# TOPOLOGICAL METHODS FOR SOCIAL AND BEHAVIORAL SYSTEMS

John Casti
*International Institute for Applied Systems Analysis, Laxenburg, Austria*

RR-83-3
February 1983

Reprinted from *International Journal of General Systems*, volume 8, number 4 (1982)

**INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS**
**Laxenburg, Austria**

*Research Reports*, which record research conducted at IIASA, are independently reviewed before publication. However, the views and opinions they express are not necessarily those of the Institute or the National Member Organizations that support it.

---

# FOREWORD

Many of the problems facing societies today emerge from structures that can be thought of as systems combining people and the natural environment with various artifacts of man and his technology. Systems analysis addresses problems emerging from the behavior of such systems.

In the early part of its history, systems analysis usually dealt with those systems that to a large extent were dominated by the technical artifacts involved, either of hardware or structure, and made good use of extensions of classical tools, while introducing a few new ones as the problems prompted their invention. However, since the frontiers of systems analysis have shifted more toward systems dominated by social or natural factors, these classical tools have become less appropriate and useful – new ones must be sought if the problems are to be addressed effectively.

Thus, as part of its work, the International Institute for Applied Systems Analysis explores new concepts and tools that can help systems analysis to advance on these current frontiers.

In this article, the research of which was partially supported by the Institute, John Casti describes two relatively new theories, catastrophe theory and $q$-analysis, and suggests both how their structures may offer useful insights on social and behavioral systems and how their formulations could potentially offer computational schemes that could support problem solutions. Finally, he outlines an approach that could be taken to develop a theory of surprises for social and behavioral systems.

HUGH J. MISER
*Leader*
The Craft of Systems Analysis

*Int. J. General Systems*, 1982, Vol. 8, pp. 187–210
0308-1079/82/0804-0187 $06.50/0



# TOPOLOGICAL METHODS FOR SOCIAL AND BEHAVIORAL SYSTEMS†

JOHN CASTI

*Department of Systems and Industrial Engineering, University of Arizona, Tucson, Arizona, U.S.A. and International Institute for Applied Systems Analysis, Laxenburg, Austria.*

*(Received August 7, 1981)*

Methods based in algebra and geometry are introduced for the mathematical formulation of problems in the social and behavioral sciences. Specifically, the paper introduces the main concepts of singularity theory, catastrophe theory and $q$-analysis for the characterization of the global structure of social systems. Applications in urban land development, electric power generation and international conflict are given to illustrate the methodology. The paper concludes with an outline for a general mathematical theory of surprises, together with a program for investigating the systemic property of resilience.

INDEX TERMS: Singularity theory, catastrophe theory, resilience, $q$-analysis.

## 1 MATHEMATICAL MODELING IN THE SOCIAL SCIENCES

Stimulated by the (partial) successes of their cousins in the economics community, social and behavioral scientists have been increasingly adopting the tools of applied mathematics to formulate and analyze various models of human behavior. Especially in the past decade there has been a veritable explosion of papers, books and lecture notes advocating the uses of linear programming, graph theory, regression analysis, and Markov chains, to name but a few approaches, for the study of such assorted social ills as the criminal justice system, populations migration, public health facilities and automobile parking space allocations. Interesting surveys of some of this literature are in Refs. 1–3. While we do not wish to minimize the importance or relevance of this work in any way, the fact still remains that most of the modeling efforts in the social and behavioral areas leave the practitioners and decision makers with a strong feeling of unease and dissatisfaction. The general view is that once one steps away from a very localized situation, such as the microeconomy of a firm, and tackles a large, complex system, such as the national economy, the tools and methods of classical applied mathematics fail to adequately cope with many of the essential ingredients of the problem. In short, the tools developed around the physics-based paradigm of classical mechanics and its minor extension into engineering, are no longer appropriate for capturing the structural aspects of large social systems. In the Kuhnian sense, a new paradigm has been created, requiring its own blend of mathematical concepts and tools. But, what are the distinguishing features of this paradigm which the mathematics must strive to capture?

One of the central foundations upon which the physics-based theory of modeling rests is the assumption of a basic "law" governing the relationship between the variables of the problem. Such a law may be something rather elementary like Ohm's Law or quite elaborate such as the selection rules of quantum mechanics, but the essential assumption underlying all modeling efforts is the existence of such a law. We have argued elsewhere[4] that in the social and behavioral realms there are no such laws, at least not in the sense in which the term is used in physics. While the arguments of Ref. 4 do not bear repeating here, it is worth noting that, in our view, the only possibility for freeing the modeling process from an overdependence upon

†Partial support for this paper was provided by the U.S. Department of Energy under Contract No. DE-AC01-80RA50260.

laws is to convert the laws into purely mathematical axioms. The search for laws is then replaced by the construction of appropriate mathematical structures. Thus, any mathematical theory of modeling addressed to the social sciences must not *rely* upon the presence of laws for its utilization.

A common consequence of the employment of a physical law in classical modeling is that the model then refers only to the behavior of the system in some localized spatio-temporal region. For instance, Newton's 2nd law postulates the relationship between the force exerted upon a *point* particle at a specific *moment* in time and the particle's acceleration at that same *instant*. The implicit hope in such a modeling scheme is that the local information provided by the law can somehow be pieced together to provide a global picture of the particle's behavior throughout space and time. Such a procedure works reasonably well in physics, probably due to the fact that all laws of physics can technically be expressed in terms of invariants of certain groups of coordinate transformations and it turns out that these groups have *analytic* representations. This fact has the consequence that if we know the experimental data in a local spatio-temporal region, then we can extend it by analytic continuation to other regions without having to take into account what is happening "at infinity". In short, the analyticity forces a certain "rigidity" upon the process which seems to be essential for the existence of a physical law. Needless to say, in the social sciences the absence of laws casts serious doubt over one's ability to employ the above sort of local-to-global, reductionist philosophy of modeling. Any mathematical approach to social science modeling must contain within its framework the ability to capture the global structure of a situation *without* first having to decompose the system into elementary "atoms".

The ability to express basic concepts in a morphological way is one of the strengths of classical modeling theory. Here by morphological, we mean being able to exhibit the concept in mind by a simple geometrical form. Thus, in classical physics we speak of *point* particles, *elliptical* planetary orbits, *spherically-symmetric* gravitational fields and so on. In the social sciences, concepts are used which cannot be expressed in a morphological way. For instance, notions such as "power", "status", "ideology", etc. seem difficult to identify with any "forms" from elementary geometry. Mathematical methods for modeling in the behavioral and social spheres must be capable of explicit geometric characterization and manipulation of morphogenetic fields if social science modeling is to be made into a scientific discipline.

The Cartesian/Newtonian world view, upon which classical modeling is based, says that "space" is *a priori* and that "objects" sit in it. Similarly, this view asserts the existence of some "absolute time" to go along with the "absolute space". Although the Einsteinian revolution abolished both such absolutes, it did so in a most peculiar way: by invoking the existence of another absolute, the velocity of light in a vacuum. It is our contention that the social sciences cannot abolish one type of absolutism by replacing it with another and that a coherent theory of modeling in these areas will have to appeal to the *relational* philosophy of Aristotle and Leibniz. In this world-view, the concept of space is developed via the notion of *relation* between observed objects, i.e. our awareness of space comes through our awareness of the relation between objects. In a similar view, time is then the manifestation of *relations between events*. The idea of absolute space and time is thus completely absent from the Aristotelian/Leibnizian framework, thereby providing the basis for a *holistic* rather than reductionistic theory of modeling.

The task that remains is to translate the foregoing desiderata into a specific mathematical form which will then supply the needed tools for modeling in the social sphere. While this program is as yet far from complete, the situation has progressed beyond that of mere armchair philosophy. In what follows, we shall explore two methodological directions which have been pursued with the above goals in mind: *catastrophe theory* and *q-analysis*. Each of these methodological approaches to modeling follow the same conceptual approach, namely, to map a given situation or process in the external world onto a well-defined and well-understood abstract geometrical form. In the case of catastrophe theory these forms are the geometrical objects (fold, cusp, butterfly, etc.) resulting from the Thom Classification Theorem for smooth functions. For $q$-analysis, the standard form is a simplicial complex (or collection of such complexes), which is associated in a well-defined way with the data sets and relations of the given problem. Thus, both catastrophe theory and $q$-

analysis provide us with a *language of structure*, enabling us to speak in a rather precise, mathematical and morphological way about the global and local connective structure present in any particular situation. In addition, many of the qualitative features observed in the social and behavioral sciences such as discontinuities, "surprises", hysteresis effects, subjective time scales and so on, which cause some nontrivial modeling difficulties using physics-based methods, can be approached in a mathematical way using the catastrophe theory and $q$-analysis machinery.

Following a brief outline of both catastrophe theory and $q$-analysis in the next two sections, we shall then illustrate the employment of these tools in a variety of social and behavioral settings involving electric power networks, international conflict, and land use development. These examples serve not only to indicate the power and scope of the catastrophe theory/$q$-analysis "language", but also to suggest certain extensions of the "vocabulary" and "grammar" needed for a deeper understanding of such human phenomena.

## 2 CATASTROPHE THEORY

So much has been written about catastrophe theory in the past few years that we shall refrain from a detailed exposition here, contenting outselves with the bare essentials needed for what follows. For the interested reader we recommend a recent book[5] as the best elementary introduction to the subject. Other volumes[6,7] can also be recommended for a wealth of interesting theory and examples and, of course, the original source which ignited the catastrophe theory explosion is Thom's treatise.[8] Finally, for some adverse views on the subject, particularly focusing upon some of its early application in the social sciences, see Ref. 9.

At the mathematical level, catastrophe theory is involved with the problem of classification of singularities of smooth (i.e. $C^\infty$) functions. From the standpoint of applications, the utility of the mathematical theory hinges upon being able to identify the equilibria states of the system under study with the critical points of some parametrized *family* of $C^\infty$ functions.

Roughly speaking, catastrophe theory addresses itself to the question: given a $C^\infty$ function $f(x)$, $x \in R^n$, when can we find a smooth coordinate transformation $y = h(x)$ such that in the $y$ variables $f$ is *exactly* represented by a *finite* segment of its Taylor series expansion in the neighborhood of a critical point? In the event such a transformation $h$ is possible, a secondary question then arises: is there a smoothly parametrized *family* of functions $F$ containing $f$ such that the above "finite truncation" property holds for each $f \in F$ and, if so, what does such a family look like? The Thom–Mather theorem answers the above questions in terms of certain integers computable from the function $f$ in a neighborhood of the critical point. Looking at the question from the other end, catastrophe theory also answers (partially) the question: in a $k$-parameter family of functions, which local types do we typically meet? For applications either form of the question may arise, although the latter seems to be more common. Now let us be a bit more specific about the foregoing matters.

Consider a smooth function $f(x_1, x_2, \ldots, x_n)$ in a neighborhood of the origin. We write $j^k f$ to denote the *k-jet* of $f$ at 0, i.e. the Taylor series expansion of $f$ to terms of order $k$. Thus,

$$f(x_1, x_2, \ldots, x_n) = j^k f + \hat{f}(x_1, x_2, \ldots, x_n),$$

where $\hat{f}$ is of order $k+1$. So, $j^k f$ is a polynomial function of degree $\leq k$. We say a function $f$ is *k-determinate* at 0 if whenever $j^k f = j^k g$ for some smooth $g$, there is a smooth change of coordinates $x \to y$ such that

$$f(x_1, x_2, \ldots, x_n) = g(y_1(x_1, \ldots, x_n), y_2(x_1, \ldots, x_n), \ldots, y_n(x_1, \ldots, x_n)).$$

In such a case, we say $f$ and $g$ are *right-equivalent*, denoted $f \sim_R g$. Note that the coordinate change must be *regular*, i.e.

$$\det\left(\frac{\partial y_i}{\partial x_j}\right) \neq 0$$

at the origin and it must leave the origin fixed. It need not be more than local, i.e. defined only in some neighborhood of 0. If we choose $g(x) = j^k f$, then clearly $j^k f = j^k g$, which implies that if $f$ is $k$-determinate then there is a local coordinate system such that in the $y$ coordinates $f$ is expressed *exactly* by the polynomial $j^k f$. The smallest $k$ such that $f$ is $k$-determinate at 0 is called the *determinacy* of $f$, denoted $\sigma(f)$.

