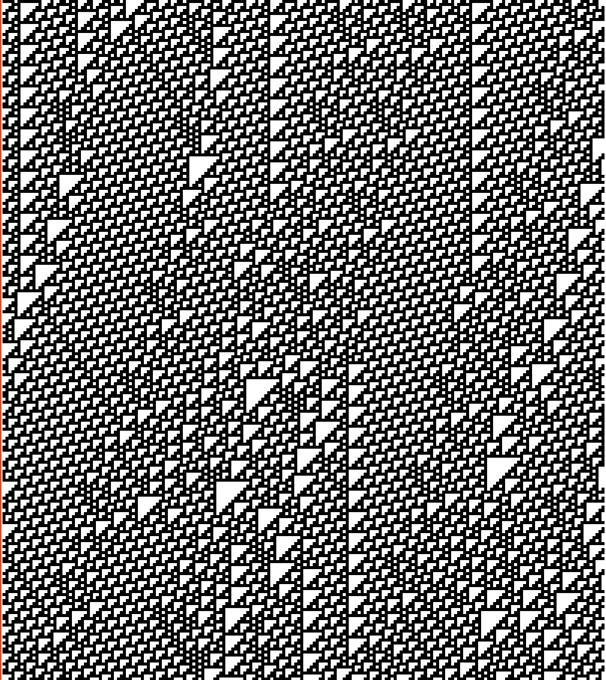
Generalised Measures of Multivariate Information Content

Information Processing in Complex Systems 2019

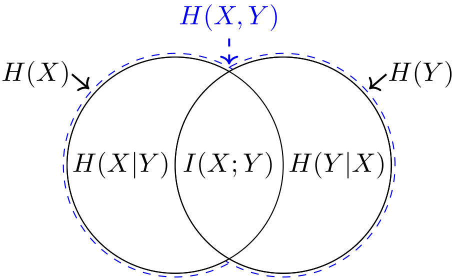
Conor Finn

October 2019

THE UNIVERSITY OF SYDNEY



Venn diagram for two entropies



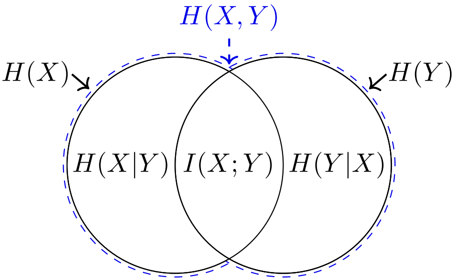
$$H(X) + H(Y) \geq H(X,Y) \geq H(X),\ H(Y) \geq 0$$

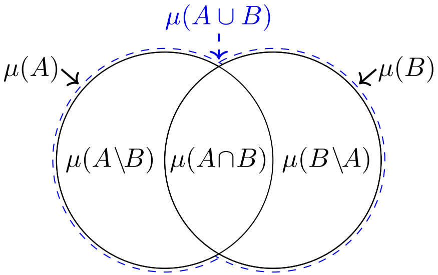
$$H(X|Y) = H(X,Y) - H(Y) \geq 0$$

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

$$I(X;Y) = H(X) + H(Y) - H(X,Y) \geq 0$$

Venn diagram for two entropies





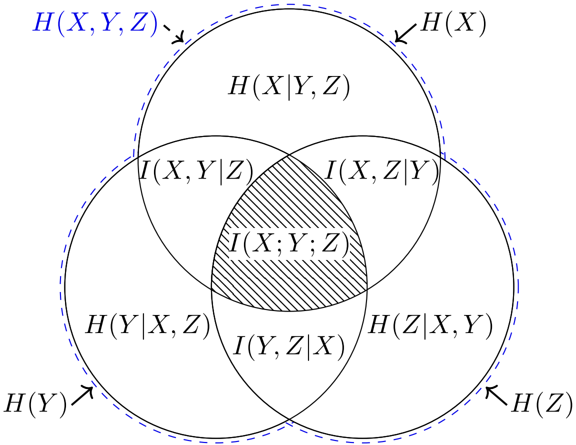
$$H(X) + H(Y) \geq H(X,Y) \geq H(X),\ H(Y) \geq 0$$

$$H(X|Y) = H(X,Y) - H(Y) \geq 0$$
$$H(Y|X) = H(X,Y) - H(X) \geq 0$$
$$I(X;Y) = H(X) + H(Y) - H(X,Y) \geq 0$$

$$\mu(A) + \mu(B) \geq \mu(A \cup B) \geq \mu(A),\ \mu(B) \geq 0$$

$$\mu(A \backslash B) = \mu(A \cup B) - \mu(B) \geq 0$$
$$\mu(B \backslash A) = \mu(A \cup B) - \mu(A) \geq 0$$
$$\mu(A \cap B) = \mu(A) + \mu(B) - \mu(A \cup B) \geq 0$$

