# Artificial Intelligence and Consciousness

Daniel Verdejo

***Abstract*— The following paper will delve into Artificial Intelligence (AI) and its potential impact in the years to come, and specific industries or topics. The first topic will review artificial intelligence in video games, how it is currently implemented across different genres of games, how it enriches or enhances the overall experience of a title for the end user, and where it could potentially go in the future if developers begin utilizing technologies such as neural networks and machine learning. The next topic will review artificial consciousness, how can researchers test to see if a machine is self-aware or conscious of its current situation. Could machines be considered living entities in the future? Could our limitations, whether they are related to knowledge, ethics, social, or political, be the reason that artificial intelligence will struggle, or take longer than it would naturally, to reach its true potential, with no restrictions applied externally by us humans, or is it a technical restraint such as hardware?**

## I. INTRODUCTION

In 1950 Alan Turing, an English Mathematician, published a paper titled "Computing Machinery and Intelligence". This would be the foundation upon which the field, later known as Artificial Intelligence (AI), a term coined by John McCarthy, would begin. Turing opened his paper by asking a simple, yet complex question: "Can machines think?"[1]. In the paper, Turing continued to propose a way in which we would be able to test if a machine can exhibit thought, similar to, or indistinguishable from a human brain, which in his paper he called 'The Imitation Game'[1] later known as the Turing test.

While the Turing Test is a general examination to determine whether a machine exhibits intelligence similar to that of a human, chess a strategy game which requires two players, is also a valid test of intelligence. Claude Shannon researched the idea of developing a program capable of playing chess [2]. In his paper he described two different approaches to program chess:

- **Type A:** In this approach the computer would examine thousands of moves and use a min-max search algorithm to play. While that sounds acceptable because of processing power of machines in 2017, Claude estimated that the program would play slowly and poorly, as he put it "It would be slow since even if each position were evaluated in one microsecond (very optimistic) there are about

  $$(10^9)$$

  evaluations to be made after three moves (for each side). Thus, more than 16 minutes would be required for a move or 10 hours for its half of a 40-move game."[2]. Due to limitations of computers during the 1950s and 1960s, this would not be the favourable approach, as the program's skill would be directly tied to that of the speed of the machine it is running on. Where the limitations of a computer may have been a concern back then, in more modern times this is less of a concern. In 1997 the best known Type-A program: Deep Blue, developed by IBM, defeated the world chess champion at the time, Gary Kasparov. The match was not a definitive victory with a score of 3.5/2.5, but with the increasing power of machines, it would be an insight of what was to come.
- **Type B:** In this approach the computer would focus on only a few, key possible moves, utilizing a specially designed heuristic (the ability to learn or discover) and strategic artificial intelligence. As the power of a machine had no correlation to the skill the Type B program would exhibit, it would be the preferred method to develop chess playing programs back in the 1950s, 1960s, and even 1970s, but it would need to be taught new strategies, and rules in order for it to become better at the game.
  As computers became more powerful, Type B programs would become less desirable and developers began moving towards Type A programs, as it is easier to develop and less volatile. The authors of the 'Chess' series of programs expressed their stress when their Type B program having hard-coded rules would act in a volatile manner during tournaments. Thus in chess 4.0, they moved to a Type A program, which went on to win ACM computer chess championships, according to Wikipedia [3].

Around the 1960s and 1970S, a subset of artificial intelligence called Expert Systems became more widespread within' industry applications. Some of the most well-known expert systems: MYCIN a medical diagnosis system, PROLOG a system which handled linguistics, primarily natural language processing, and DRENDAL a tool to analyse molecules, to name a few, were considered revolutionary at the time, but they would also be considered a threat to humanity for the very same reason[4]. More recently the same feeling was portrayed by theoretical physicist, Stephen Hawking during an interview with the BBC (British Broadcasting Corporation) in 2014[7], and at a talk, he delivered at Web Summit 2017[8]. The current Russian President Vladimir Putin said the following: "Whoever becomes the leader in this sphere will become the ruler of the world" while speaking about Artificial Intelligence to

students during an open lesson on the 1st of September 2017[5]. Elon Musk, the chief executive officer of one of the largest companies leading research and development in technology, reacted to a statement made by Putin, stating that: the competition or race to AI superiority on a national level would be the cause of World War Three[6].

