Essentially Conditional Information Inequalities

Tarik Kaced

(part of this work is joint work with Andrei Romashchenko)

Post-doctoral fellow at the Institute of Network Coding
The Chinese University of Hong Kong

April 17, 2013

First Workshop on Entropy and Information Inequalities

香港中文大學
The Chinese University of Hong Kong

INC

The “Messy” Part

Tarik Kaced

(part of this work is joint work with Andrei Romashchenko)

Post-doctoral fellow at the Institute of Network Coding
The Chinese University of Hong Kong

April 17, 2013

First Workshop on Entropy and Information Inequalities

香港中文大學
The Chinese University of Hong Kong

INC

What this talk is about

1. **Information Inequalities**
2. **Equivalence of Two Proofs Systems**
3. **Essentially Conditional Inequalities**
4. **Going further**

Information Inequalities

Shannon's Information Measures

Compulsory Slide

Conditional Entropy:

$$H(X|Y) = H(XY) - H(Y)$$

Mutual Information:

$$I(X{:}Y) = H(X) + H(Y) - H(XY)$$

Conditional Mutual Information:

$$I(X{:}Y|Z) = H(XZ) + H(YZ) - H(XYZ) - H(Z)$$

Shannon's Information Measures

Compulsory Slide

Conditional Entropy:

$$H(X|Y) = H(XY) - H(Y) \geq 0$$

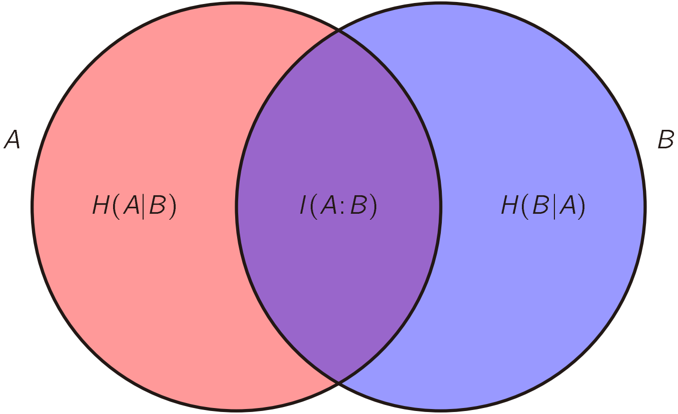
Mutual Information:

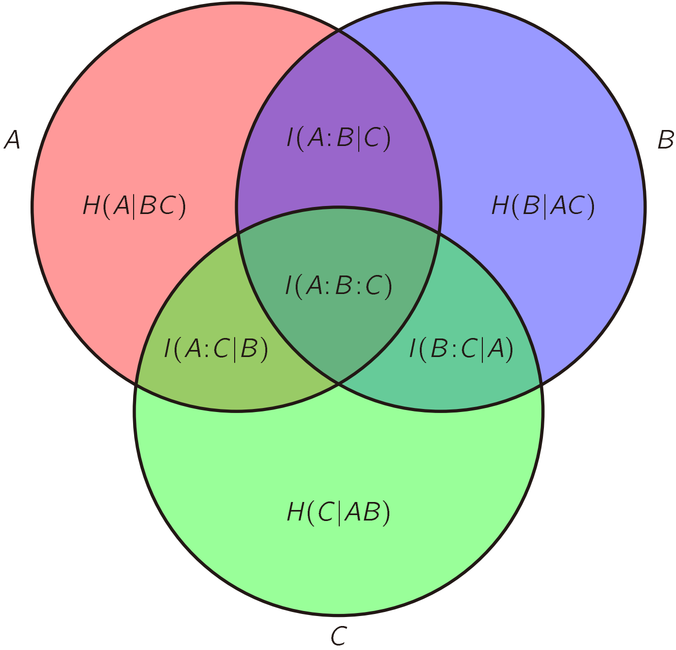
$$I(X:Y) = H(X) + H(Y) - H(XY) \geq 0$$

Conditional Mutual Information:

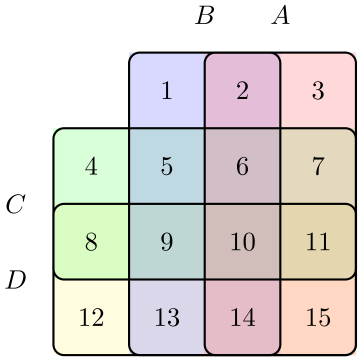
$$I(X:Y|Z) = H(XZ) + H(YZ) - H(XYZ) - H(Z) \geq 0$$

