On the Definition of Self-Organizing Systems*

George G. Lendaris
Defense Research Laboratories
General Motors Corporation, Santa Barbara, CA

(*Received October 21, 1963)

In recent years, much effort has been directed toward developing what are called “adaptive systems,” “self-organizing systems,” “learning machines,” etc. In a recent communication, Zadeh commented that it is difficult to find a precise definition of adaptation in the literature; hence, he formulated a short, mathematically precise (and quite general) definition of adaptation. Fairly precise definitions of self-organizing systems can be found in the literature, however, but they are embedded in detailed philosophical discussion and/or definitions.

In addition to defining a self-organizing system, it is desirable to know what the basic elements of such a system are. The purpose of this communication is 1) to present a set of concise definitions leading to a definition of self-organizing systems, and 2) to extract from these definitions knowledge of the structure of a self-organizing system, to be given as a theorem.

To define a self-organizing system, the term, system, must first be defined. Abstractly, for an entity to merit the name system (as generally applied) it must in some way process information. That is, starting with some kind of “inputs,” it performs some operations of these inputs, and yields the consequences of these operation (called “outputs”)--this is all that is necessary to a system. This applies, for example, to a simple R-C network, to a radio containing the network as a component, or to an entire broadcasting system containing both network and radio as components. In each case, the system (or respective component-system) has what may be called parameters which affect the way in which the system operates (for example, in the R-C network, the values of R and C affect its operating characteristics; and in the radio, the volume and tuning controls affect the way the radio system operates--each affecting the operation of the system in which it is contained.)

With this discussion as a motivation, the following sequence of definitions and axioms is presented.

**Definition 1:** A system $S$ is a triplet: $S=\{I, O, R\}$, where $I=I_r \times I_p$ (Cartesian product of $I_r$ and $I_p$). $I_r = \{U_s(t)\}$, the set of all $U_s(t)$--running over some index set; $U_s(t)$ is a vector time function defined on, say, $t \geq 0$; and the components of $U_s(t)$ are the *Regular Inputs* of $S$. $I_p = \{V_s(t)\}$, the set of all $V_s(t)$--running over some index set; $V_s(t)$ is a vector time function defined on, say, $t \geq 0$;
and the components of $V_s(t)$ are the *Parameter Inputs* of $S$. $O = \{C_s(t)\}$, the set of all possible $C_s(t)$-s running over some index set; $C_s(t)$ is a vector time function defined on, say, $t \geq 0$; and the components of $C_s(t)$ are the *Outputs* of $S$. $R$ is a relation between $I_r$ and $O$, i.e., a subset of the Cartesian product $I_r \times O$, and $R$ is dependent upon the parameter-inputs. If $R$ is fixed, the system is said to be fixed. If $R$ is variable, the system is said to be variable. (A variation in $R$ is effected by a variation in a parameter-input.\(^4\)) That is, let $B = \{R|(R \subset I_r \times O)\}$, the set of all possible relations between $I_r$ and $O$, and let $F$ be a function from $I_p$ to $B$; we have $R = F(V_s)$. Now, partition the set of parameter-inputs into two subsets: let $\Gamma$ be the set of those parameter-inputs which will be called “intentional,” and $\Lambda$ the set of all remaining parameter-inputs. Let $\gamma$ be a subset of $\Gamma$ and $\lambda$ be a subset of $\Lambda$.

**Definition 2:** A system $A$ is said to *dominate* a system $B$ when the output of system $A$ is a parameter-input of system $B$, and no output of $B$ is a parameter-input of $A$.

**Definition 3:** A system self-organizing with respect to a relation $D$, call it $S_D$, is a 4-tuple: $S_D=[I, O, R, D]$, where $R$ goes to $D$ ($R \rightarrow D$) with time, for at least two distinct $D$s, and there are no $\Gamma$-parameter-inputs to $S_D$. $I$, $O$, $R$ are as in Definition 1. $D$ is a specified relation between $I_r$ and $O$.

As a first step, the necessary axiomatic assumptions regarding the existence of $S$ and $S_D$ are made. Assuming that an $S_D$ exits, then by Definition 3, $R \rightarrow D$ with time; hence it follows that $R$ is variable (except for the trivial case where $R=D$).

Since $R \rightarrow D$ with time, it is reasonable to assume that $R$ varies purposefully (in some sense) at least part of the time, because it does not seem probable that $R \rightarrow D$ solely by random changes in the parameters. With this argument as motivation, the following axiom is stated.

