*Astounding*
# SCIENCE FICTION
Reg. U. S. Pat. Off.

## CONTENTS

AUGUST, 1950 VOL. XLV, NO. 6



*Editor*
JOHN W. CAMPBELL, JR.
*Assistant Editor*
C. TARRANT
*Adv. Mgr.*
WALTER J. MCBRIDE

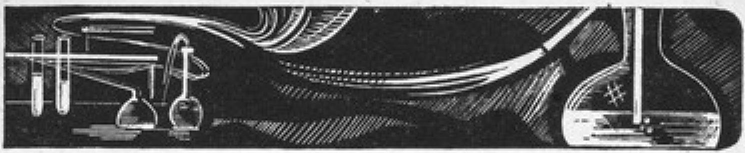
COVER: BY MILLER

Illustrations by Brush, Cartier, Miller and Noyga



Monthly publication issued by Street & Smith Publications, Incorporated at 775 Lidgerwood Avenue, Elizabeth, New Jersey. Reentered as Second Class matter October 19, 1948 at the Post Office at Elizabeth, New Jersey, under Act of Congress of March 3, 1879. Copyright, 1950, in U. S. A. and Great Britain by Street & Smith Publications, Inc. Gerald H. Smith, President; Henry W. Ralston, Vice-President; Ralph R. Whittaker, Jr., Vice-President; Arthur P. Lawler, Vice-President and Secretary; Thomas H. Kaiser, Treasurer. Subscriptions to Countries in Pan American Union, $2.75 per year; $3.00 per year in Canada, elsewhere, $3.25 per year. All subscriptions should be addressed to P.O. Box 494, Elizabeth, N. J. We cannot accept responsibility for unsolicited manuscripts or artwork. Any material submitted must include return postage.
General and Executive offices at 122 East 42nd Street, New York 17, New York.
$2.50 per Year in U. S. A. Printed in 16 the U. S. A. 25c per Copy

NEXT ISSUE ON SALE AUGUST 16, 1950 AST—1C



# PERFECT THINKING MECHANISM

In this issue is J. J. Coupling's article on building—rather, designing—a perfect thinking mechanism. Coupling wrote this before Hubbard's article on dianetics appeared, of course; Coupling's article is strictly from the cybernetics side. But I feel that Coupling, in this article, has made a major, basic contribution to the whole philosophy of thought and the mechanism of mind.

Hubbard's approach to the mind, as he specifically stated, was purely a matter of finding what worked in trying to straighten out minds—not an effort to find out the structure of the mind. Coupling's current article is, in essence, an effort to get some inkling of the sort of mechanism required to think—a totally different approach to the general field of mind and thought.

But some of the corollary of Coupling's highly important three basic rules for a perfect thinking mechanism are most suggestive. The perfect thinking mechanism, as Coupling shows, must *not* learn the first time, and *must* be able to ignore or forget a previously learned pattern of response.

Now while such a mind-mechanism leads to perfect thinking, it would, in an organism, lead very quickly to death. An animal equipped only with such a mind would have to be clawed by a lion five times or so before learning that lions are dangerous! The baby would have to burn itself five times before learning to dread fire.

Every school of psychology has maintained that there are at least two levels of mind in man, and that these levels of mind are in conflict—but none has given a really satisfactory, logical reason *why* there were two or more levels. Why a mechanism that led to conflict and inefficient operation had not been eliminated by the megayears of evolution. Coup-

*(Continued on Page 161)*

(*Continued from Page 4*)
ling's article suggests a very direct reason; the *two types of mind are essential to survival.* There *has* to be the two sharply different types of mind in an organism that is to survive.

