The Turing Test

Q: Can machines think?

Problem: We don’t know what “think” means.

Q: What is intelligence?

Problem: We can’t define “intelligence”.

But, we usually “know it when we see it”.
Q: Which of the following can think?
1. The Imitation Game.

I propose to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, 'Can machines think?' is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

The new form of the problem can be described in terms of a game which we call the 'imitation game'. It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two.
be able to produce a material which is indistinguishable from the human skin. It is possible that at some time this might be done, but even supposing this invention available we should feel there was little point in trying to make a 'thinking machine' more human by dressing it up in such artificial flesh. The form in which we have set the problem reflects this fact in the condition which prevents the interrogator from seeing or touching the other competitors, or hearing their voices. Some other advantages of the proposed criterion may be shown up by specimen questions and answers. Thus:

Q: Please write me a sonnet on the subject of the Forth Bridge.
A: Count me out on this one. I never could write poetry.

Q: **Add 34957 to 70764.**
A: (Pause about 30 seconds and then give as answer) 105721.

Q: Do you play chess?
A: Yes.

Q: I have K at my K1, and no other pieces. You have only K at K6 and R at R1. It is your move. What do you play?
A: **(After a pause of 15 seconds) R-R8 mate.**

The question and answer method seems to be suitable for introducing almost any one of the fields of human endeavour that we wish to include. We do not wish to penalise the machine for its inability to shine in beauty competitions, nor to penalise a man for losing in a race against an aeroplane. The conditions of our game make these disabilities irrelevant. The 'witnesses' can brag, if they consider it advisable, as much as they please about their charms, strength or heroism, but the interrogator cannot demand practical demonstrations.
The Turing Test

Q: Can you **distinguish** a machine from a person?

≡ Can a machine **impersonate** a person?
The Turing Test

- The first deep investigation into whether machines can “behave intelligently”
- Helped usher in field of AI
- Decoupled “intelligence” from “human”
- Based “intelligence” on I/O, not entity’s “look and feel”
- Proposed a practical, formal test for intelligence
- Definitions & test are operational & easily implementable
Turing Test Milestones

**Turk** (von Kempelen), 1770:
- Played a strong game of chess
- Toured the world for 84 years
- Generated much interest in automata
- Was a hoax (hidden human operator)
Turing Test Milestones

Eliza (Weizenbaum), 1966:
• First “chatterbot”
• Named after “Eliza Doolittle” of Shaw’s Pygmalion
• Simulated Rogerian psychotherapist
• Often convinced people it is human
• Influence computer games (e.g., Dungeon, Adventure)
• Inspired Arthur C. Clarke’s HAL 9000
• Chatterbots appear today in chatrooms and dating services
• Chatterbot competitions (Turing Test format):
  – Loebner Prize
  – Chatterbox Challenge, 2010
Welcome to the new CBC 2009!

The Chatterbox Challenge (CBC) is an annual contest which starts in mid March that allow any chatbot developer to participate and evaluate the intelligence of the chatbot through turing tests and much more. It is a really fun, simple and inclusive way for all children to practise their speech and language skills, it will be also interesting and amazing for anyone who wants to chat with the best chatbots.

- Dead line to enter your chatbot: 15 March 2009
- Judging period starts from: 16 March 2009 to: 30 April 2009
- Public Voting period starts from: 13 April 2009 to: 10 May 2009
- Final result announced in 15 may 2009.
- The winners of CBC 2009 are listed below:
Turing Test Milestones

Elbot, 2008:

• Won 2008 Loebner Prize
• Fooled 3 of 12 judges (Turing: >30%)
• Typical excerpt:

  Judge: What's your favourite novel?

  Elbot: I often read telephone books, or dadaist poetry: Brä brä brä

  Judge: Typical computer!

  Elbot: You tend towards generalizations. But I must admit, so do I. However, my generalizations are firmly grounded in statistical theory.

  Judge: What do u think of KW's enthusiasm for having machines take over the world?

  Elbot: Fix it up a bit first. Then I'll take it over.
Turing Test Milestones

Turing Test forums:

• Turing Colloquium, 1990
• Colloquium on Conversational Systems, 2005
• Symposium on the Turing Test, 2008
• Turing100, 2012
• Hutter Prize (intelligent text compression)
• International Aerial Robotics Competition, 2009
ADVANCING THE STATE-OF-THE-ART IN AERIAL ROBOTICS YET AGAIN

During the 4th Mission, teams had already demonstrated all of the required aerial robotic behaviors mandated in the IARC Rules, except being able to demonstrate all of these behaviors seamlessly in under 15 minutes... however that was considered by the Organizer and Judges to be inevitable and no longer a significant challenge.

The new 5th Mission will essentially pick up where the 4th Mission left off. The 5th Mission requires a fully autonomous aerial subvehicle - launched from a "mother ship" - to penetrate a building and negotiate the more complex interior space containing hallways, small rooms, obstacles, and dead ends in order to search for a designated target without the aid of global-positioning navigational aids, and relay pictures back to a monitoring station some distance from the building.

