# A CATEGORICAL APPROACH TO GENERAL SYSTEMS

Susanna Ginali

Dept. of Computer Science, Illinois Institute of Tech.

Chicago, Illinois, U.S.A.

Joseph Goguen

Dept. of Computer Science, University of California

Los Angeles, California, U.S.A.

## 1. INTRODUCTION

The purpose of this paper is to sketch a simplified version of Goguen's general systems theory [4], [2], [3] using graphs and sets instead of categories insofar as possible, accompanied by intuitive motivation and simple examples. We also introduce basic theorems relating some fundamental concepts of system theory but space limitations preclude much background information and all proofs. These details will appear elsewhere.

Many approaches have been taken to systems theory, illuminating the subject from various interesting angles, but often with a major division between qualitative and quantitative methods of analysis. At present, many important systems are too poorly understood to construct state transition models for them and yet, their basic components and interrelationships are often known and could be expressed precisely, if a formalism which did not require complete information about all components were available. This paper presents a formalism in which qualitative information can be expressed precisely and in which it can naturally be combined with the usual quantitative information. This has been demonstrated for musical structure by Goguen [1].

The formalism presented here includes, as special cases, diverse (and apparently incompatible) quantitative models such as discrete and continuous time (or differential) systems, and input and state variable models. This is accomplished by defining a system to be an interconnection of objects, thus making no commitment to any particular system configuration and allowing the use of objects of arbitrary character. This permits a precise yet truly general systems theory.

One might suspect on the basis of past experience that such generality condemns the theory to triviality. However, this is far from true; it is only through sufficient generality that some of the most fundamental questions of systems theory can even be asked; and the answers are far from trivial.

There is no space to list all of the influential papers, but the names of Lotfi Zadeh [5], Saunders Mac Lane [6], William Lawvere, and Gordon Pask cannot be omitted.

## 2. PRELIMINARIES

We assume standard set theory and mathematical notation. In addition, the following definitions are particularly basic: a graph G consists of a set E of edges, a set V of vertices or nodes, and two functions from E to V which give (respectively) the source and target of each edge; these functions are denoted $\partial_0$ and $\partial_1$ where $\partial_0 e$ and $\partial_1 e$ denote (respectively) the source and target of edge e. A labelled graph has, in addition, a pair of labelling functions, g and h say, from E to E' and V to V' (where E' is the set of edge labels and V' is the set of vertex labels). Sometimes the labels also have a graph structure, that is, functions $\partial'_0$ and $\partial'_1$. In this case we require that $\partial'_i(ge) = h(\partial_i e)$ for i=0,1 and all e in E; this guarantees that the labelling is consistent with the source and target designations. A consistent labelling (g,h) may also be called a graph morphism.

We sometimes need the intuitive idea of a category as a class of structural models of the same kind (e.g., all sets, or automata, or vector spaces, or Markov processes) together with all structure preserving mappings among them (these can be thought of as simplifications and/or embeddings of models); the models are called objects, and the mappings are called morphisms. Now, a category has, in particular, the structure of a graph because each morphism has a source and target. If f is a morphism with $\partial_0 f$ = A and $\partial_1 f$ = B, we write f:A→B. The additional structure which constitutes a category is that morphisms f:A→B and g:B→C can be composed to produce a morphism gf:A→C, and each object B has an identity morphism $1_B$:B→B (which models B in itself) such that $1_B f$ = f and $g1_B$ = g. In addition, composition is assumed to be associative.

Categories are denoted by double underlined capitals and the class of objects of a category $\underline{\underline{C}}$ is denoted by $|\underline{\underline{C}}|$. In the text, we will mention but not explicate some important categorical constructs such as diagram, limit, and colimit. For details, see ADJ [7] or Mac Lane [6].

## 3. OBJECTS

Our approach considers object to be the most primitive concept and allows objects to be almost anything at all. By not restricting the theory to any particular class of basic objects from which to build systems, we avoid the limitations of each class (and any arguments about which class is better). Furthermore, we are free to construct for each application a class of objects particularly tailored to its needs with no awkward translations to an artificially general class of given objects. The power of this feature for general systems theory should not be underestimated.

Usually, an object will be defined in terms of a collection of observations of some selected characteristics of the object over some collection of intervals. Conversely, such a collection of observations is taken to define an object and the pool of all possible observations over all possible intervals from which the collection defining the object is taken is called its basis. In this exposition, all objects in any given system will have the same basis and, in some contexts, it will be helpful to view a basis as the range of the "variables" in the system.

