

STANFORD RESEARCH INSTITUTE

MENLO PARK, CALIFORNIA



DRAFT PROPOSAL

Engineering Sciences Division

April 12, 1963

INVESTIGATION OF A COMMON TRAINING PROCEDURE FOR REMOTE AND RAPID DESIGN,  
REPAIR AND REDESIGN OF DEVICES

Prepared by:

Nils J. Nilsson  
Research Engineer  
Applied Physics Laboratory

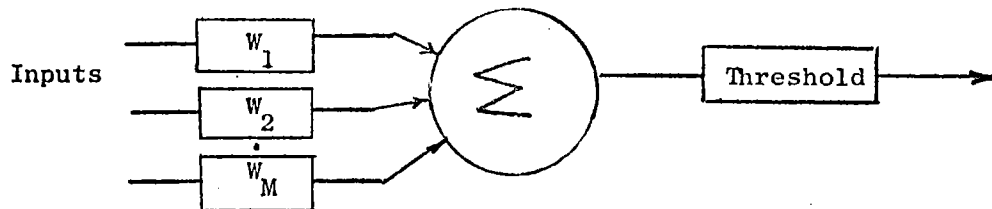
INVESTIGATION OF A COMMON TRAINING PROCEDURE FOR REMOTE AND RAPID DESIGN,  
REPAIR AND REDESIGN OF DEVICES

I OBJECT

The main object of this proposed research program is to investigate the hitherto unrealized possibility of using one training procedure as the common means to rapidly and remotely design, repair, and redesign devices. These devices would be used to perform many useful functions that are now being performed only by devices that are deterministically designed and constructed and that are repaired in a variety of ways.

II DISCUSSION

We are here concerned with "adaptive" devices which are concurrently-operating, combinational circuits built of threshold logic units (TLU's), including weights. A threshold logic unit (with say M inputs) compares a weighted sum of its inputs with a threshold. If the threshold is exceeded the output value is one; otherwise zero. A TLU is represented schematically as follows:



The weights  $w_1, \dots, w_M$  and the threshold are electronically or otherwise adjustable permitting rapid changes to be made to the input-output behavior of the unit. An adaptive device can be constructed out of a network of such elementary threshold logic units (see Appendix). Such a device will in general have M binary inputs and N binary outputs. Currently, SRI is constructing a large pattern-recognition network, MINOS II, with 6732 adjustable weights for the U. S. Army Signal Research and Development Laboratory, Contract DA 36-039 SC-78343. This network consists of essentially 66 threshold logic units whose adjustable weights are implemented by magnetic cores\*. Each step of the training procedure consists of:

- (1) Presenting the M-input signals,
- (2) Noting the actual set of N-output signals and comparing it with the desired set of N-output signals,

---

\*The development of MINOS II and its components is described in SRI Quarterly Progress Reports 1 - 11, "Graphical Data Processing Research Study and Experimental Investigation," prepared for the U. S. Army Signal Research and Development Laboratory, Ft. Monmouth, New Jersey. (Contract DA 36-039 SC-78343)

- (3) Adjusting the weights in accordance with a prescribed function of these differences.

One takes as many such steps as are required before the last presented list of the M-input signals are mapped into the desired list of the N-output signals without intervening adjustments. A device, so trained, is said to have converged to the desired mapping.

The most popular purpose to which such adaptive devices are intended to be put is to recognize not only the M-input patterns they have been shown and trained on, but also patterns they have never seen. We are not here concerned with the latter generalizing quality. We are concerned solely with training (designing in the first place, repairing, redesigning) a device to recognize (map, transform) only the M variable input pattern it was shown.

Thus, for example, to design a parallel adder of two 3-bit operands A and B to give a 6-bit sum C (the reason for 6 rather than 4 bits will be seen presently), one would train this 6 input-6 output adaptive device so that the output code C always represented the sum of A and B.

To repair this adder after it malfunctioned, one would retrain the device so that again the output code C always represented the sum of A and B.

