Adaptive Agent Models Using Temporal Discounting, Memory Traces and Hebbian Learning with Inhibition, and their Rationality

Jan Treur[a,1] and Muhammad Umair[a,b]
[a] VU University Amsterdam, Agent Systems Research Group
DeBoelelaan 1081, 1081 HV Amsterdam, The Netherlands
[b] COMSATS Institute of Information Technology, Lahore, Pakistan
Email: [a]{j.treur,m.umair}@vu.nl; [b]mumair@ciitlahore.edu.pk
URL: http://www.few.vu.nl/~{treur,mumair}

Abstract. In this paper three adaptive agent models incorporating triggered emotional responses are explored and evaluated on their rationality. One of the models is based on temporal discounting second on memory traces and the third one on hebbian learning with mutual inhibition. The models are assessed using a measure reflecting the environment's behaviour and expressing the extent of rationality. Simulation results and the extents of rationality of the different models over time are presented and analysed.

Keywords. adaptive agent model, memory traces, temporal discounting, mutual inhibition, Hebbian learning, rationality

1. Introduction

Adaptive agents develop decision making capabilities over time based on experiences with their environment. In order to do so, such agents exploit certain learning mechanisms, for example, involving emotional responses triggered by situations. By the learning processes the decision making is adapted to the characteristics of the environment, so that the decisions made are in some way rational, given the enviroment as reflected in the agent's experiences. In this paper the focus is on three of such learning mechanisms: temporal discounting, memory traces and Hebbian learning with inhibition. To assess their extent of rationality, a rationality measure is used reflecting the environment's characteristics.

The adaptation model based on temporal discounting adapts connections based on the frequency of occurrence of certain situations, thereby valuing occurences further back in time lower. A second alternative considered is a case-based memory modelling approach based on *memory traces* e.g. [20, 21]. The adaptation model based on Hebbian learning (cf. [10, 12]) with inhibition involves predictive as-if body loops through feeling states in order to make decisions (e.g., [3, 6, 8]). For the Hebbian learning different variations are considered, by applying it to different types of connections in the decision model.

In this paper, in Section 2 the basic agent model is introduced. Section 3 and Section 4 presents the adaptation model based on temporal discounting and memory traces respectively. In Section 5 the adaptation model based on Hebbian learning with mutual inhibition is introduced. Simulation results are presented in a separate appendix. In Section 6 two measures for rationality used are presented, and the different adaptation models are evaluated based on these measures. Finally, Section 7 is a discussion. Due to space limitation only the graphs for rationality over time for different scenario are presented here. The overview of other simulation results are given in Table 1 to Table 5 in the appendix.

2. The Basic Agent Model Used

This section describes the basic agent model in which other three adaptation models discussed in Section 3, Sections 4 and 5 are incorporated. It is an extension of the work presented in [22]. In this agent model emotional responses triggered by the environment play a role; see Figure 1 for an overview. More specifically, in this

[1] Corresponding Author: Jan Treur. E-mail: j.treur@vu.nl.

paper it is assumed that responses in relation to a sensory representation state roughly proceed according to the following causal chain for a *body loop* (based on elements from [4, 7, 8]):

sensory representation → preparation for bodily response → body state modification → sensing body state → sensory representation of body state → induced feeling

In addition, an *as-if body loop* uses a direct causal relation

preparation for bodily response → sensory representation of body state

as a shortcut in the causal chain; cf. [7]. This can be considered a prediction of the action effect by internal simulation (e.g., [13]). The resulting induced feeling provides an emotion-related valuation of this prediction (cf. [1, 2, 14, 15, 17]). If the level of the feeling (which is assumed positive here) is high, a positive valuation is obtained. The body loop (or as-if body loop) is extended to a recursive (as-if) body loop by assuming that the preparation of the bodily response is also affected by the level of the induced feeling:

induced feeling → preparation for the bodily response

Such recursion is suggested in [8], pp. 91-92. In this way the emotion-related valuation of the prediction affects the preparation. Through this recursive loop a high valuation will strengthen activation of the preparation. To adequately formalise such a theory the hybrid dynamic modelling language LEADSTO has been developed and used in different context; cf. [4,5]. An overview of the basic model for the generation of emotional responses and feelings is depicted in Figure. 1. This picture also shows representations from the detailed specifications explained below. However, note that the precise numerical relations are not expressed in this picture, but in the detailed specifications below, through local properties LP0 to LP6.

Note that the effector state for b_i (bodily responses) combined with the (stochastic) effectiveness of executing b_i in the world (indicated by *effectiveness rate* λ_i between *0* and *1*) activates the sensor state for b_i via body loop as described above. By a recursive as-if body loop each of the preparations for b_i generates a level of feeling for b_i which is considered a valuation of the prediction of the action effect by internal simulation. This in turn affects the level of the related action preparation for b_i. Dynamic interaction within these loops results in equilibrium for the strength of the preparation and of the feeling, and depending on these values, the action is actually activated with a certain intensity. The specific strengths of the connections from the sensory representation to the preparations, and within the recursive as-if body loops can be innate, or are acquired during lifetime.

