EVALUATION OF FUZZY RULEMAKING FOR EXPERT SYSTEMS FOR FAILURE DETECTION

Frank Laritz and Thomas B. Sheridan
Man-Machine Systems Laboratory
Massachusetts Institute of Technology

INTRODUCTION

Computer aids in the form of so called "expert systems" have been proposed repeatedly for making diagnoses of failures in complex systems.

The fuzzy set theory of Zadeh[1] has been shown to offer an interesting new perspective for modeling the way humans think and use language. In particular, we assume that real expert human operators of aircraft, power plants and other systems do not think of their control tasks or failure diagnosis tasks in terms of control laws in differential equation form, but rather keep in mind a set of rules of thumb in fuzzy form. For the reader ignorant of fuzzy sets the experiment described below communicates by example the gist of the idea.

FIRST EXPERIMENT

Five subjects repeatedly adjusted two "inputs" A and B to a "black box" to any value between 10 and 100, set a "failure mode" to any one of four available settings including "no failure", and observe two "outputs" C and D. The contents of the black box were not revealed. The subjects' task was to correlate inputs and outputs with failure modes and from this infer rules by which to assert whether and in what mode the black box had "failed" as a function of the two inputs and two outputs.

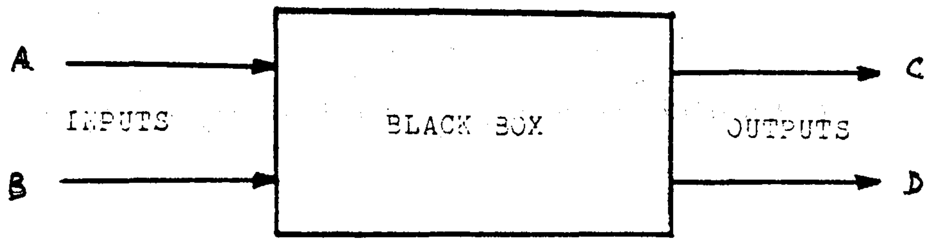
Actually the black box was a simple resistor network as shown in Figure 1 in which one of the resistors 1,2,3,4 was selectively opened (or none was).

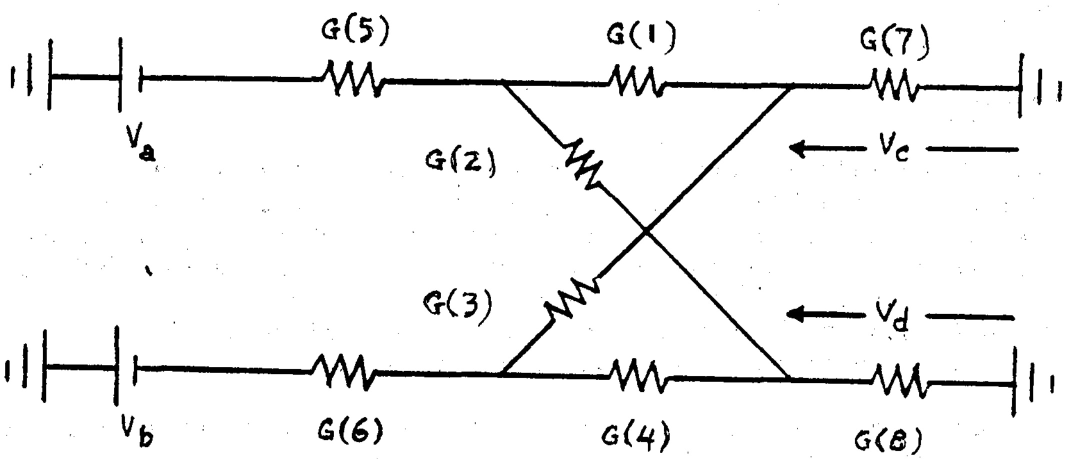
After each subject had completed a number of trials (they were all really learning trials) he was asked to formulate rules in terms of easy-to-remember descriptors for the four variables like "low", "medium" and "high" using these descriptors he was to generate rules such as:

"when A is low and B is medium or high and C is high and D is medium or high, the failure is mode 2".

