Diagnosis of Dynamic Systems Does Not Necessarily Require Simulation

Andreas Malik, Peter Struss

Department of Computer Science
Technical University of Munich
Orleansstr. 34, 81667 München, Germany
{malik, struss}@informatik.tu-muenchen.de

Abstract

We present a paradigmatic example of a feedback-controlled system: an electric motor with sensor and controller. Diagnosis of this system is performed based on a qualitative model that reflects deviations of parameters and behavior from a fixed reference state. The hypothesis that has been examined in this case study is that detection of behavior discrepancies does not necessarily require simulation of behavior, but can be done by checking (qualitative) states only. The qualitative models and the state-based diagnosis algorithm proved to establish a basis sufficient for fault detection and fault identification in the motor example. Some of the general preconditions for this are discussed.

Introduction

Dynamic systems are considered as **the** challenge for modeling, particularly for qualitative modeling. Since the evolution of characteristics over time is the crucial aspect of such systems, it is often taken for granted that computational methods for problem solving necessarily involve **simulation** of the behavior of the respective system.

Numerical simulation is **not applicable** if there is only partial, or qualitative, information about the system and its initial conditions. **Qualitative** simulation, designed for such cases, can be **complex** due to ambiguity in the predicted set of behaviors. Anyway, it would be good if one could get along without simulation.

In our project on diagnostic techniques and tools for car subsystems, we successfully demonstrated the utility for qualitative modeling and consistency-based diagnosis for tasks such as failure mode and effects analysis (FMEA) and automated generation of repair manuals (see [Struss-Malik-Sachenbacher 96]). The solutions were based on static models. However, many subsystems of vehicles that demand for automated diagnosis, such as the Anti-lock Braking System (ABS) and the Electronic Diesel Control (EDC), are dynamic feedback systems. "If you want to diagnose such dynamic systems, you need (qualitative) simulation". Our response to this prejudice was a case study checking our working hypothesis "you can achieve a lot without simulation" in preparation of work on the ABS and EDC.

The example we chose was a simple feedback system involving an electrical motor, speed sensor, and controller. Nevertheless, it constitutes a challenge in three respects:

- It is a **dynamic** system. Is it possible to diagnose it **without** having to perform some kind of **simulation**?
- It is a **continuous** system. Is it possible to diagnose it based on a **qualitative** model?
- It is a **feedback** system. Is it possible to diagnose it using **consistency-based diagnosis** (especially dependency-based diagnosis) despite the fact that each observation in the feedback loop is dependent on all components in the loop?

In this paper, we present the results of this case study. The answer to the third question, which is not the focus here, is that fault localization in feedback loops with limited observability, if possible at all, inevitably has to be based on fault models (which is necessary for fault identification for even more obvious reasons, anyway). The basis is "physical negation" ([Struss-Dressler 89]), i.e. exonerating components whose entire set of fault models is refuted by the observations.

The ultimate reason for a positive answer to the second question is that a fault is almost defined as a cause of some **qualitative** deviation from normal behavior. In our case, faults can be described as qualitative deviations of actual parameter values from the nominal ones.

The answer to the first question, which is central to this paper, is based on the following consideration: the essence of consistency-based diagnosis is to refute (correct of faulty) behaviors that are inconsistent with the observations. In qualitative reasoning, behaviors are sequences of qualitative states. If a model predicts a sequence different from the observed one it obtains an inconsistency. Predicting a sequence (or a tree) of states means performing some kind of simulation (or envisionment). However, we can do with less: if we observe a **single state** that is not consistent with the (dynamic) model of a particular behavior, this suffices to establish an inconsistency and, hence, to refute this behavior. For detecting this inconsistency, **simulation** is **not required**. All we need is to check whether the observed states are consistent with the respective behavior models.

In the following section, we describe the paradigmatic example of the controlled electric motor, the faults considered, and the diagnostic scenario. The models of the three components of the circuit are based on a formalized concept of qualitative deviations. Then we introduce the foundations for diagnosis of the circuit. Finally, we present the results of our experiment and discuss its preconditions and limitations in more detail.

The Problem

The Control Circuit with Motor

The example deals with a direct-current motor that is controlled in a feedback loop (Figure 1). The actual speed ω of the axis of motor M is measured by a revolution counter S. The respective value ω_m is fed to the controller C. Using knowledge about the present measured motor speed and the desired speed d, the controller adjusts the voltage v driving the motor.

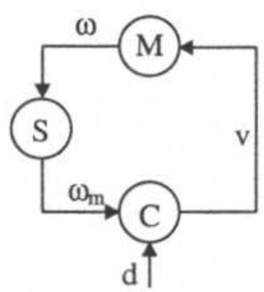


Figure 1: Control circuit with motor

The (differential) equations of the three involved components and the physical constants are given in Table 1.

Measurement of speed is achieved by means of counting pulses generated by a ferromagnetic toothed wheel („pulse wheel") and an inductive sensor. This sensor has also been modeled in more detail which is not discussed here.

