From Artificial Neural Networks to Emotion Machines with Marvin Minsky

Jozef Kelemen

Institute of Computer Science, Silesian University, Opava, Czech Republic,
and VSM College of Management, Bratislava, Slovak Republic
kelemen@fpf.slu.cz or jkeleme@vsm.sk

Abstract: In 2007, one among the founders and internationally most recognized leading pioneers of the field of Artificial Intelligence and Cognitive Science, professor emeritus at Massachusetts Institute of Technology, Marvin Minsky celebrates his 80th anniversary. Exploiting thnis opportunity, the article overviews his contribution to the above mentioned fields (and might be also to some others), and sketches an (obviously incomplet) picture of the history of Artificial Intelligence and related disciplines, and a specific way of doing science, too.

Keywords: artificial intelligence, cognitive science, artificial neural network, theory of computation, emotion, frame representation of knowledge, society theory of mind, robot

1   Introduction on Two Anniversaries

In August 9, 2007, one among the founder and one among the internationally most recognized leading pioneer of (and professional in) the field of Artificial Intelligence and Cognitive Science, MIT professor emeritus Marvin Lee Minsky celebrated his 80th birthday. This is a good opportunity to overview his professional contribution to the above mentioned fields (and might be also some others), and to sketch a particular picture of the history of the disciplines he hbave contributed, esp. of the field of artificial neural networks, of Artificial Intelligence, and of Cognitive Science.

Another relevant reason to focus to Minsky's professional activities consists in the fact, that Artificial Intelligence (AI) celebrated in the year 2006 its 50th anniversary. That anniversary is closely related to Marvin Minsky's innitiatives a half century ago. AI is a discipline initialized thanks to a 1955 initiatives of a small group of enthusiastic young professionalsformed by Marvin Minsky, Claude Shannon, the founder of the modern quantitative theory of information, John McCarthy, the author of the famous programming language LISP, the *lingua franca* of advanced AI experimentations, and Nathalien Rochester.

Figure 1

Marvin Minsky at the Department of Cybernetics, Faculty of Electro-Engineering, Czech Technical University, Prague, June 2004 (photo by J. Kelemen)

The group wrote a letter to the *Rockefeller Foundation* in the summer 1955 – see (McCarthy et al., 2006) – applying for a support of a two month ten men summer study of artificial intelligence *…which will proceed on the basis of the conjecture that every aspect of learning or any other features of intelligence can in principle be so precisely described that a machine can be made to simulate it* (cited from the application letter).

The application resulted in the possibility to organize a two months research project (a summer school) at the Dartmouth College, Hanover, NH, in the period June 19-August 16, 1956.

During the summer school, the first, and up to that time isolated, research results and projects focused to study the possibilities of computers to simulate human level intelligence have been presented to a larger community of interested professionals, e.g. to (and by, as well) Herbert A. Simon, Allen Newell, John von Neumann, Oliver Selfridge, Marvin Minsky, John McCrathy and other important persons of the history of development of cybernetics, computer science, Artificial Intelligence, and Cognitive Science of the second half of 20th century.

Because of that, the 1956 *Dartmouth Summer School on Artificial Intelligence* is now recognized as the first professional meeting of the field of AI, and the starting-point of the development of AI as a well-constituted professional discipline, as a branch of scientific research which celebrated its golden anniversary past year.

2   Artificial Neural Networks

Minsky's interests in Artificial Intelligence started with study of (artificial) naural networks. As a PhD student in mathematics at Princeton University, and then as a fresh postdoc at Harvard, under the strong influence of works of Warren McCulloch and Walter A. Pitts concerning the cybernetic conceptual (mathematical) models of neuronal cells and the neural networks; see e.g. (McCulloch, Pitts, 1943), Minsky – in 1949 – started with an attempt to construct a working artificial neural network. He names the device *SNARC* (an abbreviation for *Stochastic Neuro-Analog Reinforcement Computer*). *SNARC* was a hardwired elecrotechnical device consisting of several identical artificial neurons connected in a network which was able to learn simple concepts (Minsky, 1952), (Minsky, 1954). It has been completed in 1951.

