# Subjective Logic

José C. Oliveira

June 26, 2020

# 1 Introduction

One of the major goals of our research is to develop a quantitative logic for reasoning about belief in social networks. Subjective logic is a logic that may has the expressiveness that we look for improve our influence graph and the belief state of an agent.

Subjective logic is an extension of probabilistic logic. In probabilistic logic, we can express the truth value of a proposition by a probability distribution over a domain with disjoint events or states and reason by the axioms of probability. When we have two states of a domain $\mathbb{X}$ and the probability of a state $x \in \mathbb{X}$ can be $P(x) = 0$ or $P(x) = 1$, we have binary logic. The expression $x \wedge y$ is expressed in probabilistic logic as $P(x \wedge y) = P(x)P(y)$. Probabilistic logic is an extension of binary logic.

Subjective logic extends probabilistic logic by adding *uncertainty* and *subjectivity*. We can't express *'we don't know'* with probabilistic logic by a uniform distribution because it says that we know that the distribution over the domain is uniform. Subjective logic can express that *uncertainty* about the distribution. The *subjectivity* comes from the fact that we can assign an opinion about a proposition to an agent.

By the investigation we are doing about subjective logic, our main question is: Can we use subjective logic to improve our model of social networks? If yes, how?

This document is an introduction about subjective logic and it will present elementary definitions, types of opinions and an overview about computational trust. The main reference of this document is the book *Subjective Logic: A Formalism for Reasoning Under Uncertainty* by Audun Jøsang.

# 2 Opinion representation

The main object of subjective logic is the *opinion*, and there are many equivalent ways for opinion representation. Here I'm going to present the most used one. We represent an opinion

by $\omega_X^A$, where $A$ is the an agent and $X$ is a random variable or a *hypervariable*. This section presents the elementary definitions that compose an opinion.

## 2.1 Elementary definitions

### 2.1.1 Domain and hyperdomain

In subjective logic, a domain is a state space consisting of a set of values which can also be called states, events, outcomes, hypotheses or propositions. Those values are assumed to be exclusive and exhaustive. Let $k = |\mathbb{X}|$ be the cardinality of $\mathbb{X}$.

Suppose we have a box have balls that can red, green, or blue. Then, the domain that represents all possible outcomes is

$$\mathbb{X} = \{red, green, blue\}. \tag{2.1}$$

**Definition 2.1.** *(Hyperdomain)* Let $\mathbb{X}$ be a domain, and let $\mathscr{P}(\mathbb{X})$ denote the powerset of $\mathbb{X}$. The powerset contains all subsets of $\mathbb{X}$, including the empty set $\emptyset$, and the domain $\mathbb{X}$ itself. The *hyperdomain* denoted $\mathscr{R}(\mathbb{X})$ is the reduced powerset of $\mathbb{X}$, i.e. the powerset excluding the empty-set $\varnothing$ and the domain value $\mathbb{X}$. The hyperdomain is expressed as

$$\text{Hyperdomain: } \mathscr{R}(\mathbb{X}) = \mathscr{P} \setminus \{\mathbb{X}, \emptyset\} \tag{2.2}$$

The hyperdomain of the box is

$$\mathscr{R}(\mathbb{X}) = \{ \quad \{red\}, \{green\}, \{blue\}, \\ \{red, green\}, \{red, blue\}, \{green, blue\} \quad \}. \tag{2.3}$$

Let $\kappa = |\mathscr{R}(\mathbb{X})| = 2^k - 2.$ be the cardinality of $\mathscr{R}(\mathbb{R})$.

Every value of the hyperdomain with one value is called *singleton*. Every value with more the one value is called *composite value*. The interpretation of a composite value being TRUE, is that one and only one of the constituent singletons is TRUE. The set of all composite values $\mathscr{C}(\mathbb{X})$ is called *composite set*.

A *hypervariable* $X$ is an random variable that its values from $\mathscr{R}(\mathbb{X})$. For example, if a hypervariable takes value from the composite value $\{blue, red\} \in \mathscr{R}(\mathbb{X})$, it means that we either draw a ball blue or red.

