TROUBLE-SHOOTING BY PLAUSIBLE INFERENCE *

Leonard Friedman
Jet Propulsion Laboratory, California Institute of Technology
Pasadena, CA 91103

ABSTRACT

The PI system has been implemented with the ability to reason in both directions. This is combined with truth maintenance, dependency directed backtracking, and time-varying contexts to permit modelling dynamic situations. Credibility is propagated in a semantic network, and the belief transfer factors can be modified by the system, unlike previous systems for inexact reasoning.

I TROUBLE-SHOOTING LOGIC

The PI (for Plausible Inference) system enables a user to trouble-shoot physical systems in a very general way. The trouble-shooting process requires that our user first define a physical model which represents what takes place in normal operation with everything functioning correctly. If this physical model is translatable to the representation used by PI, it can be stored in computer memory and used to guide the search for the most likely failure. In order to make the process clearer, we shall describe a few of the many methods of reasoning employed by human beings in their trouble-shooting. These methods are the ones we can at present imitate in the PI system.

Suppose we have a desired goal state defined in our physical model, and this state depends on three conditions being true to attain the goal. If we execute the process and observe that the goal state was not attained, we conclude that at least one of the three conditions on which it depended must have been false, and all are possibly false. If we then perform some test designed to verify whether one of the conditions is actually true and the test shows that it is indeed true, we conclude that at least one of the remaining two untested conditions must be false. If all but one of the remaining conditions has been eliminated from consideration by further testing, we may conclude that the single condition remaining must be the guilty party. The process of elimination just described is the one normally employed by humans, and it is this process we have implemented on the computer. Of course, the three conditions may in turn have conditions on which they depend. In that case the method just described may be applied recursively to narrow the fault down further, at least to the granularity of the conditions employed in the representation.

This method fails if there are conditions on which the goal state depends for realization and which are not explicitly represented in the model. Nevertheless, the exercise may serve as a valuable guide to help a user to focus attention on specific, possibly false, areas as likely sources of failure. Another difficulty with the method is the fact that either a test does not exist to determine whether a specific sub-goal was reached, or the sub-goal state in question was changed by a later event occurring in the model. In this case it is difficult to verify whether the changed sub-goal state was ever achieved. Only if there were long-lasting side-effects will it be possible to verify. Such difficulties plague human trouble-shooters as well. The present implementation can not reason about such "vanished" events, in a hypothetical mode, from a past context.

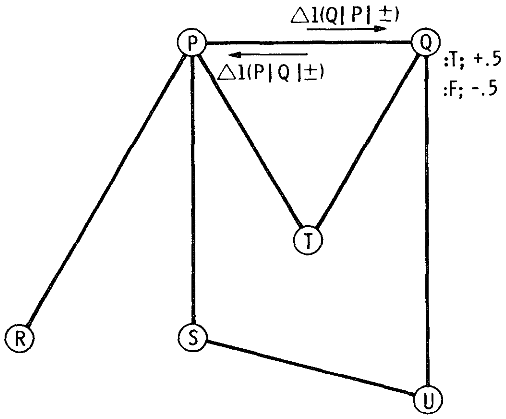
II IMPLEMENTATION AND THEORY

The PI system is part of a larger system called ERIS, named for the Greek goddess of argument. The basic module of the system, described in [1], performs deduction and modeling in the Propositional Calculus. A planning module has been built by M. Creeger by augmenting the basic deduction module with many special features, including a "causal" connective that supplements the standard logical connectives (AND, OR, etc.). Similarly, the PI module has been built by augmenting Propositional Calculus with extra rules of inference, and another belief besides truth-value associated with each assertion. This additional belief, which we call credibility, is a subjective numerical measure of the confidence in the truth-value, with values between -1 and 1.

The basic ERIS module generates a network of nodes linked by connectives as it reads in its knowledge base of assertions; this feature is retained in the other modules. The techniques used in ERIS make it possible to perform deduction without rewriting. Instead, "specialists" for each connective propagate the correct values of the beliefs to the assertions which they link. A theoretical foundation for this approach, applying

*This paper presents the results of one phase of research performed at the Jet Propulsion Laboratory, California Institute of Technology, sponsored by the National Aeronautics and Space Administration under Contract NAS 7-100.

both to Propositional Calculus and First Order Predicate Calculus, is given by the method of Analytic Tableaux" [2]. Because rewriting is totally avoided, inference, planning, model-revision, and dependency-directed backtracking can be performed in a single integrated system.



