Communication Complexity II

Amir Yehudayoff (Technion)

introduction

compression is about finding efficient representations of objects of interest (pictures, music, data,...)

the ability to compress is often an evidence that we understand something (Occam's razor, learning theory,...)

focus on compression of conversations

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.

compression?

want: length $\approx$ "true content"

definitions:

communication complexity [Yao]

information theory
[Shannon, ... ,(Bar-Yossef)-Jayram-Kumar-Sivakumar, Chakrabarti-Shi-Wirth-Yao, Barak-Braverman-Chen-Rao, ... , Bauer-Moran-Y, ...]

the model

two players

(x, y) is chosen from a known distribution μ

alice gets input x
bob gets input y

they communicate according to a randomized protocol π

the transcript $T_\pi = (\pi(x, y, r, r_a, r_b), r)$ is the "conversation"

the model

two players

(x, y) is chosen from a known distribution μ

alice gets input x
bob gets input y

they communicate according to a randomized protocol π

the transcript $T_\pi = (\pi(x, y, r, r_a, r_b), r)$ is the "conversation"

compression = an **efficient simulation** of T_π

external versus external

external compression

simulations

a protocol σ is a **external ϵ-error simulation** of π if

$$(x, y, T_\pi) \underset{\epsilon}{\approx} (x, y, d(T_\sigma))$$

i.e. ϵ-close in statistical distance, for some d

comments:

the map d can be thought of as the dictionary that translates the language of σ to that of π

the word "external" indicates that every person that hears σ can understand the meaning of $\sigma(x, y)$ as a "π conversation" as long as he/she knows the dictionary

example: one way conversations

assume alice wants to send $x \sim \mu$ to bob

what is the most efficient way?

example: one way conversations

assume alice wants to send $x \sim \mu$ to bob

what is the most efficient way?

length of π under μ

$$CC_\mu(\pi) = \mathop{\mathbb{E}}_{x\sim\mu} |\pi(x)|$$

content is measured by entropy

$$H(\mu) = \sum_x \mu(x) \log(1/\mu(x))$$

example: one way conversations

assume alice wants to send $x \sim \mu$ to bob

what is the most efficient way?

length of π under μ

$$CC_\mu(\pi) = \mathop{\mathbb{E}}_{x\sim\mu} |\pi(x)|$$

content is measured by entropy

$$H(\mu) = \sum_x \mu(x) \log(1/\mu(x))$$

theorem [Shannon, Huffman]
1. if $\sigma(x)$ determines x then $H(\mu) \leq CC_\mu(\sigma)$
2. there is σ that determines x so that $CC_\mu(\sigma) \leq H(\mu) + 1$

entropy

let X be a random variable taking values in finite set S

entropy:

$$H(X) = \sum_{x \in S} \Pr[X = x] \log(1/\Pr[X = x])$$

"the optimal expected description length of X"

entropy

let X be a random variable taking values in finite set S

entropy:

$$H(X) = \sum_{x \in S} \Pr[X = x] \log(1/\Pr[X = x])$$

"the optimal expected description length of X"

chain rule:

$$H(X, Y) = H(X) + H(Y|X)$$

entropy

let X be a random variable taking values in finite set S

entropy:

$$H(X) = \sum_{x \in S} \Pr[X = x] \log(1/\Pr[X = x])$$

"the optimal expected description length of X"

chain rule:

$$H(X, Y) = H(X) + H(Y|X)$$

properties:

$H(Y|X) \leq H(Y)$
$H(X) \leq \log |S|$, equality iff X is uniform
$H(f(X)|X) = 0$

mutual information

let X, Y be random variables taking values in finite set S

mutual information:

$$I(X;Y) = H(X) - H(X|Y)$$

mutual information

let X, Y be random variables taking values in finite set S

mutual information:

$$I(X;Y) = H(X) - H(X|Y)$$

chain rule:

$$I(X,Y;Z) = I(X;Z) + I(Y;Z|X)$$

mutual information

let X, Y be random variables taking values in finite set S

mutual information:

$$I(X;Y) = H(X) - H(X|Y)$$

chain rule:

$$I(X,Y;Z) = I(X;Z) + I(Y;Z|X)$$

properties:

