An algebra for trust dilution and trust fusion

Baptiste Alcalde and Sjouke Mauw

University of Luxembourg baptiste.alcalde@uni.lu; sjouke.mauw@uni.lu

Abstract. Trust dilution and trust fusion are two operators that are used to calculate transitive trust in a trust network. Various implementations of these operators already exist but are not fully motivated. In this paper we define the basic properties of these two operators by developing a trust algebra. We evaluate several new and existing models against the axioms of this algebra, amongst which a number of variations of the Subjective Logic. The algebra enables the comparison of models and gives more insight in the available recommendation models and their properties.

1 Introduction

Trust transitivity is defined as the possibility to use trust information from other entities in order to infer a trust evaluation to a given entity. Trust transitivity is a key concept of recommendation systems and it attracts an ever increasing interest in the very recent years [4, 6, 12, 13]. To date, we can identify two main recommendation model families. The first is qualitative and uses, for instance, modal logic [4]. The second, which is the focus of this paper, is quantitative and defines special trust operators, named fusion and dilution operators, in order to compute the resulting trust of a trust network [9, 14, 15].

Dilution is used to calculate the trust along trust chains. This operator combines agent A's trust in agent B with agent B's trust in agent C, to derive A's trust in C. Fusion is used to compute the overall trust if there are different sources of information. If agent A has two independent sources of information, say B and D, on the trustworthiness of agent C, then B and D's information can be combined using the fusion operator.

In literature, several different definitions of fusion and dilution operators have been proposed. These definitions are often mainly motivated by technical observations, rather than by strong and defendable intuition. Hence these definitions can be hard to understand or to interpret from the point of view of an outsider. Therefore, one of the main motivations for the current research is the lack of determination of the intrinsic properties of these operators.

In order to judge whether dilution and fusion definitions provide a suitable modeling of the phenomena, a description of these phenomena at a higher level of abstraction is required. In addition, such abstraction will enable to compare the relative merits of alternative definitions and extensions.

Hence, the first contribution of this paper is to develop a higher level of abstraction in the form of a *trust algebra*. This algebra consists of a number of

defining properties of the fusion and dilution operators. One of the merits of this approach is that these properties, stated in the form of equational axioms, can be motivated from the domain of trust and recommendation. Therefore, rather than proposing a new recommendation model, this paper aims at giving formal guidelines to possible implementation of recommendation models as well as a means for their comparison. To our knowledge, this approach was never explored until now.

Such an algebraic approach has shown very beneficial e.g. in the realm of parallel systems, in which the development and analysis of process algebras has added to the understanding of the many different process models. Similarly, the development of a trust algebra will help to understand the different trust models.

Using this approach, Subjective Logic (SL), as proposed in [9], can then be seen as one of a number of possible (and plausible) models for the abstract algebra. We consider the investigation of existing recommendation models, amongst which SL, and the development of new models as one of the contributions of this paper. Our study also reveals weaknesses in some of the models, and can provide a valuable feedback for the establishment of future models.

The paper is structured as follows. In Sect. 2 we clarify some general definitions and assumptions needed as a background for this research. In Sect. 3 we present the trust algebra and some extensions as well as the motivation for the rules composing this algebra. In Sect. 4 we show the applicability of the algebra through the comparison of the canonical models (with three and four elements), and Subjective Logic variations. As a result from the evaluation, we can prove impossibility results in some models, show the limitations of others, and prove the correctness of a newly crafted model. A summary of the results is provided in a table at the end of Sect. 4. In the conclusion we interpret our results and propose a number of interesting venues for future work.

2 Trust relations

Trust has been defined in several different ways. The definition of trust adopted here, first formulated by Gambetta [5], is often referred to as "reliability trust". Thus, we define *trust* as the belief or subjective probability of the *trustor* that the *trustee* will adequately perform a certain action on which the trustor's welfare depends. We also refer to the trustor and trustee as agents, which may be humans or computer programs acting on the behalf of humans. Trust is hence a quantifiable relation between two agents.

In literature, many factors have been identified that can be taken into account when calculating the trust relation between two agents (see [1] for an overview). These factors comprise e.g. the trustor's personality, the trustee's competence, contextual information such as local norms and customs, and the opinions of other agents. The introduction of opinions allows a trustor to take other agents' opinions into account when determining the trustworthiness of a trustee, thus yielding a trust network. We start off with the observation from [11] that there are different notions of trust involved. First, we make a distinction between two variants of trust, viz. functional trust and referral trust. Functional trust is the belief in an entity's ability (and willingness) to carry out or support a specific function on which the relying party depends. Referral trust is the belief in an entity's ability to recommend another entity w.r.t functional trust. The other distinction is between two types of trust, viz. direct trust and indirect trust. A direct trust relation occurs when the trustor trusts the trustee directly (without intermediaries), e.g. based on past experiences between them. An indirect trust occurs when the trustor trusts a trustee based on one or more opinions from third parties.

By combining a trust variant with a trust type we can obtain functional direct trust, functional indirect trust, referral direct trust, and referral indirect trust. For example, Alice wants to know where to find a good car mechanic to fix her car. She asks Bob's opinion because he is knowledgeable about cars (direct referral trust). Bob happens to know a good car mechanic (direct functional trust). Bob then suggests Alice the name of this car mechanic (recommendation). Alice can then bring her car to this car mechanic (indirect functional trust). We can note that after this transaction, Alice will transform her indirect functional trust into direct functional trust (since she will then have a direct experience with the car mechanic).

In addition to these definitions we set a number of assumptions. First, we assume that the trustor knows all trust relations between agents that are relevant for her own trust calculations. In literature, this rather strong assumption is often called *perfect forward*, as to indicate that all agents are willing to forward other agents' trust values without modification.

Further, we assume that each agent keeps track of his own functional and referral trust in other agents. We will use the same domain for expressing trust values of all variants and types of trust. In a given trust graph all trust relations concern referral trust, except for the arrows directly ending at the trustee, which concern functional trust. In the remainder of the paper, we will therefore not explicitly mention the type of a given trust relation if it can be derived from the context.

