Trading Information Complexity for Error

Yaqiao Li
joint work with Yuval Dagan, Yuval Filmus, Hamed Hatami

School of Computer Science
McGill University

July 8, 2017

Main results

- How much information one can save by allowing an error ϵ.

Main results

- How much information one can save by allowing an error ϵ.
- Showed a separation between two concepts in information complexity.

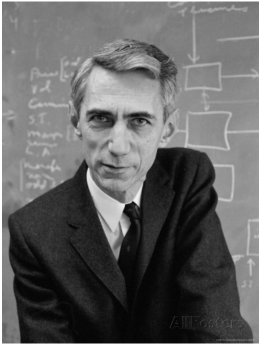
Main results

- How much information one can save by allowing an error ϵ.
- Showed a separation between two concepts in information complexity.
- Determined communication complexity of computing disjointness function with error ϵ.

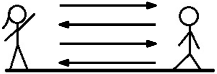
Information complexity

Extension of Shannon's information theory
towards studying communication complexity.

Shannon (1916-2001)

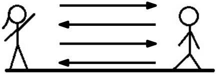


Communication complexity



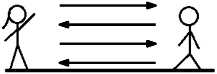
- Alice receives an input $X \in \{0,1\}^n$, Bob receives $Y \in \{0,1\}^n$.

Communication complexity



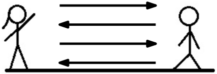
- Alice receives an input $X \in \{0,1\}^n$, Bob receives $Y \in \{0,1\}^n$.
- They want to compute $f(X,Y) : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ collaboratively using a protocol.

Communication complexity



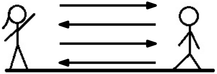
- Alice receives an input $X \in \{0,1\}^n$, Bob receives $Y \in \{0,1\}^n$.
- They want to compute $f(X, Y) : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ collaboratively using a protocol.
- A protocol π is an algorithm that defines what Alice and Bob do, in order to compute $f(X, Y)$.

Communication complexity



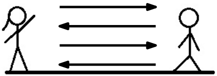
- Alice receives an input $X \in \{0,1\}^n$, Bob receives $Y \in \{0,1\}^n$.
- They want to compute $f(X, Y) : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ collaboratively using a protocol.
- A protocol π is an algorithm that defines what Alice and Bob do, in order to compute $f(X, Y)$.
- $CC(\pi) :=$ how many bits of communication are sent in π ?

Information complexity



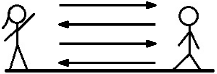
- Same setting;

Information complexity



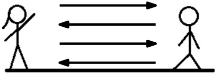
- Same setting;
- assume inputs $(X, Y) \sim \mu$, to measure information (entropy).

Information complexity



- Same setting;
- assume inputs $(X, Y) \sim \mu$, to measure information (entropy).
- How much information the players have to reveal about their inputs?

Information complexity



- Same setting;
- assume inputs $(X, Y) \sim \mu$, to measure information (entropy).
- How much information the players have to reveal about their inputs?

Information cost of a protocol π:
$\mathrm{IC}_\mu(\pi) =$ information about Y that Alice learns $+$
information about X that Bob learns.

Information cost: Example 1

AND: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X = 1] = \frac{1}{2}, \quad \Pr[Y = 1] = \frac{1}{2}$.

Information cost: Example 1

AND: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X = 1] = \frac{1}{2}, \quad \Pr[Y = 1] = \frac{1}{2}$.

Naive protocol π

- Alice sends her input X to Bob;
- Bob sends his input Y to Alice.

Information cost: Example 1

AND: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X=1] = \frac{1}{2}, \quad \Pr[Y=1] = \frac{1}{2}$.

Naive protocol π

- Alice sends her input X to Bob;
- Bob sends his input Y to Alice.

- Alice learns Bob's input: learned information $= H(Y) = 1$;

Information cost: Example 1

AND: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X = 1] = \frac{1}{2}$, $\Pr[Y = 1] = \frac{1}{2}$.

Naive protocol π

- Alice sends her input X to Bob;
- Bob sends his input Y to Alice.

