

# { A world that is controlled by algorithms }

Logical computer systems have a growing impact on everyday life



**Bruno de Pierro**

PUBLISHED IN APRIL 2018

**A**lgorithms are everywhere. When share prices rise and fall, algorithms are typically involved. According to data that were released in 2016 by the Institute for Applied Economic Research (IPEA), investment robots that are programmed to instantly react to specified scenarios account for more than 40% of stock market transactions in Brazil. In the United States, this figure is 70%. The success of a simple Google search depends on these computer programming procedures, which can filter billions of web pages in mere seconds; the importance of a website, as defined by an algorithm, is based on the quantity and quality of other pages that link to it. At the frontier of automotive engineering research, sets of algorithms are used by autonomous cars to process information that has been captured by cameras and sensors to instantly make decisions at the wheel without human intervention.



## What is an algorithm?

An algorithm is a logical sequence of steps that are used to solve a problem and are written in a computer programming language



## How to build an algorithm

1

### Identify the problem

The objective of the algorithm must be defined. If the objective is to use images to more accurately detect cancer, computer scientists can design a strategy that examines the characteristics of the tumors, the available data, and the possible methods of diagnosis

2

### Develop a solution

The next step is to establish the sequence of steps that must be performed to solve the problem. In the example of cancer diagnosis, this involves analyzing the available medical images, comparing tumors and their volumes, and collecting data on how the disease progresses

3

### Translation and development

The solution is expressed in a programming language. Each step is translated into lines of code that contain the commands that are needed to execute it. Some algorithms are composed of millions or even billions of lines of code



## Who creates an algorithm?

The development of an algorithm can involve tens or even hundreds of experts, depending on its complexity

1

### Computer scientists

Computer scientists identify the solution and determine the steps that are needed to build the algorithm. They liaise with professionals from other fields

2

### Other professionals

Computer scientists are advised by professionals who have expert knowledge of the problem. They may be sociologists, demographers, economists, or doctors—this depends on the objective

3

### Programmers

Programmers are responsible for writing parts of the algorithm in computer language, without necessarily knowing anything about the other parts

4

### Algorithms

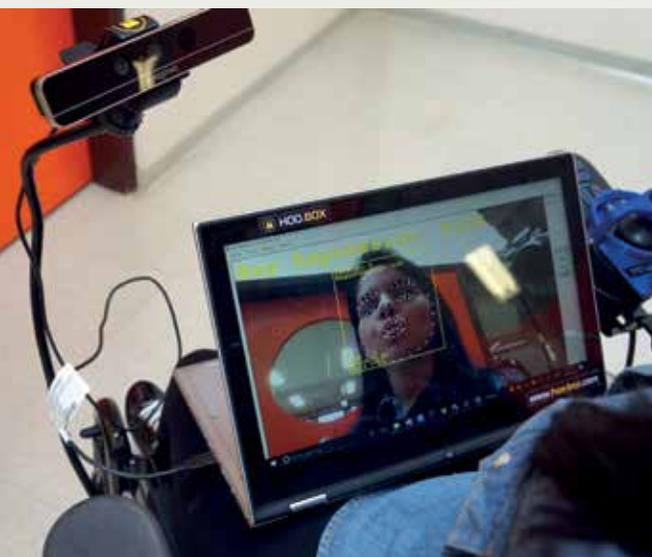
There are algorithms that are capable of developing other algorithms by modifying programming code that has been written by humans



# Facial Expression

Hoobox Robotics, which is a company that was founded by researchers from UNICAMP in 2016, has developed a system for motorized wheelchairs that enables quadriplegics to control the chair using only facial expressions. The algorithm that is used by the software, which is called Wheelie, translates up to 11 facial expressions, such as a smile or a raised eyebrow, into commands to move forward, backward, left, and right. The program is being tested by 39 patients in the USA, where the company has a research unit at the Johnson & Johnson laboratory in Houston. The system uses a 3D camera to capture dozens of facial points.

“The user can configure a command for each expression. A smile, for example, can move the chair forward, a kiss, back,” explains computer scientist Paulo Gurgel Pinheiro, who is the director of Hoobox. To learn to recognize key expressions, the Wheelie algorithm studied a set of facial data from 103 truck drivers. “We partnered with a transportation company to install cameras in trucks and record the facial expressions of volunteers over three months,” Gurgel explains.



An algorithm translates facial expressions into commands for controlling motorized wheelchairs

Although they play a role in even the most mundane tasks, such as traffic avoidance via mobile applications, algorithms are often viewed as intangible by the general population, who feel their effects but do not know or understand what they are or how they work. An algorithm is nothing more than a sequence of steps that are used to automatically solve a problem or accomplish a task, regardless of whether a dozen or a million lines of programming code are required. “It is the nucleus of any computational process,” says computer scientist Roberto Marcondes Cesar Junior, who is a researcher at the Institute of Mathematics and Statistics of the University of São Paulo (IME-USP).

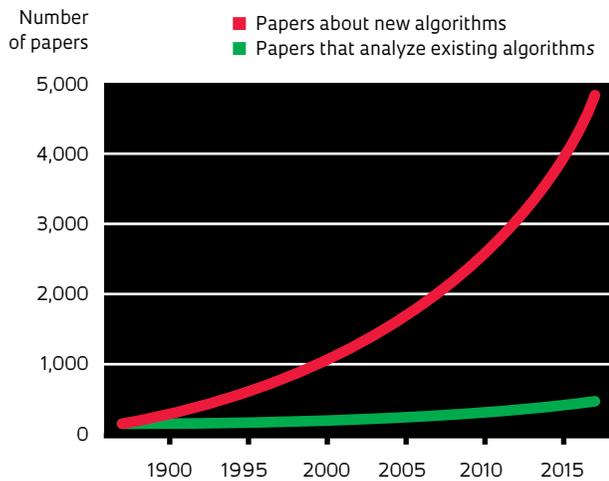
Consider the sequence of steps that are performed by the Facebook algorithm, for example. The choice of what to display in a user’s news feed is based primarily on the set of posts that have been produced by or are circulating among the user’s friends. The algorithm analyzes this information and discards posts that have been flagged as violent or inappropriate, posts that appear to be spam, and posts in which the wording is identified as “clickbait”—a form of exaggeration that is used to encourage users to click a link. Finally, the algorithm assigns a score to each post that is based on the user’s activity history and estimates how likely the user is to enjoy or share the information. The algorithm has recently been modified to reduce the reach of posts that have been made by news outlets.

## Robots are responsible for 40% of the decisions that are made on the Brazilian stock market

The development of an algorithm involves three steps (*see the infographic on page 59*): The first is to accurately identify the problem and find a solution to it. In this phase, computer programmers work with professionals who understand the task that must be performed. They could be doctors, in the case of an algorithm that analyzes imaging exams; sociologists, if the objective is to identify patterns of violence in regions of a city; or psychologists and demographers in the development of a dating application. “The challenge is to show that a practical solution to the problem exists, that it is not a problem of exponential complexity, for which the time needed to produce a response can increase exponentially, making it impractical,” explains computer scientist Jayme Szwarcfiter, who is a researcher at the Federal University of Rio de Janeiro (UFRJ).

# Rarely studied systems

An artificial intelligence journal\* was found to publish more articles that describe new algorithms than articles that study the behavior of existing ones



\* Proceedings of the Neural Information Processing Systems Conference

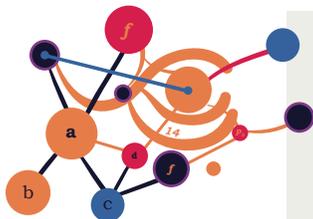
SOURCE RAHWAN ET AL. CLOSING THE AI KNOWLEDGE GAP. ARXIV, 2018

The second phase is also free of mathematical operations: it consists of describing the sequence of steps in normal language, for everyone to understand. Next, this description is translated into a programming language during phase three. Only then can the computer understand the commands—which can be simple, mathematical operations or complex algorithms within algorithms—all in a logical and precise sequence. During this stage, programmers are tasked with

writing the algorithms. On complex projects, large teams of programmers work together and share tasks.

At their origin, algorithms are logical systems that are as old as mathematics. “The expression comes from a Latinization of the name of Persian mathematician and astronomer Mohamed al-Khwarizmi, who produced famous works on algebra in the ninth century,” explains computer scientist Cristina Gomes Fernandes, who is a professor at IME-USP. They gained new impetus in the second half of the last century alongside the development of the computer, with which it was possible to create work routines for the machines. There are two reasons why algorithms are now so widely used in the real world and why they have become the basis of most complex software development: First, the increased processing power of computers has increased the speed at which complex tasks can be executed. Second, the advent of big data has made it cheaper to collect and store huge amounts of information, thereby enabling algorithms to identify patterns that are imperceptible to the human eye in a wide range of scenarios. Advanced manufacturing, which is known as Industry 4.0, promises to increase productivity by using artificial intelligence algorithms to monitor industrial plants in real time and make decisions on stock control, logistics, and maintenance.

The effect of the growing use of algorithms in computing was a boost to artificial intelligence, which is a field that was established in the 1950s and aims at developing mechanisms that are capable of simulating human reasoning. Through increasingly fast computations and the collection of data for statistical comparisons, computers can now modify their operations based on accumulated experience,



## Identifying Parasites

An IME-USP research project is being conducted in collaboration with UNICAMP's Laboratory of Image Data Science (LIDS) to improve the diagnosis of parasite infections using computer vision. Marcelo Finger, who is a computer scientist from IME, is testing an algorithm that can identify parasites by analyzing images of stool samples. “We have been able to identify 15 parasites in humans and some in animals, such as cattle, dogs, and cats,” he says. Diagnoses are currently obtained by

examining stool samples under a microscope. “A lab worker can usually analyze about six blades at a time. The aim is to automate this process,” says Finger. It seems simple; however, because algorithms operate by identifying patterns, any background noise creates an obstacle for the researchers. “It is one thing for the algorithm to be able to identify the parasite in a photo from a book; doing the same with an image in which the parasite is surrounded by dirt is quite another,” says the researcher.

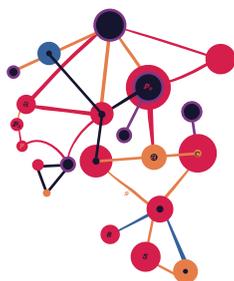
thereby improving their performance in a process that mimics learning.

Computers have proven capable of beating humans in many board games; this demonstrates how the field has evolved. In 1997, IBM's Deep Blue supercomputer succeeded for the first time in beating the world chess champion of the time, Gary Kasparov, who was from Russia. Capable of simulating approximately 200 million chess positions per second, the machine anticipated its opponent's decisions several moves ahead. However, this strategy was unsuccessful for *Go*, which is a Chinese board game, because there are too many possible moves at any time to anticipate—the number of possibilities exceeds the number of atoms in the universe. In March 2016, *Go* was finally defeated: the AlphaGo program, which was created by DeepMind, which is a subsidiary of Google, beat world champion Lee Sedol, who was from South Korea.

Instead of considering millions of possibilities, the program's algorithm used a more restricted strategy: By statistically analyzing data from previous matches between the game's best players, the program identified the most common and efficient moves, thereby resulting in a smaller set of variables, and was soon able to beat the human players. However, there was more to come. Last year, DeepMind developed a new program, namely, AlphaGo Zero, which outperformed the original AlphaGo. In the new program, the machine did not learn from humans, but from the previous versions of the program.

There are a growing number of practical applications for this type of technology. Artificial intelligence algorithms that were developed by computer scientist Anderson de Rezende Rocha, who is a professor at the Institute of Compu-

## Programmers should be aware of the implications of their work, says Nick Seaver, from Tufts University



there are files that are related to child abuse, the algorithm helps find them. “We exposed the robot to hours of pornographic videos from the internet to teach it what pornography is,” says Rocha. Then, to identify the presence of children, the algorithm needed to “watch” the videos of child abuse that were seized by the police. “This stage was carried out by police officers. Nobody at UNICAMP had access to this material,” he adds. Rocha says that these types of files were previously analyzed manually in most cases. “Automating the process makes it more efficient, giving the police more time and allowing them to examine more data.”

Many computer scientists use mathematical properties, theorems, and logic when working on algorithms, regardless of the immediate purpose

## Infinite commands

Numbers of lines of code in the algorithms that are used in various products and services

### SIMPLE IPHONE APP

10,000

### BOEING 787

14 million

### LARGE HADRON COLLIDER (LHC)

50 million

### FACEBOOK

62 million

### CAR SOFTWARE

100 million

### GOOGLE

SOURCE MCCANDLES, D.  
KNOWLEDGE IS BEAUTIFUL.  
HARPER COLLINS  
PUBLISHERS, 2014



Google's AlphaGo software beat South Korean Lee Sedol in a game of Go in 2016

of the application. In many scenarios, the only known algorithms are highly inefficient and do not perform well with large data volumes, for example, in the factorization of a number into its constituent primes (which is highly important in cryptography) or routing a welding robot through several weld points. There is little hope that efficient algorithms will be identified for these applications, which fall under unsolved problem of “P versus NP,” which is considered one of the greatest challenges in both computer science and mathematics.

Although there is more programming involved than basic science in the development of many of the algorithms that are used in everyday life, advances in knowledge are essential if new applications are to be explored in the future. Marcondes Cesar, who is from USP, is working on computer vision, which is a type of artificial intelligence that extracts information from images to simulate human vision. The technique is being explored in various industries, particularly in medical diagnoses. “Computer vision can detect anomalies more accurately and evaluate subtle details in magnetic resonance imaging, for example.”

The objective of the project, which is being carried out in partnership with the USP School of Medicine and the Children’s Institute of the

university’s teaching hospital, is to develop a mathematical model that can provide a more accurate analysis of the liver and brain in newborns. The models that are used to interpret magnetic resonance images are typically based on white adult males and have been developed in other countries, which can lead to inaccurate diagnoses in newborn babies in Brazil. However, the project’s success depends on several theoretical problems being solved first. “We do not yet know if we will be able to write an efficient algorithm. We are still studying properties based on graph theory,” he says, referring to the branch of mathematics in which the relations between objects of a specified set are studied by associating them to one another via structures that are called graphs.

The impact of algorithms has also been analyzed in other fields of knowledge. “Algorithms are already playing a moderating role. Google, Facebook, and Amazon have an extraordinary amount of power over what we are exposed to in culture today,” said Ted Striphas, who is a professor of the history of culture and technology at the University of Colorado, USA, and author of the book *Algorithmic Culture* (2015), which examines the influence of these online giants. American anthropologist Nick Seaver, who is a researcher at Tufts University, USA, is currently conducting ethnographic research and interviews with the creators of music recommendation algorithms for streaming services. His interest is in how these systems are designed to attract users and draw their attention and he is studying the interface between areas such as machine learning and online advertising. “The mechanisms that control attention and its technical mediations have become a subject of great interest. The formation of interest and opinion bubbles, as well as *fake news*, and political distractions, can be attributed to technologies designed to manipulate user attention,” he explains.

**R**ecommendation systems that are based on algorithms have become key players in the online entertainment industry. In an article that was published in the journal *ACM Transactions on Management Information Systems* in 2015, Mexican electronic engineer Carlos Gomez-Uribe described how the algorithms that are used by streaming service Netflix rank television series and movies according to the



Testing Uber's autonomous car prototype in San Francisco (USA)

individual profile of each user. The objective is to encourage customers to select a TV show to watch within 90 seconds of logging on—any longer than that and they tend to get frustrated and lose interest. The success of this ranking system gave Gomez-Uribe's career a boost and in 2017, he became head of algorithms and internet technology products at Facebook.

The influence and power of major internet companies do not depend solely on the creativity of their programmers. They are also linked to the huge volumes of data that have been accumulated and processed by their algorithms, which have generated highly valuable information. "What prevents another company from developing an application like Uber? This has already been do-

ne, in fact. But the traffic and customer behavior data that Uber has accumulated over time belongs only to them, and it is valuable," says Marcondes Cesar, who is from USP.

The recent Facebook user-data leak, which caused the value of the company to fall by US\$49 billion last month, revealed a vulnerability that was thought to be uncommon—algorithms that are used by Cambridge Analytica were able to access the behavioral data of 50 million Facebook users, which were subsequently used to influence political campaigns on social networks, including the Brexit vote and Donald Trump's ultimately successful bid to become president of the United States. The Facebook case is an example of the ethical challenges that are created by the widespread use of algorithms, although data misuse and abuse are only part of the problem. Data use has become as important for algorithms as the challenge of programming them. "Analyzing the characteristics of the data is fundamental to the construction of an algorithm; a mistake at this stage could lead to biases in the results," says Marcondes Cesar.

It is also common for algorithms to reproduce biases when they are based on human behavior. The Cloud Natural Language API, which is a tool that was created by Google that identifies the structure and meaning of texts through machine learning, has developed its own biases. A test by American website Motherboard demonstrated that when analyzing text to determine if it has a "positive" or "negative" sentiment, the algorithm classified statements such as "I'm a homosexual" and "I'm a gay black woman" as negative. "Programmers who create smart algorithms need to be aware that their work has social and politi-

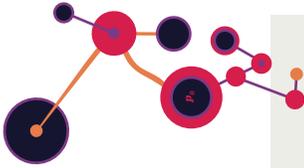
## CATTLE WEIGHT

Projeta Sistemas, which is a startup that is based in Vitória, Espírito Santo State, Brazil, has created an algorithm for assisting cattle farmers. The system, which is called Olho do Dono, uses 3D images to estimate the weights of cows. "The process of weighing cattle is very costly, time-consuming, and involves moving the animals around, which can cause stress and even weight loss," explains computer scientist Pedro Henrique Coutinho, who is the director

of the company. The software was developed based on computer vision techniques and associates the weights of the cattle with images that were captured by cameras. The system relies on a robust database. "We monitor the weighing of livestock on ranches throughout Brazil. Our algorithm is based on thousands of recorded images," says Coutinho. Development began in 2015 and the software will go to market in September.



System uses computer vision to estimate the weight of cattle



## Lost animals

CrowdPet is a smartphone application that was created by SciPet, which is a company that is based at UNICAMP, for helping find lost animals. The system uses an algorithm to compare pictures of lost pets that were provided by their owners with photos of animals on the streets that were taken by volunteers. “The application can match two images through visual recognition

methods and uses geolocation to locate where the photo of the lost animal was taken,” says Fabio Rogério Piva, who is the director of SciPet. The Animal Control Center in the municipality of Vinhedo, São Paulo, began using the application last year to register animals during welfare campaigns. SciPet has developed a prototype that can identify dogs and cats with 99% accuracy.

cal implications,” says Nick Seaver, from Tufts University. Various undergraduate and graduate computer science programs already offer classes that address computer ethics, including USP in Brazil and Harvard University and the Massachusetts Institute of Technology (MIT) in the US.

The transparency of advanced algorithms is another hot topic. The details of how these tools operate are often kept secret by developers. In some cases, the code is so complex that it is not possible to understand how the algorithm arrives at a decision and what its implications are. Systems such as these, which are opaque to external scrutiny, are known as “black box algorithms.” The debate has gained momentum after research into an experimental tool, namely, Correctional Offender Management Profiling for Alternative Sanctions (COMPAS), which is used in the US legal system to make sentencing recommendations and even to predict the risk that a defendant will reoffend. The study, which was conducted by the ProPublica organization in 2016, revealed that the COMPAS system is 77% more likely to classify black defendants as possible reoffenders than whites. Northpointe, which is the private company that created the algorithm, declined to share the code. “Algorithms used by public bodies should not be created or developed without the participation of public managers and administrators, as they are not neutral,” says Sérgio Amadeu da Silveira, who is a researcher at the Center for Engineering, Modeling, and Applied Social Sciences at the Federal University of ABC (UFABC).

In 2017, Kate Crawford, who is the head of research at Microsoft Research, and Meredith Whittaker, who is the leader of Google’s Open Research Group, founded the AI Now Institute, which is an organization that is dedicated to understanding the social implications of artificial

intelligence. Based at New York University, USA, the institute’s approach involves computer scientists, lawyers, sociologists, and economists. In October, it released a report that offered guidelines on the use of artificial intelligence algorithms. One recommendation was that public agencies such as those that are responsible for criminal justice, healthcare, welfare, and education should not use systems whose algorithms are not well known. According to the document, black-box algorithms should be subject to public auditing and validation tests to implement corrective mechanisms when necessary.

**A**nother objective of artificial intelligence algorithms is to free human beings from repetitive tasks—and there is frequent debate over the implications of AI software on the labor market. “The Future of Employment,” which is a report that was published in 2013 by economists Carl Frey and Michael Osborne from the Oxford Martin School, UK, estimated that sophisticated algorithms could soon replace 140 million professional jobs worldwide. The paper specifies examples, such as the increasing automation of decision-making in the financial market and even the impact on the work of software engineers—machine learning and algorithms can improve and accelerate various programming tasks. “Procedural intellectual activities that involve repetitive tasks, such as translating documents, have a great chance of one day being executed by computer algorithms,” says Sérgio Amadeu, who is from UFABC. It is important that we discuss the side effects of artificial intelligence, according to Marcondes Cesar, who is from USP; however, for now, they are far outweighed by the remarkable contributions that are made by these algorithms to the solution of problems of many types. ■