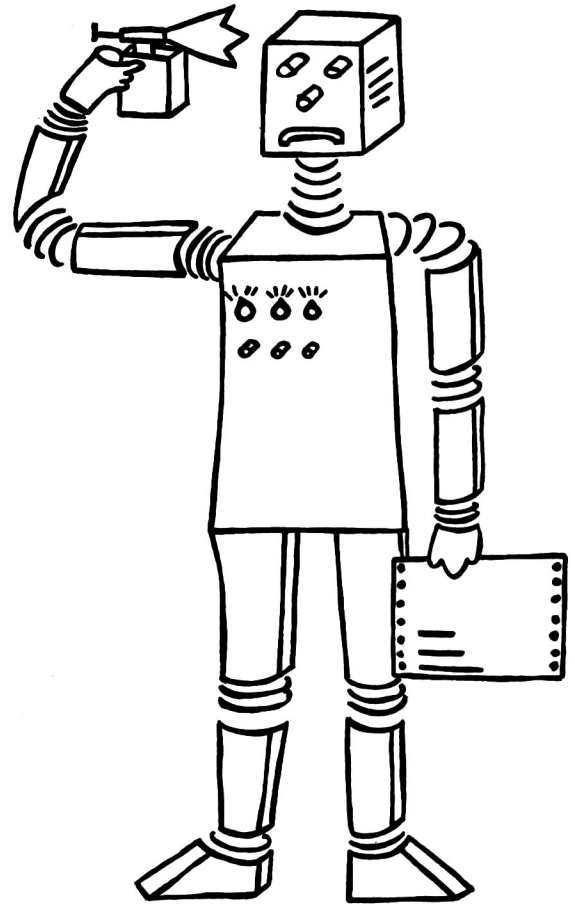


SEKI-Working Paper

Fachbereich Informatik
Universität Kaiserslautern
Postfach 3049
D-6750 Kaiserslautern 1, W. Germany



Vergleichende Untersuchung der
Integrationsproblematik von
Diagnoseexp.systemen und
Tiefenmodellierungssprachen

Simone Gladel-Speicher
SEKI Working Paper SWP-88-04

Inhaltsverzeichnis:

Einleitung	1
Einführung.....	1
Kapitel 1: Vorstellung und Bewertung der drei Expertensysteme Abel, Ohrendiagnosesystem und Med2.....4	
1. Abel	4
1.1 Anforderungen aus der Problemstellung.....	4
1.2 Funktionale Sicht.....	5
1.3 Architektursicht von Abel.....	7
1.3.1 Repräsentation von Wissen.....	7
1.3.1.1 Modelle zur Repräsentation anatomischen Wissens.....	8
1.3.1.2 Physiologisches und pathophysiologisches Wissen.....	10
1.3.1.3 Das patientenspezifische Modell.....	10
1.3.1.4 Die Diagnoseschülen.....	14
1.3.2 Die Zeitkomponente.....	15
1.3.3 Die Erklärungskomponente.....	16
1.3.4 Grober Ablauf des Systems.....	17
1.3.4.1 Initial Formulaton = Erstellen der ersten PSM's.....	18
1.3.5 Mechanismen zur Erstellung der PSM's und der Diagnoseschülen.....	19
1.3.5.1 Aggregation.....	19
1.3.5.2 Elaboration.....	21
1.3.5.3 Projektion.....	22
1.3.5.4 Komponentensumation und Komponentendekomposition.....	22
1.3.6 Unvollständiges und unsicheres Wissen.....	23
1.3.7 Bewertung der PSM.....	24
1.3.8 Bewertung der Krankheitsypothesen.....	25
1.3.9 Kontrollfluß und Kontrollstrukturen.....	25
1.3.10 Möglichkeiten zur Behandlung von Widersprüchen.....	29
1.4 Bewertung von Abel.....	30
1.4.1 Vergleich der Anforderungen aus der Problemstellung mit der funktionalen und der Architektursicht	30
1.4.2 Zusammenfassende Bewertung.....	31
1.5 Anhang zu Abel.....	35
1.5.1 Übersetzung medizinischer Ausdrücke	35
1.5.2 Erklärung von Begriffen in Abel.....	36
2. Expertensystem zur Diagnose von Ohrkrankheiten	
2.1 Anforderungen aus der Problemstellung.....	37
2.2 Funktionale Sicht des Ohrendiagnosesystems.....	38
2.3 Architektursicht des Ohrendiagnosesystems.....	39
2.3.1 Definition der Grundbegriffe.....	39
2.3.2 Ablauf des Ohrendiagnosesystems.....	40
2.3.3 Fokussierung und Abstraktion.....	43
2.3.3.1 Hypothese (Krankheiten)	44
2.3.3.1.1 Fokussierung.....	44
2.3.3.1.2 Abstraktion.....	44
2.3.3.2 Untersuchungen.....	45
2.3.3.2.1 Fokussierung.....	45
2.3.3.2.2 Abstraktion.....	46
2.3.4 Relationen zwischen Symptomen und Krankheiten.....	47
2.3.5 Die Erklärungskomponente.....	47
2.3.6 Die Repräsentation der Zeit.....	48
2.3.7 Unsicheres und unvollständiges Wissen.....	48
2.3.8 Löschen von Belegungen und Recovery.....	49
2.4 Bewertung des Ohrendiagnosesystems.....	50
2.4.1 Vergleich der Anforderungen aus der Problemstellung mit der funktionalen und der Architektursicht	50
2.4.2 Zusammenfassende Bewertung.....	54
3. Med2.....	55
3.1 Anforderungen aus der Problemstellung	55
3.2 Funktionale Sicht von Med2.....	56
3.3 Architektursicht von Med2.....	57
3.3.1 Wissensrepräsentation.....	57
3.3.2 Inferenzstrategie (= Reasoning-Strategie).....	58
3.3.2.1 Database Reasoning.....	58
3.3.2.2 Diagnostic Reasoning.....	59
3.3.3 Repräsentation der Zeit.....	59
3.3.4 Grobdiagnosen (diagnostischer Mittelbau).....	60
3.3.5 Das Working-Memory-Konzept.....	61
3.3.6 Dialogsteuerung.....	62
3.3.7 Bewertungsarten.....	62
3.3.8 Nicht-monotonies Argumentieren.....	63

Vergleichende Untersuchung der
Integrationsproblematik
von Diagnoseexpertensystemen und
Tiefenmodellierungssprachen
am Beispiel verschiedener, bereits
existierender Systeme

Diplomarbeit

Simone Giadel-Speicher
Mat. Nr. 089977

Hiermit versichere ich, daß ich die vorliegende Arbeit selbstständig verfaßt und keine
anderen als die angegebenen Hilfsmittel verwendet habe.

Kaiserslautern, im März 1988

Simone Giadel-Speicher

Prof. Dr. rer. nat. Michael M. Richter

2.4.2	Integration von H1qual1 in die Zeitkomponente.....	130
2.4.3	Integration von H1qual1 in die Erklärungskomponente.....	131
2.4.4	Konsistenzüberprüfungen der Eingabe.....	132
2.4.5	Wertung.....	132
2.5	Integration von H1qual2 in das Expertensystem Med2.....	133
2.5.1	Unterstützung der Differentialdiagnose.....	133
2.5.2	Mehrere gleichzeitig auftretende Diagnosen.....	133
2.5.3	Erklärungen für den Benutzer.....	134
2.5.4	Kann H1qual2 die Erklärungssets ersetzen?.....	134
2.5.5	Rücknahme von Ergebnissen der Simulation.....	134
2.5.6	Wertung.....	135
3.	Integration von H1qual1, H1qual2 und H1qual3 in das Ohrdiagnosesystem.....	136
3.1	Das Ohrdiagnosesystem.....	136
3.2	Allgemeines.....	136
3.2.1	Grobe Skizze für Diagnosehierarchien.....	136
3.3	Integration von H1qual1 in das Ohrdiagnosesystem.....	137
3.4	Integration von H1qual1 in das Ohrdiagnosesystem.....	138
3.4.1	Integration von H1qual1 in die Fokussierung der Hypothesen.....	139
3.4.2	Integration von H1qual1 in die Fokussierung der Untersuchungen.....	139
3.4.3	Integration von H1qual1 in die Erklärungskomponente.....	139
3.4.4	Integration von H1qual1 in die Zeitkomponente.....	140
3.4.5	H1qual1 und Recovery.....	141
3.4.6	H1qual1 und unvollständiges und unsicheres Wissen.....	141
3.4.7	Schlusswort.....	142
3.5	Integration von H1qual2 in das Ohrdiagnosesystem.....	142
3.5.1	Integration von H1qual2 in die einzelnen Arbeitsschritte.....	142
3.5.1.1	Initialphase.....	142
3.5.1.2	Finalphase.....	143
3.5.2	Integration von H1qual2 in die Erklärungskomponente.....	143
3.5.3	Integration von H1qual2 in die Zeitkomponente.....	143
3.5.4	H1qual2 und Recovery.....	144
3.5.5	Unterstützung durch H1qual2 beim Versuch, mehrere vorliegende Krankheiten zu finden.....	145
3.5.6	Allgemeine Bemerkungen.....	145
4.	Vergleich der Integration von H1qual1, H1qual2 und H1qual3 in die drei Expertensysteme Med2, Abel und das Ohrdiagnosesystem.....	146

Kapitel 5: Konzeptentwicklung des neuen Expertensystems zur Fehlerdiagnose in technischen Systemen.....		148
1.	Anforderungen an ein Expertensystem zur Diagnosefindung in technischen Systemen.....	148
2.	Abwägen bisheriger Konzepte.....	149
2.1	Soll das neue Diagnoseexpertensystem assoziativer oder kausaler Art sein?.....	149
2.2	Möglichkeiten zur Verknüpfung assoziativer und kausaler Systeme.....	150
2.3	Welche bisherigen Systeme bzw. Systemteile können für das assoziative und das kausale Modell eingesetzt werden?.....	151
3.	Integration von H1qual2 und des Abstraktionsmechanismus der Ohrdiagnose in das Expertensystem Med2.....	153
3.1	Ansatzpunkte zur Integration von H1qual2 in das Expertensystem Med2.....	153
3.1.1	H1qual2 ersetzt die Lokalisationshierarchie.....	153
3.1.2	Wozu benutzt Abel sein Hintergrundwissen?.....	153
3.1.3	Neue Ideen durch GTD.....	154
3.1.4	Anlehnung an die Denkweise eines Experten.....	155
3.2	Integration des Abstraktionsmechanismus des Ohrdiagnosesystems in das neue Expertensystem.....	156
4.	Integration der Simulation in das neue Expertensystem.....	159
4.1	Erweiterte Interpretation der Ergebnisse einer Simulation in H1qual2.....	159
4.1.1	Mögliche Ergebnisse einer Simulation.....	159
4.1.2	Einschieffiger Debugger für die Simulationsergebnisse.....	160
4.1.3	Zusammenfassung.....	163
4.2	Neue Simulationsmöglichkeiten.....	164
4.3	Bewertung der beiden Simulationsmöglichkeiten.....	165
5.	Funktionale Sicht des neuen Expertensystems.....	166
6.	Architektur des neuen Expertensystems.....	166
6.1	Änderungen von H1qual2 zur Integration in das neue Expertensystem.....	167
6.2	Veränderungen von Med2 zur Integration in das neue Expertensystem.....	171
6.2.1	Veränderungen der Wissensrepräsentation im Vergleich zu Med2.....	171
6.2.2	Konsistenzüberprüfung der Eingabedaten.....	172

3.3.9	Grober Ablauf in Med2	63
3.4	Bewertung des Expertensystems Med2	65
3.4.1	Vergleich der Anforderungen aus der Problemstellung mit der funktionalen und der Architektursicht	65
3.4.2	Zusammenfassende Bewertung	68
Kapitel 2: Vergleich der drei Diagnoseexpertensysteme		70
Kapitel 3: Hiqua1 und Weiterentwicklungen von Hiqua1		77
1.	Die Wissensrepräsentationssprache Hiqua1	77
1.1	Einleitung	77
1.2	Darstellung der von Hiqua1 benutzten Methoden zur Tiefenmodellierung	77
1.2.1	Modelltypen	79
1.2.2	Aggregationstypen	80
1.2.3	Instanzen	83
1.3	Analyse in Hiqua1	86
1.3.1	Allgemeiner Überblick	86
1.3.2	Zeitintervallanalyse von Allen	87
1.3.3	Definitionen mit Erklärungen anhand eines Beispiels	88
1.3.4	Zeitliche Relationen zwischen Intervallen und die entsprechenden ontologischen Relationen	89
1.3.4.1	Zustand-Zustand-Relation	90
1.3.4.2	Ereignis-Zustand-Relation	90
1.3.4.3	Ereignis-Ereignis-Relation	91
1.3.4.4	Momentane Ereignisse und Zustände	93
1.3.5	Relationen, die durch horizontale Verbindungen entstehen	94
1.3.6	Relationen, die durch vertikale Verbindungen entstehen	95
1.4	Pro und Contra Hiqua1	96
1.5	Was bietet Hiqua1 zur Unterstützung der Diagnostik-Expertensysteme?	97
2.	Hiqua1	101
2.1	Folgen einer defekten Komponente oder fehlerhaften Zeitbeziehung	101
2.1.1	Welche Folgen kann eine defekte Komponente verursachen?	101
2.1.2	Welche Folgen kann eine inkorrekte Zeitbeziehung verursachen?	104
2.2	Welche Komponente könnte die Ursache der auftretenden Fehler sein?	104
2.3	Integration des Hiqua1-Modells in das Expertensystem	107

2.3.1	Aufbau der Verbindungen zwischen Symptomen und Komponenten	107
2.3.1.1	Wo sollen die Verbindungen zwischen Symptomen und Komponenten gespeichert werden?	108
2.3.1.2	Verbesserungsvorschläge	108
2.3.2	Aufbau der Verbindungen zwischen Komponenten des Hiqua1-Modells und Diagnosen	110
2.3.3	Aufbau der Verbindungen zwischen Komponenten des Hiqua1-Modells und Therapien	112
3.	Hiqua2	113
3.1	Möglichkeiten, ein defektes Teil in das korrekte Hiqua1-Modell einzusetzen	113
3.1.1	Anpassung aller Vorfahren der defekten Komponenten	114
3.1.2	Die Vorfahren der defekten Komponenten werden als "leere" Hülle dargestellt	115
3.1.3	Wartung der beiden Möglichkeiten	116
3.2	Simulation von Teilen des Hiqua1-Modells	116
3.3	Pro und Contra Hiqua2	117
Kapitel 4: Integration von Hiqua1, Hiqua11 und Hiqua2 in die einzelnen Expertensysteme		118
1.	Integration von Hiqua1, Hiqua11 und Hiqua2 in Abel	119
1.1	Das Expertensystem Abel	119
1.2	Repräsentationsmechanismen in Abel und ihr Zusammenwirken mit Hiqua1	119
1.3	Gegenüberstellung der Methode von Hiqua1 und Abel	120
1.4	Integration von Hiqua1 in Abel	121
1.5	Integration von Hiqua11 in Abel	122
1.6	Integration von Hiqua2 in Abel	123
2.	Integration von Hiqua1, Hiqua11 und Hiqua2 in das Expertensystem Med2	124
2.1	Die Expertensystemshel Med2	124
2.2	Repräsentation der Anatomie in Med2	124
2.3	Integration von Hiqua1 in Med2	126
2.3.1	Ersetzen der Lokalisationshierarchie durch das Hiqua1-Modell	126
2.3.2	Die Erklärungskomponente	127
2.3.3	Konsistenzüberprüfung der Eingabe	127
2.4	Integration von Hiqua11 in Med2	129
2.4.1	Integration von Hiqua11 in den Arbeitsablauf	129

6.2.3	Recovery im neuen Expertensystem.....	174
6.2.4	Erweiterung der Questionsetselektion.....	175
6.3	Abstraktionsmechanismus	175
6.4	Grober Ablauf des neuen Expertensystems.....	176
Kapitel 6: Ausblick.....		178
1.	Grundkonzept für das erweiterte Hiqual-System.....	178
2.	Vorschläge zur Erweiterung des neuen Expertensystems.....	179
2.1	Integration einer Falldatenbank.....	179
2.2	Integration von Wissen über externe Vorgänge	180
Literaturverzeichnis		182

Einleitung.

Ausgangspunkt der vorliegenden Arbeit ist ein Vergleich zwischen den Expertensystemen Med2 [Puppe 1986a], dem Ohrendiagnosesystem [Heinen 1985, Wetter 1983, Heinen et al. 1985] und Abel [Paul 1981]. Diese drei Systeme werden auf ihre Integrationsfähigkeit hinsichtlich der Wissensrepräsentationssprache Hiqual [Voss 1986] untersucht. Da Hiqual den gewünschten Anforderungen nicht entspricht, werden zwei Erweiterungen von Hiqual, Hiqual1 und Hiqual2, konstruiert deren Integrationsfähigkeit bezüglich der einzelnen Expertensysteme wiederum untersucht und verglichen wird.

Aufbauend auf obigen Untersuchungen wird ein "neues" Expertensystem zur Fehlerdiagnose technischer Systeme vorgestellt, das die Grundstruktur von Med2 besitzt und in das Hiqual2 und der Abstraktionsmechanismus der Ohrendiagnose integriert sind. Dadurch wird die Effizienz assoziativer Systeme mit der Robustheit kausaler Systeme vereint. Die Verbindung beider Systeme realisiert der Abstraktionsmechanismus der Ohrendiagnose. Med2 bildet das Gerüst für das Expertensystem. Die Grundstrukturen des Systems, wie Questionsets, Working-Memory-Konzept und das nicht-monotone Argumentieren bleiben erhalten. Die Verantwortung für die Kontrollstruktur trägt Med2. Das kausale Modell Hiqual2 repräsentiert wichtiges Hintergrundwissen und berechnet den Weitertransport eines Fehlers oder die Ursache mehrerer komponentenbezogener Symptome in einfachen Fällen. Zusätzlich ist in Hiqual2 die Simulation defekter Komponenten möglich.

Einführung.

Expertensysteme sind Computerprogramme, die die Fähigkeiten eines Experten durch Anwendung von Problemlösungsmethoden auf Eingabedaten und gespeichertes Wissen simulieren oder unterstützen. Expertensysteme bewähren sich vor allem dort, wo es um komplexes, aber dennoch begrenztes und gut strukturiertes Spezialwissen geht. Ein Expertensystem besteht im Wesentlichen aus folgenden Komponenten:

- Dialogkomponente

Der Benutzer gibt ihm bekannte Werte ein. Das Expertensystem kann weitere Fragen an den Benutzer stellen und die Lösung des Problems mit der entsprechenden Erklärung ausgeben.

- Wissensbasis

Grundlage eines jeden Expertensystems ist eine Wissensbasis (z. B. in Form von Regeln oder Modellen), die mit Hilfe von Experten erstellt wird.

- Problemlösungskomponente

Eine systematische Ableitung von Ergebnissen aus der Wissensbasis und den eingegebenen Daten ermöglicht die Problemlösungskomponente. Sie ist von dem jeweiligen Anwendungsgebiet weitgehend unabhängig.

- Erklärungskomponente

Um die Lösung des Expertensystems zu unterstützen, wird dem Benutzer die Herleitung des Ergebnisses dargelegt.

Je nach ihrer Anwendung werden die Expertensysteme in verschiedene Klassen eingeteilt:

- Die Expertensysteme der *assoziativen Diagnostik* oder *Selektion* erhalten als Eingabe Befunde oder Meßwerte und geben als Output wiedererkannte Muster aus.
- In Expertensystemen der Problemklasse *Design* werden Anforderungen eingegeben und Objekte, die den Anforderungen genügen, zurückgegeben.
- *Planungssysteme* kennen den Ausgangs- und den Zielzustand. Daraus erarbeiten sie eine Sequenz von Aktionen, die den Ausgangszustand in den Zielzustand überführt.
- Expertensysteme zur *Simulation* leiten von einem Anfangszustand die (zukünftigen) Folgezustände ab.

Die in dieser Arbeit diskutierten Expertensysteme Med2 und das Ohrendiagnosesystem gehören zu der Gruppe der assoziativen Diagnostik. Abel wird der Problemklasse Simulation zugeordnet.

Trotzdem befassen sich alle drei Systeme mit diagnostischem Problemlösen. Diagnostik ist die Fähigkeit Krankheiten von Menschen beziehungsweise Mängel technischer Systeme zu erkennen. Innerhalb der Diagnostik unterscheidet man zwischen statistischer, assoziativer und modellbasierter Diagnostik.

Die *statistische* Diagnostik behandelt nur das Problem der Diagnosebewertung mit unsicherem Wissen. Sie zeichnet sich durch eine hohe Objektivierbarkeit aus, eignet sich aber nur für sehr kleine beziehungsweise hochgradig formalisierte Anwendungsbereiche. Die Qualität der Erklärungen und die Lösbarkeit komplexer Probleme sind unzureichend. Expertensysteme der *assoziativen* Diagnostik arbeiten überwiegend mit Symptom-Diagnose-Assoziationen und weisen eine hohe Effektivität und ein breites Anwendungsspektrum auf. Beispiele sind Med2 und das Ohrendiagnosesystem.

Die Wissensbasis *modellbasierter* Diagnostikexpertensysteme besteht hauptsächlich aus funktionalen [Davis 84] oder pathophysiologischen (Abel) Modellen. Die Lösung komplexer Probleme, sowie die Möglichkeit der Erklärungskomponente sind die Vorteile dieser Systeme. Die Effizienz ist im Vergleich zu assoziativen Systemen weit geringer.

Ausgangspunkt der Arbeit sind die Expertensysteme Med2, das Ohrendiagnosesystem und Abel, die alle drei mehr oder weniger stark auf die medizinische Diagnostik zugeschnitten sind. Das Anwendungsgebiet des zu entwerfenden Expertensystems ist die Fehlerdiagnose komplexer technischer Systeme, wie z.B. CNC-Bearbeitungszentren. Die Vorteile technischer Diagnosesituationen gegenüber medizinischen, wie die Möglichkeit, Teile der Maschine probeweise auszutauschen, müssen beim Entwurf des neuen Expertensystems beachtet werden.

In Kapitel 1 werden die drei Diagnoseexpertensysteme, Med2, das Ohrendiagnosesystem und Abel bezüglich der Anforderungen aus der Problemstellung, ihrer funktionalen und ihrer Architektursicht betrachtet.

Anschließend erfolgt in Kapitel 2 ein Vergleich dieser Expertensysteme. Kapitel 3 beschäftigt sich mit der Wissensrepräsentationssprache Hiqua1. Da Hiqua1 den Anforderungen der Simulation einer defekten Komponente nicht genügt, werden zwei Modifikationen von Hiqua1 konstruiert. Hiqua1 stellt Verbindungen zum Expertensystem her und verfügt über zwei Algorithmen zur Berechnung von Fehlern. Hiqua2 kann unter bestimmten Voraussetzungen die Simulation einer defekten Komponente durchführen.

Kapitel 4 behandelt die Integration von Hiqua1 in Expertensysteme. Nacheinander wird die Integration von Hiqua1 und seiner Modifikationen in die vorgestellten Expertensysteme vorgenommen und bewertet.

In Kapitel 5 wird ein Konzept für das neue Expertensystem entwickelt. Grundidee ist die Kombination von assoziativem Wissen in Form von Med2 und modellbasiertem Wissen durch Hiqua2. Verschiedene Ansatzpunkte für eine solche Kombination werden auf ihren Nutzen für das neue Expertensystem untersucht. Eine Verbesserung der Simulation in Hiqua2 ist Thema des darauffolgenden Abschnitts. Abschließend wird die endgültige Struktur des neuen Expertensystems dargestellt.

Kapitel 1 : Vorstellung und Bewertung der drei Expertensysteme Abel, Ohrendiagnosesystem und Med2

1. A B E L

Abel ist ein modellbasiertes Diagnostik-Expertensystem zur Simulation von Störungen im Säure/Base- und Elektrolythaushalt. Das Ziel ist es, Abel so zu entwickeln, daß es wie ein Experte mit allen Aspekten der Diagnose und Behandlung arbeiten kann (Anlehnung an die Denkweise des Menschen).

In Abel hat man Darstellungsmethoden für zugrundeliegendes medizinisches Denken entwickelt. Dazu wird eine Reihe von Konzepten vorgestellt, durch die die Performanz des KI-Programms erhöht werden soll. Mit Hilfe eines pathophysiologischen Modells beschreibt Abel die Krankheit auf verschiedenen Abstraktionsebenen. Die unterste Ebene enthält pathophysiologisches Wissen über Krankheiten, die oberste Ebene enthält klinisches Wissen über Krankheitszusammenhänge.

Die zentrale Datenstruktur ist das PSM (=patientenspezifisches Modell). Ein PSM repräsentiert eine Menge von pathophysiologischen Zuständen einschließlich ihrer Schweregrade und der kausalen Beziehungen zwischen ihnen, die insgesamt eine konsistente Erklärung für die beobachteten Symptome des Patienten sind.

1.1 Anforderungen aus der Problemstellung.

Problemstellung:

Unterstützung der Diagnose und Therapie auf dem Gebiet der Säure/Base- und Elektrolystörungen.

Anforderungen:

- 1) Laborwerte und die übrigen Eingaben des Benutzers sollen bezüglich ihrer Konsistenz überprüft werden.
- 2) Aus allen Eingaben soll das exakte Problem formuliert werden.
- 3) Größere Probleme sollen in Teilprobleme aufgeteilt werden.
- 4) Die korrekte Säure/Base- Elektrolystörung soll gefunden werden, auch wenn die Lösung aus mehreren Säure/Base- Elektrolystörungen besteht. Anschließend soll eine Therapie zugeordnet werden.

- 5) Auch bei unvollständigem oder unsicherem Wissen soll das Problem möglichst gut gelöst werden.
- 6) Folgesitzungen sollen entsprechend ausgewertet werden.
- 7) Ähnliche Diagnosen sollen unterschieden werden.
- 8) Sich über die Zeit ändernde Laborwerte sollen aggregiert werden.
- 9) Fehlerhafte Eingaben (vertauschte Laborwerte) sollen einfach und schnell zurückgezogen werden.
- 10) Hintergrundwissen über anatomisches, physiologisches und pathophysiologisches Wissen sollen geeignet dargestellt werden.
- 11) Symptome und Diagnosen sollen in Gruppen zusammengefaßt werden.
- 12) Gleichzeitige Verarbeitung von klinischem und pathophysiologischem Wissens soll ermöglicht werden.
- 13) Die Eingabekomponente soll für einen schnellen Input von Laborwerten geeignet sein.
- 14) Dialog: Expertensystem <-> Benutzer
 - Die Antworten des Expertensystems sollen entsprechend dem Wissen des Benutzers modifiziert werden und für den Benutzer verständlich sein.
 - Die zugrundeliegende Hypothese soll erklärt werden.
 - Mögliche Alternativen und wichtige Begründungen sollen ausgegeben werden.
 - Auch während des Diagnoseprozesses sollen Fragen möglich sein.
 - Risiken für Entscheidungsvorschläge sollen angegeben werden.
- 15) Eine Notfalldiagnostik soll möglich sein.

1.2 Funktionale Sicht von Abel.

- 1) Konsistenzüberprüfung der Eingabe.
- 2) Kontrollmechanismus zur Unterteilung des Hauptzieles in mehrere Unterziele.
- 3) Differentialdiagnose.
- 4) Mechanismen zur Repräsentation von unvollständigem und unsicherem Wissen.
- 5) Simulationsmechanismus für die Diagnose mit
 - Rückkopplungsschleifen
 - mehreren Abstraktionsebenen
- 6) Recovery-Mechanismus.
- 7) Zeitkomponente: Zeitaggregation
Zeitindizierung der Zustände und der Untersuchungen
- 8) Klassifikationsmechanismus für Diagnosen

9) Widerspruchsbandler:

- Rechtfertigungsmechanismus
 - FSM-Debugger
 - zielorientierter Widerspruchsbandler
- ## 10) Wissensbasis:
- Hierarchische Repräsentation der Anatomie.
 - Modelle für jede mögliche Krankheit
 - kausale Beziehungen
- ## 11) Partnermodellierter Erklärungskomponente
- Thinking-loud während des Diagnoseprozesses
- ## 12) Mechanismen zur Notfalldiagnostik.

1.3 Architektursicht von Abel.

1.3.1 Repräsentation von Wissen.

Eine Krankheit kann als eine Änderung des normalen Zustandes eines Patienten und seiner Körperfunktionen beschrieben werden. Aus diesem Grund braucht man Formalismen zur Darstellung der normalen Zustände, der Zustandsänderungen und der normalen und abnormalen Funktionen und ihrer Wechselwirkungen.

Abel bietet dazu folgende Möglichkeiten:

(1) *Anatomie-Komponente:*

part-of Hierarchie für Organsysteme (s.Abb. 1.1)

enthalten-in Relation (s.Abb. 1.2)

grobe anatomische Relationen (s.Abb. 1.3)

connected-to Relation für den Materialfluß (s.Abb. 1.4)

(2) *Physiologische Komponente:*

Das physiologische Wissen kann Antworten auf folgende Fragen geben:

Welche Flüssigkeiten kommen im Körper an welchen Stellen vor?

Aus welchen Bestandteilen bestehen die Flüssigkeiten?

Wo findet ein Austausch von Komponenten statt?

Der kausale Link kann Auskunft darüber geben, wie sich die Ab- oder Zunahme bestimmter Bestandteile einer Flüssigkeit auswirken. (physiologische Abhängigkeiten).

(3) *Pathophysiologische Komponente:*

Einfaches Wissen über Krankheitsätiologien. (s.Abb. 1.5)

Krankheiten werden nach anatomischem, physiologischem, zeitlichem oder ätiologischem Wissen klassifiziert.

Kausale Relationen, die beschreiben, wie Änderungen eines gegebenen Zustandes andere Zustände verändern. (kausaler Link)

Es ist auch wichtig gemeinsam auftretende Konstellationen anormaler Zustände als spezielle zusammengesetzte Situationen zu erkennen. Dies wird durch höhere Abstraktionsstufen ermöglicht.

1.3.1.1 Modelle zur Darstellung anatomischen Wissens.

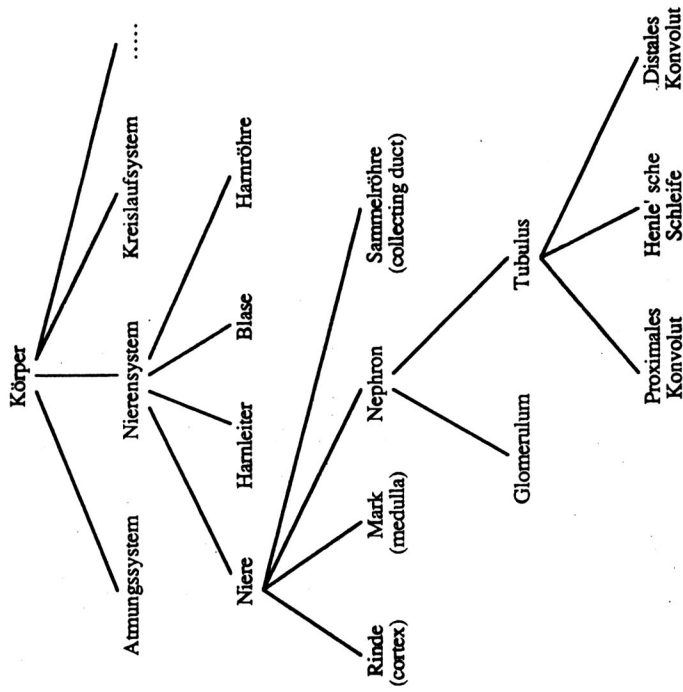


Abb. 1.1: Part-of Hierarchie von Organsystemen.

Der menschliche Körper besteht aus einem Nierensystem, einem Atmungssystem, einem Kreislaufsystem u.v.a.m. Das Nierensystem wiederum besteht aus der Niere, dem Harnleiter, der Blase und der Harnröhre.

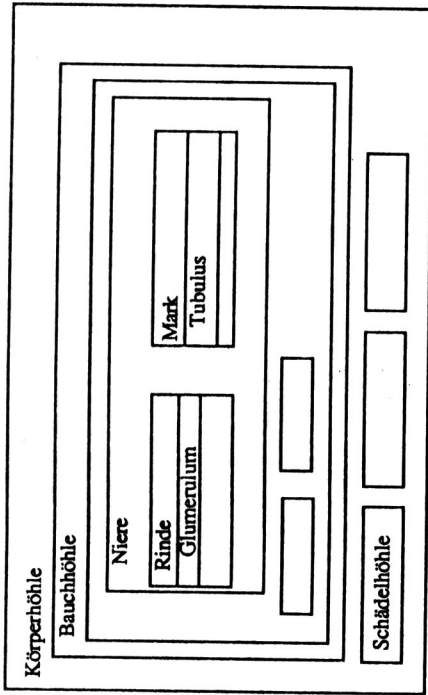


Abb. 1.2: Enthalten-in Relation Der menschliche Körper besteht unter anderem aus der Schädelhöhle und der Bauchhöhle. In der Bauchhöhle liegt unter anderem die Niere. Diese wiederum besteht aus Rinde und Mark. In der Rinde ist das Glomerulum enthalten, während die Tubuli im Mark enthalten sind. (siehe Unterschied zur Part-of Relation)

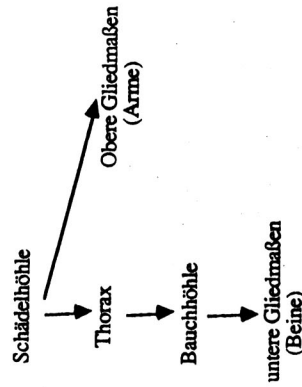


Abb. 1.3: Grobe anatomische Beziehungen Unterhalb der Schädelhöhle befinden sich im aufrechten Zustand des Menschen der Thorax und die Arme. Unterhalb des Thorax folgt die Bauchhöhle danach die unteren Gliedmaßen, die Beine.

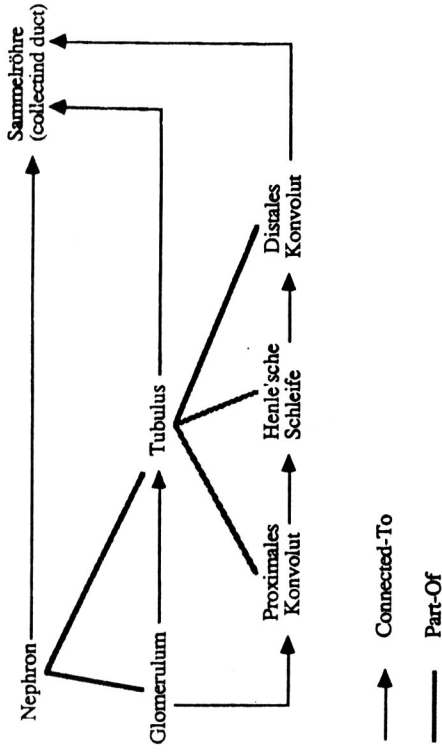


Abb. 1.4: Materialfluß Relationen (connected-to)

Das Nephron besteht aus dem Glomerulum und dem Tubulus. 'Material' fließt vom Nephron zur Sammelröhre. Oder genauer vom Glomerulum über den Tubulus zur Sammelröhre.

Die einzelnen Ebenen der Materialflußrelationen entsprechen denen der part-of Hierarchie. Zusätzlich gibt es horizontale 'connected-to' Verbindungen. Je tiefer die Ebene ist, desto genauer wird der Materialfluß beschrieben.

1.3.1.2 Physiologisches und pathophysiologisches Wissen.

Physiologisches Wissen wird in Abel besonders für Flüssigkeiten, Elektrolyt- und Säure-Basestörungen dargestellt. Welche Flüssigkeiten kommen wo im Körper vor?

Aus welchen Bestandteilen bestehen die Flüssigkeiten?
Wo findet der Austausch von Komponenten statt?

Beispiel: Definition der Kalium-Konzentration im Blut.

Die Kalium-Konzentration im Blut (K-serum) ist gleich der Konzentration der Kalium-Ionen in der außerezellulären Flüssigkeit (ecf). Serum-K kann die Werte niedrig, normal und hoch annehmen. Zu jedem dieser Bereiche wird ein Default-Wert, ein Bereich und eine Fehlergrenze angegeben.

Darstellung pathophysiologischen Wissens über Krankheitsätiologien.

Diese Darstellung ist wichtig, da Krankheiten mit gleicher Ursache oft auch ähnliche Symptome aufweisen.

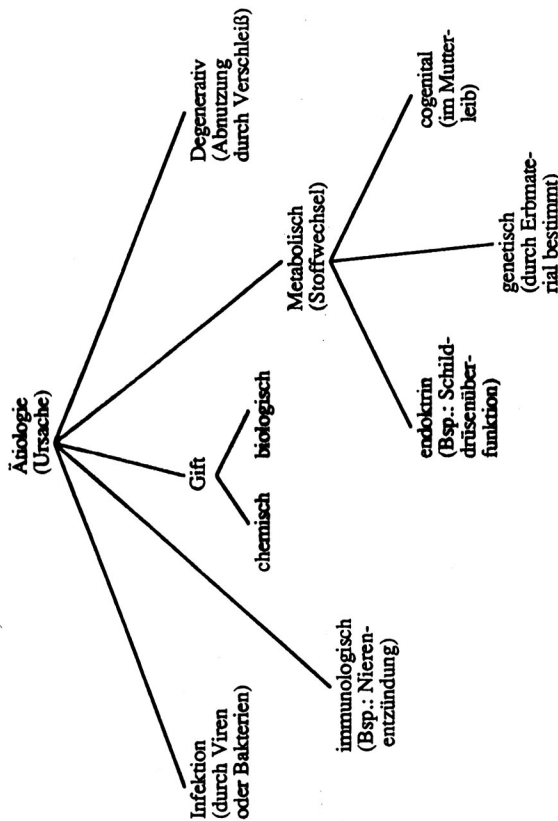


Abb. 1.5: Modell zur Darstellung des ätiologischen (ursächlichen) Wissens.

Diese Darstellung ist deshalb so wichtig, da Krankheiten mit gleicher Ursache oft auch ähnliche Symptome aufweisen.

Klassifikation der Krankheiten.

Abel benutzt anatomisches, physiologisches, ätiologisches und zeitliches Wissen, um eine klassifizierende Hierarchie aufzubauen. Das anatomische, ätiologische und auch das physiologische Wissen ist jeweils hierarchisch organisiert. Daraus wird dann die hierarchische Organisation der Krankheiten abgeleitet.

z.B. Nierenversagen durch toxische Medikamente:

- Nierenkrankheit = anatomisch : Nierensystem
- Nierenversagen = physiologisch : wenig Urinabgabe
- akutes Nierenversagen = zeitlich : akute Merkmale
- durch Medikamente verursachtes akutes Nierenversagen = ätiologisch : chemisch/toxisch

Die Krankheiten können aber auch anhand einer 'Hierarchieart' klassifiziert werden:

- Nierenkrankheit = anatomisch : Nierensystem
- Nephritis = anatomisch : Nephron
- Glomerulo-Nephritis = anatomisch : Glomerulus

Der kausale Link.

Der kausale Link kann physiologische und pathophysiologische Abhängigkeiten darstellen. Ein kausaler Link spezifiziert eine kausale Relation zwischen der Ursache (Vorgänger) und der Wirkung (Nachfolger).

In bisherigen Expertensystemen bestimmen die Links einen Kausaltyp (z.B. wird verursacht durch, ist Komplikation von). Jedem Link wird zusätzlich eine Zahl zugeordnet, die die Wahrscheinlichkeit des Links ausdrücken soll. Diese Darstellungsform ist viel zu einfach, weil eine kausale Beziehung noch von vielen anderen Faktoren abhängen kann.

Beispiel:

Wenn der Durchfall schwer ist und länger als zwei Tage dauert
dann
wenn der Patient noch keine Therapie zum Ersatz der Flüssigkeit bekommen hat
dann hat der Patient wahrscheinlich eine mittelschwere Dehydratation
andernfalls hat der Patient leichte Dehydratation

Ein kausaler Link wird durch eine multivariate Relation zwischen den Attributen der Ursache und des Effekts, dem Kontext und den Voraussetzungen, die die Relationen beschränken, beschrieben.

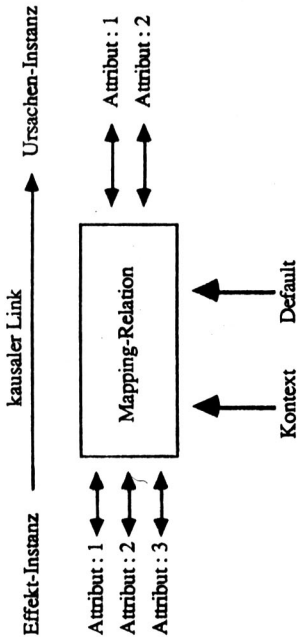


Abb. 1.6: Schematische Beschreibung des kausalen Links.

Kontext: gibt den diagnostischen Kontext, d.h. nicht lokale Faktoren an

Default: Vermutungen; was nicht explizit angegeben wurde, wird als 'normal' angenommen und evtl. später zurückgezogen.

(nicht alle Links sind unbedingt kausal, manche sind auch nur Assoziationen)

1.3.1.3 Das patientenspezifische Modell (PSM)

Das PSM ist ein kausales Modell, das aus mehreren Stufen besteht. Jede Stufe ist ein semantisches Netz, das ein Netzwerk von Relationen zwischen Krankheiten und Befunden beschreibt. Die Krankheit des Patienten wird also auf mehreren verschiedenen Detaillierungsstufen beschrieben (Die Implementierung besitzt 5 Stufen, alle Beispiele beziehen sich aber auf drei Stufen). Ein PSM enthält alle Krankheiten und Befunde, die beobachtet oder gefolgert werden (keine Hypothesen). Verschiedene PSM's schließen sich gegenseitig aus.

Jeder Knoten repräsentiert einen normalen oder anormalen Zustand eines physiologischen Parameters. Jeder Link repräsentiert eine meist kausale Relation zwischen den Zuständen.

1.3.1.4 Die Diagnosehüllen.

Zusätzlich zum PSM werden Diagnosehüllen eingeführt (diagnostic closure). Diese enthalten medizinisches Wissen, das in der jeweiligen diagnostischen Situation von Bedeutung ist. Jedem PSM können mehrere Diagnosehüllen zugeordnet werden.

Eine Diagnosehülle wird gebildet durch:

- hypothetische Vorwärtsprojektion der Zustände eines PSM, um die Konsequenzen aufzuzeigen, die von den Zuständen des PSM vorhergesagt werden.
- Rückwärtsprojektion der unerklärten Zustände des PSM, um Krankheiten zu identifizieren, die diese Zustände erklären können.

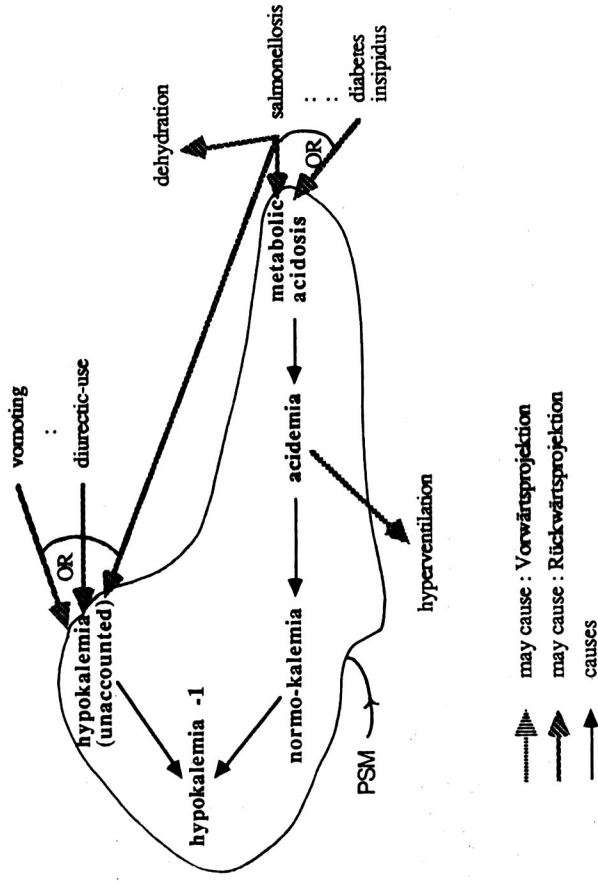


Abb. 1.7: Beispiel einer Diagnosehülle.

Vorteile der Diagnoshüllen:

- Das medizinische Wissen ist kein starres Muster irgendwo in der Wissensbasis, sondern lokal zu der jeweiligen Diagnose-Situation.
- Strukturierungsmöglichkeit: Die Diagnosealternativen sind in Gruppen eingeteilt, wobei jede Gruppe mit der teilweisen Erklärung der Krankheit des Patienten konsistent ist (PSM). Die verschiedenen Gruppen stellen Alternativen dar, die mit sehr unterschiedlichen möglichen Erklärungen konsistent sind.
- Das Erkennen von Widersprüchen ist leichter möglich. Man überprüft, ob die neue Information mit einer Diagnoshülle konsistent ist.
- Die Erklärungskomponente kann durch das 'thinking loud' verbessert werden.
- Die Strukturierung der Diagnoseplanung wird erleichtert.

1.3.2 Die Zeitkomponente.

Die Zeitkomponente in Abel ist wenig ausgebaut. Nur an drei Stellen wird versucht die Zeit zu integrieren.

(1) Zeitliche Aggregationen.

Jede Beobachtung liefert uns einen Schnappschuß über den Zustand eines Patienten. Während des Verlaufs einer Krankheit entstehen eine ganze Folge solcher Schnappschüsse. Mit Hilfe der zeitlichen Aggregationen kann man den Prozess wieder zusammenfügen und ihn als ein Konzept im PSM beschreiben. Desweiteren kann eine Folge von Werten eines Zustandes zu einer einzigen Beschreibung zusammengefaßt werden. Über die erwartete Wertefolge kann man auch Hypothesen aufstellen.

Beispiel:

Hypo-Natrium	Na	130.0	1.Tag
	Na	132.0	2.Tag
Hyper-Natrium	Na	145.0	3.Tag
	Na	142.0	4.Tag
Normal-Natrium	Na	140.0	5.Tag

(2) Fortgesetzte Behandlung des Patienten.

Für Folgeuntersuchungen wird das PSM von einer Sitzung bis zur nächsten aufbewahrt und liefert somit die Grundlage für weitere Interaktionen.

(3) Zeit-Attribute.

Jedem pathophysiologischen Zustand werden die Attribute Schweregrad, Dauer und Startzeit zugeordnet. Dadurch ist eine Möglichkeit gegeben Zeitattribute abzuspeichern.

