Knowledge-Based Systems 24 (2011) 467–477

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Maximal-confirmation diagnoses

Nico Roos [*]

Department of Knowledge Engineering, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands

ARTICLE INFO

Article history:
Received 13 January 2010
Received in revised form 3 December 2010
Accepted 4 December 2010
Available online 10 December 2010

Keywords:
Model-based diagnosis
Accuracy
Preferences
Confirmation
Simple temporal networks

ABSTRACT

Models used for Model-Based Diagnosis usually assume that observations, and predictions based on the system description are accurate. 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.[1] 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 *maximal-confirmation diagnoses*, is introduced.

This is a revised version of the original paper. Several typing errors, an error in the example presented in Section 6, and several formatting errors of the publisher have been corrected.

1 Introduction

Models used for Model-Based Diagnosis usually assume that observations, and predictions based on the system description are accurate [28, 11, 26, 12, 5, 6, 10]. Here accuracy refers to the uncertainty about the correct value.[2] This assumption is usually not stated explicitly. The assumption 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 $\{negative, positive\}$ or $\{low, high\}$ may sometimes be used [39, 40]. 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 $[lb, ub]$ may be used. Several papers deal with consistency-based diagnosis given inaccurate data [8, 18, 20, 22].

A practical problem in which we encountered the issue of inaccurate data was diagnosis of temporal constraint violations in Air Traffic Control (ATC). The temporal aspects of a plan in ATC can be described by a Simple Temporal Network (STN) [14]. An STN describes activities by start and finish events and the duration of activities by temporal constraints specifying lower and upper bounds on the temporal distance between two events. Figure 1 gives an illustration. In order to apply Model-Based Diagnosis, the temporal constraints are viewed as behavioral constraints of components and the occurrence of events as in- and outputs of components [31].

In this problem domain, generally, predictions about the occurrence of events are less accurate than the observation of these events. For instance, the duration of a fight may vary considerable because of weather conditions and congested airways. The inaccuracy of clocks that are used to register the occurrence of an event can usually be ignored. Nevertheless, observations may still be inaccurate. If we do not monitor constantly the occurrence of an event but check at different time points whether the event has occurred, our observation of the event's occurrence will still be inaccurate.

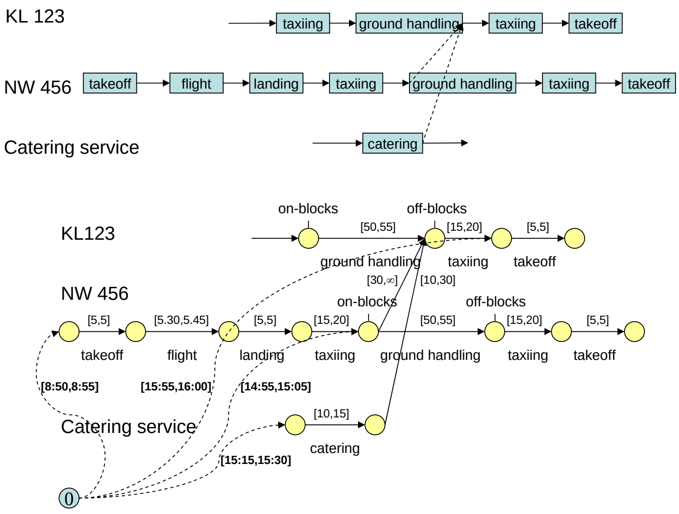

Figure 1: A plan and the corresponding STN.

[*] Tel.: +31 43 3882022.

[1] An initial version of the paper has been published in [29]

[2] Accuracy refers to the fault margin of a value while precision refers to the measurement resolution. A clock may for instance specify the time with a precision in seconds. This value need not be accurate, the clock may be running 2 minutes ahead or lagging 2 minutes behind.

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 to consistency based diagnoses since an abductive diagnosis is not only consistent with the observations made, but also explains the observed outputs of the system under diagnosis. However, in order to apply abductive diagnoses, knowledge of how components of a system may fail is required. If only partial knowledge about the normal and abnormal behavior of components is available, a weaker form of abductive diagnosis, called maximum-informative diagnosis [30, 32] can be used. A maximum-informative diagnosis tries to explain as many observed system outputs as possible. 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 $[4, 8]$. If we observe that the output value actually lays 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. For instance, we may predict that the value of some output lays in the interval $[4, 8]$, while we observe that it lays in the interval $[5, 7]$. Though the observation confirms the prediction based on Δ', Δ' is neither an abductive nor a maximum-informative diagnosis. Abductive and maximum-informative diagnosis require that the predicted value of an output is at least as accurate as the observed value. For instance, Δ' is an abductive diagnosis if we observe that the output value lays in the interval $[3, 9]$, since in that case the predicted interval $[4, 8]$ implies the observed interval $[3, 9]$.

Cordier [7] proposed to adapt the definition of an abductive diagnosis to cope with observations that are more accurate than the predicted output values. This paper, however, proposes a new notion of diagnosis, called *maximal-confirmation* diagnosis. A maximal-confirmation diagnosis is based on measuring to what extent predictions of outputs, given a diagnosis, are confirmed by the observations. A maximal-confirmation diagnosis therefore refines the new definitions of abductive diagnosis proposed by Cordier [7].

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-diagnosis. Section 4 discusses the computational complexity of identifying a maximal-confirmation and a mac-diagnosis. In Section 5, a formal underpinning of the proposed solutions is given. An application of maximal-confirmation diagnosis is described in Section 6. Section 7 describes related work and Section 8 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. There are different ways to describe a system such as Bayesian Belief Networks [38, 23, 43] and Discrete Event Systems [4, 35, 36, 2, 24]. Here, we focus on *classical* Model-Based Diagnosis (MBD) [28, 11, 26, 12, 5, 6, 10], in which the system is described in *first order logic* and either consistency-based reasoning or abductive reasoning is used to identify possible diagnoses.

2.1 Classical Model-Based Diagnosis

MBD [28, 11, 26, 12, 5, 6, 10], 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'.

In this paper we use an abstract description of the system in which we do not specify the components of the system and their behavior. Instead we assume that the system description Sd describes at least the normal behavior of the system and possibly also the abnormal behavior. Moreover, we assume a set of candidate diagnoses $\mathcal{D}$ where each diagnosis $\Delta \in \mathcal{D}$ is a set of proposition giving one possible description of the health modes of the system's components. Finally, we assume a set of possible observation sets $\mathcal{O}$ of the system. An observation set $O \in \mathcal{O}$ is a set of propositions describing the observations. Each observation in an observation set O is a proposition $o(x)$ describing a value of an in- or output x of the system. Notice that we do not consider uncertainty about which in- or output has been observed; i.e., we do not consider observations such as: $o(x) \vee o'(y)$ with $x \neq y$. The triple $Pd = \langle \mathcal{D}, Sd, \mathcal{O} \rangle$ describes our problem domain.

In consistency-based diagnosis there is no need to distinguish between observations of system inputs and observations of system outputs. However, in abductive diagnosis it is important to make this distinction. Therefore, we distinguish two subsets of $\mathcal{O}$, the set of observation sets of system inputs $\mathcal{O}^{in}$ and the set of observation sets of system outputs $\mathcal{O}^{out}$. Note that system inputs are sometimes denoted as the *context* because the system inputs are set externally, for instance by the environment. In case they are set by the environment, observations are needed to determine their values. Without loss of generality, we assume that all known values of system inputs are described by observations of the system inputs.

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[3]:

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

[3] $\Sigma \models \bot$ denotes that *false* is implied by Σ, in other words, that the set of propositions Σ is inconsistent.

