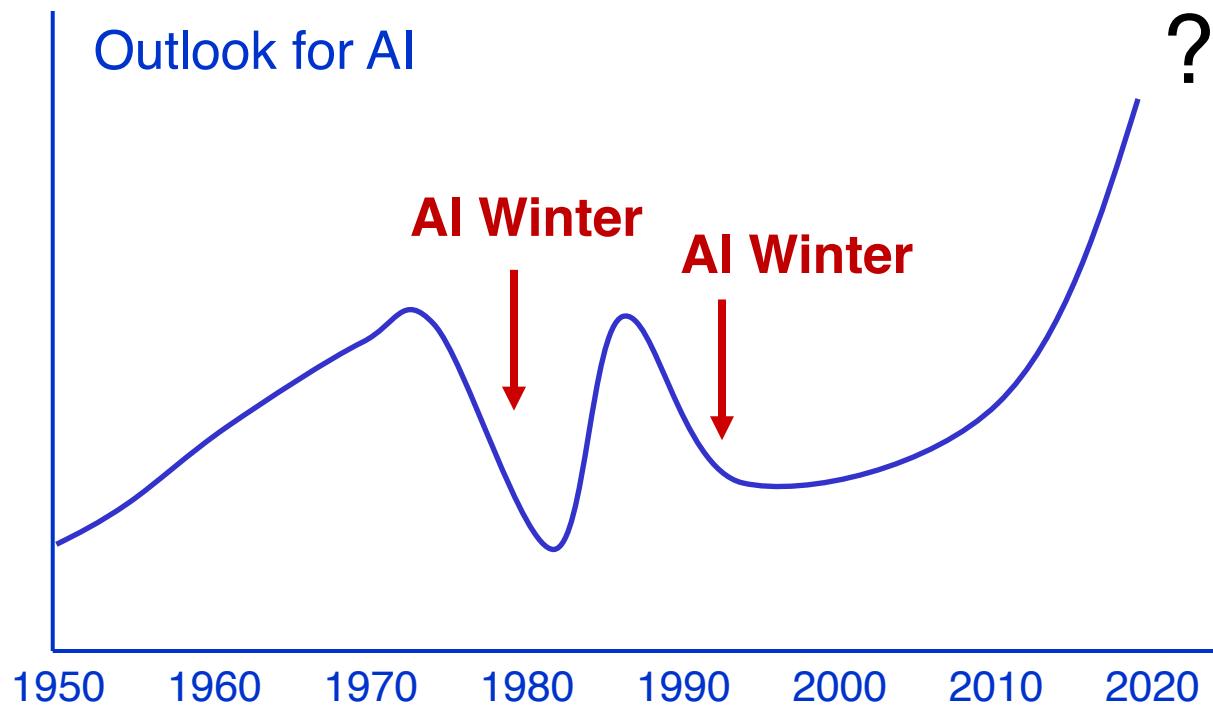


Why AI is Harder Than We Think

Melanie Mitchell
Santa Fe Institute



Search jobs

Sign in

Search ▾

US edition ▾



2015

**Self-driving cars: from 2020 you will
become a permanent backseat driver**

Search jobs

Sign in

Search ▾

US edition ▾

The Guardian

2015

**Self-driving cars: from 2020 you will
bec**

BUSINESS
INSIDER

TECH | FINANCE | POLITICS | STRATEGY | LIFE | ALL

2016

10 million self-driving cars will be on the road by 2020

Search jobs

Sign in

Search ▾

US edition ▾

The Guardian

2015

Self-driving cars: from 2020 you will
bec

BUSINESS
INSIDER

TECH | FINANCE | POLITICS | STRATEGY | LIFE | ALL

2016

10 million self-driving cars will be on the roads by 2020



2019

“A year from now, we’ll have over a million cars
with full self-driving, software, everything.”

