Title: A conceptual viewpoint on information decomposition

Date: Nov 08, 2016 03:30 PM

URL: http://pirsa.org/16110030

Abstract: <p>Can we decompose the information of a composite system into terms arising from its parts and their interactions?
For a bipartite system (X,Y), the joint entropy can be written as an algebraic sum of three terms: the entropy of X alone, the entropy of Y alone, and the mutual information of X and Y, which comes with an opposite sign. This suggests a set-theoretical analogy: mutual information is a sort of "intersection", and joint entropy is a sort of "union".
The same picture cannot be generalized to three or more parts in a straightforward way, and the problem is still considered open. Is there a deep reason for why the set-theoretical analogy fails?
Category theory can give an alternative, conceptual point of view on the problem. As Shannon already noted, information appears to be related to symmetry. This suggests a natural lattice structure for information, which is compatible with a set-theoretical picture only for bipartite systems.
The categorical approach favors objects with a structure in place of just numbers to describe information quantities. We hope that this can clarify the mathematical structure underlying information theory, and leave it open to wider generalizations.</p>

A Conceptual Viewpoint on Information Decomposition

Paolo Perrone

Max Planck Institute
Leipzig, Germany

Perimeter Institute
for Theoretical Physics
8th November 2016

Information Decomposition

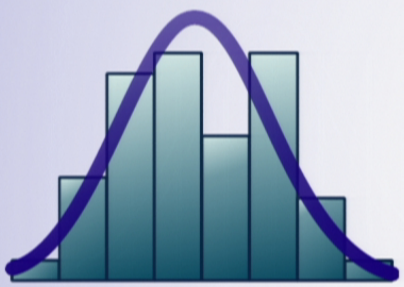
Question:

How much do parts of a system contribute to the total information?

Information Decomposition

Question:

How much do parts of a system contribute to the total information?



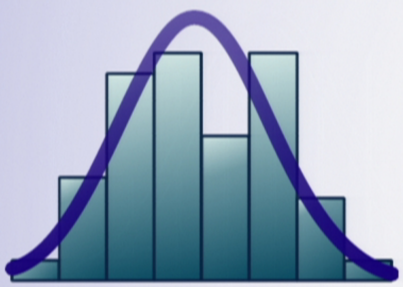
Statistics

Nonlinear higher
order correlations

Information Decomposition

Question:

How much do parts of a system contribute to the total information?



Statistics

Nonlinear higher order correlations

Game Theory

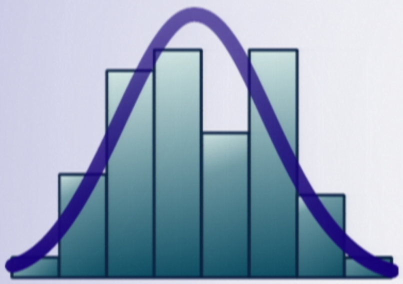
Information in games

Blackwell's theorem

Information Decomposition

Question:

How much do parts of a system contribute to the total information?



Statistics

Nonlinear higher order correlations

Game Theory

Information in games Blackwell's theorem

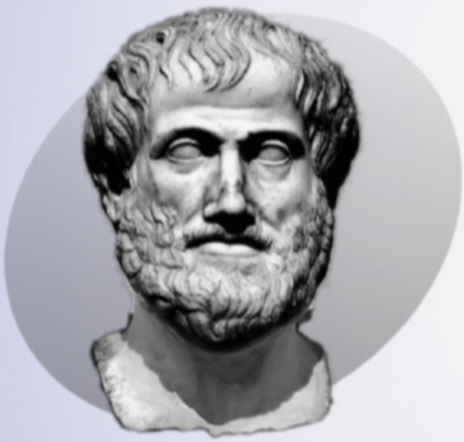
Biological Networks

Quantify complexity of neural networks.

Information Decomposition

Question:

How much do parts of a system contribute to the total information?



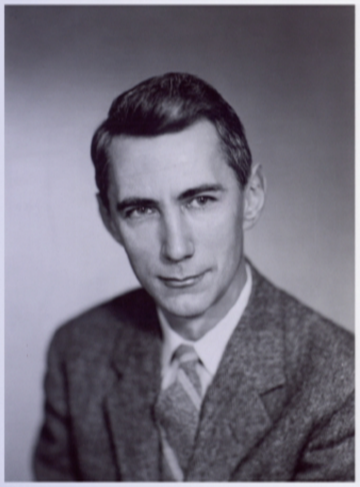
"The totality is not [...] a mere heap, but the whole is something besides the parts."

– Aristotle, Metaphysics VIII

Information Decomposition

Question:

How much do parts of a system contribute to the total information?



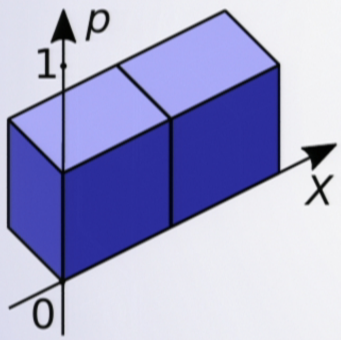
- Shannon entropy:

$$H(X) := -\sum_{x \in X} p(x) \log p(x)$$

Information Decomposition

Question:

How much do parts of a system contribute to the total information?



- Shannon entropy:

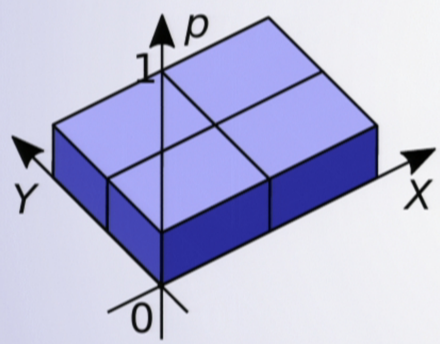
$$H(X) := -\sum_{x \in X} p(x) \log p(x)$$

In the example, $H(X) = 1$ bit.

Information Decomposition

Question:

How much do parts of a system contribute to the total information?



- Joint entropy:

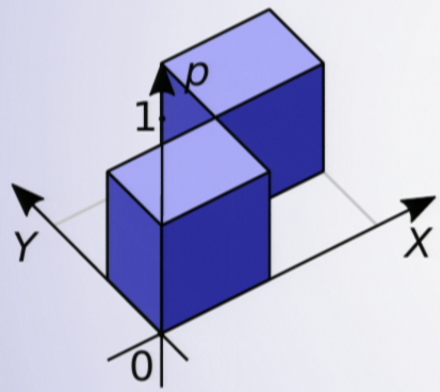
$$H(X, Y) := -\sum_{x,y} p(x, y) \log p(x, y)$$

In the example, $H(X, Y) = 2$ bits.

Information Decomposition

Question:

How much do parts of a system contribute to the total information?



- Conditional entropy:

$$H(X|Y) := -\sum_{x,y} p(x,y) \log p(x|y)$$

In the example, $H(X|Y) = 0$ bit.

Information Decomposition

Idea:

"The whole is greater than the sum of its parts."

Information Decomposition

Idea:

"The whole is greater than the sum of its parts."

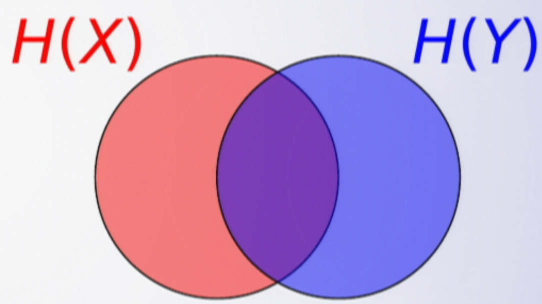
$$H(X, Y) \leq H(X) + H(Y)$$

Information Decomposition

Idea:

"The whole is greater than the sum of its parts."

$$H(X, Y) \leq H(X) + H(Y)$$



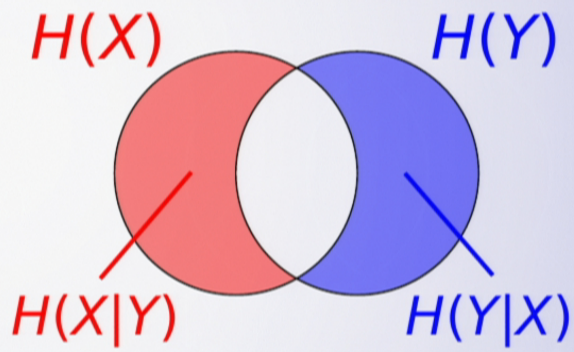
Information Decomposition

Idea:

"The whole is greater than the sum of its parts."

$$H(X,Y) \leq H(X) + H(Y)$$

$$H(X,Y) \geq H(X|Y) + H(Y|X)$$



Information Decomposition

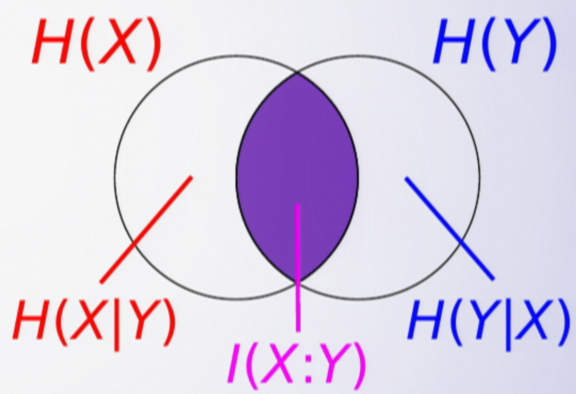
Idea:

"The whole is greater than the sum of its parts."

$$H(X, Y) \leq H(X) + H(Y)$$

$$H(X, Y) \geq H(X|Y) + H(Y|X)$$

$$\boxed{I(X : Y) \geq 0}$$



Information Decomposition

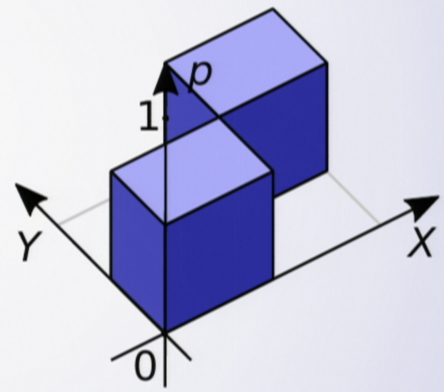
Idea:

"The whole is greater than the sum of its parts."

$$H(X,Y) \leq H(X) + H(Y)$$

$$H(X,Y) \geq H(X|Y) + H(Y|X)$$

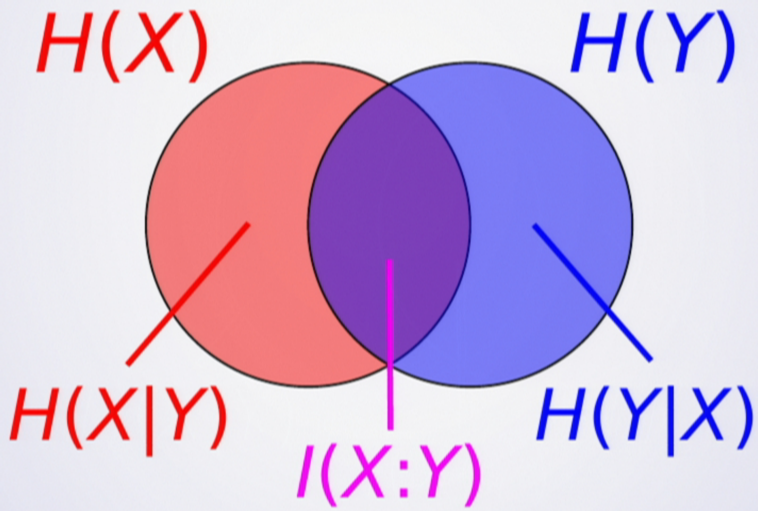
$$I(X:Y) \geq 0$$



Information Decomposition

Set-theoretical picture:

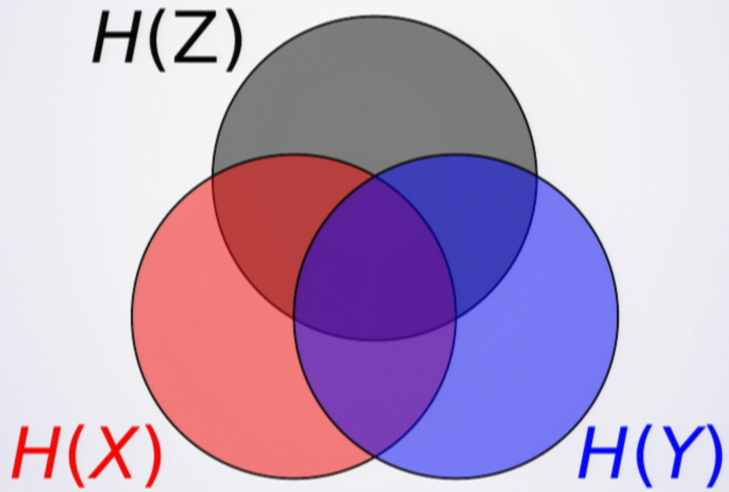
"Statistical interactions are like intersections."



Information Decomposition

Set-theoretical picture:

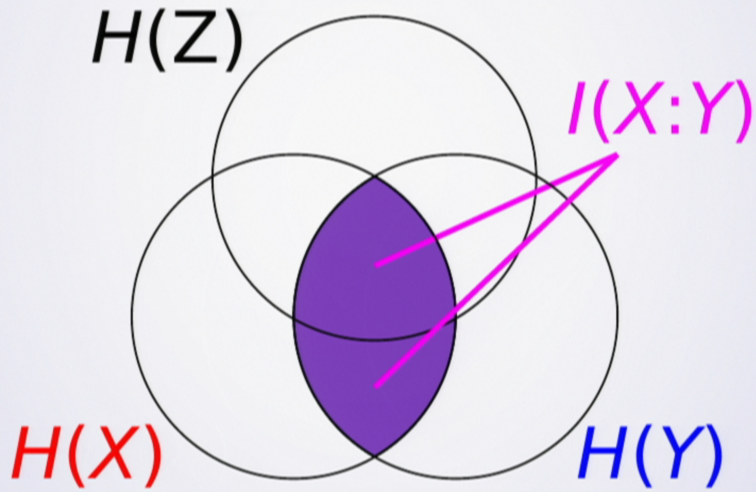
"Statistical interactions are like intersections."



Information Decomposition

Set-theoretical picture:

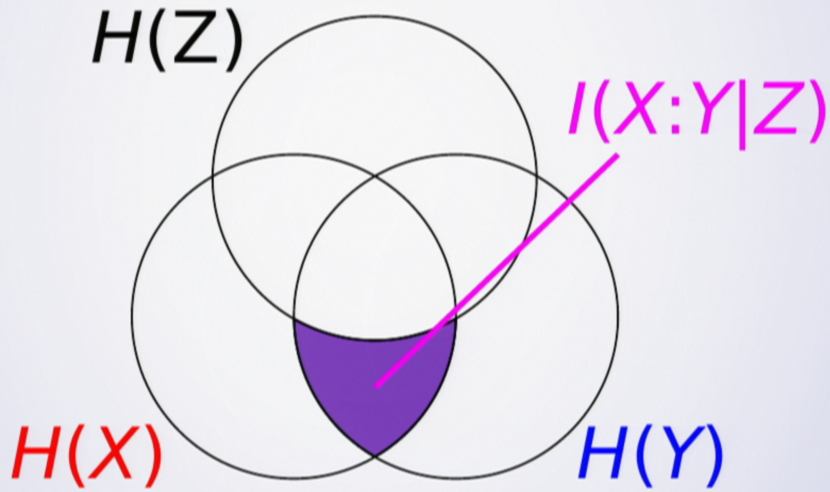
"Statistical interactions are like intersections."



Information Decomposition

Set-theoretical picture:

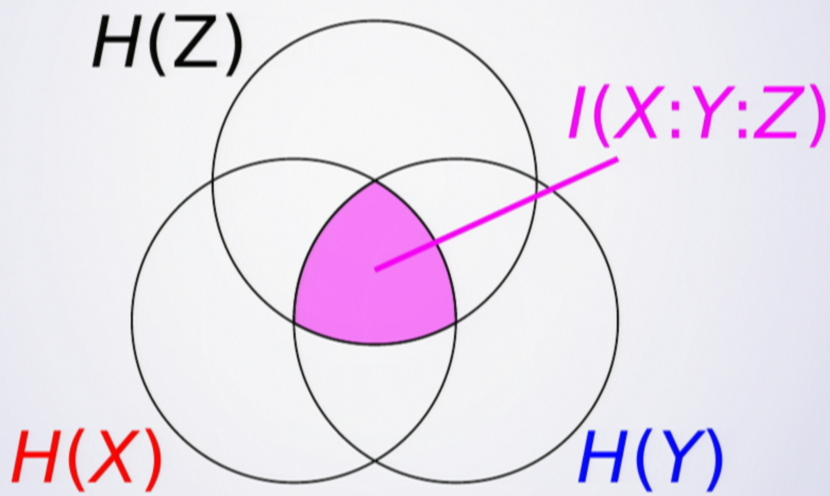
"Statistical interactions are like intersections."



Information Decomposition

Set-theoretical picture:

"Statistical interactions are like intersections."



Information Decomposition

Problem:

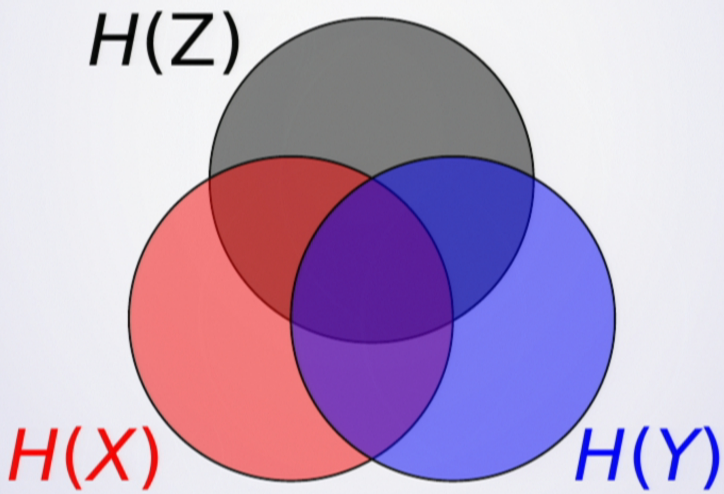
$I(X : Y : Z)$ is not in general a positive quantity.

Information Decomposition

Problem:

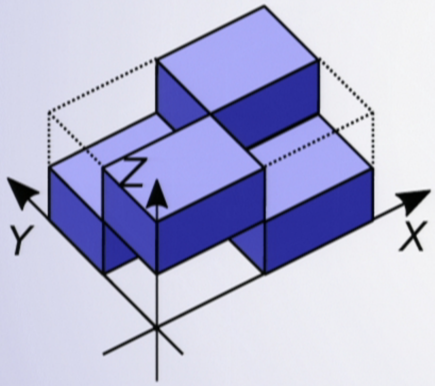
$I(X : Y : Z)$ is not in general a positive quantity.

Equivalently, it is *not* true in general that $I(X : Y \mid Z) \leq I(X : Y)$.



Information Decomposition

Problem:
$I(X : Y : Z)$ is not in general a positive quantity.
Equivalently, it is *not* true in general that $I(X : Y \mid Z) \leq I(X : Y)$.



