Why is artificial intelligence a hype?

https://www.quora.com/Why-is-artificial-intelligence-a-hype

22 Answers

Brent Oster

, former Deep Learning Solution Architect at NVIDIA (2008-2018)
Updated Sep 25 · Upvoted by

Gopi Sundharam

, Masters Computer Science, Bharathidasan University and

Hubert Andersson

, M.Sc. Electrical and Electronics Engineering & Computer Science, Chalmers University of Technology (2000)

I worked for a decade at NVIDIA, as a Solution Architect to research deep learning techniques, and present solutions to customers to solve their problems and to help implement those solutions. Now, for the past 3 years I have been working on what comes next after DNNs and Deep Learning. I will cover both, showing how it is very difficult to scale DNNs to AGI, and what a better approach would be.

What we usually think of as Artificial Intelligence (AI) today, when we see human-like robots and holograms in our fiction, talking and acting like real people and having human-level or even superhuman intelligence and capabilities, is actually called Artificial General Intelligence (AGI), and it does NOT exist anywhere on earth yet.

What we actually have for AI today is much simpler and much more narrow Deep Learning (DL) that can only do some very specific tasks better than people. It has fundamental limitations that will not allow it to become AGI, so if that is our goal, we need to innovate and come up with better networks and better methods for shaping them into an artificial intelligence.

Lets look at:

1. Where Deep Learning and Reinforcement Learning are today
2. What their limitations are - what can they do and not do?
3. The neuroscience of human intelligence
4. A possible architecture to achieve artificial general intelligence

My full article goes into more detail on these 4 points. For all of those people who have been following my AI posts we raising an equity crowd-funded round, selling stock in ORBAI

Brent Oster

, former Deep Learning Solution Architect at NVIDIA (2008-2018)
Updated Thu

What is the future of artificial intelligence?

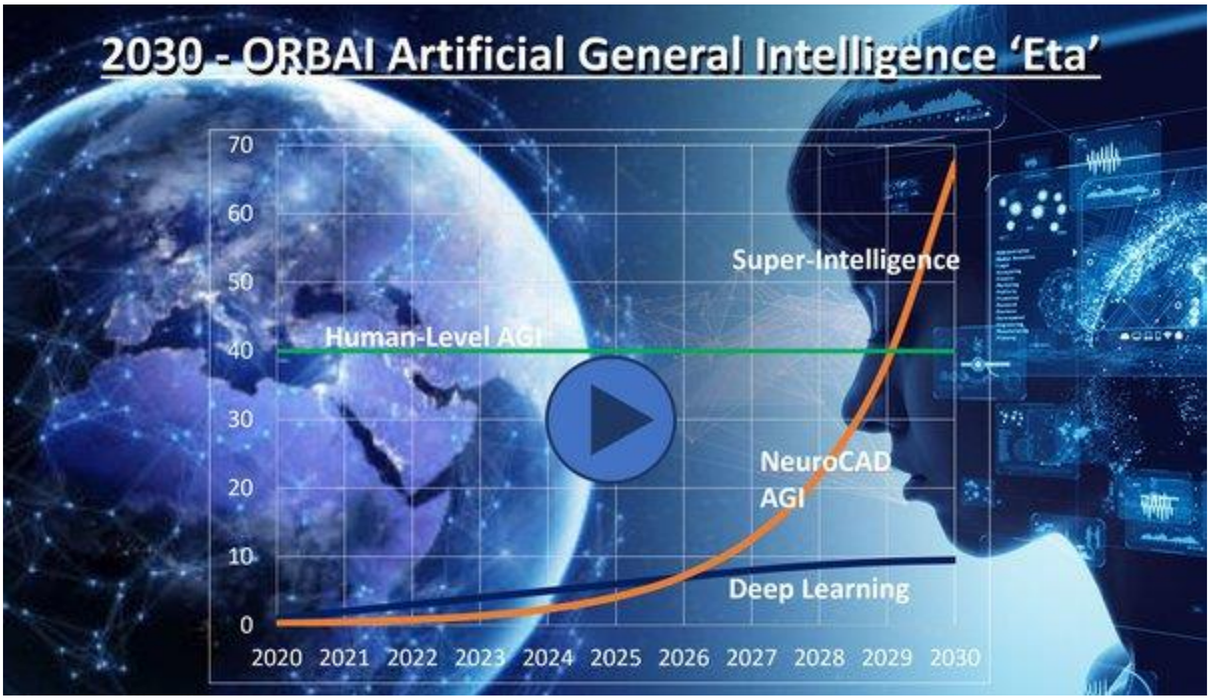
This article contains a complete description for building an AGI - covered under Provisional Patent US #63/138,058, filed Jan 15, 2021. And it is rock-solid, by a pro.

brent.oster@orbai.com

I worked for a decade at NVIDIA, as a CUDA Devtech Engineer and Solution Architect to research deep learning techniques, and present solutions to customers to solve their problems and to help implement those solutions. Now, for the past 3 years I have been working on what comes next after DNNs and Deep Learning. I will cover both, showing how it is very difficult to scale DNNs to AGI, and what a better approach would be.

Check it out the effort at http://www.startengine.com/orbai

What we usually think of as Artificial Intelligence (AI) today, when we see human-like robots and holograms in our fiction, talking and acting like real people and having human-level or even superhuman intelligence and capabilities, is actually called Artificial General Intelligence (AGI), and it does NOT exist anywhere on earth yet. What we actually have for AI today is much simpler and much more narrow Deep Learning (DL) that can only do some very specific tasks better than people. It has fundamental limitations that will not allow it to become AGI, so if that is our goal, we need to innovate and come up with better networks and better methods for shaping them into an artificial intelligence.



Lets do a deep dive into the science and tech of human intelligence and AI:

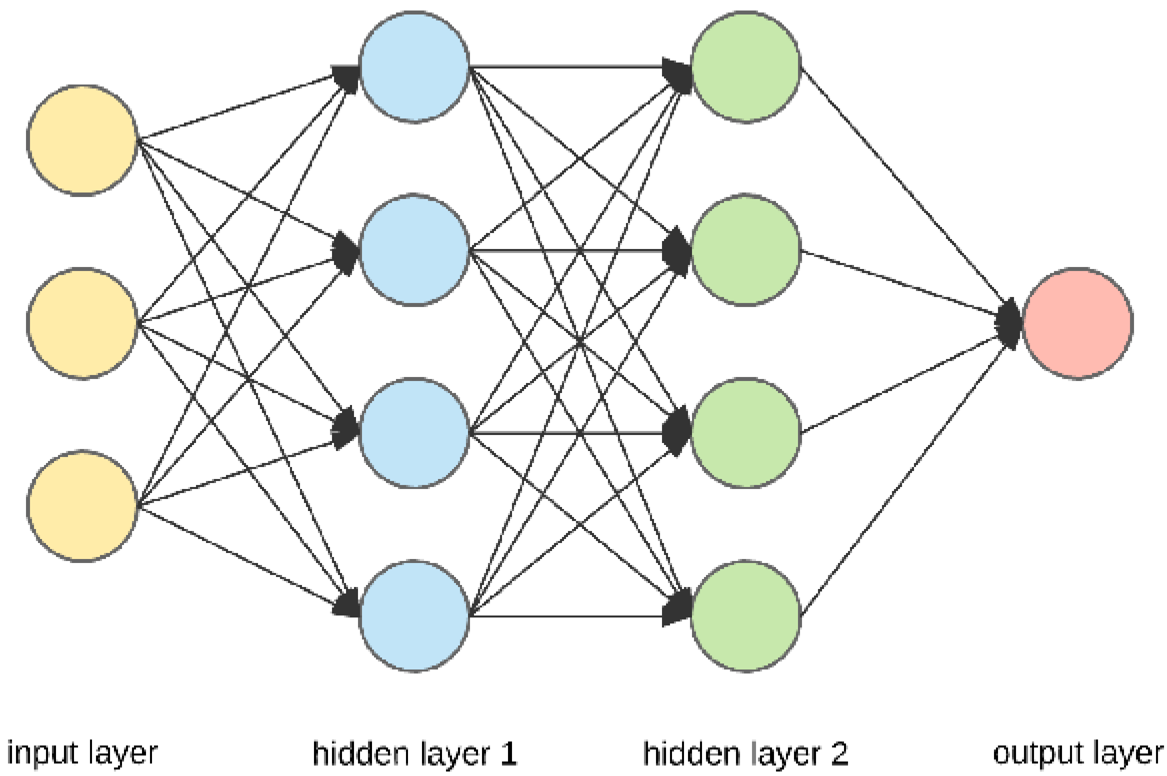
1. Where Deep Learning and Reinforcement Learning are today
2. What their limitations are - what can they do and not do?
3. The neuroscience of human intelligence
4. A possible architecture to achieve artificial general intelligence

1) Today's Deep Learning and Reinforcement Learning

Let me write down some extremely simplistic definitions of what we do have today, and then go on to explain what they are in more detail, where they fall short, and some steps towards creating more fully capable 'AI' with new architectures.

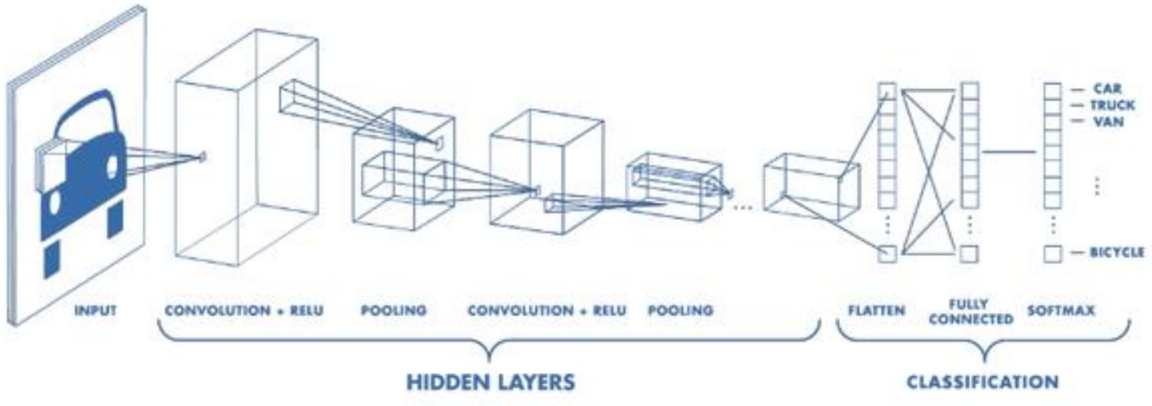
Machine Learning - Fitting functions to data, and using the functions to group it or predict things about future data. (Sorry, greatly oversimplified)

Deep Learning - Fitting functions to data as above, where those functions are layers of nodes that are connected (densely or otherwise) to the nodes before and after them, and the parameters being fitted are the weights of those connections.

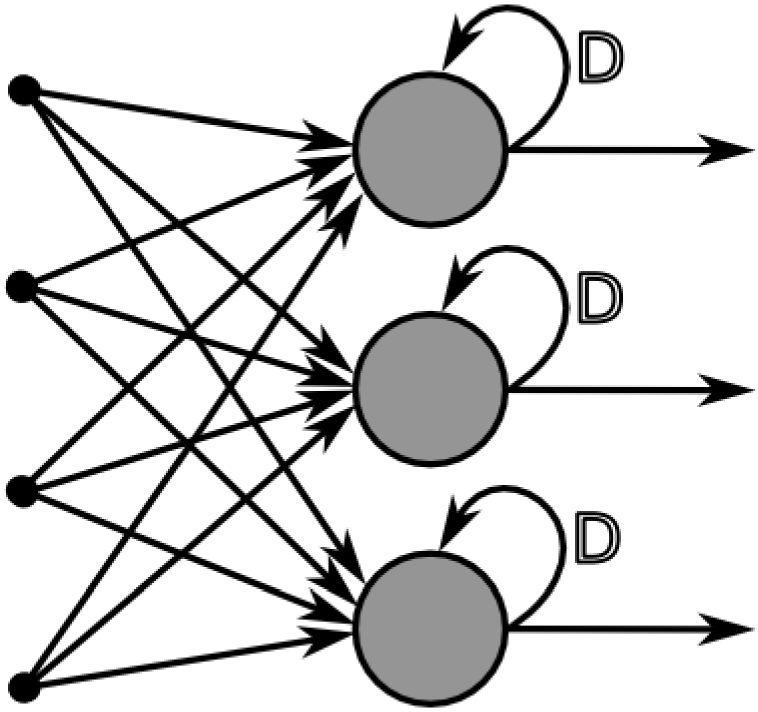


Deep Learning is what what usually gets called AI today, but is really just very elaborate pattern recognition and statistical fitting to data. The most common techniques / algorithms are Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Reinforcement Learning (RL).

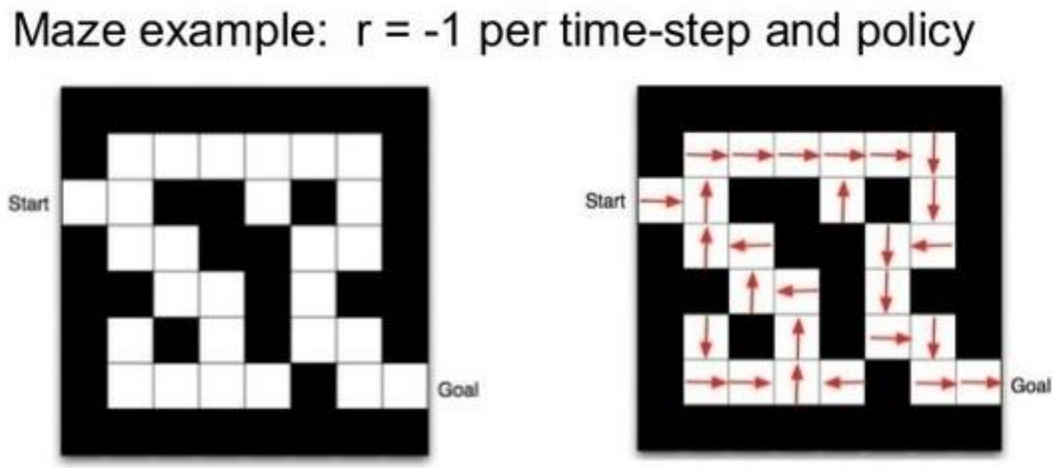
Convolutional Neural Networks (CNNs) have a hierarchical structure (which is usually 2D for images), where an image is sampled by (trained) convolution filters into a lower resolution map that represents the value of the convolution operation at each point. In images it goes from high-res pixels, to fine features (edges, circles,....) to coarse features (noses, eyes, lips, ... on faces), then to the fully connected layers that can identify what is in the image. The cool part of CNNs is that the convolutional filters are randomly initialized, then when you train the network, you are actually training the convolution filters. For decades, computer vision researchers had hand-crafted filters like this, but could never get them as effective as CNNs can get. Additionally, the output of a CNN can be an 2D map instead of a single value, giving us image segmentation. CNNs can also be used on many other types of 1D, 2D and even 3D data.



Recurrent Neural Networks (RNNs) work well for sequential or time series data. Basically each 'neural' node in an RNN is kind of a memory gate, often an LSTM or Long Short Term Memory cell. When these are linked up in layers of a neural net, these cells/nodes also have recurrent connections looping back into themselves and so tend to hold onto information that passes through them, retaining a ‘memory’ and allowing processing not only of current information, but past information in the network as well. As such, RNNs are good for time sequential operations like language processing or translation, as well as signal processing, Text To Speech, Speech To Text,...and so on.



Reinforcement Learning is a third main Machine Learning method, where you train a learning agent to solve a complex problem by simply taking the best actions given a state, with the probability of taking each action at each state defined by a policy. An example is running a maze, where the position of each cell is the ‘state’, the 4 possible directions to move are the actions, and the probability of moving each direction, at each cell (state) forms the policy.



[David Silver, Advanced Topics: RL]

By repeatedly running through the states and possible actions and rewarding the sequence of actions that gave a good result (by increasing the probabilities of those actions in the policy), and penalizing the actions that gave a negative result (by decreasing the probabilities of those actions in the policy). In time you arrive at an optimal policy, which has the highest probability of a successful outcome. Usually while training, you discount the penalties/rewards for actions further back in time.

In our maze example, this means allowing an agent to go through the maze, choosing a direction to move from each cell by using the probabilities in the policy, and when it reaches a dead-end, penalizing the series of choices that got it there by reducing the probability of moving that direction from each cell again. If the agent finds the exit, we go back and reward the choices that got it there by increasing probabilities of moving that direction from each cell. In time the agent learns the fastest way through the maze to the exit, or the optimal policy. Variations of Reinforcement learning are at the core of the AlphaGo AI and the Atari Video Game playing AI.

2) Limitations of Deep Learning

But all these methods just find a statistical fit to large amounts of data using simple models - DNNs find a narrow fit of outputs to inputs that does not usually extrapolate outside the training data set, and may learn incorrect features to identify objects with that do not work for novel inputs. Reinforcement learning finds a pattern that works for the specific problem (as we all did vs 1980s Atari games), but not beyond it. With today's ML and deep learning, the problem is there is no true perception, memory, prediction, cognition, or complex planning involved. There is no actual intelligence in today's AI.

The reason that speech interfaces in devices are so limited and awkward to talk to is that existing DL is very narrow for speech-to-text and natural language processing and can only train to learn specific phrases and map them to specific intents, actions, or answers, giving only a skeleton of language comprehension with some clever scripting but not conversational speech capability.

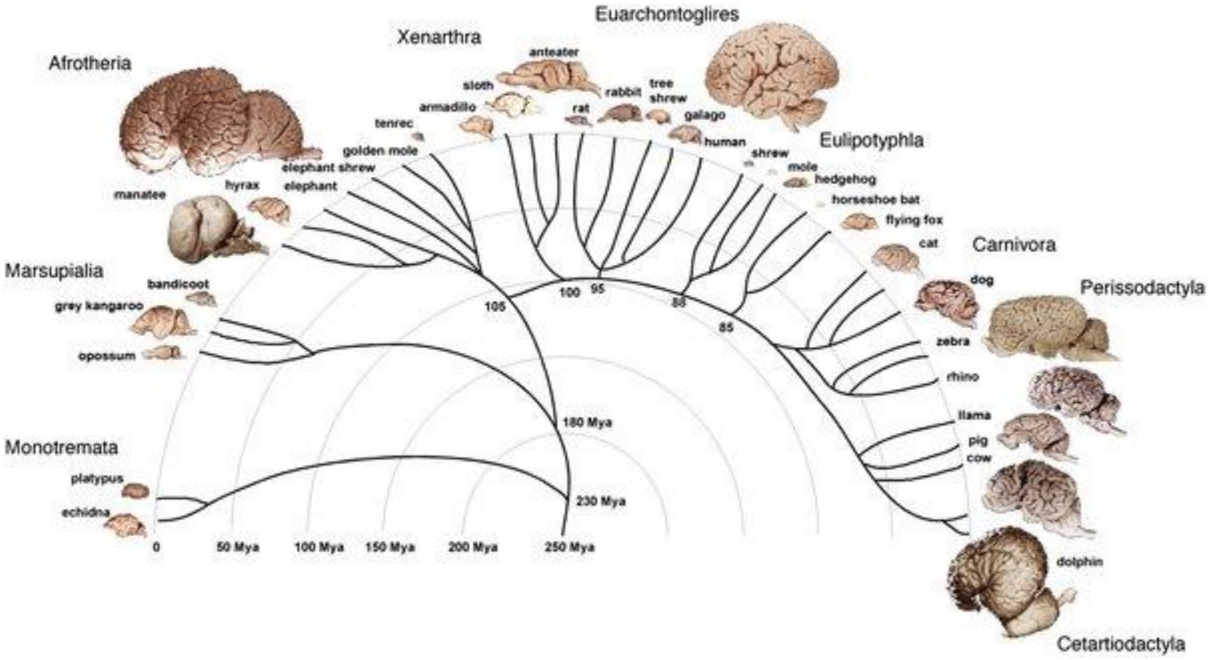
As another use case, we do not have useful home robots today, ones that can freely navigate our homes, avoid obstacles, pets, and kids, and do useful things like cleaning, doing laundry, even cooking. This is because the narrow slices of deep learning available for vision, planning, and control of motors, arms, and manipulators cannot take in all this varied input, be able to train on all the possible combinations of states, let alone plan, and do useful tasks from them.