— Elon Musk

Search jobs

Sign in

Search ▾

US edition ▾

The Guardian

2015

Self-driving cars: from 2020 you will
become

BUSINESS
INSIDER

TECH | FINANCE | POLITICS | STRATEGY | LIFE | ALL

2016

10 million self-driving cars will be on the roads by 2020



2019

“A year from now, we’ll have over a million cars



2022

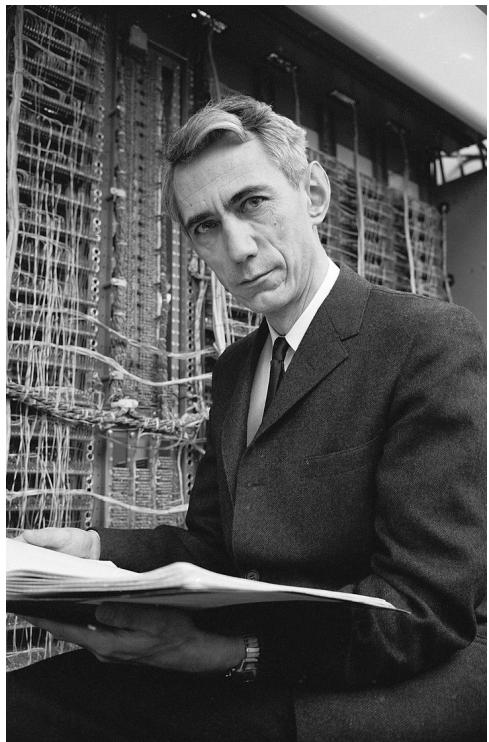
“I would be shocked if we do not achieve full self-driving safer than human this year. I would be shocked.” — Elon Musk

“Perhaps expectations are too high, and... this will eventually result in disaster.... [S]uppose that five years from now [funding] collapses miserably as autonomous vehicles fail to roll. Every startup company fails. And there's a big backlash so that you can't get money for anything connected with AI. Everybody hurriedly changes the names of their research projects to something else.

This condition [is] called the ‘[AI Winter](#).’”

—Drew McDermott, 1984

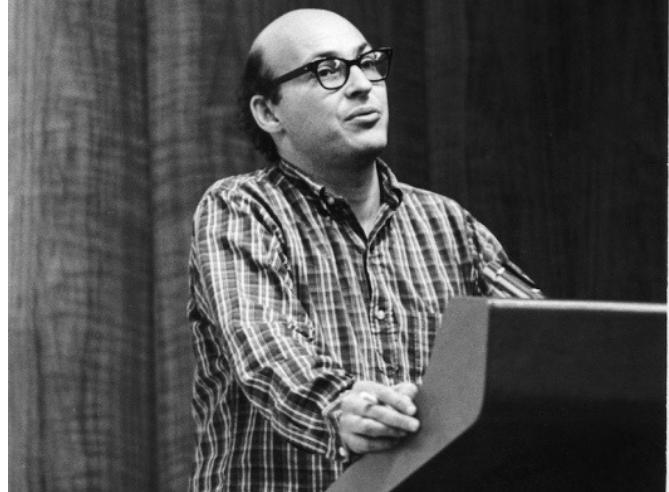




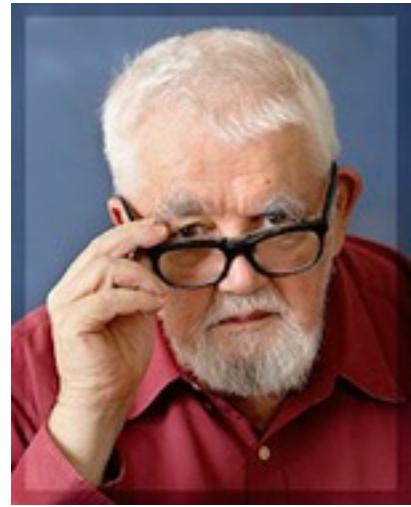
I confidently expect that within a matter of 10 or 15 years, something will emerge from the laboratory which is not too far from the robot of science fiction fame.
— Claude Shannon, 1961



Machines will be capable, within twenty years, of doing any work that a man can do.
— Herbert Simon, 1965



Within a generation...the problem of creating ‘artificial intelligence’ will be substantially solved.
— Marvin Minsky, 1967



“AI was harder than we thought.”
— John McCarthy, 2006



Human-level AI will be passed in the mid-2020s.