- Alice learns Bob's input: learned information $= H(Y) = 1$;
- Bob learns Alice's input: learned information $= H(X) = 1$;

Information cost: Example 1

AND: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X = 1] = \frac{1}{2}, \quad \Pr[Y = 1] = \frac{1}{2}$.

Naive protocol π

- Alice sends her input X to Bob;
- Bob sends his input Y to Alice.

- Alice learns Bob's input: learned information $= H(Y) = 1$;
- Bob learns Alice's input: learned information $= H(X) = 1$;
- Information cost of π:

$$\mathrm{IC}_\mu(\pi) = 1 + 1 = 2 = CC(\pi).$$

Information cost: Example 2

AND: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X=1] = \frac{1}{2}, \quad \Pr[Y=1] = \frac{1}{2}$.

Information cost: Example 2

AND: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X=1]=\frac{1}{2}$, $\Pr[Y=1]=\frac{1}{2}$.

Better protocol τ

- Alice sends her input X to Bob;
- Bob computes and outputs $\mathrm{AND}(X,Y)$.

Information cost: Example 2

AND: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X = 1] = \frac{1}{2}, \quad \Pr[Y = 1] = \frac{1}{2}$.

Better protocol τ

- Alice sends her input X to Bob;
- Bob computes and outputs $\mathrm{AND}(X, Y)$.

Note: $CC(\tau) = 2 = CC(\pi)$.

Information cost: Example 2

AND: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X = 1] = \frac{1}{2}$, $\Pr[Y = 1] = \frac{1}{2}$.

Better protocol τ

- Alice sends her input X to Bob;
- Bob computes and outputs $\mathrm{AND}(X, Y)$.

Note: $CC(\tau) = 2 = CC(\pi)$.

Why is the protocol τ better than π?

Information cost: Example 2 - continued

better protocol τ

- Alice sends her input X to Bob;
- Bob computes and outputs $\mathrm{AND}(X, Y)$.

Information cost $\mathrm{IC}_\mu(\tau) = ?$

Information cost: Example 2 - continued

better protocol τ

- Alice sends her input X to Bob;
- Bob computes and outputs $\mathrm{AND}(X, Y)$.

Information cost $\mathrm{IC}_\mu(\tau) =?$

- Bob always learns X : learned information $= H(X) = 1$;

Information cost: Example 2 - continued

better protocol τ

- Alice sends her input X to Bob;
- Bob computes and outputs $\mathrm{AND}(X, Y)$.

Information cost $\mathrm{IC}_\mu(\tau) =$?

- Bob always learns X : learned information $= H(X) = 1$;
- When $X = 1$, Alice learns Y;

Information cost: Example 2 - continued

better protocol τ

- Alice sends her input X to Bob;
- Bob computes and outputs $\mathrm{AND}(X, Y)$.

Information cost $\mathrm{IC}_\mu(\tau) = ?$

- Bob always learns X : learned information $= H(X) = 1$;
- When $X = 1$, Alice learns Y;
- When $X = 0$, Alice learns nothing;

Information cost: Example 2 - continued

better protocol τ

- Alice sends her input X to Bob;
- Bob computes and outputs $\mathrm{AND}(X, Y)$.

Information cost $\mathrm{IC}_\mu(\tau) = ?$

- Bob always learns X : learned information $= H(X) = 1$;
- When $X = 1$, Alice learns Y;
- When $X = 0$, Alice learns nothing;
- Alice learns information $= \frac{1}{2}H(Y) = 0.5$.

Information cost: Example 2 - continued

better protocol τ

- Alice sends her input X to Bob;
- Bob computes and outputs $\mathrm{AND}(X, Y)$.

Information cost $\mathrm{IC}_\mu(\tau) = ?$

- Bob always learns X : learned information $= H(X) = 1$;
- When $X = 1$, Alice learns Y;
- When $X = 0$, Alice learns nothing;
- Alice learns information $= \frac{1}{2} H(Y) = 0.5$.
- Information cost of the protocol τ:

$$\mathrm{IC}_\mu(\tau) = 1 + 0.5 = 1.5 < 2 = \mathrm{IC}_\mu(\pi) \implies \tau \text{ is better!}.$$

Information complexity of a function f

Definition

$$\mathrm{IC}_\mu(f,0) := \inf_\pi \mathrm{IC}_\mu(\pi),$$

where π computes f with no error.

