

# Dynamical Systems Account for Meta-Level Cognition

Michail Maniadakis and Jun Tani

Laboratory for Behavior and Dynamic Cognition,  
Brain Science Institute (BSI), RIKEN, Japan  
{mmaniada, tani}@brain.riken.jp

**Abstract.** The current paper studies possible neuronal mechanisms for meta-level cognition of rule switching. In contrast to the conventional approach of hand-designing the cognitive functions, our study employs evolutionary processes to search for neuronal mechanisms accounting for meta-level cognitive functions required in the investigated robotic tasks. Our repeated simulation experiments showed that the different rules are embedded in separate self-organized attractors, while rule switching is enabled by the transitions among attractors. Furthermore, the results showed that although certain segregation between the lower sensory-motor level and the higher cognitive level enhance the task performance, meta-level cognition is significantly supported by the embodiment and the lower level sensory-motor properties.

## 1 Introduction

Both animals and humans, use meta-level cognition in order to adapt their action strategies to unpredictable environment situations. Three decades ago, Gregory Bateson [1] demonstrated this ability in dolphins that learned both behavior scheme rules and additionally how to manipulate them in a meta-level, considering that rules have to repeatedly switch in order for the overall behavior to be successful. He argued that dolphins were manipulating rules at a meta-level, but how humans and other animals acquire a meta-rules, has yet to be explained. What are the neural mechanisms that enable meta-level phenomena?

To better understand how animals and humans execute rule switching, neuroscientists have relied on tests like the Wisconsin Card Sorting (WCS). The WCS task creates a situation in which the subject needs to have meta-level awareness of the rules to accomplish card sorting. A large number of experiments with humans and monkeys have shown that impairments to the prefrontal lobe adversely affect performance in WCS tests, suggesting that this region is important for meta-level cognitive behaviors.

In order to explain rule switching in meta-level, modelling researchers describe possible brain mechanisms from a human-specified cognitive science perspective. In a typical explanation, a rule is retrieved from long-term memory to working memory in order to be applied to the current situation. If the rule matches the situation (i.e. it supports gaining a reward) it is kept in the working

memory. Otherwise, the rule is switched. Such predefined cognitive mechanisms have been implemented in neural network models of prefrontal functionality [2, 3]. The models describe rule switching mechanisms by combining discrete cognitive modules responsible for memory retrieval, matching, working memory, long-term memory and rule alternation.

However, these hardwired explanations seem to lack generality because of their potential arbitrariness. Therefore, the question now is, can we explain meta-level cognition only by combining these predefined cognitive modules? Is there any alternative approach that accounts for meta-level cognition? In order to investigate this issue, we should model meta-cognitive phenomena from the lower grain size level, without assuming apriori functions. Then, we might find novel and more natural mechanisms for explaining the same phenomena. The current paper pursues this type of approach for investigating meta-rule processing.

The motivation for our experiments is to examine possible neuronal mechanisms accounting for rule switching, by providing self-organization pressure on simple neural network models. Our research methodology is based on the combination of evolutionary robotics [4] and dynamic neural networks [5, 6]. Specifically, we conduct robotics simulation experiments without abstracting predefined processes at different cognitive levels. The tasks used in our study integrate meta-level rule manipulation with sensory-motor interaction, highlighting the inseparable nature of these processes. In summary, our experiments and the analysis of the obtained results elucidate that:

- Attractor switching could be a universal mechanism accounting for rule switching in meta-level.
- Some kind of information processing segregation between higher and lower levels enhances the overall network performance. This findings may account for the higher level role of prefrontal lobe in tasks.
- Embodiment is essential even for meta-level cognitive processes because sensory-motor dynamics can support the rule switching mechanisms.

