SEEM3510 Human-Computer Interaction

Helen Meng

Department of Systems Engineering
& Engineering Management

Philip Fu

Department of Computer Science
& Engineering

Spring 2022
Week 1

Housekeeping

- Course Overview
- Academic Honesty

  https://www.cuhk.edu.hk/policy/academichonesty/

Outline

- Motivating example
- Introduction to HCI
- AI and HCI

Motivating Example

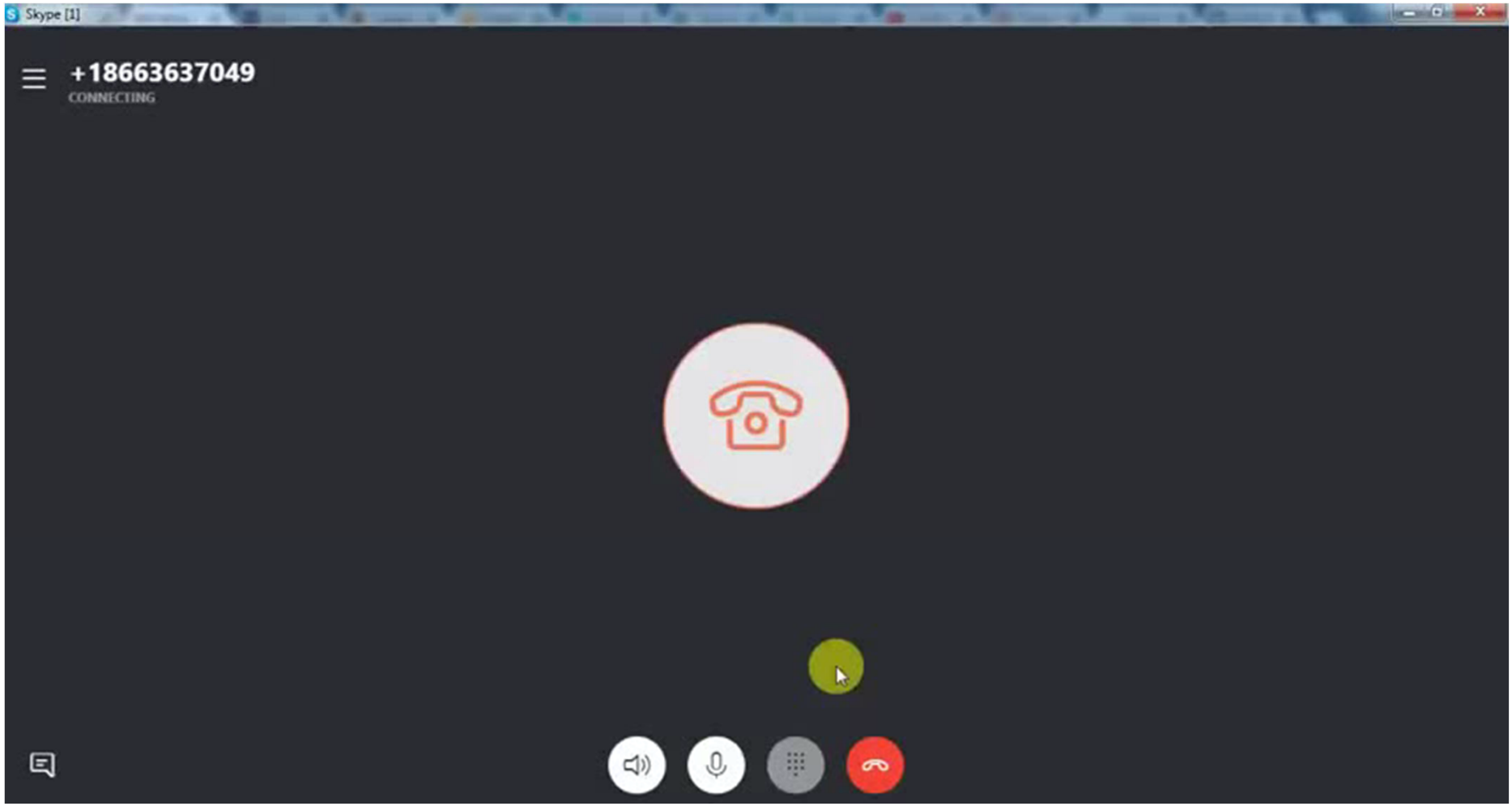
1. Interactive Voice Response (IVR) System
2. IVR DTMF demonstration
3. Intelligent assistant
4. Quick discussion
5. Some fun

Graphic: Haptik.ai

1. What is an Interactive Voice Response (IVR) System?

Source: Neotel

2. IVR DTMF demonstration



Source: Konnect247

3. Intelligent Voice Assistant

Source: Google Assistant

4. Some Fun

Source: Seinfeld

IVR DTMF

System: Welcome to Hilton Suites…. Select language Press 1…. Press 2….
User: Press
System: Hi this is Amy…. Set preferred language Press 1
User: Press
System: For hotel reservation Press 1, for modification Press 2, for cancellation Press 3, for other services Press 4
User: Pressed something
System: Enter your arrival date, example 10.01
User: Pressed something
System: Enter your departure date, example 10.01
User: Pressed something
System: OK. Confirm your arrival date is… departure date is…. To Confirm Press 1, To Re-enter Press 2
System: Enter # adults to stay <barge in>
User: Pressed something
System: Enter # children to stay <barge in>
User: Pressed something
System: OK, advise room types and cost Press 1 for standard room, Press 2 for superior room, Press 3 for Deluxe

Intelligent Voice Assistant

User: requested for “hotels in Barcelona”
System: suggested hotels, asked about dates
Users: spoke about dates and # nights
Systems: showed what was available, confirmed dates, asked about more specs
User: requested “5 star, swimming pool”
System: showed matching hotels
User: spoke about price “< 300dollars”
System: confirmed and showed updated options
User: selected one option
System: asked for confirmation with map and key details, asked for confirmation
User: YES
System: confirmed with all details, asked about booking
User: selected booking
System: confirmed with booking details and costs, asked for confirmation
User: confirmed ready to book
System: asked for CVC, faceID (on device)
User: confirmed
System: advised completion, remind about confirmation email.

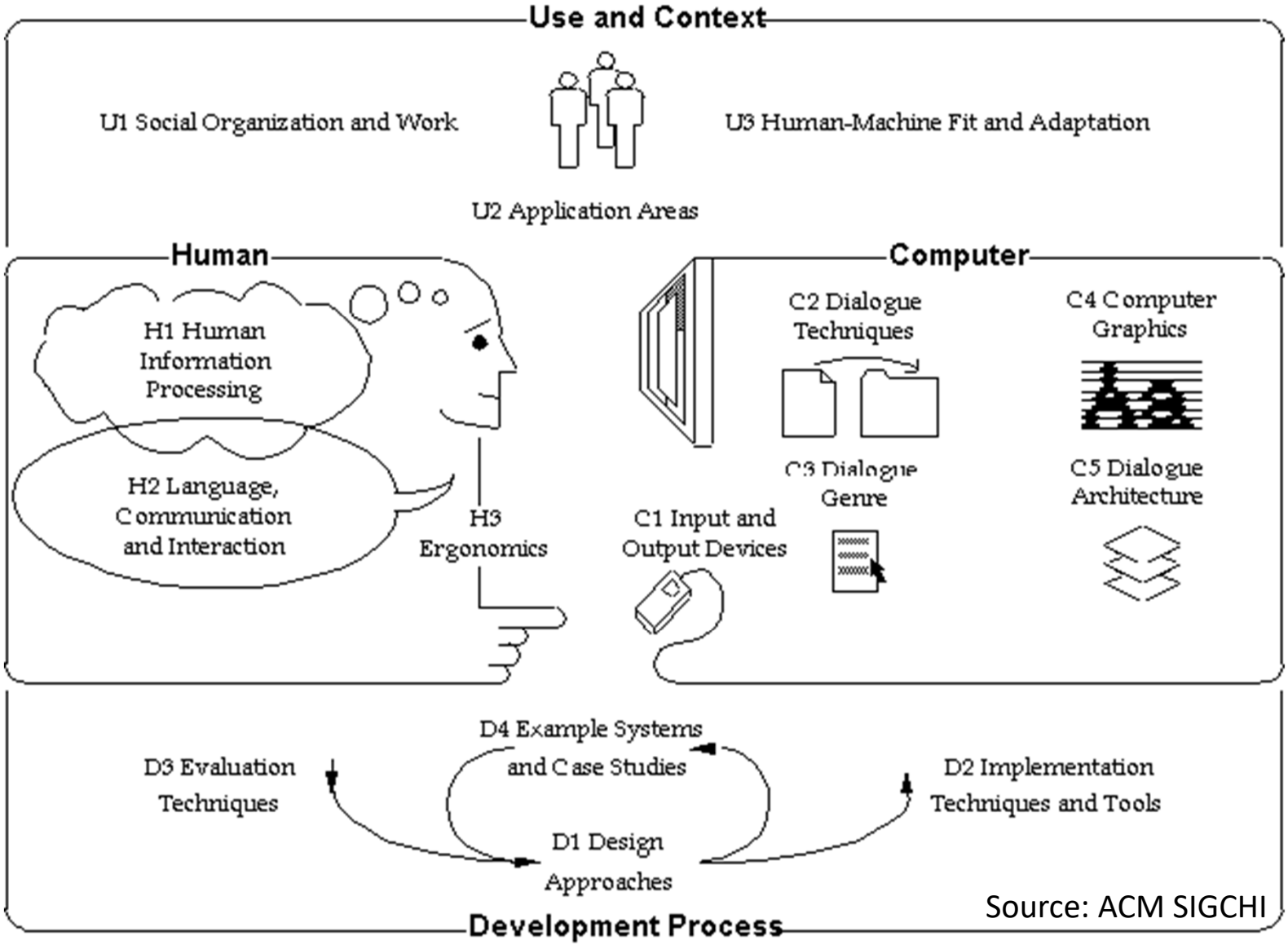
Discussion

Definition of HCI

- Human-Computer Interaction is a discipline concerned with the design, evaluation, and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them