### 2.1.2 Belief mass distribution and uncertainty mass

*Belief mass* can be distributed over a (hyper)domain. Assign belief mass to a singleton value $x \in \mathbb{X}$ expresses support for $x$ being TRUE. Assign belief mass to a composite value $x \in \mathscr{C}(\mathbb{X})$ express support for one of the singleton values contained in $x$ being TRUE. Belief mass is sub-additive and is complemented by *uncertainty mass $u_X$* and it represents vacuity of evidence, i.e. the lack of support for the variable $X$ to have any specific value.

**Definition 2.2.** *(Belief Mass Distribution)* Let $\mathbb{X}$ be a domain with corresponding hyperdomain $\mathscr{R}(\mathbb{X})$, and let $X$ be a variable over those domains. A belief mass distribution denote $\mathbf{b}_X$ assigns belief mass to possible values of the variable $X$. In the case of a random variable $X \in \mathbb{X}$, the belief mass distribution applies to domain $\mathbb{X}$, and in the case of a hypervariable $X \in \mathscr{R}(\mathbb{X})$ the belief mass distribution applies to hyperdomain $\mathscr{R}(\mathbb{X})$. This is formally defined as follows.

$$\text{Multinomial belief mass distribution: } \mathbf{b}_X : \mathbb{X} \to [0, \ 1],$$
$$\text{with the additivity requirement: } u_X + \sum_{x \in \mathbb{X}} \mathbf{b}_X(x) = 1. \tag{2.4}$$

$$\text{Hypernominal belief mass distribution: } \mathbf{b}_X : \mathscr{R}(\mathbb{X}) \to [0, \ 1],$$
$$\text{with the additivity requirement: } u_X + \sum_{x \in \mathscr{R}(\mathbb{X})} \mathbf{b}_X(x) = 1. \tag{2.5}$$

### 2.1.3 Base rate distribution

Base rate distribution represents a *prior* probability distribution over a domain, i.e., when we are not considering any evidence and we are completely uncertain.

**Definition 2.3.** *(Base Rate Distribution)* Let $\mathbb{X}$ be a domain, and let $X$ be a random variable in $\mathbb{X}$. The base rate distribution $\mathbf{a}_X$ assigns base rate probability to possible values of $X \in \mathbb{X}$, and is an additive probability distribution, formally expressed as:

$$\text{Base rate distribution: } \mathbf{a}_X : \mathbb{X} \to [0, \ 1],$$
$$\text{with the additivity requirement: } \sum_{x \in \mathbb{X}} \mathbf{a}_X(x) = 1. \tag{2.6}$$

**Definition 2.4.** *(Base Rate Distribution over Values in a Hyperdomain)* Let $\mathbb{X}$ be a domain with corresponding hyperdomain $\mathscr{R}(\mathbb{X})$, and let $X$ be a variable over those domains. Assume the base rate distribution $\mathbf{a}_X$ over the domain $\mathbb{X}$ according to Definition 2.3. The base rate $\mathbf{a}_X$ for a composite value $x \in \mathscr{R}(\mathbb{X})$ can be computed as follows:

$$\text{Base rate over composite values: } \mathbf{a}_X(x_i) = \sum_{\substack{x_j \in \mathbb{X} \\ x_j \subseteq x_i}} \mathbf{a}_X(x_j), \ \forall x_i \in \mathscr{R}(\mathbb{X}). \tag{2.7}$$

**Definition 2.5.** *(Relative Base Rate)* Assume a domain $\mathbb{X}$ of cardinality $k$, and the corresponding hyperdomain $\mathscr{R}(\mathbb{X})$. Let $X$ be a hypervariable over $\mathscr{R}(\mathbb{X})$. Assume that a base rate distribution $\mathbf{a}_X$ is defined over $\mathbb{X}$ according to Definition 2.4. Then the base rate of a value $x$ relative to a value $v_i$ is expressed as the relative base rate $\mathbf{a}_X(x|x_i)$ defined below.

$$\mathbf{a}_X(x|x_i) = \frac{\mathbf{a}_X(x \cap x_i)}{\mathbf{a}_X(x_i)}, \ \forall x, x_i \in \mathscr{R}(\mathbb{X}), \text{ where } \mathbf{a}_X(x_i) \neq 0. \tag{2.8}$$

In the case when $\mathbf{a}_X(x_i) = 0$, then $\mathbf{a}_X(x|x_i) = 0$. Alternatively it can simply be assumed that $a_X(x_i) > 0$, for every $x_i \in \mathbb{X}$, meaning that everything we include in the domain has a non-zero base rate of occurrence in general.

