A Neural-Group Basis for Evolving and Developing Neural Networks

Keith L. Downing
Complex Adaptive Organically-Inspired Systems Group (CAOS)
The Norwegian University of Science and Technology
Trondheim, Norway
keithd@idi.ntnu.no

ABSTRACT

This work investigates an intermediate abstraction level, that of neural groups, for modelling the development of complex artificial neural networks. Based on Neural Darwinism [5], Displacement Theory [4] and The Neuromeric Model [17], our DEACANN system avoids the complexities of axonal and dendritic growth while maintaining key aspects of cell signalling, competition and cooperation that appear to govern the formation of neural topologies in nature. DEACANN also includes a genetic-algorithm for evolving developmental recipes, and the mature networks can employ several forms of learning.

1. INTRODUCTION

Nervous systems of higher vertebrates have a clearly hierarchical structure, with sensory inputs translated into motor outputs in a fast, purely reactive manner at the lowest levels, but with these same inputs propagating to higher neural structures, whose delayed effects upon motor activity reflect more advanced processes such as sensory integration, memory retrieval, prediction and even planning. However, bi-directional signalling is extremely prevalent among cerebral regions, indicating more of a heterarchical than purely hierarchical organization, as shown in Figure 1.

Many modern Artificial Intelligence (AI) systems that attempt to mimic aspects of higher intelligence exploit similar heterarchical organization. Unfortunately, the closest contemporary computational analog to cerebral mechanisms, artificial neural networks (ANNs), are difficult to train on heterarchical structures. Many researchers have successfully utilized evolutionary algorithms (EAs) to evolve weight vectors for hierarchical and heterarchical ANNs [12, 20, 14], but these direct codings (one gene per weight) do not scale well to the sizes of networks needed for complex, beyond-reactive tasks.

Developmental approaches, in which EA genomes specify a recipe, not a blueprint, for ANN formation, arose to combat the scalability problem [10, 2, 8, 19]. In a seminal article on developmental EAs (DEAs) [16], Stanley and Miikkulainen attack the standard classification of DEA's as too superficial and implementation specific. As an alternative, they propose a biologically-based set of dimensions for distinguishing DEAs. Although somewhat orthogonal to the classic developmental processes cited by Wolpert [18] (cleavage division, pattern formation, morphogenesis, differentiation and growth), Stanley and Mikkulainen's dimensions (cell fate, targeting, heterochrony, canalization and complexification) also focus on the cellular (and even genetic) levels.

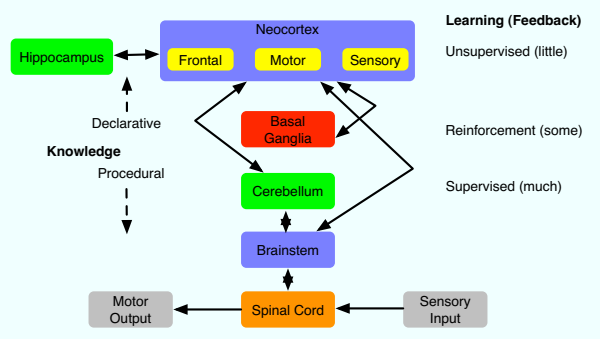

Figure 1: Mammalian sensorimotor control hierarchy

Unfortunately, the distance between these low levels and the control heterarchies seen in the brain is great. Even incremental complexification approaches, as in NEAT [9], cannot achieve the multi-level neural topologies seen in nature. Although Stanley and Miikkulainen argue for abstractions well above the level of cellular migration, chemical signalling and neuritic growth, this paper argues for yet another step up from the cellular level.

In [16], Stanley and Miikkulainen discuss heterochrony: the effects of timing upon developing phenotypes. The strong adaptability of developing embryos normally pre-

AAAI-Devp October 12–15, 2006, Washington, DC, USA.

vents timing changes from being fatal, but instead, allows greater exploration of phenotypic space via small genotypic (timing) variations. Exploration via heterochrony is reasonably safe during early and late development, when intercellular communication is relatively low. However, the middle, *phylotypic* stage is much less flexible due to a high level of global chemical signalling.

Interestingly enough, this phylotypic stage is precisely the stage during which the embryos of many different species look alike. In fact, by examining neural structures during the phylotypic stage, one finds structural similarities that can form the basis for a general model of the development of heterarchical ANNs.