Here the candidate diagnosis $\Delta^{nor} \in \mathcal{D}$ 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 diagnosis 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 [28, 11, 10].

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 explain the observed system outputs [26, 5, 27].

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 observations of the system inputs $O^{in} \in \mathcal{O}^{in}$ and system outputs $O^{out} \in \mathcal{O}^{out}$ such that $O = O^{in} \cup O^{out}$.*

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. Suppose Δ is not a consistency-based diagnosis. Then, for *no* semantic interpretation I, $I \models \Delta \cup Sd \cup \mathcal{B} \cup O$. Since Δ is an abductive diagnosis, $\Delta \cup Sd \cup \mathcal{B} \cup O^{in} \not\models \bot$ implies that there is an interpretation I such that $I \models \Delta \cup Sd \cup \mathcal{B} \cup O^{in}$. Moreover, since $\Delta \cup Sd \cup \mathcal{B} \cup O^{in} \models O^{out}$, $I \models O^{out}$. Therefore, $I \models \Delta \cup Sd \cup \mathcal{B} \cup O$. Contradiction. Hence, Δ is also a consistency-based diagnosis.

The converse does not hold. Abductive diagnosis requires knowledge about the correct and incorrect behaviors of components in order to explain the observations made in all circumstances. The descriptions of the incorrect behaviors of components are called: *fault models*.

> We say that *the set of fault models is complete* iff we can make a prediction for any system output given a diagnosis and observations of all system inputs.

In the absence of fault models, we cannot always determine an abductive diagnosis. However, we are still able to determine a consistency-based diagnosis.

Although a consistency-based diagnosis is, in general, not an abductive diagnosis, there is one exception. If a complete set of fault models is available, and if the set of possible values of system in- and outputs do not overlap, then a consistency-based diagnosis is an abductive diagnosis.

> The set of possible observation sets $\mathcal{O}$ does not allow for overlapping values iff for every $O, O' \in \mathcal{O}$ and for every in- or output x such that $o(x) \in O$ and $o'(x) \in O'$: if $\{o(x), o'(x)\} \cup \mathcal{B} \not\models \bot$, then $\mathcal{B} \models o(x) \equiv o'(x)$.

We cannot state that an in- or output has two different values unless the values overlap. Therefore, a set of non-overlapping values must be inconsistent.

Note that in case the set of observation sets $\mathcal{O}$ does not allow for overlapping observations, we cannot use different accuracies to describe a system in- or output. The following three propositions describe overlapping values:

$value(x, [4, 6]), value(x, [1, 8]), value(x, [5, 9])$

Proposition 1 *A consistency-based diagnosis is an abductive diagnosis if the set of fault models is complete, if all system inputs are observed and if the set of possible observation sets $\mathcal{O}$ does not allow for overlapping observations.*

Proof.
Let Δ be a consistency-based diagnosis: $\Delta \cup Sd \cup \mathcal{B} \cup O \not\models \bot$.

Suppose that $\Delta \in \mathcal{D}$ is no abductive diagnosis. Then: $\Delta \cup Sd \cup \mathcal{B} \cup O^{in} \not\models O^{out}$ with $O = O^{in} \cup O^{out}$. Since knowledge about the systems behavior is complete (the set fault models is complete) and since all system inputs are observed, there is an $O' \in \mathcal{O}^{out}$ such that: $\Delta \cup Sd \cup \mathcal{B} \cup O^{in} \models O'$ and O' describes a value for every system output. Clearly, for some output x and $o(x) \in O^{out}$, $o'(x) \in O'$, and $\mathcal{B} \not\models o(x) \equiv o'(x)$. Therefore, since set of observation sets $\mathcal{O}$ does not allow for overlapping values, $\{o(x), o'(x)\} \cup \mathcal{B} \models \bot$. This implies that: $\Delta \cup Sd \cup \mathcal{B} \cup O^{in} \cup O^{out} \models \bot$. Hence, Δ cannot be a consistency-based diagnosis, contradicting our starting point. □

In the literature, often a *'non-overlapping values' assumption* is used without specifying this explicitly.

2.2 Representing inaccurate values

In the examples we have represented inaccuracy by specifying an interval of possible values. This is not the only possibility of specifying inaccuracy. Other possibilities are: (*i*) an average value together with some distance-value, (*ii*) if values are normally distributed, an average value and a standard deviation, (*iii*) some abstract symbolic value such as 'high'and 'low', (*iv*) fuzzy sets, etc. In all cases, the representation of inaccurate values of in- and outputs can be described by propositions.

The propositions describing inaccurate values of in- and outputs can be order with respect to the accuracy of the described value. Propositions about accurate values imply propositions about a less accurate values. For instance, $value(x, [3, 4])$ implies $value(x, [2, 8])$ and $value(x, [-\infty, +\infty])$ given sufficient background knowledge about numbers. We use this property to define an accuracy ordering over the set of observation sets.

Definition 3 *Let $\mathcal{O}$ be the set of possible observation sets and let $O, O' \in \mathcal{O}$ be two observation sets. Moreover, let $\mathcal{B}$ be general background knowledge.*

> *O is at least as accurate as O', denoted by $O \sqsubseteq O'$ iff $O \cup \mathcal{B} \models O'$.*

Note that the accuracy ordering over $\mathcal{O}$ defines a *lattice* with bottom element $\bot$ and top element $\top$. Since observations are described by proportions, given our background knowledge $\mathcal{B}$ describing which observations are possible, the bottom $\bot$ and top $\top$ element of the lattice also correspond to *false* and *true*, respectively.[4]

In order to guarantee the correctness of the maximal-confirmation diagnosis defined in the next section, we must enforce some additional structure on the set of observation sets $\mathcal{O}$. Let $\overline{O} = \bigcup \mathcal{O}$ be all the propositions about in- or outputs. Then we require that for every output x that:

> if $o(x), o'(x) \in \overline{O}$, then $o(x) \wedge o'(x) \in \overline{O}$ and $o(x) \wedge \neg o'(x) \in \overline{O}$.

Moreover, for every $O \subseteq \overline{O}$:

> $O \in \mathcal{O}$ iff for no in- or output x: $\{o(x), o'(x)\} \subseteq O$ and $o(x) \neq o'(x)$.

2.3 *MBD and uncertainty*

Although the logical model used to describe a system does not address uncertainty, it is possible to take into consideration the a priori probabilities of faults before a diagnosis is made and the a posteriori probabilities after a diagnosis is made given the observations [9]. Usually the relation between a logical representation and the probability that a proposition holds, is not specified explicitly. Therefore, we first address this issue. We do not consider a translation of MBD to a probabilistic framework such as BBNs [16]. We only need the standard way of assigning probabilities to logical propositions.

Probability theory assumes a probability distribution over some sample space. In case of propositional and first order logic, the set of semantic interpretations is used as the sample space [34, third edition, Chapter 13.2]. Over the set of interpretations $\mathcal{I}$, we define a probability distributions $p : \mathcal{I} \to [0, 1]$ such that $\sum_{I \in \mathcal{I}} p(I) = 1$. The probability of a proposition can now be defined as the probability of the set of interpretations satisfying the proposition. Formally,

$$P(\varphi) = \sum_{I \in \mathcal{I}, I \models \varphi} p(I)$$

Moreover, given our knowledge Σ describing our diagnostic problem, the a posteriori probability that φ holds is given by:

$$P(\varphi \mid \Sigma) = \frac{\sum_{I \in \mathcal{I}, I \models \varphi, \Sigma} p(I)}{\sum_{I \in \mathcal{I}, I \models \Sigma} p(I)}$$

Note that the denominator normalizes the probability distribution w.r.t. interpretations that satisfy our knowledge Σ.

