

FOUNDATIONS OF ARTIFICIAL INTELLIGENCE - EXERCISE 1  
(50% coursework assessment; 25% course mark)

You should complete the exercises reproduced below. This assignment will be peer assessed. - [10 marks will be allocated for participation in peer marking activity following submission]  
Although you may find it helpful, the exercises do NOT require you to necessarily read the Russell and Norvig book - any introductory AI text, along with the supplementary reading indicated, will suffice. Also remember to make use of material available on the internet, where possible.

PART 1 - [60 marks]

Read Turing's original paper on AI (Turing, 1950). In the paper, he discusses several potential objections to his proposed enterprise and his test for intelligence. Below are several multiple choice questions which you should attempt to answer having read the paper. All questions are equally weighted.

- 1) In his 1950 paper Turing proposes the 'imitation game' as :-
  - a) An objective test for intelligence which can be applied to both humans and computers.
  - b) A test for whether or not a computer can be said to be intelligent.
  - c) A demonstration of how people's attitudes can change.
  - d) A test for whether or not the goal of Artificial Intelligence has actually been achieved.
  
- 2) Turing describes success in the 'imitation game' as being able to deceive a human interrogator :-
  - a) 30% of the time in five minutes questioning.
  - b) 50% of the time in five minutes questioning.
  - c) 70% of the time in five minutes questioning.
  - d) 100% of the time in five minutes questioning.
  
- 3) The sorts of 'thinking machines' that Turing will allow to take part in the imitation game' are best described nowadays as :-
  - a) All types of computers.
  - b) Neural nets only.
  - c) Electronic computers with a single CPU only.
  - d) Whatever types of computers are developed over the time period Turing specified.
  
- 4) The machines described in question three are formally known as:-
  - a) Turing machines.
  - b) Imitation machines.
  - c) Discrete state machines.
  - d) Discrete electronic machines.
  
- 5) 'Lady Lovelace's objection' is best described as the claim that:-
  - a) It is morally wrong to build intelligent machines.
  - b) Machines will never succeed in 'the imitation game' because they cannot think for themselves.
  - c) Machines will never succeed in the imitation game because they can only do what they have been programmed to do.
  - d) Machines will never succeed in they imitation game because they are incapable of artistic creativity.
  
- 6) The objection to his view that Turing seems to take the most seriously as threatening the validity of the 'imitation game' is:-
  - a) The Theological Objection.
  - b) The Argument from Extra Sensory Perception.
  - c) The Mathematical Objection.
  - d) Arguments from various Disabilities.

7) The Argument from Continuity in the Nervous System is best summarised as:-

- a] Humans can do things which they have not been programmed to do.
- b] The nervous system is continually changing, so we can't build a computer to perform in a similar way.
- c] The nervous system is continuous, but computers need separate storage and processing capability.
- d] The nervous system is continuous, so it is different from a discrete state machine.

8) If someone were to say that "computers could never really think because they don't care about what they do" this would fall into Turing's category:-

- a] 'Heads in the Sand' Objection.
- b] Theological Objection.
- c] Argument from Continuity in the Nervous System.
- d] Arguments from Various Disabilities.

9) Turing dismisses 'Lady Lovelace's objection' by asserting:-

- a] One day we will be able build to built computers which can learn to `originate`
- b] When we think of new ideas we are often unaware of the processes which leads us to them.
- c] Philosophers and mathematicians are subject to the fallacy of believing that computers can never surprise us.
- d] When we think of new ideas we do not have to work through all the steps in the way that a computer would have to.

10) Turing suggests including a `random element` in learning machines because:-

- a) This is how humans learn.
- b) In some types of search problems it is quickest to try solutions at random.
- c) Only with the 'random element' can discrete state machines hope to imitate human performance.
- d) A random answer might help to fool the interrogator.

11) The `Mathematical Objection` is best described as :-

- a] Godel's theorem proves that machines will never think.
- b] Discrete state machines cannot think because their performance is based on total mathematical accuracy whereas humans do not always use mathematics in their thinking processes.
- c] There are known formal limitations to the power of discrete state machines which don't apply to humans.
- d] Machines are subject to Godel's theorem.

12) The objection attributed to Professor Jefferson is known as:-

- a] Professor Jefferson's Objection.
- b] The Theological Objection.
- c] The Argument from Consciousness.
- d] The Argument from Continuity in the Nervous System.

13) Turing's dismissal of the `Mathematical Objection` is best summarised as:-

- a] Discrete state machines can be regarded as `universal machines` therefore the limitations of Godel's theorem do not apply to them.
- b] Turing's own paper (Turing 1937) provides the best refutation of this objection, since it deals specifically with machines.
- c] Humans give wrong answers sometimes so this objection does not matter in practice.
- d] There is no evidence that humans are not subject to the limitations talked about in this objection.

14) Turing's dismissal of the claim that a machine `cannot be the subject of its own thought` is best summarised as:-

- a] If you accept the `argument from consciousness`, then you are trapped into solipsism.

- b] If one accepts that a machine can have any thoughts about anything, then this is no problem.
- c] The problem is simply one of storage capacity.
- d] The 'imitation game' itself is the best refutation of this claim.

15) Turing thought that success in the imitation game would occur:-

- a) In 2000.
- b) Sometime in the next century.
- c) About 50 years from now.
- d) When we succeed in building an intelligent computer.

#### PART 2 - [20 marks]

AI can be characterised along a variety of dimensions. One dimension is whether we are concerned primarily with thought processes and reasoning, or whether we are concerned with behaviour and performance. Another is whether we measure success in terms of human performance or against an ideal, rational concept of intelligence. A third is whether we are interested in theoretical results or in practical applications. A fourth is whether we intend our intelligent computers to be conscious or not. Philosophers have had a lot to say about this issue, and although most AI researchers are happy to leave the questions to the philosophers, there has been some heated debate. The claim that machines can be conscious is called the strong AI claim; the weak AI position makes no such claim. Characterise the definitions reproduced below according to the four dimensions mentioned above, plus any other ones you feel are helpful. Give your own (preferred) definition of AI, categorise it, and say why you prefer it.

