

EINFÜHRUNG IN DIE MEDIZINISCHE INFORMATIK

Verfaßt von: H. Grabner
Erweitert und ergänzt von: E. Schuster
Version 1.0

TEIL 3

1. ANWENDUNGSSYSTEME IN DER MEDIZIN.....	3
1.1 MEDIZINISCHE INFORMATIONSSYSTEME.....	3
1.2 MEDIZINISCHE INFORMATIONSDATENBANKEN	4
1.3 MEDIZINISCHE EXPERTENSYSTEME	4
1.3.1 Überblick über einige medizinische Expertensysteme.....	5
1.3.1.1 Monitoringssysteme	6
1.3.1.2 Laborbefundinterpretationssysteme	6
1.3.1.3 Medizinische Konsultationssysteme	7
1.3.1.4 Elektronische Lehrbücher	7
1.4 MEDICAL DECISION SUPPORT	8
1.5 COMPUTERGESTÜTZTE MEDIZINISCHE DIAGNOSESYSTEME (CAMD-SYSTEME).....	13
1.5.1 Struktur von CAMD-Systemen.....	13
1.5.2 Bewertung von CAMD-Systemen	15
1.5.3 Zukunft von CAMD-Systemen	15
2. EXPERTENSYSTEME (WISSENSBASIERTE SYSTEME).....	16
2.1 EXPERT SYSTEM SHELLS (ENTWICKLUNGSSYSTEME).....	18
2.2 WISSENSVERARBEITUNG \Leftrightarrow INFORMATIONSVERARBEITUNG	19
2.3 KOMPONENTEN EINES EXPERTENSYSTEMS	21
2.3.1 Wissenserwerbskomponente.....	22
2.3.2 Die Wissensbank	23
2.3.3 Wissensrepräsentation (interne Darstellung von Wissen).....	24
2.3.3.1 Formallogik / Prädikatenlogik.....	24
2.3.3.2 Prozedurale Repräsentation / Prozedurales Wissen.....	24
2.3.3.3 Deklarative Formalismen	25
2.3.3.3.1 Semantische Netze.....	25
2.3.3.3.2 Frames	27
2.3.3.4 Hybride Repräsentation.....	28
2.3.4 Die Inferenzkomponente (Problemlösungskomponente).....	29
2.3.4.1 Vorwärtsverkettung (forward chaining, Daten-orientiert).....	30
2.3.4.2 Rückwärtsverkettung (backward chaining, Ziel-orientiert).....	31
2.3.5 Benutzungsoberfläche	32
2.3.6 Erklärungskomponente.....	32
2.4 KLASSEN VON DEN EXPERTENSYSTEMEN	34
2.4.1 Diagnose-Systeme	34
2.4.2 Prognose- Systeme	34
2.4.3 Bildverstehende Expertensysteme	34
2.4.4 Medizinische Diagnosesysteme (D-XPSe).....	35
2.4.5 Statistische und fallvergleichende Diagnostik-Ansätze.....	36
2.4.5.1 Fallvergleichende Diagnostik.....	37

2.4.5.2	Dempster-Shafer-Theorie	38
2.4.6	<i>Assoziative Diagnose-Expertensysteme</i>	41
2.4.6.1	MYCIN/EMYCIN	41
2.4.6.2	EXPERT	44
2.4.6.3	INTERNIST	46
2.4.6.4	HEPAXPERT-I (automatische Interpretation der Hepatitis-A- und -B-Serologie)	49
2.4.6.5	CADIAG-2 (differentialdiagnostische Entscheidungsunterstützung in der Internen Medizin)	53
2.4.6.5.1	Diskussion	59
2.4.7	<i>Modellbasierte Diagnostik-Expertensysteme</i>	60
2.4.7.1	CADUCEUS	61
2.4.7.2	ABEL	62
2.5	STATISTISCHE EXPERTENSYSTEME	64
2.5.1	<i>Diskussion</i>	65

1. ANWENDUNGSSYSTEME IN DER MEDIZIN

Im zunehmenden Maße werden Computersysteme im Krankenhaus, in der Arztpraxis und im Laboratorium nicht nur zur Bewältigung der administrativen und betriebswirtschaftlichen Aufgaben, sondern auch zur Unterstützung der eigentlichen medizinischen Arbeit, der Patientenbehandlung, herangezogen.

3 Arten von Computersystemen zur direkten Unterstützung der Patientenbehandlung:

1.1 Medizinische Informationssysteme

Ziel:

Einheitliche computergestützte Erfassung und Speicherung Darstellung der Patientendaten einer medizinischen Einrichtung (Krankenhaus, Ambulanz, Arztpraxis) mit dem Ziel, sie zu jeder Zeit und an jedem Ort der jeweiligen medizinischen Einrichtung übersichtlich und in einheitlicher Form präsentieren zu können.

Medizinische Informationssysteme tragen wesentlich zur Systematisierung und Standardisierung der ärztlichen Tätigkeit bei und können dadurch eine Verbesserung der Qualität der Patientenbehandlung bewirken. Sie erleichtern

- die Organisation der Behandlung des Patienten
- die Kommunikation der behandelnden Ärzte und des Pflegepersonals über den Patienten („shared care“).

Medizinische Informationssysteme müssen jedoch ein hohes Maß an Flexibilität beinhalten, um den täglichen Anforderungen in der praktischen Medizin gerechtzuwerden. In einige Studien konnten nachgewiesen werden, dass sich aus der Computerisierung der Patientenkrankengeschichten meßbare Verbesserungen bei der Patientenbehandlung ergaben.

Im Krankenhaus sind medizinische Informationssysteme zumeist in Krankenhausinformationssysteme (KIS, HIS) eingebettet; sie sind häufig mit einem eigenen Laborinformationssystem (LIS) gekoppelt. Der Umfang der Erfassung und Speicherung der Patientendaten in einem medizinischen Informationssystem reicht

- von der alleinigen Erfassung und Speicherung der medizinischen Basisdokumentationen
- bis hin zur umfassenden computergestützten Dokumentation aller anfallenden medizinischen Daten (incl. Bilddaten \Rightarrow PACS¹) der Patienten.

Als Kern verfügt ein medizinisches Informationssystem über eine Patientendatenbank, in der die Gesamtheit der patientenspezifischen administrativen und medizinischen Daten gespeichert ist.

Die Patientendatenbank eines medizinischen Informationssystems ist oft Ausgangspunkt für medizinisch-wissenschaftliches Arbeiten, wobei retrospektiv angelegte epidemiologische Fragestellungen im Vordergrund stehen. Computersysteme zur retrospektiven Auswertung einer Patientendatenbank können auch mit Systemen zur computergestützten Durchführung klinischer Studien gekoppelt und damit als breite Hypothesenbasis für prospektive Fragestellungen verwendet werden.

¹ Picture Archiving and Communication System

1.2 Medizinische Informationsdatenbanken

Medizinische Informationsdatenbanken, die über Internet angeboten oder auf CDs erhältlich und subskribiert werden können, stellen umfangreiche Sammlungen medizinischer Informationen dar, die der praktisch tätige Arzt zum Nutzen des Patienten anwenden kann.

Beispiel:

Das vom amerikanischen National Cancer Institute und der National Library of Medicine gemeinsam betriebene System **PDQ** (Physician Data Query).

Der Kern von PDQ ist eine Datenbank, die mehr als 1.000 Therapieprotokolle aus klinischen Versuchsreihen zur Behandlung von an Krebs erkrankten Patienten bereithält. Der Arzt kann über PDQ jene Protokolle auswählen und einsehen, die auf seinen Patienten anwendbar sind und damit einen optimalen Behandlungserfolg erzielen.

Überblick über verschiedene medizinische Informationsdatenbanken:

<http://www.akh-wien.ac.at/medtools/medlist/databases.htm>

1.3 Medizinische Expertensysteme

Ziel:

Unterstützung des Arztes

- beim Finden der richtigen Diagnose
- bei der Auswahl einer optimalen Therapie
- bei der Einschätzung der Prognose
- bei der Patientenführung

Bisher gibt es noch kein Programm, welches all diese Eigenschaften in sich vereinigt. Die derzeitigen Expertensysteme sind zumeist nur für einen kleinen Teilbereich dieser Aufgaben geschaffen worden. Sie dienen jedoch alle der direkten patientenbezogenen Unterstützung der ärztlichen Tätigkeit.

Medizinische Expertensysteme werden zur Verfügung gestellt:

- im Rahmen von medizinischen Informationssystemen
- als Einzelsysteme (⇒ Dateneingabe!)

Expertensysteme sind Computerprogramme, die über das Spezialwissen von Experten verfügen und dieses in einer konkreten Entscheidungssituation anwenden können. Der Kern eines Expertensystems ist eine Wissensbank, in der das Wissen über das jeweilige Fachgebiet in kompakter, oft hochstrukturierter Form gespeichert ist. Der Umfang solch einer Wissensbank kann größer als das Wissen eines einzelnen Experten sein.

Primär dienen Expertensysteme der Unterstützung von Nicht-Spezialisten - das können jedoch sehr wohl "Fachpersonen" sein - , aber auch ein Spezialist kann sich infolge der Mächtigkeit eines Expertensystems von diesem Hilfestellung holen.

Man erwartet sich von medizinischen Expertensystemen eine qualitative Verbesserung in vielen Fragen der diagnostischen und therapeutischen Entscheidungsfindung im konkreten Einzelfall und damit eine Erhöhung der Qualität der Krankenbehandlung.

Die meisten der existierenden medizinischen Expertensysteme sind eigenständige Computersysteme, die sich teilweise schwer in die klinische Praxis integrieren lassen, da die Patientendateneingabe meist zu langwierig ist. Die direkte Übernahme der Patientendaten aus einer zentralen Patientendatenbank eines Krankenhaus- oder medizinischen Informationssystems ist von großem Vorteil und wird bei den Systemen HELP und CADIAG-2 angewandt.

Der wissenschaftlich tätige Arzt, der bei der Durchführung und Auswertung medizinisch-wissenschaftlicher Studien biostatistische Methoden benötigt, kann ebenfalls durch statistische Expertensysteme unterstützt werden, die bei der Auswahl

- der korrekten statistischen Präsentation der Daten
- des korrekten Verfahrens zur statistischen Beurteilung der Studienergebnisse unterstützen können.

1.3.1 Überblick über einige medizinische Expertensysteme

System	Bereich	Entwickler
Monitoringsysteme		
HELP	<ul style="list-style-type: none"> • Krankenhausinfektion • Antibiotika-Therapie • Digoxin-Therapie 	LDS Hospital Salt Lake City
Laborbefund-Interpretationssysteme		
PUFF	Lungenfunktionsdiagnostik	Pacific Medical Center San Francisco
SPE	Serum-Protein- Elektrophorese	Rutgers University New Jersey
Konsultationssysteme		
QMR	Interne Medizin ²	University of Pittsburgh School of Medicine
AI/RHEUM	Rheumatologie ³	National Library of Medicine Bethesda
ONCOCIN	Klinische Onkologie	Stanford University School of Medicine
DXplain	Gesamte Medizin	Harvard Medical School Boston
Elektronische Lehrbücher der Medizin		
RECONSIDER	Gesamte Medizin ⁴	University of California San Francisco

Tabelle 1: Expertensysteme in der Medizin

² Bankowitz, R.A., McNeil, M.A., Challinor, S.M., Parker, R.C., Kapoor, W.N. & Miller, R.A. (1989): *A Computer-Assisted Medical Diagnostic Consultation Service*. Annals of Internal Medicine 110, 824-832.

³ Kingsland III, L.C., Lindberg, D.A.B. & Sharp, G.C. (1983): *AI/RHEUM - A Consultant System for Rheumatology*. Journal of Medical Systems 7, 221-227.

⁴ Nelson, S.J., Blois, M.S., Tuttle, M.S., Erlbaum, M., Harrison, P., Kim, H., Winkelmann, B. & Yamashita, D. (1985): *Evaluation RECONSIDER - A Computer Programm for Diagnostic Prompting*. Journal of Medical Systems 9, 379-388

1.3.1.1 Monitoringssysteme

Monitoringssysteme werden zum kontinuierlichen Monitoring medizinischer Daten verwendet. Zeigen sich Daten außerhalb des medizinisch gewünschten Bereiches, so werden sie dem verantwortlichen Arzt oder der betreuenden Schwester angezeigt. Die Monitorfunktionen werden von Programm-Alarmmoduln wahrgenommen, die dann aktiviert werden, wenn bestimmte pathophysiologische Bedingungen in den Patientendaten erfüllt sind, die als Regeln oder andere Entscheidungskriterien in der Datenbank des Expertensystems gespeichert wurden.

Beispiel:

HELP (Health Evaluation through Logical Processing) ist ein integriertes Krankenhaus-informationssystem mit eingebauter medizinischer Entscheidungslogik.

Im praktischen Einsatz sind Alarmmodule zum computergestützten Monitoring von

- sich häufenden Krankenhausinfektionen⁵
- Patienten mit Antibiotika-Therapie⁶
- Patienten mit mit Digoxin-Therapie⁷

1.3.1.2 Laborbefundinterpretationssysteme

Laborbefundinterpretationssysteme werden verwendet zur

- Interpretation von Laborbefundergebnissen
- Erstellung zusammenfassender Laborergebnisberichte

Beispiel:

- **PUFF** (Pulmonary Function test results")
System zur automatischen Lungenfunktionsdiagnostik⁸
- **SPE** (Serum Protein Electrophoresis)
System zur automatischen Interpretation von Serum-Elektrophorese-Tests⁹
- **HEPAXPERT-I** (HEPAtitis eXPERTe, Version 1)
System zur automatischen Interpretation der Hepatitis-A- und -B-Serologie.

⁵ Evans, R.S., Gardner, R.M., Bush, A.R., Burke, J.P., Jacobson, J.A., Larson, R.A., Meier, F. & Warner, H.R. (1985): *Development of a Computerized Infectious Disease Monitor*. Computers and Biomedical Research 18, 103-113.

⁶ Pestotnik, S.L., Evans, R.S., Burke, J.P. Gardner, R.M. & Classen, D.C. (1990): *Therapeutic Antibiotic Monitoring Surveillance Using a Computerized Expert Systems*. The American Journal of Medicine 88, 43-48.

⁷ White, K.S., Lindsay, A., Pryor, T.A., Brown, W.F. & Walsh, K. 1984: *Application of a Computerized Medical Decision-Making Process to the Problem of Digoxin Intoxication*. JACC 4, 571-576.

⁸ Aikins, J.S., Kunz, J.C., Shortliffe, E.H. & Fallat, R.J. (1983): *PUFF An Expert System for Interpretation of Pulmonary Function Data*. Computers and Biomedical Research 16, 199-208.

⁹ Weiss, S.M., Kulikowski, C.A. & Galen, R.S. 1983: *Representing Expertise in a Computer Program. The Serum Protein Diagnostic Program*. The Journal of Clinical Laboratory Automation 3, 383-387.

1.3.1.3 Medizinische Konsultationssysteme

Medizinische Konsultationssysteme dienen der Konsultation durch den Arzt im konkreten Einzelfall. Der Arzt wird ein solches System als Hilfsmittel heranziehen, wenn es ihm

- die gestellte Diagnose oder ausgewählte Therapie entweder bestätigen oder
- neue diagnostische oder therapeutische Möglichkeiten aufzeigen kann.

Ein Konsultationssystem kann

- auf seltene Diagnosen hinweisen
 - Diagnosen begründen
 - Vorschläge zur weiteren Untersuchung des Patienten geben
 - getroffene Entscheidungen auf Vollständigkeit überprüfen
-
- **QMR** ist das derzeit größte Konsultationssystem mit etwa 600 Krankheitsprofilen aus dem Gesamtbereich der Internen Medizin.
Eine Reihe von Teststudien bestätigen die Brauchbarkeit des Programms.
 - **AI/RHEUM** (Artificial Intelligence/RHEUMatology) enthält einen Kriterienkatalog für etwa 30 rheumatologische Diagnosen und ist für Ärzte ohne spezielle Ausbildung in der Rheumatologie gedacht¹⁰.
 - **ONCOCIN** (ONCOlogy und MYCIN) dient zur Therapieüberwachung und -korrektur in einer onkologischen Ambulanz¹¹ und wird mit gutem Erfolg eingesetzt.
 - **DXplain** ("eXplained by Disease") enthält die Beschreibungen einer Vielzahl von Krankheiten und wird über ein landesweites Computernetz angeboten¹².
 - **CADIAG-2** zur Unterstützung der Differentialdiagnose in der Internen Medizin (für den Bereich der Rheumatologie und Teilbereiche der Gastroenterologie); mit mehreren hundert Krankheitsfällen getestet.

1.3.1.4 Elektronische Lehrbücher

Elektronische Lehrbücher der Medizin sind Computersysteme, in denen bewußt der Lehrbuchcharakter erhalten wurde. Sie dienen als elektronische Nachschlagewerke für Symptome, Krankheiten, Therapien, ggf. Kontraindikationen und Nebeneffekte von Therapien usw.

Beispiel:

RECONSIDER wurde als Diagnose-Anzeige-System entwickelt, das auf Grund von eingegebenen Symptomen alle möglichen Diagnosen ausgibt. Es enthält mehr als 2.500 Krankheitsbeschreibungen, die dem von der American Medical Association herausgegebenen Buch CMIT¹³ entnommen wurden.

¹⁰ Kingsland III, L.C., Lindberg, D.A.B. & Sharp, G.C. (1983): *AI/RHEUM - A Consultant System for Rheumatology*. Journal of Medical Systems 7, 221-227.

Kingsland III, L.C., Lindberg, D.A.B. & Sharp, G.C. (1986): *Anatomy of a Knowledge-Based Consultant System AI/RHEUM*. M.D. Computing 3, 18-26.

¹¹ Shortliffe, E.H., Scott, A.C., Bischoff, M.B., Campell, A.B., van Melle, W. & Jacobs, C.D. 1984: *An Expert System for Oncology Protocol Management*. In Buchanan, B.G. & Shortliffe, E.H. (Eds.) *Rule-Based Expert Systems - The MYCIN Experiments of the Stanford Heuristic Programming Project*. Addison-Wesley Publishing Company, Reading/Massachusetts, 653--665.

¹² Hupp, J.A., Cimino, J.J., Hoffer, E.P., Lowe, H.J. & Barnett, G.O. 1986: *DXplain - A Computer-Based Diagnostic Knowledge Base*. In Proc. MEDINFO 86, Elsevier Science Publishers B.V. (North-Holland), 117-121

¹³ Gordon, B.L. (Ed.) 1971: *Current Medical Information and Terminology*. American Medical Association, Chicago.

1.4 Medical Decision Support

Das Interesse der Menschen an

- der medizinischen Entscheidungsfindung
 - der Etablierung von Richtlinien ("formalen" Anweisungen, guide lines) für Ärzte
- läßt sich bis in das Altertum zurückverfolgen. Die ältesten diesbezüglichen Dokumente stammen aus der Neu-Babylonischen Zeit (um 650 v.Chr.); sie geben Ärzte Ratschläge für Untersuchungen, Diagnosen und Prognosen von Krankheiten.

So wurden Tafeln gefunden, auf denen konkrete Handlungsanweisungen für das alltägliche Leben geschrieben sind. Darunter fanden sich auch Anweisungen zur medizinischen Untersuchung und zur Diagnose und Prognose von Krankheiten.

(leider unvollständige) Beispiele aus J.Wilson: Two Medical Texts from Nimrud (1956):

"When you are about to examine a sick man ..."
"If after a day's illness, he begins to suffer from headache ..."
"If a mother conceives again, her scalp and forehead becoming yellowish in colour ..."
"If ... head he will die suddenly"

In neuerer Zeit hat man erkannt, daß Ärzte von Computern im diagnostischen und therapeutischen Prozeß unterstützt werden können.

