Similarity Studies. 1. The Necessity for Analogies in the Development of Science

D. H. Rouvray

Department of Chemistry, University of Georgia, Athens, Georgia 30602

Received May 18, 1993*

An analogy is a specific kind of similarity that involves the interpretation of a given system in terms of some other system that is supposedly better understood. The analogy has played and continues to play a vital role in the development of all the sciences. Here we present a definition of the analogy in mathematical terms and examine two key instances of its application in chemistry. In the first instance we explore analogies that exist between quantum-mechanical systems and systems that can be modeled by means of Bayesian probability theory. This leads us to a new version of quantum theory that is free from wave–particle duality, collapsing wave functions, and random phenomena. In the second instance we reassess the well-known principle of analogy that like structures exhibit like behavior. As this principle is already being used in drug design among other areas, we suggest how it might be profitably extended in the more general context of molecular design studies.

INTRODUCTION

Analogies play a vital role in the cognitive processes we rely on to make sense of the natural world around us. In everyday life, for instance, where analogies are commonly referred to as metaphors, the analogy has the effect of altering the context of our perception. This makes it possible to interpret aspects of the world in a new light. In the scientific domain, it is no exaggeration to say that analogies play an indispensable role and that, without their continual application, modern science could not have developed and flourished as it has, especially in the twentieth century. In recent times, the analogy has come to be regarded as a key weapon in the armamentarium of every scientist, and one that promises spectacular results if used with some caution.[1] Appropriate analogies serve to elucidate both phenomena and events that would otherwise remain intractable and obscure. Here we shall focus on the part that analogies have played and continue to play in the chemical sciences. After a brief historical introduction to our theme, we shall explore how the analogy can be expressed in formal mathematical language. We then proceed to examine in some detail how two analogies of major significance can be applied in current chemical thinking. One of these applications is comparatively new and potentially very exciting, whereas the other is well-established though still revelant to contemporary research activity.

The essential role of the analogy in the sciences has been appreciated and commented upon by a host of observers of and participants in the scientific arena. Joseph Priestley, for instance, the discoverer of oxygen, asserted[2] that "if we could content ourselves with the bare knowledge of new facts, and suspend our judgment with respect to their causes, till, by their analogy, we were led to the discovery of more facts, of a similar nature, we should be in a much surer way to the attainment of real knowledge". The French scientist Auguste Laurent, who drew the first benzene hexagon in 1854, put forward the general principle[3] that "when two analogous bodies are submitted to the same metamorphoses, the new products will still be analogous" after carrying out detailed studies on homologous series of compounds. This, of course, was the first statement of the now widely used principle of analogy.[4] Several decades later, in proposing his kinetic theory of osmotic pressure in dilute solutions, the physical chemist Jacobus van't Hoff averred[5] that he was "not dealing with a fanciful analogy, but with one that is fundamental; for the mechanism which, according to our present conceptions, produces gaseous pressure, and in solutions osmotic pressure, is essentially the same". It is thus hardly surprising that the analogy has been described[6] as "essential in stimulating research and in scientific creativity."

One of the greatest triumphs in the use of analogy came in the second half of the nineteenth century with the construction of the periodic table. This table was progressively elaborated over a time span of more than 50 years as an ever increasing number of analogies were shown to exist among the chemical elements. The credit for the ultimate form that the table took is nowadays ascribed to the Russian chemist Dmitri Mendeleev, though other chemists set up similar schemes before him.[7] Moreover, numerous workers played a part in the overall development of the table; the evolution of the table has been comprehensively charted by van Spronsen[7] and the various forms the table has assumed over the years have been classified by Mazurs.[8] Some of the earliest basic work was performed by Döbereiner, who in 1817 established that salts of the elements calcium, strontium, and barium displayed analogies in their behavior and so could be grouped together in a triad.[9] Some twelve years later, he extended this same idea to triads of elements, such as that of chlorine, bromine, and iodine.[10] The sequence of key ideas that were eventually to lead to construction of the full periodic table is presented in Table I, which lists both the key contributors and their discoveries. In the present century, numerous periodic tables have been constructed for a wide variety of molecular structures; these tables have been comprehensively reviewed by Hefferlin and Babaev.[11]

GENERAL DESCRIPTION OF ANALOGY

The analogy is but one of several different kinds of similarity, among which may be mentioned complementarity, equivalence relations, modeling, and scaling. Most of these various manifestations of similarity have been extensively exploited in the physical sciences.[11] Analogy theory is seldom taught or analyzed because analogies are widely assumed to involve human skills that are innate and entail no more than a straightforward extension of common sense reasoning.[12] All analogies possess certain features in common that are

* Abstract published in *Advance ACS Abstracts*, February 15, 1994.

