Different Views of Information
Part I

***

Geometry in Cryptography and Communication

Andrei Romashchenko

UiT The Arctic University of Norway, October 2025

measure of information: Hartley's approach

Definition: the *amount of information* in a finite set A is $\chi(S) = \log \mathrm{Card}(S)$.

measure of information: Hartley's approach

Definition: the *amount of information* in a finite set A is $\chi(S) = \log \mathrm{Card}(S)$.

Observation 1: For all S

$$\chi(S^k) = k \cdot \chi(S).$$

measure of information: Hartley's approach

Definition: the *amount of information* in a finite set A is $\chi(S) = \log \mathrm{Card}(S)$.

Observation 1: For all S

$$\chi(S^k) = k \cdot \chi(S).$$

Observation 2: Let $A =$ set of binary strings of length N, with $p_0 N$ *zeros* and $p_1 N$ *ons* ($p_0 + p_1 = 1$). Then

$$\chi(A) = \log \frac{N!}{(p_0 N)! \cdot (p_1 N)!}$$

measure of information: Hartley's approach

Definition: the *amount of information* in a finite set A is $\chi(S) = \log \mathrm{Card}(S)$.

Observation 1: For all S

$$\chi(S^k) = k \cdot \chi(S).$$

Observation 2: Let $A =$ set of binary strings of length N, with $p_0 N$ *zeros* and $p_1 N$ *ons* ($p_0 + p_1 = 1$). Then

$$\chi(A) = \log \frac{N!}{(p_0 N)! \cdot (p_1 N)!} = \left(p_0 \log \frac{1}{p_0} + p_1 \log \frac{1}{p_1} \right) N + O(\log N)$$

measure of information: Hartley's approach

Definition: the *amount of information* in a finite set A is $\chi(S) = \log \mathrm{Card}(S)$.

Observation 1: For all S

$$\chi(S^k) = k \cdot \chi(S).$$

Observation 2: Let $A =$ set of binary strings of length N, with $p_0 N$ *zeros* and $p_1 N$ *ons* ($p_0 + p_1 = 1$). Then

$$\chi(A) = \log \frac{N!}{(p_0 N)! \cdot (p_1 N)!} = \left(p_0 \log \frac{1}{p_0} + p_1 \log \frac{1}{p_1} \right) N + O(\log N)$$

Observation 3: Let $B =$ set of strings of length N in the alphabet $\{a, b, \ldots, z\}$, with 12.7% letters 'e', 9.1% letters 't', 8.2% letters 'a', 7.5% letters 'o', ... Then

$$\chi(B) \approx 4.14N$$

measure of information: Hartley's approach

Definition: the *amount of information* in a finite set S is $\chi(S) = \log \mathrm{Card}(S)$.

measure of information: Hartley's approach

Definition: the *amount of information* in a finite set S is $\chi(S) = \log \mathrm{Card}(S)$.

Definition: If $S \subset \mathbb{N} \times \mathbb{N}$, then

$$\chi_1(S) = \log \mathrm{Card}(\pi_1 S), \quad \chi_2(S) = \log \mathrm{Card}(\pi_2 S)$$

where $\pi_i S$ denotes the projection of S onto the i-th coordinate.

measure of information: Hartley's approach

Definition: the *amount of information* in a finite set S is $\chi(S) = \log \mathrm{Card}(S)$.

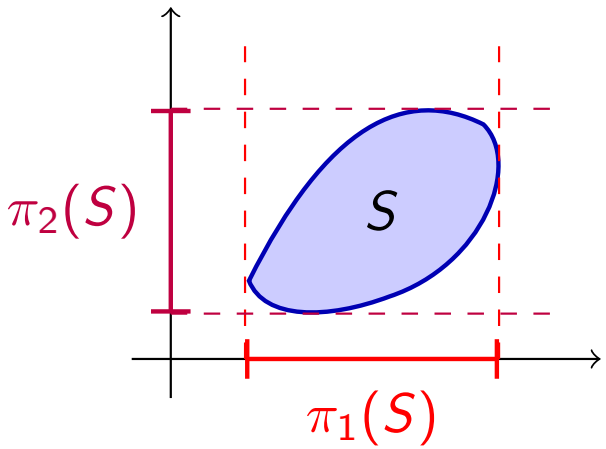
Definition: If $S \subset \mathbb{N} \times \mathbb{N}$, then

$$\chi_1(S) = \log \mathrm{Card}(\pi_1 S), \quad \chi_2(S) = \log \mathrm{Card}(\pi_2 S)$$

where $\pi_i S$ denotes the projection of S onto the i-th coordinate.

Observation 1: $\chi(S) \leq \chi_1(S) + \chi_2(S)$

and $\chi(S) = \chi_1(S) + \chi_2(S)$ iff $S = (\pi_1 S) \times (\pi_2 S)$



measure of information: Hartley's approach

Definition: the *amount of information* in a finite set S is $\chi(S) = \log \mathrm{Card}(S)$.

Definition: If $S \subset \mathbb{N} \times \mathbb{N}$, then

$$\chi_1(S) = \log \mathrm{Card}(\pi_1 S), \quad \chi_2(S) = \log \mathrm{Card}(\pi_2 S)$$

where $\pi_i S$ denotes the projection of S onto the i-th coordinate.

Observation 1: $\chi(S) \leq \chi_1(S) + \chi_2(S)$

and $\chi(S) = \chi_1(S) + \chi_2(S)$ iff $S = (\pi_1 S) \times (\pi_2 S)$

Observation 2: if $S \subset \mathbb{N} \times \mathbb{N} \times \mathbb{N}$, then $\chi(S) \leq \chi_1(S) + \chi_2(S) + \chi_3(S)$.

measure of information: Hartley's approach

Definition: the *amount of information* in a finite set S is $\chi(S) = \log \mathrm{Card}(S)$.

Definition: If $S \subset \mathbb{N} \times \mathbb{N}$, then

$$\chi_1(S) = \log \mathrm{Card}(\pi_1 S), \quad \chi_2(S) = \log \mathrm{Card}(\pi_2 S)$$

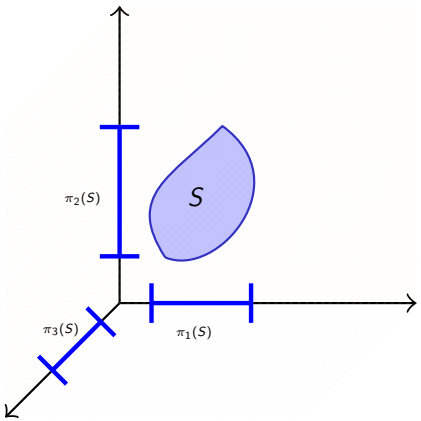
where $\pi_i S$ denotes the projection of S onto the i-th coordinate.

Observation 1: $\chi(S) \leq \chi_1(S) + \chi_2(S)$

and $\chi(S) = \chi_1(S) + \chi_2(S)$ iff $S = (\pi_1 S) \times (\pi_2 S)$

Observation 2: if $S \subset \mathbb{N} \times \mathbb{N} \times \mathbb{N}$, then $\chi(S) \leq \chi_1(S) + \chi_2(S) + \chi_3(S)$.

This is rephrasing of the claim $\mathrm{Card}(S) \leq \mathrm{Card}(\pi_1 S) \cdot \mathrm{Card}(\pi_2 S) \cdot \mathrm{Card}(\pi_3 S)$.



measure of information: Hartley's approach

Definition: the *amount of information* in a finite set S is $\chi(S) = \log \mathrm{Card}(S)$.

Definition: If $S \subset \mathbb{N} \times \mathbb{N}$, then

$$\chi_1(S) = \log \mathrm{Card}(\pi_1 S), \quad \chi_2(S) = \log \mathrm{Card}(\pi_2 S)$$

where $\pi_i S$ denotes the projection of S onto the i-th coordinate.

Observation 1: $\chi(S) \leq \chi_1(S) + \chi_2(S)$

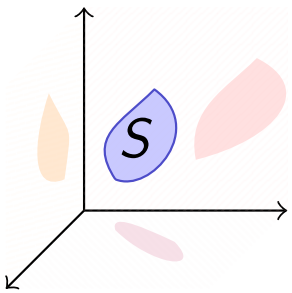
and $\chi(S) = \chi_1(S) + \chi_2(S)$ iff $S = (\pi_1 S) \times (\pi_2 S)$

Observation 2: if $S \subset \mathbb{N} \times \mathbb{N} \times \mathbb{N}$, then $\chi(S) \leq \chi_1(S) + \chi_2(S) + \chi_3(S)$.

This is rephrasing of the claim $\mathrm{Card}(S) \leq \mathrm{Card}(\pi_1 S) \cdot \mathrm{Card}(\pi_2 S) \cdot \mathrm{Card}(\pi_3 S)$.

Theorem *(Loomis–Whitney)* If $S \subset \mathbb{N} \times \mathbb{N} \times \mathbb{N}$, then

$$2 \cdot \chi(S) \leq \chi_{12}(S) + \chi_{23}(S) + \chi_{13}(S).$$



measure of information: Hartley's approach

Definition: the *amount of information* in a finite set S is $\chi(S) = \log \mathrm{Card}(S)$.

Definition: If $S \subset \mathbb{N} \times \mathbb{N}$, then

$$\chi_1(S) = \log \mathrm{Card}(\pi_1 S), \quad \chi_2(S) = \log \mathrm{Card}(\pi_2 S)$$

where $\pi_i S$ denotes the projection of S onto the i-th coordinate.

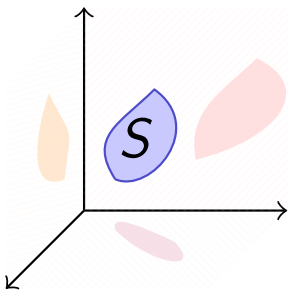
Observation 1: $\chi(S) \leq \chi_1(S) + \chi_2(S)$

and $\chi(S) = \chi_1(S) + \chi_2(S)$ iff $S = (\pi_1 S) \times (\pi_2 S)$

Observation 2: if $S \subset \mathbb{N} \times \mathbb{N} \times \mathbb{N}$, then $\chi(S) \leq \chi_1(S) + \chi_2(S) + \chi_3(S)$.

This is rephrasing of the claim $\mathrm{Card}(S) \leq \mathrm{Card}(\pi_1 S) \cdot \mathrm{Card}(\pi_2 S) \cdot \mathrm{Card}(\pi_3 S)$.

Theorem *(Loomis–Whitney)* If $S \subset \mathbb{N} \times \mathbb{N} \times \mathbb{N}$, then



$$2 \cdot \chi(S) \leq \chi_{12}(S) + \chi_{23}(S) + \chi_{13}(S).$$

This inequality basically claims that

$$\mathrm{Card}(S)^2 \leq \mathrm{Card}(\pi_{12} S) \cdot \mathrm{Card}(\pi_{13} S) \cdot \mathrm{Card}(\pi_{23} S).$$

measure of information: Hartley's approach

Definition: the *amount of information* in a finite set S is $\chi(S) = \log \mathrm{Card}(S)$.

Definition: If $S \subset \mathbb{N} \times \mathbb{N}$, then

$$\chi_1(S) = \log \mathrm{Card}(\pi_1 S), \quad \chi_2(S) = \log \mathrm{Card}(\pi_2 S)$$

where $\pi_i S$ denotes the projection of S onto the i-th coordinate.

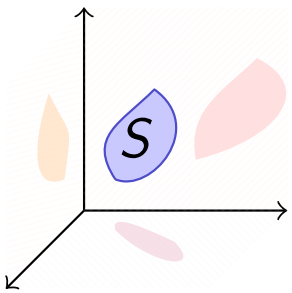
Observation 1: $\chi(S) \leq \chi_1(S) + \chi_2(S)$

and $\chi(S) = \chi_1(S) + \chi_2(S)$ iff $S = (\pi_1 S) \times (\pi_2 S)$

Observation 2: if $S \subset \mathbb{N} \times \mathbb{N} \times \mathbb{N}$, then $\chi(S) \leq \chi_1(S) + \chi_2(S) + \chi_3(S)$.

This is rephrasing of the claim $\mathrm{Card}(S) \leq \mathrm{Card}(\pi_1 S) \cdot \mathrm{Card}(\pi_2 S) \cdot \mathrm{Card}(\pi_3 S)$.

Theorem *(Loomis–Whitney)* If $S \subset \mathbb{N} \times \mathbb{N} \times \mathbb{N}$, then



$$2 \cdot \chi(S) \leq \chi_{12}(S) + \chi_{23}(S) + \chi_{13}(S).$$

Continuous version:

$$\mathrm{volume(S)}^2 \leq \mathrm{area}(\pi_{12} S) \cdot \mathrm{area}(\pi_{13} S) \cdot \mathrm{area}(\pi_{23} S).$$

Hartley's information: toy applications

Question 1: There are 25 identical-looking coins, one of which is counterfeit and lighter than the others. Using a balance scale without additional weights, determine the least number of weighings needed to identify the counterfeit coin.

Hartley's information: toy applications

Question 1: There are 25 identical-looking coins, one of which is counterfeit and lighter than the others. Using a balance scale without additional weights, determine the least number of weighings needed to identify the counterfeit coin.

Simple: a strategy with 3 weighings.

Hartley's information: toy applications

Question 1: There are 25 identical-looking coins, one of which is counterfeit and lighter than the others. Using a balance scale without additional weights, determine the least number of weighings needed to identify the counterfeit coin.

Simple: a strategy with 3 weighings.

Proof of a lower bound: For a strategy with k operations

$$\chi(\text{outcome}) \leq \chi(1\text{st weighing}) + \ldots + \chi(k\text{st weighing}) \leq \underbrace{\log 3 + \ldots + \log 3}_{k}$$

so $k \geq (\log 25)/(\log 3)$

Hartley's information: toy applications

Question 1: There are 25 identical-looking coins, one of which is counterfeit and lighter than the others. Using a balance scale without additional weights, determine the least number of weighings needed to identify the counterfeit coin.