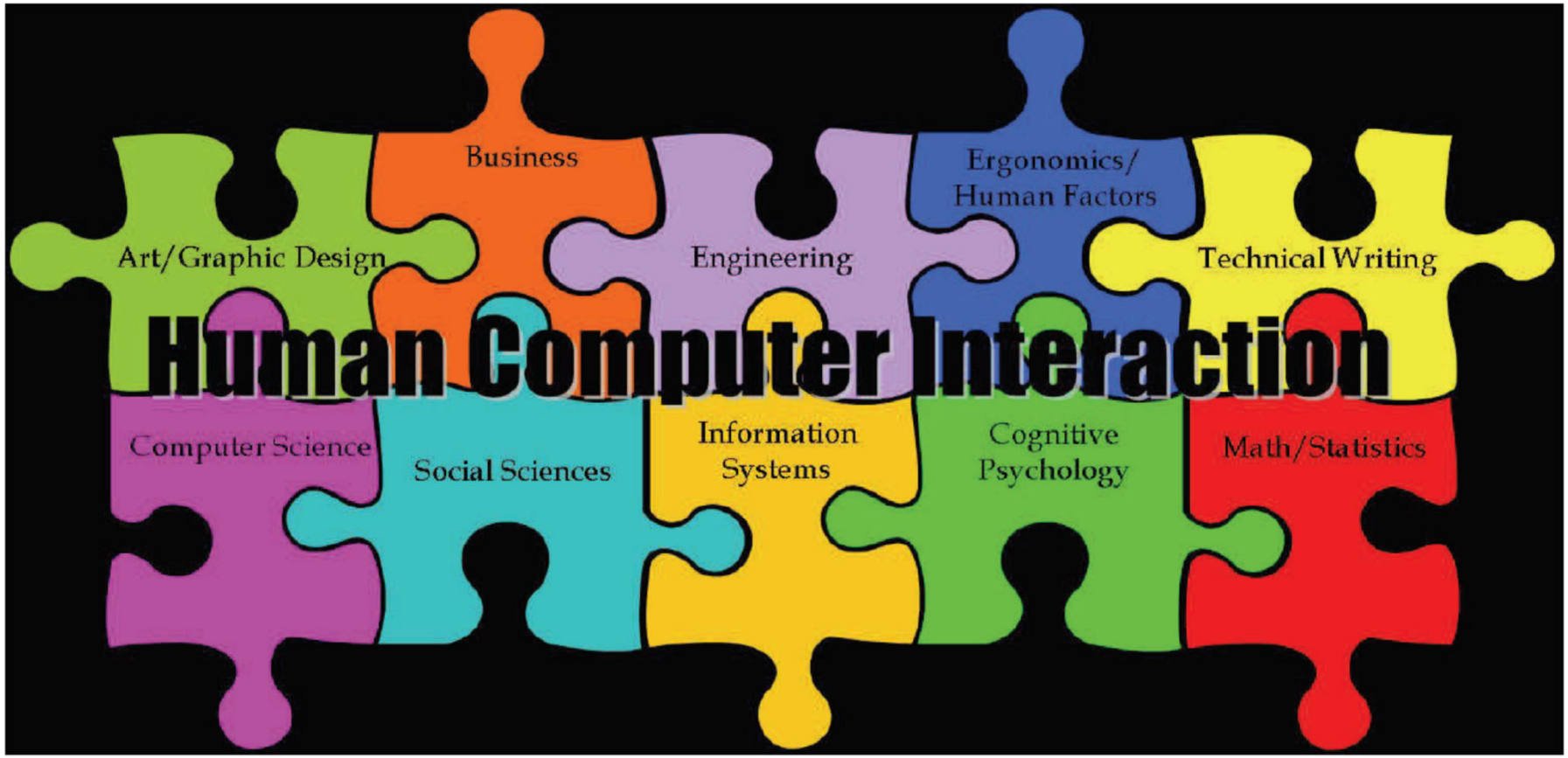
–ACM SIGCHI Definition, Thomas Hewett 1992

Taxonomy of HCI



Source: ACM SIGCHI

An Interdisciplinary Field



Graphic: Lester (2008) ISBN: 978-953-7619-15-2, InTech

History – As We May Think (1945)

- Vannevar Bush, “As We May Think” – published in 1945
  - American engineer, inventor and science administrator
  - VP of MIT and Dean of Engineering
  - Head of US OSRD (Office of Scientific Research and Development), carried out wartime military R&D

"As We May Think" predicted (to some extent) many kinds of technology invented after its publication, including hypertext, personal computers, the Internet, the World Wide Web, speech recognition, and online encyclopedias such as Wikipedia: "Wholly new forms of encyclopedias will appear, ready-made with a mesh of associative trails running through them, ready to be dropped into the memex and there amplified." Bush envisioned the ability to retrieve several articles or pictures on one screen, with the possibility of writing comments that could be stored and recalled together. He believed people would create links between related articles, thus mapping the thought process and path of each user and saving it for others to experience. Wikipedia is one example of how this vision has in part been realized, allowing elements of an article to reference other related topics. A user's browser history maps the trails of possible paths of interaction, although this is typically available only to the user that created it. Bush's article also laid the foundation for new media. Doug Engelbart came across the essay shortly after its publication, and keeping the memex in mind, he began work that would eventually result in the invention of the mouse, the word processor, the hyperlink and concepts of new media for which these groundbreaking inventions were merely enabling

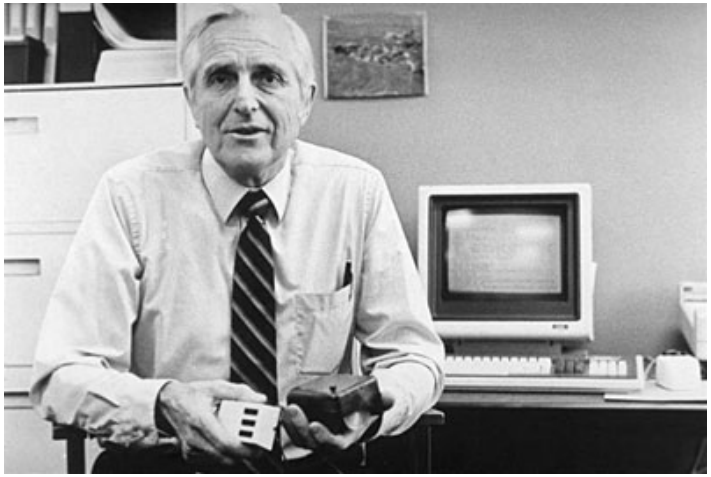
Source: Wikipedia

History – Sketchpad (1962)

- Ivan Sutherland's PhD research in Electrical Engineering in MIT
- "Heretofore, most interaction between man and computers has been slowed by the need to reduce all communication to written statements that can be typed. In the past we have been writing letters to, rather than conferring with, our computers [Sutherland, 1963, p. 329]
- With Sketchpad, objects were drawn, resized, grabbed and moved, extended and deleted using a lightpen
- Videos
  - https://www.youtube.com/watch?v=vPsFPmgT0YM
  - https://www.youtube.com/watch?v=5RyU50qbvzQ
  - https://www.youtube.com/watch?v=6orsmFndx_o

History – Mouse (1963)

- Douglas Engelbar's invention in Stanford Research Institute
- Replaces the light pen (user fatigue)
- Conducted user study [English, Engelbart and Berman, IEEE 1967]
- Many other innovations, Turing Award Winner 1997



Source: Wikipedia

History – Xerox Star (1981)

- Personal computer, desktop metaphor, for office automation
- Videos:
  - https://www.youtube.com/watch?v=Cn4vC80Pv6Q, https://www.youtube.com/watch?v=ODZBL80JPqw

Source: Wikipedia, DigiBarn

Birth of HCI (1983)