If we let $f_{,i}$ denote the partial derivative $\partial f / \partial x_i$, then we say that a smooth function $\phi(x)$ is *generated* by the $f_{,i}$ if there exist $n$ smooth functions $\{\psi_i(x)\}$ such that

$$\phi(x)=\sum_{i=1}^{n}\psi_i(x)f_{,i}(x).$$

Define the non-negative integer *codim* $f$ = the number of terms which are independently *not* generated by the $f_{,i}$. This number is called the *codimension* of $f$. For example, if $f(x_1,x_2)=x_1^3+x_2^3$, then codim $f=3$, since $x_1$, $x_2$ and $x_1x_2$ are not generated (by convention, we do not consider the constant term 1 in these computations). On the other hand, if $f(x_1,x_2)=x_1^2x_2$, then codim $f=\infty$, since $x_2^k$ is not generated by the $f_{,i}$ for any $k$.

Unfortunately, it may require an infinite number of computations to decide if a particular $f$ is finitely-determined, so we introduce the related concept of *k-completeness*. The function $f(x)$ is *k-complete* if every $\phi(x)$ such that $\phi=0(|x|^k)$ is generated by $f_{,i}$ using functions $\psi_i(x)=0(|x|)$. In other words, if $\phi$ is of order $k$ and we can generate $\phi$ by $f_{,i}$ using multipliers containing no constant terms, then $f$ is $k$-complete. For example, the function $f(x_1,x_2)=x_1^4+x_2^4$ is 5-complete, but not 4-complete since terms of the type $x_1^2x_2^2$ cannot be obtained from the $f_{,i}$, although we could obtain pure quartics like $x_1^4$ and $x_2^4$. It is reasonably clear that $k$-completeness can be decided in a finite number of steps. In fact, to prove $k$-completeness it is sufficient to show that any $\phi(x)=0(|x|^k)$ can be written as

$$\phi(x)=\sum_{i=1}^{n}\psi_i(x)f_{,i}(x)+0(|x|^{k+1})$$

for some smooth functions $\psi_i(x)=0(|x|)$.

The relationship between $k$-completeness, $k$-determinacy and codimension is contained in the following theorems, due primarily to Mather and Thom.

THEOREM 1 *$f$ k-complete implies $f$ is k-determinate.*

THEOREM 2 *$f$ k-determinate implies $f$ is $(k+1)$-complete.*

THEOREM 3 *codim $f<\infty$ if and only if $f$ is finitely determinate.*

The above theorems enable us to conclude that almost every smooth function is right-equivalent to a polynomial and the only smooth functions which are not are those with codim $f=\infty$. So, if $f$ is finitely determinate, we can introduce a local coordinate system near 0 such that the behavior of $f$ in this neighborhood is *entirely and exactly* given by its $k$-jet, i.e. by a finite segment of its Taylor series expansion.

The next question to be addressed is whether or not a small perturbation of $f$ introduces any essential change into the above results. To answer this stability question we need the concept of a *universal unfolding* of $f$. Let the function $f$ have codim $f=c<\infty$, and let $u_j(x)$, $j=1,2,\ldots,c$ be independent functions *not* generated by $f_{,i}$. The function

$$f(x)+\sum_{j=1}^{c}a_ju_j(x),$$

where $\{a_j\}$ are constants, is called a *universal unfolding of f.*

Now assume that the original function $f$ is perturbed by some smooth functions $\{\phi_k(x)\}$, i.e. the new function considered is

$$f(x)+\sum_{j=1}^{n}a_j\phi_j(x),$$

which we can write as

$$f(x)+\sum_{j=1}^{c}a_ju_j(x)+\sum_{j=c+1}^{n}a_j\phi_j(x), \qquad (*)$$

where the $\phi_i(x)$, $j=c+1,\ldots,N$, *are* generated by $f_{,i}$. The main structural stability result is

THEOREM 4. *In the expression (*), the functions $\phi_j(x)$, $j=c+1,\ldots,n$, can be removed by a smooth coordinate transformation.*

Thus, the universal unfolding of $f$ represents the most general type of smooth perturbation to which $f$ can be subjected. It then follows that to study the effect of local perturbations on $f$, it suffices to study the properties of a universal unfolding.

The last ingredient we need in order to state the Thom–Mather Classification Theorem is the idea of the *corank of f.* Define the Hessian matrix, $H$ of $f(x)$ at 0 by

$$H=\left[\frac{\partial^2 f}{\partial x_i\partial x_j}\right]_{x=0}$$

The integer $r=n-\text{rank } H$ is called the *corank of* $f$ at 0. The importance of the corank resides in the following result.

SPLITTING LEMMA *The function* $f(x)$ *is right-equivalent to the function* $g(x_1, x_2, \ldots, x_r) + q(x_{r+1}, \ldots, x_n)$, *where* $g=0(|x|^3)$ *and* $q$ *is a non-degenerate quadratic form.*

Thus, the Splitting Lemma enables us to separate the variables $x_1, \ldots, x_n$ into 2 classes: the "essential" variables entering into intrinsically nonlinear behavior, and the $n-r$ "inessential" variables which play no role in bifurcations and "catastrophes".

The claim of utility of catastrophe theory in the social and behavioral sciences rests heavily upon the Splitting Lemma in the following sense. Investigations in the social sciences usually involve large numbers of variables and traditional methods usually attempt to control all but a small number and analyze the interrelations of those remaining. On the other hand, in catastrophe theory the primary focus is upon the codimension, i.e. the number of assignable parameters. If this is small, which is usually required for any decent theory, then the corank is also small.† Hence, the Splitting Lemma then insures that the number of mathematically relevant state variables is small, usually 1 or 2. All the other state variables are well-behaved in a neighborhood of the critical point.

Finally we can state the basic classification result of Thom.

CLASSIFICATION THEOREM *Up to multiplication by a constant and addition of a non-degenerate quadratic form in other variables, every smooth function of codimension* $\leqq 6$ *is right-equivalent to one of the universal unfoldings listed in Table I.*

If we denote the universal unfoldings of Table I by $y(x)$, then we define the bifurcation set $B$ to be

$$B=\left\{a_k: \frac{\partial y}{\partial x_i}=0, \det\left[\frac{\partial^2 y}{\partial x_i \partial x_j}\right]=0, \begin{array}{l} i,j=1,2,\ldots,n \\ k=1,2,\ldots,c \end{array}\right\}.$$

†More precisely, $\binom{r+1}{2} \leqq c$.

In the above set-up, the parameters $\{a_k\}$ are usually thought of as assignable "control" variables and the $x_i$ are smooth functions of the $a_k$ except at points on $B$. So, as the parameters slowly change there *may* be a sudden change in the $x_i$ as the controls pass across $B$. Such a discontinuity is what is usually termed a "catastrophe" in the popular literature.

In regard to applications of catastrophe theory, we can distinguish two approaches. The first is when we actually know some physical law governing the process under study. In this case, we can take the known law as our function $f(x)$ and subject it to the machinery outlined above to reduce it to one of the standard forms of Table I. Such an approach is most typical of the physical sciences and has been used with some success in mechanics, geometrical optics and elasticity theory.[6] Interesting applications in biology and ecology using this approach have also been reported.[7]

The second "metaphysical" approach to the use of catastrophe theory is to postulate *a priori* that the unknown process governing the system under investigation meets the assumptions of the theory, e.g. that there exists some underlying potential function which the system locally (or globally) moves so as to minimize. This approach is more characteristic of applications in the social and behavioral areas and is the line which we shall follow in the latter sections of this paper.

## 3 *Q*-ANALYSIS

Catastrophe theory focuses upon the structure present in smooth functions of several variables and provides a geometric language for characterizing this structure. The language termed "$q$-analysis",[10] or "polyhedral dynamics",[11] offers a similar approach to the study of binary relations between finite sets of data. Thus, while catastrophe theory with its emphasis upon smooth functions, is heavily-flavored by the analytic tools of *differential* topology, $q$-analysis relies upon the ideas and methods of *algebraic* topology.

Consider two finite sets

$$X=\{x_1, x_2, \ldots, x_n\}, \quad Y=\{y_1, y_2, \ldots, y_m\},$$

and a binary relation $\lambda \subset Y \times X$. As is well-known, we can represent $\lambda$ by an $m \times n$ incidence

TABLE I
Functions of codimension $\leqq 6$

| Corank | Codimension | Function | Universal unfolding | Name |
|---|---|---|---|---|
| 1 | 1 | $x^3$ | $x^3+a_1x$ | Fold |
| 1 | 2 | $x^4$ | $x^4+a_1x^2+a_2x$ | Cusp |
| 1 | 3 | $x^5$ | $x^5+a_1x^3+a_2x^2+a_3x$ | Swallowtail |
| 1 | 4 | $x^6$ | $x^6+a_1x^4+a_2x^3+a_3x^2+a_4x$ | Butterfly |
| 1 | 5 | $x^7$ | $x^7+a_1x^5+a_2x^4+a_3x^3$ $+a_4x^2+a_5x$ | Wigwam |
| 1 | 6 | $x^8$ | $x^8+a_1x^6+a_2x^5+a_3x^4$ $+a_4x^3+a_5x^2+a_6x$ | Star |
| 2 | 3 | $x_1^3-3x_1x_2^2$ | $x_1^3-3x_1x_2^2+a_1(x_1^2+x_2^2)$ $+a_2x_1+a_3x_2$ | Elliptic umbilic |
| 2 | 3 | $x_1^3+x_2^3$ | $x_1^3+x_2^3+a_1x_1x_2+a_2x_1$ $+a_3x_2$ | Hyperbolic umbilic |
| 2 | 4 | $x_1^2x_2+x_2^4$ | $x_1^2x_2+x_2^4+a_1x_1^2+a_2x_2^2$ $+a_3x_1+a_4x_2$ | Parabolic umbilic |
| 2 | 5 | $x_1^2x_2+x_2^5$ | $x_1^2x_2+x_2^5+a_1x_1^2+a_2x_2^2+a_3x_1$ $+a_4x_2+a_5x_2^3$ | 2nd hyperbolic umbilic |
| 2 | 5 | $x_1^2x_2-x_2^5$ | $x_1^2x_2-x_2^5+a_1x_1^2+a_2x_2^2+a_3x_1$ $+a_4x_2+a_5x_2^3$ | 2nd elliptic umbilic |
| 2 | 5 | $x_1^3+x_2^4$ | $x_1^3+x_2^4+a_1x_1+a_2x_2+a_3x_1x_2$ $+a_4x_2^2+a_5x_1x_2^2$ | Symbolic umbilic |
| 2 | 6 | $x_1^3+x_1x_2^3$ | $x_1^3+x_1x_2^3+a_1x_1+a_2x_2+a_3x_1x_2$ $+a_4x_1^2+a_5x_1^2x_2+a_6x_2^2$ | (None) |
| 2 | 6 | $x_1^2x_2+x_2^6$ | $x_1^2x_2+x_2^6+a_1x_1^2+a_2x_2^2+a_3x_1$ $+a_4x_2+a_5x_2^3+a_6x_2^4$ | 2nd parabolic umbilic |

matrix $\Lambda$ defined as

$$[\Lambda]_{ij}=\begin{cases}1, \text{ if } (y_i, x_j)\in\lambda\\ 0, \text{ if } (y_i, x_j)\notin\lambda\end{cases}.$$

Associated with the relation $\lambda$ are two *simplicial complexes* $K_Y(X;\lambda)$ and $K_X(Y;\lambda^*)$ defined in the following fashion: in $K_Y(X;\lambda)$ we identify the elements of the set $X$ with the vertices of the complex and let the elements of $Y$ represent the simplices. Thus, $y_i$ is the $p$-simplex consisting of the vertices $x_{i_1}, x_{i_2},\ldots, x_{i_{p+1}}$ if and only if $(y_i, x_{i_j})\in\lambda$ for $j=1,2,\ldots,p+1$. The *conjugate complex* $K_X(Y;\lambda^*)$ is formed by interchanging the roles of the sets $X$ and $Y$, which in terms of the incidence matrix $\Lambda$ involves using $\Lambda'$, the transpose of $\Lambda$. In this fashion, we can associate a standard geometrical form, namely a simplicial complex, with every binary relation Λ and use the structural properties of this form to tell us something about $\lambda$. This is the essential idea underlying $q$-analysis.

