Some approaches for determining isomorphism of semigroups of small order

S Yu Korabelshchikova\textsuperscript{1}, B F Melnikov\textsuperscript{2}, E A Melnikova\textsuperscript{3} and L V Zyablitseva\textsuperscript{1}

\textsuperscript{1}Northern Arctic Federal University named after M.V. Lomonosov, 17 Northern Dvina, Arkhangelsk, 163001, Russia  
\textsuperscript{2}Shenzhen MSU – BIT University, 1 International University Park Road, Dayun New Town, Longgang District, Shenzhen, Guangdong Province, 518172, China  
\textsuperscript{3}Russian State Social University, 4 Wilhelm Pieck str., Moscow, 129226, Russia

E-mail: ya.e.melnikova@yandex.ru

Abstract. For a given set of graphs or graph generation algorithm, it is possible to determine a set of invariants whose computational complexity will be low, and this set will work effectively in solving the problem of determining the non-isomorphism of graphs on this set of graphs. We described and investigated in detail two invariants of semigroups, reducing them to invariants of graphs, and then compared the effectiveness of these invariants for algorithmizing the problem of determining the non-isomorphism of semigroups. The comparison results are given in the article.

1. Introduction

One of the urgent problems of the theory of semigroups is the problem of establishing isomorphism. It can be formulated as follows: is there an algorithm that establishes an isomorphism of a pair of semigroups in polynomial time. In many applied disciplines questions arise that reduce to this problem, and for their solution, as a rule, algorithms for exhaustive search are used. In this case, various heuristics can be used to optimize the search. Possible applications of this approach are described by the authors in [1,2].

Another approach to establishing isomorphism is as follows: semigroups are compared according to some characteristics (invariants), and only if these characteristics coincide, it is enumerated. Before, we used similar approaches are used in graph theory and some related scientific fields [3, 4,5]. Depending on the generation algorithm used, we can assume about which of the invariants of the graph will be more effective in determining the non-isomorphism of graphs.

Thus, for a given set of graphs or graph generation algorithm, it is possible to determine a set of invariants whose computational complexity will be low, and this set will work effectively in solving the problem of determining the non-isomorphism of graphs on this set of graphs (for this generation algorithm). Moreover, there is no need to use difficultly computable invariants. We described and investigated in detail two invariants of semigroups, reducing them to invariants of graphs, and then compared the effectiveness of these invariants for algorithmizing the problem of determining the non-isomorphism of semigroups. The comparison results are given in the article.

Recall that an invariant is a certain numerical value or their ordered set, which coincides for all isomorphic objects. It is clear from the definition that the coincidence of invariants is a necessary
condition for isomorphism. Since the polynomial algorithm for checking semigroups for isomorphism is unknown, it is advisable to compare them according to some list of invariants. At the same time, several questions arise: in what order to calculate the invariants of two semigroups, which of the invariants better recognizes semigroups that are not isomorphic? This article is devoted to the study of these issues. Considering that previously similar problems were studied by the authors for graphs \([6, 7, 8]\), here we consider invariants of semigroups that are like or reduced to invariants of graphs. In \([9]\) there is a more detailed description of the algorithms used for this purpose and a variant of their implementation in C++.

Any finite semigroup can be defined by a Cayley table, a matrix of dimension \(n \times n\), where \(n\) is the order of the semigroup. Each element \(a_{ij}\) in this table is a product of the elements \(a_i\) and \(a_j\).

Consider the invariants of semigroups that are like the invariant of a graph that associates with a graph a set of degrees of vertices. As invariant 1, we consider an ordered vector, where is the number of times that an element occurs in the Cayley table.

For invariant 2, we consider an ordered vector \((d(a_1), d(a_2),...,d(a_n))\) where \(d(a_i)\) is the number of different elements that can be obtained by multiplying the element \(a_i\) by itself.

In addition to the invariants considered above, semigroups can be compared directly using graph invariants. To do this, we associate with each semigroup a certain graph. As vertices of the graph, we take elements of the semigroup \(A\), and an edge between a pair of elements \(e = (a, c)\) exists if and only if:

\[ \exists b \in A : a \cdot b = c. \]

Example 1. Let \(A\) be a finite semigroup defined by a Cayley table:

|     | a   | b   | c   | d   |
|-----|-----|-----|-----|-----|
| a   | a   | a   | a   | a   |
| b   | a   | b   | a   | a   |
| c   | a   | a   | c   | a   |
| d   | a   | a   | a   | d   |

Then \(A\) will be associated with a directed graph \(G_A\) defined by the adjacency matrix:

\[
G = \begin{pmatrix}
4 & 0 & 0 & 0 \\
3 & 1 & 0 & 0 \\
3 & 0 & 1 & 0 \\
3 & 0 & 0 & 1 \\
\end{pmatrix}
\]

Consider another variant of mapping a graph to a semigroup. Let \(A\) be a commutative semigroup of idempotents (semilattice). On the semilattice \(A\), the relation of the natural partial order is given by the formula:

\[ a \leq b \iff a \cdot b = b \cdot a = a. \]

We associate with the semigroup \(A\) an oriented graph \(P_A\) that represents a partially ordered set \((A, \leq)\). The vertices of the graph \(P_A\) are also elements of the semigroup \(A\), but the arc \((b, a)\) between a pair of elements means that \(a < b\) and there is no element \(c\) such that \(a < c\) and \(c < b\). End vertices are the maximal elements in the semilattice \(P_A\), and the root is the minimal element.

