ENTROPY AND THE COMPLEXITY OF GRAPHS: I.
AN INDEX OF THE RELATIVE COMPLEXITY OF A GRAPH

Abbe Mowshowitz
Mental Health Research Institute,
The University of Michigan,
Ann Arbor, Michigan

The structural information content (Rashevsky, 1955; Trucco 1956a, b) \( I_g(X) \) of a graph \( X \) is defined as the entropy of the finite probability scheme constructed from the orbits of its automorphism group \( G(X) \). The behavior of \( I_g \) on various graph operations—complement, sum, join, cartesian product and composition, is examined. The principal result of the paper is the characterization of a class of graph product operations on which \( I_g \) is semi-additive. That is to say, conditions are found for binary operations \( \circ \) and \( \vee \) defined on graphs and groups, respectively, which are sufficient to insure that \( I_g(X \circ Y) = I_g(X) + I_g(Y) - H_{XY} \), where \( H_{XY} \) is a certain conditional entropy defined relative to the orbits of \( G(X \circ Y) \) and \( G(X) \vee G(Y) \).

1. Introduction. The present work is addressed to the problem of measuring the relative complexity of graphs. The central idea in our approach to this problem can be characterized as follows. Let \( X \) be a finite graph given by a set of \( V = V(X) \) of \( n \) vertices together with a set \( E = E(X) \) of edges (or adjacencies). If \( \{V_i\}_{i=1}^h \) is a decomposition of \( V \) into equivalence classes \( V_i \) containing \( n_i \) vertices (\( 1 \leq i \leq h \)), then a finite probability scheme can be constructed by assigning the probability \( p_i = n_i/n \) to \( V_i \) for each \( i = 1, 2, \ldots, h \). The entropy of a finite probability scheme associated with a graph in this manner can then be viewed as a measure of the complexity of the graph relative to the given decomposition of its vertex set.

There are, of course, many ways of obtaining a decomposition of the set of vertices of a graph. In this and subsequent papers (Mowshowitz 1968a, b, c),
we will be concerned with two decompositions, both of which have been the object of considerable study by graph theorists. The first, which will command most of our attention, is given by the set of orbits of the automorphism group of a graph. The second is called a chromatic decomposition and is determined by a coloring of a graph.

So far as this author is aware, the idea of defining an entropy measure on graphs first arose in a paper of Rashevsky (1955), dealing with the complexity of organic molecules. Rashevsky made use of the fact that a structural formula of organic chemistry can be regarded as a graph whose points represent physically indistinguishable atoms and whose edges represent chemical bonds. With this interpretation, one can consider the topological properties of a molecule in a graph theoretic setting. Of particular interest here is the automorphism group of a graph.

An automorphism of a graph $X$ is a one-one mapping of $V(X)$ onto itself which preserves adjacency. The set of all automorphisms of $X$ forms a group, the orbits of which constitute a decomposition of $V(X)$. Suppose $X$ is a graph corresponding to the structural formula of a molecule $M$. If points $a$ and $b$ are in the same orbit of the group of $X$, it is clear that when the atoms corresponding to $a$ and $b$ are interchanged, the resulting molecule has the same structure as $M$. So, if $X$ has $n$ points and the orbits of the group of $X$ contain $n_i$ elements for $i = 1, 2, \ldots, h$, the ratio $n_i/n$ can be interpreted as the probability that a particular atom of $M$ (in a set of topologically equivalent atoms) will be involved in a given chemical reaction. If the (topological) complexity of a molecule $M$ is viewed in terms of the freedom with which it can interact with other molecules, Rashevsky's topological information content (i.e. the entropy of the finite probability scheme constructed from the orbits of the group of $X$) is appropriate as a measure of complexity, since it depends both on the number of equivalence classes and on their respective cardinalities.

Using Rashevsky's measure, Karreman (1955) examined the change in information content produced by chemical reactions (which in this context correspond to simple operations on graphs). As one might expect, it was found that information is sometimes lost and sometimes gained in such reactions. We will be concerned with the general analogue of this problem when we examine the behavior of the information measure on various graph products.

One can easily imagine other contexts in which an entropy measure defined on graphs can serve as a measure of complexity. Of course, the way such a measure is defined (i.e. the underlying decomposition) will depend on the particular structural characteristics of interest in that context. For example, Rashevsky's topological information content might be useful in the study of friendship relations in certain populations. On the other hand, it is not clear