In literature, a distinction is made between two ways of composing trust values, viz. fusion and dilution. Trust fusion occurs if there are multiple trust paths from a trustor to a trustee, meaning that the trustee is recommended by several agents. In order to calculate his trust in the trustee, the trustor then has to fuse the trust values of these other agents. The fusion of trust values does not necessarily lead to a higher level of trust. The dilution of trust occurs if there is a trust chain from the trustor to the trustee. Every link in the chain reduces (or dilutes) the overall trust of the trustor in the trustee implied by the trust chain.

3 A trust algebra

In this section we develop the algebra of *trust expressions*, which is the first contribution of this paper. This abstract algebra is partly based on the more concrete operators found in literature (see e.g. [14, 15, 8]).

We develop our algebra in four layers. The first layer introduces the basic constants and operators and their basic properties. The second layer provides an extension of this algebra which allows us to compare trust expressions. In the third layer we express the duality of belief and disbelief, while in the fourth layer we treat the special case of full direct functional belief and disbelief.

Basic Fusion and Dilution Algebra. A trust expression is obtained by (recursively) applying some *trust operators* to a number of *trust atoms* (or trust values). See the upper frame of Fig. 1 for the signature of basic trust expressions. The set of trust expressions is denoted by \mathcal{T} and the set of trust atoms by \mathcal{A} . We consider two basic trust operators: trust fusion (denoted by $_+_$) and trust dilution (denoted by $_-_$). The set of trust atoms is not specified in detail. We require that it contains at least the three constants v, β , and δ . Constant v denotes full uncertainty, i.e. the absence of any information that can help to assess the trustworthiness of the trustee. Constant β denotes full trust of the trustor in the trustee, without any uncertainty. Constant δ is the dual of β . It denotes full distrust of the trustor in the truste of the trust expressions.

An example of a trust expression is $(\beta \cdot \beta) + (v \cdot \delta)$. This expresses that, although the trustor has no direct trust relation with the trustee, he knows two independent sources that have a direct functional trust relation with the trustee. The trustor has full direct referral trust in the first source, who has full direct functional trust in the trustee $(\beta \cdot \beta)$. Further, the trustor is completely uncertain whether to trust his second source or not. His second source does not trust the trustee at all; he has full distrust in the trustee $(v \cdot \delta)$.

In order to simplify trust expressions we assume that the dilution operator $_._$ binds stronger than the fusion operator $_+_$. We will often omit the $_._$ operator from expressions if no confusion can arise. In this way, the above example can be simplified to $\beta\beta + v\delta$.

| \mathcal{T} | set of trust expre | essions $_\cdot_:\mathcal{T}$ | $\times \mathcal{T} \to \mathcal{T}$ dilution (| (operator) |
|--|--------------------|---------------------------------|---|---------------------------------|
| $\mathcal{A}\subseteq \mathcal{T}$ | set of trust atom | s $v \in A$ | full uncer | rtainty (constant) |
| $x, y \in \mathcal{T}$ | variables | $\beta \in \mathcal{A}$ | full belief | f (constant) |
| $_+_:\mathcal{T}\times\mathcal{T}\to\mathcal{T}$ | fusion (operator) | $\delta \in \mathcal{A}$ | full disbe | lief (constant) |
| (B1) $x + y = y +$ | <i>x</i> (| C1) $x + v = x$ | (C4) $\beta + \beta = \beta$ | (C6) $\delta + \delta = \delta$ |
| (B2) $x + (y + z)$ | =(x+y)+z (| C2) $x \cdot v = v$ | (C5) $\beta \cdot x = x$ | (C7) $\delta \cdot x = v$ |
| (B3) $x(yz) = (xy)$ |)z (| C3) $v \cdot x = v$ | | |

Fig. 1. Basic Fusion and Dilution algebra (BFD)

The properties of these constants and operators are expressed by a set of axioms, which we call the *Basic Fusion and Dilution* (BFD) algebra (Fig. 1). The first three axioms express properties of the basic operators. The fusion operator is commutative (B1) and associative (B2), since the order in which the trustor receives independent recommendations is irrelevant. Calculating transitive trust along a trust chain is also associative, so dilution is an associative operator as well (B3). However, dilution is not commutative. This can be seen by a simple example. Assume that agent A fully trusts agent B's opinion on agent C and assume that B fully distrusts C. Then A should also fully distrust C. However,

if we swap the values, i.e. A has full distrust in B, who fully trusts C, then A should not necessarily (dis)trust C, so $\beta \delta \neq \delta \beta$.

Axioms C1–C7 define the properties of the three constants. The uncertainty constant v behaves like a zero element. Adding a fully uncertain opinion to an opinion x does not give any extra information, so x + v = x (axiom C1). By combining this with axiom B1 we obtain the symmetric case v + x = x. Axioms C2 and C3 express that full uncertainty in a trust chain annihilates any other information in this chain, so $x \cdot v = v \cdot x = v$. If we fuse full belief with itself, it remains full belief (axiom C4). Axiom C5 expresses that the full belief constant β behaves as a left-unit for dilution. This follows from the fact that if we fully belief another agent, we adopt his opinion without any hesitation. Clearly, β is not a right-unit, so we don't have $x \cdot \beta = x$. If A distrusts B and B trusts C, then this does not mean that A should distrust C, $\delta \beta \neq \delta$. The disbelief constant δ behaves similar to β in a fusion context: if we get our full disbelief confirmed by another source, the fusion is still full disbelief (axiom C6). Finally, if we consider the opinion of somebody whom we disbelief, it will give us no information at all, so $\delta \cdot x = v$ (axiom C7). Obviously, the converse, $x \cdot \delta = v$, does not hold, since e.g. $\beta \cdot \delta = \delta$. Later we will come back to expressions of the form $x \cdot \beta$ and $x \cdot \delta$.

Jøsang [11] also mentions interpretations which are different from the intuition sketched above. If we assume that "the friend of my enemy is my enemy", then the interpretation of β as a right-unit would make sense. This interpretation would also have consequences for axiom C7, since then we would have $\delta\beta = \delta$. However, following Jøsang, we consider these interpretations as rather exotic and we will leave them for future study.

