

An Axiomatic Account of Similarity

Supplementary Material: Formal Proofs

Enrique Amigó, Julio Gonzalo, Fernando Giner and Felisa Verdejo

enrique,julio,felisa@lsi.uned.es
National Distance University
C Juan del Rosal 16
Madrid 28040

ACM Reference format:

Enrique Amigó, Julio Gonzalo, Fernando Giner and Felisa Verdejo. 2017. An Axiomatic Account of Similarity Supplementary Material: Formal Proofs. In *Proceedings of SIGIR'17 Workshop on Axiomatic Thinking for Information Retrieval and Related Tasks, Tokyo Japan, August 2017 (ATIR)*, 3 pages. DOI:

PROOF 1. “The Tversky’s monotonicity axiom is not compatible with the dependency constraint”.

The Monotonicity Axiom states that if:

$$X \cap Y \supset X' \cap Y' \wedge X \setminus Y \subset X' \setminus Y' \wedge Y \setminus X \subset Y' \setminus X'$$

then

$$Sim(X, Y) \leq Sim(X', Y')$$

In the case of the dependency constraint, according to its definition, if Z and Z' are disjoint, then:

$$X \cap Y = (X \cup Z) \cap (Y \cup Z')$$

$$X \setminus Y \subset (X \cup Z) \setminus (Y \cup Z')$$

$$Y \setminus X \subset (X \cup Z) \setminus (Y \cup Z')$$

Therefore, according to the Monotonicity axiom:

$$Sim(X, Y) \leq Sim(X \cup Z, Y \cup Z')$$

which contradicts the dependency constraint situation.

PROOF 2. SIM is equivalent to say that there exists a positive similarity increase when both the single information quantity increase and their sum are higher than the join information quantity.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ATIR, Tokyo Japan

© 2017 ACM. 978-x-xxxx-xxxx-x/YY/MM...\$15.00
DOI:

$$\begin{aligned} \Delta I(X) + \Delta I(Y) &\geq I(X \cup Y) \equiv \\ &\equiv \log\left(\frac{1}{P(X')}\right) + \log\left(\frac{1}{P(Y')}\right) - \log\left(\frac{1}{P(X)}\right) + \log\left(\frac{1}{P(Y)}\right) \geq \\ &\log\left(\frac{1}{P(X' \cup Y')}\right) - \log\left(\frac{1}{P(X \cup Y)}\right) \\ &\equiv \log\left(\frac{P(X)P(Y)}{P(X')P(Y')}\right) \geq \log\left(\frac{P(X \cup Y)}{P(X' \cup Y')}\right) \\ &\equiv \frac{P(X)P(Y)}{P(X')P(Y')} \geq \frac{P(X \cup Y)}{P(X' \cup Y')} \equiv \frac{P(X' \cup Y')}{P(X')P(Y')} \geq \frac{P(X \cup Y)}{P(X)P(Y)} \\ &\equiv PMI(X', Y') \geq PMI(X, Y) \equiv \Delta PMI(X, Y) \geq 0 \end{aligned}$$

The other two conditions are also equivalent.

$$\begin{aligned} \Delta I(X) &\geq I(X \cup Y) \equiv \\ &\equiv \log\left(\frac{1}{P(X')}\right) - \log\left(\frac{1}{P(X)}\right) \geq \log\left(\frac{1}{P(X' \cup Y')}\right) - \log\left(\frac{1}{P(X \cup Y)}\right) \\ &\equiv \log\left(\frac{P(X)}{P(X')}\right) \geq \log\left(\frac{P(X \cup Y)}{P(X' \cup Y')}\right) \equiv \frac{P(X)}{P(X')} \geq \frac{P(X \cup Y)}{P(X' \cup Y')} \\ &\equiv \frac{P(X' \cup Y')}{P(X')} \geq \frac{P(X \cup Y)}{P(X)} \equiv P(Y'|X') \geq P(Y|X) \\ &\equiv \Delta P(Y|X) \geq 0 \end{aligned}$$

And in the same way:

$$\Delta I(Y) \geq I(X \cup Y) \equiv \Delta P(X|Y) \geq 0$$

PROOF 3. Satisfying SIM implies satisfying the identity axiom.