An animal that has to be injured many times before learning danger won't survive; but slow-learning is essential in the development of thinking. The only solution to such a problem would be the installation of two different types of control mechanism. One type will have to be an instant-learner, a mechanism that learns the first time, and learns infallibly, while the other will be a thinking mechanism that learns slowly. The simplest possible type of setup imaginable would be something like this:

When confronted with serious danger, the thinking mechanism works out, hurriedly, a possible solution, and throws it into action. Simultaneously, a recording mechanism—picture it mechanistically as a punched-tape, device that makes an automatic sequence control tape recording – permanently records the entire sequence of events. If the solution hastily worked out by the thinking mechanism is successful, the organism – and its punched-tape recording – survives. Evidently, the punched-tape mechanism will, then, have only successful survival solutions on record. They may not be good solutions, but they worked at least once. Perhaps they aren't the optimum—but it's something, anyhow.

The thinking mechanism will not learn that solution the first time; confronted by the same danger again, the thinking mechanism would have to resolve the problem of survival. Consider a monkey who has been attacked and clawed by a lion, but escaped by climbing a tree. The first time, he had to think of an escape. The next time the sight-sound-smell stimulus of lion shows up, the sensory perceptions will automatically restimulate the sight-sound-smell index punchings on that punched-tape of "How To Survive When Attacked By Lions"; the high-speed, quick-action punched tape goes into action right then—and the monkey's halfway up the nearest tree by the time the infinitely finer, but slower, thinking mechanism gets into gear.

Once it is realized that a perfect thinking mechanism *has* to be a slow learner, the absolute necessity for a separate, quick-learned mechanism is obvious. The quickest possible type of quick-learner would, evidently, be a recording mechanism capable of replaying survival data. But a quick-learner by its nature will be nonintelligent, and will act on the basic proposition that the recorded solution it has on hand is *the*, and *the only* solution. It, unlike a slow-learner thinking mechanism, is incapable of altering its solution; it doesn't think, it records. It, unlike the thinking mechanism, is *subject to frustration* when its recorded "only solution" cannot be put into action.

Such a setup of two mind mechanisms would fit beautifully into Hubbard's concept of the "reactive mind" and the "analytical mind." And it would explain *why* there had to be two separate minds in conflict in order for an organism to survive.

It also explains why unconsciousness and pain are the critical factors in engrams. Evidently the recorder-mechanism's solutions will have higher priority when they represent solutions to problems of maximum urgency. Presumably there's a recorded solution for the problem of how to open an egg without breaking the yolk; it won't have much priority value—won't be able to overrule the analytical mechanism. But there is also a recording, we'll say, of what to do knocked down and chewed by a savage dog at age two. This recording contains intense pain, deep unconsciousness—a close approximation of death. This data has Grade A Crash Priority—this, the recording insists, is the solution for a problem of absolute maximum urgency, a problem involving severe damage to the organism, a close approach to death—unconsciousness—in the presence of violent antagonism.

The recorded "solution" may be extremely stupid—but it carries the Grade A Crash Priority tabs that give it power to overrule and cut out the thinking mechanism at any later time when a similar situation arises.

To me, at least, it seems that J. J. Coupling's three rules for the perfect thinking mechanism lead directly, for the first time, to a clear, logical understanding of the long-observed fact that Man has two mind-levels, and that one of them is stupid, and capable of being frustrated.

THE EDITOR.

# HOW TO BUILD A THINKING MACHINE

## BY J. J. COUPLING

***This highly original article contains specifications for building a machine that thinks—solves problems by trial and learning—and does not merely do arithmetic. And it also contains some fascinating analysis of what constitutes an ideal "thinker". See the Editorial.***

There is a lot of guff in the talk of these days about "thinking machines". When I attended a conference on large-scale digital computers at Harvard, I found that people who had built operating digital computers, or had almost completed them, were very modest in their statements. Those who were just starting had larger ideas, and for the men with plans only the sky was the limit.

Electronic computer projects have come up hard against the fact that it is a man-sized job just to make a lot of vacuum tube circuits operate reliably, a job a lot bigger than most enthusiasts had realized. Many of the grander ideas and speculations of a year or so ago have been put on the shelf, and the young geniuses have buckled down to the truly stupendous problem of making exceedingly large and complicated electronic devices do even very simple-minded computing jobs.

