Meeting of the Aristotelian Society at 21, Bedford Square, London, W.C.1, on 14th December, 1959, at 7.30 p.m.

V — ON DEFINING ANALOGY*

By MARY B. HESSE

I

IT is generally held that the rôle of analogies and models in the construction of scientific theories has to be regarded as a topic in the psychology of discovery, and not as a subject for logical analysis. This state of affairs seems to arise more from the absence of any satisfactory analysis of the analogy-concept itself than from any strong convictions that its rôle in science is unimportant. Indeed it may be that attempts to explain it away are made so persistently just because it appears impossible to explain it. I shall not attempt here to give an analysis which is adequate to deal with these problems in the philosophy of science, but shall merely investigate in a preliminary way a definition of a four-term analogy-relation which has some of the desirable formal properties of analogy, and from which an algorithm can be derived for finding a fourth term (which need not be unique) given the other three.

Any attempt to discuss the properties of an analogy-relation must begin by facing two opposite objections to the whole notion of analogy: the first on the ground that it is merely a trivial relation holding between certain pairs of terms, and the second on the ground that the relation exists between any two pairs of terms whatever, and is therefore useless.

The first objection has recently been made by Richard Robinson. He remarks that analogy in any sense other than mathematical proportionality " is merely the fact that some relations have more than one example,"[1] and regards the assertion of an analogy " *a* is to *b* as *c* is to *d* " as merely equivalent to

* I should like to express my gratitude to Professor R. B. Braithwaite for helpful discussions of Sections I to IV of this paper. Any defects which may remain in the paper are entirely my own responsibility.

[1] Comments on J. F. Anderson: " Some Basic Propositions Concerning Metaphysical Analogy ", *Review of Metaphysics*, V, 1952, p. 466.

asserting the existence of an R such that *a*R*b* and *c*R*d*. This, however, is hardly an adequate analysis of the analogy-concept as traditionally understood, for the assertion in the case of an analogy of a relation between *a* and *b*, and a relation between *c* and *d*, was always qualified by the remark that these two relations are not *identical*, but only *similar* in some relevant respect: sight is not related to the eye *exactly* as understanding to the mind, but only in the way that something can be related to the eye, a way which is similar to the way that something else can be related to the mind. Thus it seems that we should make the analogy equivalent to the existence of two *similar* relations R, R′, such that *a*R*b* and *c*R′*d*.

But as soon as we ask what constitutes an analogy-generating similarity of relations we meet with the second objection, for in the usual sense of similar, it seems to be possible to find relations similar *in some respect* between any two pairs of terms whatever. The definition of ordinal similarity between the relations R and R′, for example, requires a one-one relation S which correlates *a* with *c* and *b* with *d*, and so on throughout the fields of R and R′. But without further specification of S this does not help us with the present problem, because any possible method of correlating *a* with *c* and *b* with *d* would be a possible S, and thus, by choosing an appropriate S with a little ingenuity, any two pairs of terms could be shown to be related by analogy.

In order to proceed it is therefore necessary to make some further specification of R and R′ and of S. In classical discussions of analogy, this seems to have been provided by the ontology presupposed. For example, it might be assumed that both fields of relata of R and R′ were independently ordered by a *scala natura*, making possible a " proportion " between pairs of relata of R and corresponding relata of R′. Another method was to assume a common relation holding between *a* and *b* and between *c* and *d*, for example a cause-effect or a being-becoming relation. It was usually assumed in these cases that the common relation was a generalised form of the relation as usually understood; that is, it was not assumed that, for example, God causes the world in exactly the same sense as a sculptor causes a statue, but that some generic likeness could be seen between these two examples of cause-effect, and yet that this likeness was not so general as to hold of any pair of terms whatever.

Our present problem is to see whether anything can replace such ontologies in the specification of the relations in virtue of which the analogy holds. In what follows I shall try, first, to improve upon the rather simple relations presupposed by the classical analogies, by showing that a more general analogy-relation can be defined if its field has the structure of an algebraic lattice, and secondly, I shall show that this is the structure of at least one field in which analogy has been traditionally applied, namely that of a classification in terms of properties. Then I shall consider whether a classification of the words of a natural language can be found which gives rise to the required structure, so that an analogy-relation can be defined between words in virtue of their uses in discourse.

The use of the algebra of lattices in this definition of analogy was suggested to me by the work of the Cambridge Language Research Unit, where lattice-theory is being used to develop a structure for the semantic content of language.[2] The suggestion that such semantic content can be given mathematical expression, and algorithms developed which will provide determinate answers to questions about meaning in languages, is a revolutionary one. It is, however, also a very powerful one, for it opens up the possibility that all those operations with language which have hitherto been regarded as " intuitive " and unamenable to rational description, may after all be explicable in terms of a general theory of language.