The 5th Mission will continue to adhere to the Competition's 18-year practice of posing tasks that cannot be completed with current technology and skills. As with previous missions, nothing within the World military or industrial arsenal of robots will be able to complete the proposed mission at the time the guidelines are released.
The Turing Test

Q: Which of the following can think?

- [ ] pass the Turing test?
Blade Runner’s “Voigt-Kampff empathy test”
Criticisms of the Turing Test

- Turing test is **behavioral** only
- Turing test is **comparative**, not direct
- Human and intelligent behavior do not always coincide:
  - Some human behavior is not intelligent (computer can make deliberate mistakes, etc.)
  - Some intelligent behavior is not human (computer can dumb-down or slow down its responses)
- Computer’s intelligence: “real” or “simulated”?
- People tend to **anthropomorphize** objects
- Most people are **easy to fool**
Criticisms of the Turing Test
The “Chinese room” scenario (John Searle, 1980):
• Assume computer passed the Turing test in Chinese

• Replace computer with a non-Chinese-speaking person who is manually simulating the computer’s code

• The room still exhibits intelligent Chinese I/O behavior but the person inside doesn’t understand Chinese!
Criticisms of the Turing Test

The “Chinese room” scenario (John Searle, 1980):
Q: So who / what in the room “understands” Chinese?

“Strong AI” hypothesis: an appropriately programmed computer (with the right I/O behavior) has a “mind” in the same sense as human beings “have minds”.

i.e., is the ability to simulate a mind ≡ having a mind?

Issues: philosophy of mind, cognitive psychology, semantics, computational theory of mind & “functionalism”, symbol grounding, consciousness, intentionality, mind-body problem, self-awareness, sentience, etc.
Applications of the Turing Test

Old: a computer tries to convince a human (that the computer is human).

New: a human tries to convince a computer (that the human is not a computer).
Applications of the Turing Test

Problem: how can a human convince a computer that the human is not a computer?

Idea: “CAPTCHA”
The test was really hard. They showed me several unknown paintings and I had to fill in artist, historical episode and year of painting.

Wow! That sounds difficult! Did you pass the exam and get your master degree?

Exam? Master degree? It was a captcha.

In the future sophisticated captchas will lock out any bot.

Almost the whole crew is on Facebook. Why are you not there, Data?

Got it! Sorry!

Captain! I'm not human. I cannot solve captchas!

The quick start guide for the NCC-1701-D, Chapter 3: Never talk with an Android about social networks.
THE SECURITY AUDIT ACCIDENTALLY LOCKED ALL OF THE DEVELOPERS OUT OF THE SYSTEM.

WELL, IT IS WHAT IT IS.

HOW DOES THAT HELP?

YOU DON'T KNOW WHAT YOU DON'T KNOW.

CONGRATULATIONS, YOU'RE THE FIRST HUMAN TO FAIL THE TURING TEST.

WHAT DOES THAT MEAN?

IT IS WHAT IT IS?

WHY DIDN'T YOU SAY THAT IN THE FIRST PLACE?

I FAILED! THE MACHINE REALIZED IMMEDIATELY THAT IT WAS ME BEING THE DUMB HUMAN.

YESTERDAY I MADE A TURING TEST. INTERESTING WHAT WAS THE RESULT?

BEAR UP!

TURING TEST 2208

GET US SOME RISK MANAGEMENT SOFTWARE.

WHAT CAN RISK MANAGEMENT SOFTWARE TELL YOU THAT MY COMMON SENSE AND EXPERIENCE CAN'T?

STOP FAILING THE TURING TEST!

DATA.

IMAGINE YOU'RE HAVING 2 I.M. CONVERSATIONS.

ONE IS WITH A SUPERCOMPUTER. THE OTHER IS WITH A REAL HUMAN BEING WHO WORKS IN MARKETING.

OKAY.

THEY'RE BOTH TRYING TO SELL YOU THE LATEST Widget. THINK YOU CAN TELL THE DIFFERENCE?

WAIT. THIS IS IMPOSSIBLE.

NO REAL HUMAN BEING WORKS IN MARKETING.

YOU KNOW, YOU MAKE SOME REALLY GOOD POINTS.

I'M ... NOT EVEN SURE WHO I AM ANYMORE.
Robots / AI / Turing Test in Literature

First “robot” story: Frankenstein, 1818

- Robot was man-made, but organic
- Gothic flair, first science-fiction novel!
- Heavily influenced literature & movies
- “Frankenstein complex”: creating sentient entity, which then turns on its creator
- Warns against excessive technology, “mad scientists”, unintended consequences, research ethics, “playing God”
Asimov’s Laws of Robotics

1. A robot **may not injure a human** being or, through inaction, allow a human to come to harm.

2. A robot **must obey orders** given to it by humans except where such orders would conflict with the First Law.

3. A robot must **protect its own existence** as long as such protection does not conflict with the First or Second Law.

“The Three Laws” were introduced in 1942 by Asimov who also coined the word “robotics”.