Venn diagram for three entropies



- Multivariate mutual information
$$
\begin{aligned}
I(X;Y;Z) = {} & H(X) + H(Y) + H(Z) \\
& - H(X,Y) - H(X,Z) - H(Y,Z) \\
& + H(Z,Y,Z),
\end{aligned}
$$
- Yeung: correspondence between entropy and signed measure
- Multivariate mutual information has "no intuitive meaning"

Marginal information sharing

	Johnny	Alice	Bob
			
Observations:	(X, Y)	X	Y
Knows:	$P(X, Y)$	$P(X)$	$P(Y)$

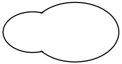
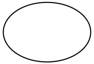
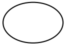
Marginal information sharing

	Johnny	Alice	Bob
			
Observations:	(X, Y)	X	Y
Knows:	$P(X, Y)$	$P(X)$	$P(Y)$
Realisation:	(x, y)	x	y

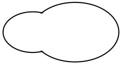
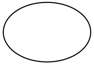
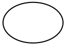
Marginal information sharing

	Johnny	Alice	Bob
			
Observations:	(X, Y)	X	Y
Knows:	$P(X, Y)$	$P(X)$	$P(Y)$
Realisation:	(x, y)	x	y
Information:	$h(x, y)$	$h(x)$	$h(y)$

Marginal information sharing

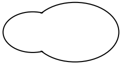
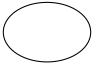
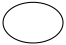
	Johnny	Alice	Bob
			
Observations:	(X, Y)	X	Y
Knows:	$P(X, Y)$	$P(X)$	$P(Y)$
Realisation:	(x, y)	x	y
Information:	$h(x, y)$	$h(x)$	$h(y)$
Venn diagram:			

Marginal information sharing

	Johnny	Alice	Bob
			
Observations:	(X, Y)	X	Y
Knows:	$P(X, Y)$	$P(X)$	$P(Y)$
Realisation:	(x, y)	x	y
Information:	$h(x, y)$	$h(x)$	$h(y)$
Venn diagram:			

$$h(x, y) \geq h(x),\ h(y) \geq 0$$
$$h(x|y) = h(x, y) - h(y) \geq 0$$
$$h(y|x) = h(x, y) - h(x) \geq 0$$

Marginal information sharing

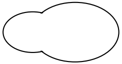
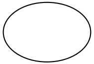
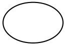
	Johnny	Alice	Bob	Eve
				
Observations:	(X, Y)	X	Y	-
Knows:	$P(X, Y)$	$P(X)$	$P(Y)$	
Realisation:	(x, y)	x	y	
Information:	$h(x, y)$	$h(x)$	$h(y)$	
Venn diagram:				

$$h(x, y) \geq h(x),\ h(y) \geq 0$$

$$h(x|y) = h(x, y) - h(y) \geq 0$$

$$h(y|x) = h(x, y) - h(x) \geq 0$$

Marginal information sharing

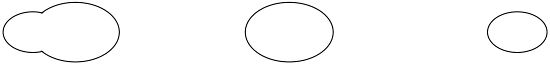
	Johnny	Alice	Bob	Eve
				
Observations:	(X, Y)	X	Y	-
Knows:	$P(X, Y)$	$P(X)$	$P(Y)$	$P(X)$ & $P(Y)$
Realisation:	(x, y)	x	y	(x, y)
Information:	$h(x, y)$	$h(x)$	$h(y)$	
Venn diagram:				

$$h(x, y) \geq h(x),\ h(y) \geq 0$$

$$h(x|y) = h(x, y) - h(y) \geq 0$$

$$h(y|x) = h(x, y) - h(x) \geq 0$$

Marginal information sharing

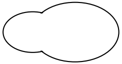
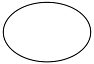
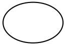
	Johnny	Alice	Bob	Eve
Observations:	(X, Y)	X	Y	-
Knows:	$P(X, Y)$	$P(X)$	$P(Y)$	$P(X)$ & $P(Y)$
Realisation:	(x, y)	x	y	(x, y)
Information:	$h(x, y)$	$h(x)$	$h(y)$	?
Venn diagram:				?

