On Gödel's incompleteness theorem(s), Artificial Intelligence/Life, and Human Mind

V. Christianto[1] & Florentin Smarandache[2]

1 – http://www.sciprint.org, email: victorchristianto@gmail.com

2 – Dept. Mathematics & Sciences, University of New Mexico, Gallup, NM, USA, email: smarand@unm.edu, or fsmarandache@gmail.com

Abstract

In the present paper we have discussed concerning Gödel's incompleteness theorem(s) and plausible implications to artificial intelligence/life and human mind. Perhaps we should agree with Sullins III, that the value of this finding is not to discourage certain types of research in AL, but rather to help move us in a direction where we can more clearly define the results of that research. Gödel's incompleteness theorems have their own limitations, but so do Artificial Life (AL)/AI systems. Based on our experiences so far, human mind has incredible abilities to interact with other part of human body including heart, which makes it so difficult to simulate in AI/AL. However, it remains an open question to predict whether the future of AI including robotics science can bring this gap closer or not. In this regard, fuzzy logic and its generalization –neutrosophic logic- offer a way to improve significantly AI/AL research.[15]

Introduction

In 1931 a German mathematician named Gödel published a paper which included a theorem which was to become known as his Incompleteness Theorem. This theorem stated that:

> "To every w-consistent recursive class k of formulae there
>
> correspond recursive class-signs r, such that neither v Gen r
>
> nor Neg (v Gen r) belongs to Flg(k) (where v is the free variable of r)" [9].

In more common mathematical terms, this means that "all consistent axiomatic formulations of number theory include undecidable propositions."[9]

Another perspective on Gödel's incompleteness theorem can be found using polynomial equations [10]. It can be shown that Gödel's analysis does not reveal any essential incompleteness in formal reasoning systems, nor any barrier to proving the consistency of such systems by ordinary mathematical means.[10] In the mean time, Beklemishev discusses the limits of applicability of Gödel's incompleteness theorems.[11]

Does Gödel's incompleteness theorem limit Artificial Intelligence?

In the 1950s and 1960s, researchers predicted that when human knowledge could be expressed using logic with mathematical notation, it would be possible to create a machine that reasons, or artificial intelligence. This turned out to be more difficult than expected because of the complexity of human reasoning.[12]

Nowadays, it is widely accepted that general purpose of artificial intelligence (AI) is to develop (1) conceptual models (2) formal rewriting processes of these models and (3) programming strategies and physical machines to reproduce as efficiently and thoroughly as possible the most authentic, cognitive, scientific and technical tasks of biological systems that we have labeled Intelligent [5, p.66].

According to Gelgi, Penrose claims that results of Gödel's theorem established that human understanding and insight cannot be reduced to any set of computational rules [1]. Gelgi goes on to say that:

> "Gödel's theorem states that in any sufficiently complex formal
>
> system there exists at least one statement that cannot be proven to
>
> be true or false. Penrose believes that this would limit
>
> the ability of any AI system in its reasoning. He argues that there
>
> will always be a statement that can be constructed which is unprovable
>
> by the AI system."[1]

The above question is very interesting to ponder, considering recent achievements in modern AI research. There are ongoing debates on this subject in many online forums, see for instance [5][6][7][8][9]. Here we give a summary of those articles and papers in simple words. Hopefully this effort will shed some light on this debatable subject. Those arguments basically stand either on the optimistic side (that Gödel's theorems do not limit AI), or on the pessimistic side (that Gödel's theorems limit AI).

Mechanism and reductionism in biology and implications to AI/AL

It is known that mechanistic or closely related reductionistic theories have been part of theoretical biology in one form or another at least since Descartes.[8] The various mechanistic and reductionistic theories are historically opposed to the much older and mostly debunked theories of vitalism (see Emmeche, 1991). These theories (the former more than the latter), along with formism, contextualism, organicism, and a number of other "isms" mark the major centers of thought in the modern theoretical biology debate (see Sattler, 1986).[8]

Such mechanistic and reductionistic view of the world were discussed by F. Capra in his book: The Turning Point [13].