Information complexity of a function f

Definition

$$\mathrm{IC}_\mu(f, 0) := \inf_\pi \mathrm{IC}_\mu(\pi),$$

where π computes f with no error.

Example

We saw for μ being the uniform distribution,

$$\mathrm{IC}_\mu(\mathrm{AND}, 0) \leq \mathrm{IC}_\mu(\tau) = 1.5$$

Information complexity of a function f

Definition

$$\mathrm{IC}_\mu(f, 0) := \inf_\pi \mathrm{IC}_\mu(\pi),$$

where π computes f with no error.

Example

We saw for μ being the uniform distribution,

$$\mathrm{IC}_\mu(\mathrm{AND}, 0) \leq \mathrm{IC}_\mu(\tau) = 1.5 \qquad \text{optimal?}$$

Information complexity of AND

Theorem (BGPW'13)

Let μ be the uniform distribution, then

$$\mathrm{IC}_\mu(\mathrm{AND}, 0) \approx 1.36...$$

Information complexity of AND

Theorem (BGPW'13)

Let μ be the uniform distribution, then

$$\mathrm{IC}_\mu(\mathrm{AND}, 0) \approx 1.36...$$

Theorem (BGPW'13)

$$\max_\mu \mathrm{IC}_\mu(\mathrm{AND}, 0) \approx 1.49...$$

Information complexity: Example 3

XOR: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X = 1] = \frac{1}{2}, \quad \Pr[Y = 1] = \frac{1}{2}$.

Information complexity: Example 3

XOR: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X = 1] = \frac{1}{2}, \quad \Pr[Y = 1] = \frac{1}{2}$.

Naive protocol π

- Alice sends her input X to Bob;
- Bob sends his input Y to Alice.

Information complexity: Example 3

XOR: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X=1] = \frac{1}{2}$, $\Pr[Y=1] = \frac{1}{2}$.

Naive protocol π

- Alice sends her input X to Bob;
- Bob sends his input Y to Alice.

Information complexity $\mathrm{IC}_\mu(\mathrm{XOR}, 0)$

$\mathrm{IC}_\mu(\pi) = 1 + 1 = 2.$

Information complexity: Example 3

XOR: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X = 1] = \frac{1}{2}$, $\quad \Pr[Y = 1] = \frac{1}{2}$.

Naive protocol π

- Alice sends her input X to Bob;
- Bob sends his input Y to Alice.

Information complexity $\mathrm{IC}_\mu(\mathrm{XOR}, 0)$

$\mathrm{IC}_\mu(\pi) = 1 + 1 = 2$. optimal!

Information complexity: Example 3

XOR: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X=1]=\frac{1}{2}, \quad \Pr[Y=1]=\frac{1}{2}$.

Naive protocol π

- Alice sends her input X to Bob;
- Bob sends his input Y to Alice.

Information complexity $\mathrm{IC}_\mu(\mathrm{XOR}, 0)$

$\mathrm{IC}_\mu(\pi) = 1 + 1 = 2$. optimal! $\Longrightarrow \mathrm{IC}_\mu(\mathrm{XOR}, 0) = 2$.

Information complexity: Example 3

XOR: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X = 1] = \frac{1}{2}$, $\Pr[Y = 1] = \frac{1}{2}$.

Naive protocol π

- Alice sends her input X to Bob;
- Bob sends his input Y to Alice.

Information complexity $\mathrm{IC}_\mu(\mathrm{XOR}, 0)$

$\mathrm{IC}_\mu(\pi) = 1 + 1 = 2$. optimal! $\Longrightarrow \mathrm{IC}_\mu(\mathrm{XOR}, 0) = 2$.

Q: what if we want to use less information?

Information complexity: Example 3

XOR: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X=1] = \frac{1}{2}, \quad \Pr[Y=1] = \frac{1}{2}$.

Naive protocol π

- Alice sends her input X to Bob;
- Bob sends his input Y to Alice.