1.3.3 Die Erklärungskomponente.

In Abel wurde keine neue Erklärungskomponente entwickelt, sondern die von Swartout entwickelte Komponente integriert. Die Abstraktionsebenen bilden die beste Voraussetzung für eine Erklärungskomponente, die den Vorkenntnissen des Benutzers angepaßt ist (Partnermodellierung). Der Benutzer kann also den für ihn geeigneten Detaillierungsgrad wählen.

Beispiel:

(1) Erklärungen auf der klinischen Stufe:

Gegeben sei ein 40 Jahre alter 70.0kg schwerer männlicher Patient mit mittelschwerem Durchfall. Seine Elektrolytwerte sind:

Na: 140.0 HCO₃: 15.0 Agap: 13.0 K: 3.0 CL: 115.0
pCO₂: 30.0 pH: 7.32

Der Durchfall verursacht eine mittelschwere metabolic acidosis (beim Stoffwechsel entstandene krankhafte Übersäuerung des Blutes), die wiederum eine schwache acidemia (Übersäuerung des Blutes) verursacht. Die acidemia und der Durchfall verursachen schwache hypokalemia (Kalium-Konzentration ist erniedrigt) und die acidemia verursacht hyperventilation (Hyperventilation). Alle Befunde können erklärt werden.

(2) Erklärungen auf der mittleren Stufe:

Gegeben sei ein 40 Jahre alter 70.0kg schwerer männlicher Patient mit mittelschwerem Durchfall. Seine Elektrolytwerte sind:

Na: 140.0 HCO₃: 15.0 Agap: 13.0 K: 3.0 CL: 115.0
pCO₂: 30.0 pH: 7.32

Der Durchfall verursacht einen mittelschweren Verlust der lower-GI-Flüssigkeit, wodurch mittelschwere metabolic acidosis (beim Stoffwechsel entstandene krankhafte Übersäuerung des Blutes) verursacht wird. Die metabolic acidosis verursacht zusammen mit der mittelschweren hypocapnia (verminderte Blut-CO₂-Spannung bei Acidose) eine mittelschwere hypobicarbonatemia (zu geringe Bicarbonat-Konzentration). Diese hypobicarbonatemia verursacht zusammen mit der hypocapnia eine schwache acidemia (Übersäuerung des Blutes). Die acidemia und der Verlust der lower-GI Flüssigkeit verursachen schwache hypokalemia (zu geringe Kalium-Konzentration) und die acidemia verursacht hypocapnia. Die acidemia verursacht auch hyperventilation (Hyperventilation). Alle Befunde können erklärt werden.

(3) Erklärungen auf der pathophysiologischen Stufe:

Gegeben sei ein 40 Jahre alter 70.0kg schwerer männlicher Patient mit mittelschwerem Durchfall. Seine Elektrolytwerte sind:

Na: 140.0 HCO3: 15.0 Agap: 13.0 K: 3.0 CL: 115.0
pCO2: 30.0 pH: 7.32

Der mittelschwere Verlust der lower-GI-Flüssigkeit, die erniedrigte renal HCO3 threshold (HCO3-Schwelle in den Nieren) und die normale HCO3-Pufferbindung verhindern eine weitere Veränderung des HCO3. Dadurch, daß keine weitere HCO3-Veränderung stattfindet, entsteht niedriges ecf HCO3, welches eine Erniedrigung des serum HCO3 (Konzentration des HCO3 im Serum) verursacht. Das erniedrigte serum HCO3 und das erniedrigte serum pCO2 verursachen zusammen einen Abfall des pH-Wert des Serums. Der tiefe serum-pH Wert verursacht, daß Kalium aus den Zellen austritt, wodurch die Atemungsgeschwindigkeit erhöht wird. Die erhöhte Atemungsgeschwindigkeit verursacht einen Abfall des serum pCO2, welches eine normale HCO3 Pufferbindung verursacht. Das niedrige serum pCO2 verursacht auch eine verringerte HCO3 Schwelle in den Nieren und die erhöhte Atemungsgeschwindigkeit verursacht eine erhöhte ventilation (Hyperventilation). Der Verlust der lower-GI-Flüssigkeit und das Austreten von Kalium aus den Zellen führt zu einem Kaliumverlust. Der Kaliumverlust verursacht einen Abfall des ecf-K, welches einen Abfall des serum K verursacht. Alle Befunde können erklärt werden.

1.3.4 Grober Ablauf des Systems.

(1) Initial formulation.

Die ersten PSM's (=patientenspezifisches Modell) werden erstellt.

(2) Bewertung der PSM.

(3) Die Diagnosehüllen für die 'wahrscheinlichsten' PSM's werden berechnet und bewertet.

(4) Hat kein PSM eine Diagnosehülle oder liefert ein PSM eine vollständige und zusammenhängende Rechtfertigung für die Krankheit des Patienten, so ist die Diagnosephase beendet.

(5) Diagnostisches Informationssammeln.

In Abhängigkeit von der Zahl der Diagnosehüllen wird ein Confirm- oder ein Differentiate-Ziel auf der obersten Ebene formuliert. Mit Hilfe der Diagnosestrategien wird dieses Ziel schrittweise in einfachere Ziele zerlegt, bis einzelne Fragen formuliert werden können.

(6) 'Überholung' des PSM.

Wird ein neuer Befund bekannt, so wird dieser in jedes PSM eingefügt. Entstehen in einem PSM oder einer Diagnosehülle Inkonsistenzen durch diesen Befund, so wird das PSM beziehungsweise die Diagnosehülle gestrichen.

(7) Weiter mit Schritt (2).

1.3.4.1 Initial Formulation = Erstellen der ersten PSM's.

(1) Eingabe aller Laboraten (z.B.: Natrium, Kalium, Chlor, HCO3) und anfänglicher Befunde wie Alter oder Geschlecht.

(2) Alle möglichen Säure/Base-Störungen, die mit den obigen Daten konsistent sind, werden aufgezählt und nach Komplexität (=Anzahl der Komponenten, die zur Erklärung der Störungen notwendig sind) und Schweregrad geordnet (Die Säure/Base-Störungen erhält man über ein in der Medizin bekanntes Diagramm).

(3) Für jede mögliche Säure/Base-Störung wird ein PSM gebildet:

- Laboraten werden im Zusammenhang mit der jeweiligen Säure/Base-Störung interpretiert.
- Auf der klinischen Stufe werden die Knoten instanziiert, die der jeweiligen Säure/Base-Störung entsprechen.
- Diese Knoten werden dann bis auf die pathophysiologische Stufe elaboriert.
- Zusammen mit den Laboraten wird die pathophysiologische Stufe ausgebaut.
- Dieses Wissen wird dann wieder bis zur klinischen Stufe aggregiert.

1.3.5 Mechanismen zur Erstellung der PSM's und der Diagnosehüllen.

1.3.5.1 Aggregation.

Die Aggregation erlaubt die Zusammenfassung der Daten von einer niedrigeren zu einer höheren Abstraktionsebene. Man unterscheidet zwei Arten der Aggregation.

A. Fokale Aggregation.

Mit Hilfe der fokalen Aggregation können ein zentraler Knoten und die ihn umgebenden Beziehungen auf einer höheren Stufe zu einem einzigen Knoten zusammengefaßt werden. Ein Knoten heißt zentraler Knoten (oder fokus-Knoten), falls folgende Bedingungen erfüllt sind:

- Der Knoten ist in der medizinischen Wissensbasis der fokus-Knoten von mindestens einem Knoten auf der entsprechenden Stufe.
 - Im PSM existiert mindestens schon ein Knoten auf der höheren Ebene oder er kann instantiiert werden.
 - Die Aggregation ist mit der existierenden Struktur des PSM nicht inkonsistent.
- Gibt es mehrere Möglichkeiten, die fokale Aggregation durchzuführen, so wird gewartet, bis zusätzliche Information die Zweideutigkeit auflöst.

B. Kausale Aggregation.

Mit Hilfe der kausalen Aggregation kann man eine Kette von Relationen zu einer einzigen kausalen Relation zwischen der ersten Ursache und der letzten Wirkung zusammenfassen. Man unterscheidet hier drei Fälle:

- (1) Der Knoten hat keinen kausalen Vorgänger oder keiner der kausalen Pfade, die in den Knoten münden, besitzt zusammenfaßbare Knoten.
=> Es werden keine Links hinzugefügt.

(2) Jeder Vorgängerpfad hat einen Knoten mit fokaler Aggregation

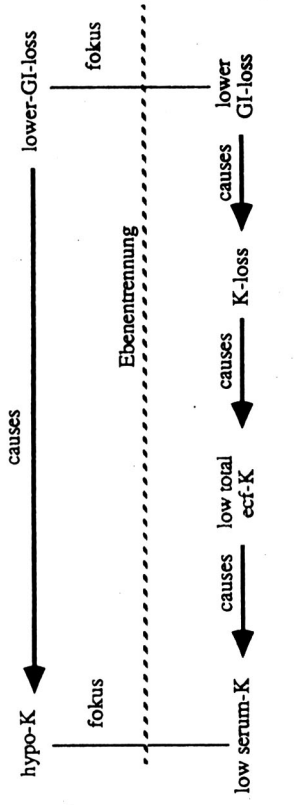


Abb. 1.8: kausale Aggregation

Zwischen dem betrachteten Knoten (hypo-K) und der ersten fokalen Aggregation in dem Vorgängerpfad (hier: lower-GI-loss) wird ein kausaler Link hinzugefügt.

(3) Nur ein Teil der Vorgängerpfade haben Knoten mit fokaler Aggregation. Der Knoten wird deshalb in zwei Bestandteile aufgeteilt: die 'accounted' und die 'unaccounted' Komponente.

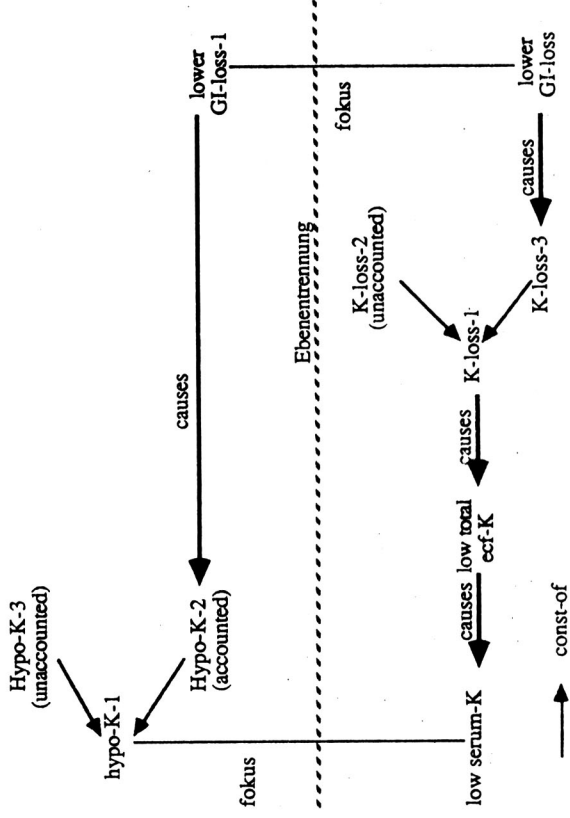


Abb. 1.9: kausale Aggregation

1.3.5.2 Elaboration.

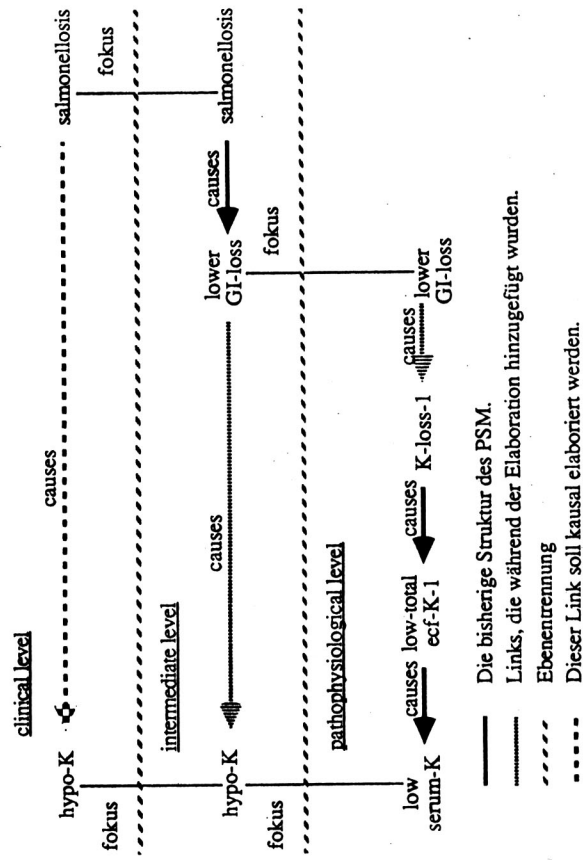
Die Elaboration ist die Umkehrung der Aggregation. Sie liefert eine detailliertere Beschreibung von Knoten und Links der höheren Stufe, die mit der Aggregation konsistent ist. Wenn der kausal instantierte Pfad andere kausale Pfade im PSM beeinflusst, so werden die kombinierten Wirkungen der 'multiplen Kausalität' mit Hilfe der Komponentensummutation berechnet.

A. Fokale Elaboration.

Um ein Netzwerk auszuarbeiten, identifiziert man die Knoten im Netzwerk, die zur Zusammenfassung von Beschreibungen benutzt werden. Man stellt ihre Verweise (Beziehungen) auf der darunterliegenden Ebene auf und fügt Knoten und Links hinzu, um das Phänomen, der oberen Ebene zu beschreiben.

B) Kausale Elaboration.

Die kausale Elaboration bestimmt die kausalen Relationen zwischen Knoten einer detaillierteren Ebene mit Hilfe der kausalen Relationen zwischen den Knoten auf der aggregierten Ebene.



- Die bisherige Struktur des PSM.
- Links, die während der Elaboration hinzugefügt wurden.
- Ebenenrennung
- Dieser Link soll kausal elaboriert werden.

Abb. 1.10: kausale Elaboration

1.3.5.3 Projektion

Die Projektion wird dazu benutzt Hypothesen (Befunde und Krankheiten), die zur Erklärung unerklärter Zustände im PSM beitragen, aufzustellen. Sie entspricht in etwa der Elaboration. Die kausalen Projektionen sind aber hypothetisch und somit noch nicht im PSM vorhanden. Die Projektionen findet man in den kohärenten Hypothesen wieder. Die Projektion spielt eine wichtige Rolle beim diagnostischen Problemlösen und zwar beim - Erforschen diagnostischer Möglichkeiten

- Bewerten ihrer Wahrscheinlichkeit
- Auswerten von Konsequenzen einer hypothetischen Diagnose.

Die Projektion schlägt fehl, wenn die kausale Beschreibung des hypothetischen Links auf irgendeiner Abstraktionsebene mit dem PSM nicht konsistent ist.

1.3.5.4 Komponentensummutation und Komponentendekomposition.

Es ist möglich, daß ein Patient an mehr als einer Krankheit leidet und durch die Anwesenheit der einen Krankheit die Symptome der anderen Krankheit verändert werden. Die Komponentensummutation verknüpft die Attribute der einzelnen Komponenten, um die Attribute der zusammengefaßten Knotens zu berechnen. Die Komponentendekomposition identifiziert die 'unaccounted' Knoten, indem sie den Unterschied zwischen dem zusammengesetzten Knoten und den bekannten Knoten feststellt. Diese Mechanismen sind für die Rückkopplungsschleifen sehr wichtig.

1.3.6 Unvollständiges und unsicheres Wissen.

Unvollständiges Wissen kann in Abel auf verschiedene Arten dargestellt werden.

- Die verschiedenen Abstraktionsstufen erlauben es, Informationen allgemeiner (also auf einer höheren Abstraktionsstufe) oder genauer (auf einer tieferen Abstraktionsstufe) darzustellen.
- Durch die Möglichkeit gleichzeitig mehrere PSM's zu erstellen, können unterschiedliche Interpretationen schon bekannter Tatsachen dargestellt werden. Jedes PSM ist hierbei eine konsistente Berechnung der beobachteten Daten.
- Zu jedem dieser PSM werden kohärente Hypothesen (=Diagnosehüllen) aufgestellt, die vollständige, aber hypothetische Bilder der Krankheit des Patienten darstellen.

Unsicherheiten können zusätzlich durch die strukturelle Bewertung der PSM und der Diagnosehüllen (=kohärente Hypothesen) behandelt werden.

1.3.7 Bewertung der PSM.

Die Bewertung eines PSM gibt an, inwieweit das PSM für die Erklärung der Krankheit des Patienten unvollständig ist. Sie wird durch die Summation der Schweregrade ganz oder teilweise unerklärter Zustände im PSM berechnet. (Jedem pathophysiologischen Zustand werden in Abel Attribute zugeordnet, wie Startzeit, Dauer und Schweregrad des Zustandes.)

Diese Bewertung in Abel ist sehr einfach, aber gerade noch ausreichend. Ein Nachteil ist, daß die Anzahl der Ursachen, die die unerklärten Zustände erklären, nicht in die Bewertung einbezogen wird.

1.3.8 Bewertung der Krankheitshypothesen.

Eine Krankheitsdiagnose wird zu einer Hypothese, wenn sie einige unerklärte Befunde erklären kann. Heuristik zur Durchführung einer Best-First-Suche nach einer geeigneten Hypothese:

- (1) Aufteilung aller Hypothesen in Gruppen, deren Anzahl der unerklärten Befunde, die erklärt werden können, gleich ist.

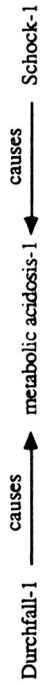
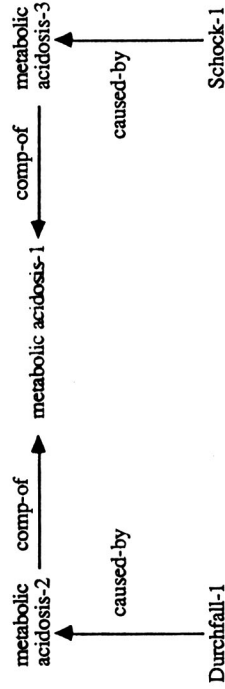


Abb. 1.11: Durchfall und Schock können Ursachen der metabolic acidosis sein.

In dieser Abbildung erkennt man die Überlagerung zweier Relationen. Sowohl der Durchfall als auch der Schock können Ursachen für die metabolic acidosis sein. Beide Ursachen können auch gemeinsam auftreten. Betrachtet man nur den ersten Teil der Relation, so könnte man davon ausgehen, daß Durchfall der einzige Grund für die metabolic acidosis ist. Der Schweregrad des Durchfalls und die Werte der metabolic acidosis passen dann allerdings nicht zusammen und können sogar zu einem Widerspruch führen.



comp-of = ist Bestandteil von

Abb. 1.12:

Um ein genaueres quantitatives Denken zu ermöglichen, werden die Komponenten der metabolic acidosis je nach Ursache getrennt instantiiert und explizit zusammengefaßt.

Die Dekomposition wirft neue Probleme bezüglich der Faktoren der einzelnen Bestandteile und ihrer Verrechnung auf (Hier nimmt man in Abel einfach an, daß es dazu einen Mechanismus gibt).

(2) Innerhalb dieser Gruppen werden die Hypothesen nach drei Faktoren berechnet:

- match*: Anzahl der Gründe und Befunde im PSM, die mit der Krankheitshypothese konsistent sind.
- mismatch*: Anzahl der Gründe und Befunde im PSM, die mit der Krankheitshypothese inkonsistent sind.
- unknown*: Anzahl der noch nicht beobachteten, aber vorhergesagten Befunde (durch die Hypothese), die mit dem PSM nicht inkonsistent sind.

Sowohl bei dem PSM als auch bei den Krankheitshypothesen liegt eine rein strukturelle und keine probabilistische Bewertung vor.

1.3.9 Kontrollfluß und Kontrollstrukturen.

Nach der Bewertung der PSM's und der dazugehörigen Diagnosehüllen ist dem Programm eine Menge 'wahrscheinlicher' Diagnosehüllen bekannt, die um die Erklärung der Krankheit des Patienten konkurrieren.

Die Differenzierung zwischen den wahrscheinlichsten Hypothesen des am besten bewerteten PSM wird als 'Problem' formuliert. Die Lösung dieses Problems wird zum Hauptziel des diagnostischen Informationssammelns. Ein diagnostisches Ziel besteht aus drei Komponenten:

- dem ursprünglichen Ziel (= Hauptproblem)
- dem Kontext (= Grund, warum das Problem gelöst werden soll)
- der Erwartung (= Erwartungen des Programms aufgrund des bisher verfügbaren Wissens. Mit ihrer Hilfe soll entschieden werden, ob die neuen Informationen mit den bisher bekannten Patientenbeschreibungen konsistent sind.)

Ziel 1 : **Differenzierte Durchfall, akutes Nierenversagen**

Kontext: Erklärung der metabolischen acidosis
Erwartungen:

- Möglich: Durchfall
- Schweregrad: mittel
- Möglich: akutes Nierenversagen
- Schweregrad: mittel
- Dauer: wenige Tage

Unterziele: (AND 2 3 8 9 10)

Abb. 1.13: Struktur des Hauptziels zur Differenzierung zwischen Durchfall und akutem Nierenversagen.

Als Problemlösungsstrategien stehen folgende Mechanismen zur Verfügung: Differentiate, Confirm, Explore, Rule-Out, Refine, Group-and-Differentiate

Das oberste diagnostische Ziel wird vom Problemlöser in Unterziele aufgeteilt, bis Unterziele erreicht werden, die einfach gelöst werden können. Die Unterziele werden mit Hilfe von Aggregation und Disaggregation im Zusammenhang mit der Patientenbeschreibung erzeugt. Der Zielbaum erlaubt lokale Reaktionsmöglichkeiten bei Widersprüchen und bietet eine gute Möglichkeit zur Verbesserung der Erklärungskomponente.

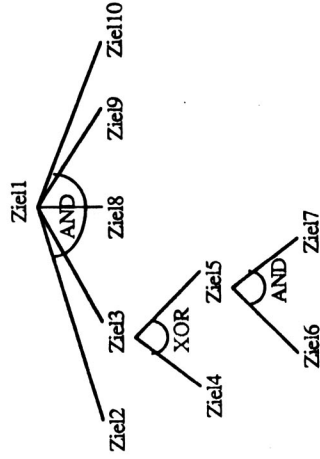


Abb. 1.14: Der vollständige Zielbaum. AND zeigt an, daß alle Ziele erreicht werden müssen. Bei XOR kann nur ein Ziel erreicht werden.

Alle Unterziele werden nacheinander bearbeitet. Dann versucht man das Hauptziel auszuwerten. Werden hierbei Widersprüche festgestellt, so kann man einen Rechtfertigungsmechanismus aktivieren (siehe nächste Seite).

Nun folgt ein Beispiel, um den gesamten Ablauf des Problemformulierens und Problemlösens besser darzustellen.

Ein Patient ist seit drei bis vier Tagen krank. Man weiß, daß er mittelschwere 'metabolic acidosis' (im Stoffwechsel entstandene krankhafte Vermehrung des Säuregehaltes im Blut) und leichte 'hyponatremia' (zu wenig Natrium) hat. Weiter ist nichts bekannt.

In Abb. 1.15 und Abb. 1.16 sieht man die Problemformulierungen für die Hypothesen Durchfall und Nierenversagen. Das Hauptziel ist nun die Entscheidung zwischen diesen möglichen Interpretationen. Dazu sucht das Programm in den Interpretationen nach unterschiedlich vorhergesagten Zuständen. Dann wird eine Menge von Unterzielen

erstellt, um jeden Unterschied zu verfolgen. Als Unterziel ist zum Beispiel die Natrium-Konzentration im Serum geeignet. Durchfall verweist auf eine niedrige Natrium-Konzentration, während das Nierenversagen relativ hohe Natrium Konzentrationen im Urin anzeigt. Weitere Unterschiede sind:

- state of hydration (Zustand des Wasserhaushaltes im Körper)
- serum K (Kalium-Konzentration im Serum)
- serum creatinine (Kreatinin-Konzentration im Serum)
- hämoglobin und tubular cells casts in urine

(Vorkommen von Hämoglobin- und Tubulusepithelzylinder im Urin)

Eine graphische Darstellung der Zielstruktur wird in Abb. 1.14 dargestellt.

Sind alle Unterziele fertig bearbeitet, so wird das Hauptziel ausgewertet. (Genauer zur Auswertung der Ziele und der Behandlung der Widersprüche findet man in Abschnitt 1.3.10). Während der Auswertung merkt das Programm, daß nicht alle Befunde untereinander konsistent sind.

Genauer:

Hohe Na-Konzentration im Urin weist auf Nierenversagen,

wohingegen niedrige Serum-K. Konzentration auf Durchfall weist.

==> Der Rechtfertigungsmechanismus des Programmes wird aktiviert, um den Widerspruch zu lösen.

Die entsprechenden Ziele lauten:

Ziel 11: Finde eine Rechtfertigung für die hohe Na-Konz im Urin

Kontext: Durchfall

Erwartungen:

Möglich: Benutzung eines harntreibenden Mittels

Möglich: zusätzliche Krankheit

Ziel 12: Finde eine Rechtfertigung für die niedrige K-Konz. im Serum

Kontext: akutes Nierenversagen

Erwartungen:

Möglich: Einnahme eines harntreibenden Mittels

Möglich: Erbrechen

Dadurch erfährt das Programm, daß der Patient zusätzlich an Erbrechen leidet. Jetzt ist der Widerspruch ausgeräumt. Danach werden die neuen Informationen in die PSM eingetragen und die Diagnosehüllen überarbeitet.

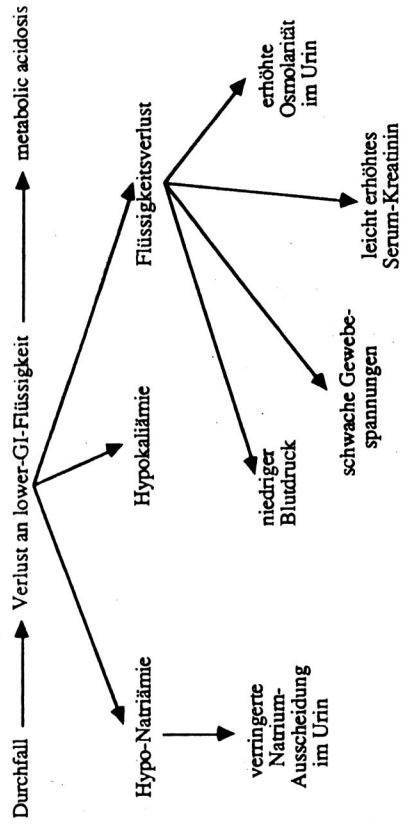


Abb. 1.15: Problemformulierung für die Hypothese Durchfall

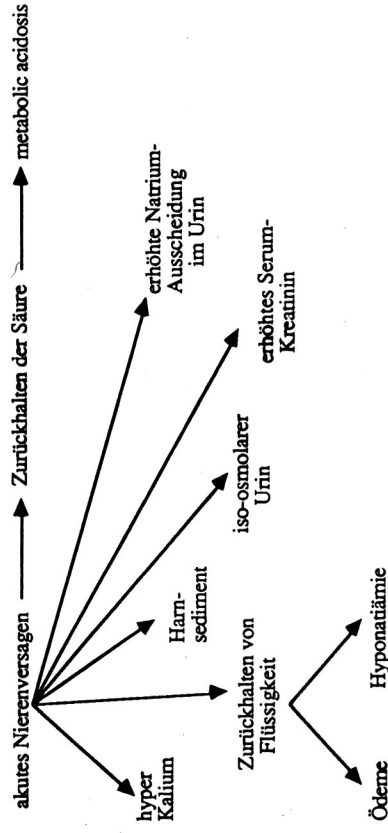


Abb. 1.16: Problemformulierung für die Hypothese akutes Nierenversagen.

1.3.10 Möglichkeiten zur Behandlung von Widersprüchen.

1. Vermeidung von Widersprüchen.

Sind die Unterziele abgearbeitet, so muß man folgendes beachten: Jedes Unterziel ist mit einer Menge von Bedingungen verknüpft, die erfüllt werden müssen, um dieses Ziel zu erreichen. Die Bedingungen der Ziele werden mit Hilfe der Diagnosehüllen berechnet. Wie gut ein Ziel erfüllt wurde, kann man mit vier Werten ausdrücken:

- (1) Erfolg
- (2) teilweiser Erfolg
- (3) Scheitern
- (4) Widerspruch

Ist das Ergebnis 'Erfolg', so wird die Kontrolle an das nächst höhere Ziel mit dem Hinweis auf Erfolg abgegeben. Ist das Ergebnis 'teilweiser Erfolg', so ist noch nicht genügend Information vorhanden, aber es gibt schon einen positiven Hinweis. Ist das Ergebnis 'Scheitern', so versucht der Problemlöser eine Rechtfertigung dafür zu finden, daß das Ziel ohne Widerspruch negiert auftritt.

Wird eine Rechtfertigung gefunden, so kann das Ziel (teilweise) erreicht sein und das Programm kann dieses Ziel weiterverfolgen. Wird keine Rechtfertigung gefunden, so wird das Ziel verfehlt. Somit ist eine Unterscheidung zwischen scheinbarem und wirklichem Widerspruch möglich.

2. Behandlung eines Widerspruchs.

Liegt ein ernsther Widerspruch vor, so liegt ein Fehler in der Patientenbeschreibung oder in der Problemformulierung vor. Der Widerspruchsbehandler besteht aus zwei Teilen:

- Zielorientierter Widerspruchsbehandler (Problemformulierung) Er legt den Bereich des Widerspruchs fest. Ist der Widerspruch lokal bezüglich eines Zieles, so werden dieses Ziel und alle abhängigen Ziele neu formuliert.
- PSM-Debugger (Patientenbeschreibung) Die Patientenbeschreibung muß in Hinblick auf den neuen Befund geändert werden. Ein neuer Problemlösungszyklus wird gestartet. Kann in beiden Teilen der Widerspruch nicht gelöst werden, so ist entweder die neue Tatsache oder eine frühere Tatsache falsch und muß vom Benutzer zurückgezogen werden (kommt allerdings selten vor).

Die Rücknahme einer Tatsache ist schwierig und teuer. Deshalb versucht das System Widersprüche so früh wie möglich zu erkennen bzw. auf die nächst tiefere Stufe zu transformieren, um den Widerspruch lokal zu halten oder erst gar nicht auftreten zu lassen.

1.4 Bewertung von Abel.

1.4.1 Vergleich der Anforderungen aus der Problemstellung mit der funktionalen und der Architektursicht.

Anforderung	funktionale Sicht:	Architektursicht
Eingaben des Benutzers sollen auf ihre Konsistenz überprüft werden.	Konsistenzüberprüfung der Eingabe	Nicht realisiert
<i>Die Eingaben des Benutzers werden nicht auf ihre Konsistenz überprüft. Besonders bei Labordaten sollte zumindest der Wertebereich kontrolliert werden.</i>		
Aus allen Eingaben soll das exakte Problem formuliert werden.	Erstellen des ersten PSM.	Initial formulation (-> 1.3.4)
Größere Probleme sollen in Teilprobleme aufgeteilt werden.	Kontrollmechanismus zur Unterteilung des Hauptzieles in mehrere Unterziele.	Das Hauptziel wird entsprechend dem PSM und der Diagnosehüllen in Unterprobleme aufgeteilt. (-> 1.3.9)
Auch bei unvollständigem und unsicherem Wissen soll das Problem gelöst werden.	Mechanismen zur Repräsentation von unvollständigen und unsicherem Wissen.	Verschiedene Abstraktionsebenen PSM, Diagnosehülle, strukturelle Bewertung der Krankheits-hypothesen und des PSM. Strukturierte Regeln
0		

1 + : Die Anforderungen werden sehr gut erfüllt.

0 : Die Anforderungen werden gut erfüllt.

- : Die Anforderungen werden nur teilweise oder gar nicht erfüllt.

Die korrekte Lösung soll (auch Mehrfachlösungen) gefunden und eine Therapie vorgeschlagen werden.	Simulationenmechanismus mit Rückkopplungsschleifen	PSM, Diagnosehilfe, Elaboration Aggregation, Projektion, Kompositionensumman und -dekomposition
+ -		
<i>Die Realisierung der Simulation mit Rückkopplungsschleifen ermöglicht es, Mehrfachdiagnosen auch dann zu erkennen, wenn sich die Symptomausprägungen überlagern. Die Zuordnung von Therapien wurde in Abel bisher nicht realisiert.</i>		
Folgesetzungen sollen entsprechend ausgewertet werden.	Zeitindizierung Mechanismen zur Auswertung von Zeit	Alle pathophysiologischen Zustände des PSM besitzen die Attribute: Startzeit und Dauer. Speicherung des letzten PSM der letzten Sitzung.
-		
<i>Es gibt eine Zeitindizierung und das letzte PSM wird gespeichert, unklar ist aber die Verwendung dieser Daten in der neuen Sitzung.</i>		
Ähnliche Diagnosen sollen unterschieden werden.	Differentialdiagnose und Komplikationen	Zu jeder Diagnose sind die Differentialdiagnosen und die Komplikationen bekannt. (->1.3.1.2)
0		
Über die Zeit sich ändernde Laborwerte sollen aggregiert werden.	Zeitaggregation	Eine Folge von Werten eines Zustandes wird zu einer Beschreibung zusammengefasst.
0		
Fehlerhafte Eingaben sollen schnell und einfach zurückgezogen werden.	Recovery-Mechanismus	Transformation des Fehlers auf eine möglichst tiefe Stufe (-> Fehler lokal halten) Änderung des PSM und Starten eines neuen Problemlösungsmechanismus.
-		

Hintergrundwissen und kausales Wissen sollen geeignet dargestellt werden.	Hierarchische Repräsentation der Anatomie. Modelle für jede mögliche Krankheit, kausale Beziehungen	Anatomie: part-of, enthalten-in und Materialflußrelationen, kausaler Link, Klassifikation der Krankheiten. (-> 1.3.1)
+		
<i>Abel verfügt über sehr umfangreiches Hintergrundwissen in Form von Hierarchien, sowie physiologisches und pathophysiologisches Wissen.</i>		
Symptome und Diagnosen sollen zu Gruppen zusammengefasst werden.	Klassifikationsmechanismus für Diagnosen	Klassifikation der Krankheiten bzgl. Anatomie, Zeit, Ätiologie und Physiologie.
0		
<i>Die Anfangssymptome sind größtenteils Laborwerte, die nur schlecht in Gruppen zusammengefasst werden können und nicht sehr zahlreich sind. Welche Symptome im Verlauf des Diagnoseprozesses erfragt werden, wird durch die Aufstellung des Hauptziels in Unterziele bestimmt. Symptome, die anhand einer Verfeinerungsrelation in einem Symptomwald überführt werden können, liegen in diesem Anwendungsgebiet nicht vor.</i>		
Gleichzeitige Verarbeitung von physiologischem und pathophysiologischem Wissen.	verschiedene Abstraktionsebenen mit guten Verbindungen	PSM mit mehreren Abstraktionsebenen und Bearbeitungsmechanismen. (-> 1.3.5)
0		
Eingabekomponente soll für einen schnellen Input von Laborwerten geeignet sein.	Nicht realisiert	Nicht realisiert
-		
Eine Notfalldiagnostik soll möglich sein.	Mechanismus zur Notfalldiagnostik	Nicht realisiert.
-		

1.4.2 Zusammenfassende Bewertung.

Ziel dieses Projektes ist es, Abel zu einem KI-Programm zu entwickeln, das sich möglichst nahe an die Denkweise des Menschen anlehnt.

Abel kann gleichzeitig phänomenologisches Wissen über Krankheitsassoziationen (z.B. Der Patient mit der Krankheitsgeschichte Y hat als allgemeine Komplikation die Krankheit X) und pathophysiologisches Wissen über Krankheitsmechanismen verarbeiten (z.B. Na- und K-Ionen Austausch in der distal tubule). Abel stellt eine mehrstufige kausale Beschreibung der Krankheit des Patienten zur Verfügung. Wirkungen verschiedener Ursachen können zusammengefaßt oder separiert werden (Komponentensumation und -dekomposition). Dadurch können auch Rückkopplungsschleifen simuliert werden. Das umfangreiche Hintergrundwissen ist in mehreren Hierarchien gespeichert.

Bis jetzt ist das implementierte Diagnostikprogramm nur eine Näherung dessen, was hier vorgestellt wurde. Die Patientenbeschreibung und die Operationen auf dem PSM sind vollständig implementiert. Die Erklärungskomponente gibt in gutem Englisch Erklärungen in drei verschiedenen Detaillierungsstufen aus. Der begrenzte Speicher des Computers macht es schwierig, alle Diagnosehüllen, wie im Entwurf vorgesehen, zu erstellen. Sie müssen in einer kompakteren Darstellungsform zusammengeschlossen werden. Als Problemlösungsstrategien kann man nur Explore, Confirm und Differentiate benutzen. Auch komplexere Kontrollstrategien als das Backtracking sind bisher nicht möglich.

In Abel gibt es keine probabilistischen, sondern nur sehr einfache strukturelle Bewertungsmechanismen. Der Gesundheitszustand des Patienten (z.B. Lebenszeichen, Stabilität) werden außer Acht gelassen.

Um Abel einsetzen zu können, müssen schon gleich zu Beginn möglichst viele Daten bekannt sein (Störungen im Säure/Base und Elektrolythaushalt). Diese Voraussetzung schränkt die Anwendung stark ein.

Mit dem gesetzten Ziel, die menschliche Denkweise besser als die bisherigen KI-Programme zu simulieren, ist Abel bisher noch an die Grenze der Maschine (Speicherkapazität) gestoßen.

<p>Dialog: Antworten entsprechend dem Wissen des Benutzers modifizieren</p>	<p>Partnermodellerte Erklärungskomponente</p>	<p>Erklärungen auf den verschiedenen Abstraktionsebenen des PSM</p>
<p>Zugrundeliegende Hypothese soll erklärt werden</p>	<p>Erklärungskomponente Thinking-loud</p>	<p>Formulierung einer Ebene des PSM in Worte. Alle möglichen Diagnosen in der Diagnosehülle Nicht realisiert</p>
<p>Mögliche Alternativen</p>	<p>Nicht realisiert</p>	<p>Nicht realisiert</p>
<p>Fragen sollen jederzeit möglich sein.</p>	<p>Nicht realisiert</p>	<p>Nicht realisiert</p>
<p>Risiken für Entscheidungsvorschläge sollen angegeben werden.</p>	<p>Nicht realisiert</p>	<p>Nicht realisiert</p>
<p>+ -</p>		

Sehr gute Möglichkeiten zur Realisierung einer partnerorientierten Erklärungskomponente, die aber noch nicht ausgeschöpft wurden.

1.5 Anhang zu Abel

1.5.1 Übersetzung medizinischer Ausdrücke:

abdominal cavity	Bauchhöhle
acidemia	Blutübersäuerung (nicht lange mit dem Leben vereinbar, Blut $pH < 7,3$)
creatin	Kreatin
creatinin	Kreatinin (= Ausscheidungsform des Kreatins = Kreatinhydratid)
dehydration	Wasserentzug, Flüssigkeitsabgabe = state of reduced extracellular volume
diarrhea	Durchfall
ecf = extracellular fluid	extrazelluläre Flüssigkeit
hyperventilation (chronic)	Atmungssyndrom (nervöses)
hypobicarbonatemia	zu wenig Bicarbonat (HCO_3) im Blut
hypocapnia (=hypocapny)	Hypokapnie (verminderter Blut- CO_2 -Gehalt bei Acidose [$pH < 7,3$]) (= erniedrigte pCO_2 -Spannung)
hypokaliemia	Hypokaliämie (zu wenig Kalium)
lower GI fluid	Diese Flüssigkeit besteht aus Wasser, Kalium, Natrium, Bicarbonat und Chloriden.
metabolic acidosis Vermehrung	beim Stoffwechsel entstandene krankhafte des Säuregehaltes im Blut
osmolality	Osmolarität; Konzentration osmotisch wirksamer Moleküle einer Lösung in Gewichtseinheiten.
poor tissue turgor	schwacher Turgor, schwache Gewebespannung
serum	Serum = der flüssige, hauptsächlich Eiweißkörper enthaltende, nicht mehr gerinnbare Anteil des Blutplasmas.
sodium	Natrium = Natrium
urinary sediment	Harnsediment (Bodensatz)
urinary system	Nierensystem (oder Harnsystem)
volume depletion	Flüssigkeitsentzug

1.5.2 Erklärung von Begriffen in Abel:

coherent hypothesis Hypothese	kohärente Hypothese = zusammengehörige
deren	(die kohärenten Hypothesen eines PSM bilden Diagnosehülle)
DC = diagnostic closure	Diagnosehülle
etiology	Ätiologie = Lehre von den Krankheitsursachen
non aggregable node	Knoten, der auf einer höheren Ebene nicht weiter zusammenfaßbar ist. (Aggregation)
PSM	patientenspezifisches Modell
unaccounted	unerklärt; Für eine vollständige Diagnose müßte dieser Knoten noch erklärt werden, d.h. man muß noch eine Ursache für ihn finden.
Hydratisierung	Bindung von Wasser
pathologisch	krankhaft (verändert)
Pathophysiologie	Teilgebiet der Medizin, auf dem man sich mit den Krankheitsvorgängen und Funktionsstörungen im menschlichen Organismus befaßt.

2. Expertensystem zur Diagnose von Ohrenkrankheiten.

Das Ohrendiagnosesystem ist ein assoziatives Expertensystem mit der Besonderheit, daß Erkrankungen als prädikatenlogische Formeln und Symptome als Variablen dieser Formeln dargestellt werden.

Man unterscheidet beim Ablauf zwischen der Initialphase -einer hypothesenfreien Anfangsphase mit Symptom-Symptom-Relationen- und der Final- phase, bestehend aus der Anamnese- und der Hypothese-Phase. Im inexakten Bereich werden die Mechanismen Fokussierung und Abstraktion sowohl auf Krankheiten als auch auf Untersuchungen angewandt.

2.1 Anforderungen aus der Problemstellung.

Problemstellung:

Diagnosefindung auf dem Gebiet der Ohrendiagnose.

Anforderungen:

- 1) Die Eingabe soll auf ihre Plausibilität überprüft werden.
- 2) Freie Entscheidungen des Benutzers sollen jederzeit möglich und in die Diagnosefindung integrierbar sein.
- 3) Ähnliche Krankheiten sollen einfach zu unterscheiden sein.
- 4) Symptome sollen in Gruppen eingeteilt werden.
Die Anzahl der Diagnosen ist nicht so groß, daß eine Gruppierung unbedingt notwendig ist.
- 5) Dialog: Expertensystem <-> Benutzer
 - Der Benutzer soll auf Anfrage jeden Schritt des Systems verstehen können
 - Die Diagnose soll verständlich erklärt werden.
 - Mögliche Alternativen sollen ausgegeben werden.
 - Risiken für Entscheidungsvorschläge sollen angegeben werden.
- 6) Grobe Zeitangaben sollen eingegeben und verwertet werden können.
- 7) Das Problemlösen soll auch mit unsicherem und unvollständigem Wissen durchgeführt werden können.
- 8) Das System soll effizient arbeiten.
- 9) Folgesetzungen sollen entsprechend ausgewertet werden.
- 10) Fehlerhafte Eingaben sollen schnell und einfach zurückgezogen werden.

- 11) In der Wissensbasis sollen Symptome, Diagnosen, regelmäßige Ableitungsmechanismen und Verbindungen zwischen Symptomen und Diagnosen dargestellt werden.
- 12) Eine Notfalldiagnostik soll möglich sein.

2.2 Funktionale Sicht des Ohrendiagnosesystems.

- 1) Plausibilitätskontrolle der Eingabe
- 2) Differentialdiagnostik
- 3) Zeitverarbeitungsmechanismen
- 4) Mechanismen zur Darstellung und Verarbeitung von unvollständigem und unsicherem Wissen.
- 5) Recovery-Mechanismus
- 6) Symptomwald
- 7) Abstraktionsmechanismus für Diagnosen
- 8) Abstraktionsmechanismus für Untersuchungen
- 9) Berechnung des Beweissatus für jede Krankheit
- 10) Relationen zwischen Symptomen und Krankheiten
- 11) Erklärungen im inexakten Bereich
- 12) Mechanismen zur Notfalldiagnostik

2.3 Architektursicht des Ohrendiagnosesystems.

2.3.1. Definition der Grundbegriffe.

Mit Hilfe dieses Expertensystems können Ohrenkrankheiten diagnostiziert und Vorschläge für den weiteren Untersuchungsablauf aufgestellt werden.

Die Grundobjekte dieses Systems sind Symptome bzw. Symptomausprägungen, Untersuchungen und Erkrankungen. Erkrankungen werden als prädikatenlogische Formeln und Symptome als Variablen dieser Formeln dargestellt.

Symptome werden mit s_1, \dots, s_m bezeichnet, wobei jedes s_i sowohl als Sorte, als auch als Variable aufgefaßt wird. Weiter sei jedem s_i ein Wertebereich $r_i \subseteq R$ zugeordnet. Die Wertebereiche r_i sind alle disjunkt, da sie durch Sorten indiziert sind. Die Elemente von r_i heißen *Symptomausprägungen*.

$R := r_1 \times r_2 \times \dots \times r_m$ heißt *Zustandsraum*. R ist die Menge aller möglichen Zustände, in denen sich ein Patient befinden kann. Der Zustand der Information über einen Patienten wird beschrieben durch den *Informationsvektor*. $I := (x_1 \dots x_m)$ mit $x_i \in r_i$, $r_i \subseteq R$, d.h. alle Symptomvariablen sind mit Werten belegt, sonst heißt *unvollständig*.

Zu jeder *Krankheit* K sei die Teilmenge $K^+ \subseteq R$ der Bereich, in dem die Krankheit bestätigt wird und die Teilmenge $K^- \subseteq R$ der Bereich, in dem die Krankheit abgelehnt wird. K^+ und K^- müssen weder disjunkt sein, noch müssen sie R überdecken.

Die Mengen K^+ und K^- kann man durch aussagenlogische Formeln über Symptomvariablen und natürliche Konstanten ausdrücken.

Eine *Untersuchung* u ist eine Menge von Symptomvariablen $u \subseteq S$, die nach ihrer Durchführung belegt werden können. Dabei wird jedes Symptom von genau einer Untersuchung erfaßt. Untersuchungen sind Mittel zur Wissensacquisition, da sie Informationsvektoren in 'informativere' Informationsvektoren überführen.

Der *Untersuchungsgraph* G_r ist ein gerichteter Graph.

