

Modelling Trust Dynamics from a Neurological Perspective

Mark Hoogendoorn, S. Waqar Jaffry, Jan Treur

VU University Amsterdam, Department of Artificial Intelligence

De Boelelaan 1081, 1081 HV Amsterdam

Email: {mhoogen, swjaffry, treur}@few.vu.nl

URL: <http://www.few.vu.nl/~{mhoogen, swjaffry, treur}>

Abstract. Trust is often assumed to depend on experiences. Models for the dynamics of trust in relation to experiences usually have a cognitive nature, leaving affective aspects out of consideration. However, neurological findings show more and more how in mental processes cognitive and affective aspects are intertwined. In this paper, by adopting neurological theories on the role of emotions and feelings, a model for trust dynamics is introduced incorporating the relation between trust and feeling. The model makes use of a Hebbian learning principle and describes how trust does not only depend on experiences viewed as information obtained over time, but also on emotional responses and feelings related to experiences.

1 Introduction

For a human performing certain task, an important aspect is the trust in various information sources that can potentially aid in performing the task. For instance, the person may have experienced in the past that the manual for the task contains a lot of flaws, whereas it was experienced that a knowledgeable colleague can immediately aid in an adequate fashion. As a result the trust level for the manual will be much lower than the trust level for the colleague, and a support system offering the manual will most likely be ignored. This illustrates two main functional properties of trust states as cognitive states that are often considered (e.g., Jonker and Treur, 2003):

- (1) A trust state accumulates experiences over time
- (2) Trust states are used in decision making by choosing more trusted options above less trusted options

In trust research, a variety of computational models have been proposed for human trust, see e.g., (Jonker and Treur, 1999, 2003; Falcone and Castelfranchi, 2004; Hoogendoorn, Jaffry, and Treur, 2008). Such models often assume that trust values over time are defined by a certain trust function, which expresses the dependency on the experiences of the human with the specific object of trust. In (Hoogendoorn, Jaffry, and Treur, 2008) an additional factor is taken into account, namely how the object of trust performs relative to its competitors. Most of such models consider experiences and trust as cognitive concepts and exclude affective factors. This contrasts with how persons in general experience trust and trust-affecting experiences, which, for example, may go hand in hand with strong feelings of disappointment or insecurity. Much work reports interactions between cognitive and affective aspects for a variety of cases (e.g., Eich, Kihlstrom, Bower, Forgas, and Niedenthal, 2000; Forgas, Laham, and Vargas, 2005; Forgas, Goldenberg, and Unkelbach, 2009; Niedenthal, 2007; Schooler and Eich, 2000; Winkielman, Niedenthal, and Oberman, 2009), without relating this explicitly to neurological findings or theories. In the current paper, neurological theories on emotion and feeling are adopted as a basis for a computational model for trust dynamics that models such interactions. The computational model, which is based on neurological theories on the embodiment of emotions as described, for example, in (Damasio, 1994, 1996, 1999, 2004; Winkielman, Niedenthal, and Oberman, 2009), describes how trust dynamics may not only depend on (external) informational sources, but may also take into account how with a certain strength emotional responses occur that lead to certain feelings. More specifically, in accordance with, for example (Damasio, 1999, 2004), for feeling the emotion associated to a mental state, a converging recursive body loop is assumed. In addition, based on Hebbian learning (cf. Hebb, 1949; Bi and Poo, 2001; Gerstner and Kistler, 2002) for the strength of the connections to the emotional responses an adaptation process is introduced, inspired by the Somatic Marker Hypothesis (Damasio, 1994, 1996). Compared to more detailed neural models, the model presented here abstracts from more fine-grained descriptions of neurons, and, for example, their biochemistry and spiking patterns. States in the presented model may be viewed as abstract descriptions of states of neurons or rather of groups of neurons. Such a more abstract representation provides a less complex and more manageable model, while the assumption is that it can still show the essential dynamics from a higher level perspective.

In this paper, first in Section 2 Damasio's theory on the generation of feelings based on a body loop is briefly introduced. In Section 3 the model is described in detail, including the adaptation based on Hebbian learning. Section 4 presents some simulation results. In Section 5 it is discussed how functional properties of trust states as cognitive states were formulated and automatically verified. Finally, Section 6 is a discussion.

2 On the Interaction Between Affective and Cognitive Aspects

Cognitive states of a person, such as sensory or other representations often induce emotions felt within this person, as described by neurologist Damasio (1999, 2004); for example:

'Even when we somewhat misuse the notion of feeling – as in “I feel I am right about this” or “I feel I cannot agree with you” – we are referring, at least vaguely, to the feeling that accompanies the idea of believing a certain fact or endorsing a certain view. This is because believing and endorsing *cause* a certain emotion to happen. As far as I can fathom, few if any exceptions of any object or event, actually present or recalled from memory, are ever neutral in emotional terms. Through either innate design or by learning, we react to most, perhaps all, objects with emotions, however weak, and subsequent feelings, however feeble.' (Damasio, 2004, p. 93)

In some more detail, emotion generation via a body loop roughly proceeds according to the following causal chain; see Damasio (1999, 2004):

cognitive state → preparation for the induced bodily response → induced bodily response → sensing the bodily response → sensory representation of the bodily response → induced feeling

As a variation, an 'as if body loop' uses a direct causal relation

preparation for the induced bodily response → sensory representation of the induced bodily response

as a shortcut in the causal chain. The body loop (or as if body loop) is extended to a recursive body loop (or recursive as if body loop) by assuming that the preparation of the bodily response is also affected by the state of feeling the emotion:

feeling → preparation for the bodily response

as an additional causal relation. Such recursiveness is also assumed by Damasio (2004), as he notices that what is felt by sensing is actually a body state which is an internal object, under control of the person:

'The brain has a direct means to respond to the object as feelings unfold because the object at the origin is inside the body, rather than external to it. The brain can act directly on the very object it is perceiving. It can do so by modifying the state of the object, or by altering the transmission of signals from it. The object at the origin on the one hand, and the brain map of that object on the other, can influence each other in a sort of reverberative process that is not to be found, for example, in the perception of an external object.' (...)

