Decision Rule Systems and Decision Trees over Finite Binary Information Systems (extended abstract)

Mikhail Moshkov[1] and Beata Zielosko[2]

[1] Computer, Electrical and Mathematical Sciences & Engineering Division
King Abdullah University of Science and Technology
Thuwal 23955-6900, Saudi Arabia
mikhail.moshkov@kaust.edu.sa

[2] Institute of Computer Science, University of Silesia
39, Będzińska St., 41-200 Sosnowiec, Poland
beata.zielosko@us.edu.pl

In this paper, we compare time complexity of decision rule systems and decision trees over finite binary information systems.

A *finite binary information system* $U = (A, F)$ consists of a set A and a finite set of attributes F which are defined on A and have values from the set $\{0, 1\}$.

The notion of a *problem over the information system* U is defined as follows. Take some attributes $f_1, \ldots, f_n$ from F. These attributes divide the set A into classes (for each class, values of the attributes are constant on elements from the class). These classes are numbered such that different classes can have the same number. The number of a class is the decision corresponding to elements of the class. For a given element a from A, it is required to recognize the number of a class which contains a. The number n is called the *dimension* of the considered problem.

As algorithms for problem solving, *decision trees* and *decision rule systems* over U are considered which use arbitrary attributes from the set F. The *depth* of a decision tree (the maximum length of a path from the root to a terminal node) and the *length* of a decision rules system (the maximum length of a rule from the system) are considered as time complexity measures.

For the information system U, two *Shannon type functions* $S_{t,U}(n)$ and $S_{r,U}(n)$ are considered which characterize the growth in the worst case of minimum depth of decision trees and minimum length of decision rule systems with the growth of problem dimension n.

A subset of F is called *independent* if no one attribute from this subset can be represented as a function depending on other attributes from the subset. We denote by $in(U)$ the maximum cardinality of an independent subset of the set F.

A problem over U is called *tree-uncancellable* if the minimum depth of a decision tree over U for this problem is equal to its dimension. We denote by $tu(U)$ the maximum dimension of a tree-uncancellable problem over U. A problem over U is called *rule-uncancellable* if the minimum length of a decision rule system over U for this problem is equal to its dimension. We denote by $ru(U)$ the maximum dimension of a rule-uncancellable problem over U. One can show that

$$ru(U) \leq tu(U) \leq in(U).$$

The following statement, which characterizes the behavior of the Shannon function $S_{t,U}$ for decision trees, was published in [1].

Theorem 1. *Let $U = (A, F)$ be a finite binary information system such that $f \not\equiv \text{const}$ for any $f \in F$. Then, for any natural n, the following statements hold:*
a) if $n \leq tu(U)$ then $S_{t,U}(n) = n$;
b) if $tu(U) \leq n \leq in(U)$ then

$$\max\{tu(U), \log_2(n+1)\} \leq S_{t,U}(n) \leq \min\left\{n, 8(tu(U)+1)^5(\log_2 n)^2\right\};$$

c) if $n \geq in(U)$ then $S_{t,U}(n) = S_{t,U}(in(U))$.

We prove the the next statement, which characterizes the behavior of the Shannon function $S_{r,U}$ for decision rule systems.

Theorem 2. *Let $U = (A, F)$ be a finite binary information system such that $f \not\equiv \text{const}$ for any $f \in F$. Then, for any natural n, the following statements hold:*
a) if $n \leq ru(U)$ then $S_{r,U}(n) = n$;
b) if $n \geq ru(U)$ then $S_{r,U}(n) = ru(U)$.

The book [2] contains more examples of comparative study of Shannon functions for decision trees and decision rule systems.

References

1. M. Moshkov, On global Shannon functions of two-valued information systems, in Proceedings of Fourth International Workshop on Rough Sets, Fuzzy Sets, and Knowledge Discovery, Tokyo University, 1996, pp. 142-143.
2. M. Moshkov, B. Zielosko, Combinatorial Machine Learning – A Rough Set Approach, Studies in Computational Intelligence 360, Springer, Berlin, 2011.