

# A Cognitive Agent Model Using Inverse Mirroring for False Attribution of Own Actions to Other Agents

Jan Treur<sup>1</sup>, Muhammad Umair<sup>1,2</sup>

<sup>1</sup>Vrije Universiteit Amsterdam, Department of Artificial Intelligence  
De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands

<sup>2</sup>COMSATS Institute of Information Technology, Lahore, Pakistan  
Email: <sup>1</sup>{j.treur, m.umair}@vu.nl; <sup>2</sup>mumair@ciitlahore.edu.pk  
URL: <http://www.few.vu.nl/~{treur, mumair}>

**Abstract.** This paper presents a cognitive agent model capable of showing situations where self-generated actions are attributed to other agents, as, for example, for patients suffering from schizophrenia. The mechanism underlying the model involves inverse mirroring: mapping preparation states onto sensory representations of observed actions. It is shown how this mechanism can develop based on Hebbian learning. The model provides a basis for applications to human-like virtual agents in the context of for example, training of therapists or agent-based generation of virtual stories.

**Keywords:** action attribution, cognitive agent model, schizophrenia

## 1 Introduction

For the development of agent models as a basis for virtual agents in serious or nonserious gaming, an often used criterion is that they show realistic human-like behaviour. One of the ways to obtain such human-like agent models is to exploit the fast growing amount of neurological literature, so that models are developed that have biological plausibility. In addition, to obtain realistic virtual agents, not only ideally functioning persons should be considered but also persons with deviant behaviour, in order to cover larger parts of the variety in types of behaviour as occurring naturally in the overall human population. This paper addresses an agent model for such naturally occurring deviant behaviour in attribution of self-generated actions to other agents. False attribution of self-generated (e.g., manual or verbal) actions to other agents is a common symptom occurring in patients with schizophrenia. One explanation put forward for the phenomenon that self-generated actions are not attributed to oneself is that self-attribution depends on prediction and monitoring of effects of these actions, and this does not function well for persons with schizophrenia; see, for example [6], [7], [9], [10]. However, in other work it is debated whether this is an appropriate explanation. For example, in [8] experimental work is reported that indicates that differences in these respects between patients with schizophrenia and a control group are not very convincing. In [16] it is argued that a more important role is played by what is called ‘the sense of agency’ (which is at a more conscious, personal level) than action effect prediction and monitoring (which is at a unconscious, subpersonal level).

Note that the issue of not attributing a self-generated action to oneself, as addressed in the literature as mentioned above, is not the same as attributing such an action to another agent, as in order to create a mental image of somebody else performing the action requires a shift from a representation of an action from a first-person to a representation from a third-person perspective (mental rotation). Patients with schizophrenia do not only fail to attribute self-generated actions to themselves, they also attribute them to other agents (which can be real or imaginary). Using neurological literature on mirroring [4], [14], [15], [20] and self-other differentiation [17], in this paper this form of false attribution to other agents is addressed.

In this paper, in Section 2 the ideas from the neurological literature are briefly discussed. In Section 3 the cognitive agent model is introduced. Section 4 presents a number of simulation results. In Section 5 a mathematical analysis is made. Finally, Section 6 is a discussion.

## 2 The Cognitive Agent Model for False Attribution of Actions

In this section the cognitive agent model and its detailed specifications are presented. First the modelling some background knowledge is briefly discussed, next the format used is briefly introduced, and the example scenario used is described, and finally the agent model is addressed in detail.

**Background knowledge** One of the recent neurological findings concerns the *mirroring function* of certain neurons; e.g., [4], [14], [15], [20]. Mirror neurons are active not only when a person prepares for a specific action or body change, but also when the person observes somebody else intending or performing this action or body change. This includes expressing emotions in body states, such as facial expressions. The idea is that these neurons and the neural circuits in which they are embedded play an important role in social functioning and in (empathic) understanding of others; e.g., [4], [14], [15], [20]. Their discovery is often considered a crucial step for the further development of the discipline of social cognition; cf. [15]. When states of other persons are mirrored by some of the person's own states that at the same time are connected via neural circuits to states that are crucial for the person's own feelings and actions (shared circuits), then this provides an effective basic mechanism for how in a social context persons fundamentally affect each other's actions and feelings; e.g. [14].

