Normality and Faults in Logic-Based Diagnosis*

David Poole
Department of Computer Science,
University of British Columbia,
Vancouver, B. C., Canada, V6T1W5
poole@cs.ubc.ca

Abstract

Is there one logical definition of diagnosis? In this paper I argue that the answer to this question is "no". This paper is about the pragmatics of using logic for diagnosis; we show how two popular proposals for using logic for diagnosis, (namely abductive and consistency-based approaches) can be used to solve diagnostic tasks. The cases with only knowledge about how normal components work (any deviation being an error) and where there are fault models (we try to find a covering of the observations) are considered as well as the continuum between. The result is that there are two fundamentally different, but equally powerful diagnostic paradigms. They require different knowledge about the world, and different ways to think about a domain. This result indicates that there may not be an axiomatisation of a domain that is independent of how the knowledge is to be used.

1 Introduction

If someone comes up to you and says that they have used logic to come up with a selection of diagnoses, it is reasonable to ask "exactly what is it that you have proven?". This paper analyses two answers that have been given to that question.

1.1 Representational Methodologies

Many people have argued for using logic in Artificial Intelligence. As with any tool, if we want people to take logic seriously, we have to show how to use it to solve the sort of problems we want to solve.

A theory that says we can model some task using logic is vacuous unless it contains a methodology of how to use the system. As all logic-based representation languages that incorporate definite clauses are equally powerful computationally, any problem that can be specified in one can be specified in another. Thus, when comparing representation systems, we have to compare how problems can be "naturally represented". To do this we need to consider how to go about representing problems in a representation language. Only when we know how to use tools can we compare them in any meaningful way (theoretically or empirically); if we don't use the methods in appropriate ways any comparisons will be vacuous. This paper is about the pragmatics of using diagnostic frameworks.

We also take seriously the *Logic + Control* distinction of Kowalski [1979] (or the epistemic and heuristic components of McCarthy [1969]). In this paper we concentrate on how to "program" the logic, independently of the control.

1.2 Diagnosis

Diagnosis is the problem of trying to find what is wrong with some system based on knowledge about the design/structure of the system, possible malfunctions that can occur in the system and observations (symptoms, evidence) made of the behaviour of the system.

The proposals to formalise the notion of diagnostic reasoning have generally considered two extremes of the diagnosis problem:

1. There is knowledge about how components are structured and work normally. There is no knowledge as to how malfunctions occur and manifest themselves. Diagnosis consists of isolating deviations from normal behaviour. This has normally been the preserve of consistency-based[1] approaches [Genesereth, 1984, de Kleer, 1987].

2. We have just information on faults (diseases) and their symptoms, and want to account for abnormal observations. This has traditionally been

*This research was supported under NSERC grant OG-POO44121. Thanks to Randy Goebel, Greg Provan, George Tsiknis, Alex Kean and Alan Mackworth for many valuable comments and discussions.

[1] This term and "abduction" are used as technical terms defined in section 2.

the preserve of abductive approaches [Popl, 1973, Reggia, 1983, PGA, 1987, Cox, 1987, Poole, 1989].

In this paper, we consider how two logic-based models of diagnosis can use each sort of knowledge. This is important if we want to use them for the continuum of cases between the two extremes. In between these extremes, we have knowledge of how a normal system works as well as knowledge of what faults (diseases, malfunctions) can occur and how they manifest themselves.

2 Two Models of Diagnosis

In this paper we cast two models of diagnosis into the Theorist framework of hypothetical reasoning [PGA, 1987, Poole, 1988a]. This formalism is well suited to the task as both paradigms can be naturally represented in the simple formal framework.

Theorist [Poole, 1988a] is defined as follows. The user provides F, a set of closed formulae[2] (called the *facts*) and H, a set of open formulae (called the *possible hypotheses*). A **scenario** is a set $D \cup F$ where D is a set of instances of elements of H such that $D \cup F$ is consistent. An **explanation** of formula g is a scenario that logically implies g. An **extension** is the set of logical consequences of a maximal (with respect to set inclusion) scenario.

Definition 2.1 (Consistency-Based Diagnosis) A **consistency-based diagnosis** is minimal set of abnormalities such that the observations are consistent with all other components acting normally [Reiter, 1987].

In terms of the Theorist framework,

F is the domain model together with the observations.