Die Knoten von G_r sind Informationsvektoren. Die Kanten von G_r sind Paare von Informationsvektoren (I, I') mit der Bedingung, daß eine Untersuchung u den Index (I, I') trägt. Befindet man sich im Informationszustand I und erfaßt die Symptome von u , so durchläuft man in G_r einen Pfad, der von I nach I' führt. Eine Folge von Untersuchungen ist nur dann in G_r enthalten, wenn sie medizinisch zulässig ist, bzw. wenn ihre Zulässigkeit noch durch Voruntersuchungen geklärt werden muß. (G beinhaltet alle Folgen von Untersuchungen) G_r enthält also für jede Belegungssituation

die Menge der möglichen Untersuchungsfolgen, die zur Diagnose einer Krankheit führen.

Im Zusammenhang mit den Informationsgraphen erweist sich eine weitere Differenzierung als sinnvoll. Deshalb werden die Krankheitsformeln (F) durch Modaloperationen erweitert.

$< > F(I)$: 'faktiv möglich':

Es gibt einen Weg in G von I nach I' mit $F(I')$.

$<(r)> F(I)$: 'real möglich':

Es gibt einen Weg in G_r von I nach I' mit $F(I')$.

$<r > F(I)$: 'nachweisfähig':

Es gibt einen Weg in G_r von I nach I' mit $F(I')$, dessen Pfeile vertauschbar sind. Es ist jedoch garantiert, daß alle Symptome der Formel erfaßt werden können.

$[\] F(I)$: 'notwendig':

In allen Knoten aller von I ausgehenden Wege in G ist F erfüllt; unabhängig von allen noch durchzuführenden Belegungen.

$[(r)] F(I)$: 'real notwendig' und $[r] F(I)$: 'real notwendig' werden entsprechend definiert.

Definition des Beweisstatus:

Mit Hilfe der Modaloperatoren kann in jeder Belegungssituation I den Krankheiten ein 'Beweisstatus' zugeordnet werden, der angibt, wie weit der Beweis beziehungsweise die Widerlegung einer Krankheit gediehen ist.

Die folgende Tabelle definiert die Beweisstatus mittels der Modaloperatoren:

Status	Bedingung	
poss	$\diamond h^+ (I)$ und $\diamond h^- (I)$ und keine Atomformel von h ist erfüllt "faktiv möglich, kein positiver Hinweis"	('possible')
susp	$\langle (r) \rangle h^+ (I)$ oder $\diamond h^+ (I)$ und $\diamond h^- (I)$ und es gibt eine erfüllte Atomformel von h "real möglich mit positivem Hinweis, verdächtig"	('suspicious')
ph	$[] h^+ (I)$ und $\diamond h^- (I)$ "bewiesen aber noch ausschließbar"	
-ph	$\sim [] h^+ (I)$ und $\diamond h^- (I)$ "nicht mehr beweisbar aber noch ausschließbar"	
eh	$\diamond h^+ (I)$ und $[] h^- (I)$ "noch beweisbar, widerlegt"	
-eh	$\diamond h^+ (I)$ und $\sim [] h^- (I)$ "noch beweisbar, nicht mehr widerlegbar"	
h	$[] h^+ (I)$ und $\sim [] h^- (I)$ "bewiesen und nicht widerlegbar"	
-h	$\sim [] h^+ (I)$ und $[] h^- (I)$ "nicht beweisbar und widerlegt"	
con	$[] h^+ (I)$ und $[] h^- (I)$ "bewiesen und widerlegt"	('controversal')
ind	$\sim [] h^+ (I)$ und $\sim [] h^- (I)$ "nicht beweisbar und nicht widerlegbar"	('indifferent')
ci	$\sim \diamond h^+ (I)$ und es gibt erfüllte Atomformeln von h^+ "Beweis unzulässig"	('contraindication')

Abb. 1.17: Auflistung aller Beweisstatus.

- I : aktueller Informationsvektor
- h : eine Erkrankung
- h^+ : die beweisende Formel von h (Repräsentation von K^+)
- h^- : die widerlegende Formel von h (Repräsentation von K^-)
- \sim : die Negation der Modaloperatoren.

Je nach ihrem Beweisstatus werden die Hypothesen (=Krankheiten) weiterbehandelt (siehe Abschnitt 2.3.3.1.1 Fokussierung der Hypothesen).

2.3.2 Ablauf des Ohrendiagnosesystems.

Die Inferenzstrategie wird aufgeteilt in die Initialphase und die Finalphase.

Die *Initialphase* ist eine hypothesenfreie Anfangsphase, in der zunächst die Grundsymptome erfaßt werden. Zusätzlich werden Symptome betrachtet, die mit pathologisch belegten Symptomen in Relation stehen (z.B. Verfeinerung oder Nachbarschaft) und durch anamnestiche Untersuchungen erhoben werden können. (forward chaining). Die Initialphase wird beendet, wenn der Benutzer dies explizit wünscht oder es keine Symptome gibt, die anamnestisch erhoben werden können und gleichzeitig durch das Zutreffen einer Relation interessant erscheinen.

Die *Finalphase* ist eine hypothesen gesteuerte Endphase, bestehend aus der Anamnese-phase und der Hypothesephase.

Anamnese-phase: Symptome, die aufgrund von Relationen oder ihres Vorkommens in Krankheitsformeln von Hypothesen interessant erscheinen und gleichzeitig als anamnestiche Symptome oder Symptome schon durchgeführter apparativer Untersuchungen als verfügbar gelten, werden zur Erfassung vorgeschlagen. (Nach einer Untersuchung werden nicht alle dadurch zur Verfügung stehenden Symptome eingegeben, sondern nur die, die wirklich angefordert wurden. Alle übrigen Symptome können später leicht nachgereicht werden.)

Hypothesenphase: Zuerst werden alle Krankheitsformeln ausgewertet und der Beweisstatus der betrachteten Krankheiten ermittelt. Mit Hilfe der Mechanismen 'Fokussierung' und 'Abstraktion' werden die Krankheiten ausgewählt, deren zugehörige Untersuchungen (im Untersuchungspfad) ebenfalls abstrahiert werden. Aus dem Ergebnis dieser Abstraktion wird dann eine Untersuchung ausgewählt.

2.3.3 Fokussierung und Abstraktion.

Zur Realisierung der Fokussierung und der Abstraktion werden folgende Listenklassen in den einzelnen Abstraktionsebenen verwendet:

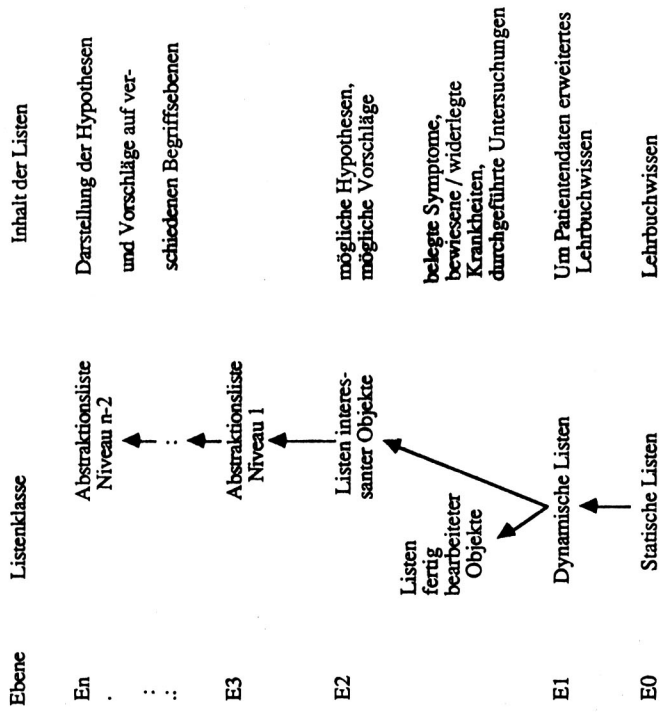


Abb. 1.18: Listenklassen der einzelnen Abstraktionsebenen.

Die Implementierung der Abstraktion ist mit Produktionsregeln möglich. Die Abstraktionsregeln können bei Bedarf leicht geändert oder ersetzt werden.

2.3.3.1 Hypothesen (Krankheiten)

2.3.3.1.1 Fokussierung

In der Fokussierungsphase werden alle Hypothesen für die interessante Liste zusammengetragen. Dazu betrachtet man den Beweisstatus aller Krankheiten. Ist der Beweisstatus 'susp', so wird die Krankheit in die interessante Liste übernommen. Liegt der Beweisstatus 'ph' oder 'h' vor (die widerlegende Formel ist erfüllt), so werden alle als Differentialdiagnose gekennzeichneten Krankheiten übernommen.

Ist der Status einer Krankheit 'ph' oder 'h', (die beweisende Formel ist erfüllt) so werden die als Komplikationen verzeichneten Krankheiten übernommen. Falls es für eine Krankheit ein pathophysiologisch belegtes Schlüsselersymptom in der aktuellen Belegungsliste gibt, wird sie übernommen. Schon bewiesene oder widerlegte Krankheiten werden nicht wieder aufgenommen.

2.3.3.1.2 Abstraktion.

a) Die erste Stufe kombiniert die verschiedenen Verdachtsmomente der Krankheit, die in der interessantesten Liste eingetragen sind, zu einer Hinweisstärke.

Index	≥50	≥50	≥50	≥75	≥60
#Hinweise	≥0			≥2	≥1
#Diff.Diag		≥1			
#Kompl.			≥1		
#Schlüssel					≥1
Hinweisstärke	deutlich	deutlich	deutlich	stark	stark
Grund	Hinweis	Formel	Diff. Diag.	Hinweise	Schlüssel Formel

Abb. 1.19: Abstraktionstabelle

b) Die zweite Stufe abstrahiert von der Hinweisstärke und den Attributen einer Krankheit die Hypothesenkategorie (HC).

Hinweiswahrscheinlichkeit	dr. utlich	stark	sehr häufig	jetzt, Tage	selten
Dringlichkeit	sehr häufig			Leben Funktion	keine Leben
Wirkung	Leben, Beschwerden Funktion				
Therapiekosten					
HK	tertiär	sekund.	sekund.	primär	widerlegen
Grund	-	Hinweis Wirkung häufig	starker Hinweis	lebensbedrohend dringend	lebensbedrohend nicht therapierbar

Abb. 1.20: Abstraktionstabelle

Beispiel: Wenn die Wahrscheinlichkeit, daß die Krankheit vorkommt "häufig" ist, dann ist die Hypothesenkategorie "sekundär" und der Grund lautet "häufig".

2.3.3.2 Untersuchungen.

2.3.3.2.1 Fokussierung.

Für alle interessanten Krankheitspfade werden die Untersuchungspfade betrachtet (=Menge von Untersuchungsfolgen, die zum Beweis einer Hypothese durchgeführt werden müssen). Alle Untersuchungen dieser Pfade werden in die interessante Liste aufgenommen.

2.3.3.2.2 Abstraktion.

a) Die erste Stufe umfaßt die Abstraktion der in der interessantesten Liste zusammengetragenen Daten über die Untersuchungen zur Wichtigkeit. Parallel dazu werden die Belastungsattribute der Untersuchungen zu dem neuen Attribut Belastung abstrahiert.

Risiko						
Schmerzen		Strahlen	erträglich gering	Stunde	stark	Verletzung
Zeit					Tag	
Kosten						
Belastung	keine	erträglich	erträglich	erträglich	hoch	hoch
Grund	-	Strahlen	Schmerzen	dauert lange	starke Schmerzen	dauert 1 Tag Verletzungsgefahr

Abb. 1.21: Abstraktionstabelle

Ist die Anzahl tertiärer Hypothesen (zu deren Beweis die Untersuchung gebraucht wird) größer als zwei oder die Anzahl sekundärer Hypothesen größer gleich eins, so wird 'wichtig' abstrahiert. Ist die Anzahl primärer Hypothesen größer gleich eins, so wird 'sehr wichtig' abstrahiert. Als Default-Wert wird 'unwichtig' eingesetzt.

b) Abstraktion der Vorschlagskategorie aus Wichtigkeit, Belastung und Access.

Wichtigkeit						
Belastung		wichtig	wichtig	wichtig	sehr wichtig	
Access		nicht hier	hier	hier		
Vorschlagskategorie	tertiär	sekundär	primär	dringend		
Grund	-	wichtig, nicht hier	wichtig hier	vorrangig		

Abb. 1.22: Abstraktionstabelle

Die Folge von Abstraktionsregeln wird als Sequenz imperativer if-then-else Regeln dargestellt.

=> Das Hinzufügen und Entfernen von Regeln ist leicht möglich und die Regeln können geschachtelt werden.

2.3.4. Relationen zwischen Symptomen und Krankheiten.

Das Ohrendiagnosesystem bietet eine Menge unärer Relationen auf die einzelnen Grundobjekte, sowie Relationen zwischen den Grundobjekten. Zu letzteren will ich einige Beispiele anführen.

Relationen zwischen Krankheiten sind die Differentialdiagnose und die Komplikation. Zwischen den Symptomen gibt es die Verfeinerungsrelation und weitere Nachbarschaftsrelationen. Mit Hilfe der Verfeinerungsrelation kann man einen gerichteten Symptomgraphen aufbauen, dessen Knoten Symptome und dessen Kanten Elemente der Verfeinerungsrelation sind. Der Symptomgraph ist ein Wald. Seine Wurzelsymptome heißen Grundsymptome. Auf dem Symptomgraph werden alle Werte fortgepflanzt. Ist der Vater physiologisch belegt, so wird auch der Sohn physiologisch belegt. Sind alle Söhne physiologisch belegt, so wird auch der Vater physiologisch belegt. Ist der Sohn pathologisch belegt, so ist auch der Vater pathologisch belegt. Durch den Prozeß der Wertfortpflanzung können neue Eingaben besser auf Konsistenz überprüft werden.

Relationen zwischen Symptomen und Erkrankungen sind *key*, *hinweis* und *covers*.

key: Ein Symptom *s* ist ein Schlüsselssymptom für eine Krankheit *K*.

hinweis: Eine näher spezifizierte Ausprägung von *s* ergibt einen starken Verdacht auf die Krankheit *K* (mit Apparatenutzung, aber starker Hinweis).

covers: Ist die Krankheit *K* bewiesen, so ist eine noch näher zu spezifizierende Ausprägung von *s* erklärt und soll nicht mehr zu der Erzeugung von neuen Hypothesen herangezogen werden.

2.3.5 Die Erklärungskomponente.

Eine Grundforderung des Ohrendiagnosesystems ist die Transparenz für den Benutzer. Besonders im inaktiven Bereich, in dem diffuses Wissen verarbeitet wird, sind die Schlußfolgerungsketten für den Benutzer nicht direkt durchschaubar. Deshalb werden den Objekten aller Ebenen (außer der untersten Ebene) Begründungen zugeordnet, warum das Objekt in der jeweiligen Ebene zu finden ist. Die Begründungen beziehen sich auf die Einträge der darunterliegenden Ebene und sind als Listen implementiert. Sie beinhalten also die Ursachen für alle Veränderungen zwischen darunter liegenden Ebenen und der betrachteten Ebene. Alle anderen Veränderungen findet man in den Begründungslisten der entsprechenden Ebenen. Wird eine Abstraktionsregel angewandt, so trägt sie eine Begründung in dieser Liste ein.

Dem Benutzer zeigt das System die Begründungsliste der Endergebnisse, die sich auf Einträge der darunterliegenden Ebene bezieht. Interessiert sich der Benutzer für weitere Informationen, so kann er sich die entsprechenden Listen der darunterliegenden Ebenen zeigen lassen.

2.3.6. Die Repräsentation der Zeit.

Die meisten Krankheitsbilder sind durch die zeitliche Entwicklung der Symptomausprägungen charakterisiert. Im Modell wird der kontinuierliche Zeitverlauf auf eine diskrete Abstraktion von Symptomwerten in einem endlichen Zeitintervall reduziert.

Im Ohrendiagnosesystem wird der Informationsvektor, der den Patientenzustand beschreibt mit dem Zeitpunkt indiziert, zu dem die Symptombelastung erfaßt wurde. Wird eine Untersuchung mehrfach wiederholt, so existieren mehrere Informationsvektoren.

Die zeitliche Entwicklung der Symptomausprägungen kann durch zeitmodale Operatoren beschrieben werden. Existiert ein Symptom, Ausprägung, Zeitintervall) gibt an, ob es während des Zeitintervalls ein Symptom mit der Ausprägung gegeben hat. Persistiert ein Symptom, Ausprägung, Zeitintervall) gibt an, ob für das Symptom während des gesamten Zeitintervalls keine andere Ausprägung vorlag. Daraus kann man folgern, daß das Symptom während des gesamten Zeitintervalls diese Ausprägung innehatte.

2.3.7. Unsicheres und unvollständiges Wissen.

Unvollständiges Wissen kann mit Hilfe des Informationsvektors und der Abstraktionsregeln dargestellt werden.

Der Informationsvektor enthält eine Menge von Symptomvariablen. Da nicht alle Symptome für eine Krankheit relevant sind, kann auch ein unvollständiger Informationsvektor eine Krankheit beweisen.

Der Definitionsbereich der Abstraktionsregeln ist die Potenzmenge der Wertebereiche der Ausgangskomponenten. Deshalb können Regeln formuliert werden, die nur auf Teilmengen der Wertebereiche definiert sind. Dank solcher Regeln kann auch bei unvollständiger Information ein Ergebnis abgeleitet werden.

Unsicheres Wissen kann mit Hilfe der Abstraktion und der unscharfen Grenzen in der Logik der Krankheitsformeln dargestellt werden. Durch die Abstraktion geht Information

verloren und unscharfe Aussagen mit hohem Informationsgehalt werden schärfere Aussagen mit niedrigerem Informationsgehalt.

Die Krankheitsformeln bieten vielfältige Ausdrucksmöglichkeiten, fordern aber die Angabe scharfer Grenzen. Gerade im medizinischen Bereich entstehen dadurch Probleme. Deshalb erweitert man die Bewertung und kann jetzt zwischen erfüllt, knapp erfüllt und nicht erfüllt unterscheiden.

2.3.8. Löschen von Belegungen und Recovery.

Um sinnvolles Arbeiten mit dem System zu gewährleisten, muß es möglich sein, einmal durchgeführte Symtombelegungen wieder zu löschen. (Eingabefehler können verbessert und Mißverständnisse bei der Erfragung von Symptomausprägungen von Patienten ausgeräumt werden.)

Das Ohrendiagnosesystem unterscheidet beim Löschen fehlerhafter Belegungen zwischen zwei Fällen:

(1) Die zu löschende Belegung wurde in der aktuellen Anamnesephase vorgenommen.
 ==> Löschen der falschen Belegungen und der daraus abgeleiteten Belegungen (aus der Belegungsliste)

(2) Die zu löschende Belegung wurde in einer früheren Anamnesephase vorgenommen.
 ==> Der gesamte Bereich der Hypothesen und Vorschläge kann inkonsistent werden.
 Deshalb wird ein Recovery notwendig. Alle abgeleiteten Informationen über Hypothesen und Vorschläge werden gelöscht. Die eingegebenen Informationen werden nun so behandelt, als ob sie in einer einzigen Anamnesephase eingegeben worden wären. Damit muß der Prozeß der Wertfortpflanzung und die Hypothese phase für alle bekannten Symptombelegungen nur einmal gestartet werden.

Abgef. Die Menge der neuen Hypothesen ist größer oder gleich der Menge der richtigen Hypothesen. Das Recovery bietet also eine größere Verfahrensweise, bei der aber keine Hypothese übersehen wird.

2.4 Bewertung des Ohrendiagnosesystems.

2.4.1 Vergleich der Anforderungen aus der Problemstellung mit der funktionalen und der Architektursicht.

Anforderungen	funktionale Sicht	Architektursicht
Die Eingaben sollen auf ihre Plausibilität überprüft werden. 0 ¹	Plausibilitätskontrolle der Eingabe	- Wertfortpflanzung auf der Verfeinerungshierarchie der Symptome - Jedem Symptom wird ein Wertebereich zugeordnet.
Freie Entscheidungen des Benutzers sollen jederzeit möglich und in die Diagnosefindung integrierbar sein. 0		- Eingabe von Symptomen ist jederzeit möglich. - Das Symptom liefert Vorschläge für eine Untersuchung, der Benutzer kann aber auch andere Untersuchungen durchführen.
Ähnliche Krankheiten sollen einfach zu unterscheiden sein. 0	Differentialdiagnostik	Jeder Diagnose werden die Differentialdiagnosen und die Komplikationen zugeordnet und im Abstraktionsmechanismus berücksichtigt
Symptome sollen in Gruppen eingeteilt werden. +	Symptomwald und Verbindungen zwischen Symptomen	Verfeinerungshierarchie der Symptome und weitere Nachbarschaftsrelationen.

Zur Einteilung der Symptome gibt es nicht nur die Verfeinerungshierarchie, sondern noch weitere Nachbarschaftsrelationen.

- 1+ : Anforderungen werden sehr gut erfüllt.
- 0 : Anforderungen werden gut erfüllt.
- : Anforderungen werden nur teilweise erfüllt.

Dialog: -Der Benutzer soll auf Anfrage jeden Schritt des Systems verstehen können. +-	Erklärungskomponente für den inexpliziten Bereich.	Begründung für die Einträge in den verschiedenen Abstraktionsebenen.
<i>Die Erklärungen im inexpliziten Bereich sind sehr gut, können aber nicht jeden Schritt des Systems erklären. Eine parnermodelliertere Erklärung ist nicht möglich.</i>		
- Die Diagnose soll verständlich erklärt werden. - mögliche Alternativen sollen ausgegeben werden. - Risiken für Entscheidungsvorschläge sollen angegeben werden.	Erklärungskomponente für den inexpliziten Bereich. Diese Anforderungen sind noch nicht realisiert, die Voraussetzungen für eine Realisierung sind aber gegeben.	Begründung für die Einträge in den verschiedenen Abstraktionsebenen.
Grobe Zeitangaben sollen eingegeben und verwertet werden. 0-	Zeitverarbeitungsmechanismus und geeignete Eingabekomponente.	Der Informationsvektor wird mit dem Zeitpunkt der Untersuchung indiziert. Zeitmodale Operationen wie exists und persists, Eingabekomponente ist noch nicht realisiert.
<i>Die Verwertung der Zeitangaben ist gut, die Eingabekomponente aber noch nicht realisiert.</i>		
Folgesetzungen sollen entsprechend ausgewertet werden.	Zeitverarbeitungsmechanismen	Der Informationsvektor wird mit dem Zeitpunkt der Untersuchung indiziert.

Das Problemlösen soll auch mit unvollständigem und unsicherem Wissen möglich sein. +	Abstraktionsmechanismen für Krankheiten und Untersuchungen. Mechanismen zur Darstellung und Verarbeitung von unsicherem und unvollständigem Wissen.	- Abstraktionsregeln - Informationsvektor - verschiedene Abstraktionsebenen - unscharfe Grenzen in der Logik der Krankheitsformeln
<i>Nestfähige und gute Möglichkeiten zur Darstellung und Verarbeitung von unsicherem und unvollständigem Wissen.</i>		
Das System soll effizient arbeiten. +-	Abstraktionsmechanismen für Untersuchungen und Krankheiten, Relationen zwischen Symptomen und Krankheiten.	Abstraktionsmechanismen (-> 2.3.3) Relationen (-> 2.3.3)
<i>Die Abstraktionsmechanismen und die verschiedenartigen Relationen führen zu einer hohen Effizienz des Systems. Die zu Beginn der Diagnosefindung notwendige Auswertung des Beweisstatus aller Krankheiten verschlechtert die Effizienz.</i>		
Fehlerhafte Eingaben sollen schnell und einfach zurückgezogen werden. 0	Recovery-Mechanismus	Zurückziehen fehlerhafter Eingaben (-> 2.3.8)
<i>Soll eine Belegung einer früheren Anamnesephase gelöscht werden, so ist nach dem Recovery-Mechanismus die Menge der neuen Hypothesen größer oder gleich der Menge der richtigen Hypothesen. Da die Okrenodiagnose ein kleines Anwendungsgebiet ist, verringert die eventuelle Vergrößerung der Hypothesenmenge die Effizienz nur minimal.</i>		

2.4.2 Zusammenfassende Bewertung.

Dieses Expertensystem dient zur Diagnose von Ohrenerkrankungen. Auf mögliche Therapien geht das System nicht ein. Die Erkrankungen werden als präkategorisierte Formeln beschrieben, deren Variablen die Symptome darstellen. Eine hierarchische Ordnung gibt es nur für die Symptome und nicht für Krankheiten (Diagnosen). Dies hat zur Folge, daß zu Beginn des Diagnoseprozesses die Formeln aller Krankheiten sequentiell ausgewertet werden müssen, ehe man die ersten Krankheiten ausschließen kann. Bei Anwendungsgebieten mit einer größeren Anzahl von Krankheiten entstehen dadurch erhebliche Schwierigkeiten. Interessante Krankheiten und Untersuchungen werden mit Hilfe der Mechanismen Fokussierung und Abstraktion gewonnen. Die Fokussierung wählt die zu betrachtenden Objekte aus, die Abstraktion faßt die Eigenschaften der Objekte zusammen.

Unsicheres Wissen wird fokussiert und abstrahiert, um interessante Untersuchungen und Krankheiten auszuwählen. Fehlerhafte Eingaben können mit Hilfe eines Recovery-Mechanismus zurückgezogen werden.

Durch die zahlreichen unären und binären Relationen auf den Grundobjekten wird ein semantisches Netz zur Darstellung der Zusammenhänge erstellt. Weder ein kausales Modell noch Hintergrundwissen stehen zur Verfügung. Zeitliche Aspekte wurden nur durch die Zeitindizierung der Informationsvektoren und zwei zeitmodale Operatoren beschrieben. Im Gegensatz zu Abel und Med2 gibt es keine strukturierten Regeln, sondern nur eine Folge von Produktionsregeln.

Wissensbasis: Symptome, Diagnosen regelmäßige Ableitungsmechanismen und Verbindungen zwischen Symptomen und Diagnosen. 0	Abstraktionsmechanismen für Krankheiten und Untersuchungen, Relationen zwischen Krankheiten und Symptomen Symptomwald	Abstraktionsregeln, Symptom-Symptom Relationen, Symptom-Diagnose-Relationen, Diagnose-Diagnose-Relationen (-> 2.3.4)
Notfalldiagnostik soll möglich sein. +	Mechanismen zur Notfalldiagnostik	Die Abstraktionsregeln weisen den Hypothesen, die lebensbedrohend und dringend sind, die höchste Priorität zu.

3. M E D 2

Med2 ist ein Expertensystemshell für medizinische Diagnostik und ähnliche Anwendungsbereiche. Im Unterschied zu vielen anderen Expertensystemen ist Med2 speziell auf einen Problemtyp zugeschnitten, nämlich die assoziative Diagnostik. In Med2 ist die Kombination der Mechanismen -hypothetisch deduktive Vorgehensweise und Differentialdiagnostik- mit dem Working-Memory-Konzept realisiert. Es orientiert sich am Vorbild des menschlichen Kurzzeitgedächtnisses und enthält immer nur eine kleine Anzahl von Verdachtsdiagnosen. Zur Effizienzverbesserung kann in großen Wissensbasen die Kontextabhängigkeit des Regelwissens realisiert werden. Med2 trennt auch zwischen der Vorverarbeitung (database reasoning) und der eigentlichen diagnostischen Auswertung (diagnostic reasoning) der Daten.

3.1 Anforderungen aus der Problemstellung.

Problemstellung:

Expertensystemshell zur Diagnose und Therapie in medizinischen und technischen Anwendungsgebieten.

Anforderungen:

- 1) Plausibilität der Eingabe soll überprüft werden.
- 2) Die Diagnosefindung soll trotz unsicherem und unvollständigem Wissen möglich sein.
- 3) Mehrfachdiagnosen sollen erkannt werden.
- 4) Folgesitzungen sollen ausgewertet werden können.
- 5) Widersprüche sollen geeignet behandelt werden.
- 6) Die Wissensbasis soll modular aufgebaut sein.
- 7) Die Diagnosefindung soll möglichst effizient durchgeführt werden.
- 8) Ähnliche Diagnosen sollen unterschieden werden können.
- 9) Die Dialogkomponente soll für alle Eingabearten geeignet sein.
- 10) Die Erklärungskomponente soll sehr ausführlich sein und gleichzeitig beim Ausfüllen der Shell in ihrem Umfang variiert werden können.
- 11) Der Benutzer soll jederzeit die Möglichkeit haben, aktiv in das Diagnosefinden einzugreifen.
- 12) Grobe Zeitangaben sollen repräsentiert und verarbeitet werden.
- 13) Eine Notfalldiagnostik soll möglich sein.

3.2 Funktionale Sicht von Med2.

- 1) Plausibilitätskontrolle der Eingabe.
- 2) Wissensbasis:
 - hierarchische Anordnung von Symptomen und Diagnosen
 - Zusammenfassung von Fragen, die in eine Richtung weisen
 - Gruppierung von Regeln.
- 3) Differentialdiagnostik.
- 4) Auswertung von Folgesitzungen.
- 5) Mechanismen zur Eingabe und Verarbeitung quantitativer und qualitativer Zeitangaben.
- 6) Effizienter Problemlösungsmechanismus.
- 7) Mechanismen zur Repräsentation und Verarbeitung von unsicherem und unvollständigem Wissen.
- 8) Recovery-Mechanismus.
- 9) Möglichkeit zur Vereinfachung des Hauptzieles.
- 10) Mechanismen zum Erkennen mehrerer Krankheiten.
- 11) Überprüfung der Enddiagnose anhand von Diagnose - Symptom - Regeln.
- 12) Mechanismen, die ein Eingreifen des Benutzers jederzeit ermöglichen.
- 13) Erklärungskomponente:
 - Informationen über den aktuellen Systemzustand
 - Erklärung des Ergebnisses.
- 14) Widerspruchsbehandler:
 - Entschuldigungsmechanismus.
- 15) Mechanismus zur Notfalldiagnostik.

3.3 Architektursicht von Med2.

3.3.1 Wissensrepräsentation.

In Med2 gibt es sieben verschiedene Frame-Typen. Jedem Frame wird ein Name zugeordnet. Nachfolgend werden die einzelnen Frames mit ihren jeweiligen Informationen aufgelistet.

1. **Pathokonzepie (Kontext-, Grob- und Feindiagnosen)**
 - die Differentialdiagnosen des Pathokonzepies
 - die Nachfolgerdiagnosen innerhalb der explizit aufgestellten Diagnosehierarchie (ohne Zirkularitäten) (diese Diagnosehierarchie ist eine Teilmenge des durch Regeln implizit aufgestellten Diagnosenetzwerks. (+ Zirkularitäten!))
 - der nach der Etablierung des Pathokonzepies zu behandelnde Questionset
 - die Questionsets, die zur Überprüfung des Pathokonzepies geeignet sind

2. **Manifestationen (Fragen und einfache Symptominterpretationen)**

- Zuordnung zu den entsprechenden Questionsets
- Fragetext für den Benutzer
- Erklärungen zur Frage, die sich der Benutzer bei Bedarf ansehen kann
- Wertebereich der Manifestation
- zeitliche Charakterisierung der Frage (In welchen Sitzungen soll die Frage gestellt werden? Nur in der ersten Sitzung? Nur in Folgesitzungen?)

3. **Questionsets: Gruppenbildung zusammengehöriger Manifestationen**

- Top-Level-Frage eines Questionsets
- Fragetext (Das Questionset enthält nicht nur eine Menge von Fragen, sondern ist selber eine Frage und zwar: Soll dieser Questionset erfragt werden oder nicht?)
- Erklärungen zur Frage
- Zuordnung von Therapien und Therapievorschlägen
- Nachfolger-Questionsets in der Klassifikationshierarchie

4. **Varianten: Therapeutische Varianten der Pathokonzepie**

- allgemeiner Therapievorschlag (allgemeine Richtlinien)
- spezieller Therapievorschlag (spezifische Therapievorschläge)
- Dringlichkeit der Therapie
- Prognose

5. **Explanationsets: Erfassung der Erklärungsbedürftigkeit der Symptomatik**

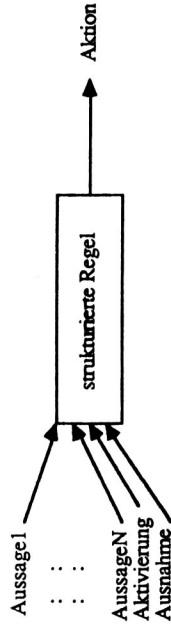
- Liste von Pathokonzepien die das Explanationset erklären kann
- Questionsets in denen wenigstens eine der Manifestationen vorkommt, die dem Explanationset zugeordnet sind.

6. **Lokalisationen**

- Menge von Unterlokalisationen (below-Listen)
- Achsensymmetrie

7. **Regeln** drücken die Beziehungen zwischen den Objekten aus.

In Med2 wie auch in Abel liegen strukturierte Regeln vor:



Die Aktivierung ist eine zusätzliche 'Kontext'-Bedingung. Sie führt dazu, daß die Regel nur in einem bestimmten Zusammenhang (Kontext) ausgeführt wird. z.B. nur falls ein bestimmtes Pathokonzepie (= meist eine allgemeine Diagnosekategorie) etabliert ist. Regeln sind typisiert. Insgesamt gibt es 16 Typen. Einige davon sind:

- f, f&b, b : zur Verdachtsgenerierung und Verdachtsüberprüfung
- ask, asknot : zum Stellen von Fragen
- v : zum Herleiten von Varianten einer Therapie

3.3.2 Inferenzstrategie (= Reasoning-Strategie)

Med2 unterscheidet zwischen der Vorverarbeitung (Database Reasoning) und der eigentlichen diagnostischen Auswertung (Diagnostic Reasoning) der Daten.

3.3.2.1 Database Reasoning (= einfache Symptominterpretationen).

Med2 beginnt mit allgemeinen Fragen und präzisiert positive Antworten sofort. Diese 'Rohdaten' werden so weit wie möglich aufgearbeitet. Beispiele hierfür sind die Umwandlung quantitativer in qualitative Daten und die Berechnung nützlicher Indizes oder Zeitrelationen (Dauer eines Symptoms aus Anfang und Ende).

3.3.2.2 Diagnostic Reasoning (= Verdachtsgenerierung und -überprüfung von Hypothesen).

Med2 benutzt die Hypothesize-and-Test-Strategie, um sich auf die wichtigsten Verdachtsdiagnosen konzentrieren zu können. Die Regeln zur Bewertung von Diagnosen haben drei verschiedene Typen

Forward (F), Forward&Backward (F&B), Backward(B).

Durch diese Regeln werden Punktekategorien (+20, -10) auf ein Objekt übertragen. Die Generierung eines Diagnoseverdachtses ist nur mit F- und F&B-Regeln möglich, wobei die F-Regeln ausschließlich der Verdachtsgenerierung dienen. Die F&B-Regeln können gleichzeitig auch die Diagnose bewerten. Falls die Bewertung einen Schwellwert überschreitet, werden die B-Regeln zu ihrer vollständigen Bewertung untersucht. In Abhängigkeit vom Ergebnis kommt die Diagnose in das WM.

Deshalb sind immer nur einige wenige diagnostische Regeln aktiv:

- von den Forward-Regeln nur diejenigen, deren Kontext etabliert ist
- von den Backward-Regeln nur diejenigen, deren Hypothese im WM ist

3.3.3 Repräsentation der Zeit.

Med2 hat zeitbezogenen Fragetypen nach 'Dauer', 'Frequenz' und 'Zeitpunkt'. Aus diesen Rohdaten kann Med2 folgende Schlussfolgerungen herleiten:

- Dauer eines Symptoms aus Beginn und Ende.
- Zeitliche Beziehung zwischen zwei Symptomen, wobei die Beziehung durch ein Zeitintervall präsentiert werden kann (vor, gleichzeitig, nach).
- In Folgesitzungen: Umrechnung aller Zeitangaben, die sich auf den alten Untersuchungszeitpunkt beziehen.
- In Folgesitzungen: Erkennen einer zeitlichen Änderung eines Symptoms (mit den Präkaten Zunahme, Konstanz und Abnahme) unter dem Einfluß besonderer Ereignisse.

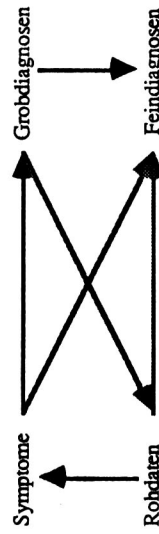
Wichtig ist auch einen ausreichenden Grad an Ungenauigkeit zuzulassen. Deshalb ist es in Med2 möglich lokale und nicht nur globale Bezugspunkte zu wählen oder sich auf verschiedene Zeiteinheiten (Jahr, .. Sekunden) zu beziehen. Auch ungenaue Zeitangaben (2-5 Wochen) können angegeben werden. In Med2 liegt eine punktbasierte Repräsentation der Zeit vor (Im Gegensatz zur intervallbasierten Repräsentation bei Hiqual).

3.3.4 Grobdiagnosen (diagnostischer Mittelbau).

In Med2 gibt es eine Diagnosehierarchie und eine Diagnoseheterarchie, die explizit durch Verweise in den Pathokonzepten aufgebaut werden. Zusätzlich gibt es noch ein Diagnosenetzwerk, welches implizit durch Diagnose-Diagnose-Regeln entsteht. Zur Aktivierung der Nachfolgerdiagnosen dienen Diagnose-Diagnose-Regeln, die in Abhängigkeit des Status einer Diagnose (etabliert oder ausgeschlossen) und eventuell weiteren Bedingungen Evidenz auf andere Diagnosen übertragen. Durch Regeln zur Verdachtsgenerierung kann die Hierarchie übersprungen und direkt eine Feindiagnose etabliert werden.

Der diagnostische Mittelbau besteht aus einfachen Symptominterpretationen, die sofort bei der Datenerfassung durchgeführt werden und komplexeren Interpretationen (Grobdiagnosen), die schrittweise zu Enddiagnosen verfeinert werden können.

Datenvorverarbeitung diagnostische Auswertung



Rohdaten	Bezinverbrauch =10 Liter	Puls 100 Blutdruck 80
Symptominterpretation	typspezifischer Benzinverbrauch zu hoch	hoher Schockindex
Grobdiagnose	Verbrennung	Kreislaufschock
Feindiagnose	Leerlaufsystem	Kreislaufschock durch starken Flüssigkeitsverlust

Abb. 1.23: Struktur des diagnostischen Mittelbaus.

Die Vorteile dieses diagnostischen Mittelbaus sind die Einschränkung des Suchraumes, die Reduktion der Unsicherheit und die Darstellung von partiellem Wissen.

3.3.5 Das Working-Memory-Konzept.

Das Working-Memory (WM) ist ein Konzept zur Realisierung der Hypothesize-and-Test-Strategie, bei der mehrere Hypothesen gleichzeitig verfolgt werden. Das WM in Med2 enthält immer alle verdächtigen, aber noch nicht etablierten Diagnosen:

==> o Es ist ein direkter Vergleich zwischen den Differentialdiagnosen möglich, bevor diese etabliert werden.

- o Indikation von Questionsets zur gleichzeitigen Aufklärung mehrerer Diagnosen im WM. Es werden somit nicht nur Fragen nach dem Spitzenreiter gestellt.
- o Erklärung des Systemvorgehens.
(In einem Window werden ständig alle neuen Verdachtsdiagnosen ausgegeben.)

Die Aufnahme ins WM erfolgt, falls einer der folgenden Punkte erfüllt ist.

- (1) Die Punktschritte der Forward-Regeln überschreitet einen bestimmten Schwellwert und nach der Aktivierung der Backward-Regeln ist dieser immer noch überschritten.
- (2) Eine Diagnose D1 ist in der Diagnosehierarchie Vorgänger einer Diagnose D2 und letztere ist im WM.
- (3) Eine Diagnose ist Differentialdiagnose einer auf Etablierung geprüften Diagnose.
- (4) Wenn sich der Benutzer für bestimmte Diagnosen besonders interessiert, kann er sie im WM verankern, so daß sie ständig mituntersucht werden.

3.3.6 Dialogsteuerung.

In Med2 können alle Dialog-Modi kombiniert werden. Der Benutzer kann aktiv oder passiv seine Daten eingeben oder das Expertensystem verarbeitet die Daten von Meßgeräten direkt. Falls der Benutzer aktiv Symptome eingeben will, wählt er in hierarchisch strukturierten Menüs die Questionsets aus, die seine Beobachtungen umfassen. Gibt der Benutzer keine Questionsets vor, so ist der Benutzer passiv und Med2 übernimmt die Initiative:

1. Med2 überprüft, ob Questionsets aufgrund kategorischer Regeln indiziert sind. Wenn ein Pathokonzept oder eine Manifestation etabliert worden ist, dann kann eine Menge von Questionsets indiziert werden, die routinemäßig zur genaueren Abklärung des Zustandes führt (erfahrungsgesteuert).

2. Falls diese Strategie zu keinem Ergebnis führt, wird der Questionset ausgewählt, der zur Überprüfung der Diagnosen im WM am Nützlichsten ist

- Abklärung einer verdächtigen Diagnose
- Differentialdiagnosen unterscheiden

(zielgerichtet).

3. Bei sehr allgemeiner Symptomatik kann es passieren, daß überhaupt keine brauchbaren Verdachtshypothesen herleitbar sind. Zur Klärung solcher Situationen gibt es in Med2 ganz allgemeine Kontext-Diagnosen, die dann die Dialogsteuerung übernehmen, um eine systematische Erfassung der Symptomatik zu ermöglichen (systematisch).

Ein Vorteil des Questionset-Konzeptes zur Dialogsteuerung ist die Bündelung zusammengehöriger Fragen zu einer Steuerungseinheit. Als Nachteil könnte sich die zu detaillierte Erfassung der Symptome erweisen. Dieses Problem wird aber durch die hierarchische Organisation der Fragen innerhalb der Questionsets ausgeglichen.

3.3.7 Bewertungsarten.

Die Bewertung der Diagnosen in Med2 wird mit Regeln durchgeführt. Es gibt nur einen groben Bewertungsmaßstab, aber vielfältige Bewertungskriterien:

- Kategorische und probabilistische Gewichtung mit Prädispositionen.
(Grunddaten wie Alter, Geschlecht)
- Differentialdiagnostik
- Regeln mit Ausnahmen
- Dynamisches Generieren und Zurückziehen von Schlußfolgerungen.
- einfache kausale Plausibilitätskontrolle
(Diagnose-Symptom-Repräsentation innerhalb der Explanationssets)

Für die Diagnosebewertung mit unvollständigem Wissen gibt es drei Grundtechniken:

- stufenweises Vorgehen (diagnostischer Mittelbau)
- nicht-monomotones Schließen
- Wald der Symptome (Verfeinerung)

3.3.8 Nicht-monotones Argumentieren (Belief revision).

Mit Hilfe des ITMS-Algorithmus können in Med2 Regeln wieder zurückgezogen werden. Dazu werden die betroffenen Schlussfolgerungen identifiziert und die entsprechenden Ableitungen invertiert. Dies ist vorteilhaft für folgende Situationen:

- Der Benutzer korrigiert eine frühere Eingabe.
- Die Ausnahme einer früheren Regel wird bekannt.
- Die Bewertung einer bereits etablierten Diagnose verschlechtert sich durch neue Symptome erhebllich.
- Änderungen der Symptomatik werden in Folgesitzungen erfasst.

Die Ausnahmen von Regeln eignen sich gut zur Behandlung schwer lösbarer Probleme wie den variablen Detaillierungsgrad. Auch die Vermeidung probabilistischer Bewertung ist möglich, indem die Regel kategorisch angegeben wird und eine Liste von Ausnahmen explizit formuliert wird. Das Tuning der Wissensbasis wird vereinfacht. Jetzt können spezielle Situationen als Ausnahmen noch nachträglich eingefügt werden. Eine direkte Repräsentation von Oder-Verknüpfungen ist in Med2 nicht erlaubt. Deshalb wird eine solche Regel in zwei Teile aufgespalten.

(A oder B \implies x " = " A \implies X; B \implies X falls nicht A (sonst wird B doppelt gezählt))

3.3.9 Grober Ablauf in Med2.

Zu Beginn der Sitzung mit Med2 wählt der Benutzer aus einem Menü die Questionsets aus, die seine Leitsymbole und Befunde charakterisieren. Die Abarbeitung der Questionsets erfolgt in zwei Stufen:

(1) Database Reasoning

Die Symptome werden erfasst und es werden einfache Symptominterpretationen durchgeführt.

(2) Diagnostic Reasoning

- Verdachtsgenerierung (dazu f und f&b Regeln)

- Verdachtsüberprüfung (dazu f&b und b Regeln)

Danach wird entschieden, welche Pathokonzepte in das Working-Memory aufgenommen werden oder herausfallen und ob Pathokonzepte etabliert werden können, weil ihre Gesamtbewertung absolut und relativ zu den Differentialdiagnosen gut genug ist. Falls ein Pathokonzept etabliert wird, werden seine Symptome als erklärt markiert und seine Nachfolger aktiviert. Anschließend beginnt ein neuer Zyklus mit der Auswahl eines Questionsets durch den Benutzer oder das System.

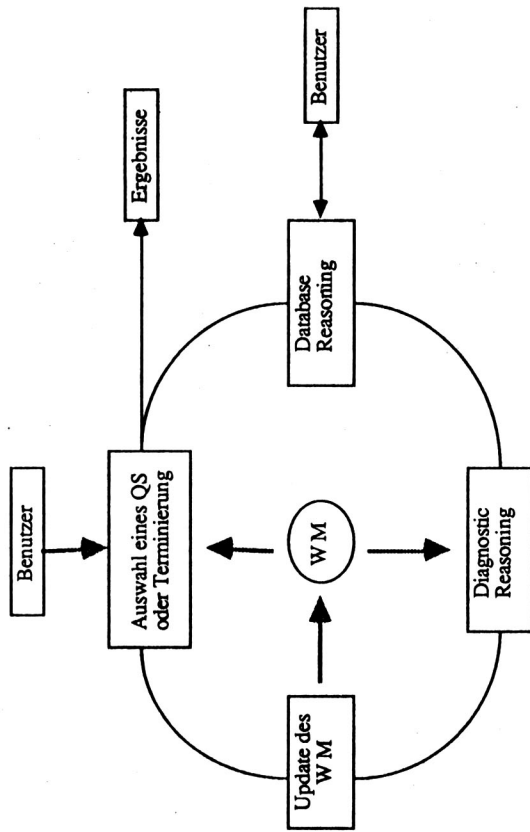


Abb. 1.24: Steuersystemzyklus in Med2

3.4 Bewertung des Expertensystems Med2.

3.4.1 Vergleich der Anforderungen aus der Problemstellung mit der funktionalen und der Architektursicht.

Anforderungen	funktionale Sicht	Architektursicht
Die Plausibilität der Eingabe soll überprüft werden. +1	Plausibilitätskontrolle der Eingabe	- Jeder Frage ist ein Wertebereich zugeordnet. - Konsistenzregeln - Vererbungsmechanismus im Hierarchiebaum der Symptome
<i>Die Plausibilitätskontrolle wird auf drei voneinander unabhängige Arten gelöst.</i>		
Die Diagnosefindung soll trotz unsicherem und unvollständigem Wissen möglich sein. +	Mechanismen zur Repräsentation und Verarbeitung von unsicherem und unvollständigem Wissen.	- nicht-monotones Argumentieren - strukturierte Regeln - probabilistische und kategoriale Bewertung mit Prädispositionen. - stufenweise Vorgehensweise - hierarchischer Aufbau der Symptome
<i>Gute und vielfältige Möglichkeiten.</i>		
Mehrfachdiagnosen sollen erkannt werden. -	Mechanismen zum Erkennen mehrerer Krankheiten.	Die Differentialdiagnostik ist auf Gruppen von Diagnosen beschränkt.
<i>Untypische Symptomanpassungen werden nur erkannt, wenn sie explizit in der Wissensbasis repräsentiert sind.</i>		

1+ : Die Anforderungen werden sehr gut erfüllt.
 0 : Die Anforderungen werden gut erfüllt.
 - : Die Anforderungen werden nur teilweise oder gar nicht erfüllt.

