Interactive compression to entropy

Shay Moran (Technion)

based on work with

Balthazar Bauer (ENS) and Amir Yehudayoff (Technion)

Compression

Compression is about identifying the essential parts of objects of interest (and discarding the rest).

Compressing is often an evidence of understanding

Occam's razor - "Simpler is better".
Einstein's razor - "Make things as simple as possible, but not simpler".

Here we focus on compression of conversations.

A conversation that is easy to compress

Alice: hi. Bob: hi. Alice: how are you? Bob: good. how are you? Alice: good. how are you? Bob: good. how are you? Alice: good. how are you? Bob: good. how are you? Alice: good. how are you? Bob: good. how are you? Alice: good. how are you? Bob: good. how are you? Alice: good. how are you? Bob: good. how are you? Alice: good. how are you? Bob: good. how are you? Alice: good. how are you? Bob: good. how are you? Alice: good. how are you? Bob: good. how are you? Alice: great. how are you? Bob: great. how are you? Alice: good. how are you? Bob: good. how are you? Alice: good. how are you? Bob: good. how are you? Alice: good. how are you? Bob: good. bye. Alice: bye.

Model

Communication complexity [Yao '79].

There are 2 players: Alice and Bob.
Alice gets input x and Bob gets input y.

They communicate according to a *protocol* π:
Alice says m_1,
Bob says m_2,
Alice says m_3, ...

The transcript is denoted by M_π.

What does "compression" mean?

Roughly, we want shorten the conversation
as much as possible while keeping its content.

There are several ways to measure information
[Shannon, ... ,(Bar-Yossef)-Jayram-Kumar-Sivakumar,
Chakrabarti-Shi-Wirth-Yao,
Barak-Braverman-Chen-Rao, ... , Bauer-M-Yehudayoff, ...].

Information complexities: external and internal information

The input (X, Y) is drawn from a known distribution μ.

The external information of π is

$$I_\mu^{ext} = I(M_\pi : XY)$$

Number of bits an external observer learns on the input from the transcript

The internal information of π is

$$I_\mu^{int} = I(M_\pi : Y|X) + I(M_\pi : X|Y)$$

Number of bits that Alice learns on input from the transcript

$+$

Number of bits that Bob learns on input from the transcript

Information complexities: external and internal entropies

The input (X, Y) is drawn from a known distribution μ.

The external entropy of π is

$$H_\mu^{ext} = H(M_\pi)$$

Number of bits required to describe the transcript to an external observer

The internal entropy of π is

$$H_\mu^{int} = H(M_\pi|X) + H(M_\pi|Y)$$

Number of bits required to describe the transcript to Alice
+
Number of bits required to describe the transcript to Bob

Information versus entropy

$H_\mu^{ext} \geq I_\mu^{ext}$ and $H_\mu^{int} \geq I_\mu^{int}$.

Both are lower bounds on communication.

Information satisfies direct sum.

[Braverman-Rao]
Internal information = Amortized communication.

$H_\mu^{ext} = I_\mu^{ext}$ and $H_\mu^{int} = I_\mu^{int}$ for protocols without private randomness.

In this talk all protocols are deterministic

External simulation

A protocol σ is an external ϵ-error simulation of π if there exists a dictionary D such that the distribution of (x, y, M_π) is ϵ-close in statistical distance to that of $(x, y, D(M_\sigma))$.

An external observer can infer M_π from M_σ.

Internal simulation

A protocol σ is an internal ϵ-error simulation of π if there exists private dictionaries D_{Alice}, D_{Bob}, such that the distribution of (x, y, M_π) is ϵ-close in statistical distance to that of $(x, y, D_{Alice}(M_\sigma, x))$ and $(x, y, D_{Bob}(M_\sigma, y))$.

Alice and Bob can infer M_π from M_σ.
This is not necessarily true for an external observer!

Lower bounds

If σ is an external 0-error simulation of π then

$$CC_\mu(\sigma) \geq H_\mu^{ext}(\pi).$$

If σ is an internal 0-error simulation of π then

$$CC_\mu(\sigma) \geq H_\mu^{int}(\pi).$$

Are these lower bounds tight?

