Spontaneous modular NeuroEvolution arising from a life/dinner paradox

Kevin Godin-Dubois, Sylvain Cussat-Blanc and Yves Duthen

University of Toulouse, IRIT - CNRS UMR 5505, 2 rue du Doyen Gabriel Marty, 31042 Toulouse, France
{kevin.dubois, sylvain.cussat-blanc, yves.duthen}@irit.fr

Abstract

Understanding the neurological implementation of emotions is a major research subject from biology to computer sciences that, in the latter case, takes many shapes: from accurate detection of human emotions to the emulation of plausible responses to stimuli. There is, however, room for a more bottom-up approach in which we would thrive to recreate emotions from undifferentiated elementary building blocks.

In this article, we used virtual creatures that interact with their environment through a low-level perception/cognition/action loop to demonstrate their potential for fear responses. Embedded in a physical environment in a typical prey/predator setting, they develop strategies for foraging while minimizing their exposure to danger.

By monitoring the neural activities of these subjects, we were able to highlight the regularities induced by an ES-HyperNEAT encoding and their eventual mapping into "mental states". We further emphasize the potential of this approach by clustering these ANNs and showing their resulting complexity in terms of conspecific identification, communication, and functional modularity. Indeed, through functional equivalence across numerous topologies, we identify a fear-related neural cluster that serves as a primitive defensive survival circuit.

Context

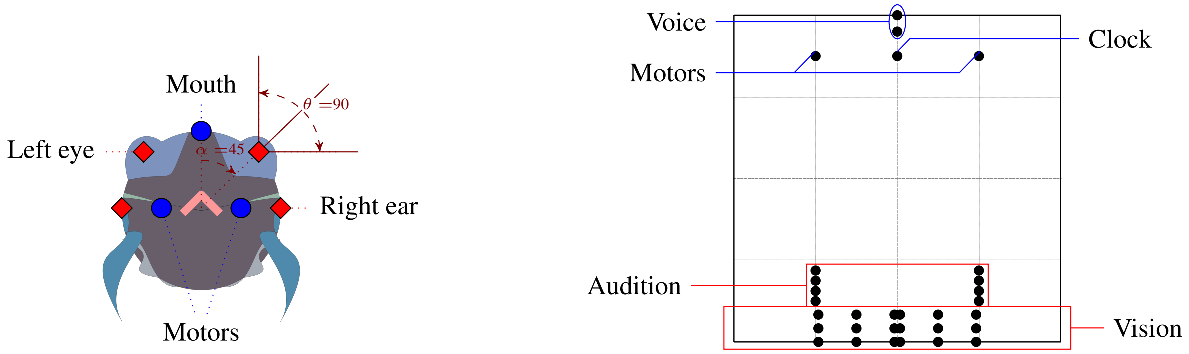
Emotions provide contextual information to better guide individuals towards favorable regions of their behavioral space. On this basis, fear is expected to result from predatory defense systems (Ohman, 1986) intended to produce incentives in fight or flight situations. Indeed based on fear "being caused by particular patterns of threat-related stimuli, and in turn causing particular patterns of adaptive behaviors to avoid or cope with that threat"(Adolphs, 2013), it can be seen as an evolutionary advantage, increasing an individual's chances of survival. Especially in the case of the early mammals, which were exposed to extreme selective pressure by a saurian-dominated environment, the ability to respond quickly and efficiently to a perceived threat is of paramount importance to ensure continued survival. As summarized in (Treccani, 2020), it has been argued that this pressure persists nowadays most notably with the "snake detection theory" (Van Le et al., 2013) which emphasizes the influence of snakes on the evolution of the primate brain.

Combinations of neurological and cognitive studies have shown that, in the human brain, fear relies on numerous interconnected neural circuits (Ledoux and Brown, 2017). While conscious fear results from their (partial) interactions, each such circuit encodes parts of that feeling: the memory of fearful events, unpleasant visual cues, or an individual's flight/fight tendency.

Indeed, thanks to extensive optimizations in complex environments, the animal brain is a marvel of modularity and distributed computing. Researchers from the field of Artificial Intelligence have used such findings to promote similar levels of complexity either in terms of discrimination, to provide emotionally aware virtual agents (Hendy and Farag, 2013; Bălan et al., 2019), or as a control mechanism that would improve upon an agent's plausibility. Classical experiments in the latter setting involve a predetermined ordering of multiple modules in generally fixed topologies. These modules aim to emulate specific regions of the human brain with varying levels of complexity (Armony et al., 1997; Delgado-Mata et al., 2007; Lotfi and Akbarzadeh-T., 2014). Furthermore, while these instances cover the connectivist side of the argument, there is a similar investment resting on the symbolist side, e.g. (de Freitas et al., 2007).

However, despite arguments in favor of bridging the gap between biological brains and Artificial Neural Networks in terms of shared characteristics instead of dividing capabilities (Treccani, 2020), biologically-oriented studies of ANNs are under-represented in the literature. This is even more marked in the context of communication between grounded virtual agents, for which emotions could serve as a facilitating means to improve joint fitness (Hesp et al., 2020).