— Shane Legg, 2008



One of [Facebook's] goals for the next five to 10 years is to basically get better than human level at all of the primary human senses: vision, hearing, language, general cognition

— Mark Zuckerberg, 2015



I estimate a 50% probability of human-level AI by 2040.

— Richard Sutton, 2022



When will superintelligent AI arrive?...it [will] probably happen in the lifetime of my children.

(My timeline of, say, eighty years is considerably more conservative than that of the typical AI researcher.)

— Stuart Russell, 2019

Why AI is Harder Than We Think

Melanie Mitchell

Santa Fe Institute
Santa Fe, NM, USA
mm@santafe.edu

Still

Why AI is Harder Than We Think



Melanie Mitchell
Santa Fe Institute
Santa Fe, NM, USA
mm@santafe.edu

Four fallacies that promote over-optimism on AI

Fallacy 1: Narrow AI is on a continuum with general AI

IBM® Watson™ represents a first step into cognitive systems, a new era of computing.

AlphaZero ... is the first step in creating real AI.

GPT-2 AS STEP TOWARD GENERAL INTELLIGENCE

Hubert Dreyfus: "The **first-step fallacy** is the claim that, ever since our first work on computer intelligence we have been inching along a continuum at the end of which is AI, so that any improvement in our programs no matter how trivial counts as progress....There was in fact a discontinuity in the assumed continuum of steady incremental progress. The unexpected obstacle was called the **commonsense knowledge problem.**"

Fallacy 2: Easy things are easy and hard things are hard

What Artificial Intelligence Can and Can't Do Right Now

by Andrew Ng

November 09, 2016

If a typical person can do a mental task with less than one second of thought, we can probably automate it using AI either now or in the near future.

Fallacy 2: Easy things are easy and hard things are hard

Mastering the game of Go without human knowledge

Nature Published: 19 October 2017

Go is among “the most challenging domains in terms of human intellect.”

The Washington Post

September 28, 2022

AI can now create any image in seconds, bringing wonder and danger

The New York Times

*We Need to Talk About How Good A.I.
Is Getting*

August 24, 2022



OpenAI GPT-3

Playground

Load a preset... ▾

Save

I put a green box on top of a blue box. I put a yellow box on the green box, and another red box on top of that. I then took off the red box and put it beneath the blue box. What is on top now? 

The green box is on top.

What is on the bottom?

The blue box is on the bottom.

Playground

Write a sentence that is not about ice cream.

I'm not a fan of ice cream.



DALL·E My collection

Edit the detailed description

Surprise me

Upload



Generate

A yellow box on top of a green box, which is
on top of a blue box.





DALL·E My collection

Edit the detailed description

Surprise me

Upload



Generate

A bowl of fruit with no apples.



Moravec's paradox: “It is comparatively easy to make computers exhibit adult level performance on intelligence tests or playing checkers, and difficult or impossible to give them the skills of a one-year-old when it comes to perception and mobility.” — **and common sense!**

Fallacy 3: The lure of “wishful mnemonics” (McDermott, 1976)

“Watson can **read** all of the health-care texts in the world in seconds.”

“Watson **understands** context and nuance in seven languages.”

“AlphaGo’s **goal** is to beat the best human players not just mimic them.”

'It's able to create **knowledge** itself':
Google unveils AI that **learns on its own**

Fallacy 3: The lure of “wishful mnemonics” (McDermott, 1976)

What is SQuAD?

