Formalizing Trust-based Decision Making in Electronic Commerce Transactions

Tyrone Grandison[1], Han Reichgelt[2]

[1] IBM Almaden Research, 650 Harry Road, San Jose, CA 95120, USA
tyroneg@us.ibm.com

[2] Georgia Southern University, Statesboro, GA 30460, USA
han@georgiasouthern.edu

Abstract. Trust is the cornerstone of traditional business transactions and a lack of trust is one of the main inhibitors on the future growth of Electronic Commerce. However, trust is a nebulous social construct that is further complicated by its context dependence. Our assertion is that through study of the contextual artifacts and the process of trust formation in traditional business settings, we can devise a logic-based model that adequately captures the constructs needed to enable the automated creation of trusted Electronic Commerce transactions. In this paper, we introduce a model, called E2T2, which enables the creation of trust-based relationships between online merchants and their consumers.

Keywords: Logic, Trust, Trust Management.

1 Introduction

Trust is important to the transition to the next level of Electronic Commerce innovation. Without technology assurances that software is built on robust, verifiable trust technology, further consumer acceptance of online commercial innovations is likely to be slow.

The state of the art in commercial trust technology, e.g. KeyNote [1] and REFEREE [2], is very effective at mapping security credentials to access rights in static environments [3]. However, security constraints are only one facet of the trust problem. Other facets that are often ignored by security-focused trust solutions include reputation, risk propensity and past behavior [4]. Though systems have been built to address each of these and many other facets individually, there is no single underlying and unifying model that incorporates all the important aspects needed for Electronic Commerce.

The current set of logic-based trust models [5-7] suffers the same fate as their 'authorization-centric' trust management engineering counterparts. They focus on specific aspects of the trust problem, without taking all the dimensions of a business transaction into consideration.

The aim of our model, E2T2 (Epistemic Event Temporal Trust), is to show that it is possible to create a model that encapsulates the decision factors and that can be

implemented for any arbitrary E-Commerce application environment. While E2T2 is not complete, we believe that the model is flexible enough to allow adaptation to arbitrary application domains and facilitate the inclusion of emerging issues and concepts as they materialize.

We present the background notions needed for this discussion in section 2 and describe the basics of the E2T2 model in section 3. In section 4, we present how trust representation for Electronic Commerce is done in E2T2 and then round out the discussion on the auxiliary concepts needed for E-Commerce decision making (section 5). We conclude in section 6.

2 Background

Our work is grounded in several fields, namely: trust management [8,9], Electronic Commerce [4] and modal logic [10-15]. We introduce the base concepts in each of these areas.

2.1 Trust

Not all the definitions of trust in the computer science literature are applicable to Electronic Commerce [16]. For our context, we borrow from the definition used in [4] for trust in the E-Services environment, which states that *"trust is the quantified belief by a trustor with respect to the competence, honesty, security and dependability of a trustee within a specified context"*, where dependability encapsulates timeliness and reliability.

A trustor is defined as the subject of the relationship and a trustee the object. For example, if Luke wishes to initiate the purchase of a large shipment of fertilizer from RedCo Chemicals, then Luke is the trustor in this instance and the object is RedCo Chemicals. Luke's decision to trust could be based on one or a combination of a myriad of different factors, e.g. RedCo's reputation in the business community, a recommendation from one of Luke's friends, RedCo's SEC credentials or maybe even Luke's personal past dealings with RedCo. These decision criteria are critical to the problem of establishing and evolving trust.

Currently, there are two prevailing views of trust management. The first is the one defined by Blaze et. al. [1] that advocates an authorization-based approach to trust management. The other is a business-oriented approach [16] that examines trust management through the core mechanism used in successful E-Commerce applications, e.g. reputation. We augment the latter approach in this paper.

2.2 Decision-Making in Electronic Commerce

When two entities, who may have no prior knowledge of each other, wish to engage in an Electronic Commerce transaction they must each ask themselves *"Should I trust this entity and engage in this transaction (or not)?"* The problem arises when each entity must choose the indicators that should be used to make this decision and request it from the other entity.

The common solution has been to use security credentials, such as authentication tokens and attributes, to establish either identity or determine the allowable trusted

tasks. However, security credentials are as good as the issuer of these credentials, do not give any indication of past or present behavior and only partially resolve the problem of how to interact with strangers. In a nutshell, a lot of contextual information is missing. This highlights the fact that the common solutions need to be enhanced.

In a standard electronic transaction e, where a wants to engage b, a must determine the level of trust or distrust and use this to take the most prudent course of action. Apart from assurances on the competence, honesty, security, timeliness and reliability of b, a has business considerations to include in his trust decision. From a purely commercial perspective, a favorable decision may depend on:

1. Intrinsic properties of the transaction.
2. Assurances about respecting a's privacy.
3. a's experience with b (with respect to similar transactions).
4. b's past experience with others (who may be like a), with respect to similar transactions.
5. b's reputation in the community of interest.
6. A recommendation on b's behalf.

Generally, these factors can be classified as either transaction-specific or event-driven. Let's briefly present each.

2.2.1 Transaction-Specific Factors

Transaction-specific trust decision criteria are intrinsic properties of a transaction, and include the associated risk, the transaction value, the insurance coverage for the transaction and any associated economic incentive (or disincentive) associated with performing the transaction.

[4] purports that there are intuitive correlations between these factors and the final decision made. For example, transactions with a low risk value (in comparison to one's risk threshold) tend to lead to a positive trust decision. Under normal circumstances, a rational a considers a low-valued transaction trivial to his or her bottom-line and may not mind engaging with b, even if b is malicious and shatters all of a's positive expectations. Generally, low-valued transactions are deemed as having very little risk. Normally, this is independent of an explicitly stated risk value for the transaction. Thus, a low-valued transaction has a high probability of being viewed as beneath one's risk threshold and often leads to positive trust decisions.

A transaction that is insured, at least to the transaction value, can be recovered (in the worst case). Having insurance lowers the risk threshold of a transaction. Normally, decisions involving insured transactions will not take the transaction's value into account. Instead the other factors involved in the decision will determine the outcome. If insurance is the only (or dominant) factor, then there is a higher chance of a favorable trust decision.