According to Sullins III [8], AL (Artificial Life) falls curiously on many sides of these debates in the philosophy of biology. For instance AL uses the tools of complete mechanization, namely the computer, while at the same time it acknowledges the existence of emergent phenomena (Langton, 1987, p. 81). Neither mechanism nor reductionism is usually thought to be persuaded by arguments appealing to emergence. Facts like this should make our discussion interesting. It may turn out that AL is hopelessly contradictory on this point, or it may provide an escape route for AL if we find that Gödel's incompleteness theorems do pose a theoretical road block to the mechanistic-reductionistic theories in biology.

Sullins III also writes that most theorists have outgrown the idea that life can be explained wholly in terms of classical mechanics.[8] Instead, what is usually meant is the following (paraphrased from Sattler, 1986):

1) Living systems can and/or should be viewed as physico- chemical systems.

2) Living systems can and/or should be viewed as machines. (This kind of mechanism is

also known as the machine theory of life.)

3) Living systems can be formally described. There are natural laws which fully describe

living systems.

According to Sullins III[8], reductionism is related to mechanism in biology in that mechanists wish to reduce living systems to a mechanical description. Reductionism is also the name of a more general world view or scientific strategy. In this world view we explain phenomena around us by reducing them to their most basic and simple parts. Once we have an understanding of the components, it is then thought that we have an understanding of the whole. There are many types of reductionist strategies.[8]

According to Sullins III [8], reductionism is a tool or strategy for solving complex problems. There does not seem to be any reason that one has to be a mechanist to use these tools. For instance one could imagine a causal reductionistic vitalist who would believe that life is reducible to the elan vital or some other vital essence. And, conversely, one could imagine a mechanist who might believe that living systems can be described metaphorically as machines but that life was not reducible to being only a property of mechanics.

Sullins III [8] also asserts that the strong variety of AL does not believe that living systems should only be viewed as physico-chemical systems. AL is life-as-it-could-be, not life-as-we-know-it (Langton, 1989, p. 1), and this statement suggests that AL is not overly concerned with modeling only physico-chemical systems. Postulates 2 and 3 seem to hold, though, as strong AL theories clearly state that the machine, or formal, theory of life is valid and that simple laws underlie the complex, nonlinear behavior of living systems (Langton, 1989, p. 2).

Sullins III [8] goes on with his argument, saying that at least one of the basic qualities of our reality will always be missing from any conceivable artificial reality, namely, a complete formal system of mathematics. This argument tends to make more sense when applied to strong AI claims about intelligent systems understanding concepts (see Tieszen, 1994, for a more complete argument as it concerns AI). He also concludes that it is impossible to completely formalize an artificial reality that is equal to the one we experience, so AL systems entirely resident in a computer must remain, for anyone persuaded by the mathematical realism posited by Gödel, a science which can only be capable of potentially creating extremely robust simulations of living systems but never one that can become a complete instantiation of a living system.[8]

However, Sullins III [8] also writes that the value of this finding is not to discourage certain types of research in AL, but rather to help move us in a direction where we can more clearly define the results of that research. In fact, since one of the above arguments rests on the assumption that the universe is infinite and that some form of mathematical realism is true, if we are someday able to complete the goal advanced in strong AL it would seem to cast doubt on the validity of the assumptions made above.

For a recent debate on this issue in the context of fuzzy logic, see for instance Yalciner et al. [5]. The debates on the possibility of thinking machines, or the limitations of AI research, have never stopped. According to Yalciner et al. (2010), these debates on AI have been focused on three claims:

- An AI system is in principle an axiomatic system.

- The problem solving process of an AI system is equivalent to a Turing machine.

- An AI system is formal, and only gets meaning according to model theoretic semantic (Wang 2006).[16]

More than other new sciences, AI and philosophy have things to say to one to another: any attempt to create and understand minds must be of philosophical interest.[5]

May be we will never manage to build real artificial intelligence. The problem could be too difficult for human brain over to solve (Bostrom, 2003).