Arcs link antecedent/consequent in implication
Fig. 1 Transfer Factors in a Semantic Net

Plausible Inference introduces two rules of inference besides Modus Ponens and Modus Tollens, called the Method of Confirmation and the Method of Denial. These extra modes permit the propagation of truth values, true or false, even in directions forbidden by Propositional Calculus. Simultaneously, the associated credibilities are propagated through the net, employing all four modes as appropriate. A calculus of Credibility transfer between arbitrary logical expressions has been worked out to specify exactly the process of Credibility propagation through the net. The calculus is described in [3], and is based on equations employed in MYCIN [4].

The basic quantities controlling propagation between antecedents and consequents in implication are transfer factors or DELTA1´s, and there are four for each antecedent/consequent pair, one for each mode. (See Fig. 1) Both MYCIN and PROSPECTOR are limited to a single reasoning mode and transfer factor for each implication and use a static transfer factor structure that is specified by human "experts". The PI system, in trouble shooting, recalculates the appropriate transfer factors on the basis of incoming evidence. In addition to a dynamic transfer factor structure, PI also incorporates the use of default values for the transfer factors in trouble-shooting when the users do not have better information.

III APPLICATION EXAMPLE

Our example of automated trouble-shooting uses a toy case selected from the application domain, the mission control of spacecraft. It is a simplified representation of signal transmission from a space craft to Earth. The desired goal state is "Signal Received" from the s/c. Fig 2 shows a plan to accomplish this which has been generated by the ERIS planner. The user supplies three basic "action CAUSES state" relations: (1) "Ground Antenna Receiving" CAUSES "Signal Received", (2) "Point s/c" CAUSES "Pointed Correctly", and (3) "Transmit Signal" CAUSES "Signal Transmitted". The pre-conditions for the first relation are the end states achieved in the

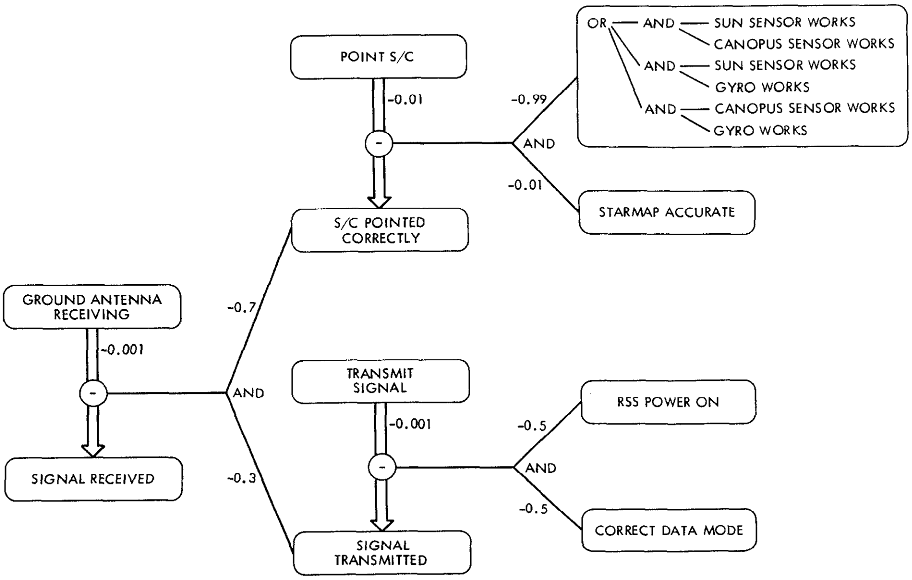


Fig. 2 Plan for Command Sequence Generation

second and third. In effect we are saying, "In order for an action to achieve the goal state certain pre-conditions must be true. In order to make those pre-conditions true certain actions will cause them which may also have pre-conditions to succeed". Thus complex sequences may be built up. The ERIS planner links the basic strategies in the manner shown, using the pre-conditions as hooks. It then collects the actions in a list, and supplies that list as the desired command sequence.