H is the set of normality assumptions.

a diagnosis corresponds to an extension (in particular, the set of abnormalities in an extension) [Reiter, 1987, theorem 6.1].

Definition 2.2 (Abductive Diagnosis) An **abductive diagnosis** is a minimal set of assumptions which, with the background knowledge implies the observations [PGA, 1987, Poole, 1989].

In terms of the Theorist framework,

F is the domain model.

H is the set of normality and fault assumptions.

a diagnosis is a minimal explanation of the observations.

The main difference is, in abduction the diagnoses entail the observations, whereas in the consistency based model the observations entail the (disjunct of the) diagnoses. As one would expect the sort of knowledge that has to be specified for each is different.

2.1 What does "normal" mean?

There are two different meanings that have been associated with the notion of normality.

1. A component is normal if it works correctly all of the time [de Kleer, 1987]. We never conclude some component is normal (it may act abnormally in the next observation). Saying a component is abnormal gives no information.
2. We localise our discussion to a particular case. By normal we mean that the correct answer is being produced in this case. We should parameterise normality by the case we are considering.

Example 2.3 Suppose we have a simple digital circuit consisting of two inverters in series, and we put 1 as input and observe 1 as output. If we follow the first definition of abnormal, either inverter could be abnormal. Nothing can be concluded. If we use the second definition of abnormal, it is not the case that one inverter is normal and the other is abnormal; they are either both normal or both abnormal.

The empty diagnosis (with no abnormalities) is a consistency-based diagnosis. If we use the first definition of normal (as in [de Kleer, 1987]) this diagnosis is interpreted as meaning that anything could be abnormal. If we use the second definition of normal, we have to appeal to a sense of parsimony to conclude that both inverters are working correctly [Reiter, 1987]. The advantage of the first interpretation, where any superset of the abnormalities could be abnormal, is that the correct state is covered by the set of diagnoses. In the second case, it could be that both inverters are broken, but we have to assume that they are not broken. Thus the diagnoses are not necessarily covering.

For abductive diagnosis, we have to prove the observations. Suppose we have one inverter, and observed that when a 1 was input, a 1 was output. The first definition is not adequate; saying that the inverter is abnormal does not imply the observation. Abductive diagnosis requires the second definition of abnormal, and requires that we parameterise hypotheses by the case we are considering.

2.2 Consistency-Based Diagnosis

One sort of diagnosis is to have knowledge about how normal components work and discover deviations from normality. This has traditionally been the domain of the consistency-based diagnostic methods [Genesereth, 1984, Reiter, 1987, de Kleer, 1987]. In these models, we axiomatise the normal operating conditions and any deviation is an abnormality.

The following methodology is used [Reiter, 1987]:

1. For each component C that could possibly be faulty we have the hypothesis $\neg ab(C)$.

[2] We assume the underlying logic is the first order predicate calculus. We follow the Prolog convention of variables being in upper case.

2. We write as facts implications that state what follows from assumptions of normality. We thus write knowledge that would allow us to conclude a disjunct of abnormality from observations of incorrect behaviour.

The following example is used extensively in this paper to illustrate the differences between the frameworks. It was chosen because it is simple enough to distinguish many features of the systems, and it is not obvious how to apply each diagnostic model to it.

Example 2.4 Consider a battery, which, operating normally, has a voltage between 1.2 volts and 1.6 volts. Suppose that we are only considering one instant in time, and use $volt(B, V)$ to mean that the voltage across battery B is V.

Following the methodology above, we use the hypothesis $\neg ab(B)$ to mean that battery B is working correctly. We write facts that allow us to prove abnormality based on the voltage:

$$\forall B \neg ab(B) \Rightarrow \exists V(V > 1.2 \wedge V < 1.6 \wedge volt(B, V))$$

together with

$$\forall B \forall V_1 \forall V_2 \; volt(B, V_1) \wedge volt(B, V_2) \Rightarrow V_1 = V_2$$

The following shows the representation of how batteries in series work:

$$\begin{aligned} &\forall B_1 \forall B_2 \forall V \exists V_1 \exists V_2 (volt(series(B_1, B_2), V) \\ &\quad \Rightarrow \quad (V = V_1 + V_2) \wedge volt(B_1, V_1) \wedge volt(B_2, V_2)) \end{aligned}$$

The implication is this way around because we have to prove abnormalities from the observation of the series voltage.

