Developmental Neural Networks

Dennis G. Wilson
Sylvain Cussat-Blanc
Julian F. Miller

10/09/2019

Developmental Neural Networks Workshop

Developmental Neural Networks Workshop

- First edition at PPSN 2018
- Second edition at ALIFE 2019
- Third edition at ALIFE 2020?
- Neuroscience keynote, submitted works, panel discussions
- `irit.fr/devonn/`

2019: ALIFE

14:00-14:20	**Introduction to Developmental Neural Networks**
	Dennis G. Wilson, Julian F. Miller, and Sylvain Cussat-Blanc
14:20-15:30	**Connectome Development – From Local Neuronal Links to Global Fibre Tract Brain Networks**
	Roman Bauer, Marcus Kaiser
15:30-16:00	**Break**
16:00-16:20	**Normalisation of Weights and Firing Rates in Spiking Neural Networks with Spike-Timing-Dependent Plasticity**
	Kasia Kozdon, Peter Bentley
16:20-16:40	**Evolving an artificial brain to solve multiple problems**
	Julian Miller
16:40-17:00	**Learning to Talk in a Gesture-Rich World: Application in Cognitive Robotics**
	Gabriella Pizzuto
17:00-17:30	**Panel discussions**

Why should we study Developmental Neural Networks?

Adult songbirds

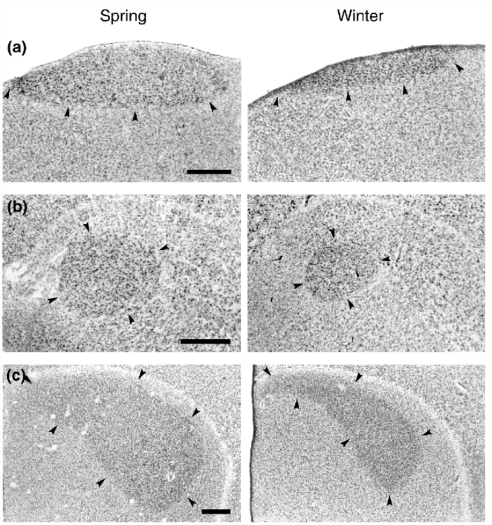


Image: [Tramontin and Brenowitz, 2000]

London taxi drivers

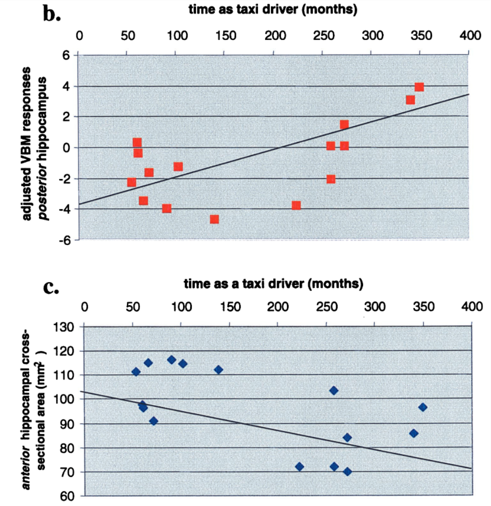


Image: [Maguire et al., 2000]

Adult owl monkeys

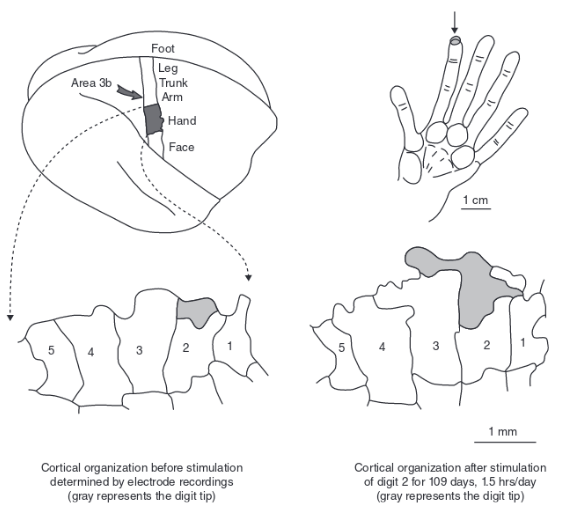
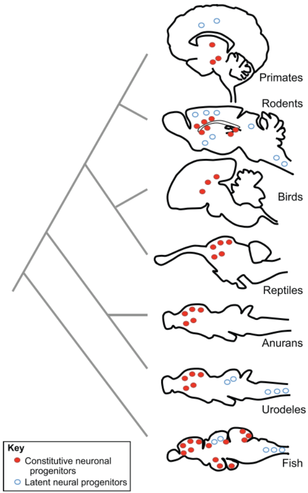