The objective of this paper is thus to address the issue of emergent brain structures in an artificial substrate in two ways: a) does the use of unconstrained topology precludes the production of modular neural clusters? b) if structures do arise, how could they be described in relationship with observed trends in biological brains?



(a) Anatomy of a creature: inputs (red) and outputs (blue)

(b) Neural substrate before hidden nodes instantiation

Figure 1: Morphological and neurological components of an agent. The location-dependent inputs/outputs are mapped onto a Cartesian plane with coordinates in $[-1, 1]^2$ for use with ES-HyperNEAT

Model

The creatures used in this experiment are elementary in their composition: a circular body, a set of physical inputs (vision/audition), and a handful of actuators both for locomotion (motors, clock speed) and interaction with neighbors (voice). Owing to the straightforward nature of the experiment, the metabolic and morphological aspects of these creatures were drastically limited. In the former case, they are endowed with a single monotonically decreasing energy reserve, filled at the beginning of each evaluation and a metabolic clock in the range $[c_0, c_1]$. In the latter case, the body plan is fixed to the central circular body of figure 1a with a genetically controlled color and none of the additional artifacts. As a result, the only variable part of their morphology is the position of eyes, controlled via several angles (main angle α, relative angle β set to 0 in the illustration and total width θ) and a precision parameter p.

The value of p is used to generate $2p+1$ rays on both sides of the creature's front, as per the illustration, each producing an RGB triplet. As such, these individuals "see" in a similar manner to humans, albeit in only one dimension and with a much coarser resolution. Still, as argued in (Treccani, 2020), the process of trial and error experienced by such creatures can produce useful insights into associated dynamics of the biological substrate.

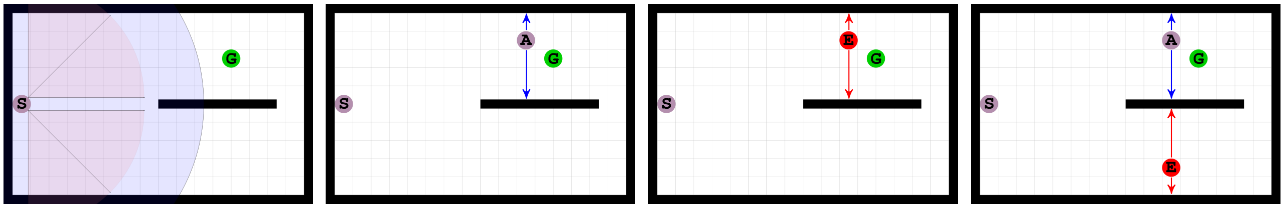
Auditory signals are another way through which these individuals can perceive their environment, although in this case only the biotic component is concerned. Each ear receives inputs from the surrounding area on two types of channels: one is dedicated to "noise", i.e. with an unspecified emitter, while the other three correspond to vocal channels, i.e. voluntary communication. Due to a linear attenuation effect, sounds are better perceived at close range resulting in both left and right ears receiving slightly different values for each stimulus, except in the case where the source is collinear with the creature's orientation. Although this lacks the refinement of extent echolocation, the variation in input levels can be leveraged to determine the relative position of the emitter and derive appropriate responses.

On the effector side of the creatures, three types of outputs are used: motion, metabolic and vocal. The first type is implemented as a pair of "motors" able to produce an effective output in the range $[-1, 1]$. In effect, this results in individuals behaving in a tank-like manner through forward/backward motion, on-the-spot rotation, and all alternatives in-between. The sole metabolic output is used to control the "clock speed" c as defined by $c(x) = x\ c_1$ if $x \leq 0$ and $c(x) = -x\ c_0$ otherwise. This clock speed is then used as a multiplicative factor of both passive energy consumption and motor strength, thereby equating low values with a quiescent state and higher values with sprint behavior. As the range $[c_0, c_1]$ is variable through genetic mutation, cumulative selection can favor specific directions both in terms of amplitude and bias towards one end. Finally, the vocal effectors are decomposed between the channel of emission and the output volume thereby allow arbitrarily complex messages to be produced.

The connection between inputs and outputs is carried out by an Artificial Neural Network based on the ES-HyperNEAT algorithm (Risi and Stanley, 2012). The use of a CPPN as the genomic basis for the neural controller has been shown to be a very compact and efficient model for geometric inputs/outputs. Furthermore, the Evolvable Substrate extension alleviates the need to manually define the hidden neurons, relying only on the connectivity pattern to convert high-density regions of information into processing nodes. This makes it possible to have structures emerge as a result of evolutionary constraints which is a more biomimetic approach to NeuroEvolution than fixed topology networks. In this specific experiment, all nodes use a custom soft-signed activation function[1]. Additionally, the CPPN has a maximal connection weight of 3, uses a link expression output, and also codes for a per-neuron bias in $[-1, 1]$.