The plan generated in this way is a descriptive model of the signal transmission process and constitutes our trouble-shooting knowledge base. The propagation of beliefs that takes place with the CAUSES connective is identical to the belief propagations of an implication as defined in [3], although the timing of belief propagation of CAUSES are different. We define the belief propagation-equivalent implication form of the CAUSES relation as (IMPLIES (AND action precondition1 .. preconditionN) goal-state). At the start, the assumption is made that all states are true.

Suppose that the sequence is executed and the ground station fails to receive the signal. Then "Signal Received" is false, and this can be entered into the data base. The effects of this change of belief are propagated through the data base by a modified Modus Tollens, making all the events on which "Signal Received" depends Possibly-False (or PF). If a test is then performed by the human controllers like causing the s/c to roll while transmitting, and a signal is received during the roll, we may conclude that the action "Point space craft" worked, "Signal Transmitted" is true, and "Ground Antenna Receiving" has been verified. Inputting these facts into the data base causes the PI system to do two things:

(1) Those preconditions required by "Signal Transmitted" are changed from PF to T by Modus Ponens. (Possibly False to True)

(2) "Pointed Correctly" is changed to False, F, rather than PF. In addition, the PI system raises the credibility of failure for those events on which "Pointed Correctly" depends. Their truth value remains Possibly False because there are multiple possibilities.

If one of these latter events is shown to be true by testing, the remaining one may be the only possibility left. For example, if the Sun Sensor and Canopus Sensor can be shown to work, and their truth status is input, the system will conclude that the Starmap must be at fault, even though the a priori credibility of such a mistake was extremely low.

How the credibilities change at various stages of operation can be described now. At the start there are two possibilities: either a priori credibilities may be entered or default credibilities generated. Figure 2 shows a priori credibilities entered on the branching lines. These are subjective measures of the likelihood of failure of the respective events to which they lead given that the state which depends on them is false. Thus, for "Signal Received", "Pointed Correctly" has an a priori credibility associated with false of -.7, "Signal Transmitted" a value of -.3 and "Point space craft" a value of -.001 .

When we start, the assumption is that every state is true (at the appropriate time) with a credibility of 1.0 (equivalent to certainty). At the next stage, when all we know is that "Signal Received" is false, the a priori credibilities are assigned to all the states. If we used the default mode, credibility would be assigned equally; i. e., for two events, each would get .5, for three .33, for four .25, etc. Whenever an event is eliminated as true, the remaining credibilities are raised by an empirical formula that reflects a reasonable sharing of suspicion, based either on the a priori splits or an equal partition. Thus, in our example, "Starmap Accurate" went from true (cred 1.0) to Possibly False (cred -.007) to Possibly False (cred -.01) to False (cred -1.0).

This, in a simplified way, describes the operation of PI in trouble-shooting using reasoning by Plausible Inference. Of course, humans employ many other methods in trouble shooting, such as analogy. For example, a person may say "This problem resembles one I encountered in another area. Maybe it has the same cause I deduced then." By such techniques, humans can often vector in on a problem, bypassing step-by-step elimination. We hope to implement some of these techniques eventually.

References

[1] Thompson A. M., "Logical Support in a Time-Varying Model", in Lecture Notes in Computer Science, Springer-Verlag, Berlin, 1980. Fifth Conf. on Automated Deduction, July 8-11, Les Arcs, Savoie, France.

[2] Smullyan R. M, First-Order Logic, Springer-Verlag, Berlin, 1968

[3] Friedman L., "Reasoning by Plausible Inference" in Lecture Notes in Computer Science Springer-Verlag, Berlin, 1980. Proc. Fifth Conf. on Automated Deduction, July 8-11, Les Arcs, Savoie, France.

[4] Shortliffe E. H. and Buchanan B. G., "A Model of Inexact Reasoning in Medicine", Math. Biosci. 23, pp. 351-379, 1975. Also chapt. 4 of Computer-Based Medical Consultations: MYCIN, Elsevier, New York, 1976.