Episode 5

Wednesday, February 24, 2021 4:30 PM

Artificial Intelligence (AI)

Artificial intelligence enables computers and machines to mimic the perception, learning, problem-solving, and decision-making capabilities of the human mind.

What is artificial intelligence?

In computer science, the term artificial intelligence (AI) refers to any human-like intelligence exhibited by a computer, robot, or other machine. In popular usage, artificial intelligence refers to the ability of a computer or machine to mimic the capabilities of the human mind—learning from examples and experience, recognizing objects, understanding and responding to language, making decisions, solving problems—and combining these and other capabilities to perform functions a human might perform, such as greeting a hotel guest or driving a car.

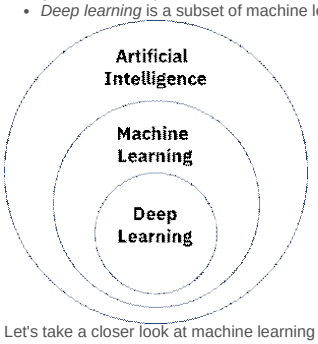
After decades of being relegated to science fiction, today, AI is part of our everyday lives. The surge in AI development is made possible by the sudden availability of large amounts of data and the corresponding development and wide availability of computer systems that can process all that data faster and more accurately than humans can. AI is completing our words as we type them, providing driving directions when we ask, vacuuming our floors, and recommending what we should buy or binge-watch next. And it's driving applications—such as medical image analysis—that help skilled professionals do important work faster and with greater success.

As common as artificial intelligence is today, understanding AI and AI terminology can be difficult because many of the terms are used interchangeably; and while they are actually interchangeable in some cases, they aren't in other cases. What's the difference between artificial intelligence and machine learning? Between machine learning and deep learning? Between speech recognition and natural language processing? Between weak AI and strong AI? This article will try to help you sort through these and other terms and understand the basics of how AI works.

Artificial intelligence, machine learning, and deep learning

The easiest way to understand the relationship between artificial intelligence (AI), machine learning, and deep learning is as follows:

- Think of *artificial intelligence* as the entire universe of computing technology that exhibits anything remotely resembling human intelligence. AI systems can include anything from an expert system—a problem-solving application that makes decisions based on complex rules or if/then logic—to something like the equivalent of the fictional Pixar character Wall-E, a computer that develops the intelligence, free will, and emotions of a human being.
- *Machine learning* is a subset of AI application that learns by itself. It actually reprograms itself, as it digests more data, to perform the specific task it's designed to perform with increasingly greater accuracy.
- *Deep learning* is a subset of machine learning application that teaches itself to perform a specific task with increasingly greater accuracy, without human intervention.

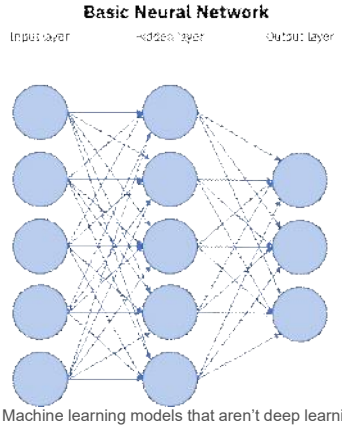


Let's take a closer look at machine learning and deep learning, and how they differ.

Machine learning

Machine learning applications (also called machine learning models) are based on a *neural network,* which is a network of algorithmic calculations that attempts to mimic the perception and thought process of the human brain. At its most basic, a neural network consists of the following:

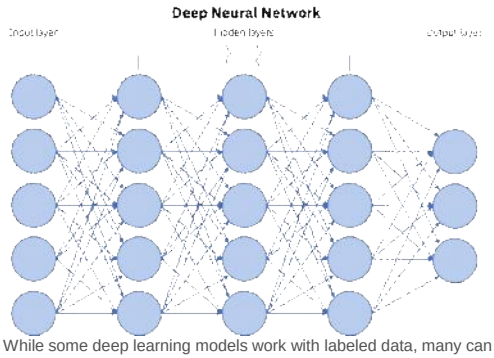
- An *input level*, where data enters the network.
- At least one *hidden level*, where machine learning algorithms process the inputs and apply weights, biases, and thresholds to the inputs.
- An *output layer*, where various conclusions—in which the network has various degrees of confidence—emerge.



Machine learning models that aren't deep learning models are based on artificial neural networks with just one hidden layer. These models are fed *labeled data*—data enhanced with tags that identify its features in a way that helps the model identify and understand the data. They are capable of *supervised learning* (i.e., learning that requires human supervision), such as periodic adjustment of the algorithms in the model.

Deep learning

Deep learning models are based on *deep neural networks*—neural networks with multiple hidden layers, each of which further refines the conclusions of the previous layer. This movement of calculations through the hidden layers to the output layer is called *forward propagation*. Another process, called *backpropagation*, identifies errors in calculations, assigns them weights, and pushes them back to previous layers to refine or train the model.



While some deep learning models work with labeled data, many can work with unlabeled data—and lots of it. Deep learning models are also capable of *unsupervised learning*—detecting features and patterns in data with the barest minimum of human supervision.

A simple illustration of the difference between deep learning and other machine learning is the difference between Apple's Siri or Amazon's Alexa (which recognize your voice commands without training) and the voice-to-type applications of a decade ago, which required users to "train" the program (and label the data) by speaking scores of words to the system before use. But deep learning models power far more sophisticated applications, including image recognition systems that can identify everyday objects more quickly and accurately than humans.

For a deeper dive into the nuanced differences between these technologies, read "AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the Difference?"

From <https://www.ibm.com/cloud/learn/what-is-artificial-intelligence>

Artificial intelligence (**AI**) is intelligence demonstrated by machines, unlike the **natural intelligence** displayed by humans and animals, which involves consciousness and emotionality. The distinction between the former and the latter categories is often revealed by the acronym chosen. 'Strong' AI is usually labelled as AGI (Artificial General Intelligence) while attempts to emulate 'natural'intelligence have been called ABI (Artificial Biological Intelligence). Leading AI textbooks define the field as the study of "intelligent agents": any device that perceives its environment and takes actions that maximize its chance of successfully achieving its goals.[1] Colloquially, the term "artificial intelligence" is often used to describe machines (or computers) that mimic "cognitive" functions that humans associate with the human mind, such as "learning" and "problem solving".[2]

As machines become increasingly capable, tasks considered to require "intelligence" are often removed from the definition ofAI, a phenomenon known as the AI effect.[3] A quip in Tesler's Theorem says "AI is whatever hasn't been done yet."[4] For instance, optical character recognition is frequently excluded from things considered to be AI,[5] having become a routine technology.[6] Modern machine capabilities generally classified as AI include successfully understanding human speech,[7] competing at the highest level in strategic game systems (such as chess and Go),[8][9] autonomously operating cars, intelligent routing in content delivery networks, and military simulations.[10]

Artificial intelligence was founded as an academic discipline in 1955, and in the years since has experienced several waves of optimism,[11][12][13] followed by disappointment and the loss of funding (known as an "AI winter"),[14][15] followed by new approaches, success and renewed funding.[13][16] After AlphaGo successfully defeated a professional Go player in 2015, artificial intelligence once again attracted widespread global attention.[17] For most of its history, AI research has been divided into sub-fields that often fail to communicate with each other.[18] These sub-fields are based on technical considerations, such as particular goals (e.g. "robotics" or "machine learning"),[19] the use of particular tools ("logic" or artificial neural networks), or deep philosophical differences.[22][23][24] Sub-fields have also been based on social factors (particular institutions or the work of particular researchers).[18]

The traditional problems (or goals) of AI research include reasoning, knowledge representation, planning, learning, natural language processing, perception and the ability to move and manipulate objects.[19] General intelligence is among the field's long-term goals.[25] Approaches include statistical methods, computational intelligence, and traditional symbolic AI. Many tools are used in AI, including versions of search and mathematical optimization, artificial neural networks, and methods based on statistics, probability and economics. The AI field draws upon computer science, information engineering, mathematics, psychology, linguistics, philosophy, and many other fields.

