Information Decomposition

Conor Finn* and Joseph Lizier†

Centre for Complex System, Faculty of Engineering & IT *†

Data 61, CSIRO*

THE UNIVERSITY OF SYDNEY

Abstract

What are the distinct ways in which a set of predictor variables can provide information about a target variable? When does a variable provide unique information? When do variables provide shared information? And when do variables combine to provide synergistic information? Information decomposition aims to address these questions.

The Problem

Shannon inequalities:

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

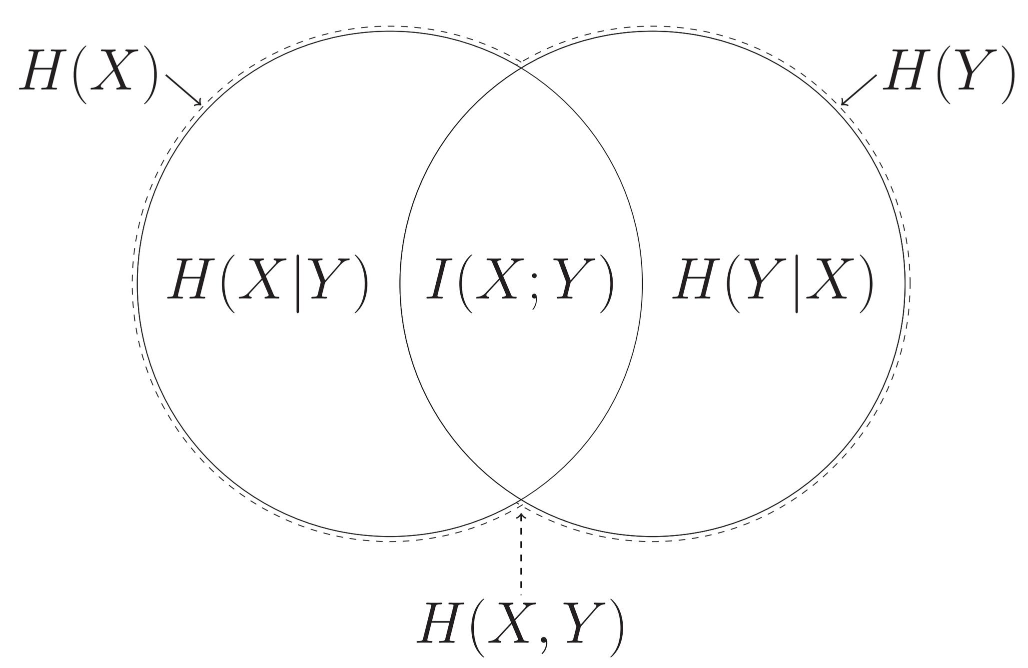


Figure: A Venn diagram can be used to represent the relationship between the entropy of two variables—this depiction is justified by the Shannon inequalities.

No inequalities require the multivariate mutual information

$$I(X;Y;Z) = I(X;Y) - I(X;Y|Z)$$

to be non-negative. This quantity has "*no intuitive meaning*" [1].

