LOCAL METHODS FOR LOCALIZING FAULTS IN ELECTRONIC CIRCUITS

by
Johan de Kleer

Abstract:

The work described in this paper is part of an investigation of the issues involved in making expert problem solving programs for engineering design and for maintenance of engineered systems. In particular, the paper focuses on the troubleshooting of electronic circuits. Only the individual properties of the components are used, and not the collective properties of groups of components. The concept of propagation is introduced which uses the voltage-current properties of components to determine additional information from given measurements. Two propagated values can be discovered for the same point. This is called a coincidence. In a faulted circuit, the assumptions made about components in the coinciding propagations can then be used to determine information about the faultiness of these components. In order for the program to deal with actual circuits, it handles errors in measurement readings and tolerances in component parameters. This is done by propagating ranges of numbers instead of single numbers. Unfortunately, the comparing of ranges introduces many complexities into the theory of coincidences. In conclusion, we show how such local deductions can be used as the basis for qualitative reasoning and troubleshooting.

INTRODUCTION

Troubleshooting involves determining why a particular correctly designed piece of equipment is not functioning as it was intended; the explanation for the faulty behavior being that the particular piece of equipment under consideration is at variance in some way with its design. To troubleshoot, a sequence of measurements must be made to localize this point of variance, or fault. The problem for the troubleshooter is to determine what a particular measurement tells him and what measurement to make next.

This paper investigates how local knowledge about the circuit can be used to answer these two questions. By *local*, we mean that only one particular component in the circuit will be considered at one time and any interactions between larger collections of components will be ignored. The teleology of collections of more than one component will not be discussed; instead only the characteristics of the individual components will be used (such as their VIC's -- the voltage-current characteristics).

The central goal of this research is to achieve a better understanding of troubleshooting. One role for this new knowledge is in an expert problem solving program. However, it can also be used in the expert component of an ICAI tutoring system. <Brown *et.al.*, 74> This means that there has to be some communication between the troubleshooting strategy and the human student. In fact, this is also true if we wanted the expert problem solver to explain its deductions. Therefore we have imposed the constraint that our troubleshooter's deductions be explainable. This constraint has motivated many of the design choices in the implementation of this theory as a program (INTER). In this paper we also include some comments about how the theory can be used in a tutoring context.

The way to obtain new information about the circuit is to make a measurement. In troubleshooting, new information is provided by coincidences. In the most general sense a *coincidence* occurs when a value at one particular point in the circuit can be deduced in a number of different ways. Such a coincidence provides information about the assumptions made in the deductions. A coincidence can occur in many different ways; it can be the difference between an

expected value and a measured value (e.g. expected output voltage of the power supply and the actual measured value); it can be the difference between a value predicted by Ohm's law and a measured value; or it can be the difference between an expected value and the value predicted by the circuit designer. There are numerous other possibilities.

A troubleshooting investigation into a particular circuit proceeds in two phases. The first involves discovering more values such as currents and voltages occurring at various points in the circuit, and the second involves finding *coincidences*. The usefulness of coincidences is based on the fact that nothing can be discovered about the correctness of the circuit with a measurement unless something is known about the value at that point of the circuit in the first place. If nothing is known about that point, a measurement will say nothing about the correctness of the components. One actual measurement implies many other values in the circuit. The first phase of the investigation involves discovering many such values in the circuit, and the second involves making measurements at those points for which we know the implied values so that we can see whether the circuit is acting as it should, or if something is wrong.

We will call such an implication a *propagation* and the discovery of a value a point at which we already know a propagated value for a *coincidence*. When these two values are equal, we will call such a coincidence a *corroboration* and when they are different we will call it a *contradiction*.

Information about the faultiness of components in the circuit can only be gained through coincidences. Propagations involve making certain assumptions about the circuit and then predicting values at other points from these. These assumptions can be of many kinds. Some of them involve just assuming the component itself is working correctly. For example, we can derive the current through a resistor from the voltage across it. Others require knowing something about how the circuit should work, thus predicting what values should be. For example, knowing the transistor is acting as a class A amplifier, we can assume it is always forward-biased. Coincidences between propagated values and new measurements provides information about the assumptions made in the propagation.

Coincidences between propagated values and values derived from knowing how the circuit should work require a teleological description of the circuit. As indicated earlier, this paper does not

investigate these latter kinds of assumptions. Research into this area was pursued by Brown <Brown, 74> <Brown, 76>. Instead, this paper investigates propagations employing only assumptions about the components themselves. Although, at first sight, the teleological analysis of troubleshooting is the more interesting, it cannot proceed without being able to propagate measurements in the circuit.

It may appear that this kind of circuit reasoning is essentially trivial and thus should not be investigated. This paper will show that the issues of local nonteleological reasoning are, in fact, very difficult. Some of the problems arise because the nonteleological knowledge should interact with the teleological knowledge. A particularly difficult problem which will arise again and again is the question of how far to propagate values. Often the propagations will be absurd, and only a small amount of teleological knowledge would have pruned out these uninteresting propagations. Part of the effort of this paper is directed into determining what other kinds of knowledge and interaction is required, aside from the nonteleological, in order to troubleshoot circuits effectively.

The sections that follow present an evolution of the knowledge required. The first sections will present a simple theory about local reasoning and troubleshooting. Next the problems of the approach will be investigated, and some of them answered by a more sophisticated theory. Finally the deficiencies of the theory and how it must interact with more teleological knowledge will be discussed.

SIMPLE LOCAL ANALYSIS

The domain of electronics under consideration will be restricted to DC circuits. These are circuits consisting of resistors, diodes, zener diodes, capacitors, transistors, switches, potentiometers and DC voltage sources. All AC effects will be ignored although an analogous type of analysis would work for AC circuits. It will be assumed that the topology of the circuit does not change so that wiring errors or accidental shorts will not be considered as possible faults.

In this section we will present a simple theory of propagation. Initially, only numeric values will be propagated. Interacting local experts produce the local analysis. Each kind of component has a special expert which, from given input conditions on its terminals, computes voltages and

currents on other terminals. For example, the expert for a transistor might, when it sees a base-emitter voltage of less than .55 volts, infer a zero current through the collector.

This propagation scheme is very similar to that used in EL <Sussman & Stallman, 75> <Stallman & Sussman, 76>. Although similar in that they are both based on propagation of constraints, the different goals of analysis and troubleshooting lead to many differences in the details of the two propagation schemes. Therefore, we include a very terse description of our propagation scheme, and the reader is referred to the two EL papers for a deeper explanation of propagation of constraints.

