Design and Integration of Partial Brain Models Using Hierarchical Cooperative CoEvolution

Michail Maniadakis and Panos Trahanias

Institute of Computer Science, Foundation for Research and Technology-Hellas (FORTH), Crete, Greece

Department of Computer Science, University of Crete, Heraklion, Crete, Greece

E-mail: {mmaniada, trahania}@ics.forth.gr

Abstract

The current work addresses the problem of designing and integrating brain-inspired artificial cognitive systems. Specifically, we introduce a new computational framework for modelling partial brain areas following a coevolutionary agent-based approach. Properly formulated neural agents are employed to represent brain areas. A cooperative coevolutionary method, with the inherent ability to co-adapt substructures, supports the design of the models, and additionally provides a consistent methodology to achieve their integration. The implemented models are successfully embedded in a simulated robotic platform which supports environmental interaction. The proposed approach is utilized to design two distinct models: one for primary motor cortex able to drive the robot in a purposeless wall avoidance mode, and one for hippocampus which supports self-localization. These models are further integrated by adding at the same time prefrontal structures, in order to drive the robot in a purposeful mode, accomplishing a DMS task in a cross (+) maze.

Introduction

The long-term vision of developing artificial organisms with mammal-like cognitive abilities, has recently given impetus in the design of brain-inspired systems. The brain of mammals consists of interconnected modules with different functionalities, implying that models with distributed architecture should be designed. In this context, a modular design approach is followed by (Krichmar and Edelman, 2003) and (Kozma et al., 2005), to develop distributed brain-like computational models.

The construction of large scale models is difficult to accomplish by developing from scratch complicated structures consisting of hundreds of modules. An alternative approach could be based on implementing separate models of partial brain areas which are further integrated in gradually more complex ones. Along this line, existing approaches suffer in terms of scalability, because they lack a systematic procedure to support the incremental integration of substructures. In contrast, they follow a manual design approach and thus they can not be used as a long-term modelling framework.

We have recently proposed a new method to design distributed partial brain models (Maniadakis and Trahanias, 2005b). Specifically, the model consists of a collection of neural agents, each one representing a brain area. The performance of agents is specified by means of their interaction with other agents, and also by the interaction of the composite model with the external world, simulating epigenetic learning. The self-organization dynamics of epigenetic learning are designed by an evolutionary process which simulates phylogenesis (Rolls and Stringer, 2000). Following the phylogenetic/epigenetic approach, the objective adopted during the evolution of agents, is to furnish them with abilities to develop similar performance to the respective brain areas, after a certain amount of environmental interaction. Instead of using a unimodal evolutionary process, we employ a Hierarchical Cooperative CoEvolutionary (HCCE) approach which is able to highlight the specialties of brain areas represented by distinct agents (Maniadakis and Trahanias, 2005a). By means of this distributed design method, composite models consisting of independent still cooperating modules can be easily formulated.

The present study illustrates the ability of the proposed computational framework to facilitate incremental modelling. This is achieved by combining the benefits of the distributed model with the benefits of the distributed design methodology. At first, the agent-based representation of brain areas enforces the autonomy of substructures, supporting problem decomposition in small tractable and progressively solved tasks. Additionally, the HCCE-based design method, with the inherent ability to combine distributed substructures, provides a mechanism of consistent integration of partial models. Thus, existing structures can be re-utilized in order to develop gradually more complex ones.

The rest of the paper is organized as follows. In the next section, we present the neural agent structures used to represent CNS areas, and the hierarchical cooperative coevolutionary scheme which supports the design of agents. Then, we present the results of the proposed approach on the design of a partial brain computational model of Hippocampal - Prefrontal - Primary Motor cortical interaction. Finally, conclusions and suggestions for future work are drawn in the last section.