Simple: a strategy with 3 weighings.

Proof of a lower bound: For a strategy with k operations

$$\chi(\text{outcome}) \leq \chi(\text{1st weighing}) + \ldots + \chi(k\text{st weighing}) \leq \underbrace{\log 3 + \ldots + \log 3}_{k}$$

so $k \geq (\log 25)/(\log 3)$

Question 2: There are 14 identical-looking coins, one of which is counterfeit and differs in weight (either lighter or heavier) from the others. Using a balance scale without additional weights, determine the least number of weighings needed to identify the counterfeit coin.

Hartley's information: toy applications

Question 1: There are 25 identical-looking coins, one of which is counterfeit and lighter than the others. Using a balance scale without additional weights, determine the least number of weighings needed to identify the counterfeit coin.

Simple: a strategy with 3 weighings.

Proof of a lower bound: For a strategy with k operations

$$\chi(\text{outcome}) \leq \chi(\text{1st weighing}) + \ldots + \chi(k\text{st weighing}) \leq \underbrace{\log 3 + \ldots + \log 3}_{k}$$

so $k \geq (\log 25)/(\log 3)$

Question 2: There are 14 identical-looking coins, one of which is counterfeit and differs in weight (either lighter or heavier) from the others. Using a balance scale without additional weights, determine the least number of weighings needed to identify the counterfeit coin.

This is an exercise!

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

with the usual convention $0 \cdot \log \frac{1}{0} = 0$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

with the usual convention $0 \cdot \log \frac{1}{0} = 0$

Properties:

- $\mathrm{H}(X) \geq 0$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

with the usual convention $0 \cdot \log \frac{1}{0} = 0$

Properties:

- $\mathrm{H}(X) \geq 0$, with equality iff

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

with the usual convention $0 \cdot \log \frac{1}{0} = 0$

Properties:

- $\mathrm{H}(X) \geq 0$, with equality iff $p_i = 1$ for some i

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

with the usual convention $0 \cdot \log \frac{1}{0} = 0$

Properties:

- $\mathrm{H}(X) \geq 0$, with equality iff $p_i = 1$ for some i
  (immediately form the definition)

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

with the usual convention $0 \cdot \log \frac{1}{0} = 0$

Properties:

- $\mathrm{H}(X) \geq 0$, with equality iff $p_i = 1$ for some i
  (immediately form the definition)
- $\mathrm{H}(X) \leq \log k$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

with the usual convention $0 \cdot \log \frac{1}{0} = 0$

Properties:

- $\mathrm{H}(X) \geq 0$, with equality iff $p_i = 1$ for some i
  (immediately form the definition)
- $\mathrm{H}(X) \leq \log k$, with equality iff

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

with the usual convention $0 \cdot \log \frac{1}{0} = 0$

Properties:

- $\mathrm{H}(X) \geq 0$, with equality iff $p_i = 1$ for some i
  (immediately form the definition)
- $\mathrm{H}(X) \leq \log k$, with equality iff $p_1 = \cdots = p_k = \frac{1}{k}$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

with the usual convention $0 \cdot \log \frac{1}{0} = 0$

Properties:

- $\mathrm{H}(X) \geq 0$, with equality iff $p_i = 1$ for some i
  (immediately form the definition)
- $\mathrm{H}(X) \leq \log k$, with equality iff $p_1 = \cdots = p_k = \frac{1}{k}$
  (proof: concavity of the logarithm + Jensen's inequality).

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

with the usual convention $0 \cdot \log \frac{1}{0} = 0$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

with the usual convention $0 \cdot \log \frac{1}{0} = 0$

two classical theorems:

- for all uniquely decodable binary code $c_1, \ldots, c_k$

  $$\sum_{i=1}^{k} p_i \cdot \mathrm{length}(c_i) \geq \mathrm{H}(X)$$

- there exists a uniquely decodable binary code $c_1, \ldots, c_k$ such that

  $$\sum_{i=1}^{k} p_i \cdot \mathrm{length}(c_i) < \mathrm{H}(X) + 1$$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

with the usual convention $0 \cdot \log \frac{1}{0} = 0$

two classical theorems:

- for all uniquely decodable binary code $c_1, \ldots, c_k$

  $$\sum_{i=1}^{k} p_i \cdot \mathrm{length}(c_i) \geq \mathrm{H}(X)$$

- there exists a uniquely decodable binary code $c_1, \ldots, c_k$ such that

  $$\sum_{i=1}^{k} p_i \cdot \mathrm{length}(c_i) < \mathrm{H}(X) + 1$$

slightly informally: the value of $\mathrm{H}(X)$ gives the optimal compression rate

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

cf. the proof in suppl. materials:
$2\mathrm{H}(X, Y, Z) \leq \mathrm{H}(X, Y) + \mathrm{H}(X, Z) + \mathrm{H}(Y, Z)$.

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

cf. the proof in suppl. materials:
$2\mathrm{H}(X, Y, Z) \leq \mathrm{H}(X, Y) + \mathrm{H}(X, Z) + \mathrm{H}(Y, Z)$.

Loomis–Whitney revisited: if $S \subset \mathbb{N} \times \mathbb{N} \times \mathbb{N}$, then

$$\mathrm{Card}(S)^2 \leq \mathrm{Card}(\pi_{12} S) \cdot \mathrm{Card}(\pi_{13} S) \cdot \mathrm{Card}(\pi_{23} S).$$

$$2 \cdot \chi(S) \leq \chi_{12}(S) + \chi_{23}(S) + \chi_{13}(S).$$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

cf. the proof in suppl. materials:
$2\mathrm{H}(X, Y, Z) \leq \mathrm{H}(X, Y) + \mathrm{H}(X, Z) + \mathrm{H}(Y, Z)$.

Loomis–Whitney revisited: if $S \subset \mathbb{N} \times \mathbb{N} \times \mathbb{N}$, then

$$\mathrm{Card}(S)^2 \leq \mathrm{Card}(\pi_{12} S) \cdot \mathrm{Card}(\pi_{13} S) \cdot \mathrm{Card}(\pi_{23} S).$$

$$2 \cdot \chi(S) \leq \chi_{12}(S) + \chi_{23}(S) + \chi_{13}(S).$$

Sketch of the proof: sample (X, Y, Z) uniformly in S

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

cf. the proof in suppl. materials:
$2\mathrm{H}(X, Y, Z) \leq \mathrm{H}(X, Y) + \mathrm{H}(X, Z) + \mathrm{H}(Y, Z)$.

Loomis–Whitney revisited: if $S \subset \mathbb{N} \times \mathbb{N} \times \mathbb{N}$, then

$$\mathrm{Card}(S)^2 \leq \mathrm{Card}(\pi_{12} S) \cdot \mathrm{Card}(\pi_{13} S) \cdot \mathrm{Card}(\pi_{23} S).$$

$$2 \cdot \chi(S) \leq \chi_{12}(S) + \chi_{23}(S) + \chi_{13}(S).$$

Sketch of the proof: sample (X, Y, Z) uniformly in S

$$2 \cdot \log \mathrm{Card}(S) \quad = \quad 2\mathrm{H}(X, Y, Z)$$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

cf. the proof in suppl. materials:
$2\mathrm{H}(X,Y,Z) \leq \mathrm{H}(X,Y) + \mathrm{H}(X,Z) + \mathrm{H}(Y,Z)$.

Loomis–Whitney revisited: if $S \subset \mathbb{N} \times \mathbb{N} \times \mathbb{N}$, then

$$\mathrm{Card}(S)^2 \leq \mathrm{Card}(\pi_{12}S) \cdot \mathrm{Card}(\pi_{13}S) \cdot \mathrm{Card}(\pi_{23}S).$$

$$2 \cdot \chi(S) \leq \chi_{12}(S) + \chi_{23}(S) + \chi_{13}(S).$$

Sketch of the proof: sample (X,Y,Z) uniformly in S

$$\begin{aligned} 2 \cdot \log \mathrm{Card}(S) &= 2\mathrm{H}(X,Y,Z) \\ &\leq \mathrm{H}(X,Y) + \mathrm{H}(X,Z) + \mathrm{H}(Y,Z) \end{aligned}$$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

cf. the proof in suppl. materials:
$2\mathrm{H}(X,Y,Z) \le \mathrm{H}(X,Y) + \mathrm{H}(X,Z) + \mathrm{H}(Y,Z).$

Loomis–Whitney revisited: if $S \subset \mathbb{N} \times \mathbb{N} \times \mathbb{N}$, then

$$\mathrm{Card}(S)^2 \le \mathrm{Card}(\pi_{12}S) \cdot \mathrm{Card}(\pi_{13}S) \cdot \mathrm{Card}(\pi_{23}S).$$

$$2 \cdot \chi(S) \le \chi_{12}(S) + \chi_{23}(S) + \chi_{13}(S).$$

Sketch of the proof: sample (X,Y,Z) uniformly in S

$$\begin{aligned} 2 \cdot \log \mathrm{Card}(S) &= 2\mathrm{H}(X,Y,Z) \\ &\le \mathrm{H}(X,Y) + \mathrm{H}(X,Z) + \mathrm{H}(Y,Z) \\ &\le \log \mathrm{Card}(\pi_{12}(S)) + \log \mathrm{Card}(\pi_{23}(S)) + \log \mathrm{Card}(\pi_{13}(S)) \end{aligned}$$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

with the usual convention $0 \cdot \log \frac{1}{0} = 0$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

with the usual convention $0 \cdot \log \frac{1}{0} = 0$

For jointly distributed X, Y we have:

$$\mathrm{H}(X),\ \mathrm{H}(Y),\ \mathrm{H}(X, Y).$$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

with the usual convention $0 \cdot \log \frac{1}{0} = 0$

For jointly distributed X, Y we have:

$$\mathrm{H}(X),\ \mathrm{H}(Y),\ \mathrm{H}(X, Y).$$

Properties:

- $\mathrm{H}(X, Y) \leq \mathrm{H}(X) + \mathrm{H}(Y)$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

with the usual convention $0 \cdot \log \frac{1}{0} = 0$

For jointly distributed X, Y we have:

$$\mathrm{H}(X),\ \mathrm{H}(Y),\ \mathrm{H}(X, Y).$$

Properties:

- $\mathrm{H}(X, Y) \leq \mathrm{H}(X) + \mathrm{H}(Y)$
- $\mathrm{H}(X, Y) = \mathrm{H}(X) + \mathrm{H}(Y)$, iff X and Y are independent

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

with the usual convention $0 \cdot \log \frac{1}{0} = 0$

For jointly distributed X, Y we have:

$$\mathrm{H}(X),\ \mathrm{H}(Y),\ \mathrm{H}(X, Y).$$

Properties:

- $\mathrm{H}(X, Y) \leq \mathrm{H}(X) + \mathrm{H}(Y)$
- $\mathrm{H}(X, Y) = \mathrm{H}(X) + \mathrm{H}(Y)$, iff X and Y are independent

An **Exercise** !

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

Definition: For jointly distributed X, Y, for each fixed value b of Y we have

$$\mathrm{H}(X \mid Y = b).$$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

Definition: For jointly distributed X, Y, for each fixed value b of Y we have

$$\mathrm{H}(X \mid Y = b).$$

Conditional Shannon entropy of X given Y is

$$\mathrm{H}(X \mid Y) := \sum_{b} \mathrm{H}(X \mid Y = b)\ \Pr[Y = b].$$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

Definition: For jointly distributed X, Y, for each fixed value b of Y we have

$$\mathrm{H}(X \mid Y = b).$$

Conditional Shannon entropy of X given Y is

$$\mathrm{H}(X \mid Y) := \sum_{b} \mathrm{H}(X \mid Y = b)\ \Pr[Y = b].$$

Properties:

- $\mathrm{H}(X, Y) = \mathrm{H}(X \mid Y) + \mathrm{H}(Y)$,
- $\mathrm{H}(X \mid Y) \geq 0$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

Definition: For jointly distributed X, Y, for each fixed value b of Y we have

$$\mathrm{H}(X \mid Y = b).$$

Conditional Shannon entropy of X given Y is

$$\mathrm{H}(X \mid Y) := \sum_{b} \mathrm{H}(X \mid Y = b)\ \Pr[Y = b].$$

Properties:

- $\mathrm{H}(X, Y) = \mathrm{H}(X \mid Y) + \mathrm{H}(Y)$,
- $\mathrm{H}(X \mid Y) \geq 0$, with $\mathrm{H}(X \mid Y) = 0$ iff $X = \mathrm{Function}(Y)$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

Definition: For jointly distributed X, Y, for each fixed value b of Y we have

$$\mathrm{H}(X \mid Y = b).$$

Conditional Shannon entropy of X given Y is

$$\mathrm{H}(X \mid Y) := \sum_{b} \mathrm{H}(X \mid Y = b)\ \Pr[Y = b].$$

Properties:

- $\mathrm{H}(X, Y) = \mathrm{H}(X \mid Y) + \mathrm{H}(Y)$,
- $\mathrm{H}(X \mid Y) \geq 0$, with $\mathrm{H}(X \mid Y) = 0$ iff $X = \mathrm{Function}(Y)$

An **Exercise** !

measure of information: Shannon's approach

Definition: (Shannon entropy)

For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

Definition: For jointly distributed X, Y, the *information on Y contained in X* is

$$\mathrm{I}(X : Y) = \mathrm{H}(Y) - \mathrm{H}(Y \mid X).$$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

Definition: For jointly distributed X, Y, the *information on Y contained in X* is

$$\mathrm{I}(X : Y) = \mathrm{H}(Y) - \mathrm{H}(Y \mid X).$$

Properties:

- $\mathrm{I}(X : Y) = \mathrm{I}(Y : X) = \mathrm{H}(X) + \mathrm{H}(Y) - \mathrm{H}(X, Y)$,

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

Definition: For jointly distributed X, Y, the *information on Y contained in X* is

$$\mathrm{I}(X : Y) = \mathrm{H}(Y) - \mathrm{H}(Y \mid X).$$

Properties:

- $\mathrm{I}(X : Y) = \mathrm{I}(Y : X) = \mathrm{H}(X) + \mathrm{H}(Y) - \mathrm{H}(X, Y)$,
- $\mathrm{I}(X : Y) \leq \mathrm{H}(X)$, and $\mathrm{I}(X : Y) \leq \mathrm{H}(Y)$,

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

Definition: For jointly distributed X, Y, the *information on Y contained in X* is

$$\mathrm{I}(X : Y) = \mathrm{H}(Y) - \mathrm{H}(Y \mid X).$$

Properties:

- $\mathrm{I}(X : Y) = \mathrm{I}(Y : X) = \mathrm{H}(X) + \mathrm{H}(Y) - \mathrm{H}(X, Y)$,
- $\mathrm{I}(X : Y) \leq \mathrm{H}(X)$, and $\mathrm{I}(X : Y) \leq \mathrm{H}(Y)$,
- $\mathrm{I}(X : Y) \geq 0$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

Definition: For jointly distributed X, Y, the *information on Y contained in X* is

$$\mathrm{I}(X : Y) = \mathrm{H}(Y) - \mathrm{H}(Y \mid X).$$

Properties:

- $\mathrm{I}(X : Y) = \mathrm{I}(Y : X) = \mathrm{H}(X) + \mathrm{H}(Y) - \mathrm{H}(X, Y)$,
- $\mathrm{I}(X : Y) \leq \mathrm{H}(X)$, and $\mathrm{I}(X : Y) \leq \mathrm{H}(Y)$,
- $\mathrm{I}(X : Y) \geq 0$, with $\mathrm{I}(X : Y) = 0$ iff $X \perp\!\!\!\perp Y$ (independent),

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

Definition: For jointly distributed X, Y, the *information on Y contained in X* is

$$\mathrm{I}(X : Y) = \mathrm{H}(Y) - \mathrm{H}(Y \mid X).$$

Properties:

- $\mathrm{I}(X : Y) = \mathrm{I}(Y : X) = \mathrm{H}(X) + \mathrm{H}(Y) - \mathrm{H}(X, Y)$,
- $\mathrm{I}(X : Y) \leq \mathrm{H}(X)$, and $\mathrm{I}(X : Y) \leq \mathrm{H}(Y)$,
- $\mathrm{I}(X : Y) \geq 0$, with $\mathrm{I}(X : Y) = 0$ iff $X \perp\!\!\!\perp Y$ (independent),
- $\mathrm{I}(X : Y) = \mathrm{H}(X)$ if and only if $X = \text{Function}(Y)$,

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

Definition: For jointly distributed X, Y, the *information on Y contained in X* is

$$\mathrm{I}(X : Y) = \mathrm{H}(Y) - \mathrm{H}(Y \mid X).$$

Properties:

- $\mathrm{I}(X : Y) = \mathrm{I}(Y : X) = \mathrm{H}(X) + \mathrm{H}(Y) - \mathrm{H}(X, Y)$,
- $\mathrm{I}(X : Y) \leq \mathrm{H}(X)$, and $\mathrm{I}(X : Y) \leq \mathrm{H}(Y)$,
- $\mathrm{I}(X : Y) \geq 0$, with $\mathrm{I}(X : Y) = 0$ iff $X \perp\!\!\!\perp Y$ (independent),
- $\mathrm{I}(X : Y) = \mathrm{H}(X)$ if and only if $X = \text{Function}(Y)$,
- $\mathrm{I}(X : X) = \mathrm{H}(X)$.

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

Definition: For jointly distributed X, Y, the *information on Y contained in X* is

$$\mathrm{I}(X : Y) = \mathrm{H}(Y) - \mathrm{H}(Y \mid X).$$

Properties:

- $\mathrm{I}(X : Y) = \mathrm{I}(Y : X) = \mathrm{H}(X) + \mathrm{H}(Y) - \mathrm{H}(X, Y)$,
- $\mathrm{I}(X : Y) \leq \mathrm{H}(X)$, and $\mathrm{I}(X : Y) \leq \mathrm{H}(Y)$,
- $\mathrm{I}(X : Y) \geq 0$, with $\mathrm{I}(X : Y) = 0$ iff $X \perp\!\!\!\perp Y$ (independent),
- $\mathrm{I}(X : Y) = \mathrm{H}(X)$ if and only if $X = \text{Function}(Y)$,
- $\mathrm{I}(X : X) = \mathrm{H}(X)$.

An **Exercise** !

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

Definition: conditional mutual information between X and Y given Z

1st definition: $\mathrm{I}(X : Y \mid Z) := \sum_c \mathrm{I}(X : Y \mid Z = c) \Pr[Z = c]$,

2nd definition: $\mathrm{I}(X : Y \mid Z) := \mathrm{H}(Y \mid Z) - \mathrm{H}(Y \mid X, Z)$.

3rd definition: $\mathrm{I}(X : Y \mid Z) := \mathrm{H}(X \mid Z) + \mathrm{H}(Y \mid Z) - \mathrm{H}(X, Y \mid Z)$.

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

Definition: conditional mutual information between X and Y given Z

1st definition: $\mathrm{I}(X : Y \mid Z) := \sum_c \mathrm{I}(X : Y \mid Z = c) \Pr[Z = c]$,

2nd definition: $\mathrm{I}(X : Y \mid Z) := \mathrm{H}(Y \mid Z) - \mathrm{H}(Y \mid X, Z)$.

3rd definition: $\mathrm{I}(X : Y \mid Z) := \mathrm{H}(X \mid Z) + \mathrm{H}(Y \mid Z) - \mathrm{H}(X, Y \mid Z)$.

Exercise: these three definitions are equivalent.

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

Definition: conditional mutual information between X and Y given Z

1st definition: $\mathrm{I}(X : Y \mid Z) := \sum_c \mathrm{I}(X : Y \mid Z = c) \Pr[Z = c]$,

2nd definition: $\mathrm{I}(X : Y \mid Z) := \mathrm{H}(Y \mid Z) - \mathrm{H}(Y \mid X, Z)$.

3rd definition: $\mathrm{I}(X : Y \mid Z) := \mathrm{H}(X \mid Z) + \mathrm{H}(Y \mid Z) - \mathrm{H}(X, Y \mid Z)$.

Exercise: these three definitions are equivalent.

Properties:

- $\mathrm{I}(X : Y \mid Z) \geq 0$,
- $\mathrm{I}(X : Y \mid Z) = \mathrm{I}(Y : X \mid Z)$,
- $\mathrm{I}(X : Y \mid Z) = \mathrm{H}(X, Z) + \mathrm{H}(Y, Z) - \mathrm{H}(X, Y, Z) - \mathrm{H}(Z)$.

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

Another Definition (triple mutual information):

$$\begin{aligned} \mathrm{I}(X:Y:Z) &:= \mathrm{I}(X:Y) - \mathrm{H}(X:Y \mid Z) \\ &:= \mathrm{I}(XY:Z) - \mathrm{I}(X:Z \mid Y) - \mathrm{I}(Y:Z \mid X) \\ &:= \mathrm{H}(X) + \mathrm{H}(Y) + \mathrm{H}(Z) - \mathrm{H}(X,Y) - \mathrm{H}(X,Z) - \mathrm{H}(Y,Z) \\ &\quad + \mathrm{H}(X,Y,Z) \end{aligned}$$

measure of information: Shannon's approach

Definition: (Shannon entropy)
For a random variable X taking k values with probabilities $p_1, \ldots, p_k$

$$\mathrm{H}(X) := \sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

Another Definition (triple mutual information):

$$\begin{aligned} \mathrm{I}(X:Y:Z) &:= \mathrm{I}(X:Y) - \mathrm{H}(X:Y \mid Z) \\ &:= \mathrm{I}(XY:Z) - \mathrm{I}(X:Z \mid Y) - \mathrm{I}(Y:Z \mid X) \\ &:= \mathrm{H}(X) + \mathrm{H}(Y) + \mathrm{H}(Z) - \mathrm{H}(X,Y) - \mathrm{H}(X,Z) - \mathrm{H}(Y,Z) \\ &\quad + \mathrm{H}(X,Y,Z) \end{aligned}$$

Exercise: these three definitions are equivalent.

Shannon entropy: *entropic profiles*

* For a random variable X the value $\mathrm{H}(X)$ can be any non-negative number

Shannon entropy: *entropic profiles*

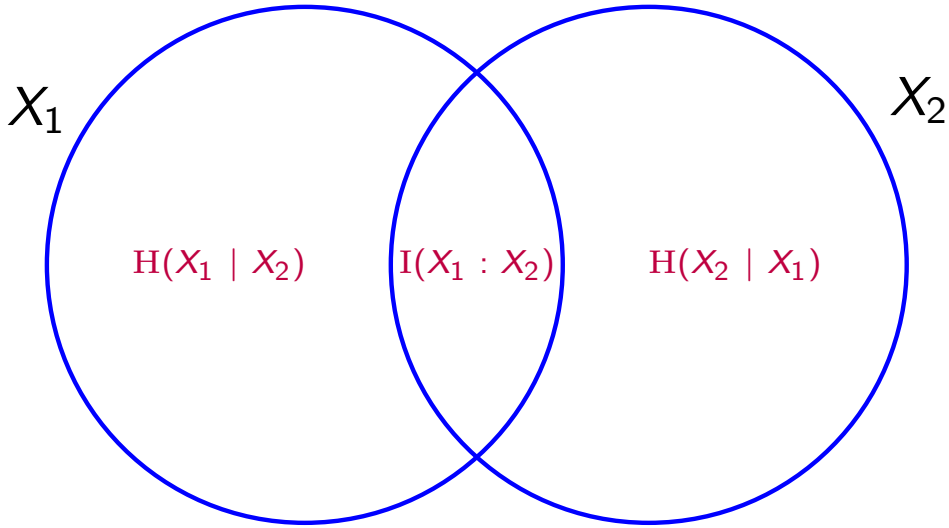
* For a random variable X the value $\mathrm{H}(X)$ can be any non-negative number

* For jointly (X_1, X_2) we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2),\ \mathrm{H}(X_1, X_2)$

$\mathrm{H}(X_1 \mid X_2),\ \mathrm{H}(X_2 \mid X_1)$

$\mathrm{I}(X_1 : X_2)$



Shannon entropy: *entropic profiles*

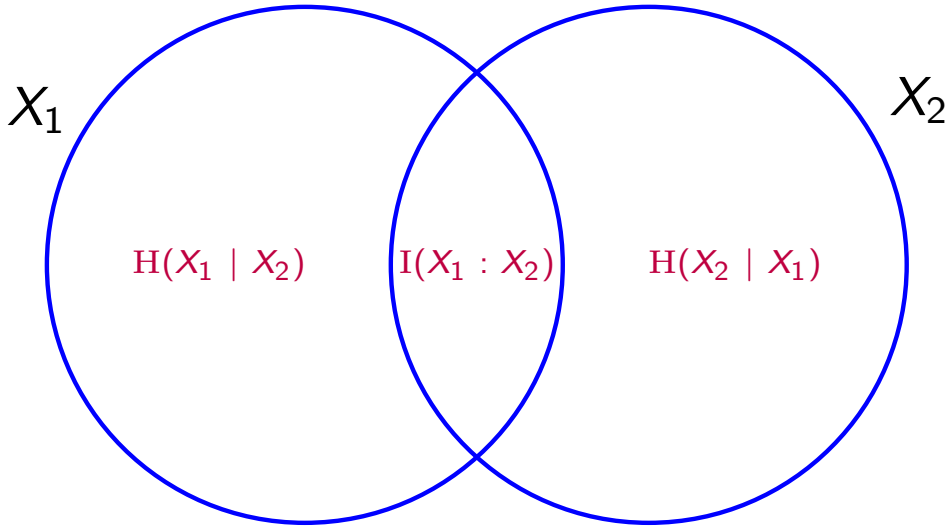
* For a random variable X the value $\mathrm{H}(X)$ can be any non-negative number

* For jointly (X_1, X_2) we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2),\ \mathrm{H}(X_1, X_2)$

$\mathrm{H}(X_1 \mid X_2),\ \mathrm{H}(X_2 \mid X_1)$

$\mathrm{I}(X_1 : X_2)$



but only **3 parameters** are enough

Shannon entropy: *entropic profiles*

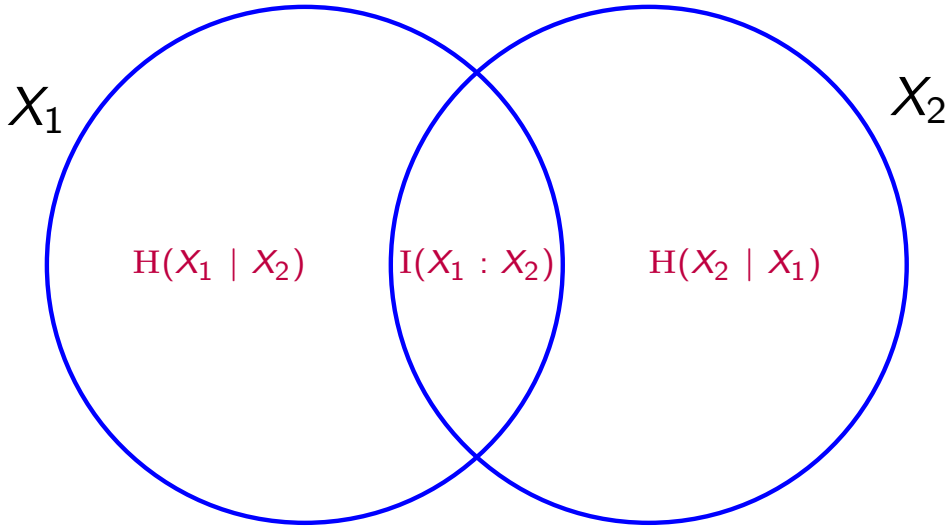
* For a random variable X the value $\mathrm{H}(X)$ can be any non-negative number