If b, or a third party, is willing to compensate a for engaging in e, to a level that a views as significant, then a will view the transaction as being below his risk threshold and thus be inclined to trust b and engage in the transaction. Similar reasoning can be applied to the disincentive case, where a third party pays a not to proceed with the transaction.

2.2.2 Event-Driven Factors

Event-driven factors relate to activity, agreements, arrangements or information that occur either prior to or outside the scope of the current trust decision with consequences that extend beyond the existing decision. For example, a's body of knowledge concerning past encounters with b as it pertains to e constitutes information that occurred prior to the current decision and whose influence will extend beyond the decision. Thus, experience, reputation, recommendation and privacy concerns all fall into this category. We will discuss these later in this paper. Grandison [4] offers a discussion on the treatment of the correlations between these factors and the trust decision outcome. We will only state here that in the case of privacy constraints, a's rights can be protected by 1) a not releasing any private information or releasing a limited view [17], or 2) b asserting that it won't violate a's privacy and a actively monitoring the disclosures to and the activity performed by b, which seems to be infeasible in today's computing environment. If a privacy violation occurs, then this event will negatively impact the trust relationship. If no violation happens, then the relationship remains in tact.

2.3 Logic

In this section, we introduce the central concepts from epistemic logic and temporal logic.

2.3.1 Epistemic Logic

Epistemic logic [10,11] allows one to model complex scenarios involving knowledge and belief, by augmenting non-modal logic and leveraging the epistemic operators K_c and B_c such that K_cA means Agent c knows A and B_cA means Agent c believes A for some arbitrary proposition A. The fundamental concept behind any modal logic is that of a possible world, each of them essentially being a model (in the logical sense) of the underlying non-modal logic. In addition, each modal logic defines an accessibility relation between possible worlds, and different modal logics differ in the interpretation of this accessibility relation. Thus, in an epistemic logic, two worlds w and w' are accessible from each other for an agent c if both of them are compatible with c's information state. The notation used is R_cww'. World w' is said to be an *epistemic* or a *doxastic* alternative to world w for agent c, depending on whether knowledge or belief is the attitude under consideration.

2.3.2 Temporal Logic

Temporal logic [14,15] allows the representation and reasoning about propositions that are qualified with a time component. There are a number of different approaches in Computer Science to allow for temporal reasoning. They range from the situation calculus of McCarthy and Hayes, which essentially captures the temporal dimension by adding a temporal term to each function and predicate symbol, to modal temporal logic, to reified temporal logic. [19] provides an overview of the different approaches to incorporating a temporal dimension in a logical framework.

Because of its greater expressive power, we adopt a reified approach in E2T2. We distinguish between states and events, and we introduce the predicates HOLDS and OCCURS to indicate that a particular state is true over a period time, or that an event occurs at some point in time.

3 The Basics of the Model

From our definition of trust, we see that a trust decision is based on an entity's beliefs. Thus, our model must be founded on epistemic logic.

As there are two parties in a typical Electronic Commerce transaction, it is safe to assume that each party has their own set of beliefs and that these will evolve over time. Thus, the trust relationship is dynamic, in that the criteria or set of beliefs and one's behavior may change over time based on trigger events. This highlights the need to incorporate some form of temporal logic.

We will use a modal approach to modeling belief. However, since trust is a quantified belief, we cannot rely on the standard modal operator BEL to model belief. Instead, we introduce a modal operator schema BEL_{str}. By convention, str $\in$ <0, 1], i.e., a number larger than 0 and less than or equal to 1, and indicates the level of confidence that the agent has in the belief.

We opt for a reified approach to temporal reasoning [18]. For our ontology, we adopt the formalism developed in [20], which introduces event and state tokens. We ignore subtleties regarding the choice of temporal primitives. For an in-depth discussion, see [21]. For the purposes of this paper, we model time as a set of points.

In the following sections, we introduce the basic vocabulary, individual terms and sentence construction in E2T2, then we present the semantic interpretation of the language constructs and highlight implications of our model's design assumptions.

3.1 The Language

We will start with the basics, i.e. presenting the syntactic primitives of E2T2 and showing proper sentence construction in this formalism.

3.1.1 The Basic Vocabulary

When representing trust, we need to be able to talk about different types of entity. We therefore define E2T2 as a sorted logic with the following sorts:

- I, Individuals ($i, j, k, i_1 \ldots i_n$)
- A, Agents, a sub-sort of individuals, ($a,b,c,a_1 \ldots a_n$)
- T, Times ($t_1 \ldots t_n$)
- S, States ($s_1 \ldots s_n$)
- E, Events ($e_1 \ldots e_n$)
- N, the set of numbers between 0 and 1.

E2T2 contains a number of function symbols. Each n-place function symbol f has a signature Σ_f: $v_1,..,v_{n-1} \rightarrow v_n$ where $v_1,..,v_n$ are sorts and $v_1,..,v_{n-1}$ specify the sorts of the input arguments and v_n the sort of the output.

E2T2 contains a number of predicate symbols. Each n-place predicate p has a signature Π_p $v_1,..,v_n$ where $v_1.,v_n$ are sorts. E2T2 contains the following special predicates:

- BEF, with signature Π_{BEF}:T,T. Intuitively, BEF(t_1, t_2) is true iff t_1 is earlier than t_2. Although a full temporal logic will have to contain more predicates between times, this one temporal predicate is sufficient for our current purposes.
- HOLDS, with signature Π_{HOLDS}:S,T,T. Intuitively, HOLDS(s, t_1, t_2) is true iff state s is true between time t_1 and t_2.

- OCCURS, with signature Π_{OCCURS}:E,T. Intuitively, OCCURS(e, t) is true iff event e occurs at time t.

Finally, E2T2 contains the 4–place modal operator BEL whose first three arguments are of type A, N and T respectively and whose final argument is a proposition. Intuitively, BEL(a,str,t,p) indicates that a believes p with confidence level str at time t.

