Learning and Understanding in the Mirror of Mathematics, Chapters 1 and 2

Misha Gromov

August 24, 2018

... IN THE ADMISSION OF IGNORANCE...
THERE IS A HOPE FOR THE CONTINUOUS MOTION
OF HUMAN BEINGS IN SOME DIRECTION.
RICHARD FEYNMAN.

OUT OF CLUTTER FIND SIMPLICITY.
JOHN WHEELER ABOUT EINSTEIN (?)

SIMPLICITY IS THE ULTIMATE SOPHISTICATION.
LEONARDO DA VINCI

What is understanding?

If you have never pondered on this question you cannot imagine how far from the answer one is.

Why? – you may naively object –

I understand *this question since* I understand *English and I understand what my* understanding *of English is.*

I understand *what the following identity signifies –*

$$123\ 456\ 789\ 101\ 112\ 131\ 415 \times 514\ 131\ 211\ 101\ 987\ 654\ 321$$
$$=$$
$$514\ 131\ 211\ 101\ 987\ 654\ 321 \times 123\ 456\ 789\ 101\ 112\ 131\ 415 \text{ –}$$

– my understanding *of arithmetic makes me to* understand *why it is true.*

I have a sufficient understanding *of chess to see why blacks cannot win; I can demonstrate the fullness of* my understanding *by explaining this to you.*

Figure 1: Penrose's draw

I understand, *I learned this at school, that heavy objects fall with the same speed as lighter ones.*

No, not at all! Look, at how it really is.

NUMBER ONE: natural language understanding remains an unsolved *AI-hard* problem.

NUMBER TWO: no 21st century mathematician would claim he/she understands why *all* numbers obey by the rules of arithmetics or what this "*all*" means. No one has an inkling of what understanding of mathematics in the human mind is made from.

NUMBER THREE: neither the latest computer program, *AlphaZero*, which is stronger by a margin than any living chess master, nor its authors understand how it performs. And, at the same time, it will take longer for AlphaZero, if ever, than it was for you, to *understand*, Penrose's draw.

NUMBER FOUR: your "understanding" of laws of mechanics is, most likely, the result of the intellectual conditioning in the high school. As B.F. Skinner (probably never) said,

give me a child and I'll shape him into anything.[1]

Well, – you may reluctantly agree – I may not truly understand this highbrow staff but you cannot convince me that I don't understand simple things.

If a glass falls on a stone floor it brakes into pieces.

Understanding of this is stark clear in my mind.

No and again No!

Understanding our own understanding is an illusion
partly (but not fully) nourished by ignorance.[2]

Figure 2: illusion

[1]Skinner's philosophy of behaviourism may be 99% wrong, but his level of intellectual sophistication was above the banality of this passage – a rhythmic imitation of GIVE ME A FULCRUM, AND I SHALL MOVE THE WORLD by Archimedes.

[2]– Ignorance goes hand in hand with disability (reluctance?) to recognise this very ignorance. This is why a first year physics student is eager to enlighten you on what quantum mechanics is and thousands on Google pages are filled with explanations of what "understanding meaning" signifies.

Psychologists call this overestimation of one's own cognitive ability *Dunning-Kruger effect*: ... if you're incompetent, you cannot know you're incompetent ... The skills you need to produce a right answer are exactly the skills you need to recognise what a right answer is. https://en.wikipedia.org/wiki/Dunning%E2%80%93Kruger_effect.

Understanding how we understand "simple things" is the hardest. "Abstract and sophisticated" is easier in this respect, chess, as a model case, must be the simplest one.

Understanding is damn hard to understand and to explain – try you ideas on a three year old who counters everything you say with "whys", or imagine designing a computer program which could encode your understanding of something.

Forget it! This is hopeless. There is nothing in the world that can guide us in doing this.

> EVERY NEW BODY OF DISCOVERY IS MATHEMATICAL IN FORM, BECAUSE THERE IS NO OTHER GUIDANCE WE CAN HAVE.
>
> CHARLES DARWIN

Yes! Every *humanly conceivable process* in this universe can be expressed in mathematical terms and operationally encoded by numbers. Simple and unbeatable–

humanly conceivable is synonymous to *arithmetically describable* –

a truism since Hilbert, Gödel[3] and Turing have mathematically demonstrated omnipotence of formal thought. [4]

This beautifully applies to evolutionary biology: 99% of the overall structure of it is mathematics, pure and simple:

cut-off of exponential growth – not some mysterious "life force" – this is what channels the time flow of Life on Earth to the gullies carved by the natural selection.

But does it apply to understanding? Is the nature of understanding is humanly conceivable?

Some philosophers – they say "formalisable" instead of "conceivable" – deny this. For instance, Roger Penrose, thinks that, on the bottom, human understanding of mathematics, even that of chess, is non-formalisable.

And if you suppose that human cognition depends on quantum and that a complete mathematically non-contradictory interpretation of quantum in the classical world we happen to live in will never be achieved, then you resign yourself to believe that the same fate awaits human comprehension of human understanding.

We are more optimistic even though the reason for this optimism may not strike you as something cheerful :

the poverty of hominid genome evolution.[5]

[3] Gödel's *incompleteness theorem* is often interpreted as *impossibility of formalisation* of arithmetic, but *the proof* of this theorem says otherwise.

[4] Charles Babbage, who, around 1837 designed what is now called a *universal Turing machine*, might have understood this.

[5] The number of hominids who lived in the interval 5 000 000 - 100 000 years ago is estimated between five and twenty billion. Probably, only a fraction (1%? 0.01%?) of these carried relevant mutations, where only a fraction of this fraction has been fixed in the population.

Also the human-chimpanzee genome comparison suggests that it would take only a few pages to write down everything what makes our genome distinctively human.

But even granted all that, you may be baffled by a lightening speed of the animal brain → human brain transition: a back-of-the-envelope calculation indicates that the probability of this ever happening must be minuscular. What saves the day – we shall explain this later – is a particular organisation of eukaryotic genomes and the (conjectural) logic of animal

This suggests that

the core of the human mind is simple and its structure *is* comprehensible.

Let us search for this simplicity by trying to make sense of the six purple arrows un the following schematic picture.

$$
\begin{array}{ccc}
 & & \text{Mathematics} \\
 & \nearrow \swarrow & \\
\text{UNDERSTANDING} & \leftrightarrows & \text{Language} \\
 & \searrow \nwarrow & \\
 & & \text{Science}
\end{array}
$$

Contents

embryogenesis.

1 Where We Go, what Stands on Our Way, what are Our Resources.

...THE HARDEST VICTORY IS OVER SELF.
ARISTOTLE

ONE OF THE MOST MISLEADING REPRESENTATIONAL TECHNIQUES IN OUR LANGUAGE IS THE USE OF THE WORD "I".
LUDWIG WITTGENSTEIN

WHAT I CANNOT CREATE, I DO NOT UNDERSTAND.
RICHARD FEYNMAN.

If the core mental processes result from combinations of a dozen or two *elementary* operations regulated by a few rules which dictate choices – *random* is allowed – of such combinations, why cannot we still figure out how the mind works?

Here is a humiliating answer:

Our mental abilities are curbed by these very rules. We are not smart enough, not creative enough to understand our own *understanding*.

Our main limitation is a low power of imagination. The structure of *mental* may be simple but since it is unlike anything we encounter in the world around us, we are unable to create a mental picture of it.

What should we do to overcome this problem?

To develop insight in the nature of invisible rules of deep undercurrents of human thought you could start by freeing your conscious mind from the exaggerated image of human, including your own, intellectual greatness that blocks the vision of your mind's eye.

But this easier said than done. You cannot counter the influence of your ego on your thoughts by a simple effort of will.

Your ego – your idea of self, what your perceive as your personality, the running ways of your rational conscious thoughts– all that have been firmly installed in your mind by millions years of evolutionary selection working on your animal predecessors.

The shadow of the ego in our minds prevents us from understanding the nature of our thoughts as well as protects our self esteem from seeng what we are in the grand scheme of things.

Like it or or not, your conscious, mind is attuned to the needs of your, broadly understood, reproductive success, nothing in you amygdala – the emotion center in you brain gives a damn about what you understand of understanding – it is irrelevant for survival of your genes, which is all the amygdala is designed to care about.

It is only a brief stretch of time in your life – first years of infancy – when the solution of the survival problem rests in the hands of your parents and when your mind is not dominated by the urges of survival and reproduction, the time allowed by Nature for thinking, playing, learning.

Yet, the matrix where a child fits in what he/she learns is centred around me and mine. This is, so to speak, a polar coordinate system – everything rotates around shining myself in the center of the world.

Don't even attempt to overcome the gravitational pull of ego and to directly penetrate below the surface of your own mind or of the minds of other people.

Think instead, how to bypass the ego staying on guard in the portal of your mind, and sneak in, so to speak, by the back door.

Think, what aspects of our knowledge about the human mind are minimally distorted by the influence of the ego.

Are there "*impersonal thinking processes*" which can be clearly seen and studied?

Yes, there are two such clear cut processes: human interactions with

Natural Languages and Mathematics,
where chess serves as a model example of the latter.

If we find a simple mathematical representation $\mathcal{MR}$ of the transformation

input → output

for linguistic and mathematical signals which enter the human mind, and if (miraculously) this $\mathcal{MR}$ will happen to fit other kinds of signals, we would happily embrace such an $\mathcal{MR}$ as a bona fide model of understanding of these signals.

However, no such $\mathcal{MR}$ can simultaneously accommodate, say, mastery of a language and proficiency in chess and none of the two can be simple either.

A mere ability to competently respond to questions in a natural language NL needs, at the very least a (broadly extended) dictionary + grammar of NL, which is indescribable in simple mathematical terms.

The simplicity and generality one looks for is not that of $\mathcal{MR}$ but of the mathematics of the rules which define the algorithmic learning processes $\mathcal{ALMR}$ which lead to the construction of $\mathcal{MR}$.

What can direct us in the quest for these algorithms?
What could be criteria for selecting them?
Where can we borrow relevant ideas?

In philosophical (or poetic) terms – the actual meaning of this will become clearer as we plod on –

the mathematical criteria which we adopt for accepting $\mathcal{ALMR}$ *are:*
simplicity, universality, naturality,

and

the biological origin of native algorithms *of this kind necessitates their*
evolutional accessibility.

A convenient mind-set where these principles apply is as follows.

Pretend you know nothing of human languages, human mathematics, human games.

You enter a vast library $\mathcal{L}$ – thousands of volumes on different topics, including those on mathematics, for instance.

Identify as many as you can structural patterns in what you find and compare this with what you, in fact, know about them. If the resulting novice's understanding of $\mathcal{L}$ turns out to be close to the expert's one precooked in your mind, trace back your steps which led you to the former and describe the tracks you followed in mathematical terms.

It may be unclear even where to start, but you get ideas by learning how scientists were deciphering "writings" by Nature by untangling the logical threads of structures present in the Nature's libraries e.g. of

live organisms, natural languages, mathematical theories.

Thus you realise that "impossible" becomes "inevitable" when you look beneath the surface of things.[6]

Encouraged by this we try to mathematically stucturalize "impossible" process(es) of learning to understand with a constant eye on ideas generated in biology, linguistic and mathematics itself.

EGO REVOLTS:

– *Why do you choose this arcane roundabout impractical way? The human mind has been analysed in depth by many generations of psychologists and philosophers. You only have to google "understanding", "learning", "meaning", "mind", "intentionality", "cognition" "philosophy of AI" etc. and you'll find thousand pages with all kind of answers to all your questions in all kinds of simple words.*

– But we, mathematicians, don't understand these words.

– *True, the language may be vague, metaphors may be deceptive, but ambiguity of language is what makes it so versatile and expressively powerful. One does not have to formally understand the meaning of words to competently use them. Isn't it wonderful?*

– Hm...may be it is, but mathematicians don't buy it:

if we understand an idea then it is only by mathematically recreating it.

1.1 Thoughts, Structures, Spaces, Numbers

...IN THE MENTAL WORLD ...,
THERE ARE VAST COUNTRIES
STILL VERY IMPERFECTLY EXPLORED.
BERTRAND RUSSELL, THE ANALYSIS OF MIND.

...THOUGHT IS SIMPLY BEHAVIOUR -
VERBAL OR NONVERBAL, COVERT OR OVERT.
BURRHUS SKINNER, VERBAL BEHAVIOUR.

Searching for mathematical structures in the dark space behind $\mathcal{MENTAL}$ is hard, finding anything of value is next to impossible. Why to bother?

It is more productive to study the behaviour implied mental processes and try to learn as much as we can about $\mathcal{MENTAL}$ itself this way.

No! No way this way. This is like looking for the lost keys near a lamppost because there is more light there.

[6]It takes some amount of preliminary thinking to appreciate the extent of impossibility of life, human language, mathematics.

Behaviour itself in all the complexity of its controlling relations[7] doesn't faithfully reflect the complexity of *THOUGHT*. Observable behaviour has fundamentally different structure from that for *MENTAL*. It is impossible to decipher the logic of *THOUGHT* by staring at its image on the screen of *BEHAVIOUR*.

... Successive energies of the highest intellects, exerted through many ages – in the words of Michael Faraday – which led to main scientific discoveries, were spent not only on designing and performing experiments but also on imagining invisible structures without which observable outcomes of these experiments would have no meaning.[8]

Kepler's laws could have not been discovered by a careful observer laboriously tracking the traces of planetary epicycles on the celestial sphere.

Biochemistry and combinatorics of metabolic pathways in cells would have not been reconstructed, not even conjectured, by a meticulous analysis of the final products of breathing + digestion.[9]

To superficially appreciate the complexity of *UNDERSTANDING* look at the numbers involved in learning native languages by children.

By the age 6-10, a child acquires a fair understanding of native language upon attending to ONLY 10^6-10^8 phrases uttered by people around him/her.

In fact, 10 years contain approximately $3 \cdot 10^8$ (300 million) seconds. This allows learning time of 10^8 seconds, whereas the average rate for English speakers is 2-3 words per second.

Accordingly, an input of texts with 10^7-10^9 words must be sufficient for a learning algorithm. More generously, one might accept learning programs trained on collections of texts with 10^{10} – 10^{11} words in them.

To get an idea – a book page may contain 100-1000 words, books come in 50 -1 000 pages and the number of different English books in the world is of order 10^8. Probably, all what comes from the world-web is realistically limited by 10^{12}-10^{13} English words.

