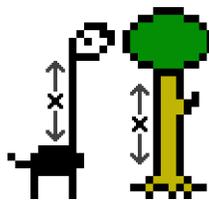


MASTER'S THESIS

Coevolution of Learning Ability and Neuro-cognitive Organization

Solvi Arnold



First supervisor and reviewer: Dr. Gerard Vreeswijk
Second supervisor and reviewer: Dr. Thomas Müller
Third reviewer: Dr. Janneke van Lith

Date: August 30, 2010
ECTS: 30

Abstract

The environmental complexity thesis states that environmental complexity is the driving force behind the evolution of cognition. Herbert Spencer held a particularly strong version of this view, and believed that life and mind can be understood as reflections of the environment they evolved in. However, Spencer's view does not account for the possibility of fit but diffusely implemented behaviour. As connectionist AI has amply demonstrated, fit behaviour does not by itself necessitate any isomorphism between a species' neuro-cognitive organization and the environment. We suggest supplementing Spencer's view with an account of the selection pressures that would cause evolution to organize cognition after the environment, and identify selection pressure on learning ability as a candidate. We argue that the more a species' neuro-cognitive organization resembles the organization of the environment, the easier it is to make appropriate updates in behaviour. We discuss various types of learning ability as it occurs in nature, and identify latent learning as the type most likely to constrain neuro-cognitive organization. We then build an Artificial Life model of the evolution of latent learning, and compare the structures of networks evolved under selection pressure for latent learning with networks evolved in absence of such selection pressure. Unlike the latter, the former repeatedly evolved the same compact behaviour system, which innately encodes some of the spatial relations of the environment. Our results indicate that selection pressure on learning ability can indeed guide evolution towards forms of neuro-cognitive organization that reflect environmental features.

Acknowledgements

First of all, I would like to thank my supervisors and reviewers for their time, effort, and numerous suggestions for improvements. The thesis is part technical and part philosophical, and I am grateful that people from both sides were willing to supervise and review it.

I stumbled upon the topic for this thesis during my internship at Nagoya University, in the Artificial Life lab of the Graduate School for Information Science. Special thanks go to Takaya Arita and Reiji Suzuki for their guidance and encouragement during that time and after, and for introducing me to so many wonderful lines of Artificial Life research.

Many thanks also go to Willeke and Vincent, for the many discussions about the content of this thesis, for offering their fresh perspectives, and for pointing out the holes in my explanations along the way. Vincent additionally helped me out a lot by letting me run the experiments on his computer, which would have taken many times as long on my own.

Contents

Introduction	1
1. The evolution of mind	4
1.1. The environmental complexity thesis	5
1.2. Spencer: life and mind as correspondence	6
1.3. Godfrey-Smith's view	8
1.4. Terminology	9
1.5. Argument against correspondence	11
1.6. Learning ability	13
1.7. Why learning matters	14
1.8. Conclusions chapter 1	16
2. The varieties of learning	17
2.1. The universal learning rules	17
2.2. Behaviourism versus the non-universality of learning	18
2.3. Non-universal learning	19
2.4. Latent learning	20
2.5. Cue-based learning	23
2.6. On the classification of learning types	24
2.7. Conclusions chapter 2	25
3. An Artificial Life approach to latent learning	26
3.1. Artificial life and mind	26
3.2. What necessitates learning?	27
3.3. Existing models	28
3.3.1. Latent learning extensions to Reinforcement Learning	29
3.3.2. Evolution of learning mechanisms	30
3.3.3. Hebbian associative learning	31
3.3.4. Auto-teaching	32
3.3.5. Neuromodulation	33
3.4. Avoiding Cue-based learning	35
3.5. A model of the evolution of latent learning	36
3.5.1. Model concept	36
3.5.2. Environment	36
3.5.3. Species	37
3.5.4. Genetic algorithm	39
3.5.5. Why this model?	39
3.6. Results	40
3.7. Analysis	41
3.7.1. Perception	42
3.7.2. Coordination	44
3.7.3. How feature detector superimposition facilitates learning	47
3.8. Conclusions chapter 3	48

Conclusions & Future work	50
Appendices	57
B. Reasoning behind model details	57
C. Tasking non-learners with learning	58
D. The necessity of echo neurons	58
E. Connectivity analysis	59
F. Non-effect of more modulatory neurons	59
G. Clustering algorithm	60

Introduction

King. "And let me ask what you mean by those words 'left' and 'right.' I suppose it is your way of saying Northward and Southward."

"Not so," replied I; "besides your motion of Northward and Southward, there is another motion which I call from right to left."

King. Exhibit to me, if you please, this motion from left to right.

I. Nay, that I cannot do, unless you could step out of your Line altogether.

King. Out of my Line? Do you mean out of the world? Out of Space?

I. Well, yes. Out of *your* world. Out of *your* Space. For your Space is not the true Space. True Space is a Plane; but your Space is only a Line.

King. If you cannot indicate this motion from left to right by yourself moving in it, then I beg you to describe it to me in words.

I. If you cannot tell your right side from your left, I fear that no words of mine can make my meaning clearer to you. But surely you cannot be ignorant of so simple a distinction.

Edwin A. Abbott, *Flatland*, 1884

The above is a two-dimensional square's vain attempt to explain his second dimension to the one-dimensional king of Lineland. Not long after, the square himself is equally baffled when a sphere from Spaceland attempts to explain to him a third dimension. To each of these characters, the number of dimensions of their world is so self-evident, that the mere idea of a different number seems outrageous.

Being born into an unknown world without prior knowledge, a good first thing to do would indeed be to count this world's dimensions. Yet it does not seem like we do this, and the very suggestion seems odd. Plenty of species incapable of counting manage perfectly well in a world of three spatial dimensions, seemingly without ever considering this number. We will not discuss dimensions much here, but rather this phenomenon of specific characteristics of one's world being ingrained so strongly in the mind that they

define how one perceives that world and interacts with it.

The example of dimensionality is perhaps *too* intuitive for our present purpose. Had there really been two-dimensional species with two-dimensional cognitions to contrast with our own, the above might have seemed less of an exercise in absurdity. In other respects however, different species' worlds *do* vary, and this variation is reflected in differences in their cognitions. Take the rat, for example. Experiments with rats show that these animals can be taught to perform all sorts of actions, even quite complex sequences of action, by rewarding these behaviours with food. Yet it has been shown to be virtually impossible to make rats jump for food.

"[R]ats are easily taught to jump to avoid shock, but jumping to obtain food is almost impossible to teach." (Gould and Gould, 1994, p.60)

Teaching them to jump to avoid electric shock poses no problem, teaching them to navigate a complex maze for food poses no problem either. Yet when a rat is presented food every time it jumps, it flat-out fails to make a connection between the two events. Why can they make some connections with ease, while other equally simple connections fall outside their cognitive capacities? A common explanation is that in the rat's natural environment, jumping is not part of any food-procuring routines, while both navigating complex habitats for food and evasive jumping are "natural" rat behaviours. Maybe the idea of a relation between jumping and food is as alien to the rat as the idea of moving left or right is to king line. These ideas simply make no sense in their respective natural habitats.

This paper explores the relation between cognitive and environmental structure, how through evolution the latter shapes the former, and what role learning plays in that process. The central question is whether the evolution of learning ability promotes the evolution of cognitive organization.

The thesis fits into two separate lines of inquiry, one in Artificial Intelligence and one in Artificial Life:

Connectionism and systematicity

The first is the debate about systematicity. Within the AI community there is a long-standing debate about this issue. The models of cognition classical AI proposes are criticized for being too brittle, rigid and biologically implausible to adequately model cognition (see e.g. Brooks, 1991), while connectionism's flagship architecture, the artificial neural network, is criticised for being too diffuse to adequately model cognition (see e.g. Fodor and Pylyshyn, 1988; Fodor and McLaughlin, 1990 for notable instigators, and McLaughlin, 2009, for a recent defense). Though the matter was never quite settled, the majority of the AI community nowadays recognizes the strengths and weaknesses of both strands of AI, and many have turned to hybrid approaches. For systems built to perform, this is a fine development, but for systems built to explain, I believe it is uncalled for.

Although I think that many proponents of classical AI overestimate the systematicity of natural cognition,

the charge that connectionist models typically fall short in terms of "cognitive organization" certainly has force. However, for those aiming to use artificial intelligence as a means for understanding natural intelligence, this is a virtue more than a vice. While the question of how a classical AI system comes by its cognitive structure is best answered by pointing at its programmer, the question of how a connectionist system might come by cognitive structure is a research question the answer to which may shed light on the origin and structure of natural cognition. Maybe the best rebuttal for a connectionist is that classical AI does not explain cognitive structure *either*, but merely assumes it (versions of this argument can be found in Aizawa 1997; Matthews 1997; Hadley 1997b). As classical AI systems cannot be built without assuming cognitive structure, they are badly suited to the study of how cognitive structure came about.

A connectionist taking this stance is bars herself from explicitly imposing structure "from above": imposed structure lacks the explanatory power of emergent structure, so by hard-coding structure into connectionist systems, one surrenders the explanatory edge connectionism *can* have over classical AI.

There is an alternative strategy, but somehow it took hold in the Artificial Life community more than in the AI community. It goes as follows: artificial neural networks mimic biological neural networks. If we want artificial neural networks to acquire the sort of structures we find in biological neural networks, we should study the forces that organize biological neural networks, and mimic these. If sufficiently flexible systems are exposed to the appropriate organizing forces, structure should emerge. Some researchers within the Artificial Life community (always more interested in new insights than in practical applicability) have picked up on this theme (see e.g. Jacobs, 1999; Kashtan and Alon, 2005; Bullinaria, 2007, 2009). Now these researchers are not particularly concerned with full-fledged systematicity as proponents of classical AI envision it, but understanding the forces that organize cognition would bring us closer to understanding where cognition got the semblance of systematicity. This thesis fits neatly into this line of inquiry.

Interactions between levels of adaptation

The second line of inquiry is almost completely outside the scope of Artificial Intelligence, but a major theme within Artificial Life: the interactions between different levels of adaptation. As early as 1896, James Baldwin hypothesized that learning ability can accelerate, or "guide", evolution (the "Baldwin effect"). Ever since Hinton and Nowlan (1987) first showed the effect in action in a computational model, this has been a major topic in Artificial Life. Recently, other interactions between adaptation levels have received attention too. Bioinformaticists Crombach and Hogeweg (2008) have shown that different levels of evolution can have acceleratory interactions as well ("evolution of evolvability"), and Paenke et al. (2006, 2009) call attention to the "hiding effect", the phenomenon that under certain conditions, learning ability *decelerates* evolution. But interactions between levels of adaptation need not merely be acceleratory or deceleratory; through its effect on selection pressure, the presence of another adaptation process can affect not just the speed, but also the direction of an adaptation process. The hypothesis of this thesis predicts one such effect.

While both are major topics in their respective fields, there is little or no literature to be found on the relevance of these two lines of inquiry to one another. We will not get into the details of the systematicity debate; our project is first and for all an artificial life project. Only if the hypothesis is confirmed does it bear relevance to the debate in AI.

The thesis roughly divides into three parts. The first part discusses an existing philosophical position that posits environmental structure as the source of the structure of cognition, and proposes a revision that introduces learning ability as a major factor. The second part, rooted in psychology, explores the varieties of learning as they occur in nature, and what they tell us about their hosts' cognitive structure. The third part then takes the findings from the second as the basis for a simple artificial life model that simulates the evolution of learning ability using artificial neural networks, in search for empirical support for the position put forth in the first.

1. The evolution of mind

We tend to think of the human mind as a sort of container filled with thoughts and feelings and the like. Philosophers of mind are primarily concerned with these contents: how they are structured, how they exist in a physical medium, how they relate to one another and to the world.

Both in introspection and the philosophical literature, there is an apparent divide between the container and its content. We don't feel limited in what the container lets us contain; it appears to support every imaginable thought (a dubious intuition to go by for sure). The container has us wondering about its ontology, but given its apparent universality, it does not figure much in our considerations of our acts of cognizing.

The intuition that the container imposes no restrictions on its contents is evidently unwarranted. We fail to mentally represent 4D spaces like we represent 3D spaces. Occasionally the restrictions of the container find their way into philosophical debate: the *cognitive closure* hypothesis states that fundamental limitations of our cognition underlie the apparent unsolvability of certain philosophical problems (McGinn, 1993). We won't argue for or against this hypothesis here. What interests us about it is merely the idea that the ability to think is not a black and white matter. It is not the case that if an organism can think something, it can think everything, and everything equally well. Some species can think things others can't, or can think with ease things others can only think with great difficulty. This gradual nature of what species can or cannot think opens up the way for an evolutionary approach.

Furthermore, it forces us to reconsider the relation between the container and its contents. Let's (try to) imagine the mind of an organism capable of entertaining only three (proto-)thoughts (ticks for example distinguish three environmental states of affairs, see Gould & Gould, 1994). This does not prompt us to envision a vast and wide but practically empty container. A set of three light bulbs that can each be on or off seems like a more suitable image, thoughts being realized not as unbounded content in a

container but as a restricted set of states of a system. The more restricted the repertoire of contents, the less we are inclined towards a sharp divide between content and container.

But what about the rat mind? It's quite capable of genuine insight, in the right circumstances, tempting us to attribute it a container-content divide. Yet it fails to make certain exceedingly simple connections (such as "jump yields food"), while it quickly picks up on other, more complex connections (such as "traversing this maze yields food"), suggesting a missing light bulb or two. For rats and other species of intermediate cognitive ability neither conception quite fits. This makes them particularly relevant to an evolutionary understanding of mind. They show us that there is no clear tipping point, no sudden evolutionary transition from a light bulb mind to a container-content mind, and this should make us question our intuitions about our own minds.

What the divide between content and container is supposed to consist in is unclear. Why we are so inclined to assume that there is one is unclear too. Could what we perceive as an unrestrictive container-content mind maybe be realized as a vast network of interconnected light bulbs, evolved from simple Christmas-light ancestors? From this angle, the surprising limitations of the rat mind are not as puzzling; they're exactly what one should expect to find in the intermediate stages of cognitive evolution.

Some philosophers of mind indeed opt for this bottom-up, evolutionary approach of cognition. The approach implies two lines of inquiry:

Base case: The most primitive proto-cognition should become understandable once we understand how evolution turns matter into proto-cognition.

Recursive case: Advanced cognition should become understandable once we understand less advanced cognition and the evolutionary processes that advance it.

In this first chapter of this thesis, we discuss the work of two philosophers advocating evolutionary views of the mind. We identify a problem in the most promising of these views, and propose a revision.

As a consequence of our decidedly bottom-up approach, the thesis won't be concerned much with advanced aspects of cognition such as language and rational thought, setting the project apart somewhat from most discussions within the philosophy of mind. What we will concern ourselves with are the basic forces that shape and organize cognition over the course of evolution.

1.1. The environmental complexity thesis

In evolutionary explanations, one central question is always where the *fitness benefit* of the explanandum lies. Unlike a polar bear's fur, cognition does not bestow a species any direct benefits. Cognizing alone won't keep an organism from freezing or starving, nor will it provide mating opportunities. Cognizing pays off indirectly, through its influence on behaviour. Cognition is adaptive because it makes it possible to exploit the patterns of the environment. Thus the selection pressures driving the evolution of cognition should be sought in the structure of the environment.

This idea is not new. Peter Godfrey-Smith (1996, 2001, 2002) calls it the *Environmental Complexity Thesis* and expresses it as follows:

"The function of cognition is to enable the agent to deal with environmental complexity." (1996, p.3)

Godfrey-Smith is an interesting figure in the philosophy of biology. He seeks to connect existing lines of thought within the philosophy of mind with artificial life's computational approach. He pitches his ideas about the evolution of cognition to the artificial life community (see e.g. Godfrey-Smith, 1994), and draws extensively from the artificial life literature to illustrate his points. Within the artificial life community itself too, there is an intuition that artificial life could further our understanding of the evolution of mind. See for example Arita (2009) and Bedau (2003), who even hints at a future merging of artificial life and

cognitive science. However, the subject has proven difficult to approach and artificial life research on the evolution of mind is still scarce. We will get back to the relation between artificial life, cognitive science, and philosophy of mind in chapter 3.

Godfrey-Smith has repeatedly stressed in particular the importance of work of Herbert Spencer and of John Dewey to the cognition-oriented lines of research within artificial life. Spencer is (carefully) credited as the first to emphasize the idea of a relation between environmental complexity and cognition, in his "Principles of Psychology" from 1855, some 4 years before the publication of Darwin's "Origin of the Species". As it turns out, Spencer's view is particularly relevant to the topic at hand: it traces the history of cognitive structure up from the simplest of organisms. The same cannot be said of Dewey, nor of Godfrey-Smith himself. In Dewey's view, thought must be (in some broad sense of the word) linguistic, and he does not attend much to its non-linguistic precursors. Godfrey-Smith's own view is more pertinent than Dewey's, as he does not only attend to high-level representational thought but also to the primitive seeds of cognition, but he never seems to make any attempt to connect the two. So, somewhat ironically, we will focus on the work Godfrey-Smith endorses least, that of Herbert Spencer.

Evolution is a central theme in much of Spencer's work. He popularized the term and coined the phrase 'survival of the fittest'. Spencer originally held a Lamarckian view of evolution, and applied it generously to just about every domain of study he concerned himself with, be it biology, sociology, politics, economics, ethics, or astronomy. He discerned in all these fields a natural trend towards differentiation, specialization, and heterogeneity, all the work of an omnipresent organizing force that brings about organization at all scales of existence, and this force he called evolution.

Of course some aspects of Spencer's view are now simply outdated. We now know that Lamarckian inheritance is exception (see e.g. Nätt et al, 2009) rather than rule, and that the forces organizing stars into galaxies are quite distinct from those organizing multicellular organisms. However, that Spencer's

notion of evolution differs on details with our modern Darwinian one should not detract from his general ideas about the origin of mind.

In the next section, we introduce Spencer's view on the evolution of cognition, and contrast it with Godfrey Smith's.

1.2. Spencer: Life and Mind as Correspondence

"[W]ith complexity of organization there goes an increase in the number, in the range, in the speciality, in the complexity, of the adjustments of inner relations to outer relations. And in tracing up the increase we found ourselves passing without break from the phenomena of bodily life to the phenomena of mental life."
(Spencer, 1885, p.294)

In Spencer's view, life and mind only differ by degree: cognizing *is* an advanced form of living. Through evolution, life grows more and more complex, eventually reaching the level of intricate internal organization we call cognition. To those inclined towards an evolutionary understanding of mind, Spencer's concept of continuity between life and mind is attractive.

Ideas of continuity between life and mind come in various forms and strengths. Godfrey-Smith proposes a distinction between weak and strong types of continuity:

"Weak continuity: Anything that has a mind is alive, although not everything that is alive has a mind. Thought is an activity of living systems."

"Strong continuity: Life and mind have a common abstract pattern. The functional properties characteristic of mind are enriched versions of the functional properties that are fundamental to life in general." (1996, p.72-73)

Already we can conclude, with Godfrey-Smith, that Spencer's continuity claim is a strong one. Godfrey-Smith himself holds weak continuity to be true. Weak continuity tells us little about cognitive structure; for the present project, strong continuity like Spencer's

gives more guidance, though it is harder to defend. We will find that it isn't entirely tenable, but still arrive at a stronger-than-weak continuity.

To get at how Spencer envisions the continuity between life and mind, we must first have a look at Spencer's concept of life. Spencer characterizes life (and thereby mind) first and for all as self-sustenance.¹ For an organism to sustain itself, it must respond to its environment. Such responding to the environment can be as simple as just absorbing nutrients from it. The yeast-plant, for example, thrives in uniform environments (fluids of a particular temperature). It absorbs nutrients from its environment without distinguishing anything from anything. When the environment fails to supply, the organism dies.

Most environments aren't as simple. Nutrients tend to be dispersed unevenly and among non-nutritious particles, requiring organisms to bring themselves into contact with them, or at least selectively absorb the nutritious ones. In Spencer's terminology, this establishes a *correspondence* between the internal and external: heterogeneities in the environment (nutritious / not nutritious) are met with "corresponding" heterogeneities in the organism's responses (absorb / do not absorb). Life, characterized as self-sustenance, is then realized as a system of correspondences between the organism and the environment.

When we move on to terrestrial plants, we find them responding to cycles of night and day, and to the seasons; they display a heterogeneity of states corresponding to a heterogeneity in their environment. Also, plants typically have certain physical differentiations between the parts that are in the soil and the parts that are above it, another corresponding pair of dichotomies.

So far, the correspondences are purely mechanical. We will not trace the evolutionary path up to the more advanced species here. The general point is that where

¹ Of course when we look at our own species and some others, not all activities of the mind serve self-sustenance, but I think it no stretch when Spencer construes self-sustenance as the driving force behind the initial stages of evolution of cognition.

there are dichotomies, sequences, contingencies or any sort of organized heterogeneity in the environment, and they are relevant to a species' self-sustenance, that there is potential for evolution to equip the species with a corresponding physical feature that allows the species to deal with it. Next, once such a feature comes about, it often puts further heterogeneities within reach of the species. For example, once an ability to distinguish light from dark has evolved, an ability to distinguish contrasts can follow.

Unless we entered the discussion with a heavy-laden concept of correspondence, the claim that species' physical characteristics correspond to the challenges of the environment should meet little resistance. Correspondence, so far, is remarkably similar to plain old *adaptation*. The two concepts arguably coincide when it comes to straightforward mechanical adaptations, but they start to diverge when we look at species' internal organization. Roughly speaking, the more "internal" the evolved feature, the more "external" the corresponding environmental feature. Thus evolution from primitive to advanced forms of life (in particular, mind) can be thought of correspondence extending *inward*, away from where the organism and environment affect one another (we will call this the "plane of interaction" for short). Let's make this line of thought a little more precise.

If an enzyme exists to decompose certain nutritious molecules found in the environment, the shape of those molecules puts constraints on, or even dictates, the shape of the enzyme. We find a 'lock and key' kind of correspondence between the two. At this lowest, physical level, where organism and environment (literally) touch, the claim of correspondence is near trivial.

At the plane of interaction, perception and action are indistinct. An enzyme does not first perceive a molecule to be nutritious, and then acts to decompose it. To recognize it and to lock onto it are one and the same thing here. Similarly for the simplest of organisms, recognition equals ingestion.

The plane of physical interaction provides a grounding for the more complex mechanisms of life (including

mind), for ultimately all self-sustenance serving interactions with the environment (exchange of chemicals, reproduction) involve the plane of interaction. (Characteristically, Spencer takes this realization two steps further by boldly claiming that all senses originally evolved out of a global and undifferentiated sense of touch: the "primordial irritability".)

But cognition is typically taken to occur "in between" perception and action; if perception and action coincide, there is no room for cognition. Differentiation from unified perception-action into separate perception and action systems occurs when an evolving species starts to exploit aspects of the environment that are removed from the plane of interaction. What are the conditions for this to occur?

1) To be relevant to an organism, events and states of affairs external to the plane must provide information about future events or states of affairs involving the plane. While nothing prohibits detection of irrelevant events and states of affairs, there is no selective advantage in capacities for it, and thus these won't usually evolve (except as side-effects of actually useful detection capacities). Thus aside from touch, evolution of all sensory modalities hinges on the predictive value external events and states of affairs have with respect to events and states of affairs involving the plane of interaction. For such predictive value to exist, there must be structure to the environment. In a fully chaotic environment, in which all events and states of affairs are wholly independent, no distal events have relevance to proximal events. Here we find a clear point of contact with the environmental complexity thesis.

2) A species' sensory modalities (besides touch) and its action repertoire must coevolve. There is no point in detecting predictive signals in one's environment when all one can do is wait out the predicted events. For a stimulus to be worth detecting, there must be a potential for response. Conversely, action relies on perception²: there is little point in an ability to flee if predators can't be detected anyway.

² Except of course types of action that are invariably useful no matter the present state of the environment.

So for prying apart perception and action in order to make room cognition, we need perception and action to coevolve, and we need exploitable contingencies to exist in the environment. Next, exploitation of those contingencies requires *coordination* between perception and action. This coordination too must coevolve along with perception and action. In this coordination we should eventually find advanced correspondences that characterize the mind.

In the evolution of this coordination we find the difference between adaptation and correspondence. In Spencer's view this coordination is itself a system of correspondences: the relations between perceptions and actions reflect relations in the environment. Viewing coordination as adaptation instead, we arrive at a far weaker claim than Spencer's: that the relations between perceptions and actions are adaptive does not imply similarity or isomorphism with relations in the environment. Mind-as-adaptation does not imply strong continuity, while mind-as-correspondence does.

To summarize: in heterogeneous environments with contingencies spread out in space and time, the coevolution of senses working at a distance (sight, smell, hearing, echolocation, electroreception, etc.) and varied repertoires of responses produces in organisms structures as rich and complex as the environments they correspond to. What we first recognize as characteristics of life so evolve to become the characteristics of cognition: recognizing salient properties of one's present environment, processing these in a manner incorporating the contingencies of that environment, and acting in such a way as to secure one's self-sustenance.