Information complexity $\mathrm{IC}_\mu(\mathrm{XOR}, 0)$

$\mathrm{IC}_\mu(\pi) = 1 + 1 = 2$. optimal! $\Longrightarrow \mathrm{IC}_\mu(\mathrm{XOR}, 0) = 2$.

Q: what if we want to use less information?
Great Idea: ALLOW ERROR!

Trading information complexity for error!

Recall

$$\mathrm{IC}_\mu(f, 0) := \inf_\pi \mathrm{IC}_\mu(\pi),$$

where π computes f with no error.

Trading information complexity for error!

Recall

$$\mathrm{IC}_\mu(f, 0) := \inf_\pi \mathrm{IC}_\mu(\pi),$$

where π computes f with no error.

Definition

$$\mathrm{IC}_\mu(f, \epsilon) := \inf_\pi \mathrm{IC}_\mu(\pi),$$

where π computes f

Trading information complexity for error!

Recall

$$\mathrm{IC}_\mu(f, 0) := \inf_\pi \mathrm{IC}_\mu(\pi),$$

where π computes f with no error.

Definition

$$\mathrm{IC}_\mu(f, \epsilon) := \inf_\pi \mathrm{IC}_\mu(\pi),$$

where π computes f with error ϵ

Trading information complexity for error!

Recall

$$\mathrm{IC}_\mu(f, 0) := \inf_\pi \mathrm{IC}_\mu(\pi),$$

where π computes f with no error.

Definition

$$\mathrm{IC}_\mu(f, \epsilon) := \inf_\pi \mathrm{IC}_\mu(\pi),$$

where π computes f with error ϵ for every input (X, Y).

Trading information complexity for error!

Recall

$$\mathrm{IC}_\mu(f, 0) := \inf_\pi \mathrm{IC}_\mu(\pi),$$

where π computes f with no error.

Definition

$$\mathrm{IC}_\mu(f, \epsilon) := \inf_\pi \mathrm{IC}_\mu(\pi),$$

where π computes f with error ϵ for every input (X, Y).

Q: How much information can one save?

$$\mathrm{IC}_\mu(\mathrm{XOR}, 0) - \mathrm{IC}_\mu(\mathrm{XOR}, \epsilon) = ?$$

Trading information complexity for error - example

XOR: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X = 1] = \frac{1}{2}, \quad \Pr[Y = 1] = \frac{1}{2}$.

Trading information complexity for error - example

XOR: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X=1] = \frac{1}{2}, \quad \Pr[Y=1] = \frac{1}{2}$.

protocol π_ϵ

Trading information complexity for error - example

XOR: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X=1]=\frac{1}{2}, \quad \Pr[Y=1]=\frac{1}{2}$.

protocol π_ϵ

- Alice flips her input X to X' with probability ϵ;
  Alice sets $X' = 1 - X$ with probability ϵ, otherwise $X' = X$;

Trading information complexity for error - example

XOR: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X=1] = \frac{1}{2}$, $\Pr[Y=1] = \frac{1}{2}$.

protocol π_ϵ

- Alice flips her input X to X' with probability ϵ;
  Alice sets $X' = 1 - X$ with probability ϵ, otherwise $X' = X$;
- Alice sends X' to Bob;

Trading information complexity for error - example

XOR: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X=1] = \frac{1}{2}, \quad \Pr[Y=1] = \frac{1}{2}$.

protocol π_ϵ

- Alice flips her input X to X' with probability ϵ;
  Alice sets $X' = 1 - X$ with probability ϵ, otherwise $X' = X$;
- Alice sends X' to Bob;
- Bob sends his input Y to Alice;

Trading information complexity for error - example

XOR: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X = 1] = \frac{1}{2}$, $\Pr[Y = 1] = \frac{1}{2}$.

protocol π_ϵ

- Alice flips her input X to X' with probability ϵ;
  Alice sets $X' = 1 - X$ with probability ϵ, otherwise $X' = X$;
- Alice sends X' to Bob;
- Bob sends his input Y to Alice;
- Alice and Bob compute $\mathrm{XOR}(X', Y)$,

Trading information complexity for error - example

XOR: $\{0, 1\} \times \{0, 1\} \to \{0, 1\}$

Product distribution μ: $\Pr[X = 1] = \frac{1}{2}, \quad \Pr[Y = 1] = \frac{1}{2}$.

protocol π_ϵ

- Alice flips her input X to X' with probability ϵ;
  Alice sets $X' = 1 - X$ with probability ϵ, otherwise $X' = X$;
- Alice sends X' to Bob;
- Bob sends his input Y to Alice;
- Alice and Bob compute $\mathrm{XOR}(X', Y)$, having error at most ϵ!

