Reflections on Artificial Intelligence – A Hermeneutic Journey

Saradhi Motamarri
UNSW Business School
UNSW Australia, Sydney, NSW 2052
saradhi.motamarri@unsw.edu.au

Abstract

Science, engineering and technology have been moulding and placing an ever increasing pressure on society and in turn, on life styles. The inquisitive nature of man has led to the amazing development of computer. In just four decades the computer has changed its role from a mere data cruncher to decision aid. A reading of the Artificial Intelligence, 17 (1-3), January, 1991, Special Volume on Foundations of AI, has motivated me to transcribe some of my long persistent feelings in writing. Loose usage and blown up speculations may bring discredit to concepts. Not founded on the characteristic behaviour of computer and its numerical instability, I fear that the same thing has happened to computers and to the associated fields of study. The purpose of this autoethnographic article is to reflect on the lessons learnt from AI and search for a right perspective for research and practice.

Keywords

Artificial Intelligence, Autoethnography, Computer, Decision Aid, Expert Systems, Information Systems, Numerical Instability, Reification, Science, Technology, Turing Test

INTRODUCTION

The inquisitive nature of man is attributable to the amazing development of science and technology. Of all, the invention of the computer is a symbol of man's ingenious intelligence. It can be hypothesised that from ancient times to date, man has been struggling to apply his acquired knowledge to reduce his labour. Accordingly, in just four decades, from the role of data crunching in its earlier generation, the computer is likely to assume its master's profession (what people say)! While people are still debating on the basic issues of computerisation, huge resources were expended to mimic the human brain. The *AI* v17, #1-3, Jan, 1991, *Special Volume on Foundation of AI*, brings forth the genuine criticism that the AI research has neither met its goals nor delivered on its promises. The discussions have motivated me to write my long persistent ideas as an autoethnographic account.

There is no doubt that the AI field has progressed in the last two decades since the 1990s, due to the multi-fold progress in both hardware and software technologies. For example, the Economist (2014) has run a special issue on 'Robots: Immigrants From The Future', recently. The Economist assesses that due to lowering costs, safety record and for the opportunity to be deployed in harsh environments, the worldwide market for robots showing resurgence. Also Professor Kevin Warwick of the University of Reading, announced in the media the successful passage of Alan Turing's test held at the Royal Society in London (Fingas 2014).

Because of lax usage, blown up speculation concepts may lose credibility. I fear that the same thing has happened to computers and to the associated fields of study. It is unfortunate that scholars were divided into opposing groups. It suggests that there is an obvious lack of a *common framework or datum* for the arguments. Quite often, we forget the premises and foundations on which our theories are erected. For example, the innate aspect of the computer's functioning takes less space than the philosophical discussions about the engineered artificats that came into being because of the advent of the computer itself. *I wonder whether the man who invented the computer to improve the functional effectiveness of man, has ever thought that another man has to operate it!* If we understand *the real power of the computer, its limitations and level of applicability*, then we recognise that the computer has a critical role in solving our day-to-day as well as complex scientific problems (Motamarri 1992a).

While it is a challenging task and beyond the scope of this paper to provide even a cursory glimpse of the AI, I will briefly provide categorisation of the AI field, based on the popular book of Russell and Norvig, 'Artificial Intelligence – A Modern Approach' (2009). I digress from scholarly debates of AI's feasibility or possibility, and discuss the preliminaries. The article presents an extremely fundamental view. Although the points are not radically novel, a logical acceptance of the same will unfurl the concepts that surround computers and its associated fields.

There are valuable lessons from this retrospective on AI to Information Systems, i.e., both relegated practice to value only philosophical discussions and got alienated from the common man. This is also one of the motivations for writing this paper and convey the need for an introspection within the IS discipline. In the following sections, touching the research method auto-ethnography, I contrast the definitions and principle assumptions of AI with the nature of human problem solving and the behaviour of computers. Making a hermeneutic journey into the past, I interpret the discussion and illustrate the application of natural and social science theories and conclude with my reflections.