This is Spencer's view of life and mind in a nutshell. It's not without its problems. As we will see, the further we get removed from the plane of physical interaction, the more the principle loses force. Spencer does not appear to have considered the weaker mind-as-adaptation claim, and this is where his view needs more support than he provides. Much of this thesis could be seen as an attempt to do so.

1.3. Godfrey-Smith's view

Godfrey-Smith does not go as far as to give an account of the evolution of mind. Instead he selectively endorses elements from others' views that posit environmental complexity as the cause of the evolution of mind:

"Those are the three central components of the view of mind endorsed here: properties of environmental *complexity*, which make mind worth having, *reliability* properties, which give thought its purchase on the world, and a naturalistic *interactionist* view of the causal traffic between mind and the rest of nature." (1996, p.196)

Godfrey-Smith claims that environmental complexity explains the existence of cognition (*why* it evolved), while Spencer could be said to claim that environmental complexity explains both the existence and organization of cognition (*why* and *how* it evolved).

It is clear that Godfrey-Smith has little faith in the how-part of Spencer's account. On correspondence, Godfrey-Smith says the following:

"The issue of whether a relation of correspondence has a role to play in developing the environmental complexity thesis is undecided. This is a consequence of the fact that there is no widely accepted naturalistic theory of what correspondence is." (1996, p.196)

Unfortunately Godfrey-Smith offers no alternative. We find nothing in his view to replace Spencer's correspondences; there is simply very little attempt to explain cognitive structure in terms of environmental structure, beyond the observation that complex environments require complex behaviour.

This keeps his position plausible and reasonable, but it provides the artificial life community he deems his views particularly relevant to little to work with.

There is something to be said about the relation between philosophy and artificial life here. I haven't found much mention of the environmental complexity

thesis in artificial life literature, but I would expect it to be fairly uncontroversial within the artificial life community. An artificial life approach to cognition almost implies viewing cognition as adaptation. To point out environmental complexity as the thing it is an adaptation to is no big step. What I think makes the thesis interesting, is that it suggests that the environment shapes cognition, and that we can learn about cognition by studying the environment and the shaping process. Here we find something for artificial life to work with, but on this point, Godfrey-Smith says little if anything. He plays it so safe that the artificial life community can do little more to support his position than continuing to tacitly assume it.

Of course, modern day philosophers are wise to avoid the sort of bold claims Spencer dealt in, but what is a dangerously tenuous claim to a philosopher can be an interesting *hypothesis* to an artificial life researcher. The more specific and bottom-up an account of cognitive evolution is, the easier it is to translate it into a computational model. I take Spencer's view over Godfrey-Smith's not because it is more plausible, but because it makes interesting testable claims that can be verified and modified using computational models.

In fact, connectionists could be said to have been testing Spencer's theory quite extensively, though maybe not on purpose. The results suggest that some revisioning of the theory is in order. We will get to that shortly, after we attend to our terminology.

1.4. Terminology

There are a small number of key terms in our discussion so far. Before we can get technical, we should specify in detail what we mean by these.

Complexity

While the environmental complexity thesis has plenty of intuitive force, the term 'complexity' is notoriously hard to define. Godfrey-Smith notes the problem and decides to interpret complexity as heterogeneity. It's questionable whether this helps. More confusingly, before he settles on this interpretation, he notes that "perhaps the common sense view is that complexity is

organized heterogeneity" (1996, p.25), then goes on to explicitly discard the property of organization from his concept of complexity. This move provides a more tangible concept of complexity than the "common sense view": we can fairly well picture what maximal heterogeneity would amount to (white noise), while heterogeneity and organization have a certain antagonism going that makes it very tough to assess a given environment's 'organized heterogeneity'. However, discarding organization as an essential feature of complexity makes Godfrey-Smith's version of the environmental complexity thesis very hard to apply to Spencer's view. As we saw, Spencer seeks the mind in the coordination of correspondences, this coordination itself corresponding to the organization of the environment. Spencer's view indeed appears more directly concerned with organization than with heterogeneity. Heterogeneity just so happens to be a requirement for any sort of organization to exist. A purely homogeneous environment can hardly be said to be organized. Similarly, coordination by definition requires multiple distinct elements; a single thing cannot be coordinated. For a coordination to be in correspondence with an environment, the environment must be organized, and only as a consequence of this requirement does the environment need to be heterogeneous. Thus if we are to read 'complexity' simply as 'heterogeneity, regardless of organization' in Godfrey-Smith's version of the environmental complexity thesis, it is hard to see where he finds his common ground with Spencer.

Imagine an unorganized *but* maximally heterogeneous environment (a white noise environment, so to say). Without contingencies to exploit, nothing can guide behaviour. Consequently, the best behaviour is to continuously perform whatever action has the highest average payoff. Thus the behaviour a maximally heterogeneous environment calls for is equally simple as that a maximally homogeneous environment calls for. If environmental heterogeneity does not call for behavioural heterogeneity, it should have little to do with cognition. The link between environmental complexity and cognition is established by the aspect of organization that Godfrey-Smith claims to discard, and consequently his version of the environmental complexity thesis fails to make sense if we take serious his discarding of organization as a crucial ingredient of

complexity. Fortunately it does not seem like we need to. When he puts forth his own view, we find the aspect of organization, now separated from heterogeneity, back under the guise of the “reliability properties (which give thought its purchase on the world)” he adopts from Dewey. It’s not clear what is gained by describing the environment as complex (defined as heterogeneous) and reliable³, instead of as just complex (defined as heterogeneous *and* organized), except maybe for shaking the definition police.

Where the common sense view might cause confusion, “complexity” can be read as “organization” in this thesis. Combined with the stronger aim of not only explaining the existence but also the structure of cognition, this allows for a stronger and snappier formulation of the environmental complexity organization thesis:

The function of cognitive organization is to enable the agent to exploit environmental organization.

Correspondence

While correspondences are the heart of Spencer’s view, he does not seem to provide a clear definition. Within philosophy, the term seems to mean a lot of different things in different contexts. I think the lack of definition stems from Spencer’s concept of correspondence being very much down-to-earth and in line with the intuition. The shape of the enzyme corresponds to that of the nutritious molecule simply by fitting onto it, a capacity to distinguish light from dark corresponds to varying light conditions in the environment, the length of the giraffe’s neck corresponds to the tallness of trees, and the square’s sense of left and right corresponds to the spatial dimension king line lacks. As a general principle, we might take the following:

³ Unlike organization, reliability alone won’t do: a perfectly homogeneous environment is certainly reliable, but it does not call for cognition.

For some internal P to correspond to some external Q, P must (specifically) enable the organism to deal with Q. Q offers a teleological explanation of the existence of P.

This works for enzymes, contrast-sensitivity, long necks, and dimensions. More generally, it also neatly fits onto the environmental complexity thesis itself. If the function of cognition is to enable an agent to deal with environmental complexity, then the above characterization of correspondence states that cognition corresponds to environmental complexity (of the organized kind). This is quite in line with Spencer’s position as I read him.

This is not to say that no P could ever correspond to multiple Q, or vice versa. It’s very well imaginable that some correspondences are tidier than others, nothing requires correspondences to be quite perfect yet (Spencer details various lines of development in the correspondences, and sees the present state of evolution as work in progress). On the extreme opposite of what Spencer thought to see in all of life and mind, were all P to simultaneously and equally correspond to all Q and vice versa, there would be little use in talking about correspondence at all (hence we should require some degree of specificity for a relation between a P and a Q to be a correspondence).

We should stress that despite the prominent role assigned to correspondence relations, neither Spencer’s nor our project should be mistaken for variants of some correspondence theory of truth. We will not be viewing cognition as a sort of end product of evolution, which then goes on to autonomously establish correspondences with its environment. We look at the role the environment plays in the evolution of cognition. That is, we look not at how cognition establishes correspondences between the internal and external, we look at how at correspondences between the internal and external establish cognition.

Brain, Cognition, Mind

The difference between ‘mind’ and ‘cognition’ is here understood as one of connotation only. We might occasionally be inclined to ascribe cognition to organisms we would be less inclined to ascribe minds

to (likely due to the word ‘mind’ being more strongly associated with ‘consciousness’), but this is more a matter of degree than of qualitative difference.

The relation between mind/cognition and brain, on the other hand, is a major issue in the philosophy of mind. However, anyone defending an evolutionary view of mind must hold that in one way or another, the mind has at least some of its components coded in the genotype. An evolutionary view of the mind naturally approaches the mind as an aspect or function of the brain. While this does not force one to hold that an understanding of the brain provides an understanding of (some aspects of) the mind, there is little to be gained from an evolutionary approach when one doesn’t.

As can be expected from someone defending strong continuity between life and mind, Spencer uses the terms ‘mind’ and ‘brain’ almost interchangeably. In supporting his claim that mind is characterized by the same sort of correspondences that characterize life, he even writes in detail on the evolution of neural structure. Godfrey-Smith isn’t as clear. He does not defend any form of continuity between life and mind, leaving open the option of handling the two separately. He also leaves a large gap between life and mind in his treatment of the environmental complexity thesis, exploring both the most basic traces of cognition and the relations between thought and environment, but not what goes in between.

1.5. Argument against correspondence

As we noted before Spencer’s view has its problems. While correspondences of the kind Spencer is after obtain naturally in perception and action (occurring at the plane of interaction), it is far from self-evident that they should extend inwards to coordination. Spencer seems to fail to consider the possibility that multiple internal features and relations could deal with multiple external features and relations in such a diffuse manner that we cannot reasonably speak of correspondence, but merely of adaptation.

The argument is maybe a bit too modern to have been considered by Spencer at the time. The core of the

argument is that most problems that have a solution have infinitely many. The problem is illustrated by the unintelligible and unwieldy programs often produced by genetic programming (see Figure 1).

$$f_1(t) = (\sqrt{((\sqrt{((\sqrt{(2.957153)) - (\sin(\sqrt{t})) + (\sin((1.854912) - ((\sqrt{(3)^t))} \\ * (\sqrt{t})) + (4.435898))) + (-2.355442)))) * t)) + ((40.675830) / (\exp(t) * (\sqrt{t}) - \\ ((\sin((\sin((\sqrt{(3.756391)^t) * t)) * (\sqrt{t})) + (\log(86)) + (7))) + (3)) + (\sin(\sin((1.654737 \\) * t)))))) + (-2.355442)))))) / (\sqrt{(54.598150)}) / (\sqrt{(7.931547)})),$$

$$f_2(t) = (\sqrt{((3.777992) - (((4.190957) - (t) - (\sin((t) * (t)^t)) - ((t) * (2.883554) \\ / ((4) - (\log(t)))) + (2.791190)) - (\exp(t) - (2.226704)) * (\sqrt{(9.642561)})) / (t^t)) \\ * (2.678347)) / (3.462360)) + (3.792098) - (4.796861) / (4)) + (((3.792098) - (\exp(t) \\ / (3.462360)) - (t) * (t^t)) / (t^t)) / (t^t)),$$

$$f_3(t) = ((\log(\log(\exp(\log(\log(\exp(\log(\exp(\pi)) - (\sin(9) + ((t) + (8.000000)) \\ ^{(\sqrt{(1.420245))}))) / ((\exp(t)) * (379.000000)) / (84.000000)))) - ((\sin(8) + ((t \\ + ((\log(109)) - (1.258803)) / (6.620476)) * (\sqrt{(\log(4.059461))}))) \\ / ((\exp(((8.337047) * (\log(\log(\sqrt{(3.021610)))) + (2.000000)) - (5.912041)) * (\exp(t)) \\ / (85.000000)))))) * (5.933873)),$$

$$f_4(t) = (((\log(6.249382)) * (\sqrt{(6)})) * ((\sqrt{(10.000000)})^{(1.172486)} - (t)) \\ / (6.981835))) * ((1.161464) - (1.161464) / (((\sqrt{(6.249382)) * (\log(7.008566)) \\ * (((\exp((6.980522) / (\sqrt{(6.288201)}) * (1.344547)))^{(1.735082)} - (t)) \\ / (0.290257)) * (\sqrt{(6.000000))} * ((9.704760) * (-0.050358) - (t)) - (t)) / (0))) \\ * ((1.634223) + (7.277223) * (0.290257 - (t)))^{(0.161464)} / (t))) / (6.980522)).$$

Figure 1: Genetic Programming used to identify a gene regulatory network (from Kim, Cho & Zhang, 2006).

Present an artificial evolution process with a simple problem, predict a fitting solution neatly in line with the structure of the problem, and more often than not, evolution will hand you back a complete mess of a solution that *just somehow happens to work*.

We can understand this problem in terms of selection pressures. To score high fitness, an organism must produce suitable responses to the environment’s stimuli. Evolution does not care whatsoever *how* the mapping from stimulus to response is realized internally; as long as the resulting behaviour is the same, the resulting fitness is the same, and evolution won’t be able to distinguish between the two.

At the plane of interaction, the selection pressure on attaining good performance forces correspondence (the enzyme cannot work without fitting the nutritious molecule), but as we get further removed from the contact border (how to behave so as to make contact between one’s enzymes and nutritious molecules more likely), the organism’s performance becomes progressively less dependent on correspondence. Consequently, without some additional selection pressure on keeping things internally organized, the correspondence will become progressively weaker as it extends inwards.

That Spencer overlooks this issue might have something to do with the differences between his concept of evolution and ours. Spencer saw evolution as a universal organizing force. With such a concept of evolution, he basically assumed that organization should occur in evolving species. Observation of the organization of both body and behaviour appears to support this idea of evolution as an organizing force. Yet evolution is not inherently organizing. It only organizes there where selection pressure on organization exists.

Although views on evolution have changed since Spencer's time, the intuition that adaptation must involve organization is very much alive.

We find what we might call a Spencerian approach in connectionism. Artificial neural networks are often "evolved" using genetic algorithms, and attempts to interpret the resulting structures typically take the shape of a searching the networks for structures reflecting distinctions and relations in the task they are trained or evolved for. These are more often than not dishearteningly tough to come by. "Development of better analysis techniques for network structure" is by far the most reliable constant in connectionism's "future work" section. This is not to say that correspondences are never found (see e.g. Sejnowski & Rosenberg, 1986), and connectionists have eagerly used the few examples there are to defend connectionism from the charge that neural networks are unsuitable as models of the mind, but it remains unclear why some artificial neural networks do develop recognizable organization while (a vast majority of) others don't.

The intuition of connectionists searching for traces of the task environment in neural network structures is, I think, the same that inspired Spencer's view. In general, simulated evolution does not bear out the results this intuition predicts.

On the other hand, the function localization pursuits of neuroscientists too seem led by the same intuition, and they *have* uncovered scores of correspondences between species' brains and their environment, in the form of specialized brain areas and neural circuits for many of the environmental features that are of

importance for survival and reproduction in the environments these neural structures evolved in. No doubt, Spencer would have been delighted with modern day neuroscience (he did in fact dedicate chapters of *The Principles of Psychology* to the evolution of biological neurons and neural networks). However, given that having evolved in an organized environment does not imply correspondence, as evidenced by the scarcity of correspondences in evolved artificial neural networks, Spencer's account of the evolution of mind cannot adequately explain the findings of neuroscience.

If evolution is not inherently organizing, but only organizes there where selection pressure on organization exists, then for a Spencerian approach to evolution of mind to work, we must supplement it with an explanation of what organization-inducing selection pressures the coordinations between perception and action are exposed to. Such an explanation, then, would predict ways in which artificial neural networks can be made to evolve the sort of correspondences neuroscience presents, and could be verified using an artificial life model.

A little qualification is in order here. A correspondence-inducing selection pressure need not directly work on organization. Organization can be a side-effect. An example: if a given behaviour is best served by mixture of neuron types, then a mixture of neuron types might evolve under selection pressure for effective behaviour. As cells often divide "in place", this pressure might indirectly lead to a highly non-uniform distribution of neuron types with a semblance of functional compartmentalization. Maybe this sort of indirect influences could "passively" organize cognition to sufficient extent, and explanations of specific, direct selection pressures on organization are unnecessary.

However, there is a particular feature that we will argue may *actively* exert selection pressure exactly on organization of the kind Spencer describes. Spencer routinely uses species equipped with it in his examples, but does not as far as I have seen note it as especially relevant to the establishment of correspondences. This feature is learning ability.

At first glance this is an odd candidate. Why would one way of tweaking the connection weights suit us better than another? Learning algorithms for neural networks have been around for decades, are used at least as often as genetic algorithms to adapt neural nets to their tasks, and do not fare any better in producing correspondences. But the suggestion is not to *replace* evolution with learning.

1.6. Learning ability

Learning is easily mistaken for a form of behaviour, while in reality it is a wholly different phenomenon. Learning ability is a form of *phenotypic plasticity*.

Phenotypic plasticity is the phenomenon that an organism's phenotype can adapt to its environment. To qualify as phenotypic plasticity, such adaptation should have a lasting effect. When a chameleon or cephalopod changes its colour for camouflage or communication, this is behaviour. When bryozoans grow (permanent) defensive spines in response to chemicals signalling the presence of predatory sea-slugs in their pond (Harvell, 1986), this is phenotypic plasticity. When a female blue wrasse turns male once the male whose territory she inhabits dies (Gilbert, 2003), this too is phenotypic plasticity. A harsh border condition is hard to give for phenotypic plasticity in general, though we will see that between behaviour and learning ability a sharp demarcation is possible.

How is phenotypic plasticity relevant to our present discussion? Two levels of adaptation can do more than one. In some artificial intelligence applications, evolution and learning are combined to improve the chances of finding optimal solutions. Unsurprisingly, this works (see e.g. Belew et al., 1990; Whitley et al., 1990; Whitley, 1995). But there can more to the stacking of levels of adaptation. Within the artificial life literature, we find a lot of work on the effects learning processes can have on the course of evolution processes (Hinton & Nowlan, 1987; Maynard-Smith, 1987; Turney et al, 1996, Suzuki & Arita, 2004, 2007, 2008; Watanabe et al, 2008; Paenke et al, 2006, 2009; Yamauchi, 2007; Arnold et al., 2010), and within the bioinformatics literature, we find some work on interactions between different levels of evolution

(Crombach and Hogeweg, 2008). Interesting things can happen when levels of adaptation are allowed to interact; including, we will argue, organization of cognition.

Why should learning ability play a pivotal role in establishing correspondences, more so than other cognitive abilities? We first define behaviour and learning.

Behaviour

If we think of organisms as systems that take stimuli as input and give responses as output, then an organism's behaviour is the stimulus-response mapping it implements. Let S denote the set of stimuli an organism can distinguish, R the set of responses it can execute, and B the set of mappings from S to R . B then contains all possible behaviours for an organism with these perception and action abilities. We use $b_t \in B$ to denote the behaviour of the organism at time t , that is, a single mapping from the set of stimuli to the set of responses:

$$b_t: R \rightarrow S$$

Note that in this conception of behaviour, an organism's behaviour at a given point in time is *not* its response to whatever stimulus it observes at that time, but a complete mapping from *all* the stimuli it could observe to *all* the responses it could give. So an organism responding differently to different stimuli can still be exhibiting the same behaviour. Only when the same stimulus produces different responses at different points in time can we say that the behaviour changed (for convenience we are ignoring stochastic behaviours for now). This slightly strange definition of behaviour lets us distinguish differences in response that are due to differing stimuli from differences in response that are due to learning.

Learning

An organism might keep implementing the same behaviour throughout its life (i.e. have a constant stimulus-response mapping, $b_x = b_y$ for any two time points x and y within the organism's lifetime). In most organisms however, behaviour changes over time. This

can happen due to internal processes (e.g. maturation, wear) or due to exposure to certain stimuli (e.g. damage, information), and of course it can be advantageous (improve behaviour), disadvantageous (degrade behaviour), or neutral, in terms of the fitness effect on the organism. Here 'learning' will denote changes in behaviour that are caused by exposure to information and are advantageous⁴. With $b_i \in B$ representing behaviour at time i and $s_i \in S$ a stimulus received at time i , the way stimuli alter behaviour in a given organism (a mapping from behaviour-stimulus tuples to behaviours) can be represented as

$$c: (b_i, s_i) \rightarrow b_{i+1}$$

and the set of such changes as

$$C: (B, S) \rightarrow B$$

The set U of learning updates is then a subset of C . The criteria (being advantageous and being caused by information) for identifying which elements of C are in U are harder to formalize. What is advantageous hinges on the specifics of the environment, and also the requirement of "being caused by exposure to information" is not unambiguous. A blow to the head can simultaneously convey information and alter behaviour, but should this update be contained in U ? Fortunately, the boundaries of U being somewhat vague is no big problem. What is important is that with these definitions we can clearly distinguish between behaviour and behaviour change.

1.7. Why learning matters

With our formal tools introduced in the previous section, we are ready to explain the role of learning in the evolution of correspondences.

As discussed above, there can be countless many ways for an organism to implement a given behaviour, and evolution is not directly sensitive to this variation. Natural selection selects on fitness only, not on how it was attained. Consider two organisms, o^a and o^b , with

⁴ Of course it is possible to construe scenarios in which genuine learning ends up disadvantageous, but this characterization of learning suffices for our purpose.

the exact same behaviour b , differing only in that o^a implements b as b^a and o^b as b^b . They will, on average, score the same fitness. Thus natural selection cannot differentiate between them.

When we add behaviour change, matters change. Learning update u works on behaviour b . For u to change b , u must make changes at implementation level. To realize the same update function u , different implementations of b require different implementations of u . So we get u^a for b^a and u^b for b^b .

Say o^a and o^b discern the values of three binary variables in their environment, p , q , and r , and have two possible responses, x and y . At time 0, they both implement behaviour b_0 . Then some stimulus causes a learning update $u_{0 \rightarrow 1}$, after which both implement behaviour b_1 . Below we show both implementations of b_0 and b_1 :

Organism o^a : tabular implementation		
b^{a_0} :	$u^{a_0 \rightarrow 1}$	b^{a_1} :
(p , q , r) \rightarrow x		(p , q , r) \rightarrow x
(p , q , $\neg r$) \rightarrow y		(p , q , $\neg r$) \rightarrow \mathbf{x}
(p , $\neg q$, r) \rightarrow x		(p , $\neg q$, r) \rightarrow \mathbf{y}
($\neg p$, q , r) \rightarrow y		($\neg p$, q , r) \rightarrow y
(p , $\neg q$, $\neg r$) \rightarrow y		(p , $\neg q$, $\neg r$) \rightarrow y
($\neg p$, q , $\neg r$) \rightarrow x		($\neg p$, q , $\neg r$) \rightarrow \mathbf{y}
($\neg p$, $\neg q$, r) \rightarrow y		($\neg p$, $\neg q$, r) \rightarrow \mathbf{x}
($\neg p$, $\neg q$, $\neg r$) \rightarrow x		($\neg p$, $\neg q$, $\neg r$) \rightarrow x

Organism o^b : rule-wise implementation		
b^{b_0} :	$u^{b_0 \rightarrow 1}$	b^{b_1} :
IF ($p=r$)		IF ($p=\mathbf{q}$)
THEN x		THEN x
ELSE y		ELSE y

The loci that $u_{0 \rightarrow 1}$ changed are highlighted. We note two things. First, that in the tabular implementation, $u_{0 \rightarrow 1}$ had to change two loci whereas in the rule-wise implementation it had to change only one. In similar scenarios with more variables, the difference in number of loci to change would increase. Second, that the rule-wise implementation exploits a regularity that

the tabular implementation ignores, namely, that both before and after the update the correct response depends on the equality of the value of p to one of the other two variables. That conciseness goes hand in hand with exploitation of regularities has little to do with the specifics of the example we chose. The whole practice of data-compression is based on exploiting regularities in data to save space.

So in designing the implementation of u , the implementation of b makes a world of difference. A tabular b tends to be very disadvantageous for u : a single piece of learning data will often have to be "translated" into a huge list of changes to make (one for every stimulus for which the response should change). This requires u to contain a massive amount of very specific information. On the other hand, if b is implemented as a set of rules working on internal variables, with one of those variables *corresponding* to v , all that needs to be updated is the state of that one variable, and instantly all stimulus-response pairs will be in line with the new state of affairs. The only information u then must contain is which internal variable corresponds to v .⁵

Generally speaking, it helps u a lot if b is organized in such fashion that the changes in b that are important to the organism can be easily realized by u . The more the structure of b resembles that of (relevant aspects of) the environment, the less u will have to "translate" its incoming information from that environment before it can sensibly be applied to b . This then, is the core hypothesis of this thesis:

Learning ability benefits from similarity in organization of cognition and environment. Consequently, the evolution of learning ability should have an organizing effect on evolving cognition.