On the fringe of the activities, a few still talk about giant electronic brains. When, however, one of these, who has written a book on the subject, asked a seasoned applied mathematician, "Dr. Blank, do you believe that electronic computers think?" he was misunderstood to say "stink." The harried mathematician replied, "Now that you put the words into my mouth, I rather think that they do." There followed a short conversation completely at cross purposes.

This doesn't mean that "thinking machines" in the form of immensely

28. Thus, the bus can go negative only if the binary number preset on switches 24 is received from the entity. When bus 20 goes negative the inverting amplifier 25 applies a positive triggering pulse to blocking oscillator 26, and thus the blocking oscillator delivers a short negative green pulse at the output 27. In connection with Figure 5, it is important to note that the ground of the circuits of Figure 4 is connected (22) to a positive terminal in Figure 5, a terminal marked +ref.

Figure 5 shows one of the n elements of the combination selector in the entity. The halves of a double triode, 30 and 31, are commonly biased a little negative by batteries 43 acting through very high resistances 42, so that after a long spell of inactivity each tube will have the same negative bias. The tubes 30 and 31 have a common cathode resistance 32 which is high enough so that if the grids are made positive with respect to ground, they will still not be positive with respect to cathode, and so that if one grid is made appreciably more positive than the other, only the tube with the more positive grid will have plate current.

The grid of tube 30 is driven by noise generator 36, and this tends to make the grid of 30 negative or positive with respect to the grid of 31 in a random manner. A red pulse applied to lead 33 is transmitted to the grids of both 30 and 31 by capacitors 34 which are of such capacitance as to offer a high impedance to signals of the frequencies supplied by noise generator 36 but a reasonably low impedance to the higher frequencies of the red pulse. If at the time the red pulse appears the grid of 30 is more positive — less negative — than that of 31, because of the signal from the noise generator 36, tube 30 will conduct and a negative pulse will be sent out on the "1" wire 38 of the output channel, and no pulse will be sent out over the "0" wire 37. If the grid of 30 is more negative than that of 31 when the red pulse arrives, a negative pulse will go out on 37 and no pulse on 38.

Now, when the pulses sent out by the entity accidentally form the combination preset in the environment, a green pulse arrives on lead 39. This is applied through capacitors 40 to crystal diodes 41 in such a direction as to tend to make the diodes conduct. The capacitors 40 are of low capacitance, so that they transmit the short green pulses but are essentially open circuits to the long red pulses. The diodes 41 are also biased by voltages derived from the plates of tubes 30 and 31, so that only the diode connected to the plate of the tube with plate current—the lower-potential plate—will be made conducting by a green pulse. Thus, when a green pulse comes along, one of the diodes is made conducting and this tends to put a negative charge on one of the capacitors 34. The capacitor which is charged negatively is that associated with the nonconducting tube. Hence, a green pulse changes the relative biases of the tubes 30 and 31 in such a way as to bias the grid of the nonconducting tube more negative. Thus, for instance, if a pulse goes out over the "0" wire and if a green pulse is elicited from the environment, the tubes are biased so as to favor production of a pulse on the "0" wire the next time a red pulse arrives. If, however, the entity is left inactive or without green pulses for a time—if the entity is turned off or if the present combination in the environment is changed—the biases will gradually equalize because of resistors 42 and the entity will no longer be predisposed toward sending some particular pulse combination.

Returning to Figure 3, the "paralyzing" or inactivating circuit consisting of elements 15, 16 and 17 prevents the entity from elicting a rapid succession of green pulses from the environment. It also gives the entity freedom from red pulses for a time as a reward for acting properly on the environment.

THE END

---

# THE ANALYTICAL LABORATORY

For once, the article in an issue really got comments! As mentioned in Brass Tacks, some 2,000-plus letters, in fact. And I think one should properly, in consequence, put "Dianetics" in first place. Incidentally, a note to would-be advertisers: The *only* serious mention of dianetics up to the moment of this writing has been in Astounding SCIENCE FICTION. And 12,000 copies of the book were sold in the days following publication. Maybe this indicates that the production of an article everybody needs is usually met with immediate success!

