EMERGENT NOVELTY AND THE MODELING OF SPATIAL PROCESSES

J. Casti
Systems Engineering Department, University of Arizona, Tuscon, AZ 85721, USA,
and
International Institute for Applied Systems Analysis, Laxenburg, Austria

RR-83-27
December 1983

Reprinted from *Kybernetes*, volume 12 (1983)

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.

Reprinted with permission from *Kybernetes* 12:167–175.

Printed by Novographic, Vienna, Austria

FOREWORD

Contributions to the Metropolitan Study: 2

The project "Nested Dynamics of Metropolitan Processes and Policies" was initiated by the Regional and Urban Development Group in 1982, and the work on this collaborative study started in 1983. This paper deals with one of the central theoretical issues in the project: the relation between different scales of aggregation of urban dynamics.

An important aspect of many urban modeling situations is to ascertain the ways in which interactions at the level of individuals are combined to produce observable global spatial and/or temporal patterns. Of equal interest is the inverse question: given an observed global pattern, classify all local interactions of choice processes that could give rise to the specified patterns.

Such questions form an important part of the Metropolitan Study inasmuch as they involve the development of housing, employment, and economic patterns in time as functions of individual choices. This paper provides the systems-theoretic background needed to investigate the evolution of such patterns and the bifurcations that generate a shift from one pattern to another.

A related theme of interest is the stabilization of a derived pattern by means of social policies. The paper shows that stabilization questions are bound up with the twin problems of adaption and anticipation in the policy-making process, and several suggestions for accommodating these features into the mathematical model of the process are discussed.

BÖRJE JOHANSSON
Acting Leader
Regional and Urban Development Group

Kybernetes
1983, Vol. 12, pp. 167–175

EMERGENT NOVELTY AND THE MODELING OF SPATIAL PROCESSES†

J. CASTI

Systems Engineering Department, University of Arizona, Tucson, AZ 85721 (USA), and IIASA, Laxenburg, A-2361 (Austria)

(*Received: January 7, 1983*)

In this paper we consider the complementary questions: in what sense do local dynamics prescribe global spatial patterns and to what extent does a global pattern impose constraints on local interactions. From the standpoint of results from mathematical system theory, it is argued that a modeling approach starting from observed patterns and passing to local dynamics is vastly to be preferred to proceeding in the opposite direction, the usual approach mimicking the procedure followed in the physical sciences. The paper concludes with a discussion of the role of anticipatory decision making and adaptation in the stabilization of certain properties of dynamical spatial processes.

1 EMERGENCE OF FORM

One of the most vexing questions in all of biology is the determination of the mechanism whereby new "forms" spring-up into existence almost overnight, seemingly with no direct connection to previous lines of development. Such a discontinuous morphogenesis is often termed "emergent novelty" in theoretical biology and its presence is certainly one of the strongest arguments against the classical Darwinian theory of evolution. In more general system-theoretic terms, the problem of emergent novelty may be re-stated in the following manner: under what circumstances do "small" changes in the structure of a system result in "large" changes in the pattern or form of the observed output? Despite the obvious vagueness of this problem statement, the relevance to spatial choice models is fairly clear, since in spatial choice phenomena one of the key issues is to what extent the microbehaviour of individuals influences the overall spatial pattern. We are also interested in the converse question: in what fashion does a global spatial pattern constrain the local actions of individuals. To deal with these questions in more specific terms, we must examine the relationship between local and global models and spatial patterns.

For us, a spatial pattern will consist simply of a distribution of population throughout some predefined spatial region. If we describe the population density at a point (x, y) at time t by $a(x, y, t)$, then a *local* model for the change of pattern will consist of a description of how the population at (x, y) interacts with that at nearby points $(x \pm \Delta x, y \pm \Delta y)$ and the manner in which this interaction, together with the external influences through the boundary $\partial\Omega$ determines the density at time $t + \Delta t$. Usually this local description is given by a partial differential equation of diffusion-type. For example we could use the equation:

$$\frac{\partial a}{\partial t} = a[f(\rho) - ag(\rho)] + D_1 \nabla^2 a - D_2 \nabla^2 \rho, \quad (1)$$

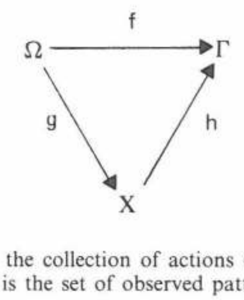
where $\rho(x, y)$ is the "affinity" of the point (x, y) and f and g are functions specifying the effect of ρ on the stimulation and inhibition of proliferation. Here the basic mechanisms accounting for the change of spatial pattern are proliferation, diffusion and environmental affinity, with diffusion acting to homogenize the pattern, while proliferation and affinity tend to make the pattern more heterogeneous. The net effect of these counteracting factors is determined by the particular functions f and g, and the parameters D_1 and D_2.