In order to prove that $Sim(X, X) > Sim(X, XY)$ is enough to say that information quantity increase (from left to right) of the first component is zero. The increase of the second component is negative, as well as the increase of the union:

$$I(X) - I(XY) < 0$$

$$I(X \cup X) - I(X \cup XY) = I(X) - I(XY) < 0$$

And equal to the information quantity increase of the second component. Therefore, the increase of the second component and the sum of increases in both components is equal than the union, while the increase of the first component is strictly higher. Therefore, the SIM conditions hold.

In order to prove that: $Sim(XY, XY) > Sim(XY, X)$ it is enough to say that the information quantity increase of the first component

is zero (from left to right):

$$I(\mathcal{X}Y) - I(\mathcal{X}Y) = 0$$

as well as the union:

$$I(\mathcal{X}Y \cup \mathcal{X}Y) - I(\mathcal{X}Y \cup \mathcal{X}) = I(\mathcal{X}Y) - I(\mathcal{X}Y) = 0$$

and the information quantity increase second component is positive:

$$I(\mathcal{X}Y) - I(\mathcal{X}) > 0$$

Therefore, the increase of the second component is equal than the union, while the increase of the first component and the sum of increases in both components is strictly higher. Therefore, the SIM conditions hold.

PROOF 4. *SIM axiom captures the identity specificity axiom.*

We want to prove that

$$Sim(\mathcal{X}Y, \mathcal{X}Y) > Sim(\mathcal{X}, \mathcal{X})$$

From left to right, the information quantity increase for the first and second components is positive:

$$I(\mathcal{X}Y) - I(\mathcal{X}) = k > 0$$

and equal than the information quantity increase of the union:

$$I(\mathcal{X}Y \cup \mathcal{X}Y) - I(\mathcal{X} \cup \mathcal{X}) = I(\mathcal{X}Y) - I(\mathcal{X})$$

Therefore, the sum of information quantity increases of both components is $2k$ which is higher than the union. Therefore, the SIM conditions hold.

PROOF 5. *SIM captures the unexpectedness axiom.*

We need to prove that if $P(\mathcal{Y}|\mathcal{X}) > P(\mathcal{Y}'|\mathcal{X})$ then

$$Sim(\mathcal{X}, \mathcal{X}Y) > Sim(\mathcal{X}, \mathcal{X}Y')$$

The information quantity increase for the first component is zero. Regarding the second component and the union:

$$\begin{aligned} I(\mathcal{X}Y) - I(\mathcal{X}Y') &= -\log\left(\frac{P(\mathcal{X}Y)}{P(\mathcal{X}Y')}\right) = -\log\left(\frac{P(\mathcal{X})P(\mathcal{Y}|\mathcal{X})}{P(\mathcal{X})P(\mathcal{Y}'|\mathcal{X})}\right) \\ &= -\log\left(\frac{P(\mathcal{Y}|\mathcal{X})}{P(\mathcal{Y}'|\mathcal{X})}\right) = \log\left(\frac{P(\mathcal{Y}'|\mathcal{X})}{P(\mathcal{Y}|\mathcal{X})}\right) < \log(1) < 0 \end{aligned}$$

Therefore, the increase of the second component is equal than the union, while the increase of the first component and the sum of increases in both components is strictly higher than the union. Therefore, the SIM conditions hold.

PROOF 6. *The Similarity Information Monotonocity axiom captures the dependency axiom.*

We need to prove that one of the SIM condition is enough to prove that if

$$P(\mathcal{X}Z|\mathcal{Y}Z') > P(\mathcal{X}|\mathcal{Y}) \text{ and } P(\mathcal{Y}Z'|\mathcal{X}Z) > P(\mathcal{Y}|\mathcal{X})$$

then

$$Sim(\mathcal{X}Z, \mathcal{Y}Z') > Sim(\mathcal{X}, \mathcal{Y})$$

Let us prove it.

