HS8.102-Phil.IHS/Assignment-1/2020114001/CLD

Part 1

We wish to analyse whether a computer which can *distinguish* a human and another computer attempting to be human should be considered intelligent. I feel that the answer to this question is a yes – only being with a human or near-human level of intelligence can effectively judge whether a certain behaviour is "human" or not. At the very least, I believe that a human can judge the distinction more effectively than a computer.

Quoting Lake *et al.*, 2017[1],

> Despite recent computational achievements, people are better than machines at solving a range of difficult computational problems, including concept learning, scene understanding, language acquisition, language understanding, speech recognition, etc. Other human cognitive abilities remain difficult to understand computationally, including creativity, common sense, and general purpose reasoning.

The above listed cognitive abilities, naturally, include those necessary for a computer to arbitrate a Turing test successfully. The arbiter of a Turing test, would need to identify whether the response made by the subject is relevant and common-sensical, contains the required amount of information, and presents the suitable emotion. Such questions are hard to define rigorously – they are (I believe) of the "I know it when I see it" category. Being able to answer such questions with some semblance of consistency is, in my opinion, necessary to be considered intelligent.

The rebuttal might be raised, however, that a computer that arbitrates a Turing test successfully knows only how to *identify* intelligent behaviour, which does not mean it can replicate it completely (possibly analogous to the way a person can identify a language without necessarily knowing to speak it).
This is certainly a possibility. However, carrying the analogy further, proficiency in a language is not a binary question – there are continuously varying measures of fluency a person can have in a language. A person who can identify a language almost certainly has *some* proficiency in it, even if it is minimal. In the same way, considering intelligence to be a spectrum and not a discrete binary characteristic, a computer that can successfully carry out the arbitration can be considered to be intelligent to some degree, without being *completely* intelligent. Replicating the intelligent behaviour that it can identify, is, naturally, an entirely different matter.

[1] Lake, Brenden M., et al. "Building machines that learn and think like people." *Behavioral and brain sciences* 40 (2017). Page 3. The paper can be found **here**.

Part 2

We want to know if there are any other potential tests to determine if a system is "intelligent", beyond the Turing test. To this end, we first note that the Turing test is a functional test of intelligence – a computer that acts intelligently will be considered intelligent. Could we possibly find a mechanistic criterion instead? I believe we cannot at present, since our understanding of the physical workings of the human brain is inadequate. We do not know enough of how the brain operates to be able to effectively compare it to the internals of a computer program. Moreover, we would have to define a rigorous notion of "similarity", which itself is by necessity a vague idea.
Thus, here too, we assume a functional definition of "intelligence". What we need to give examples of is intelligence by performance, and not intelligence by constitution.

I feel that any task that can set a computer apart as "intelligent" must involve natural language in some way. Language is considered to be one of human beings' most innate qualities, and for intelligent behaviour to include more than computational power (which computers surpass humans in, in any case), I feel that it must involve language skill (communicative competence) in some way (see the quoted passage above). This, I believe, is necessary and not merely sufficient.

Another important aspect of intelligence (generalising from humans) is the ability to learn. The intelligence of a computer can be measured by its performance on previously unseen task, quickly and/or on fewer inputs. In other words, I believe that a system can be considered intelligent if it can easily be instructed to do a task it has never been programmed for.

Considering the first point above, this instruction must be via natural language and not some artificial communication (like a programming langauge). For example, a description (in English, say) of the task of text summarisation given to the system should be enough for it to know how to summarise a text.