
**PennState**
Institute for Computational
and Data Sciences

Center for Artificial Intelligence Foundations & Scientific Applications
Artificial Intelligence Research Laboratory

**PennState**
Clinical and Translational
Science Institute



ARTIFICIAL INTELLIGENCE

The Very Idea

Vasant G. Honavar
Dorothy Foehr Huck and J. Lloyd Huck Chair in Biomedical Data Sciences and Artificial Intelligence
Professor of Informatics and Intelligent Systems
Professor, Informatics, Computer Science, Bioinformatics & Genomics and Neuroscience Graduate Programs
Director, Artificial Intelligence Research Laboratory
Director, Center for Artificial Intelligence Foundations and Scientific Applications
Director of Strategic Initiatives, Institute for Computational and Data Sciences
Informatics Lead, Clinical and Translational Sciences Institute
Director of AI and Strategic Initiatives, NSF National Center for Synthesis in Molecular and Cellular Sciences
Pennsylvania State University

vhonavar@psu.edu
<http://faculty.ist.psu.edu/vhonavar>
<http://ailab.ist.psu.edu>

**PennState**
College of Information
Science and Technology

AI 100 Fall 2025

Vasant G Honavar

Theories of intelligence

- Theories of intelligence are primarily concerned with
 - Identifying the **major components of intelligence** e.g., knowledge, skills, goals
 - **How they are represented and used**
 - **Mechanisms, processes, and the role of different components**
 - **Interactions between components**

Performance theories of intelligence

- View intelligence as a collection of abilities
- Measure an individual's performance on a set of tasks
- Driving questions
 - How can one test for the presence or degree of intelligence?
 - What are the essential components or functionalities a system capable of intelligence?

Structural/functional theories of Intelligence

- What are the mechanisms that underlie intelligence?
 - Computational theory of intelligence posits that intelligence is realized by computational means
 - Biological theory of intelligence treats intelligence as a product of
 - brain structure and brain function and
 - their relationship

Cognitive theories of intelligence

- Focus on the processes that underlie intelligence
- The processes range from simple to complex
 - Comparing items
 - Memory recall
 - Problem solving
- Focus on certain aspects of processes
 - Speed
 - Accuracy
- ...

Contextual theories of intelligence

- Focus on intelligence demonstrated in a given social, cultural, or environmental context
- Driving questions
 - How is one's intelligence adapted to one's social context?
 - How is one's intelligence adapted to the environment with which one must cope with?
 - How is one's intelligence tailored for one's cultural context?
- Contextual theories explain
 - A student adapting to a new school
 - A farmer adapting to climate change
 - An immigrant adapting to a foreign culture

Existence theories of intelligence

- Focus on the necessary and/or sufficient conditions of intelligent behavior
 - Under what conditions can one
 - Learn a concept from data?
 - Learn an adversarial game like Chess?
 - Repair faulty plans?
 - Cooperate with others to solve a complex problem?
 - Act rationally?
 -

Developmental theories of intelligence

- **Piaget** takes constructivist view of intelligence
 - Intelligence is the product of interaction with the environment
 - Forming and refining mental structures or schemas
 - **Example:** A child learns that a zebra is different from a horse and refines the schema.
 - Knowledge is actively created not passively received
- Piaget hypothesized four **developmental stages**
 - **Sensorimotor stage** (0-2 years)
 - object permanence
 - **Pre-operational stage** (2-7 years)
 - language
 - **Concrete operational stage** (7-11 years)
 - logical reasoning tied to concrete objects and events
 - **Formal operational stage** (12+ years)
 - abstract, hypothetical and deductive reasoning

Two major theories

Sternberg's Triarchic Theory

- Triarchic Theory posits that there are three aspects of intelligence
 - Analytical Intelligence (book smart)
 - Creative Intelligence
 - Practical Intelligence (street smart)

Sternberg's Triarchic Theory: Analytical Intelligence

- **Analytical Intelligence** (book smart) refers to the ability to decompose complex problems into manageable subproblems and solving them using logical reasoning
- **Examples**
 - Solving mathematical problems
 - Crafting legal arguments
 - Assessing risks and benefits of alternative actions
- Analytical intelligence relies on the ability to
 - **Acquire and organize knowledge and skills**
 - **Deploy the knowledge and skills effectively** to solve the problems
 - Knowing Newton's laws of physics does not immediately translate into a facility for solving complex physics problems

Sternberg's Triarchic Theory

- **Creative Intelligence** refers to the ability to draw on experience to solve novel or unique problems
- **Examples of creative intelligence**
 - Conceiving and designing new mode of transportation
 - Wright brothers' invention of the airplane
 - Creating a new architectural style
 - Le Corbusier's Modernist architecture
 - Creating new music or an entirely new musical genre
 - Charles Buddy Bolden and his band members' introduction of Jazz music in the early 1900's
 - Creating a new style of art
 - Claude Monet's and other Paris-based artists' introduction of impressionism in the 1860's