## II. Artificial Intelligence in Video Games

In recent times, video or computer games have become a large part of the computer and entertainment industries. In 2016 the gaming industry made over one-hundred billion dollars[9]. It is projected to do so again in 2017, and continue to do so in the following years, as anticipated by a report from NEWZOO in April 2017[9]. So why is this important to artificial intelligence research? The majority of games will implement some sort of artificial intelligence, whether it be simple, or complex, there is an artificial intelligence in the game to control the non-player characters. The games industry has identified the need for superior artificial intelligence in games, beyond what has already been implemented, and devotes a large amount of research into developing artificial intelligence[10]. In their current state the typical non-player controlled AI will be able to carry out a number of actions, such as: navigating themselves around the virtual world they inhabit, react to what is happening around them, seek the player, interact with or attack the player, make decisions during combat based off what is around them, for example hide behind a pillar if they are being shot at. That all sounds very impressive, and it is, but they still act differently to how humans play games, they still lack the human element of unpredictability. Why is this?

Game designers and developers would typically be experienced and knowledgeable in computer art or programming, but not specifically in the field of artificial intelligence[10]. Also, artificial intelligence to be able to function appropriately requires an abundance of computational power from the machine it is running on[10]. The typical target hardware for the majority of games developed in modern day would be a games console. Games consoles normally would not be the most powerful machines in order to keep the cost of the device down but have made many advancements, especially in the last five years. Games consoles are coming closer to matching the power of personal computers that are a lot more powerful as far as the components its built with are concerned, built specifically with the intent of playing games and running other high-intensity applications. The games consoles hardware architecture is optimized to run games and other entertainment applications flawlessly, this indicates that the hardware should be capable of running games that have non-player characters utilizing compelling artificial intelligence[10].

### A. Artificial Intelligence: A living entity?

One of the more sophisticated examples of a non-player character AI in recent years is in 'Alien Isolation', a game inspired by Ridley Scott's: 'Alien', developed by Creative Assembly. The developers created a sophisticated artificial intelligence for the primary enemy which the player would encounter continuously throughout their play-through of the game. It behaves differently than your average in-game AI, it is unpredictable. Its actions are carried out based on the AIs neural network of behaviours. This means it is not scripted to do a certain action when the player enters a certain area or carries out a certain task. The further the player progresses into the game, the more you encounter this AI, and the more the AI learns about the player's habits, and how the player behaves. If the player has a tendency to repeat tactics frequently to evade the enemy, but it learns of these tactics by witnessing the player doing said tactics multiple times, it will begin to investigate those actions when seeking the player throughout the environment. The AI adapts and learns how the player behaves and reacts to those behaviours by forming behaviours of its own. The Design Lead of 'Alien: Isolation' Clive Lindop said the following about the alien during an interview with CVG in 2014, Clive Lindop: "He is a living entity for us". The idea of an AI being a 'living entity' creates a narrative of its own, could an AI become 'self-aware', or gain 'consciousness'?

## III. Artificial Consciousness

What is consciousness? According to the Oxford dictionary: "The state of being aware of and responsive to one's surroundings"[14]. There is a field of research which is related to both artificial intelligence, and cognitive robotics, named Artificial consciousness (also referred to as "machine consciousness" or "synthetic consciousness")[12]. The aim of the theory according to a retired professor of Neural Systems, Igor Aleksander: "to define the characteristics of state structure that are necessary for and specific to organisms that are said to be conscious."[15]. In the field of Artificial consciousness, researchers attempt to take advantage of the attributes of a human being to map the prospective for synthetic beings[11]. A recent approach called 'neuromorphic engineering' (NE) suggested by Carver Mead in 1989, described how one would build an electronic neural system, where the architectural principles would imitate that of a biological neural system[12]. While NE has made considerable developments, a monumental conceptual challenge lays ahead before any more substantial progress can be made in conscious neuromorphic machines. The aforementioned challenge is to bridge the gap between 'Weak Artificial Consciousness' and 'Strong Artificial Consciousness'[13][12].

