The Design and Use of Hazard-Free Switching Networks*

DAVID A. HUFFMAN

Massachusetts Institute of Technology, Cambridge, Massachusetts

I. Introduction

When combinational switching networks of relay contacts or of electronic gate elements are analyzed or synthesized, the network components are usually idealized in such a way that they may be adequately described by the Boolean algebra. It is usually postulated [1] that, at all times, all the normally open (or normally closed) contacts on a given relay open and close in synchronism with each other, and that each normally open contact is open (closed) when the normally closed contacts are closed (open), and vice versa.

Similarly in a network of electronic gate elements it is usually postulated [2] that an input variable can affect the network output with no intervening time lag. These assumptions, when put to use in synthesis procedures, lead to networks which behave correctly for steady-state situations but which may not for transient conditions (during which a network input variable is changing from one of its binary states to the other).

When combinational networks which do not behave ideally (in a sense to be considered below) during changes of input state are incorporated into larger switching networks which have sequential action (that is, which act as if they had memories) a hazard [3] exists and the sequential circuit may not operate as it was meant to by the designer. The significance of a network hazard can be substantially reduced and sometimes eliminated by "smoothing" of the network output [4]. In a relay circuit, for example, the contact networks are used to control relays which in turn contribute contacts to the various other contact networks. If the response time of a given relay is increased so as to increase its smoothing action, the effect of a hazard in its controlling network may be eliminated. But now contacts from the given relay which appear in other networks of the circuit may behave even less ideally than they did originally, thus creating new hazards in the other networks. Moreover, in both relay and electronic circuits, the stratagem of smoothing at critical points in the circuit will always increase the reaction time of the circuit. In most applications this is undesirable.

This paper suggests a method for the elimination of hazards without resort to signal smoothing. It is shown here that it is always possible to synthesize a combinational switching network which behaves ideally even though its individual components do not. The technique proposed here does not involve the use of a new switching algebra. The terminal behavior of the resulting hazard-free net-

[•] Presented at the meeting of the Association, Sept. 14–16, 1955. This work was supported in part by the Signal Corps; the Office of Scientific Research, Air Research and Development Command; and the Office of Naval Research.

works are independent of the timing of the contacts (in the case of a relay circuit) or of the delays along the signal paths (in the case of an electronic gate circuit).

II. The Basic Hazards in Contact Networks

2.1. In the treatment of contact networks in this paper, the complementary symbols 0 and 1 will stand for open circuits and closed circuits, respectively. For this *transmission* concept the operations of logical addition and logical multiplication must be associated with paralleling or cascading of contacts or networks [4]. The contacts on a given relay, R, will all have the same alphabetic designator, which will either be unprimed (r) or primed (r') depending upon whether the contact is normally-open or normally-closed, respectively. When the relay, R, is unoperated (operated) the normally-open contact, r, is open (closed), and we may write r = 0 (r = 1). Of course, when r = 0, r' = 1 and when r = 1, r' = 0.

In the transmission concept, therefore, the two networks of Fig. 1-a have transmissions of $T_1 = a'b + ac$ and $T_2 = (a + b)(a' + c)$. By means of the Boolean algebra we may prove that these two expressions are equivalent. That is, when the relays A, B, and C are each in steady-state conditions (either operated or unoperated), the two networks cannot be distinguished by measurements made at their terminals.

This equivalence can also be seen by showing, for each of the $2^3 = 8$ possible combinations of values of the three switching variables, what the network transmissions are. A convenient way of doing this is in a rectangular matrix, called a map [5]. In Fig. 1-b the upper right-hand corner of the maps, for instance, repre-



FIG. 1. Illustrating static hazards in contact networks

sents the situation for which a = 0 (a' = 1), b = 1, and c = 0. If we evaluate the transmissions of each of the two networks when the *a* contact is open, the *a'* contact closed, the *b* contact closed, and the *c* contact open, we find that terminal-to-terminal paths exist in each network. Thus the corresponding entry in each map has been made equal to 1.