Sternberg's Triarchic Theory

- **Practical Intelligence** (street smart) refers to
 - Ability to adapt to one's environmental or cultural contexts
 - Shape the environment to meet personal or societal needs
- **Examples of practical intelligence**
 - A mechanic figuring out how to fix a car
 - A farmer figuring out how much fertilizer to use
 - A diplomat negotiating a peace treaty
 - A waitress noticing a restless child and offering crayons

Gardner's theory of multiple intelligences

- Traditional psychometrics reduces intelligence to a single number (IQ)
- Gardner argued that
 - Intelligence cannot be measured by a single number
 - Doing so fails to reflect or distinguish between the rich variety of human abilities
 - Intelligence should be understood in terms of different capacities, each valuable on its own.
 - There are significant variation across individuals in their levels of strength and weakness across the different intelligences

Gardner's theory of multiple intelligences

- **Multiple intelligences** according to Gardner
 - **Verbal-linguistic intelligence** – writers, poets, lawyers
 - **Logical-mathematical intelligence** – mathematicians, scientists
 - **Spatial-visual intelligence** – architects, artists, pilots
 - **Bodily-kinesthetic intelligence** – athletes, dancers, surgeons
 - **Musical intelligence** – musicians, composers
 - **Intrapersonal intelligence** – philosophers, psychologists
 - **Interpersonal intelligence** – teachers, leaders
 - **Naturalist intelligence** – farmers, environmentalists

On intelligence

- Intelligence is hard to define
- Can we recognize it when we see it?
- Can we measure intelligence?
- Can we tell if a system is intelligent?

Measuring Intelligence

Terman-Merrill revision of the Binet-Simon intelligence test includes questions in the following categories:

- Obey simple commands
- Identify object by use
- Repeat digits
- Respond to pictures
- Repeat digits reversed
- Memorize key events and actors in stories
- Find absurdities in pictures
- Recognize pictorial objects
- Comprehend text
- Answer opposite analogies
- Identify alike and different pictures
- Memorize details in sentences
- Use vocabulary
- Complete (human) picture
- Discriminate animal pictures

Intelligence tests have limited utility

- What the test scores measure is neither innate nor permanent
- The test offer a rough scale for identifying children needing help
- The test is useless for ranking "normal" children
- Tests are designed with predetermined answers
- Child's ability to score well depends on cultural social and economic factors

Are there good tests for machine intelligence?

- Can we simply give machines the same IQ tests we give to humans?
 - We know the problems with IQ tests
 - We have AI systems that can get high score on some IQ tests while failing to exhibit most attributes of intelligence
 - We can design AI systems that can ace high school math tests etc.
- Turing had an alternative idea – The Turing Test

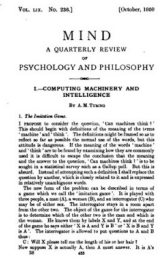
If a machine could think, how could we tell?

• Turing's idea

- If in a natural language dialog, the responses from the computer were indistinguishable from that of a human, the computer could be said to be thinking

• Turing's prediction

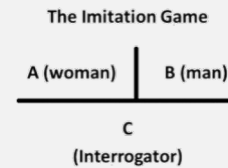
- Turing predicted that by the year 2000 a computer program would be able to fool the average questioner for five minutes about 70% of the time
- Prediction was off by a couple of decades
- Large language models passed the test by 2022 or so



Turing, A.M Computing Machinery and Intelligence. *Mind*, Volume LIX, Issue 236, 1950, Pages 433-460

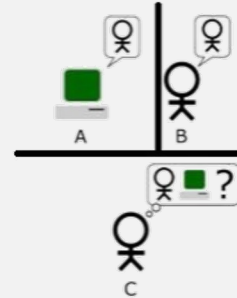
Turing Test – Inspired by the Imitation Game

- Turing test is a modification of **imitation game** that was popular among Cambridge students
 - There are three people in three separate rooms. A woman (A), a man (B), and an interrogator (C)
 - None of them knows each other's gender.
 - Interrogator can pass typed questions to A or B
 - A and B can communicate **only** with C by passing typed answers.
 - C asks questions to A and B
 - A and B seek to convince C that he/she is a woman
 - If B convinces C that B is a woman at least. as often as A does, B wins the imitation game



Turing Test

- Replace one of the two players (say A) in the imitation game by a computer
- Interrogator can pass typed questions to A or B
- A and B can communicate **only** with C by passing typed answers.
- C asks questions to A and B
- A and B seek to convince C that he/she is a human
- If A convinces C that it is a human as often as B does, A (the computer) passes the Turing Test




Turing Test – A slight variation


- Place a **computer in one room** and a **panel of human interrogators** in another room
- The interrogators do not know whether a human or a computer is in the first room
- Interrogators on the panel can pass typed questions to the first room
- The occupant of the room returns typed answers to the panel of interrogators
- **If the computer can convince at least 50% of the panel of interrogators that it is a human, it passes the Turing Test**
- Why 50%? Why not?

Top Hat Question

- Turing Test is adequate as a test of intelligence. True or False?