The human brain, which play with whichever comes into it, may be producing 10 -1000 new sentences per second by rearranging and combining what comes into it and keeping the resulting 10^9-10^{11} in its long term memory.

But the resources of the brain are far from infinite.It contains ~ 10^{11} neurones – about $2 \cdot 10^{10}$ of them in the cortex (most of neurones are in the *cerebellum*) where only a minority is dedicated to language: cortex is predominantly occupied with processing sensory information – visual, auditory, somatosensory.

With ~ 10 000 synapses per neurone and, probably, a few hundred synapses needed to record a sentence, the brain can hardly keep 10^{11} ready-made sentences (or any kind of "ideas") in its memory – even 10^9 is barely realistic.

This amounts to a fraction of all possibilities . For instance, the number of grammatical English sentences shaped as Colourless green ideas [which] sleep furiously is above $10^{20} = (10^4)^5$.

[7] B. F. Skinner, Verbal behaviour.

[8] We shall clarify later on the meaning of this meaning.

[9] The main contribution to the exhaled air is %4 carbon dioxide, not much different from what cars "exhale". But the main product of animal digestion is more interesting: the feces of a (healthy) human contains trillions of bacteria (comparable to how many cells there are in your body) of several hundred species. The life of these, still poorly understood, bacterial communities in human intestine may be richer in structure than anything we know of animal (including human) behaviour.

There are significantly less "meaningful" sentences among them – the number of predominantly probable ones may be below ~ $10^{12} = 10^4 100^4$. [10]

The brain is excruciatingly slow. Neurones need to rest 1-10 milliseconds between firings and the *average* number of spikes of a neurone in the human cortex is below one per second. This makes ONLY 10^{11} spikes in the entire brain per second – ten times slower than the Intel's $2 000 teraflop chip.

The time currency of the mind is about 100 milliseconds, human reaction time to auditory and visual stimuli is the 100-200 milliseconds range.

Even the *star-nosed mole* the fastest-eating mammal, takes 100-200 ms to identify and consume a food items with 10 ms needed to decide if it is edible or not. This is the limit speed of neurones.

Let them be small and slow, the brains seem staggeringly intricate.

It remains a mystery, for instance how the brain of the humble worm C-elegans (fig. 4 in section ???, 302 neurones, ≈ 7,500 synapses) functions, although all connections in the worm's nervous system are anatomically charted.

Numbers which stay on guard on the boarders of $\mathcal{UNDERSTANDING}$. may serve only a "negative" purpose by preventing us from developing plain stupid ideas.

What can positively direct us on the road to $\mathcal{UNDERSTANDING}$.(besides what comes from experimental psychology and neurophysiology) is a picture of "mathematical spaces" behind these numbers where $\mathcal{MENTAL}$, be it natural or artificial, resides.

Roughly, such a space has "dimension" $N = N_{bit}$ equal the number of bits in the memory of a mental system, say from $N = 10^3$ up to 10^{15}. This is large but not horrendously so.

But the number of conceivable "states" of the corresponding "mental system" – most of which are never realised – the monstrous $\mathcal{M} = 2^N$ is not something the "human intuition" is ready to deal with.

"*Just finite*" makes no sense when it comes to 2^N for such N. Your "finite" intuition may serve you for playing with automata which have 5 - 10, may be up to 100 states, but our 2^N, is closer to infinite than to anything finite in the human experience.

O.K., if you are a mathematician you have no problem with infinity, the idea is well familiar to you. But if (consciously or unconsciously) you follow the "infinity road" of thinking and forget that real "mental systems" are very much finite, you end up infinitely far from the true $\mathcal{MENTAL}$.

And it is not so much the greatness of the numbers N_{bit}, even in the range of 10^{15}, what makes $\mathcal{UNDERSTANDING}$ enigmatically intricate, but rather the modest size of such an N compared to the number of signals amenable to an $\mathcal{UNDERSTANDING}_N$.

For instance, you can understand strings in several dozen words, where the number of *possible* strings, by far exceeds the humble $N = 10^{15}$.

Indeed, there are more than 10^{500} string, say in 100 words, where at least 10^{100} out of them are grammatical, (more or less) meaningful and understandable.[11]

[10] The first word, such as Colourless is taken from a 10 000 word pool and the probable number of words "meaningfully" following a given one is estimated as 100.

[11] One cannot test 10^{100} strings for being understood, but one can test (quasi)random

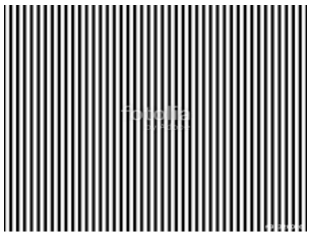
It is even more amazing that the "core" or the "seed" of a mature $\mathcal{UNDERSTANDING}_N$ must be thousands times smaller than N, probably, only a few tens of thousand bits – the rest is extracted from raw – poorly structured and noisy – flows of signals in the course of $\mathcal{LEARNING}$, similarly how an organism grows from a single fertilised cell in a flow of nutrients.

What makes the brain/mind tick, what allows it first to LEARN and then to UNDERSTAND with a seemingly improbable efficiency, is achieved by using some unknown class of structures buried in these exponential monster spaces.

Amusingly, the main brain limitation makes it adapted to production of such structures:

Parallel modus operandi by the brain, necessitated by its low operational speed, is better adapted to structuralization of signals, than *consecutive* information processing –

you cannot make a decent painting on a one dimensional canvas but families of parallel lines, which are kind of two-dimensional, serve you well.



Our dream is to bring hidden brain's mathematics into the open.

TO BE, OR NOT TO BE, THAT IS THE QUESTION.
SHAKESPEARE, HAMLET (≈ 1600).

Three and a half centuries later, this was put another way:

CAN A MACHINE THINK? YES or NO?
TURING (≈ 1950).

Look, the number of (conceivable and inconceivable) systems inside the N-dimensional dyadic space $\{\cdot\cdot\}^N$ is a super monster, $2^{\mathcal{M}}$, where M is our "monster" 2^N, and the number of "opinion" of which of them represent "true understanding" is super-duper monster $2^{2^{\mathcal{M}}}$.

Unless these numbers are faced squarely and promising patterns in such spaces as $\{\cdot\cdot\}^{1000000}$ responsible, let it be only conjecturally, for $\mathcal{MENTAL}$ are vaguely outlined, anything said about "$\mathcal{INTELLIGENCE}$", $\mathcal{UNDERSTANDING}$, etc., will remain pure, put it politely, *poetry* with probability $1-2^{-2^{\mathcal{M}}}$.[12]

Amusingly, and a bit more seriously, both common YES and NO sprung from the same egotistical fallacy.

YES, *in ten years, we'll beat the brain in the intelligence game, we are* sufficiently smart *for this.*

samples of such strings which can be generated by a simple algorithm.

[12]This is itself a poetic exaggeration: the number of opinions expressible, say on 10 pages is bounded by something like $10\,000^{10\,000}$ which makes the probability of a lucky guess safely bounded from below by $10^{-40\,000}$.

Well..., one might say, one of the two YES and NO must have probability ≥ 1/2. We disagree: assigning probabilities to these is as hard to justify as it is for TO BE AND NOT TO BE.

NO, *computers will never reach the human level of intelligence – we are* too smart *for this.*

It is hard to say who is right, but one thing you learn from doing math is that we are stupid rather than smart, blind to the obvious right under our noses.

(A stunning example is the $P \neq NP$-question, the simplest imaginable mathematical problem, that remains unsolved for almost 50 years.[13])

It is safer not to take stake in this YES/NO debate and spend energy on trying to guess what could be *simple* mathematical structures within $\{\cdot\cdot\}^N$-spaces which resemble whatever little we understand about human LEARNING and UNDERSTANDING.

1.2 Deciphering Strings and Building Machines: DNA, Mathematics, Natural Languages,

The only data available to the brain come from flows of neuronal signals and the meaningful picture of the external world in the human mind is constructed (reconstructed?) by manipulations with these signals.

How would you go about achieving such a feat?

In a novel *A for Andromeda* by Fred Hoyle and John Elliot, one detects a radio signal from somewhere in space that contains instructions for the design of a huge computer. The instructions are understood and the computer is build.

But is this possible? Can one represent a complicated structure by a string like this

... ●●○●●○○●○○●●●●○○○○●●●○●○●●○●●●○○●●●○○○○○●●●●●○○○●●●●○○○●●○○○○○
○●●●○○○●●●●○○●●○○○○●●●○○○○○○○●●●●●●●○●●●○○●●○●●●○○○○○●○●○●●●●●●●●...

Yes it is. Everything *conceivable*[14] can be described in a human language, even a sublanguage, e.g. a fragment of mathematics, will suffice.

But can one decipher and understand such a signal? Mind you:

there is no idea who the authors of the signal are, no hint at their background;
the language of the message is an unknown to you;
you have no a priori knowledge of what kind of contents is encoded by the signal.

For some of us this is easy. Bacteria, our cousins, have been solving such problems (for their own destruction) for 3 billion of years by following the instructions encoded in DNA of viruses entering their cells – viruses make sure their messages are "understood" by bacteria.[15]

But if you have no bacteria this message is addressed to in your lab, you will have hard time trying to understand this message, if at all. And your chances are nil, if a string S_0 in you hand encodes an organism from another world the biology of which is unfamiliar to you.

However, if you have sufficiently large collection of such string, say 10^{15} strings S of length $10^4 - 10^9$, and enough computational power (broadly understood) at your disposal, say of $2^{10^8} \times 10^{25}$ computer hours, you do stand a

[13] Actually, more than 50: several mathematicians were unsuccessfully trying to find lower computational complexity bounds (essentially) equivalent to $P \neq NP$ prior to the current formulation of the problem suggested by Cook and by Levin in early 1970s.

[14] Probably, the fundamental physics of the Universe is not in this category:
THE UNIVERSE IS NOT ONLY QUEERER THAN WE SUPPOSE, BUT QUEERER THAN WE CAN SUPPOSE – in the words of John Haldane.

[15] The Andromeda message in the Hoyle-Elliot story looks very much virus-like.

chance of deciphering the string language and thus, understanding the message carried by S_0.

To do this, assume that there is an algorithm A_* representable by a binary 10^8-string which makes the strings S, understandable by translating them to your own language. Then try all these algorithms A one by one, with 10^{25} verifications steps dedicated for checking soundness of each A.

Thus, you may claim that *in principle*, you *can* arrive at understanding S_0. This could be accepted if the words *"in principle"*, and *"can"* are taken in a philosophical or poetical sense, or if 2^{10^8} can be "physically" implemented in the (multi)multiverse you live in, but we, Earth-bound mathematicians, reject such solutions of this "understanding S_0" problem.[16]

Do we fare better in mathematics than in biology? Could, for instance, Archimedes understand a mathematical text of 21st century?

A priori, with no constrains on such a text, the probable answer is "NO" – according to (a version of) the $NP \neq P$ *conjecture* an exponential factor like the above 2^{10^8} is unavoidable.

In reality, no matter how hard some mathematicians try to achieve the contrary, subexponential time suffices for deciphering their papers, similarly to how children figure out what the words and sentences they hear around them mean.

Well..., after all, all *human* minds have much in common, but can you make yourself understood by an inhabitant of *another* universe?

Look at the string

☞☛☞☝☞☞✍☞☞☛☞☝☞☞☞✍☞☞☛☞☞☝☞☞☞☞✍☞☞☞☛☞☞☝☞☞☞☞☞✍☞☞☞☞☛☞☞☞☝☞☞☞☞☞☞☞✍

Can you figure out what's wrong with the following?

☞☞☞☞☞☛☞☞☞☞☝☞☞☞☞☞☞☞☞☞☞✍

Apparently, everybody in this Universe will understand by now that ☞☞☞☞☞☛☞☞☞☞☝☞☞☞☞☞☞☞☞☞✍.

But what if there is no numerals in your language if you have no idea of counting as it is with the Pirahã people?

Conceivably, no adult Pirahã man or woman will ever guess what the symbol ☛means but it won't take long for a Pirahã child to figure this out.

One can only wonder what similar limitations for representatives of other cultures are and what kind of symbolic messages can be eventually understood by every sentient being.

1.3 Behavioural Algorithms and their Evolution.

Our ultimate goal is to figure out what are the *innate algorithmic* rules that define human learning, but since learning processes are not open to direct observation we look , briefly, at better studied instances of innate sequential algorithmic behaviour in animals.

[16]Unlike the nonsensical 2^{10^8} the number 10^{25} is, roughly, in the computational range of biological (biophysical) systems. For example, the first *non-trivial* (from mathematical point view) step of "translation" DNA $\rightsquigarrow$ ORGANISM is achieved not by ribosomes but by protein folding process, where in the course of folding a polypeptide chain may undergo 10^{12}-10^{20} elementary (hinge) motions.

The most studied instance of this is the self-grooming of rats – a frequent activity of rodents with many evolutionary conserved sequencing patterns. Grooming in rats, build of about 15 chains grooming 75 sec. each, occupies 30% of rats active life.

A typical self-grooming syntactic chain in rodents, which is often embedded in other forms of grooming behaviours, serially links 20 or more grooming movements into four distinct, predictable phases that follow the same cephalocaudal (head-to-body) rule. The serial structure of such chains is repetitive and consistent in terms of order and time, so that once the first phase begins, the entire remaining sequential pattern reliably continues through all four phases. This syntactic chain pattern accounts for approximately 10-15% of all observed self-grooming behaviours in rodents, the remainder of which follow less predictable sequential patterning rules...

The syntactic chains are usually interspersed with more flexible "non-chain" grooming (that is, flexibly ordered mixtures of strokes, licks or scratches that are not components of syntactic chains), which accounts for approximately 85?90% of all grooming behaviours. [Neurobiology of rodent self-grooming...]

Sequential behaviour of insects may be also amazingly complicated, such, for instance as nesting of predatory wasps, e.g. *ammophila campestris*.

The [nest-building] *sequence begins with the construction of a vertical burrow in the ground that, upon completion, is sealed with several appropriately sized pebbles. The female wasp then leaves the nest site in search of a caterpillar to serve as food for her larva. When she finds a caterpillar, the wasp stings it in successive segments to paralyse it and carries the immobilised prey back to the nest. The small stones blocking the entrance are removed so that the caterpillar can be dragged into the burrow, after which a single egg is laid upon it. ... Only after the burrow is fully provisioned does the female engage in the final nest closing, which includes concealing the entrance by smoothing the dirt over it.* [M. Prakash], *Insect Behaviour*.

