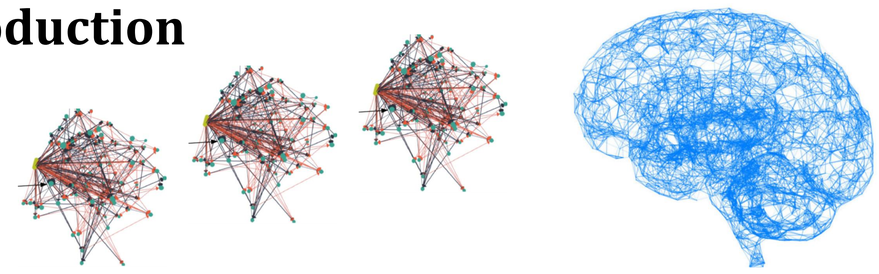
Neuromorphic Computation and the Reservoir: From Theory to Device

By Jack Mayo & Rex Albert Villahermosa
In collaboration with:
Gabriel Leuenberger
Faculty of Science and Engineering
University of Groningen

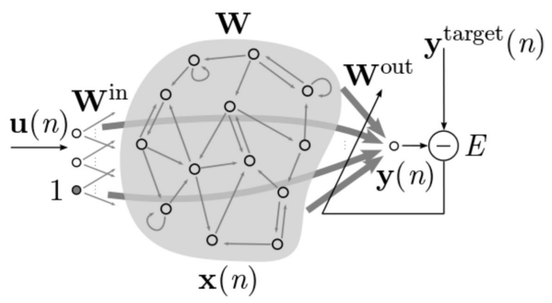
Introduction



General intelligence and cognitive learning have not yet been fully realized in today's computers, which are based on the Boolean logic and the von Neumann architecture. In contrast to biological systems, today's computers are less energy-efficient by a factor of one million to one billion for managing difficult real-world environmental problems such as image or speech recognition.

To deal with this challenge, a new computing paradigm that emulates the brain, namely, *neuromorphic computing*, is being vastly explored. This approach generally takes the advantage of massive parallelism in neural networks, with the goal of achieving energy-efficiency, robustness, and inherent error-tolerant potential.

Reservoir Computing And the Echo State Network



An Echo State Network (ESN) is one of a class of algorithms under the heading of Reservoir Computing (RC), whereby computation is performed by perturbations of a dynamical state governed by the equations.

$$\mathbf{x}(n+1) = f(W\mathbf{x}(n) + W^{in}\mathbf{u}(n) + W^{fb}\mathbf{y}(n)) \tag{1}$$

$$W^{out}\mathbf{x}(n) = \mathbf{y}(n) \tag{2}$$

Where $\mathbf{x}(t)$ denotes the internal state of the reservoir, $\mathbf{u}(t)$ the input vector, $\mathbf{z}(t)$ the output and the matrices W,W^{in} and W^{fb} encode how the respective variables influence the evolution of the state. $h(\mathbf{x}(t))$ is a readout function on the state allowing determination of the output $\mathbf{y}(t)$.

Unlike a classical recurrent network, it is only the readout matrix W^{out} which must be trained in order to achieve the necessary functionality.

The training is usually done by a linear regression, minimizing a cost function J of the form:

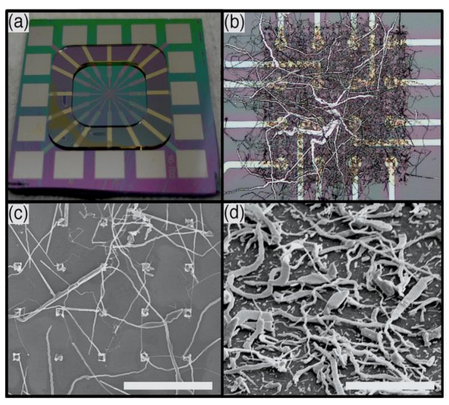
$$J = \frac{1}{2N}\sum_{\mu=1}^{P} ||\mathbf{y}_\mu - \mathbf{y}_\mu^{target}||^2 \tag{3}$$

Where $\mathbf{y}^{target}$ denotes the desired output for a particular data point μ, and the weights are updated by

$$(W^{out})_{ij} \to (W^{out})_{ij} - \eta\frac{\partial J}{\partial (W^{out})_{ij}} \tag{4}$$

where η determines the rate of learning. This process is referred to as a gradient descent.

Existing Physical Reservoirs



- **Atomic Switches**

Self-assembled Silver nanowire based atomic switch network reservoirs with one input and 16 outputs that were successfully tested in simple waveform generation tasks.

- **Memristors**

Similar to resistors, memristors are electrical components, but their resistance changes depending on the past electric flux.

