Ideas in Modern Artificial Evolution

David Buffkin

October 2020

1 Introduction

What follows is a collection and elaboration on various ideas I have had with regards to Genetic Algorithms and Neuroevolution. The majority of these ideas have been composed during the past year and are thus up to date with my current knowledge. That being said, there exists much research in these areas (though not as much as I would like) of which I am unaware. I am aware that many of these ideas have possibly been explored already, yet I find it interesting to think about them nonetheless. As I continue to review the literature, I will attempt to keep these as up to date as possible.

I first discuss a few broad yet important concepts which shape the stage of these ideas and implicate critical meaning. I then touch on a set of neat issues which I think are worthy of examination. The next two sections are the meat of the ideas, split into more classic ES and Neuroevolution. Therefore, if you are only interested in my implementation ideas, you should skip to sections 4 and 5. Ideally I would like to find the time to implement these myself. This is definitely a goal of mine, and I will see to it as I can.

Finally, do note that there are many implicit citations here. For now, this comprises a sort of to-do list and scratchpad for me to collect my thoughts. When a significant work has been made of them, I will further elaborate on references. I will say that many of the papers I have read have stemmed from links in [1].

2 Overarching Themes

Some general themes that have shaped many ideas I have had.

2.1 The Origin of Consciousness

4.10 might literally eclipse this section.

One of the reasons I suppose that ES has begun to lose interest in recent times is that the problems it is more suited to solve are not in public focus. The more algorithmic, well-defined problems (supervised) which so many seek to solve naturally have more algorithmic solutions. The stage in which evolution has proven itself a worthy strategy is far more general and less defined: the 'problem' of surviving on the planet Earth.

I believe that an ultimate goal of artificial intelligence is to mimic human consciousness. In order to do so, we must consider from what conditions this arises. This is no easy task due to the nebulous nature of consciousness, but my working theory goes something like what follows.

As a result of living in a constantly changing and challenging environment, all organisms which are alive must necessarily adapt to live on. The fitness of an individual is in direct proportion to, among other things, their ability to take advantage of the environment around them in order to live in prosperity. There must have come a point, then, in the evolutionary history of Homo Sapiens, that the entry barrier to a higher level of fitness (or a necessity to continue to procreate) was the attaining of higher thought. What I mean to say is that our consciousness is a product of the largest, most dynamic evolutionary problem in existence.

There are two possibilities. One is that I am blatantly wrong. I acknowledge yet ignore this possibility for this discussion. On the off chance that I am somewhat correct, there are two significant unfortunates. One of these is that the computational ability to reproduce consciousness will likely not exist within my lifetime. It took over 3 billion years on the smartest simulation ever to produce us; I am not optimistic for 80. The other unfortunate is that there needs to be a significant shift in attitude towards ES in the AI community. Yet, if we continue to be distracted (albeit rightfully in many cases) by other, smaller problems, this is unlikely to occur.

I guess that a conclusion of this would be that I should give up the ghost on Ex Machina. Yet, I find it fascinating, and continue to wonder in case I am wrong. Either way, though we may not get as far as to become sentient, there still lies significant promise in using ES to solve a variety of real-world problems, as so many have shown. And it is these I will focus on until further progress seems possible on the other front.

2.2 Towards Biology

It seems ironic to me that we have so much research on the creation of the smartest things we know of-humans-and yet so little in comparison is devoted towards replicating it. There is a theme in many ES related papers to use biological analogies for the techniques in the paper. However, it is often more tangentially discussed. Furthermore, rather than have the techniques be inspired by the biology, it seems that so often the technique is thought of and then attempted to be fit into a biological perspective. This is the reverse of the ideal! It seems like there is plenty of time lost in this method, whereas if we look towards the abundance of information we have on existing biological systems that work, we can likely laser in on a solution with more precision and less guesswork.

I do think, at least, that there should be more effort in strengthening the biological analogy for modern AI systems. Consider the creation of the convolutional neural net, for example. Hubel and Wiesel determined a model for vision in the human brain, which prompted a cascade of inspiration culminating in the LeNet and its scions. I think it would be wise to bring such close biological analogy to many of the techniques we use today, but more importantly, the ones we will use tomorrow.

It is by this reasoning that I place a heavy emphasis on underlying biological principles to many of the ideas I present.

