SWCP Portal
September 2019

What's really going on inside that black box?

The Promise and Perils of AI

The age of robots is getting closer every day. It seems that every company from **Alphabet** to **Uber** is working on self-driving cars or **talking software assistants** or posting videos to show off the **latest gymnastic feats** of humanoid or 4-legged robots.

But how do they work? Will these things think like we do? Can we actually depend on them finding solutions to novel problems and situations? There are a vast number of questions that must be answered.

Making machines that think

What will make these future smart machines and programs go is **artificial intelligence** (AI for short). While AI systems are based on the same fundamental electronic technologies as are all digital computers, the critical difference is in how they process data.

Digital computers depend upon **programs** to operate. These programs are lists of highly-specific commands telling the device precisely how to take in, process, and execute data, whether the machine is calculating a payroll, sending email, or running trains.

Programs are written in **code** that the machine reads which can be highly complicated, but they are highly restricted in terms of what they can do. Most programs generally do one thing and that very well, but they cannot deal with unexpected circumstances.

In other words, **"regular" programs cannot learn**. Giving computers that ability and to react independently is the great dream of AI. For over *60 years*, researchers have struggled to make it happen with far more disappointing dead-ends than successes.

Many different approaches have been tried. One was to model events by **framing** all possible outcomes of a situation. But such models were unable to incorporate every last possibility, even for such strictly ritualized behaviors such as ordering in a restaurant.

Other efforts attempted to utilize already available data in restricted settings, such as the medical wisdom of doctors, to build **expert systems**. Another approach was to apply **logic** to investigate the basis of reasoning. None were particularly successful, although in time, developers were able to construct devices that worked like human senses. They invented systems that could **see images** from TV cameras, as well as simulate touch and hearing.

Unable to create a nimble, intelligent system, the frustrated scientists increasingly turned to biological models. Humans may be unable to act as repetitively and precisely as machines, but we were born to deal with the uncertainty found in unfamiliar situations.

Brain structures – **neural networks** – were studied and carefully modeled to find out how they function. Techniques called **machine learning** evolved that could deal with uncertainties in the data. These taught computers by showing them thousands of images, "rewarding" them with confirmation whenever they made the proper identification.

This works much the same way online **Captcha identity tests** determine a if commentator is a human and not a spambot. It, too, has the person tell it which images have an item that a human should spot sooner than a machine – such as a stop sign.

The magic is all in the algorithms

The recipes that computers use for handling data, whether created by machine learning or old-fashioned programming by hand, are called **algorithms**. They form the essential basis of any AI system.

Algorithms are logical sets of instructions. Basically **mathematical equations**, they can be combined and become very complex. In AI systems, algorithms can even be tweaked by the computer as it learns.

Machine learning to create algorithms has proven very effective indeed, leading to some surprising results. By pitting algorithms against each other **adversarily**, machines were able to train themselves to play the Japanese game of Go with new moves unknown to Go masters (or other AIs). Much more sinister are **deepfake** audio or video, which turns media taken of one person into that of another.

But that's just the beginning. Right now, a vast amount of effort and money is being poured into **natural language comprehension**, absolutely vital to make talking digital assistants like Siri, Alexa, or Google's own, work. Self-driving vehicles are another

Continued on back

Continued from front

area where AI gets much publicity. Yet AI is already in the process of being **deployed** in **every field** imaginable, but in many cases, it can be hard to recognize.

For instance, **language translation** is now often done by AI online, from text or living speech. Google Maps uses smartphone data to **analyze traffic speed** and suggest faster commutes. Ridesharing apps **Uber** and Lyft use AI to determine efficient routes and minimize waiting. **UPS** uses it to navigate, route packages, and avoid delays like snowstorms.

There are extensive applications throughout aviation, starting with autopilots. Nowadays, a trip on a **Boeing jet** may involve just *7 minutes* of human-steered flight, mainly during takeoffs and landings.

AI is already finding many uses in **healthcare** from research with huge datasets to aiding diagnosis and personalizing medicine to nurse robots for the aged.

Military applications are getting a lot of attention, too – particularly in the question of **autonomous killing machines**. Fears of a future robot uprising and movements to **ban killer robots** may have slowed this arms race down, but it may be too late. For there are already **automated sentry guns** in place along the South Korean border with the North.

As might be expected, machine learning and AI are widely used in many online contexts. **Gmail's spam filters and prioritization** of messages is but one. It is in marketing and social media where AI is making the biggest impact. Photos are cataloged on **Facebook** and enhanced on **Snapchat**; AI reads **emoji in Instagram** and finds items of interest on **Pinterest**.

AI is highly involved in **banking**, and basically **runs the stock market**. Behind the scenes, AI is what is busily **mining data** on everyone to manipulate us – whether it be **recommendations** on Amazon or Netflix or even the **Russians placing fake news** on Facebook. The fact is that AI is **already everywhere**, observing, judging, and subtly influencing our world.

AI has even made discoveries of its own – like **predicting schizophrenia** in patients through voice or text samples. It can spot **plagiarism** in school essays, **finish original stories**, or write effective **ad copy**.

Inside the black box

Perceptive readers will have noticed a number of events that haven't been mentioned. Driverless cars have caused accidents – including at least one fatality due to **not recognizing** a woman crossing the road in time. Boeing's tragic difficulties with its **737 MAX aircraft** are directly due to AI complexities and a failure to train for them. AI even caused a 5-minute "**flash crash**" that terrified Wall Street in 2010.

These all involve the same enormous enigma – **external observers do not know *how* the machine reaches its conclusions**. The computer determines its *own* decision-making criteria *all by itself*, which has already had fatal outcomes in the real world.

For example, **not knowing** *how* a visual system sees a stop sign might mean a sticker on it could **fool** a self-driving car into not recognizing it. But the code is so complex that such events might not be predictable before they occur – or understandable after.

When AI systems interact, it can get even stranger – like **enormous prices** being charged on Amazon by dueling AIs, or the **two chatbots** that had to be turned off because they developed their own language that researchers could not understand.

Artificial intelligence is very powerful, but despite the ongoing **arms race**, limited. It's almost like a **magic trick** that works wonderfully until something goes wrong. That's why all the big companies have had people **listening** to their voice assistants and even **writing comedy** for them. For if machines may now learn like children, they must be as **carefully taught**.

Did you know? Temp or private emails!

Do you ever need a valid email address for a short time, say to activate an online service, but want to avoid more spam flooding your inbox? **Temporary email addresses** are freely available to our members which can be set to work for up to a month.

Also, completely secure messages can be sent using our **Burn After Reading** client. You can write a confidential message on our website, email a link to the recipient (who doesn't have to be a fellow SWCP customer). Once that person opens and reads the message, it vanishes completely without a trace.

To use either service, visit our homepage ((**https://www.swcp.com/**). Under the **`Support`** heading, click on **`Members Portal`**, and enter your username and password. On the **`Account Management`** page that comes up, both features are listed on the right. Simply click to go and compose your message.

Portal *editor/chief writer, Jay Nelson* **jnelson@swcp.com**
Click on blue terms in PDF file to open links.