The Faults Considered

The example, although simplified, is taken from a real application problem where numerical simulation alone could not meet requirements of fault prediction and fault

Motor M:	$T * \frac{d\omega}{dt} = c_M * v - \omega$
Controller C:	$c_C * v = 2 * d - \omega_m$
Revolution counter S:	$\omega_m = c_S * \omega$
ω	rotational speed of the motor [s^{-1}]
T	inertia of the motor [s]
c_M	constant of the motor [$s^{-1}V^{-1}$]
v	driving voltage [V]
c_C	constant of the controller [$s^{-1}V^{-1}$]
ω_m	measured rotational speed [s^{-1}]
d	desired rotational speed [s^{-1}]
c_S	measurement coefficient [1]

Table 1: Equations and quantities of the control circuit

identification. It includes, along with the differential equations stated above, a catalog of faults that are to be distinguished. This set of faults was extended by some less likely defects. The clear-cut success criteria were defined as to how well the diagnostic system can detect, isolate and discriminate among the faults listed in the left-hand column of Table 2.

It is obvious that not all faults can be distinguished from each other solely by the effects on the input/output variables of the components. For example, a pulse wheel with too few teeth will cause the same deviation as a constant positive slippage does.

The Diagnostic Situation

We assumed conditions similar to on-board diagnosis. The diagnostic system is situated in or near the controller and can inspect the **measured rotational speed** (ω_m), the **desired speed** (d) and possibly information about the derivatives of these quantities. The current (v) and the actual speed of the motor (ω) are not available to the diagnostic system.

The scenario we discuss in detail here is observing the response of the system to a **stepwise change in d**, i.e. a discontinuous switching from one constant value to another. The goal was to analyze if, and to what extent, the applied models and the diagnostic algorithm are suitable for

- **fault detection**, i.e. to detect **that** a fault is present,
- **fault localization**, i.e. to detect **where** the fault lies,
- **fault identification**, i.e. to determine **which** fault is present.

In the following sections, we will discuss what requirements this imposes on the applied modeling and diagnosis techniques, which solutions were chosen, and what results were achieved by their implementation with respect to these tasks.

Failure cause	Behavior model
Motor	
flow constant too low, ferro-magnetic loss to high	$[c_M] = [+]$, $[\Delta k_M] = [-]$, $[T] = [+]$, $[\Delta T] = [-]$
motor constant too low	$[c_M] = [+]$, $[\Delta k_M] = [+]$, $[T] = [+]$, $[\Delta T] = [-]$
resistance of field- and/or rotor coil too high	$[c_M] = [+]$, $[\Delta k_M] = [-]$, $[T] = [+]$, $[\Delta T] = [+]$
inertia of motor too high	$[c_M] = [+]$, $[\Delta k_M] = 0$, $[T] = [+]$, $[\Delta T] = [+]$
inertia of motor too low	$[c_M] = [+]$, $[\Delta k_M] = 0$, $[T] = [+]$, $[\Delta T] = [-]$
rotor totally jammed	$[c_M] = 0$, $[\Delta k_M] = [-]$, $[T^{-1}] = 0$, $[\Delta T] = [+]$
Controller	
controller constant too high	$[c_C] = [+]$, $[\Delta c_C] = [+]$
controller constant too low	$[c_C] = [+]$, $[\Delta c_C] = [-]$
broken wire controller-motor	$[c_C^{-1}] = 0$, $[\Delta c_C] = [+]$
Revolution Counter	
complete slippage, pulse wheel not ferromagnetic, gap pulse wheel - sensor too big, trigger threshold too high	$[c_S] = 0$, $[\Delta c_S] = [-]$, $\partial c_S = 0$
constant slippage, pulse wheel with too few teeth, assumed number of teeth too high	$[c_S] = [+]$, $[\Delta c_S] = [-]$, $\partial c_S = 0$
pulse wheel with too many teeth, assumed number of teeth too low	$[c_S] = [+]$, $[\Delta c_S] = [+]$, $\partial c_S = 0$
sporadic slippage, missing or broken tooth (teeth)	$[c_S] = [+]$, $[\Delta c_S] = [-]$ or $[\Delta c_S] = 0$
notch in tooth, ridge between teeth of pulse wheel	$[c_S] = [+]$, $[\Delta c_S] = [+]$ or $[\Delta c_S] = 0$
displaced tooth, too early or too late	$[c_S] = [+]$

Table 2: Component faults and their behavior models

The Models

The important issues to be addressed in this case study were

- Is qualitative modeling sufficient to capture the dynamic aspects of the example?
- How are the obtained models to be used for diagnosis, especially: is qualitative simulation inevitable?

The first question is vital for the robustness and generality of the diagnostic system; both questions have decisive influence on the complexity of the solution.

In the following, the main ideas underlying the modeling of the control circuit will be presented. Later, we will describe the application to diagnosis.

Qualitative Modeling of the Control Circuit

If we analyze the set of possible faults listed in the previous section, we notice that all of them can be described by a deviation of parameters occurring in the equations of the respective component. For instance, the controller constant can be too high, and a missing signal ω_m is given by $c = 0$. Note that the list of faults actually talks about **classes** of faults, rather than specifying numerical distinctions. For example, it would be highly inappropriate to try to characterize slippage by exact figures. Characterizing parameters just by the direction of their deviation appears to suffice for diagnostic purposes. This is the basis for the models we used in this case study. For each **parameter** x we introduce

$$\Delta x = x_{act} - x_{ref},$$

the deviation of the actual value from the nominal one, which characterizes normal behavior. Only $[\Delta x]$, i.e. the sign of Δx, matters. Deviation of parameters potentially produces deviations in the **system variables**. They could be defined as the difference between the dynamic quantities corresponding to the actual behavior and a reference behavior:

$$\Delta x(t) = x_{act}(t) - x_{ref}(t).$$

Determining $\Delta x(t)$ requires, besides a measurement of $x_{act}(t)$, computation of $x_{ref}(t)$ for the respective time point t, i.e. simulation. However, nothing prevents us from choosing an arbitrary quantity for reference.