For many issues, it becomes hard to tell whether replicating a specific biological idea will be beneficial, as there is a concern that the biological way is not the most efficient, and simply took that form by chance or due to other constraints. There is likely an advantage to modifying specific aspects of the biological principles to tailor it to modern-day computational power. While we can attempt to make educated guesses of which aspects to change, I think it is wiser to assume the biology is right, find out what works, and then make improvements. In the end, though, the only way for us to find that out is to try, hope for the best, and take the sum of the efforts.

2.3 Complexity

The human brain contains on the order of 85 billion neurons. The largest artificial neural network (to my knowledge) created contained approximately 16 million. This simple disparity, along with the fact that our genome contains only 3 billion base pairs, has inspired many approaches to indirect encoding in the hopes that human-level performance may be found.

An important question arises: what proportion of our neurons are necessary for higher thought? If it turns out that most of out brain is dedicated to routine maintenance of the body or otherwise non-computational externalities, then the goal of human-level AI may be closer than we think. Or, are there elements of the non-computational proportion actually necessary? Are there other parts of the body beyond the nervous system that are necessary as well?

2.4 Novelty

Contemporary results give credence to novelty search and other algorithms which abandon the fitness function and rather promote solely diversity. Such algorithms do have undeniable benefits for solutions, an almost paradoxical effect. How does one get better at a task when it is completely unknown to them?

The fact that nature is reward-based with a fitness function of survival time means that this is the correct route to intelligence. To reconcile these facts, I propose that the sheer complexity of the natural environment actually causes variety to be the source of survivability. That is, though there does exist the fitness function of surviving long enough to reproduce, the mechanism through which this is best achieved is novelty. This can be seen in the vast ilks of life throughout our planet, for which differentiation was clearly beneficial.

So why do our fitness functions used in algorithms fail to promote novelty, causing the need for novelty need to be explicitly promoted? I believe it is because the variation mechanisms we use are not powerful enough to show fruit from the perspective of the fitness function, or because our fitness functions are not complex enough to see the value in diversity *as it relates to the goal.* Therefore, to compensate for this, we must artificially inflate the need for novelty (via novelty search, etc.) to produce better results.

Yet this is not ideal. Rather, we should strive to construct a system that addresses this issue rather than bandages it. And while it is no simple task, it may yet be the right way forward.

2.5 Unification

There is quite a large variety and quantity of ideas to follow. Some of them are simple implementation tricks I have thought of, but some of them are entire paradigms that would require serious time and effort to implement. I do not know which of these ideas are actually beneficial and which are superfluous. Furthermore, some of the ideas may be incompatible with one another.

However, each of them, as previously noted, has a foundation in biology. Therefore, at least some subset of these ideas share a sort of mutual dependence; that there is a master algorithm (a unification, if you will) that takes into account the changes or addenda that many of the ideas propose *simultaneously.* Of course, such an algorithm may be very far down the road, if down the road at all. But, I would suppose that given sufficient time and thought, a complex algorithm which reconciles these ideas and puzzles them together would tend towards converging on nature itself, and thus be capable of creating true general intelligence, given the compute power.

But even combining only a few of these ideas, or using a larger one as a backbone for smaller ones and existing methods, may be worth pursuing. Thus, many of the ideas should be considered with unification as an implicit possibility.

3 Interesting Points

A few neat issues and topics I think are worth touching on.

3.1 The Gap Anomaly

It has been seen that evolutionary strategies can more easily overcome barriers to optimality of small width, while failing to proceed down simply laid-out narrow paths, where SGD would succeed. This may seem like a trade-off situation, and likely is in some aspects. However, after stopping to think, one realizes that most problems which have nontrivial solutions are practically never obvious in pattern to a goal.

The innovations of bipedalism, opposable thumbs; these kinds of adaptations to the environment which allowed the subsequent growth of the species were not linearly achievable. Rather, they are 'punctuated' moments in evolutionary history which dictate the stop-and-go of forwards progress towards dominance over the environment and higher intelligence. This is a model that parallels the prior situation in which the barrier to optimality is thin but ever-present.

This thinking seems to indicate that for the true pursuit of intelligent problem-solving, at least in problems which are more complex in nature than to have a trivial solution, strategies that model evolution will vastly outperform those of gradient descent.

3.2 Triviality of Problems and Solutions

I read a paper (which I think was by Uber) that showed that a random search algorithm was competitive with SOTA results on some problems. They suggested that perhaps the baseline problems which we use to measure performance are simply too easy.

If a random search in the search space is able to converge on a solution as well as other methods, it seems to me that the computation done by other methods is purely algorithmic, clearly excess, and entirely useless from an intelligence perspective. That is, we are not producing agents that solve the problem, but rather just solutions to the problem, if that makes any sense.