An important aspect of the effective use of the above idea is the recognition that data sets $X$ and $Y$ are often hierarchically structured. To account for this structure within the $q$-analysis language, we employ the notion of a *set cover*. We say a set $A$ *covers* a set $X$ if

i) each $a_i \in A$ is contained in $P(X)$, the *power set* of $X$;

ii) $\cup a_i = X$.

Thus, each element $a_i$ is the name of a *subset* of elements from $X$. The special case of a *set partition* occurs when $a_i \cap a_j = \{\phi\}$.

If we think of the set $X$ as being at some particular hierarchical level, say $N$, then it is natural to say that $A$ is at the $(N+1)$-level. Similarly, we could find a cover Γ of $A$ and think of Γ as existing at level $(N+2)$ and so on. Or, going the other direction, we may regard $X$ as a cover of a set $Q$ which would then be placed at the $(N-1)$-level. In this manner, the hierarchical diagram in Figure 1 could be obtained.

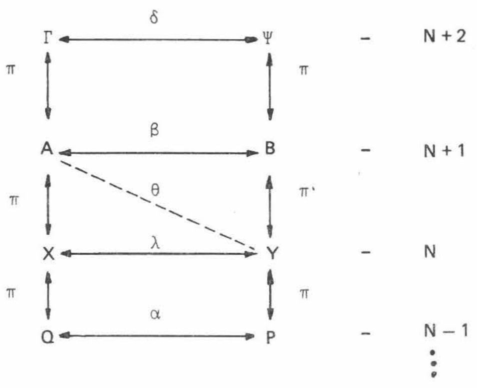



FIGURE 1

In this diagram, the relation $\pi$ represents the natural relation associating a given element at one level with its subset at the next. The diagonal relation $\theta$ taking us from one level to the next is defined in the usual set-theoretic way in order to make the diagram *commutative*.

Within the $q$-analysis framework, the basic building blocks of the relation $\lambda$ are the simplices of the complexes $K_Y(X;\lambda)$ and $K_X(Y;\lambda^*)$. We are interested in studying the way in which these pieces are "glued" together in the complex. To this end, we define a connective relation upon the simplices as follows. We say that two simplices $\sigma_p$ and $\sigma_r$ are *q-connected* if there exists a sequence of simplices $\{\sigma_{\alpha_i}\}_{i=1}^{n}$ in $K$ such that

i) $\sigma_p$ shares a face of dimension $\beta_0$ with $\sigma_{\alpha_1}$;

ii) $\sigma_r$ shares a face of dimension $\beta_n$ with $\sigma_{\alpha_n}$;

iii) $\sigma_{\alpha_i}$ shares a face of dimension $\beta_i$ with $\sigma_{\alpha_{i+1}}$;

iv) $q = \min\{\beta_0, \beta_1, \ldots, \beta_n\}$.

(Note: we shall adopt the standard notational convention that $\dim \sigma_i = i$, with $\dim \sigma = (\#$ vertices in $\sigma) - 1$. Also, $\dim K = \dim$ of highest dimensional $\sigma \in K$). It is an easy matter to verify that *q-connection* is an *equivalence relation* on $K_Y(X;\lambda)$, so we may study the equivalence classes of this relation. For each value of $q = 0, 1, \ldots, \dim K$, we define the integer

$$Q_q = \# \text{ of distinct } q\text{-classes}$$

and call the vector

$$Q = (Q_N, Q_{N-1}, \ldots, Q_0),\ N = \dim k,$$

the *structure vector* of $K$. The vector $Q$ gives us some idea of the global geometry of $K$, as it tells us how many $q$-dimensional "pieces" exist in the complex. The lower-dimensional $(<q)$ "gaps" between these pieces form an obstacle to the natural flow of information or "traffic" throughout $K$, an observation that is of some significance as our later applications will show.

While $Q$ tells us something about the complex $K$ as a whole, the relation of $q$-connection provides little information about the *individual* simplices of $K$. In particular, it is of interest to know how well a given simplex fits into the overall complex and, especially, whether or not a particular simplex should be regarded as "unusual" or "special", relative to the rest of the complex. As a measure of integration, we define the *eccentricity* of a simplex $\sigma$ as

$$\operatorname{ecc} \sigma = \frac{\hat{q} - \check{q}}{\check{q} + 1},$$

where $\hat{q} = \dim \sigma$, $\check{q}$ = highest-dimensional face which $\sigma$ shares with another distinct simplex in $K$. We remark that the above definition has the defect that the measure of a simplex's non-conforming nature depends only upon another *single* simplex in $K$ and not upon all the other members of $K$. Various alternative measures of eccentricity have been proposed to eliminate this problem but for our purposes the above definition, due to Atkin,[10] will suffice.

In order to consolidate the above notions, let us consider the following simple example. Let the incidence matrix of the relation $\lambda$ be given by

| | $\lambda$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | $x_8$ | $x_9$ | $x_{10}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $y_1$ | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| | $y_2$ | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| | $y_3$ | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| $\Lambda=$ | $y_4$ | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 |
| | $y_5$ | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 |
| | $y_6$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 |
| | $y_7$ | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 |
| | $y_8$ | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |

Geometrically, $K_Y(X;\lambda)$ has the form shown in Figure 2.

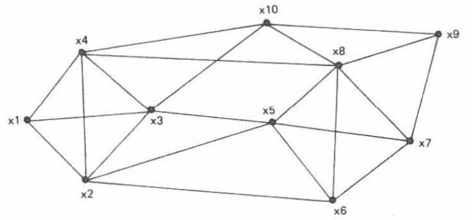



FIGURE 2

The structure vector for $K$ is $Q=\begin{pmatrix}3 & & & 0\\ 2 & 8 & 1 & 1\end{pmatrix}$. Thus there are 2 3-connected components in $K$, consisting of the single simplices $\{y_1\}$ and $\{y_4\}$. At the opposite end of the spectrum, we have $Q_0=1$ indicating that the complex is a single piece at the 0-dimensional level. The fact that $Q_2=8$ shows that $K$ splits into many disjoint pieces as far as 2-dimensional connectivity is concerned, indicating that there is a high likelihood of serious obstruction to the flow of traffic between various 2-dimensional simplices in $K$. We shall pursue these ideas further on after introducing the idea of a *dynamic* on $K$.

By "traffic" on a complex $K$, we mean anything associated with $K$ that

i) is defined on the simplices of $K$

and

ii) can be described by a *graded set function*

$$\pi=\pi^0\oplus\pi^1\oplus\ldots\oplus\pi^N$$

which we call the *pattern* of the traffic. Each

$$\pi^i\colon \{i\text{-dim. simplices}\}$$
$$\rightarrow J \text{ (a number domain).}$$

Thus, $\pi^i$ is the pattern $\pi$ restricted to the $i$-simplices of $K$. A typical example of traffic on $K_Y(X)$ would occur in a situation if we had $X$ = traffic routes thru a town, $Y$ = type of vehicles, and the traffic as the amount of goods and people carried by different vehicles.

Any change in the pattern $\pi$, which is part of a free redistribution of the values of $\pi$, means effectively that there is a free flow of numbers from one simplex to another. However, since $\pi$ is *graded* by dimensionality levels, the numbers themselves acquire a dimensional significance which must be taken into account when studying the redistribution of numbers from one simplex to another. Hence the dimensions of the common faces of two simplices is very important. If the pattern $\pi^t$ is to change freely then it needs a $(t+1)$-chain of connection to do so. Thus, the number of separate $t$-connected components in $K$ is an indication of the impossibility of free changes in $\pi^t$. For this reason we define the *obstruction vector* $\hat{Q}$ as

$$\hat{Q}=Q-U,$$

where $U$ = vector all of whose components equal 1.

Note that the above considerations regarding free changes of $\pi$ are related only to the underlying geometry of $K$. The peculiarities of some particular pattern $\pi$ might also involve internal constraints on the actual chains of connection within a single connected component which place additional obstacles in the way of the change $\pi\rightarrow\pi+\delta\pi$. For instance, we might have a "conservation law" of the form $\sum_{t=0}^{n}\delta\pi_t=0$. This is an additional constraint, above and beyond those imposed by the geometry.

A change in $\pi$ at the level $t$, i.e. $\delta\pi^t$ is associated with a "force" in $K$ at the dimension level $t$. If $\delta\pi^t > 0$ we speak of an *attractive t-force*, while $\delta\pi^t < 0$ is *t-force of repulsion*.

A great deal of the additional algebraic structure of $K$, including notions of "holes" and "loops" in the complex are discussed in some detail in Ref. 12. Of special interest for applications are the computational methods developed for patterns and their relationship with dynamics on $K$. Let us now turn to some prototypical applications of the methods introduced above.

## 4 LAND USE AND DEVELOPMENT

As a simple illustration of how catastrophe theory is sometimes applied in practice, let us consider an urban housing model, whose objective is to predict the development of a given residential area as a function of both the accessibility of the area and the number of vacant units available. More specifically, let

$\dot{N}(t)$ = rate of growth of housing units in the area at time $t$;

$a$ = *excess* number of vacant units relative to the regional norm;

$b$ = *relative* accessibility of the area to the regional population.

Our goal is to describe the variation of $N$ as a function of $a$ and $b$.

In order to justify employment of catastrophe theory we shall assume that the dynamic underlying $N$ is such that for each $(a, b)$ level, $\dot{N}(t)$ moves so as to locally maximize a potential function $V$. This assumption (or its equivalent) is often employed in land development models of the so-called "gravity" type. Furthermore, we assume that for each level of $a$ and $b$ the time-scale for the change of $\dot{N}$ is fast enough that we observe only the steady-state level of $\dot{N}$, i.e., the transient dynamics of $\dot{N}(t)$ are "fast" compared to the "slow" changes of $a$ and $b$ (for a theoretical treatment of this "delay" convention as well as a discussion of what is fast and what is slow, see Ref. 13).

Under the foregoing hypotheses, we may invoke the catastrophe theory machinery and regard $\dot{N}$ $(=\dot{N}(\infty))$ as the single "essential" variable of the Splitting Lemma, with $a$ and $b$ as two parameters. In catastrophe theory parlance, we are in the case of the *cusp* catastrophe, which has the universal unfolding $V = \pm(\dot{N}^4/4 + a\dot{N}^2/2 + b\dot{N})$, leading to the well-known picture of the equilibrium manifold $M$ for $\dot{N}$ given in Figure 3.

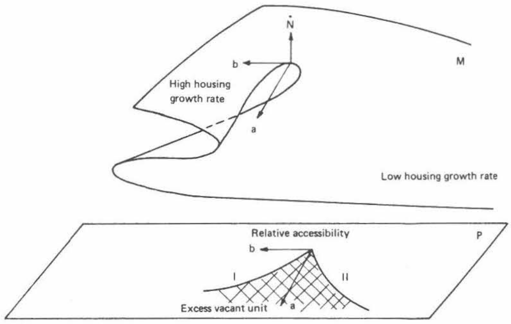



FIGURE 3

In the above canonical unfolding for the potential $V$, we would choose the negative sign since it is more reasonable to assume that for a fixed level of vacancy and accessibility, a given region will develop at the fastest, rather than slowest possible rate. Thus, our model is actually the so-called "dual" cusp.

