PI Measure: A Set-Theoretic View of Partial Information

2020 Spring CS258 Group Project Presentation

Litao Zhou, Hongbo Yang, Shiwen Dong

IEEE Class F1803016
School of Electronic Information and Electrical Engineering
Shanghai Jiao Tong University

August 13, 2020

Background

I-Measure

A theory which establishes a one-to-one correspondence between Shannon's information measure and set theory in full generality.

$$\begin{array}{ccc} H/I \leftrightarrow \mu^* & , & \leftrightarrow \cup \\ ; \leftrightarrow \cap & | & \leftrightarrow - \end{array} \tag{1}$$

Non-negative Decomposition of Multivariate Information

A measure of *redundancy information* that a set of sources provides about a given variable.
Shannon information can be decomposed into atoms of non-negative *partial information*.

Our Goal: Using the I-Measure methodology to study the redundancy measure, forming a set-theoretic view of partial information.

Definition

Definition 1 (Field)

The field $\mathcal{F}_n$ generated by sets $\tilde{X}_1, \tilde{X}_2, \cdots, \tilde{X}_n$ is the collection of sets which can be obtained by any sequence of usual set operations (union, intersection, complement, and difference) on $\tilde{X}_1, \tilde{X}_2, \cdots, \tilde{X}_n$

Definition 2 (Atom)

The atoms of $\mathcal{F}_n$ are sets of the form $\cap_{i=1}^{n} Y_i$, where Y_i is either$\tilde{X}_i$ or $\tilde{X}_i^c$, the complement of $\tilde{X}_i$

Definition 3 (Signed Measure)

A real function μ defined on $\mathcal{F}_n$ is called a signed measure if it is set-additive, i.e., for disjoint A and B in $\mathcal{F}_n$

$$\mu(A \cup B) = \mu(A) + \mu(B) \tag{2}$$

Measure Theorem

Use X_G to denote $(X_i, i \in G)$ and $\tilde{X}_G$ to denote $\cup_{i\in G}\tilde{X}_i$ for any nonempty subset G of $\mathcal{N}_n = \{1, 2, \cdots, n\}$

Theorem 4 (Completeness)

Let

$$\mathcal{B} = \left\{ \tilde{X}_G : G \textit{ is a nonempty subset of } \mathcal{N}_n \right\} \tag{3}$$

Then a signed measure μ on $\mathcal{F}_n$ is completely specified by $\{\mu(B), B \in \mathcal{B}\}$ which can be any set of real numbers.

With the terminologies of measure theory above, we can construct a correspondence between set theory and Shannon Information

Correspondence

Build Correspondence

We now construct the I-Measure μ^* on $\mathcal{F}_n$ using Theorem 4 by defining

$$\mu^*\left(\tilde{X}_G\right) = H(X_G) \tag{4}$$

for all nonempty subsets G of $\mathcal{N}_n$.
The following must hold for all (not necessarily disjoint) subsets G, G', G'' of $\mathcal{N}_n$:

$$\mu^*\left(\tilde{X}_G \cap \tilde{X}_{G'} - \tilde{X}_{G''}\right) = I(X_G; X_{G'} | X_{G''}) \tag{5}$$

Information Diagrams

With the correspondence above, it is reasonable to use an *information diagram*, which is a variation of Venn diagram, to represent the relationship between Shannon's information measures.

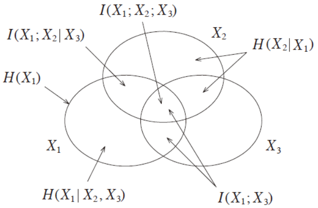


Figure: Generic Information Diagram for X_1, X_2, X_3

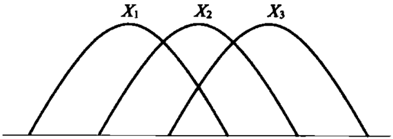
Applications

Information Diagrams for Zero Measure

In Markov Chain $X_1 \to X_2 \to X_3$, since

$$\mu^* \left(\tilde{X}_1 \cap \tilde{X}_2^c \cap \tilde{X}_3 \right) = I(X_1; X_3 | X_2) = 0 \tag{6}$$

We don't have to plot that region out in the information diagram, making the diagram simpler and more useful when reasoning about Shannon information measures.



