EVOLUTIONARY AUTONOMOUS AGENTS: A NEUROSCIENCE PERSPECTIVE

Eytan Ruppin

In this article, I discuss the use of neurally driven evolutionary autonomous agents (EAAs) in neuroscientific investigations. Two fundamental questions are addressed. Can EAA studies shed new light on the structure and function of biological nervous systems? And can these studies lead to the development of new tools for neuroscientific analysis? The value and significant potential of EAA modelling in both respects is demonstrated and discussed. Although the study of EAAs for neuroscience research still faces difficult conceptual and technical challenges, it is a promising and timely endeavour.

EVOLUTIONARY COMPUTATION
A computational paradigm that is based on searching the space of possible solutions to a problem by selectively creating new solutions from the best obtained so far.

EMBODIED AGENTS
Agents whose fitness is determined by an interaction with the environment in which they live, given a set of sensors and motors, and a controlling network that mediates between the two.

ARTIFICIAL NEURAL NETWORK
A computational model, the architecture of which is modelled after the brain. They contain idealized neurons called nodes, which are connected together in a network.

School of Computer Science and School of Medicine, Tel-Aviv University, Tel-Aviv 69978, Israel.
e-mail: ruppin@math.tau.ac.il
DOI: 10.1038/nrn729

Recent years have witnessed a growing interest in the study of neurally driven evolutionary autonomous agents (EAAs). These studies, which are part of the field of EVOLUTIONARY COMPUTATION and artificial life[1–4], involve EMBODIED AGENTS — either software programs living in a simulated virtual environment or hardware robotic devices — that 'live' in an environment and autonomously perform tasks such as gathering food, navigating, evading predators, and seeking prey and mating partners. Each agent is controlled by an ARTIFICIAL NEURAL NETWORK (ANN) 'brain'. This network receives and processes sensory inputs from the surrounding environment and governs the agent's behaviour by activating the motors that control its actions. The controlling networks are developed through GENETIC ALGORITHMS that apply some of the essential ingredients of inheritance and selection to a population of agents that undergo evolution.

A typical EAA experiment consists of a population of agents that are evolved using a genetic algorithm over many generations to survive best in a given environment (FIG. 1). In general, agents can have different kinds of controller, and can also encode sensors and motors in their genome, but I focus in this review on agents with a genome that encodes only their controlling neural network. At the beginning of each generation, a new population of agents is generated by selecting the fittest agents of the previous generation and letting them mate — that is, form new agent genomes by genetic recombination followed by mutations that introduce further variation in the population. The genomes that are formed in this process are 'transcribed' to form new agents that are placed in the environment for a given amount of time, after which each agent receives a fitness score that designates how well it performed the evolutionary task. This ends a generation cycle, and a new generation is initiated. Typically, this evolutionary 'search' process is repeated for many generations, until the agents' fitness reaches a plateau and further evolutionary adaptation does not occur. The result is a final population of best-fitted agents, the emergent behaviour and underlying neural dynamics of which can now be thoroughly studied in 'ideal conditions'. The experimenter has full control of the environment and other experimental conditions. More importantly, the experimenter also has complete knowledge of both the agent's behaviour and the controlling network's architecture and dynamics. This scenario is illustrated in a concrete example of a typical EAA navigation and foraging experiment, adapted from REF. 5.

The agents in the model live in a grid arena of size 30×30 cells that is surrounded by walls (FIG. 2). 'Poison' is randomly scattered around the arena (consuming this resource results in a negative reward). 'Food', the consumption of which results in a positive reward, is

GENETIC ALGORITHM
One of the main ways of performing evolutionary computing, based on generating new solutions from existing ones by applying to them genetically inspired operations, such as mutations and crossover. In most cases, the genotype (in which variation takes place) is different from the phenotype (in which fitness-dependent selection take place).

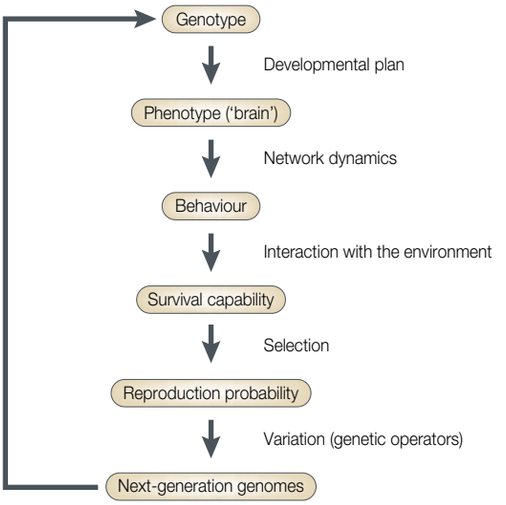

Figure 1 | **The paradigm of evolutionary autonomous agents (EAAs).**

randomly scattered in a restricted 10 × 11 'food zone' in the southwest corner of the arena. The agents' behavioural task is to eat as much of the food as they can, while avoiding the poison. The complexity of the task stems from the partial sensory information that the agents have about their environment. They are equipped with a set of sensors, motors and a fully recurrent ANN controller.

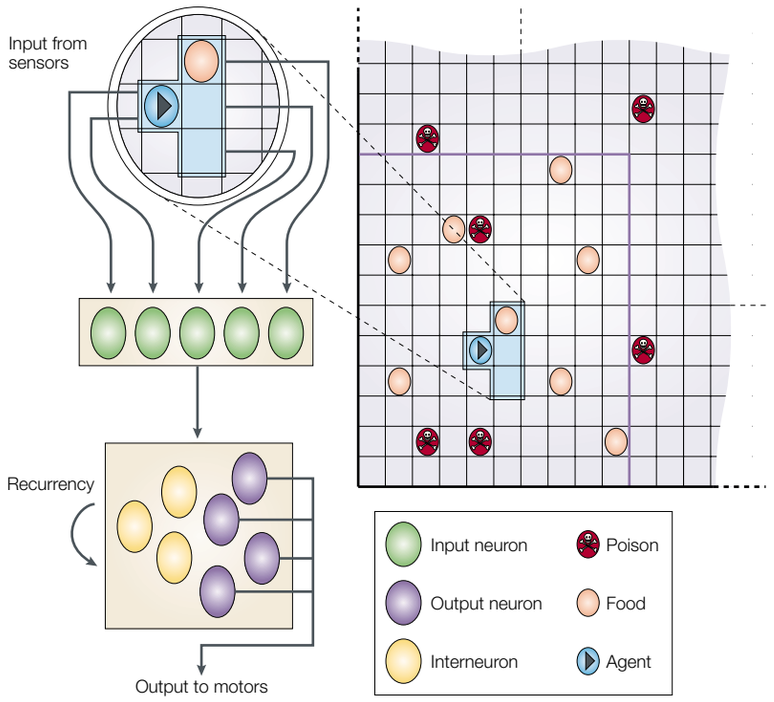

Figure 2 | **An outline of the food zone (southwest corner of the grid arena) and the agent's controlling network.** The borders of the food zone are marked (purple line), but they are invisible to the agent. The agent is marked by a small arrow on the grid, the direction of which indicates its orientation. The T-shape marking in front of the agent denotes grid cells that it senses; the curved lines indicate where in the arena each of the sensory inputs comes from. Output neurons and interneurons are all fully connected to each other. Adapted with permission from REF. 5

The neurocontroller is coded in the genome and evolved; the sensors and motors are given and constant.