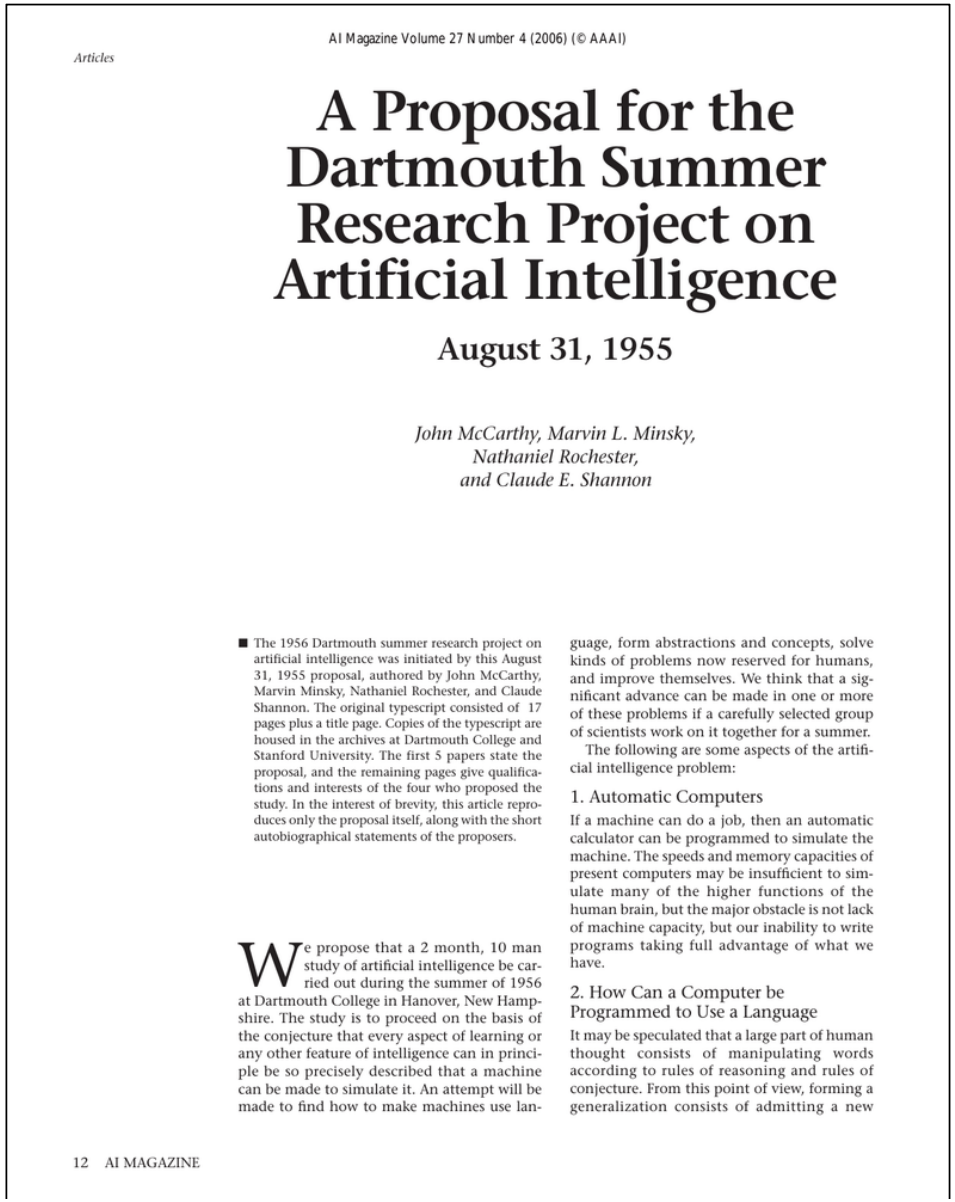
- The First ACM SIGCHI Conference
- ACM SIGCHI definition: Human-computer interaction is a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them
- Card, Moran and Newell, The Psychology of Human-Computer Interaction, 1983
  - Synergy between Psychology and Computer Science – human behavior can be understood and even modeled as an information processing activity

Graphic: sixteen-nine.net, DeWitt 2019

Artificial Intelligence

"…make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves."

[McCarthy et al. 1955]



A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence

August 31, 1955

John McCarthy, Marvin L. Minsky, Nathaniel Rochester, and Claude E. Shannon

■ The 1956 Dartmouth summer research project on artificial intelligence was initiated by this August 31, 1955 proposal, authored by John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon. The original typescript consisted of 17 pages plus a title page. Copies of the typescript are housed in the archives at Dartmouth College and Stanford University. The first 5 papers state the proposal, and the remaining pages give qualifications and interests of the four who proposed the study. In the interest of brevity, this article reproduces only the proposal itself, along with the short autobiographical statements of the proposers.

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.

The following are some aspects of the artificial intelligence problem:

1. Automatic Computers

If a machine can do a job, then an automatic calculator can be programmed to simulate the machine. The speeds and memory capacities of present computers may be insufficient to simulate many of the higher functions of the human brain, but the major obstacle is not lack of machine capacity, but our inability to write programs taking full advantage of what we have.

2. How Can a Computer be Programmed to Use a Language

It may be speculated that a large part of human thought consists of manipulating words according to rules of reasoning and rules of conjecture. From this point of view, forming a generalization consists of admitting a new

Turing Test

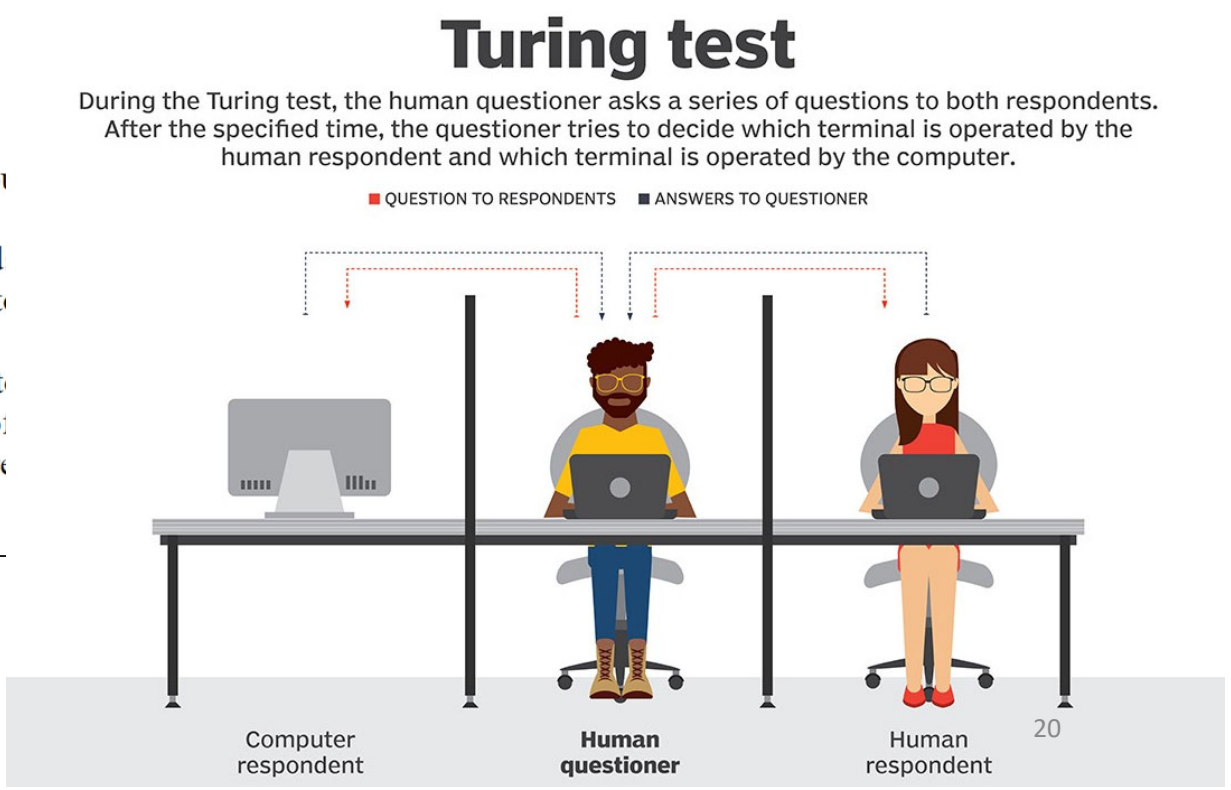
A. M. Turing (1950) Computing Machinery and Intelligence. *Mind 49:* 433-460.