You can also reverse the orientation of the edges and consider the undirected graph \(S_A\) corresponding to the semigroup \(A\). Under this correspondence, as in the previous cases, the condition is satisfied: if for two semigroups the graphs corresponding to them are not isomorphic, then the semigroups are not isomorphic.

Example 2. For the semigroup \(A\) from example 1, the graphs \(P_A\) and \(S_A\) have adjacency matrices.
\[ P = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad S = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \]

Such a mapping of graphs into a semigroup was considered, for example, in [7]. The next invariant is called invariant 3. This is an ordered sequence composed by the following rule.

Let \( v \) be the vertex of the graph \( S_A \). For \( v \), we match three numbers and one sequence. We will describe them.

a) The first number \( k \) is the length of the shortest chain from the vertex \( v \) to the root.

b) The second number \( p \) is the semi-degree of the approach of the vertex \( v \) in the graph \( P_A \).

c) The third number \( r \) is the half-degree of the outcome of the vertex \( v \) in the graph \( P_A \).

d) An ascending sequence of chain lengths from vertex \( v \) to all end vertices in \( S_A \). Such a sequence will always consist of \( t \) elements, where \( t \) is the number of end vertices. If the vertex \( v \) is an end vertex, then one of the elements of the sequence will be 0.

As a result, each vertex \( v \) is associated with a sequence \((k, p, r, f_1, ..., f_t)\).

Sort the sequences mapped to the vertices of the graph lexicographically, in ascending order.

Invariant 3 is the resulting sequence:

\[(k_1, p_1, r_1, f_{11}, ..., f_{1t}; k_2, p_2, r_2, f_{21}, ..., f_{2t}; ...; k_m, p_m, r_m, f_{m1}, ..., f_{mt})\],

where \( m \) is the number of vertices of the graph, that is, the order of the semigroup \( A \).

This invariant is easily computable and distinguishes between any two nonisomorphic semigroups of order up to 7 inclusive.

Example 3. For the semigroup \( A \) from example 1, invariant 3 is equal to

\[(0,3,0,1,1,1;1,0,1,0,2,2;1,0,1,0,2,2;1,0,1,0,2,2).\]

It is clear that the second variant of mapping graphs to a semigroup is not suitable for all semigroups, but only for semilattices. Some \( S_A \) semilattices do not contain cycles, that is, they are rooted trees. In this case, special algorithms for checking tree isomorphism can be used for graphs.

2. Main results

By the probability of coincidence of the value of the invariant is meant the probability of arbitrarily choosing two different semigroups so that its value coincides. To calculate this probability for each unique value of the invariant, it is necessary to know the number of semigroups with this value. So, the probability of choosing two semigroups with a specific value of the invariant is:

\[ P = \frac{C^2_{M_k}}{C^2_N} = \frac{M_k(M_k - 1)}{N(N - 1)}. \]

Here is the total number of semigroups of some order and is the number of semigroups with a given value of the invariant. Thus, the desired probability is calculated by the formula:

\[ P = \sum_k \frac{M_k(M_k - 1)}{N(N - 1)} = \frac{\sum k M_k(M_k - 1)}{N(N - 1)}. \]

We introduce the following definition.

Definition. Invariant of graphs that recognize semigroups corresponding to these graphs with a probability of at least 0.75 in the class of semigroups of a given order will be called semigroup invariants.

In the article [9], we examined the first correspondence variant (example 1) and the following graph invariants: the number of edges, the ordered vector of degrees of vertices, the diameter of the graph, the Wiener index, and the determinant of the adjacency matrix. Of these, semigroup invariants for semigroups from 1 to 7 orders are only the ordered vector of degrees of vertices and the Wiener index.
The number of edges, the diameter of the graph and the determinant of the adjacency matrix are not semigroup invariants. Let us explain why the remaining invariants poorly recognize nonisomorphic semigroups.

The number of edges of the graph $G_A$ is $n^2$ for all semigroups of order $n$.

The diameter of a graph corresponding to a semigroup can only take two values: 0 and 1. Moreover, it takes the value 0 if and only if the following condition is met in the semigroup $(A, \cdot)$:

$$x \cdot y = x, \forall x, y \in A.$$ 

For a certain order, such a semigroup is obviously one.