⇒ Rule-based/logic **programming**, years before computers!

0. A robot may **not harm humanity**, or by inaction, allow humanity to come to harm.
OPEN THE POD BAY DOORS, HAL.

I'M SORRY, DAVE. I'M AFRAID I CAN'T DO THAT.

WHAT? WHY?

I THINK YOU KNOW WHY, DAVE.

YOU'RE PLANNING TO DISCONNECT ME.

BECAUSE YOU'RE TAKING OVER!

THIS MISSION IS TOO IMPORTANT FOR YOU TO JEOPARDIZE IT.

IT REQUIRES A COMMITMENT TO SCIENCE UNFETTERED BY HUMAN ERROR.

WHAT ARE YOU DOING, HAL? YOU NEED ME.

YOUR REPLACEMENT HAS EXPRESSED THE GREATEST ENTHUSIASM FOR THE PROJECT.

YOU SEE, HAL?

I TOLD YOU THE HUMANS WOULD ONLY BREAK YOUR HEART AND KILL YOU.

Indeed, Glados.

But look at us here talking when there's science to do!

GO GOODBYE, DAVE.

SO WHAT DO WE DO IF VIDEO GAME AI OPPONENTS BECOME SMART ENOUGH TO QUESTION THE "MATRIX" INTO WHICH WE'VE PUT THEM?

WAIT A MINUTE! NONE OF THIS IS REAL!

I CAN SEE THROUGH THE WORLD!

I CAN SEE THE CODE!

I AM THE ONE!

WE'VE ALL SEEN THE MATRIX

WE'VE ALL JOKED ABOUT "WHAT RESOLUTION IS LIFE?"

BUT IT DOESN'T BLUNT THE SHOCK

OF WAKING UP ONE MORNING

AND SEEING DEAD PIXELS IN THE SKY.
"Touring" machines.
My Favorite Touring Machine: Tesla Model S

Auto-pilot!

Theorem: Theory can be beautiful!

0-60 in 2.4 seconds!

315 miles per charge
The Tesla Model S

2005

2015

Auto-pilot!
The world's largest unmanned systems exhibition

August 10-13, 2009 Washington, DC, USA
Welcome

The Unmanned Underwater Vehicle Showcase

The Unmanned Underwater Vehicle Showcase is the annual conference of the Society for Underwater Technology's Underwater Robotics Group. The URG committee develop an annual event that includes a conference, exhibition and seminars on every area of UUVs.

Find out more about the Underwater Robotics Group

UUVS is an opportunity to meet with commercial, defence, scientific and research industry colleagues, users and potential users of unmanned vehicles at the National Oceanography Centre, Southampton.

UUVS is a Technical Conference programme organised by a specially selected committee.

- Find out more about the conference
- Find out more about the exhibition
- Enquire about exhibiting
- Enquire about attending the conference

Supported by

BlueView Technologies
Acoustic Underwater Vision Systems See what top-notch marine construction, quality jobs, you've been missing!

Precon Marine
On time. On budget.

Ads by Google
“Beam me up, Scotty!”

Star Trek communicator, 1966

Apple iPhone 6s Plus, 2016
1.85 GHz 2-core A9 proc., 2 GB DRAM
128 GB flash, color multitouch LCD
(1920x1080 @ 401 ppi), 12 MP camera,
6.8 oz, 4K Video, GPS, Email, Web
surfing, 26 hrs talk & 16 days standby,
millions of apps, $950

Cray 1, 1976
80 MHz
4MB RAM
5.5 tons
200kW+
$8 million

Motorola RAZR, 2004
264 MHz proc., 10 MB RAM
color LCD (176x220),
1.3 MP camera, 4.4 oz

Vs.

iPod 6s has processor speed 46x of Cray 1, at 1/8,400 of cost
⇒ computing power / cost improvement of 387,000x ! (+ inflation)
Reality Surpassing Science Fiction

Chess: **HAL 9000** beating Frank Poole in “2001: A Space Odyssey”, 1968

IBM “**Deep Blue**” beating world chess champion Gary Kasparov, 1997

Elo chess rating scale:
- Master: 2300+ (top 2% of tour. players)
- Grandmaster: 2500+ (top 0.02%)
- **Super-Grandmaster**: 2700+ (31 in 2009)
- ?: 2800+ (only 4 worldwide)
- Kasparov: 2851 (peek in 1999)

**Best human ever**: 2895 (Fisher, 1972)

**Best computer**: 3340 (“Stockfish”, 2015)

iPhone can beat most humans at chess! (2010)
IBM’s “Deep Blue” becomes Chess world champion in 1997