3.1.2 Individual Terms

Individual terms in E2T2 are defined in the normal way. Thus:

- Every constant is an individual term;
- If f is a function symbol with Σ_f: $v_1,..,v_{n-1} \rightarrow v_n$ and $z_1,..,z_{n-1}$ are terms of sorts $v_1,..,v_{n-1}$ respectively, then $f(z_1.,,z_{n-1})$ is an individual term of sort v_n

3.1.3 Sentences

In defining the syntax of E2T2, we make use of the following auxiliary notion:

- If φ is a sentence and z and x terms of sort v then φ[z/x] is obtained by replacing all occurrences of z in φ by x

Sentences in E2T2 are formed in the normal way, that is:

- If p is an n-place predicate with signature Π_p $v_1,..,v_n$ and $z_1,..,z_n$ are terms of sorts $v_1,..,v_n$ then $p(z_1,..,z_n)$ is a sentence;
- If φ and ψ are sentences, then so are ¬φ, (φ → ψ), (φ & ψ) and (φ ∨ ψ);
- If φ is a sentence that contains terms z and x of sort v, then (∀x:v)[φ(z/x)] and (∃x:v)[φ(z/x)] are sentences.
- If φ is a sentence and a, str and t are terms of type A, N and T respectively, then BEL (a,str,t,φ) is a sentence.

3.2 Semantics

The semantics for E2T2 is relatively complicated as it has to combine elements from the semantics of temporal logic with elements of the semantics of epistemic logic. Moreover, because we deal with quantified belief, we cannot rely on the standard possible worlds approach to provide semantics for the belief modality. We achieve the desired effect by complicating the accessibility relation between possible worlds.

From an intuitive point of view, an E2T2 model consists of a set of individuals, suitably partitioned to take account of the different sorts in the language, a set of time intervals and a set of possible worlds. The model contains an ordering relationship < between time intervals. In addition, for each agent, we define an accessibility function that maps any two pairs of possible worlds at a particular point in time onto a number larger than 0 and less than or equal to 1. This number indicates the likelihood that a world w' is accessible from world w according to the agent at time t. It is loosely based on Lewis' complication of the standard possible world semantics to deal with counterfactuals [22] and will be used to deal with the confidence level that an agent has in different beliefs.

With this in mind, we can now formally define an E2T2 model as follows:

An E2T2 model $\mathcal{M}$ consists of

a. A tuple **<O, W, T>** where
   - **O**, the set of objects is partitioned into 3 non-overlapping sets, **I, S** and **E**, the sets of individuals, states and events respectively, and **I** contains a non-empty set **A**, the set of agents;

- **W** is the set of possible worlds
- **T** is the set of time intervals;
- **N** is the set of numbers in the interval < 0, 1]

b. A temporal ordering relation **<** over **T**.

c. A function **Acc** that associates with each **a** ∈ **A** a 3-place function from **W** X **W** X **T** into **N**.

d. An interpretation function $\Im^{\mathcal{M}}$.

The interpretation function $\Im^{\mathcal{M}}$ assigns an appropriate denotation to each expression in E2T2 in each world-time pair. We use $\Im^{\mathcal{M}}{}_{w,t}(exp)$ as a short-cut notation for $\Im^{\mathcal{M}}$ (exp, **w**, **t**).

For primitive individual terms, it is defined in the standard ways:

a. If z is a primitive term of sort v, then $\Im^{\mathcal{M}}{}_{w,t}$ (z) ∈ Corr(v), where Corr is a function mapping a sort into the corresponding set of objects. Thus,
Corr(A) = **A**, Corr(I) = **I**, Corr(S) = **S**, Corr(E) = **E**, Corr(T) = **T**,
Corr(N) = **N.**
In order to avoid the problem of trans-world identity, we stipulate that for all primitive terms $\Im^{\mathcal{M}}{}_{w,t}$ (z) = $\Im^{\mathcal{M}}{}_{w',t'}$ (z). In other words, the denotation of a primitive term does not vary from world-time pair to world-time pair.

b. If f is a function symbol with S_f: $v_1.,,v_{n-1} \rightarrow v_n$ then $\Im^{\mathcal{M}}{}_{w,t}$ (f) is a function from Corr(v_1) X … X Corr(v_{n-1}) into Corr(v_n), and $\Im^{\mathcal{M}}{}_{w,t}$ (f(v_1,..,v_{n-1})) = $\Im^{\mathcal{M}}{}_{w,t}$ (f)($\Im^{\mathcal{M}}{}_{w,t}$($v_1$),..,$\Im^{\mathcal{M}}{}_{w,t}$ (v_{n-1}))

For predicates, $\Im^{\mathcal{M}}$ is defined in the expected way as well:

a. If p is an n-place predicate with signature Π_p $v_1.,,v_n$, then, $\Im^{\mathcal{M}}{}_{w,t}$ (p) ⊆ (Corr(v_1) X … X Corr(v_n))

The only complication is that the interpretation of the special predicate BEF is the ordering relation between points in time.

In order to define $\Im^{\mathcal{M}}$ for sentences, we introduce an auxiliary notion, namely the set of all possible worlds that are accessible from a given world *w* according to actor a at time t with at least confidence level str; which is **Acc**(w, a, t, str). Formally, this notion can be defined as follows:

w' ∈ **Acc**(w, a, t, str) iff **Acc**(a)(w, w', t) ≤ str

Although this definition may at first glance seem counterintuitive, the reason for it will become clear in the following sections.

With this concept, we can now extend $\Im^{\mathcal{M}}$ to sentences. $\Im^{\mathcal{M}}$ maps each sentence to either 1 (true) or 0 (false) according to the following rules:

a. $\Im^{\mathcal{M}}{}_{w,t}$ (p(v_1,..,v_n)) = 1 iff <$\Im^{\mathcal{M}}{}_{w,t}$ (v_1),..,$\Im^{\mathcal{M}}{}_{w,t}$ (v_n)> ∈ $\Im^{\mathcal{M}}{}_{w,t}$ (p) and 0 otherwise;

b. $\Im^{\mathcal{M}}{}_{w,t}$ (¬φ) = 1 iff ($\Im^{\mathcal{M}}{}_{w,t}$ (φ) = 0) , and 0 otherwise;

c. $\Im^{\mathcal{M}}{}_{w,t}$((φ → ψ)) = 1 iff (($\Im^{\mathcal{M}}{}_{w,t}$ (φ) = 0) or ($\Im^{\mathcal{M}}{}_{w,t}$ (ψ) = 1)), and 0 otherwise;

d. $\Im^{\mathcal{M}}{}_{w,t}$ ((φ & ψ)) = 1 iff ($\Im^{\mathcal{M}}{}_{w,t}$ (φ) = $\Im^{\mathcal{M}}{}_{w,t}$ (ψ) = 1) , and 0 otherwise;

e. $\Im^{\mathcal{M}}{}_{w,t}$ ((φ ∨ ψ)) = 1 iff (($\Im^{\mathcal{M}}{}_{w,t}$ (φ) = 1) or ($\Im^{\mathcal{M}}{}_{w,t}$ (ψ) = 1)), and 0 otherwise;

f. $\Im^{\mathcal{M}}{}_{w,t}$ ((∀x:v)[φ]) = 1 iff for all $\Im'^{\mathcal{M}}$ which are exactly like $\Im^{\mathcal{M}}$ except for the value assigned to z, $\Im'^{\mathcal{M}}{}_{w,t}$ (φ[x/z]) = 1 where z is a term of sort v not occurring in φ, and 0 otherwise.

