Preferring maximum confirmation diagnoses

Nico Roos

Department Knowledge Engineering, Maastricht University
P.O.Box 616, 6200 MD Maastricht, The Netherlands
email: roos@maastrichtuniversity.nl

Abstract

Models used for Model-Based Diagnosis usually assume that the inaccuracy of data is smaller than the precision with which the data is described. In some domains, however, this assumption is invalid. Observations may not be accurate or the behavior model of the system does not allow for accurate predictions. Therefore, the accuracy of predictions, which is a function of the accuracy of the observed system inputs and the behavior model of the system, may differ from the accuracy of the observed system outputs.

This paper investigates the consequences of using inaccurate values. The paper will show that traditional notions of preferred diagnoses such as abductive diagnosis and minimum consistency-based diagnosis are no longer suited if the available data has different accuracies. A new notion of preferred diagnoses, called *maximum confirmation diagnoses*, is introduced.

1 Introduction

Models used for Model-Based Diagnosis usually assume that the inaccuracy of data is smaller than the precision with which the data is described. This assumption, which is usually not stated explicitly, implies that we can easily compare predictions and observations. In domains where data is inaccurate or where accuracy is not important, abstract values such as $\{-, 0, +\}$ or $\{low, high\}$ mays sometimes be used [17, 18]. Abstraction from specific values may reduce the diagnostic precision and may therefore be undesirable. In that case, representations that precisely express the inaccuracy, such as inequalities or intervals of values $[ul, ub]$ may be used [15]. Several papers deal with consistency-based diagnosis given inaccurate data [3, 8, 9, 10].

The use of inaccurate values raises a number of problems with respect to the notion of *preferred diagnoses*. Normally, minimal or minimum diagnoses are preferred assuming that components fail independently and that fault probabilities are low. Abductive diagnoses are preferred assuming that we know all the ways in which components may fail. Maximum-informative diagnoses [14, 16] are related to abductive diagnoses. Maximum-informative diagnoses do not require that all ways in which components may fail are known. Instead, it is based on the assumption that the probability that system outputs depending on failing components produce correct values, is small.[1] In all cases, unlikely diagnoses may be preferred if inaccurate values are used.

To give an illustration of the problem with minimum / minimal diagnoses, consider a minimal diagnosis Δ that enables us to predict that some output value lays in the interval $[3, 5]$ while a non-minimal diagnosis Δ' enables us to predict that the output value lays in the interval $[3, 7]$. If we observe that the output value must lay in the interval $[5, 7]$, then clearly Δ' should be preferred. The probability that the diagnosis Δ is correct is much smaller than the probability that Δ' is correct because in the former case, 5 is the only value on which the prediction and the observation agree, while in the latter case, they agree on the interval $[5, 7]$.

Abductive and maximum-informative diagnosis have other problems. The predicted value of some output given a diagnosis Δ may be less accurate than the observed value of that output. Though the observation confirms the prediction based on Δ, Δ is neither an abductive nor a maximum-informative diagnosis [12, 2]. Abductive and maximum-informative diagnosis require that the predicted value of an output is at least as accurate as the observed value.

[1] This assumption does not hold for systems such as digital circuits.

Cordier [2] proposed to adapt the definition of an abductive diagnosis to cope with observations that are more accurate the the predicted output values. This paper, however, proposes a new notion of diagnosis, called *maximal-confirmation* diagnosis. A maximum-confirmation diagnoses is based on measuring to what extent predictions of outputs given a diagnosis are confirmed by the observations. A maximum confirmation diagnosis therefore refines the new definitions of abductive diagnosis proposed by Cordier [2].

Maximal-confirmation diagnoses do not distinguish between a diagnosis Δ' that enables us to predict that an output value lays in the interval $[3, 7]$ and a diagnosis Δ'' that enables us to predict that an output value lays in the interval $[-\infty, +\infty]$ if we have observed that the output value lays in the interval $[5, 7]$. The diagnosis Δ' is more accurate than the diagnosis Δ''. Therefore, we propose a second preference relation, namely *maximal accuracy and confirmation* diagnosis (mac-diagnosis).