0095-2338/94/1634-0446$04.50/0

Table I. Listing of Some of the Key Discoveries and Contributions That Led to the Establishment of the Periodic Table of the Elements

discoverer	year	discovery or contribution
Döbereiner	1817	properties of strontium salts lie midway between those of calcium and barium salts
Döbereiner	1829	Law of Triads: elements form triads such as Ca, Sr, Ba and Cl, Br, I in which the middle element has the mean properties of the outer two elements
Gmelin	1843	constructed a large table of triads with electronegative elements on the left and electropositive elements on the right
Pettenkofer	1850	concluded that atomic weights of analogous elements differed by a multiple of 8
Dumas	1851	in series of analogous elements, the properties are determined by the chemical equivalent of the lowest member
Gladstone	1853	arranged the chemical equivalents of the elements in a series in numerical order, pointing out irregularities and gaps
Crookes	1854	classified the elements into seven groups on the basis of their atomic weights
Strecker	1859	pointed out that the chemical equivalents of the elements formed series that were probably not accidental
de Chancourtois	1862	arranged the elements on a helix so that analogous elements fell on the same vertical line in the helix
Newlands	1863	Law of Octaves: differences in the chemical equivalents of analogous elements are multiples of 8. Made use of atomic number for his classification and pointed out gaps
Meyer	1869	constructed a periodic table based on atomic weights which appeared in 1870
Mendeleev	1869	constructed a periodic table based on atomic weights and made predictions of missing elements that would later fill in gaps

mentioned here for our initial characterization; namely, (i) analogies are always relative and never absolute attributions; (ii) analogies refer to some specific context, model, or situation; (iii) analogies can be described in terms of a matching or mapping process from one system to some other system; and (iv) the matched constituents or component parts of the two systems are not identical but only similar. The physicist Robert Oppenheimer summed up the analogy succinctly when he stated[13] that "analogy is indeed an indispensable and inevitable tool for scientific progress...I mean a special kind of similarity which is the similarity of structure, the similarity of form, a similarity of constellation between two sets of structures, two sets of particulars that are manifestly very different but have structural parallels. It has to do with relation and interconnection."

Before delving further into the analysis of the analogy, we pause here to consider some of the analogies that have played a major role in the development of the physical sciences. Let us mention some of the manifold analogies that have been put forward to describe the atom.[14] Ancient Greek philosophers such as Leucippus and Democritus pictured atoms as tiny particles having varying shapes. Isaac Newton drew the analogy[15] that atoms are like "solid, massy, hard, impenetrable, movable particles". In more recent times the analogy of the solar system has been employed, initially by Nagoaka[16] and somewhat later Rutherford,[17] in which the atomic nucleus was seen to be analogous to the sun and the electrons to the planets. Shortly thereafter, when quantum-mechanical notions began to assume a growing importance, it was the nucleus that came to be regarded as a particle whereas the individual electrons were envisaged as analogous to a cloud in their overall behavior characteristics.[14] The above analogies for atoms are listed chronologically in Table II, along with several other examples of analogies that have been formulated over the past century or so with a view to providing a theoretical interpretation of a number of different physical systems.

Despite the fact that the writings on the subject of analogy are voluminous, very few workers have undertaken the task of classifying analogies. Thus, whereas there has been extensive coverage of topics such as the general nature of the analogy,[6,18,19] the role of the metaphor,[20–22] the exploitation of the analogy in problem solving,[23–25] and the validity of analogical reasoning,[26–28] comparatively little has been offered on the differing kinds of analogy that may be defined. Even what has been written reveals substantially different approaches to the definition. The earliest attempts to get to grips with the notion of analogy were made by a variety of philosophers, including Aristotle, Aquinas, Hume, and Kant; their contributions have been discussed by Bochenski[29] and by Sarlemijn and Kroes.[30] Among modern authors who have treated this problem within a scientific or technological context, mention may be made of Hesse,[31] Gentner,[32] Haraguchi and Arikawa,[33] Sarlemijn and Kroes,[30] Goebel,[34] Kroes,[35] and Indurkhya.[36] The changing patterns of analogy usage in the physical sciences over the centuries has been explored by Gentner and Jeziorski.[12]

Table II. Listing of Some of the Most Important Analogies That Have Been Advanced for the Interpretation of Physical and Chemical Systems

proponent(s)	year	analogy proposed
Democritus	ca. 420 B.C.	atoms are like tiny particles that differ in both shape and size
Boyle	1660	gases consist of particles that behave like miniature springs
Newton	1718	atoms are like hard, impenetrable, movable spheres
Bernoulli	1738	gases behave like billiard balls in a box
Black	1803	flow of heat through a system resembles the flow of water
Dalton	1804	molecules resemble hard spheres connected together by sticks
Davy	1806	electrolytes behave like hard charged spheres
Kekulé	1865	structure of benzene is like that of a snake biting its tail
van't Hoff	1888	dilute solutions behave like gases
Fischer	1894	a drug at a biological receptor behaves like a key turned in a lock
Thomson	1899	electrons collectively behave like clouds
Nagaoka	1904	atoms resemble planar solar systems
Rutherford	1911	atoms resemble three-dimensional solar systems
Bohr	1939	the atomic nucleus behaves like a drop of liquid
Rouse	1953	polymer molecules behave like beads on a spring
Lutz and Schmidt	1992	chemical cluster molecules behave like the atomic nucleus

