Position Paper on Appropriate Audio/Visual Turing Test

Bradley B. Custer
Screamware

bcuster@screamware.com

1. Introduction

Dr. Hugh Loebner in his 1994 article in Communications [1] makes the statement regarding future LP competitions, “the winner of the Loebner Grand Prize must develop a computer with associated hardware that can respond ‘intelligently’ to audio visual input in the manner that a human would. ... Turing wrote, ‘The question and answer method seems suitable for introducing any one of the fields of human endeavor that we wish to include.’ Well, I would like to ask questions about images and pattern recognition. If the computer answers appropriately it is intelligent.”

Some have said that requiring competitors to submit their programs to this kind of audio/visual interrogation is going overboard for what AI and robotics technology are prepared to offer [2]. The audio/visual requirement may prevent competitors from achieving the Loebner Grand Prize in the next ten years (yet I doubt it would have to take that long). However, observe how successful the scientists at Carnegie Mellon were able to be with the Navlab automated driving project in ten years [3]. Improved technologies for object and speech recognition (major component parts of the imagined program/hardware) are continuely being developed [4]. Little work has been done to consolidate neural recognition technologies and computational linguistics, to my knowledge [5]. Still, I believe that the most uncharted ground remains with determining software requirements for passing the conventional Turing Test. The ‘thinking’ element which should allow a TT to be successfully passed is still largely unknown; i.e. the modeling of executive cognitive functions.

2. A Matter of Degree

The question arises, though, as to how well sighted and acute of hearing does a machine need to be to qualify for the Loebner Prize? Drawing on an icon from science fiction, it seems to me that the intention of Loebner’s audio/visual Turing Test is to identify a machine which is in the likeness of Author C. Clark’s HAL. But does the Loebner Prize winner’s entrant need to be in the state of perfection HAL was portrayed in in 2001? I believe an initial success at linguistically identifying the quality of various simple sights and sounds would be groudbreaking. On the other hand, the principle of Turing’s imitation game is that the interviewer is prevented from distinguishing a machine’s responses from a human’s responses.

3. Possible Methods

There are several ways one might test for audial and visual comprehension. One technique would be to administer children’s intelligence tests to computers. Such as an adaptation of the WPPSI test (Wechsler Pre-school and Primary Scales of Intelligence), which requests linguistic answers to sight-vocabulary, object memory, and object relationship (e.g. which is larger) questions. Such a visually focused IQ test adaptation would not only test a machine’s object identification skills, but also it’s cognitive skills.

David Powers has recommended testing the machines ability to discuss objects of the judges choice. The objects would be placed at a given location, and the interviewee would have the chance to examine it before and during the interview. A similar arrangement could conceivably be used to test audial comprehension. A sound could be played and the interviewee asked to describe it. I would like to register support for a test after the manner which David Powers describes, believing that it would be feasible to engineer for and easy to administer.

4. Multimedia

One question looms large, though. Whether the audio and visual testing will be done discretely or together. Rather then separate the tests a single test could be administered with a noise making object. This would add another layer of complexity to contestants’ programs, matching sights to sounds. I believe deciding whether the audio/visual requirement of the Loebner Grand Prize is to be multimedia or discrete is an important issue.

If it were decided that the audio and visual tests would be combined, then another test technique would be to have the judge and contestant both watch a short multimedia clip, and conduct interviews regarding the film. Although this expectation would indeed require

Bradley B. Custer (1998) **Position Paper on Appropriate Audio/Visual Turing Test.** In D.M.W. Powers (ed.) *NeMLaP3/CoNLL98 Workshop on Human Computer Conversation,* ACL, pp 287-288.

competitors to develop recognition technology well beyond the merger of budding existing technologies (as the identifications of objects in still frames has yet a ways to go). Again, another technique might be to have the judge sit in front of a camera and microphone to conduct the interview, the computer's or confederate's responses could be returned via printed screen text. Visual questions could pertain to the judges own appearance. This could accomplish testing all three aspects of the A/V TT in one interview.

5. In Conclusion

I would like to conclude by simply thanking Hugh Loebner for the event he's started, David Powers and everyone who has volunteered time to give the contest the exciting feeling and potential it holds.

6. References

1. Loebner, Hugh G.; In Response; Communications of the ACM; June 1994/v37n6:79
2. Hutchens, Jason L; How to Pass the Turing Test by Cheating. http://ciips.ee.uwa.edu.au/~hutch/hal/Paper.txt
3. Reddy, Raj; To Dream The Possible Dream; Communications of the ACM; May 1996/v39n5:105
4. Cybenko, George; Neural Networks in Computational Science and Engineering; Computational Science and Engineering; Spring 1996/v3n1:36 also see http://www.ieee.org/nnc/
5. Lawrence, Steve; Giles, C. Lee; Fong, Sandiway; On the Applicability of Neural Network and Machine Learning Methodologies to Natural Language Processing; Tech. Report CS-TR-3479 Inst. for Advanced Comp. Studies