There could be any number of such (fuzzy) descriptors and any number of such rules, and the subjects were free to format them in tables or however they wished. They could also combine variables in forms such as C/D and C-D.

The subjects were also asked to produce functions of each descriptor (fuzzy set) defining what they "meant". Each function specified "membership" or "truth" as a function of the values of the corresponding variable (in the range 10-100). Two of the five subjects observed the black box behavior first, then devised the rules, and lastly devised membership functions.





Resistors Allowed to Fail:	G(1) = 61 mhos	
	G(2) = 87	
	G(3) = 59	
	G(4) = 76	
Other Resistors:	G(5) = 95	
	G(6) = 55	
	G(7) = 74	
	G(8) = 85	

Figure 1. Simple resistor network comprising the "black box"

The others chose to invent terms and define the membership functions first.

As an example Figure 2 lists the rules given by one subject (JR) and Figure 3 presents his membership functions. Note that certain regions of A,B and D-C were (apparantly intentionally) not covered by his membership functions (and rules). For contrast the membership functions of a second subject are also shown (Figure 4).

For each subject independently the experimenter derived the state-action matrix (failure mode as a function of input and output numerical values) using the conventioanl "max m" for "or" and min m" for "AND". He then proceeded to evaluate each resulting expert system not only against single complete failures (the basis in which the subjects made up their rules) but also on multiple complete failures and single partial failures (5% changes rather than 100% changes in resistance). For a given set of inputs and outputs each subject's expert system yielded a "truth value" for each failure mode for each combination of A,B,C,D. A simple procedure is to assert failure for that mode having the greatest truth value greater than some threshold and no failure for truth less than that threshold. Laritz used this as one decision criterion (which he called the "most true" criterion) but also counted the number of times u for each mode exceeded 0.5 (the "times true" criterion), and the sum of truth values for each mode ("truth summation" criterion). Figure 5 summarizes the rather impressive success of subject JR's expert system, and for comparison Figure 6 summarizes that of subject DM. The performances of the other fuzzy expert systems lay somewhere in between.

CONCLUSIONS FROM FIRST EXPERIMENT

From this first experiment we concluded:

1. The method of observing trends, then formulating rules, and then defining fuzzy values captures more of the human's ingenuity and pattern recognition ability and provides a better expert failue detection system than the method or creating fuzzy values, then gathering data, and then deducing rules.

2. If the second method is used, it is best to put the membeship functions for the fuzzy values on paper at the outset so that there will be no loss of information later.

3. Expert systems using non-fuzzy values require perfect failure rules. When the rules are not perfect, the expert system does not perform well.

4. Although not explicitly defined for this purpose in the investigation, the fuzzy expert systems did remarkably well in detecting and locating multiple and partial failures. This means that fuzzy methods have some robustness.

(1,1) If A is high and B is low and D is significantly greater than C, then the system is in failure mode 1.

(2,1) If A is high and B is low and C is significantly greater than D, then the system is in failure mode 2.

(3,1) If A is low and B is high and D is significantly greater than C, then the system is in failure mode 3.

(4,1) If A is low and B is high and C is significantly greater than D, then the system is in failure mode 4.

(5,1) If A is high and B is high and D is slightly greater than C, then the system is in failure mode 0.