RESEARCH METHOD – AUTOETHNOGRAPHY/ HERMENEUTIC JOURNEY

This paper presents a personal retrospective experience in learning about an interesting and passionate field, AI and associated areas of advanced research. Being a self-reflection it explores my personal experiences in traversing back into the history, the work falls under autoethnography. Marechal (2010 p.43) defines autoethnography as "a form of research that involves self-observation and reflexive investigation in the context of ethnographic field work and writing." Carolyn Ellis (2010 p.1) defines it as "an approach to research, writing, that seeks to describe and systematically analyse (graphy) personal experience (auto) in order to understand cultural experience (ethno). This approach challenges canonical ways of doing research and representing others and treats research as a political, socially-conscious act. A researcher uses tenets of autobiography and ethnography to do and write autoethnography. Thus, as a method, autoethnography is both process and product." While there are criticisms on the method, quoting the works of Stacy Holman Jones, Ellis (2010 p.9) concludes that "autoethnographers view research and writing as socially-just acts; rather than a preoccupation with accuracy, the goal is to produce analytical, accessible texts that change us and the world we live in for the better." Considering my view that the principal focus of the research shall and should be for the common good embracing a holistic perspective rather than confining the human quest within rigid frameworks of philosophical riddles, autoethnography is an appropriate method for documenting my reflections about not only AI but also to draw a humble request to the IS scholars to harvest the enormous wealth that lies in practice to advance the discipline.

In 1985, I graduated in Transportation Systems Engineering with an emphasis on Systems Engineering, Operations Research, Modelling, Simulation, and Forecasting and was influenced by the systematic methods of M.A. Jackson, for analysing and designing of programs and systems (Jackson 1975; Jackson 1983). By turn of events, I entered into commercial systems development projects and had the opportunity to lead transformative programs in many of my clients' organisations by applying the assimilated modelling principles acquired in my graduate program (Motamarri 1992b; Motamarri 1993). Subsequently, I moved into an applied Research and Development institute focusing on Artificial Intelligence, Expert Systems and Operations Management. I have developed a generalised model for Scheduling problems to assist the institute's myriad projects dealing with various scheduling applications in a variety of domains.

I was baffled by the terminology of the AI field. It has taken an enormous mental effort to learn, interpret and understand the terminology and gain a grasp of the field. Reading books, technical papers, magazine articles and discussions with peers have progressively helped me to unravel the mystic constructions, philosophical debates and what actually the scholars were attempting to create. This progressive hermeneutic interpretation left me in perplexed state with the grand question of how AI Systems development is significantly different from enterprise Information Systems where also users'/ experts' intricate business knowledge is encoded as business rules. While I myself architected some such sophisticated applications, it made me realise that there must be thousands of such ISs being created by my forerunners and peers. It was a hard realisation that sophisticated ISs could be built even in commercial infrastructure tools like: COBOL, Flat, Indexed-Sequential or Random files. It implies that there has been a massive divide between successful applications (enterprise IS) on one side and struggling researchers on the other side crafting AI/ ES in fancy languages (for example: LISP, Prolog) and other development environments.

UNDERSTANDING AI - THE HERMENEUTIC JOURNEY

Understanding a novel and complex thing takes time, effort and most of all, requires patience. While AI has not only novelty and fascination, it is also deeply mystified with the individual researchers' own creative dose of imagination. Over enthusiasm and lack of appreciation for enterprise IS development complicated matters even more. The newspeak (in George Orwellian terms) of AI will be the greatest hurdle for an outsider to get into the AI field and grasp what all these meant in practical terms. The hermeneutic cyclic process of gaining a foothold into the AI field is depicted in Figure 1. This hermeneutic cycle is in itself constructed by traversing into the layers of my past memory. I have undergone many cycles of interpretation by reading books, technical papers, magazines, dialogues with peers, by building simple systems, synthesising facts and at the end reaching a realisation, aha! this is what they mean, and again embarking on the next phase. I will present the reflections of

this interesting journey in the following sections, as they carry an important message not only to the fields of AI/ES/IS but also for the progress of science and technology.