This article examines one of the most popular (among developmental neuroscientists) characterizations of the phylotypic stage, the Neuromeric Model, and describes a DEA based upon it.

2. GROUP-LEVEL PRINCIPLES OF NEUROGENESIS

In *Principles of Brain Evolution* [17], Striedter, a comparative neurobiologist, reviews a host of useful principles, at many abstraction levels, for understanding neurogenesis, neuroevolution, and their interaction. Two of the key principles are the Neuromeric Model and Displacement Theory. The former addresses the spatial arrangement of brain regions, while the latter explains the topology of connections between them. Together, they provide a promising intermediate level of abstraction for DEAs that grow artificial neural networks.

2.1 The Neuromeric Model

In 1953, Bergquist and Kallen [3] noticed that all vertebrate embryos have a similar elongated, segmented hindbrain during the phylotypic stage. The ringed segments, termed *neuromeres*, are modular zones of high cell division (to form neurons) and radial migration (of neurons to their proper layer). Later, Puelles and Rubenstein [15] found that this pattern encompassed the midbrain and forebrain as well. They also provided genetic evidence that Hox and Hox-like genes control this segmentation, just as they control the subdivisions of the body's central axis. Hence, this revised Neuromeric Model views the entire developing brain as an elongated series of ringed modules, within which develop layers of neuron cell bodies (i.e., gray matter).

Figure 2 sketches the basic neuromeric structure of the vertebrate phylotype. The hindbrain neuromeres develop into brain regions such as the cerebellum and pons, which are tightly tied to sensory and motor systems, while the midbrain and forebrain segments become areas such as the basal ganglia, hippocampus and prefrontal cortex, all of which are involved in high-level cognitive processes. Hence, the Neuromeric Model provides the perfect developmental scaffolding for the control heterarchy of Figure 1.

2.2 Displacement Theory

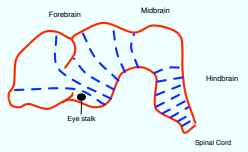

Figure 2: The contemporary neuromeric model.

The Neuromeric Model provides a useful structural bias for generating neuron sub-populations in a DEA for ANNs, while abstracting away many cellular-level details. Deacon's [4] Displacement Theory (DT) compliments the Neuromeric Model by explaining the interactions between the sizes of neural sub-populations and their inter-connectivity, again, while remaining above the cellular level of abstraction.

The basis of DT lies in Edelman's [5] Darwinistic view of neurogenesis, known as The Theory of Neural Group Selection (TNGS). In this view, neurons undergo a selective process wherein only those that grow axons to, and receive axons from, other neurons will reach maturity. Essentially, neurons are involved in a *survival of the best networkers* competition. DT expounds on TNGS by proposing that the networking competition during early development enables brains to *scale to fit* the body's sensory and motor apparatus. In short, primary sensory and motor areas of the brain are sized according to their immediate inputs or outputs, respectively. Secondary region sizes derive from those of the primary regions, and deep cortical structures grow or shrink to fit their input and output sources.

Figure 3 conveys the essence of TNGS and DT. Note a) the expansion of the 3 neuron groups along the path from the largest sensory input, S1, to the largest motor output, M2, and b) the decline of groups B and D, which lose the competition for C's dendrites and C's axons, respectively. As Deacon explains:

> So although genetic tinkering may not go on in any significant degree at the connection-by-connection level, genetic biasing at the level of whole populations of cells can result in reliable shifts in connection patterns...relative increases in certain neuron populations will tend to translate into the more effective recruitment of afferent and efferent connections in the competition for axons and synapses. So, a genetic variation that increases or decreases the relative sizes of competing source populations of growing axons will tend to **displace** (our emphasis) or divert connections from the smaller to favor persistence of connections from the larger.