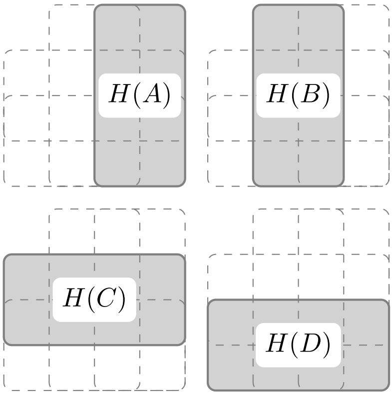
Shannon's Basic Inequality

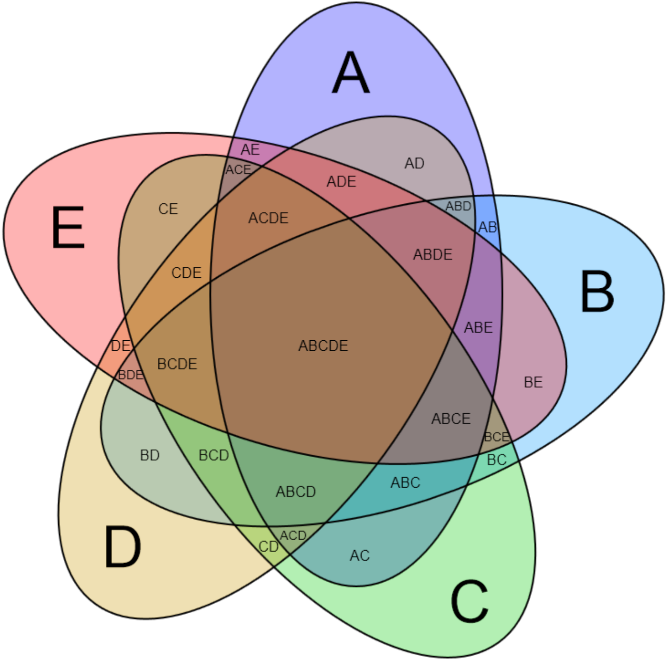

A
B
$H(A|B)$
$I(A:B)$
$H(B|A)$

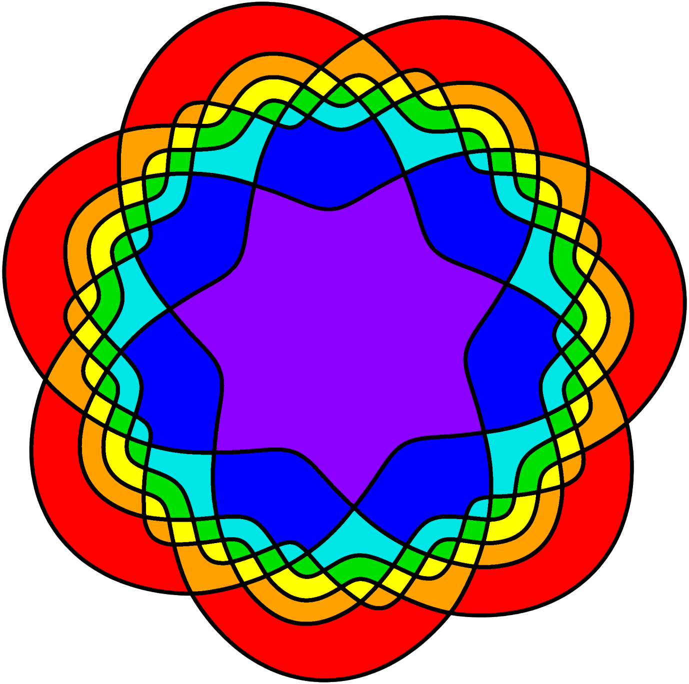

A
B
C
$H(A|BC)$
$H(B|AC)$
$H(C|AB)$
$I(A:B|C)$
$I(A:C|B)$
$I(B:C|A)$
$I(A:B:C)$

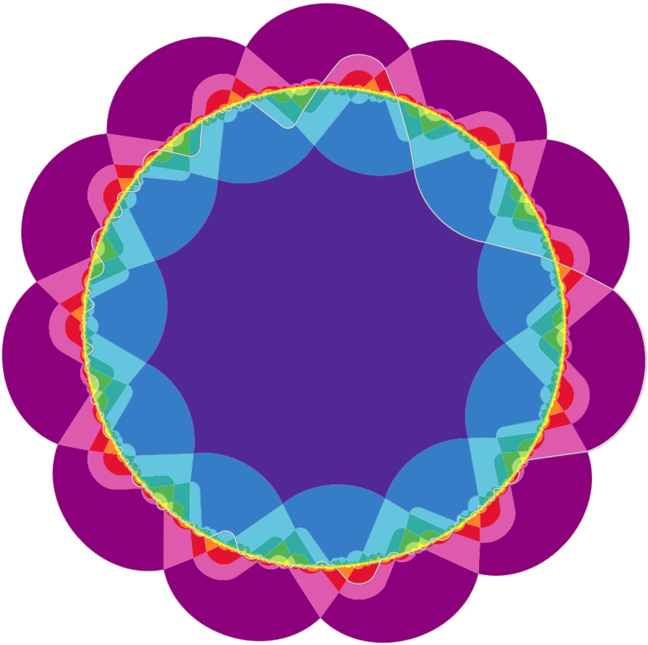
4-Information Diagram

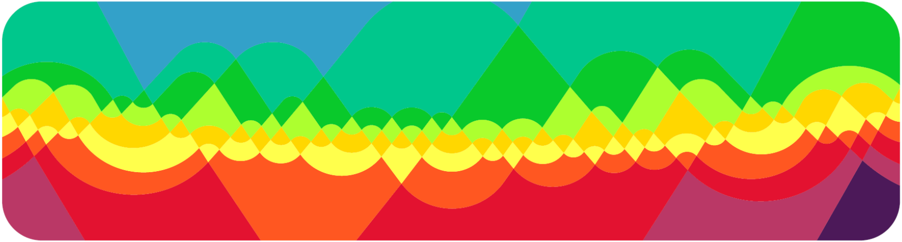





A
B
C
D
E
AE
ACE
AD
ADE
ABD
AB
CE
ACDE
ABDE
CDE
ABE
ABCDE
DE
BDE
BCDE
BE
ABCE
BCE
BC
BD
BCD
ABCD
ABC
ACD
CD
AC







Linear information inequalities

Pippenger (1986):
"What are the laws of Information Theory?"

Pippenger (1986):
“What are the laws of Information Theory?”

Basic inequality:

$$\begin{array}{ll} H(AB) \leq H(A) + H(B) & [I(A{:}B) \geq 0] \\ H(ABC) + H(C) \leq H(AC) + H(BC) & [I(A{:}B|C) \geq 0] \end{array}$$

Linear information inequalities

Pippenger (1986):
"What are the laws of Information Theory?"

Basic inequality:

$$H(AB) \leq H(A) + H(B) \qquad [I(A{:}B) \geq 0]$$
$$H(ABC) + H(C) \leq H(AC) + H(BC) \qquad [I(A{:}B|C) \geq 0]$$

Shannon-type inequalities: any positive combination of basic ineq., e.g.,

$$H(A) \leq H(A|B) + H(A|C) + I(B{:}C)$$

Linear information inequalities

Pippenger (1986):
"What are the laws of Information Theory?"

Basic inequality:

$$\begin{aligned} H(AB) &\leq H(A) + H(B) && [I(A{:}B) \geq 0] \\ H(ABC) + H(C) &\leq H(AC) + H(BC) && [I(A{:}B|C) \geq 0] \end{aligned}$$

Shannon-type inequalities: any positive combination of basic ineq., e.g.,

$$H(A) \leq H(A|B) + H(A|C) + I(B{:}C)$$

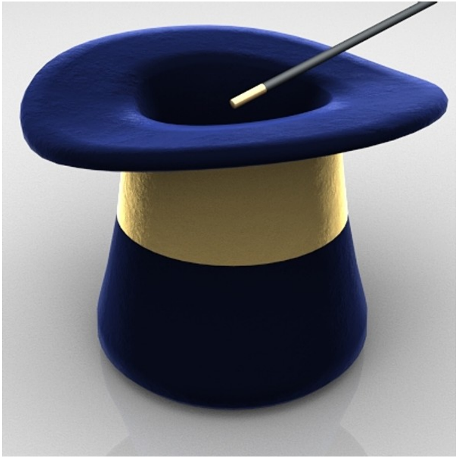
Non-Shannon-type inequalities, e.g., [Z. Zhang, R. W. Yeung, 1998] :

$$I(C{:}D) \leq 2I(C{:}D|A) + I(C{:}D|B) + I(A{:}B) + I(A{:}C|D) + I(A{:}D|C)$$

How to prove non-Shannon inequalities?

Sautéed ZY98 à la DFZ

Sautéed ZY98 à la DFZ



Sautéed ZY98 à la DFZ

The following is a Shannon-type inequality:

$$
\begin{aligned}
I(C{:}D) \leq I(C{:}D|A) + I(C{:}D|B) + I(A{:}B) +& \\
+ I(C{:}D|Z) + I(Z{:}C|D) + I(Z{:}D|C) +& \\
+ 9811 I(Z{:}AB|CD)&
\end{aligned}
$$

Sautéed ZY98 à la DFZ

The following is a Shannon-type inequality:

$$
\begin{aligned}
I(C:D) \leq I(C:D|A) + I(C:D|B) + I(A:B) +& \\
+ I(C:D|Z) + I(Z:C|D) + I(Z:D|C) +& \\
+ 3I(Z:AB|CD)&
\end{aligned}
$$

Sautéed ZY98 à la DFZ

The following is a Shannon-type inequality:

$$\begin{aligned} I(C:D) \leq I(C:D|A) + I(C:D|B) + I(A:B) + \\ + I(C:D|Z) + I(Z:C|D) + I(Z:D|C) + \\ + 3I(Z:AB|CD) \end{aligned}$$

IDEA: Take Z to be a B-copy of A over CD:

Sautéed ZY98 à la DFZ

The following is a Shannon-type inequality:

$$
\begin{aligned}
I(C\!:\!D) \leq I(C\!:\!D|A) + I(C\!:\!D|B) + I(A\!:\!B) &+ \\
+ I(C\!:\!D|Z) + I(Z\!:\!C|D) + I(Z\!:\!D|C) &+ \\
+ 3I(Z\!:\!AB|CD)&
\end{aligned}
$$

IDEA: Take Z to be a B-copy of A over CD:

$$
\begin{aligned}
I(C\!:\!D) \leq I(C\!:\!D|A) + I(C\!:\!D|B) + I(A\!:\!B) &+ \\
+ I(C\!:\!D|A) + I(A\!:\!C|D) + I(A\!:\!D|C)&
\end{aligned}
$$

MMRV

The following is a Shannon-type inequality:

$$H(Z) \quad \leq I(C:D|A) + I(C:D|B) + I(A:B) + 2H(Z|C) + 2H(Z|D)$$

MMRV

The following is a Shannon-type inequality:

$$\begin{aligned}
H(Z) &\leq I(C{:}D|A) + I(C{:}D|B) + I(A{:}B) + 2H(Z|C) + 2H(Z|D) \\
H(Z) &\leq H(Z|A) + H(Z|B) + I(A{:}B) \\
H(Z|A) &\leq H(Z|C) + H(Z|D) + I(C{:}D|A) \\
H(Z|B) &\leq H(Z|C) + H(Z|D) + I(C{:}D|B)
\end{aligned}$$