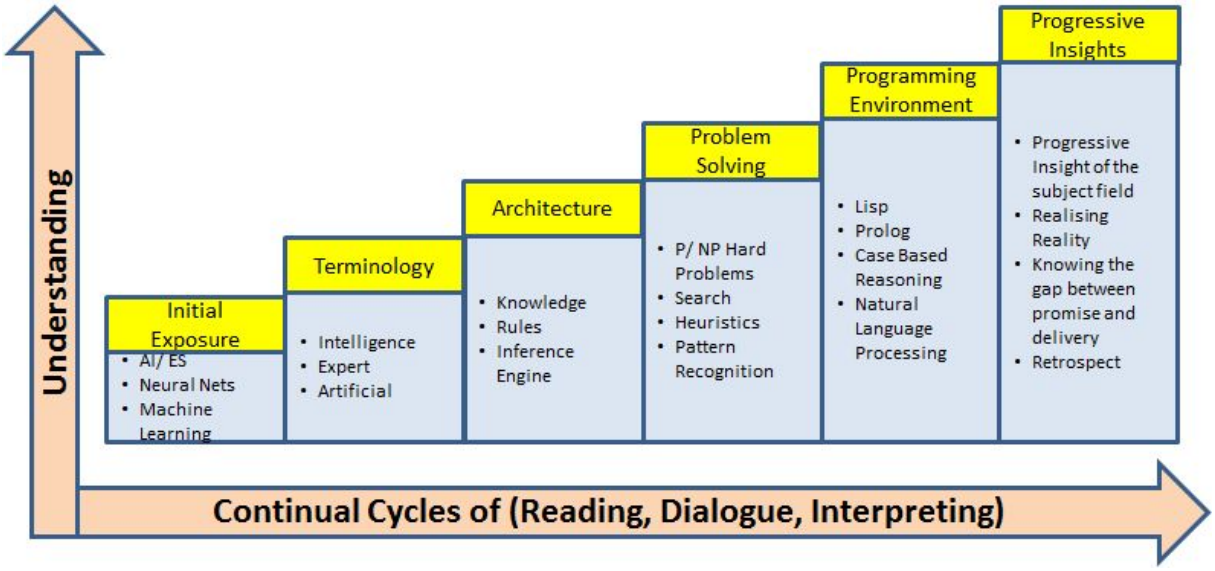

Figure 1: The Hermeneutic Progression in Understanding AI Discipline

THE PROMISE OF AI, THE AFTERMATH

AI is a broad discipline and has been pursued by different scholars in different ways. It is interesting to understand these themes, but it goes contra to the objective of this paper which is to reflect on AI approaches in general and draw some lessons from this introspection. In order to bridge this dilemma, I first present the categorisation of AI approaches from the book of (Russell and Norvig 2009) as shown in Figure 2, then move on to my introspection.

	Human Performance	Rationality
Thought Processes and Reasoning	**Thinking Humanly** • "The exciting new effort to make computers think… *machines with minds*, in the full and literal sense." (Haugeland, 1985) • "[The automation of] activities that we associate with human thinking , activities such as decision-making, problem solving, learning…" (Bellman, 1978)	**Thinking Rationally** • "The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985) • "The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)
Behaviour	**Acting Humanly** • "The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990) • The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	**Acting Rationally** • "Computational Intelligence is the study of the design of intelligent agents." (Poole et al., 1998) • "AI … is concerned with intelligent behaviour in artifacts." (Nilsson, 1998)

Figure 2: AI Categorisation (based on Figure 1.1 of (Russell and Norvig 2009 p.2) refrences as per original)

Russell and Norvig summarise that historically, all four approaches to AI, as shown in Figure 2 have been followed by different scholars with different methods. The top approaches to AI are concerned with human thought processes and reasoning while the bottom approaches are centred on human behaviour. The left approaches being human centred are in part followed empirical science, involving observations and hypotheses about human behaviour. The rationalist approach involves a combination of mathematics and engineering. Russell and Norvig also point out that these groups have both disparaged and helped each other.