Die Entwicklung von computergestützten Diagnosemethoden und medizinischen Entscheidungshilfen begann in den Jahren 1958/59 mit

- dem Vorschlag von LIPKIN & HARDY zur computergestützten Differentialdiagnose von Blutkrankheiten
- der bedeutenden Arbeit über "Reasoning Foundations of Medical Diagnosis" von LEDLEY & LUSTED.

Seither wurde eine große Anzahl von Computerprogrammen, die Diagnose- und Therapievorschlüsse unterbreiteten, entwickelt und praktisch eingesetzt.

Die erste Generation von Systemen zur medizinischen Entscheidungsunterstützung erschien in den späten 50-iger Jahren. Diese Systeme basierten hauptsächlich auf Entscheidungsbäumen und Wahrheitstabellen. Später erschienen auf statistischen Methoden basierte Systeme, gefolgt von Expertensystemen.

Die frühesten Modelle blieben jedoch nur Prototypen.

Bis zu Ende der 60-iger Jahre lag das Schwergewicht bei Erstellung von Computersystemen zur medizinischen Entscheidungsunterstützung auf der Anwendung mathematischer Methoden. Zur Anwendung kamen hauptsächlich probabilistische Verfahren:

- Verfahren, die sich auf das Bayes-Theorem stützten
- diskriminanzanalytische Verfahren
- faktorenanalytische Verfahren
- Verfahren der numerischen Mustererkennung
- Verfahren der Entscheidungsanalyse
- lernende abstrakte Neuronale Netze

Die Anwendung dieser Systeme war zumeist auf kleine medizinische Gebiete beschränkt, die maximal 20 - 30 unterschiedliche Entscheidungsklassen enthielten. Diese Verfahren lassen sich nach wie vor gut in überschaubaren Bereichen einsetzen. Sie verlangen aber als Grundlage präzise Angaben über die Häufigkeit des Auftretens von Symptomen, von Kombinationen von Symptomen, von Krankheiten; oder sie erfordern repräsentative Fallstichproben oder benötigen genaue Informationen über kausale und/oder statistische (bedingte oder unbedingte) Zusammenhänge zwischen den medizinischen Objekten.

In der praktischen Medizin ist dies jedoch zumeist nicht gegeben - man denke nur an seltene Erkrankungen -, wie dies das nachfolgenden Zitat anschaulich beschreibt:

.. my friends who are expert about medical records tell me that the attempt to dig out from even the most sophisticated hospital's record the frequency of association between any particular symptom and any particular diagnosis is next to impossible - and when I raise the question of complexes of symptoms, they stop speaking to me.

For another thing, doctors keep telling me that diseases change, that this year's flu is different from last year's flu, and so that symptom-disease records extending back over time are of very limited usefulness.

All these arguments against symptom-disease statistics as a basis of diagnosis are perhaps somewhat overstated. Where such statistics can be obtained and believed, obviously they should be used. But I argue that usually they cannot be obtained, and even in those instances in which they have been obtained, they may not deserve belief.

Seit Ende der 60-iger / Anfang der 70-iger Jahre werden medizinische Expertensysteme unter Berücksichtigung der oben erwähnten Problematik entwickelt. So erfolgten keine rein numerisch-mathematische Behandlung des Problems mehr, sondern man wählte einen symbolisch-logischen Lösungsansatz, wobei man auch **heuristische Lösungsmethoden** zuließ.

Unter Heuristiken versteht man Wissens Elemente und Problemlösungsmethoden, die sich auf keine fundierte Theorie stützen können - nicht weil man eine theoretische Fundierung ablehnt, sondern weil der medizinische Erkenntnisstand diese noch nicht erbracht hat -, die aber in der Praxis standardmäßig erfolgreich eingesetzt werden.

Mit diesen medizinischen Expertensystemen wurde es möglich,

- Entscheidungsunterstützung für medizinische Anwendungsbereiche mit unscharfem, unsicherem und unvollständigem Wissen zu entwickeln
- komplexe Gebiete mit bis zu mehreren hundert Entscheidungsklassen zu bearbeiten und gute Ergebnisse zu erzielen

Beispiel:

QMR (Quick Medical Reference) mit fast 600 Diagnosen

Auf Grund der Erfahrungen mit früheren Programmen zur medizinischen Entscheidungsunterstützung fordert man für medizinische Expertensysteme folgende zusätzliche Eigenschaften.

- Möglichst vollständige inhaltliche Transparenz
Dies ist eine unabdingbare Voraussetzung für eine breitere medizinische Akzeptanz. Medizinische Expertensysteme sollen also ihre gefundene Problemlösung durch Angabe des benutzten Wissens erklären können.
- Flexibilität
Die dem Programm zugrundeliegenden einzelnen Wissens Elemente müssen leicht ergänzt, geändert oder gelöscht werden können.
- Benutzerfreundlichkeit

Der behandelnde Arzt trägt Verantwortung für den Patienten und konsequenterweise für die Diagnose. Computer werden nicht die alleinige Verantwortung für die Diagnose übernehmen. Jene Teile der Diagnose, welche sich mit einzigartigen und individuellen Problemen (die nicht generalisiert werden können) des Patienten befassen, können nicht einer Maschine übergeben werden. Der wissenschaftlich basierte Teil der Diagnose kann jedoch dem Computer überlassen werden. Klinisches Wissen und Erfahrung der Ärzte können strukturiert und zumindest partiell und bis zu einem gewissen Grad einsatzbereit im Computer gemacht werden.

Herausforderung:

Wie kann man medizinisches Wissen so formalisieren, dass es zur **Entscheidungsunterstützung** (decision support) verwendet werden kann?

Definition:

Clinical Decision Support Systems (CDSSs) are knowledge systems which use two or more items of patient data to generate case-specific advice.¹⁴

Decision Support Modelle

- Quantitativer Decision Support:
basiert auf wohl-definierten statistischen Methoden, benützt training sets von Patienten Daten. Dabei werden die Merkmale hauptsächlich durch statistische Methoden selektiert. A-priori Wahrscheinlichkeiten für das Auftreten einer bestimmten Krankheit sind in die statistischen Modelle integriert (z.B. Bayes' rule)
- Qualitativer Decision Support:
Benützt Merkmale, die von Medizinern allgemein vorgeschlagen werden, basierend auf klinischen Studien. Benützen normalerweise "symbolic reasoning" Methoden (z.B. Forward/Backward Reasoning).
Wissensbasierte Systeme sind typischerweise auf qualitativen Methoden basiert.

Hauptkomponenten von klinischen DSS:

- Medizinisches Wissen:
variiert von System zu System
- Patientendaten:
werden interpretiert vom medizinischen Wissen, das im DSS enthalten ist
- Fall-spezifischer Rat:
Resultat der Anwendung des med. Wissens auf die Patientendaten

CDSS wurden für verschiedene Domänen entwickelt:

- für sehr spezielle diagnostische Probleme
Beispiel:
De Dombal's System für akuten Bauchschmerz
(frühe 70er, basiert auf statistischen Methoden, 13 verschiedene Diagnosen);
- zur Interpretation von EKGs;
- für große Bereiche klinischer Behandlung
Beispiel:
INTERNIST (später QMR) mit hunderten von Diagnosen versucht, die gesamte Interne Medizin abzudecken;

CDSS unterscheiden sich auch im Hinblick auf die Entscheidung, die sie unterstützen:

- diagnostische Entscheidungen
(schlagen eine Diagnose auf Basis der Patientendaten und Testergebnissen vor)
- therapeutische Entscheidungen

Für vollkommen automatisierte Systeme sind selten;
in der Praxis findet man i. R. "Computer-assistierte" Systeme und kaum solche, die ohne jegliches menschliche Zutun operieren.

¹⁴ J.H. van Bommel (Hrsg.), *Handbook of medical informatics*, Springer, 1997.

Beispiele:

- ACORN chest pain decision aid¹⁵:

Ursprüngliches Ziel:

Reduzierung der Zahl der Untersuchungen von Patienten, die mit Brustschmerzen in die Notaufnahme kamen; als man bemerkte, dass die Tests ohnehin nicht exzessiv waren, wurde das Ziel verlagert:

Neues Ziel:

Reduktion der an die Herzüberwachung anzuschließenden Patienten.

Ursprünglich: Bayes'scher Ansatz; eine große Datenbank mit Patienten-Befunden und gold Standards wurde erstellt. Dabei stellte man fest, daß zuviele Risikopatienten irrtümlicherweise wieder nach Hause geschickt wurden;

Nun wurde nicht nur das Ziel erneut verlagert, man wechselte auch zu einem Regel-basierten System. Nach sorgfältiger Evaluierung der Funktionen von ACORN wurde es in einer Notaufnahme installiert mit dem Ziel, das medizinische Personal auf Risikopatienten aufmerksam zu machen....

Gründe für den darauffolgenden Fehlschlag waren u.a.:

- primitives Interface, welches zu Dateneingabe-Fehlern führte
- das System war viel zu langsam für den Routineeinsatz in der Notaufnahme
- das System gab zu detaillierte Ratschläge...

- Acute Abdominal Pain¹⁶:

8-center study (250 Ärzte, 16.737 Patienten):

Nach Einführung des Systems

- stieg die Genauigkeit der Diagnose von 46% auf 65%
- die negative Laparotomierate sank um die Hälfte
- die Rate der perforierten Appendices sank um die Hälfte
- die Sterblichkeitsrate sank um 22%.

Auch der sogenannte "Checklist-Effekt" wurde demonstriert:

die bloße Einführung von Formularen für die Aufnahme der Symptome resultierte in einer wesentlichen Verbesserung:

Die diagnostische Präzision stieg von 46% auf 57%.

- anderes Bsp:

O. Aase, *Clinical experience with a Decision Support Computer Program Using Bayes' theorem to Diagnose Chest Pain Patients*, Cardiology 1999; 92: 128-134.

Ergebnisse

- B. C. Delaney et al., *Changing the doctor-patient relationship: Can computerised decision support systems deliver improved quality in primary care?*, BMJ 1999, 319:1281.
- M. E. Johnston et al., *A critical appraisal of research on the effects of computer-based decision support systems on clinician performance and patient outcomes*. Ann Intern Med 1994;120:135-42.
- G. Marckmann, *Recommendations for the Ethical Development and Use of Medical Decision-Support Systems*, MedGenMed, June 20, 2001.
- F.T. de Dombal, *Computer-assisted Diagnosis in Europe*, N Engl J Med 1994 ; 331 (18) : 1238.
- E. S. Berner et al., *Performance of four computer-based Diagnostic Systems*, N Engl J Med 1994, 330 (25):1792-96.

Erwartungen

- Verbesserung der Qualität der Behandlung durch

¹⁵ H.A. Heathfield, J. Wyatt, *Philosophies for the Design and Development of Clinical Decision-support models*, Methods Inf Med 1993; 32(1) : 1-8.

¹⁶ J.H. van Bommel (Hrsg.), *Handbook of medical informatics*, Springer, 1997

- Bereitstellung von genauen, effektiven und zuverlässigen Diagnosen bzw. Behandlungen
- Vermeidung von Fehlern aufgrund von unzureichendem Wissen des Arztes
- Reduktion der Kosten im Gesundheitswesen durch spezifischere und schnellere Diagnose und Reduktion von Konsultationen eines Spezialisten

Befürchtungen

- Ärzte werden mehr Zeit damit verbringen, Daten zu analysieren als ihre Patienten zu verstehen;
- Menschen, die starke Schmerzen und persönliche Konflikte haben, werden zu document files, interessanten Fällen oder überhaupt nur zu Nummern.

Realität

- CDSS sind ein rasch weiterentwickelndes und unreguliertes Gebiet
- Immer mehr Systeme werden entwickelt, aber nur wenige sind wirklich im Einsatz
- Viele wurden zumindest teilweise zu Lernsystemen umfunktioniert
- CDSS begünstigen eine Krankheits-zentrierte Medizin, welche die psychologische und soziale Individualität des Patienten nicht erfassen kann.
- Es gibt kaum Patienten-orientierte Studien
- Einige Studien legen nahe, dass CDSS die Leistung von Ärzten verbessern
- CDSS haben großes Potential für die Erstversorgung (konnten dies aber noch nicht zur Gänze entfalten, hauptsächlich weil die Bedürfnisse der Benutzer nicht adäquat untersucht wurden.)
- Ärzte haben oft nicht die Zeit, die benötigt würde, solche Systeme zu benützen.
- ...

Regionale Unterschiede

- UK hat den am meisten computerisierten Gesundheitssektor der Welt;
Hauptabsicht: Unterstützung bei klinischen Entscheidungen;
dennoch: Verwendung von Computerised DSS nicht commonplace
(Hauptgründe:
 - Fehlen nationaler Standards
 - adäquate Untersuchung der Benutzer-Bedürfnisse
 - Fülle verschiedener Systeme, die nicht miteinander kommunizieren)
- Die meisten Artikel zu computer-unterstützten Diagnose findet man in den USA, gefolgt von Europa, wo der Schwerpunkt eher auf spezifischen klinischen Problemen liegt.
- In Asien findet man bei weitem nicht so viel, aber mehr als in Afrika, Südamerika und Australien zusammen.

1.5 Computergestützte medizinische Diagnosesysteme (CAMD-Systeme)

Diagnose:

Feststellung, Bestimmung einer körperlichen oder psychischen Krankheit (durch den Arzt)¹⁷

Das Interesse an computergestützten medizinischer Diagnose (CAMD) ergibt sich aus der in den letzten Jahren steigenden Anzahl zu stellender Diagnosen bei gleichzeitig beinahe gleichbleibender Anzahl ausgebildeter Mediziner.

Traditionell, d.h. ohne Computerunterstützung durchgeführte Diagnosen sind

- personal-, zeit- und/oder kostenintensiv
- teilweise unzuverlässig
- für den Patienten eine starke Belastung, unter Umständen sogar gefährlich

CAMD-Systeme sollen diese Nachteile umgehen oder vermindern. Sie haben folgende

Ziele¹⁸:

- Zeit oder Kosten zu sparen
- Schneller eine höhere Qualität der Diagnose zu erreichen
- Vorteile der Telemedizin in Krankenhäusern zu nutzen
- Expertenwissen via Netzwerk-Kommunikation zu erhalten
- Allen Patienten dieselbe Servicequalität durch ständig verfügbare Expertennetzwerke zur Verfügung zu stellen
- Automatische und frühe Erkennung von Anomalien
- Zentrale Wissensbasen zu schaffen

Unterstützung des Mediziners durch CAMD-Systeme in Form von:

- Entscheidungsunterstützung
(Einschränkung auf engeren Kreis, Hinweis auf leicht Übersehbares)
- automatische Diagnose ("Ersetzung" des Experten, second opinion)
- Vorschläge zur weiteren Untersuchung zur Einengung des Befundes
- Hervorhebung von Regions of Interest (ROI), Segmentierung anatomischer Strukturen¹⁹

1.5.1 Struktur von CAMD-Systemen

Input in CAMD-Systeme

- Krankheitssymptome
- Rohdaten (z.B. Röntgen-, US- oder CT- Bilder), aus denen sie Symptome ableiten

Der Output variiert

- von bloßen Hinweisen auf mögliche Krankheitsbilder, z.B. Knötchen auf Bildern von Brustuntersuchungen
- bis hin zu einer vollständigen Diagnose samt Behandlungsvorschlag

Rohdaten als Input müssen vom System zuerst verarbeitet werden. Im Falle von Bildern werden Methoden aus der Mustererkennung (z. B. Filter, Segmentierungsalgorithmen, feature analysis) angewandt, um für die Diagnose relevante Merkmale zu extrahieren.

Beispiel (Brustkrebsuntersuchungen):

Merkmale sind z. B. Knötchen oder Regionen mit krankhafter Veränderung des Brustgewebes (Texturanalyse).

¹⁷ www.duden.com

¹⁸ Michael Fatzi: Diplomarbeit *Computer Aided Diagnosis of Breast Cancer*

¹⁹ <http://www.deustech.com/products.html>

Manche CAMD-Systeme arbeiten nicht nur mit den direkten Input-Daten, sondern auch mit Daten

- aus früheren Diagnosen (sowohl vom System selbst gestellt als auch von Ärzten)
- über Krankheiten und deren Symptome.

Sie verwenden diese Wissensbank, um Diagnosen mit höherer Sicherheit und Genauigkeit stellen zu können.

Je nach Art des Systems ist die Wissensbank:

- eine relationale Datenbank
- ein wissensbasiertes System
- ein neuronales Netz

Die Entscheidungsfindung erfolgt über

- **Suchalgorithmen**

Basierend auf einer relationalen Datenbank werden Datensätze mit bestimmten Eigenschaften gesucht und zurückgeliefert.

Beispiel:

Alle Krankheiten, die bestimmte Symptome aufweisen.

Beispiel:

Das System *Aide au Diagnostic Médical*²⁰ stellt eine Datenbank mit Krankheiten, Symptomen und Behandlungsmethoden im Internet zur Verfügung.

- **Regelbasierte Inferenzmechanismen**

In logikorientierten Wissensbasen werden Daten in Form von Fakten und Regeln gespeichert. Durch Inferenz können "neue" Daten, in diesem Fall Diagnosen, abgeleitet werden.

Vorteil:

Umgang mit unvollständigem Wissen (z. B. fehlendes Alter des Patienten). Nicht bekannte (fehlende) Daten können unter Umständen aus den bestehenden abgeleitet werden, was bei einfachen relationalen Datenbanken nicht so einfach möglich ist.

Beispiel

MYCIN²¹, das Mitte der 70er Jahre entwickelt und eingesetzt wurde.

- **Neuronale Netze**

Diese Systeme sind angelehnt an die Neuronen-Struktur des Hirns. Neuronen sind miteinander verknüpft und leiten Signale an ihre Nachbarneuronen weiter, wenn ein bestimmter Reizschwellwert überschritten wird. Durch Rückkopplungsmechanismen entwickeln sich die Gewichte der Verbindungen mit jedem Input. Die momentane Struktur bestimmt die Reaktion auf bestimmten Input.

Neuronale Netze werden über Lernbeispiele "trainiert".

Nach der Trainingsphase sind sie einsatzbereit und lernen mit jedem Einsatz weiter.

Neuronale Netze werden meist dort verwendet, wo es keine algorithmische Lösung oder gibt oder eine solche zu komplex wäre²².

Da derzeitige neuronale Netze nur geringe Input-Datenmengen parallel verarbeiten können, müssen in der Regel die Inputdaten auf wenige signifikante Parameter heruntergebrochen werden.

Beispiel:

Das System von INSERM²³ verwendet einen einzigen Aktivierungswert zwischen 0 und 1, der aus Bildern der Größe 32x32 berechnet wird.

- **Fuzzy Logic**

Dieses Modell basiert auf der Theorie der unscharfen Mengen (Fuzzy-Set-Theorie), so daß auch Befunde im Grenzbereich zwischen „normal“ und „pathologisch“ in die Entscheidung miteinbezogen werden können²⁴. Wertebereiche, z. B. für Fieber oder Leberwerte, werden in mehrere, oft trapez- oder dreieckförmige Mengen eingeteilt. Die Ermittlung des Ergebnisses erfolgt durch Inferenz mit Regeln, die in einer

²⁰ <http://www.med.univ-rennes1.fr/plaq/proj/adm.html>

²¹ Uwe Egly, Thomas Eiter, Hans Tompits: Skriptum zur Vorlesung *Wissensbasierte Systeme*

²² <http://www.emsl.pnl.gov:2080/proj/neuron/briefs/nose.html>

²³ www.inserm.fr

²⁴ http://www.akh-wien.ac.at/mes/abstracts/ste99_e.html

Regelbasis abgelegt sind. Das Ergebnis ist in der Regel eine unscharfe Menge (z. B. "500 - 700 mg Paracetamol"), die nach bestimmten Methoden defuzzifiziert werden muss ("genau 570 mg")²⁵.