CLASSIFICATION OF ANALOGIES

As indicated, a number of fundamentally different conceptions as to what constitutes an analogy are to be found in the relevant literature, each based on its own logical or set-theoretical definition. To simplify matters as far as possible, we shall follow, at least in broad outline, the approach advocated by Sarlemijn and Kroes.[30] These workers argued that the analogy could be divided up into five basic types and went on to indicate how each type could be delineated in a technological setting. We adopt their prescription of five types, though we give new examples of each that are more appropriate for our chemical context. The five types of analogy (in alphabetical order) are as follows: (1) attributive analogy, (2) functional analogy, (3) inductive analogy, (4) proportional analogy, and (5) structural analogy. For the sake of completeness, we now define each of these, even though our current interest does not extend beyond the first three of these analogy types. In passing, it should be mentioned that analogy type 4 has been used primarily by ancient Greek philosophers and is now largely of historical interest; a detailed account of the early use of this type of analogy has been provided by Szabo.[37]

(1) Attributive Analogy. Refers to objects A and B that have the respective properties or attributes α and β. When one property represents the other, or in general there is some transference between the properties, the two properties are said to be attributively analogous. An example is provided by a thermometer that is used to measure the boiling temperature of a liquid. There exists an attributive analogy between the length of the thermometer's mercury column and the boiling point of the liquid; i.e. α is analogous to β.

(2) Functional Analogy. Refers either to different objects A and B that have the same function or to different systems in which certain of the component parts play the same role. This involves relations between relations in the latter instance. As examples, we mention that both chromic acid and caustic soda, though very different chemically, are both powerful cleaning agents or that parts of the surface of quite different molecules may fit into the same biological receptor.

(3) Inductive Analogy. Refers to a series of a objects, A, B, C, D, etc., all of which exhibit the attributes P and Q. If the objects A, B, and C also exhibit the attribute R, by inductive logic we may argue that object D will also probably exhibit attribute R. An example is provided by the various penicillin molecules that are known to be bioactive and nontoxic. By inductive analogy, a new penicillin that is synthesized and found to be bioactive will probably also be nontoxic.

(4) Proportional Analogy. Refers to properties of objects that are commensurate or to numbers that are in the same ratio. For objects A, B, C, and D with respective attributes α, β, γ, and δ, the analogy is expressed by the notation $\alpha{:}\beta{::}\gamma{:}\delta$ which is read "α is to β as γ is to δ." An example is the following analogy: an atom is to an element as a molecule is to a compound.

(5) Structural Analogy. Refers to two systems, one of which is a scale model of the other. Both systems can, however, be described by the same set of equations. This type of analogy is sometimes described as isomorphism. An example is the analogy that exists between a laboratory distillation apparatus and a large industrial distillation plant.

ATTRIBUTIVE ANALOGY

We first focus our attention on the attributive analogy and show how it can be applied in a striking way in the study of the behavior of quantum-mechanical systems. These systems often have analogies to classical systems, an example of which is the canonical momentum of a classical particle which is represented by the operator $(h/2\pi i)\ \partial/\partial x$. However, it is important to remember that classical systems do not always have quantum-mechanical analogies and vice versa.[38] The statistics used to describe both types of system has in most cases been the same; that is, traditional statistics based on the notions of the frequency of an ensemble or of reasonable expectation of outcome. That these two notions do not always lead to identical results in physical systems has been demonstrated by Cox.[39] Because traditional statistics applied to quantum systems leads to paradoxes such as collapsing wave functions,[40] Schrödinger's cat,[41] and the often cited Einstein–Podolsky–Rosen paradox,[42] many attempts have been made to improve matters by changing the underlying analogies that have been used.[43] One such attempt that we outline here claims to overcome the paradoxes by using the attributive analogy that quantum-mechanical systems are analogous to Bayesian systems rather than to those describable in terms of traditional statistics.[43]

Bayes' theorem, the theoretical foundation of Bayesian statistics, was first published in 1746, though there is evidence to suggest that the true discoverer of the theorem was not Bayes but rather Saunderson.[44] The Theorem is a rule for the manipulation of probabilities and concerns the updating of a probability that some event will occur when new data bearing on the event become available. Let us start with the two basic rules of probability, viz., the sum and product rules that all schools of thought accept as being valid. In symbols, these rules assume the form:

$$p(X|Y) + p(\sim X|Y) = 1 \qquad (1)$$

$$p(XY|Z) = p(X|Z)\, p(Y|XZ) \qquad (2)$$

where X, Y, and Z represent statements or propositions, the symbol $\sim$X denotes not X, and the symbol $p(X|Y)$ is read as the probability that X is true given that Y is true. Now, because the product rule (2) is symmetric in X and Y, we may write:

$$p(XY|Z) = p(X|Z)\, p(Y|XZ) = p(Y|Z)\, p(X|YZ) \qquad (3)$$

This equation can be rearranged to yield:

$$p(Y|XZ) = p(Y|Z)\, p(X|YZ)/p(X|Z) \qquad (4)$$

provided that $p(X|Z)$ represents a finite probability, i.e. is not equal to zero. Equation 4 is the symbolic form of Bayes' theorem if we make the identifications that X represents a new set of data that comes to hand, Y is some hypothesis or analogy that we formulate, and Z is the initial information that we possessed. In words, the theorem states that the new probability that hypothesis Y is true (assuming that both the initial information Z and the new data X are valid) is equal to the product of the prior probability that hypothesis Y is true (based on the initial information Z) and the predicted probability that the data X are valid (assuming that the hypothesis Y and the initial information Z are true). This product is then divided by the probability that the new data are true given that the initial information was valid.