**Axiom 1:**

$(R \rightarrow D \text{ with time }) \implies (R = R(\gamma, \lambda), \text{ where } \gamma \neq \phi \text{ and } R \text{ is not a constant function of } \gamma)$

Axiom 1 requires the existence of some parameter-inputs which are changed intentionally (i.e., $\gamma \neq \phi$). Definition 3 asserts that there are no such parameter-inputs to $S_D$; therefore, there must be a source of the $\gamma$ within $S_D$, that is, a component-system (call it $W$) within $S_D$ with outputs $\gamma$ (viz., $W = [I_w, \gamma, R_w]$). Since these parameters serve to change $R$, component-system $W$ must dominate some other component-system (call it $V$) within $S_D$--one that is characterized by this $R$ (viz.
$V = [I_v, O, R(\gamma, \lambda)]$, where $R$ is the same relation as that for $S_D$; therefore, the regular inputs to $V$ are the same as those to $S_D$; i.e., $I_v = I_r \times I_{p_v}$. Furthermore, the system $W = [I_w, \gamma, R_r]$ must be such that $R(\gamma, \lambda) \rightarrow D$.

These considerations lead to the conclusion that $S_D$ must have a component-system of the form $V = [I_v, O, R(\gamma, \lambda)], \gamma \neq \phi$, and a component system of the form $W = [I_w, \gamma, R_r]$ where $W$ dominates $V$ in such a way that $R(\gamma, \lambda) \rightarrow D$. (By this last requirement, $W$ in essence represents $D$.) The requirement that $W$ causes $R \rightarrow D$ implies that $W$ receives some information regarding $R$; since $R$ is a relation between $I_r$ and $O$, this means that $W$ receives as its regular-inputs at least some of the regular-inputs of $V$ and at least some of the outputs of $V$.

These deduced necessary requirements are represented in Figure 1 via standard block diagram notation. This block diagram represents a minimum necessary interconnection pattern for a system to be self-organizing with respect to a relation $D$ as defined in Definition 3. The word subset is used in the diagram to indicate that at least some, or possibly all, of the respective information is included in the labeled path. No attempt was made to include the unintentional-parameter-inputs, since they can come from anywhere.

In practice, determination of which inputs to the system are to be called regular-inputs and which ones are to be called parameter-inputs is determined by $D$. This is because $D$ expresses a relation between certain inputs and the outputs; thus in specifying $D$, a set of inputs are specified. These latter inputs are the regular-inputs and all others are the parameter-inputs.

For the converse of the above arguments, if a system $S = [I, O, R]$ has a component-system of the form $V = [I_v, O, R(\gamma, \lambda)]$ and a component-system of the form $W = [I_w, \gamma, R_r]$, where $W$ dominates $V$ in such a way that $R(\gamma, \lambda) \rightarrow D$, where $D$ is a specified relation between $I_r$ and $O$, and this is true for at least two distinct $D$’s, then $S$ is self-organizing with respect to $D$. This follows from Definition 3. Thus, an alternate definition of a self-organizing system is given by the following theorem.

**Theorem:** A system $S = [I, O, R]$ is self-organizing with respect to a specified relation $D$ between $I_r$ and $O$ if, and only if, $S$ has a component-system of the form $V = [I_v, O, R(\gamma, \lambda)]$, $I_v = I_r \times I_{p_v}$, and a component-system of the form $W = [I_w, \gamma, R_r]$, where $W$ dominates $V$ in such a way that $R(\gamma, \lambda) \rightarrow D$ with time, for at least two distinct $D$’s.
The forte of this latter definition, over Definition 3, is its display of some (minimum necessary) properties of the structure of a self-organizing system.

Acknowledgment

The author wishes to acknowledge the contributions made by his colleague G.L. Stanley and by Dr. A. Nordsieck during the course of many discussions they had regarding this material.

Figure 1. A minimum necessary interconnection pattern for a system to be self-organizing with respect to a relation $D$, as defined in Definition 3.

---

4. A system $S$ can be composed of other systems, say $S_1$ and $S_2$, each defined by its respective triplet, $S_1 = [I_1, O_1, R_1]$ and $S_2 = [I_2, O_2, R_2]$. $S_1$ and $S_2$ will be called component-systems.
5. $[I, O, R, D] \rightarrow [I, O, D, D] \equiv [I, O, D]$