Experience and Trust: A Sytems-Theoretic Approach

Norman Foo

The School of Computer Science and Engineering
University of New South Wales, Sydney NSW 2052, Australia
email: norman@cse.unsw.edu.au

Jochen Renz

Research School of Information Science and Engineering
Australian National University, Canberra ACT 2600, Australia
email: jochen.renz@anu.edu.au

Technical Report
UNSW-CSE-TR-0717
August 2007

THE UNIVERSITY OF
NEW SOUTH WALES

School of Computer Science and Engineering
The University of New South Wales
Sydney 2052, Australia

Abstract

An influential model of agent trust and experience is that of Jonker and Treur [Jonker and Treur 99]. In that model an agent uses its experience of the interactions of another agent to assess its trustworthiness. We show here that a key property of that model is subsumed by a result of classical mathematical systems theory. Using the latter theory we also clarify the issue of when two experience sequences may be regarded as equivalent. An intuitive feature of the Jonker and Treur model is that experience sequence orderings are respected by functions that map such sequences to trust orderings. We raise a question about another intuitive property — that of continuity of these functions, viz. that they map experience sequences that resemble each other to trust values that also resemble each other. Using fundamental results in the relationship between partial orders and topologies we also show that these two intutive properties are essentially equivalent.

1 Introduction

Before launching into a formal treatment of experience, trust and reputation it is perhaps helpful to see how these *qualities* of an agent are manifested in a contemporary application, viz., auctions on the internet in an eBay-like system. We will only outline the features essential to our discussion in the sequel. Users are given a login identification — their *ids* — which allow them to electronically buy and sell all manners of items using auctions. Each completed buying or selling transaction between agents A and B is an *interaction* or *experience*, and each can evaluate the quality of the other with respect to this interaction. The system has some way of aggregating such evaluations over time and scoring each agent on its reliability — its notion of *trust*. Thus both experience and trust are dynamic in such a system.

In the formalization of the qualities of experience and trust one persuasive view which we accept is the following. We first describe the ideas informally, then introduce the formalism. In an interacting community of cognitive agents the trustworthiness of an individual agent emerges from the perceived quality of its interactions with the rest of the community, i.e., an agent A evaluates the trustworthiness — briefly the *trust* — of another agent B based on A's *experience* of past interactions with B. As can be seen, this trust evolves with more experience.

We are aware that trust interacts with other social properties like *reputation* and *social norms*. For instance, suppose every trust evaluation of any given agent by all other agents is common knowledge to the entire community. This may then be used by agents to further evaluate the *reputation* of that agent. Observe that the reputation of an agent B as evaluated by agent A with common knowledge input may be different from that by agent C. Social norms are in fact based on trust that the agents in a community are rational enough not to deviate from the norm because it would hurt their self-interest if they did so. Therefore the evolution of an agent community toward a norm rests on the emergence of such trust. In addition the agents interacting within a social norm are aware that it is common knowledge that every agent can reason about the beliefs of other agents. This paper does not address such issues but is confined to models of agent assessment of trust without seeking to understand its deeper cognitive basis. Instead, it merely seeks to understand better the constraints of how trust may emerge from experience. However, we are currently formalizing the above intuitions on the mutual dependencies among experience, trust and reputation.

2 Formalization

We take as a starting point in the formalization of experience and trust the work of Jonker and Treur [Jonker and Treur 99] and adapt their notation. A sequel to that work is that by Treur [Treur 07] on properties of states arising from it. After summarizing their work we make a detour into mathematical systems theory. We use this systems theory to (i) connect established propositions with their work, (ii) show constraints on trust structure imposed by experience structure, (iii) suggest a way to topologize these and other derivative structures, and show that it implies the desirable order-preservation property of the maps between them.

2.1 Summary of Jonker and Treur

Jonker and Treur (op.cit) introduced most of these concepts and notations. Underlying the structures below is N, the set of natural numbers, used to model discrete time points at which interactions take place.

Experience E — A poset ($\leq$) of experience classes, e.g, $\{+,-\}$, $[-1,1]$ with $- < +$ and the usual ordering on the reals in $[-1,1]$; may be partitioned into E_{pos}, 0 and E_{neg} with implied order. The positive experience is when agent B judges the behavior of agent A to be "good" at that interaction, and the converse for negative. If 0 is in the poset it connotes a "neutral" experience.