The picture emerging from Figure 3 is that the housing rate will grow discontinuously only if a combination of high vacancy and high accessibility (probably strongly positively correlated with desirability) takes place in such a way as to cross the fold line (I) moving from right to left in the parameter space $P$. Similarly, we can expect the growth rate to "crash" if the vacancy/accessibility combination crosses the fold line (II) from left to right. To *insure* a smooth development of housing, it is necessary to take steps which prevent entering the shaded cusp region. The simplest way to accomplish this is to keep $a < 0$, i.e. make sure that the vacancy rate of the particular area is no greater than that of the regional average. Zoning regulations, preferential tax rates, restrictions on building permits and/or housing subsidies could all contribute toward keeping $a$ small.

On the other hand, should we wish to stimulate a sluggish housing market and promote a boom in development of a particular area, it would be necessary to have $a>0$ and to make $b$ large, i.e. the model suggests actions such as subsidies for construction, building of new roads to the area, encouragement of development of local shopping areas and so forth. All actions of this type would work toward forcing the system to cross the fold line (I), consequently increasing the growth rate discontinuously.

To transform the above static model into a dynamical description, we can use the assumption that $\dot{N}(t)$ moves so as to maximize the potential $V(N)$. This leads to the dynamical equation for $\dot{N}$ as

$$\ddot{N}=\frac{d}{dt}[\dot{N}(t)]=[\dot{N}^3+a\dot{N}+b].$$

Now, of course, $a$ and $b$ must also be regarded as time-varying functions $a(t)$, $b(t)$ satisfying their own differential equations

$$\frac{d}{dt}a(t)=G_1(a,b,\dot{N}),$$

$$\frac{d}{dt}b(t)=G_2(a,b,\dot{N}).$$

The functions $G_1$ and $G_2$ are not dictated by the catastrophe theory methodology and must be determined through understanding of the particular process and utilization of measured data, if available. We note in closing that in order to have the dynamical model merge into the earlier static one, it is necessary to choose the functions $G_1$ and $G_2$ so that the time-scales of $N$ and $a$ and $b$ differ significantly. In other words, we cannot use functions $G_1$ and $G_2$ which would cause $a$ and $b$ to change at more or less the same rate as $\dot{N}$. This constraint can be easily met, however, by first choosing physically meaningful $G_1$ and $G_2$, then multiplying these functions by a small parameter $\varepsilon \ll 1$, which would act to slow down the time-scale in $(a, b)$ space.

Shifting now to the problem of land use, let us consider the employment of $q$-analysis for the study of how the types of activities of a given town interconnect with the physical space available. Assume that the town has a certain set of geographically-distinct areas which form the members of a set $X$. For instance, in Manhattan we might have

$X=$ {Upper East Side, Upper West Side, Harlem, Midtown, Theatre District, Garment District, Chelsea, Greenwich Village, Soho, Chinatown, Financial District}

$$=\{x_1, x_2, \ldots, x_{11}\}.$$

We also have a collection of activities which may take place in the locations of $X$. Such activities form the elements of a set $Y$. Let us take

$Y=$ {retail trade, cultural amenities, residential, entertainment, light manufacturing, heavy industry, finance/business}

$$=\{y_1, y_2, \ldots, y_7\}.$$

An obvious relation $\lambda$ on $X \times Y$ is

$\lambda$: $(x_i, y_j) \in \lambda$ if and only if activity $y_j$ takes place in area $x_i$.

A plausible incidence matrix for the relation $\lambda$ using the above sets $X$ and $Y$ is

| $\lambda$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | $x_8$ | $x_9$ | $x_{10}$ | $x_{11}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $y_1$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| $y_2$ | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| $y_3$ | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 |
| $y_4$ | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| $y_5$ | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 |
| $y_6$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $y_7$ | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

The structure vector for the complex $K_Y(X;\lambda)$ and its conjugate $K_X(Y;\lambda^*)$ are

$$Q=\begin{pmatrix} 10 & & & & & & & & & & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$Q^*=\begin{pmatrix} 4 & & & & 0 \\ 1 & 1 & 2 & 1 & 1 \end{pmatrix}.$$

Thus, in terms of the activities taking place in the city we see that there is a high degree of connectivity at every dimension level. This is accounted for by the fact that retail trade and residential activity ($y_1$ and $y_3$) take place in so many sections of the city. In fact, from dimension 4 to 10, these are the only activities we "see" taking place in the city. So, if we are considering activities which occur in 5 or more neighborhoods, we would view Manhattan as consisting only of retail trade and residential property. Furthermore, we note that heavy industry ($y_6$) is a $(-1)$-dimensional simplex showing that it does not belong to the complex at all and could be eliminated from our analysis altogether.

Looking at the eccentricities of the simplices, only retail trade ($y_1$) has a nonzero eccentricity ($\text{ecc}\, y_1 = 4/7$), indicating that there is really no single activity that is carried out in isolation in the city.

Viewing the city from the standpoint of neighborhoods rather than activities, the conjugate complex shows that Soho ($x_9$) is the most diverse neighborhood with 5 different activities taking place there, followed closely by Greenwich Village ($x_8$) with 4. In terms of overall cohesion via activities, $Q^*$ shows us that the city is well-connected at all levels except for the small fragmentation at $q=2$. A more detailed look at this separation shows that the 2-connected components are Midtown ($x_4$) and the collection of neighborhoods $N=\{x_3, x_5, x_8, x_9, x_{10}\}$. This indicates that there is some combination of 3 activities happening in Midtown that is not shared by the neighborhoods $N$. Inspection of the situation shows that this is due to the fact that the vertex $y_7$, finance/business, does not occur in any part of $N$. Other than this small anomaly, the view of Manhattan as a collection of neighborhoods suggests that the activities act to "cement" the neighborhoods together in a very solid fashion. This feeling is further borne out by the fact that the eccentricities of the neighborhoods are all very small, with only Soho and Midtown being nonzero, and even these two are quite insignificant ($\text{ecc}\, x_9 = 1/4$, $\text{ecc}\, x_4 = 1/2$).

In conclusion, the overall picture that emerges of Manhattan from the above analysis is just that which one obtains intuitively, namely, a collection of individual neighborhoods well-connected to each other through a broad array of urban activities. Furthermore, the activities themselves are well-distributed throughout the city justifying what every New Yorker knows that you can live your whole life in your own neighborhood and not feel that you're missing anything!

Should we wish to take a more detailed view of the above relation $\lambda$, we could employ the set cover idea to decompose the $N$-level sets $X$ and/or $Y$ into their $(N-1)$-level components. For example, the set $X$ may be thought of as a cover for a new set $U$ consisting of elements

$$U = \left\{ \begin{array}{l} E59\text{th St.} - E96\text{th St.}, E96\text{th St.}, \\ E125\text{th St., above } E125\text{th St.}, E42\text{nd} \\ \text{St.} - E59\text{th St.}, E14\text{th St.} - E42\text{nd St.}, \\ \text{Canal St.} - E14\text{th St., below Canal} \\ \text{St.}, \ldots . \end{array} \right\}$$

$$= \{u_1, u_2, \ldots, u_m\}.$$

Thus, each of the elements $x_i = \bigcup u_j$, where the union is over all elements of $U$ corresponding to the particular region $x_i$. In this way we can take a more detailed look at how local neighborhoods relate to the human activities taking place within them and also how the activities work to tie neighborhoods together.

On the other hand, should we wish to examine the activities in greater detail then we would use the set $Y$ to cover a collection of activities. For instance, the element $y_1$, retail trade, may act as a name for the set {butcher, bookshop, supermarket, jeweler, dept. store, camera store, pizzeria, barber}. Thus, each $y_i = \bigcup_j z_j$ and the elements $z_j$ form a new set $Z$ covered by $Y$. So, we would have the hierarchy shown in Figure 4.

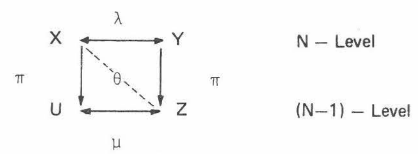



FIGURE 4

The natural projections $\pi$, together with the relations $\lambda$ and $\mu$, defined at their respective levels, enable us to construct the relation $\theta$ linking the micro to the macro view of Manhattan.

Some typical patterns which may be defined on the complex $K_Y(X; \lambda)$, include

1) amount of money spent/year in activity $y$,
2) number of people employed in activity $y$,
3) tax base provided by activity $y$.

On the complex $K_X(Y;\lambda^*)$, some possible patterns of interest include

4) number of people living in area $x$,
5) number of square blocks included in area $x$,
6) political voting distribution in area $x$,
7) consumer-good spending in area $x$.

Thus, we see that the $q$-analysis language provides us with a very flexible tool for looking at many facets of the urban structure present in a given town and gives a basis for a rational plan of land use development.

## 5 INTERNATIONAL CONFLICTS AND CRISES

Crisis has been referred to as both the actual prelude to war and the averted approaches. The current international situation certainly makes the importance of crisis perception and management clear, but a definite conceptualization of crises has so far eluded students of the subject. Perhaps the vagueness of the term "crisis" is to blame. Nonetheless, if we assume the validity of McClelland's definition:[14] "A crisis is, in some way, a change of state in the flow of international political actions", then catastrophe theory suggests itself as a possible language with which to distinguish crisis from noncrisis periods.

Since there is no readily identifiable "potential" function governing the dynamics of crisis onset and disappearance, we shall employ the "metaphysical way" of catastrophe theory and postulate the existence of such a potential. Furthermore, we shall also assume that the coordinate system chosen to verbally describe the situation is such that we can appeal to the Splitting Lemma and separate the many variables involved in a crisis into "essential" and "inessential" variables, with the essential variables corresponding to our perceived reaction to the crisis. In this example, there will be only the single essential variable, *military action*. This is equivalent to stating that our postulated potential function is of corank 1. The control parameters used in our model of crisis will be perceived *decision time* and perceived *threat*. Choice of these variables implies that we are assuming our potential function to be of codimension 2 which, by the Classification Theorem, implies that the crisis situation can be represented by the *cusp* catastrophe.

As an aside, we note that the control parameters are consistent with those advocated by Hermann[15] in his work on crisis detection. In his case there are three control dimensions, the element of *surprise* being added to the two variables time and threat. Hermann represents these three dimensions in a crisis cube (Figure 5).

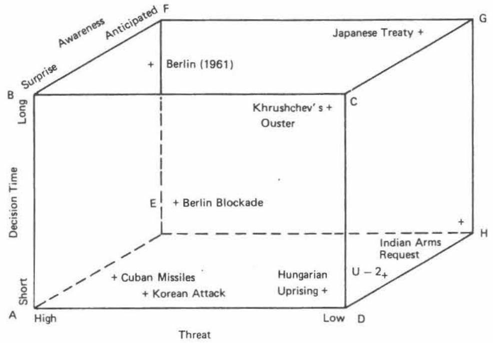



FIGURE 5

According to Hermann's theory of crisis, the vertices of the cube correspond to various levels of crisis with a high crisis situation being characterized by points near $A$ (high threat/short time/surprise) and the routine situations being at $G$ (low threat/extended time/anticipated).

In our simplified cusp model, decision time increases the relative amount of time available for choosing alternative behaviors. The zero point represents normalcy, or average decision time, using everyday standard operating procedures. On the low end, decision time is a matter of minutes such as reaction to a nuclear attack. On the high end, actions need not be taken for several days or weeks.

