

From Machine Learning to Machine Understanding

Dr Yan M Yufik

Scientia

FROM MACHINE LEARNING TO MACHINE UNDERSTANDING

Despite dramatic advances in neuroscience and biology in the 20th and 21st centuries, our understanding of the brain remains very limited. **Dr Yan M Yufik**, Head at Virtual Structures Research Inc, USA, is a physicist and cognitive scientist who has spent over 20 years combining experimental findings and theoretical concepts in domains as diverse as neuroscience and thermodynamics to form a theory of the brain. His focus has been on elucidating the mechanisms underlying human understanding and applying the results to the design of machines that can not only learn but understand what they are learning.

What is Machine Learning?

Artificial intelligence (AI) is concerned with designing computing machines that can replicate or even amplify human cognitive capacities. Recent news has been awash with examples of artificial intelligence that seem to come straight from science fiction, including self-driving cars, facial recognition, and virtual assistants such as Siri and Alexa. AI approaches are abundant in military applications and can also be found in less obvious places, powering algorithms used by online marketing companies, social media websites, financial institutions, and medical diagnostics.

These accomplishments are due to advances in machine learning, built predominantly on the idea of neural networks that originated in the middle of the last century. Three factors contributed to the success of this idea: large financial investment in the development of learning algorithms, recent theoretical breakthroughs in the development of these, and nearly a billion-fold increase in the efficiency of computing devices.

This increase in the efficiency of computing devices was critical: machine learning demands massive amounts of computation. The implementation of neural networks began around 1960 using the computers available at that time. Matching computing power available today for \$500 to that available back in 1960 would require an array of 1960-type computers at the cost of about \$9 trillion (adjusted for inflation). The development of AI technology has relied on the rapidly increasing efficiency of computing machines. Dr Yan M Yufik, Head at Virtual Structures Research Inc, USA, argues that the development of human intelligence followed a path orthogonal to that pursued in AI.

Machine learning employs a variety of statistical learning methods. Despite some differences, all these methods are based on the same principle: learning in the absence of understanding. Take, for example, a database of medical records containing data about conditions and the corresponding diagnosis in a multitude of patients. Machine learning can determine patterns in

the data allowing the machine to make responses to combinations of conditions that would appear to represent meaningful diagnostic decisions. In principle, the same learning algorithms can be applied to a database holding records of chess games, allowing the machine to produce responses to chess positions having the appearance of meaningful moves. However, in both cases, learning was concerned only with determining patterns in the arrays of symbols and was oblivious to meaning: no explicit knowledge of diseases or chess rules and strategies has been entered into the machine.

Learning algorithms make it possible to train machines to respond adequately to different inputs while being clueless about the meaning of either the inputs or the responses. For many people, the prospect of 'clueless' machines is a frightening one, especially if they are delegated important decisions with life-or-death consequences, as in medical treatment, car driving, or weapons control.



According to Napoleon, 'the art of war consists, with a numerically inferior army, in always having larger forces than the enemy at the point which is to be attacked or defended.' As per Dr Yufik's theory, thermodynamics enforces 'Napoleonic strategy' in the evolution of the brain: mechanisms are formed allowing allocation of neuronal resources sufficient for dealing successfully with a growing variety of changing conditions while, at the same time, minimising energy expenditures incurred in the operation of those mechanisms. This evolutionary development culminated in the mechanisms of mental modelling in humans (for example, understanding situations in a battlespace requires constructing models capturing fluid relations between battlespace entities. Successful models enable the commander to plan anticipatory deployment and manoeuvring, as required by Napoleon's winning formula)

For the past 20 years, Dr Yufik has pioneered research in the field of machine understanding. His objective is to design machines endowed with a degree of understanding that is sufficient to enable them to carry out complex tasks under novel and unforeseen conditions and to explain their actions and decisions in a manner that is comprehensible and compelling to human users. Reciprocally, he believes that such machines should be able to accept user feedback in a format meaningful to humans and apply it directly when organising their internal processes.

What is 'Understanding'?

Learning is crucial for survival. Even the simplest of organisms can associate conditions with responses and consequences such that when conditions recur, the beneficial responses are reproduced, and the harmful ones avoided. Learning serves well for as long as the conditions recur but fails when they change. Such failures can be particularly damaging if the learned behaviour persists after the changed conditions start penalising responses that were previously rewarded.

According to Dr Yufik, human understanding is an adaptive mechanism serving to overcome the inertia of learning, which includes the ability to timely detect and prioritise changes, allowing the construction of responses to unfamiliar conditions. Understanding is the product of brain activity (i.e., thinking, mental modelling) that is temporarily decoupled from sensory inputs and involves selecting and re-combining memory elements to form new structures (mental models) that help us to both anticipate future changes and accommodate the unanticipated ones.