Developmental neuroscience clearly supports and employs the key tenets of DT. For example, Fuster [7] documents the earlier maturation of posterior brain regions

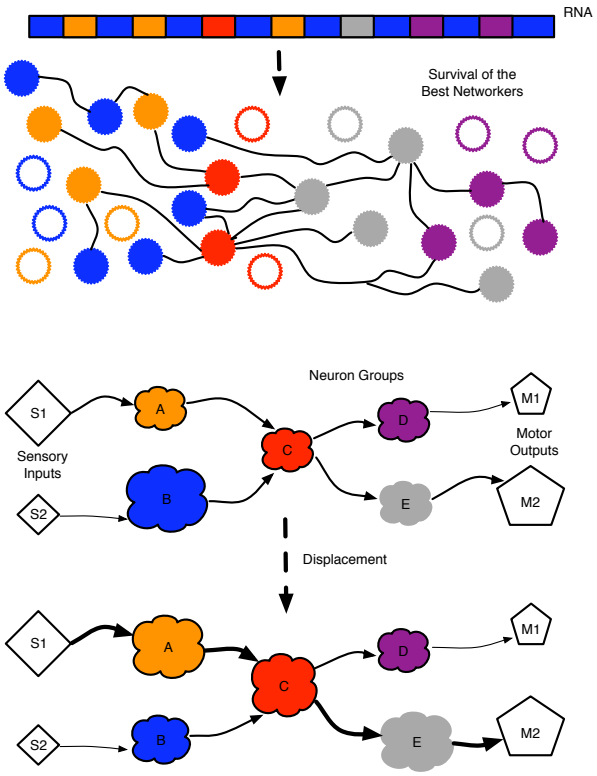

Figure 3: (Above) The Theory of Neural Group Selection (TNGS), wherein neurons survive only if they establish efferent and afferent connections. Dying neurons are unfilled in the figure. (Below) Displacement Theory (DT), in which neuron groups with good networking possibilities expand, while others shrink. Sizes of sensory, motor, and neuron group icons depict sizes of the corresponding pools.

(used in early sensory processing) and late maturation of frontal areas such as the prefrontal cortex (PFC). Striedter [17] combines DT with the work of Finlay and Darlington [6] to explain the trend of greater neocortical (and especially PFC) control of lower brain regions in higher organisms. First, Finlay and Darlington show that *late equals large* in neurogenesis: larger brain regions are those that mature later in development. Second, a key corollary of Deacon's DT is that *large equals well-connected*: big brain regions send many axons to other regions and thereby have a significant amount of control over them. Together, these show how small changes in developmental timing (in higher mammals) have enabled the frontal regions to mature later, hence grow larger, and hence exhibit greater control over a wide variety of cortical areas. And greater frontal control correlates well with behavioral sophistication, as illustrated by Nakajima et al.'s [13] comparisons of manual dexterity vis-a-vis frontal control of motor areas in mammals such as cats, monkeys and humans.

Together, these theories paint neurogenesis as a self-organizing process in which genetics determines the neuromeric structure, the basic properties of neurons in different layers of the neuromeres, and the maturation speed of neural regions, but the final sizes of these regions and their interconnectivity stem from self-organizing processes wherein neuronal subpopulations essentially compete for the right to cooperate (via synaptic signalling) with other subpopulations.

3. GROUP-LEVEL DEVELOPMENT IN DEACANN

Designed to emulate the essential elements of Neuromere formation, Neural Group Selection and Displacement Theory, while avoiding the computationally intensive simulation of axon growth, the developmental algorithm employed in DEACANN involves three phases.

In phase I, **Translation**, the binary genome is converted into a set of neuromeres, each containing one or more neuron groups, whose basic properties are genetically determined.

In phase II, **Displacement**, the sizes of each neuron group and intra- and inter-group invasion strengths undergo repeated cycles of modification. Since (in both TNGS and DT) group size affects invasion strength, and vice versa, several iterations are required to translate the initial group masks, inter-group distances, growth limits and proximities (to sensory inputs and motor outputs) into group sizes and relative degrees of inter-group connectivity.

In phase III, **Instantiation**, the final group sizes and invasion strengths are used to generate a) populations of neurons for each group, and b) connections between individual neurons in the same and different groups. Although not detailed in this paper, some of these connections specify neuromodulatory signalling pathways, wherein single neurons send messages to entire neuron

groups.