Suppose we observe $volt(series(b_1, b_2), 1.456)$. If we assume $\neg ab(b_1)$, we conclude $ab(b_2)$. There are thus two diagnoses:

$$\{ab(b_1)\}, \{ab(b_2)\}$$

At the other extreme of the diagnostic task, we have no "normality" knowledge, only knowledge of possible faults and we want to conclude faults. Following Reiter's [1987] representation of the set covering model [Reggia, 1983], we treat the negation of faults as the "normal" case. Thus the methodology becomes

1. Make the negation of faults possible hypotheses. Thus assume that there are no faults unless it can be shown there are faults.
2. Axiomatise how symptoms imply faults.

If $c_1, ..., c_n$ are the possible causes for symptom s, we write

$$s \Rightarrow c_1 \vee ... \vee c_n.$$

For each c_i that always produce symptom s we add

$$c_i \Rightarrow s.$$

We can then use the absence of symptom s to eliminate c_i as a possible cause.

2.3 Abductive diagnosis

The abductive approaches naturally view the world in terms of causes and effects. The methodology is:

1. The possible hypotheses are the possible causes (faults, diseases) parameterised by the values on which they depend.
2. We axiomatise how symptoms follow from causes. These axioms should be facts if the symptom is always present given the cause and be possible hypotheses otherwise.

Poole[1988b] discusses the propositional fault-based case, and shows that the abductive and consistency-based approaches require different knowledge but get the same diagnoses.

For the normality-based case, we have hypotheses that the device is normal and hypotheses that the device is faulty. One possible cause is that a device is working normally, and one is that the device is broken. The possible hypotheses are parameterised enough to allow the hypothesised causes to be specific enough to imply the observations (e.g., just knowing that a battery is flat does not imply any particular voltage).

Example 2.5 To represent the battery of example 2.4, we parameterise the hypothesis by the battery and the voltage; thus we have the possible hypothesis

$$battOK(B, V)$$

which means that the battery B is working normally and producing voltage V. As facts we state what the assumption implies:

$$\forall B \forall V \; battOK(B, V) \Rightarrow volt(B, V) \wedge V > 1.2 \wedge V < 1.6$$

To predict the voltage of battery b_1, the best we could do is to assume that the battery is working normally, and predict what is in all extensions:

$$\exists V \;\; V > 1.2 \wedge V < 1.6 \wedge volt(b_1, V)$$

Note that this is exactly the knowledge that follows from the assumption of normality in the consistency-based approach.

If we observe that the voltage of b_1 is 1.403, there is one explanation:

$$battOK(b_1, 1.403)$$

which, together with the facts, logically implies the observation.

We still do not have enough information to fully characterise a diagnosis, as we cannot explain an abnormal output. We need to state how abnormal devices work. Using the second definition of abnormal, an abnormal device produces an output that is different to the normal output. We define a hypothesis $battAB(B, V)$ meaning that the battery B is broken and producing voltage V. We state that a broken battery outputs a voltage less than or equal to 1.2 volts or greater than or equal to 1.6 volts.

$$\forall B \forall V \; battAB(B,V) \Rightarrow volt(B,V) \wedge (V \leq 1.2 \vee V \geq 1.6)$$

For the abductive method, the knowledge about how batteries in series work is represented as:

$$\forall B_1 \forall B_2 \forall V \forall V_1 \forall V_2 (volt(B_1,V_1) \wedge volt(B_2,V_2) \\ \wedge V = V_1 + V_2 \Rightarrow volt(series(B_1,B_2),V))$$

Suppose $volt(series(b_1,b_2),1.456)$ is observed. There are three classes of diagnoses:

$\{battOK(b_1,V_1), battAB(b_2,V_2)\}$
for $1.2 < V_1 < 1.6 \wedge V_2 = 1.456 - V_1$

$\{battAB(b_1,V_1), battOK(b_2,V_2)\}$
for $1.2 < V_2 < 1.6 \wedge V_1 = 1.456 - V_2$

$\{battAB(b_1,V_1), battAB(b_2,V_2)\}$ for $(V_1 \geq 1.6 \vee$
$V_1 \leq -0.144 \vee 0.256 \leq V_1 \leq 1.2) \wedge V_2 = 1.456 - V_1$

In the general case, there are inputs and outputs to a system. In this case, the hypotheses must depend on all of the relevant parameters.