Method

The design of brain-inspired structures is based on the argument that the behavior of animals is a result of phylogenetic evolution, and epigenetic environmental experience (Geary and Huffman, 2002). Phylogenetic evolution is facilitated by the HCCE design approach, while epigenetic learning is facilitated by the self-organization dynamics of the computational model. Both of them are described below.
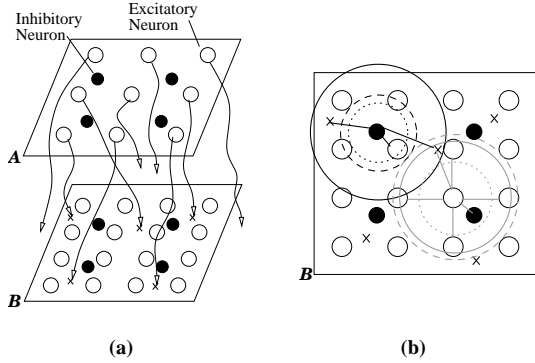
Figure 1: Schematic representation of the computational model. Part (a) illustrates a link agent which supports information flow from cortical agent A to B. Part (b) illustrates synapse definition in cortical agent B.

Computational Model

Two different neural agents, provide a computational framework which supports modelling: (i) a cortical agent to represent brain areas, and (ii) a link agent to support information flow across cortical modules.

Link Agent. The structure of the link agent is properly designed to support connectivity among cortical modules. Using link agents, any two cortical modules can be connected, simulating the connectivity of brain areas.

Each link agent is specified by the projecting axons between two cortical agents (Fig 1(a)). Its formation is based on the representation of cortical modules by planes with excitatory and inhibitory neurons (see below). Only excitatory neurons are used as outputs of the efferent cortical agent. The axons of projecting neurons are defined by their (x, y) coordinates on the receiving plane. Cortical planes have a predefined dimension and thus projecting axons are deactivated if they exceed the borders of the plane. This is illustrated graphically in Fig 1(a), where only the active projections are represented with an $\times$ on their termination. As a result, the proposed link structure facilitates the connectivity of sending and receiving cortical agents supporting their combined performance.

Cortical Agent. Each cortical agent is represented by a rectangular plane. A cortical agent consists of a predefined population of excitatory and inhibitory neurons, which all follow the Wilson-Cowan model with sigmoid activation. Both sets of neurons, are uniformly distributed, defining an excitatory and an inhibitory grid on the cortical plane. On the same plane there are also located the axon terminals from the projected cortical agents.

All neurons receive input information either from i) projecting axons, or ii) excitatory neighbouring neurons, or iii) inhibitory neighbouring neurons. The connectivity of neurons follows the general rule of locality. Synapse formation is based on circular neighbourhood measures. A separate radius for each of the three synapse types, defines the connectivity of neurons. This is illustrated graphically in Fig 1(b), which further explains the example of Fig 1(a). Neighbourhood radius for i) axons

is illustrated by a solid line, for ii) excitatory neurons by a dashed line, and for iii) inhibitory neurons by a dotted line. Sample neighbourhoods for excitatory neurons are illustrated with grey, while neighbourhoods for inhibitory neurons are illustrated with black.

The performance of cortical agents is adjusted by environmental interaction, similar to epigenetic[1] learning (Cotterill, 2001). To enforce experience-based subjective learning, each set of synapses is assigned a Hebbian-like learning rule defining the self-organization dynamics of the agent. This is in contrast to the most common alternative of genetically-encoded synaptic strengths which prevents experience based learning. We have implemented a pool of 10 Hebbian-like rules that can be appropriately combined to produce a wide range of functionalities (Maniadakis and Trahanias, 2005b).

Hierarchical Cooperative CoEvolution

Similar to a phylogenetic process, the structure of agents can be specified by means of an evolutionary method (Rolls and Stringer, 2000). However, using a unimodal evolutionary approach, it is not possible to explore effectively partial components, which represent brain substructures. To alleviate that, coevolutionary algorithms have been recently proposed that facilitate exploration, in problems consisting of many decomposable components (Potter and De Jong, 2000). Specifically, coevolutionary approaches involve many interactive populations to design separately each component of the solution. These populations are evolved simultaneously, but in isolation to one another. Partial populations are usually referred as *species* in the coevolutionary literature, and thus this term will be employed henceforth.

The design of brain-inspired structures fits adequately to coevolutionary approaches, because separate coevolved species can be used to perform design decisions for each substructure representing a brain area. As a result, coevolution is able to highlight the special features of each brain area, and additionally the cooperation within computational modules.