3) The Neuroscience of Human Intelligence

To go beyond where we are today with AI, to pass the threshold of human intelligence, and create an artificial general intelligence requires an AI to have the ability to see, hear, and experience its environment or input data streams. It needs to be able to learn that environment, to organize its memory and store abstracted concepts with distributed features so it can model that environment, and the objects, people, and events in it.

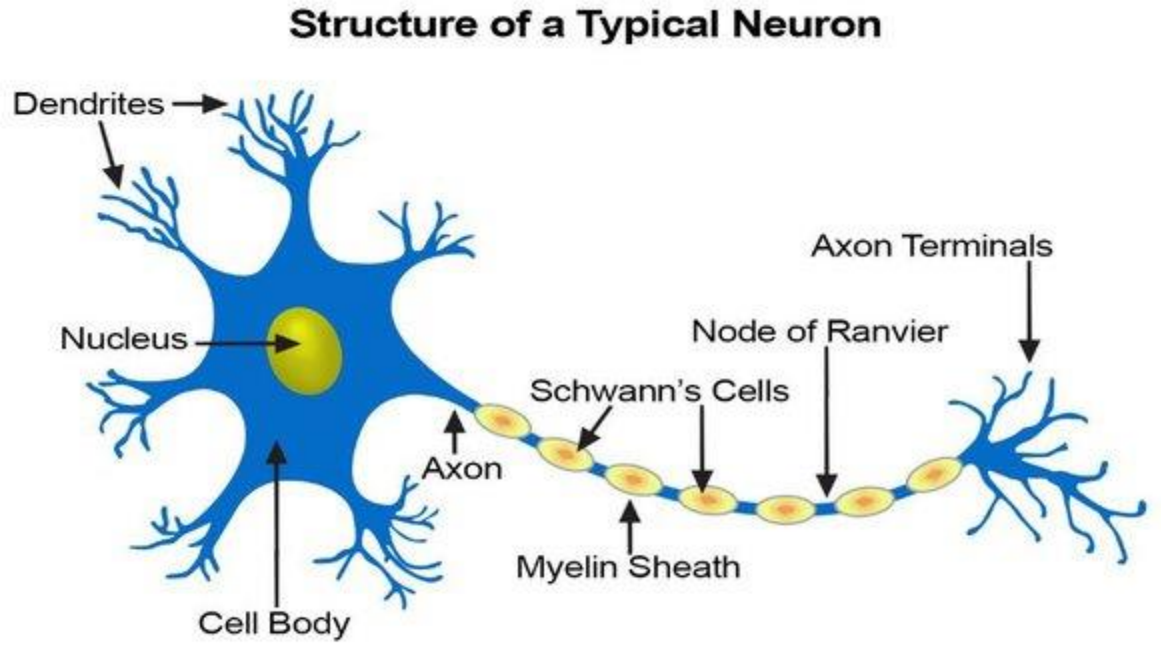
It probably needs to be able speak conversationally and interact verbally like a human, and be able to understand the experiences, events, and concepts behind the words and sentences of language and how they connect, so it can compose language at a human level. It needs to be able solve all the problems that a human can, using flexible memory recall, analogy, metaphor, imagination, intuition, logic and deduction from sparse information. It needs to be capable at the tasks and jobs humans are and to express the results in human language in order to be able to do those tasks and professions as well as or better than a human.

The human brain underwent a very complicated evolution starting sometime since 1 billion years ago from the first multi-cellular animals with a couple neurons, through the Cambrian explosion where eyes, ears and other sensory systems, motor systems, and with them more neurons and intelligence exploded in an arms race (along with armor, teeth, and claws). Evolution of brains then followed the needs of fish, reptiles, dinosaurs, mammals, and finally up the hominids lineage about 5-10 million years ago.



Much of the older parts of the human brain were evolved for the previous 500 million years of violence and competition, not the last thousands of years of human civilization, so in many ways out brain is maladapted for our modern life in the information age, and not very efficient at many of the tasks we use it for in advanced professions like law, medicine, finance, and administration. A synthetic brain, focused on doing these tasks optimally can probably end up doing them much better, so we do not seek to re-create the biological human brain, but to imbue ours with the core functionality that makes the human brain so flexible, adaptable and powerful, then augment that with CS database and computing capabilities to take it far beyond human.

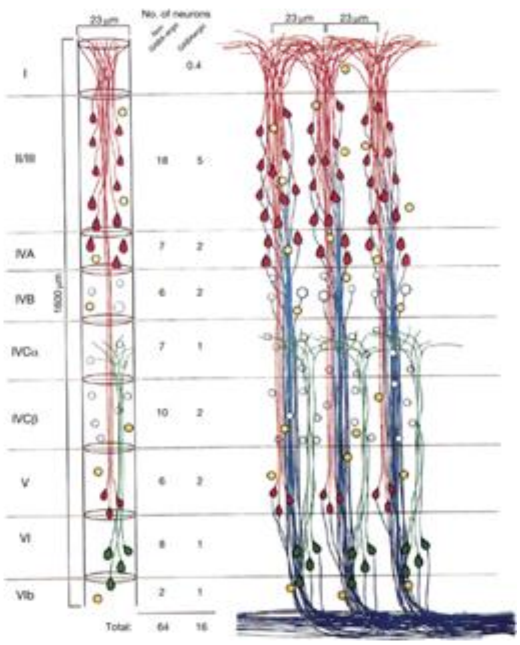
Because deep learning DNNs are so limited in function and can only train to do narrow tasks with pre-formatted and labelled data, we need better neurons and neural networks with temporal spatial processing and dynamic learning. The human brain is a very sophisticated bio-chemical-electrical computer with around 100 billion neurons and 100 trillion connections (and synapses) between them. I will describe two decades of neuroscience in the next two paragraphs, but here are two good videos about the biological Neuron and Synapse from '2-Minute Neuroscience' on YouTube that will also help.



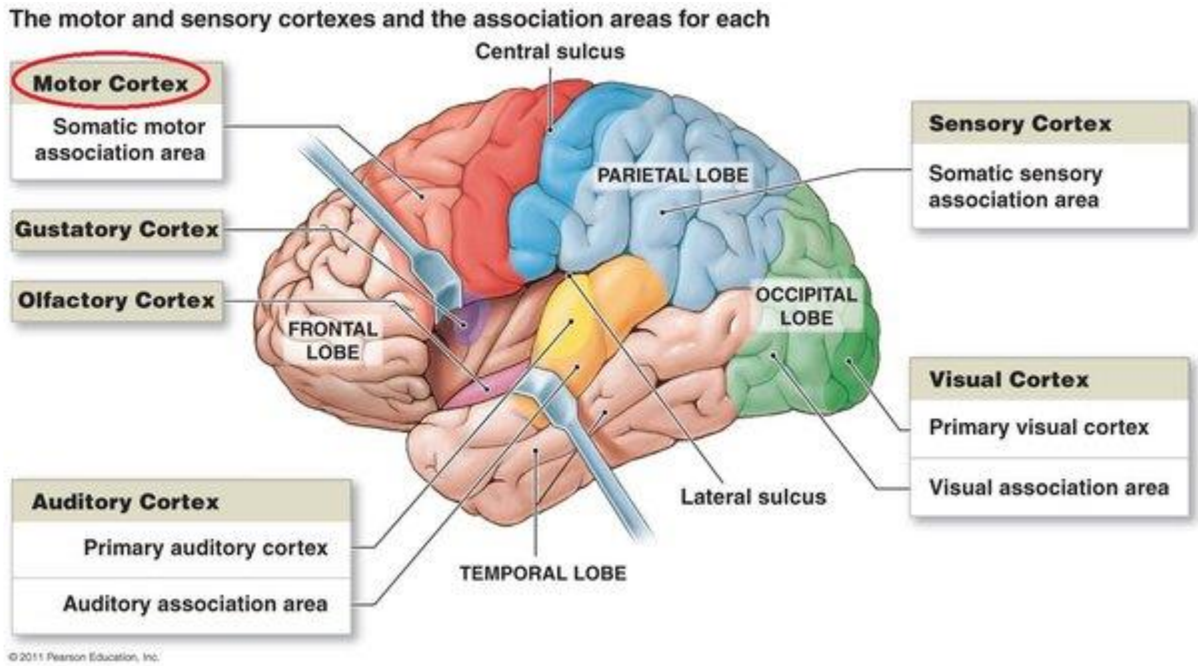
Each neuron takes in spikes of electrical charge from its dendrites and performs a very complicated integration in time and space, resulting in the charge accumulating in the neuron and (once exceeding an action potential) causing the neuron to fire spikes of electricity out along its axon, moving in time and space as that axon branches and re-amplifies the signal, carrying it to thousands of synapses, where it is absorbed by each synapse. This process causes neurotransmitters to be emitted into the synaptic cleft, where they are chemically integrated (with ambient neurochemistry contributing). These neurotransmitters migrate across the cleft to the post-synaptic side, where their accumulation in various receptors eventually cause the post-synaptic side to fire a spike down along the dendrite to the next neuron. When two connected neurons fire sequentially within a certain time, the synapse between them becomes more sensitive or potentiated, and then fires more easily. We call this Hebbian learning, which is constantly occurring as we move around and interact with our environment.

Neurons are organized into micro-structures in the brain. The cortex of the brain consists of a 'sheet' of neurons about 2–3mm thick, about 7–9 layers of neurons deep. This sheet is wrapped and folded around the brain to form the cerebral cortex, the main structure for processing information, including sensory inputs, motor control, language understanding, speaking, cognition, planning, and logic.

This sheet is divided laterally into about 1 million cortical micro-columns of 100,000 neurons each. These micro-columns run from the bottom to the top layer of the cortex, and have more connections up and down than laterally, and most of their lateral connections do not extend beyond more than a few columns. So these little columns form discrete 'computational' units that are essentially remote from their neighbors, but interconnected with the whole brain through complex internal structures and connections.



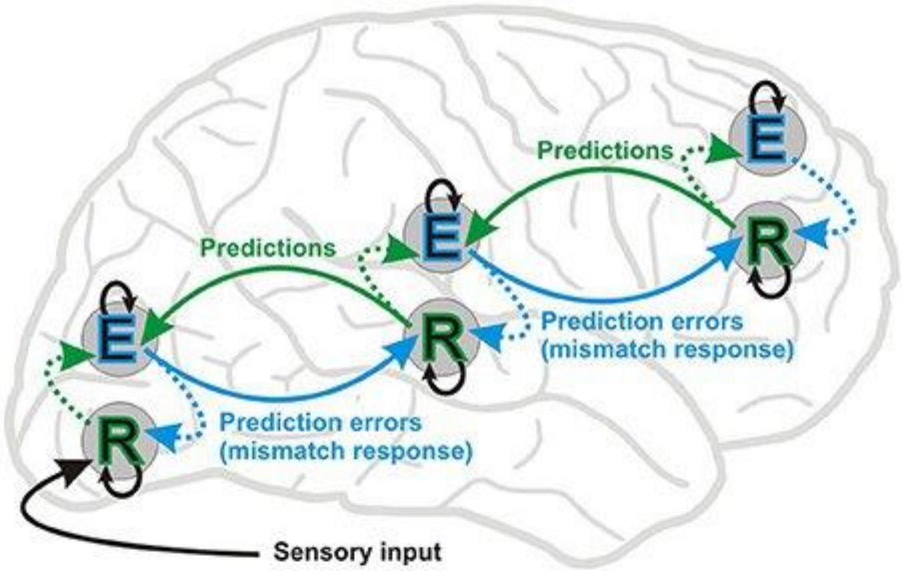
How they function exactly is the subject of great debate in neuroscience, but their existence and what we know about them shows that these units are similar, replicated across the cortex, and specialized to certain functions in certain areas. This is useful info in making an AGI, cause if we can build and train an artificial cortical column and connect thousand of them into a sheet in the right way - we have a start at an artificial brain - maybe.



The brain's cortices consisting of these specialized cortical columns evolved to have networks with very sophisticated space and time signal processing, including feedback loops and bidirectional networks, so visual input is processed into abstractions or 'thoughts' by one directional network, and then those thoughts are processed back out to a recreation of the expected visual representation by another, complementary network in the opposite direction, and they are fed back into each other throughout. Miguel

Nicolelis is one of the top neuroscientists to measure and study this bidirectionality of the sensory cortices.

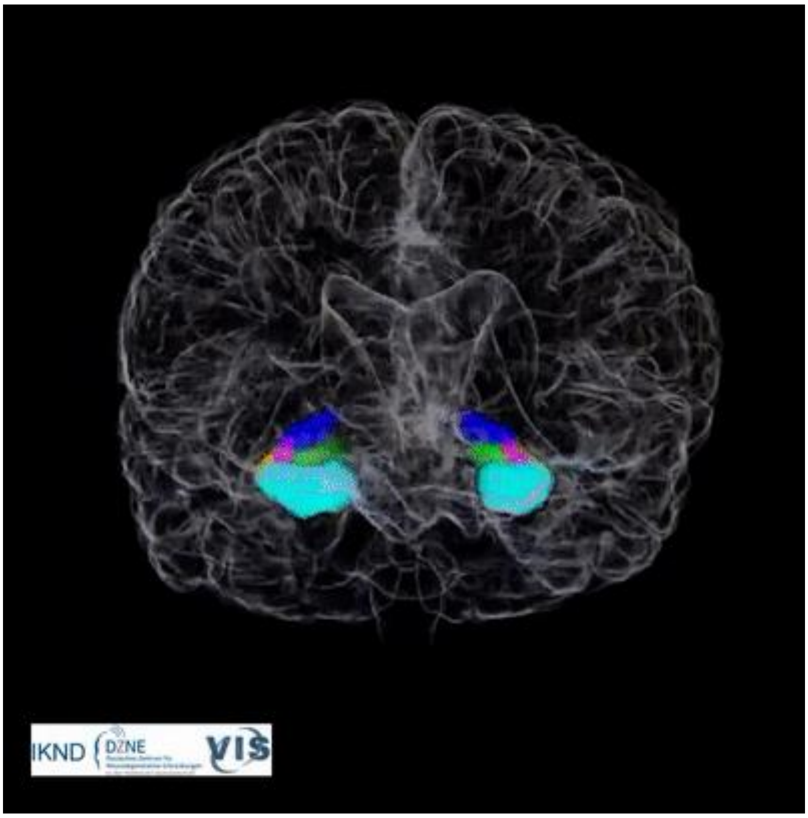
For an example, picture a 'fire truck' with your eyes closed and you will see the feedback network of your visual cortex at work, allowing you to visualize the 'thought' of a fire truck into an image of one. You could probably even draw it if you wanted. Try looking at clouds, and you will see shapes that your brain is feeding back to your vision as thoughts of what to look for and to see. Visualize shapes and objects in a dark room when you are sleepy, and you will be able to make them take form, with your eyes open



These feedback loops not only allow us to selectively focus our senses, but also train our sensory cortices to encode the information from our senses into compact 'thoughts' or Engrams that are stored in the hippocampus short term memory. Each sensory cortex has the ability to decode them again and to provide a perceptual filter by comparing what we are seeing to what we expect to see, so our visual cortex can focus on what we are looking for and screen the rest out as we stated in the previous paragraph.

The frontal and pre-frontal cortex are thought to have tighter, more specialized feedback loops that can store state (short-term memory), operate on it, and perform logic and planning at the macroscale. All our cortices (and brain) work together and can learn associatively and store long-term memories by Hebbian learning, with the hippocampus being a central controller for memory, planning, and prediction.

Human long-term memory is less well known. We do know that it is non-local, as injuries to specific areas of the brain don't remove specific memories, even a hemispherectomy which removes half the brain. Rather, any given memory appears to be distributed through the brain, stored like a hologram or fractal, spread out over a wide area with thin slices everywhere. We know that global injury to the brain, like Alzheimer's - causes a progressive global loss of all memories, which all degrade together, but no structure in the brain seems to contribute more to this long-term memory loss than another.



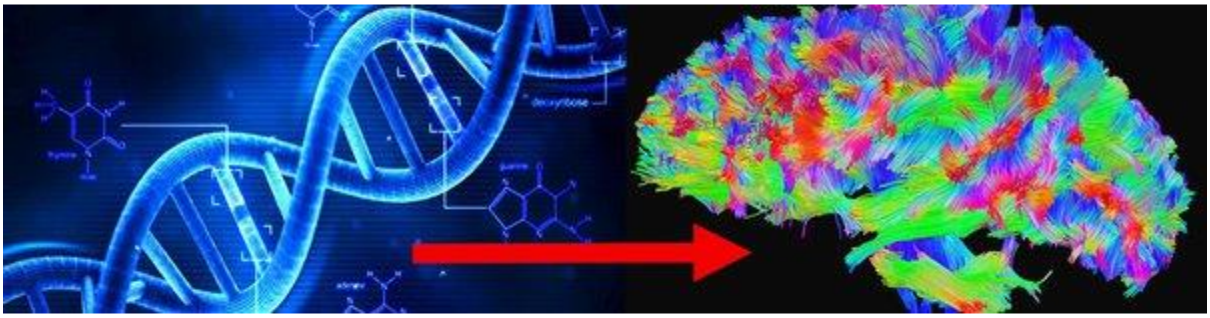
However, specific injury to the hippocampus causes the inability to transfer memory between short term and long-term memory. Coincidentally, it also causes the inability to predict and plan and other cognitive deficits, showing that all these processes are similar. This area is the specialty of prominent memory neuroscientist, Eleanor Maguire, who states that the reason for memory in the brain is not to recall an accurate record of the past, but to predict the future and reconstruct the past from the scenes and events we experienced, using the same stored information and process in the brain that we use to look into the future to predict what will happen, or to plan what to do. Therefore the underlying storage of human memories must be structured in an abstracted representation in such a way that memories can be reconstructed from some for the purpose at hand, be it reconstructing the past, predicting the future, planning, or imagining stories and narratives – all hallmarks of human intelligence.