Yalciner et al. [5] also write that a fundamental problem in artificial intelligence is that nobody really knows what intelligence is. The problem is especially acute when we need to consider artificial systems which are significantly different to humans.

Human mind is beyond machine capabilities

According to Gelgi [1], it follows that no machine can be a complete or adequate model of the mind, that minds are essentially different from machines. This does not mean that a machine cannot simulate any piece of mind; it only says that there is no machine that can simulate every piece of mind. Lucas says that there may be deeper objections. Gödel’s theorem applies to deductive systems, and human beings are not confined to making only deductive inferences. Gödel's theorem applies only to consistent systems, and one may have doubts about how far it is permissible to assume that human beings are consistent. [1]

Therefore it appears that there are some characteristics of human mind which go beyond machine capabilities. For example there are human capabilities as follows:

a. to synchronize with heart, i.e. to love and to comprehend love;

b. to fear God and to acknowledge God: “The fear of the LORD is the beginning of knowledge” (Proverbs 1:7)

c. to admit own mistakes and sins

d. to repent and to do repentance

e. to consider things from ethical perspectives.

All of the above capabilities are beyond the scope of present day AI machines, i.e. it seems that there is far distance between human mind capabilities and machine capabilities. However, we can predict that there will be much progress by AI research. For instance, by improving AI-based chess programs, one could see how far the machine can go.

Furthermore there are other philosophical arguments concerning the distinction between human mind and machine intelligence. Dreyfus contends that it is impossible to create intelligent computer programs analogous to the human brain because the workings of human intelligence are entirely different from that of computing machines. For Dreyfus, the human mind functions intuitively and not formally. Dreyfus‘s critique on AI proceeds from his critique on rationalist epistemological assumptions about human intelligence. Dreyfus‘s major attack targets the rationalist conception that human understanding or intelligence can be “formalized”.[5, p.67]

We agree with the content related to the distinctions between Human and Computer. Yet, we think that the differences (Love, God, Own mistakes, Repentance, Ethical) between Human and Computer will be in the future little by little diminished, since it would be possible to train a computer at least for partial adjustments in each of them.

In addition to the fuzzy logic in AI, neutrosophic logic provides besides truth and falsehood a third component, called indeterminacy that can be used in AI, since many approaches of reality that AI has to model or describe involve a degree of uncertainty, unknown. Neutrosophic logic is a generalization of intuitionistic fuzzy logic.[15] We have a lot of unknown and paradoxist, contradictory information that AI has to deal with in our world.

The above argument can be seen as stronger than Penrose's.

However, one should admit the differences between human intelligence and machine intelligence. There are fundamental differences between the human intelligence and today‘s machine intelligence. Human intelligence is very good in identifying patterns and subjective matters. However, it is usually not very good in handling large amounts of data and doing massive computations. Nor can it process and solve complex problems with large number of constraints. This is especially true when real time processing of data and information is required. For these types of issues, machine intelligence is an excellent substitute.[5]

Concluding remarks

In the present paper we have discussed concerning Gödel’s incompleteness theorem(s) and plausible implications to artificial intelligence/life and human mind.

Perhaps we should agree with Sullins III, that the value of this finding is not to discourage certain types of research in AL, but rather to help move us in a direction where we can more clearly define the results of that research. Gödel’s incompleteness theorems have their own limitations, but so do Artificial Life (AL)/AI systems. Based on our experiences so far, human mind has incredible abilities to interact with other part of human body including heart, which makes it so difficult to simulate in AI/AL. However, it remains an open question to predict whether the future of AI including robotics science can bring this gap closer or not. In this regard, fuzzy logic and its generalization –neutrosophic logic- offer a way to improve significantly AI/AL research. [15]

1st version: march 8, 2013. 2nd version: march 18, 2013. 3rd version: march 30, 2013

