Getting emotional with evolutionary simulations: the origin of affective processing in artificial neural networks
Heerebout, B.T.

Citation for published version (APA):
Abstract
In human evolutionary history, working memory has probably developed to temporarily maintain information about the environment when the corresponding input is no longer perceived. We investigated whether a working memory capacity could develop in evolutionary simulations, with agents collecting food while avoiding predators. The artificial neural networks, controlling the agents, contained two oscillation generators that could synchronize during evolution. We expected that the oscillations in the two generators would decouple, when the stimulus would disappear from the environment. After 20,000 generations of evolution, there was only limited synchronization between the two oscillations, and more so in response to predator than to food stimuli. With predators, the oscillation in the second recurrent layer synchronized with the first. Because the second layer lagged one time step behind, it probably reduced distraction by food. After stimulus removal, mutual excitation between nodes ensured high output activation levels. Subsequently networks would behave as if detecting a predator, regardless of the previously presented stimulus. In a small group of agents the structures that maintained activity were, however, influenced by stimulus specific properties. These findings suggest that those positive feedback loops may be the ‘spandrel’ that was later co-opted for the use of a working memory for specific stimuli.
4.1 Synchronous Oscillations in Evolutionary Simulations

When we perceive the world around us and close our eyes for a moment, we are not surprised to find it is still there after we have opened our eyes again. Yet, for very simple organisms this constancy is not so obvious. Imagine a creature that only has the ability to respond directly to its sensory input. For such a being, the discontinuation of its input would seem as if the world had vanished. Somewhere in our ancestral lineage, the capacity developed to retain information temporarily about our environment, although it may no longer be perceived directly. Computational simulations can serve to investigate the evolutionary development of the underlying neural mechanisms. We extended the artificial neural networks of previous simulations (Heerebout & Phaf, 2010ab) with an additional recurrent layer. The recurrent connections developed, as expected, with a limited capacity to maintain a signal, although a full working memory function did not arise. Evidence is presented however, that the networks’ recurrent structures, which initially function to enable fast switching to new input, may eventually be co-opted for assimilation in a working memory configuration.

In previous evolutionary simulations, recurrent connections evolved to function as negative feedback loops (Heerebout & Phaf, 2010a). These loops caused the activation of the nodes in the networks to oscillate. Because the nodes were thus regularly deactivated, competition in the network could be reset within short intervals. Hence, the oscillations allowed the networks’ competition to switch winners fast in response to new input. Additional simulations confirmed the relation between oscillations and behavioral flexibility. The networks of foraging agents exhibited faster oscillations than those of agents that were fleeing from predators. This makes sense from an evolutionary perspective because a foraging agent should be able to respond fast to a sudden threat and a fleeing agent should not be distracted by food on its path. The oscillatory mechanism and its function provided a novel interpretation suggested by the simulations.

The agents’ inheritable attributes coded for the configuration of their network and did not code directly for specific behaviors (i.e., no ‘genes-for-behavior’ type modeling). Addressing the neutral neural level in the simulations allows for the evolution of unexpected neural mechanisms. The oscillations were not implemented by design, but were an emerging property of the combination of the initial conditions and
selection pressures in the evolutionary process. Because all behaviors follow from the same underlying neural structures, the networks will produce behaviors that are per balance beneficial for the agent. Nevertheless, specific isolated behaviors, produced by the same network may be detrimental. Another advantage of addressing the neural level in the simulations is its capacity to reuse of these functional structures for novel adaptations. Gould and Lewontin (1979) coined the term ‘spandrel’ for a property that is a by-product of an adaptation, rather than an adaptation by itself. The term originally refers to the space between an arch and a rectangular enclosure, often filled decoratively in renaissance architecture. Analogous to the evolutionary process, the space is an architectural necessity and is then co-opted to hold decorations. The space was not intentionally built in to hold the decorations. In case of the simulations, it may be that the recurrent loops in the networks first evolved to facilitate behavioral flexibility. Of these loops, some caused mutual excitation between nodes. They may be co-opted to maintain activation in neural structures that gave rise to a working memory function.