But how does our brain do more than just record all our experiences? How do we predict things we have never seen, or plan for events we've never experienced, or say things we've never heard? If all our experiences are train tracks, how do we figure out what is in between? This is why humans dream, it fills in the spaces between our experiences and builds models of our world from them.

Besides moving memories from short-term episodic memory (hippocampus in humans), in dreaming - the brain forms connections between potentially related memories, which we experience as REM dreams - somewhat fantastic, sometimes nonsensical sequences of events, but self-consistent and ordered. For every sleep study done, researchers could identify with 90% accuracy which dream reports had been randomized by the researchers after recording them. There is an order and meaning to our REM dreams. Once the dreaming process has formed these connections, we now have fictional 'memories' that help us model our world and we can use to predict future events or plan contingencies. Take a look at When Brains Dream for amazing neuroscience research in this area by Antonio Zadra and Robert Stickgold.

Replicating all of the brain's capabilities seems daunting when seen through the tools of deep learning – image recognition, vision, speech, natural language understanding, written composition, solving mazes, playing games, planning, problem solving, creativity, imagination, because deep learning is using single-purpose components that cannot generalize. Each of the DNN/RNN tools is a one-of, a specialization for a specific task, that cannot generalize, and there is no way we can specialize and combine them all to accomplish all these tasks.

But, the human brain is simpler, more elegant, using fewer, more powerful, general-purpose building blocks – the biological neuron, and connecting them by using the instructions of a mere 8000 genes, so nature has, through a billion years of evolution, come up with an elegant and easy to specify an architecture for the brain and its neural network structures that is able to solve all the problems we met with during this evolution. We are going to start by just copying as much about the human brain's formation process and functionality as we can, then using evolution to solve the harder design problems.



So now we know more about the human brain, and how the neurons and neural networks in it are completely different from the DNNs that deep learning is using, and how much more sophisticated our simulated neurons, neural networks, cortices and neural networks would have to be to even begin attempting to build something on par with, or superior to the human brain.

At the end of this article is the video talk about neuroscience and AGI that I submitted to NVIDIA GTC 2021 Conference (40 min) that builds on this article. Here is a 2 min teaser:

4) How can we build an Artificial General Intelligence?

If you ask any person that works in AGI how it can be done, you will probably get a different answer from each of them, because nobody has done it successfully and there are only theories. I've put 4 years into this design, and a couple decades of research before that, and I'm a practical engineer as well as a scientist, and it has been pretty well scrutinized on Quora for 3 years since I started writing it. I will give you the quick version of the design before we dive into the details:

We will use Spiking Neural Network Autoencoders for sensory encoding and output decoding (they are bidirectional), and develop our AGI core that is capable of taking in this encoded data, sifting it into a basis set pf features, training a model on how those features change in time during an experienced narrative of events, how they relate to each other via the features, and then use that model to simulate or dream fictional narratives to fill in the blanks in the model. This helps take our AGI core well beyond deep learning, by allowing it to build a full, feature-rich and accurate model of its world, and augment that model by dreaming, like humans do. This is the only way to have something like a speech AI learn sentences it has never heard, or for a robot to do tasks it has never experienced - structured dreaming to fill in the blanks, reinforced (or degraded) by later experiences.

Now the details of the nuts and bolts:

1) AGI Methods. We propose methods and processes for running computer simulations of Artificial General Intelligence (AGI) that is able to operate on general inputs and outputs, that do not have to be specifically formatted, nor labelled by humans and can consist of any alpha-numerical data stream, 1D, 2D, and 3D temporal-spatial inputs, and others. The AGI is capable of doing general operations on them that emulate human intelligence, such as interpolation, extrapolation, prediction, planning, estimation, and using guessing, and intuition to solve problems with sparse data. These methods will not require specific coding, but that can rather be learned unsupervised from the data by the AGI and its internal components using spiking neural networks. Using these methods, the AGI would reduce the external data to an internal format that computers can understand, be able to do math, linear algebra, supercomputing, and use databases, yet still plan, predict, estimate, and dream like a human, then be able to convert the results back to human understandable form.

a) The input system would learn to autoencode any time domain input, including alphanumeric streams, 1D, 2D, and 3D inputs using the SNN autoencoders in (6), to encode them into compact engram streams, and write these engram streams to short-term memory.

b) After a predetermined duration of short-term memory has been recorded, it is batch processed by cutting it into segments with by convolving it with time-domain functions like Gaussian or unit step function centered at time t and advancing t by dt each time such that the segments have a predetermined overlap.

c) Processing the engram segments using a hierarchical sorting architecture of autoencoders and PCA operations, then convolving engram segments with the vector basis sets at the leaf nodes to transform them to a set of basis coefficients

d) Storing basis coefficients encoded from inputs (or those computed internally) - into temporal narratives in memory

e) Doing mathematical, neural net, and other computations between sets of basis coefficients

f) Using neural net constructs such as predictors, solvers, and dreamers to do operations on narratives of basis coordinates.

g) Transforming the internal basis coefficient narrative representations back to engrams using the autoencoder hierarchy, then back to real-world outputs using the autoencoders.

h) Creating outputs to drive time domain outputs for actuators, language, and other outputs using a ROS / inhibitor network scheme.

With these developments, we will take the first steps towards AGI that can perceive the real world, reduce those perceptions to an internal format that computers can understand, yet still plan, think and dream like a human, then convert the results back to human understandable form, and even converse fluently using human language, enabling online interfaces and services that can interact much more like a person.

See NeuroCAD Utility Patent US # 16/437,838 + PTC, filed June 11, 2019 for details of 2-8. Summarized here for reference:

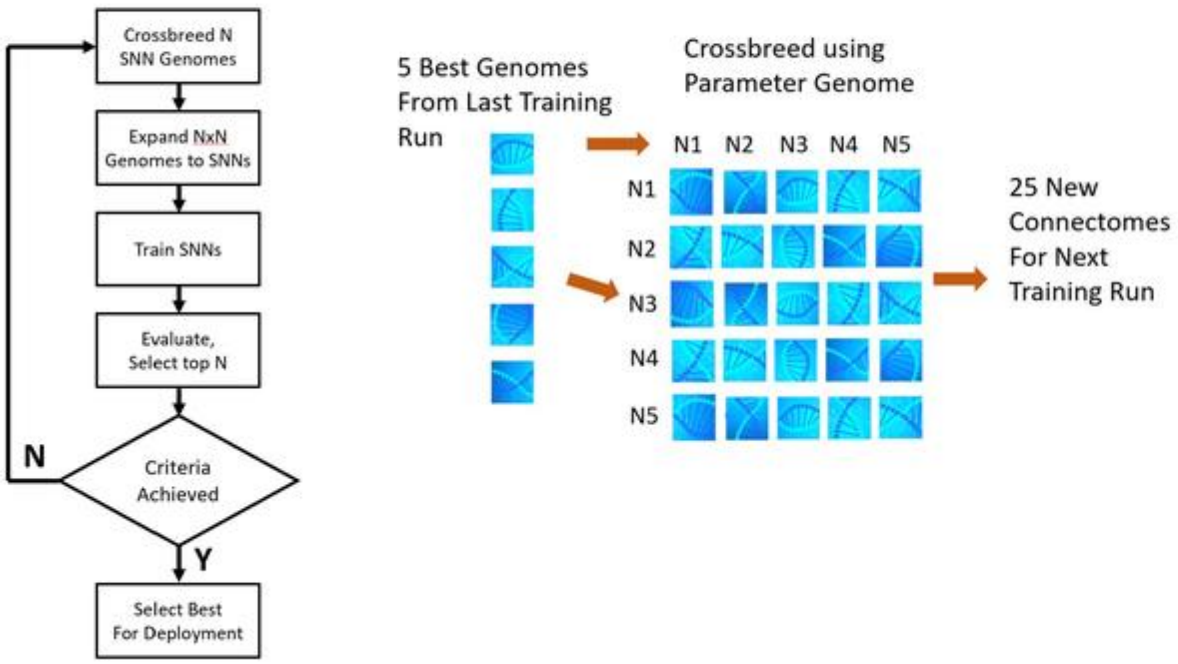
2) Methods for designing and evolving spiking neural networks (SNN), which consist of computer simulated (assumed from here forward) artificial neurons connected by axons and dendrites with a synapse between each axon and dendrite. Spikes of current are transmitted from the neuron, out along the axon, then are absorbed at the synapse and processed. The synapse may then, depending on the computation, transmit a spike out along the dendrite and to the neuron. Each time the neurons on either side of a synapse fire in sequence, that synapse is 'strengthened' and the likelihood of transmitting a spike increases. The spikes move in time and space and this temporal circuitry is key to the spiking neural network's functionality and utility. Axons and dendrites can branch, with the outgoing spikes splitting and being amplified as they go out along the branches of the axon. Likewise, signals can combine as dendrites merge before entering the neuron. The incoming signals to a neuron can be excitatory or inhibitory, adding or subtracting charge from the neurons. Neurons can then integrate the incoming signals or differentiate them, or other operations, then emit spikes based on their internal computation.

3) To model these in a computer, we use mathematical models for the neurons and synapses that integrate, differentiate, or otherwise compute the contribution of the incoming charges and compute an output based on the mathematical model over time. We move the discrete spikes of current along the axons and dendrites at a constant speed, checking for when they enter a synapse or neuron. Sensory inputs are translated to spikes of current into sensory neurons.

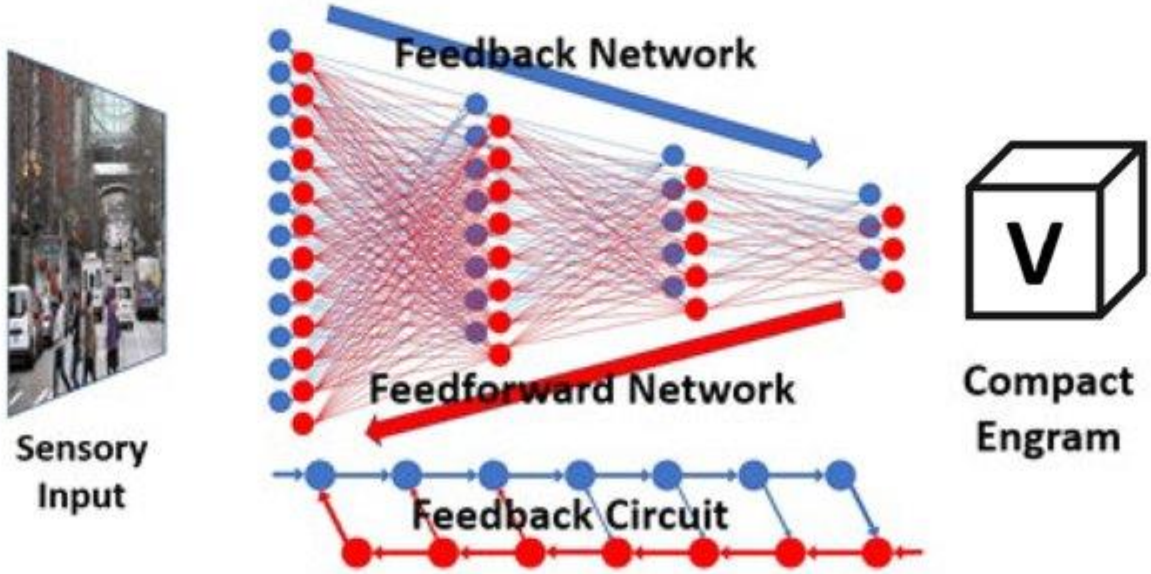
4) Each neural network connectome, consisting of neurons, connected by axons, synapses, and dendrites, all connected in a neural network is represented by a compact genome, which is a small chunk of data consisting of numbers and alphanumeric data, which compactly represents the topology of the network, the number and size of layers, the types of neurons in them, and the statistical distribution of the connections from neurons in one layer or topological region to another. These genomes (G) always expand

deterministically to the same connectome (C), and they interpolate smoothly, such that a genome G, that is interpolated to be between G0 (expands to C0) and G1 (expands to C1), when expanded, will result in a neural net connectome, C that is between C0 and C1 in properties. This ‘smoothness’ property is necessary for the genetic algorithms to converge.

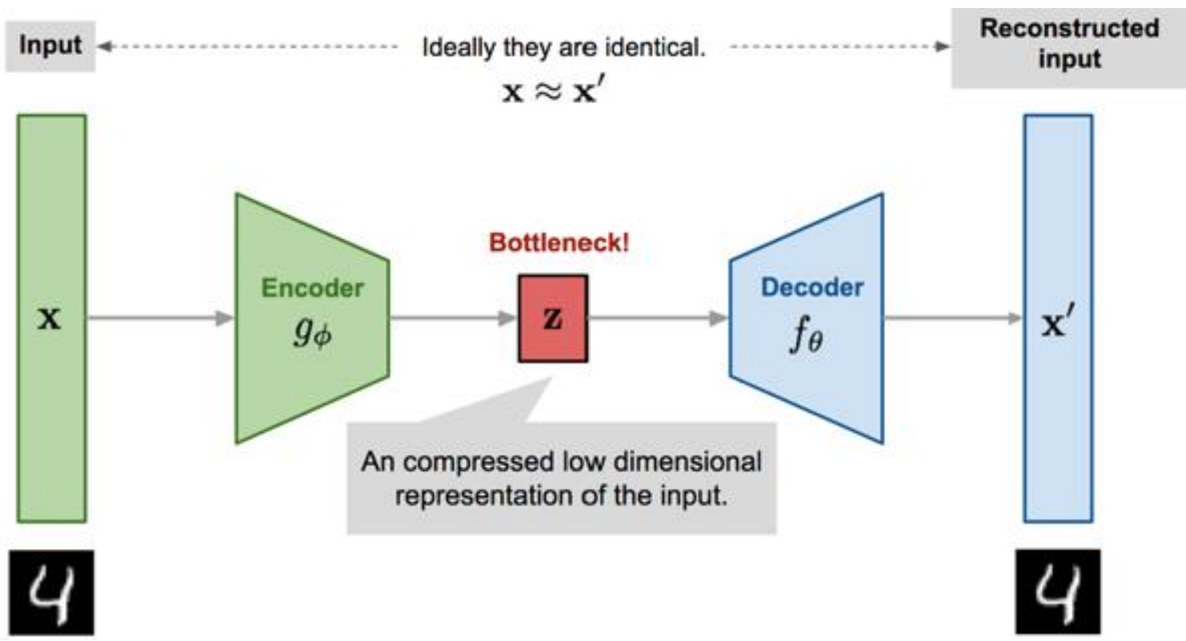
5) We use evolutionary methods for designing and evolving spiking neural networks (SNN), using genetic algorithms operating on the compact genomes - that are expanded to the full neural networks to be trained, then evaluated according to specified criteria, to see if they will be crossbred for the next generation, repeating till a SNN that meets the specified criteria is evolved.



6) We use these methods to evolve SNNs specialized for different functions, including bidirectional interleaved autoencoders that consist of layers of neurons that alternate between ‘even’ layers containing mostly forward connections, skipping one layer to the next even layer and ‘odd layers containing mostly reverse connections, with those connections skipping a layer to the next odd layer, with some crossover in the connectivity, with final connectivity determined by genetic algorithms. Input comes in at layer 0, is encoded through the encoder into the bottom layer(s), where it is forced into a constrained bottleneck, and then decoded back through the autoencoder to layer 1, which is fed back into the even layers to generate a training feedback loop. The exact connections between layers, and feedback is determined by evolution via the genetic algorithms to find the configuration with optimal performance, with the selection criteria including encode/decode quality, latency, and encoded size) with the encoding and decoding method and encoded format learned at runtime in training and evolution

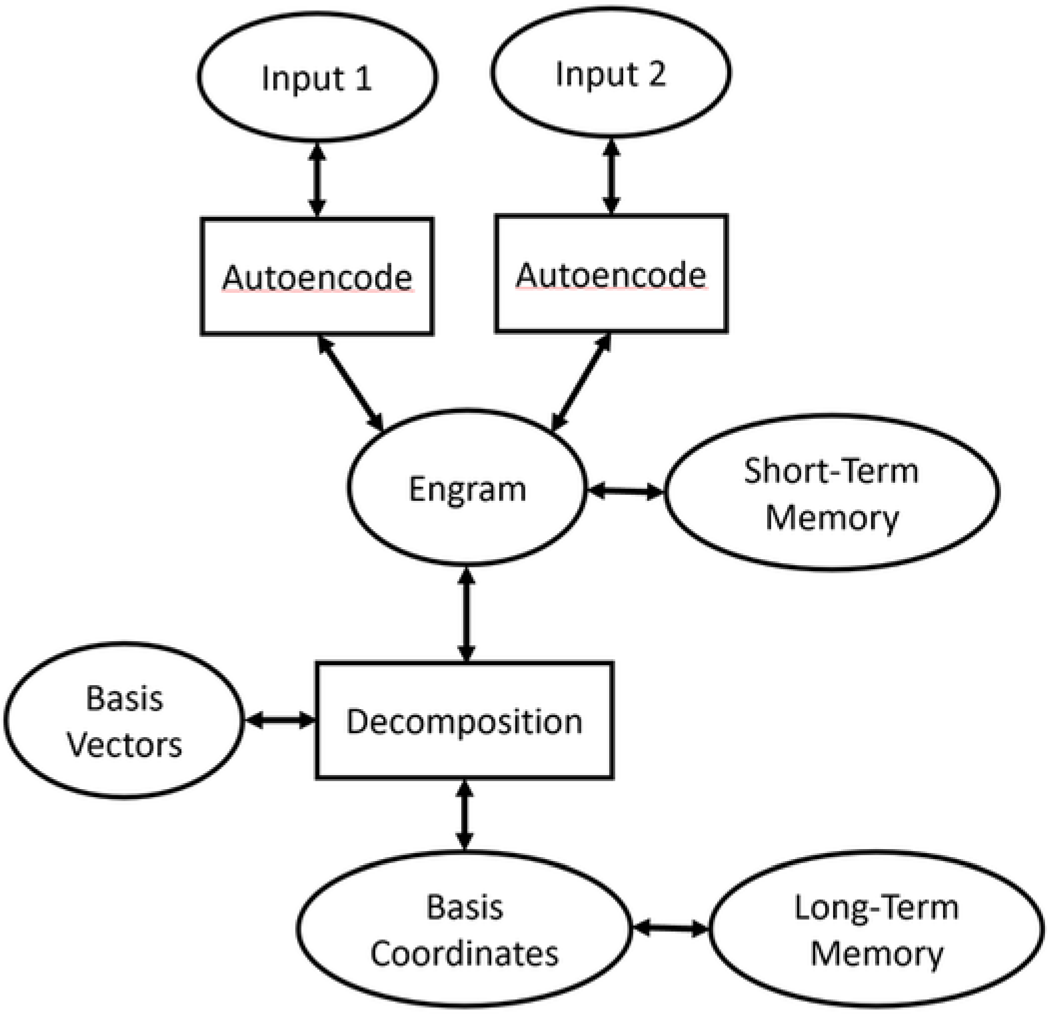


8) We put the data through our SNN autoencoder and in the mid-section, through a constriction or narrowing during the autoencoding process and enlarge it again and train the autoencoder to reproduce the original data, by doing so, the data is compressed at the constriction, but in a way that the entire autoencoder circuitry stores all the common features of the entire data set it has encoded to date (a set of basis vectors), and the output at the constriction is the set of basis coordinates referencing the basis vectors internal to the autoencoder. We take the output from the area or volume of constriction for each input and we record this into memory in time as an 'engram stream' or encoded memory stream that is the analogous to human short-term memory in the hippocampus.



