File with: Unpublished papers by JNW
WRITTEN IN 1958, at Purdue University

Version 3

ON THE DESIGN OF A CREATIVE MACHINE

John N. Warfield

"In general we mean by any concept nothing more than a set of operations; the concept is synonomous with the corresponding set of operations..."

--Percy Bridgman

"The advantage of confining attention to a definite group of abstractions is that you confine your thoughts to clear-cut definite things, with clear-cut definite relations...If the abstractions are well founded...the scientific thought which confines itself to these abstractions will arrive at a variety of important truths relating to our experience of nature...

--A. N. Whitehead

"Perhaps the most glaring shortcoming of today's art is our lack of quantitative understanding of human beings as part of a complete system that includes both the machine and the human being."

-- S. Ramo

1. The Creative Machine--Obstacles to its Realization. One might suppose that there are many obstacles to the design of a creative machine. However one need not be unduly inhibited by obstacles. The modern digital computer did not exist until fairly recently. Several obstacles had to be overcome before the digital computer could be designed. When all is said and done the primary obstacles which existed were mainly the following:

a) absence of a two-state storage element

b) absence of devices for performing logic operations

When it was recognized in the 1940's that relays could perform both these functions, the design of such machines was carried out. Shortly thereafter when it was recognized that electron tubes could also perform both these functions, an electronic machine was designed.

Perhaps the most notable aspect of these obstacles is their simplicity. Typically, once a major "breakthrough" is made in science or in design, it is found that the obstacles which prevented this breakthrough are basically simple and readily understood. It is before the fact when the problem appears obscure, and the simple relationships that are involved tend to be heavily masked and withdrawn from ready observation. Could this be true of the "creative machine" as well? Are the obstacles basically simple, just awaiting our careful appraisal, or are they shrouded in mystery and susceptible to all manner of speculation?

Is it not probable that one of the major obstacles, and perhaps the only obstacle to the design of a creative machine, is the lack of a clear-cut operational concept of what constitutes creativity?

2. Only Living Things are Creative. It is clear that the discussion of the design of a creative machine cannot begin with the assumption that there is agreement on what is meant by creativity, and indeed the whole question may hinge on such an agreement. Let us, therefore, begin where agreement seems likely with a postulate

Where there is no life, there is no creativity (P-1)

3. Operational Concepts are Necessary. Let us make a further postulate which has to do with what seems to be an essential prerequisite to the design of anything, let alone a creative machine

That which cannot be defined in terms of a set of physical and logical operations cannot be designed or constructed (P-2)

4. What is Life? In the absence of life, there is no creativity. Let us examine closely what is meant by life, for in attempting to clarify this concept we strike toward the heart of the issue of creativity. We make a third postulate:

Life and death are complementary events forming a total and complete set; any object is either alive or dead (P-3)

If postulate P-3 is accepted, it is then possible essentially to define life by defining its complement death. In defining death, we make use of a concept which involves perhaps as much obscurity as creativity--namely thought. However we propose to define thought later.

Death is the absence of thought or action.

Let us now resort to a form of Boolean algebra commonly used in the design of digital machines. This measure is taken as a step in the direction of satisfying the requirement imposed by postulate P-2.

In terms of Boolean algebra, postulate P-3 can be written as

$$L + D = 1 \tag{1}$$

where L represents life and D represents death. Alternatively we can write

$$L = \overline{D} \tag{2}$$

The definition of death can be written as

$$D = \overline{(T + A)} = \overline{T}\,\overline{A} \tag{3}$$

where T represents thought and A represents action.

Combining equations (2) and (3) yields

$$L = T + A \tag{4}$$

which states that life is present if thought is present or if action is present. Strictly speaking, both T and A are undefined concepts at this point. We hope to remedy this deficiency shortly.

It is interesting to note that the life equation, Eq. (4) would suggest that many devices may be declared to be alive; for example

an exploding atomic bomb

an electrical motore

a storage battery

an electric light

a hydroelectric installation

a moving gear

The reader who objects to such a classification might wonder at the colloquial expressions a "live" bomb, a "live" wire, a "dead" battery, "kill" the light. Also while a "living" rose is called a "plant", so is a hydroelectric installation. Perhaps our intuition knows more than it is telling about what constitutes "life".

