A Personal View on AI, Rough Set Theory and Professor Pawlak

Toshinori Munakata

Computer and Information Science Department, Cleveland State University
Cleveland, Ohio 44115, U.S.A.
t.munakata@csuohio.edu

It is an honor to contribute my short article to this special issue commemorating the life and work of Professor Zdzisław Pawlak. In this article I would like to discuss my encounters with the field of artificial intelligence (AI) in general, and how I see rough set theory and Professor Zdzisław Pawlak in this context. I have been fortunate to know some of the greatest scholars in the AI field. There are many of them, but if I had to choose the three I admire most, they are: Professors Zdzisław Pawlak, Lotfi Zadeh and Herbert A. Simon. There are common characteristics among all of them. Although they are the most prominent of scholars, all are frank and easy and pleasant to talk with. All are professionally active at ages where ordinary people would have long since retired.

I became interested in the field of AI in the mid 70s. I have observed many ups and downs of the field in terms of the popularity since then - a common phenomena in any field. The AAAI (American Association for Artificial Intelligence) was inaugurated and the first issue of the AI Magazine was published in the spring 1980. The timing of the birth of rough set theory was soon after this event. At this time many people in the world were becoming interested in the field of AI, while there were only a handful researchers when the field started in the 1950s. In the spring of 1986, the first issue of the *IEEE Expert* (now *IEEE Intelligent Systems*) was inaugurated. I served as an Associate Editor of this magazine for two terms from 1987 to 1991. In terms of public popularity AI was flourishing in this eras.

During many years of the 70s and 80s, I observed that despite media hype and claims for break-thorough technologies, most AI techniques were not practical. Here “practical” means “having real-world commercial and industrial applications on an everyday basis.” For example, I could not find cases where machine learning techniques discussed in textbooks such as “learning from examples” and “learning from analogy” were actually employed at industrial plants or commercial banks. The same were true for other AI techniques such as blackboard and neural networks. After Minsky’s pessimistic view on the field, the U.S. government funding ceased, and only a handful researchers remained active in the field. The field revived in the mid to late 80s, and became quite popular. However, I could not find a single case where neural networks were actually used every day for commercial and industrial applications. For all of these observations I could be wrong because there could have been exceptions I was not aware of, but I was certain that these exceptions were few, if any.

J.F. Peters et al. (Eds.): Transactions on Rough Sets VI, LNCS 4374, pp. 247–252, 2007.

This situation of impracticality of most AI techniques appeared to start to change around 1990. That is, many AI techniques were becoming truly practical and their application domains much broader. Of course, there were practical AI techniques before 1990. Robotics was one. The first major industrial fuzzy control was implemented in Denmark in 1982, followed by the famous Hitachi's Sendai subway control in 1986. There had been practical expert systems in the 80s. However, the repertories of AI techniques and their application domains were becoming much more extensive around 1990.

With this background, in March 1993 I made a proposal to the *Communications of the ACM* (*CACM*) for a Special Issue entitled "Commercial and Industrial AI." The *CACM* was a primary computer science magazine subscribed by some 85,000 professionals worldwide at that time. Its readers went far beyond the computer science discipline including fields such as engineering, social science and education. The proposal was accepted and I selected the most practical or promising AI areas with the help of many experts. The Special Issue was published in March 1994 [3] and was divided into four major sections with 11 articles. They are: I. "Knowledge Engineering Systems" with two articles – an article on general expert systems and an article on case-based reasoning. II. "Perception, Understanding, and Action" with three articles on vision, speech and robotics. III. "Fuzzy Systems" with two articles – an overview and a soft computing article by Professor Zadeh. IV. "Models of the Brain and Evolution" with four articles - two articles by Rumelhart, Widrow, et al., an article on neural networks in Japan, and an article on genetic algorithms. There were many behind-the-scene stories and one of them was that my original plan was to have only one article by Rumelhart, et al. After much delay, they had an article twice as long as originally planned, and I suggested splitting the manuscript into two parts.

In the Guest Editors Introduction, I wrote:

The practical application of artificial intelligence (AI) has been the center of controversy for many years. Certainly, if we mean AI to be a realization of real human intelligence in the machine, its current state may be considered primitive. In this sense, the name artificial "intelligence" can be misleading. However, when AI is looked at as "advanced computing," it can be seen as much more. In the past few years, the repertory of AI techniques has evolved and expanded, and applications have been made in everyday commercial and industrial domains. AI applications today span the realms of manufacturing, consumer products, finance, management, and medicine. Implementation of the correct AI technique in an application is often a must to stay competitive. Truly profitable AI techniques are even kept secret.