In a word, with the set-theoretic view of information diagram, we can discover many useful properties.

Redundancy Measure

Let $A_1, A_2, \ldots, A_k$ be nonempty and potentially overlapping subsets of R, which we call sources. We introduce $I_{\min}$ to measure the information that all sources provide about S.

Definition 5 (Redundancy Measure)

$$I_{\min}(S; \{\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k\}) = \sum_s p(s) \min_{\mathbf{A}_i} I(S = s; \mathbf{A}_i) \tag{7}$$

where the domain of $I_{\min}$ $\mathcal{A}(\mathbf{R}) = \{\alpha \in \mathcal{P}_1(\mathcal{P}_1(\mathbf{R})) : \forall \mathbf{A}_i, \mathbf{A}_j \in \alpha, \mathbf{A}_i \notin \mathbf{A}_j\}$

Properties of Redundancy Measure

- $I_{\min}$ is non-negative
- $I_{\min}$ is less than or equal to $I(S; \mathbf{A}_i)$ for all $\mathbf{A}_i$'s
- For a given source A, the amount of information redundant with $\mathbf{A}$ is maximal for $I_{\min}(S; \{\mathbf{A}\}) = I(S; \mathbf{A})$.

Relations between Redundancy Measures

Note that $I_{\min}$ can measure redundancy between collections of sources like $\{\{R_1\}, \{R_2, R_3\}\}$ denoted as $\{1\}\{23\}$. Certain relations exists between measures.

Redundancy Order

We can define a partial order over the elements of $A(R)$ such that one is considered to precede another if and only if the latter provides any redundant information that the former provides. Formally,

$$\forall \alpha, \beta \in \mathcal{A}(\mathbf{R}),$$
$$\alpha \preccurlyeq \beta \Leftrightarrow \forall \mathbf{B} \in \beta, \exists \mathbf{A} \in \alpha, \mathbf{A} \subseteq \mathbf{B} \tag{8}$$

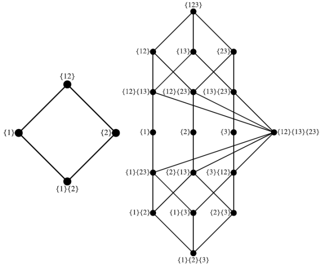


Figure: Redundancy lattice for (A) 3 and (B) 4 variables.

Partial Information Decomposition

We would like to decompose redundancy measure into non-intersecting fragments. For a collection of sources $\alpha \in \mathcal{A}(\mathbf{R})$, the PI-function denoted Π_{R}, is defined implicitly by

$$I_{\min}(S;\alpha) = \sum_{\beta \preccurlyeq \alpha} \Pi_{\mathrm{R}}(S;\beta) \tag{9}$$

Properties of Redundancy Measure

- II_{R} is non-negative
- II_{R} can be calculated recursively.
- The relationship between these measures can be shown in a partial information (PI) diagram, which we will discuss later.

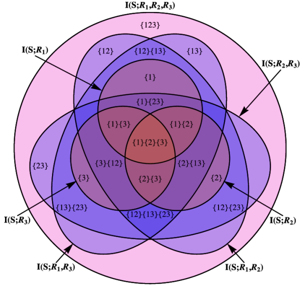


Figure: PI diagram for 4 variables.

Formulation

The sets used to generate fields, and their correspondence to redundancy measures are defined as follows.