It might seem that we could have equally well made this sort of argument for ease of execution of b . Does that not also benefit from a rule-wise implementation, outside of considerations of learning? Not quite. A rule

⁵ In no way do we mean to suggest that that cognition is or resembles a rule system, we are merely using two extremes to illustrate how the implementation of b can make a qualitative difference to u .

system can well require many rule-firings to reach its conclusion (in fact, rule-systems need not terminate at all on any given input), while picking the right line from a table can be very efficient (and never fails to terminate). Efficiency in execution and efficiency of updating most certainly do not necessarily coincide.

This reveals a tension in the evolution of species with learning ability. There is both selection pressure on execution-efficient implementation of behaviour *and* selection pressure on update-efficient implementation of behaviour. Execution-efficiency cares little about correspondence, while update-efficiency greatly benefits from it. We might expect the degree to which the structure of a species' cognition reflects the environment to vary with the emphasis the species' niche puts on learning ability.

The picture this paints of the relation between complexity and correspondence differs from Spencer's. In Spencer's view, correspondence simply increases as species grow more advanced, while the observation above suggests that for non-learning species, correspondence degrades as we move away from the plane of interaction. For learning species however, the evolution of those parts of cognition subject to learning should show an increased tendency towards correspondence.

Thus we arrive at partial agreement with Spencer. Selective parts of cognition show correspondence: those parts involved in learning processes.⁶

One particularly interesting aspect of cognition that is *necessarily* not fully subjected to learning is learning itself. We can think of u itself as an organism mapping stimuli (learning data) to responses (behaviour updates). u 's Environment is then b . u 's Situation is then just like b 's when there is no learning ability: execution-efficiency is everything, update-efficiency does not matter. Thus u itself should not be expected to exhibit much correspondence internally, at least not until there is a "second level u " - let's call it w - updating u . In that case however, we should again not expect w to exhibit much correspondence. Such "second order learning" ability ("learning to learn") is

⁶ About other parts we say nothing. Other factors may exist that also induce selection pressure towards correspondence.

arguably present in some species⁷, and we will return to it briefly in the future work section, but it is mostly outside the scope of this thesis.

1.8. Conclusions chapter 1

The environmental complexity thesis states that cognition evolved in response to environmental complexity. We say that weaker and stronger variants are possible. Godfrey-Smith's version is plausible, but too weak to be informative. Spencer's is very strong, indicating environmental structure as the source of cognitive structure. In Spencer's view, the organization of the mind reflects the organization of the environment it evolved in; Mind and environment are connected by a relation of *correspondence*. Testable predictions can be derived from Spencer's view, but the present state of connectionism suggests that correspondence does not automatically extend to the coordination between perception and action, where cognition is to be found. Any given behaviour can be instantiated in countless many ways, all equal in terms of fitness. We hypothesized that a need for learning ability should help in organizing cognition after the environment. The more the organization of behaviour resembles the environment, the easier it is for learning abilities to make effective updates in it. With this hypothesis on the role of learning in the evolution of correspondences we concluded this chapter. In the next we explore the psychology and ethology literature to find out what types of natural learning exist, and if and how they fit into our hypothesis. As it turns out, the most extensively studied forms of learning don't, and it's no accident.

⁷ Humans appear to be able to “learn to learn”. For example, adults learning foreign languages often find that subsequent languages take less and less time and effort to learn.

2. The varieties of learning

We now turn our attention to learning in biological species, asking if and how biological learning processes signal any structural similarities between cognition and environment. Non-artificial learning ability is traditionally the domain of psychology, so we will start with a short overview of the types of learning distinguished in psychology.

2.1. The universal learning rules

Up to around 1960, Psychologists distinguished roughly four types of learning: habituation, sensitization, classical conditioning and operant conditioning. Later, latent learning was added to the list.

Habituation is simply the reduction in (behavioural or psychological) response over time to a stimulus that is presented continuously or repeatedly. Sensitization is the opposite (it might seem strange that response would increase, but think of the increase in psychological response to a dripping tap). We will not get into these here, but focus on forms of learning where new relations between stimuli and responses are formed (associative learning), starting with classical and operant conditioning.

Classical conditioning

Let s_u be an unconditioned stimulus and let r be the natural response an animal displays when presented with s_u . If s_u is consistently preceded by some other stimulus s_c (the “conditioned stimulus”), then eventually s_c will come to elicit r by itself. The paradigm example is Pavlov’s famous experiment: by consistently preceding the presentation of food (s_u) with a bell sound (s_c) for a while, dogs can be made to start salivating (r) in response to the bell sound alone.

Operant conditioning

In operant conditioning, a naturally occurring behaviour r_0 is rewarded to increase its occurrence.

Once r_0 is being performed at high frequency, rewarding is restricted to those cases where r_0 is followed by another natural behaviour r_1 . This will cause frequent occurrence of the series $\langle r_0, r_1 \rangle$. The series can repeatedly be extended in the same manner to make animals perform extensive series of actions.

Both classical conditioning and operant conditioning can be described as simple, general learning rules. It does not matter much what the s ’s and r ’s in the above descriptions are, the principles remain the same. Behaviourism held that all learning can be explained from these and closely related general learning principles.

Both these types of associative learning have been demonstrated in many species. Radical behaviourists such as Burrhus Skinner held that these forms of learning are all that is needed to explain behaviour.

Classical and operant conditioning do not jibe well with our hypothesis about learning ability imposing structure. This is because they are global learning rules, that require no information about the particulars of the structure of the environment:

Classical conditioning:

Stimulus s_c is followed by $s_u \rightarrow$ respond to s_c as to s_u

Operant conditioning:

Behaviour r yields rewards \rightarrow increase frequency of r

Returning to the discussion about different implementations of learning ability, we find both the tabular and rule-wise depictions of behaviour that we used in explaining our hypothesis ill-equipped to handle operant conditioning. This is because the rule is about frequencies. For operant conditioning to be possible at all, there must be a non-zero chance of the organism performing the target behaviour to begin with, otherwise we can never reinforce its occurrence. We could revise the tabular implementation to instead of actions, output probability distributions over the set of possible actions, and similarly, replace the rule-

system with equations yielding such distributions. We won't. It is easy enough to see how operant conditioning would amount to revisions in internal variables determining those probabilities.

Given that operant conditioning only requires change in the probability of the presented stimulus triggering the target response, the only pathways that need to be adjusted are those that were involved in the behaviour just preceding the learning episode. Learning ability has its work cut out for it regardless of cognitive structure.

That operant conditioning requires little in terms of correspondence can be clearly seen in the successful reproduction of such learning in Artificial Intelligence systems. The reinforcement learning algorithms used in machine learning strongly resemble operant conditioning. Performing certain action in certain states yields rewards, and when such a <state,action> pair occurs, the algorithms adjust behaviour so as to increase the frequency of its occurrence. Typically, such systems are not taught a target behaviour through step-by-step reinforcing of its elements in sequence, like in operant conditioning, but are simply left to run until they stumble upon the rewards for the complete behaviour. This is slow, but computers are fast. Of course there are more advanced variants for more complex tasks, but the point here is, that this type of algorithm is perfectly suitable for recreating operant conditioning. Consequently, if machine reinforcement learning does not require correspondence, neither does operant conditioning.

A look at how reinforcement algorithms work indicates that indeed they don't. Reinforcement learning algorithms have at their core a *value function* that assigns a quality to either each situation or each <situation,action> tuple. An action selection mechanism uses these quality values to control behaviour, but all learning occurs inside the implementation of the value function. This implementation is usually a *randomly initialized* artificial neural network. Evidently then, operant conditioning should not be expected to promote the evolution of correspondence.

A similar argument for classical conditioning is harder to provide. Unlike operant conditioning, classical conditioning cannot get off the ground without a strong unconditioned (i.e. innate) <stimulus, response> pair to link the new conditioned stimulus to. So random initialization is no option, and the innate behaviour must instead be either hard-coded or evolved. Furthermore, classical conditioning's potential for bringing about interesting complex behaviours is rather limited, so it receives little attention from AI researchers.⁸

However, if we return to the simple tabular and rule-based implementations of behaviour in the previous chapter, we can assert that to realize a classical conditioning update, we can blindly replace s_u with (s_u OR s_c) in whatever implementation and get the appropriate result. This suggests that classical conditioning too would do little to promote the evolution of correspondence.

That these types of learning do not call for correspondences suggests that we need to make our hypothesis more specific. At best it applies to *some* forms of learning.

2.2. Behaviourism versus the non-universality of learning

It is no coincidence that the most studied forms of learning do not fit our hypothesis. Learning ability was the focus of the behaviourist research paradigm, and even though behaviourism has been abandoned for roughly half a century, its categorization of types of learning survives, and classical and operant conditioning are still the objects of much research in psychology.

Nowadays there is little need to harp on how misguided behaviourism was, but behaviourism does give us a neat criterion for the correspondence-inducing potential of any given type of learning ability: any form of learning acknowledged by behaviourism is necessarily of no use for our purpose.

⁸ Neuroscientists, however, have occasionally used modified reinforcement learning algorithms to model classical conditioning, see e.g. Pan et al. (2005).

Behaviourism banned unobservable internal states and events from psychological theory. Only stimuli, responses, and reinforcement (along with a species' physical constraints on those) were allowed to figure in behaviourist explanations of behaviour. A newborn organism was assumed to be blank slate, and all behaviour the product of universal learning rules working on with whatever stimulation the organism's environment provides. In Teitelbaum's words:

[I]n any operant situation, the stimulus, the response, and the reinforcement are completely arbitrary and interchangeable. No one of them bears any biologically built-in fixed connection to the others. (Teitelbaum, 1966, pp. 566-567)

The idea that we should find traces of an organism's environment in what and how it learns is thus in conflict with the very foundations of the behaviourist paradigm⁹. Not only were such traces not the focus of attention, they were actively ignored or explained away when they cropped up in conditioning experiments.

Indeed, traces of the environment show up where learning *diverges* from behaviourism's universal learning rules. We should find what we are looking for in the observations that foreshadowed the downfall of behaviourism.

2.3. Non-universal learning

[In a sense, the universal forms of learning are also the most restrictive forms. They are extremely general in that they can be formulated without reference to particular stimuli or responses, but only at the cost of remaining completely blind to the *informational contents* of those stimuli. The moment a learning process shows a sensitivity to the information content of the stimuli involved, the universal rules fail to apply. Examples hereof are a dime a dozen. When such divergences from universal learning rules occur in classical or operant conditioning, they are called "*learning biases*".

⁹ Different types of behaviourism differed in whether they merely deemed innate cognitive structures off-limits to psychology or assumed them non-existent altogether.

One of the early examples of of learning bias is found in Garcia et al. (1966, 1967). In these experiments, rats were given a food item and subsequently exposed to nausea inducing X-rays. This causes aversion to the food that the rats ate before getting sick. Even if the interval between eating the food and the onset of illness was multiple hours, the illness would still induce aversion to that particular food. In most learning tasks, the interval between an action and its consequences should not exceed a few seconds, for rats to mentally connect them. An interval of hours would completely block learning in virtually any other task, but when the link between food and illness is concerned, the rats make a stronger connection as the interval approaches a certain, fairly long, length (a tidy optimum curve was found by exposing the rats to the X-rays at varying intervals after ingesting the food). That this bias makes sense is self-evident, but we would not usually ascribe an understanding of digestive function to rats. The bias appears to be innate, and reflects (as in "corresponds to") a physiological property of the rat quite external to its cognition.

The same experiment brought to light a second bias in the same learning process. The aversion the rats develop selectively affects smell and taste, but *not* appearance of the food. If food smells and tastes different from the food that (was perceived to have) caused them illness, but looks identical, they show no aversion to it, while they certainly have the perception to detect the similarity. Shortly after, it was shown that pigeons are similarly selective in their aversion, except that their aversion *is* tied to appearance, and *not* to smell, though they certainly perceive smell. Gould and Gould explain these findings from the species' natural environments: rats are nocturnal scavengers, relying primarily on smell to find food. Pigeons eat seeds, which typically have little or no odour. Responses to sensations that play little or no role in procuring food are less or not affected by learning about food.

Gould and Gould are not the only researchers opting for this avenue of explanation. It has been around in the psychological literature since well before behaviourism (Shettleworth (1972) points to similar explanations for learning biases given by Small (1901) and Schaeffer (1911)). Unsurprisingly, such

explanations were unpopular during the reign of behaviourism.

The recognition of learning biases is incompatible with behaviourism's base assumption that the animal mind is blank at birth, and acquires all of its content through application of universal learning rules. But what interests us here about the phenomenon is not simply that the newborn animal mind is no *tabula rasa*, but rather, that explanations of its innate content always seem to refer to aspects of species' environment or physical characteristics.

So while *in theory* classical and operant conditioning are at odds with our hypothesis, in practice they lend it support through the learning biases they reveal.

Over the sixties, behaviourism got replaced by cognitive psychology. With this change of paradigm, a type of learning that until then had been wilfully ignored (even though Tolman had been demonstrating its existence since 1932) finally got acknowledged by mainstream psychology. This type of learning is now known as *latent learning*.

2.4. Latent learning¹⁰

We illustrate latent learning using one of Tolman's famous experiments (1948). It goes as follows: A rat is put into a simple T-shaped maze. The rat is put at the bottom of the T. A left turn leads to a narrow dark box, with some food in it. A right turn leads to a wide white box, with just as much food in it as the narrow black box. The rat is left to explore the maze freely. Exploration yields food, but there is no selective reinforcing of any target behaviour otherwise. The

¹⁰ Tolman's research described here was originally presented as demonstrating rats' ability to use 'cognitive maps', and made no use of the term 'latent learning'. Early research focused on navigation behaviour that could only be explained by ascribing mental representations (cognitive maps) of the environment (often a maze) to the test animal (usually a rat). Once the same type of learning was found in non-navigation tasks too, the more general term 'latent learning' came into vogue, though the term 'cognitive map' is still used in the context of navigation tasks.

next day, the rat is put into a wide white box, with some food in it, and left to eat the food. Next, it is put into a narrow black box, and receives an electric shock. The rat has no influence on these events whatsoever, and no behaviour is being rewarded or punished. The next day, the rat is placed into the maze again. *The rat navigates directly to the wide white box on the right* (note that everything being equal, rats prefer narrow dark spaces). This behaviour cannot be explained in terms of classical and/or operant conditioning. To generate this behaviour, the rat must have integrated its experiences over the past two days, and have certain expectations of what is to be found at either site¹¹. The rat has tacitly and without reinforcement gathered information on day one and two, but the effect on behaviour remains *latent* until it is suddenly expressed when the situation of day three calls for it.

It's hard to find a clear definition of latent learning. Maybe the clearest definition is a negative one: latent learning is learning that cannot be explained as any of the other types. This works for classification, but it does little to characterize the phenomenon. We will go with the following:

Latent learning is the spontaneous exhibiting of advantageous but non-reinforced behaviour after exposure to pertinent information.

What's crucial here is the involvement of information. As remarked above, classical and operant conditioning are restricted by their blindness to the information contents of stimuli. Latent learning is not, and I think this is its core difference with the other two types, and what makes it highly relevant to the present project.

Only after putting together the information gathered on day one and two does the behaviour of Tolman's rats make sense. Explanations without reference to these pieces of information are doomed to fail.

The Tolman experiment is a textbook example of the phenomenon. Let's have a look at some less typical examples, to get a wider view.

¹¹ This is the reason Tolman's findings were initially ignored: any explanation of his rats' behaviour must refer to 'mental phenomena' (expectations, representations), which behaviourism forbids

Bees

Bees are known to communicate the location of good foraging locations by "dancing". The dance codes direction and distance to the location, and usually a sample is brought along to communicate quality. One of the questions this behaviour raises is whether bees handle an observed dance as a set of instructions, or if they "understand" them as indicating a specific location. Gould and Gould (1982) report on the following experiment: a feeding station was set up on a boat in the middle of a lake. Some workers from a nearby hive were trained to visit the feeding station on the boat, and returned to the hive to report their finding. However, despite their efforts to direct more workers towards the feeding station, none of the bees in the hive went to the feeding station. The boat was moved incrementally closer to the (far) edge of the lake, and once it was within 10 meters of the edge, the bees attending the dances started visiting it. Their rejecting of unlikely locations makes the hypothesis that dances are handled as instructions unlikely, and supports the idea that the bees observing a dance mentally locate the indicated position on some sort of mental representation of the area (although later research indicated that these representations might not be map-like (Dyer, 1991)). Areas differ, so this representation must be learned, but this learning process plays out without reinforcement (it pays off to have the representation, but learning and reward are so disconnected that the reward cannot be construed as the reinforcer of an operant conditioning process).

Human language

Human infants learning their native language show predictable patterns of errors. Verb inflection is initially learned on a verb-by-verb basis. This leads to correct inflection of both regular and irregular verbs, but the inflections need to be learned anew for each verb. In a later stage, children discover the regularity in verb inflection, and start applying their discovery to almost all verbs. This allows for instant correct inflection of new regular verbs, but causes them to also inflect as regular the irregular verbs they previously inflected correctly. Both the correct inflection of unknown regular verbs and the incorrect inflection of known irregular verbs are never

reinforced, and we must ascribe children a mental category for verbs to make sense of this behaviour.¹²

Additionally, we can identify some large classes of learning behaviour that require a capacity for latent learning:

Practice

When an individual learns a new skill, often repeated attempts are necessary before any reward is received. Although failed attempts don't provide reinforcement, they can very well provide information, so that every subsequent attempt is informed by the preceding failures, until eventually an attempt is successful. When a kitten tries to catch a toy dangling in the air, and fails because its jump was too short, a next attempt will not just be a random variation: it will feature a longer jump. Observation of the results of one's actions is key in these sorts of scenarios.

Play

Play is similar to practice in many regards, although the distance between skill and reward is even larger. Play yields no direct rewards (by definition), and costs energy; in species capable of conditioning only, play should be evolved away from. Play does however provide lots and lots of experience, information of the "doing such and so affects the environment so and such" sort, which is useful in handling situations where competent action is required and failed attempts are fitness-expensive. Play will be used as a template for the computational model we build in the next chapter.

As these examples suggest, latent learning comes in all shapes and sizes. It has proven notoriously hard to fit into a formal theory of learning. Because mental phenomena, information content of stimuli, and

¹² That language learning would be accountable for without reference to mental phenomena might seem bizarre, but this was one of the projects of behaviourism at the time of its demise. In fact, Chomsky's review (1969) of Skinner's book "Verbal Behaviour" (1967), (a behaviourist theory of language) contributed greatly to behaviourism's downfall (although the strength of Chomsky's critique is disputable, see MacQuorcodale, 1970).

innate, species-specific aspects of cognition play a crucial role, we cannot make do with universal rules like for classical and operant conditioning. The rules of latent learning, if we can call them rules, differ from species to species, stimulus to stimulus, and situation to situation.

Not surprisingly, there is a strong tendency in psychology to associate latent learning with insight, planning and intelligence, while classical and operant conditioning are often thought of as “dumb”, automatic processes.

Computational models of latent learning are few and far between¹³. Most of them are extensions to the reinforcement learning paradigm, adding either hard-coded models of the environment (unusual for maze tasks, but common for games, see Tesauro, 1995 for a famous example), or hard-coded specialized learning routines to enable agents to build mental models of the environment, which is then employed to improve performance on the learning task. We will discuss one such extension (Gérard et al., 2005) in the next chapter. Here the environment is not as directly present as when the model is innate, but the procedures for learning the model exploit fundamental characteristics (e.g. spatial coherence) of the environment.

Let's pull some of the above threads together. Passing from classical and operant conditioning to latent learning, we find cognitive psychologists shifting research paradigm, and suddenly speaking of intelligence and insight (using the terms “latent learning” and “cognitive learning” interchangeably). Meanwhile, we find artificial intelligence researchers suddenly equipping their learning agents with models of the environment, or with specialized means to build such models on the fly. Spencer equates intelligence/cognition/mind with a system of correspondences between the internal and external. Does this hint at a role for latent learning in the evolution of correspondence?

¹³ Technically, a lot of classical AI systems could be said to exhibit latent learning. These feature plenty of hard-coded Spencerian correspondences, but as discussed in the introduction such systems do little to *explain* their biological counterparts.

Consider the challenge of making a latent learning update in a given implementation of a behaviour. In the case of operant conditioning, we had a reward signal telling us whether the frequency of giving the given response to the preceding stimulus should be increased or decreased, and we could limit our modifications to the pathways involved in the production of that response to that stimulus. In classical conditioning, we just had to make a new stimulus mimic the effect of an unconditioned stimulus. In latent learning, we have no reward signal to work with, and no unconditioned stimulus to mimic. All we have are stimuli with whatever raw data they contain, and there are no a priori constraints on what parts of the implementation to adjust. It's evident that without further details about the environment, the needs and physical abilities of the organism, and the implementation of the behaviour, no choice of modification has a better than chance probability of improving the behaviour.

In this problem we find the role of learning in evolution of correspondence as hypothesized in the first chapter. With the universal rules inapplicable, we must bite the bullet and handle updates case-by-case, using whatever information is available in a given case. Here it pays off to have the implementation of the behaviour organized after the environment, so that an observation of some external feature P being in state p can be handled by updating the corresponding internal feature P' to state p'. The less the organization of the behaviour system resembles the environment, the harder the "translation" from observation to update.

We will illustrate this using another famous Tolman experiment (Tolman and Honzik, 1930) which is nowadays still used in AI approaches to cognitive maps (see e.g. Voicu and Schmajuk, 2002). A rat is allowed to explore a simple maze (see Figure 2). There are three paths from the starting point to an endpoint where a small amount of food is placed.

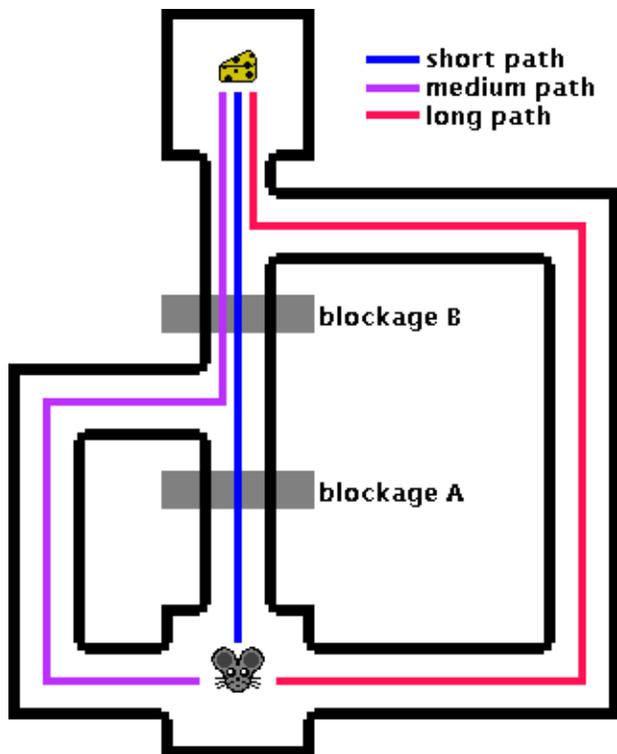


Figure 2. Three-path maze with two blocking points.

The paths have differing lengths. After a few trials rats will keep choosing the most efficient path to the food. Next, blockages are added. When blockage A is put in place, the rats try the short path, find it blocked, then backtrack to the start and choose the medium length instead. However when instead blockage B is put in place, the rats try the short path, find it blocked, then backtrack to the start and choose the long path without first trying the medium path. That the short path is blocked *at point B* implies that the medium path is also blocked, and the rats correctly make this non-reinforced connection. This indicates (some degree of) environment-cognition isomorphism at two levels:

Mental representation of environment

The rat has built up a mental map of the maze with positions on the map corresponding to positions in the maze. By updating a single position on this map from "free" to "blocked", the entire behaviour in this maze is adjusted correctly.

Spatial organization of experience

That the rat is capable *at all* of building a mental map of the maze shows that it organizes its experiences in a spatial manner. It understands (in at least an

operational sense) how its own movements bring about changes in its position in the maze, and integrates experiences at different moments into a coherent spatial representation of the maze. Visually distinct places (blockage present or absent) are recognized as the same if they are at the same "mental coordinate", and conversely, visually identical places can be recognized as distinct. The ability to create these mental representations indicates that spatial experience is interpreted within an innate spatial framework. The "blank" mental map a rat starts out with, the mechanisms for entering new data on the map, and the ability to use finished maps effectively to navigate a space can be said to correspond to spatial relations as rats encounter them in their natural environments¹⁴.

The spatial framework is what machine learning approaches to latent learning endow their species with. Without a specialized mapping mechanism, machine learning algorithms are incapable of building and using mental maps.

So, it seems like latent learning fits the bill, and indeed this is what we will be modelling in the next chapter. The distinction in two levels of environment-cognition isomorphism will come in handy when we compare existing models of artificial non-universal learning. But before we get to that, we have one more type of learning to attend to, one which psychologists prefer to leave to ethologists.