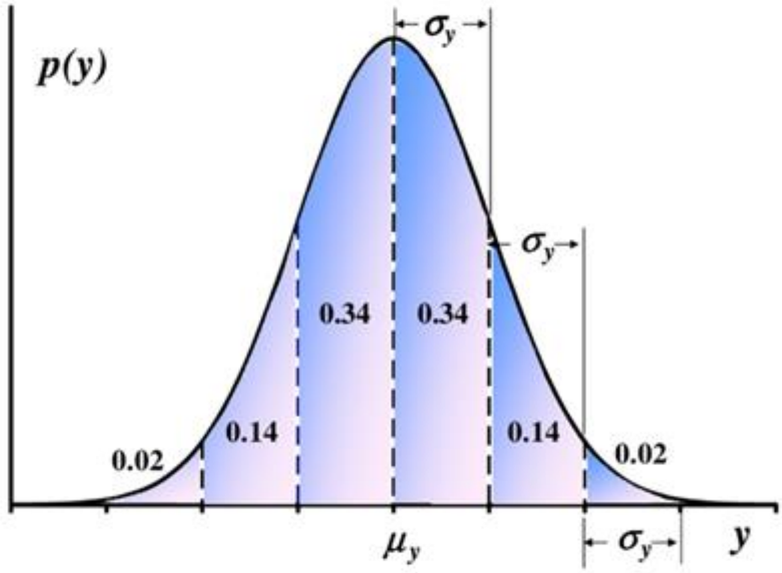
9) We autoencode these input streams into engram streams in a process where there can be one or more such inputs generating multiple engram streams. Inputs may be encoded into one engram stream each, both into the same engram stream with interleaving,

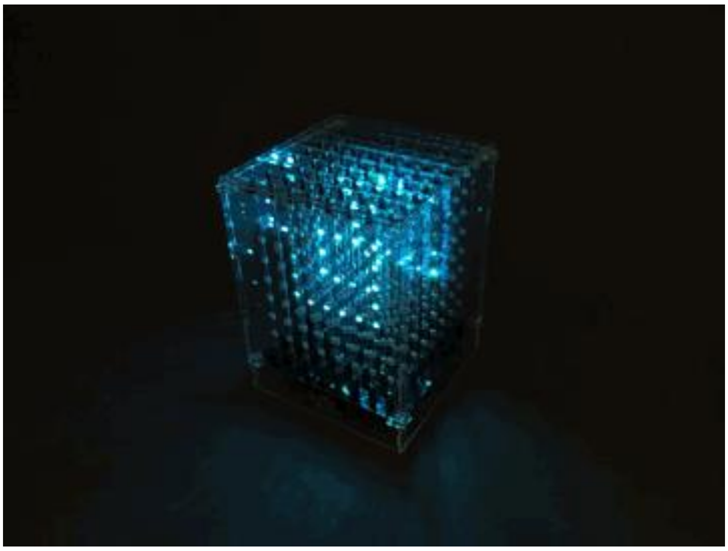
convolution, addition or other operations, or a hybrid where they are encoded into their own engram streams and both encoded into a hybrid engram stream with interleaving, convolution, addition or other operations.



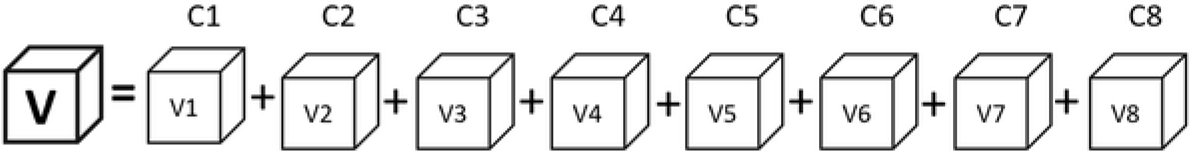
10) Our encoded engrams will be in the form of these low-dimensional constricted volumes with the time dimension they were recorded in - termed an engram stream, representing a compressed record of the inputs in time, that can be reconstituted or decoded back into the original input. We have a starting point with being able to reduce the sensory inputs for our AGI to this compressed format, but they are still unwieldly, and we cannot do useful operations on them, except for volumetric convolutions to test them against other engrams. We need a better basis set.

11) By convolving the engram stream with a function (such as a step function or Gaussian) at multiple points in the time domain, we can cut it into intervals or segments, at different time intervals (t0 + dt * j) with j spanning the engram stream to be segmented. By choosing the convolution function parameters, and the value of dt, this will determine the overlap in time of the volumes. By doing this, we reduce the engram stream to a set of unit volumes in four dimensions, a little ‘swirl’ of reality in each volume.

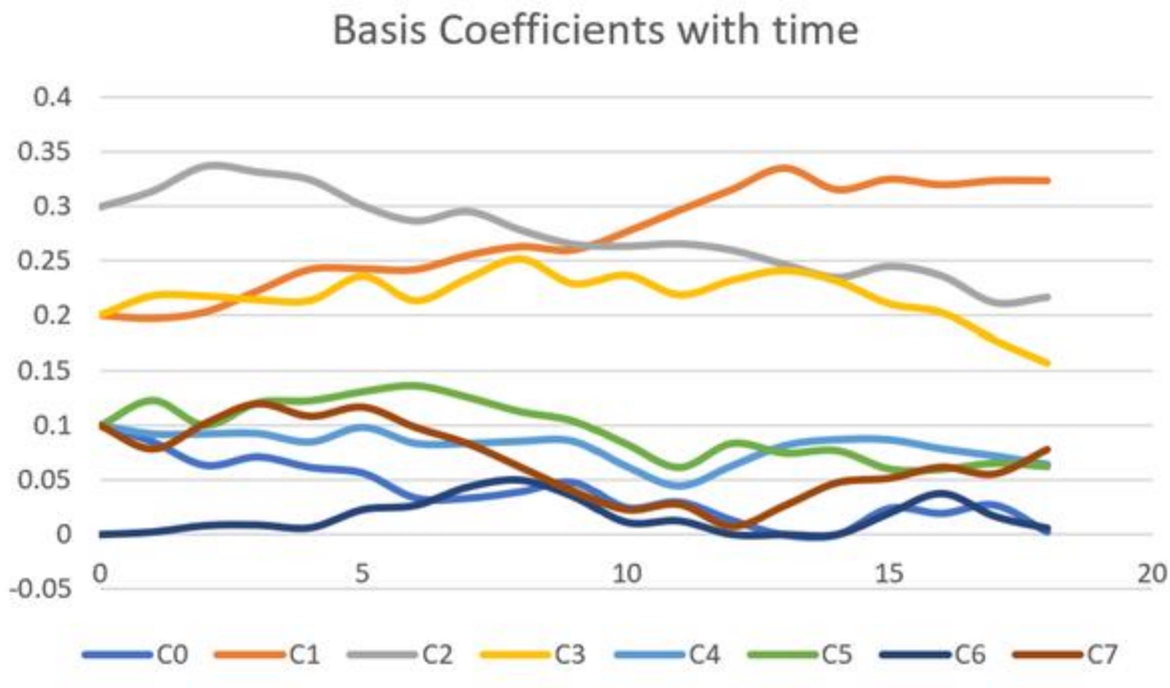




12) These volumes are a more compact and useful format that either the raw inputs or the longer engram streams. We can create an orthogonal basis set of such volumes - essentially a set of orthogonal basis vectors that spans the space of previously experienced engram segments, then any engram segment can be decomposed into a linear combination of the vectors of this basis set by convolution with each vector of the basis set in a reversible process. Then when the basis vectors are each multiplied by the corresponding basis coordinates and they are linearly combined, the original engram segment is reconstituted.

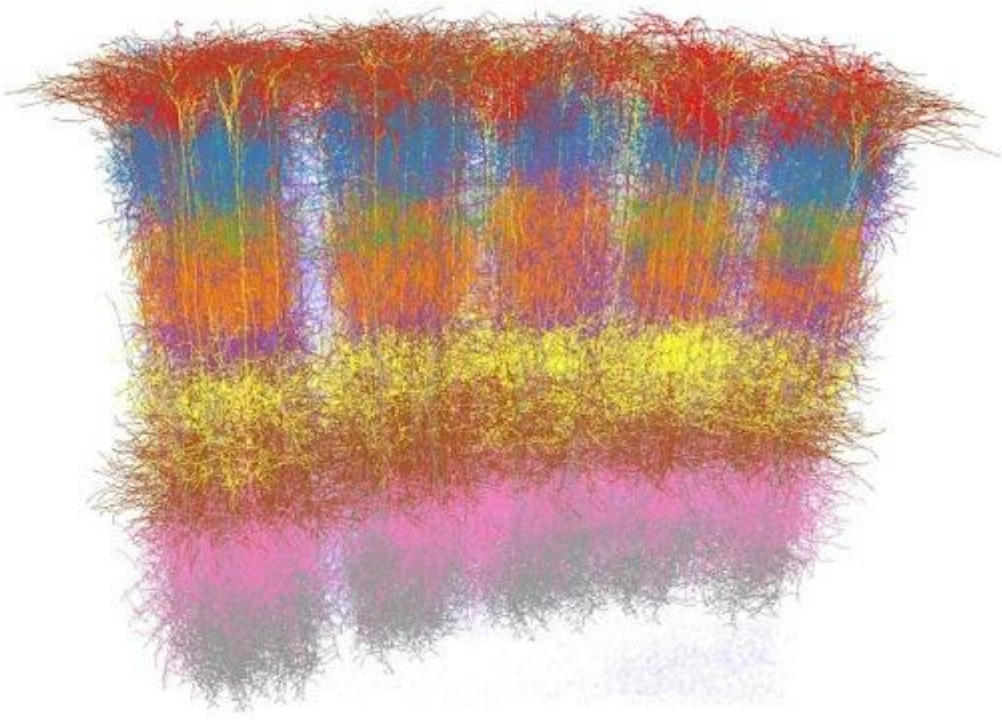


We will term this set of basis coordinates in time a 'narrative' in 'long-term' memory, essentially a stream of numbers that are much more useful for doing calculations on.

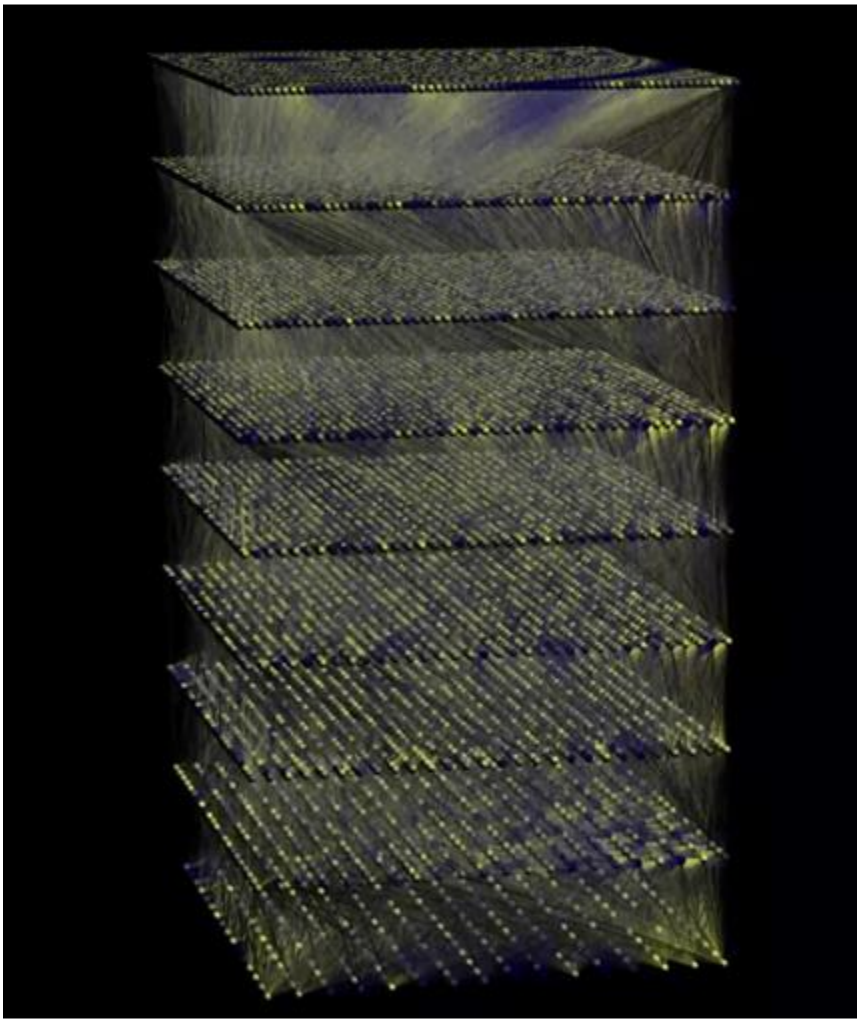


13) We still have not determined how we will compute our orthogonal basis set. Usual methods like Gram-Schmidt would be too costly, because for our system, we need a basis set than can potentially span all of visible or audio reality. The size of the basis set, and the mechanism for computing it would be immense and computationally prohibitive with these methods. We need to be able to work in parallel.

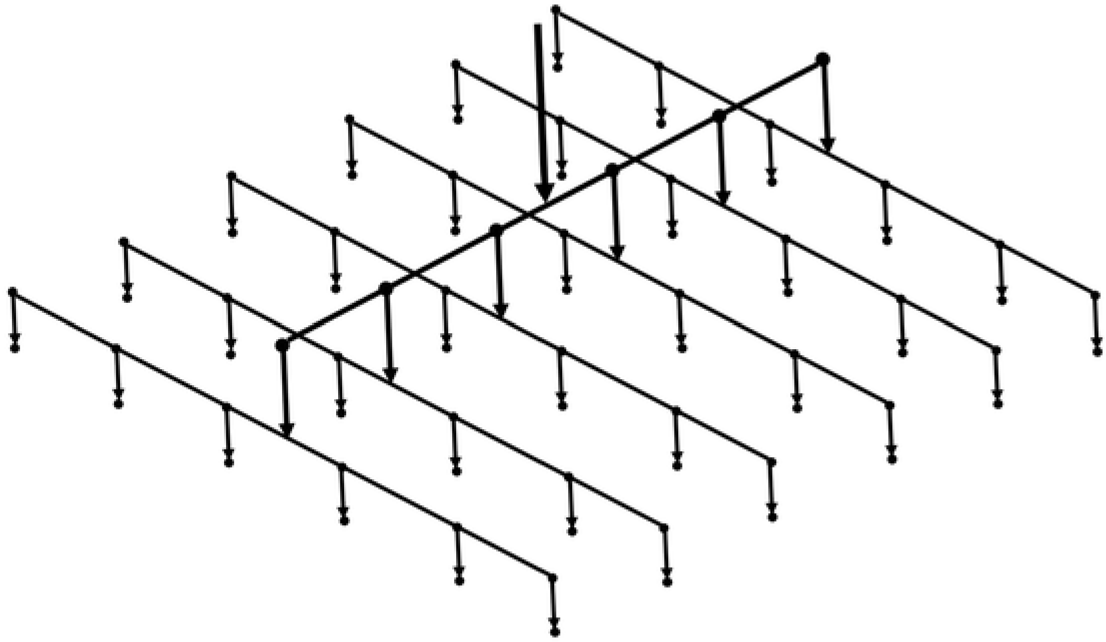
14) Fortunately, we already have an analogy of such a system in the human brain. The human brain has analogous structures. The cerebral cortex is a sheet about 4mm thick wrapped and folded around the outside of the brain, that consists of cortical micro columns, each containing about 100,000 neurons, 7 neuron layers deep. The cortical columns are depicted like this:



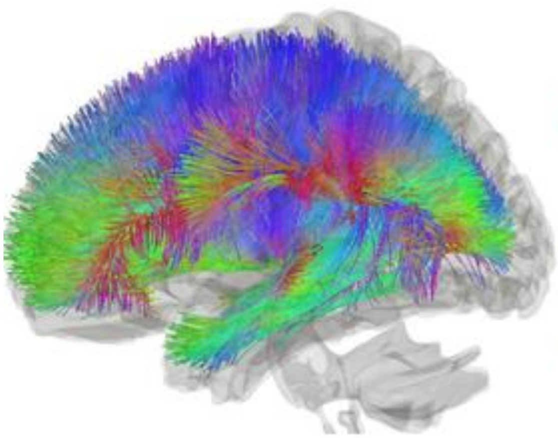
15) These cortical columns look a lot like our autoencoders, which take in inputs like vision, audio,... and encode them to a compact engram, storing the common information about all the inputs it has seen in the autoencoder circuitry and the unique information about each input in the engram. Done in a hierarchy of autoencoders, in a cascade, this would gradually separate the similarity out of the engrams as they pass down the hierarchy, and leave us with orthogonal basis engrams at the leaf nodes.

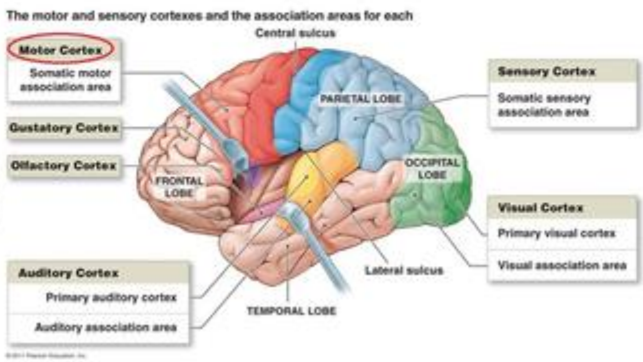


16) More formally: a method for dynamically creating an orthogonal basis set of engram vectors (12), by submitting a batch of engrams and spreading them along an axis, sorted by a specific feature, forming clusters of engrams along the axis. These clusters are then each encoded by a specific autoencoder, removing their common feature, and then the resulting engrams are spread out on new axes sorted by new features. This is done recursively till one engram remains in each cluster, giving us a set of leaf nodes that constitute an orthogonal basis set of vectors. New engram batches can be added later to create new clusters, autoencoders, axes, and basis vectors, making it dynamic and able to learn.