MMRV

The following is a Shannon-type inequality:

$$
\begin{aligned}
H(Z) &\leq I(C{:}D|A) + I(C{:}D|B) + I(A{:}B) + 2H(Z|C) + 2H(Z|D) \\
H(Z) &\leq H(Z|A) + H(Z|B) + I(A{:}B) \\
H(Z|A) &\leq H(Z|C) + H(Z|D) + I(C{:}D|A) \\
H(Z|B) &\leq H(Z|C) + H(Z|D) + I(C{:}D|B)
\end{aligned}
$$

IDEA: Take Z to be a common information between C and D

$$
H(W|C) = H(W|D) = 0
$$

$$
H(W) = I(C{:}D)
$$

The following is a Shannon-type inequality:

$$
\begin{aligned}
H(Z) &\leq I(C{:}D|A) + I(C{:}D|B) + I(A{:}B) + 2H(Z|C) + 2H(Z|D)\\
H(Z) &\leq H(Z|A) + H(Z|B) + I(A{:}B)\\
H(Z|A) &\leq H(Z|C) + H(Z|D) + I(C{:}D|A)\\
H(Z|B) &\leq H(Z|C) + H(Z|D) + I(C{:}D|B)
\end{aligned}
$$

IDEA: Take Z to be a common information between C and D

$$
\begin{aligned}
H(W|C) = H(W|D) &= 0\\
H(W) &= I(C{:}D)
\end{aligned}
$$

$$I(C{:}D) \leq I(C{:}D|A) + I(C{:}D|B) + I(A{:}B)$$

MMRV

The following is a Shannon-type inequality:

$$
\begin{aligned}
H(Z) &\le I(C{:}D|A) + I(C{:}D|B) + I(A{:}B) + 2H(Z|C) + 2H(Z|D)\\
H(Z) &\le H(Z|A) + H(Z|B) + I(A{:}B)\\
H(Z|A) &\le H(Z|C) + H(Z|D) + I(C{:}D|A)\\
H(Z|B) &\le H(Z|C) + H(Z|D) + I(C{:}D|B)
\end{aligned}
$$

IDEA: Take Z to be a common information between C and D

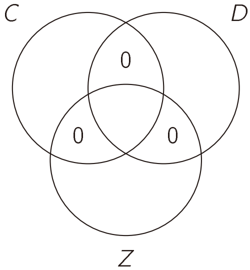
$$
\begin{aligned}
H(W|C) = H(W|D) = 0\\
H(W) = I(C{:}D)
\end{aligned}
$$

$$I(C{:}D) \le I(C{:}D|A) + I(C{:}D|B) + I(A{:}B)$$

Wait...Such a common information does not exist in general!

Extractability Criterion for Triples

(Romashchenko)

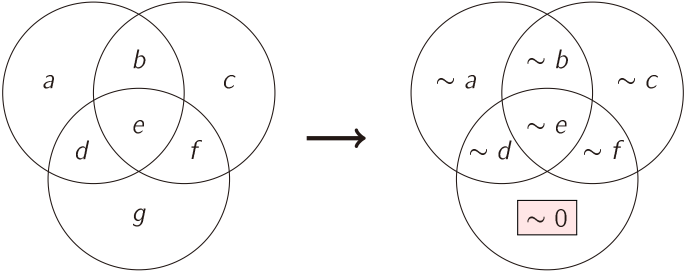


$\Leftrightarrow$

There is a common information W for the random variables C, D, Z.

Common Information vs. Mutual Information 2/2

We can still "approximately" extract mutual information.
(indep. by Ahlswede/Gacs/Korner, Wyner rediscovered by Zhang, Romashchenko, Chan...)

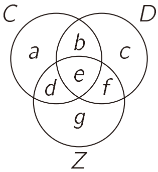


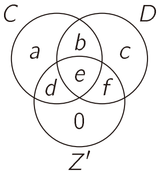
- From the diagram on the left, we can get the diagram of the right up to any given precision.
- Can be generalized to n variables.

MMRV (really)

The following is a Shannon-type inequality:

$$H(Z) \leq I(C{:}D|A) + I(C{:}D|B) + I(A{:}B) + 2H(Z|C) + 2H(Z|D)$$



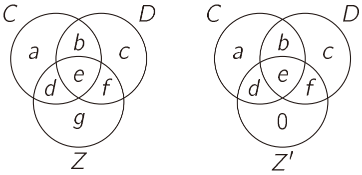


IDEA: Take Z' as in diagram

MMRV (really)

The following is a Shannon-type inequality:

$$H(Z) \leq I(C{:}D|A) + I(C{:}D|B) + I(A{:}B) + 2H(Z|C) + 2H(Z|D)$$



IDEA: Take Z' as in diagram

$$I(C{:}D) \leq I(C{:}D|A) + I(C{:}D|B) + I(A{:}B) + I(C{:}D|Z) + I(Z{:}C|D) + I(Z{:}D|C)$$

Equivalence of Two Proofs Systems

Rule ZY

(A) Let f and g be linear maps on entropies such that

$$f(X_{\mathcal{N}}, Y_{\mathcal{M}}) + g(Y_{\mathcal{M}}, Z) + \alpha I(Z : X_{\mathcal{N}} | Y_{\mathcal{M}}) \geq 0,$$

for some $\alpha \geq 0$;

(B) then the following (stronger) inequality is also valid:

$$f(X_{\mathcal{N}}, Y_{\mathcal{M}}) + g(Y_{\mathcal{M}}, Z) \geq 0.$$

Inference Rules 1/2

Rule ZY

(A) Let f and g be linear maps on entropies such that

$$f(X_{\mathcal{N}}, Y_{\mathcal{M}}) + g(Y_{\mathcal{M}}, Z) + \alpha I(Z{:}X_{\mathcal{N}}|Y_{\mathcal{M}}) \geq 0,$$

for some $\alpha \geq 0$;

(B) then the following (stronger) inequality is also valid:

$$f(X_{\mathcal{N}}, Y_{\mathcal{M}}) + g(Y_{\mathcal{M}}, Z) \geq 0.$$

$$\begin{aligned} & I(C{:}D|A) + I(C{:}D|B) + I(A{:}B) + \\ & \quad + I(C{:}D|Z) + I(Z{:}C|D) + I(Z{:}D|C) - I(C{:}D) + \\ & \quad\quad + 3I(Z{:}AB|CD) \geq 0 \end{aligned}$$

Inference Rules 2/2

Rule MMRV

(A) Let f and g be linear maps on entropies such that

$$f(X_{\mathcal{N}}, Y_{\mathcal{M}}) + g(Y_{\mathcal{M}}, Z) \geq 0;$$

(B) then the following (stronger) inequality is also valid:

$$f(X_{\mathcal{N}}, Y_{\mathcal{M}}) + g(Y_{\mathcal{M}}, Z) - r_Z H(Z|Y_{\mathcal{M}}) \geq 0,$$

where r_Z is the sum of coefficients invoving Z.

Inference Rules 2/2

Rule MMRV

(A) Let f and g be linear maps on entropies such that

$$f(X_{\mathcal{N}}, Y_{\mathcal{M}}) + g(Y_{\mathcal{M}}, Z) \geq 0;$$

(B) then the following (stronger) inequality is also valid:

$$f(X_{\mathcal{N}}, Y_{\mathcal{M}}) + g(Y_{\mathcal{M}}, Z) - r_Z H(Z|Y_{\mathcal{M}}) \geq 0,$$

where r_Z is the sum of coefficients invoving Z.

$$I(C{:}D|A) + I(C{:}D|B) + I(A{:}B) + 2H(Z|C) + 2H(Z|D) - H(Z) \geq 0$$

Proof systems

Definition

A *proof system* (for inequalities) consists of a *pool* P of inequalities and a rule T. A (computation) *step* in a proof system is described as follows:

1. Pick an inequality (A) from the convex closure of P;
2. Apply rule T to (A) and infer inequality (B);
3. Add (B) to the pool P.