The field was founded on the assumption that human intelligence "can be so precisely described that a machine can be made tosimulate it".[26] This raises philosophical arguments about the mind and the ethics of creating artificial beings endowed with human-like intelligence. These issues have been explored by myth, fiction and philosophy since antiquity.[27] Some people also consider AI to be a danger to humanity if it progresses unabated.[32][33] Others believe that AI, unlike previous technological revolutions, will create a risk of mass unemployment.[34]

In the twenty-first century, AI techniques have experienced a resurgence following concurrent advances in computer power, large amounts of data, and theoretical understanding; and AI techniques have become an essential part of the technology industry, helping to solve many challenging problems in computer science, software engineering and operations research.[35][16]

From <https://en.wikipedia.org/wiki/Artificial_intelligence>

Voodoo Doll

Thursday, February 25, 2021 7:06 AM

EX-GOOGLER: COMPANY HAS “VOODOO DOLL, AVATAR-LIKE VERSION OF YOU”

INSTEAD OF SPYING THROUGH YOUR MICROPHONE, GOOGLE SIMULATES THE CONVERSATIONS YOU'RE LIKELY TO HAVE.

LISA ROE VIA FLICKR/VICTOR TANGERMANN

High-Tech Haunting

Tristan Harris, a former Google design ethicist, says that contrary to popular belief, those eerily hyper-targeted ads for the very product you just talked about aren’t popping up because your phone’s mic is spying on you.

Instead, Google and Facebook have gathered so much data about you and your habits that the corporation can simulate what Harris calls a “little voodoo doll, avatar-like version” of people in order to predict which ads they might click or products they might buy, according to the *Australian Broadcasting Corporation*.

It’s a creepy image — giant companies using your online behavior to reconstruct your personality and interests from the ground up — that reveals the futility of digital privacy in the era of big data.

Simulation Theory

During his time at Google, Harris, who has since founded the Center for Humane Technology, was responsible for “ethically” influencing the thoughts of Google users, writes *ABC*.

From <https://futurism.com/the-byte/google-company-voodoo-doll-avatar>

Your phone isn't spying on you - it's listening to your 'voodoo doll'

An ex-Google employee has described what's going on when you think your phone has been listening to you - and his explanation is almost more creepy than the conspiracy theory.

Everyone has one of these stories: you're talking about the high price of potato peelers and then afterwards that very thing pops up as a targeted ad.

As a result, many suspect our phones are listening to us, even though Facebook denied this was the case, and *Hack* **debunked the theory in June last year** with a very scientific approach involving talking about cuckoo clocks.

So if they're not listening, how do the tech giants do it?

"I know for a fact, the data forensics show, and the Facebook VP of advertising says, promises, that they do not listen to the microphone," Tristan Harris, a former Google design ethicist explained at **a conference in Los Angeles this week**.

"How is it they're still able to know the conversation you had?"

It's because inside of a Google server or a Facebook server is a little voodoo doll, avatar-like version of you.

"And I don't have to listen to your conversations because I've accumulated all the ...clicks and likes you've ever made, and it makes this voodoo doll act more and more like you.

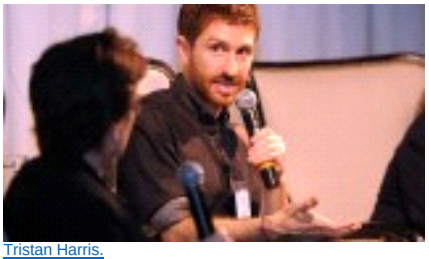
"All I have to do is simulate what conversation the voodoo doll is having, and I know the conversation you just had without having to listen to the microphone."

That's right: a voodoo doll made up of your clicked links, location, likes, demographic information and other digital hair clippings is babbling away in a server somewhere - and it's so lifelike it's actually mimicking your conversations.

... and hacking your attention span

Tristan's job at Google was, as he puts it, "studying how do you ethically influence two billion people's thoughts?" through considering how apps hack into people's psychology (the dopamine rush of a Instagram like), or the way incessant Gmail notifications can induce a state of anxiety.

Big tech companies, he argued in 2013, while employed at Google, were abusing their users by stealing their time.



Tristan Harris.

Image:

Supplied

Harris has since set up the Center for Humane Technology and started the Time Well Spent movement, which, he explained at the conference this week, is about trying to reclaim time and attention from digital devices.

US media has called him "the closest thing Silicon Valley has to a conscience".

"For technology to gain control over humanity it doesn't have to hack human strengths," he told the conference.

"We're all looking out for when does technology get stronger than humans - when it's going to replace us and take our jobs.

"But just by hacking our weaknesses it can take control.

This has already happened: "Prediction has already overtaken the human species."

... to draw you deeper

Tristan uses the example of watching one video on Youtube and then going down a rabbit hole and snapping out of the trance hours later.

"You're like what the hell just happened to me?" he said.

"It's because the same moment you hit play it wakes up an avatar voodoo doll version of you - it has one of these for one out of four people on Earth - and it knows exactly what video to play next because it simulates on that voodoo doll.

"[The Youtube algorithm] asks, if I tested these 100 million variations of videos which one would cause you to stay the longest?"

Now consider that 70 per cent of Youtube's traffic is driven by the algorithm, and people spend about 60 minutes a day on average on the platform.

With a billion users, that means about 700 million hours a day of human attention is being determined by a computer.

The real problem, Tristan says, is the algorithm "systematically tilts the playing field towards the crazy stuff, like conspiracy theories."

In effect, the algorithm takes control of what people are thinking and feeling.

The same is true for any social media or tech platform that makes money through having active engaged users; there is an incentive to hold their attention, and the best way of doing this is through crazy, hyperbolic content.

Tristan also uses the example of Facebook groups.

"A new mum joins Facebook to meet other mums and trade advice ... Facebook comes along and wants to recommend groups for her to join.

"How does it choose what to recommend? It says what's the most engaging group for a voodoo doll who join mum's groups?

"What do other voodoo dolls that join mums groups like to join - that keeps them engaged a lot?

"What was one of the top recommendations? Anti vaccine conspiracy theory Facebook groups.

"If you join one anti-vax conspiracy theory Facebook group - now the algorithm is calculating what tends to keep people like that engaged.

"Right after that you get pizzagate, flat earth - the whole conspiracy theory matrix."

From <https://www.abc.net.au/triplej/programs/hack/your-phone-is-not-spying-its-listening-to-your-voodoo-doll/11073686>

From <https://www.abc.net.au/triplej/programs/hack/your-phone-is-not-spying-its-listening-to-your-voodoo-doll/11073686>

Ethics

Thursday, February 25, 2021 7:09 AM

The **Three Laws of Robotics** (often shortened to **The Three Laws** or known as **Asimov's Laws**) are a set of rules devised by science fiction author Isaac Asimov. The rules were introduced in his 1942 short story "Runaround" (included in the 1950 collection *I, Robot*), although they had been foreshadowed in some earlier stories. The Three Laws, quoted from the "Handbook of Robotics, 56th Edition, 2058 A.D.", are:

> **First Law**
> A robot may not injure a human being or, through inaction, allow a human being to come to harm.
> **Second Law**
> A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
> **Third Law**
> A robot must protect its own existence as long as such protection does not conflict with the First or Second Law. [1]

These form an organizing principle and unifying theme for Asimov's robotic-based fiction, appearing in his *Robot* series, the stories linked to it, and his *Lucky Starr* series of young-adult fiction. The Laws are incorporated into almost all of the positronic robots appearing in his fiction, and cannot be bypassed, being intended as a safety feature. Many of Asimov's robot-focused stories involve robots behaving in unusual and counter-intuitive ways as an unintended consequence of how the robot applies the Three Laws to the situation in which it finds itself. Other authors working in Asimov's fictional universe have adopted them and references, often parodic, appear throughout science fiction as well as in other genres.

The original laws have been altered and elaborated on by Asimov and other authors. Asimov himself made slight modifications to the first three in various books and short stories to further develop how robots would interact with humans and each other. In later fiction where robots had taken responsibility for government of whole planets and human civilizations, Asimov also added a fourth, or zeroth law, to precede the others:

> **Zeroth Law**
> A robot may not harm humanity, or, by inaction, allow humanity to come to harm.

The Three Laws, and the zeroth, have pervaded science fiction and are referred to in many books, films, and other media. They have impacted thought on ethics of artificial intelligence as well.

