# Deep neural networks: the only show in town? A position paper for the workshop on *Can Deep Neural Networks (DNNs) provide the basis for Artificial General Intelligence (AGI)* at AGI 2016, July 2016


Leslie S. Smith
Computing Science and Mathematics
University of Stirling, Stirling FK9 4LA, UK
email l.s.smith@cs.stir.ac.uk


July 13, 2016

## 1 Introduction

*Can Deep Neural Networks (DNNs) provide the basis for Artificial General Intelligence (AGI)* is the topic for this workshop. We do not expect to come to a specific conclusion, but what we provide here are some talks, and space for discussion. However, the speakers seem to have lined up on one side of the division (from my reading of their abstracts), so I propose to take the other side!

But what is a DNN, exactly? Goertzel [Goe15] reckons that it is

> a learning system consisting of adaptive units on multiple layers, where the higher level units recognize patterns in the outputs of the lower level units, and also exert some control over these lower-level units.

Where do the ideas come from? The best place to look is Schmidhuber's detailed paper on DNNs [Sch15], which traces the history a long way back. But in terms of networks with multiple layers that actually learn (as opposed to networks that could learn, like the original multi-layered or gamba perceptron discussed in Minsky and Papert's 1969 book [MS69]) the earliest is probably Fukushima's Neocognitron of 1980 [Fuk80].

But what might really be understood by the phrase *deep neural network* itself? I suggest

> A set of neurons (but what exactly is a neuron in this context?) connected together in a *deep* way (which should mean that quite a few of the neurons are neither used for input nor for output and are so-called *hidden* neurons): it could be taken to mean that there are sequences of neurons that connect only to other hidden neurons.

Often the neurons are taken to be simple entities, generally rather more complex than a simple linear or linear threshold unit, and usually some form of unit that takes a weighted sum of inputs, and then applies a non-linear squashing function like a tanh or a logistic function. But this is not implied in the phrase *deep neural network* itself.

If I restrict the topology to be feed-forward, such DNNs are really not very different from the multi-layered perceptron systems discussed in the mid 1990s and trained using various versions of the back-propagated delta algorithm [RHW86]. Deep versions of these were not successful at the time because they were very difficult to train, as a result of the error surface having an almost null gradient over a large area of the space, making training very slow. In addition, computers of the time were relatively slow. More recent developments [Sch15] do some additional processing too: maximal pooling and/or convolution allows internal layers to summarise the previous layers, and this really helps to get the maximal usefulness from the multiple internal layers. Additionally, some different non-linear functions have been tried, such as rectifier units (where the output stops and stays at 0 for negative activations). DNNs of different types are also in use for reinforcement learning, and for classification using unsupervised learning.

Am I suggesting that these types of DNNs can really *solve the problems* of AGI? Not really, though they can certainly be useful system components. But must the concepts of DNNs stop there? What's missing, and might be added to this mix, whilst retaining the overall meaning of the phrase *deep neural network*?

## 2 Optional additions to the feedforward DNN.

Below, I discuss some additions to the usual meaning of the phrase *deep neural network*.

### 2.1 Time and recurrence

As described by Goertzel, and largely as used, DNNs map vectors to vectors without considering the immediate history of recent inputs. Each vector is considered independently of the others (excepting for long-term history, used in the training of the network). There is a long history of adding time to neural networks and DNNs, in ways ranging from windowing the inputs (running a window of a certain size through the time sequence points), to special algorithms developed for training recurrent networks. Time matters in AGI (and indeed, in many problems) because both the signals from the environment and any outputs that act on the environment are time-varying. There is a long history of techniques for processing time-varying signals, most often trying to predict what the future output will be. The idea is often to form some type of internal model of the problem being solved, and then use this model to make predictions. In that sense, DNNs are being used in a similar way to many other different types of prediction system.

Any application that really is going to be considered as AGI will have to work with time-varying inputs to produce time-varying outputs: the world exists in time, and the reaction of a system exhibiting AGI really has to include time as well. (Of course, one can recognise faces, or handwritten characters without this, but is that intelligence?)