1.5.2 Bewertung von CAMD-Systemen

Die Bewertung von CAMD-Systemen erfolgt durch Vergleich ihrer Diagnosen mit denen, die von Experten gestellt worden sind. Oft stehen Datenbanken mit Bildern und/oder Meßwerten von normalen und pathologischen Fällen zur Verfügung, die als Testfälle herangezogen werden können. Um wirklich aussagekräftige Ergebnisse liefern zu können, sollte jedes System mit zumindest mehreren hundert Testfällen getestet werden. Systeme, die in 10 Testfällen eine 90-prozentige Erfolgsrate haben, sind sehr wahrscheinlich für 10.000 Fälle nicht gleich erfolgreich.

Bei Systemen, die zur Aufgabe haben, Vorhandensein einer bestimmten Krankheit festzustellen, ist es wichtig, anzugeben

- wieviele Fälle richtigerweise
- wieviele Fälle fälschlicherweise positiv/negativ diagnostiziert wurden.

Viele Systemhersteller von CAMD-Systemen geben die Effizienz ihres Systems durch den Az-Wert an, d.h. die Fläche unter der Receiver Operator Characteristic (ROC)-Kurve.

1.5.3 Zukunft von CAMD-Systemen

CAMD-Systeme lassen sich so erweitern, dass sie Behandlung und benötigte Medikamente selbst vorschlagen.

Mit der steigenden Verfügbarkeit von CAMD-Systemen wird die Verpflichtung wachsen, sie zu verwenden. Spätestens wenn Kläger nachweisen können, dass bei Konsultierung eines CAMD-Systems eine andere Behandlung mit weniger nachteiligen Folgen vorgeschlagen worden wäre, wird der Gebrauch von CAMD-Systemen obligatorisch sein²⁶.

²⁵ Gerhard-Helge Schild, Wolfgang Kastner: *Prozeßautomatisierung*, Springer-Verlag

²⁶ Metzger MC. Legal implications of computer-aided medical diagnosis. *Journal of Legal Medicine*, 1988 Jun 9(2):313-328

2. EXPERTENSYSTEME (WISSENSBASIERTE SYSTEME)

Die Entwicklung von Expertensystemen, die über Wissen und über Problemlösungsmethoden verfügen, war ein praktisches Ergebnis der interdisziplinären Forschung auf dem Gebiet der Artificial Intelligence, die etwa Mitte der 50er Jahre initiiert wurde.

1956	John McCarthy ²⁷ , Marvin Minsky, Nathaniel Rochester und Claude Shannon organisieren die berühmte <u>Konferenz von Dartmouth</u> . John McCarthy prägt den Begriff <u>Artificial Intelligence</u> (Künstliche Intelligenz).
1957	Allan <u>Newell</u> und Herbert <u>Simon</u> stellen den ersten Prototyp ihres General Problem Solvers (GPS) fertig. GPS ist ein Programm, das über allgemeine Methoden und Strategien zur Lösung von Problemen verfügt, die von einem <u>speziellen Problem weitgehend unabhängig sind</u> .

Erste Ansätze finden sich schon im Altertum.

Allgemein ist "Künstliche Intelligenz":

.. the part of computer science concerned with designing intelligent computer systems, that is, systems that exhibit the characteristics we associate with intelligence in human behavior - understanding language, learning, reasoning, solving problems, and so on.
(A.Barr und E.Feigenbaum: The Handbook of Artificial Intelligence 1981)

Die grundsätzliche Erreichbarkeit dieses Zieles ist nach wie vor Gegenstand wissenschaftlicher Debatte, jedoch haben Erkenntnisse aus diesem Bereich wesentlich zur Entwicklung praktisch brauchbarer Computersysteme zur Entscheidungsunterstützung in den verschiedensten Bereichen beigetragen.

2 Hauptaspekte (der Künstlichen Intelligenz):

- Schaffung intelligenter Systeme
d. h. praktisch einsetzbare Systeme, die Aufgaben lösen, die bisher dem Menschen vorbehalten waren
- Simulation menschlichen Verhaltens mittels Computer
Sie soll Aufschlüsse über das menschliche Denken, Handeln und die menschliche Existenz liefern.
⇒ Nachahmung von natürlichen Systemen
Vgl. Fliegen (Daedalus und Ikarus):
Vogelfedern so anordnen wie bei Vögeln ⇒ man kann fliegen ⇒ **Irrtum**

Künstliche Intelligenz = Rekonstruktion intelligenten Verhaltens lebender Systeme anhand von Modellen/Theorien.

Definition:

Expertensysteme sind **wissensbasierte** Computersysteme, in denen

- das Sach- und Erfahrungswissen von Experten gespeichert ist

²⁷ Erfinder der symbolverarbeitenden Programmiersprache LISP, die bis heute (wenn auch in einer Vielzahl von Dialekten) die Standardprogrammiersprache der AI geblieben ist

- **Problemlösungskomponenten** implementiert sind
Das in ihnen gespeicherte Fachwissen von Experten ist abrufbar (Konsultationssysteme) und auf individuelle Problemstellungen anwendbar.

Synonyme

- „*Wissensbasierte Systeme*“
- „*Entscheidungsunterstützungssysteme*“
In Expertensystemen werden logische Schlußfolgerungsregeln bezüglich bestimmter Sachverhalte durchgeführt und ein Ergebnis mit einer bestimmten Wahrscheinlichkeit geliefert.

Symbolische Programmiersprachen wie *LISP* oder *PROLOG* (für: PROgrammieren in LOGik), mit deren Hilfe die ersten Expertensysteme realisiert wurden, ermöglichen flexible und effektive Problemlösungen. Anstelle von Computerinstruktionen werden logische Aussagen formuliert. Dadurch können Expertensysteme rascher und direkt entwickelt werden.

Arten von Expertensystemen:

- eingebettete Systeme (Interpretation von Daten, Messungen)
- interaktive Systeme (Prognose, Diagnose, Planung, Entwurf)

Vorteile von Expertensystemen:

- Allgemein
 - Wissen ist immer verfügbar
 - billig
 - sie machen keinen Fehler (werden nicht müde)
 - lösen Routineprobleme sicher
 - geben Auskunft über den Lösungsweg
 - sind auf große Datenmengen (hoch dimensionale Daten) anwendbar
 - überprüfen alle Hypothesen gleichermaßen
 - erweiterbar
- aus Anwenderumfrage
 - Fachwissen des Unternehmens wird dokumentiert
 - Erhöhung der Transparenz eines Vorgangs, einer Produktion usw.
 - Entlastung der Experten
 - Expertensysteme wurden vorzugsweise für Routineprobleme eingesetzt
 - ein Teil der Expertenkompetenz wird für viel mehr Mitarbeitern im Unternehmen zugänglich, was zu einer weiteren Entlastung des Experten führt.
 - In keinem Fall wurde eine Experte ersetzt

Nachteile:

- Wissensakquisition = Flaschenhals
Abhilfe: Wissen aus Datenbanken
- Wie soll Wissen dargestellt werden
(Es gibt kein Kriterium für die Wahl der Darstellungsform)
- Robustheit nicht leicht erreichbar
(d. h. auch für etwas anders gelagerte Probleme)

Der **Einsatz** von Expertensystemen ist **sinnvoll, wenn:**

- menschliche Experten sollen von unnötigen Arbeiten (Routine-Arbeiten) entlastet werden
- Es sind wenige Experten vorhanden
- Es werden viele Beratungen (Expertisen) gebraucht

- Verringerung der Variation in der Qualität von (medizinischen) Befunden und Entscheidungen²⁸
- Es wird Spezialwissen benötigt
- optimales Wissen soll dezentral zur Verfügung stehen
- Expertenwissen soll auch dort zur Verfügung stehen, wo menschliche Experten nicht verfügbar sind (z. B.: entlegene Spitäler)
Alternative: Tele-Medizin
- Transparenz des Lösungswegs ist notwendig (Schulung)

Die **zentrale Frage** bei der Entwicklung von medizinischen Expertensystemen ist, inwieweit sie eine echte Unterstützung der eigentlichen medizinischen Arbeit, der Patientenbehandlung, bieten.

Der Computer fungiert aber "nur" als Werkzeug, wenn auch als äußerst mächtiges. Die endgültige Entscheidung in jeder Phase der Patientenbehandlung obliegt nach wie vor dem verantwortlichen Arzt.

2.1 Expert system shells (Entwicklungssysteme)

Es gibt kommerzielle **Entwicklungssysteme** - sogenannte expert system shells.

Definition:

Ein allgemeiner, wiederverwendbarer Mechanismus zum

- Schlußfolgern
- Erklären / Nachvollziehen eines Lösungsweges
- Dokumentieren
- Führen des Dialogs mit dem Benutzer

heißt Experten-Shell

Ein Shell ist also ein wissensunabhängiger Rahmen zur Erstellung von Expertensystemen. Expert shells sind leere Expertensysteme, die mit dem entsprechenden Wissen für die gewünschte Applikation gefüllt werden müssen. Dies geschieht in Form von Regeln²⁹ und Fakten. Die übrigen Komponenten sind bereits vorhanden.

Beispiel:

- EMYCIN
- EXPERT

Für die Entwicklung solcher Systeme ist mit einem Personalaufwand von 0.5 bis 3 Mannjahren zu rechnen, vorausgesetzt, es existieren bereits fundierte EDV-Kenntnisse.

²⁸ In den USA fallen etwa 650 Millionen medizinische Röntgenbilder pro Jahr an und man schätzt, daß 30% der Anomalien: bei konventioneller Röntgendiagnose unerkant: bleiben. D.h. eine Entlastung des Personals durch Computer wäre wünschenswert.

²⁹ Brauchbare Expertensysteme brauchen ungefähr 500 bis 2.000 Regeln

2.2 Wissensverarbeitung ↔ Informationsverarbeitung

- **Informationsverarbeitung** erfolgt streng nach einem (vorher) festgelegten Programmablauf (Algorithmus). D. h. für jede Situation, die das Programm behandeln soll, wird ein vorgegebener Lösungsweg verfolgt. Der Lösungsweg ist vollständig geplant und jede Abweichung davon stellt einen Fehlerzustand dar.
- In **wissensbasierten Systemen** ist das entscheidungstreibende Wissen vom eigentlichen Programmablauf getrennt. Der Programmablauf ist so stark verallgemeinert, daß er nicht vom Entwickler erstellt werden muß, sondern er existiert in Form eines Schlußfolgerungsmechanismus (Inferenz-Komponente). So werden aus vorgegebenen Daten mittels
 - Faktenwissen
 - Heuristiken
 - vagem Wissen
 - Regelwissen (Ablaufverfahren)selbständig Schlüsse gezogen.

Da der Programmablauf nicht mehr (vollständig) vorhersagbar ist, gibt es eine **Erklärungskomponente**, die dem Benutzer die tatsächliche Ableitungskette verständlich darstellt.

Gleichheit von Expertensystemen und herkömmlichen Softwaresystemen:

- Beide rekonstruieren einen Ausschnitt der Wirklichkeit basierend auf einem Modell dieses Ausschnittes
Folgerungen:
 - Jedes Modell ist weniger als die Wirklichkeit.
In einem Modell geht nicht alles, was in der modellierten Wirklichkeit geht, und umgekehrt.
 - Die Modelle müssen mithilfe von Programmiersprachen vollständig formalisiert werden.
- Beide müssen programmiert, getestet und gewartet werden
Es muß qualifiziertes Personal verfügbar sein, um ein Expertensystem einsatzfähig zu halten.
- Beide sind fehleranfällig
- Beide veraltern nach wenigen Jahren

Kosten (von Expertensystemen):

Entwicklung	250.000.-- Euro
Entwicklungszeit	22 Mann-Monate
Wartung	3-5 Mann-Tage pro Monat
Lebensdauer	8 Jahre
Amortisation	2 Jahre
Benotung	2.5 (nach Schulnotensystem)

Voraussetzungen (für die Machbarkeit eines Expertensystems):

- gut abgegrenztes (d. h. kein Allgemeinwissen erforderlich) und überschaubares Gebiet
- hauptsächlich heuristisches Wissen vorhanden
- Verfügbarkeit von Expertenwissen (kooperative Experten, Fachliteratur)
- gute Testmöglichkeiten
- einfach verbalisierbare Ein-/Ausgabe

Ein **Experte** ist jemand, der eine Ausbildung erhalten hat, in der

- das allgemein bekannte und akzeptierte Fachwissen
- die Art und Weise, wie dieses für das Problemlösen eingesetzt werden soll vermittelt wird.

Der Unterschied zwischen Berufsanfänger und Experten kommt daher, daß die berufsbedingten praktischen Erfahrungen und die intensive Beschäftigung mit dem Gebiet das in der Ausbildung erworbene Wissen reorganisieren und ergänzen.

Beobachtet man einen Experten bei seiner Tätigkeit, so fällt auf:

- Routinefälle werden gleichsam ohne Nachzudenken gelöst.
- In schwierigen Fällen werden sehr schnell nur einige wenige erfolgversprechende Lösungsansätze identifiziert.
- Es fällt dem Experten ausgesprochen schwer, zu beschreiben, wie er Probleme löst. (Er handelt intuitiv, basierend auf seiner früheren Erfahrung in ähnlichen Situationen).

Fähigkeiten menschlicher Experten:

1. Schnelles Finden von Problemlösungen (gute Performance)
2. Erklären der Problemlösung
3. Erweiterung des Wissens durch Lernen
4. Einschätzen, ob ein gegebenes Problem in seinen Kompetenzbereich fällt
5. Berücksichtigen von zahlreichen Ausnahmen zu einer Regel
6. stetiger Leistungsabfall an den Grenzen des Kompetenzbereichs

Fähigkeiten (bestehender) computer-basierte Expertensysteme:

- (1)..... hauptsächlich
(2), (3)..... weniger gut
(4), (5), (6)..... überhaupt nicht

Schwierigkeiten (bei der Erstellung von Expertensystemen)

- Der Experte selbst wehrt sich dagegen
- Mangelhafte Entwicklungswerkzeuge
- Shell-Chaos
- Zu hohe Erwartungshaltung an die Expertensysteme ("KI-Systeme sind klüger als die Experten")
- Fehlen von Wissenserwerbswerkzeugen

2.3 Komponenten eines Expertensystems

Die wichtigsten Komponenten eines Expertensystems sind:

- Wissenserwerbskomponente
- Wissensbank
- Inferenzkomponente
- Benutzungsoberfläche (Dialogkomponente)
- Erklärungskomponente

Struktur von Expertensystemen:

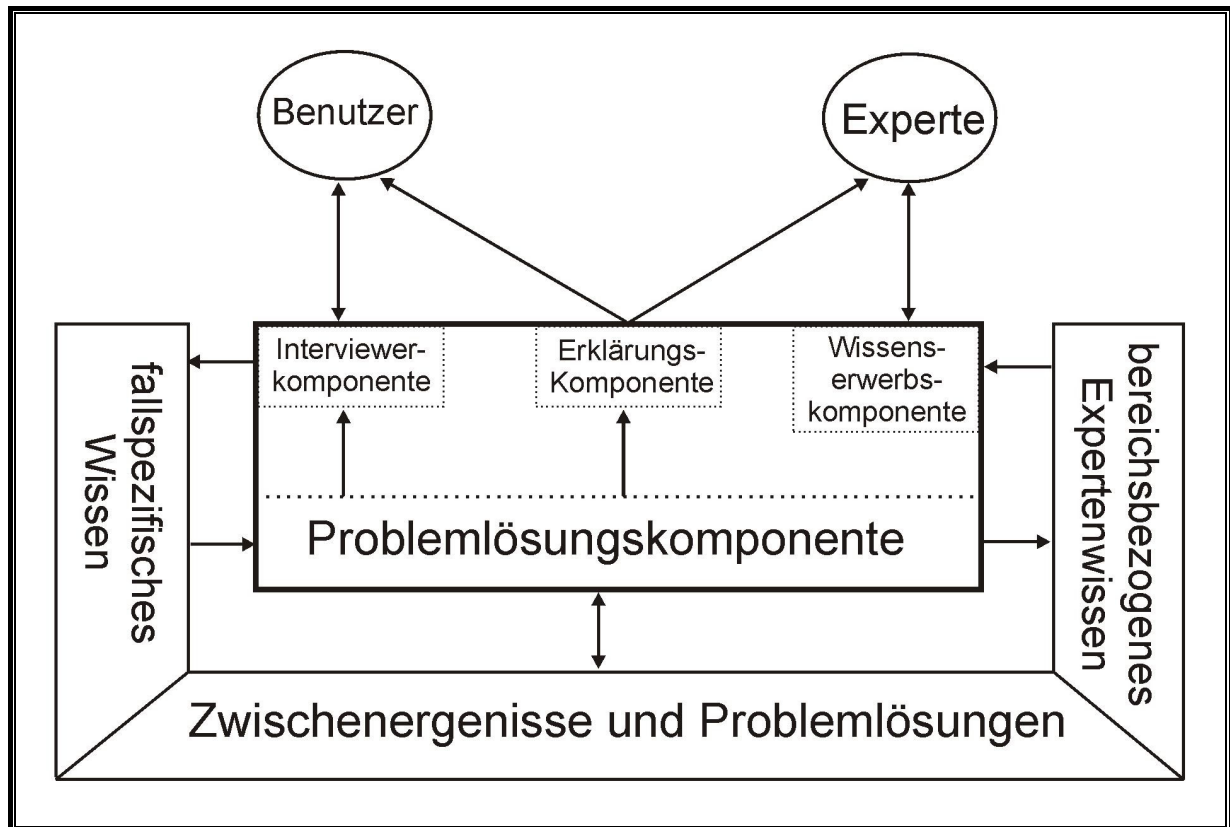


Abb. 1.1. Allgemeine Architektur eines Expertensystems [Puppe (1991) 13]

2.3.1 Wissenserwerbskomponente

Aufgabe:

Erheben, Erweitern und Übertragen des Expertenwissens (Beobachtungswissen, Fachliteraturwissen oder Expertenwissen) aus verschiedenen Wissensquellen in die Wissensbank, d.h.

- Eingeben von Wissen
- Ändern von Wissen
- Ergänzen von Wissen
- Erweitern von Wissen
- Einschränken von Wissen
- Löschen von Wissen

Die Wissenserwerbskomponente soll einen interaktiven Dialog auf einer Ebene ermöglichen, die weit vom Implementierungsniveau entfernt ist.

Die Wissenserwerbskomponente sollte in der Lage sein, das gespeicherte Wissen auf

- Konsistenz
 - Vollständigkeit
- zu überprüfen.

Mögliche Wissensquellen:

- Experten
- Bücher
- Artikel
- Datenbanken

Expertenwissen zerfällt in

- Sachwissen
 - Faktenwissen (Datenbank)
 - Regelwissen (Ablaufverfahren)
- Erfahrungswissen
 - Analogien (Assoziationen)
 - Generalisierungen
 - hypothetische Verknüpfungen (Intuition)
 - Heuristiken (Daumenregeln)
 - "fuzzy logic"³⁰
 - die gesamten erfaßten Erfahrungen

Die Wartung der Wissensbank ist Teil der Wissenserwerbskomponente. Da sich das Expertenwissen ständig verändert, müssen diese Veränderungen in die Wissensbank übernommen werden; diese Wartungen müssen schon beim Entwurf der Wissensbank berücksichtigt werden.