In any case, the stories in the May issue rated as follows:

| Place | Story | Author | Points |
|---|---|---|---|
| 1. | The Helping Hand | Poul Anderson | 1.74 |
| 2. | The Wizard of Linn | A. E. van Vogt | 2.16 |
| 3. | The Potters of Firsk | Jack Vance | 2.58 |
| 4. | The Apprentice | Miles M. Acheson | 3.63 |

Because of the length of the Dianetics article, there were only four stories in the May issue.

THE EDITOR.

complicated electronic systems will never be built. They have just been pushed a little farther into the future. We know, for instance, that machines of present design could—if they presently existed—play chess. Perhaps playing chess isn't thinking in a true sense of the word, whatever thinking means in a "true" sense, or in any other sense, for that matter. It will do for the Sunday supplements, however. But even for this sort of thing we will have to wait until more machines are available. For anything beyond there will be a long wait indeed.

Perhaps, however, this is an unduly gloomy view. What is thinking, anyhow? Is thinking something that necessarily involves cubic yards of organized gadgets? Or, will the historian of the year 3,000 look back and say, there, in that simple device built in 194—, or in that patent of 193— the thinking robot began. Without an historian's hindsight, we cannot know, but we are free to speculate. What, then, is thinking?

I certainly wish I knew what "thinking" is. Unfortunately, I don't know, and I rather think that if I did know I wouldn't be writing this article for Astounding SCIENCE FICTION. In a rough preliminary way, however, it may be possible to assign certain general characteristics to a thinking machine, although a given machine would not necessarily have to embody all of these.

First of all, I think that a thinking machine's behavior should not be entirely predictable. The machine should have a random element, so to speak. Partly, the justification for this is that people who think are in a measure unpredictable. There is a stronger justification, however. We might illustrate this by an example.

Suppose we asked a thinking machine to make its way through a maze, and suppose we offered it a maze with two paths through; one long, involving many choices at branches; and the other shorter, involving fewer choices. Such a maze is shown in the figure. Suppose we describe a path through the maze by a sequence of letters, R and L, *right* and *left*, thus giving the order in which the machine should turn right or left in solving the maze. In the maze shown in the figure, one solution—the solid line—is, R,R. Another—the dashed line—is L, R, L, R, R, L. Certainly, we would think more highly of a machine's intelligence if it made its way through the maze in the first manner rather than the second.

If, however, the machine makes choices according to some fixed pattern in trying to solve a maze, then for some mazes it will always choose the longer way. But, if there is a random element in the making of choices, the machine will more likely find the shorter path in any maze. Further, if the machine is so constructed that it learns the way through the maze only after several successes, it will almost always learn the shorter way.

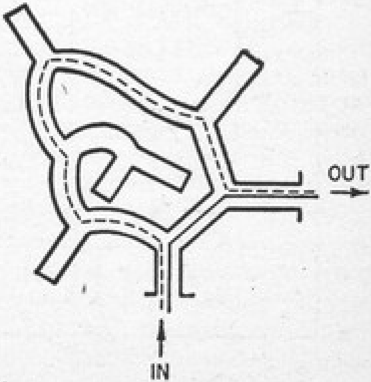
Thus, by considering the question of solving a maze, we are led to believe that (1) a thinking machine should have a random element, and (2) it should learn only after succeeding several times. This is appealing in a human sense, too, for certainly we fumble around a lot in solving our problems, and by and large we don't learn by succeeding just once. I think that these criteria can probably be justified for machines designed to solve any but the simplest and most routine problems.