Equation (1) is a typical example of a local model in that the only interactions utilized in the formulation of the model are those in an arbitrarily small neighborhood of the point (x, y). These local interactions are dictated by the choice of f, g, D_1 and D_2 with the hope that if the selection of these elements is made judiciously, then the global spatial pattern and its dynamics can be determined from Eq. (1). Such a "bottom-up" approach is, of course, very familiar from classical physics and forms the mathematical modeling manifestation of a part of the reductionist view of natural phenomena. In the next section we shall consider in more detail some of the pitfalls inherent in such a "myopic" view of system modeling. Here we only note that the credibility of

† Work partially supported by the National Science Foundation under Grants # CEE 8110778, "Modern Stability and Numerical Concepts in Water Resource Management" and # CEE 8100491, "Global Stability Analysis of Environmental Systems".

any such local model is totally dependent upon whether or not the basic ingredients (the elements f, g, D_1 and D_2) can be accurately determined from the spatial process and the available data. In physics the approach works (usually) since we have the luxury of basic "laws" to fall back upon and such laws can be invoked to give precise structure to the model elements. In the social sciences no such laws exist and, as a result we are on much more shaky ground when it comes to choosing the necessary components of the local model. This point has been extensively pursued in reference 1, so we shall not belabor it here, other than to say that the absence of natural laws in the social sciences casts a shadow over any modeling effort relying upon detailed knowledge of the form of local interactions. Without further effort to discover such laws (if indeed they exist, at all), the "physics-envy" of the social scientist has basically the same degree of substance as the smile of the Cheshire cat.

If we accept the position that local interactions are unknown, then the only recourse is to a "top-down" modeling effort in which we attempt to *construct* the local interactions on the basis of the *observed* temporal evolution of the spatial pattern. In other words, we seek to deduce the laws of microbehavior from the observed pattern. This approach is more in line with the classical tradition in experimental science, where a controlled experiment is performed, the data is analyzed and a theory (read: model) is constructed to explain the experimental evidence. The essence of such a "top-down" view of modeling is to regard the experimental data (the system inputs and outputs) as constituting the *observables* of the process and to consider the local interactions as taking place through the system *states*, which are constructed from the input/output data. What is important to note here is the idea that the state variables are not given *a priori*; they have no intrinsic physical significance and are interpretable only through their effect on the observed outputs and by the way they interact with the applied inputs to produce outputs. In short, the states only exist at the microlevel and cannot, in general, be seen as part of the system "pattern." Diagramatically, we have the following situation:



Here Ω is the collection of actions (inputs) to the system, Γ is the set of observed patterns (outputs) and f is the observed input/output relationship. The set X represents the microstates of the process, with g and h being the system input and output maps, respectively. The general modeling problem is then: given Ω, Γ and f, find X, g and h such that the above diagram commutes.

The overall problem of pattern dynamics, and especially the phenomenon of emergence, is now easy to interpret: under what circumstances can the dynamics in Γ change discontinuously as a result of changes in Ω, X, g and/or h. Obviously this is a very general statement of the problem and we shall take-up special cases later but it is already clear that the phenomenon of emergence is not dependent on any one factor. It is a complicated interplay between the internal dynamics in Ω and X, as well as the way in which these dynamics impose dynamics in Γ. Furthermore, since the sets Ω, X and Γ are, in general, quite distinct we see that it is necessary to account for the possibility of several different time scales in the process, at least one for each of these sets and maybe more if the basic sets themselves are stratified in some way.

A simple illustration of how emergent phenomena can arise is provided by some empirical work done by Schelling[2] on the problem of choice in the overall racial integration of urban neighborhoods. The approach involved dividing the population into two discriminatory groups (blacks and whites) and to impose a rectangular cellular grid upon the neighborhood. Each grid cell represented one urban dwelling and was either occupied by a black (#), a white (0) or was vacant. The local dynamics were defined by the simple principle that each racial group would prefer to have a certain percentage of its immediate neighbors being of the same group and, if this was not the case, then that party would move to the nearest grid location where the percentage of like neighbors was acceptable. In order to have a reasonable choice of where to move, it was empirically observed that around 25%–30% of the grid locations should be vacant. Starting with the initial distribution of Fig. 1, the steady-state pattern

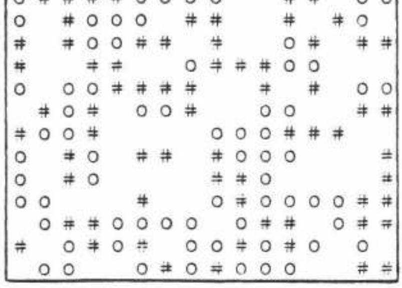


FIGURE 1 Initial distribution.