$$h(x, y) \geq h(x),\ h(y) \geq 0$$

$$h(x|y) = h(x, y) - h(y) \geq 0$$

$$h(y|x) = h(x, y) - h(x) \geq 0$$

Marginal information sharing

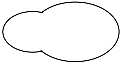
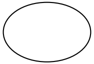
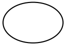
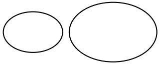
	Johnny	Alice	Bob	Indy
				
Observations:	(X, Y)	X	Y	-
Knows:	$P(X, Y)$	$P(X)$	$P(Y)$	$P(X)$ & $P(Y)$
Realisation:	(x, y)	x	y	(x, y)
Information:	$h(x, y)$	$h(x)$	$h(y)$	
Venn diagram:				

$$h(x, y) \geq h(x),\ h(y) \geq 0$$

$$h(x|y) = h(x, y) - h(y) \geq 0$$

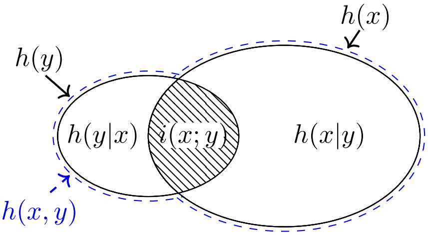
$$h(y|x) = h(x, y) - h(x) \geq 0$$

Marginal information sharing

	Johnny	Alice	Bob	Indy
				
Observations:	(X, Y)	X	Y	-
Knows:	$P(X, Y)$	$P(X)$	$P(Y)$	$P(X)$ & $P(Y)$
Realisation:	(x, y)	x	y	(x, y)
Information:	$h(x, y)$	$h(x)$	$h(y)$	$h(x) + h(y)$
Venn diagram:				

$$h(x, y) \geq h(x),\; h(y) \geq 0$$
$$h(x|y) = h(x, y) - h(y) \geq 0$$
$$h(y|x) = h(x, y) - h(x) \geq 0$$

Venn diagram for information content



- Johnny has at least as much information as Alice and Bob

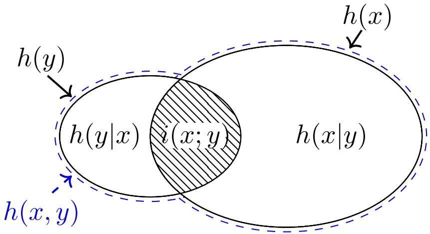
$$h(x,y) \geq h(x),\ h(y) \geq 0$$

$$h(x|y) = h(x,y) - h(y) \geq 0$$
$$h(y|x) = h(x,y) - h(x) \geq 0$$

- Indy can have more or less information than Johnny, i.e. the pointwise mutual information is not non-negative

$$i(x;y) = h(x) + h(y) - h(x,y)$$

Venn diagram for information content



- Johnny has at least as much information as Alice and Bob

$$h(x,y) \geq h(x),\ h(y) \geq 0$$

$$h(x|y) = h(x,y) - h(y) \geq 0$$
$$h(y|x) = h(x,y) - h(x) \geq 0$$

- Indy can have more or less information than Johnny, i.e. the pointwise mutual information is not non-negative

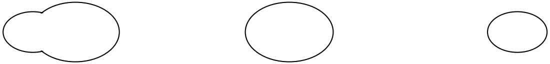
$$i(x;y) = h(x) + h(y) - h(x,y)$$

- Sometimes Indy thinks he has more information than Johnny despite knowing less

Marginal information sharing

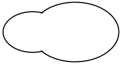
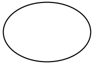
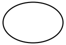
	Johnny	Alice	Bob	Eve
Observations:	(X, Y)	X	Y	-
Knows:	$P(X, Y)$	$P(X)$	$P(Y)$	$P(X)$ & $P(Y)$
Realisation:	(x, y)	x	y	(x, y)
Information:	$h(x, y)$	$h(x)$	$h(y)$	

