PID

Jeffrey Emenheiser, RGJ, JPC

UC Davis

Beyond Shannon
UvA-IAS
6-8 May 2019

Distributions of Interest

Giant Bit

X_0	X_1	Y	Prob
0	0	0	1/2
1	1	1	1/2

Distributions of Interest

Giant Bit

X_0	X_1	Y	Prob
0	0	0	1/2
1	1	1	1/2

1bit Copy

X_0	X_1	Y	Prob
0	0	0	1/4
0	1	0	1/4
1	0	1	1/4
1	1	1	1/4

Distributions of Interest

Giant Bit

X_0	X_1	Y	Prob
0	0	0	1/2
1	1	1	1/2

1bit Copy

X_0	X_1	Y	Prob
0	0	0	1/4
0	1	0	1/4
1	0	1	1/4
1	1	1	1/4

XOR

X_0	X_1	Y	Prob
0	0	0	1/4
0	1	1	1/4
1	0	1	1/4
1	1	0	1/4

Distributions of Interest

Giant Bit

X_0	X_1	Y	Prob
0	0	0	1/2
1	1	1	1/2

1bit Copy

X_0	X_1	Y	Prob
0	0	0	1/4
0	1	0	1/4
1	0	1	1/4
1	1	1	1/4

XOR

X_0	X_1	Y	Prob
0	0	0	1/4
0	1	1	1/4
1	0	1	1/4
1	1	0	1/4

AND

X_0	X_1	Y	Prob
0	0	0	1/4
0	1	0	1/4
1	0	0	1/4
1	1	1	1/4

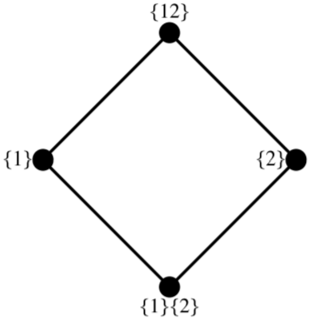
Distributions of Interest

Giant Bit	Redundant
1bit Copy	Unique
XOR	Synergistic
AND	???

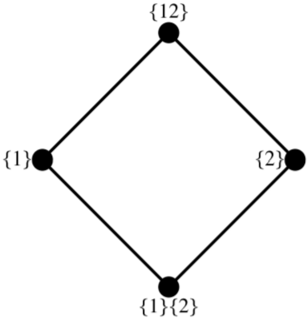
Williams and Beer 2010

"Our goal is to decompose the information that $\mathbf{X}$ provides about Y in terms of the partial information contributed either individually or jointly by various subsets of $\mathbf{X}$."

$I_\cap$



I_∂



$$I_\cap(\alpha \to Y) = \sum_{\beta \preceq \alpha} I_\partial(\beta \to Y)$$

Axioms

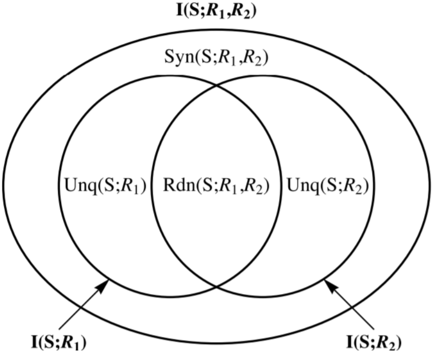
0. **Lattice** (L): Partial information is shared among subsets of source variables.

1. **Symmetry** (S): Redundancy values are invariant to permutations of sources.
2. **Self Redundancy** (SR): Single source redundancy values are equal to the mutual information
3. **Monotonicity** (M): Lattice provides a faithful ordering of redundancy values

4. **Local Positivity** (LP): Partial Information values are nonnegative
5. **Identity** (Id): Redundancy of two sources to their concatenation is the mutual information between the sources

Back to Shannon

$$I_\cap(\{01\} \to Y) = I(X_0 X_1 : Y)$$
$$I_\cap(\{0\} \to Y) = I(X_0 : Y)$$
$$I_\cap(\{1\} \to Y) = I(X_1 : Y)$$



I_{min}

Average pointwise minimum over target events.

$$I_\cap(\alpha \to Y) = \sum_{y \in Y} p(y) \min_{A \in \alpha} I(Y = y : A)$$

Williams and Beer. Nonnegative Decomposition of Multivariate Partial Information. *arxiv* 2010.