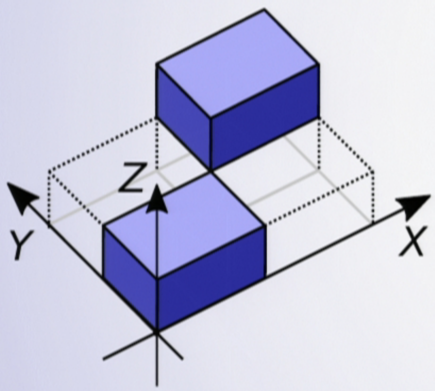
- For example, if $X = Y \oplus Z$:

$$I(X : Y) = 0$$

$$I(X : Y \mid Z) = 1$$

Information Decomposition

Problem:
$I(X : Y : Z)$ is not in general a positive quantity.
Equivalently, it is *not* true in general that $I(X : Y \mid Z) \leq I(X : Y)$.



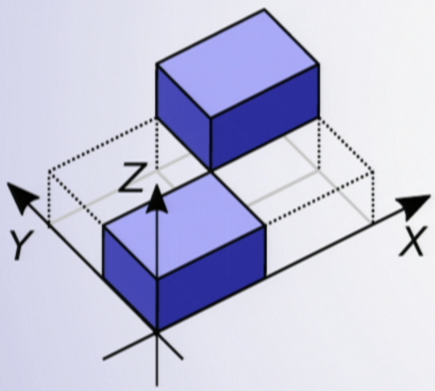
- For example, if $X = Y = Z$:

$$I(X : Y) = 1$$

$$I(X : Y \mid Z) = 0$$

Information Decomposition

Problem:
$I(X : Y : Z)$ is not in general a positive quantity.
Equivalently, it is *not* true in general that $I(X : Y \mid Z) \leq I(X : Y)$.



- For example, if $X = Y = Z$:

$$I(X : Y) = 1$$

$$I(X : Y \mid Z) = 0$$

Schneidmann, Bialek, Berry [1] and Williams, Beer [2].

Information Decomposition

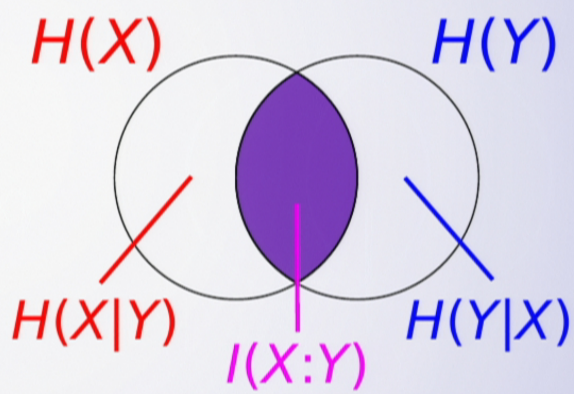
Idea:

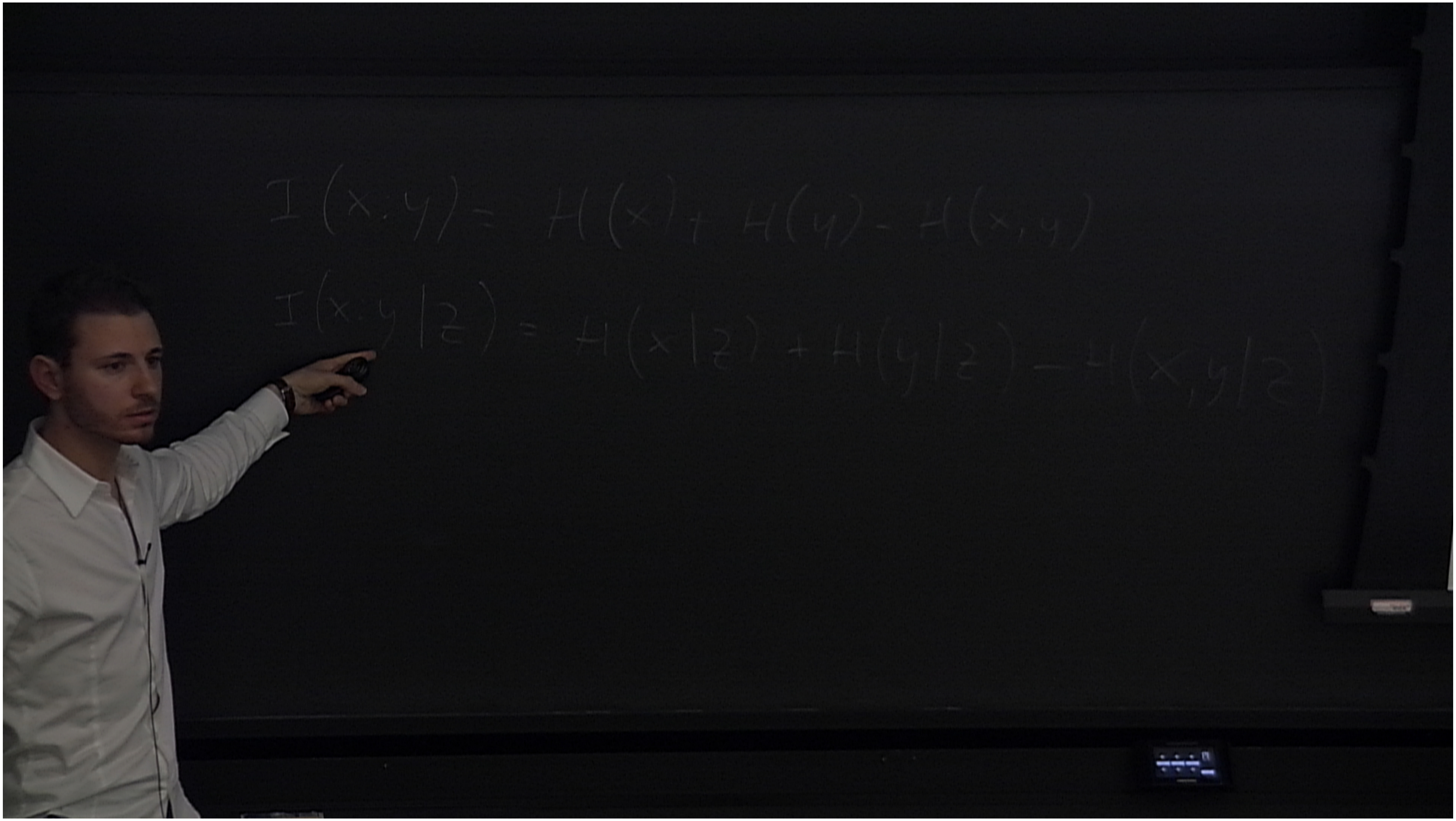
"The whole is greater than the sum of its parts."

$$H(X, Y) \leq H(X) + H(Y)$$

$$H(X, Y) \geq H(X|Y) + H(Y|X)$$

$$\boxed{I(X : Y) \geq 0}$$




$I(x:y) = H(x) + H(y) - H(x,y)$
$I(x:y|z) = H(x|z) + H(y|z) - H(x,y|z)$

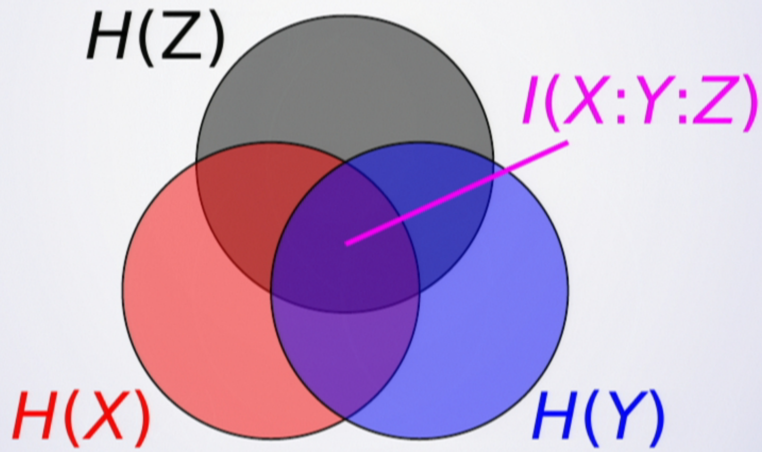
$$I(x:y) = H(x) + H(y) - H(x,y)$$

$$I(x:y|z) = H(x|z) + H(y|z) - H(x,y|z)$$

$$I(x:y:z) = I(x:y) - I(x:y|z)$$

Information Decomposition

Problem:
$I(X : Y : Z)$ is not in general a positive quantity.
Equivalently, it is *not* true in general that $I(X : Y \mid Z) \leq I(X : Y)$.

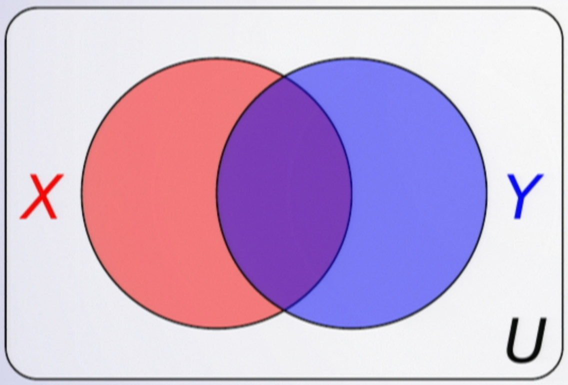


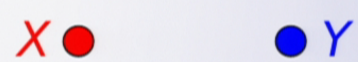
Lattices of Sets

Given a set U, we take a finite set of subsets $L \subseteq P(U)$ closed under union and intersection.

Lattices of Sets

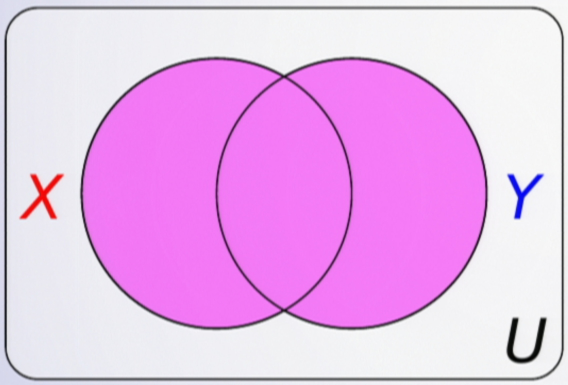
Given a set U, we take a finite set of subsets $L \subseteq P(U)$ closed under union and intersection.

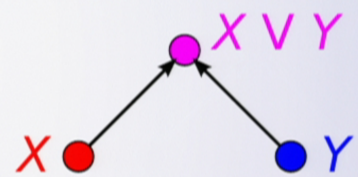




Lattices of Sets

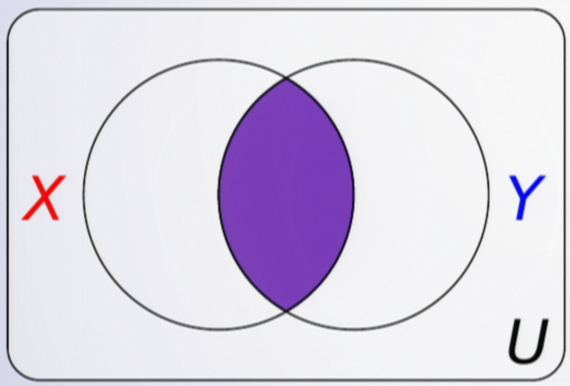
Given a set U, we take a finite set of subsets $L \subseteq P(U)$ closed under union and intersection.

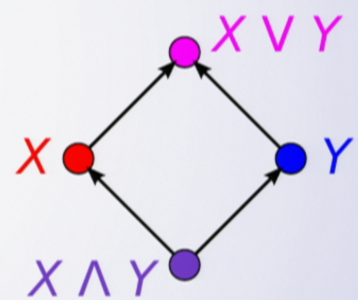




Lattices of Sets

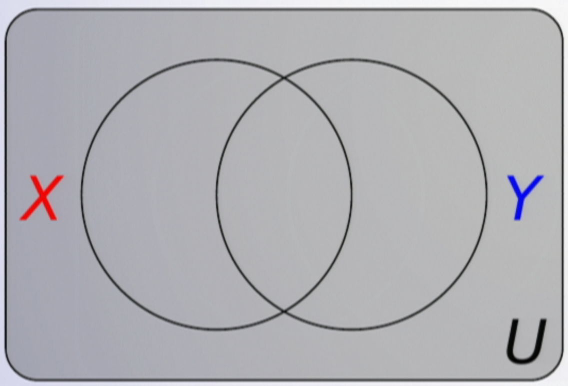
Given a set U, we take a finite set of subsets $L \subseteq P(U)$ closed under union and intersection.

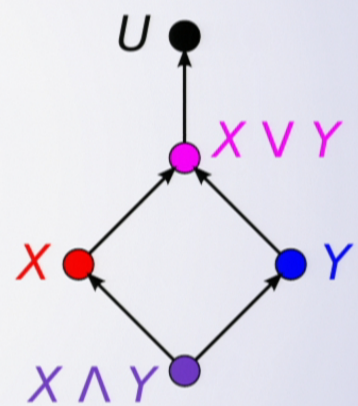




Lattices of Sets

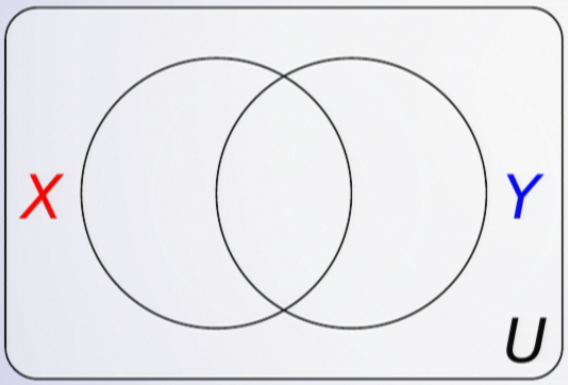
Given a set U, we take a finite set of subsets $L \subseteq P(U)$ closed under union and intersection.

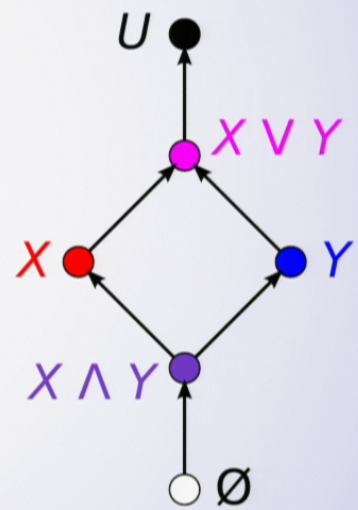




Lattices of Sets

Given a set U, we take a finite set of subsets $L \subseteq P(U)$ closed under union and intersection.





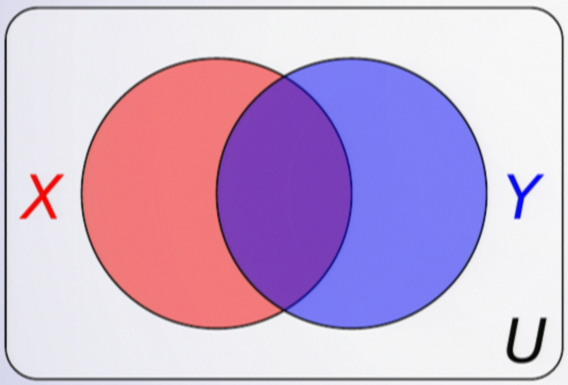
Lattices of Sets

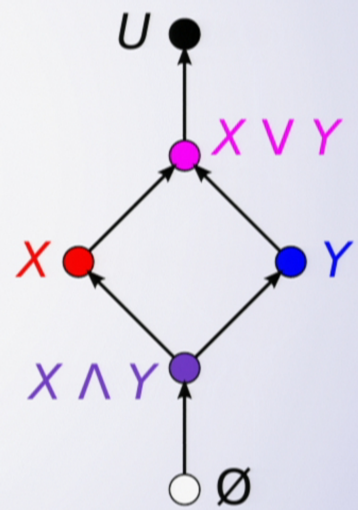
Given a set U, we take a finite set of subsets $L \subseteq P(U)$ closed under union and intersection.

Definition:
Let (U, Σ, μ) be a measure space. A *lattice of sets* in (U, Σ, μ) is a finite subset of Σ closed under union and intersection.

Lattices of Sets

Given a set U, we take a finite set of subsets $L \subseteq P(U)$ closed under union and intersection.





Lattices of Sets

Given a set U, we take a finite set of subsets $L \subseteq P(U)$ closed under union and intersection.

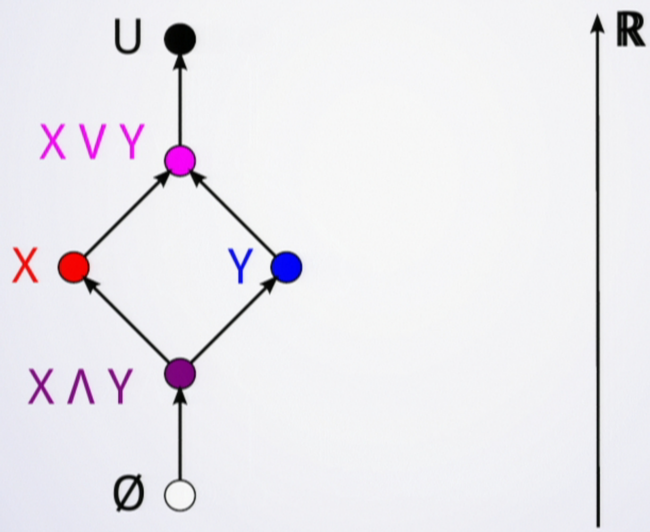
Definition:

Let (U, Σ, μ) be a measure space. A *lattice of sets* in (U, Σ, μ) is a finite subset of Σ closed under union and intersection.

This way, the measure μ restricted to L is a monotone function.

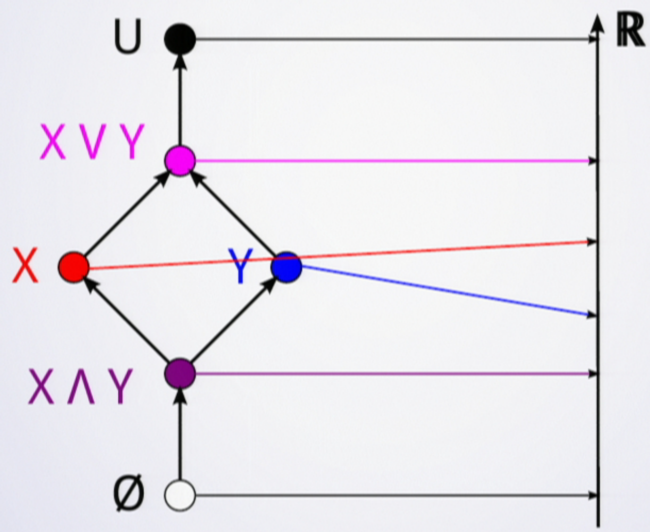
Lattices of Sets

Given a set U, we take a finite set of subsets $L \subseteq P(U)$ closed under union and intersection.