4.2 Synchronizations in Working Memory

A key aspect of working memory is its capacity to briefly hold information, even without the corresponding input. According to Wolters and Raffone (2008) this capacity is derived from recurrent connections within the pre-frontal cortex as well as with posterior parts of the brain. Through these recurrent connections, cell assemblies can coherently hold the information available with synchronized firing patterns. They emphasized that the amount of concurrently retained information is limited by how well the cell assemblies can segregate the synchronized patterns and not the size of the assemblies.

Interestingly, long-range recurrent connections from the frontal eye field, an area within the prefrontal cortex, to the visual cortex have been found to bring about synchrony between the oscillations in those areas (Gregoriou, Gotts, Zhou & Desimone, 2009). When a stimulus was attended, the neurons in these areas that responded to the stimulus, showed enhanced oscillatory couplings, particularly at the gamma range. The time lag due to the synaptic delays suggested that the coupling was initiated by the frontal eye field. Attention and working memory are closely linked and mutually influence each other. Selective attention may decide what is maintained in working memory and the contents of working memory can in turn direct attention (Awh &
Spawning Spandrels for Working Memory

Jonides, 2001; Downing, 2000). The process in the latter direction may well be taken care of by the synchronization of visual cortex with the pre-frontal cortex, in the manner that that Wolters and Raffone envisaged.

**Figure 1.** Schematic view of the network architecture with two levels of recurrent connections. Each layer consists of four nodes, except for the output layer which has two nodes projecting to the agent’s left and right motor actuator. Initially the network has zero-weight connections. The connections, depicted by the arrows between layers, indicate the potential weights that can develop between nodes of these layers.

The aim of the current study is twofold. First, to examine whether stimulus specific synchronizations could evolve in networks with two separate oscillation generators. Second, to examine whether structures would emerge that kept activation after stimulus disappearance. Previous simulations (Heerebout & Phaf, 2010ab) could not develop synchronization between the oscillatory activations, because these networks possessed only one level of recurrent connections. The initial network architecture in
the current study was therefore extended with the opportunity for an additional recurrent loop, so that oscillations could be generated by two, initially independent, generators (see Figure 1). The additional oscillation generator is, moreover, not directly addressed by the input and constitutes a parallel pathway to the output, which may enhance the chances of developing oscillation maintenance of stimulus termination.

4.3 Simulation methods
The simulation setup employed for these studies was, with exception of the initial network architecture, identical to the design of den Dulk, Heerebout, and Phaf (2003). Here, we will describe the method only briefly and refer to the appendices for further detail.

The agents in the simulations had to collect food and avoid predators. To assess their performance each agent was put in the virtual environment individually, together with six predators and ten plants. All initial locations were random. The predators would always move in the direction of the agent. The agent could smell the plants and the predators. The behavior of the agents, the mapping of the smells at different intensities to activation levels of its motor actuators, depended on the configuration of the agent’s artificial neural network. The number of nodes was constant. The weight of the connections between the nodes developed during the evolutionary process.

Initially, all connections were near zero. Agents that performed better on the task gained higher odds for producing offspring (non-randomly; see Appendix A for details). To create offspring, two parent agents were selected. Their inheritable properties, specifying the connections of the networks, were copied, recombined and slightly mutated. The reproductive process ensured slight random variations, so that configurations that helped performance would be developed further. The selection eliminated connection sets that impaired performance. This process of producing a new generation, testing it, and selecting individuals that remained and reproduced, was repeated 20,000 times per simulation.

An agent’s movement was determined by the level of activation of its motor actuators, located on each side of its round body. When both were active, the agent moved forward. The trajectory would be curved if the force exerted was not equal. The motor actuators consumed the agent’s energy (see Appendix B for details). When an
agent collided with a plant, energy was gained. If this happened, the plant disappeared and another plant would instantly appear at a new random location. An agent’s performance on the task, i.e. its fitness, was measured as the product of the agent’s energy level and the number of time cycles before the agent eventually collided with a predator. Because the entities’ initial random location influenced the result, an agent’s performance was measured as the average over twelve tests.