It also means that the benchmark problems that we evaluate methods by are far too simple. This addresses a trend in this paper towards more complex environments; it is likely high time that we move towards truly complicated problems, perhaps focusing on those an algorithm is not good at rather than those it is.

Also, if you actually watch an agent preform a task, in some domains, the behavior is very clearly trivial and cheaty, and not at all intelligent. For example, algorithms often find and exploit bugs in the code of the domain. This also suggests that to promote intelligent behavior, more thorough and complex tasks should be the benchmark.

3.3 Learning through Living

I wonder if a human infant which was removed from other humans at birth (yet remained alive-if that is even possible) would be able to attain clever thought. Is it the case that humans have the inherent ability to solve problems, or the inherent ability to learn to solve problems?

Our existing EAs model learning by collecting genetic fitness over generations. Yet, there is so much to learning beyond simple genetics. How do we model that?

3.4 The Fitness Curve

The shape of the fitness function is of utmost importance to convergence on a solution-not only on the rate, but also the value. I will focus more on the value, as this is where truly interesting horizons lie.

Generally, if the path to a solution is too winding, it will not be achieved. Yet, this is almost always the case in any difficult problem. The best we can do, then, is try to ensure that steps are taken in the right direction, even if they are locally suboptimal. This actually goes directly against the concept of gradient descent.

What this means in practice parallels curriculum learning. An agent needs to be rewarded for solving small tasks in order to solve a larger one, even if solving this small task is not of use or makes the agent worse at the larger one. The paper on Enhanced-POET from Uber AI is almost exactly what I am referring to here. Of particular importance is the fact that other algorithms completely failed to solve the domains created and solved by the poet.

Thus, if we are to create intelligence, it seems we would need to do it incrementally and step by step. Yet, I am troubled by the lack of biological analogue here. Humans were not led step by step through the harsh environment of planet Earth, yet we stumbled upon intelligent thought through generations and generations. Is such a brute-force technique really the best way to go about it?

Actually, this aspect I think represents the greatest challenge to my principle of letting biology guide the computer science. There is, as I have said, the possibility that biology did not take the most efficient route to intelligence. The poet result shows that curriculum learning can vastly out-preform pure search. It may turn out that this is the case for finding intelligence, and that the world only found it in the way it did both because it was the only way it could and because it had sufficient time. If so, I foresee quite a challenge in our future as to find the correct way to produce intelligence without the guiding light of biology and our past. However, I do continue to think that such is our best way forward currently.

3.5 Mutation and Crossover

There has been a recent trend to abandon the use of crossover in favor of solely mutation in ES. It seems to me that this is actually a step away from biology and towards random search.

Biologically, the change produced by crossover is vastly more immediate than mutative change. While mutation is clearly necessary to create variation, I think crossover can be even more effective, given the prerequisite that new genetic material is created via mutation. Modulation of the ratio of variation which arises from these two mechanisms over the course of a run may prove performant.

Sexual reproduction is a highly inefficient trait, requiring the evolution of more complex reproductive structures and also the difficulty of finding and courting a mate. Thus, it must be of some use in achieving higher fitness. I would suggest this use is through crossover.

Most asexual mutation operations are pretty simple. And, in the case they aren't, they are intricately crafted. It seems almost like such variation mechanisms are untrue; almost cultivated towards a specific structure. I am not sure what true variation would look like, or if crossover accomplishes it, but I would wager it is a step in the right direction.

3.6 Differentiability

Often times, a new algorithm is theorized, and immediately it is attempted to be made differentiable so that it can be trained using the slew of SGD based techniques. Sometimes this adaptation is costly and even prohibitive. It seems to me that with the vast array of ES that can be applied to non-differentiable domains, there should be more adaptation of ES in these scenarios. I do not know why so many balk at the idea of EAs, but I think it is an error. In fact, there are many problem domains that simply aren't able to be differentiated (and most open-ended problems fall into this category). I think it would be wise to at least attempt to integrate ES into existing systems that struggle for differentiability. One day, perhaps we may forget about the need to differentiate at all.

3.7 The Goal to Survive

I claim name to a thought experiment called Buffkin's Phantom. If you were offered one dollar to experience significant pain for 10 seconds, yet be the same as you are now afterwards, would you do it? There is clearly a tendency towards no, but Buffkin's phantom asks why not. Generally, it wonders that if there is ever an end to things, then why do anything at all, if you will eventually be no different because of it.