'In other words, feelings are not a passive perception or a flash in time, especially not in the case of feelings of joy and sorrow. For a while after an occasion of such feelings begins – for seconds or for minutes – there is a dynamic engagement of the body, almost certainly in a repeated fashion, and a subsequent dynamic variation of the perception. We perceive a series of transitions. We sense an interplay, a give and take.' (Damasio, 2004, pp. 91-92)

Within the model presented in this paper both the bodily response and the feeling are assigned a level or gradation, expressed by a number, which is assumed dynamic; for example, the strength of a smile and the extent of happiness. The causal cycle is modelled as a positive feedback loop, triggered by a mental state and converging to a certain level of feeling and body state. Here in each round of the cycle the next body state has a level that is affected by both the mental state and the level of the feeling state, and the next level of the feeling is based on the level of the body state.

Another neurological theory addressing the interaction between cognitive and affective aspects can be found in Damasio's Somatic Marker Hypothesis; cf. (Damasio, 1994, 1996; Bechara and Damasio, 2004; Damasio, 2004). This is a theory on decision making which provides a central role to emotions felt. Within a given context, each represented decision option induces (via an emotional response) a feeling which is used to mark the option. For example, a strongly negative somatic marker linked to a particular option occurs as a strongly negative feeling for that option. Similarly, a positive somatic marker occurs as a positive feeling for that option. Damasio describes the use of somatic markers in the following way:

'the somatic marker (...) forces attention on the negative outcome to which a given action may lead, and functions as an automated alarm signal which says: beware of danger ahead if you choose the option which leads to this outcome. The signal may lead you to reject, *immediately*, the negative course of action and thus make you choose among other alternatives. (...) When a positive somatic marker is juxtaposed instead, it becomes a beacon of incentive. (...) on occasion somatic markers may operate covertly (without coming to consciousness) and may utilize an 'as-if-loop'.' (Damasio, 1994, pp. 173-174)

Usually the Somatic Marker Hypothesis is applied to provide endorsements or valuations for options for a person's actions, thus shaping the decision process. Somatic markers may be innate, but may also be adaptive, related to experiences:

'Somatic markers are thus acquired through experience, under the control of an internal preference system and under the influence of an external set of circumstances which include not only entities and events with which the organism must interact, but also social conventions and ethical rules. (Damasio, 1994, p. 179)

In the model introduced below, this adaptive aspect will be modelled as Hebbian learning; cf. (Hebb, 1949; Bi and Poo, 2001; Gerstner and Kistler, 2002). Viewed informally, in the first place it results in a dynamical connection strength obtained as an accumulation of experiences over time (1). Secondly, in decision making this connection plays a crucial role as it determines the emotion felt for this option, which is used as a main decision criterion (2). As discussed in the introduction, these two properties (1) and (2) are considered two main functional, cognitive properties of a trust state. Therefore they give support to the assumption that the strength of this connection can be interpreted as a representation of the trust level in the option considered.

3 The Detailed Model for Trust Dynamics

Informally described theories in scientific disciplines, for example, in biological or neurological contexts, often are formulated in terms of causal relationships or in terms of dynamical systems. To adequately formalise such a theory the hybrid dynamic modelling language LEADSTO has been developed that subsumes qualitative and quantitative causal relationships, and dynamical systems; cf. (Bosse, Jonker, Meij and Treur, 2007). This language has been proven successful in a number of contexts, varying from biochemical processes that make up the dynamics of cell behaviour (cf. Jonker, Snoep, Treur, Westerhoff, Wijngaards, 2008) to neurological and cognitive processes (e.g., Bosse, Jonker, Los, Torre, and Treur, 2007; Bosse, Jonker, and Treur, 2007, 2008). Within LEADSTO a temporal relation $a \rightarrow b$ denotes that when a state property a occurs, then after a certain time delay (which for each relation instance can be specified as any positive real number), state property b will occur. In LEADSTO both logical and numerical calculations can be specified in an integrated manner; a dedicated software environment is available to support specification and simulation.

3.1 Incorporating affective aspects in a trust model

An overview of the model for how trust dynamics emerges from the experiences is depicted in Figure 1. How decisions are made, given these trust states is depicted in Figure 2. These pictures also show representations from the detailed specifications explained below. However, note that the precise numerical relations between the indicated variables V shown are not expressed in this picture, but in the detailed specifications of properties below, which are labeled by LP1 to LP11 as also shown in the pictures. The detailed specification (both informally and formally) of the model is presented below. Here capitals are used for (assumed universally quantified) variables. First the part is presented that describes the basic mechanisms to generate a belief state and the associated feeling. The first dynamic property addresses how properties of the world state can be sensed.

LP1 Sensing a world state

If world state property W occurs of strength V
then a sensor state for W of strength V will occur.
 $world_state(W, V) \rightarrow sensor_state(W, V)$

Note that this generic dynamic property is used for a specific world state, for experiences with the different options and for body states; to this end the variable W is instantiated respectively by w , $exp1$ and $exp2$, $b1$ and $b2$. From the sensor states, sensory representations are generated according to the dynamic property LP2. Note that also here for the example the variable P is instantiated as indicated.