After the publication of the first monograph by Frank Rosenblatt (Rosenblatt, 1961) devoted to artificial neural networks, and under the rather unbounded Rosenblatt's optimism concerning the learning and recognition capabilities of the neural networks he proposed, Minsky and his collaborator from (and co-director of) the MIT Artificial Intelligence Laboratory, Seymour Papert, published the monograph (Minsky, Papert, 1969).

Minsky's and Papert's monograph reflected the authors' criticism of the Rosenblatt's concept of neural networks. It contains some mathematically proved theorems concerning some bounds of the learning and recognizing capacities of networks proposed by Rosenblatt. The criticism was based on the fact, that in certain cases it is impossible to learn or to recognize holistic concepts on the base of results of local observations. A simple illustration perhaps provides an intuitive picture of the core of troubles.
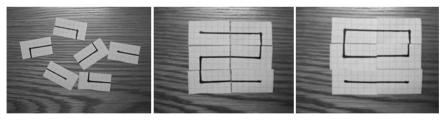

Figure 2

Two principally different 'global' drawings constructed from the same 'local' pieces; taked from (Kelemen, 2004)

Fig. 2 illustrates that on the base of the 'local observations' (represented by the isolated drawings in the picture on left), the drawings from the second and the third picture cannot be recognize. It is obvious, that the picture in left hand side contains the same 'local information' in both of cases of observing the parts of the second or of the third picture locally.

Minsky's and Papert's main contribution to better understanding of the recognition power of artificial neural networks consists in discovering of the *importance of the representation of whole concepts through local propositions*. This discovery then influenced Marvin Minsky's next scientific activities and professional views in many important ways.

Perhaps the first example of that influence is Minsky's relation and innovative views to computation and to abstract computating machines.

3   Finite and Infinite Machines

In 1967 Minsky published his influential and internationally recognized university text-book (Minsky, 1967).

The book starts with explaining finite-state machines as simple neural networks, so as devices constructed from large number of some small variety of simple computing units, as a distributed, multi-processor devices, in principle. Then Minsky continues to discuss the realism of the traditional Turing Machine (Turing, 1936) as a suitable formal model of computation and computers. He criticizes some of the more or less unrealistic features of the Turing Machines, and replaces the traditional thinking on computation as the function of the Turing Machine by another concept of the machinery, based on more realistic architecture.

The formal model of the Turing Machine memory in the form of a potentially infinite tape is in (Minsky, 1967) replaced, according his inspiration by the construction of the real hardwired (and because of that technical embodiment necessarily finite) computers working with partitioned memories, by a machine working with a finite number of so-called *registers*. The *Register Machine* is proposed as an alternative model of computing device, and it is proved its equivalent with the traditional Turing Machine with respect its computational power. In other words: What is computable by a Turing Machine is computable also by a Register Machine, and *vice versa*. The registers have infinite capacity each, but for performing the computation, only the question whether a register is empty or not is important! So, the Register Machine seems to be much more realistic, finite in certain sense, device, than the traditional Turing Machine which must deal with the infinity of its tape.

The Register Machine is a living concepts in the theoretical computer science up to now, and it makes many of proofs concerning the universality of the differently motivated computing devices in many cases much simpler; an example of such proof can be found e.g. in (Csuhaj-Varjú et al., 2006).

Another influence of the Minsky's conviction in *the power of well-mutable connections of simple parts* has been reflected in the architectural principle used

by Daniel Hillis, a post-doc at MIT AI Lab in that time, in the design of the famous *Connection Machine* (Hillis, 1985).

The first model CM 1 has been completed in 1986 as the collection of 65 536 simple one-bite processors organized in the form of a 12 dimensional hyper-cube with the memory bandwidth of 300 Gigabits per second. Different models of the Connection Machine have then been produced then by the Thinking Machines Corporation, Inc. founded by D. W. Hillis, M. Minsky, and others in 1986, in Cambridge, MA.

4   Knowledge Representation

The Minsky's conviction in the *importance of the representation of whole concepts through local proposition* is evidently present in his concept of the so-called *frame representation of knowledge*, a very influential knowledge representation concept in AI starting with its proposition by Minsky in June 1974 (in the form of the MIT AI Laboratory Memo No. 306). The first regularly published version of the proposal appeared then as (Minsky, 1975), and then many times in different wordings as a part of different edited books.