Why do creatures on earth strive to survive? Of course, natural selection simply dictates that those that survive... survive, and thus there is a tendency towards creatures that live. But what is the point? I think sometimes that happiness is at least a valid argument. I enjoy listening to music, and so I will continue to live on. But what is my objective? And furthermore, what of all the animals and plants that live with no concept of happiness? For what do they live other than because they are alive? It seems almost futile.

I do not know what the ultimate goal-the ultimate fitness function-should be. But I do know that living is a prerequisite to higher thought. I only hope that this hidden secret of life is not another one, as it would likely prove disastrously difficult to implement.

4 Strictly Evolutionary Ideas

Ideas I have for implementing various alterations of evolutionary algorithms, with no regard towards neural networks specifically. However, if they are promising, they will likely be of help in neuroevolution as well.

4.1 n-sexual reproduction

Synapsis and crossing-over have been implemented in a variety of genetic algorithms to recreate sexual reproduction. The idea is that variation through the mingling genetics of two fit individuals will give rise to beneficial alterations.

There has been a recent trend towards abandoning sexual reproduction in favor of asexual. This seems to be in error, which I would justify by pointing out that sexual reproduction is vastly more difficult than asexual reproduction. This loss in fitness due to the need to find a mate, for example, clearly must be made up for somehow by sexual reproduction, lest it would never have come about. Also, on the same time-scale, variation through mutation (the main mechanism of asexual variation) occurs far more slowly than variation through crossover.

I think it would be quite interesting, however, to explore the other direction: tri-sexual reproduction, and more generally, n-sexual. It seems that the variation from considering two genomes would pale in comparison to drawing from n. The mechanics would be interesting to implement, but I think a good place to start would be modelling trait dominance (and other Mendelian properties) over certain alleles.

I am unaware of any natural examples of more than two genomes producing one offspring, but I would suggest that this is possibly due to the structural difficulty of such a process, and the likely probability that we have arrived at the evolutionary point at which we are able to consider it well before it would happen. Also, the overhead of finding mates and evolving n-sexual reproductive structures is mitigated by the algorithm-based nature of our implementation, which perhaps means we can reap the benefits of it without facing the drawbacks.

4.2 Continuity and lifespan

The common approach to an EA is to split the evolution into discrete generations. While this does provide a nice benchmark number for experiments, it seems like needless structure, and certainly does not model what happens in nature.

What I think would be more realistic is for the algorithm to run continuously, with each individual responding to the passage of time separately. Under this paradigm, individuals would have a *lifespan* proportional to their fitness. During this lifespan, their reproduction can be continuously modelled, which would also be a function of fitness.

The process of evaluation is necessarily different in this model. Where does the fitness of an individual arise? Perhaps instead fitness is a function of lifespan, and performance over lifespan leads to reproduction. This would actually eliminate the need for a fitness function, or implicitly make it survival time. Also, initially, the reproductive rules must be relaxed else the entire population die out before they can live long enough to produce offspring.

I think that to maintain elitism, it may be prudent to give the most fit individual an infinite lifespan until it is replaced, at which point it would start to age. This may, however, be unnecessary, but it would address the initial petering out of population.

I am yet to come up with a reason as to why this is actually any better than discrete generations.

4.3 Carrying Capacity and Population Size

In a continuous model like the one above, or even a classic generation-based model, the constraint of population size is analogous to the carrying capacity of the environment. Yet, as seen through the exponential population growth of humans in recent times and other numerous examples within nature, it is often the case that certain evolved traits unlock an ability to use resources more efficiently. Therefore, it would seem like a good idea to impose a *maximum* population size constraint rather than constant one, the value of which is in proportion to the fitness of the population or species.

Generally, then, the maximum population size of any one species would be a function of their fitness. There are various implementation details of this which will probably produce different results; perhaps newly born individuals which would bring the population count above the maximum only persist if their fitness is significant, for example.

In any case, the shape of this population-fitness curve is of utmost importance. Clearly, a positive relationship is required. What I think would be more interesting would be to have various plateaus, which require large fitness increases to surpass, but produce large benefits to the max population size-think a shape like $\lfloor f \rfloor$, for example. This could model and promote real-world innovations such as opposable thumbs or the use of tools which produced huge benefits to the fitness of the population. I think this also parallels the idea of punctuated equilibrium.

Of course, different species must have different functions to represent their different niches and adaptations.

4.4 r-select vs K-select

The protection of novelty is an important current topic. NEAT, for example, uses speciation to maintain not yet useful innovation. This is definitely useful (as shown in the paper), but there are further ideas that can be applied here.

