Conscious Learning Semiotics Systems to Assist Human Persons (CLS^2H)

Gerd Doeben-Henisch
University of Applied Sciences
Department of Intelligent Systems
Nibelungenplatz 1, 60318 Frankfurt am Main
Germany
Email: doeben@fb2.fh-frankfurt.de

Giuseppe Abrami, Marcus Pfaff, Marvin Struwe
University of Applied Sciences
Department of Intelligent Systems
Nibelungenplatz 1, 60318 Frankfurt am Main
Germany
Email: http://www.basys.fh-frankfurt.de/dokuwiki/en:simulator2.0

***Abstract*—Challenged by the growing societal demand for Ambient Assistive Living (AAL) technologies, we are dedicated to develop intelligent technical devices which are able to communicate with human persons in a truly human-like manner. The core of the project is a simulation environment which enables the development of conscious learning semiotic agents which will be able to assist human persons in their daily life. We are reporting first results and future perspectives.**

I. Introduction

Our research context is strongly influenced by the growing societal demand for Ambient Assistive Living (AAL) technologies[1]. One central requirement here is the need for intelligent technical devices which are able to communicate with human persons in a truly human-like manner. This embraces e.g. speech, gestures, texts, pictures, music, and dialogues based on a sufficient understanding of the thinking of the participating human persons. At the time of this writing this is a *vision* guiding our work. But, nevertheless, it is important to include in this vision right from the beginning the challenge to support *real human persons*. Otherwise there is no need for certain strong requirements like language, consciousness, certain kinds of memory, etc.

To reach this goal we are using a general *engineering paradigm*, which allows us to approach the main goal stepwise. Within the general paradigm of systems engineering[2] we work in parallel on conceptual models (theories), on simulations[3], and as well on real world experiments.

A. Conceptual Framework

The *conceptual framework* is relying on the cognitive sciences as far as they are based on observable behavior enhanced with models of phenomenological experience and additional heuristics provided by the biological sciences including brain science. We are *not* interested in *copying* the *structures* of the brain but we are interested to re-engineer those *functions* of the brain which show up as *consciousness* and *consciousness-based behavior*. Everything which is not an experience *within* the consciousness (often called 'qualia' or 'phenomenon') is belonging to a 'machinery' behind the consciousness which is responsible for the emergence of the phenomenal experience. Whether this enabling machinery is realized with some kind of neural structures or with different kinds of rule systems or with special digital or biological circuits or something else does not matter as long as the phenomena of the consciousness show a *dynamic* which is sufficient similar to what is know from human consciousness.

The dimension of *behavior* includes the dimensions of *phylogenetic* development (evolution), of *ontogenetic* growth as well as of *learning*. Although our first models do not yet include phylogenetic and ontogenetic development we have the requirement to include these dimensions in the project during the next years. This is important to allow in the future the development of complex structures without explicit programming.

B. Semiotics

The main challenge why *semiotics* is intentionally used within the project is given by the fact that *communication* with human persons in a human like fashion requires different kinds of *sign processes* which presuppose sign systems (languages). This is traditionally the domain of semiotics. Although the basic ideas of semiotics are as old as philosophy the more modern concepts are connected with names like Peirce (1839 - 1914), Saussure (1857 - 1913), and Morris (1901 - 1979)[4]. Since the times of these pioneers of semiotics the field of semiotics has grown a lot. A straightforward extension of semiotics is *computational semiotics* combining semiotics with computation. Starting in the 1960s in the realm of *control theory* (cf. overview [23]) the applications are widening their scope in the direction of human computer interaction (HCI)[2], text linguistics [23], organizational Semiotics [23], as well as semiotic agents[15]. It was especially Gudwin working with Queiroz who investigated different possibilities to use the sign

[1]Cf. the annual German AAL congress [1] and the AAL web site of the German government

[2]Cf. the paper Erasmus/ Doeben-Henisch, where we propose a new theoretical paradigm for the systems engineering process [19].

[3]The authors of the used simulator software as well of the different experiments are Abrami, Pfaff, and Struwe. A more detailed description of the software and the experiments can be found on the web page of the simulator project.

[4]For a more broader view see Noeth [39], [40], Bouissac [8]

concept of Peirce within computational semiotics (cf. [22], [24], [25]. [26], [27]). Another subject within computational semiotics is labeled *language games*. Strongly influenced by the talking heads experiments from Steels 1995[5] [52], [50], this topic grows very fast (cf. [53], [54], [55], [34], [7], [6]. [41]). Doeben-Henisch has shown, how it is possible to formalize the concept of grounding – which is part of the concept of a language game – within the concept of semiosis of Peirce [16]:124ff.

C. Learning and Memory

Intelligent technical devices which shall assist human persons including full dialogues have to be *learning* semiotic systems. Following the main lines of experimental psychology and ethology we interpret the meaning of the term *learning* as the ability of a system S to change its system function f into some 'improved' system function f^+ depending from the actual environment in a way, that the system response based on f^+ is 'better' then the system answer with the original function f (cf. [9]. The meaning of the term 'better' depends from the special properties of the biological system under investigation as well as from the environmental conditions. These can change. The main requirement for a biological population is to survive, but this allows for variants in the behavior because there can be different strategies how to reach this goal. Furthermore is the 'agent' of a survival primarily not the individual biological system but the *population* as a whole; without a population there is no survival.