AI is ...

- (i) "the exciting new effort made to make computers think...machines with minds, in the full and literal sense" (Haugeland, 1985)
- (ii) "[the automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning..." (Bellman, 1978)
- (iii) "the study of mental faculties through the use of computational models" (Charniak and McDermott, 1985)
- (iv) "the study of the computations that make it possible to perceive, reason and act" (Winston, 1992)
- (v) "the art of creating machines that perform functions that require intelligence when performed by people" (Kurzweil, 1990)
- (vi) "the study of how to make computers do things at which, at the moment, people are better" (Rich and Knight, 1991)
- (vii) "a field of study that seeks to explain and emulate intelligent behaviour in terms of computational processes" (Schalkoff, 1990)
- (viii) "the branch of computer science that is concerned with the automation of intelligent behaviour" (Luger and Stubblefield, 1993)
- (ix) "a collection of algorithms that are computationally tractable, adequate approximations of intractably specified problems" (Partridge, 1991)
- (x) "the enterprise of constructing a physical symbol system that can reliably pass the Turing Test" (Ginsberg, 1993)
- (xi) "the field of computer science that studies how machines can be made to act intelligently" (Jackson, 1986)
- (xii) "a field of study that encompasses computational techniques for performing tasks that apparently require intelligence when performed by humans" (Tanimoto, 1990)
- (xiii) "a very general investigation of the nature of intelligence and the principles and mechanisms required for understanding or replicating it" (Sharples et al, 1989)
- (xiv) "the getting of computers to do things that seem to be intelligent" (Rowe, 1988)

#### PART 3 - [10 marks]

There are well-known classes of problems that are intractably difficult for computers, and other

classes that are provably undecidable by any computer. Does this mean that AI is impossible?

## References

The sources below provide the full references for the citations above. In addition, I have included some useful sources for pondering the nature of intelligence, and what computers may or may not achieve, that you may find helpful to take a look at. You are NOT expected to read EVERYTHING listed below. Look through textbook bibliographies, and those contained in the sources below, for additional sources on the topic you are pursuing.

Bellman, RE (1978) *An Introduction to Artificial Intelligence: Can Computers Think?* San Francisco: Boyd and Fraser Pub Co.

Charniak, E and McDermott, D (1985) *Introduction to Artificial Intelligence*, Addison Wesley.

Dennett, DC (1978) *Brainstorm: Philosophical Essays on Mind and Psychology*. MIT Press.

Dreyfus, HL (1979) *What Computers Can't Do*. Harper and Row.

Ginsberg, M (1993) *Essentials of Artificial Intelligence*. San Mateo, Calif: Morgan Kaufmann.

Haugeland, J (Ed) (1985) *Artificial Intelligence: The Very Idea*. Cambridge, Ma: MIT Press.

Jackson, P (1986) *Introduction to Expert Systems*. Reading, Ma: Addison-Wesley.

Korf, RE (1988) Search: a survey of recent results. in Shrobe, HE (ed) *Exploring Artificial Intelligence: Survey talks from the National Conferences on AI*. Morgan Kaufmann.

Kurzweil, R (1990) *The Age of Intelligent Machines*. Cambridge, Ma: MIT Press.

Luger G.F. and Stubblefield, W.A (1993), *Artificial Intelligence: structures and strategies for complex problem solving*, Benjamin Cummings Inc.

McCorduck, P (1979) *Machines who think: a personal enquiry into the history and prospects of artificial intelligence*. WH Freeman and Co.

Nilsson, NJ (1995). Eye on the prize. *AI Magazine*, Vol 16(2), pp9-17. [QE 1 Aim]

Partridge, D (1991) *A New Guide to Artificial Intelligence*. Norwood, NJ: Ablex

Rich, E. and Knight, K, *Artificial Intelligence*, McGraw-Hill, 1991.

Rowe, NC (1988) *Artificial Intelligence through Prolog*. Englewood Cliffs, NJ: Prentice-Hall.

Russell, S and Norvig P (1995) *Artificial Intelligence: A Modern Approach*. Englewood Cliffs, NJ: Prentice-Hall. (Or second 2003 edition.)

Schalkoff, RJ (1990) *Artificial Intelligence: An Engineering Approach*. NY: McGraw-Hill

Sharples, M, Hogg, D, Hutchinson, C, Torrance, S and Young, D (1989) *Computers and Thought: A Practical Introduction to Artificial Intelligence*. Cambridge, Ma: MIT Press.

Tanimoto, S (1990) *The Elements of Artificial Intelligence Using Common LISP*. Rockville, ML: Computer Science Press.

Turing, AM (1950) Computing machinery and intelligence. *Mind*, Vol.59 pp433-60. Also available at <http://www.abelard.org/turpap/turpap.htm> (which includes the bibliography and footnotes). However, you will find the version at <http://www.loebner.net/Prizef/TuringArticle.html> more printable.

See also:

<http://www.turing.org.uk/turing/>

<http://www.turing.org.uk/turing/scrapbook/test.html>

<http://www.cs.unr.edu/~sushil/class/ai/papers/coffeehouse.html>

And the philosophers among you might want to try this

<http://users.ox.ac.uk/~jrlucas/Godel/turitest.html>

Weizenbaum, J (1976) *Computer Power and Human Reason*. Pelican.

Winston, PH (1992) *Artificial Intelligence*. Reading, Ma: Addison-Wesley. 3rd Edition.