



Philosophy

CS51A
David Kauchak
Spring 2022

Adapted from notes from:
Sara Owsley Sood

1




Admin

No normal mentor hours this week

Sunday, 12-2pm mentor session for last-minute final questions

No office hours tomorrow for Dr. Dave

2



Some of the big questions of AI philosophy...

What is AI, really?

- What does an intelligent system look like?
- Does an AI entity 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?

3



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)

4

Weak AI: Turing Test

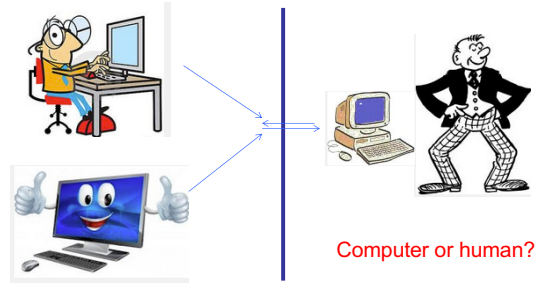
Alan Turing, 1950:

(<http://www.loebner.net/Prizef/TuringArticle.html>)

"I propose to consider the question, 'Can machines think?'"

5

Weak AI: Turing Test



6

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

7

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.

8



CyberLover

That our computers have seduced us has long been a truism. Now, thanks to the ever-inventive internet mafia, it is becoming a literal truth. Russian cyber-crooks have reportedly unleashed a software robot, or bot, that poses as a would-be paramour in sex chatrooms. It entices randy gentlemen to reveal personal information, such as their address or birthday, or even to submit photographs of themselves. The information can then be used to break into bank accounts or carry out other forms of fraud.

It was probably inevitable. As one of Tony Soprano's sidekicks observed in a classic episode of the TV series, the two most resilient sectors of the economy are organised crime and "certain aspects of showbusiness". The aspects, that is, known as the world's oldest profession - now mixing it with the world's newest technologies.

CyberLover, as the dirty-mouthed bot is called, is quite a sophisticated piece of software. It can take on a number of different guises depending on the proclivities of its target, according to security experts at the software company PC Tools. It can play the role of a romantic lover, for instance, or masquerade as a sexual predator.
<http://www.guardian.co.uk/technology/2007/dec/13/internet.crime>

9



Cleverbot

<https://www.cleverbot.com/>

<https://www.youtube.com/watch?v=WnzIbYtZsQY>

10



Can we ever achieve AI?

11



Some successes

What are some human-oriented tasks that computers can do better than people?

- Play chess, checkers and other games
- Inspect parts on assembly lines
- Check the spelling of text
- Steer cars and other vehicles (helicopters, planes, etc.)
- 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

12



13

Strong AI

Can machines really think?

What does it mean to think?

Do we have to have a brain to have a mind? to think?

14

“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 some rule:
HungryFor (Me, Pizza)

Could you tell the difference?

15

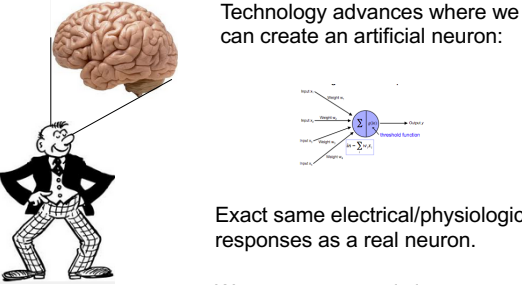
“brain in a vat” experiment

Moravec (robotics researcher/functionalist) is convinced that his consciousness would remain unaffected

Searle (philosopher and biological naturalist) is equally convinced his consciousness would vanish

16

“brain prosthesis” experiment



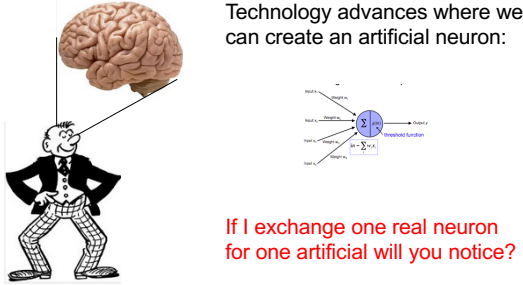
Technology advances where we can create an artificial neuron:

Exact same electrical/physiological responses as a real neuron.

We can copy an existing neuron.

