Comparing a Cognitive and a Neural Model for Relative Trust Dynamics

S. Waqar Jaffry, Jan Treur

VU University Amsterdam, Department of Artificial Intelligence
De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands
Email: {swjaffry, treur }@few.vu.nl
URL: http://www.few.vu.nl/~{swjaffry, treur }

Abstract. Trust dynamics can be modelled in relation to experiences. Both cognitive and neural models for trust dynamics in relation to experiences are available, but were not yet related or compared in more detail. This paper presents a comparison between a cognitive and a neural model. As each of the models has its own specific set of parameters, with values that depend on the type of person modelled, such a comparison is nontrivial. In this paper a comparison approach is presented that is based on mutual mirroring of the models in each other. More specifically, for given parameter values set for one model, by automated parameter estimation processes the most optimal values for the parameter values of the other model are determined to show the same behavior. Roughly spoken the results are that the models can mirror each other up to an accuracy of around 90%.

1 Introduction

A variety of computational models have been proposed for the dynamics of human trust in relation to experiences; see e.g., (Jonker and Treur, 1999, 2003; Falcone and Castelfranchi, 2004; Hoogendoorn, Jaffry, and Treur, 2008). Usually such models consider experiences and trust as cognitive concepts, and depend on values for a set of parameters for specific (cognitive) characteristics of a person, such as trust flexibility vs rigidity. Recently also neural models for trust dynamics have been introduced. An example of such a neural model, in which in addition a role for emotional responses is incorporated, is described in (Hoogendoorn, Jaffry, and Treur, 2009). Also the latter model includes a specific set of parameters for (neurological) characteristics of the person modelled. As the set of parameters of this neural model has no clear connection to the parameters in cognitive models such as in (Hoogendoorn, Jaffry, and Treur, 2008), and the behavior of such models strongly depends on the values for such parameters, a direct comparison is impossible.

Therefore in this paper, a more indirect way to compare the models is used, by mutual mirroring them in each other. This mirroring approach uses any set of values that is assigned to the parameters for one of the models to obtain a number of simulation traces. These simulation traces are approximated by the second model, based on automated parameter estimation. The error for this approximation is considered as a comparison measure. In this paper this mirroring approach is applied to the two models for the dynamics of relative trust described in (Hoogendoorn, Jaffry, and Treur, 2008) and (Hoogendoorn, Jaffry, and Treur, 2009). It is applied in two directions, and also back and forth sequentially by using the estimated parameter values for the second model to estimate new parameter values for the first model.

In the paper, first in Section 2 the cognitive model is briefly summarised, and in Section 3 the neural model. In Section 4 the mirroring approach is discussed and the automated parameter estimation method. Section 5 reports the outcome of some of the experiments performed. Finally, Section 6 is a discussion.

2 A Cognitive Model for the Dynamics of Relative Trust

The cognitive model taken from (Hoogendoorn, Jaffry, and Treur, 2008) is composed from two models: one for the positive trust, accumulating positive experiences, and one for negative trust, accumulating negative experiences. First the positive trust is addressed. The human's relative positive trust on an option *i* at time point *t* is based on a combination of two parts: the *autonomous* part, and the *context-dependent* part. For the latter part an important indicator is $\tau_i^+(t)$: the ratio of the human's trust of option *i* to the average human's trust on all options at time point *t*. Similarly the human's relative negative trust of option *i* at time point *t* ($\tau^-_i(t)$) is the ratio between human's negative trust of the option *i* and the average human's negative trust of the options at time point *t*. These are calculated as follows:

$$\tau_i^+(t) = \frac{T_i^+(t)}{\sum_{j=1}^{n} T_j^+(t) / n} \qquad\qquad \tau_i^-(t) = \frac{T_i^-(t)}{\sum_{j=1}^{n} T_j^-(t) / n}$$