3

Defining base rate distribution over a domain depends of the background of the given situation. For our box, the base rate distribution is uniform. Now suppose that we have a disease that happens in 10% of the population. Let $\mathbb{Y} = \{y, \overline{y}\}$ be the domain representing the occurrence or not of the disease on a person. Let $Y$ be a random variable over $\mathbb{Y}$. The base rate distribution for a untested person for the disease is $\mathbf{a}_Y = (0.05, 0.95)$.

### 2.1.4 Projected probability

The *projected probability* is a *posterior* probability distribution over the domain that considers the base rate distribution (*prior*), the belief mass distribution (evidence) and the uncertainty mass (lack of evidence). The projected probability distribution is different for every type of opinion, and it is explained below.

## 2.2 Binomial opinions

A binary opinion states about a domain with $k = 2$, and any multinomial opinion, with $k > 2$, can be considered binary when seen as a binary partition consisting of a proper subset $x \subset \mathbb{X}$ and its complement $\overline{x}$.

**Definition 2.6.** *Binomial Opinion* Let $\mathbb{X} = \{x, \overline{x}\}$ be a binary domain with binomial random variable $X \in \mathbb{X}$. A binomial opinion about the truth/presence of value $x$ is the ordered quadruplet $\omega_x = (b_x, d_x, u_x, a_x)$, where the additivity requirement

$$b_x + d_x + u_x = 1 \tag{2.9}$$

is satisfied, and where the respective parameters are defined as

- $b_x$: *belief mass* in support of $x$ being TRUE (i.e. $X = x$),

- $d_x$: *disbelief mass* in support of $x$ being FALSE (i.e. $X = \overline{x}$)

- $u_x$: *uncertainty mass* representing the vacuity of evidence,

- $a_x$: *base rate*, i.e. prior probability of $x$ without any evidence.

The projected probability distribution of binomial opinions is defined by:

$$P(x) = b_x + a_x u_x. \tag{2.10}$$

Note that the more the opinion relies on uncertainty, the more weight the base rate will have on the projected probability.

Let $\mathbb{Z} = \{red, green\}$ be a domain the represents the outcomes of drawing a ball from a box. Let $Z$ be a random variable over $\mathbb{Y}$. Let the base rate be uniform. Then, the opinion for drawing a red ball (and green ball too) is:

$$\omega_Z = (0, 0, 1, 0.5). \tag{2.11}$$

Now, I told you that there is 10 balls in the box, 4 of then is red and non is know about the other 6. The opinion for drawing a red ball is:

$$\omega_Z = (0.4, 0, 0.6, 0.5). \tag{2.12}$$

The projected probability distribution is:

$$
\begin{aligned}
P(red) &= 0.4 + 0.5 \cdot 0.6 &= 0.7, \\
P(green) &= 0 + 0.6 \cdot 0.5 &= 0.3.
\end{aligned}
\tag{2.13}
$$

## 2.3 Multinomial opinions

Multinomial opinions represent the generalisation of binomial opinions. Multinomial opinions apply to situations where a random variable $X$ over $\mathbb{X}$ can take one of multiple different values.

**Definition 2.7.** *(Multinomial Opinion)* Let $\mathbb{X}$ be a domain larger than binary, i.e. so that $k = |X| > 2$. Let $X$ be a random variable in $\mathbb{X}$. A multinomial opinion over the random variable $X$ is the ordered triplet $\omega_X = (\mathbf{b}_X, u_X, \mathbf{a}_X)$ where

- $\mathbf{b}_X$ is a belief mass distribution over $X$,

- $u_X$ is the uncertainty mass which represents the vacuity of evidence,

- $\mathbf{a}_X$ is a base rate distribution over $\mathbb{X}$,

and the multinomial additivity requirement of Eq.(2.4) is satisfied.