LP2 Generating a sensory representation for a sensed world or body state

If a sensor state for world state or body state property P with level V occurs,
then a sensory representation for P with level V will occur.
 $sensor_state(P, V) \rightarrow srs(P, V)$

For a given world state representations for a number of options are activated:

LP3 Generating an option for a sensory representation of a world state

If a sensory representation for w with level V occurs
then a representation for option o with level V will occur
 $srs(w, V) \rightarrow rep(o, V)$

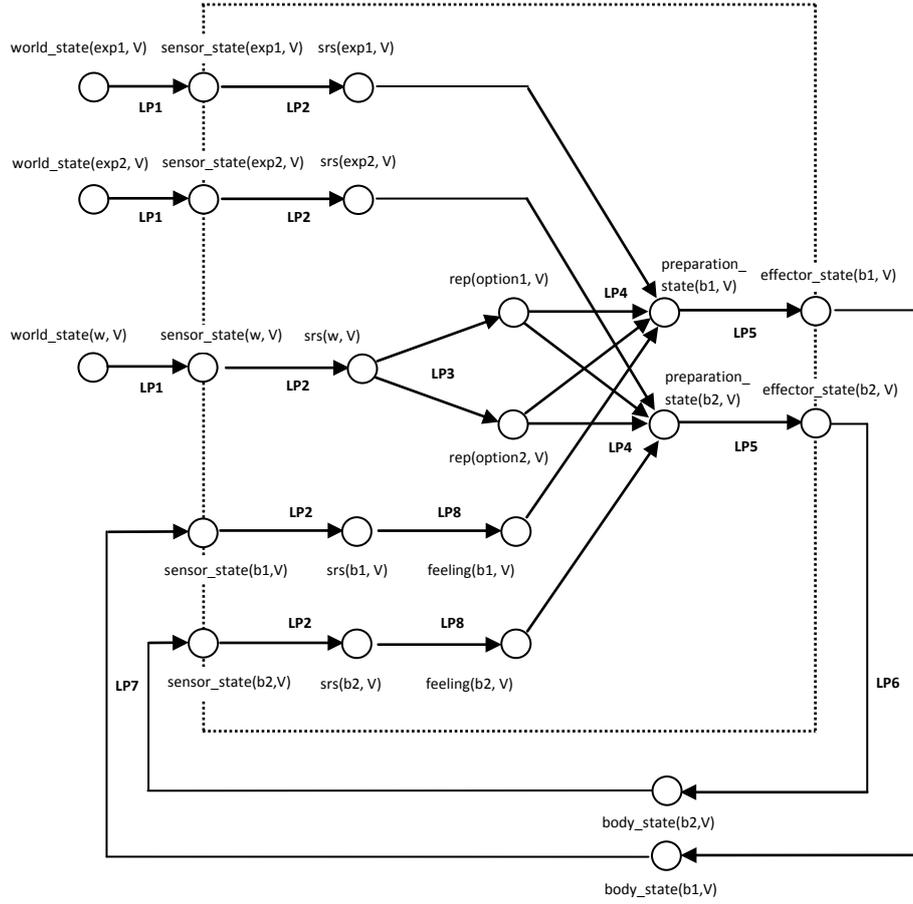


Figure 1: Overview of the model for trust dynamics

Dynamic property LP4 describes the emotional response to the person's mental state in the form of the preparation for a specific bodily reaction. Here the mental state comprises a number of cognitive and affective aspects: options activated, experienced results of options and feelings.

This specifies part of the loop between feeling and body state. This dynamic property uses a combination model based on a function

$$g(\beta_1, \beta_2, V_1, V_2, V_3, \omega_1, \omega_2, \omega_3)$$

including a threshold function. For example,

$$g(\beta_1, \beta_2, V_1, V_2, V_3) = th(\beta_1, \beta_2, \omega_1 V_1 + \omega_2 V_2 + \omega_3 V_3)$$

with V_1, V_2, V_3 activation levels and $\omega_1, \omega_2, \omega_3$ weights of the connections to the preparation state, and

$th(\beta_1, \beta_2, V) = 1/(1 + e^{-\beta_2(V - \beta_1)})$ a threshold function with threshold β_1 and steepness β_2 .

LP4a From option activation and experience to preparation of a body state (non-competitive case)

If option o with level V_1 occurs
 and feeling the associated body state b has level V_2
 and an experience for o occurs with level V_3
 and the preparation state for b has level V_4
 then a preparation state for body state b will occur with level $V_4 + \gamma(g(\beta_1, \beta_2, V_1, V_2, V_3, \omega_1, \omega_2, \omega_3) - V_4) \Delta t$.
 $rep(o, V_1) \ \& \ feeling(b, V_2) \ \& \ srs(exp, V_3) \ \& \ preparation_state(b, V_4)$
 $\rightarrow preparation_state(b, V_4 + \gamma(g(\beta_1, \beta_2, V_1, V_2, V_3, \omega_1, \omega_2, \omega_3) - V_4) \Delta t)$

For the competitive case also the inhibiting cross connections from one represented option to the body state induced by another represented option are used. In this case a function involving these cross connections can be defined, for example for two considered options

$$h(\beta_1, \beta_2, V_1, V_2, V_3, V_{21}, \omega_1, \omega_2, \omega_3, \omega_{21}) = th(\beta_1, \beta_2, \omega_1 V_1 + \omega_2 V_2 + \omega_3 V_3 - \omega_{21} V_{21})$$

with ω_{21} the weight of the suppressing connection from represented option 2 to the preparation state induced by option 1.

LP4b From option activation and experience to preparation of a body state (competitive case)

If option o1 with level V_1 occurs
 and option o2 with level V_{21} occurs
 and feeling the associated body state b1 has level V_2
 and an experience for o1 occurs with level V_3
 and the preparation state for b1 has level V_4
 then a preparation state for body state b1 will occur with
 level $V_4 + \gamma(h(\beta_1, \beta_2, V_1, V_2, V_3, V_{21}, \omega_1, \omega_2, \omega_3, \omega_{21}) - V_4) \Delta t$.
 $\text{rep}(o1, V_1) \ \& \ \text{rep}(o2, V_{21}) \ \& \ \text{feeling}(b1, V_2) \ \& \ \text{srs}(\text{exp1}, V_3) \ \& \ \text{preparation_state}(b1, V_4)$
 $\rightarrow \text{preparation_state}(b1, V_4 + \gamma(h(\beta_1, \beta_2, V_1, V_2, V_3, V_{21}, \omega_1, \omega_2, \omega_3, \omega_{21}) - V_4) \Delta t)$

Dynamic properties LP5, LP6, and LP7 together with LP2 describe the body loop.