There are many concepts available how to generate such an improved function f^+. We have opted for those concepts which are based on the main biological strategies like *genetic algorithms* (GAs)[20], [45], and *learning classifier systems* (LCS)[61], [48] [6]. We have combined the concept of learning classifier systems with with the concept of *built-in emotions* (eLCS). This concept of 'built-in emotions' is not new [15], but received in the last years a better support from findings in neuropsychology, which revealed, that the body itself has several 'built-in' evaluation processes which manage the survival while communicating with the brain[49]. Additionally one can observe today a growing number of examples using built-in values within agents or robots (cf. [30], [31],[32], [33], [58]). We have then extended the classifier concept further to represent a *conscious* agent (ecLCS) (see below). In this case the classifiers are turned into a system with an artificial consciousness additionally supported by a memory with different levels.

D. Consciousness

Although the topic of 'consciousness' has a long tradition in philosophy, the discussion gained a strong new momentum when Chalmers introduced 1995 the labels 'hard' and 'easy' to classify scientific problems with regard to the discussion about the phenomenon of 'consciousness' (cf. [11], [12]). For Chalmers the hard problem exists because we can notify the brute fact that we are *conscious about something*. These kind of data are *not reducible*. To construct some mapping from third person measurements to phenomenal data (first person data) one has to elaborate these phenomenal data explicitly as a structure. The best known methodology today to exploit phenomenological data is demonstrated by *phenomenology* (cf. [12]:413f, [60]). To construct functions based on the neuronal data – like e.g. the global workspace theory of Baars and others (cf. [3], [4], [5]) – will not suffice to explain the hard problem according to Chalmers.

Knowing this background and accepting the behavioral sciences Doeben-Henisch has worked out a conceptual framework which is intended to include the hard problem. In [15], [16] we can distinguish at least three fundamental kinds of data which are not reducible to each other: Behavioral data (SR), neurophysiological data (NN), as well as phenomenological data (PH). Each of these data sets allows for formal model (theory) building T_{SR}, T_{NN}, and T_{PH}. Done in the right way one can then map the structures of these models onto each other. This is *not* to misunderstand as a kind of reductionism, rather it can show whether there could be some *structural coupling* between structures. Thus if one assumes that events within the phenomenal $consciousness_{PH}$ are 'caused' by certain neural events, then it would be important to identify within a formal neural theory T_{NN} those parts which can be 'correlated' with phenomenal events and which thereby could be labeled *neural correlates of consciousness* as $consciousness_{NN}$.

There is still a very broad ongoing discussion about the relationship between consciousness and neural events (Cf. e.g. [29]). While e.g. Revonsuo describes the relationship between the consciousness and the brain in a way which supports a reduction of the consciousness to physiological states ('naturalizing phenomenology')[46], does MacLennan argue against reductionism[7]. But even MacLennan thinks he has to assume 'below' the level of phenomena some 'protophenomena'[8] which can not be perceived 'consciously' but which can be 'postulated' to allow a mapping onto elementary neural events. From the theoretical point of a formal mapping this is *not* necessary. It is even implausible to do this because it could be – and in most cases this seems to be highly probable – that a phenomenon which appears as a phenomenological 'simple' phenomenon can correspond to highly complex neural processes distributed over many areas in the brain showing complex timing patterns too. Thus if we assume that the evolutionary progress has to be identified with this kind of 'abstraction' mapping complex neural processes into 'simple'

[5]Steels states in [57] that his own experiments are strongly influenced from the experimental work with the robot Sharkey 1966 - 1972 under the guidance of Nilsson (cf. [38])

[6]Because classifiers can formally be 'rewritten' as fuzzy-rules or even neurons we assume that the decision for classifiers does not 'limit' the scope of this approach. In a future phase we want to replace the classifiers explicitly by fuzzy-rules, neurons, as well as simple automata and then compare the behavior of these different implementations.

[7]"to attempt to reduce first-person phenomena to the third-person objects and properties" is a "category mistake"[35]:437. This criticism has been raised by other authors too, e.g. Nagel 1986 [37].

[8]"...protophenomena, which are theoretical entities hypothesized as the elementary constituents of phenomena (conscious experiences)". [35]:438

phenomenological phenomena then the assumption of a reduction would be misleading; it is not a mere one-to-one mapping.

Furthermore if we distinguish formally between a phenomenological theory T_{PH} and a neural theory T_{NN} then we can replace the neural theory T_{NN} by any other theory T_X which 'does the job'. Thus if we can construct a formal model T_{SW} of a software agent using e.g. classifiers which are different from neurons and this model generates the 'same' phenomena as a neural theory T_{NN}, then this will not change our theory of the consciousness as part of a learning semiotic system. It only changes the concrete machinery 'behind' the consciousness$_{PH}$ which is responsible for the computation. Within the paradigm of conscious learning semiotic systems (CLSS) we will therefore distinguish only two parts: the *conscious* part and the *non-conscious* part. The latter is responsible for the needed computations.