The protection in the aforementioned nature is against genetic non-viability, and so is not directly analogous to protecting genetically viable but young individuals. In terms of K-select species, it seems interesting to implement (perhaps via fitness sharing) further protection of novel innovations through parenting. This would likely be of more use in a continuous model. The seemingly equivalent protection for r-select species is simply mass-produced offspring giving a high probability of survival. In implementation, this actually just takes the form of normal reproduction of a normal number of offspring, and I believe already exists in most current algorithms.

4.5 Competition and Symbiosis

NEAT justifies speciation by filling niches in order to only compete intra-species, yet there are so many varieties of symbiosis which play a role in ecosystems. The difference between niches in NEAT is entirely superficial; the different species have an equivalent fitness evaluation-they don't actually perform different functions.

I wonder how an algorithm could go about modelling relationships like mutualism, commensalism or parasitism, or in general other inter-species relations. This may require a sufficiently graded fitness function (or novel search) so that different species can solve different parts of a solution, and would also entail that the ultimate solution produced be multi-individual (unless some form of endosymbiosis is implemented, see later). Even simple predator-prey relationships and direct niche competition for resources could be a powerful motivator to induce useful disruptive selection and produce unique, particularly novel, solutions.

Yet the intra-species relationships also have a place. Two individuals of the same species fill a closer niche and thus may compete more directly. However, there is an entire taxon of species that rely on herd or swarm behavior, and exhibit hive-mind qualities. The production of a multi-agent solution which parallels the intricate jobs which keep a bee colony alive would be something quite interesting, and possibly effective, indeed.

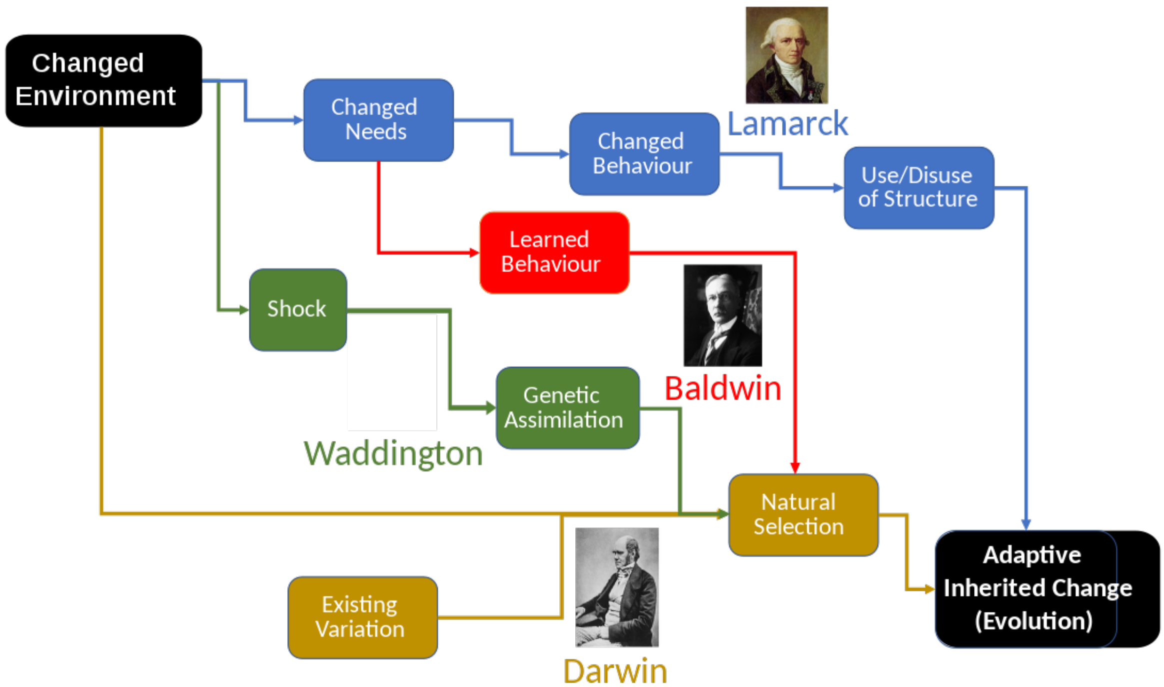
Intra-species interactions allowed humans to, for example, specialize in labor. Yet we also required the interspecies interactions afforded by livestock, horses, etc.