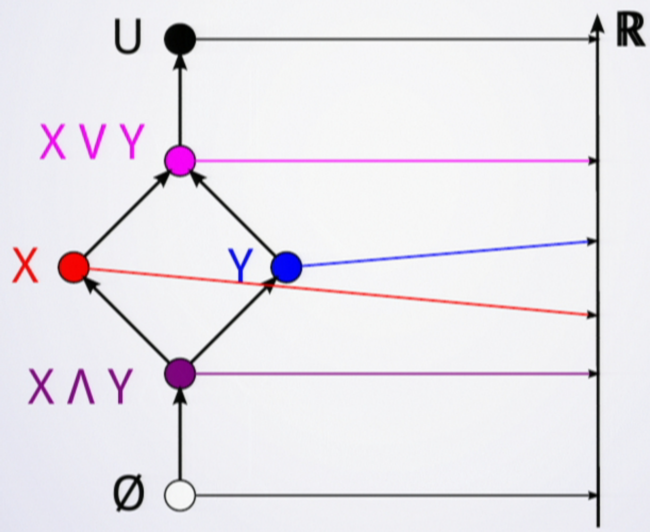
Lattices of Sets

Given a set U, we take a finite set of subsets $L \subseteq P(U)$ closed under union and intersection.



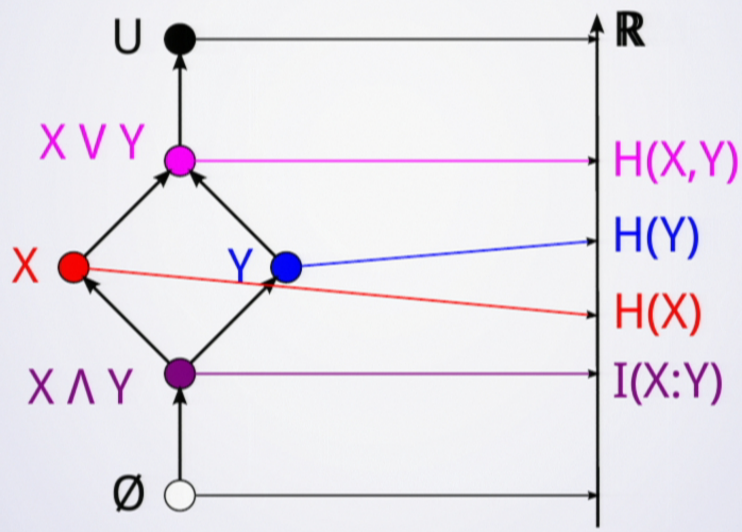
Lattices of Sets

Given a set U, we take a finite set of subsets $L \subseteq P(U)$ closed under union and intersection.



Lattices of Sets

Given a set U, we take a finite set of subsets $L \subseteq P(U)$ closed under union and intersection.



Information Lattice

Question:

Does information naturally form a lattice?

Information Lattice

Question:

Does information naturally form a lattice?

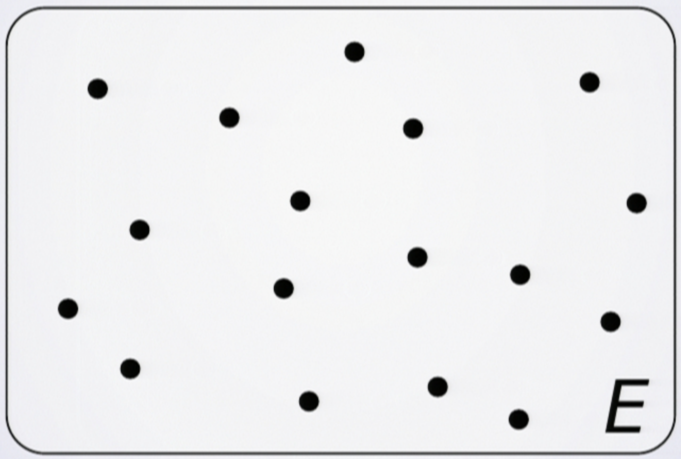
Idea:

An observable Y is *more informative* than an observable X if it allows us to distinguish between more possible states.

Information Lattice

Question:

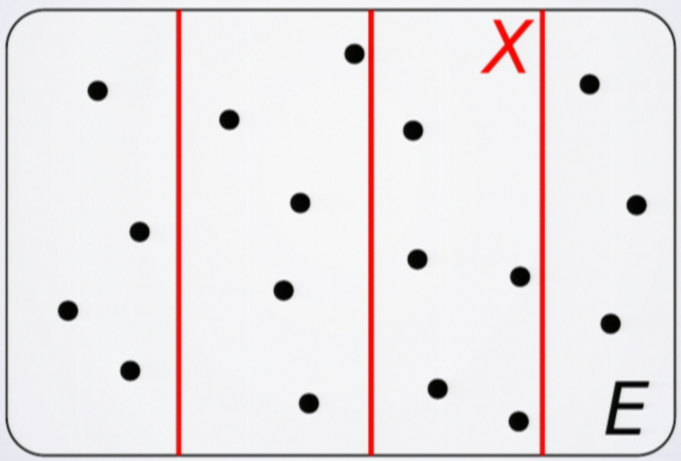
Does information naturally form a lattice?



Information Lattice

Question:

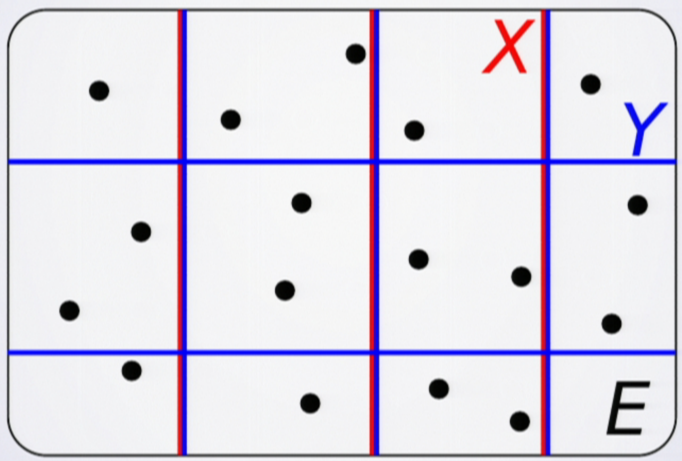
Does information naturally form a lattice?



Information Lattice

Question:

Does information naturally form a lattice?



Information Lattice

Question:

Does information naturally form a lattice?

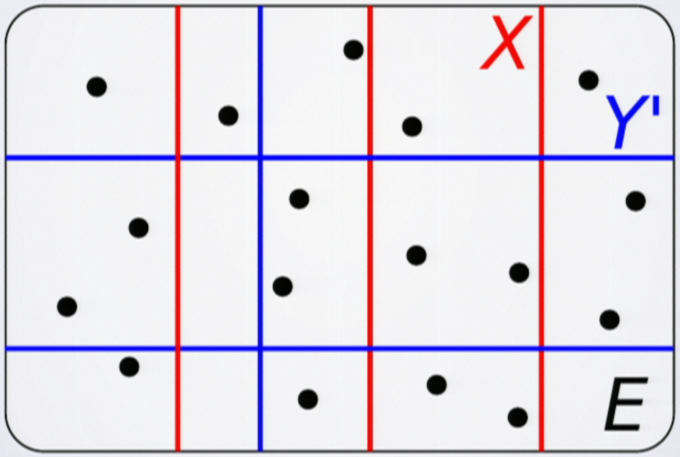
Idea (refined):

An observable Y is *more informative* than an observable X if it induces a finer partition on the outcome space.

Information Lattice

Question:

Does information naturally form a lattice?



Information Lattice

Question:

Does information naturally form a lattice?

Idea (refined):

An observable Y is *more informative* than an observable X if it induces a finer partition on the outcome space.

- The relation "more informative" forms only a *partial* order!

Information Lattice

Question:

Does information naturally form a lattice?

Idea (refined):

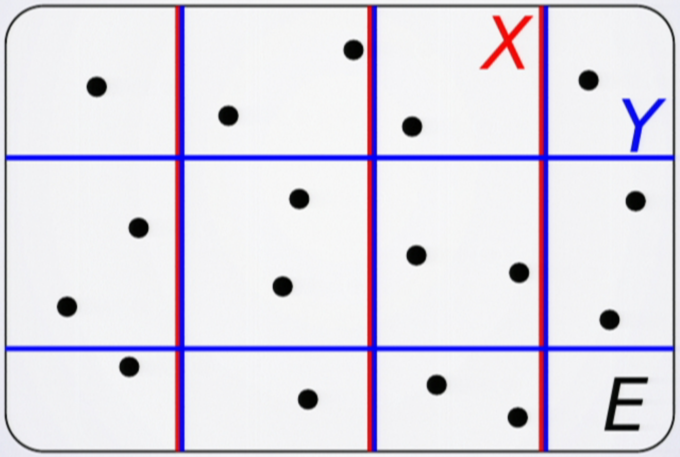
An observable Y is *more informative* than an observable X if it induces a finer partition on the outcome space.

- The relation "more informative" forms only a *partial* order!
- This order relation does not depend on the sample distribution.

Information Lattice

Question:

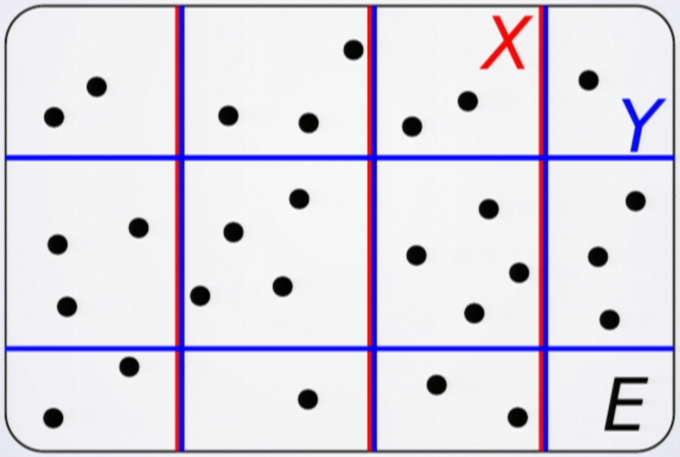
Does information naturally form a lattice?



Information Lattice

Question:

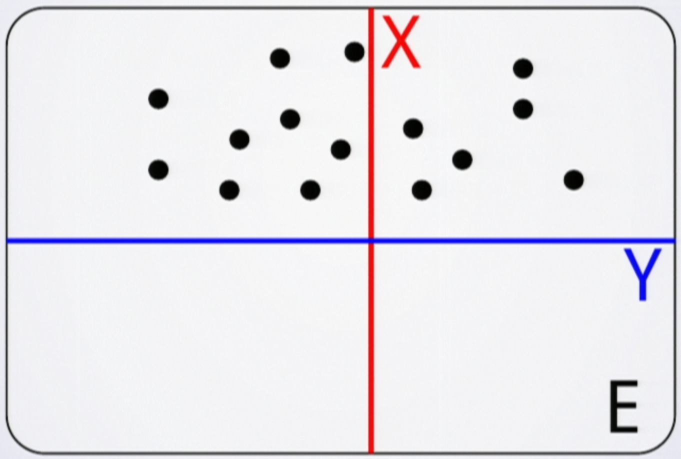
Does information naturally form a lattice?



Information Lattice

Question:

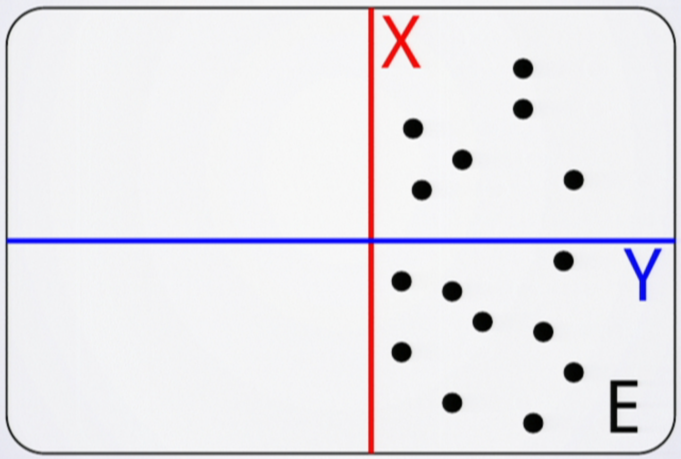
Does information naturally form a lattice?



Information Lattice

Question:

Does information naturally form a lattice?



Information Lattice

Question:

Does information naturally form a lattice?

Idea (refined):

An observable Y is *more informative* than an observable X if it induces a finer partition on the outcome space.

- The relation "more informative" forms only a *partial* order!
- This order relation does not depend on the sample distribution.
- We will write in this case $Y \geq X$.

Information Lattice

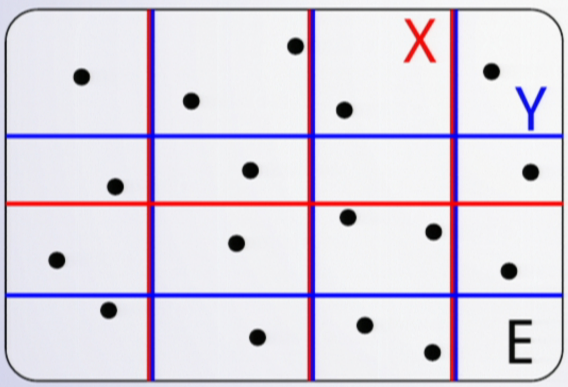
Definition (Li, Chong [3]):

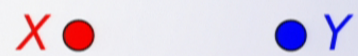
The *information lattice* of a set of observables $X, Y, \ldots$ is the sublattice of the partition lattice of E generated by the partitions induced by $X, Y, \ldots$.

Information Lattice

Definition (Li, Chong [3]):

The *information lattice* of a set of observables $X, Y, \ldots$ is the sublattice of the partition lattice of E generated by the partitions induced by $X, Y, \ldots$.

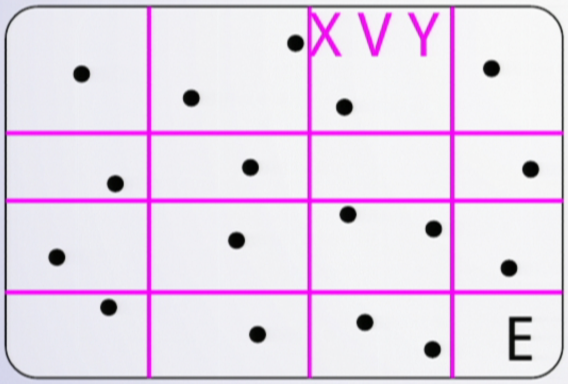


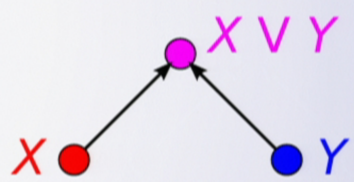


Information Lattice

Definition (Li, Chong [3]):

The *information lattice* of a set of observables $X, Y, \ldots$ is the sublattice of the partition lattice of E generated by the partitions induced by $X, Y, \ldots$.

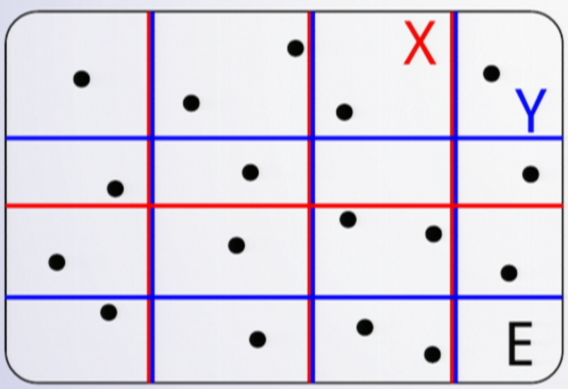


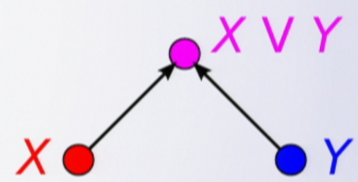


Information Lattice

Definition (Li, Chong [3]):

The *information lattice* of a set of observables $X, Y, \ldots$ is the sublattice of the partition lattice of E generated by the partitions induced by $X, Y, \ldots$.

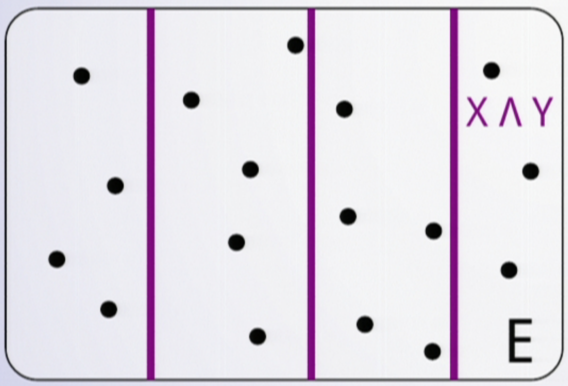


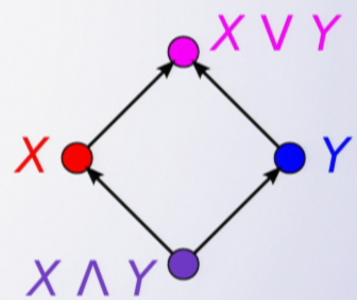


Information Lattice

Definition (Li, Chong [3]):

The *information lattice* of a set of observables $X, Y, \ldots$ is the sublattice of the partition lattice of E generated by the partitions induced by $X, Y, \ldots$.

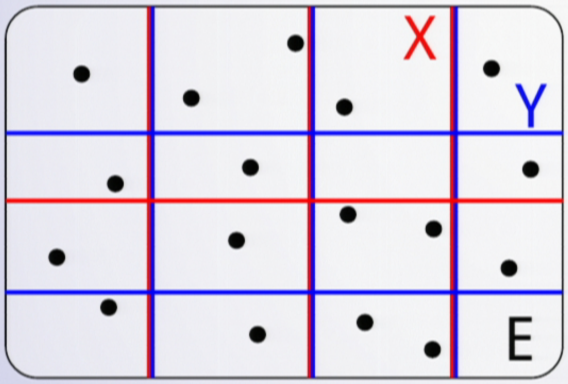


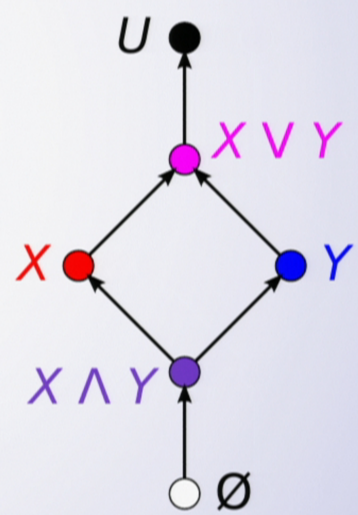


Information Lattice

Definition (Li, Chong [3]):

The *information lattice* of a set of observables $X, Y, \ldots$ is the sublattice of the partition lattice of E generated by the partitions induced by $X, Y, \ldots$.





Information Lattice

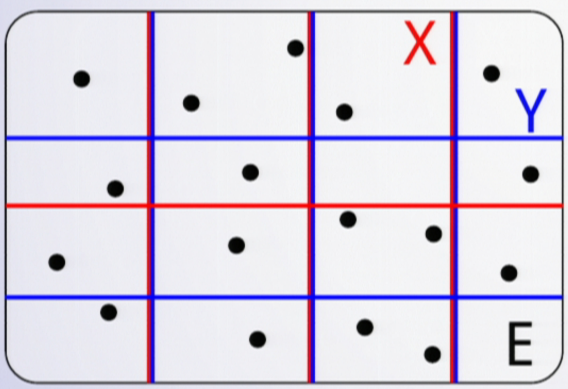
Question:

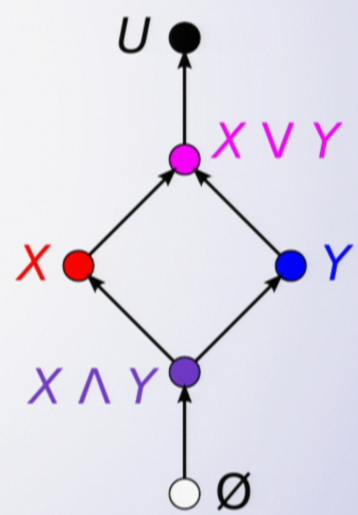
Does the information lattice have anything to do with entropy?

