KASPAROV VERSUS DEEP BLUE: AN ILLUSTRATION OF THE LUCAS GÖDELIAN ARGUMENT

Alex Hankey

ABSTRACT: This article concerns philosophy of mind, in particular the question, is the human mind more than a digital machine? A previous article took up that question from the perspective of the Lucas Gödelian Argument, which proposed an answer to the question from the point of view of metamathematics. That article pointed out that, the present full scientific status for direct mind-to-mind communication, demonstrates a field of phenomena where no computer can hope to imitate human and animal abilities. This article aims to disprove the argument put forward by the IBM team that programmed the Deep Blue computer, that because their computer beat World Champion at Chess, Garry Kasparov, Deep Blue was more powerful than the human mind. To do this, the article recounts the events of the 1997 IBM challenge, following Kasparov's 1996 defeat of Deep Blue. It seeks to support Kasparov's own account of the two chess matches, and also to present a brief version of the reasoning presented in the previous article.

KEYWORDS: Lucas; Gödelian Argument; Artificial intelligence; Kasparov; Deep Blue

INTRODUCTION – THE GÖDELIAN ARGUMENT

In 1961, philosopher J.R. Lucas presented a paper[1] to the Oxford Philosophical Society on the thesis that Kurt Gödel's system of metamathematics[2-4] implies that the human mind is more than a digital machine. Metamathematics demonstrates in a fully logical fashion that it is possible to reason about mathematics, and derive a valid statement, without carrying out detailed logical deductions using the set of axioms defining the field concerned, as long as that field contains arithmetic. [4] Such statements are independent of the system of mathematics itself, In contrast, if a mathematician wants to extend a particular field of mathematics, then he / she has the option of introducing a new axiom, which describes a new property of the mathematical entities concerned. Increasing their structure makes them represent new mathematical entities requiring new names. In the field of abstract algebra[5], this is well known to undergraduates in mathematics: Groups can be transformed into Rings, and then into Fields by the addition of appropriate axioms.

Lucas's argument[1] concerned the analogy between computers as logical machines; machines that can process statements in a logical way. To many philosophers and philosophers of mind in the 1960's, the ability of digital computers to process statements in a self-consistent way, in accordance with the rules of logic, was regarded as proof that: (a) such a digital machine can be at least as powerful as the human mind, if not more so; and (b) that the mind is no more than a digital machine.

One such was the Professor of Philosophy at Harvard, Willard Quine[6], a not unjustly famous figure in his field, but now known to have many wrong opinions concerning work in other fields, e.g. about free-will, which he resolutely refused even to discuss with Lucas. Another case is astrology, where the work of my colleague, Rameshrao, advancing Indian Jyotisha astrology to a science[7], has showed Quine's dichotomy, that, "one must either reject astrology, or accept astrology but reject all established scientific disciplines that are incompatible with astrology", to be completely erroneous.

In the field of computer science, a computer can only be programmed to emulate an ability of the human mind one fact at a time.[1,8] Each improvement in computer performance can be added with a single line of programming, as it were (could be quite a long 'line'). A very good example of such limitations was provided by the challenge of IBM's team of programmers[9,10] for its supercomputer, Deep Blue, to world chess champion, Garry Kasparov, for Chess Match over six games in 1996[11], and again in 1997[12], with an upgraded version of Deep Blue.

In 1996, Deep Blue won the first game on 10th February, against Kasparov playing White.[13] The game was noteworthy for being the first chess game in which a computer had beaten a reigning World Chess Champion under match conditions with time controls. The following day, on 11th February, Kasparov, now playing white, won the second match, when the IBM team intervened for Deep Blue, which was 3 pawns to 1 down after 73 moves. The third and fourth matches on 12th and 13th February were both drawn, leaving Kasparov and Deep Blue on 2 points each. Game 5 on 14th February[14] is considered the turning point of the match, with Kasparov adopting a strategy of blocking all Deep Blue's attempts at developing its pieces. The game ended in a position so hemmed in that its two castles were essentially worthless. The Deep Blue team, which had refused the offer of a draw on move 23, resigned when an easy opportunity for queen exchange would have left Kasparov in a dominant position with a passed pawn threatening Deep Blue's two castles on its back rank (a8 & c8). This was the only game won by Black.