g. $\Im^{\mathcal{M}}{}_{w,t}$ ((∃x:v)[φ]) = 1 iff there is an $\Im'^{\mathcal{M}}$ which is exactly like $\Im^{\mathcal{M}}$ except for the value assigned to z such that $\Im'^{\mathcal{M}}{}_{w,t}$ (φ[x/z]) = 1 where z is a term of sort v not occurring in φ.

h. $\Im^{\mathcal{M}}_{w,t}$ (BEL(a, str, *t*, φ) = 1) iff (for all w' ∈ Acc(w, a, t', str), $\Im^{\mathcal{M}}_{w',t}$ (φ) = 1) where **t'** = $\Im^{\mathcal{M}}_{w,t}$ (*t*)) and 0 otherwise.

The decisions taken in designing the syntax and semantics have led to interesting consequences, which are presented in Appendix A.

4 Representing Trust

In this section we use the logic framework of E2T2 to formalize the notion of trust. However, before we do so, we first introduce some special function symbol and predicates.

4.1 Auxiliary Function and Predicate Symbols

The first function symbol that we need is CANINDUCE with signature $\Sigma_{CANINDUCE}$: A, E → S. Intuitively, CANINDUCE(a,e) describes the state of a being able to bring about event e, either directly or through some third party Often, the event will be some action that is performed by a.

A second required function symbol is REQUEST with signature $\Sigma_{REQUEST}$: A, A, E → E Intuitively, REQUEST maps two agents and an event into an event, namely the event of the first agent requesting the second agent to bring about the event.

A third function symbol is COMMIT with signature Σ_{COMMIT}: A, E, T → E Intuitively, COMMIT maps an agent, an event and a time into an event, namely the event of the agent committing itself to bring about event before that time.

The final function symbol that we need is slightly more complicated, namely TPER, with the signature Σ_{TPER}: A, A, E, T → T. TPER takes two agents a and b, an event e, and a time t and maps this into a second time period t'. We assume that agent b has committed to agent a to perform e at time t. TPER returns the time by which agent a can reasonably expect agent b to have delivered on its commitment. It reflects the intuition that, when an agent commits to do something, one can reasonably expect him or her to perform the action within a certain time period. The time period that is reasonable clearly depends on the action committed to. For an event like passing the salt at a dinner table, the period will be fairly short; for a promise made to an incoming student in a four year program to help her complete her degree successfully, the time period is longer. We stipulate that if $x = TPER(a_1, a_2, e_1, t_1)$ then $BEF(t_1,x)$. This function is used to model timeliness, which is a crucial element of the trust formalism.

4.2 Trust

We are now in a position to give a preliminary definition of trust in terms of E2T2. We previously stated that for purposes, trust is the quantified belief by a trustor with respect to the competence, dependability, security and honesty of a trustee with respect to the trustee bringing about some event or state of affairs and in a specified context.

With this in mind, we introduce a predicate TRUST with the signature Π_{ITRUST}: A, N, A, T, E. Intuitively, TRUST(a,str,b,t,e) means that a trusts b at trust level str at time t to bring about event e. Naively, we see that this trust relationship will only be

established when a believes that b is currently competent to bring about e (in a timely manner), when a believes that b will always carry out commitments with respect to e when requested (in a timely manner), when a believes that b is currently secure and when a believes that b is currently honest. It becomes clear that timeliness is an intrinsic requirement for all the other attributes, rather than a single explicit constraint.

Further examination of the clauses reveals that there are multiple dimensions of time present. For example, with the competence and security clauses, we assume that the trustee is competent and secure for a reasonable period, while for the dependability clause we assume that whenever the trustor makes a request of the trustee to bring about some event, the trustee forms, within a reasonable time of the request being made, the intention of bringing about e in a timely manner. All of this raises the question *what time period is considered reasonable?* The answer lies in a number of factors. One factor is the agent who considers the period to be reasonable, or not. Some agents are more patient than others. The second factor is the person who is to bring about the event. Although one might trust both a seven year old child and an adult to set the table, when asked to do so, one would expect the child to take longer than the adult. The final factor is the event itself. One would expect it to take longer to make a gourmet meal than it takes to make a cup of coffee.

Let's discuss the formulations of the exact clauses for each attribute.

Competence: A competent entity is one that is able to perform or lead to the performance of an event. In our framework, this translates to a's belief that b is capable of bringing about e in a timely manner. We formally define this as:

$$BEL(a, str_1, t, [(\forall x{:}T)\ (\forall x'{:}T)\ BEF(x, x') \ \&\ BEF(x', TPER(a,b,e,x)) \rightarrow HOLDS(CANINDUCE(b,e), x', TPER(a,b,e,x))]$$

a does not have to believe that b is capable of bringing about e in perpetuity. a merely has to believe that b can bring about event e within the period in which it is reasonable for a to expect b to bring about e. Note that we also do not mean to imply that a needs to believe in b's competence for the entire period in which it is reasonable for a to expect b to bring about e. a can change its opinion at any time. The clause merely states that at the time at which a trust b to perform e, a believes that b is competent to bring about e in a reasonable timeframe.