Agents that performed better on the task not only gained higher odds for reproduction, but also for lasting another generation (see Appendix C for details). The reproductive process ensured slight random variations, so that configurations that helped performance would be developed further. The selection eliminated connection sets that impaired performance. This process of producing a new generation, testing it, and selecting individuals that remained and reproduced, was repeated 20,000 times per simulation.

The networks basic architecture (i.e. the nodes and the connection weights) consisted of an input layer, an output layer and two hidden layers in between, which do not directly receive input or produce output. The input layer (receiving input from the agent’s sensors) projected directly to the output layer (feeding into the agent’s motor actuators) and to the first hidden layer (see Appendix D for details on the node activation functions). The first layer also projected directly to the output layer and to the second hidden layer. The second hidden layer only projected to the output layer. In addition, both hidden layers were associated with their own ‘recurrent’ layer, which was exclusively and recursively connected to its hidden layer. With negative feedback from positively excited nodes in these recurrent layers, the hidden layers can generate oscillatory activation patterns though a kind of ‘flip-flop’ mechanism (see Heerebout & Phaf, 2010ab).

For optimization purposes, the connections in the network followed a left-right symmetry. This means that that an agent’s left response would be identical to the agents right response if its input would be mirrored exactly. For the evolutionary process the assignment of an exact mirror connection for each connection in the network meant that the number of dimensions along which ‘solutions’ were considered, was halved. This design choice greatly enhanced the odds of discovering adaptive networks within a feasible timeframe.
To investigate the behavioral consequences and environmental relevance of the potential double oscillation generating structures, ten evolutionary simulations were performed. The simulations were repeated (with identical initial parameters) because the evolutionary development of the networks can get stuck in a local optimum. In that case all mutations seemingly lead to worse performance but nevertheless a ‘better’ weight configuration exists that cannot be reached. Moreover, even though the networks’ basic architecture was fairly simple, the recurrent connections of the two hidden layers and their interactions could lead to runaway processes and chaotic behavior. Due to this complexity, the ‘solutions’ produced by the evolutionary process may, in principal, differ substantially per simulation. The small mutations of the weights per generation and the rigorous deletion of the ‘unfit’, however, reduced the variation within populations. In the simulation results, there never were agents with radically deviating networks.

To analyze the emerged solutions, the oscillatory properties of the networks of all the agents of each last generation (169 networks in total) were determined. For the first aim of this study, the level of synchronization between the separate oscillation generators was determined, and whether the synchronizations were stimulus specific. To this end, the nodes activations were recorded with a fixed input pattern of a stimulus (either a plant or a predator) at 45 degrees left of the agent at a distance of ten length units between the rim of the agent and the stimulus. The fixed input caused the networks to develop steady oscillation patterns. After ten time steps, the stimulus would disappear and the input would be set to zero. For the second aim of the study, it was then investigated whether the network could hold activation after stimulus termination. Moreover, the symmetry of the nodes’ activations was analyzed. Because the networks’ connections were symmetrical with respect to the left and right side, the activations could only be asymmetrical with asymmetrical input. However, symmetrical input (which also applies for zero input) could only correspond to asymmetrical activations when representations of previously detected spatially asymmetrical stimuli were still maintained. Asymmetry could thus provide important insights into a network’s capacity to temporarily hold information.
4.4 Simultaneous oscillations in the two generators

As in the previous simulations, all agents successfully evolved to approach food and avoid predators. The fitness development over the 20,000 generations varied in tempo, but all simulations showed an asymptotic curve that reached its maximum around 16,000. The average fitness of the last 50 generations of each simulation (mean = 16,199; SD = 3776.0) showed no particular outliers (more than 2.5 SD’s). First the average oscillatory properties were determined per network (of all layers except the input layer) in response to the fixed stimulus input. Consistent with our previous findings (Heerebout & Phaf, 2010b) analysis of the last generation revealed that the networks oscillated more often in response to a plant (n = 149) than to a predator (n = 117). Additionally, the frequency of the oscillations, measured as the number of cycles per time step (ranging from 0 to 0.5) over all agents was significantly higher (F(1, 168) = 13.535, p < 0.001, ηp² = 0.07), in response to a plant (f = 0.17, SD = 0.10 time step-1) than to a predator (f = 0.13, SD = 0.10 time step-1).

