# No Faults in Structure?
# How to Diagnose Hidden Interactions

Claudia Bottcher*
LIPN - university Paris Nord
Av. J.B. Clement
93430 Villetaneuse
France
email:claudia@u07.univ-parisl3.frral5

## Abstract

A model based diagnosis procedure traces connections between components only where these are provided explicitly in the system description. Consequently structure faults fall between the meshes. This problem has been known since research started in this field ([Davis 84]), but no general solution has been presented so far. We present a procedure to diagnose structure faults, based on a scheme to detect hidden interactions guided by the observation that structure faults lead to discrepancies in apparently unrelated areas and which in contrast to [Preist. Welham 90] modifies the system description dynamically. Like Davis' approach the one presented in the paper is based on the principle that an interaction can occur only where components are adjacent in some way ([Davis 84]). Unlike Davis[1] approach we introduce an explicit representation scheme for hidden interactions. A *hidden interaction model* links a required contextual, behaviour independent constellation to the impact of the interaction on the overall system behaviour. In order to control hidden interaction hypotheses we exploit the structure of diagnoses based on behavioural mode assignments.

## 1. Introduction

Look at figure 1. Pitilessly it reveals that kind of sore spot of model based diagnosis systems we want to tackle in this paper: how to cope with *structural faults* in a model based diagnosis framework. It shows a ballast tank system as needed on off-shore plants or ships to keep the balance. Some of the tanks are built on top of each other. A hole in the bottom/top of such a pair of tanks caused by corrosion results in an additional flow between them and hence in a change of the structure of the system. A conventional consistency-based diagnosis system (GDE. [deKleer, Williams 87]) generates diagnosis candidates based solely on a model of the *correct* function of the system. From a functional point of view the particular constellation of the tanks is not relevant. The model of the correct function will treat the two tanks as independent entities. Therefore an additional connection causing one tank to act as a source of a material flow into the other leads to discrepancies in apparently unrelated areas: *„tankl* loses water" and *„tank2* gains water". The corresponding conflicting assumption sets imply a multiple fault candidate. The actual *single fault location* is hidden.

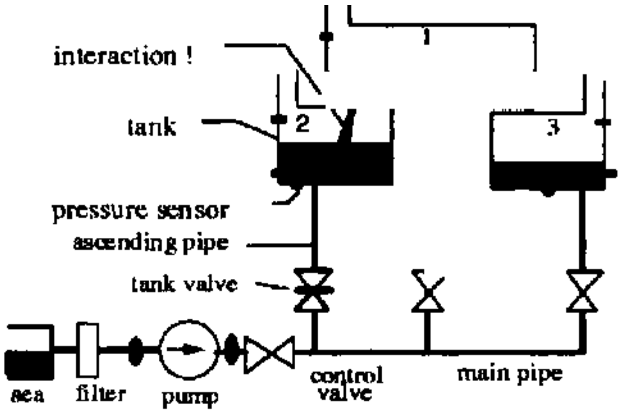


figure 1: A defective ballast system

However, the information about hidden interaction paths can be extracted from these multiple fault candidates. We presuppose, that a particular type of interaction between components requires some specific kind of neighbourhood depending on its type and results in changing the behaviour of the components to specific modes. We present *HiDe&Seek (Hidden* Interaction Detection), an approach to extend model-based diagnosis frameworks to include such a procedure.

The problem of dealing with structural faults has been recognised, as soon as research started in this field ([Davis 84]). Similar to the approach proposed by [Davis 84] HiDe&Seek is based on the integration of contextual knowledge to hypothesise structural model modifications. Davis' approach however is tailored to the diagnosis of TTL-circuits based on a constraint suspension framework and has not been adapted to other domains. Furthermore, we can profit from recent work refining the notion of diagnosis ([deKleer, Williams 89]). Instead of just stating which component correctness assumptions are consistent, we exploit a notion of diagnosis which associates consistent behavioural modes to each component.