Many of the statements I wrote here are still basically true today. This Special Issue turned out to be a big hit. The ACM printed 1,000 extra copies of this issue for back orders, but they sold out less than a month. A person from a Japanese company wanted to purchase a box of fifty copies of this issue, but it was too late. The issue became one of the most cited *CACM*, for not only within computer science but also some unexpected places such as *Scientific American*, the Berkeley Law School, the Stanford Philosophy Department, etc. The US Air Force has issued interesting predictions in the past. Around 1947, the Air Force predicted likely technologies for the next 50 years. They included jet rather than propeller powered, and supersonic airplanes. They became the reality by 1997. In 1997, they issued predictions for the next 50 years, i.e., by 2047. Between these major predictions, they published future perspectives for shorter time ranges. My Special Issue was cited in a document within a report entitled "Air Force 2025" that describes 30-year predictions by the US Air Force [1].

When I was preparing my first *CACM* Special Issue, I knew there were other AI areas that were not covered. As soon as the first issue was nearly complete, I started working on a follow-up Special Issue entitled "New Horizons in Commercial and Industrial AI." In the "Editorial Pointers" in the first issue, Executive Editor Diane Crawford wrote: "He has assembled some of the foremost minds in AI to author and/or review the 11 articles presented here. If that weren't enough, he's already digging into another proposed issue for *Communications* to appear early next year, where he hopes to address new horizons and applications in other AI-related fields."

For the first Special Issue I received many responses. One of them was a letter from Professor Herbert A. Simon of Carnegie-Mellon University, a prominent scholar in AI with a Turing Award and a Nobel Prize in economics. Basically, he stated: "The first issue was well done, although if I were the Guest Editor I would have had less emphasis on neural networks and included an article on machine learning." He suggested placing an article on machine learning in the second issue. I replied to him saying I had already planned that and asked him to write one, and subsequently he co-authored an article. I was lucky to be able to have close contact with Professor Simon. When IBM's Deep Blue defeated the human chess champion Garry Kasparov in 1997, he and I co-authored a commentary article on the significance of this event on AI [14]. He was a pleasant person to talk with. He was a fulltime professor and active until two weeks before his death in 2001 at age 84. For the second Special Issue I had planned to include certain topics from a very early stage. They included symbolic machine learning, natural language processing (e.g., machine translation) and logic programming. Also, I wanted to include articles addressing the commonsense problem, although I did not expect that this area would have many commercial or industrial applications in the near term.

At a later stage of preparation of the second issue, I searched for additional areas appropriate for the issue, and found rough set theory. I was not familiar with this area, but from what I found I thought it was a promising technique, appropriate for the second issue. Perhaps it could complement other AI techniques. I contacted Professor Pawlak and asked him whether he was interested in contributing an article to such a *CACM* Special Issue. These were my first encounters with rough set theory and Professor Pawlak. He kindly accepted my invitation and contributed an article co-authored with Professors Jerzy GrzymalaBusse, Roman Slowinski and Wojciech Ziarko [8]. This was my first acquaintance with the rough set community. As said earlier, *CACM* has a large number of audience worldwide and its impact is high. I don't know how much the appearance of this article has influenced the promotion of this theory, but I think at least it helped to introduce the term "rough sets" worldwide.

Incidentally, when I studied practically successful symbolic machine learning techniques for the first time, such as ID3, I was a bit disappointed. From the term "learning," I expected some elements of human-like learning. For example, given a specific experience the machine would abstract it, generalize it and be able to use it for similar circumstances in the future. I did not see such human-like learning in ID3. Rather, it simply classifies data based on entropy in information theory. The characteristics of the target data seemed to be too simple. Perhaps the term "learning" was misleading, and probably I expected too much on what we could do from this kind of data. Both ID3 and rough sets can learn from data, but probably ID3 had attracted more attention than rough sets in the scientific community, at least during the 80s and 90s. Why? One reason might be that ID3 appears to have been in the main stream in the machine learning community, and had received more support from its early introduction. Professor Simon was one of it's supporters, and he was a great scientist as well as a good salesman to promote his beliefs. For example, he called a software system he developed a "general problem solver," which implied, with a bit of exaggeration, the system would solve every problem on earth. He was also an optimist. In the late 1950s he predicted that a computer would defeat a human chess champion within 10 years. We waited 10 years, another 10 years, and so forth for the next 40 years. In contrast, Professor Pawlak was a humble and modest scientist and perhaps not particularly a good salesman. In my opinion, rough set theory was not as widely recognized in the AI and CS fields as it should have been.

After my first encounter with the rough set community through my *CACM* second special issue, I have been fortunate to be able to work in this field together with these people. I attended several rough set related conferences after my first encounter [4, 5, 6]. To promote rough sets, I could think of two among many possibilities. One was to have promotional articles in journals of large audience like the *CACM*. Another area was to have a rough set application with a high social impact. For the latter, rough control might be a good candidate, I thought. Fuzzy set theory became a hot topic after Hitachi successfully applied fuzzy logic to Sendai subway control. I tried to push rough control, and I was Chair of the rough control interest group. The basic idea of rough control is to employ rough sets to automatically generate input-to-output control rules [7, 9]. The idea was not particularly new, but breakthrough applications would place rough set theory in the spotlight. A long time rough set activist Professor T.Y. Lin financially supported me for this endeavor. Although we have not observed a major breakthrough yet, I think possibilities are still there. In the communication with Professor Pawlak, he suggested presenting a co-authored conference paper [13]. When I published an AI book from Springer, I included a chapter for rough sets [10]. When I served as Guest Editor for third time for *CACM* Special Section on knowledge discovery [11], I asked Professor Ziarko to contribute an article.