Trading information complexity for error - example

XOR: $\{0,1\} \times \{0,1\} \to \{0,1\}$

Product distribution μ: $\Pr[X = 1] = \frac{1}{2}$, $\Pr[Y = 1] = \frac{1}{2}$.

protocol π_ϵ

- Alice flips her input X to X' with probability ϵ;
  Alice sets $X' = 1 - X$ with probability ϵ, otherwise $X' = X$;
- Alice sends X' to Bob;
- Bob sends his input Y to Alice;
- Alice and Bob compute $\mathrm{XOR}(X', Y)$, having error at most ϵ!

Information cost $\mathrm{IC}_\mu(\pi_\epsilon) = ?$

Analysis of π_ϵ

- Product distribution μ: $\Pr[X = 1] = \frac{1}{2}, \quad \Pr[Y = 1] = \frac{1}{2}$.

Analysis of π_ϵ

- Product distribution μ: $\Pr[X = 1] = \frac{1}{2}, \quad \Pr[Y = 1] = \frac{1}{2}$.
- After flipping, X' is still uniform: $\Pr[X' = 1] = \frac{1}{2}$;

Analysis of π_ϵ

- Product distribution μ: $\Pr[X = 1] = \frac{1}{2}, \quad \Pr[Y = 1] = \frac{1}{2}$.
- After flipping, X' is still uniform: $\Pr[X' = 1] = \frac{1}{2}$;
- On receiving X', Bob is not sure what Alice's true input is.

Analysis of π_ϵ

- Product distribution μ: $\Pr[X=1]=\frac{1}{2}, \quad \Pr[Y=1]=\frac{1}{2}$.
- After flipping, X' is still uniform: $\Pr[X'=1]=\frac{1}{2}$;
- On receiving X', Bob is not sure what Alice's true input is.
- Bayesian:

$$\Pr[X=1|X'=1]=\frac{\Pr[X'=1|X=1]\Pr[X=1]}{\Pr[X'=1]}=1-\epsilon.$$

Analysis of π_ϵ

- Product distribution μ: $\Pr[X = 1] = \frac{1}{2}, \quad \Pr[Y = 1] = \frac{1}{2}$.
- After flipping, X' is still uniform: $\Pr[X' = 1] = \frac{1}{2}$;
- On receiving X', Bob is not sure what Alice's true input is.
- Bayesian:

$$\Pr[X = 1 | X' = 1] = \frac{\Pr[X' = 1 | X = 1] \Pr[X = 1]}{\Pr[X' = 1]} = 1 - \epsilon.$$

- Bob learns less information about Alice's input. Indeed,

$$\mathrm{IC}_\mu(\pi_\epsilon) = 1 + (1 - H(1 - \epsilon)) = 2 - H(\epsilon).$$

Analysis of π_ϵ

- Product distribution μ: $\Pr[X=1]=\frac{1}{2}, \quad \Pr[Y=1]=\frac{1}{2}$.
- After flipping, X' is still uniform: $\Pr[X'=1]=\frac{1}{2}$;
- On receiving X', Bob is not sure what Alice's true input is.
- Bayesian:

$$\Pr[X=1|X'=1]=\frac{\Pr[X'=1|X=1]\Pr[X=1]}{\Pr[X'=1]}=1-\epsilon.$$

- Bob learns less information about Alice's input. Indeed,

$$\mathrm{IC}_\mu(\pi_\epsilon)=1+(1-H(1-\epsilon))=2-H(\epsilon). \implies \mathrm{IC}_\mu(\mathrm{XOR},\epsilon)\le 2-H(\epsilon).$$