Here the denominators express the average positive and negative trust over all options at time point *t*. The context-dependent part is designed in such a way that when the positive trust is above the average, then upon each positive experience it gets an extra increase, and when it is below average it gets a decrease. This principle is a variant of a 'winner takes it all' principle, which for example is sometimes modelled by mutually inhibiting neurons. This principle has been modelled by basing the change of trust upon a positive experience on $\tau_i^+(t) - 1$, which is positive when the positive trust is above average and negative when it is below average. To normalise, this is multiplied by a factor $T_i^+(t) * \left(1 - T_i^+(t)\right)$. For the autonomous part the change upon a positive experience is modelled by $1 - T_i^+(t)$. As η indicates in how far the human is autonomous or context-dependent in trust attribution, a weighted sum is taken with weights η and $1\text{-}\eta$ respectively. Therefore, using the parameters defined in above $T_i^+$is modelled by the following differential equation:

$$\frac{dT_i^+(t)}{dt} = \beta * [(\eta * \left(1 - T_i^+(t)\right) + (1-\eta) * (\tau_i^+(t) - 1) * T_i^+(t) * \left(1 - T_i^+(t)\right)] * E_i(t) * \left(1 + E_i(t)\right)) / 2$$

$$- \gamma * T_i^+(t) * \left(1 + E_i(t)\right) * \left(1 - E_i(t)\right)$$

Similarly, for negative trust:

$$\frac{dT_i^-(t)}{dt} = \beta * [\eta * \left(1 - T_i^-(t)\right) + (1-\eta) * (\tau_i^-(t) - 1) * T_i^-(t) * \left(1 - T_i^-(t)\right)] * E_i(t) * \left(1 - E_i(t)\right) / 2$$

$$- \gamma * T_i^-(t) * \left(1 + E_i(t)\right) * \left(1 - E_i(t)\right)$$

The trust $T_i(t)$ of option *i* at time point *t* is a number between [-1, 1] where -1 and 1 represent minimum and maximum values of the trust respectively. It is the difference of the human's positive and negative trust of option *i* at time point *t*: $T_i(t) = T_i^+(t) - T_i^-(t)$. For more details, see (Hoogendoorn, Jaffry and Treur, 2008).

3 A Neural Model for the Dynamics of Relative Trust and Emotion

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). Emotion generation via a body loop roughly proceeds according to the following causal chain:

cognitive state $\rightarrow$ preparation for the induced bodily response $\rightarrow$ induced bodily response $\rightarrow$
sensing the bodily response $\rightarrow$ sensory representation of the bodily response $\rightarrow$ induced feeling

As a variation, an 'as if body loop' uses a direct causal relation preparation for the induced bodily response $\rightarrow$ 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 $\rightarrow$ preparation for the bodily response as an additional causal relation. Such recursiveness is also assumed by Damasio (2004, pp. 91-92), as he notices that what is felt by sensing is actually a body state which is an internal object, under control of the person. 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. 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 by adaptive, related to experiences (Damasio, 1994, p. 179). In the model used below, this adaptive aspect is 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.

The neural 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. They are explained below.