[Preist. Welham 90] propose a solution based on the explicit representation of all potential unintended interaction paths. Due to the increase of model complexity, this

---

This work was started at the Fraunhofer Institute for Information- and Dataprocessing (FhG/IITB) in Karlsruhe, Germany, supported by the Bundesminister fur Forschung und Technologic joint project PT-FKZ:01 IW 203 B:8 and finished during a temporary research stay at the LIPN, supported by the EU under the contract CHBLCT941409 in the context of the human capital and mobility program.

approach has a very limited scope of applicability. One has to pay for potential faults with increased computational cost in *every* diagnosis case even if structure faults are known to appear seldom. In contrast HiDe&Seek modifies the system description *at runtime* and only when it has to consider a structural fault.

HiDe&Seek is based on the two presuppositions, that unintended interactions lead to apparently unrelated discrepancies and that the law of parsimony is a reasonable diagnosis preference criterion. Structural faults are considered only, if no single fault can be found. If forced to, those multiple faults which match a structure fault constellation are preferred.

To this end we introduce a representation for patterns of structural faults called *hidden interaction models,* A hidden interaction model has to serve two tasks. First, it specifies the structural fault constellation to be matched. This former task has two further aspects. On the one hand it specifies a required contextual constellation, e.g. a description of the physical layout of the components involved. The constellation specification acts as an interface to a constellation recognition module. Usually such kind of information is much easier to get than behavioural models. On the other hand, besides the constellation, the behavioural modes of the respective components have to be matched. E.g. a structural fault "leak between tanks" should be hypothesised only when the behavioural modes corresponding to a material flow /leak are actually consistent with the current observations. As a second task, the hidden interaction model provides basic information for how the system description has to be changed, i.e. a characterisation how the interaction affects the overall system behaviour.

In section 2 we introduce a ballast system as an example to be used throughout the paper. The diagnostic framework is sketched in section 3. Section 4 describes the extension of the diagnostic procedure to include the method sketched above and introduces the definition of a hidden interaction model. In section 5 we show how our procedure copes with the application example and finally, we reconsider the notorious bridge fault [Davis 84].

## 2. The Example

A ballast system consists of a set of ballast tanks distributed over the plant, connected by pipes with pumps. In order to stabilise the plant, valves are opened and closed to pump sea water in or out of the tanks according to the external requirements (sea motion, changing load). A conventional system consists of up to 40 ballast tanks. We shall restrict ourselves to a minimal configuration which is quite representative, however, since only a small number of tanks is allowed to be connected by open valves at any time.

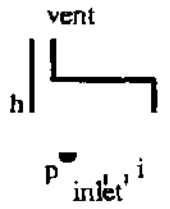
The main components included in our simplified presentation comprise ballast tanks, tank vent pipes, valves which can be opened/closed automatically or by hand, pipes, a pump, a filter, pressure sensors, control valve, and the sea tank (see figure 1, for details see [Ruemelin 921).

The following gives a rough approximation of the behaviour of each component. Besides the correct behaviour two fault modes are relevant for each. We shall discuss only the fault mode of a leaking tank, since it gives some insight into the problem to be solved.

The task of a tank is to keep water. For that purpose it is provided with one inlet and a vent pipe. Its behaviour is characterised by the water level. The water level changes according to the flow across the inlet. The vent pipe allows adjusting the pressure of the air in the upper part of the tank. So the static pressure at the tank bottom corresponds to the height of the water contained (see figure 2).

To represent a leaking tank, an additional flow $i_{leak}$ influencing the water level has to be considered. $h' = i/F$ has to be replaced by $h' = (i + i_{leak})/F$. Note, however that the actual law guiding the behaviour, i.e. consistency of flows, has not changed: $h' = \Sigma i / F$, where $\Sigma i$ denotes the sum of all flow influences on the tank. Only the number of relevant influences changed.[1]

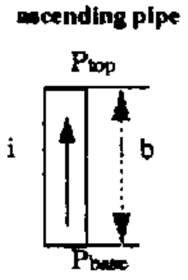
A pressure sensor reports the static pressure at the bottom of a tank.