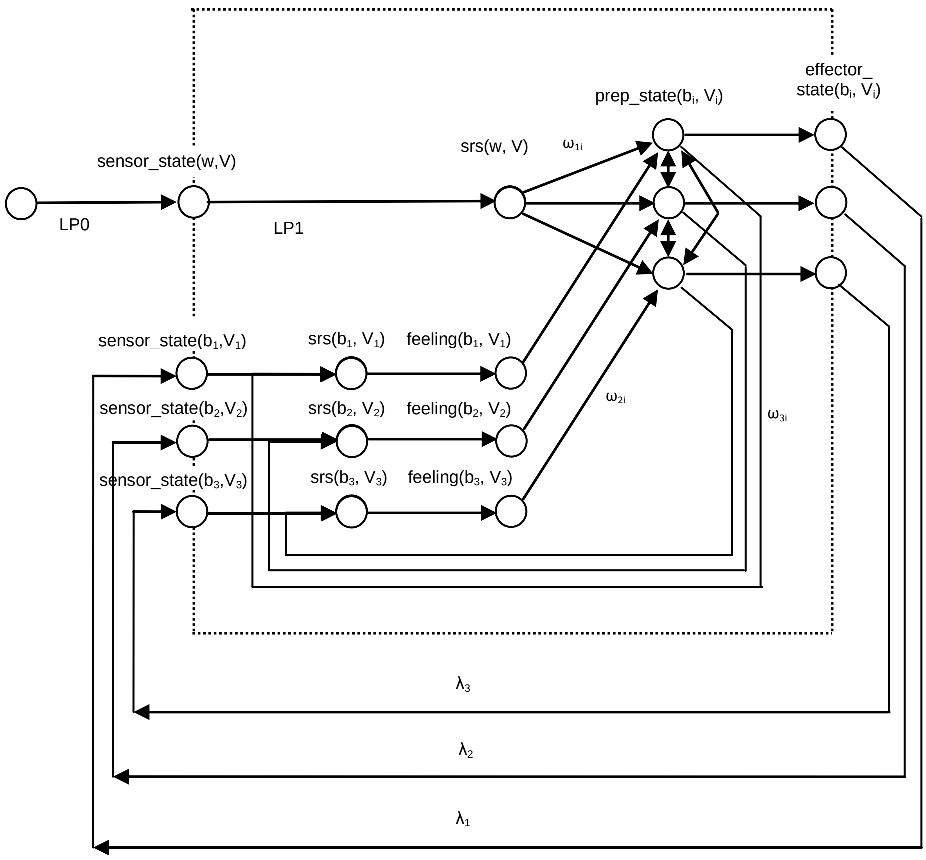

Figure. 1. Overview of the basic model with mutual inhibition

The computational model is based on such neurological notions as valuing in relation to feeling, body loop and as-if body loop. In this paper the considered adaptation mechanisms for the model are based on Hebbian learning with mutual inhibition (Section 5), temporal discounting (Section 3), and memory traces (Section 4). The detailed specification of the basic model is presented below starting with how the world state is sensed.

LP0 Sensing a world state

If world state property w occurs of level *V*
then the sensor state for w will have level *V*.

world_state(W, V) ↠ sensor_state(W, V)

From the sensor state a sensory representation of the world state is generated by dynamic property LP1.

LP1 Generating a sensory representation for a sensed world state

If the sensor state for world state property w has level *V*,
then the sensory representation for w will have level *V*.

sensor_state(W, V) ↠ srs(W, V)

The combination function *h* to combine two inputs which activates a subsequent state is used along with the threshold function *th* to keep the resultant value in the interval [0, 1] as follows:

$$h(\sigma, \tau, V_1, V_2, \omega_1, \omega_2) = th(\sigma, \tau, \omega_1 V_1 + \omega_2 V_2)$$

where V_1 and V_2 are the current activation level of the states and ω_1 and ω_2 are the connection strength of the link between the states; here

$$th(\sigma, \tau, V) = \left(\frac{1}{1+e^{-\sigma(V-\tau)}} - \frac{1}{1+e^{\sigma\tau}}\right) * (1 + e^{-\sigma\tau})$$

where σ is the steepness and τ is the threshold of the given function. Alternatively for higher value of στ , following threshold function might be used.

$$th(\sigma, \tau, V) = \frac{1}{1+e^{-\sigma(V-\tau)}}$$

Dynamic property LP2 describes the generation of the preparation state from the sensory representation of the world state and the feeling.

LP2 From sensory representation and feeling to preparation of a body state

If a sensory representation for w with level *V* occurs
and the feeling associated with body state b_i has level V_i
and the preparation state for b_i has level U_i
and ω_{1i} is the strength of the connection from sensory representation for w to preparation for b_i
and ω_{2i} is the strength of the connection from feeling of b_i to preparation for b_i
and σ_i is the steepness value for preparation of b_i
and τ_i is the threshold value for preparation of b_i
and γ_1 is the person's flexibility for bodily responses
then after Δt the preparation state for body state b_i will have level $U_i + \gamma_1 (h(\sigma_i, \tau_i, V, V_i, \omega_{1i}, \omega_{2i}) - U_i) \Delta t$.

srs(w, V) & feeling(b_i, V_i) & preparation_state(b_i, U_i) &
has_connection_strength(srs(w), preparation(bi), ω_{1i}) & has_connection_strength(feeling(b_i), preparation(b_i), ω_{2i}) &
has_steepness(prep_state(b_i), σ_i) & has_threshold(prep_state(b_i), τ_i)
↠ preparation(b_i, $U_i + \gamma_1 (h(\sigma_i, \tau_i, V, V_i, \omega_{1i}, \omega_{2i}) - U_i) \Delta t$)

Dynamic property LP3 describes the generation of sensory representation of a body state from the respective preparation state and sensory state.

LP3 From preparation and sensor state to sensory representation of a body state

If preparation state for b_i has level X_i
and sensor state for b_i has level V_i
and the sensory representation for b_i has level U_i
and ω_{3i} is the strength of the connection from preparation state for b_i to sensory representation for b_i
and σ_i is the steepness value for sensory representation of b_i
and τ_i is the threshold value for sensory representation of b_i
and γ_2 is the person's flexibility for bodily responses
then after Δt the sensory representation for body state b_i will have level $U_i + \gamma_2 (h(\sigma_i, \tau_i, X_i, V_i, \omega_{3i}, 1) - U_i) \Delta t$.

preparation_state (b_i, X_i) & sensor_state(b_i, V_i) & srs (b_i, U_i) &
has_connection_strength(preparation(bi), srs(bi), , ω_{3i}) & has_steepness(srs(b_i), σ_i) & has_threshold(srs(b_i), τ_i)
↠ srs(b_i, $U_i + \gamma_2 (h(\sigma_i, \tau_i, X_i, V_i, \omega_{3i}, 1) - U_i) \Delta t$)

Dynamic property LP4 describes how the feeling is generated from the sensory representation of the body state.

LP4 From sensory representation of a body state to feeling

If the sensory representation for body state b_i has level *V*,
then b_i will be felt with level *V*.

srs(bi, V) ↠ feeling(bi, V)

LP5 describes how an effector state is generated from respective preparation state.

LP5 From preparation to effector state

If the preparation state for b_i has level V,
then the effector state for body state b_i will have level V.

preparation_state(bi, V) ↠ effector_state(bi, V)

Dynamic property LP6 describes how a sensor state is generated from an effector state.

LP6 From effector state to sensor state of a body state

If the effector state for b_i has level V_i,
and λ_i is world characteristics/ recommendation for the option b_i
then the sensor state for body state b_i will have level $\lambda_i V_i$

effecor_state(bi,V) & has_contribution(effecor_state(bi), sensor_state(bi), λ_i)
↠ effector_state(bi, V) ↠ sensor_state(bi, $\lambda_i V_i$)

For the case studies addressed three options are assumed available in the world for the agent and the objective is to see how rationally an agent makes decisions using a given adaptation model (under static as well as stochastic world characteristics). As mentioned the current paper focuses on three learning mechanisms, i.e., temporal discounting, memory traces and Hebbian learning with mutual inhibition; these will be discussed in details in Section 3, Section 4 and Section 5.

3. An Adaptation Model based on Temporal Discounting

In this section the model is extended with the ability of the adaptation of the links (ω_{1i}, ω_{2i}, and ω_{3i}; see Figure. 1) among different state using the temporal discounting approach. Appendix shows some simulation results in detail. The dynamic properties LP7 to LP9 describe the mechanism of learning of the connection strengths between sensory representation of the world state to preparation, feeling to preparation, and preparation to sensory representation using this approach.