To examine whether stimulus specific synchronizations occurred between the two separate recursive structures their oscillation frequencies were compared. A correlation between the oscillation frequencies provides a first crude measure of synchronization. If the oscillations that occurred simultaneously in these layers would have had the same frequency, they would correlate perfectly. However, the results show that although the frequencies of the first recurrent layer and the second recurrent layer correlate positively, both with food (r = 0.08) and with predators (r = 0.12), these correlations did not reach significance, neither with plants (p = 0.31) nor with predators (p = 0.13). Thus, it seems that agents with oscillatory couplings did not have a definite advantage over their rivals. To further examine the role of synchronization between the recurrent layers, we measured the time shift for the subgroup of agents that did have both recurrent layers oscillating in the same frequency.

The signal of a stimulus reaches the second recursive layer one time step after it reached the first recursive layer, but their output is propagated to the output nodes simultaneously. Therefore, if both layers oscillate with the same frequency and phase, the synchronization cannot depend on the input and must be the result of an evolutionary predisposition. Alternatively, if the second recursive layer lags one time step behind, the synchrony depends on the input being processed over excitatory connections between the layers.
It was found that 108 agents had both recursive layers oscillating with the same frequency in response to plants and 42 agents in response to predators. On average, the second recursive layer lagged 1.22 ($SD = 0.38$) time steps behind with plants and 1.04 ($SD = 0.37$) time steps with predators. A planned one-tailed t-test revealed that the phase shift between the layers differed significantly from zero, both with plants ($t(107) = 33.7, p < 0.0001$) and with predators ($t(41) = 18.2, p < 0.0001$). It appears that with plants the phase shift was not caused by the excitatory connections, because it also differed significantly from one time step ($t(107) = 5.96, p < 0.0001$). There may have been no synchronization because with plants the corresponding frequencies were coincidental.

With predators the phase shift did not differ significantly from one time step ($t(41) = 0.747, p = 0.459$). The oscillations in the second recurrent layer were, thus, directly coupled to the oscillations in the first recurrent layer. After synchronization, the second layer responded similarly to the environment as the first, but lagged one time step behind. The one-step phase shift may well have been advantageous in evolution. Because input was processed later, the second recurrent layer remained unaffected to distracting stimuli longer. When the first-layer oscillation is in a through, the second-layer oscillation is still near a top, and therefore less susceptible to switches. The phase difference thus leads to more steadfast behavior of the agent when it is faced with a predator. Hence, the coupled synchronization was beneficial when a predator was detected and this was probably because of the resulting reduced distractibility.

The advantage of the synchronization of the two recurrent structures, with a lag of one time step, is also confirmed in the comparison of the average fitness values. The fitness of the agents of which the recurrent layers oscillate synchronized in response to predators ($n = 42$; average fitness = 24,879; $SD = 21,825$ energy × time step) reached significantly higher levels of fitness [$F(1,167) = 16.36, p < 0.0001$] than the non-synchronized agents ($n = 127$; average fitness = 13,686; $SD = 12,844$ energy × time step). With food the comparison of the synchronized agents ($n = 108$; average fitness = 16,471; $SD = 16,830$ energy × time step) with the non-synchronized agents ($n = 61$; average fitness = 15,282 energy × time step) showed no such effect [$F(1,167) < 0.0001, p = 1.00$].