Since EL is primarily interested in analysis, it must discover every value in the circuit. When conventional numeric propagation fails it resorts to propagating variables and solving algebraic equations. Since we are mainly interested in explaining and not analysis the propagation of variables and solving of equations is not done.

In order to give explanations for deductions, a record is kept as to which expert made the particular deduction. Most propagations make assumptions about the components involved in making it, and these are stored on a list along with the propagated value. Propagations are represented as:

    (<type> <location> (<local-expert> <component> <arg>) <assumption-list>)
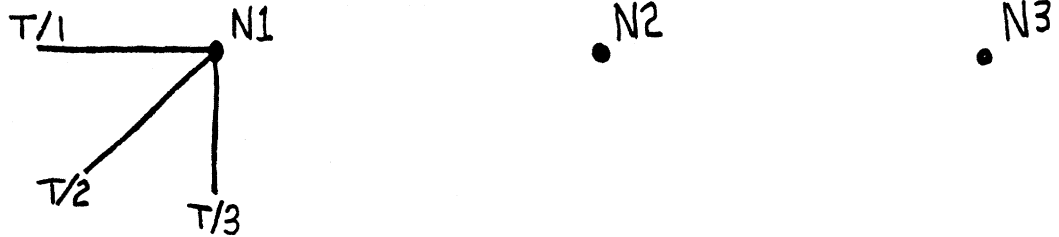
<type> is VOLTAGE or CURRENT.

<location> is a pair of nodes for a voltage and a terminal for a current.

Note that every such propagation has a value associated with it. For those examples where the exact numerical value is important, exact numbers will be included.

The simplest kinds of propagations require no assumptions at all. These are the Kirchoff voltage and current laws.

The circuit consists of components such as resistors and capacitors etc., *terminals* of these components are connected to *nodes* at which two or more terminals are joined. In the above diagram T/1, T/2 and T/3 are terminals and N1, N2 and N3 are nodes. Currents are normally associated with terminals, and voltages with nodes.

Kirchoff's current law states that if all but one of the terminal currents of a component or node is known, the last terminal current can be deduced.

```
(CURRENT T/1)

(CURRENT T/2)

(CURRENT T/3 (KCL N1) NIL)
```

Since faults in circuit topology are not considered, KCL makes no new assumptions about the circuit.

Kirchoff's voltage law states that if two voltages are known relative to a common point, the voltage between the two other nodes can be computed:
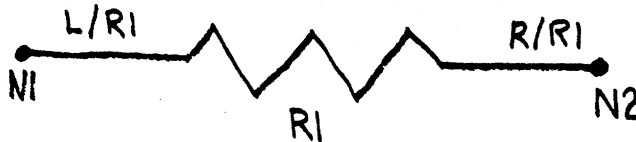
```
(VOLTAGE (N1 N2))

(VOLTAGE (N2 N3))

(VOLTAGE (N1 N3) (KVL N1 N2 N3) NIL)
```

As with KCL, KVL makes no new assumptions about the circuit.

One of the most basic types of the circuit elements is the resistor. Assuming the resistance of

the resistor to be correct, the voltage and current can be deduced from each other using Ohm's law:


```
(CURRENT R1)

(VOLTAGE (N1 N2) (RESISTORI R1) (R1))


(VOLTAGE (N1 N2))

(CURRENT R1 (RESISTORV R1) (R1))
```

(In all the example propagations presented so far it was assumed that the prerequisite values had no assumptions, otherwise they would have been included in the final assumption list.)

These three kinds of propagations suggest a simple propagation theory. First, Kirchoff's voltage law can be applied to every new voltage discovered in the circuit. Then for every node and component in the circuit, Kirchoff's current law can be applied. Finally, for every component which has a newly discovered current into it or voltage across it, its VIC is studied to determine further propagations. If this produces any new voltages or currents, the procedure is repeated.

The current through a capacitor is always zero, so the current contribution of a capacitor terminal to a node can always be determined.

```
(CURRENT C (CAPACITOR C) (C))
```

Similarly, the voltage across a closed switch is zero.

(VOLTAGE (N1 N2) (SWITCH VR) (VR))

The remaining components are semiconductor devices and these are very different from those previously discussed. Although the VIC's for transistors, diodes and zener diodes can be modeled by one nonlinear equation, these devices are usually thought of as having a number of distinct regions of operation, each region having a simple linear VIC. The region of operation must be determined before any VIC can be used.

The diode is the simplest kind of semiconductor device. The only thing we can say about it in our simple propagation theory is that if it is back biased, the current through it must be zero.

(CURRENT D (DIODEV) (D))

For the zener diode we can propagate more values. If the current through a zener diode is greater than some threshold, the voltage across it must be at its breakdown voltage.

(VOLTAGE Z (ZENERI) (Z))

If the voltage across a zener diode is less than its breakdown voltage, the current through it must be zero.

(CURRENT Z (ZENERV) (Z))

The transistor is the most difficult of all devices to deal with. This is both because it has the peculiar discontinuous characteristics of a semiconductor device and because it is a three-terminal device. If the current through any of the transistor's terminals is known, the current through the other terminals can be determined using the beta characteristics of the device (except in the case in which it is saturated). Furthermore, if the voltage across the base-emitter junction is less than some threshold (.55 volts for silicon transistors), the current flowing through any of its terminals should be zero also.

(CURRENT C/Q1 (BETA Q1 B/Q1) (Q1))

(CURRENT C/Q1 (TRANOFF Q1) (Q1))

Having experts for each component type as has been just described makes it possible to propagate measurements throughout the circuit. As an example, consider the following circuit fragment:
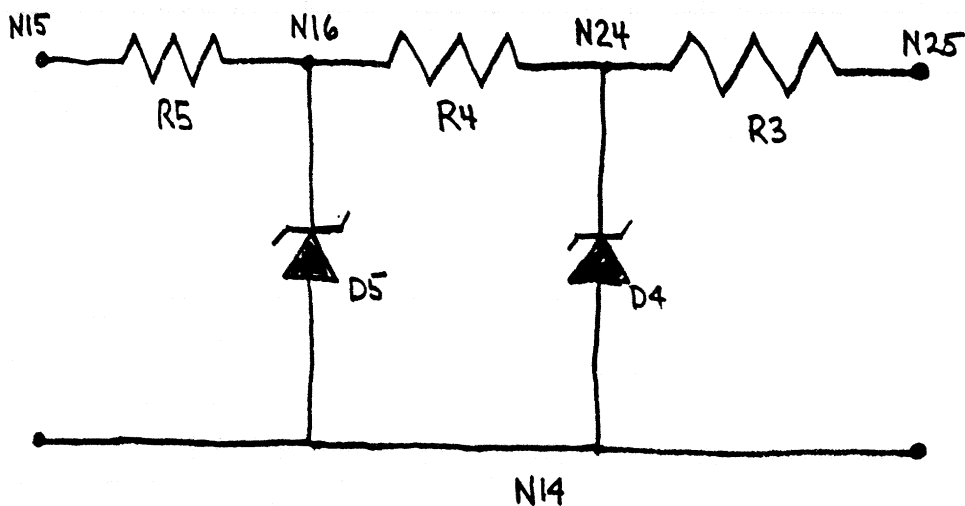
Assume that the fault in this circuit is that D4 has a breakdown voltage too low. This causes the voltage across D5 to be less than its breakdown. Assume the following measurements are made:

        (VOLTAGE (N15 N14))

        (VOLTAGE (N16 N14))

propagations:

        (VOLTAGE (N16 N15) (KVL N16 N14 N15) NIL)

        (CURRENT R5   (RESISTORV R5) (R5))

        (CURRENT D5   (ZENERV D5) (D5))

            the voltage across the zener is less than its breakdown

        (CURRENT R4   (KCL N16) (R5 D5))

        (VOLTAGE (N24 N16) (RESISTORI R4) (R4 R5 D5))

        (VOLTAGE (N24 N14) (KVL N24 N16 N14) (R4 R5 D5))

        (VOLTAGE (N24 N15) (KVL N24 N16 N15) (R4 R5 D5))

        (CURRENT D4   (ZENERV D4) (D4 R4 R5 D5))

            the voltage across the zener is less than its breakdown.

        (CURRENT R3   (KCL N24) (D4 R4 R5 D5))

        (VOLTAGE (N24 N25) (RESISTORI R3) (R3 D4 R4 R5 D5))

```
(VOLTAGE (N25 N14) (KVL N25 N24 N14) (R3 D4 R4 R5 D5))
(VOLTAGE (N25 N16) (KVL N25 N24 N16) (R3 D4 R4 R5 D5))
(VOLTAGE (N25 N15) (KVL N25 N24 N15) (R3 D4 R4 R5 D5))
```

The propagation proceeds one deduction at a time; never is it necessary to make two simultaneous assumptions in order to get the next step in the propagation chain, since the propagation can always go through some intermediate step.

A SIMPLE THEORY OF TROUBLESHOOTING

This section examines how the propagation strategy of the previous section can be used to troubleshoot the circuit. The ideas of contradictions and corroborations between propagations will be used to show how the propagator can be used to help in troubleshooting the circuit. In this simple theory we will assume that coincidences occur only between propagated values and actual measurements.

The meaning of the coincidences depends critically on the kinds of assumptions that the propagator makes. For the coincidences to be of interest every assumption made in the derivation must be mentioned, and a violation of any assumption about a component must mean that component is faulted. Then, when a contradiction occurs, one of the components of the derivation must be faulted. Furthermore, if the coincidence was a corroboration, all the components about which assumptions were made are probably unfaulted.

The usefulness of the coincidence depends critically on how many faults the circuit contains. The usual case is that there is only one fault in the circuit. Even the case where there is more than one fault in the circuit, the approach of initially assuming only a single fault in the circuit is probably a good one.

If there is only one fault in the circuit, all the components not mentioned in the derivation of the contradiction, must be unfaulted. If a corroboration occurs, all the components used in the derivation can be assumed to be unfaulted. In a multiple fault situation these would be invalid deductions: in a contradiction only one of the faulted components need be involved and in a corroboration, two faults could cancel out each other to produce a correct final value.

If, in the propagation example of the previous section, the voltage between N25 and N14 was discovered to contradict with the propagated value, one of R3, D4, R4, R5 and D5 must be faulted. But, if the values were in corroboration, all the components would have been determined to be unfaulted.

Now that the fault has been reduced to one of R3, D4, R4, R5 and D5, the propagations can be used to determine what measurement should be taken next. The best sequence of measurements to undertake is, of course, the one which will find the faulted component in the fewest number of new measurements. Assuming that the relative probability of which component is faulted is not known, the best strategy is a binary search. This is done by examining all propagations in the circuit, eliminating from their assumption lists components already determined to be correct, and picking a measurement to coincide with that propagation whose number of assumptions is nearest to half the number of possibly faulted components.

In the example there are five possibly faulted components, hence the best propagations to choose, are those with two or three assumptions. That means either measuring the current through R4, voltage across D4, the voltage across R4 or the voltage between N24 and N15.

```
(CURRENT R4        (KCL N16) (R5 D5))

(VOLTAGE (N24 N16) (RESISTORI R4) (R4 R5 D5))

(VOLTAGE (N24 N14) (KVL N24 N16 N14) (R4 R5 D5))

(VOLTAGE (N24 N15) (KVL N24 N16 N15) (R4 R5 D5))
```

All the other measurements, in the worst case, can eliminate only one of the possibly faulted components from consideration.

The current through R4 is measured. This coincidence is a corroboration; so R5 and D5 are verified to be correct. Therefore one of R3, D4 and R4 must be faulted. This leaves the following interesting propagations.

```
(VOLTAGE (N24 N16) (RESISTORI R4) (R4))

(VOLTAGE (N24 N14) (KVL N24 N16 N14) (R4))

(VOLTAGE (N24 N15) (KVL N24 N16 N15) (R4))

(CURRENT D4        (ZENERV D4) (D4 R4))
```

(CURRENT R3   (KCL N24) (D4 R4))

At this point there are too few possible faults to make a binary search necessary. Any measurement which would coincide with any propagation having R3, D4 or R4 as assumptions, but not all three at once, is a good one. One such measurement is the current through D4. In the actual circuit D4 has its breakdown voltage too low so it is drawing a great deal of current. The propagator deduced the current should be zero. This contradiction would indicate that R3 was verified since it was not involved. Two possible faults remain; R4 and D4. R4 could be faulted high. D4 could be faulted low. Measuring anyone of the following will indicate that D4 is faulted:

(VOLTAGE (N24 N16) (RESISTORI R4) (R4))

(VOLTAGE (N24 N14) (KVL N24 N16 N14) (R4))

(VOLTAGE (N24 N15) (KVL N24 N16 N15) (R4))

UNEXPECTED COMPLEXITIES OF THE SIMPLE THEORY

The discussion of the previous section presents an interesting and, on the surface, very simple scheme for troubleshooting. Unfortunately, the entire approach is fraught with difficult problems! This section deals with some of these problems and attempts to provide a solution to them within the original framework. Such an investigation will clarify the deficiencies of using only local circuit knowledge for troubleshooting.