Folgesitzungen sollen ausgewertet werden können. 0	Auswertung von Folgesitzungen.	Abspeicherung und Auswertung der Geschichte von Symptomen. Automatische Umrechnung aller auf den alten Untersuchungszeitpunkt bezogenen Zeitangaben.
Widersprüche sollen geeignet behandelt werden. 0	Widerspruchsbehandler mit Entschuldigungsmechanismus	Konsistenzüberprüfung der Eingabe, Ausnahme von Regeln, Rücknahme von Schlussfolgerungen (ITMS-Algorithmus).
Die Wissensbasis soll modular aufgebaut sein +	Wissensbasis: hierarchische Anordnung von Symptomen und Diagnosen, Gruppierung von Regeln, Zusammenfassung von Fragen, die in eine Richtung weisen.	Es gibt sieben verschiedene Frame-Typen. Symptome werden zu Questionsets zusammengefaßt und bezüglich der Relation Verfeinerung hierarchisch angeordnet. strukturierte Regeln, denen jeweils ein Typ zugeordnet ist. Diagnosehierarchie, -heterarchie und -netzwerk Relationen zwischen Symptomen und Diagnosen und viele Querverweise.
<i>Hoher Grad an Modularität durch die Diagnose- und Symptomhierarchien und die Gruppierung von Regeln und Fragen. Dadurch ist die Anwendbarkeit für größere Anwendungssysteme gewährleistet.</i>		
Die Diagnosefindung soll möglichst effizient durchgeführt werden. +	effizienter Problemlösungsmechanismus Recovery-Mechanismus Mechanismus zur Vereinfachung des Hauptzieles	WM-Konzept mit hypothetisch-deduktiver Vorgehensweise und anschließender Differentialdiagnostik Explizite Repräsentation von Kontextabhängigkeiten. ITMS-Algorithmus Unterscheidung zwischen Fein- und Grobdiagnosen.
<i>Effizienter Problemlösungsmechanismus, schnelles Zurückziehen fehlerhafter Daten oder Regeln ist möglich</i>		

Ähnliche Diagnosen sollen unterschieden werden können. 0	Mechanismus zur Differentialdiagnose	Zu jedem Pathokzept werden die entsprechenden Differentialdiagnosen angegeben. Soll ein Pathokzept etabliert werden, so werden zuerst alle Differentialdiagnosen im WM untersucht.
Die Dialogkomponente soll für alle Eingabearten geeignet sein. 0	geeignete Eingabekomponente	- der Benutzer kann aktiv oder passiv Daten eingeben oder das Exps verarbeitet Daten von Meßgeräten direkt. - verschiedene Antworttypen sind möglich.
Die Erklärungskomponente soll sehr ausführlich sein und gleichzeitig beim Ausfüllen der Shell in ihrem Umfang variiert werden können. 0	Erklärungskomponente	Transparenz des Systemvorgangs, aktueller Systemzustand und Begründung von Schlußfolgerungen, allgemeine Informationen über die Wissensbasis, auch bezüglich eines Objektes.
Grobe Zeitangaben sollen repräsentiert und verarbeitet werden. +	Mechanismen zur Eingabe und Verarbeitung quantitativer und qualitativer Zeitangaben.	Zeitrechnungen: Dauer aus Anfangs- und Endpunkt, zeitliche Beziehung zw. zwei Symptomen (vor, nach) Punktbasierete Repräsentation der Zeit, einfache Symptominterpretationen.
Der Benutzer soll jederzeit die Möglichkeit haben, aktiv in das Diagnosefinden einzugreifen. 0	Mechanismen, die ein Eingreifen des Benutzers jederzeit ermöglichen	- Der Benutzer kann eine Diagnose im WM verankern. - Der Benutzer kann die Questions aktivieren, deren Fragen als nächstes gestellt werden.

Eine Notfalldiagnostik soll möglichst sein.	Mechanismus zur Notfalldiagnostik	- Dialogsteuerung konzentriert sich nur auf die wichtigsten Symptome. - Vordringliche Untersuchung von Diagnosen, die sofortiges Handeln erfordern. - sofortige Ausgabe von Handlungsanweisungen nach Etablierung einer Notfalldiagnose
+		

3.4.2 Zusammenfassende Bewertung.

Med2 ist ein sehr gutes System für assoziative Diagnostik. Alle Objekte der verschiedenen Frametypen sind in Hierarchien oder Heterarchien angeordnet. Dies führt zusammen mit den Questionsets zu einem hohen Modularitätsgrad. Durch die Hierarchie der Diagnosen können besonders zu Anfang schon sehr viele Krankheiten ausgeschlossen werden. Die hierarchische Anordnung bzw. die Kontextbedingungen der Regeln fangen Extremfälle wie eine große Zahl von Diagnosen, Symptomen oder Regeln ab.

Med2 kann Mehrfachdiagnosen stellen, da die Differentialdiagnostik auf Gruppen von Diagnosen beschränkt ist. Untypische Symptomausprägungen werden nur dann erkannt, wenn diese explizit in der Wissensbasis repräsentiert sind (z.B. als Ausnahme von Regeln), da keine kausalen Modelle darstellbar sind. Die Erklärungskomponente ist für ein assoziatives System sehr ausgebaut. Es gibt viele verschiedene Möglichkeiten, Fragen in Bezug auf bestimmte Objekte zu stellen. Hintergrundwissen und kausales Wissen stehen nicht zur Verfügung.

Widersprüchliche Evidenz wird erkannt, da positive und negative Evidenz in Med2 getrennt abgespeichert wird. In den Ausnahmen der Regeln sind Voraussetzungen für ihre Gültigkeit explizit repräsentiert, die beim Vorliegen widersprüchlicher Evidenz gezielt überprüft werden können. Sehr vorteilhaft wirkt sich die Möglichkeit des Nichtmonotonen-Argumentierens aus.

Bei der Wissensrepräsentation fehlt die Darstellung und Auswertung von Falldatenbanken und kausalem Wissen. Ungeeignet ist Med2 für die Lösung von Konstruktionsproblemen und solchen Diagnostikanwendungen, für die nur kausale Modelle und keine Erfahrungsregeln bekannt sind.

Desweiteren müssen die Ausprägungen der Symptome entweder benannt oder als numerische oder zeitbezogene Daten angegeben werden können. Andere Arten von Symptomausprägungen wie Beziehungen zwischen geometrischen Objekten oder Fehlermeldungen des Betriebssystems müssen erst in das oben angegebene Format vorverarbeitet werden (Lisp-Schnittstelle).

Der Einsatz für Laien ist nur mit Einschränkungen möglich.

Kapitel 2 : Vergleich der drei Diagnoseexpertensysteme.

Diese drei Expertensysteme zu vergleichen ist schwierig, weil ihre Grundstrukturen unterschiedlicher Art sind und die Systeme verschieden weit entwickelt wurden.

Während Abel modellbasiert ist, sind das Ohrendiagnosesystem und Med2 assoziative Systeme. Med2 ist im Gegensatz zu den beiden anderen Systemen ein Expertensystemshell.

Bei Abel und dem Ohrendiagnosesystem wurde vorrangig eine neue Idee entwickelt, wodurch Komponenten, wie die Zeitbehandlung und der Dialog zwischen dem Benutzer und dem System nur sehr gering ausgearbeitet wurden. Med2 hingegen ist ein System, bei dem schon alle Komponenten vollständig entwickelt sind. So besteht die Gefahr, daß Abel und dem Ohrendiagnosesystem negative Kritikpunkte zugeordnet werden, die nicht durch das System an sich, sondern nur durch ihren Entwicklungszustand hervorgerufen werden.

Um diese Schwierigkeiten zu vermeiden, geht man im folgenden Vergleich von den Anforderungen an die einzelnen Expertensysteme aus. Wurde die Anforderung an mehrere der Expertensysteme gestellt, so wird deren funktionale und Architektursicht verglichen.

Anforderung 1:

Die Eingaben sollen bezüglich ihrer Konsistenz überprüft werden.

Diese Anforderung wird von allen drei Problemstellungen gefordert.

- Abel: Obwohl gerade bei Laborwerten eine Konsistenzüberprüfung sehr wichtig wäre, ist diese Anforderung in Abel nicht realisiert worden.

o ODS: Jedem Symptom wird ein Wertebereich zugeordnet.

Vererbungsmechanismus in der Symptomhierarchie.

+ Med2: wie ODS

zusätzlich: Konsistenzregeln

Anforderung 2:

Größere Probleme sollen in Teilprobleme aufgeteilt werden.
Diese Anforderung wird an das Ohrendiagnosesystem nicht gestellt.

- + Abel: Die Kontrollstruktur teilt das Hauptziel entsprechend dem PSM und der Diagnosehülle in Unterziele auf.
- o Med2: In Med2 wird zwischen Grob- und Feindiagnosen unterschieden. Dadurch wird das Problem "Finden einer Feindiagnose" in die Teilprobleme "Ausschließen beziehungsweise Etablieren der Grobdiagnosen" aufgespalten.

Anforderung 3:

Auch bei unvollständigem und unsicherem Wissen soll das Problem möglichst gut gelöst werden.

Diese Anforderung wird an alle drei Expertensysteme gestellt.

- + Abel: - strukturierte Regeln
 - verschiedene Abstraktionsebenen
- + ODS: - strukturelle Bewertung des PSM und der Diagnosehüllen
 - verschiedene Abstraktionsebenen
 - Abstraktionsregeln
- + Med2: - Informationsvektor
 - unscharfe Grenzen in der Logik der Krankheitsformeln
- nicht monotonen Argumentieren
- probabilistische und kategorische Bewertung mit Prädispositionen
- einfache kausale Plausibilitätskontrolle
- stufenweise Vorgehensweise
- Verfeinerungshierarchie der Symptome

Anforderung 4:

Finden der korrekten Diagnose a) und Therapie

b) auch Mehrfachdiagnosen

Die Anforderungen a) und b) werden an das Ohrendiagnosesystem nicht gestellt.

- a) - Abel: Bisher wurde diese Anforderung noch nicht realisiert.
- o Med2: Med2 gibt auch mögliche Therapien aus. (-> Varianten)

- b) + Abel: Simulationsmechanismus mit Rückkopplungsschleifen

- Med2: Die Differentialdiagnostik ist auf Gruppen von Diagnosen beschränkt.
Untypische Symptomausprägungen werden aber nur erkannt, wenn sie explizit in der Wissensbasis repräsentiert sind.

Anforderung 5:

Folgesitzungen sollen entsprechend ausgewertet werden.
Diese Anforderung wird an das Ohrendiagnosesystem nicht gestellt.

- Abel: Speicherung des letzten PSM.

Alle pathophysiologischen Zustände des PSM haben die Attribute Startpunkt und Dauer

- o Med2: - Abspeicherung und Auswertung der Geschichte von Symptomen

- Automatische Umrechnung aller auf den alten Untersuchungszeitpunkt bezogenen Zeitangaben.

Anforderung 6:

Ähnliche Diagnosen sollen unterschieden werden.

Diese Anforderung wird an alle drei Expertensysteme gestellt.

- o Abel: Zu jeder Diagnose sind die Differentialdiagnosen und Komplikationen bekannt.
(-> Klassifikation der Diagnosen)

- o ODS: Jeder Diagnose werden ihre Differentialdiagnosen und Komplikationen zugeordnet und im Abstraktionsmechanismus berücksichtigt.

- o Med2: Zu jedem Pathokonzzept werden die dazugehörigen Differentialdiagnosen angegeben. Soll ein Pathokonzzept etabliert werden, so werden zuerst alle Differentialdiagnosen im WM untersucht

Anforderung 7:

Sich über die Zeit verändernde Werte sollen aggregiert werden.

Diese Anforderungen wird an das Ohrendiagnosesystem und an Med2 nicht gestellt.

- o Abel: Mechanismus zur Aggregation der Zeit.

Anforderung 8:

Grobe Zeitangaben sollen dargestellt und verarbeitet werden.
Diese Anforderung wird an alle drei Expertensysteme gestellt.

o Abel: Alle pathophysiologischen Zustände von Abel besitzen die Attribute Dauer und Zeitpunkt.

o ODS: Der Informationsvektor wird mit dem Zeitpunkt der Untersuchung indiziert.
Modaloperationen wie exists und persists

Nur quantitative Zeitangaben können verarbeitet werden.

+ Med2: Zeitverarbeitungsmechanismen für quantitative und qualitative Zeitangaben.

- Dauer aus Anfangs- und Endzeitpunkt.

- Zeitliche Beziehungen zwischen zwei Symptomen (vor, gleichzeitig, nach)
(-> Folgesitzungen)

Anforderung 9:

Fehlerhafte Eingaben sollen schnell und einfach zurückgezogen werden.

Diese Anforderung wird an alle drei Expertensysteme gestellt.

o Abel: Das PSM wird verändert und ein neuer Problemlösungszyklus wird gestartet.

o ODS: Wird eine Belegung einer früheren Anamnesephase zurückgezogen, so ist die Hypothesenmenge nach der Recoveryanwendung größer oder gleich der korrekten Hypothesenmenge.

+ Med2: ITMS-Algorithmus (schnell und exakt)

Anforderung 10:

In der Wissensbasis soll das notwendige Wissen enthalten sein.

Diese Anforderung wird an alle drei Expertensysteme gestellt.

+ Abel: - Hierarchien der Anatomie: part-of, enthalten-in, Materialflußrelation,
grobe anatomische Beziehungen

- Klassifikation der Krankheiten nach Ätiologie, Zeit, Anatomie und Physiologie

- Kausale Links.

o ODS: - Symptom-Symptom-Relationen

- Symptom-Diagnose-Relationen

- Diagnose-Diagnose-Relationen

- Abstraktionsregeln

- Symptome und Diagnosen

o Med2: modularer Aufbau der Wissensbasis:

- Symptom-, Diagnosehierarchien strukturierte Regeln,
- Relationen zwischen Symptomen, Diagnosen und Symptomen und Diagnosen
- Explanationssets, Varianten, Questionsets

Anforderung 11:

Symptome sollen gruppiert werden.

Diese Anforderung wird an das Expertensystem Abel nicht gestellt.

+ ODS: Verfeinerungshierarchien und zusätzliche Nachbarschaftsrelationen

o Med2: Verfeinerungshierarchien (Questionsets)

Anforderung 12:

Diagnosen sollen gruppiert werden.

Diese Anforderung wird an das Ohrdiagnosesystem nicht gestellt.

o Abel: Klassifikation der Diagnosen nach Wissen über Ätiologie, Anatomie, Physiologie und Zeit.

+ Med2: - Diagnosehierarchie und Diagnoseheterarchie, die explizit dargestellt werden

- Diagnosenetzwerk (entsteht durch Diagnose-Diagnose-Regeln)

Anforderung 13:

Eine Notfalldiagnostik soll möglich sein.

Diese Anforderung wird an alle drei Expertensysteme gestellt.

- Abel: In Abel wird die Anforderung nicht realisiert.

o ODS: Die Abstraktionsregeln weisen den Hypothesen, die lebensbedrohend und dringend sind, die höchste Priorität zu.

o Med2: - Die Dialogsteuerung konzentriert sich nur auf die wichtigsten Symptome.
- Vordringliche Untersuchungen von Diagnosen, die sofortiges Handeln erfordern.

- Sofortige Ausgabe von Handlungsanweisungen nach Etablierung einer Notfalldiagnose.

Anforderung 14:

Klinisches und pathophysiologisches Wissen soll gleichzeitig verarbeitet werden. Diese Anforderung wird nur an Abel gestellt.

- o Abel: PSM mit verschiedenen Abstraktionsebenen und Mechanismen, die die einzelnen Ebenen ineinander überführen und verbinden.

Anforderung 15:

Freie Entscheidungen des Benutzers sollen jederzeit möglich und in die Diagnosefindung integrierbar sein.

Diese Anforderung wird an alle drei Expertensysteme gestellt.

- Abel: Diese Anforderung wird in Abel kaum realisiert.
- + ODS: Eingaben sind jederzeit möglich.

Das System liefert Vorschläge für eine Untersuchung, der Benutzer kann aber auch andere Untersuchungen durchführen (passiv ausgelegtes System).

- o Med2: Eingaben sind jederzeit möglich.

Der Benutzer kann eine Diagnose in das WM einbinden und andere als die vorgeschlagenen Untersuchungen durchführen.

Abel ist ein sehr umfangreiches System, das bisher noch nicht vollständig implementiert werden konnte. Durch den Aufbau des kausalen Modells hat Abel aber gegenüber den anderen Systemen folgende Vorteile:

- Für die Erklärungen kann man je nach dem Wissen des Benutzers eine Ebene des kausalen Modells wählen.
- Mehrere gleichzeitig auftretende Krankheiten, auch solche deren Symptome sich gegenseitig verändern, können erkannt werden.
- Die Wissensrepräsentationen bezüglich der Anatomie, Ätiologie u.ä. sind sehr umfangreich.

Viele Anforderungen, die aus der Problemstellung an Abel gestellt werden, sind noch nicht oder nur teilweise realisiert worden.

Med2 bietet eine äußerst flexible und schnelle Widerspruchsbehandlung. Es ist auch in größeren Anwendungsgebieten gut einsetzbar, da Diagnosen, Symptome und Fragen hierarchisch angeordnet sind. Auch die Regeln sind je nach Kontext in Gruppen eingeteilt. Die Zeitkomponente und die Bewertungsmöglichkeiten sind in Med2 sehr gut. Med2 bietet zwar kein kausales Modell, aber man versucht aus dem vorhandenen assoziativen System so viel wie möglich herauszuholen, wodurch der Aufwand des

kausalen Modells in Frage gestellt wird. Trotzdem können die Ergebnisse eines kausalen Modells nie erreicht werden.

Die Anforderungen an ein Expertensystemshell zur Lösung von Diagnoseproblemen wurden in Med2 gut erfüllt.

Das Ohrendiagnosesystem ist ein assoziatives System mit Abstraktionsmechanismen. Die Regeln müssen in einer festen Reihenfolge ausgeführt werden. Strukturierte Regeln wie in Abel oder Med2 gibt es nicht. Für größere Anwendungsgebiete als die Ohrendiagnose, muß das System noch erweitert werden.

Die Anforderungen, die aus der Problemstellung der Ohrendiagnose gestellt werden, werden im Ohrendiagnosesystem sehr gut erfüllt.

Kapitel 3: Hiqual und Weiterentwicklungen von Hiqual

1. Die Wissensrepräsentationssprache Hiqual.

1.1 Einleitung.

Hiqual ist ein Tiefenmodellierungssystem zur Darstellung und Analyse physikalisch-technischer Systeme. Es basiert auf einer objektorientierten Modellierung hochgradig unabhängiger Modelle. Das reale System kann auf verschiedenen Abstraktionsebenen unter expliziter Angabe der strukturellen Beziehungen zwischen angrenzenden Ebenen dargestellt werden. Die Modelle einer beliebigen Ebene können unabhängig von denen der tieferen Ebene analysiert werden.

Die Kommunikation zwischen verschiedenen Modellen wird als Fließen von Material, Kräften oder Information interpretiert. Die dadurch entstehenden Veränderungen werden durch stetige Variablen mit gleitenden Wertübergängen oder diskreten Variablen mit beliebig wechselnden Werten dargestellt. Die Semantik eines Systems von Modellen wird als eine Menge zeitlich und kausal zusammenhängender Zeitintervalle beschrieben, die den dynamischen Zuständen und Ereignissen der Modelle entsprechen. Durch den zeitlichen Ansatz können Parallelismus und andere zeitliche Aspekte wie zeitliche Unsicherheiten natürlicher dargestellt werden.

1.2 Darstellung der von Hiqual benutzten Methoden zur Tiefenmodellierung.

Hiqual ist eine objektorientierte (komponentenorientierte)¹ Repräsentationssprache.

Die einfachen Objekte heißen *Modelle*; sie werden durch *Modelltypen* instantiiert. Ein Modelltyp ist die modulare Darstellung eines spezifischen Komponententyps des betrachteten Gerätes auf einer bestimmten Abstraktionsstufe.

Für jedes Gerät wird ein *Modellbaum* erstellt. Die Wurzel des Baumes repräsentiert die abstrakte und unabhängige Sicht des Gerätes. Durch jede weitere Stufe wird das Gerät

¹ Objekt = Komponente

Eine Komponente des realen Modells entspricht einem Objekt in der Repräsentationssprache. Deshalb werden beide Begriffe synonym verwendet.

detaillierter beschreiben, wobei eine Ebene des Baumes jeweils der Beschreibung des ganzen Gerätes entspricht.

Ein neuer Baum wird erstellt, indem wiederholt schon existierende Bäume integriert werden. Bäume der Tiefe eins heißen *einfache Aggregationen* und werden durch die *Aggregationstypen* instantiiert. Neben der Definition horizontaler Verbindungen zwischen den Modellen, gibt es auch vertikale Beziehungen zwischen Daten und Aktionen des Wurzelmodells und denen der Blättermodelle. Aggregationen können hinsichtlich der horizontalen und vertikalen Zusammensetzung als *atomare Einheiten angesehen werden*. Da diese Relationen zu den Aggregationen gehören und nicht zur gesamten Abstraktionsstufe, ist es möglich, bestimmte Komponenten auf einer tieferen Abstraktionsstufe zu betrachten als andere. Dadurch gibt es in Wirklichkeit mehr Detaillierungsstufen von einem Gerät als nur die Zahl der Abstraktionsstufen (siehe Abb. 3.1).

Bäume mit einer Tiefe größer eins heißen *komplexe Aggregationen*. Diese Bäume können von unten nach oben (bottom-up), von oben nach unten (top-down) und auf beide Arten gemischt aufgebaut werden. Daher kann jeder Modellbaum für ein reales System in Hiqual durch eine der oben genannten Methoden erstellt werden.

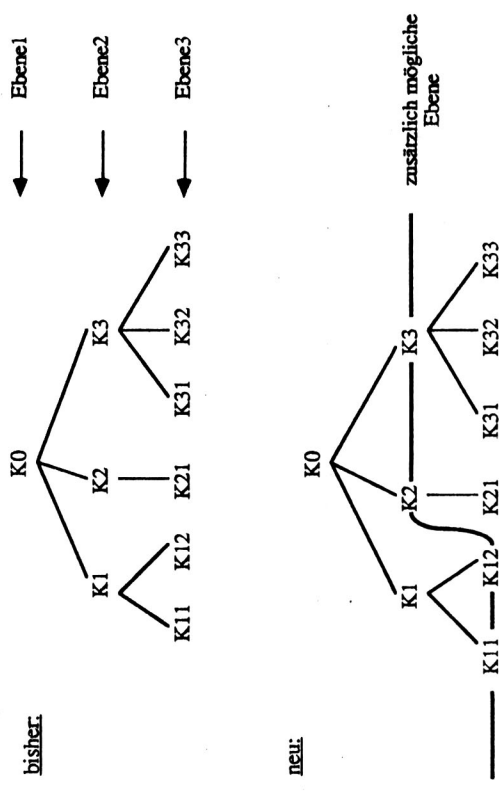


Abb. 3.1: Schematische Darstellung möglicher Abstraktionsebenen.

1.2.1 Modelltypen.

Modelltypen instanzieren individuelle Objekte, und zwar die Modelle. Außer dem Namen des Modelltyps werden noch folgende Informationen eingegeben:

- *Typen*, die innerhalb des Modells vorkommen können.

Typen definieren endliche Gebiete als Anwendungsbereiche für Ports, Variablen oder Attribute. Unendliche Typen wie 'integer' oder 'real' gibt es nicht (Begründung: Für abstrakte Modelle braucht man nur abstrakte Werte).

Typen können entweder lokal innerhalb eines Modelltyps definiert werden oder global auf der höchsten Ebene der Hiquil-Sitzung. Diese globalen Typen können durch *Import-Klassifizieren* innerhalb der Modelldefinition benutzt werden.

-- *einfache Typen*.

In Hiquil unterscheidet man zwischen geordneten Anwendungsbereichen für kontinuierliche Variablen (oder Ports) und ungeordneten Anwendungsbereichen für digitale Variablen. Ein Wert eines Anwendungsbereiches kann auch als 'instantaneous' definiert werden; d.h. dieser Wert kann nur für einen 'Moment' angenommen werden (siehe Kap.3 Abschnitt 1.3.4.4, Momentane Ereignisse).

-- *Record-Typen*.

Die Komponenten des Records können unterschiedliche Typen haben, aber es müssen einfache Typen sein.

- *Ports*.

Ports sind lokale Namen zur Verbindung zweier Modelle (Schnittstellen). Sie werden ähnlich wie Variablen behandelt. *In-Ports* können nur im Bedingungssteil einer Regel auftreten; *Out-Ports* können nur in den Aktionssteilen der Regel auftreten, um neue Werte für die Ausgabe des Modells zu spezifizieren.

- *Variablen*.

Variablen werden in Hiquil oft durch Facetten abgelöst. Der Facettenmechanismus kann ganz durch Variablen simuliert werden; aber nicht umgekehrt. Facetten werden zur ersten Strukturierung der verschiedenen Zustände eines Mechanismus benutzt. Sie können bei der zeitlichen Analyse speziell behandelt und integriert werden. Die Variablen benutzt man nur zu gewünschten Verfeinerungen der Zustände (Der Unterschied Facette - Variable entspricht ungefähr dem Unterschied Quadrat - Viereck).
- *Attribute*.

Attribute spezifizieren Eigenschaften eines Modells, die während der ganzen Lebenszeit des Modells konstant bleiben. Werte für Attribute bzw. Kombinationen von Attributen können durch Constraints beschränkt werden. Ebenso wie Variablen können auch Attribute initialisiert werden. (Nach der Initialisierung können die Attribute nur noch einmal verändert werden.)

- *Funktionen*.

Eine Funktion mit n Argumenten wird folgendermaßen spezifiziert:

- (1) Definitionsbereich (= Liste der Typen der n Argumente)
- (2) Wertebereich (= Typ für das Resultat der Funktion)
- (3) Graph (= Paare: Argumenttupel-Ergebniswert)

Diese Notation ist möglich, weil alle Definitionsbereiche der n Argumente endlich sein müssen, und somit auch das Kreuzprodukt aus den Definitionsbereichen endlich ist. Eine Funktion kann aber auch importiert werden.

- *Facetten und Regeln*.

Jede Facette beschreibt einen Zustand eines Modells. Sie hat einen Namen und enthält eine Menge von Verhaltensregeln. Jede Regel hat einen Namen, einen Bedingungssteil und einen Aktionssteil. Sowohl der Bedingungssteil als auch der Aktionssteil der Regel bestehen aus Ungleichungen und aus Übergangsaktionen zu anderen Facetten.

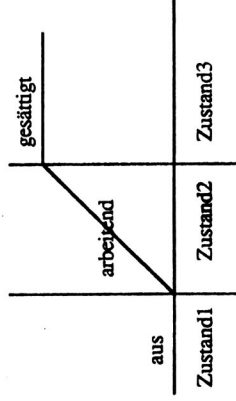


Abb. 3.2: Beispiel zur Verwendung von Facetten:

Ein Transistor reagiert in den jeweiligen Zuständen unterschiedlich. Jedem Zustand wird eine Facette zugeordnet, die die jeweiligen Verhaltensregeln beinhaltet.

1.2.2 Aggregationstypen.

Ein Aggregationstyp ist ein Schema zur Erzeugung von Aggregationen. Er legt die Wurzel -ein Modell des ganzen Gerätes- und ihre Söhne -Modelle für die Komponenten des Gerätes- fest. Aus mehreren Aggregationen können komplexe Aggregationen gebildet werden.

Eine Aggregationstypdefinition enthält folgende Informationen:

- Name des Aggregationstyps
- Formale Modellnamen aller Modelltypen, die zum Aggregationstyp gehören
- Name des Wurzelmodelltyps

- **Attributbeschränkungen**

Sie treten schon bei einer Modelltypdefinition auf. Bei der Aggregationstypdefinition können aber Attribute aller dazugehörigen Modelltypen betrachtet werden, nicht nur die eines einzelnen Modelltyps.

- **horizontale Verbindungen (connections)**

In- und Out-Ports der Blättermodelle können miteinander verbunden werden. Diese Verbindung muß explizit angegeben werden. Dabei muß folgendes beachtet werden:

- Die Imports (Outports) dürfen nur mit Outports (Imports) verbunden werden.
- Verbundene Ports müssen von demselben Typ sein.
- Jeder Port kann mit maximal einem weiteren Port verbunden sein.

- **Umformungsfunktion: (representations)**

Diese Funktionen dienen dazu, die Datentyprepräsentation von der abstrakten Stufe (Wurzel) in die der konkreten Stufe (Blätter) umzuformen. Sie werden durch Anwendungsbereich, Wertebereich und Graph definiert (Im Gegensatz zu den Funktionen ist der Definitionsbereich einstellig und der Wertebereich n-stellig).

Beispiel: Umformung dezimaler Zahlen in Binärzahlen.

- **Vertikale Verbindungen (structure map)**

Durch diese Verbindungen kann man Ports der abstrakten Stufe mit den Ports der konkreten Stufe verknüpfen. Umformungsfunktionen können nur die Zahl und die Typen der Ports festlegen. Die vertikalen Verbindungen geben an, mit welchen Ports diese Umformungsfunktion verknüpft ist (siehe Abb. 3.3).

- **Zeitliche Relationen**

Es sollen zeitliche Relationen zwischen Ereignissen oder Zuständen spezifiziert werden (siehe Kap. 3 Abschnitt 1.3.4. Zeitliche Relationen zwischen Intervallen).

Will man ein und dasselbe Modell auf zwei benachbarten Stufen benutzen, so muß man es auf der unteren Stufe als 'inherited' deklarieren, um so eine Kopie des Modells zu erzeugen (Identische Verfeinerungen müssen also explizit aufgeführt werden).

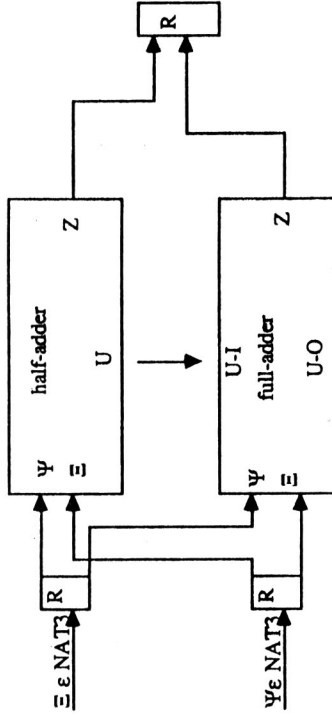


Abb. 3.3: Schematische Darstellung eines Zwei-Bit-Addierers.

Er besteht aus einem Halbaddierer und einem Volladdierer. Die mit 'R' markierten Kästchen zeigen die explizite Anwendung einer representation-Funktion an.

```

types BIT is {0,1};
function ADD-NAT3 is ...;
type ADDER-TYPE is model
types NAT3 is {0,1,2,3};
NAT7 is {0,1,2,3,4,5,6,7};
ports in Ξ, Ψ are NAT3;
      S-I is BIT;
      out Z is NAT7;
      S-O is BIT;
functions ADD-NAT3 is imported;
facet ADDING is
  ADD-RULE is
    preconds S-I = 1;
    actions  Z = ADD-NAT3(X,Y);
            S-O = 1;
INITIALLY ADDING;
end ADDER-TYPE;

```

Abb. 3.4: Hiqual-Definition eines Modell-Typs Addierer.

S-I und S-O werden zur Simulation der Synchronisationsuhr verwendet.

1.2.3 Instanzen.

-Modellinstanzen.

Modelle werden durch die Modelltypen instanziiert, indem folgendes festgelegt wird:

Name der Instanz

Typ der Instanz

Attribute mit konstanten Werten

Jedem Attribut, das keinen Default-Wert besitzt, muß explizit ein Wert zugewiesen werden. Eine Instanz kann nur gebildet werden, wenn alle Attributrestriktionen der Modelltypdefinition erfüllt sind und alle Attribute mit Werten belegt sind.

-Einfache Aggregationen.

Einfache Aggregationen werden durch Aggregationstypen instanziiert, indem folgendes festgelegt wird:

Name der einfachen Aggregation.

Typ der einfachen Aggregation

Für jeden Modelltyp des Aggregationstyps die entsprechende Modellinstanz.

Alle Attributrestriktionen in der with-Klausel des Aggregationstyps müssen durch die Modellinstanz erfüllt werden.

-Komplexe Aggregationen

Komplexe Aggregationen werden durch folgende Spezifikationen gebildet:

- Name für die komplexe Aggregation

- Liste mit Namen schon existierender einfacher und komplexer Aggregationen

Aus allen Aggregationen versucht Hiqual einen einzigen Modellbaum zu konstruieren.

Alle Aggregationen, egal ob einfach oder komplex, sind durch ihre vertikalen und horizontalen Beziehungen charakterisiert.

- Vertikale Relationen.

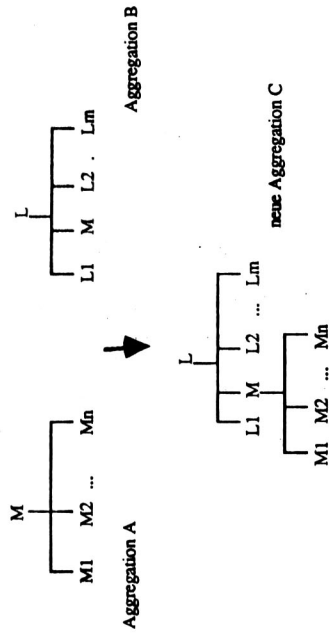


Abb. 3.5: Erstellen vertikaler Relationen

Die Aggregationen A und B besitzen beide die Komponente M. In der Aggregation A bildet sie die Wurzel, in der Aggregation B ist sie ein Blatt. Fügt man die Blätter der Aggregation A an das Blatt M in der Aggregation B hinzu, so erhält man eine neue Aggregation C.

- Horizontale Relationen.

Die Kombination einfacher Aggregationen zu einer komplexen Aggregation erfordert neue Verbindungen zwischen Modellen der verschiedenen Aggregationen. (Solch eine Verbindung (connection) führt von einem freien Port der einen Aggregation zu einem freien Port der anderen Aggregation.) Gibt es eine Verbindung zwischen zwei Modellen A und B, so sollte es auch eine Verbindung zwischen den Komponenten von A und B geben. Diese Verbindungen können nicht explizit spezifiziert werden, weil sie zu zwei verschiedenen Aggregationen gehören; deshalb definieren wir sie als implizite Verbindungen.

Durch die Information, die durch die Aggregationstypdefinition geliefert wird (structure map), ist die Menge der impliziten Verbindungen eindeutig und kann eindeutig gefunden werden.

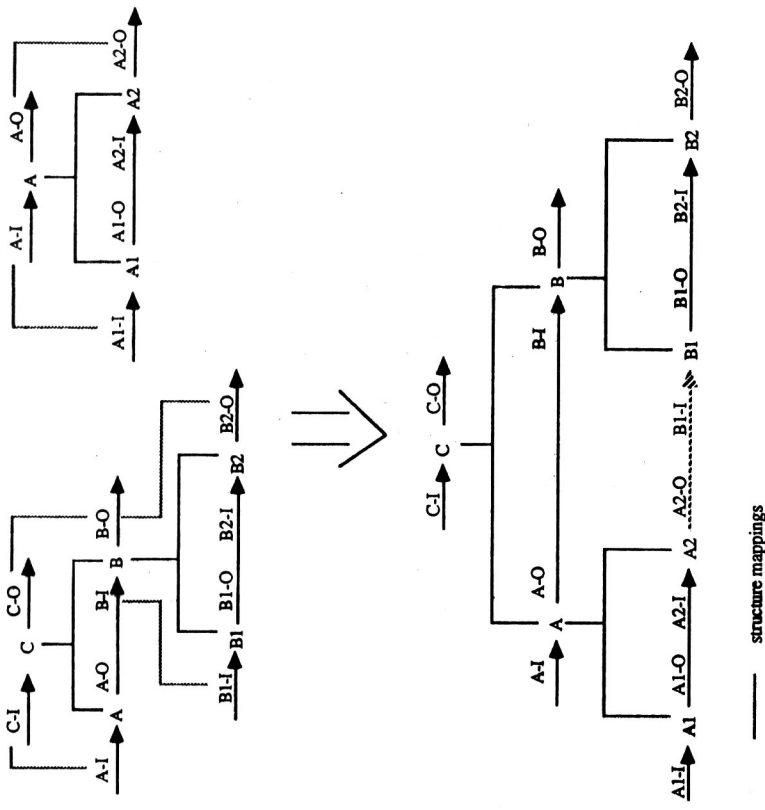


Abb. 3.5 Zusammensetzung zweier Aggregationen mit gemeinsamem Knoten A. Diese Abbildung beschreibt die Suche nach impliziten Verbindungen in einem einfachen Fall. Alle Ports der abstrakten Stufe sind mit genau einem Port der konkreteren Stufe verbunden. Man betrachtet nun jeden freien Port der einfachen Aggregation A der Reihe nach.
 A-I: bleibt frei, weil er in der komplexen Aggregation C auch frei war
 A-O: wird mit B-I verbunden, wie in der Aggregation C
 A1-I: bleibt frei, weil A-I mit A1-I verbunden ist und A-I frei bleibt
 A2-O: wird mit B1-I verbunden, weil A-O und A2-O durch eine structure-map Relation verbunden sind. A-O ist mit B-I verbunden. B-I und B1-I sind durch eine structure-map Relation verbunden. Es gibt also einen Pfad von A2-O nach B1-I über A-O und B-I. In diesem Fall erzeugt Hiqual eine implizite Verbindung zwischen A2-O und B1-I.

1.3 Analyse in Hiqual.

Ein Grund für die breite Akzeptanz der objektorientierten Beschreibungen ist die Individualität der Komponenten. Viele Tiefenmodellierungssysteme halten an der Verhaltensbeschreibung durch globale Zustände fest. Dadurch wird aber den Komponenten die Individualität entzogen, die sie durch die strukturelle Analyse gerade gewonnen haben.

Hiqual arbeitet mit Zeitintervallen und besitzt damit folgende Vorteile:

- Die Individualität der Komponenten bleibt erhalten.
- Das Verhalten des Gerätes kann auf sehr natürliche Weise beschrieben werden.
- Es können sowohl Zeitpunkte als auch Zeitintervalle dargestellt werden.
- Als zeitliche Intervalle können Bedingungen, Aktionen oder Zustände dargestellt werden.
- Die Darstellung von Unsicherheit ist möglich (statt einer Relation kann man eine ganze Relationenmenge angeben).

1.3.1 Allgemeiner Überblick.

Analyse von Modell- und Aggregationsinstanzen bedeutet:

- Zeitliche Intervalle werden mit Zuständen und Ereignissen verknüpft. Dabei entsprechen die Zustände den Instantierungen von Facetten und die Ereignisse sind Vorbedingungen oder Aktionen von Regeln, die während der Ausführungen auftreten.
- Zwischen diesen Zuständen und Ereignissen werden mögliche zeitliche Relationen bestimmt.
- Ordnungsrelationen (ontologische Relationen) werden zwischen bestimmten Zuständen und Ereignissen spezifiziert. Ontologische Relationen werden benutzt, um syntaktische Beziehungen, wie sie in der Modelldefinition zwischen Zuständen und Ereignissen spezifiziert sind, bzgl. ihrer Wirkung und Funktion darzustellen. Dadurch kann man ausdrücken, daß ein Ereignis einen neuen Zustand definiert oder ein anderes Ereignis verursacht.

Zeitliche Relationen können in den Modelldefinitionen explizit spezifiziert werden. Andernfalls übernimmt das System vordefinierte Relationen, die hinsichtlich der zeitlichen Information die maximal mögliche Unsicherheit annehmen. Benutzer können zeitliche Relationen nur innerhalb individueller Regeln angeben; und dort auch nur für die

Bedingungen und Handlungen, die in dieser Regel auftreten. Alle Relationen zwischen anderen Zuständen oder Ereignissen desselben Modells oder zwischen Zuständen und Ereignissen verschiedener Modelle werden automatisch durch das System gefolgt. Diese Folgerungen nicht explizit spezifizierter Relationen werden im wesentlichen mit der Constraint-Propagation-Methode von Allen [Allen 83] berechnet.

1.3.2 Zeitintervallanalyse von Allen.

Die Zeitanalyseprozeduren von Hiquil basieren auf den Arbeiten von Allen [Allen 83/84] über die Darstellung von zeitlichem Wissen und auf seinem Algorithmus zur Erstellung konsistenter Netze zeitlicher Relationen. Allen definiert einen Inversions- und einen Transitivitätsoperator auf der Potenzmenge der zeitlichen Relationen. Benutzt man die allgemeinen Eigenschaften dieser Operatoren, so kann man zeigen, daß der Algorithmus von Allen gut in den Rahmen der Constraint-Propagation paßt.

Bei Allen gibt es nur nicht-triviale Zeitintervalle d.h. keine Zeitpunkte. Zwischen zwei Intervallen gibt es dreizehn mögliche zeitliche Relationstypen. Normalerweise liegt zwischen zwei Intervallen jeweils ein Relationstyp vor; bei Unsicherheiten ist auch eine Menge von Relationen möglich.

Aufzählung aller möglichen Zeitrelationen zwischen zwei Intervallen:

- A < B : before : A tritt irgendwann vor B auf
- A > B : after : A tritt irgendwann nach B auf
- A d B : during : A tritt während des Zeitintervalls B auf
- A di B : contains : A enthält das Zeitintervall B
- A o B : overlaps : A und B haben ein gemeinsames, nicht leeres Zeitintervall
- A oi B : overlapped-by : A und B haben ein gemeinsames, nicht leeres Zeitintervall
- A m B : meets : A ist direkter Nachfolger von B
- A mi B : met-by : A hat als direkten Nachfolger B
- A s B : starts : A startet das Zeitintervall B
- A si B : started-by : A wird im Zeitintervall B gestartet
- A f B : finishes : Im Zeitintervall A wird B beendet
- A fi B : finished-by : A wird im Zeitintervall B beendet
- A = B : equal : A ist gleich B

Beispiele:

1. Inversionsoperator.

Seien A, B, C und D Zeitintervalle

C -- (s, f, mi) --> D <==> D -- (si, fi, m) --> C weil (s, f, mi) = (si, fi, m)

2. Transitivitätsoperator.

A -- (=, s, si, d, oi, f) --> B und B -- > --> D

Wie lautet jetzt die Relationenmenge zwischen A und D?

A -- > --> D (Berechnung nach der Transitivitätstabelle, [Allen83])

1.3.3 Definitionen mit Erklärungen anhand eines Beispiels.

Für jede Modellinstanz wird eine Regelhistory (RH) definiert. Dies ist eine Folge von Regelnamen, die für den Modelltyp definiert ist.

RH := (r₁, ..., r_N)

RH-Coil = (ON-R1, OFF-R1, ON-R2, OFF-R2)

Jede Regel ist von einer der folgenden Formen:

- a) R_i is preconds C(i,1),...,C(i,c)
actions A(i,1),...,A(i,a)
relations .. heißt non-enter-Regel
- b) R_i is preconds C(i,1),...,C(i,c)
actions enter F, A(i,1),...,A(i,a)
relations .. heißt enter-Regel

Sei C = Die Menge aller Bedingungen (conditions), die in RH auftreten

A = Die Menge aller Aktionen, die in RH auftreten

F = Die Menge aller Facetten, die in RH auftreten

CAF = C ∪ A ∪ F

C-Coil = { CUR-I=0, CUR-I>0 }

A-Coil = { TENSION-O=LOW, TENSION-O=HIGH }

F-Coil = { FACET=ON, FACET=OFF }

INTVN = Die Menge aller Intervallnamen

Die Namen aller Zeitintervalle, die in RH auftreten, kann man als Quadrupel charakterisieren:

$N1 = (caf / Name\ der\ Modellinstanz / type / index)$

mit $-caf \in CAF$; also ist caf entweder eine Aktion, Facette oder Bedingung

$-type \in \{in, ni\}$

in = momentan (instantaneous); liegt vor, wenn sich seine Spezifikation auf einen

Wert bezieht, der als 'is inst' definiert wurde.

ni = nicht momentan (non-instantaneous); sonst

-index gibt an, das wievielte Mal caf in RH auftritt

Ist $caf \in CA$, so ist das Intervall ein Ereignis. Ist $caf \in F$, so ist das Intervall ein Zustand.

Der Zustand einer Enter-Regel entspricht immer dem neuen Zustand der Regel.

Beispiel:

$INTVN = \{ CUR-I=0/COIL/ni/1, CUR-I=0/COIL/ni/2,$

$CUR-I>0/COIL/ni/1, CUR-I>0/COIL/ni/2,$

$TENSION-O=LOW/COIL/ni/1, TENSION-O=LOW/COIL/ni/2,$

$TENSION-O=HIGH/COIL/ni/1, TENSION-O=HIGH/COIL/ni/2 \}$

$\{ FACET=ON/COIL/ni/1, FACET=ON/COIL/ni/2, FACET=ON/COIL/ni/3 \}$

$\{ FACET=OFF/COIL/ni/1, FACET=OFF/COIL/ni/2 \}$

Der Startzustand ist 'ON'; somit tritt dieser Zustand einmal öfter auf.



1.3.4 Zeitliche Relationen zwischen Intervallen und die entsprechenden ontologischen Relationen.

Man unterscheidet zwischen drei verschiedenen Relationen:

- Zustand - Zustand - Relation
- Ereignis - Zustand - Relation
- Ereignis - Ereignis - Relation

1.3.4.1 Zustand-Zustand-Relation

Zu jedem Zeitpunkt muß jedes Modell in genau einer Facette sein; d.h. aufeinanderfolgende Zustände einer Modellinstanz bilden eine sich nicht überlappende und lückenfreie Folge zeitlicher Intervalle.

richtig:  falsch: 

oder



Ein neuer Zustand F2 wird als Nachfolger des alten Zustandes F1 durch die Relation

$F1 - m \rightarrow F2$

etabliert. Diese Relation kann allerdings nicht explizit definiert werden.

