The Past, Present, and Future of Computer Systems and Artificial Intelligence - Practical Applications and Ethical Implementation and Innovation

Shane Breitstein (Author)
Computer Science and Engineering Department
University of North Texas Denton, Texas
ShaneBreitstein@gmail.com

Abstract— **In a world where artificial intelligence (AI) is becoming a widespread global phenomenon with endless possibilities of practical applications, engineers are faced with the responsibility of expanding and innovating this field, while also ensuring its ethical and responsible application. Moore's law, Dennard scaling, and Amdahl's law defined the innovation and development of computers from 1965 to today. Alan Turing led the development of computer science and machine intelligence, later followed by advancements including ELIZA, Apple's Siri and more. The development of computing architecture and chips specifically designed to optimize artificial intelligence applications allows for the power of AI to increase, leading the way for efficient uses of AI in the future, in the medical field, education, and many other facets of life. While this innovation occurs, strategies for responsible innovation must also be put in place to ensure a fair and ordered future.**

Keywords— Artificial Intelligence, Neural Network, Computation, Moore's law, ChatGPT, ethical AI usage

I. Introduction

In recent years, global buzz around artificial intelligence (AI) has grown by an extraordinary amount. With a rich and promising future, AI has a wellspring of practical and useful applications in subjects ranging from medical care, mathematics, entertainment, and so much more. In addition to these applications comes countless irresponsible and immoral uses of artificial intelligence, including deepfakes, misleading and inaccurate information from generative AI, biased facial recognition AI, and energy-intensive AI models (with the energy consumed during training emitting hundreds of tons of carbon dioxide) [1]. These unethical applications underscore the imperative of balancing technological progress with ethical considerations and the preservation of our shared values and principles. With these challenges present, engineers face the challenge of continuing to expand and innovate the field of artificial intelligence, while also considering the ethical dilemmas that come with it.

Furthermore, this rapid demand for innovation is occurring at the same time as Moore's Law decline. Moore's law, the 1965 observation that the number of transistors in integrated circuits doubles each year, eventually lowering to doubling every two years in 1975, has defined the progress of computing throughout its development [2]. While Moore's law has been a driving force behind computing progress, its recent slowdown does not signify that progress will stop, just that it must occur in a different manner.

This paper will explore the development of computer systems and artificial intelligence, the changes that allowed artificial intelligence to improve dramatically in recent years, the future trajectory of artificial intelligence, and the ethical hardships that face the innovators of a field as important as artificial intelligence.

II. Development of Computers

During the later half of the 1900's, a surge in computational performance unfolded. Moore's Law, which led this computational revolution, eventually resulted in Dennard scaling and Amdahl's law, shaped the new era of computational innovation, explain the reason computers have gotten smaller, faster, and more power efficient.

Dennard scaling, named after Robert H. Dennard, complemented Moore's law by leveraging the fact that smaller transistors required less voltage and current to operate. This breakthrough allowed for a dramatic increase in clock speeds within computers [3]. Figure 1 illustrates Moore's law and Dennard scaling in effect, showcasing the correlation between transistor count and clock speed from the start of the graph to 2004.

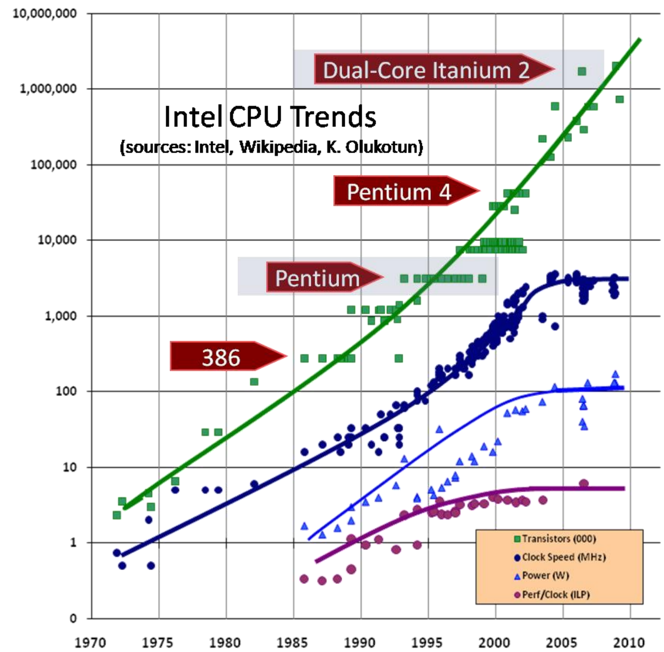

Figure 1: Intel CPU scaling, power consumption and efficiency from 1970 to 2010 [3]