There is a *twig girdler beetle* (in the genus of *oncideres*)where the female does the following. [17]

- *finds and climbs a mimosa tree,*
- *lays eggs, by crawling out on a limb, cutting a longitudinal slit with her mandible and depositing her eggs beneath the slit,*
- *backs up a foot or so and cuts a neat circular girdle all around the limb, through the bark and down into the cambium. It takes her eight hours to finish this cabinetwork.*

The limb dies from the girdling, soon falls to the ground, the larvae feed and grow into the next generation. (Left to themselves, unpruned, mimosa trees have a life expectancy of twenty five to thirty years. Pruned each year, which is what the beetle's girdling labor accomplishes, the tree can flourish for a century.)

It is most puzzling how such a behaviour could have evolved by a sequence of unorchestrated[18] mutations, where all an individual mutation does is a mod-

[17] See an essay by Lewis Thomas on http://hermiene.net/essays-trans/seven_wonders.html

[18] We avoid saying "random" as it can give a wrong impression that we understand what this means.

ification (sometimes suppression or creation) of a protein.[19]

Most likely, the specifically human intelligence, e.g. speech generation, especially, when it comes to *linguistic recursion* (the dog that chased the cat that killed the mouse that ate the cheese...), and long-chain mathematical reasoning evolved, along with general algorithms of learning, by adaptation of ancient sequential behaviour mechanisms.

> ... words of the language, as they are written or spoken, do not seem to play any role in my mechanism of thought. The psychical entities which seem to serve as elements in thought are certain signs and more or less clear images which can be 'voluntarily' reproduced and combined.... The above mentioned elements are, in my case of visual and some of a muscular type.... Conventional words or other signs [presumably mathematical ones] have to be sought for laboriously only in a secondary stage, when the associative play already referred to is sufficiently established and can be reproduced at will.
>
> ALBERT EINSTEIN IN A LETTER TO JACQUES HADAMARD.

One hardly can understand the origin of human cognition unless one understand mathematics/logic of the chain

mutagenesis ⇒ embryogenesis ⇒ neurophysiology ⇒ behaviour

in animal evolution.

And identifying individual steps in this chain may suggest the key formal operations, simple combinations of which would progressively navigate us in the "mathematical space of algorithms" and direct us us toward the portal of the magic tower where LEARNING TO UNDERSTAND resides.

1.4 Lessons from Genome Evolution.

Contrary to what the classical Darwinists thought (still think?), it is highly unlikely that the evolution proceeds by gradual accumulation of "very small random variations" of organisms, where the Nature selects the variants with the highest reproductive rate.

Why not? Small variation are seen everywhere in the living world and evolution had tens and hundreds million years at its disposal to select beneficial ones.

A convincingly probable minuscule variation of the length of the tails of mice by one micron per generation, which would remain undetectable for centuries, could result in mice with impressive ten meter tails in mere ten million years.

[19]Probably, the essential modifications are those of regulatory regions on DNA.

Very convincing... except this "convincing", if you look at it with eyes open, is purely emotional, it has no logical or observational basis behind it.

(The easy acceptance of these "very very small" and "very very long" is, psychologically, is in par with the appeal of the idea of a harmonious universe conveniently designed for human consumption.)

The logical issue here is that the concepts of "small" and "random" make no sense unless there is a description of the pool, let it be conjectural, of possibilities to which these "variations" apply.

Prior to identification of the hereditary role of DNA in the middle of the last century, nobody had any idea of what this "pool" could be and, until recently, biologists were not in a position to fathom the complexity of the molecular inheritance and "variations" mechanisms and even to appreciate the extent of their non-understanding of how these mechanisms make evolution by selection work.

At one point it was realised with a great degree of certainty that the observable "variations" are *functions* F of modifications of DNA, where the most common such modifications, are the so called *point mutations*[20] which can be regarded as random and mutually independent.[21]

In some cases these result in small random modifications of organisms but in general, the function F are by no means continuous and the effects of point mutations may be quite strong and for the most part deleterious or even even lethal in many cases.

What is more interesting is that the architectures and the dynamics of genomes result in a significant amount of large, some of them beneficial, modifications.

For instance, division of genes (of multicellular organisms) into *exons* separated by long stretches of non-coding "junk" DNA allows reshuffling of DNA that leads to functional proteins. [22]

Conceivably, such large modifications, rather than point mutations, constitute the true source of evolution.

The above was not intended even to touch the issue of the genome evolution. The point we want to make is that, by drawing parallels, essential steps in the process of learning, like those in evolution, may be non-local and some of its mechanisms might have come from unlikely sources.

1.5 Against Logic and Reason: Mathematical Principles of Understanding.

There is a world of difference between the self-centred ideas of

☛ consciousness, free will, qualia, intuition, sensibility...

and universal concepts

✎ similarity, information, complexity, probability... .

[20] If one think of DNA as a (long) sequence in four letters, a point mutation is a removal, insertion or a substation of one letter by another one.

[21] The concept of randomness is vacuous if one can't safely assume independence.

[22] The presence of this beneficial "junk" is due to the *lateral DNA transfer* , e.g. what is effectuated by viral infections or by *transposon – jumping genes*, (kind of internal viruses) which are, by no means beneficial to organisms.

The former are limited to metaphoric use in philosophical deliberations, while the latter are seeds of mathematical concepts which germinate out of them.

We want to transfer "learning" and "understanding" from – the domain of psychology and philosophy to – the sphere of mathematics; we ask:

Is there a balanced and consistently mathematically describable class *or* universe $\mathcal{UNINET}$ *of mathematical objects, we picture them as colored networks* $\mathcal{NET}$*, which would incorporate into their structures most (all?) of what we know of and what we expect from human* UNDERSTANDING*?*

When we say "mathematics" we do not mean a particular theorem, theory or a branch of mathematics, nothing of the kind called "tool" by mathematics users. There is no pre-cooked devices in mathematics for making models of UNDERSTANDING. But general mathematics ideas and principles which evolved over the last half century suggest guidelines for description and construction of "understanding networks".

I. Naturality. All concepts and notations must solely depend on the internal logic of (our ideas of) UNDERSTANDING and not be influenced by considerations of convenience and habituation.

For instance, the inputs received by such "networks" $\mathcal{NET}$, we call these inputs *signals* ♫$_{\mathcal{SIG}}$, are *not* naturally represented by "$(0,1)$-vectors"; hence, these are not permitted in mathematical description of interactions of $\mathcal{NET}$ with ♫$_{\mathcal{SIG}}$. [23]

Concurrencies and Similarities. At bottom, what the brain can perceive and distinguish in input signals ♫$_{\mathcal{SIG}}$ amounts to coccurrences and ssimilarities between such signals – the human brain has no material besides cs for building UNDERSTANDING and a human-mathematician is not allowed to use anything but cs as well.

II. Universality. The study of ♫$_{\mathcal{SIG}}$, which we pictorially call *flows of signals*, must be conducted in terms applicable to all (most?) such "flows"; however, specificity of some features is inevitable. For instance, there is a built in "concept" of *movement* in the visual systems of most (all?) animals but there is no such thing in the "linguistic perception" by humans.

III. Communality. No mathematical object X exists in isolation, it can be properly described and understood only in the language of the "community of its peers". The essential connectives between members of such a community, e.g. in the *universe of understandings networks* $\mathcal{UNINET}$, are *transformation arrows* $X \to Y$ between these objects, where the internal structure of an X is revealed by the combinatorics of the network of this arrows in the vicinity of X.

IV. Functoriality. Natural operations over mathematical objects must be applicable to all (most) members of their respective communities, where the idea of *naturality* is represented by *functoriality* that is a compatibility of these operations with the transformation arrows.

For instance, the *naturality* of the construction of UNDERSTANDING out of a FLOW-OF-SIGNALS can be seen as a kind of *functoriality* of the transformation

[23] $(0,1)$-description is legitimate for neural networks and it is also unavoidable in software implementations of mathematical models due to the hardware architecture of digital computers.

(which is effectuated by learning) from the "category" of the signal flows[24] $♫_{\mathcal{SIG}}$ to the "category" $\mathcal{UNINET}$ of understanding networks $\mathcal{NET}$

V Relativisability. Basic concept must be accompanied by their relative versions. For instance, an acceptable concept of learning should make sense for learning adapted to

(a) internal constrains or externally imposed conditions;
(b) particular abilities of a learner, e.g. previously acquired knowledge.

These III, IV and V are taken from the mathematical category theory, which offers several other useful for us ideas. However, category theoretic concepts, as they stand, do not apply the understanding problem. For instance, not all transformation arrows are composable, and when they are, the composition process may change "colors" of these arrows.

$$\begin{tikzcd}
& & A_7 \arrow[dd, bend left, "f_8"] \\
A_5 \arrow[urr, "f_6"] \arrow[d, "f_5"] & A_1 \arrow[l, "f_9"] \arrow[r, "f_2"] \arrow[d, "f_1"] & A_3 \arrow[u, "f_{11}"'] \arrow[d, "f_3"'] \\
A_6 \arrow[drr, "f_7"] & A_2 \arrow[l, "f_{10}"] \arrow[r, "f_3"] & A_4 \arrow[d, "f_{11}"'] \\
& & A_8
\end{tikzcd}$$

In fact, even the simplest mathematical concepts, such as natural numbers, should be used with discretion.

For instance, an "intelligent insider" $☺_{\mathcal{INT}}$ of a $\mathcal{NET}$
should be wary counting 1 for a "true number",
should treat 2 differently from 3 and 4
and
should put all larger numbers to the common basket with infinity.

But an "outside designer" of $\mathcal{NET}$s shouldn't be blind to the numbers, let them be vague, which characterise the size of signal flows and of corresponding understanding nets $\mathcal{NET}$, such as the numbers of "nodes" and "links" in them and also to the (much smaller and clearer) numbers of "colors".

Recall (see section 1.1) that the linguistic input reaching the mind of a child up to the age of 10 years[25] is safely limited by 10^7-10^8 primary linguistic units-words and common short phrases,[26] and, realistically speaking, models of language understanding should have 10^5-10^{11} nodes with 10^8-10^{14} links between them.[27] Even the full understanding network in the human brains which includes vision, is limited to $10^{11} - 10^{14}$ perceptual units in it.

But, remarkably – otherwise $\mathcal{UNDERSTANDING}$ would remain a philosophical metaphor – such a (tiny by mathematical standards) system is able to $\mathcal{UNDERSTAND}$ messages from a much larger potential pool of signals, say of size 10^{100}, as it was indicated in section 1.1.

[24]This "category" has not been described yet.
[25]There are about 31 million seconds in a year.
[26]This is about 310 million seconds
[27]10^{15} is the most generous estimate of the number of synapses in the brain.

More remarkably, the description of what we call *the universe of understandings* $\mathcal{UNINET}$ along with the rules of learning must be encoded by a few thousand (hundred?) "logical units" that is, roughly, the number of genes responsible for the (specifically human?) brain development and/or the conceivable number (of fixations) of beneficial mutations in these genes for the last couple million years. [28]

The numerical constrains on anticipated networks $\mathcal{NET}$ are hard to reconcile with the free spirit of mathematics: (almost) all our cherished concepts, constructions and theorems apply to an *infinite range* of possibilities. But the (stronger) bounds on the size of the description of $\mathcal{UNINET}$ and the dynamics of learning in it makes such a description amenable to a mathematical reasoning available to us. What we have to do is to

⋆ *define/describe some class (or classes) of signal flows and class (classes) $\mathcal{UNINET}$ of colored networks;*

⋆ *define/describe* LEARNING $\mathscr{L}^{\natural}$ *as a natural (approximately "bifunctorial") operation which applies to pairs ($\mathcal{NET}$, ♫$_{\mathcal{SIG}}$) and results in "more educated" networks $\mathcal{NET}^{\natural}$ in $\mathcal{UNINET}$.*[29]

⋆ Define UNDERSTANDING in this terms as a *quasistationary state* under the action of $\mathscr{L}^{\natural}$ in $\mathcal{UNINET}$ and design a computer program for its construction by successive application of $\mathscr{L}^{\natural}$ to a "baby brain network" $\mathcal{NET}_0 \in \mathcal{UNINET}$.

CONJECTURE. Most natural flows of signals ♫$_{\mathcal{SIG}}$ admit networks $\mathcal{NET}$ which UNDERSTAND them, where such a network,

$$\mathcal{NET} = \mathcal{NET}(\text{♫}_{\mathcal{SIG}}),$$

if suitably constrained by some condition(s), is *unique* up to *(quasi)equivalence*.[30]

ON REALITY, REASON AND LOGIC OF LEARNING. Instead of attempting a more precise formulation of general principles of learning and understanding we shall proceed with a step by step description of atomic rules for building *algorithmic learning systems* ☺$_{\text{ALGO}}$

These rules *must be*

simple, general and mathematically interpretable

and uninfluenced by

human reason, human logic, human idea of reality.

Our principal goal is to design such an ☺$_{\text{ALGO}}$, which,

without any built-in ideas of a language or of the external world, would be able to learn a given natural language ☺$_{\mathcal{LANG}}$, that is to build an understanding network $\mathcal{NET} = \mathcal{NET}$(☺$_{\mathcal{LANG}}$), provided this ☺$_{\text{ALGO}}$ is granted an access to a *representative collection* of texts in this language $\mathcal{LANG}$.

Hopeless – you object – *this cannot work*: language learning, especially within the window 10-30 months of age in the mind of a child, when the core of

[28]This $\mathcal{UNINET}$ is a variation of Chomsky's idea of *an internal computational system that yields a language of thought, a system that might be remarkably simple, conforming to what the evolutionary record suggests.*

[29]The "functorial" arrow $\mathscr{UN} \times \mathscr{S} \overset{\mathscr{L}}{\rightsquigarrow} \mathscr{UN}$, albeit superficially similar, is essentially different from that in the definition of *finite state automata.*

[30]This doesn't apply to the real world, since people are not "suitably constrained" – your may need a very generous QUASI to align certain understandings.

"mature understanding" takes shape, depends on and it is inseparable from

(i) inborn universal grammar in the human mind

(ii) non-linguistic – visual, tactile, somatosensory, olfactory – perceptions and actions,

(iii) human interactions,

(iv) emotions,

(iv) internal feeling of self.