We have talked about the machine's "learning" its way through the maze. That means that it must adjust itself to its environment—the problem it solves—so that it is not the same after having worked a problem as it was before. In other words, it must learn from experience. This implies a criterion of value. Such a criterion may be either built-in or external. We human beings have both kinds. When a baby takes its hand out of the fire, it is relying on an internal standard; it doesn't want to be burned. When a baby takes its hand out of the cookie jar, it is using an external criterion of value belonging to its mother, a criterion which may make itself felt through a sharp tone of voice, or otherwise.

Conceivably, we could reward our thinking machine in some way when it pleased us, perhaps by pressing a button on it labeled, maybe, "pat on the back." In a simple machine we would probably merely make it tend to remember a course of action which achieved the solution of some problem. The machine's pleasure could, if we wished, be made manifest by the flashing of a green light or by the wagging of a mechanical tail.

Does this sound interesting? Would an unpredictable machine which solved and remembered the solution to a maze be of interest to the average reader of science fiction? Suppose that, in addition, we added the feature that if the maze were changed, the machine would gradually forget the older, now invalid, course of action and learn how to solve the new maze. Would that help to make the machine interesting? I don't know. Offhand, the answer might seem to be, yes. However, I would advise the reader to suspend judgment.

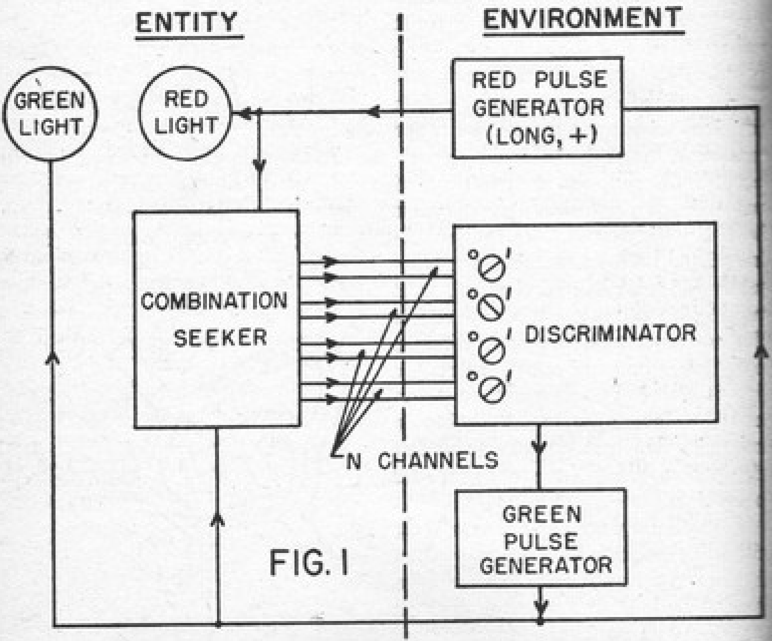
In the meantime, here is a description of a simple device which I thought up while considering maze-solving machines. While it hasn't



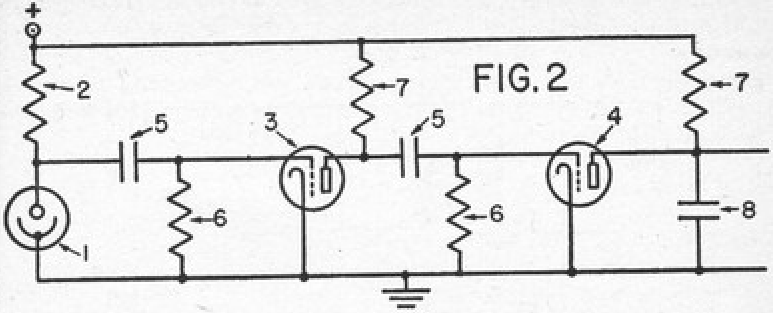
been built and debugged, I'm quite sure that it is operable and could be made by following the description and circuits given in the Appendix.