Figure 2. **Fuzzy decision rules inferred by subject JR**

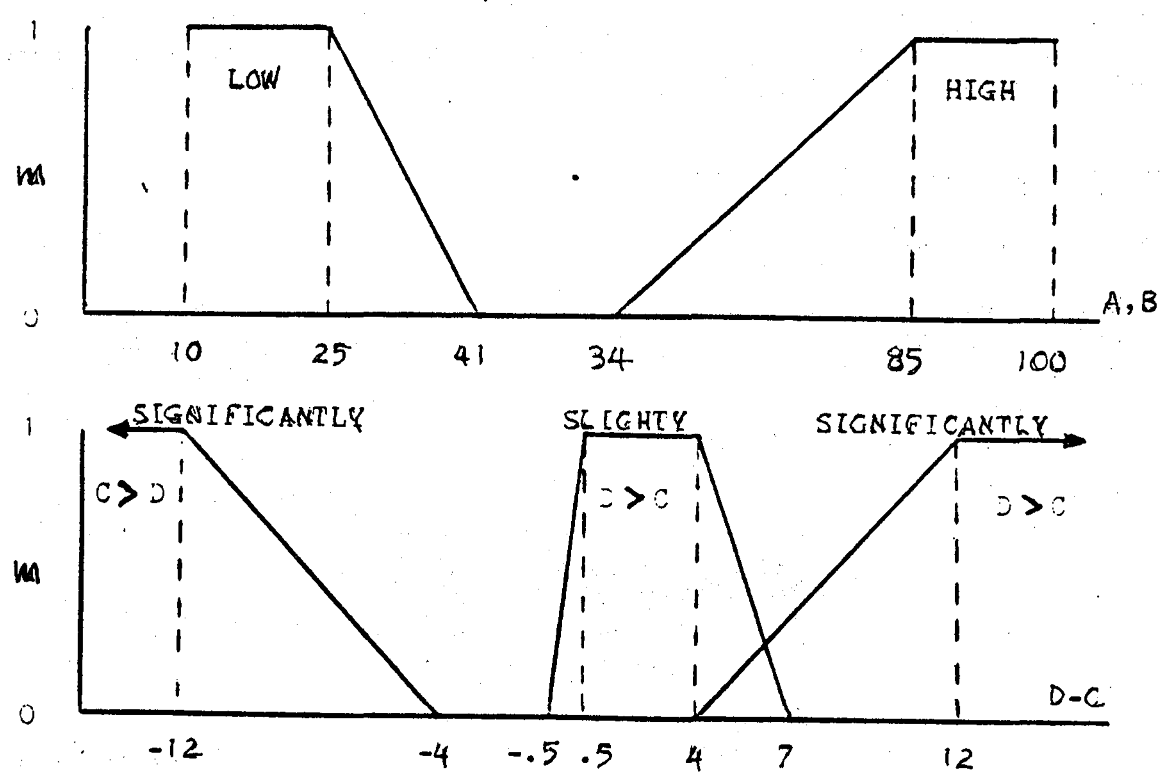


Figure 3. **Membership functions devised by subject JR**

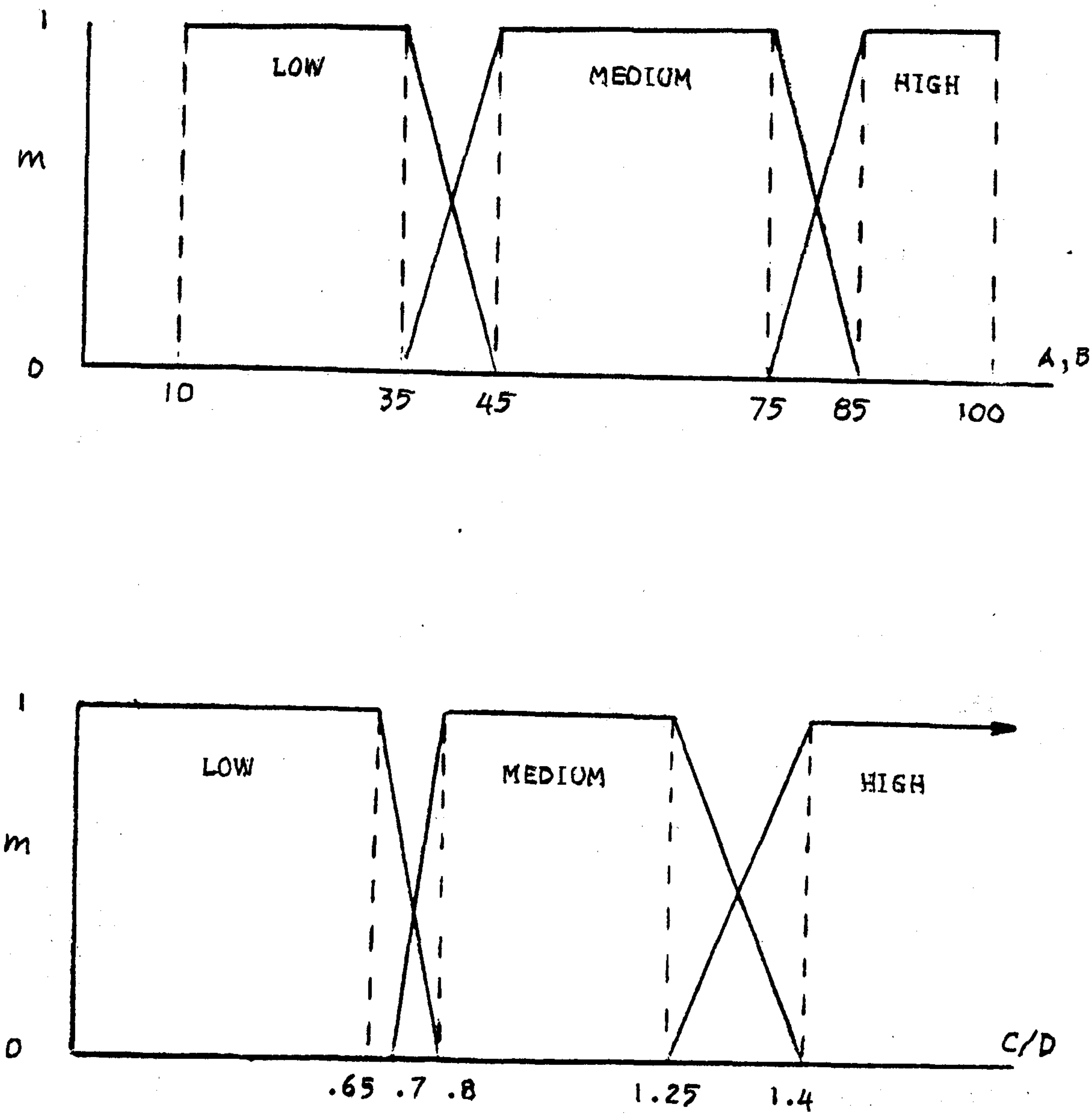


Figure 4. Membership functions devised by subject DM

TEST 1: SINGLE COMPLETE FAILURE

ACTUAL FAILURE	IDENTIFIED FAILURE		
	MOST-TRUE	TIMES-TRUE	TRUTH-SUMMATION
1	1	1	1
2	2	2	2
3	3	3	3
4	4	4	4
0	0	0	0
	---	---	---
SCORE:	5/5	5/5	5/5

TEST 2: MULTIPLE COMPLETE FAILURES

ACTUAL FAILURES	IDENTIFIED FAILURE		
	MOST-TRUE	TIMES-TRUE	TRUTH-SUMMATION
1,3	3	3	1
1,4	1,4	1	1
2,3	2,3	3	3
2,4	2,4	2	2
	---	---	---
SCORE:	4/4	4/4	4/4

TEST 3: SINGLE PARTIAL FAILURE