5. What is Action? It is not our purpose here to redefine all of science. Rather we seek to convert concepts which are vague into concepts which are sharper. As used in the life equation, action means energy conversion, thus as a logic variable A = 1 for any object or system for which not all the rates of energy vanish; that is, power exists in some form.

6. What is Thought? Supposedly one could give many definitions of thought. It is intended here to give a comprehensive and broad definition of thought by expressing it as the union of a number of constituents all of which are required before thought is identified to be present. We define thought as the logic intersection of seven components:

$$T = T_1 T_2 T_3 T_4 T_5 T_6 T_7 \qquad (5)$$

where

$T_1 = 1$ implies that a question is formed with a criterion or criteria for solution

$T_2 = 1$ implies that the logic variables required in considering the question are enumerated

$T_3 = 1$ implies that the logic tree which forms the set of possible answers is established

$T_4 = 1$ implies that the possible answers are assigned weights based on the solution criteria

$T_5 = 1$ implies that a decision or choice is made to accept one of the set of possible answers

$T_6 = 1$ implies that the action to be taken as a result of the answer is defined

$T_7 = 1$ implies that the information concerning the action is put into a form which can cause the action to be accomplished

Again we note some common colloquialisms which suggest that the above ingredients of thought agree with intuition;

T_1: "a thought-provoking question"

T_2: "he didn't take everything into account"

T_3: "he shinnied up the wrong branch"

T_4: "a weighty problem"

T_5: "he had no choice"

T_6: "he lays it on the line"

T_7: "he can't translate his thoughts into action"

With the concept of thought expressed in Eq. (5), it is possible to write the life equation in the form

$$L = T_1 T_2 T_3 T_4 T_5 T_6 T_7 + P \qquad (6)$$

where we have replaced A with P to correspond to the more scientific term power.

It will be appreciated that although Eq. (6) appears as a combinational logic equation, it is not implied that all elements of (6) will be present simultaneously, but rather that over a suitable observation time all can be found. This observation time would vary with circumstances.

7. Some Implications of T_1 Thought has its operational origin in the posing of a question. However a question cannot be asked without:

a) A language from which the words can be extracted to form the question

b) A depository or memory to hold the language

The first implication of T_1 is that access to a memory is essential at the outset of the thought process, and this memory cannot be empty.

The existence of a question strongly implies that uncertainty is present, though this is not a necessity. Any reduction of uncertainty which may result from answering the question is equivalent in information theory to the generation of information. What is the site of the uncertainty? When one wishes to compute the information generated by the resolution of a question, it is necessary to postulate an observer to whom the uncertainty is attached, and to whom information is presented as a result of the thought process. Thus while one may not argue conclusively that an observer is essential to determine whether thought is taking place, it is easy to argue conclusively that an observer is essential to the measurement of the results of thought. As far as the design of creative machines is concerned,

if thought is involved, and if it is desired to evaluate the information generated, the observer is essential and plays a role similar to that played by a wattmeter in measuring electrical energy. The second implication of T_1 is that the results of thought are measurable only relative to an observer. There is an obvious analogy here to relativity theory, where the results of motion depend upon the observer, insofar as the expression of any measurement is concerned. Here, though, we deal with information rather than motion.

8. What is Creativity? We have argued that life is essential for creativity, that life can be expressed in terms of thought and action, and that the information output of the thought process could be determined only through reference to an observer. Now we argue that

Thought, Action, (hence life) are necessary but not sufficient for creativity

and that the missing ingredient in these statements is the process of generation of information. This missing ingredient can be entered in the life equation as follows:

$$L = T(H + \overline{H})(P + \overline{P}) + P \qquad (7)$$

where the equation has been formed by multiplying thought by unity twice. Here H = 1 if the thought process results in a non-zero information rate as measured by the observer. We may rewrite Eq. (7) as

$$L = THP + P + T \qquad (8)$$

Note that the first term is completely redundant in Eq. (8), yet the equation is correct with this term present. We identify this redundant term as the "crdativity term in the life equation", that is C = THP. We believe that the redundant aspect of the creativity term is a principal reason for the difficulty in gaining a good understanding of it.