COMPUTING MACHINERY AND INTELLIGENCE

By A. M. Turing

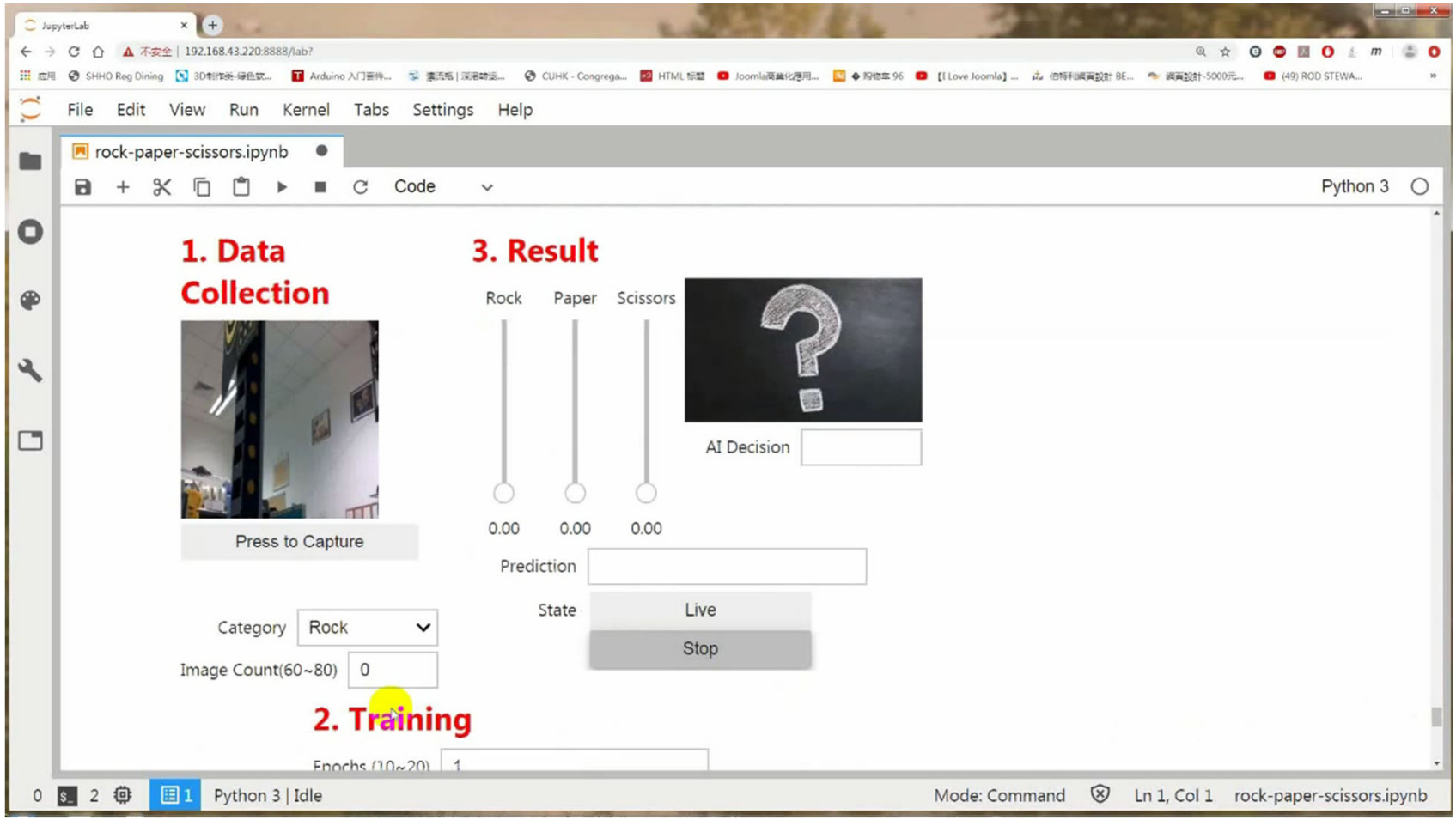
1. The Imitation Game

I propose to consider the question, "Can machines think?" This shou
definitions of the meaning of the terms "machine" and "think." The
framed so as to reflect so far as possible the normal use of the word
dangerous, If the meaning of the words "machine" and "think" are t
examining how they are commonly used it is difficult to escape the
meaning and the answer to the question, "Can machines think?" is t
statistical survey such as a Gallup poll. But this is absurd. Instead o
definition I shall replace the question by another, which is closely re
expressed in relatively unambiguous words.



Breakthroughs in AI enabled by…

Illustration: Rock Paper Scissors

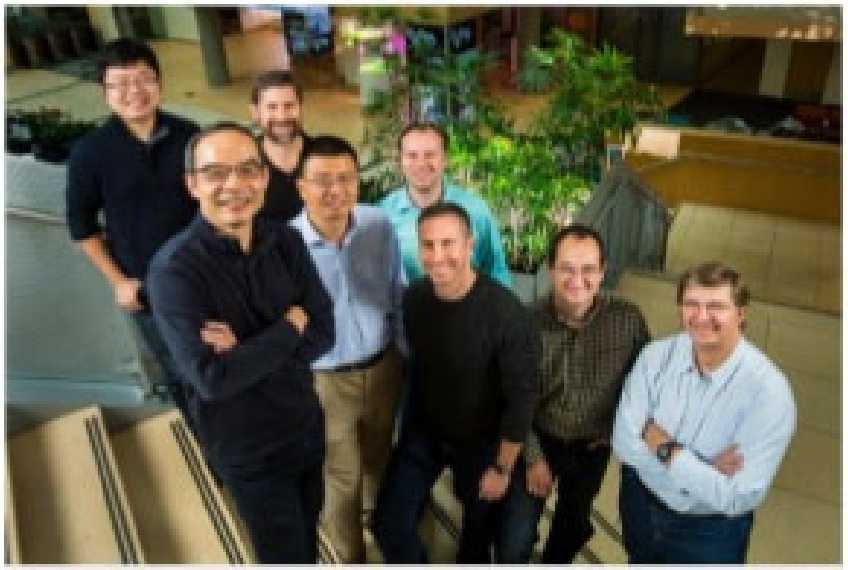


Credits: Derek Wong, Mingzhi Lyu

Illustration: Automatic Speech Recognition

Microsoft's new speech breakthrough

Historic Achievement: Microsoft researchers reach human parity in conversational speech recognition



Microsoft researchers from the Speech & Dialog research group include, from back left, Wayne Xiong, Geoffrey Zweig, Xuedong Huang, Dong Yu, Frank Seide, Mike Seltzer, Jasha Droppo and Andreas Stolcke. (Photo by Dan Deluing)

5.9% word-error rate – Human Parity

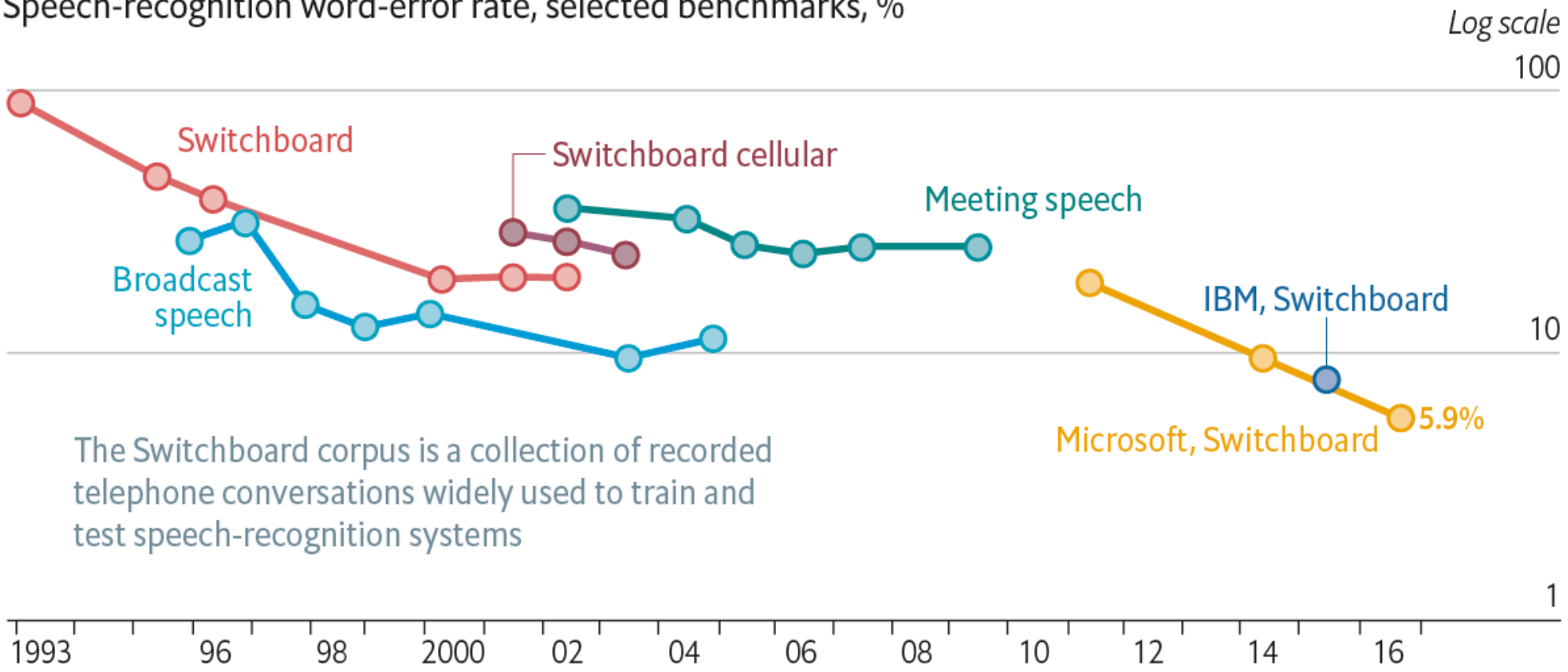
All experiments were run on CNTK

[W. Xiong, J. Droppo, X. Huang, F. Seide, M. Seltzer, A. Stolcke,
D. Yu, G. Zweig: "The Microsoft 2016 Conversational Speech Recognition System,"
http://arxiv.org/abs/1609.03528]

http://blogs.microsoft.com/next/2016/10/18/historic-achievement-microsoft-researchers-reach-human-parity-conversational-speech-recognition