$$P(\mathcal{X}Z|\mathcal{Y}Z') > P(\mathcal{X}|\mathcal{Y}) \equiv P(\mathcal{X}|\mathcal{Y}) < P(\mathcal{X} \cup \mathcal{Z}|\mathcal{Y} \cup \mathcal{Z}')$$

which is equivalent to:

$$\begin{aligned} \frac{P(\mathcal{X} \cup \mathcal{Z}|\mathcal{Y})}{P(\mathcal{Y})} &< \frac{P(\mathcal{X} \cup \mathcal{Z} \cup \mathcal{Y} \cup \mathcal{Z}')}{P(\mathcal{Y} \cup \mathcal{Z}')} \equiv \frac{\frac{1}{P(\mathcal{X} \cup \mathcal{Z} \cup \mathcal{Y} \cup \mathcal{Z}')}}{\frac{1}{P(\mathcal{X} \cup \mathcal{Z})}} < \frac{\frac{1}{P(\mathcal{Y} \cup \mathcal{Z}')}}{\frac{1}{P(\mathcal{Y})}} \\ &\equiv \log\left(\frac{\frac{1}{P(\mathcal{X} \cup \mathcal{Z} \cup \mathcal{Y} \cup \mathcal{Z}')}}{\frac{1}{P(\mathcal{X} \cup \mathcal{Z})}}\right) < \log\left(\frac{\frac{1}{P(\mathcal{Y} \cup \mathcal{Z}')}}{\frac{1}{P(\mathcal{Y})}}\right) \\ &\equiv I(\mathcal{X} \cup \mathcal{Z} \cup \mathcal{Y} \cup \mathcal{Z}') - I(\mathcal{X} \cup \mathcal{Z}) < I(\mathcal{Y} \cup \mathcal{Z}') - I(\mathcal{Y}) \end{aligned}$$

Therefore, the information quantity increase of the second component is bigger that the information quantity increase of the union.

In the same way,

$$P(\mathcal{Y}|\mathcal{X}) < P(\mathcal{Y} \cup \mathcal{Z}'|\mathcal{X} \cup \mathcal{Z})$$

is equivalent to:

$$I(\mathcal{X} \cup \mathcal{Z} \cup \mathcal{Y} \cup \mathcal{Z}') - I(\mathcal{X} \cup \mathcal{Z}) < I(\mathcal{X} \cup \mathcal{Z}') - I(\mathcal{X})$$

Therefore, the information quantity increase of the first component is bigger that the information quantity increase of the union. And then, the conditions of SIM holds.

PROOF 7. *Assuming independence between intersection and difference set componentes:*

$$I(\mathcal{X} \cup \mathcal{Y}) = I(\mathcal{X} \cap \mathcal{Y}) + I(\mathcal{X} \setminus \mathcal{Y}) + I(\mathcal{Y} \setminus \mathcal{X})$$

then the SIM axioms are equivalent to say that:

$$\Delta I(\mathcal{X} \cap \mathcal{Y}) \geq 0 \wedge \Delta I(\mathcal{X} \setminus \mathcal{Y}) \leq 0 \wedge \Delta I(\mathcal{Y} \setminus \mathcal{X}) \leq 0$$

Let us prove it. Assuming independence:

$$\begin{aligned} \Delta I(\mathcal{X}) + \Delta I(\mathcal{Y}) &\geq \Delta I(\mathcal{X} \cup \mathcal{Y}) \equiv \\ &\equiv 2\Delta I(\mathcal{X} \cap \mathcal{Y}) + \Delta I(\mathcal{X} \setminus \mathcal{Y}) + \Delta I(\mathcal{Y} \setminus \mathcal{X}) \geq \\ &\quad \Delta I(\mathcal{X} \cap \mathcal{Y}) + \Delta I(\mathcal{X} \setminus \mathcal{Y}) + \Delta I(\mathcal{Y} \setminus \mathcal{X}) \\ &\equiv \Delta I(\mathcal{X} \cap \mathcal{Y}) \geq 0 \end{aligned}$$

and

$$\begin{aligned} \Delta I(\mathcal{X}) &\geq \Delta I(\mathcal{X} \cup \mathcal{Y}) \\ &\equiv \Delta I(\mathcal{X} \cap \mathcal{Y}) + \Delta I(\mathcal{X} \setminus \mathcal{Y}) \geq \\ &\quad \Delta I(\mathcal{X} \cap \mathcal{Y}) + \Delta I(\mathcal{X} \setminus \mathcal{Y}) + \Delta I(\mathcal{Y} \setminus \mathcal{X}) \\ &\equiv 0 \geq \Delta I(\mathcal{Y} \setminus \mathcal{X}) \end{aligned}$$

