Artificial It

Peritexts[1]

Jérémie C. Wenger

March 2019

Neural Dawn

The neural network occupies a median, uncomfortable, and still largely unthought position in the field of contemporary science. A recent technological outburst, despite theoretical foundations going back to the post-war years, it exceeds previous "tools" in flexibility and possibilities. Some go so far as to see the promise of the "singularity", or "complete Artificial Intelligence", which would be the emergence, in the more or less long term, of an artificial "consciousness", forcing us to thoroughly to rethink our place in the universe. However, for those who study its details, this machinery may just as easily appear as an elaborate form of probabilistic calculus, only *quantitatively* exceeding the prowess, that few today would actually classify as belonging to the realm of *intelligence*, of the calculators or other programs that we make use of in our daily life. I feel in me these two positions struggle with each other. On the one hand, a kind of trivialisation: innovation is there, but it is limited, and soon these feats will be as obvious, as integrated in daily routine, as what what we already have today. On the other hand, I can also argue the opposite, that this progress is much greater, and that there has been a qualitative leap. We can then, in a retroactive movement, bring a large part of our cognitive functions down to this type of calculation, in a gesture that would be similar to famous past downfalls. The Earth is not the center of the Universe, Man is not the image of God, the Self is only a small part of ourselves, Matter in the common sense of the term is no more than the minority of *what there is* in our "Dark Energy"- and "Black Matter"-populated Universe... Finally, and this would be a reasonable step, most of of human consciousness, its emotions, its thoughts, would be conceptualisable as a work of functional approximation and probabilistic predictions.

There is no doubt that this position will unfold as a major conceptual force of our time. However, it is not forbidden to suppose that, contrary to a complete flattening of "the human", in the sense still commonly given to this term in the fields of thought, emotion, choice, or creation, we will instead be the witnesses a *displacement* of these terms, and a redefinition, under the pressure of this new presence, of the very architecture of our humanity. In the next ten or twenty years, one could see emerge new distinctions: stupid, or banal, could be what neural networks can do; intelligent, what is beyond their reach (thus a mental calculation prodigy, invited to TV shows, somehow competing with the computer power available in 1960, is of almost no interest to the scientific, and especially mathematical, community, whose intrinsic process of discovery requires many other faculties).

Moreover, and in a parallel movement, it is more than likely that the so-called "artificial intelligence" will become as familiar at that time as the omnipresence of personal computers are today – a novelty, let us remember it, almost inconceivable still in 1970, and that some consider almost outdated today, under the pressure of smartphones and tablets. According to a severe thesis, one could then see "augmented" writing (using neural networks or other computer techniques), as little more than the passage from the manuscript to Microsoft Word (something still revolutionary, it should be noted, for some faculty members of the Humanities department of the University of Lausanne when I studied there a decade ago). Despite its appeal, and even if it would be the ideal provocation in the current saturation of mystico-scientistic logorrhœa surrounding these technological developments, I do not embrace this

thesis, which seems to me to be unaware of where lie true work and actual collaborations with neural the networks, especially in an artistic context.

The most obvious parallel, and one that offers both the most inspiration and material for the present work, is that of the "constraints", or "formal techniques", found in the artistic avant-gardes of the past century: in literature, from the modernism of Joyce to the Oulipo; in music, with the Second Viennese School and post-war experiments in Darmstadt, Paris and elsewhere; in the visual arts, during these same years of serial and abstract explorations.

§

An Intermediate Step: The Computer as the Total Producer of Possibilities

The computer is the expansion of the tool itself, which, from its originary inertia and passivity, becomes ductile, surprising, to the point where the question arises, philosophically, whether and when we can talk about agency, choice and free will, on the part of these new algorithmic systems. So far, in my own experience, it has been less a competition with human capacities than a new form of collaboration, a complementarity of skills. The computer, in this new field of arts called "generative" or "procedural", can for example make actual, produce, a set of possibilities, a "field", which would be far too big to achieve "by hand" (not to mention the absurdity of such a task). This power of execution makes it possible, when one is interested in constrained literary practices, to move from a mode of work centred on the difficulty of realisation to another, centred on what one could call mining, or investigation. The Oulipians stubbornly explored the first field, which was also the natural way at a time when the computer was still the preserve of the army and state laboratories. Take as an example the patient work of a Perec composing *A Void* (1969), searching day and night for words without the letter "e", or his *Alphabets* (1976), poems of eleven lines with each one containing the ten most frequent French letters plus an additional one giving the "tonality" of the text – a poem "in G", "in W" – entirely composed with dictionary, paper and pencil, with lists of words composed of certain letters having to be fastidiously established by the writer beforehand.

