**Clever computers**

The dawn of artificial intelligence

Powerful computers will reshape humanity's future. How to ensure the promise outweighs the perils

May 9th 2015 | From the print edition

"THE development of full artificial intelligence could spell the end of the human race," Stephen Hawking warns. Elon Musk fears that the development of artificial intelligence, or AI, may be the biggest existential threat humanity faces. Bill Gates urges people to beware of it.
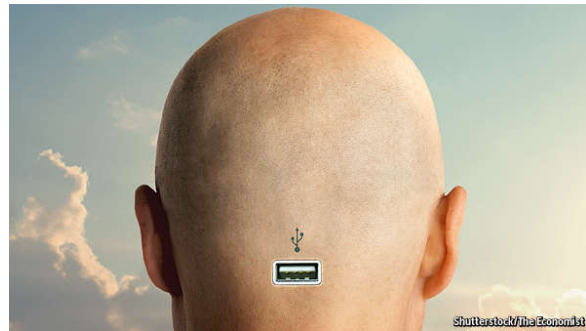
Dread that the abominations people create will become their masters, or their executioners, is hardly new. But voiced by a renowned cosmologist, a Silicon Valley entrepreneur and the founder of Microsoft—hardly Luddites—and set against the vast investment in AI by big firms like Google and Microsoft, such fears have taken on new weight. With supercomputers in every pocket and robots looking down on every battlefield, just dismissing them as science fiction seems like self-deception. The question is how to worry wisely.

You taught me language and...

The first step is to understand what computers can now do and what they are likely to be able to do in the future. Thanks to the rise in processing power and the growing abundance of digitally available data, AI is enjoying a boom in its capabilities (see article (http://www.economist.com/news/briefing/21650526-artificial-intelligence-scares-peopleexcessively-so-rise-machines)). Today's "deep learning" systems, by mimicking the layers of neurons in a human brain and crunching vast amounts of data, can teach themselves to perform some tasks, from pattern recognition to translation, almost as well as humans can. As a result, things that once called for a mind—from interpreting pictures to playing the video game "Frogger"—are now within the scope of computer programs. DeepFace, an algorithm unveiled by Facebook in 2014, can recognise individual human faces in images 97% of the time.

Crucially, this capacity is narrow and specific. Today's AI produces the semblance of intelligence through brute number-crunching force, without any great interest in approximating how minds equip humans with autonomy, interests and desires. Computers do not yet have anything approaching the wide, fluid ability to infer, judge and decide that is associated with intelligence in the conventional human sense.

Andy Potts

Briefing: AI is improving rapidly, but remains more useful than terrifying (http://www.economist.com/news/briefing/21650526-artificial-intelligence-scares-peopleexcessively-so-rise-machines)

Yet AI is already powerful enough to make a dramatic difference to human life. It can already enhance human endeavour by complementing what people can do. Think of chess, which computers now play better than any person. The best players in the world are not machines however, but what Garry Kasparov, a grandmaster, calls "centaurs": amalgamated teams of humans and algorithms.

Such collectives will become the norm in all sorts of pursuits: supported by AI, doctors will have a vastly augmented ability to spot cancers in medical images; speech-recognition algorithms running on smartphones will bring the internet to many millions of illiterate people in developing countries; digital assistants will suggest promising hypotheses for academic research; image-classification algorithms will allow wearable computers to layer useful information onto people's views of the real world.

Even in the short run, not all the consequences will be positive. Consider, for instance, the power that AI brings to the apparatus of state security, in both autocracies and democracies. The capacity to monitor billions of conversations and to pick out every citizen from the crowd by his voice or her face poses grave threats to liberty.

And even when there are broad gains for society, many individuals will lose out from AI. The original "computers" were drudges, often women, who performed endless calculations for their higher-ups. Just as transistors took their place, so AI will probably turf out whole regiments of white-collar workers. Education and training will help and the wealth produced with the aid of AI will be spent on new pursuits that generate new jobs. But workers are doomed to dislocations.

Surveillance and dislocations are not, though, what worries Messrs Hawking, Musk and Gates, or what inspires a phalanx of futuristic AI films that Hollywood has recently unleashed onto cinema screens. Their concern is altogether more distant and more apocalyptic: the threat of autonomous machines with superhuman cognitive capacity and interests that conflict with those of *Homo sapiens*.

Such artificially intelligent beings are still a very long way off; indeed, it may never be possible to create them. Despite a century of poking and prodding at the brain, psychologists, neurologists, sociologists and philosophers are still a long way from an understanding of how a mind might be made—or what one is. And the business case for even limited intelligence of the general sort—the sort that has interests and autonomy—is far from clear. A car that drives itself better than its owner sounds like a boon; a car with its own ideas about where to go, less so.

...I know how to curse

But even if the prospect of what Mr Hawking calls "full" AI is still distant, it is prudent for societies to plan for how to cope. That is easier than it seems, not least because humans have been creating autonomous entities with superhuman capacities and unaligned interests for some time. Government bureaucracies, markets and armies: all can do things which unaided, unorganised humans cannot. All need autonomy to function, all can take on life of their own and all can do great harm if not set up in a just manner and governed by laws and regulations.

These parallels should comfort the fearful; they also suggest concrete ways for societies to develop AI safely. Just as armies need civilian oversight, markets are regulated and bureaucracies must be transparent and accountable, so AI systems must be open to scrutiny. Because systems designers cannot foresee every set of circumstances, there must also be an off-switch. These constraints can be put in place without compromising progress. From the nuclear bomb to traffic rules, mankind has used technical ingenuity and legal strictures to constrain other powerful innovations.

The spectre of eventually creating an autonomous non-human intelligence is so extraordinary that it risks overshadowing the debate. Yes, there are perils. But they should not obscure the huge benefits from the dawn of AI.

 From the print edition: Leaders