The final game on 15th February, had Kasparov playing White. From the present perspective, the 6th game is noteworthy because Kasparov won by adopting a strategic approach, which made specific tactical sequences that Deep Blue could have used to its advantage secondary or impossible to implement. [15] Strategy of this kind requires thinking about the overall state of a chess board from an holistic perspective. Despite such a consideration coming effortlessly to the human mind, it is not something that computer programmers can achieve. Computer programs can put value on specific objectives like winning a pawn or a piece, or on the advantage of an exchange. But they cannot put values on the overall structure of a position with the same uncanny insight of human experience.

This illustrates the fact that logical processes can be used to solve closed problems with well-defined solutions, but not open problems where relative values play roles in decision-making processes. Kasparov's success may be attributed to his converting what was seemingly only a sequence of closed problems, how to win pieces without losing any oneself, to an open problem. There, the means to implement a specific strategy outweigh all tactical considerations – a real act of genius. Deep Blue might be a master of the first kind of game, but was nothing more than a helpless infant when faced by the second kind of game, as became evident as game 6 unfolded[15], and Deep Blue clearly had no idea how to proceed, unable to prevent itself being pinned down, and then shot to pieces.

Despite this demonstration of the inherent superiority of Mind over Machine, IBM renewed Deep Blue's challenge to Kasparov the following year. In 1997, the match lasted from 3rd May to 11th May with the 5th, and 8th and 9th May being omitted. [12] Kasparov won the first game, on 3rd May, possibly because Deep Blue could not decide on a move towards the end, and made a fatal error in trying to select a 'Fail-Safe' move. [16] However, the computer, now playing White on 4th May, came back and took the second game, but its win was not without controversy. [17,18] Kasparov pointed out that one particular move was 'far too sophisticated' for a computer to make[17], and accused Deep Blue's team, which included a Grandmaster rival, of inputting the move into the computer. [19] Also, Kasparov missed an opportunity to force a draw by perpetual check which was pointed out to him the following day. [12]

The score of 1-1 after the second game is therefore questionable. Our viewpoint once again points out that a computer is incapable of the kind of strategic consideration, which goes into making the kind of move that defeated Kasparov[17], and which his rival had given Deep Blue. Kasparov was only too aware of this non-human limitation, and played to exploit it. But even he missed the simple solution to his apparently lost position. [12]

In the 3rd game on 6th May, Kasparov used an irregular opening, the Mieses opening, reasoning that the computer would not be able to figure out how to handle an opening game not in its memory. Deep Blue, however, did well, and managed to force the game to a draw. The 4th game was played the following day, on 7th May, with Deep Blue playing White. Kasparov chose the Caro-Kann defence, but did not play his best late in the game due to time problems. As a result, a potential win ended in a draw. [12] The game was unfortunate in far greater ways, however, because it set Deep Blue's team thinking about possible variations of the Caro=Kann defence. After 4 games the match stood at 2-2.

In the 5th game played on May 10th, Kasparov, playing White, secured an excellent position in the mid-game after the opening, but Deep Blue was outstanding in the endgame and managed to secure a draw after all. [12] The moral of games 4 and 5, in both of which Deep Blue did excellently in the endgames, is that a computer does far better when detailed tactics are the most important factor, and there are little of no strategic considerations involved – or if there are, lengthy logical analysis of tactical moves can overcome them. After the 5th game, the score therefore stood at 2½-2½.

In the 6th and final game on 11th May, 1997, Kasparov playing Black again used the Caro-Kann defence. [12] He utilized a variation, which though fatal, requires a move from White, which would be impossible for a computer to find through selection of the kind that was at its disposal. As it turned out the IBM team had input that very variation into Deep Blue's memory that morning, possibly because Kasparov's previous use of the Caro-Kann defence in Game 4 had drawn their attention to it, and they wanted to guard against its repeated use. Faced with the fatal variation, Kasparov resigned after 20 moves, and with it, the 1997 challenge.

From our perspective however, the result of the 6th game is completely consistent with the Lucas argument. [1,8] A computing machine can only acquire new skills by being programmed one line at a time – and in this case that line, the particular variation of the Caro-Kann defence proved fatal to Kasparov, and won both game and match for Deep Blue.

DISCUSSION

Overall, the 1996[11] and 1997[12] challenges ended one point in Kasparov's favour; however, the press covering the event decided to interpret the result of the 1997 challenge on its own, and conclude that computers more powerful than the human mind can be constructed. Such an interpretation fails to take into account the facts that the unlikely move in 1997, Game 2 was human input by an advisor; Kasparov's strategic method of winning 1996, Game 6, was something for which the computer had no answer; that, in 1997, Game 1, the computer had been left like Pavlov's Dog, unable to make a decision; and that the victory of Game 6 in 1997 was entirely due to the additional line of program entered into Deep Blue's memory that very morning. Without that input, the one move turning Kasparov's choice into a fatal error would have been impossible for Deep Blue to find and select. [19]

In these ways, a sequence of the kind proposed by J.R. Lucas[1,8] was exemplified in Deep Blue's 1996 and 1997 challenges to Garry Kasparov. [11,12] There is no doubt that Kasparov's 1996 win led to improvements in Deep Blue. [10] Then, to continue outwitting the upgraded Deep Blue, Kasparov implemented new, deep strategies. Kasparov[19] was then right to assert that the IBM team had utilized Grand Master input in Game 2[17]; and in that same game, he had missed a draw by perpetual check[12]; also that, in the final game, Game 6, the Deep Blue team had, that same morning, anticipated Kasparov's tactic in the Caro-Kahn defence, by inputting the very variation he selected into Deep Blue's database.[19]

All this is consistent with Lucas's Gödelian Argument that digital machines can only be improved a single skill at a time, represented by one line of program: ***in order to win, the computer had to be updated one fact at a time***. This means that the IBM claim that Kasparov versus Deep Blue proves the 'Superiority of Machine over Mind'[10] fails. Rather, it offers an example of Lucas's Gödelian argument.

The above descriptions show that Deep Blue had no strategic understanding from which to reason about the game of chess, a point made strongly in the book, Gödel Escher Bach by Douglas Hofstadter. [20] Digital Machines like Deep Blue are limited to performing algorithmic calculations within the scope of their programming, and can therefore only be improved one calculation at a time.

Kasparov's book, Deep Thinking – When Machine Intelligence Ends and Human Creativity begins[19], concerning creativity as an aspect of human intelligence: makes exactly this point. Today's computers are far more powerful than Deep Blue, the most powerful thought to be the Supercomputer at MIT's Lincoln Laboratory. But even such behemoths are still capable of making strategic errors of the kind made by Deep Blue a quarter-century ago, exemplified by its failure to resign in Game 6 of the 1996 challenge. Deep Blue's team had to resign for the machine, because the Machine itself could not 'see' that its position was hopeless! [15]

A related example was given by the 2020 Nobel Laureate in physics, Sir Roger Penrose, who pictured a position where a computer would automatically choose to win a rook for a pawn and go on to lose the game, because it was unable to 'see' that such a move would inevitably transform a drawn position in a lost one. [21]

In the earlier article on Lucas's Gödelian Argument[1], the complexity biology model of Self-Awareness[22] was shown to employ perfectly self-observing loops, so that normal systems of such perfectly self-observing loops can encode ideas[23] or 'forms'. [24] Another article has showed how such loops enable human consciousness to 'witness' its functioning[25], as described by Vedic texts from ancient India[26,27], and supremely trained modern athletes, who have reported, 'As if seeing themselves from the outside', when performing a skill for which they have extreme training[28]; and that, through the witness function, the Mind can 'Reason about a Problem from the Outside'[29], in the way that Lucas proposed[1] when putting forward his Gödelian Argument.