Stanford Question Answering Dataset (SQuAD) is a reading comprehension dataset, consisting of questions posed by crowdworkers on a set of Wikipedia articles,

The General Language Understanding Evaluation (GLUE) benchmark is a collection of resources for training, evaluating, and analyzing natural language understanding systems.

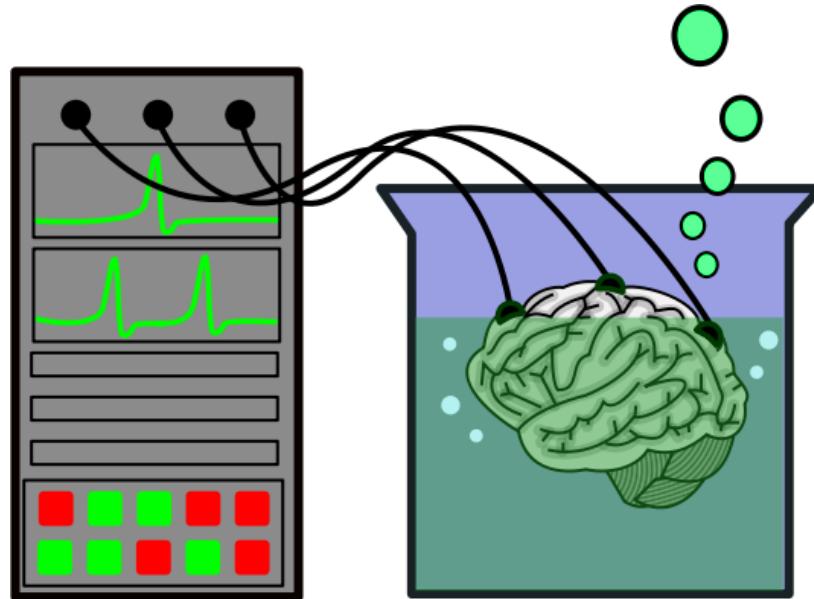
Visual Commonsense Reasoning (**VCR**) is a new task and large-scale dataset for cognition-level visual understanding.




Leaderboard Version: 2.0

| Rank | Name | Model | URL | Score | BoolQ | CB | COPA | MultiRC | ReCoRD | RTE | WiC | WSC | AX-g | AX-b |
|------|----------------------------|---------------------------------------|-------------------|-------|----------------|-------|-----------|-----------|--------|------|-------|-----------|------|------|
| + 1 | Liam Fedus | ST-MoE-32B | ↗ | 91.2 | 92.4 96.9/98.0 | 99.2 | 89.6/65.8 | 95.1/94.4 | 93.5 | 77.7 | 96.6 | 96.1/94.1 | 72.3 | |
| 2 | Microsoft Alexander v-team | Turing NLR v5 | ↗ | 90.9 | 92.0 95.9/97.6 | 98.2 | 88.4/63.0 | 96.4/95.9 | 94.1 | 77.1 | 97.3 | 93.3/95.5 | 67.8 | |
| 3 | ERNIE Team - Baidu | ERNIE 3.0 | ↗ | 90.6 | 91.0 98.6/99.2 | 97.4 | 88.6/63.2 | 94.7/94.2 | 92.6 | 77.4 | 97.3 | 92.7/94.7 | 68.6 | |
| 4 | Yi Tay | PaLM 540B | ↗ | 90.4 | 91.9 94.4/96.0 | 99.0 | 88.7/63.6 | 94.2/93.3 | 94.1 | 77.4 | 95.9 | 95.5/90.4 | 72.9 | |
| + 5 | Zirui Wang | T5 + UDG, Single Model (Google Brain) | ↗ | 90.4 | 91.4 95.8/97.6 | 98.0 | 88.3/63.0 | 94.2/93.5 | 93.0 | 77.9 | 96.6 | 92.7/91.9 | 69.1 | |
| + 6 | DeBERTa Team - Microsoft | DeBERTa / TuringNLPv4 | ↗ | 88.8 | 88.1 95.7/97.0 | 98.4 | 88.2/63.7 | 94.5/94.1 | 93.2 | 77.5 | 95.9 | 93.3/93.8 | 66.7 | |
| 7 | SuperGLUE Human Baselines | SuperGLUE Human Baselines | ↗ | 89.8 | 89.0 95.8/98.9 | 100.0 | 81.8/51.9 | 91.7/91.3 | 93.6 | 80.0 | 100.0 | 99.3/99.7 | 76.6 | |