Information Lattice

Definition (Li, Chong [3]):

The *information lattice* of a set of observables $X, Y, \ldots$ is the sublattice of the partition lattice of E generated by the partitions induced by $X, Y, \ldots$.





Information Lattice

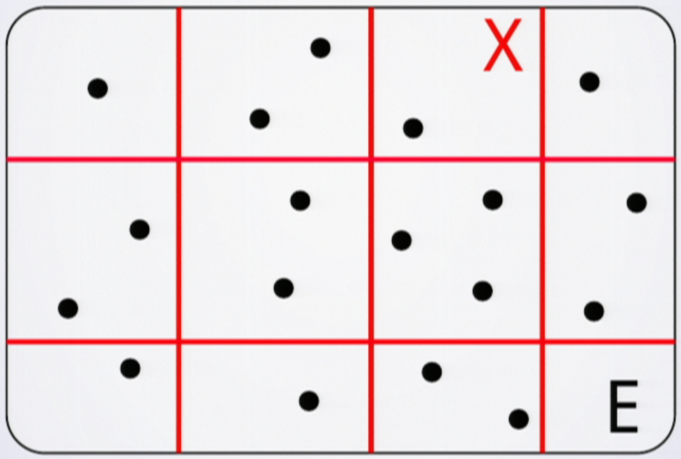
Question:

Does the information lattice have anything to do with entropy?

Information Lattice

Question:

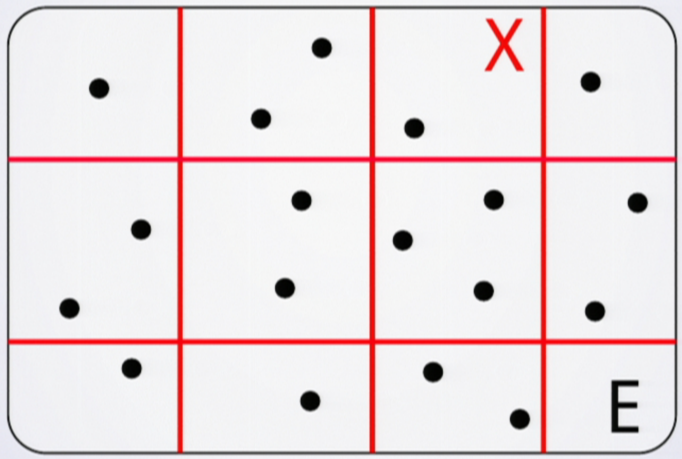
Does the information lattice have anything to do with entropy?



Information Lattice

Question:

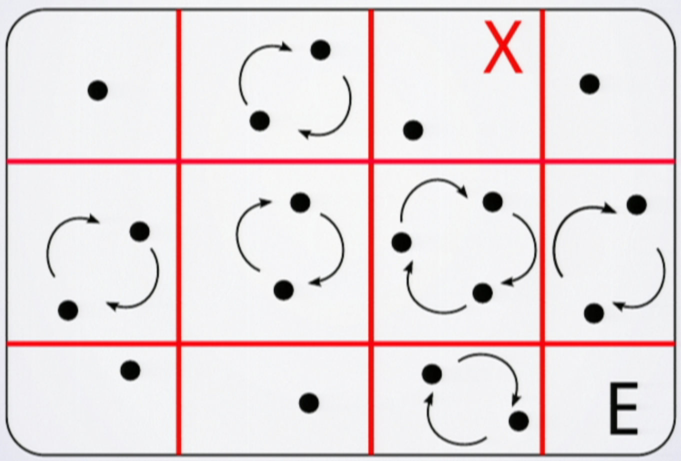
Does the information lattice have anything to do with entropy?



Information Lattice

Question:

Does the information lattice have anything to do with entropy?



Information Lattice

Question:

Does the information lattice have anything to do with entropy?

Definition:

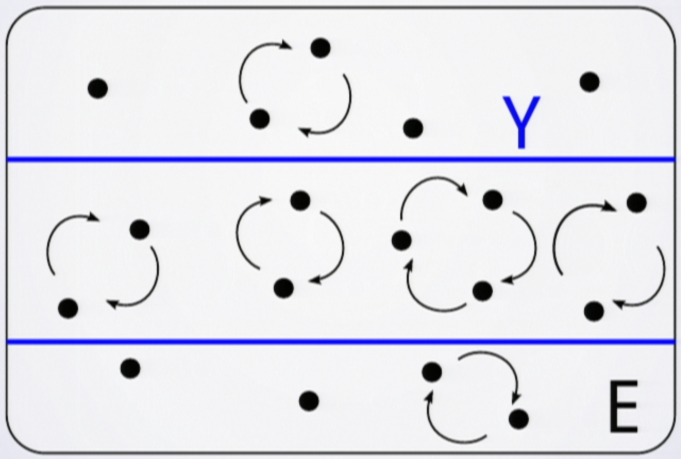
The permutation group $\mathrm{Aut}(X)$ of a partition $p : E \to X$ is given by:

$$\mathrm{Aut}(X) := \prod_{x \in X} \mathrm{Aut}\left(p^{-1}(x)\right)$$

Information Lattice

Question:

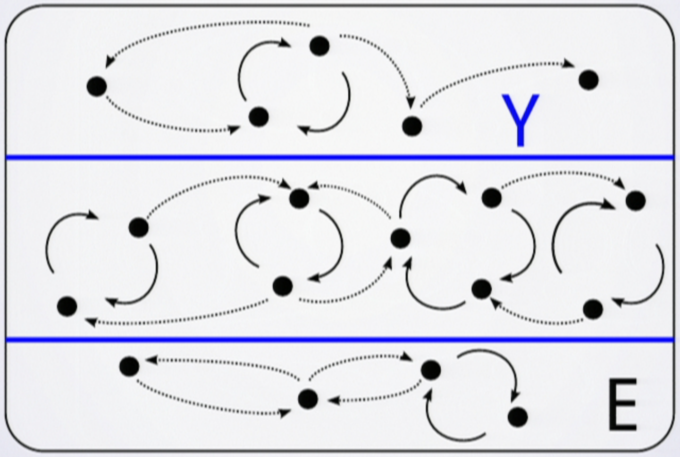
Does the information lattice have anything to do with entropy?



Information Lattice

Question:

Does the information lattice have anything to do with entropy?



Information Lattice

Question:

Does the information lattice have anything to do with entropy?

Definition:

The permutation group $\mathrm{Aut}(X)$ of a partition $p : E \to X$ is given by:

$$\mathrm{Aut}(X) := \prod_{x \in X} \mathrm{Aut}\left(p^{-1}(x)\right)$$

If $X \geq Y$, $\mathrm{Aut}(X)$ is a subgroup of $\mathrm{Aut}(Y)$.

Information Lattice

Question:

Does the information lattice have anything to do with entropy?

- If $X \geq Y$, $\mathrm{Aut}(X) \leq \mathrm{Aut}(Y)$.

Information Lattice

Question:

Does the information lattice have anything to do with entropy?

- If $X \geq Y$, $\mathrm{Aut}(X) \leq \mathrm{Aut}(Y)$.
- If $\emptyset$ is the trivial partition (no divisions), for each X we have $\mathrm{Aut}(X) \leq \mathrm{Aut}(\emptyset)$.

Information Lattice

Question:

Does the information lattice have anything to do with entropy?

- If $X \geq Y$, $\mathrm{Aut}(X) \leq \mathrm{Aut}(Y)$.
- If $\emptyset$ is the trivial partition (no divisions), for each X we have $\mathrm{Aut}(X) \leq \mathrm{Aut}(\emptyset)$.
- Subgroups of $\mathrm{Aut}(\emptyset)$ and inclusions form a lattice $S(\mathrm{Aut}(\emptyset))$.
- The information lattice is isomorphic to the *dual* of a sublattice of $S(\mathrm{Aut}(\emptyset))$.

Information Lattice

Question:

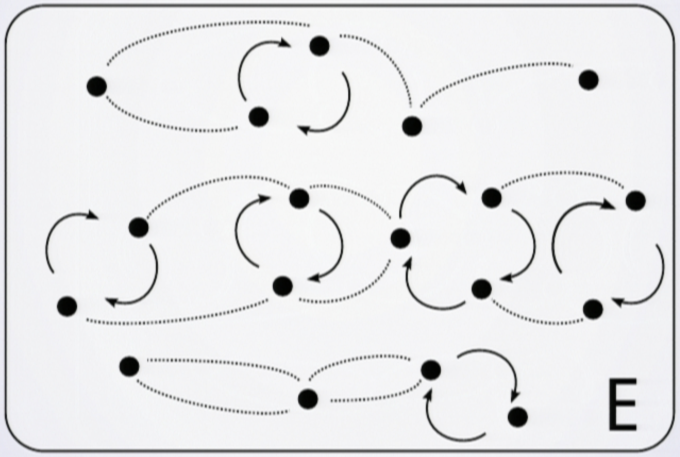
Does the information lattice have anything to do with entropy?

- If $X \geq Y$, $\mathrm{Aut}(X) \leq \mathrm{Aut}(Y)$.
- If $\emptyset$ is the trivial partition (no divisions), for each X we have $\mathrm{Aut}(X) \leq \mathrm{Aut}(\emptyset)$.
- Subgroups of $\mathrm{Aut}(\emptyset)$ and inclusions form a lattice $S(\mathrm{Aut}(\emptyset))$.
- The information lattice is isomorphic to the *dual* of a sublattice of $S(\mathrm{Aut}(\emptyset))$.
- We therefore take the lattice of *quotients*.

Information Lattice

Question:

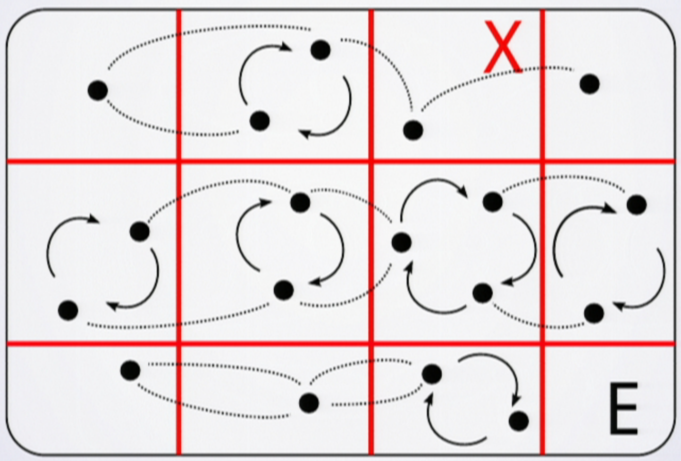
Does the information lattice have anything to do with entropy?



Information Lattice

Question:

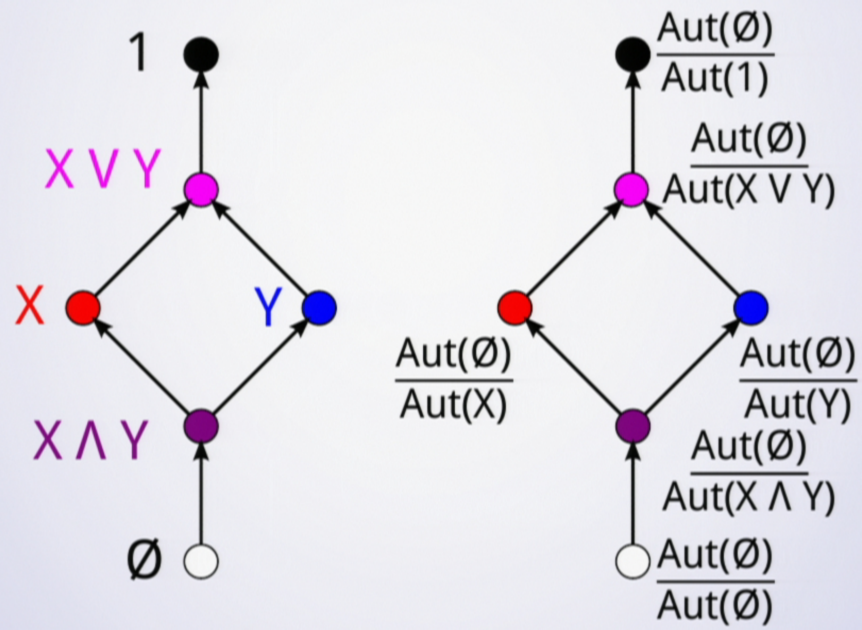
Does the information lattice have anything to do with entropy?



Information Lattice

Question:

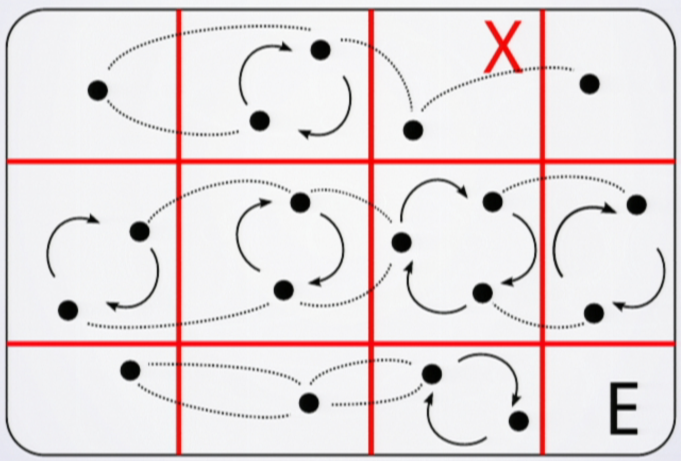
Does the information lattice have anything to do with entropy?



Information Lattice

Question:

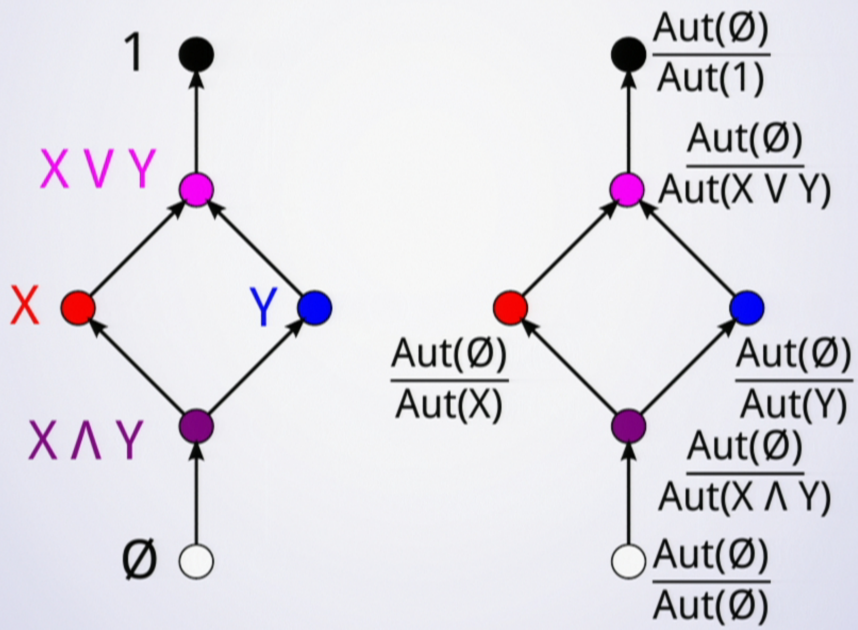
Does the information lattice have anything to do with entropy?



Information Lattice

Question:

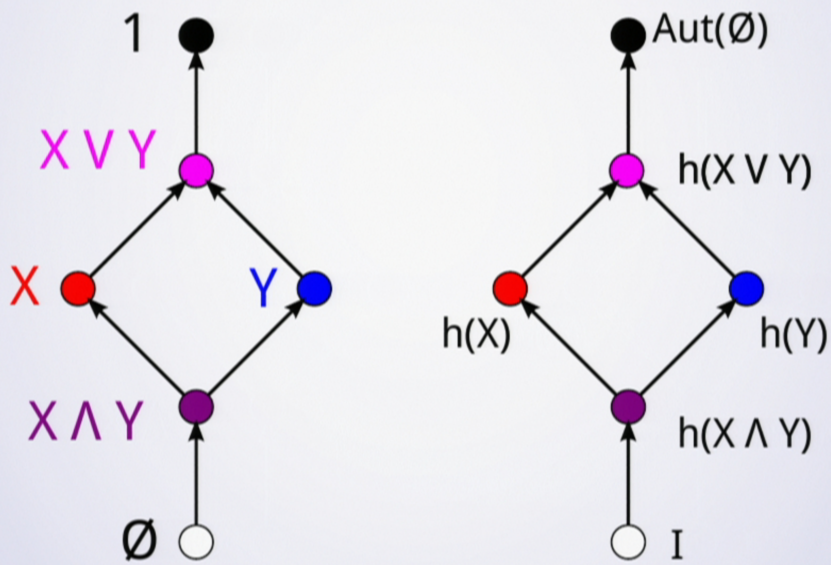
Does the information lattice have anything to do with entropy?



Information Lattice

Question:

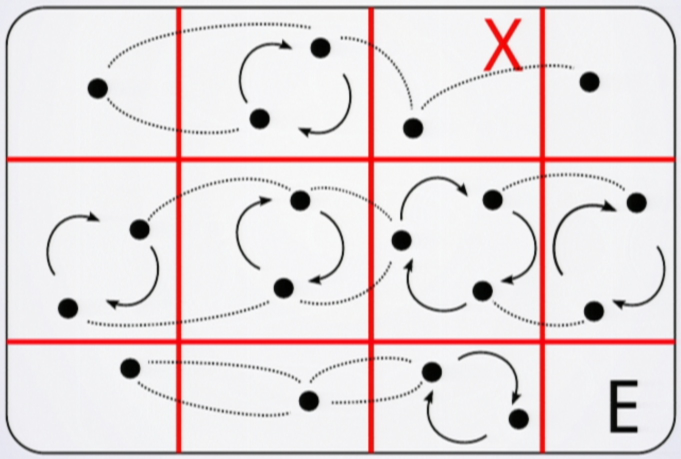
Does the information lattice have anything to do with entropy?

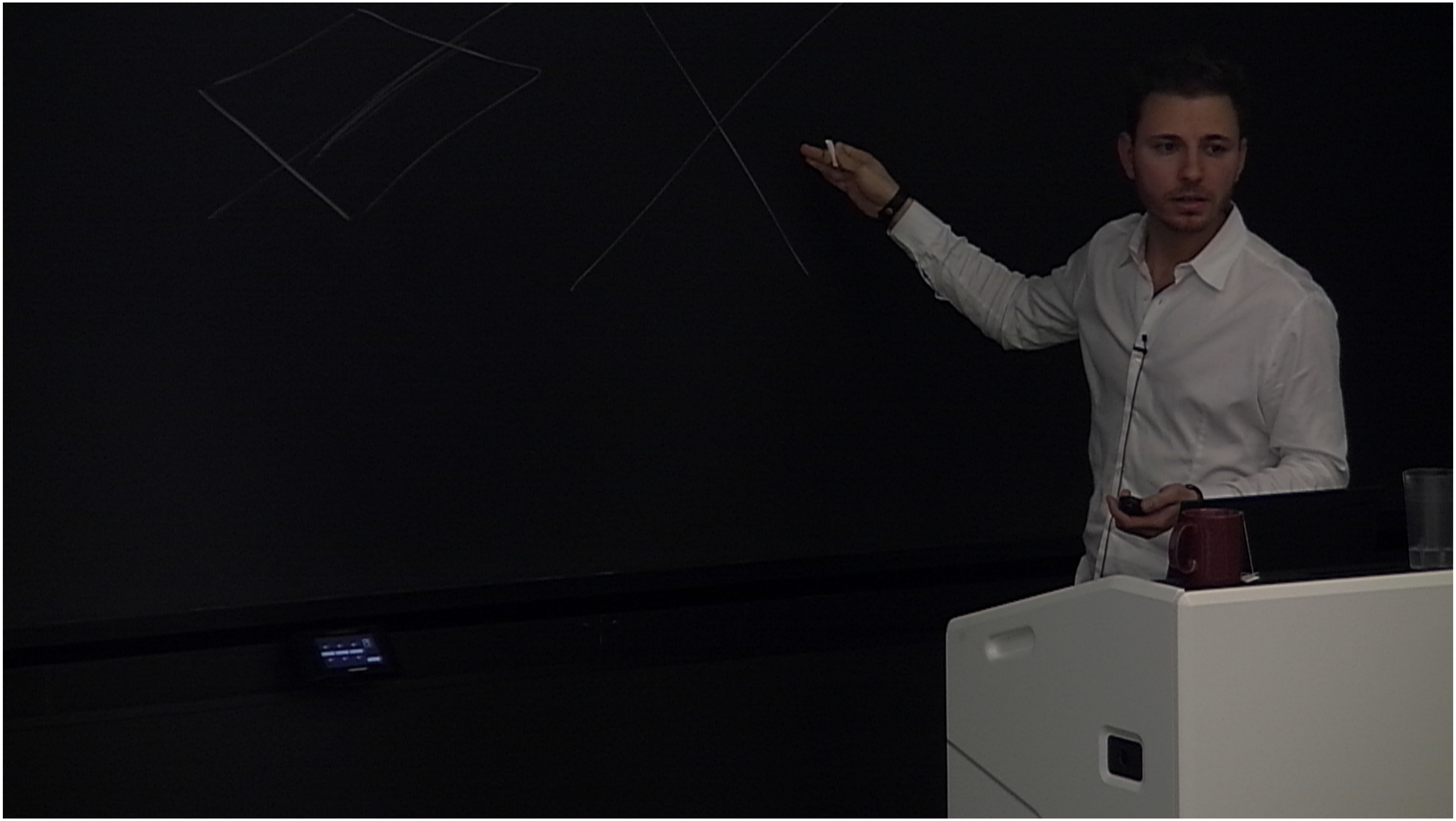


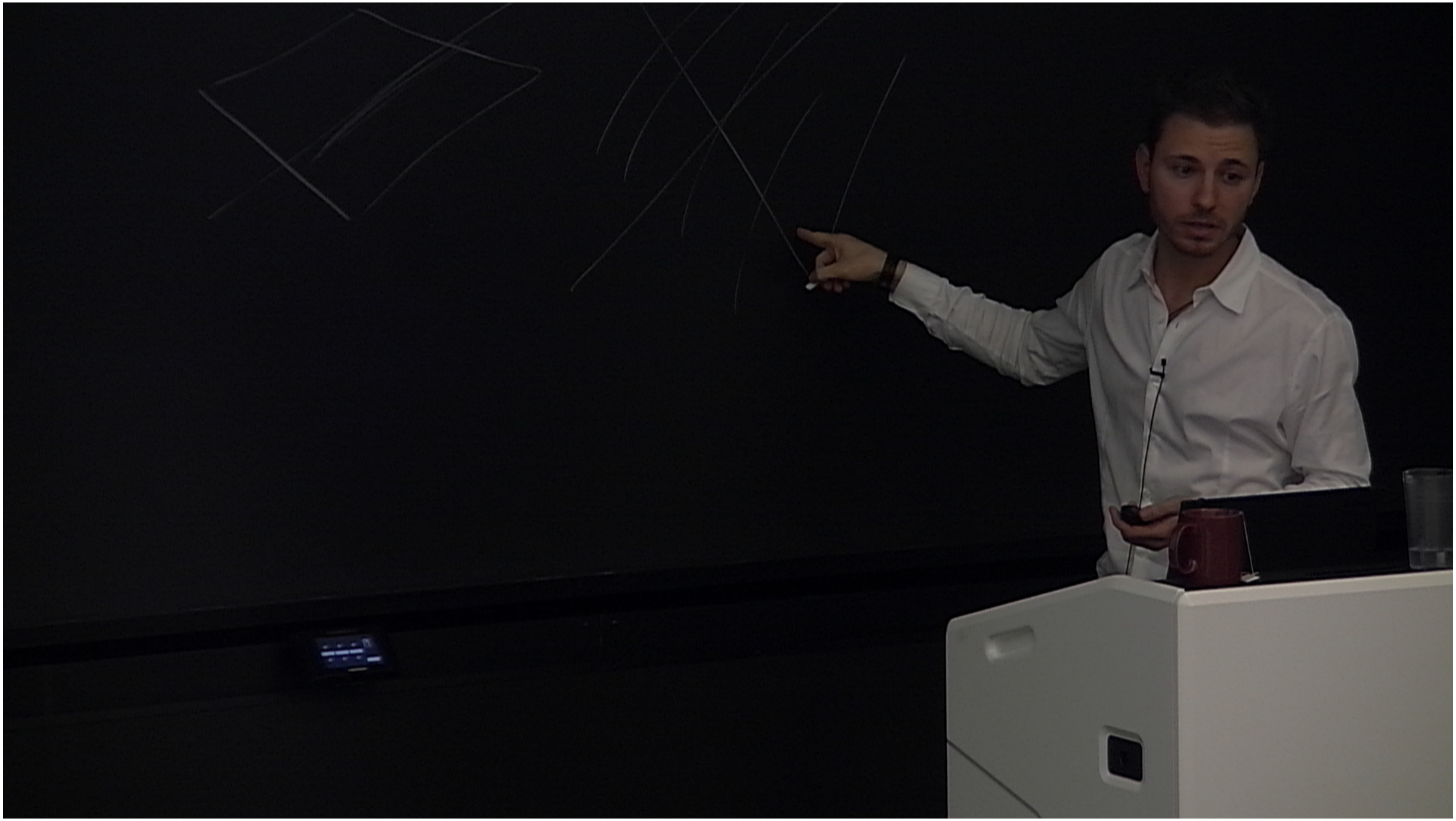
Information Lattice

Question:

Does the information lattice have anything to do with entropy?



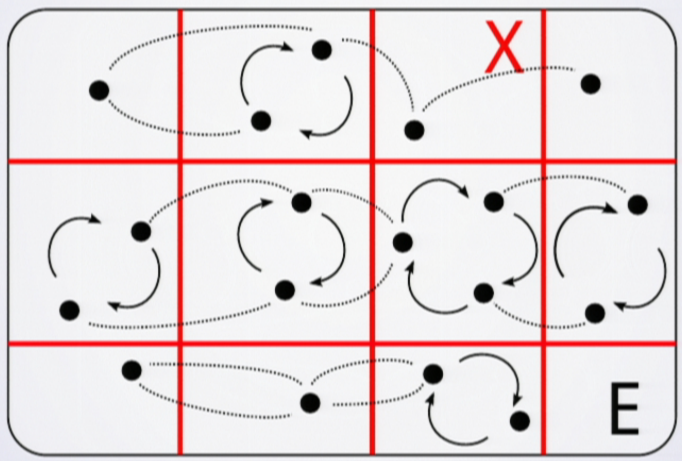




Information Lattice

Question:

Does the information lattice have anything to do with entropy?



Information Lattice

Question:

Does the information lattice have anything to do with entropy?

(Even if the elements of the lattice are sets, the lattice of quotients is not a lattice of sets as defined before, the join is not the union.)

Definition:

The *coarse entropy* of X is the log-cardinality of the quotient $h(X)$:

$$H_c(X) := \log |h(X)| = \log \frac{|E|!}{\prod_{x \in X} |p^{-1}(x)|!}$$

This is a natural monotone function on the information lattice.

Information Lattice

Question:

Does the information lattice have anything to do with entropy?

(Even if the elements of the lattice are sets, the lattice of quotients is not a lattice of sets as defined before, the join is not the union.)

Definition:

The *coarse entropy* of X is the log-cardinality of the quotient $h(X)$:

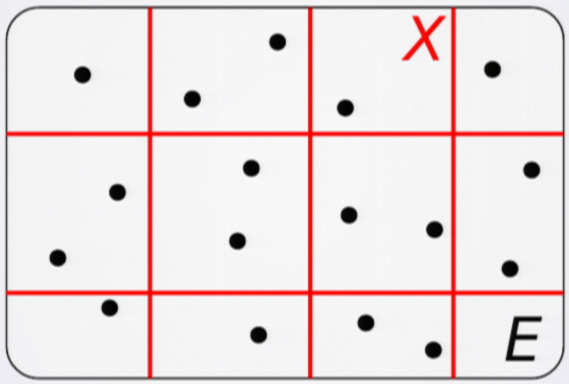
$$H_c(X) := \log |h(X)| = \log \frac{|E|!}{\prod_{x \in X} |p^{-1}(x)|!}$$

This is a natural monotone function on the information lattice.

Information Lattice

Question:

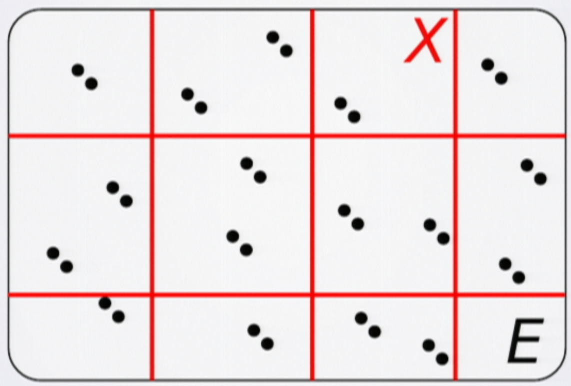
Does the information lattice have anything to do with entropy?



Information Lattice

Question:

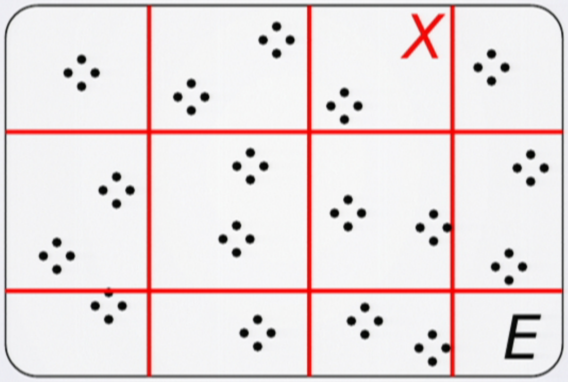
Does the information lattice have anything to do with entropy?



Information Lattice

Question:

Does the information lattice have anything to do with entropy?



Information Lattice

Question:

Does the information lattice have anything to do with entropy?

Stirling's approximation:

$$H_c(X) = \log \frac{|nE|!}{\prod_{x \in X} |np^{-1}(x)|!}$$

Information Lattice

Question:

Does the information lattice have anything to do with entropy?

Stirling's approximation:

$$H_c(X) = \log \frac{|nE|!}{\prod_{x \in X} |np^{-1}(x)|!} \xrightarrow{n \to \infty} \log \frac{|E|^{|nE|}}{\prod_{x \in X} p_x^{np_x}}$$

Information Lattice

Question:

Does the information lattice have anything to do with entropy?

Stirling's approximation:

$$H_c(X) = \log \frac{|nE|!}{\prod_{x\in X} |np^{-1}(x)|!} \xrightarrow{n\to\infty} \log \frac{|E|^{|nE|}}{\prod_{x\in X} p_x^{np_x}}$$

$$= -n \sum_{x\in X} p_x \log \frac{p_x}{|E|}$$

Information Lattice

Question:

Does the information lattice have anything to do with entropy?

Stirling's approximation:

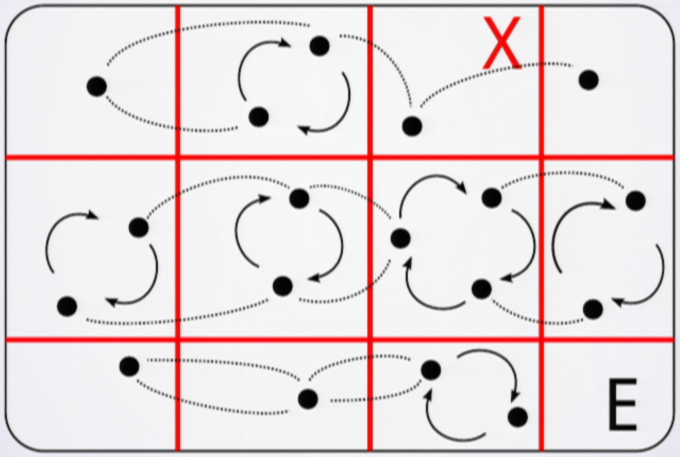
$$\lim_{n\to\infty} \frac{H_c(X)}{|E|} = H(X)$$

Up to the constant, the limit rate is the Shannon entropy of X.

Information Lattice

Question:

Does the information lattice have anything to do with entropy?



Information Lattice

Question:

Does the information lattice have anything to do with entropy?

Stirling's approximation:

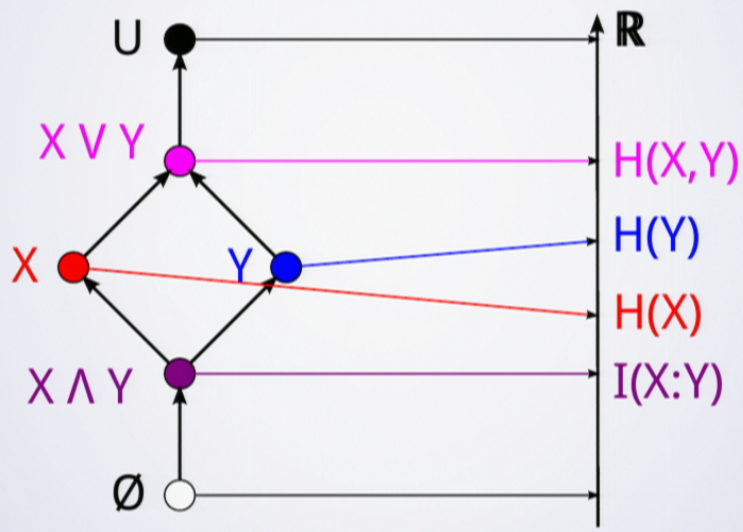
$$\lim_{n\to\infty} \frac{H_c(X)}{|E|} = H(X)$$

Up to the constant, the limit rate is the Shannon entropy of X.
Shannon-type inequalities do not depend on the distribution.
(Rigorous mathematical treatment in Chan [4] and Yeung [5].)

Information Lattice

Question:

Does the information lattice have anything to do with entropy?



Representability

Question:

Is the information lattice always isomorphic to some lattice of sets?

Representability

Question:
Is the information lattice always isomorphic to some lattice of sets?

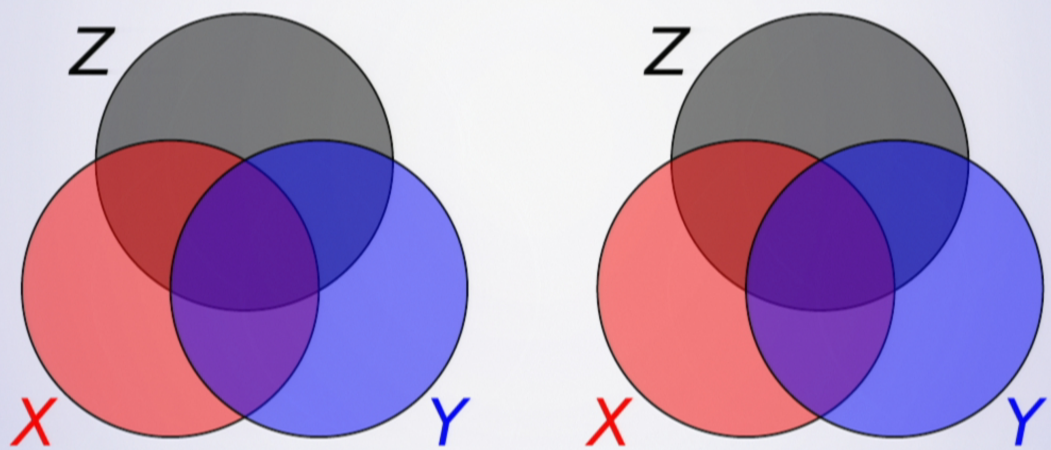
Remark:
A lattice of sets L satisfies the distributive law for every $x, y, z \in L$:

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$$

Representability

Question:

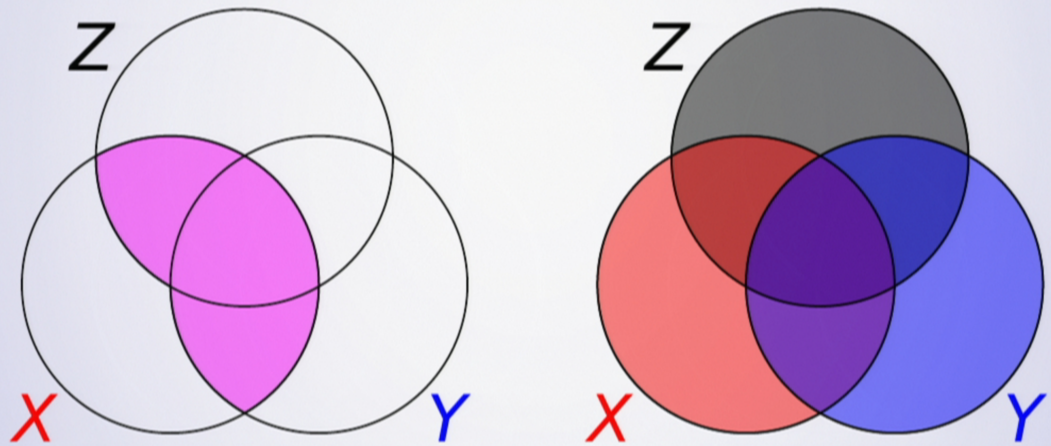
Is the information lattice always isomorphic to some lattice of sets?