³⁰ N. H. Phuong, V. Kreinovich, *Fuzzy logic and its applications in medicine*, Int J Med Inf. 2001;62(2-3):165-73.

2.3.2 Die Wissensbank

Das Sachwissen in einem Expertensystem wird Wissensbank genannt.

Definition:

Die Abbildung des Wissen (eines Experten oder anderer Quellen) auf die Wissensbank des Expertensystems mit dem Schwerpunkt auf Wissenerfassung und Wissenrepräsentation, heißt knowledge engineering

Der Wissensingenieur ist die Verbindung zwischen dem Experten und dem Expertensystem. Er erstellt aus der beobachtbaren Kompetenz eines Experten die Wissensbasis, in dem sich diese Kompetenz auf bestimmt Art und Weise widerspiegelt.

Bei der Erstellung einer Wissensbank kreiert der Wissensingenieur eine geeignete rechnerinterne Darstellung für das umgangssprachlich beschriebene Expertenwissen.

Einen besonderen Stellenwert kommt dem Experteninterview zu, das mit größter Sorgfalt und viel Einfühlungsvermögen durchgeführt werden soll. Die Hauptschwierigkeit ist es, die Vollständigkeit des Expertenwissens zu erfassen und umzusetzen.

Ein schwieriger Aspekt des Wissensingenieurs ist es, dem Experten zu helfen, sein Fachwissen

- zu identifizieren
- zu strukturieren
- zu formalisieren
- in geeignete Konzepte aufzugliedern

Elemente

- **Regelwissen**
Produktionsregeln sind Zusammenhänge zwischen generellen Fakten.
Durch Regeln können
 - logische Abhängigkeiten zwischen den Attributen des jeweiligen Fachgebietes
 - konsultative Vorgehensweisenbeschrieben werden.
- **Tabellenwissen**
Wissen kann in Tabellenform angeordnet und mittels Tabelleneinträgen und Data-Dictionary-Einträge (Schnittstelle zwischen Tabelle und Inferenzkomponente) beschrieben werden.
- **Prozedurales Wissen**
Darstellung von fachspezifischen Algorithmen in Form von Prozeduren
- **Taxonomiewissen**
 - begriffliches Taxonomiewissen
 - o Auflistung der zulässigen Objekte, Attribute und Werte
 - o Beschreibung ihrer gegenseitigen Zuordnung
 - Hierarchisches Taxonomiewissen
Es beschreibt die hierarchischen Beziehungen zwischen den Objekten
 - konsultatives Wissen
Konsultative Taxonomie-Elemente ermöglichen es, die Konsultation zu fokussieren.
- **Wissen über natürlichsprachliche Formulierungen**
Dieses Wissen kann definiert werden in Form von
 - Frageformulierungen
 - Ergebnisformulierungen
 - Objektinstanzierungen
- **Benutzerhilfen**
Der Wissensingenieur kann zu jedem Attribut zusätzlich Anwerthilfen definieren.

2.3.3 Wissensrepräsentation (interne Darstellung von Wissen)

2.3.3.1 Formallogik / Prädikatenlogik

Bei der *Formallogik* spricht man von einer Aussagenlogik, wobei die Sätze durch Variablen ersetzt werden. Die Sätze werden den Wahrheitswerten „wahr“ oder „unwahr“ zugeordnet. Die Aussagen werden verknüpft und den Verknüpfungen ebenfalls (nach einem festgelegten Schema) Wahrheitswerten zugeordnet.

Mögliche Verknüpfungen sind:

- Disjunktion („oder“)
- Konjunktion („sowohl als auch“)
- Subjunktion („wenn, dann“)
- Bijunktion („genau dann, wenn“)
- Negation („nicht“)

2.3.3.2 Prozedurale Repräsentation / Prozedurales Wissen

Die *prozedurale Repräsentation* in Form von Produktionsregeln und Skripten ist umgangssprachlich formuliert. Eine Produktionsregel (*Wenn-Dann-Regel*)

"wenn (Prämissen) dann (Schlußfolgerungen)"

besteht aus einem Bedingungsteil und einem Schlußfolgerungsteil. Durch Konjunktions- und Negationsverknüpfungen können alle Prämissen verbunden werden.

Der Schlußfolgerungsteil kann aus Hypothesen und Aktionen bestehen.

Die (dynamische) Wissensbank enthält alle Fakten, die zu einem bestimmten Zeitpunkt gültig sind. Jede Regel kann angewandt werden, deren Prämisse(n) als Fakten in der dynamischen Wissensbank aufscheinen. Wenn mehrere Regeln angewandt werden können, so bestimmt die Problemlösungskomponente, welche Regel angewandt wird. Die Schlußfolgerung aus dieser Regel wird dann als weiteres Faktum in die dynamische Wissensbank aufgenommen.

Regeln können zu Gruppen zusammengefaßt werden, die durch Metaregeln ausgewählt werden. Metaregeln fassen die zu einem Problemkreis gehörigen Regeln in der Form

"wenn (Prämissen) dann (Regelgruppe i)"

zusammen. Dadurch ist eine gewisse Strukturierung der Regeln und eine Fokussierung bei der Auswahl der nächsten anzuwendenden Regel gegeben.

Bestimmte Operationen (etwa: Verfahren der digitalen Bildverarbeitung) oder korrekte Abfolgen sind prozedural dargestellt. Ihre Aufgabe ist es, bestimmte Operationen (z. B.: Kantenfindung, Segmentierung) durchzuführen.

Prozeduren können auf **3 Ebenen** ausgeführt werden:

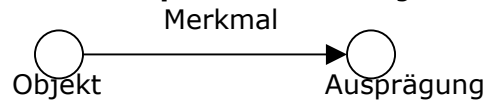
- ohne Operator Intervention
- mit Operator Intervention
- vom Operator selbst
wenn die zumindest computergestützten Verfahren nicht zielführend sind.

2.3.3.3 Deklarative Formalismen

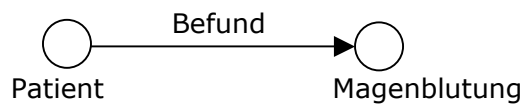
Deklarative Formalismen basieren auf Beziehungen zwischen verschiedenen Objekten. Frames und semantische Netze ermöglichen die deskriptive Repräsentation von Wissen.

2.3.3.3.1 Semantische Netze

Ein Objekt kann zusammen mit einem bestimmten Merkmal und einer spezifischen Ausprägung im Sinne der **Graphentheorie** dargestellt werden.

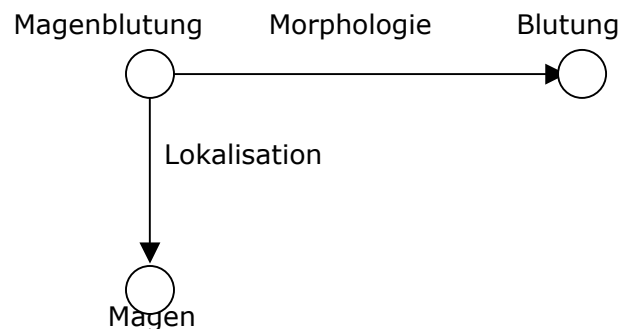


Beispiel:



Das Merkmal '*Magenblutung*' kann weiter aufgespalten werden in:

- Morphologie
- Lokalisation



Ein semantisches Netz ist definiert als ein (markierter) gerichteter Graph:

- **Knoten**, die bezeichnet / unbezeichnet³¹ sein können, entsprechen Objekten (semantische Einheiten) oder Klassen von Objekten
- **Kanten**, die bezeichnet sind, entsprechen (binäre) Relationen zwischen Objekten. In den Knoten selbst ist keine Information enthalten, alles Wissen über einen Knoten wird durch die Verbindungen repräsentiert, die von ihm ausgehen.

Die einzelnen Knoten können wieder semantische Netze sein
 ⇒ es entstehen hierarchische semantische Netze.

Unterschied zum hierarchischen Modell:

Auf einen Knoten können **mehrere** Beziehungen unterschiedlicher Art verweisen!

Semantische Netze erlauben die Speicherung von

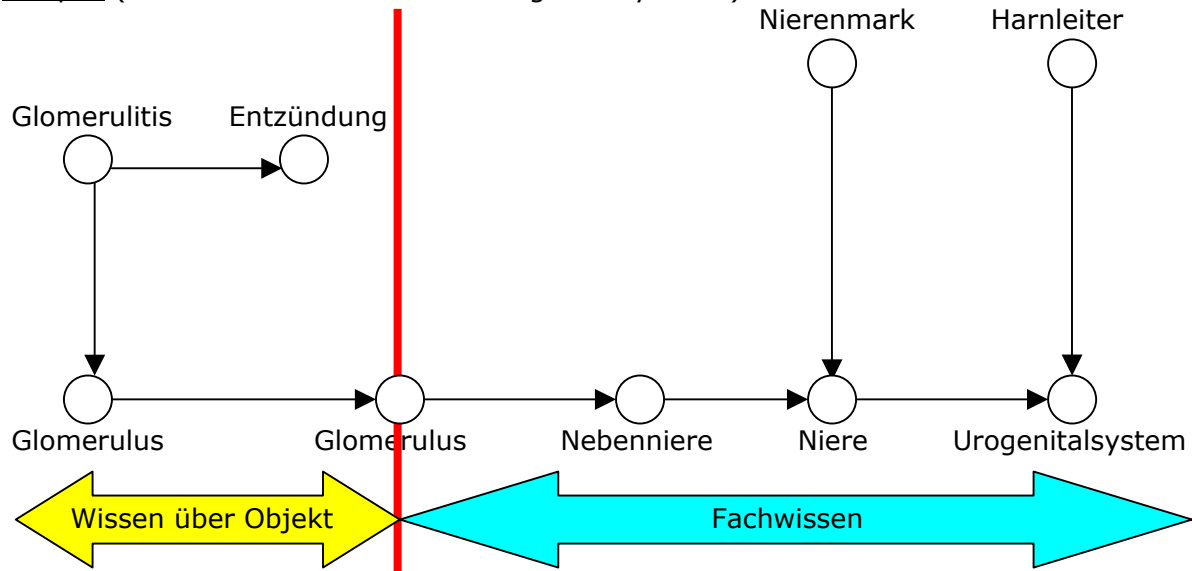
- Fakten
- assoziativen Verbindungen zwischen den Fakten (Zugriff auf benachbartes Wissen)

³¹ Bei komplexen Fakten bleiben die zugehörigen Knoten im allgemeinen unbenannt

Semantische Netze beruhen auf der Idee, daß das Gedächtnis mit assoziativ verknüpften Konzepten arbeitet. Der Psychologe Ross Quillian entwickelte sie 1968 ursprünglich als psychologisches Modell für menschliches Erinnerungsvermögen.

Semantische Netze werden auch zum Verknüpfen von Fachwissen mit (aktuellem) speziellem Wissen über ein bestimmtes Objekt verwendet.

Beispiel (anatomische Struktur des Urogenitalsystems):



Einstieg vom Wissen über das Objekt zum Fachwissen erfolgt über den Begriff 'Glomerulus'³²¹

Theoretisch sind viele Kantentypen denkbar, aber für eine effektive Verarbeitung werden sie auf einige Typen eingeschränkt:

- („ist – ein“) repräsentiert eine Relation zwischen Klasse und ein Einzelfall (Instanz).
- („hat – ein“) ist die Relation zwischen Knoten, die ihrerseits Eigenschaften anderer Knoten sind.
- Kanten, die eine definierende Funktion haben.
- Kanten, die heuristisches Wissen darstellen.

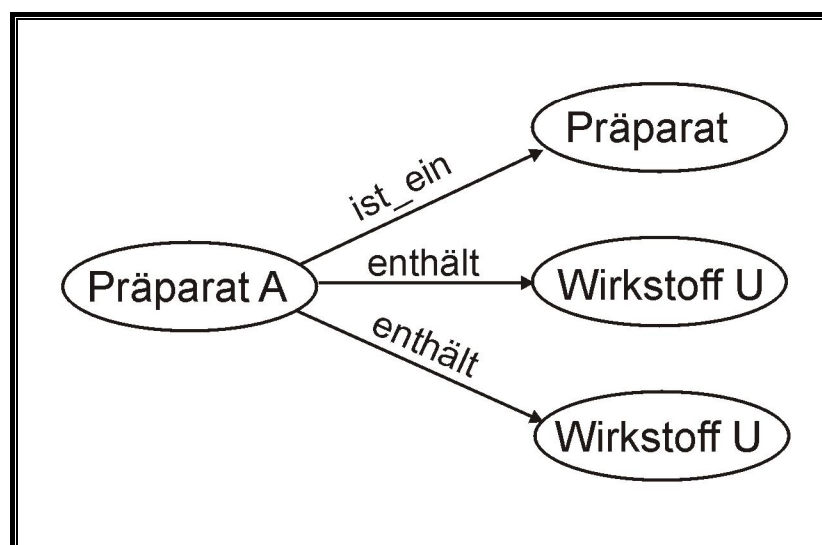


Abb. 1.2. Beispiel für ein semantischer Netz

Vorteil (der semantischen Netze):

- Flexibilität
neue Knoten und Kanten können definiert und hinzugefügt werden
- Vererbung
d.h. ein Knoten kann die Eigenschaften anderer Knoten erben, mit denen er verbunden ist

Nachteil dieser Netze:

- Reduzierung von Relationen auf einige Grundtypen. In medizinischen Expertensystemen kann so relevantes Wissen und Beziehungen verlorengehen.

2.3.3.3.2 Frames

Die Repräsentationsform *Frame* dient dazu, stereotypes Wissen zu verarbeiten und stellt eine objektorientierte Wissensrepräsentation dar. Dabei werden Objekte durch Frames beschrieben (dargestellt). Die Eigenschaften des Objekts werden in Feldern gespeichert, die Slots genannt werden. D.h. in einem Frame ist die gesamte verfügbare Information (Eigenschaften, Formen, Ausprägungen) über ein Objekt in einzelnen Slots gespeichert. Ein Frame kann aus beliebig vielen Slots bestehen, die ihrerseits wiederum aus beliebig vielen Subslots oder Facetten (Facets) bestehen können.

Jeder Frame besitzt unterschiedliche Arten von Informationen:

- die Information, wie man ihn benutzt
- die Information, was als nächstes zu erwarten ist
- die Information, was zu tun ist, wenn diese Erwartung enttäuscht wird.

Die Frames können Informationen über Relationen zwischen Objekten und zwischen Slots darstellen.

Es können auch **reservierte Relationen** vergeben werden (wie: 'Element-von', was einer Oder-Baumstruktur entspricht). Je mehr reservierte Relationen vergeben werden, umso mehr Wissen kann explizit dargestellt werden, desto inflexibler wird aber das Gesamtsystem.

Frames sind definiert durch

- einen eindeutigen Bezeichner
- ihre Position in einer Hierarchie
- ihre Relation zu anderen Frames

Da jeder Frame über seinen Namen angesprochen wird, ist das in ihm abgelegte Objekt eindeutig adressierbar.

Dasselbe gilt für Slots und Facets.

Wichtiges **Merkmal** (von Frames):

Möglichkeit, Information zu vererben,

d. h. Informationen können an alle untergeordneten Begriffe weitergegeben werden. darstellen.

Für das Füllen der Slots werden Bedingungen gestellt. Falls keine Information vorhanden ist, werden Defaultwerte angegeben. Das geschieht meist durch Prozeduren, die beim Füllen oder Ändern der Slots aktiv werden.

In einigen medizinischen Expertensystemen (z.B. *INTERNIST*³³, *PIP*³⁴) werden Krankheiten als stereotype Situationen betrachtet und durch Frames repräsentiert. Um eine Krankheit zu selektieren, die als Frame dargestellt wird, werden sogenannte Trigger-

³² Glomerulus, lat. Knäuel; Kapillarknäuel der Malpighi-Körperchen der Nierenrinde

³³ Unterstützung bei der Differentialdiagnostik in der Internen Medizin

³⁴ Present Illness Program (Diagnose nephrologischer Erkrankungen)

Manifestationen benutzt. Diese Krankheitsframes können dann für die Befragung weiterer Symptome eingesetzt werden, die zu dieser Krankheit gehören. Im Falle unzureichenden aktuellen Wissens können zusätzlich die Defaultwerte angegeben werden. Es können auch hierarchische Frame-Systeme mit Vererbungseigenschaften aufgebaut werden.

Das Konzept von Frames, auf neue Situationen durch einen Vergleich von Merkmalen unbekannter Objekte mit schon Bekanntem einzugehen, gilt als vielversprechend, solange neue Ereignisse nicht zu stark von den programmierten Stereotypen abweichen.

Die Frames werden durch den sogenannten Frame-Keeper aktiviert und gesteuert. Er

- betreut das prozedurale Wissen
- überprüft die Übereinstimmung der aktuellen Daten mit den durch den Frame gestellten Bedingungen.

Die Frame-Struktur kann als Mischung des deklarativen und prozeduralen Aspekts bezeichnet werden.

2.3.3.4 Hybride Repräsentation

Hybride Repräsentation ist eine Kombination aus prozeduralen und deskriptiven Repräsentationsformen. D. h. bei der hybriden Repräsentation werden durch den Einsatz semantischer Netze bzw. Frames die Vorteile beider Formen genützt.

Bei der Realisierung werden entweder

- semantische Netze für die Darstellung von Faktenwissen im „*Wenn-Teil*“ einer Produktionsregel verwendet oder
- Slots eines Frames mit prozeduralen Strukturen besetzt.

Menschliches Expertenwissen kann jedoch wegen seiner Komplexität und Vielfalt von Assoziationen durch hybride Repräsentation nie vollständig in einer Wissensbank dargestellt werden.

2.3.4 Die Inferenzkomponente (Problemlösungskomponente)

Ein Expertensystem muß sich intelligent verhalten. Das erfordert die Fähigkeit, aus der zur Verfügung stehenden Wissensbank Inferenzen (Schlüsse) zu ziehen, d. h. Probleme zu lösen.

3 notwendige Strukturelemente (zur Lösung eines Problems):

- **Wissensbank**, die alle Fakten enthält
- **Operatoren**, die Wissen ableiten
- **Steuerungsstrategie**, die entscheidet, welcher Operator anzuwenden ist

Die Inferenzverfahren hängen eng mit der gewählten Form der Wissensrepräsentation (Frames, semantische Netze, Prädikatenlogik) zusammen.

Ein Problem lösen bedeutet also, durch Anwendung der verfügbaren Operatoren

- einen Zustand zu erreichen, der die Zielbedingung erfüllt, oder
- Fakten in der Wissensbank zu finden, die das Ziel bestätigen.

Die *Inferenzmaschine* interpretiert anhand einer Problemstellung das in der Wissensbank enthaltene Expertenwissen und zieht daraus Schlüsse. In den Regeln der Wissensbank ist festgelegt, was in einer bestimmten Situation zu tun ist; die Inferenzmaschine bestimmt, welche Regel und in welcher Reihenfolge die Regeln verwendet werden sollen, um ein Problem zu lösen.

Für die Inferenz von Produktionsregeln werden 2 grundlegende Algorithmen verwendet, nämlich die Vorwärts- und die Rückwärtsverkettung.

In der Praxis wird eine Verbindung von daten- und zielorientierte Verfahren angestrebt, um die Zahl der zu untersuchenden Situationen möglichst gering zu halten und damit effizient und rasch zu einer Lösung zu kommen (bi-directional).