We have presented a new evolutionary scheme to improve the performance of cooperative coevolutionary algorithms, employed in the context of designing brain-inspired structures (Maniadakis and Trahanias, 2005b; Maniadakis and Trahanias, 2005a). We employ two different kinds of species to support the coevolutionary process encoding the configurations of either a Primitive agent Structure (PS) or a Coevolved agent Group (CG). PS species specify partial elements of the model, encoding the exact structure of either cortical or link agents. A CG consists of groups of PSs with common objectives. Thus, CGs specify configurations of partial solutions by encoding individual assemblies of cortical and link agents. The evolution of CG modulates partly the evolutionary process of its lower level PS species to enforce their cooperative performance. A CG can also be a member of another CG. Consequently several CGs

---

[1]Epigenesis here, includes all learning processes during lifetime.
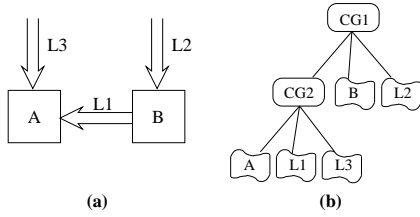
Figure 2: Hierarchical cooperative coevolutionary design. Part (a) represents schematically a hypothetical connectivity of agents. Part (b) represents the hierarchical coevolutionary scheme utilized to evolve partial structures.
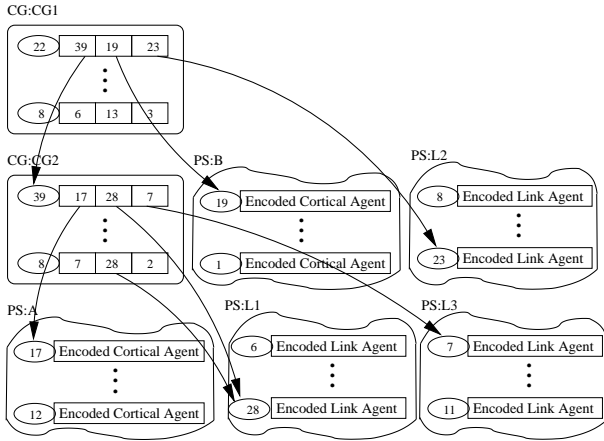

Figure 3: An overview of the hierarchical coevolutionary scheme, with CG species tuning the evolutionary processes of PS species.

can be organized hierarchically, with the higher levels enforcing the cooperation of the lower ones.

The HCCE-based design method for brain modelling is demonstrated by means of an example (Fig 2). We assume the existence of two cortical agents connected by three link agents representing their afferent and efferent projections (Fig 2(a)). One hypothetical HCCE process employed to specify agent structure is illustrated in (Fig 2(b)). CGs are illustrated with oval boxes, while PSs are represented by ovals.

All individuals in all species are assigned an identification number which is preserved during the coevolutionary process. The identification number is employed to form individual assemblies within different species. Each variable in the genome of a CG is joined with one lower level CG or PS species. The value of that variable can be any identification number of the individuals from the species it is joined with. PSs encode the structure of either cortical or link agents. The details of the encoding have been presented in (Maniadakis and Trahanias, 2005b), and thus they are omitted here due to space limitations. A snapshot of the exemplar HCCE process described above is illustrated in (Fig 3). Identification numbers are represented with an oval. CGs enforce cooperation of PS structures by selecting the appropriate

cooperable individuals among species.

In order to test the performance of a complete problem solution, populations are sequentially accessed starting by the higher level. The values of CG individuals at various levels are used as guides to select cooperators among PS species. Then, PS individuals are decoded to specify the structure of cortical and link agents, and the performance of the proposed overall solution is tested on the desired task.

Furthermore, the proposed HCCE scheme allows the employment of separate fitness measures for different species. This matches adequately to the distributed agent-based modelling of brain areas, because different objectives can be defined for each partial structure, preserving their autonomy. For each species s, a fitness function f_s is designed to drive its evolution. All PS species under a CG share a common f_s. Specifically a partial fitness function $f_{s,t}$ evaluates the ability of an individual to serve task t, while the overall fitness function is estimated by:

$$f_s = \prod_t f_{s,t} \tag{1}$$

Furthermore, the cooperator selection process at the higher levels of hierarchical coevolution will probably select an individual to participate in many assemblies. (e.g. the case of individual 28 of PS species L1, of Fig 3). Let us assume that an individual participates in K assemblies which means that it will get K fitness values $f_{s,t}$. Then, the ability of the individual to support the accomplishment of the t-th task is estimated by:

$$f_{s,t} = max_k\{f_{s,t}^k\} \tag{2}$$

where $f_{s,t}^k$ is the fitness value of the k-th solution formed with the membership of the individual under discussion.

