Paper Title: Human Language and the Limits of Artificial Intelligence. A New Religion-Science Relation
Author: Trausan-Matu, Stefan
Institutional Affiliation: Computer Science Department, "Politehnica" University of Bucharest and Math Forum @ Drexel University

Abstract:
The paper discusses some achievements and failures of artificial intelligence (AI) from the perspective of the relation between (Christian Orthodox) religion and science. A special emphasis will be given to automated language understanding in AI, and, in contrast to dialogism, in Mikhail Bakhtin's conception, with direct link to the fundamental dimension of human being as it appears in Christian theology. The theological dimension of Bakhtin's dialogism has been discussed previously by other authors but its relation with AI was not analyzed elsewhere.

Biography:
Dr. Stefan Trausan-Matu, is a Fulbright Scholar at Drexel University, Philadelphia. He is Professor at the Computer Science Department of the "Politehnica" University of Bucharest, and senior researcher at the Romanian Academy Institute for Artificial Intelligence.

He authored 9 books, 7 book chapters and more than 120 book papers and reports, has participated in 10 international and more than 15 national research projects on artificial intelligence and e-learning. He has been invited professor in universities in US, France, UK, Netherlands etc., has organized conferences and workshops, is a committee program member and reviewer at international journals and conferences, had tutorials at national and international conferences. He is Vice-chair of the ACM Romanian Special Interest Group on Computer Human Interaction and a member of the Romanian Academy Committee for the History and Philosophy of Science and of the Commission of the Romanian Academy for Language Technology.

His special research interests are: human language understanding, dialogism, computer supported collaboration, knowledge-based systems, intelligent e-learning, philosophy, and the religion-science relation

Paper Text:
Some of the most common patterns of the last centuries are the supremacy of science and the need of a clear cut between science and religion. The supporters of such positions state that science does not need religion, that anything may be explained solely by science, without considering God.

One of the achievements that would definitively prove the above ideas would be the creation of artificial life. This myth appeared recurrently in history under different

shapes: the golem, Frankenstein and, in the last century it even entered into the most prestigious scientific circles under the form of the idea of building humanoid robots, or, more general, artificial intelligence or artificial life.

From a Christian perspective, however, such ideas are nothing else than the well known sin of trying to do what only God can, to be His equal. However, potential sins may eventually turn into exactly the opposite. In the case of artificial intelligence, the expected (from the religious perspective) failure of obtaining an artificial human-like thinking robot may give some extremely interesting insights about the limits of science and technology and about the relation religion-science. Moreover, science, in general, and AI, in particular, may beneficiate from fundamental theological ideas. This possibility is clear and even needed due to the problems of AI, for example, and suggested by the influential dialogical theory of M. Bakhtin (Bakhtin, 1981, 1986).

Artificial intelligence (AI) is a research domain whose results can contribute for studying the religion-science relation because its main objective may be seen as the (sinful, from a theological perspective) challenge of imitating God. The main goal of AI is to develop machines (computer-based) that could perform tasks that, if performed by a human, we would call that human as an intelligent person. This goal is some times extended to include all psychological dimensions of a human being: learning, emotion, intentions, even language understanding and consciousness. We may say that AI is a climax of positivism, because if we could develop such an intelligent machine, it could probably explain logically anything we ask it. From another perspective, such an achievement would be a proof that humans are no more than machines and religion is only a byproduct of mental, physical processes.

There are already AI computer programs that may recognize scenes, may memorize more than a human being and may develop much more complex reasoning chains. However, other human activities like machine learning, intelligent problem solving and human language processing in computers are only partially successful, at about half a century after AI research started. Moreover, other phenomena like consciousness, intuition, creation, art, are far from being obtained in AI.

One of the most important and, meanwhile, representative sub domains of AI, especially for our religion-science relation, is human (natural) language understanding. Even if there are computer programs for translation, summarization and other powerful language processing, phenomena like metaphors, poetry, jokes, and sometimes, even mundane dialogs, are usually hardly "understood" by AI programs. For example, I can say that you cannot trust to use, without a validation, automated translation systems that are available now. Metaphorical texts are extremely difficult to translate by computers and it is not clear whether it will be much easier in the future. Even if the future programs will be able to translate metaphors (by storing large lists of usual metaphors, for example), it is a great challenge to say whether they will really understand or recognize a new metaphor. The case of poetry is even no taken into account by researchers.

Human language was in the focus of AI research from the beginning of its history. One of the first researchers that introduced the artificial intelligence ideas, twenty five years ago, was Alan Turing (Turing, 1950), the mathematician who formally described the idea of computation and that conceived the so-called "Turing Machine", the abstract model of digital computers. This model and the formalism behind it probably helped Turing to the celebrated deciphering, during the Second World War, of the German secret codes generated by their Enigma machine. We could trace the genealogy of AI, and in particular of human language processing, for example machine translation, to the success of Enigma code deciphering, which, in fact, is also a form of translation, from the encrypted form, into the actual text. From this achievement is only a step to the hope that such machinery could translate from one language to another or even could answer questions, entering into a dialogue. This is exactly what Turing proposed: He said that it is possible to implement a computer program that exhibits artificial intelligence and even he proposed a test for its validation. The so-called "Turing Test" states that if somebody writes a question on a sheet of paper and gets the answer on another sheet of paper without seeing if there was a computer program or a human who answered and he cannot discern that it was a machine who answered, that program has artificial intelligence (Turing 1950).