The projected probability distribution of multinomial opinions is defined by:

$$\mathbf{P}_X(x) = \mathbf{b}_X(x) + \mathbf{a}_X(x)u_X, \ \forall x \in \mathbb{X}. \tag{2.14}$$

Let $\mathbb{X} = \{red, green, blue\}$ be the domain representing the outcomes of drawing a ball from a box. Let $X$ be a random variable over $\mathbb{X}$. Let the base rate distribution be uniform. Now, suppose that we know that there is 3 red balls, 2 blue balls and we know nothing about the other 5 balls. Then, the opinion is:

$$
\omega_X = \left(
\begin{aligned}
\mathbf{b}_X(red) &= 0.3, & \mathbf{a}_X(red) &= 1/3, \\
\mathbf{b}_X(green) &= 0, & \mathbf{a}_X(green) &= 1/3, \\
\mathbf{b}_X(blue) &= 0.2, & \mathbf{a}_X(blue) &= 1/3, \\
u_X &= 0.5.
\end{aligned}
\right)
\tag{2.15}
$$

The projected probability distribution is:

$$
\begin{aligned}
\mathbf{P}(red) &= 0.3 + 1/3 \cdot 0.5 &\approx 0.466, \\
\mathbf{P}(green) &= 0 + 1/3 \cdot 0.5 &\approx 0.166, \\
\mathbf{P}(blue) &= 0.2 + 1/3 \cdot 0.5 &\approx 0.366.
\end{aligned}
\tag{2.16}
$$

## 2.4  Hypernomial opinions

Hyper-opinions represents the generalisation of multinomial opinions. In the case of a domain $\mathbb{X}$ with hyperdomain $\mathscr{R}(\mathbb{X})$, it is possible to obtain evidence for a composite value $x \in \mathscr{R}(\mathbb{X})$, which translates into assigning belief mass to that composite value.

**Definition 2.8.** *(Hyper-opinion)* Let $\mathbb{X}$ be a domain of cardinality $k > 2$, with corresponding hyperdomain $\mathscr{R}(\mathbb{X})$. Let $X$ be a hypervariable in $\mathscr{R}(\mathbb{X})$. A hyper-opinion on the hypervariable $X$ is the ordered triplet $\omega_X = (\mathbf{b}_X, u_X, \mathbf{a}_X)$ where

- $\mathbf{b}_X$ is a belief mass distribution over $\mathscr{R}(\mathbb{X})$,

- $u_X$ is the uncertainty mass which represents the vacuity of evidence,

- $\mathbf{a}_X$ is a base rate distribution over $\mathbb{X}$,

and the hypernomial additivity of Eq.(2.5) is satisfied.

The projected probability distribution of hyper-opinions is defined by:

$$\mathbf{P}_X(x) = \sum_{x_i \in \mathscr{R}(\mathbb{X})} \mathbf{a}_X(x|x_i)\mathbf{b}_X(x_i) + \mathbf{a}_X(x)u_X, \ \forall x \in \mathbb{X}. \tag{2.17}$$

Let $\mathbb{X} = \{red, green, blue\}$ be the domain representing the outcomes of drawing a ball from a box. The hyperdomain is:

$$\mathscr{R}(\mathbb{X}) = \{ \ \ \{red\}, \{green\}, \{blue\}, \\ \{red, green\}, \{red, blue\}, \{green, blue\} \ \ \}. \tag{2.18}$$