The continuum for perceived threat will range from strategic dominance at the low end, to strategic impotence at the high end. It should be kept in mind that we are speaking here of *perceived* threat, which may be a quite different matter than actual threat.

The behavioral output variable, military action, lies on a continuum going from complete passivism on the one hand to nuclear attack on the other. Military operations begin at the zero point and build to nuclear attack at the extreme.

Putting all the foregoing assumptions and definitions together we arrive at the cusp geometry of Figure 6 for characterizing the crisis situation.

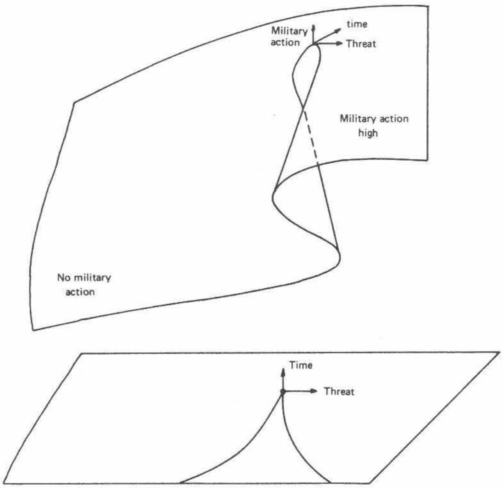



FIGURE 6

Since the literature[16] on crisis management indicates that the jump from a non-crisis mode to crisis is distinct from the jump from crisis mode to non-crisis, we employ the so-called "delay" convention for interpretation of the model.

The geometry of Figure 6 strongly suggests that when the system is not already involved in military action, it will not choose to engage in such action until the threat is extremely high. But if military action is already being taken, the system will continue such action until the threat is fairly low (e.g. occupation forces). The actual levels of where the threat is perceived to be high or low can only be determined by empirical means.

The divergence in moving from military action to non-action becomes larger as decision time becomes shorter. This implies that when decision time is short, the decision to initiate hostilities may be less routine than the cessation of military behavior.

A very interesting extension of the above model is reported in Ref. 17, where, in addition to the variables of the cusp model, an additional behavior (output) variable termed "operational preparedness" is introduced, along with a third input parameter "degree of uncertainty". The uncertainty in a situation is a replacement for the variable "surprise" in Hermann's crisis cube, and is justified on the grounds that if one defines a crisis situation only in terms of events with surprises, a great deal of important events and situations are eliminated which have the potential of being described as crises. The original output variable military activity acts as a measure of the influence of the system on its external environment. On the other hand, the new output operating procedures act as a measure of change in the internal environment of the system to meet the perceived threat.

Accepting the above variables as those which offer promise in describing crisis and crisis situations, their interrelationship is described by the *elliptic umbilic* catastrophe, having 2 behavioral outputs (military action and operational preparedness) and 3 inputs (perceived threat, decision time and degree of uncertainty). A detailed discussion of the implications of this model is given in Ref. 17 and will not be repeated here. Let us just sketch the bifurcation set $B$ in parameter space of this model (see Figure 7).

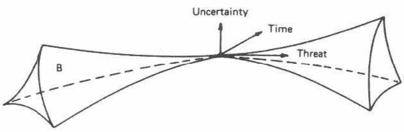



FIGURE 7

To enter $B$ is to create potential crises, while to leave $B$ is to create a shift in behavior space, i.e. a crisis. Note, however, that the above model distinguishes between a crisis and a crisis situation. A crisis is an instantaneous change in behavior, i.e. a discontinuity in behavior. A crisis situation places the system in a state of "alert". Basically, any point in $B$ corresponds to a crisis situation, while crossing the boundary of $B$ *may* bring on the crisis.

Before closing this example, it is worthwhile to point out that the ideas sketched above apply in *any* crisis management situation, not just in the military context we have chosen. By replacing our output variable "military action", with another "action" variable, the formalism of the above model goes over for other crisis situations such as business management, psychological traumas or epidemic disease control.

Now let us turn attention to another type of international conflict which can be analyzed using the topological tools of $q$-analysis. Consider the long-standing Arab-Israeli dispute over territory in the Middle East. At the most basic level, this dispute can be viewed as a relationship between the countries involved (i.e. the Arab nations, Israel and the PLO) and the various issues (e.g. Israeli occupation of the West Bank, free access to Jerusalem, return of the Golan Heights, etc.). Thus, the basic sets $X$ and $Y$ for our $q$-analysis will be taken to be

$$X=\{x_1, x_2, \ldots, x_{10}\}=\{\text{issues}\},$$

where

$x_1=$ autonomous Palestinian state in the West Bank and Gaza,

$x_2=$ return of the West Bank and Gaza to Arab rule,

$x_3=$ Israeli military outposts along the Jordan River,

$x_4=$ Israel retains East Jerusalem.

$x_5=$ free access to all religious centers,

$x_6=$ return of Sinai to Egypt,

$x_7=$ dismantle Israeli Sinai settlements,

$x_8=$ return of Golan Heights to Syria,

$x_9=$ Israeli military outposts on Golan Heights,

$x_{10}=$ Arab countries grant citizenship to Palestinians choosing to remain within their borders.

The set of participants is

$$Y=\{y_1, y_2, \ldots, y_6\},$$
$$=\{\text{Israel, Egypt, Palestinians, Jordan, Syria, Saudi Arabia}\}.$$

The relation $\lambda \subset Y \times X$ which we shall employ is

$$(y_i, x_j) \in \lambda \leftrightarrow \text{participant } y_i \text{ is neutral or favorable toward goal } x_j.$$

The incidence matrix for $\lambda$ is

| $\lambda$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | $x_8$ | $x_9$ | $x_{10}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $y_1$ | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 |
| $y_2$ | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 |
| $y_3$ | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| $y_4$ | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 |
| $y_5$ | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 |
| $y_6$ | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |

Examination of the complex $K_Y(X;\lambda)$ shows that the most likely negotiating partner for Israel is Saudi Arabia, which is neutral or favorable on all issues except one. However, both Egypt and the Palestinians are nearly as likely candidates since they are simplices of dimension only one less than Saudi Arabia. As the Camp David talks demonstrated, Egypt is indeed a favored negotiating partner due also to psychological and other factors not incorporated into the above relation $\lambda$.

Focusing upon goals and issues, we find the high-dimensional objects in $K_X(Y;\lambda^*)$ being $x_2$ = return of the West Bank and Gaza to Arab rule, $x_5$ = free access to religious centers and $x_3$ = return of the Sinai to Egypt. These goals are viewed as neutral or favorable by all 6 participants. Therefore, they provide a good basis for a negotiated settlement of the conflict. This observation has been borne out by the Camp David talks, as well as by subsequent developments.

In addition to the above relation $\lambda$, three other cases were considered in Ref. 18: favorable only, unfavorable only and neutral only. The results of these studies confirmed that (1) Israel is highly disconnected from the other parties in the dispute, (2) Saudi Arabia is the most moderate of the Arab states, (3) Syria is by far the most rigid and inflexible and (4) the single issue which tends to bring all the parties together is free access to all religious centers.

## 6 ELECTRIC POWER SYSTEMS

As an interesting example of a physical process for which the dynamical equations are known, we consider the behavior of a collection of generators forming an electric power supply network. For a network with $n$ generators and zero transfer conductance, the equations as developed in Ref. 19 are

$$M_i \frac{d\omega_i}{dt} + d_i\omega_i = \sum_{\substack{j=1 \\ j\neq i}}^{n} E_i E_j B_{ij} \times [\sin \delta_{ij}^* - \sin \delta_{ij}],$$

$$\frac{d\delta_i}{dt} = \omega_i, \qquad i = 1, 2, \ldots, n.$$

Here

$\omega_i$ = angular speed of rotor $i$,

$\delta_i$ = electrical torque angle of rotor $i$,

$M_i$ = angular momentum of rotor $i$,

$d_i$ = damping factor for rotor $i$,

$E_i$ = voltage of generator $i$,

$B_{ij}$ = short circuit admittance between generators $i$ and $j$,

$\delta_{ij} = \delta_i - \delta_j$,

$\delta_{ij}^*$ = the stable steady-state value of $\delta_{ij}$,

$\omega_{ij} = \omega_i - \omega_j$.

Out interest is in studying the behavior of the equilibrium values of $\omega_i$ and $\delta_i$ as a function of the parameters $M_i$, $E_i$, $B_{ij}$, and $d_i$.

If we define $a_{ij} = d_i - d_j$, $b_{ij} = E_i E_j B_{ij}$, then it can be shown that the function

$$V(\omega_{ij}, \delta_{ij}) = \sum_{i=1}^{n-1} \sum_{k=i+1}^{n} [1/2 M_i M_k \omega_{ik}^2 - a_{ik}\delta_{ik} - b_{ik}(M_i + M_k)\cos \delta_{ik}$$

$$- \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n-1} \sum_{\substack{k=j+1 \\ k\neq i}}^{n} M_i b_{jk} \cos \delta_{jk} + K,$$

is a Lyapunov function for the above dynamics. We may now use the function $V$ as our basic potential function and investigate the possibility of representing $V$ by a polynomial canonical form in a neighborhood of an equilibrium.

To illustrate the catastrophe theory approach with a minimum of notational complexities, consider the simple case of $n = 2$ generators. In this case, we have only 2 basic variables $\omega_{12} \doteq x_1$, $\delta_{12} \doteq x_2$. The function $V$ then becomes

$$V(x_1, x_2) = 1/2 M_1 M_2 x_1^2 - a_{12} x_2 - b_{12}(M_1 + M_2)\cos x_2 + K$$
$$= 1/2\alpha x_1^2 - \beta x_2 - \gamma \cos x_2 + K.$$

Since addition of the constant $K$ to $V$ does not affect our problem, we set $K = 0$. Upon computing grad $V$, we find the critical points of $V$ are

$$x_1^* = 0, \; x_2^* = \sin^{-1} \beta/\gamma.$$

Computing the 4-jet of $V$ at the critical point yields

$$j^4 V(x_1, x_2) = -[\beta \sin^{-1} \beta/\gamma + \sqrt{\gamma^2 - \beta^2}] + 1/2[\alpha x_1^2 - \sqrt{\gamma^2 - \beta^2}\, x_2^2] + \frac{\beta x_2^3}{3!} + \sqrt{\gamma^2 - \beta^2}\, \frac{x_2^4}{4!}.$$

Again we drop the constant term and examine

$$J^4 V(x_1, x_2) = 1/2(\alpha x_1^2 - \sqrt{\gamma^2 - \beta^2 x_2^2}) + \frac{\beta x_2^3}{3!} + \sqrt{\gamma^2 - \beta^2}\, \frac{x_2^4}{4!}.$$

Here we see that the function $V$ is 2-determinate with codim $V = 0$ if $\alpha \neq 0$ and $\gamma \neq \pm\beta$. The condition on $\alpha$ is necessary for the problem to make sense, so the only interesting possibility for a degeneracy in $V$ occurs when $\gamma = \beta$. If $\gamma \neq \beta$, then $V$ is equivalent to a Morse function in a neighborhood of its critical point and can be replaced by its 2-jet

$$1/2\alpha x_1^2 - \sqrt{\gamma^2 - \beta^2}\, x_2^2,$$

a simple Morse saddle. So, let us assume that $\gamma = \beta$.

The function $V$ now assumes the form

$$V(x_1, x_2) = 1/2\alpha x_1^2 - \beta(x_2 + \cos x_2).$$

Computing $V_{,x_1}$ and $V_{,x_2}$ we have

$$V_{,x_1}=\alpha x_1, V_{,x_2}=-\beta+\beta\sin x_2.$$

Thus, no term of the type $\{\cos x_2\}$ can be generated by $V_{,x_1}$ and $V_{,x_2}$. All other smooth terms can be generated and we can easily find that the Hessian of $V$ has corank 1. Thus, we are in the case $r=1$, $c=1$ of Table I, which is the so-called "fold" catastrophe. We may identify the single essential variable with $x_2$, as is seen by examination of the jet $J^4V$ when $\gamma=\beta$. So, a universal unfolding of the potential function $V$ in the critical case is

$$\tilde{V}=\frac{x_2^3}{3}+tx_2,$$

where $t$ is the unfolding parameter.