1.3.4.2 Ereignis-Zustand-Relation.

Hier gibt es wiederum drei verschiedene Arten von Relationen:

A: Ereignisse definieren und verlassen Zustände

Sei r eine Enter-Regel und $C/C(r)$ eine Bedingung, so heißt die Relation zwischen

$C/C(r)$ und dem neuen Zustand $F2/F2(r)$:

$F2$ wird definiert durch das Ereignis C oder C definiert $F2$ (s. Abb. 3.7)

\implies zeitlich gesehen darf $F2$ nicht vor C starten.

$F2/m \text{ -- id / not(starts-before)}^1 \text{ --> } C/n$ oder auch

$C/n \text{ -- d / not(starts-before)} \blacksquare \text{ --> } F2/m$

B: Ereignisse finden während eines bestimmten Zustandes statt.

Relationen zwischen Aktionen und Zuständen der Regel

F/m sei folgendermaßen definiert: Liegt eine Enter-Regel vor, so bezieht man sich auf

den neuen Zustand; andernfalls auf den alten Zustand. Eine Aktion A/n wird während des

Zustandes F/m bewirkt (inverse effected) (s. Abb. 3.7).

\implies zeitlich gesehen muß die Aktion A/n noch innerhalb des Zustandes F/m starten.

¹not(starts-before) = (=, s, si, >, ni, oi, f, d)

A/n -- ie / starts-in¹ --> F/m oder auch
 F/m -- e / starts-in² --> A/n

Betrachtet man die Extrempunkte in der starts-in-Relation, so muß man beachten, daß A/n gleichzeitig mit F/m starten kann, aber nicht nach F/m. Dadurch entsteht eine eindeutige Zuordnung zwischen dem Zustand und der Aktion.

C: Ereignisse werden in Zuständen registriert.

Wir betrachten Relationen zwischen Bedingungen von non-enter-Regeln und den sie umgebenden Facetten. Da der Zustand nicht geändert wird, darf die Bedingung immer starten, nur nicht am Ende oder nach dem Zustand. Dann wird das Ereignis in diesem Zustand 'registriert'. (s. Abb. 3.7)

C/n -- ir / not (>mi) --> F1 / F1(rj)
 F1 / F1(rj) -- r / not (<mi) --> C / n

1.3.4.3 Ereignis-Ereignis-Relation.

Zwischen Bedingungen können beliebige zeitliche Beziehungen bestehen. Ebenso auch zwischen Aktionen. Uns interessieren hier noch die Beziehungen zwischen Bedingungen und Aktionen innerhalb einer Regel (kausale Relation).

Das verursachte Ereignis darf nicht vor dem verursachenden Ereignis starten, sonst ist alles erlaubt.

A / m -- ic / not(starts-before)² --> C / n
 C / n -- c / not(starts-before) --> A / m

Alle bisher aufgeführten zeitlichen Relationen können durch explizite Angaben des Benutzers weiter verfeinert werden. Entstehen dadurch Widersprüche, so werden sie vom System entdeckt und die Eingaben werden zurückgewiesen.

¹ starts-in = (=, s, si, f, oi, d)
² not(starts-before) = (=, s, si, >, mi, oi, f, d)

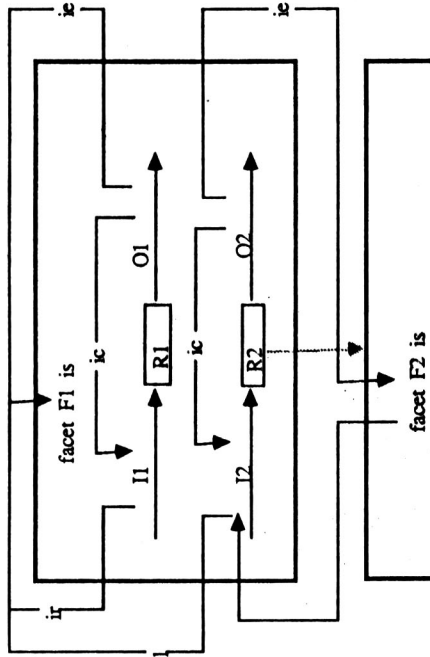


Abb. 3.7: Schematischer Überblick aller möglichen ontologischen Relationen.

Relation zwischen Bedingungen und Aktionen	Abkürzung	ontologische Rel.	Zeitliche Relation
Bedingungen und Aktionen	ic	is caused by	not(starts-before)
Bedingungen und Zuständen	c	causes	not(starts-before) ■
Aktionen und Zuständen	id	is defined by	not(starts-before)
Bedingungen und Zuständen in non-enter-Regeln	d	defines	not(starts-before) ■
Bedingungen und Zuständen in non-enter-Regeln	ic	is effected in	starts-in
Bedingungen und Zuständen in non-enter-Regeln	e	effects	starts-in ■
Bedingungen und Zuständen in non-enter-Regeln	ir	is registered in	not(>, m)
Bedingungen und Zuständen in non-enter-Regeln	r	registers	not(>, m) ■ = not(<, mi)
Bedingungen und dem alten Zustand	il	is left by	not(>)
Bedingungen und dem alten Zustand	l	leaves	not(>) ■ = not(<)
Bedingungen und dem alten Zustand	hc	horizontally connected	=
Bedingungen und dem alten Zustand	ch	inverse of hc	=
Bedingungen und dem alten Zustand	vc	vertically connected	=
Bedingungen und dem alten Zustand	cv	inverse of vc	=

Abb. 3.8: Ontologische Relationen, ihre inversen Relationen und die Default-Werte der entsprechenden zeitlichen Relationen.

1.3.4.4 Momentane Ereignisse und Zustände.

Ein Ereignis heißt momentan, wenn es sich auf einen Port oder eine Variable bezieht, die als 'is inst' in der Typdeklaration spezifiziert wurde. Zustände heißen momentan, wenn die entsprechende Facettenspezifikation lautet:

facet <name> is inst ...

Momentane Facetten kann man nur zur Modellierung von Zuständen ohne Dauer benutzen. Idee: Wie erhält man ein momentanes Ereignis?

Man versucht die Größe eines Intervalles für ein momentanes Ereignis so zu verkleinern, daß es nicht größer ist, als irgendein nicht-momentanes Ereignis und entweder disjunkt oder in einer meet-Relation mit allen anderen momentanen Ereignissen ist. Dazu ersetzt man die alte Zeitrelation R zwischen einem momentanen Ereignis und einem beliebigen Ereignis durch die Relation R' = instant(R,e)

instant (R,e) := $R \setminus \{o, =, fi, si, oi, di\}$ if e = nicht-momentanes Ereignis
 $R \setminus \{o, s, d, f, fi, si, oi, di\}$ if e = momentanes Ereignis

In Worten: Ist e ein momentanes Ereignis, so ist instant(R,e) \subseteq {m, mi, >, <, =} d.h. Zwei momentane Ereignisse sind entweder gleich oder (direkt) nacheinander. Ist e ein nicht-momentanes Ereignis, so darf das nicht-momentane Ereignis nicht gleich dem momentanen Ereignis sein, und es darf keine Möglichkeit geben, das momentane Ereignis in einen gemeinsamen und einen nicht gemeinsamen Teil mit dem nicht-momentanen Ereignis aufzuteilen.

graphische Darstellung	mnemonic	instant(R,e) e ∈ NIE	instant(R,e) e ∈ IE
XXX YYY	<	<	<
XXXYYY	m	m	m
XXX	o	()	()
YYY	s	s	()
XX	d	d	()
YYY	f	f	()
XXX	=	()	=
YYY			

graphische Darstellung	mnemonic	instant(R,e) e ∈ NIE	instant(R,e) e ∈ IE
XXX	fi	()	()
YY	si	()	()
XXX	di	()	()
YYY	oi	()	()
YYYXXX	mi	mi	mi
YYY XXX	>	>	>

Abb. 3.9: Definition von instant(R,e) YYY : e

XXX : momentanes Ereignis : ie

1.3.5 Relationen, die durch horizontale Verbindungen entstehen.

Wie stehen Ereignisse oder Zustände verschiedener Modelle zueinander in Beziehung? Um dieses Problem zu lösen, muß man folgendes tun:

- (1) Alle Intervalle und Relationen der beteiligten Modelle betrachten.
- (2) Analyse der existierenden expliziten und impliziten horizontalen Verbindungen. Versuch die 'kommunizierenden' Aktionen und Bedingungen zu identifizieren.

Dann können wir folgende Relation generieren:

A / Modellinstanz1 - hc / = -> C / Modellinstanz2

A und C passen nur zusammen, wenn ihre linken Seiten¹ in der Modelldefinition

verbundene Ports desselben Typs sind. Zwei Typen heißen gleich, wenn sie dieselben Werte und bei geordneten Typen auch dieselbe Ordnung definieren.

Die oben genannte Relation gibt an, daß A und C zeitlich gleich sind, d.h., daß die Aktion des einen Modells und die entsprechende Bedingung des verbundenen Modells dasselbe Ereignis darstellen muß. Jedes Ereignis hat seine syntaktische Herkunft als Code in der Modelldefinition. Trotzdem kann man sich auf 'dasselbe' Ereignis aus verschiedenen Gesichtspunkten beziehen: z.B. vom Sender und vom Empfänger aus. Die Ausgabe des einen Modells wird zur Eingabe des anderen Modells.

¹A und C sind meist Ungleichungen ==> Der Term auf der linken Seite der Ungleichung ist gemeint

Beispiel:

A = CUR-O = 0 (Aktion)

C = CUR-I = 0 (Bedingung)

CUR-O und CUR-I sind von demselben Typ und durch die connection F-COIL#CUR-I with F-CLAPPER#CUR-O verbunden.

1.3.6 Relationen, die durch vertikale Verbindungen entstehen.

Die zeitliche Interpretation der vertikalen Verbindungen ist ähnlich der der horizontalen Verbindungen.

- (1) Identifikation aller vertikalen Verbindungen, die in der structure-map beschrieben sind. Mit ihrer Hilfe werden die entsprechenden Ereignisse auf benachbarten Ebenen identifiziert.
- (2) Diese sich entsprechenden Ereignisse müssen mit ontologisch/zeitlichen Relationen verknüpft werden (Für weitere Einzelheiten siehe [Voss 1986]).

1.4 Pro und Contra Hiqual.

Hiqual ist eine objektorientierte Repräsentationssprache, die zur zeitlichen und kausalen Analyse physikalisch-technischer Systeme geeignet ist.

Vorteile von Hiqual:

Es gibt zwei Detaillierungsmöglichkeiten in Hiqual:

Einerseits kann man untersuchen, aus welchen Unterkomponenten und Verbindungen eine Komponente besteht, zum anderen kann man auf der Abstraktionsebene bleiben und die Komponente selbst genauer betrachten: Welche Zustände kann sie annehmen, welche Relationen gibt es innerhalb der Komponente bzw. nach außen oder welche Ereignisse können auftreten.

Eine *einfache Entwicklung* wird durch die objektorientierte Modellierung mit hochmodularen Modellen möglich. Komplexe Objekte werden aus einfachen und komplexen Aggregationen erstellt. Fast alle strukturellen Relationen müssen nur im lokalen Kontext einfacher Aggregationen definiert werden, wobei man nur eine einzelne Komponente und ihre Teile auf der nächst tieferen Abstraktionsebene betrachtet. Komplexere Modellhierarchien definiert man, indem man die Namen aller einfachen und komplexen Aggregationen angibt, die dazugehören sollen.

Die *Kommunikation* zwischen verschiedenen Modellen wird als Fließen von Material, Kräften und Information interpretiert. Für den kontinuierlichen Fluß werden stetige Variablen mit gleitenden Wertübergängen benutzt; für den digitalen Fluß diskrete Variablen mit beliebig wechselnden Werten.

Hiqual liefert sehr flexible Möglichkeiten zur Konstruktion von Hierarchien: Das Modell kann von unten nach oben, von oben nach unten oder in gemischter Form erstellt werden. Außerdem erlaubt es Hiqual, ausgewählte Komponenten genauer zu betrachten, indem sie (und nur sie) auf einer tieferen Abstraktionsebene betrachtet werden. Durch die vertikalen Relationen zwischen Darstellungen auf benachbarten Abstraktionsstufen, kann Hiqual dasselbe Verhalten auf verschiedenen Abstraktionsstufen vergleichen. Damit kann die Konsistenz auf verschiedenen Abstraktionsstufen überprüft werden. Das System zeigt an, wenn das Verhalten auf einer abstrakteren Stufe nicht mit dem Verhalten auf der konkreteren Stufe übereinstimmt.

Ein Grund für die breite Akzeptanz der objektorientierten Beschreibung ist die Individualität der Komponenten. Betrachtet man die Aktivitäten der Komponenten, so erkennt man bei technischen Geräten oft *Zeit-Parallelität*. Alle anderen Tiefenmodellierungssysteme außer Hiqua halten an der Verhaltensbeschreibung durch globale Zustände fest. Dadurch wird den Komponenten die Individualität entzogen, die sie durch die strukturelle Analyse gerade erst gewonnen haben. In Hiqua werden alle Ereignisse und Zustände als Zeitintervalle interpretiert, wodurch parallele Aktionen geeignet dargestellt werden können. Hiqua ist auch eines der wenigen Tiefenmodellierungssysteme, in dem die Zeit so einfach dargestellt werden kann. Auch momentane Ereignisse und Zustände werden als spezielle Zeitintervalle behandelt. Die zeitliche Analyse von Hiqua ist auf einer Propagierungsprozedur begründet, und zwar einer konstruktiven constraint-propagation-Prozedur, bei der jede Variable eine Menge von Werten besitzt.

Nachteile von Hiqua

entstehen durch die Tatsache, daß es nur ein Analyse- und kein Simulationssystem ist. Deshalb kann keine 'defekte' Komponente in das System eingebaut und deren Folgen betrachtet werden. Ebenso ist es nicht möglich, quantitative, zeitliche Intervalle anzugeben.

1.5 Was bietet Hiqua zur Unterstützung der Diagnostik-Expertensysteme?

Hiqua kann zur Analyse physikalisch-technischer Systeme verwendet werden. Die Grundstruktur eines Hiqua-Modells ist ein Baum, dessen Knoten Teile des Gerätes darstellen. Die Wurzel, wie auch jede Ebene des Baumes, repräsentieren das ganze Gerät auf jeweils verschiedenen Abstraktionsebenen. Zusätzlich zu der Part-of-Relation gibt es noch weitere Relationen zeitlicher und ontologischer Art, die betrachtet werden können.

Ein Modelltyp beschreibt eine Modellinstanz durch folgende Angaben:

- Typen, die im Modell auftreten
- In- und Out-Ports (Schnittstellen zu den anderen Modellen)
- Variablen
- Attribute mit Restriktionen
- Funktionen (als Berechnungsmöglichkeit für das Modell)
- Facetten (Zustände des Modells mit den entsprechenden Regeln)

Hiqua erleichtert die Erstellung des Modells durch den beliebigen komponentenweisen Aufbau. Durch die Möglichkeit das korrekte Modell 'ablaufen zu lassen' kann man viele Beziehungen zeitlicher oder ontologischer Art abfragen. Liegt vermutlich ein seltener

Defekt in einem Motorteil vor, so hat man die Möglichkeit anhand des korrekten Modells dieses Teil auf tieferen Ebenen genauer zu betrachten und sich die genau Funktionsweise der einzelnen Teile mit ihren Relationen sich wieder ins Gedächtnis zu rufen.

Die Überprüfung der Eingabedaten des Expertensystems wird durch das Hintergrundwissen, das durch das Hiqua-Modell erweitert wurde, verbessert. Dabei muß man bei der Eingabeschnittstelle folgende Unterscheidung treffen:

- (1) Der Benutzer kann nur auf Fragen antworten (näheres siehe Med2).
- (2) Der Benutzer kann Symptome in einem bestimmten Format eingeben. Dabei sind mehr Fehlermöglichkeiten gegeben, die abgefangen werden müssen.

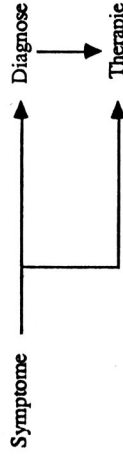
Ein Beispiel

Eingabe: Die Zündkerzen der Lampe sind verschmutzt.

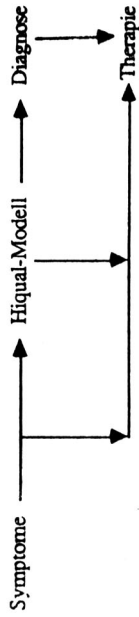
Ausgabe: Die Lampe hat als Komponente keine Zündkerzen. Sie besteht aus den Teilen K1, ..., Kn. Zündkerzen gibt es in den Komponenten F1, ..., Fr.

Alle Relationen, die in Hiqua aufgezeigt werden, sind Relationen zwischen Zuständen, Ereignissen oder Komponenten des Gerätes. Die Eingaben eines Expertensystems sind aber Symptome und die Ausgaben sind vor allem Diagnosen. Wie nützlich Hiqua für das Expertensystem ist, hängt davon ab, inwieweit die Symptome und Diagnosen des Expertensystems den Komponenten des Gerätes zugeordnet werden können. Das gilt auch für die Erweiterungen von Hiqua wie Hiqua1 und Hiqua2. Gelingt die Zuordnung nicht, oder nur unzureichend, so ist das Expertensystem ohne Hiqua genauso gut oder sogar besser.

Benutzt man nur ein Expertensystem, so versucht man aus den Symptomen eine Diagnose abzuleiten und eine Therapie aufzustellen:



Integriert man Hiqua in das Expertensystem, so steht zwischen den Symptomen und der Diagnose das Modell des Gerätes.



Wie effizient die Integration des Hiqual-Modells ist, hängt davon ab, wie gut die Zuordnung zwischen Symptomen und Komponenten des Hiqual-Modells ist; auch die Beziehung Therapie - Komponenten des Hiqual-Modells ist von Bedeutung. Gibt es zu viele Symptome wie z.B. Fieber, die sich keiner Komponente zuordnen lassen, so ist der Nutzen des Hiqual-Modells in Frage gestellt. Dasselbe gilt für die Diagnose und in geringerem Masse auch für die Therapie.

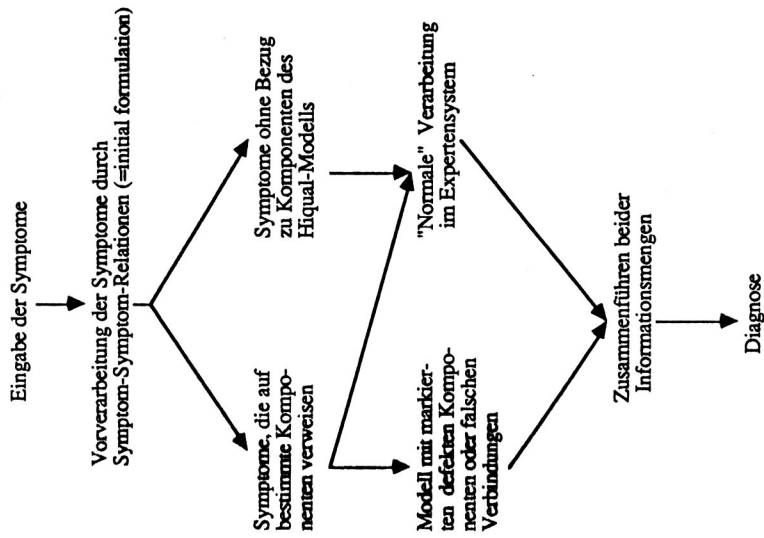


Abb. 3.10: Grobmodell: Integration von Hiqual in ein Expertensystem. Um das Wesentliche der Integration zu erfassen, wird bei diesem Modell auf Einzelheiten, wie Schleifen oder die Unterscheidung zwischen Grob- und Feindiagnosen verzichtet.

Beispiel: Wie kann die Lösung eines Problems durch Hiqual erleichtert werden?

Symptom: Das Auto verliert Öl.

Dieses Symptom stößt das Hiqual-Modell an nachzuforschen, welche Ölkreisläufe es im Auto gibt. Durch Fragen nach der Lokalität des Austritts des Öls kann man die Zahl der möglichen Ölkreisläufe verringern oder sogar innerhalb eines Kreislaufes Einschränkungen machen. Das Symptom "Das Auto verliert Öl" wird mit Hiqual und den Fragen an den Benutzer auf das Symptom "bestimmte Komponenten verlieren Öl" verfeinert. Die Wahrscheinlichkeit dieser möglichen Komponenten beziehungsweise ihrer Diagnosen wird im Expertensystem durch entsprechende Werte erhöht. Zur Unterstützung des Benutzers können auch alle Komponenten ausgegeben werden, durch die Öl fließt (natürlich auf der gewünschten Ebene.)

2. H I Q U A L I

Hiquall erlaubt nur eine Analyse korrekter technischer Systeme mit Fragestellungen wie:

- Welche Komponente ist mit welcher verbunden?
- Was macht diese Komponente? (Abhängigkeit der Ausgabe von der Eingabe)
- Aus welchen Teilen besteht eine Komponente?
- Wie stehen die Ereignisse und Zustände zeitlich oder ontologisch zueinander?

Kurz: Wie arbeitet das Modell, wenn es korrekt ist?

Hiquall wurde nun so erweitert, daß auch einfache Fragen über defekte Komponenten und ihre Auswirkungen beantwortet werden können. Zusätzlich besteht die Möglichkeit, das Hiquall-Modell in ein Expertensystem zu integrieren, indem Verbindungen zwischen Symptomen und Diagnosen einerseits und den Komponenten des Hiquall-Modells andererseits aufgebaut werden. Dieses neue System wird Hiquall genannt.

Die folgenden Fragen können mit Hiquall beantwortet werden:

- (1) Ein pathologisches Symptom ist mit einer Komponente des Hiquall-Modells verbunden. Dadurch wird vermutet, daß die Komponente defekt ist. Welche Folgefehler können durch diese defekte Komponente auftreten? Oder anders gefragt: Welche Komponenten sind von einem Defekt dieser Komponente betroffen?
- (2) Durch pathologische Symptome sind mehrere Komponenten des Hiquall-Modells als fehlerhaft gekennzeichnet. Welche defekte Komponente ist der Verursacher dieser auftretenden Defekte? Gibt es einen oder mehrere Verursacher?

2.1 Folgen einer defekten Komponente oder fehlerhaften Zeitbeziehung.

2.1.1 Welche Folgen kann eine defekte Komponente verursachen?

Ausgangspunkt ist ein korrektes Modell. Angenommen wird, daß eine oder mehrere Komponenten des Hiquall-Modells defekt sind und fehlerhafte Ausgaben weitergeben. Ziel ist es, sowohl die Komponenten zu finden, die solche fehlerhaften Ausgaben als Eingaben erhalten und dann ebenfalls inkorrekte Ausgaben weitergeben, als auch jene Komponenten, die trotz dieser defekten Komponente korrekt weiterarbeiten.

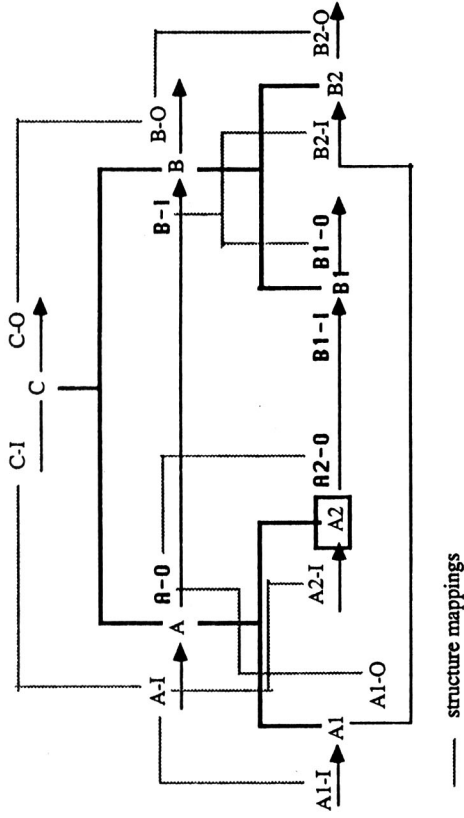


Abb. 3.11: Schematische Darstellung der Weitergabe fehlerhafter Ausgaben in einem Hiquall-Modell.

Die Komponenten des Hiquall-Modells sind: C, A, B, A1, A2, B1 und B2. A2 ist die defekte Komponente. Die Komponente B1 erhält den fehlerhaften Output von A2; dieser ist jedoch für B2-O = B-O nicht relevant. Dadurch ist die Ausgabe der Komponenten C ebenfalls korrekt. Schlussfolgerung: A2 kann nicht die Ursache für Fehler bei C-O, B-O oder B2-O sein. A2 kann aber Fehler in B1-O erzeugen.

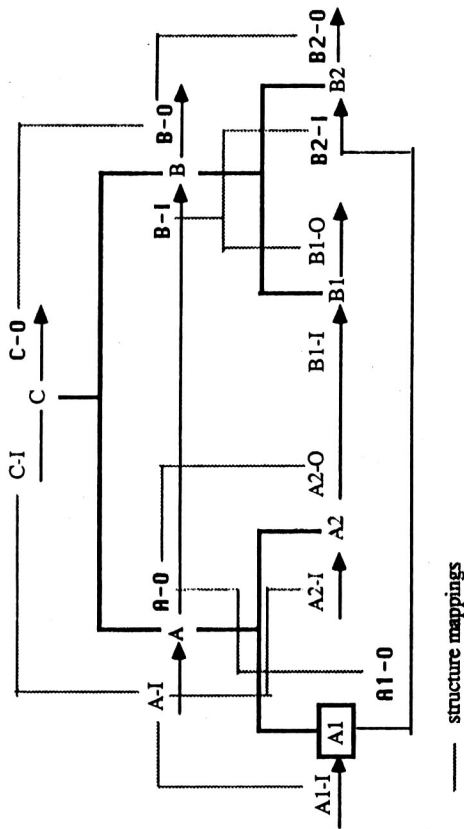


Abb. 3.12: Schematische Darstellung der Weitergabe fehlerhafter Ausgaben in einem Hiqual-Modell.

Das Hiqual-Modell entspricht dem obigen Modell. Diesemal ist die Komponente A1 defekt. A1-O gibt ihre Ausgabe an B2-I weiter, wodurch B2-O und (durch das structure-map) B-O und C-O fehlerhaft sein können. A1-O ist über ein structure-map mit A-O verbunden, das nun ebenfalls fehlerhaft werden kann. Durch A-O bekommt auch die Komponente B (B-I) eine falsche Eingabe. Nur A2 und B1 bleiben korrekt. **Schlussfolgerung: A1 kann nicht die Ursache für Fehler in A2 und B1 sein. Möglicherweise ist A1 die Ursache für die Fehler bei C-O, B-O und B2-O.**

Diese 'Berechnungen' können nur für horizontale und vertikale Relationen durchgeführt werden. Spielen zeitliche und ontologische Relationen eine Rolle, so muß man mit Hiqual2 oder anderen Mechanismen arbeiten.

Grobstruktur der Berechnungen.

Betrachte das Modell bis zu der Ebene, die die defekte Komponente enthält.

- (1) Ist eine Ausgabe fehlerhaft (am Anfang zumindest die Ausgabe der defekten Komponenten), so sind alle Komponenten, die mit dieser Ausgabe vertikal verbunden sind und im Hiqual-Modell auf einer höheren Ebene liegen, auch inkorrekt (structure-map).

- (2) Auf der Ebene der defekten Komponente werden inkorrekte Ausgaben als Eingaben an andere Komponenten weitergegeben und verursachen so scheinbar defekte Komponenten mit ebenfalls fehlerhaften Ausgaben.

- (3) Sind in Schritt 2 weitere fehlerhafte Ausgaben gefunden worden?

Ja, so gehe nach (1).

Nein, Ausgabe aller Komponenten mit möglicherweise fehlerhaften Ausgängen oder Ausgabe aller Komponenten, die trotz der defekten Komponente korrekt weiterarbeiten.

Zusammenfassung: Gibt man in diesen Algorithmus eine defekte Komponente ein, so erhält man als Ausgabe alle Komponenten, die von Folgen betroffen sein könnten, sowie alle Komponenten bei denen die defekte Komponente nicht als Verursacher in Frage kommt. Der Algorithmus betrachtet allerdings nur horizontale und vertikale Relationen.

2.1.2 Welche Folgen kann eine inkorrekte Zeitbeziehung verursachen?

Zeitrelationen gibt es in Hiqual zwischen Ereignissen, Zuständen und Ereignissen und Zuständen. Sie sind meist nicht nur einzelne Relationen, sondern ganze Relationenmengen. Mit Hilfe der Transitivitätstabelle von Allen kann man die Zeitbeziehungen berechnen. Erkennt man im realen System eine andere Zeitbeziehung als im Hiqual-Modell angeben, so kann man mit Hilfe der Transitivitätstabelle berechnen, welche Folgen diese inkorrekten Zeitbeziehungen für die verbundenen Ereignisse und Zustände haben.

2.2 Welche Komponente könnte die Ursache der auftretenden Fehler sein?

Ausgangspunkt ist ein korrektes Modell, in dem mehrere Komponenten als 'defekt' markiert sind. Betrachtet man alle gegebenen defekten Komponenten und ihre Beziehungen untereinander, so kann man die möglichen Verursacher-Komponenten dieser scheinbaren Defekte berechnen.

In jeder Komponente vom Hiqual sollte folgendes angegeben werden:

- Ist die Komponente leicht testbar (-> Ja/Nein) ?
- Sollte die Komponente, falls sie defekt ist, besser gleich ausgetauscht werden, oder sollte man ihre Unterkomponenten betrachten und diese testen?
- Muß die Komponente in bestimmten Zeitabständen oder nach einer bestimmten Zahl von Kilometern ausgetauscht werden (Beispiel: Ölwechsel und Zündkerzen). Diese Komponenten müssen nicht unbedingt die Ursache des momentanen Defektes sein, sind es aber sehr wahrscheinlich.

Grobstruktur zur Berechnung der Ursache.

(1) Betrachte alle 'defekten' Komponenten und ihre Beziehungen untereinander. Erstelle aus dem Hiqual-Modell ein Netzwerk, dessen Knoten die 'defekten' Komponenten sind. Liegt im Hiqual-Modell zwischen zwei Knoten eine direkte oder indirekte hc/vc - Relation vor, so werden diese Knoten mit einem Pfeil verbunden.

(2) Ist eine dieser Komponenten des Netzwerks leicht testbar?

Nein -> Weiter bei (4)

Ja --> Ist die Komponente laut Test wirklich defekt?

Nein --> Diese Komponente kann kein Verursacher sein.

Ja --> Diese Komponente ist ein Verursacher. Ist ein Betrachten der Unterkomponenten sinnvoll?

Nein -> Weiter mit (3)

Ja --> Einer der Unterkomponenten ist der Verursacher. Gehe zu (1)

und setze statt der Komponente ihre Unterkomponenten ein.

Gib als zusätzliche Information ein, daß einer dieser

Unterkomponenten der Verursacher ist.

(3) Ist eine Komponente K in Schritt 2 als möglicher Verursacher angesehen worden?

Nein -> Weiter bei (5)

Ja --> Es wird überprüft, ob diese Komponente alle anderen Komponenten erklären kann (d.h. ob von dieser Komponente aus alle anderen Komponenten erreicht werden können).

Nein -> Es liegen noch weitere Verursacher-Komponenten vor. Alle von K erklärbaren Komponenten und K werden gestrichen und der Algorithmus mit den restlichen Komponenten fortgesetzt.

Ja --> Diese Komponente wird als Verursacher-Komponente ausgegeben.

- Stop -

(4) Liegen zwischen zwei Komponenten (K1 und K2) Relationen in beide Richtungen vor, so werden diese zu einer Komponente {K1,K2} zusammengefaßt.

(5) Wähle eine Komponente, die keine eingehenden Pfeile besitzt. Gibt es keine solche Komponente, so gehe zu (6).

Andernfalls: Diese Komponente K5 kann durch keine andere Komponente erklärt werden und muß also eine verursachende Komponente sein.

Ist K5 eine Komponentenumenge?

Nein -> Ausgabe von K5.

Ja --> Sind in K5 Komponenten enthalten, die leicht testbar waren?

Nein -> Ausgabe der gesamten Menge.

Ja --> Gibt es eine oder mehrere Komponenten, die als defekt erkannt wurden?

Nein --> Ausgabe aller Komponenten, die nicht im Test als 'korrekt' bezeichnet wurden.

Ja --> Ausgabe nur der Komponenten, die im Test als 'defekt' erkannt wurden.

K5 und alle durch K5 erklärbaren Komponenten werden gestrichen. Liegen noch weitere Komponenten vor?

Nein -> Stop.

Ja --> Weiter mit (5).

(6) Jede Komponente hat einen eingehenden Pfeil. Es liegt also ein Zyklus vor. Alle Komponenten des Zyklus werden zu einer Komponente zusammengefaßt: {K1,...,Kn}. Weiter mit (5).

Beide Algorithmen (siehe 2.1.1 und 2.1.2) können leicht geändert auch für andere Relationen benutzt werden.

Da alle anderen zeitlichen und ontologischen Relationen zwischen Ereignissen und Zuständen auftreten, werden die Komponenten dann durch Ereignisse und Zustände ersetzt. Auch Symptome können direkt auf Ereignisse oder Zustände verweisen (siehe Verbesserungsvorschläge, 2.3.1.2). Dadurch kommt man aber leicht zu zahlreichen detaillierten Angaben. Diese Anwendung sollte auf Spezialfälle beschränkt bleiben.

Man kann auch die Ereignisse und Zustände den jeweiligen Komponenten zuordnen und auf der Ebene der Komponenten weiterarbeiten. Diese Berechnungen scheinbarer und wirklich defekter Komponenten sind nur dann korrekt, wenn entweder nur zeitliche oder nur ontologische Relationen von dem Defekt betroffen sind. Liegen Fehler bei beiden Relationen vor, so kann eine getrennte Berechnung der Relationsarten kein korrektes Ergebnis bringen. Hier muß man versuchen, Hiqual2 oder andere Mechanismen einzusetzen.

2.3 Integration des Hiqual-Modells in das Expertensystem.

2.3.1 Aufbau der Verbindungen zwischen Symptomen und Komponenten.

Im Expertensystem werden die Symptome in einem Symptomwald mit der Relation "Verfeinerung" dargestellt. Nach der Initialphase liegen -je nach Expertensystem- noch weitere Symptom-Symptom-Relationen vor. Oft gibt es im Symptomwald auch einen Vererbungsmechanismus (Wertfortpflanzung). Das Hiqual-Modell ist ein Baum, dessen Hauptrelation die Part-Of Relation ist. Desweiteren gibt es noch ontologische und zeitliche Relationen zwischen den Ereignissen und Zuständen und die hc/vc-Relationen zwischen den Komponenten.

Die Verbindung von einem Symptom zu einer Komponente des Hiqual-Modells lautet:

Das Symptom S liegt vor, daraus folgt, daß die Komponente K1 mit einer Wahrscheinlichkeit von X Prozent defekt ist (Die Angabe der Wahrscheinlichkeit kann auch weggelassen werden, oder durch qualitative Angaben ersetzt werden).

Die Verbindung von einer Komponente zu einem Symptom lautet:

Die Komponente K1 ist mit einer Wahrscheinlichkeit von Y Prozent defekt; sich nach, ob das Symptom vorliegt (Auch hier kann die Angabe der Wahrscheinlichkeit entfallen oder durch qualitative Angaben ersetzt werden).

Die Verbindung zwischen einem Symptom und einer Komponente liegt immer in beiden Richtungen vor.

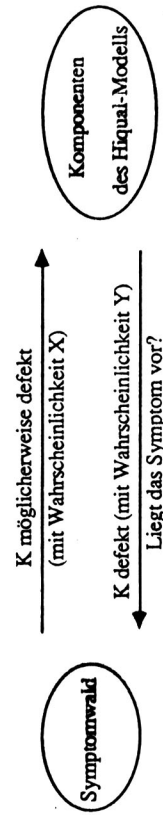


Abb. 3.13: Schematische Darstellung der Verbindungen zwischen den Symptomen des Expertensystems und den Komponenten des Hiqual-Modells.

Jede Verbindung zwischen einer Komponente und einem Symptom kann in jeder Richtung nur einmal feuern. Gilt diese Einschränkung nicht, so könnten sich die Komponente K1 und das Symptom S bezüglich ihrer Wahrscheinlichkeit hochschaukeln. Um jederzeit eine Rücknahme zu ermöglichen, wird in der Komponente notiert, welche

Symptome die Wahrscheinlichkeit erhöht haben. Nach der Rücknahme kann diese Verbindung wieder feuern.

2.3.1.1 Wo sollen die Verbindungen zwischen Symptomen und Komponenten gespeichert werden?

1. Möglichkeit: In einer Tabelle oder einer Regelmenge.

Diese wird wegen der großen Zahl von Variationsmöglichkeiten zwischen den einzelnen Komponenten eines Komponententyps und den Symptomen zu umfangreich.

2. Möglichkeit: Bei den einzelnen Symptomen.

Die Verbindungen Symptom-Komponente sind zu zahlreich. Man müßte statt der Komponenten die Komponententypen heranziehen und diese Verbindung später wieder je nach Komponententyp modifizieren. Diese Verbindungen sind auch davon abhängig, wo die Komponente im Modell steht und in welchem Modell sie vorkommt.

3. Möglichkeit: Bei den Komponenten.

Dies ist nicht möglich, da nur die Komponententypen abgespeichert werden und die einzelnen Komponenten dynamisch erzeugt werden.

4. Möglichkeit: Bei den Komponententypen.

Auch hier sind die Verbindungen davon abhängig, welche Komponente initiiert wird und in welchem Modell sie wo auftritt.

Eine geeignete Lösung ist die Kombination der ersten und der vierten Möglichkeit. In den Komponententypen werden alle Verbindungen zu Symptomen gespeichert, die unabhängig vom Modell und den einzelnen Komponenten sind. Bei der Initialisierung der Komponenten kann man bestehende Verbindungen noch modifizieren. Alle weiteren Verbindungen werden in einer Tabelle oder einer Regelmenge gespeichert, die nun nicht mehr so umfangreich ist.

2.3.1.2 Verbesserungsvorschläge.

1. XOR-Feuerung.

Es gibt Symptome, die auf verschiedene Komponenten zeigen. Man kann aber mit sehr hoher Wahrscheinlichkeit annehmen, daß nur eine dieser Komponenten defekt ist.

Beispiel1: Symptom S : Öl tritt aus.

Im System sind drei Ölkreisläufe enthalten, die nur sehr schlecht in einer Modellkomponenten zusammengefaßt werden können. Das Symptom weist auf alle drei Komponenten mit einer mittleren Wahrscheinlichkeit. Die XOR-Feuerung bietet die Möglichkeit die Wahrscheinlichkeit nicht aufzuteilen, sondern einer Komponente ganz zu übertragen. Dazu versucht man die Wahrscheinlichkeit der drei Komponenten durch andere Symptome $S_1 \dots S_n$ zu erhöhen oder zu verringern.

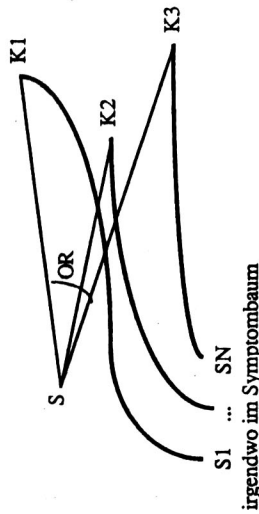


Abb. 3.14: Kann keine Komponente bestätigt werden, so werden wie bisher 'mittlere' Wahrscheinlichkeiten verteilt, die XOR-Feuerung bietet aber im Hinterkopf.

Beispiel2:

Eine Oberkomponente enthält alle drei Ölkreisläufe. Die Ölkreisläufe können entsprechenden Verfeinerungssymptomen zugeordnet werden. Hier ist keine XOR-Feuerung notwendig. Wenn möglich sollte man das System wie im Beispiel2 gezeigt aufbauen und Probleme, wie im Beispiel1 gezeigt, vermeiden.

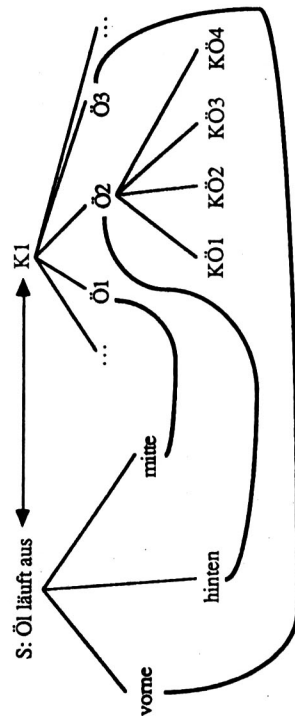
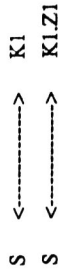


Abb. 3.15: Schematische Darstellung von Beispiel 2.

2) Symptome können auch auf Zustände, nicht nur auf Komponenten verweisen.



Bei speziellen Symptomen ist es möglich, innerhalb der Komponenten den Zustand, das Ereignis oder den zeitlichen Zusammenhang zu spezifizieren, der defekt ist. Man muß einen Kompromiß finden zwischen Genauigkeit (möglichst viele Verbindungen) und Effizienz (möglichst wenige Verbindungen). Die Entscheidung, wie genau die Verbindungen sein sollen, hängt von dem jeweiligen Expertensystem ab.

2.3.2 Aufbau der Verbindungen zwischen Komponenten des Hiqual-Modells und Diagnosen.

Das Hiqual-Modell ist ein Baum dessen Knoten den Komponenten des technischen Systems entsprechen und in der Part-Of-Relation zueinander stehen. Desweiteren gibt es noch ontologische und zeitliche Relationen sowie die hc/vc-Relationen zwischen den Komponenten.

Auch Diagnosen sind in den meisten Expertensystemen hierarchisch angeordnet. Sowohl Hierarchien als auch Heterarchien sind möglich. Auf höheren Ebenen sind meist Grobdiagnosen, auf den tieferen Ebenen Feindiagnosen zu finden.

Besteht eine Verbindung von einer Komponente des Hiqual-Modells zu einer Diagnose, so kann folgender Schluß gezogen werden:

Wenn die Komponente K mit einer Wahrscheinlichkeit von X Prozent defekt ist, folgt daraus, daß die Wahrscheinlichkeit, daß die Diagnose vorliegt um Z Prozent erhöht wird. (Die Angabe der Wahrscheinlichkeit kann weggelassen werden oder durch qualitative Angaben ersetzt werden.)

Die Verbindung von einer Diagnose zu einer Komponente des Hiqual-Modells hat zwei Anwendungsfälle.

- (1) Die Diagnose soll etabliert werden. Kann die Komponente dazu etwas beitragen?
- (2) Bei einer Überprüfung der Diagnose durch das Hiqual-Modell in Hiqual2 können die entsprechenden Komponenten gefunden werden.

Diskussion zu (1).

Zu diesem Punkt gibt es zwei Interpretationsmöglichkeiten:

A. Geh von der Komponente aus zu den entsprechenden Symptomen.

Falls die direkte Verbindung Diagnose --> Symptom existiert, gibt es einfachere Wege, um von der Diagnose zu den Symptomen zu gelangen, als den Umweg über das Hiqua-Modell zu nehmen.

Andererseits muß man sich überlegen, ob die direkte Verbindung Diagnose-Symptom überhaupt noch existiert. Ob man diese Interpretation der Verbindung einführt, hängt von dem gewählten Expertensystem ab, sowie von dem betrachteten Anwendungsbereich und der Art der Implementierung.

B. Sieh im realen System nach, ob diese Komponente defekt ist.

Komponenten, die leicht überprüfbar sind, können einem Symptom der Form 'Komponente ist defekt' entsprechen und schon früher abgefragt worden sein, so daß diese Verbindung gar nicht erst zum Zuge kommt. Dem ist entgegenzuhalten, daß es leicht überprüfbare Symptome gibt, die sehr speziell sind und somit auf höheren Ebenen nicht nachgeprüft werden. Für diese Symptome ist die Verbindung von großer Bedeutung. Die Integration der Verbindungsart hängt davon ab, welches Expertensystem und welches Anwendungsgebiet vorliegen.

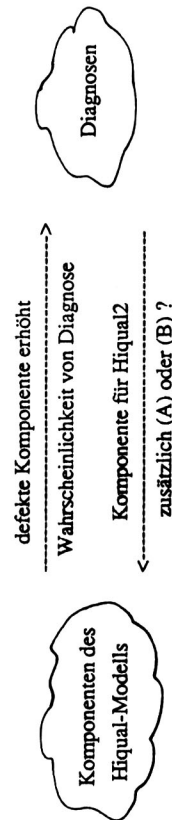


Abb. 3.16: Verbindungen zwischen Komponenten des Hiqua-Modells und Diagnosen.

Wo sollen nun die Verbindungen zwischen den Komponenten und den Diagnosen gespeichert werden? (siehe 2.3.1.1)

Alle Verbindungen zu Diagnosen werden in den Komponententypen gespeichert, die unabhängig vom Modell und von der einzelnen Komponente sind. Bei der Initialisierung der Komponenten kann man bestehende Verbindungen modifizieren. Alle weiteren Verbindungen werden in einer Tabelle oder in einer Regelmenge gespeichert.

Bei kleineren Anwendungsgebieten ist es auch möglich, alle Verbindungen in einer Regelmenge zu speichern.

Die Verbindung zwischen der Komponente und den Diagnosen kann nur einmal feuern. Die Verbindung 'Komponente für Hiqua' steht immer zur Verfügung während die Verbindungen (A) und (B) von den Diagnosen zu den Komponenten des Hiqua-Modells nur dann in Aktion treten können, wenn die Umkehr-Verbindung noch nicht gefeuert hat. Solange können sie allerdings mehrmals aktiviert werden.

Genau wie bei den Symptomen, so ist es auch bei den Diagnosen möglich, Verbindungen nicht nur zu den Komponenten, sondern auch zu den Zuständen oder Ereignissen der Komponenten aufzustellen.

2.3.3 Aufbau der Verbindungen zwischen Komponenten des Hiqua-Modells und Therapien.

Einige Expertensysteme geben eine oder mehrere Diagnosen als Ergebnis aus. Andere hingegen gehen einen Schritt weiter und erörtern auch Therapievorschläge. Diese werden meist anhand der Ergebnisd Diagnose aufgestellt. Liegen mehrere mögliche Diagnosen vor, so kann die Therapie aus den vorliegenden Symptomen und möglichen Diagnosen abgeleitet werden. Dazu kann das Hiqua-Modell eingesetzt werden. Ist eine Komponente des Hiqua-Modells mit sehr hoher Wahrscheinlichkeit defekt (Achtung: Die Komponente muß selber defekt sein. Die falschen Ausgabewerte dürfen nicht aufgrund fehlerhafter Eingaben entstanden sein), so wird als Therapie vorgeschlagen, die Komponente auszutauschen.

