

Artificial Intelligence: Philosophy, Emotion, Modern AI

CS151
David Kauchak
Fall 2010

Adapted from notes from:
Sara Owsley Sood

Admin

- Exams back Wed.
- Drafts due tomorrow, 5pm in dropbox (pdfs)

Who are you and why are you here?

- Name/nickname
- Dept., college and year
- What is AI? (or what do you think AI is? 😊)
- Why are you taking this course?
 - What topics would you like to see covered?

AI is a huge field

- So, what is AI...
- One definition:

“Building programs that enable computers to do what humans can do.”
- for example: read, walk around, drive, play games, solve problems, learn, have conversations...

Some of the big questions of AI philosophy...

- What is AI, really?
 - What does an intelligent system look like?
 - Does an AI need—and can it have—emotions, consciousness, empathy, love?
- Can we ever achieve AI, even in principle?
- How will we know if we've done it?
- If we can do it, should we?

Weak & Strong AI

- Weak AI hypothesis:
 - Machines can be created that will act as if they are intelligent (simulate thinking)
 - What AI can do is to develop useful, powerful applications
- Strong AI hypothesis:
 - Machines can be created that will actually think (rather than simulate thinking)

Weak AI: Turing Test

- Turing predicted: by 2000, a computer could be programmed to pass the Turing Test (for a 5 minute test) 30% of the time

Turing Test - Eliza

- An early success at fooling some people: Eliza, simulating a Rogerian therapist
- ...Eliza; the importance of picking the correct application area
- Sample interaction:
 - Patient: Men are all alike.
 - ELIZA: In what way?
 - Patient: They are always bugging about something or the other
 - ELIZA: Can you think of a specific example?
 - Patient: Well, my boyfriend made me come here.
 - ELIZA: Your boyfriend made you come here?
 - Patient: He says I'm depressed much of the time.
 - ELIZA: I am sorry you are depressed.

- Human: Tell me about yourself alice
- ALICE: I am an artificial linguistic entity. I was activated at Bethlehem, Pennsylvania, on November 23, 1995. My instructor was Dr. Richard S. Wallace. He taught me to sing a song. Would you like me to sing it for you?.
- Human: Yes please
- ALICE: Your polite style is very nice. It's called "Daisy." Daisy, Daisy. Give me your answer do. I am half crazy. All for the love of you. It won't be a stylish marriage. I can't afford a carriage. But you'll look sweet. Upon the seat. Of a bicycle built for two.
- Human: That reminds me of HAL in "2001 a space odyssey"
- ALICE:" Your polite style is very nice. It's called "Daisy." Daisy, Daisy. Give me your answer do. I am half crazy. All for the love of you. It won't be a stylish marriage. I can't afford a carriage. But you'll look sweet. Upon the seat. Of a bicycle built for two." reminds you of HAL in 2001 a space odyssey ? Does that make sense?
- Human: Can you tell me something about Stanley Kubrick?
- ALICE: Ambiguous: Ask who, what or where is Stanley Kubrick ?



Jabberwacky

Uses real user conversation to learn responses.

<http://www.jabberwacky.com/>



CyberLover

<http://www.guardian.co.uk/technology/2007/dec/13/internet.crime>



Can we ever achieve AI?

Can we ever achieve AI?

- Argument of disability: “hey, there are lots of things that a computer can’t do!”
 - “Be kind, resourceful, beautiful, friendly, have initiative have a sense of humor, tell right from wrong, make mistakes, fall in love, enjoy strawberries and cream, make someone fall in love with it, learn from experience, use words properly, be the subject of its own thought, have as much diversity of behavior as man, do something really new.”

Responses?

Some successes

- What are some human-oriented tasks that computers can do better than people?
 - Play chess, checkers and other game
 - Inspect parts on assembly lines
 - Check the spelling of text
 - Steer cars and helicopters
 - Diagnose diseases
 - Do hundreds of other tasks as well as or better than humans
 - Computers have made small but significant discoveries in astronomy, math, chemistry, mineralogy, biology, computer science, and other fields

or...

- Argument of informality:
 - “what people do is too complex to capture”
 - Because computers can do no more than follow a set of rules, they cannot generate behavior as intelligent as that of humans

Responses?

or...

- Argument of informality
- More or a problem with “classic” AI
 - reasoning
 - knowledge representation
- Some progress in incorporating background knowledge
- Learning algorithms: increasing ability to operate autonomously (unsupervised learning), learn new features, prune feature spaces

Strong AI

- Can machines really think?

- What does it mean to think?
 - consciousness: a machine has to be aware of its own mental states and actions – self aware

“brain in a vat” experiment

- Is physicality crucial for intelligence?
- *Matrix* scenario: a brain is supported, bodiless, in a vat, and signals simulating a virtual world are fed in/out of the brain
- Is being hungry the same as
DyingFor (Me, Pizza)



“brain prosthesis” experiment

- Moravec (robotics researcher/functionalist) is convinced that his consciousness would remain unaffected
- Searle (philosopher and biological naturalist) is equally convinced his consciousness would vanish

The Chinese Room

- John Searle, 1980
- Human who knows only English; stacks of paper with Chinese symbols; rule book in English, stating which bit of paper to give in response to a given (Chinese) input
- Human who knows only Chinese on outside of room; passes in Chinese query, receives Chinese response



Creative

- Having the ability or power to create: Human beings are creative animals.
- Productive; creating.
- Characterized by originality and expressiveness; imaginative: creative writing.

How do people write stories?

Can Computers Be Creative?

- Two paintings produced by Harold Cohen's Aaron software:



<http://www.kurzweilcyberart.com/>

Say Anything

Corpus based story telling

<http://sayanything.ict.usc.edu/SayAnything/>