I_{proj}

minimum distance between target conditional distributions

$$p_{(x \searrow Y)}(Z) := \pi_{C_{\text{cl}}(\langle Y \rangle_Z)}(p(Z|x))$$

$$I_Z^\pi(X \searrow Y) := \sum_{z,x} p(z,x) \log \frac{p_{(x \searrow Y)}(z)}{p(z)}$$

$$I_{\text{red}}(Z;X,Y) := \min\left\{ I_Z^\pi(X \searrow Y), I_Z^\pi(Y \searrow X) \right\}$$

Harder; Salge; Polani. "Bivariate measure of redundant information". *Phys. Rev. E* 2013.

I_{BROJA}

(*) assumption: find distribution with no synergy in class

$I_\partial(0 \to Y)$ and $I_\partial(1 \to Y)$ depend only on $p(Y), p(X_0|Y), p(X_1|Y)$.

$I_\partial(i \to Y) = \min_{Q \in \Delta_P} I(X_i : Y | X_{\setminus i})$

Bertschinger; Rauh; Olbrich; Jost; Ay. "Quantifying unique information". *Entropy* 2014

Gács-Körner Common Information

$$I_\cap(\{0\}\{1\} \to Y) = I(V_{GK(X_0, X_1)} : Y)$$

Griffith; Chong; James; Ellison; Crutchfield. "Intersection information based on common randomness". *Entropy* 2014.

I_{ccs}

Categorize joint events by effect

$\Delta_s h(x)$	$\Delta_s h(y)$	$\Delta_s h(x,y)$	$-i(x;y;s)$	**Interpretation**
$+$	$+$	$+$	$+$	redundant information
$+$	$+$	$+$	$-$	synergistic information
$-$	$-$	$-$	$-$	redundant misinformation
$-$	$-$	$-$	$+$	synergistic misinformation
$+$	$+$	$-$	$\ldots$	?
$-$	$-$	$+$	$\ldots$	?
$+/-$	$-/+$	$\ldots$	$\ldots$	?

Ince. "Measuring multivariate redundant information with pointwise common change in surprisal". *Entropy* 2017

I_{dep}

Minimum change to maxent with addition of pairwise joint constraint

$$\mathrm{I}_{\mathrm{dep}}[i \to Y] = \min_{(\sigma_1,\sigma_2) \in E(X_i Y)} \left\{ \Delta^{\sigma_1}_{\sigma_2} \mathrm{I}[X_{0:n} : Y] \right\}$$

James; Emenheiser; Crutchfield. “Unique information via dependency constraints”. *J. Phys. A* 2019.

Pointwise modified coinformation

$$r^{+}_{min}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k \to t) = \min_{\boldsymbol{a}_i} i^{+}(\boldsymbol{a}_i \to t) = \min_{\boldsymbol{a}_i} h(\boldsymbol{a}_i)$$

$$r^{-}_{min}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k \to t) = \min_{\boldsymbol{a}_i} i^{-}(\boldsymbol{a}_i \to t) = \min_{\boldsymbol{a}_j} h(\boldsymbol{a}_j | t)$$

$$r(s_1, s_2 \to t) = r^{+}(s_1, s_2 \to t) - r^{-}(s_1, s_2 \to t).$$

Finn; Lizier. “Pointwise partial information decomposition using the specificity and ambiguity lattices”. *Entropy* 2018

$I_{\leftarrow}$

Secret key agreement rate, target may publicly communicate

$$I_\partial(\{i\} \to Y) = S(Y \to X_i | X_{\setminus i}) = \max_{B-A-X_i-X_{\setminus i}Y} I(X_i : A|B) - I(X_{\setminus i} : A|B)$$

James; Emenheiser; Crutchfield. "Unique information and secret key agreement". *Entropy* 2019.

Measures

I_{min}	average pointwise minimum over target events
I_{proj}	information geometry
I_{BROJA}	(*) assumption: find distribution with no synergy
$I_{\wedge}$	Gács Körner common information
I_{ccs}	categorize joint events by effect
I_{dep}	minimum change with pairwise constrinat
$I_{\pm}$	pointwise modification to coinformation
$I_{\leftarrow}$	secret key agreement rate

