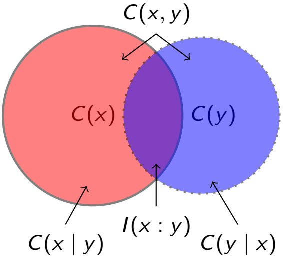
An operational characterization of mutual information in algorithmic information theory

Andrei Romashchenko   Marius Zimand

Univ Montpellier, CNRS, Montpellier   Towson University

ICALP, Prague, July 10, 2018

Two strings x, y, and the information therein



$C(x)$ = length of a shortest description of x.
$C(x \mid y)$ = length of a shortest description of x given y.

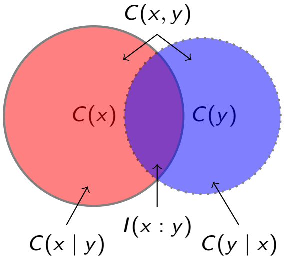
⋮

Mutual information of x and y is defined by a formula:
$I(x:y) = C(x) + C(y) - C(x,y)$.
Also, $I(x:y) =^+ C(x) - C(x \mid y)$,
$I(x:y) =^+ C(y) - C(y \mid x)$

($=^+$ hides $\pm O(\log n)$)

All the regions except the center have an operational meaning.

Two strings x, y, and the information therein



$C(x)$ = length of a shortest description of x.
$C(x \mid y)$ = length of a shortest description of x given y.

⋮

Mutual information of x and y is defined by a formula:
$I(x : y) = C(x) + C(y) - C(x, y)$.
Also, $I(x : y) =^{+} C(x) - C(x \mid y)$,
$I(x : y) =^{+} C(y) - C(y \mid x)$

($=^{+}$ hides $\pm O(\log n)$)

All the regions except the center have an operational meaning.

Does $I(x : y)$ have an operational meaning?

This work in one slide

- Question: Can mutual information be “materialized”?

This work in one slide

- Question: Can mutual information be "materialized"?
- Answer: YES.

This work in one slide

- Question: Can mutual information be "materialized"?
- Answer: YES.
- Mutual information of strings x, y = length of the longest shared secret key that Alice having x and Bob having y can establish via a randomized protocol.

This work in one slide

- Question: Can mutual information be "materialized"?
- Answer: YES.
- Mutual information of strings x, y = length of the longest shared secret key that Alice having x and Bob having y can establish via a randomized protocol.
- This was known in the setting of Information Theory (Shannon entropy, etc.) for memoryless and stationary ergodic sources.
- We show it in the framework of Kolmogorov complexity (it has been an open folklore question since the '70s).

This work in one slide

- Question: Can mutual information be "materialized"?
- Answer: YES.
- Mutual information of strings x, y = length of the longest shared secret key that Alice having x and Bob having y can establish via a randomized protocol.
- This was known in the setting of Information Theory (Shannon entropy, etc.) for memoryless and stationary ergodic sources.
- We show it in the framework of Kolmogorov complexity (it has been an open folklore question since the '70s).
- We also have analog results for multiparty secret key agreement protocols.
- We present matching upper/lower bounds for the communication complexity of 2-party secret key agreement protocols, in the public randomness model.

IT vs. AIT

IT (à la Shannon)

- Data is the realization of a random variable X.
- The model: a stochastic process generates the data.
- Amount of information in the data: $H(X)$ (Shannon entropy).

AIT (Kolmogorov complexity)

- Data is just an individual string x
- There is no generative model.
- Amount of information in the data: $C(x) =$ minimum description length.

IT vs. AIT

IT (à la Shannon)

0000000000000000

- Data is the realization of a … X.
- The model: a stochastic process generates the data.
- Amount of information in the data: $H(X)$ (Shannon entropy).

AIT (Kolmogorov complexity)

- Data is just an individual string x
- There is no generative model.
- Amount of information in the data: $C(x) =$ minimum description length.