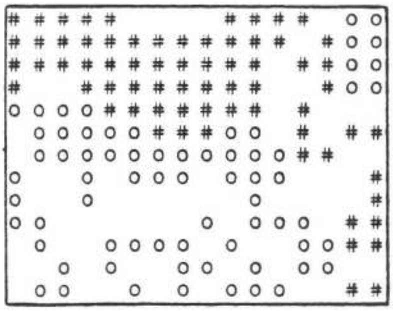


FIGURE 2 Steady-state distribution when at least half of the neighbors are the same.

of Fig. 2 was obtained under the assumption that at least half of one's immediate neighbors be of the same color while when the demand is only at the level of 1/3 of one's neighbors be of like color, the pattern of Fig. 3 emerges.

Many other variants of the basic parameters are reported in [2], with the general conclusion being that no simple correspondence of individual incentives to the final collective pattern could be discerned. Exaggerated patterning and separation results from the dynamics of movement and that the overall pattern did not in general allow any inferences to be drawn concerning individual motives. Such a conclusion, while somewhat pessimistic from a sociologist's vantage point, has a rather straightforward interpretation and explanation when viewed in the light of mathematical system theory. We shall take this point up in Section III. For now, it is sufficient just to note the main moral of Schelling's work: local dynamics cannot be *uniquely* inferred from observed patterns. However, in Section III we shall show that the non-uniqueness can be removed by imposing some very reasonable natural conditions on the local model.

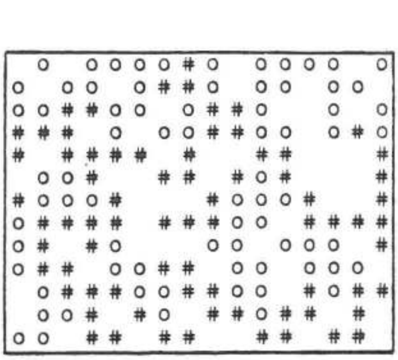


FIGURE 3 Steady-state distribution when at least 1/3 of the neighbors are the same.

II MICROBEHAVIOR TO GLOBAL PATTERNS

In order to study the effect of microbehavior upon global patterns of development, we shall consider various cases of the dynamical system (N)

$$\dot{x} = f(x, u), \quad x(0) = x_o$$
$$y(t) = h(x, u),$$

where x is an n-dimensional vector of microstates, y is a p-dimensional vector of observed outputs, u is an m-dimensional vector of control actions (including environmental parameters) and f and h are functions defining the internal dynamics and output, respectively. Usually in "bottom-up" modeling, especially in the social sciences, it is *assumed* that h = identity, implying that what we can actually measure is the microstate itself. Actually, the situation is somewhat worse since in most local models the microstates are *defined* to be those quantities like population for which data is available. Since one of our principle points is that it is important on system-theoretic grounds to distinguish between variables which can and cannot be measured directly, we shall insist upon not confusing the concept of system *state* with that of measured system *output*, a confusion which is inherent in implicity taking h as the identity function.

Now let us see the "kinds" of patterns which can arise in simple local models by examining a few cases from the literature. Our objective will be to show by concrete illustration that seemingly minor variations on the assumptions defining the local behavior can generate global patterns which have no apparent relations to each other.

Example 1: The development of socioeconomic inequalities[3]

Here we consider a model proposed to explain how inequalities can arise and be maintained in a given social system. Let the population be divided into n subgroups and define the variables

$a_i(t)$ = self-enhancing advantage for subgroup i, (e.g. generalized "wealth"),
$s_i(t)$ = resources available to subgroup i, (e.g. energy, manpower, etc.)
$p(a_i, s_i)$ = production of advantage for group i,
$r(a_i, s_i)$ = removal of advantage from group i, (by depreciation and consumption),
$D_a(a_i)$ = re-distribution of the overall advantage to group i (e.g. taxation),
$q(a_i, s_i)$ = depletion of resource by subgroup i.

Under the assumption that there is a uniform availability of limited resources and a depletion effect q proportional to production p, the following

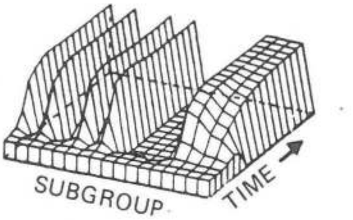


FIGURE 4 Pattern when $d_a = 0$ and there is a uniform initial advantage.

local model for the change of advantage is proposed

$$\frac{da_i}{dt} = \rho_i + \frac{a_i s}{1 + a_i/A} - a_i^v - d_a\left(a - \frac{1}{n}\sum_{j=1}^{n} a_j\right), \quad (\dagger)$$

where ρ_i is a small basic production term, A is a saturation constant, v is the overall societal wealth, d_a is a parameter reflecting re-distribution of $a'a$ is a parameter representing "reinvestment" of advantage for future generation and

$$s = \text{const} \Bigg/ \left(\sum_j \frac{a_j}{1 + a_j/A}\right).$$

In order to see how differential advantages develop, consider Figs. 4–6, which show the results of integrating the model (†) with $n = 17$ groups, $v = 0.5$ for various values of re-distribution d_a. Figure 4 shows the pattern that develops when there is a near uniform initial advantage with $d_a = 0$. A similar pattern arises in Fig. 5 if the initial advantages form a shallow gradient while, finally, Fig. 6 shows that for a sufficiently dominant d_a, the inequality generating mechanism is qualitatively changed.