And in the same way:

$$\begin{aligned} \Delta I(\mathcal{Y}) &\geq \Delta I(\mathcal{X} \cup \mathcal{Y}) \\ &\equiv \Delta I(\mathcal{X} \cap \mathcal{Y}) + \Delta I(\mathcal{Y} \setminus \mathcal{X}) \geq \\ &\quad \Delta I(\mathcal{X} \cap \mathcal{Y}) + \Delta I(\mathcal{X} \setminus \mathcal{Y}) + \Delta I(\mathcal{Y} \setminus \mathcal{X}) \\ &\equiv 0 \geq \Delta I(\mathcal{X} \setminus \mathcal{Y}) \end{aligned}$$

PROOF 8. *According to the ratio contrast model, whenever $\alpha_1 = \alpha_4$, the relative ordering of similarity instance values is not affected by α_1 :*

$$Sim(\mathcal{X}, \mathcal{Y}) > Sim(\mathcal{X}', \mathcal{Y}') \equiv$$

$$\begin{aligned} & \frac{\alpha_1 f(\mathcal{X} \cap \mathcal{Y})}{\alpha_2 f(\mathcal{X} \setminus \mathcal{Y}) + \alpha_3 f(\mathcal{Y} \setminus \mathcal{X}) + \alpha_1 f(\mathcal{X} \cap \mathcal{Y})} > \\ & \frac{\alpha_1 f(\mathcal{X}' \cap \mathcal{Y}')}{\alpha_2 f(\mathcal{X}' \setminus \mathcal{Y}') + \alpha_3 f(\mathcal{Y}' \setminus \mathcal{X}') + \alpha_1 f(\mathcal{X}' \cap \mathcal{Y}')} \equiv \\ & \frac{1}{\frac{\alpha_2 f(\mathcal{X} \setminus \mathcal{Y})}{\alpha_1 f(\mathcal{X} \cap \mathcal{Y})} + \frac{\alpha_3 f(\mathcal{Y} \setminus \mathcal{X})}{\alpha_1 f(\mathcal{X} \cap \mathcal{Y})} + 1} > \frac{1}{\frac{\alpha_2 f(\mathcal{X}' \setminus \mathcal{Y}')}{\alpha_1 f(\mathcal{X}' \cap \mathcal{Y}')} + \frac{\alpha_3 f(\mathcal{Y}' \setminus \mathcal{X}')}{\alpha_1 f(\mathcal{X}' \cap \mathcal{Y}')} + 1} \equiv \\ & \frac{\alpha_2 f(\mathcal{X} \setminus \mathcal{Y})}{\alpha_1 f(\mathcal{X} \cap \mathcal{Y})} + \frac{\alpha_3 f(\mathcal{Y} \setminus \mathcal{X})}{\alpha_1 f(\mathcal{X} \cap \mathcal{Y})} < \frac{\alpha_2 f(\mathcal{X}' \setminus \mathcal{Y}')}{\alpha_1 f(\mathcal{X}' \cap \mathcal{Y}')} + \frac{\alpha_3 f(\mathcal{Y}' \setminus \mathcal{X}')}{\alpha_1 f(\mathcal{X}' \cap \mathcal{Y}')} \equiv \\ & \frac{\alpha_2 f(\mathcal{X} \setminus \mathcal{Y}) + \alpha_3 f(\mathcal{Y} \setminus \mathcal{X})}{f(\mathcal{X} \cap \mathcal{Y})} < \frac{\alpha_2 f(\mathcal{X}' \setminus \mathcal{Y}') + \alpha_3 f(\mathcal{Y}' \setminus \mathcal{X}')}{f(\mathcal{X}' \cap \mathcal{Y}')} \end{aligned}$$

PROOF 9. *The ratio contrast model, does not capture the identity specificity counter sample.*

Assuming that f is zero for an empty set, the self similarity is fixed for any object:

$$\begin{aligned} Sim(\mathcal{X}, \mathcal{X}) &= \frac{\alpha_1 f(\mathcal{X} \cap \mathcal{X})}{\alpha_2 f(\mathcal{X} \setminus \mathcal{X}) + \alpha_3 f(\mathcal{X} \setminus \mathcal{X}) + \alpha_4 f(\mathcal{X} \cap \mathcal{X})} \\ &= \frac{\alpha_1 f(\mathcal{X})}{\alpha_2 f(\emptyset) + \alpha_3 f(\emptyset) + \alpha_4 f(\mathcal{X})} \\ &= \frac{\alpha_1 f(\mathcal{X})}{\alpha_4 f(\mathcal{X})} = \frac{\alpha_1}{\alpha_4} \end{aligned}$$