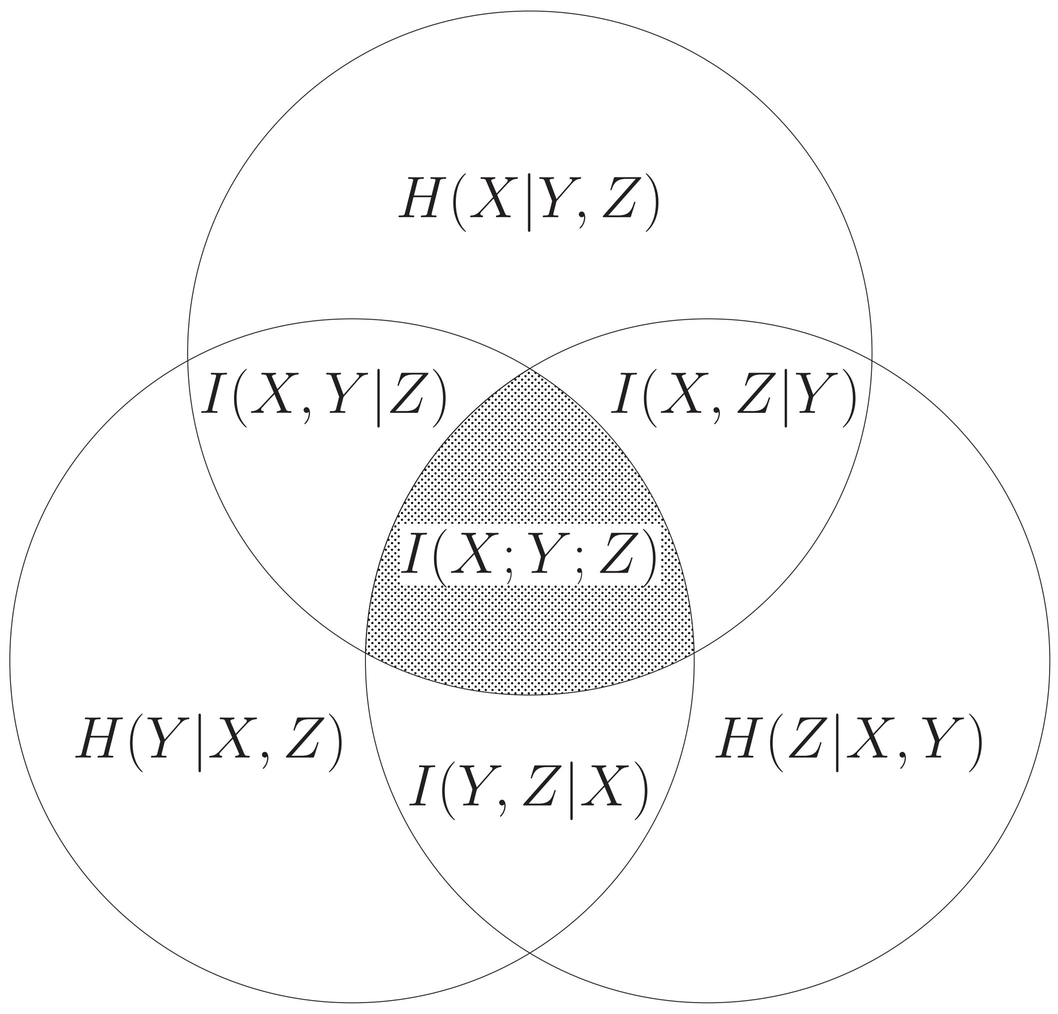


Figure: A Venn diagram does not provide a valid representation of the relationship between the entropy of three variables as $I(X;Y;Z)$ is not non-negative [2].

References

[1] Imre Csiszar and János Körner. *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Academic Press, Inc., 1981.

[2] David JC MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge university press, 2003.

[3] Paul L Williams and Randall D Beer. Nonnegative decomposition of multivariate information. *arXiv preprint arXiv:1004.2515*, 2010.

[4] Conor Finn and Joseph T Lizier. Pointwise partial information decomposition using the specificity and ambiguity lattices. *Entropy*, 20(4), 2018.

[5] Joseph T Lizier, Nils Bertschinger, Jürgen Jost and Michael Wibral. Information decomposition of target effects from multi-source interactions: Perspectives on previous, current and future work. *Entropy*, 20(4), 2018.