Basic Sets Correspondence

Denote the source variables domain as

$$\Gamma = \{\beta \in \mathcal{P}_1(\mathbf{R}) : \forall \mathbf{A}_i, \mathbf{A}_j \in \beta, \mathbf{A}_i \not\subseteq \mathbf{A}_j\} \tag{10}$$

For every random variable collection in the domain Γ, its self-redundancy is defined as $I_{\min}(S; \{A_i, \ldots, A_j\})$. We use $\left\{\tilde{A}_i, \ldots, \tilde{A}_j\right\}$, or abbreviation $\tilde{A}_{i\ldots j}$ to represent the basic set its self redundancy corresponds to.

Formulation

To build fields, define the correspondence rules for operations on basic sets.

Field Correspondence

- For every basic sets $\tilde{A}_{i\ldots j}$,

$$\mu^*\left(\tilde{A}_{i\ldots j}\right) = I_{\min}\left(S; \{A_i, \ldots, A_j\}\right). \tag{11}$$

- For any two sets αin the field, $\mu^*(\alpha^c) = \sum_{\alpha \prec \gamma} \Pi_{\mathbf{R}}(S;\gamma)$
- For any two sets α and β in the field, $\mu^*(\alpha \cap \beta) = \sum_{\gamma \preceq \alpha \text{ and } \gamma \preceq \beta} \Pi_{\mathbf{R}}(S;\gamma)$
- For any two sets α and β, $\mu^*(\alpha \cup \beta) = \sum_{\gamma \preceq \alpha \text{ or } \gamma \preceq \beta} \Pi_{\mathbf{R}}(S;\gamma)$
- For any two sets α and β, $\mu^*(\alpha - \beta) = \sum_{\gamma \preceq \alpha \text{ and } \beta \prec \gamma} \Pi_{\mathbf{R}}(S;\gamma)$

This definition is based on the partial order given in the redundancy lattice. However, such definition is not very useful in practice.

Simplification

We try to eliminate redundant fields in order to avoid the exponential increase of fields.

Example of Two Source Variables

To fix ideas, we take the multivariate information of three variables $I(S; \mathbf{A}_1, \mathbf{A}_2)$ as example, note that $\left\{\tilde{A}_1\right\} \subseteq \left\{\tilde{A}_1, \tilde{A}_2\right\}$. It follows that

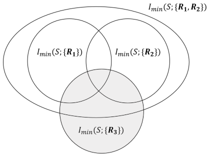
$$\begin{aligned}\mu^*\left(\left\{\tilde{A}_1\right\} \cap \left\{\tilde{A}_1, \tilde{A}_2\right\}^C\right) &= \mu^*\left(\left\{\tilde{A}_1\right\}\right) - \mu^*\left(\left\{\tilde{A}_1\right\} \cap \left\{\tilde{A}_1, \tilde{A}_2\right\}\right) \\ &= I_{\min}\left(S; \{\mathbf{A}_1\}\right) - I_{\min}\left(S; \{\mathbf{A}_1\}, \{\mathbf{A}_1, \mathbf{A}_2\}\right) = 0\end{aligned} \tag{12}$$

Similarly, with the observation that $\left\{\tilde{A}_2\right\} \subseteq \left\{\tilde{A}_1, \tilde{A}_2\right\}$, we have $\mu^*\left(\left\{\tilde{A}_2\right\} \cap \left\{\tilde{A}_1, \tilde{A}_2\right\}^C\right) = 0$. To make our diagram more legible, we can erase these regions out.

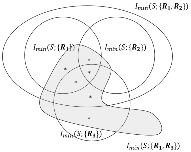
Simplification

Theorem 6 (Field Elimination)

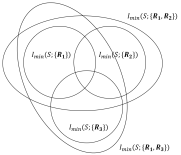
For any two basic sets $\tilde{A}_{i\ldots j}, \tilde{A}_{i'\ldots j'} \in \Gamma$, if $\tilde{A}_{i\ldots j} \subseteq \tilde{A}_{i'\ldots j'}$, then the measure of any set intersection with $\tilde{A}_{i\ldots j} \cap \tilde{A}^C_{i'\ldots j'}$ is zero.