Loud and clear

Speech-recognition word-error rate, selected benchmarks, %



Sources: Microsoft; research papers

Update: [August 2017] Microsoft announced reaching 5.1% error rate, comparable to professional transcribers

Illustration: Graphics Generation

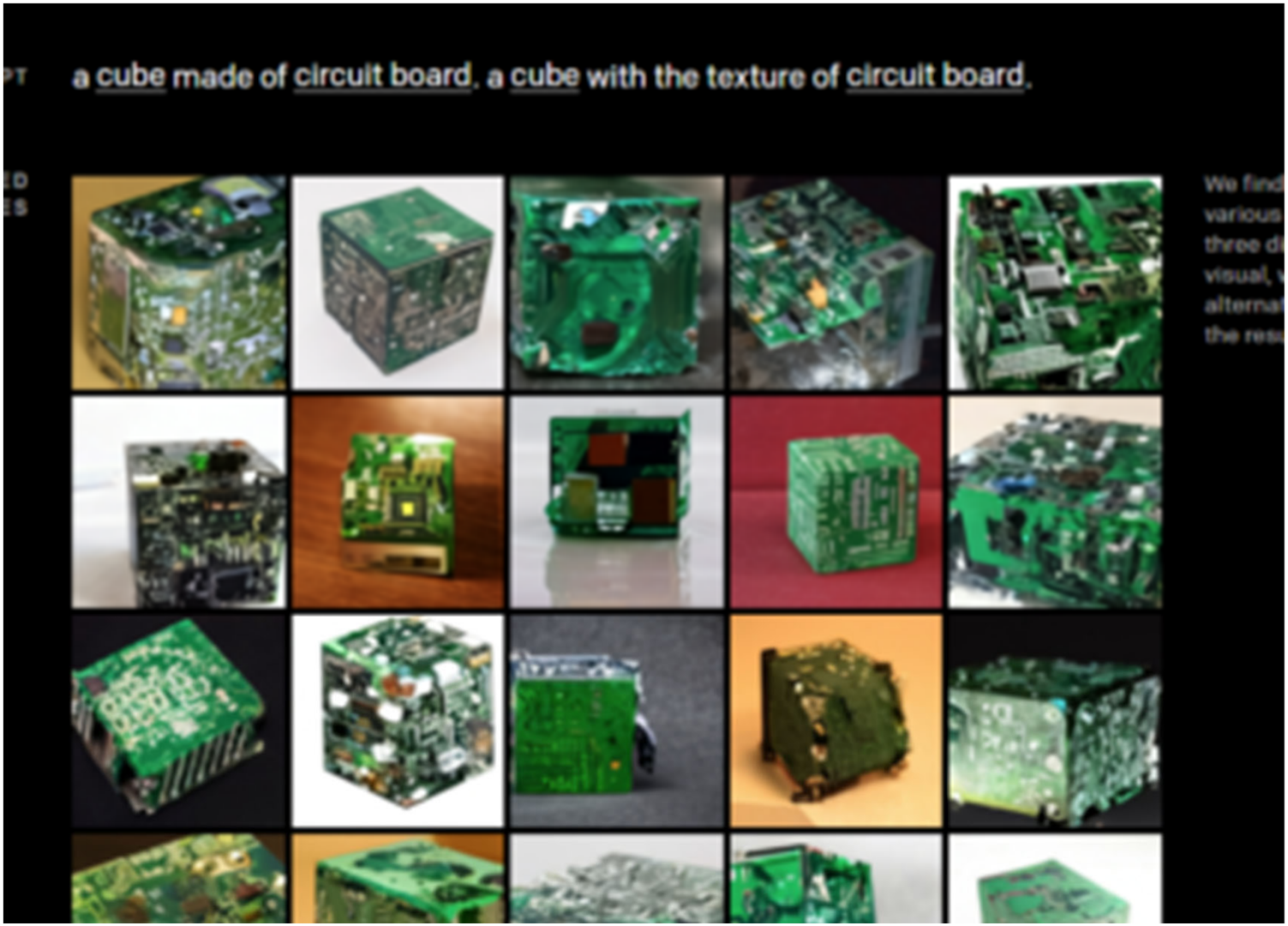


Illustration: Language Modelling

Language Models are Few-Shot Learners

Tom B. Brown* Benjamin Mann* Nick Ryder* Melanie Subbiah*

Jared Kaplan† Prafulla Dhariwal Arvind Neelakantan Pranav Shyam Girish Sastry

Amanda Askell Sandhini Agarwal Ariel Herbert-Voss Gretchen Krueger Tom Henighan

Rewon Child Aditya Ramesh Daniel M. Ziegler Jeffrey Wu Clemens Winter

Christopher Hesse Mark Chen Eric Sigler Mateusz Litwin Scott Gray

Benjamin Chess Jack Clark Christopher Berner

Sam McCandlish Alec Radford Ilya Sutskever Dario Amodei

OpenAI

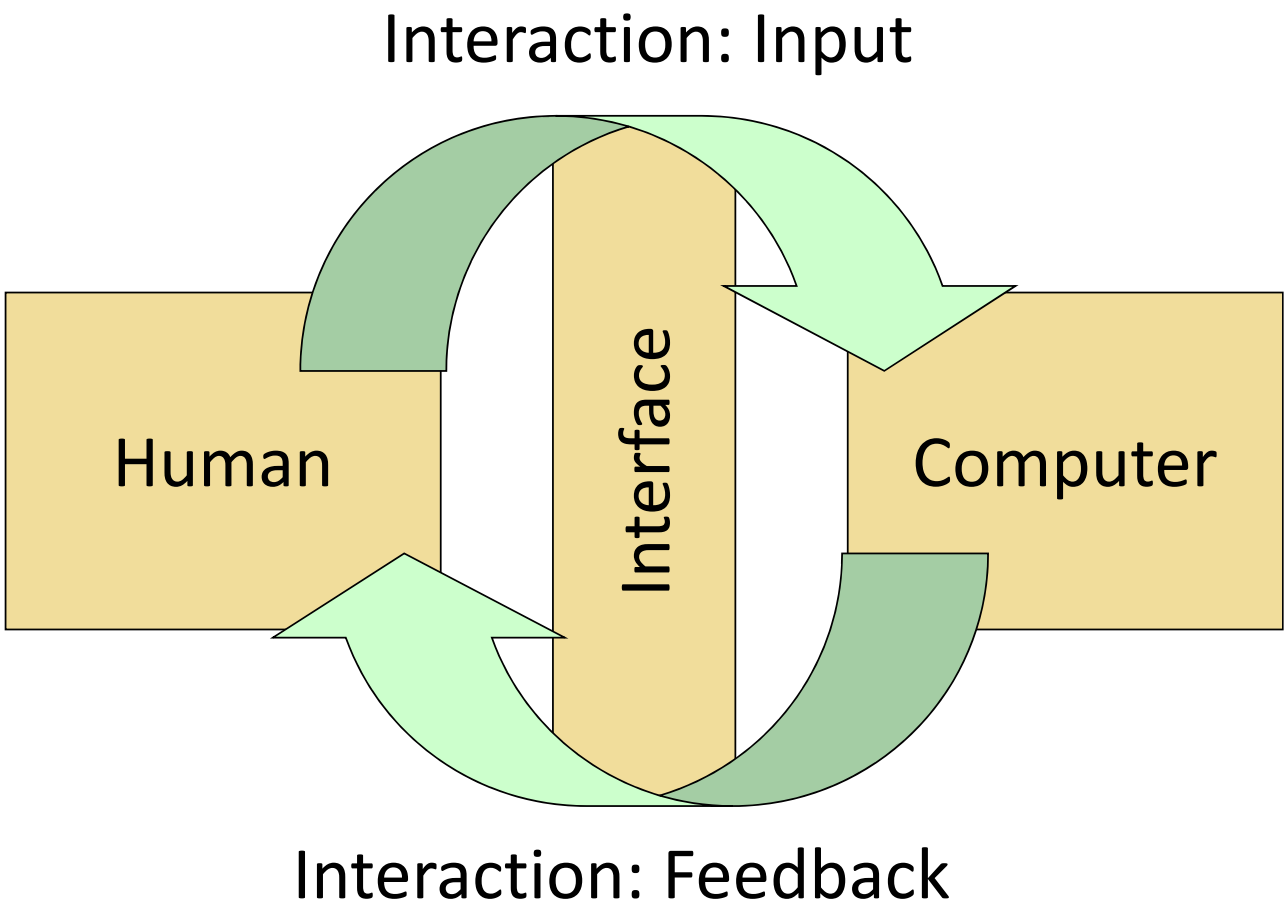
Abstract

Recent work has demonstrated substantial gains on many NLP tasks and benchmarks by pre-training on a large corpus of text followed by fine-tuning on a specific task. While typically task-agnostic in architecture, this method still requires task-specific fine-tuning datasets of thousands or tens of thousands of examples. By contrast, humans can generally perform a new language task from only a few examples or from simple instructions – something which current NLP systems still largely struggle to do. Here we show that scaling up language models greatly improves task-agnostic, few-shot performance, sometimes even reaching competitiveness with prior state-of-the-art fine-tuning approaches. Specifically, we train GPT-3, an autoregressive language model with 175 billion parameters, 10x more than any previous non-sparse language model, and test its performance in the few-shot setting. For all tasks, GPT-3 is applied without any gradient updates or fine-tuning, with tasks and few-shot demonstrations specified purely via text interaction with the model. GPT-3 achieves strong performance on many NLP datasets, including translation, question-answering, and cloze tasks, as well as several tasks that require on-the-fly reasoning or domain adaptation, such as unscrambling words, using a novel word in a sentence, or performing 3-digit arithmetic. At the same time, we also identify some datasets where GPT-3's few-shot learning still struggles, as well as some datasets where GPT-3 faces methodological issues related to training on large web corpora. Finally

Discussion

- Human-AI Interaction

HCI


Interaction: Input
Human
Interface
Computer
Interaction: Feedback

HCI

- **Human**
  - Users, single, group working together, etc.
  - His/Her mind: User tries to complete a task
  - Stationary, mobile, remote, etc.
- **Computer**
  - Not just desktop computers
  - Screen sizes? Immersive, huge, medium, small, phones, screenless, 3D
  - Systems: super computer, tablet, phone, wearables, implants, robot, chatbot, smart buildings, digital human, embodied
- **Interaction**
  - Between two or more parties
  - Communication, direct/indirect
  - Task-oriented, chatting
  - Many possibilities!