To investigate the possible evolution of a working memory function, network dynamics before and after stimulus removal were examined. Out of the ten simulations,
the agents from eight simulations exhibited steady oscillations after all input values were set to zero. Differentiation between stimulus types turned out that there was a strong correlation of the oscillation frequencies during the predator stimulus with those after the stimulus removal.

The first hidden layer ($r = 0.22, p < 0.005$) and the second hidden layer ($r = 0.30, p < 0.0001$) showed such a positive correlation. For the plant stimulus however, these correlations were not significant. Moreover, in contrast to the frequencies during the stimulus presentation, the post-food oscillation frequencies ($f = 0.08, SD = 0.10$ time step$^{-1}$) and the post-predator oscillation frequencies stimulus ($f = 0.08, SD = 0.09$ time step$^{-1}$) no longer differed significantly. Both after the plant and predator presentation the agent remained in a low-frequency ‘predator-mode’. In our particular simulation setup this makes sense because a predator still could be in the vicinity although its exact whereabouts were not detected at the time. Conversely, after a plant disappears, it is not likely to reappear in the same spot, because the agent ‘ate’ it. In that case working memory for its position (i.e., the angle from the agent) would serve no purpose.

The activations of a typical network after the stimulus was removed, is shown in Figure 2. When presented with a constant input resembling a predator at 45 degrees left of the agent, the nodes in the first hidden layer (Hidden layer 1) and its recurrent counterpart (Recurrent layer 1) were activated at constant level (only showing minute oscillatory variations). The nodes of the second recursive structure (Hidden layer 2 with Recursive layer 2) oscillated at a stable frequency of 0.17 cycles per time step. For the output layer, this resulted in high activation for the output node projecting to the left output motor actuator and fluctuating activation of the output right node. The activation of the left output node was always higher, and so the agent would have moved to the right, away from the predator. After T8, when the input was set to zero, the input nodes lost all activation. After the next time steps oscillatory activity remained in both recursive structures producing similar activations in the output nodes. Thus, until the agent would receive new input from a specific direction, it always continued to make pulsed movements in a straight line, away from where it last detected the predator.
Figure 2. An example of the activations of a network’s nodes over time. The activations are ordered per layer and all range between 0 and 1. The nodes in the input layer were active until T8. The network responded to a predator stimulus which was removed after T7. The output nodes were activated by both hidden layers. The nodes of both the first and second hidden layer had mutual excitatory connections with their recursive counterparts and continued to produce oscillatory output, even after stimulus withdrawal.

The oscillatory activations in the zero-input networks apparently could code for the predator stimulus. The recurrent connections not only consisted of negative feedback loops causing the oscillations, but also of positive feedback loops that sustained the nodes driving the oscillations. A simple model of the mechanism that was responsible for these network characteristics is depicted in Figure 3. The idealized connection scheme was abstracted from the actual networks by generalizing over similarities and deleting non-effective connections and nodes that were never activated. The node labeled $N_1$ receives external activation from an input node. The nodes $N_1$ and $N_2$ have mutual excitatory connections and $N_2$ implements a “flip-flop” mechanism with $N_3$, because after $N_3$ is activated, it inhibits $N_2$. Identical mechanisms were described in Heerebout and Phaf (2010b) and interestingly, similar structures have also been described by Ritz and Sejnowski (1997) and Duprett, Pleydell-Bouverie, & Csicvari
Spawning Spandrels for Working Memory

(2008). After input is received the mutual excitation between N1 and N2 ensures sustained activation. This continues cycle can be reset by the modulating role of the inhibitory connection from N3. In these simulations the continued activations did not do much more than merely driving the agent on in a particular mode. A ‘working memory’ function, with nodes keeping a representation of a specifically detected stimulus, i.e., its angle relative to the agent, cannot be assigned to those networks.

\[ \text{Figure 3. The network’s substructure containing both mutual excitation between N1 and N2 (solid, arrowhead lines) and inhibition by N3 (dashed, globule-head line). The positive feedback loop sustains post stimulus activation, which can be reset because of the oscillatory modulation resulting from the negative feedback.} \]