4.6 Natural Disasters

When a volcanic eruption destroys sections of a forest, sometimes what follows is an adaptive radiation. The genetics at work here are too nuanced for my understanding, but the general idea is that in such a different tabula-rasa environment, there is an augmented possibility space for individuals to succeed. I think that implementing natural disasters and other large events may contribute to convergence on a solution, as counter-intuitive as it seems.

Large scale extinctions like the K-T extinction might simply take the form of starting over with a new run. Realistically, though, it isn't a total reset like we do now; there is at least some genetic material conserved between events. This also seems like an effective way to prune the search space and maintain computable complexity. Lastly, this could play to the theory of punctuated equilibrium, which merits further investigation.

4.7 Alternatives to Natural Selection



Sometimes a picture is worth a thousand words. Exploring implementations of various parts of the diagram above may yield cool or useful results.

4.8 Cold Coevolution

During competitive coevolution, it is possible for one species to evolve more specifically to coutner the other species than to achieve a goal. As noted before (4.5), this is possibly advantageous, but it can lead to individuals that are not as optimized for the solution as the changing of both species causes unforeseeable disturbances in the convergence of the algorithm.

Perhaps coevolution could rather take place by evolving only a single species at a time, keeping the genetics and behavior other species frozen (or 'cold') while only the currently considered species is changed (or 'hot'). Ideally this would prevent destructive interference between individuals, allowing the hot species to move beyond just competing with the cold specie(s) and towards increasing relative fitness-thus maintaining the novelty and difficulty brought about by competition, but also focusing on solving the problem.

4.9 Genetic Drift

Akin to 4.6, attempting to model the effects of genetic drift may aid in convergence. For example, the founder effect could be replicated by constricting a segment of the population from breeding with the whole (and altering their fitness function). This could also aid in efficiency. The bottleneck effect I think is encapsulated in 4.6. Yet, there are other topics in genetic drift that may apply to EAs.

4.10 A Plastic Environment

I think one of the most significant issues in EAs in general is the static nature of their fitness function. The world around us, whether through mechanisms of competition and biotic factors, or multitudes of abiotic factors, is in a constant shift. Here is a paragraph I wrote some time ago:

Had our environment remained constant, we may still be simple creatures. Perfectly adapted to this exact environment, perhaps, but, depending on the complexity of it, still as simple as a DNN trained to solve an Atari game. It was the changing of our environment—the accumulation of traits and features that can be adapted to shifting fitness requirements—that eventually allowed for the necessity and thus achievement of higher thinking. It is through this that a changing environment evolves a sense of plasticity; that the more attuned to approaching change a species becomes, the more able to be clever and adapt. In this lies the potential for true intelligence.

In *On the Origin of Species*, Darwin noted that "Changed conditions of life are of the highest importance in causing variability". I think the general gist of this as applicable to EAs is that the fitness function should be ever-mutating. This is, I think, the goal of ideas like 4.5 that invoke other species in the evaluation of a singular-because other species are dynamic, this brings an ephemeral nature to the fitness. But, there are more biotic factors that play a role in this, as well as abiotic factors.

There is a paper on Enhanced POET from Uber AI labs which I think is really heading in the right direction with regards to environmental variation. I think the variation results produced in this paper should be construed as abiotic variation. Specifically, there lies abiotic variation in the simple fact that the world is vast relative to the experience of one species. Different regions have different topography and resources and thus produce different fitness functions; akin to the changing terrain in the POET paper.

Yet often species are limited geographically. One would think that if environmental variation was key in evolution, all extant species would be cosmopolitan (i.e. geographically ubiquitous). I would note that humans are generally ubiquitous, and also the most intelligent known species, which in fact counters this point by implying that abiotic variation is actually a stepping stone to intelligence, and the reason other species are so lacking in it comparatively.

Lastly, I think that when a specific species encounters different environments and attempts to adapt structures to fit these environments, it can often hinder them from living in altered environments. This must in some way be the cause of the limited region of so many species. After all, there is no one physical feature set which suits both the desert and the ocean. In fact, it becomes apparent that the only way to realistically adapt to survive in a varying environment is not structural at all, but is instead *mental*. Because it is impossible to have one structure and simple behavior be fit enough to survive in multiple environments, there is a need for an individual to think and solve each environment it encounters, taking into consideration the tools they do possess and using them to maximize survival. But this is no simple task, and requires intelligent thought.

Through this mechanism it seems that variation of environment is the root of intelligent thought. This I think might be the most important result contained in this paper I have written.