- **Weak Artificial Consciousness** This would imitate or simulate cognitive functions that are often associated to consciousness[13].
- **Strong Artificial Consciousness** Absolute consciousness[13].

### A. Consciousness in synthetic beings

The Rensselaer Artificial Intelligence and Reasoning Lab (RAIR) have been researching the subject matter, "how expressive a knowledge representation must be before

it can perform the sorts of tasks associated with self-awareness."[16]. A previous successful test carried out by the group was the 'mirror test', on a robot named Cogito. The mirror test also known as the mirror self-recognition test was developed by psychologist Gordon Gallup Jr. in 1970. The mirror test is a behavioural technique to determine if a non-human being such as an animal, or in this case an AI or Robot, has the ability of visual self-recognition. This test is a classic approach to measure self-awareness but has yet to be validated as a true indicator of self awareness[17]. More recently, the RAIR Lab group were able to successfully test another technique of testing self-consciousness. They posed a task which involved three NAO robots, giving two of the robots what they called 'dumbing pills', which would disable them from speaking, and the third a placebo. The three robots were made aware of the rules of the task prior to carrying out the test. RAIR made a video documenting this test.

The video begins with the tester activating the three robots by carrying out the action of 'giving each of them their respective pill'. After this, the tester poses the question to the three robots: "Which pill did you receive?", to which the robot given the placebo, stands, and responds "I don't know". The robot pauses for a moment, upon realizing that it had just spoke, it then waves and speaks again: "Sorry! I know now, I was able to prove that I was not given the dumbing pill."[16][18]. This successful test was achieved using a hierarchy of Cognitive Calculi, a group of formalisms which includes the Deontic Cognitive Event Calculus (DCEC*)[19]. The description of this as per the formal specification is: "DCEC* (deontic cognitive event calculus) is a multi-sorted quantified modal logic that has a well-defined syntax and a proof calculus. The syntax of the language of DCEC* and the rules of inference for its proof calculus are shown in Figure 1. DCEC* syntax includes a system of sorts S, a signature f, a grammar for terms t, and a grammar for sentences ;"[19]. While this may seem simple to the average human being, it is not a task easy for a robot to complete. The AI first has to listen to the question and understand what is desired from it. The robot then needs to hear its own voice and recognise that it is noticeably different from the other robots on the table. Once it has done this it then needs to correlate its realisation to the original question, and form an answer to respond with. This test is circumstantial, and it does not represent the same magnitude of self-awareness which human beings possess, but still represents a progression in the fields of both artificial consciousness and artificial intelligence.

## IV. Conclusion

Is artificial intelligence finally coming to the point where it could be a powerful tool to assist and push humanity forward? With the fast advancement of computational power within' the last ten to twenty years, it is now more feasible than it was in the 1950s during its inception. Does this mean we could see growth in research of artificial intelligence? With video games becoming a larger part of the entertainment industry, and their need for more sophisticated AI, like that what is in 'Alien: Isolation', this could be a good indication that we may see growth. The point at which the NAO robot pauses after it firsts speaks, during its moment of realization, is it thinking? Could this little robot have answered the question originally posed by Turing: "Can machines think?"? The successful test by RAIR Labs could bring us closer to a definitive answer to this question. If the robot were actually thinking, could this see artificial intelligence and robotics research descend into developing conscious synthetic beings that could be considered "a living entity", as Lindop described the non-player character AI in 'Alien: Isolation'. If so, how will this impact humanity?

## References


[1] A. M. Turing (1950) Computing Machinery and Intelligence. Mind 49: 433-460.

$$http://cogprints.org/499/1/turing.HTML$$

,Date accessed: 21 - Nov - 2017.

[2] C. E. Shannon (1950) Philosophical Magazine, Ser.7, Vol. 41, No. 314 - March. XXII. Programming a Computer for Playing chess 1,

$$https://users.dcc.uchile.cl/\ cgutierr/cursos/IA/shannon.txt$$

, Date accessed: 21 - Nov - 2017.

