

Case Base Indexing Using Unique Attribute Combinations Based on Parallel Power Set Tree Algorithm

Kareem Amin^{1,4}, Khaled El-Bahnasy², Klaus-Dieter Althoff^{1,3}, and Andreas Dengel^{1,4}

¹ German Research Center for Artificial Intelligence, Knowledge Management, Trippstadter Strae 122, 67663 Kaiserslautern, Germany,

kareem.amin,klaus-dieter.althoff,andreas.dengel@dfki.uni-kl.de,

² Ain Shams University, Khalifa El-Maamon st, Abbasiya sq., Cairo, Egypt, khaled.elbahnasi@cis.asu.edu.eg

³ Institute of Computer Science, Intelligent Information Systems Lab, University of Hildesheim, Hildesheim, Germany

⁴ Kaiserslautern University, P.O. Box 3049, 67663 Kaiserslautern, Germany

Abstract. The high complexity of cases and attributes and the growing amount of knowledge lead to latency in retrieving cases in reasonable time. In this paper we present a power set tree indexing methodology to find the unique attributes combinations. A power set tree can despite its high complexity, generate all possible attributes combinations to find the unique ones, then use these unique combinations as references to the original cases which can reduce the number of comparisons required to retrieve the most similar cases. We will build the parallel power set tree for each case using the adding-feature methodology to overcome the limitations of hardware resources. A lot of unbalanced consumption of resources obstacles have been resolved through our current and last researches. At the end, a better and consolidated parallel algorithm has been built to balance the resources consumptions and harness them to serve the main purposes of finding the unique combinations for each case. The maximum number of attributes we have worked with so far is 37. The experiments are carried out on case bases received from the Egyptian Agricultural Research Center (ARC). The results have shown a significant reduction in number of attributes per case, and some cases have no unique set of attributes. We are discussing the results and reasons of why that happened in details.

Keywords: Case-Based Reasoning, Indexing, Power Set Tree, Parallelization