdot K = \mathrm{poly}(n) \cdot 2^k$ points, and Π given y can be specified by a number of Π in the enumeration order (if we consider only rectangles that cross the line), i.e., by at most $k + O(\log n)$ bits. The complexity of the enumeration is $O(\mathrm{C}(C)) + O(\log n)$, so we get a contradiction, assuming that the constant in the margin expression is large enough. □

We do not know a combinatorial proof of the existence of a winning strategy for **M** in the game described (for any polynomial in n factor). However, the existence of such a strategy follows from the inequalities for complexities proven in Proposition 3, as we will show now.

4 Complexity proof of a combinatorial statement

Consider the game played inside $\mathbb{B}^n \times \mathbb{B}^n$. In this game **M** should ensure that each horizontal line crosses at most $\mathrm{poly}(n) \cdot K$ horizontal rectangles, and each vertical line crosses at most $\mathrm{poly}(n) \cdot L$ vertical rectangles.

Proposition 5. *For some polynomial* $\mathrm{poly}(n)$ *and for all* n, K, L *this game has a winning strategy for* ***M***.

Proof sketch. We will derive this result from the corresponding complexity statement, Proposition 3.

Imagine that for some n, K and L the statement is false and **N** has a winning strategy. Since we accept a polynomial margin, we may assume that K and L are powers of 2, i.e., $K = 2^k$, $L = 2^l$. Also we may assume that $KL \leqslant 2^{2n}$ (since this is the bound for the number of disjoint rectangles in $\mathbb{B}^n \times \mathbb{B}^n$). Consider a winning strategy for **N** (we may assume that this strategy has logarithmic complexity, since it can be found by an exhaustive search when n, k, l are known), and let it play with the following strategy for **M**.

When getting a rectangle Π, **M**athematician waits until either

- Π becomes simple with respect to all points in its second projection, i.e., it turns out that $C(\Pi|y) \leqslant l$ for all y in the second projection of Π;
- Π becomes simple with respect to all points in its first projection, i.e., $C(\Pi|x) \leqslant k$ for all x in the first projection.

In the first case, the rectangle is declared as horizontal, in the second case it is declared as vertical. (Ties are resolved arbitrarily.)

If neither of these events happens (i.e., there are $(x, y) \in \Pi$ such that $C(\Pi|x) > k$ and $C(\Pi|y) > l$), **M**athematician never makes her move and loses the game. But in any case every horizontal line (for every y) has at most L intersections with horizontal rectangles (they are all simple given the coordinate of this line), and every vertical line has at most K intersections with vertical rectangles. Therefore, since **N** uses the winning strategy, the game should come to the point where **M** loses because she does not reply to some rectangle Π produced by **N**.

What do we know about this rectangle? First, its complexity $C(\Pi)$ is $k + l + O(\log n)$, since Π is determined by the numerical parameters and the number of move (which does not exceed $KL = 2^{k+l}$). Second, since **M** never replied, there is a point $(x, y) \in \Pi$ such that $C(\Pi|x) > k$ and $C(\Pi|y) > l$. This almost contradicts the statement of Proposition 3. Indeed, $C(\Pi|(x, y))$ is $O(\log n)$, since knowing (x, y) we can just watch the game until the rectangle covering (x, y) appears. The same argument shows also that $i(\Pi) = O(\log n)$. Therefore, the inequality of Proposition 3 says that $C(\Pi)$ should be at least $C(\Pi|x) + C(\Pi|y) + O(\log n)$. At the same time, $C(\Pi) \leqslant k + l + O(\log n)$ while $C(\Pi|x) > k$ and $C(\Pi|y) > l$. To get a contradiction, we should just move the thresholds in the definition of strategy by $c \log n$ for large enough c (making them $k + c \log n$ and $l + c \log n$), keeping the rest of the argument unchanged. □

Remarks. 1. The argument above in fact does not require that Π is determined uniquely by the strategy and (x, y); it is enough to have polynomially many candidates for Π given (x, y). So the proof gives us a stronger statement: we may consider a game where **N** has more freedom and is not obliged to produce disjoint rectangles, it is enough that every point is covered by a polynomial number of **N**'s rectangles. This game is a direct counterpart of the complexity statement of Proposition 3.