2.2. In order to see indications of possible hazards in the map for a given network, it is necessary to include, in the map, representations of the cut-sets and tie-sets determined from the topological structure of the network. By a cut-set we mean here a minimum set of contacts which, when open, insure that the network transmission is zero (that the network terminals are cut apart), even if all other contacts are closed. Analogously a tie-set means a minimum set of contacts which, when closed, insure that the network transmission is unity (that the network terminals are tied together) even if all other contacts are open.

In N_1 the cut-sets are (a, b), (a', c), (b, c), and (a, a') and the tie-sets are (a', b)and (a, c). For N_2 the cut-sets are (a, b) and (a', c) and (a, c), (a', b), (b, c), and (a, a') are the tie-sets. We eliminate from consideration the cut-set (a, a') in N_1 because, when the A relay is in a stable state these two complementary contacts can never be open simultaneously. In $N_2(a, a')$ is dismissed as a tie-set since these two contacts can never be closed at the same time if A is in a stable state.

The cut- and tie-sets just found are shown in Fig. 1-c. The tie-set (a', b) will be closed if a = 0 (a' = 1) and b = 1. Therefore it is shown as a grouping of the right pair of entries in the top row of the maps. Similarly, the left pair of entries in the top row, for which a = b = 0, stands for the cut-set (a, b).

In N_1 assume that the *b* and *c* contacts are closed. As the relay *A* changes its state of operation, the transmission path through the network changes from the upper branch to the lower, or vice versa. The idealization which is usually made is that the *a* and *a'* contacts always have complementary transmissions and that, for example, just as the *a* contact is opening the *a'* contact is closing. Actually there may be a short interval of time for which both *a* and *a'* are open. In this case, N_1 will have a momentarily incorrect transmission of zero. The existence of this hazard in N_1 can also be seen indicated by the arrow in the third column of the left map in Fig. 1-c. Note that there is no one tie-set of contacts which is certain to be closed during the transition. The hazard in N_1 can be eliminated by timing of the contact operations so that *a* and *a'* are both closed momentarily when *a* is unstable.

In N_2 another kind of hazard exists when b = c = 0 and A is in an unstable state. It can be seen from the left column in the corresponding map in Fig. 1-c that the transmission should be zero both before and after the transition. But if a and a' are both momentarily closed during the transition, an incorrect unity transmission will occur. This could be eliminated by timing of the contact operations so that a and a' are simultaneously open when A is unstable.

It is possible to derive networks which are terminally equivalent to those of Fig. 1 and in which no hazards are present regardless of the relative timing of the contacts on any of the relays. Four such networks are shown in Fig. 2-a, b, c, and d. There are present in each of them the three cut-sets (a, b), (a', c), and (b, c)



(e)

FIG. 2. Four equivalent hazard-free networks and their map representation

and the three tie-sets (a', b), (a, c), and (b, c). These are plotted on the map. Note that any two entries which are either both zeros or both ones, and which differ in the value of one variable only, are included in a common cut- or tie-set, respectively. Consequently, during the corresponding instability of any one of the relays A, B, or C, there is either a cut-set with each of its variables having the value zero firmly established, or a tie-set with each of its variables having the unity value firmly established.

2.3. The hazards which were described and eliminated in the examples of the preceding section are what might be called static hazards, since during each change of a single input variable it was presumed that the network transmission was to remain static at either zero or one. Another kind of hazard can occur when the network transmission is meant to change, either from zero to one, or from one to zero, when some switching variable changes. We call this kind of hazard a dynamic hazard.

Consider the network in Fig. 3-a. (The subscripts on the three contacts on the D relay may be neglected for the present time.) Its cut- and tie-sets are shown in the map of Fig. 3-b. The two dotted arrows indicate the static hazards in the network. One of the possible dynamic hazards is shown by the heavy arrow. During the corresponding transition the only contacts which have changing transmissions are the contacts on the D relay. The a and b' contacts are closed and the c contact is open since a = 1, b = 0, and c = 0. Effectively then d_3 is in parallel with the series combination of d_2' and d_1 . (See Fig. 3-c.) Assume that when the D relay operates the three contacts change their states in the order of their sub-



FIG. 3. Illustrating a dynamic hazard

scripts. When d_1 closes the network closes; when d_2' opens the network reopens; and when d_3 closes the network closes again. This multiple change of transmission before the network assumes an ultimate transmission which is the complement of the original transmission, constitutes its dynamic hazard.

