

Artificial Intelligence

PHIL 101
Self, World, & God

What is the imitation game?

**Will computers ever pass
the "Turing test?"**

*Will it prove anything
if they do?*



Alan Turing



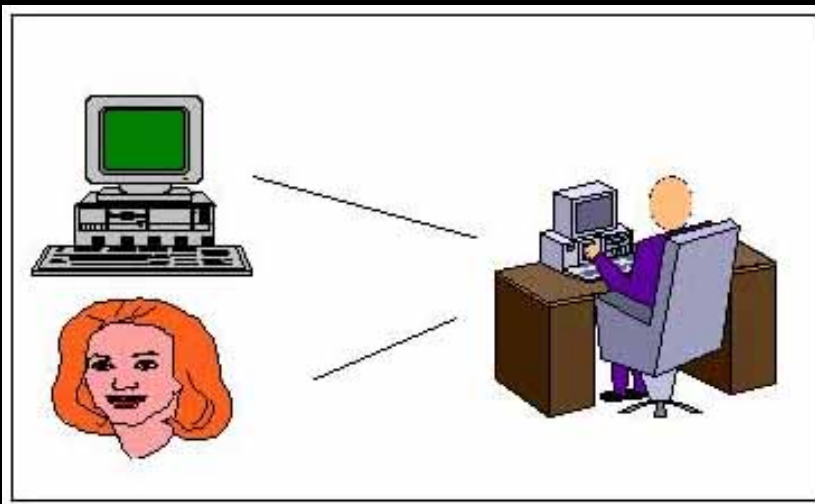
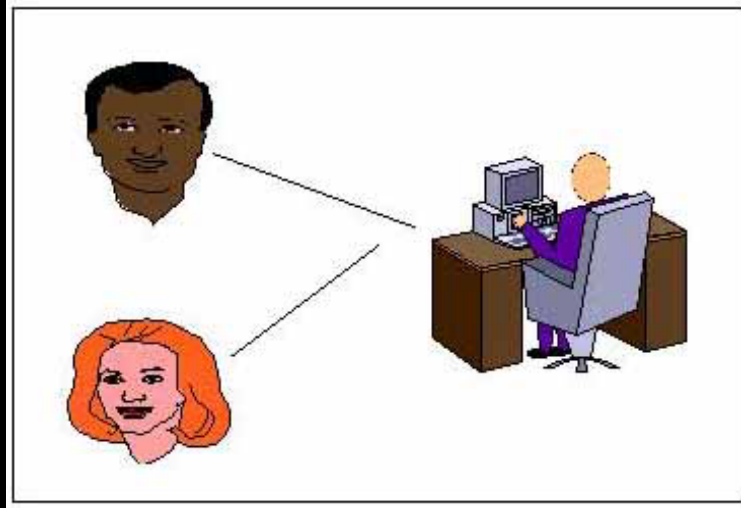
- **1912-54**
- **Childhood**
- **Cambridge**
- **1936 paper**
- **Enigma project**
- **Computer design**
- **1950 paper**
- **Arrest**
- **Death**

Can Machines Think?

- **“Computing Machinery and Intelligence” *Mind* 1950**
- **Turing’s question**
- **What is a ‘machine’?**
- **A digital computer**
- **What is ‘thinking’?**
- **Passing the “Turing test”**
- **Turing’s prediction**

The Imitation Game

- **The “Turing test”**
- **Participants in the Imitation Game**
 - **A – a man**
 - **B – a woman**
 - **C – the interrogator, either sex**
- **Goals (winning)**
 - **A – be misidentified as the woman**
 - **B – be correctly identified as the woman**
 - **C – correctly identify the woman**
- **Anything goes for all three**



The Turing Test

- **Strategies for detecting the machine**
- **Things machines do too well**
- **Emotional things**
- **Perceptual things**
- **Artistic and aesthetic things**
- **Humor**
- **Normal conversation**
- **Turing's prediction**
- **What went wrong?**