Generally, we do not know the probability distribution $p : \mathcal{I} \to [0, 1]$; i.e., the available knowledge about the probabilities does not guarantee the existence of a unique probability distribution. Although there is an extensive amount of literature addressing preferred probability distributions among the possible probability distributions (see for instance [1, 19, 21, 33]), we will not consider such preferences. Here, we will only consider conclusions about probabilities of propositions that hold for any probability distribution satisfying the knowledge (and assumptions) about the diagnostic problem.

Given the above described relation between probabilities and propositions, the validity of preferred diagnoses can be proved. Assuming that components fail independently, it is now easy to verify that subset-minimal diagnoses should be preferred if fault probabilities of components are less than 0.5. Let Σ denote all our knowledge about our diagnostic problem and let $Nor(\Delta)$ denote the set of components of the system that behave *normally* according to the candidate diagnosis Δ. Then, $P(\Delta \mid \Sigma) > P(\Delta' \mid \Sigma)$ if Δ and Δ' are diagnoses of the system and $Nor(\Delta') \subset Nor(\Delta)$. Moreover, cardinality-minimum diagnoses should be preferred if fault probabilities of components are very small. That is, $P(\Delta \mid \Sigma) > P(\Delta' \mid \Sigma)$ if Δ and Δ' are diagnoses of the system and $|Nor(\Delta')| < |Nor(\Delta)|$.

3 Inaccurate predictions and observations

The use of inaccurate values implies that the 'non-overlapping values' assumption is invalid. When using inaccurate values we can state for instance that for some in- or output x both the propositions $value(x, [2, 4])$ and $value(x, [3, 6])$ hold. Of course, if both propositions are true, the value of x must lay in the interval $[3, 4]$.

3.1 *Preferences and accuracy*

Giving up the 'non-overlapping values' assumption has no influence on the definition of consistency-based and abductive diagnosis. However, it does influence the preferred diagnoses among the set of consistency-based diagnoses. We will illustrate this with a simple example shown in Figure 2. Suppose that Anna is flying from Paris to Amsterdam. We know that the flight takes 60 to 70 minutes. Moreover, getting from Schiphol airport in Amsterdam to Anna's home takes 30 to 35 minutes. The official boarding time of Anna's plane is scheduled at 9:30. Every aircraft that departs from an airport in Europe is assigned slot for takeoff by EUROCONTROL in Brussel. We know that the assigned slot for takeoff in Paris is from 10:00 till 10:15. If the airplane does not depart in the assigned slot because of an incident such as a passenger no-show, a new slot has to be requested and the next available slot will be from 10:30 till 10:45. This new slot is our fault model for the takeoff-action of the airplane.

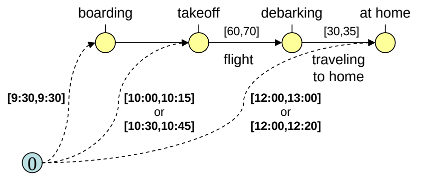

Figure 2: Anna's traveling schedule.

Based on this information we can predict that Anna will arrive at home between 11:30 and 12:00 if no incident occurs. Now suppose that we check just before 12:00 whether Anna is at home and we do it again at 13:00. The first time she has not yet arrived but the second time she has. So, we know that Anna arrived at home between 12:00 and 13:00. Assuming that no incident occurred and that the airplane departed from Paris in the designated slot; i.e., our diagnosis states that everything went as planed, the prediction of Anna arriving at home between 11:30 and 12:00 is consistent with the observation that she arrived at home between 12:00 and 13:00. Hence, this is a consistency-based diagnosis. But is it a probable diagnosis?

Of all the possible takeoff times, flight durations and traveling times from the airport to Anna's home, there is *only one*

[4] Note that the join-operator, which is often denoted by '$\vee$', does not correspond with the logical $\vee$-operator because we do not consider uncertainty about which output has been observed; i.e., we do not consider observation such as: $o(x) \vee o'(y)$ with $x \neq y$.

possible scenario that is consistent with our observation, namely, that the plane's takeoff time is at 10:15, the flight duration is 70 minutes and the travel time to her home is 35 minutes. The probability that this is what has happened is not very high.

Abductive diagnosis gives a different answer to our diagnostic problem. Only an incident that forces the plane to depart in the next available time slot from 10:30 till 10:45 can explain our observation. If the plane departs in the next time slot, Anna will arrive at home between 12:00 and 12:30, which *explains* our observation of arriving at home between 12:00 and 13:00. Note that abductive diagnosis allows for all the uncertainty that is present in problem description. This suggest that the diagnosis in which the plane's uses the takeoff slot from 10:30 till 10:45 is more probable. Therefore, in this example it is reasonable to prefer the abductive diagnosis although it is neither a minimum nor minimal consistency-based diagnosis.

Now suppose that we made a more accurate observation about Anna's arrival at her home. If our second check was not at 13:00 but at 12:20 and if she was at home when we checked at 12:20, then abductive diagnosis cannot explain our observation that Anna arrived at home between 12:00 and 12:20. Although fault models are available, we cannot explain the observation because the observation is more accurate than any prediction we can make. However, the observation *confirms* the prediction of Anna arriving at home between 12:00 and 12:30 if the plane departed at the next time slot because of some incident.

Observations that confirm the predictions based on a diagnosis, maximizes the possible scenarios that are allowed by the uncertainty in the problem description. This suggests that the diagnosis in which an incident causes the plane to use the takeoff slot from 10:30 till 10:45 is to be preferred to the minimum consistency-based diagnosis in which there is no incident. The observation only confirms one of the predicted possible values, namely Anna arriving at home at 12:00, in case of the minimum consistency-based diagnosis. The prediction is only partially confirmed if no incident delaying the takeoff has occurred.

3.2 *Maximal-confirmation diagnosis*

The idea that is put forward in this section is to prefer *maximal-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.

We will use the accuracy ordering on the set of observation sets defined in Subsection 2.2 to give a general definition of maximal-confirmation diagnoses, which is not limited to intervals of values. We say that a set of observations O *strongly* confirms a set of predicted values O' iff $O \sqsubseteq O'$; i.e., the observations O imply the predicted system outputs O'. A set of observations O *weakly* confirms a set of predicted values O' iff $O' \sqsubseteq O$. We can also define a notion of partial confirmation. A set of observations O *partially* confirms a set of predicted values O' iff $\perp \sqsubset O \wedge O'$, $O \not\sqsubseteq O'$ and $O' \not\sqsubseteq O$. Figure 3 gives an illustration using one observed output and three predictions given three diagnoses. The accuracy is described by intervals.

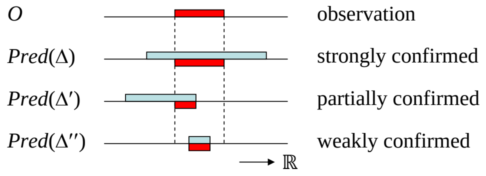

Figure 3: Different confirmation degrees.

It is clear that there are different degrees in which an observation can confirm a predicted value. The accuracy order over the set of possible observation sets $\mathcal{O}$ can be used to order diagnoses with respect to the degree of confirmation. To measure the degree of confirmation, given a diagnosis, we determine all possible values of the system outputs that agree with the observations made. This corresponds to the meet of observations and the predicted output values. Of course, we must use the most accurate predictions for the output values. The most accurate prediction may only be confirmed by some observations while the least accurate prediction is confirmed by any observation. Therefore, the most accurate predictions will give us the most relevant information.

