Proceedings of the Federated Conference on Computer Science and Information Systems pp. 189–190
DOI: 10.15439/2016F002
ACSIS, Vol. 8. ISSN 2300-5963

Working with Zdzisław Pawlak – Personal Reminiscences

Victor Marek
Department of Computer Science
University of Kentucky
Lexington, KY 40506, USA

***Abstract*—This paper accompanies panel contribution of the author to the session devoted to personal reminiscences of Professor Zdzisław Pawlak, a computer scientist and engineer. In particular we discuss some aspects of the work of Pawlak and researchers in his circle of collaborators in 1960ies, and especially, 1970ies. Given the lack of archival materials, the author bases this writing on personal recollections which may, at places, be imprecise.**

***Index Terms*—Zdzisław Pawlak, Information Storage and Retrieval Systems, Rough Sets**

I. How Did It Start

GIVEN that this text refers to personal recollections, I need to introduce myself to the reader. I finished High School in 1960, and started studies in Mathematics at Warsaw University. Soon, I met socially Andrzej Ehrenfeucht and he suggested attending a seminar that a group of scientists conducted at the Mathematical Institute of the Academy of Sciences. It was a series of meetings that can be characterized by a completely informal atmosphere and truly interdisciplinary subject of talks. Among the leaders of the groups were Robert Bartoszyński, Andrzej Ehrenfeucht, Zdzisław Pawlak and other future leaders of Computer Science in Warsaw. The atmosphere was very informal and the hierarchical relationship so often visible in the university education and research was, simply, absent. I was the youngest person attending the seminar. This informal atmosphere was very different from the seminars lead at Warsaw University by Professors Andrzej Mostowski (the head logician in Warsaw) and by Professor Helena Rasiowa. The main topics of these informal seminars were foundations and applications of computers. Of course, there were very few actual computers (but there were some) and the topics were an interesting mixture of entirely theoretical topics (for instance Turing Machines) and Hao Wang's experiments with proving theorems of "Principia Mathematicae" on IBM machines. There was strong presence of individuals interested in biomedical information (or, to be precise, what today is called such). This resulted from the research of Zdzisław Pawlak who, at the time, was interested in DNA as built by formal grammar. Eventually, Pawlak wrote a book (in Polish! [1]) devoted to this research. From my own perspective, the important aspect of the seminar was opening of a perspective – computers and their foundational issues.

In 1964 I completed my M.Sc. in Mathematics studies and joined Professor Mostowski group. This was Foundations of Mathematics group but also comprised of algebraists. Soon, Andrzej Ehrenfeucht emigrated to United States eventually settling in Boulder, CO (where he still teaches). Mostowski and his research group got involved in research on a new (and very exciting) topic called "forcing". This was, at the time completely mysterious, technique for independence proofs in Set Theory. In my case, I worked in this area for a number of years. But in 1970 I went for a post-doc appointment in Utrecht, Holland, in a group of Professor Dirk van Dalen. Their interests were different and were motivated by different aspects of Foundations. A strong influence on my thinking was exercised by Henk Barendregt and his research on λ-calculus. Even more importantly, as I was returning from Holland, I visited Arhus University in Denmark (where my colleague Dr. Janusz Onyszkiewicz was a post-doc) and noticed that logicians there were mainly interested in Computer Science considerations. This made me thinking that maybe it is time to look again at computers.

II. Changed Perspective

The year 1970 and the revolt of workers resulted in significant changes in Poland. From the point of view of Computer Science studies there were significant changes, too. Warsaw University reorganized its science programs. Mathematicians and physicists parted ways. The Mathematics and Physics faculty divided and Mathematics faculty became Mathematics, Mechanics and Computer Science. Computing Center of the Academy of Sciences evolved (first informally, then formally) into the Institute of Computer Science. Mathematicians in the Academy created a venue for Mathematical research called Banach Center. This place welcomed not only "pure" mathematicians but also computer scientists. Yet another important change was creation of "Technical Physics and Applied Mathematics" program at Warsaw Technical University. This program attracted students that were interested in applications, and became a premier program in Computer Science. The alumni of this program included Witold Lipski, Tomasz Imieliński, Mirosław Truszczyński and several other outstanding researchers. At the same time Professors Pawlak and Rasiowa created a publication venue; a journal *Fundamenta Informaticae*. This journal, associated with Polish Mathematical Society became, eventually one of premier places for the publication of Computer Science research. The very name of that journal (with the word *Fundamenta*) alluded to the great

978-83-60810-90-3/$25.00

traditions of foundational research, as done in Poland after WWI.