Bayes' theorem has been applied in the physical sciences to an ever increasing extent in recent years as it began to emerge that classical statistical arguments had their limitations. In particular, it became evident that such arguments did not always lead us to correct answers, especially in the case of quantum-mechanical systems. This has led to the search for more appropriate statistical arguments and thus to growing use of Bayes' theorem. Of course, the theorem has not been without vociferous critics, who maintain that the theorem attempts to rationalize scientific judgments by characterizing them in terms of subjective probabilities instead of objective probabilities.[45,46] This type of reasoning has been denounced by Jaynes[47] who countered that probabilities need not always be interpreted as frequencies of event occurrences and that doing so puts orthodox practitioners of classical statistics in a straitjacket that renders them incapable of coming to terms with the real problems of current science and technology. Whatever the merits of this longstanding dispute may be, it cannot be denied that Bayesian methods are now widely used in the sciences and that they appear to offer certain advantages over the classical approach.[44] Such methods have proved their value in the elucidation and sometimes the full resolution of paradoxes in the physical sciences. This fact, of course, serves to underscore just how powerful a new analogy or, more generally, a new model can be in helping to revise our description of physical reality.

As indicated in our introduction, a new analogy can have the effect of radically transforming our perception of reality. Examples of the resolution of problems or paradoxes and instances of fundamental changes in our understanding of natural phenomena that have been achieved by the application of Bayes' theorem include the derivation of a general theoretical formula for the diffusion coefficient in gases, liquids, and solids,[44] the reconstruction of sharp images of objects that are characterized by noisy data,[47] the analysis of nuclear magnetic resonance signals,[48] and a determination of the thermodynamic efficiency of muscle systems in animal species.[44] The striking illustration of the power of the analogy referred to above is that developed by Youssef[49] who attempted a reformulation of quantum mechanics by making use of the new attributive analogy that quantum systems can be modeled as Bayesian systems rather than in terms of classical statistics. On this basis he was able to demonstrate that the reformulated version not only yielded the same predictions as standard quantum theory but also gave a satisfactory explanation of quantum paradoxes. Thus, paradoxes such as wave–particle duality, the collapse of the wave function when measurements are made, the Einstein–Podolsky–Rosen paradox, Schrödinger's cat, and the role of the observer in the theory have all been elucidated. All arise from mistaking the wave function for the state of the system; in Bayesian analysis, amplitudes do not constitute a state of the system but merely represent the best estimates of certain truths about the system given some prior information. Wave–particle duality, for instance, becomes unnecessary since the apparent wave nature of particles is explained in terms of a breakdown in probability theory.[49]

FUNCTIONAL ANALOGY

At this point we turn our focus of attention to the functional analogy and it applications. Both the functional analogy and the inductive analogy considered in the next section can be conveniently discussed within the broad framework of molecular design studies. This area embraces among other things the design and development of pharmaceutical drugs, agrochemicals, fuels, and replacement body fluids.[50] One of our reasons for selecting this area is the major import it has for the commercial and industrial life of many nations. Let us start our analysis by considering some system that we can regard as effectively isolated from the rest of the world. In our case that system is one individual molecule which we suppose may be depicted by a structural formula or chemical graph

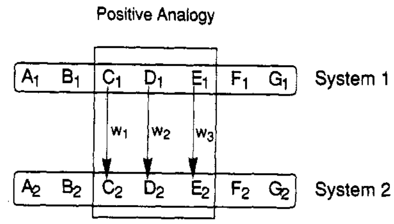

Figure 1. Schematic representation of a functional analogy that is defined between two differing systems.

with the vertices representing the atoms and the edges the bonds connecting the atoms. Such a structure may be characterized in a variety of ways, among which we might include (i) the measured chemical or other properties of the bulk material, e.g. its boiling point, refractive index, or octanol–water partition coefficient, (ii) theoretical parameters derived from the chemical graph, e.g. topological indices, information-theoretical indices, or complexity measures, (iii) quantum-chemical parameters calculated for the molecule, e.g. bond orders, charge densities, or HOMO–LUMO separations, and (iv) thermodynamic parameters, e.g. heats of formation, entropies, or heats of reaction.