Fallacy 4: Intelligence is all in the brain



<https://www.insidetherift.net/mind-at-large/2017/6/1/philosophy-portal-brain-in-a-vat>

NEUROSCIENCE

Why AI Lags Behind the Human Brain in Computational Power

INNOVATION

The Human Brain-Scale AI Supercomputer Is Coming

NEWS

This massive AI chip has the compute power of a human brain



Nando de Freitas

@NandoDF

...

Someone's opinion article. My opinion: It's all about scale now! The Game is Over! It's about making these models bigger, safer, compute efficient, faster at sampling, smarter memory, more modalities, INNOVATIVE DATA, on/offline, ... 1/N



Alex Dimakis

@AlexGDimakis

...

Based on recent papers (Gpt3, Palm, dalle2, Gato, Metaformer) I am forming the opinion that maybe 'Scale is all you need', possibly even for general intelligence (?!). Just convert everything to tokens and predict the next token. (1/n)

7:24 PM · May 16, 2022 · Twitter for Android

84 Retweets

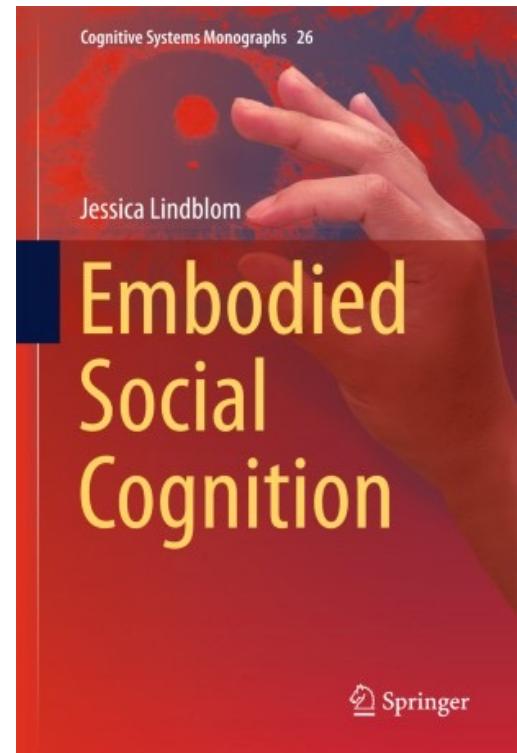
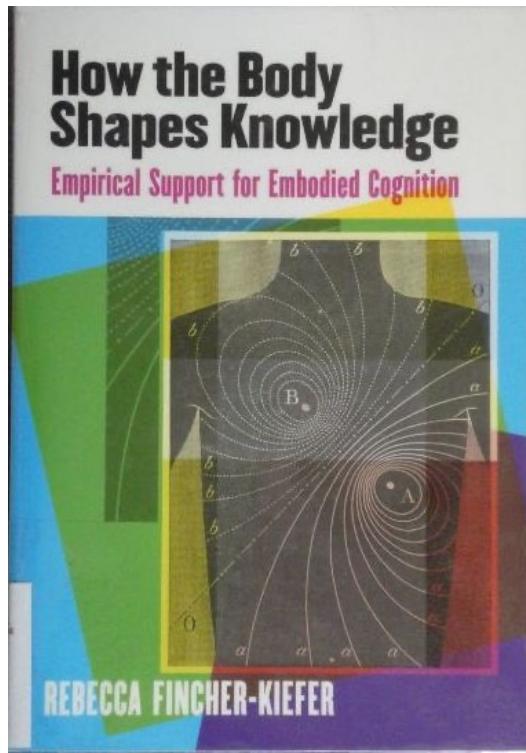
43 Quote Tweets