Venn diagram:



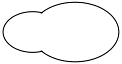
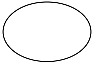
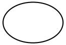
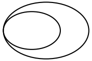
- Eve should have at least as much information as Alice and Bob, but no more than Johnny

Marginal information sharing

	Johnny	Alice	Bob	Eve
				
Observations:	(X, Y)	X	Y	-
Knows:	$P(X, Y)$	$P(X)$	$P(Y)$	$P(X)$ & $P(Y)$
Realisation:	(x, y)	x	y	(x, y)
Information:	$h(x, y)$	$h(x)$	$h(y)$	$\max\big(h(x), h(y)\big)$
Venn diagram:				

- Eve should have at least as much information as Alice and Bob, but no more than Johnny
- It is not difficult to see that Eve's information is given by $h(x \sqcup y) := \max\big(h(x), h(y)\big)$

Marginal information sharing

	Johnny	Alice	Bob	Eve
				
Observations:	(X, Y)	X	Y	-
Knows:	$P(X, Y)$	$P(X)$	$P(Y)$	$P(X)$ & $P(Y)$
Realisation:	(x, y)	x	y	(x, y)
Information:	$h(x, y)$	$h(x)$	$h(y)$	$\max\big(h(x), h(y)\big)$
Venn diagram:				

- Eve should have at least as much information as Alice and Bob, but no more than Johnny
- It is not difficult to see that Eve's information is given by $h(x \sqcup y) := \max\big(h(x), h(y)\big)$

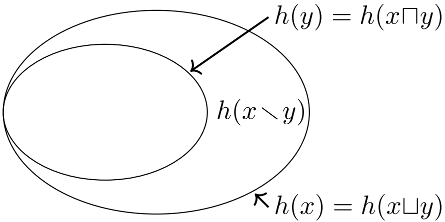
Union and intersection information content

- Union information content

$$h(x \sqcup y) = \max\big(h(x), h(y)\big)$$

satisfies

$$h(x) + h(y) \geq h(x \sqcup y) \geq h(x),\ h(y) \geq 0$$



Union and intersection information content

- Union information content

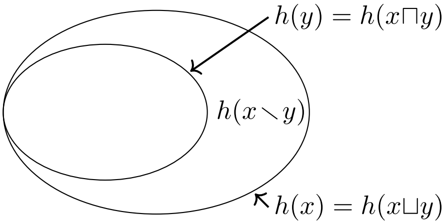
$$h(x \sqcup y) = \max\big(h(x), h(y)\big)$$

satisfies

$$h(x) + h(y) \geq h(x \sqcup y) \geq h(x),\ h(y) \geq 0$$

- Unique information content

$$\begin{aligned} h(x \smallsetminus y) &= h(x \sqcup y) - h(y) \\ &= \max\big(h(x) - h(y), 0\big) \geq 0 \\ h(y \smallsetminus x) &= h(x \sqcup y) - h(x) \\ &= \max\big(0, h(y) - h(x)\big) \geq 0 \end{aligned}$$



Union and intersection information content

- Union information content

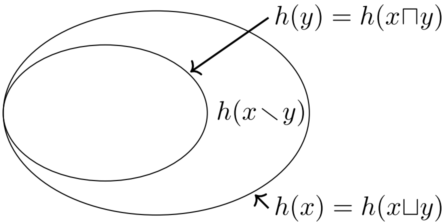
$$h(x \sqcup y) = \max\big(h(x), h(y)\big)$$

satisfies

$$h(x) + h(y) \geq h(x \sqcup y) \geq h(x),\ h(y) \geq 0$$

- Unique information content

$$\begin{aligned} h(x \setminus y) &= h(x \sqcup y) - h(y) \\ &= \max\big(h(x) - h(y), 0\big) \geq 0 \\ h(y \setminus x) &= h(x \sqcup y) - h(x) \\ &= \max\big(0, h(y) - h(x)\big) \geq 0 \end{aligned}$$



- Intersection information content

$$\begin{aligned} h(x \sqcap y) &= h(x) + h(y) - h(x \sqcup y) \\ &= \min\big(h(x), h(y)\big) \geq 0. \end{aligned}$$