2. How all these statements can be adapted to the case of communication complexity with public and private randomness?

5 Multidimensional case

Now we forget about the initial communication complexity motivation, and look just at the inequality

$$C(\Pi) \geqslant C(\Pi|x) + C(\Pi|y) - O(\log n + i(\Pi))$$

that holds for every combinatorial rectangle $\Pi \subset \mathbb{B}^n \times \mathbb{B}^n$ and every point (x, y) in it. What is its natural generalization for three (or more) dimensions? We cannot hope that

$$C(\Pi) \geqslant C(\Pi|x) + C(\Pi|y) + C(\Pi|z) - O(\log n + i(\Pi))$$

for every parallelepiped Π and every point (x, y, z) in it: if Π is just one point, the value $i(\Pi)$ is zero, and it would give $C(x, y, z) \geqslant C(y, z|x) + C(x, z|y) + C(y, z|x)$

which is false for independent random x, y, z. But we can prove the following weaker inequality.

Proposition 6.

$$C(\Pi) \geqslant C(\Pi|x, y) + C(\Pi|y, z) + C(\Pi|x, z) - O(\log n + i(\Pi))$$

for every point (x, y, z) in every combinatorial parallelepiped $\Pi \subset \mathbb{B}^n \times \mathbb{B}^n \times \mathbb{B}^n$.

Proof sketch. First, we may combine two coordinates into one and use Proposition 3. (We get a rectangle in $\mathbb{B}^n \times \mathbb{B}^{2n}$, but this does not create any problems.). In this way we get

$$C(\Pi) \geqslant C(\Pi|x) + C(\Pi|y, z) - O(\log n + i(\Pi))$$

It remains to show that

$$C(\Pi|x) \geqslant C(\Pi|x, y) + C(\Pi|x, z) \tag{3}$$

(with the same error term). The proof follows the two-dimensional case, treating x as a kind of oracle. More precisely, we keep $\Pi = X \times Y \times Z$ and x fixed and construct the set Y' of all $y' \in Y$ such that all the complexities (for objects, pairs, triples with every possible combination of conditions) that include Π, x, y do not increase when y is replaced by y'. As before, we note that the log-size of Y' cannot be much smaller than $C(y|\Pi, x)$, and for a uniformly random $y' \in Y'$ with high probability we have $C(y'|\Pi, x) = C(y|\Pi, x)$, and therefore $C(\Pi, x, y) = C(\Pi, x, y')$ (with required precision). As before, we conclude that all the informational quantities related to Π, x, y remain the same. For example, if we look at $C(\Pi|y)$, we may write

$$C(\Pi, x, y) = C(y) + C(\Pi|y) + C(x|\Pi, y);$$

when y is replaced by y', the sum remains the same and the summands can only decrease, so they also remain the same.

We repeat this construction for z and get the set Z'. Then we note that

$$C(\Pi|x, y) + C(\Pi|x, z)$$

does not change (much) when we replace our given y and z by (uniformly) random pair (y', z') in $Y' \times Z'$. With high probability this pair has complexity (given Π, x) equal to log-size of $Y' \times Z'$, i.e., $C(y|\Pi, x) + C(z|\Pi, x)$, so y' and z' are with high probability independent random strings with respect to Π, x. So it remains to prove the required inequality (3) with additional assumption that $I(y:z|\Pi, x) = 0$, and this can be done as before in Section 3, we just add x as a condition everywhere. □

Remark. Note that in this proof we use the artificial independence trick twice in a different ways. First, we make x and yz independent given Π, and prove the inequality $C(\Pi) \geqslant C(\Pi|x) + C(\Pi|yz)$. Then we make y and z independent given Π, x (and this is a different thing) to prove that $C(\Pi|x) \geqslant C(\Pi|x, y) + C(\Pi|x, z)$.