Let $X$ be a hypervariable over $\mathscr{R}(\mathbb{X})$. Let the base rate distribution be uniform. Now, suppose that we know that the box have one red ball, 2 blue balls, and 3 balls green or blue. We know nothing about 4 balls. Then, the hyperopinion is:

$$\omega_X = \begin{pmatrix} \mathbf{b}_X(\{red\}) & = 0.1, & \mathbf{a}_X(\{red\}) & = 1/3, \\ \mathbf{b}_X(\{green\}) & = 0, & \mathbf{a}_X(\{green\}) & = 1/3, \\ \mathbf{b}_X(\{blue\}) & = 0.2, & \mathbf{a}_X(\{blue\}) & = 1/3, \\ \mathbf{b}_X(\{red, green\}) & = 0, & \mathbf{a}_X(\{red, green\}) & = 2/3, \\ \mathbf{b}_X(\{red, blue\}) & = 0, & \mathbf{a}_X(\{red, blue\}) & = 2/3, \\ \mathbf{b}_X(\{green, blue\}) & = 0.3, & \mathbf{a}_X(\{green, blue\}) & = 2/3, \\ u_X & = 0.4. \end{pmatrix} \tag{2.19}$$

The projected probability distribution is:

$$\begin{aligned} \mathbf{P}(\{red\}) &\approx 0.2333, & \mathbf{P}(\{red, green\}) &\approx 0.5166, \\ \mathbf{P}(\{green\}) &\approx 0.2833, & \mathbf{P}(\{red, blue\}) &\approx 0.7166, \\ \mathbf{P}(\{blue\}) &\approx 0.4833, & \mathbf{P}(\{green, blue\}) &\approx 0.7666. \end{aligned} \tag{2.20}$$

# 3 Trust opinion

## 3.1 Notion of Trust

Trust is a directional relationship between two parties that can be called *trustor* and *trustee*. A trust relationship has a *scope*, meaning that it applies to a specific purpose or domain of action, such as 'being authentic' in the case of an agent's trust in a *cryptographic* key or 'providing reliable information' in the case of a person's trust in the correctness of an entry in Wikipedia.

There are two main different interpretations of trust. The one we are using were is the reliability trust defined below.

**Definition 3.1.** *(Reliability trust)* Reliability trust is the subjective belief with which an entity, A, expects that another entity, B, performs a given action on which A's welfare depends

For the purposes of the research about group polarization in social networks, we are working in *reliability trust* because we are focusing in the only in the influence of an agent over another. The other main interpretation is *decision trust* and it considers with utility, trustor's risk attitude and environmental factors.

This model of trust is said to be *computational* because we are reasoning about trust with a formal model. Subjective logic's trust model is not the only one in the literature. Also, there is no single correct formalism about trust because is not possible to validate with real people. In his book, Audun Jøsang states that his model about computational trust in subjective logic is specifically addressed to build systems.

## 3.2 Definition

Trust is a binomial opinion that an agent has about another. Let $\mathbb{T} = \{t, \bar{t}\}$ be a domain and let $T$ be a variable over $\mathbb{T}$. Let $B$ be an agent. A general interpretation for $\mathbb{T}$ about the action of the agent $B$ can be:

$$\text{Trust domain } \mathbb{T} : \begin{cases} t : \text{``Agent } B \text{ does perform the action as expected''} \\ \bar{t} : \text{``Agent } B \text{ does not perform the action as expected''} \end{cases} \tag{3.1}$$

Let $A$ be an agent. The opinion that $A$ have about $B$ is denoted by $\omega_{t_B}^A$. For a more directed notation, we assume that $\omega_{t_B}^A \equiv \omega_B^A$. If agent $A$ has absolute trust in $B$, then $\omega_B^A = (1, 0, 0, a_B)$, for some base rate $a_B$.

