Towards Argumentation with Symbolic Dempster-Shafer Evidence

Yuqing Tang [a,b] Chung-Wei Hang [c] Simon Parsons [b,d] Munindar Singh [c]
[a] *Robotics Institute, Carnegie Mellon University*
[b] *Department of Computer Science, Graduate Center, City University of New York*
[c] *Department of Computer Science, North Carolina State University*
[d] *Department of Computer & Information Science, Brooklyn College, Cty University of New York*

Abstract. This paper is concerned with the combination of argumentation with the Dempster-Shafer theory of evidence. We are motivated by the use of argumentation to reason about trust, and the desire to combine argumentation with an existing approach to propagating numerical measures of trust that is based on the Dempster-Shafer theory. We show how logical elements of evidence, associated with numerical degrees of belief, can be combined into arguments, and how these arguments can be related to the standard Dungian argumentation semantics.

1. Introduction

Trust is a mechanism for managing the uncertainty about autonomous entities and the information they deal with. As a result, trust can play an important role in any decentralized system, and particularly in multiagent systems, where agents are often engage in competitive interactions. Following Castelfranchi and Falcone [1], we believe that trust is based on reasons. We interpret this to mean that there is an advantage in clearly identifying the sources of information and relating these to the conclusions drawn from them, the sources and their connections to the conclusion being the reasons. In prior work [17] we have described how to track these relationships using argumentation, and to summarize the resulting connections as a graph. These connections can then be presented to individuals who have to make decisions based on information that comes from acquaintances of varying trustworthiness. A prototype system implementing these ideas is described in [18].

The formal system in [17] combines work on propagating trust through a social network with argumentation, showing how the results of this propagation can be linked to Dung-style [3] argumentation — where the arguments are structured as in [4,14]. The result is a system in which one agent can reason using information from other agents that it knows through the social network, assigning belief to that information depending on how much the agents that are the source of the information are trusted (as in [11]). A key issue in this work is the propagation of numerical measures of trust, which in our work is input from the social network, through the resulting argumentation. The system in [17] uses an approach based on possibility theory [2]. In this paper we consider how we can combine our system of argumentation with the Dempster-Shafer theory of evidence [15], in a way that seamlessly connects with the use of Demspter-Shafer theory to model trust in social networks [7].

2. Basic notation

2.1. A logical language

A predicate language $\mathcal{L}$ based on a set $\mathcal{P}$ of symbols with standard connectives $\wedge$, $\vee$, $\rightarrow$, $\neg$ and standard semantics is assumed in this work. We further constrain the domain of any term of a predicate in $\mathcal{P}$ to be finite and no functional symbols are allowed for any term of a predicate in $\mathcal{P}$. In this way, we will have a finite set of grounded predicates. For notational convenience, we also use $\mathcal{P}$ to denote the set of all grounded predicates.

The set of truth assignments to all ground predicates is denoted by $\Omega = 2^{\mathcal{P}}$ where Ω is taken as the *frame of discernment*. Following the standard semantics, every formula $\theta \in \mathcal{L}$ can be interpreted into a subset of truth assignments to $\mathcal{P}$, $\mathcal{I}(\theta) \subseteq \Omega$. Two special symbols for "false" and "true" are FALSE with $\mathcal{I}(\text{FALSE}) = \emptyset$, and TRUE with $\mathcal{I}(\text{TRUE}) = \Omega$.

Two formulae ϕ and φ, denoted by $\phi \equiv \varphi$, are equivalent iff $\mathcal{I}(\phi) = \mathcal{I}(\varphi)$, and an inference rule δ for $\mathcal{L}$ is of the form:

$$\delta = \frac{p_1, ..., p_m}{c}$$

where $p_1, .., p_m, c \in \mathcal{L}$. The p_i are the set of *premises* of the rule, and a specific p_i is denoted by $p_i(\delta)$. c is the *conclusion* of the rule, and is denoted by $c(\delta)$. The set of all valid rules is denoted by Δ.

2.2. Representing evidence

We start with a knowledge base $\mathbf{K} = \langle \Sigma, \Delta \rangle$ consisting of a set of formulae and a set of rules for reasoning with the formulae. $\Sigma = \{\langle h, E \rangle\}$ is the set of formulae, where each formula h is associated with some supporting evidence E, and $\Delta = \{\langle \delta, E \rangle\}$ is a the set of rules, where each rule δ is also associated with some supporting evidence. Our key notion is that of the *evidence argument*:

Definition 1 *An* evidence argument *is a pair $\langle h, E \rangle$, where h is a formula in $\mathcal{L}$ and $E = \{e_1, \ldots, e_n\}$ is a set of formulae in $\mathcal{L}$.*

E is called the *supporting evidence* for h, denoted by $E(h)$. Every $e_i \in E(h)$ is a indivisible chunk of information in the evidence for h, therefore it is called a *focal element* of the evidence h[1]. It is possible that $\{\langle h, E_1 \rangle, \langle h, E_2 \rangle\} \subseteq \Sigma$ for the same

[1]The term "focal element" is appropriated from Dempster-Shafer theory [15] since the e_i end up playing the same role as the focal elements do in that theory.

formula $h \in \mathcal{L}$ with two different sets of evidence E_1 and E_2. If such distinguishable repetitions occur, we assume that we can identify different occurrences of h in Σ as different pieces of information with the associated evidence denoted by $E(h)$.