The above equations describe fitness assignment in each species of the hierarchical coevolutionary process. Just after testing the assemblies of cooperators and the assignment of their fitness values, an evolutionary step is performed independently on each species, to formulate the new generation of its individuals. This process is repeated for a predefined number of evolutionary epochs, driving each species to the accomplishment of each own objectives and additionally enforcing their composite cooperative performance.

Results

The effectiveness of the proposed approach is illustrated on the design of a partial brain computational model, which simulates Parietal - Prefrontal - Premotor - Primary Motor - Hippocampal interactions, emphasizing on working memory (WM) usage.

The connectivity of brain areas (Fig 4) has been extracted from the detailed diagram presented in (Cotterill, 2001). Several biological experiments highlight the behavioral organization of these brain areas. Especially for rats, they are usually based on Delayed Matching to Sample (DMS) tasks which require to retain memory relative to a sample cue for a brief period, in order to decide upon future behavioral response (e.g. Ragozzino
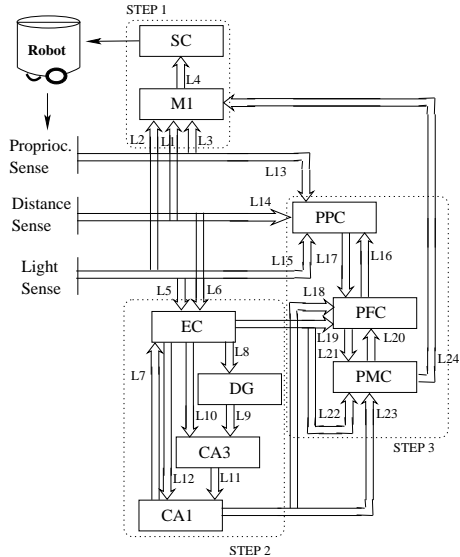
Figure 4: A schematic overview of the composite computational model. Cortical agents are illustrated with blocks, while link agents are illustrated with a double arrow.
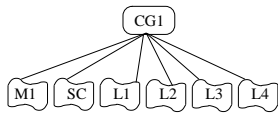

Figure 5: Graphical representation of the process used to design M1-SC model.

and Kesner, 2001). Hippocampal substructures process spatial information to identify the current location of the animal. On the other side, Primary Motor cortex (M1) encodes primitive motor commands which are expressed to actions by means of Spinal Cord (SC). Prefrontal cortex (PFC) is reciprocally connected to Posterior Parietal cortex (PPC) encoding working memory. Hippocampal activity is projected on prefrontal (PFC) and premotor (PM) structures which combine localization and working memory information formulating plans of purposeful motion. PM activation is then passed to M1 which modulates its performance accordingly in order to execute higher level orders.

We note that the present series of experiments is an extension of our previous work which demonstrated the development of distinct models of hippocampus and primary motor cortex (Maniadakis and Trahanias, 2005b). Due to space limitations, the parts which are similar to our previous work will be described here in short, namely in the paragraphs labelled "step 1" and "step 2", in the rest of the current section. The hippocampal and primary motor cortex models are integrated in the present study, adding also computational structures representing prefrontal activity, in order to design a composite model with enhanced performance. Paragraph labelled "step 3" presents in detail current experiments regarding
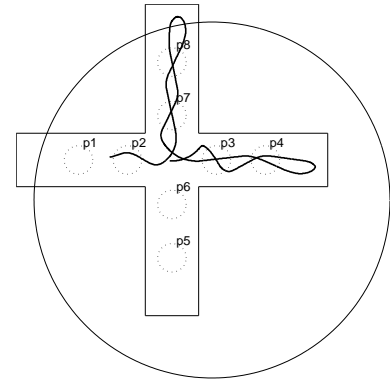

Figure 6: Solid line illustrates robot performance on wall avoidance navigation. The big circle illustrates the range of the light source, while the 8 doted circles illustrate the areas where place-cell development is considered.

incremental integration.

Both partial and composite models are embedded on a simulated mobile robot to furnish it with cognitive abilities and prove the validity of results. We employ a two wheeled robotic platform equipped with 8 uniformly distributed distance and light sensors.