The core of the concept of *frame representation of knowledge* consists in integration of the positive sides of the three previously developed representational schemes.

The historically first, the *declarative* representational scheme, is based mainly on the idea of adaptation of traditional formalisms developed in the field of the formal logic for AI purposes. The representations are logical formulas, and the problem-solving procedures are based on automation of theorem-proving.

The second scheme, the *procedural* one, was inspired by the experiences in computer programming. It is focused to developing advanced programming environments enabling the end-users to solve problems by their declarative specification only, without the necessity of programming any procedural parts of the problem-solving procedures. Programming tools – like the Prolog programming languages family, for instance – are good examples of the influence of this representational ideology.

The inspiration for the third representational scheme, the *associative* one, originated in the circles of linguists converted to AI research because of their conviction that AI research is the best way to make computers able to understand natural languages and use them for communication with human beings and for mechanizing the translation of texts from one natural language into another one.

The problems concerning mechanization of translation leads immediately to the recognition of the importance of contexts in which human beings interprets natural

language texts in order to understand (and translate) them, so to the problems concerning the semantic meaning of the syntactic units forming the texts.

The traditional approaches to deal with meaning of formulas used in formal logic, as well as the philosophical conceptions of extensionality of meaning was recognized as not suitable enough, and the interest of linguists has been shifted to search for approaches which emphasize the intensionality. This attempt resulted in formalism of associative (often called also semantic) networks.

The renaissance of the approach is in present time recognizable in researches oriented to specification of web ontologies, for instance; see e.g. (Fensel, 2004).

Minsky in his approach to knowledge representation found a productive way of how to combine the positive sides of the previously developed representational schemes, and how to enlarge the resulting representational structure by some other practically required and psychologically reasonable properties, like representation of *expectations* of events and situation (as so called *default values* in some parts of the frame representational structures) and ways to implement the possibility of the so-called *non-monotone reasoning* recognized by psychologists in human problem-solving activities but not reflected in traditional formal logics up to the invention of frames (then the non-monotone reasoning became a part of research in formal logic, esp. thanks to initiatives of John McCarthy).

The influence of the idea of frames as representational structures to the practical programming is evident in the so-called object-oriented programming, and in object-oriented programming tools like C++ or Java, for instance, where the simplified concept of frames is present as the concept of objects; see e.g. (Coad, Nicola, 1993) for details concerning the object-oriented programming.

5   Architecture of Mind

Another consequence of Minsky's conviction in *the power of well-mutable connections of simple parts* is his concept of the mind as a society of interrelated and mutable society of simpler units called agents.

Agents in Minsky's understanding are organizational units with the ability to sense and act in their environments created by other agents, so be active in a multi-agent system as its parts, and create larger structures having the same behavioral properties as the agents have. Starting with the simplest agents having the power of simple computational devices Minsky proposes the way of the evolution of much more complicated societies of agents with the power comparable with the power of human minds.

This concept – the *society theory of mind* – is perhaps the most important Minsky's contribution to the development of cognitive science.

The first description of the structure of these agencies has been proposed in (Minsky, 1980), a paper focused to the problem of specific form of functioning of the human memory. Then human memory, in the contrast with the memories of artificial computing devices, enables reconstruct memorized events or objects incompletely, stressing only the contextually relevant (important) parts of the happened events or observed objects, without all of previously sensed but in the moment of re-construction of the passed events or sensed object contextually unimportant details.

The concept of mind as a society of interrelated developing system of agents is in the form of a monograph presented in (Minsky, 1986), and then it was massively propagated using different media and narratives. It has been explained in the form of an interactive multi-medial CD in (Minsky, 1996), In the context of a thrilling sci-fi story it was presented in the form of a techno-thriller in (Harrison, Minsky, 1992). In (Singh, 2003) the development and the influence of the theory is recapitulated with the complete professional rigor.

Roughly speaking, the importance of the society theory of mind is twofold. First, it consists in emphasizing the concept of agents in cognitive science as a useful and productive concept for explaining of the phenomenon of mind and some of the mind's capacities. Second, the concept of organized societies of agents became inspiring, broadly applicable, and also very trendy, during the 90ties of the past century. It happened in the field of practical applications of computing technology as one of the basic concept for dealing with computing machineries as specific types of multi-agent systems; see e.g. (D'Inverno, Luck, 2004).