CONCLUSION

Deep Blue's win in IBM's 1997 Challenge to Garry Kasparov demonstrated that, in simple situations like endgames, a computer of that time could perform algorithmic calculations well enough to compete with a world champion chess player. However, because of Lucas's argument, the challenge really showed that a computer can only do as well as its program, which has to be upgraded one line of program at a time, as it were. A computer cannot develop conceptual strategies in the same way as the human mind; therefore, digital machines cannot be compared to mind, let alone be considered superior to mind.

alexhankey@gmail.com

REFERENCES

1. Lucas JR. Minds, machines and Gödel. *Philosophy*. 1961; 36(137):112-27.
2. Gödel K. *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. Monatshefte für mathematik und physik.* 1931 Dec 1; 38(1):173-98.
3. Gödel, K. (1992). On formally undecidable propositions of *Principia Mathematica and related systems*. Courier Corporation.
4. https://en.wikipedia.org/wiki/Metamathematics
5. Grillet, P. A. (2007). *Abstract Algebra.* Springer Science & Business Media.
6. https://en.wikipedia.org/wiki/Willard_Van_Orman_Quine
7. Rameshrao N. Hankey A. (2020) *The Science of Medical Astrology: the experimental proof of Jyotisha.* Notion Press, Chennai.
8. Lucas JR. *The Gödelian Argument: Turn over the page.* 1989 See: https://core.ac.uk/download/pdf/41174228.pdf Retrieved 2021-05-24
9. Campbell M, Hoane AJ, Hsu FH. *Deep blue. Artificial intelligence.* 2002;134:57-83

10. Hsu FH. Behind *Deep Blue: Building the computer that defeated the world chess champion*. Princeton University Press; 2004.
11. https://en.wikipedia.org/wiki/Deep_Blue_versus_Garry_Kasparov#1996_match
12. https://en.wikipedia.org/wiki/Deep_Blue_versus_Garry_Kasparov#1997_match
13. https://en.wikipedia.org/wiki/Deep_Blue_versus_Garry_Kasparov#Game_1
14. https://en.wikipedia.org/wiki/Deep_Blue_versus_Garry_Kasparov#Game_5
15. https://en.wikipedia.org/wiki/Deep_Blue_versus_Garry_Kasparov#Game_6
16. Hornyak T. "Did a bug in Deep Blue lead to Kasparov's defeating it?". cnet.com. Retrieved 2021-05-24
17. "Deep Blue's cheating move". *Chess News*. 2015-02-19. Retrieved 2021-05-24
18. "Kasparov on the future of Artificial Intelligence". *Chess News*. 2016-12-29 Retrieved 2021-05-24
19. Kasparov G. *Deep Thinking: Where Artificial Intelligence Ends And Human Creativity Begins*. John Murray, London, UK, 2017.
20. Hofstadter D. *Gödel Escher Bach, an Eternal Golden Braid*. Basic Books, New York, NY, 1979.
21. https://en.chessbase.com/post/a-chess-problem-holds-the-key-to-human-consciousness
22. Hankey, A. (2014). Complexity biology-based information structures can explain subjectivity, objective reduction of wave packets, and non-computability. *Cosmos and History: The Journal of Natural and Social Philosophy*, 10(1), 237-250.
23. Hankey A. (2015) A complexity basis for phenomenology: How information states at criticality offer a new approach to understanding experience of self, being and time. *Progress in biophysics and molecular biology*, 119(3),288-302.
24. Thom R. *Structural Instability and Morphogenesis: an essay on the mathematical properties of models*. Addison Wesley, Reading, Mass 1975.
25. Shetkar R. Hankey A. Nagendra H.R. (2017) First Person Accounts of Yoga Meditation yield clues to the Nature of Information in Experience. *Cosmos and History: The Journal of Natural and Social Philosophy*, 13, 252-53.
26. Mundaka Upanishad in Radhakrishnan, S. (1953). *The Principal Upanishads*. London: Allen.
27. Soneji R. Hankey A. Sridhar M.K. Nagendra H.R. Essential but Seldom Taught Yoganga. Submitted to Yoga Mimansa, February 2021.
28. https://en.wikipedia.org/wiki/Simone_Biles
29. Hankey A. (2021) Bringing the Lucas Argument to Completion, submitted to *Cosmos and History: The Journal of Natural and Social Philosophy*, 17. May, 2021.