PROOF 10. *PMI satisfies the dependency constraint.*

If $P(\mathcal{XZ}|\mathcal{Y}Z') > P(\mathcal{X}|\mathcal{Y})$, then

$$\begin{aligned} PMI(\mathcal{XZ}, \mathcal{Y}Z') &= \left(\frac{P(\mathcal{X}YZZ')}{P(\mathcal{XZ}) * P(\mathcal{Y}Z')} \right) = P(\mathcal{XZ}|\mathcal{Y}Z') \frac{1}{P(\mathcal{XZ})} > \\ P(\mathcal{X}|\mathcal{Y}) \frac{1}{P(\mathcal{XZ})} &= \frac{P(\mathcal{X}Y)}{P(\mathcal{Y})P(\mathcal{XZ})} > \frac{P(\mathcal{X}Y)}{P(\mathcal{Y})P(\mathcal{X})} = PMI(\mathcal{X}, Y) \end{aligned}$$

PROOF 11. *PMI does not capture the unexpectedness axiom.*

Given a text $\mathcal{X} \cup \mathcal{Z}$ subsuming \mathcal{X} :

$$\begin{aligned} PMI(\mathcal{XZ}, \mathcal{X}) &= \log \left(\frac{P(\mathcal{XZ} \cup \mathcal{X})}{P(\mathcal{XZ})P(\mathcal{X})} \right) = \log \left(\frac{P(\mathcal{XZ})}{P(\mathcal{XZ})P(\mathcal{X})} \right) = \\ &= \frac{1}{P(\mathcal{X})} = \frac{P(\mathcal{X})}{P(\mathcal{X}P(\mathcal{X}))} = PMI(\mathcal{X}, \mathcal{X}) \end{aligned}$$

Therefore PMI is constant and the unexpectedness axiom can not be satisfied.

PROOF 12. *ICM satisfies the SIM axiom whenever $\alpha_1 + \alpha_2 > \beta > \alpha_1 > \alpha_2 > 0$.*

If $I(\mathcal{X} \cup \mathcal{Y}) \geq 0$ then, given that:

$$\Delta ICM_{\alpha_1, \alpha_2, \beta}(\mathcal{X}, \mathcal{Y}) = \alpha_1 \Delta I(\mathcal{X}) + \alpha_2 \Delta I(\mathcal{Y}) - \beta \Delta I(\mathcal{X} \cup \mathcal{Y})$$

Given that $\alpha_1 > 0$, $\alpha_2 > 0$, $I(\mathcal{X}) \geq I(\mathcal{X} \cup \mathcal{Y})$ and $I(\mathcal{Y}) \geq I(\mathcal{X} \cup \mathcal{Y})$:

$$\begin{aligned} \Delta ICM_{\alpha_1, \alpha_2, \beta}(\mathcal{X}, \mathcal{Y}) &\geq \alpha_1 \Delta I(\mathcal{X} \cup \mathcal{Y}) + \alpha_2 \Delta I(\mathcal{X} \cup \mathcal{Y}) - \beta \Delta I(\mathcal{X} \cup \mathcal{Y}) = \\ &(\alpha_1 + \alpha_2 - \beta) \Delta I(\mathcal{X} \cup \mathcal{Y}) > \Delta I(\mathcal{X} \cup \mathcal{Y}) > 0 \end{aligned}$$

In other case, if $I(\mathcal{X} \cup \mathcal{Y}) < 0$ then

$$\begin{aligned} \Delta ICM_{\alpha_1, \alpha_2, \beta}(\mathcal{X}, \mathcal{Y}) &= \alpha_1 \Delta I(\mathcal{X}) + \alpha_2 \Delta I(\mathcal{Y}) - \beta \Delta I(\mathcal{X} \cup \mathcal{Y}) \geq \\ &\alpha_1 \Delta I(\mathcal{X}) + \alpha_2 \Delta I(\mathcal{Y}) - \beta \Delta I(\mathcal{X}) - \beta \Delta I(\mathcal{Y}) = \\ &(\alpha_1 - \beta) \Delta I(\mathcal{X}) + (\alpha_2 - \beta) \Delta I(\mathcal{Y}) = \\ &-(\beta - \alpha_1) \Delta I(\mathcal{X}) - (\beta - \alpha_2) \Delta I(\mathcal{Y}) \geq \end{aligned}$$

$$\begin{aligned} -(\beta - \alpha_1) \Delta I(\mathcal{X} \cup \mathcal{Y}) - (\beta - \alpha_2) \Delta I(\mathcal{X} \cup \mathcal{Y}) &= \\ -(2\beta - \alpha_1 - \alpha_2) \Delta I(\mathcal{X} \cup \mathcal{Y}) \end{aligned}$$

