UNICIENCIA
Vol. 2. Nº. 1.
Enero-junio de 1985.
Divulgación científica

# AN EQUIVALENT FORMULATION FOR THE GRAPH ISOMORPHISM PROBLEM

*Osvaldo Skliar*
*Víctor Medina*
*Tatiana Láscaris*
Universidad Nacional.
Heredia. Costa Rica.

A classical problem in graph theory is to find an invariant that characterizes the set of graphs isomorphic to a given one. "No decent complete set of invariants for a graph is known." [ 1, p. 11 ]. In this note we discuss, in an intuitive manner, a solution to this problem that is not well–known in the scientific community, probably because it was published in Spanish [ 3 ].

A *directed graph* can be defined as a triple G= (V; E; $\varphi$), where V and E are finite sets, and $\varphi$ is a one to one function that assigns to every element of E an ordered pair of elements of V. Geometrically, W is the set of *edges (arrows)*, V the set of *vertices (nodes)*, and the map $\varphi$ assigns to every edge its vertices according to some orientation of the edge. Thus, for the directed graph shown in figure 1,

V=[1, 2, 3, 4, 5], E=[a, b, c, d, e, f, g, h, i],
and $\varphi$ (a)=(1, 2),

$\varphi$(b)=(2, 1), $\varphi$(c)=(2, 2), $\varphi$(d)=(3, 2), $\varphi$(e)=(2, 4), $\varphi$(f)=(3, 4), $\varphi$(g)=(5, 3),

$\varphi$(h)=(3, 5), $\varphi$(i)=(5, 5).

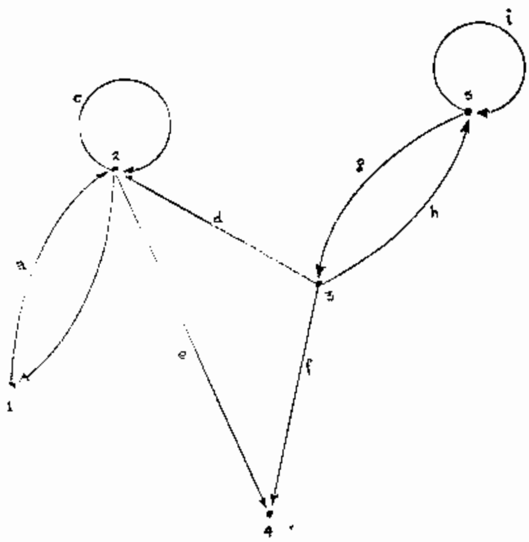


Figure 1

A graph G can also be represented by a matrix, called the *adjacency matrix*. To do this, we number the graph's vertices starting from l, say V = [ 1, 2, . . ., n ], and define

$$a_{ij} = \begin{cases} 1, \text{ if there is an arrow from vertex i to vertex j.} \\ 0, \text{ if there is no such an arrow.} \end{cases}$$

Thus, the adjacency matrix corresponding to the graph shown in figure l is

$$\begin{array}{c c} & \begin{array}{ccccc} V_1 & V_2 & V_3 & V_4 & V_5 \end{array} \\ \begin{array}{c} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \end{array} & \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} \end{array} \qquad (1)$$

It is clear that the adjacency matrix of a given graph depends on the way we number the graph's vertices (*vertex designation*); for example, if we number the vertices of the graph shown in figure l as in figure 2, then we obtain the matrix

$$\begin{array}{c c} & \begin{array}{ccccc} V_1 & V_2 & V_3 & V_4 & V_5 \end{array} \\ \begin{array}{c} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \end{array} & \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{pmatrix} \end{array} \qquad (2)$$

Two graphs $G_1$ and $G_2$ are *isomorphic* if there exists a permutation of the vertex designation of one of them such that, when we redesignate its vertices according to this permutation, its new adjacency matrix coincides with the adjacency matrix of the other graph. For example, the graphs $G_1$ and $G_2$ of figure 1 and 3, respectively, whose adjacency matrices are the matrices (1) and (2), are isomorphic. In fact if we redesignate the vertices of $G_1$ according to the permutation

$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 5 & 2 & 4 \end{pmatrix}$, then its adjacency matrix is (2).

