Diagnose eingebetteter Systeme

Björn Schieffer und Günter Hotz

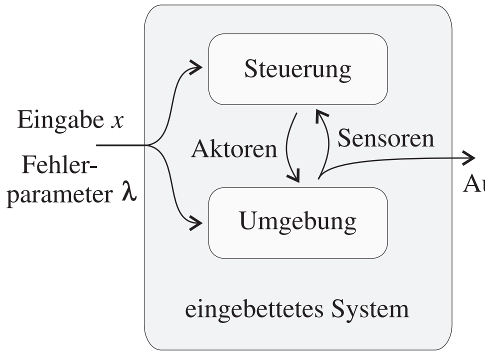

Abbildung 1: Schema eines eingebetteten Systems

1 Motivation

Wir beschäftigen uns in dieser Arbeit mit der Diagnose eingebetteter Systeme. Diese bestehen aus einem elektronischen Steuerungsteil und dessen mechanischer und hydrodynamischer Umgebung. Die Steuerung bildet über Sensoren und Aktoren mit ihrer Umwelt einen Regelkreis der gegenseitigen Beeinflußung. Beispiele eingebetteter Systeme sind moderne Automobilmotoren mit integrierter Steuerelektronik.

Abbildung 1 zeigt das Schema eines eingebetteten Systems. Ausschlaggebend für sein Verhalten sind Eingaben $x \in X$ aus dem Eingaberaum X und Fehlerparameter $\lambda \in C$ aus dem Konfigurationsraum C. Die Fehlerparameter spiegeln die Störungen des Systems wider. Wir nehmen an, es gäbe n verschiedene potentielle Störungen. Der Konfigurationsraum C ergibt sich dann als das Kreuzprodukt der Konfigurationsräume der Einzelstörungen:

$$C = C_1 \times C_2 \times \ldots \times C_n$$

Das System soll r Sensoren besitzen, von deren Meßdaten die Ausgabe $y = (y_1, y_2, \ldots, y_r) \in Y$ aus dem Ausgaberaum Y gebildet wird.

Als *Simulation* bezeichnen wir die Aufgabe, y zu bestimmen, wenn x und λ gegeben sind. Eine Simulation ist also eine Abbildung

$$sim: \qquad X \times C \rightarrow Y$$

Die prinzipielle Lösbarkeit dieses Problems zeigt uns die Natur, die die zugrunde liegenden Differentialgleichungen beim Betrieb des Systems löst. Wenn man in der Lage ist, eventuell auftretende numerische Schwierigkeiten zu beherrschen, so kann man diese *natürliche Simulation* nachvollziehen.

Wesentlich schwieriger jedoch ist das umgekehrte Problem, aus gegebener Systemausgabe y auf mögliche Eingaben x und λ zu schließen:

$$sim^{-1}: \qquad Y \rightarrow X \times C$$

Das Problem bleibt im wesentlichen das Gleiche, wenn wir die uns bekannten Eingaben x als gegeben betrachten und nur die uns unbekannten Parameter λ suchen. Wir nennen dies eine *Diagnose*. Falls sim nicht injektiv ist, so muß die Diagnose nicht eindeutig sein, es können dann mehrere Parameter λ das Systemverhalten y erklären. Eine *vollständige Diagnose* ist daher gegeben durch

$$diag: \qquad X \times Y \rightarrow \mathcal{P}(C)$$

Man spricht hierbei von einem *inversen Problem*. Im Gegensatz zur Simulation gibt es hierzu kein natürliches Vorbild. Inverse Probleme sind sehr intensiv untersucht für den Fall linearer Systeme. Ein Überblick über dafür existierende Lösungsansätze gibt Louis in [Lou89].

Der Regelkreis zwischen Steuerung und Umgebung führt dazu, daß eingebettete Systeme im allgemeinen kein lineares Verhalten zeigen. Das inverse Problem nicht-linearer Systeme ist jedoch derart schwierig, daß bis heute noch keine allgemeine Lösung gefunden werden konnte. Es existieren lediglich anwendungsspezifische Ansätze. Wir beschränken uns daher in dieser Arbeit auf die Diagnose der in [DBMB93] eingeführten Tank-Ballast-Systeme.

Da eingebettete Systeme eine Steuereinheit besitzen, wäre es naheliegend, zu ihrer Diagnose Methoden heranzuziehen, die zur Diagnose von Schaltwerken entwickelt wurden. Eingebettete Systeme haben jedoch im Gegensatz zu Schaltwerken einen kontinuierlichen Konfigurationsraum C, da schon die Konfigurationsräume C_i einzelner Störungen kontinuierlich sein können. Störungen in Schaltwerken – oft modelliert als Haftfehler von Leitungen – führen hingegen zu Veränderungen in einem endlichen, binären Zustandsraum, was sicherlich leichter zu diagnostizieren ist, als Veränderungen in einem kontinuierlichen Raum. Die Diagnose eingebetteter Systeme geht daher über die Diagnose von Schaltwerken hinaus und bedarf eigener Methoden.

Die Schwere des Diagnoseproblems führt nicht nur dazu, daß man sich spezielle Anwendungen herausgreift und nur für sie nach einer Lösung sucht. Man versucht auch, das Problem durch eine grobere Aufgabenstellung zu vereinfachen. Wir nennen diese Aufgabenstellung *qualitative Fehlereingrenzung*. Sie will klären, welche der n potentiellen Störungen auftreten, nicht aber ihre genaue Ausprägung bestimmen. Hintergrund davon ist die Annahme, daß zum untersuchten Zeitpunkt nicht alle potentielle Störungen des Systems auftreten, sondern nur einige wenige. Für den speziellen Fall, daß nur von ein oder zwei auftretenden Störungen ausgegangen wird, spricht man auch von der *Einzelfehlerannahme* beziehungsweise der *Zweifachfehlerannahme*. Eine Aussage darüber, welche der n potentiellen Störungen auftreten und welche nicht, heißt *Fehlerkandidat*. Die Menge $Cand$ aller Fehlerkandidaten

$$Cand \quad := \quad \mathcal{P}\big(\{1, 2, \ldots, n\}\big)$$

vereinfacht sich für die Einzel- und die Zweifachfehlerannahme zu den Mengen $Cand_1$ und $Cand_2$:

$$\begin{aligned} Cand_1 &:= \{1, 2, \ldots, n\} \\ Cand_2 &:= \big\{\{a, b\} \bigm| 1 \leq a < b \leq n\big\} \end{aligned}$$

Die qualitative Fehlereingrenzung qfr versucht, möglichst viele Fehlerkandidaten einer Menge $Cand' \in \{Cand, Cand_1, Cand_2\}$ auszuschließen, sodaß nur einige wenige Kandidaten, im Idealfalle sogar nur ein einziger Kandidat, übrigbleiben.

$$qfr: \qquad X \times Y \rightarrow \mathcal{P}(Cand')$$

Kern der vorliegenden Arbeit sind zwei neu entwickelte Verfahren zur qualitativen Fehlereingrenzung und zur Diagnose von Tank-Ballast-Systemen. Sie unterscheiden sich daher in der Präzision ihrer Aussage. Die qualitative Fehlereingrenzung ist in der Lage, den Fehlerort grob zu lokalisieren. Das zweite Verfahren dient dazu, den quantitativen Fehlerzustand exakt zu bestimmen. Es nutzt dazu die qualitative Fehlereingrenzung, um den zu untersuchenden Suchraum vorab stark zu reduzieren.