Unique, Redundant and Synergistic Information

Consider the information provided by variables S_1 and S_2 about T:

- Unique information

$$\left.\begin{aligned} I(S_1;T) &= 1 \\ I(S_2;T) &= 0 \\ I\big((S_1,S_2);T\big) &= 1 \end{aligned}\right\} U(S_1 \backslash S_2 \to T) = 1$$

p	s_1	s_2	t
1/4	0	0	0
1/4	0	1	0
1/4	1	0	1
1/4	1	1	1

- Redundant information

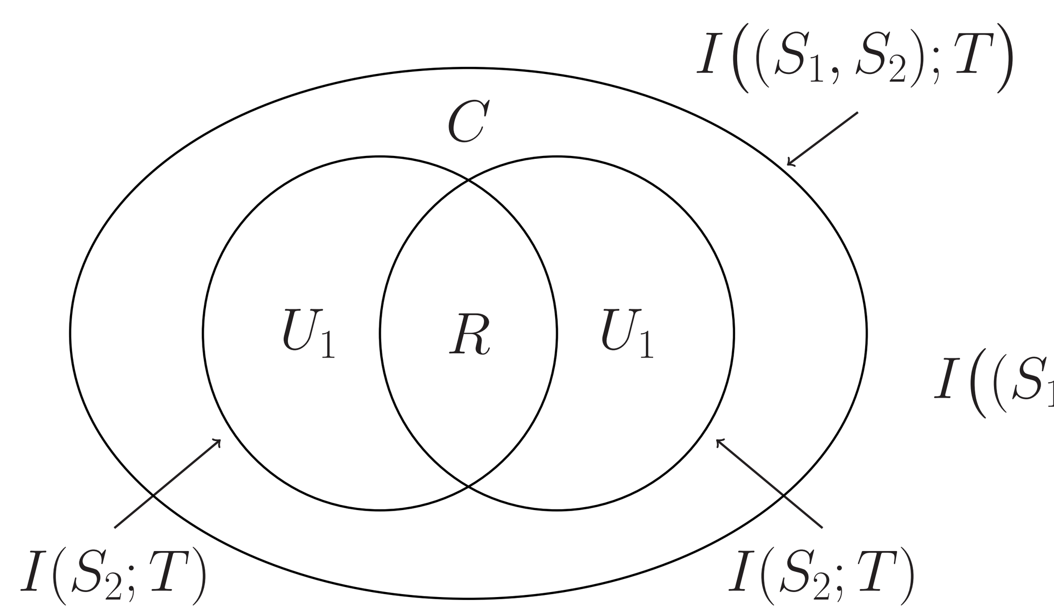
$$\left.\begin{aligned} I(S_1;T) &= 1 \\ I(S_2;T) &= 1 \\ I\big((S_1,S_2);T\big) &= 1 \end{aligned}\right\} R(S_1, S_2 \to T) = 1$$

p	s_1	s_2	t
1/2	0	0	0
1/2	1	1	1

- Synergistic information (e.g. XOR logic gate)

$$\left.\begin{aligned} I(S_1;T) &= 0 \\ I(S_2;T) &= 0 \\ I\big((S_1,S_2);T\big) &= 1 \end{aligned}\right\} C(S_1, S_2 \to T) = 1$$

p	s_1	s_2	t
1/4	0	0	0
1/4	0	1	1
1/4	1	0	1
1/4	1	1	0



$$\begin{aligned} I(S_1;T) &= U_1 + R \\ I(S_2;T) &= U_2 + R \\ I\big((S_1,S_2);T\big) &= U_1 + U_2 \\ &\quad + R + C \end{aligned}$$

Figure: The relationship between the unique, redundant, and synergistic information.

Multivariate mutual information conflates redundancy and synergy

$$I(X;Y;Z) = R(Y,Z \to X) - C(Y,Z \to X)$$

Partial Information Decomposition [3]

Redundant information should be analogous to set intersection:

$$\left.\begin{array}{l} \text{1. Commutativity} \\ \text{2. Monotonically decreasing} \\ \text{3. Idempotency} \end{array}\right\} \implies \text{Redundancy lattice}$$