Basically, three kinds of problems arise. First, the handling of corroborations and contradictions leads to faulty assertions in certain situations and thus must be examined much more closely. Second, it will be shown that the propagation scheme, the knowledge contained in the experts, and the troubleshooting strategy are all incomplete. Each of them cannot make certain kinds of deductions which one might expect of them in the framework that has been outlined. Finally, accuracy is a problem; all components and measurements have an error associated with them (if only a truncation or roundoff error), and these cause many kinds of difficulties.

The nature of corroborations requires closer scrutiny. It has already been shown that every component on which a derivation depends is in the assumption list of that derivation, so a contradiction localizes the faulted component to one of those mentioned in the assumption list. For

corroborations, the simple troubleshooting scheme used the principle that a coincidence indicated that all of the components in the assumption list were cleared from suspicion. This principle must be studied with much greater scrutiny, as there are a number of cases for which it doesn't hold.

In order to do this we must examine the precise nature of the propagations, and, more importantly, examine the relation between a single value used in a propagation with the final propagated value. Consider a propagated value derived from studying the component D. Let the resulting current or voltage value be $f(D)$. The propagator is entirely linear; so the propagated value at any point can be written as a linear expression of sums of products involving measured and propagated values. For every component, current and voltage vary directly with each other and not inversely. Hence, in the expression for the final propagated value, $f(D)$ can never appear in the denominator. So the final value can be written as:

$$value = f(D)\ a + b$$

Where a and b are arbitrary expressions not involving D. The relation between $f(D)$ and the final propagated *value* is characterized by a. By studying the nature of component experts, the structure of a can be determined. Every expert derives $f(D)$ either by multiplying the incoming value $v(D)$ by a parameter, or by applying a simple comparison test to the $v(D)$. As many such comparison tests can be involved in a single propagation, each propagation can have a predicate associated with it indicating what conditions must be true for the propagation to hold. With both kinds of propagations there is a problem if a is zero. In that case, $f(D)$ has no influence on the final value and so a coincidence says nothing about the validity of $f(D)$.

A corroboration with a propagation involving a predicate only indicates that the incoming value $v(D)$ of the predicate lies within the tested range, thus saying little about the assumptions which were used to derive $v(D)$. Note, however, that in a contradiction the predicate may be testing an erroneous value, and thus $v(D)$ might be incorrect. We shall call these assumptions, which corroborations do not remove from suspicion, the *secondary* assumptions of the propagation, and the remaining, the *primary* assumptions.

The situation for which a is zero can be partially characterized. Using the same assumption more than once in a propagation is relatively rare. In such a single-assumption propagation a must

be a single term, consisting of a product of parameters (resistances, betas, etc.) or their inverses, and since no circuit parameter is zero, a cannot be zero.

If multiple assumptions about D are made in a single propagation a may become a sum, and hence possibly zero, so another argument must be used. Every occurrence of an assumption about D in a propagation possibly introduces another term to a. Each of these terms must itself be a product of parameters. Unfortunately, we cannot prove that $a=0$ is impossible, but can only appeal to a somewhat heuristic argument. Consider the case where a is zero. By the previous argument a is only a function of circuit parameters and so is independent of any measurements. That means whatever value $f(D)$ has, or even whatever value is actually measured; that value, no matter how extreme, has absolutely no influence in our propagation scheme on the final propagated value. That seems absurd, so a must never be zero. In other words, a specifies the degree of coupling between two values in the circuit and it seems impossible that two values in the circuit are completely decoupled. In the case where a is small but not zero (i.e. weak coupling) accuracy issues become critical, but these will be discussed later.

The propagation scheme cannot make all the propagations that one might reasonably expect. Incompleteness of this type manifests itself in two ways. One is just a problem of circuit representation, and the other is an inherent problem of the propagator. In both certain obvious propagations are not made.

Kirchoff's current law can apply to collections of components and nodes, not just single components and nodes. Recognizing relevant cutsets in the topology of the circuit is a tedious (yet performable) task. Circuit diagrams usually present a visual organization so that such cutsets (and teleological organization) become clear.

The process of propagation as outlined consists of using a newly discovered value to call an expert which can use that value to make new discoveries. The expert then looks at the environment, and from this deduces new values for the component about which it is an expert. The communication with the environment always involves numeric values. Experts cannot communicate with each other, nor can they handle abstract quantities. Furthermore, propagation

stops when a coincidence occurs and iteration toward an accurate solution is never attempted.

This entire scheme is motivated by what we see in human troubleshooters, yet the strategy has some very surprising limitations. The fact that only one expert is invoked at any one time means that only one assumption can be made at any step in the propagation process. This means that propagations which require two simultaneous assumptions cannot be made. Most propagations which require more than one assumption do not require simultaneous assumptions since they can be derived using some intermediate propagation (e.g. all the previously discussed examples).

One such case requiring simultaneous assumptions is the voltage divider.
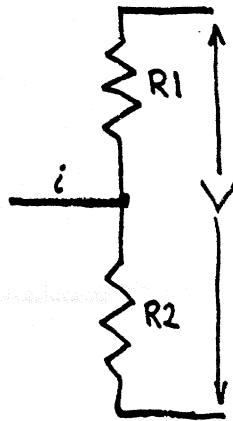
Supposing V and i are known, the current through R1 (and hence through R2) can be propagated by simultaneously assuming the correctness of both R1 and R2.

$$V = i_1 \; R1 + i_2 \; R2$$

$$i = i_1 - i_2$$

$$i_1 = (V - i \; R2)/(R1 + R2)$$

Admittedly, the voltage divider is an important enough entity that it should be handled as a special case pattern, but this kind of incompleteness will arise in other situations, and it will not be possible to design a special case pattern for each of them.

If multiple faults are allowed, simultaneous assumptions must be handled with even greater caution. For example, a propagation involving a simultaneous assumption can propagate a correct

value even though both components involved in the assumptions were faulted. In the case of a voltage divider, the resistance of both R1 and R2 could shift without affecting the voltage at the tap, yet the voltage divider would present an erroneous load to the voltage source to which it was connected.

Due to this inherent incompleteness in the propagator, coincidences can also occur between propagated values. This is much more complicated than the coincidences we have been considering since both propagations have assumptions that have to be examined. If one of the propagations has no unverified assumptions, the coincidence can be handled as if it were between a propagated value and an actual measurement. However, if both propagations have unverified assumptions the coincidence becomes far more difficult to analyze. The effects of such coincidences depend critically on whether the intersection of the unverified assumptions in each propagation is empty or not. If the intersections is empty, a contradiction reduces the list of possible faults to the union of the assumptions used in the propagations, and a corroboration indicates that the value in question is the correct one, and can be treated as two separate corroborations between propagated and measured values.