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

where O^{out} is a set of observations and $Pred(\Delta)$ is the set of the most accurate predicted output values of the system.

$$\begin{aligned}
&Pred(\Delta) = O \text{ iff}\\
&\quad O \in \mathcal{O}^{out},\\
&\quad \Delta \cup Sd \cup \mathcal{B} \cup O^{in} \models O, \text{ and}\\
&\quad \text{for no } O' \in \mathcal{O}^{out}\text{: } O' \sqsubset O \text{ and } \Delta \cup Sd \cup \mathcal{B} \cup O^{in} \models O'
\end{aligned}$$

Note that $CD(\Delta) \equiv O^{out}$ if the predictions are strongly confirmed by the observations, and that $CD(\Delta) \equiv Pred(\Delta)$ if Δ is an abductive diagnosis. Also note that if for some output, the observed value does not agree with a predicted value, $CD(\Delta) \equiv \perp$. In that case Δ is *no* consistency-based diagnosis.

The confirmation degree $CD(\cdot)$ can be used to define the *maximal-confirmation* diagnoses. We prefer those diagnoses that maximize the predicted set of possible output values that agree with the observations made.

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

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

1. *$\perp \sqsubset CD(\Delta)$,*
2. *for* no *diagnosis Δ': $CD(\Delta) \sqsubset CD(\Delta')$.*

Note that an abductive diagnosis need not be a maximal-confirmation diagnosis, or vice versa. A maximal-confirmation diagnosis is however always a consistency-based diagnosis.

Proposition 2 *Let Δ be a maximal-confirmation diagnosis of the diagnosis problem $P = (Pd, O)$.*

Then Δ is also a consistency-based diagnosis of diagnosis problem $P = (Pd, O)$.

Proof. Suppose that Δ is no consistency-based diagnosis. Then $\Delta \cup Sd \cup \mathcal{B} \cup O^{in} \cup O^{out} \models \perp$. This implies that $\Delta \cup Sd \cup \mathcal{B} \cup O^{in} \models \bigvee_{o(x) \in O^{out}} \neg o(x)$. Hence, for some $o(x) \in O^{out}$, $\Delta \cup Sd \cup \mathcal{B} \cup O^{in} \models \neg o(x)$.

Let $o'(x) \in Pred(\Delta)$. Since $o'(x) \wedge \neg o(x) \in \overline{O}$, $\Theta = (Pred(\Delta) - \{o'(x)\} \cup \{o'(x) \wedge \neg o(x)\}) \in \mathcal{O}$. Since $\perp \sqsubset Pred(\Delta)$, $\Theta \sqsubset Pred(\Delta)$. Moreover, since $\Delta \cup Sd \cup \mathcal{B} \cup O^{in} \models \neg o(x)$, $\Delta \cup Sd \cup \mathcal{B} \cup O^{in} \models \Theta$. Hence, $Pred(\Delta)$ is not the most accurate prediction of the output values. Contradiction. □

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. The question is whether diagnoses that are strongly confirmed by the observations are all equally preferred. To answer this question, consider again the plan for Anna's travel from Paris to her home. Suppose that we have a second fault model for the takeoff. The takeoff may be postponed for an unknown period of time if the preflight check fails. So, in case of a preflight check failure, Anna may arrive at home between 11:30 and ∞. Clearly, our observation that Anna arrived at home between 12:00 and 12:20 also confirms the diagnosis which the preflight check failed as is illustrated by Figure 4. Both diagnoses, the diagnosis in which an incident causes the plane to use the next takeoff slot and the diagnosis in which a preflight check fails, are strongly confirmed by our observation that Anna arrived at home between 12:00 and 12:20. Therefore, both diagnoses have the same confirmation degree. The difference between the two diagnoses is that the prediction based on the diagnosis in which an incident causes the plane to use the next takeoff slot is more accurate. This diagnosis can therefore more easily be disconfirmed. In other words, observing that Anna arrived at home between 20:10 and 20:50 does not confirm an incident that causes the plane to use the next takeoff slot but still strongly confirms a preflight check failure. This clearly shows that there is a difference between the two diagnoses. We therefore propose a second ordering principle, preferring diagnoses that give the most accurate predictions. Together, this results in preferring *maximal accuracy and confirmation* (mac-) diagnoses.

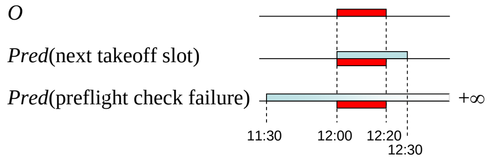

Figure 4: Strongly confirmed, different accuracies.

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 Computational complexity

Maximal-confirmation diagnosis is a special case of consistency-based diagnosis. Friedrich et al. [17] proved for consistency-based diagnosis that deciding whether there exists a diagnosis for a diagnosis problem is already NP-complete if fault models are introduced in a Horn clause system description. Since the existence of a maximal-confirmation diagnosis implies the existence of consistency-based diagnosis, finding a maximal-confirmation diagnosis is NP-hard in general.

Abductive diagnosis is a special case of maximum-confirmation diagnosis. Therefore, in the worst case, identifying a maximal-confirmation diagnosis is at least as hard as abduction. Bylander et al. [3] have determined the computational complexity of several classes of abduction problems. Relevant for us are the *independent*, *monotonic* and *incompatible* abduction problems. An abduction problem is an independent abduction problem iff the set of hypotheses H explains the same data as the union of the data explained by each individual hypothesis $h \in H$. An abduction problem is monotonic iff a larger set of hypotheses explains at least the same data. An abduction problem is an incompatible abduction problem iff some combination of hypotheses is not allowed.

We can establish a relation between maximal-confirmation diagnosis and abduction by giving an new interpretation to the function $e(H) = D$, which denotes that the set of hypotheses H explains the data D [3]. We replace the set of hypotheses by a diagnosis Δ and the data by a the set of observations. Now, the function $e(\Delta) = O$ denotes that the predictions based on the diagnosis Δ are partially confirmed by the observations O. Given this interpretation, the results and algorithms for classes of *independent*, *monotonic*, and *incompatible* abduction problems also apply to the corresponding classes of confirmation diagnosis problems. Identifying the *best explanation* of an abduction problem corresponds to identifying a *maximum-confirmation* and *mac*-diagnosis. Three other classes identified by Bylander et al., namely cancelation, the best small plausibility criterion, and the ordered abduction problems have no corresponding confirmation diagnosis problems. The inaccuracy caused by one element of a diagnosis will not cancel the inaccuracy caused by another element. Moreover, the plausibility of a diagnosis (hypotheses) depends on the confirmation degree and not on predefined preferences over hypotheses.

To summarize, the paper of Bylander et al. [3] implies the following results:

Class of problems	Condition to achieve		
	Finding all confirmation diagnoses	Finding a confirmation diagnosis	Finding a maximal confirmation or mac-diagnosis
Independent	NP	P	?
Monotonic	NP	P	?
Incompatible	NP	NP	NP

The table shows that generally, finding a maximal-confirmation diagnosis is an NP-hard problem. It also shows that Bylander et al. do not specify a general result in case diagnoses are ordered. The computational complexity of ordered diagnoses depends on the fault models.