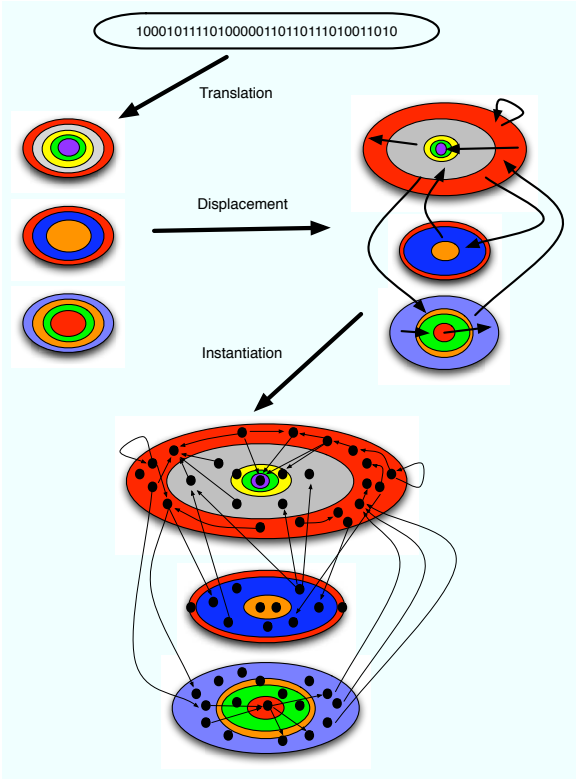

Figure 4: DEACANN's 3-stage developmental model. Each circle denotes a neuromere, while each concentric ring represents a neuron group, with ring width proportional to group size. In the upper right, arrows between neuron groups indicate biases (quantified as invasion factors - see below) that generally lead to axonal connections during the instantiation phase. In the bottom section, small dots are neurons, and their axons are thin, outward-bound arrows.

3.1 Translation

The DEACANN genome needs to support the development of neuromere chains, where each segment may consist of several distinct neuron groups. All neurons in a group should be similar, and a group type may repeat several times in the same or in different neuromeres. For example, a layer of lateral inhibitors may occur several places in the developed neural network.

Although one could rely on evolution to duplicate the genes that encode these groups, the probability of such copying (in a linear genome) decreases dramatically as gene complexity rises. Although genetic programming with subroutines [11] provides one method of modularizing reusable genes, the computational overhead of GP seems unnecessary in this case, given the objective of evolving sets of neuron-group parameters, not actual growth procedures.

In this special case, where neither the number of neuromeres nor the number of groups per neuromere is fixed, we can achieve modularity and reuse within a linear GA chromosome by simply removing the standard GA constraint of fixed gene locations and adopting a tag-based addressing system, as shown in Figure 5.

In this indirect representation, binary tags denote the start of a neuromere specification. In the example of Figure 5, this tag is 11111, with length k = 5. The user can specify the degree to which any k-length segment of the chromosome must match this tag to qualify as a hit. A matching segment is called a *neuromere header*. Here, we assume a match degree of 100 % for neuromere headers, i.e., they must match the neuromere tag exactly.

Once a neuromere header is found, the bits directly following it are interpreted as the *neuromere specification*. In the current version of DEACANN, this specification is simply a single value denoting a neuron-group tag. In Figure 5, the single neuromere specification (denoted by the pentagon labelled "N") has 01010 following its header. Thus, the neuron-group tag for this neuromere is 01010.

The translator then scans the (entire) chromosome in search of segments that match the neuron-group tag, 01010. In this example, we assume that a 80% or more of a chromosome segment's bits must match the tag. Each such matching segment denotes the header of a neuron-group specification, and the bits following the header are translated into the various parameters of the group, such as axon and dendrite masks, learning rates, etc. (which are detailed below).

Thus, for each neuromere specification header that is found, a neuron group tag is read from the chromosome and used as the basis for a complete scan of the chromosome for the specifications of the neuron groups of that neuromere. During a scan, the group specifications cannot overlap. However, the group specification for one neuromere may be ignored or interpreted differently during the scan for another neuromere's groups. Similarly, the (rather short) neuromere specifications cannot overlap.

A match degree less than 100% allows neuromeres to share some, but not necessarily all, group specifications, since each neuromere has a potentially unique group tag. For example, assuming a common group-tag-match degree of 75%, if neuromere N1 has group tag 1111 and neuromere N2 has group tag 1010, then both neuromeres would share any group specification with header 1011 or 1110, but only N1 would include groups with header 0111 and only N2 would select groups headed with 1000.

This representation supports complexification via the well-known combination of genetic duplication and differentiation [1]. Being relatively short, neuromere coding regions are frequently copied in their entirety, thus

creating additional neuromeres in later generations. Any mutation to the new neuromere's group tag then provides the potential for differentiation, since this enables the new neuromere to inherit some, but not all, of the neuron-group types from its ancestor neuromere.