Zum Beurteilen der vorgestellten Verfahren nutzen wir das vom Behavior-Projekt in [DBMB93] vorgestellte Tank-Ballast-Beispiel. Für dieses System wurden dreitausend zufällig erzeugte Störungen untersucht. All diese Störungen, die sowohl aus Einzel- als auch aus Doppelfehlern bestehen, konnten mit den vorgestellten Verfahren korrekt diagnostiziert werden.

Abbildung 2: Die Brent Spar

2 Die Anwendung

Tank-Ballast-Systeme sind Bestandteil sehr großer Schiffe und schwimmender Plattformen, wie die in Abbildung 2 dargestellte *Brent Spar* oder die *Micopery 7000*, die Lasten von bis zu 14 000 Tonnen aus dem Wasser heben kann. Um dabei die Balance halten zu können, verfügt sie über 57 Ballasttanks, die mit Hilfe zweier Pumpen mit Seewasser gefüllt werden können. Auch das Verlagern des Schwerpunktes der Plattform durch Umpumpen von Wasser zwischen den einzelnen Ballasttanks ist möglich. Durch einen elektronischen Leitstand kann dieser Lastausgleich gesteuert werden.

Der Leitstand hat viele Arbeiter vor Ort ersetzt, die früher Ventile noch manuell bedient haben. Bei einem Schaden in der Anlage macht sich das Fehlen dieser *menschlichen Sensoren* bemerkbar. Fehler, die früher vom Personal schnell erkannt wurden, sind heute nur durch ein unerwartetes Verhalten der Plattform bemerkbar und können oft nicht gleich lokalisiert werden. Ein unkorrekter Ballastausgleich kann aber schwere Folgen, bis hin zum Untergang der Plattform, haben. Daher besteht ein großer Bedarf an automatischen Verfahren, die in der Lage sind, mit Hilfe der vorhandenen Sensordaten eine Diagnose zu erstellen.

Tankballastsysteme umfassen sehr unterschiedliche Komponenten. Daher müssen nicht nur mechanische Teile, wie Pumpen und Ventile mit Motoren zu ihrer Steuerung berücksichtigt, sondern auch Flüssigkeitsbewegungen analysiert und elektronische Steuerelemente beachtet werden.

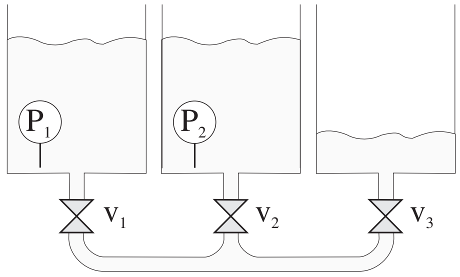

Abbildung 3: Begleitendes Beispiel

Wir extrahieren uns für spätere Demonstrationen ein kleines, aber dafür sehr anschauliches Beispiel aus dieser Anwendung. Es besteht aus drei Behältern oder Tanks, die durch Leitungen miteinander verbunden sind. Jeder Tank kann durch ein Ventil vom restlichen System entkoppelt werden. Zwei der Tanks enthalten einen Drucksensor. Schematisch ist dies in Abbildung 3 dargestellt.

Als *Eingabe* an das System betrachten wir die initialen Füllstände der Tanks und die Stellungen der drei Ventile.

$$X \quad := \quad \mathbb{R}^3 \times \{offen, geschlossen\}^3$$

Die potentiellen Störungen des Systems seien Verschiebungen dieser Ventilstellungen. Es gilt also $n = 3$. Wir wollen auch beliebige Zwischenstellungen der Ventile, ungleich der beiden Extremstellungen *offen* und *geschlossen*, erlauben:

$$C \quad := \quad [0,1] \times [0,1] \times [0,1]$$

Die Konfiguration $(0,0,0) \in C$ entspricht hierbei dem Sollzustand, das heißt dem ungestörten System.

Die Ausgabe $y = (y_1, y_2)$ des Systems sind die Druckmessungen der beiden Sensoren P_1 und P_2. Wir beobachten die Sensoren über einem Zeitintervall $[t_{start}, t_{end}]$. Jeder Sensor liefert daher eine Abbildung von $[t_{start}, t_{end}]$ nach $\mathbb{R}$.

$$M \quad := \quad Abb([t_{start}, t_{end}], \mathbb{R})$$

$$Y \quad := \quad M^2$$

Wir nehmen an, uns stünde ein Simulator $sim : X \times C \to Y$ zur Verfügung, der zu vorgegebenen initialen Füllständen und den Ventilstellungen die sich ergebenden Druckverläufe berechnet und schreiben

$$(y_1^{x,\lambda}, y_2^{x,\lambda}) \quad := \quad sim(x, \lambda)$$

$$y_1^{x,\lambda},\, y_2^{x,\lambda} : \qquad [t_{start}, t_{end}] \to \mathbb{R}$$

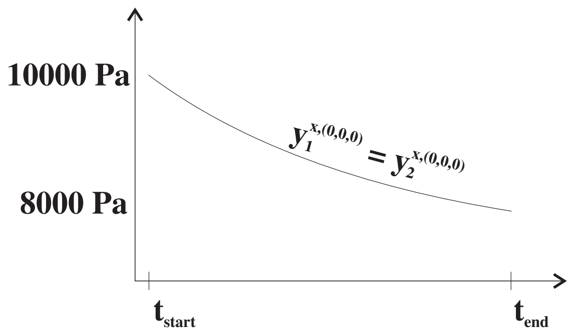

Abbildung 4: Das korrekte Systemverhalten

Unsere Aufgabe ist es nun, für eine vorgegebene Eingabe x und eine konkrete Messung der Ausgabe, $\hat{y} = (\hat{y_1}, \hat{y_2})$, herauszufinden, erstens welche Ventile verstellt sind und zweitens in welchem Maße sie verstellt sind.