The initial population consists of 100 agents with random neurocontrollers. Each agent is evaluated in its own environment, which begins with 250 poison items and 30 food items. The life cycle of an agent (an epoch) is 150 time steps, in each of which one motor action takes place. At the beginning of an epoch, the agent is introduced to the environment at a random location and orientation. At the end of its life cycle, each agent is assigned a fitness score, which is calculated as the total amount of food it has consumed minus the total amount of poison it has eaten. Simulations last for between 10,000 and 30,000 generations.

Each agent is controlled by a fully recurrent binary neural network that consists of 15–50 neurons (the number is fixed within a given simulation run). Of these, five are dedicated sensory neurons, the values of which are clamped to the sensory input and which have no input from other neurons. Four output motor neurons control the agent's motors. Network updating is synchronous. In each step, a sensory reading occurs, network activity is then updated, and a motor action is made according to the resulting activity in the designated output neurons.

The agents are equipped with a basic sensor that consists of five probes. Four probes sense the grid cell in which the agent is located and the three grid cells immediately ahead of it (FIG. 2). These probes can sense the difference between an empty cell, a cell containing a resource (either poison or food, with no distinction between the two), and the arena boundary. The fifth probe can be thought of as a smell probe, which can discriminate between food and poison if these are present in the cell that is occupied by the agent. The motor system allows the agent to go forward, turn 90° in each direction, and attempt to eat. Eating is a costly process, as it requires a time step with no other movement, in a lifetime of limited time steps.

Each agent carries a chromosome that defines the structure of its N-neuron controlling network, consisting of $N(N-5)$ real numbers that specify the synaptic weights. At the end of a generation, a phase of sexual reproduction takes place, comprising 50 reproduction events. In each of these events, two agents from the parent population are randomly selected with a probability that is proportional to their fitness (the amount of food minus poison eaten). Then, their chromosomes are crossed over and mutated to obtain the agents of the next generation. A point-crossover with a probability of 0.35 is used, after which point mutations are randomly applied to 2% of the locations in the genome. These mutations change the pertaining synaptic weights by a random value between –0.6 and +0.6. The resulting chromosomes define two agents of the next generation. Essentially, the genetic algorithm performs a search for the best synaptic-weight values in the space of all possible network architectures that might be composed of the controlling neurons. FIG. 3 shows a typical evolutionary run. The average initial population fitness is very low (around –0.05). As evolution proceeds, better controllers

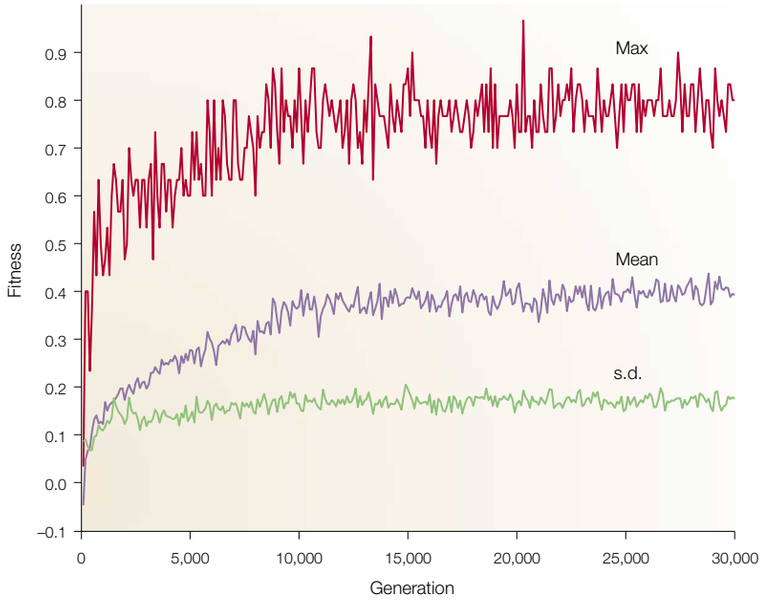

Figure 3 | **A typical evolutionary run.** Maximum (Max), mean and standard deviation (s.d.) of normalized fitness in a population plotted over 30,000 generations of an evolutionary run. Values are plotted every 100 generations. The fitness is evaluated over a single epoch for each agent, and the mean is the average of the fitness in the population (100 agents). Adapted with permission from REF. 5 © 2001 Massachusetts Institute of Technology.

emerge and both the best and average fitness in the population increase until a plateau is reached.

Current EAA studies have been able successfully to evolve artificial networks of several dozen neurons and hundreds of synapses, controlling agents that perform non-trivial behavioural tasks (see REFS 6–10 for recent reviews). These networks are less biased than conventional neural networks used in neuroscience modelling, because in many cases their architecture is not predesigned. They are the emergent result of a simplified and idealized process that models the evolution of intelligent, neurally driven life-forms. This fundamental property naturally raises the possibility of using these agents as a vehicle to study basic questions about neural processing. This potential is substantiated by two further observations. First, feasibility: the small size of the evolved networks, the simplicity of their environments and behavioural tasks, and the full information available about the model dynamics, help to make the analysis of the network's dynamics an amenable task. Second, relevancy: as the networks are evolved in biologically motivated ANIMAT environments, their analysis might potentially reveal interesting insights into the workings of biological systems.

The use of EAA models as a neuroscience research tool is a very complex and challenging scientific endeavour. Even the seemingly simple task of understanding the neural processing of small, fully transparent EAA agents is very difficult. Furthermore, the relevance of findings in EAA studies should be considered with caution (as is perhaps true of neural modelling studies in general). We should always be aware of the many simplifications that are involved in these models; although these allow us to address systems that would otherwise be too complex to investigate, it might well be that interesting and even vital components of a system are missed. But some of the results already obtained testify to the beneficial role of EAAs as a fundamental and simple method of neuroscience investigation. In the rest of this article, I review these results, many of which have not yet been brought to the attention of neuroscientists, owing to the different scientific communities involved. I attempt to answer two fundamental questions. Do EAA studies produce results that bring new insights to neuroscientific issues? And can EAA studies lead to the development of new tools for neuroscience research?

EAA studies: a neuroscience perspective

EAA studies typically evolve neurally driven computer-simulated animats or robots that solve a variety of cognitive and behavioural tasks. As such, they represent an intuitively appealing approach to modelling and studying biological nervous systems. However, do current studies really begin to realize this potential? And what can be learned from these studies? Here, I selectively review a few studies that explore specific questions that are of relevance to neuroscience. I begin with studies that have modelled simple animal systems, and proceed with models of evolution and learning. Finally, a description of evolutionary-computation investigations of cortical organization leads to a brief review and discussion of various existing models of GENOTYPE-TO-PHENOTYPE ENCODING.

Existing evolutionary-computation systems are not powerful enough to evolve neural-network controllers of detailed models of animal systems, but a few interesting investigations in this direction have already been carried out. One study has evolved neural networks that reproduce the habituation of the touch-sensitive behaviour of the nematode *Caenorhabditis elegans*[11]. The worms move backwards in response to stimulation of the head and forwards when the tail is stimulated. A comparison of data gathered from real worms and from lesion analysis of the evolved networks showed that lesioning the corresponding artificial and biological interneurons in both systems had similar global regulatory effects, resulting in disturbances in habituation. But a mismatch in the processing of the corresponding artificial and biological sensory neurons has led to the formulation and testing of a more refined EAA model, which includes the simulation of developmental events during network formation. The revised model resolved the mismatch, manifesting emergent sensory neurons that have an important role in mediating touch sensation, and others that have only a minor role, in a manner similar to that observed in real worms. This new model also indicates that such a partially asymmetrical neural mechanism might be responsible for motor response habituation, a prediction that remains to be tested.