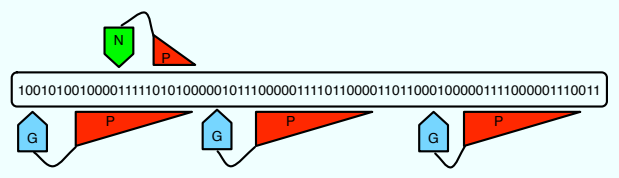

Figure 5: Translation of a bit string into the parameters for the neuron groups of a neuromere. Pentagons point to the start of neuromere (N) and group (G) tags, while triangles denote the extent of the parameter bits (P) that follow each tag. In this example, the tag for neuromeres is 11111 and must be matched exactly, while group tags must match on at least 4 of their 5 bits. The single parameter for the neuromere is the tag for its groups, 01010, which appears under the triangle attached to pentagon N. Hence, all strings that match 80% of 01010 mark the start of group specifications.

The neuron-group specification consists of a contiguous string of bits that encode the following parameters:

1. Axon mask - Masks are abstractions of cadherins, ephrins, [18] and other chemicals that govern the attraction and repulsion of axons during their migration toward dendritic targets. In general, the attractiveness of one neuron group for another is directly proportional to the complementarity of the axon and dendrite masks in the two groups, and inversely proportional to the distance between the groups.
2. Dendrite mask
3. Axon sharing - The sharing parameters indirectly control the patterns of connectivity formed between two groups during the Instantiation phase of development.
4. Dendrite sharing
5. Postsynaptic effect - This indicates whether the neurons in the group have excitatory or inhibitory effects upon their targets.
6. Neuromodulator sent - This governs whether neurons in the group send neuromodulators instead of normal action potentials when excited, along with coding the type of neuromodulator. If the group is excitatory (inhibitory), then it's neuromodulator will always have an excitatory (inhibitory) effect upon targets.
7. Neuromodulator received - This indicates the neuromodulator (if any) to which the group neurons are sensitive.
8. Growth limit - The number of developmental rounds in which the group neurons will participate.
9. Learning rate - Degree to which synaptic-strength is modified on a single training trial during post-developmental adaptation.

3.2 Displacement

Phase II of development is the heart of the DEACANN approach. It simulates the interaction between neuron groups, both within and between neuromeres, but without simulating the actual growth of axons and dendritic trees. Instead, three interacting factors are iteratively updated: 1) the cardinality of the neuron sets in each group, S_i, 2) the invasion strength of each group relative to itself and others, $I_{i,j}$, and 3) the connectivity of each neuron group, C_i.

Invasion strength, $I_{i,j}$, represents the propensity of neurons in group i for sending axons to invade targets in group j. It is a function of the distance between the two groups, $D_{i,j}$, the compatibility of the axon mask for i with the dendritic mask for j, $M_{i,j}$, and the sizes of groups i and j, S_i and S_j. The basic update formula is given in equation 1, in which α_m and α_d are weighting constants with typical values of 0.5 and 1, respectively, and T is the total elapsed time since the beginning of Phase II. The inclusion of T allows groups that are distant from one another but otherwise compatible, i.e. high $M_{i,j}$, to eventually hook up if both groups can maintain a reasonable size during the early stages of displacement.

$$I_{i,j} = S_i S_j (T+1) \frac{\alpha_m M_{i,j}}{1 + \alpha_d D_{i,j}} \tag{1}$$

For each group, i, the total invasion force into, I_i^{in}, and out from i, I_i^{out} are derived from the pairwise invasion strengths as:

$$I_i^{in} = \sum_{j \in G} I_{j,i} \tag{2}$$

and

$$I_i^{out} = \sum_{j \in G} I_{i,j} \tag{3}$$

Neuron groups compete for targets via the following modification to invasion strength:

$$I_{i,j} \leftarrow (1-\gamma) \frac{I_j^{in}}{N_g} + \gamma I_{i,j} \tag{4}$$

Here, $\frac{I_j^{in}}{N_g}$ is the average intensity of invasions into group j; N_g is the total number of neuron groups in the ANN. γ denotes the competition intensity, a constant with a range of [0 1] and typical value of 0.5. Higher values of γ will penalize intensities that are below the average, thus squeezing out weak invaders. Conversely, a γ closer to 0 will push all invasion strengths closer to the average, thus more evenly distributing access to group j's dendrites.