Wir begnügen uns mit einer einzigen Systemeingabe x. Die Füllhöhe der ersten beiden Tanks sei darin ein Meter und die des dritten Tanks 0.4 Meter. Die Ventile seien alle geöffnet.

$$x \quad := \quad (1, 1, 0.4, offen, offen, offen)$$

Wir nehmen an, das Produkt aus der Dichte der Flüssigkeit mit der Erdbeschleunigung betrage gerade 10000 Pa pro Meter. Dann sinkt im fehlerfreien System der Druck der beiden Meßgeräte von anfänglichen 10000 Pa auf 8000 Pa. Das entspricht dem Druck, bei dem alle Tanks die gleiche Füllhöhe haben. Die entsprechende Ausgabe y ist in Abbildung 4 skizziert.

3 Fehlereingrenzung

Wir gehen im folgenden von der *Einzelfehlerannahme* aus, nehmen also an, daß nicht mehrere Ventile gleichzeitig verstellt sein können. Wir stellen uns die Aufgabe, zu bestimmen, welches Ventil defekt ist. Wir betrachten also die drei Fehlerkandidaten $Cand_1 = \{1, 2, 3\}$.

Wir gehen an die Aufgabe heran, indem wir uns überlegen, wie sich nun die Ausgabe in unserem Beispiel verändert, wenn eines der drei Ventile nicht vollständig geöffnet ist.

Falls das Ventil v_1 etwas geschlossen ist, so bietet es dem Wasserstrom einen zusätzlichen Widerstand, der erste Tank leert sich langsamer. Dadurch hat der zweite Tank die Möglichkeit, sich schneller zu entleeren. Die Druckmessungen von P_1 und P_2 sind also größer bzw. kleiner als die Sollwerte aus Abbildung 4.

Für das Ventil v_2 kehrt sich das Ganze gerade um, während sich bei einem teilweise geschlossenen Ventil v_3 beide Tanks langsamer entleeren, die Druckmessungen also beide größer als ihre Sollwerte sind.

Dieser qualitative Vergleich der Messungen mit ihren Sollwerten liefert uns also ein unterschiedliches Profil aller Fehlerkandidaten. Dies versetzt uns in die Lage, bei einer gegebenen Messung zu unterscheiden, welcher der drei Fehler für diese Messung in Frage kommt.

Kandidat	Profil
1	P_1 zu hoch, P_2 zu niedrig
2	P_1 zu niedrig, P_2 zu hoch
3	P_1 zu hoch, P_2 zu hoch

Der Vergleich einer Messung $\hat{y}_i$ mit ihren Sollwerten $y_i^{x,(0,0,0)}$ für $i \in \{1, 2\}$ ist ein Vergleich zweier Elemente des Raumes M. Um Elemente aus M miteinander vergleichen zu können, normieren wir M, etwa durch

$$\forall\, x \in M, \quad \|x\| \quad := \quad \int_{t_{start}}^{t_{end}} |x|\, dt$$

Diese Norm induziert dann die irreflexible, partielle Ordnung $<$ durch

$$\forall\, x, y \in M, \quad x < y \quad :\Longleftrightarrow \quad \|x\| < \|y\|$$

Weitere Kriterien zur Unterscheidung von Fehlerkandidaten, die wir neben dem Vergleich der Messungen mit den Sollwerten betrachten können, sind qualitative Eigenschaften der Messungen, wie Monotonie oder Krümmung. So ist zum Beispiel der Druckverlauf in Abbildung 4 linksgekrümmt. Hätten wir zum Füllstandsausgleich nicht nur die Schwerkraft genutzt, sondern uns einer zusätzlichen Pumpe bedient, so wäre der Druckverlauf hingegen linear. In diesem Falle könnte uns eine Krümmung im Druckverlauf den Ausfall der Pumpe anzeigen.

Wir nennen eine Auswahl der uns interessierenden qualitativen Eigenschaften über den Messungen einen *Spot*. Wenn wir einen solchen Spot für einen Fehlerkandidaten auswerten, das heißt, wenn wir testen, ob die darin enthaltenen Eigenschaften für den Kandidaten erfüllt sind, so erhalten wir ein *Profil* des Kandidaten bezüglich des Spots. Unterscheidet sich dieses Profil von demjenigen, das man durch die erhaltenene Messung $\hat{y}$ gewinnen kann, so darf man den Kandidaten verwerfen, denn er kann für die Messung $\hat{y}$ nicht verantwortlich sein. Zusammenfassend erhält man folgende Vorgehensweise zur qualitativen Fehlereingrenzung:

1. Wähle einen Spot sp aus
2. Bestimme das Profil jedes Kandidaten zu sp
3. Bestimme das Profil $\hat{p}$ zur Messung $\hat{y}$
4. Eliminiere alle Kandidaten, deren Profil ungleich dem Profil $\hat{p}$ ist.

Unklar ist hierbei noch der zweite Punkt: Wie lassen sich Profile zu Fehlerkandidaten bestimmen? Denn ein Kandidat faßt das Verhalten vieler Ausprägungen eines Fehlers zusammen. In unserem Beispiel steht ein Kandidat für jede mögliche Verschiebung einer Ventilstellung. Jede dieser Stellungen führt jedoch zu einem unterschiedlichen Systemverhalten. Wir interessieren uns also für die Frage, ob eine Eigenschaft für alle diese unterschiedlichen Verhalten erfüllt ist. Um diese zu klären, betrachten wir drei Möglichkeiten:

1. Man nutzt das vorhandene Wissen des Systemdesigners und läßt ihn die notwendigen Profile von Hand erstellen.
2. Man versucht, dessen Schlußfolgerungen zu automatisieren und die Profile durch *qualitative Simulation* zu erhalten.
3. Man legt ein genügend feines Raster über den Konfigurationsraum des Kandidaten und simuliert für jeden Gitterpunkt des Rasters das entsprechende Systemverhalten. Erfüllen all diese Systemverhalten die untersuchte Eigenschaft, so macht man die Annahme, daß sie auch für die Zwischenpunkte erfüllt ist und damit für den gesamten Kandidaten.

Da die erste Möglichkeit den Entwerfer benötigt und die zweite Möglichkeit für Systeme, die größer sind als unser begleitendes Beispiel, nicht zu befriedigenden Ergebnissen führt, verwenden wir in unseren Diagnosetools die dritte Möglichkeit.