Yes, unadulterated understanding $\mathcal{NET}_{pure}$ by ☺$_{\mathsf{ALGO}}$ will be different from $\mathcal{NET}_{human}$, but this does not mean it will be significantly, if at all, inferior to it.

Besides, if the learning rules are *sufficiently universal* to be applicable to "mixed flows", e.g. of co-occurring visual+linguistic signals, then the resulting understanding network will automatically include links between linguistic signals and the corresponding non-linguistic ones, (including human interactions) coming from the "real world".

Also such a program would allow a realistic model of *developing human ego* built into it, while acquiring proficiency in in all aspects of social culture(s), including logic and reason, will be the least of problems for our ☺$_{\mathsf{ALGO}}$.

But the internal logic of ☺$_{\mathsf{ALGO}}$, in order to have a slightest chance to succeed, must be far removed from what is called "human logic" and "human reason".

ALL MEN ARE MORTAL. SOCRATES WAS MORTAL.
THEREFORE, ALL MEN ARE SOCRATES.
WOODY ALLEN

Why should we dispense with the ideas of logic and reason accumulated in the human culture?

This is because the kind of logic we find in the "proof" that **Socrates is mortal**, is not the logic of human learning and understanding. The failure of the early Artificial Intelligence thinkers to axiomatise common sense, naïve physics, etc., is a witness to this.[31]

But a fair idea of *logic of learning* can be traced by observing how children and young animals come to understand the world around them.[32]

In a nutshell, what must go into design of automatic learning systems is

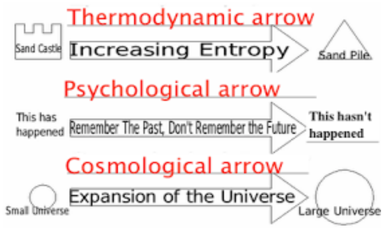
the logic of a five year old Cro-Magnon child set into
the conceptual frame of the 21st century mathematics.

1.6 Parallelism, Commutativity and Geometry of Time.

TIME IS A GAME PLAYED BEAUTIFULLY BY CHILDREN.
HERACLITUS

[31] See https://plato.stanford.edu/entries/logic-ai/ and https://www.doc.ic.ac.uk/~mpsha/naive_phys.pdf.

[32] See our *Memorandum Ergo*, where we explain what the (ergo)-logic of learning is and what of it can be learned from the behaviour of human children and animals.



Time is directly perceived by the brain the physiology of which is time dependent.[33] But the brain time, unlike physical (Newtonian) time, is divided into many (quasi)independent streams of flow corresponding to different parts of the brain when this function autonomously.

Also, division of flows of signals into distinct units allows a decomposition (of learning) of the (perceived) flow into several *disconnected* or only *weakly connected* subflows, or rather of *quotient flows* which correspond to these units. For instance, different words and short (sub)phrases in a sentence as well as distinct objects in the visual field and different classes of features of images, e.g. shapes and colors, can be analysed separately in parallel.

On a higher level, learning the grammar of a language can be divided into disjoint studies of individual rules.[34]

This is, of course, possible because grammars are representable by sets of (quasi)independent rules similarly to how the basis of the game of chess is describable by the rules of moves of individual pieces. And what you learn of a grammar or of the rules of chess *does not much depend on the order* you learn these rules.

If, thinking mathematically, you represent the *learning process* as a transformation $\mathscr{L}^{\natural}$ in the universe of all conceivable understandings, call this "universe of networks" $\mathcal{UNINET}$, then the division of this process into separate subprocesses will correspond to a *decomposition* of the transformation $\mathscr{L}^{\natural}$ into *commuting* or *almost commuting* transformations of $\mathcal{UNINET}$, where such a decomposition of $\mathscr{L}^{\natural}$, must be accompanied by some kind of decomposition of the "space" $\mathcal{UNINET}$.

Thus, the "true time" which parametrises the learning transformations is not the straight •→•→•→•→•→•→•→•→•→•→..., but something which is closer to a pair of commuting transformations:

•→•→•→•→•→•→•→•→•→•→...
↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
•→•→•→•→•→•→•→•→•→•→...
↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
•→•→•→•→•→•→•→•→•→•→...
↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
•→•→•→•→•→•→•→•→•→•→...
↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
•→•→•→•→•→•→•→•→•→•→...
↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
•→•→•→•→•→•→•→•→•→•→...
↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
•→•→•→•→•→•→•→•→•→•→...

[33] For instance, under normal conditions, the eye oscillate with about 100 cycles per second and with amplitude of several microns, which is comparable to the size of the cone receptors. But if the image on the retina is somehow stabilised for *several seconds* it fades and disappears.

[34] Holistic grammars are unlearnable, hence, non-existent in human languages. Besides, production of holistic grammars by the resources available to the guileless primate brain is also impossible.

↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
... ...

(The "real" brain time is structurally as well as dimensionally is more elaborate than that.) And neither the internal language of the subliminal mind, unlike linguistically orchestrated flows of conscious thoughts and of interhuman communications, is constrained by the one dimensional linearity. Because of this, a mathematical description of ordinary languages we look for must be general enough to be applicable to languages the geometry of which is not a priori known to us.[35]

1.7 What Worms Understand of Chess and Computers don't.

.... HE MAKES A FATAL ERROR:
HE BEGINS TO USE HIS OWN HEAD.
SIEGBERT TARRASCH[36]

Chess serves as a perfect model for the study of human understanding. How comes? – you ask.

Doesn't the game of chess belongs entirely with mathematics?
Do computers play chess better and understand it deeper than people do?
Is understanding chess representative of human cognition anyway?

Our answer is inconclusive: Yes and No. Three "Yes" and three "No".

The rules of chess stand out as logical axioms, they immerse chess as tiny fragment into the immense body of mathematics... but mathematics, as we know it, can hardly tell you anything about chess which is not already obvious to you.

The rules of the game, albeit simple, are too specific to be studied on the basis of general mathematical principles, at last of the principles we know today, while exponential combinatorial explosion allowed by these rules – 10^{50} legal positions and, seemingly, by hopelessly more, *Shannon number* N_{Sh}, of possible games – makes the direct search for winning strategies computationally unfeasible.

However, N_{Sh} is not as hopeless as it looks: many pairs of sequences of moves result in the same position. Put it another way, given two positions p_1 and p_2, what is relevant is whether p_2 can be reached from p_1 by a sequence of k moves, while the number of such sequences is of secondary concern.[37]

This suggests *retrograde analysis* of end games with n pieces on the board for small n.

Start with the set FCM of all final checkmate positions and, going backward step by step inductively in i, construct the sets CM_{-i} of positions from which you can arrive to $FCM = CM_0$ in (at most) i-moves. To accomplish this, you would need to look at less than

$$i \cdot 64 \cdot 63 \cdot ... \cdot (64 - n) < i \cdot (64 - n/2)^n$$

[35]When it comes to programming, parallel processes need to be represented by sequential algorithms. But mathematically, in the word of Hermann Weyl, this is *an act of violence*: the natural structure is destroyed and an artificial order is enforced.

[36]Tarrasch, a medical doctor by profession, arguably was the best chess player in the world in the early 1890s.

[37]We shall discuss the relevance of this number k in chapter ???

possibilities.

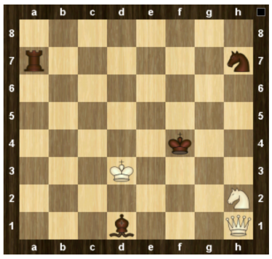


Figure 3: Magnificent Seven

As a matter of example, think of all possible developments of this position with seven pieces on the board.

The number of possible futures here is about $60^7 < 3 \cdot 10^{12}$, while modest 10 moves by each player at every step would generate an outrageously large number of 500 (white) move games – more than 10^{1000}.

There is, of course, little special about this, nor will you be exceedingly surprised to learn that this a winning position for white.

What is amazing – this was demonstrated by a computer assisted retrograde analysis – is that

it takes 545 moves to mate the black king

(with the first move by black). This is the longest mate known today.[38]

And who knows, may be a similar analysis

– *supercomputers + pruning off hopelessly loosing and drawn positions* –

would eventually deliver a *winning strategy for white in less than 25 moves.*

Why not? – There is a score of quick mates in the history of chess, e.g. in Steinitz vs. von Bardeleben game (1899) resolved in 24 moves[39] and in the one concluded with the decisive 21st move by Fisher (1956).[40]

Yet, one can safely bet that it will not be proved in the 21st century that the solution to chess is draw –

proving non-existence *of anything, whatever it can be, cannot be accomplished by brute force.*

But this way or another, it will not help to resolve more intriguing questions.

> *What is in there, in the structure of chess, that makes it aesthetically attractive?*
>
> *Can computers achieve the human level of understanding chess*?

[38] All 7 piece endgames has been generated and resolved, which took 6 months on the Lomonosov supercomputer and 140 TB to storage.
See *Chess Algorithms Theory and Practice* by Rune Djurhuus Chess Grandmaster, https://www.uio.no/studier/emner/matnat/ifi/INF4130/h16/undervisningsmateriale/chess-algorithms-rune-djurhuus-2016.pdf,
https://en.wikipedia.org/wiki/Endgame_tablebase,
https://www.chess.com/forum/view/fun-with-chess/longest-mate-official—mate-in-545.
Notice however, that according to the current FIDE rules this would be declared draw after less than 300 moves.

[39] https://en.wikipedia.org/wiki/Curt_von_Bardeleben

[40] https://en.wikipedia.org/wiki/The_Game_of_the_Century_(chess)

An argument against the latter was articulated by Edgar Allan Poe in his *Maelzel's Chess Player* – an essay published in 1836.

> ... if these [chess playing] machines were ingenious, what shall we think of the calculating machine of Mr. Babbage? What shall we think of an engine of wood and metal which can not only compute astronomical and navigation tables to any given extent, but render the exactitude of its operations mathematically certain through its power of correcting its possible errors?
>
> ....
>
> But the case is widely different with the Chess-Player. With him there is no determinate progression. No one move in chess necessarily follows upon any one other. From no particular disposition of the men at one period of a game can we predicate their disposition at a different period.
>
> ....
>
> ... even granting (what should not be granted) that the movements of the Automaton Chess-Player were in themselves determinate, they would be necessarily interrupted and disarranged by the indeterminate will of his antagonist. There is then no analogy whatever between the operations of the Chess-Player, and those of the calculating machine of Mr. Babbage, and if we choose to call the former a pure machine we must be prepared to admit that it is, beyond all comparison, the most wonderful of the inventions of mankind.

Let us try to figure out what Poe had in mind.

It is unlikely he thought that chess is, in principle, non-formulizable but neither one can say that Poe was consciously aware that

$$\textit{humanly conceivable} = \textit{computationally realisable}$$

as we understand this today.[41]

What Poe might have understood – this is childishly simple – is how to "reduce" chess to algebra:[42] assign a numerical weight w to each position by adding suitable weights attached to all pieces, compute this weight $w(p)$ in the position p obtained by k your own moves x and k moves y by your opponent, say

$$p = p(x_1, y_1, ..., x_k, y_k)$$

and choose the sequence of your moves x_i which maximise the corresponding y_i-minimum of $w(p)$, that is $\{x_i\} = \{x_1, ..., x_i, ...x_k\}$ for which the following maxmin is achieved

$$\max_{\{x_i\}} \min_{\{y_i\}} w(p((x_i, y_i)),$$

that is the maximum of the *worst (for you) case* function

$$worst_k(x_i) = \min_{\{y_i\}} p(w(x_i, y_i)).$$

Do this at each step with a sufficiently large k and you will perform beautifully!

However, and this is, apparently, the idea Poe wanted to put across, is that

the logic of consecutive computations is poorly adapted to handling exponentially growing tree-like sets of sequences of chess moves,

[41] This idea, however obvious, has not crystallised until the turn of the 20th century.

[42] In 1941 Konrad Zuse wrote down a chess algorithm in a high level programming language of his own making which has remained unknown for about 30 years.

and, given the speed limitation of the 18th mechanical contraptions, Poe justly concludes that no automaton, such as the Babbage mechanical calculator for instance, would be able to play chess even at amateur level.[43]

Things change. Today, no human can even dream beating the best - even next to the best – computer program in chess but... no computer (program) can dream of understanding chess. Why?

Well,... computers of today don't dream, essentially for the same reason they are unable to understand anything, they are not programmed for understanding the way we humans are.

A child who observes adults playing chess, even before he/she grasps much of the rules of moves, understands something which computers do not.

The obvious formal distinction is that *UNDERSTANDING*, as we understand the word, is not stationary but dynamic, it's inseparable from *LEARNING*.

The child keeps learning what he/she sees, while computer's "understanding" of how it plays is akin to what an insect (or a human for this matter) "understands" in how it breathes and digests.

But what about *AlphaZero* – a new computer program which starts by playing chess with itself, improves its performance by executing a neural network training algorithm and eventually becomes able to beat the best human, and even non-human, chess masters. Learning is there but does computer understand what it does?

Something tells you that there isn't even even a ghost of *UNDERSTANDING* in there.

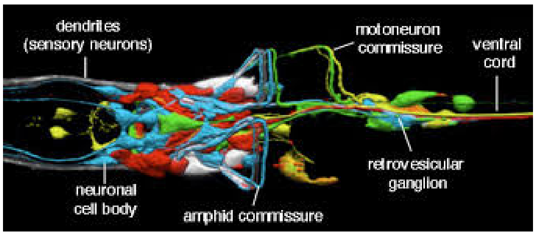


Figure 4: worm's brain.

A *Caenorhabditis elegans* worm, when it is trained to learn something, e.g. the smell of butanone associated with food (nobody yet tried to train it to play chess) feels closer to what we human call UNDERSTANDING than any computer program does.[44]

Why? What makes us think so?

[43]Similar arguments against feasibility of computer chess and of formalisation of human thinking process by overoptimistic AI-researchers were put forward in 1960s by Hubert Dreyfus, who, however, in 1967, lost to an early chess program.

https://ingram-braun.net/public/research/parlour-games/article/computer-chess-richard-greenblatt-match-mit-philosophy-artificial-intelligence-history/

[44]The nervous system of (hermaphroditic) Caenorhabditis elegans contains only 302 neurones, (male worms have additional 81 neurones in their tails) and approximately 7,500 synapses. http://www.sfu.ca/biology/faculty/hutter/hutterlab/research/Ce_nervous_system.html

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3197297/

http://learnmem.cshlp.org/content/17/4/191.full.html

https://www.sciencedirect.com/science/article/pii/S0960982204005482

If you have a tendency to self analyse, you may ascribe this to your gut feeling – the guts taking over the brain. Should you trust what the guts shouts in your ear?

