

Heuristische Diagnose mit Assoziationsregeln

Vom Fachbereich 17 (Mathematik / Informatik)
der Universität Paderborn
zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
genehmigte Dissertation

von
Uwe Husemeyer

Heuristische Diagnose mit Assoziationsregeln

Vom Fachbereich 17 (Mathematik / Informatik)
der Universität Paderborn
zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
genehmigte Dissertation

von
Uwe Husemeyer

Tag der mündlichen Prüfung: 9. Mai 2001

Referent: Prof. Dr. Hans Kleine Büning
Korreferent: Prof. Dr. Franz Josef Rammig

Vorwort

Die vorliegende Arbeit entstand während meiner Zeit als wissenschaftlicher Mitarbeiter in der Arbeitsgruppe »Wissensbasierte Systeme« an der Universität Paderborn. Ich möchte mich an dieser Stelle bei Herrn Prof. Dr. Hans Kleine Büning, dem Leiter der Arbeitsgruppe, herzlich bedanken für die Unterstützung, die er mir zukommen ließ, und für die Freiheit, die er mir während der Forschungsarbeit gewährte – sowohl in wissenschaftlicher als auch in zeitlicher Hinsicht.

Obwohl eine Dissertation die Befähigung eines Einzelnen zu selbständiger wissenschaftlicher Arbeit unter Beweis stellen soll, wäre diese Arbeit nicht ohne die Mitwirkung von Kollegen und Freunden zustande gekommen: Neben anderen danke ich daher insbesondere Benno Stein, der mir während unserer engen Zusammenarbeit immer wieder wertvolle Denkanstöße gab, sowie Theodor Lettmann und Oliver Niggemann, mit denen ich viele fruchtbare Diskussionen führte. Nicht nur hierfür, aber besonders für die Hilfe bei der anstrengenden Arbeit des Korrekturlesens und der Herstellung der Kompatibilität zur deutschen Sprache danke ich meiner Freundin Christina Meyer sowie André Schekelmann und Michael Suermann.

Darüber hinaus gilt mein Dank allen Mitgliedern der Prüfungskommission, insbesondere Herrn Prof. Dr. Franz Josef Rammig für die Übernahme des Korreferats.

Uwe Husemeyer

Paderborn, im Mai 2001

Inhaltsverzeichnis

1	Einleitung	1
1.1	Kontext der Arbeit	2
1.2	Aufbau der Arbeit	3
2	Wissensakquisition mit Data Mining	5
2.1	Automatische Wissensverarbeitung	5
2.1.1	Wissen	5
2.1.2	Wissensbasierte Systeme	8
2.1.3	Wissensakquisition	9
2.2	Wissensentdeckung in Datenbanken und Data Mining	11
2.2.1	Definition	12
2.2.2	Ziele und Verfahren	14
2.2.3	Einzelschritte im Wissensentdeckungsprozeß	16
2.2.4	Data-Mining-Schritt	19
2.3	Neuer Ansatz zur Akquisition von Diagnosewissen	22
2.3.1	Diagnose technischer Anlagen	22
2.3.2	Wissensbasierte Diagnoseansätze	25
2.3.3	Integration von Data Mining in die Wissensakquisition	28
3	Assoziationsregeln	33
3.1	Regelkonzepte	34
3.1.1	Regeln	35
3.1.2	Assoziationsregelkonzept	36
3.1.3	Assoziationsregelarten	39
3.2	Boolesche Assoziationsregeln	42
3.2.1	Definition und Eigenschaften	42

INHALTSVERZEICHNIS

3.2.2	Assoziationsregelproblem	44
3.2.3	Basialgorithmus zur Regelerzeugung	47
3.3	Kategoriale Assoziationsregeln	51
3.3.1	Definition	52
3.3.2	Erzeugung kategorialer Assoziationsregeln	53
3.4	Klassifikationsregeln	55
3.4.1	Klassifikationsregelproblem	55
3.4.2	Klassifikation mit Assoziationsregeln	56
4	Diagnose hydraulischer Systeme mit Assoziationsregeln	61
4.1	Einführende Bemerkungen zur Hydraulik	62
4.1.1	Hydraulische Anlagen	62
4.1.2	Physikalische Größen in der Hydraulik	65
4.1.3	Aufgaben einer rechnergestützten Diagnose	68
4.2	Gesamtkonzept des Diagnoseansatzes	70
4.2.1	Voraussetzungen	70
4.2.2	Einzelschritte	71
4.2.3	Repräsentation der Diagnoseregeln	73
4.2.4	Zielsetzung	75
4.3	Simulation	75
4.3.1	Erstellung eines Meßplans	75
4.3.2	Erstellung eines Simulationsplans	77
4.4	Symptomerkennung	79
4.5	Diskretisierung	81
4.5.1	Motivation und Definition	81
4.5.2	Bekannte Verfahren	84
4.5.3	Neuer Diskretisierungsalgorithmus	85
4.6	Meßstellenauswahl	89
4.6.1	Motivation und Vereinbarungen	90
4.6.2	Abhängigkeitsanalyse	91

4.6.3	Bewertungsfunktionen für die Meßstellenauswahl	98
4.7	Diagnoseregelerzeugung	106
4.7.1	Vereinbarungen	106
4.7.2	Algorithmus	107
4.7.3	Regelmengenoptimierung	109
4.8	Hypothesengenerierung	111
4.8.1	Rahmenkonzept und Problemfelder	112
4.8.2	Konfidenzverrechnung in MYCIN	114
4.8.3	Globaler Ansatz zur Konfidenzverrechnung	117
5	Evaluierung des Diagnoseansatzes	121
5.1	Testanlage und Parameter für die Lernphase	122
5.2	Testverfahren für die Diagnosephase	125
5.3	Testergebnisse	127
5.3.1	Regelanzahl	127
5.3.2	Diagnosegüte	130
5.3.3	Fazit	134
6	Zusammenfassung und Ausblick	137
A	Mathematischer Anhang	139
A.1	λ -Maße	139
A.2	Informationsgehalt	140
B	Ergebnistabellen	143
	Literaturverzeichnis	151

INHALTSVERZEICHNIS

1 Einleitung

Innerhalb eng begrenzter Aufgabengebiete werden für Problemstellungen, zu deren Lösung umfangreiches Spezialwissen erforderlich ist, typischerweise sogenannte *wissensbasierte Systeme* oder *Expertensysteme* eingesetzt. Die wichtigsten Faktoren für das »intelligente« – d. h. dem Menschen nachempfundene – Problemlösungsverhalten dieser Systeme sind zum einen die Eignung der eingesetzten Schlußfolgerungsmethoden und zum anderen der Umfang, die Qualität und die Repräsentation des zur Verfügung stehenden Domänenwissens.

Die vorliegende Arbeit beschäftigt sich schwerpunktmäßig mit dem zweiten Aspekt: der Akquisition und Verarbeitung von Domänenwissen in Expertensystemen. Das Anwendungsgebiet ist die Fehlererkennung (*Diagnose*) für komplexe hydraulische Anlagen. Um hier zu einem leistungsfähigen Gesamtkonzept zu gelangen, wird die Verknüpfung von Techniken aus den Gebieten der wissensbasierten Systeme und des Data Minings vorgenommen. Das Ergebnis ist ein neuer Ansatz für die vollständig automatische Generierung von Diagnosesystemen, in denen das anlagenspezifische Diagnosewissen durch heuristische Assoziationsregeln repräsentiert wird.

Inhaltlich gliedert sich die Arbeit in drei Teile:

1. Motivation eines neuen Ansatzes zur automatischen Akquisition von heuristischem Diagnosewissen und Diskussion von hierfür geeigneten Assoziationsregelalgorithmen.
2. Einbettung des Akquisitionsansatzes in ein vollständiges Konzept zur Diagnose komplexer hydraulischer Anlagen und detaillierte Beschreibung der Einzelschritte.
3. Evaluierung des Ansatzes durch die Auswertung von Ergebnissen, die mit einer prototypischen Implementierung zur Lösung von Diagnoseproblemen erzielt wurden.

In Abschnitt 1.1 wird der Kontext der Arbeit kurz erläutert, und in Abschnitt 1.2 wird ihr Aufbau im Überblick dargestellt.

1.1 Kontext der Arbeit

Die Diagnose, auch Diagnostik genannt, ist ein Prozeß des Rückschlusses von beobachtbaren Wirkungen (Symptomen) auf ihre unbekannten Ursachen (Fehler oder Krankheiten). Im technischen Anwendungsbereich kann ein Diagnoseproblem zum Beispiel darin bestehen, die Ursachen für die Fehlfunktion (Störung) einer technischen Anlage zu finden. In der Regel ist ein defektes Bauteil für eine solche Störung verantwortlich. Zur Identifikation dieses Bauteils können verschiedene Strategien verfolgt werden, etwa die Messung von charakteristischen physikalischen Größen und darauf basierende Schlußfolgerungen. Dieser Vorgang wird *Diagnoseprozeß* genannt und erfordert von einem Fachexperten üblicherweise viel Erfahrungswissen, insbesondere, wenn es sich um eine komplexe Anlage mit vielen Bauteilen handelt.

Für die automatische Lösung von Diagnoseproblemen werden sogenannte *Diagnosesysteme* eingesetzt. Dies können wissensbasierte Systeme (Expertensysteme) sein, die das Wissen eines oder mehrerer Fachexperten abbilden und durch geeignete Schlußfolgerungsmethoden selbständig zu Problemlösungen gelangen. Da das Diagnosewissen gewöhnlich spezialisiertes Fachwissen darstellt, ist der Wissenserwerb die Hauptschwierigkeit bei der Entwicklung von Expertensystemen (vgl. [PGPB96]). Aber auch die Wissensverarbeitung hat ihre Grenzen: Am häufigsten werden modellbasierte oder heuristische Diagnoseansätze gewählt, wobei beide Ansätze Vor- und Nachteile aufweisen. Modellbasierte Diagnosesysteme verarbeiten kausale Ursache-Wirkungs-Zusammenhänge des Problembereichs; hierdurch erlangen sie eine hohe Problemlösungsfähigkeit, sind aber sehr laufzeitintensiv. Dagegen basieren heuristische Systeme auf der Verarbeitung von assoziativem Wissen, das aus einfachen Erfahrungsregeln für das Vorliegen von Fehlern bei Beobachtung bestimmter Symptome besteht. Heuristische Verfahren sind zwar schnell, allerdings ist der Wissenserwerb oft aufwendig und fehleranfällig (vgl. [Pup91]).

In diesem Spannungsfeld zeigt die vorliegende Arbeit einen neuen Ausweg: Durch die Anwendung von Data-Mining-Verfahren werden die Vorteile modellbasierter und heuristischer Diagnoseansätze miteinander verbunden. Dazu wird ein Wissensakquisitionskonzept entwickelt, das in einem zweistufigen Prozeß auf kausale Verhaltensmodelle des Problembereichs zurückgreift, um das hierin implizit vorhandene diagnoserelevante Wissen in explizite heuristische Assoziationsregelmodelle zu transformieren. Diese Modelltransformation ist durch den Einsatz von Simulations- und Data-Mining-Verfahren vollständig automatisierbar und wird in dieser Arbeit für den Bereich hydraulischer Anlagen detailliert beschrieben.

Abbildung 1.1 stellt das entwickelte Wissensakquisitions- und Diagnosekonzept im Überblick dar. Der linke Teil zeigt die wichtigsten Phasen beim Aufbau eines Diagnosesystems (Simulations- und Regellernphase), der rechte Teil zeigt die

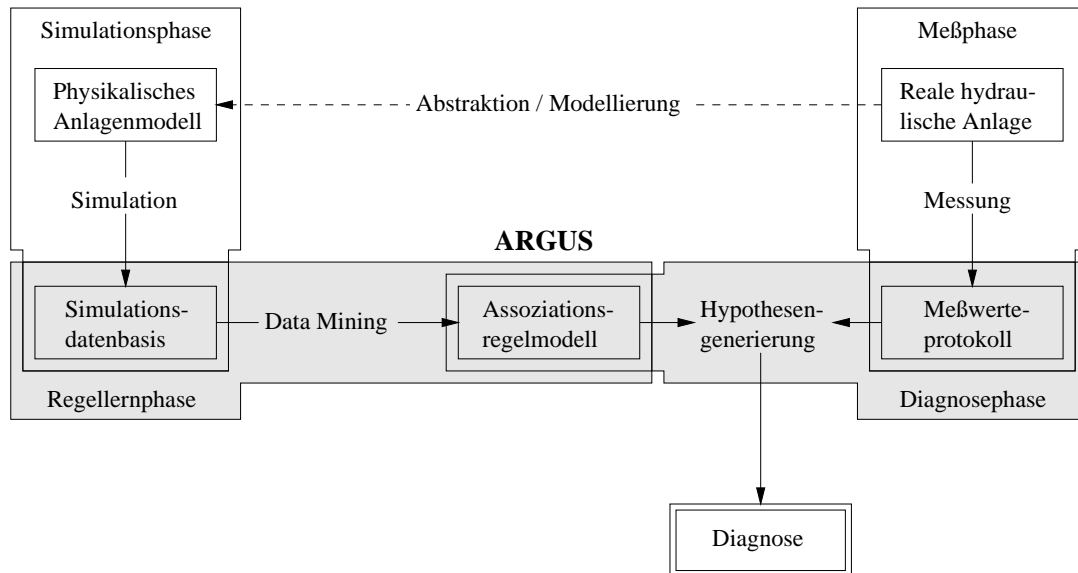


Abbildung 1.1: Automatische Generierung und Anwendung eines heuristischen Diagnosesystems unter Verwendung von Data-Mining-Verfahren

wichtigsten Phasen bei seiner Anwendung im Störfall (Meß- und Diagnosephase). Die vorliegende Arbeit behandelt schwerpunktmäßig den grau hinterlegten Bereich. Sie beschreibt dabei neue Lösungen für die Datengewinnung, die Modelltransformation sowie die Regelanwendung und faßt die Ergebnisse zusammen, die im Rahmen einer Evaluierung mit dem prototypischen Diagnosesystemgenerator ARGUS (Assoziationsregelgenerator für die heuristische Diagnose) erzielt wurden.

1.2 Aufbau der Arbeit

Kapitel 2 führt zunächst in die Themenfelder automatische Wissensverarbeitung, Wissensentdeckung in Datenbanken und Data Mining ein. Anhand des Problems der Diagnose technischer Systeme wird anschließend das Konzept zur automatischen Akquisition von heuristischem Diagnosewissen für komplexe technische Anlagen motiviert.

Kapitel 3 befaßt sich mit Assoziationsregeln als Repräsentation für heuristisches Diagnosewissen. Es werden verschiedene aus der Literatur bekannte Regelkonzepte gegeneinander abgegrenzt und hinsichtlich ihrer Ausdrucksmöglichkeiten systematisch eingeordnet. Als weiterer wichtiger Aspekt wird die Erzeugung vollständiger Regel-

1 EINLEITUNG

mengen betrachtet, die den Kriterien zur Anwendung für Diagnoseprobleme genügen.

Kapitel 4 stellt die Umsetzung des neuen Ansatzes zur automatischen Generierung von heuristischen Diagnosesystemen für hydraulische Anlagen vor. Es erfolgt zunächst eine Einführung in die Besonderheiten der Hydraulikdomäne. Einer Kurzdarstellung des vollständigen Diagnosekonzepts schließt sich die detaillierte Betrachtung seiner Einzelschritte an; dabei werden sowohl theoretische als auch praxisrelevante Aspekte beschrieben.

Kapitel 5 evaluiert den Diagnoseansatz. Dazu wird die objektive Problemlösungsfähigkeit von Diagnosesystemen bewertet, die mit dem Demonstrationsprototyp ARGUS für hydraulische Beispielanlagen automatisch erzeugt wurden. Es wird insbesondere auf die Variation der Parameter zur Regelgenerierung sowie ihre Auswirkungen auf die Regelanzahl und die Diagnosegüte eingegangen.

Kapitel 6 faßt die Ergebnisse dieser Arbeit zusammen und gibt einen Ausblick auf mögliche Erweiterungen der dargestellten Konzepte.

Im **Anhang** werden die Herleitungen für einige aus der Literatur bekannte und in dieser Arbeit benutzte Formeln aus den Bereichen Informationstheorie und Statistik ausgeführt, um die formalen Argumentationsketten aus den vorangegangenen Kapiteln zu vervollständigen. Des weiteren sind die Ergebnisse der Evaluierungsphase in Tabellenform abgebildet.

2 Wissensakquisition mit Data Mining

Dieses Kapitel motiviert den Gegenstand der vorliegenden Arbeit: einen neuen Ansatz zur automatischen Wissensakquisition für Expertensysteme. Hierbei werden Techniken aus den Gebieten der wissensbasierten Systeme und des Data Minings zu einem Konzept verknüpft, das die vollständig automatische Generierung von Diagnosesystemen für komplexe technische Anlagen ermöglicht.

Einteilung des Kapitels: Zunächst wird in die beteiligten Themenfelder eingeführt. Dazu beschreibt Abschnitt 2.1 die Wissensakquisition als eines der wichtigsten Problemfelder beim Aufbau von Expertensystemen. Abschnitt 2.2 behandelt die Wissensentdeckung in Datenbanken bzw. das Data Mining als effiziente Möglichkeit, Wissen aus Daten zu extrahieren. Auf der Grundlage dieser Ausführungen wird in Abschnitt 2.3 der neue Wissensakquisitions- und Diagnoseansatz entwickelt.

2.1 Automatische Wissensverarbeitung

Wissensverarbeitung ist ein Teilgebiet der künstlichen Intelligenz und befaßt sich als Neuroinformatik mit der Abbildung menschlicher Denkprozesse sowie im Bereich der wissensbasierten Systeme mit der Manipulation von formalisiertem Wissen (vgl. [Hel96]). Diese Arbeit beschäftigt sich nur mit dem zweiten Gebiet.

Da für den grundlegenden Begriff »Wissen« keine allgemeingültige Definition existiert, erläutert Abschnitt 2.1.1 seine Verwendung im Kontext der vorliegenden Arbeit. Abschnitt 2.1.2 skizziert die Besonderheiten wissensbasierter Systeme. Diese Gruppe von Softwaresystemen ist darauf spezialisiert, Wissen explizit zu repräsentieren und zur Lösung von Problemen automatisch zu verarbeiten. Eine wesentliche Voraussetzung für die Problemlösungsfähigkeit wissensbasierter Systeme ist die Akquisition geeigneten Wissens; hierauf geht Abschnitt 2.1.3 näher ein.

2.1.1 Wissen

Im Bereich der Datenverarbeitung wird in der Regel zwischen den drei Ebenen *Daten*, *Information* und *Wissen* unterschieden (siehe z. B. [Ums92]). Eine übliche Abgrenzung dieser Begriffe ist die folgende:

- *Daten* sind Zeichen- bzw. Bitfolgen, die definierten Syntaxregeln unterliegen und der interpretationsfreien Beschreibung von Objekten oder Ereignissen dienen.
- *Informationen* sind interpretierte Daten und stellen einzelne Aussagen zu Sachverhalten in ihrem jeweiligen Kontext dar. Der Kontext ist für alle Informationsempfänger identisch.
- *Wissen* ist verknüpfte und verstandene Information, und wird durch Erfahrung, Logik oder durch Kausalität gewonnen. Es ist das Ergebnis von Erkenntnisprozessen und kann zur Lösung von Problemen genutzt werden.

Beispiel 2.1: In Anlehnung an [HSA99] lassen sich die Begriffe Daten, Information und Wissen anhand eines Telefonbuchs verdeutlichen: Daten über Anschlußnummern sind in Telefonbüchern systematisch zusammengefaßt und stellen in diesem Kontext Informationen dar. Aber erst mit dem Wissen darüber, wie man in einer alphabetisch sortierten Liste sucht, wird es möglich, eine Nummer effizient zu finden – beispielsweise mit binärer Suche. □

Die drei Ebenen haben verschiedene Eigenschaften, von denen aus Informatik Sicht die Ausdrucks- und Nutzungsmöglichkeiten sowie der notwendige Verarbeitungsaufwand besonders wichtig sind. Die Ausdrucks- und Nutzungsmöglichkeiten nehmen in der genannten Reihenfolge zu: Daten sind rein deskriptiv, Informationen sind bereits systematisiert und Wissen ist operationalisierbar, so daß Erkenntnisse genutzt werden können (z. B. um zukünftige Ereignisse vorherzusagen oder um Probleme effizient zu lösen). Gleichzeitig nimmt aber auch der Verarbeitungsaufwand zu. Die automatische Erhebung und Verarbeitung von Daten und Informationen kann in der Regel noch effizient durchgeführt werden. Auch Wissen kann – wenn es geeignet formalisiert ist – automatisch verarbeitet werden, der dazu benötigte Aufwand ist jedoch üblicherweise sehr viel höher. Abbildung 2.1 faßt die Ausführungen zusammen.

Wissensarten

Die Unterteilung des Wissensbegriffs in verschiedene Wissensarten ist von der jeweiligen Sichtweise sowie der Zielrichtung der Wissensverwendung abhängig. Hierzu existieren in der Literatur (siehe z. B. [DD87], [TS89], [HL90], [KL90]) zahlreiche Unterscheidungsmerkmale, von denen im folgenden diejenigen kurz erläutert werden, die im Zusammenhang mit der automatischen Wissensverarbeitung von Bedeutung sind.

Implizites vs. explizites Wissen: Implizites Wissen ist zum einen unbewußtes Wissen, das sich nicht oder nur schlecht verbal bzw. formal beschreiben läßt; es ist daher

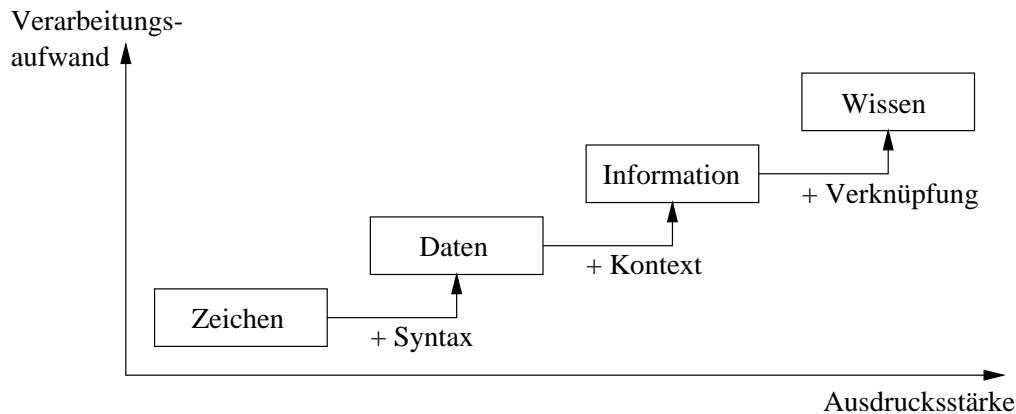


Abbildung 2.1: Zum Wissensbegriff und seiner Charakterisierung

besser mit dem Begriff *Können* ausgedrückt. Zum anderen bezeichnet es formales Wissen, das zwar vorhanden, aber nicht unmittelbar erkennbar ist (z. B. Problemlösungswissen, das in großen Datensammlungen oder in einem Algorithmus »versteckt« ist). Explizites Wissen hingegen ist jede Art von Wissen, das in einer zur automatischen Wissensverarbeitung geeigneten Wissensrepräsentationsform vorliegt.

Allgemeines vs. spezielles Wissen: Allgemeinwissen ist Hintergrundwissen, das in der Breite unterschiedliche Wissensgebiete abdeckt und kann oft in unvorhergesehenen Situationen genutzt werden. Spezialwissen geht dagegen in die Tiefe und dient nur zur Lösung von definierten Aufgaben innerhalb eines eng umgrenzten Problem-bereichs (z. B. zur Fehlerdiagnose von hydraulischen Anlagen mit einem bestimmten Aufbau).

Deklaratives vs. prozedurales Wissen: Deklaratives Wissen (*»knowing that«*) ist das Wissen um Fakten und deren Beziehungen untereinander, also die Klassifizierung und Strukturierung von Wissens-elementen. Prozedurales Wissen (*»knowing how«*) verkörpert konkrete Handlungsanweisungen zur Lösung eines Problems unter gegebenen Voraussetzungen. Hierbei kann es sich z. B. um Algorithmen (implizites prozedurales Wissen) oder um Regeln mit Aktionsteil (explizites prozedurales Wissen) handeln.

Im Kontext der vorliegenden Arbeit wird noch die folgende Unterscheidung wichtig sein (vgl. z. B. [Pup91]).

Kausales vs. heuristisches Wissen: Kausales (*»tiefes«*) Wissen repräsentiert Zusammenhänge, die aus grundlegenden, wissenschaftlich anerkannten Prinzipien gewonnen werden und besitzt dadurch für Schlußfolgerungen eine zuverlässige Erklärungs-

fähigkeit. Es basiert z. B. auf Modellen zur Beschreibung sich ändernder physikalischer Größen, wie etwa Differentialgleichungssystemen. Heuristisches (»flaches«) Wissen ist dagegen auf Erfahrungswissen gegründet, das beim Lösen von konkreten Problemen erworben wurde. Es kann dabei (eventuell unsichere) Beziehungen zwischen Beobachtungen repräsentieren, ohne diese zu erklären. Beispiele hierfür sind Fallsammlungen und Diagnoseregeln, die von Symptomen auf Fehler schließen lassen.

2.1.2 Wissensbasierte Systeme

Üblicherweise lassen sich schlecht strukturierte Problembereiche, in denen anstatt einer präzisen Theorie eher fragmentarisches bzw. empirisches Wissen vorliegt, mit konventionellen Datenverarbeitungsprogrammen nicht oder nur unzulänglich bearbeiten (vgl. [Wac93], [HL90]). Hier ist das implizit in Form von relativ starren Algorithmen abgelegte Problemlösungswissen prozedural und damit zu unflexibel, um zu befriedigenden Ergebnissen gelangen zu können.

Im Gegensatz dazu sind für »diffuse« Problembereiche sogenannte *wissensbasierte Systeme* (WBS) besser geeignet, weil sie explizites, meist deklaratives Wissen auswerten und über Verfahren verfügen, die eine Verknüpfung von Wissenselementen zu neuen Schlußfolgerungen ermöglichen (siehe [Pup91], [HSA99]). Durch eine klare Trennung zwischen dem zur Problemlösung benötigten Wissen (*Wissensbasis*) einerseits und der Wissensverarbeitungsstrategie (*Inferenzkomponente*) andererseits können Wissensinhalte leicht geändert und zur flexiblen Steuerung der Lösungssuche eingesetzt werden.

Eine spezielle Art von wissensbasierten Systemen sind die sogenannten *Expertensysteme*, die über zusätzliche Komponenten zur Benutzerinteraktion, insbesondere zur Wissensakquisition und zur Erklärung der erzielten Ergebnisse, verfügen ([RP99]). Innerhalb eines definierten und eng begrenzten Problembereichs wird von Expertensystemen das Spezialwissen und die Schlußfolgerungsfähigkeit von qualifizierten Fachleuten (Experten) nachgebildet, so daß ihr Verhalten manchmal als »intelligent« bezeichnet wird.

Aufgrund der Trennung von Wissensbasis und Inferenzkomponente ergeben sich einige Eigenschaften, die nach [Pup91] als typisch für wissensbasierte Systeme angesehen werden:

- *Transparenz*: Das Systemverhalten ist in relativ einfacher Weise nachvollziehbar und durch das benutzte Wissen erklärbar.
- *Flexibilität*: Wissen kann bei Bedarf hinzugefügt, verändert oder entfernt werden.

- *Benutzerfreundlichkeit*: Der Anwender benötigt üblicherweise kein programmiersprachliches Vorwissen.
- *Kompetenz*: Wissensbasierte Systeme verfügen in ihrem Anwendungsbereich über eine hohe Problemlösungsfähigkeit.

Anwendungsgebiete wissensbasierter Systeme

Die wichtigsten Anwendungsgebiete wissensbasierter Systeme befinden sich in den drei Bereichen Analyse, Synthese sowie Mischformen hiervon (vgl. [Boo88], [HK89]). Sie werden im folgenden kurz erläutert.

Analyse: Die Analyse umfaßt die Untersuchung und Identifizierung von existierenden Objekten anhand ihrer Eigenschaften. Dazu wird aus einer endlichen Menge vorgegebener Alternativen eine Lösung ausgewählt. Die wichtigsten Analyseaufgaben bestehen in der Klassifikation (Zuordnung von Objekten zu Klassen) und in der Diagnose (Zuordnung von Symptomen zu Fehlern oder Krankheiten).

Synthese: Bei der Synthese werden vorher noch nicht existierende Objekte durch die Kombination von vorgegebenen Einzelteilen oder Teillösungen zusammengesetzt, wobei die so entstandenen Lösungen in der Regel bestimmten Randbedingungen genügen müssen. Es existieren oft sehr viele oder sogar unendlich viele Lösungsalternativen. Zu den wichtigsten Syntheseaufgaben zählen die Konfiguration (Kombination relativ weniger Komponenten), das Design (Kombination vieler oder Entwurf neuer Komponenten) und das Planen (Zusammenstellen von Aktionen zu Handlungsabläufen).

Mischformen: Die Mischformen enthalten sowohl Analyse- als auch Syntheseanteile, wie z. B. die Reparatur (Diagnose und anschließende Aufstellung eines Reparaturplans) oder die Simulation (Synthese von Verhaltenskomponenten und Ableitung des Gesamtverhaltens).

Für die Erstellung wissensbasierter Systeme existieren eine Reihe von allgemeinen Methoden und Techniken, die sich in die Bereiche Wissensrepräsentation (formalisierte Darstellung), Wissensakquisition (Wissenserhebung) und Wissensverarbeitung (Verknüpfung und Schlußfolgerung) einteilen lassen. Im konkreten Fall erweist sich die Wissensakquisition häufig als Hauptschwierigkeit (siehe z. B. [HSA99]). Dieses Problem wird im folgenden Abschnitt 2.1.3 näher besprochen.

2.1.3 Wissensakquisition

Mit dem Begriff *Wissensakquisition* wird im allgemeinen ein Prozeß von der Erhebung problembereichsbezogenen Wissens aus verschiedenen Wissensquellen bis hin

zur Umsetzung in eine operationale Wissensbasis und deren inhaltlicher Wartung bezeichnet (vgl. [Wac93]). Die Wissenserhebung dient im Umfeld der Expertensysteme zur Formalisierung von spezialisiertem Expertenwissen. Hierzu können nach [Pup91] drei grundsätzliche Vorgehensweisen unterschieden werden:

- *Indirekte Wissensakquisition*: Ein sogenannter Wissensingenieur, der über entsprechende Kenntnisse der Wissensmodellierung verfügt, befragt einen Experten und formalisiert anschließend das so gewonnene Wissen.
- *Direkte Wissensakquisition*: Über eine Wissensakquisitionskomponente des Expertensystems formalisiert der Experte sein Wissen selbst.
- *Automatische Wissensakquisition*: Das Expertensystem nutzt maschinell verarbeitbare Daten- bzw. Wissensquellen und extrahiert hieraus mit automatischen Lernverfahren das Wissen selbständig.

Bei der indirekten Akquisition liegt das Hauptproblem für den Wissensingenieur darin, ein geeignetes mentales Modell des Anwendungsgebiets und der Schlußfolgerungsprozesse des Experten zu konstruieren (vgl. [KL90]). Der Wissensingenieur benötigt neben Allgemeinwissen auch viel fachbezogenes Vorwissen, um mit dem Experten überhaupt kommunizieren und die grundlegenden Zusammenhänge erkennen zu können. Des weiteren sind Experten häufig nicht in der Lage, ihr Wissen explizit zu machen: Sie artikulieren sich eher in erlebten Situationen und Fallbeispielen als in Abstraktionen (vgl. [Wac93]). Die indirekte Wissensakquisition kann hierdurch sehr zeitaufwendig und fehleranfällig werden.

Aber auch die direkte Wissensakquisition ist nicht unproblematisch: Sie setzt neben einer leistungsfähigen und komfortablen Akquisitionskomponente insbesondere die Fähigkeit des Experten voraus, sein Wissen zu formalisieren und zu strukturieren. An die Werkzeuge müssen daher hohe Ansprüche gestellt werden: Sie sollten eine dem Experten bekannte Wissensrepräsentation verwenden, die Eingabe unterstützen sowie die Konsistenz des Wissens gewährleisten. Dies kann jedoch in der Regel nur durch eine weitgehende Spezialisierung der Werkzeuge auf den jeweiligen Anwendungsbereich erreicht werden (vgl. [Pup90]).

Liegen maschinell verarbeitbare Daten als Wissensquelle vor, kann der Einsatz automatischer Wissensakquisitionsmethoden sinnvoll sein. Im übrigen stellt dies die einzige Möglichkeit zum Aufbau einer Wissensbasis dar, falls keine (geeigneten) Experten zur Verfügung stehen. Bei der automatischen Wissensakquisition wird der Wissenserwerb allein mit Hilfe maschineller Lernverfahren durchgeführt. Ein generelles Problem der klassischen maschinellen Lernverfahren ist allerdings, daß sie wegen ihrer Laufzeitkomplexität nicht für den Umgang mit sehr großen Datenmengen

geeignet sind (vgl. z. B. [DRSS97], [BM99]). Um auch in solchen Fällen den automatischen Aufbau einer möglichst vollständigen und qualitativ hochwertigen Wissensbasis gewährleisten zu können, müssen neue Lösungen gesucht werden.

Ein Forschungsbereich, der sich auf die Gewinnung von Wissen aus sehr großen Datenbeständen spezialisiert hat, ist die sogenannte *Wissensentdeckung in Datenbanken* bzw. das *Data Mining*. Diese Arbeit beschäftigt sich mit der Frage, ob die hier eingesetzten Modellierungs- und Entdeckungstechniken einen Ausweg für das im vorherigen Absatz geschilderte Problem bieten und zur automatischen Wissensakquisition für Expertensysteme nutzbar gemacht werden können. Bevor dieser Gedanke in Abschnitt 2.3 zu einem neuen Wissensakquisitionskonzept für heuristisches Diagnosewissen konkretisiert wird, gibt der folgende Abschnitt 2.2 zunächst einen einführenden Überblick in die Themengebiete KDD und Data Mining.

2.2 Wissensentdeckung in Datenbanken und Data Mining

In den verschiedensten informationsverarbeitenden Aufgabengebieten entstehen enorme Mengen von Daten, z. B. in der Wissenschaft, in der Wirtschaft, in der öffentlichen Verwaltung und zukünftig möglicherweise auch im privaten Bereich. Viele dieser Daten sind jedoch nicht unmittelbar von Nutzen, da das eigentlich enthaltene Wissen von uninteressanten Daten »verdeckt« wird, so daß Zusammenhänge oder Strukturen nicht direkt erkennbar sind. Diese Erkenntnis war die Motivation für die Entwicklung effizienter Konzepte zur automatischen Datenanalyse mit dem Ziel, das in den Daten verborgene implizite Wissen aufzufinden und explizit darzustellen.

Viele dieser Verfahren werden seit Beginn der neunziger Jahre unter dem Begriff *Wissensentdeckung in Datenbanken* (engl. *Knowledge Discovery in Databases*, *KDD*) zusammengefaßt. Es handelt sich hierbei nicht nur um neue Methoden – im Gegenteil: Viele sind auch in den Gebieten Statistik, Mustererkennung, künstliche Intelligenz, maschinelles Lernen, Datenvisualisierung und Datenbanken bekannt. Neu ist vielmehr ihre konsequente Ausrichtung auf die effiziente Verarbeitung sehr großer Datenmengen.

Während mit der Wissensentdeckung in Datenbanken ein ganzheitlicher, interaktiver Prozeß bezeichnet wird, der sich im Dialog mit dem Benutzer von der Auswahl der Datenquellen bis zur Interpretation der Ergebnisse erstreckt, ist *Data Mining* hierin die zentrale Phase der automatischen Erkennung von Mustern und Zusammenhängen. Da fast alle Data-Mining-Verfahren nur spezielle Eingabeformate verarbeiten können, sind sie in zusätzliche Aktivitäten zur Datenvorverarbeitung und -nachbearbeitung eingebettet.

Die weitere Einteilung ist wie folgt: Abschnitt 2.2.1 gibt eine gängige Definition für die Wissensentdeckung in Datenbanken und grenzt KDD gegenüber anderen automatischen Datenanalyseverfahren ab. Die Ziele und Verfahren des KDD-Prozesses erläutert Abschnitt 2.2.2. Abschnitt 2.2.3 geht auf den schrittweisen Ablauf zur Lösung einer Wissensentdeckungsaufgabe im ganzen ein, bevor in Abschnitt 2.2.4 der Data-Mining-Schritt detaillierter betrachtet wird.

2.2.1 Definition

Mit dem Begriff *Wissensentdeckung* wird ein Prozeß mit dem Ziel des Erkenntnisgewinns bezeichnet. Liegt diesem Prozeß eine Datenbasis in automatisch verarbeitbarer Form zugrunde, kann ein Softwaresystem den Erkenntnisgewinn des Anwenders unterstützen. Da ein Programm keine kreativen Fähigkeiten besitzt, beruht dieser Prozeß lediglich auf der expliziten Darstellung der impliziten Wissensinhalte. Das so herausgearbeitete Wissen ist unter dieser Annahme also bereits in den Daten vorhanden und läßt sich durch Transformationen wie Interpretation, Verknüpfung oder Mustererkennung aufdecken.

Diese Sichtweise wird auch bei der Modellvorstellung zur Wissensentdeckung in Datenbanken eingenommen: Hiermit werden die Schritte eines Transformationsprozesses bezeichnet, der mit (semi-) automatischen Verfahren wertvolles Wissen *explizit* zu machen sucht, das in unter Umständen sehr großen Datenbeständen *implizit* enthalten ist. Die in der Literatur allgemein anerkannte Festlegung des Begriffs »Wissensentdeckung in Datenbanken« stammt von Fayyad et al. [FPSU96]; er wird in der folgenden Definition 2.1 nach [Wro98] zitiert.

Definition 2.1 (Wissensentdeckung in Datenbanken (KDD)): *Wissensentdeckung in Datenbanken* ist der nichttriviale Prozeß der Identifikation gültiger, neuer, potentiell nützlicher und schlußendlich verständlicher Muster in (großen) Datenbeständen. □

Anmerkung: Häufig wird der Begriff *Data Mining* synonym zu Wissensentdeckung in Datenbanken verwendet, vor allem im kommerziellen Bereich. Wie bereits angedeutet wurde, ist Data Mining aber eigentlich nur ein Teilschritt innerhalb des KDD-Prozesses, nämlich der Analyseschritt zur Mustererkennung.

Die im Wissensentdeckungsprozeß erzeugten Muster stellen eine zusammenfassende Aussage über eine Untermenge der Daten dar. Ihre Repräsentationsform hängt von der Art des eingesetzten Analyseverfahrens ab; so können zum Beispiel Regeln, Gruppenbildungen, Entscheidungsbäume, Vorhersagemodelle etc. generiert werden. Muster sind also (höherwertige) Ausdrücke über die zugrundeliegenden Daten, die

in einem nichttrivialen Prozeß, also durch Suche oder Inferenz, gefunden wurden (vgl. [FPS96]). Sie müssen nach Definition 2.1 die folgenden Voraussetzungen erfüllen, um als zulässig zu gelten:

1. *Gültigkeit*: Ein Muster ist dann gültig, wenn es die aktuelle Datenbasis bzw. die zugrundeliegenden Gesetzmäßigkeiten konsistent beschreibt. Die Gültigkeit kann sich im Zeitablauf ändern und zum Beispiel durch eine Aktualisierung der Datenbasis verloren gehen. Davon betroffene Muster sollten aber auch für neue Daten noch ein hohes Maß an Gültigkeit besitzen.
2. *Neuartigkeit*: Die entdeckten Muster sollen für den Benutzer – zumindest aber für das KDD-System – neu sein, damit ein Erkenntnisgewinn möglich wird.
3. *Nützlichkeit*: Die Nützlichkeit steht in Bezug zu den Zielvorstellungen des Anwenders. Sie ist außerdem abhängig vom Problembereich und muß in der Regel individuell definiert werden, etwa durch die Angabe von Schwellenwerten für bestimmte Zielgrößen.
4. *Verständlichkeit*: Analog zur Nützlichkeit zielt die Verständlichkeit auf den menschlichen Anwender ab. Sie erfordert sowohl eine geeignete Präsentation der entdeckten Muster als auch die Einhaltung von *Ockham's Razor*, ein der Philosophie des Minimalismus entstammendes Prinzip, welches verlangt, daß ein Sachverhalt so einfach wie möglich beschrieben wird.

Abgrenzungen

Im folgenden werden einige Begriffe aus dem Bereich der automatischen Datenanalyse kurz erläutert und gegenüber der Wissensentdeckung in Datenbanken abgegrenzt.

KDD vs. Data Warehousing: Wird in einem Unternehmen die Extraktion von Daten aus verschiedenen, oft heterogenen operationalen Datenbanksystemen sowie das Zusammenführen und geeignete Speichern dieser Daten für Auswertungszwecke konsequent betrieben, dann nennt man diesen Vorgang *Data Warehousing* und den entstandenen Datenbestand *Data Warehouse* (vgl. [Inm96], [Kir97]). Das Ziel ist häufig, durch den Einsatz geeigneter Datenanalyseverfahren, wie z. B. KDD, eine unternehmerische Wissensbasis (»Unternehmensgedächtnis«) aufzubauen (vgl. [GG00]). Data Warehousing kann in solchen Fällen als Vorbereitungsphase zur Wissensentdeckung angesehen werden.

KDD vs. Online Analytical Processing: Unter dem Begriff *Online Analytical Processing* (abgekürzt *OLAP*) werden Datenanalyseverfahren zusammengefaßt, die in der Lage sind, im Dialog mit dem Benutzer aus großen betrieblichen Datenbeständen wie Data Warehouses verdichtete Informationen zu gewinnen. OLAP konzentriert

sich bei der Analyse hauptsächlich auf das Berichtswesen; die Daten werden multidimensional aufbereitet und interaktiv in Berichten mit Tabellen und Grafiken verdichtet. Im Unterschied zum KDD handelt es sich bei OLAP um eine hypothesengestützte Datenanalyse (d. h., der Anwender weiß bereits, wonach er sucht und überprüft seine Hypothesen), während KDD eine möglichst automatische, hypothesenfreie Wissensentdeckung anstrebt (vgl. [FPS96], [GG00]).

KDD vs. Pattern Recognition: Sehr viel länger als KDD existiert die Forschungsrichtung *Mustererkennung* (engl. *Pattern Recognition*). Sie befaßt sich mit der automatischen Wiedererkennung von Mustern und ihrer Zuordnung zu bekannten Kategorien oder Klassen. Häufig werden zur Mustererkennung von Sensoren stammende Meßwerte physikalischer Größen analysiert; Anwendungsbeispiele sind die Erkennung von Schriftzeichen, die Lauterkennung oder die Bildverarbeitung. Die klassische Mustererkennung löst ausschließlich Klassifizierungsaufgaben (siehe [Mer80]), während für KDD die Klassifizierung nur ein Problemfeld unter mehreren ist.

2.2.2 Ziele und Verfahren

Ein wichtiges Ziel des KDD-Prozesses ist es, die Wissensentdeckung möglichst selbstständig, d. h. mit möglichst wenig Benutzereingriffen durchzuführen. Trotzdem ist die Wissensentdeckung in Datenbanken üblicherweise kein vollständig automatischer Vorgang oder eine generelle Problemlösungsstrategie. Vielmehr soll der Anwender durch die Interaktion mit dem KDD-System in die Lage versetzt werden, die richtigen Analyseverfahren anzustoßen, statt wie bei herkömmlichen Datenbank- und Statistikwerkzeugen nur Anfragen zu formulieren bzw. Hypothesen testen zu lassen.

Der Benutzer muß also lediglich bestimmte Vorstellungen von der Art der anzuwendenden Modelle haben, und das KDD-System ermittelt dann mit geeigneten (semi-) automatischen Datenanalyseverfahren *alle* hierzu passenden Muster oder Wissenselemente. Grundsätzlich lassen sich zwei übergeordnete Ziele unterscheiden (vgl. [FPS96], [WI98]):

- Gewinnung von *Strukturwissen* und
- Gewinnung von *Vorhersagewissen*.

Das Strukturwissen ist deskriptiv und wird durch Muster repräsentiert, die in verständlicher Form signifikante Zusammenhänge in den Daten beschreiben. Es ist vom Problembereich abhängig, wie diese Muster zu interpretieren sind (z. B. als Korrelationen oder Kausalitäten). Dagegen beschreibt das Vorhersagewissen Zusammenhänge, die innerhalb der Daten in bezug auf eine vorbestimmte Zielgröße vorhanden

sind. Bei diesen Mustern steht nicht die verständliche Form im Vordergrund, sondern ihre Anwendung auf neue, zukünftige Datensätze, um den dann unbekannten Wert der Zielgröße zuverlässig bestimmen zu können. Die folgenden Unterabschnitte beschreiben die beiden Ziele der Wissensentdeckung in Datenbanken genauer und geben einen Überblick über die dazu eingesetzten Verfahren.

Strukturerkennung

Strukturwissen wird in der Regel mit unüberwachten Lernverfahren¹ gewonnen; diese Verfahren können auch dann sinnvoll eingesetzt werden, wenn die Zielgrößen nicht oder bestenfalls in groben Bereichen (z. B. durch Schwellenwerte oder relative Aussagen) definierbar sind. Aus diesem Grund liegt hier der Analyseschwerpunkt in der Entdeckung von in den Daten enthaltenen Beziehungen und Regelmäßigkeiten.

Abweichungsentdeckungsverfahren: In einer Datenbasis wird nach Mustern gesucht, die verschiedene Arten statistisch auffälliger Abweichungen von Untermengen im Verhältnis zur Gesamtmenge beschreiben. Das Wissen über solche Unregelmäßigkeiten oder Ausnahmen kann in einigen Problembereichen für sich allein bereits wertvoll sein, während es in anderen Bereichen zur Eliminierung von Ausreißerdaten genutzt werden kann.

Abhängigkeitsentdeckungsverfahren: In einer Datenbasis wird nach statistischen Abhängigkeiten innerhalb der Datensätze gesucht. Dies können z. B. Assoziationsregeln zur Beschreibung von häufig gemeinsam auftretenden Merkmalskombinationen oder Wahrscheinlichkeitsnetze zur Erfassung von probabilistischen Zusammenhängen zwischen den Merkmalswerten sein. Auf einer strukturellen Modellebene werden lokale Abhängigkeiten beschrieben, auf einer quantitativen Ebene darüber hinaus die Stärke dieser Abhängigkeiten.

Gruppenbildungsverfahren: Die Datenbasis wird so in homogene Gruppen aufgeteilt, daß eine verfahrensspezifische oder benutzerdefinierte Ähnlichkeitsfunktion zwischen den Datensätzen innerhalb einer Gruppe maximiert und zwischen verschiedenen Gruppen minimiert wird. Die dabei gefundenen Ähnlichkeitsgruppen (engl. *cluster*) sind nicht vorgegeben und werden vom Verfahren ermittelt.

Aggregations- / Charakterisierungsverfahren: Hierbei geht es um die Zusammenfassung genereller Eigenschaften der Objekte in der Datenbasis, wie z. B. durch charakterisierende Regeln in verschiedenen Abstraktionsebenen. Die aggregierten Repräsentationen sind oft aussagekräftiger oder leichter verständlich als die Ausgangs-

¹ Im Gegensatz zu überwachten Lernverfahren, die Trainingsbeispiele mit von außen vorgegebenen Lösungen auswerten, versuchen unüberwachte Lernverfahren selbständig, bestimmte statistische Eigenschaften in den Trainingsbeispielen zu entdecken.

daten, weil sie nur die zur jeweiligen Problemlösung als relevant betrachteten Charakteristika der Daten beschreiben.

Vorhersage

Im Gegensatz zum Strukturwissen wird Vorhersagewissen meistens mit überwachten Lernverfahren ermittelt, das heißt, die vorgegebene Datenbasis enthält genaue Werte für die gewünschten Zielgrößen. Die Verfahren extrahieren aus der Datenbasis die relevanten Informationen, um bei einer Anwendung auf neue Eingaben konkrete Vorhersagen für die Zielgrößen zu ermöglichen.

Klassifikations- / Regressionsverfahren: Anhand einer Trainingsmenge von Datensätzen, bei denen der gewünschte Wert einer Zielfunktion vorgegeben ist, erfolgt die Ableitung allgemeiner Funktionsbeschreibungen oder Entscheidungsvorschriften (z. B. Entscheidungsregeln oder -bäume), die dann bei neuen Datensätzen zur Vorhersage des Zielfunktionswerts genutzt werden. Ist der Zielfunktionswert ein symbolischer Wert, spricht man von einer *Klassifikation*; ist er ein Zahlenwert, spricht man allgemein von einer *Regression*.

Prognoseverfahren: Prognoseverfahren erweitern die Klassifikation bzw. Regression um eine zeitliche Komponente. Hierbei wird eine Datenbasis ausgewertet, die Veränderungen gleicher Objekte im Zeitverlauf abbildet. Geht man davon aus, daß die in solchen Zeitreihen entdeckten Regelmäßigkeiten auch zukünftig Gültigkeit besitzen, sind Prognosen für die Werte der definierten Zielgrößen möglich.

2.2.3 Einzelschritte im Wissensentdeckungsprozeß

Der Wissensentdeckungsprozeß als Ganzes ist iterativ und interaktiv. Die einzelnen Schritte werden in der Regel nicht in linearer Abfolge durchlaufen, sondern es können sich Schleifen bzw. Rücksprünge ergeben. So kommen zum Beispiel Situationen vor, bei denen zu Beginn der Datenerforschung die verfolgten Ziele noch nicht exakt festgelegt werden können oder bei denen sich während des Prozesses vorher nicht bedachte interessante Unterziele ergeben. Auch kann die Art der entdeckten Muster überraschen und das Augenmerk in eine neue Richtung lenken. Durch Interaktion mit dem Benutzer müssen insbesondere die Datenaufbereitungs- und analysearbeiten auf den jeweiligen Anwendungszweck abgestimmt werden.

Zwar existiert für den Einsatz der verschiedenen Techniken bisher kein generelles Vorgehensmodell, es haben sich aber in der Literatur und in der Anwendungspraxis die im folgenden erläuterten fünf Ablaufschritte weitgehend etabliert (vgl. [Man97], [Düs99]).

Schritt 1: Entwicklung der Zielvorstellungen

Als erstes werden die geeigneten Datenbereiche und Musterklassen sowie Kriterien zur Beurteilung der Interessanztheit von erkannten Mustern ermittelt. In dieser Planungsphase wird eine grobe Zielvorstellung entwickelt, mit der der KDD-Prozeß in die gewünschte Richtung gelenkt werden kann.

Schritt 2: Auswahl und Aufbereitung der Datenbestände (Preprocessing)

Diesem Schritt kommt in bezug auf Anwendbarkeit und Effizienz des anschließenden Data-Mining-Schritts eine besondere Bedeutung zu. Die verwendeten Algorithmen benutzen in der Regel zwar sehr große, letztendlich aber homogene Datenbestände (meistens in Tabellenform). Im einzelnen existieren die folgenden Aufgabenbereiche für das Preprocessing:

Datenselektion: Häufig sind die benötigten Datenquellen verteilt; insbesondere im kommerziellen Bereich geschieht die Datenerhebung und -speicherung an unterschiedlichen Stellen des betrieblichen Transformationsprozesses. In solchen Fällen muß in der Preprocessing-Phase eine geeignete Datenselektion und -zusammenführung stattfinden.

Datenreinigung: Gerade bei großen Datenbanken tritt oft das Problem auf, daß fehlende, mehrdeutige oder widersprüchliche Informationen vorhanden sind. Diese Datenbestände müssen dann im Hinblick auf ihre Qualität überarbeitet werden, indem z. B. statistische Ausreißer oder inkonsistente Datensätze entfernt oder fehlende Merkmalswerte ergänzt werden.

Datenreduktion: Neben den Operationen zur Nutzbarmachung unterschiedlicher Datenquellen übernimmt das Preprocessing Aufgaben der Datenreduktion, die auf eine verbesserte Effizienz der Data-Mining-Verfahren abzielt. Hierzu kann z. B. innerhalb der Datenbasis die Zusammenfassung mehrerer Merkmale, die Ermittlung und Eliminierung überflüssiger Merkmale, die Diskretisierung und Gruppierung von Merkmalswerten oder eine auf den zu verwendenden Algorithmus abgestimmte Transformation des Eingaberaums gehören.

Schritt 3: Musterentdeckung (Data Mining)

In diesem Schritt geschieht die eigentliche Wissensentdeckung durch Verfahren, die *selbständig* Hypothesen über eine Datenbasis generieren und dem Anwender die relevanten Ergebnisse in verständlicher Form präsentieren. Die wichtigste Anforderung an die genutzten Algorithmen ist ihr günstiges Laufzeitverhalten bei der Analyse *sehr*

großer Datenmengen. Daher stellt die Entwicklung von effizienten Datenstrukturen und Data-Mining-Algorithmen den Schwerpunkt innerhalb der wissenschaftlichen KDD-Forschung dar (vgl. [DRSS97]).

Stehen die für das jeweilige Analyseziel geeigneten Modelle und Data-Mining-Algorithmen sowie deren korrekte Parametrisierung fest, kann die Musterentdeckung autonom und effizient erfolgen. Im allgemeinen Fall muß ein versierter Anwender diese Festlegung vornehmen, oft in einer dynamischen und iterativen Vorgehensweise: Abhängig von den Ergebnissen erster Analysen werden andere Datenbereiche oder -repräsentationen gewählt oder es kommen andere Analyseverfahren bzw. andere Parametrisierungen zum Einsatz. Ist der Anwendungsbereich jedoch begrenzt und relativ unveränderlich, kann die Auswahl der optimalen Algorithmen und Parameterbereiche bereits im Vorfeld, also beim Aufbau des KDD-Systems, erfolgen. In diesem Fall ist ein vollständig automatisch ablaufender Data-Mining-Schritt möglich.

Anmerkung: Die vollständige Automatisierbarkeit des Data-Mining-Schritts stellt eine notwendige Voraussetzung für dessen Einsatz zur automatischen Wissensakquisition für Expertensysteme dar (vgl. Abschnitt 2.1.3). Die erfolgreiche Umsetzung dieser Verknüpfung wird am Beispiel der Gewinnung von heuristischem Diagnosewissen für hydraulische Anlagen in Kapitel 4 gezeigt.

Schritt 4: Nachbearbeitung der entdeckten Muster (Postprocessing)

Im vierten Schritt werden die gewonnenen Ergebnisse rücktransformiert und in vom Benutzer interpretierbare Aussagen umgewandelt. Die Nachbearbeitung umfaßt somit alle Datenverarbeitungstätigkeiten, die dazu dienen, anwendbarkeits- oder effizienzbedingte Vorverarbeitungsschritte rückgängig zu machen, sowie die Ergebnisse in ihrer Aussagekraft zu bewerten und dem Benutzer in verständlicher Form zu präsentieren. Sind im KDD-Prozeß mehrere geeignete Data-Mining-Verfahren angewandt worden, können die Einzelergebnisse in diesem Schritt gegenübergestellt werden.

Das Resultat ist also im Erfolgsfall das aus den Ursprungsdaten abgeleitete explizite Wissen. Die Darstellung des Wissens ist anwendungsabhängig und kann z. B. textuelle Beschreibungen, Tabellen, Regeln und graphische Darstellungsformen wie Graphen oder Diagramme umfassen. Im Idealfall ist das Wissen als vollständiger Report aufbereitet, in dem nicht nur die Ergebnisse, sondern z. B. auch zusätzliche Angaben über deren Qualität enthalten sind.

Schritt 5: Nutzbarmachung der erzielten Ergebnisse

Der letzte Schritt besteht aus der Auswertung, Operationalisierung und Nutzbarmachung des entdeckten Wissens, entweder direkt durch einen dafür zuständigen Ent-

scheidungsträger oder in einem weiteren Programm (z. B. Entscheidungsunterstützungssystem oder Expertensystem). Hierzu kann auch die Konsistenzprüfung in bezug auf das bereits vorhandene Wissen gehören.