Interactive products and services

- Smartphone, tablet, computer, remote control, coffee machine, ATM, ticketing machines, printer, GPS, ereader, TV, radio, games console
- Apps, internet webpages, social media
- Robots, smart environments (ambience)
- Many possibilities!

Examples

- Accessibility for the physically challenged (https://www.youtube.com/watch?v=R2mC-NUAmMk)

Disruptive Changes!

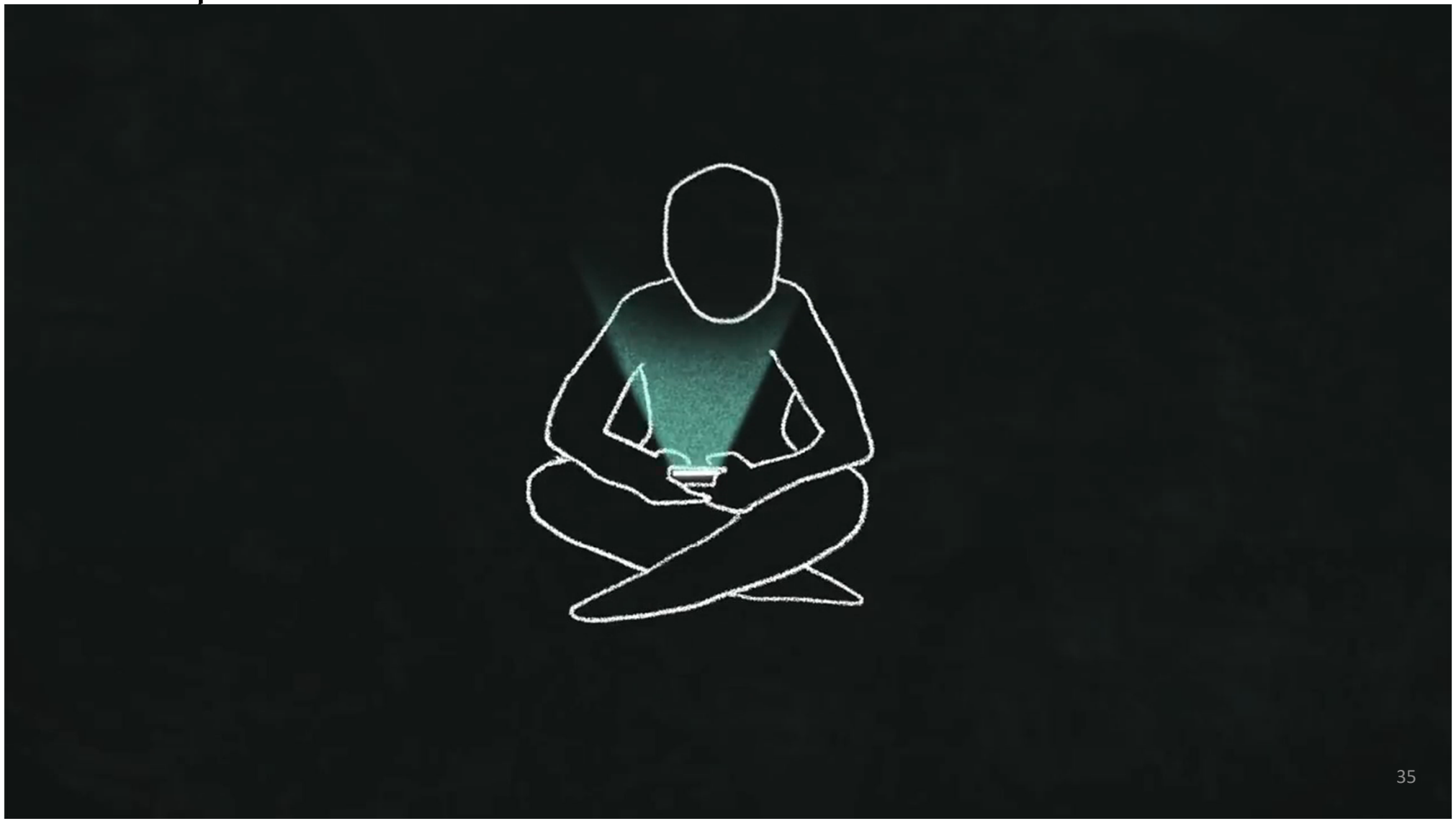
Intelligent Assistant – Embodiment

Deep Fakes

Embodiment

- https://www.youtube.com/watch?v=9owTAISsvwk (digital human)
- https://www.youtube.com/watch?v=_TmOOGx6lz4 (background)
- https://www.youtube.com/watch?v=nRyM52GQ7Ks (entertainment)
- https://www.youtube.com/watch?v=abqUMxj-1WI (pet)