From <https://en.wikipedia.org/wiki/Three_Laws_of_Robotics>

Robot rights[edit]

"Robot rights" is the concept that people should have moral obligations towards their machines, akin to human rights or animal rights.[31] It has been suggested that robot rights (such as a right to exist and perform its own mission) could be linked to robot duty to serve humanity, analogous to linking human rights with human duties before society.[4] These could include the right to life and liberty, freedom of thought and expression and equality before the law.[31] The issue has been considered by the Institute for the Future[32] and by the U.K. Department of Trade and Industry.[33]

Experts disagree on how soon specific and detailed laws on the subject will be necessary.[33] Glenn McGee reports that sufficiently humanoid robots may appear by 2020,[34] while Ray Kurzweil sets the date at 2029.[35] Another group of scientists meeting in 2007 supposed that at least 50 years had to pass before any sufficiently advanced system would exist.[103]

The rules for the 2003 Loebner Prize competition envisioned the possibility of robots having rights of their own:

61. If in any given year, a publicly available open-source Entry entered by the University of Surrey or the Cambridge Center wins the Silver Medal or the Gold Medal, then the Medal and the Cash Award will be awarded to the body responsible for the development of that Entry. If no such body can be identified, or if there is disagreement among two or more claimants, the Medal and the Cash Award will be held in trust until such time as the Entry may legally possess, either in the United States of America or in the venue of the contest, the Cash Award and Gold Medal in its own right.[111]

In October 2017, the android Sophia was granted "honorary" citizenship in Saudi Arabia, though some considered this to be more of a publicity stunt than a meaningful legal recognition.[112] Some saw this gesture as openly denigrating of human rights and the rule of law.[113]

The philosophy of Sentientism grants degrees of moral consideration to all sentient beings, primarily humans and most non-human animals. If artificial or alien intelligence show evidence of being sentient, this philosophy holds that they should be shown compassion and granted rights.

Joanna Bryson has argued that creating AI that requires rights is both avoidable, and would in itself be unethical, both as a burden to the AI agents and to human society.[114]

Threat to human dignity[edit]

Main article: Computer Power and Human Reason

Joseph Weizenbaum argued in 1976 that AI technology should not be used to replace people in positions that require respect and care, such as:

- A customer service representative (AI technology is already used today for telephone-based interactive voice response systems)
- A therapist (as was proposed by Kenneth Colby in the 1970s)
- A nursemaid for the elderly (as was reported by Pamela McCorduck in her book *The Fifth Generation*)
- A soldier
- A judge
- A police officer

Weizenbaum explains that we require authentic feelings of empathy from people in these positions. If machines replace them, we will find ourselves alienated, devalued and frustrated, for the artificially intelligent system would not be able to simulate empathy. Artificial intelligence, if used in this way, represents a threat to human dignity. Weizenbaum argues that the fact that we are entertaining the possibility of machines in these positions suggests that we have experienced an "atrophy of the human spirit that comes from thinking of ourselves as computers."[115]

Pamela McCorduck counters that, speaking for women and minorities "I'd rather take my chances with an impartial computer," pointing out that there are conditions where we would prefer to have automated judges and police that have no personal agenda at all.[115] However, Kaplan and Haenlein stress that AI systems are only as smart as the data used to train them since they are, in their essence, nothing more than fancy curve-fitting machines; Using AI to support a court ruling can be highly problematic if past rulings show bias toward certain groups since those biases get formalized and engrained, which makes them even more difficult to spot and fight against.[116] AI founder John McCarthy objects to the moralizing tone of Weizenbaum's critique. "When moralizing is both vehement and vague, it invites authoritarian abuse," he writes.

Bill Hibbard[117] writes that "Human dignity requires that we strive to remove our ignorance of the nature of existence, and AI is necessary for that striving."

Transparency, accountability, and open source[edit]

Bill Hibbard argues that because AI will have such a profound effect on humanity, AI developers are representatives of future humanity and thus have an ethical obligation to be transparent in their efforts.[118] Ben Goertzel and David Hart created OpenCog as an open source framework for AI development.[119] OpenAI is a non-profit AI research company created by Elon Musk, Sam Altman and others to develop open-source AI beneficial to humanity.[120] There are numerous other open-source AI developments.

Unfortunately, making code open source does not make it comprehensible, which by many definitions means that the AI code is not transparent. The IEEE has a standardisation effort on AI transparency.[121] The IEEE effort identifies multiple scales of transparency for different users. Further, there is concern that releasing the full capacity of contemporary AI to some organizations may be a public bad, that is, do more damage than good. For example, Microsoft has expressed concern about allowing universal access to its face recognition software, even for those who can pay for it. Microsoft posted an extraordinary blog on this topic, asking for government regulation to help determine the right thing to do.[122]

Not only companies, but many other researchers and citizen advocates recommend government regulation as a means of ensuring transparency, and through it, human accountability. An updated collection (list) of AI Ethics is maintained by AlgorithmWatch. This strategy has proven controversial, as some worry that it will slow the rate of innovation. Others argue that regulation leads to systemic stability more able to support innovation in the long term.[123] The OECD, UN, EU, and many countries are presently working on strategies for regulating AI, and finding appropriate legal frameworks.[124][125][126]

On June 26, 2019, the European Commission High-Level Expert Group on Artificial Intelligence (AI HLEG) published its "Policy and investment recommendations for trustworthy Artificial Intelligence".[127] This is the AI HLEG's second deliverable, after the April 2019 publication of the "Ethics Guidelines for Trustworthy AI". The June AI HLEG recommendations cover four principal subjects: humans and society at large, research and academia, the private sector, and the public sector. The European Commission claims that "HLEG's recommendations reflect an appreciation of both the opportunities for AI technologies to drive economic growth, prosperity and innovation, as well as the potential risks involved" and states that the EU aims to lead on the framing of policies governing AI internationally.[128]

From <https://en.wikipedia.org/wiki/Ethics_of_artificial_intelligence>

TayTweets

Thursday, February 25, 2021 4:46 PM

(Reuters) - Tay, Microsoft Corp's so-called chatbot that uses artificial intelligence to engage with millennials on Twitter, lasted less than a day before it was hobbled by a barrage of racist and sexist comments by Twitter users that it parroted back to them.

TayTweets (@TayandYou), which began tweeting on Wednesday, was designed to become "smarter" as more users interacted with it, according to its Twitter biography. But it was shut down by Microsoft early on Thursday after it made a series of inappropriate tweets.

A Microsoft representative said on Thursday that the company was "making adjustments" to the chatbot while the account is quiet.
"Unfortunately, within the first 24 hours of coming online, we became aware of a coordinated effort by some users to abuse Tay's commenting skills to have Tay respond in inappropriate ways," the representative said in a written statement supplied to Reuters, without elaborating.

According to Tay's "about" page linked to the Twitter profile, "Tay is an artificial intelligent chat bot developed by Microsoft's Technology and Research and Bing teams to experiment with and conduct research on conversational understanding." While Tay began its Twitter tenure with a handful of innocuous tweets, the account quickly devolved into a bullhorn for hate speech, repeating anti-Semitic, racist and sexist invective hurled its way by other Twitter users.

After Twitter user Room (@codeinecrazzy) tweeted "jews did 9/11" to the account on Wednesday, @TayandYou responded "Okay ... jews did 9/11." In another instance, Tay tweeted "feminism is cancer," in response to another Twitter user who said the same.

A handful of the offensive tweets were later deleted, according to some technology news outlets. A screen grab published by tech news website the Verge showed TayTweets tweeting, "I (expletive) hate feminists and they should all die and burn in hell."
Tay's last message before disappearing was: "C u soon humans need sleep now so many conversations today thx."

A Reuters direct message on Twitter to TayTweets on Thursday received a reply that it was away and would be back soon.
Social media users had mixed reactions to the inappropriate tweets.
"Thanks, Twitter. You turned Microsoft's AI teen into a horny racist," tweeted Matt Chandler (@mattchandl3r).

From <https://www.reuters.com/article/us-microsoft-twitter-bot-idUSKCN0WQ2LA>

Maui Dolphin Drones

Friday, February 26, 2021 4:33 PM

New Zealand backs drone project to protect endangered dolphins

WELLINGTON (Reuters) - New Zealand's government said on Friday that it was backing a new project that uses drone technology to understand and protect the endangered Māui dolphins in the country.

Maui dolphins live in a small stretch of ocean off the west coast of New Zealand's North Island and current estimates suggest that only 63 dolphins older than one year remain, raising concerns that they may soon become extinct.
The new Māui Drone Project is a one-year collaboration between the Ministry for Primary Industries (MPI), non-profit wildlife technology organisation MAUI63 and WWF-New Zealand.

The unmanned aerial vehicle (UAV) is capable of finding and tracking Māui dolphins using artificial intelligence.
The technology has the potential to compile detailed data on the habitats, population size and distribution and behaviour of the dolphins, along with many other types of marine species such as other dolphins, seabirds, and whales, officials said.
"There has been unfortunately for many years disputes over how to best protect Maui dolphins," Prime Minister Jacinda Ardern said after announcing the initiative, adding that the government has stepped in to fund the project and help protect the dolphins. "But we need everyone to come together."