In general it can be shown that a contact network has a dynamic hazard if and only if it is possible to fix all but one of the variables (all but, say, x) at some set of values such that in the network some x and x' contacts are effective in a tie-set and some (not the same two) x and x' contacts are simultaneously effective in a cut-set. (In Fig. 3 the d_1 and d_2' contacts were the ones which were effective in the tie-set and d_3 and d_2' were effective in the cut-set.) Therefore a necessary condition for a dynamic hazard is that some relay contributes at least three contacts (not all the same type) to the network.

The elimination of static hazards does not necessarily guarantee that dynamic hazards have been eliminated. However, from a practical point of view, dynamic hazards actually cause far less difficulty in sequential circuits than do static hazards. Consequently the cut- and tie-sets which eliminate static hazards may hereafter be called hazard-preventing even though their inclusion in a network does not always prevent dynamic hazards.

D. A. HUFFMAN

III. Derivation of Canonic Forms for Hazard-Free Contact Networks

3.1. We define a hazard-free contact network as one which, for single changes in its switching variables, does not change its transmission at all if the steadystate transmissions before and after the variable change are the same, and which changes its transmission only once if the initial and final steady-state transmissions are complementary to each other. It is assumed that no contact changes its transmission more than once, if at all, during the instability of the corresponding relay (the contact does not bounce), but that the several contacts on that relay may change their states of transmission in any order whatsoever. In other words: From its terminals a hazard-free network looks like a contact which does not bounce if only the contacts which form the network do not bounce. The restriction to single change of switching variable is not a serious one since sequential circuits must be operated in essentially this manner anyhow [4, 6].

3.2. We have seen (Section 2.2) that static hazards may be prevented by providing within a network both certain cut-sets as well as certain tie-sets. We now prove that if a network is composed of a parallel combination of cascade connections of contacts corresponding to the specific tie-sets which prevent static hazards when the network transmission is unity, it will also be free from static hazards when the transmission is zero. As an example, note the network of Fig. 4-a which has the hazard-preventing tie-sets shown in the corresponding map. The network can be closed only when one (or more) of its parallel branches is closed. If the network transmission is zero, both before and after a single variable change, then each tie-path is open independent of the value of the changing variable. That is, each tie-path contains at least one contact which does not belong to the unstable relay and which is open while that relay is changing state. It is clear that no momentary unity transmission for the entire network is then possible. Thus the canonic network which consists of paralleled hazard-preventing tie-paths is free from all static hazards, and incidentally, from dynamic hazards as well.

A similar argument proves that the network which is formed from a cascade combination of parallel contacts corresponding to the cut-sets which prevent static hazards when the network transmission is zero, will also be free from all hazards. (See Fig. 4-b.)

3.3. Networks in which all static hazards have been eliminated can be used for the control of the secondary (memory) devices in a sequential switching circuit with the assurance that no first-order hazards will then be possible. (For the definition of an n^{th} -order hazard see the summary at the end of this paper.) However, if the network (as in Fig. 4-a, b) contains *only* those tie- and cut-sets which eliminate static hazards it can further be guaranteed that the network will be free from dynamic hazards as well. For, if, during the change of state of one of the variables in this latter kind of network, one of its tie-sets (cut-sets) becomes closed (open) then this tie-set (cut-set) cannot immediately reopen (reclose) to cause a dynamic hazard since only one contact from the unstable relay is in that tie-set (cut-set), and since it has been assumed that that contact does not bounce. It is, however, possible to have a network free of dynamic hazards even



FIG. 4. Canonic forms of hazard-free contact networks and an equivalent hazardous network.