A *derivation* is a sequence of valid steps in a system. An inequality $(\mathcal{I})$ is *provable in system* S if it belongs to the convex closure of the pool of S after some derivation.

- SYSTEM ZY: the system using RULE ZY.
- SYSTEM MMRV: the system using RULE MMRV.

More on Information Inequalities

Theorem (Balanced Inequalities, Chan)

1. *The inequality*

$$\sum_{\emptyset \neq J \subseteq \mathcal{N}} c_J H(X_J) \geq 0$$

is a valid information inequality.

2. *The inequality*

$$\sum_{\emptyset \neq J \subseteq \mathcal{N}} c_J H(X_J) - \sum_{i \in \mathcal{N}} r_i H(X_i | X_{\mathcal{N}-i}) \geq 0,$$

where r_J is the sum of all c_J involving J, is a valid balanced information inequality.

In other words, we can always assume that

$$\forall i \in \mathcal{N}, \; r_i = \sum_{i \in J \subseteq \mathcal{N}} c_J = 0.$$

Toy example:

Proposition

The basic inequality

$$I(X_I\!:\!X_J|X_K) \geq 0$$

is balanced iff $I \cap J \subset K$.
Any instance can be put in balanced form as a some of two other instances.

Balanced Shannon-type Inequalities

Toy example:

Proposition

The basic inequality

$$I(X_I : X_J | X_K) \geq 0$$

is balanced iff $I \cap J \subset K$.
Any instance can be put in balanced form as a some of two other instances.

- Simpler and direct proof that Shannon-type inequalities can be "balanced".
- Note: The elemental inequality $H(X_i | X_{N-i}) \geq 0$ is not balanced.

Equivalence Modulo Balancing

Theorem (informal)

SYSTEM ZY *and* SYSTEM MMRV *prove the same balanced inequalities*

Equivalence Modulo Balancing

Theorem (informal)

SYSTEM ZY *and* SYSTEM MMRV *prove the same balanced inequalities*

$$\begin{tikzcd}
(A_1) \arrow[dr, "\mathrm{ZY}"] & & (A_2) \arrow[dl, "\mathrm{MMRV}"'] \\
& (B_{[b]}) &
\end{tikzcd}$$

Equivalence Modulo Balancing

Theorem (informal)

SYSTEM ZY *and* SYSTEM MMRV *prove the same balanced inequalities*

$$
\begin{tikzcd}
& (A_b) \arrow[dl, Rightarrow] \arrow[dr, Rightarrow] & \\
(A_1) \arrow[dr, "\mathrm{ZY}"] & & (A_2) \arrow[dl, "\mathrm{MMRV}"'] \\
& (B_{[b]}) &
\end{tikzcd}
$$

Essentially Conditional Inequalities

Framework definitions

For n random variables, there $2^n - 1$ possible entropies.
When $n = 3$, there are 7 possible joint entropies:

$$(H(A), H(B), H(C), H(AB), H(AC), H(BC), H(ABC)) \in \mathbb{R}^7$$

Such a vector of entropies is called an **entropic point**.
An **almost entropic point** is the limit of a sequence of entropic points.

Disclaimer: This notation is handy so we (ab)use it and use, e.g., $H(AB)$ for the corresponding value of an almost entropic vector (even if it does not correspond to any distribution entropy)

Story / Motivation

- Closed pointed convex cones
- Points inside the cone are entropic, *i.e.* the difference is at the boundary.
- How much difference?
- Hard: The difference can be significant
- Simpler: What do the faces looks like?

What are the valid (conditional) inequalities on faces (subcone)?

Conditional information inequalities

If [some linear constraints for entropies]
then [a linear inequality for entropies].

Conditional information inequalities

If [some linear constraints for entropies]
then [a linear inequality for entropies].

- **Example 1 (trivial):** If $I(B:C) = 0$, then $H(A) \leq H(A|B) + H(A|C)$.
  Explanation:
  $H(A) \leq H(A|B) + H(A|C) + I(B:C)$.

Conditional information inequalities

If [some linear constraints for entropies]
then [a linear inequality for entropies].

- **Example 1 (trivial):** If $I(B{:}C) = 0$, then $H(A) \leq H(A|B) + H(A|C)$.
  Explanation:
  $H(A) \leq H(A|B) + H(A|C) + I(B{:}C)$.
- **Example 2 (trivial):** If $I(C{:}D|E) = I(C{:}E|D) = I(D{:}E|C) = 0$, then
  $I(C{:}D) \leq I(C{:}D|A) + I(C{:}D|B) + I(A{:}B)$.
  Explanation:
  $I(C{:}D) \leq I(C{:}D|A) + I(C{:}D|B) + I(A{:}B) + I(C{:}D|E) + I(C{:}E|D) + I(D{:}E|C)$.

Conditional information inequalities

If [some linear constraints for entropies]
then [a linear inequality for entropies].

- **Example 1 (trivial):** If $I(B{:}C) = 0$, then $H(A) \leq H(A|B) + H(A|C)$.
  Explanation:
  $H(A) \leq H(A|B) + H(A|C) + I(B{:}C)$.

- **Example 2 (trivial):** If $I(C{:}D|E) = I(C{:}E|D) = I(D{:}E|C) = 0$, then
  $I(C{:}D) \leq I(C{:}D|A) + I(C{:}D|B) + I(A{:}B)$.
  Explanation:
  $I(C{:}D) \leq I(C{:}D|A) + I(C{:}D|B) + I(A{:}B) + I(C{:}D|E) + I(C{:}E|D) + I(D{:}E|C)$.

- **Example 3 (nontrivial) [Zhang–Yeung 1997]:** If $I(A{:}B) = I(A{:}B|C) = 0$, then
  $I(C{:}D) \leq I(C{:}D|A) + I(C{:}D|B)$.
  Any explanation???

Other Conditional Inequalities

Theorem (Matus)

$$
\begin{aligned}
I(C:D) \leq\ & I(C:D|A) + I(C:D|B) + I(A:B) \\
& + I(A:C|E) + I(A:E|C) + \frac{1}{k} I(C:E|A) + \frac{k-1}{2}[I(A:D|C) + I(A:C|D)].
\end{aligned}
$$

Corollary

If $I(A:C|D) = I(A:D|C) = 0$ *then*

$$I(C:D) \leq I(C:D|A) + I(C:D|B) + I(A:B) + I(A:C|E) + I(A:E|C).$$

This conditional inequality hold (not only for entropic but also) for almost entropic points.

Other Conditional Inequalities

Theorem (Matus)

$$I(C:D) \leq I(C:D|A) + I(C:D|B) + I(A:B) + I(A:C|E) + I(A:E|C) + \frac{1}{k} I(C:E|A) + \frac{k-1}{2}[I(A:D|C) + I(A:C|D)].$$

Corollary

If $I(A:C|D) = I(A:D|C) = 0$ *then*

$$I(C:D) \leq I(C:D|A) + I(C:D|B) + I(A:B) + I(A:C|E) + I(A:E|C).$$

This conditional inequality hold (not only for entropic but also) for almost entropic points.