Wir werden bei der Beurteilung der qualitativen Fehlereingrenzung sehen, daß diese in der Lage ist, recht viele Kandidaten auszuschließen. In vielen Fällen jedoch bleiben mehr als ein einziger Kandidat übrig. Dies alleine ist schon eine genügende Motivation für ein feineres Verfahren, das weitere Kandidaten ausschließen kann. Hinzu kommt aber noch, daß ein feineres Verfahren auch mehr Information liefern kann. Dies kann erforderlich sein, wenn man entscheiden will, ob eine laufende Operation trotz der Störung weitergeführt werden kann, oder ob man sie abbrechen muß. Wir wollen daher versuchen, die Fehlerparameter präziser zu bestimmen, als dies die qualitative Fehlereingrenzung leisten kann.

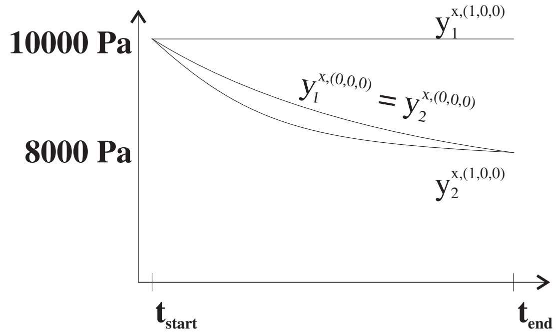

Abbildung 5: Systemverhalten für Kandidat v_1

4 Diagnose

Wir betrachten die Bestimmung der Fehlerparameter $\lambda \in C$, die eine Messung $\hat{y} \in Y$ erklären können, als einen Suchprozeß im Konfigurationsraum C. Mit $0 \leq \mu_i \leq \mu_i' \leq 1$ für $1 \leq i \leq n$ heißt

$$[\mu_1, \mu_1'] \times [\mu_2, \mu_2'] \times \cdots \times [\mu_n, \mu_n']$$

ein *Suchraum* $S \subseteq C$. Als *Eckpunkte* eines Suchraumes bezeichnen wir Punkte $(\nu_1, \nu_2, \ldots, \nu_n) \in C$ mit $\nu_i \in \{\mu_i, \mu_i'\}$ für $1 \leq i \leq n$. Die *Dimension* d eines Suchraumes ist die Anzahl der nicht fixierten Parameter

$$d \quad := \quad \#\{i \mid \mu_i \neq \mu_i'\}$$

Wir werden ein Verfahren vorstellen, das den gesuchten Wert $\lambda \in C$ durch fortwährende Verkleinerungen von Suchräumen, die λ enthalten, approximiert. Um die Suche zu beschleunigen, schalten wir die eben beschriebene, qualitative Fehlereingrenzung davor, die die Suche auf die Konfigurationsräume einiger weniger Fehlerkandidaten einschränkt.

Wir betrachten hier nur die Konfigurationsräume von Einzel- und von Doppelfehlern. Prinzipiell ist die Vorgehensweise auch auf Mehrfachfehler höherer Ordnung übertragbar. Ob dies mit vernünftigen Laufzeiten möglich ist, ist allerdings noch nicht erprobt.

4.1 Einzelfehler

Wir wollen wieder unser begleitendes Beispiel aufgreifen und einen der drei Einzelfehler exemplarisch untersuchen. Wir interessieren uns dafür, ob es eine Stellung des Ventils v_1 bei korrekter Ventilstellung von v_2 und v_3 gibt, die für eine erhaltene Messung $\hat{y} = (\hat{y}_1, \hat{y}_2)$ verantwortlich sein kann.

$$\begin{aligned} \text{Suche } \lambda \in S &:= [0,1] \times [0,0] \times [0,0] \text{ mit} \\ (y_1^{x,\lambda}, y_2^{x,\lambda}) &= (\hat{y}_1, \hat{y}_2) \end{aligned}$$

Da nur der erste Parameter eine Rolle spielt, sprechen wir bei S und seinen Teilsuchräumen auch von Intervallen. Betrachten wir uns für $\mu = (\mu_1, 0, 0) \in S$ das Systemverhalten $(y_1^{x,\mu}, y_2^{x,\mu})$. In Abbildung 5 sind die Extremfälle für $\mu = (0,0,0)$ und $\mu = (1,0,0)$ dargestellt. Diese gehen jeweils durch Variation des Parameters μ_1 zwischen 0 und 1 stetig ineinander über. Dabei gilt, daß der Druckverlauf von y_1 um so höher und der Druckverlauf von y_2 um so niedriger ist, je größer μ_1 ist:

$$\begin{aligned} \mu_1 \leq \mu_1' \;\Rightarrow& \\ y_1^{x,(\mu_1,0,0)} &\leq y_1^{x,(\mu_1',0,0)} \\ y_2^{x,(\mu_1,0,0)} &\geq y_2^{x,(\mu_1',0,0)} \end{aligned}$$

Diese Gesetzmäßigkeit stellt eine *Monotonie* des Systemverhaltens dar. Mit ihr folgt für einen beliebigen Teilsuchraum $[\mu_1, \mu_1'] \times [0,0] \times [0,0] \subseteq S$

$$\begin{aligned} \mu_1 \leq \lambda_1 \leq \mu_1' \;\Rightarrow& \\ y_1^{x,(\mu_1,0,0)} &\leq \hat{y}_1 \leq y_1^{x,(\mu_1',0,0)} \\ y_2^{x,(\mu_1,0,0)} &\geq \hat{y}_2 \geq y_2^{x,(\mu_1',0,0)} \end{aligned}$$

Diese *Schrankenbedingung* wollen wir bei der Suche ausnutzen. Sie erlaubt uns, eine *Binärsuche* zur Bestimmung von λ anzuwenden.

Der initiale Suchraum für λ ist S. Wir halbieren S fortwährend und entscheiden uns jeweils, in welchem der beiden Teilsuchräume wir die Suche fortsetzen wollen. Wir entscheiden uns für denjenigen Teilsuchraum $[\mu_1, \mu_1'] \times [0,0] \times [0,0]$, der die Schrankenbedingung für die Messungen $\hat{y}_1$ und $\hat{y}_2$ erfüllt.