Either and Gottlob [15] studied the complexity of logic-based abduction. Since we placed no restrictions on the expressiveness of propositional or first-order logic w.r.t. the specification of the system description, their complexity results are also relevant for identifying a maximum-confirmation and mac-diagnosis. Either and Gottlob show that in case an unrestricted propositional theory is used for specifying the system description, the computational complexity lays at the second level of the polynomial hierarchy. The use of priorities may raise the computational complexity to the third level of the polynomial hierarchy.

Fortunately, we normally do not need an unrestricted propositional theory for describing a system. Often the above mentioned function $e(\Delta) = O$ specified by the system description can be computed in polynomial time. In that case, the the complexity results of Either and Gottlob do not apply. Hence, the complexity result of Bylander et al. more relevant for maximal-confirmation and mac-diagnosis.

The complexity of identifying a maximal-confirmation or a mac-diagnosis might be reduced by offline compilation of the system description. For instance Ordered Binary Decision Diagrams might be exploited for the efficient computation of maximal-confirmation and a mac-diagnoses [41]. Further research in this area is required.

There are also special cases in which a maximal-confirmation and a mac-diagnosis can be identified in polynomial time. The example in Section 6 gives an illustration.

5 A probabilistic justification of preferred diagnoses

One of the motivations of preferring maximal-confirmation diagnoses is because they are expected to be more probable than other diagnoses. We will now give evidence for this preference.

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;
- components fail independently;
- the predicted value of a system output is either equal to an observation or is unknown.

Using inaccurate values, the third 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 first part of the example in Subsection 3.1, 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 formally[5]:

$$P(\Delta \mid O) = P(O \mid \Delta) \cdot \frac{P(\Delta)}{P(O)} \tag{1}$$

$$= P(O \mid \mathit{Pred}(\Delta)) \cdot P(\mathit{Pred}(\Delta) \mid \Delta) \cdot \frac{P(\Delta)}{P(O)} \tag{2}$$

$$= P(\mathit{Pred}(\Delta) \mid O) \cdot \frac{P(O)}{P(\mathit{Pred}(\Delta))} \cdot P(\mathit{Pred}(\Delta) \mid \Delta) \cdot \frac{P(\Delta)}{P(O)} \tag{3}$$

$$= P(\mathit{Pred}(\Delta) \mid O) \cdot P(\mathit{Pred}(\Delta) \mid \Delta) \cdot \frac{P(\Delta)}{P(\mathit{Pred}(\Delta))} \tag{4}$$

$$= P(CD(\Delta) \mid O) \cdot P(\Delta \mid \mathit{Pred}(\Delta)) \tag{5}$$

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

1. The conditional probability $P(O \mid \mathit{Pred}(\Delta))$ in equation 2 is conditionally independent of the diagnosis Δ.
2. The conditional probability $P(\mathit{Pred}(\Delta) \mid O)$ in equations 3 and 4 is equal to *confirmation probability* $P(CD(\Delta) \mid O)$ in equation 5 since $\mathit{Pred}(\Delta) \wedge O^{out} \equiv CD(\Delta)$.
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. The *explanation probability* $P(\Delta \mid \mathit{Pred}(\Delta))$ in equation 5 expresses the conditional probability that Δ is a diagnosis given the system outputs $\mathit{Pred}(\Delta)$ that can be *explained* by Δ. If the observed and predicted values are accurate, the diagnosis is an abductive diagnosis and the explanation probability is the probability that the abductive diagnosis is correct.
6. Since $\mathit{Pred}(\Delta)$ describes the system outputs given a diagnosis Δ, clearly, $P(\mathit{Pred}(\Delta) \mid \Delta) = 1$. Hence, the explanation probability can be written as:

   $$P(\Delta \mid \mathit{Pred}(\Delta)) = \frac{P(\Delta)}{P(\mathit{Pred}(\Delta))} \tag{6}$$
7. Let $\neg\Delta$ denote the diagnosis covering all diagnoses that are mutually independent of Δ. Then:

   $$P(\mathit{Pred}(\Delta)) = P(\Delta) + P(\mathit{Pred}(\Delta) \mid \neg\Delta) \cdot P(\neg\Delta) \tag{7}$$

   Equations 6 and 7 imply that the explanation probability $P(\Delta \mid \mathit{Pred}(\Delta))$ is equal to 1 if there is no diagnosis Δ' that is mutually independent of Δ such that $P(\mathit{Pred}(\Delta) \mid \Delta') > 0$. This is the case if Δ is the least accurate diagnosis which results in predicting the value *unknown* for every system output. Moreover, this may be the case if the diagnosis Δ is sufficiently accurate and no other diagnosis can (partially) explain $\mathit{Pred}(\Delta)$.

   Note that a less accurate diagnosis results in a less accurate prediction $\mathit{Pred}(\Delta)$. This will increase the chance that other diagnoses can (partially) explain $\mathit{Pred}(\Delta)$ and thereby decrease $P(\Delta \mid \mathit{Pred}(\Delta))$. However, if diagnoses become less and less accurate, then, eventually, $P(\Delta \mid \mathit{Pred}(\Delta))$ will start to increase and will become equal to 1 for the least accurate diagnosis.

5.1 Discussion

To maximize the probability of a diagnosis Δ given the observations O; i.e. maximizing $P(\Delta \mid O)$, we have to maximize both the *confirmation probability* and the *explanation probability*. By preferring strongly confirmed diagnoses, we maximize $P(CD(\Delta) \mid O)$.

The explanation probability Δ is maximal if there is no mutually independent diagnosis Δ' that can partially explain $\mathit{Pred}(\Delta)$. This is the case if Δ is the least accurate diagnosis. Moreover, it might be the case if Δ is a sufficiently accurate diagnosis. The former diagnosis, which always exists, does not provide us with any useful information. The latter diagnosis does provide useful information, but may not exist.

The usefulness (or *utility*) of a diagnosis increases with its accuracy. Without numerical utility values, we cannot make exact statements about which diagnoses to prefer. However, a *maximal accuracy and confirmation* diagnoses seems to balance the utility and probability of a diagnosis. In other words, mac-diagnoses seem to maximize the expected utility.

If $u(\Delta)$ is the utility of the diagnosis Δ, then the expected utility is:

$$\begin{aligned} Eu(\Delta) &= P(\Delta \mid O) \cdot u(\Delta) \\ &= P(CD(\Delta) \mid O) \cdot P(\Delta \mid \mathit{Pred}(\Delta)) \cdot u(\Delta) \end{aligned}$$

[5] Note that every probability below is conditionalized w.r.t. the knowledge $Sd \cup \mathcal{B}$. To improve readability, the knowledge $Sd \cup \mathcal{B}$ is not mentioned explicitly. We write for instance $P(\Delta \mid O)$ instead of $P(\Delta \mid O, Sd \cup \mathcal{B})$.

The term $P(\Delta \mid \mathit{Pred}(\Delta)) \cdot u(\Delta)$ is fixed given a diagnosis Δ. It only depends on the structure of the system to be diagnosed. The utility of a diagnosis increases with the accuracy of the diagnosis. We assume that the increase of the utility is the dominant factor and therefore that $P(\Delta \mid \mathit{Pred}(\Delta)) \cdot u(\Delta)$ is a monotonic function in the accuracy of Δ. This gives us an incentive to prefer more accurate diagnoses.

The confirmation probability $P(CD(\Delta) \mid O)$ is an independent factor in the expected utility of Δ. The confirmation probability monotonically decreases in the accuracy of Δ if the observations weakly or partially confirm the diagnosis. Since the inaccuracy of the observations will generally be small with respect to the whole range of observable values, we assume that the decrease of $P(CD(\Delta) \mid O)$ in the accuracy of Δ dominates the increase of $P(\Delta \mid \mathit{Pred}(\Delta)) \cdot u(\Delta)$. Therefore, mac-diagnoses maximize the expected utility of a diagnosis, which justifies our preference for mac-diagnoses.

