Goals of AI

COMPSCI 570

Ron Parr

What is a Computation?

- Alonso Church: Anything that can be expressed in a formal language (not entirely unlike today's programming languages)
- Alan Turing: Anything that can be computed by a specific type of machine (more on this later)
- Stephen Kleene: Anything that can be computed by a recursive function

- It turns out that all of these are equivalent

Church-Turing Thesis

- C-T thesis is a hypothesis, not a provable statement

- Paraphrased: "Every effectively calculable function is computable by a machine"

- Effectively calculable = has some reasonable set of rules describing the computation

Tangent: Not everything is computable

- Church-Turing thesis is not vacuous
- It turns out that some things you might want to do are not computable in any finite amount of time

- Halting problem: Is there a general procedure for determining if another program will terminate for a particular input?
- Proofs in first order logic: Is there a general mechanism that will prove any logical statement (with quantification) true or false?

- The above problems are "semi-decidable"
- We can write a program that will search for answers, but if the answer doesn't exist, the program may run forever.

What is a Turing Machine?

- Ingredients of a Turing Machine
  - An infinite tape
  - A "head" that reads symbols on the tape
  - A finite program:
    - Program has a finite number of states
    - Each state can inspect symbols on the tape
    - (Conditionally) rewrite symbols on the tape
    - Move the head

- Turing machines are abstractions of every modern computer, except that actual computers have finite tapes (finite memory)

Who Was Alan Turing?

- British mathematician
- Code breaker who developed machines for breaking the famous German "enigma" code in WWII, providing allies with key strategic advantage, e.g., locating u-boats
- Considered a father of the modern notion of computation, **and** of AI

- Committed suicide (at 41) after being prosecuted for being homosexual and forced to endure hormone treatment (or face imprisonment)

Turing's Straw Men(?)

- Theological Argument – untestable
- Head in the Sand – wishes != reality
- Mathematical (completeness) – not relevant
- Consciousness – untestable
- Arguments from disabilities – disguised form of earlier arguments, or just begs the question
- Originality/creativity (Lovelace's argument) – Applies to people too?
- Analog/digital distinction – not relevant
- Informality of behavior – begs the question
- ESP – given more credibility than it deserved

Tangent: What are Reason and Science?

- A key element of scientific/rational behavior is that it make testable claims about the world
- Subjective claims may still have some value:
  - "Roses are beautiful."
  - "Chicken is delicious."
- Untestable claims may have personal value to the speaker, e.g., "My deity forbids the consumption of bananas."

- Untestable claims cannot be a rational basis for resolving disputes, e.g., "This box is not intelligent because XXX," where XXX is untestable is not a rational way to convince somebody of your position.
- Why? If both parties offer untestable claims to support their position, there is no basis for choosing one over the other!

Behavior and Testability

- What are some other objective criteria that might be relevant for assessing intelligence aside from behavior? (For example, would it matter if a machine that randomly combined words coincidentally passed the test?)

- Do the inner workings of the machine matter?

Watson

- Arose as a "challenge problem"
- Jeopardy viewed as requiring intelligence

- Question answering:
  - Viewed as an independently useful task
  - Previous work existed but was weak
  - Partly embodied in google (now), Siri today
  - See Wolfram alpha
  - Watson is now used to help doctors

- Previous work was promising but weak

Why Is NLP Hard?

- Inherent ambiguity of language
- Headline: "Police help dog bite victim"
- "One morning I shot an elephant in my pajamas. How he got in my pajamas, I don't know." (Fodor)
- "Time Flies Like and arrow; Fruit flies like a banana." (Groucho Marx)

Watson Hardware

- Massively parallel cluster of machines

- Not quite a top 500 supercomputer

- Arguably, less computational ability than a brain

Why Didn't This Happen Sooner?

- Availability of the internet
- Advances in statistical parsing
- Move away from monolithic knowledge bases
- Advances in machine learning

- Speed, memory

Lessons Learned From Watson

- Know what you don't know!

- Value of ensembles

- Value of ontologies

- Deep and shallow knowledge both have roles