Example 1 Let $\underline{R}$ denote the set of real numbers and let $\underline{\underline{I}}$ be the set of subsets of $\underline{R}$ of the form $\{r/a\leq r\leq b\}$; these are called closed real intervals. $\overline{R}$ is defined to be the basis $\{f:I\to\underline{R}/I\in\underline{\underline{I}}$ and f has continuous derivatives of all orders$\}$.

Example 2 Let $\omega$ denote the set of non-negative integers and let $\underline{\underline{I}}$ be the set of subsets of $\omega$ of the form $\{i/0\leq i\leq n$ for $n\in\omega\}$. Suppose A is an alphabet. The basis $A^* = \{w:I\to A/I\in\underline{\underline{I}}\}\cup\{e\}$ where e denotes the unique function with empty domain and A as co-domain.

Example 3 Here we construct a basis for objects which will be used to describe a (restricted) model of piano music as in Goguen [1]. We first define sets of designators for pitch, volume, and duration as follows: $P = \{1,2,\ldots,88\}$, for the keys of a standard piano; $V = \{p,f,m,mp,mf,pp,ff,ppp,fff\}$; and $D = \underline{N}\cup\{j/2^n/n\in\underline{N}$ and $1\leq j<2^n\}$, where $\underline{N}$ is the set of positive integers. Then $E = \overline{P\times V\times D}$ is the set of elementary musical "events" (i.e. "notes") out of the context of a score and $E^*$ is the set of all sequences of notes, i.e., "pieces of music." The elements of our basis will be situated sequences of notes; to provide a notion of context, we shall enrich our representations with a pair of integers, indicating the bound-

aries of the subsequence under particular consideration. Thus the basis Ph = $\{(M,n,p)/M\in E^*, 1\leq n\leq p\leq \#M$ where $\#M$ is the length of $M\}$ is the set of phrases within the context of a musical score. In particular, (M,n,p) is the phase consisting of the n+1'st to the p'th note in the sequence M.

Example 4 In the examples above, observations were of single attributes only. The following basis has as its collection of selected attributes all finite sets. Taking $\underline{\underline{I}}$ as in Example 2, we define the basis Fin = $\{f:I\to A/I\in\underline{\underline{I}}$ and A is a finite set$\}$.

In our set and graph theoretic framework, we define a B-object to be a subset of a product $B^n$ of the basis B. Actually, we must allow objects to be subsets of arbitrarily nested products of B but this technicality results from the limitations of set theory and does not arise in a categorical setting. Also, we are primarily interested in those sets of observations drawn from the basis that are somehow "coherent." The categorical setting is needed for a satisfying description of coherence and the interested reader is referred to Goguen [2].

Example 5 Some R-objects.

a. An adder.

$$+ = \{(x,y,z)/x,y,z:I\to R \text{ and } z = x+y \text{ on } I\in\underline{\underline{I}}\}$$

b. An a-scalor for $a\in R$.

$$a = \{(x,y)/x,y:I\to R \text{ and } y = ax \text{ on } I\in\underline{\underline{I}}\}$$

c. A differentiator.

$$D = \{(x,dx/dt)/x:I\to R \text{ on } I\in\underline{\underline{I}}\}$$

d. A capacitor of C Farads.

$$C = \{(e_1,e_2,i)/e_1,e_2,i:I\to R \text{ and } i = C(d(e_1-e_2)/dt) \text{ on } I\in\underline{\underline{I}}\}$$

Example 6 Some A*-objects.

a. A delayor with initial condition $a_0$ where $a_0\in A$.

$$\text{Del} = \{(x,x')/x,x':I\to A \text{ and } x'(0) = a_0, x'(t) = x(t+1) \text{ on } I\in\underline{\underline{I}}\}$$

b. To show that objects need not necessarily be causal, we consider a device whose output anticipates its input (from alphabet A).

$$An = \{(x,x')/x,x':I\to A \text{ and } x'(t) = x(t+1) \text{ on } I\in\underline{\underline{I}}\}$$

Example 7 The input-output observations of a finite state machine $M = (X,S,Y,\delta,\lambda)$ where X,S,Y are the input, state, and output sets respectively and $\delta:S\times X\to S$ and $\lambda:S\to Y$ are the state transition and output functions, is a Fin-object defined by

$$E = \{(x,y)/x:I\to X,y:I\to Y \text{ and there exists } s:I\to S \text{ with } s(t+1) = \delta(s,(t),x(t)) \text{ and } y(t) = \lambda(s(t)) \text{ on } I\in\underline{\underline{I}}\}$$

