Otto Lappi 2017-09-01

"Twenty Questions Exam" on the Philosophy of Artificial Intelligence 5cr

The questions are based on the following references, but best results come from integrating information from across the references (maybe even other material you have read).

1. The following articles from The Stanford Encyclopedia of Philosophy (http://plato.stanford.edu):
1.1. Turing Test
1.2. The Chinese Room Argument
1.3. Intentionality
1.4. The Frame Problem

2. Z. Pylyshyn: *The Robot's Dilemma*.

3. selected bits from J.Copeland: *Artificial Intelligence, a Philosophical Introduction*. (you can use this textbook to get an overall view of the literature, if you are not familiar with it)

Three of the following twenty questions will be randomly selected for answering in the exam. You will be required to answer two of the questions chosen by the computer.

Your answers should be succinct and to the point, given in plain text form ("essay-type", not bullet points). You will have three hours in the exam. You are not expected to give exhaustive and thorough answers - answers that demonstrate you clearly understand the question and can give a well-articulated and unambiguous response will bring maximum points.

Study hint: You may want to print out each question at the top of a separate blank sheet of A4 paper, and then add your notes and comments to the sheets as you go along (filling the other side or more paper if needed!). Then, once you've gone through the chapters you can continue working on highlighting and organizing and expanding on your notes, and cross-referencing the answers. You will be using the notes as your primary study material, returning to the book (or other sources) just to check things or fill in things you might have missed, if needed.

The Questions:

1. What was the original frame problem in situation calculus (in your answer, refer to the formula below):

$$(x)(y)(c)(s)\ (True\_in(s, Color(x,c)) \rightarrow True\_in(result(s,go(x,y)), Color(x,c))))$$

2. Define the frame problem (in some way of your own choice), then discuss at least two problems in Pylyshyn's book, that according to your definition above

are *not* the frame problem. Explain how they are *related* to the frame problem and to one another, but how they *differ* from the frame problem.

3. What is the 'epistemological' frame problem (in Dennett's *Cognitive Wheels: the Frame Problem of AI*).

4. How does the simple everyday task of preparing a sandwich illustrate the frame problem? (in Dennett's *Cognitive Wheels: the Frame Problem of AI*).

5. What is the story of the "product development" from R1 to R2D2 and the moral behind it? (in Dennett's *Cognitive Wheels: the Frame Problem of AI*).

6. What is the Turing Test? Explain the original Imitation Game idea and more general interpretation of the concept.

7. What is the Chinese Room argument? Run through the argument, and explain what it is for and what it is against.

8. How does the Chinese Room Argument argument relate to *intelligence* on the one hand, and *intentionality* on the other? Be careful to explain these terms.

9. Run through at least two counter-arguments to the Chinese Room argument.

10. On balance, how compelling do you find the Chinese Room argument? Weigh considerations both for and against it.

11. How is the question can a machine *think* different from can a machine be *conscious*? (Be careful to explain what you mean by "conscious")

12. Does a) intelligence b) intentionality c) cognition require consciousness? Give a reasons to think it would, and/or counterexamples why would not.

13. Do digital computers' internal states have intentional objects? Discuss.

14. What is the physical symbol system hypothesis?

15. What does the term GOFAI refer to? Explain what it means and give examples.

16. What research tradition in cognitive science is GOFAI most closely associated with? What research tradition(s) have criticized the approach, and how?

17. What was the CYC project? How is it related to the Frame Problem of AI?

18. What could it mean to say "the human brain is a Turing machine" or

"the human brain is *not* a Turing machine", outline at least two interpretations for each.

19. In your balanced opinion, is *artificial* intelligence fundamentally different from *natural* intelligence? If so, how? If not, why do you think there is so much philosophical discussion about the special philosophical problems related to AI

20. What kinds of physical systems can exhibit *general intelligence*? Discuss.