To redesign this adder to (say) a multiplier of two 3-bit operands A and B to give a 6-bit product C (now it may be seen why the adder had a 6-bit output), one would retrain the device so that the output code C always represented the product of A and B.

In principle, then, this example serves to illustrate the possibility that training a concurrently-operating M input-N output combinational circuit, is a common means for designing a device in the first place; for repairing it; and for redesigning it to perform an entirely different function.

To cite some examples where devices with such features might be useful and economic:

- (1) Variable digital-code converters--A particular computer code converter can, by retraining, be converted to other particular code converters. For example, an IBM 7090 magnetic tape to UNIVAC 1107 magnetic tape code converter, can, by retraining, be made into a Burroughs 220 magnetic tape to a teletype paper tape code converter.
- (2) In satellites, failed equipment might be remotely and rapidly designed, repaired, and redesigned--e.g., a sequencer might have its "wired-in" program changed remotely; or it might be remotely repaired, by retraining.

### III RESEARCH PROGRAM

- (A) To demonstrate the technical feasibility of using one training procedure as the common means to rapidly and remotely design, repair, and redesign such adaptive devices, a 6 input-6 output demonstrator machine will be trained (designed in the first place) to be an adder, multiplier, code converter, and an associative memory. Each of these will be repaired by training. Finally, the demonstrator will be redesigned (retrained) to perform one or the other of these four functions. It is anticipated that MINOS II can be used for the demonstrator with the consent of the Signal Corps.
- (B) To extend our knowledge of trainable M input-N output concurrently operating, combinational circuits, theoretical studies will be focused on the following questions:

- (1) How many threshold units and weights would be required in a device; what training procedure would be used; how long would it take to train the device using this procedure; to perform any given M-N switching function with specified reliability (e.g., Mean Time Between Failures)?

Conversely, given a device and a training procedure, what switching functions could it be trained and retrained to perform? How long would the training take? What reliability would it have? Which failures could be repaired by retraining; which could not?

- (2) For certain functions such as code conversion, associative memory, addition, multiplication, and others that a sponsor might suggest, hardware cost, volume, weight, power, reliability, and repair time will be compared in conventional deterministically-designed devices and in devices studied in this program.

### IV ESTIMATED TIME AND COSTS

The estimated time required to explore this research program and report its results would be approximately one year. We could start this program within one month after a contract is made. The attached sheet details our estimated costs.

### V PERSONNEL

Biographies of key personnel follow:

Nilsson, Nils J. - Research Engineer, Applied Physics Laboratory

In August 1961 Dr. Nilsson joined the staff of Stanford Research Institute

where he is participating in the studies of pattern recognition and self-organizing machines. During the summer of 1962, he taught a graduate course in Learning Machines at Stanford University.

Dr. Nilsson received an M.S. degree in Electrical Engineering in 1956 and a Ph.D. degree in 1958, both from Stanford University. While a graduate student at Stanford he held a National Science Foundation Fellowship. His graduate field of study was the application of statistical techniques to radar and communications problems.

In July 1961 Dr. Nilsson completed a three-year term of active duty as a Lieutenant in the United States Air Force. He was stationed at the Rome Air Development Center, Griffiss Air Force Base, New York. His duties entailed research in advanced radar techniques, signal analysis, and the application of statistical techniques to radar problems. He has written several papers on various aspects of radar signal processing. While stationed at the Rome Air Development Center, Dr. Nilsson held an appointment as Lecturer in the Electrical Engineering Department of Syracuse University.

Dr. Nilsson is a member of Sigma Xi, Tau Beta Pi, and the Institute of Electrical and Electronic Engineers.

Forsen, George E. - Research Engineer, Applied Physics Laboratory

Mr. Forsen received both the S.B. and the S.M. degree in Electrical Engineering from the Massachusetts Institute of Technology in 1957, and the degree of Electrical Engineer from M.I.T. in 1959.