EAA models were also used to study the evolution and development of the central pattern generator for swimming in the lamprey[12]. The evolved neural networks guide the agents' swimming by forming connections to muscles that are located along both sides of the agent's body, and generating the coordinated oscillatory patterns

ANIMATS
Artificial embodied agents.

GENOTYPE-TO-PHENOTYPE ENCODING
A mapping (transformation) that specifies the phenotype(s) created from a given genotype.

required for propulsion. These oscillatory patterns could be modulated by varying the external excitation applied to the network, resulting in varying directions and speeds of swimming. The best-evolved controllers cover a broader range of frequencies, phase lags and speeds than the original, handcrafted model[13]. Using the EAA approach, it was also possible to develop biologically plausible controllers with ranges of oscillation frequency that were closer to those actually observed in the lamprey than those produced by the handcrafted model. In agreement with the experimental findings of REF. 14, these oscillations could be produced without excitatory interneurons. Finally, the synaptic connections formed in some of the evolved agents were, at least to some extent, similar to those observed in the real lamprey[12].

These results were obtained using fitness (optimization) functions that explicitly express the desired outcome. However, interesting biological-like phenomena can emerge in EAA models even when they are not modelled explicitly. The model of REF. 5 (described above) shows this potential, in which the agents' emergent controller networks are analysed using conventional neuroscience methods of lesioning and receptive-field measurements. This analysis reveals a command neuron, the firing state of which triggers a switch between navigation and foraging behaviours that occurs immediately after the agent ingests food (recall that the agents are placed at a random location on the grid, and that they first have to navigate and find the food zone and then remain and forage in that zone). The firing of this command neuron (or a few such neurons) essentially switches between two evolved input–output subnetworks that control navigation and foraging. These EAA findings closely resemble the findings of command neurons in animals, including crayfish[15], *Aplysia*[16–18], *Clione*[19], crabs[20,21] and lobsters[22]. In some cases, the animals' behaviour is modulated by the command neurons on the basis of certain sensory stimuli[16], and in particular, as in the EAA simulation, by food arousal[17,18]. This activity has been shown to control a variety of motor repertoires, mainly by inducing different activity patterns in the same network by modulating neuronal activity[21,22] — again, in a manner similar to that found in the EAA study. Obviously, biological reality is much more complex and, despite the resemblance, there are many significant differences between these models and real biological systems. For example, chemical neuromodulation is important in command neuron activity[19,23], but is absent from the current model (this is not, however, an inherent limitation, as such neuromodulation has been incorporated in other EAA studies; for example, REFS 24,25). Even though biological reality is much richer than that of a simple EAA model, these studies show that networks that emerge in EAA systems can manifest interesting biological-like characteristics and provide new computational insights.

There are two main motivations for studying biological systems by using EAA models. The first stems from the observation that biologically relevant neural-network models should be studied in a comprehensive system that contains not only the networks themselves, but also the 'bodies' in which they reside; that is, the agents' sensors and motors, and the environment in which the agents act. As shown in various EAA studies, this embodiment is of paramount importance in providing constraints that reduce the degeneracies involved in the neural-to-behavioural mappings[26]. Moreover, EAA agents can use the evolved motor behaviours to augment their sensory processing; for example, by moving and turning around such that important objects in the environment are viewed from a fixed angle, making their recognition much simpler[27,28].

The second motivation stems from the recognition that EAA studies represent a natural computational framework in which to study the interaction between learning and evolution, two prime adaptation mechanisms that occur on different timescales, but are interleaved and interact in complex ways (see REF. 29 for a review). A primary focus of EAA studies of learning and evolution has been the Baldwin effect, in which learning can influence the course of evolution even if learned traits are not inherited[30,31]. Below, we review just one example of EAA research into this interplay between learning and evolution.

In this study, a population of EAAs was subject to both evolutionary and learning processes[32]. Each agent's controller is composed of two subnetworks; one network guides its sensorimotor processing and the other is a 'teacher' network. Both networks are encoded in the agent's genome, and receive the same set of inputs from the environment. The teacher network processes these inputs to calculate desired responses, which are then used to modify the synaptic connections of the sensorimotor network using a supervised learning algorithm. In a typical exploration task, the combination of learning and evolution in these agents allows them to achieve significantly higher performance than agents with a similar genetically encoded sensorimotor subnetwork but without learning; that is, without the teacher subnetwork. The key to the success of the learning-able agents is their ability to develop a genetically inherited predisposition for learning. This predisposition stems from the selection of initial weights at birth, which guides behaviour to select the right set of inputs, so 'channelling' learning in a successful direction. This power of evolution to select specific emergent learning predispositions points to the potential pitfalls of studying learning in isolation, as is done at times in conventional neural networks and connectionist models.

EAA studies have been used to study the evolution of learning itself. Such an investigation can span several levels of neuroscientific research[33]. This is shown in a study that evolved (near-)optimal neuronal learning rules in a simple EAA model of reinforcement learning in bumblebees[34]. After the neural modelling study of REF. 35, this EAA investigation studied bee foraging in a simulated three-dimensional arena of blue and yellow flowers. During its flight, the bee processes visual information about the flower arena that it sees to determine its next flight heading. On landing, it receives a reward according to the nectar produced by the flower on which it has landed. This reward is then used by the bee as a reinforcement learning cue to modify the synaptic efficacies

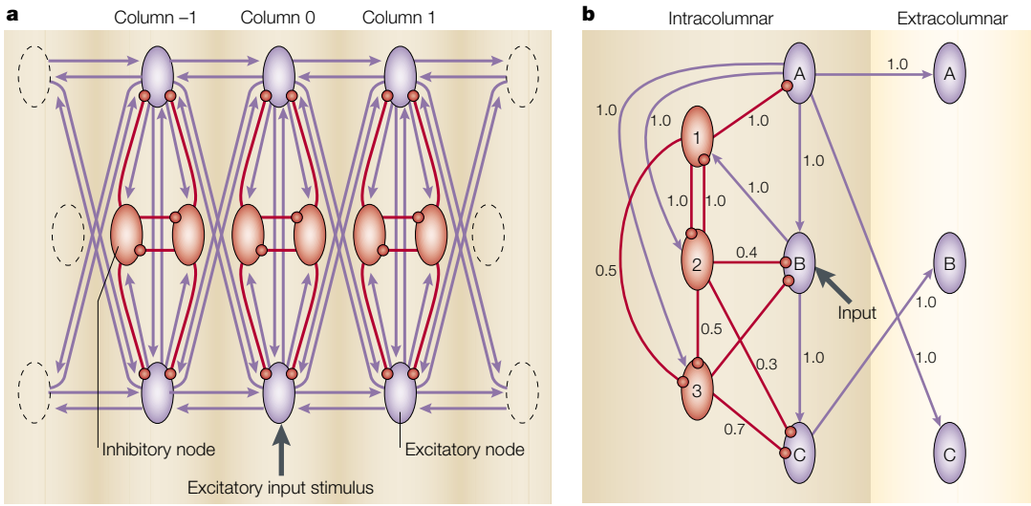

Figure 4 | **Evolution of cortical circuits. a** | An example initial cortical circuit before evolution with two excitatory and two inhibitory nodes per column. Each node represents a small population of similar neurons. Initial synaptic strengths were random, and different evolutionary runs started with different numbers of excitatory/inhibitory nodes. **b** | A diagram of an evolved cortical circuit that produces a 'Mexican hat pattern' of activity to a point stimulus in the absence of horizontal, intercolumnar inhibitory connections. Only the magnitudes of inhibitory weights are shown (red connections); they are multiplied by a negative gain constant when used. Connections with weights of <20% of the maximum were omitted. Connections for all columns are the same. Lateral inhibition arises because excitation of some excitatory neurons (nodes B and C) in a stimulated column causes other excitatory neurons (node A) in the same column (those sending lateral excitatory connections to other columns) to turn off, decreasing lateral excitation of adjacent columns and causing their mean activation levels to fall. Inhibitory neurons evolved not only to inhibit intracolumnar excitatory neurons, but also to inhibit each other in a highly specific pattern. Adapted from REF. 43 © 2001 IEEE Press.