Experience sequences ES — set of experience sequences $e = \langle e_i \rangle_{i \in N}$ partially ordered by $e \leq f \Leftrightarrow \forall i \; e_i \leq f_i$. Note that e is intuitively an infinitely long sequence of experiences.

Segments For $e \in ES$ and $k \in N$, $e|_k$ means $\langle e_1, \ldots e_k \rangle$. Call this a *k-initial segment* of e. Although it was not needed in their work, we extend their notation by defining the dual of initial segments. By $e|^k$ is meant the *k-final segment* of e, viz., the (possibly infinite) sequence $\langle e_k, e_{k+1}, \ldots \rangle$. The latter notation is useful for the next subsection.

Trust T — a poset of trust values, e.g.
$unconditional\ distrust < conditional\ distrust < conditional\ trust$
$< unconditional\ trust$. T may partition into $T_{neg} \cup \{0_T\} \cup T_{pos}$, with implied order, analogously with E.

Trust sequences TS — set of sequences of trust values $ts = \langle tv_i \rangle_{i \in N}$ where each tv_i is some value in T.

Jonker and Treur connect ES with TS in two ways. One is simply to say that there is a function $\phi : ES \rightarrow TS$ that satisfy some intuitively pleasing properties, e.g. if e^1 is better than e^2 in the sense that for all (most?) of the time then $\phi(e^1) = ts^1$ should also be better than $\phi(e^2) = ts^2$ all (most?) of the time. But another way is to think of how the "real world" works: agents can only experience interactions from a beginning time to the present time. Hence an agent at any point in time has only seen a finite initial segment $e|_k$ of a potentially infinite experience sequence. The agent has to use this to *update* its prior judgement of the trustworthiness of the other agent which is the subject of $e|_k$. Thus, another view of the function $\phi : ES \rightarrow TS$ is that it is a gradually unfolding function that does not depend on the future, i.e., if $e_1|_k = e_2|_k$ then $\phi(e_1)|_k = \phi(e_2)|_k$. This property is formalized below in definition 1. This second perspective is called *trust update* by Jonker and Treur. Their basic result is that a ϕ has a trust update version if and only if it is does not depend on the future.

2.2 Mathematical Systems Theory

It is evident that the formal model above is a *time-dynamic system* in mathematical terms. In the most general sense such systems can be regarded as "black boxes" that map input sequences (e.g. experience sequences) into output sequences (e.g. sequences of trust values). A natural model of the input space of sequences is that it is a set of functions on the natural numbers N, and similarly for the output space. The discipline that studied the structure and properties of these systems is *mathematical systems theory*, and with refinements on the black box model it has been shown to subsume much of classical engineering models and automata theory. It is to this system theory that we now appeal to examine the Jonker and Treur model in a new light.

Mathematical systems theory addresses general properties of time-dynamic systems to unify notions that are often independently repeated in different domains of application without awareness that similar (even isomorphic, homeomorphic) notions have been investigated elsewhere. We use it to place the work of Jonker and Treur in a more general setting. The results below were known in parts across a number of disciplines — engineering, automata theory, dynamics, etc. Two books, with expositions of results that most resemble the development we present below, are [Padulo and Arbib 74] and [Zeigler, et.al. 2000]. We have however distilled and collected together those aspects of systems theory that are relevant to our topic, and placed them within a unifying framework. Our other contribution is to introduce topological notions into this framework which has traditionally been principally algebraic. Using topology we will show that the intuitive requirement of continuity for the mapping from experience sequences to trust values implies the preservation of ordering from the former to the latter.

In mathematical systems theory the most basic structure is a formalization of the concept of a "black box" which accepts inputs produces outputs. See figure 2.1 for a diagrammatic representation. In case the inputs are functions on time then so are the outputs. In the discrete time model such functions are simply (possibly infinite) sequences, say, ω in the *input space* Ω and λ in the *output space* Λ. More explicitly $\omega(n)$ (resp. $\lambda(n)$) is the n-th component of the input (resp. output) sequence. Each ω maps N to a *value space* V_{in}, e.g., if ω is an experience sequence in the above sense, then V_{in} can be the poset E of experience classes; similarly each λ maps N to a value space V_{out}, e.g., if the black box converts experience sequences into trust ranks as above then V_{out} can be the poset T of trust. The notions of initial and final segments above apply in the obvious manner to such sequences, e.g., $\omega|_k$ is the function on N that is defined only for arguments $n \leq k$ and in that range it agrees with ω. The black box is modeled as a function $F : \Omega \to \Lambda$ and we will refer to such systems as *I-O systems*.