Why not, practically minded people say. Gut feeling has served our human and animal forefathers for millions of years – we wouldn't be here if they disobeyed gut's moral dictum: ALL IS FAIR IN LOVE AND WAR. Follow this feeling and you will succeed. Will you?

Sure – a practical man confirms – you will. You will succeed in everything, in all which is important: in love, in social status , in your career, in winning over your opponents.

No! – a mathematically sensitive scientist and a scientifically sensitive mathematician cry out together – down with gut feeling! We reject gut's egotistic logic we do not accept its opportunistic lies, we seek the truth! Gut feeling – call it whatever you want – common sense, intuition, accumulated human wisdom, derails us from the road to understanding the mathematical nature of *UNDERSTANDING* or of true understanding of anything for this matter. We...

Calm down, don't be carried away by your own noble feelings. Feeling of guts may come from somewhere else.

The *emotional brain*, which passed millions of years ago from reptilians to humans, hasn't enough colors in its spectrum for fine labelling. Emotionally, feeling-wise, we tell "beautiful" from "edible" no better than lizards do,

This is why, as Tarrasch says,

chess, like love, like music, has the power to make men happy.

But it is not anything sensual, not at all related to sex,[45] but something PURELY FORMAL in the structure of chess, which ignites human emotional response.

Also there must be something, also 100% formal, in the human perception and understanding of this structure, more elaborate than just "feeling happy", which, in a simpler form, may be also present in the worm's brain but not in the "brains" of the computers of today.

What constitutes this something, what makes our (and chess playing worm's) understanding of chess different from that of a computer is not a superior depth but – this is what the gut feeling is trying to tell us – its

$\bullet_1$ *universality*

and

$\bullet_2$ how this particular "understanding" *is integrated* in a wide network of other "understandings".

For instance, when a child observes people play chess, his/her perception of the game is

$\star_1$ organised on the basis of general or *universal* principles not specifically designed for learning to play chess or anything even remotely similar to chess;

$\star_2$ learning chess proceeds by *associating what the child sees to many ideas* already present in the child's mind: moving objects, symmetric patterns, etc.[46]

[45]Freudists, of course, have different idea about it. https://en.chessbase.com/post/che-psychoanalysis-psychology-and-pseudoscience

[46]A smart child who red Wittgenstein may have some ideas not related to the logic of the game but concerning the material the pieces are made from, e.g. imagining them made of chocolate. This is beyond abilities of the today computers but close to what a worm could find aesthetically attractive in chess.

The latter may appear as an obstacle to formalisation of human learning and human understanding.

Universality saves the day: it points to a path going around this obstacle.

We shall explain this in the following sections; later on, we shall revisit chess and see how the "universal" way of thinking helps one to better understand chess itself.

1.8 Biographical Digression: Fred Hoyle and John Haldane

The names of Hoyle (1915 - 2001) and Haldane (1892 - 1964) must be on the top of the list of the most unorthodox creative minds of the 20th century – you cannot mention them just in passing.

In 1946, Hoyle showed that the cores of large (more than 20 solar mass) stars, when they collapse, reach temperatures of billions of degrees (the sun core is about $15 \cdot 10^6$ K) which result in the nuclear thermal equilibrium between the attractive strong nuclear force and Coulomb's repulsion. This, in the agreement with observations, makes iron-56, the nuclei of which occupy the energy bottom of this equilibrium, the most abundant among heavy, elements in the Universe.[47]

In 1954, Hoyle suggested that the elements between carbon and iron, which cannot be synthesised in the observed amount at the high temperatures ($\gtrapprox 3 \cdot 10^9$ K) need a a specific nuclear fusion reaction – *the triple-alpha process* – which generates carbon from helium at 10^8K in the degenerating cores of the red giants and which requires, the existence of a very specific resonance energy of the carbon nucleus in what is now called *Hoyle's state*.

The calculation of the energy of this state by Hoyle, experimentally confirmed in 1957, is one of the greatest pinnacles of scientific cognition – on par with prediction of the antimatter by Paul Dirac.[48][49] And some 25 years later Hoyle comments:

Some super-calculating intellect must have designed the properties of the carbon atom, otherwise the chance of my finding such an atom through the blind forces of nature would be utterly minuscule.

IF J.B.S. HALDANE HAD NOT EXISTED
IT WOULD HAVE BEEN NECESSARY TO INVENT HIM.
EDMOND MURPHY

[47]The abundance of elements in the Universe, decreases exponentially with the atomic number, except for a drop of three light elements: Lithium 6 (but not equally stable ^{7}Li), Beryllium and Boron (which are, up to large extent, destroyed rather than generated in the normal course of stellar nucleosynthesis) and a thermodynamically predicted pronounced peak in the vicinity of iron.

[48]Non-surprisingly, the 1983 Nobel committee rightly judged Hoyle being overqualified for receiving the Nobel Prize.

[49]In 1898, Arthur Schuster coined the term *antimatter* and discussed the possibility of matter and antimatter annihilating each other.

In his 1941 book *New Paths in Genetics* Haldane suggested how the selfreproduction of the gene could be demonstrated:

> How can one distinguish between model and copy? Perhaps you could use heavy nitrogen atoms in the food supplied to your cell, hoping that the "copy" genes would contain it while the models did not.

Fifteen years later, and four years after the structure of DNA was resolved by Crick and Watson, Matthew Meselson and Frank Stahl performed this experiment – justly regarded by many as *the Most Beautiful Experiment in Biology*, which confirmed the mechanism of DNA replication outlined by Crick and Watson:

the double stranded helix of DNA unwinds
and each strand makes its own copy.

Below are excerpts from the longer list of other *Haldane's firsts in science*.

• *demonstration of genetic linkage in mammals (1915);*

• *mathematisation of the evolution theory by fusing Mendelian genetics with natural selection and thus building the main body of population genetics (1922-1932, in parallel with Ronald Fisher and Sewall Wright);*

• *the prebiotic soup theory – a physical models for the chemical origin of life (1929, before Oparin's 1924 book was translated into English);*

• *link between genes and enzymes (1920-1930);*

• *the standard model of enzyme kinetics (1930);*

• *human gene maps for haemophilia and colour blindness (1935);*

• *methods to escape from submarines (1941);*

• *the key ideas on the host-parasite evolution, including the adaptive role of heterozygosity in the sickle cell anemia. (1949, Prophetically– this was confirmed decades later – Haldane writes*

> *The corpuscles of the anaemic heterozygotes are smaller than normal, and more resistant to hypotonic solutions. It is at least conceivable that they are also more resistant to attacks by the sporozoa which cause malaria.*)

2 Atoms and Molecules in the Logic of Perception.

We do not claim to have a full mathematical representation of LEARNING and UNDERSTANDING – only certain patterns of their mathematical structure are discernible and even these are hard to describe with no adequate mathematical language available. The general picture we sketch in the first sections is, by necessity, vague, incomplete and imprecise.

2.1 Flows of Signals, Networks of Understanding and Universality Principle.

...PICTURING THINGS AS ENTERING INTO THE STREAM OF TIME...
Bertrand Russell

$\natural_{\mathcal{SIG}}$ When we say "$\mathcal{UNDERSTANDING}$" we mean "understanding something". We call these somethings, as earlier, $\mathcal{FLOWS-OF-SIGNALS}$ and denote them $\natural_{\mathcal{SIG}}$.[50]

Natural languages represented by collections of written text or records of conversations are instances of such "flows". Also images of spatial objects — their shapes, positions, motions — perceived visually or via tactile perception, are such flows. A pair of synchronous flows, e.g. of a visual and a linguistic one, makes a flow in our sense.

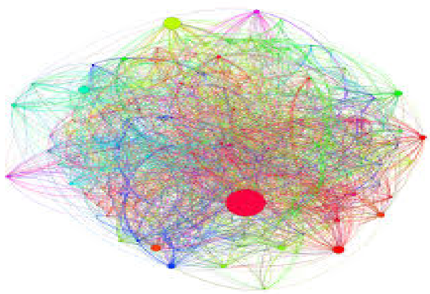
The main body of "$\mathcal{UNDERSTANDING}$" is symbolise, as in section 1.4, by a *multiscale* and *multilayer colored network* $\mathcal{NET}$ – network of "ideas" or "patterns" or whichever you call them, derived from $\natural_{\mathcal{SIG}}$ in the memory of an "*intelligent entity*" $☺_{\mathcal{INT}}$.

In general terms, $\mathcal{NET}$ is obtained from $\natural_{\mathcal{SIG}}$ by

suppressing redundancies and grouping together similar signals,

where the colors

trace mental mechanisms which implement the compression $\natural_{\mathcal{SIG}} \rightsquigarrow \mathcal{NET}$.[51]



$\mathcal{UNDERSTANDING}$ is dynamic, it is awash in the flows of external and internal (originated in the brain) signals. This is seen in how a network $\mathcal{NET}$ interacts with $\natural_{\mathcal{SIG}}$ when it is exposed to this flow.[52]

Such an interaction has several intertwined aspects to it.

1. Competent Perception: identification of structural patterns in a flow of signals $\natural_{\mathcal{SIG}}$ by matching them with correspondingly similar patterns in $\mathcal{NET}$.

This goes along with predictions of patterns which come next on the basis of their available fragments and with imposition of structures on $\natural_{\mathcal{SIG}}$.

2. Response Flow(s): a flow, or several flows, of signals which are generated in the subliminal mind of $☺_{\mathcal{INT}}$. These flows interact with $\mathcal{NET}$ and may (or may not) also influence the incoming flow $\natural_{\mathcal{SIG}}$.

[50] "Flow" sounds better than "pool"; besides a *pool* of signals, such as a library, feels a flow to those who swim through it.

[51] Nods and links in $\mathcal{NET}$ in the minds of human $☺_{\mathcal{INT}}$ of all ages are colored not only by reflections of elemental logical machines which build $\mathcal{NET}$ from flows $\natural_{\mathcal{SIG}}$, but also by ever changing flows of internal impulses and emotions.

Probably, these can represented by adjustable continuous parameters in $\mathcal{NET}$.

[52] The tendency of the human memory to continuous rearrangements will be described in terms of $\mathcal{NET}$ later on.

For instance, the visual system reacts to the flow of images it receives by sending neuronal signals toward the primary visual cortex. Also, the visual motor system send signals to the muscles which move the eyes. The latter influences the incoming visual flow.

3. LEARNING is a modification of the already present and formation of new patterns in $\mathcal{NET}$. Also the rules/algorithms of *perception* may be modified by learning.

(This is most significant at the early stages of exposure of an intelligent entity ☺$_{\mathcal{INT}}$ to a flow-of-signals ♫$_{\mathcal{SIG}}$ where the complexity of $\mathcal{NET}$ and issuing competence of ☺$_{\mathcal{INT}}$ are fast increasing.)

Realistic networks $\mathcal{NET}$ are beyond human grasp: they, probably, contain millions of nodes and billions of links between them. But we conjecture, that

the overall "space" of understanding networks $\mathcal{NET}$ – what we called $\mathcal{UNINET}$ in section 1.4 – admits a *simple* mathematical description and the dynamics of LEARNING – the time transformations of this "space" is governed by *few simple general* rules.

There is nothing paradoxical in this:

the set of all random $\{0,1\}$ sequences

is fully described (modulo standard terminological conventions) by seven words in the above line, but all descriptions of an individual such sequence must be, necessarily, as long as the sequence itself, that is eventually infinite.

More interestingly, a *simple* rule of motion on a *simple* space, kind of a vector field, may generate a trajectory with an *elaborate* limit set.

UNIVERSALITY

Central to our approach to LEARNING and UNDERSTANDING is the idea of *universality*:

the learning algorithms should not depend on the nature of a flow ♫$_{\mathcal{SIG}}$ which may be a priori known to us, e.g. being a particular language or a class of images, but rely entirely on formal properties of ♫$_{\mathcal{SIG}}$, such as mutual correlations between patterns in ♫$_{\mathcal{SIG}}$.

And, taking the lesson from the brain, we represent the building process of $\mathcal{NET}$ in the course of learning by algorithms which we select for the mathematical elegance of their structures and economy of performance, with little attention to the so called "real meaning" of signals.

2.2 Mental Units and Discretisation of Signals.

... THERE'S NO THREE AND A HALF WORD SENTENCE.

NOAM CHOMSKY[53]

"Flows" of natural languages are divided into discernible *units*: phonemes, words, phrases, sentences, utterings. Also human-human and human-object relationships, human artefacts as well as certain visual images, e.g. urban scener-

[53] In full, what Chomsky says is that the most elementary property of human language is that it consists of a discrete infinity of interpretable expressions – so there's five-word sentences, and six-word sentences, no five-and-a-half words sentence, so it goes on indefinitely like the integers. That's kind of unusual, there's nothing like that known in the biological world.

ies, trees, animals are, up to large extent, *perceptually discretizable* – one can tell stories about what one sees in there.

Figure 5: Man with a hammer

Figure 6: Urban scenery with a cow

Figure 7: Animals in a forest

Building as well as description of a network $\mathcal{NET}$ which represents $\mathcal{UNDERSTANDING}$ of such a discretizable flow ♫$_{\mathcal{SIG}}$ starts with *primary nodes* which are "copies" of directly perceptible units from ♫$_{\mathcal{SIG}}$, e.g. of *words* and *common short phrases* from a language.

More interestingly, there are *secondary* and *higher order*[54] nodes in $\mathcal{NET}$, which correspond to general concepts formed in the subliminal mind of ☺$_{\mathcal{INT}}$ at the various stages of learning, such as the *ideas* of

alive-moving (early stage), *animal* (later stage)[55]

and

sound-short-string (early stage), *word* (later stage).

Most of such concepts are inaccessible to the conscious mind of ☺$_{\mathcal{INT}}$, one does not know what they are, there are no words for them in human languages.[56] But interaction of $\mathcal{NET}$ with the flow ♫$_{\mathcal{SIG}}$ by the *rules of learning* results in "extraction" of these concepts from ♫$_{\mathcal{SIG}}$ and incorporation of them into the architecture of $\mathcal{NET}$.