ACTUAL FAILURE	IDENTIFIED FAILURE		
	MOST-TRUE	TIMES-TRUE	TRUTH-SUMMATION
1	1 (55%)	1 (50%)	1 (55%)
2	2 (75%)	2 (70%)	2 (55%)
3	3 (70%)	3 (70%)	3 (65%)
4	4 (80%)	4 (75%)	4 (65%)
	---	---	---
SCORE:	4/4	4/4	4/4
TOTAL SCORE:	13/13	13/13	13/13

* = INCORRECT DECISION

Figure 5. Results of applying JR's expert system

TEST 1: SINGLE COMPLETE FAILURE

ACTUAL FAILURE	IDENTIFIED FAILURE		
	MOST-TRUE	TIMES-TRUE	TRUTH-SUMMATION
1	1	1	1
2	* 1,2	2	2
3	* 1,3	* 1	* 1
4	* 1,4	4	4
0	0	0	0
SCORE:	2/5	4/5	4/5

TEST 2: MULTIPLE COMPLETE FAILURES

ACTUAL FAILURES	IDENTIFIED FAILURE		
	MOST-TRUE	TIMES-TRUE	TRUTH-SUMMATION
1,3	1,3	1	1
1,4	* 1,4,0	1	1
2,3	* 2,3,4,0	2	2
2,4	2	2	2
SCORE:	2/4	4/4	4/4