The case of a nonempty intersection is the most difficult. If the coincidence was a corroboration, a fault in the intersection could have caused both propagations to be incorrect yet corroborating. Even so, something can be said about the disjoint assumptions in the propagations, since if there was a fault in one of the disjoint primary assumptions it must have caused a contradiction; thus all the disjoint primary assumptions can be verified to be correct. If the coincidence was a contradiction, the list of possibly faulty components can be reduced to the union of the assumptions. In this case it is very tempting to remove from suspicion all those components mentioned in the intersection, because this would capture the notion that correct propagations from a single (albeit incorrect) value must always corroborate each other or, equivalently, that each point in the circuit has only two values associated with it: a correct value and a faulted value (which is predicted by the propagator).

Unfortunately that analysis is not valid. Consider a feed-back loop. A faulted value is propagated into this feed-back loop, the feed-back loop propagates a value completely around the

loop and contradicts with the value we entered the loop with. Either the feed-back loop is faulted, or the initial value we entered the loop with was incorrect, thus by the nature of feed-back giving a contradiction when that value was propagated completely around the loop. (Not every feed-back loop exhibits this property, however, although it is easy enough to construct one that does.)

All measurements in the circuit and all circuit parameters have errors associated with them. Even if perfect measurements are assumed, truncation and roundoff errors would still cause problems. One way to view the problem is to study the size of a relative to the error in b. If a is smaller than the error in b, a large error in some $f(D)$ could be undetected. Again we see the greatest problem lies with corroborations. In a corroborating coincidence we must make absolutely sure that an error in any of the verified assumptions could have been detected in the *value* (i.e., a is not too small).

There is a simple partial solution that works in most cases. Instead of propagating numeric values through the circuit, we propagate values and their tolerances, or just ranges of values. Each measurement and circuit parameter could have a tolerance associated with it, and the arithmetic operations could be modified to handle ranges instead of numeric values. Instead of computing a and its tolerance, the propagator could note whenever an error in some incoming value could be obscured in larger errors in other values. This is required since errors in parameters and measurements are usually percentages, and thus adding a large value and a small value will often obscure an error in the small value. Since such problems occur only with addition and subtraction of ranges, KVL and KCL are the only experts which need to be directly concerned with the accuracy issue.

Assuming that errors in values are roughly proportional to their magnitude, those propagations involved in a sum whose magnitude is less than the error in the final result should not be verified in a corroboration of the final value. (As this assumption is not always true, some assumptions may not be verified in a corroboration when they should be.) KVL and KCL can easily check for such propagations. Fortunately, a category for assumptions which should not be verified in a corroboration has already been defined: the secondary assumptions. So, primary

assumptions of the incoming values into a Kirchoff law expert may become secondary assumptions of the final result.

As usual, this theory of handling accuracy has subtle problems. If the only possible effect of a particular $f(D)$ was described in a propagation, then no matter how insignificant its contribution was to the final value, a coincidence should verify D since it wouldn't matter in such a case if D were faulted or not. Furthermore, the propagation through certain components is so discontinuous that no matter how insignificant its propagatory contribution is, a fault in the final value would so greatly affect the propagation that the assumption in question should really be treated as a major assumption. An example of the former is a switch in series with a resistor, and an example of the latter is a zener diode contributing zero current to a node.

Consider the case of a resistor in series with a switch. The only contribution of that switch to the circuit is in the voltage across the switch and the resistor. A voltage across a closed switch is zero; so unless the resistance of the resistor is zero, the switch becomes a secondary assumption of the final voltage. Unfortunately, a corroboration with that voltage should indicate the switch was acting correctly.

Similarly, a zener diode contributing zero current to a node will always become a secondary assumption of the KCL propagation. But, a corroboration should indicate that zener was functioning correctly. That is because this propagation would not even have been possible if the voltage across the zener was near its breakdown. A heuristic solution to this problem is not to secondarize propagations with zero value which were just propagated from discontinuous devices. This, of course, makes the teleological assumption that the discontinuous component makes a significant contribution whenever it is contributing a non-zero value, as is almost always the case with the switch, diode, zener diode and transistor.

Accuracy brings along other problems, as testing for equality between ranges becomes a rather useless concept. A simple workable strategy is to use a rough approximation measure such as accepting two ranges as equal if the corresponding endpoints of the two ranges are within a certain percentage of each other. More satisfactorily, the actual width of the range should also enter into consideration so that if one end of the range is extremely small relative to the other, a much more

liberal percentage is used to compare the smaller endpoints. One certainly would want the range [0 , 1] to be roughly equal to [1E-6 , 1]. A coincidence can thus be of three kinds; either the ranges can be approximately equal (or just significantly overlapping), which is a corroboration, or the ranges can be disjoint, which is a contradiction, or the ranges can overlap but not significantly, which provides no information at all.

The following simple algorithm implements these ideas. A tolerance for the comparison is computed by choosing the minimum width if the widths are very different and choosing half the width if the widths are approximately the same. Depending on the circuit and whether the coincidence is between voltages or currents a minimum tolerance is specified. The minimum tolerance for a typical circuit is .1 microamperes and .1 volts. Then the differences between the corresponding ends of the ranges are determined. If both differ within the tolerance, the values are determined to be corroboratory. For example, [.1 , .2] volts and [.15 , .3] volts are judged to be corroboratory. If only one side is within tolerance the tolerance is relaxed by 50% and the failing side is checked again. If this still does not match, we cannot really claim a corroboration; instead we can only say that one value *splits* the other. For example, [0 , 1] splits [0 , 10]. The two remaining cases occur when the values are completely disjoint (e.g. [0 , 1] and [3 , 4]) and when they contain each other (e.g. [0 , 6] and [3 , 4]). The containment case is treated as a split. Ranges are considered disjoint only if the they differ by greater than the tolerance. If none of these conditions are met, the coincidence is neither a corroboration nor a contradiction. For example, [0 , .1] volts and [.2 , .3] neither contradict nor corroborate. This algorithm is only a simple attempt at defining equivalence of ranges, and some of the parameters may have to be tuned for specific circuits.