(a) $\tilde{A}_3$ produces no extra redundant fields



(b) Adding Basic Set $\tilde{A}_{12}$ produces extra redundant fields as has been starred



(c) After Elimination

Figure: Field Elimination Example for 3 Source Variables

Simplification

We try to find alternative, useful expressions of fields by means of set theory.

Observation on Atoms

- The measure of *atoms* in $\mathcal{F}_n$ is a perfect matching to the partial information $\Pi_{\mathbf{R}}$.
- All redundancy measures can be written as sum of non-negative $\Pi_{\mathbf{R}}$s.
- All atoms can be written as sets of $\cap_{i \in \mathcal{N}_n} \tilde{Y}_i (\tilde{Y}_i = \tilde{A}_i$ or $\tilde{A}_i^c)$, where the index set $\mathcal{N}_n = \mathcal{P}_1 \left(\{1, 2, \ldots, n\}\right)$

In order to completely write measure expressions, we only need to study two operations (complement and intersection) on basic sets $Y_{i \in \mathcal{N}_n}$.

Simplification

$$\mu^*\left(\tilde{A}_{N_1}\cap\ldots\cap\tilde{A}_{N_k}\cap\tilde{A}_{M_1}^C\cap\ldots\cap\tilde{A}_{M_j}^C\right) \tag{13}$$

$$=\mu^*\left(\tilde{A}_{N_1}\cap\ldots\cap\tilde{A}_{N_k}\right)-\sum_{i=1}^{k}\mu^*\left(\tilde{A}_{N_1}\cap\ldots\cap\tilde{A}_{N_k}\cap\tilde{A}_{M_1}\cap\ldots\cap\tilde{A}_{M_i}\right) \tag{14}$$

Note that in Equation 14, all the sets being measured are intersections of basic sets. By applying Equation 11, we can write them out in the form of redundancy measure $I_{\min}$.

Conclusion

- The measure of any atom can be written as a ± 1 combination of some redundancy measures $I_{\min}$.
- The measure of any field can be written as a linear combination of redundancy measures $I_{\min}$.

Application

An application of our simplification is to calculate partial information in two ways.

Method 1

On the one hand, we can using the definition of Π_{R} to calculate, applying Eq.15.

$$\Pi_{\mathbf{R}}(S;\alpha) = I_{\min}(S;\alpha) - \sum_{s} p(s) \max_{\beta\in\alpha^{-}} \min_{\mathbf{B}\in\beta} I(S=s;\mathbf{B}) \tag{15}$$

Method 2

On the other hand, we can calculate Π_{R} by converting it to the linear combination of $I_{\min}$.

Example with Four Variables

Example

Consider three equiprobable cases of S as follows:

$$\begin{aligned}\mathbf{S} = 0 &: (\mathbf{R}_1 = 0, \mathbf{R}_2 = 0, \mathbf{R}_3 = 0)\\ \mathbf{S} = 1 &: (\mathbf{R}_1 = 0, \mathbf{R}_2 = 1, \mathbf{R}_3 = 1)\\ \mathbf{S} = 2 &: (\mathbf{R}_1 = 1, \mathbf{R}_2 = 0, \mathbf{R}_3 = 0)\end{aligned}$$

We want to calculate the value of $\Pi_{\mathrm{R}}(S; \{2\}, \{13\})$.