External compression: Upper bounds

Huffman code: If in π just Alice speaks

$$\forall \mu \ \exists \text{ external 0-error simulation } \sigma\text{:}$$
$$CC_\mu(\sigma) \leq H^{ext}_\mu(\pi) + 1.$$

[Dietzfelbinger-Wunderlich]

$$\forall \pi, \mu \ \exists \text{ external 0-error simulation } \sigma\text{:}$$
$$CC_\mu(\sigma) \leq 2.18 H^{ext}_\mu(\pi) + 2.$$

The 2.18 factor can be improved
[Kushilevitz - private communication]

Internal compression: Upper bounds

[Barak-Braverman-Chen-Rao*,
Brody-Buhrman-Koucky-Loff-Speelman-Vereshchagin, Pankratov]

$$\forall \pi, \mu, \epsilon > 0 \ \exists \text{ interrnal } \epsilon\text{-error simulation } \sigma:$$

$$CC_\mu(\sigma) \le O_\epsilon\big(H_\mu^{int}(\pi) \log CC_\mu(\pi)\big).$$

[Bauer-M-Yehudayoff]

$$\forall \pi, \mu, \epsilon > 0 \ \exists \text{ interrnal } \epsilon\text{-error simulation } \sigma:$$

$$CC_\mu(\sigma) \le O_\epsilon\big((H_\mu^{int}(\pi))^2 \log\log CC_\mu(\pi)\big).$$

A corollary: There is a private coin protocol with information $O(k)$ and communication 2^{2^k} such that every public coin simulation of it with information $O(k)$ has an exponential loss in communication [Ganor-Kol-Raz].

External simulation [Dietzfelbinger-Wunderlich]

[Dietzfelbinger-Wunderlich]

$$\forall \pi, \mu \ \exists \text{ external 0-error simulation } \sigma\text{:}$$
$$CC_\mu(\sigma) \leq 2.18 H_\mu^{ext}(\pi) + 2.$$

Proof:
Let π be a deterministic protocol.
We want to construct a protocol σ which simulates π with 0-error and $CC_\mu(\sigma) = O(H_\mu^{ext}(\pi))$.

Idea: Communicate 2 bits and convey 0.5 bits of information.

Meaningful vertices

Let π be a protocol and μ a distribution on inputs.

For every vertex v in π the set

$$R_v = \{(x, y) : v \text{ is an ancestor of } M_\pi(x, y)\}$$

is a rectangle.

v is meaningful if either
- $1/3 \leq \mu(R_v) \leq 2/3$, or
- v is a leaf and $\mu(R_v) \geq 2/3$.

Lemma: For all μ, π there exists a meaningful vertex!

External simulation [Dietzfelbinger-Wunderlich]

Description of σ on input (x, y):

- Let v be a meaningful vertex. Alice and Bob communicate 2 bits to determine if $(x, y) \in R_v$ and update μ accordingly.
- Either v was a leaf and the simulation is over or they learned 0.5 bits of information (since v is "meaningful").

An external compression: Summary

[Dietzfelbinger-Wunderlich]

$$\forall \pi, \mu \ \exists \text{ external 0-error simulation } \sigma\text{:}$$
$$CC_\mu(\sigma) \leq 2.18 H_\mu^{ext}(\pi) + 2.$$

Idea: Communicate 2 bits and convey 0.5 bits of information.

Can something similar work for internal entropy?

What is a meaningful vertex in this case?

Internal compression - meaningful vertices

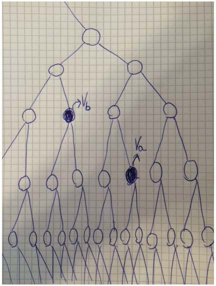
Let (x, y) be an input and let v be a vertex in T_π.

- $\mu_x(R_v) = \mu(R_v \mid X = x)$ Alice's perspective
- $\mu_y(R_v) = \mu(R_v \mid Y = y)$ Bob's perspective

Alice has a meaningful vertex v_a with respect to μ_x.
Bob has a meaningful vertex v_b with respect to μ_y.

Internal compression

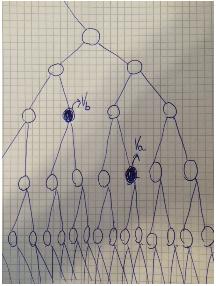
Alice has a meaningful vertex v_a.
Bob has a meaningful vertex v_b.



If both knew v_a or v_b $\implies$ communicate 2 bits to learn 0.5 bits.

Problem: Alice doesn't know v_b, Bob doesn't know v_a

Internal compression



Problem: Alice doesn't know v_b, Bob doesn't know v_a

Observation: Enough that they agree on a vertex v such that
- v is an ancestor of v_a, and
- v is not an ancestor of v_b.

Can be done with $O_\epsilon\left(H_\mu^{int}(\pi)\log\log CC_\mu(\pi)\right)$ communication:
- [Feige,Peleg, Raghavan, Upfal] protocol for finding the first difference,
- A variant of sampling from [Braverman-Rao].

Summary

Defined a couple of meanings for “compressing a conversation.”

- Optimal external compression [Dietzfelbinger-Wunderlich].
- Efficient internal compression [Bauer-M-Yehudayoff].

We still do not have a full understanding.