TEST 3: SINGLE PARTIAL FAILURE

ACTUAL FAILURE	IDENTIFIED FAILURE		
	MOST-TRUE	TIMES-TRUE	TRUTH-SUMMATION
1	1 (80%)	1 (55%)	1 (50%)
2	* 4	* 4	* 4
3	3 (70%)	3 (65%)	3 (60%)
4	* 1,4	4 (30%)	* 1
SCORE:	2/4	3/4	2/4
TOTAL SCORE:	6/13	11/13	10/13

* = INCORRECT DECISION

Figure 6. Results of applying DM's expert system

5. The decision method can be chosen to suit the strength and tightness of the rules. Stronger rules require less margin for error.

6. Expert systems which have approximately the same number of rules for each failure mode perform better than those with an uneven distribution

SECOND EXPERIMENT

As a second experiment the first author used himself as a subject on a black box resistor network that was much more complex (sufficiently so that he had no advantage over a subject who did not know what was inside). Again there were two adjustable inputs and two resulting outputs but this time eight failure modes. The first author experimented and observed, then derived his rules, (Figure 7) then defined his membership functions, and finally derived an expert system on the same basis as before. Results showed that the expert system worked perfectly on complete failures but faltered on multiple complete failures and partial failures (Figure 8). Further attempts to refine his decision rules showed little gain in discriminabililty.

CONCLUSIONS FROM SECOND EXPERIMENT

One can conclude from these results that a computer, given relatively little knowledge in fuzzy form from persons who are "expert" in the behavior of a sufficiently simple system under complete failures, can perform very well in identification of such failures. But when the system is complex <u>and</u> failures are multiple or partial <u>and</u> the expert's knowledge is not derived on the basis of experiencing such failures, an expert system cannot be expected to perform very well.