References:

[1] Gelgi, F. (2004) Implications of Gödel's Incompleteness Theorem on A.I. vs. Mind, *NeuroQuantology*, Issue 3, 186-189, URL: http://www.neuroquantology.com/index.php/journal/article/download

[2] Chalmers, DJ. ( ) *Minds, Machines, and Mathematics*. URL: http://psyche.cs.monash.edu.au/v2/psyche-2-09-chalmers.html

[3] Gödel, K. (1931) On Formally Undecidable Propositions of Principia Mathematica and related systems I. *Monatshefte fur Mathematik und Physik*, vol. 38 (1931), pp. 173-198. URL: http://www.ddc.net/ygg/etext/godel/ or http://courses.arch.ntua.gr/fsr/113698/Goedel--OnFormally.pdf or http://www.research.ibm.com/people/h/hirzel/papers/canon00-goedel.pdf

[4] Pelletier, F.J. (2000) Metamathematics of fuzzy logic, *The Bulletin of Symbolic Logic*, Vol. 6, No.3, 342-346, URL: http://www.sfu.ca/~jeffpell/papers/ReviewHajek.pdf

[5] Yalciner, A.Y., Denizhan, B., Taskin, H. (2010) From deterministic world view to uncertainty and fuzzy logic: a critique of artificial intelligence and classical logic. TJFS: *Turkish Journal of Fuzzy Systems* Vol.1, No.1, pp. 55-79.

[6] Karavasileiadis, C. & O'Bryan, S., (2008) Philosophy of Logic and Artificial Intelligence, http://rudar.ruc.dk/bitstream/1800/3868/1/Group%206-Philosophy%20of%20Logic%20and%20Artificial%20Intelligence-Final%20Hand%20In%285.1.2009%29.pdf

[7] Collins, J.C. (2001) On the compatibility between physics and intelligent organisms, arXiv:physics/0102024

[8] Sullins III, J.P. (1997) Gödel's Incompleteness Theorems and Artificial Life - Digital Library & archives, Society for Philosophy and Technology, URL: http://scholar.lib.vt.edu/ejournals/SPT/v2n3n4/sullins.html

[9] Makey, J. (1995) Gödel's Incompleteness Theorem is Not an Obstacle to Artificial Intelligence, URL: http://www.sdsc.edu/~jeff/Godel_vs_AI.html

[10] Norman, J.W. (2011) Resolving Gödel's incompleteness myth: Polynomial Equations and Dynamical Systems for Algebraic Logic, arXiv:1112.2141 [math.GM]

[11] Beklemishev, L.D. ( ) Gödel incompleteness theorems and the limits of their applicability. I, Russian Math. Surveys.

[12] URL: http://www.wikipedia.org/Logic

[13] Capra, F. (1982) *The Turning Point*. Bantam Books.

[14] Chaitin, G.J. (1999) A century of controversy over the foundations of mathematics, arXiv: chao-dyn/9909001

[15] Schumann, A. & Smarandache, F. (2007) *Neutrality and many-valued logics*. American Research Press. 121 p.

[16] Wang, P. (2006) Three Fundamental Misconceptions of Artificial Intelligence, URL: http://www.cis.temple.edu/~pwang/Publication/AI_Misconceptions.pdf

[17] URL: http://www.wikipedia.org/Philosophy_of_artificial_intelligence

[18] Straccia, U. (2000) On the relationship between fuzzy logic and four-valued relevance logic, arXiv:cs/0010037

[19] Born, R.P. (2004) Epistemological Investigations into the Foundations of Artificial Intelligence, URL: http://www.iwp.jku.at/born/mpwfst/04/0401Turing_engl_1p.pdf

[20] Chrisley, R. (2005) Simulation and Computability: Why Penrose fails to prove the impossibility of Artificial Intelligence (and why we should care), URL: http://www.idt.mdh.se/ECAP-2005/articles/COGNITION/RonChrisley/RonChrisley.pdf