It is important to notice that there exist expressions that are not equal (after applying the axioms) to a constant. An example is $\beta + \delta$ which expresses that via one route we obtain the information that the trustee can be trusted without any uncertainty, while via another independent route we learn that the trustee must be distrusted without any uncertainty. There are different ways to interpret the fusion of such dogmatic opinions, some of which are discussed in [10]. In order to allow for such different interpretations, we decided to not settle for a fixed interpretation in the algebra. Alternative interpretations can then be expressed by defining different models of the algebra.

As discussed by Jøsang [8] the fusion and dilution operators do not distribute. For instance, xz + yz = (x + y)z is not a desired property, because in the lefthand side of this equation the two occurrences of z represent two *independent* opinions, which must both be taken into consideration in the fusion. Hence, they will reinforce each other. However, in the right-hand side of the equation, opinion z is only considered once. For the same reason, idempotence of the fusion operator (x + x = x), is not a required property either.

Comparing trust expressions. In the following, we impose some additional structure on trust expressions by introducing a number of auxiliary operators. The first extension of the basic algebra allows us to compare trust expressions.

In order to evaluate the results of a trust calculation, one must be able to compare trust values. This will, for instance, allow one to select an agent that he considers most trusted for a specific task. Given the three-valued basis (v, β, δ) of our algebra, a one-dimensional measure on trust values will be insufficient. Therefore, we will introduce three different measures, one for each of the components. These measures will be formally modeled as total orders on trust expressions: \leq_u, \leq_b , and \leq_d (see Fig. 2).

The inequality $x \leq_u y$ expresses that the uncertainty component in expression x is at most as high as the uncertainty component in y. The inequality $x \leq_b y$ expresses that the belief component in expression x is at most as high as the belief component in y. Likewise, the inequality $x \leq_d y$ expresses that the disbelief component in expression x is at most as high as the disbelief component in expression x is at most as high as the disbelief component in y.

| $\leq_{u}: \mathcal{T} \times \mathcal{T}$ $\leq_{b}: \mathcal{T} \times \mathcal{T}$ | $\leq_{u}: \mathcal{T} \times \mathcal{T} \qquad \text{compare uncertainty(total order)} \\ \leq_{u}: \mathcal{T} \times \mathcal{T} \qquad \text{compare belief (total order)}$ | | | | |
|---|--|------------------------|----------------------------|--|--|
| $\leq_d : \mathcal{T} \times \mathcal{T}$ | compare disbelief (total order) | | | | |
| (T1) $x \leq_u v$ | (T4) $v \leq_b x$ | (T7) $v \leq_d x$ | (T10) $x + y \leq_u x$ | | |
| (T2) $\beta \leq_u x$ | (T5) $x \leq_b \beta$ | (T8) $\beta \leq_d x$ | (T11) $x \leq_u x \cdot y$ | | |
| (T3) $\delta \leq_u x$ | (T6) $\delta \leq_b x$ | (T9) $x \leq_d \delta$ | (T12) $y \leq_u x \cdot y$ | | |

Fig. 2. Axioms for the total orders (TO)

Axiom T1 states that full uncertainty v is the top element in the uncertainty order \leq_u , since it dominates all other elements. Axioms T2 and T3 state that full belief β and full disbelief δ do not express any uncertainty, and hence they are bottom elements. Axioms T4–T9 specify similar properties for the belief order (in which β is the top element) and the disbelief order (in which δ is the top element). Axiom T10 expresses a basic property of trust fusion, namely that uncertainty does not increase if we receive more information on the trustworthiness of a trustee. In presence of the symmetry axiom B1, this axiom is equivalent to $x+y \leq_u y$. Axioms T11 and T12 state a similar basic property for trust dilution: along a trust chain, uncertainty can only grow. The set of axioms T1–T12 forms the TO (for Total Order) extension of the BFD algebra.

The duality of belief and disbelief. The second extension of the algebra serves to express the duality of belief and disbelief. In order to express this duality, we introduce the *inverse* operator \overline{x} (see Fig. 3). This operator swaps the belief and disbelief components of a trust expression. Axiom I1 expresses the basic inversion property. Distributivity of inversion over fusion is expressed in axiom I2. This means that belief and disbelief are treated similarly when fusing trust opinions. Distributivity of inversion over dilution, $\overline{x \cdot y} = \overline{x} \cdot \overline{y}$, does not hold, because in a trust chain belief and disbelief are not each other's duals. Axiom I3 states that if we have full uncertainty (so no belief nor disbelief), the inverse operator has no effect. This also stresses that v is a zero element.

Axiom I4 expresses the duality of belief and disbelief. In presence of axiom I1 this axiom is equivalent to $\overline{\beta} = \delta$. Axioms I5 and I6 state that uncertainty is orthogonal to belief and disbelief. Finally, axiom I7 states that the inverse function swaps the belief and disbelief values of a trust expression. The set of

axioms I1–I7 forms the INV (for INVerse) extension of the BFD algebra. Using

| $\overline{}:\mathcal{T}\to\mathcal{T}$ | invert belief/disbelief (operator) | | |
|---|------------------------------------|------------------------------|--|
| (I1) $\overline{\overline{x}} = x$ | (I3) $\overline{v} = v$ | (I5) $x \leq_u \overline{x}$ | (I7) $x \leq_b y = \overline{x} \leq_d \overline{y}$ |
| (I2) $\overline{x+y} = \overline{x} + \overline{y}$ | (I4) $\overline{\delta} = \beta$ | (I6) $\overline{x} \leq_u x$ | |

Fig. 3. Axioms for the inverse operator (INV)

the dilution and inverse operators we now have that one constant suffices to define the other two. For instance, $v = \delta \cdot \delta$ and $\beta = \overline{\delta}$ (or $v = \overline{\beta} \cdot \overline{\beta}$ and $\delta = \overline{\beta}$). Further, by using I2 we achieve equivalence of axioms C4 and C6.