---

[1] The code for this experiment is publicly available on github



(a) 0+ (b) 1+ (c) 2+ (d) 3+

Figure 2: Left-hand side scenarios. The subject (S) is placed on the left-most border (in black) of a closed arena with a single food source (G, in green) and a partially walled right-most section. The eventual ally (A) and predator (E, in red) are placed near the goal and perform an endless patrol (bidirectional arrows). The visual and auditory ranges are displayed in the first case.

Evolutionary protocol

Given the objective of studying the emergence of fear on an unbiased substrate, we devised a straightforward experiment that pitted efficient foraging against avoidance strategies. Through a classical evolutionary algorithm, individuals were evaluated on 8 scenarios, each emphasizing different skill-sets. Figure 2 summarizes the left half of these scenarios while the other half is obtained by axial symmetry along the x-axis (e.g. in 0- the food is on the lower side).

In all cases, the subject (S) is placed on the left-most side of the arena with its visual and auditory ranges (displayed, in figure 2a, as disk slices and a larger disk, respectively) not reaching quite far enough to instantaneously detect either the food, other individuals, or central obstacle. In both confrontation cases (1, 2) the other agent (A or E) is placed near the food source (the goal G), leading to two possible outcomes: either S boldly races for G, at the risk of encountering E, or it goes the long way around through the small opening on the opposite side. In the last case (3), S is faced with a simple choice: going for the ally or the enemy, given that the goal is *always* on the former side.

The ally is based on the subject's genotype with no modification: it possesses the same color and neural controller. However, to provide dynamical conditions, its motion is overridden so that it performs a linear patrol: if colliding with an obstacle (wall, food, creature) the clone performs a 180° rotation, otherwise, it moves forward at its maximal speed. The rest of its neural outputs are left untouched, especially its vocal organs allowing communication both spontaneously and in response to auditory cues (e.g. produced by the subject). The predator is a slightly different case as its genome is fixed for all evaluations with forward-facing overlapping arcs of vision. In addition to the patrolling and obstacle avoidance procedures, it also uses visual and auditory data (in the same form as that obtained by regular individuals) to adjust its trajectory so that it can detect and ultimately catch poorly performing individuals. To limit its threat level, the maximum speed it can reach is capped to 75% of a creature's base speed.

Thus each type of scenario favors complementary aspects of the skill-set: foraging (0-3), conspecific recognition (1,3), threat detection/avoidance (2,3). The life/dinner paradox arises from the scoring methodology defined as follow:

$$score(e,d) = \begin{cases} -e & \text{if touched by the predator,} \\ 1-d & \text{if not touching the goal,} \\ 9e+1 & \text{if successful.} \end{cases} \quad (1)$$

where e is the normalized remaining energy (time limiting factor) and d is the distance to the goal. This produces a cost/reward matrix, in the sense of evolutionary game theory, where individuals have a defection strategy (ignoring the goal), a positive outcome (reaching the goal, however late), and a highly penalized negative outcome (being captured by the enemy). While providing evolutionary gradients along which behavioral optimization can occur, this piecewise evaluation favors adaptive solutions to the foraging problem: retrieving the food while evading danger.

To promote robust behavior, evolutionary fitness is the raw sum of an individual's score in each scenario. This allows for some error margin with initial generations typically only finding viable strategies for one specific type of scenario. Cumulative selection however can capitalize on isolated strategies to favor more all-round individuals which show robust demeanor across all categories. This aggregated score is used in conjunction with a novelty metric based on the final positions across all evaluation types thus producing a vector in $\mathbb{R}^{16}$. Indeed preliminary experiments showed that premature convergence on sub-optimal strategies (e.g. only exploring the left side) was occurring at dangerously high frequencies.

In terms of evolutionary parameters, we performed 50 replicates with varying seeds for the initial populations. Each run consisted of at least[2] 1000 generations with a population size of 200 and no crossover. Four elites were conserved across generations and selection was based on random tournaments of size 4.

---

[2] Additional generations (up to 8 in practice) were allowed for populations solely composed of empty ANNs, i.e. those for which ES-HyperNEAT did not find valid hidden nodes

Type	Periods					
	1	2	3	4	5	6
I	∅	A	∅	A	∅	A
II	∅	E	∅	E	∅	E
III	∅	E	A	E	A	E

Table 1: Distribution of presented stimuli in each type of evaluation. A: Ally, E: Enemy, ∅: None

Modularization

While the exploration of neural structures both in-vitro and in-silico has been the subject of extensive studies, these rely on the result of millions of years of evolution, in the specific conditions of biological life. Indeed either through analysis of the neural pathways of the animal brain (Ledoux, 1998), the use of identified key cerebral regions to produce plausible artificial behavior (de Freitas et al., 2007; Delgado-Mata et al., 2007; Lotfi and Akbarzadeh-T., 2014) or mathematical approaches (Broekens et al., 2015), all such methodologies rely on the initial bias of our evolutionary history. As previously argued, we postulate that the emergence of structures in an arbitrary substrate is not limited to the variations around our biological example. Indeed, especially with such highly regular ANNs, we can expect different regions of our artificial brains to specialize in different types of stimuli.