A break from Dennard scaling occurred in 2004, as physical limits reached their climax, seen from the fact that from 1994 to 1998, computer processing units (CPU) clock speeds rose by 300%, while from 2007 to 2011, they only rose by 33% [4]. Engineers responded to this by increasing the number of cores in a computer, whose efficiency is defined by Amdahl's law. [4]

III. Development of Artificial Intelligence

Throughout the development of computer systems, the concept of artificial intelligence has coexisted. Pioneering the field of AI was Alan Turing, who, in his 1950 article "Computing Machinery and Intelligence," detailed both how to create an intelligent machine, and also a test that measures a machine's intelligence, now famously known as the Turing Test [5].

The term artificial intelligence was officially coined for the first time at the 1956 *Dartmouth Summer Research Project on Artificial Intelligence (DSRPAI)*, whose primary object was to "reunite researchers … to build machines that simulate human intelligence" [5]. Two decades of research resulted from the DRSPAI, leading to the development of the ELIZA computer program from 1964 to 1966, the first program to pass the Turing Test [5].

It wasn't all successes, however, because in 1973, the U.S. congress criticized the expensive nature of AI research [5]. Furthermore, the computers at the time simply weren't powerful or quick enough to meet the high demands of researchers at the time, leading to an AI winter, where funds were low, and little research was done [5].

AI research continued in the 1980's, originating from an expansion of algorithm tools and an increase in funds. "Deep learning" artificial intelligence algorithms that learned from experience were the primary focus of AI research during this time [5]. Additionally, Edward Feigenbaum created Expert Systems, a program which questioned experts about how they would respond to a situation and advised non-experts on those situations [5]. The Japanese government joined this 80's AI advent, investing $400 million dollars from 1982 to 1990, revolutionizing computer processing, logic programming, and artificial intelligence through their research [5]. AI research continued into the 1990s, despite a lack of government funding of public attention. In 1997, IBM's Deep Blue chess-playing system defeated chess champion Gary Kasparov, the first machine to defeat a grandmaster in chess [5], which was initially predicted by 1950's computer scientists would occur by 1967 [6].

Artificial Intelligence advancement has skyrocketed in the 21st century, as computer clocks reached gigahertz speeds. The new millennium saw the expansion of autonomous artificial intelligence that interacted with humans, such as robots and self-driving cars. Honda was first to create a first two-legged, walking robot, called the ASIMO [7]. Verbal human interaction became the next focus of AI researchers, where MIT researchers developed the Kismet, which was capable of social interaction and demonstrating human emotions [7]. The 2010's saw a wave of daily AI interaction in the form of personal assistants, such as Apple's iPhone 4s released with Siri, allowing the user to communicate with their smartphone with vocal commands to call, text, create reminders and more [7]. Amazon would go on to compete with Siri, creating the Amazon Echo with its Alexa AI [7].

IV. Optimizing Modern Artificial Intelligence

The primary issue underlying typical computer systems is their lack of optimization for the operations performed by artificial intelligence algorithms [8]. This is known as the von Neumann bottleneck, which limits the speed of artificial intelligence as a result of limited data throughput speed between the CPU and memory [8]. To combat this, engineers have begun devising engines and computational methods with the sole purpose of optimizing artificial intelligence computing. These include the Transformer architecture, and AMD's XDNA Dynamic AI Architecture, which debuted at the Consumer Electronics Show (CES) in 2023. Both of these are specifically optimized for their own specific type of artificial intelligence procedure, allowing for improved efficiency of AI calculations.

Apple's Neural Engine sets out to optimize the Transformer architecture in Apple devices, initially developed when Apple noticed a trend of their machine learning models using this architecture [9]. The Transformer architecture is "the first transduction model relying entirely on self-attention to compute representations of its input and output without using sequence-aligned [recurrent neural networks] or convolution" [9].

Attention is the "mapping of a query and set of key-value pairs to an output" [9]. Essentially, the data takes in a matrix of queries, keys, and values, and performs linear operations (mapping) to encode the data, then performs operations that then decodes it. An encoder layer is a matrix that maps an input, and a decoder layer is a matrix that "unmaps" an input. The Transformer architecture uses attention in "encoder-decoder attention" layers, where queries come from a previous decoder layer and keys and values come from an output of the encoder,

Table 1. "Maximum path lengths, per-layer complexity and minimum number of sequential operations for different layer types. n is the sequence length, d is the representation dimension, k is the kernel size of convolutions and r is the size of the neighborhood in restricted self-attention."