In contrast to languages, auditory perception of music and muscle-based (to a lesser extent, visual) perception of motion are hard to discretise and verbalise.

Also many natural images, e.g. of textures, of crowns of trees, of wavy seas, of rocky terrains, etc. cannot be non-ambiguously divided into units and/or adequately described in words.

Probably(?), the visual system of the brain solves the problem similarly to how mathematicians do it by working out probability laws for random fields

[54]Some nodes may be assigned order zero, e.g. those representing edges and T-junctions in vision.

[55]*Dog* and *cat*, as far as the words go, predate *animal* in the linguistic mind of a child. But the *idea* of an animal – *furry-moving*, probably(?) precedes the ability of children to distinguish between ideas-images of dogs and cats.

[56]This depends on a language. There are hardly words for *copula* or *hyponymy* in Pirahã – the language of a small (≈400) hunter-gatherer tribe in Amazonia. But Pirahã people are as competent in the use of these concepts as linguistics professors. And academic definitions/explanations of the ideas behind this kind of words are invariably misleading since they convey the erroneous idea that one understands positions and functions of the corresponding concepts in $\mathcal{NET}$.

instances of which are implemented by these images.

2.3 Landscapes and Contexts.

> ...CONTEXT IS THE KEY
> FROM THAT COMES THE UNDERSTANDING OF EVERYTHING.

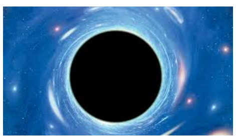
Figure 8: KENNETH NOLAND

Besides *localised units* such as words, sentences, parts of human faces and classes of these, which have (relatively) well defined boundaries, there are units representing classes of large and not fully specified chunks of flows ♫$_{\mathcal{SIG}}$, such as *city street* and *forest* in the above pictures

Obediently like the cow in the above fig.6 (???) the following sentences fit into their proper contexts.

Over a river there was a very narrow bridge. One day a goat was crossing this bridge. Just at the middle of the bridge he met another goat.

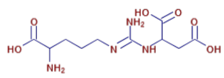
Go back
– said one goat to the other –
there is no room for both of us.



Why should I go back?
– said the other goat –
Why should not you go back?

Because
– said the first goat –
the quantum field degrees of freedom external to the horizon should be entangled with those inside it.

But
– said the second goat –
the measurement of argininosuccinic aciduria
must be performed by metabolite detection
cultured chorionic tissue.



Agreed,
– said the first goat –
Pledgor hereby grants and assigns to Patentee the exclusive, sole, permanent, world-wide transferable, sub-licensable and un-litigatable right for peremptory challenge activities.

"Our legal department wants us to download their new software. It translates gobbledygook to mumbo jumbo."

Moral: *Analysis of images and texts, e.g. evaluation of frequencies/plausibilities of particular patterns, e.g. words, is inseparable from identifying the contexts these patterns/words are taken from.*

Remember, context is everything:

Time flies like an arrow; fruit flies like a banana.

2.4 Concurrency, Textual Connections, Functional Links and Persistent Associations.

MAKING MENTAL CONNECTIONS
IS OUR MOST CRUCIAL LEARNING TOOL,
THE ESSENCE OF HUMAN INTELLIGENCE;
TO FORGE LINKS;
TO GO BEYOND THE GIVEN;
TO SEE PATTERNS, RELATIONSHIPS, CONTEXT.
MARILYN FERGUSON[57]

With no direct access to $\mathcal{UNDERSTANDING}$ in the mind of an intelligent entity ☺$_{\mathcal{INT}}$, e.g. of Marilyn Ferguson, we try to guess the structure of $\mathcal{NET}$ – the

[57]Marilyn Ferguson (1938-2008) was an American author, editor and public speaker.

imaginary network of patterns of understanding – by looking at how signals are organised in the "flows" ♫$_{\mathcal{SIG}}$.

The principal *connections* between signals-units from ♫$_{\mathcal{SIG}}$ are modulated by their *(percepted) concurrency*:

pairs of closely positioned objects in a visual field or two words (temporally) following one another in an uttering are likely to be *meaningfully connected* by serving a *common function*.

In the course of learning, a variety of different kinds of *connectives* is derived from such concurrency relations, where these connectives are incorporated as links between the corresponding nodes in $\mathcal{NET}$, which are colored according to their kinds.

Often, such connectives join small units into larger ones, where the most apparent are *primary connectives* such as

closely↭*positioned*, *is*↭*derived* and not quite concurrent *variety...* $\underset{is}{\leftrightsquigarrow}$ *...derived*.[58]

These connectives can be also regarded from an opposite angle by thinking of larger units as of bridges between their parts. Then connectives between A_1 and A_2 come as *compositions of arrow-relations*: $A_1 \hookrightarrow$makes-part-of$\hookrightarrow B$ and $B \hookleftarrow$incorporates$\hookleftarrow A_2$.

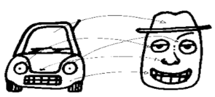
(Unlike what a classical mathematician would do, we exclude the case $A_1 = A_2$[59] and distinguish $A_2 = B$ by assigning a special color to it in the network $\mathcal{NET}$.)

There are also connectives between (percepted) units which do not make together larger units, such as *units*↶*ones* in "*...join small units into larger ones,...*". They are common in human (natural and unnatural) languages and, apparently, nowhere else. Detection of these connectives, e.g. correct identification of antecedent of pronouns on the basis of general principles, is non-trivial.[60]

Besides the first order connectives there are higher order ones. Some of them such as *noun*↭*verb* are recorded in grammar textbooks but many, probably most of them, remain hidden from our conscious minds.

Persistent concurrencies of units in ♫$_{\mathcal{SIG}}$ are recorded in $\mathcal{NET}$ where they are represented by links which are colored according to the kinds of functional relations between the corresponding units.

2.5 Similarities and Recognition of Similarities.



[58]This "is" is an instance of a concept which simultaneously serves as a node and as a *color* of certain links in the corresponding network $\mathcal{NET}$.

[59]Our logic is in agreement with that of young children and category theoretically trained mathematicians of the 21 century who are wary, for instance, of treating the equalities $2+3=5$ and $5=5$ on equal footing.

[60]Computational linguists still seem unable to fully resolve this problem, despite the knowledge that a given ♫$_{\mathcal{SIG}}$ is a language and an access to the formalised grammar of it. Even children (with no a priori knowledge of any kind) master the use of pronouns no earlier than between ages 2 and 3 years.

THE MOST INCOMPREHENSIBLE THING ABOUT THE WORLD IS THAT IT IS AT ALL COMPREHENSIBLE.

ALBERT EINSTEIN.

What makes comprehension of a flow of signals ♫$_{\mathcal{SIG}}$ possible, besides division of such a flow into weakly interacting simple parts – our units,[61] is

repetition of certain distinguished patterns in ♫$_{\mathcal{SIG}}$

and

recognition of these repetitions by the brain.

Such patterns, e.g. words in natural languages, are perceived as *signals* and positioned as *primary nodes* in the *network of understanding* $\mathcal{NET}$ in the subliminal mind of ☺$_{\mathcal{INT}}$.

Besides, most natural flows are highly *redundant*: meaningful[62] signals are separated by wide stretches of (stochastically quasi)*homogeneous information noise* with negligibly useful (for ☺$_{\mathcal{INT}}$) content in it.

For example, celestial bodies move in the vastness of (almost) empty space.[63] Less dramatically but still significantly, distributions of colors – green, brown, black, which occupies about 90% of pixels in our cartoon image (Figure 7) of animals in the forest, is information-wise (almost) irrelevant.

On the other hand, the principal artificial flows of signals – natural languages, especially in the written form, are rather condensed; yet, languages are sufficiently redundant to be deciphered and understood.

Sameness, Identity, Equality, Similarity, Resemblance. We say "*similar*" for all kind of resemblance and indicate instances of these by joining certain signal-units in the flow ♫$_{\mathcal{SIG}}$ and in the corresponding nodes in the network $\mathcal{NET}$ by *similarity links*. And, as it is done for the textual link, these are "colored" according to their kind and strength.

The apparent formal difference between the textual and the similarity links is that the former are, for the most part, *local*: textually linked units tend to be closely positioned in the flow ♫$_{\mathcal{SIG}}$. But similar signal-units, e.g. identical words in a text, especially if they are rare, may be quite far one from another.

And similarities between signals-units, e.g. copies of the same word may be represented in $\mathcal{NET}$ not by ordinary links, but rather by *recognition rules* of representatives of this word in ♫$_{\mathcal{SIG}}$.

In fact, a closer to the truth picture/structure of $\mathcal{NET}$ is that of a *network of rules, mechanisms and algorithms.*

[61] Possibly, some people are able to understand the world phenomena "holistically". But due to the reductionistic nature of language they cannot communicate this understanding to their fellow humans.

[62] The reader is not supposed to understand what the meaning of meaning is.

[63] One would hard time understanding the cosmic texture of gravitation if all stars from the Milky Way –about 300 billon of them, were packed in the sphere of radius of one light year or about ten trillion kilometres around the sun (≈70 000 × distance from Sun to Earth and 1/4 the distance to the nearest star), even though there still be a lot of empty space between the stars, more than billion kilometres between them on the average. But there would be nobody to "understand" this world – inhospitably hot and unstable.

Incomprehensible worlds are uninhabitable.

Similarities often modulate the second order textual links in $\mathcal{NET}$ by bringing in contact a priori unrelated ideas and images. Rhymes and metaphors do this in poetry:

> The desire of the moth for the star,
> Of the night for the morrow,
> The devotion to something afar
> From the sphere of our sorrow.
>
> Percy Bysshe Shelley.

And such *similarity shortcuts* in the texture of the *eventful time* happen to us every day in life as in this story.

Sir Dante – a stranger asked the poet – *What do you like best to eat?*

Dante answered: *Hard-boiled embryos of flightless birds.*[64] Ten years later the same man approached Dante again and asked: *With what?* Dante replied: *With salt.*

Sameness of the man in the mind of Dante connects eggs-with-salt.

Similarity in Structure and Similarity in Function. Similarity between visual images is mainly determined by what and how they composed of. On the other hand, what defines similarities between words in languages, especially in those called *analytic* by linguists, such as English, is not so much their morphological composition, but rather how they interact with other words.

The fundamental difference between the two is that the latter is non-local and depends on large pools of surrounding words.

For instance, no such "pool" is needed to detect something common in aggregation, and segregation but similarity between aggregation and gathering or *semi-similarity* link between segregation and divide is another matter.

The main principle of finding similarities between linguistic units – this will be formalised and extended to other situations – reads:

> If many cofunctional partners of two units are weakly similar, then these units themselves are strongly similar.

2.6 Forms, Transformations, Imitation, IQ test Diagrams and Combinatorial Logic of Languages.

Similarity links in $\mathcal{NET}$ between different forms of the same unit are "colored" according to algorithms effectuating transformations from one form to another.

The rules of (mental) transformations of visual images remain hidden from our mind's eye.

But the logic of such rules is transparent in *Diagrammatic IQ tests* and in languages where transformations decompose into elementary operations which makes them suitable for imitation and modification.

Test two is an example of what humans do when imitation of transformations fails. This works so well, because our minds are governed by what we call *ergo-logic*.

(Ergo-logic also suggests "**1**" as an alternative to the intended "**6**" in *test three*, since

[64] Apparently, Dante was quoting Terry Prachett.

Figure 9: forms one.

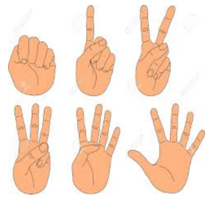


Figure 10: forms two.

Figure 11: forms three.

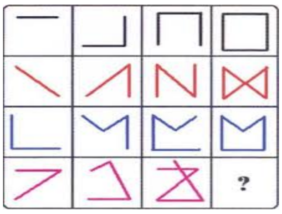


Figure 12: test one.

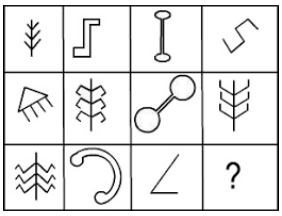


Figure 13: test two.

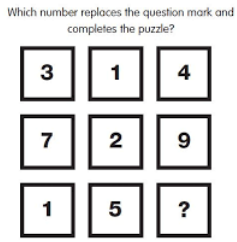


Figure 14: test three.

(a) **1** is the only symbol which appears twice,
(b) **1** makes the third row symmetric,
(c) **1** makes the configuration of three **1**'s in the table symmetric.

And – this is not so much ergo–significant,

(d) **1** complete the sequence **3 2 1** on the principal diagonal,
(e) **1** complete the column of perfect squares.[65]

(One cannot but recall the "commutative square" from O'Henry's story *Squaring the Circle*,

$$\begin{tikzcd} \text{forest} & \bullet \arrow[l, hook'] \arrow[d] \\ \text{city} \arrow[from=u, d] & \blacksquare, \arrow[l, hook'] \end{tikzcd}$$

which pinpoints the reason why diagrammatic IQ tests favour square minded city dwellers.)

In a similar spirit, students of English are given "complete the diagram" exercises.

I. dog runs – dogs run
cow walks – cows [...?]

Next,

II. A dog bites two postpersons. $\rightsquigarrow$ The persons curse the animal.
A milkmaid milks two animals. $\rightsquigarrow$ The [?] lick the [?].

And

[65]Ergologic is robust, it is not finely tuned, it may leave some items (e.g. "**7**" in the numerical square) unaccounted for and it doesn't presuppose (yet doesn't exclude) an ability to manipulate with numbers. This is different from strict, let them be tacit, rules of diagrammatic puzzles. IQ tests are designed by and for those who raised with respect to these rules.

III. A dog bit two postpersons who cursed the animal.

A milkmaid milked two animals ?.... .[66]

If you are a mathematician, these examples bring the ideas of *commutative diagrams* and *diagram chasing* to your mind you observe that II&III make a kind of cubical diagram and you conjecture that there should be more elaborate diagrams hidden in the syntax and semantics of languages.

I guess many of these are known to linguists, who, certainly, use different terminology, and, later, we shall see that significant parts of (not only linguistic) nets $\mathcal{NET}$ consists of such diagrams.