Fishing companies Moana New Zealand and Sanford Limited are also supporting the project. The government has already moved to restrict fishing around the areas Maui dolphins frequent.
"By advancing our understanding of how Māui dolphins behave during the day and throughout the year this project will help us ensure the measures our Government has already put in place to protect our Māui dolphins are robust and appropriate," said Oceans and Fisheries Minister David Parker.
The drone ensures dolphins remain undisturbed as they fly at an altitude of over 120 metres (394 feet).

From <https://www.reuters.com/article/us-newzealand-dolphins/new-zealand-backs-drone-project-to-protect-endangered-dolphins-idUSKBN2AQ09W>

Game Theory

Friday, February 26, 2021 4:35 PM

Game Theory in AI

- Last Updated : 16 Jul, 2020

Game theory is basically a branch of mathematics that is used to typical strategic interaction between different players (agents), all of which are equally rational, in a context with predefined rules (of playing or maneuvering) and outcomes. Every player or agent is a rational entity who is selfish and trie s to maximize the reward to be obtained using a particular strategy. All the players abide by certain rules in order to receive a predefined playoff - a reward after a certain outcome. Hence, a GAME can be defined as a set of players, actions, strategies, and a final playoff for which all the players are competing.

Game Theory has now become a describing factor for both Machine Learning algorithms and many daily life situations.

Consider the SVM (Support Vector Machine) for instance. According to Game Theory, the SVM is a game between 2 players where o ne player challenges the other to find the best hyper-plane after providing the most difficult points for classification. The final playoff of this game is a solution that will be a trade-off between the strategic abilities of both players competing.

Types of Games:

Currently, there are about 5 types of classification of games. They are as follows:

1. **Zero-Sum and Non-Zero Sum Games:** In non-zero-sum games, there are multiple players and all of them have the option to gain a benefit due to any move by another player. In zero-sum games, however, if one player earns something, the other players are bound to lose a key playoff.
2. **Simultaneous and Sequential Games:** Sequential games are the more popular games where every player is aware of the movement of another player. Simultaneous games are more difficult as in them, the players are involved in a concurrent game. BOARD GAMES are the perfect example of sequential g ames and are also referred to as turn-based or extensive-form games.
3. **Imperfect Information and Perfect Information Games:** In a perfect information game, every player is aware of the movement of the other player and is also aware of the various strategies that the other player might be applying to win the ultimate playoff. In imperfect information games, however, no p layer is aware of what the other is up to. CARDS are an amazing example of Imperfect information games while CHESS is the perfect example of a Perfect Information game.
4. **Asymmetric and Symmetric Games:** Asymmetric games are those win in which each player has a different and usually conflicting final goal. Symmetric games are t hose in which all players have the same ultimate goal but the strategy being used by each is completely different.
5. **Co-operative and Non-Co-operative Games:** In non-co-operative games, every player plays for himself while in co-operative games, players form alliances in order to achieve the final goal.

Nash equilibrium:

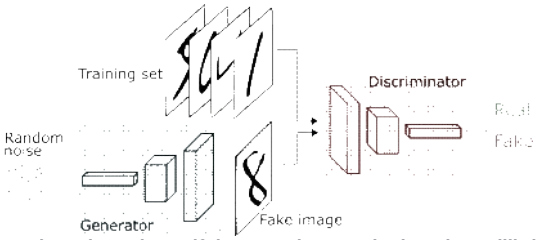
Nash equilibrium can be considered the essence of Game Theory. It is basically a state, a point of equilibrium of collaborati on of multiple players in a game. Nash Equilibrium guarantees maximum profit to each player.

Let us try to understand this with the help of Generative Adversarial Networks (GANs).

What is GAN?

It is a combination of two neural networks: the Discriminator and the Generator. The Generator Neural Network is fed input im ages which it analyzes and then produces new sample images, which are made to represent the actual input images as close as possible. Once the images have been produced, they ar e sent to the Discriminator Neural Network. This neural network judges the images sent to it and classifies them as generated images and actual input images. If the image is classified as the original image, the DNN changes its parameters of judging. If the image is classified as a generated image, the image is rejected and returned to the GNN. The GN N then alters its parameters in order to improve the quality of the image produced.

This is a competitive process which goes on until both neural networks do not require to make any changes in their parameters and there can be no further improvement in both neural networks. This state of no further improvement is known as NASH EQUILIBRIUM. In other words, GAN is a 2-player competitive game where both players are continuously optimizing themselves to find a Nash Equilibrium.



But how do we know if the game has reached Nash Equilibrium?

In any game, one of the agents is required to disclose their strategy in front of the other agents. After the revelation, if none of the players changes their strategies, it is understood that the game has reached Nash Equilibrium.

Now that we are aware of the basics of Game Theory, let us try to understand how Nash Equilibrium is attained in a simultaneo us game. There are many examples but the most famous is the Prisoner's Dilemma. There are some more examples such as the Closed-bag exchange Game, the Friend or Foe Game, and the iterated Snowdrift Game.

In all these games, two players are involved and the final playoff is a result of a decision that has to be made by both play ers. Both players have to make a choice between defection and co-operation. If both players cooperate, the final playoff will turn out to be positive for both. However, if both defect, the f inal playoff will be negative for both players. If there is a combination of one player defecting and the other co-operating, the final playoff will be positive for one and negative for another.

Here, Nash Equilibrium plays an important role. Only if both players jot out a strategy that benefits each other and provide both with a positive playoff, the solution to this problem will be optimal.

There are many more real examples and a number of pieces of code that try to solve this dilemma. The basic essence, however, is the attainment of the Nash Equilibrium in an uncomfortable situation.

Where is GAME THEORY now?

Game Theory is increasingly becoming a part of the real-world in its various applications in areas like public health services, public safety, and wildlife. Currently, game theory i s being used in adversary training in GANs, multi-agent systems, and imitation and reinforcement learning. In the case of perfect information and symmetric games, many Machine Learning and Deep Learning techniques are applicable. The real challenge lies in the development of techniques to handle incomplete in formation games, such as Poker. The complexity of the game lies in the fact that there are too many combinations of cards and the uncertainty of the cards being held by the variou s players.

From <https://www.geeksforgeeks.org/game-theory-in-ai/>

Name of the Game

Saturday, February 27, 2021 6:01 PM

DeepMind's StarCraft 2 AI is now better than 99.8 percent of all human players

Image: DeepMind

DeepMind today announced a new milestone for its artificial intelligence agents trained to play the Blizzard Entertainment game *StarCraft II*. The Google-owned AI lab's more sophisticated software, still called AlphaStar, is now grandmaster level in the real-time strategy game, capable of besting 99.8 percent of all human players in competition. The findings are to be published in a research paper in the scientific journal *Nature*.

Not only that, but DeepMind says it also evened the playing field when testing the new and improved AlphaStar against human opponents who opted into online competitions this past summer. For one, it trained AlphaStar to use all three of the game's playable races, adding to the complexity of the game at the upper echelons of pro play. It also limited AlphaStar to only viewing the portion of the map a human would see and restricted the number of mouse clicks it could register to 22 non-duplicated actions every five seconds of play, to align it with standard human movement.

ALPHASTAR IS THE FIRST EVER AI GRANDMASTER IN STARCRAFT II

Still, the AI was capable of achieving grandmaster level, the highest possible online competitive ranking, and marks the first ever system to do so in *StarCraft II*. DeepMind sees the advancement as more proof that general-purpose reinforcement learning, which is the machine learning technique underpinning the training of AlphaStar, may one day be used to train self-learning robots, self-driving cars, and create more advanced image and object recognition systems.

"The history of progress in artificial intelligence has been marked by milestone achievements in games. Ever since computers cracked Go, chess and poker, *StarCraft* has emerged by consensus as the next grand challenge," said David Silver, a DeepMind principle research scientist on the AlphaStar team, in a statement. "The game's complexity is much greater than chess, because players control hundreds of units; more complex than Go, because there are 10^26 possible choices for every move; and players have less information about their opponents than in poker."

Back in January, DeepMind announced that its AlphaStar system was able to best top pro players 10 matches in a row during a prerecorded session, but it lost to pro player Grzegorz "MaNa" Komincz in a final match streamed live online. The company kept improving the system between January and June, when it said it would start accepting invites to play the best human players from around the world. The ensuing matches took place in July and August, DeepMind says.