Full direct functional belief and disbelief. The third, and final, extension of our algebra concerns the meaning of full direct functional belief and disbelief. This means that we consider trust chains ending in β or δ , as in $x \cdot \delta$ and $x \cdot \beta$, which capture the situation that the last agent in a chain has full belief or disbelief in the trustee. Although it is tempting to set e.g. $x \cdot \beta = x$, we consider this as a too strong axiom. The belief component of $x \cdot \beta$ is clearly identical to the belief component of x, but their disbelief components are not. Therefore, we weaken this axiom to $x \cdot \beta =_b x$ (see axiom R1 in Fig. 4). Here we use $x =_b y$ as a shorthand notationfor $x \leq_b y \wedge y \leq_b x$. Thus, $x =_b y$ means that trust expressions x and y express equal belief. Likewise, we define $=_d$ and $=_u$.

Intuitively, axiom R1 states that if the last element in a chain has full belief, then belief of the whole chain is determined by the remainder of the chain. Axiom R2 states that in this case there is no disbelief. If we consider a trust chain ending in full disbelief, then the whole chain does not express any belief (axiom R3) and the disbelief expressed in the whole chain is exactly the belief that we have in the last agent before the trustee (axiom R4). We consider these axioms as a separate module since they are of a less basic nature. The set of axioms R1–R4 forms the RM (for Right Multiplication) extension of the BFD algebra.

| | (R1) $x \cdot \beta =_b x$ | (R2) $x \cdot \beta =_d \beta$ | (R3) $x \cdot \delta =_b \delta$ | (R4) $x \cdot \delta =_d \overline{x}$ |
|--|----------------------------|--------------------------------|----------------------------------|--|
|--|----------------------------|--------------------------------|----------------------------------|--|

Fig. 4. Axioms for right-multiplication (RM)

In the remaining of the paper we will refer to combinations of the rules of the basic algebra BFD and one or more of the extensions. For instance, BFD+TO+INV denotes the algebra consisting of the basic rules of BFD, and the extensions of TO and INV.

4 Models

In this section we investigate possible models of the algebra. First we look at small models containing three and four elements, and next we consider some models derived from the Subjective Logic. A distinguishing factor is the interpretation of the term $\beta + \delta$. The main purpose of this chapter is to show how

trust evaluation algorithms can be validated. In particular, several published variants of the Subjective Logic will be studied. We propose two new variants of the Subjective Logic which satisfy a larger set of axioms than the existing variants.

Given an algebra (Σ, E) , consisting of a signature Σ and a set of equations E, a model is a mathematical structure which interprets the sorts and functions of Σ as sets and (total) functions on these sets. This interpretation must be such that all equations of E are valid in the model. An equation t = t' (where t and t' are terms over the signature, possibly containing variables) is valid in a model if for every instantiation of the variables the interpretation of t is the same element as the interpretation of t'.

4.1 Three-element models

First, we investigate the canonical model of three distinct elements $\{b, d, u\}$ with interpretation $\beta \mapsto b, \delta \mapsto d, v \mapsto u$. This interpretation does not uniquely define the model, since there is still freedom in choosing a suitable interpretation of the fusion and dilution operators. Nevertheless, in order to satisfy the axioms, there is only little freedom left, namely in defining the outcome of b + d, which we consider a parameter σ of the model. Thus, we introduce the class of possible models $M_3(\sigma)$ in which $\sigma \in \{b, d, u\}$ represents the fusion of b and d. The interpretation of the operators is given in Fig. 5. The tables in this figure must be read in the order "row-operator-column". For instance, in the second table we can look up the value of $b \cdot d$ by crossing the row labeled b (first row) with the column labeled d (second column). This yields $b \cdot d = d$. The fusion table is determined, up to σ , by axioms B1, C1, C4, and C6. The dilution table is fully determined by axioms C3, C5, and C7. The inversion table is determined by axioms I1, I3, and I4.

$$\begin{array}{c|cccc} + & b & d & u \\ \hline b & b & \underline{\sigma} & b \\ d & \underline{\sigma} & d & d \\ u & b & d & u \end{array} \begin{array}{c} \cdot & b & d & u \\ \hline b & b & d & u \\ u & u & u & u \\ u & u & u & u \\ \end{array}$$

Fig. 5. The three-element models $M_3(\sigma)$

Next, we investigate possible choices for σ . $M_3(u)$ is not a model of BFD, because associativity of fusion yields the following derivation b = b + u = b + (b+d) = (b+b) + d = b + d = u. Thus $M_3(u)$ does not satisfy axiom B2.

On the other hand, a simple case distinction suffices to verify that $M_3(b)$ and $M_3(d)$ are indeed models of BFD. Surprisingly, these are not models of the extended algebra BFD+TO+INV (irrespective of the definition of the total orders). Using axiom I2 we can, e.g., derive the following equality for $M_3(d)$: $b = \overline{d} = \overline{b} + \overline{d} = \overline{b} + \overline{d} = d + b = d$.

The origin of the problem is in the requirement that belief and disbelief are treated equally by the fusion operator (axiom I2), which cannot be realized with three elements. By generalizing this reasoning we obtain the following impossibility result. **Theorem 1.** There does not exist a (non-trivial) three-element model of BFD+TO+INV.

4.2 Four-element models

As a consequence of the previous observations, we investigate somewhat richer models consisting of four elements $\{u, b, d, i\}$, where *i* denotes *inconsistency* or *contradiction*. The element *i* is used to give a meaning to $\beta + \delta$. The underlying idea is that if we receive fully certain, but contradictory information, we cannot combine this in a consistent way.

We interpret the constants as before $(\beta \mapsto b, \delta \mapsto d, v \mapsto u)$ and the operators as in Fig. 6. The table for the fusion operator follows from the fusion axioms in BFD. Observe that inconsistencies in a fusion are persistent (yielding e.g. i+u = i). The table for the dilution operator has three parameters, π , ρ , and σ . The other values are determined by the dilution axioms in BFD. The inequalities are straightforward; they express the inconsistent nature of i by assigning it minimal uncertainty, minimal belief and minimal trust.