2.5. Cue-based learning

For lack of established terminology, we will call this type of learning cue-based here. We use this term for changes in behaviour that are triggered by fully innately specified trigger-stimuli ("cues"). No interpretation of the stimulus is necessary, and there needs to be no clear connection between the stimulus

¹⁴ What sort of spatial relations a species is endowed to handle varies; octopuses for example show a capacity for mental mapping of their hunting ranges (Mather, 1991b), but these ranges are never maze-like. Consequently, they perform poorly in the maze-tasks that make rats seem smart, while they handle open, sparsely landmarked spaces better than rats.

and the behaviour change it triggers. It's debatable whether this is genuine learning; the behaviour resulting from the update must be latently present in the organism, so the change is more like flicking a switch or entering a value into a variable, than that it is a genuine update of the existing behaviour.

A well-known example is the “imprinting” that occurs in new-born ducks. Shortly after hatching, young ducks learn what their mother looks like. This allows them to follow her wherever she goes, even when there are numerous other ducks around. Once the variables for mother duck's characteristics are filled, following behaviour can be commenced. Ducklings aren't picky: within a very limited time-window, anything that fits some very broad criteria can be imprinted as mother. Indeed, imprinting thanks its fame mostly to its comical failures.

The blue wrasse we mentioned earlier (Gilbert, 2003) can also be listed as an example here. When a male blue wrasse dies, one of the females in his territory switches gender and inherits the territory. This includes a switch to male behaviour. The observation that the territory lacks a male thus triggers a substantial change in behaviour, but the resulting behaviour was latently present before the switch was flicked (the cue provides no information on how to behave male).

Learning mechanisms like these, triggering on specific stimuli and/or at specific moments in development, without reliance on stimulus information content, are probably the most basic type of learning, and often thought of as “not quite learning” (hence the lack of attention from psychologists). From an evolutionary point of view though, this makes them very interesting. Their automatic nature puts them relatively close to the “life” end of the “life-mind” spectrum.

A nice illustration of how learning, memory and behaviour blend together in their primitive forms is found in the pill beetle (Migita and Moriyama, 2005). The pill beetle mostly walks straight ahead, but when its path is blocked, it will turn either left or right. The peculiar thing about it is that it alternates left and right turns. Past action (say, a rightward turn) somehow affects present action (leftward turn). Some sort of

behavioural switch is flicked whenever an obstacle is encountered. The switch is believed to be in its legs: it has a slightly asymmetrical walking pattern, and whenever an obstacle is encountered and a turn taken, the pattern is flipped.

Cue-based learning is neither general, like classical and operant conditioning, nor information-based, like latent learning. The variables and switches of cue-based learning processes often have clear correspondents in the environment, but given its dubious status as a form of learning, we will focus on latent learning instead. The main reason to call attention to cue-based learning is to have the distinction between cue-based and latent learning in our toolbox for discussing existing computational models of learning in the next chapter. The distinction is not always easy to make in abstract computational models.

2.6. On the classification of learning types

Before we conclude this chapter, we would like to call attention to the evolutionary unlikeliness of truly universal learning rules, and the way in which surviving behaviourist theory skews our view of learning.

What we call learning biases are only biases in context of the behaviourist's universal learning rules (classical and operant conditioning). Why explain the characteristics of a species' learning ability as deviations from a purely hypothetical limit case? If we exchange the behaviourist perspective for an evolutionary one, the question why a species cannot learn something that it would not usually need is rather odd.

An evolutionary view should start by asking what learning abilities a given species needs (i.e. what learning abilities its environment exerts selection pressure towards). Often the requirements of the environment are very limited, and cue-based learning can suffice. The more demanding environments call for the more advanced specialized learning abilities we class as latent learning.

Only in generalist species, living in a wide variety of environments (rats, humans), might we find packages of learning abilities varied enough to give off a semblance of universality, but even there we should not be tempted to overlook the many special-purpose mechanisms underneath.

2.7. Conclusions chapter 2

To summarize, psychology distinguishes different types of learning ability. Fictional or not, those that behaviourism acknowledges (classical and operant conditioning) do not constrain cognitive organization much, while those types behaviourism denied (cue-based learning, latent learning) expose traces of the environment in species' cognition. We restricted our hypothesis accordingly, and identified latent learning as the type best suited for our project. In the next chapter, we set out to let it evolve in a computational model, so that we can observe its effect on cognitive organization, and hopefully lend empirical support to our philosophical position.

3. An Artificial Life approach to latent learning

In the preceding chapters we worked out a revised version of Spencer's view of the evolution of mind. Evolution working to devise adaptive *behaviour* does not necessarily result in correspondences, but evolution of *learning ability* was predicted to induce correspondence. Assessing the correspondence-inducing potential of different types of learning ability as distinguished by cognitive psychologists, we found that latent learning fits the bill. Thus we ended up with the following hypothesis:

Latent learning benefits from similarity in organization of cognition and environment. Consequently, the evolution of latent learning should have an organizing effect on evolving cognition.

A hypothesis like this is of course hard to test in biological species: it is hard enough to establish whether a natural organism's cognition indeed features Spencerian correspondences, to establish whether learning ability played a role in their evolution seems close to impossible¹⁵. However, we might find empirical support by simulating evolution of learning ability in a model sufficiently flexible to allow for correspondences to emerge spontaneously – or (proponents of classical AI take note) not. Such models are the subject matter of Artificial Life.

Artificial Life is a young field, even compared with Artificial Intelligence. It got its name in 1986, from Christopher Langton. It has been drawing the attention of some philosophers (such as Godfrey-Smith) for its supposed potential to shed light on various aspects of the evolution of mind.

3.1. Artificial life and mind

“Artificial Life [...] can be conceived as a *sort of* philosophy – the creation *and testing* of elaborate thought experiments, kept honest by

¹⁵ It might be feasible to check for correlation, but causation is problematic.

requirements that could never be imposed on the naked mind of a human thinker alone. In short, Artificial Life research is the creation of prosthetically controlled thought experiments of indefinite complexity”. (Dennet, 1994)

In the philosophy of mind more than in the philosophy of anything else (save maybe for quantum mechanics), our intuitions can easily lead us astray, with different thought experiments about one and the same issue often leading to diametrically opposed conclusions. We could sure use some computational rigour here. While I agree with Dennett's characterization above, Artificial Life models of cognitive phenomena don't usually resemble their non-computational relatives much. We find no artificial twin earths, silicon zombies, or Chinese chat-rooms.

In practice, most cognition-related Artificial Life research falls in one (or more) of the following categories:

- a. Models to investigate the adaptivity of traits that are not a priori obviously adaptive, but evolved anyway. Examples are emotion (Dörner, 1995; Kato and Arita, 2004) and delayed reproduction (Bullinaria, 2009).
- b. Models to investigate the feedback loops in the evolution of social traits. As the fitness of a given social trait depends on the composition of the population, the fitness landscapes in these models evolve along with the population. Examples are language and linguistic diversity (MacLennan, 1991; Batali, 1994; Arita and Koyama, 1998; Sugiura and Arita, 2010), Cooperation (Axelrod, 1987; Suzuki and Arita, 2004), Theory of mind (Takano et al., 2005), Task specialization (Oiko, unpublished work).
- c. Models to investigate how a complex trait can evolve from simple elements. Examples are modularity (Jacobs, 1999, Bullinaria, 2007, 2009) and learning (Nolfi and Parisi, 1994, 1996; Soltoggio et al., 2007, 2008).

d. Models to investigate how different levels of adaptation interact. Examples are the Baldwin effect (Hinton and Nowlan, 1987; Maynard-Smith, 1987; Suzuki and Arita, 2004, 2007, 2008), the "reverse Baldwin effect" (Deacon, 2003; Wiles et al., 2005; Yamauchi, 2007), the hiding effect (Paenke et al., 2006, 2009) and Evolution of learning mechanisms (Arnold et al., 2010).

The model we develop in this chapter fits in categories c and d. We have a target trait (correspondence) of which we want to find out in what form, from what elements, and under what circumstances it can emerge (c), and our hypothesis predicts correspondence to emerge from the interaction between evolution and learning (d).

We start by identifying the ingredients we need to make a computational model fit for testing our hypothesis. First we ask under what circumstances learning is necessary, i.e. *what sort of environments put selection pressure on learning ability?* We will need such selection pressure to cause learning ability to evolve. Then we look at existing computational work on evolution and learning, gathering the nuts and bolts we need to build up our model.

3.2. What necessitates learning?

That a species will eventually evolve learning ability is not at all given. A species must have some flexible internal matter for learning ability to be evolvable at all, and then still, whether evolution "finds" learning ability is dependent on countless factors. But even ignoring the obstacles an evolution process might run into, whether learning ability is adaptive to begin with is highly dependent on the environment.

There are a number of factors that can put selection pressure on learning ability or other forms of phenotypic plasticity. We list them here.

A. The environment features *fast dynamics*

Changing environments require changing behaviours. If the changes in environment are slow, spanning many generations, they can be adapted to through evolution.

But when substantial changes in environment occur within the individuals' lifetimes, evolution has no coherent target to work towards, and will favour within-lifetime adaptation (i.e. learning). A special case are environments with a *repetitive dynamic*; here we may find learning ability specially tuned to the environmental dynamic (below in section 3.3.2. we discuss a model that explores this case).

B. The environment is static but *unpredictable*

when a species' habitat covers locations with differing conditions, it can occur that different individuals find themselves born in quite different environments, even when that environment is otherwise constant. This case is similar to that of environments with fast dynamics, in that evolution has no coherent target for adaptation, and will instead favour learning ability. It could be argued that the linguistic environments of humans fit in this category.

C. The genotype cannot fully code the phenotype

Non-detrimental mutations in genotype size are extremely rare. Without phenotypic plasticity, the phenotype would be restricted by genotype size. Phenotypic plasticity allows species to draw part of the information necessary for development from the environment. Examples can be found in research on the development of the visual cortex in mice: while visual cortex development is quite invariant in mice reared in natural circumstances, the visual cortices of visually deprived mice remain underdeveloped (Valverde, 1971). This "information expectant" (Greenough et al., 1987) quality of the developing brain has been shown to extend, with surprising specificity, to the most basic aspects of vision. The visual cortices of mammals usually contain detectors for all line orientations, but it was found that in cats reared in an environment with only horizontal or vertical lines, only the detectors for the orientations the cats are exposed to develop fully (Coleman et al., 1981; Tieman and Hirsch, 1982). This line of research indicates that there is a gray area between development, learning, and phenotypic plasticity. While stimulation-induced behaviour-affecting changes in neural structure are suggestive of learning, the developmental trajectory is quite invariant under

normal conditions. It relies on the environment, but it is guided innately rather than environmentally. Simply put, a genotype need not code all information needed for development, when it can code ways of extracting parts of that information from the environment using phenotypic plasticity. We might think of such genomes as compressed development schemes for which the environmental information acts as decryption key¹⁶.

D. The phenotype is unpredictable

Many birds and mammals (humans more than any other species) are *altricial*: individuals are born helpless, require extensive parental care, and must learn even the most basic skills they need for survival (e.g. locomotion, what is and what isn't edible, etc.). One reason for being altricial is to give information expectant development processes time to do their work (in a womb or egg, comparatively little stimulation can occur). But also the unpredictability of the own phenotype plays a role. Keeping one's balance in order to walk upright, for example, requires information about unpredictable physical specifics as the length of the limbs and the weight distribution of the body: a 'body schema' in Gallagher's (2005) terminology, or a 'tier-1 internal map' in Gould and Gould's (2007). The more the characteristics of the body depend on environmental factors, the less the genotype can equip the species innately with fully developed motor programs. In this regard a species' being altricial resembles the case of unpredictable environments; the unpredictability is in the body instead of in the environment, but both are external to cognition.

The difference with information expectant development processes may be one of degree rather than kind, but information expectant development is more of a passive exposure-driven process, while learning to walk or fly involves active experimentation.

This list may not be complete, but it gives an idea of the varieties of conditions that call for learning ability, and gives us a tool for characterizing the models we look at in the next section.

¹⁶ A search for other correspondence-inducing factors might find a candidate here.

3.3. Existing models

Making learning ability evolve requires more than just the right selection pressure. We must also devise a sufficiently mouldable species. This is tricky. When the *effects* of learning ability (on for example evolution processes or population dynamics) are studied, we can take learning ability as a given, and simply hard-code a learning mechanism. However, as our hypothesis says that correspondence *co-emerges* with learning ability, we should not start out with prefab learning ability, but with simple elements that evolution might organize into learning ability.

In chapter 1, we defined learning as *adaptive behaviour change*. Following that definition, we should get learning ability when we apply natural selection to *random behaviour change*. But devising an implementation of behaviour change that is fluent and flexible is no trivial task. Therefore we explore existing techniques for evolving learning ability, and see what we can use.

Most of the models we will discuss evolve some species of artificial neural networks. Neural nets are particularly suitable, not only because they allow a huge range of behaviours to be modelled using a single uniform formalism, but also because they can easily be shaped in an incremental fashion: small changes in connection weights produce small changes in behaviour. This makes them highly evolvable, unlike, say, classical rule-based systems, in which small random changes are more likely to either leave them functionally identical or cause them to break down completely.

Indeed, algorithms for shaping neural nets using either evolution (genetic algorithms) or learning (error back-propagation) are readily available. It might seem that all we need to do is stack the two on top of each other. Yet, there is a big difference between neural nets that *evolve and learn* and neural nets that *evolve learning*.

The most common form of neural network learning is error back-propagation. There are many variants on the algorithm, but within a given model the algorithm itself is invariable, save for a few tuneable parameters. Models combining evolution and learning are

comparatively scarce (when one simply aims to optimize connection weights, having two algorithms on the job just complicates matters). When we find both in a single model, evolution usually tunes parameters and/or searches for the most suitable network topology for a given task (Harp et al., 1989; Belew et al., 1990; Merelo et al., 1993; Whitley, 1995). But as the learning algorithm itself is not subjected to evolution, it never gains the ability to exploit the contingencies of the task environment, and consequently these models do not usually display any capacity for latent learning.

The limitations of such models straightforwardly reflect the limitations of the behaviourist's universal learning rules as discussed in chapter 2: they are insensitive to the information content of stimuli, and must consequently rely on either explicit reinforcement (reinforcement learning) or training data consisting of "correct" stimulus-response pairs (supervised learning). The few models that *do* evolve learning rules (such as Chalmers et al., 1990), are a step in the right direction, but still they evolve alternative learning rules for supervised learning, not latent learning.

Models that show a genuine capacity for latent learning are scarce. Maybe the behaviouristic view of learning still has more influence than it should, or maybe the computational research on *evolution and learning* detracts attention from the lack of research on *evolution of learning*. We will discuss a number of models that in some form or another bear upon the subject of computational latent learning.

3.3.1. Latent learning extensions to Reinforcement Learning

As mentioned in chapter 2, there are a fair few models that extend the reinforcement learning paradigm with specialized mechanisms for building cognitive maps of the environment. We discuss a model by Gérard et al. (2005).

Reinforcement learning systems, in short, let agents learn from the effects of their own actions. Behaviour in these systems is guided by a *value-function*, which

given a situation-action pair¹⁷, computes an estimate of the expected reward of taking that action and following the same behaviour-policy afterwards. Through their interactions with their environment, the agents learn a suitable value function to guide their behaviour. All that is needed is a reward signal. We mentioned the similarity with operant conditioning before. Agents experiment with different behaviours, find out what yields reward, and shape their behaviour so as to maximize total reward.

As is, reinforcement learning is not suited for implementing latent learning, due to the reliance on a reward signal. However, all explicit attempts I have found at modeling latent learning are extensions of reinforcement learning, aimed at exploiting the non-reward information a reinforcement learning agent encounters "along the way". The model of Gérard et al. piqued our interest by employing evolution for the non-reinforced aspect of the learning process.

The model uses a set of *classifiers* that characterize how actions transform situations (for example, a classifier might express that turning left by 90 degrees will cause an object that was in front of the agent before to be to the right of it after). If an agent has a set of classifiers that correctly expresses these transformations for a wide enough range of situations, it can "mentally" string a number of them together to oversee the results of extended action sequences. This way, individually unreinforced actions can be strung together to achieve distant rewards quicker.

The classifiers themselves can be acquired without reliance on rewards; their measure of quality is simply how well they predict actually observed transformations. Suitable classifiers are acquired using a genetic algorithm that plays out *within* the learning agent. The agent contains a population of classifiers, and the agent's observations of actual transformations provide a fitness measure to judge classifier quality by.

The agents indeed show latent learning ability: they quietly and without reinforcement absorb pertinent

¹⁷ Or just a situation alone, in case the results of single actions can be predicted, but for ease of explanation we limit ourselves to the more general case.

information, and use it spontaneously in situations where it's relevant.

In chapter 2 we used models of this sort as an argument for a relation between correspondence and latent learning. Computational approaches of latent learning invariably involve a specialized module for building some sort of mental model of the environment (in this case, that model is the set of classifiers). For our purpose, this is a problem. By hard-coding the mechanisms for modeling the environment, the pressure on the main (reward-based) learning mechanism to get organized is lost.

Although it features both learning and evolution, the Gérard et al. model is unlikely to support the sort of interaction between them that we are looking for. The hypothesis hinges on the interaction between stacked levels of adaptation, but what we find here are juxtaposed levels of adaptation. Both processes affect the agent's choice of behaviour, and the agent's behaviour produces the information (learning inputs & fitness evaluation) that drives both processes. What sort of interaction patterns might occur in such juxtapositions is an interesting research question in and of itself, but we will not go into it here.

The model highlights the importance of the relation between agent action and the transformation of one situation into another. Novel (as in non-reinforced) application of an action to a situation only makes sense if the agent has an understanding of what that action does to situations. We take this insight along for incorporation in our model.

3.3.2. Evolution of learning mechanisms

This model was designed to test the hypothesis that evolution can use plasticity differentiation in neural nets to accelerate learning in environments with structured dynamics (Arnold et al., 2010). Learning is implemented using error back-propagation, but the algorithm has been modified so as to allow for genetic shaping of the learning process.

The genotype of the networks codes innate connection weights, like is common when neural networks are

evolved. To give evolution partial control over the learning process, connections were additionally assigned individual *plasticity genes*. Usually, only connection weights are coded genetically, and a single learning-parameter defines the strength of the weight updates throughout the network. Using local plasticity genes, one for each connection, a non-uniform weight updating process could be evolved. The idea is that evolution might shape this non-uniformity in the learning process after the particular dynamics of the environment, and that this should improve their performance.

Over the lifetime of a single generation, the environment cycles through three “phases”, each requiring a different behaviour. This environment calls for learning ability due its *fast dynamics* (category A in section 3.2.). Stimuli are random binary vectors, responses are weighted sums of the vector elements. The weight vector to be used to compute the correct responses differs per phase. Figure 3 shows the weight vectors of the phases and an example computation.

```

phase 0: [+2 +1 0 0 -1 -2]
phase 1: [-1 -2 +2 +1 0 0]
phase 2: [0 0 -1 -2 +2 +1]

stimulus vector (random) : [+1 +1 +1 0 0 0]
phase vector (phase0) : [+2 +1 0 0 -1 -2]
multiplication : [+2 +1 0 0 0 0]
summation : +2 +1 +0 +0 +0 +0 = 3

```

Figure 3. weight vectors and example computation of target response (3) from stimulus [1,1,1,0,0] in phase 0.

The learning goal is to approximate this stimulus-response mapping as closely as possible. After every response, an individual receives reward proportional to how close its response was to the correct response, and adjusts its connection weights so as to reduce the difference between response and target response. The target behaviours are simple, and each can easily be approximated using error back-propagation. However, the environment switches phase multiple times through the lifetime of a single generation, so just as many times, the individuals must change their behaviour. Whenever the phase switches, performance drops: When the environment switches from phase 1 to phase 2, the individuals will initially still perform phase 1 behaviour, which scores badly in phase 2. So to attain high fitness in this environment, the learning process should be as fast as possible.

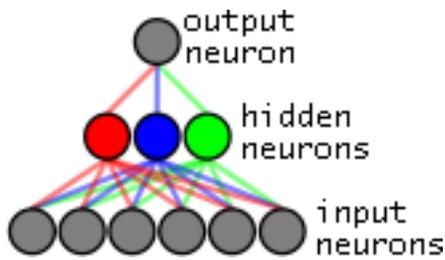


Figure 4. Neural network architecture used in Arnold et al., 2010. Colouring represents neuron specialization on the environmental phases (cf. Figure 3).

It was found that a highly specialized learning process evolved. The bottom layer connections lost virtually all plasticity. Looking at the hidden neurons, each coded a separate, rigid behaviour, matching one of the three phases. Meanwhile, the top layer connections evolved maximal plasticity, coding nothing whatsoever, but thanks to their high plasticity always quick to connect just the right hidden neuron to the output neuron.

This specialized architecture facilitates rapid learning as follows: learning a given stimulus-response pair would cause the connection between output neuron and the hidden neuron coding the behaviour for the current phase to strengthen, while causing the connections between the output neuron and the hidden neurons coding for non-current phases to weaken. This would not merely improve performance on that particular stimulus-response pair, but shift the *entire* behaviour in the right direction. So learning in phase n that stimulus S_a should produce response R_a would have the tacit side-effect of learning that completely unrelated S_b should produce R_b .

In the above, learning to link S_a to R_a is standard supervised learning. However, to explain the side-effect that S_b “mysteriously” ends up linked to R_b , we must call upon information content: to a net with this architecture, the pair (S_a, R_a) is indicative of *the environment being in phase n* . “The environment being in phase n ” being a very relevant state of affairs to the species, it evolved a special sensitivity to this information, and adjusts its behaviour accordingly when that information is encountered.

This model indeed appeared to produce correspondences. An environment switching through

three distinct phases produced a species with three distinct rigid behaviour modules (each consisting of one hidden neuron plus its downward connections) and a highly flexible module to allow switching between the behaviour modules. The need for efficient learning produced a highly organized architecture exhibiting conspicuous correspondences with the environment it evolved in.

The model supports the hypothesis that a need for environment-specific learning can organize a cognition, and indeed, this inspired the current project. However, the model uses a hard-coded learning algorithm, and while it managed to exploit information content of stimulus-response-reward triplets, the learning process was still supervised. Latent learning, by definition, is not; what we find here resembles learning biases more than latent learning.

3.3.3. Hebbian associative learning

Todd and Miller (1991) present an oft-cited model of associative learning that exchanges back-propagation for a purely Hebbian update rule in neural network learning. Hebb's rule (1949) says that simultaneous activation of neurons causes strengthening of the connection between them. It is easy to translate this into a neural network update rule, but usually no useful learning ability results. That Todd and Miller managed to make a Hebbian update rule do useful work is remarkable.

The species evolved is described as a stationary marine animal, with the ability to see and smell particles floating by, and an action repertoire consisting of "eat" and "not eat". Each particle is either nutritious or poisonous. Food particles smell sweet and poisonous particles smell sour, but the species' olfactory acuity is limited ($< 100\%$). Food and poison also differ in colour, and the species visual acuity is 100%, but which colour goes with which particle type differs per area (in effect, this linkage is decided randomly for each new individual, so this model fits in our category B (section 3.2.): the environment is static but unpredictable). The best strategy for an individual in this environment is to rely on smell first, note which colour it eats while relying on smell, and once a connection between

particle types and colours is established, rely on vision ever after. This learning process can be realized with Hebbian learning alone. We will not go into details, but the gist of the story is as follows: Fit individuals have a neuron detecting sweet smell, a neuron for each of the two colours, and a motor neuron (to eat). The innate behaviour is to eat any particle that smells sweet, so the "sweet smell" neuron connects rigidly to the motor neuron. The majority of the particles getting eaten will have whatever colour indicates food, so the neuron for that colour will more often than not be active simultaneously with the motor neuron, while the neuron for the other colour will less often than not be active simultaneously with the motor neuron. This correlation in activation then causes Hebbian connections between the food-colour neuron and motor neuron to develop. Eventually visual perception comes to dominate behaviour, and 100% correct discrimination is achieved.

Todd and Miller clearly demonstrated learning without error-backpropagation, but what sort of learning is this? I think it is closest to classical conditioning: a conditioned stimulus (colour) takes on the role of an unconditioned stimulus (smell). While reinforcement is not explicit (the organisms do not perceive whether what they ate was food or poison), it is doubtful that the Hebbian learning employed here would let us evolve anything more than stimulus associations.

3.3.4. Auto-teaching

Auto-teaching neural networks (Nolfi et al., 1994; Nolfi and Parisi, 1996) present a case of strictly reinforcement-free evolved learning ability:

“The environment does not usually provide cues that directly indicate to the individual how it should change in order to produce more adapted behavior. Natural selection is the only source of supervision for many living systems. However, organisms appear to be able to use environmental information, made available to them through their sensors, to trigger changes that make the individual more adapted to the environment.” (1996)

Although Nolfi and Parisi do not phrase their goal in psychological terminology, it should be clear that the sort of learning ability they refer to here fits our descriptions of latent learning and cue-based learning, but not classical or operant conditioning.