Layer Type	Complexity per Layer	Sequential Operations	Maximum Path Length
Self-Attention	$O(n^2 \cdot d)$	$O(1)$	$O(1)$
Recurrent	$O(n \cdot d^2)$	$O(n)$	$O(n)$
Convolutional	$O(k \cdot n \cdot d^2)$	$O(1)$	$O(log_k(n))$
Self-Attention (restricted)	$O(r \cdot n \cdot d)$	$O(1)$	$O(n/r)$

allowing for every position in the decoder to view all positions of an input sequence [9]. Additionally, the encoder contains self-attention layers, where all keys, values, and queries come from the same place, the output of the previous encoder layer, allowing for each position in the encoder to attend to all positions of the previous encoder layer [9]. Finally, self-attention layers in the decoder let each position access all positions up to and including that position," prohibiting information from flowing leftward in the system [9]. Table 1 shows the algorithmic advantages of self-attention over recurrent neural network and convolution artificial intelligence models, primarily through the big O notation. The exponential orders in the first 3 rows of "Complexity per Layer" means those types become exponentially less efficient as the sequence length increases. The lack of an exponential order in restricted self-attention conveys its significantly better efficiency over the other three artificial intelligence architectures.

The Transformer architecture utilizes a complex understanding of many fields of mathematics, such as linear algebra and functional analysis to encode and decode a set of information, allowing simplification of computational processes in its implementation and avoiding complex calculations, such as convolution. Apple's Neural Engine optimizes machine learning(ML) workloads by transferring non-ML workloads to the central processing unit and graphics processing unit [9]. Additionally, a set of four principles are detailed by Apple to allow developers to further optimize transformers for use in the Neural Engine [8]. These optimizations allows engineers to maximize the power of artificial intelligence, while keeping it time and energy efficient.

A primary limitation in artificial intelligence research is seen by its difficulty to quantify its performance. As artificial intelligence grows, measurements on the efficiency and performance of an AI are imperative [7]. Furthermore, there is not standard way to determine whether an AI is making progress, with varying means of applying these [7]. Going forward, it is important to determine standardized ways of measuring the efficacy and progress of artificial intelligence systems, to allow for

V. The Future of Artificial Intelligence

A. Significance of the History of Computers and Artificial Intelligence

The history of computing and artificial intelligence, having been laid out above, will direct our discussion of the future of the field of artificial intelligence. By understanding the history of computing and artificial intelligence, we can begin to understand different ways this field can improve, and how it has already been strengthened in the past.

In order to discuss the future, we must also define the different classification of artificial intelligence based on their capabilities. These include narrow AI, which is capable of intelligently performing a specific task, general AI, machines that mimic human intelligence, and super AI, or self-aware AI with intelligence greater than humans, and are typically able to perform any task possible by humans [10].

B. The Near Future of Artificial Intelligence

The promising future AI applications within the next few years spans many fields not entirely related to computing, including healthcare, education, and transportation, and will lead to a dramatic shift in the way work is done in professional and personal settings.

The future of artificial intelligence in the medical field will allow a healthier global population. For example, in India, a large majority of citizen do not have access to quality healthcare, a combination of a lack of qualified doctors, and undeveloped infrastructure [10]. Although many citizens can't go to the doctor, artificial intelligence can diagnose diseased based on symptoms, and wearing fitness monitors [10]. These AI's can also recommend different medications the user can take to mitigate these symptoms [10].

The development of drug property and activity in humans. Machine learning techniques trials are currently underway, developing ways to access how a human absorbs, distributes, metabolizes, and excretes drugs taken [11]. Figure 2 shows the future path this technology hopes to follow.

In education, widespread worry about ChatGPT and its harm on the learning of young students simply using it to get an assignment over with exists. While this is a valid concern, there are many positive and practical applications of other forms of artificial intelligence on education. With artificial intelligence's ability to self-learn, an education system devised that is tailored for a specific students abilities and learning style has a wellspring of positive benefits for a learning population [10].

C. Long-Term Goals of Artificial Intelligence

While the field of artificial intelligence is advancing, the mind begins to wonder what long term possibilities exist for this revolutionary technology. While many ideas may be rooted in science fiction, many long-term applications of artificial intelligence begin to be imagined.