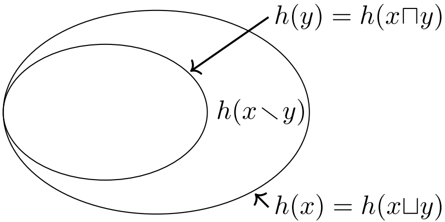
Union and intersection information content

- Union information content

$$h(x \sqcup y) = \max\big(h(x), h(y)\big)$$

satisfies

$$h(x)+h(y) \geq h(x \sqcup y) \geq h(x),\ h(y) \geq 0$$



- Unique information content

$$\begin{aligned} h(x \setminus y) &= h(x \sqcup y) - h(y) \\ &= \max\big(h(x) - h(y), 0\big) \geq 0 \\ h(y \setminus x) &= h(x \sqcup y) - h(x) \\ &= \max\big(0, h(y) - h(x)\big) \geq 0 \end{aligned}$$

- Intersection information content

$$\begin{aligned} h(x \sqcap y) &= h(x) + h(y) - h(x \sqcup y) \\ &= \min\big(h(x), h(y)\big) \geq 0. \end{aligned}$$

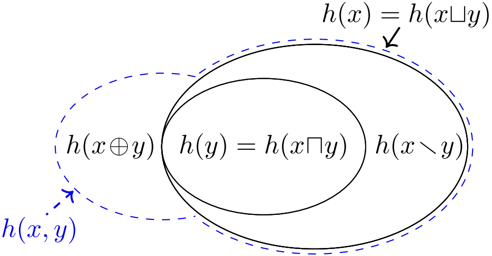
- Decomposition

$$h(x \sqcup y) = h(x \sqcap y) + h(x \setminus y) + h(y \setminus x)$$

Synergistic information content

- Eve has no more information than Johnny

$$h(x,y) \geq h(x \sqcup y)$$



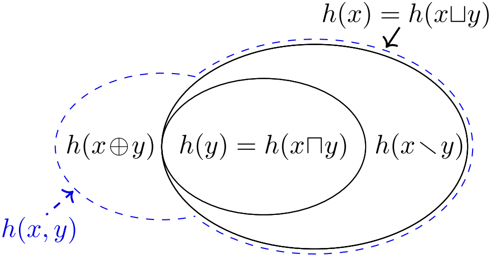
Synergistic information content

- Eve has no more information than Johnny

$$h(x,y) \geq h(x \sqcup y)$$

- Synergistic information content

$$\begin{aligned} h(x \oplus y) &= h(x,y) - h(x \sqcup y) \\ &= \min\big(h(y|x), h(x|y)\big) \geq 0 \end{aligned}$$



Synergistic information content

- Eve has no more information than Johnny

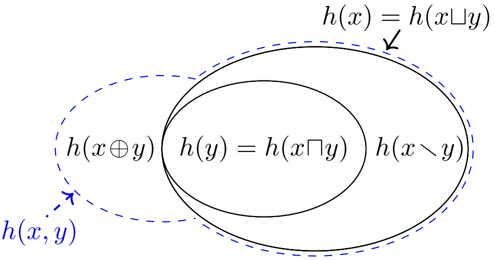
$$h(x, y) \geq h(x \sqcup y)$$

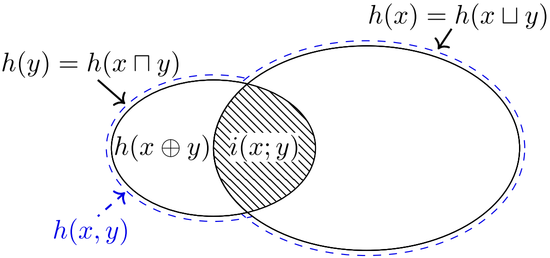
- Synergistic information content

$$\begin{aligned} h(x \oplus y) &= h(x, y) - h(x \sqcup y) \\ &= \min\big(h(y|x), h(x|y)\big) \geq 0 \end{aligned}$$

- Mutual information content

$$i(x; y) = h(x \sqcup y) - h(x \oplus y)$$





Synergistic information content

- Eve has no more information than Johnny

$$h(x,y) \geq h(x \sqcup y)$$

- Synergistic information content

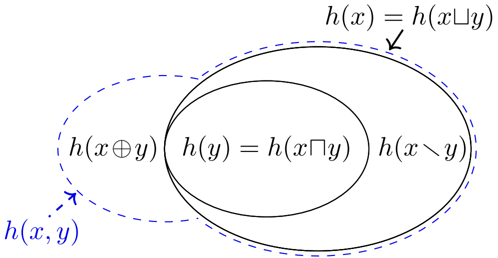
$$\begin{aligned} h(x \oplus y) &= h(x,y) - h(x \sqcup y) \\ &= \min\big(h(y|x), h(x|y)\big) \geq 0 \end{aligned}$$