17

“brain prosthesis” experiment

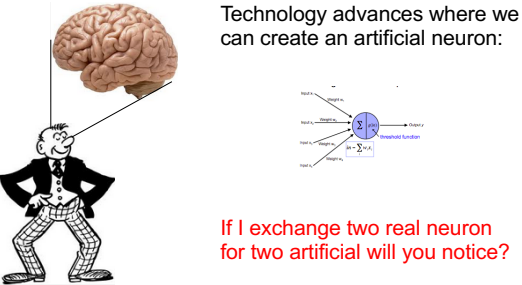


Technology advances where we can create an artificial neuron:

If I exchange one real neuron for one artificial will you notice?

18

“brain prosthesis” experiment

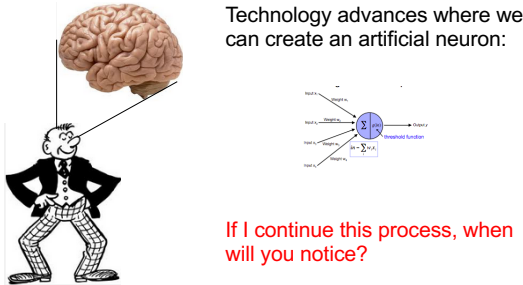


Technology advances where we can create an artificial neuron:

If I exchange two real neuron for two artificial will you notice?

19

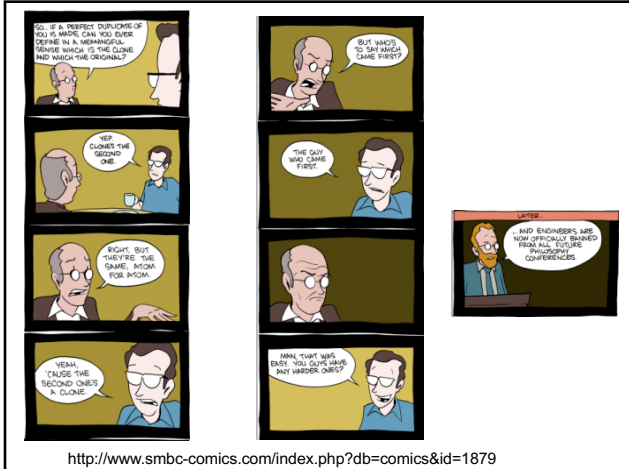
“brain prosthesis” experiment



Technology advances where we can create an artificial neuron:

If I continue this process, when will you notice?

20



21

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?

24

Can Computers Be Creative?

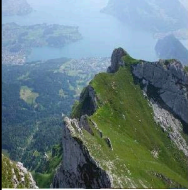
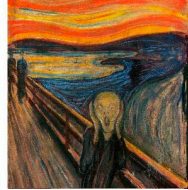
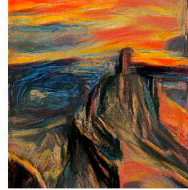
Two paintings produced by Harold Cohen's Aaron software:




<http://www.kurzweilcyberart.com/>
http://www.kurzweilcyberart.com/aaron/aim_clip_cohen.html