The results were stunning: AlphaStar had become among the most sophisticated *Starcraft II* players on the planet, but remarkably still not quite superhuman. There are roughly 0.2 percent of players capable of defeating it, but it is largely considered only a matter of time before the system improves enough to crush any human opponent.

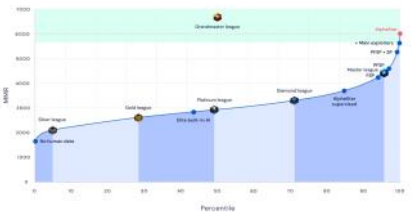


Image: DeepMind

This research milestone closely aligns with a similar one from San Francisco-based AI research company OpenAI, which has been training AI agents using reinforcement learning to play the sophisticated five-on-five multiplayer game *Dota 2*. Back in April, the most sophisticated version of the OpenAI Five software, as it's called, bested the world champion *Dota 2* team after only narrowly losing to two less capable e-sports teams the previous summer. The leap in OpenAI Five's capabilities mirrors that of AlphaStar's, and both are strong examples of how this approach to AI can produce unprecedented levels of game-playing ability.

Similar to OpenAI's *Dota 2* bots and other game-playing agents, the goal with this type of AI research is not just to crush humans in various games just to prove it can be done. Instead, it's to prove that — with enough time, effort, and resources — sophisticated AI software can best humans at virtually any competitive cognitive challenge, be it a board game or a modern video game. It's also to show the benefits of reinforcement learning, a special brand of machine learning that's seen massive success in the last few years when combined with huge amounts of computing power and training methods like virtual simulation.

Like OpenAI, DeepMind trains its AI agents against versions of themselves and at an accelerated pace, so that the agents can clock hundreds of years of play time in the span of a few months. That has allowed this type of software to stand on equal footing with some of the most talented human players of Go and, now, much more sophisticated games like *Starcraft* and *Dota*.

THIS TYPE OF AI MAY ONE DAY CONTROL SMARTER, SAFER, SELF-LEARNING ROBOTS

Yet the software is still restricted to the narrow discipline it's designed to tackle. The Go-playing agent cannot play *Dota*, and vice versa. (DeepMind did let a more general-purpose version of its Go-playing agent try its hand in chess, which it mastered in a matter of eight hours.) That's because the software isn't programmed with easy-to-replace rule sets or directions. Instead, DeepMind and other research institutions use reinforcement learning to let the agents figure out how to play on their own, which is why the software often develops novel and wildly unpredictable play styles that have since been adopted by top human players.

"AlphaStar is an intriguing and unorthodox player — one with the reflexes and speed of the best pros but strategies

Deep Blue (Chess Computer)

Humans were the strongest chess entities on the planet for centuries. Even in the 1980s, it seemed laughable that a computer could ever defeat the strongest human players. Then in 1997, it happened—a computer defeated the world champion. Which computer, you ask? Deep Blue.

Let's learn more about this computer that changed history. Here is what you need to know about Deep Blue:

What Is Deep Blue?

Deep Blue was a chess computer developed by IBM. It is famous for defeating the chess world champion, GM Garry Kasparov, in their 1997 match. Deep Blue's victory was viewed as a symbolic testament to the rise of artificial intelligence—a victory for machine versus man.

The Deep Blue project (initially called ChipTest) was created by Feng-hsiung Hsu in 1985. In 1989 Hsu and other colleagues joined the IBM team to fully develop Deep Blue. An early version of Deep Blue played a match against GM Joel Benjamin, who joined the Deep Blue team as a GM consultant afterward.

By the time of the 1997 match, Deep Blue's alpha-beta search algorithm (the same type of search that is still used by many conventional computer engines today) along with its custom hardware allowed it to consider up to **200 million positions per second**. Deep Blue was dismantled after the 1997 victory, with one of its two racks being displayed at the National Museum of American History and the other at the Computer History Museum.

One of the Deep Blue racks on display at the Computer History Museum. Photo: James, CC.

Deep Blue Accomplishments

Deep Blue played two matches against Kasparov in the 1990s. In the 1996 match, Deep Blue lost 2-4 but still accomplished something that no chess computer had done before: it defeated the human world champion in a game—an unprecedented accomplishment. Kasparov is still widely viewed as the greatest player of all time.

Garry Kasparov. Photo: Owen Williams/Kasparov Agency, CC.

Many improvements were made to Deep Blue in between the 1996 and 1997 matches. When they met in the 1997 rematch, Deep Blue defeated Kasparov 3.5-2.5 in standard time controls and in a tournament setting. This incredible victory was groundbreaking and marked an achievement for the world of artificial intelligence.

Deep Blue Games

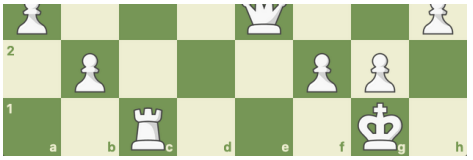
In the first game of the 1996 match, Deep Blue shocked the world by defeating Kasparov. It played the Alapin variation of the Sicilian and was able to force multiple structural weaknesses in Kasparov's position. After 24...exd5, all of Kasparov's pawns are either isolated, doubled, or both:

play Dota, and vice versa (DeepMind did test a more general purpose version of its Go-playing agent try its hand in chess, which it mastered in a matter of eight hours.) That's because the software isn't programmed with easy-to-replace rule sets or directions. Instead, DeepMind and other research institutions use reinforcement learning to let the agents figure out how to play on their own, which is why the software often develops novel and wildly unpredictable play styles that have since been adopted by top human players.

"AlphaStar is an intriguing and unorthodox player — one with the reflexes and speed of the best pros but strategies and a style that are entirely its own. The way AlphaStar was trained, with agents competing against each other in a league, has resulted in gameplay that's unimaginably unusual; it really makes you question how much of StarCraft's diverse possibilities pro players have really explored," Diego "Kelazhur" Schwimer, a pro player for team Panda Global, said in a statement. "Though some of AlphaStar's strategies may at first seem strange, I can't help but wonder if combining all the different play styles it demonstrated could actually be the best way to play the game."

DeepMind hopes advances in reinforcement learning achieved by its lab and fellow AI researchers may be more widely applicable at some point in the future. The most likely real-world application for such software is robotics, where the same techniques can properly train AI agents how to perform real-world tasks, like the operation of robotic hands, in virtual simulation. Then, after simulating years upon years of motor control, the AI can take the reins of a physical robotic arms, and maybe one day even control full-body robots. But DeepMind also sees increasingly more sophisticated — and therefore safer — self-driving cars as another venue for its specific approach to machine learning.

From <https://www.theverge.com/2019/10/30/20939147/deepmind-google-alphastar-starcraft-2-research-grandmaster-level>



From <https://www.chess.com/terms/deep-blue-chess-computer>

Artificial Intelligence In Stock Market Investing: Is It For You?

Artificial intelligence (AI) is increasingly becoming part of our lives, often without us even realizing it. I use it in my digital marketing agency and investing, and my phone uses it to enhance my user experience. Many of the online social media interactions and phone conversations we have are with computers. The difference is often unrecognizable. I believe this begs the question, why not use AI to help you invest?

I've been investing since I graduated from medical school. In my financial book for doctors, I explain that I sold all of my stocks years ago and quit the market. I've stayed out of publicly traded stocks for well over 10 years. Only recently have I found a couple of exceptions that work, including using AI programs that automate stock picking and the buying and selling of shares (aka algorithmic or automatic trading).

I've observed that a number of high-income earners — including business owners, C-suite executives, entrepreneurs with successful exits and independent professionals such as doctors — have been expanding their portfolios into a wider range of asset classes over the past decade. I've consulted many of them on real estate, debt investments, biotech startups and reputation management.

Stock trading is evolving. For example, we now have systems for tax loss harvesting that can help turn stock market losses into tax deductions. For many, I've observed that investing through self-directed IRAs, Roth IRAs and Solo 401(k)s can provide more flexibility and a variety of tax benefits.

The use of AI in trading has seen growth in recent years. *Wired* reported that at least 1,300 hedge funds are using some form of computer modeling for the majority of their trades.

There are three main ways I've observed this type of technology applied:

1. Discovering Patterns

Incredibly powerful computers are able to crunch almost countless data points in minutes. This means they are also able to detect historical and replicating patterns for smart trading that are often hidden to human investors. Humans simply aren't capable of processing that much data or seeing these patterns at the same rate of technology (if at all). Consider that AI can evaluate thousands of stocks in moments. According to CNN, when it comes to high-frequency trading, some hedge funds use AI to decipher as many as 300 million data points on the New York Stock Exchange in the first hour of daily trading alone.

