ELL788/HSL622: Minor Quiz

Sumeet Agarwal

16 March 2021

1. Suppose we go out for dinner, and after the main course, I ask you which ice cream flavour you would like to have for dessert: chocolate or vanilla? You hesitate for a few moments, then say, "I think I'll go for chocolate". I respond: "Aha! Surely that shows you have free will, that's why you took a few seconds to make up your mind! If the mind was deterministic, you would have known what you wanted instantly!"

   Is this a valid argument in favour of free will? Why or why not? Let's say you are a determinist because you believe the mind is essentially a kind of computer; how would you respond to my assertion, and what would be your explanation for why you hesitated? **[3]**

   Key points:

   - Not really a valid argument; just the fact that a decision took time does not imply that it involved the action of 'free will'
   - A deterministic physical or algorithmic process would in general also take some time; one billiard ball set in motion would take some time to hit another
   - The hesitation could be seen as reflecting the time taken by an algorithmic decision-making process which takes various factors into account to arrive at an outcome

2. In the context of their physical symbol system hypothesis (PSSH), Newell and Simon say that

   Symbols lie at the root of intelligent action, which is, of course, the primary topic of artificial intelligence. For that matter, it is a primary question for all of computer science.

   What exactly do they mean by this? Are they suggesting that at some level, the disciplines of artificial intelligence and computer science are fundamentally the same? Explain, in your own words, what the PSSH implies for the relationship between these two disciplines. **[5]**

   Key points:

- The PSSH implies that a symbol manipulation system with enough capacity is both necessary and sufficient for intelligence; since computers are also such systems, this means computers in general have the potential for intelligent action
- So yes, the PSSH implies that at some level, AI and computer science are fundamentally about the same kinds of processes and phenomena, *i.e.*, symbol manipulation processes in general, which equate to processes that can achieve intelligence
- The judgement of when the behaviour of such systems crosses the threshold of 'intelligence' is subjective and dependent on some notion of complexity (as per humans); so when these systems are carrying out less complex algorithmic tasks, tasks with less variability or unpredictability, then we may not regard them as 'intelligent' as their capacity for intelligence may be obscured
- But the same kinds of systems and processes, once applied to tasks which involve adaptation to variation and contingency, exhibit their capacity for intelligent action; and hence fundamentally the study of intelligence is just the study of such symbolic, computational processes
- Newell and Simon also suggest that computer science is the *empirical* inquiry of how symbols operate in nature, which implies that they regard computer science as being about the same themes as the study of biological information-processing or cognitive/intelligent systems

3. Suppose that the PSSH is true. What would that imply for the Mind-Body problem? How would a believer in the PSSH characterise the relationship between mind and body? Explain why. **[3]**

   Key points:

   - The PSSH says that intelligence/cognition necessarily and wholly arises from a *physical* symbol system, hence it clearly implies a physicalist position on the mind-body problem
   - As per the PSSH, any such system with sufficient capacity can exhibit intelligent behaviour, and so such behaviour should not be dependent on the specific physical structure or composition of the brain/body
   - Hence this implies a *functionalist* position, where the mind is seen as a functional kind rather than a physical kind, which could have multiple physical realisations such as human brains/bodies or modern digital computers

4. Consider the game of Connect Four, which is played by two players (red and yellow) on a 6x7 grid which is vertically suspended. At each turn, each player drops a disk of their respective colour into one of the 7 columns, and it falls to the lowest unoccupied slot in that column. The game continues

until one player gets four of their disks in a line (vertical, horizontal, or diagonal), in which case that player wins; or until no more moves are possible (the grid is full). The below animation shows an example of the gameplay for a game where the red player wins.

[https://en.wikipedia.org/wiki/Connect_Four#/media/File:Connect_Four.gif]

We would like to think of the process of playing this game (*i.e.*, the cognitive processes of the players of the game) as a production system. Recall from class that we can think of production systems as having 3 components: a global database, a set of production rules, and a control structure.

(a) What should the global database consist of here? Try to describe it as precisely as you can. **[2]**

(b) Give two examples of what you think might be reasonable production rules used by the players of the game. Try to think of rules which a good player of the game might make use of. **[2]**

(c) What do you think constitutes the control structure for this production system? **[1]**

(a) All information about the current state of the game; basically, everything needed to capture a single frame/snapshot of the provided animation – so the global database needs to store, for each of the 42 grid locations (or symbolically, elements in a 6x7 array), its current state: *vacant*, *occupied by red disk*, or *occupied by yellow disk*

(b) *E.g.,*

- IF a column has 3 consecutive disks of my colour, and the space immediately above them is vacant, THEN drop a disk into that column
- IF a column has 3 consecutive disks of the other colour, and the space immediately above them is vacant, THEN drop a disk into that column

(c) An optimisation process (probably involving lookahead) in the player's mind, which searches over all applicable production rules in the current game state and picks the one which maximises the (estimated) likelihood of that player winning the game