There are other methods of adding time to this system. One way is to use an event (or spike) coded system, rather than a fully time-varying input. This is attractive partly because this does seem to be how data in real neural systems is encoded, and partly because event-based coding can allow a degree of pre-processing to take place first (that is, to determine when an event has taken place). In real neural systems, the nature of the event is coded largely by which axons (wires) the spikes (events) occurred in, at least at the sensory input to the brain. Much of the earlier signal processing at the optical and olfactory sensory surfaces is analogue, with spiking only arriving a little later: that can be contrasted with the auditory domain, where the signal is spike coded after what is largely mechanical processing, though even here, there is adjustment to this coding from a fed-back spiking signal.

Are such systems still DNNs? It seems unreasonable to suggest that adding recurrence, or using spike-coding stops a network from being a DNN, and indeed Goertzel's description above appears to include the recurrence aspect.

### 2.2 Connection to the outside world.

There are different philosophical views about whether one can have an isolated intelligence, one that is not embodied. Without getting embroiled in this discussion, it seems clear that an AGI needs to behave in a domain, that is, it needs some form of embodiment. Of course DNNs (and other systems) are generally connected to the world, and receive (generally rather highly processed) input. This might be visual, or auditory or olfactory or tactile, or some mixture of these sensory systems.

As noted above, the outside world does not provide input and output vectors in isolation: it is a dynamic environment, where the signals arising from the sensory systems are time-varying and where actions that the system might make in this world are also dynamic and time-varying. Processing these time-varying signals is often difficult, primarily because the signals arise from many different sources, and are all mixed together (considers listening to a single speaker when there are many speakers in a reverberant room, or focussing on one object in a visual field that contains many overlapping objects). Many current applications of DNNs avoid these issues by using human preprocessing, or careful selection of the entities to be interpreted. The inputs to cortical neural processing have already undergone considerable neural pre-processing, so that in reality, cortical input signals are signals inside a (real) deep neural network. I suggest that the pathological brittleness that Goertzel [Goe15] comments on is primarily due to a lack of preprocessing of the visual input.

### 2.3 Simple and not-so-simple neurons and interconnections

Even the simplest animals, animals without any neural system, such as single-celled animals like amoebae and paramecia react dynamically to their time-varying environment. They do so using chemical sensors and actuators embedded in their cell membrane. One can argue that rather than having no nervous system, the single cell itself is an unspecialised cell, one that is gut, limbs and nervous system all in one. Multicellular animals usually have specialised cells, and the basic active nerve cell seems to have been found in nature quite early on in evolution.

However anyone who has ever looked into neuroscience knows that real neurons are not simple: the activity-based single value characterisation of a neuron used by McCulloch and Pitts initially, and many others thereafter is a very simple cartoon of a neuron. But between that and the enormous complexity (and, indeed, variation) to be found in real neurons there are many other possible models of intermediate complexity. How much complexity is required? Which of the many aspects of real neural systems really matter for intelligence?

And as if the bewildering complexity and variety of neurons were not enough, the interconnections between neurons are also highly complex, with different chemical cascades governing both the release and the uptake and processing of neurotransmitters that implement the connection between neurons across synapses. Not only that, but there are also gap junctions which connect some neurons, acting to slowly equilibrate the ionic concentrations (and hence the voltage-based activity) between them.

There are are other issues that arise here: in real neural systems, not only are there a range of interconnection types, but not all of them are point-to-point. In traditional DNNs, all the connections simply connect one neuron to another. (There can also be additional less neurally inspired interconnection types that lead to winner-takes-all or convolutional processing.) However, certain real neurons produce specific chemicals and these have an effect on (modulate) synapses (and perhaps neurons as well) in their neighbourhood through diffusion. Depending on the size of the molecules involved (for example, NO is a very small molecule, and will therefore diffuse over larger distances), this modulation can take place over a relatively large distance. This form of long-range modulation is certainly important in real neural systems, but whether it is required for AGI is not clear.

## 3 Depth, or depth and width?

Recent work in neural networks has not been restricted to *deep* NNs, but has also considered what might be called *wide* NNs. These are systems in which the input is projected into a much higher dimensional space, so that this very high dimensional data can then be mapped (perhaps linearly) to an appropriate output. Early work in this area includes radial basis function neural networks, and these were used quite effectively as early forms of single-hidden-layer neural network [Hay99]. Like the early BP networks, these were vector to vector mapping systems.