We shall denote these models by $M_4(\pi, \rho, \sigma)$. They show some resemblance with Belnap's four-valued logic [2], but since the operators $_+_$ and $_\cdot_$ are different from logical conjunction and disjunction, $M_4(\pi, \rho, \sigma)$ is not isomorphic to Belnap's logic.

Fig. 6. The four-element models $M_4(\pi, \rho, \sigma)$

Using tool support to exhaustively verify all possible instantiations of π , ρ and σ , we found six different models satisfying all axioms: $M_4(u, u, u)$, $M_4(i, u, u)$, $M_4(i, u, i)$, $M_4(b, d, i)$, $M_4(i, d, b)$, and $M_4(i, d, i)$. The last three models, however, require a simplified definition of the total order, viz. one in which all elements are equivalent (e.g. $u =_b d =_b i =_b b$). The model $M_4(b, d, i)$ is isomorphic to a model defined by Gutscher [7]. The variety of models implies that there are several different ways to interpret the proliferation of inconsistencies in a dilution context.

Theorem 2. $M_4(u, u, u)$, $M_4(i, u, u)$, and $M_4(i, u, i)$ are models of BFD+ TO+ INV+ RM.

Proof. We will sketch the proof for $M_4(i, u, i)$. Axioms B1, C1–C7, T1–T9, I1, and I3–I6 follow easily by inspecting the tables. For instance, B1 follows from the symmetry of the table for +. Axiom B2 x + (y + z) = (x + y) + z follows from a simple case distinction. If any of x, y or z equals i or u, then associativity clearly holds. Next if x, y and z are all b or all d, then associativity is also simple. Finally,

if there is at least one b and at least one d and no u, then the associativity holds because the outcome is always i. The verification of axiom B3 (associativity of dilution) follows in a similar way, but needs some more case distinctions. Axiom T10 $x + y \leq_u x$ is true because every element of a row in the + table is \leq_u dominated by the left-hand argument. A similar check of the \cdot table suffices to verify axioms T11 and T12. Axiom I2 follows from a straightforward verification of all cases (ten cases, using symmetry of +). Axiom I7 clearly holds if x = u or x = i because they are minimal w.r.t. \leq_b and \leq_d . If x = b or x = d it follows from the duality of b and d. Axioms R1-R4 follow by simple inspection of the \cdot table. \Box

4.3 Subjective Logic

Subjective Logic [9] is a framework to compute the trust between two agents in a trust network. In its simplest form, a trust value is represented by a triplet (b, d, u), representing *belief*, *disbelief*, and *uncertainty*, respectively. These values satisfy $b, d, u \in [0, 1]$ and b + d + u = 1. Each such triplet can be represented as a point in a triangle.

Coordinate b of point p = (b, d, u) (see Fig. 7) determines the (perpendicular) distance between p and side DU. Likewise, d determines the distance between p and side BU, and u the distance between p and BD. Some examples: point B has coordinates (1, 0, 0) and represents full belief, the middle point between B and D is $(\frac{1}{2}, \frac{1}{2}, 0)$ and represents the fully certain opinion that there is as much belief as disbelief in the trustee. In



Fig. 7. The SL triangle

this framework, the fusion and dilution operators are called *consensus* (notation \oplus) and *conjunction* (notation \otimes). Reformulated in our notation, the operators are defined as follows.

$$(b, d, u) \oplus (b', d', u') = \left(\frac{bu' + b'u}{u + u' - uu'}, \frac{du' + d'u}{u + u' - uu'}, \frac{uu'}{u + u' - uu'}\right)$$
$$(b, d, u) \otimes (b', d', u') = (bb', bd', d + u + bu')$$

The fusion operator is undefined if and only if u = u' = 0 (assuming $u, u' \in [0,1]$). This indicates that the fusion of two dogmatic opinions (e.g. $(1,0,0) \oplus (0,1,0)$) is not straightforward. In order for the Subjective Logic to serve as a model of our algebra, the definition of the fusion operator must be extended. We investigate several extensions in the following sections.

The model SL_{γ} Recent versions of the Subjective Logic [13, 10] use a limit construction to define fusion for u = u' = 0.

$$(b,d,0) \oplus (b',d',0) = \left(\frac{\gamma b + b'}{\gamma + 1}, \frac{\gamma d + d'}{\gamma + 1}, 0\right)$$

According to [13], γ is defined by $\gamma = \lim_{u,u'\to 0} \left(\frac{u'}{u}\right)$. It expresses the relative dogmatism between the expressions (b, d, 0) and (b', d', 0) (or rather, between the agents expressing these dogmatic opinions). The higher the value of γ , the higher the relative weight of opinion (b, d, 0) in a fusion with (b', d', 0). The "default value" of γ is 1, meaning that in general dogmatic values are averaged. We shall denote this model by SL_{γ} .

In [10] it is stated that "in case of dogmatic opinions the associativity of the consensus operator does not emerge directly". Indeed, taking $\gamma = 1$, we can use associativity to derive $(\frac{1}{2}, \frac{1}{2}, 0) = (1, 0, 0) \oplus (0, 1, 0) = (1, 0, 0) \oplus ((0, 1, 0) \oplus (0, 1, 0)) \oplus ((0, 1, 0)) \oplus (0, 1, 0)) \oplus (0, 1, 0) = (\frac{1}{2}, \frac{1}{2}, 0) \oplus (0, 1, 0) = (\frac{1}{4}, \frac{3}{4}, 0)$. This gives the following negative result.

Theorem 3. SL_{γ} is not a model of the algebra BFD.

As a possible solution, Jøsang et al. [10] introduce an algorithm which can be used to calculate the fusion of three or more dogmatic beliefs. Expressions like $((1,0,0)\oplus(0,1,0))\oplus(0,1,0)$ are interpreted by applying a ternary fusion operator to these three arguments. This approach consists in fact of the introduction of a collection of n-ary fusion operators, each of which is still not associative. An additional problem, overlooked by the authors, is the fact that expressions must first undergo some kind of normalization procedure before their algorithm can be applied correctly. For instance, the expression $(0,1,0) \oplus ((1,0,0) \otimes ((\frac{1}{4},\frac{3}{4},0) \oplus (\frac{3}{4},\frac{1}{4},0)))$ must first be reduced to $(0,1,0) \oplus (((\frac{1}{4},\frac{3}{4},0) \oplus (\frac{3}{4},\frac{1}{4},0)))$ and not to $(0,1,0) \oplus ((1,0,0) \otimes (\frac{1}{2},\frac{1}{2},0))$. The former expression equals $(\frac{4}{12},\frac{8}{12},0)$, while the latter equals $(\frac{1}{4},\frac{3}{4},0)$. This normalization is not a simple innermost-first rewriting, but must make use of the property $(1,0,0) \oplus x = x$, which coincides with axiom C5 of our algebra.

