The Economist

Briefing Algorithms | Business by Numbers

Consumers and companies increasingly depend on a hidden mathematical world

ALGORITHMS sound scary, of interest only to dome-headed mathematicians. In fact they have become the instruction manuals for a host of routine consumer transactions. Browse for a book on Amazon.com and algorithms generate recommendations for other titles to buy. Buy a copy and they help a logistics firm to decide on the best delivery route. Ring to check your order's progress and more algorithms spring into action to determine the quickest connection to and through a call-centre. From analysing credit-card transactions to deciding how to stack supermarket shelves, algorithms now underpin a large amount of everyday life.

Their pervasiveness reflects the application of novel computing power to the age-old complexities of business. "No human being can work fast enough to process all the data available at a certain scale," says Mike Lynch, boss of Autonomy, a computing firm that uses algorithms to make sense of unstructured data. Algorithms can. As the amount of data on everything from shopping habits to media consumption increases and as customers choose more personalisation, algorithms will only become more important.

Algorithms can take many forms. At its core, an algorithm is a step-by-step method for doing a job. These can be prosaic—a recipe is an algorithm for preparing a meal—or they can be anything but: the decision-tree posters that hang on hospital walls and which help doctors work out what is wrong with a patient from his symptoms are called medical algorithms.

This formulaic style of thinking can itself be a useful tool for businesses, much like the rigour of good project-management. But computers have made algorithms far more valuable to companies. "A computer program is a written encoding of an algorithm," explains Andrew Herbert, who runs Microsoft Research in Cambridge, Britain. The speed and processing power of computers mean that algorithms can execute tasks with blinding speed using vast amounts of data.

Some of these tasks are more mechanistic than others. For instance, people often make mistakes when they key in their credit-card numbers online. With millions of transactions being processed at a time, a rapid way to weed out invalid numbers helps to keep processing times down. Enter the Luhn algorithm (see above), named after its inventor, Hans Luhn, an IBM researcher. The numbers on a credit card identify

Do try this at home
The Luhn algorithm for validating credit cards

1. Take original number	3	7	7	9	5	6	5	7	0	9	4	4	7	2	6
2. Reverse number	6	2	7	4	4	9	0	7	5	6	5	9	7	7	3
3. Double every second number		4		8		18		14		12		18		14	
4. Add numbers up*	6	4	7	8	4	9	0	5	5	3	5	9	7	5	3

Answer =80. (Valid if divisible by ten)

Source:www.darkcoding.net   *If more than nine, subtract nine

the card type, the issuer and the user's account number. The last number of all is set to ensure that the Luhn algorithm produces a figure divisible by ten. If it is, the card number has been properly entered and the processing can go ahead.

The Luhn algorithm performs a simple calculation. But the real power of algorithms emerges when they are put to work on much more complex problems. As far as most businesses are concerned, these problems typically fall into two types: improving various processes, such as how a network is configured and a supply chain is run, or analysing data on things such as customer spending.

UPS uses algorithms to help deliver the millions of packages that pass through its transportation network every day in the most efficient way possible. The simplest routes are easy to draw up. If a driver has only three destinations to visit, he can take only six possible routes. But the number of possible routes explodes as the destinations increase. There are more than 15 trillion, trillion possible routes to take on a journey with just 25 drop-off points—and an average day for a UPS driver in America involves 150 destinations. The picture is further complicated by constraints such as specified drop-off and pick-up times for drivers or runway lengths and noise restrictions for aircraft. "Algorithms provide benefits when the choices are so great that they are impossible to process in your head," says UPS's Jack Levis.

Go here, go there

Solving this "travelling-salesman problem" means a lot to UPS. For its fleet of aircraft in America, the company uses an algorithm called VOLCANO (which stands for Volume, Location and Aircraft Network Optimiser). Developed jointly with the Massachusetts Institute of Technology (MIT), it is used by three different planning groups within

UPS—one to plan schedules for the following four to six months, one to work out what kind of facilities and aircraft might be needed over the next two to ten years, and one to plan for the peak season between Thanksgiving and Christmas. Getting the scheduling wrong imposes a heavy cost: flying half-empty planes or leasing extra aircraft is an expensive business. UPS reckons that VOLCANO has saved the company tens of millions of dollars since its introduction in 2000.