of its controlling network. Unlike the study of REF. 35, in which the synaptic learning rules that govern these modifications are pre-specified, the learning rules themselves (and not just the synaptic weights) undergo evolution. These evolved synaptic-plasticity rules include a new component of heterosynaptic Hebbian learning[36,37] that was not specified in the previous handcrafted solution[35], giving rise to varying exploration/exploitation levels. On a higher level of description, these synaptic microdynamics give rise to risk-aversive behaviour, providing a novel, biologically founded, parsimonious explanation for risk aversion. So, even simple EAA models can produce complex emergent behaviours that were not explicitly specified in the model.

Evolutionary-computation optimization methods can be used directly to challenge a variety of important open questions in neuroscience. For example, it is well known that a localized excitatory stimulus applied directly to the cerebral cortex can produce a surrounding peristimulus inhibitory zone (a 'Mexican hat pattern' of activity). This has long been viewed as surprising by some, because lateral intracortical connections within the 150–250-μm distances involved are predominantly excitatory (asymmetrical synapses: pyramidal/stellate neuron to pyramidal/stellate neuron)[38–42]. In this context, a recent study used a genetic algorithm to investigate whether a neuronal circuit for cortical columns could be evolved that produces peristimulus inhibition under the constraint that the only horizontal, intercolumnar synaptic connections that are permitted are excitatory[43]. Starting with columnar neuronal circuits that have arbitrary, randomly generated excitatory and inhibitory synaptic strengths, it proved possible to evolve, within at most a few thousand generations, neuronal circuits that produce Mexican hat patterns of activity (FIG. 4). The apparent lateral inhibitory effects that evolved were due to the turning off of baseline horizontal spread of excitatory activity between neurons in neighbouring columns. This result is interesting, not just because the evolutionary process discovered a novel candidate neural circuit for the cerebral cortex, but more importantly because it indicates a new approach to generating hypotheses about the nature of complex neural circuitry in general, through a simulated evolutionary process.

A kind of 'reverse' approach has shown that 'indirect', biologically inspired genotype-to-phenotype encoding allows the successful evolution of a variety of basic neural-network architectures that are assumed to participate in cortical neural processing[44]. In contrast to 'direct' encodings, in which every gene explicitly specifies a synaptic connection, 'indirect' encodings include a program for determining the network architecture or its connections' weight values in a more efficient, compact manner. This approach serves a dual role. From a computational perspective, its aim is to find efficient genotype-to-phenotype encodings, a key to the evolution of smart agents. It provides compact genetic encodings of complex networks, and filtering out of genetic changes[45]. From a neuroscience perspective, its aim is to test different hypotheses about how the architecture and operation of various biological neural networks might be specified by the genome. The work on developing indirect encodings has attracted ample efforts, focusing on a few different avenues (BOX 1).

Box 1 | **Indirect genotype-to-phenotype encodings**

'Grammar rewriting encodings' use a set of rewriting rules that are encoded in the genome. For example, in REF. 68, the genome contains numerous blocks. Each block has five elements and is interpreted as a rewriting rule that states that the first element in the block should be transcribed to a matrix composed of the next four elements in the block. Through such an iterative decoding process, a matrix that specifies the network architecture is formed (and the synaptic efficacies are then determined by other mechanisms, such as learning). Simple grammar rewriting encodings typically generate restricted tree-like network architectures, but by using graph grammars, it is possible to develop more general, recurrent networks[69,70]. The latter have been used to develop encodings that lead to the emergence of modular subnetworks — repetitive building blocks, mimicking cortical columnar structures. Such grammar-like encodings generate fairly compact genomes and hence reduce the search space of possible solutions. A variant in REF. 71 allows the evolution of compact target networks by including a complexity term in the fitness function.

In developmental, ontogenetic encodings ('geometric grammars'), the genome expresses a program for cell division and axonal migration that determines the phenotypic neural architecture[72–74]. In these encodings, the objects that are undergoing development are localized on a two- or three-dimensional space, allowing for context-dependent effects from neighbouring neurons, and the developmental program has a more biological flavour. However, the genomes generated are less compact than those generated by encoding graph grammars (scaling linearly with network size), and do not strongly bias towards the evolution of modular networks. The temporal dimension of such ontogenetic development has been studied by encoding 'maturation times' that regulate the expression of different blocks in the genome[74]. This mechanism can successfully select genomes that dictate early maturation of functional versus dysfunctional blocks. 'Layering' ontogenesis in time delays full functionality to later stages in the agent's life, and so might damage its fitness, but it might enhance genetic variation by sustaining mutations that occur in genetic blocks expressed late in maturation. A few developmental encodings have aimed at developing more biologically motivated 'regulatory' encodings, specifying the development of a multicellular organism[45,75,76]. The identity of the subset of genes that is active at each moment in a given cell is determined by a complex interaction with the 'transcription factors' that it receives from the environment and from other cells. However, these models require extensive computational resources and are still unable to evolve agents that solve complex tasks.

'Compound encodings' are newly developed genetic encodings that flexibly encode both the synaptic weights and their learning rules. They have been shown to be superior to direct synaptic encodings, but these results await further corroboration in a wider set of EAA models[46,77]. The efficacy of such encodings might be enhanced by self-organizing processes that refine the developing networks by exploiting regularities in the environment; for example, by incorporating activity-based pruning into the developmental programs[78]. Their efficacy might be further enhanced by evolutionary autonomous agent (EAA) models that incorporate neuromodulatory agents. Such a diffusible agent has already been used to dynamically modulate the neurons' transfer function in a concentration-dependent manner, enabling efficient evolution of more compact networks than those obtained with direct encodings[24]. In a manner analogous to REF. 34, neuromodulation could be harnessed to guide learning in agents. Last, self-organizing compression encodings could be used adaptively to perform the evolutionary search in a low-dimensional subspace of the original search space, thus expediting convergence to good solutions (S. Bushy and E. R., unpublished observations).

However, finding efficient indirect encodings still remains an extremely important open problem. The superiority of existing indirect encodings over direct ones has not yet been shown in a convincing manner, partly because of the absence of examples of agents with indirect encodings that solve complex tasks that were otherwise unsolvable (with direct encodings or by handcrafted solutions)[46]. In addition, some of these encodings do not scale up well with network size[47]. The idea that an encoding that successfully captures some of the essential computational principles of biological encodings could lead to a breakthrough in our ability to evolve complex agents is certainly compelling, but it remains to be seen whether this can be done. If, however, some success is obtained, then simulating such developmental processes in EAAs will probably teach us a lot about the organization and functioning of biological neural networks.