A notable application, driverless cars on the main roads, are becoming a closer reality each and every day [6]. Many cars are

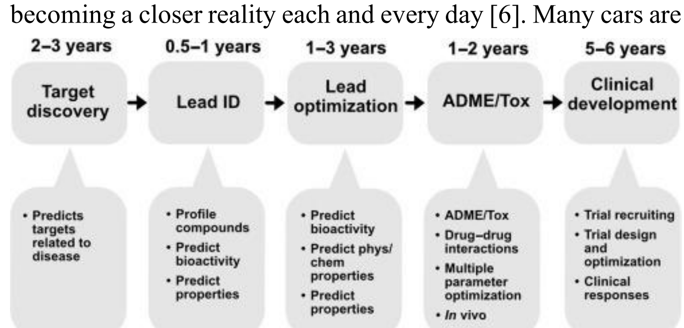

Figure 2[11]. Machine learning prospects of small molecule drug discovery already equipped with some self-driving applications, however if all cars were self-driving, and able to communicate with each other in some meaningful way, a world can be imagined where there are no accidents and reliable commuting from point A to point B will be worry-free.

VI. Responsible Expansion of Artificial Intelligence

As artificial intelligence begins to invade most aspects of everyday life, it is important to ensure this expansion occurs in a responsible and ethical manner. From students using ChatGPT to complete their assignments, biased AI systems that filter out job applicants, and personal assistance AI such as Amazon's Alexa with the potential to invade home privacy, the number of irresponsible uses for artificial intelligence is impossible to quantify. Therefore, global cooperation on the ethical incorporation of AI into our daily lives is imperative going forward with AI expansion.

Many strategies can and should be put in place to ensure that the widespread and increased implementation of artificial intelligence remains ethical, including ethics workshops, risk assessment, and bias analysis. While it is never possible to remove all the dangers coming with an innovating world, we must ensure that these precautions are put in place to maintain a fair and just society.

References

[1] "Responsible and Ethical AI: A Comprehensive Guide by Xomnia." *Xomnia*, www.xomnia.com/responsible-and-ethical-ai-comprehensive-guide/. Accessed 4 Sept. 2023.

[2] Rotman, David. "We're Not Prepared for the End of Moore's Law." *MIT Technology Review*, MIT Technology Review, 21 Apr. 2021, www.technologyreview.com/2020/02/24/905789/were-not-prepared-for-the-end-of-moores-law/.

[3] Sutter, Herb. *The Free Lunch Is over: A Fundamental Turn toward Concurrency in Software*, Dec. 2004, www.gotw.ca/publications/concurrency-ddj.htm.

[4] Hruska, Joel. "The Death of CPU Scaling: From One Core to Many -- and Why We're Still Stuck." ExtremeTech, 1 Feb. 2012, www.extremetech.com/computing/116561-the-death-of-cpu-scaling-from-one-core-to-many-and-why-were-still-stuck.

[5] Haenlein, Michael, and Adreas Kaplan. "A Brief History of Artificial Intelligence: On the Past, Present, And ..." Research Gate, July 2019, www.researchgate.net/publication/334539401_A_Brief_History_of_Artificial_Intelligence_On_the_Past_Present_and_Future_of_Artificial_Intelligence.

[6] Anyoha, Rockwell. "The History of Artificial Intelligence." Science in the News, 28 Aug. 2017, sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/.

[7] Peterson, Amanda. "How AI Has Advanced during the 21st Century and Where It's Headed." *Pro*, 4 May 2023, www.pro-sapien.com/blog/how-ai-has-advanced-during-21st-century-and-where-its-headed/.

[8] G, Alok. Architecture Apocalypse Dream Architecture for Deep Learning Inference and Compute -VERSAL AI Core, www.xilinx.com/content/dam/xilinx/support/documents/white_papers/EW2020-Deep-Learning-Inference-AICore.pdf.

[9] "Deploying Transformers on the Apple Neural Engine." *Apple Machine Learning Research*, machinelearning.apple.com/research/neural-engine-transformers. Accessed 5 Sept. 2023.

[10] Duggal, Nikita. "Future of AI (Artificial Intelligence): What Lies Ahead?" Simplilearn.Com, Simplilearn, 21 Aug. 2023, www.simplilearn.com/future-of-artificial-intelligence-article#:~:text=leap%20of%20faith.-,Future%20of%20Artificial%20Intelligence,healthcare%2C%20banking%2C%20and%20transportation.

[11] Bohr, Adam, and Kaveh Memarzadeh. "The Rise of Artificial Intelligence in Healthcare Applications." Artificial Intelligence in Healthcare, U.S. National Library of Medicine, 2020, www.ncbi.nlm.nih.gov/pmc/articles/PMC7325854/.