Representability

Question:

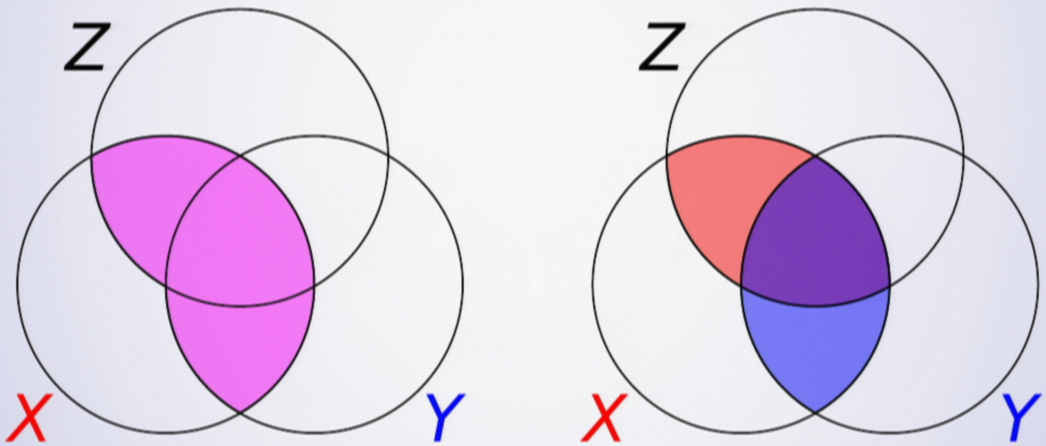
Is the information lattice always isomorphic to some lattice of sets?



Representability

Question:

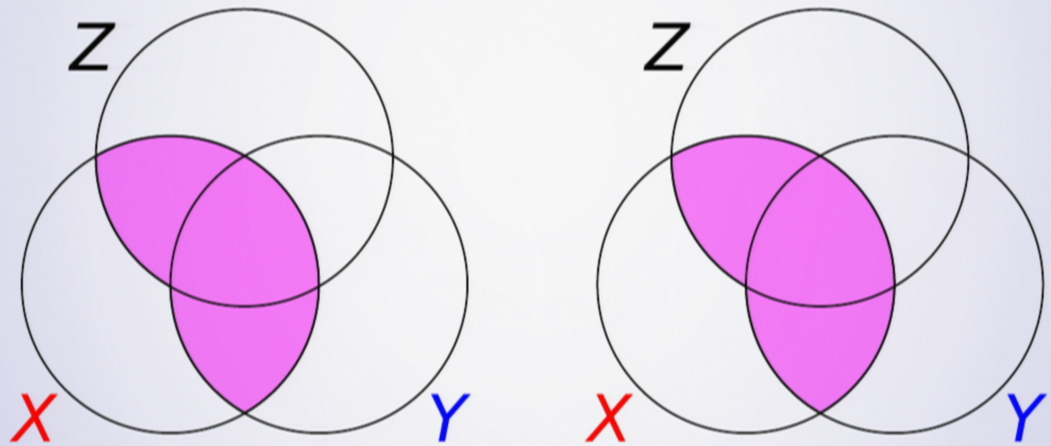
Is the information lattice always isomorphic to some lattice of sets?



Representability

Question:

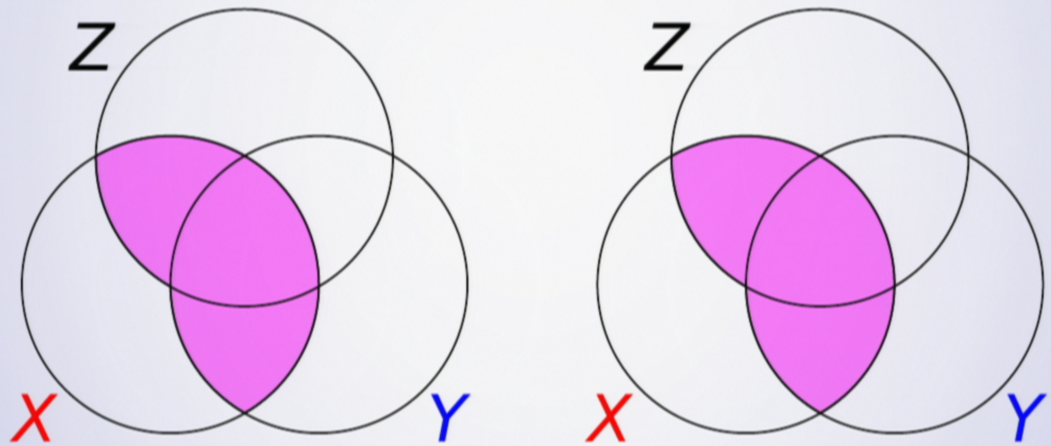
Is the information lattice always isomorphic to some lattice of sets?



Representability

Question:

Is the information lattice always isomorphic to some lattice of sets?



Representability

Question:

Is the information lattice always isomorphic to some lattice of sets?

Remark:

A lattice of sets L satisfies the distributive law for every $x, y, z \in L$:

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$$

Theorem (Birkhoff):

A finite lattice is isomorphic to a lattice of sets if and only if it is distributive. (Sometimes called "representable lattice".)

Representability

Question:

Is the information lattice always isomorphic to some lattice of sets?

Representability

Question:

Is the information lattice always isomorphic to some lattice of sets?

Theorem (Pudlák-Tůma):

Every finite lattice can be embedded in a finite partition lattice.

Representability

Question:

Is the information lattice always isomorphic to some lattice of sets?

Theorem (Pudlák-Tůma):

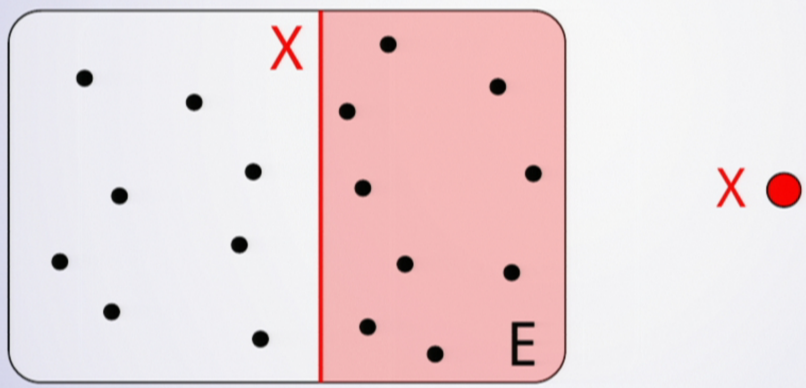
Every finite lattice can be embedded in a finite partition lattice.

Since not every finite lattice is distributive, *some* sublattice of a partition lattice will not be representable.

Representability

Question:

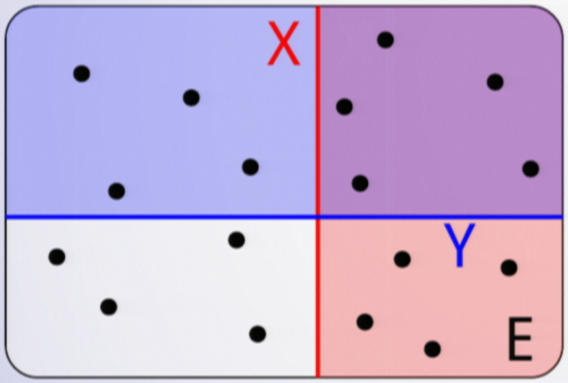
Is the information lattice always isomorphic to some lattice of sets?

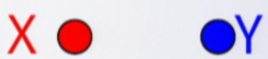


Representability

Question:

Is the information lattice always isomorphic to some lattice of sets?

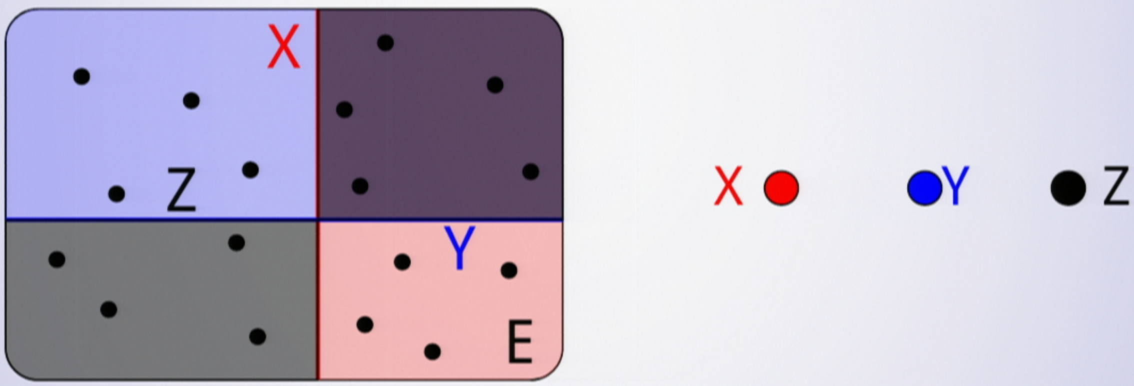




Representability

Question:

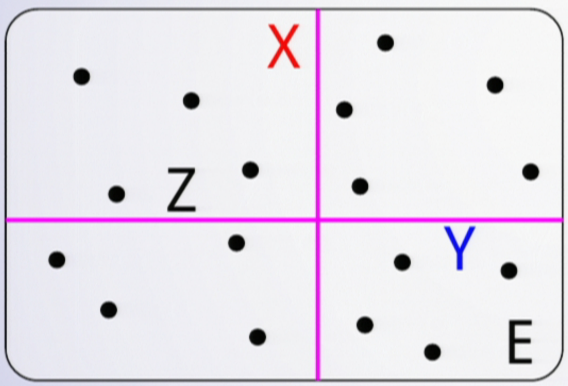
Is the information lattice always isomorphic to some lattice of sets?

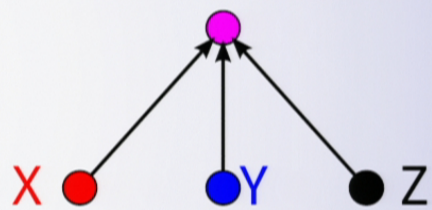


Representability

Question:

Is the information lattice always isomorphic to some lattice of sets?

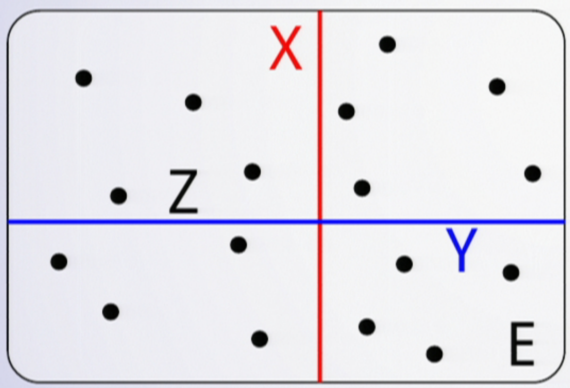


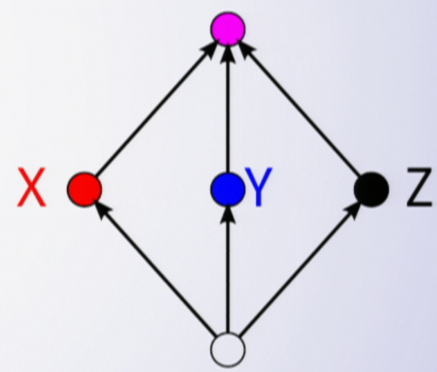


Representability

Question:

Is the information lattice always isomorphic to some lattice of sets?

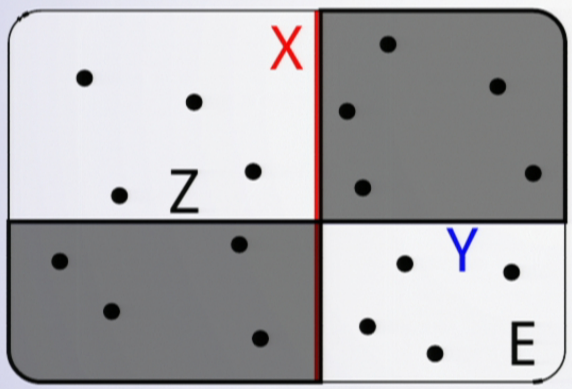


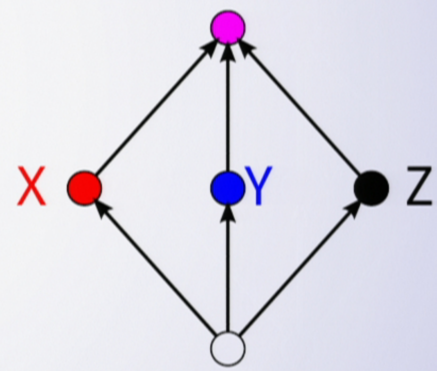


Representability

Question:

Is the information lattice always isomorphic to some lattice of sets?

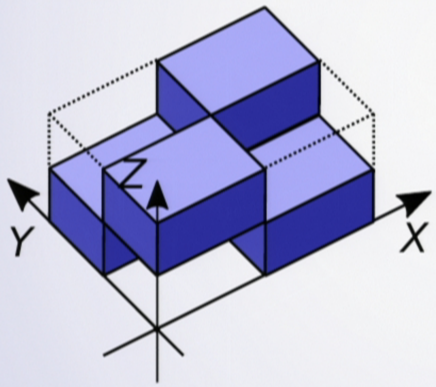


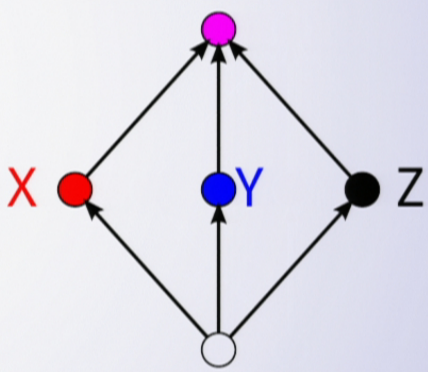


Representability

Question:

Is the information lattice always isomorphic to some lattice of sets?

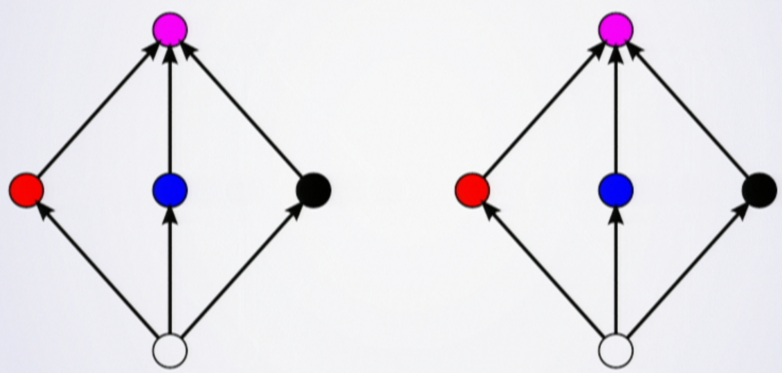




Representability

Question:

Is the information lattice always isomorphic to some lattice of sets?

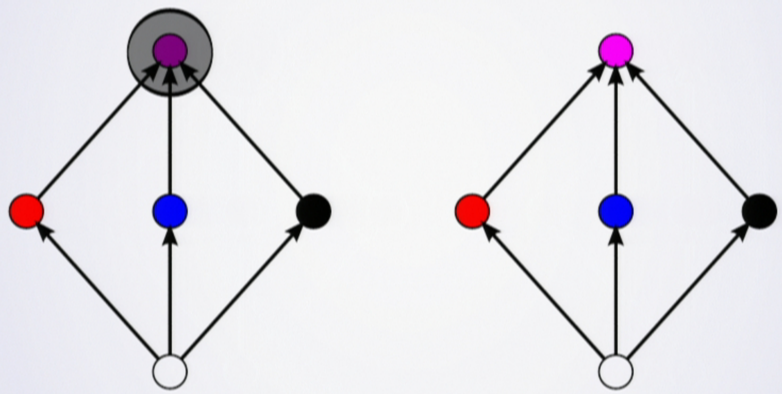


$X \wedge (Y \vee Z)$ $(X \wedge Y) \vee (X \wedge Z)$

Representability

Question:

Is the information lattice always isomorphic to some lattice of sets?

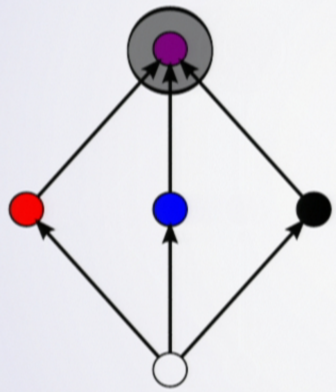


$X \wedge U$ $(X \wedge Y) \vee (X \wedge Z)$

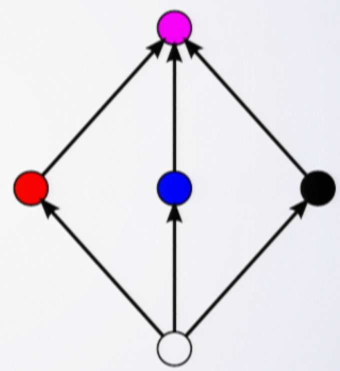
Representability

Question:

Is the information lattice always isomorphic to some lattice of sets?



X ∧ U

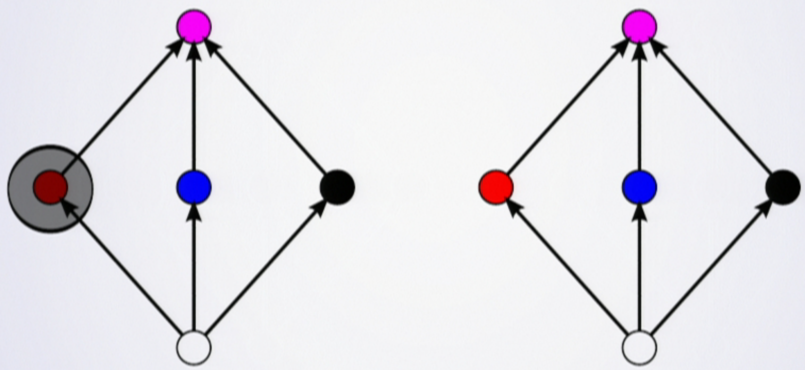


(X ∧ Y) ∨ (X ∧ Z)

Representability

Question:

Is the information lattice always isomorphic to some lattice of sets?

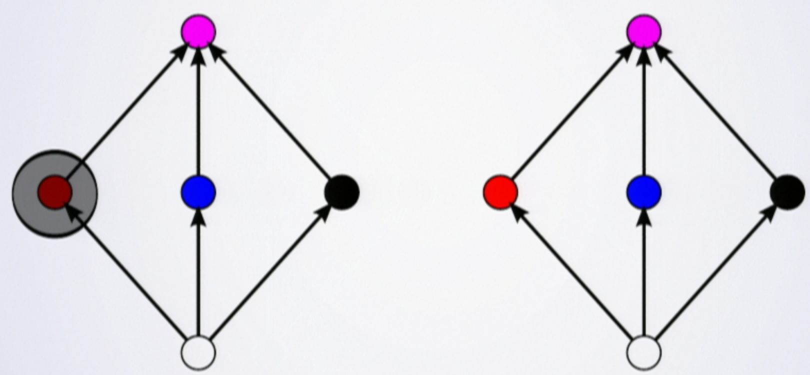


X $\qquad$ $(X \wedge Y) \vee (X \wedge Z)$

Representability

Question:

Is the information lattice always isomorphic to some lattice of sets?