A comparison test between two ranges can have five results: (1) values contradict, (2) values corroborate, (3) first value splits second, (4) second value splits first, and (5) no comparison possible. The last alternative raises the possibility that it may be useful to propagate two independent values for the same quantity! The splitting possibilities can be intelligently dealt with. If the value for A splits the value for B, then if A is valid, B must be valid, but not conversely. For example, since A:[3 , 4] splits B:[0 , 10], the validity of A implies the validity of B. But if B were valid, A might be [7 , 8] which still splits B but contradicts with the original [3 ,4]. If A is not known to be valid, we

must wait till it is proven before using this information. However, in a single fault theory a very interesting deduction can still be made. It is easier to see in formal terms: A splitting B really says $valid(A) \supset valid(B)$, while A corroborating B says $valid(A) \equiv valid(B)$. Consider $valid(A) \supset valid(B)$. If the assumptions of A and B are not disjoint, construct a $B*$ that does not mention the common assumptions. Now $valid(A) \supset valid(B*)$ also implies $invalid(B*) \supset invalid(A)$. But the assumptions of $B*$ and A are disjoint and the circuit can have only one fault. Hence $B*$ must be perfectly correct. In summary, the split of B by A in a single fault theory implies all the assumptions involved with B are correct (i.e. a corroboration of B with truth) and nothing about the assumptions of A. This corresponds with our intuition; a split is a kind of corroboration in which one of the propagations is much stronger than the other, and as such the corroboration only comments on the weaker of the two propagations.

Although the range mechanism was introduced to handle errors in measurements and component parameters, it can also be used to deal with new kinds of propagations that would have been impossible in the simple scheme. Noticing that the collector current of a transistor is large leads to the deduction that its base-emitter voltage must be between .5 and 1 volt. With the range mechanism this kind of propagation can now be included: propagate the range [.5 , 1]. There are many possible uses for this idea. Every diode could propagate a non-negative current through itself. Every transistor could propagate a base-emitter voltage of less than 1 volt. The voltage at every node could be asserted to be less than the sum of the voltage sources in the circuit. More interestingly, it could handle the problem of having a range propagated over a discontinuous device: a [-1 , +1] current range propagated into a diode should have its lower limit modified to 0 (i.e. [0 , +1]).

When a significant propagation occurs which overlaps a test point of a discontinuous component, the best strategy is to interpret that measurement to have too wide an error associated with it and stop the propagation there. In general, when error tolerances in propagated values become absurd (a significant fraction or multiple of the central value) the propagation should be artificially stopped.

When a coincidence occurred in the old propagation scheme the propagations stopped. There was no advantage in also propagating the new value. However, when ranges are involved, the new propagation might be better than the old one. The range with the smallest error is the better of the two. For example, the values [0, 10] corroborates with [1, 2], yet the latter value would provide much more information if it were propagated. This means that when a coincidence between ranges occurs, the better of the two propagations must not be stopped from propagating.

There remain certain characteristics of the devices that are not captured in the propagation scheme. These are the maximum ratings of the components. The power dissipation of a transistor cannot exceed its power rating, the voltage across a capacitor cannot exceed its breakdown voltage, the power dissipation in a resistor cannot exceed its wattage rating, etc. To a large extent these can be captured by simple modifications of the component experts. Each expert could check whenever it was invoked whether any ratings about the component were exceeded. If the component expert detects that a rating has been exceeded it must treat it as a contradiction. The maximum rating, of course, depends only on the component itself.

A contradiction casts suspicion on all the assumptions of the contradicting propagations. More careful examination of the contradiction may restrict the possible faults even further. Knowing that the current in a resistor is higher than expected indicates that its resistance has shifted downwards. If a contradiction suggests there is too little current through a capacitor, we know the capacitor cannot be contributing to the fault.

We must tackle the problem of how to scan back through the propagation to determine what faults in the components could have caused the final contradiction. Of course, a straightforward way to do this would be to compute a for every component $f(D)$ involved in the propagation. For every two-terminal component the possible fault can be immediately determined from a (unless of course we have the inaccurate case where the range for a includes zero). The only three-terminal device, the transistor, requires a more careful examination as it has many possible fault modes, and a single consideration of a propagation from it may not uniquely determine its fault mode.

Continuing in the spirit of the original propagation scheme, a method different from that of computing a should be used. The following simple scheme has difficulties only in certain kinds of multiple assumption propagations. The contradiction indicated that the propagation was in error by a shift in value in a certain direction. This shift can be propagated backwards through all the experts except KCL and KVL. The Kirchoffs' laws experts involve addition, so each of the original contributors to the sum must be examined. For those contributors whose (unverified) assumption list does not intersect with any of the other assumption lists, the shift can be propagated back, after adding the appropriate shift caused by the remaining contributors. For those contributors with intersecting contributions, it must be determined for each of the intersecting components whether all contributions of all the possible faults do not act against each other (e.g. will a shift in the resistance of the component both increase a current contribution to a node and decrease it through another path?). For such canceling intersections, nothing can be said about the intersecting component. All this does is capture qualitatively whether the signs of the terms of a are different and thus canceling. It should be noted, that if it really turns out to be the case that a a can be zero, such a scheme could be used at least to eliminate faulty verifications from taking place, again at the cost of sometimes not verifying provably unfaulted components.

Incompleteness in the propagation scheme introduces incompleteness in the troubleshooting scheme. Even if the propagation scheme were complete the troubleshooting scheme would be incomplete, since the earlier answer to what is the next best measurement is inaccurate. The measurement which reduces the list of possible faults by the greatest number is not necessarily the best measurement. Future measurements must also be taken into consideration, a poor first measurement may set the stage for an exceptionally good second measurement.

The choice of best measurement depends of course on what is currently known about the circuit. The most general approach would be to try every possible sequence of hypothetical measurements and choose the first measurement of the best sequence as the next measurement. Again, that would be an incredible, and unnatural computation task. The current troubleshooting scheme does not try to generate all possible sequences, but only considers making those

measurements about which it already knows something (so to produce a coincidence).

Since only measurements at points about which something is explicitly known are considered, the information provided by coincidences between solely propagated values (the result of incompleteness in the propagator) cannot enter into consideration. Thus the basic approach of the troubleshooter is to make no hypothetical measurements and look only at those propagations with unverified assumptions as predictions to try to coincide with. Unexpected information, such as that provided by coincidences between propagated values, cannot be considered in that paradigm (although making hypothetical measurements would handle this problem).

If we are only prepared to look ahead one measurement, our original search scheme remains reasonable. The binary search for the best measurement must, of course, be reorganized. Since a corroboration may eliminate different numbers of components from suspicion than a contradiction, the search is not purely binary. A workable solution is to just take the average of the number of components which would be verified in each case as the measurement's score. Then that measurement whose score was nearest to half the number of faulted components could be chosen as the next measurement.