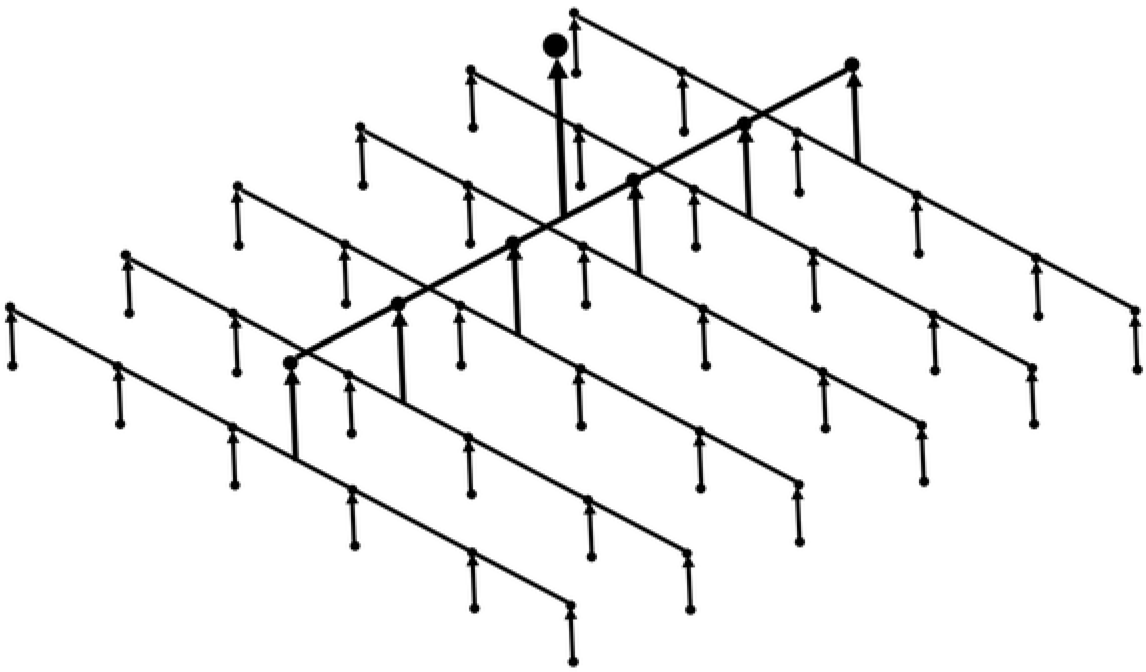
17) There is a biological analogy for this hierarchy, the thalamocortical radiations, a neural structure that branches out like a bush from the thalamus (the main input / output hub of the brain for the senses, vision, audio and motor outputs) with the finest branches terminating at the cerebral cortex, feeding input from the senses to the cortical columns. The cortical columns of the cerebral cortex are analogous to our terminal layer of autoencoders, whose purpose is to store the orthogonal basis vectors for reality and do computations against them, including computing basis coordinates from input engrams. Each section of the cortex is specialized for a specific type of input (visual, auditory, olfactory,...) or output (motor, speech), and our design will have a separate hierarchy and autoencoder basis set for each mode of input, to generate basis coordinates for that input / output mode.



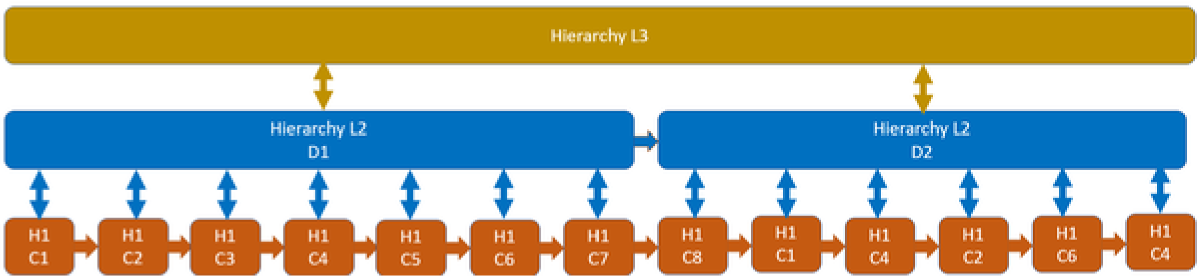


18) To determine the basis coefficients for an input engram segment, the engram segment is passed through the process specified in (16), but singularly, splitting it off into engram segments that traverse the correct portion of the hierarchy, until convolution with the basis vectors at the leaf nodes determines the basis coefficients for that engram. This process can be used in reverse, multiplying the basis coefficients with the basis vectors at the leaves and passing them back up through the hierarchy of autoencoders to

reconstruct the original engram. Basically we have a system that can deconstruct reality to numerical basis coordinate vectors and back. That will come in handy later for computations.



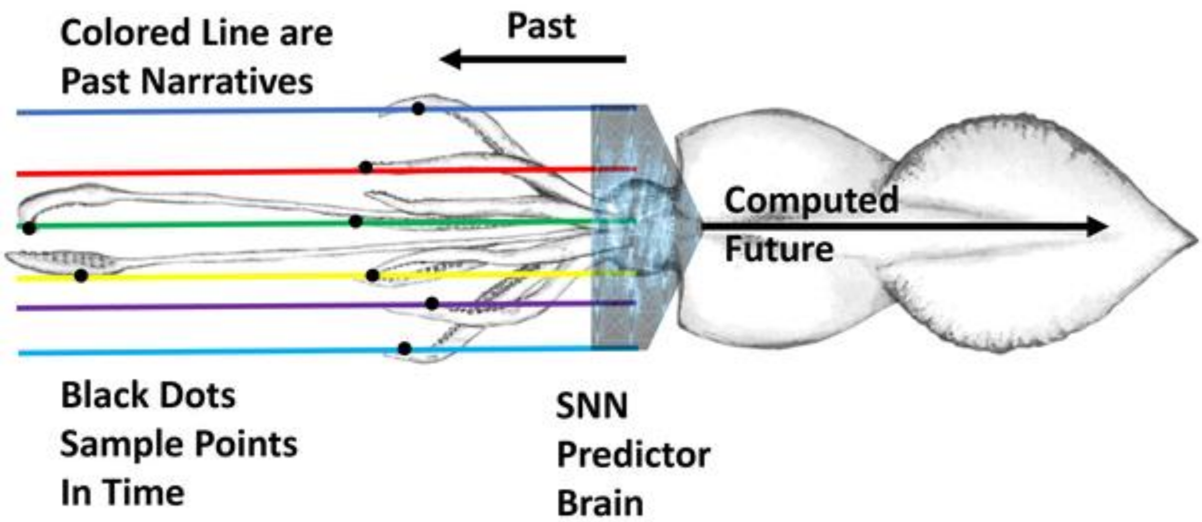
19) A method for organizing the narratives of basis coefficients (12) into segments, which can then be arranged hierarchically and/or connected to other segments and/or hierarchies to form composite structures in memory. Narrative segments and hierarchy structures can be collapsed and instanced, with multiple points in multiple hierarchies referring to the same instance of the segment (or hierarchy structure). These repeated child segments or hierarchies need only be stored once in memory, relationships with them and other data need only be computed once, reducing space and computational requirements.



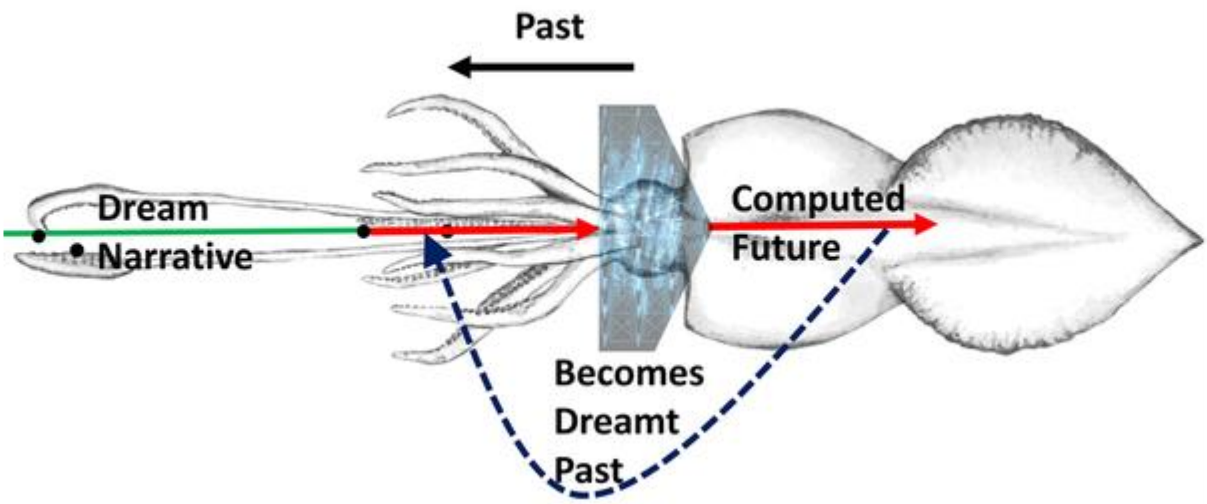
20) A method for sampling basis coordinates from a plurality of encoded narratives (N), from a plurality of intervals in time (t-j,... t-2, t-1, t0, t1, t2,..tk)N, with these coordinates as input to a spiking neural network (SNN), as defined in (5) which computes the values of one or more sets of basis coordinates for one or more points in time (t-j,... t-2, t-1, t0, t1, t2,..tk) in a computed narrative as output. Before use, this SNN is trained on known inputs and outputs, and also evolved to perform its operations optimally. This allows us to do arbitrary computing using temporal narratives of basis sets (and the reality they represent) as our inputs and outputs.

21) A method for sampling basis coordinates from a plurality of intervals in time (t-j,... t-2, t-1, t0, t1, t2,..tk)N from a plurality of encoded narratives (N), with these coordinates as input to a spiking neural network, as defined in (5) which computes the values of one or more sets of basis coordinates for one or more points in time (t0, t1, t2,..tk) in a future narrative as output. We will term this method a “predictor model”. Before use, this SNN predictor is trained and evolved as per (5) on known inputs and outputs, and also evolved optimally to predict.

Concept diagram of predictor model, sampling past narratives, and SNN brain computing the future narrative.

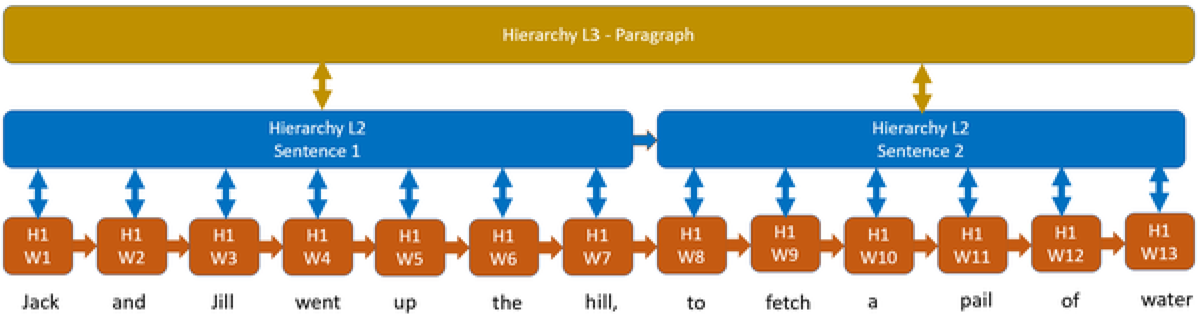


22) A method for simulating memory narratives or ‘dreaming’ by allowing the predictor (21) to start its prediction inputs on existing memories, then move forward in time, detaching from the narrative to compute its future predictions based on using input from its just-generated predicted memories, creating a fictional or dream narrative (shaped by its model of reality) in the memory narrative behind it in time. Optionally, it can ‘attach’ to an existing narrative and detach to dream, multiple times. This is repeated to create dream narratives that form a web connecting experienced narratives, to augment them.

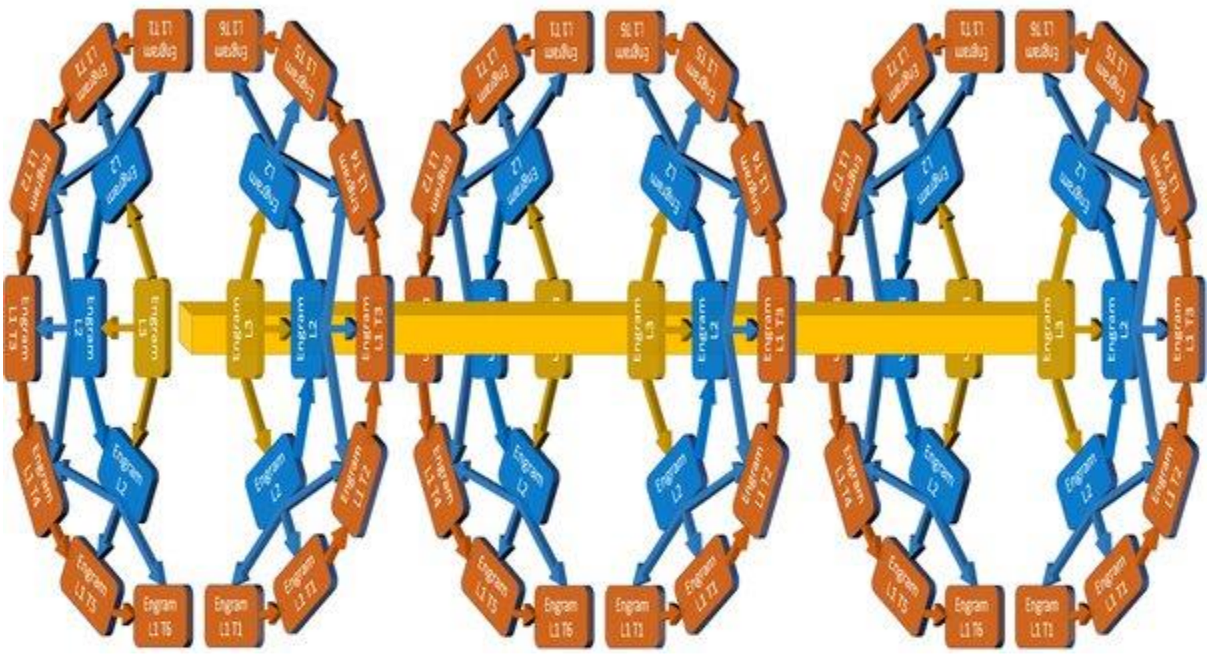


23) A method of continuously evaluating the dreamed memories as they are laid down and as they are traversed later, to decide if they should be attenuated or amplified, depending on their conformance to real memories and to the predictor model, or by reconstructing them into their corresponding engrams, and/or output data format to be evaluated.

24) A method of organizing language narratives structurally as in (19) for human language, such that each character for written language forms a basis vector, instanced in the language narrative by a basis coordinate. Spaces delineate segments consisting of words, punctuation delineates a segment (a hierarchy spanning words) of sentences, and CR characters delineate a segment (hierarchy of sentences) defining paragraphs. Similar organization is done for spoken language by having phonemes as the basis vectors, referred to by a basis coordinate, pauses delineating words, and longer pauses delineating sentences and paragraphs. Symbolic languages will be similarly structured in narratives and organized into hierarchy according to their written and spoken structure.

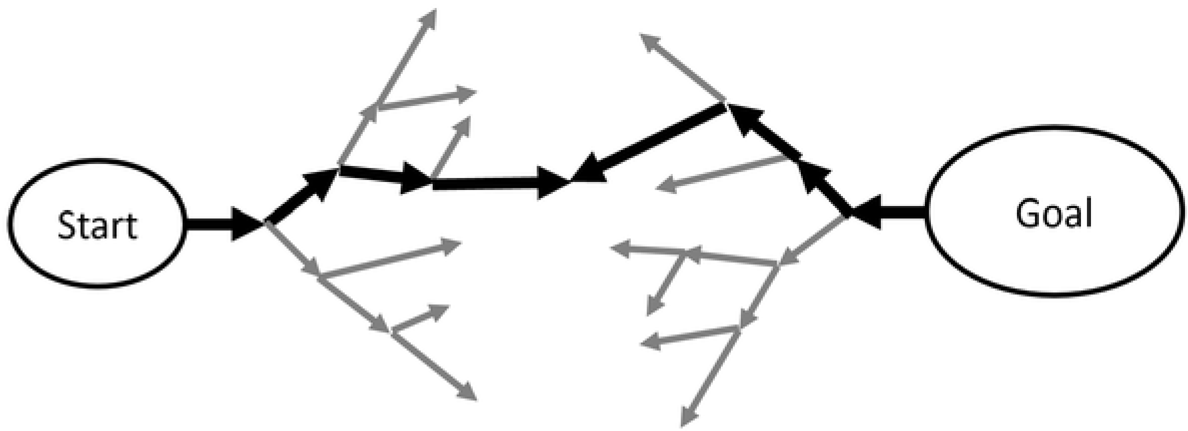


25) A method of connecting narratives (24) derived from different input types, such as visual and language, at coordinates where they are temporally, spatially or conceptually related, such that processing of one type of narrative can reference the related information in the other type of narrative as input or output in the processing. The method would also be able to connect between the higher levels of the hierarchies (9) to allow more abstract operations between the different levels. Making language the backbone for our AGI’s memory and cognition by connecting words to form references to visual objects and sounds, sentences to form abstractions for sequences of visual and audio events, and paragraphs to form abstractions scenarios and stories in memory, with each word, sentence and paragraph connected to one or more memory.

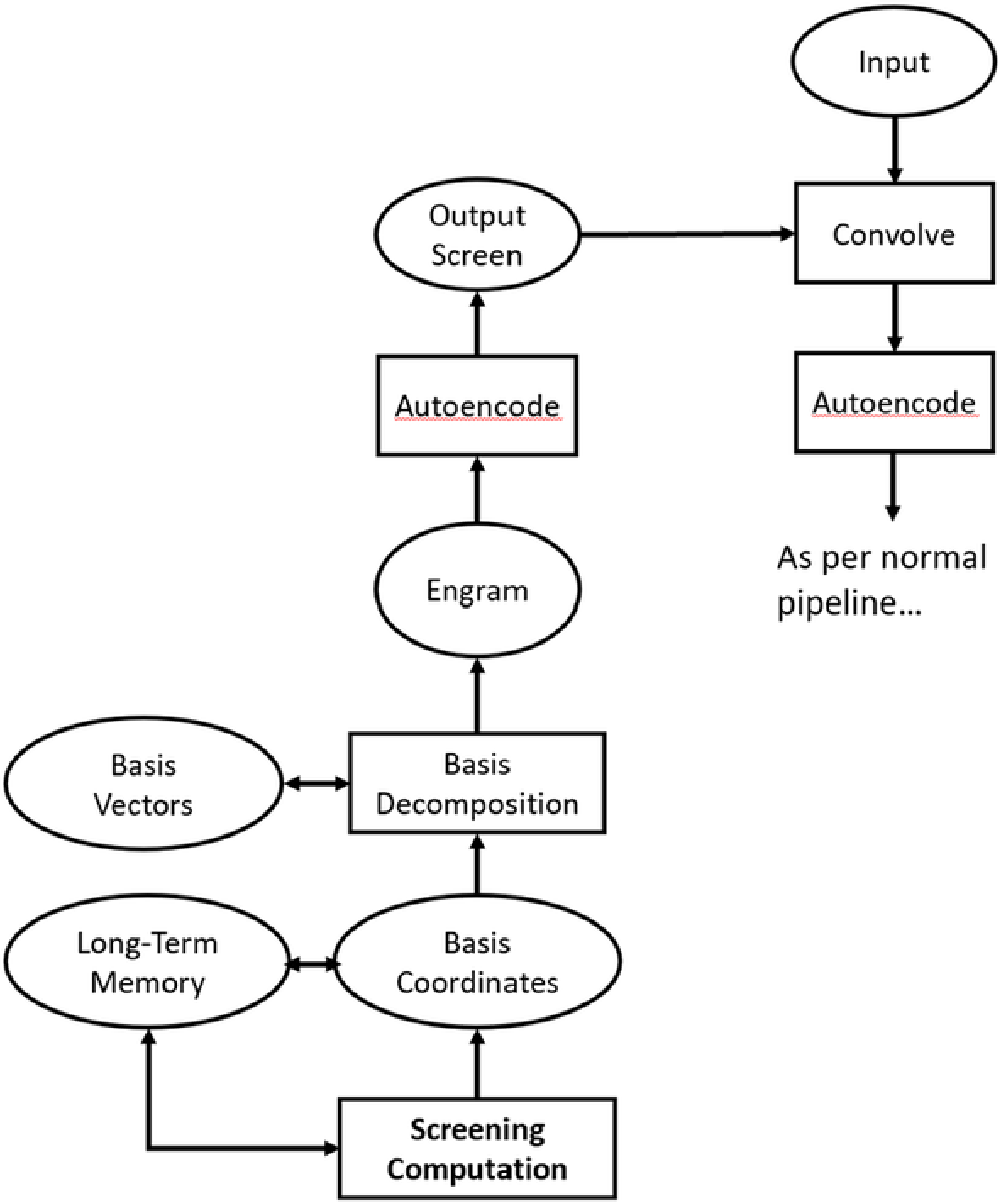


26) A method for the AI to converse naturally with a human, by training a set of dreaming predictors (22) evolved (5) to learn human language by training and evolution (5) on a plurality of human conversations, where they both learn proper responses, grammar, and composition from each training on one person’s side of the conversation, on the hierarchy of words and sentences (24) stated by each person in an alternating conversation. Then when actually conversing, the AI uses each predictors to predict what the other person will say next, and what it should say now, pulling words and phrases from previous segments of the conversation to incorporate them where appropriate, and dreaming where they need to ad-lib the conversation. Each predictor would also have connections to the information about other modalities (25) and their hierarchies, including visuals, audio, date, time, location to give the words context, and to interface with peripherals.

