Analytically deriving partial information decomposition for affine systems of stable and convolution-closed distributions

Carnegie Mellon University

THE FOR ALL LAB

Chaitanya Goswami and Amanda Merkley

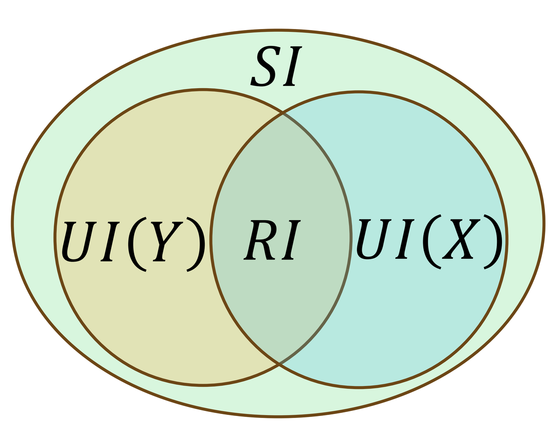
Electrical & Computer Engineering, Carnegie Mellon University, Pittsburgh, PA

Motivation

- Partial Information Decomposition (PID) is a promising tool for analyzing the information about a **target variable M** contained in the interactions between two **source variables X** and **Y** [1].
- PID is being used in *neuroscience, multimodal learning, fairness in AI, and economics*.
- Significant progress has been made in *numerically* calculating PID, but the only known *analytical solutions* remain for jointly Gaussian systems [2].
- We present new analytical PID for a much larger set of distribution families/systems, where the source variables have an affine dependence on the target variable.
- We also provide an analytical upper bound for systems with non-affine dependence on the message.

PID Background

- PID expresses mutual information under distribution $P(M,X,Y)$ as unique (UI), redundant (RI), and synergistic (SI) information terms



PID framework: 3 equations + 4 unknowns
$I_P(M;X,Y) = UI(X) + UI(Y) + RI + SI$
$I_P(M;X) = UI(X) + RI$
$I_P(M;Y) = UI(Y) + RI$

- BROJA-PID [1] is one measure of unique information that solves PID:

$$UI(Y) = \min_{Q \in \Delta_P} I_Q(M;Y|X)$$

Set of all distributions $Q(M,X,Y)$ s.t. marginals are fixed to $P(M,X)$ and $P(M,Y)$

- Our results are applicable for broader set of PID definitions, namely Blackwellian PIDs, and PID definitions satisfying assumption (*) of [1].

Analytical Solutions for PID of affine stable and conv-closed systems

Systems for which PID is analyzed:

(i) $X|M$ and $Y|M\sim$ stable or convolution-closed distribution,

(ii) X and Y have an affine dependence on M

Brief Proof Sketch

Key Property:

For $X_1|M, X_2|M\sim$ stable or convolution-closed distribution, and a linear/affine dependence on M: $I(M;X_1+X_2) = I(M;[X_1,X_2])$, or X_1+X_2 acts as a "sufficient statistic" for $[X_1,X_2]$

1. Decompose the more "informative" X *as* $Y'+X'$, where $Y' =^d Y$.
2. From key property: X is equivalent to $[Y',X']$.
3. Comparing Y and $[Y',X'] \to Y$ has zero UI about M as $Y' =^d Y$, hence UI(Y)=0.
4. Rest of the PID terms obtained by solving the three linear equations.

Analytical PID: Example Affine Systems

Example systems in which $UI(Y) = 0$. Such systems depend on M in an affine manner.

Ex. 1: Poisson
- $M\sim P(M)$. Let $a > b > 0$
- $P(X|M) = Poisson(aM)$
- $P(Y|M) = Poisson(bM)$

Ex. 2: General Gaussian
- $M\sim P(M)$. Let $|a|/b > |c|/d$
- $P(X|M) = N(aM, b\sigma^2)$
- $P(Y|M) = N(cM, d\sigma^2)$

Generalizes [2] since $P(M)$ is arbitrary

Ex. 3: Exponential + Geometric
- $M\sim P(M)$ supported on $\mathbb{R}^+$
- $P(X|M) = Exponential(M)$
- $P(Y|M) = Geometric(M/(\tau+M))$

Ex. 4: Uniform
- $M\sim P(M)$ supported on $\mathbb{R}^+$
- $P(X|M) = Uniform(0,aM)$
- $P(Y|M) = Uniform(0,bM)$