The remainder of the paper is organized as follows. In the next section, we start with introducing our diagnostic framework. Section 3 discusses the problems with preferred diagnoses and offers a solution in the form of maximal-confirmation and mac-diagnoses. In Section 4, a formal underpinning of the proposed solutions is given. Section 5 describes related work and Section 6 concludes the paper.

2 The diagnostic setting

Model-based diagnosis starts from a description of a system to be diagnosed. The system description specifies the normal behavior of the system and possibly also the abnormal behavior. Classical Model-Based Diagnosis (MBD) [13, 5, 4] describes a system of connected components. Each component has a number of inputs and outputs. The values of a component's outputs are a function of the values of the component's inputs and the component's health mode. A model of the component describes this function. The description may be partial, but it will always contain the components normal behavior; i.e., the behavior description given the health mode 'normal'.

This paper we abstract from the exact description of the system and the hypotheses describing the health modes of components. Instead we assume that the system description Sd is given and that it describes at least the normal behavior of the system. Moreover, we assume a set of candidate diagnoses $\mathcal{D}$ where each diagnosis $\Delta \in \mathcal{D}$ gives a possible description of the health modes of the components of the system. Finally, we assume a set of possible observations $\mathcal{O}$ of the system. So, the triple $Pd = \langle \mathcal{D}, Sd, \mathcal{O} \rangle$ describes our problem domain.

When making observations $O \in \mathcal{O}$ about the behavior of the system, the observations made may not correspond with the normal behavior of the system:

$$\Delta^{nor} \cup Sd \cup \mathcal{B} \cup O \models \bot$$

Here $\Delta^{nor} \in \mathcal{D}$ denotes the hypothesis that every component behaves normally, and $\mathcal{B}$ denotes the general background knowledge. If the expected behavior of the system does not correspond with observations made, we would like to identify the components that behave abnormally, giving us the diagnostic problem: $P = (Pd, O)$.

The two main forms of diagnosis are *consistency-based diagnosis* and *abductive diagnosis*. In consistency-based diagnosis, we search for a diagnosis such that the system description and the observations of systems are consistent [13, 5, 4].

Definition 1 *Let $P = (Pd, O)$ be a diagnosis problem. Moreover, let $\Delta \in \mathcal{D}$ be a candidate diagnosis.*
Δ *is a* consistency-based diagnosis *of the diagnosis problem $P = (Pd, O)$ iff*

$$\Delta \cup Sd \cup \mathcal{B} \cup O \not\models \bot$$

Abductive diagnosis uses a stronger requirement. Given the observed system inputs and a diagnosis, we must be able to predict the observed system outputs [11, 1].

Definition 2 *Let $P = (Pd, O)$ be a diagnosis problem. Moreover, let $\Delta \in \mathcal{D}$ be a candidate diagnosis. Finally, let the observations O be partitioned into system inputs O^{in} and systems outputs O^{out}.*
Δ *is an* abductive diagnosis *of the diagnosis problem $P = (Pd, O)$ iff*

$$\Delta \cup Sd \cup \mathcal{B} \cup O^{in} \models O^{out}$$
$$\Delta \cup Sd \cup \mathcal{B} \cup O^{in} \not\models \bot$$

It is not difficult to see that an abductive diagnosis is always a consistency-based diagnosis. It follows from the property *cumulativity* of predicate logic:

$$\Delta \cup Sd \cup \mathcal{B} \cup O^{in} \subseteq \Delta \cup Sd \cup \mathcal{B} \cup O \subseteq Cn(\Delta \cup Sd \cup \mathcal{B} \cup O^{in})$$
$$\text{implies } Cn(\Delta \cup Sd \cup \mathcal{B} \cup O^{in}) = Cn(\Delta \cup Sd \cup \mathcal{B} \cup O)$$

Here, $Cn(\Sigma) = \{\varphi \mid \Sigma \vdash \varphi\}$ denotes the set of consequences.

The converse does not hold. Abductive diagnosis requires knowledge about the faulty behavior of components (also called: *fault models*). Without this knowledge we cannot determine an abductive diagnosis. However, we will still be able to determine a consistency-based diagnosis. If fault models are available, and if the set of possible values of system in- and outputs do not (partially) overlap, then a consistency-based diagnosis is an abductive diagnosis. Values do not overlap iff;

for every $O, O' \in \mathcal{O}$ if $O \cup O' \cup \mathcal{B} \not\models \bot$, then $O = O'$.