It is not to be expected that such a suggestion can immediately be expressed in a form which satisfies all the criteria for a testable empirical theory, and it is still at present in the speculative stage. The Cambridge Language Research Unit, which developed it, has, however, applied it to problems of retrieving documents from libraries in response to requests; to mechanical translation; and to mechanical abstracting. In all of these applications,

[2] " Potentialities of a Mechanical Thesaurus ", Margaret Masterman, with Appendix by A. F. Parker-Rhodes, read at the *International Conference on Scientific Information*, October, 1956; " The Analogy between Mechanical Translation and Library Retrieval ", Margaret Masterman, R. M. Needham, and K. I. B. Spärck-Jones, *Proceedings of the I.C.S.I.*, Washington, 1958; " The Thesaurus Approach to Information Retrieval ", T. Joyce and R. M. Needham, *American Documentation*, IX, July, 1958; " Essays on and in Machine Translation ", Report by the *Cambridge Language Research Unit*, 1959, unpublished.

the problem is to transform input information into a desired output by making use of the structure of the system in question: in the first case, that of a library classification system, and in the second and third cases, that of the semantic structure of language. In the course of this work, a library classification system has been constructed having the structure of a lattice, so that there is one application which rigorously exemplifies lattice-structure. Moreover, the existence of synonym-dictionaries, or thesauri, as for instance Roget's Thesaurus, whose structure is a crude approximation to lattice-structure, suggests that this is also the required structure of language, and by making this assumption, some satisfactory algorithms for effecting translations and abstracts have been developed.

I shall here consider the problem of determining analogies as that of finding another such lattice-algorithm, for this problem is another in which, as we have seen, more orthodox logical treatment fails, and which has therefore been abandoned to the sphere of intuition or guess-work. The definition of analogy suggested here, however, will give a lattice-algorithm which enables the fourth term of a four-term analogy to be calculated, and the results of this algorithm, like those of the translation- and abstracting-algorithms, may then be used to provide tests of the suggested lattice-structure of language, where this structure is regarded as an empirical hypothesis.

II

I shall first introduce the axioms of a *partially-ordered system* by considering the structure of a certain kind of classification. We shall suppose that there are in the world a number (not necessarily finite) of existing individuals each having a finite number of distinct properties which are such that:

(i) It can be empirically decided whether a specified individual has a specified property or not.

(ii) The possession of a property by any individual does not entail the possession of any other, that is, the properties are mutually independent. Thus, " green " and " coloured " would not both be properties in this sense, since " being green " entails " being coloured " and so they are not independent, but " green " and " hard ", for example, are independent.

(iii) The total number of properties in the world is finite.

(iv) Of any two properties, at least one is possessed by some individual which does not possess the other.

Then classes of properties[3] are defined by the following two postulates:

C(1) Any set of properties which together hold of some individual, and between them exhaust the properties of that individual, forms a class.

C(2) If two classes, x and y, contain properties $p_1 \ldots p_r$, and only these, in common, the logical product of x and y (the set $\{p_1, \ldots p_r\}$) is a class, and the logical sum of x and y is a class. (C(2) then defines more classes recursively.)

No set of properties not satisfying either C(1) or C(2) is a class.

Classes defined in this way differ from Boolean classes in the following respects (not all of which are independent):

(*a*) The class-members are *properties*, not individuals.

(*b*) The sum of two classes is not a class unless their intersection is non-empty.

(*c*) If a set of properties makes up a class, it is not necessarily the case that any sub-set of the same properties makes up a class.

(*d*) Complementary classes are not defined.

The reason why it is convenient to work in a system satisfying less restrictive conditions than the Boolean system will emerge later when the definition of an analogy-relation in the system is given.

Classes thus defined by C(1) and C(2) are members of a system partially ordered by the set-theory inclusion relation $\geq$, which is such that if $x \geq y$, the properties making up the class x include or are identical with the properties making up the class y. This relation satisfies the axioms of partially-ordered systems:

P(1): For all x, $x \geq x$,

P(2): If $x \geq y$ and $y \geq x$, then $x = y$,

P(3): If $x \geq y$ and $y \geq z$, then $x \geq z$.

---

[3] " Classes of properties " will be used here in the restricted sense which is defined by C(1) and C(2), and which is *not* the sense of Boolean classes.

That is, the partial-ordering relation is reflexive, asymmetric and transitive, but it is not a simple-ordering relation, since it is not the case that for all x and y, either $x \geq y$ or $y \geq x$.

It is now convenient, and possible without serious loss of generality, to suppose that all the classes defined in C(1) have one property in common (according to II (iv) they cannot have more than one property in common). This assumption is equivalent to requiring the initial individuals to be chosen out of the world in virtue of some common property, and we are not usually in practice interested in classifications where this is not the case. The common property, p_0, now makes up a class O (by C(2)) which is included in every class. Since every class contains p_0, the logical sum of all the classes is a class, say I, which includes every class and contains all the properties in the classification. If I is represented by the highest point of a diagram, and O by the lowest point, all other classes can be represented by points lower than I and higher than O, in such a way that of two distinct classes x and y related by inclusion, the point representing the including class is higher than that representing the included class. The inclusion-relation is then conveniently represented by a line joining the points representing x and y.[4]

When p_0 is taken as one of the properties, the classification satisfies not only the postulates of a partially-ordered system, but also those of a *distributive lattice*. The lattice postulates assert that, for any pair of classes x and y of a partially-ordered system of classes, there is a unique *join* (written $x \cup y$), and a unique *meet* (written $x \cap y$), where the join of the two classes is a class which is included in every class which includes them both, and the meet is a class which includes every class which they both include. In the classification system, C(2) implies the existence for every pair of classes of a unique join and a unique meet, for if a class x contains properties, $p_0, p_1, \ldots p_r, \ldots p_l$, and a class y contains $p_0, p_1 \ldots p_r, p_{l+1}, \ldots p_m$ $(m > l > r)$,[5] their join is the class containing $p_0, p_1, \ldots p_1, \ldots p_m$, and their meet is the class containing $p_0, p_1, \ldots p_r$. Hence this classification system forms a lattice, which we shall call the

[4] See for example, the figure on page 86, below.