Example 8 Some interesting Ph-objects are those which correspond to relations among situated sequences of notes, for example:

a. Succession is a relation defined by the Ph-object

$$; = \{((M,n,p),(M,p,q),(M,n,q))/M\in E^*, 0\le n\le p\le q\le \#M\}$$

This puts together the consecutive phrases (M,n,p),(M,p,q) of M to form the phrase (M,n,q).

b. Similarly, repeat is defined by

$$: = \{((M_0MM_1,\#M_0,\#M_0M),(M_0MMM_1,\#M_0,\#M_0MM))/M_0,M,M_1\in E^*\}$$

We will see later how combining objects like these into systems yields structured descriptions of pieces of music in which the relational objects behave in the expected way. See Goguen [1] for a detailed description of this class of objects.

Many further examples are found in Goguen [2] which also introduces various formalizations of the notion of object. Two of the most relevant ideas are

i) A given class of objects should have morphisms associated with it so as to yield an object category.

ii) Many classes of objects have a naturally associated structure category. For example, the R-objects in Example 5 all have a certain kind of linear (or vector space) structure.

## 4. SYSTEMS

A system arises by connecting together objects. Consider the musical form AAB where A and B are variables ranging over all musical phrases. This form consists of the phrase A repeated and then followed by the phrase B. The system for the form AAB, written in an informal "black box" notation, is shown in Figure 1. The directions of the arrows are meant to indicate input-output

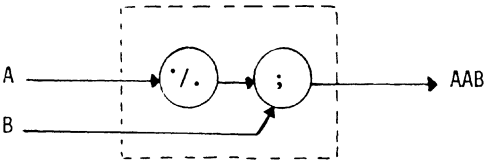



Figure 1 Black box representation of the musical form AAB.

relations; for example, the arrows out of A and B suggest that A and B are input variables and the arrow into AAB suggests that it is an output variable. However, not all systems express input-output or causal relationships and therefore we abandon black box notation and develop instead a more general, formal notation for systems.

A B-system, where B is a basis, as represented in this notation, is a graph whose vertices are labelled by B-objects, where an occurrence of the basis B itself may be viewed as a variable. (For convenience, we often use a hybrid notation, where occurrences of B are labelled instead by the particular variable corresponding to that occurrence.) The arrows of the graph are labelled by projections whose domain is the label of the arrow's source and whose co-domain is the label of its target. Thus, only vertices labelled B may have arrows into them and we further require that any vertex whose label is an n-ary object must have n arrows emanating from it. (Obviously this definition assumes objects which are subsets of a simple product $B^n$ but the extension to subsets of arbitrarily nested products of B is straightforward.) The formal representation for the musical form AAB as a Ph-system is shown in Figure 2a and its representation in the hybrid notation where each occurrence of Ph is replaced by its corresponding variable is given in Figure 2b.

Example 9 Some continuous time (or differential) R-systems.

a. The parallel connection of two capacitors represented in black box notation is given in Figure 3a. The same system in hybrid R-system notation is given in Figure 3b where $C_1$ and $C_2$ are capacitor R-objects and + is the adder. The projections labelling the arrows are left implicit.

b. A predator-prey system. Let X = number of wolves, Y = number of rabbits, a = natural growth rate of wolves, d = natural growth rate of rabbits, b = induced growth rate of wolves from eating rabbits, c = induced death rate of rabbits from being eaten by wolves (a negative number). This system is conventionally represented by the pair of differential equations

$$dX/dt = aX + bY$$
$$dY/dt = cX + dY$$

The R-system representation for this system is given in Figure 4 where a,b,c,d are scalors, + is the adder, and D is the differentiator.

Example 10 Let $M = (X,S,Y,\delta,\lambda)$ be a finite state machine as in Example 7. The state variable model M represented as a Fin-system is shown in Figure 5. Note that X*, S*, Y*, the free monoids generated by X, S, and Y respectively, are Fin-objects (see Example 2). $\delta^*$ and $\lambda^*$ are the Fin objects

$$\delta^* = \{(x,s,s')/x:I\to X,s,s':I\to S \text{ where } s'(t+1) = \delta(s(t),x(t)) \text{ on } I\in\underline{I}\}$$

$$\lambda^* = \{(s,y)/y:I\to Y,s:I\to S \text{ and } y(t) = \lambda(s(t)) \text{ on } I\in\underline{I}\}$$

The B-system formalism is not merely a matter of notation for it specifies a system as a diagram in an object category; that is, a graph whose vertices and edges are consistently labelled by objects and morphisms (respectively) drawn from the object category. As a consequence, very natural and powerful categorical definitions of the behavior of a system and of an interconnection of systems are available to us along with relevant categorical theorems. Two such theorems, relating interconnections of systems and their behavior, will be introduced in later sections of this paper.