Image: [Jenkins et al., 1990]

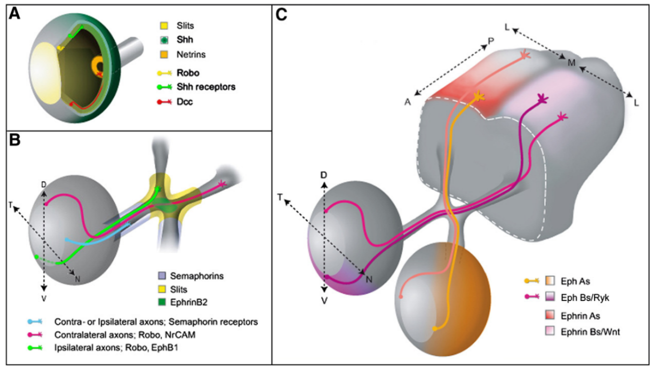
Neurogenesis



		Adult neurogenesis	Latent progenitors	Reparative neurogenesis
Primates	F	+	+	–
	M	–	?	–
	Sc	–	?	–
	R	–	?	–
Rodents	F	+	+	–
	M	–	+	–
	Sc	–	+	–
	R	–	+	+/–
Birds	F	+	?	+
	M	–	–	–
	Sc	–	–	–
	R	–	+	+
Reptiles	F	+	?	+
	M	–	?	–
	Sc	–	?	–
	R	–	–	–
Anurans	F	+	?	–
	M	?	?	–
	Sc	–	?	–
	R	+	+	+
Urodeles	F	+	?	+
	M	–	+	+
	Sc	+	?	+
	R	+	+	+
Fish	F	+	?	+
	M	+	?	?
	Sc	–	+	+
	R	+	+	+

Image: [Alunni and Bally-Cuif, 2016]

Eye-specific patterning



Eye-specific patterning depends on neural activity [Pfeiffenberger et al., 2007]

Image: [Erskine and Herrera, 2007]

Benefits of development

Development can:

- Enable multi-task and lifelong learning
- Optimize structure based on learning
- Ensure robustness through redundancy
- Allow for network healing and repair
- and more!

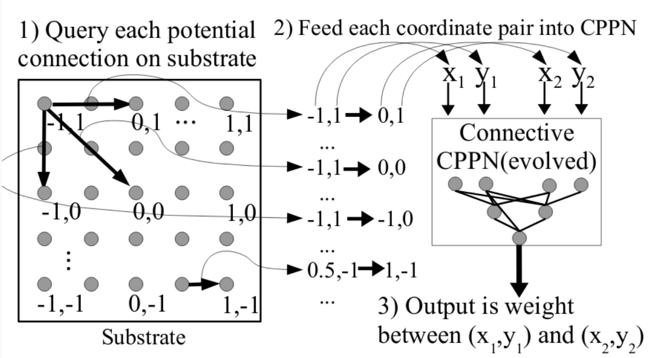
Development in Artificial Neural Networks

Many different methods, which over the years have been called:

- Developmental
- Plastic
- Constructive
- Self-organizing

Our focus:
Artificial Neural Network models which include structural change as a part of learning

Indirect Encoding Neuroevolution



The difference: learning is an epigenetic factor

Image: [Stanley et al., 2009]

Why not study Developmental Neural Networks

- Static Deep Neural Networks already work
- Deep Neuroevolution / Neural Architecture Search is exciting
- It's hard: "The use of developmental strategies for artificial learning systems has shown to be a very complex practice.... It remains unclear how to systematically define developmental stages on the basis of the interaction between innate structure, embodiment, and (active) inference." [Parisi et al., 2019]
- You hadn't yet attended the Developmental Neural Networks workshop

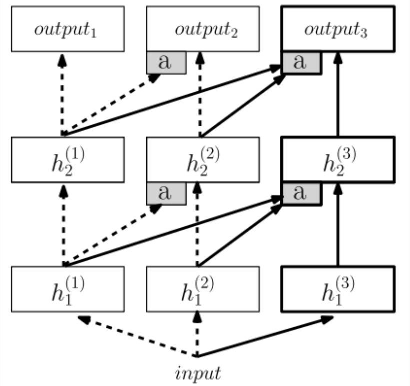
Designed Developmental Rules

Designed Developmental Rules

Developmental Neural Networks which use a heuristic function to determine when structural changes should be made.