2.3.4.1 Vorwärtsverkettung (forward chaining, Daten-orientiert)

Bei der *Vorwärtsverkettung* werden alle Regeln, deren Voraussetzungen durch eine gegebene Datenbasis erfüllt sind, verarbeitet, bis keine mehr angewendet werden kann. Der Interpret zieht aus den bekannten Fakten herleitbare Schlußfolgerungen. Das System arbeitet rekursiv, bis keine Regel aus bekannten Fakten mehr ableitbar sind.

Beispiel:

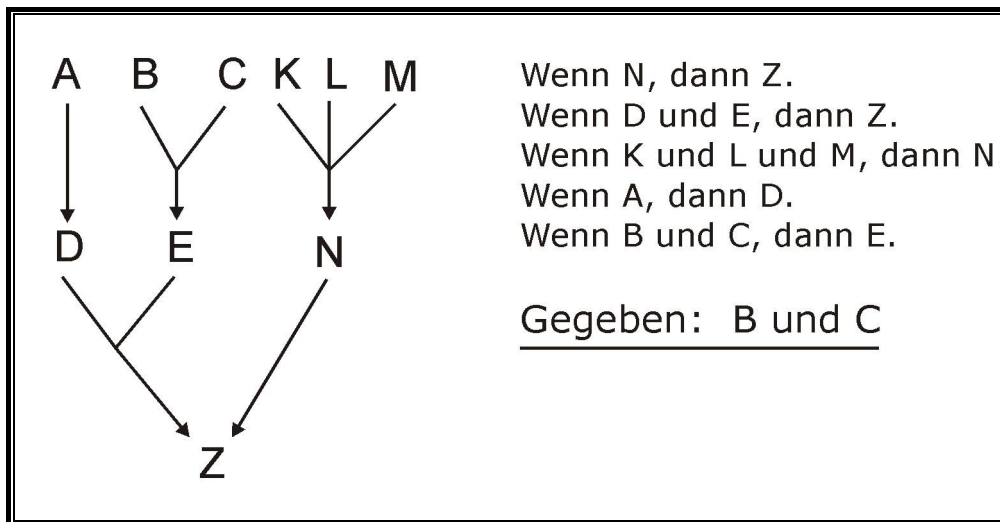


Abb. 1.3. Regelsystem mit 5 Regeln

Der Regelbaum (Abb. 1.3) besteht aus 5 Regeln und 10 Fakten, wobei die Fakten B und C gegeben sind. Das System überprüft ohne vorgegebenes Ziel alle Regeln, um herauszufinden, welche Regeln abgearbeitet werden können. Im Beispiel ist die Regel

„wenn B und C, dann schließe E“

erfüllt. Das System fügt daher E als neues Faktum den gegebenen Fakten B und C hinzu. Da keine weitere Regel ausgeführt werden kann, beendet die Vorwärtsverkettung ihre Arbeit.

In vorwärtsverketteten Systemen wird unterschieden zwischen

- Wissensbank
Sie enthält Regelprämissen, die mit dem Inhalt des Arbeitsspeichers verglichen werden. Trifft eine Regel zu, werden eine oder mehrere Schlußfolgerungen in den Arbeitsspeicher übernommen.
- Arbeitsspeicher
Er enthält Fakten, die bei der Abarbeitung einer Regel benötigt werden.

2.3.4.2 Rückwärtsverkettung (backward chaining, Ziel-orientiert)

Ausgehend von einem Ziel werden nur jene Regeln überprüft, die das Ziel im Aktionsteil enthalten, d.h. die Aufmerksamkeit des Regelinterpreters beschränkt sich auf die Regeln, die direkt mit dem Ziel in Verbindung stehen.

Sind die Prämissen nicht in der Menge der bekannten Fakten vorhanden, entscheidet der Regelinterpreter zuerst, ob es abgeleitet werden kann oder erfragt werden muß.

Ein Faktum kann dann abgeleitet werden, wenn zumindest eine Regel existiert, in der das Faktum im Aktionsteil vorkommt. Existiert keine solche Regel, bleibt dem System nur die Möglichkeit, dieses Faktum vom Benutzer zu erfragen.

Im Falle einer Ableitung werden alle Regeln, die im Aktionsteil das Ziel enthalten, abgearbeitet. Dann werden die Vorbedingungen der Regeln überprüft. Sollte dabei der Parameter einer Regel unbekannt sein, so wird zu seiner Bestimmung ein Unterziel generiert. Hierzu wird die *Rückwärtsverkettung* rekursiv eingesetzt.

Das Endresultat sieht folgendermaßen aus:

Ein Wert für das vorgegebene Ziel sowie für alle Unterziele wird bestimmt, die relevanten Regeln werden evaluiert und die nötigen Fragen werden gestellt.

Bei der Rückwärtsverkettung gibt es eine Dialogsteuerung. Die Reihenfolge der Fragen hängt dabei ab von der:

- Reihenfolge der Regeln zur Herleitung eines Parameters
- Reihenfolge der Aussagen in der Vorbedingung einer Regel

Beispiel:

2 Wege führen in Abb. 1.3. zum Zielfaktum Z:

- Z über N nach K, L und M
- Z über E und D nach B und C und A

Weg 1:

Der Regelinterpreter findet eine Regel, die Z beweisen kann:

„Wenn N, dann Z“

Da die Menge der bekannten Fakten nicht N enthält, wird N als neues Ziel gesetzt und die Rückwärtsverkettung auf das neue Ziel angewendet.

Die Regel

„wenn K und L und M, dann N“

existiert für N, aber keine der Fakten ist in der Menge der bekannten Fakten vorhanden; der Regelinterpreter muß daher diese Fakten ermitteln. Da keine Regel zur Ermittlung dieser Fakten vorhanden ist, werden diese Fakten vom Benutzer erfragt:

- Werden alle Fakten vom Benutzer als „wahr“ übermittelt, so wird N zu „wahr“ und in der Folge auch Z zu „wahr“.
- Gibt aber der Benutzer eines der Fakten K, L oder M als *nicht bekannt* ein, kann N nicht bewiesen werden und folglich auch Z nicht.

Dann muß der Regelinterpreter den zweiten Weg durchlaufen, um Z zu beweisen.

Weg 2:

Zum Beweisen von Z über den zweiten Weg müssen D und E bekannt sein.

Aus den gegebenen Fakten B und C kann E ermittelt werden. Zur Ermittlung von D muß das Faktum A vom Benutzer als „wahr“ gegeben werden. Ist das der Fall, ist D wahr, und aus den wahren Fakten D und E kann Z bewiesen und in der Menge bewiesener Fakten eingetragen werden.

2.3.5 Benutzungsoberfläche

Die Benutzungsoberfläche steuert die Interaktion zwischen Benutzer und System.

Basisanforderungen an die Benutzungsoberfläche:

- hohe Interaktionsgeschwindigkeit
- geringe Einarbeitungszeit
- leichte Bedienbarkeit
- hohe Fehlertoleranz

Die Gestaltung der Benutzungsoberflächen für *Monitoring-Systeme* (z.B. VM¹) unterscheidet sich von einem *Konsultationssystem* (z.B. MYCIN):

- Ein Konsultationssystem erfragt zuerst vom Benutzer (in natürlicher Sprache) grundlegende Daten. Sind die Daten zum Lösen des vorliegenden Problem ausreichend, gibt das System die Schlußfolgerung aus. Sonst werden weitere Daten erfragt.
- Monitoring-Systeme arbeiten meistens in der Echtzeit. Die meisten Daten werden vom System automatisch als Meßwerte erfaßt. Weitere Informationen müssen aus zeitlichen Gründe möglichst schnell eingegeben werden. Bei *Monitoring-Systemen* ist das rasche und sichere Ablesen von Ergebnissen von großer Bedeutung.

2.3.6 Erklärungskomponente

Die Erklärungskomponente steht im engen Zusammenhang mit der Inferenzkomponente; sie erzeugt verständliche und dem jeweiligen Dialogzustand angemessene Erklärung als Antwort auf eine 'Warum'-Frage.

Aufgabe:

Die *Erklärungskomponente*

- stellt dem Benutzer Informationen über die Vorgehensweise und die Wissensbank des Expertensystems zur Verfügung
- liefert Informationen über die gefundenen Lösungen
- begründet die gefolgerten Fakten

Dazu greift die Erklärungskomponente auf die von der Inferenzkomponente hinterlegten Protokolle und aufgebauten Argumentationskette zurück.

Vorteile:

- Höhere Transparenz und Steuerfähigkeit des Systems.
- Der Benutzer hat die Möglichkeit, eine Antwort auszuwerten und zu prüfen
→ Erhöhung der Akzeptanz (vor allem bei Verantwortung gegenüber Dritten)
- Für den Experten
Gezielte Rückverfolgung von Verantwortungsprozessen und Fehlerursachen

Bestandteile einer 'Warum'-Erklärung:

- Angabe des verfolgten Zieles
- Angabe der dazu ausgewählten Regeln
- Kennzeichnung der bereits erfüllten Prämissen

Eine **"Wie"-Erklärung** zeigt

- die benutzten Wissensquellen
- die relevanten Herleitungskontexte
Der Herleitungskontext besteht seinerseits aus Fakten, die durch weitere "Wie"-Erklärungen hinterfragt werden können.

¹ VM ist ein medizinisches Expertensystem zur Kontrolle bei künstlicher Beatmung

Angestrebte Eigenschaften:

- korrekt
Richtige Wiedergabe des zugrundeliegenden Inferenzprozesses
- datenunabhängig
Allgemeines Erklärungsverfahren für beliebigen Inhalt der Wissensbank
- verständlich
Klare sprachliche Struktur
Berücksichtigung des vermuteten Vorwissens beim Benutzer
- flexibel aufrufbar
großes Repertoire sprachlicher Formulierungen und kommunikativer Bedingungen zur Auslösung der Erklärungen
- dialogfähig
auf Anfrage sollten zusätzliche Erklärungen angeboten und bereits gegebene Erklärungen weiter ausgeführt werden.
- kurze Antwortzeit
damit die Komponente für die in Echtzeit ablaufenden natürlichsprachlichen Dialoge mit dem Benutzer eingesetzt werden kann
- kohärent
Bezug zum vorangegangenen Dialog

Erklärungskomponenten sind in einem Expertensystem in bezug auf Verständnis, Fehlerbeseitigung, Ausbildung und Akzeptanz sehr wichtig.

Das Verständnis der Wissensstruktur und Wissensanwendung führt zu einer besseren und effektiveren Nutzung des vorhandenen Wissens in einem Expertensystem. Um das zu erreichen, sind unter anderem

- die Dokumentation des Expertenwissens
- automatische Anzeigen von Zwischenergebnissen
- Erklärungen auf Fragen des Benutzers

sehr hilfreich, da Fehlschlüsse, die wegen einer inkonsistenten Wissensbank vorkommen, erkannt und beseitigt werden können.

Expertensysteme mit entsprechender Erklärungskomponente können auch für Lernzwecke eingesetzt werden, besonders dann, wenn die Benutzer die Entscheidungen nachvollziehen bzw. überprüfen können.

Für die Akzeptanz eines Expertensystems ist die Transparenz sehr wichtig. Dabei hat sich als sehr nützlich erwiesen:

- Kommandos auf einige wesentliche zu reduzieren
- Abfragen in natürlicher Sprache zu stellen
- die wichtigsten Informationen ohne spezielle Anforderung bereitzustellen

2.4 Klassen von den Expertensystemen

2.4.1 Diagnose-Systeme

Diagnose-Systeme versuchen, von gegebenen Beobachtungen auf die Ursachen zu schließen.

Beispiel:

- medizinische Diagnose
- Fehlerdiagnose bei technischen Systemen

2.4.2 Prognose- Systeme

Hier versucht man umgekehrt, aus gegebenen Beobachtungen Folgerungen für die Zukunft abzuleiten.

Beispiel:

Wetterprognose

2.4.3 Bildverstehende Expertensysteme

Ziel:

Ein bildverstehendes System soll die in einem Bild vorhandenen Objekte selbständig erkennen und bekannten Klassen zuordnen.

Eingabe:

diskretisiertes Bild (beliebiger Auflösung)

Ausgabe (der Objekte, die durch das System im Bild erkannt hat):

- graphischer Darstellung
- Beschreibung (z. B: Ausgabe in ein PACS)

Wissensbank:

- Statische Komponente
Wissen über den aktuellen Problemkreis **vor Beginn** der laufenden Konsultation
- Dynamische Komponente
Jenes Wissen, das **während** eines Problemlösungsvorganges angesammelt wurde:
 - bereits gefundene Objekte
 - Hypothesen über Objekte
 - bisheriger Wissenspfad

Angestrebte Eigenschaften:

- Allgemein einsetzbar
d.h. auf Bilder beliebigen Inhalts anwendbar
- Neben der allgemeinen Fähigkeit, Objekte zu erkennen, sollte ein bildverstehendes System auch spezifische Aufgaben lösen können
(z. B. nur Objekte mit bestimmten Eigenschaften zu detektieren).
- Interaktive Fehler-Korrektur
Der Anwender soll jederzeit in eine laufende Interpretation eingreifen können.
Das System soll nach jeder Intervention normal fortsetzen.
- Die Erklärungskomponente kann dem Benutzer zu jedem Zeitpunkt während und nach Abschluß einer Konsultation alle bisher erfolgten Schritte erklären. Sie stellt das in der dynamischen Wissensbank angesammelte Wissen für den menschlichen Benutzer verständlich dar.

2.4.4 Medizinische Diagnosesysteme (D-XPSe)

Je nach Art des verwendeten Wissens werden medizinische Diagnosesysteme unterteilt in:

- Statistische D-XPSe²
 - Assoziative D-XPSe
 - Modellbasierte D-XPSe
- Statistischen D-XPSe behandeln in der Hauptsache das Problem der Diagnosebewertung mit unsicheren Daten.
Vorteil:
 Hohe Objektivierbarkeit, wenn die vorhandenen Daten die statistischen Voraussetzungen erfüllen.
Nachteil:
 Die Erklärungsfähigkeit und Prognosekapazität ist gering.
- Assoziative D-XPSe basieren auf dem Erfahrungswissen von Experten.
Vorteile
 Sie können alle Aspekte des diagnostischen Problemlösens behandeln und sind breit einsetzbar, da kaum prinzipielle Voraussetzungen erfüllt sein müssen. Das Wissen assoziativer D-XPSe besteht überwiegend aus Symptom-Diagnose-Assoziationen.
- Modellbasierten D-XPSe basieren auf kausalem Ursache-Wirkungs-Wissen, so daß eine Störung zu bestimmten Symptomen führt. In modellbasierten Diagnostik-Expertensystemen wird versucht, eine Erklärung für die beobachteten Symptome durch Simulation der Auswirkungen einer vermuteten Störung zu finden. Die monokausalen Modelle können analog zu assoziativen Repräsentationen betrachtet werden. Der Unterschied wird dann deutlich, wenn die Wirkung verschiedener Ursachen behandelt wird. Dies kann in assoziativen Expertensystemen nur durch die Annahme und Verwendung zusätzlicher Regeln gelöst werden, wobei eine kombinatorische Explosion der Regelmenge vorkommen kann. Die modellbasierten Expertensysteme müßten in der Lage sein, den Kombinationseffekt zu leisten.

Vergleich der 3 Diagnostikansätze						
	Wissen	Lösbarkeit komplexer Probleme	Anwendungsspektrum	Effizienz	Qualität der Erklärung	Objektivierbarkeit
Statistische & fallvergl. D-XPSe	Falldaten	+	+(+++)	++	+	+++ (+)
Assoziative D-XPSe	Erfahrungsregeln	++	+++	+++	++	+
Modellbasierte D-XPSe	Kausale Modelle	+++	++	+	+++	++

Tab. 1.5. Vergleich zwischen statistischer, assoziativer und modellbasierter Diagnostik
 (+ = gering, ++ = mittel, +++ = hoch), [Puppe (1991) 32]

² Diagnostische Expertensysteme

2.4.5 Statistische und fallvergleichende Diagnostik-Ansätze

Das Theorem von Bayes für nur eine Diagnose D und ein Symptom S lautet:

$$p(D | S) = \frac{p(D) \times p(S | D)}{p(S)} \quad (2.5.1)$$

P(D).....a-priori-Wahrscheinlichkeit der Diagnose D

P(S|D).....bedingte Wahrscheinlichkeit des Auftretens der Symptomatik S,
falls D gegeben ist

P(D|S).....bedingte Wahrscheinlichkeit der Diagnose D,
falls die Symptomatik S gegeben ist

Die A-priori-Wahrscheinlichkeit P (D) lässt sich für alle Diagnosen D leicht aus Statistiken ermitteln. Dies gilt für P (S|D) nicht so einfach, da jeder Fall eine individuelle Symptomatik hat

Beispiel:

Bei 100 Symptomen mit je 5 Ausprägungen gibt es 10.000.000.000 verschiedene Symptomatiken.

Um dieses Problem zu lösen, wird die Symptomatik S in ihre Einzelsymptome S_1, \dots, S_n differenziert. Dann wird die Wahrscheinlichkeit P (S_i |D) für jede Einzelsymptom-Diagnose-Kombination ermittelt.

Das ergibt in obigem Beispiel eine Reduktion der erforderlichen bedingten Wahrscheinlichkeiten für eine Diagnose von 10.000.000.000 auf 500.

Die Wahrscheinlichkeit der Gesamtsymptomatik für eine Diagnose P (S|D) wird durch Multiplikation der Einzelwahrscheinlichkeiten errechnet:

$$P(S|D) = P(S_1|D) * P(S_2|D) * \dots * P(S_n|D)$$

Voraussetzung (für diese Vereinfachung):

Die Wahrscheinlichkeit des Auftretens eines Symptoms ist unabhängig von allen übrigen Symptomen.

Diese Annahme ist jedoch meist nicht erfüllt. Denn je mehr Symptome erfaßt werden (d.h. je feiner der Detaillierungsgrad ist), desto ungenauer wird das Ergebnis, das die Anwendung von Bayes' Theorem liefert, da die Abhängigkeit der Symptome untereinander zunimmt.

In der Praxis gibt es aber eine Menge von Symptomen S_1, \dots, S_m .

Für die Berechnung der wahrscheinlichsten Diagnose D_i aus den Symptomen S_1, \dots, S_m wird die Formel (2.5.1) erweitert auf:

$$p(D_i | S_1 \& \dots \& S_m) = \frac{p(D_i) \times p(S_1 | D_i) \times \dots \times p(S_m | D_i)}{\sum_{j=1}^n p(D_j) \times p(S_1 | D_j) \times \dots \times p(S_m | D_j)}$$

wenn folgende **Voraussetzungen** erfüllt sind:

- Die Symptome S_1, \dots, S_m dürfen nur von der Diagnose D abhängen
- Die Symptome S_1, \dots, S_m müssen untereinander unabhängig sein.
- Die Diagnosemenge muß vollständig sein.
- Die Diagnosen müssen sich wechselseitig ausschließen (single fault-assumption). Diese Annahme ist nur in relativ kleinen Diagnosebereichen anwendbar.
- Die a-priori-Wahrscheinlichkeiten der Diagnosen und die bedingten Symptom-Diagnose-Wahrscheinlichkeiten müssen durch weitgehend fehlerfreie und vollständige Statistiken ermittelt werden.
- konstante Wahrscheinlichkeiten
Dies ist deshalb problematisch, weil sich

- die Erscheinungsbilder der Diagnosen
- die Diagnostikmethoden verändern können

Schwierigkeiten (bei der Anwendung der Bayes'schen Formel):

- Erkennen von Mehrfachdiagnosen, da die Bayes'sche Formel für jede Diagnose nur eine Wahrscheinlichkeit liefert und nichts darüber aussagt, ob die am besten bewerteten Diagnosen sich entweder wechselseitig ausschließen oder komplementär sein können.
- Die a-priori-Wahrscheinlichkeiten der Diagnosen können sich von einer Klinik (Ort, Land, Berufsgruppe, ...) zur anderen signifikant ändern.
- Für seltene Diagnosen sind verlässliche Daten meist nicht verfügbar; die Schätzung der Wahrscheinlichkeiten durch Experten ist meist kein vollwertiger Ersatz.
- Je mehr Symptome erfaßt werden (d.h. je feiner der Detaillierungsgrad ist), desto ungenauer wird das Ergebnis, da die Abhängigkeit der Symptome untereinander zunimmt.