E. Measurement

The last main requirement for our project is *measurement*. Within the engineering context we are obliqued to follow an engineering process model[9]. Minimally this includes the following phases: introduction of a *problem* (P); transformation of the problem during requirements engineering into an appropriate *behavior model* M_{SR}; creation of a *design model* M_D during synthesis; *verification* of the design model against the behavior model; transformation of the design model into an *implemented model* M_I; *validation* of the implemented model against the behavior model:

$$\begin{aligned} requengineering &: P \longmapsto M_{SR} & (1)\\ synthesis &: M_{SR} \longmapsto M_D & (2)\\ verification &: M_{SR} \times M_D \longmapsto V1 & (3)\\ implementation &: M_D \longmapsto M_I & (4)\\ validation &: M_{SR} \times M_I \longmapsto V2 & (5)\end{aligned}$$

Today in most cases the behavior model M_{SR} belongs to a class of simulation models $M_{SR} \in SIM$. A simulation is a sequence of states generated by operators typical for a certain model. The set of all possible simulations of a certain behavior model M_{SR} would then represent the space of all possible behaviors of this model, written as $[M_{SR}]$.

If we can assume that validation of a certain implemented model $\mathbf{m}_I$ compared to a behavior model $\mathbf{m}_{SR}$ is 'true' with regard to the agreed criteria then one has to assume that all those properties which one wants to be 'part' of the 'intended behavior' $[\mathbf{m}_{SR}]$ have to be 'part' of this set of all possible behaviors[10].

[9]Although this minimal model is written as a sequence of phases in practice there are several iterations possible between different points of this process. Cf. our more elaborated and modified proposal for a theory of the Systems Engineering Process [19]

[10]This is the same paradigm which Turing used in his famous paper how to measure the intelligence of a computer [59]. Instead of giving a detailed description of some structural properties of the machine he took a 'reference object' for intelligence – the observable behavior of a human person – and required the facilitator of the test so compare the behavior of the computer with the behavior of a human person.

Another interesting case is the *evaluation* of an implemented system $\mathbf{m}_I$ embedded in a target environment E which is also part of the behavior model M_{SR}.

$$evaluation \quad : \quad E \times M_I \longmapsto V \qquad (6)$$

In this case one can e.g. measure parameters like *success* in finding some 'food' to gain energy, the *number of moves* or the *energy consumed* between the necessity to find food and the 'success' to find it.

This kind of evaluation can also include complex *psychological intelligence tests* and it can be used to compare different systems with regard to the same *test environment* E. Furthermore it would be possible to compare the behavior of an implemented system $\mathbf{m}_I$ with the behavior of human persons either directly or mediated by a simulation environment within which the user acts as an *avatar*.

II. Outline of a CLS^2H-Theory

As explained in the introduction we are developing a theory, a simulation framework and real applications for **C**onscious **L**earning **S**emiotic **S**ystems **for H**umans, abbreviated as CLS^2H.

A. Simulation Framework (SF)

The simulation framework for CLS^2H is given by at least one *environment (E)* and at least one CLS^2H, further abbreviated as 'A' for agent[11]. We start with the simple case of one 2D-environment $e \in E$ and one agent $a \in A$.

$$\begin{aligned} SF &\subseteq E \times A & (7)\\ E &\subseteq POS \times 2^{PROP} & (8)\\ POS &\subseteq REAL^2 & (9)\\ ainp &: E \times POS \times 2^{PROP} \longmapsto \Sigma^* & (10)\\ aout &: A \times \Xi^* \longmapsto E \times POS \times 2^{PROP} & (11)\end{aligned}$$

Thus the 2D-environment E has positions and properties related to these. There is a mapping from a an environment E and some position with it's properties into a finite string $\sigma \in \Sigma^*$ which is the possible input which an agent A can receive from the world E. Vice versa can an agent A send an output string $\xi \in \Xi^*$ through the mapping function aout into the environment E thereby generating some change in the Environment E at position POS. From a purely *behavioral* point of view this is all information which can be observed.

As pointed out above a *simulation* $s \in S$ can be understood as a sequence of states $\langle s_1, s_2, \cdots \rangle$ with $s_i \in s$. Every state s_i contains the elements $(e_{jk}, \langle \sigma, \xi \rangle)$ with $e_{j.k}$ as the k-th state of environment $e_j \in E$. Thus for every step in the simulation one has the input and output of the agent which can be translated according to the mapping functions 'ainp()' and 'aout()'.

[11]The concept of 'agent' has no clearcut definition in the community. In our theory the term 'agent' is part of a formal structure which as a whole provides the meaning space for this term. We start with a very general concept for an agent which then will be enhanced stepwise with additional properties.

An extension of the simulation concept is the inclusion of some conscious parameter of the agent. If one assumes some internal *drives/ emotions* $\delta \in PH_A$ with PH_A as the set of possible phenomena of the consciousness of agent A then one could extend the concept of the simulation as follows:

$$s \in S^n \quad (12)$$
$$s_i \in E \times (\Sigma^* \times \Xi^* \times \Delta^d) \quad (13)$$

With these extensions one can view the observable behavior of the agent in connection with certain internal parameter δ of the agent. The length 'n' of a simulation s can be *infinite*. If one assumes a *death*-criterion for agents – e.g. energy level below some threshold θ – then one can define a *halt criterion* to stop the simulation. It would then be possible to measure *how long* an agent can survive in a certain environment. Assuming different meta parameter $\pi^i \in \Pi$ defined over a simulation s would allow to measure different kinds of behavior properties (see below the description of different experiments).