LP7a Temporal discounting learning rule for sensory representation of world state

If the connection from sensory representation of w to preparation of b_i has strength ω_{1i}
and the sensory representation for w has level V and $V>0$
and the discounting rate from sensory representation of w to preparation of b_i is η_i
and the extinction rate from sensory representation of w to preparation of b_i is ζ_i
then after Δt the connection from sensory representation of w to preparation of b_i will have strength $\omega_{1i} + (\eta_i(V- \omega_{1i})- \zeta_i \omega_{1i}) \Delta t$.

has_connection_strength(srs(w), preparation(b_i), ω_{1i}) & srs(w, V) & V>0 &
has_discounting_rate(srs(w), preparation(b_i), η_i) & has_extinction_rate(srs(w), preparation(b), ζ_i)
↠ has_connection_strength(srs(w), preparation(b_i), $\omega_{1i} + (\eta_i (V- \omega_{1i}) - \zeta_i \omega_{1i}) \Delta t$)

LP7b Temporal discounting learning rule for sensory representation of world state

If the connection from sensory representation of w to preparation of b_i has strength ω_{1i}
and the sensory representation for w has level 0
and the extinction rate from sensory representation of w to preparation of b_i is ζ_i
then after Δt the connection from sensory representation of w to preparation of b_i
will have strength $\omega_{1i} - \zeta_i \omega_{1i} \Delta t$.

has_connection_strength(srs(w), preparation(b_i), ω_1) & srs(w, 0) & has_extinction_rate(srs(w), preparation(b_i), ζ_i)
↠ has_connection_strength(srs(w), preparation(b_i), $\omega_{1i} - \zeta_i \omega_{1i} \Delta t$)

LP8a Temporal discounting learning rule for feeling of b_i

If the connection from feeling of b_i to preparation of b_i has strength ω_{2i}
and the feeling for b_i has level V_i and $V_i >0$
and the discounting rate from feeling of b_i to preparation of b_i is η_i
and the extinction rate from feeling of b_i to preparation of b_i is ζ_i
then after Δt the connection from feeling of b_i to preparation of b_i will have strength $\omega_{2i} + (\eta_i(V_i - \omega_{2i})- \zeta_i \omega_{2i}) \Delta t$.

has_connection_strength(feeling(b_i), preparation(b_i), ω_{2i}) & feeling(b_i, V_i) & V_i >0 &
has_discounting_rate(feeling(b_i), preparation(b_i), η_i) & has_extinction_rate(feeling(b_i), preparation(b_i), ζ_i)
↠ has_connection_strength(feeling(b_i), preparation(b_i), $\omega_{2i} + (\eta_i (V_i - \omega_{2i}) - \zeta_i \omega_{2i}) \Delta t$)

LP8b Temporal discounting learning rule for feeling of b_i

If the connection from feeling of b_i to preparation of b_i has strength ω_{2i}
and the feeling for b_i has level 0
and the extinction rate from feeling of b_i to preparation of b_i is ζ_i
then after Δt the connection from feeling of b_i to preparation of b_i will have strength $\omega_{2i} - \zeta_i \omega_{2i} \Delta t$.

has_connection_strength(feeling(b_i), preparation(b_i), ω_{2i}) & feeling(b_i, 0) &
has_extinction_rate(feeling(b_i), preparation(b_i), ζ_i)
↠ has_connection_strength(feeling(b_i), preparation(b_i), $\omega_{2i} - \zeta_i \omega_{2i} \Delta t$)

LP9a Temporal discounting learning rule for preparation of b_i

If the connection from preparation of b_i to sensory representation of b_i has strength ω_{3i}
and the preparation for b_i has level V_i and $V_i > 0$
and the discounting rate from preparation of b_i to sensory representation of b_i is η_i
and the extinction rate from preparation of b_i to sensory representation of b_i is ζ_i
then after Δt the connection from preparation of b_i to sensory representation of b_i will have strength $\omega_{3i} + (\eta_i(V_i - \omega_{3i}) - \zeta_i \omega_{3i}) \Delta t$.

has_connection_strength(preparation(b_i), srs(b_i), ω_{3i}) & preparation(b_i, V_i) & $V_i > 0$ &
has_discounting_rate(preparation(b_i), srs(b_i), η_i) & has_extinction_rate(preparation(b_i), srs(b_i), ζ_i)
$\twoheadrightarrow$ has_connection_strength(preparation(b_i), srs(b_i), $\omega_{3i} + (\eta_i (V_i - \omega_{3i}) - \zeta_i \omega_{3i}) \Delta t$)

LP9b Temporal discounting learning rule for preparation of b_i

If the connection from preparation of b_i to sensory representation of b_i has strength ω_{3i}
and the preparation for b_i has level 0
and the extinction rate from preparation of b_i to sensory representation of b_i is ζ_i
then after Δt the connection from preparation of b_i to sensory representation of b_i will have strength $\omega_{3i} - \zeta_i \omega_{3i} \Delta t$.

has_connection_strength(preparation(b_i), srs(b_i), ω_{3i}) & preparation(b_i, 0) &
has_extinction_rate(preparation(b_i), srs(b_i), ζ_i)
$\twoheadrightarrow$ has_connection_strength(preparation(b_i), srs(b_i), $\omega_{3i} - \zeta_i \omega_{3i} \Delta t$)

4. An Adaptation Model based on Memory Traces

The dynamic properties LP7 to LP9 describe the mechanism of learning of the connection strengths between sensory representation of the world state to preparation, feeling to preparation, and preparation to sensory representation using this approach.

LP7a Discounting memory traces from srs(w) to prep(bi)

If the sensory representation for "w" has strength V
and the preparation of "*bi*" has strength V_i
and the discounted number of memory traces with state srs(w) are X
and the discounted number of memory traces with state srs(w) and successor state preparation(*bi*) are Y
and the discounting rate from sensory representation of w to preparation of "*bi"* is α_i
and the extinction rate from srs(w) to preparation of *bi* is ζ_i
then the discounted number of memory traces with state srs(w) is $X + \alpha_i V - \zeta_i X$
and the discounted number of memory traces with state srs(w) and successor state preparation(*bi*) is $Y + \alpha_i VV_i - \zeta_i Y$

srs(w, V) & preparation(b_i, V_i) & has_discounting_rate(srs(w), preparation(b_i), α_i) &
has_extinction_rate(srs(w), preparation(b_i), ζ_i) & memory_traces_including(srs(w), X) &
memory_traces_including_both(srs(w), preparation(b_i), Y) $\twoheadrightarrow$ memory_traces_including(srs(w),
$X + \alpha_i V - \zeta_i X$) & memory_traces_including_both(srs(w), preparation(b_i), $Y + \alpha_i VV_i - \zeta_i Y$)

Given these numbers the induction strength of the connection from sensory representation to preparation state is determined as Y/X.