Here, *both* $UI(X) = 0$ and $UI(Y) = 0$

Analytical Upper Bound: Non-Affine Systems

We also provide analytically solvable upper-bound to calculate the UI for certain non-affine convolution-closed systems

- Split X and Y:

$$X = X' + Y'' + n_X$$
$$Y = Y' + n_Y$$

Noise terms

where Y' and Y'' have the same distribution

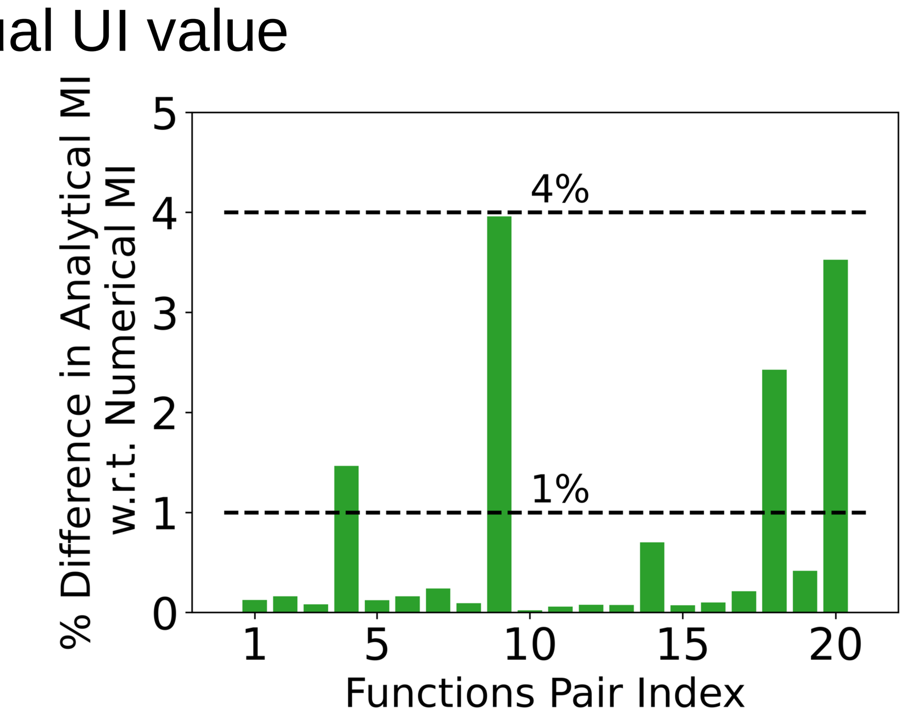
- By data processing inequality,

$$I_{\bar{Q}}(M;X',Y'',Y',n_X,n_Y) \geq I_Q(M;X,Y) \quad \text{for } Q \in \Delta_P$$

Can be analytically minimized

Ex. 5: The upper-bound is tight for various non-affine Poisson system

Tested 20 non-affine Poisson systems. The proposed upper-bound was within **4%** of the actual UI value



Distribution Families

1. Stable distributions
- X is stable if, for 2 independent copies, $X_1 + X_2$ is distributed as a scaled and translated X
- Continuous stable described by stability, skew, scale, and location: $f(\alpha,\beta_i,\gamma_i,\mu_i)$, e.g., Gaussian, Cauchy
- Discrete stable described by rate and exponent: $P(\nu,\tau)$, e.g., Poisson
- Multivariate extensions exist

2. Convolution-closed distributions
- $\mathcal{F}$ is a set of distributions with elements $X_1\sim f(\delta_1)$ and $X_2\sim f(\delta_2)$, and is convolution-closed if $(X_1+X_2)\sim f(\delta_1+\delta_2)$. Can be thought of as an extended exponential-family distribution. E.g., Gamma, Exponential, Binomial

Conclusion

- We show that many systems (stable and convolution-closed) exhibit a "zero unique information" property, generalizing the previously only known Gaussian result [2], providing valuable theoretical insight into PID computation.
- Our results solve PID for systems with "tricky" fat-tailed distributions, e.g., Cauchy, and provide a diverse benchmark against which numerical estimators can be tested.
- Promising future work in neuroscientific applications modeled by Poisson, binomial, and Cauchy systems.

References

[1] N. Bertschinger et al., Entropy, 2014.
[2] A. B. Barrett, Phys. Rev. E., 2015.

Acknowledgements

We thank Pulkit Grover, Praveen Venkatesh, Alireza Chamanzar, and Neil Ashim Mehta for insightful discussions. This research was supported by funding from the National Science Foundation.