A Brief History of AI

- **Early successes**
 - **Mathematical problems**
 - **Checkers and chess**
- **General intelligence**
- **Problem of language**
- **Problem of common sense**
- **Frame problem**
- **Modest reconsideration of AI**
- **Expert systems**
- **Eastern's grant in 1987**

John Searle



- 1932 -
- U. C. Berkeley
- Moral philosophy
- Philosophy of language
- Philosophy of mind
- AI
- "The Chinese Room"
- Weak AI
- Strong AI



文學、藝術、人文類
 大千世界 讀書人 橄欖樹 花招 拱橋樓 饒恒遠曲 蕭軍38小書
 新語集 杏花村 皇冠 盲口 無聊眼 五言六社雜志 大衆舞臺圖展

政經類
 北京之春 大西 東亞外交與國際關係 民主中國 四大商場 臺北政壇
 中國與世界 當代中國研究 中國之春 真言 新世紀 經濟學報(周刊)

The Chinese Room

- Thought experiments
- Jeremy doesn't know Chinese
- Receives the message
- Instruction book is in English
- Produces message
- Passes Turing test
- But, Jeremy doesn't understand
- He has simply manipulated symbols
- Strong AI is impossible

John Searle's "Chinese Room"

Chinese symbols received on paper:

文學、藝術、人文類
 大千世界 讀書人 撒撒樹 花招 拱橋樓 絨帽道曲 蕭瑟38小書
 新語集 杏花村 皇冠 寶口 無聊眼 五言六社雜志 大衆舞臺圖風



Person in box:
 1. Looks up symbols in instruction book
 2. Follows directions for which symbols to write down
 3. Passes written symbols back outside the box
 This is equivalent to executing a program.

(According to Searle)
 Person does not understand Chinese

Response passes Turing Test:

政經類
 北京之春 區區 東亞外交與國際關係 民主中國 四大商場 臺北政壇
 中國與世界 當代中國研究 中國之春 共言 新世紀 經濟學報(周刊)

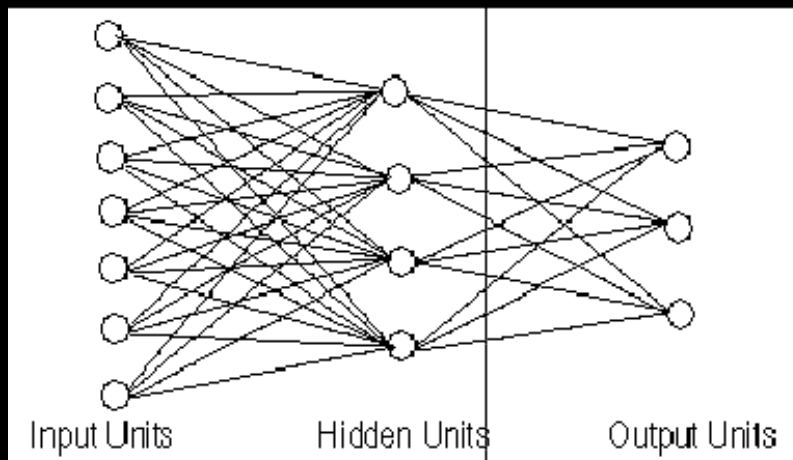
The AI Debate

- **The optimists**
 - Alan Turing
 - Doug Lanat
 - The “connectionists”
- **The pessimists**
 - Joe Weizenbaum
 - Hubert Dryfus
 - John Searle
- **What’s at stake?**
- **How humans think**
- **The representational theory of mind**

Return to the Chinese Room

- **Empirical problems**
- **Explanatory problems**
- **Logical problems**
- **Passing the Turing test would never count as strong AI**
- **Responses**
 - Systems reply
 - Luminous room argument
- **What would we think if a computer did pass the Turing test?**

Connectionism



Connectionism

- **McCulloch and Pitts**
- **Fickle finger of fate**
- **Resurgence in the 1980s**
- **Net-talk**
- **Past tenses**
- **Mines versus rocks**

Two On-going Projects

- **Cyc**
 - **Doug Lanet**
 - **My grant**
 - **The project**
- **Cog**
 - **M.I.T.**
 - **The project**
- **Any reason for optimism?**