[3] Computer chess (21 Nov 2017) Wikipedia

$$https://en.wikipedia.org/wiki/Computer_chess$$

Date accessed: 21 - Nov - 2017.

[4] C. Smith, B.McGuire, T. Huang, G. Yang, (2006) University of Washington, The History of Artificial Intelligence,

$$https://courses.cs.washington.edu/courses/csep590/06au/projects/histo$$

, Date accessed: 15 - Nov- 2017.

[5] RT, (1 Sep 2017) 'Whoever leads in AI will rule the world: Putin to Russian children on Knowledge Day

$$https://www.rt.com/news/401731-ai-rule-world-putin/$$

Date accessed: 22 - Nov - 2017.

[6] A. Hern, The Guardian (4 Sep 2017) 'Elon Musk says AI could lead to third world war'

$$https://www.theguardian.com/technology/2017/sep/04/elon-musk-ai-$$

Date accessed: 22 - Nov - 2017.

[7] R. Cellan-Jones, BBC (2014) Stephen Hawking warns artificial intelligence could end mankind

$$http://www.bbc.com/news/technology-30290540$$

Date accessed: 22 - Nov - 2017.

[8] S. Hawking, (2017) Talk on AI at Web Summit

$$https://www.youtube.com/watch?v=FjgQOOVgIoM$$

, Date accessed: 22 - Nov - 2017.

[9] E. McDonald, Newzoo (2017) THE GLOBAL GAMES MARKET WILL REACH $108.9 BILLION IN 2017 WITH MOBILE TAKING 42%

$$https://newzoo.com/insights/articles/the-global-games-market-will-$$

, Date accessed: 22 - Nov - 2017.

[10] D. C. Moffat, D. M. Romano, Artificial Intelligence and Games: Applications and New Directions, Date accessed: 20 - Nov - 2017.

[11] R. Manzotti, V. Taglias (2008) Artificial consciousness: A discipline between technological and theoretical obstacles,

$$http://www.sciencedirect.com/science/article/pii/S0933365708000912$$

Date accessed: 25 - Nov - 2017.

[12] A. A. Fingelkurts, A. A. Fingelkurts, C. F.H. Neves, (2011), "Machine" Consciousness and "Artificial" Thought: An Operational Architectonics Model Guided Approach,

$https://pdfs.semanticscholar.org/c8e1/22c0d819c914f66ec2685c17d3cb89a3c5a0.pdf$

, Date accessed: 25 - Nov - 2017.

[13] O. Holland, (2003), Machine consciousness. New York: Imprint Academic;

[14] Oxford dictionary,

$https://en.oxforddictionaries.com/definition/consciousness$

Date accessed: 25 - Nov - 2017.

[15] I. Aleksander, (1995), Artificial Neuroconsciousness: An Update,

$https://web.archive.org/web/19970302014628/http://www.ee.ic.ac.uk/research/neural/publications/iwann.html$

, Date Accessed: 25 - Nov - 2017.

[16] RAIR Lab, Moral Reasoning, Decision Making, and Moral Dilemmas with Cognitive Calculi,

$http://rair.cogsci.rpi.edu/projects/muri/$

, Date Accessed: 26 - Nov - 2017.

[17] G. G. Gallup Jr. (1970), Chimpanzees: Self-Recognition, Science, Volume 167, Issue 3914, pp. 86-87.

$http://science.sciencemag.org/content/167/3914/86$

, Date Accessed: 26 - Nov - 2017.

[18] RAIR Lab, (2015), Self Consciousness with NAO Bots,

$https://www.youtube.com/watch?v=MceJYhVD_xY$

Date Accessed: 26 - Nov - 2017.

[19] S. Bringsjord, N. Sundar G., (2013), Deontic Cognitive Event Calculus (Formal Specification),

$https://pdfs.semanticscholar.org/8043/adb9ef06b171813ad15ee7581218b796ecf0.pdf$

, Date Accessed: 26 - Nov - 2017.