Ring Perception: Proof of a Formula Calculating the Number of the Smallest Rings in Connected Graphs

Michel Petitjean, Bo Tao Fan,* Annick Panaye, and Jean-Pierre Doucet

Institut de Topologie et de Dynamique des Systèmes, CNRS UPRES-A 7086, Université Paris 7, 1 rue Guy de la Brosse, 75005 Paris, France

Received January 30, 2000

A general mathematical proof of a formula proposed and used by Fan *et al.* for calculating the number of the smallest rings in their smallest set of the smallest rings search algorithm is reported. This proof generalizes this formula to all connected cyclic graphs.

The enumeration of rings in a given structure plays an important role for any chemical computer system. Among the different categories classified by Downs *et al.*,[1] the smallest set of the smallest rings (SSSR) is often the preferred type of ring set for chemists, because it is the basic ring set, and all other types of ring sets can be generated from SSSR.

The calculation of the number of the smallest rings in a SSSR can be used to control and check the search process. The Cauchy formula is generally used to calculate the set cardinality n_R in a graph:

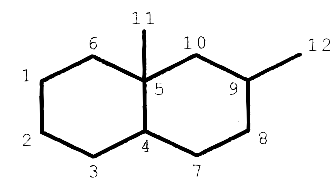
$$n_R = n(\text{edges}) - n(\text{vertices}) + n(\text{components}) \quad (1)$$

In the case where the graph contains only a component, eq 1 can be reduced to

$$n_R = n(\text{edges}) - n(\text{vertices}) + 1 \quad (2)$$

In the algorithm developed by Fan *et al.*,[2] a formula has been proposed as an alternative method to evaluate the number of rings in a complete cyclic system in which all nodes are involved in at least one cycle.

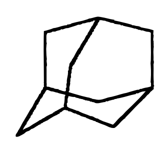
Before introducing the proposed formula, we recall some terms and definitions used in the Fan algorithm: (a) A *ring* is represented as R(n_1,n_2,...), where n_i are atom indices in a molecule. A SSSR is noted as S(m_1,m_2,...), where m_i are the sizes of the rings. (b) Each atom is a *node*. (c) The first selected node is a *root*. A *closed path* is a walk that starts and ends at a root. (d) The *connectivity* of a node is defined as the number of its links with other nodes. (e) A *tree* is a structure derived from a root. (f) If a node has only unit connectivity, it is called *terminal*. (g) A *block* is a group of atoms such that all links between them are involved in one or more rings. (h) The *real ring connectivity* of a node is the number of links that are the edges of one or more rings in a block. It is denoted as N_m (m = 2, 3, 4, ...). In the following graph, for instance, the connectivity for node 5 is 4, and for node 9 it is 3, but their real cyclic connectivities are 3 and 2, respectively.



The proposed alternative method is described by the following formula:

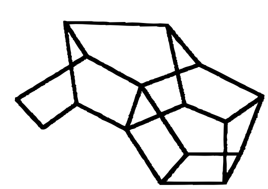
$$n_R = 1 + \frac{n(N_3)}{2} + n(N_4) \quad (3)$$

where n_R is the number of the smallest rings to be found in the cyclic block, N_i means the node with real ring connectivity i, and $n(N_i)$ represents the number of nodes with connectivity i. This formula was proposed on the basis of the fact that generally the rings of cyclic compounds consist of carbon, nitrogen, and sulfur atoms in chemistry. The valences of these atoms are less than or equal to 4. For example, in adamantane, three smallest rings should be found according to the formula



$$n_R = 1 + \frac{4}{2} = 3$$

In the following N_4 containing structure, $n(N_3) = 12$, $n(N_4) = 2$, and 9 smallest rings should be found.



$$n_R = 1 + \frac{12}{2} + 2 = 9$$

Equation 1 was extended later to the cyclic structures with

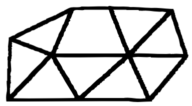
* Corresponding author. Phone: 33-1-4427 4412. Fax: 33-1-4427 6814. E-mail: fan@paris7.jussieu.fr.

10.1021/ci000004n CCC: $19.00
Published on Web 05/23/2000

all possible connectivities[3]

$$n_R = 1 + \frac{1}{2}n(N_3) + \frac{2}{2}n(N_4) + \frac{3}{2}n(N_5) + ... = 1 + \sum_{i=0} \frac{i}{2} n(N_{i+2}) \quad (4)$$

An example is given below. The following structure contains 7 N_3, 1 N_4, 1 N_5, and 1 N_6. The number of the smallest rings is evaluated as 9.



$$n_R = 1 + \frac{1}{2} \times 7 + \frac{2}{2} \times 1 + \frac{3}{2} \times 1 + \frac{4}{2} \times 1 = 9$$

Equation 4 indicates that the number of rings can be evaluated by counting the number of nodes with different connectivities. We need not count the number of edges and the number of all nodes as requested in the Cauchy formula. Now the question is whether this formula can be really generalized for all cyclic structures. In this paper, we give the mathematical proof for showing that this formula is universal.