The objective in molecular design studies is to produce molecules that display behavioral characteristics that are similar to those of molecules already known to be effective in some specific area of application. Thus, if it is desired to synthesize a new drug, the resulting structure should be able to mimic or even supersede the existing drug in terms in its pharmacological activity. Many different measures of structural similarity have been introduced in recent years, especially in the context of drug design.[51,52] Of those that involve the direct comparison of two structures to determine their degree of similarity, we would mention as a typical example the calculation of the Tanimoto coefficient.[53] Although such coefficients cannot provide us with absolute measures for the similarity of two structures, they do afford a means of assessing the relative similarities of structures. The Tanimoto coefficient, T, is defined as follows:[51]

$$T = \frac{n_c}{n_a + n_b + n_c} \tag{5}$$

where n_c is the number of structural fragments that a pair of molecules have in common, n_a is the number of structural fragments found in the first molecule but not in the second molecule, and n_b is the number of structural fragments found in the second molecule but not in the first molecule. To gain greater insight into the significance of such a calculation, we move on now to an in-depth analysis of the mapping processes involved.

This analysis may be conveniently carried out by referring to the diagram in Figure 1. The figure shows two systems designated as system 1 and system 2 which in the present context would represent two different molecules. System 1 is comprised of the components (atoms) A_1, B_1, C_1, etc., and system 2 of the components (atoms) A_2, B_2, C_2, etc. If system 1 is similar to system 2, the components in system 1 that are analogous to components in system 2 may be mapped in one-to-one mappings as indicated by the arrows. The system from which such arrows start is referred to as the source and that on which the arrows terminate is known as the target system. The domains that can be mapped in this way are called the positive analogy; those that cannot be so mapped are said to be the negative analogy. Corresponding components (atoms) need not be identical and may be only similar; e.g. the elements in a given column of the periodic table are analogous, so a

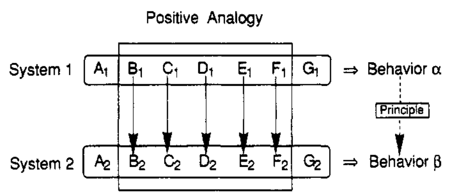

Figure 2. Schematic depiction of an inductive analogy that is defined in terms of the behavior of two differing systems by means of the principle of analogy.

potassium atom might be mapped on to a sodium atom. In cases where the components are not identical, it may be necessary to introduce some weightings into the mapping process; such weightings are indicated in our figure by symbols w_1, w_2, and w_3. Weighting procedures are used to reflect differences in the mapped components and are often rather arbitrary in nature.[51] The various connections between the components in a system, i.e. the chemical bonds in our case, can be viewed as mappings from one component to another. If the connections are also included in the intersystem mapping process, the resultant positive analogy then becomes a mapping of a mapping.

INDUCTIVE ANALOGY

Let us now explore how this detailed description of the functional analogy can be applied in the design of molecules. To do this it is necessary to take the functional analogy one stage further and to move on to a consideration of the inductive analogy. This latter type of analogy makes it possible to infer further analogies in addition to those already established by the functional analogy. To illustrate how this may be accomplished, we present in Figure 2 a diagrammatic representation of the inductive analogy. As will be clear, this figure resembles Figure 1 apart from two significant differences. The first difference is that the positive analogy is now much larger and embraces most of the system (molecule) in each case. (The weighting factors have been omitted here for the sake of simplicity.) The second difference is that a new inference is made based on the principle of analogy.[3] This inference is that, because a large positive analogy exists between the two systems (molecules), their behavioral characteristics will be similar. Thus, if system 1 exhibits a behavior α in a given setting and system 2 exhibits a behavior β in the same setting, we may conclude that the two behaviors α and β will be similar. The behavior of special interest in drug design studies is, of course, pharmacological or physiological activity, and the inductive analogy has been widely exploited in this context. The inductive analogy is however not restricted in its scope to this specific type of application and may be applied to any kind of systemic behavior.

We might mention, for instance, that a variety of workers have made implicit use of inductive analogy in studies on many different kinds of systems. Wold and Sjöström[54] used it in a study of ^{13}C nuclear magnetic resonance data that investigated whether patterns existed that could be exploited in the assignment of NMR spectra to both exo and endo molecular structures. Wipke and Hahn[55] developed a computer program that automatically invoked inductive analogy in the construction of minimum energy molecular models without resort to any minimization procedure. Okada and Kawai[56,57] utilized inductive analogy in their method of estimating the properties of molecules from their structure and of designating which component in a molecule was particularly responsible for the manifestation of a property. At this point it would seem reasonable to inquire on what basis such inferences are being made and whether these can be justified by logical analysis. Although manifold inferences of this type have been made and will doubtless continue to be made, it has to be admitted that there is no logical basis to demonstrate the validity of inductive analogical arguments. In fact the best that can be done is to point to several studies carried out within the framework of cognitive science or artificial intelligence research that apparently provide convincing testimony for the usefulness of the inductive analogy in a scientific context.