While learning is tied to past experiences, understanding can deviate from them. Take, for example, the earliest known artefact dated approximately 30,000 years BC which is a figurine depicting a creature with a human body and lion head. Perhaps, imagining such creatures expressed a primitive understanding of important realities, serving a dual purpose of indicating the possibility of encountering opponents with extraordinary (lion-like) strength and ferocity, and allowing advance preparation for such encounters. From imagining chimeric creatures to imagining and designing intelligent machines, the process of understanding involves selective adjustment and re-combination of previously formed memory structures to produce new ones.

Meaning is imputed to such combinations when relations between components are apprehended: first, behavioural repertoires are attributed to memory elements representing objects, followed by imagining how the behaviour of one object can impact the behaviour of other ones. Imagine, for example, a drawing depicting a vase lying on the floor next to a stand and a cat sitting nearby. Imagining a cat jumping and knocking the vase down imputes meaning to the drawing which would have otherwise remained a meaningless aggregation of objects. Note that neither the jumping cat nor



Carl Friedrich Gauss, one of the greatest mathematicians of all time, surprised his teacher in elementary school when quickly adding integers from 1 to 100. While the other students were laboriously moving along the number row, young Gauss grasped the relations across the row ($1 + 100 = 101$, $2 + 99 = 101$, and so on) thus reducing computing the sum to finding the product of 101×50 . Thus, using mental models to capture global relations in the number series yields the problem solution more quickly and less laboriously.

the standing vase is in the drawing, and that cats usually do not attack inedible stationary objects, so the imagined behaviour is a product of adjustment and re-combination rather than merely a recollection of past experience.

The importance of mental modelling has been long recognised in cognitive psychology, but Dr Yufik proposes a specific and central role for it in his theory. Note that you have probably observed different forms of cat behaviour (climbing, running, sitting, jumping, lying on a side, sleeping, reaching, eating, and so on) but images of a sitting cat floating through the air and hitting the vase or other such choices were unlikely to cross your mind. Dr Yufik hypothesises that mental models are synergistic memory structures where all components are amenable to mental variation and are all mutually constrained and coordinated. Such mutual constraining has the dual effect of limiting the range of plausibly imaginable variations and causing co-variations across the structure consistent with any local change. For example, if you imagine varying the height of the stand, your image of plausible cat behaviour will vary accordingly.

The experience of attaining understanding accompanies the formation of mental models where cross-coordination between all components radically reduces the number of degrees of freedom available to them. These benefits of understanding might not be apparent when dealing with a few objects but become obvious in multi-object situations affording many choice combinations, such as in playing chess or fighting battles. Chess machines have to reach the speed of hundreds of million decisions per second in order to compete with humans capable of at most a few decisions per second. Master players compensate for the disadvantage in speed by forming synergistic models of chess positions that confine the analysis to, figuratively, a hair-thin path in the combinatorial space the size of the Pacific Ocean. As a result, bad moves are kept outside such analytic paths and do not come to mind in players who understand the game (no more than illegitimate moves would come to the mind of novices familiar with the rules) while the way forward can be envisioned to a substantial distance (e.g., the astonishing 15 moves 'look-ahead' analysis reported by chess champions).

Step-by-step computing (analysis) is time- and energy-demanding while cross-coordination in mental models is simultaneous and computation-free and, thus, energy-inexpensive. Embracing a physics perspective, Dr Yufik argues that thermodynamics enforces energy efficiency in neuronal systems, and thermodynamic pressure propelled evolutionary transition from protohumans (a hypothetical prehistoric primate) to humans with a desire to understand themselves and their world.

According to Dr Yufik's theory, the basic functional units in the brain are not neuronal networks but neuronal packets which are groups of tightly associated neurons underlying the perception of objects. Inducing different firing patterns inside packets underlies apprehending and imagining behaviour variations (e.g., imagining a sitting, running, or jumping cat involves inducing different firing patterns in the 'cat packet'). Dr Yufik further proposes that apprehending relations between objects involves establishing co-ordination between successions of firing patterns in the corresponding packets. Packets form as a result of self-organisation in associative networks, not unlike the formation of raindrops in water vapour. Energy barriers at the packet boundary (again, not unlike boundary surfaces in raindrops) make packets stable and amenable to composition into models. Crucially, forming coordinated compositions of packets and manipulating them is more energy efficient than manipulating individual packets.