**tank** : height h, height change h'
flow i, static pressure p
$h = p/(\rho g)$
$h' = p' / (rg)$
$h' = i / F$
where F is the tank's bottom surface,
$\rho$ is the density of sea water and
g is gravity acceleration

figure 2: tank behaviour description

A pump provides a flow into or out of the system. This flow is caused by producing a pressure drop and opening the control valve accordingly. The pressure values before and behind the pump are reported by two further sensors. Similarly, the behaviour of a pipe is characterised by the flow caused by the pressure drop between its ends. For an ascending pipe the static pressure has to be taken into account additionally (see figure 3).



pressure $p_{base}$ $p_{top}$
flow i
resistance R

$$p_{base} = p_{top} + i^2 R_{pipe} + \rho g b$$

where,
b is the height of the ascending pipe.
r is the density of sea water, and
g the gravity acceleration

figure 3: pipe behaviour

A valve establishes a common pressure point between two partial systems or keeps them independent of each other. If it is closed, the flow across a valve is zero (see figure 4). If a tank pressure sensor indicates that the water level is at its maximum, the corresponding valve is closed automatically. The state of a valve is continuously reported to the control system.

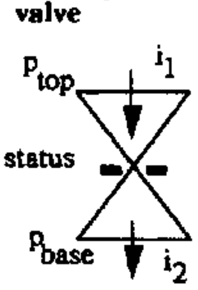
## 3. Diagnostic Framework

We follow a framework similar to those of [deKleer,Williams 89], [Hamscher 88], [Dressler, Struss 92] and [Böttcher, Dressler 94]. That is we perform *consis-*

---

[1] Note that the behavioural mode "leaks" covers both a loss and a gain of fluid.

***tency based diagnosis* based on a computational model of the device. i.a. the system description (SD). The basic idea of these approaches is to guide the search for diagnoses by discrepancies found between behaviour predictions based on the system description and the observations of the actual system.**



status, command
flow $i_1$ $i_2$
pressure $p_{base}$ $p_{top}$
status = command
if status = :open then $i_1 = -i_2$ and $p_{top} = p_{base}$
if status = :closed then $i_1 = i_2 = 0$

figure 4: valve behaviour

The system description consists of a collection of local behavioural models for each component and descriptions of the connections between the components. [deKleer,Williams 89] and [StrussJDressler 89] show how to profit from a system description differentiating between the different possible *behavioural modes* of each component. In the context of HiDe&Seek this kind of structuring the system description into component modes and models is crucial. It provides a basis for controlling when to hypothesise a structural fault as we shall see in section 4.

Then a diagnosis is defined as an assignment II of behavioural modes to the components of the system such that it is consistent with the current set of observations (Obs) and the system description.

For complex systems the system description may contain even more structure. Each component can be described at different levels of structural detail and each mode can be described at different levels of abstraction ([Struss 92}). It is profitable to exploit such structure during a diagnosis process in form of model switch operations such as *structural decomposition* or *behavioural refinement* ([Hamscher et al. 92}). To control the activation of those parts of the system description relevant in a particular stage of the diagnosis process, we adopt a technique as described in [Bottcher, Dressier 94]. As the other approaches mentioned above, this one views a diagnosis as an assignment of behavioural modes to components. But additionally it introduces a handle to control the partial activation of the system description and, hence, the diagnosis process. This will be the starting point to control the process of hypothesising hidden interactions.

The approach assumes that each piece of model can be addressed in a form such as

$$modg \wedge mode(\text{component}) \rightarrow \langle\text{model constraint}\rangle,$$

where *modg* is a conjunction of *working hypotheses* characterising that piece of model and mode(component) corresponds to the assumption that the component is in that particular behavioural mode. Adding the attached working hypotheses to a *current set of working hypotheses (w-hyp)* activates the related model constraint and allows testing the component mode assumption.

For example consider modelling a tank as used in the ballast system (section 2). If a tank is ok, the quantitative description of the water level *h* can be inferred from the static pressure *p* based on the gravity acceleration *g* and the density of water p:

*quantitative-view* A *ok(tank)* —> *h = pl(pg)*

By adding the working hypothesis *quantitative-view* to the current set of working hypotheses, the related pieces of the model are activated and the component mode assumption is tested.

Focusing strategies based on simplifying diagnostic hypotheses (e.g. „no multiple faults*') can be easily integrated into such a control framework as well.