Logistics firms are far from the only ones working on "optimisation" algorithms. Telecoms operators use algorithms to establish the quickest connections for phone calls through their networks or to retrieve web pages speedily from the internet. Manufacturers and retailers use them to fine tune their supply chains. Call centres decide where to place an incoming call, based on things such as the customer's location, the length of queues that operators have to deal with and the reason for people calling.

Jeff Gordon, who looks after innovation for Convergys, a call-centre operator, says that the efficiency of algorithms is as crucial to his industry as the quality of call agents: "If you get the algorithm wrong and put customers into the wrong hands you degrade the experience. No one likes being handed off to someone else."

The most powerful algorithms are those that cope with continual changes. The delivery schedules for online grocers have huge "feedback loops" in which the delivery times chosen by customers affect the routes that vans take, which in turn affects the choice of delivery slots made available to customers. UPS is working on a real-time algorithm for its drivers that can recalibrate the order of deliveries on the fly, in much the same way that satellite-navigation systems in cars adjust themselves if a driver chooses to ignore a suggested route.

In the world of the internet, operators are looking at ways of marrying up the algorithms that find the shortest path through a network and those that control the speed with which information flows. At the moment, the routing algorithm does not talk to the flow-control algorithm, which means paths do not change even when there is congestion. According to Marc Wennink, a researcher at Britain's BT, combining the algorithms would mean that tasks such as downloading files could become much more resilient to network disruption. It would also allow BT to make better use of its existing network capacity.

Airports also have a keen interest in dynamic algorithms. Passengers at London's Heathrow and other congested airports often sit in a long queue of planes waiting near the runway to depart. Delays happen because air-traffic controllers need to leave a safety margin between aircraft as they take off. This margin depends on the size and speed of an aircraft, and re-ordering the queue can minimise the delay before all the planes get into the air (mathematicians call this the departure problem). Air-traffic controllers have always reordered planes in the departure queue manually, but researchers are working on algorithms that would be more efficient.

Just as optimisation algorithms come in handy when people are swamped by vast numbers of permutations, so statistical algorithms help firms to grapple with complex datasets. Dunnhumby, a data-analysis firm, uses algorithms to crunch data on customer behaviour for a number of clients. Its best-known customer (and majority-owner) is Tesco, a British supermarket with a Clubcard loyalty-card scheme that generates a mind-numbing flow of data on the purchases of 13m members across 55,000 product lines. To make sense of it all, Dunnhumby's analysts cooked up an algorithm called the rolling ball.

It works by assigning attributes to each of the products on Tesco's shelves. These range from easy-to-cook to value-for-money, from adventurous to fresh. In order to give ratings for every dimension of a product, the rolling-ball algorithm starts at the extremes: ostrich burgers, say, would count as very adventurous. The algorithm then trawls through Tesco's purchasing data to see what other products (staples such as milk and bread aside) tend to wind up in the same shopping baskets as ostrich burgers do. Products that are strongly associated will score more highly on the adventurousness scale. As the associations between products become progressively weaker on one dimension, they start to get stronger on another. The ball has rolled from one attribute to another. With every product categorised and graded across every attribute, Dunnhumby is able to segment and cluster Tesco's customers based on what they buy.

Where to put the biscuits

The rolling-ball algorithm is in its fourth version. Refinements occur every year or two, to add new attributes or to tweak the maths. All these data then feed into a variety of decisions, such as the ranges to put into each store and which products should sit next to each other on the shelves. "All this sophisticated data analysis and it comes down to where you put the biscuits," laments Martin Hayward, director of consumer strategy at Dunnhumby.