LP5 From preparation to effector state for body modification

If preparation state for body state B occurs with level V ,
 then the effector state for body state B with level V will occur.
 $\text{preparation_state}(B, V) \rightarrow \text{effector_state}(B, V)$

LP6 From effector state to modified body state

If the effector state for body state B with level V occurs,
 then the body state B with level V will occur.
 $\text{effector_state}(B, V) \rightarrow \text{body_state}(B, V)$

LP7 Sensing a body state

If body state B with level V occurs,
 then this body state B with level V will be sensed.
 $\text{body_state}(B, V) \rightarrow \text{sensor_state}(B, V)$

LP8 From sensory representation of body state to feeling

If a sensory representation for body state B with level V occurs,
 then B is felt with level V .
 $\text{srs}(B, V) \rightarrow \text{feeling}(B, V)$

Alternatively, dynamic properties LP5 to LP7 can also be replaced by one dynamic property LP9 describing an as if body loop as follows.

LP9 From preparation to sensed body state

If preparation state for body state B occurs with level V ,
 then the effector state for body state B with level V will occur.
 $\text{preparation_state}(B, V) \rightarrow \text{srs}(B, V)$

For the decision process on which option O_i to choose, represented by action A_i , a winner-takes-it-all model is used based on the feeling levels associated to the options; for an overview, see Figure 2. This has been realised by combining the option representations O_i with their related emotional responses B_i in such a way that for each i the level of the emotional response B_i has a strongly positive effect on preparation of the action A_i related to option O_i itself, but a strongly suppressing effect on the preparations for actions A_j related to the other options O_j for $j \neq i$. As before, this is described by a function

$$h(\beta_1, \beta_2, V_1, \dots, V_m, U_1, \dots, U_m, \omega_{11}, \dots, \omega_{mm})$$

with V_i levels for representations of options O_i and U_i levels of preparation states for body state B_i related to options O_i and ω_{ij} the strength of the connection between preparation states for body state B_i and preparation states for action A_j .

LP10 Decisions based on felt emotions induced by the options

If options O_i with levels V_i occur,
 and preparation states for body state B_i related to options O_i occur with level U_i ,
 and the preparation state for action A_i for option O_i has level W_i
 then the preparation state for action A_i for option O_i will occur
 with level $W_i + \gamma(h(\beta_1, \beta_2, V_1, \dots, V_m, U_1, \dots, U_m, \omega_{11}, \dots, \omega_{mm}) - W_i) \Delta t$.
 $\text{rep}(O_1, V_1) \ \& \ \dots \ \& \ \text{rep}(O_m, V_m) \ \& \ \text{preparation_state}(B_1, U_1) \ \& \ \dots \ \& \ \text{preparation_state}(B_m, U_m) \ \& \ \text{preparation_state}(A_i, W_i)$
 $\rightarrow \text{preparation_state}(A_i, W_i + \gamma(h(\beta_1, \beta_2, V_1, \dots, V_m, U_1, \dots, U_m, \omega_{11}, \dots, \omega_{mm}) - W_i) \Delta t)$

LP11 From preparation to effector state for an action

If preparation state for action A occurs with level V ,
 then the effector state for action A with level V will occur.
 $\text{preparation_state}(A, V) \rightarrow \text{effector_state}(A, V)$

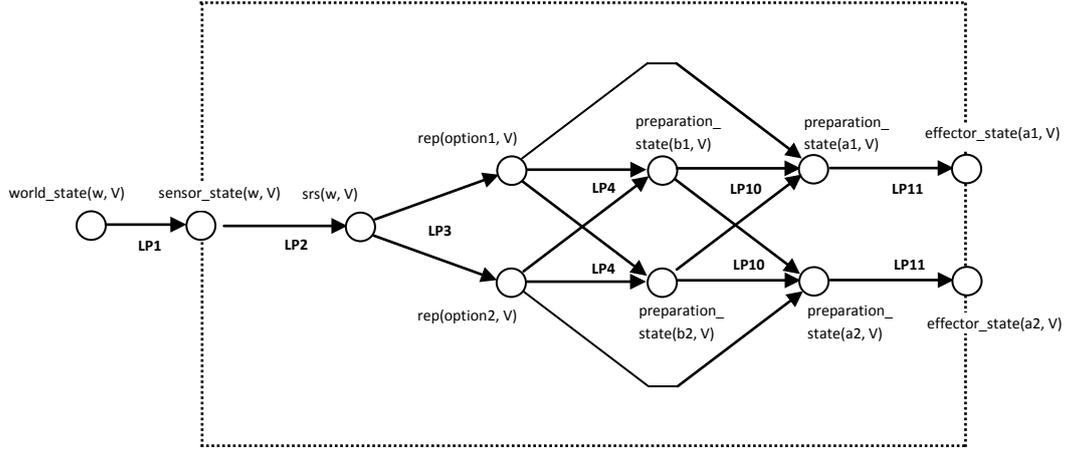


Figure 2: Overview of the model for trust-based decision making

3.2 The Hebbian adaptation process

From a neurological perspective the strength of a connection from an option to an emotional response may depend on how experiences are felt emotionally, as neurons involved in the option, the preparation for the body state, and in the associated feeling will often be activated simultaneously. Therefore such a connection from option to emotional response may be strengthened based on a general Hebbian learning mechanism (Hebb, 1949; Bi and Poo, 2001; Gerstner and Kistler, 2002) that states that connections between neurons that are activated simultaneously are strengthened, similar to what has been proposed for the emergence of mirror neurons; e.g., (Keyser and Perrett, 2004; Keyser and Gazzola, 2009). This principle is applied to the strength ω_i of the connection from option 1 to the emotional response expressed by body state b1. The following learning rule takes into account a maximal connection strength I , a learning rate η , and an extinction rate ζ .