The competition-modified invasion strengths are re-summed to update the invasion forces, and the estimated - remember, no actual connections have yet been formed - connectivity of each neuron group is then computed as:

$$C_i = I_i^{in} + I_i^{out} \tag{5}$$

To reflect the basic tenet of Neural Darwinism - the best networkers proliferate, while poorly-connected neurons die - neuron-group sizes are updated as a function of connectivity:

$$\triangle S_i = \alpha_g(C_i - \overline{C}) \tag{6}$$

where $\overline{C}$ is the average connectivity over all neuron groups, and α_g is a growth constant, typically 0.1. Notice that S_i does not occur in this update formula; its effect is already present in the contribution of I_i^{out} to C_i, since each addend of I_i^{out} involves S_i.

These updates of $I_{i,j}$, C_i and S_i are repeated a maximum of R (a user-defined parameter typically in the range [5 10]) developmental rounds. Any group with a growth limit, G_i less than R will not participate in the final R-G_i rounds. This embodies the *late equals large* principle, since neuron groups with longer growth limits will tend to have higher connectivity - remember the affect of T in equation 1 - and hence larger sizes.

Initially, each group size is 1, with the exception of the input and output groups, whose sizes are constant and equal to the number of sensors and affectors, respectively. Sensory inputs are assumed to enter through the outermost group of neuromere 0, while outputs exit through its innermost group. Groups that are spatially adjacent or otherwise attractive as targets for the input and output groups will have high incoming invasion forces, and thus high connectivity and increasing sizes, exactly as prescribed by Displacement Theory.

3.3 Instantiation

Given the S_i and $I_{i,j}$ values for each group and group pair, respectively, Phase III of development uses these values to bias the generation of neurons for each group and connections between them.

To generate neurons, all group sizes, except the input and output groups, are normalized to produce distribution fractions for a fixed total number of neurons, N.

Connections are formed by considering each group, i, and all invasion factors into that group. These factors are normalized to produce the values $\tilde{I_{j,i}}$ for all groups j. For each potentially invading group j, corresponding neurons in groups i and j are considered. For each pair, a connection from the j neuron to the i neuron is formed with probability $\tilde{I_{j,i}}$.

Once formed, a connection from the kth neuron of group j to the kth neuron of group i neuron can spawn further connections of a convergent, divergent or parallel form, depending upon the axonal sharing factor of group j, λ_j^{out} and the dendritic sharing factor of group i, λ_i^{in}. In the process described below, the kth neuron of group is denoted $G_i(k)$

- $k_i = k$, $k_j = k$
- Repeat
  - Create connection $G_j(k_j) \rightarrow G_i(k_i)$
  - Generate random fractions r_i and r_j
  - If $(r_i \leq \lambda_i^{in} \wedge r_j \leq \lambda_j^{out})$ then $k_i \leftarrow k_i + 1$ and $k_j \leftarrow k_j + 1$
  - Else if $r_i \leq \lambda_i^{in}$ then $k_i \leftarrow k_i + 1$
  - Else if $r_j \leq \lambda_j^{out}$ then $k_j \leftarrow k_j + 1$
- Until $(r_i > \lambda_i^{in} \wedge r_j > \lambda_j^{out})$

Briefly, if both sharing factors are high, then the first conditional will often be true and both indices will be incremented, causing the next pair of corresponding neurons to be paired. Several rounds of this will create a set of parallel connections, similar to a topological map. If dendritic sharing is high but axonal sharing low, then the second, but not the first, conditional will often be true, thus causing the presynaptic neuron to remain fixed, while the postsynaptic neuron becomes the neighbor to $G_i(k_i)$. A few repeats of this situation forms a divergent connection pattern from group j to i. Conversely, if dendritic sharing is low but axonal sharing high, the third condition is frequently triggered and a convergent pattern results. Finally, if both sharing values are low, then the loop exits early and only one or a few connections are created. These outcomes are summarized in Figure 6.

4. PRELIMINARY RESULTS

To date, the system has only been used to evolve ANNs to meet simple structural goals; the nets have not been used as actual controllers. The ANNs of Figures 7 and 8 were evolved using a population size of 20 over 20 generations, with a mutation rate of 0.1 (per individual) and a single-point crossover rate of 0.5. Fitness-proportionate selection was used, with single-individual elitism. Genomes were 100 bits long, and the developmental process had a maximum duration of 5 rounds.