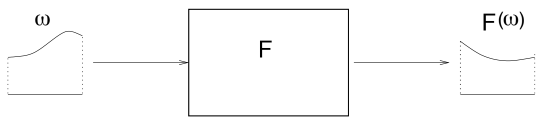

Figure 2.1: Input-Output System

The I-O systems of interest to us are the *causal* ones. Informally these are systems that "do not look into the future" to produce outputs. They correspond to the Jonker and Treur systems that have trust updates, and we will show below that their result is subsumed by a standard proposition of mathematical systems theory.

Definition 1 *An I-O system is causal if its input-output function F satisfies:* $\forall k, \omega_1|_k = \omega_2|_k \Rightarrow F(\omega_1)|_k = F(\omega_2)|_k$.

Causal I-O systems are precisely those whose black boxes can be ascribed an internal state structure. Although versions of this construction are implicit in the the above cited references ([Padulo and Arbib 74] and [Zeigler, et.al. 2000]), for completeness we will outline it using our notation and formalism.

Let $\bar{\Omega} = \{\omega|_k \mid \omega \in \Omega \; and \; k \in N\}$, i.e. the collection of all initial segments of Ω. Analogously $\bar{\Lambda}$ is the collection of all initial segments of Λ. An algebraic structure can be imposed on $\bar{\Omega}$ (respectively, $\bar{\Lambda}$) by concatenation of segments, i.e., contiguously placing one segment after another. A little bit of machinery is necessary to express the notion of translating the second segment into the proper place on the domain N. Let $len(\omega)$ be the length of the domain of ω, the difference between its two end points of definition. The *shift function* S_k applied to ω yields the translation of ω to the right by k units, i.e., $S_k(\omega)(n)$ is undefined for $n < k$ and $n > (k + len(\omega))$, and for $k \le n \le (k + len(\omega))$ its value is $\omega(n - k)$. Then the *concatenation* of initial segment ω_1 by initial segment ω_2 (the latter following the former) is the new initial segment $\omega_1 \circ \omega_2$ defined by $(\omega_1 \circ \omega_2)(n) = \omega_1(n)$ if $n \le len(\omega_1)$, and $(\omega_1 \circ \omega_2)(n) = S_{len(\omega_1)}\omega_2(n)$ if $n > len(\omega_1)$. With concatenation $\bar{\Omega}$ becomes a monoid semigroup $\langle \bar{\Omega}, \circ \rangle$ in the obvious way, with the null initial segment ϵ as the identity. Similar comments apply to $\langle \bar{\Lambda}, \circ \rangle$. For brevity below we will write $\bar{\Omega}$ (respectively, $\bar{\Lambda}$) below for this semigroup.

Corollary 1 *$F : \Omega \to \Lambda$ is causal if and only if it induces a function $\bar{F} : \bar{\Omega} \to \bar{\Lambda}$.*

The forward implication of Corollary 1 follows directly from definition 1. The backward direction can be shown by constructing a simple example of a non-causal F that would make $\bar{F}$ non-functional. This corollary exhibits the essence of what "no look ahead" means.

Definition 2 *The segments ω_1 and ω_2 in $\bar{\Omega}$ are Nerode-equivalent, written $\omega_1 \equiv \omega_2$, if $\bar{F}[(\omega_1 \circ \mu)(k + len(\omega_1))] = \bar{F}[(\omega_2 \circ \mu)(k + len(\omega_2))]$ for $1 \le k \le len(\mu)$ and for every $\mu \in \bar{\Omega}$.*

Notation 1 *Using the earlier notation $e|^k$ to denote the segment beyond the initial k-length segment of sequence e, we may re-write this less formally but more succinctly as $\bar{F}(\omega_1 \circ \mu)|^{len(\omega_1)+1} = \bar{F}(\omega_2 \circ \mu)|^{len(\omega_2)+1}$ for every $\mu \in \bar{\Omega}$.*

Figure 2.2 graphically illustrates the idea in definition 2. Informally $\omega_1 \equiv \omega_2$ if $\bar{F}$ produces exactly the same outputs — viewed after the end of the segments ω_1 and ω_2 — due to any segment is concatenated to them. An informal way to think about this is that the segments ω_1 and ω_2 cannot be distinguished once their end points are reached, which is indeed the intuition for the state construction justifying proposition 1.