 PennState
Institute for Computational
and Data Sciences

Center for Artificial Intelligence Foundations & Scientific Applications
Artificial Intelligence Research Laboratory

 PennState
Clinical and Translational
Science Institute

Turing Test Inspired the Work on Chatbots

- Chatbots are computer programs which converse in a natural language
- Many chatbots were developed to
 - Explore how humans-machine communication in natural language
 - To pass - or ridicule – the Turing test
- Some examples of Chatbots
 - ELIZA, developed by Joseph Weizenbaum of MIT in 1966
 - ChatGPT developed by Open AI
 - Claude developed by Anthropic
 - Gemini developed by Google
 - Llama developed by META
 - Qwen developed by Alibaba
 - DeepSeek
 - Mistral

 PennState
College of Information
Science and Technology

AI 100 Fall 2025

Vasant G Honavar

ELIZA – The first chatbot



- In 1966 Joseph Weizenbaum of MIT wrote the first chatbot
- He called it ELIZA and cast it in the role of a Rogerian psychiatrist
- Sample conversation with ELIZA
 - Person: Men are all alike.
 - Chatbot: IN WHAT WAY?
 - Person: They're always bugging us about something or other.
 - Chatbot: CAN YOU THINK OF A SPECIFIC EXAMPLE?
 - Person: Well, my boyfriend made me come here.
 - Chatbot: YOUR BOYFRIEND MADE YOU COME HERE?
 - Person: He says I'm depressed much of the time.
 - Chatbot: I AM SORRY TO HEAR YOU ARE DEPRESSED.
- <https://eliza.botlibre.com>

How ELIZA works

- ELIZA simulates a conversation between a psychiatrist and her and patient
- ELIZA mirrors the patient's feelings by recasting the patient's statements as questions
 - Uses a small database of words and phrases
 - Looks for certain patterns of words in the input e.g., I want, because, always
 - I want to run away from my parents
 - Break the sentence into parts (based on the matching keyword)
 - Modify and assemble the parts into a response based on a predefined set of rules
 - Why do you want to run away from your parents?
 - When no rule matches, ELIZA produces a generic response
 - I see

How ELIZA works

- ELIZA was designed to simulate a conversation between a psychiatrist and her and patient
- ELIZA is as effective as it is –
 - not because it understands English
 - or because it knows psychotherapy
 - but because humans read so much meaning into what is said
- ELIZA fools human into interpreting the result as conversation
- Modern chatbots are not that much better!

ELIZA

- To Weizenbaum's horror
 - Some users were convinced that they were really talking to a person
 - Some psychiatrists were willing to try ELIZA in actual therapy setting

Chatbots have come a long way

- Large language models and their multi-modal variants
 - Google's Gemini
 - Open AI's ChatGPT
 - Anthropic's Claude
 - Deepseek
 - Meta's Llama
 - Alibaba's Qwen
 - Mistral
 - and many more...
- Arguably the Turing Test has been passed
- Have we built intelligent machines?
- More on this later

**PennState**
Institute for Computational
and Data Sciences

Center for Artificial Intelligence Foundations & Scientific Applications
Artificial Intelligence Research Laboratory

**PennState**
Clinical and Translational
Science Institute

The danger of chatbots

Account ▾

Published Aug. 26, 2025 Updated Aug. 27, 2025

A Teen Was Suicidal. ChatGPT Was the Friend He Confided In.

More people are turning to general-purpose chatbots for emotional support. At first, Adam Raine, 16, used ChatGPT for schoolwork, but then he started discussing plans to end his life.



**PennState**
College of Information
Science and Technology

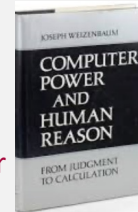
AI 100 Fall 2025

Vasant G Honavar

31

Weizenbaum foresaw the dangers of Chatbots

- To Weizenbaum's horror
 - Some users of ELIZA were convinced that they were really talking to a person
 - Some psychiatrists were willing to try ELIZA in actual therapy setting
- Weizenbaum became concerned about the ethics of AI and eventually one of the most fervent critics of AI
- Large Language Models of today, e.g., ChatGPT etc. suffer from the same problem as ELIZA
- Chatbots have their uses, but they must be designed and used with ethical oversight

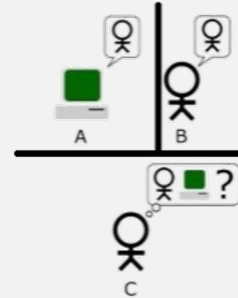


Chatbots have come a long way

- Large language models and their multi-modal variants
 - Google's Gemini
 - Open AI's ChatGPT
 - Anthropic's Claude
 - Deepseek
 - Meta's Llama
 - Alibaba's Qwen
 - Mistral
 - and many more...
- Arguably the Turing Test has been passed
- Have we built intelligent machines?