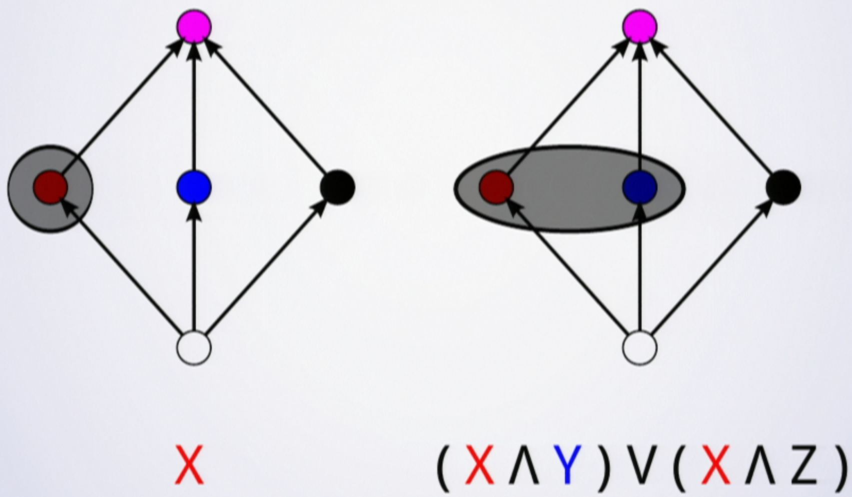


X   (X ∧ Y) ∨ (X ∧ Z)

Representability

Question:

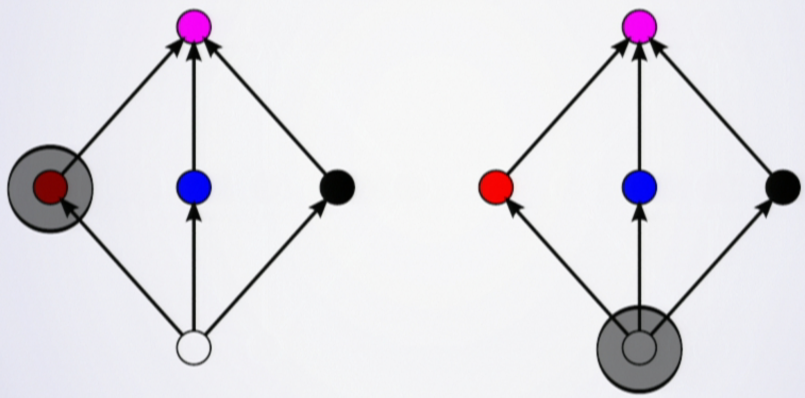
Is the information lattice always isomorphic to some lattice of sets?



Representability

Question:

Is the information lattice always isomorphic to some lattice of sets?

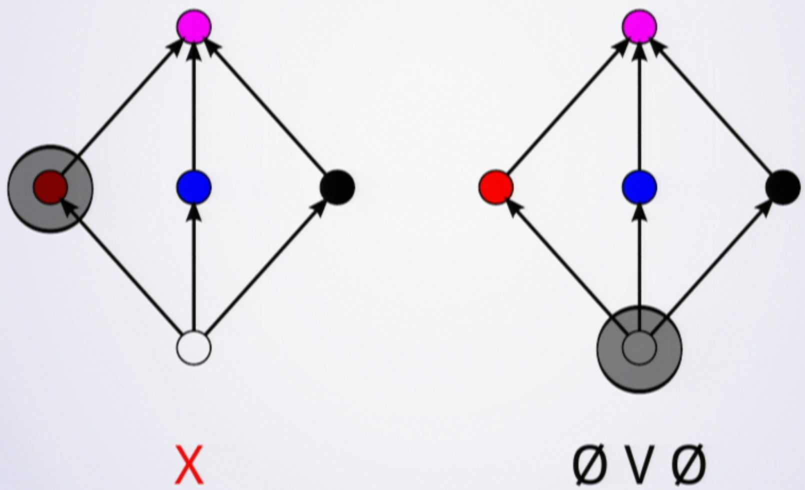


X $\quad\quad$ Ø $\vee$ (X $\wedge$ Z)

Representability

Question:

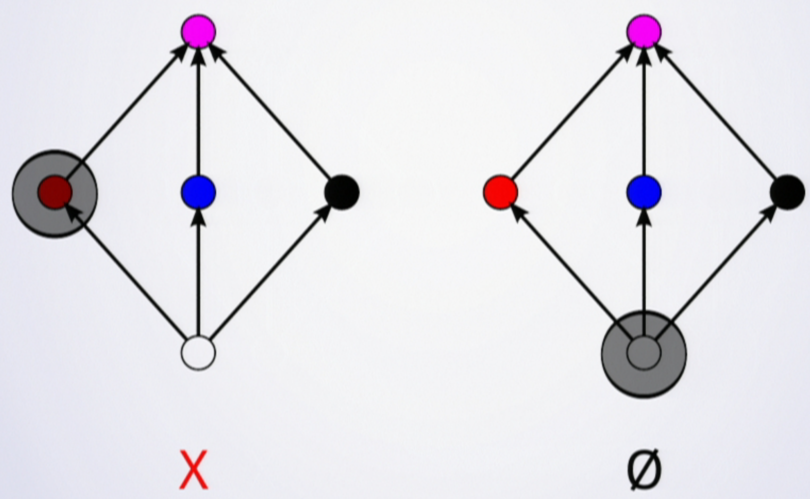
Is the information lattice always isomorphic to some lattice of sets?



Representability

Question:

Is the information lattice always isomorphic to some lattice of sets?



Conclusion

- The information of N observables forms naturally a lattice, as a sublattice of partitions of the outcome space.
- Entropy is a monotone function on the information lattice, obtained from permutations in the large numbers limit.
- The information lattice is not distributive for $N > 2$.
- Therefore, Shannon-type inequalities can be represented faithfully with Venn diagrams only for $N \leq 2$ random variables.

Conclusion

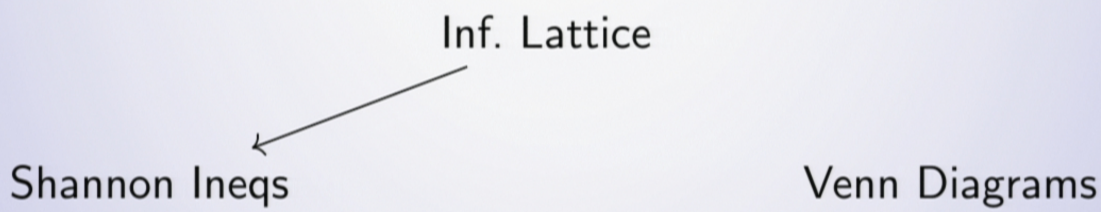
- The information of N observables forms naturally a lattice, as a sublattice of partitions of the outcome space.
- Entropy is a monotone function on the information lattice, obtained from permutations in the large numbers limit.
- The information lattice is not distributive for $N > 2$.
- Therefore, Shannon-type inequalities can be represented faithfully with Venn diagrams only for $N \leq 2$ random variables.

Inf. Lattice

Shannon Ineqs   Venn Diagrams

Conclusion

- The information of N observables forms naturally a lattice, as a sublattice of partitions of the outcome space.
- Entropy is a monotone function on the information lattice, obtained from permutations in the large numbers limit.
- The information lattice is not distributive for $N > 2$.
- Therefore, Shannon-type inequalities can be represented faithfully with Venn diagrams only for $N \leq 2$ random variables.



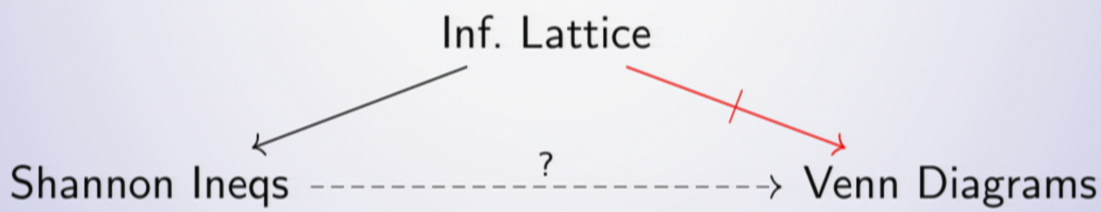
16 of 18

Conclusion

- The information of N observables forms naturally a lattice, as a sublattice of partitions of the outcome space.
- Entropy is a monotone function on the information lattice, obtained from permutations in the large numbers limit.
- The information lattice is not distributive for $N > 2$.
- Therefore, Shannon-type inequalities can be represented faithfully with Venn diagrams only for $N \leq 2$ random variables.

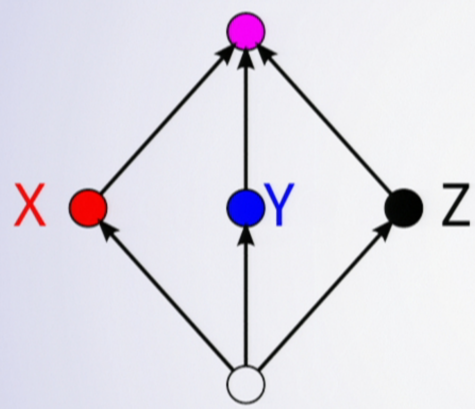
Conclusion

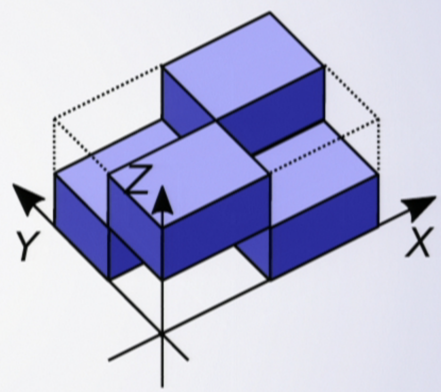
- The information of N observables forms naturally a lattice, as a sublattice of partitions of the outcome space.
- Entropy is a monotone function on the information lattice, obtained from permutations in the large numbers limit.
- The information lattice is not distributive for $N > 2$.
- Therefore, Shannon-type inequalities can be represented faithfully with Venn diagrams only for $N \leq 2$ random variables.



Further Questions

- Synergy and M_3: do they imply one another?



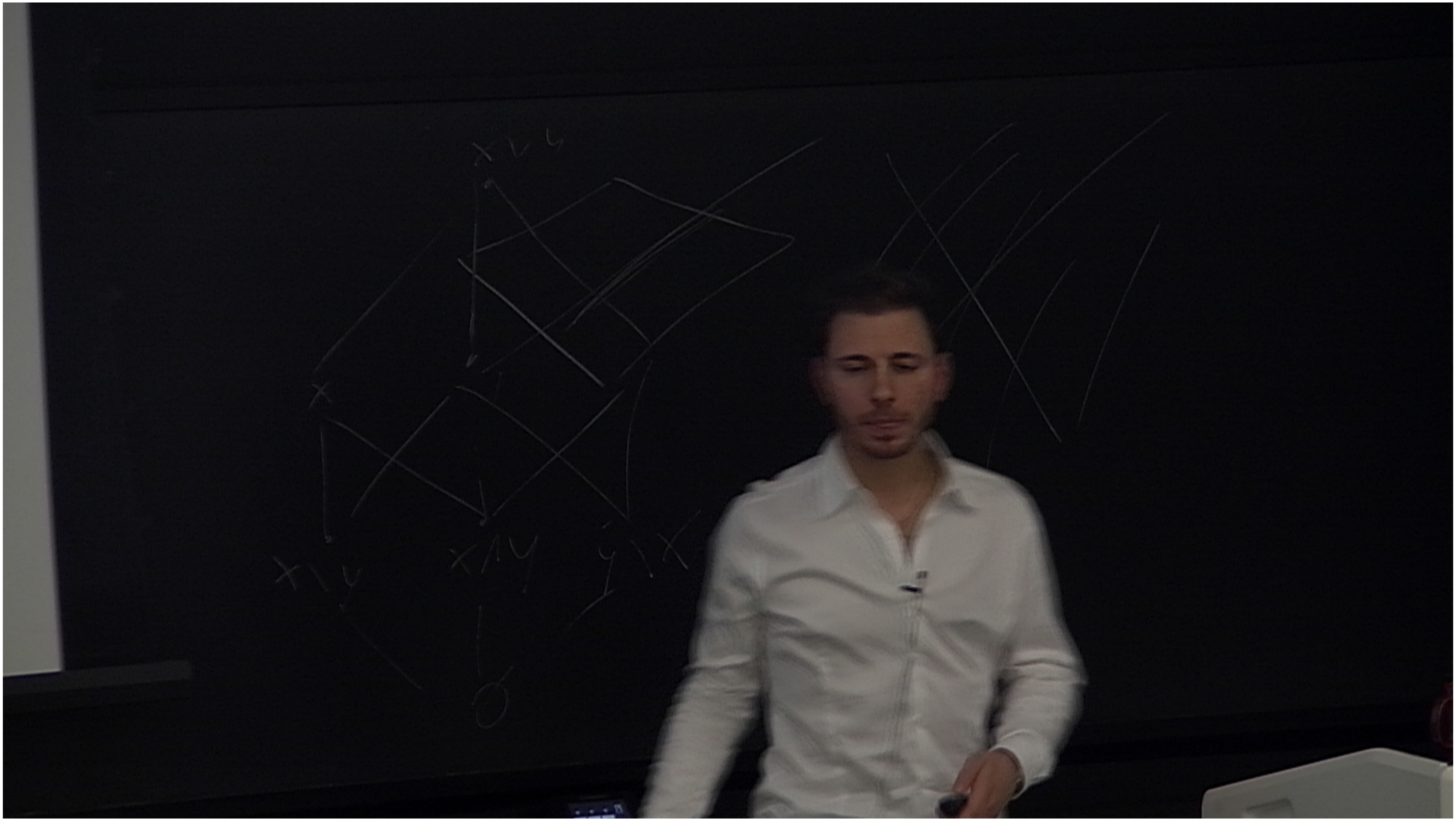


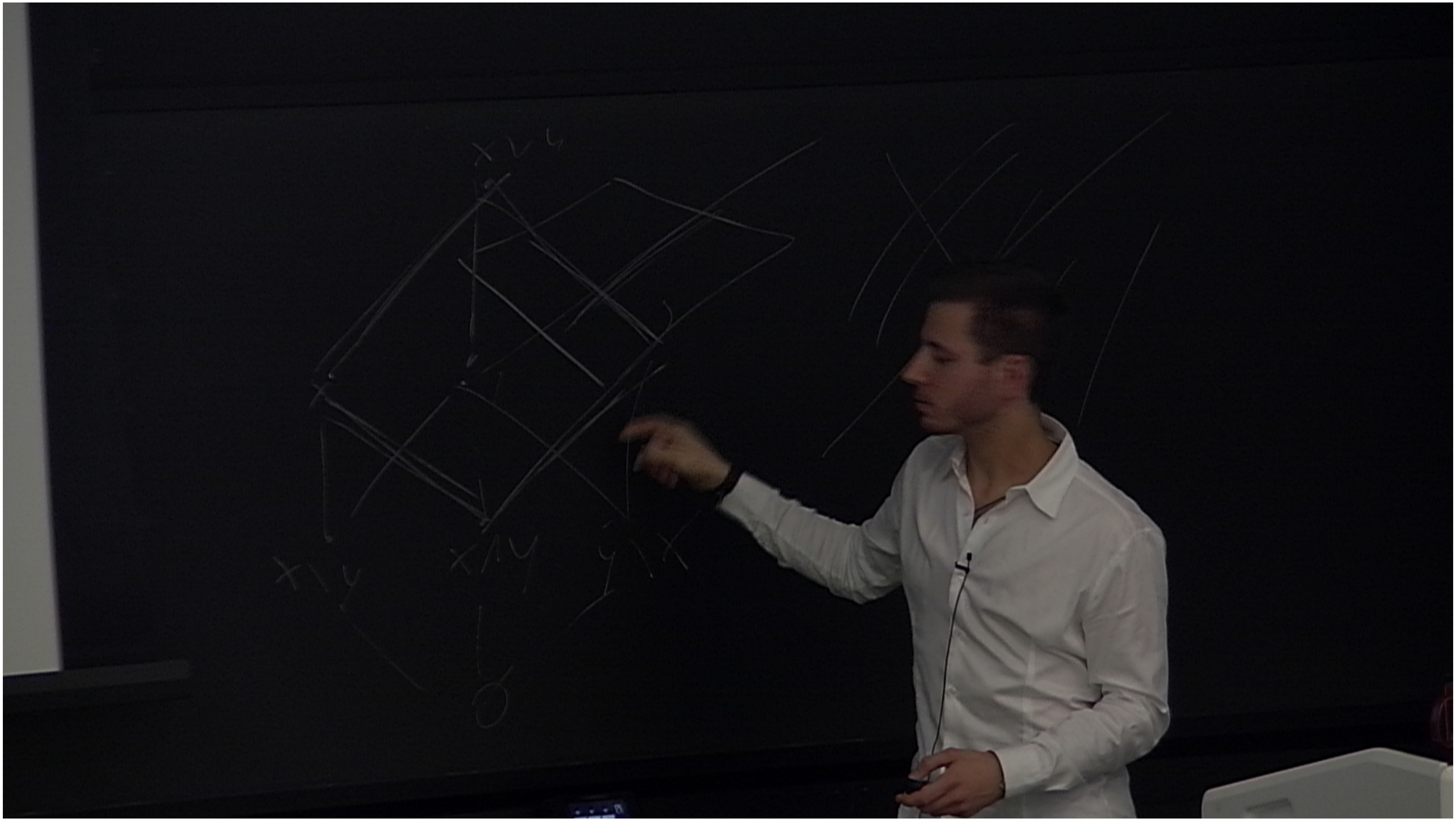
- More in general, is synergy linked to non-distributivity?