## 5. BEHAVIOR

One fundamental problem of systems theory is analysis; that is, given a system-objects and connections-what can it do? It is an important corroboration of our notion of a system that it has a very natural notion of behavior which we will introduce via an auxiliary notion called a consistent net of points. Suppose S is a (finite) B-system with underlying graph G, where x and y are vertices in G, and $f:x\to y$ is an edge in G. Let $S_x$ and $S_y$ denote the objects labelling x and y and let $S_f:S_x\to S_y$ denote the projection labelling edge f. A consistent net of points in S is a tuple $p = (p_1,\ldots,p_n)$ such that $p_x\in S_x$ for each vertex x in G and

if $f:x \to y$ is an edge in G, then $S_f(p_x) = p_y$.

Example 11 As an illustration we will work out a consistent net of points for the musical form AAB whose representation as a Ph-system is given in Figure 2. The underlying graph for this system is given in Figure 6. Suppose

$$p_1 = (M_1AM_3, \#M_1, \#M_1A)$$

$$p_2 = (M_2BM_4, \#M_2, \#M_2B)$$

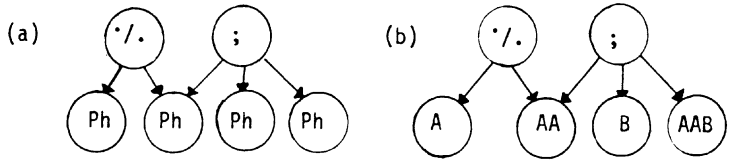



Figure 2 Formal (a) and hybrid (b) Ph-system representation of the musical form AAB.

In order for $(p_1,\ldots,p_6)$ to be a consistent net of points, it must be that

$$p_3 = ((M_1AM_3, \#M_1, \#M_1A), (M_1AAM_3, \#M_1, \#M_1AA))$$

$$p_4 = (M_1AAM_3, \#M_1, \#M_1AA)$$

$$p_5 = ((M_1AAM_3, \#M_1, \#M_1AA), (M_2BM_4, \#M_2, \#M_2B), (M_1AABM_4, \#M_1, \#M_1AAB))$$

where, because of the definition of ; , $M_2 = M_1AA$ and $M_3 = BM_4$.

$$p_6 = (M_1AABM_4, \#M_1, \#M_1AAB)$$

The projections meeting at $S_4$ force the projection of the second component of $p_3$ to be identical with the projection of the first component of $p_5$.

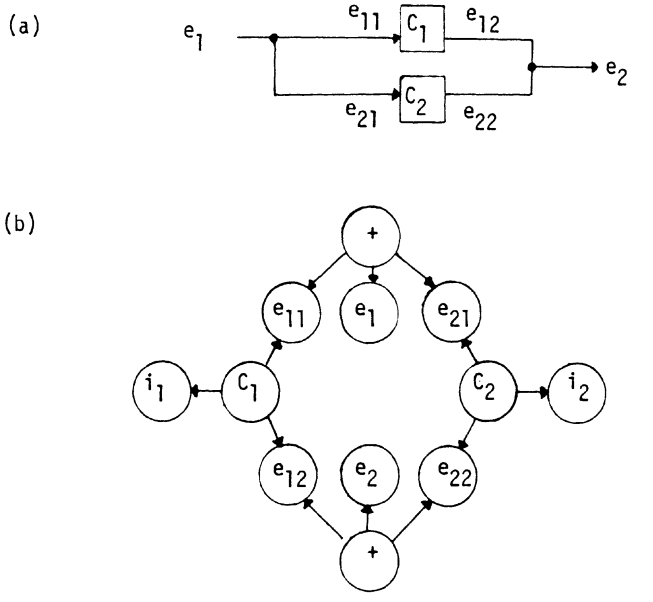



Figure 3 Black box (a) and hybrid R-system (b) representations of the parallel connection of two capacitors $C_1$ and $C_2$.