Step 1. The first experiment, accounts for the development of M1-SC computational model which should accomplish primitive motion abilities without purposeful planning. For mobile robots, a task with the above characteristics is wall avoidance navigation. Since M1-SC are the brain modules which serve basic motor commands, it is assumed that they are relevant for the accomplishment of wall avoidance task, $T1$.

M1-SC interactions are modelled by means of a coevolutionary process illustrated in Fig 5. The success of wall avoidance task is evaluated by the fitness measure $E^{M1,SC}$, described in detail in (Maniadakis and Trahanias, 2005b), directing forward robot motion without bumps on the walls. Following the formulation introduced in eqs.(1),(2), the fitness function which guides the coevolutionary process is:

$$f_{CG1} = f_{CG1,T1} \text{ with } f_{CG1,T1}^k = E^{M1,SC}$$

where k represents each membership of an individual in a proposed solution. A sample result of robot wall avoidance motion is illustrated in Fig 6.

Step 2. Our study on hippocampus is focused on the entorinal cortex (EC), dentate gyrus (DG), and Amon's horn structures CA3, CA1. Place-cell activation has been detected in all these structures. The Hippocampal model is evolved by the hierarchical coevolutionary scheme illustrated in Fig 7. The process is joined with the ten best individuals of $CG1$ (which are not evolved in the present step), in order to enforce the synchronization of the robot's wheel speed with the change rate of hippocampal neurons.

The second task $T2$ aims at successful localization by means of place cell activation in hippocampal structures. In order to test the development of place cells, we define
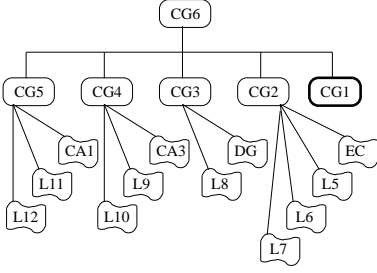
Figure 7: Graphical representation of the process used to design Hippocampal model. $CG1$ (Fig 5), is not evolved in the present coevolutionary process.
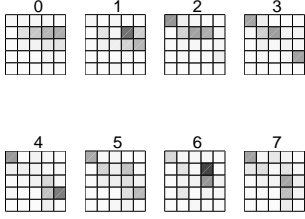

Figure 8: The development of place-cells in CA1 module, with respect to the locations defined in Fig 6.

$P = 8$ areas in the environment (see Fig 6), where the activation of hippocampal excitatory neurons is observed. The successful development of place cell activity in cortical agents evolved by CG_i, $i \in \{2...6\}$ are evaluated by the measures E^j, $j \in \{EC, DG, CA3, CA1\}$ which seek for place-cells with increased separability among different locations. These measures are described in detail in (Maniadakis and Trahanias, 2005b). Following the formulation introduced in eqs.(1),(2), the fitness functions which guide the hierarchical coevolutionary process are:

$$f_{CG2} = f_{CG2,T2} \text{ with } f^k_{CG2,T2} = E^{EC},$$
$$f_{CG3} = f_{CG3,T2} \text{ with } f^k_{CG3,T2} = E^{DG},$$
$$f_{CG4} = f_{CG4,T2} \text{ with } f^k_{CG4,T2} = E^{CA3},$$
$$f_{CG5} = f_{CG5,T2} \text{ with } f^k_{CG5,T2} = E^{CA1},$$
$$f_{CG6} = f_{CG6,T2} \text{ with } f^k_{CG6,T2} = E^{EC} \cdot E^{DG} \cdot E^{CA3} \cdot E^{CA1}$$

where k is as above. The results of place cell development at CA1 are illustrated in Fig 8.

Step 3. When the first two processes are completed, a third coevolutionary scheme commences to design premotor, prefrontal and parietal structures integrating the performance of the two partial models in a composite one (Fig 9). The scheme under $CG6$ is not evolved. The ten best individuals of $CG6$ species are used as indicative hippocampal - primary motor pairs, to form a basis for the construction of the global model.

The successful interaction of substructures is demonstrated by means of a Delayed Matching to Sample (DMS) task, $T3$. Similar to biological experiments with rats, we test the performance of the model in a cross ($+$) maze, highlighting the development and manipulation of working memory (WM), together with the utilization of a place consideration strategy to solve the task.
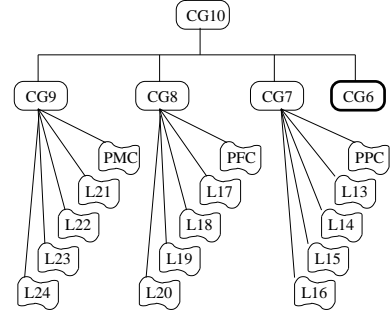

Figure 9: The coevolutionary process utilized to integrate partial models.