Dependability: Dependability refers to reliability and timeliness. A reliable entity is one that performs an event when it is requested to do so. A dependable entity is a reliable one that executes in a timely manner. We formaslize this as:

$$BEL(a, str_2, t, (\forall x{:}T)\ [OCCURS(REQUEST(a,b,e), x) \rightarrow (\exists x'{:}T)\ (\exists x''{:}T)\ [BEF(x, x') \ \&\ BEF(x', TPER(a,b,REQUEST(a,b,e),x)) \ \&\ BEF(x', x'') \ \&\ OCCURS(COMMIT(b,e, TPER(a,b,e,x)), x'')]])$$

The clause states that a believes that whenever it requests of b to bring about e, then b will, within a reasonable period of the request being made, form the intention to bring about event e in a timely manner. We use the function symbol TPER twice, to reflect two subtly different forms of timeliness. In order to be considered reliable, a trustee has to meet two requirements.

First, whenever the trustor makes a request, the trustee must, in a timely manner, respond to the request by making some commitment. How long a period is reasonable depends on the request. One would expect a friend to respond to a request to make a cup of coffee more quickly than he or she might respond to a request to

marry you. This element of timeliness is modeled in the first use of TPER. However, it is not enough for the trustee to merely make the commitment in a timely manner. The trustee must also commit to bringing about whatever the trustor requests within reasonable timescales. We model this aspect through the second use of TPER.

An example may further clarify the distinction. Assume that a trusts waitress b in Frank's Diner to serve a meal. a believes that whenever he orders a meal from b, i.e. requests b to serve him a meal, b will form the commitment to bring the meal after a makes the request. However, a would not expect b to bring it the meal immediately, although there clearly is some limit to the amount of time that a would be prepared for the meal to arrive.

Security: A secure entity is one that provides assurances on its safe operation and execution.

$$BEL(a, str_3, t, (\forall x{:}T)\ (\forall x'{:}T)\ [\ BEF(x, x') \ \&\ BEF(x', TPER(a,b,e,x)) \rightarrow HOLDS(SECURE(b,e), x', TPER(a,b,e,x))])$$

As was the case with our analysis of competence, a does not have to assume that b will be secure with respect to e for ever, but merely for the period in which it is reasonable for a to expect b to perform e after a has requested b to perform e. Again, as in the case of the competence attribute, it is also true that a can change its view about b's security at any time.

Honesty: A honest entity is one that is sincere with regards to its interactions, i.e. it does what it commits to doing.

$$BEL(a, str_4, t, (\forall x{:}T)(\forall x'{:}T)\ [OCCURS(COMMIT(b,e,x'), x) \rightarrow (\exists x''{:}T)[BEF(x, x') \ \&\ BEF(x', x'') \ \&\ OCCURS(e, x'')]])$$

The above states that a believes that whenever b has committed to bringing about e before some time, e will indeed occur before that time. Thus, b is sincere with respect to its commitments.

The Complete Formulation: Putting all the pieces together, we get the following formulation of trust:

$$\begin{aligned} &TRUST(a, CALC(str_1, str_2, str_3, str_4), b, t, e) \leftrightarrow \\ &\ BEL(a, str_1, t, [\ (\forall x{:}T)\ (\forall x'{:}T)\ BEF(x, x') \ \&\ BEF(x', TPER(a,b,e,x)) \rightarrow \\ &\qquad HOLDS(CANINDUCE(b,e), x', TPER(a,b,e,x))\ ] \\ &\ \&\ BEL(a, str_2, t, (\forall x{:}T)\ [OCCURS(REQUEST(a,b,e), x) \rightarrow \\ &\qquad (\exists x'{:}T)\ (\exists x''{:}T)\ [BEF(x, x') \ \&\ BEF(x', TPER(a,b,REQUEST(a,b,e),x)) \ \& \\ &\qquad BEF(x', x'') \ \&\ OCCURS(COMMIT(b,e, TPER(a,b,e,x)), x'')]]) \\ &\ \&\ BEL(a, str_3, t, (\forall x{:}T)\ (\forall x'{:}T)\ [\ BEF(x, x') \ \&\ BEF(x', TPER(a,b,e,x)) \rightarrow \\ &\qquad HOLDS(SECURE(b,e), x', TPER(a,b,e,x))]) \\ &\ \&\ BEL(a, str_4, t, (\forall x{:}T)(\forall x'{:}T)\ [OCCURS(COMMIT(b,e,x'), x) \\ &\qquad \rightarrow (\exists x''{:}T)[BEF(x, x') \ \&\ BEF(x', x'') \ \&\ OCCURS(e, x'')]]) \end{aligned} \quad \textbf{(1)}$$

There are a number of issues that we wish to draw attention to.

First, the term $CALC(str_1, str_2, str_3, str_4)$ indicates that the trust level is dependent on the four quantified beliefs that make up the definition of trust. Clearly, this raises the question about the actual definition of CALC, which may be domain-specific and or entity-related. The exact mathematical formulae used to calculate the trust value may be as simple as arithmetic average or as complex as weighted statistical computation. This paper is not focused on trust calculation and we will leave this issue until another occasion. However, there are a number of stipulations that we need to put on the function CALC. First, we assume that the range of the function CALC is the interval [0,1]. Second, whatever function is chosen to implement CALC, it

must be monotonically increasing. We want to make sure that, as the confidence level in the trustor's beliefs of either the trustee's competence, dependability, security, or honesty with respect to e increases, then its trust level in the trustee to bring about e minimally does not decrease, and should probably increase.

Second, Note that the model theory and the increasing monotonicity of CALC imply that if a trusts b at time t to bring about e at trust level n, then a also trusts b at time t to bring about e at trust level n' where n' is less than n. After all, if a holds any belief at confidence level c, a also holds that belief at any confidence level less than c. If a trusts b to bring about e at level n, where n is the result of applying the function CALC on the confidence levels of the beliefs in the competence, dependability, security and honesty of b, then a also believes in the competence, dependability, security and honesty of b at lower confidence levels. Applying CALC to these lower confidence levels must, because of the increasing monotonicity of CALC, result in a value that is less than the original value. Thus,

$$\text{TRUST}(a, n, b, t, e) \rightarrow (\forall n')[\ (0 < n' \leq n) \rightarrow \text{TRUST}(a,n',b,t,e)] \qquad \textbf{(2)}$$

Since n takes on a value greater than 0 and less than or equal to 1, it also follows that

$$\neg\text{TRUST}(a, n, b, t, e) \rightarrow (\forall n')[\ (n \leq n' \leq 1) \rightarrow \neg\text{TRUST}(a,n',b,t,e)] \qquad \textbf{(3)}$$

Our formulation of trust lends itself to the modeling of constructs needed for Electronic Commerce, which for purposes of brevity are presented in Appendix B.