Dieser Prozeß ist ein $\mathbb{R}$-analytischer Algorithmus im Sinne von [HVS95] und approximiert die Lösung λ durch die Grenzen der immer kleiner werdenden Suchräume. Gibt es jedoch zu der vorgegebenen Messung $(\hat{y}_1, \hat{y}_2)$ und dem initialen Suchraum $S \subseteq C$ keine Lösung $\lambda \in S$, so können zwei Fälle auftreten. Zum einen kann im Laufe der Berechnung die Schrankenbedingung für beide Teilintervalle verletzt werden. Daran kann man die Nichtexistenz einer Lösung im Suchraum erkennen und anzeigen. Zum andern kann der Algorithmus einen Wert λ' approximieren, der keine Lösung ist, das heißt, $(y_1^{x,\lambda'}, y_2^{x,\lambda'}) \neq (\hat{y}_1, \hat{y}_2)$.

Um dies zu erklären, betrachten wir wieder den Raum M. Zu der oben eingeführten Norm definieren wir die Äquivalenzrelation $\sim$ der unvergleichbaren Elemente in M durch

$$a \sim b \quad :\Longleftrightarrow \quad \int_{t_{start}}^{t_{end}} |a|\, dt \;=\; \int_{t_{start}}^{t_{end}} |b|\, dt$$

Diese Äquivalenzrelation liefert eine Restklasse $M_{/\sim}$, die die Ordnung $<$ des Raumes M in wohldefinierter Weise übernehmen kann.

$$\begin{aligned} \forall\, a \in M,\ \ [a] \quad &:= \quad \{ b \in M \mid a \sim b \} \\ M_{/\sim} \quad &:= \quad \{ [a] \mid a \in M \} \end{aligned}$$

$$\forall\, [a], [b] \in M_{/\sim},\ \ [a] < [b] \quad :\Longleftrightarrow \quad a < b$$

Bei dem oben beschriebenen Suchprozeß fallen die Entscheidungen zwischen den einzelnen Teilsuchräumen durch Vergleiche in M zwischen dem Verhalten an den Eckpunkten der Suchräume und den $\hat{y}_i$. Der analytische Algorithmus verzweigt daher für alle Eingaben $\big([\hat{y_1}], [\hat{y_2}]\big)$ in gleicher Weise, wie für die Eingabe $(\hat{y_1}, \hat{y_2})$ und berechnet folglich ein $\lambda' \in S$ mit

$$(y_1^{x,\lambda'}, y_2^{x,\lambda'}) \in \big([\hat{y_1}], [\hat{y_2}]\big)$$

Für sein Resultat λ' muß daher noch die Gleichheit von $y^{x,\lambda'}$, mit $\hat{y}$ überprüft werden.

4.2 Doppelfehler

Bei einem Doppelfehler tragen zwei Parameter eines Systems Werte ungleich Null. Dadurch werden aus den eindimensionalen Suchräumen der Einzelfehler nun zweidimensionale Suchräume. Die für Einzelfehler vorgestellte Binärsuche soll hierauf übertragen werden. Anstatt die Suchräume zu halbieren, müssen wir die Suchräume nun vierteln. Für jeden dieser Teilsuchräume muß wiederum entschieden werden, ob die Suche darin fortgesetzt wird oder nicht. Wir wollen das Entscheidungskriterium, das wir dazu verwenden, am Beispiel entwickeln.

Dazu suchen wir nun nach den Ventilstellungen der beiden Ventile v_1 und v_2, die ein gemessenes Systemverhalten $\hat{y} = (\hat{y_1}, \hat{y_2})$ erklären können. Wir nehmen dabei an, das dritte Ventil arbeite korrekt.

$$\text{Suche } \lambda \in [0,1] \times [0,1] \times [0,0] \text{ mit}$$
$$(y_1^{x,\lambda}, y_2^{x,\lambda}) \quad = \quad (\hat{y_1}, \hat{y_2})$$

Die im letzten Abschnitt vorgestellte Monotonie des Systemverhaltens bei der Variation des ersten Ventils gilt auch dann, wenn sich das zweite Ventil in einer beliebigen, aber festen Position befindet. Ebenso gilt die symmetrische Aussage für Variationen des zweiten Ventils bei fester Position des ersten Ventils:

$$\mu_1 \le \mu_1' \ \Rightarrow \forall\, \mu_2 \in [0,1]:$$
$$\begin{aligned} y_1^{x,(\mu_1,\mu_2,0)} \quad &\le \quad y_1^{x,(\mu_1',\mu_2,0)} \\ y_2^{x,(\mu_1,\mu_2,0)} \quad &\ge \quad y_2^{x,(\mu_1',\mu_2,0)} \end{aligned}$$

$$\mu_2 \le \mu_2' \ \Rightarrow \forall\, \mu_1 \in [0,1]:$$
$$\begin{aligned} y_1^{x,(\mu_1,\mu_2,0)} \quad &\ge \quad y_1^{x,(\mu_1,\mu_2',0)} \\ y_2^{x,(\mu_1,\mu_2,0)} \quad &\le \quad y_2^{x,(\mu_1,\mu_2',0)} \end{aligned}$$

Wenn wir diese Abschätzungen über Variationen in jeweils einem Parameter miteinander kombinieren, so erhalten wir eine Abschätzung über die gleichzeitige Variation der beiden Parameter:

$$\mu_1 \le \mu_1',\ \ \mu_2 \le \mu_2' \ \ \Rightarrow$$
$$\begin{aligned} y_1^{x,(\mu_1',\mu_2,0)} \quad &\ge \quad y_1^{x,(\mu_1,\mu_2',0)} \\ y_2^{x,(\mu_1',\mu_2,0)} \quad &\le \quad y_2^{x,(\mu_1,\mu_2',0)} \end{aligned}$$

Daraus können wir nun wieder eine Schrankenbedingung ableiten: Sei der Lösungspunkt λ Element des Suchrechtecks $[\mu_1, \mu_1'] \times [\mu_2, \mu_2'] \times [0,0]$. Dann findet man unter den vier Eckpunkten des Suchraumes sowohl eine obere als auch eine untere Schranke für $\hat{y_1}$ und $\hat{y_2}$.

$$\mu_1 \le \lambda_1 \le \mu_1',\ \ \mu_2 \le \lambda_2 \le \mu_2' \ \ \Rightarrow$$
$$\begin{aligned} y_1^{x,(\mu_1',\mu_2,0)} \quad &\ge \ \hat{y_1} \ \ge \quad y_1^{x,(\mu_1,\mu_2',0)} \\ y_2^{x,(\mu_1',\mu_2,0)} \quad &\le \ \hat{y_2} \ \le \quad y_2^{x,(\mu_1,\mu_2',0)} \end{aligned}$$

Diese Schrankenbedingung ermöglicht wiederum die Entscheidung, ob die Suche in einem Teilsuchraum fortgesetzt oder abgebrochen werden soll.