I will continue to ponder this. For example, what role does structural evolution play at all, then? Is it simply to generate the tools which are best fit for the average environment, thus maximizing the ability to use these tools under smart thinking? If this is the case, we may be closer to higher thought than we think. We are likely at a point at which we are intelligent enough to derive these tools without evolution. Thus, giving a robot some hands and sticking them in a changing environment might be enough to make them conscious, skipping the billions of years needed to figure out that opposable thumbs are important.

4.11 A Plastic Genome

There are a variety of aspects of most multicellular genomes that go beyond simple rote expression. I think it would pose an interesting venture to attempt to create an encoding that embodies things like transposons or epigenetics, as well as splicing characteristics and intron/exons.

5 Neuroevolutionary Ideas

Ideas that apply more specifically to neuroevolution, though may have other ramifications.

5.1 The Circle of Life

To address the issue in 3.3, I propose that the most logical step, and strongest biological analogue, to reproducing learning via living to be the interleaving of EAs and gradient descent. The general idea would be to evolve only structure, which represents genetic learning, then train individuals using SGD, which represents learned behavior (potentially Hebbian-like), and then rinse and repeat.

This seemingly could be applied pretty easily to existing structural NE techniques. But I do wonder how it would fit into a continuous model as proposed in 4.2. Perhaps the learning is done such that it is proportional to the amount of time alive, as in one backprop per timestep?

The amount of training done per iteration between evolutions probably plays an important role. One would think to just train to convergence (which may well work), but it may be wise to try other degrees. Training a small amount each iteration may help, for example, to achieve that so-desired well shaped fitness curve, leading to more incremental yet positive structural changes. Also, transferring trained weights to offspring would both help maintain the training and also mimic learned behaviors in young.

I think that one of the most important facets of this idea is that it lies almost perfectly at the intersection of differentiable problem domains like deep learning and ES like neuroevolution. If a result is produced here it could work wonders towards bringing these communities together.

5.2 Musical Chairs Effect

I recently read up on Cascade Correlation Networks. One of the inspirations behind them is the musical chairs effect-the idea that each neuron has a specific job to fill, but by training them all at once, there is a sort of bum rush for every job by every node. The goal of CCNs is to reduce the search for a job to one node at a time, which happens to decrease the training time while maintaining roughly approximate accuracy (or did, anyway).

The problem with this I would think is that there are almost inevitably a large variety of jobs that are not solvable by a single neuron. Many specific issues probably require two or more single perceptrons in order to resolve. For CCNs, the implication is that these problems cannot be solved under the current model, and would require the adding of two or more dependent neurons at once rather than a single one.

For neuroevolution, however, I think this takes more of the form of the necessity of node-add mutations to be able to add more than one neuron. The probability of such a mutation being beneficial feels astronomically low, but is necessary to solve such a 'dual-node' problem. Also, under a model like 5.1, the mutation would likely be either embraced as useful or weeded out very quickly; no harm, no foul.

5.3 Recurrent Representation

Memory is clearly a crucial part of human intelligence. The exact role it plays continues to be nebulous, but we should strive to model it nonetheless. Simple RNNs are often used as an analogue for memory solely because their recurrence allows transfer of past information to the future, but this transfer is limited by the vanishing and exploding gradient issue. While the LSTM model does take steps to overcome this, both of these methods, and recurrence in general, seem to lack an elegance that may be required to replicate human use of memory; rather they have this predetermined, algorithmic bottleneck that would feign to scale well beyond time-series or other simple data formats. I am not really sure how this could be amended. I think that the natural progression evolved in a Markov brain is definitely more powerful. In general I would suspect that certain aspects of memory are instant-dependent, requiring interesting new models like the spiking neural net. Using autapses in an snn may be cool. But there are other aspects which are more static and almost inherited. Genetic memory would be a challenge to implement, yet may fruit fantastic results.

5.4 Super-HyperNEAT

I think that evolving both the CPPN *and* the substrate could lead to interesting results via the Hyper-NEAT algorithm. Perhaps this could take the form of cold coevolution, or even classical coevolution (though it may be a bit wacky, for lack of a better word, if both are changed simultaneously). Furthermore, this might be a great way to model epigenetics.

5.5 Speciation

While I applaud NEAT for its technique of using innovation number to determine individual compatibility, I must wonder if it is the best idea. I should do more thinking as to an alternative to innovation number, both for its chromosomal compatibility testing capability and the speciation capability.

Regarding the latter, I think it could do better to parallel the actual mechanism of speciation. While it is true that two individuals are generally considered different species when they can not produce fertile offspring, the NEAT speciation seems to be exclusively sympatric. I wonder what an allopatric equivalent would be for neuroevolution.