Proposition 6 (as before) can be used to prove the existence of a winning strategy in a game, now in three dimensions. The game has three parameters K_1, K_2, K_3. **N**ature provides (one by one) $K_1K_2K_3$ disjoint combinatorial parallelepipeds in $\mathbb{B}^n \times \mathbb{B}^n \times \mathbb{B}^n$. When a next parallellepiped is given, **M** classifies it into one of three categories, corresponding to three different coordinate axes. The winning condition is that for every (one-dimensional) line parallel to some coordinate axis i, the number of rectangles of type i that intersect this line is bounded by $\text{poly}(n) \cdot K_i$. Repeating the proof of Proposition 5, we get the following result:

Proposition 7. *There is some polynomial* poly(n) *such that for every n and every K_1, K_2, K_3 **Mathematician** has a winning strategy in this game.*

As before, we do not really need the rectangles produced by **N** to be disjoint. It is enough to assume that every point in $\mathbb{B}^n \times \mathbb{B}^n \times \mathbb{B}^n$ is covered by polynomially many rectangles (and this polynomial in n is fixed in advance, before we choose the polynomials in Proposition 7).

Remarks. 1. The same argument can be extended to arbitrary dimension (though the constant in $O(\cdot)$-notation in this argument grows rather fast with the dimension). For that, having dimension k, we first apply the $(k-1)$-dimensional statement grouping two strings x, y into one. Then we finish the argument by making x and y independent gives Π and all other inputs. In total, for dimension k we need to apply artificial independence trick $k-1$ times.

2. For three dimensions the winning strategy for **M** is not clear even for the case when all parallelepipeds are given in advance (the induction argument does not work as is).

We may also keep individual strings as conditions, but then we have to increase the coefficient in the left hand side. For a singleton parallelepiped we have

$$2\,\mathrm{C}(x,y,z) \geqslant \mathrm{C}(y,z|x) + \mathrm{C}(x,z|y) + \mathrm{C}(x,y|z) - O(\log n)$$

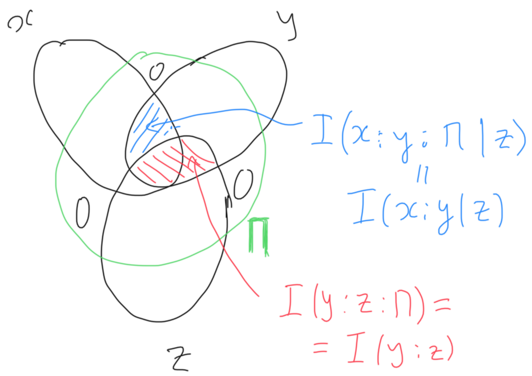
for strings x, y, z of length at most n. We can rewrite the right hand side as $\mathrm{C}(x,y,z|x)+\mathrm{C}(x,y,z|y)+\mathrm{C}(x,y,z|z)$ and note that $\mathrm{C}(x,y,z|x) = \mathrm{C}(x,y,z)-\mathrm{C}(x)$, etc. This reduces the inequality to $\mathrm{C}(x,y,z) \leqslant \mathrm{C}(x)+\mathrm{C}(y)+\mathrm{C}(z)$ and finishes the proof for the singleton case.

For the general parallelepiped $\Pi \subset \mathbb{B}^n \times \mathbb{B}^n \times \mathbb{B}^n$ and a point (x,y,z) in Π we get the following inequality:

Proposition 8.

$$2\,\mathrm{C}(\Pi) \geqslant \mathrm{C}(\Pi|x) + \mathrm{C}(\Pi|y) + \mathrm{C}(\Pi|z) - O(i(\Pi) + \log n)$$

Proof. We start by making x, y, z independent with respect to Π keeping unchanged all the information quantities that involve Π and x, or Π and y, or Π and z. It is done as usual, so we may assume that x, y, z are independent (not only pairwise-independent!) given Π. This allows us to draw the simplified version of information diagram for Π, x, y, z where parts of x, y and z outside Π are disjoint, something like this:



Then we can see what is the difference between both sides, and represent it as a sum of regions of two types shown in the picture — and both are positive. But it is better to use a simple computation (that will work for arbitrary dimension, too). We need to prove that

$$2\,\mathrm{C}(\Pi) \geqslant \mathrm{C}(\Pi|x) + \mathrm{C}(\Pi|y) + \mathrm{C}(\Pi|z)$$