LP12 Hebbian learning rule for the connection from option to preparation

If the connection from option o1 to preparation of b1 has strength ω_i
and the option o1 has strength V_1
and the preparation of b1 has strength V_2
and the learning rate from option o1 to preparation of b1 is η
and the extinction rate from option o1 to preparation of b1 is ζ
then after Δt the connection from option o1 to preparation state b1 will have
strength $\omega_i + (\eta V_1 V_2 (I - \omega_i) - \zeta \omega_i) \Delta t$
has_connection_strength(rep(o1), preparation(b1), ω_i) & rep(o1, V_1) & preparation(b1, V_2) &
has_learning_rate(rep(o1), preparation(b1), η) & has_extinction_rate(rep(o1), preparation(b1), ζ)
→ has_connection_strength(rep(o1), preparation(b1), $\omega_i + (\eta V_1 V_2 (I - \omega_i) - \zeta \omega_i) \Delta t$)

By this rule through their affective aspects, the experiences are accumulated in the connection strength from option o1 to preparation of body state b1, and thus serves as a representation of trust in this option o1. A similar Hebbian learning rule can be found in (Gerstner and Kistler, 2002, p. 406).

4 Example Simulation Results

The model described in Section 3 has been used to generate a number of simulation experiments for non-competitive and competitive cases (see Figure 3 for some example results). To ease the comparison between these cases the same model parameter values were used for these examples (see Table 1). In Figure 3a) example simulation results are shown for the non-competitive case. Here the subject is exposed to an information source that provides experience values 0 respectively 1 alternating periodically in a period of 200 time steps each. In this figure it can be observed that change in experience leads to changes in the connection strengths (representing trust) as well as the action effector states. Furthermore, the decrease in the connection strengths representing trust due to a bad experience (0) takes longer than the increase due to a good experience (1), which can be explained by the higher value of the learning rate than of the extinction rate.

Parameter	Value	Meaning
β_1	0.95	threshold value for preparation state and action effector state
β_2	10, 100	steepness value for preparation state, action effector state
γ	0.90	activation change rate
η	0.80	learning rate of the connection from option representation to preparation
ζ	0.10	extinction rate of the connection from option representation to preparation
Δt	0.90	time step
ω_s, ω_a	0.50	suppressing weight from option representation to preparation state and from preparation state to the action state (competitive case)

Table 1: Parameter values used in the example simulations

In Figures 3b), c) and d), the simulation results are shown for the competitive case with two competitive options having suppression weight 0.5 from option representation to preparation state and from preparation state to the action state. In this case the subject is exposed to two information sources that provides experience values 0 respectively 1 alternating periodically in a period of 200 time steps each, in a reverse cycle with respect to each other (see Figure 3b)). Here it can also be noted that change in experience changes the connections representing trust as well as the action effector states. Moreover, in comparison to the non-competitive case, it can be observed that the learning is slow while decay is fast, which is due to the presence of competition. Finally Figure 3 shows that the connection strengths in the presented model exhibit the two fundamental functional properties of trust discussed in Section 1, namely that trust is based on accumulation of experiences over time (see Figure 3c)) and that trust states are used in decision making by choosing more trusted options above less trusted ones (see Figure 3d)).

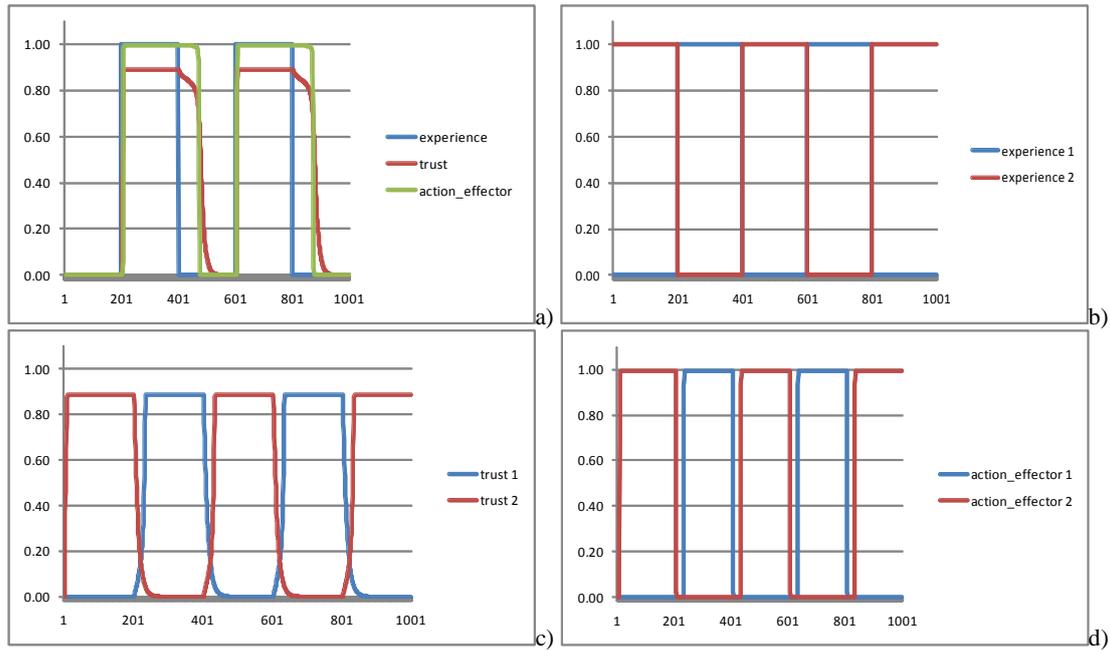


Figure 3: Simulation Results: experience, connection representing trust, and action effector state for a non-competitive case in a) and a competitive case in b), c) and d) respectively

5 Verification of Functional Properties of Trust for the Neurological Model

For the presented neurological model for trust, two main functional properties of trust states have been formulated in the introduction:

- (1) A trust state accumulates experiences over time
- (2) Trust states are used in decision making by choosing more trusted above less trusted options

These properties characterise trust states from a functional, cognitive perspective. Therefore any model or computational or physical realisation claimed to describe trust dynamics has to (at least) satisfy these properties. Such properties can be formalized in a temporal logical language, and can be automatically verified for the traces that have been generated using the proposed model. In this section this verification of properties is discussed. First, the language used to verify these properties is explained in more detail. Thereafter the properties and the results of the verification are discussed.