The key idea here is that the evidence associated with a formula $\theta \in \mathcal{L}$ or a rule $\delta \in \Delta$ summarises the data that shows the rule or formula holds. When we reason with the formulae, which we do by using the rules, we will then propagate the evidence, and so obtain the evidence that supports any conclusions. For every pair $\langle h, E \rangle$ it is then the case that:

1. $h = \theta \in \mathcal{L}$ or $h = \delta \in \Delta$; and
2. $E = \{e_1, ..., e_n\}$ is a set of evidence for h such that $e_1, ..., e_n \in \mathcal{L}$, $e_i \not\equiv e_j$ for any $i \neq j$.

In addition we assume the existence of a *probability mass function* $m(E) : E \mapsto [0, 1]$ defined on E which satisfies the constraint:

$$m(E, e_1) + ... + m(E, e_n) = 1$$

and for all $\phi \notin E$, we set $m(E, \phi) = 0$. In other words we associate some measure of belief $m(\cdot)$ wth every item of evidence, with the goal that from these we can calculate a measure for every h.

In the original Dempster-Shafer theory [15], it is this probability mass that is the focus, and it is the probability mass that constitutes the evidence. In the work we present here, the evidence is, as just outlined, a combination of logical statements over which a probability mass can be defined. This is the reason that we distinguish our work as being about *symbolic* Dempster Shafer theory. As in standard Dempster-Shafer theory, we use the probability mass to determine how much certain interesting hypothesese are believed. In our case, these hypotheses are the conclusions of arguments

Definition 2 *Given an evidence argument $A = \langle h, E \rangle$ for a formula $h \in \mathcal{L}$, the* belief $b(h)$, disbelief $d(h)$, *and the* uncertainty $u(h)$ *of h are computed as follows:*

$$\begin{aligned} b(h) &= \Sigma_{\mathcal{I}(e_i) \subseteq \mathcal{I}(h)} m(E, e_i) && = \Sigma_{e_i \vdash h} m(E, e_i) \\ d(h) &= \Sigma_{\mathcal{I}(e_i) \cap \mathcal{I}(h) = \emptyset} m(E, e_i) && = \Sigma_{e_i \vdash \neg h} m(E, e_i) \\ u(h) &= \Sigma_{\mathcal{I}(e_i) \cap \mathcal{I}(h) \neq \emptyset} m(E, e_i) && = \Sigma_{e_i \nvdash h \textit{ and } e_i \nvdash \neg h} m(E, e_i) \end{aligned}$$

In other words (in the formulation on the left above), the belief in h is the sum of the mass of the all focal elements in E. that are part of the evidence for h, the disbelief in h is the sum of all the mass for all the focal elements that are evidence for $\neg h$, and the uncertainty is the sum of all the mass that for evidence that is assigned to neither h or $\neg h$. Equivalently (in the formulation on the right above), the belief in h is the sum of the mass of all the formulae that imply h, the disbelief in h is the sum of the mass of all the formulae that imply $\neg h$, and the uncertainty is the sum of the mass of the formulae that imply neither h nor $\neg h$.

	p	q	
$\mathcal{I}(\neg p \wedge \neg q)$	0	0	
$\mathcal{I}(\neg p \wedge q)$	0	1	
$\mathcal{I}(p \wedge \neg q)$	1	0	$\}\mathcal{I}(h_1)$
$\mathcal{I}(p \wedge q)$	1	1	

(a)

	p	q	
$\mathcal{I}(\neg p \wedge \neg q)$	0	0	
$\mathcal{I}(\neg p \wedge q)$	0	1	$\}\mathcal{I}(h_2)$
$\mathcal{I}(p \wedge \neg q)$	1	0	
$\mathcal{I}(p \wedge q)$	1	1	$\}\mathcal{I}(h_2)$

(b)

Figure 1. Truth tables for Example 1. (a) Truth table for h_1. Obviously, $b(h_1) = m(E_1, p)$ because $\mathcal{I}(p) \subseteq \mathcal{I}(h_1)$, and $d(h_1) = m(E_1, \neg p \wedge q)$ because $\mathcal{I}(\neg p \wedge q) \cap \mathcal{I}(h_1) = \emptyset$. (b) Truth table for h_2. We can see that $b(h_2) = m(E_2, \neg p \wedge q)$ because $\mathcal{I}(\neg p \wedge q) \subseteq \mathcal{I}(h_2)$, $d(h_2) = m(E_2, \neg q)$ because $\mathcal{I}(\neg q) \cap \mathcal{I}(h_2) = \emptyset$, and $u(h_2) = m(E_2, p)$ because $\mathcal{I}(p) \cap \mathcal{I}(h_2) \neq \emptyset$.