Then a diagnosis under w-hyp is an assignment n of behavioural modes to the components of the system such that,

$SD \cup \Pi \cup Obs \cup w$-hyp is consistent

In general given a set of observations many mode assignments may fulfil this requirement. We assume that for each component each of its behavioural modes is weighted according to its diagnostic relevance (e.g. probability, risk), i.e. for every component a partial ordering is defined on the set of all its modes. The mode of the correct function precedes all other. Then only the most plausible or critical mode assignments in this sense satisfying the condition above are of real interest. Hence, the diagnosis search will focus on the so-called preferred diagnoses.

A mode assignment $\Pi = \{m_{ij}(c_i)\}$ is preferred over another $\Pi' = \{m_{ij}'(c_i)\}$ if. for all components Ci n assigns a mode mji whicE precedes or is equal to m;;* assigned by

II

Then a diagnosis is a preferred diagnosis under w-hyp. if no other diagnosis is strictly preferred over it ([Dressier, Struss 92]).

This characterisation implies that as long as no discrepancies have been detected, the mode assignment ..all components work correctly" is the most preferred one. Only if a preferred mode assignment has been refuted, i.e. turned out to imply discrepancies, successors are considered. Then the diagnostic process has to identify those mode assignments which deviate as few as possible from the previous one while being consistent with the observations, the system description and the current set of working hypotheses. That is. a preferred diagnosis stating "component c is consistent with mode m" actually says "c is consistent with a mode m and we don't know about less preferred ones yet".

The key to control the diagnosis process is the set of working hypotheses. Changing it activates different parts of the system description and results in a different set of diagnoses. A catalogue of partial preferences between subsets of working hypotheses representing general diagnostic and domain specific control knowledge, is used to guide the choice of an adequate set of working hypotheses in a given situation ([Bottcher, Dressier 94]).

## 4. Integrating Interaction Detection into the Diagnosis Process

The problem of the diagnosis approach sketched above is, that it depends on the system description which may be inadequate. Besides specifying how components behave and are connected, the system description states *implicitly* where interactions are *not* allowed. Structural faults violate the closed world assumption inherent in the model.

In order to integrate the consideration of structural faults two issues have to be addressed

- *when* to consider a violation of the closed world assumption and how to control it,
- how to *change* the system description to reflect the violation and how to control that.

Our answer is based on the presumption that an interaction hidden with respect to a model results in apparently unrelated discrepancies and, hence, in multiple fault candidates. As long as no multiple fault diagnosis has to be generated (according to the known discrepancies and current working hypotheses), the diagnosis process proceeds in the usual way[2]. If we are forced to consider multiple faults, we should still prefer a single explanation. Therefore, we shall look for an interaction acting as a single cause.

### 4.1 Hidden Interaction Models

For that purpose we introduce an explicit representation for interaction types: the *hidden interaction model.* The question is. what has to be included in such a model.

Consider the ballast system example. If working correctly, the water level of each tank can be changed only by a flow across the inlet (see figure 2 in section 2): h'=i/F. Now the structure of the system is changed by an additional flow-interaction between two tanks. This means, the behavioural modes of each of the tanks is changed. Each behaves according to an additional influence on the water level. Due to the hidden interaction path these additional influences constrain each other. Finally the character of the hidden interaction is a result of the particular construction of the tanks, i.e. one on top of the other, so as to share a steel plate between them. Considering all this, the flow interaction between the tanks is represented in the following way.

$$\text{leak-structure-fault}(t1,t2) \rightarrow$$
$$\mathit{below}(t1,t2) \wedge \mathit{distance}(t1,t2,0.0)$$
$$\wedge\ \text{leaks}(t1) \wedge \text{leaks}(t2)$$
$$\wedge\ t1.i_{leak} = -\,t2.i_{leak}$$

**In general a hidden interaction model is given as**

$$\mathit{hidden\text{-}interaction\text{-}hypothesis}(c_1,..c_n) \rightarrow$$
$$\mathit{constellation}(c_1,...,c_n)$$
$$\wedge\ \mathit{mode\text{-}constellation}(c_1,...,c_n)$$
$$\wedge\ \mathit{connection}(c_1..c_n)$$

The *constellation description* marks the interface to a module for recognising contextual, behaviour independent constellations such as a facility to represent and reason about geometrical or spatial information. What kind of constellation description is required or from what properties a constellation specification is made up of depends only on the application domain. The basic hidden interaction procedure described herein is independent of a particular choice. The only requirement is a facility to answer queries about which interaction types are consistent with a given set of components.