- Cascade-correlation [Fahlman and Lebiere, 1990]
- Upstart [Frean, 1990]
- Self-organizing Neural Networks [Horzyk and Tadeusiewicz, 2004]
- Constructive Neural Networks workshop at International Conference on Artificial Neural Networks (ICANN 2008)
- Adaptive Neuron Apoptosis [Siegel et al., 2016]
- AdaNet [Cortes et al., 2017]

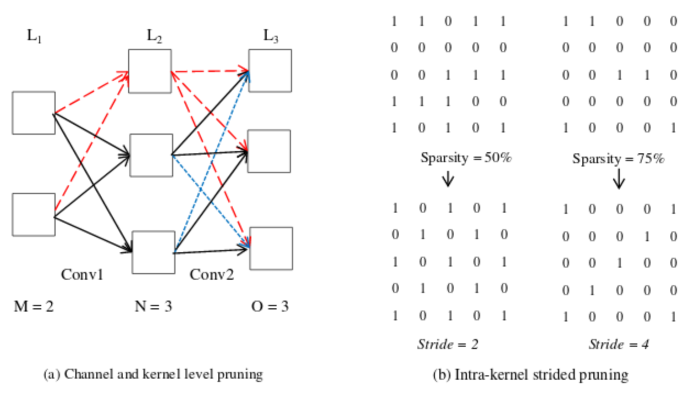
Progressive Neural Networks



Progressive Neural Networks add a new parallel architecture for each new task

Image: [Rusu et al., 2016]

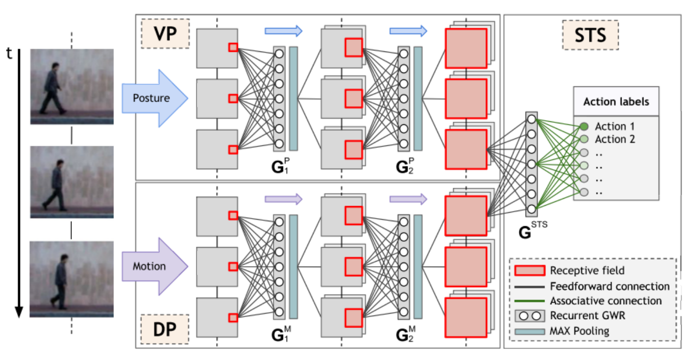
Structured Pruning



Structured pruning removes redundant network weights

Image: [Anwar et al., 2017]

Recurrent GWR network



Recurrent GWR layers modify their neurons and connections based on neural activation

Image: [Parisi et al., 2017]

Recurrent GWR network

Algorithm 1 Associative Gamma-GWR (AG-GWR)

1: Start with a set of two random neurons, $A = \{\mathbf{w}_1, \mathbf{w}_2\}$ with empty context vectors $\mathbf{c}_i^k$ for $k = 1, ..., K$, $i = 1, 2$.
2: Initialize an empty set of connections $E = \emptyset$.
3: [AG-GWR only] Initialize an empty label matrix $H(i, l) = \emptyset$.
4: Initialize K empty global contexts $\mathbf{C}_k = 0$.
5: At each iteration, generate an input sample $\mathbf{x}(t)$ with label ξ.
6: Select the best and second-best matching neurons (Eq. 8):
$b = \arg\min_{i \in A} d_i(t)$, $s = \arg\min_{i \in A/\{b\}} d_i(t)$.
7: Update context descriptors:
$\mathbf{C}_k(t) = \beta \cdot \mathbf{c}_{b(t-1)}^k + (1 - \beta) \cdot \mathbf{c}_{b(t-1)}^{k-1}$.
8: Create a connection $E = E \cup \{(b, s)\}$ if it does not exist and set its age to 0.
9: If $(\exp(-d_b(t)) < a_T)$ and $(\eta_b < f_T)$ then:
a: Add a new neuron r $(A = A \cup \{r\})$:
$\mathbf{w}_r = 0.5 \cdot (\mathbf{x}(t) + \mathbf{w}_b)$, $\mathbf{c}_r^k = 0.5 \cdot (\mathbf{C}_k(t) + \mathbf{c}_i^k)$, $\eta_r = 1$.
b: Update edges between neurons:
$E = E \cup \{(r, b), (r, s)\}$ and $E = E/\{(b, s)\}$.
c: [AG-GWR only] Associate the sample label ξ to the neuron r:
If $(\xi \neq \emptyset)$: $H(r, \xi) = 1$, $H(r, l) = 0$, with $l \in L/\{\xi\}$.
10: If no new node is added:
a: Update weight and context of the winning neuron and its neighbors:
$\Delta\mathbf{w}_i = \epsilon_i \cdot \eta_i \cdot (\mathbf{x}(t) - \mathbf{w}_i)$, $\Delta\mathbf{c}_i^k = \epsilon_i \cdot \eta_i \cdot (\mathbf{C}_k(t) - \mathbf{c}_i^k)$.
b: [AG-GWR only] Update label values of b according to the sample label ξ:
If $(\xi \neq \emptyset)$: $\Delta H(b, \xi) = \delta^+$, $\Delta H(b, l) = -\delta^-$, with $l \in L/\{\xi\}$.
11: Increment the age of all edges connected to b of 1.
12: Reduce the firing counters of the best-matching neuron and its neighbors:
$\Delta\eta_i = \tau_i \cdot \kappa \cdot (1 - \eta_i) - \tau_i$.
13: Remove all edges with ages larger than μ_{max} and remove neurons without edges.
14: If the stop criterion is not met, repeat from step 5.