5.2 Additional preference

Do minimum or minimal diagnoses play no role in diagnosis problems with inaccurate observations?

To answer this question, suppose that $\Delta_1, \dots, \Delta_k$ are mac-diagnoses Moreover, suppose that these mac-diagnoses exactly confirm the observations O; i.e., $\mathit{Pred}(\Delta_i) = O$. In other words, $\Delta_1, \dots, \Delta_k$ are the least accurate abductive diagnoses. Then the confirmation probability is equal to 1 and the explanation probability can written as:

$$P(\Delta_i \mid \mathit{Pred}(\Delta_i)) = \frac{P(\Delta_i)}{\sum_{j=1}^{k} P(\Delta_j)} = \alpha \cdot P(\Delta_i)$$

since $P(\mathit{Pred}(\Delta_i) \mid \Delta_j) = P(O \mid \Delta_j) = P(\mathit{Pred}(\Delta_j) \mid \Delta_j) = 1$. Here, α is a normalization factor, which is the same for every diagnosis Δ_i with $i \in \{1, \dots, k\}$. Therefore, the mac-diagnoses can be ordered with respect to their a priori probability. Hence, we may prefer the minimum or minimal diagnoses among the mac-diagnoses.

6 An application

This section illustrates the use of mac-diagnoses in the identification of constraint violations in a Simple Temporal Networks (STN) [14]. This illustration generalizes the maximum confirmation diagnoses of STNs proposed by Roos & Witteveen [31].

6.1 Simple Temporal Networks

An STN $(\mathcal{E}, \mathcal{C})$ describes a plan and its schedule by a set of events $\mathcal{E}$ and a set of constraints $\mathcal{C}$ over the events. Events denote such things as the start $start(s)$ of a plan step s and the finish $\mathit{finish}(s)$ of s. The constraints are used to specify the durations of plan steps, the precedence relations between plan steps, and the plan's schedule. It is also possible to specify requirements such as the requirement that a plan step that must start within x minutes after the finish of its preceding plan step.

To describe a constraint, we associate a variable t_e with each event $e \in \mathcal{E}$. These variables take values in some dense time domain $Time$. We assume $Time$ to be a total order with element 0 and maximum element ∞. A constraint $c \in \mathcal{C}$ specifies the allowed temporal difference between two events: $lb \le t_e - t_{e'} \le ub$ where e and e' are events in $\mathcal{E}$, $lb, ub \in Time$ and $0 \le lb \le ub$.

Relating an STN to a traditional plan description $Pl = (S, \prec)$, the *duration* of a plan step $s \in S$ is described by $0 < lb \le t_{\mathit{finish}(s)} - t_{start(s)} \le ub$. A *precedence constraint* $s \prec s'$ is described by $lb \le t_{start(s')} - t_{\mathit{finish}(s)} \le ub$. Note that in the standard interpretation of a precedence constraint, $lb = 0$ and $ub = \infty$.

A *schedule* is a placement of events on the timeline. To describe a schedule we need a special event '0' marking start of the timeline; i.e., $t_0 = 0$. This enables us to schedule the period in which an event $e \in \mathcal{E}$ should take place: $lb \le t_e - t_0 \le ub$; i.e.: $lb \le t_e \le ub$.

6.2 Constraint violations and candidate diagnoses

Unforseen circumstances may cause the violation of constraints of an STN. To indicate the degree in which a constraint $c = (lb \le t_e - t_{e'} \le ub) \in \mathcal{C}$ is violated, we replace the constraint c by:

$$lb \le t_e - t_{e'} - \delta_c \le ub \tag{8}$$

Note that $\delta_c = 0$ corresponds with the normal 'behavior' of a constraint while each $\delta_c \neq 0$ corresponds with a different fault model of the constraint.

A diagnosis specifies exactly whether and how much each constraint is violated. If a constraint c is violated, then there often exists an interval $[l, u]$ such that for each $v \in [l, u]$ there is a maximal-confirmation diagnosis specifying $\delta_c = v$. We therefore introduce a generalized diagnosis in which we can specify multiple constraint violations: $(\delta_c \in [l, u]) \in \Delta$. This generalized diagnosis specifies a cartesian product of *maximum-confirmation diagnoses*[6]:

$$\Delta' \in \left\{ \bigcup_{c \in \mathcal{C}} \{\delta_c = v\} \,\middle|\, (\delta_c \in [l, u]) \in \Delta, v \in [l, u] \right\}$$

Maximum confirmation diagnoses cannot decrease the accuracy of the predictions. Therefore, maximal-confirmation diagnoses will always be mac-diagnoses. This also implies that in case a prediction is more accurate than an observation involving the same events, the observation can only weakly confirm the prediction given a maximal-confirmation diagnosis. We could modify the Equation 8 to decrease the accuracy of a constraint. Because we introduced generalized diagnoses, we can express multiple accurate constraint violations. This makes it unnecessary to be able to express a single inaccurate constraint violation.

Below, a candidate diagnosis Δ will be a generalized diagnosis containing a proposition $\delta_c \in [l, u]$ for every constraint $c \in \mathcal{C}$. Δ^{nor} will be used to denote the diagnosis in which no constraint is violated. So, for every constraint $c \in \mathcal{C}$, $(\delta_c \in [0, 0]) \in \Delta^{nor}$.

6.3 Observations

During the execution of a plan *observations* can be made. These observations may pertain to the time difference observed between two events e and e' as specified in the plan or may pertain to the time at which a certain event $e \in \mathcal{E}$ takes place.

We assume that the first type of observation is described by some constraint $a \le t_e - t_{e'} \le b$ indicating that we have observed that event e occurred at least a time steps, but within b time steps after e'.

[6] The use of intervals to describe multiple constraint violations is a generalization w.r.t. [31].

The second type of observation is given by a constraint $a \leq t_e - t_0 \leq b$ indicating that e occurred after a time units but before b time units have been passed (after the occurrence of the time reference event '0'). The set of observations containing these constraints is denoted by O.

6.4 Semantics

The constraints of an STN place restrictions on the way a plan may be executed; the *execution schedule*. An execution schedule for the set of events $\mathcal{E}$ of an STN $(\mathcal{E}, \mathcal{C})$ is a function $\sigma : \mathcal{E} \to Time$.[7] We say that an execution schedule σ *satisfies* the constraints $\mathcal{C} \cup O$ given the generalized diagnosis Δ, denoted by $\sigma \models \mathcal{C} \cup O \cup \Delta$, iff

- for every $(lb \leq t_e - t_{e'} \leq ub) \in O$, $lb \leq \sigma(e) - \sigma(e') \leq ub$ holds,
- for every $(lb \leq t_e - t_{e'} \leq ub) \in \mathcal{C}$ and for every $v \in [l, u]$ with $(\delta_c = [l, u]) \in \Delta$, $lb \leq \sigma(e) - \sigma(e') - v \leq ub$ holds.

The identification of an allowable execution schedule is called a Simple Temporal Problem (STP) [14].

We say that a constraint $c : a \leq t_e - t_{e'} \leq b$ *is entailed by* a set of constraints C, denoted by $C \models c$, iff every allowable schedule for C satisfies c.