For the ballast system domain we used a special tool "Space-box" which provides a facility to recognise spatial configuration patterns. For the ballast system domain simple qualitative spatial relations such as "on-top-of', and quantitative relations such as "distance" were sufficient (for details see [Schick et al. 94], [Bottcher, Schick 94]). A more complex application might require additional technical properties such as „the plate connecting the tanks has to be made of steel", and, hence, a more complex tool. Other applications may come along with simpler languages and reasoning to check a constellation.

### 4.2 Controlling Hidden Interaction Hypotheses

Forced to consider multiple faults, but provided with a set of hidden interaction models representing relevant interaction types, we want to look for diagnoses which can be related to a single problem source. That is, if a preferred diagnosis has been inferred which *matches* the *structure fault pattern* of a hidden interaction model, we want to explore the consequences of the respective interaction, i.e. *to change* the system description so as to represent the influence of the hidden interaction on the overall system behaviour.

Given the structure of a hidden interaction model as introduced above and the framework to control the diagnosis as introduced in section 3, the system description change is realised by activating the related working hypothesis, in this case the related *hidden interaction hypothesis.* Activating a working hypothesis means adding it to the current set of working hypotheses. The decision when to switch to what set of working hypotheses is guided by a catalogue of partial preferences between subsets of working hypotheses ([Bottcher, Dressier 94]). This is the place where to state when to hypothesise a hidden interaction.

Before specifying the control preference rule to add, we have to precise, what is meant by a *matching structure fault pattern.* We introduce a predicate *potential-interaction* to describe that a particular hidden interaction hypothesis /z(ci,...,Cn) is a potential candidate to represent an interaction between suspect components ci,...,$c_n$ in the context of a mode assignment n. To be valid, we require on the one hand that the components are arranged corresponding to the constellation description of the hidden interaction type in question. Furthermore, we exploit the fact, that preferred diagnoses indicate which behavioural modes have been refuted already. We require that the behavioural modes assigned to the components by the diagnosis in question are "consistent" with those required by the interaction. Given the definition of preference between mode-assignments and of preferred diagnoses (section 3), "consistent" has to be read as "for each component the mode assigned by the diagnosis has to be more preferred or equal to that mode resulting from the interaction wrt. the partial mode ordering local to that component":

$$\mathit{potential\text{-}interaction}(h,\Pi,c_1,...,c_n) \rightarrow$$
$$\text{mode}_{\Pi}(c_i) <=_{c_i} \text{mode}_{h\text{-mode-constellation}}(c_i)$$

**where $\text{mode}_M(c)$ denotes the behavioural mode assigned to component c by (partial) mode-assignment M.**

**Refering to *potential-interactions*$(\Pi,c_1,...,c_n)$ as the set of all hidden interaction hypotheses for which *potential-interaction*$(h,\Pi,c_1,...,c_n)$ is valid. we add a control preference rule P**

---

[2] Single fault diagnoses which are related to an unintended interaction between a component and an external part are not covered by this scheme. But in principle they could be abstracted into a single faulty behavioural mode.

**if** $\exists\ \Pi \in \textit{diagnoses(w-hyp)} : \exists\ c_1,\dots,c_n : n > 1$
$\quad \wedge \forall\, i,j \in \{1,\dots,n\} : c_i \neq c_j$
$\quad \wedge\ \forall i \in \{1,\dots,n\} : \mathrm{ok}(c_i) \neq \mathrm{mode}_{\Pi}(c_i)$
$\quad \wedge\ \exists\ h(c_1,\dots,c_n) \in \textit{potential-interactions}(\Pi, c_1,\dots,c_n)$
$\qquad$ **then** $\{h(c_1,\dots,c_n)\} >_P \{\}$

to the catalogue of partial preferences between subsets of working hypotheses. It states, that whenever in a multiple fault situation a valid potential hidden interaction hypothesis can be found for some of the suspect components, adopting it is preferred over not doing so.