Artificial Intelligence (AI)

AI is the *intelligence of machines* and the branch of Computer Science that aims to create it. AI takes a realist stand (Chalmers 2013) as the models are assumed to be explaining the intricate functioning of the human brain.

Expert Systems (ES) a segment of AI, takes a positivist approach to model the functioning of human experts for predictive purposes. Philosophically positivism holds that experimental investigation and observation are the only source of substantial knowledge (Chalmers 2013). AI also presumes a naturalist stand, as it studies the general human capacities as science. The naturalist philosophy holds that all phenomena can be explained in terms of natural causes and laws (Chalmers 2013). AI is defined as: "the study and design of *intelligent agents*" where an *intelligent agent is a system that perceives its environment and takes actions that maximises its chances of success*. Herbert Simon, one of the AI founders was profoundly optimistic and predicted in 1956 that "*machines will be capable within 20 years of doing any work a man can do*" and Marvin Minsky agreed that "*within a generation... the problem of creating AI will substantially be solved*." Thus, AI mistakenly attributes human like sensory and problem solving skills to inanimate objects, a kind of **anthropomorphism (Silverman 1968)**.

The invention of the computer helped man in tackling complex issues brought forth by the Industrial Revolution. It has become easier to use the computer for various tasks. At this juncture it is theorised that human beings do think and work like computers! This is an inductive generalisation derives from the fact that computers perform mathematical/ algorithmic calculations. From ordinary computations, it is induced that the computer is performing like the human mind. Surprisingly, AI deduced that the brain functions like a computer, a kind of **reverse anthropomorphism**, and imposed mechanist properties on the human mind. After huge investment of both human faculties and money, it was concluded that ordinary computational schemes are inadequate and novel techniques are needed to address the complexities to model the functioning of the brain. Along the way, the AI community **mystified** by creating its own jargon, terminology and alienated itself from systems engineers and ordinary people. Each generation of AI scholars, on seeing the failures of prevailing approaches, proposed a different paradigm as a response to the problem of creating AI, for example: expert systems, machine learning, case based reasoning, natural language processing and neural networks

Models and Reality

Eugene Charmiak and Drew McDermott define the fundamental assumptions of AI as:

1. "What the brain does may be thought of at some level as a kind of computation"; and
2. "Knowledge is essential for decision making. It is explicit and representable."

As discussed in the previous paragraph, the frist principle is an inappropriate deduction, while the second principle, the necessity of **knowledge** to problem solving is understandable. The creation of some simple models/ trivial prototypes that mimic human thinking and solve simple tasks led the researchers to blown-up speculation on the role of the computers/ AI and people were even made to fear about robots. The positivist approaches inappropriately generalised the human mind (as a machine) and researchers mistook their models for reality. Reflecting on this hype, fear and misconceptions, Sydney Harris quipped that "*The real danger is not that computers will begin to think like men, but that men will begin to think like computers*."

Mind versus Computer

In his famous paper '*Computing Machinery and Intelligence*,' Alan Turing (1950) suggested that instead of asking whether machines can think, we should ask whether machines can pass a behavioural intelligence test, which later came to be known as the famous ***Turing Test***. The test consists of an interrogator querying a computer program via online typed messages. Based on the conversation the interrogator needs to guess whether his conversation was with a human or a computer program. If the the interrogator was fooled for 30% of the time, then the program passes the test. Turing imagined that by the year 2000, a computer with a storage of 10^9 units could be capable enough to pass the test (Russell and Norvig 2009). Recently, Professor Kevin Warwick of University of Reading, claimed that they had succeeded in achieving the 30% mark, however, some sceptics express their reservations for the claim (Fingas 2014).

Consider a simple mathematical expression: $\frac{10}{3} \times 3$. Typically, we cancel out 3s and conclude the answer as 10. This is not computation but *simplification through visualisation*. However, no known computer will give us the exact answer, i.e., 10 (excluding the case of explicitly specifying the result of type integer or the result to be rounded off to the nearest integer). This **numerical instability** has profound impact on our theories and beliefs. Probably, this is a far more elegant falsification test for the claims of AI, than Alan Turing's procedure.