In more recent years, however, a new form of this type of network has been developed: reservoir networks [MNM02][VSDS07]. In these, a large number of neurons (generally quite simple neurons, though they can also include spiking neurons) are put together into a connected system, in a random way. A smaller number of inputs are used to drive these neurons, again through a set of weights which differ from neuron to neuron, and the output is then classified (or used to produce a temporal prediction) through a further, often linear, network. These networks are *wide*, in the sense that they make the inputs diverge to a large number of neurons.

Interestingly, there are now suggestions that granule cells of the cerebellum are implementing something of this form [CAW$^+$15]. These cells (which are actually the most numerous neurons in humans: they are very small) receive inputs from a small number of mossy fibers that arise from many parts of the brain. The outputs from these granule cells are collected by the much larger Purkinje cells that collect input from a large number of granule cells. The cerebellum is implicated in the fine timing of motor outputs in a very large number of animals: this is surely part of what one thinks of as AGI, so perhaps one should consider deep *and* wide neural networks.

## 4 In conclusion

This workshop poses a difficult question: difficult not just because of the question itself, but because of the difficulty in pinning down the meanings of some of the words in the question. It has long been realised that intelligence is a slippery customer: at one time a program that could play chess or translate from one language to another would have been seen as a good example of an intelligent system, but now the community has realised that intelligence is much more multi-faceted than that, and this has led to the relatively recent term *Artificial General Intelligence*, which was required because of the debasement of the original term *Artificial Intelligence*.

Here, I have tried to show that the concept of a deep neural network extends beyond how it is currently used. To this author, the word *neural* would lead one to expect that the elements in the network can be either as simple as the McCulloch-Pitts neurons of the 1950s or as complex as a real neuron, and that the interconnections within the *network* might be characterisable as a single value (a weight), or might be a whole lot more complex, perhaps including timing as well as the possibility of modulation depending on the state of nearby neurons. In addition, I have suggested that width as well as depth can be useful: the inputs to the (wide layer of) cerebellar granule cells are not (in general) coming directly from the sensory system, but have already been through some layers of processing, so that the actual neural system is both wide and deep.

So what do I mean when I write *Deep neural networks: the only show in town*?

There are clearly other views on ways forward in AGI, such as logical constructions, reasoning systems, etc., and this meeting is the place to find them. But what sort of AGI am I considering here? If it is animal-like, in the sense of surviving and thriving in unpredictable situations, perhaps with some cognition as well: if one is looking at real intelligence, and trying to replicate (or even better) it, then DNNs, including the extended meaning of DNN that I have discussed here, are currently the only techniques that have promise of robustness. Of course such DNNs are not homogenous: there can be structures within such a (deep and wide) DNN that are describable in completely different ways, ways that reflect logic, reasoning and other constructions. Indeed, given that human brains are really DNNs in the above sense, and yet can also be considered logical and/or reasoning devices, I suggest that it is indeed true that deep neural networks are the only show in town.

## References


[CAW+15] François P Chabrol, Alexander Arenz, Martin T Wiechert, Troy W Margrie, and David A DiGregorio. Synaptic diversity enables temporal coding of coincident multisensory inputs in single neurons. *Nature neuroscience*, 18(5):718–727, March 2015.

[Fuk80] Kunihiko Fukushima. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36(4):193–202, April 1980.

[Goe15] Ben Goertzel. Are There Deep Reasons Underlying the Pathologies of Today's Deep Learning Algorithms? In *Artificial General Intelligence*, pages 70–79. Springer International Publishing, Cham, July 2015.

[Hay99] S Haykin. *Neural Networks: a Comprehensive Foundation.* Prentice Hall, 2nd edition, 1999.

[MNM02] Wolfgang Maass, T Natschläger, and H Markram. Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural computation*, 14(11):2531–2560, 2002.

[MS69] M Marvin and P Seymour. *Perceptrons.* MIT Press, 1969.

[RHW86] D E Rumelhart, G E Hinton, and D Williams. Learning representations by back-propagating errors. *Nature*, 323:533–536, October 1986.

[Sch15] Jürgen Schmidhuber. Deep learning in neural networks: an overview. *Neural networks : the official journal of the International Neural Network Society*, 61:85–117, January 2015.

[VSDS07] D Verstraeten, B Schrauwen, M D'Haene, and D Stroobandt. An experimental unification of reservoir computing methods. *Neural Networks*, 20(3):391–403, 2007.