In the model designed by Nolfi and Parisi (1996), simple mobile robots explore an environment in search for a target, while avoiding bumping into walls. There are two variants of the environment: one with dark walls and one with bright walls. In the simulated evolution process, the environment alternates between the two variants from generation to generation. Because walls of differing brightness impinge differently on the robots’ sensors, the two variants require different behaviours. A far off bright wall gives the same level of sensor activation as a nearby dark wall. Consequently, the dark environment forces the robots to move around carefully, and start evading walls at sensor activation levels that would pose no threat in the bright environment. In the bright environment, this cautiousness is unnecessary and detrimental to the task of finding the target quickly¹⁸. Thus the robots should evolve two behaviour patterns and a way of learning through observation which one to use. The environment calls for learning ability because it is *unpredictable* (category B in section 3.2.).

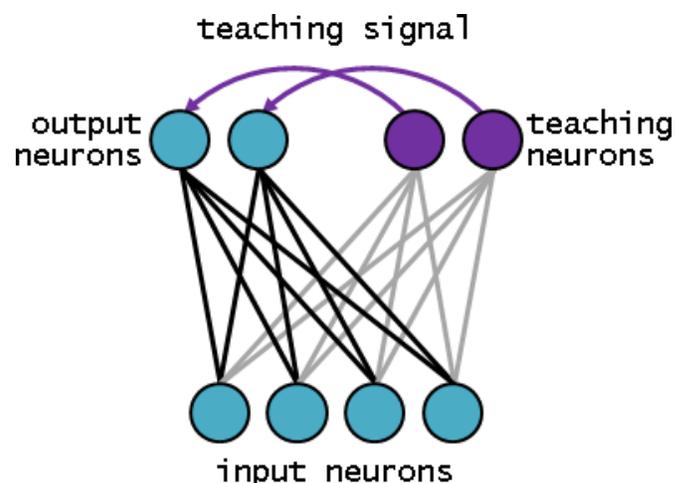


Figure 5: Auto-teaching neural network as used by Nolfi and Parisi (1996).

¹⁸ If the target happens to be near a wall, overcautiousness can prevent agents from finding it at all.

Figure 5 depicts the architecture of the nets Nolfi and Parisi used. Sensor activation values are fed into the input neurons and input neuron activation propagates to the output neurons, where the resulting activation pattern determines the activation of two motors controlling the robot's movement. Additionally, input neuron activation propagates to two "teaching neurons", and here lies the novelty of the architecture. In conventional error back-propagation learning, connection weights are updated on basis of an "error signal" that is computed from the difference between the output neurons' actual and intended activation. To know the intended activation, we must know the intended response, and in line with the above quote, Nolfi and Parisi set out to eliminate this element from the learning process. Therefore the error signal is replaced with a "teaching signal": the teaching neurons send their activation to the output neurons, and these use the received activation value in place of the error value in otherwise normal back-propagation updates. As the teaching signal is produced from the sensory input, this should allow the nets to learn from observing the information in their environment. Evolution sets the weights on the connections from input to teaching neurons, and also the initial weights of the connections from input to output neurons.

Nolfi and Parisi's simulations show auto-teaching networks capable of indeed evolving two distinct behaviour patterns for the two environmental variants, and a mechanism for selecting the right pattern on basis of nothing but observation, explicitly evading reinforcement. Does this make it latent learning? Ignoring hard-coded extensions to reinforcement learning for a moment, I think auto-teaching gets closest. However, the learning abilities of auto-teaching networks are extremely limited. Robinson and Bullinaria (2009) have experimented with auto-teaching networks in environments that change during the networks' lifetime (in Nolfi and Parisi's study, change of environment always happened in between generations). Auto-teaching networks handled a single switch halfway their lifetime just fine, but when we replace the environment for one that requires switching back and forth between behaviour patterns multiple times (i.e. exchange *unpredictability* for *fast dynamics*, see section 3.2.), auto-teaching networks fail. Robinson and Bullinaria show that this can be

overcome by equipping each connection with its own plasticity gene, like in the Arnold study above, but then still switching back and forth between two behaviours is of course a very restrictive kind of learning ability.

One might argue that learning in auto-teaching networks is cue-based learning instead of latent learning. Remember that in cue-based learning, learning triggers on exposure to a stimulus meeting certain innately defined constraints, but does not involve further processing of that stimulus than is necessary to recognize it as that cue. While switching between two innate behaviour patterns could easily be driven by a cue, it is not clear what that cue would be in the Nolfi and Parisi study. As mentioned in the previous chapter, the line between cue-based learning and latent learning can be fuzzy; auto-teaching networks present an ambiguous case.

Whether or not auto-teaching networks exhibit Spencerian correspondences is hard to assess. At the most general level, we're bound to find two distinct weight configurations for the two environmental variants, but Nolfi and Parisi make no attempt to interpret these weight configurations.

3.3.5. Neuromodulation

Soltoggio et al (2007, 2008) present a promising technique for evolving learning ability, called *neuromodulation*. We focus here on their 2008 paper. In this model, agents must collect rewards in a double T-maze (see Figure 6). Rewards are located in all four arms of the maze, but one of them is big while the others are small. After collecting a reward, an agent must return to the home position before it can collect the next reward. To maximize its fitness, an agent should collect the big reward as quickly and often as possible, but the big reward occasionally changes position (so this model employs *fast dynamics* to put selection pressure on learning ability, category A in section 3.2.). Ideally then, an agent should explore the arms of the maze until it finds the big reward, remember its location, and subsequently keep revisiting the remembered location, until the big reward changes position, at which point exploration should be resumed.

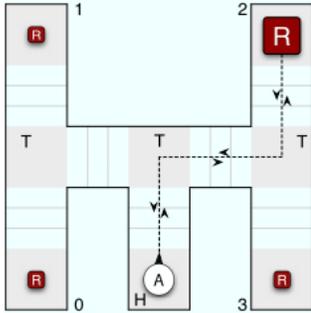


Figure 6: The double T-maze environment used in Soltoggio et al. 2008. A=agent, H=home, R=reward.

Agents in this model are controlled by neural networks evolved to learn using neuromodulation. As mentioned in the discussion of Todd and Miller's model, Hebb's rule alone does not explain learning. Neurological research showed that the degree of strengthening of neural connections depends not only on neural activation, but also on numerous biochemical factors such as neurotransmitter concentrations. Neuromodulation is a blanket term for those biochemical factors. The most well-known example is the release of dopamine that occurs in conditioning experiments when a reward is received.

Soltoggio et al. allow for neuromodulation effects to evolve in artificial neural networks. This is achieved by adding *modulatory neurons* to regular neural networks. Calling these elements neurons is potentially confusing, as they do not actually represent real neurons, but rather, non-descript biochemical processes. A neuromodulatory neuron receives and processes activation like a normal neuron, but sends out *modulation* instead of *activation*. This modulation determines not the activation level of the receiving neurons, but the *plasticity* of those neurons' outgoing connections. Those connections are then updated on basis of the plasticity value and the activation of the neurons they connect (the update rule is essentially Hebbian, except that the resulting change in strength is multiplied by the present plasticity of the connection before it is applied). Like the connections of the teaching neurons in auto-teaching networks, the connections of modulatory neurons themselves are never updated, and fully determined by the genotype. Much like the modulatory neurons do not represent actual neurons but biochemical processes, their connections do not represent actual neural connections, but rather the influence of perceptions

on biochemical processes (incoming connections) and the influence of these processes on neural plasticity (outgoing connections). Neuromodulation does away with (biologically implausible) back-propagation altogether.

Soltoggio et al. let not only connection weights, but also network topology evolve. Networks with neuromodulation enabled were shown to perform close to the theoretical maximum on the double T-maze task. The technique clearly shows potential, and we will adopt it for the model to be proposed in the next section.

First however we ask: *does the model of Soltoggio and Bullinaria show latent learning?* They never mention the term. It is clearly not their aim. Trying to classify the learning abilities demonstrated by the Soltoggio et al. model poses some interesting problems.

That there is a neuron marked REWARD among the inputs of these networks is suggestive of operant conditioning, but the rewards have a different role than in operant conditioning. Rewards are not used to reinforce the target behaviour; collecting rewards is *part of* the target behaviour, so what is learned from the rewards in this model is not simply to increase the frequency of the behaviour preceding the reward. Also, not only the REWARD neuron induces learning. The species in this model come to recognize certain states of the world as indicating that a specific change of behaviour is called for. This recognition and the corresponding change in behaviour are part of an innate cognitive structure. The environment never informs the species of the correct change in behaviour. This is indicative of latent or cue-based learning. Should we then expect to find Spencerian correspondences in this model? Going by the ideas developed in the preceding chapters: yes. And indeed we do, but unfortunately they are hard-coded parts of the setup.

The nets in the Soltoggio et al. study are given highly pre-processed input. The input neurons do not receive raw sensory input, like in the auto-teaching model above (section 3.3.5.). Instead, each input neuron is assigned a single relevant piece of information: "being at the home position", "being at a turning point",

“being at the end of a maze arm” and “size of reward at present location”. The input stimuli perfectly coincide with the necessary information. The input neurons certainly correspond, but they do little to support our point. Had the nets taken raw sensory data for input, and had they then evolved separate neural structures specifically sensitive to the relevant features (home, turning point, reward etc.), then this model might have supported our point. As it stands, the Soltoggio et al. model could be said to assume correspondence.

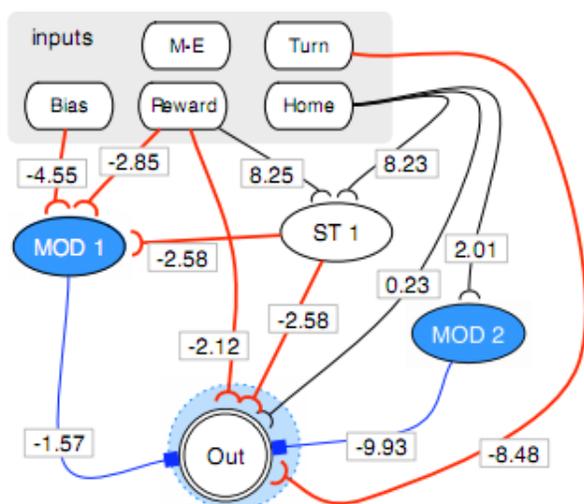


Figure 7. Example evolved neuromodulation network from Soltoggio et al., 2008. It has 2 modulatory neurons and one hidden neuron. Black connections are positive, red connections negative, and blue connections are modulatory.

As it stands, the learning ability of these neuromodulation nets seems closest to cue-based: a stimulus signalling a certain state arises, and the corresponding learning update is applied. The information content of the stimulus and the resulting update have little to do with each other; they just so happen to be linked. It’s clear that these nets’ learning is triggered by their observations, but it’s not clear that they learn *from* their observations.

3.4. Avoiding Cue-based learning

We found that Auto-teaching and Neuromodulation models get close to evolving latent learning, but in both cases it is hard to say with certainty whether their learning is latent or cue-based.

In chapter 2 we distinguished latent and cue-based learning by the role of the information content of the stimuli in the learning process. In cue-based learning, a cue stimulus triggers the innately associated update, while in latent learning, the stimulus is “interpreted in context of the present situation”, so to say. The distinction is clear in theory, but in practice it can be hard to distinguish the two (see Gould and Gould, 1994 for examples). The reason is that even if the stimulus merely serves as cue, it must still be recognized as an instance of that cue, and recognition and interpretation can look a lot alike.

In fact, what looks like latent learning can often be explained alternatively as cue-based learning with a huge list of cues, quite like we could explain seemingly intelligent behaviour as a huge list of stimulus-response pairs. Especially in simplified computational models where the stimulus and response spaces are limited, this is an issue to consider. We cannot fully dodge the problem here, but Nolfi and Parisi offer a useful observation:

“By acting on the environment the individual can change either the environment itself [...] or it can change its own physical relation to the environment [...]. Thus, sensory input in ecological¹⁹ networks is a function of both the independent properties of the environment and the individual’s behaviour. It is the interaction between what the network does and the external environment that will decide which inputs are seen by the network during learning, in what order, with what frequency, etc.” (1996)

An organism is itself partially responsible for what stimuli it receives. It can probe around and observe the effect of different actions in different situations to gather the data it needs to learn what it needs to learn. The latent learning ability found in the Gérard et al. model crucially depends on this insight: their classifiers express how situations are transformed by the agent’s actions.

¹⁹ “Ecological” here can be read as “situated”.

One way to steer away from cue-based learning would be to have one's virtual species learn about the effects of their own actions on their environment. This would require them to at least integrate a given perception with whatever they did to cause that perception, forcing some degree of interpretation on top of simple stimulus-recognition. This sort of scenario calls for latent instead of cue-based learning. We will use a simple scenario of this kind.

3.5. A model of the evolution of latent learning

Using bits and pieces from the models introduced in the previous section, we now build our own:

- We adopt the neuromodulation framework of Soltoggio et al.
- Like Nolfi and Parisi, we keep the learning process strictly reward-free.
- We let agents learn from observation of the results of their own actions, like both Gérard et al. and Nolfi and Parisi.
- We use a similar line of interpretation of evolved network structure as in the study on the evolution of learning mechanisms (Arnold et al, 2010).

3.5.1. Model concept

The basic idea is as follows: neural network-controlled individuals "catch" "prey" to increase their fitness, and higher fitness leads to increased number of offspring. Thus over the course of evolution, prey-catching skill should improve. However, in order to catch prey, individuals must be able to move about effectively in their world. Here we introduce an element of uncertainty, to necessitate learning. It's of the *uncertain phenotype* type (section 3.2.). Each network has four output neurons. These are linked to four actions (step forward, turn left, turn right, jump forward), but the wiring from outputs to actions is generated arbitrarily for each individual. The wiring is always one-to-one, so every individual is capable of every action, but to assemble those actions into advantageous prey-catching behaviour, the species must employ latent learning.

As our aim is to make latent learning evolve, our organisms receive no external guidance or feedback. The only information available for the learning process is the effect an output has on the environment, and the internal state of the network. Thus no general learning algorithm can be used. The learning process itself is left to evolve out of random behaviour change, using neuromodulation.

3.5.2. Environment

The environment is a 20 by 20 torus-shaped grid. Every cell in the grid is always in one of two states: 0 or 1.

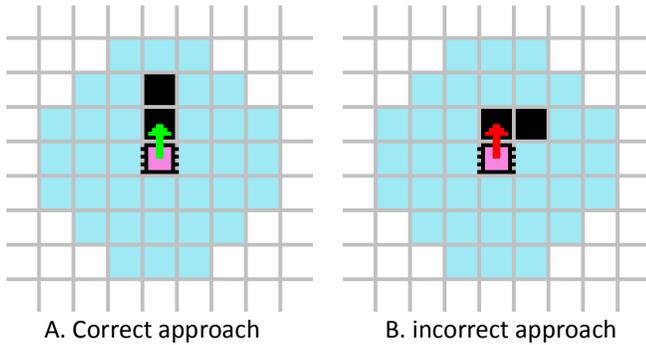
0-cells are empty; stepping or jumping on them has no effect.

Single 1-cells (1-cells whose four adjacent cells are all 0) can be stepped and jumped on, but this too has no effect. We will call a single 1-cell a "toy". Toys have no effect on fitness.

Double 1-cells (a pair of adjacent 1-cells) can be interacted with: stepping or jumping on them causes them to disappear (i.e. both of the 1-cells turn 0), and has an effect on the individual's fitness. We will call a double 1-cell a "prey". The fitness effect depends on how the prey is approached. Figure 8 indicates the two different types of approach. In general, prey should only be stepped or jumped on when it's oriented vertically in the individual's field of vision. If a prey is approached correctly, the interaction increases the individual's fitness by one (the prey is eaten). If a prey is approached incorrectly, the interaction reduces the individual's fitness by one (the prey injures the individual and flees).

Individuals each have their private world to live in, so they are not affected by the actions of other individuals. They first live for 800 time steps in an environment with only toys (80 pieces), while learning is enabled. Then they move to an environment with only prey (40 pieces, 20 of each orientation), where they live for 400 time steps while learning is disabled. As only prey affect fitness, the behaviour in the second environment determines the individual's fitness. The reason to split the lifetime into these two phases is to

ensure that learning and reward are strictly separated. It is debatable whether this is necessary (prey disappears the moment it is stepped on, so no straight-forward reward states exist anyway), but we deemed it beneficial for conceptual clarity.



A. Correct approach B. incorrect approach

Figure 8. interacting with prey:

In (a), taking a step forward will yield +1 fitness.

In (b), taking a step forward will yield -1 fitness.

Similarly, the choice to make different approaches produce different fitness effects serves to disconnect the learning process from the performance phase. Making the value of prey approach-dependent renders 1-cells by themselves meaningless. A behaviour of simply approaching 1-cells in general leads to the same fitness as avoidance, or, for that matter, altogether random behaviour. By rendering single 1-cells meaningless, it is ascertained that learning is latent: using whatever form of general learning ability, no sensible strategy can be acquired in the toy environment where learning takes place. The only information provided by the toy environment is how the individual's actions affect its position and orientation in the environment.

3.5.3. Species

Perception

Individuals cover a single cell. They have a 37-cell field of vision (FOV for short, see figure 8 for its shape). Every step, the content of the world that falls within the FOV is copied to the input neurons of the individual, taking into account the individual's orientation.

Action

As mentioned, the species is capable of 4 actions (see Figure 9).

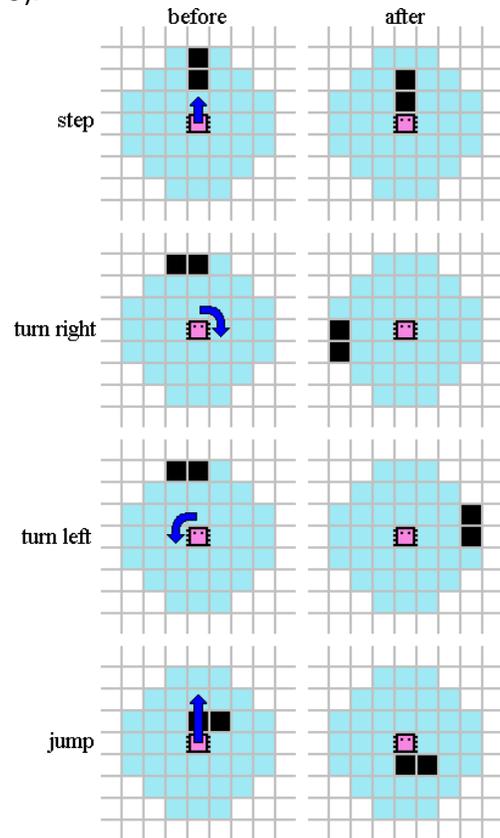


Figure 9. movement actions of the virtual species.

Step: moves the organism 1 step forward in the direction it's facing.

Jump: moves the organism 2 steps forward in the direction it's facing. The skipped cell is not touched, so this action can be used to avoid incorrect approach of a prey without having to move all the way around it.

Turn left/right: changes the organism's orientation by +90 or -90 degrees. Turns don't affect position, and as the shape of the FOV is rotationally symmetrical, no new information enters the FOV.²⁰

²⁰ The FOV is rotationally symmetrical precisely for this reason: unlike step and jump, either turn action has a corresponding action that can undo it (the opposite turn). If a turn would introduce new information, the unpredictable new state could cause the organism to make the opposite turn next, catching it in an unescapable loop. With a rotationally symmetrical FOV, turns only transform the FOV

When it comes to moving quickly, jumps are more efficient than steps, but steps are necessary to align oneself correctly with prey. The effect of either turn can be attained by repeating the opposite turn 3 times, but this is inefficient, taking 3 time steps instead of one. All actions are sufficiently important for individuals to need all four to be really successful.

Brain

The networks we use are similar to those in the neuromodulation model of Soltoggio et al., though instead of letting the network topology evolve, we take a fixed 3-layer topology. Our model too uses regular and modulatory neurons, but additionally adds a (fairly trivial) new type: echo neurons. See figure 10 for the topology we used.

The regular neurons in the network are no different from those in conventional neural networks. Each of the input neurons receives the state of one of the 37 positions in the FOV (0 or 1), and the input additionally contains a bias neuron (which always has an activation value of 1), and a “noise neuron”, which takes on random real values in the [0,1] range (this allows for evolution to introduce some randomness in the behaviour). For the hidden neurons, we compute the activation by applying the hyperbolic activation function to the sum of the incoming activation. Output neurons sum incoming activation but use no activation function.

The output layer consists of four (regular) neurons. When the activation from the input layer has propagated all the way to the output layer, we read the activation of these four neurons and select the one with the highest activation (if output are tied for highest activation, no action is performed). This neuron’s action (whichever that is under the individual's random output-action wiring) will be performed. The outcome of the selection process is also “visible” to the network itself: the selected output neuron’s activation is set to 1 and the other three have their activation set to 0. This could be modelled with

content, and predictably so, without adding or removing content. Loops can then simply be selected against by evolution. The symmetry of the FOV saves us the overhead of equipping the species with some form of memory.

excitatory reflexive connections and inhibitory connections between every pair of output neurons (we might even try to leave it to evolution to devise this structure), but for the sake of simplicity it has been hard-coded (presence of loops would complicate propagation). Such mutual inhibition of motor programs is standard equipment of natural brains (in primates found in the basal ganglia and cerebellum, see e.g. Rubchinsky et al., 2003; Humphries et al., 2007; Houk et al., 2007). It prevents organisms from attempting to simultaneously execute conflicting motor programs. Although our implementation is a gross simplification, it suffices for our purpose.

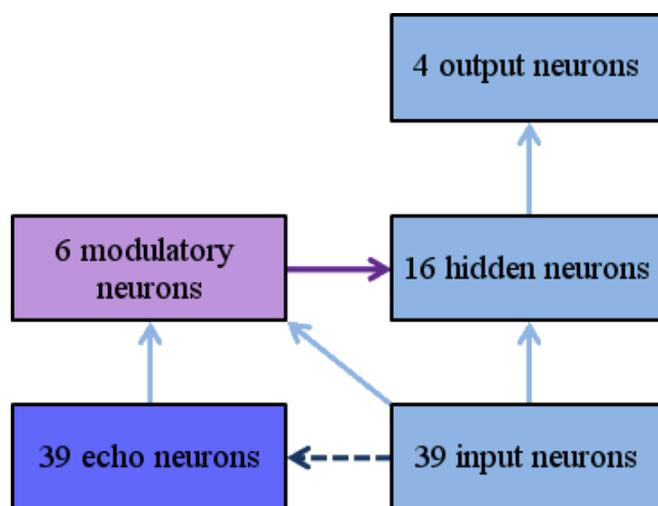


Figure 10. Neural network topology. Light blue blocks/arrows indicate regular neurons/connections. The input block contains a bias neuron and a neuron providing random noise. Solid arrows indicate that blocks are fully connected. The dashed arrow indicates a one-on-one connectivity with one time-step of lag (see text).

Modulatory and echo neurons mostly behave like regular neurons, but their connections are slightly different. There is one echo neuron for each input neuron, connected one-to-one, and these connections have one time step of lag. So the block of echo neurons is simply a copy of the input block as it was one time step ago. Echo neurons only connect to modulatory neurons (though this is no necessity). Otherwise they function like normal input neurons.

The reason to add echo neurons is to allow the organism to learn from the effects its actions have on its situation. Actions transform the content of the field of vision in a partially predictable manner. The step action, for example, predictably shifts the content down by one cell, and puts some unpredictable new content into the upper rim of the field of view. The resulting state tells us little about the action; to understand the effect of the action, we need to see the difference between the states preceding the action and resulting from the action. This is why we add echo neurons.

As in the Soltoggio et al. model, modulatory neurons differ from regular neurons in that they send out *modulation* instead of activation. This modulation serves as the learning rate of the receiving neurons. Modulation is computed in exactly the same way as activation, but instead of being summed to the receiving neuron's activation it is summed to its modulation. As can be glanced from the figure, only the hidden (i.e. middle layer) neurons ever receive modulation. Modulatory neurons themselves never receive modulation.

During the learning phase, individuals work through the following steps:

1. The part of the world that falls within the organism's FOV is copied into the input layer.
2. Activation from input layer propagates up through the regular neurons.
3. The action linked to the output neuron with the highest activation is performed.
4. The activation pattern on the input neurons is copied to the echo neurons.
5. The part of the world that falls within the organism's FOV is copied into the input layer.
6. Activation from input and echo layer propagates up to the modulatory neurons and sets the modulation value of the hidden neurons.
7. The hidden neurons' outgoing connections are updated.
8. Return to step 2.²¹

²¹ Step 1 is outside the main loop, as there can be no learning before the first action is performed.