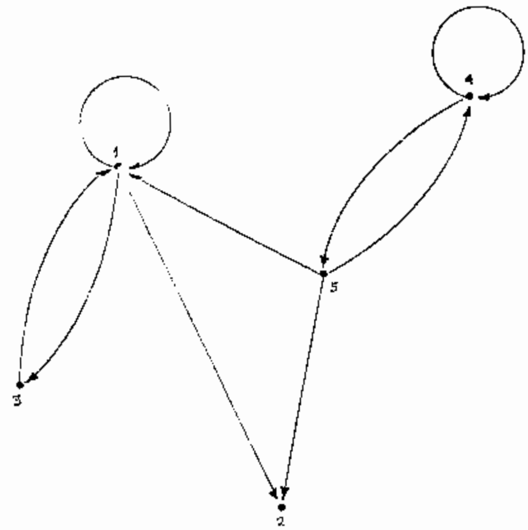


Figure 2

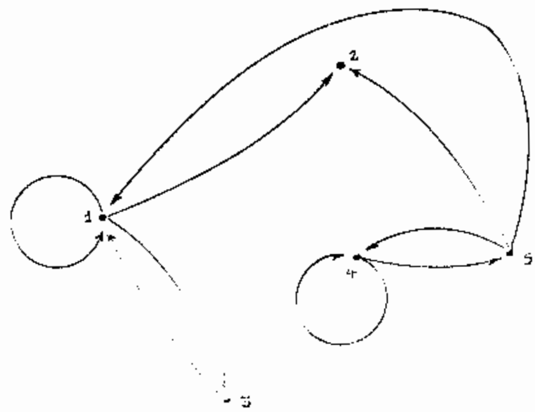


Figure 3

The problem that we are dealing with can be phrased as follows: Look for an algorithm by means of which we can obtain for any given graph an "entity" (*the invariant*) that remains unchanged with respect to any variation on the graph's vertex designation. In other words, to a given pair of graphs, such an algorithm ought to assign the same invariant to each graph if and only if these graphs are isomorphic.

Some authors have called this problem the "coding problem". Read and Corneil [ 2 ] characterized the problem as "that of finding a good algorithm for determining whether two given graphs are isomorphic. A closely related problem is the 'coding problem' – that of finding a good algorithm for assigning to every graph a 'code', i.e., a string of symbols, in such a way that two graphs are assigned the same code if and only if they are isomorphic."

We proceed to construct one algorithm of the type mentioned, utilizing the graph $G_1$ shown in figure 1, whose adjacency matrix is the matrix (1). We define the *characteristic function* of the graph $G_1$ as follows: Assign to the vertices 1, 2, 3, 4, 5 the real variables $X_1, X_2, X_3, X_4, X_5$; now if the i–th row of the adjacency matrix of $G_1$ has ones in the columns $j_1, \ldots j_r$, we assign to this row the function

$X_i^{X_{j_1}.X_{j_2}\ldots X_{j_r}}$. In our case, the first row of matrix (1) corresponds to the function $X_1^{X_2}$;

the second to $X_2^{X_1X_2X_4}$; the third to $X_3^{X_2X_4X_5}$;

and the fifth to $X_5^{X_3X_5}$. In case some row has no ones in it, we assign to this row its corresponding variable with exponent one; thus, the fourth row of matrix (1) corresponds to the function $X_4^1 = X_4$.