III. Information Storage and Retrieval Systems

So, after I came back from post doctoral stint in Holland, and seeing that many logicians started to do research in Computer Science, I was certainly open to look at the problems that were grounded in applications of computers. Given the academic system in Poland where one had to write *Habilitation* (a degree established in Germanic and other countries of Europe) took time, sometimes around 1973 I was ready to expand my interests into Computer Science. Then came a series of phone calls for Professor Pawlak (i.e. Zdzisław, recall the 1960ies seminar). After some time I understood the idea. It was a model of databases. Recall that since the work of E.F. Codd on the relational model of database, computer scientists investigated logical formalisms that would eliminate need for the fluency in data structures (such as double linked lists) to operate databases. Eventually several proposals for logic-based languages were proposed and SQL and many of its variations were adopted as a logic-based language for what is known as relational databases. But the Pawlak's proposal [2] differed from the Codd's proposal in several important details. The single table idea was closely related to so-called universal instance. In fact significant effort of a large group of researchers of relational model was spent on decomposing tables, since universal instances were too big to be effectively processed. This research and the enormous effort spent by the database community is at least partly forgotten now, as, with the increased processing capacity of computers, often one denormalizes databases. Other differences of Pawlak's proposal versus SQL-based databases was that SQL admitted null values (thus is, really, based on multivalued logic). But the real difference (and the one that was quietly incorporated into SQL-based systems) was the fact that Information Storage and Retrieval Systems (ISRs), from the very beginning admitted duplicate objects. In other words, different objects could have the same descriptions. This property of ISRs implied existence of (potentially nontrivial) equivalence relation on the set of objects, $\equiv$, namely: $o_1 \equiv o_2$ iff description of o_1 is the same as description of o_2.

The studies of ISRs were conducted by many researchers in Poland and other countries, and the reason for it was that the logicians were ready to investigate ISRs. Let me mention that the group of researchers in Warsaw alone consisted of over 10 individuals. Led mostly by Witold Lipski it included people in the Technical University, Warsaw University and the Academy of Sciences: Tomasz Imieliński, Paweł Traczyk, Michał Jeagermann, Cecylia Rauszer, Andrzej Jankowski, myself, and many others. Lipski and Imieliński tied ISR to the relational model and very soon we were aware of similarities and of differences with Relational Model.

The motivation of ISR came, not surprisingly, from the work of Pawlak with the physicians, more generally, biomedicine researchers. Once one thinks about it, it is clear that computerized medical records of patients may be quite similar or even (after elimination of personal identifiable information) identical. Since the idea in the background was to "mine" the information and find dependencies present in such data, the presence of duplicates was, actually, an interesting information.

IV. Rough Sets, Handling Different Description Languages

The ISRs were based on a first-order language; the objects possessed descriptions. But one observation (motivated by potential biomedical applications) was that there may be more than one language associated with a set of objects. To give one example let us look at a collection of patients. We may have a language associated with symptoms exhibited by the patients but there may be another language to describe the patients, one based on objective data such as various biochemical data (levels of enzymes, presence or absence of specific genes, etc.). The question asked by Pawlak, and one that lies behind the concept of rough sets [3] is this: do we need to specify the description language for the set of objects, or could we, instead, abstract from the specific language and consider equivalence relation "having the same description"? Once this fundamental question was asked, the concept of rough sets, and the associated "interior" and "closure" operations became very natural. Moreover, such approach opened, immediately, the need to look at numerical parameters associated with rough sets (for instance: measures of roughness such as various ratios (for instance of the boundary to the closure), and various characteristics of the equivalence relation (such as the discrepancy - the ratios of sizes of "large" equivalence classes to the small ones). The rest is history; rough sets were introduced, studied, and – most importantly – applied. Like many other fundamental concepts, they were specialized, generalized, investigated for relationship with other concepts (from logic, topology, universal algebra, and combinatorics). Contributions of many researchers, from many countries, and with many interests witness to importance of Pawlak's intuitions.

References

[1] Z. Pawlak, *Grammar and Mathematics (In Polish)*. PZWS, 1965.
[2] V. W. Marek and Z. Pawlak, "Information Storage and Retrieval Systems: Mathematical Foundations," *Theor. Comput. Sci.*, vol. 1, no. 4, pp. 331–354, 1976.
[3] Z. Pawlak, *Rough Sets – Theoretical Aspects of Reasoning about Data*. Kluwer, 1991.