Two 4-variable conditional inequalities **are valid** for all almost entropic points

Other Conditional Inequalities

$$\underbrace{I(A{:}B) = I(A{:}B|C) = 0}_{\text{[Zhang–Yeung 97]}}$$

$$\underbrace{\begin{array}{c} I(A{:}B|C) = I(B{:}D|C) = 0, \text{ or} \\ I(A{:}C|D) = I(A{:}D|C) = 0, \text{ or} \\ I(A{:}C|D) = I(C{:}D|A) = 0 \end{array}}_{\text{[Matúš 99/2007]}}$$

$$\underbrace{H(C|A,B) = I(A{:}B|C) = 0}_{\text{[Romashchenko,K. 2011]}}$$

$$\underbrace{I(C{:}D) \leq I(C{:}D|A) + I(C{:}D|B) + I(A{:}B)}_{\text{[Ingleton 69]}}$$

Other Conditional Inequalities

$$\underbrace{I(A{:}B) = I(A{:}B|C) = 0}_{\text{[Zhang–Yeung 97]}} \qquad \underbrace{\begin{array}{l} I(A{:}B|C) = I(B{:}D|C) = 0, \text{ or} \\ I(A{:}C|D) = I(A{:}D|C) = 0, \text{ or} \\ I(A{:}C|D) = I(C{:}D|A) = 0 \end{array}}_{\text{[Matúš 99/2007]}} \qquad \underbrace{H(C|A,B) = I(A{:}B|C) = 0}_{\text{[Romashchenko,K. 2011]}}$$

$$\searrow \qquad \downarrow \qquad \swarrow$$

$$\underbrace{I(C{:}D) \leq I(C{:}D|A) + I(C{:}D|B) + I(A{:}B)}_{\text{[Ingleton 69]}}$$

Theorem (Romashchenko, K. 2011/2012)

All of these statements are essentially conditional inequalities.

Essentially Conditional Inequalities

- **Z. Zhang, R. W. Yeung 97:**
  if $I(A{:}B) = I(A{:}B|C) = 0$, then $I(C:D) \leq I(C:D|A) + I(C:D|B)$.

Essentially Conditional Inequalities

- **Z. Zhang, R. W. Yeung 97:**
  if $I(A{:}B) = I(A{:}B|C) = 0$, then $I(C{:}D) \leq I(C{:}D|A) + I(C{:}D|B)$.

- **Theorem** [Romashchenko, K. 2011] This inequality is *essentially conditional*, i.e., for all κ_1, κ_2 the inequality:

$$I(C{:}D) \leq I(C{:}D|A) + I(C{:}D|B) + \kappa_1 I(A{:}B) + \kappa_2 I(A{:}B|C)$$

  is not valid.

Proof by ad-hoc example

Claim: For any κ_1, κ_2 there exist (A, B, C, D) such that:

$$I(C{:}D) \not\leq I(C{:}D|A) + I(C{:}D|B) + \kappa_1 I(A{:}B) + \kappa_2 I(A{:}B|C)$$

Proof by ad-hoc example

Claim: For any κ_1, κ_2 there exist (A, B, C, D) such that:

$$I(C{:}D) \nleq I(C{:}D|A) + I(C{:}D|B) + \kappa_1 I(A{:}B) + \kappa_2 I(A{:}B|C)$$

Proof:

a	b	c	d	$\text{Prob}[a, b, c, d]$
0	0	0	1	$(1-\varepsilon)/4$
0	1	0	0	$(1-\varepsilon)/4$
1	0	0	1	$(1-\varepsilon)/4$
1	1	0	1	$(1-\varepsilon)/4$
1	0	1	1	ε

Proof by ad-hoc example

Claim: For any κ_1, κ_2 there exist (A, B, C, D) such that:

$$I(C{:}D) \not\leq I(C{:}D|A) + I(C{:}D|B) + \kappa_1 I(A{:}B) + \kappa_2 I(A{:}B|C)$$

Proof:

a	b	c	d	Prob$[a, b, c, d]$
0	0	0	1	$(1-\varepsilon)/4$
0	1	0	0	$(1-\varepsilon)/4$
1	0	0	1	$(1-\varepsilon)/4$
1	1	0	1	$(1-\varepsilon)/4$
1	0	1	1	ε

$$\begin{array}{ccccccccc} I(C{:}D) & \not\leq & I(C{:}D|A) & + & I(C{:}D|B) & + & \kappa_1 I(A{:}B) & + & \kappa_2 I(A{:}B|C) \\ \| & & \| & & \| & & \| & & \| \\ & & 0 & & 0 & & & & 0 \end{array}$$

Proof by ad-hoc example

Claim: For any κ_1, κ_2 there exist (A, B, C, D) such that:

$$I(C{:}D) \nleq I(C{:}D|A) + I(C{:}D|B) + \kappa_1 I(A{:}B) + \kappa_2 I(A{:}B|C)$$

Proof:

a	b	c	d	$\mathrm{Prob}[a, b, c, d]$
0	0	0	1	$(1-\varepsilon)/4$
0	1	0	0	$(1-\varepsilon)/4$
1	0	0	1	$(1-\varepsilon)/4$
1	1	0	1	$(1-\varepsilon)/4$
1	0	1	1	ε

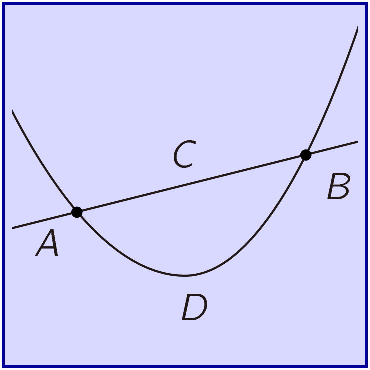
$$\begin{array}{ccccccccc} I(C{:}D) & \nleq & I(C{:}D|A) & + & I(C{:}D|B) & + & \kappa_1 I(A{:}B) & + & \kappa_2 I(A{:}B|C) \\ \| & & \| & & \| & & \| & & \| \\ \Theta(\varepsilon) & \nleq & 0 & + & 0 & + & O(\kappa_1 \varepsilon^2) & + & 0 \end{array}$$

Proof by geometric example

Construction of $(A, B, C, D)_q$

On the affine plane over $\mathbb{F}_q$:

1. Pick a random a non-vertical line C.
2. Pick two random points A and B on C.
3. Pick a random non-degenerate parabola D intersecting C exactly at A and B.

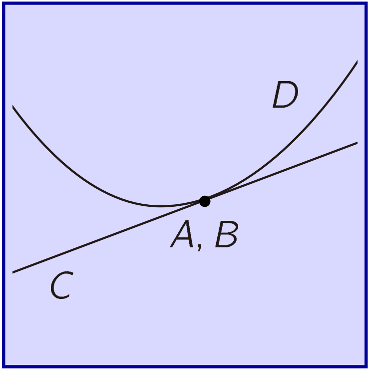


Proof by geometric example

Construction of $(A, B, C, D)_q$

On the affine plane over $\mathbb{F}_q$:

1. Pick a random a non-vertical line C.
2. Pick two random points A and B on C.
3. Pick a random non-degenerate parabola D intersecting C exactly at A and B.

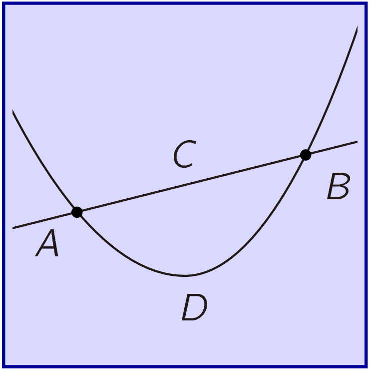


Proof by geometric example

Construction of $(A, B, C, D)_q$

On the affine plane over $\mathbb{F}_q$:

1. Pick a random a non-vertical line C.
2. Pick two random points A and B on C.
3. Pick a random non-degenerate parabola D intersecting C exactly at A and B.

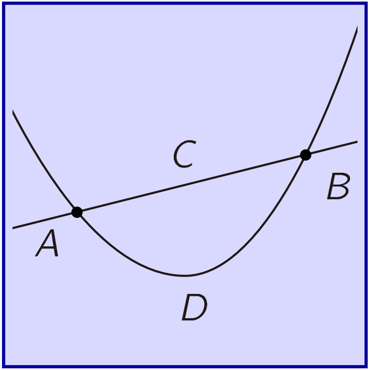


Proof by geometric example

Construction of $(A, B, C, D)_q$

On the affine plane over $\mathbb{F}_q$:

1. Pick a random a non-vertical line C.
2. Pick two random points A and B on C.
3. Pick a random non-degenerate parabola D intersecting C exactly at A and B.