[5] Here and elsewhere, where the relata of $>$ are numbers, this symbol stands for " greater than ".

C-lattice. Furthermore, the C-lattice satisfies the following necessary and sufficient condition for distributivity:

For all x, y, z, $x \cap (y \cup z) = (x \cap y) \cup (x \cap z)$.[6]

It is important to notice that an element of the C-lattice must be determined empirically. A C-lattice whose elements are classes of properties in the world does not usually contain all possible combinations of the properties concerned,[7] but only such combinations as occur in some individual, together with their joins and meets, and other joins and meets defined recursively. It may therefore happen that some properties appear in the lattice only in classes containing more than one property. This will be the case, for instance, if an individual having more than one property is the the only individual having a particular property, say p_s, for then the class x_s made up of the properties of this individual will contain p_s, and so will the classes which include x_s; but no classes (other than x_s) which x_s includes will contain p_s, because p_s is not a property common to x_s and any other class which does not include x_s. Other properties, however, may be contained in classes containing only one property other than p_0. This will happen, for example, when two individuals have just one property other than p_0, say p_r, in common, for then the meet of the pair of classes defined by the properties of the two individuals will contain just p_0 and p_r.

If it does happen that all possible combinations of properties $p_1, \ldots p_n$ are found in the world, then the C-lattice is the Boolean lattice of degree n, and this is the C-lattice with the maximum number of elements constructed out of $p_0, p_1 \ldots p_n$.

I shall now describe those mathematical characteristics of the C-lattice which are relevant to the definition of analogy.

[6] Let x contain $p_0, p_1 \ldots pr, pr_{+1}, \ldots ps, ps_{+1}, \ldots, pt,$
$pu_{+1}, \ldots pl$;
y contain $p_0, p_1, \ldots pr, pr_{+1}, \ldots ps, pt_{+1}, \ldots, pu,$
$pl_{+1}, \ldots pm$;
and z contain $p_0, p_1, \ldots pr, ps_{+1}, \ldots pt, pt_{+1}, \ldots pu,$
$pm_{+1}, \ldots pn$;
$(n > m > l > u > t > s > r > 0)$.

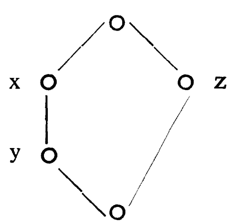
Then $(x \cap y) \cup (x \cap z) = (p_0, p_1, \ldots pr, pr_{+1}, \ldots ps) \cup (p_0, p_1, \ldots, pr,$
$ps_{+1}, \ldots pt)$
$= p_0, p_1, \ldots pr, pr_{+1}, \ldots ps, ps_{+1} \ldots pt$
$= x \cap (y \cup z)$

[7] *i.e.*, it is not a *Boolean lattice.*

(1) The elements of the lattice may be arranged in levels depending on the number of properties in each class. A level of classes each containing r properties will be said to be of *dimension* $r-1$. The dimension of a class x is written $d[x]$. Thus $d[O]=0$ and $d[I]=n$ for a lattice of $n+1$ properties.

This ordering of classes into levels implies that the similarity between two classes is represented in the lattice by the relations between the dimensions of the classes and their join. Consider, for simplicity, two classes in the same level, each containing m properties. Then if they have r of these properties in common, their join will contain $2m-r$ properties. If they are very similar, that is, if r is large compared to m, their join will be in a level not far above them, whereas if they are very dissimilar, that is, if r is small, their join will be in a level nearly twice as high as their own level. The extent of the similarity between two classes is important in defining an analogy-relation of which they are terms, and will be represented in the definition by the dimension-intervals between the classes and their join.

(2) The C-lattice satisfies the *chain condition*, namely, every chain of elements connecting I and O by means of the inclusion relation has the same number of elements.[8] Furthermore, each chain has an element in each level, as can be proved by using II(iv).



(3) It follows easily from C(2) that, for any two classes x and y, their dimensions $d[x]$ and $d[y]$ and the dimension of their meet and join are related by the equation $d[x]+d[y]=d[x \cap y] + d[x \cup y]$, which we shall call the *dimension-interval relation.* It states that any pair of classes, not necessarily in the same level, have a join and a meet in levels symmetrically above and below the levels of x and y.

[8] For if not, it would contain the five-element sub-lattice in the figure, and since in this sub-lattice $x \cap (y \cup z) = x$, and $(x \cap y) \cup (x \cap z) = y$, distributivity would not be satisfied.

Conditions (2) and (3) are implied by the distributivity of the C-lattice, but they are not sufficient for distributivity. They are in fact a set of necessary and sufficient conditions for a finite lattice to be a *modular* lattice,[9] and they have been explicitly mentioned here because they are required for the definition of the analogy-relation given below. Thus it appears that a lattice satisfying a weaker set of conditions than C(1) and C(2) will allow this analogy-relation to be defined within it, so long as it is at least a modular lattice.