All the networks’ connections were symmetrical with respect to the agents’ left and right. Activations can therefore only be asymmetrical when the input is asymmetrical (e.g., a plant is close by on the left resulting in greater sensory readings on the agent’s left side and greater output values on the agent’s right side). Asymmetric activations must thus stem from asymmetric input. Once all input had ceased, a ‘memory’ activation of such an asymmetric stimulus (45 degrees left of the agent) should still maintain asymmetric activations. Such networks would have had the ability to keep a turn towards a stimulus, or away from it, after the removal. A comparison of
the agents’ left and right sided post-stimulus activations revealed that these networks were rare. Yet, the evolutionary algorithm was principally capable of producing that type of configured connections. Out of the 169 agents analyzed, fourteen agents were discovered with networks that kept asymmetric activations after removal of the plant stimulus and six agents after removal of the predator stimulus. The resulting output difference moved the agents towards the direction where the plant had been and away from the location where the predator had been, but was probably too small to substantially enhance the agents’ fitness (the nodes activations ranged between 0 and 1; the average difference in activation was 0.09 with both stimuli). Further analyses revealed that the output difference was exclusively caused by the recurrent loops to the second hidden layer. The small number of ‘working memory’ networks does not permit strong conclusions. However, the results strengthen the conclusion that the ‘memory’ based activations were more likely to emerge in the second hidden layer than in the first, due to its larger distance from the input and its parallel contribution to the output.

4.5 A spandrel for residual post stimulus activation

In sum, the basic feature of switching enhancing oscillations evolved in all networks. The two oscillation generators exhibited coupled oscillations in response to a predator. The lag of the second recurrent structure relative to the first probably lowered the influence of potential food distractions when the agent was trying to escape a threat. After termination of the predator stimulus, networks could retain their oscillatory activity, particularly in the second recurrent layer. The sense of direction, however, was generally lost. Nevertheless, a few agents had developed networks in which the stimulus caused a directed movement, even after the stimulus was no longer detected. But because the output difference was so small the directed movement was probably not substantive. Although the underlying feedback loops were thus most likely inconsequential for the agents’ reproductive success, the structures show similarities to the recursive loops that Wolters and Raffone (2008) have proposed to underlie our working memory. The oscillatory structures may thus not have evolved directly for the purpose of a working-memory function, but more likely improved flexibility with fitness-enhancing stimuli, and attentional locking with fitness-reducing stimuli. The maintenance of these oscillations after input has ceased, however, does not seem to play a role in attentional flexibility. The results from the simulations suggest that the positive
Spawning Spandrels for Working Memory

feedback loops enabling such maintenance may have been a neural ‘spandrel’, which were later co-opted for the use of a working memory for specific stimuli. It may even have been an exaptation.

The environment, as created for these simulations, was not conducive for the evolution of a working memory capacity in the networks. In light of these results one could hypothesize that also in our own evolutionary history the recurrent connections in our brains were first installed because of their switching advantages. If so, they would be a precursor to working memory and likely have formed a spandrel necessary for the reuse in a mechanism with a working memory function. However, based on the results from the simulations, it is likely that additional elements from the environment, like more detailed (i.e., than only angle and distance from the agent) location-specific properties and the division of the environment into sub-domains, also formed a necessary precondition.

Further studies with simulated evolutions might reveal what the necessary ingredient, missing from our current setup, could be. The environment in the present simulations has no features that can be kept active in working memory. Pilot simulations have already been performed where the position was also coded by the agent, but recurrent networks could not yet be tested in this new setup. It is expected that, when this is done the position of the predator may be kept active in such a working memory, even when the agent has successfully fled it. The internal representation of the predator position could then serve as an avoidable stimulus in the future, particularly when the movement of the predator is restricted to a specific quarter of the virtual world. Eventually, the emergence of a working memory capacity would show that even high-level “cognitive” functions have clear evolutionary roots and that these processes can be analyzed with the help of evolutionary computation.