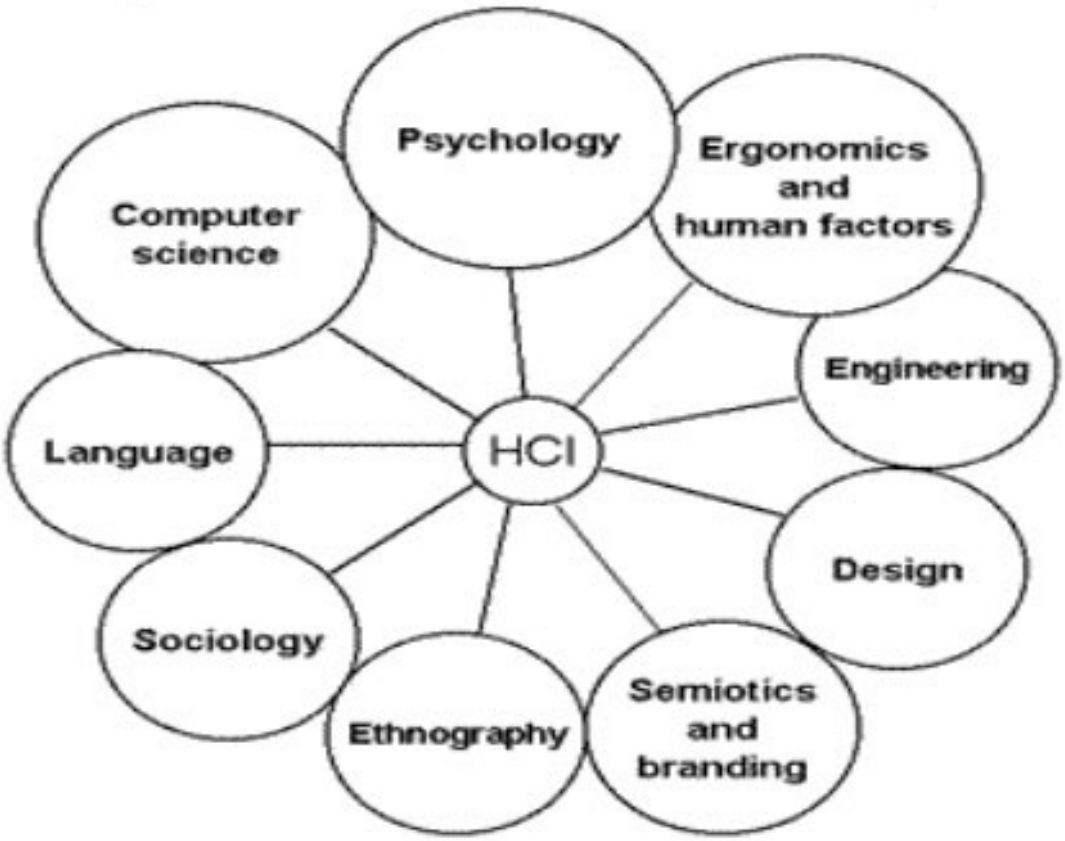
Companion

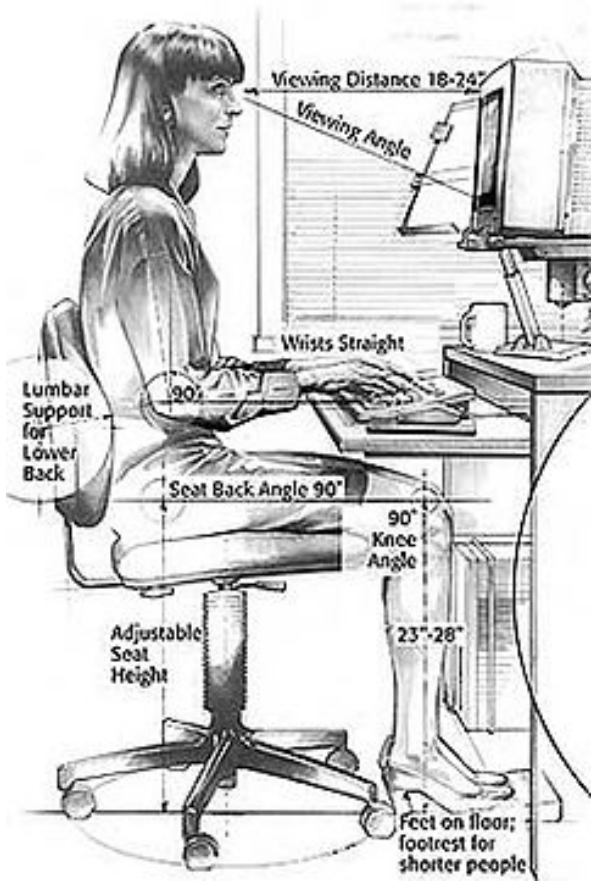


Augmented Reality / Virtual Reality

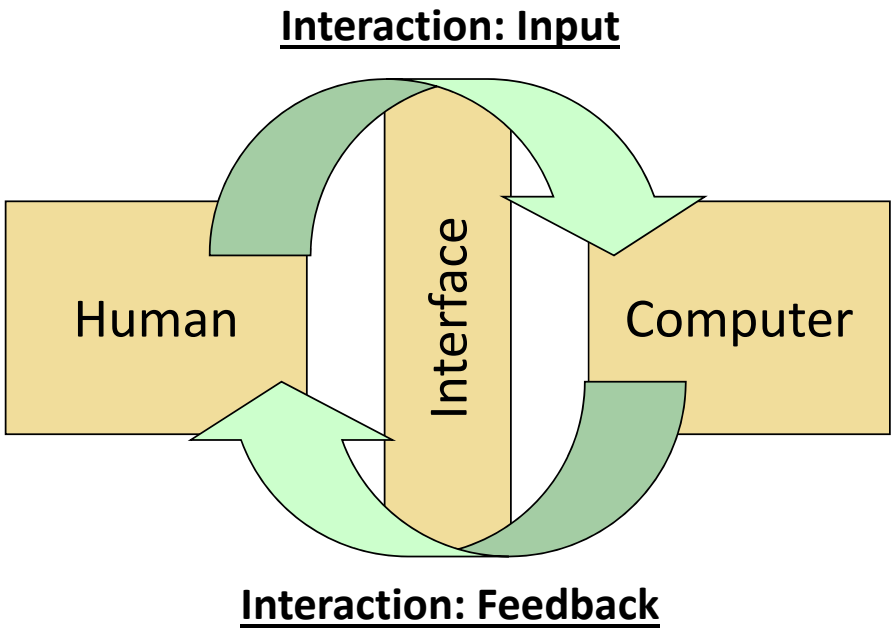
- https://www.youtube.com/watch?v=UmJriqzDUTo (Wanna Kicks)
- https://www.youtube.com/watch?v=1pWfFsX9lEM (Wanna Nails)
- https://www.youtube.com/watch?v=6OUsYMVc1Qw (Jewellery)
- https://www.youtube.com/watch?v=cDqACT5sc7U (Hololens)

A Multi-disciplinary Field





HCI (cont)

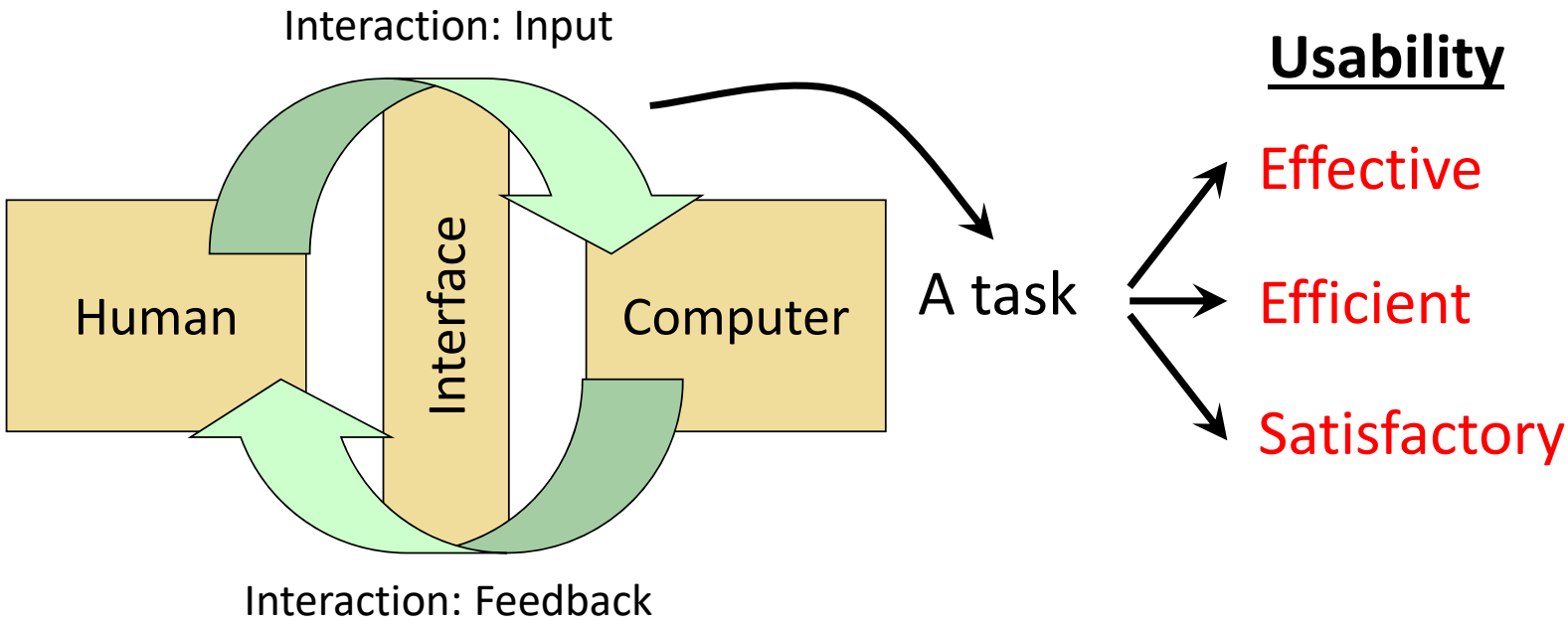


Continuous Cycle!

- Think about those games that you like
- Think about successful software / products

Possible design options?

HCI (cont)



In this course, we will look at different parts of this diagram

Usability

- **Intuitive design**: a nearly effortless understanding of the architecture and navigation of the site
- **Ease of learning**: how fast a user who has never seen the user interface before can accomplish basic tasks
- **Efficiency of use**: How fast an experienced user can accomplish tasks
- **Memorability**: after visiting the site, if a user can remember enough to use it effectively in future visits
- **Error frequency and severity**: how often users make errors while using the system, how serious the errors are, and how users recover from the errors
- **Subjective satisfaction**: If the user likes using the system

http://www.usability.gov/basics/index.html

User Experience

- **Useful**: Your content should be original and fulfill a need
- **Usable**: Site must be easy to use
- **Desirable**: Image, identity, brand, and other design elements are used to evoke emotion and appreciation
- **Findable**: Content needs to be navigable and locatable onsite and offsite
- **Accessible**: Content needs to be accessible to people with disabilities
- **Credible**: Users must trust and believe what you tell them