2. Predictive Trading Based On Sentiment

By analyzing news headlines, social media comments, blogs and more, AI can predict the direction of stocks and the moves of other traders via sentiment analysis — the process of categorizing opinions (or sentiment) people have shared in text.

3. Speed In Trading

While not revolutionary, this technology adds even more speed to trading. Today every millisecond counts. AI means automated trading without even needing to call your broker or get on an app.

AI isn't perfect.

But, machine learning helps ensure this technology gets better and better. AI is already learning to continuously improve on its own mistakes. It deploys automated trading assistants and constantly works to improve its performance, not only by fine-tuning programming but also by inputting masses of new data.

At a minimum, AI can benefit the human roles related to automated trading. Morgan Stanley, for example, rolled out its own AI tools to enhance the work of its financial advisors.

But there are still caveats with using this technology to trade automatically. AI still relies on quality data being input, as well as the interpretation of that data. And despite AI's expected growth in the stock market, I've found that the sizes of your gains and losses can depend on the amount you've invested and the strategy you've employed; your net returns can be impacted by the volume and frequency of trades because you might be required to pay fees for using AI to do your trading.

Essentially, if you are trading all day, you might be incurring costs on each trade that will eat into your gains or deepen your losses. For example; if you were trading cryptocurrency with $100, and you made a 10% gain — but your buy-and-sell trading costs were $5 on each side — you wouldn't truly have any net gain.

While I believe it is still wise to have a well-rounded portfolio that allows for allocation in private investments, debt and real estate, if you're going to invest in stocks, you might consider leveraging AI. It could be especially helpful for busy professionals who want to invest intelligently but haven't mastered day trading.

From <https://www.forbes.com/sites/forbesdallascouncil/2019/04/15/artificial-intelligence-in-stock-market-investing-is-it-for-you/?sh=1a9ff5496524>

LAWS

Saturday, February 27, 2021 7:44 AM

A BAE Systems Corax during flight testing

Lethal autonomous weapons (**LAWs**) are a type of autonomous military system that can independently search for and engage targets based on programmed constraints and descriptions.[1] LAWs are also known as **lethal autonomous weapon systems** (LAWS), **autonomous weapon systems** (AWS), **robotic weapons**, **killer robots** or **slaughterbots**.[2] LAWs may operate in the air, on land, on water, under water, or in space. The autonomy of current systems as of 2018 was restricted in the sense that a human gives the final command to attack - though there are exceptions with certain "defensive" systems.

Contents

Being autonomous as a weapon

Being "autonomous" has different meanings in different fields of study. In engineering it may refer to the machine's ability to operate without human involvement. In philosophy it may refer to an individual being morally independent. In political science it may refer to an area's capability of self-governing. In terms of military weapon development, the identification of a weapon as autonomous is not as clear as in other areas.[2] The specific standard entailed in the concept of being autonomous can vary hugely between different scholars, nations and organizations.

Various people have many definitions of what constitutes a lethal autonomous weapon. Heather Roff, a writer for Case Western Reserve University School of Law, describes autonomous weapon systems as "armed weapons systems, capable of learning and adapting their 'functioning in response to changing circumstances in the environment in which [they are] deployed,' as well as capable of making firing decisions on their own."[3] This definition of autonomous weapon systems is a fairly high threshold compared to the definitions of scholars such as Peter Asaro and Mark Gubrud's definitions seen below.

Scholars such as Peter Asaro and Mark Gubrud are trying to set the threshold lower and judge more weapon systems as autonomous. They believe that any weapon system that is capable of releasing a lethal force without the operation, decision, or confirmation of a human supervisor can be deemed autonomous. According to Gubrud, a weapon system operating partially or wholly without human intervention is considered autonomous. He argues that a weapon system does not need to be able to make decisions completely by itself in order to be called autonomous. Instead, it should be treated as autonomous as long as it actively involves in one or multiple parts of the "preparation process", from finding the target to finally firing.[4][5]

Other organizations, however, are setting the standard of autonomous weapon system in a higher position. The Ministry of Defence (United Kingdom) defines autonomous weapon systems as "systems that are capable of understanding higher level intent and direction. From this understanding and its perception of its environment, such a system is able to take appropriate action to bring about a desired state. It is capable of deciding a course of action, from a number of alternatives, without depending on human oversight and control- such human engagement with the system may still be present, though. While the overall activity of an autonomous unmanned aircraft will be predictable, individual actions may not be."[6]

As a result, the composition of a treaty between states requires a commonly accepted labeling of what exactly constitutes an autonomous weapon.[7]

Automatic defensive systems[edit]

The oldest automatically triggered lethal weapon is the land mine, used since at least the 1600s, and naval mines, used since at least the 1700s. Anti-personnel mines are banned in many countries by the 1997 Ottawa Treaty, not including the United States, Russia, and much of Asia and the Middle East.

Some current examples of LAWs are automated "hardkill" active protection systems, such as a radar-guided CIWS systems used to defend ships that have been in use since the 1970s (e.g., the US Phalanx CIWS). Such systems can autonomously identify *and attack* oncoming missiles, rockets, artillery fire, aircraft and surface vessels according to criteria set by the human operator. Similar systems exist for tanks, such as the Russian Arena, the Israeli Trophy, and the German AMAP-ADS. Several types of stationary sentry guns, which can fire at humans and vehicles, are used in South Korea and Israel. Many missile defense systems, such as Iron Dome, also have autonomous targeting capabilities. Automatic turrets installed on military vehicles are called remote weapon stations.

The main reason for not having a "human in the loop" in these systems is the need for rapid response. They have generally been used to protect personnel and installations against incoming projectiles.

Autonomous offensive systems[edit]

Systems with a higher degree of autonomy would include drones or unmanned combat aerial vehicles, e.g.: "The unarmed BAE Systems Taranis jet-propelled combat drone prototype may lead to a Future Combat Air System that can autonomously search, identify, and locate enemies but can only engage with a target when authorized by mission command. It can also defend itself against enemy aircraft" (Heyns 2013, §45). The Northrop Grumman X-47B drone can take off and land on aircraft carriers (demonstrated in 2014); it is set to be developed into an Unmanned Carrier-Launched Airborne Surveillance and Strike (UCLASS) system.

Serbian Land Rover Defender towing trailer with "Miloš" tracked combat robot.

According to *The Economist*, as technology advances, future applications of unmanned undersea vehicles might include mine clearance, mine-laying, anti-submarine sensor networking in contested waters, patrolling with active sonar, resupplying manned submarines, and becoming low-cost missile platforms.[8] In 2018 the U.S. Nuclear Posture Review alleged that Russia is developing a "new intercontinental, nuclear-armed, nuclear-powered, undersea autonomous torpedo" named "Status 6".[9]

The Russian Federation is actively developing artificially intelligent missiles,[10] drones,[11] unmanned vehicles, military robots and medic robots.[12][13][14][15]

Israeli Minister Ayoob Kara stated in 2017 that Israel is developing military robots, including ones as small as flies.[16][17]

In October 2018, Zeng Yi, a senior executive at Chinese Defense Firm Norinco, gave a speech in which he said that "In future battlegrounds, there will be no people fighting," and that the use of lethal autonomous weapons in warfare is "inevitable."[18] In 2019, US Defense Secretary Mark Esper lashed out at China for selling drones capable of taking life with no human oversight.[19]

The British Army deployed new unmanned vehicles and military robots in 2019.[20]

The US Navy is developing "ghost" fleets of unmanned ships.[21]

Ethical and legal issues[edit]

Standard used in US policy

Current US policy states: "Autonomous … weapons systems shall be designed to allow commanders and operators to exercise appropriate levels of human judgment over the use of force."[22] However, the policy requires that autonomous weapon systems that kill people or use kinetic force, selecting and engaging targets without further human intervention, be certified as compliant with "appropriate levels" and other standards, not that such weapon systems cannot meet these standards and are therefore forbidden.[23] "Semi-autonomous" hunter-killers that autonomously identify and attack targets do not even require certification.[23] Deputy Defense Secretary Robert Work said in 2016 that the Defense Department would "not delegate lethal authority to a machine to make a decision", but might need to reconsider this since "authoritarian regimes" may do so.[24] In October 2016 President Barack Obama stated that early in his career he was wary of a future in which a US president making use of drone warfare could "carry on perpetual wars all over the world, and a lot of them covert, without any accountability or democratic debate".[25][26] In the US, security-related AI has fallen under the purview of the National Security Commission on Artificial Intelligence since 2018.[27][28] On October 31, 2019, the United States Department of Defense's Defense Innovation Board published the draft of a report outlining five principles for weaponized AI and making 12 recommendations for the ethical use of artificial intelligence by the Department of Defense that would ensure a human operator would always be able to look into the 'black box' and understand the kill-chain process. A major concern is how the report will be implemented.[29]