Abbildung 2.2 faßt in Anlehnung an [FPS96] die grundlegenden Transformations-schritte zur Wissensentdeckung zusammen, die von einem KDD-System zwischen der Planungsphase und der Auswertungsphase durchgeführt bzw. unterstützt werden.

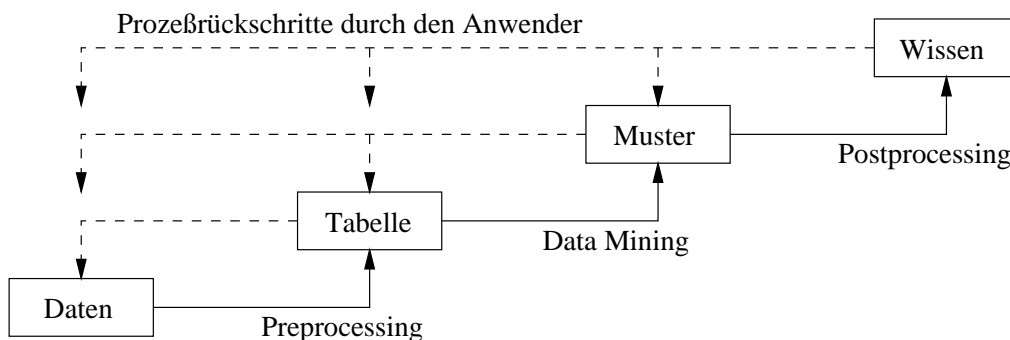


Abbildung 2.2: Wissensakquisition durch KDD-Prozeß

2.2.4 Data-Mining-Schritt

Dieser Abschnitt geht näher auf den Datenanalyseschritt (Schritt 3 des KDD-Prozesses) ein. Eine allgemeine Definition hierfür stammt von Holsheimer et al. [HS94], nach der bereits jeder Prozeß des maschinellen Lernens als Data Mining zu bezeichnen ist, wenn die zum Lernen ausgewerteten Daten in einer Datenbank verwaltet werden. Dieser Systematik folgend, existiert in der Literatur eine Begriffsfamilie für Verfahren, die nach der Art oder dem Inhalt der zu analysierenden Daten unterteilt werden:

- *Database Mining* (Das »klassische« Data Mining: Auswertung von Datenbanken),
- *Text Mining* (Auswertung und Klassifizierung von Textdokumenten),
- *Image Mining* (Auswertung von Bilddaten),
- *Spatial Data Mining* (Auswertung von raumbezogenen und geographischen Daten),

- *Web Mining* (Auswertung von Datenquellen aus dem Internet).

Die vorliegende Arbeit befaßt sich ausschließlich mit dem klassischen Data Mining. Im weiteren wird die datenbankorientierte Sichtweise genauer beschrieben und einige Vereinbarungen getroffen. Dazu definieren die folgenden Unterabschnitte das Eingabeformat von Data-Mining-Algorithmen sowie ihre Ausgabe.

Datenbasis

Die meisten Data-Mining-Verfahren greifen auf eine homogene, nicht verteilte Datenbasis zurück. Liegen die Ursprungsdaten nicht in der benötigten Repräsentation vor, müssen sie im Preprocessing geeignet transformiert werden (Schritt 2 des KDD-Prozesses). Aufgrund des großen Datenvolumens ist eine wichtige Anforderung an die Datenrepräsentation, daß sich ein effizienter Datenzugriff realisieren läßt. In dieser Arbeit soll jedoch von den technischen Aspekten des Datenzugriffs abstrahiert werden, und es werden zwei Annahmen getroffen:

1. *Die Datenbasis besteht aus einer einzigen Tabelle.*

Hierdurch wird ein schneller Zugriff auf die Gesamtheit der Informationen des Problembereichs ermöglicht. Diese Annahme stellt in der Regel keine Einschränkung dar, weil sich üblicherweise alle benötigten Einzeltabellen mit den entsprechenden Datenbankoperationen (zumindest virtuell) zu einer einzigen Tabelle zusammenführen lassen.

2. *Jede Tabellenzeile repräsentiert ausschließlich atomare Eigenschaften eines Objekts.*

Es dürfen also z. B. keine Referenzen zwischen Datensätzen oder komplexe Strukturen als Einzeleigenschaften modelliert werden. Durch diese Annahme ist gewährleistet, daß mit der einfachen Auswertung eines einzigen Datensatzes alle relevanten Informationen zu dem betreffenden Objekt bekannt sind.

Im weiteren wird also davon ausgegangen, daß eine Datenbasis, die mit einem Data-Mining-Verfahren analysiert werden soll, als einzige »flache« Tabelle vorliegt. In der Tabelle sind die Merkmale (Attribute) spaltenweise und die Datensätze zeilenweise definiert.

Definition 2.2 (Datenbasis für Data Mining): Es sei $A = \{A_1, \dots, A_m\}$ eine Menge von $m \in \mathbb{N}$ Attributen über nominal-, ordinal- oder kardinalskalierte Wertebereiche $dom(A_1), \dots, dom(A_m)$. Weiter seien Datensätze d_1, \dots, d_n für $n \in \mathbb{N}$ mit den Elementen $a_{i,j} \in dom(A_j)$ für $i \in \{1, \dots, n\}$ und $j \in \{1, \dots, m\}$ gegeben. Dann ist der

Aufbau einer Datenbasis D als Eingabe für Data-Mining-Algorithmen mit

	A_1	A_2	\cdots	A_m
d_1	$a_{1,1}$	$a_{1,2}$	\cdots	$a_{1,m}$
d_2	$a_{2,1}$	$a_{2,2}$	\cdots	$a_{2,m}$
\vdots	\vdots	\vdots	\ddots	\vdots
d_n	$a_{n,1}$	$a_{n,2}$	\cdots	$a_{n,m}$

definiert. □

Jede Zeile einer Datenbasis D aus Definition 2.2 stellt einen Datensatz fester Länge dar, der aus den Attributwerten eines Objekts, Ereignisses oder Falls besteht. Somit kann ein Datensatz $d_i \in D$ als ein Tupel aus dem Kreuzprodukt der Attribut-Wertebereiche dargestellt werden:

$$d_i = \langle a_{i,1}, a_{i,2}, \dots, a_{i,m} \rangle \in \text{dom}(A_1) \times \cdots \times \text{dom}(A_m).$$

Die Anzahl der möglichen Ausprägungen eines Attributs A_j sei in dieser Arbeit bezeichnet mit

$$|A_j| := \text{card}(\text{dom}(A_j)).$$

Data-Mining-Problem

Auf den Aspekt der Gewinnung von *vollständigem* Wissen zielt Definition 2.3 ab. Unter Berücksichtigung der im letzten Unterabschnitt erläuterten Datenrepräsentation spezifiziert sie das im Rahmen dieser Arbeit betrachtete generische Data-Mining-Problem.

Definition 2.3 ((Generisches) Data-Mining-Problem): Es sei D eine Datenbasis, die gemäß Definition 2.2 aufgebaut ist. Dann ist ein *Data-Mining-Problem* das Problem, aus D alle Muster zu erzeugen, die neben den in Definition 2.1 genannten Eigenschaften insbesondere die folgenden aufweisen:

1. Die Muster machen eine Aussage über die (beobachteten) Zusammenhänge zwischen den Attributwerten einer Teilmenge der Datensätze aus D und
2. die Muster sind im betrachteten Problembereich hinreichend interessant. □

Anmerkungen: (i) Die Beschreibung von Algorithmen zur Lösung einer konkreten Problemausprägung erfolgt in Kapitel 3. Hier geht es um die Erzeugung von Mustern, die durch sogenannte *Assoziationsregeln* repräsentiert werden.

(ii) Die Beurteilung, welche Muster »hinreichend interessant« sind, ist im allgemeinen subjektiv und daher anwender- bzw. anwendungsabhängig. Für den in Kapitel 4 betrachteten Problembereich der automatischen Akquisition von Diagnosewissen ist ein Muster dann interessant, wenn es einen positiven Informationsgewinn in bezug auf die Zielwissensbasis beiträgt, oder – anders ausgedrückt – wenn seine Nichterzeugung einen Informationsverlust bedeutet. Die Höhe des Informationsgewinns oder -verlusts läßt sich aus der jeweiligen Diagnosegüte ermitteln, die durch die Anwendung des Diagnosewissens auf reale Störungsfälle erzielt wird.

Innerhalb eines *benutzergesteuerten* KDD-Systems spielt die Anzahl der entdeckten Muster für die Lösung des Data-Mining-Problems eine wichtige Rolle: Werden dem Anwender zu viele Einzelmuster präsentiert, kann der hiermit erzielte Nutzen aufgrund der fehlenden Übersichtlichkeit gering sein. In diesem Fall ist eine Anpassung des Interessanzkriteriums oder eine zusätzliche Aggregation der Muster notwendig, wodurch jedoch ein beträchtlicher Informationsverlust entstehen kann. Werden dagegen die erzeugten Muster *automatisch* weiterverarbeitet, ist die Anzahl der Muster nicht entscheidend, und die gesamte Information kann ausgenutzt werden. Dieser Aspekt ist insbesondere für den Aufbau vollständiger Wissensbasen für Expertensysteme wichtig.

2.3 Neuer Ansatz zur Akquisition von Diagnosewissen

Im folgenden wird die Akquisition von Diagnosewissen als Anwendung des neuen Konzepts zur automatischen Wissensakquisition behandelt. Zu Beginn führt Abschnitt 2.3.1 in das Problemfeld der Diagnose technischer Anlagen ein, und Abschnitt 2.3.2 diskutiert die klassischen, aus der Literatur bekannten Lösungsansätze. Diese Ausführungen motivieren in Abschnitt 2.3.3 die Konkretisierung des automatischen Wissensakquisitionsansatzes.

2.3.1 Diagnose technischer Anlagen

Technische Systeme bestehen häufig aus einer großen Anzahl miteinander verbundener Einzelkomponenten, deren Zusammenwirken ein genau spezifiziertes Systemverhalten (das sog. *Sollverhalten*) gewährleisten soll. Als Beispiel seien in Autowerkstätten anzutreffende hydraulische Hubvorrichtungen genannt: Sie müssen in der Lage sein, Lasten bis zu einem festgelegten Höchstgewicht in einer bestimmten Zeit in die Höhe zu drücken.

Das Verhalten eines technischen Systems kann entweder direkt beobachtet werden (z. B. das Ausfahren eines Zylinders) oder indirekt mit Hilfe von Meßgeräten für die

2.3 NEUER ANSATZ ZUR AKQUISITION VON DIAGNOSEWISSEN

relevanten physikalischen Größen (z. B. Kraft oder Druck) bestimmt und protokolliert werden. Eine fehlerhaft arbeitende Komponente des Systems läßt sich in der Regel nicht unmittelbar lokalisieren. Der Fehler kann sich aber in einem vom Soll-Verhalten abweichenden Verhalten des Gesamtsystems bemerkbar machen. Diese Verhaltensanomalien werden *Symptome des Fehlers* genannt. Das Soll-Verhalten und die Symptomausprägungen sind – abgesehen vom Fehler selbst – auch von der vorliegenden *Betriebssituation* abhängig, die durch den aktuellen Betriebszustand (Phase) und die aktuellen Systemanforderungen (Belastungsstufen) gekennzeichnet ist.

Die Fehlererkennung, auch *Diagnose* genannt², ist ein abduktiver Prozeß des Rückschlusses von den beobachteten Wirkungen auf ihre nicht beobachtbaren Ursachen. Hierzu wird eine (mentale) »Diagnosefunktion«

$$diag : S \times B \times Z \rightarrow F$$

benötigt, die den Kombinationen aus Symptomen S , Belastungsstufen B und Betriebszuständen Z einen oder mehrere Fehler F zuordnet (siehe auch Abbildung 2.3).

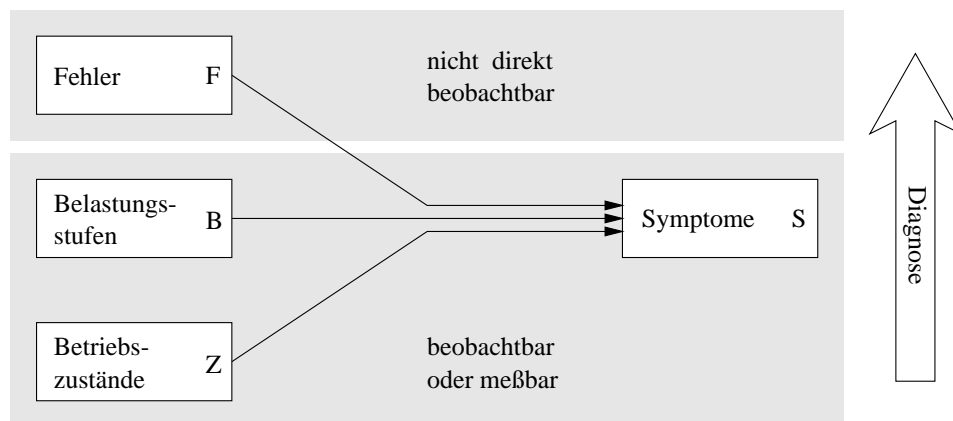


Abbildung 2.3: Ursache-Wirkungs-Zusammenhänge und Diagnoseproblem

Da sich technische Anlagen in ihrem Aufbau aus Komponenten und Verbindungen unterscheiden können, ist für jeden Anlagentyp die Anwendung einer spezialisierten Diagnosefunktion notwendig. Gute Ingenieure sind durch ihren Erfahrungsschatz und das grundsätzliche Verständnis der Arbeitsweise eines technischen Systems in der

2 Mit dem Begriff »Diagnose« wird sowohl der Prozeß der Fehlererkennung als auch der Fehler als Ergebnis dieses Prozesses bezeichnet. Die jeweilige Bedeutung geht dabei aus dem aktuellen Kontext hervor.

Lage, übliche Fehler sogar in für sie neuen Anlagen zu finden, wenn ihnen Unterlagen wie Schaltpläne, Meßwerte, Störungstabellen usw. zur Verfügung stehen. Oft liegen auch Meßprotokolle für Soll-Werte vor, oder die Beschreibung von Tätigkeiten, die zur Störungssuche auszuführen sind (z. B. Ablaufdiagramme, siehe [DD98]).

Mit einer zunehmenden Komplexität wird jedoch die Fehlersuche für menschliche Experten immer aufwendiger und schwieriger, selbst wenn es sich um Spezialisten für einen bestimmten Anlagentyp handelt. Parallel zur Größe des Systems steigen in der Regel auch die Anschaffungs- und Betriebskosten, so daß die Ursachen einer Störung um so schneller lokalisiert und behoben werden müssen. Dabei sollte ein unüberlegtes Auswechseln von vermeintlich defekten Komponenten nicht nur aus Zeit- und Kostengründen unbedingt vermieden werden, sondern auch wegen der Gefahr der Entstehung zusätzlicher Folgeschäden.

Aus diesen Gründen spielt die rechnergestützte Diagnose in technischen Anwendungsbereichen eine große Rolle und hat sich als eines der Haupteinsatzgebiete für Expertensysteme entwickelt (vgl. [PGPB96]). Ein Expertensystem zur Lösung des in Abbildung 2.3 dargestellten Diagnoseproblems wird auch *Diagnosesystem* genannt. Im technischen Umfeld sind an ein solches Expertensystem besondere Anforderungen zu stellen:

- *Effiziente Verarbeitung des Diagnosewissens*: Tritt im Betrieb einer technischen Anlage eine Störung auf, sollte das Diagnosesystem ihre Ursachen möglichst schnell ermitteln, damit sofort geeignete Gegenmaßnahmen ergriffen werden können.
- *Korrektheit und Vollständigkeit der Wissensbasis*: Der Einsatz einer geeigneten, umfangreichen Wissensbasis ist die Voraussetzung dafür, daß die Wahrscheinlichkeit von Fehldiagnosen sowie nicht erkannten Störungsursachen möglichst gering ist.
- *Möglichkeit der automatischen Wissensakquisition*: Das Diagnosesystem sollte bereits bei Inbetriebnahme einer neuen Anlage einsatzbereit zur Verfügung stehen.

Existierende Diagnoseansätze zur Repräsentation und automatischen Verarbeitung von Diagnosewissen erfüllen diese Forderungen in unterschiedlichem Maße. Im folgenden Abschnitt werden die wichtigsten bekannten Ansätze vorgestellt und im Hinblick auf den hier betrachteten Einsatzbereich bewertet.

2.3.2 Wissensbasierte Diagnoseansätze

Die ersten Diagnosesysteme wurden für den medizinischen Bereich entwickelt (vgl. [Pup91]). Es stellte sich jedoch heraus, daß die hier verwendeten Lösungsansätze auf technische Anwendungen übertragbar und wegen der verstandenen Ursache-Wirkungs-Prinzipien gut formal zu fassen waren. Im weiteren wird insbesondere ihre Eignung für die Diagnose *komplexer* technischer Anlagen diskutiert.

Üblicherweise erfolgt eine Einteilung der wissensbasierten Diagnoseverfahren in folgende »klassische« Kategorien (vgl. [Pup90], [PGPB96]): Einfach, probabilistisch, fallbasiert, heuristisch und modellbasiert.

Einfache Diagnose

Bei der *einfachen Diagnose* werden Entscheidungsbäume oder Entscheidungstabellen ausgewertet. Sie enthalten hierarchische Fragen zu den beobachteten Symptomen; ihre Beantwortung führt direkt zu einer Aussage über die Fehlerursache. Die einfache Diagnose ist daher nur für gut verstandene, stark begrenzte Anwendungsbereiche mit einfachen kausalen Zusammenhängen geeignet.

Probabilistische Diagnose

Die *probabilistischen Diagnoseansätze* (z. B. das Bayes-Theorem oder die Dempster-Shafer-Theorie) beruhen auf der Anwendung von statistischen Methoden zur Ableitung von Aussagen über die typische Verteilung möglicher Diagnosen. Auch sie lassen sich nur bei speziellen Problemstellungen anwenden (vgl. [Pup91]). Zum einen sind es restriktive Voraussetzungen, insbesondere die Forderung einer statistischen Unabhängigkeit der Symptome und des gegenseitigen Ausschlusses von Fehlern, zum anderen eine exponentielle Laufzeitkomplexität in der Anzahl der möglichen Fehler, die eine breite Anwendung der probabilistischen Diagnose verhindern.

Die ersten beiden Diagnoseansätze werden im folgenden nicht mehr berücksichtigt, denn es ist erkennbar, daß ihre Mächtigkeit nicht ausreicht, um eine Fehlerdiagnose für komplexe technische Anlagen erfolgreich durchführen zu können.

Fallbasierte Diagnose

Die Grundidee des *fallbasierten Diagnoseansatzes* ist es, zur Lösung eines neuen Problems auf Diagnosewissen zurückzugreifen, das in Form von gelösten Aufgaben des gleichen Problembereichs (z. B. für die selbe technische Anlage) vorliegt. Dieses

Wissen wird in einer sogenannten Falldatenbasis gesammelt; dabei besteht jeder Fall aus einer Problembeschreibung und der zugehörigen Lösung, der Diagnose. Liegt eine Diagnosesituation (Störungsfall) vor, wird aus der Fallbasis ein vergleichbarer, möglichst ähnlicher Fall herausgesucht und dessen Diagnose auf das aktuelle Problem übertragen ([Pup91]).

Die Problemlösungsfähigkeit von fallbasierten Diagnosesystemen hängt von zwei Faktoren ab:

1. Von der Anzahl und der Qualität der Fälle in der Fallbasis und
2. von der Güte des sogenannten *Ähnlichkeitsmaßes*.

Die Fallbasis muß eine hinreichend große Auswahl typischer Fehlersituationen enthalten, um das Fehlerspektrum geeignet abdecken zu können. Der Aufbau einer solchen Fallbasis kann sich im technischen Bereich, insbesondere für komplexe technische Anlagen, als schwierig erweisen. Mit dem Ähnlichkeitsmaß wird der Grad der Übereinstimmung (Ähnlichkeit) der alten Fälle mit dem aktuellen Fall berechnet. Auch dies ist für komplexe Anlagen in der Regel schwierig zu entwickeln. Außerdem muß dem Problembereich eine gewisse »Stetigkeit« zugrunde liegen, damit die Übertragbarkeit der Lösung des alten Falls auf das aktuelle Problem gewährleistet ist.

Heuristische Diagnose

Heuristische Diagnosesysteme basieren auf der Auswertung von assoziativem Wissen, das in Form von Heuristiken, also »Daumen-« oder Erfahrungsregeln, in einer Regelbasis repräsentiert wird. Die Regeln beschreiben direkt und ohne Begründung, welche Symptomkombinationen mit welcher Sicherheit auf welche Ursache hindeuten, sie haben also die Form

Wirkung → Ursache.

Da diese Regeln jedoch mit Unsicherheit behaftet sein können und jede Wirkung verschiedene Ursachen haben kann, kommen in einer konkreten Diagnosesituation üblicherweise mehrere Regeln zur Anwendung. Diese müssen dann geeignet verrechnet werden, um eine Aussage über die wahrscheinlichste Ursache machen zu können. Als Diagnose wird dann die Ursache mit der im Hinblick auf die beobachteten Symptome höchsten Bewertung betrachtet. Da das nach diesem Ansatz verarbeitete Wissen keine kausalen Zusammenhänge des Problembereichs berücksichtigt, wird es auch »flaches« Wissen genannt.

Die Diagnosegüte wird beim heuristischen Ansatz im wesentlichen von der Qualität der Regelbasis bestimmt. Häufig werden die Regeln durch einen sogenannten *Wissensingenieur* ermittelt, der einen Experten befragt und dessen Erfahrungen in das Diagnosesystem einpflegt (vgl. Abschnitt 2.1.3). Sie führen damit zwar zu einer klaren, verständlichen und leicht verarbeitbaren Wissensdarstellung, der Aufwand zum Aufbau, zur Pflege und zur Konsistenzerhaltung einer Regelbasis kann jedoch schnell mit ihrer Größe steigen. Ein weiteres Problem ist die Vollständigkeit der Regelmenge, denn es existiert kein etabliertes Verfahren zu ihrer systematischen und ingenieurmäßigen Konstruktion.

Als erfolgreiches heuristisches Diagnosesystem ist z. B. MYCIN ([Sho76]) für eine medizinische Anwendung bekannt geworden. Auch für den technischen Bereich existieren einige Systeme, hierbei werden jedoch nur kleine Problemgrößen behandelt oder auf zusätzliche Ansätze zurückgegriffen (z. B. MOLTKE, [PR93]). Für komplexe technische Anlagen reicht dieser Ansatz alleine nicht aus, da der »manuelle« Aufbau der Regelbasis zu aufwendig und fehleranfällig ist.

Modellbasierte Diagnose

In der *modellbasierten Diagnose* wird die zu diagnostizierende Domäne durch explizite Verhaltens- und Strukturmodelle repräsentiert (sogenanntes »tiefes« Wissen), mit denen das Systemverhalten simuliert werden kann. Zur Fehlerdiagnose werden Hypothesen generiert und deren Auswirkungen durch Simulation im Modell hergeleitet. Stimmen die erwarteten Symptome mit den bei der realen Anlage beobachteten Symptomen überein, kann die zugrundeliegende Hypothese als Diagnose angesehen werden, andernfalls muß eine neue Hypothese geprüft werden.

Als Voraussetzung für die Anwendung modellbasierter Diagnoseansätze muß der zugrundeliegende Problembereich gut und vollständig verstanden sein, andernfalls ist das Aufstellen der Verhaltens- und Strukturmodelle nicht möglich. Diese Voraussetzung kann für viele technische Bereiche als erfüllt angesehen werden: Aufgrund der oft komponentenorientierten Systemmodellierung sind kontextfreie, lokale Beschreibungen von Funktion und Arbeitsweise der Einzelteile möglich. Ein physikalisches Gesamtmodell für eine spezielle Anlage entsteht dann durch die Synthese der Komponentenmodelle sowie ihrer Verbindungen. Durch geeignete Werkzeuge (z. B. ^{art}*deco* für hydraulische und pneumatische Anlagen, siehe [KCH⁺95], [Cur96] und [SCH98]) kann die Modellbildung automatisch »im Hintergrund« durchgeführt werden, wenn der Ingenieur eine neue Anlage mit graphischen Operationen konstruiert.

Der wesentliche Nachteil des modellbasierten Diagnoseansatzes ist jedoch, daß nur

Beziehungen der Form

Ursache → Wirkung

zur Verfügung stehen. Für die Lösung eines konkreten Diagnoseproblems ist aber genau der Umkehrschluß notwendig. Deshalb kommt der Hypothesengenerierung und der Ermittlung ihrer Auswirkungen innerhalb des Modells eine entscheidende Bedeutung zu. Da die Zahl zu prüfender Hypothesen (Fehlerszenarien) in komplexen technischen Systemen sehr groß sein kann, ist der Aufwand zur Fehlerdiagnose üblicherweise zu hoch.

Der modellbasierte Ansatz hat aber auch einen entscheidenden Vorteil: Durch die Verwendung von »tiefem« Wissen liegt ein relativ mächtiges Verfahren vor, mit dem ein breites Spektrum an Fehlersituationen abgedeckt werden kann. Aus diesem Grund existieren einige modellbasierte Diagnosesysteme, wie z. B. die GDE (*General Diagnostic Engine*) von de Kleer et al. ([dKW87],[FdK93]) oder FEMO von Puppe ([Pup90]).

2.3.3 Integration von Data Mining in die Wissensakquisition

Die Diskussion der bekannten klassischen Diagnoseansätze hat gezeigt, daß kein Ansatz alle der in Abschnitt 2.3.2 motivierten Anforderungen an ein technisches Diagnosesystem erfüllen kann. Abbildung 2.4 faßt die Ergebnisse noch einmal vereinfacht zusammen.

Diagnoseansatz	effiziente Verarbeitung	Korrektheit bzw. Vollständigkeit	automatische Wissensakquisition
fallbasiert	+	-	-
heuristisch	+	-	-
modellbasiert	-	+	+

Abbildung 2.4: Bewertung der klassischen Diagnoseansätze im Hinblick auf die Diagnose technischer Anlagen (+ : möglich, - : nicht möglich bzw. sehr aufwendig)

Wünschenswert ist ein Diagnoseansatz, der die Stärken der klassischen Ansätze vereint und auf diese Weise alle genannten Anforderungen bestmöglich erfüllt. Ein Weg kann in der Kombination zweier Ansätze liegen, indem aus dem »tiefen« kausalen Wissen des modellbasierten Ansatzes automatisch »flaches« assoziatives Wissen in

Form von Regeln oder Fallbeispielen erzeugt wird. In der Literatur wird diese Vorgehensweise auch mit *Wissenskompilierung* oder *Modelltransformation* bezeichnet (vgl. [CM83], [PR93], [Ste01]³).

Leschka untersucht in [Les96] am Beispiel des fallbasierten Störungsmanagements in flexiblen Fertigungssystemen die modellbasiert-fallbasierte Kombination. Sein Ansatz ist auf die direkte Beteiligung eines Experten bei der Wissensakquisition ausgerichtet. Dagegen wird in der vorliegenden Arbeit die modellbasiert-heuristische Kombination am Beispiel der Fehlerdiagnose für hydraulische Anlagen verfolgt, wobei die Kernfrage lautet: Wie läßt sich eine *vollständig automatische* Modelltransformation durchführen, die auch für komplexe Anlagen eine hohe Problemlösungsfähigkeit ermöglicht? Dieser Frage wird im weiteren nachgegangen.

Neuer Wissensakquisitionsansatz

Wird für die zu diagnostizierende Anlage durch die Simulation von Fehlerzenarien eine Simulationsdatenbasis erzeugt, enthält sie Wissen über mögliche Komponentenfehler und ihre Auswirkungen auf das Anlagenverhalten in Form von Ursache→Wirkung-Zusammenhängen. Diese Simulationsdatenbasis ist jedoch für eine heuristische Fehlerdiagnose nicht direkt von Nutzen, weil diese auf Wirkung→Ursache-Regeln angewiesen ist (vgl. Abschnitt 2.3.2). Daher ist ein weiterer Transformationsschritt erforderlich, in dem mit einem Regellernverfahren das implizite Diagnosewissen der Simulationsdatenbasis in eine explizite heuristische Darstellung überführt wird.

Die durch Anwendung des heuristischen Diagnosewissens erzielbare Diagnosegüte, also der Anteil der korrekt diagnostizierten Störungssituationen, ist dabei insbesondere abhängig von der Korrektheit und Vollständigkeit der erzeugten Regelbasis. Eine Voraussetzung ist daher, daß die Simulationsdatenbasis bereits ein möglichst vollständiges Spektrum der Ursache / Wirkungs-Zusammenhänge der Anlage abdeckt. Bei komplexen technischen Anlagen ist also eine große Anzahl von Simulationsläufen notwendig, was zwangsläufig zu einer sehr großen Datenbasis führt.⁴

Für das Problem der Wissensakquisition aus sehr großen Datenbeständen ist bereits in Abschnitt 2.1.3 eine mögliche Lösung angeregt worden, die den Einsatz von Data-Mining-Verfahren vorsieht. Als konkrete Umsetzung dieses Lösungsvorschlags wird

3 Insbesondere Stein befaßt sich in [Ste01] ausführlich mit der Erzeugung und Transformation von Modellen sowie deren Einsatz für Diagnose- und Designaufgaben.

4 Der Simulationsaufwand für die zu diagnostizierende Anlage ist sowohl von der Anzahl der Einzelkomponenten als auch der möglichen Betriebsszenarien abhängig, wie in Abschnitt 4.3.2 näher ausgeführt wird.

2 WISSENSAKQUISITION MIT DATA MINING

in den weiteren Kapiteln dieser Arbeit die Integration von Data Mining in den automatischen Akquisitionsprozeß für Diagnosewissen behandelt. Das neue Konzept umfaßt zwei zeitlich aufeinanderfolgende Schritte (vgl. Abbildung 2.5)⁵:

1. Erzeugung einer Simulationsdatenbasis durch *Simulation* der kausalen Anlagenmodelle und
2. Erzeugung einer heuristischen Wissensbasis durch *Data Mining* in der Simulationsdatenbasis.

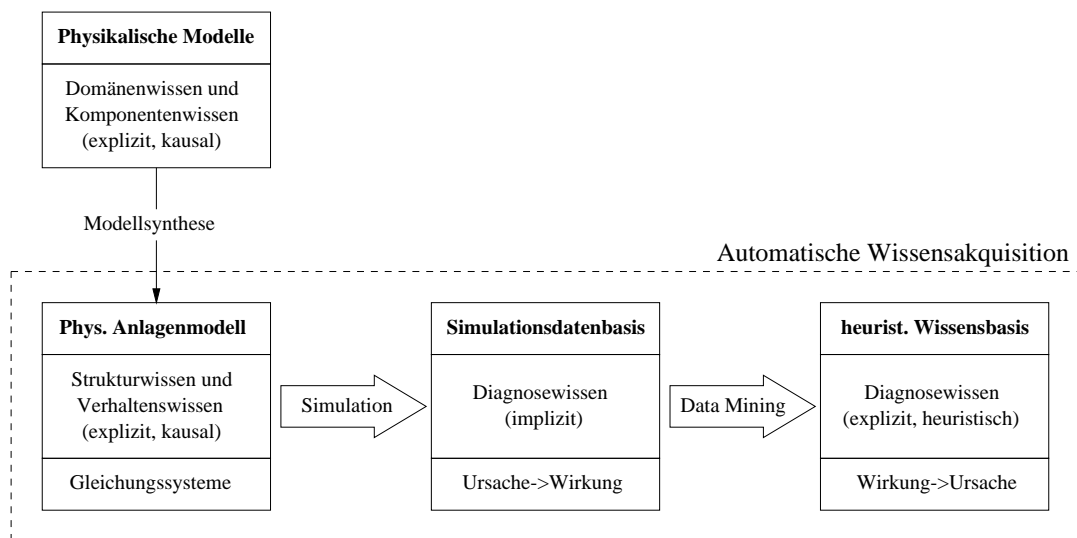


Abbildung 2.5: Wissenstransformationsschritte zur automatischen Akquisition von heuristischem Diagnosewissen

Da die eigentliche Verhaltenssimulation (Schritt 1) durch die Anwendung bestehender Simulationswerkzeuge vorgenommen werden kann, konzentriert sich die vorliegende Arbeit auf die Erzeugung der Diagnoseregeln (Schritt 2). Als Repräsentation für die heuristischen Diagnoseregeln werden in Kapitel 3 die sogenannten *Assoziationsregeln* ausführlich behandelt. Mit Assoziationsregeln lassen sich die für den heuristischen Diagnoseansatz benötigten unsicheren Zusammenhänge in der Form

$$\text{Symptom } 1 \wedge \dots \wedge \text{Symptom } m \rightarrow \text{Diagnose}$$

⁵ Im Unterschied zu Console et al. ([CPDT93]) oder Hesse ([Hes99]), die *hybride* Ansätze zur verzahnten Ausführung von Modellanalysen und Regelanwendungen zur Diagnosezeit vorschlagen, wird nach dem in dieser Arbeit vorgestellten Ansatz ein *rein heuristisches* Diagnosesystem generiert, in dem bereits vor der ersten Diagnosesituation das in den Ausgangsmodellen enthaltene Wissen in Regelform vorliegt.

2.3 NEUER ANSATZ ZUR AKQUISITION VON DIAGNOSEWISSEN

darstellen, und aus dem Data-Mining-Bereich existieren effiziente Lernalgorithmen zu ihrer Erzeugung.

In Kapitel 4 wird dieser Ansatz in ein vollständiges Konzept zur automatischen Erzeugung von Diagnosesystemen für hydraulische Anlagen eingebettet und detailliert beschrieben. Die mit einer prototypischen Realisierung erzielten Ergebnisse (vgl. Kapitel 5) zeigen die Vorteile des Ansatzes:

- Es können keine in bezug auf die Ausgangsdatenbasis inkonsistenten Regelmengen entstehen, weil die (korrekten) Verhaltens- und Strukturmodelle des Anwendungsbereichs als Wissensgrundlage dienen.
- Die Gefahr der Unvollständigkeit ist klein, weil durch die Simulation ein hoher Grad der Überdeckung von Fehlersituationen erreicht werden kann.
- Der eigentliche Diagnoseprozeß im Störfall ist sehr effizient durchführbar, weil in dieser Phase durch einfache Regelanwendung der direkte Schluß von Symptomen auf ihre Ursachen möglich ist.

2 WISSENSAKQUISITION MIT DATA MINING

3 Assoziationsregeln

In diesem Kapitel wird mit den sogenannten *Assoziationsregeln* (engl. *association rules*) ein Konstrukt zur Beschreibung von strukturellen Zusammenhängen in Daten diskutiert. Sie wurden 1993 von Agrawal, Imielinski und Swami in [AIS93] als Mittel zur Warenkorbanalyse¹ eingeführt. Die automatische Entdeckung von Assoziationsregeln ist heute ein wichtiges Ziel innerhalb des Data Minings, denn die Regeln lassen sich auch aus großen Datenmengen effizient lernen und sind intuitiv interpretierbar.

Allgemein beschreiben *Assoziationen* eine über die Zufallshäufigkeit hinausgehende gemeinsame Auftretswahrscheinlichkeit zweier oder mehrerer Muster bzw. Gedanken. Dazu sind in der Psychologie sogenannte Assoziationsexperimente bekannt, in denen Versuchspersonen auf ein einzelnes vorgegebenes Wort mit demjenigen anderen Wort antworten sollen, das ihnen zuerst einfällt. Hieraus ergibt sich eine Häufigkeitsverteilung für paarweise Wortnennungen, die als Beweis für die assoziative Arbeitsweise des menschlichen Gehirns gilt (vgl. [Rap96]).

Ähnlich ist das Prinzip der Assoziationsregeln: Sie bilden Zusammenhänge zwischen Elementen einer Datenbasis ab, die nicht unmittelbar sichtbar sind, sondern erst durch eine Datenanalyse gewonnen werden. Dabei stellen die Regeln keine exakten Implikationen dar, d. h. sie haben eher probabilistischen oder korrelativen Charakter. Sind P und K Aussagen über zugrundeliegende Datensätze, so hat eine Assoziationsregel die Gestalt

$$P \rightarrow K \text{ mit Wahrscheinlichkeit } c \text{ und relativer Häufigkeit } s$$

und sagt aus, daß bei denjenigen Datensätzen, auf die Aussage P zutrifft, die Aussage K mit einer Wahrscheinlichkeit von c assoziiert ist; die Regel läßt dabei auf einen Anteil s der Datensätze korrekt anwenden. Meistens sind die Aussagen innerhalb der Regeln auf einfache Vergleiche von Attributwerten sowie ihre konjunktive Verknüpfung beschränkt.

Im Gegensatz zu vielen anderen Regellernverfahren geht es im Data-Mining-Umfeld üblicherweise um das Problem, *alle* Assoziationen zu entdecken, die bestimmte Kriterien erfüllen. Gerade dieser Aspekt macht ihre Verwendung bei der Modelltransformation zur Akquisition von Diagnosewissen interessant. Hier repräsentieren die erzeugten Assoziationsregeln alle heuristischen Symptom→Diagnose-Zusammenhänge, die aus den Verhaltens- und Fehlermodellen einer technischen Anlage herleitbar sind.

¹ In der Warenkorbanalyse wird das Kaufverhalten von Kunden untersucht, zum Beispiel durch die Ermittlung von Artikelmenngen, die häufig gemeinsam gekauft wurden.

Die weiteren Ausführungen dieses Kapitels beschäftigen sich daher mit verschiedenen, aus der Literatur bekannten Repräsentationsformen für Assoziationsregeln sowie ihrer effizienten Generierung und diskutieren jeweils ihre Eignung als heuristische Diagnoseregeln.

Einteilung des Kapitels: Abschnitt 3.1 behandelt generelle Konzepte, kausale oder korrelative Zusammenhänge mit herkömmlichen Regeln sowie mit Assoziationsregeln darzustellen. Der Schwerpunkt liegt dabei auf der Beschreibung semantischer Aspekte, unabhängig von einer konkreten Regelsyntax oder möglichen Regelverarbeitungsstrategien. Die anschließenden Abschnitte gehen näher auf spezielle Assoziationsregelarten ein, die sich durch Variationen der Regelstruktur und / oder der Wertebereiche der zugrundeliegenden Datenbasis ergeben. Im einzelnen werden behandelt: boolesche Assoziationsregeln für binäre Wertebereiche (Abschnitt 3.2), kategoriale Assoziationsregeln für diskrete Wertebereiche (Abschnitt 3.3) und Klassifikationsregeln für einelementige Konklusionen zur Vorhersage von Klassenzugehörigkeiten (Abschnitt 3.4). Neben der Darstellung der formalen Definitionen und Eigenschaften von Assoziationsregeln wird insbesondere auf ihre automatische Erzeugung eingegangen.

3.1 Regelkonzepte

Jedes regelbasierte System ist zweigeteilt und setzt sich aus einer Wissensbasis und einer Inferenzkomponente zusammen. Die Wissensbasis wiederum besteht aus einer Faktenbasis, die Wissen über konkrete Sachverhalte des Problemkreises repräsentiert, und einer Regelbasis, die Wissen über allgemeine Beziehungen zwischen Sachverhalten enthält. Die wesentliche Aufgabe der Inferenzkomponente ist es, neues Wissen herzuleiten, indem mit Hilfe der Regeln aus den vorhandenen Fakten neue Fakten geschlußfolgert werden.

In diesem Abschnitt wird näher auf die Regelbasis eingegangen. Dazu erfolgt die Erläuterung der Ausdrucksmöglichkeiten von Regeln, insbesondere zur Modellierung von unsicheren Schlußfolgerungen (Abschnitt 3.1.1). Anschließend wird in Abschnitt 3.1.2 auf das Konzept einer speziellen Regelfamilie, die Assoziationsregeln, eingegangen und das sogenannte Assoziationsregelproblem definiert. Abschnitt 3.1.3 nimmt eine systematische Einordnung der in der Literatur bisher isoliert betrachteten Assoziationsregeltypen vor.

3.1.1 Regeln

Jede Regel besteht mindestens aus einem Bedingungsteil P , der *Prämisse*, und einem Folgerungsteil K , der *Konklusion*. Zusätzlich kann einer Regel eine Menge Q von Eigenschaften zugeordnet sein, die diese Regel charakterisieren. Die zulässigen Regelstrukturen und die Bestandteile der Eigenschaftsmenge sind vom jeweiligen Regelkonzept abhängig. Sollen im weiteren syntaktische Aspekte unberücksichtigt bleiben, wird eine Regel r durch das Tripel

$$r = \langle P, K, Q \rangle$$

oder durch die Darstellung

$$r = P \rightarrow K \mid Q$$

repräsentiert, wobei die Eigenschaftsmenge Q leer sein darf.

Der Bedingungsteil P einer Regel ist ein Ausdruck, der sich über einer gegebenen Datenbasis immer zu einem Wahrheitswert evaluieren läßt, während im Folgerungsteil K zusätzlich noch prozedurale Elemente oder Anweisungen erlaubt sein können (z. B. in Produktionsregelsystemen). Eine Regel stellt somit eine Folgerungsbeziehung dar und besitzt die folgende Semantik:

Wenn P wahr ist, dann gilt K mit den Eigenschaften Q .

Zur Anpassung an den jeweiligen Einsatzbereich bzw. an den Regelverarbeitungsmechanismus existieren zahlreiche Regelvarianten. Generell lassen sich die verschiedenen Regelstrukturen und -ausdrucksmöglichkeiten gegeneinander abgrenzen, indem ihre syntaktischen und semantischen Eigenschaften betrachtet werden. In Abschnitt 3.1.3 wird dazu eine Hierarchie der wichtigsten (Assoziations-) Regelarten beschrieben.

Ein wesentlicher Aspekt bei der Abbildung von heuristischen Symptom→Diagnose-Folgerungen ist die Repräsentation von *Unsicherheit*. Hierfür werden Regelkonzepte benötigt, die Unsicherheitswerte bezüglich der Aussagen über Schlußfolgerungen von Prämissen auf Konklusionen enthalten. Durch eine Regel wird die Konklusion dann u. U. nicht exakt, sondern nur mit einer gewissen Evidenz oder Wahrscheinlichkeit hergeleitet; die Konklusion ist also nur in einem Teil der Anwendungsfälle gültig.

Zur Modellierung des unsicheren Schließens lassen sich je nach Anwendungsbereich Ursachen berücksichtigen, die auf subjektiven oder auf objektiven Unsicherheiten beruhen (vgl. [Ric89], [Lus90]). Subjektive Unsicherheiten entstehen durch menschliche Unzulänglichkeiten (z. B. durch begriffliche Ungenauigkeiten oder subjektive Wahrscheinlichkeiten) und werden häufig auch mit *Vagheit* bezeichnet. Objektive

Unsicherheiten hingegen liegen vor, wenn durch scharfe Interpretationen die Exaktheit von Aussagen zwar scheinbar vorhanden ist, jedoch inhärente Eigenschaften des Problembereichs oder Informationsmangel sichere Schlußfolgerungen verhindern (z. B. durch Meßfehler, numerische Ungenauigkeiten, falsche Aggregation, Zufallseinflüsse oder Unkenntnis von Parametern und allgemeinen Zusammenhängen).

Sind Unsicherheitsphänomene vorhanden, so müssen diese geeignet modelliert und den Regeln zugeordnet werden. Dies kann durch die bereits oben erwähnte Eigenschaftsmenge Q geschehen. Abbildung 3.1 listet einige Regelkonzepte für unsicheres bzw. vages Schließen sowie ihre charakterisierenden Eigenschaften auf. Auch das Problemfeld der heuristischen Diagnose ist üblicherweise mit objektiver Unsicherheit behaftet. In Kapitel 4 wird gezeigt, daß hierfür die Assoziationsregeln eine geeignete Repräsentationsform darstellen können. Ihre Semantik und Erzeugung wird daher in den folgenden Abschnitten ausführlich erläutert.

Regelart	Charakterisiert durch die Eigenschaft(en)
1. Probabilistische Regel	bedingte Wahrscheinlichkeit
2. Produktionsregel	Sicherheitsfaktor (engl. <i>certainty factor</i>)
3. Fuzzy-Regel	Zugehörigkeitsfunktion(en)
4. Assoziationsregel	relative Häufigkeit, Konfidenz (s. Abschnitt 3.1.2)

Abbildung 3.1: Regeltypen für unsicheres Schließen

3.1.2 Assoziationsregelkonzept

Zunächst wird das generelle, nicht anwendungsabhängig eingeschränkte Konzept der Assoziationsregeln definiert. Jede Regel repräsentiert ein Muster (Beziehung zwischen Attributen) innerhalb einer Datenbasis, deren Aufbau Definition 2.2 auf Seite 20 entspricht. Die Muster können dabei in Abhängigkeit des betrachteten Problemkreises Funktionalitäten, Korrelationen oder zufällige Phänomene ausdrücken. Die folgende Definition 3.1 ist angelehnt an [AIS93].

Definition 3.1 (Assoziationsregel): Es sei D eine Datenbasis mit der Attributmenge A . Für eine Teilmenge $X \subseteq A$ sei C_X ein boolescher Ausdruck, wobei C_X (evtl. logisch verknüpfte) Aussagen über genau die Attribute in X beinhaltet. Weiter seien $P \subset A$ und $K \subset A$ Attributmengen, für die gilt: $P \neq \emptyset$, $K \neq \emptyset$ und $P \cap K = \emptyset$.

Dann heißt eine Regel der Form

$$C_P \rightarrow C_K \mid c, s$$

Assoziationsregel über D mit der Konfidenz $c \in [0, 1]$ und der relativen Häufigkeit $s \in [0, 1]$. \square

Anmerkungen: Sei $r = C_P \rightarrow C_K$ eine Assoziationsregel. Durch die Einschränkungen $P \neq \emptyset$, $K \neq \emptyset$ und $P \cap K = \emptyset$ werden die Ausdrucksmöglichkeiten nicht verringert, denn es gilt:

1. bei Verletzung von $P \neq \emptyset$ oder $K \neq \emptyset$:
 $\emptyset \rightarrow C_K$ und $C_P \rightarrow \emptyset$ haben innerhalb des Assoziationsregelkonzepts keine Aussagekraft.

2. bei Verletzung von $P \cap K = \emptyset$:

In diesem Fall können folgende Ersetzungen so lange ausgeführt werden, bis $P \cap K = \emptyset$ gilt: Sei $g \in P \cap K$ ein Attribut, das sowohl in der Prämisse als auch in der Konklusion auftritt. Weiter seien $P' := P \setminus \{g\}$ und $K' := K \setminus \{g\}$. Dann ist die betrachtete Regel r entweder tautologisch, oder sie kann so umgeformt werden, daß g nicht mehr in der Schnittmenge vorkommt², denn r besitzt eine der vier Formen:

- (a) $r = C_{P'} \wedge C_g \rightarrow C_g \wedge C_{K'}$: Dann ist r äquivalent zu der Regel $C_{P'} \wedge C_g \rightarrow C_{K'}$.
- (b) $r = C_{P'} \vee C_g \rightarrow C_g \wedge C_{K'}$: Dann ist r äquivalent zur Konjunktion der drei Regeln $C_{P'} \rightarrow C_g$, $C_{P'} \rightarrow C_{K'}$ und $C_g \rightarrow C_{K'}$.
- (c) $r = C_{P'} \wedge C_g \rightarrow C_g \vee C_{K'}$: Dann ist r tautologisch.
- (d) $r = C_{P'} \vee C_g \rightarrow C_g \vee C_{K'}$: Dann ist r äquivalent zu der Regel $C_{P'} \rightarrow C_g \vee C_{K'}$.

Eine Assoziationsregel $C_P \rightarrow C_K$ kann gewöhnlich nicht als strenge Folgerung aufgefaßt werden, sondern beschreibt unsichere Zusammenhänge, d. h., in der Datenbasis existieren sowohl Datensätze, für die die Regelbeziehung zutrifft (positive Beispiele), als auch Datensätze, für die sie nicht zutrifft (negative Beispiele). Zu jeder Regel wird daher die in der Datenbasis gültige Unsicherheitssituation mit zwei statistischen Größen aggregiert: relative Häufigkeit s und Konfidenz c einer Regel, die in Anlehnung an [KMRV94] gemäß der folgenden Definition 3.2 berechnet werden.

Definition 3.2 (Relative Häufigkeit und Konfidenz einer Assoziationsregel): Es sei $r = C_P \rightarrow C_K \mid c, s$ eine Assoziationsregel über einer Datenbasis D .

² Ein Beweis der Äquivalenzen kann leicht über Wahrheitstabeln geführt werden.

3 ASSOZIATIONSREGELN

Dann ist die *relative Häufigkeit* (engl. *support*) s der Regel r definiert als

$$s = \frac{|\{d \in D \mid (C_P \wedge C_K) \text{ ist wahr für } d\}|}{|\{d \mid d \in D\}|}.$$

Die *Konfidenz* (engl. *confidence*) c der Regel r wird mit

$$c = \frac{|\{d \in D \mid (C_P \wedge C_K) \text{ ist wahr für } d\}|}{|\{d \in D \mid C_P \text{ ist wahr für } d\}|}$$

berechnet.

Für die relative Häufigkeit s und die Konfidenz c der Regel r ist auch die funktionale Notation $s(r)$ bzw. $c(r)$ gebräuchlich. \square

Beispiel 3.1: Es sei eine Datenbasis D mit den Attributen $A = \{A_1, \dots, A_5\}$ über der Menge der natürlichen Zahlen wie folgt gegeben:

	A_1	A_2	A_3	A_4	A_5
d_1	1	4	6	6	1
d_2	1	2	6	1	1
d_3	3	3	5	2	5
d_4	1	3	5	6	1
d_5	4	1	1	2	4

Dann gilt für die relative Häufigkeit s und die Konfidenz c ausgewählter Assoziationsregeln r_1, \dots, r_4 :

$$\begin{aligned} r_1 &= (A_1 = 1) \rightarrow (A_5 = 1) & | & s = \frac{3}{5}, c = 1 \\ r_2 &= (A_4 = 6) \vee (A_5 \geq 3) \rightarrow (A_3 \in [4, \dots, 8]) & | & s = \frac{4}{5}, c = \frac{3}{4} \\ r_3 &= \neg(A_2 < 5) \rightarrow (A_1 > 0) & | & s = 0, c = \text{undef.} \\ r_4 &= (A_1 < 5) \rightarrow (A_3 > 5) \wedge (A_4 = 1) & | & s = \frac{1}{5}, c = \frac{1}{5} \end{aligned}$$

\square

Die relative Häufigkeit s einer Assoziationsregel r gibt an, für welchen Anteil der Datensätze einer Datenbasis D die Regel r gültig ist; sie wird in der Literatur manchmal auch mit *rule frequency* bezeichnet. s ist normiert auf das Intervall $[0, 1]$ und kann als quantitatives Maß zur Bewertung einer Regel angesehen werden. Welche relative Häufigkeit für eine Regel günstig ist, läßt sich nur anwendungsabhängig beurteilen. Tendenziell gilt aber, daß Regeln mit einem Wert nahe bei 0 Ausnahmefälle und Regeln mit einem Wert nahe bei 1 triviale oder offensichtliche Zusammenhänge beschreiben.

Dagegen ist die Konfidenz c ein Maß für die Güte einer Regel $r = C_P \rightarrow C_K$, denn sie gibt für r den Anteil der positiven Beispiele in der Datenbasis an. c wird manchmal auch als Regelstärke (engl. *rule strength*) bezeichnet, weil – bezogen auf die Gesamtheit der Datensätze aus D – die Konfidenz einer Regel ein Maß für die Sicherheit ist, mit der von einem Ausdruck C_P auf einen Ausdruck C_K geschlossen werden darf. Auch dieses Maß ist auf das Intervall $[0, 1]$ normiert. Bei einem Wert $c < 0.5$ kann sicherer auf die Negation der Konklusion geschlossen werden, denn aus Definition 3.2 folgt unmittelbar:

$$c(C_P \rightarrow C_K) = 1 - c(C_P \rightarrow \neg C_K).$$

Kann der Benutzer eines Regelsystems Hypothesen über die in der Datenbasis vorhandenen Beziehungen anstellen, so ist die Erzeugung und Bewertung der entsprechenden Regeln einfach. Im Sinne des Data Minings ist jedoch vielmehr das Problem der *automatischen* Erzeugung *aller* hinreichend häufigen und hinreichend konfiden-ten Assoziationsregeln interessant (vgl. Definition 2.3 auf Seite 21). Dieses läßt sich allerdings nicht effizient und problemangemessen lösen, solange die booleschen Ausdrücke in den Regelprämissen und -konklusionen als *beliebige* Aussagen über die Attribute sein dürfen.

Erst wenn der Kontext des Regeleinsatzes bekannt ist, wird es möglich, Systeme zu erstellen, die eine Unterscheidung zwischen sinnvollen und nicht sinnvollen Regeln vornehmen können und somit zur Regelerzeugung nicht mehr auf eine Hypothesenformulierung durch den Benutzer angewiesen sind. Aus diesem Grund existieren Assoziationsregeltypen, die das Konzept anwendungsbezogen vereinfachen, indem sie die Wertebereiche der Attribute einschränken und / oder die Regelstruktur anpassen. Die im Hinblick auf ihre Eignung als Diagnoseregeln wichtigen Konzepte werden im folgenden Abschnitt behandelt.

3.1.3 Assoziationsregelarten

In diesem Abschnitt wird eine integrierende Darstellung und begriffliche Abgrenzung der wichtigsten aus der Literatur bekannten Assoziationsregelarten vorgenommen. Die Darstellung liefert zudem einen Überblick über die Unterschiede in der Regelmodellierung und verweist auf diagnoserelevante Aspekte. Im allgemeinen hat die Festlegung auf eine Regelart einen entscheidenden Einfluß auf die Konzeption einer Problemlösung: Mit der Art der Regelmodellierung werden die wesentlichen Ausdrucks- und Verarbeitungsmöglichkeiten des Problemlösungswissens bestimmt. Abbildung 3.2 zeigt eine Hierarchie der im weiteren behandelten Assoziationsregeltypen.

Obwohl mit der ersten Erwähnung des Begriffs »Assoziationsregel« in [AIS93] die restriktivste Regelart, die booleschen Assoziationsregeln, eingeführt wurde und erst

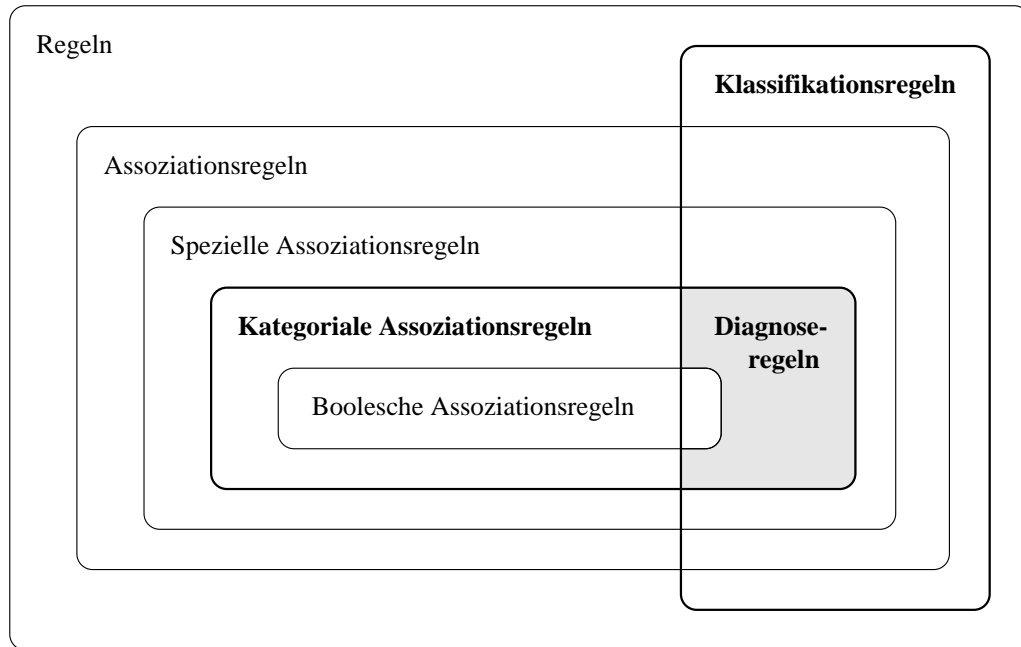


Abbildung 3.2: Begriffliche Hierarchie der Assoziationsregeltypen (Unterscheidung nach syntaktischen Restriktionen)

anschließend durch andere Autoren Erweiterungen des Konzepts stattfanden, wird im folgenden die Hierarchie von außen nach innen dargestellt.