## 5. Hidden Interaction Detection in Action

### 5.1 The Ballast System Example

Consider the following scenario (see figure 5). *Tank*$_1$ and *tank*$_3$ have to be filled. For that purpose *valve*$_2$ is closed, *valve*$_1$ and *valve*$_3$ opened, and the *pump* is set in state :fill. The valve states are reported accordingly. The bottom of *tank*$_1$ is perforated connecting it to *tank*$_2$.

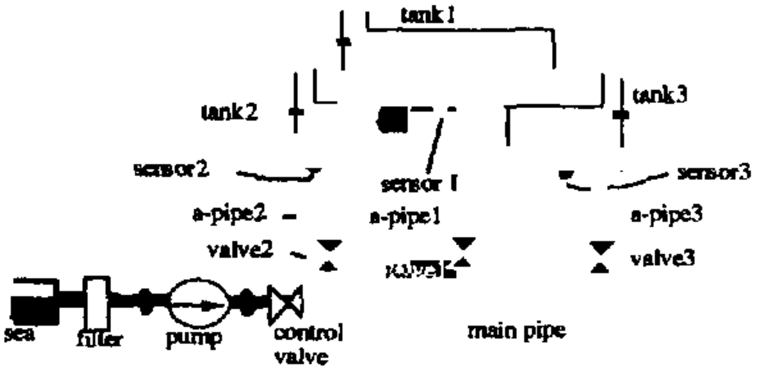


figure 5: a fault scenario

Given the scenario and the models sketched in section 2 the following sets of component correctness assumptions lead to discrepancies

- {ok(*tank*$_2$), ok(*sensor*$_2$), ok(*a-pipe*$_2$), ok(*valve*$_2$)}
  Since *valve*$_2$ is closed, the flow across it should be zero. This flow value should be the same as that across *a-pipe*$_2$, and as that out of *tank*$_2$. But *sensor*$_2$ reports a pressure change, and hence a flow unequal to zero.
- {ok(*tank*$_1$), ok(*sensor*$_1$), ok(*a-pipe*$_1$), ok(*valve*$_1$), ok(*tank*$_3$), ok(*sensor*$_3$), ok(*a-pipe*$_3$), ok(*valve*$_3$), ok(*main-pipe*)}
  This inconsistency is the result of comparing the values of the rate of change reported by *sensor*$_1$ and *sensor*$_3$. Since *valve*$_1$ is open, each of these values can be used to calculate the pressure below *valve*$_1$. The respective calculations are based on the closed world assumption that the rate of change reported by a tank sensor is directly related to the flow out of the tank. But in the case of *tank*$_1$ this is not valid, since it loses water on another way. As a result different, contradictory values are calculated for the pressure below *valve*$_1$.
- {ok(*tank*$_1$), ok(*sensor*$_1$), ok(*a-pipe*$_1$), ok(*valve*$_1$), ok(*pump*), ok(*control-valve*), ok(*main-pipe*)}
  Supposing that the rate of change reported by the pressure sensor at the *pump* corresponds to the flow across the *main pipe*, the pressure value below *valve*$_1$ can be calculated on a third way. Again this value is contradicting that one calculated on the base of *sensor*$_1$.

So the hypothesis that everything is ok, has been refuted. Hence other combinations of behavioural mode assignments have to be tested, the conflicting sets indicating which correctness assumptions have to be changed. The conflicting assumption sets above implicate at least a double fault of one of 16 component pairs or even a triple fault, where for each component two fault modes are relevant (see [Böttcher, Schick 94] for a more detailed discussion). Even if the partial order of preferred diagnoses is considered, the remaining search space is huge.

One of the resulting preferred diagnoses is {*leaks*(*tank*$_1$), *leaks*(*tank*$_2$)}. According to the constellation recognition this is the only double fault which could be caused by an additional interaction, since the components *tank*$_1$ and *tank*$_2$ satisfy the condition of being on top of each other, and because the behavioural modes assigned to them are equal, hence consistent with the "*leaks* "mode required by a hidden flow interaction. Therefore, the hidden interaction hypothesis *leak-structure-fault(tank*$_1$, *tank*$_2$) is added to the set of current working hypotheses which in turn enforces the corresponding behavioural modes and activates the connection $tank_1.i_{leak} = -\ tank_2.i_{leak}$.