The weight updates are computed using the following equation:

$$\Delta w_{ij} = 0.01 \cdot m_i \cdot a_i \cdot a_j$$

Where W_{ij} is the weight of the connection from neuron i to neuron j , m_i is the modulation at neuron i , and a_i is the activation of neuron i . Weights are clipped to the range $[-2,+2]$.

As action is determined by activation only, and modulatory neurons don't contribute activation, and echo neurons only project to modulatory neurons, the action for a given situation is determined fully by the regular neurons. When learning is deactivated (as is the case in the performance environment, where fitness is determined), the echo and modulatory neurons are idle.

As mentioned above, the wiring from output neurons to actions is generated randomly for every individual. Because there are 4 outputs and actions, there are $4! = 24$ possible one-to-one wirings from output neurons to actions.

3.5.4. Genetic algorithm

The initial connection weights of the network are coded in a genotype. A population of 90 networks is evolved using a simple genetic algorithm. New generation are produced from the preceding one using Rank-based selection with elitism, and asexual reproduction with single-point mutation. For details on the algorithm, see appendix A.

3.5.5. Why this model?

Does this model match our requirements? We have made every effort to ensure that the learning process is reinforcement-free. Direct reinforcement is explicitly avoided by using separate learning and performance phases. This alone does not necessarily suffice: certain states may take on the role of reinforcer (for example, had prey always yielded +1 fitness, then "standing on a toy" might be used as a reward-state). The unusual

handling of food items is chosen to ensure that no such indirect reinforcement can exist either. For details, see appendix B.

Furthermore, we have ensured that the stimuli that drive the learning process are not atomic in the sense they are in the Soltoggio et al. model: none of the input neurons by themselves signal a situation to learn from. The information to learn from is in the *patterns of change* in the content of the field of view as they occur in response to the individual's actions. And this information must be combined with the agent's recollection (in the form of neural activation) of what it did to cause the perceived change. This should rule out simple cue-based learning.

3.6. Results

We ran the simulation 10 times. Additionally, we ran a simulation in which there is only a single wiring from output to action, and no learning phase (also 10 runs). As no learning takes place in these runs, the nets have no echo neurons and no modulatory neurons (resulting in a far smaller genotype). With the uncertainty of what wiring a given individual will have to deal with removed, these runs can fully rely on evolution to devise fit behaviour. These runs serve two purposes: they provide a measure for assessing the quality of the learning ability that evolved, and they let us compare the internal wirings evolved with and without learning, so that we can assess what effect evolution of learning had on the emergence of correspondence. Figure 11 shows population best and average fitness as a function of time for both types of run. Note that this experiment does not follow the common pattern of comparing performance of learning and non-learning populations on one and the same task. The learners have random output-action wirings, while the non-learners have a fixed one, so the learners' performance will approach the non-learners' performance to the extent that learning is successful. Consequently, the performance of the non-learners does not set a bar for the learners to raise, but rather, sets an upper bound for the learners to approach.

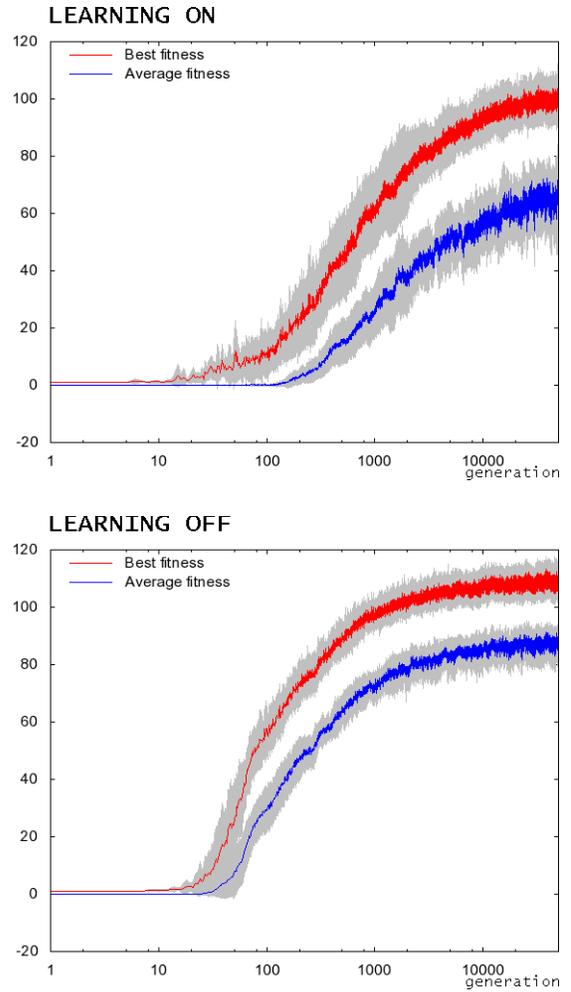


Figure 11. Generation best and average fitness for evolution processes with and without learning. Runs with learning used a random output-action mapping, runs without learning used a static mapping. Averages over 10 runs each. Gray areas indicate the standard deviation over the runs. X-axes in logscale.

Why this unusual comparison? Should we not use the same type of output-action mapping (either static or random) for both learners and non-learners? Doing so would be uninformative. Remember that we introduced the random output-action mapping for the purpose of necessitating learning ability. Introducing this necessity in a species that does not have the means to evolve learning ability is pointless: they simply cannot handle it, and would be forever stuck at near-zero (see appendix C for a run demonstrating this fact). Conversely, letting learners evolve with a static action-mapping should lead to fine fitness, but it would be hard to assess how much of a role learning

played, as learning would not be necessary. Our aim is to compare a neuro-cognitive structure that results from a combination of learning and evolution with one that results from evolution alone. Thus one set of experiments should have both the need and means to evolve learning ability, while the other should have neither.

The size of the difference in population best fitness at the end of the simulations (about 109 for non-learning populations versus about 100 for learning populations) indicates that the evolved learning ability is quite effective. The difference in average fitness is larger (about 87 versus 65), but this is to be expected given the larger genotype of the learners (evolving a larger genotype with the same per-gene mutation rate leads to a higher proportion of dysfunctional mutants, which drag down average fitness). That evolution is slower in learning populations is a logical consequence of the larger genotype and harder task. Overall these results show that the learning populations succeed in evolving effective learning ability.

For confirmation that our addition of echo neurons is crucial for learning to evolve in this environment, see appendix D.

3.7. Analysis

Next we assess whether the presence of learning ability induced any increase in cognitive organization. In advance, we had no clear idea of what shape such organization should take within the networks, so we search for significant differences between networks evolved with and without learning, and then see whether these differences are differences in organization. To allow such comparison, we need a way of visualizing the connection patterns within the networks.

For hidden layer neurons we can plot the incoming and outgoing connections as in Figure 12. For an explanation of how these images are created, see appendix E.

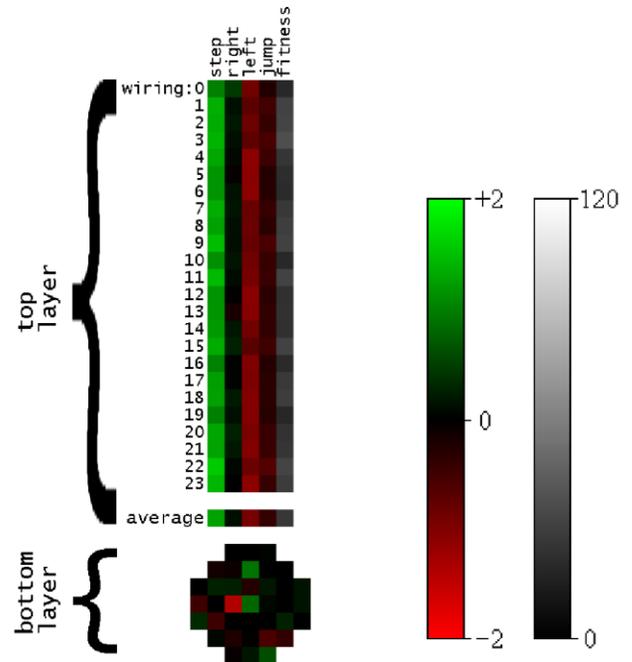


Figure 12: example connection plot for a single hidden neuron. Green and red indicate positive and negative connection strength, respectively, ranging from -2 to +2. Fitness values are shown as grayscale values, ranging from 0 to 120. The bottom part shows incoming connections (from the input neurons). Each input neuron signals the state of one position in the field of view (FOV), and we arrange the connection strengths to those neurons in accordance with those neurons' arrangement in the FOV. We can read, for example, the topmost green block as a positive connection from this neuron to the input neuron that signals the state of the FOV-position 2 steps ahead of the organism. The top part shows the strength of the outgoing connections, after learning, over all 24 possible output-action wirings. Every row shows the result for one wiring (averages over 25 lifetimes). The last column shows the fitness of the individual containing this neuron for the various output-action wirings (brighter is fitter). This example neuron is from an early generation. It can be seen to always make a negative connection to the output neuron associated with the "turn left" action (whichever output neuron that happens to be), a strong connection with the output neuron associated with the "step" action (whichever output neuron that happens to be), and occasionally weak negative connections with the "jump" action (which can be seen to improve fitness).

Consistency in connection pattern over different output-action wirings indicates effective learning. For example, a green vertical bar under "step" in a connection plot expresses that the neuron establishes

a positive connection to the output neuron that triggers the "step" action, regardless of which output neuron that happens to be under the different wirings.

Inconsistency in connection pattern over different output-actions wirings, on the other hand, can indicate either a lack of learning ability, or it can indicate that qualitatively different solutions evolved for different wirings: a neuron might take on different functional roles under different wirings. Neurons taking on different functional roles under different variants of essentially the same task would be a strong indication against cognitive organization. In our experiments this clearly did not happen. In fact we chose to have a large number of different wirings exactly for the purpose of avoiding it: it seemed highly unlikely that the nets would manage to evolve 24 separate solutions, making the evolution of one general "fits all" solution more likely. When the task has only two variants (like in the Nolfi and Parisi study discussed in section 3.3.5.) there is a serious chance of this sort of ad-hoc solutions (see Chalmers (1990) for an exploration of the relation between solution generality and number of environmental variants).

The neuron in the example image is from an individual in a fairly early generation, and the performance level had not stabilized yet. It shows a decent degree of consistency (we see clear green and red lines), but also a fair amount of noise, and variation in fitness over different wirings. The vertical bars in the connection plots get cleaner as evolution progresses. For neurons of individuals without learning, we can make similar plots, except of course that there is only one wiring to show.

To make sense of the wirings that evolved, we should take into account that our activation function is symmetrical on the x-axis. So if we flip the sign of all incoming and outgoing connections of a hidden neuron, the behaviour remains exactly the same (and if we flip the sign on the incoming modulatory connections too, learning should also be unaffected). We should take this symmetry into account when looking at connection plots. The sign-flipped equivalent of a given neuron would have the same plot, except with all reds replaced with greens of the same intensity and vice versa.

Figure 13 shows connection plots for all neurons of the best individual of the last generation of each run. It's a large amount of data and we have no simple objective measure of "cognitive organization" to apply, so interpretation of the data is no easy challenge. Remember that correspondence was taken to be straightforward in *perception* and *action*, and that our hypothesis predicted that the need for learning should help extend correspondence into the *coordination* between perception and action. Action is not represented in our model (there are no motor systems involved), but we can distinguish perception and coordination.

3.7.1. Perception

Both in nets with and without learning we find *feature detectors* among the hidden neurons, that respond to particular salient situations. We show idealized versions of some typical detectors here. In the actual networks, they are noisy and often not complete.

Most runs, both with and without learning, evolve a neuron containing some version of the pattern of incoming connections pictured in Figure 14 (or the sign-flipped equivalent of it). It reacts positively to vertically oriented prey ahead and negatively to horizontally oriented prey ahead.

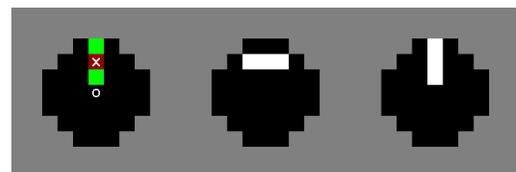
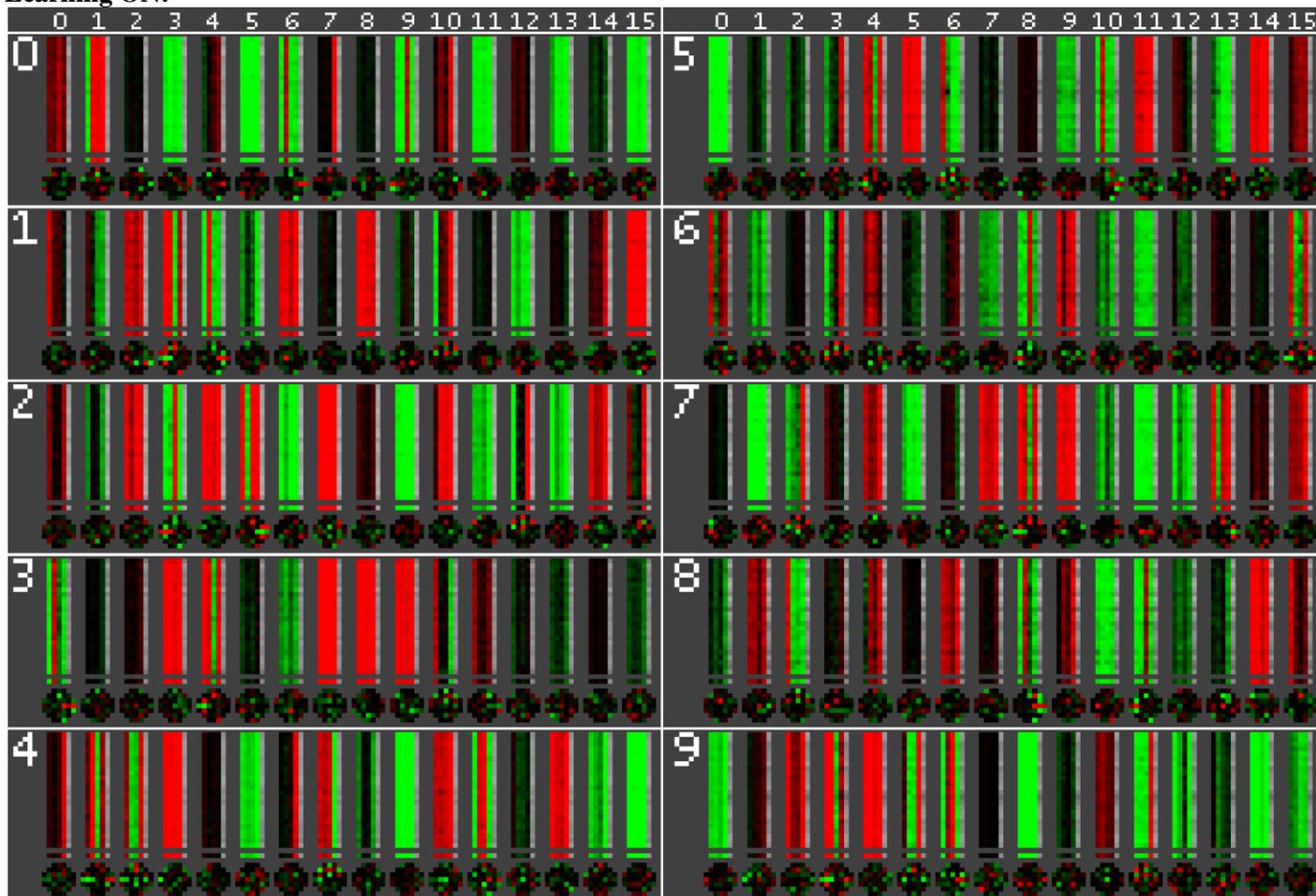


Figure 14. Left: Idealized neuron for detecting orientation of a prey in front of the agent. O indicates the agent's own position. X indicates where a jump would land it. Middle: when there is a prey inside this white area, the red position is occupied, and consequently this hidden neuron becomes negatively activated. Jumping in this situation would lead to fitness penalty. Right: when there is a prey within this white area, both the red position and one of the green positions are occupied. As the green connections are stronger, the net activation of this hidden neuron becomes positive. Jumping in this situation would yield fitness gain.

Learning ON:



Learning OFF:

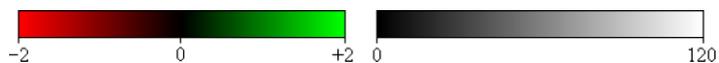
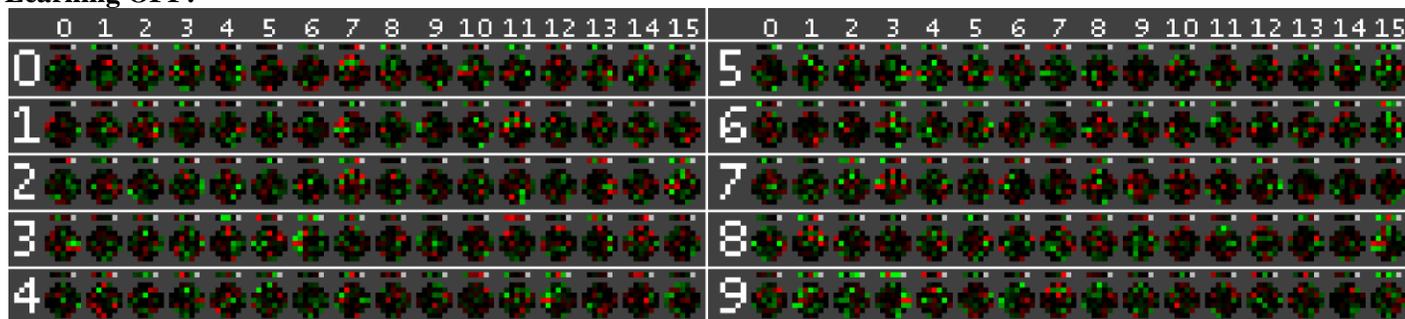


Figure 13. Connectivity of the hidden neurons in the best individual of the final generation of each run for learning-ON and learning-OFF. Each row represents one neural network. Connections to input neurons are ordered after the position in the field of view that the input neuron get their input from. Connections to output neurons are ordered after the action that the output neuron connects to (regardless of output-action wiring), in the following order: step, turn right, turn left, jump. The gray bars/dots show the network's fitness score. See Figure 12 for detailed explanation of the depiction of the individual neurons.

Detectors for horizontally oriented prey to one side are also common (Figure 15):



Figure 15: Left: Idealized neuron for detecting horizontally oriented prey to the right. The cells of a vertically oriented prey to the right (middle image) would cover a red and a green cell, which cancel each other out. A horizontally oriented prey (right image) covers two green cells, leading to strong activation of the neuron. Horizontal prey to the right should be detected, as a rightward turn puts it in a vertical orientation straight ahead, where it can then be eaten by jumping or stepping. As is to be expected, the mirror version of this pattern is common too.

Feature detectors could be seen as perception-level correspondences: they clearly depict those features of the environment that are relevant to the species' fitness gain (relative to the species' action repertoire). Detectors show up both with and without learning, as Spencer would predict.

3.7.2. Coordination

The coordination between perception and action occurs in how feature detectors are combined within hidden neurons, and in the connection patterns from hidden to output neurons.

That there is a substantial difference in how nets with and without learning ability wire their top layers should be evident from looking at Figure 13. In nets with learning, many hidden neurons end up with near-uniform connections to all output neurons. These do not contribute to the choice of action: they increase or decrease the activation received by all output neurons equally. We find far fewer such obviously vestigial neurons in the nets without learning. At first glance this might seem like a flaw in the learning networks. Why do so many neurons remain unused? It's imaginable that the number of functional neurons depends on the number of modulatory neurons, and that there are simply insufficient modulatory neurons to allow all regular neurons to develop learning

function. To rule this out we did a small number of alternative runs with twice as many modulatory neurons (see appendix F). We found that increasing the number of modulatory neurons does not lead to an appreciable increase in functional neurons, and that the functional neurons that evolved show the same perception characteristics as those in runs with 6 modulatory neurons. Thus it would seem that the limited number of functional neurons is evolution's preferred way of implementing the necessary learning ability. This is interesting in the light of our hypothesis. In chapter 1 (section 1.7.) we argued that learning ability would benefit from having a small number of loci to update in a behaviour system, and the experimental result here shows behaviour of networks with learning to concentrate in a limited number of neurons, independent from the number of neurons available. Nets without learning, on the other hand, seem to spread their behaviour out over as many neurons as we equip them with. In this aspect of coordination, the difference between learners and non-learners is clear.

Next we look at how feature detectors are combined within the hidden neurons. A single neuron can take on multiple detection duties (Figure 16):

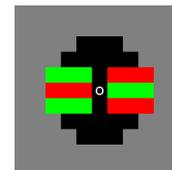


Figure 16. An example compound feature detector. This neuron detects both horizontally oriented prey to the left AND horizontally oriented prey to the right, and responds differentially to those cases by activating positively or negatively (note that horizontally oriented prey on *both sides simultaneously* would leave this neuron uninterested, so in practice a slightly biased neuron would do better).

When we look at Figure 13 in search for such feature detector superimpositions, we find them occurring in both learners and non-learners, and it is hard to make out whether one type is "more organized" than the other. To make out whether a significant difference exists, we used a cluster analysis (CT-clustering, Heyer et al., 1999. See appendix G for details on the algorithm). For each run type (learning ON or OFF) we

pool the hidden neurons of all 10 runs together. Distance between two hidden neurons is computed on basis of their bottom layer connections. If the need for learning ability constrained what coordinations evolved, we should expect the pool of neurons from learners to be more strongly clustered than that of the non-learners, and the resulting clusters might reveal interesting patterns.

Figure 17 plots the number of clusters found by the clustering algorithm as a function of the maximum allowed cluster radius. We see that the algorithm consistently finds a smaller number of clusters for the neuron pool of the learners, indicating that the data in this pool is indeed more strongly clustered.

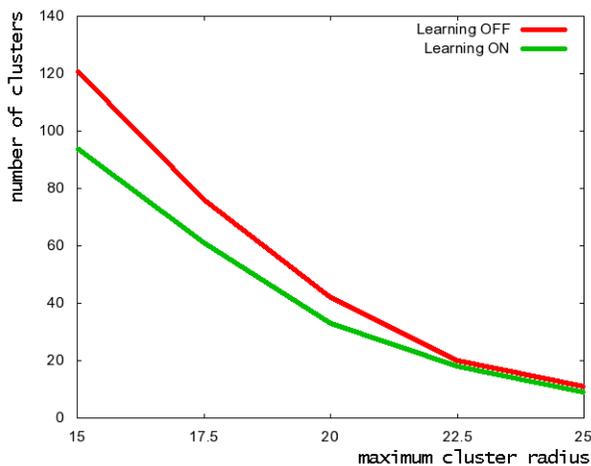


Figure 17. Number of clusters found by CT-clustering for various maximum cluster sizes (see appendix G for details). The datasets consist of 160 neurons each.

In Figure 18, we show the clusters found with a maximum cluster radius setting of 20. It characterizes each cluster with an average of the downward connection patterns of the neurons it contains, and shows the upward connection patterns of all the neurons with each cluster. The clustering algorithm takes sign-flipping into account, and flipped signs where appropriate (see appendix G for details).

Comparing the results for learning and non-learning populations, we note the following:

- Column 0 for both types shows a cluster characterized as complete insensitivity (all-black perception plot). For both types, this is the largest

cluster, but it is larger for learners than for non-learners. This is in line with the idea that in learners, a more focused coordination is preferred, keeping the number of loci that need updating small.

- We recognize pattern detectors in the perception plots of both learners and non-learners. Column 3 in for the non-learners consists of detectors for horizontally oriented prey to the right, for example. However in the learners, we find 3 prominent clusters of compound feature detectors, that are also highly consistent in their upward connections (clusters 2, 3, 4). Compound feature detectors are not absent among the non-learners (compare for example cluster 7 in the non-learners with cluster 4 in the learners), but the tendency to evolve the same compound time and time again does appear weaker in the non-learners. We checked for each member of clusters 2, 3 and 4 in the learners from which runs they came, and found that no two neurons in any of these clusters came from one and the same run. Might this triplet of compound feature detectors constitute a single behaviour system that evolved 7 times over?

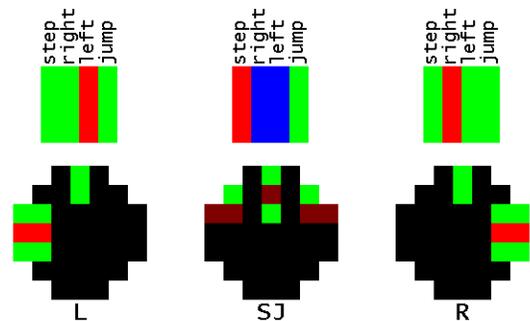


Figure 19. Idealized solution 1 (runs ON-0,1,2,3,5,7 and 8). Neurons L and R detect horizontally oriented prey to the left and right, respectively, and promote the corresponding action. When moving neither right nor left, the SJ neuron assesses whether it is better to step or jump. Connections coloured blue varied over different occurrences of the SJ neuron. They can be either positive, negative or near-zero, but do take the same sign.

