Summary of the doctoral dissertation: *Medical diagnosis support based on fuzzy if-then rules*
Sebastian Porębski

Dissertation deals with the topic of medical diagnosis support. Diagnosis of the disease based on its symptoms, medical test results. It is the domain of diagnosticians because they are responsible for the diagnostic decisions. Nevertheless, part of the responsibility, although indirectly, is borne by a medical apparatus designers: electronic engineers, IT specialists and constructors. For a long time, medicine makes use of technical achievements such as magnetic resonance imaging or measuring the body's electrical activity. However, all measurements of the these signals or diagnostic images of the patient's body are imprecise.

The development of computer infrastructure in hospitals allows recording and processing information about patients under observation or therapy. It is possible to use knowledge engineering to extract new and useful knowledge from collected medical data. In the dissertation, it is postulated to model this knowledge in the form of the fuzzy if-then rules. Knowledge in the form of diagnostic rules can form the basis of the electronic diagnosis support system.

Providing a useful diagnostic support system may increase the diagnosis efficiency, reduce its time and cost, and help to objectify the diagnosis itself. Choosing the fuzzy if-then rules as a model of the knowledge extracted from medical data results from the fact that a useful diagnostic support system is one that provides a clear knowledge base, an understandable mechanism of inference and easy communication with the user. As will be explained, the fuzzy if-then rules and the chosen method of diagnosis support using the belief measure of the Dempster-Shafer theory successfully meet the above requirements. The Dempster-Shafer theory is chosen to diagnosis support because it is recognized as an effective method of decision support in the face of uncertainty. The belief measure, generated from the basic probability assignment, as proven by the presented results, allows using available, although imperfect, medical information and efficiently support a diagnosis.

The main element of the thesis is the introduction of rule extraction algorithms, focusing on the obtaining the compromise between the complexity of the extracted rules and their efficiency. The most important and unambiguously positive effect of the proposed methods is that the extracted diagnostic rules based on the training data, obtain satisfactory generalization efficiency while maintaining readability (a small number of logical conditions) and ease of interpretation.

As part of the dissertation, the research steps are described that led to implementation of the diagnostic rule extraction algorithms. An example of the described rule selection algorithm was presented with the use of a non-medical "iris flowers" dataset only due to its popularity. Later, the proposed methods were verified using five medical benchmark databases. Also, dissertation presents the results of the extraction of diagnostic rules from the actual medical dataset related to the diagnosis of liver fibrosis collected by physicians during several years of therapy of patients with the hepatitis C virus. Finally, the obtained results were compared with several of the latest reference rule extraction methods. A detailed comparison allows to conclude that the use of Dempster-Shafer theory and rule selection algorithms improve the quality of the medical knowledge and diagnosis support.