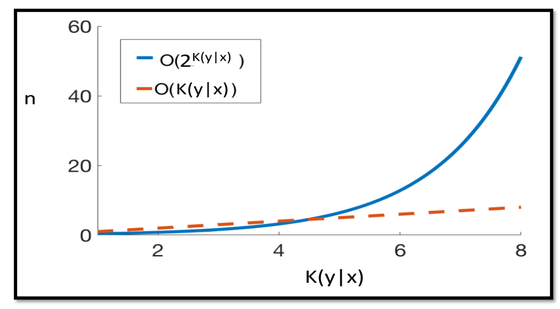
- **Spintronics**

Spintronics is similar to electronics but it additionally exploits the spin of the electron.

- **Photonics**

Photonic or optical circuits use light for computation. Alloptical high-speed reservoirs were successfully tested in 8-bit-pattern-recognition tasks with bit rates of up to 160 Gbps.

Limitations of Classical Reservoir Computing



Despite excellent efficiency on many basic tasks, the ESN has a scaling problem, formulated in the following manner. Generally speaking in a supervised learning task, a program p should be found which maps an input x to the desired output y, aka the target. We can represent programs as bit strings. Let $K(y|x)$ denote the length of the shortest bit string that maps the input x to the target y (Kolmogorov complexity of y relative to x). Considering the reservoir as a randomly chosen program, and noting that there exist $2^{K(y|x)}$ programs of length $K(y|x)$ we may estimate (using Levin's Coding theorem) the probability of the reservoir yielding the correct program via a single readout as:

$$P_1 = 2^{-K(y|x)} \tag{5}$$

If we have n readouts from which we select only the one with the corret output, the probability of success becomes:

$$P_n = 1 - (1 - P_1)^n \leq n \cdot P_1 = n \cdot 2^{-K(y|x)}, \tag{6}$$

and so:

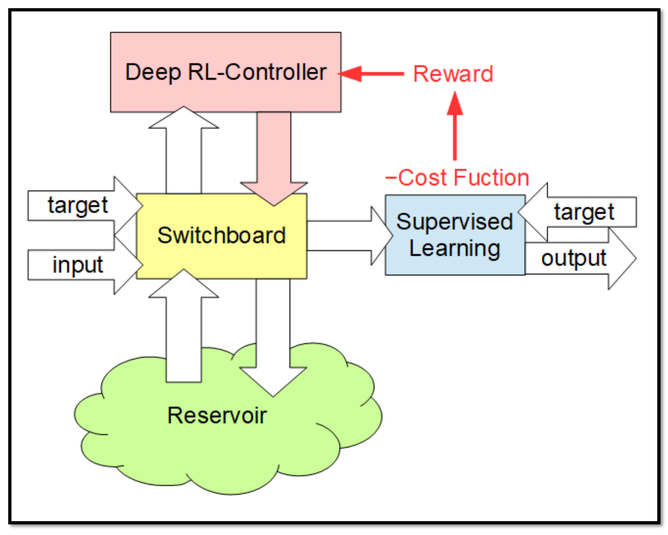
$$n \geq P_n \cdot 2^{K(y|x)}. \tag{7}$$

Meaning that, if we keep P_n at a constant satisfactory level, and we scale the size of our system n linearly, $K(y|x)$ can only grow logarithmically.

In classic reservoir computing these readouts are optimally linearly combined to produce a better output. However the aforementioned logarithmic limitation can only disappear under the special assumption that y can be written as a linear combination of outputs of subprograms that are sufficiently short to all appear in the reservoir by chance.

Thus, when scaled up, the classic reservoir computing algorithms can only succeed at learning functions of a logarithmically limited depth given by subprogram size, effectively making it a form of shallow learning as opposed to deep learning algorithms that successfully scale to the depths required for their impressive results.

A Potential Solution: Controlling Internal Dynamics



To overcome the limitations of classical reservoir computing, one central connectivity matrix could can be employed that takes in the input, the target, and the output of the reservoir, and gives out information to the controller, the supervised learning algorithm, and back to the reservoir. The weights of this matrix would be changed continuously by the controller aiming to optimize the computations to increase its future rewards. We call this matrix the switchboard. This architecture is a general framework for the optimal exploitation of large fixed neuromorphic reservoirs. Its success depends strongly on the quality of the employed reinforcement learning algorithm.

References

Sillin, H. et al. (2013). *A theoretical and experimental study of neuromorphic atomic switch networks for reservoir computing.* Nanotechnology, 24(38), 384004.

Steil, J. J. (2004). *Backpropagation-decorrelation: online recurrent learning with o(n) complexity.* In Neural networks, 2004 proceedings.

Yu, S. (2017). *Neuro-inspired computing using resistive synaptic devices.* Springer.

Acknowledgements

To Prof. Michael Biehl and Prof. Patrick Onck.
For the great mentorship and support throughout this project.