Example with Four Variables

Method 1

Firstly, we can applying Eq.15 to solve the problem:

$$\Pi_{\mathbf{R}}(S;\{2\},\{13\}) = I_{\min}(S;\{2\},\{13\}) - \sum_{s} p(s) \max_{\beta\in\alpha^{-}} \min_{\mathbf{B}\in\beta} I(S=s;\mathbf{B}) \quad (16)$$

where $\beta \in \{\{\{1\}\{2\}\},\{\{2\}\{3\}\}\}$. Then we just need to expand the formula to calculate. We have

$$I_{\min}(S;\{2\},\{13\}) = \log 3 - \frac{2}{3}\log 2 \quad (17)$$

$$\sum_{s} p(s) \max_{\beta\in\alpha^{-}} \min_{\mathbf{B}\in\beta} I(S=s;\mathbf{B}) = \log 3 - \frac{2}{3}\log 2 \quad (18)$$

So $\Pi_{\mathrm{R}}(S;\{2\},\{13\})$=0.

Example with Four Variables

Method 2

based on the observations in FIG.3, we can express the field using Eq.19:

$$\mu^* \left(\tilde{A}_{i \dots j} \cap \dots \cap \tilde{A}_{m \dots n}\right) = I_{\min}\left(S; \{\mathbf{A}_i, \dots, \mathbf{A}_j\}, \dots, \{\mathbf{A}_m, \dots, \mathbf{A}_n\}\right) \quad (19)$$

$$\begin{aligned}
&\mu^* \left(\left\{\tilde{R}_1\right\}^C \cap \left\{\tilde{R}_2\right\} \cap \left\{\tilde{R}_3\right\}^C \cap \left\{\tilde{R}_{12}\right\} \cap \left\{\tilde{R}_{13}\right\} \cap \left\{\tilde{R}_{23}\right\} \cap \left\{\tilde{R}_{123}\right\}\right) \\
&= I_{\min}\left(S; \{2\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\right) \\
&\quad - I_{\min}\left(S; \{1\}, \{2\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\right) \\
&\quad - I_{\min}\left(S; \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\right) \\
&\quad + I_{\min}\left(S; \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\right) \\
&= I_{\min}\left(S; \{2\}, \{1,3\}\right) - I_{\min}\left(S; \{1\}, \{2\}\right) - I_{\min}\left(S; \{2\}, \{3\}\right) \\
&\quad + I_{\min}\left(S; \{1\}, \{2\}, \{3\}\right)
\end{aligned}$$

Example with Four Variables

Method 2

We can also list all $I_{\min}$ of the sets which have the partial order relation with $\{\{2\},\{13\}\}$ to express $\Pi_{\mathbf{R}}$ using the linear combination of $I_{\min}$.:

$$I_{\min}(S;\{2\},\{13\}) = \Pi_{\mathbf{R}}(S;\{2\},\{3\}) + \Pi_{\mathbf{R}}(S;\{1\},\{2\}) + \Pi_{\mathbf{R}}(S;\{2\},\{13\}) + I_{\min}(S;\{1\},\{2\},\{3\}) \tag{20}$$

$$I_{\min}(S;\{2\},\{3\}) = \Pi_{\mathbf{R}}(S;\{2\},\{3\}) + I_{\min}(S;\{1\},\{2\},\{3\}) \tag{21}$$

$$I_{\min}(S;\{1\},\{2\}) = \Pi_{\mathbf{R}}(S;\{1\},\{2\}) + I_{\min}(S;\{1\},\{2\},\{3\}) \tag{22}$$

Combine Eq.20,21,22, then we have the save result as before.
Calculating by the linear combination of $I_{\min}$, we can also have the same outcome of $\Pi_{\mathrm{R}}(S;\{2\},\{13\})$.

Summary

Previous Work

- Correspondence between Shannon information and set theory
- Redundancy measure and partial information decomposition methods for multivariate information

Achieved Goals

- Formalize the correspondence between redundancy measure and set theory
- Prove the completeness of the correspondence relation
- Simplify the correspondence expressions
- Propose a top-down strategy to calculate redundancy measure

Bibliography

[1] Paul L. Williams, Randall D. Beer.
Nonnegative Decomposition of Multivariate Information[J]
arXiv preprint arXiv:1004.2515, 2010.

[2] Yeung R W.
Information theory and network coding[M]
Springer Science & Business Media, 2008.