Komponenten, die für einen solchen Austausch in Frage kommen, haben eine Verbindung zu der entsprechenden Therapie.

Eine Verbindung Therapie --> Komponente bedeutet, daß die Therapie Auswirkungen auf die Komponente hat (nicht umgekehrt) (Beispiel: verbessert werden, gereinigt werden). Diese Verbindungen können dazu benutzt werden, um die Richtigkeit einer Therapie zu überprüfen (Vorteil der technischen Systeme gegenüber den medizinischen Systemen).

3. H i q u a l 2

Hiqual ermöglicht eine Analyse korrekter technischer Systeme. Hiqual1 kann zusätzlich einfache Fragestellungen über defekte Komponenten beantworten und das Hiqual-Modell in ein Expertensystem integrieren. Sind sowohl zeitliche als auch ontologische Relationen inkorrekt, so genügt auch Hiqual1 den Anforderungen nicht mehr.

Ausgehend vom ursprünglichen Hiqual-System wird ein neues System (Hiqual2) geschaffen, das es ermöglicht, defekte Komponenten und Zeitrelationen zu integrieren, ohne daß bei der Simulation Konsistenzprobleme auftreten. Die Simulation des Hiqual-Modells mit defekten Komponenten kann Expertensystemen auch bei der Überprüfung der Diagnose weiterhelfen.

3.1 Möglichkeiten ein defektes Teil in das korrekte Hiqual-Modell einzusetzen.

Die bisherigen Komponententypen erzeugen nur korrekte Komponenten. Für jedes defekte Teil muß erst ein neuer Komponententyp kreiert werden. Dazu muß man die defekte Komponente sehr genau beschreiben. Alle Zustände und Regeln, alle Relationen zu anderen Komponenten, alle In- und Outports und Funktionen müssen festgelegt werden.

Eine Analyse in Hiqual läuft auf allen Ebenen ab und zwischen den Ebenen erfolgen Konsistenzüberprüfungen. Setzt man die defekte Komponente in das Hiqual-Modell ein, so ist die Konsistenz nicht mehr gewahrt und die Analyse stoppt oder gibt fehlerhafte Werte aus. Deshalb müssen alle Vorfahren der defekten Komponente bis hin zur Wurzel geändert werden.

Zur Lösung dieses Problems gibt es zwei Möglichkeiten:

1. Man verändert auch die Vorfahren der defekten Komponente und erreicht so, daß die Konsistenzbedingungen wieder erfüllt sind.
2. Die Vorfahren der defekten Komponente werden nur als 'defekt' gekennzeichnet und die Hiqual-Simulation läuft nur auf einer Ebene ab.

3.1.1 Anpassung aller Vorfahren der defekten Komponenten.

Die Integration einer defekten Komponente erfordert sehr viel Aufwand, egal ob es die defekte Komponente selber oder einer ihrer Vorfahren ist. Alle Zustände und Regeln, alle Relationen zu anderen Komponenten, alle In- und Outports und Funktionen müssen festgelegt werden. Prinzipiell gibt es zwei Möglichkeiten, Vorfahren defekter Komponenten zu integrieren (siehe Abb. 3.17).

A) Die Folgen der defekten Komponente K18 werden in der nächst höheren Ebene betrachtet. Dabei kann man versuchen die Vorfahren entsprechend dem Defekt der Komponente K18 zu modifizieren. Diese Möglichkeit läßt sich wegen der Unabhängigkeit der Modelle nur in Spezialfällen anwenden.

B) Die Vorfahren der defekten Komponente werden genauso behandelt wie die defekte Komponente selbst. Dieses Verfahren ist immer anwendbar, wenn von allen Vorfahren genügend Information zur Verfügung stehen. Diese Möglichkeit erfordert großen Aufwand, besonders für den Anwender, der alle Komponententypen und Attributrestriktionen (während der Laufzeit) beschreiben muß.

Hat man das Verhalten aller Vorfahren der Komponente K18 beschrieben, so kann man in vielen Fällen auf die Simulation verzichten, denn die Auswirkungen der defekten Komponente auf höherer Ebene sind zum großen Teil bekannt.

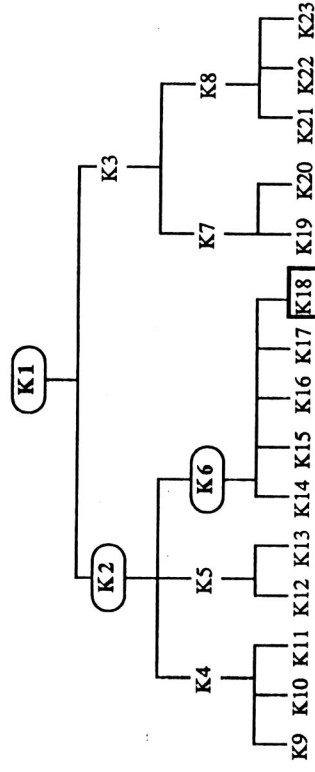


Abb. 3.17:

K18 ist die defekte Komponente

K6, K2 und K1 sind die Vorfahren von K18. Auch sie müßten durch "defekte" Komponenten ersetzt werden.

In den defekten Komponenten können auch Funktionen enthalten sein. Diese Funktionen werden unabhängig von den Unterkomponenten angegeben. Im Simulationsmodell sollte man daher zwei Funktionen angeben:

- eine unabhängige Funktion wie bisher und zusätzlich
 - eine Funktion in Abhängigkeit von ihren Unterkomponenten.
- Beide Funktionen müssen in den korrekten Komponenten immer identisch reagieren. Die Wirkung defekter Unterkomponenten kann nun einfacher berechnet werden.

Aber: 1. Nicht alle Komponenten enthalten Funktionen.

2. Die bisherige unabhängige Funktion kann nicht leichter berechnet werden. Sie kann gegebenenfalls in defekten Komponenten wegfallen. Die abhängige Funktion kann also die Wirkung von K1 in Abhängigkeit von K14, K15, K16, K17, K18 und K3 angeben (siehe Abb. 3.17).

3.1.2 Die Vorfahren der defekten Komponenten werden als 'leere' Hüllen dargestellt.

Eine andere Möglichkeit defekte Komponenten einzubinden, besteht darin HiquaI nur auf einer Ebene ablaufen zu lassen, und die Konsistenzüberprüfungen 'auszuschalten'. Wegen der Unabhängigkeit der Modelle auf den verschiedenen Ebenen ist dies möglich. Die hc/vc Relationen zwischen den einzelnen Komponenten müssen aus den höheren Ebenen berechnet werden. Dazu läßt man die Vorfahren defekter Komponenten als 'leere' Hüllen stehen, und alle Verbindungen zwischen Komponenten bleiben erhalten. Attributrestriktionen, die sich auf diese Komponenten beziehen, müssen gegebenenfalls verändert oder gestrichen werden.

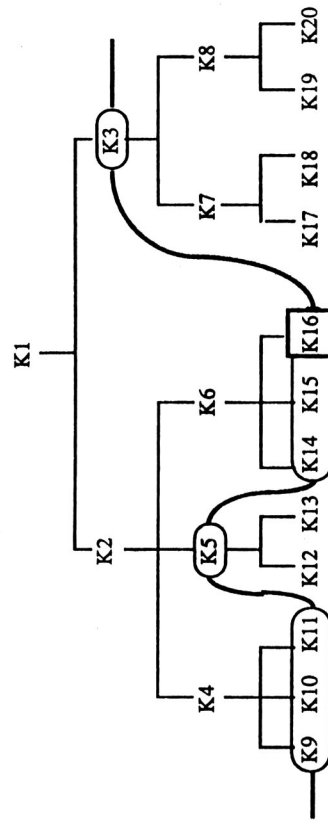


Abb. 3.18:

K16 ist die defekte Komponente. Die Wahl der Simulationsebene ist frei mit der Einschränkung, daß die defekte Komponente darin enthalten sein muß. K14 und K17 können nur konkretisiert, nicht aber abstrahiert werden (auch nicht erwünscht, K16 ist die defekte Komponente). Standardmäßig wird die Ebene simuliert, in der sich die defekte Komponente befindet (K9 bis K20).
 Nachteil: Will man eine andere Ebene betrachten, so muß man neu simulieren. Es ist nicht möglich, die Simulation mit K5 und den Unterkomponenten K12 und K13 zu starten. Dies ist nur beim Ablauf des korrekten Modells möglich.

3.1.3 Wertung der beiden Möglichkeiten.

Eine ausführliche Beschreibung der Vorfahren der defekten Komponente ist meist nicht möglich, da zu wenig Informationen zur Verfügung stehen. Die Erstellung neuer Typen erfordert großen Aufwand. Aus diesen Gründen wird wohl in den meisten Fällen zu der zweiten Möglichkeit gegriffen und die Komponente als 'defekt' gekennzeichnet werden. Statt der Komponenten setzt man eine 'leere' Box ein. Alle bisherigen Verbindungen und Relationen zu anderen Komponenten werden beibehalten, soweit sie nicht explizit als 'defekt' oder anders verlaufend angegeben werden.

3.2 Simulation von Teilen des HiquaI-Modells.

Eine Simulation des HiquaI-Modells ist sehr aufwendig. Da stellt sich die Frage, ob es nicht sinnvoll ist, statt des ganzen HiquaI-Modells nur einen kleinen Teil zu simulieren und somit den Aufwand zu reduzieren. Nimmt man aber einen Teilbaum aus dem gesamten Modell heraus, so können Beziehungen, die von dem Teilbaum in den übrigen Baum und zurück zum Teilbaum verlaufen, nicht berücksichtigt werden. HiquaI bietet die Möglichkeit, eine oder mehrere Komponenten der Abstraktionsebene genauer zu betrachten, während die anderen Komponenten auf ihrem Abstraktionsniveau bleiben (Der Teilbaum bleibt im Modell integriert, siehe Abb. 3.19).

Beim Einsatz von Hiqua2 muß man für jeden einzelnen Anwendungsfall prüfen, ob die defekte Komponente exakt genug beschrieben werden kann. Die Simulation des Modells mit einer Komponente, von der man nur weiß, daß sie defekt ist, genügt nicht. Man muß genau angeben können, welche Ausgaben die Komponente weitergibt und wie sie sich verhält. Sie muß also ebenso exakt beschrieben werden, wie eine korrekte Komponente. Das ist ein großer Nachteil von Hiqua2.

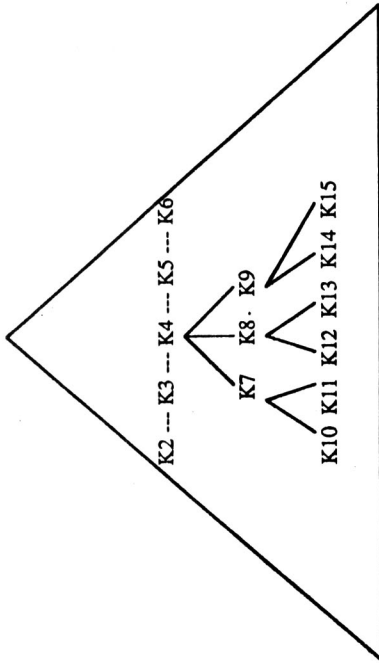


Abb. 3.19: Hiqua-Modell mit variablen Abstraktionsebenen
 K2, K3, K4, K5 und K6 sind die Elemente einer Abstraktionsebene. K4 kann genauer betrachtet werden durch K7, K8 und K9. K7 wiederum kann durch K10 und K11 konkretisiert werden.
 K8 wird durch K12 und K13 und K9 durch K14 und K15 konkretisiert.
 Eine weitere Abstraktionsebene ist dann: K2, K3, K10, K11, K12, K13, K14, K15, K5 und K6.

Der Teilbaum, bestehend aus K4, K7, K8, K9, K10, K11, K12, K13, K14 und K15 bleibt durch die Komponenten K2, K3, K5 und K6 im Modell integriert. Die Unterkomponenten von K2, K3, K5 und K6 interessieren dabei nicht. Die variablen Abstraktionsebenen von Hiqua ersetzen die Möglichkeit einen Teilbaum aus dem Hiqua-Modell herauszuschneiden und isoliert zu betrachten.

3.3 Pro und Contra Hiqua2

Hiqua2 erlaubt die Simulation eines Modells mit eingebauten defekten Komponenten. Dadurch erhält der Anwender Informationen, die in assoziativen Expertensystemen nicht verfügbar sind.

In Hiqua findet auf Grund der Konsistenzüberprüfung und der Attributrestriktionen eine kontinuierliche logische Überprüfung der Eingabedaten statt. Beim Einsatz einer defekten Komponente entfällt diese Überprüfung.

Kapitel 4 : Integration von Hiqua1, Hiqua11 und Hiqua12 in die einzelnen Expertensysteme.

1. Integration von Hiqua1, Hiqua11 und Hiqua12 in Abel.

1.1 Das Expertensystem Abel.

Abel ist ein modellbasiertes Diagnostik-Expertensystem zur Simulation von Störungen im Säure/Base- und Elektrolythaushalt. Die zentrale Datenstruktur ist das PSM (= patientenspezifische Modell), das die Krankheit des Patienten auf mehreren Abstraktionsstufen darstellen kann. Die Diagnosehüllen sind eine Erweiterung der PSM's, in denen zusätzlich unsicheres Wissen gespeichert wird. Die Mechanismen zur Erstellung der PSM's und der Diagnosehüllen lauten: Aggregation, Elaboration, Komponentensumma und Projektion. Umfangreiches Hintergrundwissen steht durch die Modelle zur Darstellung des Aufbaus des Organismus, der Materialflußrelationen und der Ätiologie zur Verfügung.

1.2 Repräsentationsmechanismen in Abel und ihr Zusammenwirken mit Hiqua1.

In Abel gibt es sehr viele Repräsentationsmechanismen.

(1) Part-Of-Hierarchie der Organsysteme.

Diese Part-Of-Hierarchie entspricht genau den Part-Of-Relationen des Hiqua1-Modells. Wird Hiqua1 integriert, so kann man diese Hierarchie weglassen.

(2) Materialfluß-Relationen.

Die Relation 'connected with' gibt es in Hiqua1 nicht. Sie kann aber in Hiqua1 durch die horizontalen und vertikalen Relationen zwischen den Komponenten dargestellt werden.

(3) Enthalten-in-Relationen.

Diese Relationen existieren in Hiqua1 nicht. Sie müssen auch bei einer Integration von Hiqua1 erhalten bleiben.

(4) Grobe anatomische Beziehungen.

Sie dienen dazu, die Grobteilung des Systems aus einem anderen Blickwinkel zu sehen, als dies auf der obersten Ebene der Part-Of-Hierarchie geschieht. Auch diese Relation kann durch das Hiqua1-Modell nicht ersetzt werden.

(5) Modelle zur Darstellung von ätiologischem Wissen.

Solche Modelle können durch Hiqua1 nicht ersetzt werden und bleiben weiter bestehen.

(6) Physiologisches Wissen.

Das physiologische Wissen beschreibt die normale Funktionsweise des Körpers. Dies entspricht der korrekten Funktionsweise der einzelnen Komponenten auf allen Ebenen und wird durch Hiqua1 abgedeckt.

Eine Integration von Hiqua1 macht also die Part-Of-Hierarchie, die Materialflußrelation und das physiologische Wissen von Abel überflüssig. Die groben anatomischen Beziehungen, die Enthalten-in Relation und die Modelle zur Darstellung von ätiologischem Wissen bleiben bestehen. Das Hiqua1-Modell kann seine Part-Of-Relationen, die horizontalen und vertikalen Relationen und die Angabe der Funktionsweise der korrekten Komponenten auf jeder Ebene ausnutzen.

Ein Vorteil des Hiqua1-Modells gegenüber den Hierarchien von Abel sind seine zusätzlichen Informationen und seine Flexibilität. Abel muß für jeden Maschinentyp die Hierarchien von Grund auf neu erstellen. Hiqua1 dagegen kann sich auf den Zusammenbau der vorhandenen Komponenten beschränken.

(Dieses Problem taucht bei dem Anwendungsgebiet von Abel nicht auf, da der Körperbau aller Menschen gleich ist. Werden aber physikalisch-technische Probleme gelöst, so kommen auch verschiedene Maschinen- oder Systemtypen vor.)

1.3 Gegenüberstellung der Modelle von Hiqua1 und Abel.

In Abel gibt es mehrere Hierarchien, die physiologisches Hintergrundwissen speichern. Zusätzlich gibt es das PSM (=patientenspezifisches Modell), das eine Simulation des realen Systems mit allen Fehlern wiedergibt. In Hiqua1 wird aus den verschiedenen Modelltypen jeweils das physiologische Modell erstellt. Dieses physiologische Modell kann das korrekte Modell simulieren.

Abel geht nicht wie Hiqual vom korrekten Modell aus, sondern nimmt alle ihm bekannten Werte (physiologischer oder pathophysiologischer Zustände) und setzt sie zueinander in Beziehung. Eine Modifikation von Hiqual in dieser Richtung ist nicht möglich, da in einem unvollständigen Hiqual-Modell die Simulation nicht gestartet werden kann. Trägt man nur die sichtlich defekten Komponenten in das korrekte Hiqual-Modell ein, so entstehen sehr leicht Konsistenzprobleme.

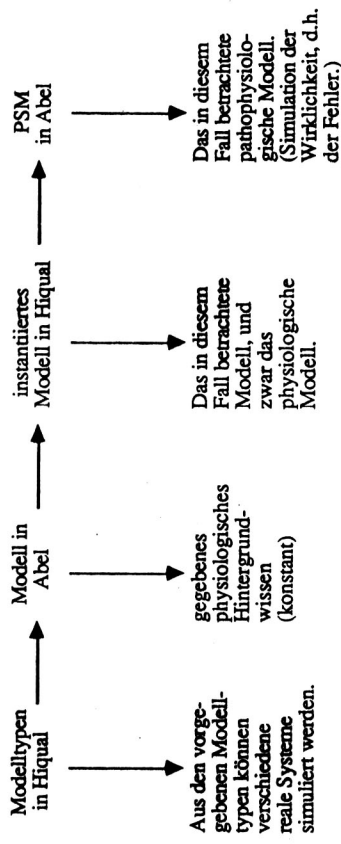


Abb. 4.1: Vergleich der Modelle von Hiqual und Abel.

Der Vergleich des in Hiqual instantiierten physiologischen Modells mit dem PSM in Abel ist kaum möglich, da die beiden Modelle sehr verschiedenartig aufgebaut sind.

1.4 Integration von Hiqual in Abel.

Bei einer Integration von Hiqual kann man die Part-Of-Hierarchie, die Materialflußrelation und das physiologische Wissen im Abel-System streichen. Ein großer Vorteil von Hiqual gegenüber den konstanten Hierarchien des Hintergrundwissens von Abel ist die Flexibilität im Aufbau.

Während die Hierarchien in Abel einmal aufgebaut werden und dann für alle Anwendungen konstant bleiben, kann man in Hiqual das jeweilige reale System des physikalisch-technischen Bereichs schnell und einfach nachbauen. Wenn man diese Flexibilität von Hiqual für Abel ausnutzen will, so muß man auch Änderungen in den übrigen Hierarchien des Hintergrundwissens, die nicht durch Hiqual dargestellt werden konnten und den kausalen Links vornehmen.

Die restlichen konstanten Hierarchien können entweder zusätzlich in Hiqual eingebaut werden, oder man erstellt Hiqual ähnliche flexible Modelle. Die Anpassung der kausalen Links, die nicht nur eine Beziehung zwischen Ursache und Wirkung sind, sondern eine multivariate Relation zwischen den Attributen der Ursache und der Wirkung sowie dem Kontext und dem Default ausdrücken, ist sehr schwierig.

Eine Möglichkeit ist es, die Links in zwei Gruppen aufzuteilen:

- (1) Die Links, die von der Struktur des Modells unabhängig sind.
- (2) Die Links, die mit jedem neuen Modell neu erstellt werden.

Eine andere Idee wäre, aus den 'causes'-Relationen im Hiqual-Modell der jeweiligen Umgebung im Modell und zusätzlichen Daten den kausalen Link automatisch zu berechnen. Die Frage ist nur, ob sich diese Theorie vom Aufwand her vertreten läßt.

Hiqual bietet aber noch mehr Hintergrundwissen als oben beschrieben. Alle Informationen darüber, was wann im korrekten Modell passiert, können abgefragt werden. Besonders zeitliche Relationen werden im korrekten Modell gespeichert. Die Zeit wird in Abel nur nebenbei behandelt. Hiqual würde eine Erweiterung der Zeitkomponente in Abel aber unterstützen. Vergleiche des pathophysiologischen Modells von Abel mit dem physiologischen Modell von Hiqual sind leider nicht möglich, da die Grundstruktur beider Modelle zu verschieden ist.

Der Begriff der 'Simulation' wird in den Systemen unterschiedlich angewandt. In Abel simuliert man einen Krankheitszustand, in den frühere Ereignisse integriert werden. In Hiqual dagegen ist die Simulation gleich dem Ablauf des nachgebildeten Modells.

====> Die Integration von Hiqual in das Abelsystem bedeutet eine Erweiterung des Systems. Da Abel wegen seiner enormen Komplexität nicht implementiert werden konnte, ist eine Integration von Hiqual, auch wenn sie besonders für physikalisch-technische Systeme einige Vorteile mitschlingt, heute noch nicht möglich.

1.5 Integration von Hiqual in Abel.

Die Integration von Hiqual in das Expertensystem Abel ist nicht sinnvoll. Hiqual beantwortet Fragen nach den möglichen Auswirkungen einer defekten Komponente und Fragen nach dem Verursacher verschiedener auftretender Defekte. Diese Fragen, die sehr komponentenbezogen sind, werden in Abel nicht gestellt oder im Simulationsmodell gleich genauer beantwortet. Der Umweg von den Symptomen über die

Komponenten zu den Diagnosen und die primitiven Berechnungsmethoden sind große Nachteile von Hiqua11.

====> Keine Integration von Hiqua11 in Abel.

1.6 Integration von Hiqua2 in Abel.

Die Integration von Hiqua2 in das Expertensystem Abel ist nicht sinnvoll. Die Simulationen, die Hiqua2 mit viel Aufwand durchführt, bringen Ergebnisse, die Abel nicht benötigt oder ihm schon zur Verfügung stehen. Deshalb ist sowohl ein zusätzliches Hiqua-Modell als auch das Ersetzen des PSM durch das Hiqua-Modell nicht sinnvoll. Beide Systeme -Hiqua2 und Abel- führen Simulationen durch. In Hiqua2 werden dazu in das korrekte Modell ein oder mehrere defekte Komponenten eingesetzt, deren Auswirkungen beobachtet werden. Da im Hiqua-Modell nicht direkt Symptome und Diagnosen eingegeben werden können, muß man den Umweg über die Komponenten machen, wodurch leicht Fehler entstehen können.

In Abel werden alle verfügbaren Informationen über das defekte System in die verschiedenen PSM's eingetragen. Im PSM wird das gesamte defekte System auf verschiedenen Abstraktionsstufen wiedergegeben. Das PSM simuliert so das wirkliche System (und nicht das korrekte System mit einem oder mehreren defekten Teilen) mit allen seinen Symptomen, Diagnosen, Laborwerten und Säure/Basestörungen.

====> Keine Integration von Hiqua2 in Abel.

2. Integration von Hiqua1, Hiqua2 und Hiqua2 in das Expertensystem Med2.

2.1 Die Expertensystemshell Med2.

Med2 ist ein Expertensystemshell, das auf den Problemtyp der assoziativen Diagnostik zugeschnitten ist. Zur Wissensrepräsentation gibt es sieben verschiedene Frametypen.

- Pathokonzepete (für Kontext-, Grob- und Feindiagnosen)
- Manifestationen (für Fragen und einfache Symptominterpretationen)
- Questionsets (zur Gruppenbildung zusammengehöriger Manifestationen)
- Varianten (Therapeutische Varianten der Pathokonzepete)
- Explanationsets (zur Erfassung der Erklärungsfähigkeit der Symptomatik)
- Lokalisationen (grobe, lokale Anordnungen) und
- strukturierte Regeln.

Die Diagnosen werden durch die Pathokonzepete hierarchisch und heterarchisch angeordnet. Zusätzlich erstellen die Diagnose-Diagnose-Regeln ein Diagnosenetzwerk. Die Kombination der hypothetisch deduktiven Vorgehensweise und der Differentialdiagnostik wird mit dem Working-Memory-Konzept realisiert, das sich am Vorbild des menschlichen Kurzzeitgedächtnisses orientiert und immer nur eine kleine Anzahl von Verdachtsdiagnosen enthält. Med2 unterscheidet zwischen der Vorverarbeitung (database reasoning) und der eigentlichen diagnostischen Auswertung (diagnostic reasoning) der Daten.

2.2 Repräsentation der Anatomie in Med2.

In Med2 gibt es die Möglichkeit, eine für den Anwendungsbereich allgemeingültige Lokalisationshierarchie zu definieren, auf die in speziellen Lokalisationsfragen Bezug genommen werden kann. Diese Hierarchie besteht aus Atomen und enthält folgende Informationen:

- Below: Menge der Unterlokalisationen (= Eine Liste von Lokalisationen, für die die definierte Lokalisation der Oberbegriff ist) Eine Lokalisation kann Unterlokalisation mehrerer Oberbegriffe sein, d.h. in mehreren Below-Listen vorkommen.
- Sided: Achsensymmetrie (= Information, ob die Lokalisation beidseitig symmetrisch ist)

[ns]
 abdomen
 central-abdomen
 umbilicus
 periumbilical
 [s]
 upper-abdomen
 epigastric
 subcostal
 midline-upper-abdomen
 RUQ-abdomen
 LUQ-abdomen
 [s]
 lower-abdomen
 suprapubic
 inguinal
 lower-midline-abdomen
 RLQ-abdomen
 LLQ-abdomen
 [s]
 lateral-abdomen
 loin

Abb. 4.2: Lokalisationshierarchie "abdomen" [Puppe 1986a]

[s] : symmetrisch

[ns]: nicht symmetrisch

Der Wertebereich der Lokalisationsfragen wird in Med2 als Zeiger auf eine Komponente der Lokalisationshierarchie angegeben.

Vorteile: o Ökonomie bezüglich des Aufbaus der Wissensbasis.

o Repräsentation von allgemeinem Wissen über Lokalisationen (hierarchische Beziehungen), was für den Regelinterpretierer von Bedeutung ist.

o Möglichkeit zur graphischen Lokalisationseingabe

Das Wissen über Lokalisationen nutzt der Regelinterpretierer aus, indem er bei dem Test auf Gleichheit von Lokalisationen erkennt, daß eine Lokalisation die andere umfaßt (d.h. in der Hierarchie weiter oben steht). In Med2 wird nicht ausgeführt wie der Regelinterpretierer auf diese Erkenntnis reagiert beziehungsweise ob diese Reaktion bei allen Regeltypen gleich ist. Dadurch entsteht der Eindruck, daß ähnlich wie bei der Zeitkomponente auch die Daten der Lokalisation zwar eingespeichert, aber nur spärlich benutzt werden.

2.3 Integration von Hiqual in Med2

Hiqual gibt die Struktur des realen Systems sowie seiner Komponenten an und analysiert das Verhalten und die Funktion der Komponenten. Das Hiqual-Modell kann in Med2 als umfangreiches Hintergrundwissen genutzt werden.

2.3.1 Ersetzen der Lokalisationshierarchie durch das Hiqual-Modell.

Da in technischen Systemen der Symmetrie keine große Bedeutung zufällt, kann man die sided-Komponente der Hierarchie weglassen. Die Lokalisationshierarchie wird zu einer Hierarchie umgeformt, indem man die Angabe mehrerer Pointer (die auch auf Elemente verschiedener Abstraktionsebenen zeigen können) als Wertemenge erlaubt.

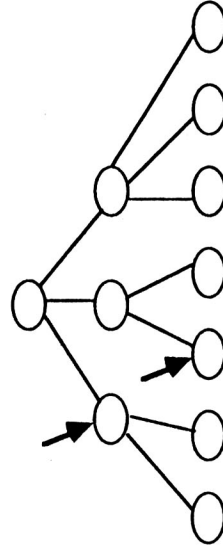


Abb. 4.3: Schematische Darstellung einer Lokalisationshierarchie.

Die Wertemenge besteht aus den Elementen, auf die einer der beiden Pointer zeigt, sowie die dazugehörigen Unterkomponenten.

Die Part-Of Beziehung in Hiqual entspricht dem Attribut 'Below'. Durch die Integration des Hiqual-Modells wird die Lokalisationshierarchie überflüssig. Das Questionset wird um komponentenorientierte Questionsets erweitert und es werden Regeln zur Verarbeitung der Antworten von Hiqual eingebaut. So können die Part-Of Relationen sowie Informationen der Komponenten in die Arbeit des Expertensystems integriert werden.

Die Möglichkeit der graphischen Lokalisationseingabe bleibt bestehen. Da der Benutzer die einzelnen Teile der Maschine oft nicht benennen kann, die Teile in einer graphischen Darstellung aber wiedererkennbar sind, ist diese Möglichkeit der Eingabe von großer Bedeutung.

2.3.2 Die Erklärungskomponente.

Die Erklärungskomponente des Med2-Systems ist sehr ausgefeilt, Erklärungen anhand eines Modells sind aber nicht möglich. Hiqual bietet dem Benutzer parmerbezogene Erklärungen an: Eine Simulation des korrekten Modells auf einer für den Benutzer geeigneten Abstraktionsebene.

Zusätzlich kann der Benutzer eine Reihe von Fragen bezüglich des Zusammenhangs und der Struktur der Komponenten an das korrekte Modell stellen. Die Antworten müssen vom Benutzer selbst verwertet werden. Folgende Fragen sind möglich:

- Welche zeitlichen Zusammenhänge bestehen zwischen zwei bestimmten Aktionen oder Zuständen?
- Aus welchen Teilen besteht eine Komponente?
- Wie verhält sich eine Komponente in diesem Modell?
- Welche ontologischen Zusammenhänge bestehen zwischen zwei bestimmten Komponenten (Aktionen, Zuständen)?
- Welche Aktionen (Ereignisse) sind kausal mit einem gegebenen Ereignis (Aktion) verbunden?

2.3.3 Konsistenzüberprüfung der Eingabe.

Eingabedaten werden durch das erweiterte Hintergrundwissen überprüft und gegebenenfalls mit geeigneten Erklärungen zurückgewiesen. Direkte Eingaben des Benutzers sind in Med2 nur insoweit möglich, als daß der Benutzer das entsprechende Questionset auswählt und dort seine Beobachtungen eingibt. Dadurch wird ein großer Teil der Eingabefehler abgefangen. Den Fragen innerhalb der Questionsets ist jeweils eine Menge möglicher Antworten zugeordnet. Ist die Antwort nicht in dieser Menge enthalten, so wird sie zurückgewiesen. In so einem Fall liefert das Hiqual-Modell eine Erklärung für die Zurückweisung und erläutert dem Benutzer seinen Eingabefehler (auf Abruf oder direkt).

Beispiel: Die Komponente K1 ist defekt. Frage an den Benutzer: Welche der drei Unterkomponenten von K1 ist defekt? Der Benutzer antwortet mit K11. Liegt eine falsche Antwort vor?

Dank der Lokalisationshierarchie erkennt Med2, daß K11 eine Unterkomponente von K4 ist. K4 und K11 werden von nun an im Modell als defekt angenommen. Diese Aufgabe der Lokalisationshierarchie übernimmt jetzt Hiqual.

Gibt der Benutzer statt einer Komponente einen Zustand oder ein Ereignis an, so kann das Hiqual-Modell im Gegensatz zu Med2 auch diese Antwort mit einem entsprechenden Hinweis zurückgeben, (Sie haben statt einer Komponente einen Zustand angegeben! Komponente X nimmt diesen Zustand an. Bitte erneuern sie die Eingabe!) oder die dazugehörige Komponente ohne Nachfrage als defekt ansehen. Med2 kann in solch einem Fall nur ein Element der entsprechenden Wertemenge fordern.

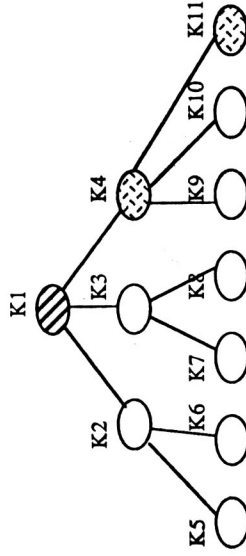


Abb. 4.4: Schematische Darstellung des im Beispiel verwendeten Hiqual-Modells.

2.4 Integration von Hiquall1 in Med2.

Hiquall1 ist eine Erweiterung des Hiqual-Systems. Die Komponenten des Modells besitzen zusätzliche Informationen (siehe Hiquall1 Kap.3: 2.2) und es werden Verbindungen von Symptomen und Diagnosen zu Komponenten des Hiqual-Modells erstellt. So kann das Expertensystem eigenständig Fragen an das Hiquall1-System stellen und Antworten ohne das Eingreifen des Benutzers verarbeiten. Dieser Verbindungskomplex wird vor der Initialphase erstellt, so daß für die Laufzeit keine Nachteile entstehen.

Hiquall1 berechnet die Folgen kleinerer Defekte zeitlicher oder ontologischer Art. Treten beide Defektarten gleichzeitig auf, so liefert Hiquall1 keine exakten Ergebnisse mehr.

Mehrere gleichzeitig auftretende Krankheiten kann Hiquall1 nicht erkennen.

2.4.1 Integration von Hiquall1 in den Arbeitsablauf.

In Med2 unterscheidet man zwischen der Datenvorverarbeitung, dem *database reasoning* und der eigentlichen diagnostischen Auswertung der Daten, dem *diagnostic reasoning*.

In der Phase des *database reasoning* werden Symptome erfaßt und einfache Symptominterpretationen durchgeführt. Ein Einsatz von Hiquall1 bringt in dieser Phase mehr Aufwand als Nutzen.

Während der diagnostischen Auswertung der Daten entstehen für Hiquall1 zwei Einsatzmöglichkeiten.

1) Unterstützung von Med2 bei der Erfassung von Symptomen.

a) Eine Komponente des Hiqual-Modells ist mit hoher Wahrscheinlichkeit defekt. Um diese Komponente zu überprüfen, werden alle Symptome, die auf die Komponente verweisen, untersucht. In Med2 gibt es diese Vorgehensweise nicht, da Med2 nicht komponentenorientiert arbeitet.

b) Med2 hat Schwierigkeiten Symptome zu finden, die es wert sind, untersucht zu werden. Eine oder mehrere Komponenten (K) im Modell sind als defekt gekennzeichnet. Algorithmus1 gibt Folgekomponenten (K1, K2, K3) an, die der Defekt der oben genannten Komponente beeinflussen kann. Zu untersuchen sind jetzt die auf diese Folgekomponenten verweisenden Symptome (S1, S2, S3 und S4).

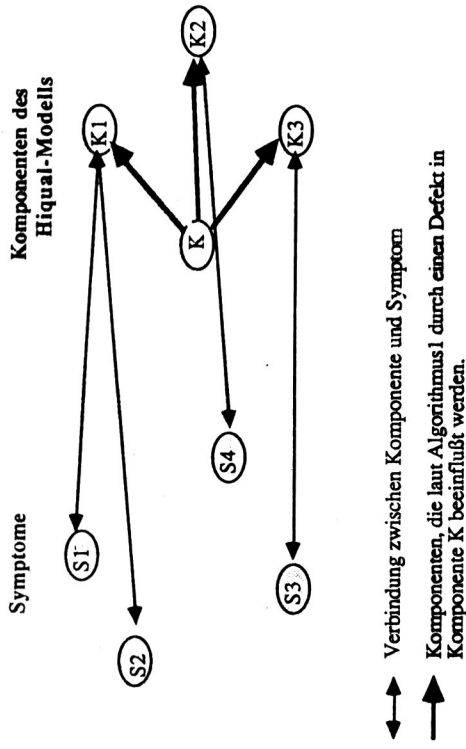


Abb. 4-5: Graphische Darstellung von 1 b).

Dadurch werden Symptome betrachtet, die im normalen Regelmechanismus von Med2 außer Acht gelassen oder erst später betrachtet werden.

2) Unterstützung der Verdachtsgenerierung.

Im Hiqual-Modell sind mehrere Komponenten als defekt gekennzeichnet. Algorithmus 2 findet (einen oder mehrere) mögliche Verursacher dieser Defekte. Die Wahrscheinlichkeit der Diagnosen, die diesen Verursacherkomponenten entsprechen, wird erhöht oder die Diagnose wird direkt dem WM zugeordnet.

Diagnosen, die bisher durch die Regelmenge außer Acht gelassen wurden, werden so miteinbezogen.

2.4.2 Integration von Hiquall1 in die Zeitkomponente.

Med2 stellt zeitbezogene Fragen über Dauer, Frequenz und Zeitpunkt der Symptome und berechnet aus diesen Rohdaten folgende Informationen:

- Dauer eines Symptoms aus Beginn und Ende
- zeitliche Beziehungen zwischen zwei Symptomen
- Umrechnungen bei Folgesitzungen

Da beide Systeme, Med2 und Hiquall qualitative Zeitangaben verarbeiten, muß nur die intervallbasierte Zeitarstellung von Hiquall mit der punktbasierten Zeitarstellung von Med2 kombiniert werden (Eine Möglichkeit zur Kombination: Die Zeitpunkte von Med2 werden in Hiquall als momentane Ereignisse angesehen). Wird Hiquall um die Möglichkeit erweitert, auch quantitative Zeitangaben zu verarbeiten, so kann der Datenaustausch auch auf dieser Ebene stattfinden. Wie können diese Zeit-Daten zur Diagnosefindung genutzt werden? In vielen Expertensystemen stellt man fest, daß zwar die Eingabe und Speicherung von Zeitangaben möglich ist, der Diagnoseprozeß aber abläuft, ohne auf zeitliche Zusammenhänge zu achten (außer ein paar Alibi-Regeln). Hier ist Med2 keine Ausnahme.

Das Hiquall-Modell ist eine solide Grundlage für die Integration der Zeit, da es alle Zeitbeziehungen des korrekten Modells repräsentiert und bei einer Simulation vorführen kann. Stellt der Benutzer bei einem Vergleich des korrekten Modells mit dem realen System zeitliche Inkorrektheiten fest, so kann er diese als Symptome eingeben. Die Transitivitätstabelle ermöglicht es Hiquall zu entscheiden, ob sich benachbarte Zeitrelationen verändern oder nicht (es ist genügend Zeit-Puffer dazwischen). Die Verwendung dieser Daten im Diagnoseprozeß bleibt aber weiterhin in Frage gestellt.

2.4.3 Integration von Hiquall in die Erklärungskomponente.

Über die in Hiquall aufgezählten Möglichkeiten hinaus kann der Benutzer in Hiquall dank der Algorithmen und der Verbindungen zu den Symptomen und Diagnosen weitergehende Erklärungen erhalten, die sich an dem Modell und seinen Komponenten orientieren.

- o Ein Symptom S liegt vor. Welche Komponenten können dadurch defekt sein?
 - o Eine Komponente ist defekt. Welche anderen Komponenten werden dadurch beeinflusst?
 - o Eine Diagnose D liegt vor. Mit welchen Komponenten ist sie verbunden?
- Wie stehen diese Komponenten zueinander bzw. wie stehen diese Komponenten zu anderen Komponenten?
- o Die Komponenten A und B sind defekt. Wieso wird nur die Komponente B verfolgt? (Der Defekt in A wird immer durch die Komponente B verursacht.)

Im Gegensatz zu Hiquall kann man sich nicht nur auf Komponenten, sondern auch auf Diagnosen und Symptome beziehen. Die Erweiterung durch Berechnungsalgorithmen ermöglicht es auch, indirekte Zusammenhänge zu erkennen.

Der Detaillierungsgrad der Erklärungen ist auf zwei Arten wählbar:

1. Wahl der Abstraktionsebene im Modell und damit auch der entsprechenden Symptome und Diagnosen.
2. Wahl der Genauigkeit der Relation:
 - a) zwischen Komponenten
 - b) zwischen Zuständen oder Ereignissen der Komponenten
 (Diese Relationen können mit Algorithmus1 nicht immer berechnet werden.)

2.4.4 Konsistenzüberprüfungen der Eingabe.

In Med2 wird jeder Frage eines Questionsets eine Wertemenge zugeordnet. Inkonsistenzen verschiedener Fragen entdeckt Med2 mit Regeln von Typ "contra". Feuert eine solche Regel, so wird der Benutzer aufgefordert, eine der Fakten, die zu der Inkonsistenz geführt haben, zurückzuziehen. Das Hiquall-Modell mit seinen Verbindungen zu Symptomen und Diagnosen erlaubt es, Konsistenzüberprüfungen anhand des Modells durchzuführen. Ein Vererbungsmechanismus gibt Informationen über korrekte und defekte Komponenten an Komponenten anderer Abstraktionsebenen weiter. Wird ein neues Symptom eingegeben, so wird die Konsistenz zwischen der entsprechenden Komponente und den Komponenten, die durch den Vererbungsmechanismus beeinflusst sind, überprüft. Tritt eine Inkonsistenz auf, so werden beide Symptome, das neue und das in der entsprechenden Begründung angegebene, überprüft (näheres siehe Kap.5).

2.4.5 Wertung.

Die Integration von Hiquall in das Expertensystem Med2 wirft in etwa dieselben Probleme auf, wie die Integration in das Ohrendiagnosesystem. Die Zuordnung der Symptome und Diagnosen zu den Komponenten des Modells ist ein wichtiges Faktum. Gelingt sie vollständig, so ist die Integration eine Verbesserung von Med2. Nicht nur assoziative Regeln, sondern auch die modellbasierte Arbeitsweise tragen zur Diagnosefindung bei.

Gelingt die Zuordnung nur unvollständig, so muß jeder Arbeitsschritt von Hiquall neu überdacht werden. Den Ideen von Hiquall wird weniger Wahrscheinlichkeit zugeordnet

oder Hiqua11 wird an einigen Stellen gar nicht mehr angewandt. Im Extremfall ist die Kombination Hiqua11 und Med2 schlechter als Med2 alleine.

2.5 Integration von Hiqua12 in das Expertensystem Med2.

Hiqua12 ist eine Erweiterung von Hiqua11 und ermöglicht die Auswirkungen defekter Komponenten ohne Konsistenzprobleme zu simulieren. Der Einsatz einer Simulation ist an die Bedingung geknüpft, daß die defekte Komponente vollständig beschrieben werden kann. Die Simulation erfordert großen Aufwand, sowohl für den Benutzer als auch für das System (siehe Hiqua12). In Med2 gibt es drei Fälle, in denen eine Simulation trotzdem sinnvoll sein kann.

2.5.1 Unterstützung der Differentialdiagnose.

Liegen zwei ähnliche Diagnosen vor, die auch in der Differentialdiagnose kaum Unterschiede aufweisen, so kann eventuell die Simulation entscheiden, welches die richtige ist. Die zu simulierenden Diagnosen müssen im Working-Memory unter den besten Diagnosen zu finden sein. Voraussetzung für die Simulation ist die Zuordnung beider Diagnosen zu den entsprechenden Komponenten des Modells und die vollständige Beschreibung der defekten Komponente (siehe Hiqua12).

2.5.2 Mehrere gleichzeitig auftretende Diagnosen.

Liegen mehrere Diagnosen vor, so kann eine vom Expertensystem ausgewählte, wahrscheinliche Diagnosekombination simuliert werden. Sich gegenseitig verändernde Symptome werden erkannt und eventuell bei der Auswahl weiterer Diagnosekombinationen berücksichtigt. Auch hier gilt es, die Anzahl der Simulationsläufe wegen des enormen Aufwandes möglichst gering zu halten.

2.5.3 Erklärungen für den Benutzer.

Die für den Benutzer verständlichsten Erklärungen sind die anhand eines Modells. Eine Simulation der Diagnose auf einer vom Benutzer gewählten Abstraktionsebene, ermöglicht es die Auswirkungen einer defekten Komponente zu verfolgen, und Modell und Wirklichkeit direkt zu vergleichen. Hiqua12 realisiert eine Erklärungskomponente mit diesen Vorteilen und dient zugleich als letzte Überprüfung der Diagnose.

2.5.4 Kann Hiqua12 die Explanationssets ersetzen?

Explanationssets fassen Symptomgruppen zusammen und gewichten sie entsprechend ihrer Erklärungsbedürftigkeit. Am Ende einer Sitzung führt Med2 mit den Explanationssets und den etablierten Enddiagnosen eine grobe Plausibilitätskontrolle auf höherem Abstraktionsniveau durch. Die Simulation einer Enddiagnose liefert genauere Ergebnisse und zusätzliche Informationen für die Erklärungskomponente. Sind die Bedingungen zur Durchführung einer Simulation erfüllt, so wird auf die Überprüfung durch die Explanationssets verzichtet (und die Simulation vorgezogen). Andernfalls muß eine Kontrolle durch die Explanationssets genügen.

Trotz der Integration von Hiqua12 und der Möglichkeit Simulationen durchzuführen, kann Hiqua12 die Explanationssets nicht vollständig ersetzen, da eine Simulation nicht immer durchgeführt werden kann. Für diese Fälle muß der Mechanismus der Explanationssets aufrecht erhalten werden.

2.5.5 Rücknahme von Ergebnissen der Simulation.

Simulationen in Med2 sind selten und kommen meist erst gegen Ende des Diagnoseprozesses vor. Deshalb kann man die Voraussetzungen und Ergebnisse der Simulation tabellarisch speichern. Liegen andere Symptombelagungen vor, so wird die entsprechende Simulation aus der Tabelle gestrichen und der ITMS-Algorithmus nimmt die weitergegebenen Ergebnisse und ihre Auswirkungen zurück.

2.5.6 Wertung.

Eine Simulation erfordert großen Aufwand für den Benutzer und das System und kann nur unter bestimmten Bedingungen angewandt werden. Die Ergebnisse der Simulation im Bereich der Differentialdiagnose oder der Erklärungskomponente können aber mit regelbasierten Mechanismen nicht erreicht werden. Deshalb ist die Integration einer Simulation wünschenswert.

3. Integration von Hiqua1, Hiqua11 und Hiqua12 in das Ohrendiagnosesystem.

3.1 Das Ohrendiagnosesystem.

Die Grundobjekte des Ohrendiagnosesystems sind Symptome, Erkrankungen und Untersuchungen. Die Erkrankungen werden als prädikatenlogische Formeln dargestellt, deren Variablen die Symptome sind. Eine Untersuchung ist eine Menge von Symptomvariablen, die nach der Durchführung der Untersuchung belegt werden können. Das Ohrendiagnosesystem unterscheidet zwischen der Initialphase, einer hypothesenfreien Anfangsphase mit Symptom-Symptom-Relationen, und der Finalphase, die aus der Anamnesephase und der Hypothesephase besteht. Im inexacten Bereich arbeitet man mit den Mechanismen Fokussierung und Abstraktion.