He was employed part time in 1954-1956 by the General Electric Company, on the Cooperative Plan with M.I.T. While with G.E. he worked on non-destructive testing methods, and measurement techniques for heat flow in power transistors.

In 1958-1959 he was a Research Assistant and staff member of the Communications Biophysics Group, Research Laboratory of Electronics at M.I.T. There he designed electronic instrumentation for the study of neuroelectric and psychophysical phenomena related to nervous systems. From 1957 to 1959 he was also employed by the Electrical Engineering Department of M.I.T. as a Teaching Assistant.

In October 1959 Mr. Forsen joined the staff of Stanford Research Institute. At the Institute he is currently engaged in the study of neuron-like devices, and adaptive, cognitive systems.

Mr. Forsen is a member of the Institute of Radio Engineers and Sigma Xi.

Fein, Louis, Consultant

Louis Fein holds a Ph.D. in physics. He has been in the computer field for 15 years, working as a computer designer, project head, company president, university lecturer, consultant, and author. He built the RAYDAC computer at Raytheon; was founder and president of Computer Control Co., Inc.; has lectured at Wayne and Stanford; and has been a consultant in reliability, design, programming, and applications of computers to almost every computer organization in the United States. He has worked extensively in the reliability of computing systems.

Rosen, Charles A. - Manager, Applied Physics Laboratory

Dr. Rosen received a B.E.E. degree from the Cooper Union Institute of Technology in 1940. He received an M.Eng. in Communications from McGill University in 1950, and a Ph.D. degree in Electrical Engineering (minor, Solid-State Physics) from Syracuse University in 1956.

During 1940-1943 he served with the British Air Commission as a Senior Examiner dealing with inspection, and technical investigations of aircraft radio systems, components, and instrumentation. From 1943-1946 he was successively in charge of the Radio Department, Spot-Weld Engineering Group, and Aircraft Electrical and Radio Design at Fairchild Aircraft, Ltd., Longueuil, Quebec, Canada. During the period 1946-1950 he was a co-partner in Electrolabs Reg'd., Montreal, in charge of development of intercommunication and electronic control systems. During this period he also acted as a self-employed consulting engineer in these fields. In 1950 he was employed at the Electronics Laboratory, General Electric Co., Syracuse, New York, where he was successively Assistant Head of the Transistor Circuit Group, Head of the Dielectric Devices Group, and Consulting Engineer, Dielectric and Magnetic Devices Subsection. In August 1957 Dr. Rosen joined the staff of Stanford Research Institute, where he has been working on applied physics projects.

His fields of specialty include learning machines, dielectric and piezoelectric devices, electro-mechanical filters, and a general acquaintance with the solid-state device field. He has contributed substantially as co-author to two books, Principles of Transistor Circuits, R. F. Shea, editor (John Wiley and Sons, Inc., 1953) and Solid State Dielectric and Magnetic Devices, H. Katz, editor (John Wiley and Sons, Inc., 1959).

Dr. Rosen is a Senior Member of the Institute of Radio Engineers, a member of the American Physical Society, American Institute of Electrical Engineers, and the Research Society of America.

VI CONTRACT FORM

It is requested that this contract be written on a cost-plus-fixed-fee basis.

VII ACCEPTANCE PERIOD

This proposal will remain in effect until June 1963. If consideration of the proposal requires a longer period the Institute will be glad to consider a request for an extension of time.

## APPENDIX

### DESCRIPTION OF DEMONSTRATOR AND TRAINING RULE

#### A. Demonstrator

The demonstrator will consist of a layered configuration of threshold logic units (TLU's). (See Fig. 1) The first layer contains 30 TLU's each of which operates on the 6 inputs through weighted connections. The 210 weights (including 30 weights which act as thresholds) connecting the inputs with the 30 TLU's are adjustable. The 30 TLU's are divided into 6 groups of 5 TLU's per group. For each group, the majority response of the TLU's in the group is then taken to be the group response. The machine then has 6 group responses which are the 6 outputs. Each group response can be determined by a second layer of 6 output TLU's (called vote-takers) each of which operates on the output of 5 first-layer TLU's.