IT vs. AIT

IT (à la Shannon)

- Data is the realization of a random variable X.
- The model: a stochastic process ge[...] the data.
- Amount of information in the data: $H(X)$ (Shannon entropy).

101101000110010

AIT (Kolmogorov complexity)

- Data is just an individual string x
- [...]e is no generative model.
- [...]mount of information in the data: $C(x)$ = minimum description length.

IT vs. AIT

IT (à la Shannon)

- Data is the realization of a random variable X.
- The model: a stochastic process generates the data.
- Amount of information in the data: $H(X)$ (Shannon entropy).

AIT (Kolmogorov complexity)

- Data is just an individual string x
- There is no generative model.
- Amount of information in the data: $C(x) =$ minimum description length.

Kolmogorov complexity

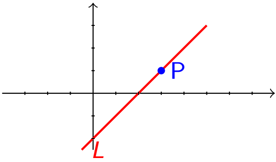
Fix U a universal Turing machine.
p is a description of x if $U(p) = x$. p is a description of x given y if $U(p, y) = x$.

$C(x) = \min\{|p| \mid p \text{ is a description of } x.\}$

$C(x \mid y) = \min\{|p| \mid p \text{ is a description of } x \text{ given } y.\}$

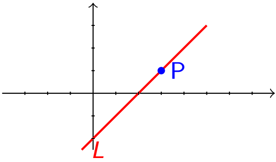
Secret key agreement protocol: warm-up example

- Alice and Bob want to agree on a secret key.
- Problem is that they can only communicate through a public channel.
- Alice knows line L : $y = a_1 x + a_0$;
  Bob knows point P: (b_1, b_2);
- L : $2n$ bits of information (intercept, slope in $\mathbb{F}_{2^n}$).
- P: $2n$ bits of information (the 2 coord. in $\mathbb{F}_{2^n}$).
- Total information in $(L, P) = 3n$ bits; mutual information of L and $P = n$ bits.



Secret key agreement protocol: warm-up example

- Alice and Bob want to agree on a secret key.
- Problem is that they can only communicate through a public channel.
- Alice knows line $L : y = a_1 x + a_0$;
  Bob knows point P: (b_1, b_2);
- L : $2n$ bits of information (intercept, slope in $\mathbb{F}_{2^n}$).
- P: $2n$ bits of information (the 2 coord. in $\mathbb{F}_{2^n}$).
- Total information in $(L, P) = 3n$ bits; mutual information of L and $P = n$ bits.



SOLUTION:

- Alice sends a_1 to Bob.
- Bob, knowing that $P \in L$, finds L.
- Alice and Bob use a_0 as a secret key.
- It works! Eve has seen a_1, but a_1 and a_0 are independent.

Main result (informally stated)

Secret key agreement protocol:

- Alice knows x
- Bob knows y
- they exchange messages and compute a shared secret key z
- z must be random conditioned by the transcript of the protocol

Main result (informally stated)

Secret key agreement protocol:

- Alice knows x
- Bob knows y
- they exchange messages and compute a shared secret key z
- z must be random conditioned by the transcript of the protocol

Our setting:

(1) Alice and Bob also know how their x and y are correlated.

Main result (informally stated)

Secret key agreement protocol:

- Alice knows x
- Bob knows y
- they exchange messages and compute a shared secret key z
- z must be random conditioned by the transcript of the protocol

Our setting:

(1) Alice and Bob also know how their x and y are correlated.
Technically, they know the complexity profile of x and y : $(C(x), C(y), C(x, y))$.

(2) Alice and Bob use randomized algorithms to compute their messages.

Main result (informally stated)

Secret key agreement protocol:

- Alice knows x
- Bob knows y
- they exchange messages and compute a shared secret key z
- z must be random conditioned by the transcript of the protocol

Our setting:

(1) Alice and Bob also know how their x and y are correlated.
Technically, they know the complexity profile of x and y : $(C(x), C(y), C(x, y))$.

(2) Alice and Bob use randomized algorithms to compute their messages.