When a young colleague approached me to work on a data mining article, I suggested employing rough sets [2]. I am currently working on another article on a rough set application with a young assistant professor.

Although we all saddened by the recent death of Professor Pawlak, I think he was fortunate to observe that his theory has been widely recognized in the scientific community worldwide. This was not necessarily the case for many great scholars in the past. During my sabbatical in the fall of 2002, I traveled to Poland, visiting Professors Slowinski, Skowron and Pawlak, and received a warm welcome. This was the last time I saw Professor Pawlak.

What are the future prospects of rough sets? No one knows, but the following is my speculation. Despite it's founder's death, the community will grow – there will be more researchers worldwide and more theoretical and application developments. But, growth in the field may level out eventually, unless we achieve major breakthroughs. As in the case of other machine learning techniques and AI in general, we don't know what, when or if such breakthroughs may come. Targeting to extremely large volumes of data (e.g., terabytes) and/or massively parallel computing alone do not look very promising, as we have observed similar attempts such as the Cyc and the Connection Machine. For knowledge discovery techniques such as rough sets, there may be a limit when we deal only with decision tables. Perhaps we should also look at other formats of data as well as other types of data, for example, non-text, comprehensive types of information, such as symbolic, visual, audio, etc. Also, the use of huge background knowledge, in a manner similar to human thought, would be necessary and effective. Human-computer interactions would also enhance the discovery processes. Other totally different domains are non-silicon based new computing paradigms. I am currently working on my fourth Special Section for the *Communications of the ACM* as a guest editor on this subject [12]. These approaches may lead to a new dimension of information processing in a wide range of application domains including rough sets. As with other scientific developments in history, such as alchemy and the first airplane, a breakthrough may come in a totally unexpected form.

References

1. Clarence E. Carter, et al. The Man In The Chair: Cornerstone Of Global Battlespace Dominance, *Air Force 2025*, 1996.
2. Brenda Mak and T. Munakata. "Rule extraction from expert heuristics: A comparative study of rough sets with neural networks and ID3," *European Journal of Operational Research*, 136(1), pp. 212-229, 2002.
3. Toshinori Munakata. "Guest Editor's Introduction," in the Special Issue "Commercial and Industrial AI Applications," *Communications of the ACM*, 37(3), pp. 2325, 1994.
4. Toshinori Munakata. "Commercial and Industrial AI: Where it is now, and where it will be," an invited talk at *the Third International Workshop on Rough Sets and Soft Computing (RSSC'94)*, San Jose, CA, Nov., 1012, pp. 5155, 1994.

5. Toshinori Munakata. "Rough Control: A Perspective," *Workshop on Rough Sets and Data Mining, the 1995 ACM Computer Science Conference*, Nashville, TN, March 2, 1995.
6. Toshinori Munakata. "Commercial and Industrial AI and a Future Perspective on Rough Sets," in *Soft Computing: the Third International Workshop on Rough Sets and Soft Computing (RSSC94)*, T.Y. Lin and A.M. Wildberger (Eds.), the Society of Computer Simulation, pp. 219222, 1995.
7. Toshinori Munakata. "Rough Control: Basic Ideas and Applications," *Workshop on Rough Set Theory, the Second Annual Joint Conference on Information Sciences*, Wrightsville Beach, NC, Sept. 28 Oct. 1, 1995.
8. Toshinori Munakata. "Guest Editor's Introduction," in the Special Issue "New Horizons for Commercial and Industrial AI," *Communications of the ACM*, 38(11), pp. 28-31, 1995.
9. Toshinori Munakata. "Rough Control: A Perspective," in T.Y. Lin and N. Cercone (Eds.), *Rough Sets and Data Mining: Analysis of Imprecise Data*, Kluwer Academic, pp. 7788, 1997.
10. Toshinori Munakata. *Fundamentals of the New Artificial Intelligence: Beyond Traditional Paradigms*, Springer-Verlag, 1998, 2007.
11. Toshinori Munakata. "Guest Editor's Introduction," in the Special Section "Knowledge Discovery," *Communications of the ACM*, 42(11), pp. 26-29, 1999.
12. Toshinori Munakata. "Guest Editor's Introduction," in the Special Section "Beyond Silicon: New Computing Paradigms," *Communications of the ACM*, to appear in 2007.
13. Zdzisław Pawlak and Toshinori Munakata. "Rough Control: Application of Rough Set Theory to Control," *EUFIT '96 the Fourth European Congress on Intelligent Techniques and Soft Computing*, Aachen, Germany, Sept. 25, 1996, pp. 209-217.
14. Herbert A. Simon and Toshinori Munakata. "Kasparov vs. Deep Blue: The Aftermath - AI Lessons," *Communications of the ACM*, 40(8), pp. 23-25, 1997.