REFERENCES

1. Zadeh, L.A., Fuzzy Sets, <u>Information and Control</u>, 8, 338-353.

2. Laritz, F.J., The Use of Fuzzy Sets in Failure Detection, MIT SM Thesis, December 1983.

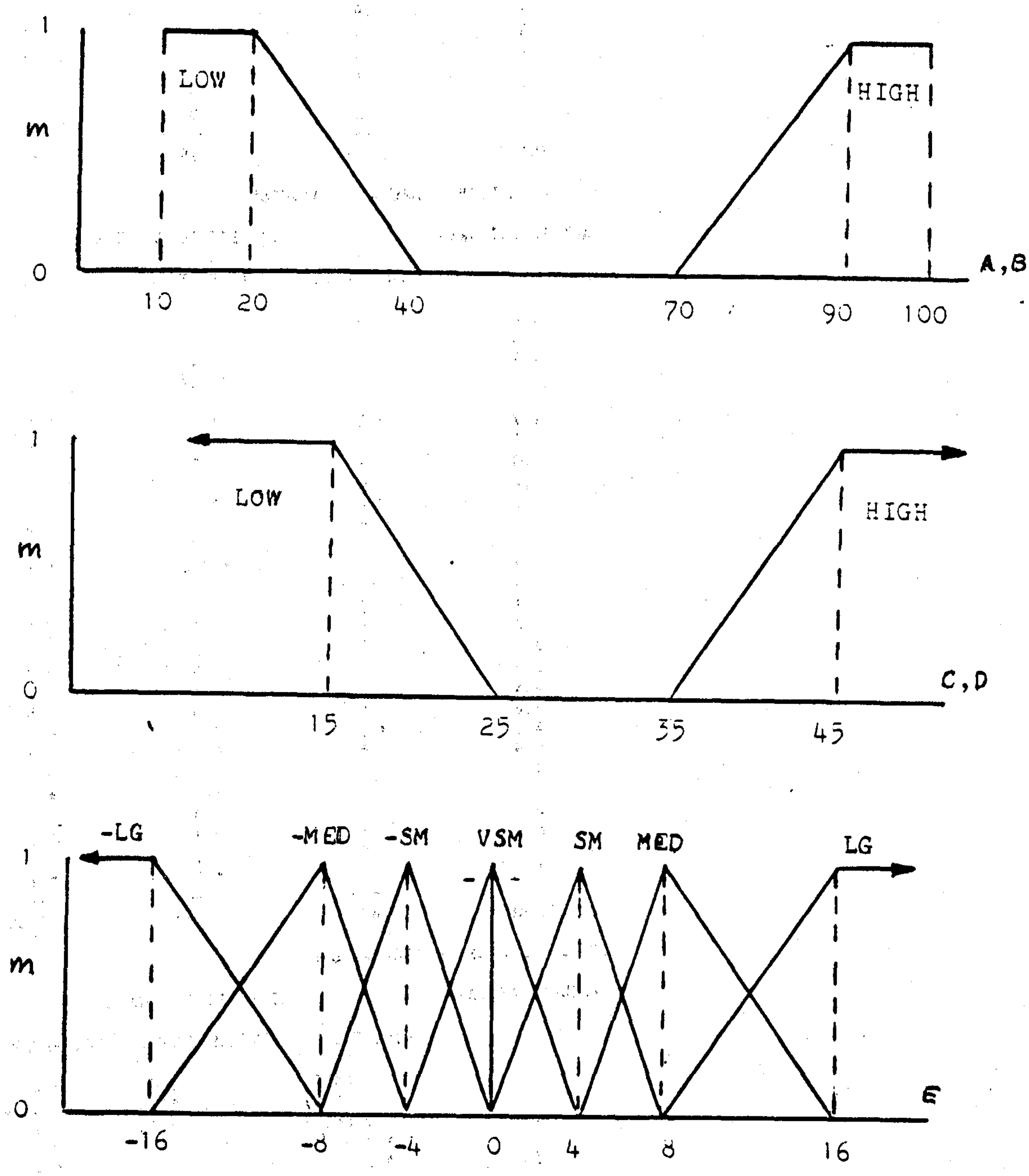


Figure 7. Membership functions devised by Laritz for variables of complex resistor network

TEST 1: SINGLE COMPLETE FAILURE

ACTUAL FAILURE	IDENTIFIED FAILURE		
	MOST-TRUE	TIMES-TRUE	TRUTH-SUMMATION
1	1	1	1
2	2	2	2
3	3	3	3
4	4	4	4
5	5	5	5
6	6	6	6
7	7	7	7
8	8	8	8
0	0	0	0
SCORE:	9/9	9/9	9/9

TEST 2: MULTIPLE COMPLETE FAILURES

ACTUAL FAILURES	IDENTIFIED FAILURE		
	MOST-TRUE	TIMES-TRUE	TRUTH-SUMMATION
1,2	2	2	2
1,3	1	1	1
1,4	1	1	1
1,5	1	* 1,7	1
1,6	1	1	1
1,7	* 5	* 5	* 5
1,8	8	8	8
2,3	2	2	2
2,4	4	4	4
2,5	* 4	* 4	5
2,6	6	6	6
2,7	7	7	2
2,8	2	* 1,2	* 1
3,4	4	4	4
3,5	5	5	5
3,6	6	6	3
3,7	7	7	7
3,8	8	8	8
4,5	* 7	* 7	* 7
4,6	6	6	6
4,7	* 5	* 5	* 5
4,8	* 6	* 6	* 6
5,6	* 1	* 1	* 1
5,7	* 0	* 0	* 0
5,8	* 2	* 0	* 2
6,7	* 4	* 8	* 8
6,8	6	6	6
7,8	* 4	* 4	* 1
SCORE:	18/28	16/28	18/28

TEST 3: SINGLE PARTIAL FAILURE

ACTUAL FAILURE	IDENTIFIED FAILURE		
	MOST-TRUE	TIMES-TRUE	TRUTH-SUMMATION
1	1 (55%)	1 (55%)	* 2
2	2 (65%)	2 (65%)	2 (25%)
3	3 (60%)	3 (60%)	* 2
4	* 3	* 3	* 2
5	* 2	* 2	* 4
6	* 8	* 8	* 8
7	* 4	* 4	* 2
8	8 (55%)	8 (55%)	* 2
SCORE:	4/8	4/8	1/8
TOTAL SCORE:	31/45	29/45	28/45

* = INCORRECT DECISION

Figure 8. Results of applying Laritz' expert system to failures of complex resistor network