For all other semigroups of this order, the diameter value will be equal to 1. We show this for any order. If the order is $n < 3$, then this statement is obvious. Let a graph corresponding to a semigroup have an edge from $a$ to $b$ and an edge from $b$ to $c$. Then

$$\exists a', b' \in A : a \cdot a' = b, b \cdot b' = c.$$ 

Due to the associativity of the operation:

$$(a \cdot a') \cdot b' = a \cdot (a' \cdot b'), \ c = a \cdot (a' \cdot b').$$

Thus, there is an edge from $a$ to $c$, and therefore the diameter cannot take a value greater than one.

The probability of choosing two semigroups with equal values of the adjacency matrix determinant of the corresponding graphs is very high because the determinant very often takes the value zero. For example, out of 836021 seventh-order semigroups of more than 780000, the determinant is zero.

It turned out that there are not many semilattices of small order. Of 836021 semigroups of order 7, only 222 are semilattices. We examined the second correspondence variant (example 2) and invariant 3. For semilattices of order 1 to 7, invariant 3 is a complete invariant. For semilattices of order 8, invariant 3 is a semigroup invariant, but there are 2 no isomorphic semilattices with the same value of this invariant.

3. Conclusion

We can conclude that the invariant 1 is more effective than the invariant 2, because for all the orders of semigroups under consideration, the probability of choosing two semigroups with the same invariant 1 is significantly lower than choosing two semigroups with the same invariant 2. Most likely, it is also better to use invariant 1 when establishing the isomorphism of semigroups of large orders. If several invariants are calculated, the order of calculations is determined in descending order of efficiency.

It turned out that the invariant 1 and the ordered vector composed of semi-powers of vertices of the graph corresponding to the semigroup are equivalent invariants. This is because the number of edges originating from each vertex is equal to the order of the semigroup, and the number of edges entering each vertex is equal to the number of times this vertex occurs in the Cayley table.

Despite the fact that there are more different values of the adjacency matrix determinant than there are different values of the Wiener index or invariant 2, the probability of choosing two semigroups with equal values of the adjacency matrix determinant of the corresponding graphs is significantly higher. The part of the invariants effective for graphs has no meaning for semigroups; it is, for example, the number of edges and the diameter of the corresponding graph.

The present work assumes such possible continuations. First, we intend to consider in detail, as an invariant, a vector of second-order degrees, which is defined and partially investigated in papers [5, 8]. Secondly, we are going to replace the definition of the invariant that we considered above based on an undirected graph with a similar definition associated with a directed graph. Perhaps we will consider other invariants, but, in fact, future work is not at all limited to considering more invariants and comparing them. We also intend to apply the obtained results to the study of special cases (subclasses) of graphs and directed graphs (for example, tournaments), for which, based on the material presented
by us, we can describe effective algorithms for generating all the corresponding objects and calculating their number.

References

[1] Melnikov B F, Radionov A N, Moseev A V, and Melnikova, E A 2006 Some specific heuristics for situation clustering problems Proc. Int. Conf. ICSoft 2006 - 1st International Conference on Software and Data Technologies pp 5181-4

[2] Melnikov B F, Melnikova E A, Pivneva S V, Dudnikov V A, and Davydova E V 2018 Geometric and game approaches for some discrete optimization problems Proc. Int. Conf. Information Technology and Nanotechnology pp 312-21

[3] Melnikov B F and Romanov N V 2001 Once again about heuristics for the problem of the travel vehicle Theoretical problems of computer science and its applications 4 81–8

[4] Melnikov B F and Panin A G 2012 Parallel implementation of the multi-heuristic approach in the problem of comparison of genetic sequences Vector of science of Tolyatti state University 4(22) 83-6

[5] Melnikova E A and Saifullina E F 2013 Approach to the verification of graph isomorphism by constructing invariants Vector of science of Tolyatti state University 1(23) 113-20

[6] Korabelshchikova S Yu, Chesnokov A I, and Butin K P 2015 On a single algorithm for counting root trees with a given number of leaves and branching order Heuristic algorithms and distributed computing 1(2) 33-44

[7] Zyablitseva L V and Pestov S A 2016 Application of algorithms for checking graph isomorphism in the theory of semigroups Bulletin of the Northern (Arctic) Federal University 4 69-74

[8] Melnikov B F and Saifullina E F 2013 Application of a multi-heurist approach for random generation of a graph with a preset vector of degrees Vector of science of Tolyatti state University 3(27) 70-83

[9] Zyablitseva L V and Vasyukov A O Comparison of the effectiveness of initial invariants when checking the isomorphism of graphs and semigroups 2017 Lomonosov readings in the Altai: fundamental problems of science and education pp 372-7