So, we do not use values of different accuracies.

Proposition 1 *A consistency-based diagnosis is an abductive diagnosis if the set of fault models is complete and if the problem domain Pd does not allow for partially overlapping values.*

Proof. Let Δ be a consistency-based diagnosis: $\Delta \cup Sd \cup \mathcal{B} \cup O \models \bot$.

Suppose that $\Delta \in \mathcal{D}$ is no abductive diagnosis. Then: $\Delta \cup Sd \cup \mathcal{B} \cup O^{in} \not\models O^{out}$. Since knowledge about the systems behavior is complete, there is a $O' \in \mathcal{O}$ such that: $O' \neq O^{out}$ and $\Delta \cup Sd \cup \mathcal{B} \cup O^{in} \models O'$. Since $O' \neq O^{out}$, the non overlapping values assumption implies that $O' \cup O^{out} \cup \mathcal{B} \models \bot$ and therefore: $\Delta \cup Sd \cup \mathcal{B} \cup O^{in} \cup O^{out} \models \bot$. Hence, Δ cannot be a consistency-based diagnosis, contradicting our starting point. $\square$

3 Inaccurate predictions and observations

The use of inaccurate values implies that the 'non-overlapping values' assumption is invalid. Giving up this assumption has no influence on the definition of consistency-based and abductive diagnosis. However, it does influences the preferred diagnoses among the set of consistency-based diagnoses.

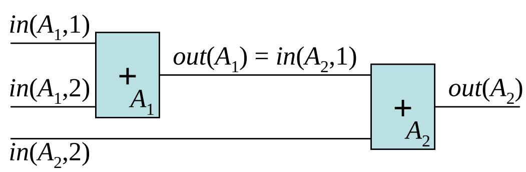

Figure 1: A system consisting of two adders.

3.1 Preferences and accuracy

If predictions and observations have different accuracies, does this impact the diagnoses that we should prefer? To investigate this question, consider a system consisting of two analog adders A_1 and A_2, shown in Figure 1. The normal behavior of the adders is described by:

$$\forall x[adder(x) \wedge m(x, nor) \rightarrow \forall l_1, l_2, l_3, u_1, u_2, u_3[v(in(x,1), [l_1, u_1]) \wedge$$
$$v(in(x,2), [l_2, u_2]) \wedge v(out(x), [l_3, u_3]) \wedge (l_1 + l_2 = l_3) \wedge (u_1 + u_2 = u_3)]]$$

Here, $m(x, h)$ denotes that h is the health mode of component x, nor is the health mode in which the component function normally, and $v(io, [lb, ub])$ denotes that the values of in- or output io lays in the interval $[lb, ub]$. Note that a predicate $v(io, [lb, ub])$ is used instead of a function $v(io) = [lb, ub]$ since we may have different intervals consistently describing the same in- or output: e.g. $v(io, [1, 5])$ and and $v(io, [4, 6])$.

For these adders we also know their faulty behavior, denoted by the health mode: offset. The output of the adder has a fixed offset of 3 given the health mode offset:

$$\forall x[adder(x) \wedge m(A, \mathit{offset}) \rightarrow \forall l_1, l_2, l_3, u_1, u_2, u_3[v(in(x,1), [l_1, u_1]) \wedge$$
$$v(in(x,2), [l_2, u_2]) \wedge v(out(x), [l_3, u_3]) \wedge (3 + l_1 + l_2 = l_3) \wedge (3 + u_1 + u_2 = u_3)]]$$

The structural description of the system specifies that the output of adder A_1 is connected to input 1 of adder A_2: $out(A_1) = in(A_2, 1)$.