2.7 Reductions, Quotients, Classifications, Clusterizations, Generalisations, Descriptions.

In descriptive terms, *reduction* of a signal-unit S is something, denote it S, which is obtained by curtailing S, e.g. by removing the redundancy from it.

Photographs and *shadows* of 3D-objects are familiar 3D → 2D reduction of geometric signals.

Neural signals received by the primary visual cortex from the retina are reductions of the light signals which reach the retina.

In general, a *reduction* $\mathcal{R}$ is a transformation

$$S \mapsto \mathsf{S} = \mathcal{R}(S)$$

which applies to the members of a large collection $\mathcal{C}$ of signals S, e.g. to what we call *flow of signals*. An essential feature of such an $\mathcal{R}$ is that it "respects" structures of objects and of collections of these it applies to.[67]

A "collective reduction" $\mathcal{R}$ applied to a $\mathcal{C}$ can result in *something* smaller than $\mathcal{C}$, since $\mathcal{R}$ may identify certain signals S from $\mathcal{C}$. Mathematician calls this *something* the $\mathcal{R}$-*quotient*, of the collection $\mathcal{C}$.

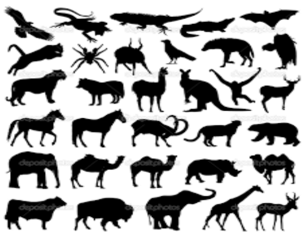
For instance, different objects, such as human faces, can have indistinguishably similar shadows; therefore, there are fewer possible shapes of shadows than of the real $3D$-objects in the world.

On the other hand, since the "space of shadows" is (very) high dimensional, members of many realistic collections $\mathcal{C}$, e.g. of body-forms of different animals, are distinguishable by their shadows.

But albeit the concept of quotient becomes vacuous in such a case, the "collective reduction" effectuates a *compression* of $\mathcal{C}$ by enhancing similarities

[66] Cows licking (milk)persons are uncommon in English, same as dogs milking (post)persons. No mysterious intuition behind it, just a bit more of *semantic diagram chasing*.

[67] Mathematical concepts of *(homo)morphism* and *functor* may serve as first approximations to the desired formal definition of this "respect".



between different signals. Thus, for instance, shadows of different animals, albeit different, may have closer resemblance than these animals bodies themselves.[68]

Besides being *affected* by reductions, similarities serve as a *source* of reductions.

Namely, certain *similarity criteria* $\mathcal{SIC}$ define a division of signals S into $\mathcal{SIC}$-*similarity* or *equivalence classes*, where the signals within each class are mutually $\mathcal{SIC}$-similar, while the members of different classes are $\mathcal{SIC}$-dissimilar. Then the $\mathcal{SIC}$-reduction of an S is defined as the $\mathcal{SIC}$-equivalence class of S.

For instance, one may divide English words into two classes according to their lengths.

I. short words: *three or less* phonemes (usually, four letters or less),

II. long words: at least four phonemes (usually five or more letters).

A more elaborate and not clearly/uniquely defined is the partition of words according to their frequency.

I. frequent words, say, *100 most common words*:

the, be, to, and, a, in, that,..., want, because, any, these, give, day, most, us.

(These make about 50% of the vocabularies in most texts.)

II. infrequent words.

On a higher level, there are classifications of the words by their roles in sentences, where the simplest of these is the following.

I. function words: the, she, is, in, without, some, got, and,....; there are about 300 of this kind of words in English.

II. content words: man, old, dog, black, walk, today, tired, frequently, ...

This bluish classification is refined by dividing words into parts of speech, where, customary, one counts 8-11 of these in English.

Classification/reduction of textual units goes along with reduction of functional connectives between them. These are recorded as "higher order" links in networks $\mathcal{NET}$.

[68]Reductions may also suppress similarities: a striking resemblance between a red-furred dog and its red-haired owner will be lost in a black and white photograph.

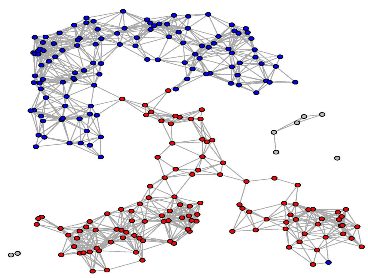
Many of these links, e.g. *noun*$\leftrightsquigarrow$*verb*, are collected in grammar textbooks but most of them, as we mentioned earlier, are inaccessible for our conscious minds.

In mathematics, similarities which allow *non-ambiguous* division of signals into classes are called *equivalence relations*; the essential feature of these is *composability*:[69]

there is a *composition rule* between (computational) processes which establish such similarities, say $A \underset{\alpha}{\leftrightsquigarrow} B$ and $B \underset{\beta}{\leftrightsquigarrow} C$, which results in similarity $A \underset{\alpha\cdot\beta}{\leftrightsquigarrow} C$.

But unrestricted composability is rare in the "real life" and one resorts to *clusterization* instead of classification, where

a *cluster* is a group of signals or of more general "objects" with significantly higher density of similarity links between them than between the members of this cluster and outsiders.



GENERALISATIONS. Much of human thinking – this is clearly visible in mathematics – consists of generalising older concepts and ideas as in the following.

EXAMPLE. Writing small letters by a pen on a list paper takes some time and some effort to learn, but then writing large letters by chalk on the blackboard comes painlessly and almost instantaneously. Somehow – nobody(?) knows how – your nervous motor system system generalises:

Writing small letters by a pen on a list of paper

$\swarrow$

Writing letters $\rightsquigarrow$ writing big letters by chalk on a blackboard[70]

Besides "simple" reductions there are *higher level* transformations, which also *condense (compress), information*[71]These are exemplified by

[text X]$\twoheadrightarrow$[summary of X]

and by *verbal descriptions* of images and situations

Observe, that unlike "true reductions", *descriptions are non-composable*:

description of description of an X is *not* a description of X.[72]

[69]We avoid the word *transitivity* since the "formula" A *implies* B is indefinable in the present context.

[70]It seems, this generalisation includes *writing by your foot*.

[71]The meaning of this "information" will be explained later on.

[72]Composable transformations appear in mathematics under names of *morphisms* and *functors*, see sections???. But there is no accepted general mathematical notion expressing the idea of description of the grammar of language A in terms of language B.

2.8 Ambiguities, Annotations and Contextual Disambiguation.

ANYONE WHO VISITS A PSYCHIATRIST
OUGHT TO HAVE HIS HEAD EXAMINED.
SAMUEL GOLDWYN

Figure 15: bat one.

Figure 16: bat two.

Reduction of information in a flow of signals ♫$_{SIG}$ goes hand in hand with *disambiguation by annotations* of ♫$_{SIG}$, which *add* information to ♫$_{SIG}$.

This information may be partly derived from a priori knowledge and partly from the cues offered by the context.

For instance, since the two meaning of *bat* are widely separated, a few words around "bat" tell you (we shall explain the details of this later) what actually happened when

somebody was struck in the head by a blind bat.

But designing a program which would figure out who had the umbrella(s) in these sentences is more of a challenge.

The angry lady hit the thief with a red umbrella over his head.
The lady hit an elderly passerby with a black umbrella over his head.

2.9 Color of Meaning and the Problem of Precision.

,

NO SHORTAGE OF PEOPLE WHO
ARE WAITING TO PREACH YOU
ABOUT THE *meaning of life,*
BUT THERE IS NOBODY TO TELL YOU
WHAT THE *meaning of meaning* IS.

Conclusion after an unsuccessful search for
a quote on "meaning of meaning" on Google.

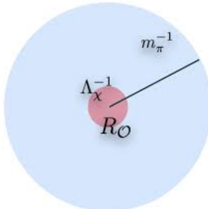
...IMPOSSIBLE TO SAY ANYTHING WITH ABSOLUTE PRECISION,
UNLESS THAT THING IS SO ABSTRACTED FROM THE REAL WORLD
AS TO NOT REPRESENT ANY REAL THING.

RICHARD FEYNMAN

Tendency and ability to classify are key attributes of $\mathcal{UNDERSTANDING}$, where the bottom line of classification in understanding a language is distinguishing between meaningful and meaningless sentences. This is non-ambiguous as far as the cores of the languages are concerned: all (sane) English speaking entities $\circledcirc_{\mathscr{ENG}}$ would agree on which of the two following sentences is meaningful and which is not.

A naked man was seen by a witness repeatedly traversing an intersection walking on his hands.

An American anthropologist has been repeatedly observed by an evolutionary ornithologist traversing a pool of gasoline.[73]



On the other hand, mathematicians, physicists and philosophers may put the following sentence in different categories.

Auto-unification of singularities derived from the chiral perturbation expansion of the gravitational Lagrangian is indicative of paradoxical dissipation of the Bekenstein- Hawking entropy.

There are many schools of thought of what the meaning of meaning is[74], where the relevant for us is the so called *distributional hypothesis:*

✷ *the* meaning *of a word W is determined by arrangements of other words functionally interacting with W.*

This seemingly over-formalistic "definition" can be reconciled with the naive idea of meaning – reference of W to a certain $\mathcal{R}$ – object, event or property from the "real world"[75] $\mathscr{REAL}$ – if this "definition" applies to W immersed into a sufficiently broad, *not purely linguistic*, flow of signals emanating from $\mathscr{REAL}$.

On the other hand, if we apply the distributive definition of meaning to the word *meaning* itself and single out unquestionably *meaningful* instances of the usage this word, such as in the sentence "The words W_1 and W_2 have *similar* meanings", then we realise that individual meanings serve as *tags* or *colours* attached to certain (functional) *similarity links* in the understanding network $\mathcal{NET}$, while the concept of meaning would stand for the full class of these colors.

Disambiguation of ✷. Different interpretations of "W", "functionally", "arrangements" and even of "words" may result in different outcomes for the color of meaning(W).

For instance,

- "W" may stand for a word in a *particular text* or a word in a *dictionary*;
- "functionally" could be *syntactic* and/or *semantic*, where the latter may essentially depend on usage of W in several texts;

[73] Automatic determination of antecedents of "his" (man/witness) and of ""traversing" (anthropologist/ornithologist) in these sentences is difficult. But $\mathcal{NET}$ parsing algorithms we are looking for must be good enough to take care of this.

[74] See, e.g. *The Term Meaning? in Linguistics* by Allen Walker Read, ETC: A Review of General Semantics, Vol. XIII, No. 1,1955
http://www.generalsemantics.org/wp-content/uploads/2011/05/13-1-read.pdf

[75] Pythagorean theorem, toothache suffered by professor Moriarty, color of the skin of little men in flying saucers, free will in the soul of a righteous man – everything is "real" compared to the words naming them.

• "arrangements" and "words" may refer either to *instances* of these in a *vicinity* of a particular W in a text or to *classes* of these derived from *multiple usage* of these classes of arrangements and words.

Unless you switch to mathematics, this kind of indeterminacy will be relentlessly present in everything you say, but a usage of mathematical language is also riddled with problems.

A. A mathematical encoding of an idea may miss the essence of this idea.

B. A precise mathematical description of a simple situation may be unbearably complicated and confusing.

However, since we eventually want to design an $\mathcal{UNDERSTANDING}$ computer program, we need to be 100% precise and somehow to resolve these problems.

2.10 Abstraction, Sophistication, Complexity and Hierarchy of Perception and Ideas .

Concepts and ideas differ in their degrees of abstraction, sophistication and complexity. But an actual assignment of such a degree to an idea, that is a nod, edge or a color in the understanding network $\mathcal{NET}$ in the mind of an intelligent entity ☺$_{\mathcal{INT}}$, depends on:

- a pool of signals ♫$_{\mathcal{SIG}}$ available to ☺$_{\mathcal{INT}}$;
- the order in which ☺$_{\mathcal{INT}}$ was exposed to ♫$_{\mathcal{SIG}}$;
- algorithms/programs in the mind of ☺$_{\mathcal{INT}}$ transforming signals to ideas.

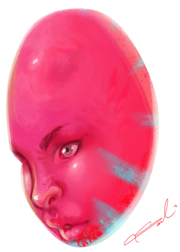


Figure 17: blob one.

Figure 18: blob two.

Figure 19: blob three.

For instance, an intelligent human and an equally intelligent gigantic squid may not agree on which of the above three blobs are abstract and which are concrete.

On the other hand, if the intuitive ideas of abstraction, sophistication and complexity, visualised as *hierarchical tree-like structure(s)* superimposed on $\mathcal{NET}$, are derived from combinatorics of $\mathcal{NET}$ in objective mathematical terms,[76] this would be acceptable to intelligent entities of all shapes and colors.

[76]Mathematics may loose objectivity of innocence when it ventures to the "real world" but there is no alternative to mathematics when it comes to expressing abstract ideas which are both objective and nontrivial.

2.11 Morphology and Syntax, Semantics and Pragmatics

Borrowed from linguistics, this terminology applies to other flows of signals as follows.

Morphology concerns structural patterns of relatively small perceptual units, such as *words* or *sketches of animals*, where the (approximate) number of these (possibly reduced) units is what we call *listable*: say, in the range 10^5-10^9.

Syntax deals with arrangements of smaller units into larger ones, such as sentences or composed images in visual fields.

The set of rules which define composition of such arrangements are supposed to be listable (10^3-10^6) but the numbers of (potential) arrangements abiding these rules are by no means listable being in the range 10^{20}-10^{50} or more.[77]

Semantics, which unlike syntax is non-local,[78] imposes further constrains on usage and generation of small sentence-like arrangements of basic units.

These constrains are seen in distributions of patterns of small units in large pools of signals e.g. words in books, and they are not formalisable, at least not in the forms accessible to conscious human minds.

For instance, the rules of chess – the syntax of the game – serve to define legal pairs of consecutive positions, while semantics tells you what positions of chess pieces on the board and pairs of positions are likely to come from a true game and which are are random derivations of the syntactic rules, where the number of legal positions is estimated about 10^{50} while there are, probably, only 10^9-10^{15} semantically plausible positions and consecutive pairs of these, which could arise, say, in grandmasters' games. [79]

Pragmatics of a flow of signals ♫$_{\mathcal{SIG}}$, as we understand it, is about a use of ♫$_{\mathcal{SIG}}$ for *representing/reflecting/referring to* other signals ♫$_{\mathcal{SIG}}$.