(we omit the error terms), or, after adding one more $\mathrm{C}(\Pi)$,

$$3\,\mathrm{C}(\Pi) \geqslant \mathrm{C}(\Pi|x) + \mathrm{C}(\Pi|y) + \mathrm{C}(\Pi|z) + \mathrm{C}(\Pi).$$

Using that $\mathrm{C}(\Pi|x) = \mathrm{C}(\Pi, x) - \mathrm{C}(x)$, we rewrite this inequality as

$$3\,\mathrm{C}(\Pi) + \mathrm{C}(x) + \mathrm{C}(y) + \mathrm{C}(z) \geqslant \mathrm{C}(\Pi, x) + \mathrm{C}(\Pi, y) + \mathrm{C}(\Pi, z) + \mathrm{C}(\Pi),$$

or

$$\mathrm{C}(x) + \mathrm{C}(y) + \mathrm{C}(z) \geqslant \mathrm{C}(x|\Pi) + \mathrm{C}(y|\Pi) + \mathrm{C}(z|\Pi) + \mathrm{C}(\Pi).$$

Now we recall our assumption: x, y, z are independent given Π. Then we can rewrite the right hand side:

$$\mathrm{C}(x) + \mathrm{C}(y) + \mathrm{C}(z) \geqslant \mathrm{C}(x, y, z|\Pi) + \mathrm{C}(\Pi),$$

or

$$\mathrm{C}(x) + \mathrm{C}(y) + \mathrm{C}(z) \geqslant \mathrm{C}(x, y, z, \Pi).$$

It remains to recall that x, y, z determine Π (i.e., $\mathrm{C}(\Pi|x, y, z)$ is small; it does not exceed $i(\Pi)$ by definition) and that

$$\mathrm{C}(x) + \mathrm{C}(y) + \mathrm{C}(z) \geqslant \mathrm{C}(x, y, z).$$

□

Remarks. 1. A similar statement (with similar proof) is true for any dimension k:

$$(k-1)\,\mathrm{C}(\Pi) \geqslant \mathrm{C}(\Pi|x_1) + \mathrm{C}(\Pi|x_2) + \ldots + \mathrm{C}(\Pi|x_k) - O(i(\Pi) + \log n)$$

for every parallelepiped $\Pi \subset \mathbb{B}^n \times \ldots \times \mathbb{B}^n$ and every its point $(x_1, \ldots, x_k)$. (Unlike in the previous inequality, we use only one artificial independence step.)

2. The inequality of Proposition 8 also has combinatorial meaning. Now **N** provides $K_1K_2K_3$ parallelepipeds, and **M** classifies them in three groups corresponding to coordinate *planes* (instead of lines), in such a way that the number of parallelepipeds intersecting some plane parallel to the coordinate plane does not exceed corresponding K_i^2 (up to some factor polynomial in n). How to prove this statement combinatorially?

References

[1] T. Kaced, A. Romashchenko, N. Vereshchagin, A Conditional Information Inequality and its Combinatorial Applications, *IEEE Transactions on Information Theory*, **64**(5), 3610–3615.

[2] A. Romashchenko, M. Zimand, An Operational Characterization of Mutual Information in Algorithmic Information Theory, *Journal of the ACM*, **66**(5), 1–42 (2019), https://dl.acm.org/doi/10.1145/3356867, Lemma 4.6

[3] A. Rumyantsev, Kolmogorov complexity, Lovasz Local Lemma and critical exponents, *Proc. 2nd Computer Science in Russia Conference* (CSR 2007), Lecture Notes in Computer Science, **4649**, 349–355, see also https://arxiv.org/abs/1009.4995

[4] A. Shen, V.A. Uspensky, N. Vereshchagin, *Kolmogorov complexity and algorithmic randomness*, American Mathematical Society (Mathematical Surveys and Monographs, volume 220), https://www.lirmm.fr/~ashen/kolmbook-eng-scan.pdf

[5] Raymond W. Yeung, *A First Course in Information Theory*, Kluwer, 2002, see also https://iest2.ie.cuhk.edu.hk/~whyeung/tempo/main.pdf