III

Since we have made use of a scheme of classification in terms of properties we may expect that most of the classical examples of analogies in the Aristotelian tradition will be definable in terms of it. Aristotle's analogies properly so-called are all suggested by the form of a mathematical proportion with four terms and are what the Scholastics called "analogy of proportionality".[10] We shall represent the analogy "x_1 is to x_2 as y_1 is to y_2" by $x_1 : x_2 : : y_1 : y_2$, or by

$$\frac{x_1}{x_2} : : \frac{y_1}{y_2} .$$

Aristotle uses the analogy of proportionality, in connexion with the classification of animal species, to denote likeness of structure or function, as for example

$$\frac{\text{MAN}}{\text{FISH}} : : \frac{\text{(HUMAN) BONE}^{11}}{\text{(FISH) SPINE}} : : \frac{\text{NOSE}}{\text{GILLS}}$$

This is easily expressible in terms of the C-lattice. The class made up of properties each of which belongs to some man

[9] See G. Birkhoff: *Lattice Theory*, New York, 1948, p. 68.

[10] The Scholastics distinguished "analogy of attribution" from "analogy of proportionality", but neither Aristotle nor the main school of Thomist interpretation regarded the former as analogy in the strict sense (*cf.* H. Lyttkens: *The Analogy between God and the World*, Uppsala, 1952, pp. 206 ff.) Since the analogy of attribution involves two-place predicates (usually causal relations), it cannot be described in terms of the C-lattice and is not considered here.

[11] According to Aristotle, no common word is properly used in Greek either for human and fish "bone", or for human and fish "spine".

(which will be called the MAN-class) will include the BONE-class and the NOSE-class, and similarly the FISH-class will include the SPINE-class and the GILLS-class. The analogies require in addition some relations of likeness between the members of each pair, and these are provided by the ANIMAL-class; which includes the MAN-class and the FISH-class; by the SKELETON-class, including the BONE-class and the SPINE-class, and the ORGAN-OF-BREATHING-class, including the NOSE-class and the GILLS-class. The relations involved can be formulated in the following table :

ANIMAL-SPECIES	SKELETON	ORGAN OF BREATHING
MAN	BONE	NOSE
FISH	SPINE	GILLS

Two kinds of analogy are represented here; one in which the upper and lower classes on the left-hand side include the upper and lower classes respectively on the right-hand side, and the other in which the upper pair are both included in the MAN-class, and the lower pair in the FISH-class. In both cases the analogy relation $(x_1 : x_2 :: y_1 : y_2)$ implies that the four joins $x_1 \cup x_2$, $x_2 \cup y_2$, $y_2 \cup y_1$, $y_1 \cup x_1$ are of lower dimension than the joins $x_2 \cup y_1$, $x_1 \cup y_2$, for if they were not, the analogy $(x_1 : y_2 :: y_1 : x_2)$ would also hold, and this is in general not the case.

Although the classical texts hardly say so explicitly, it is clear that the analogy of proportionality remains valid under alternation, that is to say, if $(x_1 : x_2 :: y_1 : y_2)$, then $(x_1 : y_1 :: x_2 : y_2)$, owing to the symmetry of the classification table.

IV

These considerations do not fully characterise the analogy-relation, however. Since one of the lattice postulates is that any pair of classes has a join, it follows that there will be a great many formal analogies satisfying the conditions already mentioned which will depend on joins very high in the lattice, and in which the terms will therefore appear to have very remote similarity. Clearly we need to define *degrees of analogy* in such a way that a close analogy has small dimension-intervals between its terms and their relevant joins.

Two simple possibilities for such a definition immediately suggest themselves, namely, to make degree of analogy a function of the *greatest* dimension-interval (g.d.i.) between any of x_1, x_2, y_2, y_1, and any of $x_1 \cup x_2$, $x_2 \cup y_2$, $y_2 \cup y_1$, $y_1 \cup x_1$, or to make it a function of the *average* dimension-interval (a.d.i.) between the same elements. Since the maximum dimension-interval in the C-lattice is n, degree of analogy is conveniently represented as the proper fraction $\frac{(n-k)}{n}$, where k is the g.d.i. in the case of the first definition, or the a.d.i. in the case of the second. Both definitions satisfy the requirement that a close analogy, with degree near to 1, is one in which the terms and their joins are close together in the lattice in the sense of having small dimension-intervals. The first definition, however, has some undesirable consequences.

Let us suppose that the join of the (ANIMAL)-BONE-class and the (ANIMAL)-HEART-class is the ANIMAL-class, and that the join of the CHASSIS-class and the PETROL-PUMP-class is the (MECHANICAL)-VEHICLE-class. (The petrol-pump referred to is the one under the bonnet, not the one from which petrol is obtained at a garage.) These classes may constitute a fairly " good " analogy:

$$\frac{\text{BONE}}{\text{HEART}} :: \frac{\text{CHASSIS}}{\text{PETROL PUMP}}$$

On the other hand, any four classes included in pairs in the joins ANIMAL-class and VEHICLE-class would, on this definition of degree, form analogies of equal degree, for example,

$$\frac{\text{HORSE}}{\text{HERRING}} :: \frac{\text{STEAM-ROLLER}}{\text{CRUISER}}$$

The reason why the first of these is a reasonable analogy while the second is not, is that the common classes other than the ANIMAL-class and the VEHICLE-class in the first case are something like the FRAMEWORK-class and the PUMP-class, which are of comparatively low dimension since they include fewer classes than do the ANIMAL- and VEHICLE- classes; whereas in the second case the second pair of common classes is, say, the SELF-MOVERS-ON-LAND-class and the SELF-

MOVERS-IN-SEA-class which are as high in the classification as the ANIMAL-class and the VEHICLE-class.