Suppose that we make the following observations:

$$v(in(A_1, 1), [1, 2]), v(in(A_1, 2), [2, 3]), v(in(A_2, 2), [3, 4]), v(out(A_2), [12, 16])$$

Clearly the system description together with the observations are inconsistent. There are two minimum consistency-based diagnoses given the observations: $\Delta_1 = \{m(A_1, \mathit{offset})\}$ and $\Delta_2 = \{m(A_2, \mathit{offset})\}$. Both diagnoses are not very likely since 12 is the only value on which the prediction and the observation of output A_2 agree. Every value in the interval $(12, 16]$ is inconsistent with: $\Delta_i \cup Sd \cup \mathcal{B} \cup O^{in}$ with $i \in \{1, 2\}$. To maximize the likelihood, we should prefer diagnoses that maximizes the number of possible output values that are consistent with the observations. In other words, we must prefer *abductive* diagnoses. The diagnosis $\Delta_3 = \{m(A_1, \mathit{offset})\}, m(A_2, \mathit{offset})\}$ is such an abductive diagnosis. The observed system outputs $O^{out} = \{v(out(A_2), [12, 16])\}$ can be derived form the system description together with the observed system inputs $O^{in} = \{v(in(A_1, 1), [1, 2]), v(in(A_1, 2), [2, 3]), v(in(A_2, 2), [3, 4])\}$ and the diagnosis Δ_3.

Since abductive diagnosis requires knowledge of the system's faulty behavior, we might conclude that if this knowledge is available, abductive diagnoses should be preferred. This conclusion is, however, premature.

To illustrate the problem, suppose that we observed for the system output: $v(out(A_2), [7, 8])$ instead of $v(out(A_2), [12, 16])$. Note that now all observations are made with the same accuracy. This is quite common if similar measurement devises are used. The accuracy with which the output value of adder A_2 can be predicted in the absence of faults is, however, lower than the accuracy of the observation. That is, the most accurate prediction is: $v(in(A_2, 1), [6, 9])$, while $v(out(A_2), [7, 8])$ has been observed. This implies that abductive diagnosis is infeasible even if the adders behave normally as might be the case in this example. Although the observation cannot be explained by the normal behavior of the adders, it does confirm the normal behavior. This suggests that we need a new notion of preferred diagnosis, namely *confirmation* diagnosis. Abduction can than be viewed as a weak form of confirmation.

3.2 Maximal confirmation diagnosis

The idea that is put forward in this section is to prefer *confirmation* diagnoses. This preference is motivated by the fact that (*i*) some minimum / minimal diagnoses can be very unlikely, and (*ii*) abductive diagnoses may not be possible even if complete information about the faulty behavior is available. Concerning the latter, since abductive diagnosis is only possible if observations are sufficiently *inaccurate*, confirmation of the predictions made is a better criterium. Of course, we must also be able to deal with partial confirmations.

In order to give a general definition of confirmation diagnoses, which is not limited to intervals of values, we assume that an accuracy ordering can be defined over the set of possible observations $\mathcal{O}$.

Definition 3 *Let $\mathcal{O}$ be the set of possible observations and let $O, O' \in \mathcal{O}$ be two observations. Moreover, let $\mathcal{B}$ be general background knowledge.*
O is at least as accurate as O', denoted by $O \sqsubseteq O'$ iff $O \cup \mathcal{B} \models O'$.

Without lost of generality, we assume that the ordering over $\mathcal{O}$ forms a lattice with bottom element $\bot$ (false) and top element $\top$ (true). Therefore, we also have a meet operator $O \wedge O'$ and a join operator $O \vee O'$.

The accuracy ordering of in- and output values enables us to formalize the confirmation of predictions made. We say that an observation O *strongly* confirms a prediction O' iff $O \sqsubseteq O'$. An observation O *weakly* confirms a predicted value O' iff $O' \sqsubseteq O$. We can also define a notion of partial confirmation. An observation O *partially* confirms a predicted value O' iff $\bot \sqsubset O \cap O'$, $O \not\sqsubseteq O'$ and $O' \not\sqsubseteq O$. Figure 2 gives an illustration using one output and possible observations described by intervals.