$I(X;Y) = 0$ iff X, Y independent
$I(X;Y) \geq I(X;f(Y))$ info processing inequality

external compression

roughly, σ is an **external compression** of π with if

1. σ externally simulates[1] π
2. $CC_\mu(\sigma) \lesssim I_\mu^{ext}(\pi)$

the **external information cost** of π is measured as

$$I_\mu^{ext}(\pi) = I(X, Y; T_\pi)$$

[1] usually error is $1/3$

external compression

roughly, σ is an **external compression** of π with if

1. σ externally simulates[1] π
2. $CC_\mu(\sigma) \lesssim I_\mu^{ext}(\pi)$

the **external information cost** of π is measured as

$$I_\mu^{ext}(\pi) = I(X, Y; T_\pi)$$

a lower bound: if σ is an external 0-error simulation of π then

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

[1] usually error is $1/3$

external compression: upper bounds

for all π, μ:

[Huffman] if π is deterministic and just alice speaks, there is an optimal external 0-error simulation σ of π so that

$$CC_\mu(\sigma) \leq I_\mu^{ext}(\pi) + 1$$

[Dietzfelbinger-Wunderlich] if π is deterministic, there is an external 0-error simulation σ of π so that

$$CC_\mu(\sigma) < 4I_\mu^{ext}(\pi) + 4$$

[Barak-Braverman-Chen-Rao] there is an external simulation σ of π so that

$$CC_\mu(\sigma) \lesssim I_\mu^{ext}(\pi) \cdot \log CC_\mu(\pi)$$

internal compression

internal simulations

a protocol σ is an **internal simulation** of π if

$$(x, y, T_\pi) \underset{\epsilon}{\approx} (x, y, d_a(x, r_a, T_\sigma)) \underset{\epsilon}{\approx} (x, y, d_b(y, r_b, T_\sigma))$$

i.e. ϵ-close in statistical distance, for some d_a, d_b

comments:

the maps d_a, d_b are private dictionaries that translate the language of σ to that of π

the word "internal" indicates that only alice and bob are guaranteed to understand the meaning of σ

internal compression

roughly, σ is an **internal compression** of π with ϵ-error if

1. σ internally ϵ-error simulates π and
2. $CC_\mu(\sigma) \lesssim I_\mu^{int}(\pi)$

the **interal information cost** of π is measured as

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

always

$$I_\mu^{int}(\pi) \leq I_\mu^{ext}(\pi)$$

Internal simulation: lower bounds

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

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

[Bauer-Moran-Y]* for every N, there is μ and a protocol π with $I_\mu^{int}(\pi) = 1$ so that every σ that simulates π with 0-error satisfies

$$CC_\mu(\sigma) \geq N$$

[Ganor-Kol-Raz] there are π, μ so that

1. $I_\mu^{int}(\pi) = k$
2. if σ is an internal simulation of π then $CC_\mu(\sigma) \geq exp(k)$

and more...

Internal simulation: upper bounds

[Brody-Buhrman-Koucky-Loff-Speelman-Vereshchagin, Pankratov] for every deterministic π and μ, there is an internal simulation σ of π so that

$$CC_\mu(\sigma) \lesssim I_\mu^{int}(\pi) \cdot \log CC_\mu(\pi)$$

[Bauer-Moran-Y] for every deterministic π and μ, there is an internal simulation σ of π so that

$$CC_\mu(\sigma) \lesssim (I_\mu^{int}(\pi))^2 \cdot \log\log CC_\mu(\pi)$$

[Barak-Braverman-Chen-Rao] for every π and μ, there is an internal simulation σ of π so that

$$CC_\mu(\sigma) \lesssim \sqrt{I_\mu^{int}(\pi) \cdot CC_\mu(\pi)} \cdot \log CC_\mu(\pi)$$

local summary

compress π = efficient simulation of π

simulations: external and internal

information costs: external and internal

comments:

external compression is easier but internal compression is more useful (e.g. direct sums)

know almost optimal external compressions but no internal ones an external compression

an external compression

theorem [Dietzfelbinger-Wunderlich]

optimal external compression of deterministic protocols

idea

try jump directly to a "meaningful state" in the conversation

recall that $\pi(x, y)$ is a walk on a tree

observe: if π is deterministic then

$$I_\mu^{ext}(\pi) = I(X, Y; \pi(X, Y)) = H(\pi(X, Y)) - H(\pi(X, Y)|X, Y)$$
$$= H(\pi(X, Y)) = \mathop{\mathbb{E}}_{u \sim \pi(\mu)} \log \frac{1}{\mu(\{(x, y) : \pi(x, y) = u\})}$$