Example 1 *Let* $\langle h_1, E_1 \rangle = \langle p, \{p, \neg p \wedge q\} \rangle$*, where* $m(E_1, p) = 0.4$ *and* $m(E_1, \neg p \wedge q) = 0.6$*. Then, as explained in Table 1(a):*

$$b(h_1) = m(E_1, p) = 0.4, d(h_1) = m(E_1, \neg p \wedge q) = 0.6, \text{ and } u(h_1) = 0.$$

Let $\langle h_2, E_2 \rangle = \langle q, \{\neg p \wedge q, \neg q, p\} \rangle$*, where* $m(E_2, \neg p \wedge q) = 0.5$*,* $m(E_2, \neg q) = 0.3$*, and* $m(E_2, p) = 0.2$*. Then, as explained in Table 1(b):*

$$b(h_2) = m(E_2, \neg p \wedge q) = 0.5, d(h_2) = m(E_2, \neg q) = 0.3, \text{ and } u(h_2) = m(E_2, p) = 0.2.$$

3. Combining evidence

In reasoning about complex situations, we often find that we have to deal with evidence from multiple sources. There are a number of ways that we can combine evidence, and different methods lead to different ways of calculating the probability mass of each focal element under the combined set of evidence. Here we discuss two ways of combining evidence. Section 3.1 describes a conjunctive combination and Section 3.2 describes disjunctive combination.

3.1. Conjunctive evidence combination

Conjunctive combination is the approach used to combine evidence in a way that has the characteristics of a logical and. For example, it is appropriate to use it to fuse data from two sensors which both have to indicate the presence of some object in order for us to infer that the object is there. Suppose one sensor collects evidence of intrusion by detecting movement, a second collects evidence based on temperature, and we need to establish the movement of a warm body to detect intrusion. In this case, conjunctive combination should be adopted to combine evidence from the two sensors.

Given a list of formula $\{p_1, ..., p_k\}$ with a set of independent pieces of evidence $\{E_1, , ..., E_k\}$ respectively, we can combine the evidence into a set of evidence for the whole list $\{p_1, ..., p_k\}$ in a conjunctive way :

$$E(p_1 \otimes p_2 \otimes ... \otimes p_k) = E_1 \otimes E_2 \otimes \ldots \otimes E_k = \{e'\}$$

such that $e' = e_{1,j_1} \wedge e_{2,j_2} \wedge ... \wedge e_{k,j_k}$ where $e_{i,j_i} \in E_i$, and

$$m(E(p_1 \otimes ... \otimes p_k), e') = \frac{\Sigma_{e' \equiv \bigwedge_{i=1,\ldots,k} e_{i,j_i}} \Pi_{i=1,\ldots,k} m(p_i, e_{i,j_i})}{\Sigma_{\langle e_{1,j_1},\ldots,e_{k,j_k}\rangle \in \Pi_{i=1,\ldots,k} E_i} \Pi_{i=1,\ldots,k} m(p_i, e_{i,j_i})},$$

Example 2 (Conjunctive combination) *Let* $E_1 = \{p, \neg p \wedge q\}$ *and* $E_2 = \{\neg p \wedge q, \neg q, p\}$ *with* $m(E_1, p) = 0.4$*,* $m(E_1, \neg p \wedge q) = 0.6$*,* $m(E_2, \neg p \wedge q) = 0.5$*,* $m(E_2, \neg q) = 0.3$ *and* $m(E_2, p) = 0.2$*. Then*

$$\begin{aligned} E_1 \otimes E_2 = \{&p \wedge (\neg p \wedge q), p \wedge \neg q, p \wedge p, \neg p \wedge q \wedge \neg p \wedge q, \neg p \wedge q \wedge \neg q, \neg p \wedge q \wedge p\} \\ = \{&\perp, p \wedge \neg q, p, \quad \neg p \wedge q, \perp, \perp\} \\ = \{&\perp, p \wedge \neg q, p, \neg p \wedge q\} \end{aligned}$$

where

$$\begin{aligned} m(E_1 \otimes E_2, \perp) &= 0.4 \times 0.5 + 0.6 \times 0.3 + 0.6 \times 0.2 = 0.5 \\ m(E_1 \otimes E_2, p \wedge \neg q) &= 0.4 \times 0.3 = 0.12 \\ m(E_1 \otimes E_2, p) &= 0.4 \times 0.2 = 0.08 \\ m(E_1 \otimes E_2, \neg p \wedge q) &= 0.6 \times 0.5 = 0.3 \end{aligned}$$