It is clear that there are different degrees in which an observation can confirm a predicted value. The above introduced order on possible observations $\mathcal{O}$ can be used to order diagnoses with respect to the degree of confirmation. The confirmation degree of a diagnosis is the meet between between the observation made and the most accurate prediction of the system outputs.

$$CD(\Delta) = O^{out} \wedge Pred(\Delta)$$

where $Pred(\Delta)$ is the most accurate predicted system output.

$$Pred(\Delta) = O \text{ iff } \Delta \cup Sd \cup \mathcal{B} \cup O^{in} \models O \text{ and for no } O' \sqsubset O\text{: } \Delta \cup Sd \cup \mathcal{B} \cup O^{in} \models O'.$$

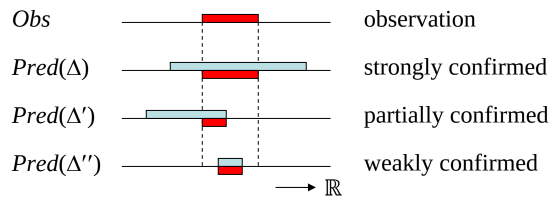

Figure 2: Different confirmation degrees.

Note that a partial or a weak confirmation of some predicted value of a system output x leads to a lower confirmation degree $CD(\Delta)$. Also note that Δ is no consistency-based diagnosis if $CD(\Delta) \equiv \perp$.

The confirmation degree $CD(\cdot)$ can be used to define the *maximal-confirmation* diagnoses.

Definition 4 *Let $P = (Pd, O)$ be a diagnosis problem.*

Δ *is a* maximal-confirmation diagnosis *of the diagnosis problem $P = (Pd, O)$ iff for* no *diagnosis Δ':*

$$CD(\Delta) \sqsubset CD(\Delta')$$

3.3 Maximal accuracy and confirmation diagnosis

Every diagnosis Δ that is strongly confirmed by the observations has the same degree of confirmation. The degree of confirmation of these diagnoses corresponds to the degree of confirmation of a diagnosis that exactly predicts the observations made. To give an illustration, reconsider the example of the system with the two adders. Suppose that we also have the general unknown behavior denoted by the health model ab (abnormal):

$$\forall x[adder(x) \wedge m(x, ab) \rightarrow \top]$$

which is equivalent to:

$$\forall x[adder(x) \wedge m(x, ab) \rightarrow v(out(x), [-\infty, +\infty])]$$

Given the observations:

$$v(in(A_1, 1), [1, 2]), v(in(A_1, 2), [2, 3]), v(in(A_2, 2), [3, 4]), v(out(A_2), [9, 12])$$

we have four maximal confirmation diagnoses: $\Delta_1 = \{m(A_1, \mathit{offset})\}$, $\Delta_2 = \{m(A_1, ab)\}$, $\Delta_3 = \{m(A_2, \mathit{offset})\}$ and $\Delta_4 = \{m(A_2, ab)\}$. For all diagnoses we have:

$$CD(\Delta_1) = CD(\Delta_2) = CD(\Delta_3) = CD(\Delta_4) = \{v(out(A_2), [9, 12])\}$$

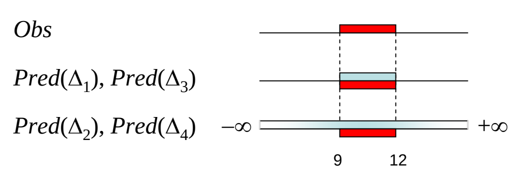

Figure 3: Strongly confirmed, different accuracies.

Diagnoses Δ_1 and Δ_3 both explain and are strongly confirmed by the observed output value, while diagnoses Δ_2 and Δ_4 are strongly confirmed by the observation but do not explain the observation. Δ_2 and Δ_4 ignore the information about the faulty behavior of the adders, They only state that one of the adders behaves abnormally. Note that an observation such as $v(out(A_2), [12, 15])$ also strongly confirms Δ_2 and Δ_4. Diagnoses Δ_1 and Δ_3 are not strongly confirmed by the latter observation. Since the predictions of diagnoses Δ_1 and Δ_3 are more accurate and can therefore more easily be disconfirmed, we propose a second ordering principle, preferring diagnoses that give the most accurate predictions. Together, this results in preferring *maximal accuracy and confirmation* (mac-) diagnoses.