The device consists of two boxes, one called the "entity" and the other the "environment." The environment spontaneously sends signals called "pain signals" to the entity. These pain signals light a red lamp on top of the entity and also cause the entity to send out a code signal of "As" and "Bs" along the "A" or "B" wires of N pairs of wires leading from the entity to the environment. Instead of As and Bs we may say instead, if we wish, *ones* and *zeros*. Such a group of ones and zeros can be thought of as forming a binary number, as, the number 1001, for instance.

At first, the binary numbers sent from the entity to the environment in response to pain signals from the environment are chosen entirely at random. However, the environment is so constructed that on receipt of one particular binary number it will send a "pleasure" signal to the entity. This pleasure signal lights a




*Block diagram of machine which can learn by trial-and-error, and remember.*




*"Noise source" generator circuit. N-plus-1 needed.*

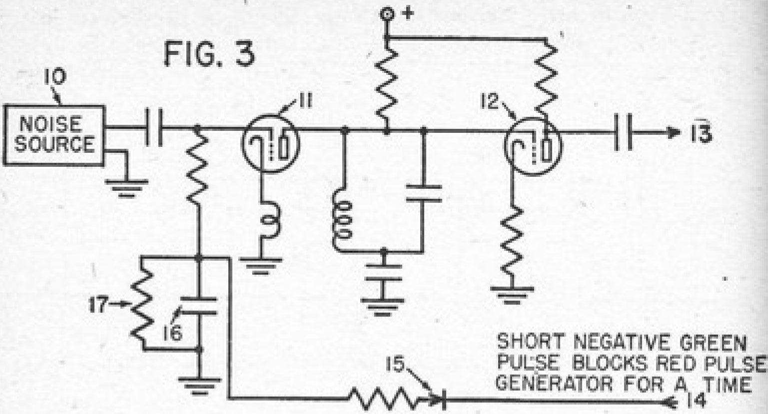
green lamp on top of the entity and also biases the tubes in the entity which send the binary numbers to the environment in such a way as to favor the last signal sent over all other combinations. After several successes, the biases will become such as to predispose the entity overwhelmingly toward sending the combination of pulses which elicits a pleasure signal. If the environment is changed so as to require a different code from the entity in order to obtain a pleasure signal, the biases will gradually leak off; the entity will gradually forget the old pattern, and then it will learn a new one.

We could make such a device learn quickly and forget quickly, or learn only after many successes and be very stubborn about forgetting, or any other combination of stubbornness or adaptability.

What about maze-solving, though? Well, we saw earlier that two solutions of the maze shown in the figure can be written in terms of turnings to the right, R, or to the left, L, as R, R and L, R, L, R, R, L. We can just as well let 1 stand for R or right and let 0 stand for L or left. Then the two solutions can be written 11 and 010110. Suppose there were six pairs of wires between our environment—which now represents the maze—and our entity—which is now a maze-solving machine. We can set the environment to give a pleasure pulse when it receives the number 010110. That takes care of the long path solution, the dotted path of the figure. We can further add circuits to the environment, duplicating some of those already there, so that it will give a pleasure pulse for any number beginning with 11. Thus, the entity will have succeeded if it learns

110000
110001
110010
110011

or, in fact, any of the sixteen six-digit binary numbers which start




*"Red pulse" generator circuit—the equivalent of pain stimulus for the electronic organism.*

with 11. No wonder the entity will be most likely to hit on the short path. There are sixteen numbers corresponding to it, and only one corresponding to the long path!

I really think that a reader who has waded through the Appendix can't deny at this point that we have a way of making a machine which can in this sense solve a maze. Further, the machine does so while satisfying some criteria which, when they were presented earlier looked, I hope, rather interesting. The machine is unpredictable. It learns after repeated trials. It tends to find the shortest way through the maze. If the maze is changed, it forgets the old and now useless route through, and it learns the new one. It even flashes a green light when it is pleased by its success.

I also really think that a reader who has got this far is bound to be disappointed. The machine doesn't look like a man, or even like a dog. If it ran through the maze on wheels, there would be some satisfaction to it, but the machine doesn't move. It is incapable of seeing, hearing, feeling or smelling the real physical maze, and it can only solve the maze when the maze is represented by a certain setting of dull-looking switches on a box called "environment" which stands for the maze. This little gadget isn't what we bargained for at all.