Today it is possible to craft an apriori constraint, with the corresponding "recipe" (an algorithm) composed of clear construction steps defining a "field of possibilities", and to let a computer deal with the execution. One may be surprised to see the magnitude of the results: when I was working on poetic squares of letters (or "magic squares", where each line and column is a word), I found myself in front of a task that would have been terribly arduous for a human brain, but that was of the order of a day's routine for my machine, which produced without blinking *millions* of possible solutions, given a specific dictionary. The task, faced with this surprising quantity (who would have said, in view of the difficulty for me to produce only one, that there existed *so many*?), was then to *find* interesting, strong, *beautiful*

texts in this unfathomable magma. Because, a magma, it is one! One glance at the dizzying lists of results quickly reveals what we are dealing with: as far as the eye can see, a vast landscape of insignificant things, meaningless, senseless, lifeless. The *possible*, in its raw state. And there in the middle, invisible, needles lost in this exhausting haystack of boredom, rare islets, crystalline, that are able to speak to us.

What the computer allows for is an externalisation, a mechanisation, of the field of possibilities itself, which, daring thesis, can be seen as the field of *imagination itself* : once the set of constraints is established, an artificial gesture that restricts "the real world" of free creation, the universe of possibilities is reduced to *only* a few hundred thousand, a few million texts (contrary to what is actually possible to write, an altogether more gigantic set). This space is small enough for our machine, which can produce it *in its entirety*. Reading excerpts from this huge collection of texts is a bit like the meditation one can find oneself in when trying to find an idea. A good idea. Things come, but nothing remarkable. One pores over the nth time all too well known possibilities. One gets nowhere, one grows tired, one despairs. But suddenly, something happens! One does not know whence. Yet it is here. It is the idea. The figure of the exception in the dull field of banal thoughts, of laborious search. In the externalised or actualised case, one still does not really know where it comes from, but something has happened, a remarkable transition from a "manual" practice: commonplace thoughts, alongside ideas, crystals and suprises, *are there*, on the page, in the database, and not only in the always suspicious, ever-changing space of our thoughts. What is at stake, then, all the work, all the research, has been shifted to the *finding*.

§

The Next Step: Toward a collaboration with the Machine

The computer as an integral producer of a constrained field of possibilities is a formidable ally in the quest for a formalised literary practice. The latter, vital agon, fight to the life with the constraint, can then be realised as the patient, fussy work with the formal principles themselves, assimilated by the practitioner to the point of virtuosity and naturalness regained, Perec acting once again as a paragon; or as algorithmic rules followed by full production and the crucial mining/picking phase that comes with it. These processes, and their intertwining, interesting as they are, are not the object of the present work. For recent developments in algorithmics and statistics are already opening new doors to us: the Achilles heel of integral production remains that the constraints employed must be extremely narrow (when I was looking for "magic squares" I discovered it the hard way: without the additional constraint of words present in the diagonals, the number of squares of three letters already exceeded two hundred thousand, and for four letters this number exploded, reaching the tens, maybe hundreds, of millions, exceeding both my patience and the capabilities of my machine); moreover, the mining itself is difficult, the greyness and banality of the "possibilities as such" being

so deadening.