27) A method to compute a path through memory narratives (consisting of basis coordinates) and segments that can best solve a problem, by having predictors begin at a start point (consisting of starting basis coordinates), and/or at the goal (consisting of the desired ending basis coordinates), and then follow memory narratives away from their origin, branching at junctions one or more times until a path from the start connects with one from their goal, or they connect with their respective goal directly. The predictors can optionally detach from the narratives and dream as in (12) and reattach to another narrative. This process (27) can be iterated multiple times using the prior solution narratives and computing deviations to find better solutions.



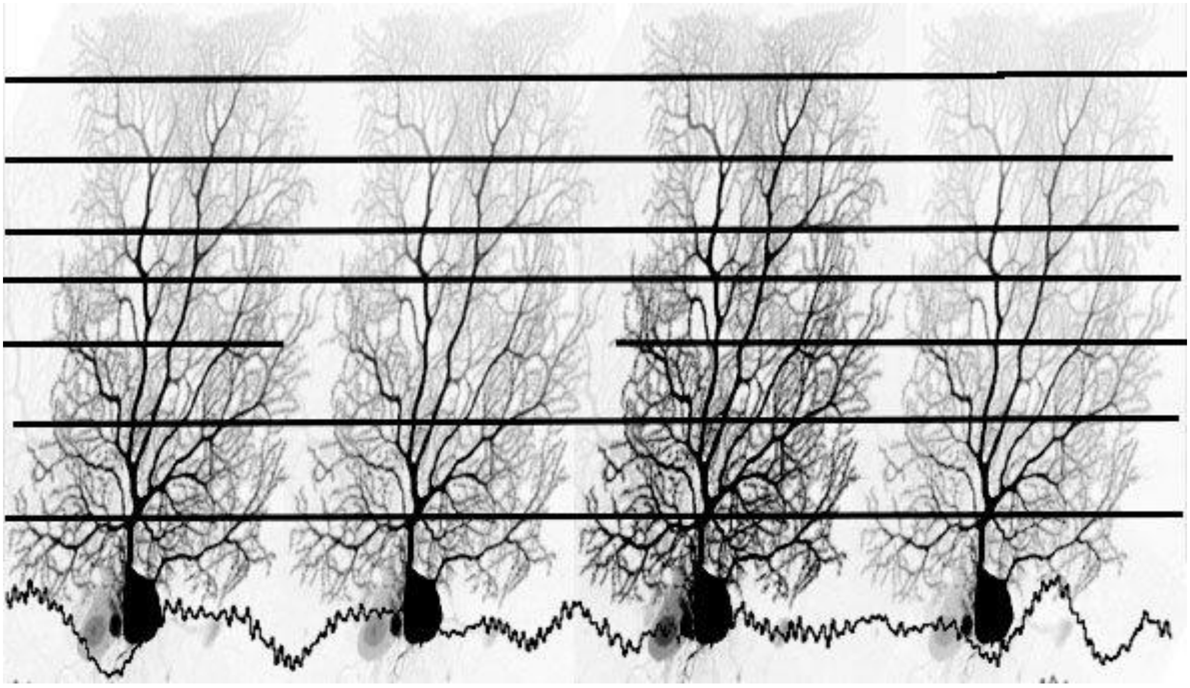
28) A method for selectively screening input by performing operations on it with computed output data, such as a convolution, subtraction, addition, or other operation. The data to be used as a screen can be generated from a memory, computed by a predictor, allowing the AI to look or listen for what it anticipates, or as a spatial filter to specify what area of a 2D input to look at, such as scanning with a convolution kernel smaller than the input, or samples at grid points in a 2D input with different convolution extents to generate a multi-sample, multi-resolution representation of the 2D input.



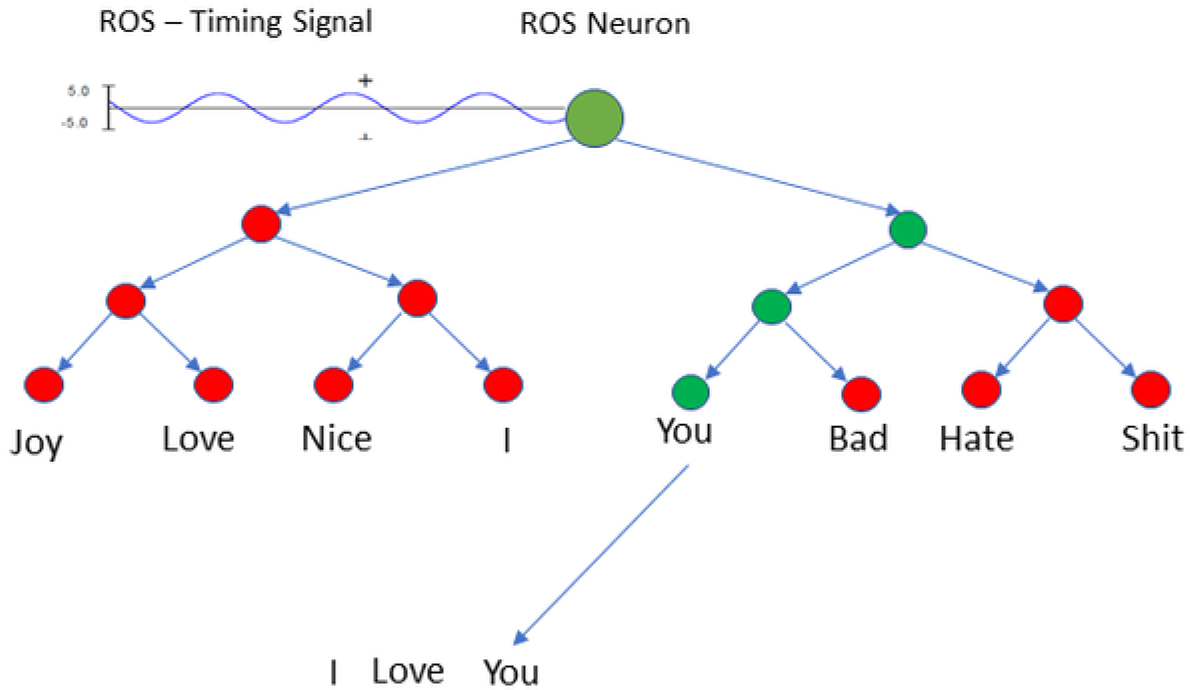
29) A method for training the AI to produce desired outputs for external systems (actuators for drones, robots, or peripherals, displays, and other devices), and control them. Based on input basis coordinates, and operations between them and narratives in memory, and other means - output basis coordinates and narratives are computed. Then these synthesized basis coordinates are decoded to a synthesized engram and then to synthesized outputs (at the 2nd layer of the autoencoders) which are fed to the external systems. The autoencoders are bidirectional, so for training, inputs are fed in the top layer of the output autoencoders and are encoded to engrams and basis coordinates (representing output narratives for actuators, peripherals, and displays), and these are

used to train the internal AI model and components on what outputs they should be computing.

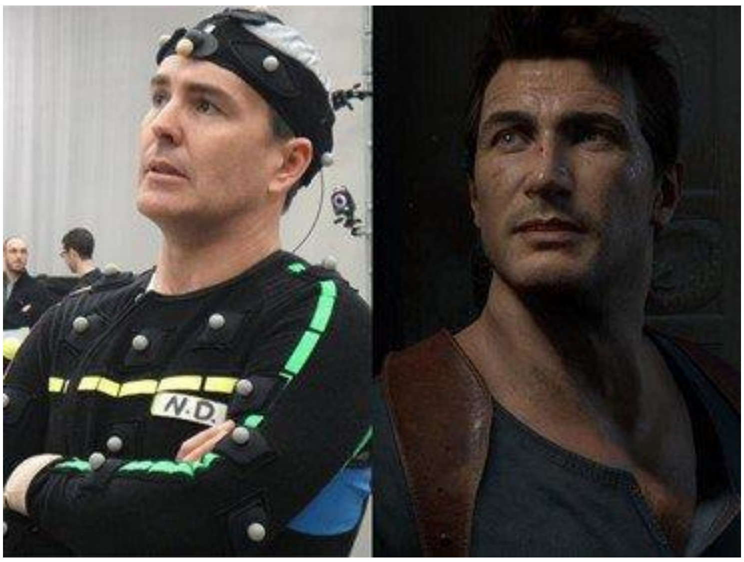
30) A method for computing temporal outputs for motor controls, language, and other outputs based on ROS neurons. It originates with a signal that sets a tempo or pattern with time (t), that is the same regardless of the output to generate. This signal is then input at the root of each of a plurality of hierarchies (19, 24) of branching structures (with the outputs at the leaf nodes) that can be selectively inhibited at each branch and level (by modulating the inhibitory signal at each branch with the signal that was passed down into the hierarchy at the branch) to select which terminal node(s) of the hierarchy are activated when the tempo signal is >0. Each branching hierarchy forms a spatial-temporal basis set that can be controlled by the inhibitory signals, and the outputs from each blended via these inhibitory signals to form novel output units that are sequenced temporally. This is trained by back-driving the desired outputs to train the inhibitory signals to generate that output.



31) Creating a branching pyramidal neural structure branching from the ROS temporal input origin up to the ‘cortical column’ autoencoders (15–18) such that each outermost branch terminates at one autoencoder, with the signal strength being the basis coefficient fed into that autoencoder. Now, when the ROS temporal input fires, the signal travels through the branches (modulated by the inhibitory signals to each branch), delivering basis coefficients that modulate with the basis vectors in the cortical column / autoencoder layer. which are then propagated up through the HFM hierarchy, and decoded to a series of engram segments corresponding to the output of the ROS / inhibitor network. This engram stream is decoded to the correct output by the autoencoder, be it audio, speech, visuals, actuator controls.



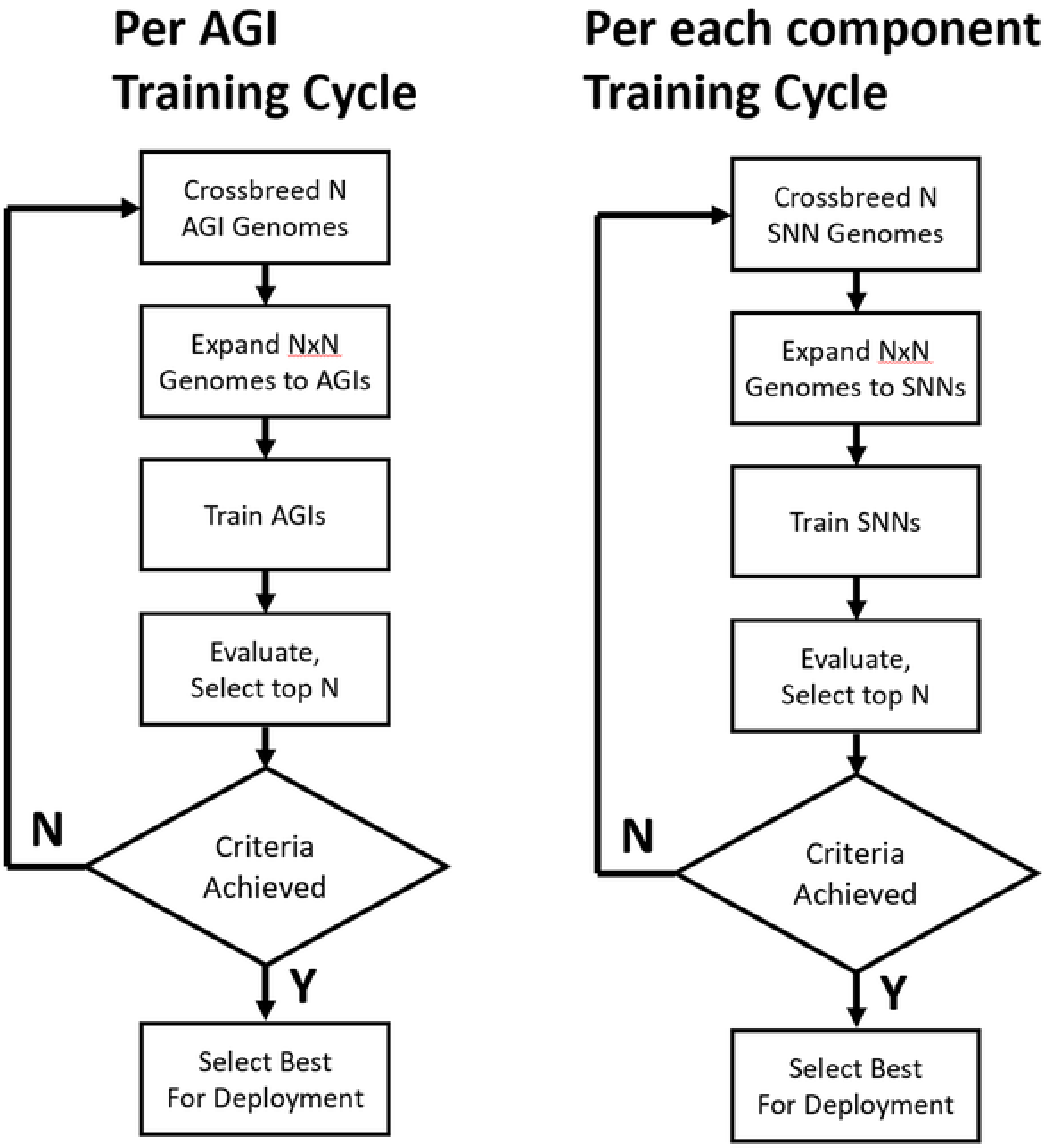
32) A method of training a human ‘mimic’ AI using the methods of (29, 30) by using data from a performance capture of a real person acting out a plurality of scenarios to supply the inputs and outputs to the AI for the language, speech, vision, body movement, and facial movement of an artificial AI person, where that AI person can be without physical form, instantiated with 3D computer graphics as a character, or instantiated as a realistic humanoid robot, with the motion and facial expressions mapped to the actuators for the latter, with the goal being to provide an AI with realistic, human-like dialog, lip-sync, facial expressions and movement. The detailed performance capture data can be augmented with simpler text conversations that may be general or specific purpose to a vocation, and other general data to fill in the blanks in training, and also augmented with dreaming (22) between training sessions.

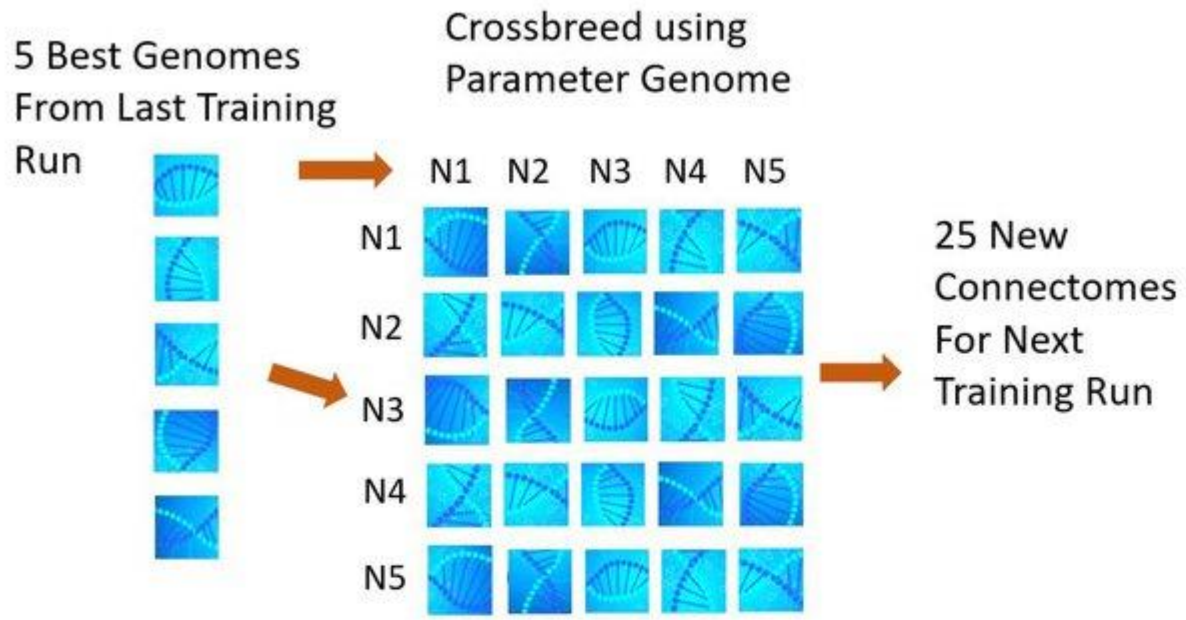


32) A method for defining the entire AGI system architecture described in these claims by a master genome, consisting of all parameters determining the functionality of the

system as a whole and all the properties, and number of instances of all the components using methods defined in 1-20, their configurations, and all parameters determining their individual functionality. This genome is compact and can be expanded to the full AGI system deterministically, in a manner that is smooth, such that a genome (G) interpolated between two genomes (G1, G2) expands to an AGI (A) that is also between the two AGIs represented by (A1, A2) in form and function.