In our models, we chose **fixed** values of the variables as a reference which could be easily determined: the value of the equilibrium state of the system under correct behavior of all components, i.e.

$$\omega_{ref} = \omega_{m\text{-}ref} = d,$$
$$v_{ref} = c_C^{-1} * d,$$

and 0 for all derivatives:

$$\left(\frac{d\omega}{dt}\right)_{ref} = \left(\frac{d\omega_m}{dt}\right)_{ref} = \left(\frac{dv}{dt}\right)_{ref} = \frac{dd}{dt} = 0.$$

In particular, we obtain

$$\Delta\omega_m = \omega_m - d,$$

which can be determined from the measurements in the diagnostic scenario as defined in earlier.

Qualitative deviations of sums and products can be expressed as qualitative sums and qualitative products. With the definition of Δx we obtain the following rules (which actually hold independently of the particular choice of reference values):

$$[\Delta(a + b)] \equiv [(a + b) - (a + b)_{ref}] \equiv [a + b - (a_{ref} + b_{ref})] \equiv [\Delta a + \Delta b] = [\Delta a] \oplus [\Delta b]$$

$$[\Delta(a * b)] \equiv [(a * b) - (a * b)_{ref}] \equiv [a * b - a_{ref} * b_{ref}] \equiv [a * \Delta b + b * \Delta a - \Delta a * \Delta b] = [a] \otimes [\Delta b] \oplus [b] \otimes [\Delta a] \ominus [\Delta a] \otimes [\Delta b]$$

The deviation of a product can be simplified if at least one of the reference values, e.g. $[b_{ref}]$, is known:

$$[\Delta(a * b)] \equiv [a * \Delta b + \Delta a * b_{ref}]$$
$$= [a] \otimes [\Delta b] \oplus [\Delta a] \otimes [b_{ref}]$$

Finally, $\left[\Delta \frac{1}{a}\right]$ can be transformed to:

$$\left[\Delta \frac{1}{a}\right] \equiv \left[\frac{1}{a} - \frac{1}{a_{ref}}\right] \equiv \left[\frac{a_{ref} - a}{a * a_{ref}}\right] \equiv - [\Delta a] \otimes [a] \otimes [a_{ref}].$$

For qualitative derivatives, we use the notation ∂:

$$\partial x := \left[\frac{dx}{dt}\right],$$

and $\partial\Delta$ denotes the qualitative deviation of a derivative, which is justified because

$$\left[\Delta \frac{dx}{dt}\right] = \left[\frac{d}{dt} \Delta x\right] = \partial\Delta x .$$

Since all component faults can be described as parameter deviations, the equations presented above hold for both the correct and the faulty behaviors. From these equations (and their derivatives) we obtain the following generic qualitative models of the respective components, comprising the qualitative and the Δ-version of the original ones:

Motor: $[T] \otimes \partial\omega \oplus [\omega] = [c_M] \otimes [v]$

$$[\Delta T] \otimes \partial\omega \oplus \partial\Delta\omega \oplus [\Delta\omega] = [\Delta c_M] \otimes [v] \oplus [\Delta v]$$

Controller: $[c_C] \otimes [v] \oplus [\omega_m] = [d]$

$$[\Delta c_C] \otimes [v] \oplus [\Delta v] \oplus [\Delta\omega_m] = [\Delta d]$$

Rev. counter: $[\omega_m] = [c_S] \otimes [\omega]$

$$[\Delta\omega_m] = [\Delta c_S] \otimes [\omega] \oplus [\Delta\omega]$$
$$\partial\omega_m = [c_S] \otimes \partial\omega \oplus \partial c_S \otimes [\omega]$$
$$\partial\Delta\omega_m = [\Delta c_S] \otimes \partial\omega \oplus \partial\Delta\omega \oplus \partial c_S \otimes [\omega]$$

Models of the correct and the faulty behavior are derived from these generic models by conjunction with their characteristic constraints on signs of parameters and their deviations. A non-negligible, but not total, constant slippage of the sensor wheel, for instance, is given by

$$[\Delta c_S] = [-],\ [c_S] = [+],\ \partial c_s = 0$$

and, hence, the behavior model

$$[\omega_m] = [\omega] = [+]$$
$$[\Delta\omega_m] = [\Delta\omega] \ominus [\omega] = [\Delta\omega] \ominus [+]$$
$$\partial\omega_m = \partial\omega$$
$$\partial\Delta\omega_m = \partial\Delta\omega \ominus \partial\omega$$

if ω is assumed positive. This model captures, for example, the information that the measured speed will be too low if the speed of the motor coincides with the reference value ($[\Delta\omega] = 0$).

The constraints for the various fault models which combine with the generic models are given in Table 2.