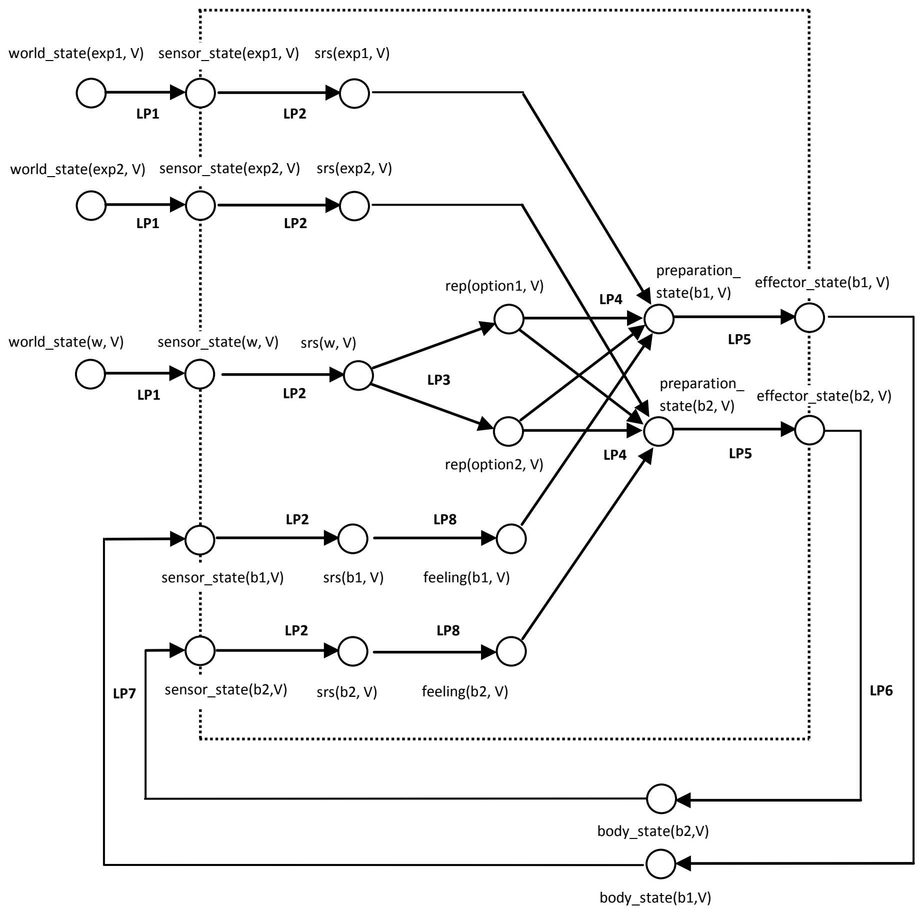


Figure 1: Overview of the neurological model for dynamics of trust and emotion

Activation level for preparation of body state: noncompetitive case

The emotional response to the person's mental state in the form of the preparation for a specific bodily reaction (see label LP4 in Figure 1) is modelled in the noncompetitive case as follows. 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

$$g(\sigma, \tau, V_1, V_2, V_3, \omega_1, \omega_2, \omega_3)$$

including a threshold function. For example,

$$g(\sigma, \tau, V_1, V_2, V_3) = th(\sigma, \tau, V_1 + \omega_2 V_2 + \omega_3 V_3)$$

with V_1, V_2, V_3 activation levels and ω_1, ω_2, ω_3 weights of the connections to the preparation state, and $th(\sigma, \tau, V) = 1/(1+e^{-\sigma(V-\tau)})$ a threshold function with threshold τ and steepness σ. Then the activation level V_4 of the preparation for an option is modelled by

$$dV_4/\mathrm{d}t = \gamma(g(\sigma, \tau, V_1, V_2, V_3, \omega_1, \omega_2, \omega_3) - V_4)$$

Activation level for preparation of body state: competitive case

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

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

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

$$dV_4/\mathrm{dt} = \gamma (h(\sigma, \tau, V_1, V_2, V_3, V_{21}, \omega_1, \omega_2, \omega_3, \omega_{21}) - V_4)$$

with V_4 the activation level of preparation for option 1.

Activation level for preparation of action choice

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 label LP10 in 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 similar function

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

as before, 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.

$$dW_i /dt = \gamma (h(\sigma, \tau, V_1, \dots, V_m, U_1, \dots, U_m, \omega_{11}, .. \omega_{mm}) - W_i)$$