Gick and Holyoak[25] were able to substantiate anecdotal evidence that analogical thinking plays a role in creative problem solving and that the analogy may be taken from a domain that is far from the problem under consideration. Moreover, they were able to show that individuals could generate analogous solutions to problems even in cases where a complete mapping from the base to the target was not possible. Weitzenfeld[58] found from his study of the validity of reasoning by analogy that the patterns of such reasoning vary greatly from field to field and concluded that the amount of commonality there can be in these patterns is uncertain. This conclusion is supported by the work of Clement,[59] who showed that individuals adopt three different methods of generating analogies, viz., (i) the use of some general principle, (ii) an association of ideas, or (iii) a transformation of some aspect of the original problem. The latter two methods were found to be especially prevalent, whereas analogies generated via some general principle occurred only rarely. This would seem on the face of it a paradoxical strategy in that to approach a solution to a problem one actually moves away from the problem. The paradox is rationalized by Clement[59] who points out that human beings appear to be constrained to build up new knowledge by starting from old knowledge. This suggests that, when confronted by some novel problem, we attempt to "make the unfamiliar familiar" in the words of Nagel.[26] The reasoning here is that searching familiar territory may well reveal similarities between the known and the unknown and hence indicate how the new problem might be tackled.

Armed with these facts about how humans operate when it comes to applying inductive analogies, we can now begin to see what measures might be necessary to streamline the molecular design process. Clearly what is needed is a set of constraints that can be imposed on the mapping process from the base system to the target system in order to optimize the process. Fortunately, the principal workers in this field are all in broad agreement on the nature of the problem. Thus, Kodratoff,[60] Collins and Burstein,[61] and Hall[62] set up schemas similar to that shown in our figure 2 in their analysis of the problem. The difficulty arises when decisions have to be made on what constraints to impose, for little work has been done on this in the physical sciences. However, several relevant studies have been carried out in the psychological sphere, and it is to these we now turn for guidance. In passing, we mention here that this area would seem to be ripe for further development and could well provide a rich vein of stimulating new research problems. Clement and Gentner[63] proposed as a basic mapping constraint the selection of "commonalities that matter" and examined constraints stemming from the goal state, the relative importance of the information in the analogous domains, and a tacit preference for a common system of relations. Their research demonstrated that individuals preferred those matches and made those predictions that maintained a highly systematic correspondence between the two analogous domains. In addition, when deciding how to compare two situations, individuals were found to select information on the basis of its connection to a larger matching structure, suggesting that higher-order interconnections play an important role. The differing possible approaches to rule acquisition for the automation of such a task have been reviewed by VanLehn.[64]

SUMMARY AND CONCLUSIONS

Analogies are very potent abstractions. Because they have the effect of altering our perception of reality, analogies have played and will certainly continue to play an indispensable role in the development of all scientific disciplines. Although the analogy itself is viewed as a special kind of similarity, it may be further broken down into at least five different types. Here we have explored the three most prevalent types in use today, namely, attributive, functional, and inductive analogies. Analogies can be understood in various ways and given a variety of differing interpretations. For instance, an analogy may be seen as a kind of symmetry since it involves the invariance of some structure or set of relationships when system components are changed.[6] It can also be interpreted as the imposition of an equivalence relation on sets of systemic elements.[6] In the case of inductive analogy, we are dealing with an essentially psychological process, for the underlying reasoning cannot be justified on logical grounds. This suggests that subjective factors are probably entering the picture whenever we resort to the inductive analogy. The worrisome prospect is then that the inductive analogy may not always be the positive, problem-solving heuristic it is usually presented to be but rather an impediment that hinders instead of enhancing our cognitive abilities.[36] To obviate any possible negative impact analogies could have, rules are currently being devised to optimize analogy making,[64] and constraints that can be imposed on the process are being explored.[63] In fact, matters have progressed so far that some workers[65] have already introduced computer simulations that claim to automate the entire process. However, it should never be forgotten that analogies are no more than representations of reality that should not be confused with the real thing. For the mapping is not the terrain, and the analogy is not the reality.

ACKNOWLEDGMENT

Early versions of this paper were presented in oral form at the Fifth International Conference on Mathematical and Computational Chemistry held at the University of Missouri, Kansas City, MO, during the week May 17–21, 1993 and at the First "Girona" Seminar on Molecular Similarity held in Hotel S'Agaró, Spain, over the period July 8–10, 1993. The author thanks Dr. Jerry R. Dias and Dr. Ramon Carbó respectively for making his attendance at these meetings possible.