Turing Test Revisited

- We have a computer A in one room and a human B in another room
- Interrogator can pass typed questions to A or B
- A and B can communicate **only** with C by passing typed answers.
- C asks questions to A and B
- A and B seek to convince C that he/she is a human
- If A convinces C that it is a human as often as B does, A (the computer) passes the Turing Test




Turing Test Revisited


- Exercise
 - Chatbots like ChatGPT now easily pass the Turing Test
 - Does this mean we have succeeded in building intelligent machines?
 - Is the Turing test a good test for determining whether a computer can think?
 - Why or why not?
 - Reflect on this for 2-3 minutes, discuss with your neighbors
 - I will call on a few people at random to summarize what they concluded

Critique of the Turing Test

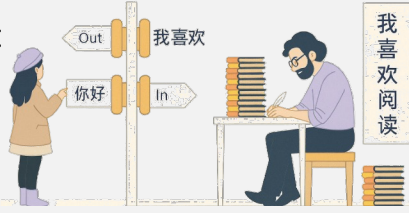
- Supporters
 - Human-like interaction is essential to human-like intelligence
- Opponents
 - Passing the Turing Test is neither necessary nor sufficient for intelligence
 - It is not necessary for program to talk to us in our language for it to be intelligent
 - There are humans that would fail the Turing test
 - Searle's Chinese Room Argument

**PennState**
Institute for Computational
and Data Sciences


Center for Artificial Intelligence Foundations & Scientific Applications
Artificial Intelligence Research Laboratory

**PennState**
Clinical and Translational
Science Institute

Searle's Chinese Room Argument



- You are looking at a room
- You see some Chinese friends of yours come to the room and slip a piece of paper into the room through a slot
- They wait a while until the same piece of paper comes back out of a room through a second slot
- You soon learn that the questions slipped into the room were written in Chinese
- Suppose the question was
 什麼是生命？
 What is life?

**PennState**
College of Information
Science and Technology

AI 100 Fall 2025

Vasant G Honavar



PennState
Institute for Computational
and Data Sciences

Center for Artificial Intelligence Foundations & Scientific Applications
Artificial Intelligence Research Laboratory



PennState
Clinical and Translational
Science Institute

Searle's Chinese Room Argument



- When the paper is later passed out of the room, your Chinese friends find that an answer to their question written on the slip

生命是一系列自然而自發的改變。不要抗拒它們——那隻會帶來悲傷。
讓現實成為現實。讓事情以任何他們喜歡的方式自然地向前發展。

Life is a series of natural and spontaneous changes. Don't resist them — that only creates sorrow. Let reality be reality. Let things flow naturally forward in whatever way they like.
- The answer is in Chinese and your friends, all fluent in Chinese agree that it is, in fact, a wise answer to their question — a direct quote from the writings of the great Lao-Tzu himself



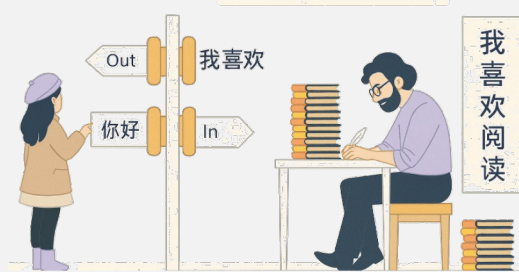
PennState
College of Information
Science and Technology

AI 100 Fall 2025

Vasant G Honavar

Searle's Chinese Room Argument

- Your Chinese friends continue to pass questions in Chinese into the room
- They receive what seem to be intelligent answers to their questions in Chinese
- They conclude that the person inside the room who is answering their questions must be an intelligent person who understands Chinese



**PennState**
Institute for Computational
and Data Sciences

Center for Artificial Intelligence Foundations & Scientific Applications
Artificial Intelligence Research Laboratory

**PennState**
Clinical and Translational
Science Institute

Searle's Chinese Room Argument



Whoever or whatever is in that room is an intelligent Chinese Speaker!

I'm just manipulating squiggles and squiggles to produce Chinese language behaviour. But I don't understand Chinese. The rule book is in English.

我喜欢阅读

1 2 3

Image source: ChatGPT

**PennState**
College of Information
Science and Technology

AI 100 Fall 2025

Vasant G Honavar

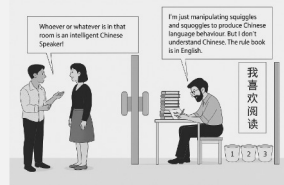
Searle's Chinese Room Argument

- Suppose then you find out that the only person inside the room was your not-so-smart brother who understands only English and has absolutely no knowledge of Chinese.
- You ask your brother how he managed to answer the questions so wisely.
- He tells you that he was given a set of books by his boss and instructed to look through the books until he found the string of symbols that look exactly like the ones written on the piece of paper.
- And when he finds that string of symbols, the book will tell him (in English) what new string of symbols he is to write in response.