“meaningful”

for every vertex v in protocol tree, define a number

$$p(v) = \mu(\{(x, y) : \pi(x, y) \in \pi_v\})$$

where π_v is the subtree rooted at v

property: if v is parent of v_1, v_2 then $p(v) = p(v_1) + p(v_2)$

“meaningful”

for every vertex v in protocol tree, define a number

$$p(v) = \mu(\{(x, y) : \pi(x, y) \in \pi_v\})$$

where π_v is the subtree rooted at v

property: if v is parent of v_1, v_2 then $p(v) = p(v_1) + p(v_2)$

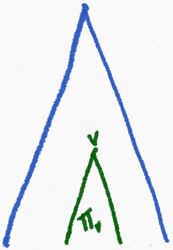
lemma (there is a meaningful state)

there exists v so that either v is a leaf with $p(v) \geq 2/3$ or

$$1/3 \leq p(v) \leq 2/3$$

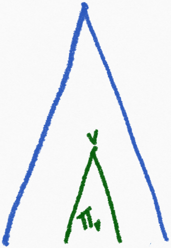
simulation

$p(v) \approx \frac{1}{2}, \; R_v = X_v \times Y_v$



simulation

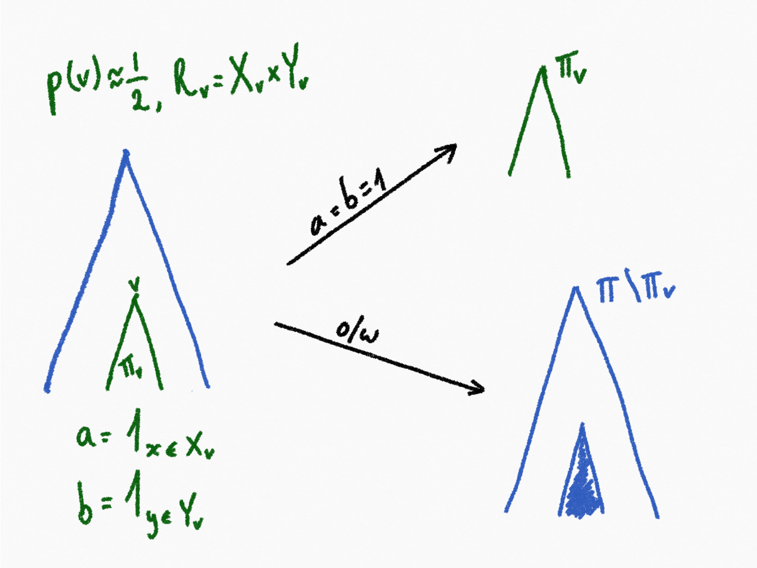
$p(v) \approx \frac{1}{2}, \; R_v = X_v \times Y_v$



$a = 1_{x \in X_v}$

$b = 1_{y \in Y_v}$

simulation



simulation

summary:

"two bits of communication give one bit of information"

simulation

summary:

"two bits of communication give one bit of information"

formally:

for fixed (x, y), the length of σ when simulating $u = \pi(x, y)$ is at most

$$2 \cdot \left\lceil \log_{3/2} \frac{1}{p(u)} \right\rceil \leq 2 \cdot \left(1 + 2 \log \frac{1}{p(u)} \right)$$

simulation

summary:

"two bits of communication give one bit of information"

formally:

for fixed (x, y), the length of σ when simulating $u = \pi(x, y)$ is at most

$$2 \cdot \left\lceil \log_{3/2} \frac{1}{p(u)} \right\rceil \leq 2 \cdot \left(1 + 2 \log \frac{1}{p(u)} \right)$$

so

$$\mathbb{E} \left| \sigma(x, y) \right| \leq 2 + 4\, \mathbb{E} \log \frac{1}{p(u)} = 2 + 4 I^{ext}(\pi)$$

an internal compression

an internal compression

theorem [Barak-Braverman-Chen-Rao]

for every π and μ, there is an internal simulation σ of π so that

$$CC_\mu(\sigma) \lesssim \sqrt{I_\mu^{int}(\pi) \cdot CC_\mu(\pi)} \log CC_\mu(\pi)$$

an internal compression

outline:

alice can sample a leaf v_a by guessing bob's messages

bob can sample a leaf v_b by guessing alice's messages

use knowledge of μ, π

an internal compression

outline:

alice can sample a leaf v_a by guessing bob's messages

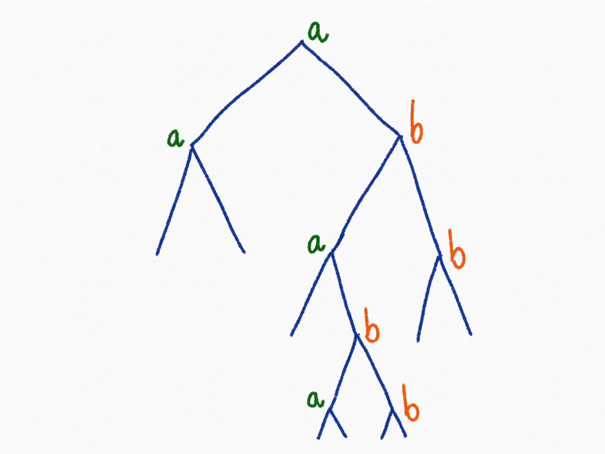
bob can sample a leaf v_b by guessing alice's messages