There may be compensations, however. If the reader is sadder at this point, perhaps he is wiser as well. Neither I nor he knows just what we mean by "thinking." And I don't think that we are going to be taken in by any attractive definitions, only to have to admit that according to such definitions a trivial gadget must be admitted to think.

---

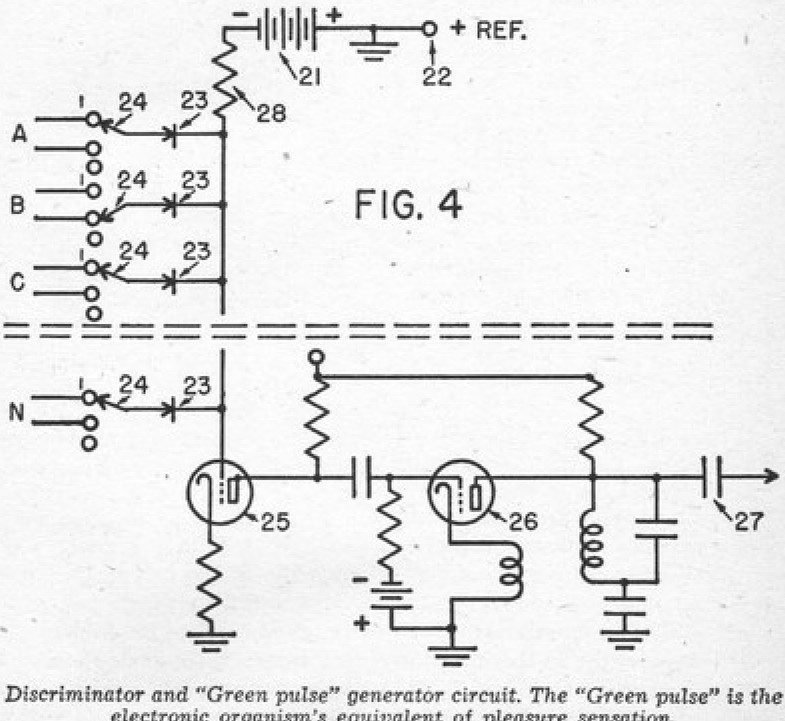
*APPENDIX*

## MACHINE WHICH ADJUSTS TO ENVIRONMENT

The purpose of this appendix is to describe in more detail a device mentioned in the text.

Figure 1 shows a block diagram of the device. It is divided into two parts shown separated by a vertical dashed line. The elements to the left constitute the *entity* and those to the right constitute the *environment*.

The environment has a *red pulse generator*, which spontaneously emits "red" pulses, which are long positive pulses. It has an input lead from a *green pulse generator*. The arrival of a "green" pulse, which is a short negative pulse, temporarily paralyzes the red pulse generator. The output of the red pulse generator goes to a




*Discriminator and "Green pulse" generator circuit. The "Green pulse" is the electronic organism's equivalent of pleasure sensation.*

*combination seeker* and to a *red light* in the entity.

The environment also has a *discriminator*. This has n input channels from the entity. The input channels consist of two wires each, the wires of each pair being labeled "0" and "1". The entity can act on the environment by sending simultaneously a negative pulse to the environment on each channel. In a given channel, the pulse can be on either the 0 or the 1 wire. Thus, the pulses on the various channels can be thought of as forming a binary number of n digits or bits. The discriminator is equipped with n single-pole-double-throw switches by means of which it can be adjusted to respond to one only of the $2^n$ possible binary numbers. When the discriminator receives the preset number from the entity, it activates the green pulse generator. This sends a short negative pulse to the red pulse generator, which temporarily paralyzes it, and also sends a short negative pulse to the combination seeker in the environment and to the *green light*.