Regeln: Hiermit ist die Gesamtheit aller Regelkonstrukte gemeint, die sich als Tripel $\langle P, K, Q \rangle$ für Prämissen P , Konklusionen K und Eigenschaftsmengen Q darstellen lassen (siehe Abschnitt 3.1.1).

Assoziationsregeln: Die allgemeinen Assoziationsregeln sind eine echte Teilmenge der allgemeinen Regeln. Sie lassen sich als Tripel $\langle P, K, \{s, c\} \rangle$ beschreiben und zeichnen sich neben den Eigenschaftszuordnungen Konfidenz c und relative Häufigkeit s durch die Beschränkung der Ausdrücke in P und K auf Aussagen über die Attributwerte der Datensätze in einer relationalen Datenbasis aus (siehe Abschnitt 3.1.2).

Spezielle Assoziationsregeln: Werden die Ausdrücke in P und K weiter eingeschränkt, die Wertebereiche der Datenbasis begrenzt oder die Eigenschaftsmenge Q erweitert, so erhält man auf bestimmte Einsatzzwecke spezialisierte Assoziationsregeln. Hierunter fallen z. B. Regeltypen wie Implikationsregeln ([BMUT97]), Fuzzy-Assoziationsregeln ([KFW98]), temporale Assoziationsregeln ([RMS98]), quantitative Assoziationsregeln ([SA96], [FMMT96], [MY97]), kategoriale Assoziationsregeln ([BW98]) oder repräsentative Assoziationsregeln ([Kry98a], [Kry98b]).

In dieser Arbeit werden die zur Lösung von Diagnoseproblemen geeigneten kategorialen Assoziationsregeln näher untersucht.

Kategoriale Assoziationsregeln: Bei dieser Regelart dürfen die Ausdrücke in P und K nur aus Gleichheitstests für Attribut/Wert-Paare bestehen, die evtl. durch Konjunktion miteinander verbunden sind. Die Wertebereiche der Attribute müssen endlich sein, und jeder Wert bezeichnet eine diskrete Kategorie. Üblicherweise werden die Kategorien durch ganze Zahlen oder durch Buchstaben dargestellt. Ein Beispiel für eine kategoriale Assoziationsregel ist

$$A_2 = 4 \wedge A_5 = 0 \rightarrow A_9 = F \mid c, s.$$

Weitere Ausführungen hierzu befinden sich in Abschnitt 3.3.

Boolesche Assoziationsregeln: Für boolesche Assoziationsregeln ist die Datenbasis per Definition auf zweielementige Wertebereiche (z. B. $\{0; 1\}$) beschränkt. Die Prämisse P und die Konklusion K bestehen dann lediglich aus Attributmengen, wobei jedes Element für einen Vergleich des Attributs auf den Wert 1 steht. Die Vergleiche innerhalb einer Attributmenge gelten als durch Konjunktion verknüpft. Eine typische boolesche Assoziationsregel könnte also unter Verzicht auf Mengenklammern lauten:

$$A_2, A_4, A_7 \rightarrow A_3, A_8 \mid c, s.$$

Kategoriale Assoziationsregeln und boolesche Regeln lassen sich ohne Informationsverlust ineinander transformieren (siehe Abschnitt 3.3.2), die beiden Regelarten besitzen also vergleichbare Eigenschaften. Die in der vorliegenden Arbeit benötigten Eigenschaften werden anhand der booleschen Assoziationsregeln in Abschnitt 3.2 beschrieben.

Klassifikationsregeln: Bei den Klassifikationsregeln ist die Konklusion K auf ein festes Vorhersageziel, die Klassenzuordnung, festgelegt. Diese Einschränkung kann mit allen anderen bisher erwähnten Einschränkungen kombiniert werden, so daß keine echte Teilmengenbeziehung zu den Assoziationsregelarten besteht (vgl. Abbildung 3.2). In Abschnitt 3.4 erfolgt die nähere Beschreibung dieser Regelart.

Diagnoseregeln: Durch eine Kombination der Definitionen von Klassifikationsregeln und kategorialen Assoziationsregeln erhält man eine zur Lösung von Diagnoseproblemen geeignete Regelart, weil sich hiermit heuristische Symptom \rightarrow Diagnose-Zusammenhänge repräsentieren lassen. Werden z. B. an zwei Meßstellen m_1 und m_2 die Symptome s_i und s_j beobachtet, die auf einen Fehler f schließen lassen, so lautet eine typische Diagnoseregeln (vgl. Definition 4.1 auf Seite 74):

$$m_1 = s_i \wedge m_2 = s_j \rightarrow f \mid c, s.$$

Da Diagnoseregeln auch allgemein zur Lösung von Klassifikationsproblemen³ geeignet sind, werden sie in diesem Kontext auch *klassifizierende Assoziationsregeln* genannt (Näheres in Abschnitt 3.4.2).

Die weiteren Abschnitte des Kapitels gehen auf einige Assoziationsregelarten detaillierter ein. Zunächst werden die booleschen Assoziationsregeln dargestellt, weil an ihnen grundlegende Eigenschaften und Algorithmen erläutert werden können. Darauf aufbauend wird die effiziente Erzeugung von kategorialen und klassifizierenden Assoziationsregeln beschrieben und zu einem Lernalgorithmus für heuristische Diagnoseregeln kombiniert.

3.2 Boolesche Assoziationsregeln

Boolesche Assoziationsregeln beschreiben strukturelle Zusammenhänge in den Datensätzen einer Datenbasis, deren Einträge einen zweielementigen Wertebereich besitzen. Weil es sich hierbei um die zuerst eingeführte Assoziationsregelart handelt ([AIS93]), werden sie häufig auch als »klassische« Assoziationsregeln bezeichnet.

In Abschnitt 3.2.1 erfolgt eine formale Beschreibung des Konzepts und der Teilmenge-eigenschaft, die sich durch die Beschränkung auf boolesche Wertebereiche ergibt. Das Problem, alle Assoziationsregeln zu erzeugen, die bestimmte Kriterien erfüllen, wird in Abschnitt 3.2.2 als Assoziationsregelproblem definiert. Abschließend beschreibt Abschnitt 3.2.3 ein aus der Literatur bekanntes Verfahren zur effizienten Lösung des Assoziationsregelproblems.

3.2.1 Definition und Eigenschaften

Zunächst werden einige Bezeichnungen vereinbart. Die zugrundeliegende Datenbasis sei D_n ; sie bestehe aus $n \in \mathbb{N}$ Datensätzen über $m \in \mathbb{N}$ Attributen $A = \{A_1, \dots, A_m\}$, die alle den Wertebereich $\mathbb{B} = \{0; 1\}$ besitzen. Dann kann D_n als Teilmenge des kartesischen Produkts aller Attributwertebereiche

$$D_n \subseteq \text{dom}(A_1) \times \dots \times \text{dom}(A_m) = \mathbb{B}^m$$

aufgefaßt werden. Wegen der Zweiwertigkeit der Datenbasis ist es sinnvoll, auch die Datensätze als Mengen anzusehen. Jeder Datensatz enthält dann alle Attribute,

³ Die Diagnose ist eine Klassifikationsaufgabe (Herleitung der Lösung aus einer Menge vorgegebener Alternativen anhand von Merkmalen, vgl. [Pup90]).

für die der Wert der entsprechenden Komponente gleich 1 ist, d. h. für die Attribute A_1, \dots, A_m und Datensätze $d_i \in D_n$ gilt (mit $i \in \{1, \dots, n\}$ und $j \in \{1, \dots, m\}$)

$$d_i = \langle a_{i,1}, \dots, a_{i,m} \rangle \Leftrightarrow d_i = \{A_j \in A \mid a_{i,j} = 1\}.$$

Der Begriff der relativen Häufigkeit einer Attributmenge $X \subseteq A$ in der Datenbasis D_n macht dann eine Aussage über den Anteil derjenigen Datensätze, in denen die Attribute aus X den Wert 1 besitzen:

Definition 3.3 (Relative Häufigkeit einer Attributmenge): Es sei D_n eine Datenbasis mit $n \in \mathbb{N}$ Datensätzen und einer Attributmenge A über einem binären Wertebereich \mathbb{B} . Weiter sei $X \subseteq A$. Dann heißt

$$h_n(X) = \frac{|\{d \in D_n \mid X \subseteq d\}|}{n}$$

relative Häufigkeit von X . □

Mit diesen Vereinbarungen ergibt sich die folgende Definition 3.4 für boolesche Assoziationsregeln.

Definition 3.4 (Boolesche Assoziationsregel): Es sei D_n eine Datenbasis mit der Attributmenge $A = \{A_1, \dots, A_m\}$, $m \in \mathbb{N}$, über einem binären Wertebereich \mathbb{B} . Weiter seien $P = \{A_{p_1}, \dots, A_{p_i}\}$ und $K = \{A_{k_1}, \dots, A_{k_j}\}$ Teilmengen von A (wobei $p_1, \dots, p_i, k_1, \dots, k_j \in \{1, \dots, m\}$), für die gilt: $P \cap K = \emptyset$. Dann wird eine Regel der Form

$$P \rightarrow K \mid c, s \quad \text{oder} \quad A_{p_1}, \dots, A_{p_i} \rightarrow A_{k_1}, \dots, A_{k_j} \mid c, s$$

boolesche Assoziationsregel über D_n genannt. Die Konfidenz c und die relative Häufigkeit s der Regel berechnen sich für $n \in \mathbb{N}$ Datensätze durch

$$c = \frac{h_n(P \cup K)}{h_n(P)} \quad \text{und} \quad s = h_n(P \cup K). \quad \square$$

Des weiteren stellt der folgende Satz 3.1 eine monotone Beziehung zwischen den relativen Häufigkeiten einer Attributmenge und ihrer Teilmengen her. Der bereits von Agrawal et al. in [AIS93] und [AMS⁺96] beschriebene Zusammenhang wird in dieser Arbeit als *Monotonieeigenschaft* bezeichnet.

Satz 3.1 (Monotonieeigenschaft): Es sei D_n eine Datenbasis mit $n \in \mathbb{N}$ Datensätzen und der Attributmenge A . Dann gilt für alle Attributmengen $X \subseteq A$ und $X' \subset X$:

$$h_n(X') \geq h_n(X). \quad \square$$

Beweis: Es sei d_i ein Datensatz der Datenbasis D_n . Aus $X' \subset X$ folgt $X \subseteq d_i \Rightarrow X' \subseteq d_i$, jedoch nicht die Gegenrichtung, d. h. $X \subseteq d_i \not\Rightarrow X' \subseteq d_i$. Auf alle Datensätze $d \in D$ bezogen gilt also $|\{d \in D_n \mid X' \subseteq d\}| \geq |\{d \in D_n \mid X \subseteq d\}|$, und unter Berücksichtigung von Definition 3.3 folgt die Behauptung. \square

Die Monotonieeigenschaft sagt also aus, daß jede Teilmenge einer Attributmenge eine relative Häufigkeit besitzt, die mindestens so groß ist wie die der Menge selbst. Daraus ergibt sich als Folgerung, daß keine Obermenge einer Attributmenge eine größere relative Häufigkeit besitzen kann als die Menge selbst. Die zweite Interpretation wird insbesondere zur Effizienzsteigerung der Assoziationsregelerzeugung genutzt (vgl. Abschnitt 3.2.3).

3.2.2 Assoziationsregelproblem

Dieser Abschnitt motiviert und formuliert das sogenannte *Assoziationsregelproblem*. Dazu wird zunächst diskutiert, welche Regelmengen überhaupt sinnvoll sind und welche Mächtigkeit sie besitzen können, bevor die Definition des Problems erfolgt.

Größe von Regelmengen

Liegt der Regelerzeugung eine Datenbasis D_n mit der aus $m \in \mathbb{N}$ Attributen bestehenden Attributmenge A zugrunde, werden laut Definition 3.4 sowohl die Prämisse P als auch die Konklusion K boolescher Assoziationsregeln aus Teilmengen von A gebildet. Somit gilt für jede durch Auswertung von D_n gewonnene Assoziationsregelmenge $R(D_n)$:

$$R(D_n) \subseteq P(A) \times P(A),$$

wobei $P(A) = \{X \mid X \subseteq A\}$ die Potenzmenge von A ist. Aufgrund der Bedingungen $P \neq \emptyset$, $K \neq \emptyset$ und $P \cap K = \emptyset$ bilden jedoch nicht alle Elemente von $P(A) \times P(A)$ eine gültige Assoziationsregel. Über alle möglichen Datenbasen D_n mit n Datensätzen über m Attributen betrachtet, läßt sich das theoretische Maximum der Regelanzahl, $\max_{D_n} \{|R(D_n)|\}$, durch die folgende Überlegung kombinatorisch ermitteln: Um eine i -elementige Attributmenge X aus A auszuwählen, gibt es $\binom{m}{i}$ Möglichkeiten. Aus einer Menge X lassen sich durch jede Bi-Partitionierung $P \cup K = X$ zwei gültige Regeln bilden, nämlich $P \rightarrow K$ und $K \rightarrow P$. Die Anzahl der möglichen Bi-Partitionierungen einer i -elementigen Menge X beträgt $2^{i-1} - 1$ (vgl. z. B. [Aig96]). Insgesamt ergibt also sich die folgende Regelanzahl:

$$\max_{D_n} \{|R(D_n)|\} = \sum_{i=2}^m \left[\binom{m}{i} \cdot 2 \cdot (2^{i-1} - 1) \right].$$

Man erkennt das exponentielle Wachstum der Regelanzahl in Abhängigkeit von der Attributanzahl m . An dem asymptotischen Wachstumsverhalten ändert jedoch auch die Verwendung einer vereinfachten Regelstruktur nichts: Bezeichnet $R^{(1)}$ die Regelmengemenge, in der die Konklusionen aus nur einem Attribut bestehen, dann können aus einer i -elementigen Attributmenge X nur i zulässige Regeln gebildet werden, und es gilt:

$$\begin{aligned} \max_{D_n} \{|R^{(1)}(D_n)|\} &= \sum_{i=2}^m \left[\binom{m}{i} \cdot i \right] \\ &= m \cdot (2^{m-1} - 1) . \end{aligned}$$

Durch die theoretische Komplexitätsbetrachtung wird deutlich, daß die Erzeugung vollständiger Regelmengen für eine höhere Attributanzahl m nicht durchführbar ist. Die folgenden Überlegungen zeigen jedoch, daß für praktische Problemstellungen diese obere Abschätzung zu hoch gegriffen ist, weil nicht alle Regeln von Interesse sind.

Abhängig von einer konkreten Datenbasis D_n können viele theoretisch mögliche Regeln ohne Informationsgehalt sein, weil sie Strukturen beschreiben, die in D_n überhaupt nicht vorhanden sind. Eine solche Regel $r = P \rightarrow K$ läßt sich z. B. an ihrer relativen Häufigkeit $s(r) = 0$ und Konfidenz $c(r) = 0$ erkennen, denn diese Werte können nur durch $h_n(P \cup K) = 0$ entstehen, also durch Attributmengenkombinationen, die in keinem der Datensätze in D_n vorhanden sind (vgl. die Definitionen 3.3 und 3.4).

Als Folge hiervon sollte jede Regel eine positive Mindesthäufigkeit und Mindestkonfidenz besitzen, um überhaupt erzeugt zu werden. Der niedrigste sinnvolle Schwellenwert berechnet sich aus dem relativen Anteil für genau einen Datensatz in D_n , so daß bei n Datensätzen für jede Regel r gilt:

$$s(r) \geq \frac{1}{n} \quad \text{und} \quad c(r) \geq \frac{1}{n} .$$

Je nach Anwendungsgebiet können aber auch höhere Schwellenwerte als $\frac{1}{n}$ sinnvoll sein: Bereits in Abschnitt 3.1.2 wurde darauf hingewiesen, daß eine Konfidenz $c(r) < 0.5$ eher *gegen* die Konklusion als *für* die Konklusion spricht. Da das boolesche Assoziationsregelkonzept keine Negation kennt, sind z. B. bei Klassifikationsproblemen für zwei Klassen (sog. dichotome Klassifikationsprobleme) nur Regeln mit einer Konfidenz $c(r) \geq 0.5$ nützlich. Ein verallgemeinertes Konzept für diese Schwellenwerte führt zur Definition 3.5 im folgenden Abschnitt.

Anmerkung: Für spezielle Problembereiche existieren in der Literatur neben Konfidenz und relativer Häufigkeit zusätzliche Maße zur Regelbewertung, wie z. B. Distanzmaße ([GO98], [DL98]), Intensity of Implication ([GGP98]), statistische Maße

([BMS97], [SBMU00], [ST96]), Fehler-basierte Maße ([KLKF98]) oder sogenannte Share-basierte Maße ([CHC97], [HCHC98]). Auf diese Erweiterungen wird jedoch in der vorliegenden Arbeit nicht näher eingegangen, weil die bisher eingeführten Eigenschaften $Q = \{s, c\}$ zur Beschreibung heuristischer Diagnoseregeln hinreichend sind.

Problemdefinition

Das Assoziationsregelproblem besteht in der Erzeugung aller für das jeweilige Problemfeld interessanten Assoziationsregeln. Es handelt sich somit um eine Konkretisierung des in Definition 2.3 auf Seite 21 eingeführten generischen Data-Mining-Problems und verwendet als Interessantheitskriterium anwendungsabhängige Schwellenwerte für die Regelkonfidenz und -häufigkeit.

Definition 3.5 (Assoziationsregelproblem): Es seien D_n eine Datenbasis über dem booleschen Wertebereich \mathbb{B} und $R(D_n)$ die vollständige Menge der booleschen Assoziationsregeln über D_n . Weiter sei $\sigma \in [0, 1]$ eine Häufigkeitsschwelle und $\gamma \in [0, 1]$ eine Konfidenzschwelle. Dann besteht das *Assoziationsregelproblem* in der Erzeugung der Regelmenge

$$R(D_n, \sigma, \gamma) := \{r \in R(D_n) \mid s(r) \geq \sigma \text{ und } c(r) \geq \gamma\}. \quad \square$$

Eine theoretische Abschätzung der Regelanzahl $|R(D_n, \sigma, \gamma)|$ in Abhängigkeit von der Häufigkeitsschwelle σ und der Konfidenzschwelle γ ist nicht möglich, ohne die Verteilung der Attributwerte in D_n genau zu kennen. Es sei daher an dieser Stelle auf die empirische Untersuchung der Regelmengen in Diagnoseanwendungen (Abschnitt 5.3.1) verwiesen.

Anmerkungen: (i) Eine Regel $r \in R(D_n, \sigma, \gamma)$ wird im folgenden auch σ -relevant und γ -konfident genannt.

(ii) Es ist offensichtlich, daß für positive Schwellenwerte $\sigma > 0$ und $\gamma > 0$ die Beziehung

$$|R(D_n, \sigma, \gamma)| \leq |R(D_n)|$$

gilt. An der Problemkomplexität ändert sich jedoch nichts, denn zu jedem Paar von Parametern σ und γ läßt sich eine Datenbasis D_n finden, für die $|R(D_n, \sigma, \gamma)| = |R(D_n)|$ gilt. Dies ist allerdings nur mit unrealistischen Datenbasen möglich, deren Einträge zum Beispiel ausschließlich aus dem Wert 1 bestehen. In praxisrelevanten Fällen ergibt sich durch die Verwendung positiver Schwellenwerte ein echter Laufzeitgewinn (vgl. z. B. [CDF⁺00]).

3.2.3 Basisalgorithmus zur Regelerzeugung

Dieser Abschnitt beschäftigt sich mit der Erzeugung von Assoziationsregeln. Zur effizienten Lösung des Assoziationsregelproblems nach Definition 3.5 wird ein an Agrawal et al. ([AMS⁺96]) angelehnter Algorithmus angegeben. Er stellt die Basis für die meisten aus der Literatur bekannten Assoziationsregelalgorithmen dar – und auch für das in Abschnitt 4.7 beschriebene Verfahren zur Generierung von Diagnoseregeln.

Da die relative Häufigkeit $s(r)$ einer Regel $r = P \rightarrow K$ ausschließlich durch die relative Häufigkeit der beteiligten Attributmengen $h_n(P \cup K)$ determiniert ist (vgl. Definition 3.4), läßt sich in Anlehnung an [AIS93] das Problem der Assoziationsregelerzeugung in die folgenden zwei Teile gliedern:

1. Erzeuge für eine vorgegebene Häufigkeitsschwelle σ alle σ -relevanten Attributmengen (engl. *frequent itemsets*), also die Menge

$$F(D_n, \sigma) = \{X \in P(A) \mid X \neq \emptyset \text{ und } h_n(X) \geq \sigma\}.$$

Somit weisen alle Regeln, die aus einer σ -relevanten Attributmenge X mit $|X| > 1$ konstruiert sind, ebenfalls die Mindesthäufigkeit σ auf.

2. Erzeuge für eine vorgegebene Konfidenzschwelle γ alle γ -konfidenten Bipartitionen (P, K) aller Attributmengen aus $F(D_n, \sigma)$, also die Menge

$$R(D_n, \sigma, \gamma) = \{P \cup K \in F(D_n, \sigma) \mid P \cap K = \emptyset \text{ und } c(P \rightarrow K) \geq \gamma\}.$$

Das Ergebnis ist die Menge aller σ -relevanten und γ -konfidenten Assoziationsregeln über D_n .

Diese Vorgehensweise ist vollständig, d.h., es kann keine Regel $r \in R(D_n) \setminus R(D_n, \sigma, \gamma)$ existieren, für die $s(r) \geq \sigma$ und $c(r) \geq \gamma$ gilt. In den folgenden Unterabschnitten werden die beiden Schritte näher erläutert. Darüber hinaus wird auf das Problem der Vermeidung redundanter Regeln eingegangen.

Ermittlung σ -relevanter Attributmengen

Es ist offensichtlich, daß ein naiver Algorithmus, der die relative Häufigkeit aller 2^m Attributteilmenen der Datenbasis D_n bestimmen müßte, für den praktischen Einsatz nicht geeignet ist. Die Monotonieeigenschaft der Attributmengenhäufigkeiten (Satz 3.1) ermöglicht es jedoch, durch eine geschickte Erzeugungsreihenfolge viele Attributmengen von der Suche auszuschließen. Die Grundidee ist dabei, stufenweise

Eingabe: Datenbasis D_n mit Attributmenge A , Häufigkeitsschwelle σ
 Ausgabe: Menge aller σ -relevanten Attributmengen $F(D_n, \sigma) =$
 $\{X \in P(A) \mid X \neq \emptyset \text{ und } h_n(X) \geq \sigma\}$

- (1) $s \leftarrow 1$
- (2) $F_s \leftarrow \{\{A_i\} \mid A_i \in A \text{ und } h_n(A_i) \geq \sigma\}$
- (3) **while** $F_s \neq \emptyset$ **do begin**
- (4) $s \leftarrow s + 1$
- (5) $C_s \leftarrow \{X = Y \cup Z \mid Y \in F_{s-1}, Z \in F_{s-1} \text{ und } |X| = s\}$
- (6) $F_s \leftarrow \{X \in C_s \mid h_n(X) \geq \sigma\}$
- (7) **end**
- (8) **return** $F(D_n, \sigma) \leftarrow \bigcup_{i=1}^s F_i$

Algorithmus 3.1: Ermittlung der Attributmengen mit Mindesthäufigkeit σ in einer Datenbasis D_n

einmal als σ -relevant ermittelte Attributmengen solange um ein zusätzliches Attribut zu erweitern, bis ihre relative Häufigkeit unter die Schwelle σ gesunken ist. Diese Überlegungen führen zu Algorithmus 3.1; er basiert auf »Apriori« aus [AMS⁺96].

Der Algorithmus 3.1 geht stufenweise vor, wobei in jeder Stufe s mit C_s die Menge der s -elementigen Kandidaten (potentiell σ -relevante Mengen) und mit F_s die Menge aller s -elementigen σ -relevanten Attributmengen bezeichnet wird. Zunächst werden mit F_1 alle Einzelattribute A_i ermittelt, die in D_n die Mindesthäufigkeit σ besitzen (Zeile 2). Danach werden diese zu zweielementigen Kandidaten C_2 kombiniert (Zeile 5), von denen die σ -relevanten Attributmengen der Menge F_2 zugeordnet werden (Zeile 6). Ist F_2 nicht leer, werden die dreielementigen Kandidaten C_3 durch Kombination der Mengen in F_2 erzeugt usw. Der Algorithmus stoppt, wenn keine σ -relevante Attributmenge der nächsthöheren Kardinalität existiert (Zeile 3); das Ergebnis ist die Vereinigung aller F_i (Zeile 8).

Laufzeitkritisch sind vor allem die Kandidatengenerierung (Erzeugung der Kandidatenmengen C_s aus F_{s-1} , Zeile 5) sowie die Ermittlung der relativen Häufigkeiten $h_n(X)$ für Attributmengen X (Zeilen 2 und 6). Eine grobe Laufzeitabschätzung ist dabei wie folgt möglich: Die Kandidatenmenge C_s kann maximal $\binom{m}{s}$ Attributmengen enthalten. Für jede Menge $X \in C_s$ ist die relative Häufigkeit $h_n(X)$ zu bestimmen; hierzu wird höchstens ein Durchlauf über die n Datensätze benötigt. In jeder Stufe s ergibt sich somit eine Laufzeit von $O(n \cdot \binom{m}{s})$. Unter der Annahme, daß in einer konkreten Anwendung die Attributanzahl m konstant ist (festes Datenbankschema),

Eingabe: Datenbasis D_n , Menge σ -relevanter Attributmengen $F(D_n, \sigma)$, Konfidenzschwelle γ
 Ausgabe: Regelmenge $R(D_n, \sigma, \gamma)$

```

(1)  $R \leftarrow \emptyset$ 
(2) for each  $X \in F(D_n, \sigma)$  do begin
(3)   if  $|X| > 1$  then
(4)     for each  $P \neq \emptyset \subset X$  do
(5)       if  $h_n(X)/h_n(P) \geq \gamma$  then
(6)          $R \leftarrow R \cup (P \rightarrow X \setminus P \mid c = \frac{h_n(X)}{h_n(P)}, s = s(X))$ 
(7)   end
(8) return  $R$ 

```

Algorithmus 3.2: Erzeugung γ -konfidenter Assoziationsregeln

läßt sich die Gesamtlaufzeit des Algorithmus als quasi-linear in n bezeichnen.

Erzeugung γ -konfidenter Regeln

Algorithmus 3.2 stellt die Vorgehensweise zur Erzeugung der Assoziationsregeln dar. Es erfolgt die Aufteilung aller σ -relevanter Attributmengen $X \in F(D_n, \sigma)$, die mindestens zweielementig sind, in die möglichen Prämissen P und Konklusionen K mit $P \cup K = X$ (Zeilen 3 und 4). Für jede so entstandene Regel $P \rightarrow K$ muß geprüft werden, ob sie die vorgegebene Mindestkonfidenz γ besitzt (Zeile 5). Ist dies nicht der Fall, wird sie verworfen, ansonsten erfolgt die Berechnung der Konfidenz und relativen Häufigkeit (Zeile 6), und sie wird zur Regelmenge R hinzugefügt. Nach der Verarbeitung aller σ -relevanter Attributmengen wird R als Ergebnis zurückgegeben.

Laufzeitkritisch sind die Berechnungen in den Zeilen 5 und 6. Da eine s -elementige Menge 2^s Teilmengen enthält, wird Zeile 5 $O(|F(D_n, \sigma)| \cdot 2^s)$ -mal ausgeführt, wobei s die Mächtigkeit der größten Attributmenge in $F(D_n, \sigma)$ ist. Für den Fall, daß die relativen Häufigkeiten $h_n(X)$ und $h_n(P)$ berechnet werden müssen, beträgt der Aufwand zur Regelerzeugung insgesamt $O(n \cdot |F(D_n, \sigma)| \cdot 2^s)$. Liegen die Werte dagegen noch aus der Phase der Attributmengenerzeugung (Algorithmus 3.1) vor, entfällt der Faktor n , und die Laufzeit ist unabhängig von der Anzahl der Datensätze.

Eliminierung redundanter Regeln

Abschließend soll auf den Aspekt der Redundanz bzw. Subsumption von Regeln eingegangen werden, der für die Effizienz von Diagnoseregeln von Bedeutung ist (vgl. Abschnitt 4.7): Eine Regel $r \in R(D_n, \sigma, \gamma)$ heißt *redundant*, wenn mit der verkleinerten Regelmenge $R(D_n, \sigma, \gamma) \setminus \{r\}$ die gleichen Aussagen wie mit der Ausgangsmenge $R(D_n, \sigma, \gamma)$ möglich sind. In diesem Fall wird r durch eine Regel $r' \in R(D_n, \sigma, \gamma)$ subsumiert, d. h., r' macht eine stärkere Aussage als r und ist für jeden Datensatz $d \in D_n$ gültig, für den auch r gültig ist.

Der Begriff der Subsumption stammt aus der Aussagenlogik und bezeichnet hier die Teilmengenbeziehung zweier Klauseln π, π' einer Formel α , die in konjunktiver Normalform vorliegt (siehe z. B. [KL94]). Gilt $\pi' \subseteq \pi$, dann ist mit jeder erfüllenden Variablenbelegung (Interpretation) für π' unmittelbar auch π erfüllt; es wird also π von π' subsumiert. Auf diese Weise führt eine syntaktische Analyse der Klauseln (Teilmengentest) zu einer semantischen Aussage (Erfüllbarkeit). Zwar können auch Assoziationsregeln als Klauseln aufgefaßt werden, da hier allerdings zusätzlich die Regelkonfidenzen berücksichtigt werden müssen, ist der Subsumptionsbegriff nicht direkt übertragbar, wie im folgenden gezeigt wird.

Eine Menge boolescher Assoziationsregeln $R(D_n)$ läßt sich in eine aussagenlogische Formel α in konjunktiver Normalform transformieren; α hat dann die Gestalt

$$\alpha = \bigwedge_{p_1, \dots, p_i \rightarrow k_1, \dots, k_j \in R(D_n)} (\neg p_1 \vee \dots \vee \neg p_i \vee k_1) \wedge \dots \wedge (\neg p_1 \vee \dots \vee \neg p_i \vee k_j).$$

Folglich ist auf diese Formel α das aussagenlogische Subsumptionskriterium anwendbar. Weiter gilt, daß jeder Datensatz $d \in D_n$ als Interpretation für α angesehen werden kann; eine *erfüllende* Interpretation ist d aber nur für eine Teilmenge der Klauseln von α bzw. der Regeln $R(D_n)$. Sei diese Teilmenge $R(d)$ genannt, dann gilt

$$R(d) := \{P \rightarrow K \in R(D_n) \mid (P \cup K) \subseteq d\} \subseteq R(D_n).$$

Subsumiert eine Regel $r' \in R(D_n)$ eine andere Regel $r \in R(D_n)$ aussagenlogisch, so müßte jede Interpretation, die für r' erfüllend ist, auch für r erfüllend sein, es müßte also gelten:

$$\forall d \in D_n : r' \in R(d) \Rightarrow r \in R(d).$$

Diese semantische Beziehung darf allerdings nur dann aus einem Teilmengentest gefolgert werden, wenn beide Regeln sichere Implikationen sind, d. h. $c(r) = c(r') = 1$. Ist mindestens eine der Regelkonfidenzen $c(r)$ bzw. $c(r')$ kleiner als 1, können Datensätze existieren, in denen r' gilt, nicht jedoch r ; somit würden zu viele Regeln

fälschlicherweise als subsumiert erkannt. Soll trotzdem nur eine syntaktische Analyse der Regeln vorgenommen werden, ist es als Näherung sinnvoll, das aussagenlogische Subsumptionskriterium um die Berücksichtigung der Regelkonfidenzen zu erweitern. Dies führt zu Definition 3.6, nach der eine Regel r nur dann als von r' subsumiert gelten darf, wenn sie nicht zuverlässiger ist, wenn also ihre Konfidenz $c(r)$ nicht größer als $c(r')$ ist.

Definition 3.6 (Subsumption von Assoziationsregeln): Es seien $r = P \rightarrow K$ und $r' = P' \rightarrow K'$ Assoziationsregeln über einer Datenbasis D_n . Dann wird r durch r' *subsumiert* ($r' \succ r$), wenn gilt:

$$P \supseteq P' \text{ und } K \subseteq K' \text{ und } c(r) \leq c(r').$$

Die Regel r wird *speziellere Regel*, und die Regel r' wird *allgemeinere Regel* genannt. \square

Beispiel 3.2: Es sei für alle Regeln dieses Beispiels eine Konfidenz von 1.0 angenommen. Dann zeigt die folgende Tabelle eine Auswahl von Regeln, die von der Regel $p_1 \rightarrow k_1, k_2$ subsumiert werden (p_2 ist ein weiteres Prämissenattribut).

Subsumierte Assoziationsregel	Aussagenlogisches Äquivalent
$p_1, p_2 \rightarrow k_1, k_2$	$(\neg p_1 \vee \neg p_2 \vee k_1) \wedge (\neg p_1 \vee \neg p_2 \vee k_2)$
$p_1 \rightarrow k_1$	$(\neg p_1 \vee k_1)$
$p_1, p_2 \rightarrow k_1$	$(\neg p_1 \vee \neg p_2 \vee k_1)$

\square

Der Algorithmus 3.2 zur Regelerzeugung läßt sich einfach zur Erzeugung nichtredundanter Regeln erweitern, indem z. B. die redundanten Regeln aus der vollständigen Regelmengende durch einen Subsumptionstest herausgefiltert werden (siehe Zeilen 8 bis 10 in Algorithmus 3.3). In der Praxis zeigt sich hierdurch eine deutliche Reduktion der Regelanahl: So wird in Abschnitt 5.3.1 beschrieben, daß mehr als 90% der Regeln eines Diagnosesystems redundant sind und eliminiert werden können.

3.3 Kategoriale Assoziationsregeln

Im Konzept der kategorialen Assoziationsregeln sind für die Attribute der Datenbasis D_n nicht nur zweielementige Wertebereiche, sondern alle endlichen Wertebereiche zulässig. Hierdurch werden ihre Ausdrucksmöglichkeiten gegenüber den boole-

Eingabe: Datenbasis D_n , Menge σ -relevanter Attributmengen $F(D_n, \sigma)$, Konfidenzschwelle γ
 Ausgabe: redundanzfreie Regelmenge $R(D_n, \sigma, \gamma)$

```

(1)  $R \leftarrow \emptyset$ 
(2) for each  $X \in F(D_n, \sigma)$  do begin
(3)   if  $|X| > 1$  then
(4)     for each  $P \neq \emptyset \subset X$  do
(5)       if  $h_n(X)/h_n(P) \geq \gamma$  then
(6)          $R \leftarrow R \cup (P \rightarrow X \setminus P \mid c = \frac{h_n(X)}{h_n(P)}, s = s(X))$ 
(7)     end
(8) for each  $r \in R$  do
(9)   if  $\exists r' = P' \rightarrow K' \in R : P \supseteq P' \wedge K \subseteq K' \wedge c(r) \leq c(r')$  then
(10)     $R \leftarrow R \setminus r$ 
(11) return  $R$ 

```

Algorithmus 3.3: Erzeugung redundanzfreier γ -konfidenter Assoziationsregeln

schen Regeln stark erweitert. Im folgenden wird eine formale Definition kategorialer Assoziationsregeln gegeben (Abschnitt 3.3.1) und ihre Erzeugung skizziert (Abschnitt 3.3.2).

3.3.1 Definition

Während boolesche Assoziationsregeln nur Aussagen über das Vorliegen von Attributwerten ungleich 0 in den Datensätzen einer Datenbasis repräsentieren, können kategoriale Assoziationsregeln auch quantitative Aspekte ausdrücken. Jedes Attribut A_i der Datenbasis D_n hat einen endlichen Wertebereich von $k_i \in \mathbb{N}$ diskreten Kategorien, die z. B. einen symbolischen Wert, eine natürliche Zahl oder ein Intervall bezeichnen. Jeder Ausdruck in der Regelprämisse bzw. Regelkonklusion besteht dann aus einem Attribut / Kategorie-Vergleich wie in Definition 3.7 angegeben. Ein Attribut / Kategorie-Vergleich wird im folgenden auch als *Tupel* bezeichnet.

Definition 3.7 (Kategoriale Assoziationsregel): Es sei D_n eine Datenbasis mit $m \in \mathbb{N}$ Attributen A_i über den endlichen Wertebereichen $\text{dom}(A_i)$. Weiter seien $P = \{A_{i_1}, \dots, A_{i_p}\}$ und $K = \{A_{u_1}, \dots, A_{u_k}\}$ Attributmengen, für die gilt: $P \cap K = \emptyset$.

Dann wird mit Attributwerten $a_{j_x} \in \text{dom}(A_{i_x})$ bzw. $a_{v_x} \in \text{dom}(A_{u_x})$ eine Regel der Form

$$A_{i_1} = a_{j_1}, \dots, A_{i_p} = a_{j_p} \rightarrow A_{u_1} = a_{v_1}, \dots, A_{u_k} = a_{v_k} \mid c, s$$

kategoriale Assoziationsregel über D_n genannt.

Ein Attribut / Kategorie-Vergleich $A_{i_x} = a_{j_x}$ wird auch als Tupel $\langle A_{i_x}, a_{j_x} \rangle$ dargestellt.

Bezeichnen

$$T_P := \{ \langle A_{i_1}, a_{j_1} \rangle, \dots, \langle A_{i_p}, a_{j_p} \rangle \} \quad \text{und} \quad T_K := \{ \langle A_{u_1}, a_{v_1} \rangle, \dots, \langle A_{u_k}, a_{v_k} \rangle \}$$

die Tupelmengen der Regelprämisse bzw. der Regelkonklusion, dann kann die Berechnung der Konfidenz c und der relativen Häufigkeit s der Regel analog zu Definition 3.4 durch

$$c = \frac{h_n(T_P \cup T_K)}{h_n(T_P)} \quad \text{und} \quad s = h_n(T_P \cup T_K),$$

erfolgen, wobei in diesem Fall h_n die relative Häufigkeit einer Tupelmenge ist (siehe Definition 4.11 auf Seite 107). \square

Anmerkungen: (i) In der Literatur werden Regeln dieser Art manchmal auch als *quantitative Assoziationsregeln* bezeichnet (siehe z.B. [SA96], [MY97] oder [BW98]).

(ii) Boolesche Assoziationsregeln können als Spezialfall der kategorialen Regeln aufgefaßt werden, wenn jeder Attributwertebereich aus den Kategorien 0 und 1 besteht.

In vielen praktischen Anwendungsbereichen haben die Datenbankattribute kontinuierliche Wertebereiche. Hier ist es notwendig, zunächst eine geeignete Diskretisierung, also Einteilung der Wertebereiche in diskrete Intervalle, vorzunehmen und anschließend jedem Intervall eine eigene Kategorie zuzuweisen. Das Ergebnis ist dann eine kategoriale Datenbasis D_n^k . In Abschnitt 4.5 wird auf diese Vorgehensweise am Beispiel der Diagnoseanwendung detailliert eingegangen.

3.3.2 Erzeugung kategorialer Assoziationsregeln

In diesem Abschnitt werden zwei Alternativen zur Erzeugung kategorialer Assoziationsregeln skizziert. Während die erste Alternative eine Transformation der Datenbasis durchführt, um die vorhandenen booleschen Regelerzeugungsalgorithmen anwenden zu können, werden als zweite Alternative diese bestehenden Algorithmen an das neue Problem angepaßt.

1. Alternative: Reduktion auf die Erzeugung boolescher Regeln

Das Problem der Erzeugung kategorialer Assoziationsregeln mit vorgegebener Mindesthäufigkeit σ und Mindestkonfidenz γ kann durch Reduzierung auf das Problem der Erzeugung boolescher Regeln (siehe Definition 3.5) gelöst werden. Dazu sind die folgenden Schritte notwendig:

1. Transformiere die kategoriale Datenbasis D_n^k in eine boolesche Datenbasis D_n^b .
2. Löse das boolesche Assoziationsregelproblem wie in Abschnitt 3.2.3 beschrieben. Das Ergebnis ist eine Regelmenge $R(D_n^b, \sigma, \gamma)$.
3. Transformiere die Regelmenge $R(D_n^b, \sigma, \gamma)$ in eine kategoriale Regelmenge $R(D_n^k, \sigma, \gamma)$ mit kategorialen Regeln.

Bei der Transformation von D_n^k nach D_n^b (Schritt 1) wird zu jeder Kategorie a_j eines Attributs $A_i \in D_n^k$ ein boolesches Attribut $A_j \in D_n^b$ erzeugt. Dies erhält in einem Datensatz $d^b \in D_n^b$ den Wert 1, wenn der korrespondierende Datensatz d^k für A_i den Wert a_j besitzt; andernfalls erhält A_j in d^b den Wert 0. Somit besteht D_n^b aus $\sum_{i=1}^m |dom(A_i)|$ Attributen über dem Wertebereich \mathbb{B} , und die Algorithmen 3.1 und 3.2 sind anwendbar (Schritt 2). Die Rücktransformation der booleschen Regeln (Schritt 3) geschieht dann durch die Umkehrabbildung der Attributwerte; Konfidenz, relative Häufigkeit und Subsumptionseigenschaften der Regeln bleiben bei der Transformation unverändert.

2. Alternative: Erweiterung der booleschen Regelerzeugung

Die erstgenannte Vorgehensweise zur Erzeugung von kategorialen Regeln ist ineffizient, weil sich zwar viele Attribute in D_n^b gegenseitig ausschließen⁴, dies aber bei der Kandidatengenerierung für σ -relevante Attributmengen (Zeile 5 in Algorithmus 3.1) nicht berücksichtigt wird. Eine bessere Alternative ist daher die Erweiterung der beiden Algorithmen aus Abschnitt 3.2.3, wobei sich die Anpassungen im wesentlichen auf zwei Aspekte beziehen:

- Ersetzung der Repräsentationen jedes Attributs A_i durch Attribut / Wert-Tupel $\langle A_i, a_j \rangle$ für alle $a_j \in dom(A_i)$, um die Verarbeitung von Tupelmengen zu ermöglichen.

⁴ Aufgrund der Konstruktion von D_n^b in Schritt 1 besitzt in jedem Datensatz genau eins der booleschen Attribute, die aus demselben kategorialen Attribut in D_n^k entstanden sind, den Wert 1. Also kann in D_n^b keine Kombination dieser booleschen Attribute mit Wert 1 existieren.

- Sicherstellung, daß bei der Kandidatengenerierung niemals zwei gleiche kategoriale Attribute in einer Tupelmengen vorkommen, d. h., für je zwei Tupel $\langle A_{i_1}, a_{j_1} \rangle$ und $\langle A_{i_2}, a_{j_2} \rangle$ aus derselben Menge muß gelten: $A_{i_1} \neq A_{i_2}$.

Durch die skizzierten Veränderungen werden dann anstelle von σ -relevanten *Attributmengen* σ -relevante *Tupelmengen* erzeugt und als Grundlage zur Bildung von Regeln verwandt (siehe auch Abschnitt 3.4.2). Ihre konkrete Realisierung wird in Algorithmus 4.7.2 (Erzeugung von Diagnoseregeln, Seite 107) dargestellt.

3.4 Klassifikationsregeln

Klassifikationsregeln können zur Lösung von Klassifikationsaufgaben für Objekte, Ereignisse oder Fälle genutzt werden, indem sie diese anhand einer Eigenschaftsbeschreibung in eine von mehreren bekannten Kategorien (Klassen) einordnen. Dabei bestehen die Prämissen aus (evtl. mehreren verknüpften) Ausdrücken über einzelne Eigenschaften, während die Konklusionen jeweils genau eine Klassenzuordnung beinhalten. Die genaue Definition der Klassifikationsregeln ist anwendungsabhängig; im Rahmen der heuristischen Fehlerdiagnose werden die Eigenschaftsbeschreibungen durch Symptome und die Klassen durch Diagnosen repräsentiert.

Im folgenden wird für die Lösung von Klassifikationsaufgaben ausschließlich die Verwendung von kategorialen Assoziationsregeln betrachtet. Zunächst wird in Abschnitt 3.4.1 das Klassifikationsregelproblem definiert, bevor in Abschnitt 3.4.2 einige Aspekte der Erzeugung und Anwendung von Assoziationsregeln für die Klassifikation angesprochen werden.

3.4.1 Klassifikationsregelproblem

Im Gegensatz zum Assoziationsregelproblem, bei dem die Erzeugung einer *vollständigen* Regelmengen mit Mindestkonfidenz und Mindesthäufigkeit im Vordergrund steht (siehe Abschnitt 3.2.2), geht es beim Klassifikationsregelproblem üblicherweise um die Erzeugung einer *kleinen* Regelmengen, die für das zugrundeliegende Problemfeld einen *geeigneten* Klassifikator bildet (vgl. [QR89], [LHM98]). Hierbei wird versucht, aus einer für den Problembereich repräsentativen Datenbasis mit bereits klassifizierten Beispielobjekten eine möglichst genaue Abbildungsvorschrift mit generalisierenden Eigenschaften abzuleiten.

Die Klassifikationsabbildung soll also in der Lage sein, auch neue Objekte korrekt in Klassen einzuordnen. Um diese Eigenschaft bei fehlenden Testobjekten überprüfen zu können, wird im allgemeinen innerhalb der Datenbasis D_n eine Teilmenge

$D_l \subset D_n$ als Lernbasis bestimmt. Die (bekannten) Objektklassifizierungen der nicht in D_l enthaltenen Datensätze dürfen dann nicht während der Lernphase berücksichtigt werden. Als Testbasis zur Überprüfung der Klassenvorhersage durch eine gelernte Klassifikationsabbildung können die Restmenge $D_n \setminus D_l$ oder die gesamte Datenbasis D_n genutzt werden. Dies führt zu der folgenden Definition des Klassifikationsregelproblems.

Definition 3.8 (Klassifikationsregelproblem, Klassifikator): Es sei D_n eine kategoriale Datenbasis mit der Attributmenge $A = \{A_1, \dots, A_m, C\}$, in der ein Attribut das ausgezeichnete Klassenattribut C ist. Dann besteht das *Klassifikationsregelproblem* darin, aus einer repräsentativen Teilmenge $D_l \subseteq D_n$ eine Klassifikationsregelmenge $R(D_l)$ zu erzeugen, mit der sich eine Abbildung

$$\kappa : \text{dom}(A_1) \times \dots \times \text{dom}(A_m) \rightarrow \text{dom}(C)$$

realisieren läßt, die in bezug auf die Gesamtdatenbasis D_n die Anzahl der korrekten Klassifizierungen maximiert:

$$|\{d = \langle a_1, \dots, a_m, c \rangle \in D_n \mid \kappa(a_1, \dots, a_m) = c\}| \longrightarrow \max.$$

κ wird auch *Klassifikator* genannt. □

Ein regelbasierter Klassifikator κ ist nicht als m -stellige mathematische Funktion aufzufassen, sondern besteht aus einer Regelmenge und einer Regelanwendungsstrategie zur Ableitung einer Klassenvorhersage aus einem m -stelligen Attributwertevektor. Hierbei kann es Problembereiche geben, in denen eine eindeutige Klassifizierung der Objekte nicht möglich ist; in diesen Fällen muß der Klassifikator für die Klassen Wahrscheinlichkeiten berechnen können bzw. nach maximaler Wahrscheinlichkeit entscheiden. Einige solcher Strategien werden in Abschnitt 4.8 mit Bezug auf die Diagnoseanwendung erläutert.

3.4.2 Klassifikation mit Assoziationsregeln

Durch die Beschränkung der Konklusion auf genau eine Klassenzuordnung erhält man Assoziationsregeln, die für die Lösung von Klassifikationsaufgaben prinzipiell geeignet sind. In Anlehnung an Liu et al., die für diese Regelart in [LHM98] die Bezeichnung *Class Association Rules (CAR)* verwenden, wird in der vorliegenden Arbeit der Begriff *klassifizierende Assoziationsregeln* gewählt.

Um zur assoziationsregelbasierten Lösung einer Klassifikationsaufgabe zu gelangen, ist die Verknüpfung des Assoziationsregelproblems (Definition 3.5) mit dem Klassifikationsregelproblem (Definition 3.8) notwendig. Dies läßt sich auf der Grundlage

einer kategorialen Datenbasis D_n , die ein Attribut C mit einer Klassenzuordnung c_i für jeden Datensatz enthält, in Anlehnung an [LHM98] durch die folgende Vorgehensweise realisieren:

1. Erzeuge zu einer Häufigkeitsschwelle σ und einer Konfidenzschwelle γ die vollständige Menge der klassifizierenden Assoziationsregeln

$$R_C(D_I, \sigma, \gamma) := \{r \in R(D_I, \sigma, \gamma) \mid r = P \rightarrow (C = c_i) \text{ und } c_i \in \text{dom}(C)\},$$

wobei $D_I \subseteq D_n$ eine repräsentative Lernbasis ist.

2. Erzeuge durch die Auswertung von Lernbasis D_I und Regelmenge $R_C(D_I, \sigma, \gamma)$ einen für D_n geeigneten Klassifikator κ .

Im Unterschied zu bekannten Klassifikationssystemen wie z. B. ID3 / C4.5 von Quinlan ([Qui93]), bei denen zur Bildung eines Klassifikators die Regelmenge durch *lokal* optimale Entscheidungen sukzessive aufgebaut wird, ist bei der obigen Vorgehensweise mit der Assoziationsregelmenge $R_C(D_I, \sigma, \gamma)$ nach Schritt 1 die Gesamtheit der klassifikationsrelevanten Strukturen in der Datenbasis bekannt. Ein wesentlicher Vorteil in bezug auf die Klassifikationsgüte entsteht somit dadurch, daß im 2. Schritt stets auf die *global* besten Regeln zugegriffen werden kann (vgl. [LHM98]). Im folgenden werden einige Aspekte der beiden Schritte näher erläutert.

Schritt 1: Erzeugung der klassifizierenden Assoziationsregeln

Auch die Erzeugung klassifizierender Assoziationsregeln läßt sich durch eine Problemreduktion vornehmen. Eine einfache, aber nicht sehr effiziente Strategie zur Erzeugung aller klassifizierenden Assoziationsregeln besteht darin, zunächst alle kategorialen Regeln mit einem in Abschnitt 3.3 beschriebenen Verfahren zu erzeugen und danach diejenigen Regeln zu entfernen, deren Konklusion nicht genau eine Klassenzuordnung darstellt. Bei der Verwendung realistischer Datenbasen ist jedoch nur ein geringer Anteil der kategorialen Regeln zur Klassifikation geeignet, so daß im ersten Schritt sehr viele unzulässige Regeln erzeugt werden.

Eine bessere Alternative besteht in der Modifizierung des Algorithmusses zur Erzeugung σ -relevanter Tupelmengen. Dabei ist die Kandidatengenerierung so vorzunehmen, daß jede Tupelmenge genau einmal das Klassenattribut C enthält, d. h., jede $p + 1$ -elementige Kandidatenmenge ist dann von der Gestalt

$$\{\langle A_{i_1}, a_{j_1} \rangle, \dots, \langle A_{i_p}, a_{j_p} \rangle, \langle C, c_i \rangle\}.$$

Hierdurch vereinfacht sich auch die Regelerzeugung: Eine σ -relevante Tupelmengemenge kann jetzt nur noch auf eine Weise sinnvoll partitioniert werden, nämlich zur Erzeugung der Regel

$$A_{i_1} = a_{j_1} \wedge \dots \wedge A_{i_p} = a_{j_p} \rightarrow C = c_i.$$

In Abschnitt 4.7 wird am Beispiel der Erzeugung von Diagnoseregeln diese Vorgehensweise näher beschrieben. Zusätzliche Maßnahmen, die bei der Generierung klassifizierender Assoziationsregeln eine Effizienzsteigerung durch Suchraumbeschneidung bewirken, finden sich z. B. in [Bay97] und [WZH00]. Diese Strategien sind jedoch teilweise nicht informationserhaltend, d. h., die Vollständigkeit der resultierenden Regelmengemenge kann von ihnen nicht garantiert werden.

Schritt 2: Erzeugung des Klassifikators

Die vollständige Menge klassifizierender Assoziationsregeln wird analysiert, um einen Klassifikator mit maximaler Klassifikationsleistung zu bilden. Wie bereits in Abschnitt 3.4.1 ausgeführt, gehören zu einem Klassifikator eine Regelmengemenge und ein Verfahren zur Regelanwendung. Das bei der Klassifikatorbildung bestehende Optimierungsproblem (siehe Definition 3.8) wird anhand der folgenden Überlegung deutlich: Angenommen, es soll ein neues Objekt klassifiziert werden. Dann steigt mit der Größe der Regelmengemenge auch die Wahrscheinlichkeit, daß mehrere Regeln mit unterschiedlichen, also widersprüchlichen Klassenvorhersagen anwendbar sind. Dem gegenüber sinkt gleichzeitig die Wahrscheinlichkeit, daß keine Regel anwendbar und damit keine Klassenvorhersage möglich ist.

Hieraus folgt, daß eine Vielzahl von Strategien denkbar sind, die sich in der Verteilung des Optimierungsaufwands unterscheiden. Die beiden Extrema, zwischen denen sich ein konkretes Verfahren befinden kann, sind dabei:

1. *Reine Regelmengemengenoptimierung*: Aus der Gesamtregelmengemenge wird eine bezüglich der Klassifikationsleistung optimale Teilmenge ermittelt. Bei der Regelanwendung erfolgt lediglich die Ermittlung einer anwendbaren Regel, die sofort zur Klassenvorhersage genutzt wird.
2. *Reine Regelanwendungsoptimierung*: Die Gesamtregelmengemenge bleibt unverändert. Mit einem optimalen Verfahren wird das hierin befindliche Klassifizierungswissen zu einer korrekten Klassenvorhersage genutzt. Zur Klassenvorhersage werden insbesondere die Regelkonfidenzen ausgewertet.

Die Ansätze zur Regelmengemengenoptimierung sind eher dem Bereich der anwendungsabhängigen Klassifikation zuzuordnen (z. B. [LHM98]: Classification Based on Associations (CBA) oder [DZWL99]: Classification by Aggregating Emerging Patterns

(CAEP)), während die Ansätze der Regelanwendungsoptimierung eher aus dem Bereich der Diagnose stammen (z. B. [Sho76]: Konfidenzverrechnung in MYCIN oder die Hypothesengenerierung aus Abschnitt 4.8.3). Ab einer bestimmten Problemgröße kann jedoch kein bekanntes Verfahren die Erzeugung eines optimalen Klassifikators, also eine Lösung des Klassifikationsregelproblems nach Definition 3.8, garantieren; in diesen Fällen wird nach Näherungslösungen gesucht.

Im Rahmen der Diagnoseanwendung erfolgt in Abschnitt 4.7.3 die nähere Beschreibung der beiden Optimierungsansätze, wobei verschiedene Verfahren zur Ermittlung eines Klassifikators diskutiert werden.

3 ASSOZIATIONSREGELN

4 Diagnose hydraulischer Systeme mit Assoziationsregeln

In diesem Kapitel wird das in Abschnitt 2.3.3 entwickelte Konzept zur automatischen Erzeugung von heuristischen Regelmodellen am Problembereich der Fehlerdiagnose für hydraulische Anlagen umgesetzt. Dieses Konzept beruht darauf, in zwei Schritten kausale Verhaltensmodelle einer technischen Anlage in ein heuristisches Assoziationsregelmodell für diese Anlage zu transformieren.

Im ersten Schritt wird mit dem Simulationswerkzeug *art deco* aus den Verhaltens- und Strukturmodellen einer Hydraulikanlage eine Simulationsdatenbasis erzeugt. Diese Datenbasis wird im zweiten Schritt unter Verwendung von Data-Mining-Verfahren in ein Assoziationsregelmodell transformiert. Die Regeln repräsentieren Diagnosewissen, mit dem in einer realen Fehlersituation durch einfache Regelanwendung schnell von den Beobachtungen auf ihre Ursachen geschlossen werden kann.