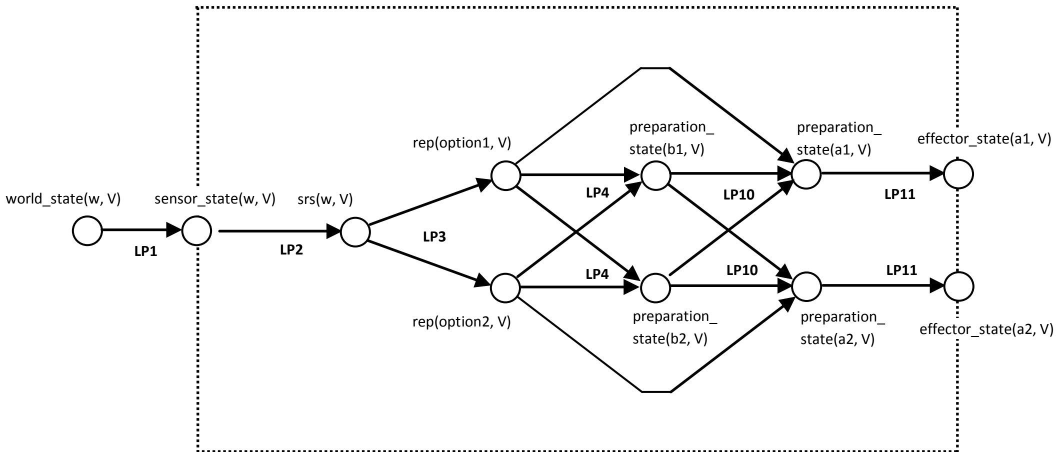


Figure 2: Overview of the neurological model for trust-based decision making

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., (Keysers and Perrett, 2004; Keysers and Gazzola, 2009). This principle is applied to the strength ω_1 of the connection from an option to the emotional response expressed by the related body state. The following Hebbian learning rule takes into account a maximal connection strength 1, a learning rate η, and an extinction rate ζ.

$d\omega_1 /dt = \eta V_1 V_2 (1 - \omega_1) - \zeta\omega_1$

Here V_1 is the activation level of the option o1 and V_2 the activation level of preparation for body state b1. A similar Hebbian learning rule can be found in (Gerstner and Kistler, 2002, p. 406). 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.

4 The Mirroring Approach to Compare the Parameterised Models

The mirroring approach used to compare the two parameterised models for trust dynamics works as follows:

- Initially, for one of the models any set of values is assigned to the parameters of the model
- Next, a number of scenarios are simulated based on this first model.
- The resulting simulation traces for the first model are approximated by the second model, based on automated parameter estimation.
- The error for the most optimal values for the parameters of the second model is considered as a comparison measure.

Parameter estimation can be performed according to different methods, for example, exhaustive search, bisection or simulated annealing (cf. Hoogendoorn, Jaffry and Treur, 2009a). As the models considered here have only a small number of parameters exhaustive search is an adequate option. Using this method the entire attribute search space is explored to find the vector of parameter settings with maximum accuracy. This method guarantees the optimal solution, described as follows:

```
for each observed behavior B
for each vector of parameter value settings P
  calculate the accuracy of P
end for
output the vector of parameter settings with maximal accuracy
end for
```

In the above algorithm, calculation of the accuracy of a vector of parameter setting P entails that agent predicts the information source to be requested and observes the actual human request. It then uses the equation for calculating the accuracy described before. Here if p parameters are to be estimated with precision q (i.e., grain size 10^{-q}), the number of options is n , and m the number of observed outcomes (i.e., time points), then the worst case complexity of the method can be expressed as $O\ ((10)^{pq}\ nm^2)$, which is exponential in number of parameters and precision. In particular, when $p=3$ (i.e., the parameters β, γ, and η), $q=2$ (i.e., grain size 0.01), $n=3$ and $m=100$, then the complexity will result in 3×10^{10} steps.

5. Comparison Results

A number of experiments were performed using the mutual mirroring approach described in Section 4 to compare the two parameterised models for trust dynamics. Experiments were set up according to two cases:

1. Two competitive options provide experiences *deterministically*, with a constant positive, respectively negative experience, alternating periodically in a period of 50 time steps each (see Figure 3).
2. Two options provide experiences with a certain *probability* of positivity, again in an alternating period of 50 time steps each.

The first case of experiments was designed to compare the behavior of the models for different parameters under the same deterministic experiences while the second case is used to compare the behavior of the models for the (more realistic) case of probabilistic experience sequences. the general configurations of the experiment that are kept constant for all experiments are shown in Table 1.

Three experiments were performed for each case: after some parameter values were assigned to the cognitive model, its behavior was approximated by the neural model, using the mirroring approach based on the automatic parameter estimation technique described in Section 4. The best approximating realization of the neural model was used again to be approximated by the cognitive model using the same mirroring approach. This second approximation was performed to minimize uni-directionality of the mirroring approach that might bias the results largely if performed from only one model to another and not the other way around.