9. Classes of Creative Machines. We have defined creativity as the intersection of thought with power resulting in the generation of information to the observer. Since the observer (information meter) is an essential aspect of creativity, and would be a part of any creative machine, it seems essential to give careful consideration tothe peculiar role of the observer. This leads us to discuss classes of creative machines.

A machine is relatively creative is there exists at least one observer for which the creativity equation is satisfied; that is for which C = THP = 1. As an example, consider the parent-child relationship. The child asks a question, the parent relieves the uncertainty of the child, the child takes some action based upon the answer. The two collectively constitute a creative machine, with the child serving as the observer. The machine is relatively creative as farsas the child is concerned, for all the elements of creativity are present, though to another observer, the machine might not be creative.

A machine is absolutely creative if the following conditions are satisfied:

a) There is a class of observers who collectively share the vocabulary of the machine

b) This total class acting together make up a joint observer

c) For this class of observers the creativityequation THP = 1 is satisfied.

A machine is a nonsense machine to an observer who does not share the language of the machine. A machine which is absolutely creative is readily transformed into a nonsense machine by the simple expedient of replacing the original observer class with a new class not acquainted with the language of the machine.

It is, therefore, to be expected that a machine would be classed as absolutely creative by one observer and as a nonsense machine by another. This natural consequence should put to rest any thought that universal agreement would be obtained upon the classification of any machine.

10. The Absolutely Creative Machine. The absolutely creative machine has a memory, and has a vocabulary which is collectively possessed by its observer class. Furthermore it satisfies the equation THP = 1 for its observer class, when this class is viewed as a joint observer.

An interesting further property of such a machine follows. If such a machine is realizable, it must be designed at least in part by its own observer class. This follows from the need to introduce the common vocabulary into the (assumed finite) memory. It is made even more evident by the nature of the thought components T_2--T_7.

11. A Small-Scale First-Stage Design.

The title of this paper is not "The Design of a Creative Machine"--it is "On The Design...", and it is certainly not intended here to try to present any design of such a machine. However it is intended to discuss the design in terms of the first stage of the thought process--namely the formation of a question. In order to illustrate this, a number of words were selected as follows:

1) The first noun appearing on pages 10, 20, 30, ..., in the MIT Applied Electronics text.

2) The first verb appearing on pages 800, 810, 820, ... in the same work.

Random formation of questions from this partial vocabulary according to a pre-determined sentence structure led to the following questions being generated, as can readily be done for example by a digital computing machine:

How can force be moved?

How can synchrotrons be simplified?

How can electrons be rotated?

How can electrons be formed?

How can paths be derived?

How can paths be computed?

Another similar selection from two other books led to:

Can time be recorded?

How can receivers influence discussion?

How can approximation influence reason?

How can criticism influence method?

How can thoughts influence philosophers?

Many of these questions appear to be quite reasonable ones to ask, and are indeed rather suggestive of potentially creative acts.

We might also mention the fact that machines are now performing a task essentially equivalent to answering the question: "What are the essential words in this technical paper?"

It can be visualized, therefore, that the first stage of a creative machine can consist of a memory with a suitable vocabulary, and provision for generating questions. The consideration of succeeding stages poses difficulties which will not be treated in this paper.

12. Conclusion. A clarification of what constitutes a creative machine has been given. Consideration of the design aspects of such a machine, though quite limited, have served to motivate the structuring of a machine into a set of operational concepts. These concepts collectively furnish a kind of model of a creative machine.

It is tempting to break into a rash of speculation at this point involving the ideas set forth. However this would defeat another purpose of the paper, which has been to furnish others with a model suitable for their own speculation, or at least suitable for criticism in evaluating the concepts presented.

Therefore such questions as: "Does the model correctly represent man as a creative machine?" or "Is the combination of man and a digital computer likely to provide an absolutely creative machine?" will not be investigated here. A central conclusion of this paper, though, is that the model of a creative machine conceptually presented furnishes a basis for rational discussion of such questions as these, and that it also may suggest ways in which other stages of such machines could be designed, in addition to the first stage discussed in Sec. 11.

Finally, though this paper doubtless leaves many questions unanswered about creative machines, and the concepts used to describe such machines, it is believed that the methods demonstrated herein may fruitfully be applied to modify these preliminary ideas, perhaps in a way to lead to a sharper and more precise idea of what constitutes a creative machine.