One has to keep in mind that the models, being derivations of the underlying (differential) equations, are considered valid at any time point, and the variables are treated as continuous functions. This is only an approximation, since, after all, the speed is computed whenever a new pulse is received and then kept constant until the next computation. Hence, ω_m is actually a step function with discontinuities at each time point corresponding to a pulse. As a result, the models presented here have limited power when it comes to discrimination between different faults within the sensor. For this purpose, a refined model of the sensor has been developed.

Foundations of the Diagnostic Approach

Consistency-based Diagnosis of Dynamic Systems

As stated in the introduction, our work is based on consistency-based diagnosis ([Dressler-Struss 96]) which checks consistency of a behavior model with a set of observations of the actual system behavior. Expressed in a more formal way, the diagnostic algorithm decides whether a set of observations, OBS, an assignment of particular modes of behavior (correct or faulty) to the components, C_i, of the system, and the structural and behavioral model of this system together form a consistent theory or entail an inconsistency:

$$\text{MODEL} \cup \{\text{mode}(C_j)\} \cup \text{OBS} \overset{?}{\models} \bot .$$

This allows for

- **fault detection**, if there is an inconsistency with the model of correct behavior,
- **fault localization** by suspecting all components whose models of correct behavior contribute to the inconsistency („dependency-based diagnosis"),
- **fault identification** by refuting component faults whose models are inconsistent with the observations.

In each case, a behavioral discrepancy is the starting point. For static models, such a discrepancy is simply given by two contradictory states, i.e. different values of one variable. For a system that changes state over time, a discrepancy is obtained if there are two conflicting predicted/observed states **for the same time points**. Often, it is concluded that, in order to detect this,

- we need simulation to derive a description of behavior over time, and
- we need numerical simulation, because different values establish a discrepancy only if they refer to the very same time point.

The example in our case study sheds a light on these hypotheses. Fault detection may work with a numerical model: initial values of ω_m and $\frac{d}{dt}\omega_m$ could be used to simulate the expected behavior based on the given (differential) equations, provided there is an appropriate way to distinguish a real behavior discrepancy from a virtual one which is due to errors of the simulation algorithm. Fault identification would be impossible, since for simulating the possible faults, initial values would be required which are simply not derivable.

If we apply **qualitative** simulation, we face the problem of **synchronization**, i.e. relating observed, real time and the qualitative representation of time in the model. There are several diagnosis systems that are based on running qualitative simulation of a system concurrently with the observation of its actual evolution (e.g. MIMIC, [Dvorak-Kuipers 92]). Again, based on the assumption of a known initial state (and real-time performance of the qualitative simulation algorithm), the detection of a (qualitative) discrepancy between predicted and observed behavior and, hence, fault detection is possible. Identification of the fault usually faces complexity problems, since simulation of many fault situations may be required, even though the problem of unknown initial conditions is somewhat relaxed: there is only a finite number of initial states to begin with. A system like MIMIC has to assume some effective heuristic for selecting fault models to try.

With the background given above, what simulation-based systems like MIMIC do can be described in a general and formal way as checking consistency of a sequence of observed states, ($S_{obs\ 1}$, $S_{obs\ 2}$, …, $S_{obs\ k}$), with the given (qualitative) dynamic model:

$$\text{MODEL} \cup \{\text{mode}(C_j)\} \cup \{(S_{obs1}, S_{obs2}, \ldots, S_{obsk})\} \overset{?}{\vdash} \bot .$$

At least theoretically, we can regard the conjunction of the mode assignment and the model as the set (disjunction) of possible behavior sequences emerging from the initial state, $S_{obs\,1}$, according to the model:

$$\{(S_{obs\,1}, S_{i\,2}, \ldots, S_{i\,n})\} \cup \{(S_{obs\,1}, S_{obs\,2}, \ldots, S_{obs\,k})\} \overset{?}{\vdash} \bot ,$$

and to be consistent with this set, the observed state sequence has to be a subsequence of at least one of the predicted ones (with or without "gaps"). This also subsumes approaches that, instead of generating the predicted sequence concurrently with the evolving real process, compile empirical knowledge or first principles into associations between modes and resulting system evolution. [Milne et al 94] presents an example which matches observations over time with given "chronicles", i.e. sequences of events including temporal constraints.

As we stated earlier, we tried to prove our working hypothesis:

Model-based diagnosis of dynamic systems does not necessarily require simulation.

The basic idea is more than simple and may appear close to naive: we ignore the chronological information, represent observations over time by a set of states, $\{S_{obs\ i}\}$, rather than a sequence, and the diagnostic engine checks each individual state

$$\text{MODEL} \cup \{\text{mode}(C_j)\} \cup \{S_{obs\,i}\} \overset{?}{\vdash} \bot \quad \text{for all i.}$$

This means checking whether there exists an observed state that is not in the **set** of states specified by the model under the mode assignment, although this set is not enumerated. Rather, the usual constraint-based consistency check of the "static" consistency-based diagnostic engine applies. We got rid of simulation, and we do not need to assume subsequent observations. Obviously, inconsistency of an individual state with the model is a **sufficient condition** for a whole sequence that contains it to be inconsistent. Hence, our diagnostic algorithm is guaranteed to produce a superset of the diagnoses (consistent mode assignments) generated by the state-sequence-based approach which, in turn, contains the actual diagnosis.