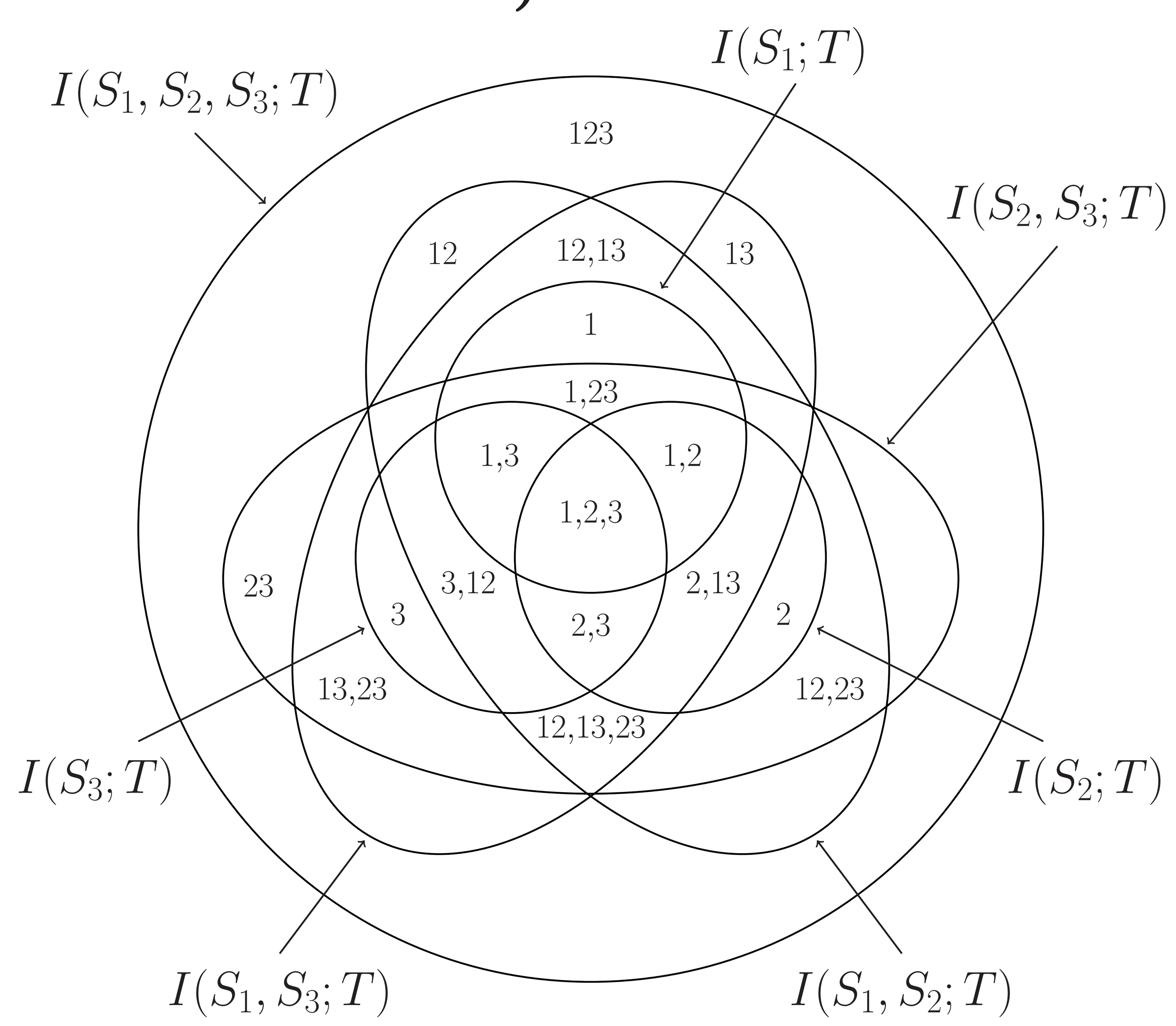


Figure: The redundancy lattice enables us to decompose multivariate information.

However, defining multivariate redundancy has been a contentious problem:

- Much ongoing research, e.g. see the special issue in *Entropy* [5].
- Our approach provides satisfying results, e.g. target chain rule [4].

conor.finn@sydney.edu.au