6 Conclusion

Epistemic Event Temporal Trust (E2T2) is a logical framework that leverages knowledge on the trusting behavior of businesses to create a unifying model that allows all the components of the trust decision to be factored into the process of trust establishment. The model's design allows it to be flexible enough to model the trusting behavior of any arbitrary business environment.

Our future works involves modeling the more troublesome notions that online entities must handle (e.g. partial information, uncertainty, etc.). We will also create E2T2 reasoning engines and evaluate their effectiveness. Finally, we will consider the best deployment mechanisms.

The overall focus on the work on E2T2 is to provide a foundation for discourse in the research community and to underscore the importance of the need to take a holistic approach to the trust problem.

References

[1] Blaze, M., J. Feigenbaum and Keromytis, A.D.: *KeyNote: Trust Management for Public-Key Infrastructures*. in Security Protocols International Workshop. 1998. Cambridge, England. http://www.cis.upenn.edu/~angelos/Papers/keynote-position.ps.gz

[2] Chu, Y.-H., J. Feigenbaum, B. LaMacchia, P. Resnick and M. Strauss: *REFEREE: Trust Management for Web Applications*. 1997, AT&T Research Labs Research Report. http://www.farcaster.com/papers/www6-referee/

[3] Blaze, Matt, Ioannidis, John, Keromytis, Angelos D.: *Experience with the Keynote Trust Management System: Applications and Future Directions*, Proceedings of Intl. Trust Management Conference (iTrust) 2003, Pg 284-300.
[4] Grandison, T. : *Conceptions of Trust: Definition, Constructs and Models* in Trust in E-Services: Technologies, Practices and Challenge, Editor: Ronggong Song. Published by IDEA Group. 2007.
[5] Xie, H.-b., Zhou, M.-t,: *PKI Trust Model Analysis Based on Probabilistic Model and Conditional Predicate Calculus Logic*, MINIMICRO SYSTEMS -SHENYANG, Vol 27, No 1, 2006.
[6] Nefti, Samia, Meziane, Farid, Kasiran, Khairudin,: *A Fuzzy Trust Model for E-Commerce*, Proceedings of the Seventh IEEE International Conference on E-Commerce Technology (CEC'05), Pages 401-404, 2005.
[7] Ramchurn, S. D., Sierra, C., Godo, L. and Jennings, N. R. : *Devising a trust model for multi-agent interactions using confidence and reputation*. International Journal of Applied Artificial Intelligence 18(9-10) pp. 833-852. 2004.
[8] Ji, Ma, Orgun, M.A. : *Trust management and trust theory revision*, IEEE Transactions on Systems, Man and Cybernetics, Vol 36, Issue 3, May 2006, Pages 451- 460.
[9] Ruohomaa, S., Kutvonen, L.: *Trust management survey*, Proceedings of the 3rd International Conference on Trust Management 2005, Pages 77–92.
[10] Georg Henrik von Wright: An Essay in Modal Logic, 1951
[11] Jaakko Hintikka: The Logic of Epistemology and the Epistemology of Logic.
[12] R. Kowalski and M.Sergot: A Logic-Based Calculus of Events New Generation Computing, vol. 4 pp. 67–95, 1986.
[13] R. Miller and M. Shanahan: The event-calculus in classical logic — alternative axiomatizations. Electronic Transactions on Artificial Intelligence, 3(1):77-105, 1999.
[14] Venema, Yde: Temporal Logic, in Goble, Lou, ed., The Blackwell Guide to Philosophical Logic. Blackwell, 2001.
[15] E.A. Emerson: Temporal and modal logic, Handbook of Theoretical Computer Science, Chapter 16, the MIT Press, 1990
[16] Grandison, T. and Sloman, M.: A Survey of Trust in Internet Applications, IEEE Communications Surveys and Tutorials, Vol. 3 No. 4. Oct-Dec 2000.
[17] K. Lefevre, R. Agrawal, V. Ercegovac, R. Ramakrishnan, Y. Xu, D. DeWitt: Limiting Disclosure in Hippocratic Databases. Proc. of the 30th Int'l Conf. on Very Large Databases (VLDB 2004), Toronto, Canada, August 2004.
[18] Vila, L. and H. Reichgelt: *The token reification approach to temporal reasoning*. Artificial Intelligence, 83 (1996), 59-74.
[19] Reichgelt, H. and L. Vila: *Temporal qualification in artificial intelligence*. In: M. Fisher, D. Gabbay and L. Vila (eds) Handbook of Temporal Reasoning in Artificial Intelligence. Amsterdam: Elsevier, 2005.
[20] Shoham, Y.: *Temporal logics in AI: Semantical and ontological considerations.* Artificial Intelligence, 33 (1987), 89-104.
[21] Vila, L.: *Formal Theories of Time and Temporal Incidence.* In: M. Fisher, D. Gabbay and L. Vila (eds) Handbook of Temporal Reasoning in Artificial Intelligence. Amsterdam: Elsevier, 2005.
[22] Lewis, D.: *Counterfactuals.* Oxford: Blackwell, 1973.

Appendix A: E2T2 Design Consequences

Our definition of belief has a number of consequences that are worth mentioning. First, notice that, as any standard epistemic logic, our logic suffers from the problem of logical omnidoxasticity:

$$(\forall a{:}A)\ (\forall n{:}N)\ (\forall t{:}T)\ [\ ((\varphi \rightarrow \psi)\ \&\ BEL(a, n, t, \varphi)\ ) \rightarrow BEL(a,n,t, \psi)\ ] \quad \textbf{(4)}$$

Second, agents can change their minds, i.e. if an agent believes at time t that some event will occur at time t' or that some state will hold at time t', then it does not follow that it must believe at some other time t" that this event occurs or state holds at t'. In other words,

$$\neg\ (\forall a{:}A)(\forall n{:}N)\ (\forall t, t', t"{:}T)(\forall e;E)\ [\ BEL(a, n, t, OCCURS(e,t'))\ \rightarrow\ BEL(a, n, t", OCCURS(e,t'))\ ] \quad \textbf{(5)}$$

$$\neg\ (\forall a{:}A)(\forall n{:}N)\ (\forall t, t, t"{:}T)(\forall s;S)\ [\ BEL(a, n, t, HOLDS(s,t, t'))\ \rightarrow\ BEL(a, n, t", HOLDS(e,t, t''))\ ] \quad \textbf{(6)}$$

After all, there is no requirement that all the worlds that are accessible from some given world at time t, are also accessible from that world at another time.