Fraud detection has a touch more glamour to it. **Chicago-based SPSS Inc.**, another data-analysis firm, uses algorithms to scrutinise customer data and to build propensity scores that predict how people will behave. One of its clients is ClearCommerce, which provides payment-processing services to online merchants. **SPSS** helped ClearCommerce to build a system that looks at a customer's past transactions and learns what hints at fraud—it might be the amount of money being spent, the shipping details and the time of day, and so on. Transactions then get a fraud-propensity score based on these characteristics; merchants decide which scores should ring alarm bells and how to respond.

Algorithms are most commonly associated with internet-search engines. "The tussle between MSN, Google and Yahoo! is about whose algorithm produces the best results to a query," observes Microsoft's Mr Herbert. Ask.com, another search engine, has even tried to popularise the term in an advertising campaign. Few other types of companies are so obviously dependent on algorithms for success, but the role that they play is rising in importance for two reasons.

The first is the sheer amount of data that is now available to companies. The information floodwaters are rising everywhere. Smart meters give utility firms data on consumption patterns inside households. Digital media will make it easier for firms such as Dunnhumby to see how what people read online and watch on television affects what they buy.

Online shopping means that internet merchants now know what customers are browsing as well as buying. Search engines are mining their own information on the relationship between queries and clickthroughs so as to improve their ranking algorithms. "For the first time in business history there is more information than many organisations' capacity to deal with it," says Dunnhumby's Mr Hayward. Algorithms are a way to cope.

The second reason why algorithms are becoming more important is that companies inevitably want to use all this new data to do more complicated things. In particular, they want to respond to each customer in a personalised way. Tesco does this by using its analysis to tailor direct-marketing offers to each Clubcard member. As well as segmenting its customers on how they

live, the data also enable the supermarket rapidly to spot shifts in their consumption patterns (caused by children going to university, say). Tesco's response rates to such targeted marketing stands at 10-20%, against an industry average of only around 1%.

Convergys wants to bring more real-time data to the operation of call-centres. Mr Gordon gives the example of a customer who calls an electricity utility from an area that has suffered a power failure and, because of where they are speaking from, is automatically put through to an operator who can deal with his queries. Such algorithms help firms to tease simplicity from complexity.

Algorithms are not for everyone. Some companies will always generate more data than others, of course: retailers, utilities and telecoms firms process many more transactions than house insurers, whose deals tend to happen once a year. Some will also be more focused than others on how algorithms can shave costs or maximise capacity. Firms that enjoy high margins and strong demand are going to be less worried about the efficiency of their supply chains, says Hau Lee, of Stanford Graduate School of Business.

Rocket science for non-boffins

What is more, lots of things have to fall into place for algorithms to work. They tend to be highly complex: it is not easy to find people with the right skills to develop and refine them. The systems within which the algorithms run—the user interface—need to be intuitive to non-boffins. **"This is rocket science but you don't have to be a rocket scientist to use it," says Jack Noonan, boss of SPSS.** The inputs have to be right. One UPS planning model routed all the packages in the system through Iowa, which perplexed everyone until they found an error in the data that made it appear to be free to send packages via Iowa. The algorithm was right, in other words, but the data were wrong. **Mr Noonan says that SPSS's "secret sauce" lies in its ability to deal with missing or unreliable data, rather than the algorithms themselves.**

Above all, human judgment still has a role—a point perhaps reinforced by the recent performance of algorithmically driven quantitative funds in the financial markets. In fraud detection, for example, algorithms can eliminate the majority of transactions that are above suspicion but a human is still best placed to analyse the dodgy ones. Dunnhumby is trying to overlay attitudinal research on top of purchasing data to understand why people buy things as well as what they buy. Even so, Autonomy's Mr Lynch is convinced that algorithms are on the march. Algorithms process data to arrive at an answer. The more data they can process the more accurate the answer. For that reason, he says, "they are bound to take over the world".

SPSS