Theorem (Characterization of the mutual information)

1. *There is a protocol that for every n-bit strings x and y allows to compute with high probability a shared secret key of length $I(x : y)$ (up to $-O(\log n)$).*
2. *No protocol can produce a longer shared secret key (up to $+O(\log n)$).*

Main result (positive part).

Theorem

There exists a secret key agreement protocol with the following property: if

- *Alice knows* x*,* ϵ*, and the complexity profile of* (x, y)*,*
- *Bob knows* y*,* ϵ*, and the complexity profile of* (x, y)*,*

then with probability $1 - \epsilon$ *they obtain a string* z *such that,*

$|z| \geq I(x : y) - O(\log(n/\epsilon))$

and $C(z \mid \text{transcript}) \geq |z| - O(\log(1/\epsilon))$.

Main result (positive part).

Theorem

There exists a secret key agreement protocol with the following property: if

- *Alice knows* x, ϵ, *and the complexity profile of* (x, y),
- *Bob knows* y, ϵ, *and the complexity profile of* (x, y),

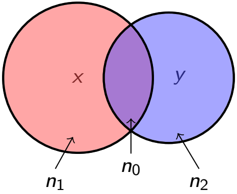
then with probability $1 - \epsilon$ *they obtain a string* z *such that,*

$|z| \geq I(x : y) - O(\log(n/\epsilon))$ /* common key of size $\geq^{+} I(x : y)$ */

and $C(z \mid \text{transcript}) \geq |z| - O(\log(1/\epsilon))$. /* no information leakage */

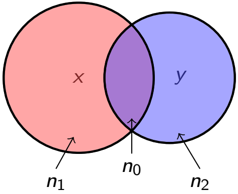
Secret key agreement: sketch of the general protocol

- Alice and Bob want to agree on a secret key.
- they can only communicate through a public channel.
- Alice knows : x; Bob knows a point y;
- $C(x \mid y) =^{+} n_1$
- $C(y \mid x) =^{+} n_2$
- $I(x : y) =^{+} n_0$.



Secret key agreement: sketch of the general protocol

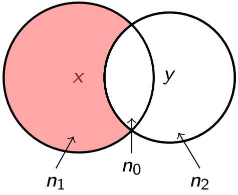
- Alice and Bob want to agree on a secret key.
- they can only communicate through a public channel.
- Alice knows : x; Bob knows a point y;
- $C(x \mid y) =^{+} n_1$
- $C(y \mid x) =^{+} n_2$
- $I(x : y) =^{+} n_0$.



Protocol:

Secret key agreement: sketch of the general protocol

- Alice and Bob want to agree on a secret key.
- they can only communicate through a public channel.
- Alice knows : x; Bob knows a point y;
- $C(x \mid y) =^{+} n_1$
- $C(y \mid x) =^{+} n_2$
- $I(x : y) =^{+} n_0$.

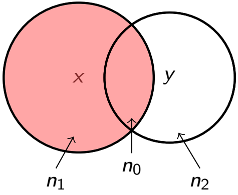


Protocol:

- Alice sends to Bob a random $\text{hash}^{(1)}(x)$ of size $\approx n_1$.

Secret key agreement: sketch of the general protocol

- Alice and Bob want to agree on a secret key.
- they can only communicate through a public channel.
- Alice knows : x; Bob knows a point y;
- $C(x \mid y) =^{+} n_1$
- $C(y \mid x) =^{+} n_2$
- $I(x : y) =^{+} n_0$.

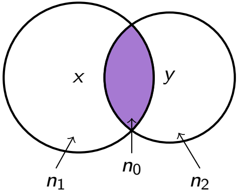


Protocol:

- Alice sends to Bob a random $\text{hash}^{(1)}(x)$ of size $\approx n_1$.
- Bob (knowing y) reconstructs x.

Secret key agreement: sketch of the general protocol

- Alice and Bob want to agree on a secret key.
- they can only communicate through a public channel.
- Alice knows : x; Bob knows a point y;
- $C(x \mid y) =^{+} n_1$
- $C(y \mid x) =^{+} n_2$
- $I(x : y) =^{+} n_0$.