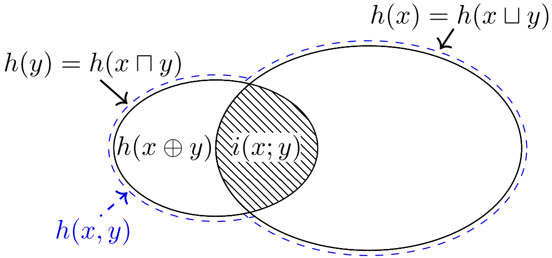
- Mutual information content

$$i(x;y) = h(x \sqcup y) - h(x \oplus y)$$

- Decomposition

$$h(x,y) = h(x \setminus y) + h(y \setminus x) + h(x \sqcap y) + h(x \oplus y)$$





Union and intersection entropy

- Union entropy

$$H(X \sqcup Y) = \mathrm{E}_{XY}\left[h(x \sqcup y)\right]$$

satisfies

$$H(X)+H(Y) \geq H(X \sqcup Y) \geq H(X),\ H(Y) \geq 0$$

- Unique information content

$$H(X \backslash Y) = \mathrm{E}_{XY}\left[h(x \setminus y)\right]$$
$$H(Y \backslash X) = \mathrm{E}_{XY}\left[h(y \setminus x)\right]$$

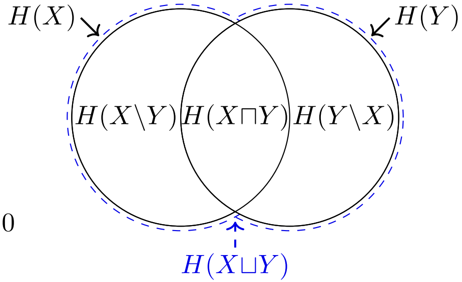
- Intersection information content

$$\begin{aligned} H(X \sqcap Y) &= H(X) + H(Y) - H(X \sqcup Y) \\ &= \mathrm{E}_{XY}\left[h(x \sqcap y)\right] \end{aligned}$$

- Decomposition

$$H(X \sqcup Y) = H(X \sqcap Y) + H(X \backslash Y) + H(Y \backslash X)$$

Union and intersection entropy



- Union entropy

$$H(X \sqcup Y) = \mathrm{E}_{XY}\left[h(x \sqcup y)\right]$$

satisfies

$$H(X) + H(Y) \geq H(X \sqcup Y) \geq H(X),\, H(Y) \geq 0$$

- Unique information content

$$H(X \setminus Y) = \mathrm{E}_{XY}\left[h(x \setminus y)\right]$$
$$H(Y \setminus X) = \mathrm{E}_{XY}\left[h(y \setminus x)\right]$$

- Intersection information content

$$\begin{aligned} H(X \sqcap Y) &= H(X) + H(Y) - H(X \sqcup Y) \\ &= \mathrm{E}_{XY}\left[h(x \sqcap y)\right] \end{aligned}$$

- Decomposition

$$H(X \sqcup Y) = H(X \sqcap Y) + H(X \setminus Y) + H(Y \setminus X)$$

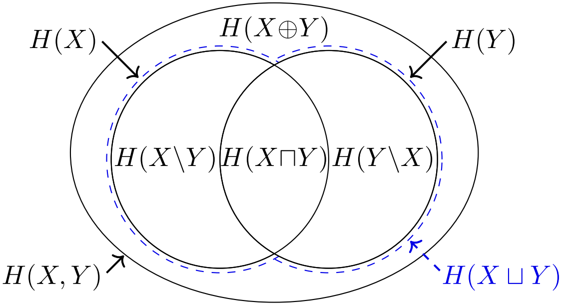
Synergistic entropy

- Synergistic entropy

$$H(X \oplus Y) = H(X,Y) - H(X \sqcup Y)$$
$$= \mathrm{E}_{XY}\big[h(x \oplus y))\big] \geq 0$$