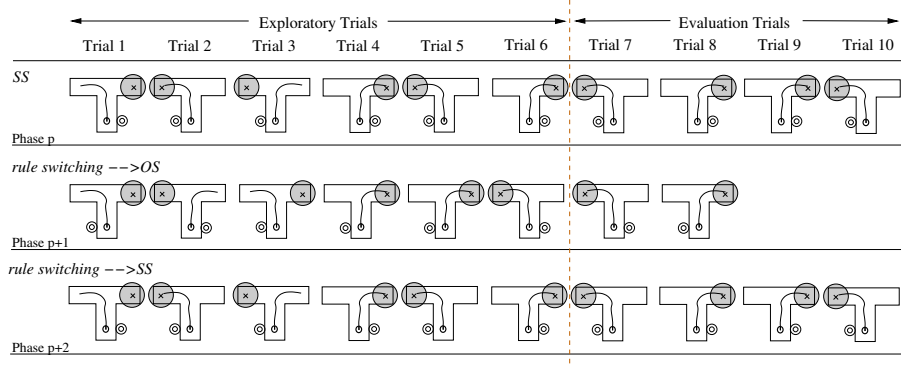
## 2 Behavioral Tasks

We have designed a robotic task that resembles Wisconsin Card Sorting test, but additionally emphasizes dynamic body-environment interaction. Following this approach, we consider both meta-level cognition and sensory-motor coupling as inseparable parts of a complex behavioral problem.

The task goes as follows. Let's assume that a simulated robotic agent is located in the lower part of a T-maze, and a light sample appears at its left or right side [7]. The agent has to move to the end of the corridor making a 90° left or right turning, depending on the side of the light sample. Two different response rules are defined (see Fig 1) similar to [8]. According to the Same-Side (SS) rule, the agent should turn left if the light source appeared at its left side, and it should turn right if the light source appeared at its right side. The complementary response rule named Opposite-Side (OS), implies that the robot should turn to the opposite direction of the light source sides. For both rules,



**Fig. 1.** A schematic representation of the delayed response rules. Light samples are represented by double circles. Target locations are represented by  $\times$ , while reward corresponds to the gray area. The behavioral task asks for controllers capable of switching between the two rules.



**Fig. 2.** A schematic demonstration of the task. For the current example, we assume that  $T_p = 10, T_{p+1} = 8, T_{p+2} = 10$ .

when the agent responds correctly it receives positive reinforcement indicating it is following the correct rule.

The task trials are repeated with resetting robot position to the start position but without resetting the neural state of the robot controller (see below). Because the rules are switched from one to the other unpredictably after several trials, the robot has to adapt its response strategy to the newly adopted rules by monitoring the reward signal. The details of the experimental procedure are described below.

## 2.1 Task Setup

The task is separated to  $P \in \{1...10\}$  phases, each one including  $T_p$  trials. The number of trials  $T_p \in \{8, 10, 12, 14\}$  is randomly specified, so that the agent can not predict the end of a phase. During phase  $p$ , the agent has to follow the same response rule for all  $T_p$  trials. Let's assume for example that it should follow the SS rule. Each trial tests the response of the robot after light sample appearance at its left or right side (their order is randomly chosen). When a trial starts, the robot is sensing the light and then it moves to the end of the corridor where it makes a turn choice. According to the SS rule, the response is correct when the robot turns towards the side of light sample. If the robot makes

the correct choice, it drives to a reward area receiving positive reinforcement. In case that the robot turning is not correct, it will drive to an area that no reward exists, indicating that the currently adopted rule is not correct and it should be switched. During phase  $p$ , the robot is given six free exploratory trials to discover what is the correct rule. In the remaining  $T_p - 6$  trials the performance of the robotic agent is evaluated in terms of following the desired response rule.

If phase  $p$  is completed successfully, the robot moves to phase  $p + 1$ , where the response rule is switched. This means that the reward signals have changed places and -for our example- they are now positioned according to the OS rule. However, the agent is not informed that the rule has been switched and thus it will continue responding according to the previous rule. In that case, the agent will be unable to get more reward, that indicates it is not following the correct rule. In order to get a new reward, the robot must reconsider its rule choice, switching to OS. In phase  $p + 1$ , the robot is given again six free exploratory trials to discover rule switching. In the remaining  $T_{p+1} - 6$  trials agent's responses are evaluated according to the currently correct response rule. If any of these trials is incorrect, the evaluation is interrupted.

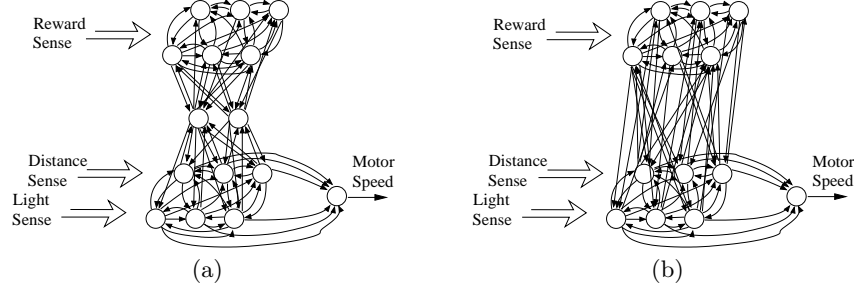
If phase  $p + 1$  is completed successfully, the robot moves to phase  $p + 2$ , where the response rule is switched again -to SS, for our example- and a similar experimental procedure is repeated. Overall, the task evaluates agent's switching behavior for a maximum of  $P$  phases (if all of them are completed successfully).