Because one may suspect that this diagnosis algorithm is too weak, we use the motor example to provide some evidence for why it might work. (In the final section, we outline a more systematic analysis of its preconditions and limitations.)

Fault Detection and Identification in the Control Circuit

Figure 5 depicts the possible qualitative behaviors of the control circuit under normal conditions and for a particular fault ("controller constant too high") for varying initial conditions.

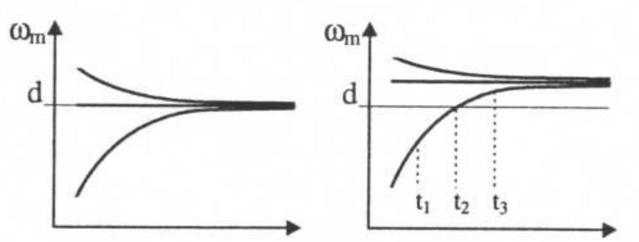


Figure 5: Development of ω_m for correct behavior (left) and controller constant too low (right) for different initial ω_m

The difference is clearly captured by the chosen qualitative representation ([$\Delta\omega_m$], $\partial\Delta\omega_m$) where (d, 0) is the reference value, leading to a compact description of the possible states as in Figure 6.

{([+], [−])	{([+], [−]), ([+], 0)
(0, 0)	([−], [+]), (0, [+])
([−], [+])}	([+], [+])}

Figure 6: Set of possible qualitative states for correct behavior (left) and controller constant too low (right), expressed in ($\Delta\omega_m$, $\partial\Delta\omega_m$)

This tells us, for instance, that state (0, 0) does not conform with the illustrated faulty behavior, whereas (0, [+]) contradicts correct behavior, and illustrates the possibility to dispense with simulation and simply check each observed state for consistency with the behavior models currently under consideration. Suppose the controller fault is present, and the diagnostic engine observes states in a snapshot manner and checks them for consistency. Starting with the observation ([–], [+]) at snapshot t_1 (in Figure 5), both cases are conceivable, but when state (0, [+]) at snapshot t_2 is reached, correct behavior can be excluded (fault detection). Because this state also contradicts, for instance, "controller constant too high", this also enables fault identification.

Practical Evaluation of the Models and the Diagnosis

In this section, we first summarize the results of applying the state-based diagnostic algorithm to the electric motor example. Because it worked remarkably well, we also provide a more detailed analysis of why and how the system achieved its results for this example.

Diagnosis of the Control Circuit - Results

The models were implemented using constraint-based component-oriented modeling and evaluated in the diagnostic tasks. The change in the desired rotational speed was chosen to be a step (i.e. $\partial d = 0$); observations were available as described above:

- d, the desired rotational speed
- ω_m, the measured rotational speed
- $\partial\omega_m$, the (sign of the) derivative of ω_m

From these observations, the system determines

- $[\Delta\omega_m] = [\omega_m - d]$, the qualitative deviation of ω_m from the desired rotational speed and
- $\partial\Delta\omega_m = \partial\omega_m$.

Application of the models for fault detection yields the following results:

- Of **26** considered faults
- **24** are detectable merely by checking state-consistency.

The fault identification results are discussed below.

How Does it Work in Detail?

The world spanned by the qualitative representation ($\Delta\omega_m$, $\partial\Delta\omega_m$) with three possible values for each variable is quite perspicuous.

		$\partial\omega_m = \partial\Delta\omega_m$		
		[–]	0	[+]
$[\Delta\omega_m]$	[–]	4	4	c
	0	2(1)	c	3
	[+]	c	5	5

Table 3: Possible observable states. c marks states that are consistent with correct behavior.
The numbers refer to categories in the text.