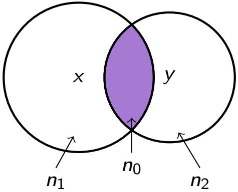


Protocol:

- Alice sends to Bob a random $\text{hash}^{(1)}(x)$ of size $\approx n_1$.
- Bob (knowing y) reconstructs x.
- Alice and Bob compute (independently) a random $z = \text{hash}^{(2)}(x)$ of size $\approx n_0$

Secret key agreement: sketch of the general protocol

- Alice and Bob want to agree on a secret key.
- they can only communicate through a public channel.
- Alice knows : x; Bob knows a point y;
- $C(x \mid y) =^{+} n_1$
- $C(y \mid x) =^{+} n_2$
- $I(x : y) =^{+} n_0$.

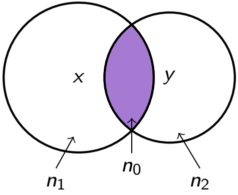


Protocol:

- Alice sends to Bob a random $\text{hash}^{(1)}(x)$ of size $\approx n_1$.
- Bob (knowing y) reconstructs x.
- Alice and Bob compute (independently) a random $z = \text{hash}^{(2)}(x)$ of size $\approx n_0$
- Adversary gets $\text{hash}^{(1)}(x)$ but learns nothing about $\text{hash}^{(2)}(x)$.

Secret key agreement: sketch of the general protocol

- Alice and Bob want to agree on a secret key.
- they can only communicate through a public channel.
- Alice knows : x; Bob knows a point y;
- $C(x \mid y) =^+ n_1$
- $C(y \mid x) =^+ n_2$
- $I(x : y) =^+ n_0$.



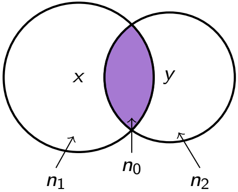
Protocol:

- Alice sends to Bob a random $\mathrm{hash}^{(1)}(x)$ of size $\approx n_1$.
- Bob (knowing y) reconstructs x.
- Alice and Bob compute (independently) a random $z = \mathrm{hash}^{(2)}(x)$ of size $\approx n_0$
- Adversary gets $\mathrm{hash}^{(1)}(x)$ but learns nothing about $\mathrm{hash}^{(2)}(x)$.

Tricky part: choose "communication-efficient" and independent $\mathrm{hash}^{(1)}$ and $\mathrm{hash}^{(2)}$.

Secret key agreement: sketch of the general protocol

- Alice and Bob want to agree on a secret key.
- they can only communicate through a public channel.
- Alice knows : x; Bob knows a point y;
- $C(x \mid y) =^{+} n_1$
- $C(y \mid x) =^{+} n_2$
- $I(x : y) =^{+} n_0$.



Protocol:

- Alice sends to Bob a random $\text{hash}^{(1)}(x)$ of size $\approx n_1$.
- Bob (knowing y) reconstructs x.
- Alice and Bob compute (independently) a random $z = \text{hash}^{(2)}(x)$ of size $\approx n_0$
- Adversary gets $\text{hash}^{(1)}(x)$ but learns nothing about $\text{hash}^{(2)}(x)$.

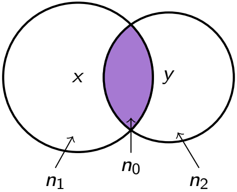
Tricky part: choose "communication-efficient" and independent $\text{hash}^{(1)}$ and $\text{hash}^{(2)}$.

Under the hood:

randomness extractors and universal hashing.

Secret key agreement: sketch of the general protocol

- Alice and Bob want to agree on a secret key.
- they can only communicate through a public channel.
- Alice knows : x; Bob knows a point y;
- $C(x \mid y) =^{+} n_1$
- $C(y \mid x) =^{+} n_2$
- $I(x : y) =^{+} n_0$.