2.4.5.1 Fallvergleichende Diagnostik

Eine Alternative zum Theorem von Bayes bezüglich der Auswertung abgespeicherter Fälle ist die Selektion solcher Fälle aus einer Datenbank, die einem zu diagnostizierenden Fall möglichst ähnlich sind. Die Ähnlichkeit wird durch den Vergleich von ausgewählten Kriterien ermittelt, die die abgespeicherten Fälle mit dem vorliegenden Fall gemeinsam haben müssen. Der Erfolg dieses Ansatzes hängt ab von

- der Formalisierung, Standardisierung und Vollständigkeit der abgespeicherten Fälle
- dem Grad der Ähnlichkeit der Fälle

Diese Technik wurde bisher

- meist zur Bestimmung von Therapie und Prognose einer gegebenen Diagnose angewendet (z.B. ARAMIS für eine Rheumatologie-Datenbank)
- nur selten zur Diagnostik, da die Symptomatik viel schwerer einheitlich formalisierbar ist, so daß ein Merkmalsvergleich möglich ist.

Falls jedoch eine Standardisierung der Symptomatik vorhanden ist, ist die fallvergleichende Diagnostik ein wissenschaftstheoretisch interessantes Instrument, da das zentrale Problem des Wissenserwerbs auf das Abspeichern von Fällen reduziert wird.

Zur Gewichtung der Unterschiede zwischen ähnlichen Fällen ist jedoch Zusatzwissen unverzichtbar.

2.4.5.2 Dempster-Shafer-Theorie

Arthur Dempster entwickelte 1960 die Grundlage seiner Theorie, die nach einer Erweiterung durch Glenn Shafer in „*A Mathematical Theory of Evidence*“ publiziert wurde.

Unterschied (Dempster-Shafer-Theorie \Leftrightarrow Theorem von Bayes):

- Es können auch Grobdiagnosen (Diagnoseklassen) und Diagnosehierarchien repräsentiert werden. Eine Diagnoseklasse wird als Menge von Feindiagnosen dargestellt.
- Die Grundbewertungen sind bedingte Symptom-Diagnose-Wahrscheinlichkeiten $p(D|S)$ statt $p(S|D)$ wie beim Theorem von Bayes.
- Die Restwahrscheinlichkeit $(1 - p)$ einer Wahrscheinlichkeit p für eine Diagnose wird nicht gegen die Diagnose bewertet, sondern als Unsicherheitsintervall interpretiert, d.h. die Wahrscheinlichkeit der Diagnose ist nicht p , sondern liegt zwischen p und 1.
- Eine Wahrscheinlichkeit gegen eine Diagnose wird als Wahrscheinlichkeit für das Komplement der Diagnose, d.h. die Menge aller übrigen Diagnosen, repräsentiert und bewirkt eine Verkleinerung des Unsicherheitsintervalls vom oberen Ende her (d.h. ihre Maximalwahrscheinlichkeit wird entsprechend kleiner als 1).

Da für Diagnosemengen

- Wahrscheinlichkeiten miteinander verknüpft werden
- die Wahrscheinlichkeit einer bestimmten Diagnosemenge aus der Verteilung der Wahrscheinlichkeiten über alle Diagnosemengen errechnet werden

ist die Verrechnung von Wahrscheinlichkeiten bei der Dempster-Shafer-Theorie komplizierter als beim Theorem von Bayes.

Vorteile (der Dempster-Shafer-Theorie):

- die Unsicherheitsintervalle,
sie sind bei wenig Informationen sehr groß und werden bei mehr Informationen immer kleiner.

Nachteil:

- Die Grundbewertung mittels der Symptom-Diagnose anstatt der Diagnose-Symptom-Wahrscheinlichkeiten erschwert die Bestimmung der Wahrscheinlichkeiten, da Falldatenbanken meist nach Diagnosen geordnet sind.
- großer Rechenaufwand
(bei n Diagnosen sind 2^n Diagnosemengen zu berücksichtigen)
- kein effizienter Mechanismus für die Berücksichtigung von Symptomkorrelationen und Mehrfachdiagnosen (ähnlich wie das Theorem von Bayes)

Beispiel (Verrechnungsschema mit 4 hierarchisch strukturierten Diagnosen (Abb. 1.6.)):

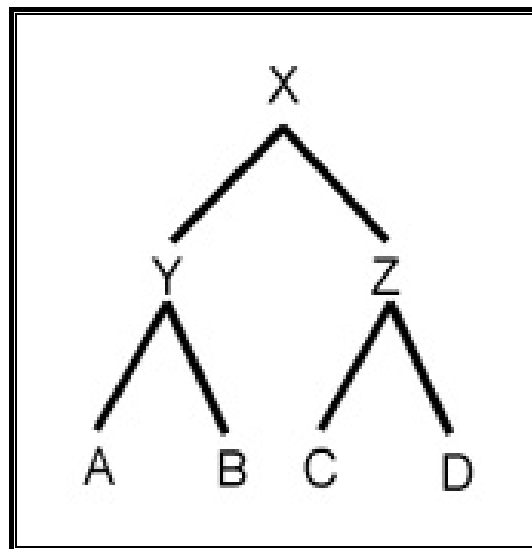


Abb. 1.6. Hierarchie: $X = \{X, Y\} = \{A, B, C, D\}$;
 $Y = \{A, B\}$; $Z = \{C, D\}$ sind Diagnoseklassen [Puppe (1991) 48]

Gegeben:

Die Wahrscheinlichkeiten p_1 und p_2 (grün unterlegt in der Matrix 1), die

- bedingte Symptom-Diagnose-Wahrscheinlichkeiten
- a-priori-Wahrscheinlichkeiten

sein können:

p_1 : 0.30 Evidenz gegen $\{A\}$

p_2 : 0.60 Evidenz für $Y = \{A, B\}$

Evidenz gegen $\{A\}$ bedeutet Evidenz für das Komplement von $\{A\}$, i.e. für die Diagnosemenge $\{B, C, D\}$.

Evidenz für Y ist Evidenz für die Menge $\{A, B\}$.

Bei der Zuweisung einer Evidenz p für eine Diagnosemenge wird die Restevidenz $(1-p)$ der Menge aller Diagnosen, also hier der Menge $\{A, B, C, D\}$, zugewiesen (gelb unterlegt in der Matrix 1).

Bei der Verknüpfung der Evidenzen (hellblau unterlegt in der Matrix 1) werden

- die Wahrscheinlichkeiten multipliziert
- die Mengen miteinander geschnitten

Daraus ergibt sich folgende Matrix:

		p2	
p1		$\{A, B\} = 0.6$	$\{A, B, C, D\} = 0.4$
	$\{B, C, D\} = 0.3$	$\{B\} = 0.18$	$\{B, C, D\} = 0.12$
	$\{A, B, C, D\} = 0.7$	$\{A, B\} = 0.42$	$\{A, B, C, D\} = 0.28$

Matrix 1

Die Gesamtwahrscheinlichkeit einer Diagnosemenge wird durch ein Unsicherheitsintervall $[a, b]$ angegeben, das die untere und die obere Wahrscheinlichkeitsgrenze abschätzt:

- a = Summe der Wahrscheinlichkeiten der Menge und aller ihrer Teilmengen

- b = Differenz zwischen 1 und der Summe der Wahrscheinlichkeiten des Komplementes der Menge und aller ihrer Teilmengen

Für die Diagnosen $\{A\}$ und $\{B\}$ ergibt sich also:

- $\{A\} = [\{A\}, 1 - (\{B, C, D\} + \{B, C\} + \{B, D\} + \{C, D\} + \{B\} + \{C\} + \{D\})] =$
 $= [\{A\}, 1 - (0.12 + 0.18)] = [0, 0.7]$
d.h. die Wahrscheinlichkeit für $\{A\}$ liegt zwischen 0 und 70%
- $\{B\} = [\{B\}, 1 - (\{A, C, D\} + \{A, C\} + \{A, D\} + \dots)] = [0.18, 1]$
d.h. die Wahrscheinlichkeit für $\{B\}$ liegt zwischen 18 und 100%

Wenn eine Evidenz

p_3 : 0.50 Evidenz für $\{D\}$

hinzukommt, ändern sich die Wahrscheinlichkeiten wie folgt:

$p_1 \text{ \& } p_2 \backslash p_3$	$\{D\} = 0.5$	$\{A, B, C, D\} = 0.5$
$\{B\} = 0.18$	$\{ \} = 0.09$	$\{B\} = 0.09$
$\{A, B\} = 0.42$	$\{ \} = 0.21$	$\{A, B\} = 0.21$
$\{B, C, D\} = 0.12$	$\{D\} = 0.06$	$\{B, C, D\} = 0.06$
$\{A, B, C, D\} = 0.28$	$\{D\} = 0.14$	$\{A, B, C, D\} = 0.14$

Die leere Teilmenge $\{ \}$ bedeutet, daß keine der Diagnosen zutrifft.

Unter der Annahme der Vollständigkeit der Diagnosemenge kann die Wahrscheinlichkeit der Diagnosemenge für die leere Teilmenge eliminiert werden, indem alle anderen Wahrscheinlichkeiten durch den Faktor

$$1 - (\text{Summe der Wahrscheinlichkeiten der leeren Teilmengen})$$

$$\text{hier: } 1 - (0.09 + 0.21) = 1 - 0.3 = 0.7$$

dividiert werden, so daß die Summe aller Wahrscheinlichkeiten ohne die leeren Menge wieder 1 ergibt.

Also gilt:

- $\{B\} = 0.09 / 0.7 \approx 13\%$
- $\{D\} = 0.2 / 0.7 \approx 28.5\%$
- $\{A, B\} = 0.21 / 0.7 \approx 30\%$
- $\{B, C, D\} = 0.06 / 0.7 \approx 8.5\%$
- $\{A, B, C, D\} = 0.14 / 0.7 \approx 20\%$

Die Wahrscheinlichkeiten für $\{A\}$ und $\{B\}$ haben sich dadurch wie folgt verändert:

$$\{A\} = [\{A\}, 1 - (\{B, C, D\} + \dots)] = [0, 0.5], \quad \text{d.h. zwischen 0 und 50\%}$$

$$\{B\} = [\{B\}, 1 - (\{A, C, D\} + \dots)] = [0.13, 0.715], \quad \text{d.h. zwischen 13 und 71.5\%}$$

2.4.6 Assoziative Diagnose-Expertensysteme

Experten können diagnostische Routineprobleme schnell lösen, weil sie

- wissen, welche Merkmale für das Problem relevant sind
- in der Lage sind, diese Merkmale mit den Diagnosen zu assoziieren.

Assoziationen lassen sich durch Produktionsregeln darstellen, deren Zuverlässigkeit kategorisch, probabilistisch oder durch Angabe der Ausnahmen ausgedrückt werden kann.

Assoziative D-XPSe besitzen durch ihre regelbasierte Wissensrepräsentation eine modulare Struktur, die den iterativen Prozeß des Wissenserwerbs wesentlich vereinfacht. Die Prüfung auf Widerspruchsfreiheit ist bei einem System mit Produktionsregeln allerdinds sehr aufwendig.

Assoziative D-XPSe kann man auch als Frame-basierte Systeme beschreiben, wobei die Diagnosen Frames sind.

Vorteile von assoziativen D-XPSe im Vergleich zum Theorem von Bayes:

- Bewertung von Symptomkombinationen in Regeln (anstelle der Beschränkung auf Bewertung von Einzelsymptomen)
- Repräsentation von Symptom- und Diagnosehierarchien (anstelle der unstrukturierten Symptom- und Diagnosegruppen)
- dynamische Erweiterungsmöglichkeiten um zusätzliche Mechanismen

Problem bei assoziativen D-XPSe:

- aufwendiger Wissenserwerb, der bisher nur über menschliche Experten möglich ist. Wenn dafür gute Voraussetzungen gegeben sind, reichen die vorhandenen Mechanismen aus, in begrenzten Anwendungsbereichen praktisch einsetzbare Systeme zu entwickeln.

2.4.6.1 MYCIN/EMYCIN

MYCIN³⁵, das bekannteste frühe D-XPS und Vorbild für zahlreiche ähnliche Systeme, wurde Anfang der 70-iger Jahre von

- Edward H. Shortliffe, Professor für Medizin und Computerwissenschaft in Stanford
- den Computerwissenschaftlern Bruce G. Buchanan und Randall Davies entwickelt. MYCIN ist ein (in LISP geschriebenes) Konsultationssystem zur Diagnose und Behandlung von Infektionskrankheiten.

MYCIN wurde später zum Shell EMYCIN weiterentwickelt.

Eigenschaften:

- getrennte Bewertung positiver und negativer Evidenz (wie die Dempster-Shafer-Theorie und INTERNIST)
- ermöglicht die Repräsentation von Symptom-Kombinationen mit Regeln

Die Auswertung der Regeln durch die Inferenzmaschine erfolgt nach dem Prinzip der Rückwärtsverkettung (nur einige wenige Regeln werden mit Vorwärtsverkettung abgearbeitet) mit Tiefensuche, d.h. nach ungefährrer Einschätzung der Krankheitsursache überprüft das System alle Hypothesen, indem es versucht, die Prämissen auf bekannte Fakten zurückzuführen. Sind nicht alle Daten vorhanden, werden fehlende Angaben vom

³⁵ Der Name „MYCIN“ leitet sich von dem Suffix zahlreicher Antibiotika ab

Benutzer erfragt. Zunächst wird versucht, eine einzelne Prämisse unter Hinzuziehung weiterer Regeln oder zusätzlicher Informationen zu bestätigen. Gelingt dies nicht oder sinkt die Wahrscheinlichkeit einer Hypothese unter „0.2“, wird die Hypothese verworfen und eine neue überprüft.

Läßt sich eine Diagnose ermitteln, erstellt MYCIN eine (algorithmisch ermittelte) Rangordnung von möglichen Medikamenten inklusive der nötigen Dosierung.

Das medizinische Wissen von MYCIN besteht zum größten Teil aus heuristischen Wenn-Dann-Regeln.

Innerhalb einer Regel dürfen die Symptome miteinander korrelieren, die Regeln als Ganzes müssen aber unabhängig voneinander sein.

Das Verrechnungsschema von MYCIN (vergleicht nicht wie INTERNIST die Diagnosen untereinander, sondern) berechnet für jede Diagnose je eine positive und negative Wahrscheinlichkeit, wobei MYCIN keine Wahrscheinlichkeiten im statistischen Sinn, sondern positive (*measure of belief*) und negative Sicherheitsmaße (*measure of disbelief*) berechnet und diese zu Sicherheitsfaktoren (certainty factor) addiert gemäß folgender Formel:

$$p = \begin{cases} p_{alt} & \text{falls } p_{neu} < 0.2 \\ p_{alt} + (1 - p_{alt}) \times p_{neu} & \text{sonst} \end{cases}$$

p_{neu} neuer Sicherheitsfaktor

p_{alt} alter Sicherheitsfaktor der Diagnose

p_{neu} Regelwahrscheinlichkeit

Die Regelwahrscheinlichkeit wird berechnet, indem die vom Experten kommende Implikationswahrscheinlichkeit mit der Wahrscheinlichkeit der vom Benutzer eingegebenen Vorbedingung multipliziert wird.

Beispiel: Regel 535 [Puppe,1991]:

Wenn	1. der Organismus gram-positiv ist,	und
	2. der Organismus in Ketten wächst,	und
	3. der Organismus die Gestalt einer Kugel hat,	
Dann	beträgt die Wahrscheinlichkeit („suggestive evidence“ ³⁶) 70%, daß der Organismus ein Streptokokkus ist.	

Angenommen, der Benutzer hat eingegeben:

Organismus gram-positiv: 1.00

Organismus wächst in Ketten: 0.60

Organismus hat Kugelform: 0.70

⇒ die Wahrscheinlichkeit der Vorbedingung ist das Minimum³⁷ der 3 Einzelwahrscheinlichkeiten:

$$\min(1.0, 0.6, 0.7) = 0.6 = 60\%$$

Daraus folgt: Die Regelwahrscheinlichkeit beträgt $0.7 * 0.6 = 0.42 = 42\%$.

Angenommen, der alte Sicherheitsfaktor von Streptokokkus, der aus anderen Regeln hergeleitet wurde, beträgt 50%, dann beträgt der neue Sicherheitsfaktor:

$$0.5 + (1 - 0.5) * 0.42 = 0.71 = 71\%.$$

³⁶ Der Ausdruck "suggestive evidence" wird als Sicherheitsgrad der Aussage verstanden. Als Sicherheitsgrade sind Werte aus dem Intervall von -1 (sicher ausgeschlossen) bis +1 (sicher bewiesen) erlaubt.

³⁷ Da eine „und“-Verknüpfungen; bei „oder“-Verknüpfungen wird das Maximum genommen

Wenn von negativen Regeln ein Sicherheitsfaktor von 0.30 errechnet wurde, dann ist der Gesamtsicherheitsfaktor von Streptokokkus:

$$0.71 - 0.3 = 0.41 = 41\%.$$

Eine Besonderheit³⁸ von MYCIN ist, daß der Benutzer die Unsicherheiten bei der Symptomerhebung sehr detailliert angeben kann. Dadurch entstehen jedoch neue Probleme, die nicht befriedigend gelöst sind:

- Unterschiedliche Situationen werden gleich behandelt:
Beispiel:
Wenn in einer Regel alle Aussagen nur zu 0.50 sicher sind, wird dies genauso behandelt, wie wenn nur eine Aussage zu 0.50 und die anderen Aussagen zu 1.00 sicher wären.
- Für den Experten wird die Abschätzung der Implikationswahrscheinlichkeit dadurch erschwert, daß er die Benutzer-Symptomwahrscheinlichkeiten mitberücksichtigen muß.

Bewertung:

MYCIN erstellt nicht nur Diagnose und Therapievorschl ag, sondern kann auch aufgrund einer umfangreichen Erkl arungskomponente

- getroffene Entscheidungen begr unden
- allgemeine Fragen zu Infektionskrankheiten beantworten

MYCIN's Diagnose- und Therapievorschl age wurden in einer Blindstudie mit jenen von Fachexperten, weiteren  rzten sowie Medizinstudenten verglichen, wobei die Akzeptanz der von MYCIN erstellten Vorschl age im Durchschnitt sowohl die der Fachexperten als auch die der anderen  rzte und der Medizinstudenten  bertraf.

MYCIN wurde jedoch nie in den klinischen Routinebetrieb  berf hrt, da das in der Wissensbank gespeicherte medizinische Wissen (mehr als 400 Regeln) noch nicht f r einen praktischen Einsatz ausreichte.

³⁸ im Vergleich vielen anderen Expertensystemen

2.4.6.2 EXPERT

EXPERT ist ein aus *CASNET*³⁹ weiterentwickeltes System, das von C. Kulikowski und dem Computerwissenschaftler S. Weiss entwickelt wurde.

Wesentliche Änderungen im Vergleich CASNET sind

- zusätzliche Produktionsregeln
- Ergänzung der kausalen Struktur durch eine *taxonomische* Hierarchie
- Vereinfachung der Wissenskategorien
Statt 4 Ebenen existieren nur noch 2 Ebenen:
 - Observationen (Befunde, findings)
 - Hypothesen
(sie umfassen jetzt auch pathophysiologische Zustände, Krankheitskategorien und Therapievorschlge)

EXPERT gehrt zu den Produktionensystemen.

