The statistical method of effective data structures designing for particular software and hardware environments

V. Shynkarenko, H. Zabula & D. Petin
Dnipropetrovsk National University of Railway Transport, Dnipro, Ukraine

ABSTRACT: The parameters of time efficiency of data structures and the methods of their assessment with the confidence intervals have been developed. The authors have shown the approaches to forming scenarios (the sequences of the data processing operations). The authors have developed the order of parameters determining on the basis of computer experiment. The results of computer experiments compared with a number of the data structure implementations have been given. The obtained objective parameters provide a basis for the automated data structure adaptation.

1 INTRODUCTION

Simple and structured data are differentiated at all the levels of data designing, i.e. at the logical and hardware ones as well as at the level of language implementation (Ziegler 1982).

The indexed array, record, list, tree, stack, deque and queue are the most widely used classical structures (Cherkassky 1999; Knuth 1997; Perevozchikova 2007). At present the associative array, collection and other structures are widespread. These structures are predefined and trivial in programming languages.

The design and development of data structures consist in choosing them from predefined structures or designing more complex ones (Bentley 1982; Cormen 2009). Designing is carried out by putting some structures into other ones, so the elements of structures are structures (i.e. Hoare 1975).

Designing is performed if it is necessary for the solution of the problem, with the solution being done algorithmically. The array of arrays, the list tree, the stack of records, the collection of array trees and so on are the examples of designed structures.

In programming practice, the development of data structures is done intuitively but if there are specific requirements to the time characteristics of the software, the objective efficiency parameters are to be taken into account.

The basis of the design of data structures at the physical level is partially specified at the logical one (Shynkarenko 2009, 2012, 2014).

The data structures are not functional therefore the time efficiency of the data structure is determined by the time efficiency of the processing operations that are implemented in the form of algorithms (Shynkarenko 2001) and can consist of the operations of access as well as changing the order and location of structured elements.

The memory location of the data structures can considerably influence the time efficiency of programs and program systems using structured data. It has been shown (Kaspersky 2003; Shynkarenko 2006, 2009) that the time span of the operation of access to a location can vary by two orders of magnitude depending on the memory location of the data. There are many different approaches and methods of analyzing the algorithm time execution using computational complexity (Chase 1990; Cormen 2009; Galil 1983; Kozen 1992).

2 FORMULATION OF THE PROBLEM

The problem of developing of time efficiency parameters of the data structures and the methods of assessing these parameters experimentally is solved in this paper.

3 THE TIME EFFICIENCY OF THE DATA STRUCTURES

We define the time efficiency of the data structures as the time efficiency of the set of operations (algorithms) of data processing (not including calculations).

Let us specify what is meant by the operations of data processing. Data processing is carried out on the basis of atomic or primitive processing operations. All primitive operations of processing data structures can be divided into: