UC SANTA BARBARA

THE Current

December 5, 2013
Julie Cohen

Algorithms Are Watching

In his prescient novel “1984,” English author George Orwell predicted a future that bears an uncanny resemblance to current reality — except for a simple twist. Rather than Big Brother watching, today we have big brothers — plural — in the form of huge Internet companies such as Google, Facebook and LinkedIn, which log every keystroke.

To paraphrase Orwell, algorithms are watching.

Almost 30 years beyond 1984, these step-by-step mathematical bits so essential to computer code rule modern life. Calculations once performed manually by statisticians and mathematicians, for example, are now done by algorithms. Financial institutions use algorithms to make 70 percent of stock trades, and mortgage and insurance companies rely on them to calculate risk.

This is not news to the many UC Santa Barbara professors whose research involving algorithms ranges from their use in data mining and the analysis of social networks and online communities to revealing how information moves within the massively large social graphs that capture relationships, transactions and social interactions among users.

“In every aspect of life, as long as an area generates data, an algorithm is required,” said Xifeng Yan, an associate professor in UCSB’s Department of Computer Science.

Algorithms, which can be formalized mathematically, are simply a finite set of precise instructions. The origin of the term comes from the name of Persian

mathematician Abu Abdullah Muhammad Ibn Musa al-Khwarizmi who introduced the algorithm concept in his ninth-century treatise on algebra.

According to Ben Zhao, associate professor of computer science and co-director of UCSB's Sand Lab, many of the classical algorithms have not fundamentally changed much in the past 10 to 20 years. What has changed, however, is the way they are applied. "We used to rely on dramatic advances in mathematical algorithms to accomplish something really fast," Zhao said. "Today, folks are much more interested in whether algorithms can lend themselves to parallel processing on an extremely large scale."

Parallel processing is the ability to carry out multiple operations or tasks simultaneously. In computing, this means using multiple CPUs or multicore processors to compute data quickly or make programs run faster.

While financial or GPS applications seem relatively benign, the use of algorithms can get downright creepy when it comes to the Internet. Amazon, Netflix and Pandora use complex algorithms to make recommendations based on what similar people like, and Facebook and Google use them to cull pertinent information from personal emails and Internet searches in order to provide unsolicited user-specific advertising.

"It's not just Facebook; it's Facebook and all its partners," Zhao said. "Facebook has a system called Facebook Connect that allows you to use your Facebook credentials to log into third-party websites and applications. Behind that is an agreement that says Facebook and its other partners will share information on your activities online with those third-party sites."

Most people don't log out of Facebook each time they visit, Zhao noted, and this is why surfing the Web for, say, a house on Zillow or Trulia produces a related Facebook ad identifying another home in the same area. "It is, in fact, very, very difficult to actually turn that off because tracking is so pervasive and so complete," Zhao said. "To actually disentangle or completely avoid tracking is far more challenging than people realize."

Staying under the tracking radar is next to impossible because so many different kinds of tracking systems are built into the basic structure of the Web browser and the computer operating system. "The Internet can infer things about you even if you turn off all actual data flow," Zhao explained. "Even if you're not logging onto

Facebook, Facebook can track you when your friends post pictures and tag you, so people know where you've been even if you don't post anything."

Zhao acknowledged the potential for violations of privacy but suggested they may not be as nefarious as one might think. "In most cases, everything is at an aggregate level," he said. "They aren't looking at you as an individual, and most of us are not interesting enough to become individual targets."

Yan concurred, adding, "There may be privacy issues, but on the positive side, your search will be utilized to help improve search results for other queries. It's important to remember that whenever you are connected, there is intelligent feedback to the system, which current learning algorithms use to improve ranking results."

As algorithms become more sophisticated, their influence over our lives increases exponentially. "Much of what we see today is customized for us because of all the data tracking done by Google and Facebook," Zhao said. "They customize everything for you because of what you've already done." He and other researchers are trying to understand just how much this impacts us and to what extent data tracking influences what we see on a daily basis.

Eli Pariser, co-founder of the Internet news site Upworthy, coined the term "filter bubble" to describe how invisible algorithmic editing selectively guesses the information that users would like to see based on their past click behavior, search history and location. The results, however, can be quite one-sided. "There's a sense of being placed in this echo chamber — a term people use a lot," Zhao said. "Whatever you already believe, whatever you already like tends to get reflected back at you. If you're a hardcore liberal Democrat, for instance, Google shows you news from blue-leaning states. If you're a conservative Republican, then you get everything that's slanted that way."

Many algorithms try to mimic human learning processes, and in certain circumstances they operate more efficiently and effectively. "For simple rule-based tasks, algorithms can outdo humans anytime, partly because they can compute and access massive quantities of data quickly," said Subhash Suri, chair of UCSB's Department of Computer Science and director of the Geometric Computing Center.

Yet for many seemingly simple functions, the best algorithms still cannot match the human brain. "It is remarkable how good humans are at some things, which are dauntingly difficult to teach an algorithm," Suri added. "Once I see your face, for

example, I will remember it and recognize it tomorrow after a quick glance, even from a different angle or under a different light. When I walk into a new building, I know to use the door, not the window. I don’t have to relearn how to walk every day.”

Translating such seemingly innate learning processes into something a machine can master is the ultimate algorithmic challenge. “Designing algorithms that can sense and interact with the physical and dynamic world is intellectually exciting with significant practical impact,” Suri said.

Nowhere is this more apparent than with personalized robots, which Suri predicts will one day be as common a household item as the personal computer. It’s a matter of creating algorithms that match the way humans learn.

While a robot has a camera that can see, it doesn’t have an understanding of physics, so it doesn’t know the wall is a barrier, Suri explained. “Teaching robots human skills is essentially making us go back and think about how we learned them in the first place,” he said.

Other areas where algorithms will shape our future include algorithmic medicine, which Suri sees as something akin to the kind of recommendation systems used by online merchants. “It’s a controversial point of view,” he said, “but there is very little your doctor does that is not automated.

“He takes your temperature, he takes your pulse, he takes a bit of your history and he asks you where your stomach hurts,” Suri explained. “It is rule-based. If somebody needs surgery or has a strange heart problem, they should go to the experts. But 90 percent of visits to the doctor are for routine things that algorithms can do more cost-effectively, probably more comfortably and much more reliably.”

While applications such as algorithmic medicine may seem more akin to science fiction than reality, the truth is that algorithms existed long before al-Khwarizmi codified them. According to Suri, DNA, which determines every individual’s physiology, anatomy and function, is itself an algorithm. “In some sense, we have an algorithm running in us,” he said. “The genome is an algorithm, and one of the great challenges is to figure out what this algorithm is doing and to what extent can we modify and learn from it. Evolution itself is an algorithm — on an even grander scale.”

About UC Santa Barbara

The University of California, Santa Barbara is a leading research institution that also provides a comprehensive liberal arts learning experience. Our academic community of faculty, students, and staff is characterized by a culture of interdisciplinary collaboration that is responsive to the needs of our multicultural and global society. All of this takes place within a living and learning environment like no other, as we draw inspiration from the beauty and resources of our extraordinary location at the edge of the Pacific Ocean.