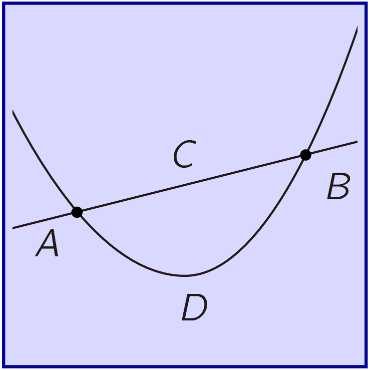
$$I(C{:}D) \leq \kappa[I(C{:}D|A) + I(C{:}D|B) + I(A{:}B) + I(A{:}B|C) + H(C|AB)]$$

Proof by geometric example

Construction of $(A, B, C, D)_q$

On the affine plane over $\mathbb{F}_q$:

1. Pick a random a non-vertical line C.
2. Pick two random points A and B on C.
3. Pick a random non-degenerate parabola D intersecting C exactly at A and B.



$$I(C{:}D) \leq \kappa[I(C{:}D|A) + I(C{:}D|B) + I(A{:}B) + I(A{:}B|C) + H(C|AB)]$$

$$1 + \frac{1}{q} \not\leq O\left(\kappa \frac{\log q}{q}\right)$$

Non-robustness of some conditional inequalities

$$I(A{:}B) = I(A{:}B|C) = 0 \Rightarrow I(C{:}D) \le I(C{:}D|A) + I(C{:}D|B) \qquad \text{(ZY97)}$$

In fact we have a stronger result. Let $\epsilon > 0$, assume

- $0 < I(A{:}B) \le \epsilon$.
- $0 < I(A{:}B|C) \le \epsilon$.
- $0 < H(ABCD) = const$.

Non-robustness of some conditional inequalities

$$I(A{:}B) = I(A{:}B|C) = 0 \Rightarrow I(C{:}D) \leq I(C{:}D|A) + I(C{:}D|B) \qquad \text{(ZY97)}$$

In fact we have a stronger result. Let $\epsilon > 0$, assume

- $0 < I(A{:}B) \leq \epsilon$.
- $0 < I(A{:}B|C) \leq \epsilon$.
- $0 < H(ABCD) = const$.

Then the ratio

$$\frac{I(C{:}D)}{I(C{:}D|A) + I(C{:}D|B)}$$

can be made arbitrarily large.

Non-robustness of some conditional inequalities

$$I(A{:}B) = I(A{:}B|C) = 0 \Rightarrow I(C{:}D) \leq I(C{:}D|A) + I(C{:}D|B) \qquad \text{(ZY97)}$$

In fact we have a stronger result. Let $\epsilon > 0$, assume

- $0 < I(A{:}B) \leq \epsilon$.
- $0 < I(A{:}B|C) \leq \epsilon$.
- $0 < H(ABCD) = const$.

Then the ratio

$$\frac{I(C{:}D)}{I(C{:}D|A) + I(C{:}D|B)}$$

can be made arbitrarily large.

Non-robustness of some conditional inequalities

$$I(A{:}B) = I(A{:}B|C) = 0 \Rightarrow I(C{:}D) \leq I(C{:}D|A) + I(C{:}D|B) \qquad \text{(ZY97)}$$

In fact we have a stronger result. Let $\epsilon > 0$, assume

- $0 < I(A{:}B) \leq \epsilon$.
- $0 < I(A{:}B|C) \leq \epsilon$.
- $0 < H(ABCD) = const$.

Then the ratio

$$\frac{I(C{:}D)}{I(C{:}D|A) + I(C{:}D|B)}$$

can be made arbitrarily large.

Non-robustness of some conditional inequalities

$$I(A{:}B) = I(A{:}B|C) = 0 \Rightarrow I(C{:}D) \leq I(C{:}D|A) + I(C{:}D|B) \qquad \text{(ZY97)}$$

In fact we have a stronger result. Let $\epsilon > 0$, assume

- $0 < I(A{:}B) \leq \epsilon$.
- $0 < I(A{:}B|C) \leq \epsilon$.
- $0 < H(ABCD) = const$.

Then the ratio

$$\frac{I(C{:}D)}{I(C{:}D|A) + I(C{:}D|B)}$$

can be made arbitrarily large.

For almost entropic points

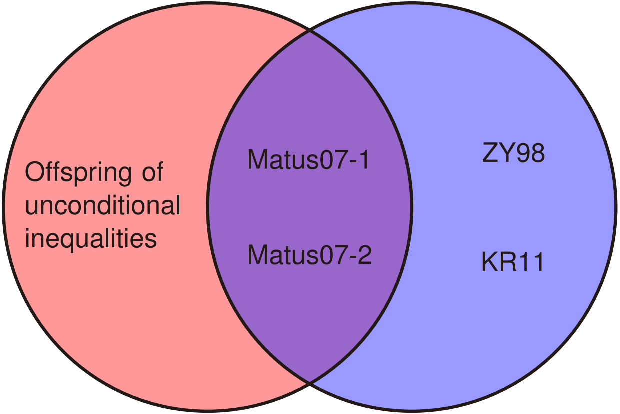
Theorem (Romashchenko, K. 2012)

Two essentially conditional inequalities ***are not valid*** *for all almost entropic points*

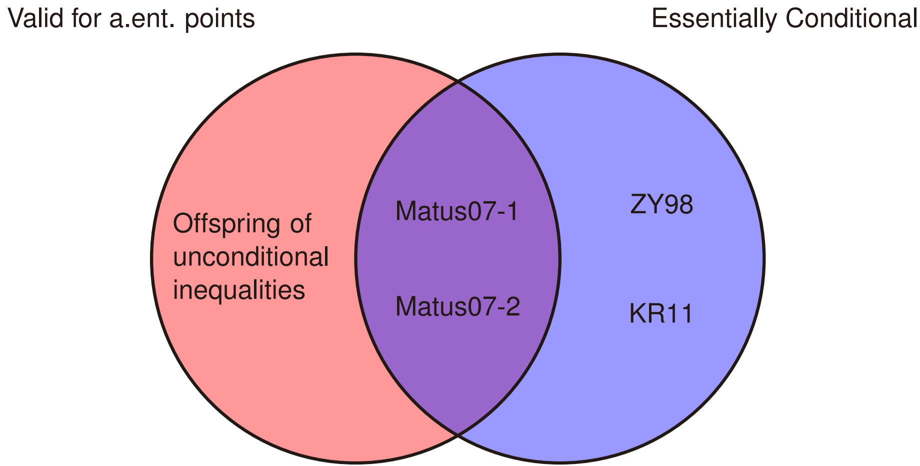
The current essentially conditional zoo

Valid for a.ent. points

Essentially Conditional



The current essentially conditional zoo



Where does Matus99 belong ?

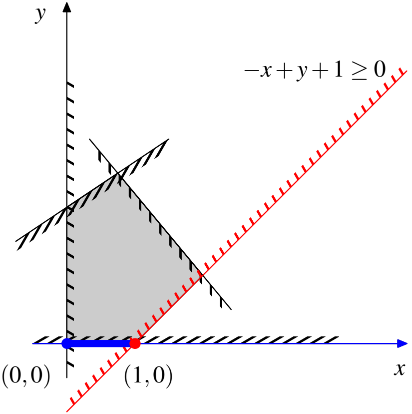
Geometric Interpretations

Convexity has an immensely rich structure and numerous applications. On the other hand, almost every "convex" idea can be explained by a two-dimensional picture.

- Alexander Barvinok

Geometric interpretation 1/3

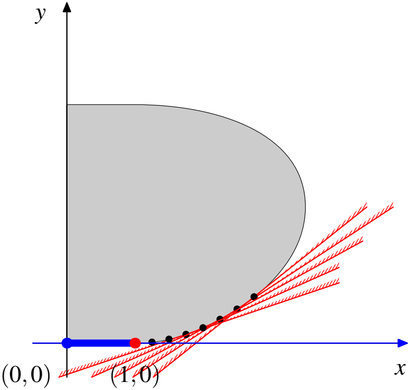
For (x, y) in the gray set: if $y = 0$ then $x \leq 1$



A trivial conditional inequality can be extended to an unconditional one.