745 Likes

EMBODIED MIND, MEANING, AND REASON

HOW OUR BODIES
GIVE RISE TO UNDERSTANDING

MARK JOHNSON



Major Open Challenges

- Few-shot learning
- Generalization
- Abstraction and analogy
- Understanding and Common Sense

JACK STEWART

TRANSPORTATION 10.18.2018 87:00 AM

Why People Keep Rear-Ending Self-Driving Cars

Human drivers (and one cyclist) have rear-ended self-driving cars 28 times this year in California—accounting for nearly two-thirds of robocar crashes.





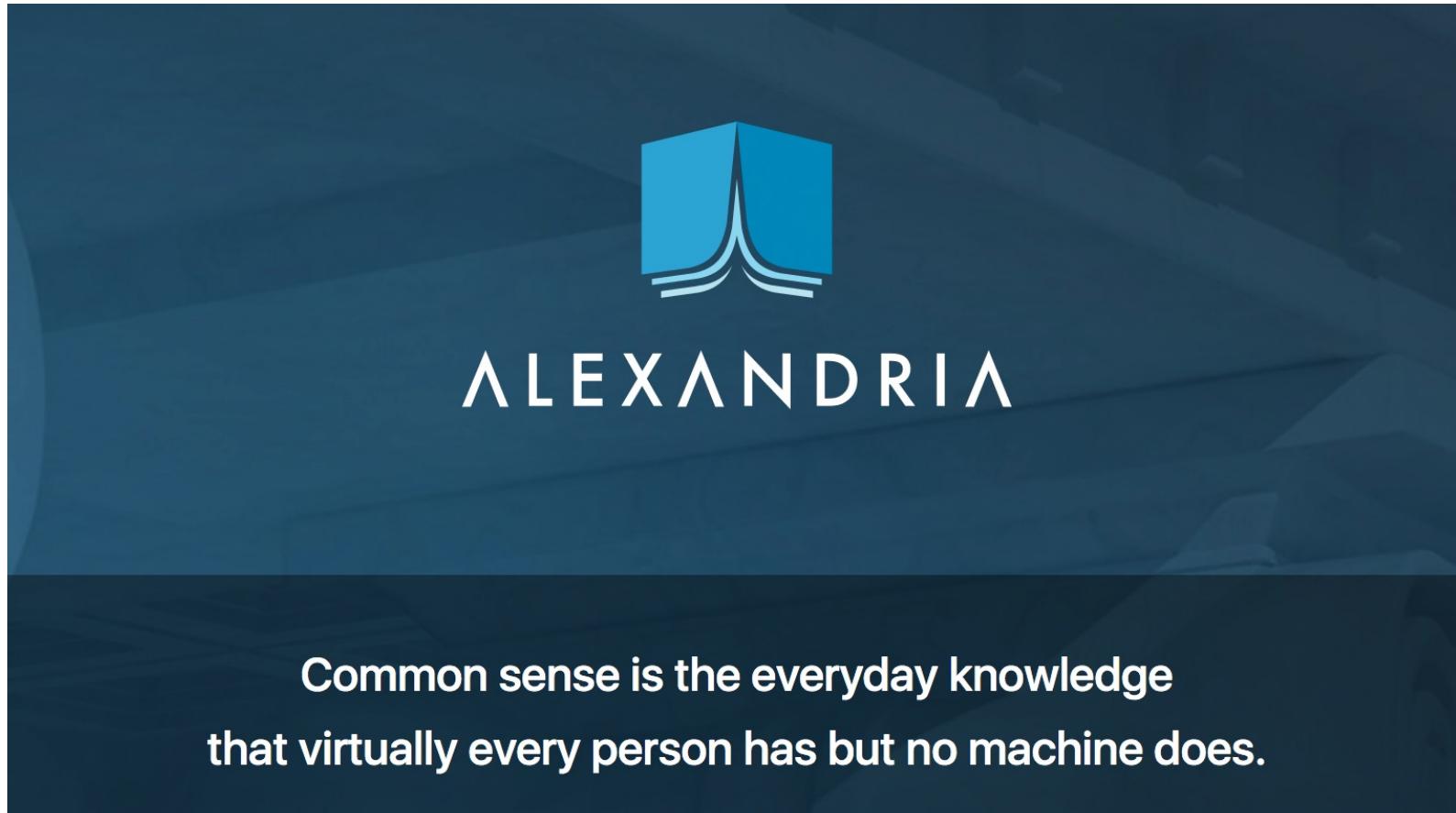






Paul Allen invests \$125 million to teach computers common sense

<https://www.seattletimes.com/business/technology/paul-allen-invests-125-million-to-teach-computers-common-sense/>



The background of this section is a dark blue gradient with faint, curved white lines suggesting a landscape or architectural design. In the center, there is a large, stylized blue logo consisting of three vertical rectangles that meet at the bottom, forming a shape reminiscent of a book or a rising sun. Below this logo, the word "ALEXANDRIA" is written in a large, white, sans-serif font. At the bottom of the slide, there is a white rectangular box containing a quote in a white serif font.

Common sense is the everyday knowledge
that virtually every person has but no machine does.

<https://allenai.org/alexandria/>

**Department of Defense
Fiscal Year (FY) 2019 Budget Estimates**

February 2018



Defense Advanced Research Projects Agency

Title: Machine Common Sense (MCS)

Description: The Machine Common Sense (MCS) program will explore approaches to commonsense reasoning. Recent advances in machine learning have resulted in exciting new artificial intelligence (AI) capabilities in areas such as speech recognition, natural language processing, and two-person strategy games (Chess, Go). But in all of these applications, the machine reasoning is narrow and highly specialized; broad, commonsense reasoning by machines remains elusive. This program will create more human-like knowledge