While Dr Yufik's ideas are speculative, modern experimental techniques have started delivering data that seem to support them. His theory has also led to conclusions consistent with a principle of brain operation advanced recently by Professor Karl Friston, asserting that processes in the brain are driven towards minimising surprises. In a joint paper, Dr Yufik and Professor Friston argue that principles of surprise minimisation



In the now-famous experiments, American psychologist Edward Thorndike placed hungry cats in cages equipped with a lever for opening the door. When frantically thrashing around, cats would accidentally push the lever and thus free themselves. After a series of repetitions, cats would learn the requisite action and, when placed in the cage, proceed to push the lever without delay. However, when one of the sides in the cage was removed, the trained cats would still implement the lever-pushing routine instead of simply walking out. Cats were learning – but failed to understand either the situation or what they had learned.

and energy cost minimisation are mutually consistent and complementary: pressure to reduce energy costs sculpts and fine-tunes mental models so that more reliable predictions are produced diminishing future correction costs. Discovering the dual benefit of packet coordination by the evolution (more accurate and reliable predictions at lower energy costs) may be responsible for the emergence of sapience about 100, 000 years ago.

According to Dr Yufik, the gradual build-up of sensory-motor coordination machinery in the brain could plausibly have brought evolution to the point where a one-step transition from protohuman to sapience was possible. Think of walking and carrying a cup of hot coffee in one hand and a pile of documents in the other: the process requires precise dynamic coordination of multiple muscle groups to avoid spilling the coffee and or the papers. Imagine reaching a door and trying to open it: the coordination pattern needs to be quickly re-organised to meet the challenge. Dr Yufik's hypothesises that the machinery of sensory-motor coordination in manipulating external objects richly developed in the protohuman was co-opted and re-purposed for manipulating internal 'objects'. And with that, understanding appeared on the scene, enabling advances in technology at a blistering pace: from improving gadgets for throwing projectiles to hit distant objects to designing space craft capable of reaching the moon.



The Gnostron Framework and the Future of Machine Understanding

We should bear in mind that Dr Yufik's theory does not align fully with mainstream AI. The neural network approach argues that intelligence derives from pattern recognition while the neuronal packet approach proposed by Dr Yufik derives intelligence from pattern coordination. Critically, the former makes predictions by extrapolating from the past while the latter derives predictions from understanding the past.

Nonetheless, Dr Yufik proposes that the approaches are not mutually exclusive and can be integrated within a unifying mathematical framework. The neuronal packet approach has been expressed in architecture and mathematical formalism dubbed 'gnostron' (which is different from and complementary to 'perceptron' formalism and architecture that initiated the development of neural networks).

Perceptron architecture includes a fixed set of interconnected neurons (i.e., a neuronal network) while gnostron processes operate on a neuronal pool, selecting and combining neurons into packets and packet compositions (models) that can be tried out and then matched dynamically against the streaming input. The process allows dynamically optimising responses in complex situations, as in the battlespace where conditions are fluid and never twice the same.

Application of the gnostron framework in machine understanding is in the early stages, requiring mathematical and, possibly, hardware engineering approaches different from those currently employed in AI. These developments aspire to deliver a new generation of AI systems capable of carrying out complex tasks on small energy budgets and in a manner the users can trust and understand. Such systems will be able to envision immediate and distant consequences of their actions and explain their decisions in terms relevant to both internal operations in the machine and human understanding: values, objects, behaviours, and relations. Developing machine understanding is a worthwhile challenge: the distance between future AI and its present version cannot be less than that between humans and their evolutionary ancestors.



Meet the researcher

Dr Yan M Yufik

Head, Virtual Structures Research Inc

Potomac, MD

USA

Dr Yan M Yufik holds a PhD in physics and received postdoctoral training in cybernetics and cognitive science. Dr Yufik currently heads Virtual Structures Research Inc, a non-profit company that aims to facilitate the study of biological and artificial intelligence. Along with colleagues, Dr Yufik is pioneering the field of machine understanding, an area of artificial intelligence that uses biophysics and neuroscience approaches to simulate human understanding in artificial systems. He holds five US patents and has published numerous papers and several book chapters on the subject.

CONTACT

E: imc.yufik@att.net

KEY COLLABORATORS

Professor Thomas B Sheridan, Massachusetts Institute of Technology (Emeritus)

Professor Karl Friston, University College London

FUNDING

US government, private sources

FURTHER READING

YM Yufik, The understanding capacity and information dynamics in the human brain, *Entropy*, 2019, 21, 308, 1–38.

YM Yufik, K Friston, Life and understanding: Origins of the understanding capacity in self-organizing nervous systems, *Frontiers in Systems Neuroscience*, 2016, 10, article 98.

YM Yufik, Understanding, consciousness and thermodynamics of cognition, *Chaos, Solitons & Fractals*, 2013, 55, 44–59.

YM Yufik, TB Sheridan, Swiss Army Knife and Ockham's Razor: Modeling and Facilitating Operator's Comprehension in Complex Dynamic Tasks, *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 2002, 32, 2, 185–199.