LP7b Generation of preparations(bi) based on discounted memory traces

If the discounted number of memory traces with state srs(w) is X
and the discounted number of memory traces with state srs(w) and successor state preparation(*bi*) is Y
then the connection strength from srs(w) to preparation(*bi*) is Y/X

memory_traces_including(srs(w), X) & memory_traces_including_both(srs(w), preparation(bi),
Y) $\twoheadrightarrow$ has_connection_strength(srs(w), preparation(bi), Y/X)

LP8a Discounting memory traces from feeling(bi) to prep(bi)

If the feeling for "b_i" has strength V_i
and the preparation of "*bi*" has strength U_i
and the discounted number of memory traces with state feeling(b_i) are X
and the discounted number of memory traces with state feeling(b_i) and successor state preparation(*bi*) are Y
and the discounting rate from feeling of b_i to preparation of "*bi"* is α_i
and the extinction rate from feeling(bi) to preparation of *bi* is ζ_i
then the discounted number of memory traces with state feeling(b_i) is $X + \alpha_i V_i - \zeta_i X$
and the discounted number of memory traces with state feeling(b_i) and successor state preparation(*bi*) is $Y + \alpha_i V_i U_i - \zeta_i Y$

feeling(b_i, V_i) & preparation(b_i, U_i) & has_discounting_rate(feeling(b_i), preparation(b_i), α_i) &
has_extinction_rate(feeling(bi), preparation(b_i), ζ_i) & memory_traces_including(feeling(b_i), X) &
memory_traces_including_both(feeling(b_i), preparation(b_i), Y) $\twoheadrightarrow$
memory_traces_including(feeling(b_i), $X + \alpha_i V_i - \zeta_i X$) &
memory_traces_including_both(feeling(b_i), preparation(b_i), $Y + \alpha_i V_i U_i - \zeta_i Y$)

Given these numbers the induction strength of the connection from feeling to preparation state is determined as Y/X.

LP8b Generation of preparations(bi) based on discounted memory traces

If the discounted number of memory traces with state feeling(b_i) is X
and the discounted number of memory traces with state feeling(b_i) and successor state preparation(*bi*) is Y

then the connection strength from feeling(b_i) to preparation(*bi*) is *Y/X*

memory_traces_including(feeling(b_i), X) & memory_traces_including_both(feeling(b_i), preparation(bi), Y) $\twoheadrightarrow$ has_connection_strength(feeling(b_i), preparation(bi), Y/X)

LP9a Discounting memory traces from prep(bi) to srs(bi)

If the preparation for "b_i" has strength V_i
and the sensory representation of "*bi*" has strength U_i
and the discounted number of memory traces with state preparation(b_i) are *X*
and the discounted number of memory traces with state preparation(b_i) and successor state srs(*bi*) are *Y*
and the discounting rate from preparation of b_i to sensory representation of "*bi*" is α_i
and the extinction rate from prepration(bi) to srs(*bi)* is ζ_i
then the discounted number of memory traces with state preparation(b_i) is $X + \alpha_i V_i - \zeta_i X$
and the discounted number of memory traces with state preparation(b_i) and successor state srs(*bi*) is $Y + \alpha_i V_i U_i - \zeta_i Y$

preparation(b_i, V_i) & srs(b_i, U_i) & has_discounting_rate(preparartion(b_i), srs(b_i), α_i) &
has_extinction_rate(feeling(bi), preparation(b_i), ζ_i) & memory_traces_including(preparation(b_i), X) &
memory_traces_including_both(preparation(b_i),srs(b_i), Y) $\twoheadrightarrow$
memory_traces_including(preparation(b_i), $X + \alpha_i V_i - \zeta_i X$) &
memory_traces_including_both(preparation(b_i), srs(b_i), $Y + \alpha_i V_i U_i - \zeta_i Y$)

Given these numbers the induction strength of the connection from feeling to preparation state is determined as Y/X.

LP9b Generation of sensory representation(bi) based on discounted memory traces

If the discounted number of memory traces with state preparation (*bi*) is *X*
and the discounted number of memory traces with state preparation(*bi*) and successor state srs(*bi*) is *Y*
then the connection strength from preparation(*bi*) to srs(*bi*) is *Y/X*

memory_traces_including(preparation(b_i), X) & memory_traces_including_both(preparation (b_i), srs(bi), Y) $\twoheadrightarrow$ has_connection_strength(preparation (b_i), srs(bi), Y/X)

5. An Adaptation Model for Hebbian Learning and Inhibition

In this section, the basic agent model described in Section 2 is extended with two additional features. One is mutual inhibition for preparation states and the other one is adaptation of connection strength using a Hebbian learning approach. Later in this section the overview of the model is discussed with detailed specifications in LEADSTO.

An overview of the extended model for the generation of emotional responses and feelings is depicted in Figure. 1 with the extension of *mutual inhibition* at preparation states. This picture also shows representations from the detailed specifications explained in Section 2. In the current section only the local property LP2 is discussed which incorporate the effect of mutual inhibition in calculation of the value of the preparation state. Moreover, local properties LP7 to LP9 will define the Hebbian learning approach.

Dynamic property LP2 describes the generation of the preparation state from the sensory representation of the world state and the feeling thereby taking into account mutual inhibition. For this particular case the combination function is defined as:

$$g(\sigma, \tau, V_1, V_2, V_3, V_4, \omega_1, \omega_2, \theta_1, \theta_2) = th(\sigma, \tau, \omega_1 V_1 + \omega_2 V_2 + \theta_1 V_3 + \theta_2 V_4)$$

where θ_{mi} is the strengths of the mutual inhibition link from preparation state for b_m to preparation state for b_i (which have negative values).

LP2 From sensory representation and feeling to preparation of a body state with mutual inhibition

If a sensory representation for w with level *V* occurs
and the feeling associated with body state b_i has level V_i
and the preparation state for each b_m has level U_m
and ω_{1i} is the strength of the connection from sensory representation for w to preparation for b_i
and ω_{2i} is the strength of the connection from feeling of b_i to preparation for b_i
and θ_{mi} is for each m the strength of the inhibition connection from preparation state for b_j to preparation for b_i
and σ_i is the steepness value for preparation of b_i
and τ_i is the threshold value for preparation of b_i
and γ_1 is the person's flexibility for bodily responses
and $j \neq i, k \neq i, j \neq k$
then after Δt the preparation state for body state b_i will have level $U_i + \gamma_1 (g(\sigma_i, \tau_i, V, V_i, U_j, U_k, \omega_{1i}, \omega_{2i}, \theta_{ji}, \theta_{ki}) - U_i) \Delta t$

srs(w, V) & feeling(b_i, V_i) & $\wedge_m$ preparation_state(b_m, U_m) &
has_connection_strength(srs(w), preparation(bi), ω_{1i}) &
has_connection_strength(feeling(b_i), preparation(b_i), ω_{2i}) &
has_steepness(prep_state(b_i), σ_i) & has_threshold(prep_state(b_i), τ_i) & j≠i, k≠i, j≠k
$\twoheadrightarrow$ preparation(b_i, $U_i + \gamma_1 (g(\sigma_i, \tau_i, V, V_i, U_j, U_k, \omega_{1i}, \omega_{2i}, \theta_{ji}, \theta_{ki}) - U_i) \Delta t$)

Moreover, the connection strength of the different links using Hebbian learning are updated according to the local properties LP7 to LP9 given below.