Mirroring involves a change of perspective from another agent (third person) to oneself (first person). This requires a nontrivial *mental rotation* transformation of the available representations (cf. [17]): sensory representations of observed actions of other agents are mapped onto representational structures for self-generated actions. *Attribution* a self-generated action *to another agent* is in fact a kind of reverse process. It requires a change of perspective from preparation for a self-generated action (first person) to another agent (third person) perspective, based on a reverse mental rotation transformation of the available representations. In fact this is *inverse mirroring*: the representational structures for self-generated actions are mapped onto sensory representations of observed actions of other agents, thus forming a mental image of somebody else performing the action. When it is assumed that such an

inverse mapping is made, a self-generated action is perceived as observed from a third person perspective, and thus it provides a mechanism for the self-generated action to be attributed to another agent.

A further question is how such a reverse mental rotation mapping can exist or develop in a neurological context. One possibility is that the mechanism is there initially, due to improper genetics. However, another possibility is that it is developed during lifetime. This is also analysed below, assuming a *Hebbian learning* principle. This is the principle that connected neurons that are frequently activated simultaneously strengthen their connecting synapse. The principle goes back to Hebb [12], but has recently gained enhanced interest by more extensive empirical support (e.g., [2]), and more advanced mathematical formulations (e.g., [11]). In the models a variant of this principle has been adopted to realise an inverse mirroring connection from preparation of an action to sensory representation of a similar observed action.

**Modelling Format Used** To formalise the agent model in an executable manner, the hybrid dynamic modelling language LEADSTO has been used; cf. [3]. Within LEADSTO the dynamic property or temporal relation  $a \rightarrow_D b$  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  $D$ ), state property  $b$  will occur. Below, this  $D$  will be taken as the time step  $\Delta t$ , and usually not be mentioned explicitly. Both logical and quantitative calculations can be specified, and a software environment is available to support specification and simulation. In most cases in the model below some form of combination function  $f$  is used. A typical example of an update rule in LEADSTO is

$$SS(w_1, V_1) \& SS(w_2, V_2) \& SR(w_3, V_3) \rightarrow SR(w_3, V_3 + \gamma [ f(\omega_1 V_1, \omega_2 V_2) - V_3 ] \Delta t)$$

which expresses that when the sensor states for  $w_1$  and  $w_2$  have values  $V_1$  and  $V_2$  respectively and the sensory representation of  $w_3$  has value  $V_3$ , then after time duration  $\Delta t$  this sensory representation will have value  $V_3 + \gamma [ f(\omega_1 V_1, \omega_2 V_2) - V_3 ] \Delta t$ . Here  $\omega_1$  and  $\omega_2$  are the connection strengths from the sensor states to the sensory representation, respectively, and  $\gamma$  is an update speed factor. Moreover,  $f$  is a combination function, for which different choices can be made, for example,

$$f(W_1, W_2) = \beta(1 - (1 - W_1)(1 - W_2)) + (1 - \beta)W_1W_2 \quad (0 \leq \beta \leq 1)$$

In simulations with the agent model, a combination function  $f(W_1, \dots, W_n)$  based on a continuous logistic threshold function  $th(\sigma, \tau, W)$  has been used of the form  $f(W_1, \dots, W_n) = th(\sigma, \tau, W_1 + \dots + W_n)$  (where  $\sigma$  is a steepness and  $\tau$  a threshold value) with:

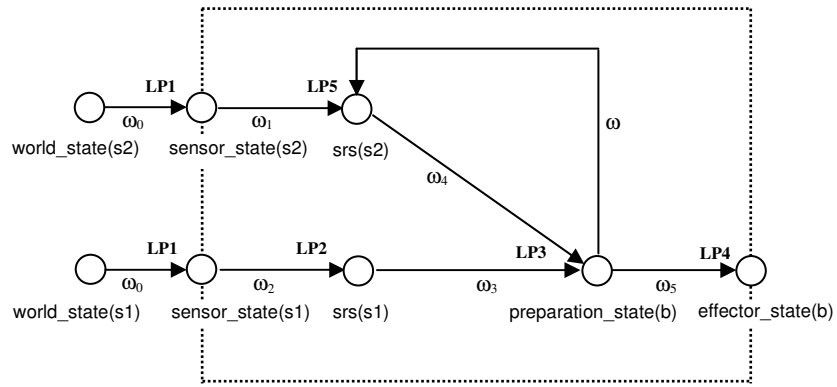
$$th(\sigma, \tau, W) = \left( \frac{1}{1 + e^{-\sigma(W - \tau)}} - \frac{1}{1 + e^{\sigma\tau}} \right) (1 + e^{-\sigma\tau})$$

For higher values of  $\sigma\tau$  this threshold function is approximated by the expression:

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

**Example Scenario** The designed agent model will be illustrated for the following scenario. A sensed stimulus  $s1$  leads to a sensory representation of this stimulus, which in turn triggers preparation and execution of an action  $b$  as a response of the agent; see the causal chain in the lower part of Fig. 1. Moreover, when another agent performs action  $b$ , this is represented as a stimulus  $s2$  that can be sensed; see upper part of Fig. 1. The sensory representation  $srs(s2)$  indicates the mental image of another

person performing the action. The agent's mirroring capability is based on the assumption that an activated sensory representation of  $s_2$  will also activate the agent's own preparation for  $b$ . This assumption has been confirmed in neurological literature such as [4], [14], [15], [20]. When this latter chain of events happens (i.e., whenever mirroring takes place), for the model introduced here it is assumed that by Hebbian learning this will strengthen the connection from preparation of  $b$  to sensory representation of  $s_2$  (observed action), thus developing inverse mirroring capabilities. When such a learning process has achieved substantial connection strength, the agent's response on stimulus  $s_1$  has changed. When  $s_1$  is sensed (in the absence of  $s_2$ ), not only will the agent trigger preparation and execution of action  $b$  as before, but in addition it will generate a mental image of another agent performing action  $b$  (the sensory representation  $srs(s_2)$ ), thus creating a third person perspective on the action.



**Fig. 1.** Overview of the cognitive agent model

**Table 1.** State properties used

notation	description	notation	description
WS(s)	world state for s	PS(b)	preparation state for b
SS(s)	sensor state for s	ES(b)	execution state for b
SRS(s)	sensory representation state of s	cs( $\omega$ )	strength of connection $\omega$ (from preparation of b to sensory representation of s2)

**Table 2.** Overview of the connections and their weights

from states	to state	weights	process	LP
WS(s)	SS(s)	$\omega_0$	sensing world state	LP1
SS(s1)	SRS(s1)	$\omega_2$	representing world state	LP2
SRS(s1), SRS(s2)	PS(b)	$\omega_3, \omega_4$	preparing action / mirroring action	LP3
PS(b)	ES(b)	$\omega_5$	action execution	LP4
SS(s2), PS(b)	SRS(s2)	$\omega_1, \omega$	representing observed action / inverse mirroring action	LP5
SRS(s2), PS(b)	cs( $\omega$ )	$\eta, \zeta$	Hebbian learning for inverse mirroring	LP6

**Detailed Specification of the Agent Model** In the detailed specification, states have been formalised as shown in Table 1. Moreover, the dynamical relationships between these states are shown in Table 2. Note that each connection has a fixed strength, except the connection from preparation of b to sensory representation of s2, which is adapted over time by the Hebbian learning. More detailed specifications of the dynamical relationships are presented below. First it is shown how in a generic manner sensor states are generated from world states.

**LP1 Generating a sensor state**

If world state s has level  $V_1$  and the sensor state for s has level  $V_2$   
then after duration  $\Delta t$  the sensor state for s will have level  $V_2 + \gamma [f(\omega_0 V_1) - V_2] \Delta t$   
 $WS(s, V_1) \ \& \ SS(s, V_2) \ \rightarrow \ SS(s, V_2 + \gamma [f(\omega_0 V_1) - V_2] \Delta t)$

Note that this applies to both s1 and s2. Activation for sensory representation for s1 has been modelled in a straightforward manner as shown in LP2.