5   Emotion Machines

Elaborating his society theory of mind Minsky discovered that there are no principal difference between the architecture of agents and agencies required by the rational side of intelligence, and that one required by the emotional one. The basic emotional states are results of activities of some simple agents sensing the chemical properties of an organism and evoking basic emotional states like hunger and thirst, or the concentrations of different hormones in the organism, and evoking emotional states of fear or wrath. Realizing this universality of the society theory of mind Minsky started at the beginning of 90^{ties} of the past century to work on his theory of emotions. The work resulted in the monograph (Minsky, 2006).

In (Minsky. 2006) emotions are treated in their abstract forms, so in the form, which does not emphasize and make no principal difference between the emotionality of human beings or machines. This view of emotionality is enabled thanks to the quit general view of the mind as a society of interrelated agents, and not as a specific phenomenological concept attributed exclusively to human beings.

The multi-agent (or resource instead of agent, as Minsky proposed instead of agents in his book) picture of the mind enables to formulate a specific approach to the functions of brains (human ones, but also artificial ones like computers, or advanced computer controlled robots). Minsky imagined brains as things which developed different ways to maintain parallel representations in multiple contexts, 'realms' as Minsky writes, so that each can apply its own resources to different aspects of the same word or topic. He names this multi-agent, or multi-resource, problem solving by a neologism *panalogy* (from *parallel analogy*).

For example, in the case described by the sentence 'John gave Mary a book' we recognize minimally three different interpretation (understanding) realms:

- the physical realm, where 'give' refers to book's motion through the physical space,

- the social realm, when 'give' refers to a specific social act, and

- the 'dominion realm', when if we heard that 'John gave Mary a book' we inferred that Mary then be holding the book and has gained permission to use it.

So, according Minsky's opinion, we have things well-represented if we have represented them in several realms at once, including the emotional realm(s). If we are watching a horror in a cinema for instance, we understood the movie not only in physical, social, common-sense etc. realms, but in certain moments necessarily also in realms of the surprise, fear, disgust, etc. The 'global' understanding of the movie emerges in certain sense from all these partial understandings, and this emerged understanding is the complete and psychically normal one.

7   Anothe Anniversary – Robots

As we have mentioned above, Marvin Minsky has contributed also to the development of the advanced cognitive robotics. Not only as the co-director of the MIT AI Laboratory, where one among the firsts projects of development of this kind of robots have been started during the end of 60ties of the past century, but also in the role of coordinator of research projects, and supervisor of many PhD projects oriented to different particular problems (e.g. in natural language understanding, manipulation with real objects, vision, hand-eye coordination, learning form experience, etc.) connected with the cognitive robots construction. Then, step-by-step he approached the cognition and emotions from a point of view from which the difference between the cognitive and emotional phenomena in biologically and culturally evolved human beings and in artificially engineered machines have no importance.

Minsky did that all as a rigorous scientist, however as a scientist with an almost unbounded ability to dream and intellectual creativity. Namely these his abilities connect him with other pioneers of robotics, and provide an opportunity for us to mention, after mentioning his anniversary, and the anniversary of AI, also a third anniversary, the anniversary of the birth of the word *robot*.

Figure 3
Karel and Josef Čapek. A photo from the twenties of the 20th century

In 2006, at least the world of theatre celebrated the 85th anniversary of the moment, when robots appeared first at the stage of the Prague's National Theatre, and started their successful march through numerous branches of culture, science and technology of the 20th century. The first night of the theatre play R. U. R. (Rossum's Universal Robots) by Karel Čapek, the play which evoked many of the appearing and appealing problems with numerous social and political dimensions of the co-existence of the human culture and the industrial production, of the co-existence of human and engineered, was in January 25, 1921.

This is also the date when the world *robot* – invented by Karel's brother Josef Čapek – was first used for denoting artificially created human-like beings. Perhaps it is not without interest that the play has been written during the summer mounts of 1920 in the Slovak spa Trenčianske Teplice, and that the neologism of the Slavonic origin, the word 'robot', has been proposed in the same place. However, this topic goes slightly out of the subject of present contribution. More about the cultural destiny of the word robot, and the cultural history of hunam-like creatures and human beings co-existence with such creatures see in (Horáková, 2006) or (Horáková, Kelemen, 2006).