Third, if an agent believes a proposition with a certain confidence level, then it will also believe that proposition with a lower confidence level. In other words:

$$(\forall a{:}A)(\forall n, n'{:}N)(\forall t{:}T)\ [\ (\ BEL(a, n, t, \varphi\ )\ \&\ (n' < n)\ )\ \rightarrow BEL(a, n', t, \varphi)\ ]\ \textbf{(7)}$$

This proposition follows directly from the way in which we defined **Acc**(w, a,,t, str) as the set of possible worlds that were accessible from w according to a at time t with a confidence level of at least str. Clearly, it follows that if a proposition φ is true in all worlds that are accessible from some world w according to a at time t with at least confidence level n, then it must also be true in all worlds that are accessible with a lower confidence level.

Fourth, while it does follow that if an agent believes the negation of a proposition with a certain confidence level, then it does not believe that proposition with any confidence level, it does not follow that if an agent does not believe a proposition with a certain confidence level, then it believes the negation of that proposition. In other words,

$$\neg(\forall a{:}A)(\forall t{:}T)(\forall n{:}N)\ [\ \neg\ BEL(a, n, t, \varphi\ )\ \rightarrow (\exists n')[\ BEL(a, n', t, \neg\varphi)\ ] \quad \textbf{(8)}$$

$$(\forall a{:}A)(\forall t{:}T)(\forall n{:}N)\quad [\ BEL(a, n, t, \neg\varphi\ )\ \rightarrow \neg(\exists n')[\ BEL(a, n', t, \varphi)\ ] \quad \textbf{(9)}$$

Again, the above follows directly from the way in which the model theory has been defined. An agent not believing a proposition with a certain confidence level merely means that there is at least one possible world that is accessible with that confidence level in which the negation of that proposition is true; it certainly does not follow that the negation of the proposition must be true in all possible worlds that are accessible with an arbitrary confidence level. On the other hand, if an agent believes the negation of a proposition with an arbitrary confidence level, then the negation of that proposition must be true in all worlds that are accessible with a certain confidence level, and that, no matter what confidence level one considers, there is always at least one possible world in which the proposition is false.

Fifth, because denotations of primitives are unique across possible worlds and function symbols and predicates, other than BEF, varies from possible world to possible world, it is possible to tailor the model to any particular world of interest (e.g. an auction website, a manufacturing system, etc).

Appendix B: Electronic Commerce Constructs in E2T2

From [4], we know that mathematical properties such as transitivity, asymmetry, etc. cannot be axioms of any trust model simply because they are not universally applicable across domains. Given the way in which we have formulated the model

theory, none of these properties are axioms. However, each property could be modeled in E2T2 if it was applicable to the domain of interest.

In this section, we highlight the notions of distrust, trust dynamism and risk.

5.1 Distrust

Distrust can be conceptually viewed as simply a lack of trust. Formally, it is the quantified belief by a trustor that a trustee is incompetent, undependable, not secure and or dishonest with respect to bringing about some event or state of affairs for a specified context. For example, if a does not believe in b's ability to competently execute event e, then a may distrust b to execute e. The same holds true for all the other attributes.

$$\begin{aligned}
&\neg \mathrm{BEL}(a, str_1, t, [\ (\forall x{:}T)\ (\forall x'{:}T)\ \mathrm{BEF}(x, x') \,\&\, \mathrm{BEF}(x', \mathrm{TPER}(a,b,e,x)) \rightarrow \\
&\qquad \mathrm{HOLDS}(\mathrm{CANINDUCE}(b,e), x', \mathrm{TPER}(a,b,e,x))\ ] \\
&\vee\ \neg \mathrm{BEL}(a, str_2, t, (\forall x{:}T)\ [\mathrm{OCCURS}(\mathrm{REQUEST}(a,b,e), x) \rightarrow \\
&\quad (\exists x'{:}T)\ (\exists x''{:}T)\ [\mathrm{BEF}(x, x') \,\&\, \mathrm{BEF}(x', \mathrm{TPER}(a,b,\mathrm{REQUEST}(a,b,e),x)) \,\&\, \\
&\quad \mathrm{BEF}(x', x'') \,\&\, \mathrm{OCCURS}(\mathrm{COMMIT}(b,e, \mathrm{TPER}(a,b,e,x)), x'')]]) \\
&\vee\ \neg \mathrm{BEL}(a, str_3, t, (\forall x{:}T)\ (\forall x'{:}T)\ [\ \mathrm{BEF}(x, x') \,\&\, \\
&\quad \mathrm{BEF}(x', \mathrm{TPER}(a,b,e,x)) \rightarrow\ \mathrm{HOLDS}(\mathrm{SECURE}(b,e), x', \mathrm{PER}(a,b,e,x))]) \\
&\vee\ \neg \mathrm{BEL}(a, str_4, t, (\forall x{:}T)(\forall x'{:}T)\ [\mathrm{OCCURS}(\mathrm{COMMIT}(b,e,x'), x) \\
&\quad \rightarrow (\exists x''{:}T)[\mathrm{BEF}(x, x') \,\&\, \mathrm{BEF}(x', x'') \,\&\, \mathrm{OCCURS}(e, x'')]]) \\
\leftrightarrow\ &\neg \mathrm{TRUST}(a, \mathrm{CALC}(str_1, str_2, str_3, str_4), b, t, e) \qquad \textbf{(10)}
\end{aligned}$$

Note that the confidence levels now measure the lack of confidence.