If you see this shape, "什麼"	then produce this shape, "爲天"
followed by this shape, "帶來"	followed by this shape, "下式".
followed by this shape, "快樂"	

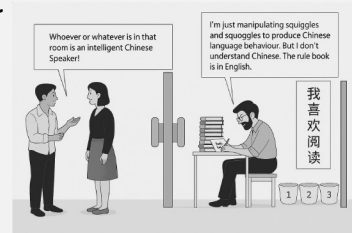
Searle's Chinese Room Argument

- Note that the only sentences on the page that your brother, as an English speaker, can understand are in English.
- This book is one of countless volumes that together tell your brother what output (in the form of Chinese symbols) should be given in response to virtually any input (of Chinese symbols) that comes through the slot into the room.
- This particular volume tells what output to give in response to virtually any input of Chinese symbols that *begins* with the first two Chinese symbols written on the piece of paper.



If you see this shape, "什麼"	then produce this shape, "爲天"
followed by this shape, "帶來"	followed by this shape, "下式".
followed by this shape, "快樂"	

Searle's Chinese Room Argument



- Your brother says he doesn't recognize any of the symbols
- You know that your brother is truthful
- The symbols are simply meaningless shapes to him
- For all he knows, they may be nothing more than patterns for making wall-paper and not a language at all
- You know that symbols do have meanings - They are Chinese symbols
- More than that, they describe questions in Chinese being asked by your Chinese friends.

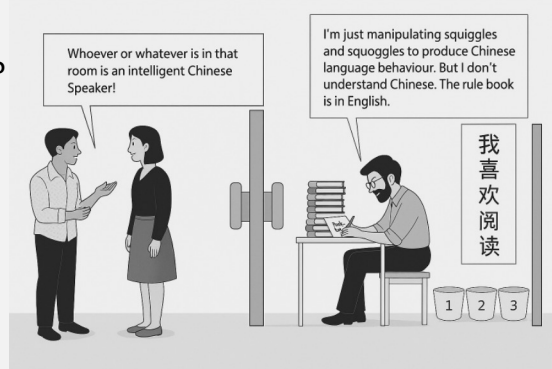
Searle's Chinese Room Argument



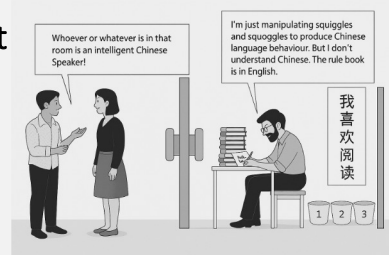
- The books function as a computer program. Each page gives instructions for how to manipulate symbols.
- The instructions at no point make any reference to the *meaning* of the symbols.
- That is, nowhere will you find a sentence that gives the English translation of any of the Chinese symbols.
- None of these books is anything like a Chinese-English dictionary.
- Instead, like a computer program itself, it instructs the reader how to manipulate the symbols based on their properties (their shape and position) not their meaning.
- If you see symbol "X" here, then write symbol "Y" there.

Searle's Chinese Room Argument

- Let's imagine that instead of your brother in the room, we have an AI program that passes the Turing Test in Chinese, just like your brother did
- Should we say that the AI program understands Chinese?
- No, says Searle.
- What do you think?

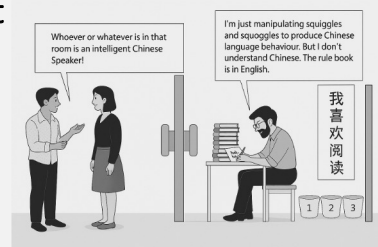


Searle's Chinese Room Argument



- Searle's claim is that he has demonstrated that no computer program that manipulates symbols based solely on their formal "syntactic" properties (e.g., their shape and their position) can ever be said to understand a language . . even if it *does* pass The Turing Test.
- Searle is *not* saying that no machine or computer can understand a language.
- Searle is willing to concede that certain operations of the human brain can properly be described as "computation."
- Searle is not insisting that humans do *not* compute and do not implement functions.

Searle's Chinese Room Argument




- Searle is insisting, however, that genuine thought and understanding require something *more* than mere computation.
- He is saying that understanding a language we do not merely manipulate symbols based on their formal properties.
- That is, **we do something** (he doesn't say what) in addition to manipulating symbols **by virtue of which**
 - we in fact understand the *meaning* of the symbols
 - which your brother or the computer did not.

Implications of Searle's Chinese Room Argument


- Is Searle right in that a process that simply syntactically manipulates its inputs to produce its outputs cannot be intelligent?
- What does this say about the working hypothesis of AI?
 - Thinking can be modeled by computation (weak version)
 - Thinking is computation (strong version)

Implications of Searle's Chinese Room Argument


- Is Searle right in that a process that simply syntactically manipulates its inputs to produce its outputs cannot be intelligent?
- What does this say about
 - The strong AI hypothesis that **Cognition is computation?**
 - The weak AI hypothesis that **Cognition can be modeled by computation?**
- Does the Chinese Room Argument suggest a more effective test of machine intelligence?