Analysis of π_ϵ

- Product distribution μ: $\Pr[X = 1] = \frac{1}{2}, \quad \Pr[Y = 1] = \frac{1}{2}$.
- After flipping, X' is still uniform: $\Pr[X' = 1] = \frac{1}{2}$;
- On receiving X', Bob is not sure what Alice's true input is.
- Bayesian:

$$\Pr[X = 1 | X' = 1] = \frac{\Pr[X' = 1 | X = 1] \Pr[X = 1]}{\Pr[X' = 1]} = 1 - \epsilon.$$

- Bob learns less information about Alice's input. Indeed,

$$\mathrm{IC}_\mu(\pi_\epsilon) = 1 + (1 - H(1 - \epsilon)) = 2 - H(\epsilon). \implies \mathrm{IC}_\mu(\mathrm{XOR}, \epsilon) \leq 2 - H(\epsilon).$$

Hence the information saved by allowing error ϵ:

$$\mathrm{IC}_\mu(\mathrm{XOR}, 0) - \mathrm{IC}_\mu(\mathrm{XOR}, \epsilon) \geq H(\epsilon).$$

Analysis of π_ϵ

- Product distribution μ: $\Pr[X = 1] = \frac{1}{2}, \quad \Pr[Y = 1] = \frac{1}{2}$.
- After flipping, X' is still uniform: $\Pr[X' = 1] = \frac{1}{2}$;
- On receiving X', Bob is not sure what Alice's true input is.
- Bayesian:

$$\Pr[X = 1 | X' = 1] = \frac{\Pr[X' = 1 | X = 1] \Pr[X = 1]}{\Pr[X' = 1]} = 1 - \epsilon.$$

- Bob learns less information about Alice's input. Indeed,

$$\mathrm{IC}_\mu(\pi_\epsilon) = 1 + (1 - H(1-\epsilon)) = 2 - H(\epsilon). \Longrightarrow \mathrm{IC}_\mu(\mathrm{XOR}, \epsilon) \leq 2 - H(\epsilon).$$

Hence the information saved by allowing error ϵ:

$$\mathrm{IC}_\mu(\mathrm{XOR}, 0) - \mathrm{IC}_\mu(\mathrm{XOR}, \epsilon) \geq H(\epsilon).$$ True for all functions f?

Main results

Theorem (Dagan-Filmus-Hatami-L'16)

For all functions f, for all distributions μ such that $\mathrm{IC}_\mu(f,0) > 0$,

$$\Omega(H(\epsilon)) \leq \mathrm{IC}_\mu(f,0) - \mathrm{IC}_\mu(f,\epsilon) \leq O(H(\sqrt{\epsilon})).$$

Theorem (Dagan-Filmus-Hatami-L'16)

For all μ such that $\mathrm{IC}_\mu(\mathrm{AND},0) > 0$,

$$\mathrm{IC}_\mu(\mathrm{AND},0) - \mathrm{IC}_\mu(\mathrm{AND},\epsilon) = \Theta(H(\epsilon)).$$

Proof idea of lower bound

Theorem (Dagan-Filmus-Hatami-L'16)

$$\mathrm{IC}_\mu(f, 0) - \mathrm{IC}_\mu(f, \epsilon) \geq \Omega(H(\epsilon)).$$

Proof idea of lower bound

Theorem (Dagan-Filmus-Hatami-L'16)

$$\mathrm{IC}_\mu(f, 0) - \mathrm{IC}_\mu(f, \epsilon) \geq \Omega(H(\epsilon)).$$

Proof idea - converting a zero-error protocol π to ϵ error π_ϵ

Proof idea of lower bound

Theorem (Dagan-Filmus-Hatami-L'16)

$$\mathrm{IC}_\mu(f, 0) - \mathrm{IC}_\mu(f, \epsilon) \geq \Omega(H(\epsilon)).$$

Proof idea - converting a zero-error protocol π to ϵ error π_ϵ

- Alice smartly flips her input X with probability ϵ to X';

Proof idea of lower bound

Theorem (Dagan-Filmus-Hatami-L'16)

$$\mathrm{IC}_\mu(f,0) - \mathrm{IC}_\mu(f,\epsilon) \geq \Omega(H(\epsilon)).$$

Proof idea - converting a zero-error protocol π to ϵ error π_ϵ

- Alice smartly flips her input X with probability ϵ to X';
- They compute f using π with (X', Y).