FIG. 5. A hazard-free network

when additional tie- or cut-sets are added to those which prevent static hazards, *if* the additional sets are carefully chosen. The network of Fig. 5 is such a network.

Implicit in the derivation of the two canonic forms of a hazard-free contact network is the conclusion that every switching function has at least one hazardfree realization.

The two canonic forms of Fig. 4-a, b may be simplified as in Fig. 4-c, d while maintaining the original cut- and tie-sets. Even these reduced networks contain twice the number of contacts than does the minimum contact realization of the function under discussion. (See Fig. 4-e.) This latter network is algebraically equivalent to the first four networks but is replete with hazards. In this example at least, the elimination of hazards required a substantial increase in the number of contacts.

Not all economical networks have hazards, nor are all hazard-free networks necessarily in series-parallel form. The network of Fig. 5, which can be shown from its map to be hazard-free, demonstrates the point. In addition, it is easy to prove that any network which contains only normally open or normally closed contacts from each of its input relays is hazard-free.

IV. The Analysis and Prevention of Hazards in Gate Networks

4.1. For the purposes of this section, the symbols 0 and 1 may be considered to represent low and high voltages, respectively. The gate networks will be composed of three fundamental kinds of elements with interconnected signal leads. These elements are the multiply, add, and inversion gates (sometimes known as the "and", "or", and "negation" gates, respectively). Other gate elements can be represented by simple combinations of these three gates.

In order to parallel the thoughts in the analysis of contact networks, we introduce here the concepts of drop-sets and lift-sets which will be related to the structure of gate networks and which are analogous to cut-sets and tie-sets, respectively.

The gate networks shown in Fig. 6 each produce the function f = a'b + ac = (a + b)(a' + c). The drop-sets of the first network are (a, b), (a', c), and (b, c) and its lift-sets are (a', b) and (a, c). For example, (b, c) is a drop-set since when b = c = 0 (when b and c are both "dropped") the output voltage, f, is certain to be dropped to the zero level also. And (a', b) is a lift-set because when a' = 1 (a = 0) and b = 1 (when a' and b are "lifted") the output voltage, f, is necessarily lifted to the one level also, no matter how the c voltage is changing.



FIG. 6. Two hazardous gate networks and their maps

Note that in the sense used here the set (b, c) is not a lift-set for the network in Fig. 6-a since even though b = c = 1 neither the signal a'b nor ac is thereby directly implied to be unity; therefore neither is the function f. This conclusion can be seen more clearly by setting b = c = 1 and changing the variable a from zero to one and back. If there were no signal delay along either of the dotted paths (or if the delays were equal) then the inputs into the final add gate would always be complementary and f would always be unity even while a was changing.

4.2. Physically, however, we are sure that the signal delays are not precisely equal. If, for instance, the delay along the upper path is greater than that along the lower, and if (while b = c = 1) a changes from one to zero, then ac will become zero shortly before a'b becomes equal to one. For this short interval the output, f, will also be zero. This hazard is shown by an arrow on the corresponding map.

It might be guessed that a greater delay in the lower signal path would eliminate the hazard above. And so it will. But if we assume that the lower path has the greater delay, consider what happens when a changes from zero to one (still keeping b = c = 1). Now a'b will become zero before ac becomes one and another hazard has been generated. In this situation, (unlike the somewhat analogous contact network of Fig. 1-a), hazards cannot be eliminated by adding delays in the signal paths.

In Fig. 6-b the gate network can be shown to have a static hazard for b = c = 0 which cannot be eliminated unless the delays along the two parallel dotted paths could be made exactly equal. If the gates are realized from physical devices, which necessarily have slowly changing characteristics, this is an impossible ideal to achieve.

4.3. With the analogy we have established between cut- and tie-sets and dropand lift-sets, it is clear that any function can be realized in a completely hazardfree gate network in forms analogous to the two canonic forms of Fig. 5-a, b. In these networks paralleling of contacts corresponds to combination of signals in an add-gate and cascading of contacts corresponds to combination of signals in a multiply-gate.