Im Rahmen dieses Kapitels wird insbesondere der zweite Schritt näher behandelt. Dazu werden die Architektur und die theoretischen Grundlagen eines Systems vorgestellt, welches durch eine Analyse der Simulationsdatenbasis

1. diagnoserelevante Designentscheidungen bei der Entwicklung neuer hydraulischer Anlagen unterstützt, indem es Vorschläge zur Wahl geeigneter Meßstellen macht,
2. anlagenspezifische Assoziationsregelmodelle als Wissensbasis für die heuristische Diagnose automatisch generiert und
3. eine assoziationsregelbasierte Fehlererkennung für hydraulische Anlagen durchführt.

Einteilung des Kapitels: Abschnitt 4.1 erläutert die für das Verständnis der weiteren Ausführungen notwendigen Grundlagen hydraulischer Anlagen. In Abschnitt 4.2 wird das Gesamtkonzept zur automatischen Generierung von Diagnosesystemen für hydraulische Anlagen dargestellt. Die Einzelaspekte dieses Konzepts werden in den Abschnitten 4.3 bis 4.8 detaillierter betrachtet.

4.1 Einführende Bemerkungen zur Hydraulik

Die *Hydraulik* ist die Lehre von den Kräften und Bewegungen, die mit Hilfe von Druckflüssigkeiten erzeugt werden; sie gliedert sich in die Bereiche *Hydrostatik* und *Hydrodynamik*, in denen die physikalischen Eigenschaften und Verhaltensweisen von Flüssigkeiten in ruhendem bzw. bewegtem Zustand untersucht werden (vgl. [DD98]).

In Abschnitt 4.1.1 wird eine kurze Einführung in den Aufbau und die Funktionsweise hydraulischer Anlagen gegeben. Abschnitt 4.1.2 beschäftigt sich mit den physikalischen Größen und der Meßtechnik in der Hydraulik, wobei jeweils nur die im Zusammenhang mit der Realisierung einer wissensbasierten Diagnose wichtigen Aspekte dargestellt werden. Abschnitt 4.1.3 behandelt die Anforderungen an eine rechnergestützte Diagnose hydraulischer Anlagen.

4.1.1 Hydraulische Anlagen

Prinzip und Aufbau

Hydraulische Anlagen (auch hydraulische Schaltkreise genannt, engl. *hydraulic circuits*) sind technische Anlagen, deren Funktionalität auf dem Prinzip der hydraulischen Leistungsumformung beruht. Dabei kommt es zu einer Umwandlung von mechanischer Leistung in hydraulische Leistung, üblicherweise durch eine Hydropumpe, die von einem Elektro- oder Verbrennungsmotor angetrieben wird. Die erzeugte hydraulische Leistung wird dann über Leitungen sowie Steuer- und Regelventile einem oder mehreren Verbrauchern, z. B. Hydromotoren oder -zylindern, zugeführt, die eine Rückwandlung in mechanische Leistung vornehmen.

Die wesentlichen Komponenten einer hydraulischen Anlage lassen sich also in Leistungsversorgungsteile, Steuerungsteile und Arbeitsteile (sogenannte Abtriebsteile) untergliedern (siehe Abbildung 4.1). Die Übertragung der hydraulischen Leistung geschieht mittels einer Druckflüssigkeit, dem Hydrauliköl, das durch Rohr- oder Schlauchleitungen fließt. Des weiteren können sich Zubehörteile wie Filter, Kühler oder Behälter zur Aufbereitung der Flüssigkeit im Leitungssystem befinden und elektrische Komponenten wie Relais, Schalter etc. zur Ansteuerung der mit Elektromagneten betätigten Schaltventile vorhanden sein (vgl. z. B. [Mat91]).¹

Ein großer Vorteil hydraulischer Anlagen ist ihre hohe Kraftdichte; so können z. B. Hydromotoren bei gleicher Leistungsabgabe wesentlich leichter und kleiner sein als Elektromotoren ([DD98]). Daneben begünstigen die umfangreichen Steuerungs-

¹ Elektrische Komponenten werden in der vorliegenden Arbeit nicht berücksichtigt. Im Fokus dieser Arbeit stehen die hydraulischen Komponenten.

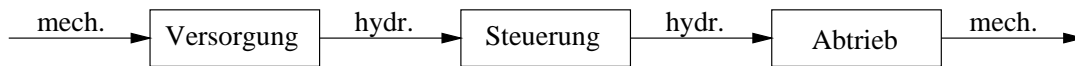


Abbildung 4.1: Komponentengruppen und Leistungstransformationen innerhalb einer hydraulischen Anlage

und Regelungsmöglichkeiten ein gutes Zeitverhalten sowie kontinuierliche Veränderungen der Übersetzungsverhältnisse auch unter Last.

Nachteilig ist jedoch insbesondere die Gefahr von unkontrollierten Leckagen, die neben ökonomischen Schäden auch Folgen für die Umwelt haben können, z. B. durch den Austritt von Hydrauliköl. Sowohl zur Vermeidung von inneren oder äußeren Leckagen als auch zur Beherrschung hoher Drücke und Kräfte ist bei der Herstellung hydraulischer Komponenten eine hohe Genauigkeit notwendig. Dieser erhöhte fertigungstechnische Aufwand schlägt sich wiederum in hohen Kosten nieder; zur Vermeidung von teuren Stillstandszeiten sind daher schnelle und genaue Methoden der Fehlererkennung notwendig.

Abbildung 4.2 zeigt im linken Teil den schematischen Aufbau einer kleinen Beispielanlage, die aus Antriebseinheit P, Steuerventil S, Hydrozylinder C, Auffangbehälter T sowie Leitungsverbindungen besteht. Die dargestellten Größenverhältnisse, Positionen und Drehungswinkel der Bauteile stimmen dabei nicht mit der Realität überein. Ein typischer Einsatzbereich für diese Anlage könnte die Hubvorrichtung in einer Hebebühne sein.

Verhalten einer Anlage

Das nach außen sichtbare Verhalten einer hydraulischen Anlage ist die Basis für die Erkennung von Fehlern. Im folgenden werden die wesentlichen Einflußfaktoren für das Anlagenverhalten besprochen, ohne näher auf die physikalischen Zusammenhänge einzugehen. Diese können der Fachliteratur für Physik (z. B. [HMS88]) entnommen werden.

Eine hydraulische Anlage befindet sich zu jedem Zeitpunkt in einem definierten *Betriebszustand*, dessen Dauer durch die Konstanz aller internen Zustände bestimmt ist. Hierfür wird im folgenden auch der Begriff *Phase* verwendet. Verändert sich mindestens ein interner Zustand, so geht das System in eine andere Phase über. Die Ursache für einen Phasenwechsel kann ein äußeres Ereignis (z. B. die manuelle oder elektrische Veränderung einer Ventilposition) oder ein systeminternes Ereignis (z. B. das Erreichen einer Kolbenanschlagposition) sein. Der zeitliche Ablauf aller Phasen wird

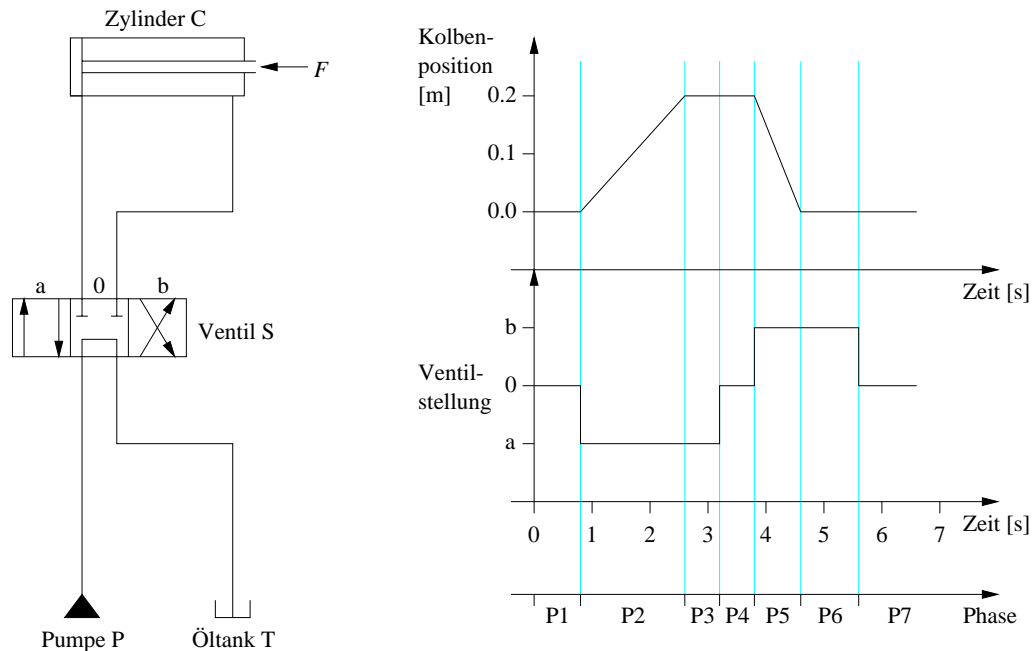


Abbildung 4.2: Schematischer Aufbau (links) und Phasen (rechts) einer einfachen hydraulischen Anlage

Phasensequenz genannt (vgl. [Hof99]).

Beispiel 4.1: Wenn keine Komponentenfehler vorliegen, lässt sich für die Beispielanlage in Abbildung 4.2 die folgende Phasensequenz identifizieren (rechter Teil der Abbildung):

- Phase 1: Ausgangszustand. Der Kolben des Zylinders C ist vollständig eingefahren. Das Steuerventil S ist in der Mittelstellung.
- Phase 2: Der Zylinderkolben fährt aus. Das Steuerventil befindet sich in der Stellung a.
- Phase 3: Der Zylinderkolben ist in der Endlage.
- Phase 4: Der Zylinderkolben bleibt in der Endlage. Das Steuerventil ist in der Mittelstellung.
- Phase 5: Der Zylinderkolben fährt ein. Das Steuerventil befindet sich in der Stellung b.
- Phase 6: Der Zylinderkolben hat wieder seine Ausgangsposition erreicht.

- Phase 7: Endzustand. Das Steuerventil ist zurück in der Mittelstellung. Dieser Betriebszustand entspricht dem Ausgangszustand (Phase 1), so daß eine neue Sequenz beginnen kann. \square

Die Phasensequenz beschreibt dabei nur das *prinzipielle* Verhalten einer hydraulischen Anlage zur Erfüllung einer definierten Aufgabe. Das *genaue zeitliche* Verhalten hängt von verschiedenen Einflußfaktoren ab, von denen die wichtigsten die Systemanforderungen und Komponentenfehler sind.

Innerhalb der Aufgabendefinition können an eine hydraulische Anlage unterschiedliche Anforderungen gestellt werden, die zu einer Veränderung des zeitlichen Verhaltens führen. Im Folgezyklus kann eine andere Belastung der Abtriebskomponenten bewirken, daß einzelne oder sogar alle Phasen in ihrer Länge variieren: Es ist zum Beispiel zu erwarten, daß das Ausfahren des Zylinders bei höherer Belastung länger dauert als bei kleinerer Belastung². Wird eine fehlerfreie Anlage innerhalb der konstruktionsbedingt vorgesehenen *Belastungsstufen* betrieben, bleibt aber der Phasenzyklus als solcher erhalten.

Das belastungsabhängige zeitliche Verhalten einer fehlerfreien Anlage wird *Soll-Verhalten* genannt. Werden Phasenausfälle, andere Phasenabfolgen oder signifikant abweichende Phasenlängen beobachtet, kann das auf mehrere Ursachen zurückzuführen sein: eine falsche Montage der Anlage, eine falsche Bedienung, der Betrieb außerhalb der Spezifikation bezüglich Umgebungstemperatur, Belastungen etc. sowie Defekte einzelner oder mehrerer Komponenten.

Üblicherweise wird bei der Störungssuche davon ausgegangen, daß die zu diagnostizierende Anlage korrekt konstruiert, montiert und bedient wurde. Unter diesen Voraussetzungen sind trotz regelmäßiger Wartung nach [DD98] verschleißbedingte Komponentenfehler die häufigste und wichtigste Störungsursache. Aus diesem Grund werden in der weiteren Arbeit alle anderen Ursachen ausgeklammert.

4.1.2 Physikalische Größen in der Hydraulik

Die genaue Ermittlung der Abweichung des Ist-Verhaltens vom Soll-Verhalten einer hydraulischen Anlage ist in der Regel nur durch die Messung bestimmter physikalischer Größen möglich. Zu diesem Thema folgt eine kurze Übersicht der diagnoserelevanten Aspekte.

2 In der Beispielanlage (Abbildung 4.2) wird die Belastung durch die von außen auf den Zylinder wirkende Kraft F dargestellt.

Größen und Einheiten

Die Bestimmung physikalischer Größen setzt die Meßbarkeit und Vergleichbarkeit der betreffenden Eigenschaften und Zustände voraus; hierfür ist ein geeignetes Meßverfahren und die Festlegung einer entsprechenden Einheit erforderlich. Aus diesem Grund umfaßt das internationale Einheitensystem (SI) sieben Basisgrößen, die durch Meß- bzw. Zählvorschriften oder Maßverkörperungen eindeutig festgelegt sind (siehe z. B. [ZS92]).

Für die Berechnung und Beschreibung hydraulischer Anlagen werden vier SI-Maßeinheiten benötigt, aus denen die Einheiten anderer wichtiger Größen abgeleitet werden können (Abbildung 4.3).

	phys. Größe	Zeichen	SI-Einheit	Zeichen
1.	Länge	l	Meter	m
2.	Masse	m	Kilogramm	kg
3.	Zeit	t	Sekunde	s
4.	Temperatur	T	Kelvin	K

	phys. Größe	Zeichen	Einheit	Zeichen	SI-Einheiten
	Druck	p	Pascal	Pa	$\text{kg} \cdot \text{m}^{-1} \cdot \text{s}^{-2}$
	Volumenstrom (Fluß)	q_V			$\text{m}^3 \cdot \text{s}^{-1}$
	Kraft	F	Newton	N	$\text{kg} \cdot \text{m} \cdot \text{s}^{-2}$
	Geschwindigkeit	v			$\text{m} \cdot \text{s}^{-1}$
	Leistung	P	Watt	W	$\text{kg} \cdot \text{m} \cdot \text{s}^{-3}$
	Umdrehungsfrequenz	n			s^{-1}

Abbildung 4.3: SI-Einheiten (oben) und abgeleitete Einheiten (unten) in der Hydraulik

Anmerkung: Die für die Diagnose wichtigsten physikalischen Größen sind Druck p und Fluß q_V . Aus Gründen der besseren Handhabbarkeit werden sie üblicherweise in SI-fremden Einheiten dargestellt. Dabei gilt:

- $[p] = \text{Bar}$, wobei $1 \text{ bar} = 10^5 \text{ Pa}$.
- $[q_V] = \text{Liter pro Minute}$, wobei $1 \text{ l} \cdot \text{min}^{-1} = 6 \cdot 10^{-2} \text{ m}^3 \cdot \text{s}^{-1}$.

Die physikalischen Zusammenhänge und die Berechnung der in Abbildung 4.3 genannten Größen kann der Fachliteratur (z. B. [HMS88]) entnommen werden. Im Rahmen dieser Arbeit ist lediglich wichtig, zwei Klassen von physikalischen Größen zu unterscheiden, weil sich hierdurch Konsequenzen für die Verrechnung während der Fehlerdiagnose ergeben:

- Die *skalaren Größen* Länge l , Masse m , Zeit t , Temperatur T , Druck p , Leistung P und Drehzahl n werden durch die Angabe ihres Zahlenwerts und ihrer Einheit charakterisiert. Zahlenwerte von Größen der gleichen Einheit dürfen beliebig verrechnet werden; im Rahmen der Diagnose ist insbesondere die Summen- bzw. Differenzbildung erlaubt.
- Dagegen werden die *vektoriellen Größen* Fluß q_V , Kraft F und Geschwindigkeit v zusätzlich durch eine Richtungsangabe charakterisiert. Hier dürfen die Zahlenwerte zweier Größen nur dann ohne Einschränkungen verrechnet werden, wenn beide Richtungsangaben übereinstimmen. Die Zahl der möglichen Richtungen kann sehr eingeschränkt sein, z. B. kommen für den Fluß in einer Rohrleitung nur zwei Richtungen in Betracht.

Messung hydraulischer Größen

Das *Messen* ist der experimentelle Vorgang zur Bestimmung des Wertes einer physikalischen Größe, die auch als *Meßgröße* bezeichnet wird. Der aktuelle *Meßwert* wird als Bruchteil oder Vielfaches einer physikalischen Einheit angegeben und durch ein *Meßgerät* ermittelt, das an einer bestimmten *Meßstelle* im Schaltkreis angebracht ist (siehe z. B. [DD98]). Ein Meßwert bezieht sich also immer auf eine physikalische Größe, einen Zeitpunkt und einen Ort.

Die Meßgeräte unterscheiden sich im eingesetzten Meßverfahren, ihrer Empfindlichkeit, ihrer Genauigkeit usw. Sie werden entweder an den Meßstellen fest eingebaut oder als transportable Handmeßgeräte bei Bedarf über bereits installierte Adapter an das System angeschlossen. Die erfaßten Meßwerte können direkt von einer Skala abgelesen werden oder, was besonders beim Einsatz von rechnergestützten Überwachungs- und Diagnosesystemen von Vorteil ist, automatisch gespeichert und ausgewertet werden.

Jede Messung hat das Ziel, den wahren Wert der jeweiligen Meßgröße zu erfassen. Dies ist jedoch stets mit einer Meßunsicherheit verbunden, so daß unter gleichen Bedingungen durchgeführte Wiederholungsmessungen voneinander abweichende Meßwerte ergeben können. Die Ursachen hierfür lassen sich in zwei Gruppen aufteilen (vgl. z. B. [HMS88]):

- Die *systematischen Abweichungen* sind abhängig vom Meßverfahren und können z. B. durch eine falsche Kalibrierung, konstruktionsbedingte Ablese- bzw. Quantisierungsfehler, falsche Montage oder Meßwertdriften (Veränderungen der Meßgröße durch die Messung) entstehen. Sie führen zu einer unsymmetrischen Häufung der Meßwerte von Wiederholungsmessungen.
- Die *zufälligen Abweichungen* sind abhängig vom Meßvorgang und werden von nicht erkennbaren Einflüssen hervorgerufen. Dabei kann es sich um menschliche Unzulänglichkeiten wie z. B. manuelle Ungeschicklichkeiten, falsche Schätzungen von Zwischenwerten auf der Meßskala usw. oder um Umwelteinflüsse handeln. Sie führen zu einer symmetrischen Häufung der Meßwerte um einen häufigsten Wert.

Die systematischen Abweichungen lassen sich nur durch die Änderung der Meßumstände, d. h. durch die Anwendung eines anderen Meßgeräts oder Meßverfahrens aufdecken und sind einer mathematischen Analyse nicht zugänglich. Zufällige Abweichungen hingegen können mit einer Fehleranalyse statistisch beschrieben und eingegrenzt werden. Das Ergebnis der Analyse ist eine Meßunsicherheit $u \in \mathbb{R}$, um die der exakte physikalische Wert x_0 vom Meßwert \bar{x} maximal abweicht, falls keine systematischen Fehler vorliegen; es gilt dann

$$x_0 \in [\bar{x} - u, \bar{x} + u].$$

Anmerkungen: (i) Trotz der unvermeidbaren Meßunsicherheit muß der gemessene Wert für den wahren Wert der Meßgröße repräsentativ sein, sonst ist das Meßverfahren ungeeignet (vgl. [NL94]). Während die generelle Eignung der eingesetzten Meßverfahren im weiteren vorausgesetzt wird, werden die Meßunsicherheiten durch die Verwendung von Meßwertintervallen statt scharfer Einzelwerte berücksichtigt (siehe Abschnitt 4.5.1).

(ii) Im Rahmen der Meßstellenauswahl (Abschnitt 4.6) werden vom hier vorgestellten Diagnosesystem ausschließlich Meßstellen betrachtet, an denen Druck- oder Flußmeßgeräte angeschlossen werden können; dies sind im allgemeinen die Leitungstücke zwischen zwei hydraulischen Komponenten. Im Simulationswerkzeug ^{art}*deco* werden hierfür die in Abbildung 4.4 dargestellten Symbole verwandt.

4.1.3 Aufgaben einer rechnergestützten Diagnose

Mit dem im Rahmen dieser Arbeit entwickelten automatischen Wissensakquisitionskonzept ist es möglich, zu einer hydraulischen Anlage ein anlagenspezifisches Diagnosesystem zu erzeugen. Dieses Softwaresystem soll einen für die Wartung der

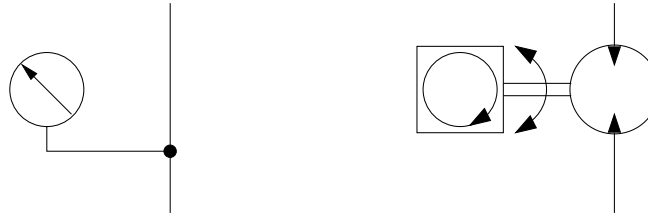


Abbildung 4.4: Symbole für Druckmeßgerät (links) und Flußmeßgerät (rechts)

Anlage zuständigen Hydraulikingenieur in einer realen Störungssituation schnell und effektiv bei der Fehlersuche unterstützen. Zur Verdeutlichung der Einsatzmöglichkeiten eines solchen Systems wird im folgenden ein Diagnoseszenario skizziert.

Tritt eine Störung auf, werden üblicherweise zunächst die Bauteile mit leicht erkennbaren Verschleißerscheinungen überprüft. Hierzu gehört die Sichtprüfung auf äußere Leckagen und Beschädigungen, Prüfung des Flüssigkeitsstands im Vorratsbehälter sowie der Sauberkeit der Filter usw. ([DD98]). Konnten diese Fehlerursachen ausgeschlossen werden, erfolgt der rechnergestützte Diagnoseprozeß.

Zunächst müssen dem Diagnosesystem die aktuellen Parameter der hydraulischen Anlage übermittelt werden. Hierzu gehören neben den Informationen über die Betriebssituation (Phase und Belastungen) insbesondere die an vorbestimmten Meßstellen gewonnenen Meßwerte. Abhängig von den technischen Voraussetzungen sind zwei Strategien der Datenübertragung möglich:

1. *Manuelle Datenübertragung:* Wenn keine Verbindung zwischen dem Diagnosesystem und der Anlage besteht, muß die Meßwerterhebung nach einem expliziten Plan für jede Meßstelle manuell durchgeführt und im System eingegeben werden. Ein Problem ist hierbei die Gefahr von Ablese- und Eingabefehlern.
2. *Automatische Datenübertragung:* Hier existiert eine Verbindung zwischen dem Diagnosesystem und der Anlage, etwa deshalb, weil die Anlage bereits rechnergestützt gesteuert und überwacht wird. Fest installierte Meßgeräte übermitteln kontinuierlich die aktuellen Meßwerte.

Auf der Basis der gesammelten Informationen ermittelt das Diagnosesystem in kurzer Zeit³ eine oder mehrere Komponenten der Anlage, die am wahrscheinlichsten als defekt angesehen werden können. Diese Komponenten werden unter Angabe der

3 Unter »kurzer Zeit« sollten wenige Sekunden verstanden werden.

berechneten Wahrscheinlichkeitswerte genannt. Der Ingenieur erhält damit eine Eingrenzung der Fehlermöglichkeiten und überprüft nur diese Komponente bzw. Komponenten, um den realen Defekt herauszufinden. Durch das Diagnosesystem wird also einerseits die Menge der in Betracht kommenden defekten Komponenten verringert und zum anderen eine Reihenfolge zur Untersuchung der Komponenten vorgegeben.

Durch eine geeignete Rechnerunterstützung des Diagnoseprozesses können die Stillstandszeiten einer hydraulischen Anlage entscheidend verkürzt werden, so daß ein hoher wirtschaftlicher Nutzen zu erwarten ist. Gerade im Hinblick auf sehr komplexe und damit störungsanfällige Anlagen wirkt sich dieser Vorteil besonders aus.

4.2 Gesamtkonzept des Diagnoseansatzes

Im folgenden wird ein Gesamtkonzept zur Erzeugung von heuristischen Diagnosesystemen für hydraulische Anlagen im Überblick vorgestellt. Es integriert und konkretisiert den in Abschnitt 2.3.3 eingeführten automatischen Wissensakquisitionsansatz und enthält Problemlösungen zur Wissensanwendung, wie zum Beispiel die Auswertung der Regelmenge im Störfall (Hypothesengenerierung). Eine prototypische Realisierung des Konzepts wurde als Diagnosesystemgenerator ARGUS vorgenommen (siehe Kapitel 5).

Zunächst werden in Abschnitt 4.2.1 die notwendigen Voraussetzungen für die Realisierung des Konzepts besprochen, bevor Abschnitt 4.2.2 die wesentlichen Einzelschritte sowie ihr Zusammenwirken darstellt. Anschließend wird in Abschnitt 4.2.3 auf den wichtigen Aspekt der Repräsentation des Diagnosewissens durch kategoriale Assoziationsregeln eingegangen und in Abschnitt 4.2.4 die mit dem Konzept verfolgte Zielsetzung definiert.

4.2.1 Voraussetzungen

Der Einsatz des in dieser Arbeit vorgeschlagenen Diagnoseansatzes ist an einige Voraussetzungen geknüpft. Sie betreffen sowohl die Wissensakquisitionsphase (*Lernphase*) als auch die Wissensanwendungsphase (*Diagnosephase*) des Diagnosesystems.

Für die Lernphase müssen geeignete Verhaltens- und Strukturmodelle aller Komponenten der zu diagnostizierenden Anlage vorhanden sein. Die Modelle müssen dabei das *korrekte* Verhalten als auch das im Sinne der Anlagenspezifikation *fehlerhafte* Verhalten jeder Komponente im Kontext der Gesamtanlage abbilden. Des weiteren

müssen die typischen Betriebsszenarien, z. B. die für die Anlage vorgesehenen Belastungsstufen, bekannt sein. Zur Auswahl der besten Meßstellen im Schaltkreis werden die Kosten, Wertebereiche und Ablesegenauigkeiten der potentiellen Meßstellen benötigt.

Für die Diagnosephase wird angenommen, daß die technischen Installationen zur Übertragung der Meßdaten im Störfall vorhanden sind. Des weiteren wird davon ausgegangen, daß jede Störung nur auf genau eine fehlerhafte Komponente zurückzuführen ist, daß also keine Mehrfachfehler vorliegen. Diese Annahme ist jedoch keine zwingende Voraussetzung zur Anwendung des Diagnoseansatzes, sie wurde lediglich getroffen, um den experimentellen Aufwand zu begrenzen. Des weiteren werden Bedienungs-, Konstruktions- oder Montagefehler grundsätzlich ausgeschlossen (siehe auch Abschnitt 4.1.1).

4.2.2 Einzelschritte

Das Gesamtkonzept zur heuristischen Diagnose von hydraulischen Anlagen läßt sich in zwei Bereiche einteilen, in deren Schnittpunkt sich eine Assoziationsregelbasis mit Symptom→Fehler-Beziehungen befindet (vgl. auch Abbildung 1.1 auf Seite 3). Der erste Aufgabenbereich umfaßt den Aufbau der Regelbasis und besteht aus einer *Simulations-* und einer *Regellernphase*. Diese Phasen werden für jede zu diagnostizierende Anlage einmal durchlaufen; sind Parameteränderungen zur Optimierung der Regelbasis notwendig, auch mehrfach. Der zweite Aufgabenbereich ist in eine *Meß-* und eine *Diagnosephase* unterteilt. Ausgehend von den gemessenen Werten wird für jeden Störfall der realen technischen Anlage die Regelbasis zur Ableitung des wahrscheinlichsten Fehlers (Diagnose) genutzt.

Abbildung 4.5 skizziert den Aufbau des Diagnosesystems. Dabei werden Parallelen zu einem KDD-System erkennbar (vgl. Abschnitt 2.2.3): Das Preprocessing besteht aus der Symptomerkennung, Diskretisierung und Meßstellenauswahl, das Data Mining ist die Regelerzeugung, und die Hypothesengenerierung ist das Postprocessing bzw. die Nutzbarmachung des Diagnosewissens. Im folgenden werden die Einzelschritte kurz beschrieben.

Simulationsphase: Die Simulationsphase dient dazu, eine Simulationsdatenbasis aufzubauen, die das Verhalten der zu diagnostizierenden hydraulischen Anlage durch eine Menge von Meßwertevektoren beschreibt. In jedem Vektor werden für eine spezifische Betriebssituation (Phase und Belastungsstufen) die simulierten Werte an den vorher definierten Meßstellen repräsentiert. Da die Protokollierung der Meßwerte nach Anlagenphasen getrennt geschieht, wird die Simulationsdatenbasis im weiteren auch *Phasenprotokollmodell* genannt. Das Phasenprotokollmodell enthält Simulationsdaten sowohl für fehlerfreie Komponenten als auch für fehlerhafte Komponenten.

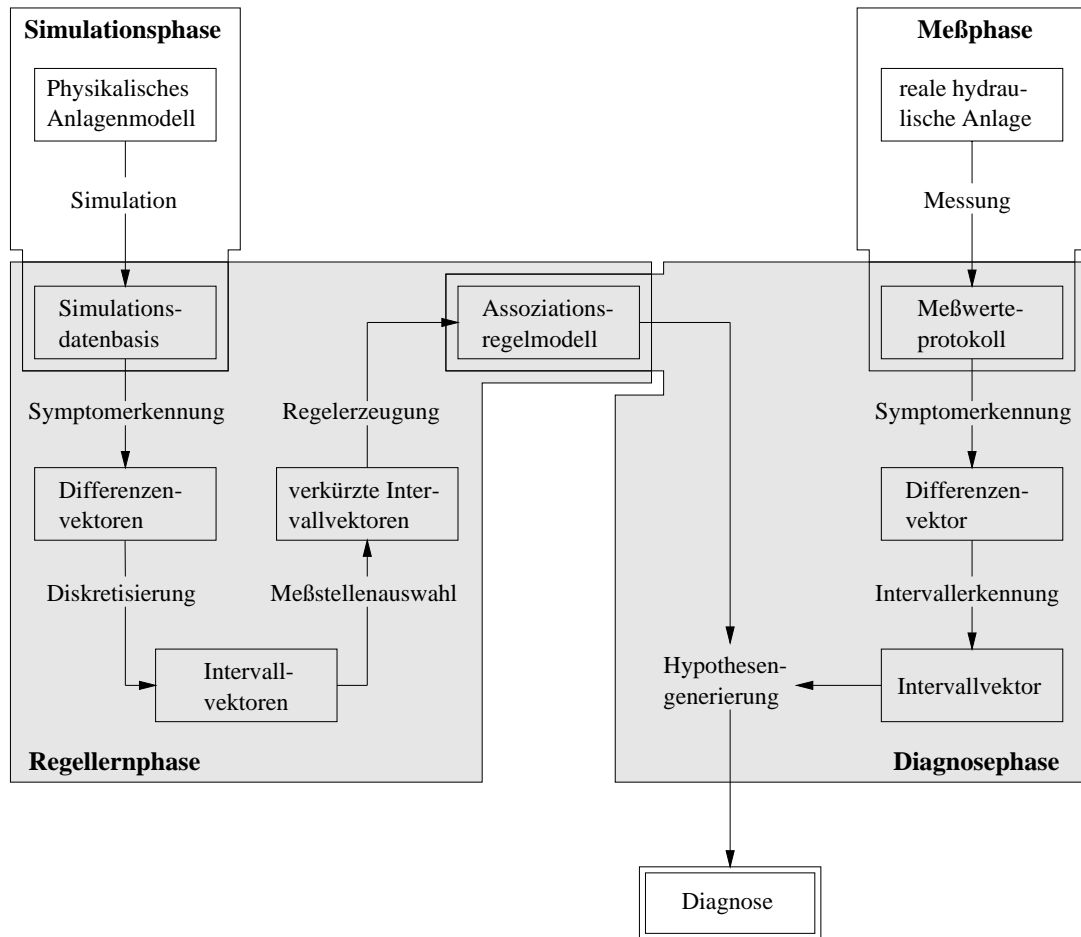


Abbildung 4.5: Einzelschritte des Gesamtkonzepts zur Fehlerdiagnose hydraulischer Anlagen

Näheres wird in Abschnitt 4.3 beschrieben.

Bei der Realisierung des prototypischen Diagnosesystemgenerators ARGUS ist die Simulation von den anderen Programmteilen entkoppelt und wird mit dem separaten Simulationswerkzeug *artdeco* durchgeführt. Hierbei handelt es sich um ein Programmsystem zur graphischen Schaltplanerstellung und Simulation fluidischer Systeme; es wird z. B. in [SL92], [KCH⁺95], [Cur96] und [SCH98] beschrieben. *artdeco* enthält bereits alle relevanten Verhaltensmodelle für die hydraulische Domäne, so daß es durch eine Analyse der Entwurfszeichnung und die Synthese der lokalen Komponentenmodelle ein globales Anlagenmodell ableiten und das Systemverhalten simulieren kann.

Regellernphase: Diese Phase umfaßt neben der Anwendung des eigentlichen Lernalgorithmus alle Vorverarbeitungsschritte, die zur Erzeugung geeigneter Diagnoseregeln notwendig sind. Des weiteren werden Meßstellenvorschläge zur Plazierung derjenigen Meßgeräte gemacht, deren Auswertung die höchste Diagnosegüte verspricht.

Zunächst wird das Phasenprotokollmodell analysiert, um für jede Phase der Anlage Symptome als Differenzen zwischen simulierten Soll- und Ist-Meßwerten zu identifizieren (siehe Abschnitt 4.4). Das Ergebnis ist eine aus Differenzenvektoren bestehende Datenbasis. Da das eingesetzte Assoziationsregellernverfahren nur diskrete Eingaben verarbeiten kann, werden anschließend die reellwertigen Differenzen in Intervalle eingeteilt und eine Datenbasis mit Intervallvektoren erzeugt (siehe Abschnitt 4.5). Eine Analyse der Diskriminierungsfähigkeit der simulierten Meßstellen führt unter Berücksichtigung von benutzerdefinierten Randbedingungen zu einer optimalen Meßstellenauswahl (siehe Abschnitt 4.6). Es wird davon ausgegangen, daß in der realen Anlage auch nur an diesen Stellen die Meßwerte erhoben werden können; aus diesem Grund können aus den Intervallvektoren alle anderen Werte entfernt werden, und bei der Regelerzeugung (siehe Abschnitt 4.7) wird nur auf Informationen zugegriffen, die in einem realen Störfall meßbar sind. Das Ergebnis des letzten Schritts sind Assoziationsregelmengen, die nach Anlagenphasen getrennt als Diagnoseregeln ausgewertet werden können.

Meßphase: Liegt ein Störfall in der realen Anlage vor, werden die aktuellen Meßwerte an den bei der Meßstellenauswahl definierten Stellen erhoben und als Meßwerteprotokoll zum Diagnosesystem geleitet. Dieser Vorgang wird im weiteren nicht näher betrachtet, da er von den technischen Voraussetzungen der zu diagnostizierenden Anlage abhängt.

Diagnosephase: Die Diagnosephase schließt sich an die Meßphase an und generiert Diagnosen durch die Auswertung des Meßwerteprotokolls und die Anwendung der Diagnoseregeln. Das Meßwerteprotokoll wird (analog zu den simulierten Meßwertvektoren) durch den Vergleich mit den Soll-Meßwerten in Differenzenvektoren und durch Diskretisierung in Intervallvektoren transformiert. Bei der Hypothesengenerierung (siehe Abschnitt 4.8) werden die anwendbaren Assoziationsregeln ermittelt und durch eine Verrechnung der Regelkonfidenzen die Wahrscheinlichkeiten für jede Fehlerklasse bestimmt. Die Fehlerklassen mit den höchsten Bewertungen stellen schließlich die Diagnose dar.

4.2.3 Repräsentation der Diagnoseregeln

Bei den im vorgestellten Ansatz verarbeiteten Diagnoseregeln handelt es sich um eine Kombination aus kategorialen Assoziationsregeln und Klassifikationsregeln (vgl. Abbildung 3.2 auf Seite 40). Die Regelprämissen bestehen aus durch Konjunktion ver-

knüpften Meßgerät / Intervall-Vergleichen $m = i$, die jeweils erfüllt sind, wenn das am Meßgerät $m \in M$ beobachtete Symptom in das Intervall $i \in I_m$ fällt. Als Konklusion ist nur jeweils eine Diagnose $f \in F$ erlaubt. Gemäß Definition 4.1 gebildete Diagnoseregeln stellen somit heuristische Symptom \rightarrow Diagnose-Zusammenhänge dar.

Definition 4.1 (Diagnoseregeln): Es sei $M = \{m_1, \dots, m_g\}$ die Menge der $g \in \mathbb{N}$ Meßstellen mit den zugehörigen Intervallmengen I_{m_1}, \dots, I_{m_g} . Weiter sei F die Menge der Fehler. Dann hat eine *Diagnoseregeln* r mit $p \leq g$ Prämissenelementen die Gestalt

$$r: m_{a_1} = i_{b_1} \wedge \dots \wedge m_{a_p} = i_{b_p} \rightarrow f \quad | \quad c, s$$

wobei $a_1, \dots, a_p \in \{1, \dots, g\}$, $i_{b_j} \in I_{m_{a_j}}$ und $f \in F$. Die Regel r wird charakterisiert durch die Konfidenz c und die relative Häufigkeit s . \square

Beispiel 4.2: In der weiteren Arbeit werden die Intervalle mit i_1, i_2, i_3, \dots und die Fehler mit Großbuchstaben A, B, C, ... oder f_1, f_2, f_3, \dots bezeichnet. Damit könnte eine typische Diagnoseregeln lauten:

$$r: m_1 = i_3 \wedge m_4 = i_4 \rightarrow E \quad | \quad c(r) = 0.8, s(r) = 0.05.$$

Mit den entsprechenden Symbolzuordnungen (Intervall i_3 für m_1 entspricht [10.0, 12.5] bar, Intervall i_4 für m_4 entspricht [-3.0, -1.0] l/min, E ist das Drosselventil) läßt sich r interpretieren als:

»Wenn die Druckdifferenz zum Normalwert an Manometer m_1 zwischen 10.0 und 12.5 bar und die Flußdifferenz zum Normalwert an Flußmesser m_4 zwischen -3.0 und -1.0 l/min beträgt, dann liegt mit der Sicherheit 80% ein Fehler im Drosselventil vor. Diese Situation tritt in 5% aller Fälle auf.« \square

Anmerkungen: (i) Mit geeigneten Bezeichnungen für die Intervalle, aus denen der Bezug zum jeweiligen Meßgerät hervorgeht, können die Diagnoseregeln auch effizienter als boolesche Assoziationsregeln repräsentiert werden (vgl. Abschnitt 3.2). Hierauf wird aber aus Gründen der Übersichtlichkeit im weiteren verzichtet.

(ii) Die Prämissen können auch Bedingungen bezüglich der aktuellen Betriebssituation enthalten, d. h., die Elemente m_i bezeichnen neben Meßgeräten auch Kräfte, die an den Zylindern anliegen. Das zugehörige Intervall repräsentiert in diesen Fällen eine Belastungsstufe. Eine Angabe der Anlagenphase ist dagegen nicht erforderlich, weil davon ausgegangen wird, daß die Zuordnung einer Regel zu der entsprechenden Phase jederzeit eindeutig möglich ist.

(iii) Die Beschränkung der Konklusion auf *eine* Diagnose $f \in F$ steht *nicht* im Zusammenhang mit der in Abschnitt 4.2.1 beschriebenen Einzelfehlerannahme, denn jede Beziehung der Form $s \rightarrow f_1, \dots, f_n$ läßt sich durch n Regeln $s \rightarrow f_1, \dots, s \rightarrow f_n$ ausdrücken.

4.2.4 Zielsetzung

Der in dieser Arbeit vorgestellte Ansatz ist ein *allgemeines* Konzept zur automatischen Generierung von Diagnosesystemen. Zu einem generierten Diagnosesystem wird es immer manuell und anlagenspezifisch optimierte Systeme mit höheren Diagnoseleistungen geben. Diese sind aber nur mit einem sehr viel höheren zeitlichen und personellen Aufwand zu erstellen. Folglich kann es nicht das Ziel des hier beschriebenen Ansatzes sein, *alle denkbaren* Komponentenfehler einer hydraulischen Anlage zu entdecken. Um den Hydraulikingenieur sinnvoll zu unterstützen, ist vielmehr die *zuverlässige* Erkennung derjenigen Fehler wichtig, die von Experten als bedeutend (d. h. als gravierend und nicht zu selten) eingestuft werden. Solche Fehler treten in der Regel an den Hauptkomponenten der Anlage auf.

Aus diesem Grund ist neben einer effizienten Verarbeitungsmöglichkeit angestrebt, eine möglichst hohe *Diagnoseleistung* für die als bedeutend eingestuften Störungursachen zu bieten: Es sollen also möglichst viele dieser Fehler erkannt und möglichst wenig Fehler falsch diagnostiziert werden. Die konkrete Bewertung der mit einem generierten Diagnosesystem erzielbaren Diagnosegüte wird in Abschnitt 5.2 erläutert.

4.3 Simulation

Im Simulationsschritt wird das fehlerfreie und das fehlerhafte Verhalten einer hydraulischen Anlage in repräsentativen Betriebssituationen ermittelt, indem durch die numerische Simulation der Verhaltens- und Strukturmodelle ein *Phasenprotokoll* erzeugt wird. Bei dem Phasenprotokoll handelt es sich um ein Verhaltensprofil der Anlage, in der für jede berücksichtigte Betriebssituation und jede Phase die simulierten Meßwerte tabellarisch als Meßwertektoren abgelegt sind.

Vor der eigentlichen Simulation sind die beiden Dimensionen des Phasenprotokolls festzulegen: Zunächst muß ein Meßplan erstellt werden, in dem bestimmt wird, welche physikalischen Größen während der Simulation aufzuzeichnen sind. Des weiteren erfolgt die Definition eines Simulationsplans, in dem die zu simulierenden Fehlervariationen und Belastungsstufen der Anlage beschrieben werden. Diese Aufgaben werden in den Abschnitten 4.3.1 und 4.3.2 näher erläutert.

4.3.1 Erstellung eines Meßplans

Ein *Meßplan* definiert die Spalten des Phasenprotokolls; er legt für eine hydraulische Anlage fest, welche simulierten Meßgrößen zu protokollieren sind. Zusätzlich

werden Informationen zu Anlagenphase, Belastungsstufen und Komponentenfehler aufgenommen, da die Meßwerte hiervon abhängig sind (vgl. Abschnitt 2.3.1). Damit besteht eine Zeile des Phasenprotokolls aus den Meßwerten und der zugehörigen Betriebssituation.

Für die Meßwerte kommen im Rahmen der Diagnose hydraulischer Anlagen lediglich die folgenden Meßstellen und Meßgrößen in Betracht (elektrische Komponenten werden in dieser Arbeit nicht berücksichtigt):

- An jedem Zylinder die Kolbenposition s ,
- an jedem Hydromotor die Drehzahl n ,
- an jedem Leitungstück zwischen zwei Komponenten der Druck p und
- an jedem Leitungstück zwischen zwei Komponenten der Volumenstrom q_V .

Aus technischen Gründen kann es vorkommen, daß nicht alle Leitungstücke zwischen zwei Komponenten als Druck- und Flußmeßstellen zulässig sind. Dies ist z. B. dann der Fall, wenn hier eine Messung nur mit unverhältnismäßig hohem Aufwand oder wegen Unzugänglichkeit überhaupt nicht möglich ist. Randbedingungen dieser Art sind beim Aufstellen des Meßplans zu berücksichtigen. Abbildung 4.6 zeigt eine hydraulische Anlage mit ausgewählten Fluß- und Druckmeßstellen.

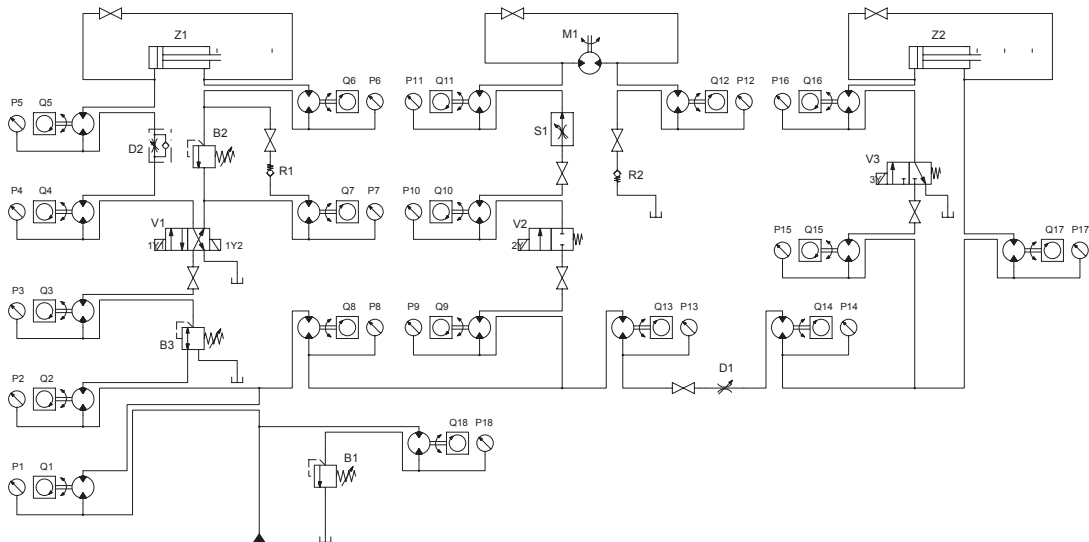


Abbildung 4.6: Schaltplan einer hydraulischen Anlage und ihrer Meßstellen laut Meßplan, dargestellt im Simulationswerkzeug *artdeco*

Die zu einem Meßwertvektor gehörige Betriebssituation wird durch die folgenden Werte gekennzeichnet:

- Ein Zeitstempel t ,
- an jedem Zylinder die belastende Kraft F und
- an einer Anlagenkomponente der Fehler f und seine Variation $v \in v_f$ (siehe Simulationsplan, Abschnitt 4.3.2).

Da die Struktur der jeweils zu diagnostizierenden Anlage in *art deco* graphisch vorliegt, ist es grundsätzlich möglich, den anlagenspezifischen Meßplan automatisch zu erstellen. Dieses ist jedoch noch nicht realisiert, so daß im prototypischen Diagnosesystem eine manuelle Festlegung der Meßpläne notwendig ist. Dazu werden mit Mausoperationen Druck- und Flußmeßgeräte an den gewünschten Stellen des Schaltplans eingefügt und die zu protokollierenden Meßgrößen festgelegt.

Anmerkungen: (i) Da für die Diagnose keine absoluten Zeitpunkte, sondern nur die Phasen einer hydraulischen Anlage relevant sind, wird nach der Simulation eine *Phasenerkennung* durchgeführt. Diese Information kann von den Meßgeräten nicht geliefert werden und wird in einem Nachbearbeitungsschritt durch die Analyse der Simulationsdatenbasis gewonnen. Die Phasenerkennung stellt darüber hinaus durch Einfügungen sicher, daß trotz fehlerverursachter Phasenausfälle jede Phasensequenz die gleiche Länge hat. Nur so ist ein phasenweiser Abgleich der Meßwerte zur Symptomerkennung möglich (vgl. Abschnitt 4.4).

(ii) Als weiterer Nachbearbeitungsschritt werden aus den Zeitpunkten t und den Zylinderpositionen s die Zylindergeschwindigkeiten $v = \frac{\Delta s}{\Delta t}$ berechnet. Diese sind auch an einer realen Anlage gut zu beobachten und können für einige Fehlerarten sehr charakteristisch sein.

4.3.2 Erstellung eines Simulationsplans

Im *Simulationsplan* werden die Zeilen des Phasenprotokolls festgelegt. Zunächst sind für die zu diagnostizierende Anlage die typischen Belastungsstufen und Fehlervariationen zu ermitteln. Sofern keine anlagen- oder anwendungsbezogenen Gründe gegen die Gleichverteilung aller Belastungsstufen und Fehlervariationen sprechen, muß mit jeder möglichen Kombination eine Simulation durchgeführt werden. Je mehr Zeit für die Simulationsphase zur Verfügung steht, desto mehr Betriebssituationen können berücksichtigt werden und desto vollständiger ist das im Phasenprotokoll abgelegte Verhaltensprofil der Anlage.

Prinzipiell ist auch die Erstellung eines Simulationsplans automatisierbar, sie wird aber im prototypischen Diagnosesystem noch manuell durchgeführt. Im folgenden werden die von einem Experten zu treffenden Entscheidungen zur Aufstellung eines Simulationsplans und der sich daraus ergebende Simulationsaufwand genauer beschrieben.

Die an einem Zylinder anliegende Kraft kann je nach Betriebssituation variieren. Sie ist eine kontinuierliche Meßgröße, deren Wertebereich anwendungsabhängig in geeignete Intervalle (Belastungsstufen) zu zerlegen ist. Ist mehr als ein Zylinder vorhanden, müssen typische Belastungsstufenkombinationen festgelegt werden. Eine weitere Festlegung ist bezüglich der Fehler notwendig: Hier sind *Fehlerklassen* und *Fehlervariationen* zu unterscheiden. Potentiell kann jede Anlagenkomponente defekt sein und damit eine Fehlerklasse f darstellen. Jede Fehlerklasse wiederum kann Fehler verschiedener Schweregrade umfassen, die zu f die Fehlervariationen v_f bilden.

Es ist offensichtlich, daß sowohl das korrekte Verhalten der Anlage als auch alle Fehlervariationen mit jeder Belastungsstufenkombination simuliert werden müssen. Damit läßt sich die Anzahl der Simulationsläufe und die Größe des Phasenprotokolls wie folgt berechnen. Sei F die Menge der Fehlerklassen (Komponenten) einer Anlage mit p Phasen, v_f die Menge der Variationen eines Fehlers $f \in F$ und b_i die Anzahl der Belastungsstufen des i -ten von $c \in \mathbb{N}$ Zylindern. Unter der Annahme der Unabhängigkeit müssen das Sollverhalten sowie alle Fehlervariationen in jeder Belastungsstufenkombination simuliert werden; es gilt somit für den maximalen Simulationsaufwand S und die Anzahl C der Meßwertvektoren im Phasenprotokoll:

$$S = \left(1 + \sum_{f \in F} |v_f| \right) \cdot \prod_{i=1}^c b_i \quad \text{und} \quad C = p \cdot S.$$

Anmerkungen: (i) Sind die Belastungsstufen voneinander abhängig, kann der Simulationsaufwand deutlich verringert werden. Ein Zahlenbeispiel für eine konkrete hydraulische Anlage befindet sich in Abschnitt 5.1.

(ii) Zur Evaluierung einer Regelmenge und ihrer Diagnoseleistung kann es sinnvoll sein, auf mehrere Simulationsdatenbasen mit unterschiedlichen Fehlervariationen zuzugreifen (siehe auch Abschnitt 5.2). Eine mögliche Charakterisierung dieser Variationismengen als leichte und schwere Variationen zeigt das folgende Beispiel 4.3.

Beispiel 4.3: Im fehlerfreien Fall betrage der Öffnungsgrad eines Drosselventils 50%. Die Fehlerklasse f ist dann »Defekt am Drosselventil« und gibt an, daß der aktuelle Öffnungsgrad vom Sollwert (50%) abweicht. Eine konkrete Fehlervariation kann numerisch z.B. einen Wert aus der Menge $v_f = \{0\%, 10\%, \dots, 40\%, 60\%, \dots, 100\%\}$ annehmen, die Höhe der Abweichung bestimmt ihren Schweregrad. Zur verbalen Beschreibung unterschiedlicher Schweregrade bietet sich die Zuordnung

von linguistischen Variablen, wie sie aus dem Bereich der *Fuzzy Logic* bekannt sind (vgl. z. B. [Zim93]), an. Abbildung 4.7 zeigt eine mögliche Gruppierung dieser Variationen durch unscharfe Definitionen der linguistischen Variablen »einfache« und »schwere Fehlervariation«. □

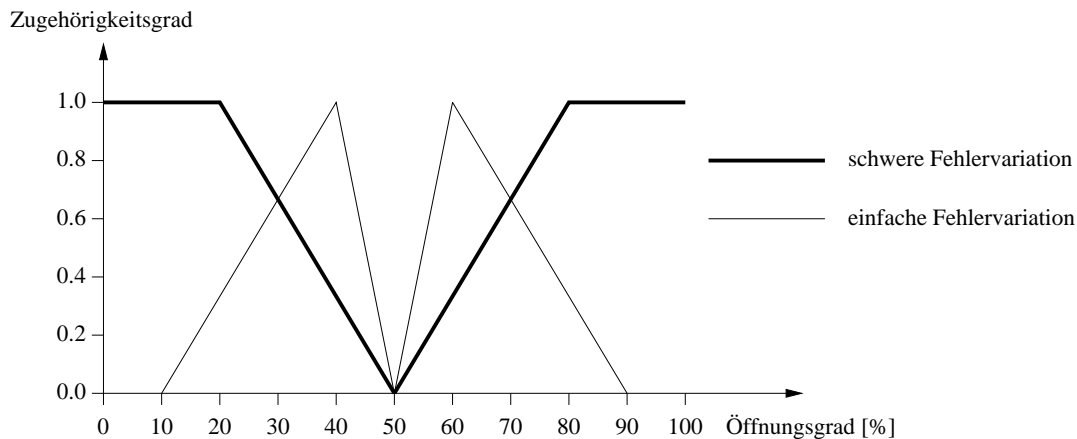


Abbildung 4.7: Definition der Fehlervariationen eines Drosselventils (zu Beispiel 4.3)

4.4 Symptomerkennung

Auf der Basis des Phasenprotokolls werden die Symptome der im Simulationsplan definierten Fehler identifiziert. Dies geschieht durch den Vergleich der Meßwertvektoren einer fehlerhaften Anlage mit den entsprechenden Vektoren der fehlerfreien Anlage. Weichen die Meßwerte voneinander ab, stellen ihre Differenzen die Symptome des Fehlers dar; eine Differenz von 0 entspricht dem Soll-Verhalten. Die so für alle Fehler entstandenen *Differenzenvektoren* oder *Symptomvektoren* werden zur weiteren Verarbeitung in einer Datenbasis abgelegt.

Es ist zu beachten, daß eine Differenzenbildung für Meßwerte von vektoriellen Größen der Hydraulik (Fluß q_V und Geschwindigkeit v) nicht numerisch erfolgen darf, wenn die Richtungen beider Werte (repräsentiert durch ihre Vorzeichen) ungleich sind. Richtungsunterschiede sind aber für einige Fehlerklassen sehr charakteristisch und im Gegensatz zu kleinen numerischen Abweichungen sogar ohne Meßgeräte während des Anlagenbetriebs leicht zu bemerken. Daher wird die Art der Abweichung im Differenzenvektor für einige Meßgrößen (z. B. Geschwindigkeiten) durch

einen symbolischen Wert beschrieben, was zu der folgenden Definition führt.