**LP2 Sensory representation for a sensor state**

If property s1 is sensed with level  $V_1$  and the sensory representation of s1 has level  $V_2$ .  
then after  $\Delta t$  the sensory representation of s will have level  $V_2 + \gamma [f(\omega_2 V_1) - V_2] \Delta t$ .  
 $SS(s1, V_1) \ \& \ SRS(s1, V_2) \ \rightarrow \ SRS(s1, V_2 + \gamma [f(\omega_2 V_1) - V_2] \Delta t)$

Next it is shown how an action preparation for b is generated. This is based on the sensory representation of stimulus s1, or based on the sensory representation of s2 representing an observed action of another agent; in the latter case mirroring takes place. Note that the two options are not exclusive.

**LP3 Preparing and mirroring an action**

If sensory representation of s1 has level  $V_1$ , and sensory representation of s2 has level  $V_2$ ,  
and the preparation for b has level  $V_3$   
then after  $\Delta t$  the preparation state for b will have level  $V_3 + \gamma (f(\omega_3 V_1, \omega_4 V_2) - V_3) \Delta t$ .  
 $SRS(s1, V_1) \ \& \ SRS(s2, V_2) \ \& \ PS(b, V_3) \ \rightarrow \ PS(b, V_3 + \gamma (f(\omega_3 V_1, \omega_4 V_2) - V_3) \Delta t)$

Action execution has been modelled in a straightforward manner as shown in LP4.

**LP4 Action execution**

If preparation for b has level  $V_1$  and the action execution state for b has level  $V_2$   
then after  $\Delta t$  the action execution state for b will have level  $V_2 + \gamma (f(\omega_5 V_1) - V_2) \Delta t$ .  
 $PS(b, V_1) \ \& \ ES(b, V_2) \ \rightarrow \ ES(b, V_2 + \gamma (f(\omega_5 V_1) - V_2) \Delta t)$

Next it is shown in LP5 how the sensory representation for s2 (a mental image of an observed action) is generated based on a sensor state for s2 or an action preparation for b (or a combination of both). In the latter case inverse mirroring take place.

**LP5 Representing a world state and inverse mirroring of an action**

If sensor state of s2 has level  $V_1$ , and preparation of b has level  $V_2$ ,  
and the sensory representation of s2 has level  $V_3$   
then after  $\Delta t$  the sensory representation of s2 will have level  $V_3 + \gamma (f(\omega_1 V_1, \omega V_2) - V_3) \Delta t$ .  
 $SS(s2, V_1) \ \& \ PS(b, V_2) \ \& \ SRS(s2, V_3) \ \rightarrow \ SRS(s2, V_3 + \gamma (f(\omega_1 V_1, \omega V_2) - V_3) \Delta t)$

Finally, it is shown in LP6 how the Hebbian learning process of the connection from preparation state for b to sensory representation s2 of an observed action was modelled. This takes place using the following *Hebbian learning rule*, with maximal connection strength  $I$ , a *learning rate*  $\eta$ , and *extinction rate*  $\zeta$  (usually taken small):

$$\Delta \omega = \gamma [ \eta V_1 V_2 (I - \omega) - \zeta \omega ] \Delta t$$

Here  $V_1$  and  $V_2$  are (time-dependent) activation levels of the connected nodes, and  $\gamma$  is an adaptation speed factor. In differential equation format it can be written as

$$\frac{d\omega}{dt} = \gamma [ \eta V_1 V_2 (1 - \omega) - \zeta \omega ] = \gamma [ \eta V_1 V_2 - (\eta V_1 V_2 + \zeta) \omega ]$$

A similar Hebbian learning rule can be found in [11], p. 406. By the factor  $(1 - \omega)$  the learning rule keeps the level of  $\omega$  bounded by  $1$ . When the extinction rate is relatively low, the upward changes during learning are proportional to both  $V_1$  and  $V_2$  and maximal learning takes place when both are  $1$ . Whenever one of them is close to  $0$ , extinction takes over, and  $\omega$  slowly decreases. This is specified as follows:

**LP6 Learning for inverse mirroring**

If the sensory representation of stimulus s2 has level  $V_1$ ,

and the preparation for b has level  $V_2$ ,

and the connection weight from preparation for b to sensory representation of s2 has level  $W$ ,

then after duration  $\Delta t$  the connection weight from preparation for b to sensory representation of s2 will have level  $W + \gamma [ \eta V_1 V_2 (1 - W) - \zeta W ] \Delta t$ .

$$\text{SRS}(s_2, V_1) \ \& \ \text{PS}(b, V_2) \ \& \ \text{cs}(\omega, W) \ \rightarrow \ \text{cs}(\omega, W + \gamma [ \eta V_1 V_2 (1 - W) - \zeta W ] \Delta t)$$

### 3 Simulation Results

A number of simulations have been performed with the focus of simulating normal functioning and deviant functioning of the model. *Normal functioning* of the agent occurs by parameter settings in which stimulus s1 does not lead to high activation of sensory representation of s2 (i.e., no mental image of somebody else performing the action is created) in absence of the stimulus s2, although stimulus s2 has occurred time and time again in the past. In contrast, *deviant functioning* occurs by parameter setting in which stimulus s1 does lead to high activation of sensory representation of s2 under similar circumstances, so in this case a mental image of somebody else performing the action is created.

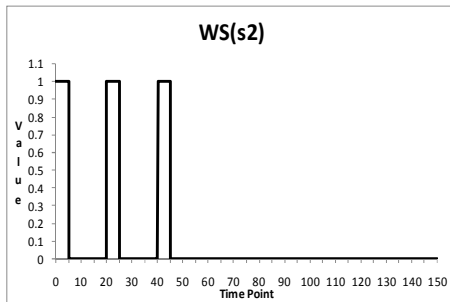


Fig. 1. World State for s2

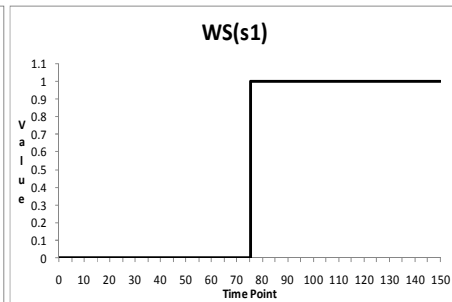
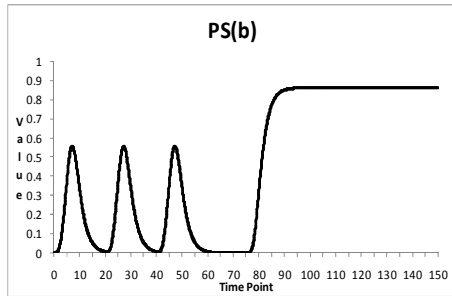


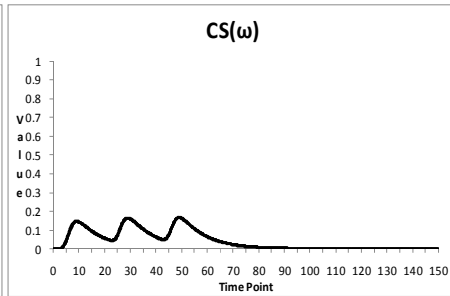
Fig. 2. World State for s1

In the simulations shown, time is on the horizontal axis and the activation level of the state properties is on the vertical axis. The connection strengths between different states were initialized with  $1$  (i.e.,  $\omega_0 = \omega_1 = \omega_2 = \omega_3 = \omega_4 = \omega_5 = 1$ ) and kept fixed throughout the simulation, except the connection strength  $\omega$  which was initialized

with 0 and adapted over time by the Hebbian learning rule given in LP6 in Section 2. Other parameters are set as:  $\Delta t = 0.1$ , learning rate  $\eta = 0.3$ , extinction rate  $\zeta = 0.2$ , speed factor  $\lambda = 0.5$ , steepness  $\sigma = 5$  and threshold  $\tau = 0.5$ . For the initial duration of 45 time units the stimulus s2 occurs three times for 5 time units alternatively, i.e., for the first 5 time units world state for s2 has value 1 and for the next 15 time units value 0, and so on (see Fig. 1). During these 45 time units the world state for s1 was kept 0 (see Fig. 2). This represents the situation in which a person observes somebody else performing some action (or bodily change) and the mirroring function of the preparation neurons makes the person prepare for this action for him or herself. The fluctuation in the activation level of the sensor state is repeating the same pattern between 0.1 to 0.9 as it only depends (via LP1) upon the world state for s2, which also is repetitive.