This methodology does not get us into the problems that [McDermott, 1987] claims are fatal flaws for abductive reasoning. His problem was that he did not allow the parameterisation of possible hypotheses.

3 Normality and Faults

3.1 Abduction

Example 3.1 Consider expanding the knowledge of example 2.4 to incorporate faults. Suppose a battery can be flat or shorted. If a battery is flat, its voltage is between 0.3 and 1.2 volts, and if it is shorted its voltage is zero.

For the abductive methods, we must prove the observations from the assumptions of faults or normality. We need the same axiomatisation of how normal batteries work, but replace the assumptions of abnormality with fault assumptions. The flat battery knowledge is represented as the hypothesis $flat(B,V)$ and the fact:

$$\forall B \forall V (flat(B,V) \Rightarrow volt(B,V) \wedge 0.3 < V \leq 1.2)$$

To represent the shorted battery we use the hypothesis $shorted(B)$ and the fact

$$\forall B \; shorted(B) \Rightarrow volt(B,0)$$

Suppose we observe that the voltage across batteries b_1 and b_2 in series is 1.517. This is the observation

$$volt(series(b_1,b_2),1.517)$$

For the abductive approach, we find sets of assumptions which imply the observation. We end up with the diagnoses:

$$\{shorted(b_1), ok(b_2,1.517)\}$$
$$\{ok(b_1,1.517), shorted(b_2)\}$$
$\{flat(b_1,V), flat(B_2,1.517-V)\}$ for $0.3 < V < 1.217$

3.2 Consistency-based approach

For the consistency based approach, we need to be able to conclude the faults from the observations. With both normality and faults, there are three choices as to what is assumed:

1. Assume normality and let faults be concluded as a side effect [Reiter, 1987, section 2.2].
2. Assume the absence of faults and let normality be concluded as a side effect.
3. Assume both normality and the absence of faults.

Example 3.2 Consider example 3.1. For the consistency-based approaches we must axiomatise how to conclude the faults from the observations. Let $flat(B)$ mean that B is flat, and $shorted(B)$ mean that B is shorted. The flat battery knowledge is represented:

$$\forall B \forall V (volt(B,V) \wedge 0.3 < V \leq 1.2 \Rightarrow flat(B))$$

Shorted batteries are represented by the fact

$$\forall B \; volt(B,0) \Rightarrow shorted(B)$$

For the consistency based diagnosis conjoin the observations to that facts and determine what can be proven from the assumptions. From the facts, the observation of the series voltage, and the assumption that b_1 is ok, we can prove:

$$\exists V \; volt(b_2,V) \wedge -0.083 < V < 0.212$$

This is inconsistent with the assumption that b_2 is OK, but it still doesn't allow us to find a fault. To allow us to find a fault we need to explicitly state that we have complete knowledge:

$$\forall B \; battery(B) \Rightarrow ok(B) \vee shorted(B) \vee flat(B)$$

With this extra knowledge, we can now prove

$$ok(b_1) \Rightarrow shorted(b_2)$$
$$ok(b_2) \Rightarrow shorted(b_1)$$

From the assumption that devices are OK, the possibility that both b_1 and b_2 could be flat is lost. This can be implied from

$$\neg shorted(b_1) \wedge \neg shorted(b_2) \Rightarrow flat(b_1) \wedge flat(b_2)$$

This indicates that the first choice (to assume normality and to conclude faults as a side effect) is wrong. We must assume the negation of faults. If we don't also assume normality, when there are no known faults for a device we must treat "abnormality" as a fault (as we did in the abductive diagnosis).

Example 3.3 Suppose we further elaborate our example, and add the fault that a battery can be overcharged, which means its voltage is between 1.6 and 2.0. Suppose we observe that the voltage of the batteries in series is 2.44. In the abductive method there are diagnoses corresponding to both batteries being OK, as well as diagnoses corresponding to one being overcharged and the other being flat.

If we use the knowledge for the consistency based approach we can prove

$(ok(b_1) \wedge ok(b_2)) \vee (flat(b_1) \wedge o\_c(b_2)) \vee (o\_c(b_1) \wedge flat(b_2))$

There is one diagnosis based on both batteries being OK. The question is how to interpret this diagnosis. If we follow the first interpretation of "normal" (section 2.1), this diagnosis means that anything could be wrong. If use the second interpretation of normality we need to appeal to parsimony (presumably due to probability and utility concerns).