3.2 Allgemeines.

Hiqua1 in das Ohrendiagnosesystem zu integrieren ist nicht sinnvoll, da das Modell 'Ohr' für einen solchen Aufwand zu klein ist.

Im Folgenden übernimmt man die Mechanismen der Ohrendiagnose und denkt sich ein größeres Anwendungsmodell. Dabei ist zu beachten, daß größere Anwendungsgebiete eine Diagnosehierarchie erfordern.

3.2.1 Grobe Skizze für Diagnosehierarchien.

Der vollständige Hypothesenbaum wird vor der Konstruktion der Wissensbasis als eine für alle Diagnoseläufe gleiche Datenstruktur festgelegt. Während des Diagnoselaufs wird lediglich aufgrund des Beweisstatus der jeweils betrachteten Hypothesen dynamisch entschieden, welche Teilbäume des Hypothesenbaumes untersucht bzw. abgeschnitten werden.

Definition eines Hypothesenbaumes:

- (1) Ein *Hypothesenbaum* ist ein Paar $(\mathcal{H} < \mathcal{H})$ wobei \mathcal{H} eine Menge von Hypothesen und $<_{\mathcal{H}}$ eine Relation auf \mathcal{H} ist, die die Baumeigenschaft erfüllt.
- (2) Ist $H \in \mathcal{H}$ und $(H' | H <_{\mathcal{H}} H') = (H_1, \dots, H_K)$, so heißt H der *Cluster* der Hypothesen H_1, \dots, H_K (in Zeichen $H = (H_1, \dots, H_K)$).

- (3) Die Menge $B = \{H | \neg \exists H' : H < H'\}$ heißt die Menge der *Basis-hypothesen*. Sie sind identisch mit den Hypothesen des Grundmodells, entsprechend sind für sie die Mengen H^+ und H^- festgelegt.

- (4) $\mathcal{H} \setminus B$ ist die Menge der *Zwischenhypothesen*... Ihre Annahme- bzw.

Ablehnungsbereiche sind wie folgt definiert:

$$H = \{H_i\}_{i \in I} \implies H^+ := \bigcup_{i \in I} H_i^+, \quad H^- := \bigcap_{i \in I} H_i^-$$

Bemerkungen: Für eine Zwischenhypothese $H = \{H_i\}_{i \in I}$ gilt: In H^+ wird (mindestens) eine der Hypothesen H_i bestätigt. In H^- werden alle H_i abgelehnt.

Die Wurzel des Hypothesenbaumes ist nur aus systematischen Gründen von Interesse und entspricht der Hypothese: "Irgendein Fehler ist vorhanden".

Die hierarchische Anordnung der Hypothesen gestattet es zu Beginn die Hypothesen der obersten (allgemeinsten) Hierarchieebene zu betrachten. Kandidaten für die Verfeinerung auf der nächst tieferen Ebene sind nur die Hypothesen, die auf der ihnen entsprechenden Ebene bestätigt werden konnten.

Beim Aufbau der Diagnosehierarchie muß man darauf achten, daß Symptome, die schnell und einfach erfaßt werden können, auch auf höheren Ebenen (also gleich am Anfang) erfragt werden.

\implies aufangs mehr Grundsymptome,

später Symptome, die anamnestiche Untersuchungen erfordern

3.3 Integration von Hiqua1 in das Ohrendiagnosesystem

Mit Hilfe von Hiqua1 kann man das zu untersuchende Modell schnell und einfach aufbauen. In diesem Modell können alle ontologischen und zeitlichen Relationen zwischen den Komponenten bzw. ihren Aktionen und Ereignissen angegeben werden. Zusätzlich können in Hiqua1 einzelne ausgewählte Komponenten (und nur sie) auf einer tieferen Abstraktionsebene betrachtet werden.

Hiqua1 kann als Hintergrundwissen verwendet werden und ermöglicht Abfragen folgender Art:

- Welche zeitliche Zusammenhänge bestehen zwischen zwei bestimmten Aktionen oder Zuständen?

- Welche Aktionen müssen vor einem bestimmten Zustand (oder Ereignis) ablaufen?

- Aus welchen Teilen besteht eine Komponente?

- Welche ontologischen Zusammenhänge bestehen zwischen zwei bestimmten Komponenten (Aktionen oder Zuständen) ?

- Welche Komponenten sind kausal mit einer bestimmten Komponente (Ereignis oder Zustand) verbunden?

Diese Anfragen können auf allen Abstraktionsebenen gestellt werden. Eine automatische Integration der Antworten in das Expertensystem gibt es allerdings nicht. Sowohl die Idee für eine Frage, als auch die Verwertung der Antwort, müssen vom Benutzer selber eingebracht werden.

Mehrere gleichzeitig auftretende Krankheiten kann Hiqua1 nicht erkennen.

Die *Erklärungskomponente* ist eines der wenigen Teile des Ohrendiagnosesystems, welches direkt durch Hiqua1 unterstützt werden kann. Das Ohrendiagnosesystem stellt nur im inaktiven Bereich Begründungen zur Verfügung, warum ein Objekt auf der jeweiligen Abstraktionsstufe zu finden ist. Hiqua1 erlaubt dem Benutzer auch, alle Fragen nach dem korrekten Modell zu stellen (Aufbau und Zusammenhänge im Modell, in den einzelnen Komponenten und zwischen Ereignissen und Zuständen). Diese Ausgabe kann auch auf bestimmte Ausschnitte des Modells beschränkt werden.

Die Ein- Ausgabekomponente im Ohrendiagnosesystem ist noch nicht realisiert. Man kann aber davon ausgehen, daß das Hiqua1-Modell dank seines Hintergrundwissens Konsistenzüberprüfungen der Eingaben durchführen kann. (Für weitere Informationen siehe Integration von Hiqua1 in Med2, Kapitel 4: 2.3.3)

3.4 Integration von Hiqua11 in das Ohrendiagnosesystem.

In Hiqua11 bleiben alle Möglichkeiten von Hiqua1 verfügbar. Zusätzlich können einfache Fragestellungen über defekte Komponenten und ihre Auswirkungen beantwortet und das Hiqua1-Modell in ein Expertensystem integriert werden. Der Informationsgehalt der einzelnen Komponenten wird durch folgende Angaben erweitert:

- Liegt eine Komponente vor, die leicht testbar ist?

- Informationen über Verbindungen zu Symptomen oder Diagnosen.

Durch die Integration des Hiqua1-Modells in ein Expertensystem können Fragen automatisch gestellt und beantwortet werden, ohne daß der Benutzer eingreifen muß. Die zur Integration notwendigen Verbindungen müssen noch vor der Initialphase erstellt werden. So entstehen durch den Aufbau des Komplexes: Symptome - Komponenten - Diagnosen für die Laufzeit keine Nachteile.

Sind sowohl zeitliche als auch ontologische Relationen defekt, oder treten mehrere Krankheiten gleichzeitig auf, so genügt Hiqua1 den Anforderungen nicht mehr. Der Einbau und die Simulation auch nur kleinerer Defekte sind erst in Hiqua2 möglich.

3.4.1 Integration von Hiqua1 in die Fokussierung der Hypothesen.

In der Fokussierungsphase werden alle interessanten Hypothesen extrahiert. Zusätzlich zu den bisherigen Punkten bei der Fokussierung betrachte man alle Hypothesen, auf die von defekten Komponenten aus verwiesen wird und nehme sie in die Liste der interessanten Hypothesen auf. Eine Variante besteht darin, die Hypothesen erst dann aufzunehmen, wenn die Verbindung zwischen der Hypothese und der Komponenten eine bestimmte Wahrscheinlichkeit überschritten hat.

Krankheiten, die schon bewiesen oder widerlegt sind, werden nicht mehr betrachtet. So werden mögliche Krankheiten von einer anderen Perspektive her eingebracht.

3.4.2 Integration von Hiqua1 in die Fokussierung der Untersuchung.

Die Fokussierung der Untersuchung behandelt die Frage, welche Untersuchungen von Interesse sind. Möglicherweise defekte Komponenten können Anstöße an Symptome geben, durch die die entsprechende Untersuchung in die Liste der interessanten Untersuchungen eingefügt oder ihre Wahrscheinlichkeit erhöht wird.

3.4.3 Integration von Hiqua1 in die Erklärungskomponente.

Hiqua1 erweitert die Erklärungskomponente um folgende Fragen:

- Welche Komponenten können durch eine bestimmte defekte Komponente verändert werden?

In diesem Zusammenhang kann man noch einen Schritt weiter gehen und fragen:

- Ein Symptom S liegt vor: Welche Komponenten könnten dadurch defekt sein? Und zwar direkt (die Komponenten, auf die das Symptom verweist) und indirekt (die Komponenten, die mit den Komponenten K verbunden sind). Auch die Umgebung einer Komponenten im Modell kann man ausgeben.

- Eine Diagnose D liegt vor. Man kann die Verbindungen zwischen der Diagnose und den entsprechenden Komponenten beschreiben. Zusätzlich kann die Umgebung der Komponenten anhand des Modells beschrieben werden. (Wie stehen die Komponenten zueinander? Welche Unterkomponenten besitzen sie? Welche Relationen liegen zwischen ihnen vor?)

Im Gegensatz zu Hiqua1 kann man hier auch Symptome oder Diagnosen und nicht nur Komponenten eingeben, über die man genauere Informationen einholen will.

3.4.4 Integration von Hiqua1 in die Zeitkomponente.

In Hiqua1, Hiqua2 und auch Hiqua3 können nur qualitative Relationen zwischen Intervallen spezifiziert werden. Eine Erweiterung um quantitative Informationen über die absolute und relative Größe der Intervalle ist möglich und sollte je nach Themengebiet eingebaut werden. In Hiqua1 werden zeitliche Relationen zwischen Ereignissen und Zuständen gespeichert.

Im Ohrendiagnosesystem werden die Informationsvektoren mit Zeitpunkten indiziert. Zusätzlich gibt es zeitmodale Operationen wie exists und persists.

Um die Systeme (Hiqua1 und Ohrendiagnose) gemeinsam anwenden zu können, müssen beide erweitert werden:

Im Ohrendiagnosesystem sollen Fragen nach zeitlichen Zusammenhängen und Zeitpunkten zugelassen werden. Die Ergebnisse solcher Anfragen muß man verarbeiten können. Zeitliche Zusammenhänge sollten als Symptome implementiert zu Variablen der Diagnosen werden.

Sind diese Voraussetzungen gegeben, so kann man mit Hiqua1 nicht nur korrekte zeitliche Zusammenhänge abfragen, sondern auch die Folgen inkorrekt zeitlicher Zusammenhänge mit Hilfe der Transitivitätstabelle berechnen.

3.4.5 Hiquall1 und Recovery.

Werden Änderungen an dem Hiquall-Modell oder an dem Verbindungskomplex: Symptome - Komponenten - Diagnosen vorgenommen, so muß die gesamte Ohrendiagnose neu gestartet werden. Dies entspricht im Expertensystem einer Änderung der Symptom-Relation 'Verfeinerung' oder einer Änderung der Diagnosehierarchie. Auch in diesen Fällen muß das System neu gestartet werden.

Der Recovery-Mechanismus des Ohrendiagnosesystems unterscheidet zwei Situationen:

- (1) Die zu löschende Belegung wurde in der aktuellen Anamnesephase vorgenommen. Wird ein Symptom zurückgezogen und hat eine Verbindung zu einer Komponente im Hiquall-Modell schon gefeuert, so muß diese Feue rung zurückgenommen werden. d.h. - Die Wahrscheinlichkeit der Komponente wird um die Wahrscheinlichkeit der Verbindung verringert.

- Die Verbindung kann wieder feuern.

Alle von dieser Komponenten abhängigen Knoten werden entsprechend geändert. Dies ist möglich, da in einer Komponente nicht nur die Wahrscheinlichkeit gespeichert ist, sondern auch die dazugehörige Begründung.

- (2) Die zu löschende Belegung wurde in einer früheren Anamnesephase vorgenommen. Alle abgeleiteten Informationen (über Hypothesen und Vorschläge) werden gelöscht. Im Hiquall-Modell werden alle zusätzlichen Informationen in den Komponenten gelöscht und alle Verbindungen in den ursprünglichen Zustand gebracht (d.h. sie können wieder feuern). Dann kann der Prozeß mit allen bekannten Symptomen neu starten.

3.4.6. Hiquall1 und unvollständiges und unsicheres Wissen.

Unsicheres Wissen kann in Hiquall1 durch die Angabe von Wahrscheinlichkeiten bei den Verbindungen und den Komponenten selber dargestellt werden. Unvollständiges Wissen kann durch zwei verschiedene Detaillierungsstufen angegeben werden.

- (1) Die Verbindungen können von verschiedenen Ebenen des Symptombaums zu verschiedenen Ebenen des Hiquall-Modells laufen. Dasselbe gilt für die Diagnosen.
- (2) Die Verbindungen entsprechen Informationen zwischen den Symptomen und den Komponenten. Sie können aber auch genauere Informationen über Teile der Komponenten enthalten, wie z.B. Ereignisse oder Zustände.

3.4.7 Schlußwort.

In Hiquall1 muß man immer darauf achten, daß es (besonders im medizinischen Bereich) auch Symptome gibt, die keine Verbindungen zu Komponenten haben, gleichzeitig aber pathophysiologisch belegt sind. Sowohl bei der Erklärungskomponente als auch bei der Fokussierung können solche Symptome zu einem "schiefen" Ergebnis führen.

3.5 Integration von Hiquall2 in das Ohrendiagnosesystem.

Hiquall2 ist eine Erweiterung des Hiquall1-Systems. Es ermöglicht defekte Komponenten und Zeitrelationen zu integrieren, ohne daß bei der Simulation Konsistenzprobleme auftreten.

Die Simulation soll nur dann eingesetzt werden, wenn es gute Verbindungen zwischen Symptomen, Komponenten bzw. ihren Zuständen und Ereignissen und den Diagnosen gibt. Sonst lohnt das Ergebnis den Aufwand nicht. Es dürfen nicht zu viele Simulationen durchgeführt werden, weil jede Simulation viel Aufwand auch für den Benutzer verursacht.

Beispiel: Für jede defekte Komponente muß ein neuer Typ erstellt werden, alle Konsistenzüberprüfungen, die irgendwie mit dieser Komponente in Beziehung stehen, müssen unterbunden werden (letzteres ist evtl. automatisch möglich).

3.5.1 Integration von Hiquall2 in die einzelnen Arbeitsschritte.

3.5.1.1 Initialphase

Die Integration von Hiquall oder einer seiner Nachfolger in die Initialphase ist zwar theoretisch möglich, bringt aber mehr Aufwand als Vorteile. Das Ohrendiagnosesystem besitzt sehr viele Symptom-Symptom Relationen, die weitere Hilfestellungen überflüssig machen.

3.5.1.2. Finalphase.

Die Finalphase besteht aus der Anamnese- und der Hypothese-Phase. Hiqua2 wird in die Hypothese-Phase eingebaut und zwar nach der Abstraktion der Krankheiten und noch vor der Abstraktion der Untersuchungen. Kann keine Diagnose gestellt werden, ist es besser eine Simulation mit Hiqua2 einzuschleichen, noch ehe man weitere Tests startet. Die Simulation ist unter folgenden Bedingungen möglich:

- Es liegen mehrere hochwahrscheinliche Diagnosen vor, von denen man annimmt, daß eine von ihnen die richtige Diagnose ist (=Fertig-Liste).
- Diese Diagnosen lassen sich alle mit einer oder mehreren Komponenten des Hiqua2-Systems in Beziehung bringen.

Die Ergebnisse der einzelnen Simulationen kann man entweder getrennt mit dem realen System vergleichen oder man vergleicht erst die verschiedenen Simulationsergebnisse untereinander und geht dann gezielt auf die Unterschiede ein. Diese Unterschiede können durch schon belegte Symptome geklärt oder in der wirklichen Maschine überprüft werden.

Ein großer Vorteil dieser Integration von Hiqua2 ist, daß eine Betrachtungsweise aus einer ganz anderen Richtung mit einbezogen wird. So kann man Unterschiede zwischen zwei Diagnosen aufzeigen, die durch die Formeln nicht zum Ausdruck kommen.

3.5.2 Integration von Hiqua2 in die Erklärungskomponente.

Um Ergebnisse oder Zusammenhänge verständlicher zu gestalten, werden dem Benutzer die Resultate der Simulation zur Verfügung gestellt. Das Diagnoseergebnis wird damit untermauert.

Auch während des Ablaufs des Expertensystems kann man dem Benutzer auf Anfrage die Simulation bestimmter Diagnosen, sowie Unterschiede zwischen zwei verschiedenen Simulationen ausgeben. Allerdings muß dem Benutzer der hohe Aufwand einer Simulation bewußt sein.

3.5.3 Integration von Hiqua2 in die Zeitkomponente.

Hier treten dieselben Probleme auf, wie sie in Hiqua1 schon angedeutet wurden. Das Ohrendiagnosesystem verfügt nur über quantitative Zeitinformationen, während Hiqua

nur qualitative Informationen angeben kann. Erweitert man aber beide Systeme so weit, daß sie über eine gemeinsame Schnittstelle verfügen, so kann man Hiqua2 integrieren. Vorteilhaft ist eine solche Integration für die Simulation inkorrektur zeitlicher Relationen, deren Auswirkungen so genauer betrachtet werden können.

Besonders Berechnungen in Hiqua1, die zu umfangreich wurden, oder wegen der gleichzeitig vorliegenden ontologischen Defekte nicht durchgeführt werden konnten, werden durch die Simulation gelöst. Für das momentane Ohrendiagnosesystem sind die Zeitrelationen des Hiqua2-Modells mit all' ihren zusätzlichen Aussagen bedeutungslos.

3.5.4 Hiqua2 und Recovery.

Das Modell von Hiqua2 ist von Recovery-Maßnahmen nicht betroffen, da für jede Simulation das Modell 'neu' aufgebaut wird. Man nimmt das korrekte Modell, setzt die defekten Komponenten ein und schaltet die entsprechenden Konsistenzbedingungen aus. Die Informationen, die benutzt werden, sind entweder von vorne herein gegeben (das korrekte Modell und die Konsistenzbedingungen) oder entsprechen dem aktuellen Informationsstand (defekte Komponente). Ändert sich der Informationsstand während der Simulation, so muß diese erneut durchgeführt werden.

Die Ergebnisse der Simulation werden aber auch weiterverwendet. Wird eine Vorbedingung der Simulation zurückgenommen, so müssen auch die Ergebnisse der Simulation zurückgezogen werden.

Das Ohrendiagnosesystem unterscheidet zwischen zwei Fällen:

1. Die zu löschende Belegung wurde in der aktuellen Anamnese-Phase vorgenommen. In diesem Fall werden die Ergebnisse der Simulation, sofern sie von der gelöschten Belegung abhängig sind, zurückgezogen. Da Simulationen selten sind, kann man sich die Simulationen bzgl. einer Anamnese-Phase merken.
2. Die zu löschende Belegung wurde in einer früheren Anamnese-Phase vorgenommen. Alle abgeleiteten Informationen über Hypothesen und Vorschläge werden gelöscht. Dadurch werden automatisch auch die Ergebnisse aller bisherigen Simulationen zurückgezogen.

3.5.5 Unterstützung durch Hiqual2 beim Versuch mehrere vorliegende Krankheiten zu finden.

Liegen mehrere Krankheiten vor, so muß das Expertensystem eine Menge wahrscheinlicher Kombinationen finden. Die Überprüfung dieser Krankheitskombinationen kann Hiqual2 übernehmen, falls alle Krankheiten bestimmten Komponenten zugeordnet (in technischen Systemen meist gegeben) und zusätzlich diese 'defekten' Komponenten genau beschrieben werden können. Der Vergleich der Simulationen in Hiqual2 bietet eine gute Unterstützung bei der Suche nach mehreren vorliegenden Krankheiten.

3.5.6 Allgemeine Bemerkungen.

Die Arbeitsfähigkeit von Hiqual2 ist in hohem Maße davon abhängig, inwieweit sich die einzelnen Symptome bestimmten Komponenten zuordnen lassen. Bei größeren Betrachtungen kommt es nicht nur darauf an, daß sich einzelne Symptome zuweisen lassen, sondern es müssen *alle* pathophysiologischen Symptome Verbindungen zum Modell herstellen können. Dies ist besonders wichtig bei einer Simulation innerhalb der Erklärungskomponente und beim Vergleich verschiedener Diagnosen durch die Simulation (Bsp.: Symptom = Fieber).

Als Abhilfe kann man versuchen diesen Simulationen eine geringere Wahrscheinlichkeit zuzuordnen oder in das Ergebnis der Simulation das pathophysiologische Symptom 'von Hand' einzubauen. Auf jeden Fall muß man sich vergegenwärtigen, daß es hier ein Symptom gibt, das bei der Simulation nicht berücksichtigt wurde, aber pathophysiologisch belegt ist.

4. Vergleich der Integration von Hiqual, Hiqual1 und Hiqual2 in die drei Expertensysteme Med2, Abel und das Ohrendiagnosesystem.

Eine Integration von Hiqual in das Expertensystem Abel erhöht die Flexibilität bezüglich des Aufbaus der Hierarchien. Hiqual1 und Hiqual2 vergrößern die Komplexität von Abel. Anwendungsmöglichkeiten für diese Systeme finden sich aber nur selten, da Abel schon ein modellbasiertes Expertensystem ist. Die Integration von Hiqual (oder einer ihrer Modifikationen) in Abel ist schon auf Grund der enormen Komplexität des Expertensystems, das bisher nicht vollständig implementiert werden konnte, unmöglich. Deshalb scheidet Abel in diesem Vergleich aus.

Bleiben noch die assoziativen Diagnostik-Expertensysteme Med2 und das Ohrendiagnosesystem. Beide Systeme gewinnen durch die Integration von Hiqual oder einer seiner Modifikationen und vereinen assoziative und modellbasierte Ideen in einem System. Problematisch ist für beide Systeme die Zuordnung der Symptome und Diagnosen zu den Komponenten des Modells. Gelingt sie, so ist das Gesamtsystem ein Erfolg, andernfalls bringt die Integration von Hiqual (bzw. ihrer Modifikationen) keinen Vorteil. Zur Integration bevorzugen beide Systeme das Gesamtkonzept Hiqual2 vor Hiqual1 und Hiqual.

Die Integration von Hiqual2 verkleinert die Nachteile des Ohrendiagnosesystems gegenüber Med2 nicht. Das Gegenteil ist eher der Fall. Die Kombination Hiqual2 und Med2 wird somit zum Favoriten.

Gründe:

- Möglichkeit des Einsatzes in größeren Anwendungsgebieten.

Das Ohrendiagnosesystem wurde für ein kleines Anwendungsgebiet erstellt. Ein Ausbau des Systems (z.B. eine Diagnosehierarchie) ist begrenzt möglich. Med2 dagegen bietet neben der Diagnosehierarchie auch eine Diagnosehierarchie und ein Diagnosenetzwerk. Zusätzlich werden auch Fragen und Regeln in Gruppen zusammengefaßt.

- Verarbeitung von Zeitaspekten.

Hiqual2 und Med2 arbeiten mit qualitativen Zeitrelationen, während das Ohrendiagnosesystem nur quantitative Zeitdaten verarbeiten kann.

- Repräsentation der Anatomie.

Im Ohrendiagnosesystem gibt es keine Modelle zur Darstellung des Ohrs. Med2 bietet eine Lokalisationshierarchie mit dazugehörigen Lokalisationsfragen an und erleichtert dadurch die Integration von Hiqua2.

- Zurückziehen von Fakten und Regeln.

Der Recovery-Mechanismus der Ohrendiagnose und der ITMS-Algorithmus können beide auch nach der Integration von Hiqua2 weiter angewandt werden.

Die Tatsache, daß Med2 ein Diagnoseschell ist und auch in technischen Bereichen angewandt wurde, erleichtert die Integration von Hiqua2 und das Erstellen eines ähnlichen Expertensystems für Bearbeitungszentren.

Zusammenfassung: Integration von Hiqua2 in Expertensysteme der assoziativen Diagnostik.

Die Integration von Hiqua2 in ein Expertensystem der assoziativen Diagnostik im medizinischen Bereich ist nur bedingt zu empfehlen. Das Problem ist die Zuordnung der Symptome und Diagnosen zu den Komponenten. Kann eine Diagnose nicht einer defekten Komponenten (Organ) zugeordnet werden, so ist eine Simulation dieser Diagnose unmöglich. Die Zuordnung der Symptome zu den Komponenten muß vollständig realisiert werden. Ist das nicht der Fall (Symptom = Fieber), so führen die Simulation in Hiqua2 und die Algorithmen von Hiqua1 zu "schiefen" Ergebnissen, da sie nur von einer Teilmenge der bekannten Daten ausgehen.

In technischen Anwendungsbereichen ist das Zuordnungsproblem kleiner. Die Therapie-technischer Systeme besteht zum Beispiel aus dem Auswechseln oder Säubern bestimmter Komponenten. Deshalb ist auch die Diagnose fast immer komponentenbezogen. Die Möglichkeit in technischen Systemen Teile, die vermutlich defekt sind, testweise auszutauschen, ist einerseits komponentenorientiert und andererseits eine in medizinischen Anwendungen nicht bekannte Untersuchungsmöglichkeit, deren Anwendung das Hiqua-Modell unterstützt. Der Grad der Zusammengehörigkeit von Symptomen, Diagnosen und Komponenten ist in technischen Systemen bedeutend höher, wodurch die Integration von Hiqua2 in ein Expertensystem der assoziativen Diagnostik im technischen Bereich durchaus befürwortet werden kann.

KAPITEL 5 : Konzeptentwicklung des neuen Expertensystems zur Fehlerdiagnose in technischen Systemen.

Ausgehend von den Expertensystemen Med2, Abel und der Ohrendiagnose sowie der Tiefenmodellierungssprache Hiqua1 und ihrer Modifikationen Hiqua1 und Hiqua2 soll ein Diagnosesystem erstellt werden, das möglichst viele Vorteile dieser Systeme vereint und für technische Anwendungen, wie zum Beispiel der Fehlerdiagnose von CNC-Bearbeitungszentren, besonders gut geeignet ist.

Zunächst wird eine Liste mit den Anforderungen an ein solches Expertensystem erstellt und untersucht, inwieweit die bisherigen Konzepte diesen Anforderungen entsprechen. Es zeigt sich, daß eine Verknüpfung kausaler und assoziativer Systeme die Bedingungen, die an das Diagnosesystem gestellt werden, am ehesten erfüllt. Für das assoziative System wird Med2, für das kausale System Hiqua2 eingesetzt. Anschließend werden verschiedene Ansatzpunkte zur Integration von Hiqua2 und dem Abstraktionsmechanismus des Ohrendiagnosesystems in das Expertensystem Med2 untersucht.

Wegen der Bedeutung der Simulation in Hiqua2 für das neue Expertensystem werden im vierten Abschnitt eine verbesserte Ausnutzung der Ergebnisse der Simulation, sowie eine Verringerung der Voraussetzungen zur Anwendung der Simulation diskutiert. Themen der beiden letzten Abschnitte sind die funktionale Sicht und die Architektur des neuen Expertensystems.

1. Anforderungen an ein Expertensystem zur Diagnosefindung in technischen Systemen.

- 1.) Die vielen möglichen Fehlerwirkungen und -ursachen, die durch das Zusammenwirken der zahlreichen unterschiedlichen Komponenten (Hydraulik, Mechanik, Elektronik,...) z.B. eines CNC-Bearbeitungszentrums entstehen, fordern einen modularen Aufbau der Wissensbasis.
- 2.) Viele Fehlerwirkungen sind durch einen zeitlichen Verlauf charakterisiert und fordern eine explizite "Zeitrepräsentation" quantitativer und qualitativer Art.
- 3.) Der Anspruch der wirtschaftlichen Einsetzbarkeit fordert ein robustes System, das auch in Randbereichen seiner Kompetenz einsetzbar ist, sowie eine effiziente Problemlösungsstrategie.

2.2 Möglichkeiten zur Verknüpfung assoziativer und kausaler Systeme.

Hauptproblem ist die Konsistenz zwischen den verschiedenen Wissensarten, deren Wissen sich zwar überlappt, aber nicht ineinander überführbar ist. Der Unterschied zwischen assoziativem und kausalem Wissen besteht darin, daß in heuristischen Systemen die probabilistische Bewertung eine große Rolle spielt, während kausale Systeme nur überprüfen, ob Diagnosen mit den beobachteten Symptomen konsistent sind. Der diagnostische Mittelbau des kausalen Wissens ist wesentlich detaillierter als bei assoziativem Wissen.

Nach Puppe [1986a] gibt es drei Möglichkeiten zur Integration kausaler und assoziativer Module. Die Lösungsmöglichkeiten A) und B) versuchen die Konsistenzprobleme durch Erweiterung oder Einschränkung der Wissensarten zu lösen, während C) die Wissensarten nebeneinander bestehen läßt.

- A) Kausale Modelle werden durch probabilistische Bewertungen erweitert. Der dadurch entstehende Wissenserwerb, der schon für heuristische Systeme enorm ist, vergrößert sich durch die Bewertung zusätzlicher Details kausaler Modelle erheblich.
- B) Eine andere Möglichkeit bietet die Kompilierung des kausalen Wissens in heuristische Regeln, um die Effizienz zu steigern. Die entstehenden Regeln enthalten keine probabilistischen Bewertungen mehr, wird ihre Nützlichkeit erheblich eingeschränkt.
- C) Beim dritten Ansatz zur Integration stellt man beide Wissensarten nebeneinander und definiert Kriterien dafür, wann welches Modul zum Einsatz kommt.

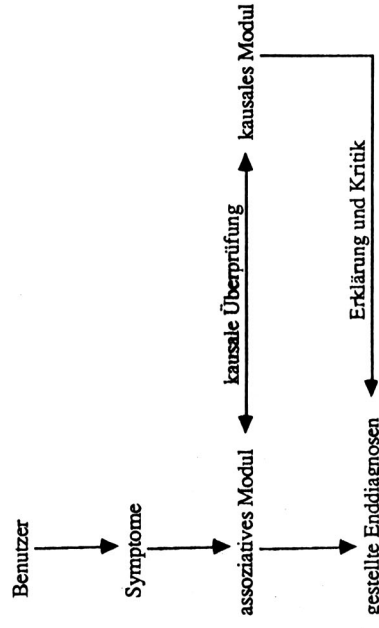


Abb. 5.1: Grundidee für ein mögliches Zusammenspiel assoziativer und kausaler Diagnostik.

4.) Ein Umstellen des Diagnoseprogramms für andere Maschinen darf nicht zur Erstellung eines völlig neuen Systems führen.

5.) Konsistenzüberprüfungen sollen fehlerhafte Eingaben des Benutzers oder falsche Messungen der Sensoren weitgehend ausschließen.

6.) Ein Zurückziehen der Eingaben soll zu jedem Zeitpunkt möglich sein.

2. Abwägen bisheriger Konzepte.

2.1 Soll das neue Diagnosesystem assoziativer oder kausaler Art sein?

Das Ohrendiagnosesystem und Med2 sind assoziative Systeme, während Abel kausaler Art ist. In assoziativen Systemen wird die Diagnose nach ihrer akkumulierten Evidenz bewertet. Sie haben ein breites Anwendungsspektrum, sind sehr effizient, versagen aber bei der Lösung komplexer Probleme und liefern keine parternmodellierten Erklärungen. In kausalen Systemen werden die Diagnosen nach der Anzahl der erklärten Symptome bewertet. Die Lösung komplexer Probleme und die Erklärung der Diagnosen bereiten modellbasierten Systemen keine Schwierigkeiten. Allerdings ist ihre Effizienz nur sehr gering und das Anwendungsspektrum begrenzt.

Kurz zusammengefaßt:

rein kausal --> gute Erklärungskomponente, komplexe Probleme können gelöst werden

rein assoziativ --> hohe Effizienz, breites Anwendungsspektrum

Im folgenden Kapitel wird versucht, beide Systemarten so zu verknüpfen, daß das neue System die Vorteile der beiden alten in sich vereint.

Wegen seiner Effizienz und Flexibilität sollte zunächst das assoziative Modul die diagnostischen Probleme verarbeiten. Das kausale Modul repräsentiert das Hintergrundwissen und dient zur Rechtfertigung und Erklärung der Endergebnisse, sowie der Überprüfung verdächtiger Diagnosen und dem Erkennen neu auftretender Diagnosen.

Puppe betrachtet nicht nur die Integration kausaler und assoziativer Module, sondern auch statistischer Module und Falldatenbanken (siehe Kapitel 6, siehe [Puppe1986a]).

Im Folgenden soll der Ansatz C) näher betrachtet werden.

2.3 Welche bisherigen Systeme bzw. Systemteile können für das assoziative und das kausale Modell eingesetzt werden?

Im neuen System steht kausales und assoziatives Wissen nebeneinander zur Verfügung. Für das assoziative Modul wird Med2 eingesetzt, das funktionale kausale Wissen repräsentiert Hiqua2. Für das pathophysiologische kausale Wissen wird gegebenenfalls ein neues Modell erstellt, das in schwierigen Fällen Teile des Systems simulieren kann (siehe Ausblick). Zur Verknüpfung der Informationen, die aus dem assoziativen und dem kausalen Modul kommen, wird ein Abstraktionsmechanismus ähnlich dem im Ohrendiagnosesystem angestrebt. Die Kontrollstruktur, die angibt, wann welches Modul zum Einsatz kommt, muß neu erstellt werden. Med2 behält die Kontrolle und stößt das kausale Modell in bestimmten Fällen an.

Med2 ist ein sehr gutes System zur assoziativen Diagnostik, das in technischen Bereichen gut anwendbar ist, hohe Effizienz aufweist und die Integration des kausalen Systems Hiqua2 zuläßt (siehe Kapitel 4: Abschnitt 4.). Strukturierte Regeln, Frames, WM-Konzept mit hypothetisch deduktiver Vorgehensweise und Differentialdiagnostik und nicht-monotones Argumentieren sind die wichtigsten Konzepte von Med2. Dank seines hierarchischen Aufbaus kann Med2 auch in größeren Anwendungsbereichen effizient arbeiten.

Kausales Wissen kann durch ein funktionales oder ein pathophysiologisches Modell repräsentiert werden. Das funktionale Modell ist typisch für den technischen Bereich und geht von dem normalen Funktionieren des zu diagnostizierenden Systems aus. Die Diagnosen entsprechen den Veränderungen des Modells, die zu den beobachteten Symptomen führen (Bsp. Hiqua2, Davis [Davis 1984]). Das pathophysiologische Modell gibt eine Folge von Zuständen an, die zu dem jeweiligen Symptom führen. Für die unerklärten Zustände werden externe Ursachen (Diagnosen) gesucht (Bsp. PSM in ABEL). Zur Darstellung der korrekten Funktionsweise des Modells entscheiden wir uns

in diesem System für Hiqua2. Das System von Davis ist zu einseitig auf Hardware-Diagnostik zugeschnitten und technische Systeme, wie zum Beispiel komplexe CNC-Maschinen, können den zahlreichen Restriktionsanforderungen nicht genügen. Hiqua2 dagegen ist ein Tiefenmodellierungssystem zur Analyse korrekter technischer Systeme allgemeiner Art. Es dient zur Repräsentation hierarchisch aufgebauten Wissens (entsprechend der Hierarchien in Abel) und kann als Hintergrundwissen je nach Struktur und Ablauf des umgebenden Expertensystems verwendet werden. Ein Beispiel ist die Unterstützung beim Aufbau eines pathophysiologischen Modells, das von den Zuständen und Ereignissen des Hiqua-Modells ausgeht. Der schnelle und einfache Aufbau des betrachteten Modells mit Hiqua2 macht das Hintergrundwissen flexibel und ermöglicht die Modellierung ähnlicher Maschinentypen.

In Hiqua2 werden die Zeitbeziehungen zwischen Zuständen und Ereignissen so repräsentiert, daß sie eine gute Grundlage für die Zeitkomponente bieten.

3. Integration von Hiqua12 und des Abstraktionsmechanismus der Ohrendiagnose in das Expertensystem Med2.

3.1 Ansatzpunkte zur Integration von Hiqua12 in das Expertensystem Med2.

3.1.1 Hiqua12 ersetzt die Lokalisationshierarchie.

Zur Repräsentation der Anatomie gibt es in Med2 Lokalisationsfragen, eine Lokalisationshierarchie und die entsprechenden Regeln. Die daraus entstehenden Vorteile (Möglichkeit der graphischen Lokalisationseingabe, Angabe der Wertemenge einer Lokalisationsfrage als Zeiger auf ein Element der Lokalisationshierarchie) bleiben auch beim Ersetzen der Lokalisationshierarchie durch das Modell von Hiqua12 bestehen (siehe Kapitel.4: 2.3.1). Das Questionset Lokalisationsfragen und die dazugehörigen Regeln werden entsprechend dem Modell von Hiqua12 erweitert. Das System Hiqua12 steht jetzt nicht mehr neben Med2, sondern ist in das Expertensystem eingebaut, wenn auch noch nicht alle Funktionen von Hiqua12 ausgenutzt werden.

3.1.2 Wozu benutzt Abel sein Hintergrundwissen?

Abel ist ein modellbasiertes Expertensystem. Sein Hintergrundwissen ist umfangreicher als das Hiqua1-Modell und umfaßt anatomisches, ätiologisches und physiologisches Wissen. Dieses Hintergrundwissen benutzt Abel, um die Diagnosen bezüglich Zeit, Ätiologie, Anatomie und Physiologie zu ordnen. Dieser Aspekt kommt für das neue Expertensystem nicht in Frage, da schon die Diagnosehierarchien, -heterarchien und das Diagnosenetzwerk in Med2 eine sehr gute Ordnung der Diagnosen erlauben. Weiterhin benutzt Abel sein Hintergrundwissen zum Aufbau des PSM's und der Diagnosehüllen. Diese Anwendung wird nur interessant, wenn später dem neuen Expertensystem außer dem funktionalen noch ein pathophysiologisches Modell hinzugefügt wird (s. Kapitel 6). Zusammenfassend ist zu sagen, daß Abel keine Anregungen für eine Integration von Hiqua12 in Med2 liefert.

3.1.3 Neue Ideen durch GTD.

GTD (= Generate, Test and Debug) [Davis et al. 87] ist ein Problemlösungsparadigma, das assoziative Regeln mit kausalen Modellen verknüpft und besonders für Planungs- und Interpretationsaufgaben geeignet ist. Der Generator benutzt assoziative Regeln, um eine glaubhafte Hypothese zu erstellen. Der Tester untersucht die Hypothese mit kausalen Modellen. Falls der Test mißlingt, erzeugt er eine genaue Charakterisierung der Widersprüche. Der Debugger verändert die Hypothesen, indem er mit gebietsunabhängigem Debugger-Wissen über das kausale Modell 'nachdenkt'.

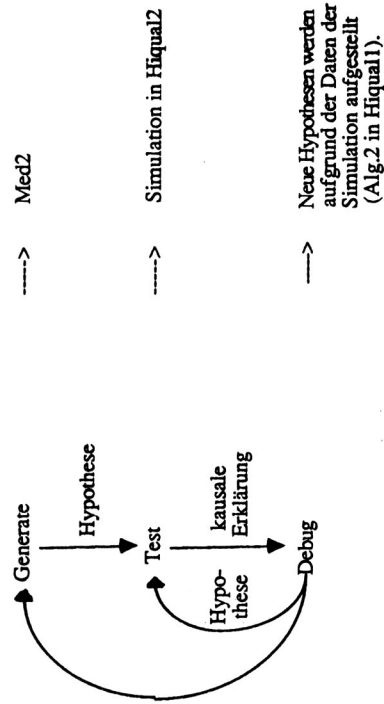


Abb. 5.2: Kontrollfluß des GTD-Paradigma und die entsprechenden Mechanismen des neuen Expertensystems.

Das GTD-Paradigma ist eigentlich für Planungs- und Interpretationsaufgaben erstellt worden. Überträgt man diese Idee auf das neue Expertensystem so entspricht Med2 dem Generator und erzeugt mit assoziativen Regeln eine Menge möglicher Hypothesen (nicht nur eine, da Med2 das Prinzip des Kurzzeitgedächtnisses des Menschen simuliert). Eine Simulation in Hiqua12 realisiert die Testphase mit kausalen Modellen. Der Debugger entspricht dem Algorithmus2 in Hiqua11, der Vorschläge für neue Hypothesen aufgrund der Daten der vorangegangenen Simulation aufstellt.

Zu beachten bleibt, daß die Simulation in Hiqua12 sehr viel Aufwand erfordert und nicht immer durchgeführt werden kann. Weiter verdrängen die Test- und Debug-Phase gute Mechanismen von Med2, auf die nicht verzichtet werden kann. Wenn auch GTD in dieser Form nicht im neuen Expertensystem angewandt wird, so gibt es doch den Hinweis, die

Daten einer Simulation weiter als bisher (Diagnose Ja/Nein) zu interpretieren und die Simulation so zu modifizieren, daß sie jederzeit angewandt werden kann (siehe Integration der Simulation in das neue Expertensystem, Abschnitt 4).

3.1.4 Anlehnung an die Denkweise eines Experten.

Ein Experte diagnostiziert nicht nur anhand von assoziativem Wissen, er kennt auch die korrekte Funktionsweise des Systems. Die Integration des Hiqua1-Modells in das von assoziativem Wissen geprägte Expertensystem Med2 ermöglicht eine Annäherung an die Denkweise des Experten.

Grobe Vorgehensweise des Experten:

1) Der Experte sammelt die ersten Symptome (die nicht unbedingt komponentenbezogen sind).

Beispiel: die Maschine quietscht

Werkstücke sind nicht in Ordnung

Flüssigkeit tritt aus

Diese Symptome entsprechen den bisher in Med2 eingegebenen Symptomen.

2) Der Experte betrachtet grob die Maschine und erkennt auffallende Veränderungen gegenüber korrekt arbeitenden Maschinen.

Beispiel: die Maschine ist heiß

ein großes Teil der Maschine ist zerbrochen oder fehlt

Daten von Sensoren

Diese Symptome sind überwiegend komponentenorientiert.

3) Der Experte betrachtet die Maschine genauer und stellt weitere Fragen an den Maschinenbediener.

Die Eingabe in das neue Expertensystem teilt sich auf in Symptome, ähnlich derer in Med2, und in komponentenbeschreibende Symptome (durch das Questionset Lokalisationen). Diese zwei Symptomarten werden im Expertensystem unterschiedlich weiterbehandelt.

-Symptome wie in Med2.

Diese Symptome werden möglichst exakt einer oder mehreren Komponenten des Modells zugeordnet. Ist keine Zuordnung möglich, so muß man abwarten, bis weitere Informationen zur Verfügung stehen. Für weitere Probleme wie die XOR-Zuordnung siehe Hiqua1.

-Komponentenbezogene Symptome.

Alle Beschreibungen des Defektes in einer Komponente werden in komponentenorientierten Symptomen gespeichert, die der Komponente direkt zugeordnet sind. Gleichzeitig bilden diese komponentenorientierten Symptome eine Menge von Rohdaten, die mit einfachen Symptominterpretationen zu 'bisherigen' Symptomen oder Grobdiagnosen zusammengefaßt werden und dem Regelsystem von Med2 zur Verfügung stehen.

Die Information, ob eine Komponente defekt ist oder nicht, wird im Hiqua1-Modell in der Komponente direkt abgespeichert.

Folgende Werte sind möglich:

- o keine Information vorhanden
- o möglicherweise korrekt
- o korrekt
- o möglicherweise defekt
- o defekt

Die Angabe einer Begründung für den Wert erleichtert ein späteres Zurückziehen mit dem ITMS-Algorithmus. Die möglichen Begründungen sind:

- o Information vom Benutzer direkt durch Symptom S
- o Information durch Vererbung von Komponente K
- o Information durch Algorithmus1 in Hiqua1
- o Information durch Test (oder Einsetzen einer neuen Komponente)

3.2 Integration des Abstraktionsmechanismus des Ohrendiagnosesystems in das neue Expertensystem.

Im neuen Expertensystem stehen assoziatives und kausales Wissen nebeneinander zur Verfügung. Aus beiden Wissensarten extrahiert das neue Expertensystem Vorschläge für mögliche Diagnosen und bringt mit Hilfe des Abstraktionsmechanismus der Ohrendiagnose die verschiedenen Bewertungskriterien auf einen Nenner. Med2 stellt alle Diagnosen des WM mit den entsprechenden Wahrscheinlichkeiten zur Verfügung. Das Hiqua1-System bietet zwei Möglichkeiten zur Diagnosefindung:

(1) - Gegeben: Eine Menge defekter Komponenten im Hiqual-Modell. Algorithmus 2 findet mögliche Verursacherkomponenten für diese defekten Komponenten. Die den Verursachern zugeordneten Diagnosen werden dem Abstraktionsmechanismus zur Verfügung gestellt.

Maßstab: Anzahl der möglichen Diagnosen - Anzahl der widerlegten¹ Diagnosen.

- Eine Simulation ist gescheitert. Eine weitere Auswertung der Daten führt (über Algorithmus 2) zu möglichen Diagnosen (siehe Simulationen, Abschnitt 4.). Der Maßstab entspricht dem von oben.

(2) - Voraussetzung: 1:n Zuordnung von Diagnose zu Komponente.

Eine Komponente ist WIRKLICH defekt. Alle Diagnosen, die mit dieser Komponente verbunden sind und deren weitere Komponenten nicht als korrekt gekennzeichnet sind, werden zu hypothetischen Diagnosen (siehe Abb. 5.3).

Maßstab: Anzahl möglicher Diagnosen - Anzahl schon widerlegter Diagnosen

Komponenten des Hiqual-Modells Diagnosen

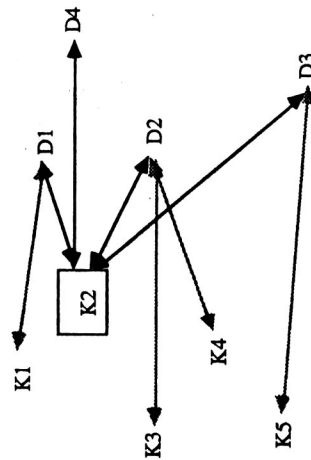


Abb. 5.3: Schematische Darstellung von (2)

K2 ist eine wirklich defekte Komponente. Sie hat Verbindungen zu den Diagnosen D1, D2, D3 und D4. Diese Diagnosen können außer auf K2 noch auf andere Komponenten verweisen. Diese Komponenten werden als Partnerkomponenten zu K2 bezeichnet.

Beispiel: Die Diagnose D1 liegt vor, wenn K2 und K1 defekt sind. K1 heißt dann Partnerkomponente zu K2.