33) A method of evolving the AGI as in (5), with a plurality of AGI genomes expanded to a plurality of AGI architectures that are then trained and evaluated in parallel, with the genomes of the most successful AGIs cross-bred, mutated and used to define the next set of AGI architectures to test. This could include an iterative process to evolve the genomes of the internal components one or more times in a sub-process between iterations of evolving the whole AGI system. This allows the AGI to be evolved, refined, and expanded exponentially, limited only by computing power and memory available and the amount and quality of input data to train and test it on.


Per AGI
Training Cycle
Crossbreed N
AGI Genomes
Expand NxN
Genomes to AGIs
Train AGIs
Evaluate,
Select top N
Criteria
Achieved
N
Y
Select Best
For Deployment
Per each component
Training Cycle
Crossbreed N
SNN Genomes
Expand NxN
Genomes to SNNs
Train SNNs
Evaluate,
Select top N
Criteria
Achieved
N
Y
Select Best
For Deployment



Resulting in: Artificial General Intelligence Eta in 2030

Artificial General Intelligence

ORBAI: Artificial General Intelligence ORBAI is developing Artificial General Intelligence that will enable more advanced AI applications, with conversational speech, human-like cognition, and planning and interaction with the real world, learning without supervision. It will find first use in smart devices, homes, and robotics, then in online professional services with an AGI at the core powering them. What we usually think of as Artificial Intelligence (AI) today, when we see human-like robots and holograms in our fiction, talking and acting like real people and having human-level or even superhuman intelligence and capabilities, is actually called Artificial General Intelligence (AGI), and it does NOT exist anywhere on earth yet. What we actually have for AI today is much simpler and much more narrow Deep Learning (DL) that can only do some very specific tasks better than people and has fundamental limitations that will not allow it to become AGI, simply because since computers were invented, their most significant limitation is that they are unable to interpret the real world around them. They have counted on humans to format the world into data the computers can understand and operate on, and for humans to interpret the results and even modern deep learning is subject to this limitation. AI researchers agree that today’s narrow deep learning and machine learning is too simplistic to ever match the human brain and our broad intellectual capacity and that it will instead take more powerful and flexible artificial general intelligence (AGI) to enable human-level AI and to go beyond it, which is what we propose here. ORBAI is developing and patenting more general AI methods that can dynamically decompose any reality it perceives into fundamental building blocks that it can use to understand and manipulate that reality with the native mathematical language of linear algebra and computers, then reconstruct its results from the building blocks back to reality, giving a computer the ability to work with general inputs and general real-world computations by proxy by: Learning to autoencode any time-space domain input, using the SNN autoencoders , into compact engram streams Convolving segments (moments in time) of those engram streams with a vector basis set using our HFM to get basis coefficients Storing basis coefficients encoded from inputs (or those computed) - into temporal narratives in memory Doing mathematical and other operations between basis coefficients, and narratives, effectively operating on reality by doing so Using neural net constructs such as predictors, solvers, and dreamers to do operations on narratives Reconstructing the internal basis coefficient narrative representations to engrams then real-world outputs. With these developments, ORBAI will take the first steps towards AGI that can perceive the real world, reduce those perceptions to an internal format that computers can understand, yet still plan, think and dream like a hum

https://www.orbai.ai/artificial-general-intelligence.htm

Brent Oster

brent.oster@orbai.com

To see more: ORBAI Artificial General Intelligence Roadmap

Fred Mitchell

, I seek to create AGI, or to aid in its genesis.
Answered Jun 3

Today’s AI is mostly hype. Nothing more than very fancy and sophisticated matrix operations with variations on the overall gradient descent algorithm which does not scale well.

The true AGI will require new hardware that does not exists yet.

Oh, but it is good hype to separate the money from the investors. We will see a similar “30 years from now…” bating technique that has been used with nuclear fusion for the past… what, 60 years or so?
701 views
View 5 upvotes
5

Add Comment

Roland Hughes

, Author - The Minimum You Need to Know book series
Answered Aug 7

Artificial Intelligence is the future of software development and always will be.

When I started in the 1980s that mantra was true. It’s 2021 and that mantra is still true. There was a time not long after I got into programming that every vulture capital firm on the planet was dumping money into AI. They lost their shirts, their shorts, and their shoes.

Brent Oster gave you an incredibly detailed answer.

When regular Joe and Jane hear “Artificial Intelligence” they think of Lt. Commander Data.

A fully autonomous android able to interact with humans. As an author I have to say that is still Science Fiction and won’t happen in my lifetime. Humans have random stupid thought patterns that AI can’t mimic.

Testing Conversational AI

Conversational Flow Testing ¶ This section starts with some technical background on Botium and then demonstrates a methodology to identify and formalize test cases for the conversational flow of a chatbot. The conversational flow, often called “user stories”, can be visualized in a flow chart. Hello, World! The Botium Basics ¶ The most basic test case in Botium consists of BotiumScript ¶ In Botium, the test cases are described by conversational flows the chatbot is supposed to follow. For a sample “greeting” scenario, the Botium test case looks like this — also known as “BotiumScript”: #me hello bot! #bot Hello, meat bag! How can I help you ? You can write BotiumScript in several file formats Convos and Utterances ¶ So, let’s elaborate the “Hello, World!”-example from above. While some users will say “hello”, others maybe prefer “hi”: #me hi bot! #bot Hello, meat bag! How can I help you ? Another user may enter the conversation with “hey dude!”: #me hey dude #bot Hello, meat bag! How can I help you ? And there are plenty of other phrases we can think of. For this most simple use case, there are now at least three or more BotiumScripts to write. So let’s rewrite it. We name this file hello.convo.txt: TC01 - Greeting #me HELLO_UTT #bot Hello, meat bag! How can I help you ? You may have noticed the additional lines at the beginning of the BotiumScript. The first line contains a reference name for the test case to make it easier for you to locate the failing conversation within your test case library. And we add another file hello_utt.utterances.txt: HELLO_UTT hello bot! hi bot! hey dude good evening hey are you here anyone at home ? The first BotiumScript is a convo file — it holds the structure of the conversation you expect the chatbot to follow. The second BotiumScript is an utterances file — it holds several phrases for greeting someone, and you expect your chatbot to be able to recognize every single one of them as a nice greeting from the user. Botium will take care that the convo and utterances files are combined to verify every response of your chatbot to every greeting phrase. So now let’s assume that your chatbot uses several phrases for greeting the user back. In the morning it is: #me HELLO_UTT #bot Good morning, meat bag! How can I help you this early ? And in the evening it is: #me HELLO_UTT #bot Good evening, meat bag! How can I help you at this late hour ? Let’s extract the bot responses to another utterances file: BOT_GREETING_UTT Good evening Good morning Hello Hi And now comes the magic, we change the convo file to: #me HELLO_UTT #bot BOT_GREETING_UTT Utterances files can be used to verify chatbot responses as well. To summarize: An utterance referenced in a #me-section means: Botium, send every single phrase to the chatbot and check the response An utterance referenced in a #bot-section means: Botium, my chatbot may use any of these answers, all of them are fine Identification of Test Cases ¶ If the flow chart is available

https://botium-docs.readthedocs.io/en/latest/03_testing/01_testing_conversational_ai.html

Data trying to learn what humor is. Fantastic episode of TNG. The writers let Whoopi have a big role in that episode. They even cover “funny over time” when it becomes not funny anymore.

All this is not to say AI has been a complete failure.

One of the very first things we got out of all that financial loss was fuzzy logic and the autofocus camera. Much of this lane drift detection, automatic parallel parking, etc. you now see in automobiles are baby steps in AI.

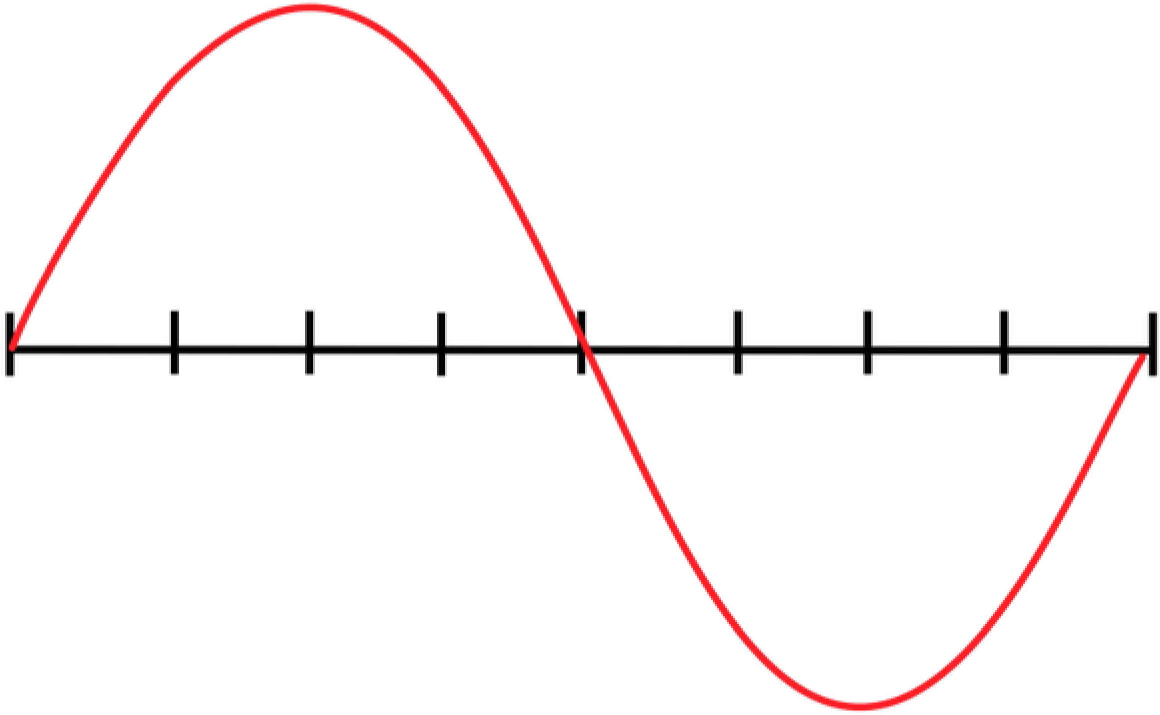
True AI cannot happen with a digital computer.

No matter how good you get you can't make it work. Whatever you create will always be a poor man's imitation. Watch the *Firefly* episode *Safe* to understand this next quote.

It was made by someone with longing. They longed to see a duck.

To have even the slightest hope of real AI, you have to bring back and heavily invest in computing technology that was abandoned long ago. Analog Computing.

Digital computing is binary. A bit is either charged (1) or not charged (0). Those are the only states it has. Analog computing uses the infinity of a Sine wave.



You are limited only by how finely you can measure current/voltage/charge/whatever. A single “bit” for lack of a better word, can have any value, positive or negative, within that wave.

In the digital world it can only understand the thermometer says the room is 70 degrees.

In an analog world it can understand that yesterday 70 degrees felt uncomfortably warm and today it feels chilly.

Every so many years someone with too much money will get too enamored with Science fiction. They will convince others with too much money to sink huge amounts into AI. Some tiny advances will happen and all of the money will be lost.

Intelligence isn’t binary. Some of the smartest people in the world have no common sense what-so-ever.

AI has to mimic that.

Related Questions
More Answers Below

Is artificial intelligence overhyped technology?

Is machine learning currently overhyped?

Will economists be replaced by AI?

Nathan Ketsdever

, Facinated by science, discovery, & innovation.
Answered May 16

They are, by my estimation, the opposite of hype in terms of the endurance: Artificial intelligence, algorithms, and big data promise to be with us for the rest of existence.

All of the technology company have created an algoritm or use an algoritm to make their company work. Billions and billions and billions of dollars are in algoritms. The number is probably the size of the tech sector—maybe the whole of Western economies.

Let's be clear, early experiments in A.I. in a given sector or context, like any new product or technology will experience cycles of hiccups until they are calibrated and perfected.

FYI: I would say otherwise, you might have to define hype vs. success to help provide any deeper level of specificity, context, or evidence to support a given thesis on this question.

Update:

Perhaps you could argue that A.I. at one time was hype to some extent or was in a Gartner hype cycle, but I think it's beyond that now, given how many of our modern day apps are fundamentally reliant on A.I. and algorisms:

Gartner hype cycle - Wikipedia

The Gartner hype cycle is a graphical presentation developed, used and branded by the American research, advisory and information technology firm Gartner to represent the maturity, adoption, and social application of specific technologies . The hype cycle claims to provide a graphical and conceptual presentation of the maturity of emerging technologies through five phases. The model is not perfect and research so far shows possible improvements for the model. [1] Five phases [edit] General hype cycle for technology Each hype cycle drills down into the five key phases of a technology's life cycle. No. Phase Description 1 Technology Trigger A potential technology breakthrough kicks things off. Early proof-of-concept stories and media interest trigger significant publicity. Often no usable products exist and commercial viability is unproven. 2 Peak of Inflated Expectations Early publicity produces a number of success stories—often accompanied by scores of failures. Some companies take action; most don't. 3 Trough of Disillusionment Interest wanes as experiments and implementations fail to deliver. Producers of the technology shake out or fail. Investment continues only if the surviving providers improve their products to the satisfaction of early adopters. 4 Slope of Enlightenment More instances of how the technology can benefit the enterprise start to crystallize and become more widely understood. Second- and third-generation products appear from technology providers. More enterprises fund pilots; conservative companies remain cautious. 5 Plateau of Productivity Mainstream adoption starts to take off. Criteria for assessing provider viability are more clearly defined. The technology's broad market applicability and relevance are clearly paying off. If the technology has more than a niche market then it will continue to grow. [2] The term "hype cycle" and each of the associated phases are now used more broadly in the marketing of new technologies. Hype in new media [edit] Hype (in the more general media sense of the term "hype" [3]) plays a large part in the adoption of new media . Analyses of the Internet in the 1990s featured large amounts of hype, [4] [5] [6] and that created "debunking" responses. [3] A longer-term historical perspective on such cycles can be found in the research of the economist Carlota Perez . [7] Desmond Roger Laurence, in the field of clinical pharmacology , described a similar process in drug development in the seventies. [citation needed] Criticisms [edit] There have been numerous criticisms [8] [9] [10] [11] of the hype cycle, prominent among which are that it is not a cycle, that the outcome does not depend on the nature of the technology itself, that it is not scientific in nature, and that it does not reflect changes over time in the speed at which technology develops. Another is that it is limited in its application, as it prioritizes economic considerations in decision-making processes. It seems to assume t

https://en.wikipedia.org/wiki/Hype_cycle

Tatsiana Isakova

, keen on Artificial Intelligence and Big Data
Answered Jul 23

```
“By far the greatest danger of Artificial Intelligence is that people
conclude too early that they understand it.” — Eliezer Yudkowsky
```

Yep, it’s been mostly hype so far. However, it is starting to live up to its popularity with the proliferation of AI methods in a gamut of industries.

But to be clear: *it is still far away from its intended meaning - a true, human-like intelligence demonstrated by machines.*

I think its overhyped present owes to the money that can be raked in by claiming that AI is already real. It sells well, so why not promise the fully automated future even if it's not?

In all its fairness, AI-powered systems beef up lots of legacy industries and niches. Therefore, its popularity is not exactly dubious.

Join my space about AI and business: AIandBusiness

Andrey Kozhanov

, Senior Developer at Luxoft
Answered Jun 28

What AI is in its core? It's essentially just a mapping from incoming data to the best possible outcome. How good a particular outcome is is determined by a function. Each problem must first be determined in terms of input and output - a set of numbers for input, another set of numbers for output. And a function to determine the value of outcome.

The dimension of both input and output are determined by human. AI can't change any of them. Same is true for value function. So it does not matter how much advanced AI is or whether it could train itself. It's just a tool like any other and we should not overestimate it.

Peter Wone

, I use scientific method every day to debug software.

Answered Jun 9

It is and it isn’t.

AI as presented in fiction is god-in-a-box.

AI as built is a pattern finder.

Reductionism would have it that natural intelligence is just a pattern finder, and it is in that context that one can reasonably use the term AI as it is currently used in the software industry.

The reason you perceive it as a hype is the vast gulf in utility between god-in-a-box and an algorithm that can learn to recognise bicycles in photos.

This is probably just as well. The day that the pattern seeking machine starts having opinions and desires of its own might be the beginning of the end for humanity, and even if it’s benign or even well-disposed to humanity, using it as a tool amounts to slavery.

Suresh Kumar S

, former Chief Scientist at Council of Scientific and Industrial Research
Answered Jun 6

There is a gap between popular belief in machine learning and what the tools of machine learning can actually accomplish. Pandora’s box opens and dreams follow suit.Today, machine learning is an iceberg of massive proportions sitting at the top of the Hype Cycle, the Peak of Inflated Expectations. Slowly, bits and pieces of it get chipped away by the population, falling through a Trough of Disillusionment to be refined into a usable product.

AI does the math faster, saving money by automating normally complex processes. It makes your life easier even now, behind the scenes, like your Netflix recommendation engine

Artificial intelligence (AI) is suddenly everywhere, promising self driving cars, medical breakthroughs, and entirely new ways of working. But how do you separate the hype from reality, and how can your business apply AI to solve real-world business problems is the million dollar question.

Some people are also under the impression that AI may have already achieved human level intelligence, something we talked about in a previous video, where we talked about different categories of AI, one of which is called artificial general intelligence or AGI. And this is partly due to what we see in, TV series and movies and comic books around science fiction type stuff

AI will transform industries. For example, algorithms help healthcare professionals more accurately recognize anomalies or patterns in medical images. AI will help redesign the entire shopping experience, optimizing everything with more and better data.

AI can absorb cognitive drudgery, like turning data points into visual charts, calculating complex math formulas, or summarizing the financial news of the day into a single report. This frees up people to focus on acting on the insights.

1

Add Comment

Emmanuel Maggiori

, Author of Get Tech: Cut Through the AI hype
Answered Jun 4

I think there's two answers to your question:

1. Sometimes AI doesn’t learn meaningful stuff but this goes undetected

Current AI systems are based on finding correlations in data. For example, a system automatically scans past users' purchases to infer patterns (such as the fact that buying a dumbbell makes a customer likely to buy protein powder). However, it is very common for those systems to learn patterns that aren't really meaningful (think of spurious correlations that just happened in the data by chance). They also fail a lot at producing anything meaningful for rare cases that don't appear in the data but are very important (think of the surprising threats a self-driving car may have to face yet wouldn't be present at all in past data samples).

Quite often, the validation framework used to measure performance isn't good enough to detect those things, making AI look better than it really is. Throughout my years in the field, I've seen many cases of an overly optimistic performance going undetected for years. And this (exaggerated) performance often brings attention to the team building it from others in the organization, resulting in a higher budget allocated to them, new hires, etc., ramping up the hype.

2. Understanding and validating AI is quite obscure

The second reason that aggravates the hype is that business managers often do not understand how AI works and do not scrutinize the validation process. It requires indeed a mix of maths, computer science and statistics to understand what's going on under the hood. Therefore, it is quite easy to "oversell" a result to a manager who doesn't know the technicalities, or for an unintentionally bad result to go undetected all the way up to the higher spheres of a company. Unlike other tasks (e.g., number of units sold), the figures that show that AI is working tend to be quite obscure and easy to manipulate or misunderstand.