B. Agents

The description of an agent A for the simulation framework of a CLS^2H-Theory does not yet deal explicitly with 'signs', 'languages' or 'communication'. Only a first general framework is presented to enable later the more advanced features (see below upcoming experiments). The *learning* is realized through the usage of a classifier system which has been transformed into a *memory* system which provides a behavior profile which can be identified with a socalled *state based episodic* memory (cf. the recent review by [14]). A more advanced *scene-based* episodic memory will be realized in the upcoming experiments with the aid of additional memory levels[12]. The term *consciousness* is introduced from the beginning following the motivation by everyday experience, phenomenology, and a Peircian-like semiotics (see the text above).

The agent A has a dual structure consisting of a *consciousness* C and the *non-consciousness* NC. For the semiotic dimension the consciousness will be of primary importance. The *non-conscious* part operates as 'generator' of the consciousness. From this point of view it does not matter, how exactly the NC 'works' or how it will be realized (neurons, classifier, rules, automata,....or some kind of special memory-mechanisms.)[13]. For a first minimal architecture of an CLS^2-agent A we opted for the following schema:

$$A(x) \quad iff \quad \langle C, \Sigma^*, \Xi^*, \Delta^n, \mu, \gamma \rangle \quad (14)$$
$$C \subseteq \Sigma^* \times \Delta^n \times \mu \times \Xi^* \quad (15)$$

[12]The agent A_2 has already a second level within its memory but the details will be described in another paper

[13]There are other strategies like e.g. da Silva and Gudwin [13] They define the consciousness and the unconsciousness primarily from an engineering point of view with the aid of a technical structure called 'global workspace' including lots of technical details. It remains an open question in the whole paper, why one should use the term 'consciousness' at all.

An agent A interacts with its environment with the aid of input strings Σ^* ('sensor input') and output strings Ξ^* ('actions'). The important internal states ('drives', 'emotions', ...) are represented in Δ^n. The memory is represented by μ. The overall behavior is organized by the system function γ. The *consciousness* C is an actual representation of input, selected system states, selected memory contents as well as the responding output.

$$\gamma \quad : \quad \Sigma^* \longmapsto \Xi^* \quad (16)$$
$$\gamma \quad = \quad xgen \oplus (store \vee find) \oplus perc \quad (17)$$
$$perc \quad : \quad \Sigma * \times \Delta^* \times \mu \longmapsto \Delta^* \quad (18)$$
$$store \quad : \quad \Sigma * \times \Delta^* \times \mu \longmapsto \mu \quad (19)$$
$$find \quad : \quad \Sigma * \times \Delta^* \times \mu \longmapsto \{\Xi^*\} \quad (20)$$
$$xgen \quad : \quad \{\Xi^*\} \longmapsto \Xi^* \quad (21)$$

To simplify the theory it is assumed here that the *consciousness* C is always representing the actual states of input, system states, memory outputs as well as responding actions. In a more elaborated version one has to specify this mapping in more detail[14]. The γ-function represents the system function of the agent. Generally it maps incoming sensory events into an output action. In more detail the system function is composed of several contributing functions. The *perception* function $perc$ maps the actual input into actual system states thereby exploiting information from the old system states and the memory. Input can be 'new' or 'known' depending from the available memory contents. *New* input will be *stored* in the memory; otherwise input will e.g. reinforce the available memory contents. If the agent is in some deprived state represented in its system states then the $find$ function can check the memory for known states where this kind of actual deprivation could be 'solved'. If the $find$ function is successful then the *action generation function* $xgen$ selects from the possible alternatives $\{Xi^*\}$ one concrete action $\xi \in \{Xi^*\}$.

The overall behavior of the first simple agent types can be characterized as either being in the *play*-mode – if all other system states are below some threshold θ – or driven by those system states which are above their thresholds. In play-mode the agent – depending from it's type (see below) – will behave randomly or according to some predefined reactions or depending from it's memory μ. During play-mode can the agents with memory extend this memory. If at least one drive is 'on' then the agent has to complete at least one critical goal. While the agents without memory can only continue to react randomly or according to some fixed reactive patterns the agents with memory can start to use the memory to *find* some past experience matching the actual situation[15](cf. for

[14]At this point of our theory development we made no final decision about the details of the mapping between the consciousness and the non-conscious parts of the system. Actually the consciousness is only a selection of important parameters to monitor the system. The final goal is to solve the 'hard' problem for a system with an artificial consciousness.

[15]Here the memory functions like a classifier system being optimized by actions instead of genetic operators; one can therefore call a memory a *memetic* machine

the term 'memetic' [42]).

As mentioned above the memory μ is a modified classifier system: using the *if*-part of a classifier as *state* and the *then*-part as *action* we have transformed the classifiers into a *graph* which represents perceivable situations and the possible actions leading from one situation to another. We classified this primary graph resulting from a classifier system as *level-1* memory which can be extended by n-many others levels operating on level 1[16]. The genetic operators usually used to optimize the classifiers are in this context replaced by the actions of the agents which induce changes in the memory graph and thereby replace the cross-over and the mutation operator.[17],[18]

III. First Experiments

According to our engineering 'philosophy' we start with most simple agents and improve these stepwise.

A. Three Experiments

According to our engineering 'philosophy' we start with most simple agents and improve these stepwise.