Analysis of neural information processing in EAAs

This section reviews research that analyses the evolved controller networks. The dual goal of this research is to uncover principles of neural processing in animat and biological nervous systems, and to develop new methods for their analysis.

A series of studies has developed a rigorous, quantitative analysis of the dynamics of central pattern generator (CPG) networks that have evolved for locomotion[26,48]. The evolved networks are very small, comprising three, four or five neurons. A high-level description of the dynamics of these CPG networks was developed, based on the concept of a dynamic module: a set of neurons that have a common temporal behaviour, making a transition from one quasi-stable state of firing to another together. The evolved networks can be decomposed to a varying number of multi-stable dynamic modules that are traversed through successive destabilizations. In some networks, the dynamic modules do not remain fixed, but change over a slow timescale. Dynamic modules give new insights into CPG operation, describing them in terms of a finite state machine, and allowing a rigorous analysis of their robustness to parameter variations. They provide one possible concrete realization of Getting's hypothesis that biological CPGs are constructed from basic dynamical 'building blocks'[49,50]. Interestingly, in some cases, the dynamic modules could be assigned specific functional roles, but in others this assignment was not possible. This observation touches on the fundamental topic of 'understanding' neural processing by localizing and assigning specific procedures and functions to component subnetworks, an issue that we shall return to later.

Examining the variety of CPG networks evolved, it becomes evident that there is degeneracy in the mapping between structural and functional levels[26]. That is, many different network architectures give rise to the same functional level of rhythmic activities and, consequently, to similar walking performance. Such degeneracies might be ubiquitous in biological systems, and should be considered by neuroscientists who aim to construct biologically realistic models of CPGs[26]: owing to technical recording difficulties, the biophysical properties of neurons are usually measured across many individuals, and models are constructed using values from several animals. In view of these degeneracies, this construction of 'chimaera-like' model networks might be artificial, and could lead to the 'brittleness' that is often observed in realistic neuronal simulations[51], in which small parameter variations can lead to large changes in model dynamics. The negative findings of this EAA study are also of interest: some of the evolved five-neuron CPGs were not functionally decomposable[26]. Similarly, in spite of the progress that has been made in neuroscience in analysing and

modelling biological CPGs, their functional decomposition remains enigmatic[50]. This is perhaps not surprising — after all, behaviour rather than circuit architecture is selected for in evolution[52].

The dynamic-module analysis was carried out in very small networks with a regular rhythmic pattern of activity, and its application to significantly larger EAA networks with less regular dynamics remains a daunting task. How can we analyse these networks? Several studies have made use of a variety of conventional neuroscience techniques to this end. Here, I briefly review a few of these studies, each involving the use of a different technique.

In REF. 53, the activity of the internal (hidden-layer) neurons in a network as a function of a robot's location and orientation was charted by a simple form of receptive-field measurement. The function of these intermediate neurons was generally highly distributed and dependent on previous states, but a certain interneuron that had an important role in path planning was also identified. Others have systematically clamped neuronal activity and studied its effects on a robot's behaviour (for example, inducing rotation, straight-line motion, or more complex behaviours such as smooth tracking of moving targets)[54]. The 'command' neurons described in the previous section[5] were discovered by single-lesion analysis. Finally, a more 'procedural' kind of ablation to the network, in which different processes (and not just units or links) are systematically cancelled out, was used recently in REF. 55. Overall, these studies have provided only glimpses of the processing in these networks. Moreover, an integrative EAA study that systematically analyses a neural network by using all these methods together and comparing them is still lacking.

In neuroscience, assessing the importance of single neurons or cortical areas to specific tasks is traditionally done either by assessing the deficit in performance after lesioning a specific area, or by recording the activity in the area during behaviour. These classical methods suffer from two fundamental flaws[56]. First, they do not take into account the probability that there are complex interactions among elements in the system. For example, if two neurons have a high degree of redundancy, lesioning either one alone will not reveal its influence. Second, they are mostly qualitative measures, lacking the ability to quantify precisely the contribution of a unit to the performance of the organism and to predict the effect of new, multiple-site lesions. The relative simplicity of EAAs and the availability of full information about the network's structure and dynamics make EAA models an ideal test-bed for studying neural processing. In this framework, a rigorous, operative definition of neurons' (or cortical regions') contributions to an organism's performance in various tasks, and a novel functional contribution algorithm (FCA) to measure them, have recently been presented[56]. This operative definition allows an accurate prediction of the performance of EAA agents after multi-lesion damage, and yields, at least in principle, precise quantification of the distribution of processing in the network, a fundamental open question in neuroscience[57,58].

Task / Neuron	1	2	…	*P*
1	C_{11}	C_{12}	…	C_{1P}
2	C_{21}	C_{22}	…	C_{2P} → S_2
3	C_{31}	C_{32}	…	C_{3P}
…	…	…	…	…
N	C_{N1}	C_{N2}	…	C_{NP}
	↓ L_1			

Figure 5 | **Contributions of neurons (units) to tasks.** Consider an agent with a controlling network of *N* interconnected units that performs *P* functional tasks. The figure shows a contribution matrix, where C_{ik} is the contribution of element *i* to task *k*, computed by measuring the agent's performance after different multiple lesions. The functional contribution algorithm (FCA) finds the contribution values, C_{ik}, that provide the best performance prediction for new, multiple-site lesions. The localization, L_k, of task *k* can now be defined as a deviation from equipotentiality along column *k* (for example, L_1); similarly, S_i, the specialization of neuron *i*, is the deviation from equipotentiality along row *i* of the matrix (for example, S_2).

To understand the concept of 'contributions', consider an agent (either natural or artificial) with a controlling network of *N* interconnected neurons (or, more generally, units) that performs a set of *P* different functional tasks in the environment in which it is embedded. When addressing this question of who does what, it is natural to think in terms of a contribution matrix, where C_{ik} is the contribution of element *i* to task *k*, as shown in FIG. 5. The data that are analysed to compute the contribution matrix are gathered by inflicting a series of multiple lesions on the agent's network. (Obviously, there are different ways of lesioning networks; for example, by knocking out neurons, or by severing all incoming or all outgoing synapses from a neuron. The latter option is assumed in the FCA.) After each lesion, the resulting performance of the agent in different tasks is measured. Given these data, the FCA finds the contribution values, C_{ik}, that provide the best performance prediction for new, multiple-site lesions. Following the spirit of REF. 59, the localization, L_k, of task *k* can now be defined as a deviation from equipotentiality along column *k* (for example, L_1 in FIG. 5); similarly, S_i, the specialization of neuron *i*, is the deviation from equipotentiality along row *i* of the matrix (for example, S_2 in FIG. 5).

The FCA algorithm was applied to the analysis of the EAA neurocontrollers evolved in REF. 5, which are recurrent neural networks with complex interactions between the elements, and provided a precise prediction of the effects of new multiple lesions[56]. This remarkable performance was obtained primarily by using a general monotonic but nonlinear performance-prediction function. However, more complex EAA neurocontrollers cannot be accurately described using an FCA based on contributions of single units only. Precise multi-lesion prediction in these networks requires the consideration of the contributions from further, functionally important conjunctions of the basic units. These findings

strongly indicate that the classic, conventional thinking in neuroscience of aiming to decompose the processing of various tasks to a set of individual distinct regions is an oversimplification. It should also be noted that the ongoing development of more efficient derivatives of the FCA algorithm would not have been possible without the body of data provided by the EAA investigations.