4.3 Verallgemeinerung

Damit der Suchprozeß nicht auf das Beispiel beschränkt bleibt, wollen wir eine allgemeine Schrankenbedingung für alle Systeme mit monotonem Verhalten formulieren:

> Sei $sim(x, \lambda) = (\hat{y_1}, \hat{y_2}, \ldots, \hat{y_r})$ für ein $x \in X$ und ein $\lambda \in C$. Sei weiterhin $S \subseteq C$ ein Suchraum beliebiger Dimension $d \le n$ mit $\lambda \in S$. Dann gibt es zu jeder Messung $\hat{y_i}$ für $1 \le i \le r$ zwei Eckpunkte des Suchraumes S, $\nu_{\min}, \nu_{\max} \in S$ mit
>
> $$sim(x, \nu_{\min}) \le \hat{y_i} \le sim(x, \nu_{\max})$$

Man sieht dies leicht ein, wenn wenn man sich vorstellt, daß man die gesuchten Eckpunkte $\nu_{\min}$ und $\nu_{\max}$ von λ ausgehend dadurch erreichen kann, indem man nacheinander jeweils einen einzelnen Parameter λ_i in Richtung einer seiner beiden Grenzen μ_i oder μ_i' wandern läßt, bis er diese erreicht. Dabei ändert sich das Verhalten monoton und man kann durch die richtige Wahl zwischen μ_i und μ_i' jeweils erreichen, daß das Verhalten nur zunehmen oder nur abnehmen kann. Hat man dies für alle n Parameter getan, so hat man den Eckpunkt $\nu_{\max}$ bzw. $\nu_{\min}$ des Suchraumes S gefunden.

Da die Schrankenbedingung für Suchräume beliebiger Dimension gilt, läßt sich das Prinzip der Binärsuche bei Systemen mit monotonem Verhalten für beliebige Mehrfachfehler anwenden. Wir wollen den analytischen Algorithmus abschließend skizzieren. Seine Eingabe ist ein initialer Suchraum $S \subseteq C$, eine Messung $\hat{y} = (\hat{y_1}, \hat{y_2}, \ldots, \hat{y_r}) \in Y$ und die Systemeingabe $x \in X$.

1. $\mathcal{S}$ sei eine Menge von Suchräumen, initialisiert mit $\{S\}$
2. Falls $\mathcal{S} = \emptyset$, dann wird abgebrochen, es konnte keine Lösung $\lambda \in S$ gefunden werden.
3. Entferne ein Suchraum S' aus $\mathcal{S}$ und bilde Teilräume von S'
4. Prüfe für jeden dieser Teilräume die Schrankenbedingung zur Messung $\hat{y}$. Nimm diejenigen Teilräume in $\mathcal{S}$ auf, für die die Schrankenbedingung erfüllt ist.
5. Setze die Berechnung mit dem zweiten Schritt fort.

Falls dieser Suchprozeß nicht abbricht, so konvergieren die Eckpunkte der Suchräume in $\mathcal{S}$ gegen einen Wert $\lambda \in S$. Wenn das Systemverhalten für diese Konfiguration der Messung $\hat{y}$ entspricht, so wurde eine Diagnose gefunden. Alle anderen Diagnosen für die Messung $\hat{y}$ bilden mit λ eine zusammenhängende Teilmenge in C, da das Systemverhalten monoton ist.

5 Reale Diagnose

Wir haben zum Zwecke der einfachen Darstellung in den vorigen Abschnitten einige Dinge idealisiert. Für einen realen Einsatz der entwickelten Verfahren müssen wir jedoch folgende Punkte beachten:

1. Meßgeräte liefern keine stetige Abbildungen aus dem Intervall $[t_{start}, t_{end}]$ nach $\mathbb{R}$, sondern eine diskrete Folge von Einzelmessungen, die Stützstellen der realen Abbildung sind. Wir nennen diese Folge von Einzelmessungen eine *Meßreihe* $m \in \mathbb{R}^s$.
2.
   - Das Modell, mit dem wir das reale System vergleichen, kann nicht vollständig sein.
   - Die Simulation dieses Modelles ist keine exakte, sondern nur eine ungefähre Berechnung.
   - Meßgeräte arbeiten nicht völlig präzise.

   Diese drei Einschränkungen führen dazu, daß sich das berechnete Systemverhalten vom gemessenen Systemverhalten auch dann unterscheiden kann, wenn das System fehlerfrei arbeitet. Wir fassen die genannten Effekte unter dem Begriff des *Rauschens* zusammen und fordern von unseren Verfahren Robustheit gegenüber einem begrenzten Maß an Rauschen.
3. In größeren Systemen als in unserem begleitenden Beispiel trifft man auch auf nicht-monotones Systemverhalten. Ebenso kann das oben erwähnte Rauschen die Monotonie stören. Die Suche darf bei solchen Verletzungen der Monotonie nicht zu fehlerhaften Ergebnissen führen.
4. Einem analytischen Prozeß steht eine unendliche Rechenzeit zur Verfügung. Eine reale Diagnose muß jedoch in endlicher Zeit eine Lösung finden.

Ad 1: Wir haben für den Raum M nur die Möglichkeit der Normenbildung gefordert. Dies ist für Meßreihen $m = (m_1, m_2, \ldots, m_s) \in \mathbb{R}^s$ durch die L_p-Normen ebenfalls möglich:

$$\|m\|_p \quad := \quad \sqrt[p]{\sum_{i=1}^{s} |m_i|^p}$$

Wir können daher $M := \mathbb{R}^s$ setzen ohne die Algorithmen ändern zu müssen. Wir wollen aber in dieser Arbeit weiterhin bei der intuitiveren Vorstellung von Messungen als Abbildungen eines Zeitintervalles nach $\mathbb{R}$ bleiben.

Ad 2: Um Rauschen berücksichtigen zu können, muß man die Relation der Gleichheit in M durch die Relation der *Ähnlichkeit* ersetzen. Man benötigt dazu ein Maß $L \in [0, 1]$, das beschreibt, mit welcher Präzision man arbeitet. Die beiden Extremfälle $L = 0$ und $L = 1$ spiegeln hierbei absolute Präzision und völlige Ungenauigkeit wider, das heißt:

$$a \approx_0 b \quad \Longleftrightarrow \quad a = b$$

$$a \approx_1 b \qquad\qquad \forall\, a, b \in M$$

Zwischen diesen beiden Extremen soll nun ein gleitender Übergang geschaffen werden. Wir unterscheiden dabei *absolute* und *relative* Ähnlichkeit. Bei letzteren wird die Größe der Abbildungen als Gewicht des Abstandes berücksichtigt.

$$a \approx_L^a b \quad :\iff \quad (1-L)\,||a-b|| \;\leq\; L$$

$$a \approx_L^r b \quad :\iff \quad ||a-b|| \;\leq\; L\,(||a||+||b||)$$

Der Begriff der Ähnlichkeit induziert die Robustheit von Eigenschaften gegenüber Rauschen.