B. Types of Agents

1) Random Agent A_0: The first agent is the agent A_0 which has been introduced as the primary 'benchmark' agent. This agents has a simple input string σ showing the area 'immediately' before him[19]. But the agent A_0 does not exploit his input. He acts completely at *random*. The agent has as system states $\Delta = \langle 'doplay', 'behungry' \rangle$. But he can only consume the food-object if he covers the cell with the food-object by chance. If the energy level reaches '0' the agent will 'die'.

2) Reactive Agent A_1: The second type of agent A_1 equals agent A_0 with the exception, that for every perception σ this agent has a pre-defined response what to do ('reactive' agent). If agent A_1 perceives a food-object with his system state $behungry$ 'active' then he will move onto that cell and then consume the food, otherwise he will look for a 'free cell' before him do some move onto it.

3) Memory-Based Agent A_2: The third agent A_2 equals agent A_1 with the exception that this agent has a simple memory μ consisting of two levels. This memory can in level-1 *store* the content of the consciousness $\langle \sigma, \delta \rangle$ as a 'node' of a graph and the action ξ based on this conscious content as an 'edge' in the graph. Simultaneously level-2 builds a simple spatial model of the grid with all the encountered objects based on the data of level-1. If agent A_2 switches into the state 'being hungry' he can look up in his memory for some node which contains 'food' correlated with a lowering of the 'intensity' of the system state 'being hungry'. The agent can then try to construct a 'path' based on his memory content. The realization of this 'planned' path is still very simple. In case of 'obstacles' he 'unlocks' his plan and tries to find some way randomly. The memory includes additionally some more parameter like the *frequency* κ as well as some *forgetting* parameter ϕ. If these parameters are set to '-1' then they are 'inactive'. In the *play*-mode the memory can be extended randomly with new nodes and edges[20].

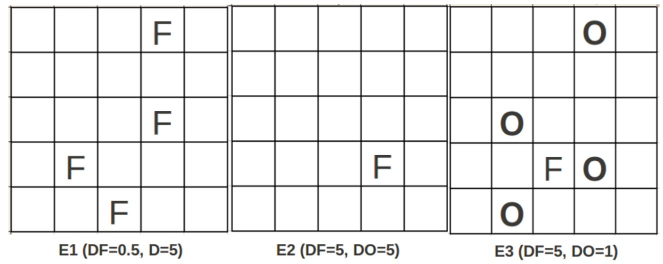

Fig. 1. Environments E1 - D3 with Food-Distance FD and Objekt-Distance OD

C. Three Experiments

We have run[21] the following experiments.

For a first test we have prepared manually three simple environments $E1, E2, E3$ (cf. figure 1) as grids with 5×5-many cells containing either 'food' (F), 'obstacles' (O) or being 'free space' (_). The agents started always in the 'upper lower corner' of the grid. For every environment we have computed the average *food distance* DF as well as the average *object distance* DO.

The working hypothesis for environment $E1$ has been, that the high density of food objects makes a survival for an agent such an easy task, that the different capabilities should not make a great difference between the three agents A_0, A_1, A_2. And – as one can see in the table of results table I – during 10 experiments with 100 moves each the agents A_1, A_2 are only 1.9 or 1.8 'better' than agent A_0 with regard to find food. While agent A_1 needs 1.4 times more moves than agent A_2, agent A_0 needs 5.9 times more moves than A_2.

The working hypothesis for environment $E2$ predicted, that the low density of food objects makes a survival for an agent searching only randomly a difficult task, whereby the improved capabilities of the agents A_1, A_2 should suffice to show a clear

[16] In the agent A_2 used in the reported experiments there is already a leve-2 memory level implemented which supports the agent with an adaptive model of the surrounding space.

[17] The structure of the memory μ seems to have some similarity with the state - action - state (SAS) structure mentioned in [36]. In our case was the decision for such a memory-structure motivated by the question, which information does an agent need to be able to orientate himself in an environment based on the information from it's memory.

[18] There is another point to clarifify for the future: as Perlosky has shown in many publications there is an interesting bottom-up and top-down mechanism called *MFT-DL* theory, which works quite stable in many different practical applications (cf. [44]). This MFT-DL model has to be compared with the concept of the multi-level memory planned for our agents.

[19] Actually those 3 cells in a row.

[20] We have been asked, why the A_2 has not been labeled as 'purpose-based' agent? We decided not to use the term 'purpose' here because this term has strong associations to many different meanings in psychology and philosophy – which perhaps can be discussed further in the future –, but here we wanted to focus on the fact that it is a certain structure of the memory which enables a 'usage of the memory to support a concrete problem'.

[21] until April-15, 2011

TABLE I
RESULTS OF EXPERIMENTS E1 – E3

	SUCCESS			MOVES			ENERGY		
	A0	**A1**	**A2**	**A0**	**A1**	**A2**	**A0**	**A1**	**A2**
E1	1	1.9	1.8	5.9	1.4	1	4.6	1.4	1
E2	1	6	8	4.4	3.1	1	3.5	3.1	1
E3	1	3.5	4.6	5.2	1.3	1	4.3	1.4	1

better performance. The table of results (cf. table I) shows indeed that the agents A_1, A_2 are 6 to 8 times 'better' than agent A_0 with regard to find food. Agent A_2 is only slightly better than A_1. Again do the agents A_0, A_1 need 3.1 and 4.4 times more moves than agent A_2, and 3.1 and 3.5 times more energy than A_2. That the difference between A_1 and A_2 is not greater is due to the fact that the size of the whole grid is not large enough to reveal the shortage of success when missing a memory.