Multi-lesion analysis algorithms such as the FCA are important in neuroscience for the analysis of reversible inactivation experiments, combining reversible neural cooling deactivation with behavioural testing of animals[60]. They can also be used for the analysis of transcranial magnetic stimulation (TMS) studies, which aim to induce multiple transient lesions and study their cognitive effects[61]. Another possible use is the analysis of functional imaging data by assessing the contributions of each element to the other; that is, extending previous studies of 'effective' connectivity that use linear models (for example, REF. 62). Applying algorithms such as the FCA should be useful for obtaining insights into the organization of natural nervous systems, and settling the debate about local versus distributed computation.

Discussion

In his illuminating treatise on the fictional autonomous 'vehicle' agents[63], Valentino Braitenberg makes two important observations: "in most cases our analysis [of 'type-6' artificial brains] would fail altogether: the wiring that produces their behaviour may be so complicated and involved that we will never be able to isolate a simple scheme. And yet it works …". Indeed, EAA models work and are a relevant neuroscience research tool. Yet, even the simple models that are studied at present can be fairly complex, and their analysis represents a significant challenge. But if this is so with regard to these much simplified systems, what about the task of understanding real, natural nervous systems? The 'central dogma' of addressing the latter challenge in neuroscience and cognitive sciences research has been the knowledge-based engineering approach. This approach aims to analyse and conceptualize neural information processing in terms of hierarchical, top–down operations that manipulate representations, as exemplified in Marr's work on vision[64]. However, the necessity and computational value of such representations have been questioned by robotics and adaptive-behaviour researchers in recent years (for example, REF. 65). An interesting review of numerous EAA studies that have evolved visual-processing agents was recently presented[66]. Like the Braitenberg vehicle robots, which do not need representations, all the agents in the studies that are surveyed in this review did not use representations in the conventional sense (that is, neural patterns of firing that correspond to concrete or abstract concepts pertaining to the agent and its environment), at least as far as the researchers' analysis techniques could tell. Rather, understanding the controller's dynamics requires treating the agent and its environment as a coupled, embedded, dynamical system[66,67]. There is considerable evidence to support the possibility that the same is true for animals[66].

Although I hope that the examples presented are convincing, we should not underestimate the difficulties of using EAAs in neuroscience — the most important being that it remains to be seen whether the EAA paradigm can generate really complex agents on the scale of ANIMATE systems, and if so, whether we will be able to analyse them. Current EAA modelling is limited in important ways: first, most current EAA models use simple binary neurons, and the investigation of EAAs driven by spiking networks and temporal synaptic dynamics seems to be a necessary step if we want to make closer contact with animate neural processing. Second, even the brains of the simplest biological organisms use hundreds and thousands of neurons with thousands and millions of synapses. It is obvious that we will not be able to evolve neural controllers of this magnitude using simple direct encodings. The development of new genotype-to-phenotype encodings is crucial for the evolution of smarter agents. Third, current EAA models use very simplistic forms of embodiment; that is, very rudimentary sensors and motors in elementary environments. More elaborate and realistic sensors and motors must be developed if we really wish to study sensorimotor processing. Moreover, these sensors and motors should probably be co-evolved with their controllers. And finally, as the scale and the complexity of the evolved networks grows, the challenge of finding new ways to analyse the evolved networks will become more complicated. In its current nascent stage, the EAA paradigm is still fairly limited, but a gradual, incremental approach for its further development is feasible and within our reach. We should also bear in mind that, until now, there have been many fewer EAA studies of neuroscience questions than more traditional computational neuroscience investigations.

This brings us to what is perhaps the chief criticism of EAA in neuroscience — the idea that evolution can take many paths and directions, and hence the findings observed in EAA models might not teach us anything significant about biological nervous systems. As outlined throughout this paper, there are various reasons to believe that this is not the case, and that biologically relevant principles of neural information processing can best be studied in these models. But the importance of EAA research goes beyond that: to my mind, its primary value for neuroscience is first and foremost its ability to serve as a very simple, but emergent, accessible and concrete, even if artificial, test-bed for thinking about neural processing principles and for developing new methods for deciphering them. EAAs represent a promising way to make neuroscience modelling as simple as possible, without simplifying it so much that it ceases to be useful. "As simple as possible" might turn out eventually to be fairly or very complex, but even so, I believe that EAA studies are one of our best bets in this quest.

Many avenues for further development of EAA studies await. But the combination of the results reviewed here, the clear challenges that await in the near future, and the continuing fast growth of computing resources that open new possibilities for more realistic EAA modelling, give us grounds for confidence that the study of EAAs as a neuroscience research methodology is a promising and timely endeavour.

ANIMATES
Biological embodied agents.