Source: https://www.usability.gov/what-and-why/user-experience.html

www.baddesigns.com

Example Design

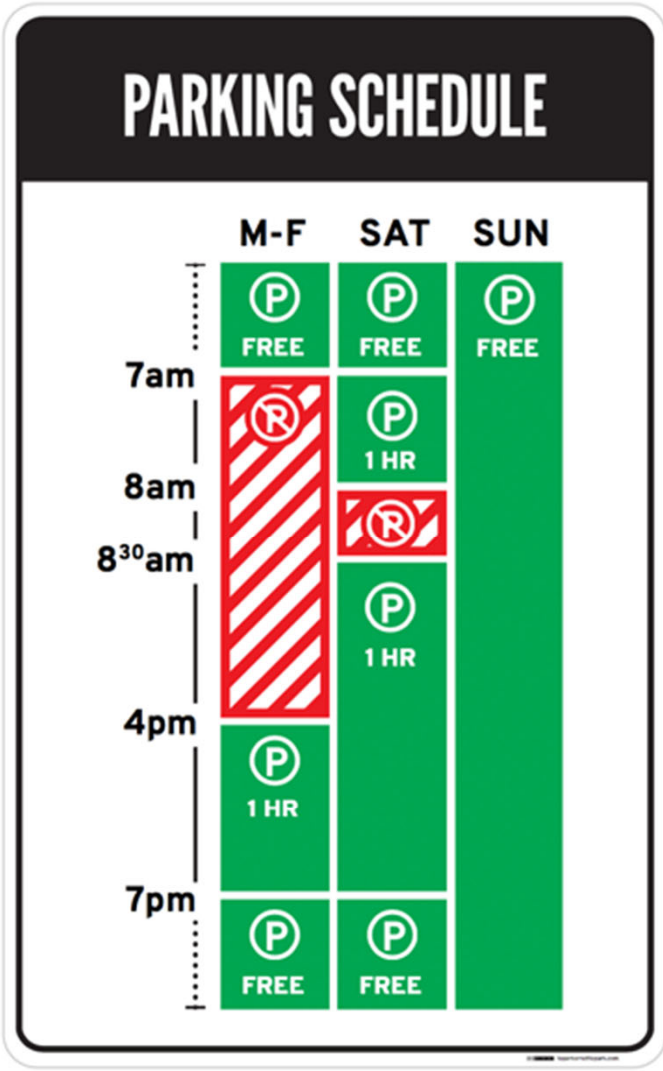
Tachometer, Speedometer
Prevent confusions

www.baddesigns.com

Example Design

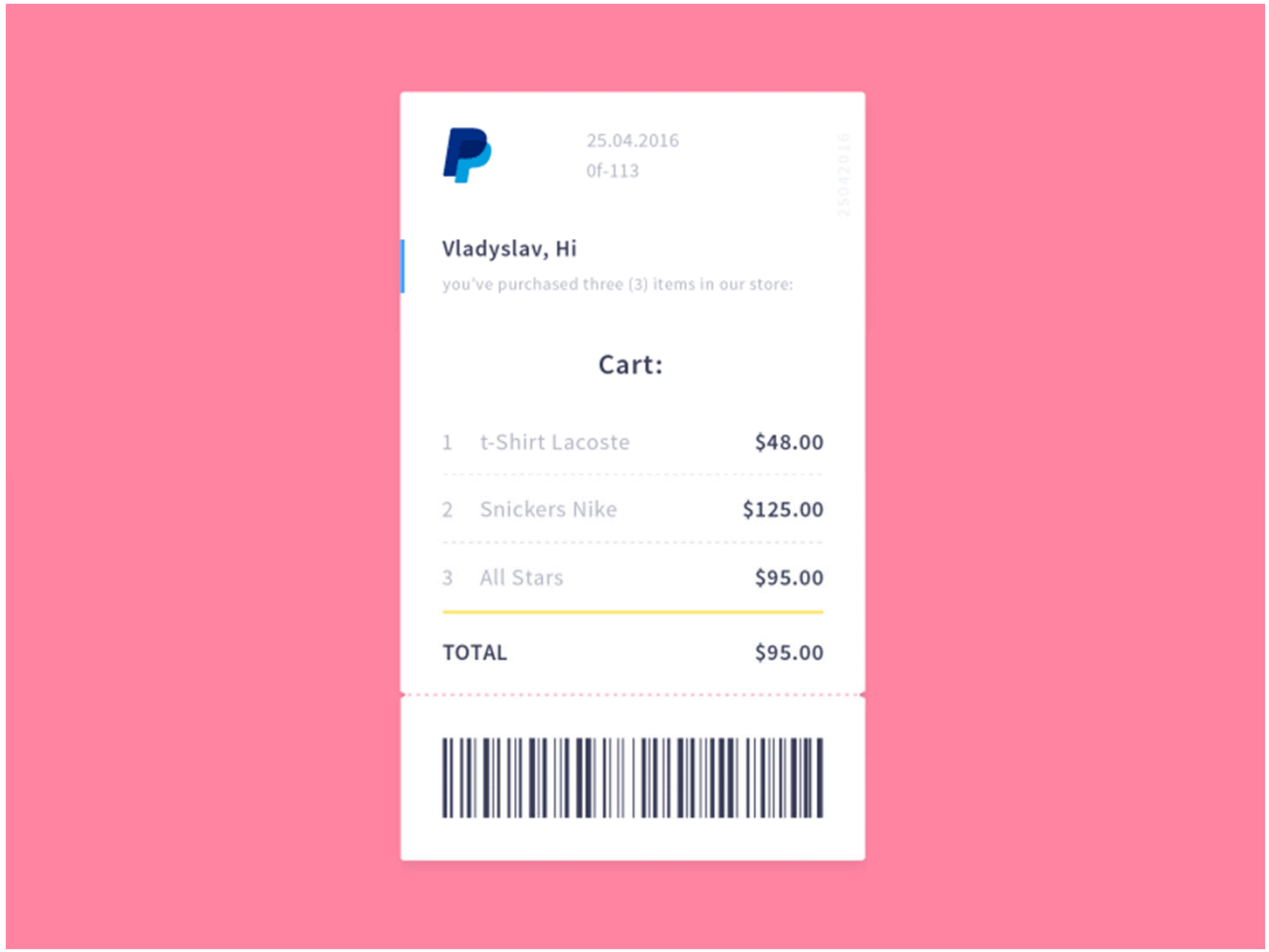
www.baddesigns.com

Example Designs



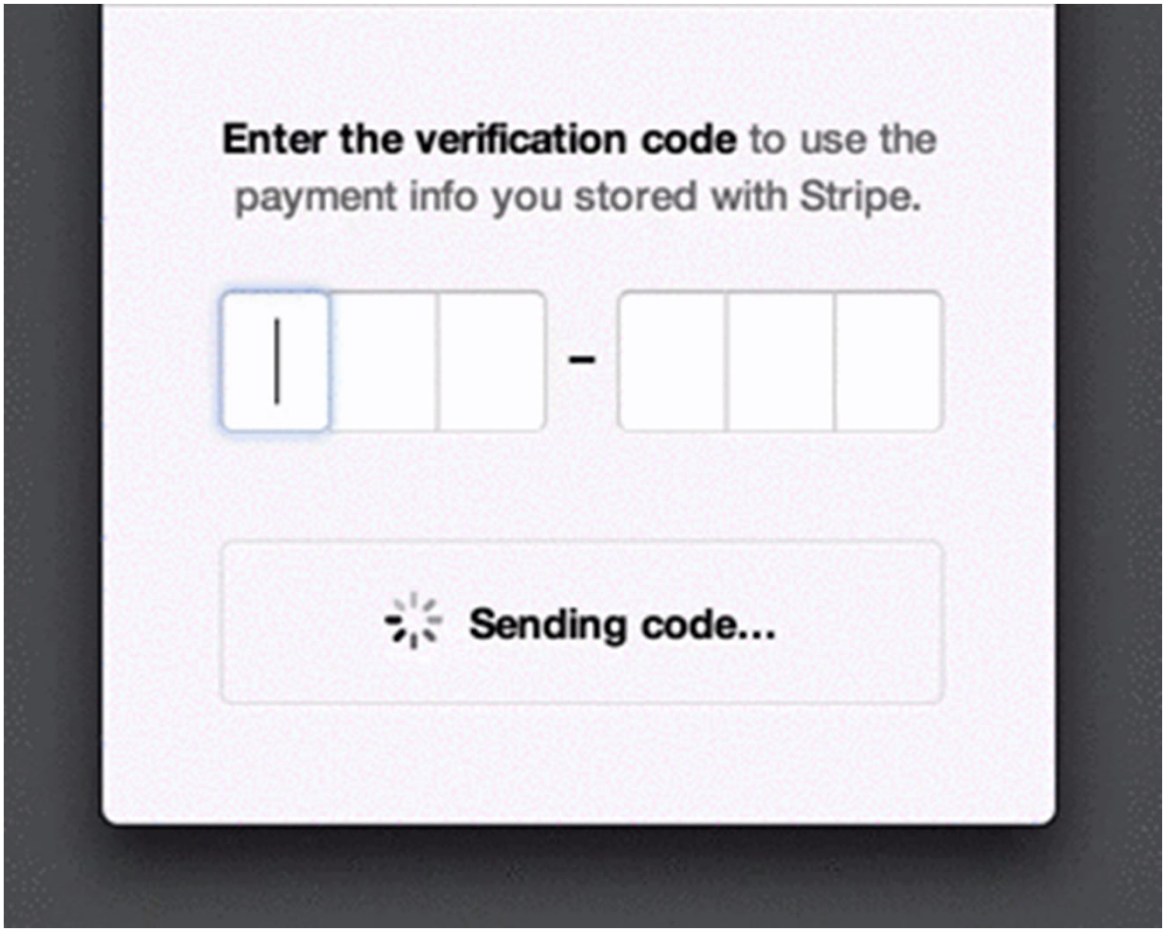
https://www.interaction-design.org

Example Design



https://www.interaction-design.org

Example Design



Basic UI Design Goal

"We all want to design good and useful systems"

- There are practical and principle ways to achieve this. The first, before any work begins, is to think through the problem/task.

- "Usability Requirements"
  - Needs *thoughtful planning*, sensitivity to *user needs*, devotion to *requirement analysis* and diligent *testing*, while keeping *within budget and schedule*
  - This is more than just having the goal of "user-friendly"