3.2. Disjunctive evidence combination

The Dempster-Shafer theory also provides a disjunctive rule for combining evidence, that is combining evidence in a way akin to the logical or operation. To extend our previous example of intrusion detection, the disjunctive rule would be appropriate if either detecting motion or detecting warmth was enough to indicate an intrusion.

$$E(p_1 \oplus p_2 \oplus ... \oplus p_k) = E_1 \oplus E_2 \oplus \ldots \oplus E_k = E_1 \cup E_2 \cup ... \cup E_k$$

such that, for each $e \in E(p_1 \oplus p_2 \oplus ... \oplus p_k)$, we set

$$m(E(p_1 \oplus p_2 \oplus ... \oplus p_k), e) = \frac{m(E_1, e) + m(E_2, e) + ... + m(E_k, e)}{\Sigma_{i=1,\ldots,k} \Sigma_{e_{i,j} \in E_i} m(E_i, e_{i,j})},$$

Example 3 (Disjunctive combination) *Let* $E_1 = \{p, \neg p \wedge q\}$ *and* $E_2 = \{\neg p \wedge q, \neg q, p\}$ *where* $m(E_1, p) = 0.4$*,* $m(E_1, \neg p \wedge q) = 0.6$*,* $m(E_2, \neg p \wedge q) = 0.5$*,* $m(E_2, \neg q) = 0.3$ *and* $m(E_2, p) = 0.2$*. Then:*

$$E_1 \oplus E_2 = \{p, \neg p \wedge q, \neg q\}$$

where

$$m(E_1 \oplus E_2, p) = \frac{0.4 + 0.2}{2} = 0.3$$

$$m(E_1 \oplus E_2, \neg p \wedge q) = \frac{0.6 + 0.5}{2} = 0.55$$

$$m(E_1 \oplus E_2, \neg q) = \frac{0.3}{2} = 0.15$$

Note that both the conjunctive and disjunctive combinations assume that evidence is independent and this is sufficient for our needs in this paper. Dempster-Shafer can handle dependent evidence, but needs to extend the approach we adopt here, for example as in [10,16].

4. Deductive reasoning with evidence

Given that we want to combine the use of evidence with logical reasoning, the combination rules introduced above do not help us directly. Rather we need to build on them to create combination rules for logical combinations. That is the subject of this section.

Property 3 Not $\neg$. *Given an evidence argument* $\langle h, E\rangle$*, we can derive an evidence argument* $\langle \neg h, E\rangle$ *for* $\neg h$ *such that:*

$$b(\neg h) = d(h) = \Sigma_{\mathcal{I}(e_i) \cap \mathcal{I}(h) = \emptyset} m(E, e_i)$$

$$d(\neg h) = b(h) = \Sigma_{\mathcal{I}(e_i) \subseteq \mathcal{I}(h)} m(E, e_i)$$

$$u(\neg h) = u(h) = \Sigma_{\mathcal{I}(e_i) \cap \mathcal{I}(h) \neq \emptyset} m(E, e_i)$$

In this case h and $\neg h$ share the same evidence, but the belief and disbelief will be computed differently given that evidence. Note that h and $\neg h$ share the same uncertainty.

The proofs of all the properties in this section follow quickly from the definitions introduced above, and are omitted in the interests of space.

Property 4 And $\wedge$. *Given two evidence arguments* $\langle h_1, E_1\rangle$ *and* $\langle h_2, E_2\rangle$ *with independent evidence, we can derive evidence argument* $\langle h_1 \wedge h_2, E\rangle$ *for* $h_1 \wedge h_2$ *where* $E = E_1 \otimes E_2$*.*

Example 4 *Following Example 1, let* $\langle h_1, E_1\rangle = \langle p, \{p, \neg p \wedge q\}\rangle$*, where* $m(h_1, p) = 0.4$ *and* $m(h_1, \neg p \wedge q) = 0.6$*, and* $\langle h_2, E_2\rangle = \langle q, \{\neg p \wedge q, \neg q, p\}\rangle$*, where* $m(h_2, \neg p \wedge q) = 0.5$*,* $m(h_2, \neg q) = 0.3$*, and* $m(h_2, p) = 0.2$*. Then,*

$$\begin{aligned}\langle h_1 \wedge h_2, E_1 \otimes E_2\rangle &= \langle p \wedge q, \{p \wedge \neg p \wedge q, p \wedge \neg q, p, \neg p \wedge q, \neg p \wedge q \wedge \neg q, p \wedge \neg p \wedge q\}\rangle \\ &= \langle p \wedge q, \{\bot, p \wedge \neg q, p, \neg p \wedge q\}\rangle\end{aligned}$$