The second suggested definition, in terms of average dimension-interval, overcomes this difficulty at least to the extent of ensuring that the first example will be of lower degree than the second. It has another advantage if we consider the degree of the examples given in Section III above. If the greatest dimension-interval is taken, the analogy between corresponding class-members and their classes

$$\frac{\text{MAN}}{\text{FISH}} :: \frac{\text{BONE}}{\text{SPINE}} \tag{4·1}$$

is of lower degree than that between the pairs of class members

$$\frac{\text{BONE}}{\text{SPINE}} :: \frac{\text{NOSE}}{\text{GILLS}} \tag{4·2}.$$

This is unsatisfactory, since it is the common relation to their classes expressed in (4·1) which constitutes the analogy in (4·2). If average dimension-interval is taken, however, assuming for simplicity that the BONE-, NOSE-, SPINE-, and GILLS-classes are in the same level, and the MAN- and FISH-classes one level higher, (4·1) and (4·2) turn out to have the same degree, namely $\frac{n-1}{n}$. We therefore provisionally adopt the second definition of degree, although it may well turn out that something more complex than either of these suggestions is more satisfactory.

We are now in a position to define an analogy-relation in the C-lattice more formally. The conditions for the analogy $(x_1 : x_2 :: y_1 : y_2)$ to hold are as follows:

A(1). If the maximum dimension-interval between any of x_1, x_2, y_2, y_1, and any of $x_1 \cup x_2$, $x_2 \cup y_2$, $y_2 \cup y_1$, $y_1 \cup x_1$, is D, and the minimum dimension-interval between any of x_1, x_2, y_2, y_1 and either of $x_1 \cup y_2$, $y_2 \cup x_1$ is D', then D' is greater than or equal to D. If D' is less than D, there is no analogy $(x_1 : x_2 :: y_1 : y_2)$.

A(2). If $D' = D$ and $y_1 = x_1$ and $y_2 = x_2$, the degree of the analogy is put equal to 1.

That is, $(x_1 : x_2 :: x_1 : x_2)$ and $(x_1 : x_1 :: x_1 : x_1)$ are analogies of degree 1.

A(3). If $D' = D$ and $y_1 \neq x_1$, $y_2 \neq x_2$, there is no analogy $(x_1 : x_2 :: y_1 : y_2)$.

A(4). If D' is greater than D, the degree of the analogy is $\frac{(n-k)}{n}$ where k is

$$\tfrac{1}{4}\left\{ d[x_1 \cup x_2] + d[x_2 \cup y_2] + d[y_2 \cup y_1) + d[y_1 \cup x_1] - d[x_1] - d[x_2] - d[y_2] - d[y_1] \right\},$$

that is, k is the average dimension-interval.

The relation has the following properties:

(i) In virtue of A(3), there are no analogies between four distinct classes of the same dimension d_1, when their six joins are all of the same dimension d_2; it follows that, if each basic property makes up a class of dimension 1 (containing that property and p_0) so that the joins of these classes in all possible pairs will have dimension 2, then there will be no analogies between any four of the classes of dimension 1. This will be the situation in the Boolean C-lattice.

(ii) If $(x_1 : x_2 :: y_1 : y_2)$ is an analogy of degree K, then $(y_1 : y_2 :: x_1 : x_2)$ is of degree K, and all analogies obtained by cyclic permutation of the terms (x_1, x_2, y_2, y_1) are analogies of degree K.

(iii) If $(x_1 : x_2 :: y_1 : y_2)$ has a.d.i. k_1, and $(y_1 : y_2 :: z_1 : z_2)$ has a.d.i. k_2, then the a.d.i. of the analogy $(x_1 : x_2 :: z_1 : z_2)$ can be calculated from the results of Section II (3) and is at most $k_1 + k_2$.

That is, the analogy relation is in a sense transitive, but in general to a lower degree.

A further point to be noticed is that although in general there are many fourth terms which form an analogy with three given terms, and although the analogy relation is not strictly transitive, the combination of a series of pairs of analogies will tend to reduce the number of solutions. If a set of values for y_2 is obtained from the relation $(x_1 : x_2 :: y_1 : y_2)$, and there is known to be an analogy $(x_1 : x_2 :: z_1 : z_2)$, a set of values for y_2 obtained from the relation $(z_1 : z_2 :: y_1 : y_2)$ may intersect with the previous set for y_2 to give a reduced number of solutions of the three relations. The process can be continued for any number of pairs known to be analogous.