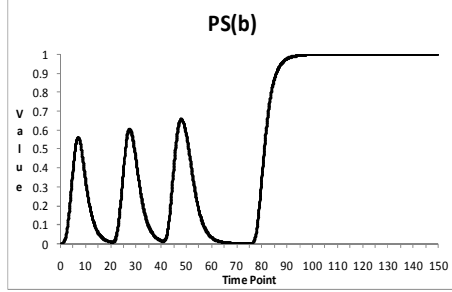
**Fig. 3.** Preparation State for b  
(Normal Functioning)



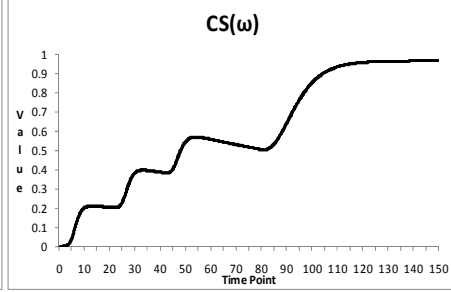
**Fig. 4.** Connection Strength  $\omega$   
(Normal Functioning)

Fig. 3 shows the activation of preparation state PS(b) resulting from the sensory representation pattern. The slight change in the strength of the connection (inverse mirroring) from preparation state PS(b) to sensory representation SRS(s2) (via LP6) shows a similar but slightly delayed fluctuating pattern; see Fig. 4. After that for about 30 time units both world states for s1 and s2 were kept 0, so that the effect of any stimulus on preparation state and execution state becomes zero, as reflected, for example, in Fig. 3. For the rest of the simulation, the world state for s2 is kept 0 while the world state for s1 is becoming 1 (see Figs 1 and 2). As soon as world state for s1 becomes 1, the values of sensor state and its sensory representation increase smoothly and become stable at a high value (of about 0.9). One of the interesting facts (in comparison to the second scenario discussed below) is that, even though a link from preparation state to SRS(s2) develops, it is not strong enough to propagate the effect of SRS(s1) via PS(b) to SRS(s2), see Fig. 4. Hence SRS(s1) only has a positive effect on the activation levels of the preparation state and execution state of b (shown in Fig. 3), and not on SRS(s2). No activation is developed of a mental image SRS(s2) of another person performing action b; this shows normal functioning.

To obtain the deviant behavior of the model, again all parameters were initialized with the same values as used for the normal behavior and mentioned earlier in current section, except the extinction rate which was taken lower this time:  $\zeta = 0.01$ . In this case by Hebbian learning the connection from preparation state of b to sensory representation of s2 achieves a substantially higher connection strength (see Fig. 6) which also induces an upward trend in the fluctuating value of the preparation state for b (see Fig. 5).