## 3.3 Trust transitivity

Lets illustrate trust transitivity with an example. Suppose that Alice's car is broken and she doesn't know any car mechanic. Alice borrow Bob's car for one day a she realizes that his car

is well maintained. Then, she trusts in Bob in matters of car maintenance. In order to have a well maintained car, Bob trusts the car mechanic Erik and asks his services. Then, Alice has a *indirect* trust in Erik in matters of car maintenance.
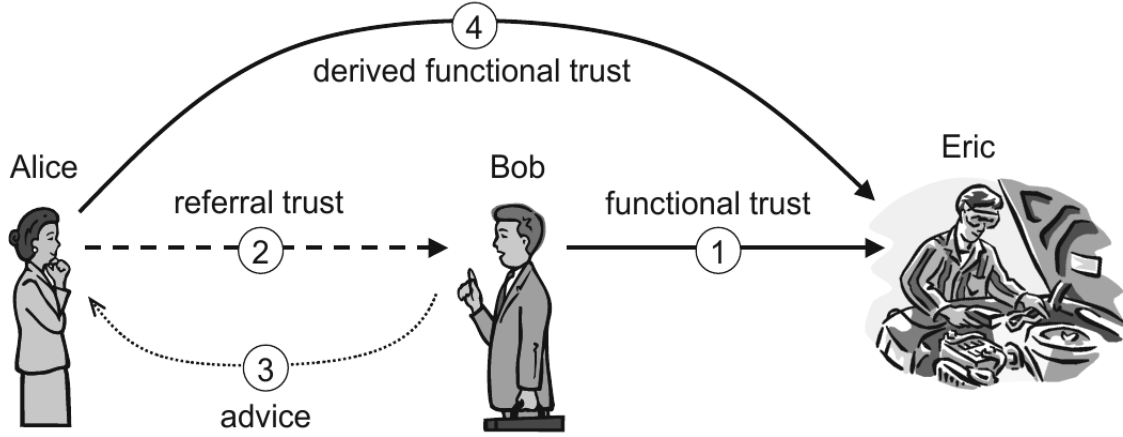


Figure 3.1: Transitive trust principle

The trust that Alice has about Bob is called *referral trust* and is denoted as $[A; B]$. The trust that Bob has about Erik is called *functional trust* and is denoted as $[B, C]$. The trust that Alice as about Erik through Bob is called *derived functional trust* as is denoted as:

$$\begin{aligned} [A, E] \quad &= [A; B] : [B, E] \\ &= [A; B, E]. \qquad \text{(Compact notation)} \end{aligned} \tag{3.2}$$

Let $\omega_B^A$ be the trust opinion that Alice has about Bob, $\omega_E^B$ be the trust opinion that Bob has about Erik, and $\omega_E^{[A;B]}$ be the trust opinion that Alice has about Erik through Bob. We obtain the opinion $\omega_E^{[A;B]}$ by the operator trust discount and is defined below.

**Definition 3.2.** (*Trust Discounting for Two-Edge Path*) Assume agents $A$ and $B$, where $A$ has referral trust in $B$ for a trust scope represented by domain $\mathbb{X}$. Let $X$ denote a variable on domain $X$, and let $\omega_X^B = (\mathbf{b}_X^B, u_X^B, \mathbf{a}_X^B))$ be agent $B$'s general opinion on $X$ as advised by $B$ to $A$. Assume further that agent $A$'s referral trust in $B$, with respect to advising belief about $X$, is expressed as $\omega_B^A$. The notation for trust discount is given by

$$\omega_X^{[A;B]} = \omega_B^A \otimes \omega_X^B. \tag{3.3}$$

The trust-discounting operator combines agent $A$'s referral trust opinion about agent $B$, denoted $\omega_B^A$, to discount $B$'s opinion about variable $X$, denoted $\omega_X^B$, to produce $A$'s derived opinion about $X$, denoted $\omega_X^{[A;B]}$. The parameters of the derived opinion $\omega_X^{[A;B]}$ are defined in the following way:

$$\omega_X^{[A;B]} = \begin{cases} \mathbf{b}_X^{[A;B]}(x) \quad &= \mathbf{P}_B^A \mathbf{b}_X^B(x), \\ u_X^{[A;B]} \quad &= 1 - \mathbf{P}_B^A \sum_{x \in \mathscr{R}(\mathbb{X})} \mathbf{b}_X^B(x), \\ \mathbf{a}_X^{[A;B]}(x) \quad &= \mathbf{a}_X^B(x). \end{cases} \tag{3.4}$$

In particular, when A fully trusts B, the opinion that A has about X will be completely accepted for A. Also, when A fully distrusts B, then opinion that A has about X will be completely rejected for A and A will have an opinion completely uncertain.