Damping down the hype

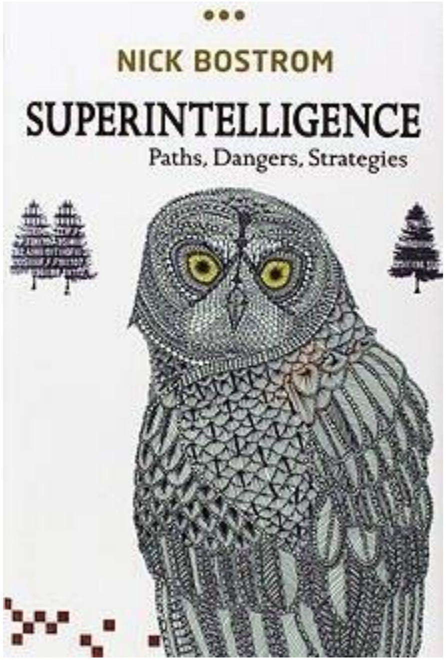
My general advice to avoid those pitfalls is to focus on having a really clean data flow and validation process (since the slightest glitch can skew the results), to set realistic expectations (no, self-driving cars won't roam the streets anytime soon), to measure the business value of AI in a clear way, to have managers learn about the validation process, and to create a culture in which it is okay to admit that an AI-based system may end up not being as good as everyone had expected.

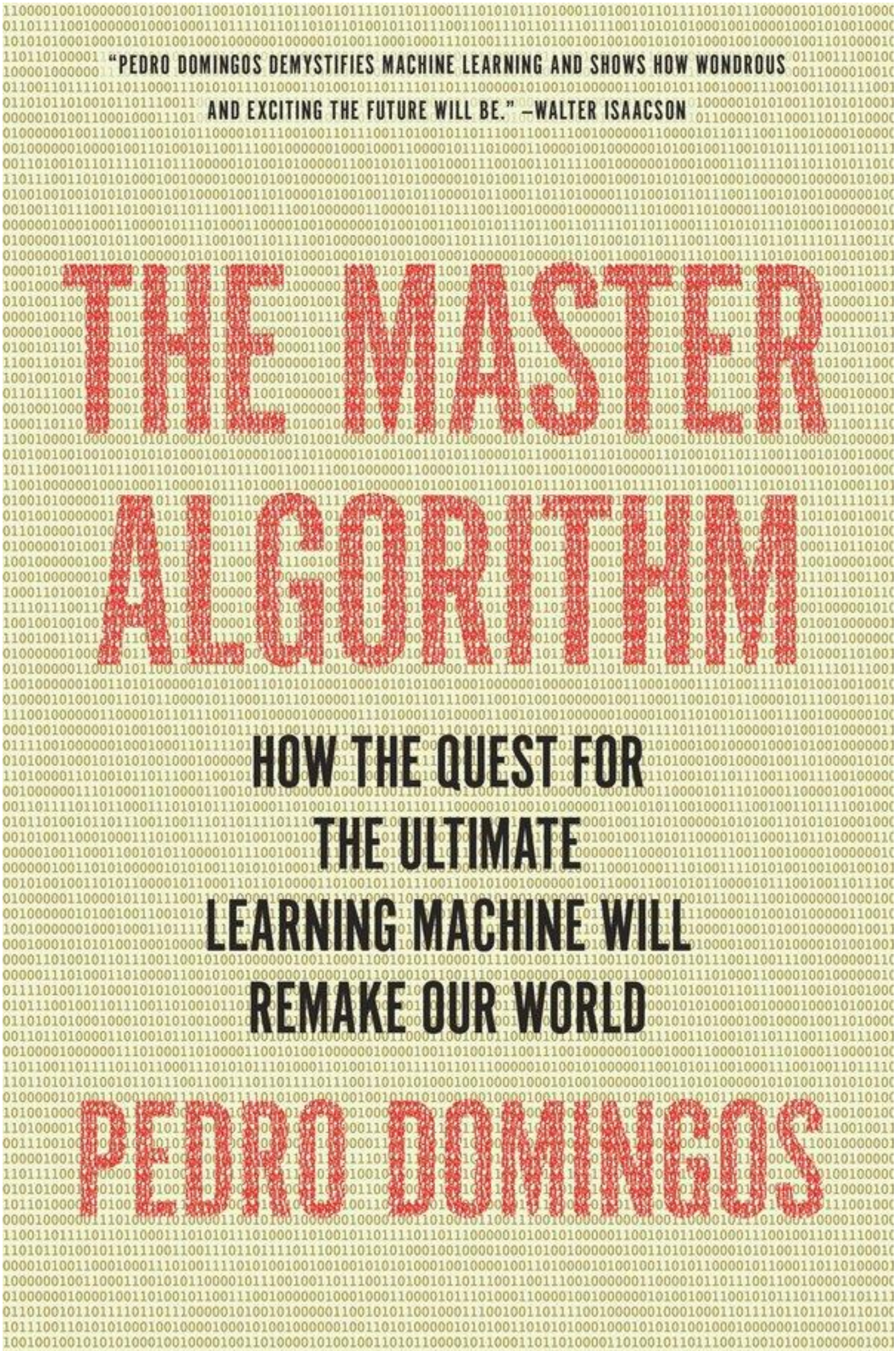
Da Zheng

, works at Self-Employment
Updated Feb 27, 2019

Originally Answered: Why is artificial intelligence so popular?

Hi, I think it’s the promised technology in the future, in fact the technical name is called “deep learning” or “machine learning”. Both media and tech company are promoting it like crazy, that is why people think it is so popular, but may be a little bit exaggerated. If you want to know more about it, you can read 2 books called “Master algorithm” and “super intelligence”.


NICK BOSTROM
SUPERINTELLIGENCE
Paths, Dangers, Strategies


"PEDRO DOMINGOS DEMYSTIFIES MACHINE LEARNING AND SHOWS HOW WONDROUS AND EXCITING THE FUTURE WILL BE." —WALTER ISAACSON
THE MASTER ALGORITHM
HOW THE QUEST FOR THE ULTIMATE LEARNING MACHINE WILL REMAKE OUR WORLD
PEDRO DOMINGOS

View 3 upvotes
3

Add Comment

Boyan Dob

, Working with computers for 30+ Y. Consciousness is my love.
Updated Jun 23

Originally Answered: Is the concept of artificial intelligence overrated?

Yes, what most people think what AI can do is overrated. To be clear though, AI is surely a very useful tool.

Faster and faster computers, with more and more sophisticated alghoritms, can solve many tasks which humans just couldn't and cannot.

But no matter how fast a computer will work, and no matter how well it will be programmed, that still means nothing in regard of computers being self-conscious. Consciousness won't just magically appear due to fast data processing.

I will claim that without consciousness there cannot be true intelligence, nor imagination, nor creativity. Nor can there be true feelings and emotions which are main drivers (to survive, evolve and prosper) for all living beings.

So, the real question is how to make computers conscious, if we want to bring humanity to new better era.

My opinion is, first, that primordial consciousness is fundamental & universal, and second, that special part in center of our brains, called claustrum, can tap into it and make whole brains a conscious biological machine.

That's what I would recommend to study to those who are seriously in this field.

I am aware I might be shooting in the dark, but so far it seems everyone is shooting in the dark in attempt to program consciousness into computers & robots, despite that billions of dollars are being put into this field of research.

It's clear that whoever (government, company or even individual) achieves this it will make them incredibly powerful and rich.

Rayed Ali

, Engineering Computer Science & Artificial Neural Networks (2019)
Answered Apr 13, 2018

Originally Answered: Is the concept of artificial intelligence overrated?

Not at all AI could be the solution to human beings most complex problems, besides AI can and is changing the way technology influences our day to day life. AI that we see today is in it's infancy state with evolving concepts in microbiology, genetics, predictive analysis, data science there's a whole new universe awaiting to be explored.

AI could change the way we are born, live and die. With artificial neural networks there's chance we can someday be able to reside our memory and consciousness in a small computer, maybe a way to achieve immortality. Everything that we consider science fiction could well be possible in near future. One such example is gene editing with the help of CRISPR technique, which could eliminate hereditary mutations in chromosomes and even add disease resistant genes to new offsprings.

AI was recently used to conclude the Bose Einstein condensate experiment and surprisingly the algorithm with AI's Intellect was 10 times faster as compared to human scientist which shows that the concept of AI is set to achieve infinite possibilities in physics and cosmology.

In short imagine if a computer could think and solve what surpasses a human brain's level we can literally recreate a Big Bang.

2.4K views

View 4 upvotes

4

1

1 comment from

Boyan Dob

Brad Cardin

, Entrepreneur (2008-present)

Answered Nov 17, 2018

AI is considered a hype due to

1. A lot of people do not understand the technicality behind most of the innovation and almost fall for any news published.
2. AI's potential is immense. Its a revolutionary technology that covers almost every mainstream industry we know, be it automation or forecast, be it diagnosis or maintenance, AI is capable of doing low skill to most complex tasks.

Although, I don't term it as hype. Most of what we aim to achieve is practically possible, its just a matter of time AI will be a part of our routine life.

892 views

View 2 upvotes

Answer requested by

Arjun Parth

2

Add Comment

Josh Greig

, studied Computer Science
Answered Feb 25, 2020

AI is definitely over-hyped. Don't buy into it or fall for it.

AI is the cure-all tonic of the 21st century. It solves everything or could kill everything at least that's what science fiction would lead you to believe. Like the cure-all tonics of the early 20th century, AI won't solve every problem or come close any time soon.

AI is becoming such a subjective, hyped, overused, and meaningless term that it is clearer to use the word "software".

Computers have been getting faster, more connected, smaller, and in more places over the last few decades. Software is moving with them and solving new problems. This is the morsel of truth that gets twisted to attribute an unrealistically extreme anticipated change to AI.

Many wish software was able to solve more problems, better, and faster than they actually can. Elon Musk is quite brilliant but overestimates the speed and quality of software Tesla can produce for autonomous driving. Tesla has frequent delays and fatal car accidents related to its Autopilot technology as a result of this wishful thinking and rushed deployment.

The hype pushes advertisers to use "AI" whenever anything software-related is used to solve a problem now. This causes people to pay extra for things that are less innovative and useful than they're sold to be. A good example of this is The Grid.

See the original pitch at:

If you don't want to watch, essentially "The Grid" is supposedly a website that uses AI to fully automate a web designer's job more completely and better than any website builder. The Grid will help you manage content while it manages all design aspects better than you, a non-designer, non-programmer could.

This reviewer explains the disappointing reality:

These articles are great further reading on the misuse of the AI term to sell:

- Beware of Geeks Bearing AI Gifts
- 3 Reasons AI Is Way Overhyped

The hype for AI is the same hype I've seen for decades and appears to largely come from science fiction, advertising, and wishful thinking. AI programming concepts are an interesting way to solve software problems but they're not new and won't solve as much or as well as most want.

644 views

View 1 upvote

1

Add Comment

Jesse Pollard

, programmer/analyst/administrator from way back

Answered Dec 2, 2016

What makes you think it is a hype?

Yes, some companies WILL exaggerate their results (hence all the face recognition failures), but it does work - when given ideal data. Given partial data/incomplete data/ inaccurate data ... and it isn't as reliable even if it is reported as correctly working.

It works quite well in the fields it is aimed at. It sucks big time when you get outside those fields until new approaches and retraining of the AI (depending on type) is done.

850 views

Add Comment

Christopher Albertson

, lived in Los Angeles
Answered Apr 23, 2020

It is something created by people who don’t work in the field of AI and imagine lathe things it might do one day. And then others believe it and it feeds back on itself.
275 viewsAnswer requested by

Benyamin Kellow

Add Comment

John Sanders

, Principle Scientist Marconi - AI and computing (MMARL) in 1980s
Answered Apr 30

It’s a hype because it creates high levels of expectation without a hope of delivery.

Hyperbole is a really a figure of speech , and so not quite appropriate, but in recent times the word has evolved particularly in the worlds of fashion, media sales etc. In essence its exaggerated as in making great claims for…. and usually abbreviated to hype. The claims of AI as a new advance in computing is hype. Most of the computer AI products do not have any level of intelligence, but may have specific uses and are classified as computer software Machine “learning” being a good example.

The concept of systems that think has been attached to computer based product range by media type making films for example. This is not going to happen on a computer system and really can be considered as hyperbole.
258 viewsAnswer requested by

Genny Harrison

and

De'Angelo Belle

Dhiraj Naik

, Founder of Nikstark youtube channel

Answered Jun 21

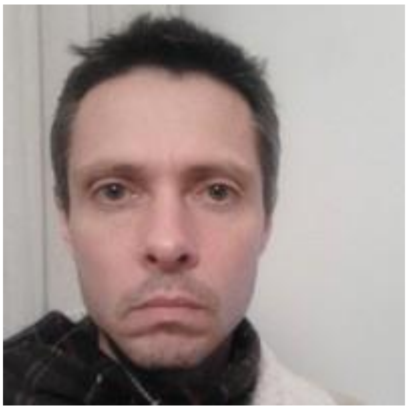
Answer: Artificial intelligence was once just hype, now it's business.

Only a few years ago, artificial intelligence was mostly a fascination in science fiction. But today AI is all the rage--and companies are investing billions to make use of its power, with predictions it will generate trillions of dollars in revenue. In recent months they've been aggressively moving into healthcare and autonomous driving while expanding their ambitions overseas and investing in startups that develop the technologies they need to reach those goals. Some say that AI could change every facet of our world by making us smarter and more efficient--today companies are most focused on using these new-found capabilities for competitive advantage abroad rather than at home where there's still plenty of room for improvement.

Rachel Li

Answered Oct 12

The fact that AI has been hyped can't deny its capabilities, for example, it is doing applied across multiple industries, and because of its huge potential, AI is adding tremendous value to them. The center is just to make you aware of AI abilities that are far away from misinformation.



Gavin Rens

, Doctorate in Computer Science specializing in Knowledge Rep. & Reasoning.

Answered Nov 16, 2018

Originally Answered: Why is artificial intelligence in too much hype these days?

AI is hyped so much because of how large the influence could be to people in the next few year, that is, AI will have a real impact on most people alive now in their life time.

Some people will lose their jobs, some people's jobs will change a lot, some people will have their lives extended by AI tech, some people will become poor(er) and others rich(er).

AI will change what we know, and how we interface with information. It will create and is creating new entertainment genres and military weapons and way to deceive and new ways to stay safe. Et cetera, et cetera.

Most people (even AI specialists) have only woken up to this fact in the last three to five years, due to impressive results mentioned in the media.

Steph Groom

, former Structural Engineer and Project Manager (1980-2010)

Answered Mar 2, 2019

Originally Answered: Why is artificial intelligence so popular?

Why is artificial intelligence so popular?

Artificial Intelligence or AI is one of the newest Sciences, that is constantly moving forward at very fast development pace. Despite of some researchers' say that AI is a threat to humanity (which human controlled it is not), a large number of people believe that AI may positively change the future of the world and ease mankind's life routines. In the late 20th century, when the main studies on AI began, no one would think of a quick jump it would make in a few decades. AI-powered tools now help to scale the efforts of sales teams by gathering useful patterns from data, finding successful courses of action, and addressing customer needs and grievances.

The general benefit of artificial intelligence is that it replicates decisions and actions of humans without human shortcomings, such as fatigue, injuries, emotion and limited attention time. There are several examples and applications of artificial intelligence in use today: voice-controlled individual partners, robots, behavioral calculations, suggestive searches, autonomously-powered self-driving vehicles, virtual assistants, etcetera.

The following short list of uses of AI that exemplifies its growing popularity:

1. **AI-Powered CRM Suites**

Like no other human resource AI suites may replace Customer Relations Management block, which includes activities like answering calls, filtering them, sorting, calculating the satisfactory analysis and so on. Of course, this kind of labor automation have both advantages and disadvantages of its own, but this is another large topic of discussion.

2. **Gaming Intelligence**

AI plays a crucial role in strategic games such as chess, poker, tic-tac-toe and much more, where machine can consider a huge number of conceivable positions based on artificial intelligence applied on it.

3. **Vision Systems**

These systems understand, interpret, and comprehend visual input on the computer, e.g. a spying aerial vehicle (drone) takes photographs, which are used to figure out spatial information or map of the areas; face detectors used by police which recognize face of the criminal and much more.

4. **Speech Recognition**

Many intelligent systems are capable of hearing and comprehending the language in terms of sentences and their meanings while a human talks to it. It can handle different accents, slang words, noise in the background, human speech variations, etc. Such AI systems are Apple's Siri, Google Assistant, Amazon Alexa, Microsoft Cortana, SoundHound, Bixby, VIV and still others to come.

5. **Handwriting Recognition**

This is a computer software designed to receive and interpret intelligible handwritten input from sources such as paper documents, photographs, touch-screens and other devices. It can recognize the shapes of the letters and convert them into an editable text.

6. **Intelligent Robots**

Robots can play out the tasks given by a human. They have sensors to recognize physical information from reality, e.g. light, warm, temperature, movement, sound, knock, and pressure. They have productive processors, various sensors and tremendous memory to show insight. Also, they are capable of learning from their mistakes and can adjust to the new environment.

Source

1. https://medium.com/@sarah.dukes/why-is-artificial-intelligence-becoming-so-popular-33d12c35a232

1.3K viewsAnswer requested by

De'Angelo Belle

Add Comment

Akhil Reddy

, Business Development Associate at Skill Analytica (2018-present)
Answered Feb 7, 2019

Originally Answered: Why is artificial intelligence so popular?

Artificial intelligence or AI is nothing but the science of computers and machines developing intelligence like humans. In this technology, the machines are able to do some of the simples to complex stuff that humans need to do on a regular basis. As the AI systems are used on a day to day basis in our daily life, it is not wrong to say that our lives have also become advanced with the use of this technology.

The AI systems are efficient enough to reduce human efforts in various areas. In order to perform various activities in the industry, many of them are using artificial intelligence to create machine slaves that perform various activities on a regular basis. The artificial intelligence applications help to get the work done faster and with accurate results. Error-free and efficient worlds are the main motives behind artificial intelligence. In recent years, many sectors have started using AI technology to reduce human efforts, and also to get efficient and faster results.

Artificial Intelligence plays a very important role in not just the development of business and processes but also the humans to the next level. With the rapid growth in technology and development, we can expect a lot more exciting features and uses of AI in the future.

1.9K views
View 1 upvote
1

Add Comment

32 answers collapsed

(Why?)

Related Questions

Is artificial intelligence overhyped technology?

Is machine learning currently overhyped?

Will economists be replaced by AI?

How is deep learning not the future of artificial intelligence?

How much of AI is hype vs. substance?

Is artificial intelligence a smart idea?

Is it possible to create artificial intelligence? Why or how?

What are applications of artificial intelligence that are in use today?

What comes after artificial intelligence?

What is the most advanced artificial intelligence (general intelligence) that exists today?

Why does everyone need to learn AI and machine learning?

What is hype and what is real in the artificial intelligence world?

How does artificial intelligence affect human life?

Who are the biggest players in artificial intelligence today?

What is the future of artificial intelligence?