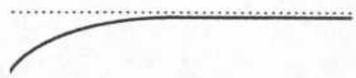
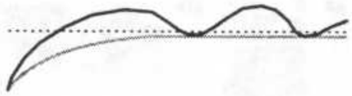
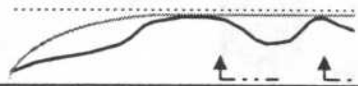
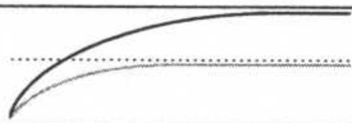
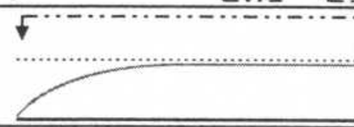
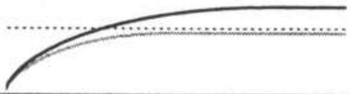
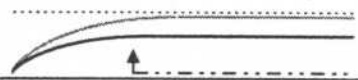
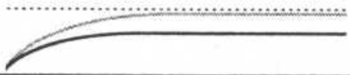
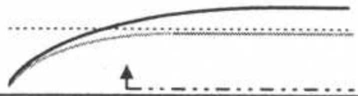
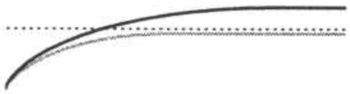
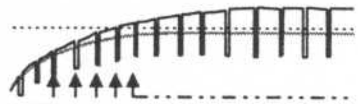
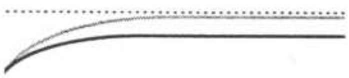
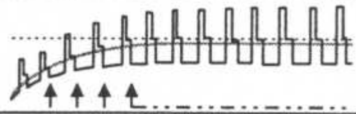
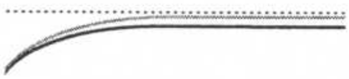
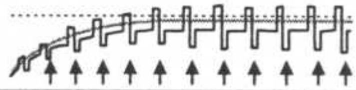
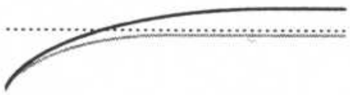
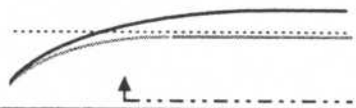
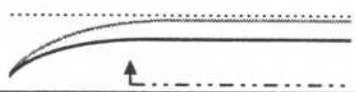
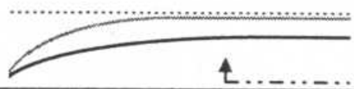
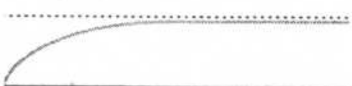
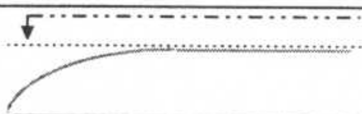
The nine theoretically possible states are listed in Table 3. Three of them (marked with "c") are consistent with correct behavior of the control circuit (but also with some faulty behaviors). The remaining cases are only consistent with faulty behavior and are hence suitable for fault detection:

1. The measured rotational speed remains **constantly zero** (specialization of 2).
2. The measured rotational speed remains **constant**, being **lower** than the desired rotational speed.
3. The measured rotational speed remains **constant**, being **higher** than the desired rotational speed.
4. The measured rotational speed **matches** or **exceeds** the desired rotational speed, but increases.
5. The measured rotational speed **matches** or **does not come up to** the desired rotational speed, but **decreases**.

Table 4 indicates for each type of fault the instances when these faults are detected because they produce one of these states. Please note that, although the evolutions of the behaviors are sketched for illustrative purpose, this is nothing actually generated by the system or given to it as an input. Where fault detection is achieved over a longer period of time this is also indicated. Notably, some faults can be detected quite early, with certainty after some period of adjustment has finished. The two unrecognized faults are due to a deviate inertia of the motor, and only manifest themselves in a slower or faster period of adjustment. As the models contain no concept of "slow" or "fast", these faults remain undetected.

Table 5 shows how the observable qualitative states contribute to fault identification (states ([–], [+]) and ([+], [–]) are omitted, because they are consistent with almost all behaviors). A tick ("✓") marks those states that are consistent with a certain cause of failure; likewise a cross ("✗") marks inconsistencies. The table makes evident that, for example, state 2 is still consistent with five different failure causes, yet with a previous observation of state 4 only three causes remain: controller constant too high, motor constant too low, and measurement coefficient too low. It should be noted though, that even numerical simulation cannot distinguish these three causes: all three components stay suspect.

Next page, Table 4: Failure causes, effects and detection in the control circuit. Legend:

Fault	Effect of fault (ω)	Fault detection (ω_m)
correct		
sporadic slippage		
total slippage, gap too large, pulse wheel not ferromagnetic trigger threshold too high		
too few teeth, too many teeth assumed, constant slippage		
too many teeth, too few teeth assumed		
missing tooth (or teeth)		
notch or ridge		
displaced tooth (too early or too late)		
controller constant too low, motor constant too high		
controller constant too high, flow constant too low, ferromagnetic loss too high		
resistance of field- and/or rotor-coil too high		
inertia of motor too high		
inertia of motor too low		
jammed motor, broken wire controller-motor		

Fault	Qualitative state 1 (— / →)	2 (→ / —)	3 (→ / —)	4 (— / ↘)	5 (↗ / —)	(0,0) →
control const too low $[\Delta c_C] = [-]$	✗	✗	✓	✗	✓	✗
control const too high $[\Delta c_C] = [+]$	✗	✓	✗	✓	✗	✗
motor const zero $[c_M] = 0$	✓	✓	✓	✗	✗	✗
motor const too low $[\Delta c_M] = [-]$	✗	✓	✗	✓	✗	✗
motor const too high $[\Delta c_M] = [+]$	✗	✗	✓	✗	✓	✗
rotor jammed „$T = \infty$"	✓	✗	✗	✗	✗	✗
inertia of motor too low $[\Delta T] = [-]$	✗	✗	✗	✗	✗	✓
inertia of motor too high $[\Delta T] = [+]$	✗	✗	✗	✗	✗	✓
measure coeff. too low $[\Delta c_S] = [-]$ $(\partial c_S = [?])$	✗	✓	✗	✓	✗	✗
measure coeff. too high $[\Delta c_S] = [+]$ $(\partial c_S = [?])$	✗	✗	✓	✗	✓	✗
correct	✗	✗	✗	✗	✗	✓

Table 5: Consistency („✓") and inconsistency of behavior and qualitative states

Extensions

The experiment can be extended to cover more general situations. Particularly, one can

- change the type of input and
- use a different controller.