To accurately detect such regions, we exposed each of our champions to "canonical" situations. In every instance, the subject is pinned in the center of an empty, wall-less arena for 10 seconds (100 simulation steps). Due to the total lack of external stimuli, most networks remain quiescent with next to no neural activation. Following this period, the subject is faced with one type of stimulus (ally or enemy depending on the evaluation) for an additional 10 seconds. This brutal switch from peaceful to potentially hostile conditions induces a change in regime in most networks, resulting in an observably different state.

The pattern of presented stimuli in each three types[3] of evaluation scenarios is summarized in table 1. Thanks to these reproducible conditions, we were able to tag timestamps with either the presence or absence of a particular type of input allowing precise monitoring of the activation state of each of a given individual's hidden neurons. Using this information, every neuron's time series is subjected to a one-sided Mann-Whitney test to determine whether it was significantly more active in response to specific stimuli[4]. This procedure is repeated across all three evaluation scenarios, thereby coupling each neuron with a tag specifying to which (if any) stimulus it responds to. Thus, starting from an arbitrarily dense ANN as depicted at the top of figure 3 one obtains three different patterns of differential activation across the network (middle row). Each tag is then combined to produce an aggregated version of the network where modules abstract a subset of its topology while maintaining its functional partitioning. The aggregation procedure is performed through an intersection operator, meaning that for a neuron to belong into the enemy (red) module, for instance, it has to be found significantly activated in both types II and III evaluations.

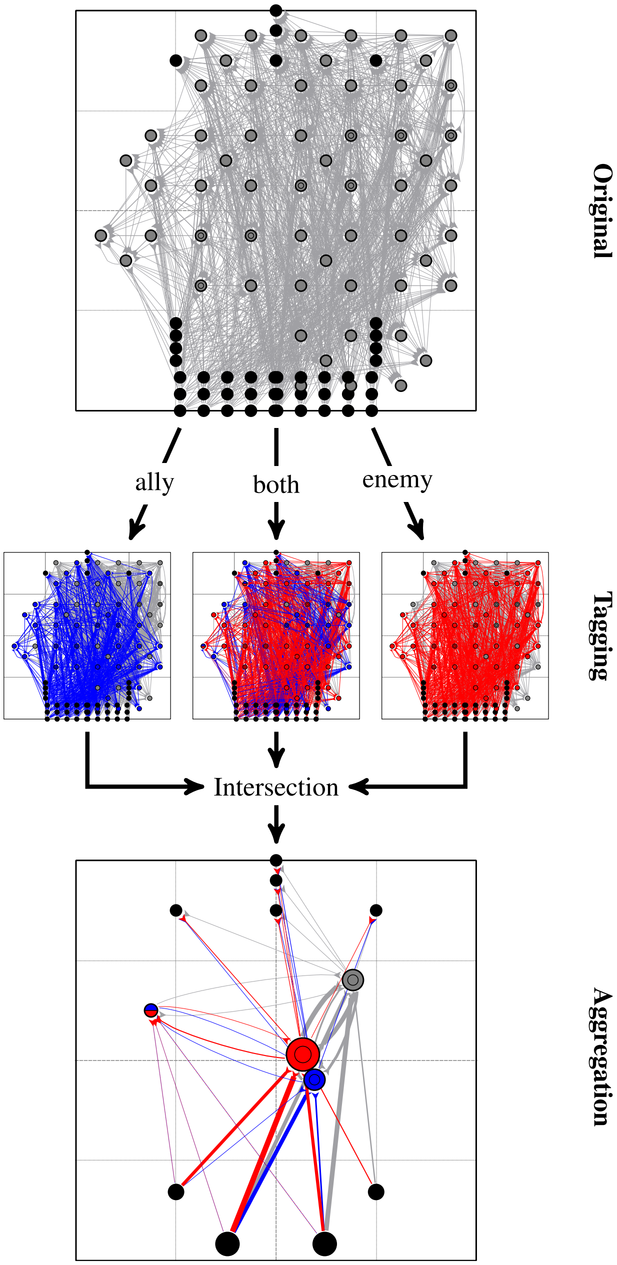
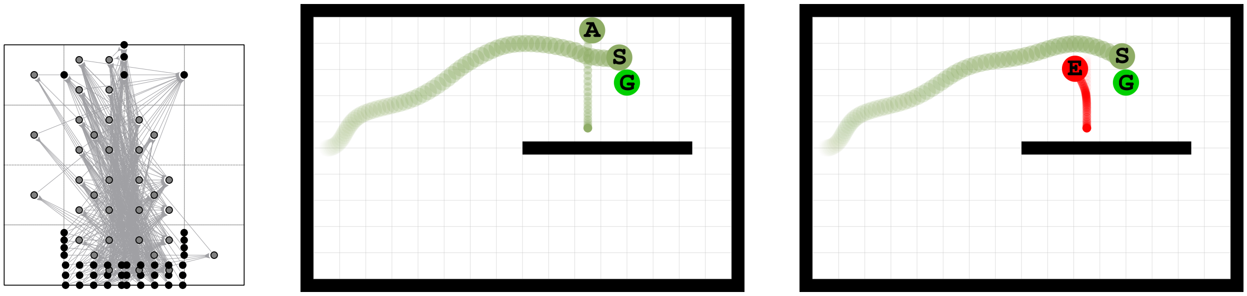