We conclude by stating that, even though the authors claim associativity of their logic, this is not the case for any of the interpretations they give. Hence, SL_{γ} is not a model of BFD.

The model SL_c Based on the idea of taking the average of conflicting dogmatic beliefs, we define the extension SL_c . Elements of this model are fourtuples (b, d, u, c), where b, d, and u play their usual role and c counts the weight of a dogmatic opinion. Thus, we have the following set of trust atoms $\mathcal{A} = \{(b, d, u, c) \mid b, d, u \in [0, 1] \land c \in \mathbb{N}^+ \land b + d + u = 1 \land (u = 0 \lor c = 1)\}$. The last condition means that only opinions with uncertainty equal to zero can have a counter different from 1. Fusion and dilution are defined as follows.

$$(b, d, u, c) \oplus (b', d', u', c') = \begin{cases} \left(\frac{bc+b'c'}{c+c'}, \frac{dc+d'c'}{c+c'}, 0, c+c'\right) & \text{if } u = u' = 0\\ (b, d, u, c) & \text{if } u = 0, u' \neq 0\\ (b', d', u', c') & \text{if } u \neq 0, u' = 0\\ \left(\frac{bu'+b'u}{u+u'-uu'}, \frac{du'+d'u}{u+u'-uu'}, \frac{uu'}{u+u'-uu'}, 1\right) \text{ else} \end{cases}$$
$$(b, d, u, c) \otimes (b', d', u', c') = \begin{cases} (b', d', u', c') & \text{if } (b, d, u, c) = (1, 0, 0, c)\\ (bb', bd', d+u+bu', 1) & \text{else} \end{cases}$$

The first case of the fusion definition clarifies the role of the counter. If two dogmatic views are combined, then the resulting belief is calculated as the weighted average of the individual beliefs. The counter of the resulting trust value is the sum of the counters of the individual trust values. The other three cases are straightforward extensions of SL. For the dilution operator we treat the case where the left operand equals (1, 0, 0, c) differently. The reason is that in this case the resulting value must inherit the counter value of the right operand. This is motivated by the fact that (1, 0, 0, c) acts as a left-unit for dilution (cf. axiom C5 and the discussion on this axiom in the previous section).

We interpret the constants as follows: $\beta \mapsto (1,0,0,1), \delta \mapsto (0,1,0,1), v \mapsto (0,0,1,1)$. The inverse operator is defined by (b,d,u,c) = (d,b,u,c), and the total orders by: $(b,d,u,c) \leq_b (b',d',u',c') \Leftrightarrow b \leq b', (b,d,u,c) \leq_d (b',d',u',c') \Leftrightarrow d \leq d'$, and $(b,d,u,c) \leq_u (b',d',u',c') \Leftrightarrow u \leq u'$.

This model, which we call SL_c , satisfies all axioms except (C4) $\beta + \beta = \beta$ and (C6) $\delta + \delta = \delta$. This is because the weight of $\beta + \beta$ is higher than the weight of β : $(1, 0, 0, 1) \oplus (1, 0, 0, 1) = (1, 0, 0, 2)$.

However, SL_c satisfies two weaker axioms T13 and T14 (see Fig. 8). We use $x =_{bdu} y$ as a shorthand notation for $x =_b y \land x =_d y \land x =_u y$. It easily follows that axiom C4 implies T13 and that C6 implies T14. If we denote by BFD⁻ the

Fig. 8. Weakening axioms C4 and C6

axiom system BFD minus axioms C4 and C6, we can formulate the following theorem.

Theorem 4. SL_c is a model of the algebra $BFD^- + TO + INV + RM + T13 + T14$.

Proof. Axiom B1 follows by observing the symmetry in the definition of the fusion operator (e.g. $\frac{bc+b'c'}{c+c'} = \frac{b'c'+bc}{c'+c}$). The proof of axiom B2 consists of a case distinction and a number of straightforward calculations. As an illustration, we show the calculation for the first component b_1^+ of $((b, d, u, c) \oplus (b', d', u', c')) \oplus (b'', d'', u'', c'')$ if $u \neq 0, u' \neq 0$.

$$b_1^+ = \frac{\frac{bu'+b'u}{u+u'-uu'}u''+b''\frac{uu'}{u+u'-uu'}}{\frac{uu'}{u+u'-uu'}+u''-\frac{uu'}{u+u'-uu'}u''} = \frac{(bu'+b'u)u''+b''uu'}{uu'+u''(u+u'-uu')-uu'u''}$$

The first component b_2^+ of $(b, d, u) \oplus ((b', d', u') \oplus (b'', d'', u''))$ is

$$b_2^+ = \frac{b \frac{u'u''}{u'+u''-u'u''} + \frac{b'u''+b''u'}{u'+u''-u'u''}u}{u + \frac{u'u''}{u'+u''-u'u''} - u \frac{u'u''}{u'+u''-u'u''}} = \frac{bu'u'' + (b'u'' + b''u')u}{u(u'+u''-u'u'') + u'u'' - uu'u''}$$

It is easy to check that the resulting expressions are equal.

For axiom B3 (associativity of dilution) the most complex case is equality of the third component. For instance, if $(b, d, u) \neq (1, 0, 0), (b', d', u') \neq (1, 0, 0),$ and $(b'', d'', u'') \neq (1, 0, 0)$, the third component of $(b, d, u, c) \otimes ((b', d', u', c') \otimes$ (b'', d'', u'', c'') is d + u + b(d' + u' + b'u''). This is equal to the third component of $((b, d, u, c) \otimes (b', d', u', c')) \otimes (b'', d'', u'', c'')$, which is bd' + (d + u + bu') + bb'u''.