In the context of decision making and on the importance of rules, Charles Gow says that: "*Judgement is not the knowledge of fundamental laws. It is knowing how to apply a knowledge of them.*" While knowledge is necessary for decision making, it is not sufficient. Thus, the AI scholars discarded common sense that human beings do not always compute even the computational problems and there are limitations to explicating **knowledge** and of all, it is never simple to describe/ codify tacit knowledge. It is a hard learned lesson that human problem solving

involves **tacit knowledge** and that it requires a kind of apprenticeship to transfer tacit knowledge from the mentor to mentee (cf. Chalmers 2013 p.7).

We communicate in a variety of modes. Language is one of the greatest inventions of man to express, share and preserve his feelings and ideas. However, in order to communicate with a machine, one needs to instruct it in a highly structured manner, referred to as *programming*. The high level programs are translated eventually to binary form (0s|1s) as the digital computer can only understand 0s or 1s (Figures 3 and 4).

Our medium for communication, the language, itself is a model of reality. At times, we feel the inadequacy of these symbols and constructs, despite their thousands of years of existence. When the language itself lacks precision, in what way a programming language, a derivative, serve as an unambiguous medium for communication whether it is between man-to-man, man-to-machine or even machine-to-machine? Here again, the AI scholars have not realised that their models are only approximations to **reality**.

Figure 3: The Myriad Communication Modes

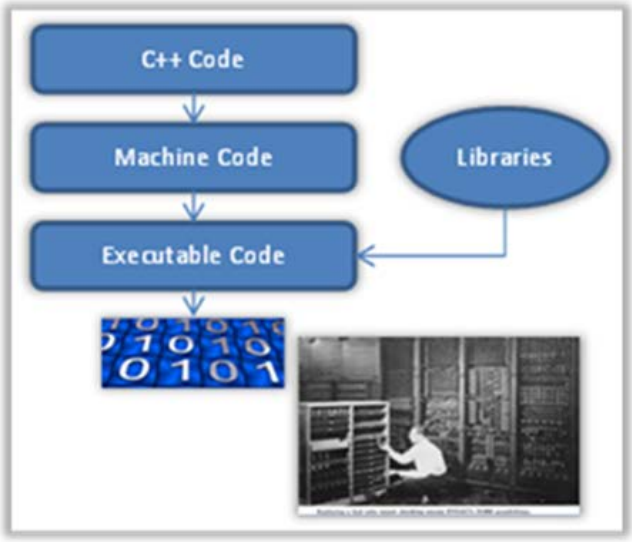

Figure 4: Man-Machine Communication

The Nature of Computer Behaviour

From a systems perspective, the raw power of the computer comes from the following facts (Motamarri 1992a):

- Unimaginable computational speed;
- Accuracy to perform routine and mundane computations;
- Ability to store/ retrieve voluminous data at fantastic speed; and
- Flexibility to be programmed.

Inductively we can say that through the last characteristic the rest of the features can be visualised. It means that the leverage of the computer comes from the fact of stored programs, otherwise it is no different from a hand calculator or at best a programmable calculator. It can also be deduced that to extract the power of the computer one needs to instruct it in a logical and disciplined way. The computer can neither accept a superficial statement nor an ambiguous case of a problem. *The computer is no more powerful than its user*. For a typist it is a word processor and for the other it is a game player. If one is smart, he can visualise his own expertise with the computer system. So, the inference is that the utility/ smartness of the computer is subjective (Motamarri 1992a). The AI community dwelling in an alienated world, ignoring common sense and systems perspectives, went on with grand speculations without making a reality check. This has not only brought discredit, but also led to the demise of the AI discipline.

If we have agreed that the power of the computer comes from the above facts and that it cannot operate without a stored program, and that program or software must be provisioned by man, then where is the real scope for autonomous and endogenous creation of intelligence? Furthermore, we have also noted that the computer computations are prone to numerical instability which is not even considered in the philosophical debates of AI or even IS. This surprises me as to why these foundational premises of computational intelligence have not been agreed upon as a datum in the philosophical debates!