Figure 3: Procedure for producing a modular ANN based on individual neurons' activity patterns in canonical conditions

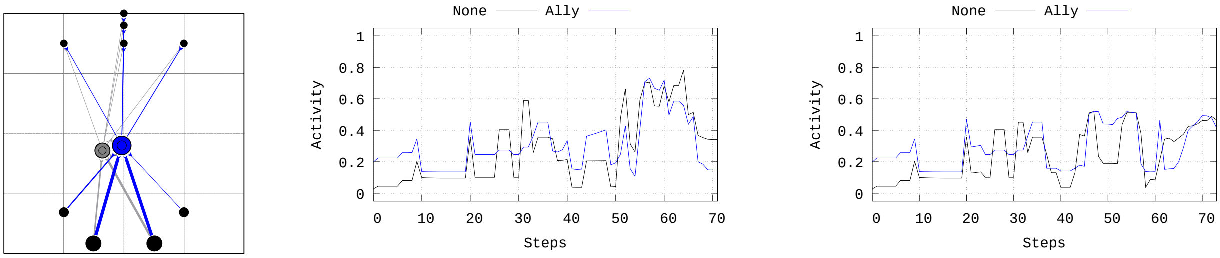
---

[3] I,..., III as opposed to evolution scenarios 0,...,3

[4] The p-value threshold was corrected for the network size n by comparing against $0.05/n$



(a) Neural Network (c) Hospitable conditions (1+) (e) Dangerous conditions (2+)



(b) Modular Network (d) Modular responses in 1+ (f) Modular responses in 2+

Figure 4: No enemy-dedicated neurons and (almost) no impact of predator presence on behavior

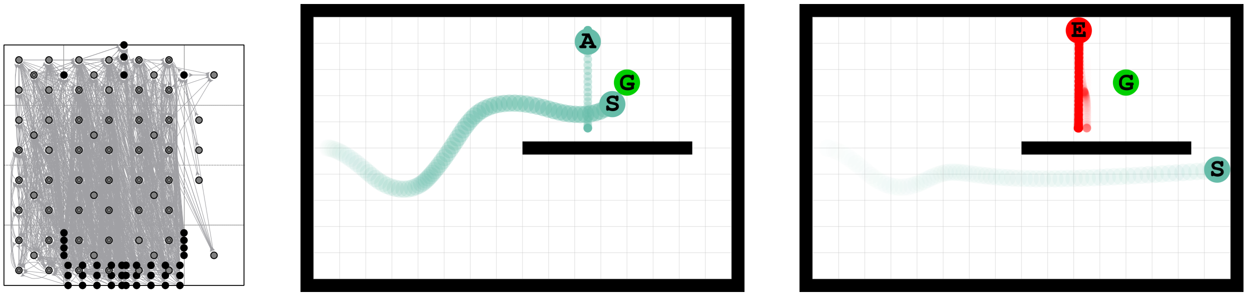
In this manner, one can extract meaningful structures from a seemingly chaotic distribution of neurons. In the following sections, we will leverage this modular transposition to better describe mid-level dynamics which will allow for the investigation of "mental states".