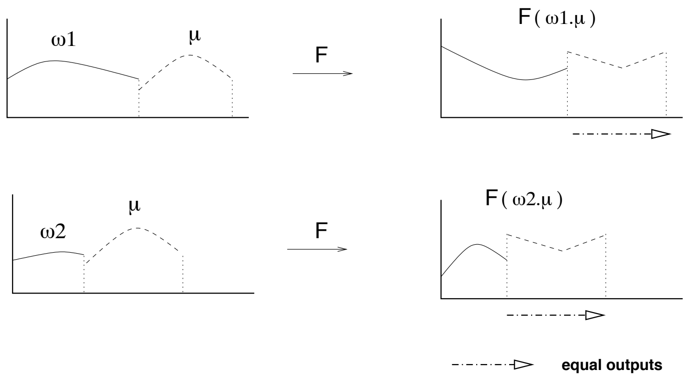

Figure 2.2: Nerode Equivalence

Corollary 2 *The Nerode equivalence is a right-congruence with respect to concatenation, i.e., if $\omega_1 \equiv \omega_2$ then $\omega_1 \circ \mu \equiv \omega_2 \circ \mu$ for any μ.*

This follows directly from the definition 2.

Definition 3 *A state realization of a function $\bar{F} : \bar{\Omega} \rightarrow \bar{\Lambda}$, with V_{out} as the value space of λ in Λ, is a tuple $\langle \bar{\Omega}, Q, q_0, \delta, \eta \rangle$ where Q is the state set, q_0 is the initial state, $\delta : Q \times \bar{\Omega} \rightarrow Q$ is the state transition function, $\eta : Q \times \bar{\Omega} \rightarrow \bar{\Lambda}$ is the output function such that $\eta(\delta(q_0, \omega)) = \bar{F}(\omega)$ and $\delta(q, \epsilon) = q$ where ϵ is the null segment.*

In the following proposition and its proof it is helpful refer to figure 2.3.

Proposition 1 *Every causal I-O function F has a state realization.*

Proof

Since a causal I-O function F is essentially the same as its induced function $\bar{F}$ we will use the latter in the proof of Proposition 1. The proof will be by constructing the elements of tuple $\langle \bar{\Omega}, Q, q_0, \delta, \eta \rangle$. The quotient map $\psi : \bar{\Omega} \rightarrow \bar{\Omega}/\equiv$ will be referred to later in the paper; it takes each input segment to its equivalence class as explained next. For Q we take the quotient space $\bar{\Omega}/\equiv$, and use the null segment as q_0; hence $\psi(\epsilon) = q_0$ and $\psi(\omega) = \omega/\equiv\, = [\omega]_{\equiv}$, i.e. the quotient class of ω under the equivalence (congruence) relation $\equiv$. The state transition function $\delta : Q \times \bar{\Omega} \rightarrow Q$ is then defined as $\delta([\omega_1]_{\equiv}, \omega_2) = [\omega_1 \circ \omega_2]_{\equiv}$. That δ well-defined follows from Corollary 2. The output function η is defined by $\eta([\omega)_1]_{\equiv}, \omega_2)$ = $\bar{F}(\omega_1 \circ \omega_2)]^{len(\omega_2)}$. (Recall the notation $e|^k$ means the k length final segment of sequence e.) We need to show this is well-defined. So suppose $\omega \equiv \omega'$. Then $\bar{F}(\omega \circ \omega_2)|^{len(\omega_2)}$ = $\bar{F}(\omega' \circ \omega_2)|^{len(\omega_2)}$ follows from the notation 1 since $\omega \equiv \omega'$. Further, $\eta([\epsilon]_{\equiv}, \omega)$ = $\bar{F}(\epsilon \circ \omega) = \bar{F}(\omega)$ as required.