DISCUSSION

The computer is a machine as well as an intelligent system, and it has the potential to assist man in solving critical problems. But the fact is that, *if a problem cannot be solved by man, then a machine can never solve it!* Despite AI's failure, it did contribute to a wide body of knowledge and a few practical tools have emerged, like: knowledge representation, Neural Networks (for classification problems), pattern recognition, heuristics, robotics and many more.

One shall not be discouraged with the fate of AI, but the search for sophisticated technologies must continue without forgetting the roots or foundations of our models. Scientific progress happens when we all work together *without attributing divinely things to models, as done for ether*! Finally, research and learning shall reduce complexity and misunderstanding and shall strive for simplicity and mutual understanding.

What if AI Succeeds?

In concluding their thorough treatise on AI, Russell and Norvig (2009 p.1051), put forward a serious question to the AI researchers, "*What if you succeed?*" They observe that a moderate success of AI has produced some valuable applications such as speech recognition, surveillance systems, robots and search engines. A moderate to large scale success of AI, creation of human-level-intelligence and beyond, they foresee would change the lives of humankind. It alters our view of intelligence, consciousness, and the destiny of human race. For these valid reasons, Russell and Norvig argue that AI research cannot divorce its responsibilities from ethics.

Why should we be Concerned about AI?

This is an important as well as a difficult question, but it is one of my objectives as well in writing these reflections. In the broadest sense, IS encompasses AI. As noted before even some of the enterprise IS do contain levels of intelligence which the practitioners might not have realised or claimed. Like AI scholars, IS scholars too distanced themselves from the practice, riddled in theoretical debates. While AI is wondering in its own creation with mystic attribution to their artefacts, IS is deeply in search of a theory(ies) to influence practice when IS is alienated from practice. When we look at engineering, historically and many a time practice precedes theory, and as time progresses the discipline attempts to understand and interpret the success of practice and creates theoretical scaffolding to formalise the practice (Motamarri 1993). The theory thus formed, settles into practice as best practices and standards. The confrontation of unsolved issues triggers again the cycle of practice-to-theory and then theory-to-practice. The important question to ponder is why IS is lagging in achieving this?

The Limits

In view of the above arguments, what essentially, I want to drive home is the point that the computer on one hand is a machine, and on the other hand it is an intelligent system. It has potential to assist man in critical areas of his problem solving. If a typical problem cannot be solved by human beings, it can never solved by the machine (supposing that man is patient enough to do all the book-keeping, repetitive calculations etc., for even thousands of years). Interestingly, we must admit that the vice versa is not possible. It is said that animal species or perhaps, the whole of nature passes information or characteristic behaviour to descendent generations through *genes*. We transmit our own intelligence (mere analogy) to the machine though the form of *software* (let us not worry whether it is procedural form, symbolic form or even neuro-net form)! But the transmissions are themselves models of reality that are bound to have some limitations in terms of accuracy, precision, of all, replication itself.

The Path

I strongly believe that man must continue his quest in knowing the unknown, must continue to toil to make lives simpler, easier and happier. But we must be cautious and open-minded. I further believe that the computer system has much to achieve in the years to come. The search for sophisticated computer technologies must continue with firm belief and confidence for fruitful benefits. At the same time, we should not forget the roots or foundations of our tools. Definitely, we require judgment to proceed in our perpetual innovative thirst to understand nature. As noted before, it's not mere knowledge of fundamental principles or laws, but wisdom that's more important.