The verification of properties has been performed using a language called TTL (Temporal Trace Language), that features a dedicated editor and an automated checker; cf. (Bosse, Jonker, Meij, Sharpanskykh, and Treur, 2009). This predicate logical temporal language supports formal specification and analysis of dynamic properties, covering both qualitative and quantitative aspects. TTL is built on atoms referring to *states* of the world, *time points* and *traces*, i.e. trajectories of states over time. In addition, *dynamic properties* are temporal statements that can be formulated with respect to traces based on the state ontology Ont in the following manner. Given a trace γ over state ontology Ont, the state in γ at time point t is denoted by $\text{state}(\gamma, t)$. These states can be related to state properties via the infix predicate \models , where $\text{state}(\gamma, t) \models p$ denotes that state property p holds in trace γ at time t . Based on these statements, dynamic properties can be formulated using quantifiers over time and traces and the usual first-order logical connectives such as $\neg, \wedge, \vee, \Rightarrow, \forall, \exists$. For more details, see (Bosse et al., 2009).

In order to be able to automatically verify the properties upon the simulation traces, they have been formalised. From the computational verification process it was found that indeed they are satisfied by the simulation traces of the model for which they were verified. The first functional property (1), specifying that a trust state accumulates experiences over time, is split up into a number of properties. First, two properties are specified which express trust accumulation for the non-competitive case, whereby the connections for the respective trustees are not influenced by experiences with competitors.

P1.1 Connection strength increases with more positive experience (non-competitive case)

If a sensor state indicates a particular value V_1 of an experience E , and E is an experience for trustee T , and the current strength of the connection for trustee T is V_2 , and V_1 is higher than V_2 , then the connection strength will remain the same or increase.

$$\begin{aligned} & \forall \gamma: \text{TRACE}, t: \text{TIME}, E: \text{EXPERIENCE}, T: \text{TRUSTEE}, V_1, V_2, V_3: \text{VALUE} \\ & [\text{state}(\gamma, t) \models \text{sensor_state}(E, V_1) \ \& \ \text{state}(\gamma, t) \models \text{connection}(T, V_2) \ \& \ \text{state}(\gamma, t) \models \text{matches}(E, T) \ \& \ V_1 > V_2 \\ & \Rightarrow \exists V_3: \text{VALUE} [\text{state}(\gamma, t+1) \models \text{connection}(T, V_3) \ \& \ V_3 \geq V_2]] \end{aligned}$$

P1.2 Connection strength decreases with more negative experience (non-competitive case)

If a sensor state indicates a particular value V_1 of an experience E , and E is an experience for trustee T , and the current strength of the connection for trustee T is V_2 , and V_1 is lower than V_2 , then the connection strength will remain the same or decrease.

$$\begin{aligned} & \forall \gamma: \text{TRACE}, t: \text{TIME}, E: \text{EXPERIENCE}, T: \text{TRUSTEE}, V_1, V_2, V_3: \text{VALUE} \\ & [\text{state}(\gamma, t) \models \text{sensor_state}(E, V_1) \ \& \ \text{state}(\gamma, t) \models \text{connection}(T, V_2) \ \& \ \text{state}(\gamma, t) \models \text{matches}(E, T) \ \& \ V_1 < V_2 \\ & \Rightarrow \exists V_3: \text{VALUE} [\text{state}(\gamma, t+1) \models \text{connection}(T, V_3) \ \& \ V_3 \leq V_2]] \end{aligned}$$

Besides the non-competitive case, also properties have been specified for the competitive case. Hereby, the experiences with other competitive information sources are also taken into account.

P2.1 Connection strength increases with more positive experience (competitive case)

If a sensor state indicates a particular value V_1 of an experience E , and E is an experience for trustee T , and the current strength of the connection for trustee T is V_2 , and V_1 is higher than V_2 , and all other experiences are lower compared to V_1 , then the connection strength will remain the same or increase.

$$\begin{aligned} & \forall \gamma: \text{TRACE}, t: \text{TIME}, E: \text{EXPERIENCE}, T: \text{TRUSTEE}, V_1, V_2, V_3: \text{VALUE} \\ & [\text{state}(\gamma, t) \models \text{sensor_state}(E, V_1) \ \& \ \forall E' \neq E [\exists V': \text{VALUE} \ \text{state}(\gamma, t) \models \text{sensor_state}(E', V') \ \& \ V' < V_1] \ \& \\ & \ \text{state}(\gamma, t) \models \text{connection}(T, V_2) \ \& \ \text{state}(\gamma, t) \models \text{matches}(E, T) \ \& \ V_1 > V_2 \\ & \Rightarrow \exists V_3: \text{VALUE} [\text{state}(\gamma, t+1) \models \text{connection}(T, V_3) \ \& \ V_3 \geq V_2]] \end{aligned}$$

P2.2 Connection strength decreases with more negative experience (competitive case)

If a sensor state indicates a particular value V_1 of an experience E , and E is an experience for trustee T , and the current strength of the connection for trustee T is V_2 , and V_1 is lower than V_2 , and all other experiences are higher compared to V_1 , then the connection strength will remain the same or decrease.

$$\begin{aligned} & \forall \gamma: \text{TRACE}, t: \text{TIME}, E: \text{EXPERIENCE}, T: \text{TRUSTEE}, V_1, V_2, V_3: \text{VALUE} \\ & [\text{state}(\gamma, t) \models \text{sensor_state}(E, V_1) \ \& \ \forall E' \neq E [\exists V': \text{VALUE} \ \text{state}(\gamma, t) \models \text{sensor_state}(E', V') \ \& \ V' > V_1] \ \& \\ & \ \text{state}(\gamma, t) \models \text{connection}(T, V_2) \ \& \ \text{state}(\gamma, t) \models \text{matches}(E, T) \ \& \ V_1 < V_2 \\ & \Rightarrow \exists V_3: \text{VALUE} [\text{state}(\gamma, t+1) \models \text{connection}(T, V_3) \ \& \ V_3 \leq V_2]] \end{aligned}$$

Finally, a property is specified which compares different traces, as shown in P3 below.

