ELL788/HSL622: Major Test

Sumeet Agarwal

May 20, 2021

Instructions: You should primarily refer to the assigned readings and lecture materials, but may also consult any other resources you find useful. However, you should attempt the questions on your own and your answers should be written entirely in your own words. If you refer to sources outside of the assigned readings, please mention them in a list of references at the end of your answer script.

Please try to answer to the point, without getting into general discussion that is not directly responding to what has been asked. The answer scripts will need to be uploaded on Moodle, and you should do so by the end of Thursday 10th June. You can prepare your responses using any kind of software you wish, including scans/photos of handwritten text or diagrams; it would be preferable to finally combine everything into a single PDF file for submission. Please make sure that all parts of your submission are clearly legible when rendered on a typical laptop screen.

1. For this question you need to design an artificial neural network for a specific Boolean function assigned to you. Please look up your individual Boolean function at `https://docs.google.com/spreadsheets/d/1eHrj-GYmieA5-kRBkyGXlRNFW2ON7_rt7XXtAQSiwI4/edit?usp=sharing`. Recall that a Boolean function has both inputs and outputs which are 0/1.

   (a) Draw a full neural network which implements your assigned function. The inputs to the network should be x_1, x_2, and x_3, and the output should be the value of the function for the given inputs. Please clearly show all edge weights and activation thresholds as part of your diagram. **[6]**

   (b) Is your function a linear threshold function, or not? Explain why in either case. **[3]**

2. Consider a convolutional neural network (CNN) which has been trained to be able to detect whether a given image contains a human playing a piano. Explain briefly what kinds of *representations* of the image the layers of such a CNN would be constructing. What sort of information about the image would be represented in each layer, and how would this change as you went up the layers? **[2]**

3. (a) In what way did Alan Turing, in his 1950 paper, anticipate the Chinese room kind of critique of the Turing test that John Searle would make 30 years later? What was Turing's response to this line of criticism? **[2.5]**

   (b) Searle in his paper talks about Roger Schank's AI program, and how it is able to answer questions based on logical inference from natural language stories, even when the answer is not directly provided in the story. For example:

   A man went into a restaurant and ordered a hamburger. When the hamburger arrived it was burned to a crisp, and the man stormed out of the restaurant angrily, without paying for the hamburger or leaving a tip." Now, if you are asked -Did the man eat the hamburger?" you will presumably answer, ' No, he did not.'

   Why does Searle mention this specific example of AI, and in what way does he seek to use it to make his argument against the case for strong AI? **[4]**

   (c) Steven Pinker, in the assigned reading, also responds to Searle's Chinese room argument. Do you think Turing and Pinker would broadly agree on the problems with Searle's argument? Explain why or why not, based on the respective readings. **[2]**

4. Suppose a 'scoffer' in the First AI Debate comes and shows you some examples of how badly Google Translate performs on metaphors, and says, "look, this is exactly why machines can never have *real* understanding! How can you ever make sense of things like metaphors via mere symbol manipulation? It needs *intentionality*, it needs the symbols to *stand* for something in the real, physical world!"

   (a) This argument appears similar to Searle's Chinese Room argument. But there is a clear logical flaw in equating these two arguments. What is it? **[1.5]**

   (b) How might someone like Turing use the above argument (and the mentioned logical flaw) to turn the tables, and show that this way of thinking actually supports the validity of the Turing test as a means of assessing machine intelligence? **[1.5]**

5. (a) Consider the below stimuli:

A vheclie epxledod at a plocie cehckipont near the UN haduqertares in Bagahdd on Mnoday kilinlg the bmober and an Irqai polcie offceir

How do such instances relate to the Second AI Debate? Which side of the debate have they been used to argue for, and why? **[2.5]**

(b) In what way might thinking about the role of *representation* in cognition/intelligence serve as a way of bridging or unifying the two sides of the Second AI Debate? **[3]**