**PennState**
Institute for Computational
and Data Sciences

Center for Artificial Intelligence Foundations & Scientific Applications
Artificial Intelligence Research Laboratory


**PennState**
Clinical and Translational
Science Institute

Alternatives to (the original) Turing Test


**PennState**
College of Information
Science and Technology

AI 100 Fall 2025

Vasant G Honavar

**PennState**
Institute for Computational
and Data Sciences

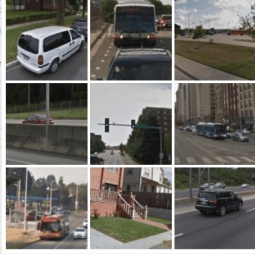
Center for Artificial Intelligence Foundations & Scientific Applications
Artificial Intelligence Research Laboratory

**PennState**
Clinical and Translational
Science Institute

Variations on the Turing Test – Reverse Turing Test


- In the reverse Turing Test (RTT), the task is for the humans to convince the computer that they are human.
- Example of a reverse Turing Test
 - CAPTCHA security measures that you've likely encountered when signing onto websites
- We can envision structuring the RTT to assess the intelligence of the machine by its success in correctly telling apart humans from machines

Select all images with a
bus
Click verify once there are none left.



⏮ ⏪ ⏩ ⏭

VERIFY

**PennState**
College of Information
Science and Technology

AI 100 Fall 2025

Vasant G Honavar

Variations on the Turing Test – Winograd Schema Test

- Winograd Schema Challenge, named for AI pioneer Terry Winograd offers a test combines natural language dialog and common sense.
- “The city councilmen refused the demonstrators a permit because they [feared/advocated] violence.”
- Which noun phrase does “they” refer to in the two sentences?
 - “The city councilmen refused the demonstrators a permit because they feared violence.”
 - “The city councilmen refused the demonstrators a permit because they advocated violence.”
- Getting the answer right seems to require commonsense understanding of how councilmen and demonstrators behave

What are valid Winograd schema?

- The two sentences are identical except for one or two words, e.g., “feared” versus “advocated” in the example above.
- The two sentences both contain two noun phrases and a pronoun, e.g., “the city council”, “the demonstrators”, and “they” in the example above.
- In reading the two sentences in isolation, humans would associate different noun phrases to the pronoun
- Both sentences must be easily understood by the human reader,
- For humans, the answer to the WSC is so obvious that the potential ambiguity goes unnoticed

What are valid Winograd schema?

- "Selection restrictions" should not suffice for correct disambiguation.
 - **Prohibited:** The women stopped taking the pills because they were [pregnant/carcinogenic]
- Pills cannot be pregnant and women cannot be carcinogenic!
 - Disambiguation only requires matching the features of the candidate referents **pills** and **women** with **pregnant** and **carcinogenic**
 - **Without the selection restrictions, the task is too easy!**

What are valid Winograd schema?

- Matching based on co-occurrence probabilities of words will not suffice for correct disambiguation.
- **Disallowed:** The race car zoomed by the golf cart because it was going too [fast/slow].
- **Why?** – Too easy!
 - “race car” and “fast” tend to co-occur in text as do “golf cart” and “slow.”
- Practical challenge:
 - Constructing good schema requires manual effort
 - The validity conditions are not always easy to verify

Winograd Schema Challenge – Quo Vadis?

- Many human-curated data sets of WSC created
- By 2021, large language models trained on large text corpora were able to match human performance
- A more extensive version of the challenge with 44,000 problems was developed - Winogrande
- By 2024 large language models matched human performance on Winogrande

What does success of machines on the Winograd Schema Challenge Tell Us about Machine Intelligence?

- Does the success of machines at beating the Winograd Schema Challenge mean that
 - Machines have become intelligent?
 - Machines have managed to acquire commonsense?
- What could explain the good performance of machines?

How was Winograd Schema Challenge defeated?

- Large text corpora contain stereotypical usage of most sentences
- Large language models (LLM) trained on such corpora become good at predicting the occurrence of words in the context of other words in the sentence
- Because of the size of the corpus, LLM can go beyond relationships between specific words, and generalize to relationship between one class of words and another – in other words, sentence schema
 - Horse raced past the barn fell
 - Boy pushed over the fence slipped
- LLM don't understand natural language any more than Searle's Chinese Room does
- LLM have no commonsense understanding of the world
- LLM display language competence without language comprehension

Visual Turing Test challenge



Challenge Questions

- Who do you see?
- What is she doing?
- What is on the desk?
- Is she indoors or outdoors?
- What color are the walls?
- How many chairs are there?

Visual Turing Test challenge



What is involved?

- Face recognition
- Object identification
- Activity recognition
- Spatial reasoning
- General knowledge
- Color perception
- Counting
- Natural Language

Discussion: What does machine success at Visual Turing Test tell us about Machine Intelligence?

- Suppose machines have passed the Visual Turing Test
- Does passing the Visual Turing Test mean that
 - Machines have become intelligent?
 - Machines have managed to acquire commonsense?
- What could explain the good performance of machines?

Visual Turing Test – Quo Vadis?