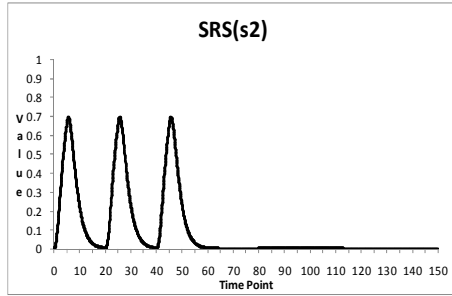


**Fig. 5.** Preparation State for b  
(Deviant Functioning)

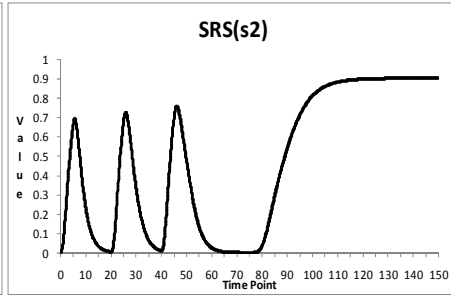


**Fig. 6.** Connection Strength  $\omega$   
(Deviant Functioning)

This achieved connection strength is sufficient to change the impact of stimulus  $s_1$  on  $SRS(s_2)$  (see Fig. 8) compared to the impact shown in Fig. 7. Even in the absence of the world state for  $s_2$ , from time 75 onwards, if the world state for  $s_1$  occurs, it leads to high activation of  $SRS(s_2)$  (see Fig. 8), which shows that the agent develops a mental image of somebody else performing action b. This contrasts the case of normal functioning in which case after time point 75 the level of  $SRS(s_2)$  stays (close to) 0 (see Fig. 7).



**Fig. 7.** Sensory Representation for  $s_2$   
(Normal Functioning)



**Fig. 8.** Sensory Representation for  $s_2$   
(Deviant Functioning)

## 4 Mathematical Analysis of the Model

The equilibrium of the connection strength from preparation state of b to the sensory representation of  $s_2$  may be found by using the equation for  $\Delta\omega$  presented in Section 2 to determine the change in the connection strength after  $\Delta t$  as follows. An equilibrium will occur if  $\Delta\omega = 0$ , so the equation can be rewritten as

$$\gamma[\eta V_1 V_2 (1 - \omega) - \zeta\omega] \Delta t = 0 \quad \text{or} \quad \eta V_1 V_2 (1 - \omega) - \zeta\omega = 0 \quad \text{or} \quad \omega = \frac{1}{1 + \left(\frac{\zeta}{\eta V_1 V_2}\right)}$$

As  $V_1 V_2 \leq 1$ , the following is an upper bound for the equilibrium value:

$$\omega \leq \frac{1}{1 + \left(\frac{\zeta}{\eta}\right)}$$

This expression gives the relation describing the maximum connection strength that may be achieved for given values of learning and extinction rates. It shows that for a smaller value of the extinction rate compared to the learning rate the connection strength will be closer to  $1$ . For the simulation results discussed in Section 3, for the second case (deviant functioning) the learning rate was  $\eta = 0.3$  and the extinction rate  $\zeta = 0.01$  which means the maximum connection strength which may be achieved is  $0.97$ . Indeed it was observed that the connection strength indeed becomes stabilized just below  $0.97$ . Similarly, for the case of normal functioning, the learning rate was  $\eta = 0.3$  and extinction rate  $\zeta = 0.2$ . Indeed the equilibrium was below the  $0.6$  indicated by the above analysis. Note that these are upper bounds resulting from maximal values  $1$  continuously for  $V_1$  and  $V_2$ . In practice this will not continuously happen, so the strength will stay lower.

## 5 Discussion

One of the recent developments in Neuroscience concerns the notion of a mirror system and its functions; e.g., [4], [14], [15], [20]. Cognitive agent models have been designed using this notion as a point of departure, and showing its role in various high-level cognitive and social capabilities such as prediction, imagination, emotion reading, empathic understanding, imitation, and attribution of observed actions; see, for example [5], [13], [18]. Mirroring is a process from an observed action or body state of another person to the person's own preparation states. As discussed in [17] this involves a mental rotation mapping sensory representations of observed actions of other agents onto the representational structures for self-generated actions. This realises a change of perspective from another agent (third-person) to perspective from oneself (first-person). Attribution a self-generated action to another agent proceeds in the opposite direction, realising a change of perspective from oneself (first-person) to another agent (third-person) perspective. This requires inverse mirroring: the representational structures for self-generated actions are mapped onto sensory representations of an observed action of another agent, thus creating a mental image of another agent performing the action. When such a mapping involving reverse mental rotation is made, a self-generated action is perceived as observed from a third person perspective. The agent model presented in this paper addresses the issue of false attribution of self-generated actions using such a mechanism. In addition it is shown how the mechanism can develop based on Hebbian learning [12], [2], [11]. Note that due to the opposite direction, an inverse mirroring process is not covered by cognitive agent models based on mirroring, such as described in [5], [13], [18].