1. Mitchell, M. *An Introduction to Genetic Algorithms* (MIT Press, Cambridge, Massachusetts, 1996).
2. Langton, C. *Artificial Life: an Overview* (MIT Press, Boston, Massachusetts, 1997).
3. Fogel, D. B. *Evolutionary Computation — Toward a New Philosophy of Machine Intelligence* (IEEE Press, Piscataway, New Jersey, 1999).
4. Adami, C. *Introduction to Artificial Life* (Springer, New York, 1998).
5. Aharonov-Barki, R., Beker, T. & Ruppin, E. Emergence of memory-driven command neurons in evolved artificial agents. *Neural Comput.* **13**, 691–716 (2001).
This study was one of the first to evolve general, recurrent controlling networks, manifesting the spontaneous emergence of biological-like command neurons that switch between distinct behavioural modes.
6. Meyer, J. A. & Guillot, A. From SAB90 to SAB94: four years of animat research. *Proc. 3rd Int. Conf. Simul. Adaptive Behav.* (eds Cliff, D., Husbands, P., Meyer, J. A. & Wilson, S. W.) 2–11 (MIT Press, Cambridge, Massachusetts, 1994).
7. Kodjabachian, J. & Meyer, J. A. Evolution and development of neural controllers for locomotion, gradient-following and obstacle-avoidance in artificial insects. *IEEE Trans. Neural Netw.* **9**, 796–812 (1998).
8. Yao, X. Evolving artificial neural networks. *Proc. IEEE* **87**, 1423–1447 (1999).
9. Guillot, A. & Meyer, J. A. From SAB94 to SAB2000: what's new, animat? *Proc. 6th Int. Conf. Simul. Adaptive Behav.* (eds Meyer, J. A., Berthoz, A., Floreano, D., Roitblat, H. L. & Wilson, S. W.) 2–11 (MIT Press, Cambridge, Massachusetts, 2000).
10. Guillot, A. & Meyer, J. A. The animat contribution to cognitive systems research. *J. Cogn. Syst. Res.* **2**, 157–165 (2001).
11. Cangelosi, A. & Parisi, D. A neural network model of *Caenorhabditis elegans*: the circuit of touch sensitivity. *Neural Process. Lett.* **6**, 91–98 (1997).
12. Ijspeert, A. J., Hallam, J. & Willshaw, D. Evolving swimming controllers for a simulated lamprey with inspiration from neurobiology. *Adaptive Behav.* **7**, 151–172 (1999).
This paper describes an elaborate study that evolved swimming controllers, showing the power of EAAs in generating biologically relevant network models that outperform existing handcrafted ones.
13. Ekeberg, O. A combined neuronal and mechanical model of fish swimming. *Biol. Cybern.* **69**, 363–374 (1993).
14. Jung, R., Kimmel, T. & Cohen, A. H. Dynamical behavior of a neural network model of locomotor control in the lamprey. *J. Neurophysiol.* **75**, 1074–1086 (1996).
15. Edwards, D. H., Heitler, W. J. & Krasne, F. B. Fifty years of command neurons: the neurobiology of escape behavior in the crayfish. *Trends Neurosci.* **22**, 153–161 (1999).
16. Xin, Y., Weiss, K. R. & Kupfermann, I. A pair of identified interneurons in *Aplysia* that are involved in multiple behaviors are necessary and sufficient for the arterial-shortening component of a local withdrawal reflex. *J. Neurosci.* **16**, 4518–4528 (1996).
17. Nagahama, T., Weiss, K. & Kupfermann, I. Body postural muscles active during food arousal in *Aplysia* are modulated by diverse neurons that receive monosynaptic excitation from the neuron CPR. *J. Neurophysiol.* **72**, 314–325 (1994).
18. Teyke, T., Weiss, K. & Kupfermann, I. An identified neuron (CPR) evokes neuronal responses reflecting food arousal in *Aplysia*. *Science* **247**, 85–87 (1990).
19. Panchin, Y. V. *et al.* Control of locomotion in the marine mollusc *Clione limacina*. XI. Effects of serotonin. *Exp. Brain Res.* **109**, 361–365 (1996).
20. Norris, B. J., Coleman, M. J. & Nusbaum, M. P. Recruitment of a projection neuron determines gastric mill motor pattern selection in the stomatogastric nervous system of the crab, *Cancer borealis*. *J. Neurophysiol.* **72**, 1451–1463 (1994).
21. DiCaprio, R. A. An interneurone mediating motor programme switching in the ventilatory system of the crab. *J. Exp. Biol.* **154**, 517–535 (1990).
22. Combes, D., Meyrand, P. & Simmers, J. Motor pattern specification by dual descending pathways to a lobster rhythm-generating network. *J. Neurosci.* **19**, 3610–3619 (1999).
23. Brisson, M. T. & Simmers, J. Neuromodulatory inputs maintain expression of a lobster motor pattern generating network in a modulation-dependent state: evidence from long-term decentralization *in vitro*. *J. Neurosci.* **18**, 2212–2225 (1998).
24. Husbands, P., Smith, T., Jacobi, N. & Oshea, M. Better living through chemistry: evolving GasNets for robot control. *Connection Sci.* **10**, 185–210 (1998).
25. Ishiguro, A. Evolving an adaptive controller for a legged robot with dynamically-rearranging neural networks. *Proc. 6th Int. Conf. Simul. Adaptive Behav.* (eds Meyer, J. A., Berthoz, A., Floreano, D., Roitblat, H. L. & Wilson, S. W.) (MIT Press, Cambridge, Massachusetts, 2000).
26. Beer, R. D., Chiel, H. J. & Gallagher, J. C. Evolution and analysis of model CPGs for walking. II. General principles and individual variability. *J. Comput. Neurosci.* **7**, 119–147 (1999).
This paper and reference 48 show the use of EAA models as a tool for generating and analysing simple and tractable dynamic network systems, introducing the concept of dynamical modules as a computational analogue of Getting's classical dynamically varying modules.
27. Scheier, C., Pfiefer, R. & Kunyioshi, Y. Embedded neural networks: exploiting constraints. *Neural Netw.* **7–8**, 1551–1569 (1998).
This paper describes a set of experiments that show the importance and significance of studying neural-network controllers of agents embodied in their environment.
28. Beer, R. D. Dynamical approaches to cognitive science. *Trends Cogn. Sci.* **4**, 91–99 (2000).
29. Nolfi, S. & Floreano, D. Learning and evolution. *Auton. Robots* **7**, 89–113 (1999).
30. Hinton, G. E. & Nowlan, S. How learning can guide evolution. *Complex Syst.* **1**, 495–502 (1987).
31. Miller, G. F. & Todd, P. Exploring adaptive agency. I. Theory and methods for simulating the evolution of learning. *Proc. 1990 Connectionist Models Summer School* (eds Touretzky, D. S., Elman, J. L., Sejnowski, T. J. & Hinton, G. E.) 65–80 (Morgan Kaufmann, San Mateo, California, 1990).
32. Nolfi, S. & Parisi, D. Learning to adapt to environments in evolving neural networks. *Adaptive Behav.* **5**, 75–98 (1997).
33. Churchland, P. S. & Sejnowski, T. J. *The Computational Brain* (MIT Press, Boston, Massachusetts, 1989).
34. Niv, Y., Joel, D., Meilijson, I. & Ruppin, E. Evolution of reinforcement learning in foraging bees in neural terms. *10th Annu. Comput. Neurosci. Meet. CNS2001* (Monterey, California, 2001). *Neurocomputing* (in the press).
35. Montague, P. R., Dayan, P., Person, C. & Sejnowski, T. J. Bee foraging in uncertain environments using predictive Hebbian learning. *Nature* **377**, 725–728 (1995).
36. Schacher, S., Wu, F. & Sun, Z.-Y. Pathway-specific synaptic plasticity: activity-dependent enhancement and suppression of long-term heterosynaptic facilitation at converging inputs on a single target. *J. Neurosci.* **17**, 597–606 (1997).
37. Vogt, K. E. & Nicoll, R. E. Glutamate and γ-aminobutyric acid mediate a heterosynaptic depression at mossy fiber synapses in the hippocampus. *Proc. Natl Acad. Sci. USA* **96**, 1118–1122 (1999).
38. Fisken, R., Garey, L. & Powell, T. Patterns of degeneration after intrinsic lesions of the visual cortex of the monkey. *Brain Res.* **53**, 208–213 (1973).
39. Hess, R., Negishi, K. & Creutzfeldt, O. The horizontal spread of intracortical inhibition in visual cortex. *Exp. Brain Res.* **22**, 415–419 (1975).
40. Douglas, R. & Martin, K. in *The Synaptic Organization of the Brain* (ed. Shepherd, G. M.) 389–438 (Oxford Univ. Press, Oxford, 1997).
41. Reggia, J., Autrechy, L., Sutton, G. & Weinrich, M. A competitive distribution theory of neocortical dynamics. *Neural Comput.* **4**, 287–317 (1992).
42. Mountcastle, V. *Perceptual Neuroscience: Cerebral Cortex* (Harvard Univ Press, Boston, Massachusetts, 1998).
43. Ayers, D. & Reggia, J. A. Evolving columnar circuitry for lateral cortical inhibition. *Proc. INNS–IEEE Int. Joint Conf. Neural Netw.* 278–283 (IEEE Press, Washington DC, 2001).
Going beyond EAAs, this study shows the power of evolutionary computation to search for candidate cortical columnar architectures that might best subserve observed patterns of cortical activity.
44. Rolls, E. T. & Stringer, S. M. On the design of neural networks in the brain by genetic algorithms. *Prog. Neurobiol.* **61**, 557–579 (2000).
45. Dellaert, F. & Beer, R. D. Toward an evolvable model of development for autonomous agent synthesis. *Proc. 4th Conf. Artif. Life* (eds Brooks. R. & Maes, P.) 246–257 (MIT Press, Cambridge, Massachusetts, 1994).
46. Floreano, D. & Urzelai, J. Neural morphogenesis, synaptic plasticity and evolution. *Theory Biosci.* **120**, 223–238 (2001).
47. Cliff, D. & Miller, G. F. Co-evolution of pursuit and evasion. II: Simulation methods and results. *Proc. 4th Int. Conf. Simul. Adaptive Behav.* (eds Maes, P. *et al.*) 506–515 (MIT Press, Cambridge, Massachusetts, 1996).
48. Chiel, H. J., Beer, R. D. & Gallagher, J. C. Evolution and analysis of model CPGs for walking. I. Dynamical modules. *J. Comput. Neurosci.* **7**, 99–118 (1999).
49. Getting, P. Emerging principles governing the operations of neural networks. *Annu. Rev. Neurosci.* **12**, 185–204 (1989).
50. Marder, E. & Calabrese, R. L. Principles of rhythmic motor pattern generation. *Physiol. Rev.* **76**, 687–717 (1996).
51. Marder, E. & Abbott, L. F. Theory in motion. *Curr. Opin. Neurobiol.* **5**, 832–840 (1995).
52. Dumont, J. P. C. & Robertson, R. M. Neuronal circuits: an evolutionary perspective. *Science* **233**, 849–853 (1986).
53. Floreano, D. & Mondada, F. Evolution of homing navigation in a real mobile robot. *IEEE Trans. Syst. Man Cybern. B* **26**, 396–407 (1996).
54. Harvey, I., Husbands, P. & Cliff, D. Seeing the light: artificial evolution, real vision. *Proc. 3rd Int. Conf. Simul. Adaptive Behav.* (eds Cliff, D., Husbands, P., Meyer, J. A. & Wilson, S. W.) 392–401 (MIT Press, Cambridge, Massachusetts, 1994).
55. Stanley, K. O. & Miikkulainen, R. Evolving neural networks through augmenting topologies. Technical Report No. AI01-290 <http://citeseer.nj.nec.com/stanley01evolving.html> (University of Texas at Austin, 2001).
56. Aharonov, R., Meilijson, I. & Ruppin, E. in *Advances in Neural Information Processing Systems* Vol. 13 (eds Leen, T. K., Dietterich, T. G. & Volker, T.) 3–9 (MIT Press, Boston, Massachusetts, 2001).
This paper uses the EAA paradigm to develop a new functional-contribution analysis method, which assesses how different behavioural tasks are localized and represented across a controller network.
57. Wu, J., Cohen, L. B. & Falk, C. X. Neuronal activity during different behaviors in *Aplysia*: a distributed organization? *Science* **263**, 820–822 (1994).
58. Thorpe, S. in *The Handbook of Brain Theory and Neural Networks* (ed. Arbib, M. A.) 549–553 (MIT Press, Boston, Massachusetts, 1995).
59. Lashley, K. S. *Brain Mechanisms in Intelligence* (Univ. Chicago Press, Chicago, 1929).
60. Lomber, S. G. The advantages and limitations of permanent or reversible deactivation techniques in the assessment of neural function. *J. Neurosci. Methods* **86**, 109–117 (1999).
61. Walsh, V. & Cowey, A. Transcranial magnetic stimulation and cognitive neuroscience. *Nature Rev. Neurosci.* **1**, 73–79 (2000).
62. Friston, K. J., Frith, C. D. & Frackowiak, R. S. J. Time-dependent changes in effective connectivity measured with PET. *Hum. Brain Imaging* **1**, 69–79 (1993).
63. Braitenberg, V. *Vehicles, Experiments in Synthetic Psychology* (MIT Press, Cambridge, Massachusetts, 1984).
64. Marr, D. *Vision* (W. H. Freeman, New York, 1982).
65. Brooks, R. A. Intelligence without representations. *Artif. Intell.* **47**, 139–159 (1991).
66. Cliff, D. & Noble, S. Knowledge-based vision and simple visual machines. *Phil. Trans. R. Soc. Lond. B* **352**, 1165–1175 (1997).
This paper reviews a series of EAA studies of visual processing to conclude that this task is performed in a distributed, complex manner that is fundamentally different from the top–down, hierarchical processing used in classical computational/engineering approaches.
67. Cliff, D. in *The Handbook of Brain Theory and Neural Networks* (ed. Arbib, M. A.) 626–631 (MIT Press, Boston, Massachusetts, 1995).
68. Kitano, H. Designing neural networks using genetic algorithms with graph generation system. *Complex Syst.* **4**, 461–476 (1990).
69. Gruau, F. Automatic definition of modular neural networks. *Adaptive Behav.* **3**, 151–183 (1994).
This paper presents graph grammar phenotype-to-genotype encodings that achieve highly compact encodings, leading to modular network architectures.
70. Kodjabachian, J. & Meyer, J. A. Evolution and development of modular control architectures for 1-D locomotion in six-legged animats. *Connect. Sci.* **10**, 211–254 (1998).
71. Zhang, B.-T. & Muhlenbein, H. Evolving optimal neural networks using genetic algorithms with Occam's razor. *Complex Syst.* **7**, 199–220 (1993).
72. Belew, R. K. in *Advances in Neural Information Processing (NIPS5)* (ed. Cowan, J.) 99–106 (Morgan Kaufmann, San Mateo, California, 1993).
73. Cangelosi, A., Parisi, D. & Nolfi, S. Cell division and migration in a 'genotype' for neural networks. *Network* **5**, 497–515 (1994).
This paper was one of the first to study biologically inspired 'developmental' encodings that are based on a genetic program composed of steps such as cell division and axonal growth. This lays the foundations for more detailed 'regulatory' encodings, in which the identity of the expressed gene subset is determined by interacting transcription factors.