Die **Wissenskomponenten** sind in:

- Befunde (finding):
(Symptome, persnliche Daten, Labordaten,.....)
- Hypothesen:
Jene Diagnosen, die hergeleitet werden knnen.
- Regeln:
Mittels Regeln werden die Hypothesen aus den Befunden hergeleitet.

Bei den **Produktionsregeln** wird unterschieden zwischen:

- Befund-Befund-Regeln (finding-finding):
Die Prmissen und Konklusionen beziehen sich nur auf Observationen, d.h. aus vorhandenen Befunden wird auf andere Befunde geschlossen.
Die Aussagen knnen nur wahr, falsch oder unbekannt sein.
- Befund-Hypothese-Regeln (finding-hypothesis):
Als Prmissen dienen die logische Kombinationen von Befunden, die eine Hypothese als Konklusion bestrken oder abschwchen.
Beispiel (Befund-Hypothese-Regel):
Wenn Observation 1 *wahr* und Observation 2 *falsch* ist, dann ist die Hypothese 1 mit Wahrscheinlichkeit 0.8 besttigt.
- Hypothese-Hypothese-Regeln (hypothesis-hypothesis):
Verknpfung von gewichteten Hypothesen.
Bei der Hypothese-Hypothese-Regeln drfen Hypothesen in der Prmissen vorkommen.
Beispiel:
Wenn der Wahrheitswert von Hypothese 1 zwischen 0.4 und 1 und der Wahrheitswert von Hypothese 2 zwischen 0.6 und 1 liegt, dann ist Hypothese 3 mit Wahrscheinlichkeit 1 besttigt

In EXPERT sind die wesentlichen Hypothesen in einer gewichteten Taxonomie geordnet, die durch kausale Verbindungen ergnzt werden kann.

Beim Inferenzproze werden die endgltigen Gewichtungen der Hypothesen sowohl aus den regelbasierten als auch aus den taxonomisch-kausalen Beziehungen ermittelt. Dazu werden nach Eingabe von grundlegenden Observationen zunchst die Befund-Befund-Regeln ausgewertet, um weitere Fakten zu erschlieen. Anschließend werden die Befund-Hypothese-Regeln unterteilt in

- besttigte (alle Prmissen erfllt)
- unbesttigte (mindestens eine Prmissen nicht erfllt)

³⁹ **CASNET** (Causal-Associational Network (Diagnose und Therapie des Glaukoms))

- unbestimmte (nicht alle Prämissen bekannt)

Nur die unbestimmten Befund-Hypothese-Regeln können ihren Status nach Eingabe weiterer Observationen noch ändern.

Die durch die Regelauswertung gewonnen Hypothesen besitzen Wahrheitswerte, die mit Hilfe der Hypothese-Hypothese-Regeln verändert werden können. Diese Regeln werden (anders als beim häufig benutzten Prinzip der Rückwärtsverkettung) in der Reihenfolge ihres Auftretens beim Entwurf des Expertensystems ausgewertet. Zusätzlich werden die Wahrheitswerte durch die taxonomisch-kausalen Beziehungen beeinflusst, so daß am Ende des Inferenzprozesses eine Hypothesenrangfolge angegeben werden kann.

Um die Ähnlichkeit mit einem ärztlichen Gutachten zu vergrößern, werden die Beziehungen zwischen den einzelnen Hypothesen auch erläutert.

Entwicklungs-Shell

Die Entwicklung eines neuen Expertensystems mit EXPERT ist vergleichbar mit der Erstellung eines Computerprogramms. Unter Verwendung eines Text-Editors werden

- die Hypothesentaxonomie
 - die Liste der Observationen zusammen mit der Art der Erfragung (Ja/Nein, Multiple-choice oder numerische Werte)
 - die Produktionsregeln eines bestimmten Krankheitsgebietes
- einggegeben und anschließend durch einen Compiler in eine maschineninterne Repräsentation übersetzt.

Zusätzlich erlaubt das System die Speicherung von Fallstudien, verbunden mit den Möglichkeiten einer empirischen Datenbankanalyse.

Beispiel:

Mit EXPERT wurde ein Programm zur Diagnose von rheumatischen Erkrankungen erstellt, das

- 8 Hypothesen
- über 150 Observationen
- ungefähr 100 Produktionsregeln

umfaßt und bei Studien mit einer Erfolgsquote von 94% sehr zuverlässig arbeitet.

Außerdem wurden medizinische Expertensysteme zur Behandlung von Augenkrankheiten sowie von Schilddrüsenstörungen entwickelt.

Bewertung:

Die freie Entscheidungsmöglichkeit des Entwicklers, den Schwerpunkt entsprechend den Erfordernissen des jeweiligen Krankheitsgebietes mehr auf taxonomische Strukturierung oder auf regelbasierte Relationen zu setzen, wird allgemein positiv beurteilt.

Diese Flexibilität kann allerdings auch zu schwer durchschaubaren Wissensbasen führen.

2.4.6.3 INTERNIST

INTERNIST ist eines der größten Diagnosesysteme und wurde von J. D. Myers und H. E. Pople entwickelt.

INTERNIST gehört zu den Produktionssystemen und enthält ca. 75% des gesamten Wissens der inneren Medizin. In INTERNIST sind ungefähr 500 Krankheiten und 3.500 Symptome gespeichert.

Als Wissensbank enthält INTERNIST **Krankheitsprofile** (disease profiles), in welchen zu jeder Krankheit die wichtigen Daten und alle Symptome sowie Labordaten gespeichert sind.

INTERNIST benutzt (wie EXPERT) zur Gliederung der Krankheiten Hierarchien. Da dabei allgemeinere Begriffe als Überbegriffe von spezielleren eingesetzt werden, können Problemfelder bei der Diagnose durch den nächst höheren, allgemeineren Krankheitsbegriff abgegrenzt werden.

Sämtliche im System gespeicherten Daten werden als Manifestationen bezeichnet. Die Manifestationen sind aufgelistet und ihre Relation zur Krankheit gespeichert.

Es gibt 3 Kategorien von Relationen zu Krankheiten :

- evoking strength (Werte von 0-5):
wie stark kann man annehmen, daß eine Manifestation die spezielle Krankheit als Ursache hat.
- frequency (Werte von 1-5):
wie häufig tritt diese Manifestation bei dieser Krankheit auf
- import (Werte von 1-5):
wie wichtig ist es, daß diese Manifestation durch eine Diagnose erklärt wird.
⇒ Dieser Wert ist nicht krankheitsspezifisch.

Beispiel:

Manifestation	Diagnose 1 (z.B. Pfortaderverschluß)		Diagnose 2	
	import	evoking strength	frequency	evoking strength
Manifestation 1 (z.B. Anämie)	3	1	3	

Verbale Beschreibung der Kategorien			
	evoking strength	frequency	import
0	extrem selten	---	---
1	selten	selten	Die Manifestation ist unwichtig
2	oft	oft	Die Manifestation könnte wichtig sein, kann aber meist ignoriert werden
3	meistens	meistens	Mittlere Bedeutung
4	in der Mehrheit der Fälle	in der Mehrheit der Fälle	Höhere Bedeutung
5	praktisch immer (pathognomonisches Symptom)	praktisch immer	Die Manifestation muß unbedingt durch eine Diagnose erklärt werden

Tab. 1.7. [Puppe (1991) 51]

Diese Eintragung bedeutet:
die Manifestation "Anämie"

- ist relativ wichtig (import=3)
- deutet selten (evoking strength=1) auf Diagnose 1, da sie viele andere Ursachen haben kann
- kommt meistens (frequency=3) bei Diagnose 1 vor

Zur Errechnung werden den Kategorien Punkte zugeordnet.

Berechnung:

Aufgetretene Symptome mit der Evoking Strength bestärken eine Diagnose, nicht-aufgetretene Symptome verringern das Verdachtsmoment, auch wenn man ihr Auftreten vermuten sollte.

Die Erklärbarkeit der Symptome durch die Diagnose wird mit Hilfe des Import-Value festgelegt.

Wenn ein Symptom auftritt, das nicht auf der Symptomliste der Diagnose steht und nicht erklärt werden kann, werden je nach Import-Value des Symptoms Punkte abgezogen (Tab. 1.8).

Kategorie	Evoking Strength	Frequency	Import-Value
0	1 Punkt	0 Punkte	0 Punkte
1	4 Punkte	1 Punkt	2 Punkte
2	10 Punkte	4 Punkte	6 Punkte
3	20 Punkte	7 Punkte	10 Punkte
4	40 Punkte	15 Punkte	20 Punkte
5	80 Punkte	30 Punkte	40 Punkte

Tab. 1.8: Bedeutung der Kategorien [Puppe (1991) 51]

Beispiel 1:

Manifestation	Import	D1		D2		D3	
		evoking strength	frequency	evoking strength	frequency	evoking strength	frequency
S1	3	3	3	2	2	1	1
S2	2	3	4	1	1	1	2
S3	4	1	2	3	3	2	1
S4	3	0	2	2	3	4	3
S5	1	1	1	2	1	1	2

Gegeben:

S1	S2	S3	S4	S5
+	+		-	+

Diagnose	S1	S2	S3	S4	S5
	+	+		-	+
D1	3 (20)	3 (20)	1 (4)	0 (1)	1 (4)
D2	2 (10)	1 (4)	3 (20)	2 (10)	2 (10)
D3	1 (4)	1 (4)	2 (10)	4 (40)	1 (4)

$$\Rightarrow D1: 20+20-1+4 = 43$$

$$\Rightarrow D2: 10+4-10+10 = 14$$

$$\Rightarrow D3: 4+4-40+4 = -28$$

Beispiel 2:

Manifestation	Import	D1		D2		D3	
		evoking strength	frequency	evoking strength	frequency	evoking strength	frequency
S1	3			2	2	1	1
S2	2	3	4	1	1	1	2
S3	4	1	2	3	3		
S4	3	0	2			4	3
S5	1	1	1	2	1	1	2

Gegeben:

S1	S2	S3	S4	S5
+	+	+	-	+

Diagnose	S1	S2	S3	S4	S5
	+	+		-	+
D1		3 (20)	1 (4)	0 (1)	1 (4)
D2	2 (10)	1 (4)	3 (20)		2 (10)
D3	1 (4)	1 (4)		4 (40)	1 (4)

$$\begin{aligned}
 \Rightarrow D1: & \quad .+20+4-1+4 & = 27 & \quad 27-10 & = 17 \\
 \Rightarrow D2: & \quad 10+4+20-.+10 & = 44 & \quad 44-0 & = 44 \\
 \Rightarrow D3: & \quad 4+4+.-40+4 & = -28 & \quad -28-20 & = -48
 \end{aligned}$$

Die Gesamtpunktebewertung hat in INTERNIST keine absolute Bedeutung, sondern wird mit der Bewertung anderer Diagnosen verglichen. In INTERNIST stehen nur die Diagnosen in Konkurrenz zueinander, die die gleichen Symptome erklären können. Innerhalb der konkurrierenden Diagnosegruppe sucht INTERNIST die beste Diagnose aus, die einen Vorsprung von mindestens 80 Punkten vor ihren Konkurrenten haben muß. Die Symptome der ausgewählten Diagnose werden als erklärt markiert. Mit den restlichen Symptomen wird nach möglichen weiteren Diagnosen gesucht.

Bewertung:

Das Modell von INTERNIST kann als sinnvoller Mittelweg benutzt werden, um einen einfachen und genauen Umgang mit unsicherem Wissen zu gewährleisten, zumal wenn eine breite und gute Wissensbank zur Verfügung steht.

2.4.6.4 HEPAXPERT-I (automatische Interpretation der Hepatitis-A- und -B-Serologie)

Ziel:

Verbesserung der Hepatitis-Serologie-Befundung durch eine klinisch relevante, textliche Interpretation der im Laboratorium ermittelten Befunde (Ergebnisse). Diese Interpretation soll dem zumeist nicht hepatologisch geschulten behandelnden Arzt beim Verständnis der Befundergebnisse helfen. Das befunderstellende Laboratorium soll die Interpretation ohne zusätzliche anamnestische, biochemische oder sonstige klinische Daten bereitstellen können.

Daten

Die qualitativen Ergebnisse der routinemäßig durch Enzymimmunoassay bestimmten Hepatitis-A- und -B-Antigene und -Antikörper, i.e.:

für die Hepatitis A:

- Gesamtantikörper (Anti-HAV)
- Antikörper der IgM-Fraktion (IgM anti-HAV)
- der Nachweis der Hepatitis-A-Viren im Stuhl (HAV)

für die Hepatitis B:

- Surface Antigen (HBsAg)
- Anti-Surface Antikörper (Anti-HBs)
- Anti-Core Gesamtantikörper (Anti-HBc)
- Anti-Core Antikörper der IgM-Fraktion (IgM anti-HBc)
- Envelope Antigen (HBeAg)
- Anti-Envelope Antikörper (Anti-HBe)

Jeder dieser Tests kann einen der 4 qualitativen Werten annehmen:

"positiv", "negativ", "grenzwertig", "nicht untersucht"

Kombiniert man jedes mögliche Testergebnis miteinander, so ergeben sich

- für die Hepatitis-A-Serologie 64 ($= 4^3$)
- für die Hepatitis-B-Serologie 4.096 ($= 4^6$)

verschiedene Befundmuster.

Ziel:

Jedes mögliche Befundmuster, auch wenn es äußerst selten auftritt, widersprüchlich oder unscharf ist, wird vollständig analysiert und hinsichtlich Virusexposition, Immunität, Krankheitsstadium, Prognose und Infektiosität interpretiert.

Analyse:

Das aktuelle Befundmuster wird mit jenen Serologiebefundmustern verglichen, die sich im Verlauf einer Hepatitis-A- und -B-Virusinfektion ergeben können. Insgesamt wurden an Hand der Literatur

- für die Hepatitis-A
der natürliche Infektionsverlauf sowie eventuelle passive Immunisierung,
- für die Hepatitis-B
4 akute und 4 chronische Verlaufsformen sowie Kombinationen dieser Verläufe und zusätzlich die Möglichkeit der aktiven oder passiven Immunisierung angenommen.

Bei 64 bzw. 4.096 möglichen Befundmuster ist zu erwarten, daß eine Reihe von verschiedenen Befundmustern unter einer Interpretation vereinigt werden können.

Der Wissenserwerbsprozeß setzte sich deshalb aus 2 Teilaufgaben zusammen:

- Auswahl jener serologischen Befundmuster, die die gleiche Interpretation erhalten;
d. h. Bestimmen der Befundklassen mit gleichem Interpretationstext.
Problem:

Bestimmen der Anzahl der benötigten Befundklassen (= der verschiedenen Interpretationstexte)

- Formulierung der jeweiligen Interpretationstexte

Ergebnis:

- 13 Befundklassen für die Hepatitis-A-Serologie
- 106 Befundklassen für die Hepatitis-B-Serologie

Jede Befundklasse wird gemeinsam mit dem dazugehörigen Interpretationstext als logische Wenn-Dann-Regel repräsentiert.

Beispiel (Regel 29 der Hepatitis-B-Serologie):

WENN

HbsAg	Anti-HBs	Anti-HBc	IgM Anti-HBc	HbeAg	Anti-HBe
+/- ?	-	-	- +/- ?	- +/- ?	- +/- ?

+ = "positiv", - = "negativ", +/- = "grenzwertig" und ? = "nicht untersucht".

DANN

Eine bestehende oder frühere Hepatitis-B Virusinfektion kann (mit Ausnahme des Inkubationsstadiums) ausgeschlossen werden.
Es besteht keine Immunität gegen das Hepatitis-B-Virus

Mehrere Testergebnisse in einem Tabellenelement sind durch "oder" verknüpft;
die Ergebnisse verschiedener Tests in einer Tabellenzeile sind durch "und" verknüpft;
und mehrere Tabellenzeilen sind wiederum durch "oder" verknüpft.

Ablauf:

Die erhobenen Laborergebnisse der Hepatitis-A- und -B-Serologie werden über eine Bildschirmmaske in das Interpretationsprogramm eingegeben. Auf Tastendruck wird der Befundbericht gedruckt.

Beispiel:

Eingabe 1:

Anti-HAV	IgM Anti-HAV	HAV (Stuhl)
+/-		

⇒ Befundbericht 1:

Der Befund ist zuwenig aussagekräftig, um hinsichtlich Hepatitis-A eine sichere Interpretation zuzulassen.

Zur Kontrolle des nicht eindeutig negativen oder positiven Befundes wird eine neuerliche Materialeinsendung empfohlen.

Eingabe 2:

HbsAg	Anti-HBs	Anti-HBc	IgM Anti-HBc	HbeAg	Anti-HBe
+	+	+	+	-	-

⇒ Befundbericht 2:

Diese Befundkonstellation (positives HBs-Antigen, positive Anti-HBs und IgM anti-HBc Antikörper bei negativem HBe-Antigen und negativen Anti-HBe Antikörpern) findet sich im Verlauf der akuten Hepatitis-B und charakterisiert die Serokonversion sowohl von HBs-Antigen zu Anti-HBs als auch von HBe-Antigen zu Anti-HBe Antikörpern. Diese Phase kann als prognostisch günstiges Zeichen hinsichtlich eines nicht-chronischen Verlaufes der Erkrankung angesehen werden, wenn sie innerhalb von 10 Wochen nach

Krankheitsbeginn auftritt. Blut und Sekrete (Speichel, Sperma, Muttermilch) des Patienten sind infektiös.

Hepaxpert-III - Netscape

File Edit View Go Communicator Help

Back Forward Reload Home Search Netscape Print Security Shop Stop

Bookmarks Location: <http://medexpert.inc.akh-wien.ac.at/hepax/default-e.htm> What's Related

Hepaxpert-III

Knowledge-Based Interpretation
of
Hepatitis A und B Serology

[Interpretation](#)

[Introduction](#)

[Med. Fundamentals](#)

[User Guide](#)

[Links](#)

[Disclaimer](#)

→ [deutsch/German](#)

e-mail: [HepMaster](#)

→ [MedExpert/WWW](#)

Input of Test Results

Hepatitis A Serology

anti-HAV ☒ positive ☐ negative ☐ borderline ☐ not tested

IgM anti-HAV ☐ positive ☒ negative ☐ borderline ☐ not tested

HAV (Stool) ☐ positive ☐ negative ☐ borderline ☒ not tested

Hepatitis B Serology

HBsAg ☐ positive ☒ negative ☐ borderline ☐ not tested

anti-HBs ☒ positive ☐ negative ☐ borderline ☐ not tested

anti-HBc ☐ positive ☒ negative ☐ borderline ☐ not tested

IgM anti-HBc ☐ positive ☐ negative ☐ borderline ☒ not tested

HBeAg ☐ positive ☒ negative ☐ borderline ☐ not tested

anti-HBe ☐ positive ☒ negative ☐ borderline ☐ not tested

anti-HBs titre units/litre (integer values only!)

By sending this form for interpretation of the entered test results all statements of the [disclaimer](#) are fully accepted.