PROOF

The graph is connected; it has only one component. Multiple edges are allowed, but self-loops are forbidden. For node j, its real cyclic connectivity is denoted as N_j, which represents in fact the number of edges of j involved in cycles, or alternatively the number of cyclic edges.

Equation 4 can be rewritten as

$$2n_R = 2 + \sum_{i=2} (i-2)n(N_i) \quad (5)$$

It is true for any monocyclic graph. Any connected cyclic graph can be generated starting from a monocyclic graph.

Adding one edge between two nodes will create one new cycle. Now we try to deduce this property from eq 5 to prove its generality. Adding one edge to the graph means that two nodes have their number of cyclic edges $j1$ and $j2$ increased by 1, because a new cycle is created. The sum of the terms $(i-2)n(N_i)$ is changed only for $i = j1, j2, j1+1$, and $j2+1$. The changes can be summarized as

$$\begin{array}{cccc} n(N_{j1}) & n(N_{j2}) & n(N_{j1+1}) & n(N_{j2+1}) \\ \downarrow & \downarrow & \downarrow & \downarrow \\ n(N_{j1})-1 & n(N_{j2})-1 & n(N_{j1+1})+1 & n(N_{j2+1})+1 \end{array}$$

and $j1 \to j1+1$ and $j2 \to j2+1$. Thus, $2n_R$ is increased by the following quantity:

$$\Delta(2n_R) = (j1-2)\{n(N_{j1})-1\} + (j2-2)\{n(N_{j2})-1\} + (j1+1-2)\{n(N_{j1+1})+1\} + (j2+1-2)\{n(N_{j2+1})+1\} - (j1-2)\{n(N_{j1})\} - (j2-2)\{n(N_{j2})\} - (j1+2)\{n(N_{j1+1})\} - (j2+1-2)\{n(N_{j2+1})\} = 2$$

meaning that n_R is increased by 1. In a particular case where $j1 = j2 = j$, the sum of the terms $(j-2)n(N_j)$ is changed only for j and $j+1$. We have

$$n(N_j) \to n(N_j) - 2, \quad n(N_{j+1}) \to n(N_{j+1}) + 2, \quad \text{and} \quad j \to j+1$$

Then $2n_R$ is increased by the following quantity:

$$\Delta(2n_R) = (j-2)\{n(N_j) - 2\} + (j+1-2)\{n(N_{j+1}) + 2\} - (j-2)\{n(N_j)\} - (j+1-2)\{n(N_{j+1})\}$$

$$\Delta(2n_R) = 2$$

meaning that n_R is also increased by 1. The formula stands therefore for any connected cyclic graph.

As pointed out by a reviewer, the use of eq 4 is limited to connected cyclic graphs, a particular class of graphs; consequently its proof should be obtained on the basis of the Cauchy formula. Another reviewer gives such an alternative proof from the Cauchy formula as follows: Assume $n(N_1) = 0$ because all nodes are involved in at least one cycle; i.e., there is no terminal. Then

$$n(\text{vertices}) = n(N_2) + n(N_3) + n(N_4) + n(N_5) + ... \quad (6)$$

$$n(\text{edges}) = \text{half the sum of the degrees of all nodes}$$

i.e.

$$n(\text{edges}) = \frac{1}{2}\{2n(N_2) + 3n(N_3) + 4n(N_4) + 5n(N_5) + ...\} \quad (7)$$

Introducing eqs 6 and 7 into eq 2 gives

$$n_R = n(\text{edges}) - n(\text{vertices}) + 1$$

$$= 1 + \frac{1}{2}n(N_3) + \frac{2}{2}n(N_4) + \frac{3}{2}n(N_5) + ... = 1 + \sum_{i=0} \frac{i}{2} n(N_{i+2})$$

CONCLUSION

The proposed formula in the calculation of the number of the smallest rings is shown to be universal for all connected cyclic graphs; it can be used as an alternative method of the Cauchy formula. In this formula, we need only count the numbers of nodes with a connectivity more than 2 for evaluating the number of independent rings. It constitutes an important control tool in the SSSR search process. If the smallest rings found by the process exceed the calculated number, it can be concluded that errors occur in the process. Of course, this formula can evaluate only the number of rings; it is not capable of giving the size of each ring.

ACKNOWLEDGMENT

We thank a reviewer for providing the alternative proof based on the Cauchy formula.

REFERENCES AND NOTES

(1) Downs, G. M.; Gillet, V. J.; Holliday, J. D.; Lynch, M. F. Review of ring perception algorithms for chemical graphs. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 172−187.

(2) Fan, B. T.; Panaye, A.; Doucet, J. P.; Barbu, A. Ring perception: a new algorithm for finding directly the SSSR from connection table. *J. Chem. Inf. Comput. Sci*. **1993**, *33* 657−662.
(3) Fan, B. T.; Panaye, A.; Barbu, A.; Doucet, J. P. Ring perception. Application of elimination technique to the SSSR search from a connection table. In *AIP Conference Proceedings 330, E.C.C.C.1, Computational chemistry*; Bernardi, F., Rivail, J. L., Eds.; American Institute of Physics: Woodbury, NY, 1995; pp 538−543.

CI000004N