The modelling format used to formally specify the agent model is based on the executable hybrid dynamical modelling language LEADSTO [3]. This hybrid language combines executable temporal logical elements [1] and numerical dynamical system elements [19]. Although the model can also be specified well in a purely numerical format, an advantage of LEADSTO is the possibility to use a logical format to describe state properties.

As a next step a validation study can be conducted to compare the agent model's behaviour to real behaviours. The agent model obtained can be used as a basis for

applications involving realistic, human-like virtual agents in the context of serious or nonserious gaming. In the area of virtual stories it can be used to create virtual characters that react in a less standard but realistic manner. Another possible application is to obtain virtual patients in the area of simulation-based training for psychotherapists.

## References

1. Barringer, H., Fisher, M., Gabbay, D., Owens, R., Reynolds, M.: *The Imperative Future: Principles of Executable Temporal Logic*, John Wiley & Sons (1996)
2. Bi, G.Q., and Poo, M.M.: Synaptic Modifications by Correlated Activity: Hebb's Postulate Revisited. *Ann Rev Neurosci* 24, 139--166 (2001)
3. 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)
4. Bosse, T., Memon, Z.A., and Treur, J.: A Cognitive and Neural Model for Adaptive Emotion Reading by Mirroring Preparation States and Hebbian Learning. *Cognitive Systems Research Journal*, in press (2011)
5. Damasio, A., and Meyer, K.: Behind the looking-glass. *Nature* 454, 167--168 (2008)
6. Farrer, C., and Franck, N.: Self-monitoring in schizophrenia. *Curr Psychiatry Rev* 3, 243--251 (2007)
7. Feinberg, I.: Efference copy and corollary discharge: Implications for thinking and its disorders. *Schizophrenia Bulletin* 4, 636--640 (1978)
8. Fournieret, P., de Vignemont, F., Franck, N., Slachevsky, A., Dubois, B., and Jeannerod, M.: Perception of self-generated action in schizophrenia, *Cogn. Neuropsych.* 7, 139--156 (2002)
9. Frith, C.D., Blakemore, S., & Wolpert, D.: Explaining the symptoms of schizophrenia: Abnormalities in the awareness of action. *Brain Research Rev.* 31, 357--363 (2000)
10. Frith, C.D. (1992). *The cognitive neuropsychology of schizophrenia*. Hove: Lawrence Erlbaum Associates Ltd.
11. Gerstner, W., and Kistler, W.M.: Mathematical formulations of Hebbian learning. *Biol. Cybern.* 87, 404--415 (2002)
12. Hebb, D.O.: *The Organization of Behaviour*. New York: John Wiley & Sons (1949)
13. Hendriks, M., and Treur, J.: Modeling Super Mirroring Functionality in Action Execution, Imagination, Mirroring, and Imitation. In: Pan, J.-S., et al. (eds.), *Computational Collective Intelligence, Proc. ICCCI'10, Part I. LNAI*, vol. 6421, pp. 330--342. Springer (2010)
14. Hurley, S.: The shared circuits model: How control, mirroring, and simulation can enable imitation, deliberation, and mind reading. *Behavioral and Brain Sciences* 31, 1--22 (2008)
15. Iacoboni M.: *Mirroring People: the New Science of How We Connect with Others*. Farrar, Straus & Giroux (2008)
16. Jeannerod, M.: The sense of agency and its disturbances in schizophrenia: a reappraisal. *Exp Brain Res* 192, 527--532 (2009)
17. Jeannerod, M., and Anquetil, T.: Putting oneself in the perspective of the other: A framework for self-other differentiation. *Social Neuroscience* 3, 356-367 (2008)
18. Memon, Z.A., and Treur, J., *Designing Social Agents with Empathic Understanding*. In: Nguyen, N.T., et al. (eds.), *Computational Collective Intelligence, Proc. ICCCI'09, Part 1. LNAI*, vol. 5796, pp. 279--293. Springer Verlag (2009)
19. Port, R.F., Gelder, T. van (eds.): *Mind as Motion: Explorations in the Dynamics of Cognition*. MIT Press (1995)
20. Rizzolatti, G, and Sinigaglia, C.: *Mirrors in the Brain: How Our Minds Share Actions and Emotions*. Oxford Univ. Press (2008)