Given a constraint $c = (a \leq t_e - t_{e'} \leq b)$ we say that $c' = (a' \leq t_e - t_{e'} \leq b')$ is a *tightening* of c, denoted by $c' \models c$ iff $a \leq a' \leq b' \leq b$. There is a sound and complete *derivation procedure* ($\vdash$) for determining the *most tightened* constraint $c = (a \leq t_e - t_{e'} \leq b)$ entailed by a set of constraints C [37]. That is: $C \vdash c$ iff $C \models c$. We can derive in polynomial time the most tightened constraint between all pairs of events [42, 25].

6.5 Diagnosis

If an STN $(\mathcal{E}, \mathcal{C})$ is not compatible with a set O of observations given the diagnosis Δ^{nor}; i.e., no execution schedule exist for $\mathcal{C} \cup O$ given Δ^{nor}, some constraints in $\mathcal{C}$ must have been violated. We need to identify a new diagnosis Δ to restore the compatibility between plan and observations. By preferring maximal-confirmation diagnosis we maximize the probability of a diagnosis.

To identify violated temporal constraints of an STN $(\mathcal{E}, \mathcal{C})$, we view the constraints $\mathcal{C}$ as components of the system under diagnosis, and the events $\mathcal{E}$ as the in- and outputs of the components. An event $e \in \mathcal{E}$ may be part of more than one constraint. Each of these constraints enforces a restriction on the occurrence of e with respect to some event e'; i.e., we can derive a constraint between e and e' using different paths through the constraint graph. It is important to note that each constraint derived using a path from e' to e determines one or more diagnoses involving the constraints on the path. Here, we assume that *at most one constraint on each path between pairs of observed events will be violated.*

Let $(l \leq t_e - t_{e'} \leq u) \in O$ be an observation, let the sequence of constraints $c_1, \ldots, c_k$ be an undirected path from event e' to event e in the constraint graph defined by the constraints $\mathcal{C}$, and let $(lb \leq t_e - t_{e'} \leq ub)$ be the most tightened constraint that can be predicted for this path from e' to e. From a temporal perspective, a constraint c_i on the path points towards e or towards e'. Let p^+ denote the constraints in $\{c_1, \ldots, c_k\}$ pointing to e and p^- the remaining constraints in $\{c_1, \ldots, c_k\}$, which point towards e'. Since we assume that only one constraint $c = (x \leq t_{e''} - t_{e'''} \leq y)$ on the path is violated, if $c \in p^+$ and if the observation is more accurate than the prediction, then we can easily show that $\delta_c = v$ is an element of a maximal-confirmation diagnosis for every $v \in [u - ub, l - lb]$. Moreover, we can show that $\delta_c = v$ is an element of a maximal-confirmation diagnosis for every $v \in [l - lb, u - ub]$ if the prediction is more accurate than the observation and if there is *only one path* from e' to e. Note that the restriction of only one path is necessary since multiple paths will result in dependencies between the paths if the observation weakly confirms the prediction given a diagnosis. Therefore, we assume that *there is only one path between the events mentioned in an observation in case the prediction along the path is more accurate than the observation.* Under this assumption,

$$\delta_c \in [\min(l - lb, u - ub), \max(l - lb, u - ub)]$$

is an element of the generalized diagnosis Δ if $c \in p^+$. If $c \in p^-$ however, then under the same assumption,

$$\delta_c \in [-\max(l - lb, u - ub), -\min(l - lb, u - ub)]$$

is an element of a generalized diagnosis Δ. Moreover, for every constraint $c' \in \{c_1, \ldots, c_k\} - \{c\}$ on the path, $(\delta_c \in [0, 0]) \in \Delta$ must hold.

Paths between observations $(l \leq t_e - t_{e'} \leq u) \in O$ may partially overlap. Our requirement that at most one constraint on a path between a pair of events mentioned in an observation may be violated implies that a constraint that is on multiple paths must ensure a maximal-confirmation diagnosis for each path on which it occurs. Let $(\delta_c \in [l_1, u_1]), \ldots, (\delta_c \in [l_n, u_n])$ be the requirements of the paths on which the constraint c occurs. Since for each value $v \in [l_i, u_i]$, $\delta = v$ is a maximal-confirmation diagnosis for the path determining $\delta_c \in [l_i, u_i]$, we should use the intersection of all intervals $[l_i, u_i]$. So,

$$\delta_c \in [\max\{l_1, \ldots, l_n\}, \min\{u_1, \ldots, u_n\}]$$

provided that $\max\{l_1, \ldots, l_n\} \leq \min\{u_1, \ldots, u_n\}$. If, however, $\max\{l_1, \ldots, l_n\} > \min\{u_1, \ldots, u_n\}$, then the constraint c cannot be violated in a maximal-confirmation diagnosis given our assumption that at most one constraint on each path is violated.

It is not difficult to create constraint graphs in which we have an exponential number of paths between two events. Just consider a constraint graph that forms a clique. In real applications, the number of paths that we must consider is usually limited. Moreover, there always exist maximal-confirmation diagnoses that can be determined in polynomial time. For instance, for each observation $(l \leq t_e - t_{e'} \leq u) \in O$, mark every constraint $(x \leq t_e - t_{e''} \leq y) \in \mathcal{C}$ as being violated if necessary. Given these violated constraints, we do not need to consider every path from event e' to e. It suffices to predict the most tightened constraint from e' to e'', which can be determined in polynomial time [42, 25].

6.6 An example

Consider the plan in Figure 5 together with the observations: 11:05 $\leq t_5 - t_0 \leq$ 11:20 and 10:46 $\leq t_6 - t_0 \leq 11 : 00$. If the

[7] Note that an execution schedule differs from scheduling constraints. An execution schedule is a semantic notion similar to an interpretation of first order logic. It describes when events are actually executed.

plan is executed normally, then the constraints entail: $10{:}35 \leq t_5 - t_0 \leq 11{:}03$ and $10{:}27 \leq t_6 - t_0 \leq \infty$, of which the former is inconsistent with the corresponding observation.

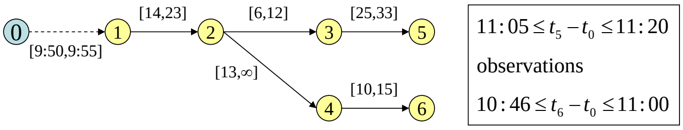

Figure 5: Maximal accuracy and confirmation diagnosis.

In this example, we have the requirement $\delta_c \in [17, 30]$ for a constraint c on the path from event 0 to event 5, and the requirement $\delta_c \in [-\infty, 19]$[8] for a constraint c on the path from event 0 to event 6. Using the requirements, we can determine two minimum (single fault) maximal-confirmation diagnoses and another four minimal (double fault) maximal-confirmation diagnoses.

- $\Delta_1 = \{\delta_{1,0} \in [17,19], \delta_{2,1} = \delta_{3,2} = \delta_{5,3} = \delta_{4,2} = \delta_{6,4} = [0,0]\}$
- $\Delta_2 = \{\delta_{2,1} \in [17,19], \delta_{1,0} = \delta_{3,2} = \delta_{5,3} = \delta_{4,2} = \delta_{6,4} = [0,0]\}$
- $\Delta_3 = \{\delta_{3,2} \in [17,30], \delta_{4,2} \in [-\infty,19], \delta_{1,0} = \delta_{2,1} = \delta_{5,3} = \delta_{6,4} = [0,0]\}$
- $\Delta_4 = \{\delta_{3,2} \in [17,30], \delta_{6,4} \in [-\infty,19], \delta_{1,0} = \delta_{2,1} = \delta_{5,3} = \delta_{4,2} = [0,0]\}$
- $\Delta_5 = \{\delta_{5,3} \in [17,30], \delta_{4,2} \in [-\infty,19], \delta_{1,0} = \delta_{2,1} = \delta_{3,2} = \delta_{6,4} = [0,0]\}$
- $\Delta_6 = \{\delta_{5,3} \in [17,30], \delta_{6,4} \in [-\infty,19], \delta_{1,0} = \delta_{2,1} = \delta_{3,2} = \delta_{4,2} = [0,0]\}$

Note that in diagnoses $\Delta_3, \ldots, \Delta_6$ we could also assume that the constraints $c_{4,2}$ and $c_{6,4}$ are not violated; i.e., $\delta_{4,2} = \delta_{6,4} = [0,0]$.