It is unfair to compare this diagnosis with the abductive diagnosis that does not appeal to such parsimony. This notion of parsimony can be applied to the abductive approaches by preferring explanations that assume less (in set theoretic terms) faults [Poole, 1989]. This should be done if we are to compare equals.

4 Evolution and Modularity

4.1 From Normality-based to Fault-based

It is anticipated that in a large system, there will be some devices for which we have fault knowledge and some devices for which we have no such knowledge. We would expect that as we observe more cases we will gain experience as to what faults occur in practice, and would expect an artificial system to evolve from being abnormality based to being fault based.

For the consistency based diagnosis, if we assume normality as well as the absence of faults, faults are used to restrict the possible abnormalities (i.e, by stating $ab(D) \Rightarrow f_1(D) \vee ... \vee f_n(D)$). For each component we have the choice or either saying anything can go wrong or saying that there can only be a restricted number of faults.

For the abductive diagnosis and consistency-based diagnoses where we just assume the absence of faults, if there is no fault information, we have to treat abnormality as a fault. We need to state that if some device is abnormal (for a particular case), it produces a value different from the normal condition. As the system evolves, we would add in extra knowledge of faults and remove the fault which says that the device is just not working.

4.2 New faults

Suppose that we want to change our example so that there is another possible problem. For example, suppose it is possible that there was an error whereby some small batteries were in the circuit instead of the regular sized batteries. These small batteries produce a voltage of between 0.5 and 0.8 volts.

To add this to the abductive knowledge base, we add the hypothesis $small(B, V)$, and the fact

$$\forall B \forall V (small(B, V) \Rightarrow volt(B, V) \wedge 0.5 < V \leq 0.8)$$

If we add this, we find that everything works.

Consider how to add this to the consistency-based knowledge base. Unfortunately, it is not nearly as simple. We assumed complete knowledge of the possible faults. The existence of such small batteries is inconsistent with the facts given to the system. To change the knowledge base to incorporate the discovery of the new problem, we must change all facts which talked about the voltage range between 0.5 and 0.8 volts. We need to replace the now false statement

$$\forall B \forall V (volt(B, V) \wedge 0.3 < V \leq 1.3 \Rightarrow flat(B))$$

with

$$\begin{aligned} &\forall B \forall V \ volt(B, V) \Rightarrow \\ &\qquad (0.8 \leq V \leq 1.3 \Rightarrow flat(B)) \wedge \\ &\qquad (0.5 < V < 0.8 \Rightarrow flat(B) \vee small(B)) \wedge \\ &\qquad (0.3 < V \leq 0.5 \Rightarrow flat(B)), \end{aligned}$$

and the fact

$$battery(B) \Rightarrow ok(B) \vee shorted(B) \vee flat(B)$$

by

$$battery(B) \Rightarrow ok(B) \vee shorted(B) \vee flat(B) \vee small(B).$$

Forcing us to assume complete knowledge makes it very difficult to add new knowledge about faults.

5 Issues

5.1 What is an observation?

It may seem that the logical formulation of an observation is uncontroversial. However we find that this is not so. Consider the following example:

Example 5.1 Suppose we have a device d with a digital port at each end, and we input a value 4 to the left hand side at time t_0, and observe 16 at the right hand side at t_1. Using the relations $lhs(D, V, T)$ (meaning the value at the left hand port of device D at time T is V) and $rhs(D, V, T)$ (meaning the value at the right hand port of device D at time T is V), what is the logical expression representing our observation?

There seems to be two answers to this question:

$$lhs(d, 4, t_0) \wedge rhs(d, 16, t_1)$$

$$lhs(d, 4, t_0) \Rightarrow rhs(d, 16, t_1)$$

The argument for the first is that the input 4 and the output 16 is what we observed. At time t_0 there was a value of 4 on the left hand side and at time t_1 there was a 16 on the right hand side of device d.

The argument for the second representation is that the input has a different status than the output. We didn't observe the 4 on the left hand side, we put it there. We only observed the 16 when we put the 4 there. This is a very different case to putting a 16 on the right hand side and observing a 4 on the left hand side. Note that material implication does not imply a causal link between the input and the output.

The interesting thing about this example is that consistency based diagnosis requires observations in the first form, whereas the abductive diagnosis requires observations in the second form (unless $lhs(D, V, T)$ is a possible hypothesis)[3]. The simplest explanation is that both systems require the input to be a fact.