Summarizing the above results, we conclude that when $\gamma\neq\beta$ the potential $V$ is equivalent to a simple Morse saddle, while in the critical case when $\gamma=\beta$, $V$ is right-equivalent to the cubic potential $\tilde{V}$. Only in the second case can we expect to find an abrupt change in the stability behavior of the power system as parameters are varied. In the canonical structure $\tilde{V}$, the unfolding parameter $t$ depends upon the physical parameters $\alpha$, $\beta$, $\gamma$, and a change in the system stability behavior will occur when $t$ passes thru the value 0. We can see the structure more clearly if we set $\gamma=\beta+\varepsilon$ and consider $J^3V$. We obtain

$$J^3V=1/2(\alpha x_1^2-\sqrt{2\beta\varepsilon+\varepsilon^2}\,x_2^2)+\frac{\beta x_2^3}{3!}.$$

Neglecting the quadratic term in $x_1$, this is an unfolding of the function $x_2^3$ and can be brought into the standard form $\tilde{V}$ by a simple coordinate change. The change will then yield $t$ as a function of $\beta$ and $\gamma$.

Adopting a more macroscopic view of the power network, we can represent the system schematically as in Figure 8. The above diagram makes it clear that the transmission network, which consists of passive elements, is a *relation* between the set of $n$ generators and the set of $m$ loads. Thus, we can apply the $q$-analysis language to describe various aspects of the connective structure of the network.

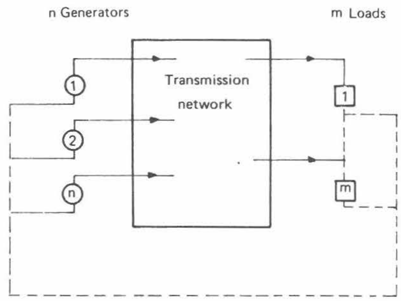



FIGURE 8

In the power system context, it is useful to think of the relation between generators and loads as a *weighted relation*, with the entries of the incidence matrix $\Lambda$ of the relation defined as

$\Lambda_{ij}$ = the fraction of the power requirement of load $j$ which is supplied by generator $i$, $i=1,2,\ldots,n$; $j=1,2,\ldots,m$.

For scaling convenience, we shall multiply each $\Lambda_{ij}$ by 100 in order to work only with integers.

To illustrate the use of $q$-analysis, let us consider a network consisting of $n=5$ generators and $m=6$ loads with the weighted incidence matrix $\Lambda$

| | $\lambda$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ |
|---|---|---|---|---|---|---|---|
| | | *X* (loads) | | | | | |
| | $y_1$ | 30 | 20 | 10 | 10 | 30 | 20 |
| *Y* (generators) | $y_2$ | 10 | 30 | 0 | 20 | 0 | 20 |
| | $y_3$ | 40 | 50 | 40 | 20 | 0 | 10 |
| | $y_4$ | 0 | 0 | 20 | 10 | 60 | 0 |
| | $y_5$ | 20 | 0 | 30 | 40 | 10 | 50 |

In order to investigate the connective structure of this distribution network, we take various views of the relation by "slicing" $\Lambda$ at different levels.

For instance, let us slice at the lowest level 1%, which consists of including any generator which supplies any amount of power to any load in our *induced* binary relation $\mu$. The relation $\mu$ is given by a binary incidence matrix $U$ as follows:

$$[U]_{ij}=\begin{cases}1, & \text{if } \Lambda_{ij}>0\\ 0, & \text{otherwise.}\end{cases}$$

Thus, at the 1% level we have

| $\mu$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ |
|---|---|---|---|---|---|---|
| $y_1$ | 1 | 1 | 1 | 1 | 1 | 1 |
| $y_2$ | 1 | 1 | 0 | 1 | 0 | 1 |
| $U(1\%)=y_3$ | 1 | 1 | 1 | 1 | 0 | 1 |
| $y_4$ | 0 | 0 | 1 | 1 | 1 | 0 |
| $y_5$ | 1 | 0 | 1 | 1 | 1 | 1 |

The complex $K_Y(X;\mu)$ has the structure vector

$$Q(1\%)=\begin{pmatrix}5 & 0\\ 111111\end{pmatrix},$$

while the conjugates complex $K_X(Y;\mu^*)$ has the structure vector

$$Q^*(1\%)=\begin{pmatrix}4 & 0\\ 11111\end{pmatrix}.$$

So, we see that both the generators and the loads are well-connected insofar as the supply and demand of some power is concerned. This fact is also seen if we calculate the eccentricities of the generators and the loads. The only elements having nonzero eccentricities are the generator $y_1$ with $\text{ecc}\, y_1=1/5$ and the load $x_4$ with $\text{ecc}\, x_4=1/4$.

Now we slice at a much higher level and consider only those connections which provide at least 20% of the total power requirement of a load. In this case, the incidence matrix $U$ is

| $\mu$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ |
|---|---|---|---|---|---|---|
| $y_1$ | 1 | 1 | 0 | 0 | 1 | 1 |
| $y_2$ | 0 | 1 | 0 | 1 | 0 | 1 |
| $U(20\%)=y_3$ | 1 | 1 | 1 | 1 | 0 | 0 |
| $y_4$ | 0 | 0 | 1 | 0 | 1 | 0 |
| $y_5$ | 1 | 0 | 1 | 1 | 0 | 1 |

In this case, the relevant structure vectors turn out to be

$$Q(20\%)=\begin{pmatrix}3 & 0\\ 3311\end{pmatrix},$$

$$Q^*(20\%)=\begin{pmatrix}2 & 0\\ 521\end{pmatrix}.$$

So, we see that at the 20% supply rate both the complexes of generators connected by loads and loads connected by generators fragment into disconnected pieces at the higher-dimensional levels. This means, for example, that any kind of load-sharing involving 20% or more of a load among 2 or more generators is impossible due to the connective structure of the network. It is also of some interest to observe the integration of individual generators into the complex $K_Y(X;\mu)$. The eccentricities at the 20% level are

$$\text{ecc}\, y_1=1,\ \text{ecc}\, y_2=1/2,\ \text{ecc}\, y_3=1/3,$$

$$\text{ecc}\, y_4=0,\ \text{ecc}\, y_5=1.$$

Thus, $y_5$ which was perfectly well integrated into the network at the 1% level, is now quite a conspicuous element at the 20% level. This indicates that $y_5$ is a much more central element in the overall power system than, say, $y_4$.

In the conjugate complex, we have

$$\text{ecc}\, x_5=1,\ \text{ecc}\, x_i=1/2,\ i=1,2,3,4,6.$$

Thus, we see that load $x_5$, with its heavy dependence upon generator $y_4$, stands out at the 20% level more than any of the other loads. This suggests that a surge in demand in load area $x_5$ is likely to cause more strain on the network than the same surge in other load areas.

The preceding ideas can be extended to a dynamical context by introducing a pattern on the complexes as was indicated in our earlier examples. For instance, a typical pattern on $K_Y(X;\lambda)$ would be the amount of power generated at each generator. The connectivity structure just presented would then enable us to study the possibility of redistributing the power from a given generator to other loads in the event of a disruption in some part of the system. Analogously, a pattern on the conjugate complex involving the demand at each load would provide us with the basis for investigation of how to reallocate power when surges in demand occur.

## 7 RESILIENCE, TIME AND SURPRISE

Lurking just below the surface in much of the preceding discussion and examples is the idea that in many social, as well as natural systems some seemingly small unknown, or even unknowable disturbances will occur and cause the system to enter into a state of structural and/or dynamical collapse. Indeed, part of the *raison d'etre* for the development of catastrophe theory was to provide a mathematical foundation for speaking of such behavioral discontinuities. It is evident that the need for such a theory of systemic stability is greater now than ever before, as mankind faces an ever-increasing sense of vulnerability to risks, uncertainties and the unknown. The Three-Mile Island and Love Canal incidents represent only the tip of the iceberg of potential disasters that lie in wait for society if we are unable to provide a design and management methodology that enables a system to not only persist and respond to the unknown, but to actually benefit from it.

It has been argued elsewhere,[20] that a comprehensive theory of system "resilience" must include predictive, regulative and adaptive components. Based upon a variety of case studies involving forest pests, fisheries, forest-fire systems, human disease and savannah ecosystems, it was concluded in Ref. 20 that "all management policies succeeded in the short term; all failed in the longer term and produced a crisis; all owed the failure to a successful effort to reduce variability; and some adapted successfully to the failure while others did not." Furthermore, in assessing the ingredients needed to avoid the types of crises mentioned above,

> ". . . the key questions are what are the sources of surprise, which ones are most difficult to deal with, how do people, and institutions interpret and respond to surprise, and how can we design and manage failure and surprise adaptively? . . . . Hence, resilience becomes the capability to adapt to surprise because of past experience of instabilities . . . ."

Thus, we see a pressing need for a systematic *theory of surprises*, warning us when our conceptions of reality fail to match the real world. When couched in these terms, it is reasonably clear that such a theory is far removed from traditional arguments in probability theory and statistics. While a comprehensive theory of surprises is far from a reality, at the present, it is our contention that many of the necessary components can be expressed through the catastrophe and $q$-analysis languages we have employed above.

Let us begin by consideration of the qualitative features of a system which "surprise theory" must capture.

A. *Bifurcation*—by definition, surprise occurs when reality and our conception of reality part company. In effect, this means that *our* model of a system and the "*real*" model bifurcate from each other. To speak meaningfully of such a bifurcation, we must develop a mathematical theory which enables us to say when one model bifurcates from another and, when this happens, to give a measure of the magnitude of the bifurcation. It will be noted below that the type of bifurcation (discontinuities) seen in catastrophe theory, are only a special case of this more general bifurcation of models.

B. *Qualitative Time*—no theory of surprise can avoid addressing the observed fact that different events in a system occur on widely-varying time-scales and, as a consequence, time and rates of change are central factors in the understanding and design of resilient systems. In fact, we shall go further and argue that the classical Newtonian view of time expressed in the *Principia* as: "Absolute, true and mathematical time, of itself, and by its own nature, flows uniformly on, without regard to anything external," is woefully inadequate for capturing the notion of *duration* between events outside the realm of classical physics. Our contention will be that the dimensional quality of events (as expressed through $q$-analysis) has temporal significance and must be taken into account when speaking of the occurrence or non-occurrence of those events.

Here again we see that classical probabilistic arguments, having their basis in the absolutism of Newtonian time, cannot possibly capture the structural difference between the natural times associated with high- and low-dimensional events. The qualitative structure of time given below also allows us to speak precisely about graded rates of change of variables, taking explicit account of their dimensional levels. Such a distinction provides a firm footing for the somewhat vague "fast-slow" distinction often seen in elementary applications of catastrophe theory, especially in the behavioral sciences.

C. *Adaptation and Evolution*—in traditional engineering design and optimization systems are developed which are unforgiving of error. Success is measured by the reduction of variability of important quantities, usually by some type of error-controlled feedback mechanism. There are two basic problems associated with such an approach: (1) the underlying system is not static and often evolves into an unexpected structure for which the regulation scheme is inappropriate, and (2) the concept of error-controlled regulation means that if there is no error, there is no regulation, i.e. if there is no variation in the state variables, then the controller become inactive and may become incapable of responding to an unexpected change in the operating environment. What is needed is a control policy that is more tolerant of error and provides for the continued exploration of alternative actions and objectives based upon a learning process stimulated by *induced* variation in the system. In short, we need an adaptive control policy.