REFERENCES AND NOTES

(1) Rouvray, D. H. In *Concepts and Applications of Molecular Similarity*; Johnson, M. A., Maggiora, G. M., Eds.; Wiley-Interscience: New York, 1990; Chapter 2, p 15.
(2) Priestley, J. *Experiments and Observations Relating to Various Branches of Natural Philosophy*; Johnson: London, 1779; Vol. 1, p x–xi.
(3) Laurent, A. *Méthode de Chimie*; Mallet-Bachelier: Paris, 1854; p 373. English translation by: Odling, W. *Chemical Method*; Harrison: London, 1855; p 308.
(4) Mehlhorn, A.; Fratev, F.; Polansky, O. E.; Monev, V. Distance Measures: A New Tool for the Analysis and the Characterization of Molecular Properties. *Math. Chem. (MATCH)* **1984**, *15*, 3–103.
(5) van't Hoff, J. The Function of Osmotic Pressure in the Analogy between Solutions and Gases. *Philos. Mag.* **1888**, *26* (5), 81–105.
(6) Rosen, J. Symmetry, Analogy, Science. *Symmetry* **1990**, *1*, 19–21.
(7) van Spronsen, J. W. *The Periodic System of Chemical Elements*; Elsevier: Amsterdam, 1969.
(8) Mazurs, E. G. *Graphic Representations of the Periodic System During One Hundred Years*; University of Alabama Press: Tuscaloosa, Alabama, 1957.
(9) Döbereiner, J. W. Stoichiometric Determinations on Mineral Water, Oxalic Acid, Celestite, Sugar and Alcohol (in German). *Ann. Phys.* **1817**, *56*, 331–333.
(10) Döbereiner, J. W. Attempt at Grouping of Elementary Materials Based on Analogy (in German). *Ann. Phys.* **1829**, *15* (91), 301–309.
(11) Hefferlin, R.; Babaev, E. Periodic Systems of Molecules and Possibly Clusters. *J. Chem. Inf. Comput. Sci.*, submitted for publication.
(12) Gentner, D.; Jeziorski, M. In *Psychology of Science: Contributions to Metascience*; Gholson, B., Shadish, W. R., Neimeyer, R. A., Houts, A. C., Eds.; Cambridge University Press: Cambridge, U.K., 1989; Chapter 11, p 296.
(13) Oppenheimer, R. Analogy in Science. *Am. Psychol.* **1956**, *11*, 127–135.
(14) Rouvray, D. H. Some Key Historical Highlights in the Evolution of the Modern Concept of Valence. *J. Mol. Struct. (THEOCHEM)* **1992**, *259*, 1–28.
(15) Newton, I. *Opticks*, 2nd ed.; Innys: London, 1718; Query 31.
(16) Nagoaka, H. Kinetics of a System of Particles Illustrating the Line and the Band Spectrum and the Phenomena of Radioactivity. *Philos. Mag.* **1904**, *7* (6), 445–455.
(17) Rutherford, E. The Scattering of α and β Particles by Matter and the Structure of the Atom. *Philos. Mag.* **1911**, *21* (6), 669–688.
(18) Agassi, J. Discussion: Analogies as Generalization. *Philos. Sci.* **1964**, *31*, 351–356.
(19) Vosniadou, S., Ortony, A., Eds. *Similarity and Analogical Reasoning*; Cambridge University Press: Cambridge, U.K., 1989; esp. Chapter 7, p 199.
(20) Lakoff, G.; Johnson, M. *Metaphors We Live By*; University of Chicago Press: Chicago, 1980.
(21) Jones, R. S. *Physics as Metaphor*; University of Minnesota Press: Minneapolis, 1982.
(22) Rothbart, D. The Semantics of Metaphor and the Structure of Science. *Philos. Sci.* **1984**, *51*, 596–615.
(23) Hadamard, J. *The Psychology of Invention in the Mathematical Field*; Princeton University Press: Princeton, 1945.
(24) Mulholland, T. M.; Pellegrino, J. W.; Glaser, R. Components of Geometric Analogy Solution. *Cognit. Psychol.* **1980**, *12*, 252–284.
(25) Gick, M. L.; Holyoak, K. J. Analogical Problem Solving. *Cognit. Psychol.* **1980**, *12*, 306–355.
(26) Nagel, E. *The Structure of Science*; Harcourt, Brace and World: New York, 1961.
(27) Carney, J. D.; Scheer, R. K. *Fundamentals of Logic*; Macmillan: New York, 1964.
(28) Copi, I. M. *Introduction to Logic*, 6th ed.; Macmillan: New York, 1982.
(29) Bochenski, I. M. In *Logico-Philosophical Studies*; Menne, A., Ed.; Reidel: Dordrecht, The Netherlands, 1962; Chapter 8, p 97.
(30) Sarlemijn, A.; Kroes, P. A. In *Technology and Contemporary Life*; Durbin, P. T., Ed.; Philosophy and Technology, Vol. 4; Reidel: Dordrecht, The Netherlands, 1988; p 237.
(31) Hesse, M. B. *Models and Analogies in Science*; University of Notre Dame Press: Notre Dame, Indiana, 1966; Chapter 2, p 57.
(32) Gentner, D. Structure-Mapping: A Theoretical Framework for Analogy. *Cognit. Sci.* **1983**, *7*, 155–170.
(33) Haraguchi, M.; Arikawa, S. In *Analogical and Inductive Inference*; Jantke, K. P., Ed.; Lecture Notes in Computer Science 265; Springer: Berlin, 1987; p 61.
(34) Goebel, R. In *Analogical and Inductive Inference*; Jantke, K. P., Ed.; Lecture Notes in Artificial Intelligence 397; Springer: Berlin, 1989; p 243.
(35) Kroes, P. Structural Analogies Between Physical Systems. *Brit. J. Philos. Sci.* **1989**, *40*, 145–154.
(36) Indurkhya, B. In *Analogical and Inductive Inference*; Jantke, K. P., Ed.; Lecture Notes in Artificial Intelligence 642; Springer: Berlin, 1992; p 214.
(37) Szabo, A. *Anfange der griechischen Mathematik (Beginnings of Greek Mathematics)*; Oldenburg: Munich, 1969.
(38) Greenberger, D. M. In *Physics as Natural Philosophy*; Shimony, A., Feshbach, H., Ed.; MIT Press: Cambridge, MA, 1982; p 178.
(39) Cox, T. R. Probability, Frequency and Reasonable Expectation. *Am. J. Phys.* **1946**, *14*, 1–13.
(40) von Neumann, J. *Mathematical Foundations of Quantum Mechanics*; Princeton University Press: Princeton, NJ, 1955; Chapter 5, p 347.
(41) Schrödinger, E. Discussion of Probability Relations between Separated Systems. *Proc. Cambridge* Philos. *Soc.* **1935**, *31*, 555–563.
(42) Einstein, A.; Podolsky, B.; Rosen, N. Can Quantum-Mechanical Description of Physical Reality be Considered Complete? *Phys. Rev.* **1935**, *47*, 777–780.
(43) Stigler, S. M. Who Discovered Bayes's Theorem? *Am. Stat.* **1983**, *37*, 290–296.
(44) Jaynes, E. T. In *Maximum Entropy and Bayesian Methods*; Skilling, J., Ed.; Kluwer: Dordrecht, The Netherlands, 1989; p 1.
(45) Urbach, P. Regression Analysis: Classical and Bayesian. *Brit. J. Philos. Sci.* **1992**, *43*, 311–342.
(46) Juhl, C. Bayesianism and Reliable Scientific Inquiry. *Philos. Sci.* **1993**, *60*, 302–319.
(47) Jaynes, E. T. In *Maximum-Entropy and Bayesian Methods in Inverse Problems*; Ray Smith, C., Grandy, W. T., Eds.; Reidel: Dordrecht, The Netherlands, 1985; p 21.