As a consequence a preferred diagnosis such as {*leaks*(*tank*$_2$), *clogged(pipe*$_1$)} is refuted, since it assigns the mode "leaks" only to *tank*$_2$, but the mode "ok" to *tank*$_1$. Activating an interaction hypothesis results in changing the system structure. A connection between the tanks requires, that a leak flow observed at *tank*$_2$ has to be observed at *tank*$_1$ as well.

Note, that as long as the new working hypothesis is not added to the current set of working hypotheses, the extension of the system description does not affect the simulation.

### 5.2 The notorious Bridge Fault

[Davis 84] revealed the problem of dealing with structural faults by introducing the classic bridge fault example. A bridge fault is an unintended connection between lines on a circuit board caused by solder splashes. As a result the behaviour of an input port is changed into one of an output port.

To illustrate the similarities and differences, we examine the bridge fault example as described in [Davis 84] in detail. A bridge may occur where two components are connected to adjacent pins. The system description consists of gates (slices) and lines. It is structured according to the functional entities. Hence, pins are not modelled explicitly. The behaviour of two connected lines is dependent on the TTL technology. It corresponds to a connection by an and-gate (see figure 6).

To represent a bridge between lines as a hidden interaction model, suppose that the physical layout of the given circuit is represented in a constellation reasoning system. Then the adjacency of pins can be directly verified via their distance, e.g. pin-neighbour($p_1$,$p_2$) ::= distance($p_1$,$p_2$) <= $d_{min}$.

*bridge-fault*(*line*$_1$, *line*$_2$) $\rightarrow \exists\ p_1, p_2$ :
$\qquad$ *internal-pin* ($p_1$ ,*line*$_1$) $\wedge$ *internal-pin* ($p_2$ ,*line*$_2$)
$\quad \wedge$ *pin-neighbour*($p_1$,$p_2$)
$\quad \wedge$ *bridge-modes*(*line*$_1$,*line*$_2$)
$\quad \wedge$ (*line*$_1$.out = *line*$_2$.out = logical-and(*line*$_1$.in,*line*$_2$.in)).

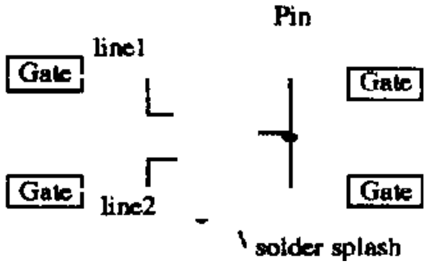


figure 6: A bridge between lines

Davis proposed a procedure to cope with bridge faults which starts with the assumption that the structure of the system is correct. If no single component fault can explain the observed behaviour, the assumption is relaxed and the procedure falls back on some special mechanism. The mechanism hypothesises structural modifications between components, if a particular pattern of values of the variables describing the components, has been derived and if they are neighboured in a particular way. The behaviour patterns required are specific for bridge faults and for the TTL technology.

The difference between the approach introduced by Davis and our lies in the notion of a diagnosis. In the former framework a diagnosis is viewed as a single component supposed to be not ok. On the one hand this implies that the single fault assumption is built into it. More important, however, is the fact, that it ignores the notion of behavioural modes. The knowledge about the behaviour of bridged components can be exploited in a "compiled" form only. That is, it is used beforehand to derive a particular pattern of variable values. So the pattern can be viewed as a result of changing the behavioural modes of the components in question and connecting them.

Common to both his and our approach is the presumption that a bridge leads to a multiple fault candidate and the insight to enhance the search for structure fault candidates by referring to the knowledge about adjacency between components.

## 6. Conclusions

We presented HiDe&Seek, a method to hypothesise hidden interactions based on an analysis of the context of components. We defined a hidden interaction model representation scheme which links a required contextual constellation to the description of the impact of an interaction on the overall system behaviour. We showed how to integrate the search for hidden interaction hypotheses into a diagnosis process control framework thereby exploiting the information about behavioural modes of components contained in candidate diagnoses to guide the search.