Definition 4.2 (Differenz zweier Meßwertvektoren): Es sei $M = \{m_1, \dots, m_s\}$ die Menge der Meßgeräte. Weiter seien $\vec{o}_1 = \langle a_{1,1}, \dots, a_{1,s} \rangle$ und $\vec{o}_2 = \langle a_{2,1}, \dots, a_{2,s} \rangle$ zwei Meßwertvektoren. Dann ist der Differenzenvektor \vec{o}_Δ definiert durch

$$\vec{o}_\Delta(\vec{o}_1, \vec{o}_2) = \langle a_{1,1} \ominus a_{2,1}, \dots, a_{1,s} \ominus a_{2,s} \rangle,$$

wobei für den Differenzoperator $\ominus : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} \cup \{0/+, 0/-, +/0, +/-, -/0, -/+\}$ gilt:

$$a_{1,i} \ominus a_{2,i} \mapsto \begin{cases} a_{1,i} - a_{2,i}, & \text{falls } m_i \text{ skalar oder } \text{sign}(a_{1,i}) = \text{sign}(a_{2,i}) \\ 0/+, & \text{falls } m_i \text{ vektoriell und } a_{1,i} = 0, a_{2,i} > 0 \\ 0/-, & \text{falls } m_i \text{ vektoriell und } a_{1,i} = 0, a_{2,i} < 0 \\ +/0, & \text{falls } m_i \text{ vektoriell und } a_{1,i} > 0, a_{2,i} = 0 \\ +/-, & \text{falls } m_i \text{ vektoriell und } a_{1,i} > 0, a_{2,i} < 0 \\ -/0, & \text{falls } m_i \text{ vektoriell und } a_{1,i} < 0, a_{2,i} = 0 \\ -/+, & \text{falls } m_i \text{ vektoriell und } a_{1,i} < 0, a_{2,i} > 0 \end{cases} \quad \square$$

Die Symptomerkenkung erfolgt für alle im Phasenprotokoll simulierten Fehler und Betriebssituationen. Hierbei wird zu einem Meßwertvektor \vec{o}_f für das Verhalten bei Vorliegen eines Fehlers f der zugehörige Meßwertvektor \vec{o}_{ok} für das Soll-Verhalten ermittelt und der Differenzenvektor $\vec{o}_\Delta(\vec{o}_f, \vec{o}_{ok})$ nach Definition 4.2 berechnet.⁴ Somit bedeutet ein positiver Symptomwert im Differenzenvektor \vec{o}_Δ , daß die entsprechende physikalische Größe im Fehlerfall zu hoch ist, während ein negativer Symptomwert eine zu niedrige physikalische Größe anzeigt.

Anmerkung: Die Einteilung und Protokollierung des Anlagenverhaltens in Phasen ermöglicht die Beschreibung von *zeitlich verteilten Symptomen* (vgl. [PR93]). Einfachere Diagnosesysteme basieren auf der Annahme, daß im Störfall der Zustand des hydraulischen Systems zu einem einzigen Zeitpunkt erfaßt werden kann und die Diagnose aus dieser statischen Beschreibung des Fehlverhaltens ableitbar ist. Im Gegensatz hierzu repräsentieren die Differenzenvektoren im hier vorgestellten Diagnoseansatz auch zeitliche Auswirkungen eines Fehlers, da diese phasengetreunt berechnet werden. Hierdurch sind Veränderungen der Symptome sichtbar, die insbesondere bei komplexen Anlagen dynamisch im Zeitablauf entstehen können.

⁴ Als »zugehörig« gilt derjenige Soll-Meßwertvektor, der durch die Simulation der gleichen Phase und Belastungsstufenkombination gewonnen wurde.

4.5 Diskretisierung

In der Diskretisierungsphase werden die Wertebereiche für die bei der Symptomererkennung berechneten Meßwertdifferenzen in diskrete Intervalle eingeteilt. Das Ergebnis ist also eine Datenbasis mit Intervallvektoren.

Zunächst werden in Abschnitt 4.5.1 die Bedingungen einer im Hinblick auf die Diagnoseanwendung sinnvollen Intervalleinteilung motiviert und definiert. Danach erfolgt in Abschnitt 4.5.2 die Diskussion der wichtigsten aus der Literatur bekannten Diskretisierungsverfahren. Abschließend wird in Abschnitt 4.5.3 ein neuer, im prototypischen Diagnosesystem realisierter Algorithmus vorgestellt und diskutiert.

4.5.1 Motivation und Definition

Zur Generierung von Diagnoseregeln wird eine Datenbasis mit Symptomvektoren ausgewertet. Sieht man von den symbolischen Darstellungen für vektorielle Meßgrößen ab, werden Symptome durch reellwertige Differenzzahlen repräsentiert (somit auch in den Prämissen der Diagnoseregeln). Aus den folgenden Gründen ist jedoch die Verwendung von diskreten Intervallen an Stelle von reellwertigen Symptomen in den Regelprämissen besser geeignet:

1. Anwendbarkeit des Regellernverfahrens:
Viele maschinelle Lernverfahren – wie auch die Assoziationsregelerzeugung – benötigen zwingend symbolische Eingaben oder sind hiermit zumindest effizienter (siehe [DKS95]).
2. Einsatzspektrum der Diagnoseregeln:
Die Diagnoseregeln sind universeller anwendbar und damit besser zur Diagnose neuer Betriebsszenarien geeignet, wenn die Prämissen keine einzelnen Symptomwerte enthalten.
3. Ungenauigkeit der Symptomwerte:
Es ist davon auszugehen, daß sowohl die gemessenen als auch die simulierten Meßwerte Abweichungen von den wahren physikalischen Werten aufweisen.

Der dritte Punkt sei noch etwas weiter ausgeführt: Da die Datengrundlage des Diagnosesystems aus Werten besteht, die aus simulierten und realen Messungen hervorgegangen sind, müssen diese nicht mit den exakten physikalischen Werten übereinstimmen. So können die in der Simulationsphase benutzten Verhaltensmodelle Modellierungsfehler enthalten oder eine zu geringe Modellierungstiefe aufweisen; als Konsequenz sind Abweichungen zwischen Simulationswerten und realen Werten

möglich.⁵ Auch in der Meßphase sind Abweichungen zu den realen Werten möglich, da physikalische Meßvorgänge i. d. R. mit systematischen und zufälligen Meßfehlern behaftet sind (vgl. Abschnitt 4.1.2). Die scharfen Differenzwerte sind daher eher als Näherungswerte anzusehen, die eine nicht vorhandene Exaktheit vortäuschen.

Zur Vermeidung der genannten Nachteile wird der Symptomwertebereich jedes Meßgeräts partitioniert, d. h. in disjunkte Intervalle eingeteilt; anschließend werden die Symptome nur noch durch das jeweils umschließende Intervall repräsentiert. Ein wichtiges Problem bei der Diskretisierung ist die Festlegung der geeigneten Intervallgrenzen, denn aufgrund der Vergrößerung kann ein Informationsverlust entstehen. Es stellt sich eine Optimierungsaufgabe: Werden die Intervalle zu groß gewählt, wird die Fehlervorhersage häufiger mehrdeutig, und die Wahrscheinlichkeit von Fehldiagnosen steigt. Dagegen bringt eine Diskretisierung in zu kleine Intervalle keinen Vorteil gegenüber der Verwendung der Ausgangswerte. Gesucht sind daher solche Partitionierungen, die eine gute Generalisierung darstellen und eine möglichst zuverlässige Fehlervorhersage erlauben.

Neben der Diagnoseleistung sind im Hydraulikbereich jedoch weitere Randbedingungen zu beachten. Durch die Einhaltung von individuellen Mindestintervallbreiten für jedes Meßgerät kann den jeweiligen Meß-, Ables- und Simulationsungenauigkeiten Rechnung getragen werden. Diese kleinstmöglichen Intervalle legen Bereiche der Ununterscheidbarkeit fest. Darüber hinaus ist eine Intervalleinteilung mit »glat-ten« Intervallgrenzen sinnvoll, die dem menschlichen Anwender die Interpretation der erzeugten Regeln erleichtern. Definition 4.3 legt die Kriterien für eine zulässige Intervalleinteilung fest.

Definition 4.3 (Partitionierung eines Meßwertebereichs): Es sei $m \in M$ ein Meßgerät, an dem die Meßwertdifferenzen $dom(m)$ beobachtet wurden und $u \in \mathbb{R}^{>0}$ ein absoluter Wert für die Meßunsicherheit von m . Dann ist eine Menge von $g \in \mathbb{N}$ Intervallen $I_m = \{i_1, \dots, i_g\}$ eine *Partitionierung des Meßwertebereichs von m* , wenn gilt:

1. $\forall j \in \{1, \dots, g\} : i_j = \begin{cases} [l_j, u_j[, & \text{falls } l_j < u_j \leq 0, \\]l_j, u_j], & \text{falls } 0 \leq l_j < u_j, \\ [0, 0] = 0 & \text{sonst.} \end{cases}$
2. $\forall j \in \{1, \dots, g-1\} : u_j = l_{j+1}$

5 So wird z. B. in ^{art}*deco* bei der Modellierung von Rohrleitungen auf Reibungszahlen, Krümmungswinkel usw. verzichtet. In der Realität eintretende Druckminderungen, die durch Reibungsverluste aufgrund von Widerständen an den Rohrwandungen oder Änderungen der Strömungsrichtung entstehen, sind also in der Simulation nicht berechenbar. Diese Fehler können sich im System fortpflanzen und gegenseitig verstärken.

3. $l_1 \leq \min_{a \in \text{dom}(m)} \{a\} < u_1$ und $u_g \geq \max_{a \in \text{dom}(m)} \{a\} > l_g$
4. $\forall j \in \{1, \dots, g\} : |u_j - l_j| \geq 2u$
5. $\forall j \in \{1, \dots, g\} : \exists x, y \in \mathbb{Z}, d \in \{1, 2, 4\}, \frac{10^y}{d} \geq 2u : l_j = \frac{x}{d} \cdot 10^y$
6. analog zu 5. für u_j . □

Anmerkung: Die Bedingungen in Definition 4.3 haben die folgende Bedeutung:

1. Festlegung der Interpretation der Intervallgrenzen; der Wert 0 bildet dabei immer ein eigenes »Intervall«, damit innerhalb eines Intervalls keine Vorzeichenwechsel vorkommen.
2. Sicherstellung, daß der Wertebereich der Symptome durch die Intervalle lückenlos abgedeckt wird.
3. Sicherstellung, daß sich die Extremwerte des Meßgeräts in den äußeren Intervallen befinden, so daß keine Bereiche ohne Symptome abgedeckt werden.
4. Definition der minimalen Intervallbreite durch die doppelte Meßungenauigkeit $2 \cdot u$; dies entspricht der Größe des Unsicherheitsintervalls für einen wahren Meßwert x_0 bei Beobachtung eines Symptomwerts s , denn es gilt: $x_0 \in [s - u, s + u]$ (vgl. Abschnitt 4.1.2).
5. bzw. 6. Festlegung, daß die Intervallgrenzen nur Vielfache der Zehnerpotenzen der einfachen Brüche $\frac{1}{1}$, $\frac{1}{2}$ oder $\frac{1}{4}$ sein dürfen. Diese werden im folgenden als *glatte Intervallgrenzen* bezeichnet.

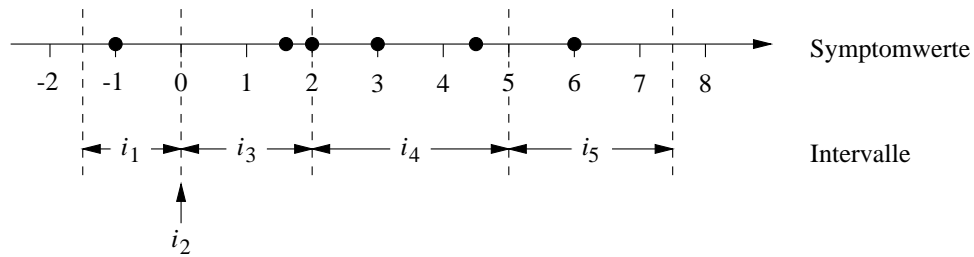
Bevor in den nächsten Abschnitten auf Algorithmen zur Suche einer zulässigen Intervalleinteilung eingegangen wird, gibt Beispiel 4.4 ein mögliches Ergebnis der Diskretisierung an.

Beispiel 4.4: Es sei m ein Meßgerät, an dem die Meßwerte mit der Meßunsicherheit $u = 0.25$ beobachtet und die Meßwertdifferenzen (Symptome) $-1.0, 1.6, 2.0, 3.0, 4.5$ und 6.0 berechnet wurden. Dann ist eine zulässige Partitionierung die Intervallmenge $I_m = \{i_1, \dots, i_5\}$ mit

$$i_1 = [-1.5, 0[, \quad i_2 = 0, \quad i_3 =]0, 2], \quad i_4 =]2, 5] \quad \text{und} \quad i_5 =]5, 7.5].$$

Sie genügt den Bedingungen der Definition 4.3; so gilt etwa für $u_5 = 7.5 = \frac{x}{d} \cdot 10^y$ mit $d = 2, y = 0$ und $x = 15$ (Bedingung 6).

Das folgende Bild veranschaulicht die Intervalleinteilung graphisch.



□

4.5.2 Bekannte Verfahren

Viele Algorithmen zum maschinellen Lernen benötigen einen diskreten, also höchstens ordinalskalierten Eingaberaum. Um sie trotzdem in Problembereichen mit kardinalskalierten Merkmalen anwenden zu können, müssen für die reellwertigen Eingabegrößen geeignete Intervalleinteilungen gefunden werden. In der Literatur sind unterschiedliche Diskretisierungsansätze untersucht worden, von denen einige auf den jeweiligen Klassifizierungs- oder Regellernalgorithmus zugeschnitten sind (z. B. »C4.5« von Quinlan [Qui93] sowie »1RD« von Holte [Hol93]), hingegen andere universell einsetzbar sind. Im weiteren werden nur die universellen Verfahren betrachtet.

In Anlehnung an Dougherty et al. ([DKS95]) lassen sich diese Verfahren hinsichtlich der folgenden drei Unterscheidungen charakterisieren:

- **Global vs. lokal:** Globale Verfahren analysieren die Eingabedaten als ganzes und erzeugen für jedes Merkmal genau eine Intervalleinteilung. Dagegen nehmen lokale Verfahren auch Diskretisierungen vor, die auf separate Teilbereiche der Eingabedaten optimiert sind; unter Umständen können also mehrere Intervalleinteilungen für ein Merkmal entstehen.
- **Unüberwacht vs. überwacht:** Unüberwachte Verfahren berücksichtigen zur Intervallbildung ausschließlich die beobachtete Verteilung der Merkmalswerte, während überwachte Verfahren zusätzlich die Information auswerten, mit welchen Klassen die einzelnen Merkmalswerte korrespondieren.
- **Statisch vs. dynamisch:** Statische Verfahren bestimmen für jedes Merkmal separat die Einteilung in eine benutzerdefinierte Anzahl von Intervallen. Im Gegensatz dazu berücksichtigen dynamische Verfahren die Auswirkungen, die die Diskretisierung eines Merkmals auf die Diskretisierung anderer Merkmale haben kann, um für jedes Merkmal die optimale Intervallanzahl und -einteilung zu erhalten.

Lokale Verfahren (z. B. Fayyad und Irani [FI92]) werden insbesondere beim Lernen von Entscheidungsbäumen eingesetzt, weil jedem inneren Knoten eine noch nicht klassifizierte Teilmenge der Eingabedaten entspricht, für die jedesmal ein lokales Diskretisierungsverfahren eingesetzt werden kann. Obwohl diese differenziertere Analyse der Datenbasis gegenüber schnelleren globalen Verfahren eine bessere Klassifikationsgüte erwarten läßt, konnte dieses von Dougherty et al. in ihrer Studie [DKS95] nicht nachgewiesen werden.

Üblicherweise teilen die unüberwachten Verfahren den Wertebereich in Intervalle auf, die entweder alle gleich groß sind oder alle gleich viele Merkmalswerte enthalten. Als Alternative kann auch eine Clusteranalyse durchgeführt werden. Dagegen bewerten überwachte Verfahren mit unterschiedlichen Maßen den Zusammenhang zwischen einem Merkmal und der Klasseninformation. Die bekannteren Ansätze verwenden dazu die Statistik (z. B. Tests basierend auf χ^2 : »ChiMerge« [Ker92], Φ : »Stat-Disc« [RR95], λ : »Zeta« [HS97]/[HS98]) oder die Informationstheorie (z. B. »D-2« [CBS91] und Pfahringers Verfahren [Pfa95]).

Konstruktionsgemäß erzielen überwachte Verfahren für Klassifikations- und Diagnoseprobleme auf die Vorhersagegüte bezogen gute Ergebnisse, es ist unter den genannten aber keines in der Lage, die besonderen Anforderungen an eine Intervalleinteilung im Zusammenhang mit der Diagnose realer hydraulischer Anlagen zu erfüllen (siehe Definition 4.3). Aus diesem Grund wird im folgenden Abschnitt ein neuer Diskretisierungsalgorithmus vorgestellt.

4.5.3 Neuer Diskretisierungsalgorithmus

Der im weiteren beschriebene Algorithmus zur Ermittlung von Intervalleinteilungen für Meßgeräte ist insbesondere für Diagnoseanwendungen im technischen Bereich geeignet. Er analysiert als überwachtes Verfahren, welche Fehler mit welchen Symptomwerten korrespondieren und hat gegenüber den bekannten Algorithmen den Vorteil, Vorgaben hinsichtlich der »glatten« Intervallgrenzen sowie hinsichtlich der Meßgerätequalitäten berücksichtigen zu können.

Zunächst werden einige diagnoserelevante Aspekte näher erläutert, bevor der eigentliche Algorithmus aufgeführt wird. Anschließend erfolgt eine Diskussion des Algorithmus sowie möglicher Verbesserungen.

Diagnoserelevante Aspekte

Zur Anpassung an unterschiedliche technische Anwendungssituationen berücksichtigt der Algorithmus benutzerdefinierte Vorgaben für das jeweilige Meßgerät. So

können als Eigenschaften eines Meßgeräts m die Meßungenauigkeit u , die maximale Intervallbreite t und die maximale Anzahl unterscheidbarer Meßbereiche bzw. Intervalle i_{max} definiert werden. In der maximalen Intervallanzahl i_{max} sind die symbolischen Intervalle (aus Definition 4.2) und das Null-Intervall (aus Definition 4.3) nicht inbegriffen. Im folgenden wird von $i_{max} > 2$ ausgegangen, da sonst kein Diskretisierungsalgorithmus benötigt wird:

- Für $i_{max} = 1$ kann nur zwischen dem obligatorischen Null-Intervall (stellt den Sollwert dar) und Differenzwerten ungleich 0 (stellt irgendein Symptom dar) unterschieden werden.⁶
- Für $i_{max} = 2$ kann nur zwischen dem obligatorischen Null-Intervall sowie positiven und negativen Differenzwerten unterschieden werden.

Ein wichtiger Aspekt zur Intervalleinteilung ist die Bestimmung der zulässigen Intervallgrenzen bzw. einer geeigneten Mindestintervallbreite. Insbesondere bei analogen Meßgeräten, die vom Ingenieur abgelesen werden, ist eine gleichmäßige Skalen- bzw. Intervalleinteilung sinnvoll. Die rechnerisch optimale Breite für gleich große Intervalle läßt sich dann aus der Intervallanzahl i_{max} und der Spannweite des Wertebereichs berechnen. Gleichzeitig müssen jedoch die Bedingungen 4 bis 6 der Definition 4.3 (Meßunsicherheit und »glatte« Intervallgrenzen) eingehalten werden. Dies führt zu der folgenden Berechnungsvorschrift für die meßgeräteabhängige Mindestintervallbreite S_m .

Es sei $m \in M$ ein Meßgerät, an dem die Symptomwerte a_1, \dots, a_s mit $a_1 < \dots < a_s$ beobachtet wurden. Weiter seien die Meßungenauigkeit u , die maximale Intervallbreite t und die maximale Intervallanzahl i_{max} gegeben. Dann läßt sich die Mindestintervallbreite S_m wie folgt berechnen:

$$S_m(u, t, i_{max}) = \begin{cases} 2u, & \text{falls } S^* < 2u, \\ t, & \text{falls } S^* > t, \\ \min_x |x - S^*|; x = \frac{10^y}{d}; y \in \mathbb{Z}; d \in \{1, 2, 4\} \text{ sonst.} & \end{cases}$$

Dabei ist S^* die Intervallbreite für äquidistante Intervalle, für die gilt:

$$S^* = \frac{a_s - a_1}{i_{max}}.$$

Beispiel 4.5: Innerhalb der Grenzen $u = 0.005$ und $t = 250$ für ein Meßgerät m ergibt sich der in Abbildung 4.8 ausschnittsweise dargestellte Zusammenhang zwischen der optimalen Intervallbreite S^* und der gerundeten Intervallbreite S_m . \square

⁶ Dieser Fall kann nicht der Definition 4.4 genügen, weil das Null-Intervall die Bedingung 2 verletzen würde.

Äquidistante Intervallbreite S^*			Gerundete Intervallbreite S_m
	> 175		250
75	$\dots 175$		100
37.5	$\dots 75$		50
17.5	$\dots 37.5$		25
	\vdots		\vdots
0.075	$\dots 0.175$		0.1
0.0375	$\dots 0.075$		0.05
0.0125	$\dots 0.0375$		0.025
< 0.0125			0.01

Abbildung 4.8: Zusammenhang zwischen optimaler und gerundeter Intervallbreite (zu Beispiel 4.5)

Die gerundete Intervallbreite S_m stellt gleichzeitig eine Mindestintervallbreite dar und wird dazu benutzt, vom Null-Intervall ausgehend in beide Richtungen ein Raster für die zulässigen Intervallgrenzen zu markieren. Hierfür wird abschließend eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ eingeführt, die für einen beliebigen Symptomwert die nächste, im Sinne der Definition 4.3 zulässige (»glatte«) Intervallgrenze berechnet:

$$f(x) \mapsto \left\lfloor \frac{x}{S_m} + \text{sign}(x) \cdot 0.5 \right\rfloor \cdot S_m.$$

Der im folgenden beschriebene Diskretisierungsalgorithmus greift auf diese Definitionen zurück.

Algorithmus

Nach der in Abschnitt 4.5.2 ausgeführten Einteilung handelt es sich bei dem neuen Intervallbildungssalgorithmus um ein globales, überwachtes und statisches Verfahren. Es geht von den Annahmen aus, daß die Symptomwerte für ein Meßgerät m etwa gleichverteilt sind und daß benachbarte Werte von unterschiedlichen Fehlern hervorgerufen wurden.

Anmerkung: Zwar können die erwähnten Annahmen im Einzelfall verletzt sein, dies muß sich aber auf die erzielbare Diagnosegüte nicht negativ auswirken, wie die folgende Überlegung zeigt. Ein Informationsverlust tritt nicht auf, wenn jedes Intervall genau ein Symptom umschließt, weil dann jedes Intervall eine eindeutige

Eingabe: Symptomwerte a_1, \dots, a_s einer Meßstelle $m \in M$, Meßunsicherheit u , maximale Intervallbreite t , Intervallanzahl i_{max}

Ausgabe: Intervallmenge I_m

- (1) Sortiere die Symptomwerte, so daß $a_1 < \dots < a_s$ gilt
- (2) Berechne die gerundete Mindestintervallbreite $S_m(u, t, i_{max})$
- (3) Führe die unterste und die oberste Intervallgrenze durch $f(a_1 - \frac{S_m}{2})$ bzw. $f(a_s + \frac{S_m}{2})$ ein
- (4) Führe 0 als Intervallgrenze ein
- (5) Für je zwei benachbarte Symptomwerte a_j und a_{j+1} :
- (6) Berechne das arithmetische Mittel $\bar{a} = \frac{a_j + a_{j+1}}{2}$
- (7) Führe $f(\bar{a})$ als neue Intervallgrenze ein
- (8) Bilde aus den Intervallgrenzen die entsprechende Intervallmenge I_m
- (9) Für jedes Intervall $i \in I_m$ ohne Symptomwert:
- (10) Ist die Anzahl der unterschiedlichen Fehlerkorrespondenzen der beiden Nachbarintervalle von i gleich?
- (11) Dann vereinige i mit dem kleineren Nachbarintervall
- (12) Sonst vereinige i mit dem Nachbarintervall, das mehr unterschiedliche Fehlerkorrespondenzen hat
- (13) Für je zwei benachbarte Intervalle in I_m :
- (14) Sind die korrespondierenden Fehlermengen gleich?
- (15) Dann vereinige diese beiden Intervalle

Algorithmus 4.1: Intervallbildung

Fehlervorhersage ermöglicht. Bei einer ungleichmäßigen Verteilung der Symptome (Verletzung der ersten Annahme) steigt aber wegen der einzuhaltenden Mindestintervallbreiten S_m die Wahrscheinlichkeit, daß einige Intervalle mehrere Symptome umschließen. Wenn nun die Symptome eines Intervalls mit dem gleichen Fehler korrespondieren (Verletzung der zweiten Annahme), ist dennoch eine eindeutige Fehlervorhersage möglich.

Der Algorithmus 4.1 zur Intervallbildung gliedert sich in zwei Hauptteile: Zunächst wird im ersten Teil (Zeilen 1 bis 8) eine grobe, aber zulässige Intervalleinteilung realisiert, die danach im zweiten Teil (Zeilen 9 bis 15) unter Berücksichtigung der Fehlerkorrespondenzen optimiert wird.

In den Zeilen 3 bis 7 werden die Intervallgrenzen ohne Berücksichtigung der Fehlerkorrespondenzen ermittelt. Dabei wird versucht, jeweils eine neue Intervallgrenze mittig zwischen zwei benachbarten Symptomwerten zu plazieren; die Anwendung der Funktion f verschiebt diese Grenze auf den nächsten zulässigen Wert (Zeilen 5 bis 7); hierdurch können leere Intervalle entstehen. Anschließend erfolgt die Auswertung der Fehlerkorrespondenzen, um leere Intervalle zu entfernen (Zeilen 9 bis 12) und um benachbarte Intervalle mit den gleichen Fehlerkorrespondenzen zu vereinigen (Zeilen 13 bis 15). Durch die genannten Optimierungen wird die Anzahl eindeutiger Fehlervorhersagen nicht verringert.

In Abhängigkeit von der Anzahl der Symptomwerte s beträgt die Laufzeit des Algorithmus $O(s \cdot \log s)$. Er wird vom Aufwand zur Sortierung der Symptomwerte in Zeile 1 bestimmt, da die Ermittlung der gerundeten Intervallbreite S_m nicht und die nachfolgenden Schleifen nur linear von s abhängig sind.

Beispiel 4.6: Es seien an einem Meßgerät $m \in M$ die folgenden Symptomwerte a_1, \dots, a_6 und einelementige Fehlerkorrespondenzen f_1, \dots, f_6 gegeben:

j	1	2	3	4	5	6
a_j	-1.0	1.6	2.0	3.0	4.5	6.0
f_j	{A}	{A}	{B}	{C}	{C}	{B}

Dann ergibt sich mit einer maximalen Intervallanzahl $i_{\max} = 7$ die Mindestbreite $S_m = 1.0$. Die mit Algorithmus 4.1 ermittelte Intervalleinteilung ist

$$i_1 = [-2, 0[, \quad i_2 = 0, \quad i_3 =]0, 2], \quad i_4 =]2, 5] \quad \text{und} \quad i_5 =]5, 7].$$

Abbildung 4.9 veranschaulicht das Ergebnis. Man erkennt, daß bis auf i_3 alle Intervalle eine eindeutige Fehlervorhersage ermöglichen. \square

4.6 Meßstellenauswahl

Die Meßstellenauswahl dient der Ermittlung derjenigen Meßstellen in der hydraulischen Anlage, die bei vorgegebenen Randbedingungen eine maximale Diagnoseleistung ermöglichen. Die Randbedingungen begrenzen dabei entweder die Anzahl einsetzbarer Meßgeräte oder den Kostenrahmen für Meßgeräte und Messungen. Ist eine optimale Meßstellenteilmenge $M^* \subseteq M$ gefunden, werden aus den vollständigen Intervallvektoren diejenigen Symptome entfernt, die an den Meßstellen aus $M \setminus M^*$ simuliert wurden. Zur Erzeugung der Diagnoseregeln wird dann die Datenbasis mit den so verkürzten Intervallvektoren analysiert.

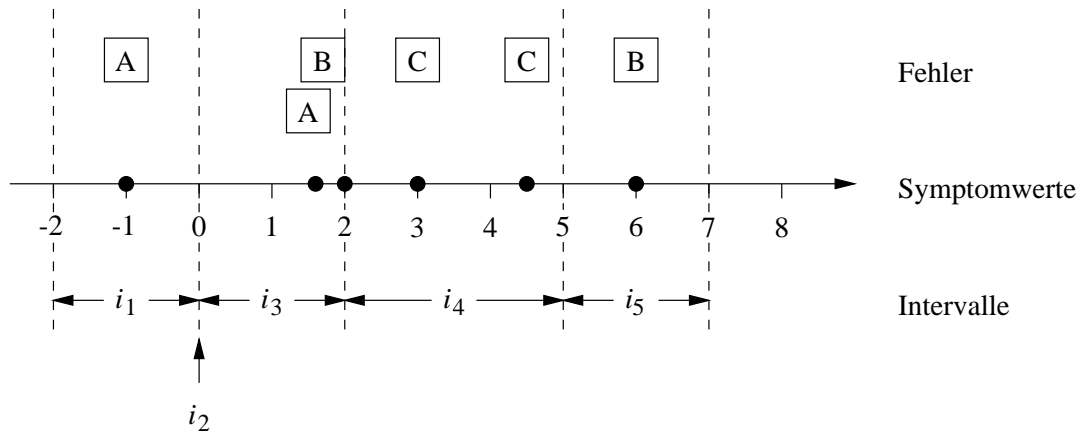


Abbildung 4.9: Ergebnis der Intervallbildung (zu Beispiel 4.6)

In Abschnitt 4.6.1 wird auf die wichtigsten Aspekte der Meßstellenauswahl eingegangen, und es werden einige vereinfachende Annahmen getroffen. In Abschnitt 4.6.2 erfolgt zunächst die Beschreibung statistischer Methoden der Abhängigkeitsanalyse zur Ermittlung redundanter Meßstellen, bevor in Abschnitt 4.6.3 Alternativen zur Bewertung von Meßstellenteilmengen entwickelt werden.

4.6.1 Motivation und Vereinbarungen

Bei der Aufstellung eines Meßplans (vgl. Abschnitt 4.3.1) werden für die Simulation alle potentiellen Meßstellen definiert. Es ist im allgemeinen aber nicht möglich, jede dieser Meßstellen in der realen hydraulischen Anlage tatsächlich zu berücksichtigen, denn häufig existieren Restriktionen bezüglich der Meßgeräteanzahl oder der Meßkosten. Die Auswahl der optimalen Meßstellen kann daher bereits die Konstruktion der zu diagnostizierenden Anlage beeinflussen und legt das Meßwertprotokoll für den Störfall fest.

Die eigentliche Meßphase ist die einzige Phase, die das Diagnosesystem nicht algorithmisch abbilden kann, denn das Meßwertprotokoll muß durch reale Messungen erhoben werden. Hierdurch werden Meßkosten verursacht, die je nach Art der Meßgröße und Position der Meßstelle sehr unterschiedlich sein können⁷. Im weiteren wird davon ausgegangen, daß zwei Kategorien von Meßstellen vorhanden sind:

⁷ Die Meßkosten entstehen in der Regel durch finanziellen und zeitlichen Aufwand (z. B. Kosten für Anschaffung des Meßgeräts, Zeit für Montage und Ablesen etc.).

1. Meßstellen, die keine oder vernachlässigbare Meßkosten verursachen.
Hierzu zählen die Meßstellen zur Bestimmung der Zylinderkräfte F (Belastungsstufen), Kolbenpositionen s und Drehzahlen n . Sie definieren die Menge der fest ausgewählten Meßgeräte M_{fix} , weil ihre Meßergebnisse ohne Aufwand zur Verfügung stehen.
2. Meßstellen, die Meßkosten über einem anwendungsabhängigen Schwellenwert verursachen.
Hierbei handelt es sich i. d. R. um Meßstellen zur Bestimmung von Drücken p und Flüssen q_V . Die entsprechenden Meßgeräte definieren die Menge der optionalen Meßgeräte M_{opt} , und sie unterliegen der Meßstellenauswahl.

Ferner soll vereinfachend angenommen werden, daß die Anzahl der auszuwählenden Meßstellen durch einen benutzerdefinierten Wert $v \leq |M_{opt}|$ ohne Differenzierung nach physikalischen Größen beschränkt ist und daß die Meßkosten an allen Meßstellen $m \in M_{opt}$ gleich groß sind. Somit ist für die Auswahl der besten v -elementigen Teilmenge von M_{opt} nur ihre Diagnoseleistung zu bewerten.

Der Aufwand zur Meßstellenauswahl kann deutlich reduziert werden, wenn als Pre-processing eine paarweise Abhängigkeitsanalyse durchgeführt wird. Von zwei als gegenseitig abhängig erkannten Meßstellen ist eine redundant, weil sie keine neue Diskriminierungsleistung erbringt. Sie kann ohne Informationsverlust von der weiteren Analyse ausgeschlossen werden, so daß sich die Anzahl der verbliebenen optionalen Meßstellen verringert. Im folgenden Abschnitt werden daher zunächst verschiedene Verfahren zur Abhängigkeitsanalyse diskutiert. Da jederzeit klar ist, daß sich alle Ausführungen auf die Menge der optionalen Meßstellen M_{opt} beziehen, wird diese auch einfach mit M bezeichnet.

4.6.2 Abhängigkeitsanalyse

Vorbemerkungen

Angenommen, die Fehler in einer hydraulischen Anlage wirken sich auf zwei Meßgeräte immer gleichartig aus (das heißt, sie verursachen immer miteinander korrespondierende Symptome, wie es zum Beispiel bei direkt benachbarten Meßgeräten der Fall sein kann); dann enthalten beide Meßstellen gemeinsam nicht mehr Diskriminierungsinformation als eine alleine. Als Konsequenz gilt, daß bei Berücksichtigung beider Meßstellen zwar die Meßkosten höher sind, der Nutzen aber nicht zunehmen kann. Sind dagegen gleichartige Symptomkorrespondenzen nicht oder nur teilweise

beobachtbar, kann ein positiver Nutzen entstehen, wie man am folgenden Beispiel erkennt.

Beispiel 4.7: Es seien die an den Meßgeräten m_1, m_2 und m_3 erkennbaren Symptome und Fehler f gemäß der Intervallvektoren in der linken Tabelle gegeben, wobei die Symptome von m_1 und m_2 eine gleiche Verteilung aufweisen und die Symptome von m_3 eine hierzu unterschiedliche:

m_1	m_2	m_3	f		Meßgerät(e)	Eindeutige Fehler
i_1	i_3	i_5	A	\Rightarrow	$\{m_1\}$	-
i_1	i_3	i_6	B		$\{m_2\}$	-
i_2	i_4	i_6	C		$\{m_3\}$	$\{A\}$
i_2	i_4	i_6	D		$\{m_1, m_2\}$	-
					$\{m_1, m_3\}$	$\{A\}, \{B\}$
					$\{m_2, m_3\}$	$\{A\}, \{B\}$

Dann können mit der Meßgerätekombination $\{m_1, m_2\}$ nicht mehr Fehler als mit nur einem Meßgerät eindeutig erkannt werden, wohl aber mit den Kombinationen $\{m_1, m_3\}$ und $\{m_2, m_3\}$ (siehe rechte Tabelle). \square

Zur Quantifizierung der Symptomverteilungen können die aus der statistischen Zusammenhanges- bzw. Abhängigkeitsanalyse bekannten Zusammenhangsmaße verwandt werden (siehe z. B. [Har89], [BB89], [Pol97]). In diesem Abschnitt wird die bivariate Datenanalyse zur Bewertung der gegenseitigen Abhängigkeit zweier Meßgeräte besprochen. Für die gleichzeitige Untersuchung von mehr als zwei Meßstellen ist eine multivariate Datenanalyse notwendig. Die Testergebnisse zeigen jedoch, daß für das Diagnosesystem eine bivariate Analyse ausreicht, weil mit der eigentlichen Regelerzeugung ein multivariates Verfahren nachgelagert ist.

Die bekannten Ansätze der bivariaten Abhängigkeitsanalyse unterscheiden sich sowohl in ihrer Komplexität als auch in ihren Voraussetzungen und Modellen. Im folgenden werden in Anlehnung an Hartung ([Har89]) mit dem χ^2 - und dem λ -Maß zwei bekannte, aber im Hinblick auf das Diagnoseproblem etwas veränderte statistische Maße zur Quantifizierung der Unabhängigkeit zweier Meßstellen beschrieben.

Bezeichnungen

Beide Maße basieren auf der Auswertung einer sogenannten Kreuztabelle, in der die beobachtete Häufigkeit aller möglichen Symptomkombinationen für zwei Meßgeräte m_1 und $m_2 \in M$ notiert wird. Bezeichnet $h(i, j)$ die gemeinsame absolute Häufigkeit

der Symptome $i \in I_{m_1}$ und $j \in I_{m_2}$ in der Datenbasis D_n , dann ist der Aufbau einer Kreuztabelle wie in Abbildung 4.10 angegeben.

		I_{m_2}			
		j_1	\cdots	$j_{ I_{m_2} }$	Σ
I_{m_1}	i_1	$h(i_1, j_1)$	\cdots	$h(i_1, I_{m_2})$	$h(i_1, \bullet)$
	\vdots	\vdots	\ddots	\vdots	\vdots
	$i_{ I_{m_1} }$	$h(I_{m_1} , j_1)$	\cdots	$h(I_{m_1} , I_{m_2})$	$h(I_{m_1} , \bullet)$
Σ		$h(\bullet, j_1)$	\cdots	$h(\bullet, I_{m_2})$	$h(\bullet, \bullet) = n$

Abbildung 4.10: Kreuztabelle für die Häufigkeiten der Symptomkombinationen zweier Meßgeräte m_1 und m_2

Als Abkürzung werden die als Zeilen- bzw. Spaltensummen definierten Randhäufigkeiten benutzt. Die Summe der Zeilenrandhäufigkeiten wie auch der Spaltenrandhäufigkeiten ergibt wieder die Anzahl der Datensätze n .

In der Statistikk-literatur ist die gemeinsame absolute Häufigkeit einer Merkmalskombination $h(i, j)$ in D_n üblicherweise datensatzbezogen definiert; es gilt also

$$h(i, j) = |\{d \in D_n \mid i \in d \wedge j \in d\}|.$$

Dahinter steht die Vorstellung, daß Merkmalsausprägungen, die häufiger gemeinsam beobachtet werden, für eine höhere Abhängigkeit der Merkmale sprechen. Im Zusammenhang mit der Diagnoseproblematik sind für die Beurteilung zweier Meßgeräte jedoch nicht die Symptomkombinationen entscheidend, die durch den *gleichen* Fehler verursacht wurden; aussagekräftiger ist, wie häufig gleiche Symptomkombinationen mit *unterschiedlichen* Fehlern korrespondieren. Eine Überlegung der Grenzfälle soll dies verdeutlichen: Angenommen, es liegen zwei verschiedene Symptomkombinationen (a, b) und (x, y) mit den folgenden Eigenschaften vor:

1. (a, b) hat die Häufigkeit $h(a, b) = 10$ und korrespondiert in allen Fällen mit dem selben Fehler f .
2. (x, y) hat ebenfalls die Häufigkeit $h(x, y) = 10$, korrespondiert aber mit 10 verschiedenen Fehlern.

Dies deutet im ersten Fall nur darauf hin, daß der Fehler f die Symptome a und b gleichzeitig verursacht, was auf dieselben physikalischen Zusammenhänge zurückzuführen ist. Dagegen ist im zweiten Fall die Wahrscheinlichkeit für abhängige Meßgeräte viel größer, weil sie sogar bei zehn verschiedenen Fehlern, also zehn verschiedenen physikalischen Zusammenhängen, die gleiche Wirkung anzeigen. Die bisherige Häufigkeitsberechnung $h(i, j)$ kann die beiden genannten Situationen jedoch nicht voneinander unterscheiden.

Einen Ausweg bietet eine neue Häufigkeitsfunktion, die das Auftreten von Symptomkombinationen in Korrespondenz zur *gleichen Fehlerart* nur einmal zählt. Auf diese Weise werden Effekte eliminiert, die allein aufgrund der gleichen physikalischen Zusammenhänge zu häufigen Symptomkombinationen führen. Die im folgenden beschriebenen statistischen Verfahren zur Abhängigkeitsanalyse sind deshalb um die neue Berechnung der gemeinsamen Häufigkeiten modifiziert worden:

$$h(i, j) := \sum_{f \in F} c_{i,j}(f) \quad \text{mit} \quad c_{i,j}(f) = \begin{cases} 1, & \text{falls } \exists d \in D_n : i \in d \wedge j \in d \wedge f \in d \\ 0 & \text{sonst.} \end{cases}$$

Anmerkung: Eine »echte« statistische Unabhängigkeit wird auch nach der neuen Häufigkeitsberechnung erkannt, weil ungleiche Symptomkombinationen zwangsläufig auch über mehrere Fehlerarten verteilt sind. Die neue Berechnung verringert aber die Gefahr, daß eine hohe statistische Abhängigkeit zum Ausschluß einer Meßstelle führt, obwohl sie einen Informationsgewinn zur Diskriminierung eines Teils der Diagnosen beitragen könnte.

χ^2 -Kontingenzmaß nach Pearson / Cramer

Innerhalb der statistischen Korrelationsrechnung sind verschiedene Maße zur Quantifizierung von Abhängigkeiten und Zusammenhängen (Interdependenzen) zweier Merkmale bekannt. Mit den hier eingesetzten Verfahren können Vergleiche zwischen den Korrelationen verschiedener Merkmalskombinationen durchgeführt werden.

Üblicherweise wird eine Einteilung der Zusammenhangsanalyseverfahren abhängig von der Skalierung der beiden Merkmale vorgenommen:

- Für nur kardinale Skalierungen: Bravais-Pearson-Korrelationskoeffizient,
- für zumindest ordinale Skalierungen: Rangkorrelationskoeffizient von Spearman und
- für alle anderen Skalierungskombinationen: Pearsonscher Kontingenzkoeffizient.

Besitzen die beiden Merkmale unterschiedliche Skalierungsarten, so muß das höhere Skalenniveau abgewertet werden, d. h., das niedrigere bestimmt das anzuwendende Verfahren. Bei dem betrachteten Diagnoseproblem bestehen die Ausprägungen neben ordinalen Intervallen auch aus nominalen Symbolen für vektorielle Meßgrößen (vgl. Abschnitt 4.4), daher wird im weiteren nur der Kontingenzkoeffizient nach Pearson beschrieben.

Der Kontingenzkoeffizient basiert auf dem sogenannten χ^2 -Maß, das für zwei Meßgeräte m_1 und $m_2 \in M$ den Unterschied zwischen der tatsächlichen Häufigkeitsverteilung und einer hypothetischen Verteilung bei angenommener Unabhängigkeit ermittelt. Dazu werden die relativen quadratischen Abweichungen zwischen den in der Datenbasis D_n beobachteten gemeinsamen absoluten Häufigkeiten $h(i, j)$ und den unter der Annahme der Unabhängigkeit berechneten gemeinsamen Häufigkeiten $\tilde{h}(i, j)$ für alle Intervallkombinationen aus $i \in I_{m_1}$ und $j \in I_{m_2}$ aufsummiert. Definition 4.4 zeigt die Berechnung in Anlehnung an [Har89].

Definition 4.4 (Quadratische Kontingenz χ^2): Es seien m_1 und $m_2 \in M$ zwei Meßgeräte mit den zugehörigen Intervallmengen I_{m_1} und I_{m_2} , deren kombinierte Häufigkeitsverteilung h gemäß Abbildung 4.10 vorliegt. Weiter sei n die Anzahl der Symptomvektoren in der Datenbasis D_n . Dann ist die *quadratische Kontingenz* $\chi^2(m_1, m_2)$ definiert als

$$\chi^2(m_1, m_2) = \sum_{i \in I_{m_1}} \sum_{j \in I_{m_2}} \frac{(h(i, j) - \tilde{h}(i, j))^2}{\tilde{h}(i, j)} \quad \text{mit} \quad \tilde{h}(i, j) = \frac{h(i, \bullet) \cdot h(\bullet, j)}{n}. \quad \square$$

Gilt $\chi^2(m_1, m_2) = 0$, dann weichen die berechneten gemeinsamen Häufigkeiten nicht von den beobachteten ab, und das Meßgerät m_1 ist statistisch unabhängig von m_2 . Der Maximalwert für die quadratische Kontingenz tritt bei vollständiger Abhängigkeit auf und beträgt n ; er ist also abhängig von der Größe der Datenbasis. Um für alle $\chi^2(m_i, m_j)$ einen einheitlichen Wertebereich zu erhalten, sind in der Literatur mehrere Normierungsansätze bekannt. Die beiden in Definition 4.5 vorgestellten Ansätze normieren den ermittelten χ^2 -Wert jeweils auf das Intervall $[0, 1]$; sie sind [Har89] entnommen.

Definition 4.5 (Kontingenzkoeffizient K^* , Cramers V^2): Es seien m_1 und $m_2 \in M$ zwei Meßgeräte, χ^2 ihre quadratische Kontingenz nach Definition 4.4. Dann sind der (korrigierte) *Pearsonsche Kontingenzkoeffizient* $K^*(m_1, m_2)$ und *Cramers Assoziationsmaß* $V^2(m_1, m_2)$ definiert als

$$K^*(m_1, m_2) = \frac{\sqrt{\frac{\chi^2(m_1, m_2)}{n + \chi^2(m_1, m_2)}}}{\sqrt{\frac{k-1}{k}}} \quad \text{und} \quad V^2(m_1, m_2) = \sqrt{\frac{\chi^2(m_1, m_2)}{n \cdot (k-1)}},$$

wobei $k := \min\{|I_{m_1}|, |I_{m_2}|\}$ gilt. \square

Mit einem normierten χ^2 -Maß kann die Abhängigkeitsanalyse für eine Meßstellenmenge M mit dem Aufwand von $O(|M|^2)$ durchgeführt werden. Dazu erfolgt die phasenweise Bewertung von je zwei Meßstellen. Ein Ergebnis nahe bei 1 zeigt für die betrachtete Phase eine starke Abhängigkeit der beteiligten Meßgeräte an. Tritt dieses Ergebnis in allen Phasen auf, so kann eine der beiden Meßstellen entfernt werden und muß bei der weiteren Analyse nicht mehr berücksichtigt werden. Der Schwellenwert, ab dem zwei Meßstellen als abhängig zu betrachten sind, hängt vom Problembereich ab. Dies führt zu der folgenden Definition der Abhängigkeit zweier Meßstellen.

Definition 4.6 (Symmetrische Abhängigkeit zweier Meßstellen): Es seien α_K bzw. $\alpha_V \in [0, 1]$ geeignete Schwellenwerte und P die Menge der Phasen der betrachteten Anlage. Dann heißen zwei Meßstellen m_1 und $m_2 \in M$ *voneinander symmetrisch abhängig*, wenn gilt:

$$\min_{p \in P} \{K^*(m_1, m_2)\} \geq \alpha_K \quad \text{bzw.} \quad \min_{p \in P} \{V^2(m_1, m_2)\} \geq \alpha_V. \quad \square$$

An der Definition der χ^2 -basierten Maße ist erkennbar, daß diese symmetrisch sind, d. h., es ist keine Aussage darüber möglich, ob die Abhängigkeiten einseitig sind. Daher ist es schwierig, bei erkannter Abhängigkeit zu entscheiden, welches Meßgerät entfernt werden soll. Auch die Definitionslücken bei $k = 1$ verhindern eine sinnvolle Entscheidung, die dennoch möglich wäre, wie Beispiel 4.8 zeigt.

Beispiel 4.8: Es seien die an den Meßgeräten m_1 und $m_2 \in M$ beobachteten Symptome wie folgt verteilt:

m_1	m_2
i_1	i_2
i_1	i_3
i_1	i_4

Dann ist wegen $k = \min\{|I_{m_1}|, |I_{m_2}|\} = 1$ kein Kontingenzkoeffizient nach Definition 4.5 berechenbar. Offensichtlich ist jedoch m_1 vollständig abhängig von m_2 , aber m_2 vollständig unabhängig von m_1 . Es sollte also m_1 entfernt werden, da die Diskriminierungsleistung von m_2 größer ist. \square

Eine Situation wie in Beispiel 4.8 ist in der Praxis durchaus häufiger möglich, weil die Symptomstatistiken nach Phasen getrennt ausgewertet werden. Besonders in zeitlich frühen sowie späten Anlagenphasen können manche Meßgeräte keine Symptominformationen liefern, so daß undefinierte Abhängigkeitswerte die Meßstellenauswahl

erschweren. Diese Überlegungen zeigen, daß ein symmetrisches Abhängigkeitsmaß für die in dieser Arbeit betrachtete Diagnoseaufgabe nicht geeignet ist; daher werden im weiteren die asymmetrischen λ -Maße besprochen.

Proportionale Fehlerreduktion (λ -Maße)

Mit einem weiteren statistischen Verfahren läßt sich ein Abhängigkeitsmaß für Meßstellen gewinnen, nämlich mit dem Modell der proportionalen Fehlerreduktion (engl. *proportional reduction of errors, PRE*), zu denen die sogenannten λ -Maße gehören ([GK54]). Hierbei wird versucht, die Intervalle eines Meßgeräts vorherzusagen, einmal ohne und einmal mit Ausnutzung der Kenntnis, wie die Intervallverteilung eines weiteren Meßgeräts ist. Der Anteil der Fehlvorhersagen nimmt dabei um so stärker ab, je mehr Information das zusätzliche Meßgerät liefern kann, oder anders ausgedrückt, je unabhängiger das erste vom zweiten ist.

Die relative Reduzierung des Vorhersagefehlers für die Intervalle eines Meßgeräts, die abhängig von der Intervallverteilung eines weiteren Meßgeräts erzielt werden kann, wird mit den λ -Maßen quantifiziert. Bezogen auf zwei Meßgeräte $m_1 \in M$ und $m_2 \in M \setminus \{m_1\}$ sind sie asymmetrisch, da die Richtung der Abhängigkeit von der Wahl des vorherzusagenden Meßgeräts abhängt. Im folgenden gibt Definition 4.7 in Anlehnung an [EMT95] die Notation und Interpretation der Abhängigkeitsmaße an.

Definition 4.7 (Abhängigkeitsmaße λ_{m_1} und λ_{m_2}): Es seien m_1 und $m_2 \in M$ zwei Meßstellen. Weiter seien $e(m_1)$ die Anzahl der Vorhersagefehler für m_1 ohne Berücksichtigung von m_2 und $e(m_1|m_2)$ die Anzahl der Vorhersagefehler für m_1 mit Berücksichtigung von m_2 . Dann sind die asymmetrischen Abhängigkeitsmaße $\lambda_{m_1}(m_2)$ und $\lambda_{m_2}(m_1)$ definiert als relative Fehlerreduzierungen mit

$$\lambda_{m_1}(m_2) = \frac{e(m_1) - e(m_1|m_2)}{e(m_1)}$$

zur Beschreibung der Abhängigkeit m_1 von m_2 ($m_1 \leftarrow m_2$) bzw.

$$\lambda_{m_2}(m_1) = \frac{e(m_2) - e(m_2|m_1)}{e(m_2)}$$

zur Beschreibung der Abhängigkeit m_2 von m_1 ($m_1 \rightarrow m_2$). □

Durch die in Anhang A.1 gezeigte Herleitung gelangt man zu der folgenden Berechnung von $\lambda_{m_1}(m_2)$:⁸

$$\lambda_{m_1}(m_2) = \frac{\sum_{j \in I_{m_2}} \max_{i \in I_{m_1}} \{h(i, j)\} - \max_{i \in I_{m_1}} \{h(i, \bullet)\}}{n - \max_{i \in I_{m_1}} \{h(i, \bullet)\}}.$$

⁸ Auf analoge Weise läßt sich die umgekehrte Abhängigkeit $\lambda_{m_2}(m_1)$ herleiten.

Die λ -Maße sind auf den Wertebereich $[0, 1]$ normiert; der Wert 0 bedeutet, daß keine Fehlerreduktion und damit keine Abhängigkeit vorliegt, der Wert 1 weist auf eine vollständige Abhängigkeit hin. Für den Fall, daß das abhängige Meßgerät m_1 genau ein Intervall liefert (vgl. auch Beispiel 4.8), existiert eine Definitionslücke für $\lambda_{m_1}(m_2)$. Sofern das andere Meßgerät mehr als ein Intervall liefert ($|I_{m_2}| > 1$), ist das λ -Maß für die entgegengesetzte Abhängigkeitsrichtung, also $\lambda_{m_2}(m_1)$, definiert.

Auch mit den λ -Maßen kann die Abhängigkeitsanalyse für eine Meßstellenmenge M mit dem Aufwand von $O(|M|^2)$ durchgeführt werden. Dazu wird phasenweise eine $|M| \times |M|$ -Tabelle mit allen Meßstellenkombinationen erzeugt, in der jede Zelle $m_{i,j}$ einen Eintrag mit $\lambda_{m_j}(m_i)$, also die Abhängigkeit m_j von m_i enthält (siehe Abbildung 4.11). Auf der Diagonale gilt trivialerweise $\lambda_{m_i}(m_i) = 1$, weil jedes Meßgerät von sich selbst abhängig ist. Ein Wert $\lambda_{m_j}(m_i)$ mit $i \neq j$, der in allen Phasen nahe bei 1 ist, läßt sich als Abhängigkeit $m_i \rightarrow m_j$ interpretieren, und die Meßstelle m_j kann entfernt werden.

	m_1	m_2	\dots	$m_{ M }$
m_1	1	$\lambda_{m_2}(m_1)$	\dots	$\lambda_{m_{ M }}(m_1)$
m_2	$\lambda_{m_1}(m_2)$	1		$\lambda_{m_{ M }}(m_2)$
\vdots	\vdots		\ddots	\vdots
$m_{ M }$	$\lambda_{m_1}(m_{ M })$	\dots		1

Abbildung 4.11: Abhängigkeitstabelle für Meßstellenpaare

Abschließend kann also mit einem anwendungsabhängigen Schwellenwert die asymmetrische Abhängigkeit zweier Meßstellen wie folgt definiert werden.

Definition 4.8 (Asymmetrische Abhängigkeit zweier Meßstellen): Es sei $\alpha \in [0, 1]$ ein geeigneter Schwellenwert und P die Menge der Phasen der betrachteten Anlage. Dann heißt eine Meßstelle $m_1 \in M$ *abhängig* von einer Meßstelle $m_2 \in M$, wenn gilt:

$$\min_{p \in P} \{\lambda_{m_1}(m_2)\} \geq \alpha. \quad \square$$

4.6.3 Bewertungsfunktionen für die Meßstellenauswahl

In diesem Abschnitt werden ein bekannter und zwei neue Ansätze zur optimierten Auswahl von $v \leq |M|$ Meßstellen beschrieben. Die Betrachtung der Ansätze zeigt,

daß ein Konflikt zwischen dem Berechnungsaufwand und der erzielbaren Diagnosegüte besteht. Während der erste Ansatz in Anlehnung an [FdK93] jede Meßstelle einzeln bewertet und dafür eine suboptimale Auswahl zuläßt, werden beim zweiten Ansatz zwar im ungünstigsten Fall alle $\sum_{v'=1}^v \binom{|M|}{v'}$ Meßstellenkombinationen untersucht, aber die beste Auswahl wird gefunden. Auch der dritte Ansatz ist ein vollständiges Verfahren und ermittelt informationstheoretisch den mit v Meßstellen zu erwartenden Diskriminierungsaufwand.

Jeder der drei Bewertungsansätze greift auf eine Funktion zurück, die zu einem einzelnen Symptom oder zu einem Symptomvektor die korrespondierende Diagnosenmenge liefert. Diese Funktion wird im folgenden mit *Diagnosefunktion* bezeichnet und in Definition 4.9 eingeführt.

Definition 4.9 (Diagnosefunktion): Es sei $M = \{m_1, \dots, m_c\}$ die Menge der Meßgeräte mit den zugehörigen Intervallmengen I_1, \dots, I_c . Weiter seien $a_1, \dots, a_s \in \{1, \dots, c\}$ Meßgeräteindizes für einen Intervallvektor und F die Menge der Fehler. Dann wird eine Funktion

$$\delta : I_{a_1} \times \dots \times I_{a_s} \rightarrow P(F),$$

die einen s -stelligen Symptomvektor auf ihre korrespondierende Fehlermenge abbildet, *Diagnosefunktion* genannt.

Es gelte für den Sonderfall der »leeren« Beobachtung ($s = 0$): $\delta(\langle \rangle) = F$. □

Anmerkung: Die Realisation der Diagnosefunktion ist recht einfach möglich. Ausgedrückt durch die Operationen *Selektion* σ und *Projektion* Π der Relationenalgebra (vgl. z. B. [KE97]) kann die zu einem Intervall- bzw. Symptomvektor \vec{s} korrespondierende Fehlermenge aus einer Datenbasis D_n durch

$$\delta(\vec{s}) = \Pi_f(\sigma_{\vec{s}}(D_n))$$

ermittelt werden, wobei f das Fehlerattribut ist.