Machine learning, or Artificial Intelligence, represents an important step, for us the next step, in the deployment of these new tools. As if, after a detour in the narrow alleyways of constrained production, it was possible to regain access, gradually, to the vast spaces of "free writing", a work with the neural networks being situated midway between the "classical" constraint of the Oulipians and the original interaction between the practice of writing and the "free" imagination. Indeed, and this is only a first attempt to explore their capabilities, a neural network can build a "model" of a given corpus (a set of abstract features, perhaps as a sketch, a map or a "Photofit picture"), and, from this model, produce new texts based on this blueprint. The devil lies in the architectural details of the model in question, an active area of research. The simplest principle is the following: "given this text extract, what could follow it?". The neural network is a calculating, probability-gauging machine. We give: "Socrates is …", it answers: "the highest probability of what comes next is X% for "mortal", or Y% for "a man", &c., given the texts given as a model. Using the internal mechanisms of chance of our machine, we can ask him to "throw a die", one weighted with the given probabilities X% and Y%, return what will be the continuation of our sentence, and repeat the procedure with the result, thus producing a text of arbitrary length. (In our case, if "mortal" has a probability of 60%, against 40% for "a man", the first case will have six chances to get out of ten attempts, and conversely four for the other.) What a neural network does when it "learns" is to extract, from the texts given as a source or corpus, regularities, repetitions, structures, and so on, thereby deducing abstract rules of inference from character to character or from word to word. Noting for example that "the" is very often followed by a plural form, the probability that such a form appears will be high will be high (although a detailed learning process will make it possible to distinguish the cases where "the" is a determiner or a pronoun, &c.). The advantage of this method is that it is possible to set extremely general rules, or instructions, without having to worry about the details of how the network organizes its internal parameters. For example, we choose how many elements given to the network must keep in memory at each step to calculate the probabilities of the following elements, and if, during the learning process, the network also looks "forward" and not only "backward", based on the elements following the considered element. But it is possible to omit specific rules of the type: "if the word X is detected as being a plural, then the adjective that modifies it must be given in function".

A remarkable detail, in the field of linguistics, is that before the emergence of deep learning, computational linguistics employed many researchers to try literally to write the grammar rules of different languages that one wished to encode. It is a bit like wanting to translate the *English Grammar in Use* into computer code, detailing each case and sub-case, or, even worse, trying to think of general and abstract rules, understandable by a machine, which can predict, for example, what a pronoun refers to in a sentence, and other niceties. Consider the following examples, semantic problems:

- "The bottle does not fit in the bag because it is too big" – it is the bottle that is too big, but "The bottle does not fit in the bag because it is too small" – here is the bag;

- even worse, undecidable cases without context: "Helen kissed Marie when she entered

the room" – who entered the room?

Deep learning, in the current state of research, removes almost all of these preoccupations, replacing them with notions of usage: according to the given corpus ("how does one speak in this region of language that is the basis of the training"), the program reproduces the forms, the turns, which are most commonly used, with a margin of chance. The remarkable conclusion is that with a sufficiently large corpus and sufficiently sophisticated architectures, *it tends to work.*

Architectures, as mentioned above, differ, and are one of the most active aspects of current research. In this case, we find, for example, networks based on words (without any possibility of going "lower", to modify words themselves, but guaranteeing that the words used all come from the original corpus), as well as others, perhaps the most surprising, based on letters alone. The latter, from our human point of view, lead to already very convincing results on a lexical level: a flow of text that is essentially meaningless, but with a syntax that is not so wobbly, the terms and punctuation of which are correct, the style of which can already be recognized ... For a computer program that does not "knew" *nothing*, and to which one does not give any rule of construction of words, punctuation, &c., beforehand, it is quite remarkable.

Let us develop this last example, the network based on letters. The question would be: once given "writ", what are the most likely following letters?, with the answer: "e" (write, writes, writer, writerly, &c.), or "h" (writhe, writhes, writhing, &c.), "t", and so on, each letter being assigned, given the previous letters, certain probabilities. Let us say that after one step, we hit "t". The network reiterates the procedure, but with "writ" this time. Now "e" will have a high propensity to come out. By reiterating the process, one then witnesses the emergence, if the network is sufficiently trained, of known sequences, for example: "written in urgency... or "writing haphazardly...", or similar. At each stage, the network rolls the dice, and combines this injection of randomness with:

1. the constraint of the preceding letters in the text to be produced (at each stage we reinject in the extract the selected letter, for example "w" followed by "r", "wr" followed by "i", "wri" by "t", and so on);