25

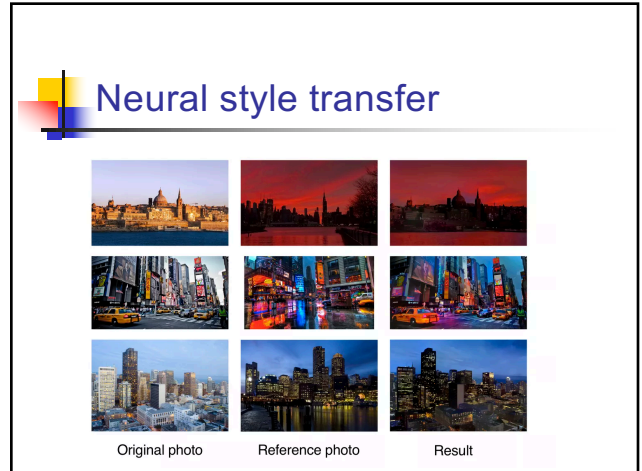
Neural style transfer


+

=


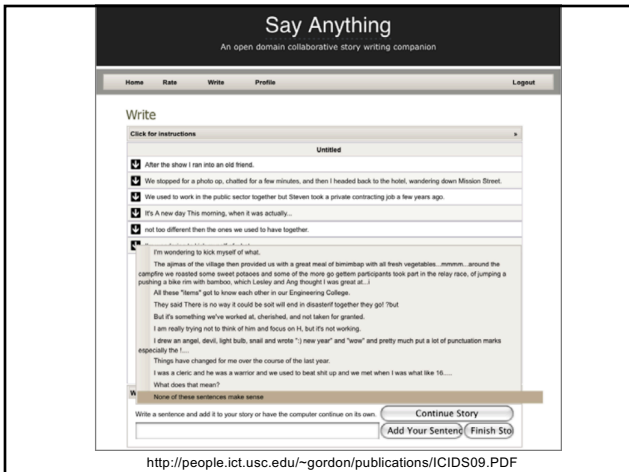
26



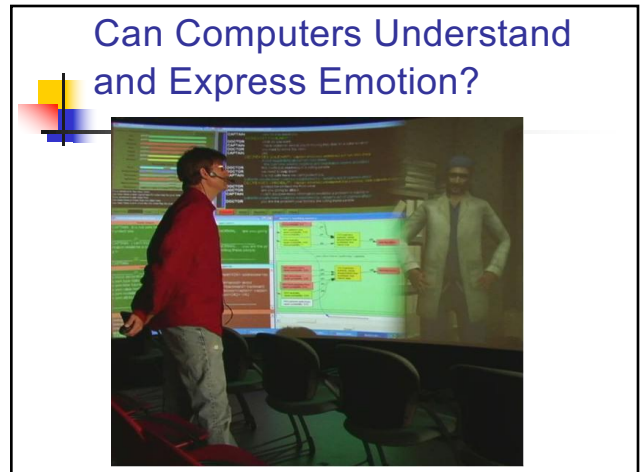
27



28



29



30

Emotion

Can we build systems to

- Detect it?
- Express it?

32

Detection

Gestures

Facial Expressions

Speech/Text

Physiological Cues

- Blood volume pressure
- Skin Conductivity
 - Glove (other WEARABLE DEVICES!)

33

Detecting emotion via wearable devices

2001 - 81% accuracy in (forced decision)

- detection of 8 emotions:
 - Neutral, anger, hate, grief, platonic love, romantic love, joy, reverence
- Person dependent - trained for at least 4 weeks
- GROUND BREAKING!

Since then, lots more work has been done

- <http://www.youtube.com/watch?v=ceP-vcbFhx0>

Applications?

34

What we covered

Python!

- variables
- functions
- loops
- conditionals
- recursion
- higher order functions
- classes
- file I/O
- many other, sub-topics
 - lists, tuples, dictionaries, ...
 - exceptions
 - turtle graphics

35

What we covered

Machine learning
Naïve Bayes model

Neural Networks

Search

- algorithms
- problem solving
- adversarial search and game playing

Web basics

Artificial Intelligence

36

Where we started

```
# This program figures out the number of hot dogs
# needed for a BBQ
teran = 1
jasmin = 2
chris = 2 * jasmin
brenda = chris - 1
grace = (brenda+1)//2 + 1 # add 1 to brenda using truncated division to round up

total_hotdogs = teran + jasmin + chris + brenda + grace
print(total_hotdogs)
```

37

Where we ended

```
class NimGame:
    """ Class to keep track of a nim game. """
    def __init__(self, starting_piles):
        """ Construct a new game with the list starting_piles
        as the piles """
        self.piles = starting_piles
    def get_piles(self):
        """ Returns a copy of the current piles """
        # return a copy to avoid anybody messing with the internal
        # state of the game
        return self.piles[:]
    def make_move(self, pile_number, num_to_remove):
        """ Move num_to_remove from pile_number.
        Return True if the move was valid, False otherwise.
        if pile_number < 0 or pile_number == len(self.piles):
            return False
        elif num_to_remove < 0 or num_to_remove > self.piles[pile_number]:
            return False
        else:
            self.piles[pile_number] -= num_to_remove
            return True
    def is_over(self):
        """ Is the game over? """
        return sum(self.piles) == 0
    def __str__(self):
        return str(self.piles)
```

38

Where we ended

```
def play_nim(player1, player2, game_state):
    # player1 will start
    player1_turn = True
    while not game_state.is_over():
        if player1_turn:
            print("Player 1's turn")
            print(game_state)
            (pile_number, num_to_remove) = player1(game_state.get_piles())
            print("Player 1:",end='')
        else:
            print("Player 2's turn")
            print(game_state)
            (pile_number, num_to_remove) = player2(game_state.get_piles())
            print("Player 2:",end='')
        print(str(num_to_remove) + " from pile " + str(pile_number))
        print()
        if not game_state.make_move(pile_number, num_to_remove):
            print("ILLEGAL MOVE!")
        player1_turn = not player1_turn
    # game's over, see who won
    if player1_turn:
        print("Player 2 won!")
    else:
        print("Player 1 won!")
```

39

Ethics



<http://www.smbc-comics.com/index.php?db=comics&id=2956#comic>

40



41