This property of the relation as defined here corresponds to the intuitive requirement that if an analogy-relation is specified by a large number of pairs already known to exhibit the analogy required, the result is somehow made more determinate. A classical example is Aristotle's attempt to explain " actuality " and " potentiality " by means of a series of pairs exhibiting the required relation:

$$\frac{\text{BUILDING}}{\text{CAPABLE OF BUILDING}} :: \frac{\text{WAKING}}{\text{SLEEPING}} :: \frac{\text{SEEING}}{\text{EYES SHUT}} :: \frac{\text{FORMED MATTER}}{\text{UNFORMED MATTER}} :: \ldots\ldots :: \frac{\text{ACTUALITY}}{\text{POTENTIALITY}}.$$ [12]

This completes the mathematical discussion of the analogy-relation, and its illustration by means of a somewhat uninteresting classification system. I now want to consider its exemplification in a lattice which is a highly simplified model of the structure of language, and whose formal properties are as far as possible isomorphic with those of the *C*-lattice.

V

" Language " is taken in this model to be a total set of words regarded as signs of a certain graphical form, ignoring words which stand for syntactical connectives. Since I shall have also to speak about words as they are used in discourse in order to discuss the interpretation of this language-model, it will be convenient to refer to words regarded as signs simply as *signs*, and to use the word *word* in a looser colloquial sense in which a word can be said to have meaning, be synonymous with other words, or be used in different contexts, that is, in discourse. In this colloquial sense, the same word may be said to have different uses or meanings. In addition, the convention will be adopted that a single sign always corresponds to a single word, whether the word has only one meaning, several connected meanings (as, for example, in metaphor), or several unconnected meanings (as in puns). This convention may or may not be colloquial use, but it is possible to adopt it in this model, because the different meanings of single words will be taken care of by the classification system of words now to be explained. The power of the model in fact depends on just this possibility of distinguishing between the set of meanings of a word, and its graphical form, and formalizing both in terms of the elements of a lattice.

[12] *Met.* 1048 a, b.

We first assume that there is some empirical procedure (the nature of which is left unspecified) which results in a classification of words according to their meanings. I shall argue in a moment that there is no need for our purpose here to give a precise account of the principles of this classification, even if these could be made precise, and in fact all I shall do is point to the classification contained in Roget's Thesaurus, and attempt to give a brief and crude description of it. This description is not to be taken as in any sense a definition of a thesaurus, but only as an impressionistic account of it, to be supplemented if required by a first-hand examination of the contents of Roget.

In Roget's Thesaurus, words which have similar meanings, or "express the same idea," or are concerned with the same topic, are grouped together in sets of words called *heads*. A head therefore contains words which are, in an extended sense, "similar in meaning" and in general a single word will be included in as many heads as it has distinguishable meanings. The relation between a head and any word contained in the head will be called the *head-inclusion relation*. Imprecise as these principles of thesaurus-construction are, there is in fact found to be considerable correspondence between thesauri constructed by different methods.

I propose to assume the existence of such classifications of language without enquiring further how the particular groupings of words which they contain came to be selected. The feature of them which is of primary concern here is the formal structure which they can be seen to exemplify. If this formal structure can be used to define algorithms which, on interpretation, can be successfully applied to problems of translation, analogy-finding,[13] and so on, these applications can be regarded as tests, both of the formal structure as a theory of language, and of the particular classification of words in the thesaurus as the interpretation of that structure. I shall therefore proceed on the

[13] Tests now being carried out in the Cambridge Language Research Unit indicate that, for example, the purely verbal (*i.e.*, not mathematical) relation

$$\frac{\text{area}}{\text{volume}} \quad :: \quad \frac{\text{flat}}{\text{solid}} \quad :: \quad \frac{\text{square}}{\text{cube}}$$

can be extracted from Roget by using an analogy-algorithm similar to the one to be described here.

assumption that the formal description of the model is independent of the difficult question of how a thesaurus should be constructed.

A sub-set of signs obtained by regarding the words in a head only as signs will be called an *aggregate* of signs. Thus the language-model consists of aggregates of signs, and in its interpretation those aggregates which are derived from the empirical grouping procedure become heads, and all the signs become words.

The simplest model of this kind which will allow analogy-relations as to be defined as in the earlier part of this paper will be one which is as far as possible isomorphic with the *C*-lattice. Certain correspondences immediately suggest themselves. The initial heads derived from the empirical procedure correspond to the initial classes containing all the properties, and only the properties, of empirically existing individuals. Having made this correlation, however, several distinctions must be made between the elements of the *C*-lattice and those of the prospective language-lattice.

(i) In the case of classes of properties it was not necessary to forbid every sub-set of properties to be a class, but in the case of heads, it is necessary to forbid every sub-set of words to be a head, because the notion of a head as just described implies that not every possible combination of words makes up a head. If every possible combination did make up a head, thesauri would be vacuous.

(ii) When the definition of an analogy-relation is considered, an important difference emerges between words and the properties of the *C*-lattice. In the *C*-lattice the elements related by analogy are classes in general, not necessarily classes made up of single properties. In the prospective language-lattice on the other hand, it is not clear what would be the significance of defining a relation in the lattice which would be interpreted as an analogy--relation between *heads*. Analogy-relations between *words*, however, can be defined, because the head-inclusion relations, and the relations between heads, in principle exhaust all that is known about the semantic content of the language, and so, if analogies are expressible at all in language, they are expressible in terms of these relations. The comparison of the language-lattice with the *C*-lattice implies that head-inclusion is a partial ordering of language, giving a classification of its words, as the class-inclusion relation gives a classification of properties in the *C*-lattice. Analogies between words are then defined in terms of relations between the heads that include them, as analogies between classes of properties are defined in terms of relations between classes that include them. In the language-lattice then, analogies will only be defined as relations between signs, and interpreted as relations between words, not between heads.

