World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:11, No:11, 2017

Hardware Implementation on Field Programmable Gate Array of Two-Stage Algorithm for Rough Set Reduct Generation

Authors : Tomasz Grzes, Maciej Kopczynski, Jaroslaw Stepaniuk

Abstract : The rough sets theory developed by Prof. Z. Pawlak is one of the tools that can be used in the intelligent systems for data analysis and processing. Banking, medicine, image recognition and security are among the possible fields of utilization. In all these fields, the amount of the collected data is increasing quickly, but with the increase of the data, the computation speed becomes the critical factor. Data reduction is one of the solutions to this problem. Removing the redundancy in the rough sets can be achieved with the reduct. A lot of algorithms of generating the reduct were developed, but most of them are only software implementations, therefore have many limitations. Microprocessor uses the fixed word length, consumes a lot of time for either fetching as well as processing of the instruction and data; consequently, the software based implementations are relatively slow. Hardware systems don't have these limitations and can process the data faster than a software. Reduct is the subset of the decision attributes that provides the discernibility of the objects. For the given decision table there can be more than one reduct. Core is the set of all indispensable condition attributes. None of its elements can be removed without affecting the classification power of all condition attributes. Moreover, every reduct consists of all the attributes from the core. In this paper, the hardware implementation of the two-stage greedy algorithm to find the one reduct is presented. The decision table is used as an input. Output of the algorithm is the superreduct which is the reduct with some additional removable attributes. First stage of the algorithm is calculating the core using the discernibility matrix. Second stage is generating the superreduct by enriching the core with the most common attributes, i.e., attributes that are more frequent in the decision table. Described above algorithm has two disadvantages: i) generating the superreduct instead of reduct, ii) additional first stage may be unnecessary if the core is empty. But for the systems focused on the fast computation of the reduct the first disadvantage is not the key problem. The core calculation can be achieved with a combinational logic block, and thus add respectively little time to the whole process. Algorithm presented in this paper was implemented in Field Programmable Gate Array (FPGA) as a digital device consisting of blocks that process the data in a single step. Calculating the core is done by the comparators connected to the block called 'singleton detector', which detects if the input word contains only single 'one'. Calculating the number of occurrences of the attribute is performed in the combinational block made up of the cascade of the adders. The superreduct generation process is iterative and thus needs the sequential circuit for controlling the calculations. For the research purpose, the algorithm was also implemented in C language and run on a PC. The times of execution of the reduct calculation in a hardware and software were considered. Results show increase in the speed of data processing.

Keywords : data reduction, digital systems design, field programmable gate array (FPGA), reduct, rough set
Conference Title : ICIIP 2017 : International Conference on Intelligent Information Processing
Conference Location : Havana, Cuba
Conference Dates : November 23-24, 2017