The above example shows the sensitivity of the global pattern to a change in only a single problem parameter. Since such parameter values can seldom be determined with a degree of accuracy anywhere near what is needed to lend credibility (and stability)

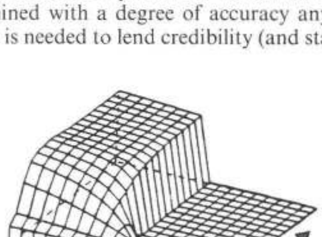


FIGURE 5 Pattern for $d_a = 0$ with a non-uniform initial advantage.

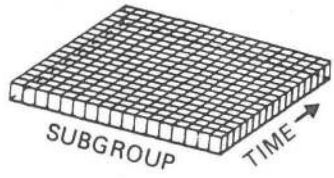


FIGURE 6 Pattern for dominant d_a.

to the model's output, the potential pitfalls of basing policy on such a model is obvious. Our next example shows that such a situation is not at all "exceptional" in the modeling of social phenomena.

Example 2: Dynamics of urban residential structure

The basic objective of this model is to describe the dynamics of residential attractiveness as a function of travel costs from one neighborhood to another, employment factors for each region and average household size. If we let

$H_i(t)$ = the residential attractiveness of zone i,
$T_{ij}(t)$ = the number of people in zone i who work in zone j,
a = employment multiplier,
b = average household size,
σ = rate of response parameter

and set

$$T_{ij}(t) = B_j E_j H_i^\gamma \exp(-\beta c_{ij}),$$

where

$$B_j = \left[\sum_i H_i^\gamma \exp(-\beta c_{ij})\right]^{-1},$$

with

E_j = number of persons working in zone j,
β = travel impedance parameter,
γ = residential attractiveness parameter,
c_{ij} = cost of travel from zone i to zone j,
$i, j = 1, 2, \ldots, n$,

then the postulated dynamics are

$$H_i(t+1) = H_i(t) + \sigma\left[\frac{a}{b}\sum_j T_{ij}(t) - H_i(t)\right] H_i(t).$$

Using the above model to study the rise and decline of neighborhoods as a function of the travel parameter β, it was seen in reference 4 that, generally speaking, two distinct patterns could emerge: for small $\beta\,(\cong 0.1)$ the population tended to cluster in the central area, while for large $\beta\,(\cong 1.0)$ the population was distributed throughout the suburbs as well as in the central area.

For us, the important feature of the above model is that it is a quadratically nonlinear difference equation. A simple example of such a function which captures all of the essential features is the logistic equation

$$x(t+1) = \mu x(t)[1 - x(t)], \quad \mu \in R^1.$$

It is well known that for values of $\mu \gtrsim 3.58$ the oscillations of definite periods between the steady-states of the logistic equation cease and the oscillations become aperiodic, i.e. *chaos* sets in.

In the "chaotic" regime, the dynamical behavior is characterized by the three properties: (1) an infinite number of periodic trajectories; (ii) an unaccountable number of nonperiodic trajectories; (iii) the instability of all trajectories. On the other hand, for $1 \leq \mu \leq 3$, there is a single stable trajectory (fixed point at $x^* = (\mu - 1)/\mu$). For $3 < \mu \lesssim 3.58$, x^* becomes unstable, but new stable steady-states come into existence. The point is that totally unexpected and qualitatively different patterns may emerge as a result of minor (unmeasurable, even) changes in basic parameters.

However, by regarding μ as a randomly fluctuating environmental parameter, an entirely different global pattern may arise as our next example shows.

Example 3: Nonequilibrium transitions induced by a noisy environment[5]

Consider the logistic equation

$$\frac{dx}{dt} = \tfrac{1}{2} - x + \beta x(1 - x),$$

where β is a parameter representing the state of the operating environment. Assume that β is a random variable with $E\beta = \beta$, $\operatorname{var} \beta = \sigma^2$ and that the correlation time for β is much less than the macro time-scale of x. This allows us to regard β as a white noise fluctuation affecting the system. If we let $P_s(x)$ represent the stationary probability density for x, an easy consequence of the Fokker–Planck equation shows[5] that the extrema of $P_s(x)$ are given by the roots of the equation

$$1/2 - x^* + \beta x^*(1 - x^*) - (1/2)\sigma^2 x^*(1 - x^*)(1 - 2x^*) = 0.$$

This is a cubic polynomial in x^* whose roots are depicted in Fig. 7 as a function of β and $\sigma^2/2$. Thus, with regard to the extrema of $P_s(x)$ we have a *cusp* catastrophe in the (β, σ^2)-plane with the cusp point at (0, 4).