Parameter	**Neural Model**	**Cognitive Model**
Number of competitive options	2	2
Time step (difference equations)	0.1	0.1
Number of time steps	500	500
Initial trust values of option 1 and option 2	0.5, 0.5	0, 0
Strength of connection from option to emotional response (ω_1)	0.5	not applicable
Strength of connection between preparation state of body and preparation state of action (ω_{ij})	0.5	not applicable
Strength of connection between feeling and preparation of body state	0.25	not applicable
Value of the world state	1	not applicable
Grain size in parameter estimation	0.05	0.01

Table 1: General Experimental Configuration

An instance of a parameterized model can uniquely be represented by a tuple containing the values of its parameters. Here the cognitive and neural models described in Section 2 and 3 are represented by value tuples for (γ, β, η) and (σ, τ, γ, η, ζ) respectively. For the sake of simplicity, a few parameters of the neural model, namely ω_1, ω_{12} and ω_{21}, were considered fixed with value *0.5*, and were not included in model representation tuple. Furthermore, the initial trust values of both models are assumed neutral (*0.0* and *0.5* for cognitive and neural model resp.), see Table 1.

Case 1

In this case the behavior of the models was compared experiences are provided deterministically with positive respectively negative experience, alternating periodically in a period of 50 time steps each (see Figure 3).

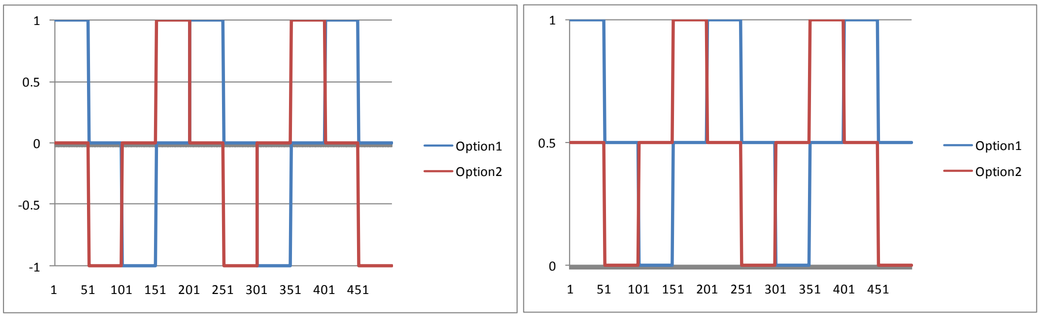


Figure 3: a) Experience sequence for cognitive model, b) Experience sequence for neural model

Here three different experiments were performed, where the parameters of cognitive model are assigned with some initial values and then its behavior is approximated by the neural model. The best approximation of the neural model against the initially set cognitive model was reused to find the best matching cognitive model. Results of the approximated models and errors are shown in Table 2 while the graphs of the trust dynamics are presented in Figure 4. Note that for the sake of ease of comparison and calculation of standard error the trust values of cognitive model are projected from the interval [*-1*, *1*] to [*0*, *1*] (see Figure 4). In Table 2, the comparison error ε is the average of the root mean squared error of all options, as defined by the following formula,