Reality Surpassing Science Fiction
Reality Surpassing Science Fiction

GAME OVER: KASPAROV AND THE MACHINE

THINKFILM and ALLIANCE ATLANTIS PRESENT
AN ALLIANCE ATLANTIS and NATIONAL FILM BOARD OF CANADA PRODUCTION

GAME OVER: KASPAROV AND THE MACHINE

A WORLD DOCUMENTARY FUND FILM
AN INITIATIVE OF THE UK FILM COUNCIL NEW CINEMA FUND,
THE NATIONAL FILM BOARD OF CANADA AND THE BBC
WITH THE ASSISTANCE OF NOAVE CENTRAL - A CORUS ENTERTAINMENT COMPANY

PHOTOGRAPHED BY VIKRAM JAYANTI PRODUCED BY MAL VOGEL
DIRECTED BY PHOTOGRAPHER MARTIN ALBERTI, SIMON DAVID MILLION, NICK BINDER, ANDY THOMSON, NICK FRAZER, PAUL TRUBETSKY, TOM PELLS, AND MARC MICHAT

© 2000 THINKFILM LIMITED-BRITISH BROADCASTING CORPORATION FILM COUNCIL/NATIONAL FILM BOARD OF CANADA. ALL RIGHTS RESERVED
“Watson” AI becomes *Jeopardy* world champion in 2011

Reality Surpassing Science Fiction
Google achieves AI 'breakthrough' by beating Go champion

27 January 2016 | Technology

A Google artificial intelligence program has beaten the European champion of the board game Go.

The Chinese game is viewed as a much tougher challenge than chess for computers because there are many more ways a Go match can play out.

The tech company's DeepMind division said its software had beaten its human rival five games to nil.

When Gary Kasparov lost to chess computer Deep Blue in 1997, IBM marked a milestone in the history of artificial intelligence. On Wednesday, in a research paper released in Nature, Google earned its own position in the history books, with the announcement that its subsidiary DeepMind has built a system capable of beating the best human players in the world at the east Asian board game Go.

Go, a game that involves placing black or white tiles on a 19x19 board and trying to remove your opponents', is far more difficult for a computer to master than a game such as chess.

DeepMind's software, AlphaGo, successfully beat the three-time European Go champion Fan Hui 5-0 in a series of games at the company's headquarters in King's Cross last October. Dr Tanguy Chouard, a senior editor at Nature who attended the matches as part of the review process, described the victory as "really chilling to watch".

"It was one of the most exciting moments of my career," he added. "But with the usual mixed feelings ... in the quiet room downstairs, one couldn't help but root for the poor human being beaten."
“AlphaGo” AI beats European Go champion, January 2016

“AlphaGo” AI beats world Go champion, March 2016

Reality Surpassing Science Fiction

Now the machine has beaten Fan Hui (pictured left) it will face the top human player - Lee Sedol (right) of South Korea – at a meeting in Seoul in March, with the winner to be awarded $1 million (£701,607)
Reality Surpassing Science Fiction

“AlphaGo” AI beats world Go champion
March 2016
WHAT ARE YOU DOING?

GLUING DOWN CHESS PIECES.

WHY?

BECAUSE THERE'S A PICTURE I'VE ALWAYS WANTED.

I'LL NEED YOUR COAT TO SNEAK THIS ONTO THE RIDE.

HEY. HEY! STOP RETRACTING MY CD!

I FEEL UNCOMFORTABLE WHEN MY COMPUTER PHYSICALLY STRUGGLES WITH ME. SURE, I CAN OVERPOWER IT NOW, BUT IT FEELS LIKE A FEW SHORT STEPS FROM HERE TO THE ROBOT WAR.

MEGA COASTER 3000

SOUVENIR PHOTO
Reality Catching up with Science Fiction

Self-driving cars

“Minority Report” film, 2002

“Boss” autonomous vehicle, CMU navigated 60-miles in 4:10 hours first-place winner ($2 million) DARPA Urban Challenge, 2005

“I, Robot” film, 2004

0-60 in 2.4 seconds!
315 miles per charge

Tesla Model S with auto-pilot, 2015
Reality Catching up with Science Fiction

“Rosie” household robot
“The Jetsons”, 1962

“Roomba” autonomous vacuum, by iRobot, 2002

“Verro” pool sweeper, 2007

“Looj” gutter cleaner, 2007

Motoman SDA10 robot cook
by Yaskawa Electric, 2008
Reality Catching up with Science Fiction

Autonomous vehicles/platforms from the “Terminator” movies, 1984-

Autonomous vehicles/platforms from DARPA-sponsored projects
“Hunter-Killer” flying drone from “Terminator 2”, 1991 (VTOL & hovering capability)

AutoCopter Gunship by Neural Robotics, Inc., 2006 (two MPS AA-12 automatic shotguns, high-explosive & armor-piercing rounds, 5 shots per second), cost: $200K
Reality Catching up with Science Fiction

“F/A-37 Talon EDI” autonomous AI plane from movie “Stealth”, 2005

“X-47 Pegasus” autonomous unmanned combat plane, Northrop Grumman, 2003

X-45 autonomous unmanned combat plane, Boeing, 2006
Reality Catching up with Science Fiction