The entity consists of the red and green lights and the combination seeker. This latter always sends a binary number to the environment when stimulated by a red pulse. Immediately after turn-on these numbers are randomly produced through the use of noise generators. However, whenever the number preset in the environment is produced by the entity, the entity receives a green pulse, and this prejudices it to a greater or lesser degree, depending on adjustment, in favor of producing the preset number the next time it is stimulated by a red pulse.

Some of the less obvious contents of the boxes of Figure 1 will now be described.

Figure 2 shows a noise source. One of these is needed in the environment to produce red pulses, and n are needed in the entity. The source of the noise is a gas tube 1 fed through a resistor 2. The noise voltage of the gas tube goes to a two-stage amplifier consisting of tubes 3 and 4, which might be halves of a double triode. Large coupling capacitors 5 and very high grid resistors 6 are used in order to obtain good low-frequency response. If the resistors 6 are 10 megohms or so, the tubes will be self-biasing. The plate resistors 7 are also high to give high gain. The noise output appears across capacitor 8, which cuts off the high frequencies. Maybe more than two stages of amplification will be needed.

Figure 3 shows a circuit for the red pulse generator. A blocking oscillator 11 is triggered by a noise generator 10. An amplifier 12 is used to invert the long negative pulse produced so as to give a long positive pulse at the output 13. Green pulses, which are short and negative, are applied to the lead 14 and act through crystal diode 15 to put a negative charge on condenser 16. This biases the oscillator 11 so that the noise generator 10 does not trigger it until part of the charge leaks off through resistor 17.

Figure 4 shows the discriminator and the green pulse generator. The n negative pulses from the entity come on wire pairs a,b,—,n to switches 24, each of which can be set to positions 0 or 1. These switches lead to crystal diodes 23, all of which have a common terminal in bus 20. The bus is connected to a negative voltage source 21 by means of a very high resistance 28. The bus cannot go negative unless all of the diodes 23 as biased negative by pulses from the entity, for the entity impedance plus the impedance of the crytsal diodes in the conducting direction is small compared with the resistance

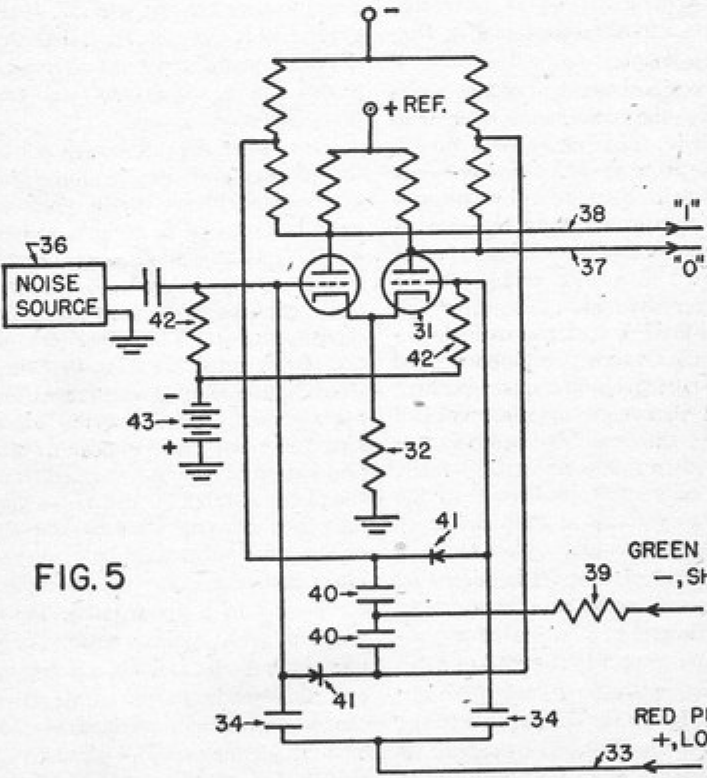



FIG. 5

*Circuit for the combination-seeker in the "entity"—N needed.*