The behavior of a system S is defined by Beh(S) = $\{(p_1,\ldots,p_n)/(p_1,\ldots,p_n)$ is a consistent net of points in S$\}$. Note that if S is a B-system, then Beh(S) is itself a B-object, a very useful property of the behavior as later theorems will demonstrate. Categorically, the behavior of a system S (now viewed as a diagram in an object category $\underline{\underline{O}}$) is the limit of the diagram S as computed in the category $\underline{\underline{O}}$. Our set theoretic behaviors always exist as objects; however, limits do not exist for all diagrams in all categories. Note that limits of (finite diagrams) in the category of finite sets are computed just as above, i.e. as the set of consistent nets of points in the diagram and hence always exist.

Clearly, the definition of Beh(S) as the set of consistent nets of points in S produces rather unwieldy sets, even for small systems. Since we are often interested in the observations of only some of the objects in a system, (for example, input-output observations of a finite state machine), we introduce relative behavior, that is, the behavior of a system relative to a set of objects in the system. Thus, let $V_0$ be a set of vertices in the underlying graph of a system S. The behavior of S relative to $V_0$

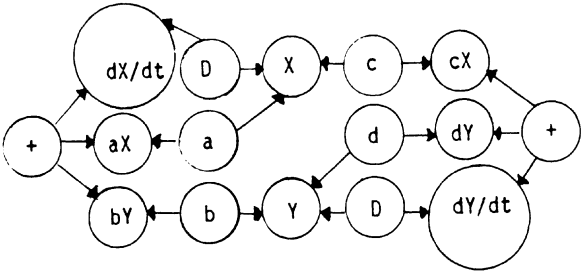



Figure 4 Hybrid R-system representation for a predator-prey system.

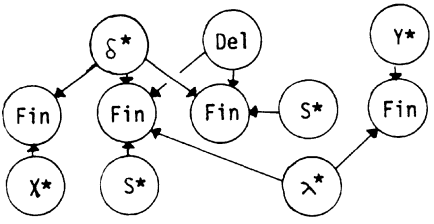



Figure 5 Fin-system representation for a finite state machine $M = (X,S,Y,\delta,\lambda)$.

is the set of restrictions of consistent nets of points in S to those components corresponding to vertices in $V_0$.

The relative behavior of systems may be compared to determine relative equivalence of systems. Let $S_0$ denote a set of objects common to the B-systems S and S'. Then S and S' are equivalent relative to $S_0$ iff the behavior of S relative to the vertices in S corresponding to the objects in $S_0$ is identical with the behavior of S' relative to the vertices in S' corresponding to the objects in $S_0$. In case $S_0$ includes all objects in S', and S and S' are equivalent relative to $S_0$, then S simulates S'.

## 6. INTERCONNECTIONS OF SYSTEMS

This paper presents the view that an understanding of a system involves an analysis of it into an interconnection of simpler components. An interconnection of systems is a special instance of a system of systems and can be readily described in the present framework given the appropriate choice of basis. A system of B-systems is simply an S(B)-system where S(B) denotes the basis whose elements are B-systems. An interconnection of systems will be done over common subsystems and the result of the interconnection of two systems over a common subsystem $S_0$ is the system obtained by "pasting" the copy of $S_0$ in one system over the copy of $S_0$ in the other system. Such a process can obviously be extended to an interconnection of a number of systems over various common subsystems.

Example 12 The musical form AABAAB is obtained by following A with itself and then with B and, finally, repeating the result. This form, shown in Figure 7b, is the result of interconnecting the three systems shown in Figure 7a. The interconnection is done over the subsystem common to $S_1$ and $S_2$ consisting of the single object Ph (corresponding to AA in $S_1$ and C in $S_2$) and over the subsystem common to $S_2$ and $S_3$ consisting of the single object Ph (corresponding to CB in $S_2$ and F in $S_3$).

In the categorical framework, an interconnection of systems with object category $\underline{\underline{O}}$ (that is, an interconnection of $\underline{\underline{O}}$-systems) is a diagram in the category $S_i(\underline{\underline{O}})$ of $\underline{\underline{O}}$-systems. The morphisms in $S_i(\underline{\underline{O}})$ are strict system morphisms (roughly, subsystem inclusion morphisms). The result of the interconnection of $\underline{\underline{O}}$-systems specified by the diagram T in $S_i(\underline{\underline{O}})$ is colimit T as computed in $S_i(\underline{\underline{O}})$. In the intuitive framework, the result of an interconnection of B-systems always exists as a B-system. Surprisingly, this is also the case for the categorical formulation as asserted by the following theorem.