where $m(E_1 \otimes E_2, \bot) = 0.5$*,* $m(E_1 \otimes E_2, p \wedge \neg q) = 0.12$*,* $m(E_1 \otimes E_2, p) = 0.8$*, and* $m(E_1 \otimes E_2, \neg p \wedge q) = 0.3$*. Then,*

$$
\begin{aligned}
b(h_1 \wedge h_2) &= 0 \\
d(h_1 \wedge h_2) &= m(E_1 \otimes E_2, \perp) + m(E_1 \otimes E_2, p \wedge \neg q) + m(E_1 \otimes E_2, \neg p \wedge q) \\
&= 0.92 \\
u(h_1 \wedge h_2) &= m(E_1 \otimes E_2, p) = 0.08.
\end{aligned}
$$

Property 5 Or $\vee$. *Given two evidence arguments $\langle h_1, E_1\rangle$ and $\langle h_2, E_2\rangle$ with independent evidence, we can derive an evidence argument $\langle h_1 \vee h_2, E\rangle$ for $h_1 \vee h_2$ where $E = E_1 \oplus E_2$.*

Example 5 *Following Example 1, let $\langle h_1, E_1\rangle = \langle p, \{p, \neg p \wedge q\}\rangle$, where $m(E_1, p) = 0.4$ and $m(E_1, \neg p \wedge q) = 0.6$, and $\langle h_2, E_2\rangle = \langle q, \{\neg p \wedge q, \neg q, p\}\rangle$, where $m(E_2, \neg p \wedge q) = 0.5$, $m(E_2, \neg q) = 0.3$, and $m(E_2, p) = 0.2$. Then,*

$$
\langle h_1 \vee h_2, E_1 \oplus E_2\rangle = \langle p \vee q, \{p, \neg p \wedge q, \neg q\}\rangle,
$$

where $m(E_1 \oplus E_2, p) = 0.3$, $m(E_1 \oplus E_2, \neg p \wedge q) = 0.55$, and $m(E_1 \oplus E_2, \neg q) = 0.15$. Then,

$$
\begin{aligned}
b(h_1 \vee h_2) &= m(E_1 \oplus E_2, p) = 0.3 \\
d(h_1 \vee h_2) &= m(E_1 \oplus E_2, \neg q) = 0.15 \\
u(h_1 \vee h_2) &= m(E_1 \oplus E_2, \neg p \wedge q) = 0.55.
\end{aligned}
$$

Property 6 Implication $\rightarrow$. *Given two evidence argument $\langle h_1, E_1\rangle$ and $\langle h_2, E_2\rangle$ with independent evidence, we can derive evidence argument for $h_1 \rightarrow h_2$: $\langle h_1 \rightarrow h_2, E\rangle$ where $E = E_1 \oplus E_2$.*

Thus evidence for implication is combined just like evidence for $\vee$.

This set of properties tells us how to construct logical combinations of arguments. From the arguments $\langle h_1, E_1\rangle$ and $\langle h_2, E_2\rangle$ we can now derive arguments for $\neg h_1$, $h_1 \wedge h_2$, $h_1 \vee h_2$ and $h_1 \rightarrow h_2$. However, this is not sufficient to allow us to do much in the way of argument construction, since that typically involves the use of rules of inference. As a result, we consider how evidence arguments combine when using modus ponens and a form of generalized modus ponens that allows inference with the rules from Δ:

Property 7 Modus Ponens. *Given two evidence argument $\langle h_1, E_1\rangle$ and $\langle h_1 \rightarrow h_2, E_2\rangle$ with independent evidences, we can derive an evidence argument $\langle h_2, E\rangle$ for h_2 where $E = E(h_1) \otimes E(h_1 \rightarrow h_2)$.*

Property 8 GMP. *Given m evidence arguments $\langle h_1, E_1\rangle$, $\langle h_2, E_2\rangle$,..., $\langle h_m, E_m\rangle$, and an evidence argument for a rule:*

$$
\langle \delta = \frac{h_1, ..., h_m}{h}, E_\delta\rangle
$$

with independent evidence, we can derive an evidence argument $\langle h, E\rangle$ for h where $E = E_1 \otimes E_2 \otimes \ldots \otimes E_m \otimes E_\delta)$.

Thus both rules combine evidence using the conjunctive rule. Note that while GMP has the form of the generalized modus ponens rule, it differs in that our rules δ are defeasible rules, not material implications.

