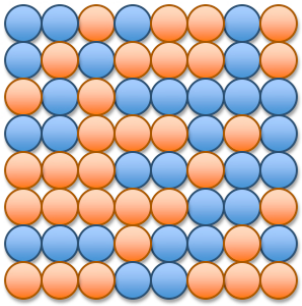
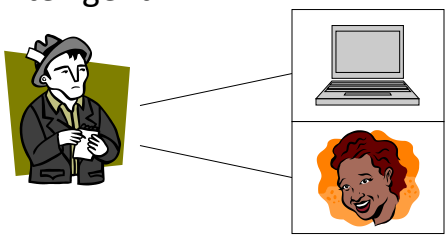


UNIT 13B

AI: Natural Language Processing

The Turing Test

- Turing publishes Computing Machinery and Intelligence in 1950.
- Describes a test (now called the Turing Test) to determine whether a computer can be called intelligent.



A machine is considered intelligent if an interrogator cannot tell if a human or a computer is answering a set of questions using typewritten responses.

Natural Language Processing

- Natural language processing involves the interactions between a computer and a human via spoken (or written) language.
- ELIZA
  - One of the first programs to build a conversation between a human and a computer, created by Joseph Weizenbaum in 1966.
  - Computer responses were created based on *scripts*.
  - Most famous example was DOCTOR which simulated a Rogerian psychotherapist.

ELIZA

H: My father wouldn't buy me a puppy.

C: Tell me more about your family.

H: I am sad I never got a puppy.

C: Can you explain what made you sad you never got a puppy?

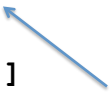
H: I am sad we never had a chance to play fetch.

C: I am sorry to hear you are sad we never had a chance to play fetch.

Regular Expressions

- A regular expression is a rule that describes the format required for a specific string.

```
>>> import re
>>> s = "how now brown cow"
>>> match = re.findall('\wow', s)
>>> match
['how', 'now', 'row', 'cow']
>>>
```



regular expression that says match any string in s consisting of 3 characters, where the first character is anything (\w) and the next 2 characters are 'o' and 'w' exactly

Wildcards

- We can use a wildcard symbol (.*) to match any number of characters.

Let `p = Pattern("I am afraid of (.*)")`

```
>>> p.add_response("Why are you afraid of $1?")
>>> p.apply("I am afraid of ghosts")
'Why are you afraid of ghosts?'
>>> p.apply("I am afraid of Tom")
'Why are you afraid of Tom?'
```

Postprocessing

- To make things more realistic, we can replace personal pronouns with their opposites.

```
"I am near my car"
=> 'Are you really near my car?'
"I am annoyed at you"
=> 'Are you really annoyed at you?'
```


Postprocessing (cont'd)

```
>>> Eliza.post["my"] = "your"
>>> Eliza.post["you"] = "me"
>>> Eliza.post
{'my' : 'your', 'you' : 'me' }
>>> p.apply("I am near my car.")
'Are you really near your car?'
>>> p.apply("I am annoyed at you.")
'Are you really annoyed at me?'
>>> p.apply("I am sad, my oh my.")
'Are you really sad, your oh your?'
```

An associative array maps one element to another.

Preprocessing

- Preprocessing is used to transform part of a sentence before pattern matching is performed.

```
>>> p = Pattern("I am afraid of (.*)")
>>> p.apply("I'm afraid of ghosts")
>>> Eliza.pre["I'm"] = "I am"
>>> p.apply("I'm afraid of ghosts")
'Why are you afraid of ghosts?'
```


Algorithm for a Conversation

1. Perform any preprocessing on the sentence.
2. Break the sentence into words.
3. If there is a rule for a word, add the rule to a priority queue.
4. Try the rules in order of decreasing priority.
5. If a rule applies to the input sentence, apply postprocessing rules to placeholder variables and return the response.

Eliza and the Turing Test

- Human (natural) language is very complex.

  "You said I wouldn't be afraid of horses if I visited a state fair, but when I was there I was chased by a pig, and now I'm having nightmares about bacon."

  Natural language processors (NLPs) must use **parsers** to break a sentence down into its constituent parts (noun, verb, etc.) to determine an appropriate response.

Eliza and the Turing Test

- Human (natural) language is very complex.

  "The lions were lying down in the grass."
  A: Go on.

  "They were quiet"
  A: Why were the lions quiet?

  Natural language processors (NLPs) must have **memory** to associate words (e.g. "they") to words from previous sentences.

Eliza and the Turing Test

- Human (natural) language is very complex.

  "The elephant was huge."

  A: How big was it?

  "The cat fell into the well."

  A: Was it able to get out?

  NLPs must use **semantics** to assign meaning to certain phrases in order to generate appropriate responses.

Eliza and the Turing Test

- Human (natural) language is very complex.

  "The dog ran across the backyard while I waited on my patio."

  A: What happened when the dog reached the patio?

  NLPs must establish **inferences** to add information to sentences that is not given but is assumed.