The characteristic function of the graph G is the product of the functions assigned to each of the rows of its adjacency matrix, that is:

$$X_1^{X_2}.X_2^{X_1X_2X_4}.X_3^{X_2X_4X_5}.X_4.X_5^{X_3X_5}$$

Finally, we restrict our variables to satisfy the following condition:

$$\begin{aligned}&(X_1-2)^2.(X_2-2)^2.(X_3-2)^2.(X_4-2)^2.(X_5-2)^2 +\\&(X_1-3)^2.(X_2-3)^2.(X_3-3)^2.(X_4-3)^2.(X_5-3)^2 +\\&(X_1-5)^2.(X_2-5)^2.(X_3-5)^2.(X_4-5)^2.(X_5-5)^2 +\\&(X_1-7)^2.(X_2-7)^2.(X_3-7)^2.(X_4-7)^2.(X_5-7)^2 +\\&(X_1-11)^2(X_2-11)^2(X_3-11)^2(X_4-11)^2(X_5-11)^2 = 0\end{aligned} \qquad (3)$$

We also could consider $\chi = [X_1, X_2, X_3, X_4, X_5]$ as a subset of the natural numbers, and in this case condition (3) can be replaced by

$$X_1.X_2.X_3.X_4.X_5 - 2.3.5.7.11 = 0 \qquad (4)$$

It is easy to see that the essential fact here is that the variables take values on the set of prime numbers, and that different variables must take different values.

In general, the characteristic function of a directed graph with n vertices can be written as

$$\mathcal{C}_G = \prod_{i=1}^{n} X_i^{\left(\prod_{j=1}^{n} X_j^{a_{ij}}\right)}$$

where $a_{ij}$ is the (i, j) –element of the adjacency matrix. In case we consider $\chi=[X_1, X_2, \ldots, X_n]$ as a subset of the real number, the restriction takes the form

$$\sum_{p\in\Phi} \prod_{j=1}^{n} (X_j - p)^2 = 0$$

where $\Phi = [2,3, \ldots, p_n]$ is the set of the first n prime numbers. If $\chi$ is considered as a subset of the natural numbers, then the restriction corresponds to the equation:

$$\prod_{j=1}^{n} X_j - \prod_{p\in\Phi} p = 0$$

The invariant we are looking for is the maximum of the characteristic function subject to the corresponding restriction. The proof of this invariance is explained in 3, but it is so direct and simple that the reader can find it for himself. In our example, this invariant is

$$\mathrm{Max}\,\mathcal{C}_{G_1} = \mathrm{Max}\ X_1^{X_2}.X_2^{X_1X_2X_3}.X_3^{X_2X_4X_5}.X_4.X_5^{X_3X_5};$$

subject to (3) if the variables take real values, or (4) if they take values in the set of natural numbers, respectively.

If we had started with the graph $G_2$ instead of $G_1$, we should get as invariant

$$\mathrm{Max}\,\mathcal{C}_{G_2} = \mathrm{Max}\ X_1^{X_1X_2X_3}.X_2.X_3^{X_1}.X_4^{X_4X_5}.X_5^{X_1X_2X_4};$$

subject to (3) or (4); and as it was pointed out above, $\mathrm{Max}\,\mathcal{C}_{G_1} = \mathrm{Max}\,\mathcal{C}_{G_2}$.

Summarizing, the maxima of the characteristic functions of two graphs agree if and only if these graphs are isomorphic. This invariant is not the only one which characterizes the set of graphs isomorphic to a given one. The minimum of the characteristic function is also an invariant in the same sense, meaning that the minima of the characteristic functions of two given graphs agree if and only if the graphs are isomorphic. It is abvious that for any *completely symmetric* graph $G_s$, i.e. for any graph whose adjacency matrix remains invariant under any permutation of the vertex designation, we should have