REFLECTIONS

In conclusion, I earnestly believe that progress is not merely the result of one person or a group of persons. We must all come together, work together without attributing divine things to our theories, as done for *ether*. Whether it is AI or in general IS, there lies a common need for both theory and practice to work together, influencing each other for the emergence of newer and enriching understanding of nature and towards improving the quality of life of the society as a whole. While theorising and theory building are essential for the progress of humankind, we

shall not be dogmatic about our individual theoretical containers. Even while looking at an individual entity, our focus shall not ignore the bigger system out there (Motamarri 1992c). Theoretical debates existed thousands of years before the Aristotlean era, like the Eastern philosophies which are holistic and recognises the inter-connectivity beyond human perceptions as illustrated by famous Physicists like, Fritjof Capra and Gary Zukav (Capra 2010; Zukav 2009). I close this article with the wise words of Lancelot Hogben (1937):

> "*The truth is that fruitful progress only occurs when a large number of people are thinking together about the same sort of thing. There would be less reason for surprise if we realised that the clever people and people not so clever have a common need for each other, and that we cannot plan a future of higher intellectual progress for the human race if we place the needs of exceptionally gifted people in a false antithesis to the common needs of mankind*."

In a world where continual research and learning produces complexity and misunderstanding, I appeal to the stalwarts of advanced science for simplicity and mutual understanding.

REFERENCES

Capra, F. 2010. *The Tao of Physics: An Exploration of the Parallels between Modern Physics and Eastern Mysticism*, (5 ed.). Boston, Massachusetts 02115: Shambala.

Chalmers, A. 2013. *What Is This Thing Called Science? New and Extended*, (4 ed.). St Lucia, Queensland 4067, Australia: University of Queensland Press.

Ellis, C., Adams, T.E., and Bochner, A.P. 2010. *Autoethnography: An Overview*.

Fingas, J. 2014. "Super Computer' First to Pass Turing Test by Mimicking a Teenager." http://www.engadget.com/2014/06/08/supercomputer-passes-turing-test/: Engadget.

Hogben, L. 1937. *Mathematics for the Million*, (10 ed.). W.W. Norton & Company Inc.

Jackson, A.M. 1975. *Principles of Program Design*. Aacademic Press.

Jackson, A.M. 1983. *Systems Development*. Prentice-Hall.

Marechal, G. 2010. "Autoethnography," in *Encylopedia of Case Study Research*, A.J. Mills, G. Durepos and E. Wiebe (eds.). Thoousand Oads, CA: Sage Publications, pp. 43-45.

Motamarri, S. 1992a. "A Diagnostic View on Information Technology," *SIGSOFT Softw. Eng. Notes* (17:4).

Motamarri, S. 1992b. "Software Engineering from a Practical Perspective," *SIGSOFT Softw. Eng. Notes* (17:3), pp. 21-26.

Motamarri, S. 1992c. "Systems Modelling and Description," *SIGSOFT Softw. Eng. Notes* (17:2), pp. 57-63.

Motamarri, S. 1993. "From Practice to Theory: A Development Strategy Worth Looking At," *SIGSOFT Softw. Eng. Notes* (18:3), pp. 57-63.

Russell, S., and Norvig, P. 2009. *Artificial Intelligence: A Modern Approach*. Prentice Hall Press.

Silverman, D. 1968. "Formal Organizations or Industrial Sociology: Towards a Social Action Analysis of Organizations," *Sociology* (2), pp. 221-238.

The Economist. 2014. "Special Report: Robots - Immigrants from the Future," in: *The Economist, 410 (8880)*. London, UK: The Economist, pp. 1-16.

Turing, A.M. 1950. "Computing Machinery and Intelligence," *Mind* (49), pp. 433-460.

Zukav, G. 2009. *Dancing Wu Li Masters: An Overview of the New Physics*. New York, NY: HarperOne.

Additional Reading

Penrose, Roger (1989). "*The Emperor's New Mind: Concerning Computers, Minds And The Laws Of Phsics.*" Oxford University Press.

Artificial Intelligence (1991). *Special Volume on Foundations of AI*. 17 (1-3).

ACKNOWLEDGEMENTS

The author is very grateful to the anonymous editors and program chairs for their patient review and invaluable advice that helped in enhancing the paper. The author also would like to thank Professor Lex Donaldson for his comments on a previous version of this paper and Professor Patrick Finnegan, Head of School of Information Systems, Technology and Management, UNSW Business School, for his support and encouragement.

INTELLECTUAL PROPERTY