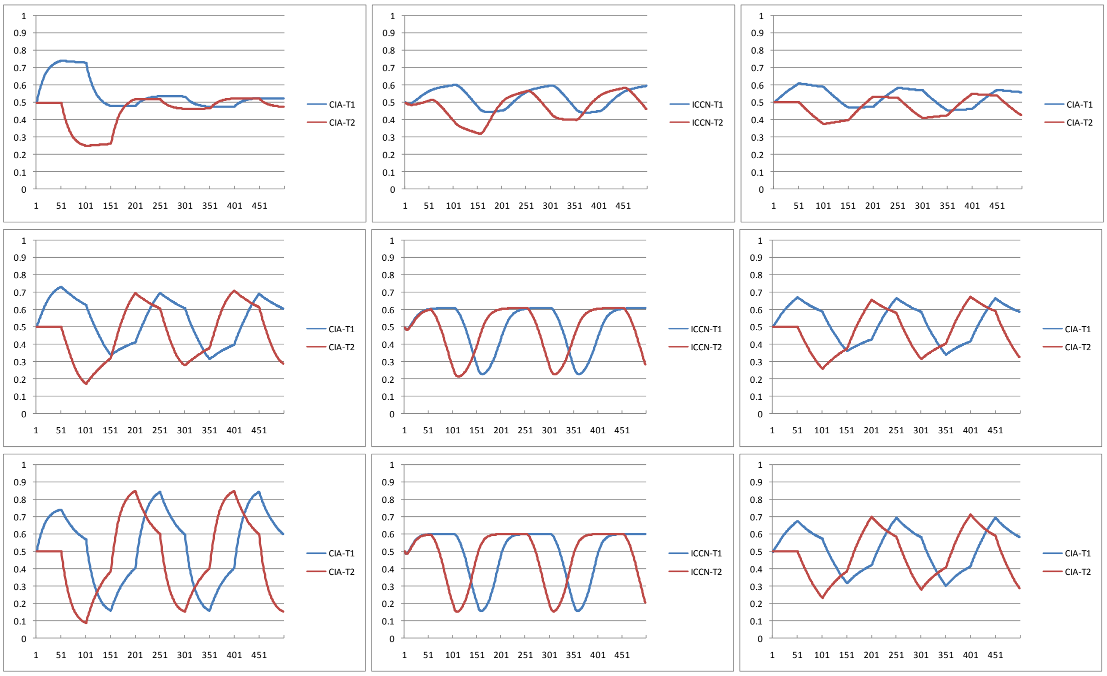
$$\varepsilon = \frac{1}{n} * \sum_{i=1}^{n} \sqrt{\sum_{j=1}^{m} (T(j)_{1i} - T(j)_{2i})^2}$$

In the above formulation, n is the number of options, m is the number of time steps while $T(j)_{1i}$ and $T(j)_{2i}$ represent trust value of option i at time point j for each model, respectively.

Experiment	Initial Model	Approximating Model using the mirroring approach	Comparison Error (*ε*)
1	Cognitive Model (*0.99, 0.75, 0.75*)	Neural Model (*0.55, 10, 0.15, 0.90, 0.50*)	*0.074050*
	Neural Model (*0.55, 10,0 .15, 0.90, 0.50*)	Cognitive Model (*0.96, 0.20, 0.53*)	*0.034140*
2	Cognitive Model (*0.88, 0.99, 0.33*)	Neural Model (*0.35, 10, 0.60, 0.95, 0.60*)	*0.071900*
	Neural Model (*0.35, 10, 0.60, 0.95, 0.60*)	Cognitive Model (*0.87, 0.36, 0.53*)	*0.059928*
3	Cognitive Model (*0.75, 0.75, 0.75*)	Neural Model (*0.30, 10, 0.95, 0.90, 0.60*)	*0.138985*
	Neural Model (*0.55, 10,0 .15, 0.90, 0.50*)	Cognitive Model (*0.83, 0.37, 0.55*)	*0.075991*

Table 2: Results of Case 1

In Table 2 for experiment 1 initially the cognitive model was set with parameters (0.99, 0.75, 0.75) which was then approximated by the neural model. The best approximation of the neural model was found to be (0.55, 10, 0.15, 0.90, 0.50) with an approximate mean of root mean squared error of all options ε value 0.074050. Then this setting of neural model was used to approximate cognitive model producing best approximate with parameter values (0.96, 0.20, 0.53) and ε 0.034140. Similarly the results of other two experiments can be read in Table 2.



a)
b)
c)

Figure 4: Dynamics of the Trust in Case 1 a) Experiment 1, b) Experiment 2, c) Experiment 3

Figure 4 represents the dynamics of the trust in the two options over time for the deterministic case. The horizontal axis represent time step while vertical axis represent the value of trust. The graphs for each experiment are represented as set of three figures, where the first figure shows the dynamics of the trust of both options by the cognitive model with an initial setting as described in the second column of the first row of each experiment of Table 2. The second figure shows the traces of the dynamics of trust by the model as described in the third column of the first row of each experiment of Table 2. Finally the third figure shows the approximation of the neural model by the cognitive model, where the neural model is described in the second column of the second row of each experiment of Table 2 (which is similar to third column of the first row of each experiment), and the approximating cognitive model is presented in the third column of the second row of each experiment.