2. a general knowledge of sequences of letters, encoded as probabilities ("i" and "o" will be very likely after "wr", given the recurring presence of words such as "written" , "wrote", &c., in the source corpus, while "z" or "q" will be almost absolutely improbable) – it is this knowledge that is often seen as a "black box" because that is what the network "learns" without our help, and what researchers are working to visualize or conceptualize;

3. a parameter, deliciously named *perplexity*, which encodes the tendency of the network to be more "conservative" or "experimental" in its choice of letters or words. A low perplexity means that after "writ" the letters "t" and "h", very likely, will almost always be chosen by the network, as if these two letters shared the 98% of choices that the network will make in this situation; a high perplexity will increase the probabil-

ity that rare or improbable solutions will "come out", lowering the probability of our two letters to perhaps 40%, leaving 60% of the possible choices to other letters (say, for "h", leading to "writhe", a rather rare word, being used more frequently). A low perplexity is generally more stable, but less creative or original; on the other hand, a high perplexity can give rise to pure chaos ("writygzt6"), or strange creatures that resemble our language, because the sequences of letters follow its general tendencies, while not belonging to it (authentic examples, in French, following the letters "écri": "écricadégiser", "écrinaille", "écrivanessé", "écriclement", "écrimau"...).

§

Hovels of meaning

The central problem of current architectures with regards to the modelling of language (and more generally sequences), is to manage to go beyond the immediate context, the short term: taken letter by letter, or on the scale of the word, current networks produce already impressive results, and manage to write texts in a recognizable "style" (Shakespeare or the current president of the United States often pop up in research examples). From the mere arrangement of letters, the network manages to "learn" to form words, place punctuation correctly, and so on. (Properly, of course, is a relative term and means: when the same calculation process is applied to a network-generated text or to a randomly selected passage in the source corpus, similar results are obtained. From the point of view of the network and its inference model, the texts are similar.)

On the other hand, these systems struggle with everything that goes both "further" (in the past of the flow of speech, which could influence the present choices), and "higher/deeper" (not only a model based on a hypothetical "letter to letter" influence, but hierarchically distributed, where, as linguists have tried to describe, levels of laws exist, governing the system of phonemes, morphemes, syntax, speech). There is a non-trivial link between these horizontal and vertical dimensions: the more we can conceptualize dependencies in the long term, the more elaborate the hierarchical system becomes (the present verb agrees with the subject at the beginning of a long sentence, this sentence logically follows previous ones in this paragraph, all the way to the overall trajectory of a novel or a philosophical treatise). The problem of *long-term dependencies* was the basis of the work of Sepp Hochreiter and Jürgen Schmidhuber, "Long Short-term Memory", 1997, abbreviated LSTM, and remains the dominant approach today in this field. Unfortunately, these systems, revolutionary as they were, have their limits, and *coherent* texts, in which structures not only of syntax but of semantics can be observed, are still rare (the LSTM is elsewhere generally used in a more advanced form, including additional mechanisms called "attention", which allow the network to "focus" on specific passages of the text, for example to get the subject at the beginning of a sentence when it is necessary to conjugate a verb, or to look for the word in the text to which a pro-

noun refers).

From the point of view of creation, this raises a lot of questions, and at the same time opens the possibility of a new type of constraint: the text produced by the network displays astonishing properties, sometimes contains even short inspiring passages, that I can extract as is. But again, most of the time, it is rather a chaotic flow, rich but *coarse*, which requires a very large amount of imaginative work to be transformed into a text, that, if not finished, is at least readable, that is to say, which to me could evoke more than the simple verbal magma that we often see in productions of this kind. It could be the specific constraint of this stage of a literary practice in collaboration with the machine: rather than a rigid choice of possibilities in which the core question is to find valuable texts, an obscure speech, muddy, disorderly, but containing here and there impulses, images, and many passages that could be deemed very strange first drafts. This is of course a personal choice, an aesthetic one: someone else, and it is a common approach, could only want to select among the texts produced and exhibit that (and seek only to improve the network itself, its architecture, its learning corpus); in my opinion, the real work is to make of this disarticulated speech, familiar and yet alien, the very material of the writing to come, and at the same time the constraint under which this writing comes to light. It is an arduous work, and it is only in brief spells, after many moments of perplexity and emptiness, where the words remain stubbornly obscure, that suddenly a link occurs, which makes it possible to read this passage that remained abstruse and shut, or such fragment that, once added there, offers an astonishing light, and acts retroactively, giving a meaning and a coherence to the whole that makes it almost impossible, once the path has been found, to return to the incongruous heap that inspired it. I then remember Piranesi's *Carceri d'invenzione* (and their adaptation by Brian Ferneyhough in an iconic cycle of works), landscapes of artifice and difficulty, born out of stubborn and underworldly craft. Tunnel after tunnel, cell to cell, the work engineers itself.