> Eine Eigenschaft einer Messung heißt genau dann robust gegenüber absolutem oder relativem Rauschen der Stärke L, wenn alle Messungen, die zu ihr absolut oder relativ ähnlich sind mit Stärke L, ebenfalls die Eigenschaft besitzen.

In [Sch96] werden Verfahren beschrieben, die die Robustheit einiger wichtiger Eigenschaften von Meßreihen, wie beispielsweise Monotonie oder Krümmung und die Robustheit des Vergleichs zweier Meßreihen bestimmen.

Ad 3: Damit Verletzungen der Monotonie im Systemverhalten nicht zu fehlerhaften Berechnungen führen, weichen wir die Binärsuche auf. Keinen Fehler macht man dann, wenn man den Suchraum vollständig nach einer Lösung durchforstet. Eine vollständige Suche ergibt sich, wenn man bei jeder Aufteilung des Suchraumes alle entstehenden Teilräume weiter verfolgt.

Da dies zu aufwendig ist, ist unser Ziel, daß die Verfahren sich zwischen den beiden Extremen der reinen Binärsuche und der vollständigen Suche bewegen und dabei je nach Problemstellung mehr die eine oder andere Ausprägung wählen. Dies erreichen wir durch die Vergabe von Strafpunkten an die einzelnen Teilsuchräume. Erhält ein Teilraum viele Strafpunkte, so kann man ihn von der weiteren Suche ausschließen. In diesem Falle ergibt sich eine schnelle Binärsuche. Erhalten hingegen mehrere Teilräume nur wenige Strafpunkte, so werden diese alle weiter verfolgt. Dies führt zu der langsameren, vollständigen Suche. Die Strafpunkte entscheiden dann nur über die Reihenfolge der weiteren Durchforstung.

Ad 4: Durch den Wechsel von der Gleichheit zur Ähnlichkeit können wir die Suche dann abbrechen, wenn das Systemverhalten eines Eckpunktes eines betrachteten Suchraumes ähnlich zu $\hat{y}$ ist. Dadurch wird bei einer erfolgreichen Suche aus dem analytischen ein endlicher Prozeß, dessen Laufzeit lediglich gegen Unendlich strebt, wenn die geforderte Präzision L gegen Null geht.

Der Prozeß bleibt jedoch auch dann mit der Ähnlichkeit anstelle der Gleichheit analytisch, wenn er gegen einen Wert konvergiert, der keine Lösung darstellt. Diese Möglichkeit des analytischen Algorithmus haben wir bei der eindimensionalen Suche aufgezeigt. Durch die geschickte Vergabe von Strafpunkten sind wir jedoch in der Lage, diesen Fall zu erkennen und die Suche in endlicher Zeit abzubrechen.

Die Strafpunkte eines Suchraumes ergeben sich hierfür aus der Distanz von $\hat{y}$ zu den Systemverhalten der Eckpunkte des Suchraumes, gewichtet mit der maximalen Distanz zwischen den Systemverhalten der Eckpunkte. Dieses Gewicht strebt bei den kleiner werdenden Suchräumen gegen Null, während die Distanz von $\hat{y}$ zu den Systemverhalten der Eckpunkte von unten beschränkt ist, wenn wir uns nicht einer Lösung annähern. Dadurch wachsen die Strafpunkte bei genügend kleinen Suchräumen ins Unendliche und die Suche wird abgebrochen.

6 Beurteilung

Das vom Bundesministerium für Forschung und Entwicklung unterstützte Verbundprojekt TEX-B hat sich von 1986 bis 1990 mit der Entwicklung qualitativer Diagnoseverfahren beschäftigt, die auf de Kleer zurückgehen [dKW87]. Das Nachfolgeprojekt BEHAVIOR untersuchte von 1990 bis 1994 die Anwendbarkeit dieser Mechanismen für Tank-Ballast-Systeme. Wir wollen dieses Anwendungsbeispiel zur Beurteilung unserer Verfahren aufgreifen.

6.1 Vergleich zu Behavior

Das Verbundprojekt BEHAVIOR hat sich mit rein qualitativen Verfahren zur Diagnose beschäftigt. Insbesondere ist das betrachtete Systemmodell und damit auch das Systemverhalten qualitativer Natur. Bezogen auf unser begleitendes Beispiel kann eine Messung dann ergeben, daß der entsprechende Druck sinkt, steigt oder unverändert bleibt. Das heißt:

$$M \;:=\; \{\mathit{sinkt}, \mathit{steigt}, \mathit{unver\ddot{a}ndert}\}$$

$$Y \;:=\; M^2$$

Damit kann man unterscheiden, ob die Ventile vollständig geschlossen sind oder nicht. Zur Diagnose von teilweise geschlossenen Ventilen braucht man hingegen Beobachtungen, wie solchen, daß der Druck von P_1 *langsamer* oder *schneller* sinkt als erwartet. Solche Aussagen benötigen aber, obwohl auch sie rein qualitativer Natur sind, den Vergleich mit dem quantitativen Verhalten des Systems. Ein Diagnoseverfahren, das ein rein qualitatives Systemmodell zugrunde legt, kann diese Fehler nicht diagnostizieren. Die hier vorgestellten Verfahren sind daher viel leistungsfähiger als die Verfahren des BEHAVIOR-Projektes.

6.2 Empirische Untersuchung

Wir wollen nun die Leistungsfähigkeit der Verfahren dadurch beurteilen, daß wir für ein Beispielsystem untersuchen, ob alle darin modellierten Störungen korrekt diagnostiziert werden können. Dazu wurde in der Dissertation des ersten Autors das von BEHAVIOR eingeführte Dreitanksystem quantitativ modelliert und ein eigenes numerisches Verfahren zu seiner Simulation entwickelt, da gebräuchliche Verfahren zum Lösen von Differentialgleichungen, wie das Runge-Kutta-Verfahren, bei dieser Anwendung scheiterten [Sch96].