The conclusion is that new transition phenomena can be induced by external fluctuations in the environment; the extrema of $P_s(x)$ are essentially different from those of the deterministic case. Hence, it makes a difference in our interpretation of the model whether we regard the parameter β as unknown, but deterministic, or as a random variable.

The purpose of introducing the above examples is not to discredit the utility of the particular models themselves, but rather to call attention to some of the inherent *potential* pitfalls that lurk in the wings. All of the types of pathologies sketched above have a single root source: an *instability* in the observed output pattern as various features of the model are changed. Unless we understand why, how and when such instabilities arise, we are "whistling in the dark" as far as being able to place any real faith in the conclusions drawn from the model. This problem is particularly acute in the use of such models in the social sciences, including spatial choice situations, since the underlying microinteractions are poorly understood. What is needed are results which

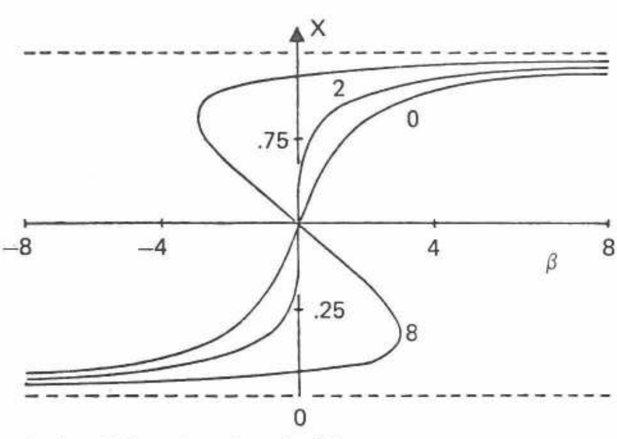


FIGURE 7 Roots x^* as a function of β for various values of $\sigma^2/2$.

enable us to assert that a particular observed output pattern is stable (persists) over some range of local interaction hypotheses. Such results require a "top-down" approach to modeling to which we devote the next section.

III LOCAL MODELS FROM MACROPATTERNS

If we renounce the modeling scheme based on deducing global patterns from assumed microbehavior, the only alternative is to construct models in which the observed micropattern imposes a *class* of local dynamics. A basic question, of both theoretical and applied concern, is to what extent the macropattern dictates a unique local dynamics. As we shall see in a moment, mathematical system theory has a rather definitive, clear-cut answer to this question: *up to an irrelevant labeling of the internal state variables, there is a single "good" local model for each observed pattern.* Now let us explain this result and its underlying assumptions in more detail.

We are concerned with the existence of "good" local models of the form

$$\frac{dx}{dt} = f(x, u), \quad x(0) = x_0,$$

(N)

$$y(t) = h(x, u),$$

when given the observed pattern consisting of the input $u(t)$ and the output $y(t)$, $0 \leq t \leq T$. In other words, given the input/output pair $(u(t), y(t))$, we consider the following basic modeling questions:

(i) under what circumstances can we *construct* a state space X and functions f and h such that (N) displays the same input/output behavior as (u, y);

(ii) if such a construction is possible, is it unique in any sense, and what are the structural features of X, f and h;

(iii) if the construction is not unique, what additional "natural" conditions must be imposed in order to remove the non-uniqueness, i.e., how do we find "good" models;

(iv) how can we make the construction of good models computationally feasible from given experimental data.

There is, as yet, no complete answer to these questions. A reasonably complete theory does exist, though, if the input/output relationship is linear[6]. In this case, it is known that a "good" model is unique up to a choice of the coordinate system in X, where "good" here means that dim X is minimal. If the input/output relationship is nonlinear, then we must be content with studying special classes of nonlinear relationships to achieve specific results (e.g., bilinear, polynomial, etc.). However, as far as the condition as to what constitutes a good model goes, the result from the linear case comes over to the *general* case almost without change. Thus, if there are two local models N and N' which both "explain" the same observed pattern, then these two models are isomorphic in the sense that there exists a diffeomorphism ϕ: $X \to X'$ such that under ϕ, $f \to f'$ and $h \to h'$. In other words, N and N' differ only through a smooth coordinate change in their state spaces. It is important to observe here that since the state space X is a mathematical *construction* obtained from the pattern, there is no possibility of ever determining from the pattern alone what the internal coordinatization of X will be. Thus, the above result is "best possible" in that once the coordinate system in X is chosen, then the local model is *uniquely* determined from the observed pattern. Many more details and additional results may be found in the survey paper.[7]

One of the main difficulties in carrying out the general modeling program discussed above is that of determining in what class of functions we should seek f and h. A hint as to how to proceed is provided by results from the theory of singularities, together with the above theorem stating that any two good models must be equivalent through a smooth coordinate change. Let us look at this idea in more detail.