5. Propagating evidence within an argumentation framework

Having shown how to combine evidence during inference, we can go on to consider how evidence is propagated as arguments are constructed. The idea here is that propagation of evidence allows us to to identify conclusions along with the evidence that supports them. This not only allows us to compute the belief in conclusions (given the probability mass on the evidence), but—unlike other approaches which only manipulate the numerical values during reasoning—allows the symbolic evidence itself to be used in reaching further conclusions. Using the framework from [17], we consider an argument to be a graph constructed by chaining rules from Δ together:

Definition 9 *A* rule network $\mathcal{R}$ *is a directed hypergraph* $\langle V^r, E^r\rangle$ *where (1) the set of vertices* V^r *are elements of* $\mathcal{L}$*; (2) the set of edges* E^r *are inference rules* δ*; (3) the initial vertices of an edge* $e \in E^r$ *are the premises of the corresponding rule* δ*; and (4) the terminal node of that edge is the corresponding conclusion* c*.*

Definition 10 *For a given knowledge base* $\mathbf{K} = \langle \Sigma, \Delta\rangle$*, a rule network* $\mathcal{R} = \langle V^r, E^r\rangle$ *is a* proof network *if and only if every premise of each* $\delta \in E^r$ *is either a member of* Σ *or the conclusion of some* $\delta' \in E^r$*.*

Definition 11 *An* tree argument A *from a knowledge base* $\mathbf{K}$ *and a rule base* Δ *is a pair* $\langle h, E\rangle$ *where* $E = \langle V^r, E^r\rangle$ *is a proof network for* h*, and* h *is the only leaf of* E*.*

In accordance with the usual terminology, $E = \langle V^r, E^r\rangle$ is the *support* of the argument, and h is the *conclusion.* $C(H)$ is the set of *intermediate conclusions* of H, the set of all the conclusions of the $\delta \in E^r$ other than h. $P(H)$ is the set of *pure premises* of H, the premises of the $\delta \in E^r$ that aren't intermediate conclusions of H.

Note that $\langle h, \langle V^r, E^r\rangle\rangle$ is just an evidence argument, albeit one with some additional structure in terms of the associated graph. This structure places some restrictions on what arguments meet the conditions of Definition 11, and not all evidence arguments will be tree arguments. Given that the construction of a tree argument is equivalent to repeated applications of the GMP rule from Property 8, we can easily obtain a form of soundness result:

Proposition 12 *If an argument* $\langle h, \langle V^r, E^r\rangle\rangle$ *is constructed from a knowledge base* $\mathbf{K}$*, then the conclusion* h *follows from a sequence of rule applications in* Δ *and the premises are grounded on facts in* Σ*. As these reasoning steps are recorded in* $\langle V^r, E^r\rangle$*, the evidence* $E(h)$ *for the conclusion* h *is combined from the evidences of the facts and rules using the combination operators in the same order as they are recorded into the argument.*

In addition, we can clearly compute the evidence associated with h from the evidence arguments for the elements of H:

Proposition 13 *If an argument $\langle h, \langle V^r, E^r\rangle\rangle$ is constructed from a knowledge base* $\mathbf{K}$*, then:*

$$E(h) = \bigotimes_{p \in P(h)} E(p) \otimes \bigotimes_{\delta \in E^r} E(\delta)$$

This follows immediately from Property 8 and the structure of $\langle h, \langle V^r, E^r\rangle\rangle$.

Finally, the system is also complete in the sense that repeated applications of the GMP scheme will allow all possible tree arguments to be constructed from the knowledge base $\mathbf{K}$. This in turn follows from the completeness of generalized modus ponens as a rule of inference for Horn clauses, the fact that our rule base Δ is a set of augmented Horn clauses, and the fact that our GMP rule provides exactly the same inference for these rules as generalized modus ponens does for standard Horn clauses.

6. Argumentation semantics

A key notion in argumentation is that arguments *defeat* one another — that is, one argument casts doubt on another by, for example, casting doubt on one of the premises of the second argument — and that it is possible to take a set of arguments that interact in this way and extract a coherent subset. In this section we describe how this coherent set may be extracted for sets of arguments constructed using our system of argumentation.

We start by distinguishing a number of ways that a defeat may occur:

Definition 14 *An argument $\langle h_1, H_1\rangle$* defeats *an argument $\langle h_2, H_2\rangle$ if it rebuts, premise-undercuts, intermediate-undercuts, or inference-undercuts it, where: (1) An argument $\langle h_1, H_1\rangle$* rebuts *another argument $\langle h_2, H_2\rangle$ iff $h_1 \equiv \neg h_2$; (2) An argument $\langle h_1, H_1\rangle$* premise-undercuts *another argument $\langle h_2, H_2\rangle$ iff there is a premise $p \in P(H_2)$ such that $h_1 \equiv \neg p$; (3) An argument $\langle h_1, H_1\rangle$* intermediate-undercuts *another argument $\langle h_2, H_2\rangle$ iff there is an intermediate conclusion $c \in C(H_2)$ such that $c \neq h_2$ and $h_1 \equiv \neg c$; and (4) An argument $\langle h_1, H_1\rangle$* inference-undercuts *another argument $\langle h_2, H_2\rangle$ iff there is an inference rule $\delta \in \Delta(H_2)$ such that $\delta = \frac{p_1, \ldots, p_n}{c}$ and $h_1 \equiv \neg(p_1 \wedge \ldots \wedge p_n \rightarrow c)$.*