7 Related work

The use of inaccurate values in diagnosis is related to, but differs from the use of value abstraction [39, 40] and domain abstraction [13]. 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, [8, 20, 18]. 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 [22].

Cordier [7] has addressed consequences of using inaccurate values for abductive diagnosis. The 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 the diagnoses.

The idea of ordering diagnoses w.r.t. the degree of confirmation was first proposed by Roos and Witteveen [31]. They describe diagnosis of a Simple Temporal Network [14], 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.

Diagnosis using Discrete Event Systems [4, 35, 36, 2, 24] and Bayesian Belief Networks [38, 23, 43] are forms of abductive diagnosis. Therefore, one could apply maximal-confirmation and mac-diagnosis to DESs and BNNs. In a BBN, system outputs are represented by variables. The domains of these variables should be extended to enable the representation of inaccurate values. In a DES, system outputs are observable events generated by the system. Inaccuracy in this context means uncertainty about which event has been observed. Hence, we would need a representation for this form of uncertainty.

BNNs offer the possibility to encode the relation between an inaccurate prediction and an observation using a conditional probability. Although we may use maximal-confirmation diagnosis in BBNs, if we can determine the conditional probability of some observation given an (inaccurate) prediction, there is no need for considering maximal-confirmation or mac-diagnosis.

8 Conclusion

This paper studies the consequences of using inaccurate data in *classical* Model-Based Diagnosis (MBD). Models used for Model-Based Diagnosis usually assume that observations and predictions based on the system description are accurate. 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, minimum, abductive 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 diagnosis of Simple Temporal Networks is discussed.

Acknowledgement I thank the anonymous reviewers for their many valuable suggestions.

[8] Unrestricted negative values in a diagnosis are usually not possible in real application domains. In this example we ignore this issue.

References

[1] F. Bacchus, A. J. Grove, J. Y. Halpern, and D. Koller. From statistical knowledge bases to degrees of belief. *Artificial Intelligence*, 87:75–143, 1996.

[2] P. Baroni, G. Lamperti, P. Poglianob, and M. Zanella. Diagnosis of large active systems. *Artificial Intelligence*, (110):135–183, 1999.

[3] Tom Bylander, Dean Allemang, Michael C. Tanner, and John R. Josephson. The computational complexity of abduction. *Artificial Intelligence*, 49(1-3):25–60, 1991.

[4] C. G. Cassandras and S. Lafortune. *Introduction to Discrete Event Systems*. Kluwer Academic Publishers, 1999.

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

[6] L. Console and P. Torasso. A spectrum of logical definitions of model-based diagnosis. *Computational Intelligence*, 7:133–141, 1991.

[7] 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.

[8] P. Dague, O. Jehl, P. Deveès, P. Luciani, and P. Taillibert. When oscillators stop oscillating. In W. Hamscher, L. Console, and J. de Kleer, editors, *Readings in Model-Based Diagnosis*, pages 235–241. 1992.

[9] J. de Kleer. Focusing on probable diagnosis. In *AAAI 91*, pages 842–848, 1991.

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

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

[12] J. de Kleer and B. C. Williams. Diagnosing with behaviour modes. In *IJCAI 89*, pages 104–109, 1989.

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

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

[15] Thomas Eiter and Georg Gottlob. The complexity of logic-based abduction. *Journal of the ACM*, 42(1):3–42, 1995.

[16] Ildikó Flesch, Peter Lucas, and Theo van der Weide. Conflict-based diagnosis: Adding uncertainty to model-based diagnosis. In *IJCAI 2007*, pages 380–385, 2007.

[17] G. Friedrich, G. Gottlob, and W. Nejdl. Physical impossibility instead of fault models. In *AAAI 90*, pages 331–336, 1990.

[18] D. J. Goldstone. Controlling inequality reasoning in a tms-based analog diagnosis system. In W. Hamscher, L. Console, and J. de Kleer, editors, *Readings in model-based diagnosis*, pages 206–211. 1992.

[19] A. J. Grove, J. Y. Halpern, and D. Koller. Random worlds and maximum entropy. *Journal of Artificial Intelligence Research*, 2:33–88, 1994.

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

[21] H. E. Kyburg. Combinatorial semantics: semantics for frequent validity. *Computational Intelligence*, 13:215–257, 1997.

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

[23] J. Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, San Mateo, California, 1988.

[24] Y. Pencolé and M. Cordier. A formal framework for the decentralised diagnosis of large scale discrete event systems and its application to telecommunication networks. *Artificial Intelligence*, 164:121–170, 2005.

[25] Léon Planken, Mathijs de Weerdt, and Roman van der Krogt. P^{3}C: A new algorithm for the simple temporal problem. In *ICAPS 2008*, pages 256–263, 2008.

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

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

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

[29] N. Roos. Preferring maximum confirmation diagnoses. In *BNAIC 2009*, 2009.

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

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

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

[33] Nico Roos. How to reason with uncertain knowledge. In B. Bouchon-Meurier, R. R. Yager, and L. A. Zadeh, editors, *Uncertainty in knowledge bases, IPMU'90*, pages 403–412. Springer-Verlag, 1991.

[34] Stuart Russell and Peter Norvig. *Artificial Intelligence, a modern approach (third edition)*. Pearson Education, Inc., 2010.

[35] M. Sampath, R. Sengupta, S. Lafortune, K. Sinnamohideen, and D. Teneketzis. Diagnosibility of discrete event systems. *IEEE Transactions on Automatic Control*, 40:1555–1575, 1995.

[36] M. Sampath, R. Sengupta, S. Lafortune, K. Sinnamohideen, and D. Teneketzis. Failure diagnosis using discrete event models. *IEEE Transactions on Control Systems Technology*, 4:105–124, 1996.

[37] Eddie Schwalb and Rina Dechter. Processing disjunctions in temporal constraint networks. *Artificial Intelligence*, 93(1-2):29 – 61, 1997.

[38] D.J. Spiegelhalter. Probabilistic reasoning in predictive expert systems. In *Uncertainty in Artificial Intelligence*, pages 47–67. 1986.

[39] P. Struss. What's in SD?: Towards a theory of modeling for diagnosis. In W. Hamscher, L. Console, and J. de Kleer, editors, *Readings in model-based diagnosis*, pages 419–449. 1992.

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

[41] G. Tota and P. Torasso. On the use of OBDDs in model-based diagnosis: An approach based on the partition of the model. *Knowledge-Based Systems*, 19:316–323, 2006.

[42] L. Xu and B. Y. Choueiry. A new efficient algorithm for solving the simple temporal problem. In *Proc. of the 10th Int. Symp. on Temporal Representation and Reasoning and 4th Int. Conf. on Temporal Logic*, page 210220. IEEE Computer Society, 2003.

[43] N. L. Zhang and D. Poole. Exploiting causal independence in bayesian network inference. *Journal of Artificial Intelligence Research*, 5:301–328, 1996.