Suppose that the global pattern is described by an equation of the form

$$f(y_1, y_2, \ldots, y_n; u_1, u_2, \ldots, u_m) = 0,$$

where $f(\cdot)$ is some smooth function relating the inputs $\{u_i\}$ and outputs $\{y_j\}$, i.e., f is the input/output relation. For simplicity, we assume here that f is scalar-valued. The vector-valued case can also be treated at the expense of substantial additional mathematical sophistication and notation. According to the basic theorems of Mather, Thom and Arnol'd, the form of f in a neighborhood of any point (y, u) is determined by two integers, the *codimension* and *corank* of f. These numbers can be computed directly from f by *linear algebraic* means and, once they are known, the theory of singularities assures us that there exist coordinate systems in (y, u)-space such that f assumes one of a small number of standard forms, the particular form depending upon codim f and corank f. What is important here is that all of these forms are low-order *polynomials* in y and *linear* in u. This result strongly suggests focusing attention upon the class of polynomially nonlinear systems in any efforts directed toward modeling local dynamics in spatial choice situations.

As an indication of how the ideas from singularity theory shed light on the development of spatial patterns, consider the continuous mapping ϕ of the x–y plane to itself

$$(x, y) \xrightarrow{\phi} (u(x, y), v(x, y)).$$

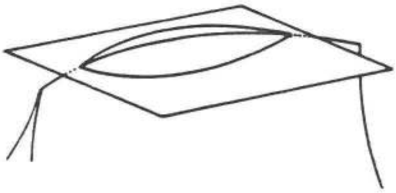


FIGURE 8 Lip events.

The singular points of ϕ are where the system Jacobian $\partial(u, v)/\partial(x, y) = 0$. These singular points form smooth curves in (x, y)-space called *fold lines* which contain special points called *cusp points*. The folds and cusps have the important character that they are *stable* features not destroyed by small perturbations of ϕ. Intuitively, we can imagine that (u, v) represents the velocity of a flow (of people, resource, etc.) at the point (x, y). If we plot the velocity at (x, y) as a dot at the point $(u(x, y), v(x, y))$, we would see the density of dots being very high along the fold lines and even higher at the cusp points. That is, the velocities are "focused" at the singularities.

Now imagine that the velocity field evolves in time, i.e., $\phi = \phi_t$. The pattern of folds and cusps will change in a continuous way, except that at certain times the folds and cusps will interact to produce an *event*. To see what might happen, consider now the map

$$(x, y, t) \overset{\phi_t}{\rightarrow} (u, v, \tau), \quad (\tau = t)$$

In general maps from $R^3 \rightarrow R^3$, the generic, stable singular sets are folds, cusps and swallowtails. However, this particular map has the property that $t = \tau$ and it turns out that this imposes certain restrictions on the orientation of the singular sets in (u, v, τ)-space. The successive planes $\tau =$ constant cut through the singular sets to produce two sorts of events as seen in (u, v): (a) when $\tau =$ constant cuts a singular point (a fold), then a swallowtail type of event occurs; (b) when $\tau =$ constant is tangent to a rib (the locus of a cusp in (u, v, τ)), then only "beak-to-beak" and "lip" events occur. These are the only types of generic events that can occur in a freely evolving flow field. Basically, these events correspond to the way an entire line of cusps can be sectioned by a plane in R^3 (see Figs. 8–9).

The implications of the foregoing result for spatial modeling is quite profound. Basically, it says that regardless of the assumptions concerning the microdynamics built into the map ϕ, the global pattern that emerges can have singularities of only a very limited type. So, if our interest is in how and where the flow field can exhibit discontinuous changes in behavior (unbounded velocities), only a small number of situations need be considered and only a handful of inequivalent geometries can occur. In fact, we can be a bit more specific about the local structure of ϕ and state the following weak form of Whitney's Theorem:[8]

Whitney's Theorem. Let p *be a singular point for the map* ϕ. *Then smooth coordinate systems* (x', y') *and* (u', v') *may be introduced about* p *and* $\phi(p)$ *such that* ϕ *takes the form*

(a) $u' = x'^2, v' = y'$, *if p is a fold point;*

(b) $u' = x'y' - x'^3, v' = y'$, *if p is a cusp point.*

Furthermore, the set of smooth maps ϕ: $R^2 \rightarrow R^2$ *having only fold or cusp singularities is dense in the set of all smooth maps* $R^2 \rightarrow R^2$.

Thus, almost *any* local dynamics, when viewed in the "right" coordinates looks like either (a) or (b) of Whitney's Theorem near the singular points. (Note that if p is a regular point, then locally ϕ takes the form $u' = x'$, $v' = y'$.)

Whitney's Theorem is another illustration of the point made earlier in the section to the effect that the global pattern actually induces a microdynamic that is unique, up to coordinate change. This result adds further strength to our earlier claim that "top-down" modeling is the preferred direction to go for characterizing situations in which there is no local "law" to fall back upon for choice of the microdynamics.