For environment $E3$ with an object distance $DO = 1$ and a food distance $DF = 5$ the working hypothesis says that it should be difficult for agent A_0 to find the hidden food and even for agent A_1 with the improved search capability it should be more difficult than for agent A_2 to find the same hidden place again. And this is what the table of results shows, especially agent A_1 is weaker than agent A_2.

IV. NEXT EXPERIMENTS

Because our primary goal is the enabling of communication we will continue with the following experiments[22]:

- Extend the varieties of environments with regard to size and object – food distributions.
- Increase the number of drives beyond play-mode and being hungry.
- Extend the memory μ with further additional levels for meta-objects which can function as 'abstractions' for objects of level-1
- Improve the handling of plans when applied to the real situation.
- Differentiate the content of the input string σ to include *signs* and *non-signs* to allow for *symbol grounding* experiments as well as for *language games*[23]
- Increase the number of participating agents above 1. Technically the simulation framework does allow for more agents but the agents $A_0, ..., A_2$ have not yet the ability to deal with other agents.
- Provide an interface for human persons that they can act in the environment like virtual agents. This would allow direct comparisons between human as avatars and artificial agents in different tests. Furthermore the human agents can function as 'trainers' in dedicated experiments.

V. DISCUSSION

A methodological framework for conscious learning semiotic agents has been introduced. A simulation framework as well as first simple agent types have been presented. First experiments demonstrated the working of the concepts and show by dedicated experiments how the different architectures of the different agents can be distinguished by distinguished parameters like 'success', 'number of moves' as well 'consumed energy'. The memory-system as an architecture supporting genetic learning on a 'higher level' ('memetic') and which allows for extensions to more advanced learning tasks. The next experiments will include simple language learning with symbol grounding. A still open question are the details of the interaction between the conscious and the unconscious part of the agent. Several experiments will be conducted. As primarily 'guide' for the experiments we will use the constraints induced by the necessity of language learning and language communication (discourse).

[22] starting during summer 2011

[23] Here we will cooperate with colleague Alexander Mehler (Johann-Wolfgang Goethe University, Frankfurt am Main, Germany)

REFERENCES

[1] AAL congress, URL: http://www.aal-kongress.de/; AAL German Government initiative URL: http://www.aal-deutschland.de/

[2] Andersen, P.B. *A Theory of Computer Semiotics. Semiotic approaches to construction and assessment of computer systems*, Cambridge: Cambridge University Press, 1990

[3] Baars, J.B. *How does a serial, integrated and very limited stream of consciousness emerge from a nervous system that is mostly unconscious, distributed, parallel and of enormous capacity* , Ciba Foundations Symposium 174, Experimental and Theoretical Studies of Consciousness, Chichester - New York - Brisbane et.al: John Wiley & Sons, 1993, pp.282–290, Panel Discussion pp.292-303

[4] Baars, J.B. *Understanding Subjectivity: Global Workspace Theory and the Resurrection of the Observing Self*, Journal of Consciousness Studies, Vol.3, No.3, 1996, pp.211-216

[5] Baars, J.B.; Franklin, S. *An architectural model of consciousness und unconscious brain functions: Global Workspace Theory and IDA*, Neural Networks, 20 (2007), pp.955–961 (Elsevier Ltd.)

[6] Baronchelli, A. *The Minimal Naming Game: A Complex Systems Approach*, in: Steels, L. (Ed.) *Experiments in Language Evolution*, Amsterdam: John Benhamins Publ.Co., Chapt.12, 2010

[7] Belpaeme, T.; Cowley, S. J.; MacDorman, K. (Eds.) *Symbol grounding*, Amsterdam, The Netherlands: John Benjamins, 2009

[8] Bouissac; P. (ed.) *Encyclopedia of Semiotics*, New York – Oxford. Oxford University Press, 1998

[9] Bower, G.H; Hilgard E.R. *Theories of learning*, 5th. ed., Prentice-Hall Inc., 1981 (with a german edition: *Theorien des Lernens, Vol.1+2*, 5th. rev.ed., Stuttgart (Germany): Klett-Cotta, 1983, translated by Aebli, H.; Aeschbache , 1983, 1984)

[10] Cangelosi, A. *The Grounding and Sharing of Symbols*, Pragmatics & Cognition, 14(2): 275 - 285, 2006

[11] Chalmers, D.J. *Facing Up to the Problem of Consciousness*, In: Journal of Consciousness Studies, 2, No.3 (1995), pp.200 – 219

[12] Chalmers, D.J. *Moving Forward on the Problem of Consciousness*, In: Journal of Consciousness Studies, 4, No.1 (1997), pp.3 – 46, repr. in Shear, J. (Ed.) *Explaining Consciousness – The 'Hard Problem'*, Cambridge (MA) – London: The MIT Press, 1997

[13] Silva, R.C.M. da; Gudwin, R.R.*Developing a Consciousness- Mind for an Artificial Creature*, Lecture Notes in Computer Science, 2011, Vol. 6404, Advances in Artificial Intelligence– SBIA 2010, Pages 122-132, 20th SBIA - Brazilian Symposium on Artificial Intelligence - October 23-28, 2010 - São Bernardo do Campo, SP - Brazil