Observation 1 *In figure 2.3 the map F' is defined from the map $\bar{F}$ in the following manner. Suppose ω is an initial segment; then $\bar{F}(\omega) = \lambda$ is an output initial segment which is some sequence $\langle v_1, v_2, \ldots, v_k \rangle$ where each v_i is in V_{out}. Then $F'(\omega) = v_k$. Conversely it is not hard to see that given F' we can define $\bar{F}$ from it. In automata theory [Padulo and Arbib 74] these equivalent descriptions of the input-output maps correspond to Moore and Mealy automata. In that figure $\gamma([\omega]_{\equiv})$ corresponds to (the singleton entry of) $\eta([\omega]_{\equiv}, \epsilon)$.*

$$\begin{tikzcd} \bar{\Omega} \arrow[rr, "F'"] \arrow[dr, "\psi"'] & & V_{out} \\ & \Omega / \equiv \arrow[ur, "\gamma"'] & \end{tikzcd}$$

$$\bar{\Omega} / \equiv \times \bar{\Omega} \xrightarrow{\eta} \bar{\Lambda}$$

Figure 2.3: State Realization – The Key Ideas

Observation 2 *The state realization in Proposition 1 is what is also known in systems theory as a* transition system *and it is easy to see that every transition system defines a causal I-O system. Hence, Proposition 1 can be strengthened to an "if and only if" statement.*

Observation 3 *It can be shown that the state realization above, call it R, is in a strong sense the most economical among all possible state representations. Formally, it is said that this realization is* canonical *in that if there is another realization R' that reproduces the same $\bar{F}$, then there is a unique homomorphism that maps R' to R. In particular a typical assumption (see e.g., Treur [Treur 07]) that input (trust, etc.) sequences and system states are both viable primitives in formalizing temporal dynamics is subject to this canonical constraint.*

One of the main Jonker and Treur results, that only non-predictive maps from experience sequences to trust sequences has a update function is now simply an interpretation of Proposition 1 in which we interpret the input segment space $\bar{\Omega}$ as initial experience sequences, and the state space Q as trust states (= trust values in their formalization). The advantage of the systems theoretic view is that it highlights constraints as well as suggests generalizations. One generalization is topologies for the experience class E and trust space T, where Proposition 1 suggests the constraint that any topology for T ("states") is inherited from the topology of E ("inputs"). We discuss this in a later section.

Let us now consider the "size" of T, knowing that it is effectively the same as $\bar{\Omega}/\equiv$. For simplicity assume that the input value space V_{in} is the ordered discrete set {*worst, bad, neutral, good, best*}, with ordering as suggested, i.e. *worst* $<$ *bad* $<$ *neutral* ... etc. We now argue that unless $\equiv$ aggregates many input segments in $\bar{\Omega}$ in each equivalence class, the trust space T will be very large. To see this, consider just a particular segment subset, denoted by Ω_k, of all the segments of length k. There are 5^k distinct segments in it, so if each of them constitutes a singleton equivalence class the number of classes is also 5^k. Over all input segments up to length n in the worst case there will be $\sum_{k=0}^{n} 5^k$ elements in T. It is easily seen that the cardinality of T is infinite in the worst case when all of $\bar{\Omega}$ is considered. The systems-theoretic result that $\bar{F} : \bar{\Omega} \rightarrow \bar{\Lambda}$ has a *finite state realization* if and only if the index[1] of $\equiv$ is finite should now be obvious.

The worst case is of course highly unlikely, as in any realistic application many distinct input segments will give rise to the same trust value. An example of how this can happen is in the *finite memory* case where an agent just looks at, say, the experience values of the most recent 10 interactions. This reduces the input segments that need to be considered to simply those in Ω_{10}. A further collapse is possible if the agent does not distinguish between permutations of the 10-sequence experience values. For ease of exposition we therefore temporarily assume that $\bar{\Omega}/\equiv$ is finite.

Assumption 1 *A neutral experience leaves the trust evaluation of the agent unchanged.*

This assumption is tantamount to the following. Suppose $e = \langle e_1, e_2, \ldots, e_k \rangle$ and $e' = \langle e_1, e_2, \ldots, e_k, neutral, \ldots, neutral \rangle$. Then $\bar{F}(e) = \bar{F}(e')$. We are aware that this assumption is not as innocent as it looks. Suppose an agent has been assessed with *good* experience up to some point in time k, and thereafter it only behaves with *neutral* experience. Might there not be grounds to suspect that its trustworthiness has

[1] Number of equivalence classes.

diminished beyond k? Moreover, although this assumption is weaker than one that simply ignores neutral values in a sequence (thereby making it equivalent to a shorter one with neutral values deleted), one can argue that an agent whose behavior oscillates between *neutral* and *best* suggests that it is erratic and therefore unreliable. In other words, it is easy to find *epistemic contexts* in which the *neutral* element is significant. These issues are currently being investigated. For this paper we confine ourselves to the case where *neutral* merely signifies "no opinion".