Even if Turing said that he believes "that in about fifty years' time it will be possible, to program computers, … to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification after five minutes of questioning" (Turing 1950), and fifty five years have passed (and even an award of $100,000 and a solid gold medal are given 'for the first computer whose responses were indistinguishable from a human's', see http://www.loebner.net/Prizef/loebner-prize.html), the Turing Test has not yet been passed by any computer program and it is not expected to be passed soon. In the decades that passed, the initial enthusiasm has disappeared when unexpected complexities of human language were discovered, when it became clear that programs for language understanding must posses a huge amount of "commonsense knowledge" and, maybe the most difficult, that even the simplest dialogues imply context problems extremely complex.

There are many artificial intelligence scholars that passed from successful advancements in their domain to very critical positions and, many of them are now not anymore researching how to build intelligent machines but how to enhance human-computer interaction. I will give here only a single name, Terry Winograd, one of the most important researchers in human language processing, knowledge representation and artificial intelligence in the seventies and first eighties, now a professor of human-computer interaction at Stanford University.

Terry Winograd is coauthor to a very influential book (Winograd and Flores, 1986) where he discusses in detail the limitations of artificial intelligence. His ideas are very clearly presented also in a research report (Winograd, 1987), where he says that artificial intelligence cannot go beyond a beaurocracy level (that, I could say, is a person without empathy, that acts according to some strict, mechanic rules). For achieving language

understanding, a computer, that "as a language machine, manipulates symbols without respect to their interpretation" (Winograd, 1987), cannot reach humans, that "create their world through language … always interpreted in a tacitly understood background." (Winograd, 1987).

Winograd proposes a constructivist-hermeneutic approach (Winograd, 1987), in the tradition of Heidegger and Habermas, in which interpretation in context is crucial. I agree with his ideas, but I consider as least as sound the approach that takes into account also the philosophy of language based on dialogism, theory introduced by Mikhail Bakhtin in direct relation to Christian Orthodox theology that obviously influenced him (Mihailovic, 1997, Ugolnik, 1989, Clark and Holquist, 1984), even if he didn't explicitly mentioned this fact (we must not forget that he lived during communism and that, as so many intellectuals in Soviet Union, he was even arrested and deported).

In Trausan-Matu (2002) I remarked another problem that limits AI researchers, what I called the complexity syndrome. Computer programs could solve the needed problems for understanding language if an exponentially increasing number of computations (a so-called "combinatory explosion") would not be needed, making practically impossible the achievement of the desired goals. From a biblical perspective, we could view this phenomenon as an instance of the Babel's Tower myth, in which, when people tried to build something that could bring them near God, they could no more understand each other, their languages got different, they lost the unity needed in language communication.

I have also remarked (Trausan-Matu, 2002) that many hard problems of artificial intelligence (like understanding, intuition, consciousness, creation, free will) are linked one to another (what some researchers name the AI-complete problem) and that an ethical "filter" would also be needed for such machines. In fact, AI applications limitations could be very well understood if we accept that humans have, in addition to machines, as the Christian religion says, something that resembles to God.

We could say that the Turing Test verifies the possibility of entering into a dialogue and implicitly states that intelligence is strongly related to dialog. I emphasized these ideas because exactly dialog is, in my opinion, one of the key issues that are separating humans from machines. The concept of dialogue and the idea of a community of believers are characteristic features of Christianity. Exactly on these concepts, Mikhail Bakhtin built his dialogism theory, considering dialogue as a fundamental philosophical category. In fact, several scholars affirm that his theory is directly related to Christian basic concepts like enfleshment, eucharist and the One-Multiple attribute of Trinity (Mihailovic, 1997, Ugolnik, 1989, Clark and Holquist, 1984).

Mikhail Mikhailovici Bakhtin (1895-1975), a Russian philosopher, linguist and philologist said that everything written or spoken has a dialogic nature, even novels. This is the basis for his dialogical, post-structuralist theory which is situated on a position different from that usually used in AI. Bakhtin's ideas (Bakhtin, 1981, 1986) could be a solution to build a bridge between religious ideas and science.

Understanding language in Bakhtin's conception is very similar to the eucharistic phenomenon. As Mihailovic states, "the participants in Bakhtinian dialogue begin to emerge as communicants in an eucharistic sense as well in a strictly linguistic one." (Mihailovic, 1997). To understand means to participate to a community of voices, to collaborate, to share. The word at Bakhtin gets a corporeal, an enfleshed nature, as in John 1:1 and 1:14 : "In the beginning was the Word and the Word was with God and the Word was God… And the Word was made flesh, and dwelt among us" (Mihailovic, 1997).

I consider, in conclusion, however, that the best position is not to blame AI, but to try to blend the creative, artistic, even, if we consider religion, divine dimensions of human beings with the mechanized reasoning power of computer programs. AI computer programs, which, in fact, are the result of human creative potential, may be seen as extensions of human mind, to be used as tools that could perform activities that do not need creative abilities.

References

Bakhtin, M.M., The Dialogic Imagination: Four Essays, University of Texas Press, 1981

Bakhtin, M.M., Speech Genres and Other Late Essays, University of Texas Press, 1986

Mihailovic, A., Corporeal Words: Mihail Bakhtin's Theology of Discourse, Northwestern University Press, 1997

Trausan-Matu, St., Psihologia robotilor, in Constandache, G.G., Trausan-Matu, St., M.

Albu, C. Niculescu, Filosofie si stiinte cognitive, MatrixRom, 2002.

Turing, A. M. (1950). Computing machinery and intelligence, *Mind, 59*, 433-460.

Clark, K., Holquist, J.M., , Mikhail Bakhtin, Harvard Univ esrity Press, 1984

Ugolnik, A., The Illuminated Icon, Eerdmans, 1989.

Winograd, T., Flores, F., Understanding Computers and Cognition, Norwood, N.J.: Ablex, 1986.

Winograd, T., Thinking machines: Can there be? Are we?, Report No. STAN-CS-87-1161, Stanford, 1987.