74. Nolfi, S. & Parisi, D. Evolving artificial neural networks that develop in time. *Eur. Conf. Artif. Life* 353–367 (Springer, Berlin, 1995).
This is one of a series of papers that exploit the potential of EAA models to study the interesting and complex interactions between learning and evolution (originally cast in the Baldwin effect).
75. Cangelosi, A. & Elman, J. L. Gene regulation and biological development in neural networks: an exploratory model. Technical Report No. CRL–UCSD <http://www.citeseer.nj.nec.com/context/15377/132530> (University of California at San Diego, 1995).
76. Eggenberger, P. Cell interactions as a control tool of developmental processes for evolutionary robotics. *Proc. 4th Int. Conf. Simul. Adaptive Behav.* (eds Maes, P. *et al.*) 440–448 (MIT Press, Cambridge, Massachusetts, 1996).
77. Floreano, D. & Urzelai, J. Evolutionary robots with online self-organization and behavioral fitness. *Neural Netw.* **13**, 431–443 (2000).
This paper explores the evolution of EAA neurocontrollers through learning in a self-organizing manner without any specification of a set of inherited initial synaptic weights.
78. Rust, A. G., Adams, R., George, S. & Bolouri, H. Activity-based pruning in developmental artificial neural networks. *Proc. 4th Eur. Conf. Artif. Life (ECAL 97)* (eds Husbands, P. & Harvey, I.) 224–233 (MIT Press, Cambridge, Massachusetts, 1997).

Acknowledgements
Supported by the FIRST grant of the Israeli Academy of Sciences, and the Adams Brain Centre, Tel-Aviv University. I thank R. Aharonov, T. Beker, N. Gal-Ruppin, D. Horn, I. Meilijson, Y. Niv, Y. Oron, J. A. Reggia and Z. Solan for careful reading of this manuscript and many helpful comments.

Online links

FURTHER INFORMATION
Encyclopedia of Life Sciences: http://www.els.net/
central pattern generators | neurons and neural networks: computational models | neural networks and behaviour
Eytan Ruppin's lab: http://www.math.tau.ac.il/~ruppin/
MIT Encyclopedia of Cognitive Sciences:
http://cognet.mit.edu/MITECS/
evolutionary computation | neural networks
Access to this interactive links box is free online.