From Table 2 and Figure 4 it can be observed that the mirroring approach based on automatic parameter estimation when used in bidirectional way gives a better realization of both models in each other, resulting in a smaller comparison error and better curve fit.

Case 2

In the second case the behavior of the models was compared when experiences are provided with a certain probability of positivity, again in an alternating period of 50 time steps each. Also here three different experiments were performed, where the parameters of the cognitive model were assigned with some initial values and then its behavior was approximated by the neural model. The best approximation of the neural model against initially set cognitive model was reused to find the best matching cognitive model. In experiment 1, 2 and 3 the option 1 and option 2 give positive experiences with (100, 0), (75, 25) and (50, 50) percent of probability, respectively. Results of approximated models and errors for this case are shown in Table 3 while the graphs of trust dynamics are presented in Figure 5. Note that for the sake of ease of comparison and calculation of the standard error, again the trust values of the cognitive model are projected from the interval [*-1, 1*] to [*0, 1*] (see Figure 4).

Experiment	**Initial Model**	**Approximating Model using the mirroring approach**	**Comparison Error (ε)**
1	Cognitive Model (0.99, 0.75, 0.75)	Neural Model (0.85, 10, 0.95, 0.20, 0.05)	0.061168
	Neural Model (0.85, 10, 0.95, 0.20, 0.05)	Cognitive Model (0.97, 0.99, 0.18)	0.045562
2	Cognitive Model (0.99, 0.75, 0.75)	Neural Model (0.40, 20, 0.90, 0.20, 0.15)	0.044144
	Neural Model (0.40, 20, 0.90, 0.20, 0.15)	Cognitive Model (0.83, 0.05, 0.99)	0.039939
3	Cognitive Model (0.99, 0.75, 0.75)	Neural Model (0.10, 20, 0.45, 0.10, 0.10)	0.011799
	Neural Model (0.10, 20, 0.45, 0.10, 0.10)	Cognitive Model (0.99, 0.50, 0.99)	0.011420

Table 3: Results of Case 2

In Table 3 for experiment 1 initially the cognitive model was set with parameters (0.99, 0.75, 0.75) which was then approximated by the neural model. The best approximation of the neural model was found to be (0.85, 10, 0.95, 020, 0.05) with an approximate mean of root mean squared error of all options ε of value 0.061168. Then this setting of neural model was used to approximate cognitive model producing best approximate with parameter values (0.97, 0.99, 0.18) and ε 0.034140. Similarly the results of other two experiments could also be read in Table 3.