Neural characterization of fear

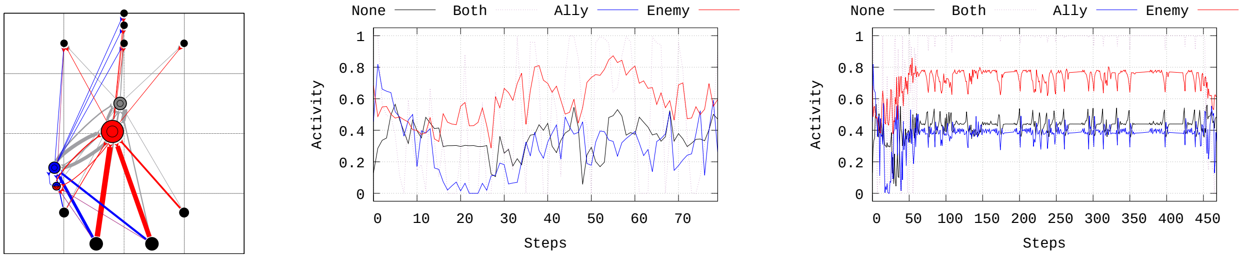
Thanks to the transformation of an arbitrary neural network into a functionally partitioned version, we can more easily map observed behavior with modular activity to determine whether or not specific types of processing occur in the network. While the clustering procedure could be performed for any stimulus encountered by the subjects, we focus mainly on the two types illustrated in the previous section: enemy (predator) which is associated with negative outcomes and ally (clone) which is more ambiguous as it shares many characteristics with the former case. More specifically, we arbitrarily define the enemy-focused module as processing the "fear" response of the individual given that, in this case, it is the sole source of danger. However, mapping the other modules is a more delicate task: while neurons in the generic module can be seen as the background processing center of the individual (food, obstacles ...) there is no clear interpretation for neurons only associated with the ally or with both types of stimuli. Informally, we could describe the former as encompassing the creature's "hope" of a positive outcome, given that the presence of an ally is always associated with proximity to the food (whereas the enemy is only next to the goal in half the scenarios). To better understand how this clustering procedure streamlines the reading of "emotional" responses we illustrate two opposed reactions: bold and fearful (figs. 4, 5).

In the former case, one can see that the neural network (fig. 4a) is of moderate density (35 hidden neurons out of a theoretical maximum of 84) and connectivity (17% of potential connections are expressed). Observing this individuals' behavior on scenarios where either the ally or enemy is present shows very little difference in terms of trajectories (figs. 4c,4e) and speed (data not shown). Indeed the "bold" characterization is based upon this distinct lack of behavioral differences between hospitable and dangerous conditions. As illustrated by the modular version of the ANN, one can see that such boldness is reflected in the functional partitioning: no neurons were found significantly more active in the presence of a predator, either by itself or intertwined with the ally stimulus. Both of the detected modules are of comparable size (40/60% in favor of the ally), highlighting the cognitive investment of this individual into conspecific detection/identification. Indeed, correctly determining the ally's position can drastically increase efficiency in scenarios of type 1 and 3 with the neurons in the remaining module providing a fall-back mechanism in both other cases. Moreover, this lack of reaction to the presence of a predator is also clearly visible by comparing the activity levels of the detected modules in both types of scenarios (figs. 4d,4f): in the former, encountering the ally induces significantly higher levels of activity in the associated module whereas in the latter case no such increase can be observed.

In the case of the fearful creature of figure 5, the trends are inverted with almost twice as many neurons and a denser connectivity pattern (25%). The difference is much more



(a) Neural Network (c) Hospitable conditions (1+) (e) Dangerous conditions (2+)



(b) Modular Network (d) Modular responses in 1+ (f) Modular responses in 2+

Figure 5: Large predator-focused module and exhibition of behavioral and neuronal fears.

marked when comparing the foraging capabilities: while it performs with similar efficiency when faced with an ally, it favors defection in the confrontational case: instead of aiming for the potentially high reward associated with food collection, it prefers to settle for the safer, albeit markedly lower, proximity-based reward. Such a strategy ensures that the penalty for capture is avoided which is a conservative approach to this evolutionary game. As could be expected, this fearful demeanor is associated with an especially large proportion of neurons (57%) linked with the predator stimulus. Unlike the previous individual, a very small bimodal module was also detected that responds to both ally and predator inputs. Due to its limited size (2 neurons), it is unclear whether or not it has a crucial role in the overall dynamics of the network and, is thus displayed in dotted patterns on figures 5d and 5f.

However, the dynamics of the other three modules are much clearer thanks to the normalization resulting from their larger size. On the one hand, when comparing the activity patterns of the neutral and ally modules, one can see that, following an initial decrease in the first steps, it quickly settles into a semi-stable state especially visible in the type 2 scenario. On the other hand, the evolution of activity levels in the predator-focused module is much more diverse with the relatively high, albeit oscillating, values observed in the hospitable case being replaced with a quickly increasing trend when in the presence of the predator. Numerically speaking, around the 50th time step this module has converged onto its maximal output of 0.8 whereas such a value is only reached two times in the, admittedly shorter, hospitable case. We can also note that the high stress level experienced by the individual only starts to decrease at the very end of the evaluation, potentially presaging a relaxation period where it could have resumed its foraging. However, due to the energy constraints of the protocol, it failed to do so in time to reap any benefits.

Lesion study

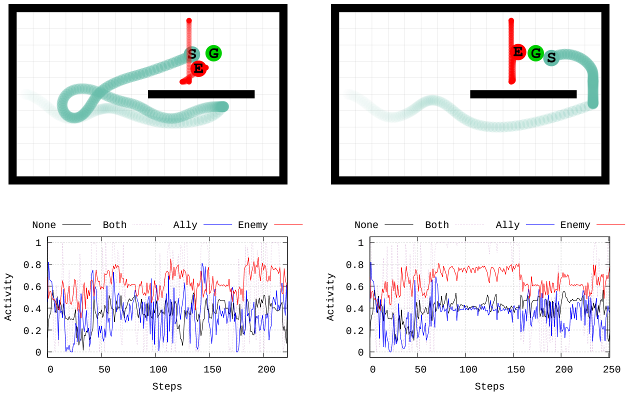
While we identified, on some networks, a specific subset of neurons that could reliably be assigned to the enemy, it remains unclear whether such a module is responsible for any form of mental state. Indeed it might only encode for a set of navigational procedures that are relevant when faced with a predator[5]. Thus, to address this question we first turn back to our previously observed fearful individual and perform additional evaluations of its behavior when two types of lesions are performed upon its network.