Instead of using a step function for d, a **ramp** could be considered. This input can be characterized qualitatively by

$$[d] = [+], \partial d = [+], \partial^2 d = 0$$

(or $\partial d = [-]$). The same reference values as before can be used:

$$\omega_{ref} = \omega_{m\text{-}ref} = d,$$
$$v_{ref} = c_C^{-1} * d,$$
$$\left(\frac{d\omega}{dt}\right)_{ref} = \left(\frac{d\omega_m}{dt}\right)_{ref} = \left(\frac{dv}{dt}\right)_{ref} = \frac{dd}{dt},$$
$$\left(\frac{d^2\omega}{dt^2}\right)_{ref} = \left(\frac{d^2\omega_m}{dt^2}\right)_{ref} = \left(\frac{d^2v}{dt^2}\right)_{ref} = \frac{d^2d}{dt^2} = 0,$$

but now d is varying over time. Furthermore, we have to extend the component models by including equations corresponding to the next order of derivatives.

Note that a deviation of ω (and ω_m) from d, i. e. $\Delta\omega_m \neq 0$, is consistent with the correct behavior (ω does not follow d instantaneously). This already suggests that we also need to include $\partial^2\omega_m := \left[\frac{d^2\omega_m}{dt^2}\right]$ in the observable states in order to obtain the discrepancies required for fault detection and fault identification. One has to be aware that determining the second derivative may constitute a practical problem.

Secondly, different controllers can be modeled. PI-controllers are described by

$$\frac{dv}{dt} = c * (d - \omega_m),$$

which results in the qualitative model

$$\partial v = [d] \ominus [\omega_m]$$
$$\partial \Delta v = [\Delta v] \ominus [\Delta\omega_m].$$

With such a controller, the speed of the motor, ω, does not approach the set point asymptotically, but exhibits "overshooting" and reaches the set point in an oscillatory manner. This implies that all states in the ($[\Delta\omega_m]$, $\partial\Delta\omega_m$) representation space can occur in a correct behavior (except for the "degenerate case" 1). As for the ramp, discrepancy detection requires observation of $\partial^2\Delta\omega_m$ and the respective extensions of the component models. Basically, the important characteristic that distinguishes correct behavior from others is

$$\partial^2\omega_m = -\Delta\omega_m.$$

Discussion

The case study presented here is meant to provide an "existence proof" (as does [Dressler 95]), not more. A systematic and formal analysis of the properties and preconditions of the state-based approach to diagnosis of dynamic systems is presented in another paper. This analysis aims at determining restrictions on both the physical system to be diagnosed and the modeling formalism (possibly including the simulation algorithm) and at characterizing different model representations and prediction algorithms with respect to their diagnostic power. In the following, we summarize some of the results.

The diagnostic power of the application of a particular algorithm to a specific physical system is its ability to distinguish the correct behavior from faulty behaviors (for fault detection) and the faulty behavior from each other (fault identification). It is influenced basically by three different conditions:

- The **physics** of the device, particularly the physics of faults, determining **distinguishability of** (the actual) **behaviors**.
- The **observability** of the device, i.e. which quantities are measurable, determining **distinguishability of** (actual) **states** (and, hence, also of behaviors).
- The **model** of the device, especially its granularity w.r.t. quantities and time, determining **distinguishability of values**, states and behaviors.

For our state-based approach to diagnosis of dynamic systems, we derive more specific criteria:

- **Subsumption of sets of states**: Fault detection does not work properly, if the states of correct behavior are a superset of the states of some faulty behavior:

  $\mathrm{STATES(FAULT_i)} \subseteq \mathrm{STATES(CORR\text{-}BEHVR)}$,

  and, likewise, for fault identification.
- **Observability of inconsistent states**: Fault detection also fails if the distinction between a faulty state that is not consistent with the correct behavior and a correct state does not manifest itself in the observable variables:

  $p_{obs}(\mathrm{STATE_k(FAULT_i)}) = p_{obs}(\mathrm{STATE_j(CORR\text{-}BEHVR)})$
  $\wedge\ \mathrm{STATE_k(FAULT_i)} \notin \mathrm{STATES(CORR\text{-}BEHVR)}$,

  where p_{obs} denotes the projection of the full state representation to the observables. Again, there is the obvious analogous criterion for fault identification.
- **Model granularity**: Even in case the distinctions between different states are observable in principle, the model may fail to reveal this distinction:

  $p_{obs}(\mathrm{STATE_k(FAULT_i)}) \neq p_{obs}(\mathrm{STATE_j(CORR\text{-}BEHVR)})$
  $\wedge\ p_{obs}(\tau_q(\mathrm{STATE_k(FAULT_i)})) = p_{obs}(\tau_q(\mathrm{STATE_j(CORR\text{-}BEHVR)})$,

  where τ_q denotes the transformation to a „coarser" (e.g. qualitative) domain. This may be because the observables are lacking a landmark corresponding to a landmark of some internal variable. The electric motor case study contains another example: the difference between the correct behavior and a motor with deviating inertia, although observable with numerical measurements, is eliminated in the sign representation.