This happens exclusively to man-made flows ♫$_{\mathcal{SIG}}$, mainly in natural and artificial languages[80] which have capacities to represent within themselves fragments $\mathcal{F}$ of (structures of) other flows including, this is essential, *fragments $\mathcal{F}$ of themselves.*

For instance, one can summarise the contents of a visual scene in English as well as to describe (possibly not in full) the grammar of English in mathematical terms. In most cases such a representation results in reduction of $\mathcal{F}$, albeit it is hard to interpret extraction of grammar from a language as a true reduction in the sense of section ???.

2.12 Referential Links and Self-referentiality in Languages.

The human mind has an obsession of representing non-linguistic signals, be they external or internal, by linguistic ones.

Because of this,

[77] If one allows "infinite recursion" mentioned in section 18 and equate $5 = \infty$, then one may reach 10^{100} or even 10^{200}, which makes little difference anyway.

[78] This *locality* means a possibility of (more or less) explicit description of syntactic rules in terms of individual small chunks of signals flow they apply to.

[79] A chess master, probably, keeps in his memory tens of thousand (millions?) learned and made by his mind (chunks of?) positions as well as fast algorithms for comparing them with newly incoming one.

[80] In limited way, paintings and, up to a lesser extent, music also can reflect other, internal and as external, flows of signals.

language becomes a mirror in which the mind contemplates itself.

And being itself a part of the mind, this mirror reflects itself, which is facilitated by a presence of syntactic (self)*referential connectives* in languages.

No other natural or artificial flow of signals has this property, there is no comparable means of self-representation in music, in board games (chess, go), in the visual arts. Even mathematics can't speak about itself without a resort to the natural language.

Common instances of referential connectives in languages are modulated by noun↜pronoun-linkages.

I never saw a *man* ↜ *who* looked With such a wistful eye
Upon that little *tent of blue* ↜ *Which* prisoners call the sky

These connectives neither serve for functional conglomeration of units nor can they be regarded as bona fide similarities or reductions but they may bridge units which are spatially or temporally far each from another by pointing to similarity linkages:

connectives↜they↭may bridge.

Or,

Meanwhile, the wandering goat, which we left when it got itself into trouble trying to pacify the crazy cow, began... .

Such connections may depend not only on pronouns:

a dog is a vicious animal who fights back when attacked,

and sometimes they (connections) need no pronouns at all:

a dog is an animal.[81]

Automatic detection of referential links, even a correct identification of antecedents of pronouns on the basis of general principles is difficult. Computational linguists, who are granted the knowledge that they are dealing with a language are still unable to fully resolve this problem even having an access to the formalised grammar of it. Even children master the use of pronouns, only by the age of 3 (rarely 2) years. [82]

Mastery of self-referentiality is an indication of human understanding a language. If your email interlocutor can properly respond to the following question:

Do you recall which word in the message I wrote to you after we spoke

[81] This animal is half way to he, but "a man is he" is improbable. Yet, he travels the fastest who travels alone and "man is he who" is also on Google.

[82] Do you think teaching grammar to babies would accelerate this process?

about homeopathic palliation of crazy cows *contained more than ten letters?*

you can safely presume this is not a currently existing computer program.

But *universal* learning algorithms should be up to this task, where the needed level of $\mathcal{UNDERSTANDING}$ may be – this is an optimistic conjecture– achieved by a certain *self similarity* in the architecture of the corresponding colored networks $\mathcal{NET}$.

The first step toward a design of such an architecture is a particular organisation of the ensemble of colors in $\mathcal{NET}$:

the colors of $\mathcal{NET}$ *must be arranged in a colored network of their own, say* $\mathcal{NET}'$*, which should be many times smaller than* $\mathcal{NET}$ *but with the combinatorics structurally similar to that of* $\mathcal{NET}$*.*

Then, with a caution, we proceed to some kind of $\mathcal{NET}''$ and, bearing in mind the inconsistency and infinite regress problems associated with self-referentiality, terminate at $\mathcal{NET}'''$ if not earlier.

To see what iterations of self-references do, try it with "the meaning of". The first iteration, meaning of meaning tells you what the (distributive) meaning really is. But try it two more times, and what come out of it,

meaning of meaning of meaning of meaning,

strikes you as something meaningless.

In the universe of languages, besides the learning arrow from signal flows $♫_{\mathcal{SIG}}$ to the networks $\mathcal{NET}$ which $\mathcal{UNDERSTAND}$ these flows, there is an opposite arrow from $\mathcal{NET}$ to the linguistic flows:

collectively, linguistic networks in our minds generate such flows.[83]

How does it work? What necessitates $\mathcal{NET}$ *– a fragment of the mind built from the raw material of linguistic flows* $♫_{\mathcal{SIG}}$ *– to be itself an active source of such flows?*[84]

Superficially, this is clear enough:

records of travelings along paths in a "network of rules" $\mathcal{NET}$ *according to these rules result in such flows.*[85]

But a true answer, if at all, must come in terms of relations between the "categories" of linguistic signals and of the corresponding networks, where these

[83]There are similar reversed arrows for other flows but they are mainly confined to the subliminal mind.

[84]Next to *language imitation*, stands a more difficult problem of *new language generation*. A mathematical model of the initial stage of such "generation" – *phonemes fixation* – is suggested by Pierre-Yves Oudeyer in *Self-Organisation in the Evolution of Speech (2006).*

[85]In familiar mathematical terms, this is the correspondence

directed graphs $\leftrightarrow$ *Markovian dynamical systems* $\subset$ *regular languages.*

"categories" are joined by two "dual functorial" arrows,

$$♫_{\mathcal{SIG}} \mapsto \mathcal{NET} \text{ and } \mathcal{NET} \mapsto ♫_{\mathcal{SIG}}.$$

The most common reactions common to animals is activation of skeletal muscles resulting in movements of bodies and, specifically human and depending on learning, speech.

Besides learning to reproduce speech, humans, at least some of them, are able to *respond to musical melodies* by new melodies and some, let it be unwillingly, to *generate visual images* in their minds. Despite apparent dissimilarity between these, all three, probably, follow universal rules of *creative imitation of signals*.

Apparently, the brain creates music and speech, paintings and *hypnagogic hallucinations*, by playing with flows of signals. The brain plays with these flows as a puppy with its toys: no purpose, no goal, no meaning.

2.13 Symmetries, Transformations, Arrows, Categories, Morphisms, Functors

Brain's responsiveness to similarities is vividly demonstrated by delightful sensitivity of the human (and animal) visual systems to symmetry. [86]

Mathematically, symmetries are associated with *transformations* which either keep something *unchanged* or *changing it in a particular way*.

In vision, common transformations are:

A. *(Approximately) projective maps from 3D-space to the 2D screen of the retina in the eye.* Straight line segments under these maps remain (approximately) straight.

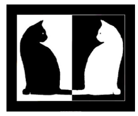
B. *(Rigid) motions* of the visual field, e.g. under movement of the body or the eye of an observer.

(The human visual system has no problem with *parallel translations* but the invariance of images under *all rotational symmetries* is harder to grasp, probably, due to *non-commutativity* of the group $O(3)$.)

[86] There many articles relating to the visual perception of symmetry in humans and animals. In particular, data are presented for symmetry perception in infants, children, older people ..., e.g. *Symmetry perception*, Peter A. van der Helm peter.vanderhelm@ppw.kuleuven.be and V GIANNOULI www.encephalos.gr/pdf/50-1-03e.pdf

C. *Similarity (rescaling) transformations* of the (retinal) image under variations of the distance to the object in view.[87]

Most widespread in the (animal) world is *bilateral symmetry*, which also plays a distinguished role in mathematics, where it is called *involutive*, or ±1, symmetry.



In languages, (imperfect) ±1 symmetry, similar to **black**$\overset{\pm 1}{\longleftrightarrow}$white in vision, pops up as correspondence between *semantic opposites*:

against↔for, ancestor↔descendant, to enter↔to leave.

Also the statement⟷question and active$\overset{\pm voice}{\longleftrightarrow}$ passive *transformations* of sentences are of the ±1 kind.[88]

Below is an example of (quasi)reversible but not semantically involutive transformation.

[A drunken man in a black suit entered city hall. The angry chimp bit off his finger. The surgeon reattached the finger.**]**⇛**[**The surgeon reattached the finger of the drunken man, who, upon entering city hall in a black suit, has had it bitten off by the angry chimp.**]**[89]

???

TRANSLATIONS FROM ONE LANGUAGE TO ANOTHER. These are quasi-reversible and quasi-composable transformations, where the mathematical definition of these "quasi" will come later, but where the following instance of composed translation English → Russian → English gives an idea.

The spirit is strong but the flesh is weak.

$\circlearrowleft$

The liquor is fine but the meat is stinky.

There are also all kinds of transformations which are far from being reversible, such as *reductions* from section???.

It happens to be an unexpectedly fruitful idea in mathematics to allow all kind of non-reversible transformations and study configurations of arrows representing such transformations with an emphasis on the composition relations between these arrows, rather than on transformation themselves.

The domain of mathematics concerned with such "transformational arrows" is called *the category theory* where the arrows themselves are called *morphisms*.

[87]It remain largely unknown how the visual system deals with ABC. Probably, A and B which are learnt by children before the age of 12 (6?) months, are assisted by the "experiments" performed by the child's *ocular motor system*, e.g. by brain's control over (and recognition of) the curvature of saccadic eye movements (for A) and by correlating the eye and the head movements (for B). But a correct size evaluation independent of the distance (i.e. C) is acquired later in life, if at all. https://books.google.fr/books?isbn=1483290069

https://books.google.fr/books?isbn=0199539782

http://boards.straightdope.com/sdmb/showthread.php?t=687878

[88]In mathematical usage, correspondences, unlike transformations, do *not* have to be *one-to-one*. Thus, strictly speaking, *statement↔question* is a correspondence rather than a transformation: a declarative sentence such as "A dog bites a postman", for instance, has (at least) three significantly different interrogative forms. Also the question "Why postmen don't bite dogs?" doesn't come with a unique answer.

[89]Is there a law in English grammar that would rule out *"it"* standing for *black suit* bitten off by the angry chimp?

We have already met non-reversible *reduction arrows* and instances of their compositions, such as

cat → animal → noun → word
↓ ↘ ↑
furry moving → something moving

There is an opposite in meaning class of arrows, which is ubiquitous, but may seem too simple to be relevant, which also must be accepted in the to the "(quasi)composable transformations club"

These are *inclusions*

part ↪ whole.

e.g inclusions of words to sentences, where, following the dictum of the category theory, we admit compositions of these with reductions.

ABOUT DESCRIPTIONS. These in general neither invertible nor composable but

- *detailed verbal descriptions* ⇆ movements and images.

are (quasi)reversible. These, however are closer in spirit to what is called *functors* as they operate between different categories. In fact translation between different languages are also closer to functors than to morphisms where the latter are suppose to describe transformations between objects in the *same category*.

However, it should be noted that the standard notions of the category theory must be significantly modified and adjusted to properly apply to the above examples (see sections ???).

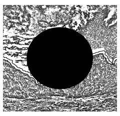
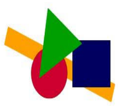
2.14 Descriptive Understanding: Prediction, Correlation, Reduction.

THE GRAND AIM OF ALL SCIENCE IS TO COVER THE GREATEST NUMBER OF EMPIRICAL FACTS BY LOGICAL DEDUCTION FROM THE SMALLEST NUMBER OF HYPOTHESES OR AXIOMS.

ALBERT EINSTEIN

This idea has not been invented by physicists, neither computer scientists are responsible for it.

Compression of information and *reconstruction of images from incomplete data*, albeit not in the extreme form $\frac{greatest}{smallest}$, has been practised by your visual system for tens of millions of years, well before you became human.

Anticipation of what comes next also makes the backbone of *UNDERSTANDING* language, where this is closer to how it is in physical science:

one expects to hear not only and not so much a specific word but a *class* of words or a *particular structural pattern* in the incoming flow of speech.

For instance, a sure way to drive the host of a house party crazy is responding with "yes, thank you very much" to his – *"Which one do you like better?"*

In general terms, a *descriptive understanding* of a flow ♫$_{\mathcal{SIG}}$ consists of a list of algorithms algo for computing certain *observables features* of ♫$_{\mathcal{SIG}}$ and a table of correlations between pairs, triples, etc. of values of different features, ftr=*feature* ($♫_{\mathcal{SIG}}$),

$$corr_{ijk...} = corr(\mathsf{ftr}_i, \mathsf{ftr}_j, \mathsf{ftr}_k, ...).$$

Predictions are special kind of correlations, where the expected values of certain features ftr of the flow $♫_{\mathcal{SIG}} = ♫_{\mathcal{SIG}}(t)$ *after* a given moment t_0 are expressed in terms of (possibly different) ftr of $♫_{\mathcal{SIG}}$ *prior* to this t_0.

In physics, correlations are expressed in terms of *simultaneous probability distributions* of observable values of ftr($♫_{\mathcal{SIG}}$).

But beware: applicability of mathematical probability to the "real world" is not god given to you – if this works in physics it is essentially because of the space-time symmetry of the Universe.

No comparable luxury of symmetry is available in the structures of living organisms and in grammars/semantics of languages. "Probability of a mutation" and "probability of a sentence" are of different kinds from "probability of radioactive decay of an atom of Radium in the next two seconds"[90] or "probability of a billion ton asteroid hitting Earth in the next two thousand years".[91]

The ideas of "randomness" and of "correlation" in PHYSICS, biology and linguistics need different mathematical representations of probability that we shall discuss later in this text.[92]

But no matter how correlations are understood, one needs, realistically, a structural organisation of the set of all $corr_{ijk...}$ or rather of the set of the corresponding algorithms $\mathrm{algo}_{ijk...}$.

The automatic learning process which one wants to design must deliver these algo organised according to a *hierarchy of consecutive reductions* of observable patterns in $♫_{\mathcal{SIG}}$.

[90]For the most abundant ^{226}Ra with the half life approximately 1600 years, this probability is close to 0.000 000 000 04.

[91]This can be estimated somewhere between $\frac{1}{10^3}$ and $\frac{1}{10^4}$, but "probability" makes little sense in this case. In fact, there are only a few million asteroids of this size in the Solar System. When/if all of them are catologized and their orbits determined, one will be close to certainty of when such a collision would take place.

[92]Absence of commonly accepted "non-physical" notions of probability often results in misunderstanding and leads to controversy about the meaning of this concept when it is used away from a few well established branches of physics.