5.2 Noise

Consider the problem of noise. Using the ideas in the preceding section, that this problem is easily surmounted.

Example 5.2 Suppose we are told that our voltage reader has a possible error of ϵ. To represent this for the abductive methods, we need to identify the assumptions to prove the observation. Let $volt(B, V)$ mean V is the actual voltage across B, and $observed\_volt(R, B, V)$ mean V is the voltage measured by device R across B. We allow the hypothesis of an error, $error(R, E)$, which means that there is an error of E in the reader R and have the fact

$$\begin{aligned}&\forall B \forall E \forall V \forall R volt(B, V) \wedge error(R, E)\\ &\quad \Rightarrow observed\_volt(R, B, V + E) \wedge |E| < \epsilon\end{aligned}$$

This allows us to predict what value will be read from the reader, as well as allowing us to hypothesise an error in order to explain our observation. For the consistency-based diagnosis, we need the fact

$$\begin{aligned}&\forall R \forall B \forall V\, observed\_volt(R, B, V) \Rightarrow\\ &\quad \exists V_1\ |V - V_1| < \epsilon \wedge volt(B, V_1)\end{aligned}$$

Both of these formulations correctly handle errors. Being explicit about assumptions is important if we want a probabilistic analysis of the diagnoses [de Kleer, 1987, Neufeld, 1987]. Much work needs to be done in this area, however.

5.3 Hierarchical Reasoning

Both models can handle hierarchical models (as in [Genesereth, 1984]). There was nothing in the preceding sections that precluded the batteries being complex power stations. By finding the diagnoses at the level of power supply, we can use these diagnoses to diagnose individual power stations. Whether one model is better in some way than the other is an open question.

5.4 Epistemological Assumptions

For the consistency based diagnosis we had to explicitly make the complete knowledge assumption. We had to explicitly state that the known faults cover all possible faults. Unfortunately, this is usually false. It is, of course, an assumption that is made to build the knowledge base, but it is unfortunate that we need to state knowledge known to be false.

For the abductive-based diagnosis we do not need to make a complete knowledge assumption as part of the knowledge base. If we want to interpret the set of diagnoses as the set all possible problems, we are making a complete knowledge assumption about the knowledge base, and is not part of the knowledge base.

Another difference arises with unanticipated observations. With abduction, if we get an observation that was not anticipated, there is no diagnosis (as we can never imply that observation). Ad hoc techniques (e.g., assuming any observation for which there are no possible explanations) are possible, but unsatisfying. Unanticipated observations for consistency based diagnoses are just ignored as they will not be inconsistent with any set of assumptions. They are unable to detect when the system's knowledge is inadequate to explain a particular observation.

5.5 Efficiency

It is interesting to consider how the methods differ in implementation. In the consistency-based methods [Genesereth, 1984, Reiter, 1987, de Kleer, 1987], one can forward chain from the observations (i.e., finds what the observations imply), collecting assumptions of normality and absence of faults. When a contradiction occurs we know we have some faults. We then try to rule out faults by trying to prove the negation of faults.

For the abductive approach [Popl, 1973, PGA, 1987], one can backward chain from the observations and collect assumptions needed to prove the goals. We try to rule out faults by trying to prove the negation of faults.

If one follows the methodology described in the sections above, one essentially searches the same search space going forward from the observations in the consistency-based approach as going backwards from the observations in the abduction approach. The same knowledge and search is involved in ruling out possible faults.

The other supposed problem with these diagnostic models is their undecidability. Undecidable means powerful, not inefficient. It means the ability (in principle) to diagnose complex algorithms. Theorist should be regarded as a programming language, that we have to learn how to program efficiently (and to implement efficiently). Complexity is a property of a problem and an algorithm, it is not a property of a programming language. We desire the ability to solve simple problems simply, not the inability to solve hard problems.

[3] It seems as though this may indicate a way to decide which view of diagnosis is more natural. We could find out which of the two formulations was seen as a more natural way to represent an observation. I carried out an informal survey, and the results were very inconclusive. The common factor amongst everyone was the denial that the opposing view was sensible.

6 Conclusion

In this paper we have presented two different logical models of diagnosis, and compared them where we have both normality assumptions and fault assumptions. The important points are:

There is no representation which can lay claim to be the "one true logical definition of diagnosis".