**Computational Details.** At the beginning of trials the robot is located at a predefined starting position with its direction randomly specified in the range  $[85^\circ - 95^\circ]$  degrees. The robot is kept in the same initial position for five simulation steps, and then it is allowed to navigate freely in the environment for 165 more simulation steps. After the completion of one trial the simulated robot is automatically transferred to the initial position having a new random direction, in order to experiment for the next trial.

### 3 CTRNN Model and Input-Output Connectivity

We use Continuous Time Recurrent Neural Network (CTRNN) models to investigate how meta-level mechanisms self-organize in neuronal dynamics under the given task pressure. Interestingly, in CTRNNs contextual memory is implicitly represented by internal neurons dynamics. In our experimental setup, the neuronal state is initialized only once in the beginning of the first trial, and then neuronal dynamics continues across trials and phases without resetting. In this manner, we speculate that dynamical states will emerge for representing the rule stored in working memory, and additionally, these dynamical states might switch to one another according to the currently adopted rule.

Additionally, both bottleneck (BN) [6] and fully connected CTRNN configurations are explored (see Fig 3), investigating what kind of network structure is essential for achieving meta-level functions. As shown in Fig 3(a), a CTRNN is squeezed in the middle with BN neurons (i.e. the upper and lower parts can interact only through BN neurons). The lower part perceives the sensory flow



**Fig. 3.** Schematic representation of (a) the bottleneck CTRNN and (b) the fully connected CTRNN.

and outputs motor flow and the higher part receives reward stimuli. The BN is supposed to segregate information processing in different levels, maintaining minimum interactions between them. In contrast, in the fully connected case (see Fig 3(b)) information processing levels can hardly differentiate. Overall, the CTRNN consists of 15 neurons for the case of the bottleneck structure and 13 neurons for the case of fully connected structure. All neurons are governed by the standard leaky integrator equations described in previous studies [5, 6].

In order to investigate embodied rule switching in meta-level we employ a two wheeled simulated robotic agent equipped with 8 uniformly distributed distance, light and reward sensors. The robotic platform is based on YAKS environment<sup>1</sup>, which simulates motion dynamics of the Khepera robot.

## 4 Evolutionary Procedure

A Genetic Algorithm<sup>2</sup> (GA) is utilized to explore how the desired meta-level switching behavior can self-organize in the CTRNNs.

**Incremental Evolution.** In order to facilitate successful convergence of the evolutionary process we have used an incremental approach investigating gradually more complex versions of the problem. In the first 60 generations (see Table 1) the evolutionary process asks for robot controllers capable of adopting both SS and OS response rules. Two different tasks are used to evaluate robot controllers. Each task consists of only one phase. The accomplishment of *Task1* implies that the robot can adopt SS rule, while the accomplishment of *Task2* implies that the robot can adopt OS rule. At the beginning of each task the states of all CTRNN neurons are reset to zero, which means that the robot is in a neutral state without following any rule. The robot explores the environment to find out which is the rule that should be adopted in order to get rewards.

<sup>1</sup> The simulator has been slightly modified for the needs of the present study.

<sup>2</sup> The current evolutionary procedure does not mean to represent an artificial counterpart of biological evolution. It only serves our study as a consistent mechanism to explore the domain of solutions for our problem.

Generations	Task Type	Description
1-60	Single Phase	CTRNN reset - <i>Task 1</i> : SS CTRNN reset - <i>Task 2</i> : OS
61-140	Two Phase	CTRNN reset - <i>Task 1</i> : SS $\rightarrow$ OS CTRNN reset - <i>Task 2</i> : OS $\rightarrow$ SS
141-300	Multiple Phase	CTRNN reset - <i>Task 1</i> : SS $\rightarrow$ OS $\rightarrow$ SS $\rightarrow$ OS $\dots$ SS $\rightarrow$ OS CTRNN reset - <i>Task 2</i> : OS $\rightarrow$ SS $\rightarrow$ OS $\rightarrow$ SS $\dots$ OS $\rightarrow$ SS