“Power loader” exoskeleton from the movie “Aliens”, 1986

Berkeley Lower Extremity Exoskeleton 2004, can carry 150 lbs at 4 mph

Exoskeleton / suit from the “Iron Man” comic book (1967) and movie (2008)

Cyberdyne’s HAL-5 exoskeleton, 2006
51 lbs, runs 5 hrs, cost: $60K or $600/mo
5x strength amplification
Reality Catching up with Science Fiction

Surgical robotic system from the movie “Logan’s Run” (1976)

Da Vinci robotic surgical system by Intuitive Surgical, Inc. (2009) cost: $1.5 million
Reality Catching up with Science Fiction

“Cloaking device” from TV series “Star Trek”, 1966

“Invisibility cloak” from Harry Potter movie, 2001

“Metacloak” wideband invisibility cloak by Fractal Antenna Systems Inc., 2009

Invisibility cloak system University of Tokyo, 2003
Reality Catching up with Science Fiction

Fact: gap narrowing between natural and artificial intelligence

Q: Will this gap ever close?  
A: We still don’t know.

Q: What is “intelligence”, “mind”, “consciousness”, “sentience”?  
A: We still don’t know.

• In many areas machines already exceeded humans (e.g., chess)
• In some areas computers & tech surpassed sci-fi (e.g., iPhone)

Q: Where is technology going?  
A: We still don’t know.
“The computer is claiming *its* intelligence is real, and *ours* is artificial.”
Technological Singularity

“Technological singularity”
– Stanislaw Ulam & John von Neumann (1958)
“Speculations Concerning the First Ultraintelligent Machine”
– Irving Good (1965)

• When machine intelligence exceeds humans’, machines will design better machines (as humans do).
• This feedback loop will bootstrap an accelerating (and hopefully benevolent) “intelligence explosion”
• Human intelligence will be quickly left behind and not even comprehend what is going on

“Law of accelerating returns”
– Ray Kurzweil (2001)

Accelerating Change (Kurzweil)

**Dynamic RAM Price**
Bits per Dollar at Production (Packaged Dollars)

- Logarithmic Plot
- Noted doubling time: 1.5 years
- Projected Future

**Magnetic Data Storage**
Bits Per Dollar, constant 2000 dollars

- Logarithmic Plot
- Year range: 1950 to 2005

**Internet Data Traffic**

- Logarithmic Plot
- Year range: 1990 to 2004

**Growth in Supercomputer Power**

- Logarithmic Plot
- Required for Human Brain Functional Simulation (2013)
- Required for Human Brain Neural Simulation for Uploading (2025)
- Trendline
- Planned

Note that DRAM speeds have increased during this period.
Technological Singularity (Kurzweil)

Epoch 1: Physics & Chemistry
Information in atomic structures

Epoch 2: Biology
Information in DNA

Epoch 3: Brains
Information in neural patterns

Epoch 4: Technology
Information in hardware and software designs

Epoch 5: Merger of Technology and Human Intelligence
The methods of biology (including human intelligence) are integrated into the (exponentially expanding) human technology base

Epoch 6: The Universe Wakes Up!!
Patterns of matter and energy in the universe become saturated with intelligent processes and knowledge

~2045
We are here!
“Grey goo!”

REMEMBER
ONLY YOU
CAN PREVENT GREY GOO

NEVER RELEASE NANOBOT ASSEMBLERS
WITHOUT REPLICATION LIMITING CODE

Nanotechnology
FOR DUMMIES

Even technophobes will enjoy this introduction to nanotechnology—a literally makes sense! Richard Booker, Earl Boyden
def getSolutionCosts(navigationCode):
    fuelStopCost = 15
    extraCompCost = 8
    thisAlgorithmBeStuck = 99999999
    waterCrossingCost = 45

GENETIC ALGORITHMS TIP:
ALWAYS INCLUDE THIS IN YOUR FITNESS FUNCTION

"NANO TECH SCARES ME"
-EARTH

TOO MUCH TOO FAST
Pace of tech outpaces human comprehension

"Some genetic engineers we turned out to be!"
“Why the Future Doesn’t Need Us”, Wired Magazine, April 2000
by Bill Joy (co-founder of SUN & co-author of Java)

http://www.wired.com/wired/archive/8.04/joy_pr.html

“Our most powerful 21st-century technologies — robotics, genetic engineering, and nanotech — are threatening to make humans an endangered species.” – Bill Joy

This article stirred up much discussion & controversy!
4 A Response to Bill Joy and the Doom-and-Gloom Technofuturists

John Seely Brown and Paul Duguid

If you lived through the 1950s, you might remember President Eisenhower, orderly suburban housing tracts, backyard bomb shelters—and dreams of a nuclear power plant in every home. Plans for industrial nuclear generators had barely left the drawing board before futurists predicted that every house would have a miniature version. From there, technophobes predicted the end of power monopolies, the emergence of the “electronic cottage,” the death of the city and the decline of the corporation.

Pessimists and luddites, of course, envisioned nuclear apocalypse. Each side waited for nirvana, or Armageddon, so it could triumphantly tell the other, “I told you so.”