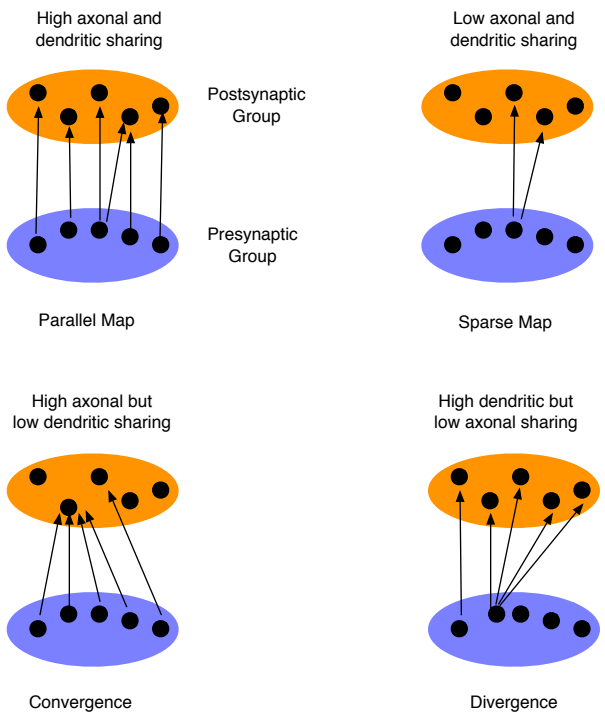

Figure 6: Inter-group connection topologies as a function of axonal and dendritic sharing factors.

The fitness functions are based on the number of neuromeres (N), the average size of each neuromere in terms of its neuron-group cardinality, $\overline{S}$, and the standard deviation of neuromere size, σ_S. In Figure 7, fitness is simply $N\overline{S}$, while in Figure 8, size diversity is also rewarded, while N receives less weight:

$$Fitness = \frac{N}{2} + \overline{S} + \sigma_S \qquad (7)$$

Although having no role in the current evolutionary task, the darker neurons (with solid lines emanating from them) in Figures 7 and 8 are excitatory, while the lighter circles (with dotted emerging lines) are inhibitory. To avoid clutter, the connections lack directional arrows.

5. DISCUSSION

We have just begun testing DEACANN on simple performance tasks but have not yet evolved useful controllers. Unfortunately, even if successful, these toy-problem tests will not convincingly validate this approach. Complex tasks that require large neural hierarchies or heterarchies as controllers must eventually be attempted. Although DEACANN's intermediate level of abstraction saves considerable computational effort during development, it still involves a complex iterative process that is quadratic in the number of neu-

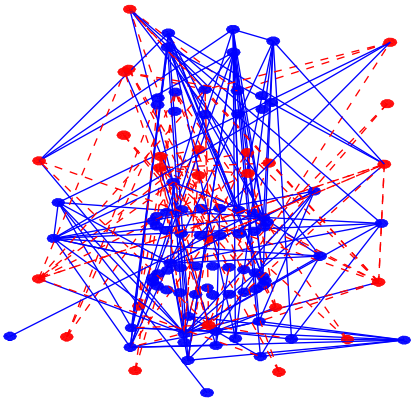

Figure 7: ANN evolved to have as many neuromeres and neuron groups as possible. The neuromeres occupy successive horizontal planes in the picture, and neuron groups are concentric rings within each plane.

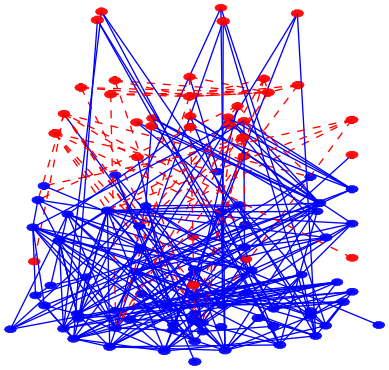

Figure 8: ANN evolved to have a diversity of neuromere sizes.

ron groups. To combat the resource demands of integrated evolutionary, developmental and learning systems (a.k.a. TRIDAP systems), we are currently exploring the use of distributed evolution, where each processor handles an individual's development, task performance and learning.