- Mutual information

$$I(X;Y) = H(X \sqcup Y) - H(X \oplus Y)$$



- Decomposition

$$H(X,Y) = H(X \setminus Y) + H(Y \setminus X) + H(X \sqcap Y) + H(X \oplus Y)$$

Generalised marginal information sharing

- Idempotent

$$h(x \sqcup x) = h(x)$$
$$h(x \sqcap x) = h(x)$$

Generalised marginal information sharing

- Idempotent

$$h(x \sqcup x) = h(x)$$
$$h(x \sqcap x) = h(x)$$

- Commutative

$$h(x \sqcup y) = h(y \sqcup x)$$
$$h(x \sqcap y) = h(y \sqcap x)$$

Generalised marginal information sharing

- Idempotent

$$h(x \sqcup x) = h(x)$$
$$h(x \sqcap x) = h(x)$$

- Commutative

$$h(x \sqcup y) = h(y \sqcup x)$$
$$h(x \sqcap y) = h(y \sqcap x)$$

- Associative

$$h(x \sqcup y \sqcup z) = h\big((x \sqcup y) \sqcup z\big) = h\big(x \sqcup (y \sqcup z)\big)$$
$$h(x \sqcap y \sqcap z) = h\big((x \sqcap y) \sqcap z\big) = h\big(x \sqcap (y \sqcap z)\big)$$

Generalised marginal information sharing

- Idempotent

$$h(x \sqcup x) = h(x)$$
$$h(x \sqcap x) = h(x)$$

- Absorption

$$h\big(x \sqcup (x \sqcap y)\big) = h(x)$$
$$h\big(x \sqcap (x \sqcup y)\big) = h(x)$$

- Commutative

$$h(x \sqcup y) = h(y \sqcup x)$$
$$h(x \sqcap y) = h(y \sqcap x)$$

- Associative

$$h(x \sqcup y \sqcup z) = h\big((x \sqcup y) \sqcup z\big) = h\big(x \sqcup (y \sqcup z)\big)$$
$$h(x \sqcap y \sqcap z) = h\big((x \sqcap y) \sqcap z\big) = h\big(x \sqcap (y \sqcap z)\big)$$

Generalised marginal information sharing

- Idempotent

$$h(x \sqcup x) = h(x)$$
$$h(x \sqcap x) = h(x)$$

- Commutative

$$h(x \sqcup y) = h(y \sqcup x)$$
$$h(x \sqcap y) = h(y \sqcap x)$$

- Associative

$$h(x \sqcup y \sqcup z) = h\big((x \sqcup y) \sqcup z\big) = h\big(x \sqcup (y \sqcup z)\big)$$
$$h(x \sqcap y \sqcap z) = h\big((x \sqcap y) \sqcap z\big) = h\big(x \sqcap (y \sqcap z)\big)$$

- Absorption

$$h\big(x \sqcup (x \sqcap y)\big) = h(x)$$
$$h\big(x \sqcap (x \sqcup y)\big) = h(x)$$

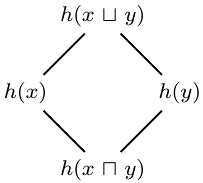
- Distributive

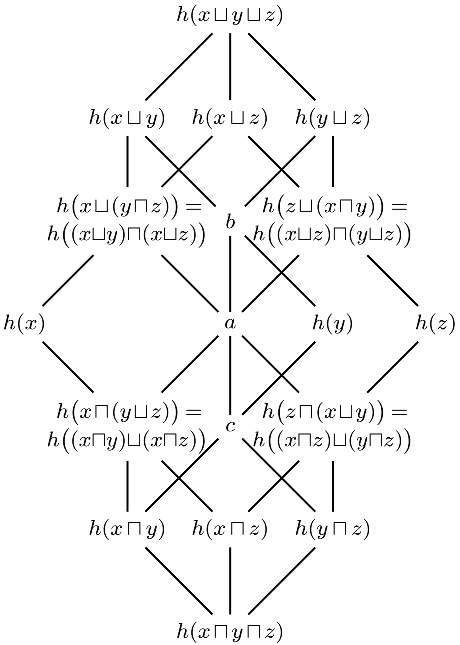
$$h\big(x \sqcup (y \sqcap z)\big) = h\big((x \sqcup y) \sqcap (x \sqcup z)\big)$$
$$h\big(x \sqcap (y \sqcup z)\big) = h\big((x \sqcap y) \sqcup (x \sqcap z)\big)$$