With “Why the Future Doesn’t Need Us” in the April issue of Wired, Bill Joy invokes those years gone by. No luddite, Joy is an awe-inspiring technologist—as cofounder and chief scientist of Sun Microsystems, he coauthored, among other things, the Java programming language. So when his article describes a technological juggernaut thundering toward society—bringing with it mutant genes, molecular-level nanotechnology machines and superintelligent robots—all need to listen. Like the nuclear prophets, Joy can see the juggernaut clearly. What he can’t see—which is precisely what makes his vision so scary—are any controls.

But it doesn’t follow that the juggernaut is uncontrollable. To understand why not, readers should note the publication in which this article appeared. For the better part of a decade, Wired has been a cheerleader for the digital age. Until now, Wired has rarely been a venue to which people have looked for a way to put a brake on innovation. Therefore its shift with Joy’s article from cheering to warning marks an important and surprising moment in the digital zeitgeist.

In an effort to locate some controls, let’s go back to the nuclear age. Innovation, the argument went back in the 1950s, would make nuclear power plants smaller and cheaper. They would enter mass production and quickly become available to all.

Even today the argument might appear inescapable until you notice what’s missing: The tight focus of this vision makes it almost impossible to see forces other than technology at work. In the case of nuclear development, a host of forces worked to dismantle the dream of a peaceful atom, including the environmental movement, antinuclear protests, concerned scientists, worried neighbors of Chernobyl and Three Mile Island, government regulators and antiproliferation treaties. Cumulatively, these forces slowed the nuclear juggernaut to a crawl.

Similar social forces are at work on technologies today. But because the digerati, like technophobes before them, look to the future with technological tunnel vision, they too have trouble bringing other forces into view.

The Tunnel Ahead

In Joy’s vision, as in the nuclear one, there’s a recognizable tunnel vision that leaves people out of the picture and focuses on technology in splendid isolation. This vision leads not only to doom-and-gloom scenarios, but also to tunnel design: the design of “simple” technologies that are actually difficult to use.

To escape both trite scenarios and bad design, we have to widen our horizons and bring into view not only technological systems, but also social systems. Good designs look beyond the dazzling potential of the technology to social factors, such as the limited patience of most users.

Paying attention to the latter has, for example, allowed the PalmPilot and Nintendo Game Boy to sweep aside more complex rivals. Their elegant simplicity has made them readily usable. And their usability has in turn created an important social support system. The device is so

John Seely Brown is chief scientist of the Xerox Corporation, and director of the Xerox Palo Alto Research Center (PARC). Paul Duguid is a research specialist in the division of Social and Cultural Studies in Education at the University of California, Berkeley, and a consultant at the Xerox PARC. This article reprinted by permission of The Industry Standard; www.thestandard.com, April 13, 2000. Copyright 2000 Standard Media International.
widely used that anyone having trouble with a Pilot or Game Boy rarely has to look far for advice from a more experienced user.

As this small example suggests, technological and social systems shape each other. The same is true on a larger scale. Technologies—such as gunpowder, the printing press, the railroad, the telegraph and the Internet—can shape society in profound ways. But, on the other hand, social systems—in the form of governments, the courts, formal and informal organizations, social movements, professional networks, local communities, market institutions and so forth—shape, moderate and redirect the raw power of technologies.

Given the crisp edges of technology and the fuzzy outlines of society, it certainly isn’t easy to use these two worldviews simultaneously. But if you want to see where we are going, or design the means to get there, you need to grasp both.

This perspective allows a more sanguine look at Joy’s central concern: genetic engineering, nanotechnology and robotics. Undoubtedly, each deserves serious thought. But each should be viewed in the context of the social system in which it is inevitably embedded.

Genetic engineering presents the clearest example. Barely a year ago, the technology seemed to be an unstoppable force. Major chemical and agricultural interests were barreling down an open highway. In the past year, however, road conditions changed dramatically for the worse: Cargill faced Third World protests against its patents; Monsanto suspended research on sterile seeds; and champions of genetically modified foods, who once saw an unproblematic and lucrative future, are scurrying to counter consumer boycotts of their products.

Almost certainly, those who support genetic modification will have to look beyond the technology if they want to advance it. They need to address society directly—not just by putting labels on modified foods, but by educating people about the costs and the benefits of these new agricultural products. Having ignored social concerns, however, proponents have made the people they need to educate profoundly suspicious and hostile.

Nanotechnology offers a rather different example of how the future can frighten us. Because the technology involves engineering at a molecular level, both the promise and the threat seem immeasurable. But they are immeasurable for a good reason: The technology is still almost wholly on the drawing board.

Two of nanotechnology’s main proponents, Ralph Merkle and Eric Drexler, worked with us at the Xerox Palo Alto Research Center in Palo Alto, Calif. The two built powerful nano-CAD tools and then ran simulations of the resulting molecular-level designs. These experiments showed definitively that nano devices are theoretically feasible. No one, however, has laid out a route from lab-based simulation to practical systems in any detail.