Bewertung einzelner Meßstellen

Die Bewertung einzelner Meßstellen basiert auf einem Vorschlag von Forbus und de Kleer in [FdK93]. Hierbei gehen sie davon aus, daß während des Fehlerdiagnoseprozesses sukzessive Messungen zur Erhebung von Symptomen durchgeführt werden. Dies geschieht solange, bis eine eindeutige Diagnose möglich ist. Jede zusätzliche Messung liefert einen Informationsgewinn, der zu einer weiteren Einschränkung der verbliebenen Fehlermenge führen kann. Damit bei einer eingeschränkten Meßgeräteanzahl die Berechnung der optimalen Teilmenge nicht zu aufwendig wird, verwenden Forbus und de Kleer ein Greedy-Verfahren, das lokal die beste zusätzliche Meßstelle bestimmt, ohne die bisherige Auswahl zu revidieren.

Statt jeder Meßstelle $m \in M$ den erwarteten Informationsgewinn zuzuordnen, wird als Vorschau der erwartete restliche Diskriminierungsaufwand für die verbliebene Fehlermenge mit einer Bewertungsfunktion

$$f: M \rightarrow \mathbb{R}$$

bestimmt. Zur Identifizierung der lokal besten Meßstelle wird dann das minimale Ergebnis der Funktion f herangezogen. Es folgt eine Erläuterung dieser Funktion sowie der Anpassungen, die im Hinblick auf die Diagnoseproblematik in dieser Arbeit notwendig sind.

Seien I_m die Menge der Intervalle einer Meßstelle $m \in M$ und δ eine Diagnosefunktion nach Definition 4.9. Angenommen, es wird das Symptom $s \in I_m$ beobachtet, dann enthält $\delta(i)$ die Menge der hiermit korrespondierenden Fehler. Es besteht kein weiterer Diskriminierungsbedarf, wenn $\delta(s)$ genau einen Fehler enthält. Für den Fall, daß mehrere Fehler möglich sind ($|\delta(s)| > 1$), muß zur Unterscheidung dieser Fehler mindestens eine weitere Meßstelle zur Messung herangezogen werden. Forbus und de Kleer bewerten die Kosten c des verbleibenden Diskriminierungsaufwands mit der Anzahl der zu erwartenden Folgemessungen. Diese Anzahl läßt sich informationstheoretisch mit $\log_k |\delta(s)|$ ermitteln, wobei $k \in \mathbb{N}$ die durchschnittliche Intervallanzahl der restlichen Meßgeräte in $M \setminus \{m\}$ ist (Zur Herleitung siehe Anhang A.2, insbesondere Definition A.1).

Zu dem erwarteten Restdiskriminierungsaufwand $f(m)$ einer Meßstelle $m \in M$ führt dann die Summe der mit den Symptomwahrscheinlichkeiten $P(s)$ gewichteten Diskriminierungskosten $c(|\delta(s)|)$:

$$f(m) = \sum_{s \in I_m} P(s) \cdot c(|\delta(s)|).$$

In [FdK93] wird die Wahrscheinlichkeit $P(s)$ über die relative Häufigkeit der verbliebenen Fehlermenge berechnet, weil in der dort beschriebenen Diagnoseproblemstellung jeder Fehler eine eindeutige Wirkung besitzt. Im Gegensatz dazu kann in der vorliegenden Arbeit ein Fehler *unterschiedliche* Symptomvektoren bewirken (zum Beispiel aufgrund unterschiedlicher Belastungsstufen), so daß für $P(s)$ die relative Häufigkeit des Symptoms s heranzuziehen ist ($P(s) = \frac{|\{d \in D_n | s \in d\}|}{n}$ für n Datensätze).

Zusammenfassung: Nach dem Ansatz der Einzelbewertung werden die v durch f niedrigst bewerteten Meßstellen aus M ausgewählt, d. h., es wird eine Teilmenge $M^* \subseteq M$ mit $|M^*| = v$ gesucht, für die gilt:

$$\max_{m \in M^*} \{f(m)\} \leq \min_{m \in M \setminus M^*} \{f(m)\},$$

wobei

$$f(m) = \sum_{s \in I_m} \frac{|\{d \in D_n \mid s \in d\}|}{n} \cdot \log_k |\delta(s)|.$$

Bewertung von Meßstellenkombinationen

Wird statt der Bewertung einzelner Meßstellen die kombinierte Bewertung von maximal v Meßstellen vorgenommen, steigt zwar der Berechnungsaufwand beträchtlich, aber die Gefahr, nur ein lokales Optimum zu finden, entfällt. Im folgenden wird daher ein neuer Ansatz entwickelt, der eine vollständige Exploration des Suchraums für maximal v -elementige Meßstellenmengen vornimmt und das Minimum der Bewertungen über alle Meßstellenkombinationen sucht.

Zur Meßstellenbewertung gemäß des neuen Ansatzes werden drei Alternativen aufgeführt, die alle einer v -elementigen Meßstellenmenge $M_{(v)} = \{m_{a_1}, \dots, m_{a_v}\}$ einen vergleichbaren Kostenwert durch eine Funktion f_1 , f_2 bzw. f_3 mit

$$f_{1,2,3} : P(M) \rightarrow \mathbb{R}$$

zuweisen. Die Alternativen unterscheiden sich in der Bewertung des nach Auswahl einer Meßstellenkombination verbliebenen Diskriminierungsaufwands.

1. Alternative: Die einfachste Möglichkeit stellt eine kanonische Erweiterung der Bewertung einzelner Meßstellen (siehe vorheriger Abschnitt) dar. Es gilt dann für die Bewertungsfunktion f_1 :

$$f_1(M_{(v)}) = \sum_{\vec{s} \in I_{a_1} \times \dots \times I_{a_v}} P(\vec{s}) \cdot \log_k |\delta(\vec{s})|,$$

wobei δ eine Diagnosefunktion nach Definition 4.9 ist und die Wahrscheinlichkeit für die Beobachtung eines Symptomvektors \vec{s} mit

$$P(\vec{s}) = \frac{|\{d \in D_n \mid \vec{s} \subseteq d\}|}{n}$$

berechnet wird.

Die Bewertung der Kosten durch eine Abschätzung des verbleibenden Diskriminierungsaufwands mit der erwarteten Anzahl der Folgemessungen hat den Nachteil, daß nur die mit einem Symptomvektor korrespondierenden möglichen Fehlerarten berücksichtigt werden, aber nicht ihre Häufigkeitsverteilung. Der tatsächliche Informationsgewinn, der durch die Beobachtung eines Symptomvektors entsteht, hängt

jedoch auch von der Verteilung der vorhersagbaren Fehler ab: Bei einer Gleichverteilung ist der Informationsgewinn minimal, weil der resultierende Informationsgehalt maximal ist (vgl. Satz A.1 im Anhang). Diese Überlegung führt zur zweiten Alternative.

2. Alternative: Wird der Shannonsche Informationsgehalt der Fehlermenge zur Abschätzung des verbleibenden Diskriminierungsaufwands herangezogen, ergibt sich die folgende Bewertungsfunktion f_2 :

$$f_2(M_{(v)}) = \sum_{\vec{s} \in I_{a_1} \times \dots \times I_{a_v}} \left(P(\vec{s}) \cdot \sum_{f \in \delta(\vec{s})} \left(P(f|\vec{s}) \cdot \log \frac{1}{P(f|\vec{s})} \right) \right),$$

wobei $P(\vec{s})$ wie oben definiert ist und $P(f|\vec{s})$ die bedingte Wahrscheinlichkeit für das Vorliegen des Fehlers f bezeichnet, wenn der Symptomvektor \vec{s} beobachtet wurde:

$$P(f|\vec{s}) = \frac{|\{d \in D_n \mid \vec{s} \subseteq d \wedge f \in d\}|}{|\{d \in D_n \mid \vec{s} \subseteq d\}|}.$$

3. Alternative: Eine weitere Möglichkeit zur Kostenbewertung von $\delta(\vec{s})$ ist die Abschätzung des erwarteten Diagnosefehlers. Der beste Schätzer zur Erzielung des minimalen Diagnosefehlers ist der Modalwert der Fehlermenge, da diese Diagnose am häufigsten korrekt ist. Die Wahrscheinlichkeit für eine falsche Fehlervorhersage beträgt dann $1 - \max_{f \in \delta(\vec{s})} \{P(f|\vec{s})\}$. Damit ergibt sich für die dritte Bewertungsfunktion f_3 :

$$f_3(M_{(v)}) = \sum_{\vec{s} \in I_{a_1} \times \dots \times I_{a_v}} \left(P(\vec{s}) \cdot \left(1 - \max_{f \in \delta(\vec{s})} \{P(f|\vec{s})\} \right) \right).$$

Das folgende Beispiel vergleicht die drei alternativen Kostenbewertungen.

Beispiel 4.9: Es sei die folgende Datenbasis D_n mit sechs Symptomvektoren $\vec{s}_1, \dots, \vec{s}_6$, die an drei unabhängigen Meßstellen m_1, m_2 und m_3 beobachtet wurden, sowie den Fehlerzuordnungen f aus der Fehlermenge $F = \{A, B, C\}$ gegeben:

	m_1	m_2	m_3	f
\vec{s}_1	i_1	i_2	i_1	A
\vec{s}_2	i_1	i_2	i_3	A
\vec{s}_3	i_2	i_1	i_2	B
\vec{s}_4	i_2	i_2	i_2	A
\vec{s}_5	i_2	i_1	i_1	B
\vec{s}_6	i_3	i_2	i_3	C

Dann kann der vollständige Suchraum zur Meßstellenauswahl wie in Abbildung 4.12 als Baumstruktur dargestellt werden. Die Kanten sind mit je einer ausgewählten Meßstelle markiert. In jedem Knoten der Tiefe t steht die Diagnosemenge $F \setminus \delta(M_{(t)})$, die nach Auswertung der in diesem Pfad ausgewählten Meßstellen $M_{(t)}$ noch zu diskriminieren ist: In der Wurzel ist dies die Gesamtfehlermenge F , in den inneren Knoten eine nichtleere Teilmenge von F und in den Blättern die leere Menge.

Abhängig von der eingesetzten Bewertungsalternative wird jedem Knoten bis zur Tiefe v , der maximalen Meßstellenanzahl, ein Wert zugewiesen. Gesucht ist dann ein minimal bewerteter Pfad, und die Meßstellen auf diesem Pfad bilden die optimale Auswahl. In Abbildung 4.12 ist ein solcher Pfad (mit den Meßgeräten m_1 und m_2) grau hinterlegt. Für den Fall $v = 2$ liefern die Bewertungsfunktionen f_1 , f_2 und f_3 die folgenden Ergebnisse:

	$\{m_1\}$	$\{m_2\}$	$\{m_3\}$	$\{m_1, m_2\}$	$\{m_1, m_3\}$	$\{m_2, m_3\}$
f_1	0.50	0.67	1.58	0.00	0.33	0.33
f_2	0.46	0.59	1.00	0.00	0.33	0.33
f_3	0.17	0.17	0.50	0.00	0.17	0.17

Die Meßstellenkombination $\{m_1, m_2\}$ diskriminiert die Menge der Symptomvektoren vollständig und wird mit jeder Bewertungsalternative als Optimum identifiziert. Auch wenn die Funktionen unterschiedliche Werte liefern, bleibt in diesem Beispiel die Rangordnung der Meßstellenkombinationen gleich. \square

Zusammenfassung: Von den v -stelligen Meßstellenkombinationen wird diejenige gewählt, die je nach angewendeter Bewertungsfunktion f_1 , f_2 oder f_3 den niedrigsten Wert erhält, d. h., es wird ein Meßstellenvektor $M_{(v)}^*$ gesucht, für den gilt:

$$f(M_{(v)}^*) = \min_{M_{(v)}} \{f_{1,2,3}(M_{(v)})\}.$$

Bewertung des erwarteten Diskriminierungsaufwands

Stein schlägt in [Ste01] einen neuen Ansatz vor, der sich grundsätzlich von den bisher vorgestellten Verfahren unterscheidet. Hierbei wird nicht eine einzelne Meßstellenmenge bewertet, sondern eine Eigenschaft der Fehlermenge F mit einer skalierbaren Bewertungsfunktion quantifiziert. Diese Eigenschaft ist der restliche Diskriminierungsaufwand, der zu erwarten ist, wenn die informationstheoretisch besten v Meß-

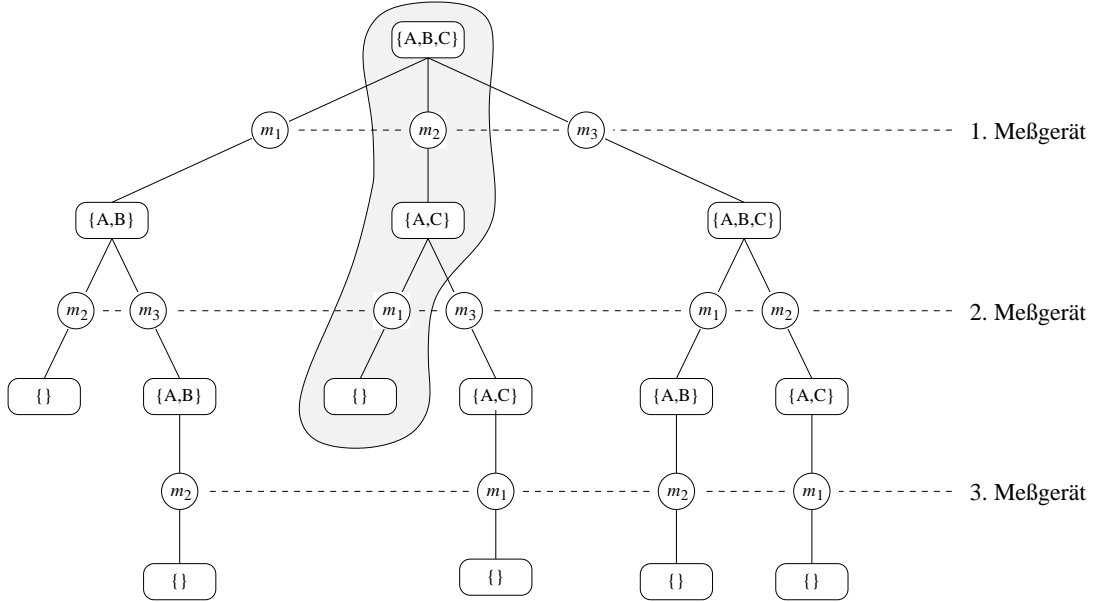


Abbildung 4.12: Suchbaum für die Meßstellenauswahl (zu Beispiel 4.9)

geräte ausgewählt wurden (Definition 4.10). Als Voraussetzung muß gelten, daß die möglichen Fehler aus F gleichverteilt sind.

Definition 4.10 (Erwarteter Diskriminierungsaufwand): Es sei $M = \{m_1, \dots, m_g\}$ eine Menge von Meßgeräten mit den zugehörigen Symptommengen I_1, \dots, I_g . Für ein Symptom $s \in I_j$ (mit $j \in \{1, \dots, g\}$) und eine Teilmenge $D \subseteq F$ der Fehler sei $P(s|D)$ die Wahrscheinlichkeit, daß das Symptom s beobachtet wird, wenn ein Fehler aus D vorliegt. Weiter sei δ eine Diagnosefunktion nach Definition 4.9 und k die mittlere Intervallanzahl der Meßstellen ($k = \frac{1}{g} \cdot \sum_{j=1}^g |I_j|$).

Dann ist der *erwartete Diskriminierungsaufwand* h , der zur Diskriminierung von D nach der Auswertung von n Meßgeräten ($1 \leq n \leq g$) noch zu leisten ist, durch

$$h(D, n) = \begin{cases} \min_{m \in M} \left\{ \sum_{s \in I_m} P(s|D) \cdot h(D \cap \delta(s), n-1) \right\}, & \text{falls } n > 0 \wedge |D| > 1 \\ \log_k |D|, & \text{sonst} \end{cases}$$

definiert. □

Von einem Algorithmus, der h mit der vollständigen Fehlermenge F und der ma-

ximalen Meßstellenanzahl v aufruft, wird der Lösungsraum rekursiv bis zur Tiefe v durchsucht und der jeweils darunter liegende Diskriminierungsaufwand anhand der verbliebenen Fehlermenge D mit $\log_k |D|$ abgeschätzt. In jeder vorherigen Rekursionsebene $t < v$ wird eine optimale Meßstellenkombination mit $v - t$ Meßgeräten als Minimum über alle mit $v - t - 1$ Meßgeräten verbliebenen Diskriminierungsaufwände berechnet. Am Ende sind die v besten Meßgeräte dem minimal bewerteten Rekursionspfad zu entnehmen und als Ergebnis der Meßstellenauswahl anzusehen.

Auch bei diesem Ansatz wird die im bezug auf ihre Diagnoseleistung beste Meßstellenkombination garantiert gefunden. Dafür müssen jedoch im schlechtesten Fall alle $\sum_{v'=1}^v \binom{|M|}{v'}$ Kombinationen bewertet werden. Der folgende Abschnitt vergleicht die vorgestellten Verfahren zur Meßstellenauswahl untereinander.

Diskussion der Ansätze

Abschließend soll eine kurze vergleichende Beurteilung der drei beschriebenen Bewertungsansätze zur Auswahl einer optimalen v -elementigen Meßstellenteilmenge erfolgen. Die wichtigsten Aspekte sind dabei ihre Laufzeit in Abhängigkeit von der Meßstellenanzahl $|M|$, ihre Ergebnisqualität in bezug auf die Diskriminierungsleistung und ihr Verhalten bei redundanten, d. h. voneinander abhängigen Meßstellen.

Wie bereits zu den Verfahren angedeutet wurde, ist die einfache Meßstellenbewertung zwar mit einer linearen Laufzeit am schnellsten, sie kann aber nicht die optimale Meßstellenauswahl garantieren, weil die Einzelbewertung keine mehrwertigen Diskriminierungsinformationen berücksichtigt. Die beiden anderen Verfahren sind vollständig und finden das Optimum, haben jedoch eine exponentielle Laufzeit, da im schlechtesten Fall alle Meßstellenkombinationen bewertet werden. Da die Verfahren im Preprocessing angewandt werden, können sie für kleinere Werte von v aber praktikabel sein. Dies gilt insbesondere, wenn vorher eine Abhängigkeitsanalyse zur Entfernung redundanter Meßstellen durchgeführt wurde, die einen quadratischen Aufwand besitzt (vgl. Abschnitt 4.6.2).

Sind in der Menge M voneinander abhängige Meßstellen enthalten, so wirkt sich dies negativ auf das Ergebnis der einfachen Meßstellenbewertung aus: Weil die einzelnen Diskriminierungsleistungen der abhängigen Meßgeräte identisch sind, werden sie auch gleich bewertet, so daß es zur Auswahl redundanter Meßstellen kommen kann. Im Gegensatz dazu wird bei den beiden anderen Verfahren die gemeinsame Diskriminierungsleistung mehrerer Meßstellen bewertet; da sie durch redundante Meßgeräte nicht erhöht werden kann, ändert sich auch das Ergebnis der Auswahl nicht.

Im prototypischen Diagnosesystem ARGUS wurde dem Argument der kürzeren Laufzeit der Vorzug gegenüber der Exaktheit des Ergebnisses gegeben und das Verfahren

der Einfachbewertung mit vorheriger Abhängigkeitsanalyse realisiert. Es hat sich herausgestellt, daß die Diagnoseergebnisse auch dann sehr gut sind, wenn die Optimalität der Meßstellenauswahl nicht garantiert werden kann (vgl. Abschnitt 5.3.2).

4.7 Diagnoseregelerzeugung

Die Regelerzeugung stellt die eigentliche Data-Mining-Phase innerhalb des Diagnosesystems dar. Aus einer Datenbasis, die Intervallvektoren mit den Symptomen der ausgewählten Meßstellen sowie die zugehörigen Betriebssituationen und Diagnosen enthält, wird für jede Anlagenphase $p \in P$ eine separate Menge von Diagnoseregeln R_p erzeugt. Durch die phasenetrennte Regelerzeugung entfällt die Notwendigkeit, daß jede Regelprämisse die Phaseninformation enthält. Im folgenden wird die Erzeugung von Diagnoseregeln näher beschrieben.

Dazu werden zunächst in Abschnitt 4.7.1 einige Vereinbarungen für die Verwendung von klassifizierenden Assoziationsregeln zur Diagnose hydraulischer Anlagen getroffen. Wie bereits in Abschnitt 3.4 erläutert, ist es für die Erzeugung von Diagnoseregeln nicht sinnvoll, allgemeine Assoziationsregelalgorithmen zu verwenden, weil bei der Ermittlung σ -relevanter Tupelmengen sehr viele Kombinationen a priori ausgeschlossen werden können; aus diesem Grund wurden einige Erweiterungen zur Laufzeitverbesserung skizziert. In Abschnitt 4.7.2 erfolgt die Beschreibung eines Diagnoseregelalgorithmus, der die früheren Überlegungen umsetzt. Abschließend wird in Abschnitt 4.7.3 die Optimierung einer erzeugten Diagnoseregelmenge durch Auswahl der »besten« Regeln diskutiert.

4.7.1 Vereinbarungen

Durch die bisherigen Vorverarbeitungsschritte Symptomerkenkung, Diskretisierung und Meßstellenauswahl ist auf der Grundlage des simulierten Phasenprotokollmodells für jede Anlagenphase eine kategoriale Datenbasis D_n erzeugt worden. Jeder Datensatz $d \in D_n$ beschreibt die Auswirkungen eines Fehlers in einer bestimmten Betriebssituation der Anlage. Befinden sich im Schaltkreis c Zylinder, h Hydromotoren, u ausgewählte Druckmeßstellen und v ausgewählte Flußmeßstellen, ist ein Datensatz mit den Bezeichnern aus Abschnitt 4.3 üblicherweise wie folgt aufgebaut:

$$d = \langle \underbrace{F_1, \dots, F_c}_{\text{Betriebssituation}}, \underbrace{s_1, \dots, s_c, v_1, \dots, v_c, n_1, \dots, n_h, p_1, \dots, p_u, q_1, \dots, q_v}_{\text{Symptome}}, \underbrace{f, v_f}_{\text{Diagnose}} \rangle$$

wobei jedes Element der Betriebssituation und jedes Symptom ein Intervall repräsentiert.

Eine Regel, die in ihrer Prämisse keine Ausdrücke über Betriebssituationen, sondern nur über Symptome enthält, ist unabhängig von einer speziellen Situation anwendbar. Diese Information spielt jedoch zum Zeitpunkt der Regelerzeugung keine Rolle, so daß zur Vereinfachung nicht mehr explizit zwischen Betriebssituationen und Symptomen unterschieden wird und beide Teile in den Regelprämissen gleichberechtigt sind. Des weiteren kann die Fehlerart f als »Intervall« für alle Variationen v_f aufgefaßt werden, so daß nur f , nicht aber die Variation v_f als Diagnose zu berücksichtigen ist. Insgesamt wird also vereinheitlichend davon ausgegangen, daß d einen Intervallvektor für Symptome an den Meßstellen $m_i \in M$ und für eine Diagnose $f \in F$ repräsentiert.

Die folgenden beiden Definitionen wurden bereits früher in dieser Arbeit angeführt, müssen aber im Hinblick auf die Erzeugung von Diagnoseregeln modifiziert werden. Zunächst wird Definition 3.3 (Seite 43) zur Bestimmung von relativen Häufigkeiten von Tupelmengen in kategorialen Datenbasen erweitert, und anschließend wird Definition 3.6 (Seite 51) zu einem Subsumptionskriterium für Diagnoseregeln vereinfacht.

Definition 4.11 (Relative Häufigkeit einer Tupelmenge): Es sei D_n eine kategoriale Datenbasis mit $n \in \mathbb{N}$ Datensätzen und einer Attributmenge $A = \{A_1, \dots, A_m\}$ über den Wertebereichen $\text{dom}(A_i)$. Weiter sei $X = \{\langle A_{i_1}, v_{j_1} \rangle, \dots, \langle A_{i_p}, v_{j_p} \rangle\}$ mit $v_{j_x} \in \text{dom}(A_{i_x})$ eine Menge von Attribut / Wert-Tupeln. Dann heißt

$$h_n(X) = \frac{|\{d = \langle a_1, \dots, a_m \rangle \in D_n \mid \forall \langle A_i, v_j \rangle \in X : a_i = v_j\}|}{n}$$

relative Häufigkeit von X . □

Definition 4.12 (Subsumption von Diagnoseregeln): Es seien $r = S \rightarrow f$ und $r' = S' \rightarrow f$ zwei Diagnoseregeln für den gleichen Fehler $f \in F$. Dann wird r durch r' subsumiert ($r' \succ r$), wenn gilt:

$$S \supseteq S' \text{ und } c(r) \leq c(r'). \quad \square$$

Der im nächsten Abschnitt entwickelte Algorithmus 4.2 greift auf die beiden Definitionen zurück.

4.7.2 Algorithmus

Im Unterschied zum Basisverfahren zur Assoziationsregelerzeugung (siehe Abschnitt 3.2.3) weist Algorithmus 4.2 drei wesentliche Eigenschaften auf, die im folgenden näher erläutert werden:

Eingabe: Datenbasis D_n mit Intervallvektoren für Symptome an den Meßstellen $m_i \in M$ und für eine Diagnose aus der Fehlermenge F , Häufigkeitsschwelle σ , Konfidenzschwelle γ

Ausgabe: Diagnoseregelmengemenge $R(D_n, \sigma, \gamma)$

```

(1)  $R \leftarrow \emptyset, s \leftarrow 1$ 
(2)  $S \leftarrow \{ \langle m_a, i_b \rangle \mid i_b \in I_{m_a} \text{ und } h_n(\langle m_a, i_b \rangle) \geq \sigma \}$ 
(3)  $F_1 \leftarrow \{ \langle F, f \rangle \mid f \in F \text{ und } h_n(\langle F, f \rangle) \geq \sigma \}$ 
(4) while  $F_s \neq \emptyset$  do begin
(5)    $s \leftarrow s + 1$ 
(6)   for each  $t = \langle m_{a_1}, i_{b_1} \rangle, \dots, \langle m_{a_{s-2}}, i_{b_{s-2}} \rangle, \langle F, f \rangle \in F_{s-1}$  do
      begin
(7)     for each  $\langle m_a, i_b \rangle \in S \mid a > a_{s-1}$  do begin
(8)        $r \leftarrow (\langle m_{a_1}, i_{b_1} \rangle \wedge \dots \wedge \langle m_{a_{s-2}}, i_{b_{s-2}} \rangle \wedge \langle m_a, i_b \rangle \rightarrow f \mid c, s)$ 
(9)       if  $s(r) > \sigma$  then begin
(10)        if  $c(r) > \gamma$  and  $\nexists r' \in R : r' \succ r$  then  $R \leftarrow R \cup r$ 
(11)        if  $c(r) < 1.0$  then  $F_s \leftarrow F_s \cup \{t \cup \langle m_a, i_b \rangle\}$ 
(12)       end
(13)     end
(14)   end
(15) end
(16) return  $R$ 

```

Algorithmus 4.2: Erzeugung von σ -relevanten und γ -konfidenten Diagnoseregeln

1. Integrierte Ermittlung σ -signifikanter Tupelmengen und γ -konfidenter Regeln,
2. optimierte Erzeugung kategorialer Diagnoseregeln und
3. Vermeidung redundanter Regeln durch zusätzliche Anwendung eines Subsumptionstests gemäß Definition 4.12.

In Anlehnung an das Basisverfahren geht Algorithmus 4.2 stufenweise vor, wobei in jeder Stufe s die Diagnoseregeln mit $s - 1$ -elementiger Prämisse erzeugt werden. Zunächst werden in Zeile 2 mit S alle σ -relevanten Meßstelle / Symptom-Tupel ermittelt, die später zur Erweiterung von Tupelmengen dienen. Da jede Diagnoseregeln in der Konklusion das Fehlerattribut besitzen muß, bildet die Menge F_1 den Ausgangspunkt

für potentiell zulässige Diagnoseregeln (Zeile 3). Jedes Tupel aus F_1 wird in der zweiten Stufe um ein Symptom-Tupel aus S erweitert; das Ergebnis ist jeweils eine Regel mit einem Prämissenelement (Zeile 8). Anschließend wird eine so erzeugte Regel auf ihre relative Häufigkeit, Konfidenz und Subsumption getestet (Zeilen 9 bis 12), wobei gilt:

- Nicht σ -relevante Regeln bzw. Tupelmengen werden verworfen.
- γ -konfidente und nicht von der bisherigen Regelmenge R subsumierte Regeln werden zu R hinzugefügt.
- Regeln mit einer Konfidenz kleiner 1.0 können evtl. durch zusätzliche Symptom-Tupel aus S eine höhere Konfidenz bekommen und stellen somit die Erweiterungskandidaten F_2 für die nächste Stufe dar.

Ist F_2 nicht leer, werden in der folgenden Stufe die γ -konfidenten Regeln mit zwei Prämissenelementen sowie die nächstgrößeren σ -relevanten Kandidaten in F_3 erzeugt usw. Der Algorithmus stoppt, wenn aus der vorherigen Stufe kein Erweiterungskandidat vorhanden ist.

Die Optimierungen bestehen erstens darin, daß nur Tupelmengen erzeugt werden, die auch das Fehlerattribut F enthalten (Zeilen 3 und 6). Zweitens wird in Zeile 7 durch die Ausnutzung einer festen Attributreihenfolge auf einfache Weise sichergestellt, daß in einer Tupelmenge kein Attribut mehrfach vorkommen kann. Als letzte Optimierung ist die Erzeugung einer Regel aus einer Tupelmenge zu nennen (Zeile 8): Hier steht die einzige sinnvolle Regelstruktur Symptom \rightarrow Diagnose unmittelbar fest, so daß keine weiteren Möglichkeiten getestet werden müssen.

Trotz der Vereinfachungen bleibt das asymptotische Laufzeitverhalten der Regelerzeugung gegenüber den Basisalgorithmen bestehen und kann unter der Annahme einer konstanten Meßstellenanzahl als quasi-linear in der Anzahl der Datensätze bezeichnet werden (vgl. Abschnitt 3.2.3).

4.7.3 Regelmengenoptimierung

In Abschnitt 3.4.2 wurde eine generelle Vorgehensweise für ein Einsatz von Assoziationsregeln zur Lösung von Klassifikationsaufgaben beschrieben; sie besteht aus der Erzeugung aller klassifizierenden Assoziationsregeln und der Bildung eines geeigneten Klassifikators. Der Klassifikator ist ein Verfahren, das zur Ableitung von Klassenvorhersagen die erzeugte Regelmenge oder eine Teilmenge hiervon auswertet. Um gute Ergebnisse zu erzielen, muß eine Anpassung an den jeweiligen Problembereich erfolgen; hierzu ist eine Optimierung der Regelmenge und / oder der

Regelauswertung möglich. Die Regelmengenoptimierung ist eine Teilaufgabe der Lernphase und wird daher in diesem Abschnitt behandelt. Im Gegensatz dazu findet die optimierte Regelauswertung in der Diagnosephase statt, so daß hierauf in Abschnitt 4.8 näher eingegangen wird.

Die Bestimmung geeigneter Schwellenwerte für die relative Häufigkeit und die Konfidenz der Diagnoseregeln stellt bereits ein wichtiges Mittel zur Regelmengenoptimierung dar. In einigen Problembereichen können die Klassifikationsergebnisse durch eine Auswahl der »besten« Regeln zusätzlich verbessert werden, wenn diese Teilmenge die zugrundeliegenden Strukturen im Vergleich zur vollständigen Regelmenge mit weniger Unsicherheiten beschreibt. Eine Unsicherheit besteht immer dann, wenn für einen gegebenen Merkmalsvektor die Regelauswertung keine eindeutige Klassenzuordnung liefert.

Die im folgenden skizzierte Vorgehensweise verbessert die Regelmenge im Hinblick auf eine naive Regelauswertung⁹. Sie wählt in Anlehnung an das Verfahren »CBA-CB« (*Classification Based on Associations, Classifier Building*) von Liu et al. [LHM98] zu einer festen Regelreihenfolge sukzessive diejenigen Regeln aus, mit denen die Lerndaten am besten klassifiziert werden.

Es sei D_n die Lerndatenbasis und R die Menge aller klassifizierenden Assoziationsregeln mit Mindesthäufigkeit σ und Mindestkonfidenz γ . Dann kann eine optimierte Regelmenge in drei Schritten ermittelt werden:

1. Sortiere die Regelmenge R nach absteigender Konfidenz, innerhalb gleicher Konfidenzen nach absteigender relativer Häufigkeit und innerhalb gleicher relativer Häufigkeiten nach zunehmender Regellänge.
2. Analysiere die Regeln in der Reihenfolge aus Schritt 1: Ist die aktuell betrachtete Regel r_i für mindestens einen Datensatz in der (Rest-)Datenbasis D_n korrekt anwendbar, dann entferne alle Datensätze aus D_n , für die r_i anwendbar ist. Andernfalls entferne r_i aus R .
3. Ermittle innerhalb der aus Schritt 2 verbliebenen Regelmenge $R = \{r_1, \dots, r_c\}$ diejenige Regel r_i , für die der Klassifikationsfehler bei naiver Auswertung der Regeln r_1 bis r_i minimal ist und entferne die Regeln r_{i+1} bis r_c aus R . Liefere die Restmenge R als Ergebnis zurück.

Das Verfahren stellt durch die Sortierung der Regeln in Schritt 1 sicher, daß immer die höchstkonfidenten Regeln zur Anwendung kommen. Darüber hinaus werden Regeln

9 Mit einer »naiven Regelauswertung« soll ein Verfahren bezeichnet werden, das zu einem gegebenen Attributwertevektor in einer geordneten Regelliste die erste anwendbare Regel identifiziert und dessen Konklusion als Klassenvorhersage nutzt.

entfernt, die auf der Lernmenge keine zusätzliche Klassifikationsleistung bringen: In Schritt 2 werden diejenigen eliminiert, die nicht anwendbar waren, in Schritt 3 diejenigen, die den Klassifikationsfehler wieder ansteigen ließen. Insgesamt wird also ein Klassifikator erzeugt, der für die definierte Regelreihenfolge eine maximale Klassifikationsgüte liefert; es handelt sich jedoch nur um ein Näherungsverfahren, das keine über alle möglichen Regelreihenfolgen optimale Regelauswahl garantieren kann.

Für Problembereiche, die auf möglichst kleine Regelmengen sowie schnelle Regelauswertungen angewiesen sind (wie zum Beispiel einfache Klassifikationsaufgaben), ist die Regelmengenoptimierung eine sinnvolle Maßnahme zur Verbesserung des Klassifikators (vgl. die Untersuchung in [LHM98]). Die Diagnose hydraulischer Anlagen erfordert jedoch eine komplexere Hypothesengenerierung, die temporale Aspekte (zeitlich verteilte Symptome) berücksichtigen und jedem möglichen Fehler einen Konfidenzwert zuweisen sollte, da sich heuristische Zusammenhänge häufig nicht als eindeutige Klassifikationen darstellen lassen. Im prototypischen Diagnosesystem ARGUS wird daher die mit Algorithmus 4.2 erzeugte Regelmenge unverändert zur Hypothesengenerierung ausgewertet.

4.8 Hypothesengenerierung

Die Hypothesengenerierung ermittelt Hypothesen bezüglich derjenigen Komponentenfehler, die das beobachtete Fehlverhalten der zu diagnostizierenden hydraulischen Anlage erklären können. Zu einem gegebenen Symptomvektor wird durch die Auswertung der anlagenspezifischen Diagnoseregeln für jede ableitbare Diagnose ein Konfidenzwert aus dem Intervall $[0, 1]$ berechnet. Üblicherweise ist es bei Diagnoseaufgaben in komplexen technischen Bereichen nicht möglich, genau einen sicheren Fehler zu identifizieren. So werden statt des Idealfalls, daß genau ein Fehler die Konfidenz 1 und alle anderen die Konfidenz 0 besitzen, mehrere Alternativen mit positivem Konfidenzwert berechnet, von denen dann eine als Diagnose zu interpretieren ist.

Im folgenden werden die Aufgaben Regelauswertung, Konfidenzberechnung und Ableitung einer Diagnose behandelt. Zunächst geht Abschnitt 4.8.1 auf ein Rahmenkonzept zur Diagnosefindung sowie einige damit verbundene Problemstellungen ein, bevor zwei konkrete Lösungsmöglichkeiten für den wichtigsten Teilaspekt, nämlich die Verrechnung von Regelkonfidenzen, beschrieben werden. Mit der lokalen Konfidenzverrechnung aus dem medizinischen Diagnosesystem MYCIN wird ein bekannter Ansatz skizziert (Abschnitt 4.8.2). Dieser weist jedoch einige Nachteile auf, so daß anschließend ein neuer Ansatz zur globalen Verrechnung entwickelt wird (Abschnitt 4.8.3).

4.8.1 Rahmenkonzept und Problemfelder

In der Diagnosephase wird zunächst das im Störfall erhobene Meßwerteprotokoll so aufbereitet, daß es zur direkten Regelanwendung geeignet ist. Anschließend erfolgt die Hypothesengenerierung mit der phasengetrennten Anwendung der Diagnoseregeln sowie mit weiteren, phasenübergreifenden Konfidenzverrechnungsschritten. Insgesamt stellt die folgende Vorgehensweise innerhalb des besprochenen Diagnoseansatzes ein Rahmenkonzept zur Lösung einer Diagnoseaufgabe dar:

1. Bestimme aus dem Meßwerteprotokoll die Symptomvektoren für alle Anlagenphasen.
2. Verrechne innerhalb jeder Phase für jeden möglichen Fehler die Konfidenzen der anwendbaren Diagnoseregeln.
3. Verrechne für jeden möglichen Fehler die in Schritt 2 berechneten Konfidenzen über alle Phasen.
4. Bestimme einen Fehler als Diagnose.

Zu den einzelnen Schritten folgen weitere Erläuterungen.

Zu Schritt 1: Für die Aufbereitung des Meßwerteprotokolls werden die gleichen Verarbeitungsschritte wie in der Lernphase für die Simulationsdaten durchgeführt, also die Symptomerkennung (vgl. Abschnitt 4.4) und die Diskretisierung der Symptomwerte (Abschnitt 4.5). Im Unterschied zur Lernphase erfolgt bei der Diskretisierung keine Bildung *neuer* Intervalle, sondern nur die Zuordnung der Symptomwerte zu den *vorhandenen* Intervallen, die bei der Generierung des Diagnosesystems mit Algorithmus 4.1 ermittelt wurden. Das Ergebnis der Vorverarbeitung ist je ein Intervallvektor pro Anlagenphase.

Zu Schritt 2: Im ersten Verrechnungsschritt müssen alle durch die Regelkonfidenzen repräsentierten Unsicherheiten problemangemessen verarbeitet werden. Anders als in Abschnitt 4.7.3 (Regelmengenoptimierung) wird hierzu die reihenfolgeunabhängige Auswertung *aller* Diagnoseregeln, die für einen Symptomvektor anwendbar sind, untersucht. Existiert nur eine anwendbare Regel $r \in R$ für einen Fehler $f \in F$, dann ist es sinnvoll, dem Fehler die Regelkonfidenz zuzuordnen (also $c(f) := c(r)$), weil dieser Wert ungefähr der relativen Häufigkeit von f bei Vorliegen der beobachteten Symptome entspricht.¹⁰ Nicht eindeutig ist die Situation immer dann, wenn mehrere Regeln mit unterschiedlichen Konfidenzen auf denselben Fehler schließen lassen.

¹⁰ Begründung: Es sei $r = s_1, \dots, s_p \rightarrow f$ die einzige anwendbare Regel für den Symptomvektor s , die auf einen Fehler $f \in F$ schließen läßt. Dann entspricht innerhalb der Lerndatenbasis die Regelkonfidenz $c(r)$ exakt der relativen Häufigkeit von f bei Beobachtung der Symptomteilmenge

In diesen Fällen ist eine gemeinsame Verrechnung der Regelkonfidenzen notwendig, wie Beispiel 4.10 zeigt.

Beispiel 4.10: Es seien im Meßwerteprotokoll die Meßstellen m_1, \dots, m_6 berücksichtigt. Weiter sei in einer Diagnosesituation für die Anlagenphase j der Symptomvektor $s = \langle i_2, i_1, i_4, i_4, i_3, i_5 \rangle$ gegeben, so daß die folgenden Assoziationsregeln anwendbar sind:

$$\begin{array}{llll}
 r_4 : & m_1 = i_2 & \rightarrow & f_2 \\
 r_{10} : & m_4 = i_4 \wedge m_6 = i_5 & \rightarrow & f_2 \\
 r_5 : & m_5 = i_3 & \rightarrow & f_2 \\
 r_6 : & m_6 = i_5 & \rightarrow & f_2
 \end{array} \left. \begin{array}{l} c = 0.50, s = 0.04 \\ c = 0.50, s = 0.02 \\ c = 0.40, s = 0.06 \\ c = 0.40, s = 0.05 \end{array} \right\} \Rightarrow c_j(f_2)$$

$$\begin{array}{llll}
 r_7 : & m_2 = i_1 \wedge m_4 = i_4 & \rightarrow & f_6 \\
 r_{14} : & m_5 = i_3 & \rightarrow & f_6
 \end{array} \left. \begin{array}{l} c = 0.75, s = 0.04 \\ c = 0.60, s = 0.06 \end{array} \right\} \Rightarrow c_j(f_6)$$

Dann wird eine geeignete Verrechnung der Regelkonfidenzen zur Ableitung der Fehlerkonfidenzen $c_j(f_2)$ und $c_j(f_6)$ benötigt, um eine Diagnose bestimmen zu können. \square

Auf das Problem der phasengetrennten Kombination aller anwendbaren Regeln mit gleicher Konklusion wird in den Abschnitten 4.8.2 und 4.8.3 näher eingegangen. Wie in den meisten Diagnoseanwendungen sind auch im vorliegenden Problembereich die Voraussetzungen für den Einsatz von statistischen Verfahren des probabilistischen Schließens nicht erfüllt (vgl. [Pup91]). Daher werden in den genannten Abschnitten zwei Verfahren vorgestellt, die zwar nicht statistisch fundiert sind, aber dennoch im jeweiligen Einsatzgebiet die Unsicherheitsphänomene angemessen behandeln.

Zu Schritt 3: Da für einen Fehler $f \in F$ in jeder Anlagenphase j durch die Anwendung phasenspezifischer Diagnoseregeln ein unabhängiger Konfidenzwert $c_j(f)$ berechnet wird, können diese Konfidenzwerte unterschiedlich sein. Somit ist auch hier eine geeignete Verrechnung notwendig, um eine auf die gesamte Diagnosesituation bezogene Fehlerkonfidenz $c(f)$ zu bestimmen. Das prototypische Diagnosesystem ARGUS verwendet hierfür das Maximum der Fehlerkonfidenzen, da die Erfahrung gezeigt hat, daß üblicherweise für jeden Fehler (mindestens) eine Anlagenphase existiert, in der sich dieser Fehler besonders sicher gegenüber den anderen abgrenzen

$\{s_1, \dots, s_p\} \subseteq s$, d. h. es ist $c(r) = h_n(f \mid s_1, \dots, s_p)$. Weiter gilt:

$$h_n(f \mid s_1, \dots, s_p) \approx h_n(f \mid s),$$

weil nach Voraussetzung keine Regel $s' \rightarrow f$ mit $s' \subseteq s \setminus \{s_1, \dots, s_p\}$ anwendbar ist. Die Gleichheit gilt, falls bei der Regelerzeugung keine Regel mit Prämissenelementen aus s die Häufigkeits- und Konfidenzschwelle unterschritten hat.

läßt. Die Anlagenphasen sind untereinander ohne Prioritäten, so daß es zur Hypothesengenerierung unerheblich ist, welche Phase bestimmend für die Fehlerkonfidenz ist.

Zu Schritt 4: Da die Fehlerkonfidenzen aus den Regelkonfidenzen abgeleitet werden (Schritt 2), gelten bei der Festlegung einer Diagnose für den aktuellen Störfall die folgenden Beziehungen: Ist die Konfidenz $c(f)$ für einen Fehler $f \in F$ nahe bei 1, so deutet dies auf eine korrekte Diagnose hin; hingegen deutet ein Wert nahe bei 0 auf eine falsche Diagnose hin. Sind die Diagnoseregeln den Problembereich repräsentativ, dann ist die Wahrscheinlichkeit, daß es sich bei f um den tatsächlichen Fehler handelt, um so kleiner, je kleiner seine Konfidenz $c(f)$ ist. Die Diagnoseleistung wird also dann maximiert, wenn – unter der Annahme, daß nur einzelne Fehler vorkommen – der höchstkonfidente Fehler f^* als Schätzer für die Diagnose herangezogen wird.¹¹

Die Ermittlung einer Diagnose $f^* \in F$ aus den Einzelkonfidenzen der Fehler (Schritte 3 und 4) wird in Abbildung 4.13 zusammenfassend dargestellt.

	Fehler f_1	...	Fehler f_c	Diagnose f^*
Phase 1	$c_1(f_1)$...	$c_1(f_c)$	
\vdots	\vdots	\ddots	\vdots	
Phase p	$c_p(f_1)$...	$c_p(f_c)$	
	\Downarrow	...	\Downarrow	
Hypothesen	$c(f_1) = \max_j \{c_j(f_1)\}$...	$c(f_c) = \max_j \{c_j(f_c)\}$	$\Rightarrow f^* \in F \text{ mit } c(f^*) = \max_i \{c(f_i)\}$

Abbildung 4.13: Hypothesengenerierung und Ermittlung einer Diagnose f^* bei c möglichen Fehlern in p Anlagenphasen

4.8.2 Konfidenzverrechnung in MYCIN

Das medizinische Diagnosesystem MYCIN wurde in den siebziger Jahren zur regelbasierten Erkennung von bakteriellen Infektionskrankheiten des Blutes und der

¹¹ Sind Mehrfachfehler möglich, können neben der Konfidenz weitere Optimalitätseigenschaften einfließen, wie zum Beispiel die Größe der Diagnosenmenge; dies wird im vorliegenden Diagnoseansatz allerdings nicht berücksichtigt (vgl. Abschnitt 4.2.1).

Gehirnhaut entwickelt (siehe z. B. [Sho76], [BS84]); es war eines der ersten erfolgreichen Expertensysteme überhaupt. Das Expertenwissen wurde in MYCIN mit konjunktiven Regeln dargestellt. Zur Repräsentation von Unsicherheit diente ein Konzept, in dem jeder Regel r ein Wert aus dem Intervall $[-1, 1]$ als sogenannter *Sicherheitsfaktor* (engl. *Certainty Factor*) $CF(r)$ zugeordnet wurde. Dies geschah üblicherweise manuell auf der Grundlage des Wissens von Fachexperten. Der Wert 1 bedeutete »definitiv wahr«, der Wert -1 »definitiv falsch«, und der Wert 0 repräsentierte völlige Unwissenheit. Durch die Anwendung eines Verrechnungsschemas (s. u.) wurden Sicherheitsfaktoren für Hypothesen abgeleitet.

Die weiteren Ausführungen beziehen sich in Analogie zum Konfidenzkonzept nur auf Sicherheitsfaktoren aus dem Wertebereich $[0, 1]$, mit denen Aussagen *für*, nicht aber *gegen* eine Hypothese möglich sind. Da beide Konzepte auf unsicheren Aussagen basierende Schlußfolgerungen zulassen, soll die Frage diskutiert werden, ob die Anwendung des Verrechnungsschemas aus MYCIN auch im vorliegenden Diagnoseansatz sinnvoll ist. Es wird jedoch deutlich, daß den Interpretationen der beiden Unsicherheitsmaße unterschiedliche Voraussetzungen zugrunde liegen und daß die im nachfolgenden Abschnitt 4.8.3 vorgestellte Alternative für den in dieser Arbeit betrachteten Problembereich besser geeignet ist.

Im Unterschied zu den statistisch ermittelten Konfidenzen werden Sicherheitsfaktoren in MYCIN als Addition von subjektiven Vertrauens- und Zweifelgraden aufgefaßt, die sich auf menschliches Hintergrundwissen (manchmal auch mit *Evidenz* bezeichnet) gründen. Dieses Wissen kann sich im Laufe der Zeit verändern, zum Beispiel durch die Berücksichtigung neuer Fakten oder Regeln. Wird eine Hypothese durch mehrere Regeln gestützt, so leistet jede Regel mit positivem Sicherheitsfaktor einen Beitrag zum Abbau der Unsicherheit. Für eine Diagnosesituation bedeutet dies, daß jede anwendbare Regel $r \in R$ mit Konklusion $f \in F$ den resultierenden Sicherheitsfaktor $CF(f)$ erhöht. In Verbindung mit einem empirisch ermittelten Mindestwert von $CF(r) = 0.2$ ergibt sich im i -ten Schritt ein neuer Sicherheitsfaktor für den Fehler f durch die Formel

$$CF_{i+1}(f) = \begin{cases} CF_i(f), & \text{falls } CF(r) < 0.2 \\ CF_i(f) + (1 - CF_i(f)) \cdot CF(r) & \text{sonst.} \end{cases}$$

Der Anfangswert $CF_0(f)$ wird üblicherweise mit 0 initialisiert. Für n anwendbare Regeln muß dann die Verrechnung der Sicherheitsfaktoren n mal wiederholt werden, wobei das Ergebnis $CF(f) := CF_n(f)$ unabhängig von der Reihenfolge des Regeleinsatzes ist.

Eine wesentliche Eigenschaft der MYCIN-Formel ist, daß die Anzahl der Verrechnungsschritte einen großen Einfluß auf das Ergebnis hat. So kann durch die Akkumulation auch niedriger Regelsicherheitsfaktoren der Sicherheitsfaktor einer Hypothese

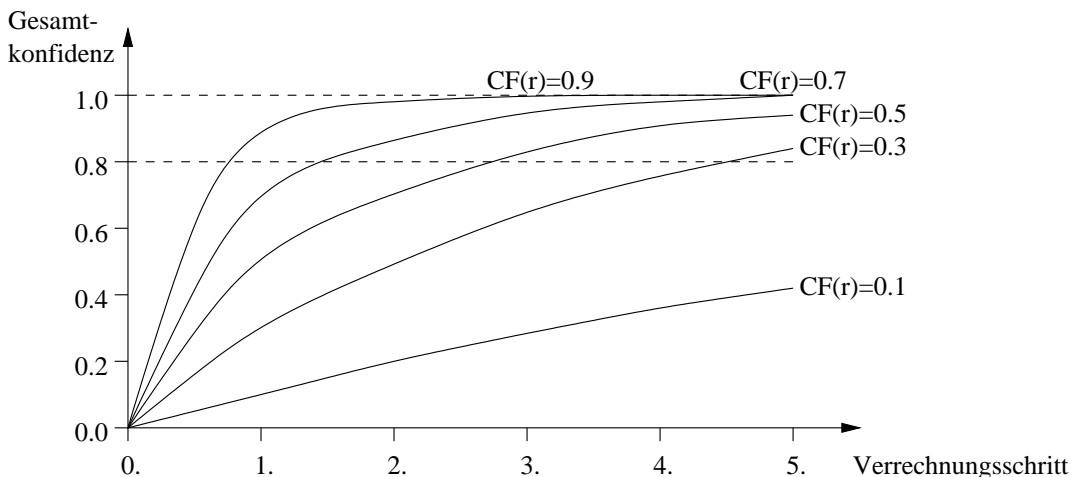
schnell gegen den Grenzwert 1 konvergieren. Beispiel 4.11 verdeutlicht dieses Verhalten.

Beispiel 4.11: (i) Für die Fehler f_2 und f_6 aus Beispiel 4.10 ergeben sich mit der MYCIN-Formel die Sicherheitsfaktoren

$$CF(f_2) = 0.91 \text{ und } CF(f_6) = 0.90,$$

obwohl alle Regelkonfidenzen zur Ableitung von f_2 deutlich kleiner sind als zur Ableitung von f_6 .

(ii) Die folgende Graphik zeigt in Anlehnung an [BS84] das Wachstum verschiedener Ergebnisse bei wiederholter Verrechnung von Regeln mit gleichem Sicherheitsfaktor. Man erkennt, daß nach nur fünf Schritten alle untersuchten Regeln mit $CF(r) > 0.3$ ein Ergebnis von über 0.8 liefern.



□

Die Frage, ob die Konfidenzverrechnung von MYCIN auch für den in dieser Arbeit besprochenen Diagnoseansatz geeignet ist, läßt sich durch die folgende Überlegung beantworten. Im MYCIN-Konzept wird jede Anwendung einer Regel mit positivem Sicherheitsfaktor als Vertrauenszuwachs bezüglich der Konklusion interpretiert, deshalb müssen zur Konfidenzverrechnung zwei Voraussetzungen erfüllt sein:

1. Die anwendbaren Regeln müssen voneinander unabhängig sein (vgl. [Pup91]) und
2. die Verrechnung muß auf wenige Regeln beschränkt sein.

Bei einer Verletzung der ersten Voraussetzung würde die errechnete Sicherheit von Diagnosen aufgrund identischer Begründungen – also ungerechtfertigt – zunehmen.

Diese Tatsache wird im vorliegenden Diagnoseansatz durch die Eliminierung abhängiger Meßstellen (vgl. Abschnitt 4.6.2) sowie der Nichterzeugung subsumierter Regeln (vgl. Abschnitt 4.7.2) berücksichtigt, so daß keine abhängigen Regeln existieren und Punkt 1 erfüllt ist.

Wäre die zweite Voraussetzung verletzt, so könnte der unerwünschte Fall aus Beispiel 4.11 (i) eintreten, in dem für viele niedrigkonfidente Regeln eine höhere Bewertung als für wenige hochkonfidente Regeln berechnet wird. Im Gegensatz zur ursprünglichen MYCIN-Anwendung, die relativ wenige, manuell implementierte Diagnoseregeln verarbeitet, werden im vorliegenden Diagnoseansatz die Assoziationsregeln automatisch und üblicherweise in großer Zahl generiert (zum Teil mehr als 4000, siehe Anhang B). Da sich Punkt 2 nicht erfüllen läßt, wird als Alternative im folgenden Abschnitt eine neue, für diesen Einsatzzweck besser geeignete Konfidenzverrechnung entwickelt.

4.8.3 Globaler Ansatz zur Konfidenzverrechnung

Der in diesem Abschnitt vorgestellte neue Ansatz nimmt eine globale Konfidenzverrechnung aller anwendbaren Regeln vor. Dabei wird die Basis für die Bewertung $c(f)$ eines Fehlers f durch die Konfidenz der höchstkonfidenten Regel mit Konklusion f gebildet. Es ist sinnvoll, eine Erhöhung dieses Werts in Abhängigkeit von zwei Einflußfaktoren vorzunehmen:

- *In Abhängigkeit vom Anteil der f vorhersagenden anwendbaren Regeln an allen anwendbaren Regeln:* Je mehr anwendbare Regeln die Konklusion f besitzen, desto höher sollte die Konfidenz von f sein.
- *In Abhängigkeit von der durchschnittlichen Konfidenz der f vorhersagenden anwendbaren Regeln:* Je höher die Konfidenzen der Regeln mit Konklusion f sind, desto höher sollte die Konfidenz von f sein.

Mit der folgenden Definition werden diese Überlegungen innerhalb eines formalen Rahmens integriert (vgl. auch [Ste01]).

Definition 4.13 (Konfidenz eines Fehlers): Es seien R die Menge der Diagnoseregeln einer Anlagenphase, $\langle a_1, \dots, a_m \rangle$ ein Symptomvektor und S eine Regelprämisse. Weiter sei $f \in F$ ein Fehler sowie

- $R^a := \{(S \rightarrow f) \in R \mid \forall (m_j = i) \in S : i = a_j\}$ die Menge der anwendbaren Regeln,
- $R_f^a := \{r \in R^a \mid r = S \rightarrow f\}$ die Menge der anwendbaren Regeln, die den Fehler f vorhersagen,

- $r^* \in R_f^a$ mit $\forall_{r \in R_f^a} : c(r) \leq c(r^*)$ eine Regel für Fehler f mit maximaler Konfidenz.