[14] Castro, E.C.de; Gudwin, R.R. *An Episodic Memory Implementation for a Virtual Creature – Studies in Computational Intelligence*, 2010, Vol. 314, Model-Based Reasoning in Science and Technology, pp. 393-406, In: L.Magnani et al. (Eds.) *Model-Based Reasoning in Science & Technology*, Berlin - Heidelbeerg: Springer Verlag, 2010

[15] Doeben-Henisch, G.; *The BLINDs WORLD I. Ein philosopisches Experiment auf dem Weg zum digitalen Bewusstsein. A Philosophical Experiment on the Way to Digital Consciousness*, In: K.Gerbel/ P.Weibel (eds.), Mythos Information. Welcome to the wired world. @rs electronica 95, Springer-Verlag, Wien, pp.227-244, 1995 (This is a German-English text).

[16] Doeben-Henisch, G. *Reconstructing Human Intelligence within Computational Semiotics. An Introductory Essay.* , In: Loula, A.; Gudwin, J.; Queiroz, J. (Eds). *Artificial Cognition Systems*, Hershey (PA): Idea Group Inc.,pp.106-139, 2007

[17] Doeben-Henisch, G. *Humanlike Computational Learning Theory. A Computational Semiotics Perspctive*, In: Proceedings IEEE Africon2009 Conference, 23-25 Sept. 2009, Nairobi (Kenia)

[18] Doeben-Henisch, G.; Wagner, M. *Validation within Safety Critical Systems Engineering from a Computational Semiotics Point of View*, In: IEEE Africon2007 Intern.Conference, Windhoek (Namibia), Sept.2007

[19] Erasmus, L.D.; Doeben-Henisch, G. *A Theory for the Systems Engineering Process*, in: IEEE Africon2011 Intern.Conference, Livingstone (Zambia) Sept.201, delivered (see also the accompanying website http://www.os-pe.org).

[20] Goldberg, D.E. *Genetic Algorithms in Search, Optimization & Machine Learning*, Reading (MA): Addison-Wesley Publ.Company, Inc., 1989

[21] Gudwin, R. *Evaluating intelligence: a computational semiotics perspective*, Systems, Man, and Cybernetics, 2000 IEEE International Conference on, pp. 2080 - 2085 vol.3, 2000

[22] Gudwin, R.; Queiroz J.; *On a computational model of the Peircean semiosis*, Integration of Knowledge Intensive Multi-Agent Systems, 2003. International Conference on, pp. 703 - 708, 2003

[23] Gudwin, R.; Queiroz J.; *Towards an introduction to computational semiotics*, Integration of Knowledge Intensive Multi-Agent Systems, 2005. International Conference on, pp. 393 - 398, 2005

[24] Gudwin, R.; Queiroz J.; *Meaningful agents: a semiotic approach*, Integration of Knowledge Intensive Multi-Agent Systems, 2005. International Conference on, pp. 399 - 404, 2005

[25] Gudwin, R.; Queiroz J.; *The emergence of symbol-based communication in a complex system of artificial creatures*, Integration of Knowledge Intensive Multi-Agent Systems, 2005. International Conference on, pp. 279 - 284, 2005

[26] Gudwin, R.; Queiroz J.; *Towards Machine Understanding: Some Considerations Regarding Mathematical Semiosis*, Integration of Knowledge Intensive Multi-Agent Systems, 2007. KIMAS 2007. International Conference on, pp. 247 - 252, 2007

[27] Gudwin, R.; Queiroz J.; (Eds.) *Semiotics and Intelligents Systems Development*, Hershey (PA): Idea Group Inc. 2007

[28] Hayashi, E.; Yamasaki, T.; Kuroki, K.; *Autonomous behavior system combing motivation with consciousness using dopamine*, Computational Intelligence in Robotics and Automation (CIRA), 2009 IEEE International Symposium on, 2009 , pp. 126 - 131

[29] JCS, Journal of consciousness studies, URL: http://www.imprint.co.uk/jcs.htmlq

[30] Kitamura, T.; Otsuka, Y.; Nakao, T.; *Imitation of animal behavior with use of a model of consciousness-behavior relation for a small robot*, Robot and Human Communication, 1995. RO-MAN'95 TOKYO, Proceedings., 4th IEEE International Workshop on, 1995 , pp. 313 - 317

[31] Kitamura, T. *Can a robot's adaptive behavior be animal-like without a learning algorithm?*, Systems, Man, and Cybernetics, 1999. IEEE SMC '99 Conference Proceedings. 1999 IEEE International Conference on, 1999 , pp. 1047 - 1051 vol.2

[32] Kitamura, T.; Otsubo, J.; Abe, M.*Emotional Intelligence for Linking Symbolic Behaviors*, Proceedings of the 2002 IEEE International Conference on Robotics 8 Automation, Washington, DC, May 2002, pp.1001-1006

[33] KUBOTA, N.; KOJIMA, F.; FUKUDA, T. *Self-Consciousness and Emotion for A Pet Robot with Structured Intelligence*, IEEE 2001, pp.2786 - 2791

[34] Loula, A.; Gudwin, J.; Queiroz, J. (Eds). *Artificial Cognition Systems*, Hershey (PA): Idea Group Inc. 2007

[35] MacLennan, B.J. *Consciousness in robots: the hard problem and some less hard problems*, Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on, 2005, pp. 434 - 439