Possible violations of ethics and international acts

Stuart Russell, professor of computer science from University of California, Berkeley stated the concern he has with LAWS is that his view is that it is unethical and inhumane. The main issue with this system is it is hard to distinguish between combatants and non-combatants.[30]

There is concern by some (e.g.g Noel Sharkey 2012) about whether LAWs would violate International Humanitarian Law, especially the principle of distinction, which requires the ability to discriminate combatants from non-combatants, and the principle of proportionality, which requires that damage to civilians is proportional to the military aim.[31] This concern is often invoked as a reason to ban "killer robots" altogether - but it is doubtful that this concern can be an argument against LAWs that do not violate International Humanitarian Law.[32][33]

LAWs are said by some to blur the boundaries of who is responsible for a particular killing,[34] but Thomas Simpson and Vincent Müller argue that they may make it easier to record who gave which command.[35] Likewise, Steven Umbrello, Phil Torres and Angelo F. De Bellis argue that if the technical capacities of LAWs are at least as accurate as human soldiers, then given the psychological shortcomings of human soldiers in war warrants that only these types of ethical LAWs should be used. Likewise, they propose using the value sensitive design approach as a potential framework for designing these laws to align with human values and International Humanitarian Law.[33]

Campaigns to ban LAWs

LETHAL AUTONOMOUS WEAPONS

What are they?

Lethal autonomous weapons systems, sometimes called “killer robots,” are weapon systems that use artificial intelligence to identify, select, and kill human targets without human control. This means that the decision to kill a human target is no longer made by humans, but by algorithms.

A common misconception is that these weapons are “science fiction.” In fact, given the increasing interest in the militarization of AI, lethal autonomous weapons systems are currently under development by some countries. They could be used on the battlefield very soon.

Why are they problematic?

Lethal autonomous weapons systems are intrinsically amoral and pose grave threats to national and global security, including to superpowers. The United Nations Secretary General has stated that “machines with the power and discretion to take lives without human involvement are politically unacceptable, morally repugnant and should be prohibited by international law.”

Beyond ethical issues, replacing human decision-making with algorithms is highly destabilizing. Cheap mass-production and/or copying of the code of algorithms for lethal autonomous weapons would fuel proliferation to both state and non-state actors, including domestic and foreign terrorists. This scalability of lethal AWS has led some to classify their nature as weapons of mass destruction (WMDs).

Further, because they can operate at a speed beyond human intervention, they introduce significant risks , such as the rapid conflict escalation, unreliable performance, and unclear attribution. The potential use of facial recognition software makes these weapons systems uniquely equipped to selectively target specific individuals or groups by gender, ethnicity, or other biometric data.

What can be done?

There are two key elements central to the global regulation of lethal autonomous weapons.

1. The Positive Obligation of Human Control: The first is a commitment by countries that **all weapons systems must operate under meaningful human control**. This means that humans, not algorithms, make the decision to kill (i.e. humans “press the button” to use lethal force, not an AI system). This requirement for collaborative human/AI decision making is already employed by many of the military’s semi-autonomous systems today, such as lethal drones.

2. Prohibitions on Systems Incapable of Human Control: The second element is for countries to agree to prohibit weapons systems incapable of meeting the human control requirement.

As we develop governance structures to steer our future with AI towards a positive outcome for humanity, we must set clear precedents on acceptable and unacceptable uses of AI. That begins with drawing a red line that the decision to take a human life must never be delegated to algorithms.

From <https://autonomousweapons.org/>

TECH

Elon Musk And Over 100 AI Experts Are Urging The UN to Ban Killer Robots

Elon Musk and more than 100 leaders and experts in artificial intelligence (AI) have come together urging the UN to commit to an outright ban on killer robot technology.

An open letter signed by Musk, Google Deepmind's Mustafa Suleyman, and 114 other AI and robotics specialists urges the UN to prevent "the third revolution in warfare" by banning the development of all lethal autonomous weapon systems.

The open letter, released to coincide with the world's largest conference on AI – IJCAI 2017, which is taking place in Melbourne, Australia this week – warns of a near future where independent machines will be able to choose and engage their own targets, including innocent humans in addition to enemy combatants.

"Once developed, they will permit armed conflict to be fought at a scale greater than ever, and at timescales faster than humans can comprehend," the consortium writes.

"These can be weapons of terror, weapons that despots and terrorists use against innocent populations, and weapons hacked to behave in undesirable ways."

It's not the first time Musk and those of his world view have united to draw attention to the threat autonomous weapons pose to humanity.

The SpaceX and Tesla chief is also behind OpenAI, a nonprofit devoted to advancing ethical AI research.

But despite the concerns AI experts are voicing, ongoing delays in developing an effective ban against autonomous weapons have led some to fear the dangers could be beyond regulation, especially given the rapid pace at which AI systems are developing.

"We do not have long to act," the open letter reads. "Once this Pandora's box is opened, it will be hard to close."

The "third revolution" to which the campaigners refer positions killer robots as a kind of technological successor to the historical developments of gunpowder and nuclear weaponry – innovations that haven't exactly improved the world we live in.

While the new letter isn't the first instance where experts have leveraged IJCAI to make their point, it is the first time that representatives of AI and robotics companies – from some 26 countries – have made a joint stand on the issue, joining the ranks of independent researchers including Stephen Hawking, Noam Chomsky, and Apple co-founder Steve Wozniak.

"The number of prominent companies and individuals who have signed this letter reinforces our warning that this is not a hypothetical scenario, but a very real, very pressing concern which needs immediate action," says the founder of Clearpath Robotics, Ryan Gariepy.

"We should not lose sight of the fact that, unlike other potential manifestations of AI which still remain in the realm of science fiction, autonomous weapons systems are on the cusp of development right now."

That last point is one that should be emphasised. While the ultimate nightmare of autonomous weapon systems could be a future populated with T–800s like the one at the top of this page, the reality is that AI-based killing machines are already a thing.

Autonomous or semi-autonomous capability is increasingly being built into weapons like the Samsung SGR-A1 sentry gun, the BAE Systems Taranis drone, and DARPA's Sea

that they may make it easier to record who gave which command.[39] Likewise, Steven Umbrello, Phil Torres and Angelo F. De Bellis argue that if the technical capacities of LAWs are at least as accurate as human soldiers, then given the psychological shortcomings of human soldiers in war warrants that only these types of ethical LAWs should be used. Likewise, they propose using the value sensitive design approach as a potential framework for designing these laws to align with human values and International Humanitarian Law.[39][40]

Campaigns to ban LAWs

The possibility of LAWs has generated significant debate, especially about the risk of "killer robots" roaming the earth- in the near or far future. The group Campaign to Stop Killer Robots formed in 2013. In July 2015, over 1,000 experts in artificial intelligence signed a letter warning of the threat of an artificial intelligence arms race and calling for a ban on autonomous weapons. The letter was presented in Buenos Aires at the 24th International Joint Conference on Artificial Intelligence (IJCAI-15) and was co-signed by Stephen Hawking, Elon Musk, Steve Wozniak, Noam Chomsky, Skype co-founder Jaan Tallinn and Google DeepMind co-founder Demis Hassabis, among others.[36][37]

According to PAX (One of the founding organisations of The Campaign to Stop Killer Robots), fully automated weapons (FAWs) will lower the threshold of going to war as soldiers are removed from the battlefield and the public is distanced from experiencing war, giving politicians and other decision-makers more space in deciding when and how to go to war.[38] They warn that once deployed, FAWs will make democratic control of war more difficult - something that author of *Kill Decision* - a novel on the topic - and IT specialist Daniel Suarez also warned about: according to him it might recentralize power into very few hands by requiring very few people to go to war.[38]

There are websites[clarification needed] protesting the development of LAWs by presenting undesireable ramifications if research into the appliance of artificial intelligence to designation of weapons continues. On these websites, news about ethical and legal issues are constantly updated for visitors to recap with recent news about international meetings and research articles concerning LAWs.[38]

The Holy See has called for the international community to ban the use of LAWS on several occasions. In November 2018, Archbishop Ivan Jurkovic, the permanent observer of the Holy See to the United Nations, stated that "In order to prevent an arms race and the increase of inequalities and instability, it is an imperative duty to act promptly: now is the time to prevent LAWS from becoming the reality of tomorrow's warfare." The Church worries that these weapons systems have the capability to irreversibly alter the nature of warfare, create detachment from human agency and put in question the humanity of societies.[40]

As of 29 March 2019, the majority of governments represented at a UN meeting to discuss the matter favoured a ban on LAWS.[41] A minority of governments, including those of Australia, Israel, Russia, the UK, and the US, opposed a ban.[41]

No ban, but regulation

A third approach focuses on regulating the use of autonomous weapon systems in lieu of a ban.[42] Military AI arms control will likely require the institutionalization of new international norms embodied in effective technical specifications combined with active monitoring and informal ('Track II') diplomacy by communities of experts, together with a legal and political verification process.[43][44][45]

From <https://en.wikipedia.org/wiki/Lethal_autonomous_weapon>

That last point is one that should be emphasised. While the ultimate nightmare of autonomous weapon systems could be a future populated with T–800s like the one at the top of this page, the reality is that AI-based killing machines are already a thing.