Given that

$$2\beta - \alpha_1 - \alpha_2 > 2\beta - \beta - \beta = 0$$

and $I(\mathcal{X} \cup \mathcal{Y}) < 0$ Therefore:

$$-(2\beta - \alpha_1 - \alpha_2) \Delta I(\mathcal{X} \cup \mathcal{Y}) > 0$$

PROOF 13. *ICM has a direct relationship with Pointwise Mutual Information and conditional probabilities.*

When $\beta = \alpha_1 = \alpha_2$ then ICM matches with the Pointwise Mutual Information.

$$\begin{aligned} ICM_{\alpha, \alpha, \alpha}(\mathcal{X}, \mathcal{Y}) &= \log \left(\frac{P(\mathcal{X} \cup \mathcal{Y})^\alpha}{P(\mathcal{X})^\alpha * P(\mathcal{Y})^\alpha} \right) = \\ &= \alpha \log \left(\frac{P(\mathcal{X} \cup \mathcal{Y})}{P(\mathcal{X}) * P(\mathcal{Y})} \right) = \end{aligned}$$

At the other extrem when $\beta = \alpha_1 + \alpha_2$, the ICM fit into the product of conditional probabilities.

$$\begin{aligned} ICM_{\alpha_1, \alpha_2, \alpha_1 + \alpha_2}(\mathcal{X}, \mathcal{Y}) &= \log \left(\frac{P(\mathcal{X} \cup \mathcal{Y})^{\alpha_1 + \alpha_2}}{P(\mathcal{X})^{\alpha_1} * P(\mathcal{Y})^{\alpha_2}} \right) = \\ &= \log \left(\frac{P(\mathcal{X} \cup \mathcal{Y})^{\alpha_1 + \alpha_2}}{P(\mathcal{X})^{\alpha_1} * P(\mathcal{Y})^{\alpha_2}} \right) = \log \left(\frac{P(\mathcal{X} \cup \mathcal{Y})^{\alpha_1} * P(\mathcal{X} \cup \mathcal{Y})^{\alpha_2}}{P(\mathcal{X})^{\alpha_1} * P(\mathcal{Y})^{\alpha_2}} \right) = \\ &= \log \left(\frac{P(\mathcal{X} \cup \mathcal{Y})^{\alpha_1}}{P(\mathcal{X})^{\alpha_1}} \frac{P(\mathcal{X} \cup \mathcal{Y})^{\alpha_2}}{P(\mathcal{Y})^{\alpha_2}} \right) = \log (P(Y|X)^{\alpha_1} P(X|Y)^{\alpha_2}) \end{aligned}$$

Being $\alpha_1 = \alpha_2 = \alpha$:

$$\alpha \log (P(Y|X) * P(X|Y))$$

PROOF 14. *Assuming independence between component sets and information quantity as salience function, both ICM and the linear contrast model are equivalent*

$$\begin{aligned} ICM_{\alpha_1, \alpha_2, \beta}(\mathcal{X}, \mathcal{Y}) &= \alpha_1 I(\mathcal{X}) + \alpha_2 I(\mathcal{Y}) - \beta I(\mathcal{X} \cup \mathcal{Y}) \\ &= \alpha_1 (I(\mathcal{X} \cap \mathcal{Y}) + I(\mathcal{X} \setminus \mathcal{Y})) + \alpha_2 (I(\mathcal{X} \cap \mathcal{Y}) + I(\mathcal{X} \setminus \mathcal{Y})) - \\ &\quad \beta (I(\mathcal{X} \cap \mathcal{Y}) + I(\mathcal{X} \setminus \mathcal{Y}) + I(\mathcal{Y} \setminus \mathcal{X})) \\ &= (\alpha_1 + \alpha_2 - \beta) I(\mathcal{X} \cap \mathcal{Y}) - (\beta - \alpha_1) (I(\mathcal{X} \setminus \mathcal{Y})) - (\beta - \alpha_2) (I(\mathcal{Y} \setminus \mathcal{X})) \end{aligned}$$