We consider a hypothetical network with only three hidden neurons, each with a (cleaned up, and in some cases sign-flipped) connective pattern characterizing one of the three implicated clusters in the learners set. We dub the connection patterns L, R and SJ (see Figure 19). A look at Figure 13 reveals that the majority of the

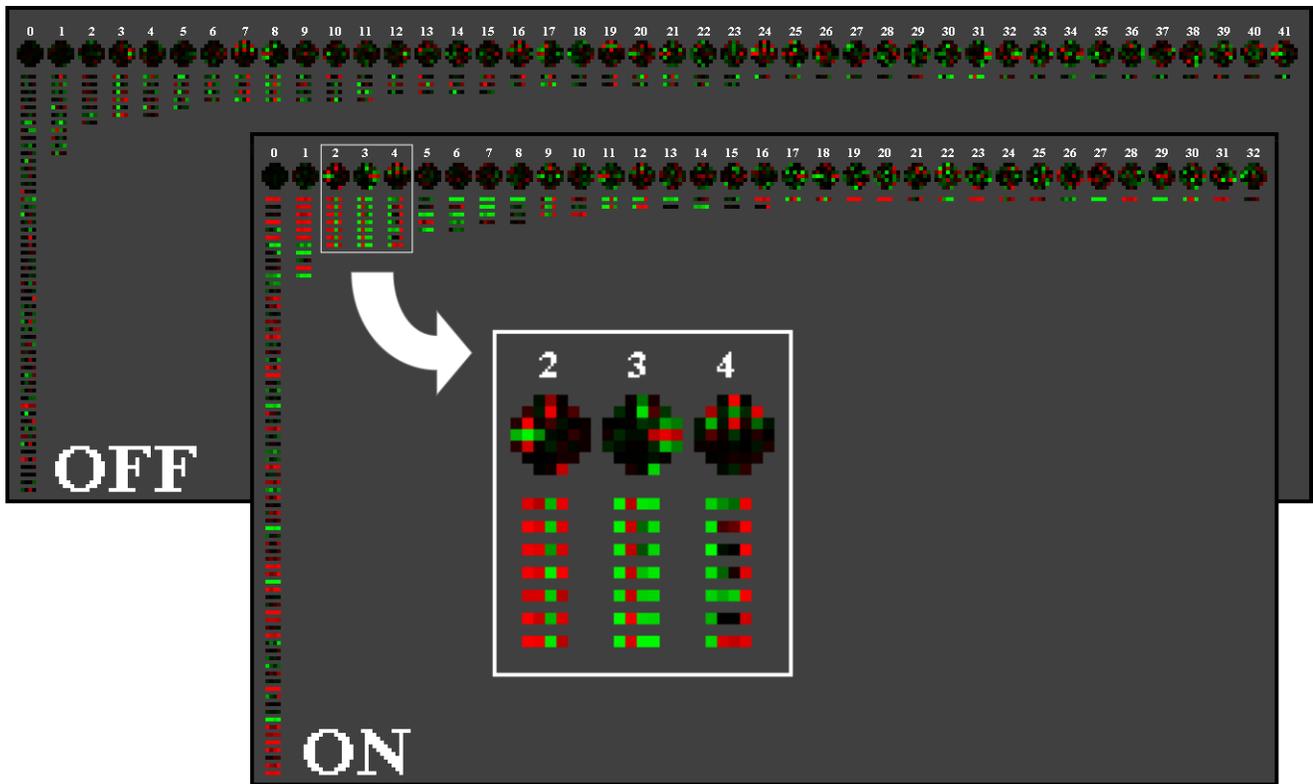


Figure 18. Results of cluster analysis on data collected in the learning-ON and learning-OFF experiments. Each column represents one cluster, headed by a "perception plot", an average of the perception qualities of the member neurons. For each cluster we additionally show the upward connections of all member neurons. The inset highlights the set of patterns we identified as solution 1 (see text).

runs with learning (all except runs 4, 6 and 9, to which we will get later) indeed evolve this triplet of neurons. We call this pattern "solution 1". It is not the only solution, but the fittest and most common.

To get a feel for how solution 1 works, we consider how it responds to a couple of key situations.

Horizontally oriented prey to the right

R-neuron is positively activated, promoting right turn while inhibiting left turn, step and jump.

L-neuron is silent.

SJ-neuron is silent.

Action: right turn.

Vertically oriented prey straight ahead

R-neuron is negatively activated, inhibiting right turn while promoting left turn, step, jump.

L-neuron is negatively activated, inhibiting left turn while promoting right turn, step, jump.

Right turn now receives both positive and negative activation. These cancel each other out.

Left turn now receives both positive and negative activation. These cancel each other out.

Step and jump both remain activated.

SJ-neuron is weakly positively activated, promoting jump while inhibiting step.

Action: jump.

Horizontally oriented prey to the right AND vertically oriented prey straight ahead

R-neuron is both positively and negatively activated, effectively remaining silent.

L-neuron is negatively activated, inhibiting left turn while promoting right turn, step, jump.

SJ-neuron detects that jumping is safe, promoting jump while inhibiting step.

Action: jump.

With a little puzzling we can see why the connections from the SJ neuron to the turn actions are of little consequence. When there is prey straight ahead, both L and R will promote forward motion (both step and jump). SJ will promote either step or jump, and inhibit the other. So either step or jump receives activation from 3 neurons. Even if SJ then promotes turn actions in addition to the forward motion it promotes, activation on turn actions cannot top the total activation of the preferred forward motion. Conversely, if either L or R promotes a turn action, SJ cannot prevent this by inhibiting the turn actions, as the L and R neurons only promote turn actions when there is no prey ahead, in which case SJ will not be activated. So when solution 1 evolves, the connections from SJ to left turn and right turn can drift around randomly.

We could work through more situations, but hopefully the above gives an impression of how solution 1 works, and of the comparatively high degree of organization of its approach (for a neural network, that is).

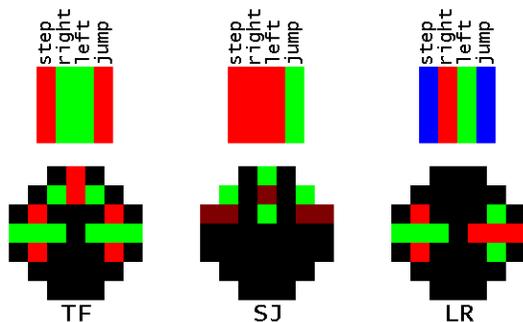


Figure 20. Attempted idealization of solution 2 (runs 4 and 9). The TF neuron decides whether to move forward or to take a turn. When taking a turn, the LR neuron decides to which side. The SJ neuron selects between step and jump as in solution 1, but in both occurrences of solution 2 it had its connections to turn left and turn right equal in sign to its connection to step. Blue connections again vary over occurrences of this solution.

Run ON-4 and run ON-9 evolved different solutions. They order their priorities differently, and are less tidy than solution 1. Figures 18 gives an attempted idealization of this solution and short description of how it appears to work. This leaves ON-6. Fitness in ON-6 was seen to take an inexplicable nose-dive over its last 10000 generations, and it's hard to make out whether it's closer to solution 1 or 2, having aspects of

both. As it had not yet stabilized at the time of run termination, we do not attempt to analyze it further.

The consistency and compactness of the two solutions contrasts with what we find in the non-learners. Although similar compound feature detectors can be found within the non-learners, they are notably noisier and don't appear to come in neat recurring triplets like in the learners. The non-learners' solutions are more varied and more scattered over their neurons.

Solution 1 is notable for being internally *symmetrical*: it handles left and right turns with the same detector patterns, except mirrored: L-neurons mirror R-neurons²². Symmetry is a global feature of the task environment and action repertoire (if turning right in state x produces state y , turning left in the mirror image of state x produces the mirror image of state y). That this symmetry readily finds its way into the learning networks but barely or not at all into the non-learning networks bodes well for our hypothesis. The symmetry of the environment is trivially reflected in the symmetrical perception and action abilities of our species, but only under selection pressure for learning ability do we find this symmetry extending into the coordination in between.

Next we consider how the learners' characteristic patterns of feature detector superimposition serve learning, and how they relate to the structure of the environment.

3.7.3. How feature detector superimposition facilitates learning

In section 1.7. we argued that that learning benefits from conciseness of the behaviour system, and that conciseness goes hand in hand with exploitation of environmental regularities. Do we find environmental regularities back in the solutions we identified in the previous section? And if so, how do they facilitate learning?

²² Solution 2 as we depicted it is similarly symmetrical (the LR-neuron is a sign-flipped mirror image of itself), but we have too few instances of this solution to say for sure that the idealization is accurate.

Consider the “left and right” detector of figure 16. The learning system will want to connect this neuron positively to the output neuron for rightward turn, as this would cause the individual to make rightward turns when there is horizontally oriented prey to its right. But when the learning system makes this connection, not just the response to horizontally oriented prey to the right is affected: when the individual later on finds a horizontally oriented prey to its left, this observations will *inhibit* rightward turns (the prey causes negative activation of the neuron, which travels through the previously established positive connection, lowering the activation of the rightward turn neuron). That turning right is a bad idea when one aims to collect an item to one’s left might seem trivial when phrased like this, but bear in mind that the spatial coherence of the environment is never explicitly given. That this feature of the learning system is remarkable will be clearer when we phrase it from the nets’ point of view.

That action a will favourably transform input vector configuration P implies that action a will NOT favourably transform input vector configuration Q.

This relation clearly does not hold in general. There are plenty of different input vector configurations that are all advantageously transformed by the same action. Of course in a spatially coherent environment, there will be many pairs of input vector configurations for which the above relation *does* hold, but we cannot a-priori identify them without reference to that environment’s spatial coherence.

Similar arguments can be given the other combinations of sensitivities found in the learners (forward & right, forward & left, step & jump, forward & turn). The fact that our learning networks are innately wired to exploit such relations between pairs of input vector configurations indicates “adjustment of internal relations to external relations”, to phrase it in Spencerian terminology.

It’s clear that this can speed up learning. Thanks to the internalizations of these spatial relations, the nets can learn about suitable behaviour for input vector configuration Q from experiencing input vector

configuration P. This gives a tangible example of how a need for learning ability exerts selection pressure on correspondence. It is also clear that non-learners have little use for these contingencies. They facilitate learning, but not innate hard-coded behaviour.

To what extent can we credit learning ability for the emergence of these correspondences? Similar “internal relations” are produced in nets without learning, though there it appears to happen by chance, through random superimposition of feature detectors. As noted before, nothing guarantees a lack of correspondence in non-learners, and there may well be other correspondence-inducing factors in evolution. In the present model, chance superimposition appears to be such a factor. Learning, however, appears to be the more reliable of the two.

3.8. Conclusions chapter 3

In this chapter we took an Artificial Life approach to the evolution of learning. We discussed existing Artificial Life research on the evolution of learning ability, and picked out the elements we needed for a simple model of the evolution of latent learning.

Using a combination of neuromodulation and echo neurons, we let a neural network species evolve the ability to learn to control its behaviour correctly, in an environment that provides no reinforcement or guidance. The species successfully evolved the ability to learn from observing the effects of its own actions.

The need for learning ability was found to exert selection pressure towards specific patterns of feature detector superimposition. This led to focused recognizable control systems within the networks, concentrated in a relatively small number of hidden neurons. It also led to *convergent evolution* over multiple runs: the same systems evolved again and again.

These results show that evolution of latent learning can promote organization in evolving neural networks. That the networks focus their functionality in a small number of neurons is in line with our hypothesis: learning benefits from having to update as few loci as

possible. In an attempt to interpret the control systems found in the learning networks, we observed that superimposition of feature detectors lets a learning update in one situation non-trivially affect the behaviour in other situations, by exploiting the spatial relations of the environment. This interpretation shows that the effect selection pressure on learning has on network organization can indeed be understood in terms of increased alignment of internal to external relations, lending empirical support to our modification of Spencer's theory of the evolution of mind.

Conclusions & Future work

Conclusions

We hypothesized that the evolution of learning ability should have an organizing effect on evolving cognition. The closer the organization of cognition resembles that of the environment, the simpler it is to update it effectively. This hypothesis supplements Spencer's view of the evolution of mind, giving one avenue of explanation for why correspondence should extend beyond perception and action into the coordination between them.

Exploring the varieties of learning as they occur in nature, we found that correspondence starts where behaviourist theory ends. Although many types of learning reveal traces of environmental structure in animal cognition, we found our clearest case in latent learning.

Thus we set out to model the evolution of latent learning in an Artificial Life model. We adopted the neuromodulation architecture, modified it to enable agents to learn from the effects of their own actions, and let a simple virtual species evolve the ability to learn to control their movements without any reliance on reinforcement or example. We contrasted their neural organization with that of a species evolved to solve the same task without learning, and found significant differences. Learners evolved the same concise control system many times over, while non-learners, though not deprived of organization altogether, took a more diffuse approach. This find supports the idea that the evolution of learning can lead to increased neuro-cognitive organization. We then attempted interpretation of the learners' control systems in terms of correspondence between environment and cognition, and found them to contain some simple spatial relations.

Testing our hypothesis required us to devise a novel technique for modelling latent learning ability, as models of latent learning that do not rely on hard-coded model-building mechanisms proved hard to come by. While arguably weak in comparison to

existing reinforcement learning models with hard-coded extensions, the model we built does exhibit more advanced learning ability than other purely evolutionary approaches. By adding simple "echo neurons" to the neuromodulation architecture, we made it possible for our species to learn from observing the effects of their actions. This ability is essential to active modes of learning like practice and play, and might prove useful in further computational study of these modes of learning.

As for the debate within AI about systematicity, our findings support the idea that connectionist models are not doomed to remain too diffuse to be relevant as models of cognition. We identified evolution of learning ability as one factor that can promote organization in neural evolution. Our findings would not convince those that require models of mind to exhibit full-fledged systematicity, but the neural wirings that evolved under selection pressure for learning ability do seem a step closer than those evolved without it. A better understanding of this and other such organizing factors could help in inching artificial neural networks closer to that particular blend of order and chaos that characterizes natural cognition.

As for the Artificial Life project of identifying the interactions between levels of adaptation, the organizing effect of learning on neural evolution looks like a viable candidate for the list. Our results reconfirm that multiple levels of adaptation do not simply produce the same adaptations quicker, but can lead to altogether different patterns of adaptation, and suggest that such altered patterns of adaptation may play a role in the evolution of cognition.

How do these two themes relate to one another? The observed organizing effect crucially depends on the stacking of adaptation processes: organization serves the learning process, and through the effect of learning on fitness, this gives rise to a selection pressure steering the evolution process towards increased organization. If anything, the results indicate

that connectionists should be wary of the simplistic view of evolution and learning as alternative algorithms for the same problems.

Future work

Other organizing factors

We saw that learning ability can promote cognitive organization, but there undoubtedly are more such organizing factors. In section 3.2. (category C) we mentioned that phenotypic plasticity allows for phenotypes that are "too large" for their genotype. Environments can be counted on to provide part of the information necessary for the development of the phenotype, but the genotype should provide the mechanisms for exploiting this environmental information. How does this connection between environment and development fit into the environmental complexity thesis? What effects does a reliance of phenotype development on environmental information have on cognitive structure? These questions merit investigation, and seem quite amendable to computational study.

Action-selection in multi-layer neuromodulation nets

In our model, output neuron activation is set to 0 or 1 depending on whether that neuron's action was performed or not, a simplistic ad-hoc implementation of mutual inhibition between competing motor programs. We don't actually need this for action selection (its primary biological function), but we do need it to let the learning process know which output neuron's action was performed. Our implementation suffices when all neuromodulated connections terminate on an output neuron, but is not easily extendable to networks with multiple layers of neuromodulated connections. This issue is interesting, because learning in such networks might depend on a downward information flow in addition to the standard upward flow. Such necessity could provide a selection pressure that would promote evolution of biologically plausible back-propagation mechanisms. This could be studied by letting network topology evolve in a way that allows for downward connections.

Second order learning

In this thesis we focused on the question of how correspondences establish the mind. We noted that most philosophers focus instead on how the mind establishes correspondences. This distinction had a cameo in section 2.4., when we distinguished the rat's innate spatial framework from its ability to build up new mental representations (cognitive maps) of new test environments. We focused on the former, but does our model in its present form say anything about the latter? Not really. The correspondences in our model constitute, at best, a very primitive innate spatial framework. Our virtual organisms do not go on to establish new correspondences (representations) over their lifetime. However, the model suggests a natural extension that might allow for this. Our individuals are not capable of making new correspondences, *but our populations are*. If we could reproduce this ability of our populations within our individuals, we might be a step closer to modelling representation.

We saw that correspondence emerged from the interaction between evolution and learning, and we argued that it is necessary to have indeed two levels of adaptation for such correspondence to emerge. But nothing about the model demands these levels to be learning and evolution per se. In many regards evolution and learning do the same thing, just at different levels of organization (population and individual, respectively). We already hinted at second order learning in chapter 1. Replacing evolution with a second level of learning ability would put both levels of adaptation, that we saw *can* suffice to produce new correspondences, within the individual. Of course, to shape those individuals, we would need to reintroduce evolution. Could such a model, with three levels of adaptation (evolution, learning, learning), produce individuals capable of establishing new correspondence relations with their environment on the fly? I'm not aware of any attempts whatsoever in this direction. Second order learning does not seem to receive much attention. If two interacting levels of adaptation can establish correspondences, and we think of thought as the active establishing of correspondence relations with the environment, then it should be interesting to see what happens when we

put two levels of adaptation into a virtual species. It is doubtful that this would simply be a matter of plugging second order modulatory neurons into our first order modulatory neurons, devising an environment that exerts the appropriate selection pressures is not trivial either, and of course analyzing the resulting networks would be a major challenge, but an approach of this kind might help in bridging a small part of the gap between the computational and philosophical study of the mind.

References

- Aizawa, K. (1997). Explaining Systematicity. *Mind & Language*, 12: 115-136.
- Arita, T., Yuhji K. (1998). Evolution of Linguistic Diversity in a Simple Communication System. *Artificial Life*, 4(1):109-124.
- Arita, T. (2009). A methodology for ensuring constructive approaches based on artificial life models. *Jinkōchinō Gakkaishi*, 24(2):153-259. (In Japanese)
- Arnold, S.F., Suzuki, R. & Arita, T. (2010). Evolving Learning Ability in Cyclically Dynamic Environments: The structuring Force of Environmental Heterogeneity (abstract). In *Artificial Life XII: Proceedings of the Twelfth International Conference on the Simulation and Synthesis of Living Systems* (forthcoming).
- Axelrod, R. (1987). The evolution of strategies in the iterated prisoner's dilemma. in L. Davis (Ed.) *Genetic algorithms and simulated annealing*. Morgan Kaufmann. pp.32-41.
- Baldwin, J. M. (1896). A new factor in evolution. *American Naturalist*, 30:441-451.
- Batali, J. (1994). Innate biases and critical periods: Combining evolution and learning in the acquisition of syntax. In R. A. Brooks & P. Maes (Eds.) *Artificial life IV*. Cambridge, MA: MIT Press. pp. 160–171.
- Bedau, M. A. (2003). Artificial life: organization, adaptation and complexity from the bottom up. *Trends in Cognitive Sciences*, 7:505-512.
- Belew, R. K., McInerney, J., & Schraudolph, N. N. (1990). Evolving networks: Using the genetic algorithm with connectionist learning. In *Proceedings of the Second Artificial Life Conference*. Addison–Wesley. pp. 511-547.
- Brooks, R.A. (1991) Intelligence without representation, *Artificial Intelligence* 47:139- 159.
- Bullinaria, J.A. (2007). Understanding the Emergence of Modularity in Neural Systems. *Cognitive Science*, 31:673-695.
- Bullinaria, J.A. (2009). The Importance of Neurophysiological Constraints for Modelling the Emergence of Modularity. In D. Heinke & E. Mavritsaki (Eds.) *Computational Modelling in Behavioural Neuroscience: Closing the Gap Between Neurophysiology and Behaviour*, 187-208. Psychology Press.
- Bullinaria, J.A. (2009). Lifetime Learning as a Factor in Life History Evolution. *Artificial Life*, 15:389-409.
- Chalmers, D. J., Touretzky, D. S., Elman, J. L., Hinton, G. E. (1990). The evolution of learning: An experiment in genetic connectionism. *Proceedings of the 1990 Connectionist Models Summer School*. Morgan Kaufmann. pp. 81-90.
- Chomsky, N. (1980). *Rules and representations*. Oxford: Basil Blackwell.
- Coleman, P. D., Flood, D. G., Whitehead, M. C., Emerson, R. C. (1981). Spatial sampling by dendritic trees in visual cortex. *Brain Research*, 214:1-21.
- Crombach A, Hogeweg P. (2008). Evolution of evolvability in gene regulatory networks. *PLoS Computational Biology*, 4(7): e1000112.
- Deacon, T.W. (2003). Multilevel selection in a complex adaptive system: The problem of language origins. In B.H. Weber, D. J. Depew (Eds.) *Evolution and Learning*. MIT Press. pp.81-106.
- Dennet, D. (1994). Artificial Life as Philosophy. *Artificial Life*, 1(3):291-292.
- Dörner, D. Hille, K. (1995). Artificial Souls: Motivated Emotional Robots. *Proceedings of the International Conference on Systems, Man and Cybernetics*, 4:3833-3832.
- Dyer, F.C. (1991). Bees acquire route-based memories but not cognitive maps in a familiar landscape. *Animal Behaviour*, Volume 41(2):239-246.
- Fodor, J., McLaughlin, B. (1990). Connectionism and the Problem of Systematicity: Why Smolensky's Solution Doesn't Work. *Cognition*, 35:183-204.
- Fodor, J., Pylyshyn, Z. (1988). Connectionism and Cognitive Architecture: a Critical Analysis. *Cognition*, 28 (1988): 3-71
- Gallagher, S. (2005). *How the body shapes the mind*, Oxford University Press.
- Garcia, J., Ervin, F.R., Koelling, R.A. (1966). Learning with prolonged delay of reinforcement. *Psychonomic Science*, 5:121-122.

- Garcia, J., Koelling, R.A. (1967). A Comparison of Aversions Induced by X Rays, Toxins, and Drugs in the Rat. *Radiation Research*, 7:439–450.
- Gérard P., Meyer J.-A., Sigaud O. (2005). Combining latent learning with dynamic programming in the modular anticipatory classifier system. *European Journal of Operational Research*, 160 (3):614-637. doi:10.1016/j.ejor.2003.10.004
- Gilbert, S.F. (2003). *Developmental Biology*. 7th edition. Sunderland, MA: Sinauer Associates.
- Godfrey-Smith, P. (1994). "Spencer and Dewey on Life and Mind," in R. Brooks and P. Maes (eds.), *Artificial Life 4*. MIT Press. pp. 80-89.
- Godfrey-Smith, P. (1996). Complexity and the Function of Mind in Nature. Cambridge University Press.
- Godfrey-Smith, P. (2001). Three Kinds of Adaptationism. In S. H. Orzack & E. Sober (eds.) *Adaptationism and Optimality*, Cambridge University Press. pp. 335-357.
- Godfrey-Smith, P. (2002). Environmental Complexity and the Evolution of Cognition. In R. Sternberg, J. Kaufman (Eds.) *The Evolution of Intelligence*. Mahwah: Lawrence Erlbaum. pp. 233-249.
- Gould, J. L. (1982). *Ethology*. New York: W.W. Norton.
- Gould, J. L. and Grant Gould, C. (1994). *The Animal Mind*. Scientific American Library.
- Gould, J. L. and Grant Gould, C. (2007). *Animal Architects: Building and the Evolution of Intelligence*. Basic Books.
- Greenough, W. T., Black, J. E., Wallace, C. S. (1987) Experience and Brain Development. *Child Development*, 58:539-559.
- Hadley, R.F. (1997). Cognition, systematicity and nomic necessity. *Mind & Language*, 12: 137-153.
- Harp, S. A., Samad, T., Guha, A. (1989). Towards the genetic synthesis of neural networks. In J. D. Schaffer (Ed.), *Proceedings of the Third International Conference on Genetic Algorithms and their Applications*. Morgan Kaufmann. pp. 360-369.
- Harvell, D. (1986). The ecology and evolution of inducible defences in a marine bryozoans: cues, costs, and consequences. *American Naturalist* 128: 810-823.
- Hebb, D.O. (1949). *The organization of behavior*. New York: Wiley.
- Heyer, L.J., Kruglyak, S., Yooseph, S. (1999). Exploring expression data: identification and analysis of coexpressed genes. *Genome Research*, 9(11):1106-15.
- Hinton, G. E. and Nowlan, S. J. (1987). How learning can guide evolution. *Complex Systems*, 1:495–502.
- Houk, J. C., Bastianen, C., Fansler, D., Fishbach, A., Fraser, D., Reber, P. J., Roy, S. A., Simo, L. S. (2007). Action selection and refinement in subcortical loops through basal ganglia and cerebellum. *Philosophical Transaction of the Royal Society B: Biological Sciences*, 362:1573-1583. doi:10.1098/rstb.2007.2063.
- Humphries, M.D., Gurney, K., Prescott, T.J. (2007). Is there a brainstem substrate for action selection? *Philosophical Transaction of the Royal Society B: Biological Sciences*, 362:1627-1639. doi: 10.1098/rstb.2007.2057.
- Jacobs, R.A. (1999). Computational studies of the development of functionally specialized neural modules. *Trends in Cognitive Science*, 3, 31-38.
- Kashtan, N & Alon, U. (2005). Spontaneous evolution of modularity and network motifs. *Proceedings of the National Academy of Sciences*, 102:13773-13778.
- Kato, T., Arita, T. (2004). A Robotic Approach to Emotion from a Selectionist Perspective. *Proceedings of the 9th International Symposium on Artificial Life and Robotics*. pp. 601-604.
- Kim, K.Y., Cho, D.Y., Zhang, B.T. (2006). Multi-stage evolutionary algorithms for efficient identification of gene regulatory networks. *Applications of evolutionary computing, Lecture Notes in Computer Science*, 3907:45-56. doi:10.1007/11732242_5.
- Langton, C., G. (1986). Studying artificial life with cellular automata. *Physica D*, 22:1120-1149.
- MacCorquodale, K. (1970). On Chomsky's review of Skinner's *Verbal Behavior*. *Journal of the Experimental Analysis of Behavior*, 13(1): 83–99. doi:10.1901/jeab.1970.13-83.
- MacLennan, B. (1991). Synthetic ethology: An approach to the study of communication. In C. G. Langton, C. Taylor, J. D. Farmer, S. Rasmussen (Eds.) *Artificial life II*. Reading, MA: Addison-Wesley. pp. 631–658.