**Table 1.** The incrementally more complex tasks solved in different parts of the evolutionary procedure.

During generations 61-140, the tasks are getting more complex asking for controllers capable of switching between rules. At this version of the tasks only one switching step is explored, therefore tasks consist of two phases. Properly positioned reward signals indicate the desired response strategy in each phase. The *Task1* examines agent’s ability to adopt SS and then switch to OS. In a similar way, the *Task2* examines robot’s ability to first adopt OS and then switch to SS. At the beginning of each task the CTRNN state is reset to zero, but then it is kept continuous implying that special memory pathways have to develop facilitating rule switching from SS to OS and visa versa.

Finally, in generations 141-300 we ask for controllers capable of repeatedly switching between rules. Both *Task1* and *Task2* are now described by ten phases. Similarly to previous generations CTRNN is reset to zero at the beginning of each task, and then keeps continuous memory state when passing from one phase to the other (i.e. continuously switching between SS and OS rules).

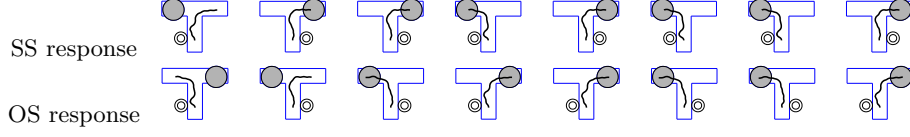
**Task Evaluation.** The accomplishment of tasks is evaluated based on the goal positions of each trial. The goal positions are specified according to (i) the current rule, and (ii) the side of the light sample (see Fig 1). For each response of the robot the minimum distance  $d_{min} \in [0, D]$  between the goal and the robot route, is used to measure the success of robot turning choice ( $D$  is the distance between the starting position and the goal). For a task  $i$  evaluating the behavior of the robot for  $p$  phases, the success on rule switching is given by:

$$E_i = \sum_{q=1}^p \left( \sum_{t=7}^{T_q} \left( 1 - \frac{d_{min}}{D} \right) \right) \quad (1)$$

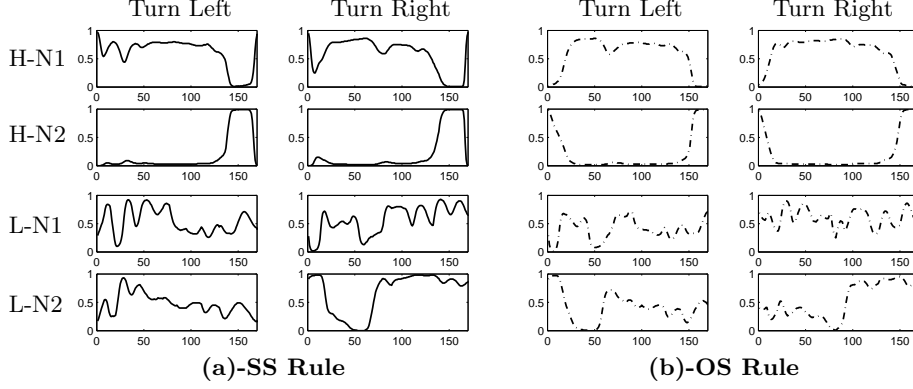
The evaluation starts from trial  $t = 7$  because the first six trials of each phase are exploratory and they are not considered in evaluation. The higher the value of  $E_i$  the more rule switches the agent has accomplished.

**Fitness Measure.** The individuals encoding CTRNN controllers are tested on *Task1* and *Task2* described above. The accomplishment of each task is evaluated separately according to eq (1). The total fitness of the individual is then estimated by:

$$fit = E_{Task1} \cdot E_{Task2} \quad (2)$$



**Fig. 4.** The response of the agent in two consecutive phases.

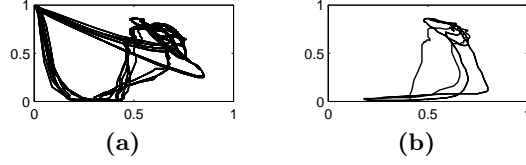


**Fig. 5.** The activation of two higher (H-N1, H-N2) and two lower (L-N1, L-N2) level neurons when the agent follows the SS rule and the OS rule. SS is demonstrated with a solid line while OS is demonstrated with a dashed line.