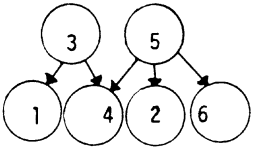



Figure 6 Underlying graph of the Ph-system in Figure 2.

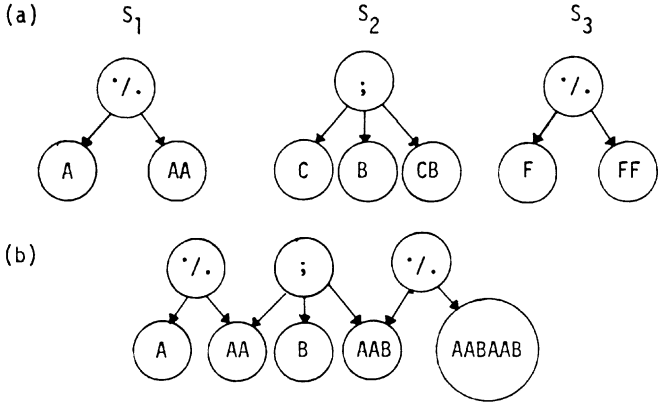



Figure 7 (b) shows the interconnection of systems $S_1$, $S_2$, and $S_3$ in (a).

Interconnection Theorem Let $\underline{\underline{O}}$ be a category and let $S_i(\underline{\underline{O}})$ denote the category of $\underline{\underline{O}}$-systems and strict system morphisms. If T is a diagram in $S_i(\underline{\underline{O}})$, then colimit T exists in $S_i(\underline{\underline{O}})$. That is, every strict interconnection of $\underline{\underline{O}}$-systems has a result as an $\underline{\underline{O}}$-system.

Our second theorem gives the method for computing the behavior of the result of an interconnection of systems from the behaviors of the component systems. This method suggests the "top-down" design characteristic of structured programming; i.e., the program design begins with an interconnection of modules, each left (temporarily) unspecified except for its behavior, and then continues by repeatedly specifying each module as an interconnection of submodules (these again left unspecified except for their behavior) until the final (lowest) level of the actual code is reached. At each successive stage of design, correctness is verified by checking that the assigned behavior of a module agrees with the behavior obtained by computing the interconnection of the behaviors assigned to the submodules. The Behavior Theorem certifies this method and, moreover, suggests that such structured programming techniques may be applied to the design of more general systems; for example, environmental or social systems.

Informally stated, the Behavior Theorem says that if T specifies an interconnection of $\underline{\underline{O}}$-systems, each of whose behaviors (limits) exist in $\underline{\underline{O}}$, then the behavior (limit) of the result (colimit T) of the interconnection specified by T exists and is the behavior (limit) of the system of behaviors of the component $\underline{\underline{O}}$-systems.

Behavior Theorem Let T be a diagram in $S_i(\underline{\underline{O}})$ such that limit $T_x$ exists in $\underline{\underline{O}}$ for each vertex x in the underlying graph of T. Then limit (colimit T) exists and equals $\text{limit}_x$ (limit $T_x$), i.e., limit (colimit T) equals the limit of the diagram in $\underline{\underline{O}}$ with the same underlying graph as T except with each $T_x$ replaced by its limit in $\underline{\underline{O}}$.

REFERENCES

1. J. A. Goguen, "Complexity of hierarchically organized systems and the structure of musical experiences," UCLA Computer Science Department Quarterly, 3, No. 4, Oct. 1975.
2. J. A. Goguen, "Objects," Int. J. General Systems, 1, pp. 237-243, 1975.
3. J. A. Goguen, "Mathematical representation of hierarchically organized systems," In: Global Systems Dynamics, edited by E. Attinger, S. Karger, Basel, Switzerland, pp. 112-129, 1970.
4. J. A. Goguen, "Systems and minimal realization," Proc. 1971 IEEE Conf. on Decision and Control, Miami Beach, Fla., pp. 42-46, 1972.

5. L. A. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," IEEE Trans. on Systems, Man, and Cybernetics, 1, pp. 28-44, 1973.
6. S. Mac Lane, Categories for the Working Mathematician, Springer Verlag, New York, 1971.
7. J. A. Goguen, J. Thatcher, E. Wagner, and J. Wright (ADJ), "A junction between computer science and category theory." IBM Research Report RC 5369, 1976.