My main thinking here is that two new species should be produced when they are sufficiently different due to being separated, not be separated due to being sufficiently different.

5.6 Efficient Evaluation

One of the common arguments against ES is that evaluation tends to be computationally prohibitive. I propose that it likely unnecessary to evaluate a model on an entire training dataset, just as an individual does not face every possible challenge in life*. In fact, it may be prudent to segregate the training data into correlational groups, and then train different subpopulations on these smaller groups; an artificial speciation of sorts. This could reveal underlying methods and data features that would otherwise remain hidden, and the individuals could be bred somehow to form an ultimate solution.

This may just be a form of overcomplicated curriculum learning. Also, separating the data may prove to be difficult. In terms of a plastic environment, simply placing the initial population into groups in different areas would achieve this effect, where true ability would arise when one species learns to live in other and then all environments. I suppose gradual environment barriers may be of importance. But I digress.

*I now see the Limited Evaluation EA, which mostly confirms this.

5.7 Evolutionary Sub-problems

To mimic certain traits (speech, for example) it is probably wise to research particularly how and why that trait evolved. Comprehension of speech is unlikely to evolve in a system in which individuals need not interact, for example.

5.8 Inherited and Learned Behavior

The idea of behaviors being either inherent or learned may be of use in NE. While I have given little thought to it, I would say that the general analogue would be that, in a cyclic algorithm like 5.1, the weights at birth would be inherited behavior, and they can be changed to reflect learning.

But this analogy lacks substance at the human-level. Babies are not as capable as their parents when born. Also, what would the equivalent of imprinting be?

5.9 Endosymbiosis and Self-replication

When a solution becomes sufficiently competent at solving a certain part of a problem, the mutation of the structure that allows this which takes place when solving the rest is generally detrimental to the convergence of the algorithm. Yet, the search space still needs to be searched.

I read a paper from Deepmind on hierarchical representations of NNs which evolved complex CNNs. Generally this is difficult due to the minutiae of the convolution operation, but the paper essentially froze small functional subunits and used them as building blocks for evolution, avoiding this issue.

Therefore, when an individual is good at a small task in the environment, it may prove wise to freeze it and assimilate it into another more complex organism (especially if the latter is yet to accomplish this small task). The inspiration for this lies in endosymbiont theory, but the analogue may extend to all cellular organelles.

An issue this poses is the need to differentate between 'tasks' in the environment and when two subsolutions solve different subproblems. The latter half of 5.6 may prove to be of use here. Or, one could make use of novelty search to ensure that solutions are of differing importance. Ideally, the algorithm would take the form of novelty search until plateau, then employ endosymbiosis and fitness to assimilate and complexify, then rinse and repeat. This could produce very powerful results, specifically with regards to open-endedness.

Previously I thought it possible to form a multi-agent solution using different sub-solutions, and this still has interesting implications (think beehive), but endosymbiosis may allow single-agent solutions. I wonder if this is even a relevant distinction?

Furthermore, most mutative operations in use are simple and singular. Yet in biology, mutations often have complex effects and replicate entire developed structures; i.e. polydactlyism. Perhaps once solutions begin to plateau, the scope of mutation may be expanded to more broader structures, implicitly implementing self-replication. This would seem to bring individuals into more and more advanced development classes. Also, this is something I think is tacitly implemented in HyperNEAT.

5.10 Evolving Inputs and Outputs

Most, if not all of the algorithms used to date have fixed input and outputs structures. Yet, through evolution, creatures create their own sensors (eyes, ears, ...) and actuators (hands, tails, vocal chords...). This is pretty tough for me to think about. I think that in order to allow alteration of IO structures, one would require an environment that is sufficiently complex, implementation wise; an environment like Earth that is truly open-ended.

Unfortunately, it seems to me that any attempt to allow creation of IO will bottleneck it too much. But I suppose I ought to keep thinking. Fortunately, it might be the case that we can take advantage of the fact that evolution already produced eyes and legs, and assume this is the best case scenario, allowing us to bypass the need to evolve IO. It is sort of boring, though, and there are likely very neat and creative alternatives which we may never know.

6 Undeveloped Keywords

The great filter, momentum in evolution, Kimura's neutral theory. modern synthesis, extended synthesis

References

[1] P. A. Vikhar, "Evolutionary algorithms: A critical review and its future prospects," *2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC)*, Jalgaon, 2016, pp. 261-265, doi: 10.1109/ICGTSPICC.2016.7955308.