There remains the issue of generating an explanation for this choice. Although the above argument for deriving a future choice of measurement could be made understandable to humans it does not always admit a very good explanation. A large part of the explanation for a future choice of measurement involves indicating why a certain component cannot be faulted. Once a component is eliminated from suspicion for any reason it is never considered again. However, a later measurement might give a considerably better explanation for its non-faultiness. The problem of generating good explanations, of course, also must take into account a model of the student and what he knows about the electronics and the particular circuit in question.

The above scheme for selecting measurements does not take into account how "close" the measurement is to the actual components in question. For example, a voltage measurement across two unverified resistors is just as good as a measurement many nodes away which also has only those two resistors as unverified assumptions. Fortunately these can be easily detected: just remove from the list of possible measurements all those which are propagated from other elements on the

list. These are the propagations which make no new assumption in their most recent propagation step and involve only one unverified propagation. For example in the first troubleshooting scenario the measuring the voltage between N15 and N24 was a candidate. Since KVL makes no assumptions and the other voltage between N15 and N16 had been already verified this suggestion should have been thrown out.

SOME ILLUSTRATIVE EXAMPLES

The following are some debugging scenarios to illustrate the ideas of the previous section. Note that primary and secondary assumption lists are kept for each propagation.

The case of R11 being high:
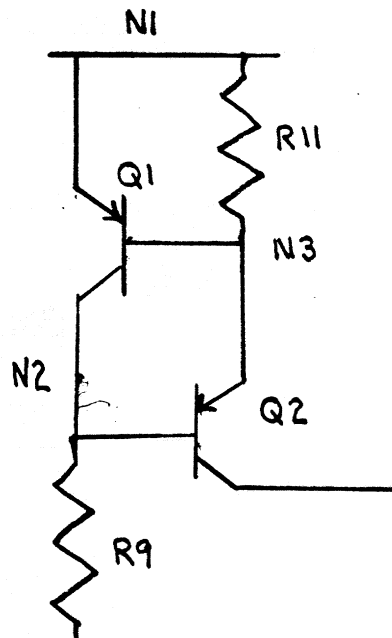
(= (CURRENT C/Q2 (MEAS M0004) NIL NIL) [.00017 , .00019])

(= (CURRENT B/Q2 (BETA Q2 C/Q2) (Q2) NIL) [1.1E-6 , 3.8E-6])

(= (CURRENT E/Q2 (BETA Q2 C/Q2) (Q2) NIL) [-.00019 , -.00017])

(= (VOLTAGE (N2 GROUND) (MEAS M0005) NIL NIL) [45 , 49])

(= (CURRENT R9) (RESISTORV R9) (R9)) NIL) [.012 , .017])

(= (CURRENT C/Q1 (KCL N2) (R9) (Q2)) [.012 , .017])

(= (CURRENT B/Q1 (BETA Q1 C/Q1) (Q1 R9) (Q2)) [8.1E-5 , 33E-5])

(= (CURRENT E/Q1 (BETA Q1 C/Q1) (Q1 R9) (Q2)) [-.017 , -.012])

(= (CURRENT R11   (KCL N3) (Q1 R9) (Q2)) [-.00015 , .00011])

(= (VOLTAGE (N1 N3) (RESISTORI R11) (Q1 R9 R11) (Q2)) [-.26 , .18])

(= (CURRENT C/Q1 (TRANOFF Q1) (R11 Q1 R9) (Q2)) [-1.E-6 , 4.0E-5])

A contradiction occurs. The new propagation is "better" than the old one. The old propagation cannot not be removed in favor of the new propagation because it is an antecedent of the new propagation. We conclude that one of R11, Q1, R9 or Q2 must be faulted.

Consider the problem of R9 being open:

(= (CURRENT C/Q2 (MEAS M0001) NIL NIL) [.00033 , .00036])

(= (CURRENT B/Q2 (BETA Q2 C/Q2) (Q2) NIL) [2.2E-6 , 7.2E-6])

(= (CURRENT E/Q2 (BETA Q2 C/Q2) (Q2) NIL) [-.00037 , -.00033])

(= (VOLTAGE (N2 GROUND) (MEAS M0002) NIL NIL) [44 , 49])

(= (CURRENT R9 (RESISTORV R9) (R9) NIL) [.012 , .016])

(= (CURRENT C/Q1 (KCL N2) (R9) (Q2)) [.012 , .016])

(= (CURRENT B/Q1 (BETA Q1 C/Q1) (Q1 R9) (Q2)) [8E-5 , .00033])

(= (CURRENT E/Q1 (BETA Q1 C/Q1) (Q1 R9) (Q2)) [-.017 , -.012])

(= (CURRENT R11 (KCL N3) (Q1 R9) (Q2)) [2.6E-6 , .0003])

(= (VOLTAGE (N1 N3) (RESISTORI R11) (R11 Q1 R9) (Q2))

   [.0036 , .475])

(= (CURRENT C/Q1 (TRANOFF Q1) (R11 Q1 R9) (Q2)) [-1.E-6 , 4.E-5])

This contradiction indicates that one of R11, Q1, R9 or Q1 is faulted.

In this example the circuit has no faults.
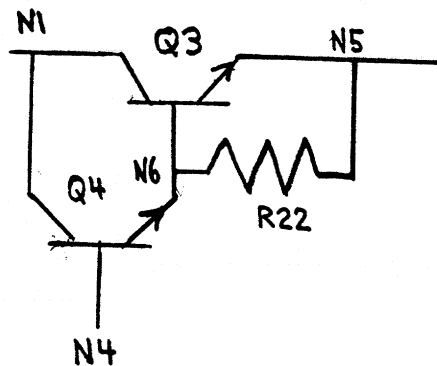
(= (CURRENT B/Q4 (MEAS M001) NIL NIL) [-.00036 , -.00032])

(= (CURRENT E/Q4 (BETA Q4 B/Q4) (Q4) NIL) [.016 , .05])

(= (CURRENT C/Q4 (BETA Q4 B/Q4) (Q4) NIL) [-.05 , .016])

(= (VOLTAGE (N6 N5) (MEAS M002) NIL NIL) [.85 , .93])

(= (CURRENT R22 (RESISTORV R22) (R22) NIL) [.0015 , .0020])

(= (CURRENT B/Q3 (KCL N6) (Q4) (R22)) [-.052 , -.014])

(= (CURRENT E/Q3 (BETA Q3 B/Q3) (Q3 Q4) (R22)) [.16 , 1.6])

(= (CURRENT C/Q3 (BETA Q3 B/Q3) (Q3 Q4) (R22)) [-1.6 , -.14])

(= (CURRENT E/Q3 (MEAS M003) NIL NIL) [.64 , .71])

This split of [.16 , 1.6] by [.64 , .71] indicates that Q3 and Q4 must be unfaulted.