Definition 5 *Let $P = (Pd, O)$ be diagnosis problem.*

Δ *is a* maximal accuracy and confirmation diagnosis *(mac-diagnosis) of the diagnosis problem $P = (Pd, O)$ iff*

- Δ *is a maximal confirmation diagnosis,*
- *for* no *maximal confirmation diagnosis Δ': $Pred(\Delta') \sqsubset Pred(\Delta)$.*

4 Justification of preferences

One of the motivations of preferring maximal-confirmation diagnoses is because they are more likely than other diagnoses. We will now formalize this notion of likelihood.

Normally, we prefer the most probable diagnoses given the observation made. Using some general assumptions, this leads to preferring minimal or minimum diagnoses. The assumptions are:

- fault probabilities of components are less than 0.5 or are very small, respectively;
- the predicted value of a system output is either equal to an observation or is unknown.

Using inaccurate values, the second assumption is no longer valid. Instead, the probability that the actual value of a system output corresponds with the observation made, is important. As we have seen in the example in the introduction of the previous section, of all the values that were possible according to the prediction, only one value was allowed by the observation. As a result the probability that this diagnosis is correct will be low whatever its a priori probability. Diagnoses that allow for more overlap between predictions and observations will have a higher probability. The following derivation shows this in a formally:

$$\begin{aligned} P(\Delta \mid O) &= P(O \mid \Delta) \cdot \frac{P(\Delta)}{P(O)} && (1) \\ &= P(O \mid Pred(\Delta)) \cdot P(Pred(\Delta) \mid \Delta) \cdot \frac{P(\Delta)}{P(O)} && (2) \\ &= P(Pred(\Delta) \mid O) \cdot \frac{P(O)}{P(Pred(\Delta))} \cdot P(Pred(\Delta) \mid \Delta) \cdot \frac{P(\Delta)}{P(O)} && (3) \\ &= P(Pred(\Delta) \mid O) \cdot P(Pred(\Delta) \mid \Delta) \cdot \frac{P(\Delta)}{P(Pred(\Delta))} && (4) \\ &= P(CD(\Delta) \mid O) \cdot P(\Delta \mid Pred(\Delta)) && (5) \end{aligned}$$

In the above derivation, the following issues should be noted:

1. The conditional probability $P(O \mid Pred(\Delta))$ in equation 2 is conditionally independent of the diagnosis Δ.
2. The conditional probability $P(Pred(\Delta) \mid O)$ in equations 3 and 4 is equal to *confirmation probability* $P(CD(\Delta) \mid O)$ in equation 5 since $Pred(\Delta) \wedge O^{out} \equiv CD(\Delta)$.
3. By preferring maximal-confirmation diagnoses, Definition 4, we maximize the confirmation probability:

   $P(CD(\Delta) \mid O) \geq P(CD(\Delta') \mid O)$ iff $CD(\Delta') \sqsubseteq CD(\Delta)$
4. If observed and predicted values are accurate, $P(CD(\Delta) \mid O)$ will either be 0 or 1, corresponding to whether Δ is an abductive diagnosis.
5. Since $Pred(\Delta)$ describes the system outputs given a diagnosis Δ, clearly, $P(Pred(\Delta) \mid \Delta) = 1$. Hence, equation 4 becomes:

$$P(\Delta \mid O) = P(Pred(\Delta) \mid O) \cdot \frac{P(\Delta)}{P(Pred(\Delta))} \qquad (6)$$

6. The *explanation probability* $P(\Delta \mid Pred(\Delta))$ in equation 5 expresses the conditional probability that Δ is a diagnosis given the system outputs $Pred(\Delta)$ that can be *explained* by Δ. The explanation probability corresponds with probability of the diagnosis of the system if observed and predicted values are accurate.

7. The explanation probability $P(\Delta \mid Pred(\Delta))$ in equation 5 is inversely propositional with the number of diagnoses Δ' such that $Pred(\Delta) \cup Pred(\Delta') \cup \mathcal{B} \not\models \bot$. The number of diagnoses Δ' that can (partially) explain $Pred(\Delta)$ decreases if $Pred(\Delta)$ becomes more accurate. If $Pred(\Delta)$ is maximally *inaccurate*, any diagnosis is possible and $P(\Delta \mid Pred(\Delta))$ will be low. If, however, $Pred(\Delta)$ is maximally *accurate*, Δ might be the only possible diagnosis. Clearly $P(\Delta \mid Pred(\Delta)) = 1$ if Δ is the only diagnosis that can explain $Pred(\Delta)$.