Image: [Parisi et al., 2017]

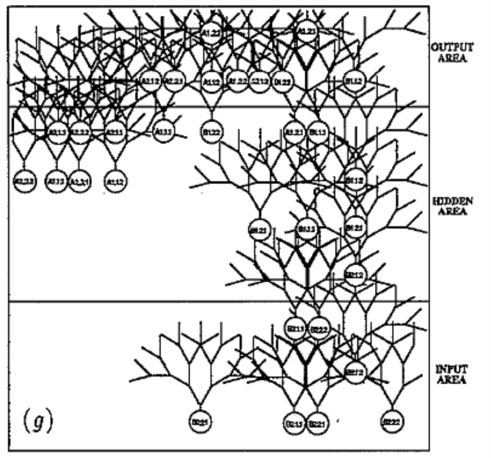
Generated Developmental Rules

Generated Developmental Rules

Generate the function which determines structural changes, often using artificial evolution, called evo-devo methods.

- Graph generation system [Kitano, 1990]
- Cellular Encoding [Gruau and others, 1994]
- Binary networks in cells [Dellaert and Beer, 1994]
- CGP-based approaches [Miller and Khan, 2011], [Miller and Wilson, 2017]
- aGRN evoluion for Spiking Neural Networks [Federici, 2005]
- Probabilistic Program Neurogenesis [Martin and Pilly, 2019]

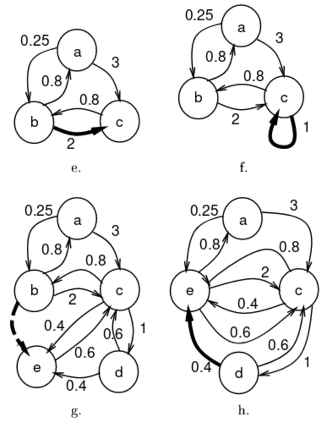
Cell division and migration



Evolved rules for cellular automaton system starting from a single cell

Image: [Cangelosi et al., 1994]

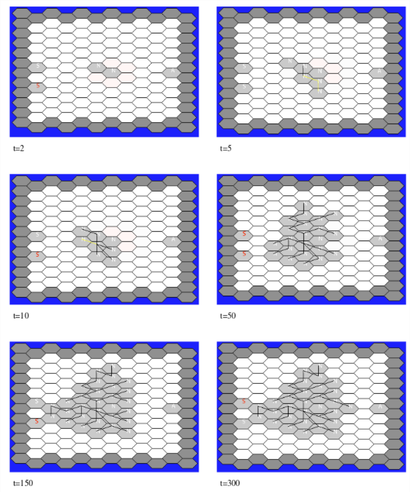
L-Systems



Brain and body morphology were co-evolved using L-Systems

Image: [Hornby and Pollack, 2001]