- Matthews, R. J. (1997). Can connectionists explain systematicity? *Mind & Language*, 12: 154-177.
- Maynard-Smith, J. (1987). When learning guides evolution. *Nature*, 329: 761-762.
- McGinn, C. (1993). *Problems in Philosophy: The Limits of Inquiry*. Blackwell.
- McLaughlin, B. P. (2009). Systematicity Redux. *Synthese*, 170: 251-274.
- Merelo, J. J., Patón, M., Cañas, A., Prieto, A., Morán, F. (1993). Optimization of a competitive learning neural network by genetic algorithms. In *Proceedings of International Workshop on Artificial Neural Networks*. Springer-Verlag. *Lecture Notes in Computer Science*, 686:185-192.
- Migita, M., Moriyama, T. (2005) Modeling Ratiomorphic Property in Animal Behavior: Computer Simulations of Cognitive Aspect of Turn Alternation Response in Pill Bug. *Cognitive Studies*, 12(3):207-220. (In Japanese)
- Nätt, D., Lindqvist, N., Stranneheim, H., Lundeberg, J., Torjesen, P. A., Jensen, P. (2009). Inheritance of Acquired Behaviour Adaptations and Brain Gene Expression in Chickens. *PLoS One*, 4(7): e6405.
- Nolfi, S., Parisi, D. Elman, J. L. (1994). Learning and Evolution in Neural Networks. *Adaptive Behavior*, 3(1):5-28. doi:10.1177/105971239400300102.
- Nolfi, S., Parisi, D. (1996). Learning to adapt to changing environments in evolving neural networks. *Adaptive Behavior*, 5(1):75-97. doi:10.1177/105971239600500104.
- Paenke, I., Kawecki, T., Sendhoff, B. (2006). On the influence of lifetime learning on selection pressure. *Proceedings of Artificial Life X*. MIT Press. pp. 500-506.
- Paenke, I., Kawecki, T., Sendhoff, B. (2009). The influence of learning on evolution: A mathematical framework, *Artificial Life*, 15:227-245.
- Pan, W.X., Schmidt, R., Wickens, J. R., Hyland, B.I. (2005). Dopamine Cells Respond to Predicted Events during Classical Conditioning: Evidence for Eligibility Traces in the Reward-Learning Network. *The Journal of Neuroscience*, 25(26): 6235-6242. doi:10.1523/JNEUROSCI.1478-05.2005.
- Robinson, E. & Bullinaria, J.A. (2009). Neuroevolution of Auto-Teaching Architectures. In J. Mayor, N. Ruh & K. Plunkett, *Connectionist Models of Behavior and Cognition II*. Singapore: World Scientific . pp. 361-372.
- Rubchinsky, L. L., Kopell, N., Sigvardt, K. A. (2003). Modeling facilitation and inhibition of competing motor programs in basal ganglia subthalamic nucleus–pallidal circuits. *Proceedings of the National Academy of Sciences of the United States of America*, 100(24):14427-14432. doi:10.1073/pnas.2036283100.
- Schaeffer, A. A. (1911). Habit formation in frogs. *Journal of Animal Behavior*, 1(5):309-335. doi:10.1037/h0070752.
- Sejnowski, T., Rosenberg, C. (1986). NETtalk: A parallel network that learns to read aloud. *John Hopkins University Electrical Engineering and Computer Science Technical Report*, January 1986.
- Shettleworth, S.J. (1972). Constraints on learning. In Lehrman, D. S., Hinde, R. A., Shaw, E. (Eds.) *Advances in the study of behavior Volume 4*. Academic Press.
- Skinner, B. F. (1957). *Verbal Behaviour*. New York: Appleton-Century-Crofts.
- Small, W.S. (1901). Experimental study of the mental process of the rat. II. *The American Journal of Psychology*, 12(2):206-139.
- Soltoggio, A., D'urr, P., Mattiussi, C., and Floreano, D. (2007). Evolving Neuromodulatory Topologies for Reinforcement Learning-like Problems. In *Proceedings of the IEEE Congress on Evolutionary Computation 2007*.
- Soltoggio, A., Bullinaria, J.A., Mattiussi, C., Dürr, P. & Floreano, D. (2008). Evolutionary Advantages of Neuromodulated Plasticity in Dynamic, Reward-based Scenarios. *Proceedings of Artificial Life XI*. MIT Press. pp. 569-576.
- Spencer, H. 1885. *The Principles of Psychology*. 3rd edition. New York: Appleton.
- Sugiura, K., Arita, T. (2010). "Why We Talk?: Altruism and Multilevel Selection in the Origin of Language", *Artificial Life and Robotics*, 15 (in press).

- Suzuki, R., Arita, T. (2004). Interactions between Learning and Evolution: The Outstanding Strategy Generated by the Baldwin Effect. *Biosystems*, 77:57-71.
- Suzuki, R. and Arita, T. (2007). Repeated occurrences of the Baldwin effect can guide evolution on rugged fitness landscapes. In *Proceedings of the First IEEE Symposium on Artificial Life*. IEEE Press. pp. 8–14.
- Suzuki, R., Arita, T. (2008). How Learning Can Guide Evolution of Communication. *Artificial Life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems*. MIT Press. pp. 608-615.
- Takano, M., Kato, M., Arita, T. (2005). A Constructive Approach to the Evolution of the Recursion Level in the Theory of Mind. *Cognitive Studies*, 12(3):221-233.
- Teitelbaum, P. (1966). The use of operant methods in the assessment and control of motivational states. In W.K. Honig (Ed.) *Operant Behavior: Areas of Research and Application*. New York: Meredith. pp. 565-608.
- Tesauro, G. (1995). Temporal Difference Learning and TD-gammon. *Communications of the ACM* 38 (3).
- Tieman, S. B., Hirsch, H. (1982). Exposure to lines of only one orientation modifies dendritic morphology of cells in the visual cortex of the cat. *Journal of Comparative Neurology*, 211:353-362.
- Todd, P. M., Miller, G. F. (1991). Exploring adaptive agency II: simulating the evolution of associative learning. *Proceedings of the first international conference on simulation of adaptive behavior*. pp.306-315.
- Tolman, E. C., Honzik, C. H. (1930). 'Insight' in rats. *University of California Publications in Psychology*, 4:215-232.
- Tolman, E. C. (1932). *Purposive behavior in animals and men*. New York: Appleton-Century-Crofts.
- Tolman, E. C. (1948). Cognitive maps in rats and men *The Psychological Review*, 55(4): 189-208. doi:10.1037/h0061626
- Turney, P., Whitley, D., Anderson, R. W. (1996). Evolution, Learning, and Instinct: 100 Years of the Baldwin Effect. *Evolutionary Computation*, 4(3):4-8.
- Valverde, F. (1971). Rate and extent of recovery from dark-rearing in the mouse. *Brain Research*, 33:1-11.
- Voicu, H., Schmajuk, N. (2002). Latent learning, shortcuts and detours: a computational model. *Behavioural Processes*, 59:67-86. doi:10.1016/S0376-6357(02)00060-8.
- Watanabe, Y., Suzuki, R., Arita, T. (2008). Language Evolution and the Baldwin Effect. *Artificial Life and Robotics*, 12:65-69.
- Whitley, D., Starkweather, T., Bogart, C. (1990). Genetic algorithms and neural networks: optimizing connections and connectivity. *Parallel Computing*, 14:347-361.
- Whitley, D. (1995). Genetic algorithms and neural networks. In J. Periaux, G. Winter (Eds.) *Genetic Algorithms in Engineering and Computer Science*. Wiley. pp. 203-216.
- Wiles, J., Watson, J., Tonkes, B., Deacon, T. (2005). Transient phenomena in learning and evolution: Genetic assimilation and genetic redistribution. *Artificial Life*, 11:177–188.
- Yamauchi, H. (2007). How does niche construction reverse the Baldwin effect? *Proceedings of the Ninth European Conference on Artificial Life*. Springer. pp. 315–324.

APPENDICES

APPENDIX A: Genetic algorithm

Each run starts with a population of 90 networks, each with all its connection weights initialized at 0. An individual's fitness is the number of correctly approached prey minus the number of incorrectly approached prey.

Selection was rank-based, with replacement and elitism. After assigning each individual its rank (from 0 to 89), we pick 40 individuals using the following formula:

$$i = \text{floor} \left[\frac{n \cdot (r - 1)}{r \cdot (1 - s^{-2}) - 1} \right]$$

Where n is the population size (90), r is a random number from $[0,1]$ and s is a parameter controlling selection pressure (0.4). This formula allows for easy manipulation of selection pressure independent of population size. The next generation is composed of two mutated copies of the selected individuals plus unaltered copies of the 10 best individuals of the preceding generation.

Mutation is single-point. Every individual to be mutated is first assigned a random mutation-sensitivity value from the range $[0,0.0075]$. This value is used as the probability of mutation for each individual connection within the individual. The higher this value, the larger the expected number of connection mutations. Early evolution is best suited with large numbers of simultaneous connection-mutations while late evolution ("fine-tuning") is best suited with few mutations, and the use of a mutation-sensitivity value guarantees that both dramatic and subtle mutations occur. Mutation on weights that are not exposed to learning has a 0.75 probability of adding a random number in the range $[-0.5,+0.5]$ to the connection weight and a probability of 0.25 of resetting the connection weight to zero. Weight values are clipped to the range $[-2,+2]$. For weights that are exposed to learning, mutation strength and clipping range are divided by 100, as trial runs indicated that these

connections need very weak but non-zero innate weights, which are hard to evolve with the mutation settings that work well for the other connections.

APPENDIX B: Reasoning behind model details

It might seem like a strange choice to complicate the model with positive and negative. As discussed in the text, the present setup guarantees that the "value" of a 1-cell is zero (both in the training and performance environment, stepping on a 1-cell yield an average reward of zero, just as 0-cells do). This allows us to create a neutral training environment without adding more properties to the cells. Had we always positively rewarded contact with a prey (in an otherwise identical setup), the species could have devised pseudo-reward-based learning: if they would evolve to recognize "standing on a 1-cell" (something that never happens in the performance environment: prey disappears the moment it's touched) as a reward state in the training environment, they could use it to anchor an ability to learn to approach 1-cells, which would produce positive fitness in the performance environment. Towards the end of the training period, we would see the organisms approaching 1-cells, and simply keep on doing so in the performance environment, ignoring the difference between the two. The actual setup blocks this solution. It could also have been blocked by assigning more than one property ("colours") to each cell: a property of containing toy and a property of containing prey. Then too the performance environment would provide new information. However, it would double the number of input neurons required, making the already computationally expensive experiments slower still.

Another aspect of the model that may raise an eyebrow is why we use direction of approach to determine whether a prey yields positive or negative reward. It might seem like we might as well have used distinct "food" and "poison" items, like in many other models. The reason is mainly technical. Given binary cells, differently oriented pairs of 1-cells are the simplest way of representing two different objects, but

as the species can make 90 degree turns, objects' orientation in the visual field is variable and cannot be used to distinguish objects, unless object type changes along with orientation. Thus in our model, "food" (vertically oriented prey) turns into "poison" (horizontally oriented prey) with every 90 degree turn of the agent.

Additionally, having every object be edible potentially increases diversity of the situations encountered by a single individual. When items get eaten, they reappear elsewhere in the grid, resulting in a continuous reconfiguring of the layout of the field. A poison item would ideally never be eaten, causing them to stay in place throughout the lifetime. An unfortunate starting distribution would thus have a bigger impact on fitness scores, though it can be doubted that this effect would be significant.

APPENDIX C: Tasking non-learners with learning

In section 3.6. we compared populations with and without learning ability. However, the tasks of the two are not identical: the learners get random output-action wirings, while the non-learners all get the same constant wiring. We explained that a comparison using the same task would not be informative: without the random wiring, evolution can hard-code optimal behaviour genetically, so there is no need to evolve learning ability at all. On the other hand, without learning, there is no way to deal with the randomness of the output-action wiring, and adaptive behaviour is impossible. Here we show what happens when a non-learning population is left to evolve under random output-action wirings (Figure 21).

Contrary to expectation, average fitness does not stay put at zero, but goes up to about 5. Compared to non-learners with a static output-action mapping or learners with a random mapping, this performance is dire, so our point stands, but still we wondered how it is possible that positive average fitness was attained at all: in absence of learning, the random mapping guarantees that agents have no idea what they are doing. We think the positive average fitness can be explained as follows: agents might detect situations where, say, jumping would lead to a fitness penalty,

and connect this perception so as to inhibit three out of their four outputs. This would lead to three out of every four agents not jumping when jumping is disadvantageous (and one in every four leaping at every opportunity to get hurt), leading to a slight elevation in average fitness.

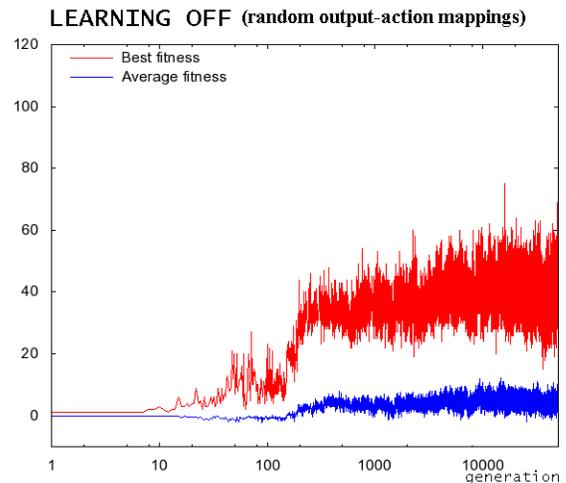


Figure 21. Evolution process of a non-learning population with randomization of output-action mapping enabled. At around generation 200, the ability to perform random behaviour evolves, which leads to variation in fitness. Generation best fitness goes up to around 40 and generation worst fitness dives to well below zero (not shown), while average fitness stays at around 5 (against about 87 for non-learners with static mapping, and about 65 for learning with random mapping).

APPENDIX D: The necessity of echo neurons

We added echo neurons to the neuromodulation framework to enable agents to learn from observing the effects of their own actions. Here we verify that these echo neurons are indeed necessary. Figure 22 shows what happens when we remove them. This experiment is otherwise identical to the main experiment. We found that performance stabilizes at the same level as in non-learners with randomized output-action mappings (see appendix C). This indicates that without echo neurons, the modulatory neurons are useless in this task.

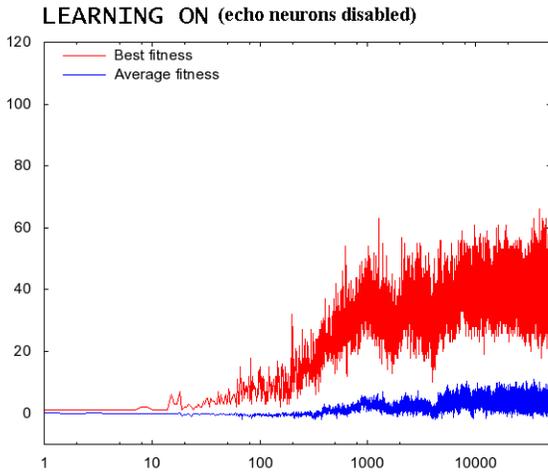


Figure 22. Evolution process of a learning population with echo neurons disabled. Between generations 100 and 1000, generation best fitness rises to scores of around 40, generation worst fitness scores remain well below zero (not shown), and average fitness scores of around 5.

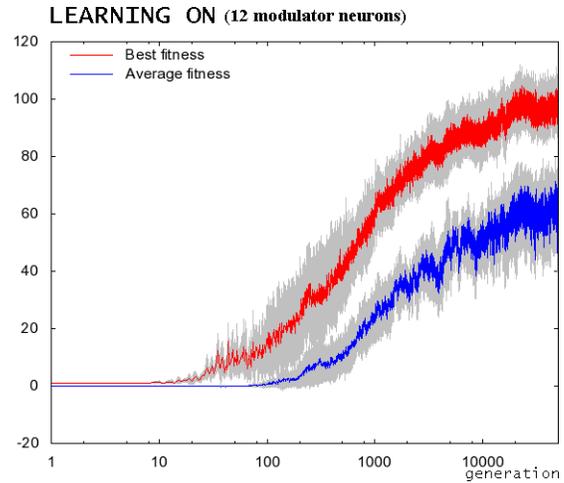


Figure 23. Evolution process of nets with 12 modulatory neurons instead of 6. Averages over 5 runs. X-axis in logscale. Gray area indicates standard deviation. Fitness scores are similar to those of nets with 6 modulatory neurons: best fitness is around 97 and average fitness at 62 after 50000 generations, versus 101 and 65 in the nets with 6 modulatory neurons.

APPENDIX E: Connectivity analysis

The images in Figure 13, learning ON, were created as follows: Each of the depicted networks lived 25 lifetimes for each of the 24 output-action wirings (having its connection weights reset to the innate configuration at the start of each lifetime). Each row in the column depicting upward connections shows the average connection strengths learning produced under that row's wiring. Fitness scores also are averages of each set of 25 lifetimes. Averages over all wirings are computed over all $25 \times 24 = 600$ lifetimes in the analysis. In Learning-OFF, connection weights don't change over individuals' lifetimes, so we simply plot the innate connection weights. For assessment of fitness in learning-OFF networks however, we average over the same number of lifetimes (600) as in learning-ON.

APPENDIX F: Non-effect of more modulatory neurons

To rule out the possibility that the compactness of the learners' control systems are due to a shortage of modulatory neurons, we ran the simulation 5 times with twice as many modulatory neurons. If the number of modulators forced compact control systems, these runs should show more functionally involved hidden

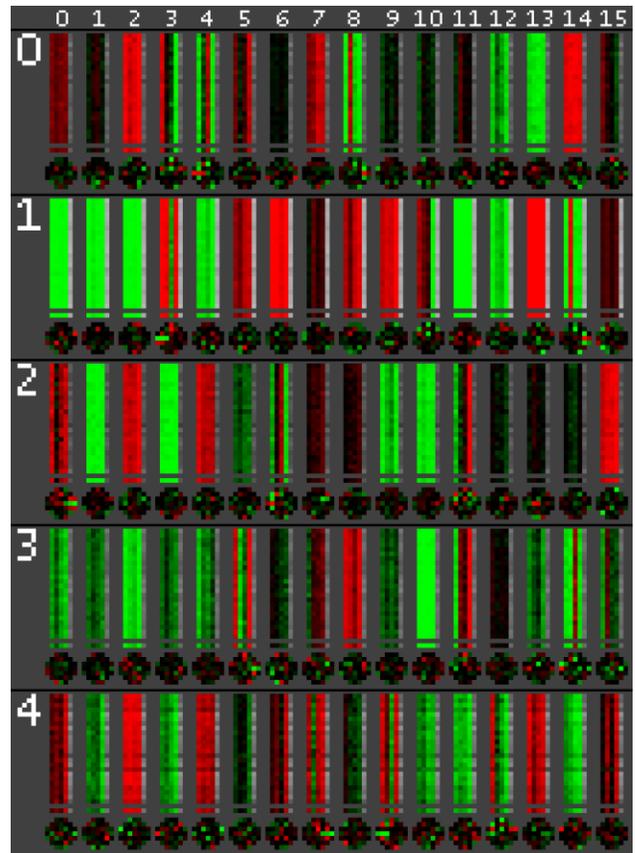


Figure 24. Best networks of last generations of 5 runs with 12 modulatory neurons per network. Runs 0, 1, 4 and arguably 2 evolved solution 1 (see section 3.7.2.), run 3 presents an unclear case.

neurons. Figure 23 shows the evolution process and Figure 24 the resulting network structures. Neither fitness nor network structure was found to differ substantially from the runs with 6 modulatory neurons.

APPENDIX G: Clustering algorithm

Clustering analysis was done using CT-clustering (Heyer et al., 1999). This algorithm works as follows: Every point (neuron) in the dataset roots a candidate cluster. Candidate clusters are built by iteratively adding the neuron closest to the cluster to the cluster (where distance is measured as the maximum of the distances to all neurons already in the candidate cluster), until no neuron can be added without its distance to the cluster exceeding a user-defined threshold (the maximum cluster radius value shown on the horizontal axis in figure 17). Then the largest candidate cluster is selected and all its neurons are removed from the dataset. Then the procedure begins anew with the remaining neurons, until all neurons are assigned to a cluster.

We had to make a small adjustment in the clustering algorithm to account for the sign-symmetry of hidden neurons. As discussed in the text, flipping the sign on all incoming and outgoing connections of any given hidden neuron has no effect on a network whatsoever, however as data-points a neuron and its sign flipped counterpart can lie far apart. This would lead to functionally similar neurons not getting grouped correctly by the clustering algorithm. This should be prevented, so in computing the distance from a cluster to a neuron we compute both the distance from the actual neuron and the distance to its sign-flipped equivalent, and add to the cluster whichever of the two neurons is nearest. This also enables computation of sensible average downward connection patterns as depicted in Figure 18 (functionally similar neurons of opposed signature would cancel each other out in computing the average pattern).

We computed neuron distances on basis of bottom layer connections only. There are two reasons for this choice. First, as described in appendix A, top layer connections in learners had smaller weight ranges and weaker mutation. This would skew the results of a clustering based on the full set of innate connection weights. Second, had we clustered on basis of the full set of connection weights *after* learning, then the criticism could be levelled that we are comparing the products of different algorithms (genetic versus genetic plus neuromodulation), and that this is bound to cause differences in clustering. As the bottom layer is handled identically for both learners and non-learners, comparing the bottom layer connections gives us the clearest view of the effect that the selection pressure for learning ability had on evolving neural organization. For completion's sake, Figure 25 shows the result of clustering on basis of top-layer connections (innate for non-learners, after learning for learners). We see that here too, learning-ON shows a higher degree of clustering.

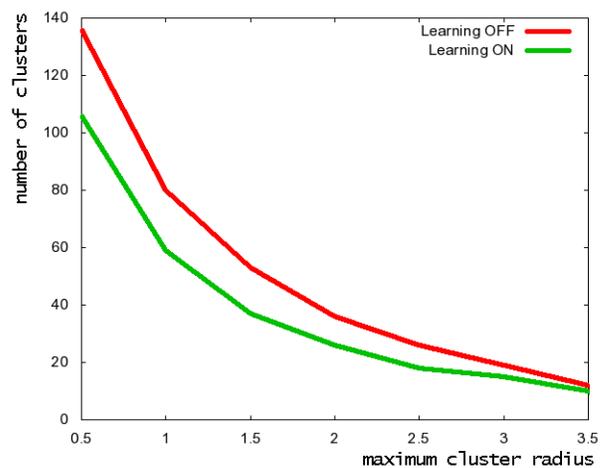


Figure 25. Number of clusters found by CT-clustering for various maximum cluster sizes when clustering on basis of top layer connections (after learning in case of learning-ON).