Definition 4 *Suppose* $\bar{\Omega}/\equiv$ *is finite. Call the* minimal k *such that* $\bar{\Omega_k}/\equiv\ = \bar{\Omega}/\equiv$ *the* saturation length *of* $\bar{\Omega}/\equiv$.

The saturation length formalizes the notion of "sufficient experience" to explore all trust values. In the finite memory example above the saturation length is 10. It is a special case of the following observation. If k is known to be the saturation length then any segment from Ω_j where $k < j$ will be equivalent to some segment in Ω_k. Therefore, since T is the surjective image of Ω_k it suffices to consider only the latter for the structure of T.

Observation 4 *If the index of* $\equiv$ *is finite then under assumption 1 there is some k which is the saturation length.*

For, if some experience sequence e of length j is mapped by ψ to a trust value tv then any experience sequence which extends e with only neutral values leaves tv unchanged. Hence $\psi(\Omega_j) \subseteq \psi(\Omega_k)$ if $j < k$, where ψ is the map that takes input segments to their quotients (see notation in the proof of Proposition 1).

3 Experience and Trust Orderings

What about the ordering *worst* < *bad* < *neutral* ... in experience values? How are they reflected in the trust space T? Consider input segment experience sequences of saturation length. Intuitively we would like to have the sequence $\langle best, best, \ldots, best\rangle$ map to the highest trust value, and $\langle worst, worst, \ldots, worst\rangle$ to the lowest trust value, with all others mapping to those in between, i.e., we want the map from experience sequences to trust to *preserve* the ordering of the the former. To formalize this intuition we first use the ordering $<$ on experience values to induce a partial order $\sqsubset_{ES}$ on the experience segments in Ω_k where k is the saturation length. One way to do this is as follows: define the partial order $\sqsubseteq_{ES}$ on ES by $\omega_1 \sqsubseteq_{ES} \omega_2$ if $\omega_1(i) \leq \omega_2(i)$ for $1 \leq i \leq k$ (whence $\omega_1 \sqsubset \omega_2$ means $\omega_1 \sqsubseteq_{ES} \omega_2$ but $\omega_1 \neq \omega_2$). But other ways are also possible, not excluding unusual ones like only considering the experience at even numbered times, or only the last 10 experiences; in many cases these will amount to modifying the range qualifier above ($1 \leq i \leq k$) to another one. Generally, the application domains and the particular models of agent cognition determine how best to define the partial order.

Example 1 *Suppose only the last 5 experiences matter. The partial order on* $\bar{\Omega}$ *is given by* $\omega_1 \sqsubseteq_{ES} \omega_2$ *iff* $\omega_1|^{len(\omega_1)-i} \leq \omega_2|^{len(\omega_2)-i}$ *for* $0 \leq i \leq 5$.
If $\omega_1 = \langle worst, bad, bad, best, good\rangle$ *and* $\omega_2 = \langle bad, bad, good, good, best\rangle$ *then the supremum of these is* $\omega_3 = \langle bad, bad, good, best, best\rangle$. *Moreover* $\omega_1 \not\sqsubseteq_{ES} \omega_2$ *and* $\omega_2 \not\sqsubseteq_{ES} \omega_1$. *Sequences longer than 5 will be equivalent to some sequence of length 5, e.g., the sequence* $\omega_4 = \langle best, good, neutral, worst, bad, bad, best, good\rangle$ *is equivalent to* ω_1 *as they have identical "most recent" five experiences.*

As suggested by the above example and discussion, an informal way to interpret the partial order $\sqsubset_{ES}$ is to think of $\omega_1 \sqsubset \omega_2$ is saying "ω_2 is a better experience sequence than ω_1". Suppose the trust space T also has a partial order $\sqsubseteq_T$ (whence $t_1 \sqsubset_T t_2$ means $t_1 \sqsubseteq_T t_2$ but $t_1 \neq t_2$). The desired preservation property is $\omega_1 \sqsubseteq_{ES} \omega_2 \Rightarrow \psi(\omega_1) \sqsubseteq_T \psi(\omega_2)$ which is to say that ψ is *monotonic* in the partial order. We argued above that this is surely the case for the best possible experience sequence mapping to the best trust, and dually for the worst cases. Monotonicity guarantees that the (partial) ordering is also observed for all other sequences — informally it says that better sequences map to better trust values. This is surely an intuitive property that one would like for any map from experience sequences to trust states. However, besides monotonicty there is another property of such maps that is intuitively appealing, viz. *continuity*. Informally, continuity of a map may be paraphrased as "things that are near to each other get mapped to things that are near to each other". Hence we would expect that two experience sequences that closely resemble each other should map to two trust values that are near to each other. Continuity needs some notion of "nearness", i.e. a *topology*. It would be satisfying if these two intuitions can be shown to arise from some fundamental connection, and it is to this that we now turn.