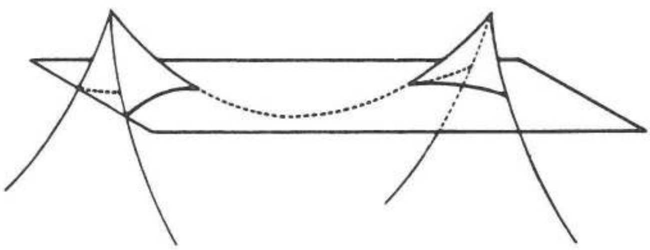


FIGURE 9 Beak-to-beak events.

IV ANTICIPATION, ADAPTATION AND MONITORING

So far we have totally ignored the question of decisionmaking, or control, in the spatial choice modeling process. One might say that we have considered only the so-called "free dynamics" of the process from the point of view of passive observers. Now we wish to become activists and consider certain aspects of the modeling operation associated with management and control of spatial choice processes.

Let us assume then that we have a working model for the free dynamics of the process given by

$$\frac{dx}{dt} = f(x),$$
$$y(t) = h(x).$$

Here x is the microstate and y is the observed output pattern. Suppose that the model predicts future outputs that are in some way undesirable and, as a result, we wish to introduce an input into the dynamics to move the system back "on course." Thus, we now have the dynamics

$$\dot{x} = g(x, u),$$

where $u(t)$ is the decision function. We note that choice of $u(t)$ is dependent upon a *prediction* of y at some future time $\bar{t} > t$ made by the original model. This is an example of *anticipatory* decisionmaking, a point we shall return to later.

There are many results in the system theory literature concerned with the way in which a feedback control $u = u(x)$ can be used to modify the original dynamics. Our concern here is with the fact that in most spatial choice phenomena, the microstate $x(t)$ needed to construct the feedback action $u(x)$ is *not available*. We have at our disposal only the macropattern $y(t) = h(x)$. This situation creates a *monitoring problem*: can we determine the microstate x from the macropattern y? The answer depends upon many factors, most importantly the function h, which specifies the manner in which the microstates are combined to form the pattern. Issues of this sort form the basis for the *theory of observability* which is well-documented in the literature.[9]

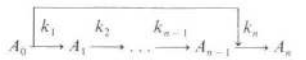
Another important aspect of the overall monitoring question is the following: is it necessary to actually construct *all* of x from y in order to form a "good" control law approximating $u(x)$ or would only a small subset of "essential" variables in x suffice? Leaving aside the issue of what constitutes a "control approximating $u(x)$," this question revolves about the issue of whether or not we can find a coordinate system in the x-space, such that in this system only a small number of original components of x appear in the feedback law. Problems of this sort were considered in reference 10 but much additional work remains. We note that the Splitting Lemma of singularity theory[8] should prove to be a valuable tool in addressing the foregoing question. The interested reader can find a more detailed treatment of many aspects of the monitoring question in the context of natural resource systems in the paper.[11]

Returning now to the issue of anticipatory decisionmaking, we have the following general situation: a model M *predicts* the future pattern $\hat{y}(T)$ of our spatial process at some time $T > t$. On the basis of this prediction, a decision $u(t)$ is taken which modifies the behavior of the system and the process unfolds to time T at which point the *actual* pattern $y(T)$ is observed. At this point the prediction $\hat{y}(T)$ and realization $y(T)$ are compared and a judgment about the credibility of the model M is made. If $\| y(T) - \hat{y}(T) \|$ is sufficiently small, M is left unchanged. If this difference exceeds some threshold, the model is *modified* on the basis of the new information $y(s)$, $t \leq s \leq T$. This process of modifying M is what we usually term *adaptation* and is, as we see, intimately connected with the ideas of prediction and observation. Under various hypotheses it can be expected that as $t \to \infty$, the difference $\| y(t) - \hat{y}(t) \| \to 0$, or at least can be reduced to some acceptable level.

An important point to note in this context is the role played by the decision $u(t)$, which is based not only upon the past observations $y(s)$, $s \leq t$, but also upon the future observations as predicted by M. In effect, the action $u(t)$ is acting as a "probe" as well as a "controller." The probing has the effect of modifying the behavior of the system in order to learn about its structure. The knowledge gained by probing is then employed in transforming $M \to M'$. Details of how this adaptation is carried out in various settings may be found in reference 12.

As a final example, let us consider in detail a particular anticipatory system of the type alluded to above in order to exhibit the manner in which such systems act to stabilize properties of given processes *without* the use of feedback loops. Consider first the chemical reaction sequence depicted in Fig. 10.

FIGURE 10 Chemical reaction sequence.



where the k_i is the reaction rate for the ith reaction. Further, we assume there is a forward activation step in this sequence, so that the concentration of A_0 serves to activate the production of A_n, i.e., the concentration $A_0(t)$ *predicts* the concentration $A_{n-1}(t+h)$ at a later time $t+h$. Thus, we choose $k_n = k_n(A_0)$ to embody the forward activation step of the system. All other reaction rates k_i are constant, $i = 1, 2, \dots, n-1$.