Distributions of Interest, advanced

2bit Copy

X_0	X_1	Y	Prob
0	0	0	1/4
0	1	1	1/4
1	0	2	1/4
1	1	3	1/4

Pointwise Unique

X_0	X_1	Y	Prob
0	1	1	1/4
0	2	2	1/4
1	0	1	1/4
2	0	2	1/4

XORCat

X_0	X_1	X_2	Y	Prob
0	0	0	0	1/4
0	1	1	1	1/4
1	0	1	2	1/4
1	1	0	3	1/4

Distributions of Interest

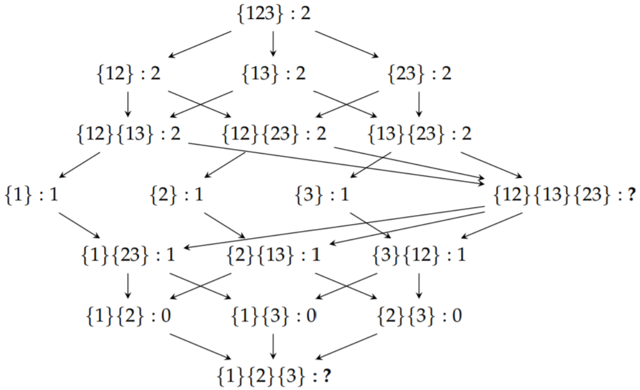
Giant Bit	Redundant
1bit Copy	Unique
XOR	Synergistic
AND	???
2bit Copy	Unique + Unique, or Redundant + Synergistic?
Pointwise Unique	Camels and Elephants!
XORCat	Multivariate Identity Axiom?

Comparisons

	I_∂	I_{min}	I_{proj}	I_{BROJA}	I_{ccs}	I_{dep}	$I_\pm$	I_{RR}	$I_\leftarrow$
AND	$X_0 X_1$	0.500 bit	0.500 bit	0.500 bit	0.292 bit	0.270 bit	0.750 bit	0.189 bit	0.500 bit
	$X_0 \setminus X_1$	0.000 bit	0.000 bit	0.000 bit	0.208 bit	0.230 bit	−0.250 bit	0.311 bit	0.000 bit
	$X_1 \setminus X_0$	0.000 bit	0.000 bit	0.000 bit	0.208 bit	0.230 bit	−0.250 bit	0.311 bit	0.000 bit
	$X_0 \cdot X_1$	0.311 bit	0.311 bit	0.311 bit	0.104 bit	0.082 bit	0.561 bit	0.000 bit	0.311 bit
PNT. UNQ.	$X_0 X_1$	0.500 bit	0.500 bit	0.500 bit	0.000 bit	0.250 bit	0.000 bit	0.333 bit	0.500 bit
	$X_0 \setminus X_1$	0.000 bit	0.000 bit	0.000 bit	0.500 bit	0.250 bit	0.500 bit	0.167 bit	0.000 bit
	$X_1 \setminus X_0$	0.000 bit	0.000 bit	0.000 bit	0.500 bit	0.250 bit	0.500 bit	0.167 bit	0.000 bit
	$X_0 \cdot X_1$	0.500 bit	0.500 bit	0.500 bit	0.000 bit	0.250 bit	0.000 bit	0.333 bit	0.500 bit
TWO BIT	$X_0 X_1$	1.000 bit	0.000 bit	0.000 bit	0.000 bit	0.000 bit	1.000 bit	0.000 bit	0.000 bit
	$X_0 \setminus X_1$	0.000 bit	1.000 bit	1.000 bit	1.000 bit	1.000 bit	0.000 bit	1.000 bit	1.000 bit
	$X_1 \setminus X_0$	0.000 bit	1.000 bit	1.000 bit	1.000 bit	1.000 bit	0.000 bit	1.000 bit	1.000 bit
	$X_0 \cdot X_1$	1.000 bit	0.000 bit	0.000 bit	0.000 bit	0.000 bit	1.000 bit	0.000 bit	0.000 bit

Discussion: Three sources

Three ways to split four events: XORCat



Rauh, "Secret sharing and shared information". *Entropy* 2017.

Discussion: Directionality

Pointwise Unique, the camel

X_0	X_1	Y	Prob
0	1	1	1/4
0	2	2	1/4
1	0	1	1/4
2	0	2	1/4

Pointwise Unique, the elephant

X_0	X_1	Y	Prob
0	1	1	1/4
1	0	1	1/4
0	2	2	1/4
2	0	2	1/4

Discussion: Resource theory

What is a resource?
What are free operations?

$$\text{2bitCopy} \sim \text{1bitCopy}_0 \otimes \text{1bitCopy}_1$$

$$Y = f(\text{2bitCopy}(X_0, X_1))$$