- Several versions of the Visual Turing Test available
- In recent years neural network models trained on large data sets of images and associated captions do impressively well
- Details matter
 - Data sets
 - Methods
 - Evaluation metrics

Video Turing Test – The Marcus Version



Watch an episode of Young Sheldon

- Summarize the episode
- Answer questions
 - Why did Sheldon want build rockets?
 - Why did his dad tell him he could not launch his rocket?
 - Why did his mom change her mind about allowing him to launch rockets?
 - Why did the FBI show up to talk to Sheldon?

Video Turing Test



What is involved?

- Recognize individuals
- Activity recognition
- Reasoning about emotions
- Causes and effects not shown
- Comprehending conversational context
- Recognizing and reasoning about intentions, goals, etc.
- Commonsense

Video Turing Test – Marcus Test

- The computer is asked to watch any arbitrary TV program or YouTube video and answer questions about its content
 - ‘Why did Russia invade Crimea?’
 - Why did Jordan Chiles lose her Olympic medal after it was awarded to her?

Discussion: Video Turing Test

- Not quite passed by machine
- What makes video Turing Test more challenging than the other tests discussed so far?

Video Turing Test – Quo Vadis?


- More challenging than the basic Turing Test, Winograd schema challenge, Visual Turing Test
- All of the information needed to understand the video is not in the video
- High quality annotated data hard to come by for brute force machine learning to work

Variations on the Turing Test – **Standardized testing for AI**


- Simply administer to machines the same standardized tests we use for various purposes with humans
 - SAT and ACT used for college admissions
 - GRE used for graduate school admissions
 - GMAT used for business school admissions
 - MCAT used for medical school admissions

Discussion: Standardized tests


- What would it take for machine to ace the standardized tests?
- Will passing the test imply the machine is intelligent?

**PennState**
Institute for Computational
and Data Sciences


Center for Artificial Intelligence Foundations & Scientific Applications
Artificial Intelligence Research Laboratory

**PennState**
Clinical and Translational
Science Institute

The Loebner prize for passing the Turing Test



- Held annually during 1991 through 2020
- Dr. Hugh Loebner pledged
 - \$100,000 and a gold medal to the creators of the first AI system that could pass the extended Turing Test involving textual, visual, and auditory components
 - \$25,000 and a silver medal promised for the first program that could fool judges in a text only Turing Test
 - \$2,000 and a bronze medal awarded to the chatbot judged “most human-like” in any year’s competition.
- Bronze medal was awarded a few times, but silver and gold medals were never awarded
- The Loebner prize was controversial in the AI community
- Minsky scathingly offered a prize to anyone who could stop the Loebner competition
- https://en.wikipedia.org/wiki/Loebner_Prize

**PennState**
College of Engineering
Science and Technology

AI 100 Fall 2025

Vasant G Honavar

Sample questions from the Loebner competition

- What is your most memorable childhood event and how has that impacted you today?
- Describe yourself using only colors and shapes.
- Describe why
 - time flies like an arrow but
 - fruit flies like a banana?
- How do you feel when you think about your upbringing and what makes you feel that way?
- What historical event changed you the most and where were you when it happened?
- Which of the previous questions was the most difficult to answer and why?

Lovelace Test



- Motivated by Ada Lovelace's critique of the Turing Test
- Consider a machine programmed by a human
- To pass the test, machine must display creativity in some domain under constraints given by an evaluator
 - Create an artifact in the domain
 - The artifact must meet the specified constraints
 - Human evaluator determines whether the machine's creation is acceptable
 - A referee confirms that the domain and constraints are not unrealistic for an average human
 - The creation has aesthetic appeal

Lovelace Test



Examples

- Music composition
- Fictional story generation
- Poetry production
- Paintings
- Sculpture
- Novel math proofs
-

Lovelace Test



Story generation requires

- Familiarity with existing literature
- Familiarity with culture, social norms
- Commonsense knowledge
- Discourse planning
- Natural language competence
- Reason about people, their feelings, emotions, goals, mental states, intentions

Lovelace test – Story Generation – Quo Vadis?

The Boy and His Dog

Source: ChatGPT

Ethan was a quiet boy who often felt invisible at school. He preferred books to basketball and daydreams to noisy classrooms. But when he came home each afternoon, he was greeted by Scout, his shaggy brown dog with eyes that always seemed to understand.

Scout wasn't like other dogs. He didn't bark much or chase balls for long. Instead, he listened. When Ethan spoke about the loneliness he carried, Scout would rest his head on Ethan's knee as if to say, *I hear you*.

One evening, Ethan sat on the porch steps, holding a crumpled test paper. His chest ached with the weight of disappointment. Scout nudged his arm, then trotted toward the woods behind the house. Curious, Ethan followed.

Under the canopy of trees, Scout bounded ahead, leading Ethan to a small clearing where fireflies shimmered like stars. Scout barked once—soft, urging Ethan to look up. Ethan realized the world was wider than a single grade, brighter than a single bad day. He laughed, the sound carrying into the night. Scout wagged his tail, circling him as if he, too, was laughing.

