Schedule for Week 2 of SI 701: Information Theory and Artificial Intelligence

Meghana & Linfeng

September 14, 2015

1 Part 1: while Paul is in the room

For the first 80 minutes, Paul will be around.

1.1 Opening: summary and stories (30 min)

1.1.1 What was the most interesting? (10-15 min)

- Inspire, and have people to talk about it. (Rephrase the question)

- Top two are needed.

(The goal is to get people as tomorrow's "PhD in Information" know about something to gossip while talking about how information theory came about.)

1.1.2 What's new: (check list)

- Entropy: measure for information.

  - A measure of information with units as 'bits'.
  - More technical details comes after the break.

- Diagram of Information Source, Transmitter, Noise Source, Receiver, Destination;

  - Compression Lossy vs Lossless
    1. Summary vs Excerpt
    2. Diff (for storage and version control)

- Further discussion on applications/understandings of Entropy (Need to post these questions and lead the discussion)

- When a message is not yet random, it contains redundancy.

  * Compression reaches zero redundancy when the message is truly random;
  * Encryption achieves the stage of "undecipherable" if the message is truly random.
  * Then, does lossless compression reach a fully-lossy consequence when pushing it to the limit?

- Underdog wins

- Redundancy does not increase or decrease entropy

- Turing test as an (old) standard for Artificial Intelligence: **Can machines think? 1950**

  1. Building on the work regarding the universal machine.
  2. Turing continued to ask What can a machine not do?
  3. Turing Test.

- Questions: Whats the turing test? How are the Turing test and Entropy related?

1.2 Paul's lecturing (15 minutes)

Paul agreed to cover:

- compression

- Comparison among: completing texts (texting on iPhone) and generating text (Turing test)

1.3 Nicole's lecturing + connecting with future materials (15 minutes)

- Modern communication theory's critique of Shannon's information theory.

2 Part 2: second 1.5 hours

Well, given a minimum of 10 minutes break in between, 80 minutes is the target here.

2.1 On definition of Entropy (20min)

- Foundation: discrete probability distribution and probability density function;

Definition 2.1. For discrete distribution,

$$H(\tilde{x}) = -\sum p_i \log_2 p_i$$

2

2.1.1 About the probability distribution:

When calculating the entropy of a random variable, we only use "true probability".

2.2 Game on word perdiction (20 min)

Would a human being be better at next word prediction? We could test this in class half the class could use phones to do the word prediction on their phones, half the class could compete what they think the next word is. Just a version of the Shannon game. (Common starting point)

Stage 1 (**Administrative**) Pair up into groups of 2;

Stage 2 (**Seeding**) One with a Smartphone comes up with a sentence, and then type the first word;

Stage 3 (**Predictin**) Both member of the group predicts. One uses the Smartphone's prediction algorithm, while the other compete through "guessing".

2.3 Discussion

- Effects of text prediction on behaviour/communication patterns? Also, ethics of training data used, biases introduced by the data perpetuating existing biases in society.

- What makes computers a possible channel of artificial intelligence? [1]

- Shall a system/algorithm/mechanism that follows **(fixed) rules** be counted as intelligence?

- What would be the problem if computers turned out to be intelligent?

- Summarize the Ashley Madison and ask people's opinions.

2.4 Closing: any changes in what was most interesting?

---

[1] Due to the fact that computers as binary calculators can practice arbitrary algorithm, as algorithms could all be decomposed into binary relations governed by conditions and coconditions ("if" statement"), as well as logic connectives ("and, or, not").