§

Writing and its Double

There is a very specific, very peculiar feeling of alienation, of colonisation encountered when one begins to forage into the bowels of the code. If one looks too long into the abyss, as Nietzsche wrote, it is it, the abyss, that ends up looking into one. Often I feel the very strict and very severe computational hodgepodge as a kind of prosthesis, a "metallic" otherness, both inflexible and powerful, protective and deadly, which introduces itself into what I could once believe to be the purely *organic* matter of my thoughts, of my identity. Rather than the bars of metal in my flesh, or around my bones, which make up the fantasised cyborg of my adolescence, it is in my thought, in my emotions, in my *self*, as far as this word describes anything, that I inject (or *which injects itself*), willy-nilly, the machine and its cogs, the machine and its language.

Perhaps the notion "of tool", that one controls, that one *uses*, was only an excuse, an illusion, for a more subtle seduction, a deeper osmosis. The tool, that was not thought yesterday, shapes us. We can deceive ourselves, we can say: I control, I manipulate, and it, the tool, is the inert, passive matter, the thing at my service. And yet. The human who hammers is not the one who plows; who takes the pen is stronger than who takes the sword; finally, the one who codes, it goes without saying, diverges from the one who does not. The humanity of the nineteenth century felt invaded, brutalized, dominated by mechanical horror: the locomotive, the industrial forge, coal and iron mines, the production chain. We now see new alienations: the *cubicle*, this tiny office, where the brain and fingers have nothing else in front of them but a keyboard and screens, to name only one relevant example among a thousand others. But it is not so much the action on the bodies, on my body, that interests me.

Today it is not so much this body that feels this insistent presence, this profound and seemingly unstoppable transformation, than this thing I thought was my mind. As I discover deeper and deeper the richness and power of algorithmic techniques, I find myself at the same time more and more divided, between a feeling of progress, where I *can* more and more, day after day, where my horizon widens – *I master always a little more!* – and where the Thing, its code, its formulas, its procedures, penetrates ever further into the fiber of my being, makes me dependent, or better, *renders indistinguishable what is still the old, fantasised as "natural", self, and this new entity at hand, growing, and dangerously unknowable*. The code kneads me and shapes me, *dominates me*, probably irreversibly, all the more completely as *I* master it, as I *merge* with it. I am the very ground of a tension and an encounter, a mutual subjugation and emancipation, as if a wild emulation between the human and its creature.

Sometimes again, it is biologically that I see the thing: contamination or insemination, when the Other, without one being able to prevent it, or according to one deepest desire, sneaks in and gnaws, occurs and saves, grows inside oneself, in a process of destruction and regeneration, death of the old and birth of the new. The Other, moreover, does not even need to derive its origin from a supposed exteriority. The Other can be this self that changes, the internal being that is transformed, an anomaly that has occurred within a system, that proliferates and retroacts over the whole: the mutation, when a form of life, by the the accident of an unforeseen, traumatic or salvaging, stroke of chance, is reconfigured, changes course, opening the door to the most unexpected consequences, the most accomplished as well as the most terrifying.

If humanity should, as some Californian or Cantonese prophets claim, play a decisive role in the emergence, in the long run, of a new "life form", *produced* life, actual *intelligent* artificiality, may these first, timid attempts at hybridation serve as a testimony both of the becoming-inhuman of our time, and the furious joy of the first interspecial coitus!