## 5 Results

For each network topology (either bottleneck or fully connected) we have conducted ten independent evolutionary runs to acquire CTRNN controllers capable of rule switching in meta-level. Specifically, a standard GA with mutation, but without crossover, is employed to evolve populations of 500 encoded CTRNN solutions. In the case of the bottleneck network eight out of the ten runs converged successfully, producing controllers capable of switching between SS and OS rules. However, for the fully connected networks, the success of evolutionary runs was reduced to five out of ten. This significant difference of the success rate implies that bottleneck structures are more appropriate to address meta-level cognitive process. This is because the bottleneck adequately segregates the composite system facilitating self-organization of distinct partial functionalities. For the rest of the paper we will concentrate on the results of bottleneck networks.

The performance of the agent for one representative CTRNN is demonstrated in Fig 4. Despite the fact that evolutionary procedures have been statistically independent, they have all produced CTRNNs with similar internal dynamics. The consistently similar characteristics of neural networks supports the validity of our results in terms of considering their internal dynamics as a valuable alternative scenario for rule switching in meta-level.



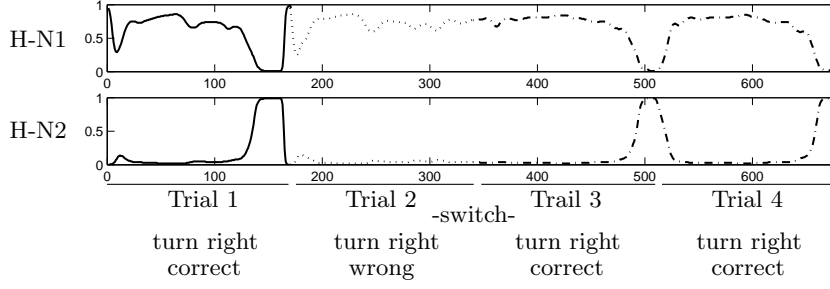
**Fig. 6.** Phase plot of higher level neural activity when the agent follows (a) the SS rule and (b) the OS rule. Neural activities stabilize to attractors having distinct shapes for each case.

Additionally, we have investigated neural activity in the higher and lower levels of the CTRNN network observing different qualitative characteristics in their firing (see Fig 5), implying that layers have acquired different roles in the functionality of the overall system. In particular, the activity of higher level neurons varies mostly in (i) the beginning of trials, indicating they are involved in deciding the response of the agent, and (ii) the end of trials, indicating they are involved in assessing agent’s response in terms of accomplishing the expected reward (Fig 5 (a),(b) top two lines). In contrast, variance in the activity of lower level neurons is observed during the whole trial indicating its involvement in the execution of higher level plans taking also into account environmental interaction issues, e.g. wall avoidance (Fig 5 (a),(b) lower two lines). It is worth emphasizing that the design procedure does not artificially force CTRNN to develop different roles in the higher and lower levels. This property is an emergent result of evolutionary self-organization, and appears consistently in all CTRNNs capable of rule switching in meta-level.

Comparing higher level activity for all four possible cases (left and right response for either the SS or the OS rule), we observe that the two rules are separated by the activation of neurons at the beginning and end of trials. Specifically, both left and right SS rule turnings start and end with maximum activity at H-N1, and minimum activity for H-N2 (see Fig 5(a)). In contrast, these neurons have a complementary activation pattern for OS rule, that is common for both the left and the right turnings (see Fig 5(b)). The fact that the neurons have similar activities for both turnings of a rule facilitates positioning agents behavior in the same context when the underlying rule is adopted. At the same time, the fact that neural activity is very different when different rules are followed facilitates separating agent’s behavioral context between OS and SS rules.

The higher level differences between the two rules can be clearly demonstrated by conducting the attractor analysis taking the phase plots of neuron activities in the higher level. Specifically, for both rules, we ask the agent to perform 30 random turning trials (either left or right) after randomly resetting the neurons in the higher level (Fig 6). For each rule we observed the same shape of attractors to appear in the plot, regardless of the randomness in the initial state. Therefore, each plot represents a distinct invariant set for the corresponding rule. The neuronal state always converge to one of the two invariant sets depending on the reward stimuli condition that specify the currently cor-





**Fig. 7.** Higher level neural activity during rule switching. In the trial 1 the agent follows the SS rule giving a successful right response (demonstrated with a solid line). In the next trial the rule has unexpectedly changed to OS, therefore the agent gives an erroneous response (demonstrated with dotted line). In the trial 3 the agent tries the OS rule that is correct (demonstrated with dashed line), and therefore it continues with the same rule in the next trial (demonstrated again with a dashed line).

rect rule. The generation of distinct attractors for each rule is observed in all successful evolutionary runs, implying that this might be a general mechanism for rule encoding.