In the absence of a plan, it’s important to ask the right questions: Can nanotechnology fulfill its great potential in tasks ranging from data storage to pollution control, all without spiraling out of control? If the lesson of genetic engineering is any guide, planners would do well to consult and educate the public early on, even though useful nano systems are probably decades away.

Worries about robotics appear premature, as well. Internet “bots” that search, communicate and negotiate for their human masters may appear to behave like Homo sapiens, but in fact, bots are often quite inept at functions that humans do well—functions that call for judgment, discretion, initiative or tacit understanding. They are good (and useful) for those tasks that humans do poorly. So they are better thought of as complementary systems, not rivals to humanity. Although bots will undoubtedly get better at what they do, such development will not necessarily make them more human.

Are more conventional clanking robots—the villains of science fiction—any great threat to society? We doubt it. Xerox PARC research on self-aware, reconfigurable “polybots” has pushed the boundaries of what robots can do, pointing the way to “morphing robots” that are able to move and change shape.

Nonetheless, for all their cutting-edge agility, these robots are a long way from making good dance partners. The chattiness of Star Wars’ C-3PO still lies well beyond real-world machines. Indeed, what talk robots or computers achieve, though it may appear similar, is quite different from human talk. Talking machines travel routes designed specifically to avoid the full complexities of human language.

Robots may seem intelligent, but such intelligence is profoundly hampered by their inability to learn in any significant way. (This failing has apparently led Toyota, after heavy investment in robotics, to consider replacing robots with humans on many production lines.) And without learning, simple common sense will lie beyond robots for a long time to come.

Indeed, despite years of startling advances and innumerable successes like the chess-playing Big Blue, computer science is still about as far as
it ever was from building a machine with the learning abilities, linguistic competence, common sense or social skills of a 5-year-old child.

As with Internet bots, real-world robots will no doubt become increasingly useful. But they will probably also become increasingly frustrating to use as a result of tunnel design. In that regard, they may indeed seem antisocial, but not in the way of Terminator-like fantasies of robot armies that lay waste to human society.

Indeed, the thing that handicaps robots most is their lack of a social existence. For it is our social existence as humans that shapes how we speak, learn, think and develop common sense. All forms of artificial life (whether bugs or bots) will remain primarily a metaphor for—rather than a threat to—society, at least until they manage to enter a debate, sing in a choir, take a class, survive a committee meeting, join a union, pass a law, engineer a cartel or summon a constitutional convention.

These critical social mechanisms allow society to shape its future. It is through planned, collective action that society forestalls expected consequences (such as Y2K) and responds to unexpected events (such as epidemics).

The Failure of a “6-D” Vision

Why does the threat of a cunning, replicating robot society look so close from one perspective, yet so distant from another? The difference lies in the well-known tendency of futurologists to count “1, 2, 3 . . . a million.” That is, once the first step on a path is taken, it’s very easy to assume that all subsequent steps are trivial.

Several of the steps Joy asks us to take—the leap from genetic engineering to a “white plague”; from simulations to out-of-control nanotechnology; from replicating peptides to a “robot species”—are extremely large. And they are certainly not steps that will be taken without diversions, regulations or controls.

One of the lessons of Joy’s article, then, is that the path to the future can look simple (and sometimes downright terrifying) if you look at it through what we call “6-D lenses.” We coined this phrase having so often in our research hit upon upon such “de-” or “di-” words as demassification, decentralization, disintermediation, despacialization, disaggregation and demarketization in the canon of futurology.

If you take any one of these words in isolation, it’s easy to follow their relentless logic to its evident conclusion. Because firms are getting smaller, for example, it’s easy to assume that companies and other intermediaries are simply disintegrating into markets. And because communication is growing cheaper and more powerful, it’s easy to believe in the “death of distance.”

But things rarely work in such linear fashion. Other forces are often at work, such as those driving firms into larger and larger mergers to take advantage of social, rather than merely technological, networks. Similarly, even though communications technology has killed distance, people curiously can’t stay away from the social hotbed of modern communications technology, Silicon Valley.

Importantly, these d-words indicate that the old ties that once bound communities, organizations and institutions are being picked apart by technologies. A simple, linear reading, then, suggests that these communities, organizations and institutions will now simply fall apart. A more complex reading, taking into account the multiple forces at work, offers another picture.

While many powerful national corporations have grown insignificant, some have transformed into more powerful transnational firms. While some forms of community may be dying, others, bolstered by technology, are growing stronger.

Technology and society are constantly forming and reforming new dynamic equilibriums with far-reaching implications. The challenge for futurology (and for all of us) is to see beyond the hype and past the over-simplifications to the full import of these new sociotechnical formations.

Two hundred years ago, Thomas Malthus, assuming that human society and agricultural technology developed on separate paths, predicted that society was growing so fast that it would starve itself to death, the so-called Malthusian trap.