It has been determined both theoretically and by digital computer simulation that the tasks of adding and multiplying two three-bit numbers are well within the combinational capability of the network described above. It will also be possible to train the machine to operate as an associative memory and a code-converter. The same machine could be used to perform any one of these tasks after being trained specifically on the job to be performed. This training is effected by a procedure which iteratively modifies the adjustable weights.

#### B. Training Rule

The machine will be trained in such a way that each of the six outputs is always correct for every possible input combination. The training process involves presenting to the machine, one pattern at a time, all of the possible input combinations (called patterns) together with the correct sets of responses. If all six bits of the machine response to a given pattern match the desired outputs, then no modification is made to the adjustable weights, and another pattern is presented to the machine. If one or more of the six outputs is in error, the error(s) are corrected by adjusting the weights before presenting another pattern. The act of adjusting the weights to correct an error is called an adaptation.

An error in any given output bit is corrected in the following way. (If two or more outputs are in error, they are corrected simultaneously.) A

determination is made of how many TLU's in the erroneously responding group must have their responses revised so that the majority will vote correctly. Suppose the minimum number of such reversals necessary in a group is equal to  $k$ . Of those TLU's voting incorrectly, one selects the  $k$  particular TLU's whose analog sums are closest to threshold and one prepares to reverse their responses. Suppose that (among others, perhaps) the response of the  $i$ th TLU is to be reversed. Then its weights must be adjusted. An increment is added to all the weights connecting the  $i$ th TLU with the 6 inputs. Those weights connected to plus one inputs are altered in a direction opposite to that of weights connected to minus one inputs. The size of the increments is the same for all weights, and the size and direction are determined by the total change needed in the analog sum to effect a reversal of the TLU binary output. This adjustment process is applied simultaneously to each of the  $k$  TLU's in the group whose responses are to be reversed.

After all errors in the machine output have been corrected, another pattern is presented. This process is repeated until all patterns elicit correct machine responses without the need for further weight adjustment. Each pass through the complete pattern set is called an iteration and the machine is said to have converged when no further weight adjustments are necessary. For the kinds of tasks to be performed by the demonstrator, convergence will typically require from 10 to 20 iterations.

### C. An Example

The following illustrative experiment was conducted on a computer simulated device which had 6 binary inputs and one binary output. The 6 inputs were used to represent two three-bit binary numbers,  $X$  and  $Y$ . The desired output was the 4-th--highest order--bit in the sum  $Z = X + Y$ . The device which was simulated was one of the six identical groups (the 4-th) of the proposed demonstrator and had 5 TLU's whose majority response was taken to be the machine output. Since each of the 6 outputs of the demonstrator is obtained by one of these groups, imagine that the other 5 groups are being simultaneously trained to give the other 5 bits of the sum  $Z = X + Y$ .<sup>\*</sup> The complete machine is being trained to be a binary adder which would require carry if done serially. Since the 4th bit of the sum involves the most complicated logic, the training experiment for this group will give an upper bound on the total number of adjustments and iterations needed to train the complete machine to be an adder.

---

\*The 5th and 6th bits of the output each have a desired response of zero.



In the simulation, it was found that convergence occurred after eight iterations of the 64 possible input combinations. During these eight iterations a total of 66 adaptations were made. Even after only one iteration, a test showed that out of the 64 possible input combinations only 10 were in error. Such rapid convergence indicates that the number of TLU's of the simulated machine far exceeds the minimum number of TLU's needed for this simple task-- a fact that augurs well for the ability of the intact parts of the machine to be easily retrained should parts of it become inoperative through component failure.