Protocol:

- Alice sends to Bob a
- Bob (knowi
- Alice and) of size $\approx n_0$
- Adversary get).

cf. Buhrman-Fortnow-Laplante 2001,
Musatov-R.-Shen 2009,
Bauwens *et al.* 2013,
Z. 2017

Tricky part: choose "com ation efficient" and independent hash$^{(1)}$ and hash$^{(2)}$.

Under the hood:

randomness extractors and universal hashing.

Main result (negative part).

Theorem

Let x and y be input strings of length n on which the protocol succeeds with error probability ϵ so that with prob $1-\epsilon$ Alice and Bob have at the end the same z, and $C(z \mid t) \geq |z| - \delta(n)$.

Then with probability $\geq 1 - O(\epsilon)$ we have
$|z| \leq I(x:y) + \delta(n) + O(\log(n/\epsilon))$.

Main result (negative part).

Theorem

Let x and y be input strings of length n on which the protocol succeeds with error probability ϵ so that with prob $1-\epsilon$ Alice and Bob have at the end the same z, and $C(z \mid t) \geq |z| - \delta(n)$.

Then with probability $\geq 1 - O(\epsilon)$ we have
$|z| \leq I(x:y) + \delta(n) + O(\log(n/\epsilon))$. */* common key of size $\leq^{+} I(x:y)$ */*

Main result (negative part).

Theorem

Let x and y be input strings of length n on which the protocol succeeds with error probability ϵ so that with prob $1-\epsilon$ Alice and Bob have at the end the same z, and $C(z \mid t) \geq |z| - \delta(n)$.

Then with probability $\geq 1 - O(\epsilon)$ we have
$|z| \leq I(x : y) + \delta(n) + O(\log(n/\epsilon))$.

Under the hood:

Conditional information inequality

Main result (negative part).

Theorem

Let x and y be input strings of length n on which the protocol succeeds with error probability ϵ so that with prob $1-\epsilon$ Alice and Bob have at the end the same z, and $C(z \mid t) \geq |z| - \delta(n)$.

Then with probability $\geq 1 - O(\epsilon)$ we have
$|z| \leq I(x:y) + \delta(n) + O(\log(n/\epsilon))$.

Under the hood:

Conditional information inequality

- simple part: if no communication, then **key** $\leq I(x:y)$

Main result (negative part).

Theorem

Let x and y be input strings of length n on which the protocol succeeds with error probability ϵ so that with prob $1 - \epsilon$ Alice and Bob have at the end the same z, and $C(z \mid t) \geq |z| - \delta(n)$.

Then with probability $\geq 1 - O(\epsilon)$ we have
$|z| \leq I(x : y) + \delta(n) + O(\log(n/\epsilon))$.

Under the hood:

Conditional information inequality

- simple part: if no communication, then **key** $\leq I(x : y)$
- still simple: with communication, **key** $\leq I(x : y \mid \mathrm{transcript})$

Main result (negative part).

Theorem

Let x and y be input strings of length n on which the protocol succeeds with error probability ϵ so that with prob $1-\epsilon$ Alice and Bob have at the end the same z, and $C(z \mid t) \geq |z| - \delta(n)$.

Then with probability $\geq 1 - O(\epsilon)$ we have
$|z| \leq I(x:y) + \delta(n) + O(\log(n/\epsilon))$.

Under the hood:

Conditional information inequality

- simple part: if no communication, then **key** $\leq I(x:y)$
- still simple: with communication, **key** $\leq I(x:y \mid \text{transcript})$
- hard part: $I(x:y \mid \text{transcript}) \leq I(x:y)$

Main result (negative part).

Theorem

Let x and y be input strings of length n on which the protocol succeeds with error probability ϵ so that with prob $1-\epsilon$ Alice and Bob have at the end the same z, and $C(z \mid t) \geq |z| - \delta(n)$.

Then with probability $\geq 1 - O(\epsilon)$ we have
$|z| \leq I(x:y) + \delta(n) + O(\log(n/\epsilon))$.