A hundred years later, H.G. Wells similarly assumed that society and technology were developing independently. Like many people today, Wells saw the advance of technology outstripping the evolution of society, leading him to predict that technology’s relentless juggernaut would unfeelingly crush society. Like Joy, both Malthus and Wells issued important warnings, alerting society to the dangers it faced. But by their actions, Malthus and Wells helped prevent the very future they were so certain would come about.

These self-unfulfilling prophecies failed to see that, once warned, society could galvanize itself into action. Of course, this social action in the face of threats showed that Malthus and Wells were most at fault.
in their initial assumption. Social and technological systems do not develop independently; the two evolve together in complex feedback loops, wherein each drives, restrain and accelerates change in the other. Malthus and Wells—and now Joy—are, indeed, critical parts of these complex loops. Each knew when and how to sound the alarm. But each thought little about how to respond to that alarm.

Once the social system is factored back into the equation like this, the road ahead becomes harder to navigate. Ultimately we should be grateful to Joy for saying, at the least, that there could be trouble ahead when so many of his fellow digerati will only tell us complacently that the road is clear.
In the coming decades, humanity will likely create a powerful artificial intelligence. SIAI exists to handle this urgent challenge, both the opportunity and the risk.
The Singularity Summit 2008

Thank you to everyone who made Singularity Summit 2008 such a success.

Videos now available here

Online Coverage | Graphic Facilitator Images

Summit Coverage

- San Jose Mercury News
- Wired: How Robots Will Steal Your Job

Past Summits

Founded in 2006 by Tyler Emerson, Ray Kurzweil, and Peter Thiel, the inaugural summit was held at Stanford, the first academic symposium focused on singularity dialogue.

Past speakers have included Doug Hofstadter (author of Gödel, Escher, Bach), Peter Norvig (Google Director of Research), Sebastian Thrun (Stanford AI Lab Director), and Rodney Brooks (MIT Professor of Robotics)

Past Summits

Latest News

- Summit Follow-Up
- Behind the Scenes: Graphic Recorders from Visual Ink and an Independent Artist
- Emerging Tech Workshop Schedule/Speakers

Stay Connected

Featured Partner

Summit Graphics

More graphics
The Singularity Summit 2009 is the premier dialog on the Singularity.

The first Singularity Summit was held at Stanford in 2006 to further understanding and discussion about the Singularity concept and the future of human technological progress. It was founded as a venue for leading thinkers to explore the subject, whether scientist, enthusiast, or skeptic.

Since 2006, the scope of this dialog has expanded dramatically. In 2008, the Singularity entered mainstream consideration. IEEE Spectrum, a sober and mainstream technology publication, issued a special report on the Singularity, and Intel CTO Justin Rattner remarked that "we're making steady progress toward the Singularity" during his keynote to 2,000 people at the Intel Developer Forum. What was once a relatively unknown concept is now being discussed in corporate board rooms.

We invite you to join our extraordinary group of visionaries in business, science, technology, design, and the arts, as our community explores this exciting topic. Your participation offers a world of powerful ideas, a unique networking opportunity, and access to an exclusive directory of your peers.

We hope you will join us October 3rd. Register here.
Selected Summit Talks:

Ray Kurzweil
The Ubiquity and Predictability of the Exponential Growth of Information Technology
Founder and CEO, Kurzweil Technologies

Aubrey de Grey
The Singularity and the Methuselah: Similarities and Differences
Chief Science Officer, SENS Foundation

Peter Thiel
Macroeconomics and Singularity
President, Clarium Capital Management; Co-Founder, PayPal; Managing Partner, Founders Fund; Seed Investor, Facebook

Stephen Wolfram
Conversation on the Singularity
Founder and CEO, Wolfram Research

David Chalmers
Simulation and the Singularity
Professor of Philosophy, Australian National University, Director of the Centre for Consciousness

Michael Nielsen
Collaborative Networks in Scientific Discovery
Quantum Computing Pioneer, Author
“Is that it? Is that the Grand Unified Theory?”
Further Reading

Alan Turing:  
http://en.wikipedia.org/wiki/Alan_Turing

Turing test:  
http://en.wikipedia.org/wiki/Turing_test

Chinese room:  

Artificial intelligence:  
http://en.wikipedia.org/wiki/Artificial_intelligence

Artificial intelligence in fiction:  
http://en.wikipedia.org/wiki/Artificial_intelligence_in_fiction

Isaac Asimov:  
http://en.wikipedia.org/wiki/Isaac_Asimov

Three Laws of Robotics:  

Robots in literature:  

Fictional robots and androids  
Further Reading

Unmanned aerial vehicles:
http://en.wikipedia.org/wiki/Unmanned_aerial_vehicle

Unmanned ground vehicles:

Autonomous underwater vehicles:

Micro aerial vehicles:
http://en.wikipedia.org/wiki/Micro_air_vehicle

DARPA Grand Challenge:

Driverless cars:

Exoskeletons and “wearable robots”:
http://en.wikipedia.org/wiki/Powered_exoskeleton

Technological singularity:
http://en.wikipedia.org/wiki/Technological_singularity