Acknowledgment

References

[1] Coad, P., Nicola, J.: *Object-Oriented Programming*. Yourdon Press, Upper Saddle River, NJ, 1993

[2] Csuhaj-Varjú, E., Kelemen, J., Kelemenová, A., Paun, Gh., Vaszil, Gy.: Computing with Cells in Environment – P Colonies. *Journal of Multiple Valued Logic and Soft Computing* **12** (2006) 201-215

[3] D'Inverno, M., Luck, M.: *Understanding Agent Systems*. Springer, Berlin, 2004

[4] Fensel, D.: *Ontologies – a Silver Bullet for Knowledge Management and Electronic Commerce*. Springer, Berlin, 2004

[5] Harrison, N., Minsky, M.: *The Turing Option*. Warner Books, New York, 1992

[6] Hillis, W. D.: *The Connection Machine*. The MIT Press, Cambridge, MA, 1985

[7] Horáková, J.: Staging Robots – Cyborg-Culture as a Context for Robots Emancipation. In: *Cybernetics and Systems, Proc. EMCSR 06, Vol. 1* (R.

Trappl, Ed.). Austrian Society for Cybernetic Studies, Vienna, 2006, pp. 312-317

[8]   Horáková, J., Kelemen, J.: From Golem to Cyborg – a Note on the Cultural Evolution of the Concept of Robots. *Human Affairs* **16** (2006) 83-98

[9]   Kelemen, J.: The Minsky Way (From Artificial Neural Networks to Emotion Machines). In: *Proc. 9th International Conference Informatics 2007* (Plander, I., Ed.). Slovak Society for Applied Cybernetics and Informatics, Bratislava, 2007, pp. 131-138

[10]   Kelemen, J.: Od umělých neuronů ke strojům s emocemi. *Computerworld* **15** (Czech Edition) No. 9 (2004) 24-25

[11]   Lúčny, A.: Building Complex Systems with Agent-Space Architecture. *Computing and Informatics* **23** (2004) 1-36

[12]   McCarthy, J., Minsky, M. L., Rochester, N., Shannon, C. E.: A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. *The AI Magazine* **27** (Winter 2006) 12-14

[13]   McCulloch, W., Pitts, W. A.: A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bull. Math. Biophys*. **5** (1943) 115-133

[14]   Minsky, M.: *The Emotion Machine*. Simon & Schuster, New York, 2006

[15]   Minsky, M.: *The Society of Mind* (Interactive CD Edition). The Voyager Company, Boston, MA, 1996

[16]   Minsky, M.: *The Society of Mind*. Simon & Schuster, New York, 1986

[17]   Minsky, M.: K-Lines – a Theory of Memory. *Cognitive Science* **4** (1980) 117-133

[18]   Minsky, M.: A Framework for Representing Knowledge. In: *The Psychology of Computer Vision* (P. H. Winston, Ed.). McGraw-Hill, New York, 1975, pp. 211-277

[19]   Minsky, M.: *Computation – Finite and Infinite Machines*. Prentice Hall, Englewood Cliffs, NJ, 1967

[20]   Minsky, M.: *Theory of Neural-Analog Reinforcement System and its Applications to the Brain-Model Problem* (Ph. D. Thesis). University of Princeton, Princeton, NJ, 1954

[21]   Minsky, M.: *A Neural-Analogue Calculator Based upon a Probability Model of Reinforcement*. Psychological Laboratories, Harvard University, Cambridge, MA, 1952

[22]   Minsky, M., Papert, S.: *Perceptrons – An Introduction to Computational Geometry*. The MIT Press, Cambridge, MA., 1969

[23]   Rosenblatt, F.: *Principles of Neurodynamics*. Spartan Books, Washington, D. C., 1961

[24]   Singh, P.: Examining the Society of Mind. *Computing and Informatics* **22** (2003) 521-543

[25]   Turing, A. M.: On Computable Numbers, with an Application to the Entscheidungsproblem. *Proc. London Math. Soc.* **42** (1936) 230-265; corrections **43** (1937) 544-546