(48) Bretthorst, G. L. In *Maximum-Entropy and Bayesian Methods in Science and Engineering*; Erickson, G. J., Ray Smith, C., Eds.; Kluwer: Dordrecht, The Netherlands, 1988; p 75.
(49) Youssef, S. A. Reformulation of Quantum Mechanics. *Mod. Phys. Lett. A* **1991**, *6*, 225–235.
(50) Rouvray, D. H. Taking a Short Cut to Drug Design. *New Sci.* **1983**, *138*, 35–38.
(51) Willett, P. *Similarity and Clustering in Chemical Information Systems*; Research Studies Press: Letchworth, England, 1987; Chapter 2, p 46.
(52) Willett, P. In *Concepts and Applications of Molecular Similarity*; Johnson, M. A., Maggiora, G. M., Eds.; Wiley-Interscience: New York, 1990; Chapter 3, p 43.
(53) Salton, G.; McGill, M. J. *Introduction to Modern Information Retrieval*; McGraw-Hill: New York, 1983; Chapter 3, p 52.
(54) Wold, S.; Sjöström, M. In *Chemometrics: Theory and Application*; Kowalski, B. R., Ed.; ACS Symposium Series 52; American Chemical Society: Washington, D.C., 1977; Chapter 12, p 243.
(55) Wipke, W. T.; Hahn, M. A. AIMB: Analogy and Intelligence in Model Building. System Description and Performance Characteristics. *Tetrahedron Comput. Methodol.* **1988**, *1*, 141–167.
(56) Okada, T.; Kawai, T. Analogical Reasoning in Chemistry. 1. Introduction and General Strategy. *Tetrahedron Comput. Methodol.* **1989**, *2*, 327–336.
(57) Okada, T.; Kawai, T. Analogical Reasoning in Chemistry. 2. DNET/MS System. *Tetrahedron Comput. Methodol.* **1989**, *2*, 337–347.
(58) Weinzenfeld, J. S. Valid Reasoning by Analogy. *Philos. Sci.* **1984**, *51*, 137–149.
(59) Clement, J. Observed Methods for Generating Analogies in Scientific Problem Solving. *Cognit. Sci.* **1988**, *12*, 563–586.
(60) Kodratoff, Y. In *Data Analysis, Learning Symbolic and Numeric Knowledge*; Diday, E., Ed.; Nova Science: New York, 1989; p 349.
(61) Collins, A.; Burstein, M. In *Similarity and Analogical Reasoning*; Vosniadou, S., Ortony, A., Eds.; Cambridge University Press: New York, 1989, p 546.
(62) Hall, R. P. Computational Approaches to Analogical Reasoning: A Comparative Analysis. *Artif. Intell.* **1989**, *39*, 39–120.
(63) Clement, C. A.; Gentner, D. Systematicity as a Selection Constraint in Analogical Mapping. *Cognit. Sci.* **1991**, *15*, 89–132.
(64) VanLehn, K. Rule Acquisition Events in the Discovery of Problem-Solving Strategies. *Cognit. Sci.* **1991**, *15*, 1–47.
(65) Falkenhainer, B.; Forbus, K. D.; Gentner, D. The Structure-Mapping Engine: An Algorithm and Examples. *Artif. Intell.* **1989**, *41*, 1–63.