Both diagnostic methods considered require a different representation of the world. Determining the method most applicable for different domains is an empirical question for which there has not been enough work. Even worse, many people do not realise they have a choice.

Both diagnostic frameworks can work in real-valued domains. Unfortunately, building general purpose theorem provers to support this is lagging behind the theory, but there is progress here, particularly in the use of constraint logic programming [Dincbas, 1988, Aiba, 1988].

One major difference between the diagnostic approaches is that in the abductive systems normality and faults are treated symmetrically, whereas in the consistency approaches they are treated as opposites.

Much of the discussion in this paper has not been specific to diagnosis, but can be applied to any recognition task, where the problem is to determine what is in a system (or a picture) based on observations of the system. For example, one can see [Kautz, 1987] as using the idea of consistency-based diagnosis with faults corresponding to plan objects. This paper indicates that there is a corresponding abductive theory of plan recognition.

The results of this study indicate that there may not be a representation of knowledge that is independent of how the knowledge is to be used. The systems will not work if some of the knowledge is in the form required by the abductive diagnostic methods and some in the form required by the consistency based methods; they do not mix well. There is no advantage gained by having the union of the knowledge needed for each model.

References

[Aiba, 1988] A. Aiba et. al., "Constraint Logic Programming Language CAL", *Proc. International Conference of Fifth Generation Computing Systems*, 263-276.

[Cox, 1987] P. T. Cox and T. Pietrzykowski, *General Diagnosis by Abductive Inference*, Technical report CS8701, School of Computer Science, Technical University of Nova Scotia, April.

[de Kleer, 1987] J. de Kleer and B. C. Williams, "Diagnosing Multiple Faults", *Artificial Intelligence, Vol. 32, No. 1*, pp. 97-130.

[Dincbas, 1988] M. Dincbas et. al., "The Constraint Logic Programming Language CHIP", *Proc. International Conference of Fifth Generation Computing Systems*, pp. 693-702.

[Genesereth, 1984] M. R. Genesereth, "The Use of Design Descriptions in Automated Diagnosis", *Artificial Intelligence 24*, 411-436.

[Kautz, 1987] H. A. Kautz, "A Formal Theory for Plan Recognition", Department of Computer Science, University of Rochester, Technical Report TR-215.

[Kowalski, 1979] R. Kowalski, "Algorithm = Logic + Control", *Comm. ACM*, 22 (7), 424-436.

[McCarthy, 1969] J. McCarthy and P. J. Hayes, "Some philosophical problems from the standpoint of artificial intelligence", in B. Meltzer and D. Michie (eds.) *Machine Intelligence 4*, Edinburgh University Press, 463-502.

[McDermott, 1987] D. McDermott, "A Critique of Pure Reason", *Computational Intelligence 3*, 151-160.

[Neufeld, 1987] E. M. Neufeld and D. Poole, "Towards solving the multiple extension problem: combining defaults and probabilities", *Workshop on Reasoning with Uncertainty*, Seattle, July.

[PGA, 1987] D. L. Poole, R. G. Goebel, and R. Aleliunas, "Theorist: a logical reasoning system for defaults and diagnosis", in N. Cercone and G.McCalla (Eds.) *The Knowledge Frontier: Essays in the Representation of Knowledge*, Springer Varlag, New York, pp. 331-352.

[Poole, 1988a] D. L. Poole, "A Logical Framework for Default Reasoning", *Artificial Intelligence*, Vol. 36, No. 1, pp. 27-47, August 1988.

[Poole, 1988b] D. Poole, "Representing Knowledge for Logic-Based Diagnosis", *Proc. International Conference on Fifth Generation Computing Systems*, Tokyo, 1282-1290.

[Poole, 1989] D. Poole, "Explanation and Prediction: An Architecture for Default and Abductive Reasoning", Tech Report 89-4, Comp. Sci., U.B.C.

[Popl, 1973] H. E. Popl, Jr., "On the Mechanization of Abductive Logic", *Proc. IJCAI-73*, 147-152.

[Reggia, 1983] J. A. Reggia, D. S. Nau and P. Y. Wang, "Diagnostic expert systems based on a set covering model", *Int. J. Man-Machine Studies 19*, pp.437-460.

[Reiter, 1987] R. Reiter, "A Theory of Diagnosis from First Principles" *Artificial Intelligence*, Vol. 32, No. 1, pp. 57-96.