Abbildung 6 stellt das Beispielsystem schematisch dar. Es enthält drei Tanks mit jeweils einem Druckmesser, einem Entlüftungsrohr, zwei Schwimmschaltern und einem Ventil im Zulauf. Dazu gibt es eine Pumpe, einen Filter, der diese vor Verunreinigungen schützt und weitere Druckmesser vor und nach der Pumpe. Eine Steuereinheit kontrolliert den maximalen Fluß in den Leitungen und den minimalen Saugdruck vor der Pumpe durch das Regelventil v_8. Weiterhin gibt es mehrere sonstige Ventile, die entscheiden, ob die Tanks gefüllt oder entleert werden.

Einer der an BEHAVIOR beteiligten Partner stellt solche Tank-Ballast-Systeme her und konnte daher typische Defekte aufzählen. Dazu gehören Lecks, Verstopfungen des Filters, die im begleitenden Beispiel betrachteten Verschiebungen der Ventilstellungen, Verstopfungen der Entlüftungsrohre, klemmende Schwimmschalter, Lufteinschluß in der Pumpe und Defekte der Drucksensoren.

In einem Fehlermodell wurden alle Einzelfehler und einige exemplarische Doppelfehler dieser typischen Defekte aufgesammelt. Für drei verschiedene Operationen des Systems, nämlich Füllen der Tanks, Entleeren und einem Druckausgleich zwischen den Tanks ohne Beteiligung der Pumpe wurden jeweils eintausend Störungen gewürfelt und durch Rauschen verfremdet.

Die qualitative Fehlereingrenzung führt in allen Fällen zu einer *korrekten Diagnose*. Darunter verstehen wir eine Diagnose, deren Liste von Fehlerkandidaten denjenigen Kandidaten enthält, der der Störung entspricht, mit der das untersuchte Systemverhalten erzeugt wurde.

In etwa einem Drittel aller Fälle führt die qualitative Fehlereingrenzung zu einer *minimalen Diagnose*; das heißt, sie enthält keine überflüssigen Kandidaten. Bei diesem Ergebnis wurde schon berücksichtigt, daß es bei einigen Störungen im Rahmen der zugrunde gelegten Ähnlichkeit $\approx^r_{0.0001}$ durchaus mehrere Erklärungen für das beobachtete Systemverhalten gibt. Dies widerspricht jedoch nicht der geforderten Minimalität.

Die überflüssigen Kandidaten, die sich durch die qualitative Fehlereingrenzung nicht eliminieren ließen, konnten alle durch die anschließenden Suchverfahren beseitigt werden. Die Suchverfahren führten ausnahmelos zu korrekten Diagnosen im obigen Sinne, obwohl die zu diagnostizierenden Messungen bis zu dem der Ähnlichkeit zugrunde liegenden Grad $L = 0.0001$ verrauscht waren.

7 Ausblick

Die entwickelten Verfahren konnten bisher nur an einem fixen Beispiel, dem Dreitanksystem erprobt werden. Durch ein Designsystem für beliebige Ballast-Tank-Systeme über einer festen Menge an Grundbausteinen, ähnlich zum System Cadic für den Design integrierter Schaltkreise [HR96], sollen weitere Probanten geschaffen werden. Dieses Designsystem soll gleichzeitig als graphische Oberfläche zur Visualisierung der Verfahren dienen.

Die Diagnoseverfahren erlauben einem Systementwickler schon vor dem Bau eines Systems dessen Diagnostizierbarkeit zu untersuchen. Er soll bei der Wahl möglicher Meßpunkte von unseren Werkzeugen unterstützt werden können.

Eine Diagnostizierbarkeit, die empirisch ermittelt wurde, läßt einen Rest an Zweifeln übrig. Durch Argumentieren mit der Beschränktheit der Variation im Systemverhalten zwischen diskreten Punkten des Konfigurationsraumes könnten hingegen echte Garantien möglich werden.

Literatur

[DBMB93] O. Dressler, C. Böttcher, M. Montag, and A. Brinkop. Qualitative and Quantitative Models in a Model-based Diagnosis System for Ballast Tank Systems. In *Proc. of the International Conference on Fault Diagnosis TOOLDIAG '93, Toulouse, France*, pages 397–405, 1993.

[dKW87] J. de Kleer and B. Williams. Diagnosing multiple faults. *Artificial Intelligence*, 32:97–130, 1987.

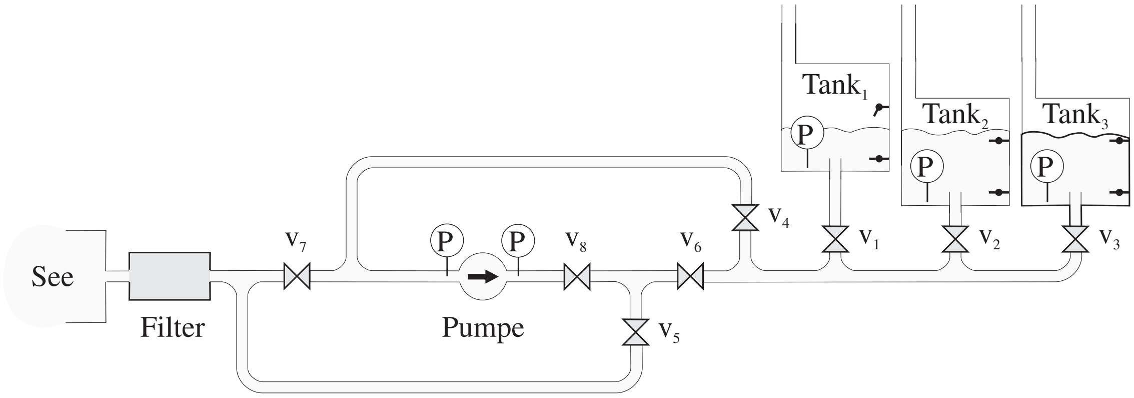

Abbildung 6: Schema eines Dreitanksystems

[HR96] G. Hotz and A. Reichert. *I. Wegener (Ed.): Highlights aus der Informatik*, chapter Hierarchischer Entwurf komplexer Systeme. Springer Verlag, 1996.

[HVS95] G. Hotz, G. Vierke, and B. Schieffer. Analytic machines. Technical Report TR95-025, Electronic Colloquium on Computational Complexity (http://www.eccc.uni-trier.de/eccc), 1995.

[Lou89] A. Louis. *Inverse und schlecht gestellte Probleme*. B.G. Teubner Verlag, Stuttgart, 1989.

[Sch96] B. Schieffer. *Diagnose komplexer Systeme am Beispiel eines Tank-Ballast-Systems*. PhD thesis, Fachbereich Informatik, Universität des Saarlandes, 66041 Saarbrücken, 1996.