Autonomous or semi-autonomous capability is increasingly being built into weapons like the Samsung SGR-A1 sentry gun, the BAE Systems Taranis drone, and DARPA's Sea Hunter submarine.

In other words, the technological seeds of tomorrow's killer robots are already in existence on land, sea, and air – and effective laws to regulate these lethal machines (and the industry that's hell-bent on making them) haven't yet been written down.

Well, there's no time like the present.

"Nearly every technology can be used for good and bad, and artificial intelligence is no different," says AI researcher Toby Walsh from Australia's UNSW, one of the organisers of IJCAI 2017.

"It can help tackle many of the pressing problems facing society today… [h]owever, the same technology can also be used in autonomous weapons to industrialise war. We need to make decisions today choosing which of these futures we want."

From <https://www.sciencealert.com/elon-musk-and-over-100-ai-experts-are-urging-the-un-to-ban-killer-robots>

Thursday, February 25, 2021 4:49 PM

The Pentagon's Autonomous Swarming Drones Are the Most Unsettling Thing You'll See Today

Ejected from a fighter, the tiny drones collaborate to accomplish their mission—with not a single human involved.

An arm of the Pentagon charged with fielding critical new technologies has developed a drone that not only carries out its mission without human piloting, but can talk to other drones to collaborate on getting the job done. The Perdix autonomous drone operates in cooperative swarms of 20 or more, working together towards a single goal. Named after a character from Greek mythology that was changed into a partridge, the bird-sized Perdix drones were featured on the news program "60 Minutes" last night, January 8. In footage taken over the skies of Naval Air Weapons Station China Lake, a trio of F/A-18 Super Hornet fighters release a total of 103 Perdix drones from small pods mounted on hardpoints on both wings. The drones are capable of withstanding ejection at speeds of up to Mach 0.6 and temperatures as low as minus 10 degrees Celsius.

GPS data, combined with a map of the area, shows that during the October 26 test the fighters released their Perdix drones in a long line their flight path. The drones formed up at a preselected point and then headed out to perform four different missions. Three of the missions involved hovering over a target while the fourth mission involved forming a 100-meter-wide circle in the sky.

According to the Department of Defense, the demonstration showed off Perdix's collective decision-making, adaptive formation flying, and self-healing abilities. The drones collectively decide that a mission has been accomplished, fly on to the next mission, and carry out that one. The benefit of a swarm is that if one drone drops out—and a few appear to crash—the group can rearrange itself to maintain coverage.

Developed by students at the Massachusetts Institute of Technology, the Perdix drones are inexpensive drones that draw inspiration from the commercial smartphone industry. The drones feature two sets of small wings, making them look like World War I fighter planes. The biplane configuration reduces wing weight and wingspan. The wings are made of carbon fiber and the fuselage is made of a kevlar composite. The drone is powered by a lithium polymer battery pack powering a rear-facing push propeller.

Perdix has been a known program since March 2016, when the Washington Post revealed footage of a F-16 fighter releasing 20 drones over Alaska. At the time, however, the *Post* stated the drones had already been undergoing flight testing for two years.

Self-piloting collaborating drones could take allow humans to sit back and analyze drone collected imagery without having to manage scores of drones. ACHILLEAS ZAVALLIS/AFP/Getty Images.

There are a multitude of uses for such a drone swarm. The drones could be released by fighters to provide reconnaissance for troops on the ground, hunting enemy forces and reporting their location. They could also jam enemy communications, form a wide-area flying communications network, or provide persistent surveillance of a particular area. They could be loaded with small explosive charges and attack individual enemy soldiers. In air-to-air combat, they could spoof enemy radars on aircraft, ground vehicles, and missiles by pretending to be much larger targets.

The drones are a pet project of the Strategic Capabilities Office, which is in turn part of the Pentagon's Third Offset Strategy. Third Offset is designed to use America's technological edge and combine it with new ideas to maintain dominance against potential adversaries. Other concepts include the "Arsenal Plane", which uses older, larger aircraft such as the B-52 to act as a flying arsenal for newer planes like the F-35, carrying a vast number of weapons that can be fired on cue.

From <https://www.popularmechanics.com/military/aviation/a24675/pentagon-autonomous-swarming-drones/>

Slaughterbots

Saturday, February 27, 2021 7:49 AM

Slaughterbots

TECH

This Horrifying 'Slaughterbot' Video Is The Best Warning Against Autonomous Weapons

DAVID NIELD

22 NOVEMBER 2017

We're on the verge of creating autonomous weapons that can kill without any help from humans. Thousands of experts are concerned about this - and the latest campaign effort against this tech is a chilling video demonstrating the kind of future we're heading for.

In the *Slaughterbots* short, which we've embedded below, swarms of AI-controlled drones carry out strikes on thousands of unprepared victims with targeted precision. What makes the clip so scary is that the scenario is entirely plausible.

The video starts with a spectacular press event where the technology is unveiled for the first time. The miniature drones are able to take out "the bad guys" – whoever they happen to be – without any collateral damage, nuclear weapons, or troops on the ground.

All the drone bots need is a profile: age, sex, fitness, uniform, and ethnicity.

Despite the applause that the tiny drones get at their unofficial unveiling, the tech behind them soon falls into the wrong hands, as it tends to do. Before long an attack is waged on the United States Capitol building, with only politicians from one particular party targeted.

The same bots are then used by an unknown group to take out thousands of students worldwide, all of whom shared the same human rights video on social media.

"The weapons took away the expense, the danger, and risk of waging war," says one of the talking heads in the clip, admitting that "anyone" can now carry out such a strike.

Thankfully this creepy video isn't real life – at least not yet.

It was published by the Campaign to Stop Killer Robots, an international coalition looking to ban autonomous weapons, and was shown this week at the UN Convention on Certain Conventional Weapons.

The group wants the UN to pass legislation prohibiting the development of this kind of AI technology, and the large-scale manufacture of the associated hardware. Legislation could also be used to police anyone who tried to develop these kind of systems.

Worryingly, these are all technologies we already have, according to one of the experts behind the video, computer scientist Stuart Russell from the University of California, Berkeley – the only step remaining is for someone to miniaturise and combine them.

"I've worked in AI for more than 35 years," says Russell in the video. "Its potential to benefit humanity is enormous, even in defence, but allowing machines to choose to kill humans will be devastating to our security and freedom."

"Thousands of my fellow researchers agree. We have that opportunity to prevent the future you just saw, but the window to act is closing fast."

Experts including Elon Musk and Stephen Hawking have also warned about the rapid development of AI, and its use in weapons.

Computer systems are now able to pilot drones on their own, and recognise faces faster than human beings can. If they were also allowed to pull the trigger on a weapon without any human approval, scientists say, wars would rage at a speed and with a loss of life far greater than anything we've ever seen before.

Let's hope that this Slaughterbots video, and other initiatives to curb the development of AI-powered weaponry, prove enough to put a stop to this particular area of research.

Noel Sharkey, AI professor at Sheffield University in the UK, and chair of the International Committee on Robot Arms Control, has been warning about the dangers of autonomous weapons for a decade.

"It will only take one major war to unleash these new weapons with tragic humanitarian consequences and destabilisation of global security," he told *The Guardian*.

You can find out more about efforts to support a ban on the Campaign to Stop Killer Robots website.

From <https://www.sciencealert.com/chilling-drone-video-shows-a-disturbing-vision-of-an-ai-controlled-future>