A straightforward solution to cope with structural faults would be to incorporate all potential pathways of interaction explicitly into the model. [Preist, Welham 90] shows that for applications where these kinds of faults are the actual problem and not the exception such a procedure is feasible but time consuming, hence, of limited scope of applicability. Similar to [Preist.Welham 90] we build our procedure on basis of an explicit representation of an "interaction". Our approach, however, dynamically activates additional interactions as needed.

Davis* approach is tailored to the diagnosis of TTL-circuits and hard to adapt to other domains ([Davis 84], section 5.2). But in some way our approach can be viewed as an extension of it. Analogous to that approach we based our procedure on the integration of contextual knowledge to hypothesise structure modifications. Additionally we exploit the structure of a diagnosis as a behavioural mode assignment.

Our procedure is based on the assumption that unintended interactions lead to discrepancies in unrelated areas. It seems that this assumption is useful to treat structural faults resulting from unintended interactions between system components in some domains. It is. however, not obvious for what kind of applications this assumption is satisfied. Characterising the difference between a multiple fault hiding an interaction between system components and one hiding a interaction with some external part, or a fault (single or multiple) hiding an interaction with an external part from an ordinary fault still needs to be done. So future work will comprise characterising classes and properties of structural faults in different domains.

The procedure has been implemented in Allegro Common Lisp on a SUN Sparc workstation and tested on the ballast system example.

## Acknowledgements

Thanks to Philippe Dague. Oskar Dressier. Francois Levy, Karsten Poeck. Peter Struss and the anonymous referees for commenting on earlier versions of this paper, and last but no least to Matthias Schick who provided the Space Box as the basis of this work.

## References

[Bottcher Dressler 94] C. Bottcher. O.Dressler. A Framework for Controlling Model-Based Diagnosis Systems with multiple Actions, in: *Annals of Mathematics and Artificial Intelligence,* Special Issue on Principles of Model-Based Diagnosis Systems, Vol. IInr.1-4,1994

[Bottcher.Schick 94] C. Bottcher. M. Schick. Diagnosing Structural Faults by Detecting Hidden Interactions. *Proc. 5th International Workshop on Principles of Diagnosis* (DX-94). New Paltz, NY, USA, 1994

[Davis 84] R. Davis. Diagnostic Reasoning based on Structure and Behavior, in: *Artificial Intelligence* (24).1984

[deKleer, Williams 89]. J. deKleer, B. Williams. Diagnosis With Behavioral Modes, in: *Proc. IJCAI-89,*1989

[deKleer. Williams 87], J. deKleer, B. Williams. Diagnosing Multiple Faults, in *Artificial Intelligence* (32), pp. 97-130. 1987

[Dressier, Struss 92] O. Dressier, P. Struss. Back to Defaults: Characterising and Computing Diagnoses as Incoherent Assumption Sets, in: *Proc. ECAI*$_t$ Vienna, pp. 719-723. 1992

[Hamscher et al. 92] W. Hamscher, L. Console. J. deKleer (eds.). *Readings in Model-Based Diagnosis,* Morgan-Kaufmann. San Mateo, CA. USA. 1992

[Preist, Welham 90] C. Preist, B. Welham. Modelling Bridge Faults for Diagnosis in Electronic Circuits, in: *Proc. First International Workshop on Princi-*

*pies of Diagnosis** Stanford, 1990

[Ruemelin 92] W. Ruemelin. Systemanalyse des Dreitank-Ballastsystems, in German, Behavior-Memo 11-92, 1992. available upon request from Fraunhofer Institute, Karlsruhe

[Schick et al. 94] M. Schick, S. Kockskaemper, B. Neumann. Analysing spatial Constellations in Technical Domains, internal report, available from the authors upon request, university of Hamburg, 1994

[Struss 92] P. Struss. What's in SD - towards a modelling theory for diagnosis, in: [Hamscher et al. 92]

[Struss. Dressier 89] P. Struss, O. Dressier. "Physical Negation "- Integrating Fault Models into the General Diagnostic Engine, in: *Proc. 12th IJCAI,* 1989