There is a close connection between partial orders — in fact pre-orders will do — and *topologies* on a space. The facts below are well-known and [Arenas 99] or [Wiki Alexandrov] may be consulted for details. For deeper topological notions the standard reference [Kelley 55] may be consulted. Given a partial order $\sqsubseteq$ on a space S, the Alexandrov topology defined by it has as open sets the so-called *up-sets*, viz., subsets θ such that $x \in \theta$ and $x \sqsubseteq z$ implies $z \in \theta$. See figure 3.1 for a geometric interpretation

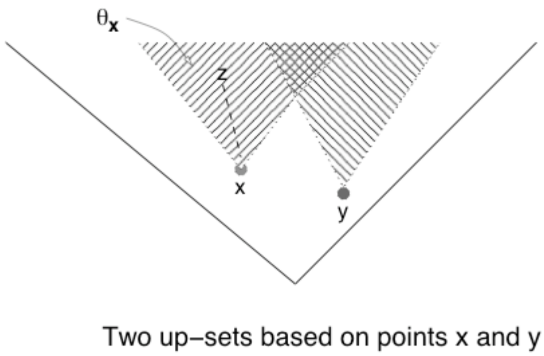

Figure 3.1: Geometric picture of up-sets

Conversely, given a topology τ on a set S, the *specialization pre-order* $\leq$ is defined by $x \leq y$ iff y is in every open set that contains x. It is easily seen that $\leq$ so defined is indeed a pre-order. If we had started with some partial order $\sqsubseteq$ and used it to define the Alexandrov topology as before, it is natural to ask what is the specialization order that arises from that topology. The answer is that we get back $\sqsubseteq$, and although there are other topologies (e.g. the Scott topology [Abramsky and Jung 94] or [Stoy 77]) that have this "reversal" property the Alexandrov topology is the finest one. In this way the partial order $\sqsubseteq_{ES}$ defines the Alexandrov topology on the input segment space $\bar{\Omega}$ (which in our context is identified with the space of experience sequences ES) and is induced by it.

The following lemma is described without proof in [Wiki Alexandrov] and is well-known, but we provide its easy proof for completeness.

Lemma 1 *If $\phi : (X, \tau_1) \to (Y, \tau_2)$ is a continuous map from topological space X with topology τ_1 to topological space Y with topology τ_2 then ϕ is order-preserving from $(X, \leq_1)$ to $(Y, \leq_2)$ where $\leq_1$ and $\leq_2$ are the specialization partial orders in the respective spaces.*

Proof
Suppose $x \leq_1 y$ in X. Let θ be an open neighborhood of $\phi(x)$ in Y. Since ϕ is continuous $\phi^{-1}(\theta)$ is open. By definition $x \in \phi^{-1}(\theta)$. Since $x \leq_1 y$, by definition of $\leq_1$ the open neighborhood $\phi^{-1}(\theta)$ of x contains y. Hence $\phi(y)$ is in θ, showing that $\phi(x) \leq_2 \phi(y)$.

Any topology that is placed on the trust space T will induce a specialization pre-order (partial orders are special cases). So what is a suitable topology for it? If we identify T with the range of ψ, i.e., $\bar{\Omega}/\equiv$, then T is the quotient space of $\bar{\Omega}$. $\bar{\Omega}/\equiv$ can thus be given the *quotient topology* [Kelley 55]. It is perhaps helpful to recall the gist of such quotient topologies in general as follows: if $f : X \to Y$ is a map from topological space X to space $Y = X/R$ where R is an equivalence relation on X, the finest topology on X/R such that f is continuous is the quotient topology.