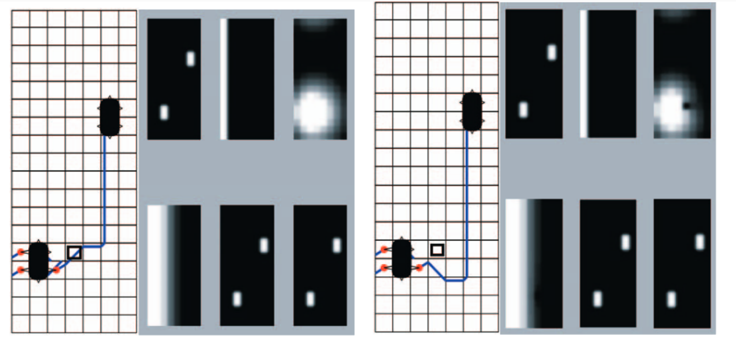
Morphogenesis



GRNs control morphogenesis and cell actions in a 2D grid

Image: [Astor and Adami, 2000]

Adaptive Spiking Neural Network Development



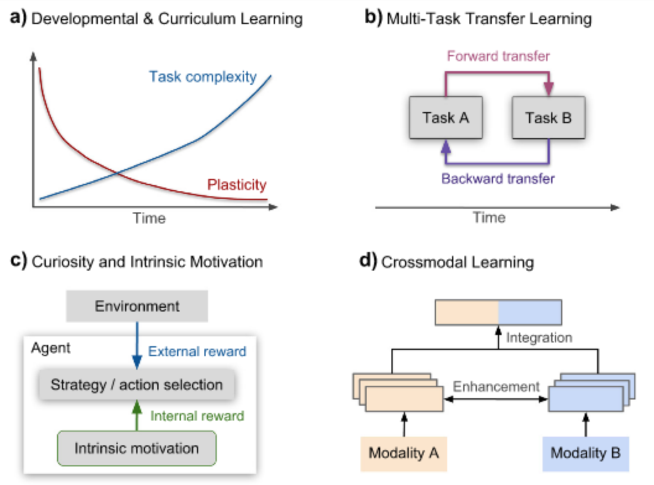
(a) (b)

Development allows for adaptation to faulty cells

Image: [Shayani et al., 2009]

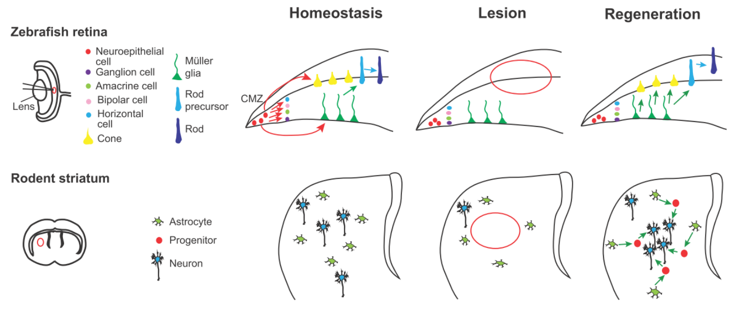
The future of Developmental Neural Networks

Applications of development



Lifelong (continuous, multi-task) learning

Image: [Parisi et al., 2019]



Network healing and regeneration

Image: [Alunni and Bally-Cuif, 2016]

Applications of development

For discussion (end of workshop):
Where can development be useful?

- Multi-task, continuous, lifelong learning
- Active Neural Architecture Search
- Regeneration, healing
- Understanding biological development
- Opening up new forms of learning

Goals and challenges of development

What challenges are there in working on development?

- Complexity, especially in Deep Neural Networks
- Dissonance between cell-based and layer-based models
- Computational cost
- Lack of clarity in biological development
- Evaluation metric, how to measure development

The future of Developmental Neural Networks

You!

Alunni, A. and Bally-Cuif, L. (2016).
A comparative view of regenerative neurogenesis in vertebrates.
Development, 143(5):741–753.

Anwar, S., Hwang, K., and Sung, W. (2017).
Structured pruning of deep convolutional neural networks.
ACM Journal on Emerging Technologies in Computing Systems (JETC), 13(3):32.

Astor, J. C. and Adami, C. (2000).
A developmental model for the evolution of artificial neural networks.
Artificial life, 6(3):189–218.

Cangelosi, A., Parisi, D., and Nolfi, S. (1994).

Cell division and migration in a 'genotype'for neural networks.
Network: computation in neural systems, 5(4):497–515.

Cortes, C., Gonzalvo, X., Kuznetsov, V., Mohri, M., and Yang, S. (2017).
Adanet: Adaptive structural learning of artificial neural networks.
In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 874–883. JMLR. org.