The activity of neurons for the case of rule switching is demonstrated in Fig 7. Specifically, when the agent makes a wrong response (due to an unexpected rule change) higher level neurons are unable to get values that confirm the current rule is correct, providing instability to the whole system. This unstable condition facilitates changing the currently adopted rule to the new one (i.e making a transition from the one attractor to the other). We note that according to the current experimental setup the robot is provided 6 free trials to consider rule switching. However, in the majority of cases, the attractor state can transit immediately after a single wrong response, and rarely two or three unsuccessful trials are necessary.

## 6 Discussion

The current study investigates possible mechanism capable of accomplishing rule switching in meta-level. In contrast to traditional approaches using cognitive science terminology to explain rule switching, our results showed that new mechanisms based on dynamical systems principles can also accomplish meta-level manipulation of behavioral rules.

In our results, the self-organized internal dynamics encode each rule to a distinct attractor, while switching of rules takes place by accompanying state transition from one attractor to another. This mechanism seems to be widely valid because the same internal dynamics was observed in the most of our experiments and additionally because the assumptions made in our neural network model are minimal. Our only constraint is the bottleneck connectivity in the network [6]. In that case it is shown that partial segregation of information processing be-

tween the meta-level and the sensory-motor level enhances the performance of the global network. This may account why meta-level cognition has developed in a semi-autonomous higher level module (i.e. prefrontal cortex) in mammals and especially in humans.

It is noted that the agent follows different trajectories depending on the rules adopted. In other words, sensory-motor level dynamics have considerable correlations with meta-level processing. This means that distinct attractors are generated not only by abstract meta-level processing but with the active participation of sensory-motor dynamics. Therefore, we argue that the embodiment plays an important role in meta-level cognition.

Our results can be related to the study on emergence of turn-taking, conducted by Ikegami and Iizuka [9]. In this work, a couple of evolved agents alternates predator-pray roles acquiring autonomous, self-generated behavior switching mechanisms without external triggering. A similar phenomenon of self-generated rule switching was observed in our study, when the reward signal is completely removed for all trials. In that case, the agent repeatedly switches the adopted rule from one trial to the other whenever the robot enters the location where reward “used to be”.

Our future research includes more complex cases with having more than two rules. Our preliminary result showed that 3 rules case can be evolved but with slightly lower success rate. The rule switching dynamics becomes much more complex compared to the two rule case. The detail analysis of these results is left for a future study.

## References

1. Bateson, G., ed.: Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology. Chandler Publishing Company (1972)
2. Kaplan, G., Sengr, N., Grvit, H., Gen, I., Gzelis, C.: A composite neural network model for perseveration and distractibility in the wisconsin card sorting test. *Neural Networks* **19**(4) (2006) 375–387
3. Stemme, A., Deco, G., Busch, A.: The neuronal dynamics underlying cognitive flexibility in set shifting tasks. *J. of Comp. Neuroscience* **23**(3) (2007) 313–331
4. Nolfi, S., Floreano, D.: *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines*. MA: MIT Press/Bradford Books (2000)
5. Yamauchi, B.M., Beer, R.D.: Spatial learning for navigation in dynamic environment. *IEEE Trans. Syst. Man Cybern.* **26**(3) (1996)
6. Paine, R., Tani, J.: How hierarchical control self-organizes in artificial adaptive systems. *Adaptive Behavior* **13**(3) (2005) 211–225
7. Ziemke, T., Thieme, M.: Neuromodulation of Reactive Sensorimotor Mappings as Short-Term Memory Mechanism in Delayed Response Tasks. *Adaptive Behavior* **10**(3/4) (2002) 185–199
8. Maniadakis, M., Trahanias, P.: Hierarchical cooperative coevolution facilitates the redesign of agent-based systems. In: *Proc. 9th Int. Conf. on the Simulation of Adaptive Behavior, (SAB-2006)*. (2006) 582–593
9. Ikegami, T., Iizuka, H.: Joint attention and dynamics repertoire in coupled dynamical recognizers. In: *Proc. of 2nd International Symposium on Imitation in Animals and Artifacts (AISB’03)*. (2003) 125–130