Dann berechnet sich die *Konfidenz des Fehlers* f durch

$$c(f) := c(r^*) + (1 - c(r^*)) \cdot \frac{1}{|R^a|} \sum_{r \in R_f^a} c(r). \quad \square$$

Anmerkung: $c(r^*)$ stellt die Basis für die Fehlerkonfidenz $c(f)$ dar. Die weiteren zwei oben erwähnten Einflußfaktoren (der Anteil der f vorhersagenden anwendbaren Regeln an allen anwendbaren Regeln $\frac{|R_f^a|}{|R^a|}$ und die durchschnittliche Konfidenz der f vorhersagenden anwendbaren Regeln $\frac{\sum_{r \in R_f^a} c(r)}{|R_f^a|}$) wurden multiplikativ miteinander verknüpft. $(1 - c(r^*))$ ist ein Normierungsfaktor, so daß gilt: $c(f) \in [0, 1]$.

Der folgende Satz zeigt die wichtigste Eigenschaft der globalen Konfidenzverrechnung: Die Konfidenz eines Fehlers nimmt nur bei Vorliegen einer *sicheren* Regel den Maximalwert 1 an.

Satz 4.1 (Maximum einer Fehlerkonfidenz): Nach dem Ansatz der globalen Konfidenzverrechnung aus Definition 4.13 kann die Konfidenz $c(f)$ eines Fehlers $f \in F$ nur dann den Wert 1 annehmen, wenn eine anwendbare Regel

$$r = S \rightarrow f \in R_f^a \text{ mit } c(r) = 1$$

existiert. □

Beweis: Zu zeigen ist, daß die Fehlerkonfidenz $c(f)$ kleiner 1 ist, falls keine Regelkonfidenz von 1 existiert. Angenommen, es ist zwar $c(f) = 1$, es existiert aber *keine* Regel $r \in R_f^a$ mit $c(r) = 1$. Dann gibt es eine Regel maximaler Konfidenz $r^* \in R_f^a$ mit $c(r^*) < 1$. Weiter gilt:

$$\begin{aligned} c(f) &= c(r^*) + (1 - c(r^*)) \cdot \frac{1}{|R^a|} \cdot \sum_{r \in R_f^a} c(r) = 1 \\ \Leftrightarrow & (1 - c(r^*)) \cdot \frac{1}{|R^a|} \cdot \sum_{r \in R_f^a} c(r) = 1 - c(r^*) \\ \Leftrightarrow & \frac{1}{|R^a|} \cdot \sum_{r \in R_f^a} c(r) = 1 \\ \Leftrightarrow & \sum_{r \in R_f^a} c(r) = |R^a| \end{aligned}$$

Zusammen mit den Randbedingungen $|R_f^a| \leq |R^a|$ und $\forall_{r \in R_f^a} : c(r) \leq 1$ folgt

$$|R_f^a| = |R^a| \text{ und } \forall_{r \in R_f^a} : c(r) = 1.$$

Dies steht jedoch im Widerspruch zur Annahme $c(r^*) < 1$. Somit gilt also die Behauptung $c(r^*) < 1 \Rightarrow c(f) < 1$. \square

Aus dem Beweis zu Satz 4.13 erkennt man weiter, daß der Wert $c(f) = 1$ nur dann erreicht wird, wenn *alle* anwendbaren Regeln *mit Sicherheit* auf den Fehler f schließen lassen. Bestimmend für die Konfidenz eines Fehlers ist also nicht die Anzahl der Konfidenzverrechnungen, sondern die Eindeutigkeit und die Güte der anwendbaren Assoziationsregeln. Das Beispiel 4.12 verdeutlicht diesen Effekt.

Beispiel 4.12: Die Konfidenzen der Fehler f_2 und f_6 aus Beispiel 4.10 nehmen in Abhängigkeit von der Anzahl der anwendbaren Regeln $|R^a|$ mit der globalen Konfidenzverrechnung nach Definition 4.13 Werte aus den folgenden Intervallen an:

$$CF(f_2) \in [0.50, 0.65] \text{ und } CF(f_6) \in [0.75, 0.81].$$

Die oberen Intervallgrenzen werden für $|R^a| = 6$, die unteren für $|R^a| = \infty$ erreicht. \square

Im folgenden werden die einzelnen Schritte des in diesem Kapitel vorgestellten Diagnoseansatzes im Zusammenhang evaluiert.

4 DIAGNOSE HYDRAULISCHER SYSTEME MIT ASSOZIATIONSREGELN

5 Evaluierung des Diagnoseansatzes

In diesem Kapitel wird der vorgestellte Diagnoseansatz experimentell bewertet. Als Evaluierungswerkzeug diente die Expertensystemshell ARGUS (Assoziationsregelgenerator für die heuristische Diagnose), mit deren Hilfe zahlreiche Testreihen durchgeführt wurden. ARGUS ist ein ebenfalls im Rahmen dieser Arbeit implementierter Demonstrationsprototyp, der es erlaubt, alle für die Vorverarbeitung der Simulationsprotokolle und für die Erzeugung der Assoziationsregeln wichtigen Parameter einzustellen sowie die Lernphase zu starten. Des weiteren ist mit dem System eine Simulation der Diagnosephase möglich, in der über benutzerdefinierte Skripte verschiedene Testreihen zur Ermittlung der Diagnosegüte durchgeführt werden können. Abbildung 5.1 zeigt die Bedienungsoberfläche von ARGUS unter Microsoft Windows 2000 sowie eine typische Regelmenge.¹

Für die Evaluierung eines Expertensystems oder eines Problemlösungsansatzes fehlen häufig geeignete Kriterien, so daß die Qualitätsbeurteilung eine schwierige Aufgabe darstellen kann (vgl. [GKP⁺83], [Pup91]). Die in den folgenden Abschnitten beschriebenen Tests beschränken sich daher auf wenige objektive und damit gut nachvollziehbare Kriterien zur Beurteilung der Leistungsfähigkeit und sind nicht auf spezielle Anwendungsbereiche zugeschnitten. Subjektive Kriterien, wie zum Beispiel Nützlichkeit für den Endbenutzer, Änderungsfreundlichkeit o. ä. werden nicht betrachtet.

Zur Ermittlung der objektiven Leistungsfähigkeit wurden mit ARGUS zahlreiche Diagnosesysteme für mehrere hydraulische Anlagen von mittlerer bis hoher Komplexität automatisch erzeugt und getestet. Obwohl die untersuchten Anlagen individuell unterschiedliche Verhaltensabfolgen besaßen, waren einheitliche Charakteristika der erzeugten Diagnosesysteme bzw. ihrer Testergebnisse identifizierbar. Stellvertretend werden sie im folgenden am Beispiel einer kleineren Testanlage dargestellt; alle konkreten Zahlenangaben dieses Kapitels beziehen sich auf diese Anlage.

Einteilung des Kapitels: Zunächst erläutert Abschnitt 5.1 den Aufbau der Testanlage und die in der Regellernphase gewählten Testparameter. Abschnitt 5.2 beschreibt daraufhin die generellen Testbedingungen in der Diagnosephase sowie die Berechnung der Diagnosegüte. Abschließend werden in Abschnitt 5.3 die Beurteilungskriterien erläutert sowie die wichtigsten Testergebnisse graphisch dargestellt.

¹ Die Evaluierung wurde unter Microsoft Windows 2000 vorgenommen, ARGUS ist aber auf jedem Betriebssystem lauffähig, das über die JAVA-Runtime-Umgebung JRE 1.3 oder höher verfügt.

5 EVALUIERUNG DES DIAGNOSEANSATZES

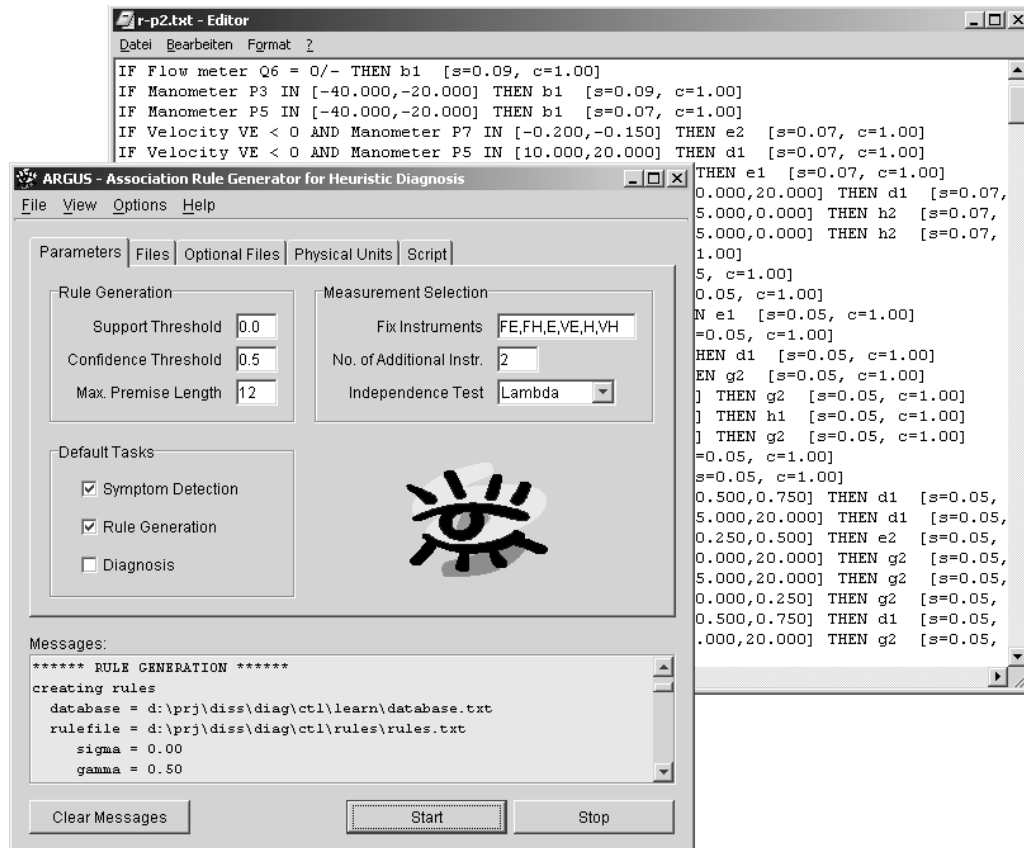


Abbildung 5.1: Oberfläche des Assoziationsregelgenerators ARGUS (vorne) und Ausschnitt einer erzeugten Regelmenge (hinten)

5.1 Testanlage und Parameter für die Lernphase

Abbildung 5.2 zeigt die Testanlage, deren Komponentenzahl mit neun Hauptkomponenten vergleichsweise niedrig ist. Durch ihre relativ komplexe Verschaltung ergibt sich jedoch ein nichttriviales Verhalten, das aus neun Phasen besteht.

Die Hauptkomponenten sind zwei Abtriebseinheiten (doppeltwirkende Hydrozylinder E und H), ein Steuerventil D, ein druckgesteuertes Abschalt- / Bremsenventil G, ein Rückschlagventil F, ein Absperrventil C, ein Druckbegrenzungsventil B, eine Pumpe A sowie ein Tank I. Da teilweise auch mehrere Fehlerarten für eine Komponente berücksichtigt werden konnten (z.B. innere und äußere Leckagen an den Zylindern), wurden im Simulationswerkzeug *art deco* insgesamt elf Fehlerklassen modelliert. Im weiteren werden die wichtigsten in der Lernphase verwendeten Parameter

5.1 TESTANLAGE UND PARAMETER FÜR DIE LERNPHASE

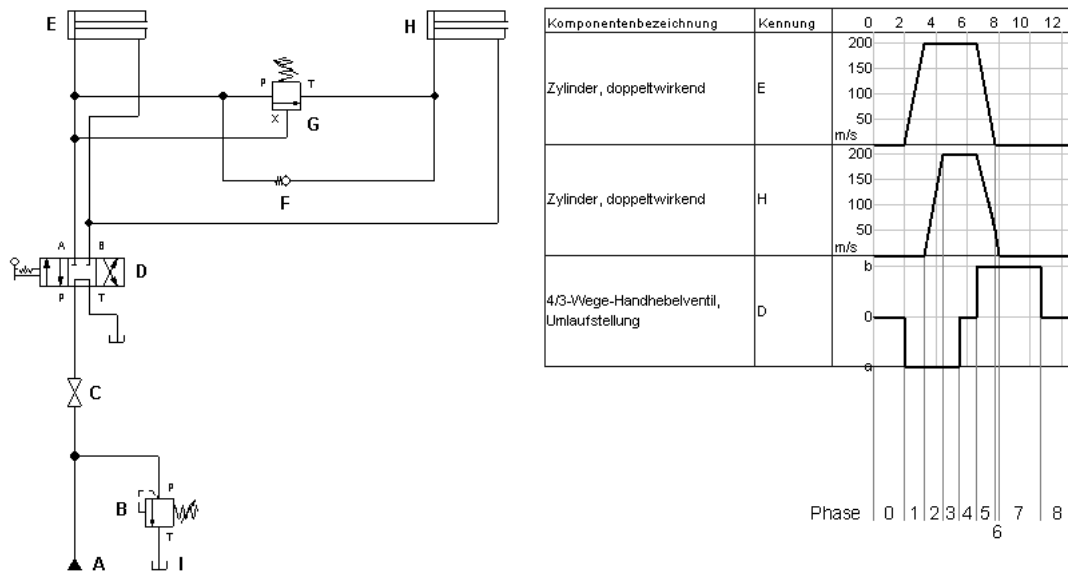


Abbildung 5.2: Schaltkreis (links) und Phasen (rechts) einer zu diagnostizierenden hydraulischen Anlage, dargestellt im Simulationswerkzeug *artdeco* (Phasenkennzeichnung und Phasennummerierung nachträglich hinzugefügt)

aufgeführt.

Simulation: Zur Aufstellung eines Meßplans erfolgte innerhalb des *artdeco*-Schaltplans die Plazierung von 38 Druck- und Flußmeßgeräten sowie für jeden der zwei Zylinder ein Meßgerät für die Zylinderposition und die anliegende Kraft. Die zulässigen Belastungen der Zylinder (Kräfte F_E und F_H) betrugen für diese Anlage konstruktionsbedingt zwischen 0 N und 900 N. Im Simulationsplan wurden für F_E/F_H die vier Belastungskombinationen 0/0, 100/100, 100/500, 900/900, elf Fehlerklassen sowie 20 leichte und 30 mittlere bis schwere Fehlervariationen definiert. Damit ergab sich ein Simulationsaufwand von

$$S = (1 + 11 + 20 + 30) \cdot 4 = 248$$

Simulationsläufen, so daß ein vollständiges Phasenprotokoll bei je neun Phasen $9 \cdot S = 2232$ Meßwertvektoren umfaßte (vgl. Abschnitt 4.3.2).

Symptomerkennung: Die Berechnung der Differenzenvektoren wurde nach Definition 4.2 (Seite 80) vorgenommen, welche die unterschiedlichen Eigenschaften der skalaren und vektoriellen physikalischen Größen berücksichtigt. Gleichzeitig erfolgte der Aufbau von drei Datenbasen mit den Symptomvektoren für unterschiedliche Feh-

lervariationen: D_{lern} für die Ausgangsfehler, D_{var_1} für die leichten Fehlervariationen und D_{var_2} für die mittleren bis schweren Variationen (vgl. Abschnitt 5.2).

Diskretisierung: Bei der Partitionierung der Wertebereiche wurden fünf Meßgeräteklassen gebildet und jeder Klasse eine maximale Intervallanzahl i_{max} zugewiesen, wobei die Drehzahlmesser für diese Testanlage nicht relevant waren. Die Intervallanzahlen wurden so gewählt, daß auch bei »manueller« Ablesung eine Unterscheidbarkeit der Intervalle gewährleistet ist (Abbildung 5.3).

Physikalische Größe	Meßgeräteart	Maximale Intervallanzahl i_{max}
Volumenstrom q_V	Flußmesser	5
Druck p	Druckmesser	10
Umdrehungszahl n	Drehzahlmesser	5
Geschwindigkeit v	(keine)	2
Kraft F	Kraftmesser	3

Abbildung 5.3: Maximale Intervallanzahlen für die verschiedenen Meßgerätearten (Anmerkung: Die Geschwindigkeiten wurden berechnet, vgl. Abschnitt 4.3.1)

Meßstellenauswahl: Die Optimierung der auszuwählenden Meßstellenkombination erfolgte ausschließlich auf der Menge der Fluß- und Druckmeßstellen; die Meßstellen für alle anderen physikalischen Größen wurden immer automatisch ausgewählt, da hier vernachlässigbare Meßkosten angenommen werden konnten (vgl. Abschnitt 4.6.1). Eine Abhängigkeitsanalyse mit asymmetrischen λ -Maßen (siehe Definition 4.7) hatte zum Ergebnis, daß etwa die Hälfte der 38 Druck- und Flußmeßstellen wegen Redundanz entfernt werden konnte. Anschließend wurden innerhalb der Testreihen von den verbliebenen – also untereinander unabhängigen – Meßstellen nach dem Ansatz der Einfachbewertung (siehe Abschnitt 4.6.3) über den Parameter v bis zu sechs Meßstellen ausgewählt.

Regelerzeugung: Zur Erzeugung der Assoziationsregelmengen wurden verschiedene Parameterkombinationen untersucht. Dabei blieb jedoch die Häufigkeitsschwelle konstant bei $\sigma = 0$, damit auch seltene Diagnosesituationen vom System erkennbar waren. Zur Ermittlung der objektiven Problemlösungsfähigkeit des Diagnoseansatzes kamen Konfidenzschwellen γ aus dem Intervall $[0.1, 1.0]$ zur Anwendung. Auf den Einsatz von $\gamma = 0.0$ wurde im Hinblick auf die allzu geringe Regelgüte verzichtet.

5.2 Testverfahren für die Diagnosephase

Für die Experimente standen keine realen hydraulischen Anlagen zur Verfügung; aus diesem Grund dienten die ^{art}*deco*-Simulationsprotokolle nicht nur zur Regelerzeugung, sondern auch als Datengrundlage für die Testreihen zur Ermittlung der Diagnosegüte. Die Simulationsprotokolle wurden hierzu in mehrere Datenbasen aufgeteilt, z. B. für die Testanlage aus Abbildung 5.2 in

- D_{lern} mit den Simulationsprotokollen für die elf Ursprungsfehler,
- D_{var_1} mit den Simulationsprotokollen für die 20 leichten Fehlervariationen und
- D_{var_2} mit den Simulationsprotokollen für die 30 mittleren bis schweren Fehlervariationen.

Während in der Lernphase ausschließlich die Datenbasis D_{lern} ausgewertet wurde, konnte in den Testreihen für die Diagnosephase auf jeweils eine der drei erzeugten Datenbasen zugegriffen werden. Zur Nachbildung eines realen zu diagnostizierenden Meßwerteprotokolls wurden aus den Symptomvektoren innerhalb der Datenbasen jedoch nur diejenigen Elemente interpretiert, die die entsprechende reale Anlage hätte liefern können. Jeder zu einer Betriebssituation gehörende phasenvollständige Satz von Symptomvektoren entsprach dabei genau einem Störfall.

Abhängig von der Wahl der zu diagnostizierenden Datenbasis sind zwei verschiedene Qualitätsaspekte des Diagnoseansatzes unterscheidbar:

1. *Eignung der Modelltransformation für Diagnoseprobleme*: Durch die Anwendung des Diagnosesystems auf das gelernte Phasenprotokoll D_{lern} kann überprüft werden, ob die Transformation der numerischen Verhaltensmodelle in heuristische Assoziationsregelmodelle hinreichend informationserhaltend ist. Dies ist der Fall, wenn alle gelernten Fälle korrekt diagnostiziert werden.
2. *Anwendbarkeit auf neue Diagnosesituationen*: Wird zum Testen eine andere Datenbasis als zum Lernen eingesetzt (also D_{var_1} oder D_{var_2} statt D_{lern}), findet die Anwendung der Diagnoseregeln auf bisher nicht analysierte Störfälle der Anlage statt. Mit der hier erzielbaren Diagnosegüte läßt sich die Generalisierungsfähigkeit der Regeln bzw. die Flexibilität des Diagnoseansatzes bewerten. Ist die Diagnosegüte hoch, steigt die Wahrscheinlichkeit, daß auch unvorhergesehene Fehlersituationen vom Diagnosesystem korrekt erkannt werden können.

Anmerkung: Zwar ist es ein Merkmal des Diagnoseansatzes, die Diagnoseregeln aus möglichst *umfassenden* Fehlersimulationen zu lernen, so daß die Wahrscheinlichkeit für das Auftreten eines vollkommen neuen Störfalles sehr gering ist. Generell

stellt aber die Diagnosegüte, die bei der Anwendung eines Diagnosesystems auf neue Fälle erzielbar ist, eine wichtige Information zu dessen Beurteilung dar. Aus diesem Grund wird der zweite Qualitätsaspekt bei der Evaluierung berücksichtigt.

Für die Diagnose eines Störfalls wurde die in Abschnitt 4.8.3 erläuterte globale Konfidenzverrechnung (Definition 4.13) angewandt. Die anschließende Beurteilung der Diagnosegüte basierte nicht nur auf dem Vergleich der höchstkonfidenten Hypothese f^* mit der »wahren« Diagnose d , sondern differenzierte im Erfolgsfall zusätzlich danach, ob für weitere Hypothesen (höchstens 3) eine ähnlich hohe Konfidenz berechnet wurde (maximale Abweichung 0.1). Mit diesen im Hinblick auf die Testanlage als sinnvoll erachteten Grenzwerten konnte zwischen falschen sowie korrekten Einzeldiagnosen und korrekten Diagnosenmengen unterschieden werden, wobei den Begriffen die Kriterien aus Definition 5.1 zugrunde lagen.

Definition 5.1 (Klassifizierung von Diagnosen): Es sei F die Menge der Fehler und $f^* \in F$ mit $c(f^*) = \max_{f \in F} \{c(f)\}$ eine Hypothese mit maximaler Konfidenz für den aktuellen Störfall. Weiter sei $d \in F$ für diesen Fall die wahre Störungsursache. Dann stellt f^* eine Diagnose aus genau einer der folgenden Klassen dar.

- *Korrekte Einzeldiagnose:* f^* ist die korrekte Diagnose d , und die Konfidenzen aller anderen Hypothesen aus der Fehlermenge F sind mindestens 0.1 niedriger als $c(f^*)$:

$$f^* = d \text{ und } |\{f \in F \setminus \{f^*\} \mid c(f^*) - c(f) < 0.1\}| = 0.$$

- *Korrekte Diagnosenmenge:* f^* ist die korrekte Diagnose d , es existieren aber bis zu drei weitere Hypothesen mit weniger als 0.1 niedrigerer Konfidenz als $c(f^*)$:

$$f^* = d \text{ und } |\{f \in F \setminus \{f^*\} \mid c(f^*) - c(f) < 0.1\}| \in [1, 2, 3].$$

- *Fehldiagnose:* Alle übrigen Fälle, d.h., es liegt eine falsche Diagnose oder eine zu hohe Mehrdeutigkeit vor (mehr als drei Hypothesen, deren Konfidenz weniger als 0.1 von $c(f^*)$ abweicht):

$$f^* \neq d \text{ oder } |\{f \in F \setminus \{f^*\} \mid c(f^*) - c(f) < 0.1\}| > 3.$$

f^* heißt *korrekte Diagnose*, wenn sie eine korrekte Einzeldiagnose ist oder in einer korrekten Diagnosenmenge ist, f^* heißt *falsche Diagnose*, wenn sie eine Fehldiagnose ist. \square

Die Bewertung der Diagnosegüte $d^* \in [0, 1]$ für eine Testreihe ist damit zweigeteilt und besteht aus der Summe des Anteils d_1 der korrekten Einzeldiagnosen und des

Anteils d_2 der korrekten Diagnosenmengen in den Testfällen. Der Anteil der Fehldiagnosen beträgt dann $1 - d^* = 1 - d_1 - d_2$.

Im praktischen Einsatz realer hydraulischer Anlagen kommen einzelne Störungen häufiger vor als andere, oder sie verursachen unterschiedliche Kosten. Da jedoch bei den Testreihen keine zusätzlichen Informationen über die A-priori-Wahrscheinlichkeiten und Schweregrade der Fehler vorlagen, wurden für die Ermittlung der Diagnosegüte d^* alle Störungsfälle gleichgewichtet, d. h., es wurde auf eine Gewichtung von Fehlern oder die Unterscheidung zwischen einer Fehldiagnose und einer Nichterkennung des Fehlers verzichtet.

5.3 Testergebnisse

In diesem Abschnitt werden einige Aspekte bezüglich der Regelerzeugung und der in den Testreihen erzielten Diagnoseergebnisse präsentiert. Als wichtigste Einflußgrößen werden die Konfidenzschwelle γ (Maß für die Regelgüte) und die Anzahl zusätzlicher Meßstellen ν (Maß für die Regelspezialisierung) untersucht. In bezug auf die Testanlage aus Abbildung 5.2 werden für verschiedene Parametervariationen die grundlegenden Trends bei der Entwicklung der Regelanzahl (Abschnitt 5.3.1) und der Entwicklung der Diagnosegüte (Abschnitt 5.3.2) graphisch aufbereitet und diskutiert. Abschnitt 5.3.3 gibt eine kurze Zusammenfassung der Evaluierungsergebnisse; die vollständigen Ergebnistabellen befinden sich im Anhang B.

5.3.1 Regelanzahl

Das erste objektive Kriterium zur Beurteilung des Diagnoseansatzes ist die Anzahl der erzeugten Assoziationsregeln $|R|$. Hier wird zunächst der Einfluß des in Abschnitt 3.2.3 besprochenen Subsumptionstests untersucht, bevor auf die Auswirkungen der Parameter ν und γ näher eingegangen wird. Nicht betrachtet werden die zur Regelerzeugung mit dem Data-Mining-Algorithmus 4.2 benötigten Zeiten, weil sie in jedem Fall wesentlich geringer sind als die Simulationszeiten.

Abhängigkeit vom Subsumptionstest

In Abbildung 5.4 ist für einige Parameterkombinationen ν / γ aufgeführt, wie sich die Anzahl der Diagnoseregeln entwickelt, wenn redundante Regeln eliminiert oder beibehalten werden. Dabei ist R^* die Menge der ohne Subsumptionstest erzeugten Diagnoseregeln und R diejenige Menge, aus der gemäß Definition 4.12 von Seite 107

subsumierte Regeln entfernt wurden. Es zeigt sich anhand der letzten Zeile, daß die Subsumptionsrate (d. h. der Anteil der subsumierten Regeln) mit regelmäßig über 90% sehr hoch ist.

ν	1	1	1	1	1	2	3	4	5	6
γ	1.0	0.7	0.5	0.3	0.1	0.5	0.5	0.5	0.5	0.5
$ R $	161	170	349	463	1254	518	728	922	1164	1322
$ R^* $	1697	1748	4522	6002	12713	9167	19531	38392	65471	125416
$1 - \frac{ R }{ R^* }$	0.91	0.90	0.92	0.92	0.90	0.94	0.96	0.98	0.98	0.99

Abbildung 5.4: Anteil subsumierter Regeln für verschiedene Kombinationen von Meßstellenanzahl ν und Konfidenzschwelle γ (R : Menge der nicht-subsumierten Regeln, R^* : Menge aller Regeln)

In der linken Tabellenhälfte von Abbildung 5.4 wird die Subsumptionsrate in Abhängigkeit von der Konfidenzschwelle γ dargestellt, in der rechten in Abhängigkeit von der Meßgeräteanzahl ν . Während die Konfidenzschwelle nur wenig Einfluß hat, führt eine höhere Anzahl zusätzlicher Meßgeräte auch zu deutlich höheren Subsumptionsraten. Durch die letztgenannte Abhängigkeit wird sogar das exponentielle Wachstum der Regelanzahl (Verdopplung mit jedem neuen Meßgerät) auf ein lineares Wachstum (jeweils Zunahme um ca. 200 Regeln) verringert.

Die Gründe für dieses Verhalten sind die folgenden. Zunächst sei daran erinnert, daß für zwei Regeln der Subsumptionstest in Definition 4.12 aus einem Strukturkriterium (Prämisse ist Teilmenge, Konklusion ist gleich) und einem Konfidenzkriterium (Konfidenz der spezielleren Regel nicht größer) besteht, wobei das Strukturkriterium die wichtigere, weil restriktivere Bedingung für eine Subsumptionskonstellation ist.

Wird die Konfidenzschwelle gesenkt, kommen neue, minderkonfidente Regeln zur Regelmenge hinzu, deren struktureller Aufbau unabhängig von den bisherigen Regeln ist. Daher wird die Erfülltheit des Strukturkriteriums und damit der Subsumptionsbedingung bei den neuen Regeln nicht wahrscheinlicher als vorher, und die Subsumptionsrate bleibt in etwa konstant. Anders ist die Situation bei einer Erhöhung der Meßgeräteanzahl. Hier entstehen neue Regeln, indem alte Regeln stufenweise um zusätzliche Prämissenelemente erweitert werden (vgl. Algorithmus 4.2). Durch diese Vorgehensweise bei der Regelkonstruktion ist das Strukturkriterium sehr viel wahrscheinlicher erfüllt, so daß die Subsumptionsrate stark zunimmt.

Im weiteren wird der Einfluß von Meßstellenanzahl und Konfidenzschwelle auf die

Diagnoseregelanzahl näher betrachtet. Dabei beziehen sich alle Ausführungen nur noch auf Regelmengen, aus denen subsumierte Regeln entfernt wurden.

Abhängigkeit von der Meßstellenanzahl

In Abbildung 5.5 wird die Entwicklung der Regelanzahl bei einer Erhöhung der Meßstellenanzahl v für zwei konstante Konfidenzschwellen $\gamma = 1.0$ bzw. $\gamma = 0.5$ graphisch dargestellt. Es läßt sich der bereits im letzten Unterabschnitt erläuterte lineare Zusammenhang zwischen $|R|$ und v erkennen.

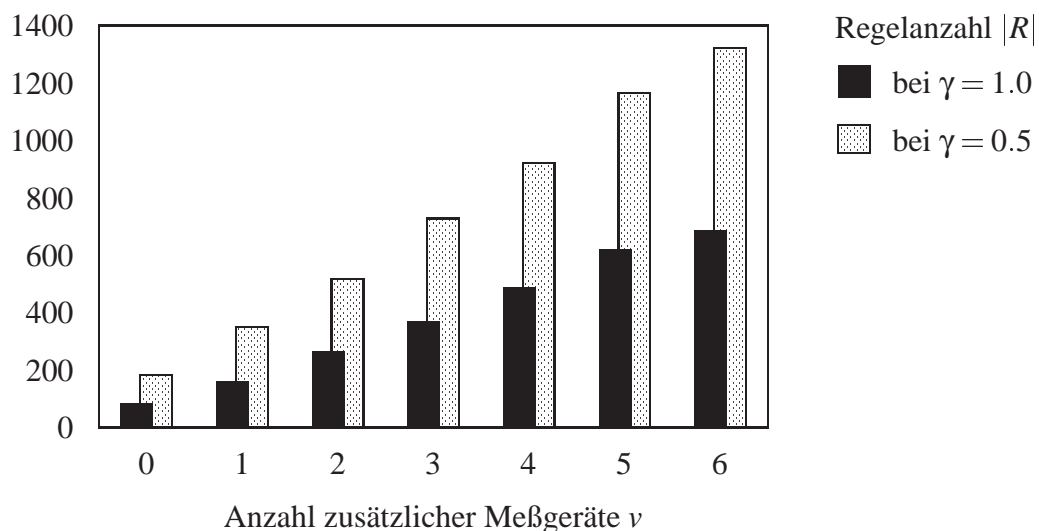


Abbildung 5.5: Entwicklung der Regelanzahl $|R|$ in Abhängigkeit von der Meßgeräteeanzahl v (bei konstanten Konfidenzschwellen γ)

Die absolute Anzahl der erzeugten sicheren Regeln (d. h. $\gamma = 1.0$) steigt von 82 bei keiner zusätzlichen Meßstelle bis auf 685 bei sechs Meßstellen an. Durch eine Halbierung der Konfidenzschwelle verdoppelt sich die jeweilige Regelanzahl. Da bei der Regelerzeugung in jedem Fall fixe Meßstellen berücksichtigt werden (vgl. Abschnitt 4.6.1), ist die Regelmenge auch bei $v = 0$ nicht leer.

Abhängigkeit von der Konfidenzschwelle

In Abbildung 5.6 wird die Entwicklung der Regelanzahl bei einer Erhöhung der Konfidenzschwelle γ für zwei konstante Meßstellenanzahlen $v = 0$ bzw. $v = 1$ graphisch dargestellt. Hier läßt sich ein exponentieller Zusammenhang zwischen $|R|$ und γ erkennen.

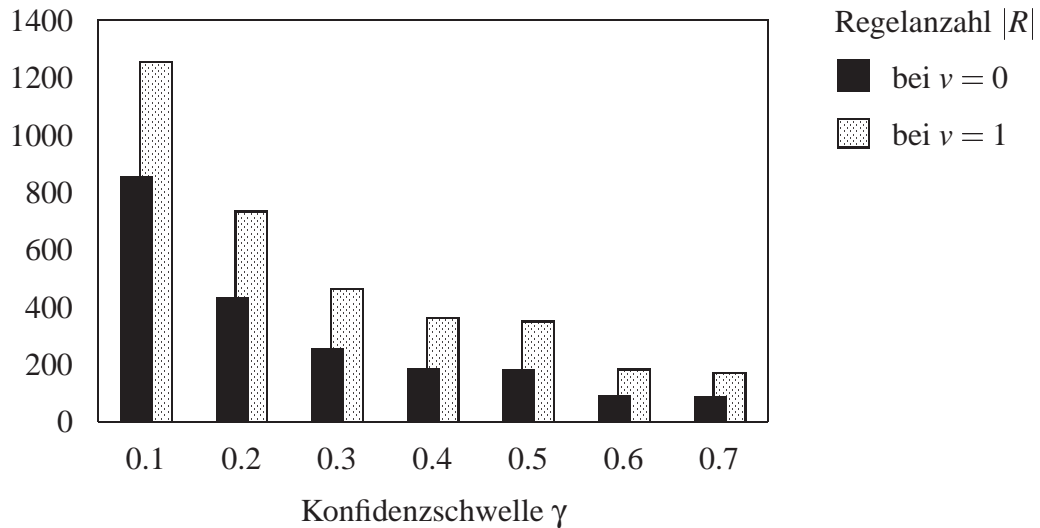


Abbildung 5.6: Entwicklung der Regelanzahl $|R|$ in Abhängigkeit von der Konfidenzschwelle γ (bei konstanten Anzahlen zusätzlicher Meßgeräte v)

So nimmt z. B. für $v = 1$ beginnend bei $\gamma = 0.1$ durch eine Erhöhung der Konfidenzschwelle die Regelanzahl von 1254 stark ab, ehe sie bei etwa $\gamma = 0.6$ mit unter 200 zu stagnieren beginnt. In weiten Bereichen gilt die »Daumenregel«, daß eine Verdoppelung der Konfidenzschwelle zu einer Halbierung der Regelanzahl führt.

Im Gegensatz zur Meßstellenanzahl, die nur linearen Einfluß hat und die voraussichtlich von außen vorgegeben wird, um den jeweiligen realen technischen Randbedingungen zu genügen, hat die Wahl der geeigneten Konfidenzschwelle für das zu erzeugende Diagnosesystem größere Auswirkungen. Zwar ist die Größe der Regelbasis nicht der entscheidende Qualitätsaspekt für ein Diagnosesystem, dennoch sind möglichst kleine Regelmengen anzustreben, um sowohl die schnelle Verarbeitung in der Diagnosephase als auch die Erklärbarkeit der getroffenen Diagnoseentscheidung zu gewährleisten.

Bei der Ermittlung der geeigneten Parameterkombination für die Generierung eines neuen Diagnosesystems ist jedoch insbesondere die Diagnosegüte ein wichtigeres Qualitätskriterium; sie wird im folgenden Abschnitt untersucht.

5.3.2 Diagnosegüte

Auch für die Entwicklung der Diagnosegüte wird nach variabler Meßstellenanzahl v (Veränderung der Regelspezialisierung) und variabler Konfidenzschwelle γ (Verände-

rung der Regelgüte) unterschieden. Die Berechnung der Diagnosegüte erfolgt gemäß Definition 5.1, also mit einer Differenzierung der korrekten Diagnosen in korrekte Einzeldiagnosen und korrekte Diagnosenmengen sowie mit einer Gleichgewichtung der Störungsursachen.

In den folgenden Unterabschnitten wird die Anwendung der Diagnoseregeln auf die Lernbasis D_{lern} nicht näher ausgeführt, weil für bereits kleine Meßstellenanzahlen bzw. hohe Konfidenzschwellen die maximale Diagnosegüte von $d^* = 1.0$ erzielt wird (z. B. für $\nu = 0/\gamma = 0.1$, $\nu = 1/\gamma = 0.5$ oder $\nu = 3/\gamma = 1.0$, vgl. Abbildung B.1, Seite 145). Ab zwei Meßstellen ist auch der Anteil der korrekten Einzeldiagnosen d_1 maximal; dies stellt das optimale Ergebnis dar. Hierdurch wird deutlich, daß der vorgestellte Wissensakquisitionsansatz zur Gewinnung heuristischer Diagnoseregeln geeignet und praktikabel ist.

Interessant sind auch die Auswirkungen der Parameterwahl auf die Generalisierungseigenschaften der Regeln; sie zeigen sich bei der Anwendung der Regeln auf *neue* Diagnosesituationen. Aus diesem Grund erfolgt im weiteren eine Analyse der Regelanwendungen auf die Datenbasen D_{var_1} und D_{var_2} mit Meßwerteprotokollen für einfache bzw. mittlere bis schwere Fehlervariationen.

Abhängigkeit von der Meßstellenanzahl

Abbildung 5.7 zeigt in vier Diagrammen die Entwicklung der Diagnosegüte bei einer Erhöhung der Meßstellenanzahl ν , getrennt nach den konstanten Konfidenzschwellen $\gamma = 1.0$ und $\gamma = 0.1$ sowie nach den beiden Datenbasen D_{var_1} und D_{var_2} . Es läßt sich erkennen, daß zwar die Gesamtdiagnosegüte d^* gegen 1 konvergiert, der Anteil der korrekten Einzeldiagnosen d_1 aber etwa gleich bleibt bzw. leicht abnimmt.

Die beiden linken Diagramme zeigen für die leichten Fehlervariationen D_{var_1} einen Anteil korrekter Einzeldiagnosen von etwa $d_1 = 0.4$, die beiden rechten Diagramme für die mittleren bis schweren Variationen D_{var_2} einen etwas niedrigeren Wert von ungefähr $d_1 = 0.3$. Mit zunehmender Meßstellenanzahl ν nimmt d_1 nach einem Maximum bei $\nu = 1$ leicht ab, dagegen nimmt der Anteil korrekter Diagnosenmengen $d_2 = d^* - d_1$ deutlich zu: Zum Beispiel steigt bei $\gamma = 1.0$ für D_{var_1} bzw. D_{var_2} der Wert von $d_2 = 0.04$ bzw. $d_2 = 0.08$ auf den Wert $d_2 = 0.6$ bzw. $d_2 = 0.7$; bei $\gamma = 0.1$ ist lediglich das Anfangsniveau höher.

Diese Ergebnisse sind wie folgt zu begründen. Durch eine höhere Meßstellenanzahl nimmt der Spezialisierungsgrad der neuen Regeln stetig zu, weil bei ihrer Konstruktion immer mehr Prämissenelemente eingesetzt werden. Je spezieller eine Regel ist, desto besser beschreibt sie eine Ausnahmesituation in der Lernbasis. Auf neue, zu einem gewissen Grad abweichende Situationen aus D_{var_1} oder D_{var_2} sind diese Regeln jedoch nicht anwendbar, oder aber sie bewirken mit höherer Wahrscheinlichkeit

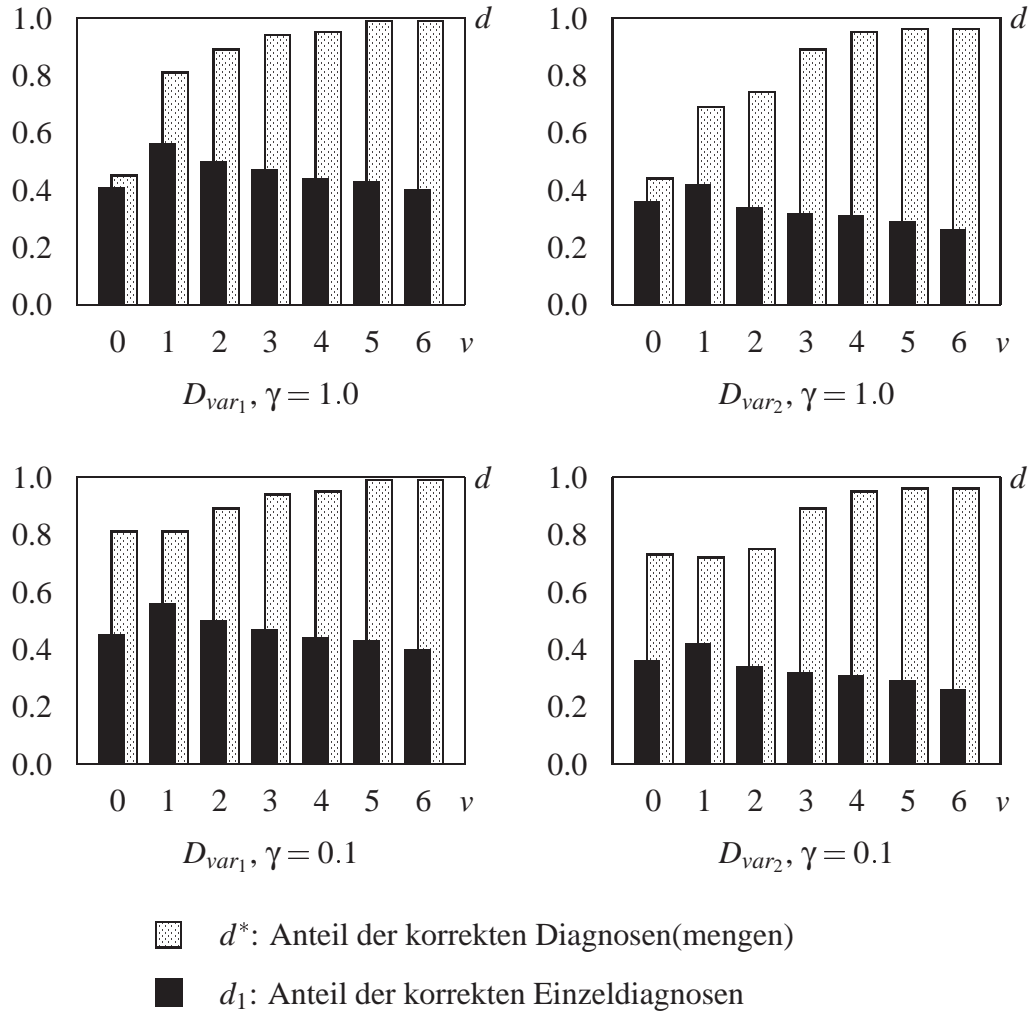


Abbildung 5.7: Entwicklung der Diagnosegüte d in Abhängigkeit von der Meßgeräteeanzahl v (bei konstanten Konfidenzschwellen γ und für zwei Datenbasen D)

uneindeutige Hypothesen (in wenigen Fällen werden sogar Einzeldiagnosen zu Diagnosenmengen).

Da die leichte Abnahme des Anteils korrekter Einzeldiagnosen d_1 durch den größer werdenden Anteil korrekter Diagnosenmengen d_2 mit steigender Meßstellenanzahl deutlich überkompensiert wird, konvergiert die Gesamtdiagnosegüte d^* gegen 1. Allerdings kann aus der leichten Abnahme von d_1 ein Effekt des Überlernens geschlossen werden.

Abhängigkeit von der Konfidenzschwelle

Abbildung 5.8 zeigt in vier Diagrammen die Entwicklung der Diagnosegüte bei einer Erhöhung der Konfidenzschwelle γ , getrennt nach den konstanten Meßstellenanzahlen $\nu = 0$ und $\nu = 2$ sowie nach den beiden Datenbasen D_{var_1} und D_{var_2} . Da die beobachteten Werte für $\gamma = 0.7$ bis $\gamma = 1.0$ identisch sind, wird in den Diagrammen nur der Bereich $\gamma \in [0.1, 0.7]$ dargestellt. Der Anteil korrekter Einzeldiagnosen d_1 bleibt stets in etwa konstant, dagegen wird die Gesamtdiagnosegüte d^* im Fall $\nu = 0$ für höhere Konfidenzschwellen deutlich kleiner, der Anteil der Fehldiagnosen steigt also.

Die Abnahme des Anteils korrekter Diagnosenmengen $d_2 = d^* - d_1$ bei steigender Konfidenzschwelle γ tritt nur auf, wenn kleine Meßstellenanzahlen ν gewählt wurden. Bei $\nu = 0$ sinken die Werte von $d_2 = 0.36$ für $\gamma = 0.1$ auf $d_2 = 0.06$ (D_{var_1}) bzw. $d_2 = 0.13$ (D_{var_2}) für $\gamma = 0.7$. Bereits ab zwei Meßstellen bleiben sie mit etwa $d_2 = 0.4$ für beide Testdatenbasen und für alle Konfidenzschwellen im wesentlichen unverändert.

Für den etwa konstanten Anteil korrekter Einzeldiagnosen sind bereits die wenigen allgemeinen, aber sicheren Regeln der Parameterkombination $\nu = 0/\gamma = 1.0$ verantwortlich. Kommt eine solche Regel zur Anwendung, dann macht sie auch eine korrekte Fehlervorhersage. Wird dagegen die Konfidenzschwelle gesenkt, kommen unsichere Regeln hinzu, die lediglich den Anteil korrekter Diagnosenmengen erhöhen können. Ähnlich ist es bei einer Erhöhung der Meßstellenanzahl: Hier kommen speziellere Regeln zur Ursprungsmenge hinzu, die – im Gegensatz zur Anwendung auf die Lernfälle – bei einer Anwendung auf die Fehlervariationen die Mehrdeutigkeiten erhöhen. Sie sind zwar für die Lernsituationen »optimiert«, aber weniger für neue Situationen geeignet.

Erkennbar ist, daß die gleichgerichteten Effekte, die durch eine Erhöhung von ν und durch eine Senkung von γ entstehen, sich nicht einfach addieren: Bei hoher Meßstellenanzahl bringt die Senkung der Konfidenzschwelle keine Zunahme des Anteils korrekter Diagnosenmengen d_2 . Dies bedeutet, daß in diesem Fall bereits das maximale Diagnosewissen erzeugt wurde. Der Informationsgewinn durch die Erweiterung der Datengrundlage (Auswertung *neuer* Meßstellen) ist also höher als der Informationsgewinn durch die tiefergehende Untersuchung der bisherigen Datengrundlage (Absenkung der Konfidenzschwelle). Als praktische Konsequenz zur Maximierung der Diagnoseergebnisse läßt sich hiermit (auch in Verbindung mit der Entwicklung der Regelanzahl, vgl. Abschnitt 5.3.1) bestätigen, daß eine möglichst hohe Meßstellenanzahl ν angestrebt werden sollte.

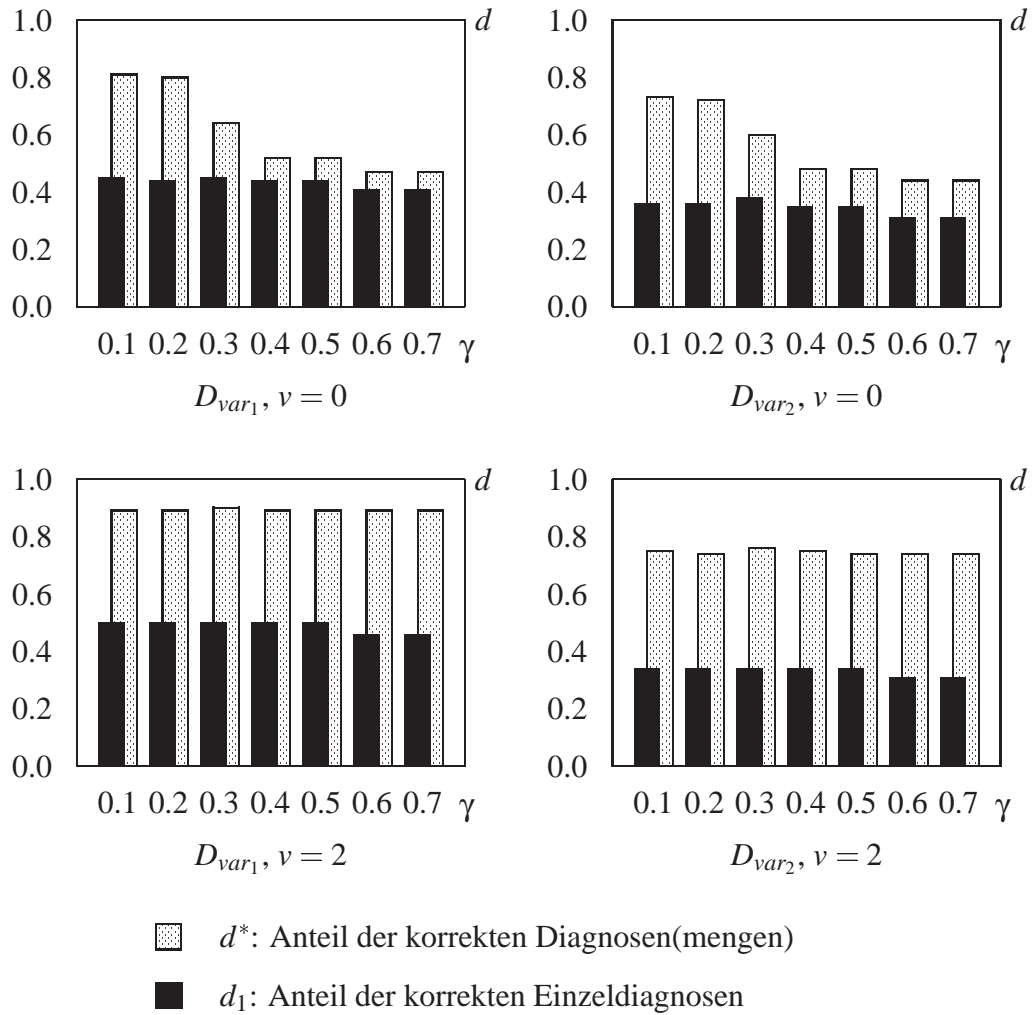


Abbildung 5.8: Entwicklung der Diagnosegüte d in Abhängigkeit von der Konfidenzschwelle γ (bei konstanten Anzahlen zusätzlicher Meßgeräte v und für zwei Datenbasen D)

5.3.3 Fazit

Abschließend seien die wichtigsten Ergebnisse der Evaluierung zusammengefaßt. Die qualitativen Aussagen treffen auf die Diagnosesysteme für alle getesteten hydraulischen Anlagen zu. Die quantitativen Aussagen beziehen sich auf die Testanlage aus Abbildung 5.2.

- Die Diagnoseregelanzahl wächst linear mit zunehmender Meßstellenanzahl.

- Die Diagnoseregelanzahl wächst exponentiell mit abnehmender Konfidenzschwelle.
- Die Gesamtdiagnosegüte ist bei kleinen Meßstellenanzahlen (kleiner zwei) von der Konfidenzschwelle abhängig: Je kleiner die Konfidenzschwelle ist, desto größer ist die Diagnosegüte.
- Die Gesamtdiagnosegüte ist bei großen Meßstellenanzahlen (ab zwei) von der Konfidenzschwelle praktisch unabhängig.
- Das durch zusätzliche Meßstellen gewonnene Diagnosewissen ist »wertvoller« als das durch niedrigere Konfidenzschwellen gewonnene Diagnosewissen.
- Für die gelernten Fehlersituationen beträgt der Anteil korrekter Einzeldiagnosen bis zu 100%.
- Für kleine Fehlervariationen beträgt der Anteil korrekter Einzeldiagnosen bis zu 56%. Inklusive der korrekten Diagnosenmengen beträgt die Gesamtdiagnosegüte bis zu 99%.
- Für mittlere bis große Fehlervariationen beträgt der Anteil korrekter Einzeldiagnosen bis zu 42%. Inklusive der korrekten Diagnosenmengen beträgt die Gesamtdiagnosegüte bis zu 96%.

Besonders unter Berücksichtigung der Tatsache, daß die Testbasen D_{var_1} und D_{var_2} wesentlich größer waren als die Lernbasis D_{lern} (D_{var_1} um 75% und D_{var_2} um 158%, vgl. Anhang B), sind die erzielten Diagnoseergebnisse als sehr gut anzusehen. Sie stützen die Erwartung, daß mit dem in dieser Arbeit vorgestellten automatischen Wissensakquisitions- und Diagnoseansatz eine effiziente Erkennung von Störfällen in hydraulischen Anlagen möglich ist.

5 EVALUIERUNG DES DIAGNOSEANSATZES

6 Zusammenfassung und Ausblick

In der Arbeit wurde ein neuer Ansatz zur automatischen Wissensakquisition für die heuristische Diagnose komplexer hydraulischer Anlagen vorgestellt. Hierbei sind Techniken aus den Gebieten der wissensbasierten Systeme und des Data Minings miteinander verknüpft worden, um zu einem Konzept zu gelangen, das die automatische Generierung von leistungsfähigen Diagnosesystemen ermöglicht. Zur Repräsentation des Diagnosewissens wurden heuristische Assoziationsregeln untersucht, die sich mit geeigneten Lernverfahren effizient aus großen Simulationsdatenbasen erzeugen lassen.

Neben einem Überblick über die Einsatzmöglichkeiten der Wissensentdeckung in Datenbanken bzw. des Data Minings wurde eine systematische Einordnung der bekannten Assoziationsregelkonzepte vorgenommen und hinsichtlich ihrer Eignung als Diagnoseregeln diskutiert. Anschließend erfolgte die detaillierte Beschreibung der theoretischen und algorithmischen Grundlagen eines neuen zweistufigen Wissensakquisitionsansatzes, der auf einer Transformation von kausalen Anlagenmodellen in heuristische Assoziationsregelmodelle zur Darstellung von Symptom→Diagnose-Zusammenhängen basiert. Anhand der Diagnosegüte, die mit dem prototypischen Diagnosesystemgenerator ARGUS für verschiedene hydraulische Anlagen in zahlreichen Testreihen ermittelt wurde, konnte die Qualität des vorgestellten Diagnoseansatzes demonstriert werden.

Der wichtigste Vorteil gegenüber existierenden Diagnoseansätzen für technische Anlagen ist die Möglichkeit zur vollständig automatischen Wissensakquisition, insbesondere auch für Anlagen von hoher Komplexität. Die hiermit verbundenen Zeit- und Kosteneinsparungen beim Aufbau eines Diagnosesystems gehen nicht zu Lasten guter Diagnoseergebnisse, obwohl im Wissensakquisitionskonzept einige Anforderungen aus der Praxis berücksichtigt wurden, die zu einem Informationsverlust führen können (z. B. Diskretisierung mit »glatten« Intervallgrenzen, Nutzung weniger Meßwertintervalle, Einsatz weniger Meßgeräte usw.). Ein weiterer Vorteil liegt in der Verwendung von heuristischen Diagnoseregeln begründet: Es ist jederzeit die Erklärbarkeit einer Diagnoseentscheidung durch die Angabe der hierfür relevanten Regeln gewährleistet.