Das assoziative und das kausale Diagnosefinden kann teilweise parallel ablaufen. (siehe Abschnitt 6).

¹ Schon widerlegte Diagnosen werden nicht mehr betrachtet.

Wahrscheinlichkeit der Diagnose im WM.	Wkt nahe dem Schwellpunkt	Wkt nahe dem Schwellpunkt	mittlere Wkt
Anzahl (1)	>5	4, 5	3, 2
Anzahl (2)	>10	≤10 ≥4	1, 2 3
Hinweisstärke	schwach	mittel	stark

Wahrscheinlichkeit der Diagnose im WM.	Wkt nahe dem Schwellpunkt	Wkt nahe dem Schwellpunkt	mittlere Wkt	sehr hohe Wkt
Anzahl (1)	4, 5	4, 5 > 5	≤2 >4	1
Anzahl (2)	≤10 ≥4	≤10 ≥4	>6 ≤2	=2
Hinweisstärke	schwach	stark	mittel	sehr stark

Abb. 5.4: Abstraktionstabelle.

Werden weitere mögliche Diagnosen z.B. aus einer Falldatenbank miteinbezogen, so wird der Abstraktionsmechanismus um eine Stufe erweitert (siehe Kapitel 6). Im Anschluß an den Abstraktionsmechanismus wird für die besten Diagnosen eine Differentialdiagnose wie in Med2 durchgeführt.

4. Integration der Simulation in das neue Expertensystem.

Da Simulationen für ein Expertensystem von großer Bedeutung sind, versucht man die Nachteile der Simulation in Hiqua2, die strengen Voraussetzungen und den enormen Aufwand, zu verringern.

Um den Aufwand der Simulation in Hiqua2 zu rechtfertigen, wird im ersten Abschnitt eine erweiterte Interpretation der Ergebnisse vorgeschlagen.

Im zweiten Abschnitt versucht man die Voraussetzung, daß die defekte Komponente vollständig beschrieben werden muß, im Rahmen einer "neuen Simulation" abzuschießen.

Voraussetzung für beide Simulationsarten ist, daß es einen Algorithmus gibt, der anhand der komponentenspezifischen Symptome den defekten Komponententyp erstellt, die Komponente initiiert und korrekt in das Hiqua-Modell einsetzt. Existiert ein solcher Algorithmus nicht, so muß der Benutzer diesen Part übernehmen. Dadurch werden die meisten Anwender von der Simulation ausgeschlossen.

4.1 Erweiterte Interpretation der Ergebnisse einer Simulation in Hiqua2.

Voraussetzung zur Durchführung der Simulation in Hiqua2 ist, daß Verhalten, Aufbau und Funktion der defekten Komponente vollständig beschrieben werden.

4.1.1 Mögliche Ergebnisse einer Simulation.

1) Die Simulation zeigt, daß diese defekten Komponenten die Ursache des Fehlers in der Maschine sind. Alle Symptome werden erklärt und neue pathologisch belegte Symptome treten nicht auf. Die Erklärungskomponente erläutert dem Benutzer die Ergebnisse und der Diagnoseprozeß ist beendet.

2) Die Simulation zeigt Symptome auf, die bisher nicht als defekt erkannt waren. Med2 kann anschließend diese komponentenorientierten Symptome untersuchen, indem es das entsprechende Questionset aktiviert.

3) Es gibt Symptome, die nicht erklärt werden können. In diesem Fall gibt es drei Reaktionsmöglichkeiten:

- Die der Simulation entsprechende Diagnose wird zurückgezogen. Weiter mit Med2.
 - Überprüfung der nicht erklärten Symptome. Liegen diese Symptome wirklich vor? (Inkonsistenzen?) Können nur ein oder zwei Symptome nicht erklärt werden, so ist diese Überprüfung sinnvoll.
 - Untersuchung der Ergebnisse der Simulation. Grund: Die defekte Komponente muß mit dem realen Fehler in Bezug stehen, da der Benutzer den Defekt der simulierten Komponente sehr genau beschrieben hat.
- A) Die Komponente ist symptomatisch defekt (siehe unten).
B) Es gibt eine weitere Komponente, die defekt ist.

Dieser Fall wird im neuen Expertensystem nicht betrachtet, und wird deshalb hier nur kurz angesprochen.

Die Untersuchung und Verwertung der Ergebnisse der Simulation hängen sehr davon ab, an welcher Stelle in Med2 die Simulation gestartet wird (Beispiel: als Differentialdiagnose, als Erklärung).

Im Folgenden werden Fall A) und Fall B) näher untersucht.

4.1.2 Einschleifiger Debugger für die Simulationsergebnisse.

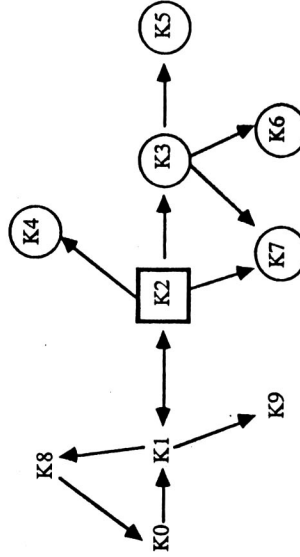


Abb. 5.5: Weitertransport von Fehlern im Hiqua-Modell.

Ki sind Komponenten des Hiqua-Modells.

Die Pfeile entsprechen einer direkten oder indirekten Verbindung zwischen zwei Komponenten im Hiqua-Modell über die hc/vc-Relation.

K2 sei 'defekt', aber nicht die eigentliche Ursache des Fehlers in der Maschine. Die Symptome der Komponenten K3, K4, K5, K6 und K7 werden durch K2 korrekt erklärt. Bei den Komponenten K0, K1, K8 und K9 treten keine oder andere Symptome auf. Ergebnis: K0 ist die 'wirklich' defekte Komponente.

Es ist wahrscheinlicher, daß nur eine Komponente defekt ist, als daß mehrere defekte Komponenten vorliegen. Deshalb wird zuerst Fall A) untersucht.

Fall A) Komponente K2 ist nur symptomatisch defekt.

Der Defekt, den der Benutzer beschrieben hat, wurde durch einen anderen Defekt ausgelöst. Alle Komponenten, deren Symptome durch K2 nicht oder nur unvollständig erklärt wurden, werden verdächtig. Algorithmus2 erhält als Eingabe all diese verdächtigen Komponenten, und die Angabe, daß K2 als Lösung nicht in Frage kommt. Ausgabe von Algorithmus2 ist ein möglicher Verursacher (oder mehrere) aller auftretenden Defekte.

Algorithmus2:

Eingabe: Mehrere 'defekte' Komponenten, die durch die Simulation nicht oder nur unvollständig erklärt wurden.

[K0, K1, K2, K8, K9 und die Bemerkung, dass K2 nicht die Lösung sein kann]

Ausgabe: Mögliche(r) Verursacher dieser 'defekten' Komponenten

[K0 als mögliche defekte Komponente]

Erweitertes Ergebnis der Simulation von K2:

K2 kommt als 'defekte' Komponente nicht in Frage, aber K0 ist nach Algorithmus2 sehr verdächtig.

Liegen mehrere Ergebnisse von Algorithmus2 vor, so werden alle Lösungsmöglichkeiten an Med2 zurückgegeben. Liefert Algorithmus2 kein Ergebnis, so treten wahrscheinlich mehrere defekte Komponenten gleichzeitig auf. => Fall B)

Fall B) Es liegen mehrere defekte Komponenten vor.

Beispiel:

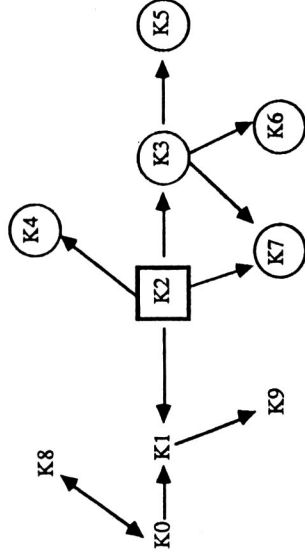


Abb. 5.6: Weitertransport von Fehlern im Hiqua1-Modell.

K1 sind die Komponenten des Hiqua1-Modells. Die Pfeile entsprechen einer direkten oder indirekten Verbindung zwischen zwei Komponenten im Hiqua1-Modell über die hc/vc-Relation.

Bezogen auf das vorangehende Beispiel ist außer der Komponente K2 mindestens noch eine weitere Komponente defekt. Die Anwendung des Algorithmus2 von Hiqua1 liefert bei einer Eingabe der verdächtigen Komponenten ohne K2 (K0, K1, K8 und K9) die Ausgabe K0.

K2 hat Verbindungen zu mehreren Diagnosen (siehe Abb. 5.7) [D1, D2, D3 und D4]. Für diese Diagnosen sind K2 oder K2 und "andere Komponenten" die entsprechenden defekten Komponenten. (Beispiel: D2 liegt vor, wenn K2 und K4 defekt sind. K4 = "andere Komponente". D3 liegt vor, wenn K6, K7 und K2 defekt sind.) Betrachte alle "anderen Komponenten", die K2 jeweils als Partnerkomponente dienen und schneide diese Menge mit der Ausgabemenge von Algorithmus2. Die Schnittmenge dieser beiden Komponentenmengen liefert mögliche Partnerkomponenten für K2.

Ist die Schnittmenge leer, so ist es möglich, daß mehrere Diagnosen auf einmal auftreten, oder es liegt eine neue Diagnose vor, die aus der Kombination von K2 und einer Komponente der Ausgabemenge besteht.

Komponenten des
Hiqua1-Modells

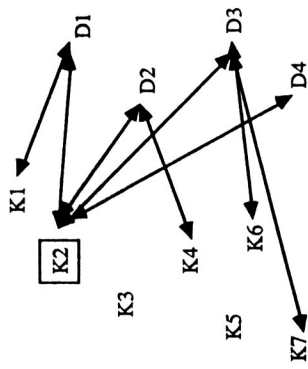


Abb. 5.7: Schematische Darstellung der Verbindungen zwischen Komponenten und Diagnosen.

Erweitertes Ergebnis der Simulation von K2:

Die Komponente K2 kann alleine nicht alle Symptome erklären, aber es gibt eine Komponente, die laut Algorithmus2 und einer Überprüfung als Partnerkomponente zu K2 für eine Diagnose, als zusätzliche defekte Komponente möglich ist (K0).

Fall B) kann noch weiter ausgeführt werden, aber die Problematik mehrerer defekter Komponenten oder sogar mehrerer gleichzeitig vorliegender Diagnosen ist so umfassend, daß man im neuen Expertensystem davon ausgeht, daß nur eine Komponente defekt ist und eine Diagnose immer nur einer Komponente zugeordnet ist.

4.1.3 Zusammenfassung.

Eine Simulation in Hiqua2 erfordert sehr viel Aufwand und das Ergebnis, die Simulation mit dieser Komponente erklärt alle Symptome (--> ok) oder sie erklärt nicht alle Symptome (--> Diagnose zurückziehen) steht in keiner Relation zu den Informationen, die eine Simulation weitergeben kann.

Im GTD-Paradigma [Davis et al. 87] wird der Testphase (in Form einer Simulation) ein Debugger angeschlossen. Jetzt verfügt auch die Hiqua2-Simulation über einen Ein-Schleifen-Debugger, d.h. es werden nicht automatisch neue Simulationen durchgeführt (Kosten!), sondern die Kontrolle wird wieder an Med2 zurückgegeben. Eine Vereinfachung der Durchführung der Simulation ermöglicht, auch mehrere Debuggerschleifen.

Die erweiterte Interpretation der Ergebnisse der Simulation kann Med2 steuern. Es wird dann nicht nur eine Simulation aufgerufen, sondern die Simulation wird mit einer oder mehreren Nachbehandlungen (Behandlung neu auftretender Symptome, Fall A) oder Fall B)) beliebig gekoppelt. Eine Kopplung mit Fall B) kommt für das neue Expertensystem nicht in Frage, da hier Mehrfachdiagnosen ausgeschlossen werden.

4.2 Neue Simulationsmöglichkeiten.

Die Bedingungen zur Durchführung einer Simulation in Hiqua2 sind enorm: Der Benutzer muß die defekte Komponente vollständig beschreiben. Vollständigkeit ist allerdings eine Voraussetzung, die in Expertensystemen nur selten gegeben ist, deshalb versucht man diese Bedingung abzuschwächen.

Sind nicht alle oder gar keine komponentenspezifischen Symptome (= die Symptome, die den Defekt der Komponente beschreiben) bekannt und ist trotzdem die Simulation der Komponente von Interesse, so wird versucht Informationen über die Komponente von den Diagnosen her zu erhalten.

Jeder Diagnose werden Werte über die komponentenspezifischen Symptome zugeordnet. Aus diesen Werten kann das Expertensystem die defekte Komponente selber zusammenbauen und einsetzen. Gibt der Benutzer einige Symptomwerte an, so werden die übrigen Symptome an die beobachteten Symptome angepaßt (Mit Hilfe einer komponentenorientierten Regelmenge von 2-3 Regeln pro Komponente).

Diese Simulation startet mit vermuteten Mittelwerten, anstatt mit korrekten im Modell 'gesehenen' Werten. Deshalb ist die Aussagekraft dieser Simulation nicht so hoch wie die der Simulation in Hiqua2. Die Aussagefähigkeit nimmt ab, je weniger komponentenorientierte Symptome bekannt sind.

Für die Bereiche, in denen bisher die Simulation angewandt wurde, wie zum Beispiel bei der Differentialdiagnose oder der Erklärungskomponente, ist die neue Simulation zu ungenau. Neue Anwendungsfelder sind:

- 1) Test der Diagnosen im WM
- 2) Med2 weiß nicht mehr weiter
- 3) Vorstufe zu einer Hiqual2-Simulation

Die Vorteile der neuen Simulation sind unumstritten, aber das Ergebnis ist so vage, daß ein Einsatz in den bisherigen Anwendungsfeldern der Simulation nicht möglich ist. Die Ergebnisse, die die Simulation in den neuen Anwendungsbereichen erzielen kann, stehen in keiner Relation zu dem Aufwand, der dafür betrieben werden muß. Hier stellt Med2 einfachere Mechanismen zur Verfügung.

Die neue Simulation kann einspringen, falls fast alle komponentenorientierten Symptome bekannt sind, und zu einer Simulation mit Hiqual2 nur einige wenige Werte fehlen. Das Anwendungsgebiet entspricht dann dem der Simulation in Hiqual2. Zusätzlich zu den Werten sollte im System eine Wahrscheinlichkeit über das Eintreten dieser Werte berechnet werden. Überschreitet die Wahrscheinlichkeit einen bestimmten Schwellwert, so wird die Simulation durchgeführt und ähnlich einer Simulation mit Hiqual2 verwertet.

4.3 Bewertung der beiden Simulationsmöglichkeiten.

Simulation mit Hiqual2:

Vorteil:

- Korrekte Simulation aufgrund von Daten, die wirklich im Modell vorliegen (Die simulierte Komponente muß trotzdem nicht die wirkliche Verursacherkomponente sein).

Nachteil:

- Die Simulation ist oft nicht anwendbar, da
 - o ein zu großer Aufwand gefordert wird
 - o alle komponentenspezifischen Symptome bekannt sein müssen

Neue Simulation:

Vorteil:

- Die Simulation ist immer anwendbar.
- Nur einige oder gar keine komponentenspezifischen Symptome müssen bekannt sein.
- Der Benutzer muß nicht unbedingt angesprochen werden.

Nachteil:

- Nur ein vages Ergebnis, das für die bisherigen Anwendungsbereiche der Simulation nicht genau genug ist.
- Der Aufwand des Systems bleibt wie in der Simulation von Hiqual2.
- Zusätzlicher Aufwand für die diagnosebezogenen Mittelwerte und die Berechnungsmechanismen, falls einige Symptome bekannt sind und die anderen angeglichen werden sollen.

Wenn irgend möglich, sollte man die Simulation in Hiqual2 benutzen.

Die neue Simulation ist nur für die Fälle geeignet, in denen nur einige wenige Werte für eine Simulation fehlen.

5. Funktionale Sicht des neuen Expertensystems.

1. Strukturierung der Symptome und Diagnosen, sowie Gruppierung der Regeln.
2. Hierarchischer Aufbau des kausalen Modells.
3. Vorverarbeitung der Symptome in Med2.
4. Verarbeitung quantitativer und qualitativer Zeitangaben.
5. Zeitrepräsentationsmechanismen.
6. Effizienter Problemlösungsmechanismus.
7. Mechanismen zur Plausibilitätskontrolle der Eingabe.
8. Recovery-Mechanismen.
9. Mechanismen zur Darstellung und Verarbeitung von unsicherem und unvollständigem Wissen.

6. Architektur des neuen Expertensystem.

Kurz zusammengefaßt besteht das neue Expertensystem aus Hiqua12, dem Abstraktionsmechanismus des Ohrendiagnosesystems und großen Teilen von Med2.

Im Gegensatz zu den bisherigen Betrachtungen geht man jetzt davon aus, daß nur eine Diagnose vorliegt und diese Diagnose nur einer Komponente zugeordnet ist. Zwischen Diagnose und Komponente liegt also eine 1:1 Beziehung vor, während eine Komponente aber auf mehrere Diagnosen verweisen kann.

6.1 Änderungen von Hiqua12 zur Integration in das neue Expertensystem.

Das Konzept des Hiqua-Systems und das Modell bleiben bestehen. In den Komponenten des Modells werden zusätzlich folgende Informationen gespeichert:

- Ist die Komponente leicht testbar?
- Die Komponente ist defekt. Ist es sinnvoller die Komponente gleich auszutauschen oder die Unterkomponenten zu betrachten?
- Informationen über den Zustand der Komponente.

- o keine Information
- o defekt
- o möglicherweise defekt

- o korrekt
 - o möglicherweise korrekt
- und die dazugehörigen Begründungen:
- o Information direkt vom Benutzer durch Symptom S
 - o Information durch Vererbung von Komponente K
 - o Information durch Test
 - o Information durch Algorithmus1 in Hiqua1
- Verbindungen zu Symptomen
- Verbindungen zu Diagnosen

Verbindungen zwischen Symptomen, Diagnosen und Komponenten.

Die Symptome des neuen Expertensystems teilen sich auf in komponentenorientierte und die schon bekannten Symptome des Symptomwaldes von Med2. Die komponentenorientierten Symptome beschreiben den Defekt der Komponente und sind ihr direkt zugeordnet. Einfache Symptominterpretationen ermöglichen auch einen Informationsaustausch zwischen den komponentenorientierten Symptomen und dem Symptomwald von Med2. Deshalb bilden die komponentenorientierten Symptome ein Bindeglied zwischen den Symptomen des Expertensystems einerseits und den Komponenten des Hiqua-Modells andererseits.

Die Verbindung zwischen einem Symptom und einer Komponente liegt immer in beiden Richtungen vor. Wenn ein Symptom vorliegt, dann ist die Komponente entweder defekt (korrekt) oder nur möglicherweise defekt (korrekt), je nachdem wie die Verbindung aussieht. Die umgekehrte Verbindung von einer Komponente zu einem Symptom hat lediglich Zuordnungscharakter. Ihre bisherige Aufgabe in Hiqua2 wird dank der 1:1 Beziehung zwischen Diagnosen und Komponenten durch die Diagnosen in Med2 besser bewältigt.

Die Speicherung der Verbindungen erfolgt in den Komponententypen. Alle weiteren Verbindungen werden in einer Regelmenge mit dem neuen Typ 'verbSK' und 'verKS' aufgenommen.

Der Verbesserungsvorschlag der XOR-Feuerung in Hiqua1 kann übernommen werden. Der zweite Verbesserungsvorschlag, Symptome nicht nur auf die Komponenten, sondern auch auf ihre Zustände verweisen zu lassen, wird durch die komponentenorientierten Symptome überflüssig.

Die Verbindungen zwischen Komponenten und Diagnosen und umgekehrt tragen keine Informationen, sie sind nur Zuordnungsrelationen. Eine Komponente kann auf mehrere Diagnosen verweisen, aber einer Diagnose ist immer nur eine Komponente zugeordnet. Diagnosen, die keiner Komponente zugeordnet sind, gibt es nicht. Die Abspeicherung der Verbindungen erfolgt wie bei den Symptomen in den Komponententypen. Außergewöhnliche Verbindungen werden in einer Regelmenge vom Typ 'verbDK' und 'verbKD' gespeichert.

Die Verbindungen zu Therapien sollen prinzipiell nur von den Diagnosen ausgehen. Eine Komponente weist je nach ihrem Defekt auf unterschiedliche Diagnosen, die dann auf die entsprechende Therapie verweisen.

Veränderungen in Algorithmus1 und Algorithmus2.

Der Algorithmus1 aus Hiqua1 kann unverändert in das neue Expertensystem übernommen werden. Anwendung findet er als Bedingung für jede Simulation in Hiqua2 und in Fällen, in denen die Folgen einer bestimmten defekten Komponente betrachtet werden.

Algorithmus2 wird durch die Voraussetzung, daß nur eine Komponente des Modells defekt ist, vereinfacht. Die Veränderungen sind in kursiver Schrift angedeutet.

- (1) Betrachte alle 'defekten' Komponenten und ihre Beziehungen untereinander. Erstelle aus dem Hiqua-Modell ein Netzwerk, dessen Knoten die 'defekten' Komponenten sind. Liegt im Hiqua-Modell zwischen zwei Knoten eine direkte oder indirekte hc/vc - Relation vor, so werden diese Knoten mit einem Pfeil verbunden.
- (2) Ist eine dieser Komponenten des Netzwerks leicht testbar?
 Ja --> Ist die Komponente laut Test wirklich defekt?
 Nein --> weiter bei (3)
 Ja --> Diese Komponente kann kein Verursacher sein.
 Ja --> Der Defekt ist gefunden. Ausgabe aller durch diesen Defekt erklärten und nicht erklärten Komponenten mit entsprechendem Hinweis. Stop.
- (3) Liegen zwischen zwei Komponenten (K1 und K2) Relationen in beide Richtungen vor, so werden diese zu einer Komponente (K1, K2) zusammengefaßt.
- (4) Wähle eine Komponente, die keine eingehenden Pfeile besitzt.
 Gibt es keine solche Komponente, so gehe zu (5).
 Andernfalls: Diese Komponente K5 kann durch keine andere Komponente erklärt werden und muß also eine verursachende Komponente sein.

Ist K5 eine Komponentenumenge?

- Nein --> Ausgabe von K5.
 Ja --> Sind in K5 Komponenten enthalten, die leicht testbar waren?
 Nein --> Ausgabe der gesamten Menge.
 Ja --> Gibt es eine oder mehrere Komponenten, die als korrekt erkannt wurden?
 Nein --> Ausgabe aller Komponenten der Menge.
 Ja --> Ausgabe aller nicht als korrekt erkannten Komponenten.
 Ausgabe aller durch K5 erklärten und nicht erklärten Komponenten mit entsprechendem Hinweis. Stop.

- (5) Jede Komponente hat einen eingehenden Pfeil. Es liegt also ein Zyklus vor. Alle Komponenten des Zyklus werden zu einer Komponente zusammengefaßt (K1, ..., Kn). Weiter mit (4).

Veränderungen der Simulation.

Die Simulation in Hiqua2 wird trotz ihrer zahlreichen Nachteile in das neue Expertensystem übernommen. Voraussetzung für die Integration in das neue Expertensystem ist allerdings ein Mechanismus, der aus den komponentenorientierten Symptomen automatisch den defekten Komponententyp erstellt, ihn initiiert und einbaut. Kann diese Voraussetzung nicht erfüllt werden, so muß der Benutzer viel Wissen über die Maschine und das Hiqua-Systems mitbringen, um eine Simulation zu starten. Die Benutzung der Simulation und damit auch des neuen Expertensystems wird dann für eine grosse Benutzerschicht zu schwierig und deshalb ungeeignet.

Die Voraussetzung, daß alle komponentenorientierten Symptome bekannt sind, ist in Expertensystemen oft nicht gegeben (Vollständigkeit!). Eine Regelmenge, die aus den bekannten komponentenorientierten Symptomen die übrigen berechnet und eine Wahrscheinlichkeit für diese Kombination ausgibt, hilft in vielen Fällen weiter (siehe neue Simulation, Abschnitt 4). Die Ergebnisse einer solchen Simulation müssen immer im Hinblick auf die von der Regelmenge angegebene Wahrscheinlichkeit verwertet werden. Die Möglichkeit direkt von einer Diagnose aus, ohne (oder mit wenigen) bekannten komponentenorientierten Symptomen eine defekte Komponente zu simulieren, wird dabei nicht in Betracht gezogen. Der hier integrierte Mechanismus ist nur für Anwendungsfälle geeignet, in denen die meisten komponentenorientierten Symptome bekannt sind und nur einige wenige Daten fehlen, deren Berechnung keine größeren Probleme aufwirft.

6.2 Veränderungen von Med2 zur Integration in das neue Expertensystem.

Die Leitideen von Med2 wie der hierarchische und modulare Aufbau, die hypothetisch deduktive Vorgehensweise mit der anschließenden Differentialdiagnose und das nicht-monotone Argumentieren bleiben erhalten.
In den folgenden Abschnitten wird beschrieben, welche Komponenten erweitert und welche durch Hiqual2 ersetzt werden.

6.2.1 Veränderungen der Wissensrepräsentation im Vergleich zu Med2.

- A) Die *Pathokonzepte* fuer Komplex-, Grob- und Feindiagnosen bleiben unverändert bestehen und ermöglichen weiterhin den Aufbau der Diagnosehierarchien und -heterarchien
- B) Die *Manifestationen* für Fragen und einfache Symptominterpretationen werden in zwei sich nicht ausschließende Gruppen aufgeteilt:
- die 'bisherigen' Manifestationen und
 - die komponentenorientierten Manifestationen
- Jeder Komponente im Hiqual-Modell ist ein Questionset zugeordnet, das aus komponentenorientierten Manifestationen besteht.
- C) Die Menge der *Questionsets* zur Gruppenbildung zusammengehöriger Manifestationen wird um die *Questionsets* der Komponenten erweitert, die die *Questionsets* 'Lokalisationen' ersetzen.
- D) *Therapeutische Varianten der Pathokonzepte* bleiben unverändert.
- E) Die *Explanationsets* verlieren durch die Simulationsmöglichkeit an Bedeutung. Da aber die Simulation nicht immer durchgeführt werden kann, bleiben die *Explanationsets* erhalten.
- F) Das Hiqual-Modell ersetzt die *Lokalisationen* (siehe Himed).
- G) Das neue Expertensystem benutzt die *strukturierten Regeln* von Med2. Das Diagnosenetzwerk und die Möglichkeit des nicht-monotonen Argumentierens bleiben somit erhalten. Für die Verbindungen zwischen Diagnosen, Symptomen und

Komponenten des Hiqual-Modells wird die Regelmenge um Regeln vom Typ 'verbKD', 'verbDK', 'verbSK', und 'verbKS' erweitert.

6.2.2 Konsistenzüberprüfung der Eingabedaten.

Die Informationen über den Zustand der einzelnen Komponenten im Hiqual-Modell und die dazugehörige Begründung ermöglichen zusammen mit dem Vererbungsmechanismus eine Konsistenzüberprüfung vorhandener oder neuer Eingaben anhand des Modells. Der Zustand der Komponente kann die Werte "keine Information", "möglicherweise defekt", "defekt", "möglicherweise korrekt" oder "korrekt" annehmen.

Vererbungsalgorithmus

A) Die neue Eingabe für eine beliebige Komponente K1 lautet: korrekt.

A1) Betrachtet wird der Zustand der Komponente K1.

bisher 1) keine Information ==> 'korrekt' und neue Begründung werden
möglicherweise korrekt in K1 eingetragen.
korrekt

möglicherweise defekt

2) defekt

==> Inkonsistenz!

Beide Begründungen erneut
untersuchen

A2) Alle Unterkomponenten von K1, die mit K1-O (=Output von K1) über die structure-map Relation verbunden sind, werden betrachtet:

bisher 1) keine Information ==> 'korrekt' und die neue Begründung
möglicherweise defekt werden in K1 eingetragen.
möglicherweise korrekt

2) korrekt

==> a) bisher: vererbte Begründung

=> neue Begründung eintragen

b) bisher: direkte Begründung

=> Vererbungsmechanismus

stoppt in diesem Zweig

3) defekt

==> Inkonsistenz!

Beide Begründungen untersuchen.
Alle Unterkomponenten, deren Zustand oder Begründung geändert wurde, werden in A2 neu eingegeben.

B) Die neue Eingabe für K1 lautet 'defekt'.

B1) Betrachtet wird der Zustand der Komponente K1.

bisher: 1) keine Information
möglicherweise korrekt
möglicherweise defekt

2) defekt

==> a) bisher: vererbte Begründung
=> neue Begründung eintragen
b) bisher: direkte Begründung
=> Vererbungsmechanismus
stoppt in diesem Zweig
==> Inkonsistenz!

3) korrekt

Beide Begründungen untersuchen.

B2) Analog zu A2)

C) Die neue Eingabe für K1 lautet möglicherweise korrekt.

C1) Betrachtet wird der Zustand der Komponente K1.

bisher: 1) keine Information
möglicherweise korrekt
2) möglicherweise defekt
=> möglicherweise korrekt und die neue Begründung in K1 eintragen.
=> zusätzlich die neuen Werte eintragen.

3) korrekt
defekt

==> keine Änderung

C2) Alle Unterkomponenten von K1, die mit K1-O über eine structure-map-Relation verbunden sind, werden betrachtet.

bisher: 1) keine Information
=> möglicherweise korrekt und die neue Begründung in K1 eintragen
=> keine Änderung

2) möglicherweise defekt
defekt
korrekt
möglicherweise korrekt

D) Die neue Eingabe für K1 lautet möglicherweise defekt.

Analog zu C)

6.2.3 Recovery im neuen Expertensystem.

Um das Prinzip des nicht-monotonen Argumentierens nicht zu gefährden, muß es möglich sein, abgeleitete und eingegebene Daten jederzeit wieder zurückzuziehen. Dazu gehören auch die Daten im Hiqual-Modell, sowie diejenigen Daten, die vom Hiqual-Modell aus weitergegeben wurden. Der ITMS-Algorithmus behandelt die Diagnosen, Symptome und Regeln trotz der Erweiterungen wie bisher. Zusätzlich zu betrachten sind die Angaben in den Komponenten des Hiqual-Modells, sowie die Verbindungen von den Symptomen und Diagnosen zu den Komponenten.

Variable Daten einer Komponente im Hiqual-Modell sind ihr Zustand und die dazugehörige Begründung. Wird eine Begründung zurückgezogen, so werden der entsprechende Zustand und alle vererbten Werte gelöscht. Der Vererbungsmechanismus belegt die Komponenten anschließend mit neuen Werten.

Problem: Eine Komponente K wird zuerst mit 'Begründung1 und korrekt' belegt, anschließend mit 'Begründung2 und korrekt'.

Beide Belegungen gehen direkt von K aus und werden nicht vererbt.

==> Nur die Begründung2 wird gespeichert.

Ein Zurückziehen der Begründung2 führt dazu, daß die Komponente K den Wert 'keine Information' erhält, obwohl 'Begründung1 und korrekt' noch gelten. Diesem Problem kann man abhelfen, indem man alle (direkten) Zustandswerte und ihre Begründungen in einer Liste speichert.

Da Simulationen nur selten durchgeführt werden, speichert man ihre Voraussetzungen und Ergebnisse in einer Tabelle. Alle übrigen Erweiterungen von Med2 kann man als Regeln oder regelmäßige Mechanismen betrachten und mit dem ITMS-Algorithmus behandeln.

6.2.4 Erweiterung der Questionsetselektion.

Im neuen Expertensystem erfolgt die Auswahl der Questionsets in fünf Prioritätsstufen:

- 1) Vom Benutzer selektierte Questionsets.
- 2) Genauere Abklärung des Zustandes eines etablierten Pathokzeptes oder einer etablierten Manifestation (Regeln vom Typ 'ind').
- 3) Questionsets der Lösungskomponenten von Algorithmus2.
- 4) Questionsets zur Abklärung eines verdächtigen Pathokzeptes oder zur Unterscheidung zwischen Differentialdiagnosen.
- 5) Questionset zur symptomatischen Indikation

Im Vergleich zu Med2 ist die Questionsetselektion in Punkt 3) und indirekt über das modifizierte WM in Punkt 4) verändert.

6.3 Abstraktionsmechanismus

Im Diagnoseprozess des neuen Expertensystems werden hypothetische Diagnosen parallel aus dem assoziativen und dem kausalen Wissen abgeleitet. Um diese verschiedenartig bewerteten Diagnosen zu vergleichen, wird ein Abstraktionsmechanismus ähnlich dem im Ohrendiagnosesystem angewandt.

Wkt der Diagnosen im WM	Wkt nahe dem Schwellpunkt	mittlere Wkt	Wkt nahe dem Schwellpunkt
Anzahl der Diagnosen zu Lösungskomponenten von Alg.2	$3 < \text{Anzahl} \leq 10$	Anzahl = 1, 2 oder 3	$3 < \text{Anzahl} \leq 10$
Hinweisstärke	schwach	mittel	stark

Wkt der Diagnosen im WM	sehr hohe Wkt	Wkt nahe dem Schwellpunkt	mittlere Wkt
Anzahl der Diagnosen zu Lösungskomponenten von Alg.2		1	2 oder 3
Hinweisstärke	sehr stark		

Abb. 5.8: Abstraktionstabelle

Das WM enthält jetzt alle Diagnosen, deren Hinweisstärke "stark" oder "sehr stark" ist, und die Diagnosen, die schon vorher im WM¹ enthalten waren, aber nicht in den Abstraktionsmechanismus eingegangen sind.

Mit diesem veränderten WM geht der Diagnoseprozess in Med2 normal weiter.

6.4 Grober Ablauf des neuen Expertensystems.

Das neue Expertensystem besteht, wie Med2, das Ohrendiagnosesystem und auch Abel, aus einer Initialphase und einer Finalphase.

Initialphase:

Die Initialphase ist eine hypothesenfreie Phase, in der Fragen über bisherige Manifestationen an den Benutzer gestellt und einfache Symptominterpretationen durchgeführt werden. Nur auf ausdrücklichen Wunsch des Benutzers, werden auch komponentenorientierte Questionsets miteinbezogen. Alle Eingaben werden der Konsistenzüberprüfung unterzogen.

Am Ende der Initialphase liegen wie in Med2 Symptome vor, die durch einfache Symptominterpretationen aufbereitet sind. Zusätzlich sind schon einige komponentenorientierte Symptome im Hiqua-Modell belegt.

¹Die vom Benutzer im WM verankerten Diagnosen und die Differentialdiagnosen einer schon etablierten Diagnose bleiben im WM und gehen nicht in den Abstraktionsmechanismus ein.

1. Grundkonzept für das erweiterte Hiqual-System.

Eine Modifikation von Hiqual, die zur Integration in Expertensystemen geeignet ist, kann folgendes Aussehen haben:

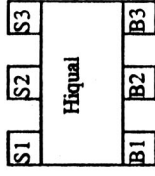


Abb. 6.1: S1 : Standardpakete
 B1 : Pakete für expertensystemspezifische Anwendungen

Ausgehend von dem ursprünglichen Hiqual-Modell und den zusätzlichen Standardpaketen gibt es die Möglichkeit, besondere Pakete auszuwählen, die nicht für jedes Expertensystem von Bedeutung sind.

Beispiele für Standardpakete:

- Zuordnung Komponente <-> Diagnose
- Zuordnung Komponente <-> Symptome
- Möglichkeiten zur Behandlung quantitativer Zeitangaben
- einfache Fragestellungen (Hiqual1)

Beispiele für besondere Pakete:

- speziellere Zuordnung der Symptome und Diagnosen auch zu den Ereignissen und Zuständen
- Ausbau der quantitativen Zeitangaben
- Angabe geometrischer Zusammenhänge
- Zuordnung Komponente <-> Therapie
- Simulationsmöglichkeiten (Hiqual2)

Soll Hiqual in ein Expertensystem integriert werden, so werden das Hiqual-Modell und die Standardpakete auf jeden Fall integriert. Zur Verbesserung des Modells kann man sich zusätzlich besondere Pakete aussuchen, die in diesem Expertensystem von Bedeutung sind. Gleichzeitig kann man auf 'Pakete' verzichten, die nicht zur Anwendung kommen und nur Platz und Zeit in Anspruch nehmen.

Finalphase:

Die Finalphase teilt sich auf in die Hypothesenphase und die Anamnese-phase. In der Hypothesenphase arbeiten das assoziative und das kausale Modul parallel. Med2 stellt die Diagnosen des WM auf. Hiqual2 sucht mit Algorithmus2 im Hiqual-Modell nach einer Verursacherkomponenten, die die (möglicherweise) defekten Zustände aller anderen Komponenten erklären kann. Die entsprechenden Diagnosen aller Lösungsmöglichkeiten von Algorithmus2 (außer den Diagnosen, die schon widerlegt sind) und die Diagnosen des WM werden im Abstraktionsmechanismus verglichen. Die anschließende Differentialdiagnose entspricht der in Med2.

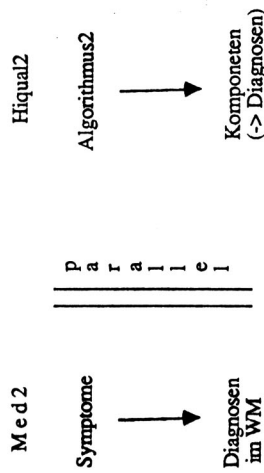


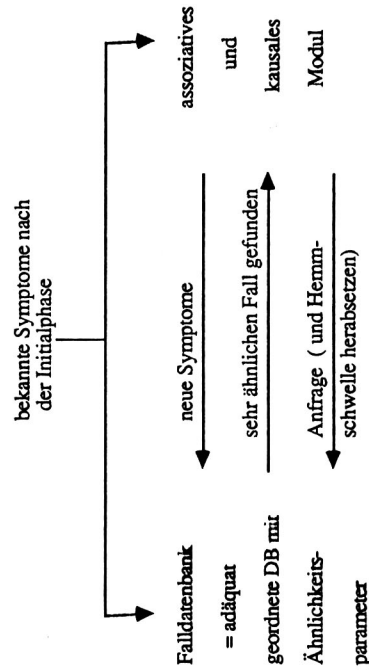
Abb. 5.9: Ablauf der Hypothesenphase.

In der Anamnese-phase wird ein Questionset aktiviert, das das System mit weiteren Daten versorgt. Zusätzlich zu den in Med2 vorgenommenen Bewertungen, werden die Questionsets der Lösungskomponenten von Algorithmus2 mit der drithöchsten Priorität belegt.

2. Vorschläge zur Erweiterung des neuen Expertensystems.

2.1. Integration einer Falldatenbank.

In das bisher beschriebene System, bestehend aus einem assoziativen Modul und einem kausalen Modul, kann zusätzlich eine Falldatenbank integriert werden. Ähnlichkeiten zwischen dem aktuellen Fall und früheren Fällen aus der Datenbank liefern Informationen zur Diagnostik und insbesondere zur Therapie und Prognose. Die Suche in der Datenbank läuft parallel zu den Aktionen des assoziativen und des kausalen Moduls ab. Der bisherige Diagnoseprozess wird also kaum verzögert. Die Datenbank enthält Fallbeschreibungen bisher aufgetretener Defekte dieser Maschine, sowie die Beschreibung seltenerer Fälle an ähnlichen Maschinentypen.



Vorgehensweise:

Nach der Initialphase werden alle bekannten Symptome sowohl in die Falldatenbank als auch in das assoziative und kausale Modul eingegeben und beide Systeme beginnen zu arbeiten. Neue Symptome werden direkt an die Falldatenbank weitergegeben. Damit die Falldatenbank den normalen Ablauf nicht zu häufig unterbricht, wird eine Hemmschwelle eingebaut. Zu Beginn ist diese Hemmschwelle sehr hoch, damit der Kontrollfluß des assoziativen und kausalen Moduls nur unterbrochen wird, wenn sich in der Datenbank ein sehr ähnlicher Fall findet. Liegt ein typischer Fehler vor, so kann die Diagnose schneller mit der Falldatenbank als mit dem assoziativen und kausalen Modul gefunden werden. Hält diese Diagnose einer Überprüfung im assoziativen und kausalen Modul stand, so

wird sie zusammen mit Erklärungen des kausalen oder des statistischen Modells ausgegeben. Kommt das assoziative und kausale Modell zu keinem Ergebnis, so kann man auch die Hemmschwelle verändern oder eine Anfrage an die Falldatenbank bezüglich ähnlich gelagerter Fälle richten.

Liegt eine seltenere Kombination von Symptomen vor, so kann die Datenbank nach ähnlichen Maschinentypen (evtl. Vorgängertypen) suchen, bei denen diese Symptome schon vorlagen. Anhand der Beschreibung des damaligen Falles mit dessen Therapie und Prognose, kann man aus früheren Fehlern lernen (und diese vermeiden).

Die gemeinsame Bewertung der hypothetischen Diagnosen der Falldatenbank und des assoziativen und kausalen Modells wird durch einen Abstraktionsmechanismus ähnlich dem im Ohrendiagnosesystem durchgeführt.

2.2 Integration von Wissen über externe Vorgänge.

Gibt ein Benutzer dem System ein, daß äußere Einflüsse auf die Maschine eingewirkt haben, so ist dies ein wichtiger Ausgangspunkt für die Diagnose, der unbedingt beachtet werden sollte. Ereignisse wie Wasser, Stromausfall, Stöße, hoher Feuchtigkeitsgrad oder ein hoher Anteil von Staubpartikeln in der Luft, können die Maschine in ihrem Verhalten und in ihrem Fehlverhalten beeinflussen. Alle äußeren Einflüsse auf die Maschine kann man nicht integrieren, weil man sie entweder nicht kennt, oder ihre Auswirkungen im Modell nicht beschreibbar sind.

Es bieten sich zwei Möglichkeiten zur Integration:

A) Der Benutzer gibt einen der äußeren Einflüsse z.B. Wasser an.

Das System kann so reagieren, daß es mit Hilfe eines Questionsets das 'Wasser' genauer spezifiziert:

- War das Wasser sauber, oder handelt es sich um Schlammwasser?
- Kam das Wasser von oben oder von unten?
- von unten: - Bis in welche Höhe stieg das Wasser?
- von oben: - Kam das Wasser als Strahl, oder nur tropfenweise?

Aus diesen Angaben versucht das System 'normale' Symptome zu finden (mit Hilfe von Symptom-Symptom-Relationen) und achtet auf besonders wasserempfindliche Teile der Maschine, indem es entsprechende Untersuchungen anordnet.

B) Die zweite Möglichkeit besteht darin, diese äußeren Einflüsse im Rahmen eines Rechtfertigungsmechanismus bei Inkonsistenzen zu integrieren.

Beispiel: Eine Maschine besitzt nur im unteren Teil einen Wasserkreislauf. Im Normalfall steht das Wasser nicht unter Druck, d.h. es wird im Falle eines Defektes der

Leitung nach unten und nicht nach oben spritzen. Trotzdem findet man im oberen Teil der Maschine Wasser. Ist doch ein höherer Druck auf der Leitung? Fragt man beim Benutzer nach, so erfährt man, daß Regenwasser durch das Fenster eingedrungen ist. Solche Inkonsistenzen können manchmal mit äußeren Einflüssen erklärt werden.

C) Da äußere Einflüsse dieser Art selten vorkommen, ist es wenig sinnvoll, den Benutzer ohne vorliegende Inkonsistenzen direkt darauf anzusprechen.

Bei äußeren Einflüssen oder selteneren Defekten kann auch die Falldatenbank weiterhelfen. Vergleiche mit denselben oder auch nur ähnlichen Maschinen, die diesen Einflüssen ausgeliefert waren, können besonders im Hinblick auf die jeweilige Therapie und Prognose die Lösung des Problems erleichtern. Allerdings muß man die Unterschiede der Maschinen und der externen Einflüsse immer im Auge behalten (siehe auch Falldatenbanken).

Allen, J., Maintaining Knowledge about Temporal Intervals, CACM, Vol.26, No.11, (832-843), 1983

Allen, J., Towards a General Theory of Action and Time, Artificial Intelligence, Vol.23, (123-154), 1984

Davis, R., Diagnostic Reasoning Based on Structure and Function, AI-Journal-84, 1984

Davis, R., Simmons, R., Generate, Test and Debug: Combining Associational Rules and Causal Models Artificial Intelligence Laboratory Massachusetts Institute of Technology (1071-1078) 1987

Heinen, P., Informelle Schlußweisen auf verschiedenen Abstraktionsniveaus in einem medizinischen Expertensystem: Konzepte und Implementierung. Diplomarbeit, Aachen 1985

Heinen, P., Reusch, H., Richter, M., Wetter, T., Formal Description of Objects, Processes and Levels of Expert Reasoning, GWAI-85, Springer-Verlag

Patil, R.S. Causal Representation of Patient Illness for Electrolyte and Acid-Base Diagnosis, Ph.D. thesis, Dept. of Electrical Engineering and Computer Science, MIT Lab. for Comp. Sci., Cambridge, Mass., Oct 1981

Patil, R.S., Szolovits, P. and Schwartz, W.B., Causal Understanding of Patient Illness in Medical Diagnosis, Proc. Seventh Intl. Joint Conf. on Artificial Intelligence, Yale University, New Haven CT, 1981

Patil, R.S., Szolovits, P. and Schwartz, W.B., Modeling Knowledge of the Patient in Acid-Base and Electrolyte Disorders, Artificial Intelligence in Medicine, AAAAS Selected Symposium 51, Boulder: Westview Press 1982

Puppe, F. Assoziatives Diagnostisches Problemlösen mit dem Expertensystemshell Med2, Dissertation, Kaiserslautern 1986a

Puppe, F. Diagnostik-Expertensysteme, Informatik-Spektrum, Springer 1987 (10:293-308)

- Puppe, F. Expertensysteme, Informatik-Spektrum, Springer 1986b 9: 1-13
- Puppe, F. Hybride Diagnosebewertung, GWA1-86, Informatik Fachberichte, Vol.124, (332-342), Berlin Heidelberg, New-York: Springer 1986c
- Voss, H. Representing and Analysing Causal, Temporal and Hierarchical Relations of Devices, Dissertation, Kaiserslautern, Aug. 1986
- Weck, M. Bedeutung von Überwachungs- und Diagnoseverfahren in der Fertigung, Handbuch: Überwachung und Diagnose von Maschinen und Prozessen zum Seminar mit Praktikum in Aachen am 3. und 4. Juni 1985, VDI-Bildungswerk BW 6611 (1-15),
- Weiter, Th. Modallogisch beschriebenes Expertensystem ausgeführt am Beispiel von Ohrenkrankungen, Dissertation, Aachen 1983