LP7 Hebbian learning rule for connection from sensory representation of stimulus to preparation
If the connection from sensory representation of w to preparation of b_i has strength ω_{1i}
and the sensory representation for w has level V
and the preparation of *bi* has level V_i
and the learning rate from sensory representation of w to preparation of b_i is η
and the extinction rate from sensory representation of w to preparation of b_i is ζ
then after Δt the connection from sensory representation of w to preparation of *bi* will have strength $\omega_{1i} + (\eta V V_i (1 - \omega_{1i}) - \zeta\omega_{1i}) \Delta t$.
has_connection_strength(srs(w), preparation(b_i), ω_{1i}) & srs(w, V) & preparation(b_i, V_i) &
has_learning_rate(srs(w), preparation(b_i), η) & has_extinction_rate(srs(w), preparation(b_i), ζ)
→ has_connection_strength(w, b_i, $\omega_{1i} + (\eta V V_i (1 - \omega_{1i}) - \zeta\omega_{1i}) \Delta t$)

LP8 Hebbian learning rule for connection from feeling to preparation
If the connection from feeling associated with body state b_i to preparation of b_i has strength ω_{2i}
and the feeling for *bi* has level V_i
and the preparation of *bi* has level U_i
and the learning rate from feeling of *bi* to preparation of b_i is η
and the extinction rate from feeling of *bi* to preparation of b_i is ζ
then after Δt the connection from feeling of *bi* to preparation of *bi*
will have strength $\omega_{2i} + (\eta V_i U_i (1 - \omega_{2i}) - \zeta\omega_{2i}) \Delta t$.
has_connection_strength(feeling(bi), preparation(b_i), ω_{2i}) & feeling(b_i, V_i) & preparation(b_i, U_i) &
has_learning_rate(feeling(b_i), preparation(b_i), η) & has_extinction_rate(feeling(b_i), preparation(b_i), ζ)
→ has_connection_strength(feeling(bi), preparation(b_i), $\omega_{2i} + (\eta V_i U_i (1 - \omega_{2i}) - \zeta\omega_{2i}) \Delta t$)

LP9 Hebbian learning rule for connection from preparation to sensory representation of b_i
If the connection from preparation of b_i to sensory representation of b_i has strength ω_{3i}
and the preparation of *bi* has level V_i
and the sensory representation of *bi* has level U_i
and the learning rate from preparation of *bi* to sensory representation of b_i is η
and the extinction rate from preparation of *bi* to sensory representation of b_i is ζ
then after Δt the connection from preparation of *bi* to sensory representation of *bi* will have strength $\omega_{3i} + (\eta V_i U_i (1 - \omega_{3i}) - \zeta\omega_{3i}) \Delta t$.
has_connection_strength(preparation (bi), srs(b_i), ω_{3i}) & preparation (b_i, V_i) & srs(b_i, U_i) &
has_learning_rate(preparation (bi), srs(b_i), η) & has_extinction_rate(preparation (bi), srs(b_i), ζ)
→ has_connection_strength(preparation (bi), srs(b_i), $\omega_{3i} + (\eta V_i U_i (1 - \omega_{3i}) - \zeta\omega_{3i}) \Delta t$)

In the Appendix an overview of simulation results is shown.

6. Evaluating the Models on Rationality

In the simulation experiments it was shown that the agent model behaves rationally in different scenarios (see Table1 to Table 5). These scenarios and its different cases are elaborated in detail in the appendix, but the results were assessed with respect to the extent of their rationality only in a rather informal manner. In the current section the rationality is determined more formally by two methods developed earlier: by one rationality measure based on a discrete scale and another one based on a continuous scale.

6.1. Method 1: Discrete Rationality Measure

The first method presented is based on the following point of departure: *an agent which has the same respective order of effector state activation levels for the different options compared to the order of world characteristics λ_i will be considered highly rational*. More specifically, the following formula is used to determine the irrationality factor *IF*.

$$IF = \sum_{i=1}^{n} abs(rank(es_i) - rank(\lambda_i))$$

where n is the number of options available. To calculate the discrete rationality factor *DRF*, the maximum possible irrationality factor *Max. IF* can be determined as follows.

$$Max.\ IF = \frac{n(n+1)}{2} - ceiling(\frac{n}{2})$$

Here $ceiling(x)$ is the first integer higher than x. Note that *Max. IF* is approximately $½n^2$. As a higher *IF* means lower rationality, the discrete rationality factor *DRF* is calculated as:

$$DRF = 1 - \frac{IF}{Max.\ IF}$$

6.2. **Method 2: Continuous Rationality Measure**

The second method presented is based on the following point of departure: *an agent which receives the maximum benefit will be the highly rational agent. This is only possible if ES_i is 1 for the option whose λ_i is the highest*. In this method to calculate the continuous rationality factor *CRF*, first to account for the effort spent in performing actions, the effector state values ES_i are normalised as follows.

$$ {}_nES_i = \frac{ES_i}{\sum_{i=1}^{n} ES_i} $$

Here n is number of options available. Based on this the continuous rationality factor *CRF* Is determined as follows, with $Max(\lambda_i)$ the maximal value of the different λ_i.

$$ CRF = \frac{\sum_{i=1}^{n} {}_nES_i \, \lambda_i}{Max(\lambda_i)} $$

This method enables to measure to which extent the agent is behaving rationally in a continuous manner.

6.3. **Rationality for Temporal Discounting**

First the temporal discounting method is discussed. In Figure 2 to Fig. 5 the first 250 time points show the rationality achieved by the agent just before changing world characteristics drastically for the simulations shown Table 1. From time point 250 onwards, it shows the rationality of the agent after the change has been made (see Table 1). It is clear from the results (Fig. 2 to Fig. 5) that the rationality factor of the agent in all four cases improves over the time for the given world.

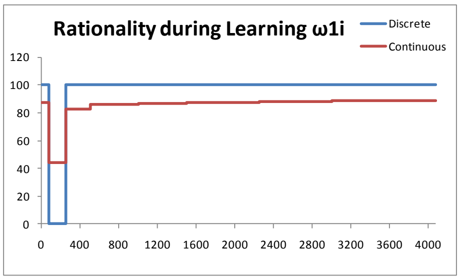

Figure. 2. Rationality during learning ω_{1i}

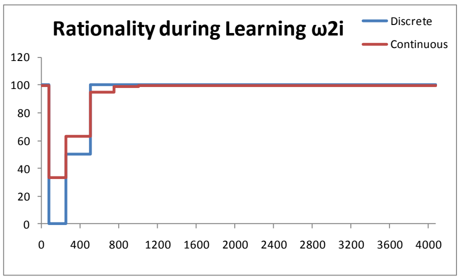

Figure. 3. Rationality during learning ω_{2i}

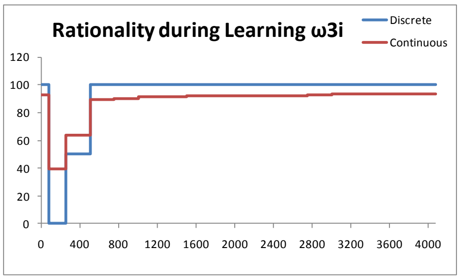

Figure. 4. Rationality during learning ω_{3i}

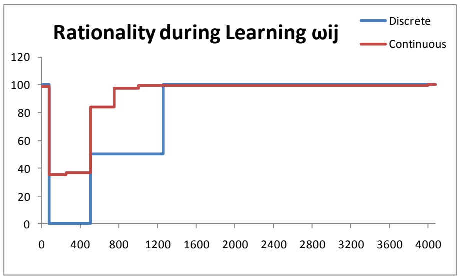

Figure. 5. Rationality during learning ω_{1i}, ω_{2i}, ω_{3i}

6.4. **Rationality for Memory Traces**

Similar set of experiments were executed for Memory Traces approach under same setting as discussed in previous section 6.3. The result for the rationality over time is given from Figure 6 to Figure 9 below. Overview of these results is given in the Table 5.

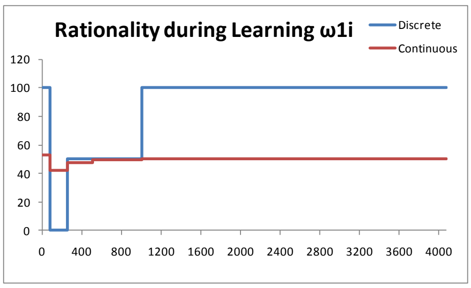

Figure. 6. Rationality during learning ω_{1i}

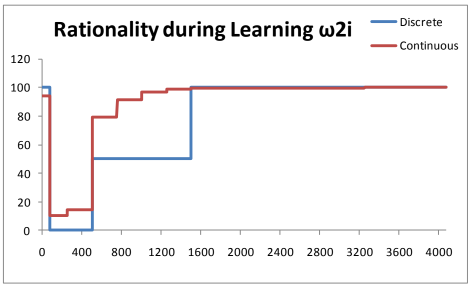

Figure. 7. Rationality during learning ω_{2i},

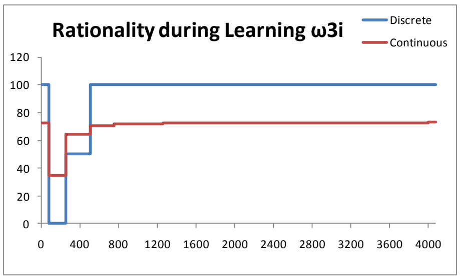

Figure. 8. Rationality during learning ω_{3i}

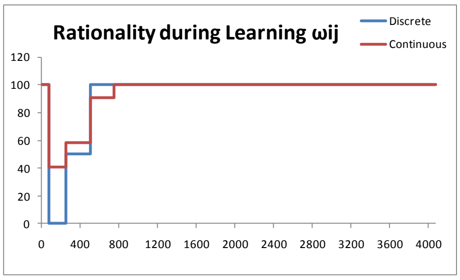

Figure. 9. Rationality during learning ω_{1i}, ω_{2i}, ω_{3i}

6.5. Rationality for Hebbian Learning with Mutual Inhibition

In this section the results for the rationality measures for Hebbian learning with inhibition are presented and it is shown how the change in mutual inhibition factor effects the rationality. For this purpose the rationality was measured for inhibition factors from *0.0* to *0.2* for all four cases of learning different links between states i.e. learning of the connection strengths between sensory representation to preparation, feeling to preparation and preparation to sensory representation one at a time and simultaneously. Due to space limitation only the rationality results related to learning all links simultaneously are presented in this paper from Figure 10 to Figure. 13.

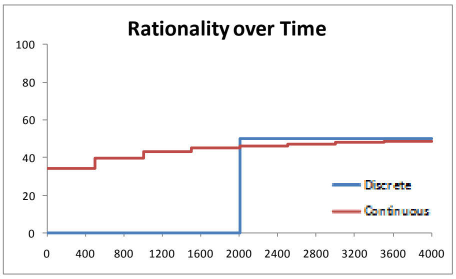

Figure. 10. $\theta_1 = \theta_2 = \theta_3 = 0.0$

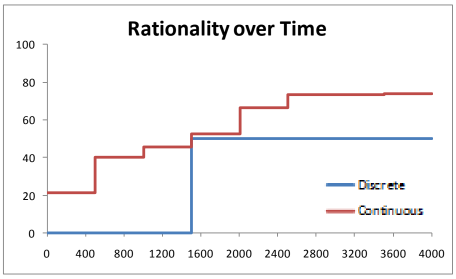

Figure. 11. $\theta_1 = \theta_2 = \theta_3 = 0.05$

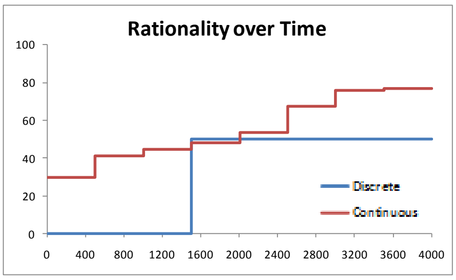

Figure. 12. $\theta_1 = \theta_2 = \theta_3 = 0.10$

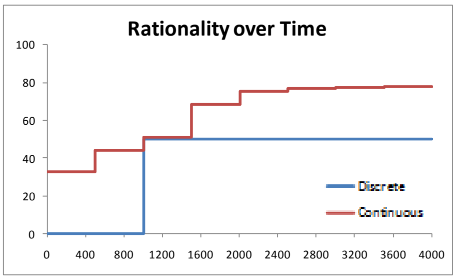

Figure. 13. $\theta_1 = \theta_2 = \theta_3 = 0.15$

The results show that the inhibition factor has a positive effect on the rationality measure in both aspects i.e. discrete and continuous. With the increase in inhibition factor either the rationality level is increased or the rationality level is achieved faster. Rationality level achieved based on CRF is 48.63, 73.81, 77.08 and 80.23 for inhibition factor 0.0, 0.05, 0.1 and 0.15 respectively. Similarly based on DRF same rationality level 50 is achieved at time point 2000, 1500, 1500, 1000 for inhibition factor 0.0, 0.05, 0.1 and 0.15 respectively.

7. Discussion

This paper focused on three different adaptive agent models: the first based on temporal discounting, the second on memory traces and the third based on Hebbian learning with mutual inhibition; cf. [10, 12]. For all described adaptive models an analysis of the extent of rationality was made. The basic agent model in which adaptation models were incorporated is based on emotion-related valuation of predictions, involving feeling states generated in the amygdala; e.g., [1, 2, 6, 8, 14, 15, 17].

The assessment of the extent of rationality with respect to given world characteristics, was based on two measures. It was shown how by the learning processes indeed a high level of rationality was obtained, and how after major world changes with some delay this rationality level is re-obtained. It turned out that emotion-related valuing of predictions in the amygdala as a basis for adaptive decision making according to temporal discounting, to memory traces and to Hebbian learning all satisfy reasonable rationality measures.

References

[1] Bechara, A., Damasio, H., and Damasio, A.R.: Role of the Amygdala in Decision-Making. Ann. N.Y. Acad. Sci. 985, 356-369 (2003)

[2] Bechara, A., Damasio, H., Damasio, A.R., and Lee, G.P.: Different Contributions of the Human Amygdala and Ventromedial Prefrontal Cortex to Decision-Making. Journal of Neuroscience 19, 5473–5481 (1999)

[3] Bosse, T., Hoogendoorn, M., Memon, Z.A., Treur, J., and Umair, M.: An Adaptive Model for Dynamics of Desiring and Feeling based on Hebbian Learning. In: Yao, Y., Sun, R., Poggio, T., Liu, J., Zhong, N., and Huang, J. (eds.), *Proceedings of the Second International Conference on Brain Informatics, BI'10*. Lecture Notes in Artificial Intelligence, vol. 6334, pp. 14-28. Springer Verlag (2010)

[4] Bosse, T., Jonker, C.M., and Treur, J.: Formalisation of Damasio's Theory of Emotion, Feeling and Core Consciousness. *Consciousness and Cognition Journal* 17, 94–113 (2008)

[5] Bosse, T., Jonker, C.M., Meij, L. van der, Treur, J.: A Language and Environment for Analysis of Dynamics by Simulation. Intern. J. of AI Tools 16, 435-464 (2007)

[6] Damasio, A.: Descartes' Error: Emotion, Reason and the Human Brain. Papermac, London (1994)

[7] Damasio, A.: The Feeling of What Happens. Body and Emotion in the Making of Consciousness. New York: Harcourt Brace (1999)

[8] Damasio, A.: *Looking for Spinoza*. Vintage books, London (2004)

[9] Forgas, J.P., Laham, S.M., and Vargas, P.T.: Mood effects on eyewitness memory: Affective influences on susceptibility to misinformation. *Journal of Experimental Social Psychology* 41, 574–588 (2005)

[10] Gerstner, W., and Kistler, W.M.: Mathematical formulations of Hebbian learning. Biol. Cybern. 87, 404-415 (2002)

[11] Ghashghaei, H.T., Hilgetag, C.C., Barbas, H.: Sequence of information processing for emotions based on the anatomic dialogue between prefrontal cortex and amygdala. Neuroimage 34, 905-923 (2007)

[12] Hebb, D.O.: *The Organization of Behaviour.* New York: John Wiley & Sons (1949)

[13] Hesslow, G.: Conscious thought as simulation of behaviour and perception. Trends Cogn. Sci. 6, 242-247 (2002)

[14] Montague, P.R., Berns, G.S.: Neural economics and the biological substrates of valuation. Neuron 36, 265-284 (2002)

[15] Morrison, S.E., and Salzman, C.D.: Re-valuing the amygdala. Current Opinion in Neurobiology 20, 221–230 (2010)

[16] Murray EA: The amygdala, reward and emotion. Trends Cogn Sci 11, 489-497 (2007)

[17] Rangel, A., Camerer, C., Montague, P.R.: A framework for studying the neurobiology of value-based decision making. Nat Rev Neurosci 9, 545-556 (2008)

[18] Salzman, C.D., and Fusi, S.: Emotion, Cognition, and Mental State Representation in Amygdala and Prefrontal Cortex. Annu. Rev. Neurosci. 33, 173–202 (2010)

[19] Sugrue, L.P., Corrado, G.S., Newsome, W.T.: Choosing the greater of two goods: neural currencies for valuation and decision making. Nat Rev Neurosci 6, 363-375 (2005)

[20] Shors, T.J. (2004). Memory traces of trace memories: neurogenesis, synaptogenesis and awareness. Trends in Neurosciences, vol.27, 2004, pp. 250-256.

[21] Weinberger, N.M. (2004). Specific Long-Term Memory Traces In Primary Auditory Cortex. Nature Reviews: Neuroscience, vol. 5, 2004, pp. 279-290.

[22] Treur, J., and Umair, M.,On Rationality of Decision Models Incorporating Emotion-Related Valuing and Hebbian Learning. In: Lu, B.-L., Zhang, L., Kwok, J. (eds.), Proceedings of the 18th International Conference on Neural Information Processing, ICONIP'11, Part III. Lecture Notes in Artificial Intelligence, vol. 7064, pp. 217–229. Springer Verlag, 2011.

Appendix Simulation Results

As mentioned earlier in this paper we are presenting an overview of the simulation results by help of Table to Table instead of detailed graphs due to space limitation. The data in the given tables shows the final connection strengths of different links along with the activation level of effector states (ES_i). Table 1 presents the summarized results for Temporal Discounting approach.

Table 1. Overview of Model using Temporal Discounting

Link	Scenario	ω_{x1}	ω_{x2}	ω_{x3}	ES_1	ES_2	ES_3	World State
A	Static	0.97	0.97	0.97	0.71	0.66	0.63	λ_1=0.9,λ_2=0.2, λ_3=0.1
	Stochastic	0.97	0.97	0.97	0.71	0.66	0.63	
	Change World	0.97	0.97	0.97	0.63	0.66	0.71	λ_1=0.1,λ_2=0.2, λ_3=0.9
B	Static	0.92	0.37	0.28	0.60	0.29	0.27	λ_1=0.9,λ_2=0.2, λ_3=0.1
	Stochastic	0.92	0.36	0.28	0.06	0.29	0.27	
	Change World	0.28	0.36	0.92	0.27	0.29	0.60	λ_1=0.1,λ_2=0.2, λ_3=0.9
C	Static	0.57	0.27	0.26	0.59	0.28	0.26	λ_1=0.9,λ_2=0.2, λ_3=0.1
	Stochastic	0.56	0.27	0.26	0.59	0.28	0.26	
	Change World	0.26	0.27	0.57	0.26	0.28	0.59	λ_1=0.1,λ_2=0.2, λ_3=0.9
ABC	Static	0.98	0.98	0.98	0.63	0.34	0.33	λ_1=0.9,λ_2=0.2, λ_3=0.1
		0.90	0.16	0.10				
		0.52	0.32	0.32				
	Stochastic	0.98	0.98	0.98	0.61	0.34	0.33	
		0.88	0.17	0.11				
		0.50	0.32	0.32				
	Changed World	0.98	0.98	0.98	0.33	0.34	0.62	λ1=0.1,λ2=0.2 ,λ3=0.9
		0.12	0.17	0.89				
		0.32	0.32	0.51				