With the above hypotheses, the rate equations for the system are

$$\frac{dA_i}{dt} = k_i A_{i-1} - k_{i+1} A_i,$$
$$\frac{dA_n}{dt} = k_n(A_0) A_{n-1}.$$
$$i = 1, 2, \ldots, n-1.$$

Let us assume that the purpose of the forward activation step is to prevent accumulation of A_{n-1} in the face of ambient fluctuations in the initial substance A_0. In fact, we shall take this requirement to be that $\frac{dA_{n-1}}{dt} = 0$, independently of A_0. Thus, from the rate equations we require that the equality

$$k_{n-1} A_{n-2} = k_n(A_0) A_{n-1}$$

hold. We shall attempt to achieve this condition by choosing the functional form of k_n, which will then embody the predictive model implicit in the forward loop.

It is easy to see that

$$A_{n-2}(t) = \int_0^t K_1(t-s) A_0(s)\, ds,$$
$$A_{n-1}(t) = \int_0^t K_2(t-s) A_0(s)\, ds,$$

where K_1 and K_2 are functions determined entirely by the rate constants k_i. These expressions show explicitly that the value of A_0 at a given moment determines the values of A_{n-1} and A_{n-2} at *later* instants.

The control condition above now becomes

$$k_n(A_0) = k_{n-1} \left[\frac{\int_0^t K_1(t-s) A_0(s)\, ds}{\int_0^t K_2(t-s) A_0(s)\, ds} \right], \quad (*)$$

i.e., the reaction rate k_n at any time t is determined by the value of A_0 at a prior instant $t - h$ or, equivalently, the value of $A_0(t)$ determines k_n at some *future* time $t + h$. Thus, we see the manner in which the initial substance A_0 serves to *adapt* the pathway so as to stabilize the condition that

$$\frac{dA_{n-1}}{dt} = 0.$$

Finally, we observe that the homeostasis maintained in the pathway is obtained entirely through the modeling relation between A_0 and A_{n-1}, by virtue of the relation (*) which links the prediction of the model to the actual rate k_n. That is, the homeostasis is preserved entirely through adaptation generated on the basis of a *predicted* value of A_{n-1}. In particular, there is no feedback in the pathway and no mechanism available for the system to "see" the value of the quantity which is in fact controlled.

While we shall not pursue the point here, it is worthwhile to note that the anticipatory linkage between A_0 and A_{n-1} will be adaptive only as long as the relation (*) holds, i.e., only as long as the linkage "wired in" by the forward activation step and the actual linkage dictated by the chemical kinetics remain the same. If there should be a departure, then the rate change given by (*) will become *maladaptive* to a degree measured by the magnitude of the departure. Since such deviations can *always* be expected to occur in real processes, we conclude that a forward activation step of the above type can retain its adaptive character only for a characteristic time dependent upon the nature of the larger system within which it is imbedded and upon the character of its interactions with that larger system. This phenomenon may be termed *temporal spanning* and has no analogue in non-anticipatory systems. It is our contention that this property may allow us to more deeply understand in a unified way many puzzling features of biological, social and behavioral processes.

REFERENCES

1. J. Casti, "Systemism, System Theory and Social System Modeling" *Reg. Sci. and Urban Econ.* **11**, 405–424 (1981).
2. T. Schelling, "Dynamic Models of Segregation" *J. Math. Sociology* **1**, 143–186 (1971).
3. A. Gierer, A. "Socioeconomic Inequalities: Adaptation and Application of a Theory of Biological Pattern Formation" in: *Pattern Formation by Dynamic Systems and Pattern Recognition* (H. Haken, ed.)(Springer, Berlin, 1979).
4. J. Beaumont, M. Clarke and A. G. Wilson, "The Dynamics of Urban Spatial Structure: Some Exploratory Results Using Difference Equations and Bifurcation Theory" *Environ. and Plan. A.* **13**, 1473–1483 (1981).
5. W. Horsthemke, "Nonequilibrium Transitions Induced by External White Noise and Colored Noise" in: *Dynamics of Synergetic Systems* (H. Haken, ed.) (Springer, Berlin, 1980).
6. J. Casti, *Dynamical Systems and their Application: Linear Theory* (Academic Press, New York, 1977).
7. J. Casti, "Recent developments and future perspectives in Nonlinear System Theory" *SIAM Rev.* **24**, 301–331 (1982).
8. Y. C. Lu, *Singularity Theory* (Springer, New York, 1976).
9. R. Hermann and A. Krener, "Nonlinear Controllability and Observability" *IEEE Tran. Auto. Control* **AC-22**, 728–740 (1977).
10. J. Casti and A. Letov, "Minimal Control Fields" *J. Math. Analysis and Applic.* **43**, 15–25 (1973).
11. J. Casti, "Forest Modeling, Monitoring and Harvesting Policies" *Applied Math. and Computation* (to appear 1983).
12. S. Yakowitz, *Mathematics of Adaptive Control Processes* (Elsevier, New York, 1969).