* For jointly (X_1, X_2) we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2),\ \mathrm{H}(X_1, X_2)$

$\mathrm{H}(X_1 \mid X_2),\ \mathrm{H}(X_2 \mid X_1)$

$\mathrm{I}(X_1 : X_2)$



but only **3 parameters** are enough

e.g. $\mathrm{H}(X_1), \mathrm{H}(X_2), \mathrm{H}(X_1, X_2)$ allow to find $\mathrm{H}(X_1 \mid X_2)$, $\mathrm{H}(X_2 \mid X_1)$, $\mathrm{I}(X_1 : X_2)$

Shannon entropy: *entropic profiles*

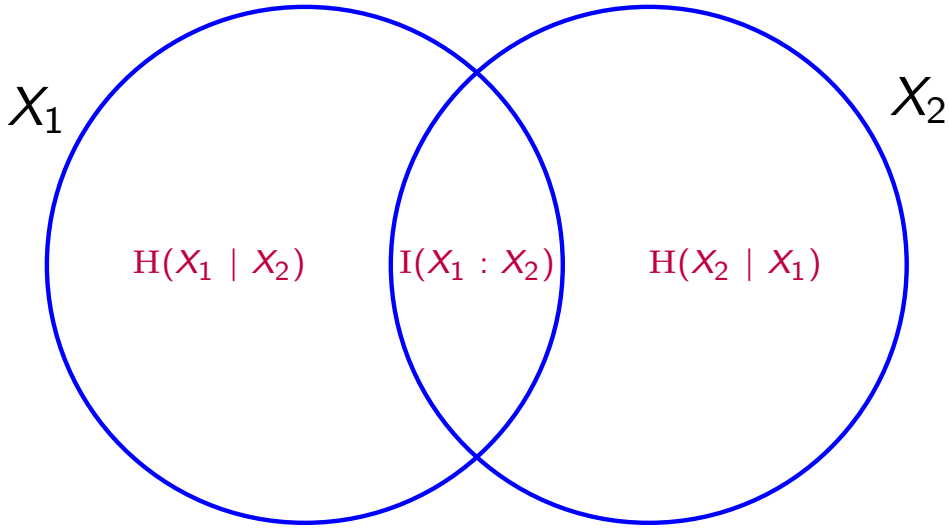
* For a random variable X the value $\mathrm{H}(X)$ can be any non-negative number

* For jointly (X_1, X_2) we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2),\ \mathrm{H}(X_1, X_2)$

$\mathrm{H}(X_1 \mid X_2),\ \mathrm{H}(X_2 \mid X_1)$

$\mathrm{I}(X_1 : X_2)$



but only **3 parameters** are enough

e.g. $\mathrm{H}(X_1), \mathrm{H}(X_2), \mathrm{H}(X_1, X_2)$ allow to find $\mathrm{H}(X_1 \mid X_2)$, $\mathrm{H}(X_2 \mid X_1)$, $\mathrm{I}(X_1 : X_2)$

$\mathrm{H}(X_1 \mid X_2) = \mathrm{H}(X_1, X_2) - \mathrm{H}(X_2)$,

$\mathrm{H}(X_2 \mid X_1) = \mathrm{H}(X_1, X_2) - \mathrm{H}(X_1)$,

$\mathrm{I}(X_1 : X_2) = \mathrm{H}(X_1) + \mathrm{H}(X_2) - \mathrm{H}(X_1, X_2)$

Shannon entropy: *entropic profiles*

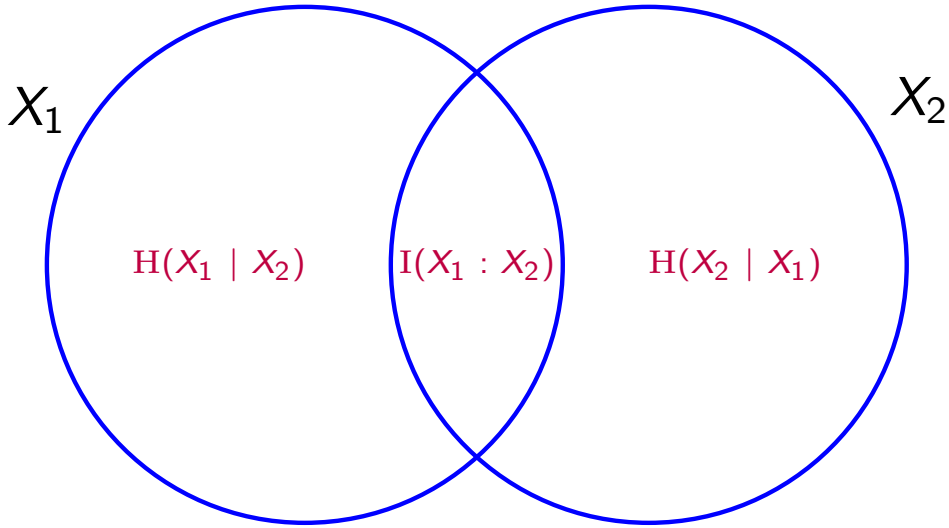
* For a random variable X the value $\mathrm{H}(X)$ can be any non-negative number

* For jointly (X_1, X_2) we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2),\ \mathrm{H}(X_1, X_2)$

$\mathrm{H}(X_1 \mid X_2),\ \mathrm{H}(X_2 \mid X_1)$

$\mathrm{I}(X_1 : X_2)$



but only **3 parameters** are enough

e.g. $\mathrm{H}(X_1), \mathrm{H}(X_2), \mathrm{H}(X_1, X_2)$ allow to find $\mathrm{H}(X_1 \mid X_2)$, $\mathrm{H}(X_2 \mid X_1)$, $\mathrm{I}(X_1 : X_2)$

or $\mathrm{H}(X_1 \mid X_2)$, $\mathrm{H}(X_2 \mid X_1)$ and $\mathrm{I}(X_1 : X_2)$ allow to find $\mathrm{H}(X_1), \mathrm{H}(X_2), \mathrm{H}(X_1, X_2)$.

Shannon entropy: *entropic profiles*

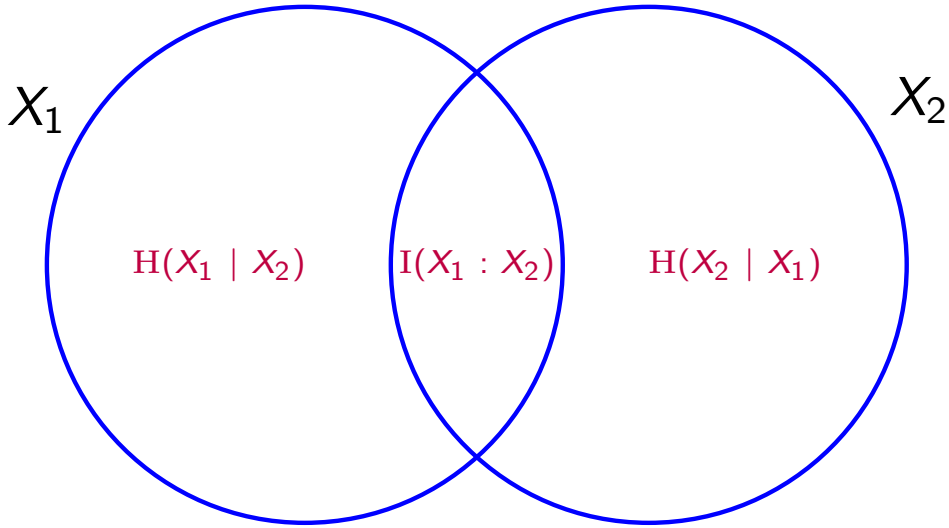
* For a random variable X the value $\mathrm{H}(X)$ can be any non-negative number

* For jointly (X_1, X_2) we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2),\ \mathrm{H}(X_1, X_2)$

$\mathrm{H}(X_1 \mid X_2),\ \mathrm{H}(X_2 \mid X_1)$

$\mathrm{I}(X_1 : X_2)$



but only **3 parameters** are enough

e.g. $\mathrm{H}(X_1), \mathrm{H}(X_2), \mathrm{H}(X_1, X_2)$ allow to find $\mathrm{H}(X_1 \mid X_2)$, $\mathrm{H}(X_2 \mid X_1)$, $\mathrm{I}(X_1 : X_2)$

or $\mathrm{H}(X_1 \mid X_2)$, $\mathrm{H}(X_2 \mid X_1)$ and $\mathrm{I}(X_1 : X_2)$ allow to find $\mathrm{H}(X_1), \mathrm{H}(X_2), \mathrm{H}(X_1, X_2)$.

$\mathrm{H}(X_1) = \mathrm{H}(X_1|X_2) + \mathrm{I}(X_1 : X_2)$,

$\mathrm{H}(X_2) = \mathrm{H}(X_2|X_1) + \mathrm{I}(X_1 : X_2)$,

$\mathrm{H}(X_1, X_2) = \mathrm{H}(X_1 \mid X_2) + \mathrm{I}(X_1 : X_2) + \mathrm{H}(X_2 \mid X_1)$

Shannon entropy: *entropic profiles*

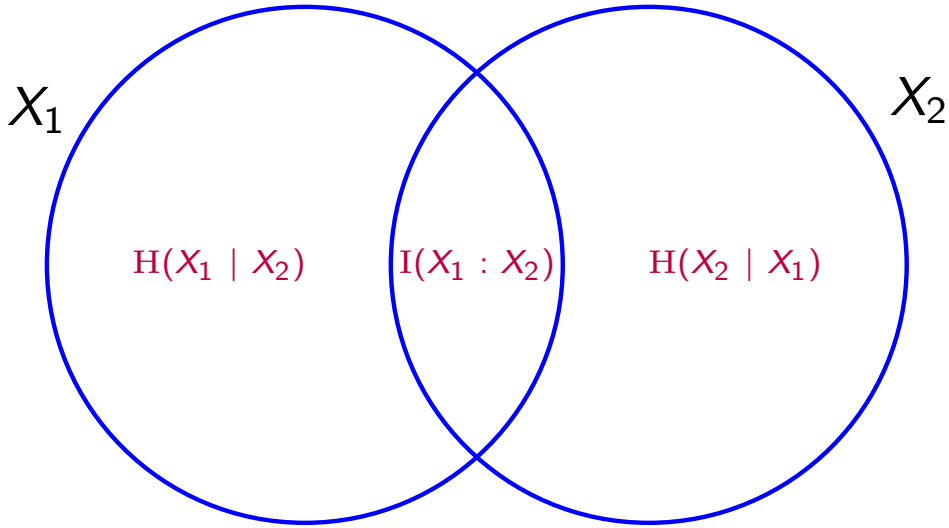
* For a random variable X the value $\mathrm{H}(X)$ can be any non-negative number

* For jointly (X_1, X_2) we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2),\ \mathrm{H}(X_1, X_2)$

$\mathrm{H}(X_1 \mid X_2),\ \mathrm{H}(X_2 \mid X_1)$

$\mathrm{I}(X_1 : X_2)$



but only **3 parameters** are enough

e.g. $\mathrm{H}(X_1), \mathrm{H}(X_2), \mathrm{H}(X_1, X_2)$ allow to find $\mathrm{H}(X_1 \mid X_2)$, $\mathrm{H}(X_2 \mid X_1)$, $\mathrm{I}(X_1 : X_2)$

or $\mathrm{H}(X_1 \mid X_2)$, $\mathrm{H}(X_2 \mid X_1)$ and $\mathrm{I}(X_1 : X_2)$ allow to find $\mathrm{H}(X_1), \mathrm{H}(X_2), \mathrm{H}(X_1, X_2)$.

constraints: $0 \leq \mathrm{H}(X_1),\ \mathrm{H}(X_2) \leq \mathrm{H}(X_1, X_2) \leq \mathrm{H}(X_1) + \mathrm{H}(X_2)$

Shannon entropy: *entropic profiles*

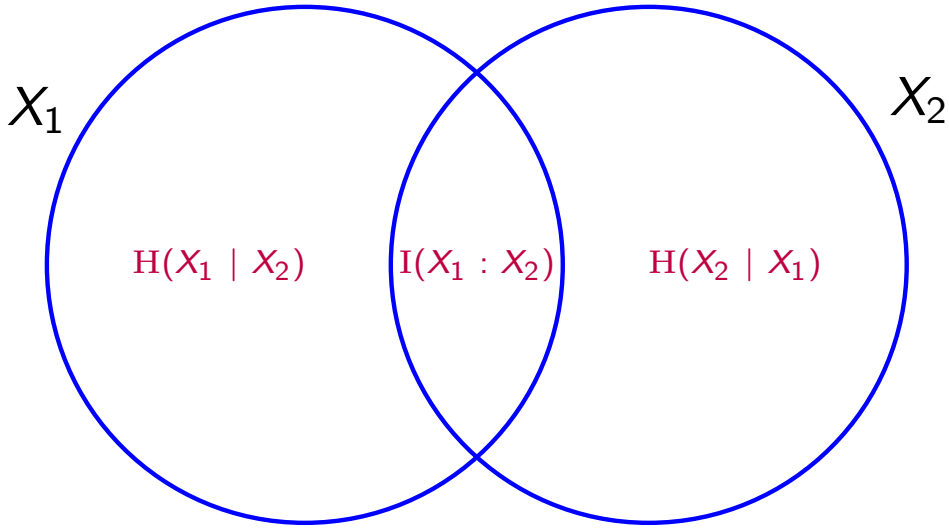
* For a random variable X the value $\mathrm{H}(X)$ can be any non-negative number

* For jointly (X_1, X_2) we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2),\ \mathrm{H}(X_1, X_2)$

$\mathrm{H}(X_1 \mid X_2),\ \mathrm{H}(X_2 \mid X_1)$

$\mathrm{I}(X_1 : X_2)$



but only **3 parameters** are enough

e.g. $\mathrm{H}(X_1), \mathrm{H}(X_2), \mathrm{H}(X_1, X_2)$ allow to find $\mathrm{H}(X_1 \mid X_2)$, $\mathrm{H}(X_2 \mid X_1)$, $\mathrm{I}(X_1 : X_2)$

or $\mathrm{H}(X_1 \mid X_2)$, $\mathrm{H}(X_2 \mid X_1)$ and $\mathrm{I}(X_1 : X_2)$ allow to find $\mathrm{H}(X_1), \mathrm{H}(X_2), \mathrm{H}(X_1, X_2)$.