Notwithstanding computational issues, the DEACANN approach deserves consideration in the general attempt to scale up evolutionary and developmental systems to complex problems that require large, well-structured controllers. It has become a somewhat abused cliche in bioinspired computing to point to nature's successful approaches as justification for similar attempts in silico. Our work, particularly in its current state, certainly has no better grounds for invoking that cliche.

However, we do feel that moving up the abstraction hierarchy has clear advantages over moving down. The neuron-group level a) still affords the cooperative and competitive interactions at the heart of Neural Darwinism, and b) appears to be the proper level at which to explain the formation of large-scale brain topologies, via principles such as *late equals large* and *large equals well connected* and structures such as laminae and neuromeres [17]. In artificial TRIDAP systems, development can afford to work primarily at this intermediate level, since learning can (and should) handle the fine-tuning of individual synapses.

Finally, it is worth noting that nature-inspired techniques in engineering, from simulated annealing to genetic algorithms to neural networks to swarm intelligence, are based on the essence, not the details, of a natural process. Since the development of complex heterogeneous control systems, i.e., brains, is best explained at an intermediate level, there is no reason to believe that it cannot be simulated there as well.

6. REFERENCES

[1] J. Allman, *Evolving Brains*, W.H. Freeman and Company, New York, NY, 1999.

[2] J. Astor and C. Adami, *A developmental model for the evolution of artificial neural networks*, Artificial Life, 6 (2000), pp. 189–218.

[3] H. Bergquist and B. Kallen, *On the development of neuromeres to migration areas in the vertebrate cerebral tube.*, Acta Anat (Basel), 18 (1953), pp. 65–63.

[4] T. Deacon, *The Symbolic Species: The Co-evolution of Language and the Brain*, W.W. Norton and Company, New York, 1998.

[5] G. Edelman and G. Tononi, *A Universe of Consciousness*, Basic Books, New York, NY, 2000.

[6] B. Finlay and R. Darlington, *Linked regularities in the development and evolution of mammalian brains*, Science, 268 (1995), pp. 1578–1584.

[7] J. Fuster, *Cortex and Mind: Unifying Cognition*, Oxford University Press, Oxford, 2003.

[8] F. Gruau, *Genetic micro programming of neural networks*, in Advances in Genetic Programming, J. Kenneth E. Kinnear, ed., MIT Press, 1994, ch. 24, pp. 495–518.

[9] Kenneth and R. Miikkulainen, *Evolving neural networks through augmenting topologies*, Evolutionary Computation, 10 (2002), pp. 99–127.

[10] H. Kitano, *Designing neural networks using genetic algorithms with graph generation system*, Complex Systems, 4 (1990), pp. 461–476.

[11] J. R. Koza, *Genetic Programming II: Automatic Discovery of Reusable Programs*, MIT Press, Cambridge, MA, 1994.

[12] D. J. Montana and L. D. Davis, *Training feedforward networks using genetic algorithms*, Proceedings the Eleventh International Joint Conference on Artificial Intelligence, (1989), pp. 762–767.

[13] K. Nakajima, M. A. Maier, P. Kirkwood, and R. Lemon, *Striking differences in transmission of coricospinal excitation to upper limb motoneurons in two primate species*, Journal of Neurophysiology, 84 (2000), pp. 698–709.

[14] S. Nolfi and D. Floreano, *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines*, The MIT Press, Cambridge, MA, 2000.

[15] L. Puelles and J. Rubenstein, *Expression patterns of homeobox and other putative regulatory genes in the embryonic mouse forebrain suggest a neuromeric organization.*, Trends in Neuroscience, 16 (1993), pp. 472–479.

[16] K. Stanley and R. Miikkulainen, *A taxonomy for artificial embryogeny*, Artificial Life, 9 (2003), pp. 93–130.

[17] G. F. Streidter, *Principles of Brain Evolution*, Sinauer Associates, Sunderland, Massachusetts, 2005.

[18] L. Wolpert, *Principles of Development*, Oxford University Press, New York, 2002.

[19] L. Yaeger, *Computational genetics, physiology, metabolism, neural systems, learning, vision and behavior or polyworld: Life in a new context*, in Artificial Life III, Proceedings Volume XVII, C. G. Langton, ed., Reading, Massachusetts, 1994, Santa Fe Institute Studies in the Sciences of Complexity, Addison-Wesley, pp. 263–298.

[20] X. Yao, *Evolving artificial neural networks*, Proceedings of the IEEE, 87 (1999), pp. 1423–1447.