use knowledge of μ, π

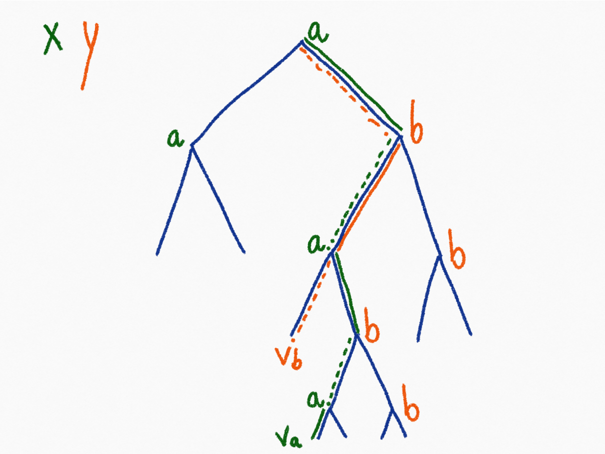
using public randomness they jointly sample (v_a, v_b)

if by chance $v_a = v_b$ then they sampled correctly

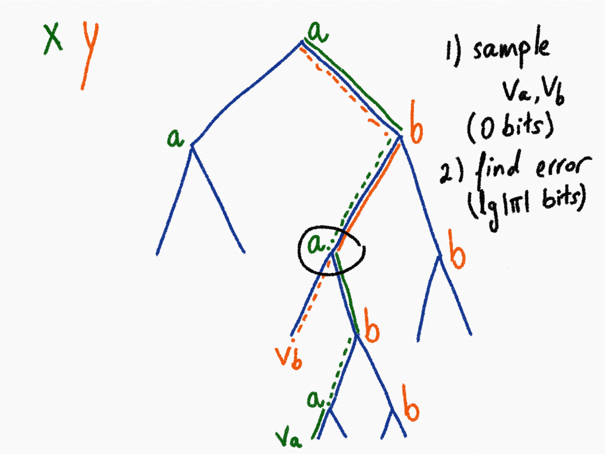
internal compression



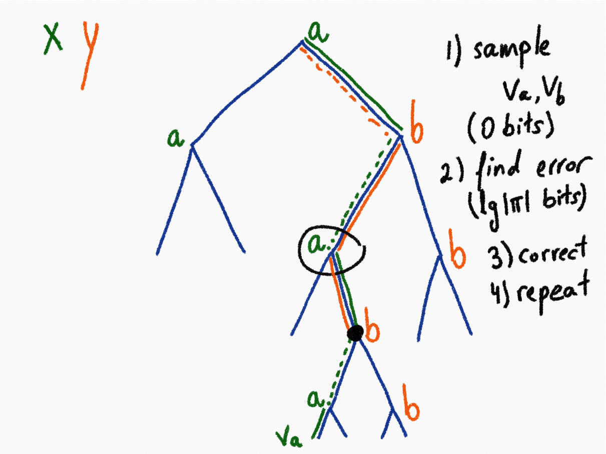
internal compression



internal compression



internal compression



internal compression

outline of simulation σ:

sample v_a, v_b (0 bits)

if $v_a = v_b$ then done, otherwise

find "error" ($\log CC_\mu(\pi)$ bits)

correct and repeat

internal compression

outline of simulation σ:

sample v_a, v_b (0 bits)

if $v_a = v_b$ then done, otherwise

find "error" ($\log CC_\mu(\pi)$ bits)

correct and repeat

$$\Rightarrow \quad CC_\mu(\sigma) \leq \log CC_\mu(\pi) \cdot \mathbb{E}\,\text{number of "errors"}$$

internal compression

outline of simulation σ:

sample v_a, v_b (0 bits)

if $v_a = v_b$ then done, otherwise

find "error" ($\log CC_\mu(\pi)$ bits)

correct and repeat

$$\Rightarrow \quad CC_\mu(\sigma) \leq \log CC_\mu(\pi) \cdot \mathbb{E}\,\text{number of "errors"}$$

lemma: $\mathbb{E}$ number of "errors" $\leq \sqrt{I_\mu^{int}(\pi) \cdot CC_\mu(\pi)}$

summary

defined internal and external compression

saw two concrete compression schemes

connections to direct sum and product

we still do not have a full understanding

spaseeba