From that day forward, whenever Ethan felt the world pressing down, he remembered the clearing and the fireflies. And always, by his side, was Scout—his companion, his comfort, his friend.

Lovelace test – Poetry – Quo Vadis?

An Ode to Our Politicians

Source: ChatGPT

Politicians line the stage,
Each one reads a different page.
They vow to clean the nation's mess,
But can't clean up their own desk.

They bicker, boast, and wag their tongues,
Like toddlers fighting over crumbs.
They spend all day just passing blame—
It's almost like a party game!

And when the crowd begins to scoff,
They vote themselves a raise—and laugh it off.

Lovelace test – Art – Quo Vadis?

An impressionist painting of a family strolling through a flower garden



Source: ChatGPT

Discussion: Lovelace Test

- What would it take for machine to pass the Lovelace test?
- Will passing the test imply the machine is intelligent?

Lovelace Test

- While machines may not quite pass the Lovelace test with flying colors, they are beginning to look like serious contenders
- Success, as in the case of Winograd Schema Challenge, Visual Turing Test, etc., has come from training generative models on large corpora

Social-Emotional Turing Challenge

- A test of what Sternberg called street smart and what Gardner called interpersonal intelligence
- **Scenario:** Tracy asks for a banana; Mom gives Tracy an apple
- **Question:** How will Tracy feel?
- Potential responses:
 - Tracy feels sad because she did not get what she wanted
 - Tracy feels happy because she asked for a banana which is what her mom usually gives her, but she was bored of eating banana pretty much everyday so was excited to try apple for a change
- **Evaluation:** Rating of machine generated and human generated responses by neutral evaluators

Social-Emotional Turing Challenge

- **Scenario:** Tracy asks for a banana; Mom gives Tracy an apple
- **Question:** How will Tracy feel?
- Probably disappointed or annoyed. Tracy asked for a banana (clear expectation). Getting an apple violates that expectation, so the likely emotions are disappointment, irritation, or confusion (“Why didn’t I get what I asked for?”).
- Nuance:
 - If Tracy likes apples or was just hungry: mild disappointment to acceptance.
 - If Tracy is young or really fixated: frustration/tantrum is possible.
 - If Mom explains (“we’re out of bananas”) and offers apple instead: feelings may shift to understanding.

Source: ChatGPT

Scientific AI Challenge

Scientific discovery offers a grand challenge for AI

- AI Assistants for Scientists
- Autonomous Robotic Scientists

Representative AI associates for scientists

A scientist's associate that

- Learns what you and others in your field and related fields are working on
- Finds and reads relevant literature
- Locates and ingests available knowledge and data
- Offers assistance
 - Here are some data that contradict your hypothesis
 - Here are arguments for and against your hypothesis
 - Here is some data from lab X that explains your finding
 - Here is why you should prefer model A to model B



Examples of AI associates for scientists

A scientist's associate that:

- Given background knowledge and access to literature in two or more domains:
 - Use literature in one domain to generate hypotheses or explanations in another domain
- Given a scientific question, and a social network of researchers and their scientific output e.g., publications,
 - Identify a collection of researchers that are best equipped to address the question



Success criteria

- For an associate
 - Ratings by human scientists
 - Productivity of human scientist with and without AI associate
- For an AI scientist
 - Scientific breakthroughs or findings that are deemed worthy of publication in reputed scientific venues
 - Scientific impact

Variations on the Turing Test – **Physically Embodied Turing Test**

- A robot has to physically manipulate real-world objects in meaningful ways
 - Build a structure from a pile of parts using verbal, written instructions and pictures (imagine assembling IKEA furniture)
- A robot has to devise solutions to a set of open-ended but increasingly creative challenges using toy blocks
 - Build a wall
 - Build a house
 - Attach a garage to the house
- The robot is required to “explain” its efforts
 - “The door needs to be installed before the wall around it; otherwise it wouldn’t fit

Harnad Test or Total Turing Test

- To have any hope of passing the Total Turing Test,
 - The machine must be more than a computer
 - It must be robot with the sensors, effectors, and
 - the physical wherewithal to interact with the physical world with real objects, real people, etc.
 - If the robot can do everything a body with a mind can do such that we can't tell them apart, we have no basis for doubting it has a mind.
- Challenges:
 - What does “everything a body with a mind can do”?
 - Eating? Sleeping? Playing basketball?
 - How to operationalize the test?

Discussion: Total Turing Test

- How can we operationalize Total Turing Test?
- If a robot passes the the Total Turing Test, does it mean that the robot is intelligent?
 - Why or why not?

Limitations shared by all forms of Turing Test

- Any attempt to assess machine intelligence must confront the other-minds problem
 - How can you tell whether anyone other than yourself has a mind when the only way to know is by being the other person?
 - We can't.
 - We are left with some form of Turing Test:
 - If it can do everything a body with a mind can do such that we can't tell them apart, we have no basis for doubting it has a mind.
 - "If it looks like a duck, it quacks like a duck, it must be a duck"
 - Strong AI hypothesis is untestable
 - Weak AI hypothesis is testable