[36] Madden,C.J.; Lallée,S.; Dominey, P.F.; *Commentary on Solutions an Open Challenges for the Symbol Grounding Problem*, Intern. J. of Signs and Semiotic Systems, 1(1), pp. 59-60, 2011

[37] Nagel, T., *The View from Nowhere*, New York - Oxford: Oxford University Press, 1986

[38] Nilsson, N. J. *Shakey The Robot*, Tech. Rep. 323, Menlo Park, CA: AI Center, SRI International, Retrieved from http://www.ai.sri.com/shakey/

[39] Noeth, W., *Handbook of Semiotics*, Bloomington, Indianapolis: Indiana Univ. Press (=Advances in Semiotics), xii + 576 pp, 1990 (Transl. from the German edition *Handbuch der Semiotik*, Stuttgart: J.B.Metzler, xxi + 560 pp, 1985)

[40] Noeth, W., *Handbuch der Semiotik*, 2. vollst. neu bearb. und erw. Aufl. mit 89 Abb. Stuttgart/Weimar: J.B. Metzler, xii + 668pp, 2000

[41] Nolfi,N.; Mirolli,M. (Eds.) *Evolution of Communication and Language in Embodied Agents*, Berlin - Heidelberg: Springer Verlag, 2010

[42] Ong, Y-S.; Lim, M.; Chen, X.; *Memetic Computation-Past, Present & Future*, Computational Intelligence Magazine, IEEE 5(2), 2010, pp.24 - 31

[43] Peirce, C. S., *Collected Papers of Charles Sanders Peirce. Vols. 1-8*, ed. by Charles Hartshorne; Pauls Weiss, Vols.7-8. ed. by Arthur W.Burks. Cambridge (MA): Harvard University Press. Citation with (CP, n.m) (n := Volume, m := Paragraph), 1931-58

[44] Perlovsky, L.I. *Neural Mechanisms of the Mind, Aristotle, Zadeh, and fMRI*, IEEE TRANSACTIONS ON NEURAL NETWORKS, 21(5), pp.718 - 733, 2010

[45] Rawlins, G.J.E. (ed.) *Foundations of Genetic Algorithms 1991 (FOGA 1) (v. 1)*, Publisher: Morgan Kaufmann; 1st edition (July 15, 1991), ISBN-10: 1558601708, ISBN-13: 978-1558601703

[46] Revonsuo, A. *Prospect for a Scientific Research Program for Consciousness*, In: Thomas Metzinger (Ed.), *Neural Correlates of Consciousness. Empirical and Conceptual Questions*, Cambridge - London: The MIT Press. pp.57-75, 2000

[47] Shear, J.*Introduction*, In: Shear, J. (Ed.) *Explaining Consciousness – The 'Hard Problem'*, Cambridge (MA) – London: The MIT Press, 1997, pp.1–6

[48] Sigaud, O.; Wilson, S.W., *Learning classifier systems: A survey*, Soft Computing, vol. 11, no. 11, September, 2007, pp. 1065-1078

[49] Solms, M.; Turnbull, O.; *The Brain and the Inner World: An Introduction to the Neuroscience of the Subjective Experience*, New York: Other Press, 2002

[50] Steels, L. *Cooperation between distributed agents through self-organisation*, Intelligent Robots and Systems '90. 'Towards a New Frontier of Applications', Proceedings. IROS '90. IEEE International Workshop on, pp. 8 - 14 supl, 1990

[51] Steels, L. *Mathematical analysis of behavior systems*, From Perception to Action Conference, 1994., Proceedings, pp. 88 - 95, 1994

[52] Steels, L. *A self-organizing spatial vocabulary*, Artificial Life Journal, 2(3), pp. 319-332, 1995

[53] Steels, L. *Language Games for Autonomous Robots*, IEEE Intelligent Systems , Volume 16 Issue 5, pp.16 – 22, 2001

[54] Steels, L. *Evolving grounded communication for robots*, Trends in Cognitive Science 7(7), July 2003, pp.308 – 312

[55] Steels, L. *Semiotic Dynamics for Embodied Agents*, Intelligent Systems, IEEE, Volume: 21 , Issue: 3, pp. 32 - 38, 2006

[56] Steels, L. *Steels, L. (The Symbol Grounding Problem has been solved. So What's Next?*, In Glenberg, A.; Graesser, A.; Vega, M. de; (Eds.), *Symbols, Embodiment, and Meaning*, Oxford University Press, pp. 506-557, 2008

[57] Steels, L.*Work on Symbol Grounding now needs concrete Experimentation*, Intern. J. of Signs and Semiotic Systems, 1(1), pp. 57-58, 2011

[58] Takeno, J.; Akimoto, S.; *A conscious robot to perceive the unknown*, Systems Man and Cybernetics (SMC), 2010 IEEE International Conference on, 2010, pp.1375 - 1379

[59] Turing, A.M. *Computing Machinery and Intelligence*, Mind 59, pp.433 - 460, 1950

[60] Varela, F.J. *Neurophenomenology. A Methodologial Remedy for the Hard Problem*, Journal of Consciousness Studies, Vol.3, No.4, 1996, pp.330–349

[61] Wilson, S.W. *ZCS: a zeroth level classifier system*, Evolutionary Computation, 2(1), 1-18 (1994)