Since prediction, regulation and adapative control have been extensively developed in the engineering and system theory literature for a number of years we shall say no more about these matters here, but focus the balance of our attention upon some system-theoretic ideas which appear relevant to developing the theories of system bifurcation and time need to formalize our concept of surprise.

Let $\Omega$ represent a collection of admissible input functions (decisions) to our system and let $f:\Omega\to\Gamma$ be an input/output map with $\Gamma$ the set of output functions. Thus, $f$ is what is termed an *external* description of the system. We can introduce an equivalence relation $R_f$ into $\Omega$ by defining $(\omega_1,\omega_2)\in R_f$ if and only if $f(\omega_1)=f(\omega_2)$. The input/output description $f$ gives rise to the set of equivalence classes of inputs $\Omega/R_f$, which is termed the *state set* $X$ of the system. Thus, the states are elements of the form $[\omega]$ denoting the class of all inputs giving rise to the same output as $\omega$. Let us assume that the set $\Gamma$ is a metric space so that we may induce a metric onto $X$ as follows: $\|[\omega_1], [\omega_2]\|_X=\|f(\omega_1)-f(\omega_2)\|_\Gamma$. Under this metric, there is a homeomorphism between $X$ and the set of values (outputs) of $f$, i.e. the values of $f$ parametrize the states $X$. Through the metric $\|\cdot\|_X$ we will then say that two states are "close" if and only if their corresponding $f$ values are close.

We can make exactly the same arguments if we have a collection of descriptions $F=\{f_1,f_2,\ldots,f_n\}$ of the system at our disposal. We generate an equivalence relation by $(\omega_1,\omega_2)\in R_F$ if and only if $f_i(\omega_1)=f_i(\omega_2)$ for each $i=1,2,\ldots,n$. We then parametrize the quotient set $\Omega/R_F$ and impose a metric on this space which shows that two states $[\omega_1]$ and $[\omega_2]$ are close if and only if $f_i(\omega_1)$ and $f_i(\omega_2)$ are close for each $i=1,2,\ldots,n$.

What is important to recognize is that any set of descriptions $F$ gives rise to a particular way of characterizing the system and captures some aspect of the reality of the system. We shall be concerned with the way in which alternate descriptions can be compared with each other.

In Thom's view of catastrophe theory, ideas of the above type are used to discuss organic form. He considers a set $E$ of "geometric objects" parametrized by a manifold $S$ through a mapping $\eta: E\to S$. Here, $E$ corresponds to our set $\Omega/R$ ($=X$) and the parameter set $S$ corresponds to the space of values of the maps $f\in F$. Thom then says that a point $a\in S$ is *generic* if, for all $a'$ close to $a$, $a'$ has "the same form" as $a$. The set of generic points is clearly open in $S$ and the complement of this open set is the closed set of *bifurcation* points.

In Thom's geometric view, we are clearly in the situation of comparing two different descriptions of the elements in $E$. On the one hand, we have a description arising from the parametrization. On the other hand, we have a tacit description summed up in the use of the words "same form". The determination of the "form" of an object in $E$ can only come about from an alternate description of those objects, one which is independent of the description arising from the parametrization.

*Example* 1 (The Cusp Catastrophe) Here we take $E$=set of all curves representing cubic equations in one variable. In appropriate coordinates we can write

$$E=\{x^3+ax+b\colon a,b \text{ real numbers}\}.$$

Thus, we can regard $E$ as being parametrized by the set

$$S=\{(a,b)\colon a,b \text{ real}\}=R^2,$$

and $\eta: E\to S$ is a map from $E\to R^2$. On the other hand, these curves may also be described by their root structure relative to the $x$-axis. They may have a single real root, repeated roots or 3 distinct real roots. We define a mapping $\phi$ from the set of all cubic curves as

$$\phi(C)=\begin{cases}0, & \text{if } C \text{ has one real root}\\ 1, & \text{if } C \text{ has a repeated real root}\\ 2, & \text{if } C \text{ has three distinct real roots.}\end{cases}$$

Then, according to the $\phi$ description, two cubic curves $C_1$ and $C_2$ are "close" if $\phi(C_1)=\phi(C_2)$.

If we use $\phi$ to define the intrinsic "topology" on $E$ and look at the generic points of $R^2$ induced by $\eta$, we find that these comprise all $(a,b)$ for which $4a^3+27b^2\neq 0$, i.e. the bifurcation points of the $\eta$ description relative to the $\phi$ description lie on the cusp, which is the complement of the generic set. On the other hand, if we use $\eta$ to define the intrinsic topology, we find that each of the three points $\{0,1,2\}$ is a bifurcation point of the $\phi$ description relative to the $\eta$ description, i.e. a point where two curves are no longer "close" in the $\phi$ description, while remaining "close" in the $\eta$ description. Here, of course, "close" in a particular description is defined in terms of the topology in $X$ associated with that description.

*Example* 2 (Dynamical Systems) Consider the class $E$ of all dynamical systems defined on some differentiable manifold $M$. Suppose that the dynamics are given (locally) by a system of differential equations

$$\dot{x}_i=f_i(x_1,x_2,\ldots,x_n),\ i=1,2,\ldots,n.$$

There are two independent ways in which such a system can be described: (1) in terms of the functions $\{f_i\}$ and (2) in terms of the asymptotic properties of the trajectories determined by the differential equations.

In the first description, two dynamical systems $\Sigma_1=(f_1,f_2,\ldots,f_n)$ and $\Sigma_2=(g_1,g_2,\ldots,g_n)$ are regarded as close if each $f_i$ is close to the corresponding $g_i$ in some appropriate norm in function space. Usually, we use the so-called Whitney $C^1$-topology, which gives the distance between $\Sigma_1$ and $\Sigma_2$ as

$$\rho(\Sigma_1,\Sigma_2)=\max_x \sum_{i=1}^{n} \left(|f_i-g_i|+\sum_{i,j=1}^{n}\left|\frac{\partial f_i}{\partial x_j}-\frac{\partial g_i}{\partial x_j}\right|\right).$$

On the other hand, in the second type of description, $\Sigma_1$ and $\Sigma_2$ are close if corresponding trajectories in $M$ are close for all $t$.

In the usual investigations of structural stability, the "intrinsic" topology on $E$ is taken to be that which is imposed by the closeness of corresponding trajectories. Comparing this description with that given by the Whitney topology, the resulting bifurcation set (i.e. the points which are *not* close in the Whitney topology but are close in the intrinsic topology) lies in a space of parameters determined by the functions $\{f_i\}$ which map $M\to R^1$. On the other hand, if we choose the Whitney topology as the intrinsic one, then the bifurcation set lies in a set of parameters determined by mappings of $R^1\to M$ (representing the corresponding trajectories).

Now let us return to the situation in which the system $\Sigma$ is described by a map $f:\Omega\to\Gamma$. Then any other description $g$, which is a continuous function of $f$ induces the same metric on $X$ ( $=\Omega/R_f$) that $f$ does, i.e. if $E$, $E'\in X$ are close under $f$, then they are also close under $g$. In this case, the description of $\Sigma$ provided by $g$ is *redundant* to that provided by $f$, since every point of $\Gamma$ is a generic point for either description with respect to the topology induced by the other. In short, we obtain no new information about $\Sigma$ by supplementing the $f$-description with the $g$-description.

The point of using an alternate description is to gain new information about $\Sigma$. Thus, another criterion for the equivalence of two descriptions $f:X\to S$, $g:X\to\bar{S}$, where $S$, $\bar{S}$ are arbitrary parameter sets, is that the bifurcation sets in $S$, $\bar{S}$ induced by the pairs $(g,f)$ and $(f,g)$, respectively, shall be empty. In general, given the descriptions $f$ and $g$, the generic points of $S$ and $\bar{S}$ represent elements of $X$ for which the two descriptions coincide. On the other hand, on the bifurcation points, the two descriptions differ and we gain information about the corresponding elements of $X$ by employing both descriptions. On the generic points of either descriptions with respect to the other, the properties of the second description are hidden; they reveal themselves only on the bifurcation points.

The question arises as to how we should go about incorporating both descriptions $f$ and $g$ into a new description which improves upon them both. The obvious way is to take their cartesian product. We form the product $S\times\bar{S}$ and describe a given element $[\omega]\in X$ by the pair $[f(\omega),g(\omega)]$. We give $S\times\bar{S}$ the product topology, so that two elements $[\omega]$, $[\omega']$ are close if and only if $f(\omega)$ is close to $f(\omega')$ in $S$ and simultaneously $g(\omega)$ is close to $g(\omega')$ in $\bar{S}$. We can now state the following abstract characterization of when the description $f$ improves upon the description $g$:

IMPROVEMENT THEOREM *$f: X \to S$ is an improvement upon $g: X \to \bar{S}$ if and only if every point of $S$ is generic relative to the topology imposed on $X$ by $g$, while the bifurcation set in $\bar{S}$ arising from the topology imposed on $X$ by $f$ is not empty.*

Let us briefly discuss the meaning of the above results for Example 1. There we had two descriptions of the family of cubic curves in a single variable: (1) via the coefficients $(a, b)$ and (2) via the numbers 0, 1, 2 representing the root structure of the curve. Thus, a combination of these two descriptions yields the cartesian product $S \times \bar{S} = R^2 \times \{0, 1, 2\}$ and a description $\theta: X \to S \times \bar{S}$. Thus, in the $\theta$-description, we associate with every cubic curve $C$ the triple $(a, b, z)$, where $(a, b)$ are the coefficients of $C$ and $z$ represents the root structure (Remark: not every such triple corresponds to a cubic curve). In the $\theta$-description, two curves $C$ and $C'$ are close if and only if $(a, b)$ and $(a', b')$ are close in $R^2$ and simultaneously $z = z'$. Thus, in the new $\theta$-description, the cubic curves corresponding to coefficients lying on the bifurcation set (i.e. on the cusp $4a^3 + 27b^2 = 0$) can be close *only to each other*, and not to any cubics determined by coefficients off the bifurcation set. Thus, we see by the Improvement Theorem that the description $\theta$ is an improvement over both the descriptions $\eta$ and $\phi$.

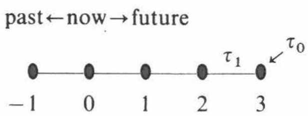
Similar remarks apply to Example 2 making use of the two alternate descriptions given for a dynamical system $\Sigma$.

Returning now to our discussion of "surprises", we noted earlier that a surprise occurs when our *conceptions* of reality fail to match reality. Interpreting a conception of reality as one description $f$ (or model) for reality, while thinking of reality itself as given by another description $g$, it is reasonably clear that a surprise occurs when the $g$-description bifurcates from the $f$-description. In other words, when two states are close in our conception of a situation, but fail to remain close in the real situation, then we may interpret this bifurcation as a surprise.

Operationally, we never really know the description of our situation which corresponds to reality. All we have in practice are two or more alternative descriptions (models) which we can compare with each other. Thus, we must choose one model as the "base" model, and interpret bifurcations of the other models from the base as the surprises. One systematic way of choosing the base model is to start with a description $f$ and then "refine" $f$ to a model $g$ so that the relation $R_g$ refines $R_f$, i.e. each $f$-equivalence class is a union of $g$-equivalence classes of inputs. In this way, two states which are close in $g$ must be close in $f$, but not conversely. Thus, the model $g$ may exhibit surprises relative to the description provided by $f$. More details about all of the above procedures, as well as a discussion of related issues may be found in Ref. 21.

Any viable theory of surprise must contain some component for dealing with the observed qualitative nature of time. The intuitively felt notions of "time flies" or "time drags" must manifest themselves on the change of pattern associated with the dynamics on the basic geometric structure representing our system. We find it convenient to represent these various types of time through the medium of $q$-analysis in the following manner.