P3.1 Higher experiences lead to higher connection strengths (non-competitive case)

If within one trace the experiences for a trustee are always higher compared to the experiences for a trustee in another trace, then in that trace the connection strengths will always be higher.

$$\begin{aligned} & \forall \gamma_1, \gamma_2: \text{TRACE}, E: \text{EXPERIENCE}, T: \text{TRUSTEE} \\ & [\gamma_1 \neq \gamma_2 \ \& \ \text{state}(\gamma_1, 0) \models \text{matches}(E, T) \ \& \ \forall t: \text{TIME} \ [\exists V_1, V_2: \text{VALUE} \ [\text{state}(\gamma_1, t) \models \text{sensor_state}(E, V_1) \ \& \\ & \quad \text{state}(\gamma_2, t) \models \text{sensor_state}(E, V_2) \ \& \ V_1 > V_2]] \\ & \Rightarrow \forall t: \text{TIME} \ [\exists V_1, V_2: \text{VALUE} \ [\text{state}(\gamma_1, t) \models \text{connection}(T, V_1) \ \& \ \text{state}(\gamma_2, t) \models \text{connection}(T, V_2) \ \& \ V_1 > V_2]] \end{aligned}$$

P3.2 Higher experiences lead to higher connection strengths (competitive case)

If within one trace the experiences for a trustee are always higher compared to the experiences for a trustee in another trace, and there are no other experiences with a higher value at that time point, then in that trace the connection strengths will always be higher.

$$\begin{aligned} & \forall \gamma_1, \gamma_2: \text{TRACE}, E: \text{EXPERIENCE}, T: \text{TRUSTEE} \\ & [\gamma_1 \neq \gamma_2 \ \& \ \text{state}(\gamma_1, 0) \models \text{matches}(E, T) \ \& \ \forall t: \text{TIME} \ [\exists V_1, V_2: \text{VALUE} \ [\text{state}(\gamma_1, t) \models \text{sensor_state}(E, V_1) \ \& \\ & \quad \forall E' \neq E \ [\exists V': \text{VALUE} \ \text{state}(\gamma_1, t) \models \text{sensor_state}(E', V') \ \& \ V' \leq V_1] \ \& \ \text{state}(\gamma_2, t) \models \text{sensor_state}(E, V_2) \ \& \\ & \quad \forall E'' \neq E \ [\exists V'': \text{VALUE} \ \text{state}(\gamma_2, t) \models \text{sensor_state}(E'', V'') \ \& \ V'' \leq V_2] \ \& \ V_1 > V_2]] \\ & \Rightarrow \forall t: \text{TIME} \ [\exists V_1, V_2: \text{VALUE} \ [\text{state}(\gamma_1, t) \models \text{connection}(T, V_1) \ \& \ \text{state}(\gamma_2, t) \models \text{connection}(T, V_2) \ \& \ V_1 > V_2]] \end{aligned}$$

The formalization of the second functional property (2), i.e., trust states are used in decision making by choosing more trusted options above less trusted options, is expressed as follows.

P4 The trustee with the strongest connection is selected

If for trustee T the connection strength is the highest, then this trustee will be selected.

$$\begin{aligned} & \forall \gamma: \text{TRACE}, t: \text{TIME}, T: \text{TRUSTEE}, V_1: \text{VALUE} \\ & [\text{state}(\gamma, t) \models \text{connection}(T, V_1) \ \& \ \text{state}(\gamma, t) \models \text{sensor_state}(w, 1) \ \& \\ & \quad \forall T' \ [\exists V': \text{VALUE} \ \text{state}(\gamma, t) \models \text{connection}(T', V') \ \& \ V' \leq V_1] \\ & \Rightarrow \exists V_3: \text{VALUE} \ [\text{state}(\gamma, t+1) \models \text{effector_state}(T, V_3) \ \& \ V_3 > 0.9] \end{aligned}$$

6 Discussion

In this paper a computational model for trust dynamics was introduced incorporating the reciprocal interaction between cognitive and affective aspects based on neurological theories that address the role of emotions and feelings. The introduced model describes more specifically how considered options and experiences generate an emotional response that is felt. For feeling the emotion, based on elements taken from (Damasio, 1999, 2004; Bosse, Jonker and Treur, 2008), a converging recursive body loop is included in the model. An adaptation process based on Hebbian learning (cf. Hebb, 1949; Bi and Poo, 2001; Gerstner and Kistler, 2002), was incorporated, inspired by the Somatic Marker Hypothesis described in (Damasio, 1994, 1996; Bechara and Damasio, 2004), and as also has been proposed for the functioning of mirror neurons; e.g., (Keysers and Perrett, 2004; Keysers and Gazzola, 2009). The model was specified in the hybrid dynamic modelling language LEADSTO, and simulations were performed in its software environment; cf. (Bosse, Jonker, Meij, and Treur, 2007).

It has been shown that within the model the strength of the connection between a considered option and the emotional response induced by it, satisfies two properties that are considered as two central functional properties of a trust state as a cognitive state (e.g., Jonker and Treur, 2003): (1) it accumulates experiences, and (2) it is a crucial factor used in deciding for the option. This provides support for the assumption that the strength of this connection can be interpreted as a representation of the trust level in the considered option.

Models of neural processes can be specified at different levels of abstraction. The model presented here can be viewed as a model at a higher abstraction level, compared to more detailed models that take into account more fine-grained descriptions of neurons and their biochemical and/or spiking patterns. States in the presented model can be viewed as abstract descriptions of states of neurons or as representing states of groups of neurons. An advantage of a more abstract representation is that such a model is less complex and therefore may be less difficult to handle, while it still shows the essential dynamics. An interesting issue for further research is how such a more abstract model can be related to more detailed models, and in how far patterns observed in more specific models also are represented in such a more abstract model.