Axioms C1, C2, C3, C5, C7 and T1–T9 can be verified easily. For axiom T10 we have to consider four cases, three of which are trivial. The fourth case $(u \neq 0, u' \neq 0)$ is treated as follows: $\frac{uu'}{u+u'-uu'} \leq u \Leftrightarrow \frac{u'}{u+u'-uu'} \leq 1 \Leftrightarrow u' \leq u+u' - uu' \Leftrightarrow 0 \leq u - uu'$, which is true for $u, u' \in [0, 1]$. The most interesting case for axiom T11 is $(b, d, u) \neq (1, 0, 0)$. We then have $u \leq d + u + bu'$, which holds for $u, b, d, u' \in [0, 1]$. Likewise, axiom T12 follows from $u' = (d+u+b)u' = du' + uu' + bu' \leq d + u + bu'$. The remaining axioms T13, T14, I1–I7, and R1–R4 are trivial.

The model SL_i Finally, we will construct a model by extending SL with a constant for inconsistency (as for the M4-models in Section 4.2). We define $\mathcal{A} = \{(b, d, u) \mid b, d, u \in [0, 1] \land b + d + u = 1\} \cup \{i\}$. The fusion and dilution operators are a merger of their definitions in SL and M4(i, u, i).

$$\begin{split} i \oplus (b, d, u) &= (b, d, u) \oplus i = i \oplus i = i \\ (b, d, 0) \oplus (b, d, 0) &= (b, d, 0) \\ (b, d, 0) \oplus (b', d', 0) &= i & \text{if } b \neq b' \\ (b, d, u) \oplus (b', d', u') &= (\frac{bu' + b'u}{u + u' - uu'}, \frac{du' + d'u}{u + u' - uu'}, \frac{uu'}{u + u' - uu'}) & \text{if } u \neq 0 \lor u' \neq 0 \\ i \otimes (1, 0, 0) &= (1, 0, 0) \otimes i = i \otimes i = i \\ (b, d, u) \otimes i &= i \otimes (b, d, u) = (0, 0, 1) & \text{if } (b, d, u) \neq (1, 0, 0) \\ (b, d, u) \otimes (b', d', u') &= (bb', bd', d + u + bu') \end{split}$$

We interpret the constants as follows: $\beta \mapsto (1,0,0), \delta \mapsto (0,1,0), \upsilon \mapsto (0,0,1)$. The inverse operator is defined by $\overline{(b,d,u)} = (d,b,u)$ and $\overline{i} = i$. The total orders are given by:

 $\begin{array}{ll} (b,d,u) \leq_b (b',d',u') \Leftrightarrow b \leq b' & i \leq_b (b,d,u) \\ (b,d,u) \leq_d (b',d',u') \Leftrightarrow d \leq d' & i \leq_d (b,d,u) \\ (b,d,u) \leq_u (b',d',u') \Leftrightarrow u \leq u' & i \leq_u (b,d,u) \end{array}$

This model, which we call SL_i , satisfies all axioms. The proof follows the same line of reasoning as the proof of Theorem 4.

Theorem 5. SL_i is a model of the algebra BFD+TO+INV+RM.

All results of the current section are gathered in the table displayed in Fig. 9. The result for M_4 holds only for certain values of the parameters and the result for SL_c only for BFD⁻, T13, and T14.

| | BFD | ΤO | INV | RM |
|----------------------------------|--------------|--------------|--------------|--------------|
| $M_3(\sigma)$ | \checkmark | × | × | Х |
| $M_4(\pi, \rho, \sigma)$ and [7] | \checkmark | \checkmark | \checkmark | \checkmark |

| | BFD | TO | INV | RM |
|---------------|--------------|--------------|--------------|--------------|
| SL | (incom | plet | e mo | del) |
| SL_{γ} | × | \times | × | × |
| SL_c | $\rm BFD^-$ | \checkmark | \checkmark | \checkmark |
| SL_i | \checkmark | \checkmark | \checkmark | \checkmark |

Fig. 9. Results of the evaluation of the models (\checkmark means satisfied)

5 Related work

There are many existing models that propose ways to combine trust values or more widely to combine opinions. In the simplest models, the trust values are discrete or continuous values on a given interval (implying at least two elements in the model, i.e. a bottom and a top element). This is the case for instance in PGP [17].

Other models are taking the uncertainty into account, such as Subjective Logic [9], Dempster-Shafer [14], or Yager [15] to name only a few. The uncertainty level adds another dimension to the trust metrics and can profitably be used in order to compute more accurate trust values. Subjective Logic was extended several times, e.g. with a limit construction enabling the fusion of two fully certain opinions [10], with an algorithm enabling the commutativity of the fusion operator [10], with different operators definitions depending on the (partial) dependence of the trust values [13].

Nevertheless, the combination of trust values in 3-elements models can also lead to further questions. For instance, this raises the question on the dogmatic belief composition [16]. The extension to a 4-element model such as Gutscher's [7] (or based on Belnap's [2] or Bergstra's [3] theories) seems to answer this issue partially.

We noticed that the available models were developed in a bottom-up fashion, i.e. starting from the model and showing which properties it satisfies or not. To our knowledge, in the domain of trust, the development of a top-down approach such as the algebra proposed in the current paper, is novel. This algebra takes the fusion and dilution operators as a starting point since these are the common point of all the available models with only differences in their naming (fusion and dilution can respectively be referred to as consensus and recommendation in some models). The developed algebra focuses specifically on the trust application domain and all axioms of the algebra are motivated in this context. For this reason, the algebra may or may not make sense for other domains.