The previous example considers the relationship $[A, E]$ as a functional trust relationship. We also can consider as a belief relationship. Suppose now that Alice need to go to a restaurant, but she doesn't know any restaurant at the town. Alice knows that Bob has good taste for restaurants and as him an advice. Bob tells Alice some restaurants to go and his favorites.

Let $\mathbb{X}$ be the domain of all restaurants in town. Let $X$ be the variable over $\mathbb{X}$ that describes the best restaurant. The relationship $[A, X]$ is not about trust but about belief. Bob acts as a *source information* for Alice about $X$. Trust discounting also applies here.
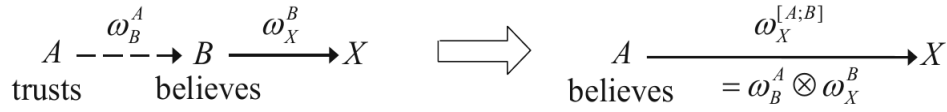
Figure 3.2: Trust discounting of opinions

## 3.4 Trust fusion

Suppose that Alice now wants the opinion of Clarie about the restaurants in town. She has different trust opinions about Bob and Clarie and she want to consider both. Belief fusion $\oplus$ is the operator that solves that issue. I don't know the details about the operator yet, because there is a whole chapter for it. Here, I explain how it works with trust opinions.

We use the symbol $\diamond$ to denote a fusion between two relations:

$$\begin{aligned} [A, E] \quad &= ([A; B] : [B, X]) \diamond ([A; C] : [C, X]) \\ &= [A; B, C] \diamond [A; C, X]. \qquad \text{(Compact notation)} \end{aligned} \tag{3.5}$$

The opinion that Alice has about X through the Bob and Clarie is:

$$\omega_X^{[A;B]\diamond[A;C]} = (\omega_B^A \otimes \omega_X^B) \oplus (\omega_C^A \otimes \omega_X^C) \tag{3.6}$$
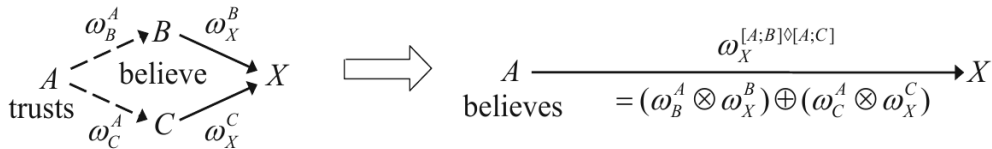
Figure 3.3: Fusion of trust-discounted opinions

# 4 The model and next questions

We have a goal to improve or model with subjective logic. At first, I was thinking of model a society with the following representations:

- The influence of an agent B over an agent A can be represented by the trust opinion $\omega_B^A$.

- Trust transitivity can be useful for update functions.

- When an agent has a trust opinion for each of the other agents of the society, trust fusion can be used to obtain the overall belief update.

The first problem is: **Trust discounting does not consider the case that the trustor already has an opinion**. Because of it, we don't have *update*. To solve it, I plan to look if someone has the same problem before at the literature and search for some operator in subjective logic that has the same properties that we expect. If I don't find any operator, may be the case of creating one, but I don't if this is *allowed* in subjective logic.

The second problem is: **How to represent influence by a trust opinion?**. Suppose that an agent B has no influence on A ($I(A, B) = 0$). I thought of two ways to represent by a opinion: a) $\omega_B^A = (0, 0, 1, a_X^A))$, agent A is completely uncertain about trusting in B, or b) $\omega_B^A = (0, 1, 0, a_X^A)$, agent A completely distrusts B. I have to understand the difference between this representations when used with trust discounting and trust fusion.

In this document, I presented what I find most useful for the research about group polarization in social networks. I studied some other topics that I didn't presented here like the difference between aleatory and epistemic opinions, decision theory in subjective logic and how opinions can be represented with Beta PDFs and Dirichlet PDFs.