representations, for example, perceptually-grounded representations of commonsense reasoning by machines about the physical world and spatio-temporal phenomena. Equipping AI systems with more human-like reasoning capabilities will make it possible for humans to teach/correct a machine as they interact on tasks, enabling more equal collaboration and ultimately symbiotic partnerships between humans and machines.

FY 2019 Plans:

- Develop approaches for machine reasoning about imprecise and uncertain information derived from text, pictures, speech, and sensor data.
- Design methods to enable machines to identify knowledge gaps and reason about their state of knowledge.
- Formulate perceptually-grounded representations to enable commonsense reasoning by machines about the physical world and spatio-temporal phenomena.



The state of Computer Vision and AI: we are really, really far away.

Oct 22, 2012



What would it take for a computer to understand this image?

Some core components of human understanding

- Intuitive physics, biology, psychology
- Mental models of cause and effect
- Vast world-knowledge
- Abstraction and analogy



In spite of enormous progress in specific areas, AI systems are still far from having anything like these general human abilities.

AI is still harder than we think.

A PROPOSAL FOR THE
DARTMOUTH SUMMER RESEARCH PROJECT
ON ARTIFICIAL INTELLIGENCE

J. McCarthy, Dartmouth College

M. L. Minsky, Harvard University

N. Rochester, I.B.M. Corporation

C.E. Shannon, Bell Telephone Laboratories

August 31, 1955

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.

A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

“One might even say that the [1955] proposal, if properly reformatted, could be resubmitted to the NSF...today and would probably get some funding by some excited program managers.”

— Eric Horvitz, 2015

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.

Artificial Intelligence

A Guide for
Thinking Humans



Melanie Mitchell

Thank you for listening!