Proposition 2 *If T has the quotient topology via ψ of ES which has the Alexandrov topology induced by its partial order $\sqsubseteq_{ES}$, then ψ is order-preserving with respect to $\sqsubseteq_{ES}$ and the specialization partial order $\sqsubseteq_T$ (of the quotient topology) on T.*

Proof
In our context, if T is given the quotient topology via ψ of the Alexandrov topology in ES, then by definition ψ is continuous. The quotient topology so induced in T defines the specialization partial order, say $\sqsubseteq_T$. Hence the map ψ can also be interpreted as one between two partial orders. Since ψ maps ES to its quotient space with the induced quotient topology, by definition it is continuous. By lemma 1 continuity of ψ implies order-preservation by ψ.

Observation 5 *It is not hard to see that $\sqsubseteq_T$ does reflect the intuition that $t_1 \sqsubseteq_T t_2$ means that $t1$ is lower in trust estimation than t_2 by considering example shown in figure 3.2.*

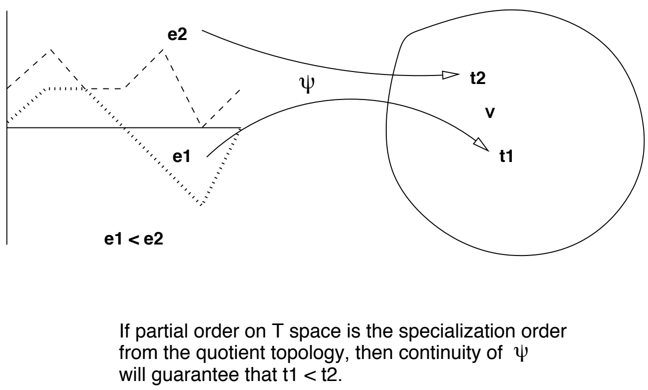

Figure 3.2: The quotient map from experience sequences to trust

Experience values in the real interval $[-1, 1]$ rather than finite or even discrete values do not alter the character of the results and observations above except that the experience and trust spaces can now be infinite and continuous. In particular everything that was said about the relationship between the orderings and the topologies still hold. However, the additional structure of the reals may afford refinement of some of the above results, an issue we postpone to future work.

4 Conclusion

We used classical mathematical systems theory to underpin the foundations of an influential model of agent trust and experience. An intuitive feature of that model is that experience sequence orderings are respected by functions that map such sequences to trust orderings. We raised a question about another intuitive property — that of continuity of these functions, viz. that they map experience sequences that resemble each other to trust values that also resemble each other. Using fundamental results in the relationship between partial orders and topologies we showed that these two intutive properties are essentially equivalent.

Bibliography

[Wiki Alexandrov] See the Wikipedia entry on Alexandrov topology: *http://en.wikipedia.org/wiki/Alexandrov_topology*.

[Arenas 99] F.G. Arenas, Alexandroff spaces, Acta Math. Univ. Comenianae Vol. LXVIII, 1 (1999), pp. 17-25

[Padulo and Arbib 74] L. Padulo and M.A. Arbib, *System theory: a unified state-space approach to continuous and discrete systems*, Philadelphia, Saunders, 1974.

[Zeigler, et.al. 2000] B.P. Zeigler, H. Praehofer and T.G. Kim, *Theory of Modeling and Simulation :integrating discrete event and continuous complex dynamic systems*, 2nd ed, Academic Press, San Diego, 2000.

[Jonker and Treur 99] C.M. Jonker and J. Treur. Formal Analysis of Models for the Dynamics of Trust based on Experiences. In: F.J. Garijo, M. Boman (eds.), Multi-Agent System Engineering, Proceedings of the 9th European Workshop on Modelling Autonomous Agents in a Multi-Agent World, MAAMAW'99. Lecture Notes in AI, vol. 1647, Springer Verlag, Berlin, 1999, pp. 221-232.

[Stoy 77] J.E. Stoy, *Denotational Semantics: The Scott-Strachey Approach to Programming Language Semantics*. MIT Press, Cambridge, Massachusetts, 1977.

[Treur 07] J. Treur. Temporal Factorisation: Realisation of Mediating State Properties for Dynamics. Cognitive Systems Research Journal, vol. 8, 2007 (to appear).

[Abramsky and Jung 94] S. Abramsky and A. Jung: Domain theory. In S. Abramsky, D. M. Gabbay, T. S. E. Maibaum, (ed), Handbook of Logic in Computer Science, vol. III. Oxford University Press, 1994.

[Kelley 55] J.L. Kelley, *General Topology*, Springer Verlag, 1955 (reprinted).