6 Conclusion

Taking the Subjective Logic as a starting point, we developed an abstract algebra expressing the basic properties of trust fusion and trust dilution. To the core of this algebra belong the three absolute trust values β , δ , and v. In a modular way, we extended this core algebra with some auxiliary operators to capture more properties of the operators involved. Since there are different ways to fuse dogmatic beliefs (such as considering $\beta + \delta$ as an inconsistency), we decided to not enforce one particular choice in the algebra. As a consequence, the algebra is not complete for any of the models studied. This is also reflected in the fact that the initial algebra (which we did not study in this paper) is not particularly interesting. An interesting next step would be to extend the algebra with additional properties (and possibly operators) that more precisely capture certain interpretations of the fusion of dogmatic beliefs, as to develop complete axiomatizations.

We studied two types of models of this algebra: canonical models with only a few elements, and models based on SL with an infinite number of elements. Partly to our surprise, there is no three-element model of the full algebra, indicating that the expression $\beta + \delta$ necessarily must be interpreted by a special fourth element.

As expected, SL with a partially defined fusion operator cannot be considered a model. This also applies to the SL model extended with a limit construction presented in [13], in contradiction with its (unproven) claim of associativity. More surprising is that the extension of SL with a limit construction is not a model because it lacks associativity of fusion. This contradicts the (unproven) claim of associativity in [13]. The algorithmic approach to associativity of fusion proposed in [10] does not imply associativity of the (binary) fusion operator either. In fact, while verifying the axioms of our algebra, it turned out that the reduction of terms according to axiom C5 is an essential, but omitted, step for the algorithm to work correctly. If this reduction step is not performed before evaluating a trust expression, then the algorithm does not take all +-related terms into account and gives the wrong result.

In order to overcome these problems, we experimented with two extensions of SL. The first extension tries to achieve the same results as SL with a limit construction by introducing a weight for dogmatic opinions. Due to this weight, which can be any positive natural number, the collection of possible interpretations of β becomes infinite and has no maximum element. As a consequence, axiom C4 which states that β is maximum, becomes invalid. This extension of SL satisfies a slightly weaker algebra. The second extension of SL concerns the introduction of a special element expressing inconsistency. This is a model of the full algebra.

In addition to this, the validation of the axioms of our algebra for SL also gives more insight in the properties that SL satisfies. Whereas e.g. associativity and commutativity have been discussed in detail by Jøsang et al., properties as expressed in e.g. T10–T12 have not been mentioned explicitly.

The proofs presented here mostly consist of a number of straightforward case distinctions. Rather than in the advanced level of the proofs, the complexity of our work lies in the design. A slight modification of the definition of e.g. M4 or SL_i will already invalidate essential properties like associativity.

An important next step is to validate other extensions of SL that were proposed in literature and to model other ways to deal with dogmatic beliefs. It is also interesting to look at more practical models, such as the model underlying PGP.

A particularly interesting model to investigate is the model of trust graphs (or transitive trust networks [11]). An open question is the reduction of such networks. Because not every trust graph can be represented as a trust expression, our theory has to be extended (e.g. with the notion of recursive equations) to deal with trust graphs.

Finally, we mention that our model does not consider dynamic aspects, such as the possible decay of trust or the occurrence of events that influence opinions. Extending our algebra in this direction would also be an interesting topic for future research.

Acknowledgment. This work as been partially funded by the Fonds National de la Recherche (Luxembourg), grant number TR-PDR BFR08-038

References

- Baptiste Alcalde, Eric Dubois, Sjouke Mauw, Nicolas Mayer, and Saša Radomirović. Towards a decision model based on trust and security risk management. In AISC 2009, volume 98. Australian Computer Society, 2009.
- N.D Belnap. A useful four-valued logic. In G. Epstein and J. Dunn, editors, Modern uses of multiple valued logics, pages 8–37. D. Reidel, Dordrecht, 1977.
- Jan A. Bergstra, Inge Bethke, and Piet Rodenburg. A propositional logic with 4 values: true, false, divergent and meaningless. *Journal of Applied Non-Classical Logics*, 5(2), 1995.
- Changyu Dong, Giovanni Russello, and Naranker Dulay. Trust transfer in distributed systems. In *Trust Management*, number 238/2007 in IFIP, pages 17–30. Springer, 2007.
- 5. Diego Gambetta, editor. Trust: Making and breaking cooperative relations. Department of Sociology, University of Oxford, 1988.
- Elizabeth Gray, Jean-Marc Seigneur, Yong Chen, and Christian Jensen. Trust propagation in small worlds. In *iTrust 2003*, volume 2692 of *LNCS*, pages 239– 254. Springer-Verlag Berlin Heidelberg, 2003.
- Andreas Gutscher. Reasoning with uncertain and conflicting opinions in open reputation systems. In STM 2008, Trondheim, Norway, 2008.
- Audun Jøsang. An algebra for assessing trust in certification chains. In Proceedings of the Network and Distributed Systems Security (NDSS), 1999.
- Audun Jøsang. A logic for uncertain probabilities. Int. J. Uncertain. Fuzziness Knowl.-Based Syst., 9(3):279–311, 2001.
- Audun Jøsang, Milan Daniel, and Patrick Vannoorenberghe. Strategies for combining conflicting dogmatic beliefs. In Proceedings of the 6th International Conference on Information Fusion, pages 1133–1140, 2003.
- 11. Audun Jøsang, Elizabeth Gray, and Michael Kinateder. Simplification and analysis of transitive trust networks. Web Intelli. and Agent Sys., 4(2):139–161, 2006.
- Audun Jøsang and Michael Kinateder. Analysing topologies of transitive trust. In Workshop of Formal Aspects of Security and Trust (FAST), pages 9–22, 2003.
- Audun Jøsang, Stephen Marsh, and Simon Pope. Exploring different types of trust propagation. In *iTrust 2006*, volume 3986 of *LNCS*, pages 179–192. Springer-Verlag Berlin Heidelberg, 2006.
- 14. G. Shafer. A Mathematical Theory of Evidence. Princeton Univ. Press, 1976.
- R.R. Yager. On the Dempster-Shafer framework and new combination rules. Information Sciences, 4:93–137, 1987.
- L.A. Zadeh. Review of mathematical theory of evidence by Glenn Shafer. AI Magazine, 5(3):81–83, 1984.
- 17. Philip R. Zimmermann. The Official PGP User's Guide. MIT Press, 1995.