Note that the issues discussed above also affect the results of simulation-based diagnosis systems. However, the latter appear stronger because of additional information about the temporal order of the states. It is possible to determine the relationship between the techniques more precisely. Let $\mathrm{DOM}(\underline{v})$ denote the state representation for a device with a vector $\underline{v}$ of all variables and T be some temporal universe. The step

$$\tau_{set}: \{\mathrm{behvr}: T \rightarrow \mathrm{DOM}(\underline{v})\} \rightarrow P(\mathrm{DOM}(\underline{v}))$$

from a representation of behaviors as state changes over time to the sets of states occurring in a behavior,

$$\tau_{set}(\mathrm{behvr}) := \mathrm{behvr}(T),$$

is a representational transformation in the sense of our theory of multiple models for diagnosis ([Struss 92], [Struss 94]), more specifically an abstraction. This theory then tells us that the diagnoses obtained by our state-based system is a superset of those generated by the simulation-based one. For the same reason, limited observablity of variables leads to a superset of diagnoses, because the projection

$$p_{obs}: \mathrm{DOM}(\underline{v}) \rightarrow \mathrm{DOM}(\underline{v}_{obs})$$

to the subvector of observables is a representational transformation.

However, neither limited observability of variables, nor ignorance of temporal information necessarily implies that the diagnostic results are weaker. For diagnosis, we need to **observe the relevant inconsistencies**. This is the intuition behind the following concept:

Definition (Complete observability of inconsistencies)

Let $(\underline{v}, DOM(\underline{v}))$ be a representational space for behavior models and

$$R(behvr) \subset DOM(\underline{v})$$

be a relational model (state set) of a behavior. The property of complete observability of inconsistencies holds iff

$$\forall FAULT_i\ \forall s \in R(FAULT_i) \backslash R(CORRECT)$$
$$p_{obs}(s) \notin p_{obs}(R(CORRECT)).$$

Proposition 1:

Fault detection is not affected by limited observability of variables as long as complete observability of inconsistencies holds.

The condition may appear quite strong (actually, it can be weakened under certain conditions by replacing the second quantifier by an existence quantifier), but the motor example fully satisfies it, at least for single faults (note that $\mathrm{FAULT_i}$ in the definition refers to a fault of the entire device, and, hence, may correspond to multiple failures). Also note that it is a condition on (sets of) states, rather than on behaviors (state sequences).

Finally, it is worth while analyzing whether or under which conditions simulation-based diagnosis can be strictly stronger than state-based diagnosis.

The key consideration (already discussed in [Dressler 95]) is that most qualitative simulation algorithms generate a state sequence $(S_1, S_2, \ldots, S_k)$ as a possible behavior, if and only if

- each S_i is consistent with the model and
- each transition (S_i, S_{i+1}) satisfies continuity conditions.

The first condition means checking states which is what our approach does, as well. The second condition seems to become obsolete, if we assume that we obtain a gapless sequence of observations of a continuously changing system. However, one has to take into account that a gapless continuous sequence of consistent **observed** states may correspond to a sequence of internal states that contains an inconsistent state or a discontinuity. The property of complete observability guarantees that „internal inconsistencies“ would be made visible by the observed states.

Proposition 2:

Assume that

- *complete observability of inconsistencies holds, and that*
- *the sequence of observations is gapless, i.e. does not miss an actual state.*

Then simulation-based and state-based diagnosis are equivalent w.r.t. their results.

But, of course, state-based diagnosis is more efficient.

Acknowledgments

We would like to thank the members of the MBSQR group Munich, O. Dressler, F. Dummert, U. Heller and M. Sachenbacher, for their valuable contributions. This work was supported in part by the German Ministry of Education and Research (# 01 IN 50941).

References

[Dressler 95] Oskar Dressler, *On-line Diagnosis of Dynamic Systems based on Qualitative Models and Dependency-based Diagnostic Engines*, in: Working Papers of the 9th International Workshop on Qualitative Reasoning about Physical Systems (QR-95), Amsterdam, 1995.

[Dressler-Struss 96] Oskar Dressler and Peter Struss, The Consistency-based Approach to Automated Diagnosis of Devices, in: G. Brewka (ed.), Principles of Knowledge Representation, CSLI, 1996.

[Dvorak-Kuipers 92] Daniel Dvorak and Benjamin Kuipers, *Model-based Monitoring of Dynamic Systems*, in: W. Hamscher et al. (eds.), *Readings in Model-based Diagnosis*, Morgan Kaufmann Publishers, San Mateo, 1992.

[Milne et al 94] Robert Milne et al, *TIGER: real-time situation assessment of dynamic systems*, in: Intelligent Systems Engineering, Vol. 3, No 3, 1994.

[Struss 92] Peter Struss, *What's in SD? Towards a Theory of Modeling for Diagnosis*, in: W. Hamscher et al. (eds.), *Readings in Model-based Diagnosis*, Morgan Kaufmann Publishers, San Mateo, 1992.

[Struss-Dressler 89] Peter Struss, Oskar Dressler, *Physical Negation - Integrating Fault Models into the General Diagnostic Engine*, in: Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI-89), Morgan Kaufmann Publishers, San Mateo, 1989.

[Struss-Malik-Sachenbacher 96] Peter Struss, Andreas Malik, Martin Sachenbacher, *Qualitative Modeling is the Key to Automated Diagnosis*, to appear in: Proceedings of the 13th IFAC World Congress, San Francisco, 1996.