constraints: $0 \leq \mathrm{H}(X_1),\ \mathrm{H}(X_2) \leq \mathrm{H}(X_1, X_2) \leq \mathrm{H}(X_1) + \mathrm{H}(X_2)$

equivalently: $\mathrm{H}(X_1 \mid X_2) \geq 0,\ \mathrm{H}(X_2 \mid X_1) \geq 0,\ \mathrm{I}(X_1 : X_2) \geq 0$

Shannon entropy: *entropic profiles*

* For jointly (X_1, X_2, X_3) we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2),\ \mathrm{H}(X_3)$

$\mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3),\ \mathrm{H}(X_2, X_3),$

$\mathrm{H}(X_1, X_2, X_3),$

$\mathrm{H}(X_1 \mid X_2),\ \mathrm{H}(X_2 \mid X_1), \ldots,$

$\mathrm{H}(X_1 \mid X_2, X_3), \ldots,$

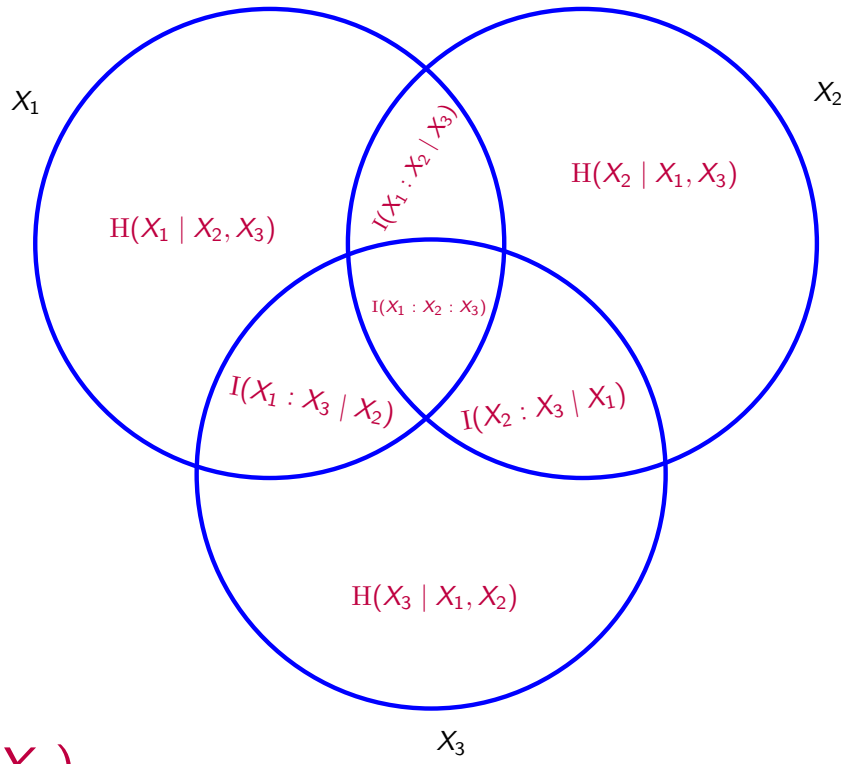
$\mathrm{H}(X_1, X_2 \mid X_3), \ldots,$

$\mathrm{I}(X_1 : X_2),\ \mathrm{I}(X_1 : X_3),\ \mathrm{I}(X_2 : X_3)$

$\mathrm{I}(X_1 : X_2 X_3),\ \mathrm{I}(X_2 : X_1 X_3),\ \mathrm{I}(X_3 : X_1 X_3)$

$\mathrm{I}(X_1 : X_2 \mid X_3),\ \mathrm{I}(X_1 : X_3 \mid X_2),\ \mathrm{I}(X_2 : X_3 \mid X_1)$

$\mathrm{I}(X_1 : X_2 : X_3)$



Shannon entropy: *entropic profiles*

* For jointly (X_1, X_2, X_3) we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2),\ \mathrm{H}(X_3)$

$\mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3),\ \mathrm{H}(X_2, X_3),$

$\mathrm{H}(X_1, X_2, X_3),$

$\mathrm{H}(X_1 \mid X_2),\ \mathrm{H}(X_2 \mid X_1), \ldots,$

$\mathrm{H}(X_1 \mid X_2, X_3), \ldots,$

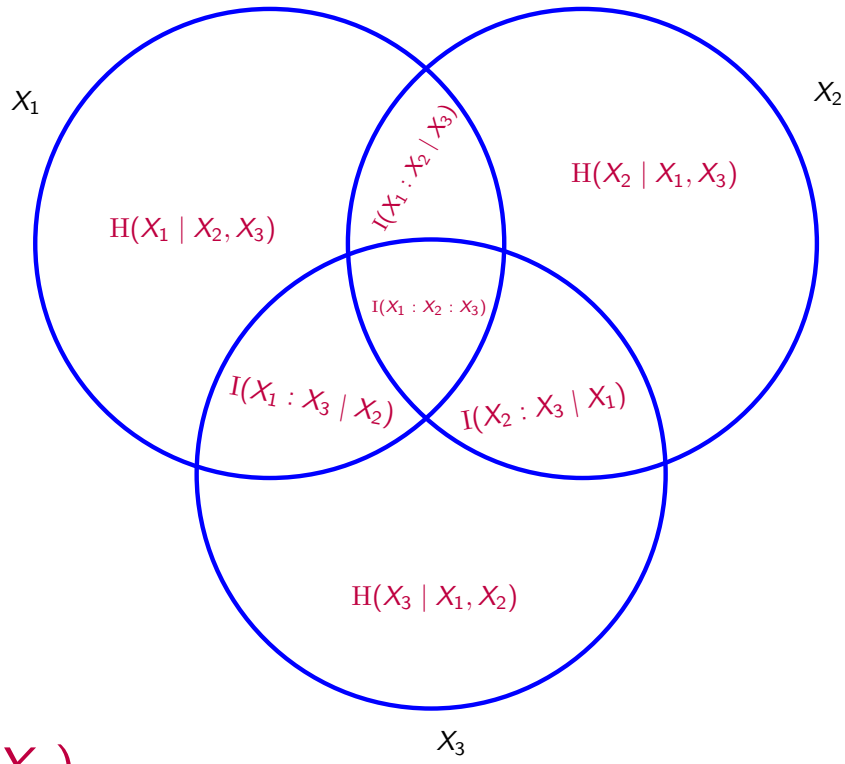
$\mathrm{H}(X_1, X_2 \mid X_3), \ldots,$

$\mathrm{I}(X_1 : X_2),\ \mathrm{I}(X_1 : X_3),\ \mathrm{I}(X_2 : X_3)$

$\mathrm{I}(X_1 : X_2 X_3),\ \mathrm{I}(X_2 : X_1 X_3),\ \mathrm{I}(X_3 : X_1 X_3)$

$\mathrm{I}(X_1 : X_2 \mid X_3),\ \mathrm{I}(X_1 : X_3 \mid X_2),\ \mathrm{I}(X_2 : X_3 \mid X_1)$

$\mathrm{I}(X_1 : X_2 : X_3)$



but only **7 parameters** are enough:

Shannon entropy: *entropic profiles*

* For jointly (X_1, X_2, X_3) we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2),\ \mathrm{H}(X_3)$

$\mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3),\ \mathrm{H}(X_2, X_3),$

$\mathrm{H}(X_1, X_2, X_3),$

$\mathrm{H}(X_1 \mid X_2),\ \mathrm{H}(X_2 \mid X_1), \ldots,$

$\mathrm{H}(X_1 \mid X_2, X_3), \ldots,$

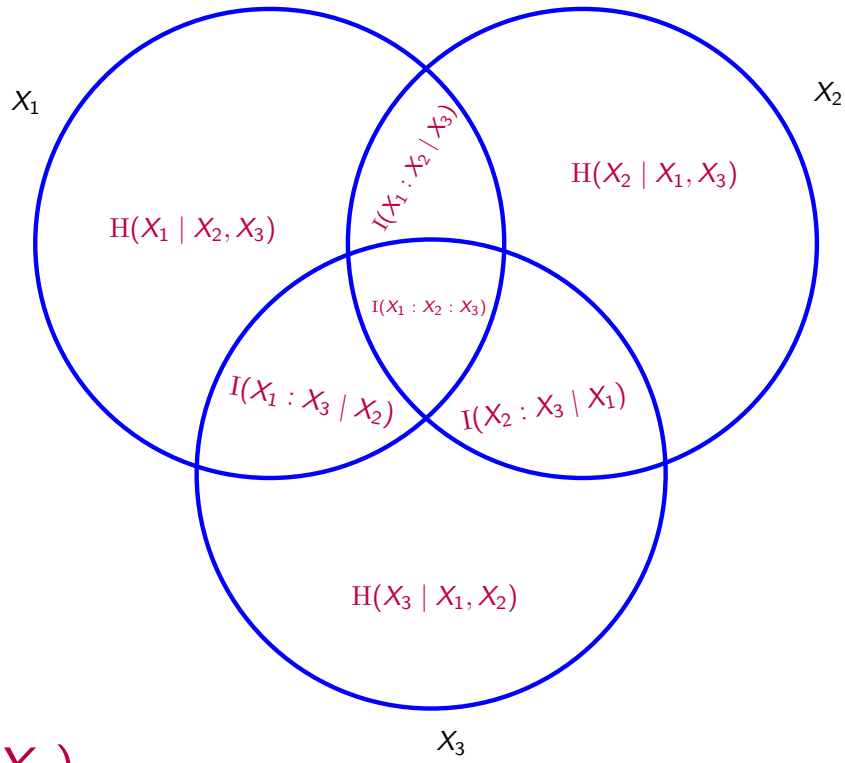
$\mathrm{H}(X_1, X_2 \mid X_3), \ldots,$

$\mathrm{I}(X_1 : X_2),\ \mathrm{I}(X_1 : X_3),\ \mathrm{I}(X_2 : X_3)$

$\mathrm{I}(X_1 : X_2 X_3),\ \mathrm{I}(X_2 : X_1 X_3),\ \mathrm{I}(X_3 : X_1 X_3)$

$\mathrm{I}(X_1 : X_2 \mid X_3),\ \mathrm{I}(X_1 : X_3 \mid X_2),\ \mathrm{I}(X_2 : X_3 \mid X_1)$

$\mathrm{I}(X_1 : X_2 : X_3)$



but only **7 parameters** are enough:

e.g. $\mathrm{H}(X_1), \mathrm{H}(X_2), \ldots$ allow to find all other quantities

Shannon entropy: *entropic profiles*

* For jointly (X_1, X_2, X_3) we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2),\ \mathrm{H}(X_3)$

$\mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3),\ \mathrm{H}(X_2, X_3),$

$\mathrm{H}(X_1, X_2, X_3),$

$\mathrm{H}(X_1 \mid X_2),\ \mathrm{H}(X_2 \mid X_1), \ldots,$

$\mathrm{H}(X_1 \mid X_2, X_3), \ldots,$

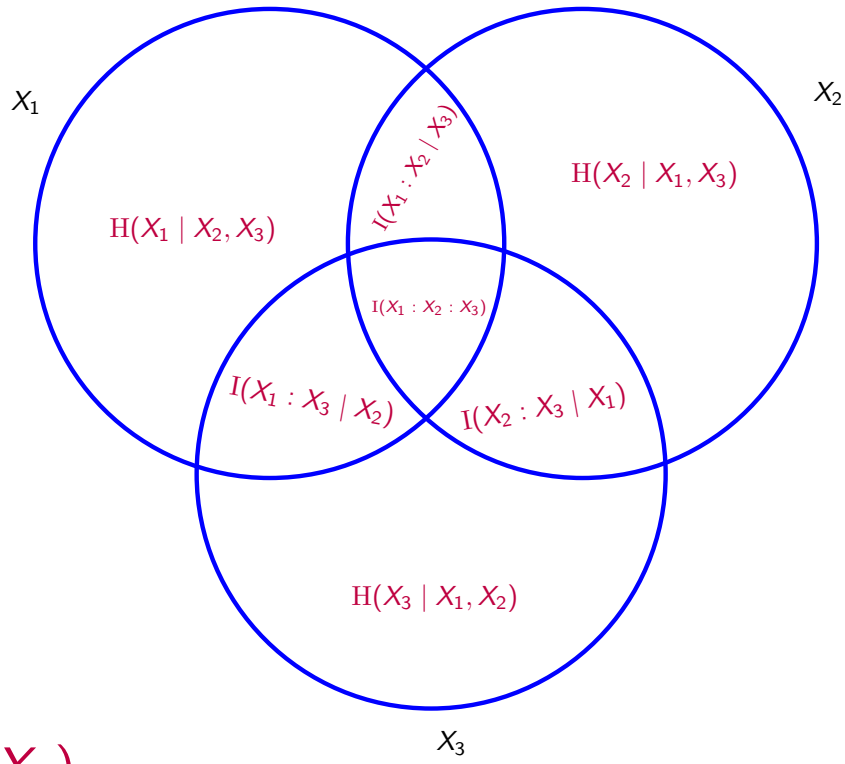
$\mathrm{H}(X_1, X_2 \mid X_3), \ldots,$

$\mathrm{I}(X_1 : X_2),\ \mathrm{I}(X_1 : X_3),\ \mathrm{I}(X_2 : X_3)$

$\mathrm{I}(X_1 : X_2 X_3),\ \mathrm{I}(X_2 : X_1 X_3),\ \mathrm{I}(X_3 : X_1 X_3)$

$\mathrm{I}(X_1 : X_2 \mid X_3),\ \mathrm{I}(X_1 : X_3 \mid X_2),\ \mathrm{I}(X_2 : X_3 \mid X_1)$