Discussion We maximize the confirmation probability in equation 5 by focussing on maximum-confirmation diagnoses. The confirmation probability becomes 1 if the observations are strongly confirmed. If observations are weakly or partially confirmed, then the confirmation probability will less than one.

Weakly confirmed diagnoses are abductive diagnoses where the predictions are more accurate than the observations. The more accurate the predictions, the lower the confirmation probability. Our preference for maximal-confirmation diagnoses is a preference for diagnoses resulting in less accurate predictions and thereby increasing the confirmation probability.

The explanation probability $P(\Delta \mid Pred(\Delta))$ in equation 5 does not increase with a decrease in the accuracy of the predictions given a diagnosis. The opposite holds, diagnoses allowing for more accurate predictions will be more probable given the predicted values because there are less diagnoses that can (partially) explain the same predicted values. So, while the confirmation probability increases, the explanation probability decreases.

The confirmation probability is the dominant factor in equation 5. $P(\Delta \mid Pred(\Delta))$ may become 1 if $Pred(\Delta)$ is very accurate. If $Pred(\Delta)$ describes a unique value while the observation describes an interval, the confirmation probability $P(CD(\Delta \mid \Delta)$ will approximate the value 0. If $Pred(\Delta)$ is equal to the observations or less accurate than the observations, then $P(CD(\Delta \mid \Delta) = 1$ while $P(\Delta \mid Pred(\Delta))$ becomes smaller but will not approximate 0. Since the confirmation probability varies over a larger range, it is generally the dominant factor. Therefore, we should prefer diagnoses that strongly confirm the observations.

Diagnoses that are strongly confirmed by the observations maximize the confirmation probability; i.e., $P(CD(\Delta) \mid O) = 1$. Among these diagnoses, the most accurate diagnoses maximize the explanation probability in equation 5; i.e., $P(\Delta \mid Pred(\Delta))$, without changing the confirmation probability. This justifies our preference for mac-diagnoses.

5 Related work

The use of inaccurate values in diagnosis is related to, but differs from the use of value abstraction [17, 18] and domain abstraction [6]. Abstraction enables us to focus on the relevant aspects while ignoring other details. We may abstract from the specific values of the in- and outputs of a system. Although the abstracted values do not accurately describe the actual in- and output values, the inaccuracy is irrelevant if the abstract values suffices to make a diagnosis. If, however, the abstract values are insufficient for making a diagnosis, the inaccurate values should be used.

Reasoning with intervals or inequations is closely related to the use of inaccurate values. Several authors have studied reasoning with intervals and inequations in a diagnosis system. See for instance, [3, 9, 8]. Reasoning with intervals and inequations turns out to be a source of computational overhead because in- and outputs of components may have multiple values. One cannot simply ignore the intervals or inequations that are subsumed by other intervals or inequations. Each derived interval or inequation may be supported by different sets of assumptions about the health modes of components. Considering the consequences of all derived intervals or inequations together with the underlying assumptions may result in a combinatory explosion. Fortunately, for diagnosis, it is not always necessary to consider all derived intervals or inequations. Unnecessary computations can be avoided by ignoring derived intervals and inequations as long as there is no evidence that they cannot be ignored, and by identifying minimal conflicts using the derivation tree for a derived inconsistency [10].

Cordier [2] has addressed consequences of using inaccurate values for abductive diagnosis. The here proposed notion of maximal confirmation diagnosis generalizes her modified definitions of abductive diagnosis by (*i*) providing a measure of confirmation, and (*ii*) using this measure to order different diagnoses. The idea of ordering diagnoses w.r.t. the degree of confirmation was first proposed by Roos and Witteveen [15]. They describe diagnosis of a Simple Temporal Network [7], a formalism for representing a plan together with the temporal constraints on plan execution. Diagnosis of temporal constraint violations raised a number of issues among which the confirmation of observations. This paper extends previous work in several directions. First, a general framework for diagnosis when using inaccurate values is introduced. Second, inaccuracy need not be described by intervals. Third, maximal-confirmation instead of maximum-confirmation diagnoses are introduced. Fourth, mac-diagnoses are introduced. Finally, a formal justification of maximal-confirmation and mac-diagnoses is given.