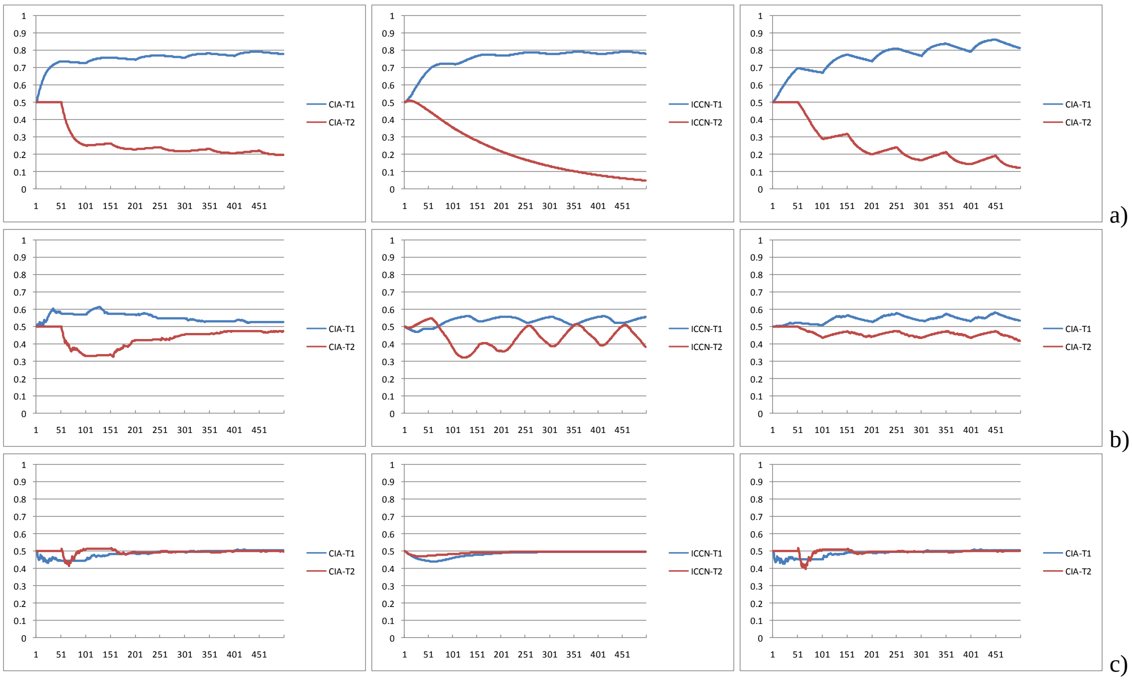


Figure 5: Dynamics of the Trust in Case 2, a) Experiment 1, b) Experiment 2, c) Experiment 3

Figure 5 represents the dynamics of the trust in the two options over time for the probabilistic case. the horizontal axis represents time while the vertical axis represents the values of trust. Here also the graphs of each experiment are represented as set of three figures, where the first figure shows the dynamics of the trust in both options by the cognitive model with an initial setting as described in the second column of the first row of each experiment of Table 3. The second figure shows the traces of the dynamics of trust by the neural model as described in the third column of the first row of each experiment of Table 3. Finally, the third figure is the approximation of the neural model by the cognitive model, where the neural model is described in the second column of the second row of each experiment of Table

3 (which is similar to third column of the first row of each experiment), and the approximating model is presented in the third column of the second row of each experiment.

As already noticed in case 1, also here it can be observed that the mirroring approach based on automatic parameter estimation when used in bidirectional way gives a better realization of both models in each other, resulting smaller comparison error and a better curve fit. Furthermore, it can also be noted that as the uncertainty in the options behavior increases, both models show more similar trust dynamics producing lower error value in comparison.

6. Discussion

In this paper two parameterised computational models for trust dynamics were compared: a cognitive model and a neural model. As the parameter sets for both models are different, the comparison involved mutual estimation of parameter values by which the models were mirrored into each other in the following manner. Initially, for one of the models any set of values was assigned to the parameters of the model, after which a number of scenarios were simulated based on this first model. Next, the resulting simulation traces for this first model were approximated by the second model, based on automated parameter estimation. The error for the most optimal values for the parameters of the second model was considered as a comparison measure. It turned out that appoximations could be obtained with error margins of about 10%. Furthermore the results for the (more realistic) case of probabilistic experience sequences have shown much better approximation than for the deterministic case. This can be considered a positive result, as the two models have been designed in an independent manner, using totally different techniques. In particular, it shows that the cognitive model, which was designed first, without taking into account neurological knowledge, can still be grounded in a neurological context, which is a nontrivial result.

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. (2008). 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. Hoogendoorn, M., Jaffry, S.W., and Treur, J., (2009a). An Adaptive Agent Model Estimating Human Trust in Information Sources. In: *Proceedings of the 9th IEEE/WIC/ACM International Conference on Intelligent Agent Technology, IAT'09*. IEEE Computer Society Press, 2009, to appear.
20. Hoogendoorn, M., Jaffry, S.W., and Treur, J. (2009b). Modelling Trust Dynamics from a Neurological Perspective. To appear.
21. 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.
22. 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.
23. 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.
24. 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.
25. Keysers, C., and Perrett, D.I. (2004). Demystifying social cognition: a Hebbian perspective. *Trends in Cognitive Sciences*, vol. 8, 2004, pp. 501-507.
26. Niedenthal, P.M. (2007). Embodying Emotion. *Science,* vol. 316, (2007), pp. 1002-1005.
27. 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.
28. 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.