$\mathrm{I}(X_1 : X_2 : X_3)$



but only **7 parameters** are enough:

e.g. $\mathrm{H}(X_1), \mathrm{H}(X_2), \ldots$ allow to find all other quantities

or $\mathrm{H}(X_1 | X_2, X_3), \ldots, \mathrm{I}(X_1 : X_2 \mid X_3), \ldots, \mathrm{I}(X_1 : X_2 : X_3)$
allow to find all other quantities

Shannon entropy: *entropic profiles*

* For jointly (X_1, X_2, X_3) we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2),\ \mathrm{H}(X_3)$

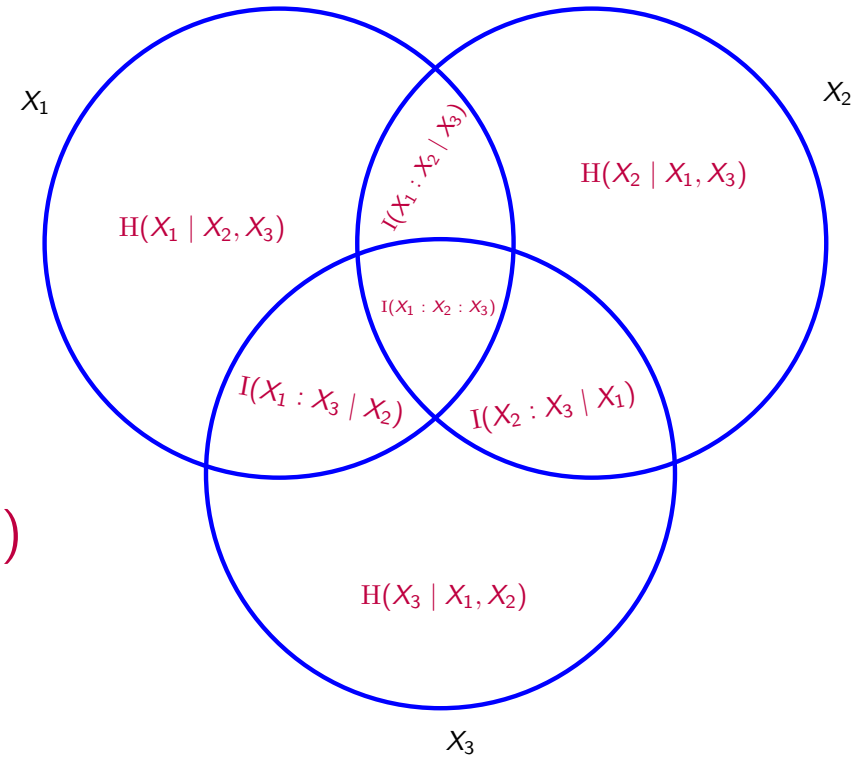
$\mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3),\ \mathrm{H}(X_2, X_3),$

$\mathrm{H}(X_1, X_2, X_3),$

$\mathrm{H}(X_1 \mid X_2, X_3), \ldots,$

$\mathrm{I}(X_1 : X_2 \mid X_3),\ \mathrm{I}(X_1 : X_3 \mid X_2),\ \mathrm{I}(X_2 : X_3 \mid X_1)$

$\mathrm{I}(X_1 : X_2 : X_3)$



7 parameters define the profile:

e.g. $\mathrm{H}(X_1), \mathrm{H}(X_2), \ldots$ allow to find all other quantities

or $\mathrm{H}(X_1|X_2, X_3), \ldots, \mathrm{I}(X_1 : X_2 \mid X_3), \ldots, \mathrm{I}(X_1 : X_2 : X_3)$
allow to find all other quantities

Shannon entropy: *entropic profiles*

* For jointly (X_1, X_2, X_3) we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2),\ \mathrm{H}(X_3)$

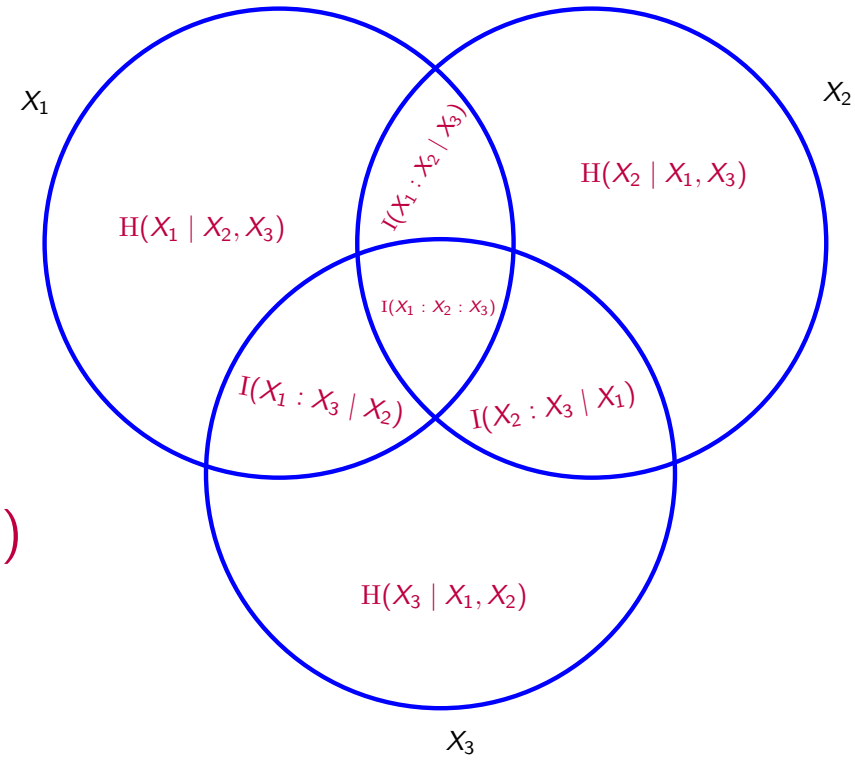
$\mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3),\ \mathrm{H}(X_2, X_3),$

$\mathrm{H}(X_1, X_2, X_3),$

$\mathrm{H}(X_1 \mid X_2, X_3), \ldots,$

$\mathrm{I}(X_1 : X_2 \mid X_3),\ \mathrm{I}(X_1 : X_3 \mid X_2),\ \mathrm{I}(X_2 : X_3 \mid X_1)$

$\mathrm{I}(X_1 : X_2 : X_3)$



7 parameters define the profile:

e.g. $\mathrm{H}(X_1), \mathrm{H}(X_2), \ldots$ allow to find all other quantities

or $\mathrm{H}(X_1|X_2, X_3), \ldots, \mathrm{I}(X_1 : X_2 \mid X_3), \ldots, \mathrm{I}(X_1 : X_2 : X_3)$
allow to find all other quantities

9 constraints: $\mathrm{H}(X_1|X_2, X_3) \geq 0, \ldots, \mathrm{I}(X_1 : X_2) \geq 0, \ldots, \mathrm{I}(X_1 : X_2 \mid X_3) \geq 0, \ldots$

Shannon entropy: *entropic profiles*

* For jointly (X_1, X_2, X_3) we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2),\ \mathrm{H}(X_3)$

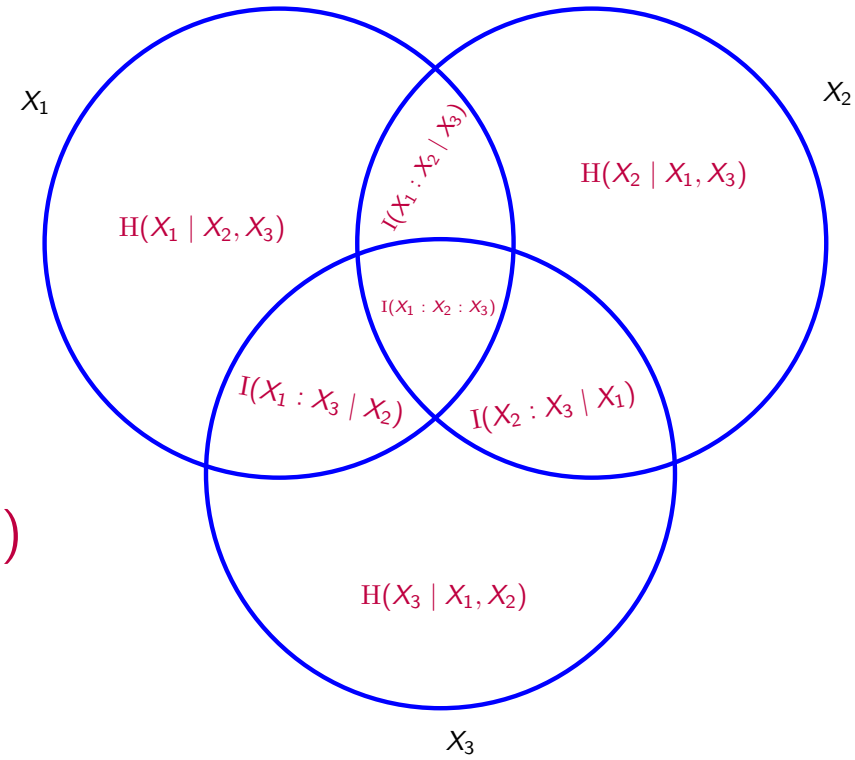
$\mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3),\ \mathrm{H}(X_2, X_3),$

$\mathrm{H}(X_1, X_2, X_3),$

$\mathrm{H}(X_1 \mid X_2, X_3), \ldots,$

$\mathrm{I}(X_1 : X_2 \mid X_3),\ \mathrm{I}(X_1 : X_3 \mid X_2),\ \mathrm{I}(X_2 : X_3 \mid X_1)$

$\mathrm{I}(X_1 : X_2 : X_3)$



7 parameters define the profile:

e.g. $\mathrm{H}(X_1), \mathrm{H}(X_2), \ldots$ allow to find all other quantities

or $\mathrm{H}(X_1|X_2, X_3), \ldots, \mathrm{I}(X_1 : X_2 \mid X_3), \ldots, \mathrm{I}(X_1 : X_2 : X_3)$
allow to find all other quantities

9 constraints: $\mathrm{H}(X_1|X_2, X_3) \geq 0, \ldots, \mathrm{I}(X_1 : X_2) \geq 0, \ldots, \mathrm{I}(X_1 : X_2 \mid X_3) \geq 0, \ldots$

Exercise: no other inequalities for entropies of (X_1, X_2, X_3)

Shannon entropy: *entropic profiles*

* For jointly distributed $(X_1, \ldots, X_n)$ we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2), \ldots$
$\mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3), \ldots,$
$\mathrm{H}(X_1, X_2, X_3), \ldots$
$\ldots$

Shannon entropy: *entropic profiles*

* For jointly distributed $(X_1, \ldots, X_n)$ we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2), \ldots$
$\mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3), \ldots,$
$\mathrm{H}(X_1, X_2, X_3), \ldots$
$\ldots$

$2^n - 1$ **parameters** that define the profile

Shannon entropy: *entropic profiles*

* For jointly distributed $(X_1, \ldots, X_n)$ we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2), \ldots$
$\mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3), \ldots,$
$\mathrm{H}(X_1, X_2, X_3), \ldots$
$\ldots$

$2^n - 1$ **parameters** that define the profile

classical constraints:

Shannon entropy: *entropic profiles*

* For jointly distributed $(X_1, \ldots, X_n)$ we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2), \ldots$
$\mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3), \ldots,$
$\mathrm{H}(X_1, X_2, X_3), \ldots$
$\ldots$

$2^n - 1$ **parameters** that define the profile

classical constraints:

- monotonicity:
  $\mathrm{H}(X_{i_1} \ldots X_{i_m}) \leq \mathrm{H}(X_{i_1} \ldots X_{i_m} X_{j_1} \ldots X_{j_s})$

Shannon entropy: *entropic profiles*

* For jointly distributed $(X_1, \ldots, X_n)$ we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2), \ldots$
$\mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3), \ldots,$
$\mathrm{H}(X_1, X_2, X_3), \ldots$
$\ldots$

$2^n - 1$ **parameters** that define the profile

classical constraints:

- monotonicity:
  $\mathrm{H}(X_{i_1} \ldots X_{i_m}) \leq \mathrm{H}(X_{i_1} \ldots X_{i_m} X_{j_1} \ldots X_{j_s})$
- concavity:
  $\mathrm{H}(X_{i_1} \ldots X_{i_m} X_{j_1} \ldots X_{j_s}) \leq \mathrm{H}(X_{i_1} \ldots X_{i_m}) + \mathrm{H}(X_{j_1} \ldots X_{j_s})$

Shannon entropy: *entropic profiles*

* For jointly distributed $(X_1, \ldots, X_n)$ we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2), \ldots$
$\mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3), \ldots,$
$\mathrm{H}(X_1, X_2, X_3), \ldots$
$\ldots$

$2^n - 1$ **parameters** that define the profile

classical constraints:

- monotonicity:
  $\mathrm{H}(X_{i_1} \ldots X_{i_m}) \leq \mathrm{H}(X_{i_1} \ldots X_{i_m} X_{j_1} \ldots X_{j_s})$
- concavity:
  $\mathrm{H}(X_{i_1} \ldots X_{i_m} X_{j_1} \ldots X_{j_s}) \leq \mathrm{H}(X_{i_1} \ldots X_{i_m}) + \mathrm{H}(X_{j_1} \ldots X_{j_s})$
- submodularity:
  $$\mathrm{H}(X_{i_1} \ldots X_{i_m} X_{j_1} \ldots X_{j_s} X_{k_1} \ldots X_{k_\ell}) + \mathrm{H}(X_{k_1} \ldots X_{k_\ell}) \leq$$
  $$\leq \mathrm{H}(X_{i_1} \ldots X_{i_m} X_{k_1} \ldots X_{k_\ell}) + \mathrm{H}(X_{j_1} \ldots X_{j_s} X_{k_1} \ldots X_{k_\ell})$$