Under the hood:

Conditional information inequality

- simple part: if no communication, then **key** $\leq I(x:y)$
- still simple: with communication, **key** $\leq I(x:y \mid \text{transcript})$
- hard part: $I(x:y \mid \text{transcript}) \leq I(x:y)$
- technical lemma: $C(\text{transcript} \mid x) + C(\text{transcript} \mid y) \leq C(\text{transcript})$

Main result (negative part).

Theorem

Let x and y be input strings of length n on which the protocol succeeds with error probability ϵ so that with prob $1-\epsilon$ Alice and Bob have at the end the same z, and $C(z \mid t) \geq |z| - \delta(n)$.

Then with probability $\geq 1 - O(\epsilon)$ we have
$|z| \leq I(x:y) + \delta(n) + O(\log(n/\epsilon))$.

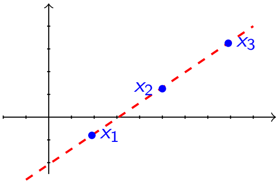
Under the hood:

Conditional information inequali

cf. Kaced-R.-Vereshchagin 2017
(Shannon's entropy version)

- simple part: if no comm
- still simple: with communicatio **key** $\leq I(x:y \mid \text{transcript})$
- hard part: $I(x:y \mid \text{transcript}) \leq I(x:y)$
- technical lemma: $C(\text{transcript} \mid x) + C(\text{transcript} \mid y) \leq C(\text{transcript})$

Result (2): Multi-party secret agreement

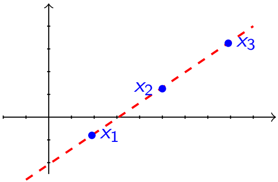


Alice:	x_1
Bob:	x_2
Charlie:	x_3

points x_1, x_2, x_3 belong to one line in the affine plane over $\mathbb{F}_{2^n}$

maximal common secret key: $n/2$ bits

Result (2): Multi-party secret agreement



Alice: x_1
Bob: x_2
Charlie: x_3

points x_1, x_2, x_3 belong to one line in the affine plane over $\mathbb{F}_{2^n}$

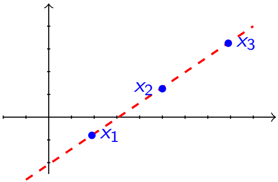
maximal common secret key: $n/2$ bits

Theorem (Informal statement)

We have ℓ parties, given inputs $x_1, \ldots, x_\ell$.

Each party also knows the complexity profile of $(x_1, \ldots, x_\ell)$.

Result (2): Multi-party secret agreement



Alice:	x_1
Bob:	x_2
Charlie:	x_3

points x_1, x_2, x_3 belong to one line in the affine plane over $\mathbb{F}_{2^n}$

maximal common secret key: $n/2$ bits

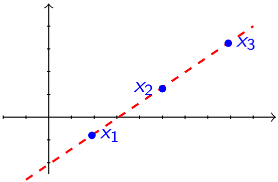
Theorem (Informal statement)

We have ℓ parties, given inputs $x_1, \ldots, x_\ell$.

Each party also knows the complexity profile of $(x_1, \ldots, x_\ell)$.

We provide an explicit formula for the maximal size of the common secret key that can be established via an open communication channel.

Result (2): Multi-party secret agreement



Alice: x_1
Bob: x_2
Charlie: x_3

points x_1, x_2, x_3 belong to one line in the affine plane over $\mathbb{F}_{2^n}$

maximal common secret key: $n/2$ bits

Theorem (Informal statement)

We have ℓ parties, given inputs $x_1, \ldots, x_\ell$.

Each party also knows the complexity profile of $(x_1, \ldots, x_\ell)$.

We provide an explicit formula for the maximal size of the common secret key that can be established via an open communication channel.

Under the hood:

- the size of the secret key → linear program → explicit (but complex) formula [Chan et. al., 2015]
- ... + the same techniques as for $\ell = 2$

Result (3): Communication complexity for secret key agreement

Fact: Our protocol for secret key agreement produces a key of length $\approx I(x:y)$ and has communication complexity $\approx \min\{C(x \mid y), C(y \mid x)\}$.