$$\mathrm{Max}\,\mathcal{C}_{G_s} = \mathrm{Min}\,\mathcal{C}_{G_s}.$$

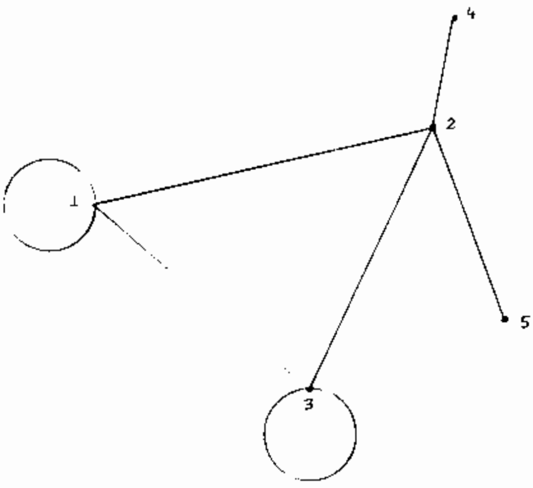


Figure 4

The characteristic function is not the only tool to construct invariants for the set of graphs isomorphic to a given one. As an example of a different method, consider the $n^2$ elements of the adjacency matrix of a given directed graph. Write them in a row according to some previously established order on the entries of the matrix; the result will be a binary number. Now to each of the n! permutations of the vertex designation corresponds one well-defined binary number. This set of n! numbers has a maximum and a minimum, and each one of them is an invariant of the type we are looking for.

Following these ideas the reader may be able to find other invariants to characterize the set of graphs isomorphic to a given one.

Aside from the great theoretical interest of finding invariants of the type mentioned, the effective computation of them via a polynomial-bound algorithm would constitue one way to solve the graph isomorphism problem: Find an efficient method to establish whether two given graphs are isomorphic. We shall treat these problems in a forthcoming article.

The solution of the graph isomorphism problem has not only academic interest, but also practical importance in such areas as: a) the establishment of a unified nomenclature in organic chemistry, b) the applicability of the analysis done one certain systems (e.g. electronic circuits) to others of the same type, and c) measuring the degree of complexity of a variety of systems.

Finally we note that the methods which we have applied to directed graphs can also be extended in a natural way to other kinds of graphs, namely non-directed graphs, weighted graphs, and labeled graphs. As an example we consider the case of the non-directed graphs. Change each are which joins distinct vertices to a pair of arrows in opposite directions; change each closed are to a single arrow. Thus, we obtain a directed graph such that any invariant (under vertex designation) for this graph is also an invariant of the same type for the non-directed graph. Following this criterion, the non-directed graph shown in figure 4 can be considered as the directed one of figure 5.

When we are working in a "universe" of non-directed graphs, there is no problem in considering as an invariant of a given graph the invariant of the corresponding directed graph.

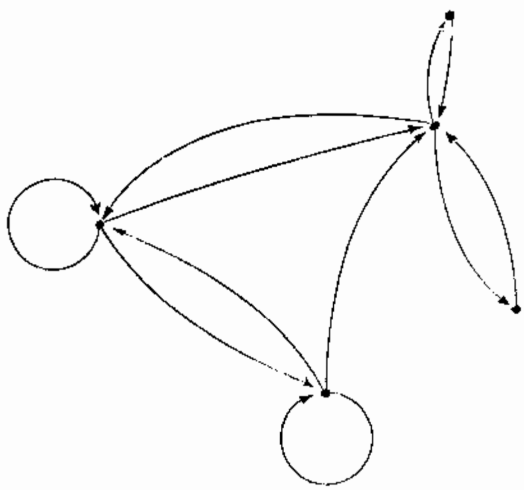

Figure 5

## REFERENCES


1. Harary, F. 1969. Graph Theory. 1st ed. Addison-Wesley. Reading. Massachusetts.

2. Read, R. C. and D. G. Corneil. 1977. The graph isomorphism disease. J. Graph Theory. 1. 339-363.

3. Skliar, O. and A. M. Haeberer. 1976. Designación, isomorfismo y complejidad estructural de grafos. Revista del Instituto de Cibernética de la Sociedad Científica Argentina. 1 17-28.