(iii) That analogy-relations can be defined only between words implies another difference between the *C*-lattice and the prospective language-lattice. In the *C*-lattice some properties may be contained in classes only together with other properties, and not by themselves in classes containing only one property. But since there is no reason why all words should not be potential terms of analogy-relations, it is desirable that each sign shall be contained by itself in some element of the language-lattice. In order to ensure this, signs must each be regarded as making up an aggregate for purposes of constructing the lattice.

(iv) Further aggregates may be derived from the initial aggregates by means of a postulate which is similar to C(2), and which states that if two aggregates have signs in common then the meet and join of the aggregates are aggregates, and so on recursively. It is not necessary to assume, however, that the further aggregates thus defined can be interpreted as heads. In order to give a justification for this postulate in terms of the interpretation it is only necessary to assume that if two heads have a large set of words in common, then the heads are likely to be " close " to each other in the sense that all the words of both heads have some similarity of meaning.

This may be illustrated by an example from Roget's Thesaurus. There the heads called ABODE and RECEPTACLE have in common the words " hole ", " cell ", " cave ", " box ", " court " —a higher than average overlap of words between two Roget heads. The signs " hole ", " cell ", " cave ", " box ", " court " will form an aggregate according to our postulate, but there is no need to interpret this as a head in the usual sense (indeed the occurrence of " court " with the other words would make it difficult to do so.) The sum of all the words in ABODE and RECEPTACLE, however, may reasonably be called a head in an extended sense, for its words are related to each other by a

L

similarity of meaning roughly described by "place to put things (including people) in." As aggregates get higher up in the lattice, however, it becomes less and less easy to interpret them as heads in however weakened a sense, because the similarity of meaning between the words becomes more and more tenuous. But it is still possible to interpret the ordering of aggregates in lattice-levels as an ordering of closeness of meaning, or similarity, between pairs of heads, just as the *C*-lattice orders similarity between pairs of classes. Thus the aggregate containing the sum of signs from the "close" heads ABODE and RECEPTACLE should appear lower in the lattice than that containing the sum of signs from the "remote" heads, such for instance as ABODE and DENSITY, which have no words in common, and are clearly rather remote in meaning.

It should be remarked that "similarity" is being used in a wider sense in this interpretation of the language-lattice than in the interpretation of the *C*-lattice. In both cases it refers to similarity between classes, or heads, in virtue of properties, or words, in common between them, but in the language-lattice it refers also to the similarity of meaning between words in virtue of which the initial heads are obtained. Thus similarity between heads is an extension of this initial similarity between the words in the initial heads. This is not the case for similarity between classes, for there it is not assumed that there is any similarity between the properties which make up an initial class, but only that the properties belong to some one empirical individual.

We are now in a position to list the postulates of the partially-ordered system of aggregates.

T(1) Each sign of the total set of signs is an aggregate.

T(1′) Each sub-set of signs grouped by the empirical procedure is an aggregate.

T(2) If two aggregates have signs $w_1 \ldots w_r$ in common their logical product and their logical sum are aggregates. (T(2) then defines other aggregates recursively.)

T(3) Not every sub-set of signs is aggregate.

Aggregates thus defined are elements of a system partially-ordered by an inclusion-relation such that, if x and y are aggregates and $x \geq y$, the aggregate x contains all the signs contained in the aggregate y.

The system is not yet a lattice, however, because although

T(2) ensures that any pair of aggregates can have at most one join and one meet, it has not yet been shown that a join and a meet exist for every pair of aggregates, that is, the elements I and O have not been defined. Here another important difference emerges between the C-lattice and the prospective language-lattice. If we suppose for a moment that the element O can be introduced into this system by means of a sign w_0 formally added to all aggregates, as a property p_0 was added to all classes in the C-lattice, then the system becomes a lattice, and its first level will contain the elements (w_0, w_1); (w_0, w_2); . . . (w_0, w_n); respectively, since all the signs in the language have to appear in level 1 according to T(1). But since all pairs of signs in level 1 now have a sign w_0 in common, it follows from T(2) that the joins in all possible pairs of (w_0, w_1); . . . (w_0, w_n) appear on level 2. Now we have seen that if the joins of all possible pairs of elements on level 1 appear on level 2, it is impossible to define analogy-relations between the elements on level 1. What we have just described are in fact the first two levels of a Boolean lattice of degree n constructed out of the signs $w_1, \ldots w_n$, and this lattice contains all possible combinations of $w_1, \ldots w_n$, which is contrary to T(3). The possibility of defining analogies between words therefore depends essentially on the postulate that not all sub-sets of signs are aggregates, and the suggested method of introducing O has led to a contradiction of this postulate.