References

1. Bechara, A., and Damasio, A. (2004). The Somatic Marker Hypothesis: a neural theory of economic decision. *Games and Economic Behavior*, vol. 52, pp. 336-372.
2. Bi, G.Q., and Poo, M.M. (2001) Synaptic Modifications by Correlated Activity: Hebb's Postulate Revisited. *Ann Rev Neurosci*, vol. 24, pp. 139-166.

3. Bosse, T., Jonker, C.M., and Treur, J. (2008). Formalisation of Damasio's Theory of Emotion, Feeling and Core Consciousness. *Consciousness and Cognition Journal*, vol. 17, 2008, pp. 94–113.
4. Bosse, T., Jonker, C.M., Los, S.A., Torre, L. van der, and Treur, J. (2007). Formal Analysis of Trace Conditioning. *Cognitive Systems Research Journal*, vol. 8, 2007, pp. 36-47.
5. Bosse, T., Jonker, C.M., Meij, L. van der, Sharpanskykh, A., and Treur, J., (2009). Specification and Verification of Dynamics in Agent Models. *International Journal of Cooperative Information Systems*, vol. 18, 2009, pp. 167 - 193.
6. Bosse, T., Jonker, C.M., Meij, L. van der, and Treur, J. (2007). A Language and Environment for Analysis of Dynamics by Simulation. *International Journal of Artificial Intelligence Tools*, vol. 16, 2007, pp. 435-464.
7. Bosse, T., Jonker, C.M., and Treur, J. (2007). Simulation and Analysis of Adaptive Agents: an Integrative Modelling Approach. *Advances in Complex Systems Journal*, vol. 10, 2007, pp. 335 - 357.
8. Damasio, A. (1994). *Descartes' Error: Emotion, Reason and the Human Brain*, Papermac, London.
9. Damasio, A. (1996). The Somatic Marker Hypothesis and the Possible Functions of the Prefrontal Cortex. *Philosophical Transactions of the Royal Society: Biological Sciences*, vol. 351, pp. 1413-1420
10. Damasio, A. (1999). *The Feeling of What Happens. Body and Emotion in the Making of Consciousness*. New York: Harcourt Brace, 1999.
11. Damasio, A. (2004). *Looking for Spinoza*. Vintage books, London, 2004.
12. Eich, E., Kihlstrom, J.F., Bower, G.H., Forgas, J.P., & Niedenthal, P.M. (2000). *Cognition and Emotion*. New York: Oxford University Press.
13. Falcone, R., and Castelfranchi, C., Trust dynamics: How Trust is Influenced by Direct Experiences and by Trust Itself, In: Proc. of AAMAS 2004, 2004, pp. 740–747.
14. Forgas, J.P., Laham, S.M., and Vargas, P.T. (2005). Mood effects on eyewitness memory: Affective influences on susceptibility to misinformation. *Journal of Experimental Social Psychology*, vol. 41, 2005, pp. 574–588.
15. Forgas, J.P., Goldenberg, L., and Unkelbach, C. (2009). Can bad weather improve your memory? An unobtrusive field study of natural mood effects on real-life memory. *Journal of Experimental Social Psychology*, vol. 45, 2009, pp. 254–257.
16. Gerstner, W., and Kistler, W.M. (2002). Mathematical formulations of Hebbian learning. *Biol. Cybern.*, vol. 87, 2002, pp. 404–415
17. Hebb, D. (1949). *The Organisation of Behavior*. Wiley, New York.
18. Hoogendoorn, M., Jaffry, S.W., and Treur, J., Modeling Dynamics of Relative Trust of Competitive Information Agents, In: Klusch, M., Pechoucek, M., Polleres, A. (eds.), Proceedings of the 12th International Workshop on Cooperative Information Agents, CIA'08. LNAI, vol. 5180. Springer, 2008, pp. 55-70.
19. Jonker, C.M., Snoep, J.L., Treur, J., Westerhoff, H.V., and Wijngaards, W.C.A. (2008). BDI-Modelling of Complex Intracellular Dynamics. *Journal of Theoretical Biology*, vol. 251, 2008, pp. 1–23.
20. Jonker, C.M., and Treur, J., Formal Analysis of Models for the Dynamics of Trust based on Experiences, In: F.J. Garijo, M. Boman (eds.), Proc. of MAAMAW'99, LNAI, vol. 1647, Springer Verlag, Berlin, 1999, pp. 221-232.
21. Jonker, C.M., and Treur, J., (2003). A Temporal-Interactivist Perspective on the Dynamics of Mental States. *Cognitive Systems Research Journal*, vol. 4, 2003, pp. 137-155.
22. Keysers, C., and Gazzola, V., (2009). Unifying Social Cognition. In: Pineda, J.A. (ed.), *Mirror Neuron Systems: the Role of Mirroring Processes in Social Cognition*. Humana Press/Springer Science, 2009, pp. 3-28.
23. Keysers, C., and Perrett, D.I. (2004). Demystifying social cognition: a Hebbian perspective. *Trends in Cognitive Sciences*, vol. 8, 2004, pp. 501-507.
24. Niedenthal, P.M. (2007). Embodying Emotion. *Science*, vol. 316, (2007), pp. 1002-1005.
25. Schooler, J.W., and Eich, E. (2000). Memory for Emotional Events. In: E. Tulving, F.I.M. Craik (eds.), *The Oxford Handbook of Memory*. Oxford University Press, 2000, pp. 379-394.
26. Winkielman, P., Niedenthal, P.M., and Oberman, L.M. (2009). Embodied Perspective on Emotion-Cognition Interactions. In: Pineda, J.A. (ed.), *Mirror Neuron Systems: the Role of Mirroring Processes in Social Cognition*. Humana Press/Springer Science, 2009, pp. 235-257.