Ausblick

Neben einer Suche nach neuen Anwendungsbereichen für den vorgestellten Diagnoseansatz (z. B. pneumatische Anlagen oder andere technische Einsatzfelder) sind für

weiterführende Arbeiten u. a. die folgenden Punkte denkbar:

- *Verknüpfung der Hypothesengenerierung mit einer Fehlersimulation:* Die aus der assoziativen Hypothesengenerierung hergeleiteten Verdachtsdiagnosen können anhand der kausalen Verhaltensmodelle durch Simulation überprüft werden; dies ist insbesondere dann sinnvoll, wenn die Konfidenzverrechnung keine eindeutige Diagnose ergibt. Eine solche Strategie wird üblicherweise als *Hypothesize-and-Test* bezeichnet (vgl. [Pup91]).
- *Automatische Optimierung von Wissensbasen:* Der Diagnosesystemgenerator ARGUS kann um die Möglichkeit erweitert werden, systematisch verschiedene Parametervariationen hinsichtlich der damit erzielbaren Diagnosegüte zu testen. Auf diese Weise kann ein selbstoptimierendes System entstehen, das über die wiederholte Erzeugung von Regelbasen sowie ihre Anwendung auf Simulationsdatenbasen zu einer optimalen Lernparametereinstellung für die zu diagnostizierende Anlage gelangt.
- *Berücksichtigung von Mehrfachfehlern:* Dem vorgestellten Diagnoseansatz liegt die Annahme zugrunde, daß im Störfall nur genau eine Komponente fehlerhaft ist. In der Praxis kann jedoch das gleichzeitige Auftreten mehrerer Fehler vorkommen, insbesondere dann, wenn durch den Defekt einer Komponente Folgeschäden verursacht werden. Da die Einzelfehlerannahme nicht konzeptionell bedingt ist (vgl. Abschnitt 4.2.1), sollte sich die Berücksichtigung von Mehrfachfehlern in den Diagnoseansatz integrieren lassen.
- *Variable Konfidenz- und Häufigkeitsschwellen in der Lernphase:* Durch neue Testreihen könnte ermittelt werden, ob zur Steuerung der Assoziationsregelkonstruktion andere Lösungen als die bisherigen starren Konfidenz- und Häufigkeitsschwellen sinnvoll sind. So besteht z. B. die Möglichkeit, die Höhe dieser Parameter von der Regellänge abhängig zu machen, wie etwa die Verwendung höherer Häufigkeitsschwellen für kurze Regeln.
- *Berücksichtigung weiterer Randbedingungen:* Für den realen Einsatz des Diagnosesystemgenerators sollte ermittelt werden, ob zusätzliche physikalische oder technische Einschränkungen zu berücksichtigen sind. Diese könnten z. B. in der Festlegung der zur Verfügung stehenden Meßgeräte oder von A-priori-Wahrscheinlichkeiten für die Komponentenfehler bestehen.

A Mathematischer Anhang

A.1 λ -Maße

In diesem Abschnitt wird die Berechnung des Abhängigkeitsmaßes λ_{m_1} , ausgehend von Definition 4.7 (Seite 97), hergeleitet.

Seien m_1 und $m_2 \in M$ Meßstellen mit den zugehörigen Intervallmengen I_{m_1} und I_{m_2} . Bezogen auf eine Datenbasis D_n mit n Symptomvektoren wird die beobachtete Häufigkeit eines Intervalls $i \in I_{m_1}$ durch $h(i, \bullet)$ angegeben (vgl. Abbildung 4.10). Gleichzeitig ist $h(i, \bullet)$ die Anzahl der richtigen Fälle, falls das Intervall i zur Vorhersage für m_1 dient. Bezogen auf m_1 ist somit der Vorhersagefehler dann am geringsten, wenn das häufigste Intervall, also der sog. Modalwert $\max_{i \in I_{m_1}} \{h(i, \bullet)\}$, als Schätzer genutzt wird. Als minimale Anzahl der Vorhersagefehler ergibt sich dann

$$e(m_1) = n - \max_{i \in I_{m_1}} \{h(i, \bullet)\}.$$

Durch die zusätzliche Auswertung der Intervallverteilung des Meßgeräts m_2 entsteht ein Informationsgewinn, der es erlaubt, zu jedem bei m_2 beobachteten Intervall den besten Schätzer für m_1 anzugeben: Falls ein Intervall $j \in I_{m_2}$ beobachtet wird, ist dies der auf j bezogene Modalwert $\max_{i \in I_{m_1}} \{h(i, j)\}$, und der Vorhersagefehler beträgt $h(\bullet, j) - \max_{i \in I_{m_1}} \{h(i, j)\}$. Die Summe der auf diese Weise für jedes Intervall von m_2 bestimmbaren minimalen Vorhersagefehler führt zu

$$e(m_1|m_2) = \sum_{j \in I_{m_2}} (h(\bullet, j) - \max_{i \in I_{m_1}} \{h(i, j)\}).$$

Somit ist $e(m_1)$ die Anzahl der Vorhersagefehler für m_1 *ohne* Berücksichtigung von m_2 und $e(m_1|m_2)$ die Anzahl der Vorhersagefehler für m_1 *mit* Berücksichtigung von m_2 . Durch Einsetzung in die Formel von Definition 4.7 gelangt man zur folgenden Berechnung von $\lambda_{m_1}(m_2)$:

$$\begin{aligned}
\lambda_{m_1}(m_2) &= \frac{e(m_1) - e(m_1|m_2)}{e(m_1)} \\
&= \frac{n - \max_{i \in I_{m_1}} \{h(i, \bullet)\} - \sum_{j \in I_{m_2}} (h(\bullet, j) - \max_{i \in I_{m_1}} \{h(i, j)\})}{n - \max_{i \in I_{m_1}} \{h(i, \bullet)\}} \\
&= \frac{n - \max_{i \in I_{m_1}} \{h(i, \bullet)\} - \left(n - \sum_{j \in I_{m_2}} \max_{i \in I_{m_1}} \{h(i, j)\} \right)}{n - \max_{i \in I_{m_1}} \{h(i, \bullet)\}} \\
&= \frac{\sum_{j \in I_{m_2}} \max_{i \in I_{m_1}} \{h(i, j)\} - \max_{i \in I_{m_1}} \{h(i, \bullet)\}}{n - \max_{i \in I_{m_1}} \{h(i, \bullet)\}}.
\end{aligned}$$

Die Abhängigkeit in der Gegenrichtung, also $\lambda_{m_2}(m_1)$, ist analog herleitbar.

A.2 Informationsgehalt

Im folgenden wird der informationstheoretische Hintergrund der in Abschnitt 4.6.3 ab Seite 98 besprochenen Bewertungsfunktionen für Diskriminierungsaufwände bei der Meßstellenauswahl umrissen.

Die von Shannon ([Sha48]) eingeführte *Informationstheorie* untersucht die Übertragung und Verarbeitung von Nachrichten und die in ihnen enthaltenen Informationen. Im Gegensatz zur Abgrenzung der Begriffe *Daten* und *Information* in Abschnitt 2.1.1, in der ein Kontextbezug als Charakteristikum für Information im Vordergrund steht, wird in der Kommunikationstheorie Information als interpretationsfrei angesehen (vgl. [SW76]: »Information in der Kommunikationstheorie bezieht sich nicht so sehr auf das, was gesagt *wird*, sondern mehr auf das, was gesagt werden *könnte*.«). Es werden also lediglich theoretische Grenzen aufgezeigt, die bei der Übertragung von Information unabhängig von ihrem Inhalt oder ihrer Bedeutung existieren.

Der Informationsgehalt eines Ereignisses (z. B. der Empfang eines Zeichens oder der Ausgang eines Zufallsexperiments) wird abhängig von der Länge einer optimalen Entscheidungskaskade, die zu seiner Identifizierung (Rekonstruktion) benötigt wird, definiert. Dazu werden in der Regel binäre Alternativentscheidungen vorausgesetzt und jede Entscheidung mit einem Bit bewertet. Werden alle Ereignisse der Grundmenge in den Blättern eines ausgeglichenen binären Suchbaums plziert, so entspricht jede Entscheidungskaskade einem Suchpfad von der Wurzel zum gesuchten Blatt. Bei

n möglichen Ereignissen sind mit Binärsuche $\log_2 n$ Entscheidungen zu treffen; dies führt zur sogenannten Hartley-Formel ([Har28]).

Definition A.1 (Informationsgehalt eines Ereignisses (Hartley-Formel)): Der Informationsgehalt eines Ereignisses e wird mit der Hartley-Formel

$$I(e) = \log_2 n \text{ Bits}$$

bemessen (auch dann, wenn n keine Zweierpotenz ist). \square

Anmerkungen: (i) Bei der Verwendung einer nichtoptimalen Suchstrategie erhöht sich der Informationsgehalt eines Ereignisses jedoch nicht; statt dessen werden die Alternativentscheidungen mit weniger als einem Bit bewertet.

(ii) Wenn nicht anders angegeben, besitzen die Logarithmen die Basis 2: $\log := \log_2$.

Bei nicht gleichverteilten Ereignissen muß eine optimale Entscheidungskaskade die Ereignismenge in jedem Schritt in zwei gleichwahrscheinliche und nicht mehr in zwei gleichgroße Teilmengen zerlegen. Dies führt zu der Beziehung

$$I(e) = \log n = \log \frac{1}{P(e)} \text{ Bits},$$

die sich aus Definition A.1 ergibt, wenn einem Ereignis e aus einer gleichverteilten Grundmenge von n Ereignissen die Wahrscheinlichkeit $P(e) = \frac{1}{n}$ zugewiesen wird.

Hiermit wird deutlich, daß die zugrundeliegende Wahrscheinlichkeitsverteilung die wichtigste Einflußgröße zur Quantifizierung von Informationsgehalten ist; nach Shannons Theorie ist Information nur hierin enthalten. Die folgende Definition gibt den Informationsgehalt beliebiger Wahrscheinlichkeitsverteilungen an.

Definition A.2 (Informationsgehalt und Informationsentropie (Shannon-Formel)): Es sei X eine diskrete Zufallsvariable mit den Ausgängen x_1, \dots, x_l . Die Wahrscheinlichkeitsverteilung sei $P(x_1), \dots, P(x_l)$, für die $\sum_{i=1}^l P(x_i) = 1$ gilt. Dann ist der in Bits gemessene Informationsgehalt eines Ereignisses x_i aus X eine Funktion $I : X \rightarrow \mathbb{R}^{\geq 0}$ mit

$$I(x_i) = \log \frac{1}{P(x_i)}.$$

Der mittlere Informationsgehalt eines Ereignisses von X ist der Erwartungswert $E(I)$ über alle möglichen Informationsgehalte. Er wird Informationsentropie $H(X)$ oder $H(P(x_1), \dots, P(x_l))$ genannt und mit der sogenannten Shannon-Formel in der Einheit Bits pro Ereignis berechnet:

$$\begin{aligned} H(X) = H(P(x_1), \dots, P(x_l)) &:= E(I) \\ &= \sum_{i=1}^l P(x_i) \cdot \log \frac{1}{P(x_i)}. \end{aligned} \quad \square$$

Die Informationsentropie entspricht der mittleren Anzahl von Alternativentscheidungen zur Identifizierung eines Ausgangs von X und ist damit ein Maß dafür, welchen durchschnittlichen Informationszuwachs die Beobachtung eines Ausgangs liefert. Es entsteht ein Informationsgewinn, weil die bisherige Unbestimmtheit über das Auftreten der Ereignisse beseitigt wurde. Die Unbestimmtheit wird dabei durch die Wahrscheinlichkeitsverteilung determiniert.

Anmerkung: Auf ein konkretes Ereignis bezogen ist der Informationsgewinn um so größer, je kleiner die Wahrscheinlichkeit des betreffenden Ereignisses ist. Allerdings ist es sinnvoll, den Entropiebeitrag des »unendlichen« Informationsgehalts $I(e) = \log \frac{1}{0}$ eines unmöglichen Ereignisses e mit $P(e) = 0$ als Null anzusehen, denn es gilt (siehe z. B. [Top74]):

$$\lim_{x \rightarrow 0^+} \left(x \cdot \log \frac{1}{x} \right) = 0.$$

Der folgende Satz gibt an, daß die Informationsentropie für eine Wahrscheinlichkeitsverteilung ihren Maximalwert bei der Gleichverteilung annimmt (ein Beweis ist z. B. in [HQ95] zu finden).

Satz A.1 (Maximale Informationsentropie): Sei X eine diskrete Zufallsvariable mit $l \in \mathbb{N}$ Ausgängen x_1, \dots, x_l und der Wahrscheinlichkeitsverteilung $P(x_1), \dots, P(x_l)$, für die $\sum_{i=1}^l P(x_i) = 1$ gilt. Dann ist stets

$$H(X) \leq \log l,$$

wobei die Gleichheit genau dann gilt, wenn $P(X)$ gleichverteilt ist, d. h. falls $P(x_1) = \dots = P(x_l) = \frac{1}{l}$. \square

B Ergebnistabellen

Die Ergebnistabellen zeigen detailliert die Resultate von mehreren Testreihen, die mit zwei Beispielanlagen durchgeführt wurden. Jeder Einzeltest einer Testreihe bestand aus den folgenden Schritten:

1. Festlegung einer Parameterkombination aus Konfidenzschwelle γ und Anzahl optionaler Meßgeräte v .
2. Erzeugung von jeweils einer Regelmenge für jede Anlagenphase unter Berücksichtigung der Parameterkombination γ / v ; Lerndatenbasis ist D_{lern} .
3. Diagnose aller Testfälle durch Anwendung der Regelmengen auf
 - (a) Lerndatenbasis D_{lern} ,
 - (b) Testdatenbasis D_{var_1} mit leichten Fehlervariationen und
 - (c) Testdatenbasis D_{var_2} mit mittleren bis schweren Fehlervariationen.

Für jede Testreihe variierte die Konfidenzschwelle γ von 1.0 bis 0.1 und die Anzahl v der optionalen Meßgeräte von 0 bis 6. Als Häufigkeitsschwelle wurde immer $\sigma = 0$ verwendet. Die wichtigsten Charakteristika der im folgenden untersuchten Testanlagen sind:

Testanlage	Abbildung	Anz. Hauptkomponenten	Anz. Abtriebskomponenten	Anzahl Phasen	Anzahl Meßstellen ¹
1	5.2, S. 123	9	2	9	6 + 38
2	4.6, S. 76	15	3	10	7 + 36

Die Datenbasen für die Simulationsprotokolle hatten nach der Aufbereitung der mit dem Simulationswerkzeug *art deco* erzeugten Daten (siehe Abschnitte 4.4 bis 4.6) folgende Eigenschaften:

¹ Feste + variable Meßstellen (vgl. Abschnitt 4.6.1).

B ERGEBNISTABELLEN

	Datenbasis	Anz. Fehler	Anz. Zeilen	Anz. Spalten ²	Ø Intervalle pro Spalte ³	Abb. Ergebnistabelle
Testanlage 1	D_{lern}	11	432	46	2.53	B.1, S. 145
	D_{var_1}	20	756			B.2, S. 146
	D_{var_2}	30	1116			B.3, S. 147
Testanlage 2	D_{lern}	14	280	45	2.49	B.4, S. 148
	D_{var_1}	28	560			B.5, S. 149
	D_{var_2}	28	560			B.6, S. 150

Auf einem K6/2-Rechner mit 400 MHz betragen die Zeiten für die Simulation einer Betriebssituation etwa 40 Sekunden für die Testanlage 1 und etwa 200 Sekunden für die Testanlage 2. Die Assoziationsregelerzeugung dauert inklusive der Datenvorverarbeitung in Abhängigkeit von der Parameterwahl etwa eine Sekunde bis etwa eine Minute, die Diagnose (Regelanwendung) pro Störungsfall nur Sekundenbruchteile.

In den Ergebnistabellen besteht jeder Eintrag aus den folgenden vier Werten:

$$\begin{array}{c} d^* \\ d_1/d_2 \\ |R| \end{array}$$

Dabei ist mit den Kriterien aus Definition 5.1 (Seite 126)

- d_1 der Anteil der korrekten Einzeldiagnosen,
- d_2 der Anteil der korrekten Diagnosenmengen,
- $d^* = d_1 + d_2$ der summierte Anteil der korrekten Diagnosen und
- $|R|$ die Anzahl der mit den Parametern γ und ν erzeugten Diagnoseregeln.

Für jede Meßstellenanzahl ν stellen die umrahmten Werte das beste Diagnoseergebnis d^* dar, wobei bei gleichen Ergebnissen dasjenige mit der kleinsten Regelanzahl $|R|$ markiert ist.

² Inklusive Fehler- und Phasenangabe.

³ Die durchschnittliche Anzahl der Intervalle pro Spalte ist über alle Meßstellen und alle Anlagenphasen gemittelt.

γ	Anzahl zusätzlicher Meßgeräte ν						
	0	1	2	3	4	5	6
1.0	0.59 0.59/0.00 82	0.95 0.95/0.00 161	0.98 0.98/0.00 263	1.00 1.00/0.00 369	1.00 1.00/0.00 486	1.00 1.00/0.00 619	1.00 1.00/0.00 685
0.9	0.59 0.59/0.00 82	0.95 0.95/0.00 161	0.98 0.98/0.00 263	1.00 1.00/0.00 369	1.00 1.00/0.00 486	1.00 1.00/0.00 619	1.00 1.00/0.00 685
0.8	0.59 0.59/0.00 82	0.95 0.95/0.00 161	0.98 0.98/0.00 265	1.00 1.00/0.00 371	1.00 1.00/0.00 488	1.00 1.00/0.00 623	1.00 1.00/0.00 689
0.7	0.61 0.57/0.05 86	0.95 0.91/0.05 170	0.98 0.93/0.05 278	1.00 0.98/0.02 388	1.00 0.98/0.02 509	1.00 0.98/0.02 649	1.00 0.98/0.02 717
0.6	0.64 0.57/0.07 92	0.95 0.91/0.05 182	0.98 0.93/0.05 298	1.00 0.98/0.02 420	1.00 0.98/0.02 547	1.00 0.98/0.02 702	1.00 0.98/0.02 775
0.5	0.66 0.61/0.05 181	1.00 0.95/0.05 349	1.00 1.00/0.00 518	1.00 1.00/0.00 728	1.00 1.00/0.00 922	1.00 1.00/0.00 1164	1.00 1.00/0.00 1322
0.4	0.66 0.61/0.05 187	1.00 0.95/0.05 362	1.00 0.98/0.02 537	1.00 1.00/0.00 760	1.00 1.00/0.00 968	1.00 1.00/0.00 1229	1.00 1.00/0.00 1402
0.3	0.82 0.64/0.18 255	1.00 0.95/0.05 463	1.00 0.98/0.02 685	1.00 1.00/0.00 966	1.00 1.00/0.00 1221	1.00 1.00/0.00 1560	1.00 1.00/0.00 1766
0.2	0.95 0.61/0.34 432	1.00 0.95/0.05 732	1.00 1.00/0.00 999	1.00 1.00/0.00 1394	1.00 1.00/0.00 1722	1.00 1.00/0.00 2126	1.00 1.00/0.00 2446
0.1	0.98 0.61/0.36 852	1.00 0.95/0.05 1254	1.00 1.00/0.00 1784	1.00 1.00/0.00 2269	1.00 1.00/0.00 2918	1.00 1.00/0.00 3385	1.00 1.00/0.00 4072

Abbildung B.1: Diagnosegüte und Regelanzahl zur **Testanlage 1** mit Lernbasis D_{lern}
(in Abhängigkeit von den Lernparametern Mindestkonfidenz γ und
Anzahl zusätzlicher Meßgeräte ν)

B ERGEBNISTABELLEN

γ	Anzahl zusätzlicher Meßgeräte ν						
	0	1	2	3	4	5	6
1.0	0.45 0.41/0.04 82	0.81 0.56/0.25 161	0.89 0.50/0.39 263	0.94 0.47/0.47 369	0.95 0.44/0.51 486	0.99 0.43/0.56 619	0.99 0.40/0.59 685
0.9	0.45 0.41/0.04 82	0.81 0.56/0.25 161	0.89 0.50/0.39 263	0.94 0.47/0.47 369	0.95 0.44/0.51 486	0.99 0.43/0.56 619	0.99 0.40/0.59 685
0.8	0.45 0.41/0.04 82	0.81 0.56/0.25 161	0.89 0.50/0.39 265	0.94 0.47/0.47 371	0.95 0.44/0.51 488	0.99 0.43/0.56 623	0.99 0.40/0.59 689
0.7	0.47 0.41/0.06 86	0.81 0.54/0.27 170	0.89 0.46/0.43 278	0.94 0.45/0.49 388	0.95 0.43/0.52 509	0.99 0.42/0.57 649	0.99 0.39/0.60 717
0.6	0.47 0.41/0.06 92	0.81 0.54/0.27 182	0.89 0.46/0.43 298	0.94 0.45/0.49 420	0.95 0.43/0.52 547	0.99 0.42/0.57 702	0.99 0.39/0.60 775
0.5	0.52 0.44/0.08 181	0.81 0.56/0.25 349	0.89 0.50/0.39 518	0.94 0.47/0.47 728	0.95 0.44/0.51 922	0.99 0.43/0.56 1164	0.99 0.40/0.59 1322
0.4	0.52 0.44/0.08 187	0.81 0.56/0.25 362	0.89 0.50/0.39 537	0.94 0.47/0.47 760	0.95 0.44/0.51 968	0.99 0.43/0.56 1229	0.99 0.40/0.59 1402
0.3	0.64 0.45/0.19 255	0.84 0.56/0.28 463	0.90 0.50/0.40 685	0.94 0.47/0.47 966	0.95 0.44/0.51 1221	0.99 0.43/0.56 1560	0.99 0.40/0.59 1766
0.2	0.80 0.44/0.36 432	0.81 0.56/0.25 732	0.89 0.50/0.39 999	0.94 0.47/0.47 1394	0.95 0.44/0.51 1722	0.99 0.43/0.56 2126	0.99 0.40/0.59 2446
0.1	0.81 0.45/0.36 852	0.81 0.56/0.25 1254	0.89 0.50/0.39 1784	0.94 0.47/0.47 2269	0.95 0.44/0.51 2918	0.99 0.43/0.56 3385	0.99 0.40/0.59 4072

Abbildung B.2: Diagnosegüte und Regelanzahl zur **Testanlage 1** mit Testbasis D_{var_1} (in Abhängigkeit von den Lernparametern Mindestkonfidenz γ und Anzahl zusätzlicher Meßgeräte ν)

γ	Anzahl zusätzlicher Meßgeräte ν						
	0	1	2	3	4	5	6
1.0	0.44 0.36/0.08 82	0.69 0.42/0.27 161	0.74 0.34/0.40 263	0.89 0.32/0.57 369	0.95 0.31/0.64 486	0.96 0.29/0.67 619	0.96 0.26/0.70 685
0.9	0.44 0.36/0.08 82	0.69 0.42/0.27 161	0.74 0.34/0.40 263	0.89 0.32/0.57 369	0.95 0.31/0.64 486	0.96 0.29/0.67 619	0.96 0.26/0.70 685
0.8	0.44 0.36/0.08 82	0.69 0.42/0.27 161	0.74 0.34/0.40 265	0.89 0.32/0.57 371	0.95 0.31/0.64 488	0.96 0.29/0.67 623	0.96 0.26/0.70 689
0.7	0.44 0.31/0.13 86	0.71 0.40/0.31 170	0.74 0.31/0.43 278	0.89 0.32/0.57 388	0.95 0.31/0.64 509	0.96 0.29/0.67 649	0.96 0.26/0.70 717
0.6	0.44 0.31/0.13 92	0.71 0.40/0.31 182	0.74 0.31/0.43 298	0.89 0.32/0.57 420	0.95 0.31/0.64 547	0.96 0.29/0.67 702	0.96 0.26/0.70 775
0.5	0.48 0.35/0.13 181	0.70 0.42/0.28 349	0.74 0.34/0.40 518	0.89 0.32/0.57 728	0.95 0.31/0.64 922	0.96 0.29/0.67 1164	0.96 0.26/0.70 1322
0.4	0.48 0.35/0.13 187	0.71 0.42/0.29 362	0.75 0.34/0.41 537	0.89 0.32/0.57 760	0.95 0.31/0.64 968	0.96 0.29/0.67 1229	0.96 0.26/0.70 1402
0.3	0.60 0.38/0.22 255	0.73 0.42/0.31 463	0.76 0.34/0.42 685	0.89 0.32/0.57 966	0.95 0.31/0.64 1221	0.96 0.29/0.67 1560	0.96 0.26/0.70 1766
0.2	0.72 0.36/0.36 432	0.71 0.42/0.29 732	0.74 0.34/0.40 999	0.89 0.32/0.57 1394	0.95 0.31/0.64 1722	0.96 0.29/0.67 2126	0.96 0.26/0.70 2446
0.1	0.73 0.36/0.37 852	0.72 0.42/0.30 1254	0.75 0.34/0.41 1784	0.89 0.32/0.57 2269	0.95 0.31/0.64 2918	0.96 0.29/0.67 3385	0.96 0.26/0.70 4072

Abbildung B.3: Diagnosegüte und Regelanzahl zur **Testanlage 1** mit Testbasis D_{var_2}
(in Abhängigkeit von den Lernparametern Mindestkonfidenz γ und
Anzahl zusätzlicher Meßgeräte ν)

B ERGEBNISTABELLEN

γ	Anzahl zusätzlicher Meßgeräte ν						
	0	1	2	3	4	5	6
1.0	0.43 0.43/0.00 54	0.79 0.79/0.00 99	0.86 0.86/0.00 172	0.86 0.86/0.00 232	0.86 0.86/0.00 286	0.86 0.86/0.00 373	0.86 0.86/0.00 422
0.9	0.43 0.43/0.00 54	0.79 0.79/0.00 99	0.86 0.86/0.00 172	0.86 0.86/0.00 232	0.86 0.86/0.00 286	0.86 0.86/0.00 373	0.86 0.86/0.00 422
0.8	0.43 0.43/0.00 54	0.79 0.79/0.00 99	0.86 0.86/0.00 172	0.86 0.86/0.00 232	0.86 0.86/0.00 286	0.86 0.86/0.00 373	0.86 0.86/0.00 422
0.7	0.43 0.43/0.00 54	0.79 0.79/0.00 99	0.86 0.86/0.00 172	0.86 0.86/0.00 232	0.86 0.86/0.00 286	0.86 0.86/0.00 373	0.86 0.86/0.00 422
0.6	0.43 0.43/0.00 57	0.79 0.79/0.00 104	0.86 0.86/0.00 181	0.86 0.86/0.00 242	0.86 0.86/0.00 298	0.86 0.86/0.00 394	0.86 0.86/0.00 446
0.5	0.75 0.43/0.32 102	1.00 0.79/0.21 199	1.00 0.89/0.11 339	1.00 0.89/0.11 450	1.00 0.89/0.11 547	1.00 0.89/0.11 718	1.00 0.89/0.11 804
0.4	0.75 0.43/0.32 104	1.00 0.79/0.21 206	1.00 0.89/0.11 356	1.00 0.89/0.11 473	1.00 0.89/0.11 574	1.00 0.89/0.11 752	1.00 0.89/0.11 838
0.3	1.00 0.43/0.57 214	1.00 0.79/0.21 375	1.00 0.86/0.14 567	1.00 0.86/0.14 740	1.00 0.86/0.14 579	1.00 0.86/0.14 1097	1.00 0.86/0.14 1206
0.2	1.00 0.43/0.57 407	1.00 0.79/0.21 654	1.00 0.86/0.14 935	1.00 0.86/0.14 1179	1.00 0.86/0.14 1370	1.00 0.86/0.14 1659	1.00 0.86/0.14 1786
0.1	1.00 0.43/0.57 699	1.00 0.79/0.21 1134	1.00 0.86/0.14 1622	1.00 0.86/0.14 2121	1.00 0.86/0.14 2526	1.00 0.86/0.14 2988	1.00 0.86/0.14 3326

Abbildung B.4: Diagnosegüte und Regelanzahl zur **Testanlage 2** mit Lernbasis D_{lern} (in Abhängigkeit von den Lernparametern Mindestkonfidenz γ und Anzahl zusätzlicher Meßgeräte ν)

γ	Anzahl zusätzlicher Meßgeräte ν						
	0	1	2	3	4	5	6
1.0	0.36 0.34/0.02 54	0.68 0.63/0.05 99	0.80 0.64/0.16 172	0.84 0.61/0.23 232	0.84 0.54/0.30 286	0.84 0.50/0.34 373	0.84 0.50/0.34 422
0.9	0.36 0.34/0.02 54	0.68 0.63/0.05 99	0.80 0.64/0.16 172	0.84 0.61/0.23 232	0.84 0.54/0.30 286	0.84 0.50/0.34 373	0.84 0.50/0.34 422
0.8	0.36 0.34/0.02 54	0.68 0.63/0.05 99	0.80 0.64/0.16 172	0.84 0.61/0.23 232	0.84 0.54/0.30 286	0.84 0.50/0.34 373	0.84 0.50/0.34 422
0.7	0.36 0.34/0.02 54	0.68 0.63/0.05 99	0.80 0.64/0.16 172	0.84 0.61/0.23 232	0.84 0.54/0.30 286	0.84 0.50/0.34 373	0.84 0.50/0.34 422
0.6	0.36 0.34/0.02 57	0.70 0.65/0.05 104	0.80 0.64/0.16 181	0.84 0.61/0.23 242	0.84 0.54/0.30 298	0.84 0.50/0.34 394	0.84 0.50/0.34 446
0.5	0.61 0.36/0.25 102	0.82 0.64/0.18 199	0.87 0.64/0.23 339	0.91 0.61/0.30 450	0.91 0.54/0.37 547	0.91 0.50/0.41 718	0.91 0.50/0.41 804
0.4	0.61 0.36/0.25 104	0.82 0.64/0.18 206	0.87 0.64/0.23 356	0.91 0.61/0.30 473	0.91 0.54/0.37 574	0.91 0.50/0.41 752	0.91 0.50/0.41 838
0.3	0.79 0.34/0.45 214	0.82 0.64/0.18 375	0.87 0.64/0.23 567	0.91 0.61/0.30 740	0.91 0.54/0.37 579	0.91 0.50/0.41 1097	0.91 0.50/0.41 1206
0.2	0.79 0.34/0.45 407	0.82 0.64/0.18 654	0.86 0.65/0.21 935	0.89 0.61/0.28 1179	0.89 0.54/0.35 1370	0.89 0.50/0.39 1659	0.89 0.50/0.39 1786
0.1	0.79 0.34/0.45 699	0.82 0.63/0.19 1134	0.86 0.65/0.21 1622	0.89 0.61/0.28 2121	0.89 0.54/0.35 2526	0.89 0.50/0.39 2988	0.89 0.50/0.39 3326

Abbildung B.5: Diagnosegüte und Regelanzahl zur **Testanlage 2** mit Testbasis D_{var_1}
(in Abhängigkeit von den Lernparametern Mindestkonfidenz γ und
Anzahl zusätzlicher Meßgeräte ν)

B ERGEBNISTABELLEN

γ	Anzahl zusätzlicher Meßgeräte ν						
	0	1	2	3	4	5	6
1.0	0.29 0.25/0.04 54	0.59 0.55/0.04 99	0.70 0.54/0.16 172	0.73 0.43/0.30 232	0.73 0.38/0.35 286	0.75 0.38/0.37 373	0.75 0.38/0.37 422
0.9	0.29 0.25/0.04 54	0.59 0.55/0.04 99	0.70 0.54/0.16 172	0.73 0.43/0.30 232	0.73 0.38/0.35 286	0.75 0.38/0.37 373	0.75 0.38/0.37 422
0.8	0.29 0.25/0.04 54	0.59 0.55/0.04 99	0.70 0.54/0.16 172	0.73 0.43/0.30 232	0.73 0.38/0.35 286	0.75 0.38/0.37 373	0.75 0.38/0.37 422
0.7	0.29 0.25/0.04 54	0.59 0.55/0.04 99	0.70 0.54/0.16 172	0.73 0.43/0.30 232	0.73 0.38/0.35 286	0.75 0.38/0.37 373	0.75 0.38/0.37 422
0.6	0.29 0.25/0.04 57	0.63 0.59/0.04 104	0.70 0.54/0.16 181	0.73 0.43/0.30 242	0.73 0.38/0.35 298	0.75 0.38/0.37 394	0.75 0.38/0.37 446
0.5	0.57 0.29/0.28 102	0.75 0.59/0.16 199	0.77 0.54/0.23 339	0.80 0.43/0.37 450	0.80 0.38/0.42 547	0.82 0.38/0.44 718	0.82 0.38/0.44 804
0.4	0.57 0.29/0.28 104	0.75 0.59/0.16 206	0.77 0.54/0.23 356	0.80 0.43/0.37 473	0.80 0.38/0.42 574	0.82 0.38/0.44 752	0.82 0.38/0.44 838
0.3	0.73 0.25/0.48 214	0.75 0.59/0.16 375	0.77 0.54/0.23 567	0.80 0.43/0.37 740	0.80 0.38/0.42 579	0.82 0.38/0.44 1097	0.82 0.38/0.44 1206
0.2	0.73 0.25/0.48 407	0.75 0.59/0.16 654	0.75 0.54/0.21 935	0.79 0.43/0.36 1179	0.79 0.38/0.41 1370	0.80 0.38/0.42 1659	0.80 0.38/0.42 1786
0.1	0.73 0.25/0.48 699	0.75 0.55/0.20 1134	0.75 0.54/0.21 1622	0.79 0.43/0.36 2121	0.79 0.38/0.41 2526	0.80 0.38/0.42 2988	0.80 0.38/0.42 3326

Abbildung B.6: Diagnosegüte und Regelanzahl zur **Testanlage 2** mit Testbasis D_{var_2} (in Abhängigkeit von den Lernparametern Mindestkonfidenz γ und Anzahl zusätzlicher Meßgeräte ν)

Literaturverzeichnis

- [Aig96] M. Aigner. *Diskrete Mathematik*. Friedr. Vieweg & Sohn Verlagsgesellschaft, Braunschweig, Wiesbaden, 1996.
- [AIS93] R. Agrawal, T. Imielinski, A. Swami. Mining association rules between sets of items in large databases. In Peter Buneman, Sushil Jajodia (Hrsg.), *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*, Washington, D. C., Mai 1993. ACM Press.
- [AMS⁺96] R. Agrawal, H. Mannila, R. Srikant, H. Toivonen, A. I. Verkamo. Fast discovery of association rules. In U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, R. Uthurusamy (Hrsg.), *Advances in Knowledge Discovery and Data Mining*, Seiten 307–328, Cambridge, 1996. AAAI / MIT Press.
- [Bay97] R. J. Bayardo Jr. Brute-force mining of high-confidence classification rules. In *Proceedings of the Third International Conference on Knowledge Discovery in Databases and Data Mining, KDD'97*, Seiten 123–126. AAAI Press, 1997.
- [BB89] G. Bamberg, F. Baur. *Statistik*. R. Oldenbourg Verlag, München, 6. Auflage, 1989.
- [BM99] P. Brockhausen, K. Morik. Wissensentdeckung in relationalen Datenbanken: Eine Herausforderung für das maschinelle Lernen. In G. Nakhaeizadeh (Hrsg.), *Data Mining, theoretische Aspekte und Anwendungen*. Physika Verlag, 1999.
- [BMS97] S. Brin, R. Motwani, C. Silverstein. Beyond market baskets: Generalizing association rules to correlations. In *Proceedings of the 1997 ACM SIGMOD Conference on Management of Data*, Seiten 265–276, Tucson, AZ, Mai 1997.
- [BMUT97] S. Brin, R. Motwani, J. D. Ullman, S. Tsur. Dynamic itemset counting and implication rules for market basket data. In *Proceedings of the ACM SIGMOD International Conference on Management of Data, SIGMOD-97*, Seiten 255–264. ACM Press, Mai 1997.
- [Boo88] J. H. Boose. A research framework for knowledge acquisition techniques and tools. In *Proceedings of the Second European Knowledge Acquisition Workshop, EKAW-88*. GMD-Studien Nr. 143, 1988.

LITERATURVERZEICHNIS

- [BS84] B. G. Buchanan, E. H. Shortliffe (Hrsg.). *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Addison-Wesley Publishing Company, Reading, Massachusetts, 1984.
- [BW98] O. Büchter, R. Wirth. Discovery of association rules over ordinal data: A new and faster algorithm and its application to basket analysis. In X. Wu, R. Kotagiri, K. B. Korb (Hrsg.), *Research and Development in Knowledge Discovery and Data Mining, Proceedings of the Second Pacific-Asia Conference PAKDD-98*, Seiten 36–47, Melbourne, Australia, April 1998. Springer-Verlag, Berlin, Heidelberg, New York.
- [CBS91] C.-C. Chan, C. Batur, A. Srinivasan. Determination of quantization intervals in rule based model for dynamic systems. In *Proceedings of the IEEE Conference on Systems, Man, and Cybernetics*, Seiten 1719–1723, Charlottesville, Virginia, 1991.
- [CDF⁺00] E. Cohen, M. Datar, S. Fujiwara, A. Gionis, P. Indyk, R. Motwani, J. Ullman, C. Yang. Finding interesting associations without support pruning. In *Proceedings of the 16th Annual IEEE Conference on Data Engineering (ICDE 2000)*, Februar 2000.
- [CHC97] C. L. Carter, H. J. Hamilton, N. Cercone. Share based measures for itemsets. In J. Komorowski, J. Zytkow (Hrsg.), *Principles of Data Mining and Knowledge Discovery: Proceedings of the First European Symposium PKDD '97*, Seiten 14–24, Trondheim, Norway, Juni 1997. Springer-Verlag, Berlin, Heidelberg, New York.
- [CM83] B. Chandrasekaran, S. Mittal. Deep versus compiled knowledge approaches to diagnostic problem solving. *AAAI-83*, Seiten 349–354, 1983.
- [CPDT93] L. Console, L. Portinale, D. T. Dupré, P. Torasso. *Second Generation Expert Systems*. Springer-Verlag, Berlin, Heidelberg, New York, 1993.
- [Cur96] D. Curatolo. *Wissensbasierte Methoden zur effizienten Simulation fluidtechnischer Systeme*. Dissertation, Universität-Gesamthochschule Paderborn, Fachbereich Mathematik/Informatik, 1996.
- [DD87] H. Dreyfus, S. Dreyfus. *Künstliche Intelligenz: Von den Grenzen der Denkmachines und dem Wert der Intuition*. Rororo, 1987.
- [DD98] A. K. Dörp, W. Dücker. *Meßtechnik in der Hydraulik*. VDE-Verlag, Berlin, Offenbach, 1998.

- [DKS95] J. Dougherty, R. Kohavi, M. Sahami. Supervised and unsupervised discretization of continuous features. In A. Prieditis, S. Russell (Hrsg.), *Proceedings of the 12th International Conference on Machine Learning*, Seiten 194–202, Tahoe City, CA, Juli 1995. Morgan Kaufmann, Menlo Park, CA.
- [dKW87] J. de Kleer, B. C. Williams. Diagnosing multiple faults. *Artificial Intelligence*, 32, Seiten 97–130, 1987.
- [DL98] G. Dong, J. Li. Interestingness of discovered association rules in terms of neighborhood-based unexpectedness. In X. Wu, R. Kotagiri, K. B. Korb (Hrsg.), *Research and Development in Knowledge Discovery and Data Mining, Proceedings of the Second Pacific-Asia Conference PAKDD-98*, Seiten 72–86, Melbourne, Australia, April 1998. Springer-Verlag, Berlin, Heidelberg, New York.
- [DRSS97] J. S. Deogun, V. V. Raghavan, A. Sarkar, H. Sever. Data mining: Research trends, challenges, and applications. In T. Y. Lin, N. Cercone (Hrsg.), *Rough Sets and Data Mining: Analysis of Imprecise Data*, Seiten 9–45. Kluwer Academic Publishers, Boston, MA, 1997.
- [Düs99] R. Düsing. Knowledge Discovery in Databases und Data Mining. In P. Chamoni, P. Gluchowski (Hrsg.), *Analytische Informationssysteme: Data Warehouse, On-Line Analytical Processing, Data Mining*, Seiten 345–353. Springer-Verlag, Berlin, Heidelberg, New York, 1999.
- [DZWL99] G. Dong, X. Zhang, L. Wong, J. Li. CAEP: Classification by aggregating emerging patterns. *Discovery Science*, 1999.
- [EMT95] A. Engel, M. Möhring, K. G. Troitzsch. *Sozialwissenschaftliche Datenanalyse*. BI-Wissenschaftsverlag, Mannheim, Leipzig, Wien, Zürich, 1995.
- [FdK93] K. D. Forbus, J. de Kleer. *Building Problem Solvers*. MIT Press, Cambridge, Massachusetts, 1993.
- [FI92] U. M. Fayyad, K. B. Irani. On the handling of continuous-valued attributes in decision tree generation. *Machine Learning*, 8, Seiten 87–102, 1992.
- [FMMT96] T. Fukuda, Y. Morimoto, S. Morishita, T. Tokuyama. Mining optimized association rules for numeric attributes. In *Proceedings of the Fifteenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, Seiten 182–191, Montreal, Canada, Juni 1996. ACM Press.

LITERATURVERZEICHNIS

- [FPS96] U. Fayyad, G. Piatetsky-Shapiro, P. Smyth. From data mining to knowledge discovery in databases. In *AI Magazine*, Band 17, Seiten 37–54. AAAI Press, Menlo Park, CA, 1996.
- [FPSU96] U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, R. Uthurusamy. *Advances in Knowledge Discovery and Data Mining*. AAAI / MIT Press, Cambridge, 1996.
- [GG00] M. Grothe, P. Gentsch. *Business Intelligence*. Addison-Wesley Verlag, 2000.
- [GGP98] S. Guillaume, F. Guillet, J. Philippe. Improving the discovery of association rules with intensity of implication. In J. M. Zytkow, M. Quafafou (Hrsg.), *Principles of Data Mining and Knowledge Discovery, Second European Symposium PKDD-98*, Seiten 318–327, Nantes, France, September 1998. Springer-Verlag, Berlin, Heidelberg, New York.
- [GK54] L. Goodman, W. Kruskal. Measure of association for cross classification. *Journal of the American Statistical Association*, 49, Seiten 732–764, 1954.
- [GKP⁺83] J. Gaschnig, P. Klahr, H. Pople, E. Shortliffe, A. Terry. Evaluation of expert systems: Issues and case studies. In F. Hayes-Roth, D. A. Waterman, D. B. Lenat (Hrsg.), *Building Expert Systems*, Seiten 241–280. Addison-Wesley Publishing Company, Reading, Mass., 1983.
- [GO98] B. Gray, M. E. Orlowska. CCAIIA: Clustering categorical attributes into interesting association rules. In X. Wu, R. Kotagiri, K. B. Korb (Hrsg.), *Research and Development in Knowledge Discovery and Data Mining, Proceedings of the Second Pacific-Asia Conference PAKDD-98*, Seiten 132–143, Melbourne, Australia, April 1998. Springer-Verlag, Berlin, Heidelberg, New York.
- [Har28] R. V. L. Hartley. Transmission of information. Technical report, Bell System Technical Journal, 1928.
- [Har89] J. Hartung. *Statistik*. R. Oldenbourg Verlag, München, 7. Auflage, 1989.
- [HCHC98] R. J. Hilderman, C. L. Carter, H. J. Hamilton, N. Cercone. Mining market basket data using share measures and characterized itemsets. In X. Wu, R. Kotagiri, K. B. Korb (Hrsg.), *Research and Development in Knowledge Discovery and Data Mining, Proceedings of the Second Pacific-Asia Conference PAKDD-98*, Seiten 159–173, Melbourne, Australia, April 1998. Springer-Verlag, Berlin, Heidelberg, New York.

- [Hel96] H. Helbig. *Künstliche Intelligenz und automatische Wissensverarbeitung*. Verlag Technik, Berlin, 2. Auflage, 1996.
- [Hes99] T. Hesse. *Zur wissensbasierten Diagnose fluidischer Systeme und automatischen Generierung von dialogbasierten Diagnoseunterstützungssystemen*. Dissertation, Universität-Gesamthochschule Paderborn, Fachbereich Mathematik/Informatik, 1999.
- [HK89] C. Hestermann, S. Karl. Methoden und Werkzeuge für die Wissensakquisition. Technical Report 7/89, Universität Karlsruhe, Fakultät für Informatik, Mai 1989.
- [HL90] D. Hartmann, K. Lehner. *Technische Expertensysteme*. Springer-Verlag, Berlin, Heidelberg, New York, 1990.
- [HMS88] E. Hering, R. Martin, M. Stohrer. *Physik für Ingenieure*. VDI-Verlag, Düsseldorf, 1988.
- [Hof99] M. Hoffmann. *Zur Automatisierung des Designprozesses fluidischer Systeme*. Dissertation, Universität-Gesamthochschule Paderborn, Fachbereich Mathematik/Informatik, 1999.
- [Hol93] R. C. Holte. Very simple classification rules perform well on most commonly used datasets. *Machine Learning*, 11, Seiten 63–90, 1993.
- [HQ95] W. Heise, P. Quattrocchi. *Informations- und Codierungstheorie*. Springer-Verlag, Berlin, Heidelberg, New York, 1995.
- [HS94] M. Holsheimer, A. Siebes. The search for knowledge in databases. Technical report, Centrum voor Wiskunde en Informatica, 1994.
- [HS97] K. M. Ho, P. D. Scott. Zeta: A global method for discretization of continuous variables. In *Proceedings of KDD-97, The Third International Conference on Knowledge Discovery and Data Mining*, Seiten 191–194, Newport Beach, CA., August 1997. AAAI Press.
- [HS98] K. M. Ho, P. D. Scott. An efficient global discretization method. In X. Wu, R. Kotagiri, K. B. Korb (Hrsg.), *Proceedings of PAKDD-98, The Second Pacific-Asia Conference on Knowledge Discovery and Data Mining*, Seiten 383–384, Melbourne, Australia, April 1998. Springer-Verlag, Berlin, Heidelberg, New York.
- [HSA99] J. Heinsohn, R. Socher-Ambrosius. *Wissensverarbeitung*. Spektrum Akademischer Verlag, Heidelberg, Berlin, 1999.

LITERATURVERZEICHNIS

- [Inm96] W. H. Inmon. *Building the Data Warehouse*. John Wiley & Sons, New York, 2. Auflage, 1996.
- [KCH⁺95] H. Kleine Büning, D. Curatolo, M. Hoffmann, R. Lemmen, M. Suermann, B. Stein. ArtDeco – Entwurfsunterstützung in der Hydraulik. *KI* 5/95, 1995.
- [KE97] A. Kemper, A. Eickler. *Datenbanksysteme*. R. Oldenbourg Verlag, München, Wien, 2. Auflage, 1997.
- [Ker92] R. Kerber. ChiMerge: Discretization of numeric attributes. In *AAAI-92, Proceedings of the Tenth National Conference on Artificial Intelligence*, Seiten 123–128. AAAI Press/MIT Press, Menlo Park, 1992.
- [KFW98] C. M. Kuok, A. Fu, M. H. Wong. Mining fuzzy association rules in databases. *SIGMOD Record*, 27, Seiten 41–46, 1998.
- [Kir97] J. Kirchner. Transformationsprogramme und Extraktionsprozesse entscheidungsrelevanter Basisdaten. In H. Mucksch, W. Behme (Hrsg.), *Das Data Warehouse-Konzept*. Gabler Verlag, Wiesbaden, 2. Auflage, 1997.
- [KL90] W. Karbach, M. Linster. *Wissensakquisition für Expertensysteme. Techniken, Modelle und Softwarewerkzeuge*. Hanser, München, 1990.
- [KL94] H. Kleine Büning, T. Lettmann. *Aussagenlogik: Deduktion und Algorithmen*. B. G. Teubner, Stuttgart, 1994.
- [KLKF98] F. Korn, A. Labrinidis, Y. Kotidis, C. Faloutsos. Ratio rules: A new paradigm for fast, quantifiable data mining. In *Proceedings of 24th International Conference on Very Large Databases (VLDB'98)*, Seiten 582–593, New York City, New York, USA, August 1998.
- [KMRV94] M. Klemettinen, H. Mannila, P. Ronkainen, A. I. Verkamo. Finding interesting rules from large sets of discovered association rules. In N. R. Adan, B. K. Bhargava, Y. Yesha (Hrsg.), *Third International Conference on Information and Knowledge Management*, Seiten 401–407, Gaithersburg, Maryland, November 1994. ACM Press.
- [Kry98a] M. Kryszkiewics. Representative association rules and minimum condition maximum consequence association rules. In J. M. Zytkow, M. Quafafou (Hrsg.), *Principles of Data Mining and Knowledge Discovery, Second European Symposium PKDD-98*, Seiten 361–369, Nantes, France, September 1998. Springer-Verlag, Berlin, Heidelberg, New York.

- [Kry98b] M. Kryszkiewicz. Representative association rules. In X. Wu, R. Kotagiri, K. B. Korb (Hrsg.), *Research and Development in Knowledge Discovery and Data Mining, Proceedings of the Second Pacific-Asia Conference PAKDD-98*, Seiten 198–209, Melbourne, Australia, April 1998. Springer-Verlag, Berlin, Heidelberg, New York.
- [Les96] S. Leschka. *Fallbasiertes Störungsmanagement in flexiblen Fertigungssystemen*. Dissertation, Universität-Gesamthochschule Paderborn, Heinz Nixdorf Institut, 1996. HNI-Verlagsschriftenreihe, Band 15; Rechnerintegrierte Produktion.
- [LHM98] B. Liu, W. Hsu, Y. Ma. Integrating classification and association rule mining. In R. Agrawal, P. Storloz, G. Piatetsky-Shapiro (Hrsg.), *Proceedings of the Fourth International Conference on Knowledge Discovery in Databases and Data Mining, KDD'98*, Seiten 80–86, New York, NY, 1998. AAAI Press.
- [Lus90] M. Lusti. *Wissensbasierte Systeme: Algorithmen, Datenstrukturen und Werkzeuge*. BI-Wissenschaftsverlag, Mannheim, Wien, Zürich, 1990.
- [Man97] H. Mannila. Methods and problems in data mining. In F. N. Afrati, P. Kolaitis (Hrsg.), *Proceedings of the 6th International Conference on Database Theory ICDT '97*, Delphi, Griechenland, Januar 1997. Springer-Verlag, Berlin, Heidelberg, New York.
- [Mat91] H. J. Matthies. *Einführung in die Ölhydraulik*. B. G. Teubner, Stuttgart, 1991.
- [Mer80] P. Mertens. Die Theorie der Mustererkennung in den Wirtschaftswissenschaften. In I. Dahlberg (Hrsg.), *Studien zur Klassifikation, Bd. 4: Klassifikation und Erkenntnis, Proceedings der 3. Fachtagung der Gesellschaft für Klassifikation e. V.*, Frankfurt, 1980. Indeks-Verlag.
- [MY97] R. Miller, Y. Yang. Association rules over interval data. In *Proceedings of the ACM SIGMOD International Conference on Management of Data, SIGMOD-97*, Seiten 452–461, Tucson, Arizona, USA, Mai 1997. ACM Press.
- [NL94] J. Niebuhr, G. Lindner. *Physikalische Meßtechnik mit Sensoren*. R. Oldenbourg Verlag, München, Wien, 3. Auflage, 1994.
- [Pfa95] B. Pfahringer. Compression-based discretization of continuous attributes. In A. Friediris, S. Russell (Hrsg.), *Proceedings of the 12th International Conference on Machine Learning*. Morgan Kaufmann, Menlo Park, CA, Juli 1995.

LITERATURVERZEICHNIS

- [PGPB96] F. Puppe, U. Gappa, K. Poeck, S. Bamberger. *Wissensbasierte Diagnose- und Informationssysteme*. Springer-Verlag, Berlin, Heidelberg, New York, 1996.
- [Pol97] W. Polasek. *Schließende Statistik*. Springer-Verlag, Berlin, Heidelberg, New York, 1997.
- [PR93] T. Pfeifer, M. M. Richter (Hrsg.). *Diagnose von technischen Systemen*. DUV, Deutscher Universitäts-Verlag, Wiesbaden, 1993.
- [Pup90] F. Puppe. *Problemlösungsmethoden für Expertensysteme*. Springer-Verlag, Berlin, Heidelberg, New York, 1990.
- [Pup91] F. Puppe. *Einführung in Expertensysteme*. Springer-Verlag, Berlin, Heidelberg, New York, 2. Auflage, 1991.
- [QR89] J. R. Quinlan, R. L. Rivest. Inferring decision trees using the minimum description length principle. *Information and Computing*, 80, Seiten 227–248, 1989.
- [Qui93] J. R. Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers, San Mateo, CA, 1993.
- [Rap96] R. Rapp. *Die Berechnung von Assoziationen: ein korpuslinguistischer Ansatz*. Dissertation, Universität Konstanz, Informationswissenschaft, 1996.
- [Ric89] M. M. Richter. *Prinzipien der künstlichen Intelligenz: Wissensrepräsentation, Inferenz und Expertensysteme*. B. G. Teubner, Stuttgart, 1989.
- [RMS98] S. Ramaswamy, S. Mahajan, A. Silberschatz. On the discovery of interesting patterns in association rules. In A. Gupta, O. Shmueli, J. Widom (Hrsg.), *VLDB'98, Proceedings of 24th International Conference on Very Large Data Bases*, Seiten 368–379, New York City, New York, USA, August 1998. Morgan Kaufmann.
- [RP99] P. Rechenberg, G. Pomberger. *Informatik-Handbuch*. Carl Hanser Verlag, München, Wien, 2. Auflage, 1999.
- [RR95] M. Richeldi, M. Rossotto. Class-driven statistical discretization of continuous attributes. In N. Lavrac, S. Wrobel (Hrsg.), *ECML-95, Proceedings of the European Conference on Machine Learning*, Seiten 335–338. Springer-Verlag, Berlin, Heidelberg, New York, 1995.

- [SA96] R. Srikant, R. Agrawal. Mining quantitative association rules in large relational tables. In H. V. Jagadish, I. S. Mumick (Hrsg.), *Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data*, Seiten 1–12, Montreal, Quebec, Canada, Juni 1996. ACM Press.
- [SBMU00] C. Silverstein, S. Brin, R. Motwani, J. D. Ullman. Scalable techniques for mining causal structures. *Data Mining and Knowledge Discovery*, 4(2/3), Seiten 163–192, 2000.
- [SCH98] B. Stein, D. Curatolo, M. Hoffmann. Simulation in FluidSIM. Technical Report tr-ri-98-194, Universität-Gesamthochschule Paderborn, Fachbereich Mathematik/Informatik, 1998.
- [Sha48] C. E. Shannon. A mathematical theory of communication. In *Bell System Technical Journal*, Band 27, Seiten 379–423, Juli 1948.
- [Sho76] E. H. Shortliffe. *Computer Based Medical Consultations: MYCIN*. Elsevier, New York, 1976.
- [SL92] B. Stein, R. Lemmen. ArtDeco: A system which assists the checking of hydraulic circuits. *Workshop for Model-based Reasoning, ECAI-92*, 1992.
- [ST96] A. Silberschatz, A. Tuzhilin. What makes patterns interesting in knowledge discovery. *IEEE Transactions on Knowledge and Data Engineering*, 8(6), Seiten 970–974, Dezember 1996.
- [Ste01] B. Stein. Model construction in analysis and synthesis tasks. Habilitationsschrift, Universität-Gesamthochschule Paderborn, Fachbereich Mathematik/Informatik, 2001.
- [SW76] C. E. Shannon, W. Weaver. *Mathematische Grundlagen der Informationstheorie*. R. Oldenbourg Verlag, München, 1976.
- [Top74] F. Topsøe. *Informationstheorie*. B. G. Teubner, Stuttgart, 1974.
- [TS89] N. H. C. Thuy, P. Schupp. *Wissensverarbeitung und Expertensysteme*. R. Oldenbourg Verlag, München, 1989.
- [Ums92] W. Umstätter. Die evolutionsstrategische Entstehung von Wissen. In Deutsche Sektion der Internationalen Gesellschaft für Wissensorganisation e. V. (Hrsg.), *Fortschritte in der Wissensorganisation*, Band 2, Seiten 1–11, Frankfurt, 1992. Indeks-Verlag.

LITERATURVERZEICHNIS

- [Wac93] I. Wachsmuth et al. Expertensysteme, Planen und Problemlösen. In G. Görz (Hrsg.), *Einführung in die künstliche Intelligenz*. Addison-Wesley Verlag, Bonn, Paris, Reading, 1. Auflage, 1993.
- [WI98] S. M. Weiss, N. Indurkha. *Predictive Data Mining*. Morgan Kaufmann, San Francisco, 1998.
- [Wro98] S. Wrobel. Data Mining und Wissensentdeckung in Datenbanken. *KI 1/1998*, Seiten 6–10, 1998.
- [WZH00] Ke Wang, S. Zhou, Y. He. Growing decision trees on support-less association rules. In *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Seiten 265–269, Boston, MA, August 2000.
- [Zim93] H.-J. Zimmermann. *Fuzzy-Technologien*. VDI-Verlag, Düsseldorf, 1993.
- [ZS92] J. Zeitler, G. Simon. *Physik für Techniker und technische Berufe*. Fachbuchverlag Leipzig, 1992.