Proof idea of lower bound

Theorem (Dagan-Filmus-Hatami-L'16)

$$\mathrm{IC}_\mu(f, 0) - \mathrm{IC}_\mu(f, \epsilon) \geq \Omega(H(\epsilon)).$$

Proof idea - converting a zero-error protocol π to ϵ error π_ϵ

- Alice smartly flips her input X with probability ϵ to X';
- They compute f using π with (X', Y).
- Show

$$\mathrm{IC}_\mu(\pi) - \mathrm{IC}_\mu(\pi_\epsilon) \geq \Omega(H(\epsilon)).$$

Tight bound for AND

Theorem (Dagan-Filmus-Hatami-L'16)

$$\mathrm{IC}_\mu(\mathrm{AND}, 0) - \mathrm{IC}_\mu(\mathrm{AND}, \epsilon) \le O(H(\epsilon)).$$

Tight bound for AND

Theorem (Dagan-Filmus-Hatami-L'16)

$$\mathrm{IC}_\mu(\mathrm{AND}, 0) - \mathrm{IC}_\mu(\mathrm{AND}, \epsilon) \leq O(H(\epsilon)).$$

Some new ideas

- study the stability of protocols;
  may be useful to study other functions.
- parametrization of all distributions by product distributions.
  applicable to all functions!

Worst case error v.s. distributional error

Recall

$$\mathrm{IC}_\mu(f,\epsilon) := \inf_\pi \mathrm{IC}_\mu(\pi),$$

where π computes f with error ϵ for all inputs (X, Y): worst case error.

Worst case error v.s. distributional error

Recall

$$\mathrm{IC}_\mu(f, \epsilon) := \inf_\pi \mathrm{IC}_\mu(\pi),$$

where π computes f with error ϵ for all inputs (X, Y): worst case error.

Definition

$$\mathrm{IC}_\mu^D(f, \epsilon) := \inf_\pi \mathrm{IC}_\mu(\pi),$$

where π computes f with distributional error ϵ:

$$\Pr_{(X,Y)\sim\mu}[\pi(X, Y) \neq f(X, Y)] \leq \epsilon.$$

Worst case error v.s. distributional error

Recall

$$\mathrm{IC}_\mu(f, \epsilon) := \inf_\pi \mathrm{IC}_\mu(\pi),$$

where π computes f with error ϵ for all inputs (X, Y): worst case error.

Definition

$$\mathrm{IC}^D_\mu(f, \epsilon) := \inf_\pi \mathrm{IC}_\mu(\pi),$$

where π computes f with distributional error ϵ:

$$\Pr_{(X,Y)\sim\mu}[\pi(X, Y) \neq f(X, Y)] \leq \epsilon.$$

Obviously, $\mathrm{IC}^D_\mu(f, \epsilon) \leq \mathrm{IC}_\mu(f, \epsilon)$, and the inequality can be strict!

Prior-free information complexity

Definition [Braverman'12]

$$\mathrm{IC}(f, \epsilon) := \max_{\mu} \mathrm{IC}_{\mu}(f, \epsilon), \qquad \mathrm{IC}^{D}(f, \epsilon) := \max_{\mu} \mathrm{IC}_{\mu}^{D}(f, \epsilon).$$

Prior-free information complexity

Definition [Braverman'12]

$$\mathrm{IC}(f,\epsilon) := \max_{\mu} \mathrm{IC}_{\mu}(f,\epsilon), \qquad \mathrm{IC}^{D}(f,\epsilon) := \max_{\mu} \mathrm{IC}^{D}_{\mu}(f,\epsilon).$$

Obviously, $\mathrm{IC}^{D}(f,\epsilon) \leq \mathrm{IC}(f,\epsilon)$.

Prior-free information complexity

Definition [Braverman'12]

$$\mathrm{IC}(f,\epsilon) := \max_{\mu} \mathrm{IC}_\mu(f,\epsilon), \qquad \mathrm{IC}^D(f,\epsilon) := \max_{\mu} \mathrm{IC}^D_\mu(f,\epsilon).$$

Obviously, $\mathrm{IC}^D(f,\epsilon) \le \mathrm{IC}(f,\epsilon)$.