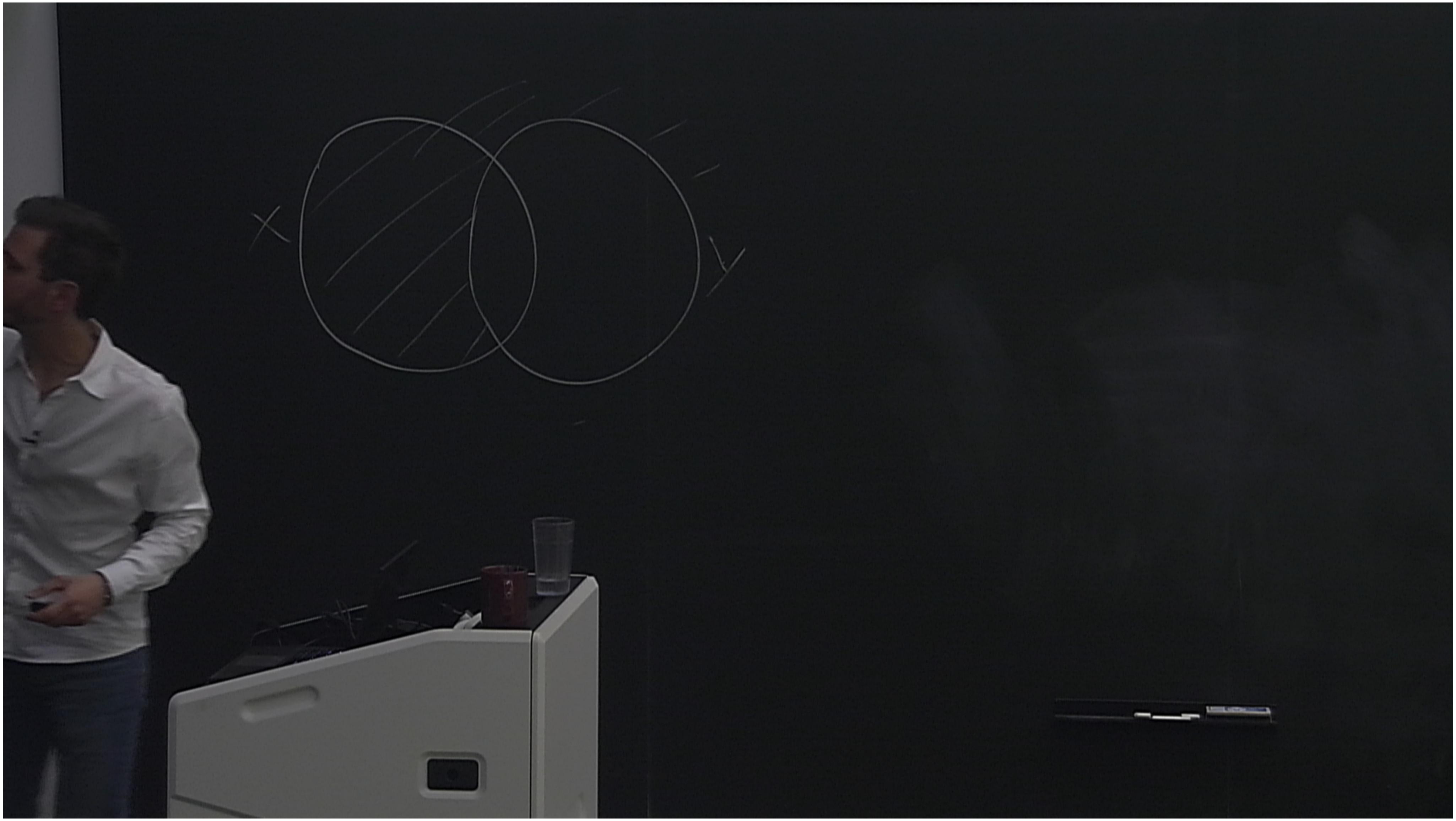
References

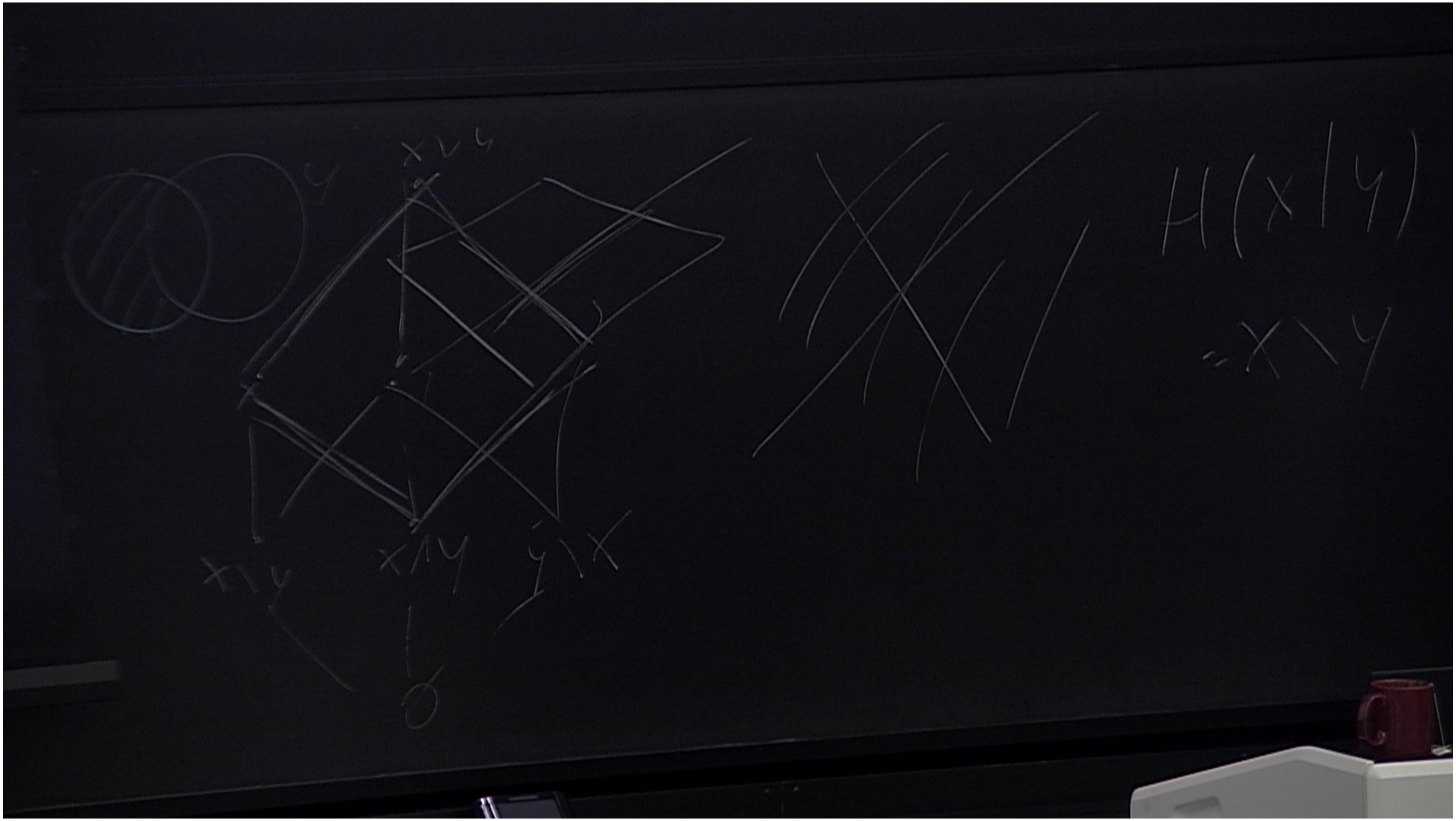
- Elad Schneidmann, William Bialek, and Michael J. II Berry.
  Synergy, Redundancy, and Independence in Population Codes.
  The Journal of Neuroscience, 17, 2003.
- Paul L. Williams and Randall D. Beer.
  Nonnegative Decomposition of Multivariate Information.
  arXiv:1004.2151, 2010.
- Hua Li and Edwin K. P. Chong.
  On a Connection between Information and Group Lattices.
  Entropy, 13:683–708, 2011.
- T. H. Chan.
  Balanced information inequalities.
  IEEE Transactions on Information Theory, IT-49:3261–3267, 2003.
- R. W. Yeung.
  Information Theory and Network Coding.
  Springer, 2008.
- J. C. Baez, T. Fritz, and T. Leinster.
  A Characterization of Entropy in Terms of Information Loss.
  Entropy, 13(11):1945–1957, 2011.
- B. Fong.
  Causal Theories: A Categorical Perspective on Bayesian Networks.
  Master's thesis, University of Oxford, 2012. Available on arXiv:1301.6201v1.
- M. Gromov.
  Symmetry, Probability, Entropy: Synopsis of the Lecture at MAXENT 2014.
  Entropy, 17(3):1273–1277, 2015.
- M. Gromov.
  In a Search for a Structure, Part 1: On Entropy.
  Preprint available at http://www.ihes.fr/gromov, 2012.
- Pavel Pudlák and Jiří Tůma.
  Every finite lattice can be embedded in a finite partition lattice.
  Algebra Universalis, 10(1):74–95, 1980.

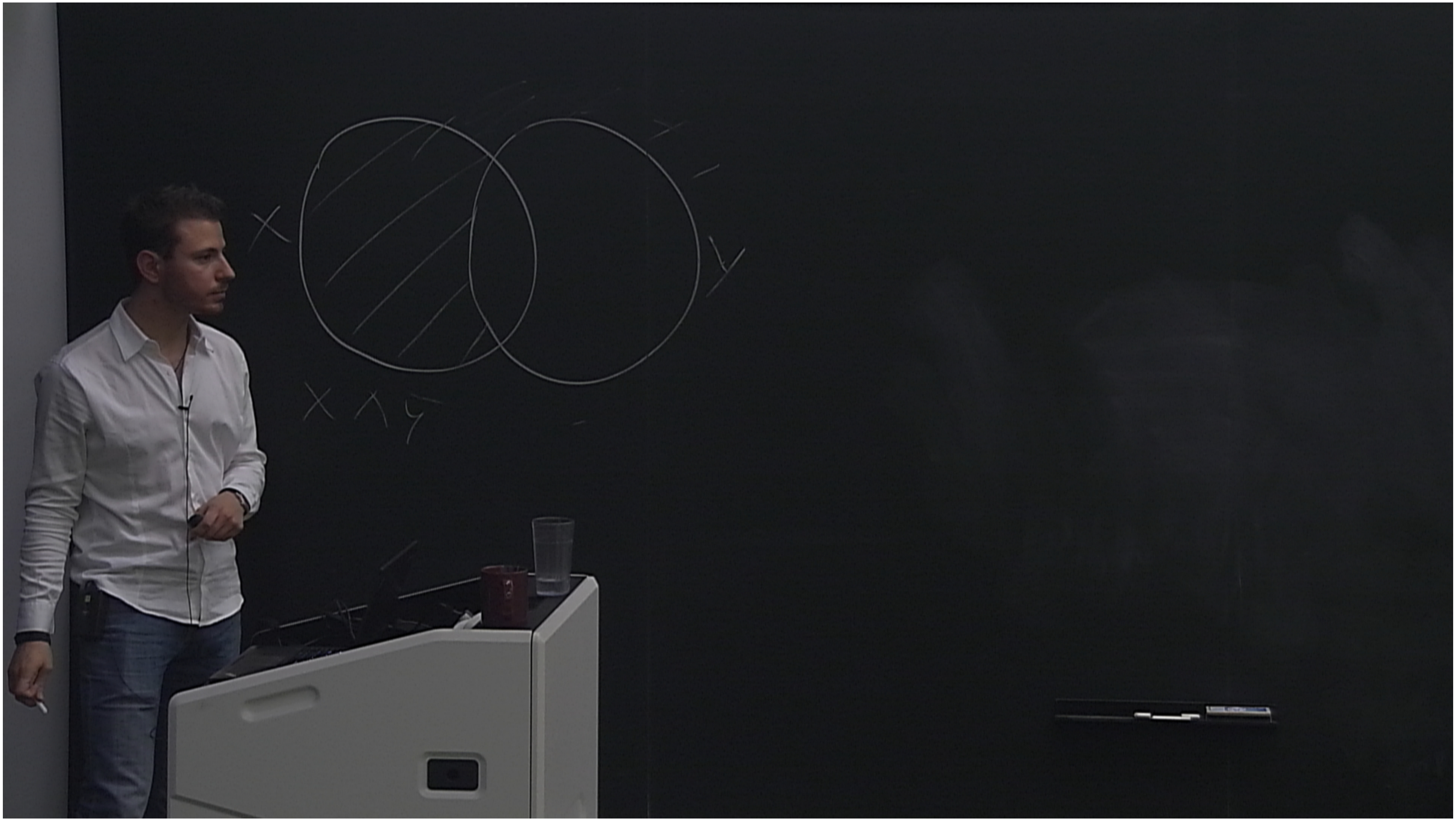

x∨y
x
x\y
x∧y
y\x
0






X
Y

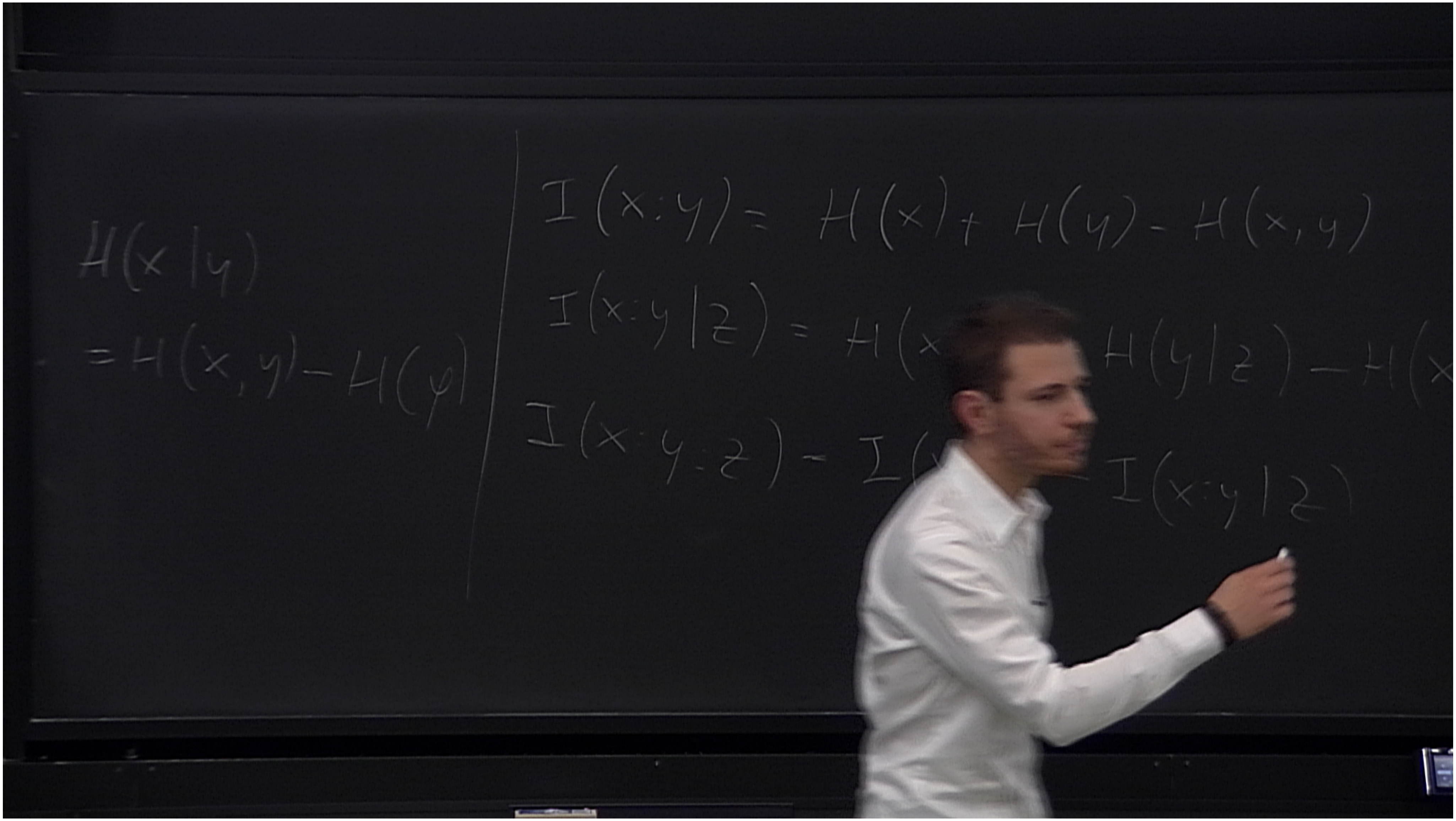

x∨y
y
x∧y
H(x|y)
= x \ y

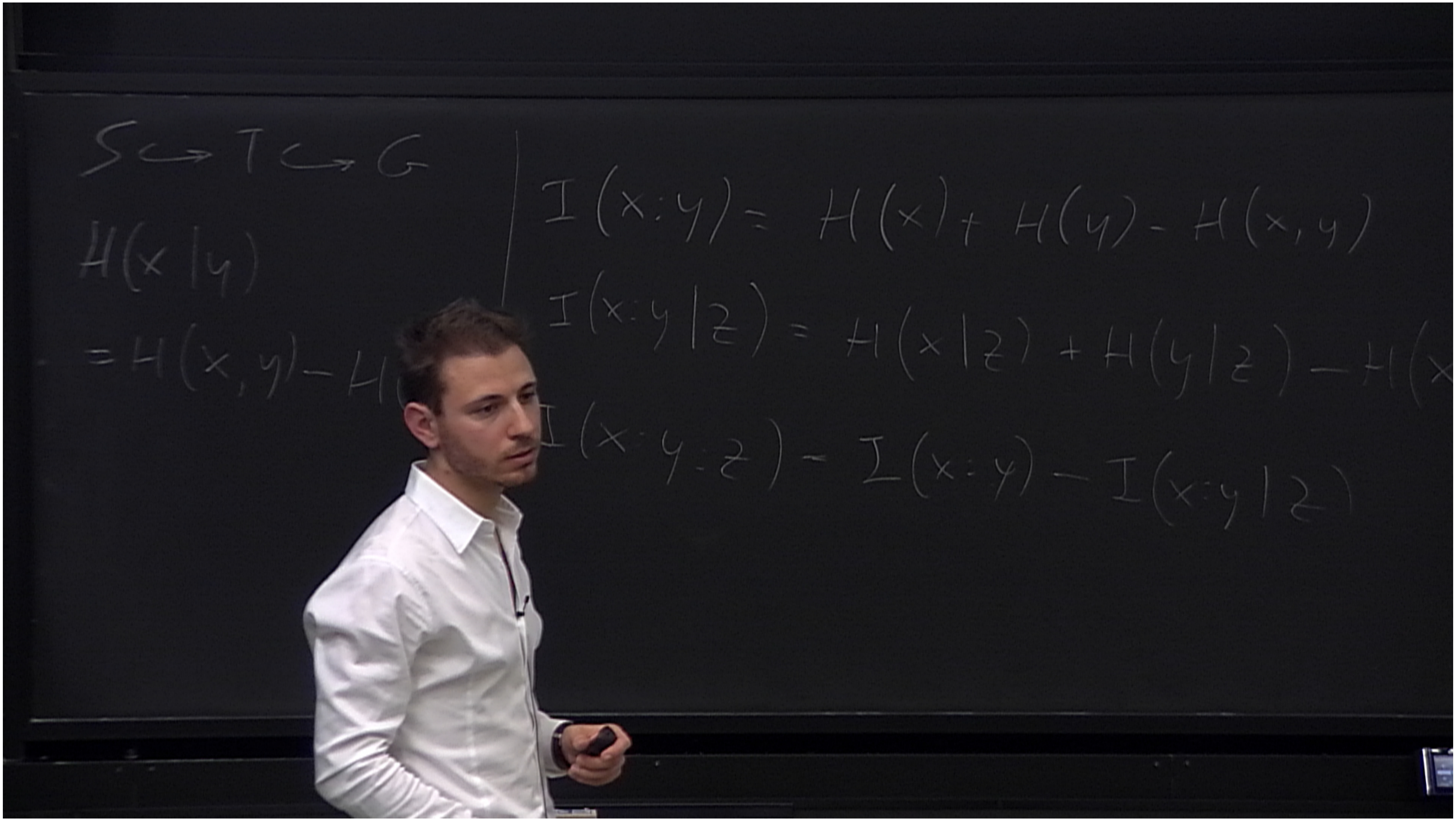

X
Y
X ∧ Y

$$I(x:y) = H(x) + H(y) - H(x,y)$$

$$I(x:y|z) = H(x|z) + H(y|z) - H(x,y|z)$$

$$I(x:y:z) = I(x:y) - I(x:y|z)$$


H(x|y)
=H(x,y)-H(y)
I(x:y) = H(x) + H(y) - H(x,y)
I(x:y|z) = H(x
H(y|z) - H(x
I(x:y:z) = I(
I(x:y|z)


S ↪ T ↪ G
H(x|y)
= H(x,y) − H(
I(x:y) = H(x) + H(y) − H(x,y)
I(x:y|z) = H(x|z) + H(y|z) − H(x
I(x:y:z) = I(x:y) − I(x:y|z)