SAND2016-3026C

Using Neuromorphic Computing Methods for General Computer Performance Growth

Erik P. DeBenedictis,
Center for Computing Research, Sandia National Laboratories

Neuromorphic computing has been identified as an option for continuing traditional growth levels of the computer industry. The computer industry grew exponentially for decades, establishing exponential growth in computer processing power as the baseline for industry health. Roadmaps project semiconductor chips will not be able to support more logic gates at the same power per chip for more than another decade, meaning per-chip processing power will flat line. If neuromorphic, quantum, and a few other less-known approaches could be better understood, they are seen as possibly enabling continuation of exponential growth.

We will demonstrate that the stall out of energy efficiency growth is not fundamental, but is simply a characteristic of Boolean logic. The pertinent characteristics of neural systems might be different because nature used an independent development path, and we will show this is in fact true. The limits of current computers were studied for power dissipation by Landauer [Landauer 61], who established a thermodynamic minimum energy on the order of kT for each function (k being Boltzmann's constant and T being the absolute temperature) – using the AND gate as a typical function. Since the kT minimum is based on thermodynamics, it is device independent and cannot be beaten by better devices alone. The kT minimum is often called the "Landauer limit," but we will show this is misleading. The source of confusion is that the kT minimum applies to the typical Boolean logic gates that were popular building blocks in the 1960s, but designing computers with a different set of basis functions can yield minimum energy $\ll$kT. The point in this talk is that neuromorphic systems can have a lower minimum energy, joining reversible logic that has had this property since its discovery in the 1970s.

The talk includes an example from neural networks that parallels the AND gate example in Landauer's paper. In lieu of the AND gate, the example is one synapse in a neural network. The state-containing synapse monitors two inputs looking for correlations, learning or remembering the correlation when it occurs. As with many neural systems, this system mostly verifies that is has learned what it needs to know and actually changes state with just a small probability p. The talk applies the same thermodynamic analysis to this modern, non-Boolean, example yet yields a minimum energy close to p kT, which smaller than the Boolean logic limit by the factor p that can be arbitrarily small.

The example generalizes. Landauer created a process for calculating the minimum energy of a computing approach, where you feed in the characteristics of the approach and it produces the theoretical minimum energy. We essentially reverse engineered Landauer's process to identify properties of computing approaches that tend to have lower minimum energies. For example, use of state-containing devices and exploitation of probabilities in the input data can be used to reduce minimum energy. With the current shift from strictly numerical computing to big data (i. e. a lot of state) and learning, the neural approach used in biology becomes a fruitful source of inspiration.

The functions in biological neurons can be reconciled with "Beyond CMOS" device characteristics using the approach described to yield engineered devices that have lower energy minimums for many problems of current interest. Lower minimum energies should yield lower practical energies.

[Landauer 61] Landauer, Rolf. "Irreversibility and heat generation in the computing process." *IBM journal of research and development* 5.3 (1961): 183-191.

Preprint submitted to 16th International Conference on Nanotechnology.
Received March 31, 2016.