Theorem (Braverman'12)

If $\epsilon = 0$, i.e., no error,

$$\mathrm{IC}(f,0) = \mathrm{IC}^D(f,0).$$

If $\epsilon > 0$, then

$$\frac{1}{2}\mathrm{IC}(f,2\epsilon) \le \mathrm{IC}^D(f,\epsilon) \le \mathrm{IC}(f,\epsilon).$$

Prior-free information complexity

Definition [Braverman'12]

$$\mathrm{IC}(f,\epsilon) := \max_{\mu} \mathrm{IC}_{\mu}(f,\epsilon), \qquad \mathrm{IC}^{D}(f,\epsilon) := \max_{\mu} \mathrm{IC}^{D}_{\mu}(f,\epsilon).$$

Obviously, $\mathrm{IC}^{D}(f,\epsilon) \le \mathrm{IC}(f,\epsilon)$.

Theorem (Braverman'12)

If $\epsilon = 0$, i.e., no error,

$$\mathrm{IC}(f,0) = \mathrm{IC}^{D}(f,0).$$

If $\epsilon > 0$, then

$$\frac{1}{2}\mathrm{IC}(f,2\epsilon) \le \mathrm{IC}^{D}(f,\epsilon) \le \mathrm{IC}(f,\epsilon).$$

Q: $\mathrm{IC}^{D}(f,\epsilon) = \mathrm{IC}(f,\epsilon)$?

Main results - continued

Theorem (Dagan-Filmus-Hatami-L'16)

For n sufficiently large,

$$\mathrm{IC}^D(\mathrm{DISJ}_n, \epsilon) < \mathrm{IC}(\mathrm{DISJ}_n, \epsilon).$$

Implication on communication complexity

Theorem (BGPW'13)

$$\lim_{\epsilon \to 0} CC(\mathrm{DISJ}_n, \epsilon) \approx 0.4827n.$$

Implication on communication complexity

Theorem (BGPW'13)

$$\lim_{\epsilon \to 0} CC(\mathrm{DISJ}_n, \epsilon) \approx 0.4827n.$$

Theorem (Dagan-Filmus-Hatami-L'16)

$$CC(\mathrm{DISJ}_n, \epsilon) \approx (0.4827 - \Theta(H(\epsilon)))n.$$

Proof: follows from results on $\mathrm{IC}(\mathrm{AND}, \epsilon)$. Conjectured by [BGPW13].

Summary and Open Problems

Theorem (Dagan-Filmus-Hatami-L'16)

- $\Omega(H(\epsilon)) \leq \mathrm{IC}_\mu(f, 0) - \mathrm{IC}_\mu(f, \epsilon) \leq O(H(\sqrt{\epsilon}))$;
- $\mathrm{IC}^D(f, \epsilon) \neq \mathrm{IC}(f, \epsilon)$, *example:* DISJ_n;
- $CC(\mathrm{DISJ}_n, \epsilon) \approx (0.4827 - \Theta(H(\epsilon)))n$.

Summary and Open Problems

Theorem (Dagan-Filmus-Hatami-L'16)

- $\Omega(H(\epsilon)) \le \mathrm{IC}_\mu(f,0) - \mathrm{IC}_\mu(f,\epsilon) \le O(H(\sqrt{\epsilon}))$;
- $\mathrm{IC}^D(f,\epsilon) \neq \mathrm{IC}(f,\epsilon)$, *example:* DISJ_n;
- $CC(\mathrm{DISJ}_n,\epsilon) \approx (0.4827 - \Theta(H(\epsilon)))n$.

Open Problem

- $\mathrm{IC}_\mu(f,0) - \mathrm{IC}_\mu(f,\epsilon) = \Theta(H(\epsilon))$?
- $\lim_{n\to\infty} \frac{\mathrm{IC}(\mathrm{DISJ}_n,\epsilon)}{n} = \max_{\mu:\mu(11)=0} \mathrm{IC}_\mu(\mathrm{AND},\epsilon,1\to 0)$?