Closer examination of the above examples reveals that more information about the faultiness of the components could have been deduced earlier. The current theory embodies only a small amount of the different reasoning strategies the student might have available. This is the subject of the subsequent sections.

THE NECESSITY AND UTILITY OF OTHER KNOWLEDGE

In this section we will attempt to characterize where and why local and nonteleological reasoning fails. Many such failures have already been demonstrated in the previous sections. Our

method of attack will be from two directions. First, problems inherent in the earlier propagation scheme can be alleviated with other knowledge about the circuit. Second, many of the kinds of troubleshooting strategies we see in humans cannot be captured even by a generalization of the proposed scheme. One of the basic issues is that of teleology. The more teleological information one has about the circuit, the more different the troubleshooting process becomes. Currently, most of the ideas presented in this paper so far have been implemented in a program so that much of the discussions derive their observations from actual interactions with the program.

The most arresting observation is that the propagator cannot propagate values very far, and at other times it propagates values beyond the point of absurdity. Examining those propagations which go too far the most dominant characteristic is that either the value itself has too high of an error associated with it, or that the propagation itself is not relevant to the issues in question. The former problem can be more easily answered by more stringent controls on the errors in propagations. The latter requires an idea of localization of interaction. This idea of a theater of interactions would limit senseless propagation; however, it requires a more hierarchical description of the circuit.

The idea that every measurement must have a purpose points out the basic problem: our troubleshooter cannot make intelligent measurements until it has, by accident, limited the number of possible faults to a small subset of all the components in the circuit. After this discovery has been made, which the troubleshooter is not given and must make by itself, fairly intelligent suggestions, can be made. However, as such a discovery is usually made when the set of possible faults is reduced to about five components, it can only intelligently troubleshoot in the last few (two or three) measurements that are made in the circuit.

Clearly, many measurements are made before this discovery and the troubleshooter cannot do anything intelligent during this period. Still, the propagation scheme and the ideas of corroborations and contradictions can be effectively used even during this period.

The only way intelligent measurements can be made during this period is by knowing something about how the circuit should be behaving. This requires teleological information about the circuit. For example, just to know that the circuit is faulted and requires troubleshooting

requires teleology. In the situations where the propagator did not propagate very far, the problem usually was that some simple teleological assumption could have been made. The voltages and currents at many points in the circuit remain relatively constant for all instantiations of the circuit, and furthermore many of them can be easily deduced (e.g. knowing certain voltage and current sources such as the power supply, knowing contributions by certain components to be small, etc.). Propagation can then proceed much further. Of course, the handling of coincidences requires modifications, and a new kind of strategy to deal with teleological coincidences needs to be developed.

Coincidences provided information only about the assumptions of the propagations involved. Since the only kind of assumptions we were considering were those about the faultedness of components, the consequences of violating assumptions were obvious. The consequences of violating a teleological assumption is not at all obvious and requires more knowledge about the circuit. The point is that the ability the propagate teleological assumptions is just a small step towards dealing with teleology.

In his thesis Brown <Brown, 76> deals primarily with how to represent and use teleological knowledge in troubleshooting. Although propagation plays only a small role in his theory, many of his ideas address the problems that we have been discussing in this section.

FUTURE RESEARCH

The previous sections have sketched out the necessity for more teleological and non-local knowledge. Since Brown addressed this problem, one obvious direction for research is to try to incorporate his ideas. This direction suffers from two difficulties. First, Brown never implemented his ideas and thus they require a major effort to become actually utilizable. (The troubleshooter based on the ideas of this paper (INTER) is working and requires a practical theory of teleology.) Second, Brown's troubleshooting theory would not be usable in a tutoring context where the expert must be able to understand the student's troubleshooting strategy.

Fortunately, there appears to be a rather simple strategy based on the existing propagator which can be used to deal with non-local knowledge. The idea is based on observations that

students often reason something like: "If the voltage limiter is off and it should be off, then the constant voltage source cannot be contributing to the observed symptom." Note that this argument is not in terms of numerical quantities, but is in terms of states of the components and sections. The component experts can be modified to determine what state the components are in. These observations could then be asserted in a data-base.

This collection of assertions forms a qualitative description of the state of the circuit. Of course, the assertions, like propagations, have their assumptions stored with them. Circuit specific theorems can then be encoded referring to assertions in the description space. The rule of the previous paragraph might be encoded as:

$$(STATE \ voltage\text{-}limiter \ off) \wedge (CORRECT\text{-}STATE \ voltage\text{-}limiter \ off)$$
$$\supset (OK \ constant\text{-}voltage\text{-}source)$$

It appears that only a small number of such theorems are necessary to determine what is known about a circuit from a set of measurements. The theorems are, of course, very circuit specific. Since only a few of them are be required for any specific circuit the principle is still usable.

The local reasoning strategy isolates the qualitative reasoner from worrying about many of the idiosyncrasies of propagating numerical values by describing the circuit in qualitative terms. This is giving us the opportunity to try many different kinds of qualitative reasoning strategies. The failings of the local troubleshooting strategy is also showing exactly where this qualitative reasoning is required.

REFERENCES:

<Brown, 74>
Brown, A.L., "Qualitative Knowledge, Causal Reasoning, and the Localization of Failures -- a Proposal for Research", Artificial Intelligence Laboratory, WP-61, Cambridge: M.I.T., 1974.

<Brown, 76>
Brown, A.L., "Qualitative Knowledge, Causal Reasoning, and the Localization of Failures", Artificial Intelligence Laboratory, forthcoming TR, Cambridge: M.I.T., 1976.

<Brown & Sussman, 74>
Brown, A.L., and G.J. Sussman, "Localization of Failures in Radio Circuits a Study in Causal and Teleological Reasoning", Artificial Intelligence Laboratory, AIM-319, Cambridge: M.I.T., 1974.

<Brown et.al., 74>
Brown, John Seely, Richard R. Burton and Alan G. Bell, *SOPHIE: A Sophisticated Instructional Environment for Teaching Electronic Troubleshooting (An example of AI in CAI)*, Final Report, B.B.N. Report 279, A.I. Report 12, March,1974.

<Stallman & Sussman, 76>
Stallman, R.S., and G.J. Sussman, "Forward Reasoning and Dependency-Directed Backtracking In a System for Computer-Aided Circuit Analysis", Artificial Intelligence Laboratory, AIM-380, Cambridge: M.I.T.,1976.

<Sussman & Stallman, 75>
Sussman, G.J., and R.M. Stallman, "Heuristic Techniques in Computer Aided Circuit Analysis", Artificial Intelligence Laboratory, AIM-328, Cambridge: M.I.T., 1975.