In any case in which $\langle h_1, H_1\rangle$ defeats $\langle h_2, H_2\rangle$, $\langle h_1, H_1\rangle$ is said to be a *defeater* of $\langle h_2, H_2\rangle$, and $\langle h_2, H_2\rangle$ is said to be the *defeatee*. The relation defeat collects all pairs $(\langle h_1, H_1\rangle, \langle h_2, H_2\rangle)$ such that $\langle h_1, H_1\rangle$ defeats $\langle h_2, H_2\rangle$. From Dung [3], we have the following component definitions, all of which hold for the system of argumentation we have described.

Definition 15 *An* argumentation framework *is a pair, $Args = \langle \mathsf{ARG}, \mathsf{DFT}\rangle$, where* ARG *is a set of arguments, and* DFT *is the binary relation* defeat *over the arguments.*

We can derive a preference relation PREF over the constructed arguments to capture relative strength of the arguments derived from the evidence:

Definition 16 *Given two arguments $A_1 = \langle h_1, H_1 \rangle$ with evidence E_1 for h_1 derived from H_1 and $A_2 = \langle h_2, H_2 \rangle$ with evidence E_2 for h_2 derived from H_2, we can define a preference PREF as $(A_1, A_2) \in \mathsf{PREF}$ iff (1) $b(E_1, h_1) > b(E_2, h_2)$, or (2) $b(E_1, h_1) = b(E_2, h_2)$ and $u(E_1, h_1) > u(E_2, h_2)$*

In other words one argument is preferred over another if the evidence for it has a higher degree of belief, or it has the same degree of belief and more uncertainty. The reason for preferring an argument with more uncertainty is that this means there is less disbelief in the argument (or, alternatively, more belief that can possibly be assigned to the argument by further evidence).

Definition 17 *Let PREF be a preference relation and DFT be a defeat relation on a set ARG of arguments. A* preference-refined defeat relation *PDFT can be defined as the following for any two arguments A_1 and A_2 in ARG*

$$(A_1, A_2) \in \mathsf{PDFT} \textit{ iff } (A_1, A_2) \in \mathsf{DFT} \textit{ but } (A_2, A_1) \notin \mathsf{PREF}.$$

With the concept of preference-refined defeat relation PDFT, we can translate a preference-based argumentation framework $\mathsf{AFP} = \langle \mathsf{ARG}, \mathsf{DFT}, \mathsf{PREF} \rangle$ into an argumentation framework $\mathsf{AFD(AFP)} = \langle \mathsf{ARG}, \mathsf{PDFT} \rangle$. in essence by discarding defeat relations where the defeater is not preferred to the defeatee. We can then quickly get to a standard definition of acceptability.

Definition 18 *Let $\langle \mathsf{ARG}, \mathsf{DFT} \rangle$ be an argumentation framework, and $S \subseteq \mathsf{ARG}$. An argument A is defended by S iff $\forall B \in \mathsf{ARG}$ if $(B, A) \in \mathsf{DFT}$ then $\exists C \in S$ such that $(C, B) \in \mathsf{DFT}$.*

Definition 19 *$S \subseteq \mathsf{ARG}$. $\mathcal{F}_{\mathcal{R}}(S) = \{A \in \mathsf{ARG} | A$ is defended by S with respect to $\mathsf{DFT}\}$.*

For a function $F : D \to D$ where D is the domain and the range of the function, a fixed point of F is an $x \in D$ such that $x = F(x)$. When the D is associated with an ordering P — for example, P can be set inclusion over the power set D of arguments — x is a *least fixpoint* of F if x is a least element of D with respect to P and x is a fixed point.

Definition 20 *Let $\langle \mathsf{ARG}, \mathsf{DFT} \rangle$ be an argumentation framework. The set of acceptable arguments, denoted by $Acc^{F}_{\mathcal{R}}$, is the least fixpoint of the function $\mathcal{F}_{\mathcal{R}}$ with respect to set inclusion.*

With this final definition we obtain a notion of acceptability for our argumentation system, and hence a connection to the standard Dungian approach to argumentation semantics [3]. Thus we have shown how the system we have introduced here can be fitted within more standard notions of argumentation.

7. Related work

This work starts with the association of basic probability mass with elements of a knowledge base and show how beliefs measures may be derived for arguments constructed from that knowledge base, and then how those measures can be used to construct a preference-based argumentation system. In this way our work connects with existing approaches to preference-based argumentation, but allows the preference order to be established from what the agent knows, rather than assuming the existence of some pre-defined order. (And as discussed in [6], the measures we use here can be learned by the agent).