How then are the elements I and O to be defined consistently with the postulates? Consider the set of initial aggregates, and suppose for a moment that they fall into two sub-sets which are such that no sign which is contained in any aggregate of one is contained in any aggregate of the other. It will then be the case that the operation of T(2) cannot generate an aggregate containing all words in the language, but can at most generate an element containing all the words of the first sub-set, and another element containing all the words of the second. If, on the other hand, such a division into sub-sets of the total language cannot be made, T(2) will itself generate the element I containing all the signs in the language. We therefore add as a condition on the initial aggregates:

T(4) The set of aggregates defined by T(1′) cannot be divided into two sub-sets such that no sign which is contained in any aggregate of the first sub-set is contained in any aggregate of the second.

If T(4) does not hold, the language falls into two or more sub-sets such that analogy-relations cannot be defined between signs of one sub-set and signs of another, although they may be defined between signs all lying within one sub-set for which a condition like T(4) does hold. This seems to be a reasonable state of affairs, since it cannot be expected that an analogy-relation holds between terms which are otherwise quite unrelated.

When T(4) is added to the postulates, the partially-ordered system consists of a level containing aggregates each made up of a single sign, together with the initial aggregates and the further aggregates defined by T(2), including the element *I*. The level containing aggregates each made up of a single sign cannot be level 1 of the lattice, as we have just seen, but a lattice may now be constructed out of the system by reflecting in this single-sign level the system of the rest of the aggregates which lies wholly above this level. The result will be a distributive lattice in virtue of T(2); the element *I* having been reflected into *O*. The aggregates in the middle level will now have meets in the lower half of the lattice, and some of these may be interpreted as "that which words have in common in virtue of being contained in the same head", but nothing depends on this interpretation, and the bottom half of the lattice need only be interpreted as a formal reflection of the top.

Let us call this lattice the *T*-lattice. Analogy-relations between aggregates made up of single signs can now be defined in it in a way similar to analogy-relations between classes in the *C*-lattice, and can be interpreted as analogies between words. All the formal characteristics of analogy-relations in the *C*-lattice can be carried over unchanged into the *T*-lattice.

VI

This simplified model of language-structure is highly unsatisfactory in several ways, but it is only possible here to mention one obvious objection to it.

In this model the units are single signs interpreted as words, and the level of an aggregate in the lattice depends on the number of signs it contains. Now it is clearly desirable that the ordering of aggregates in levels should correspond to the *generality*, in some sense, of the heads into which some aggregates are interpreted, since heads corresponding to higher levels of the lattice contain words which are less similar to each other in meaning than are contained in heads corresponding to lower levels. But there is no obvious connexion between generality and the number of words in a head, indeed a glance at Roget's Thesaurus is sufficient to indicate that heads differ widely in the numbers of words they contain, and this number does not appear to order them in any way.

The source of the trouble is the postulate T(2), which implies that the units of the lattice are signs. Fortunately, however, this is also the postulate which is stronger than is required for the definition of analogy, for it makes the lattice distributive, whereas the analogy-relation requires only that the lattice be modular. It is therefore possible to indicate in a general way how this objection may be overcome in a lattice which satisfies the conditions of modularity, but is not constructed by means of T(2). I shall here merely suggest that, on any acceptable definition of "head," it is reasonable to suppose that the conditions for modularity, namely the chain-condition and the dimension-interval relation, will be satisfied.

Take first the chain-condition. Given the idea of an inclusion-relation between heads, the chain-condition implies that, if between any two heads A and B there is a maximum chain of n heads included in A and including B, then any other chain of heads included in A and including B has the same number of heads. Thus, to take an example from Roget, one chain relating the head ORGANIC MATTER to the word TURF goes through the heads VITALITY and VEGETABLE, while another goes through SENSATION, HEAT and FUEL. Let us regard all these heads as just the sets of words which are contained in them in Roget. We find that VEGETABLE contains a sub-set of words, beginning with BUSH and ending with CROP, which includes GRASS, TURF and GRASSLAND, and which is already distinguished from the other words in VEGETABLE by being included in another head PLAIN. This sub-set can be made into a separate head, thus giving this chain the same number of elements as the other. It is reasonable to suppose that such a sub-division of heads could generally be made in a non-arbitrary way.

Now consider the dimension-interval relation. This implies that if two heads A and B, on the same level, have a meet on the level next below, they also have a join on the level next above. That is to say, if A and B meet in C which is "close" to A and B in the sense of being an element of the level next below, this indicates that A and B have a large overlap of meaning, and that therefore there will be another head equally close to A and B in the sense of being in the level next above, and containing A and B and possibly some other heads as well. That this assumption is a reasonable one for heads has already been illustrated in the case of the Roget heads ABODE and RECEPTACLE.

In summary, it has been the purpose of this paper to show that:

(1) A four-term analogy-relation having the desirable formal properties of analogy may be defined within a system having the structure of a modular lattice.

(2) Given three terms of such a relation, a fourth term (not necessarily unique) may be calculated from them by a lattice-algorithm.

(3) A particular system of classification of properties (the C-lattice), which is described, has the structure of a distributive lattice. This satisfies conditions more restrictive than those for a modular lattice. Analogies between classes can be defined within the C-lattice.

(4) A thesaurus-like classification of words can be used to set up a T-lattice which is formally similar to the C-lattice, and this enables a definition to be given of analogies between words.

(5) The T-lattice has several unsatisfactory features when regarded as a model of language-structure. It may be possible, however, to relax the conditions of the T-lattice in such a way that the language-model has the structure of a modular but not a distributive lattice, and in this case analogies will still be definable within it.