Shannon entropy: *entropic profiles*

* For jointly distributed $(X_1, \ldots, X_n)$ we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2), \ldots$
$\mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3), \ldots,$
$\mathrm{H}(X_1, X_2, X_3), \ldots$
$\ldots$

$2^n - 1$ **parameters** that define the profile

classical constraints:

for $\mathcal{M} = \{i_1, \ldots, i_m\}$ we denote $\mathrm{H}(X_{\mathcal{I}}) = \mathrm{H}(X_{i_1} \ldots X_{i_m})$

- monotonicity:
  $\mathrm{H}(X_{\mathcal{I}}) \leq \mathrm{H}(X_{\mathcal{I}\cup\mathcal{J}})$ or equivalently $\mathrm{H}(X_{\mathcal{J}} \mid X_{\mathcal{I}}) \geq 0$
- concavity:
  $\mathrm{H}(X_{\mathcal{I}\cup\mathcal{J}}) \leq \mathrm{H}(X_{\mathcal{I}}) + \mathrm{H}(X_{\mathcal{J}})$ or equivalently $\mathrm{I}(X_{\mathcal{I}} : X_{\mathcal{J}}) \geq 0$
- submodularity:
  $\mathrm{H}(X_{\mathcal{I}\cup\mathcal{J}\cup\mathcal{K}}) + \mathrm{H}(X_{\mathcal{K}}) \leq \mathrm{H}(X_{\mathcal{I}\cup\mathcal{K}}) + \mathrm{H}(X_{\mathcal{J}\cup\mathcal{K}})$
  or equivalently $\mathrm{I}(X_{\mathcal{I}} : X_{\mathcal{J}} \mid X_{\mathcal{K}}) \geq 0.$

Shannon entropy: *entropic profiles*

* For jointly distributed $(X_1, \ldots, X_n)$ we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2), \ldots$

$\mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3), \ldots,$

$\mathrm{H}(X_1, X_2, X_3), \ldots$

$\ldots$

Shannon entropy: *entropic profiles*

* For jointly distributed $(X_1, \ldots, X_n)$ we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2), \ldots$

$\mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3), \ldots,$

$\mathrm{H}(X_1, X_2, X_3), \ldots$

$\ldots$

$2^n - 1$ **parameters** that define the profile

Shannon entropy: *entropic profiles*

* For jointly distributed $(X_1, \ldots, X_n)$ we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2), \ldots$
$\mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3), \ldots,$
$\mathrm{H}(X_1, X_2, X_3), \ldots$
$\ldots$

$2^n - 1$ **parameters** that define the profile

Shannon, the 1940s:

- $\mathrm{H}(X_{\mathcal{J}} \mid X_{\mathcal{I}}) \geq 0$
- $\mathrm{I}(X_{\mathcal{I}} : X_{\mathcal{J}}) \geq 0$
- $\mathrm{I}(X_{\mathcal{I}} : X_{\mathcal{J}} \mid X_{\mathcal{K}}) \geq 0$

Shannon entropy: *entropic profiles*

* For jointly distributed $(X_1, \ldots, X_n)$ we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2), \ldots$
$\mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3), \ldots,$
$\mathrm{H}(X_1, X_2, X_3), \ldots$
$\ldots$

$2^n - 1$ **parameters** that define the profile

Shannon, the 1940s:

- $\mathrm{H}(X_{\mathcal{J}} \mid X_{\mathcal{I}}) \geq 0$
- $\mathrm{I}(X_{\mathcal{I}} : X_{\mathcal{J}}) \geq 0$
- $\mathrm{I}(X_{\mathcal{I}} : X_{\mathcal{J}} \mid X_{\mathcal{K}}) \geq 0$

Question: Are there any other inequalities?

Shannon entropy: *entropic profiles*

* For jointly distributed $(X_1, \ldots, X_n)$ we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2), \ldots$
$\mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3), \ldots,$
$\mathrm{H}(X_1, X_2, X_3), \ldots$
$\ldots$

$2^n - 1$ **parameters** that define the profile

Shannon, the 1940s:

- $\mathrm{H}(X_{\mathcal{J}} \mid X_{\mathcal{I}}) \geq 0$
- $\mathrm{I}(X_{\mathcal{I}} : X_{\mathcal{J}}) \geq 0$
- $\mathrm{I}(X_{\mathcal{I}} : X_{\mathcal{J}} \mid X_{\mathcal{K}}) \geq 0$

Question: Are there any other inequalities?

Zhang–Yeung 1998: Yes, another inequality for $n = 4$ random variables!

Shannon entropy: *entropic profiles*

* For jointly distributed $(X_1, \ldots, X_n)$ we have

$\mathrm{H}(X_1),\ \mathrm{H}(X_2), \ldots$
$\mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3), \ldots,$
$\mathrm{H}(X_1, X_2, X_3), \ldots$
$\ldots$

$2^n - 1$ **parameters** that define the profile

Shannon, the 1940s:

- $\mathrm{H}(X_{\mathcal{J}} \mid X_{\mathcal{I}}) \geq 0$
- $\mathrm{I}(X_{\mathcal{I}} : X_{\mathcal{J}}) \geq 0$
- $\mathrm{I}(X_{\mathcal{I}} : X_{\mathcal{J}} \mid X_{\mathcal{K}}) \geq 0$

Question: Are there any other inequalities?

Zhang–Yeung 1998: Yes, another inequality for $n = 4$ random variables!

Matus 2007: **infinitely many** inequalities with $n \geq 4$ random variables!!

Shannon entropy: *entropic profiles*

* For jointly distributed $(X_1, \ldots, X_n)$ we have

$$\mathrm{H}(X_1),\ \mathrm{H}(X_2), \ldots\ \mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3), \ldots,\ \mathrm{H}(X_1, X_2, X_3), \ldots\ \ldots$$

Shannon entropy: *entropic profiles*

* For jointly distributed $(X_1, \ldots, X_n)$ we have

$$\mathrm{H}(X_1),\ \mathrm{H}(X_2), \ldots\ \mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3), \ldots,\ \mathrm{H}(X_1, X_2, X_3), \ldots\ \ldots$$

this vector of $2^n - 1$ **reals** $\Rightarrow$ an **entropic profile**

Shannon entropy: *entropic profiles*

* For jointly distributed $(X_1, \ldots, X_n)$ we have

$$\mathrm{H}(X_1),\ \mathrm{H}(X_2), \ldots\ \mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3), \ldots,\ \mathrm{H}(X_1, X_2, X_3), \ldots\ \ldots$$

this vector of $2^n - 1$ **reals** $\Rightarrow$ an **entropic profile**

Notation: $\mathbf{P}_n :=$ entropic profiles for all $(X_1, \ldots, X_n)$

Shannon entropy: *entropic profiles*

* For jointly distributed $(X_1, \ldots, X_n)$ we have

$$\mathrm{H}(X_1),\ \mathrm{H}(X_2), \ldots\ \mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3), \ldots,\ \mathrm{H}(X_1, X_2, X_3), \ldots\ \ldots$$

this vector of $2^n - 1$ **reals** $\Rightarrow$ an **entropic profile**

Notation: $\mathbf{P}_n :=$ entropic profiles for all $(X_1, \ldots, X_n)$

Theorem. $\bar{\mathbf{P}}_n$ (topological closure) is a *convex cone*.

Shannon entropy: *entropic profiles*

* For jointly distributed $(X_1, \ldots, X_n)$ we have

$$\mathrm{H}(X_1),\ \mathrm{H}(X_2), \ldots\ \mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3), \ldots,\ \mathrm{H}(X_1, X_2, X_3), \ldots\ \ldots$$

this vector of $2^n - 1$ **reals** $\Rightarrow$ an **entropic profile**

Notation: $\mathbf{P}_n :=$ entropic profiles for all $(X_1, \ldots, X_n)$

Theorem. $\bar{\mathbf{P}}_n$ (topological closure) is a *convex cone*.

Exercise: prove it.

Shannon entropy: *entropic profiles*

* For jointly distributed $(X_1, \dots, X_n)$ we have

$$\mathrm{H}(X_1),\ \mathrm{H}(X_2), \dots\ \mathrm{H}(X_1, X_2),\ \mathrm{H}(X_1, X_3), \dots,\ \mathrm{H}(X_1, X_2, X_3), \dots\ \dots$$

this vector of $2^n - 1$ **reals** $\Rightarrow$ an **entropic profile**

Notation: $\mathbf{P}_n :=$ entropic profiles for all $(X_1, \dots, X_n)$

Theorem. $\bar{\mathbf{P}}_n$ (topological closure) is a *convex cone*.

Exercise: prove it.

Fact:

- the case $n \leq 3$ is simple : $\bar{\mathbf{P}}_1$, $\bar{\mathbf{P}}_2$, $\bar{\mathbf{P}}_3$ are defined by Shannon's inequalities $\mathrm{H}(X_\mathcal{J} \mid X_\mathcal{I}) \geq 0$, $\mathrm{I}(X_\mathcal{I} : X_\mathcal{J}) \geq 0$, $\mathrm{I}(X_\mathcal{I} : X_\mathcal{J} \mid X_\mathcal{K}) \geq 0$
- the case $n \geq 4$ is hard : $\bar{\mathbf{P}}_n$ is not polyhedral

  (and not filly understood)

Shannon entropy: *symmetric key encryption*

setting:

- **Sender** and **Receiver** sample a common secret **key**

Shannon entropy: *symmetric key encryption*

setting:

- **Sender** and **Receiver** sample a common secret **key**
- Sender wants to send a random **message**

Shannon entropy: *symmetric key encryption*

setting:

- **Sender** and **Receiver** sample a common secret **key**
- Sender wants to send a random **message**
- Sender transmits **ciphertext** $= \mathrm{Enc}(\textbf{message}, \textbf{key})$

Shannon entropy: *symmetric key encryption*

setting:

- **Sender** and **Receiver** sample a common secret **key**
- Sender wants to send a random **message**
- Sender transmits **ciphertext** $= \mathrm{Enc}($**message**, **key**$)$
- Receivers computes **message** $= \mathrm{Dec}($**ciphertext**, **key**$)$

Shannon entropy: *symmetric key encryption*

setting:

- **Sender** and **Receiver** sample a common secret **key**
- Sender wants to send a random **message**
- Sender transmits **ciphertext** $= \mathrm{Enc}($**message**, **key**$)$
- Receivers computes **message** $= \mathrm{Dec}($**ciphertext**, **key**$)$

requirements:

- **ciphertext** is a function of (**message**, **key**)
- **message** is a function of (**ciphertext**, **key**)
- **message** and **ciphertext** are independent

Shannon entropy: *symmetric key encryption*

setting:

- **Sender** and **Receiver** sample a common secret **key**
- Sender wants to send a random **message**
- Sender transmits **ciphertext** $= \mathrm{Enc}(\textbf{message}, \textbf{key})$
- Receivers computes **message** $= \mathrm{Dec}(\textbf{ciphertext}, \textbf{key})$

requirements:

- **ciphertext** is a function of (**message**, **key**)
- **message** is a function of (**ciphertext**, **key**)
- **message** and **ciphertext** are independent

solution: Vernam's scheme / one-time pad,
where **message**, **key**, **ciphertext** $\in \{0,1\}^n$

- **ciphertext** $=$ bitwise XOR of **message** and **key**
- **key** $=$ bitwise XOR of **message** and **ciphertext**

Shannon entropy: *symmetric key encryption*

setting:

- **Sender** and **Receiver** sample a common secret **key**
- Sender wants to send a random **message**
- Sender transmits **ciphertext** $= \mathrm{Enc}(\textbf{message}, \textbf{key})$
- Receivers computes **message** $= \mathrm{Dec}(\textbf{ciphertext}, \textbf{key})$

requirements:

- **ciphertext** is a function of $(\textbf{message}, \textbf{key})$
- **message** is a function of $(\textbf{ciphertext}, \textbf{key})$
- **message** and **ciphertext** are independent

Theorem (Shannon): $\mathrm{H}(\textbf{key}) \geq \mathrm{H}(\textbf{message})$

Shannon entropy: *perfect secret sharing* (threshold scheme)

setting: fix integer numbers n and $t \leq n$.

- **Dealer** samples a secret key S_0 and shares $S_1, \ldots S_n$
- for $i = 1, \ldots, n$ the i-th **participant** gets the share S_i
- correctness: every t **participants** together can compute S_0
  $\mathrm{H}(S_0 \mid S_{i_1} \ldots S_{i_t}) = 0$
- security: every $\leq (t-1)$ **participants** together have no information on S_0
  $\mathrm{H}(S_0 \mid S_{j_1} \ldots S_{j_{t-1}}) = \mathrm{H}(S_0)$

Shannon entropy: *perfect secret sharing* (threshold scheme)

setting: fix integer numbers n and $t \leq n$.

- **Dealer** samples a secret key S_0 and shares $S_1, \ldots S_n$
- for $i = 1, \ldots, n$ the i-th **participant** gets the share S_i
- correctness: every t **participants** together can compute S_0
  $\mathrm{H}(S_0 \mid S_{i_1} \ldots S_{i_t}) = 0$
- security: every $\leq (t-1)$ **participants** together have no information on S_0
  $\mathrm{H}(S_0 \mid S_{j_1} \ldots S_{j_{t-1}}) = \mathrm{H}(S_0)$

solution (Shamir) : fix a field $\mathbb{F}$ (with $> n$ elements)

Shannon entropy: *perfect secret sharing* (threshold scheme)

setting: fix integer numbers n and $t \leq n$.

- **Dealer** samples a secret key S_0 and shares $S_1, \ldots S_n$
- for $i = 1, \ldots, n$ the i-th **participant** gets the share S_i
- correctness: every t **participants** together can compute S_0
  $\mathrm{H}(S_0 \mid S_{i_1} \ldots S_{i_t}) = 0$
- security: every $\leq (t-1)$ **participants** together have no information on S_0
  $\mathrm{H}(S_0 \mid S_{j_1} \ldots S_{j_{t-1}}) = \mathrm{H}(S_0)$

solution (Shamir) : fix a field $\mathbb{F}$ (with $> n$ elements)
fix elements $x_0, x_1, \ldots, x_n \in \mathbb{F}$

Shannon entropy: *perfect secret sharing* (threshold scheme)

setting: fix integer numbers n and $t \leq n$.