6 Conclusion

Models used for Model-Based Diagnosis usually assume that the inaccuracy of data is smaller than the precision with which the data is described. In some domains, however, this assumption is invalid.

The use of inaccurate values raises a number of problems with respect to the notion of *preferred diagnoses*. Normally, minimal or minimum diagnoses, abductive diagnoses or maximum-informative diagnoses are preferred among the consistency-based diagnoses. We have seen that in case observations and predictions of the system's behavior are inaccurate, these preferences are no longer adequate. Instead, the paper argues for preferring *maximal-confirmation diagnoses*, and *maximal-confirmation and accuracy diagnoses*.

To summarize, a general framework for diagnosis when using inaccurate values is introduced. The inaccuracy need not be described by intervals in this framework. Problems with minimum / minimal and abductive diagnoses are demonstrated and a solution in the form of maximal-confirmation diagnoses and maximal-confirmation and accuracy diagnoses is presented. A formal justification of maximal-confirmation and mac-diagnoses is given. Moreover, the application of the results to other model-based diagnosis approaches is discussed.

References

[1] L. Console and P. Torasso. Hypothetical reasoning in causal models. *International Journal of Intelligence Systems*, 5:83–124, 1990.

[2] M.-O. Cordier. When abductive diagnosis fails to explain too precise observations: an extended spectrum of definitions based on abstracting observations. In *International Workshop on Principles of Diagnosis [DX-1998]*, pages 24–31, 1998.

[3] P. Dague, O. Jehl, P. Deveès, P. Luciani, and P. Taillibert. *When Osillators stop osillating*, pages 235–241. 1992.

[4] J. de Kleer, A.K. Mackworth, and R. Reiter. Characterizing diagnoses and systems. *Artificial Intelligence*, 56:197–222, 1992.

[5] J. de Kleer and B. C. Williams. Diagnosing multiple faults. *Artificial Intelligence*, 32:97–130, 1987.

[6] Johan de Kleer. Dynamic domain abstraction through meta-diagnosis. In *Abstraction, Reformulation, and Approximation (SARA), LNAI 4612*, pages 109–123. Springer, 2007.

[7] R. Dechter, I. Meiri, and J. Pearl. Temporal constraint networks. *Artificial Intelligence*, 49:61–95, 1991.

[8] D. J. Goldstone. *Controlling inequality reasoning in a TMS-based analog diagnosis system*, pages 206–211. 1992.

[9] Walter Hamscher. Acp: Reason maintenance and inference control for constraint propagation over intervals. In *AAAI*, pages 506–511, 1991.

[10] Jakob Mauss and Mugur Tatar. Computng minimal conflicts for rich constraint lanuages. In *Thirteenth International Workshop on Principles of Diagnosis [DX-2002]*, pages 170–181, 2002.

[11] D. Poole. Representing knowledge for logic-based diagnosis. In *International Conference on Fifth Generation Computer Systems*, pages 1282–1290, 1988.

[12] D. Poole. An methodology for using a default and abductive reasoning system. *International Journal of Intelligent Systems*, 5:521–548, 1990.

[13] R. Reiter. A theory of diagnosis from first principles. *Artificial Intelligence*, 32:57–95, 1987.

[14] N. Roos and C. Witteveen. Models and methods for plan diagnosis. In *Formal Approaches to Multi-Agent Systems (FAMAS'06)*, 2006.

[15] N. Roos and C. Witteveen. Diagnosis of simple temporal networks. In *ECAI 2008*, pages 593–597, 2008.

[16] N. Roos and C. Witteveen. Models and methods for plan diagnosis. *Journal of Autonomous Agents and Multi-Agent Systems*, 19:30–52, 2009.

[17] P. Struss. *What's in SD?: Towards a theory of modeling for diagnosis*, pages 419–449. 1992.

[18] G. Tota and P. Torasso. Automatic abstraction in component-based diagnosis driven by system observability. In *IJCAI-03*, pages 394–400, 2003.