Next Table 2 shows the results for Hebbian Learning approach without any mutual inhibition among the different links i.e . θ1 = θ2 = θ3 = 0.0

Table 2. Overview of Model using Hebbian Learning without Mutual inhibition

Link	Scenario	ω_{x1}	ω_{x2}	ω_{x3}	ES_1	ES_2	ES_3	World State
A	Static	0.78	0.53	0.52	0.56	0.15	0.14	λ_1=0.9,λ_2=0.2, λ_3=0.1
	Stochastic	0.78	0.53	0.52	0.56	0.15	0.14	
	Change World	0.40	0.38	0.80	0.09	0.09	0.58	λ_1=0.1,λ_2=0.2, λ_3=0.9
B	Static	0.89	0.58	0.46	0.65	0.39	0.31	λ_1=0.9,λ_2=0.2, λ_3=0.1
	Stochastic	0.89	0.57	0.47	0.66	0.38	0.32	
	Change World	0.43	0.56	0.88	0.31	0.37	0.65	λ_1=0.1,λ_2=0.2, λ_3=0.9
C	Static	0.88	0.29	0.23	0.63	0.28	0.26	λ_1=0.9,λ_2=0.2, λ_3=0.1
	Stochastic	0.88	0.3	0.24	0.63	0.29	0.27	
	Change World	0.05	0.08	0.87	0.25	0.26	0.63	λ_1=0.1,λ_2=0.2, λ_3=0.9
ABC	Static	0.81	0.55	0.55	0.59	0.13	0.13	λ_1=0.9,λ_2=0.2, λ_3=0.1
		0.85	0.30	0.29				
		0.85	0.30	0.29				
	Stochastic	0.81	0.55	0.55	0.58	0.13	0.13	
		0.85	0.30	0.29				
		0.85	0.30	0.29				
	Changed World	0.65	0.65	0.93	0.16	0.16	0.76	λ1=0.1,λ2=0.2 ,λ3=0.9
		0.02	0.03	0.96				
		0.02	0.03	0.96				

Table 3 shows the data of Hebbian Learning approach with mutual inhibition factor 0.1 i.e. θ1 = θ2 = θ3 = 0.1

Table 3. Overview of Model using Hebbian Learning with Mutual inhibition

Link	Scenario	ω_{x1}	ω_{x2}	ω_{x3}	ES_1	ES_2	ES_3	World State
A	**Static**	0.76	0.49	0.48	0.53	0.11	0.11	λ_1=0.9,λ_2=0.2, λ_3=0.1
	Stochastic	0.76	0.49	0.48	0.53	0.11	0.11	
	Change World	0.31	0.26	0.78	0.06	0.04	0.55	λ_1=0.1,λ_2=0.2, λ_3=0.9
B	**Static**	0.88	0.44	0.36	0.63	0.27	0.25	λ_1=0.9,λ_2=0.2, λ_3=0.1
	Stochastic	0.88	0.44	0.37	0.63	0.27	0.25	
	Change World	0.30	0.38	0.88	0.24	0.26	0.62	λ_1=0.1,λ_2=0.2, λ_3=0.9
C	**Static**	0.87	0.24	0.20	0.61	0.24	0.22	λ_1=0.9,λ_2=0.2, λ_3=0.1
	Stochastic	0.87	0.24	0.21	0.61	0.24	0.22	
	Change World	0.03	0.05	0.86	0.22	0.22	0.60	λ_1=0.1,λ_2=0.2, λ_3=0.9
ABC	**Static**	0.73	0.52	0.52	0.42	0.11	0.11	λ_1=0.9,λ_2=0.2, λ_3=0.1
		0.73	0.29	0.28				
		0.73	0.29	0.28				
	Stochastic	0.72	0.52	0.52	0.41	0.11	0.11	
		0.72	0.29	0.29				
		0.72	0.29	0.29				
	Changed World	0.61	0.61	0.95	0.14	0.14	0.79	λ1=0.1,λ2=0.2 ,λ3=0.9
		0.01	0.02	0.96				
		0.01	0.02	0.96				

During our experiment a number of simulations were run for different interesting scenarios. One of those was to study the effect of change of inhibition factors on the learning of different links among different states. The effect of change of inhibition factor from 0.0 to 0.2 on the connection strength (sensory representation and preparation states) and effector states are shown in Table 4.

Table 4. Effect of Mutual Inhibition Factors

Inhibition	ω_{11}	ω_{12}	ω_{13}	ES_1	ES_2	ES_3
0.00	0.65	0.65	0.69	0.17	0.16	0.21
0.05	0.68	0.68	0.94	0.19	0.18	0.79
0.10	0.63	0.63	0.95	0.15	0.15	0.79
0.15	0.61	0.61	0.95	0.13	0.13	0.79
0.20	0.56	0.55	0.95	0.11	0.10	0.82

The results for the Memory Traces approach are presented in Table 5 below.

Table 5. Overview of Model using Memory Traces

Link	Scenario	ω_{x1}	ω_{x2}	ω_{x3}	ES_1	ES_2	ES_3	World State
A	**Static**	0.29	0.25	0.25	0.51	0.45	0.44	λ_1=0.9,λ_2=0.2, λ_3=0.1
	Stocastic	0.29	0.25	0.25	0.51	0.45	0.44	
	Change World	0.69	0.69	0.70	0.89	0.89	0.91	λ_1=0.1,λ_2=0.2, λ_3=0.9
B	**Static**	0.86	0.03	0.02	0.99	0.04	0.04	λ_1=0.9,λ_2=0.2, λ_3=0.1
	Stochastic	0.88	0.03	0.03	0.99	0.04	0.04	
	Change World	0.03	0.03	0.98	0.04	0.04	0.99	λ_1=0.1,λ_2=0.2, λ_3=0.9
C	**Static**	0.13	0.03	0.01	0.04	0.03	0.03	λ_1=0.9,λ_2=0.2, λ_3=0.1
	Stochastic	0.13	0.03	0.01	0.04	0.03	0.03	
	Change World	0.01	0.03	0.13	0.03	0.03	0.04	λ_1=0.1,λ_2=0.2, λ_3=0.9
ABC	**Static**	0.77	0.18	0.17	0.99	0.35	0.31	λ_1=0.9,λ_2=0.2, λ_3=0.1
		0.94	0.21	0.19				
		0.99	0.08	0.03				
	Stocastic	0.77	0.18	0.17	0.99	0.36	0.32	
		0.87	0.22	0.21				
		0.98	0.08	0.04				
	Changed World	0.19	0.21	0.84	0.31	0.34	1.00	$\lambda1$=0.1,$\lambda2$=0.2 ,$\lambda3$=0.9
		0.23	0.22	0.98				
		0.09	0.08	0.99				