V. An Example of the Use of Hazard-Free Networks

5.1. We have mentioned previously that the reason for designing hazard-free networks is that their use in sequential circuits helps to insure the operation planned by the circuit designer. In this section we sketch the synthesis for a relatively simple sequential circuit to show what additional complexity in its final design might typically be required. Only the outline of the synthesis is presented here. It is but an example of the general procedure which has been proposed in [4].

The circuit we wish to design has two binary inputs, X_1 and X_2 , and one binary output, Z. (The meanings of the symbol values, 0 and 1, need not be decided upon until later.) It is assumed that only one input may change at a time. When X_2 changes from 0 to 1, the output is to change to (or remain at) the value of X_1 then existing. Otherwise, Z is to remain constant. A typical sequence of terminal values is given in Fig. 7.

The primitive flow table, which is an exact statement of the specifications above, is given in Fig. 8-a. One of two possible merged flow tables is given in Fig. 8-b along with a valid secondary state assignment. The Y and Z matrices follow in a straightforward way. They are map representations of the switching functions Y_1 , Y_2 , and Z which are dependent upon the variables x_1 , x_2 , y_1 , and y_2 (see Fig. 8-e). In the Z matrix, the parenthesized entries must be chosen as they have so as to prevent false outputs during transient circuit states. The two dashed entries correspond to transient states for which the output may be chosen as either zero or one. The blank cells in both the Y and Z matrices represent nonoccurring states and therefore the values of the functions may be chosen arbitrarily here also.

5.2. For a relay realization of the circuit, we chose the Y_1 and Y_2 functions as shown in Fig. 9-a, b, and $Z = y_2$. One economical contact network for generating the Y's is given in Fig. 9-c. The dotted y_1 contact is not necessary except to provide the dotted cut-sets in Fig. 9-a, b. With this contact, the network is hazardfree and no special timing between the x_2 and x_2' contacts is necessary. Without





(e) Functional diagram FIG. 8. Flow table synthesis of a sequential circuit

the additional y_1 contact, it would be necessary to be certain that there was a short interval of time during which both x_2 and x_2' contacts were open. The modified circuit of Fig. 9-d provides additional simplicity if it is not necessary to give isolation at inputs and output. For this realization, $Z = y_2 + x_1x_2y_1$, and the hazard-preventing cut- and tie-sets remain the same for Y_1 and Y_2 .

Note that the two hazards shown in Fig. 9-a, b correspond to transitions which may actually take place during the operation of the circuit. These transitions (from Fig. 8-b) are $\bigcirc \rightarrow 3$ and $\bigcirc \rightarrow 7$. If these changes in circuit state had not been possible, the associated hazards could have no effect on circuit behavior and would not need to be eliminated.

5.3. For an all-triode realization of the circuit, it is convenient to choose Y_1 and Y_2 in accordance with the maps of Fig. 10-a, b. The lift-sets shown dotted prevent the hazards indicated. They are generated by adding the dotted gate in the gate network of Fig. 10-c. One possible circuit realization of this network



FIG. 9. Relay realization of a sequential circuit

(Fig. 10-d) is obtained by associating inverter gates with plate-loaded triodes and multiply gates with resistive circuits for logical multiplication. These latter are chosen so that both of their input voltages must be high before the following grid voltage is raised above the cut-off level. Or, alternatively, diode "and" circuits may be used.

5.4. The gate network of Fig. 10-c has been modified slightly, rearranged, and its gates renumbered in Fig. 11-a. The changes preserve the hazard preventing lift- and drop-sets of the maps of Fig. 10-a, b. By associating the gates of the revised network with the transistor circuits [7] of Fig. 11-b, the completed sequential transistor switching circuit of Fig. 11-c results. Again we have illustrated that additional physical devices, even though not affecting the steady-state algebraic relationships within a circuit, may yet prevent the hazards which endanger proper operation during transients.