- **Dealer** samples a secret key S_0 and shares $S_1, \ldots S_n$
- for $i = 1, \ldots, n$ the i-th **participant** gets the share S_i
- correctness: every t **participants** together can compute S_0
  $\mathrm{H}(S_0 \mid S_{i_1} \ldots S_{i_t}) = 0$
- security: every $\leq (t-1)$ **participants** together have no information on S_0
  $\mathrm{H}(S_0 \mid S_{j_1} \ldots S_{j_{t-1}}) = \mathrm{H}(S_0)$

solution (Shamir) : fix a field $\mathbb{F}$ (with $> n$ elements)
fix elements $x_0, x_1, \ldots, x_n \in \mathbb{F}$

Dealer samples a random polynomial
$P(x) = a_0 + a_1 x + \ldots + a_{t-1} x^{t-1}$ over $\mathbb{F}$

Shannon entropy: *perfect secret sharing* (threshold scheme)

setting: fix integer numbers n and $t \leq n$.

- **Dealer** samples a secret key S_0 and shares $S_1, \ldots S_n$
- for $i = 1, \ldots, n$ the i-th **participant** gets the share S_i
- correctness: every t **participants** together can compute S_0
  $\mathrm{H}(S_0 \mid S_{i_1} \ldots S_{i_t}) = 0$
- security: every $\leq (t-1)$ **participants** together have no information on S_0
  $\mathrm{H}(S_0 \mid S_{j_1} \ldots S_{j_{t-1}}) = \mathrm{H}(S_0)$

solution (Shamir) : fix a field $\mathbb{F}$ (with $> n$ elements)
fix elements $x_0, x_1, \ldots, x_n \in \mathbb{F}$

Dealer samples a random polynomial
$P(x) = a_0 + a_1 x + \ldots + a_{t-1} x^{t-1}$ over $\mathbb{F}$

secret key $S_0 := P(x_0)$

Shannon entropy: *perfect secret sharing* (threshold scheme)

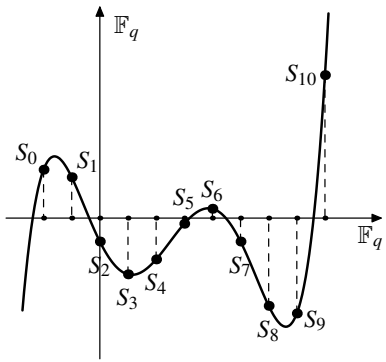
setting: fix integer numbers n and $t \leq n$.

- **Dealer** samples a secret key S_0 and shares $S_1, \ldots S_n$
- for $i = 1, \ldots, n$ the i-th **participant** gets the share S_i
- correctness: every t **participants** together can compute S_0
  $\mathrm{H}(S_0 \mid S_{i_1} \ldots S_{i_t}) = 0$
- security: every $\leq (t-1)$ **participants** together have no information on S_0
  $\mathrm{H}(S_0 \mid S_{j_1} \ldots S_{j_{t-1}}) = \mathrm{H}(S_0)$

solution (Shamir) : fix a field $\mathbb{F}$ (with $> n$ elements)
fix elements $x_0, x_1, \ldots, x_n \in \mathbb{F}$

Dealer samples a random polynomial
$P(x) = a_0 + a_1 x + \ldots + a_{t-1} x^{t-1}$ over $\mathbb{F}$

secret key $S_0 := P(x_0)$ shares
$S_i = P(x_i)$



Shannon entropy: *perfect secret sharing* (threshold scheme)

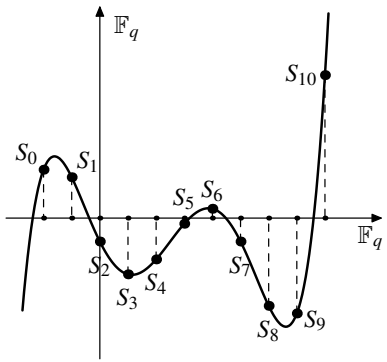
setting: fix integer numbers n and $t \leq n$.

- **Dealer** samples a secret key S_0 and shares $S_1, \ldots S_n$
- for $i = 1, \ldots, n$ the i-th **participant** gets the share S_i
- correctness: every t **participants** together can compute S_0
  $\mathrm{H}(S_0 \mid S_{i_1} \ldots S_{i_t}) = 0$
- security: every $\leq (t-1)$ **participants** together have no information on S_0
  $\mathrm{H}(S_0 \mid S_{j_1} \ldots S_{j_{t-1}}) = \mathrm{H}(S_0)$

solution (Shamir) : fix a field $\mathbb{F}$ (with $> n$ elements)
fix elements $x_0, x_1, \ldots, x_n \in \mathbb{F}$

Dealer samples a random polynomial
$P(x) = a_0 + a_1 x + \ldots + a_{t-1} x^{t-1}$ over $\mathbb{F}$

secret key $S_0 := P(x_0)$ shares
$S_i = P(x_i)$



Exercise: prove correctness and security of the Shamir scheme

Shannon entropy: *perfect secret sharing* (threshold scheme)

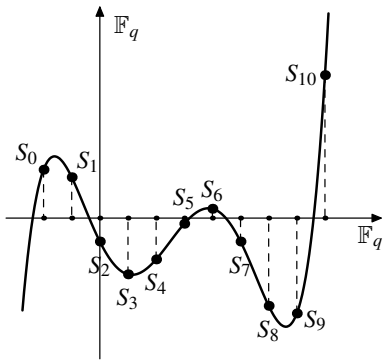
setting: fix integer numbers n and $t \leq n$.

- **Dealer** samples a secret key S_0 and shares $S_1, \ldots S_n$
- for $i = 1, \ldots, n$ the i-th **participant** gets the share S_i
- correctness: every t **participants** together can compute S_0
  $\mathrm{H}(S_0 \mid S_{i_1} \ldots S_{i_t}) = 0$
- security: every $\leq (t-1)$ **participants** together have no information on S_0
  $\mathrm{H}(S_0 \mid S_{j_1} \ldots S_{j_{t-1}}) = \mathrm{H}(S_0)$

solution (Shamir) : fix a field $\mathbb{F}$ (with $> n$ elements)
fix elements $x_0, x_1, \ldots, x_n \in \mathbb{F}$

Dealer samples a random polynomial
$P(x) = a_0 + a_1 x + \ldots + a_{t-1} x^{t-1}$ over $\mathbb{F}$

secret key $S_0 := P(x_0)$ shares
$S_i = P(x_i)$



Exercise: prove correctness and security of the Shamir scheme

Exercise: prove that in *every* correct and secure scheme of secret sharing
$\mathrm{H}(S_i) \geq \mathrm{H}(S_0)$ for $i = 1, \ldots, n$

Shannon entropy: *perfect secret sharing* (Π-scheme)

- **Dealer** samples a secret key S_0 and shares S_A, S_B, S_C, S_D for 4 participants
- correctness: the pairs $\{S_A, S_B\}$, $\{S_B, S_C\}$, $\{S_C, S_D\}$ allow to get the secret
  i.e., $\mathrm{H}(S_0 \mid S_A S_B) = 0$, $\mathrm{H}(S_0 \mid S_B S_C) = 0$, $\mathrm{H}(S_0 \mid S_C S_D) = 0$
- security: other pairs get no information on the secret
  i.e., $\mathrm{H}(S_0 \mid S_A S_C) = \mathrm{H}(S_0)$, $\mathrm{H}(S_0 \mid S_A S_D) = \mathrm{H}(S_0)$, $\mathrm{H}(S_0 \mid S_B S_D) = \mathrm{H}(S_0)$

Shannon entropy: *perfect secret sharing* (Π-scheme)

- **Dealer** samples a secret key S_0 and shares S_A, S_B, S_C, S_D for 4 participants
- correctness: the pairs $\{S_A, S_B\}$, $\{S_B, S_C\}$, $\{S_C, S_D\}$ allow to get the secret
  i.e., $\mathrm{H}(S_0 \mid S_A S_B) = 0$, $\mathrm{H}(S_0 \mid S_B S_C) = 0$, $\mathrm{H}(S_0 \mid S_C S_D) = 0$
- security: other pairs get no information on the secret
  i.e., $\mathrm{H}(S_0 \mid S_A S_C) = \mathrm{H}(S_0)$, $\mathrm{H}(S_0 \mid S_A S_D) = \mathrm{H}(S_0)$, $\mathrm{H}(S_0 \mid S_B S_D) = \mathrm{H}(S_0)$

Exercise: construct a correct and secure scheme of secret sharing such that
$\max\{\mathrm{H}(S_A), \mathrm{H}(S_B), \mathrm{H}(S_C), \mathrm{H}(S_D)\} \leq \frac{3}{2}\mathrm{H}(S_0)$

Shannon entropy: *perfect secret sharing* (Π-scheme)

- **Dealer** samples a secret key S_0 and shares S_A, S_B, S_C, S_D for 4 participants
- correctness: the pairs $\{S_A, S_B\}$, $\{S_B, S_C\}$, $\{S_C, S_D\}$ allow to get the secret
  i.e., $\mathrm{H}(S_0 \mid S_A S_B) = 0$, $\mathrm{H}(S_0 \mid S_B S_C) = 0$, $\mathrm{H}(S_0 \mid S_C S_D) = 0$
- security: other pairs get no information on the secret
  i.e., $\mathrm{H}(S_0 \mid S_A S_C) = \mathrm{H}(S_0)$, $\mathrm{H}(S_0 \mid S_A S_D) = \mathrm{H}(S_0)$, $\mathrm{H}(S_0 \mid S_B S_D) = \mathrm{H}(S_0)$

Exercise: construct a correct and secure scheme of secret sharing such that
$\max\{\mathrm{H}(S_A), \mathrm{H}(S_B), \mathrm{H}(S_C), \mathrm{H}(S_D)\} \le \frac{3}{2}\mathrm{H}(S_0)$

Exercise: prove that in *every* correct and secure scheme of secret sharing
$\max\{\mathrm{H}(S_A), \mathrm{H}(S_B), \mathrm{H}(S_C), \mathrm{H}(S_D)\} \ge \frac{3}{2}\mathrm{H}(S_0)$

Shannon entropy: *perfect secret sharing* (Π-scheme)

- **Dealer** samples a secret key S_0 and shares S_A, S_B, S_C, S_D for 4 participants
- correctness: the pairs $\{S_A, S_B\}$, $\{S_B, S_C\}$, $\{S_C, S_D\}$ allow to get the secret
  i.e., $\mathrm{H}(S_0 \mid S_A S_B) = 0$, $\mathrm{H}(S_0 \mid S_B S_C) = 0$, $\mathrm{H}(S_0 \mid S_C S_D) = 0$
- security: other pairs get no information on the secret
  i.e., $\mathrm{H}(S_0 \mid S_A S_C) = \mathrm{H}(S_0)$, $\mathrm{H}(S_0 \mid S_A S_D) = \mathrm{H}(S_0)$, $\mathrm{H}(S_0 \mid S_B S_D) = \mathrm{H}(S_0)$

Exercise: construct a correct and secure scheme of secret sharing such that
$\max\{\mathrm{H}(S_A), \mathrm{H}(S_B), \mathrm{H}(S_C), \mathrm{H}(S_D)\} \le \frac{3}{2}\mathrm{H}(S_0)$

Exercise: prove that in *every* correct and secure scheme of secret sharing
$\max\{\mathrm{H}(S_A), \mathrm{H}(S_B), \mathrm{H}(S_C), \mathrm{H}(S_D)\} \ge \frac{3}{2}\mathrm{H}(S_0)$

Theorem (Laszlo Csirmaz): one can formulate the conditions of correctness and security for n participants so that for every secret sharing scheme

$$\max_{1 \le i \le n} \mathrm{H}(S_i) \ge \Omega(n/\log n) \cdot \mathrm{H}(S_0)$$

Shannon entropy: *perfect secret sharing* (Π-scheme)

- **Dealer** samples a secret key S_0 and shares S_A, S_B, S_C, S_D for 4 participants
- correctness: the pairs $\{S_A, S_B\}$, $\{S_B, S_C\}$, $\{S_C, S_D\}$ allow to get the secret
  i.e., $\mathrm{H}(S_0 \mid S_A S_B) = 0$, $\mathrm{H}(S_0 \mid S_B S_C) = 0$, $\mathrm{H}(S_0 \mid S_C S_D) = 0$
- security: other pairs get no information on the secret
  i.e., $\mathrm{H}(S_0 \mid S_A S_C) = \mathrm{H}(S_0)$, $\mathrm{H}(S_0 \mid S_A S_D) = \mathrm{H}(S_0)$, $\mathrm{H}(S_0 \mid S_B S_D) = \mathrm{H}(S_0)$

Exercise: construct a correct and secure scheme of secret sharing such that
$\max\{\mathrm{H}(S_A), \mathrm{H}(S_B), \mathrm{H}(S_C), \mathrm{H}(S_D)\} \le \frac{3}{2}\mathrm{H}(S_0)$

Exercise: prove that in *every* correct and secure scheme of secret sharing
$\max\{\mathrm{H}(S_A), \mathrm{H}(S_B), \mathrm{H}(S_C), \mathrm{H}(S_D)\} \ge \frac{3}{2}\mathrm{H}(S_0)$

Theorem (Laszlo Csirmaz): one can formulate the conditions of correctness and security for n participants so that for every secret sharing scheme

$$\max_{1 \le i \le n} \mathrm{H}(S_i) \ge \Omega(n/\log n) \cdot \mathrm{H}(S_0)$$

Much simpler fact: for all conditions of correctness and security there exists a secret sharing scheme such that $\mathrm{H}(S_i) \le 2^{O(n)} \cdot \mathrm{H}(S_0)$ for all i