In the first case, its visual inputs are impaired by selectively removing all connections between the neurons receiving the *red* input and those of the predator module. The second case is similar, only we remove the connections from the two neurons responsible for perceiving noise (i.e. non vocalized sounds). These intrusive manipulations are designed to be as limited as can be, indeed they do not preclude predator detection by the remaining portions of the creature's network: only direct input into the investigated fear center.

As illustrated by figure 6, there is a notable change in strategy when compared with the initial choice of defection.

---

[5] Although one could argue that it would be one manifestation of fear, as in flight/fight or freezing behavior. See initial quote from (Adolphs, 2013).



(a) Vision

(b) Audition

Figure 6: Behavioral and neural variations when subjected to partial visual (left) and auditory (right) lesions in the same conditions as figure 5e.

Interestingly, both types of lesions result in drastically different variations of the outcome. When visually impaired it backtracks before diving directly towards the food source, only by doing so it fails to evade the predator. Its neural outputs, especially in the red module, paint a similar picture to that observed through its behavior: the lack of visual information being fed into the aforementioned module results in noticeably lower activation levels which can be interpreted as a lower fear response. Oppositely, in the case of auditory impairment, its outcome is better than the base version thanks to a relaxation occurring around the 150th time step. We hypothesize that, in this particular instance, the downward trend observed at the very end of the unmodified conditions was facilitated by the lack of noise perception. This induced a faster turnabout through which the individual was able to perform more efficiently or to put into more psychological terms: the fear-induced freezing state, conditioned by auditory cues, was markedly lower when disconnecting a selective region of the creature's brain.

Based on these observations, one can conclude that, in the case of this specific individual, the predator module behaved as a defensive survival circuit. Indeed, negatively associated inputs (red color, noise) induced an increase in activity while targeted lesions decreased the overall fear perception while preserving the remaining functionalities: in both cases it still show appropriate foraging behavior.

We tested whether this individual's reaction was a random, isolated event by performing the same evaluation (for scenarios of type 2 and 3, i.e. with a predator) on all 43 members of our population that exhibited a predator module, as summarized in table 2. In the "natural" conditions we find that only lesions to the auditory inputs induce a statistically significant variation in the activation level of their fear center. However, one must recall that the individual's initial strategy can provide a very strong bias as, for instance, they do not necessarily make visual contact with the enemy. To alleviate this problem we also performed the comparison on type II and III scenarios which correspond to the controlled conditions used for the clustering procedure. In this case, we find that there is a strong reaction to both types of lesions, demonstrating that this fear center behaves in a modular manner by parsing specific inputs. Moreover, we also evaluated each individual in isolation[6] to determine behavioral classes and, as can be seen in table 3, found that subgroups exist in the population with respect to their reaction to lesions. Thus, while not explicitly required by the experimental protocol, we have observed the emergence of fear-related structures in an artificial neural substrate for a fairly large number of independent runs.

Lesion	Scenarios	
	Evolution	Control
Visual	.122	.009
Auditory	.001	.000

Table 2: Variations in fear center activation in "natural" versus controlled conditions. All but one (in gray) are positive under a 0.05 p-value threshold (Wilcoxon test)

Lesion	Reaction		
	Positive	Negative	Irrelevant
Visual	16	27	7
Auditory	35	8	

Table 3: Responses to lesions of the fear center. The impact is markedly higher with auditory lesions.

Conclusion

To investigate whether structural similarities with the biological brain could emerge from an artificial substrate, we devised an experiment in which creatures only relied on a low-level perception/cognition/action loop. By using a highly indirect encoding, namely ES-HyperNEAT, we were able to abstract the computational units (hidden neurons) both in terms of position and density. Additionally, the connectivity patterns of such neural networks, being also encoded in a similarly indirect manner, were capable of highly regular topologies using geometrical (e.g. left/right motor) and modal (e.g. red, green, or blue color) information. Various strategies emerged from the evolution in a paradoxical setting which required a balance between aggressive foraging and conservative exploration. As observed, not all final creatures were concerned with the presence of a predator some, instead, relying on boldness to reach the objective be-

---

[6] under a threshold of $p < 0.0001$ to correct for the number of evaluations

fore being captured.