$$
\begin{aligned}
a &= h\big((x \sqcup y) \sqcap (x \sqcup z) \sqcap (y \sqcup z)\big)\\
&= h\Big(\big(x \sqcup (y \sqcap z)\big) \sqcap \big(y \sqcup (x \sqcap z)\big)\Big)\\
&= h\Big(\big(x \sqcup (y \sqcap z)\big) \sqcap \big(z \sqcup (x \sqcap y)\big)\Big)\\
&= h\Big(\big(y \sqcup (x \sqcap z)\big) \sqcap \big(z \sqcup (x \sqcap y)\big)\Big)\\
&= h\Big(\big(y \sqcap (x \sqcup z)\big) \sqcup \big(z \sqcap (x \sqcup y)\big)\Big)\\
&= h\Big(\big(x \sqcap (y \sqcup z)\big) \sqcup \big(z \sqcap (x \sqcup y)\big)\Big)\\
&= h\Big(\big(x \sqcap (y \sqcup z)\big) \sqcup \big(y \sqcap (x \sqcup z)\big)\Big)\\
&= h\big((x \sqcap y) \sqcup (x \sqcap z) \sqcup (y \sqcap z)\big)
\end{aligned}
$$

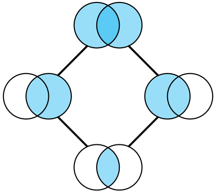
$$b = h\big(y \sqcup (x \sqcap z)\big) = h\big((x \sqcup y) \sqcap (y \sqcup z)\big)$$

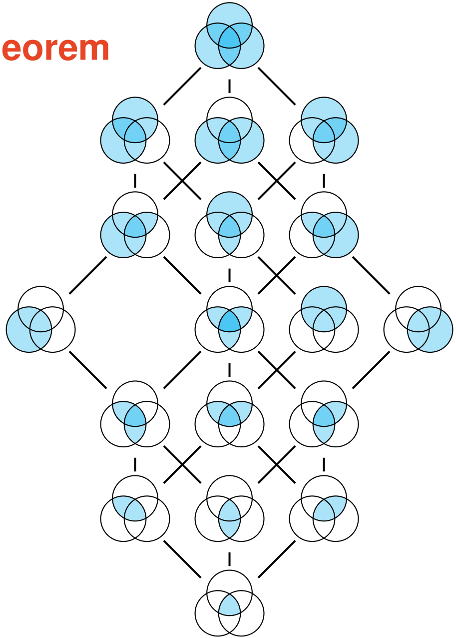
$$c = h\big(y \sqcap (x \sqcup z)\big) = h\big((x \sqcap y) \sqcup (y \sqcap z)\big)$$





Birkhoff’s representation theorem



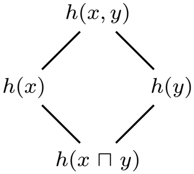


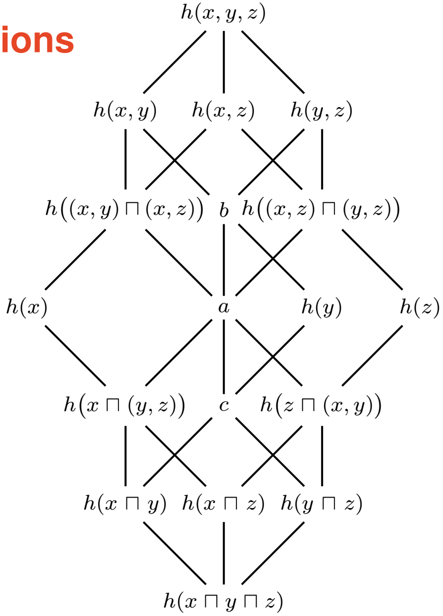
Partial information decompostions

$$a = h\big((x, y) \sqcap (x, z) \sqcap (y, z)\big)$$

$$b = h\big((x, y) \sqcap (y, z)\big)$$

$$c = h\big(y \sqcap (x, z)\big)$$





Questions?

- Paper is available on Arxiv
  – https://arxiv.org/abs/1909.12166/