Dellaert, F. and Beer, R. D. (1994).
Toward an evolvable model of development for autonomous agent synthesis.
In *Artificial Life IV, Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems*, pages 246–257. MIT press.

Erskine, L. and Herrera, E. (2007).
The retinal ganglion cell axon's journey: insights into molecular mechanisms of axon guidance.
Developmental biology, 308(1):1–14.

Fahlman, S. E. and Lebiere, C. (1990).
The cascade-correlation learning architecture.
In *Advances in neural information processing systems*, pages 524–532.

Federici, D. (2005).
Evolving developing spiking neural networks.
In *2005 IEEE Congress on Evolutionary Computation*, volume 1, pages 543–550. IEEE.

Frean, M. (1990).
The upstart algorithm: A method for constructing and training feedforward neural networks.

Neural computation, 2(2):198–209.

Gruau, F. and others (1994).
Neural Network Synthesis Using Cellular Encoding And The Genetic Algorithm.

Hornby, G. S. and Pollack, J. B. (2001).
Body-brain co-evolution using L-systems as a generative encoding.
In *Proceedings of the 3rd Annual Conference on Genetic and Evolutionary Computation*, pages 868–875. Morgan Kaufmann Publishers Inc.

Horzyk, A. and Tadeusiewicz, R. (2004).
Self-optimizing neural networks.
In *International Symposium on Neural Networks*, pages 150–155. Springer.

Jenkins, W. M., Merzenich, M. M., Ochs, M. T., Allard, T., and Guic-Robles, E. (1990).
Functional reorganization of primary somatosensory cortex in adult owl monkeys after behaviorally controlled tactile stimulation.
Journal of neurophysiology, 63(1):82–104.

Kitano, H. (1990).
Designing neural networks using genetic algorithms with graph generation system.
Complex systems, 4(4):461–476.

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S., and Frith, C. D. (2000).
Navigation-related structural change in the hippocampi of taxi drivers.

Proceedings of the National Academy of Sciences, 97(8):4398–4403.

Martin, C. E. and Pilly, P. K. (2019).
Probabilistic Program Neurogenesis.
In *The 2018 Conference on Artificial Life: A Hybrid of the European Conference on Artificial Life (ECAL) and the International Conference on the Synthesis and Simulation of Living Systems (ALIFE)*, pages 440–447. MIT Press.

Miller, J. F. and Khan, G. M. (2011).
Where is the Brain inside the Brain?
Memetic Computing, 3(3):217–228.

Miller, J. F. and Wilson, D. G. (2017).
A developmental artificial neural network model for solving multiple problems.

In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pages 69–70. ACM.

Parisi, G. I., Kemker, R., Part, J. L., Kanan, C., and Wermter, S. (2019).
Continual lifelong learning with neural networks: A review.
Neural Networks.

Parisi, G. I., Tani, J., Weber, C., and Wermter, S. (2017).
Lifelong learning of human actions with deep neural network self-organization.
Neural Networks, 96:137–149.

Pfeiffenberger, C., Cutforth, T., Woods, G., Yamada, J., Rene, C., Copenhagen, D. R., Flanagan, J. G., and Feldheim, D. A. (2007).

Ephrin-As and neural activity are required for eye-specific patterning during retinogeniculate mapping.
8(8):1022–1027.

Rusu, A. A., Rabinowitz, N. C., Desjardins, G., Soyer, H., Kirkpatrick, J., Kavukcuoglu, K., Pascanu, R., and Hadsell, R. (2016).
Progressive neural networks.
arXiv preprint arXiv:1606.04671.

Shayani, H., Bentley, P. J., and Tyrrell, A. M. (2009).
A multi-cellular developmental representation for evolution of adaptive spiking neural microcircuits in an FPGA.
Proceedings - 2009 NASA/ESA Conference on Adaptive Hardware and Systems, AHS 2009, (January 2016):3–10.

Siegel, C., Daily, J., and Vishnu, A. (2016).

Adaptive neuron apoptosis for accelerating deep learning on large scale systems.
In *2016 IEEE International Conference on Big Data (Big Data)*, pages 753–762. IEEE.

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.

Tramontin, A. D. and Brenowitz, E. A. (2000).
Seasonal plasticity in the adult brain.
Trends in neurosciences, 23(6):251–258.

Developmental Neural Networks

Dennis G. Wilson
Sylvain Cussat-Blanc
Julian F. Miller

10/09/2019