Geometric interpretation 2/3

For (x, y) in the gray set: if $y = 0$ then $x \leq 1$



This conditional inequality is implied by an infinite family of tangent half-planes.

A Corollary

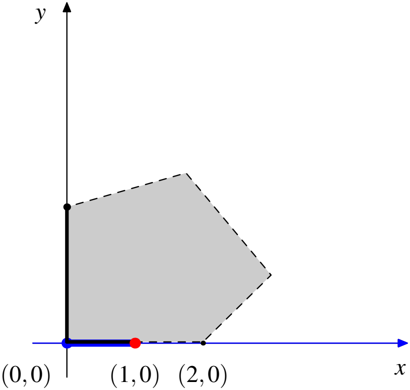
Theorem: There exist **essentially conditional** inequalities that hold for almost entropic points.

$$\Downarrow$$

Theorem [Matúš 07] **The cone** of linear information inequalities with 4 random variables **is not polyhedral**, i.e., there exist infinitely many independent linear information inequalities.

Geometric interpretation 3/3

For (x, y) in the gray set: if $y = 0$ then $x \leq 1$



For the closure of this set, with the same constraint $y = 0$ we only have $x \leq 2$.

Going further

Other Frameworks

Frameworks with the same underlying inequalities

Framework	Objects	Projection	Quantity
Quantum Entropy	systems	subsystem	Quantum Entropy
Kolmogorov	strings	subtuples	Kolmogorov Complexity
Information Theory	Random variables	subtuples	Shannon Entropy
Group Theory	groups	subgroups	log size
Combinatorial Arrays	Arrays	subarrays	log size
...	...	...	...
Vector spaces	subspaces	rank dimension	intersection

Other Frameworks

Frameworks with the same underlying inequalities

Framework	Objects	Projection	Quantity
Quantum Entropy	systems	subsystem	Quantum Entropy
Kolmogorov	strings	subtuples	Kolmogorov Complexity
Information Theory	Random variables	subtuples	Shannon Entropy
Group Theory	groups	subgroups	log size
Combinatorial Arrays	Arrays	subarrays	log size
...	...	...	...
Vector spaces	subspaces	rank dimension	intersection

- **"Reverse mathematics" (philosophical) question:**
  Is there a common set of axioms that induce these inequalities ?
- **Question (obvious extension):**
  Since unconditional inequalities are the same in every framework : What about their conditional inequalities ?

Kolmogorov Complexity

Counterpart to Kolmogorov Complexity

Fix an acceptable programming system.
For any binary strings x, y:

$$C(x) = \text{length of a shortest program printing } x,$$
$$C(x|y) = \text{length of a shortest program printing } x \text{ given input } y.$$

And up to $O(\log |xy|)$,

$$C(x) \geq 0,$$
$$C(x|y) \geq 0,$$
$$C(x) + C(y) \geq C(x, y).$$

Counterpart to Kolmogorov Complexity

Fix an acceptable programming system.
For any binary strings x, y:

$$C(x) = \text{length of a shortest program printing } x,$$
$$C(x|y) = \text{length of a shortest program printing } x \text{ given input } y.$$

And up to $O(\log |xy|)$,

$$C(x) \geq 0,$$
$$C(x|y) \geq 0,$$
$$C(x) + C(y) \geq C(x, y).$$

Theorem (Inequalities are the same, Hammer *et al*)

An inequality is valid for Shannon iff it is valid for Kolmogorov up to an additive logarithmic term

Conditional Algorithmic Inequalities

- We cannot say $C(a, b) = C(a) + C(b)$ with exact equality.
- All statements in the Kolmogorov framework are (inherently) asymptotic.
- **Need to add a precision for conditions**:
- We have "thick" faces of thickness $f(N)$
  (where N is the complexity of the tuple of strings)

Conditional Algorithmic Inequalities

- We cannot say $C(a, b) = C(a) + C(b)$ with exact equality.
- All statements in the Kolmogorov framework are (inherently) asymptotic.
- **Need to add a precision for conditions**:
- We have "thick" faces of thickness $f(N)$
  (where N is the complexity of the tuple of strings)

- Some conditional inequalities are valid up to $O(f(N))$
- Some conditional inequalities are valid up to $\theta\left(\sqrt{Nf(N)}\right)$
- Some conditional inequalities are not valid ($O(N)$ counterexample)

Secret Sharing

New parameters: the leakages.

Definition

A perfect secret-sharing scheme for Γ is a tuple of discrete random variables $(s, p_1, \ldots, p_n)$ such that :

- if $A \in \Gamma$ then $H(s|A) = 0$
- if $B \notin \Gamma$ then $I(s:B) = 0$

New parameters: the leakages.

Definition

A secret-sharing scheme for Γ is a tuple of discrete random variables $(s, p_1, \ldots, p_n)$ such that :

- if $A \in \Gamma$ then $H(s|A) \leq \underbrace{\varepsilon H(s)}_{\text{missing information}}$
- if $B \notin \Gamma$ then $I(s:B) \leq \underbrace{\delta H(s)}_{\text{information leak}}$

New parameters: the leakages.

Definition

A secret-sharing scheme for Γ is a tuple of discrete random variables $(s, p_1, \ldots, p_n)$ such that :

- if $A \in \Gamma$ then $H(s|A) \leq \underbrace{\varepsilon H(s)}_{\text{missing information}}$
- if $B \notin \Gamma$ then $I(s\!:\!B) \leq \underbrace{\delta H(s)}_{\text{information leak}}$

Parameters of a scheme:

ε : missing information ratio.

δ : information leak ratio.

ρ : information ratio $= \max\limits_{p} \frac{H(p)}{H(s)}$.

Quasi-perfect Secret Sharing

Definition

An access structure Γ can be **quasi-perfectly implemented with information ratio** ρ if there exists a sequence of secret-sharing schemes such that:

(1) the $\limsup$ of the information ratio does not exceed ρ;

(2) the missing information ratio tends to zero;

(3) the information leak ratio tends to zero.

Quasi-perfect Secret Sharing

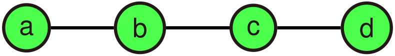
Definition

An access structure Γ can be **quasi-perfectly implemented with information ratio** ρ if there exists a sequence of secret-sharing schemes such that:

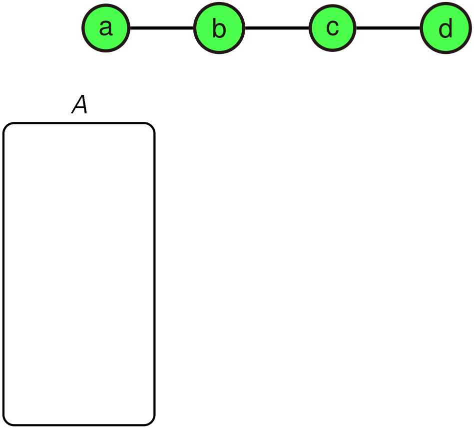
(1) the $\limsup$ of the information ratio does not exceed ρ;
(2) the missing information ratio tends to zero;
(3) the information leak ratio tends to zero.

- Almost entropic version of secret sharing.
- Closely related to a "Kolmogorovian" Counterpart of Secret Sharing.
- Question: Can they achieve better information ratios?