Result (3): Communication complexity for secret key agreement

Fact: Our protocol for secret key agreement produces a key of length $\approx I(x:y)$ and has communication complexity $\approx \min\{C(x \mid y), C(y \mid x)\}$.

Theorem (Somewhat informal)

If the communication complexity of a protocol <u>*with public randomness*</u> *is* $< 0.999 \cdot \min\{C(x \mid y), C(y \mid x)\}$, *then the size of the obtained common secret key is* $\ll 0.001 \cdot I(x:y)$.

Result (3): Communication complexity for secret key agreement

Fact: Our protocol for secret key agreement produces a key of length $\approx I(x:y)$ and has communication complexity $\approx \min\{C(x \mid y), C(y \mid x)\}$.

Theorem (Somewhat informal)

If the communication complexity of a protocol with public randomness is $< 0.999 \cdot \min\{C(x \mid y), C(y \mid x)\}$*, then the size of the obtained common secret key is* $\ll 0.001 \cdot I(x:y)$.

Under the hood:

- common information is far less than mutual information; (Gács & Körner 1970s ; Kolmogorov seminar in 1990s; Muchik & A.R. 2000s)
- opposition stochastic/nonstochastic objects (Shen 1983; Razenshteyn 2011)

Result (3): Communication complexity for secret key agreement

Fact: Our protocol for secret key agreement produces a key of length $\approx I(x:y)$ and has communication complexity $\approx \min\{C(x \mid y), C(y \mid x)\}$.

Theorem (Somewhat informal)

If the communication complexity of a protocol with public randomness is $< 0.999 \cdot \min\{C(x \mid y), C(y \mid x)\}$, *then the size of the obtained common secret key is* $\ll 0.001 \cdot I(x:y)$.

Under the hood:

- common information is far less than mutual information;
  (Gács & Körner 1970s ; Kolmogorov seminar in 1990s; Muchik & A.R. 2000s)
- opposition stochastic/nonstochastic objects (Shen 1983; Razenshteyn 2011)

Open question

What is the communication complexity for the model with private random bits?

Previous results: Shannon framework

- Ahlswede and Csiszár [1993] and Maurer [1993]:
  the optimal size of the common secret key for two parties
- Csiszár and Narayan [2004]:
  the optimal size of the common secret key for $\ell > 2$ parties
- Tyagi [2013]: communication complexity of the protocols
-

Previous results: Shannon framework

- Ahlswede and Csiszár [1993] and Maurer [1993]:
  the optimal size of the common secret key for two parties
- Csiszár and Narayan [2004]:
  the optimal size of the common secret key for $\ell > 2$ parties
- Tyagi [2013]: communication complexity of the protocols
-

formal difference
previous works: random variables & Shannon's entropy
our work: binary strings & Kolmogorov complexity

Previous results: Shannon framework

- Ahlswede and Csiszár [1993] and Maurer [1993]:
  the optimal size of the common secret key for two parties
- Csiszár and Narayan [2004]:
  the optimal size of the common secret key for $\ell > 2$ parties
- Tyagi [2013]: communication complexity of the protocols
-

1st substantial difference
previous works: (X,Y) from random memoryless / stationary ergodic sources
our work : no specific structure on X and Y

Previous results: Shannon framework

- Ahlswede and Csiszár [1993] and Maurer [1993]:
  the optimal size of the common secret key for two parties
- Csiszár and Narayan [2004]:
  the optimal size of the common secret key for $\ell > 2$ parties
- Tyagi [2013]: communication complexity of the protocols
-

2nd substantial difference
previous works: protocols work for most admissible pairs (X, Y)
our work: protocols work for all admissible pairs (X, Y)

Previous results: Kolmogorov complexity framework

• Finding an operational characterization of mutual information has been a folklore open problem.