This work also connects to approaches that combining argumentation and Dempster-Shafer theory. [5,9] showed that it was possible to associate probability mass with formulae, reason with the formulae, and compute measures like belief in the conclusions of the reasoning. However, this approach has a limited notion of an argument — an argument is just a conjunction of literals — and the work is only concerned with the construction of arguments and the computation of belief. In considering the connection to Dungian semantics, our work goes beyond [5,9].

A more recent approach to combining argumentation and Dempster-Shafer theory is [12], which builds on subjective logic [8], a logic incorporates measures from Dempster-Shafer theory. [12] established argumentation semantics solely based on the evidence and belief/disbelief/uncertainty, but its connection to Dung's argumentation semantics is not clear [12, Section 5], and, like [5,9], the focus is more on establishing the strength of individual arguments. We believe our approach has a stronger connection with the standard argumentation semantics, as well as fitting into our graphical approach to argumentation [17].

8. Summary

This paper has provided the foundations for a system of argumentation that combines logical reasoning and the Dempster-Shafer theory. The system allows arguments to be constructed from formulae in a predicate logic, each of which has a numerical measure associated with it, measures expressed using the Dempster-Shafer theory. These measures are appropriately combined as the formulae are constructed into arguments. The work is set in the context of a wider effort to use argumentation to reason in an environment when sources of information are of varying trustworthiness. Well-founded approaches to reasoning about the trust in individuals have already been established and several of us have developed an approach that quantifies trust using the Dempster-Shafer theory [7]. Other among us have described how such reasoning about trust can be used as input for argumentation-based reasoning [13]. The work decribed here will allow for a seamless integration by providing a means to propagate the values established by the trust reasoning system through the resulting argumentation.

Acknowledgements

This research was funded under Army Research Laboratory and Cooperative Agreement Number W911NF-09-2-0053. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

References

[1] C. Castelfranchi and R. Falcone. Trust is much more than subjective probability: Mental components and sources of trust. In *Proceedings of the 33rd Hawaii International Conference on System Science*, Maui, Hawai'i, January 2000. IEEE Computer Society.

[2] D. Dubois and H. Prade. *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. Plenum Press, New York, NY, 1988.

[3] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357, 1995.

[4] A. J. Garcia and G. R. Simari. Defeasible logic programming: an argumentative approach. *Theory and Practice of Logic Programming*, 4(2):95–138, 2004.

[5] R. Haenni, J. Kohlas, and N. Lehmann. Probabilistic argumentation systems. In *Handbook of Defeasible Reasoning and Uncertainty Management Systems*. Kluwer, Dordrecht, 2000.

[6] C.-W. Hang, Y. Wang, and M. P. Singh. An adaptive probabilistic trust model and its evaluation. In *Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems*, 2008.

[7] C.-W. Hang, Y. Wang, and M. P. Singh. Operators for propagating trust and their evaluation in social networks. In *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems*, 2009.

[8] A. Jøsang. A logic for uncertain probabilities. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 9:279–311, June 2001.

[9] J. Kohlas, D. Berzati, and R. Haenniz. Probabilistic argumentation systems and abduction. *Annals of Mathematics and Artificial Intelligence*, 34(1-3):177–195, 2002.

[10] N. Lehmann and R. Haenni. An alternative to outward propagation for Dempster-Shafer belief functions. In *Proceedings of the 5th European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*. Springer-Verlag, 1999.

[11] C.-J. Liau. Belief, information acquisition, and trust in multi-agent systems — a modal logic formulation. *Artificial Intelligence*, 149:31–60, 2003.

[12] N. Oren, T. J. Norman, and A. Preece. Subjective logic and arguing with evidence. *Artificial Intelligence*, 171(10-15):838–854, 2007.

[13] S. Parsons, Y. Tang, E. Sklar, P. McBurney, and K. Cai. Argumentation-based reasoning in agents with varying degrees of trust. In *Proceedings of the 10th International Conference on Autonomous Agents and Multi-Agent Systems*, Taipei, Taiwan, 2011.

[14] H. Prakken. An abstract framework for argumentation with structured arguments. *Argument and Computation*, 1:93–124, 2010.

[15] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ, 1976.

[16] P. P. Shenoy. Binary join trees for computing marginals in the Shenoy-Shafer architecture. *International Journal of Approximate Reasoning*, 17(2/3):239–263, 1997.

[17] Y. Tang, K. Cai, P. McBurney, E. Sklar, and S. Parsons. Using argumentation to reason about trust and belief. *Journal of Logic and Computation*, (to appear), 2012.

[18] Y. Tang, K. Cai, E. Sklar, and S. Parsons. A prototype system for argumentation-based reasoning about trust. In *Proceedings of the 9th European Workshop on Multiagent Systems*, 2011.