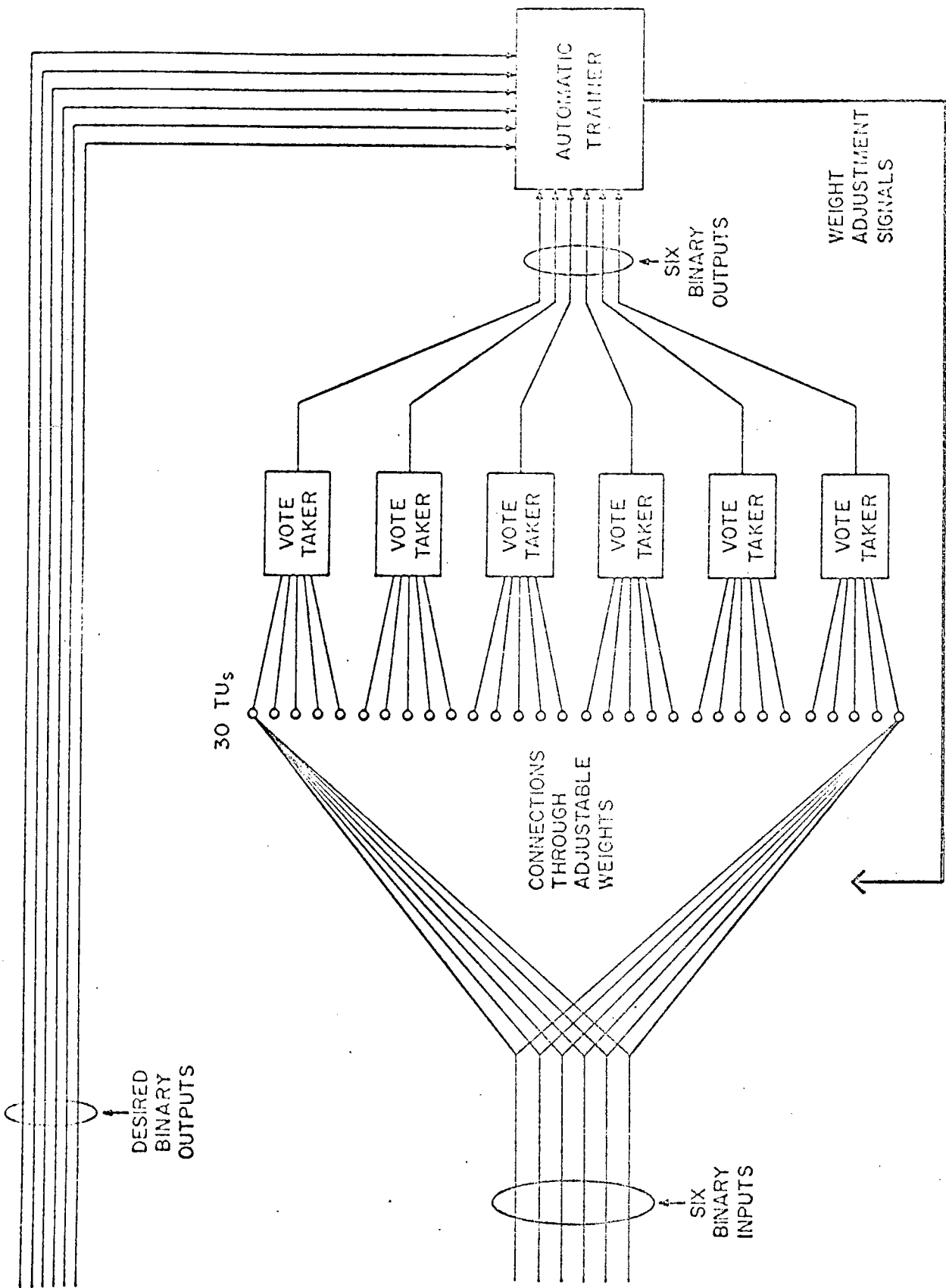


Fig. 1 Schematic Diagram of Adaptive Device