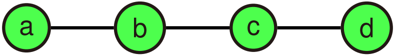
Bounds on Perfect Schemes

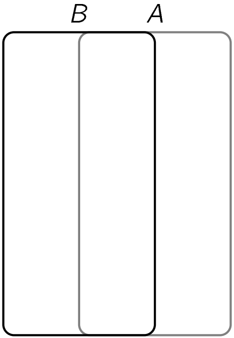


Proof by (Venn) Information Diagram

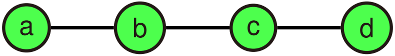


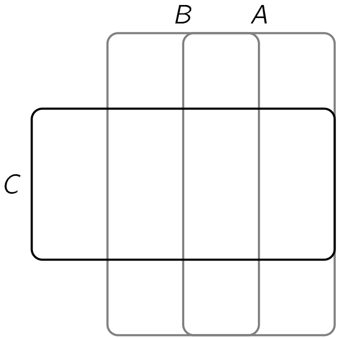
Proof by (Venn) Information Diagram



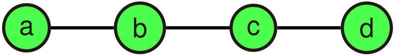


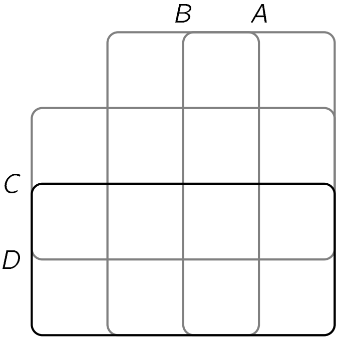
Proof by (Venn) Information Diagram



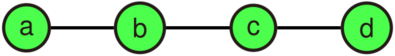


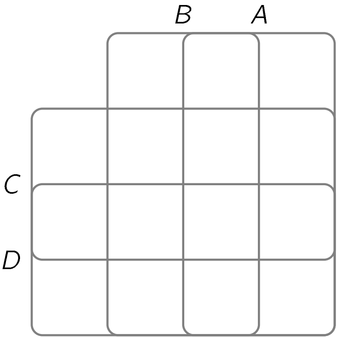
Proof by (Venn) Information Diagram



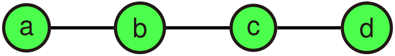


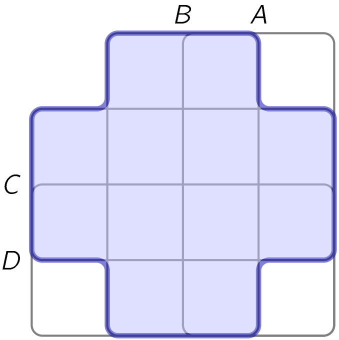
Proof by (Venn) Information Diagram





Proof by (Venn) Information Diagram

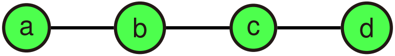


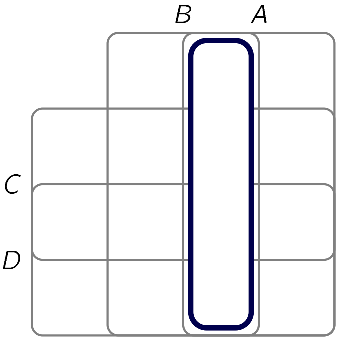


Cells contained in B or C represent:

$$H(BC)$$

Proof by (Venn) Information Diagram

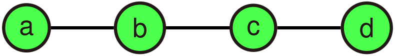


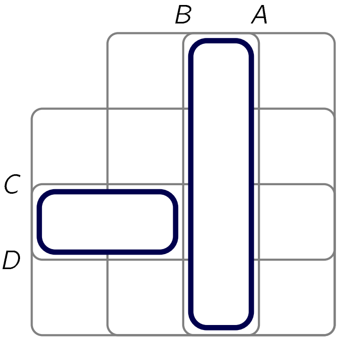


Cells contained in both A and B represent:

$$I(A:B)$$

Proof by (Venn) Information Diagram

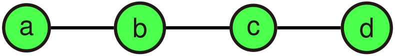


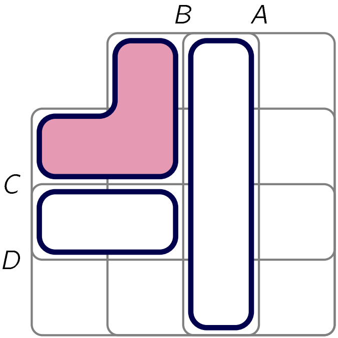


Cells contained in both C and D but not A represent:

$$I(C:D|A)$$

Proof by (Venn) Information Diagram

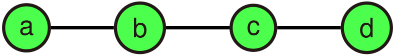


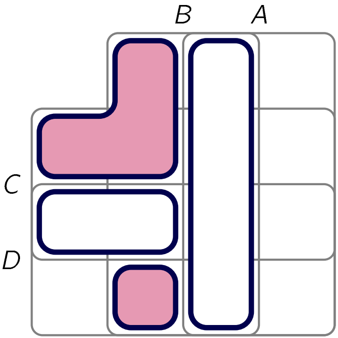


Cells contained in B or C but not A nor D represent:

$$H(BC|AD)$$

Proof by (Venn) Information Diagram

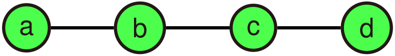


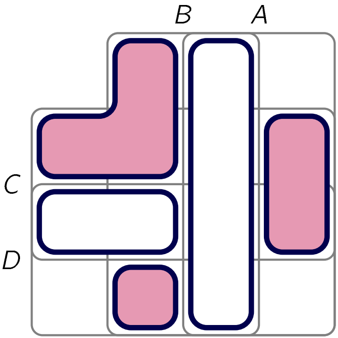


Cells contained in both B and D but not A nor C represent:

$$I(B:D|AC)$$

Proof by (Venn) Information Diagram

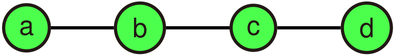


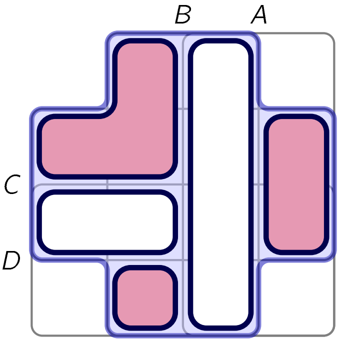


Cells contained in both A and C but not B represent:

$$I(A{:}C|B)$$

Proof by (Venn) Information Diagram

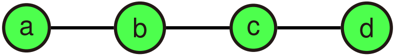


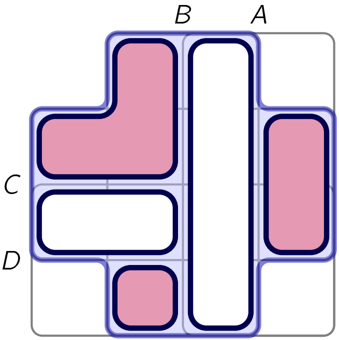


Actually, we just proved an identity without words...

$$H(BC) = I(A{:}C|B) + I(B{:}D|AC) + H(BC|AD) + I(A{:}B) + I(C{:}D|A).$$

Proof by (Venn) Information Diagram

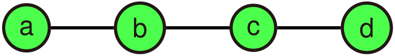


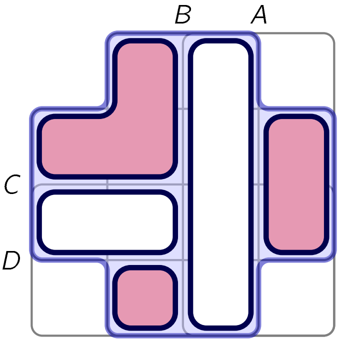


..or an inequality, since all quantities are non-negative.

$$H(BC) \geq I(A{:}C|B) + I(B{:}D|AC) + H(BC|AD).$$

Proof by (Venn) Information Diagram

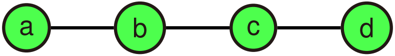


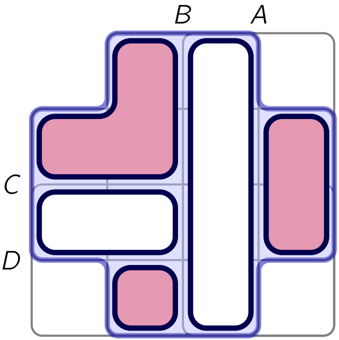


Using the perfect secret sharing requirements, we obtain:

$$H(BC) \geq 3H(S).$$

Proof by (Venn) Information Diagram





In general this is a HUGELY conditional inequality

"secret sharing requirements" $\Rightarrow H(BC) \geq 3H(S)$.

Merci de votre attention.

Des questions?