Distrust is not the same as the trust not to do something. Thus, let us define two events e and e' as incompatible if whenever event e occurs, event e' does not occur and vice versa. More formally:

$$\mathrm{INCOMPATIBLE}(e,e') \leftrightarrow (\forall t)[\ \mathrm{OCCURS}(e,t) \leftrightarrow \neg\ \mathrm{OCCURS}(e',t)] \qquad \textbf{(11)}$$

Neither of the following hold:

$$(\mathrm{INCOMPATIBLE}(e,e') \,\&\, \mathrm{TRUST}(a,n,b,t,e)) \rightarrow \neg \mathrm{TRUST}(a,n,b,t,e') \qquad \textbf{(12)}$$

$$(\mathrm{INCOMPATIBLE}(e,e') \,\&\, \neg \mathrm{TRUST}(a,n,b,t,e)) \rightarrow \mathrm{TRUST}(a,n,b,t,e') \qquad \textbf{(13)}$$

In other words, the fact that a does not trust b to bring about e, does not mean that a trusts b to bring about the opposite of e, or the fact a trusts b to bring about e does not mean that a does not trust b to bring about the opposite of e. (12) does not hold because belief on my part in your competence, dependability, security and honesty with respect to e does not imply that I must believe you are either incompetent, undependable, insecure or dishonest with respect to the opposite of e. (13) does not hold because the fact that I do not believe in your competence, dependability, security or honesty with respect to e does not imply that I believe in your competence, dependability, security and honesty with respect to the opposite of e.

5.2 Trust Dynamism

A trust relationship evolves as new experiences and new information is added and incorporated.

There are two circumstances that can lead to a change in a trustor's trust level in a trustee to bring about some event. The first circumstance is simply a change in one of the constituent beliefs of the trust relationship, i.e., the belief that the trustor is competent, dependable, secure and honest. Thus, the trust that I had in WorldCom (in

2001) morphed into distrust when I discovered that their accounting practices were dishonest and misleading.

A second circumstance in which a trustor's level may change is the occurrence of some event, which although it changes the trust level, cannot immediately be traced to one of the constituent beliefs. For example, I may lose the trust in my insurance agent to get me the best rate on my car insurance, but I may not be sure whether this is due to his incompetence, or dishonesty. In order to model this situation, we introduce the predicate IMPACTS whose signature is $\Pi_{IMPACTS}$: E,A,A,E,N. Intuitively, IMPACTS(e,a,b,e',r) is true if the occurrence of the event e impacts the level of trust that a has in b to bring about e' by r, where r is a ratio. The ratio is negatively correlated to the amount of confidence I already have. That is, the more I already trust you, the less my confidence increases when there is a new positive experience. This ensures that repeated occurrences of the same positive experience do not increase my confidence level to the same extent every time they occur. We can now introduce the following axiom to model trust dynamism:

$$(\forall t{:}T)\ [(TRUST(a, n, b, t, e) \ \&\ OCCURS(e',t) \ \&\ IMPACTS(e',a,b,e,w)) \rightarrow TRUST(a, n*(1+w), b, t+1,e)] \quad \textbf{(14)}$$

Thus, if a trusts b to perform e with trust level n at time t, and at time t some event e' occurs that impacts a's trust in b to bring about e by ratio w, then a's trust level in b to bring about e at the next point in time has changed by ratio w.

The notion of impact can also be used to distinguish between positive and negative experiences. An experience is an event that one has participated in. Trust interactions that are recorded and can be referenced in the future constitute one's experiences. The value of experience is that it can be used to enhance the quality of a decision. Experience is used to increase the probability of a favorable outcome and reduce the risk profile of a relationship. In essence, using experience to guide trust decisions makes life (and the trust decision process) simpler. A positive experience implies an increased confidence level, while a negative experience implies a lowered confidence level (even distrust). In other words, a positive experience is an event that positively impacts the trustor's trust level in the trustee, while a negative experience negatively impacts it. The trust level that a trustor has in a trustee with respect to some event is the result of the cumulative experiences that the trustor has had regarding the trustee with respect to the event.

There is a specific class of negative experiences that are worth mentioning. When a trusts b to bring about e, a believes in b's competence, dependability, security and honesty with respect to e, and a therefore expects b to behave competently, dependably, securely and honestly with respect to e. b behaving competently, dependably, securely and honestly with respect to e is therefore likely to have no impact on a's trust level. However, the same cannot be said for any violation of any of these expectations. Thus, if a experiences b behaving incompetently, undependably, unsecurely or dishonestly with respect to e, then a's trust level in b to bring about e will be negatively impacted.

Finally, the experience that impacts the trustor's trust level in the trustee with respect to some action can be a purely cognitive experience, such as learning about an entity's reputation or receiving a recommendation from a third party. The concepts of experience, reputation, recommendations, business confidence, diffidence,

expectation, reliance and deception can all be modeled using the construct defined above.

5.3 Risk

There is a clear relationship between trust and uncertainty. After all, in general, when a trustor trusts a trustee to bring about an event, the trustor cannot be absolutely certain that a trustee will bring about the event. Because of the inherent uncertainty in a trust relationship, a trust relationship carries a certain risk, where risk can be regarded as the possibility or probability of incurring harm or loss. The extent to which a trustor is willing to take the risk associated with the trust relationship clearly depends on the trustor, the trustee and the nature of a particular transaction. Some trustors have a lower risk threshold than others, i.e. they are less risk-averse than others, and some trustees are inherently more trustworthy than others, no matter the transaction. Finally, trustors are more likely to trust trustees with respect to events that are less likely to cause harm to the trustor.

In order to model the relationship between risk and trust, we introduce the function symbol THRESHOLD with signature $\Sigma_{THRESHOLD}$: A,A,E,T $\rightarrow$ N. Intuitively, THRESHOLD(a,b,e,t) describes the upper bound on a's willingness to take the risk to engage b to perform event e at time t. If a's trust in b to perform e exceeds THRESHOLD(a,b,e,t) then a is willing to engage b to perform e at time t as a is certain that any request for b to perform e will result in e occurring within a reasonable period of the request being made. In other words

$$(\forall t{:}T)[(TRUSTS(a, n, b, t, e) \,\&\, (n > THRESHOLD(a,b,e,t) \,\&\, OCCURS(REQUESTS(a,b,e), t) \rightarrow BEL(a,(\exists t'{:}T)(t' < TPER(a,b,e,t) \,\&\, OCCURS(e,t'))] \quad \textbf{(15)}$$

Risk-averse entities will trust if their thresholds are not exceeded, whereas risk-loving entities do the opposite.

The important observation here is that confidence levels and risk thresholds are in the same numerical range and can be compared because there is a semantic equivalence. It should be noted that risk thresholds can be compared to explicitly derived risk values, from transaction-specific and event-driven trust decision factors.