During the sample phase, the robot starts from position $s1$ facing the center of the maze. Then it moves along the corridor where it is appropriately driven to turn in the left or right side of maze, by utilizing a Γ-shaped wall (see cases 1,4 of Fig 10). Similar to the biological prototype, WM is encoded by the interactions of PPC, PFC which has to store what was the side of robot turning. Two different states a, b are defined associated with the two possible rotations. For each state, separate activation averages, p_l, are computed, with l identifying excitatory neurons. The formation of WM related to the side of sample turning is evaluated by considering activation in PFC, PPC structures:

$$E^j_{wm} = \left(\frac{v_a}{m_a} + \frac{v_b}{m_b} \right) \cdot \min\left\{ \sum_{p^a_l > p^b_l}^l \left(p^a_l - p^b_l \right), \sum_{p^b_l > p^a_l}^l \left(p^b_l - p^a_l \right) \right\}$$

where $j \in \{PPC, PFC\}$, and m_a, v_a, m_b, v_b are the mean and variance of average activation at states a, b. The first term enforces consistent activation, while the second supports the development of distinct activation patterns for each state a, b. Thus, E^{PPC}_{wm} and E^{PFC}_{wm} evaluate the development of distinct memories relative to the locations a and b, in PPC and PFC.

In contrast to the sample phase, during testing the robot is able to start either from $s1$ or $s2$. The cross maze is transformed in a T-maze utilizing an extra wall (see cases 2,3,5,6 of Fig 10), and the robot is allowed to act freely performing a left or right turn. In order to have a correct response, the robot should turn to the same side with the sample turning. We note that in order to reach a, the robot has to turn right starting from point $s1$, while it has to turn left starting from $s2$. As a result, the task demands the manipulation of both localization information developed in Hippocampus, and working memory stored in prefrontal-parietal structures.

The success on DMS task is evaluated by means of two subtasks related to the two starting positions $s1, s2$. Each subtask tests the responses of the robot related to two different sample phases, in order to evaluate proper target reaching of a or b. The first subtask concerns sample-test pairs 1-2 and 4-5 of Fig 10, while the second subtask concerns pairs 1-3 and 4-6 of the same figure. Two distinct evaluation measures are estimated, E^j_{dr},
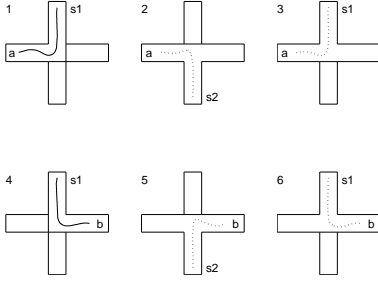
Figure 10: The performance of the robot during DMS task.

$j \in \{s1, s2\}$, one for each subtask:

$$E_{dr}^{j} = (1 + 2(1 - D_a))^3 \cdot (1 + 2(1 - D_b))^3 \cdot \left(1 - 4\frac{B}{M}\right)^2$$

where $D_a, D_b \in [0, 1]$ are the distances between targets a, b and the robot, and B is the total number of robot bumps. The first two terms enforce reaching of a, b, while the last term facilitates goal approximation without crashing on the walls. Thus, E_{dr}^{s1} and E_{dr}^{s2} evaluate the success on DMS task when robot starts from $s1$ and $s2$, respectively.

Following eqs.(1),(2), the fitness functions which guide the hierarchical process illustrated in Fig 9, are:

$$f_{CG7} = f_{CG7,T3^{s1}} \cdot f_{CG7,T3^{s2}} \quad \text{with,}$$
$$f_{CG7,T3^{s1}}^{k} = E_{wm}^{PPC} E_{dr}^{s1}, \quad f_{CG7,T3^{s2}}^{k} = E_{wm}^{PPC} E_{dr}^{s2}$$
$$f_{CG8} = f_{CG8,T3^{s1}} \cdot f_{CG8,T3^{s2}} \quad \text{with,}$$
$$f_{CG8,T3^{s1}}^{k} = E_{wm}^{PFC} E_{dr}^{s1}, \quad f_{CG8,T3^{s2}}^{k} = E_{wm}^{PFC} E_{dr}^{s2}$$
$$f_{CG9} = f_{CG9,T3^{s1}} \cdot f_{CG9,T3^{s2}} \quad \text{with,}$$
$$f_{CG9,T3^{s1}}^{k} = E_{dr}^{s1}, \quad f_{CG9,T3^{s2}}^{k} = E_{dr}^{s2}$$
$$f_{CG10} = f_{CG9,T3^{s1}} \cdot f_{CG10,T3^{s2}} \quad \text{with,}$$
$$f_{CG10,T3^{s1}}^{k} = E_{dr}^{s1}, \quad f_{CG10,T3^{s2}}^{k} = E_{dr}^{s2}$$

where k represents each membership of an individual in a proposed solution. Partial fitness functions of f_{CG7}, f_{CG8} aim at the development of working memory patterns (i.e. E_{wm}^{PPC}, E_{wm}^{PFC}) and the expression of the proper delayed responses from both starting positions (i.e. E_{dr}^{s1}, E_{dr}^{s2}). In contrast, partial fitness functions of f_{CG9}, f_{CG10} focus only on manipulating working memory in order to accomplish the DMS task (i.e. E_{dr}^{s1}, E_{dr}^{s2}).

A sample result of robot performance in the DMS task, is illustrated in Fig 10. Solid lines (cases 1,4) illustrate sample robot motion, while the respective two dotted lines (cases 2,3 and 5,6) illustrate robot response from two different starting positions. It is obvious that localization and working memory information are successfully considered to solve the task. Additionally, lower motor structures (M1, SC) adapt their performance according to the higher level orders, changing their functionality from wall avoidance to goal reaching. Therefore, in the third step, a new complex model is developed by reutilizing existing substructures.

Overall, the presented results illustrate the efficiency of the proposed computational framework in both the design and integration of brain inspired structures.

Conclusions

In the present work, we introduce a computational framework for the design and integration of partial brain models. The proposed method is based on the employment of neural agent modules to represent brain areas. The agent-based approach is in accordance to the distributed nature of mammalian CNS, and additionally supports the decomposition of the problem in small tractable and progressively solved tasks. Furthermore the HCCE-based design methodology facilitates both the design of partial models and their further integration in gradually more complex ones.

We believe that by exploiting the proposed approach, a powerful method to design brain-inspired structures can emerge. Further work is currently underway, to investigate the suitability of our approach in large scale modelling tasks.

References

Cotterill, R. (2001). Cooperation of the basal ganglia, cerebellum, sensory cerebrum and hippocampus: possible implications for cognition, consciousness, intelligence and creativity. *Progress in Neurobiology*, 64(1):1 – 33.

Geary, D. and Huffman, K. (2002). Brain and cognitive evolution: Forms of modularity and functions of mind. *Psych. Bulletin*, 128:667–698.

Kozma, R., Wong, D., Demirer, M., and Freeman, W. (2005). Learning intentional behavior in the k-model of the amygdala and ethorinal cortex with the cortico-hyppocamal formation. *Neurocomputing*, 65-66:23–30.

Krichmar, J. and Edelman, G. (2003). Brain-based devices: Intelligent systems based on principles of the nervous system. In *Proc. IEEE/RSJ Int. Conf. on Intel. Robots and Systems*, pages 940–945.

Maniadakis, M. and Trahanias, P. (2005a). Distributed brain modelling by means of hierarchical collaborative coevolution. In *Proc. IEEE Congress on Evolutionary Computation*, pages 2699–2706.

Maniadakis, M. and Trahanias, P. (2005b). Modelling brain emergent behaviors through coevolution of neural agents. *Neural Networks journal, in print.*

Potter, M. and De Jong, K. (2000). Cooperative co-evolution: An architecture for evolving coadapted subcomponents. *Evol. Computation*, 8:1–29.

Ragozzino, M. and Kesner, R. (2001). The role of rat dorsomedial prefrontal cortex in working memory for egocentric responces. *Neuroscience Letters*, 308:145–148.

Rolls, E. and Stringer, S. (2000). On the design of neural networks in the brain by genetic evolution. *Progress in Neurobiology*, 61:557–579.