However, a significant portion of the studied population did show varying degrees of cautiousness either by avoiding confrontation or by relying on round-about trajectories. By using a clustering procedure, partly inspired by the methodology of fMRI, we were able to abstract from the neural layer into a higher, functional, layer which was instrumental in highlighting the diversity of dynamics exhibited by the population. Thanks to this combined approach we were able to assert that it is possible to obtain individuals which exhibit fear-like behavior through the emergent production of neural structures.

Furthermore, by selectively disconnecting threat-related inputs from our investigated "fear center" we observed a significant neural/behavioral change with respect to the predator with little impact on their other capabilities (e.g. foraging). While this confirms that a modular organization can spontaneously appear in an unconstrained, undifferentiated ANN, it is even more interesting to note how well this fits into Higher-Order Theories of consciousness. Indeed the information deleted by such lesions corresponds to the "first-order representation of the threat" (Ledoux and Brown, 2017) which, in such elementary creatures, is the sole order of information processing.

We can thus conclude that a) functional clustering can spontaneously emerge, even without explicit selection b) this modularity shares key characteristics with the biological brain notably with respect to the separation of inputs. This further highlights the potential of artificial neural networks both as a tool to better understand our genetic heritage but also as the means to produce autonomous artificial lifeforms.

In future researches, we plan on deepening our understanding of the evolutionary dynamics leading to the emergence of such higher-level structures, most notably by increasing the dimensionality of our ANNs into the third dimension. Although an order of magnitude harder to accurately analyze, the gain in expressive power should facilitate spatial clustering of inputs/outputs. Additionally, in the current setting, modularity was an expected side-effect but was not expressively enforced, or even encouraged. Thus, we plan on investigating the differential benefits of incorporating the connection length into the CPPN inputs, as this was shown to promote short-length connectivity thereby further increasing the similarities with the biological brain. These technical improvements are expected to be instrumental in follow-up studies on emotional responses, the scope of which should encompass not only fear but also happiness and aggression.

Acknowledgment

This work was granted access to the HPC resources of CALMIP supercomputing center (allocation P16043).

References

Adolphs, R. (2013). The Biology of Fear. *Current Biology*, 23(2):R79–R93.

Armony, J., Servanschreiber, D., Cohen, J., and Ledoux, J. (1997). Computational modeling of emotion: explorations through the anatomy and physiology of fear conditioning. *Trends in Cognitive Sciences*, 1(1):28–34.

Broekens, J., Jacobs, E., and Jonker, C. M. (2015). A reinforcement learning model of joy, distress, hope and fear. *Connection Science*, 27(3):215–233.

Bălan, O., Moise, G., Moldoveanu, A., Leordeanu, M., and Moldoveanu, F. (2019). Fear Level Classification Based on Emotional Dimensions and Machine Learning Techniques. *Sensors*, 19(7):1738.

de Freitas, J. S., Imbert, R., and Queiroz, J. (2007). Modeling Emotion-Influenced Social Behavior for Intelligent Virtual Agents. pages 370–380.

Delgado-Mata, C., Martinez, J. I., Bee, S., Ruiz-Rodarte, R., and Aylett, R. (2007). On the Use of Virtual Animals with Artificial Fear in Virtual Environments. *New Generation Computing*, 25(2):145–169.

Hendy, N. A. and Farag, H. (2013). Emotion Recognition Using Neural Network: A Comparative Study. *International Journal of Computer and Information Engineering*, 7(3):433–439.

Hesp, C., Heerebout, B. T., and Phaf, R. H. (2020). Evolutionary computation for bottom-up hypothesis generation on emotion and communication. *Connection Science*, pages 1–25.

Ledoux, J. (1998). *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*.

Ledoux, J. E. and Brown, R. (2017). A higher-order theory of emotional consciousness. *Proceedings of the National Academy of Sciences of the United States of America*, 114(10):E2016–E2025.

Lotfi, E. and Akbarzadeh-T., M.-R. (2014). Practical emotional neural networks. *Neural Networks*, 59:61–72.

Ohman, A. (1986). Face the beast and fear the face: animal and social fears as prototypes for evolutionary analyses of emotion. *Psychophysiology*, 23(2):123–45.

Risi, S. and Stanley, K. O. (2012). An Enhanced Hypercube-Based Encoding for Evolving the Placement, Density, and Connectivity of Neurons. *Artificial Life*, 18(4):331–363.

Stanley, K. O., D'Ambrosio, D. B., and Gauci, J. (2009). A Hypercube-Based Encoding for Evolving Large-Scale Neural Networks. *Artificial Life*, 15(2):185–212.

Treccani, C. (2020). The brain, the artificial neural network and the snake: why we see what we see. *AI & SOCIETY*.

Van Le, Q., Isbell, L. A., Matsumoto, J., Nguyen, M., Hori, E., Maior, R. S., Tomaz, C., Tran, A. H., Ono, T., and Nishijo, H. (2013). Pulvinar neurons reveal neurobiological evidence of past selection for rapid detection of snakes. *Proceedings of the National Academy of Sciences*, 110(47):19000–19005.