VI. Summary

This paper has contained a suggestion for a method for the elimination of hazards in combinational switching networks. It is possible [4, 6] that when a combinational network is used for the control of a memory device in a sequential



(c) Final circuit FIG. 11. Surface barrier transistor realization of a sequential circuit



FIG. 12. Illustrating a third-order hazard

switching circuit more than one of its variables may be allowed to change state simultaneously. Through an extension of the procedure developed in this paper it is possible to prove that such a network can be made hazard-free for the multiple changes of variable to which it may be subject, if its tie-sets correspond to the prime implicants [8] of the transmission function expressed in sum-ofproducts form and if its cut-sets correspond to terms which are dual to prime implicants in the product-of-sums form.

When hazard-free networks are used for the control of the binary memory devices in a sequential circuit, they prevent the major cause of departure of circuit operation from the ideal: momentarily improper excitation signals to the devices.

Another related, but more obscure, cause of non-ideal circuit operation, multiple-order hazards, is illustrated in Fig. 12. (This is not meant to be a useful circuit.) Assume initially that all relays are unoperated. Grounding of the input operates the A relay. Ideally both contacts a_1 and a_2 should close simultaneously, which action would be followed by the operation of B, C, and D. If, instead, the a_1 contact closes (before the a_2 contact closes) and operates B, it is conceivable that C and D also operate before the a_2 contact closes. If this happens, the sequence of action in the B relay controlling network is: a_1 closes, d' opens, and a_2 closes. This non-ideal excitation of B constitutes what we will call a third-order hazard because it exists only if the difference in closing times of the a_1 and a_2 contacts is more than the response time of the three relays B, C, and D.

Clearly, hazards of higher and higher order are less and less likely to cause departure of circuit behavior from that planned by its designer. Even first-order hazards do not usually cause trouble because all memory devices have at least a little smoothing (inertial) action on the signals they handle. The subject of multiple-order hazards is, however, one of theoretical interest. A careful analysis of Figs. 10-c and 11-a shows that possible second-order hazards are present if there is sufficient delay at the points labeled H. The hazards may be eliminated by inserting delay (no smoothing is necessary) in the (Y_1, y_1) lead; no delay or smoothing is necessary in the (Y_2, y_2) lead.

Acknowledgments

I would like to express my appreciation to the students in the course on switching at the Massachusetts Institute of Technology upon whom some of the ideas in this paper were tested, and whose stimulation was essential in forming them. In particular, I want to acknowledge the indirect help given by Al Lett, James Bushnell, and Dan Daggett, who, at my suggestion, pursued the subject of this paper along slightly different lines.

REFERENCES

- 1. SHANNON, C. E., A Symbolic Analysis of Relay and Switching Circuits, A.I.E.E. Transactions, 59 (1938), 713-722.
 - -----, The Synthesis of Two-Terminal Switching Circuits, Bell System Technical Journal, 28, Jan. 1949.
- 2. STAFF OF THE COMPUTATION LABORATORY, The Synthesis of Electronic Computing and Control Circuits, Harvard University Press., Cambridge, Mass., 1951.
- KEISTER, W., RITCHIE, A. E., and WASHBURN, S. H., The Design of Switching Circuits, D. Van Nostrand Co., Inc., New York, 1951, 151-171.
- HUFFMAN, D. A., The Synthesis of Sequential Switching Circuits, Journal of the Franklin Institute, 257, nos. 3 and 4, March and Apr. 1954, 294-295.
- KARNAUGH, M., The Map Method for Synthesis of Combinational Logic Circuits, A.I. E.E. Technical Paper #53-217, Apr. 1953.
- HUFFMAN, D. A., A Study of the Memory Requirements of Sequential Switching Circuits, Technical Report # 293, Research Lab. of Electronics, Mass. Inst. of Tech., Apr. 1955.
- 7. BETER, R. H., BRADLEY, W. E., BROWN, R. H., and RUBINOFF, M., Surface-Barrier Transistor Switching Circuits, I.R.E. Convention Record, part 4, March, 1955, 139-145.
- QUINE, W. V., The Problem of Simplifying Truth-Functions, Amer. Math. Monthly, 59, Oct., 1952.