Consider the classical Newtonian view of time, which can be represented by the diagram




where the numbers represent the measurement of specific moments of time. The lines connecting up the vertices of time measurement represent our sense of time duration. This picture represents a very elementary sort of simplicial complex, having an infinite set of vertices with the numbers attributed to each vertex forming a pattern $\tau^0$ on the vertices, i.e. the 0-simplices. The numbers which we assign to the edges joining the vertices form another pattern $\tau^1$, referred to as the time intervals between successive moments of measurement. Thus, in the Newtonian view the time pattern is the *graded* pattern

$$\tau = \tau^0 \oplus \tau^1.$$

The representation above makes it clear why we refer to the Newtonian view of time as a *linear* concept associated with the simplicial complex $K$ consisting of a set of 1-simplices which are 0-connected. When the Newtonian time-axis is used to represent a set of observed events, the idea behind it is to somehow produce a kind of "clock", whose time moments (the vertices) can be put into a $1-1$ correspondence with the set of events. The pattern $\tau^0$ describes the "now" events, while $\tau^1$ describes the interval pattern.

In the relativity theory of Einstein, the classical Newtonian pattern above was replaced by a new and different one. The new structure was a consequence of the physical role played by the light signal and had the structure

$$\tau=\tau^3\oplus\tau^4.$$

Thus, now the time moments are represented by 3-simplices, while the time intervals are corresponding 4-simplices (see Figure 9).

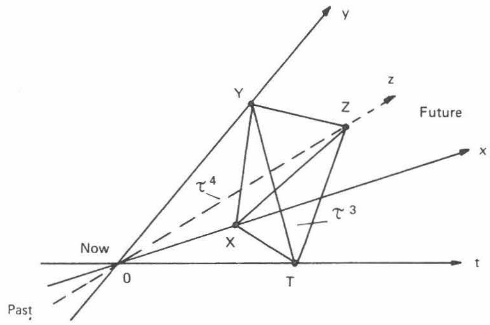



FIGURE 9

So, in the relativistic view of physics there will be 1-simplex intervals as well as 2- and 3-simplex intervals between appropriate "now" moments. A more detailed discussion of these points can be found in Ref. 22.

The foregoing considerations suggest that we introduce the following structural definition of time.

DEFINITION On any given simplicial complex $K$, *time* is a specific type of graded pattern satisfying the following conditions:

i) time imposes a *total ordering* of all $p$-simplices in $K$ (so that we can speak of one $p$-event preceding another);

ii) the time-pattern is of the form

$$\tau=\sum_{p=0}(\tau^p\oplus\tau^{p+1}).$$

The pattern $\tau^p$ describing the now-traffic of $p$-events on $K$ can be understood as a simple 0/1 function on the $p$-simplices of $K$. The location of the single non-zero value identifies the particular $p$-event which is the now-moment. Then the "next $p$-event" in $K$ is experienced by a change in this pattern via $\delta\tau^p$, so that the 1-value is found on a different $p$-simplex. This movement of values throughout $K$ clearly involves the topology (connectivity) of $K$, since one $p$-event cannot follow another unless there is an available $(p+1)$-interval connecting the two. Thus, $\delta\tau^p$ is a *p-force* in the structure of events representing our sense of moving time.

The commonly accepted assumption of Western culture is that when we use the word "time" we are referring to the Newtonian time pattern $\tau=\tau^0\oplus\tau^1$. This convention is certainly useful in providing a common frame of reference, yet it disguises the essential nature of our experience of time. Thus, suppose that an individual experiences a time traffic of dimension $p$ $(p>0)$. Then he finds it culturally necessary to replace this by the Newtonian $\tau$ as follows:

$$\tau=\tau^p\oplus\tau^{p+1}\rightarrow\tau^0\oplus\tau^1$$

(experienced time) (Newtonian time).

This means that the individual experiences a $(p+1)$-force of repulsion as far as the interval pattern is concerned, having to force his $(p+1)$-perception of the time interval down into the 1-dimensional interval of the Newtonian pattern. The experience of this force is expressed by such phrases as "time drags" or "time flies", indicating the fact that $\tau^{p+1}$ and $\tau^1$ are out of step.

Since a "$p$-event" in $K$ will generally correspond to the recognition of a $p$-simplex in $K$, let us assume, for the moment, that there is a $1-1$ correspondence between the 1-simplices of $K$ and the Newtonian reference frame. Then the time-intervals for the gap between one $p$-event and another will be proportional to the number $(p+1)(p+2)/2$ since this is the number of edges in the least connection between two $p$-simplices. Here we have assumed that a $p$-event occurs by way of the edges which make up the $(p+1)$ event which bridges the current and previous $p$-events. As an example, if the $\tau^1$ unit (the Newtonian time) is 1 day, then a 6-event would require $(6+2)(6+1)/2=28$ days to "arrive". Similarly, a $p=25$ event, would require $(27)(26)/2=351$ days, almost a full year, i.e. the interval between successive 25-events is 351 days.

We have assumed above that recognition of a $p$-event occurs by way of the edges which connect it to another (the "next") $p$-event. If, on the other hand, we assume the worst case, that we cannot recognize a $p$-event until all of its *faces* have been separately recognized, we would have to recognize all intervals between the successive $0, 1, 2, \ldots, p$ events. This number is the sum

$$\sum_{t=1}^{p+1} \binom{p+2}{t+1}\binom{t+1}{2} = (p+2)(p+1)2^{p-1}.$$

Thus, now if $\tau^1 = 1$ day, we have that a $p=5$-event would take 672 days $\cong 2$ years to arrive. Although we shall not pursue it here, an argument can be made for associating this "worst" case with the question of moving up a hierarchy from level $N$ to level $N+1$, since a $p$-event at $(N+1)$ will generally correspond to a $q$-event at level $N$, with $q \gg p$.

At a fixed $N$-level structure, the total range of time-interval patterns will be contained in the sequence of numbers 1, 3, 6, 10, 15, 21,... derived from the values of $(p+1)(p+2)/2$. These numbers represent the apparent intervals of conventional time ($\tau^1$) required for the recognition of 0-events, 1-events, etc. When we speak of the future we conventionally refer to the sequence of 0-events in the Newtonian structure $\tau = \tau^0 \oplus \tau^1$. But, in our structural view of time we can see that many $p$-events at the now-point cannot manifest themselves until much later on the Newtonian scale. Thus, there is an obvious sense in which the "future" events are already contained in the "present" events. But an ability to be sensitive to higher order $p$-times by seeing successive $p$-events would be manifest as insight into this future.

We note in closing that the multidimensional time theory outlined above can be employed in our bifurcation-based theory of surprise in at least two different ways. First of all, we can use the time theory to assign definite meaning to the problem variables as "fast", "slow", "intermediate", etc. and regard their variation as a manifestation of certain $p$-events. Since we already know where the bifurcation surfaces are in the parameter space, the time theory enables us to predict when these surfaces will be crossed, causing one description to bifurcate from another.

An alternative route would be to formulate the basic system descriptions in simplicial complex terms. The dynamics are then represented as a change of pattern

$$\delta\pi = \delta\pi^0 \oplus \delta\pi^1 \oplus \ldots \oplus \delta\pi^m$$

as already discussed. Now we could introduce the multidimensional time factor by attempting to extend the classical notion of a differential equation by setting

$$\frac{\Delta\pi^p}{\Delta\tau^p} = h_p(\sigma_p^1, \sigma_p^2, \ldots, \sigma_p^s),$$

where $\Delta\pi^p$ is the change in the $p$-pattern, $\Delta\tau^p$ is the change in the $p$-time pattern and $\{\sigma_p^i\}$ are the $p$-simplices of the complex $K$. Here, our task would be to develop the appropriate functions $h_p$ characterizing the corresponding dynamics.

## REFERENCES


1. W. Helly, *Urban System Models*. Academic Press, New York, 1975.
2. T. Saaty, *The Analytic Hierarchy Process*. McGraw Hill, New York, 1980.
3. P. Doreian and N. P. Hummon, *Modeling Social Processes*. Elsevier, New York, 1976.
4. J. Casti, "Systemism, System Theory and Social System Modeling." *Regional Science and Urban Economics*, **11**, 1981, pp. 405–424.
5. P. Saunders, *An Introduction to Catastrophe Theory*. Cambridge Univ. Press, Cambridge, 1980.
6. T. Poston and I. Stewart, *Catastrophe Theory and its Applications*. Pitman, London, 1978.
7. E. C. Zeeman, *Catastrophe Theory: Selected Papers*. Addison-Wesley, Reading, MA., 1977.
8. R. Thom, *Structural Stability and Morphogenesis*. Addison-Wesley, Reading, MA., 1975.
9. H. Sussman and R. Zahler, "A Critique of Applied Catastrophe Theory in the Behavioral Sciences." *Behavioral Science*, **23**, 1978, pp. 383–389.
10. R. H. Atkin, *Mathematical Structure in Human Affairs*. Heineman, London, 1973.
11. J. Casti, *Connectivity, Complexity and Catastrophe in Large-Scale Systems*. John Wiley, Chichester and New York, 1979.
12. R. H. Atkin, *Combinatorial Connectivities in Social Systems*. Birkhäuser, Basel, 1977.
13. R. Gilmore, "Catastrophe Time Scales and Conventions." *Physical Rev. A.*, **20**, 1979, pp. 2510–2515.
14. C. McClelland, "System Theory and Human Conflict." In: *The Nature of Human Conflict*, edited by E.-B. McNeil, Prentice-Hall, Englewood Cliffs, NJ, 1965.
15. C. F. Hermann, "International Crises as a Situational Variable." In: *International Politics and Foreign Policy*, edited by H. Rosenau, Free Press, New York, 1969.

16. W. Phillips and T. Lorimor, "The Effect of Crises Upon the Stability of the International System." Dept. of Political Science, Ohio State University, Columbus, 1972.
17. W. Phillips and R. Rimkunas, "The Concept of Crisis in International Politics." Dept. of Gov't and Politics, University of Maryland, College Park, 1976.
18. J. Casti and M. Shakun, "Goals/Values Geometry and Conflict Resolution." On-going research, Grad. School of Business, New York University, 1978.
19. *Modern Concepts of Power System Dynamics.* IEEE Special Publ. # 70MG2-PWR., IEEE, New York.
20. C. S. Holling, "Resilience in the Unforgiving Society." Report R-24, Institute of Resource Ecology, University of British Columbia, Vancouver, 1981.
21. R. Rosen, *Fundamentals of Measurement and Representation of Natural Systems.* Elsevier, New York, 1978.
22. R. H. Atkin, *Multidimensional Man.* Penguin, London, 1981.

**John Casti** is currently a research mathematician at the International Institute for Applied Systems Analysis, Laxenburg, Austria and Professor, Department of Systems Engineering, University of Arizona, Tucson, Arizona. Previously, Dr. Casti held professorships at New York University, Princeton University and SUNY, Binghampton, as well as research positions at Systems Control, Inc. and the Rand Corporation.

Dr. Casti received his Ph.D. in Mathematics (1970) at the University of Southern California under Professor R. E. Bellman and has pursued research interests in control theory, dynamic programming, numerical solution of boundary-value problems and integral equations. His current research interests are in the areas of dynamical system theory and its applications in natural resource problems are the social and behavioural sciences.

In addition to over 75 research articles, Dr. Casti has also authored several research monographs including *Dynamical Systems and their Applications in Linear Theory* (Academic Press, 1977), *Connectivity, Complexity and Catastrophe in Large Systems* (Wiley-Interscience, 1979) and *Principles of Dynamic Programming*, vols. I and II, (with R. Larson, Dekker, 1978, 1982). In addition, Dr. Casti edits the journal *Applied Mathematics and Computation* (Elsevier) and the two book series, *Control and System Theory* (Dekker) and *Frontiers of Systems Research in the Social Sciences* (Nijhoff).