• Some earlier approaches:

- Common information of x and y: longest z that can be computed from x and, separately, from y with a few help bits:
  $C(z \mid x) = O(\log n) \qquad C(z \mid y) = O(\log n)$.
- Is common information equal to mutual information?
- Gács and Körner [1973]: NO!They exhibit x, y with $I(x : y) = \Omega(n)$ and $|z| = o(n)$.
- Muchnik, Romashchenko, Chernov, Vereschagin: several papers with refinements of Gács and Körner.

  For example: there are x, y with $C(x), C(y) =^{+} n, I(x : y) = 0.99n$, $|z| = O(\log n)$.

One proof

- Upon conditioning, mutual information can increase or decrease.
- There are x, y, t_1, t_2 such that $I(x : y \mid t_1) > I(x : y)$ and $I(x : y \mid t_2) < I(x : y)$
- We show:

If $t(x, y)$ is a function with the rectangle property, then conditioning with $t(x, y)$ decreases mutual information.

- t has the rectangle property, if

$$t(x_1, y_1) = t(x_2, y_2) = t \Rightarrow t(x_1, y_2) = t.$$

- This fact is the key point in showing the negative part of our main result.
- The transcript $t(x, y)$ of a protocol on input (x, y) has the rectangle property.
- It has other applications in comm. complexity, maybe elsewhere as well.

One proof(2)

Theorem

If t has the rectangle property, then for all x, y, $I(x : y \mid t(x,y)) \leq I(x : y) + O(\log n)$.

Proof:

- Fix x, y, $t = t(x,y)$
- x' is a clone of x, if $\exists y', t(x', y') = t$ and $C(x') \leq C(x)$.
- y' is a clone of y, if $\exists x', t(x', y') = t$ and $C(y') \leq C(y)$.
- Clones_x: set of clones of x; Clones_y: set of clones of y.
- FACT: $\log |\mathrm{Clones}_x| \geq C(x \mid t) - O(\log n)$ (and similarly for Clones_y.)

  Proof: x is described by its ordinal in an enumeration of Clones_x, which can be done effectively given t and $C(x)$.

  So, $C(x \mid t) \leq \log |\mathrm{Clones}_x| + O(\log n)$.

One proof(3)

Theorem

If t has the rectangle property, then for all x, y, $I(x : y \mid t(x,y)) \le I(x : y) + O(\log n)$.

Proof (continuation):

- Take a pair $(x', y') \in \mathrm{Clones}_x \times \mathrm{Clones}_y$ with maximal $C(x', y' \mid t)$.
- $C(x', y' \mid t) \ge^+ \log |\mathrm{Clones}_x \times \mathrm{Clones}_y| =^+ C(x \mid t) + C(y \mid t)$.
- $C(x', y') = C(x', y', t)$ (because $t = t(x', y')$, using the rectangle property)
  $=^+ C(t) + C(x', y' \mid t)$ (chain rule)
  $\ge^+ C(t) + C(x \mid t) + C(y \mid t)$.
- On the other hand,
  $C(x', y') \le C(x') + C(y') \le C(x) + C(y)$ (by def. of clones)
- Combining the last two inequalities, $C(t) + C(x \mid t) + C(y \mid t) \le^+ C(x) + C(y)$.
- Now subtract $C(x, y, t)$ in the LHS, and (the smaller) $C(x, y)$ in the RHS.
- QED

Take home message

- Operational characterization of the mutual information of strings x and y:

$I(x : y)$ is equal (up to logarithmic precision) to the length of a longest secret key that two parties, one having x and the other having y, can establish via an interactive protocol on an open channel.

The protocol is probabilistic and the parties also need to know how their strings are correlated (i.e., they know the complexity profile of x and y).

- The protocol has communication complexity $\min(C(x \mid y), C(y \mid x))$.
- The communication is **optimal** for finding a secret key of maximal length, in the model with public randomness.
- We also determine the maximum length of a shared secret key in the multi-party setting.

Thank you

Full version:

- A. Romashchenko and M. Zimand, An operational characterization of mutual information in algorithmic information theory, available at ECCC `https://eccc.weizmann.ac.il/report/2018/043`