© 1998, 1999, 2000; Section on [Medical Expert and Knowledge-Based Systems](#), Department of Medical Computer Sciences, University of Vienna
WWW-Implementation: Clemens Chizzali-Bonfadin, Knowledge-Base: [Klaus-Peter Adlassnig](#), Wolfgang Horak

Document: Done

Hepaxpert-III
Knowledge-Based Interpretation of Hepatitis A und B Serology

[Interpretation](#)
[Introduction](#)
[Med. Fundamentals](#)
[User Guide](#)
[Links](#)
[Disclaimer](#)

deutsch/German

e-mail: [HepMaster](#)

[MedExpert/WWW](#)

Interpretation

Hepatitis A Serology

anti-HAV	IgM anti-HAV	HAV (Stool)
positive	negative	not tested

Positive results for total anti-HAV antibodies in combination with negative results for IgM anti-HAV antibodies indicate immunity to the hepatitis virus A and exclude the possibility of a recent hepatitis A. This immunity may either have been acquired naturally through an earlier infection or it may have been induced by active vaccination or passively acquired immunization.

Hepatitis B Serology

HBsAg	anti-HBs	anti-HBc	IgM anti-HBc	HBeAg	anti-HBe
negative	positive	negative	not tested	negative	negative

anti-HBs titre: 30 units/litre

The patient is immune to the hepatitis virus B owing to vaccination or passive immunization. (In rare cases anti-HBs antibodies may occur as the only antibodies after a hepatitis B virus infection as a primary immune response).

Vaccination Recommendation: If an indication for a hepatitis B vaccination exists, the primary course of immunization has been completed, and the last partial vaccination was given at least 1 month previously, an immediate hepatitis B booster shot is recommended to maintain the immunity. If the person is suspected of being a 'low responder' a titre check 2 months after the booster is advisable.

Automatic interpretation by Hepaxpert-III/WWW at 3:20:32PM on May 14, 2001.

© 1998, 1999, 2000; Section on [Medical Expert and Knowledge-Based Systems](#), Department of Medical Computer Sciences, University of Vienna
WWW-Implementation: Clemens Chizzali-Bonfadin, Knowledge-Base: [Klaus-Peter Adlassnig](#), Wolfgang Horak

<http://medexpert.imc.akh-wien.ac.at/hepax/>

2.4.6.5 CADIAG-2 (differentialdiagnostische Entscheidungsunterstützung in der Internen Medizin)

Ziel:

- auf Grund von vorhandenen Symptomen des Patienten sind
 - sicher bewiesene Diagnosen
 - sicher ausgeschlossene Diagnosen
 - mögliche Diagnosenanzuzeigen und zu begründen;
- seltene, aber mögliche Krankheiten sind in Evidenz zu halten und dem Benutzer zu präsentieren;
- auf die durch die ermittelten Diagnosen nicht geklärten Symptome ist gesondert hinzuweisen und eventuelle Hinweise dieser Symptome zu anderen Krankheitsgruppen sind anzugeben;
- das weitere diagnostische Vorgehen ist - nach rationalen Gesichtspunkten aufgeschlüsselt - vorzuschlagen;

Methodik

Das medizinische Wissen ist in Form von nosologischen Krankheitsprofilen in einer medizinischen Wissensbank gespeichert. Das Krankheitsprofil einer Krankheit enthält die gesamte bekannte Symptomatik dieser Erkrankung einschließlich der Wertigkeit der einzelnen Symptome für diese Erkrankung.

Falls möglich, werden auch ganze Kombinationen von Symptomen in die Wissensbank aufgenommen, sofern sie

- Krankheitsdefinitionen darstellen
- pathognomonisch
- von besonderer Wertigkeit

sind. Die Krankheitsprofile wurden auf der Grundlage von Lehrbuch- und Originalliteratur sowie unter Hinzunahme eigener klinischer Erfahrung erstellt und teilweise mit Patientendaten aus Falldatenbanken verglichen.

Damit werden (einfache oder komplexe) Zusammenhänge zwischen Symptomen und Krankheiten festgehalten, die definitorischer, kausaler oder statistischer Natur sein können. Es werden 2 klinische Aspekte des Zusammenhanges zwischen Symptomen und Krankheiten betrachtet:

- die Notwendigkeit des Auftretens des Symptoms bei der Erkrankung
- das Hinreichendsein des Symptoms für die Bestimmung der Erkrankung

Sowohl die Notwendigkeit als auch das Hinreichendsein kann definitiver Natur oder mit einer Unsicherheit behaftet sein, d.h. es können

- definitive Schlussfolgerungen (z.B. "ist pathognomonisch", "schließt aus")
- unsichere Aussagen (z.B. "tritt häufig auf", "ist hinweisend")

formuliert und in die Wissensbank eingebracht werden. In CADIAG-2 werden diese beiden Aspekte als Auftreten des Symptoms bei der Erkrankung und Beweiskraft des Symptoms für die Erkrankung bezeichnet. Dem Auftreten und der Beweiskraft werden reelle Zahlen zwischen

0	für "tritt nie auf"	bzw. "ist ausschließend"	und
1	für "tritt immer auf"	bzw. "ist beweisend"	

zugewiesen, wobei die numerischen Zwischenwerte ggf. auch durch linguistische Begriffe wie "häufig", "selten", "stark", "schwach" repräsentiert sein können.

Zusätzlich zu den gespeicherten Krankheitsprofilen und den klinisch-relevanten Symptomkombinationen werden auch definitorische, kausale und hierarchische

Beziehungen zwischen Symptomen sowie taxonomische Beziehungen zwischen Diagnosen in die medizinische Wissensbank eingebracht.

Diagnoseermittlung

Die Symptome eines Patienten werden mit den Eintragungen in der Wissensbank verknüpft und Diagnoseergebnisse abgeleitet. Die Ergebnisse gliedern sich in

- bewiesene Diagnosen
- ausgeschlossene Diagnosen
- Diagnosehypothesen, die in den nachfolgenden Schritten noch bestätigt oder verworfen werden müssen

Alle ermittelten Diagnosen werden

- detailliert begründet
- im Zweifelsfall weitere Untersuchungen zur Abklärung vorgeschlagen

Da der gesamte Diagnoselauf beliebig oft wiederholt werden kann und die jeweiligen Diagnoseergebnisse angezeigt werden, ist es möglich,

- mit wenigen, einfach zu erhebenden Symptomen die Konsultation zu beginnen, sich das entsprechende Diagnosespektrum anbieten zu lassen - wobei auch besonderer Wert auf seltene Erkrankungen gelegt wird - , dann
- iterativ mit weiteren Untersuchungen bis zur detaillierten differentialdiagnostischen Abklärung des gesamten Falles fortzufahren.

Besonderheit :

Die Möglichkeit, nicht nur "Ja/Nein"-Befunde, sondern auch grenzwertige Befunde eingeben zu können. Das betrifft

- grenzwertige Ergebnisse von Laboruntersuchungen
- anamnestisch berichtete Symptome wie "Schmerzen"
- Ergebnisse der physikalischen Untersuchung wie "Ikterus", "Leber am Rippenbogen palpabel" usw.

Die Wertigkeit dieses Symptoms in bezug auf eine Diagnose wird dann entsprechend verringert und bei der Ermittlung der Diagnoseergebnisse berücksichtigt.

Ergebnisse

In die Wissensbank des Systems wurden Krankheitsprofile und Symptomkombinationen von insgesamt 262 Krankheiten eingegeben:

- 185 rheumatologische Erkrankungen
(69 Gelenkserkrankungen, 12 Erkrankungen der Wirbelsäule, 38 Krankheiten des Binde- und Stützgewebes, 45 Knochenerkrankungen und 21 regionale Schmerzsyndrome)
- 77 Krankheiten aus der Gastroenterologie
(24 Gallenblasen- und Gallenwegserkrankungen, 10 Pankreaserkrankungen, 37 Dickdarmerkrankungen) und der Hepatologie (6 Hepatitiden)

Darunter befinden sich auch die aus den ARA (American Rheumatism Association) - Kriterien abgeleiteten Symptomkombinationen für das rheumatische Fieber, die chronische Polyarthrit, den Lupus Erythematodes Disseminatus, usw.

Die Patientendatenerfassung erfolgt mit Hilfe des KIS. Dadurch ist es möglich, alle dokumentierten Daten (Anamnese, physikalischer Status, Laborbefunde, klinische und sonstige Untersuchungsergebnisse) in der Konsultationssitzung zu nutzen.

Um eine "intelligente" Übertragung der dokumentierten Daten aus dem KIS in das CADIAG-2 zu ermöglichen, wurden mehr als 1.500 von Regeln definiert, die die im KIS erhobenen Daten für die weitere Arbeit im CADIAG-2 in einem Datenvorverarbeitungsschritt abstrahieren und aggregieren.

Beispiel:

- Bestimmung des Symptoms "symmetrischer Gelenksbefall" aus den Befällen der einzelnen Gelenke
- Bestimmung von "erhöhter Harnsäurespiegel im Serum" auf Grund des gemessenen quantitativen Befundwertes („Fuzzifizierung“)

Das System wurde an etwa 600 klinischen Fällen getestet.

Als Vergleichsdiagnosen wurden die klinischen - ggf. auch die chirurgischen oder pathologisch-anatomischen - Diagnosen der getesteten Fälle herangezogen.

Die Genauigkeit bei der Diagnoseerstellung durch CADIAG-2 lag je nach Erkrankung zwischen 80 und 95 %.

Häufigste Gründe für eine nicht erfolgreiche Computerdiagnose:

- wesentliche Symptome waren auf Grund schon verabfolgter Therapien vermindert oder annulliert
- frühe oder stabilisierte Krankheitsstadien
- unvollständige oder sehr unsichere Anamnese

Beispiel (Arbeitsweise von CADIAG-2 an Hand eines konkreten Falles):

Ersten Teil

Ausgabe der

- aus Anamnese und Status ermittelten Befunde (mit + gekennzeichnet), (die Befunde sind schon das Ergebnis einer erfolgten Datenvorverarbeitung. Diese Datenvorverarbeitung dient zur ersten Aggregation der Patientendaten. So setzt sich der Befund "Derzeitiger Gelenksbefall: symmetrischer Befall" aus einer Vielzahl von anamnestisch erhobenen Einzelbefunden zusammen).
- quantitative Befundergebnisse werden mit dem jeweiligen Befunddatum angezeigt. Bleibt das Befunddatum für mehrere Befunde gleich, so wird eine Mehrfachausgabe des Datums auf der Befundliste unterdrückt.

ANAMNESE

- + Frühere Erkrankung, Cholelithiasis
- + Krankheitsbeginn: zwischen dem 30. und 50. Lebensjahr
- + schleichend
- + Müdigkeit
- + Morgensteifigkeit der Extremitäten > 30 Min.
- + Derzeitiger Gelenksbefall: Befall
 - + Symmetrischer Befall
 - + Befall eines oder mehrerer PIP
 - + Befall eines oder mehrerer MCP
 - + Befall eines oder beider Handgelenke
 - + Befall eines oder beider Kniegelenke
- + Gelenksbefall in den letzten 3 Monaten: Befall
 - + symmetrischer Befall
 - + Befall eines oder mehrerer PIP
 - + Befall eines oder mehrerer MCP
 - + Befall eines oder beider Handgelenke
- + Gelenksbefall vor 3 Monaten: Befall
 - + symmetrischer Befall
 - + Befall eines oder mehrerer MCP

STATUS

- | | | | |
|---|---------------|--------|----------|
| + | Größe | 167 cm | 6.6.1995 |
| + | Körpergewicht | 59 kg | |
| + | adipös | | |
| + | Haut: | feucht | |

+	Lunge:	o.B.
+	Herz:	o.B.
+	Abdomen:	o.B.
+	Leber:	o.B.
+	Nierenlager:	o.B.
.....		

Zweiter Teil der erhobenen Befunde (weitere Status sowie Labor- und Röntgenbefunde).

Auch hier sind die angezeigten Befunde das Ergebnis einer erfolgten Datenvorverarbeitung. Steht statt + ein numerischer Wert, so ist dieser Befund grenzwertig, d.h. unscharf. Je näher der Wert bei +1 liegt, desto stärker trifft der Befund zu.

Beispiel:

Der Wert von 172 U/l für die Alkalische Phosphatase (AP) stellt einen grenzwertigen Befund zwischen "normal" und "erhöht" dar, wobei dieser Wert eher als "erhöht" (Wert 0.6 auf einer Skala von 0 bis +1) interpretiert wird. Ebenso ist das Befundergebnis des Waaler-Rose-Tests grenzwertig (zwischen "negativ" und "positiv").

Ein mit "-H>" gekennzeichnetes Symptom bedeutet, daß dort eine hierarchische Symptom-Symptom-Beziehung in der Wissensbank eingetragen ist, die das Setzen des dazugehörigen Oberbegriffes bewirkt.

+	Gelenksbefall: symmetrischer Befall
+	-H> Faustschlußdefizit rechts 2 cm
+	Fingergelenke: Kapselschwellung eines MCP
+	Kapselschwellung mehrerer MCP
+	Kapselschwellung eines PIP
+	Kapselschwellung mehrerer PIP
+	Druckschmerz eines MCP
+	Druckschmerz mehrerer MCP
+	Druckschmerz eines PIP
+	Druckschmerz mehrerer PIP
+	Schmerz und Schwellung eines MCP
+	Schmerz und Schwellung mehrerer MCP
+	Schmerz und Schwellung eines PIP
+	Schmerz und Schwellung mehrerer PIP
+	Handgelenke: Befall eines oder beider Handgelenke
+	Schwellung eines oder beider Handgelenke
+	Bewegungseinschränkung eines oder beider Handgelenke
+	Kapselschwellung eines oder beider Handgelenke
+	Kniegelenke: Befall eines oder beider Kniegelenke
+	Schwellung eines oder beider Kniegelenke
+	Schmerzhaftigkeit eines Kniegelenks
+	Schmerzhaftigkeit eines oder beider Kniegelenke
+	Bewegungseinschränkung eines oder beider Kniegelenke
+	Kapselschwellung eines oder beider Kniegelenke
+	Erguß eines oder beider Kniegelenke
+	Schmerzhaftigkeit eines oder beider Kniegelenke bei Bewegung
+	Druckschmerz eines oder beider Kniegelenke
+	Füße: Pes transversus
+	Gefäße: Puls o.B.
+	Varizen am rechten Unterschenkel, stark
+	Varizen am linken Unterschenkel, leicht

LABOR

+	BSG: erhöht	42 mm n.W.	6.6.95
0.40	AP: normal	172 U/l	

0.60	AP: erhöht	172 U/l
+	Kreatinin im Serum: normal	0.98 mg/dl
+	Harnsäure im Serum: normal	5.20 mg/dl
+	GOT im Serum: normal	8.00 U/l
+	GPT im Serum: normal	7.00 U/l
+	Rheuma-Latex-Test: positiv	
0.80	Waller-Rose-Test: negativ	16 Titer
	

=====>**BEWIESENE DIAGNOSE**

+ Chronische Polyarthrit

Begründung:

REGEL

WENN

mindestens 5 von 11 Kriterien:

- + Morgensteifigkeit der Extremitäten > 30 Min.
 - + Gelenkschmerzen in mindestens einem Gelenk bei Bewegung oder auf Druck
 - + Weichteilschwellung in mindestens einem Gelenk
 - + Schwellung von mindestens 2 Gelenken innerhalb von 3 Monaten
 - + Symmetrischer Gelenksbefall
 - Rheumaknoten
 - + Typische Röntgenzeichen für Arthritis
- 0.20 Positiver Rheumafaktornachweis
- ? Typische Synovialflüssigkeitsveränderungen (schwaches Mucinpräzipitat)
 - ? Typische histologische Veränderungen der Synovialmembran bei Arthritis
 - ? Hautbiopsie: Typischer Rheumaknoten

UND

nicht mindestens 1 von 26 Kriterien:

- Krankheitsbeginn: vor dem 16. Lebensjahr
- Schmetterlingsexanthem
- ? Lupus erythematoses Zellen: positiv
- Muskelerkrankung der Extremitäten
- ? Muskelbiopsie: Panarteriitis nodosa
- ? Nierenbiopsie: Panarteriitis nodosa
- ? Hautbiopsie: Sklerodermie
- Diffuse Sklerodermie
- Fleckige Sklerodermie
- Maskengesicht
- -H>Rheumatisches Fieber

0.05 Gicht

- Tophus
- Verdacht auf Psoriasis-Arthritis

0.25 Bakterielle Arthritis

0.05 Bakterielle Arthritis: TBC-Arthritis

0.10 Schulter-Arm-Syndrom

0.01 Neurogene Arthropathie

- ? Homogentisinsäure im Harn: positiv
- ? Histologischer Nachweis einer Sarkoidose
- ? Plasmozytom
- Erythema nodosum
- ? CPK im Serum: stark erhöht
- ? Thrombophlebitis
- ? Malignes Lymphom

mindestens 3 von 5 Kriterien:

- + Schmerz und Schwellung mindestens eines Gelenkes

- Konjunktivitis
- Urethritis
- Diarrhoe
- ? Balanitis oder Keratodermie

DANN

+ Chronische Polyarthrit

Die Diagnose "Chronische Polyarthrit" wird vom Konsultationssystem CADIAG-2 als bewiesene Diagnose für den Patienten ausgegeben. Als Begründung wird die für die chronische Polyarthrit von den Patientendaten erfüllte ARA-Regel aus dem Jahr 1958 angegeben, die eine Beweiskraft von +1 hat. Wegen der dort vorgesehenen Ausschlusskriterien scheint diese Regel für die Diagnosefindung günstiger als die revidierte, einfachere ARA-Regel aus dem Jahr 1987.

Die Computerausgabe zeigt explizit an, welche der geforderten ARA-Kriterien

- erfüllt sind (gekennzeichnet durch +)
- definitiv nicht erfüllt sind (gekennzeichnet durch -)
- grenzwertig erfüllt sind (gekennzeichnet durch Werte zwischen 0 und 1)
- nicht bestimmt werden konnten (gekennzeichnet durch ?)

Der Beispielpatient erfüllt von 5 notwendigen der 11 möglichen Kriterien zur Diagnose der "Chronischen Polyarthrit" 6 Kriterien definitiv (gekennzeichnet durch +) und eines grenzwertig (gekennzeichnet durch 0.20). Keines der zusätzlich angeführten 26 zu prüfenden Ausschlusskriterien ist zur Gänze erfüllt, so daß die Konklusion der Regel

Bewiesene Diagnose: "Chronische Polyarthrit"

ausgegeben werden kann.

Die insgesamt 26 Ausschlusskriterien wurden bei der automatischen Diagnose der "Chronischen Polyarthrit" als zusätzliches Sicherheitsnetz eingebaut; sie gehen teilweise über die Definitionen der ARA-Kriterien hinaus.

====>DIAGNOSEHYPOTHESEN

0.80 Seronegative Chronische Polyarthrit, Stadium I

Begründung:

LABOR

0.80 Waaler-Rose-Test: negativ 16 Titer 6.6.95

REGELWENN

+ Chronische Polyarthrit

UND

- Rheuma-Latex-Test: negativ

oder

+ Rheuma-Latex-Test: positiv
und

0.80 Waaler-Rose-Test: negativ

und nicht

- Röntgen: Arthritiszeichen, Erosionen, Usuren der Gelenke

oder

- Röntgen: Subluxation der Gelenke

oder

- Röntgen: Ankylose der Extremitätengelenke

DANN

0.80 Seronegative Chronische Polyarthrit, Stadium I

Die Diagnose "*Seronegative chronische Polyarthrit*is, Stadium I" wird als Diagnose-hypothese angeführt.

Der Waaler-Rose-Test ist bei dem Patienten grenzwertig negativ (gekennzeichnet durch die Angabe 0.80), so daß die Diagnose einer seronegativen Arthritis nur als Hypothese gestellt werden kann.

Die Hypothese „*seropositiven chronischen Polyarthrit*is" wird nicht gestellt, da die Aussage "positiver Waaler-Rose-Test" nur mit 0.20 bewertet wurde.

Die Angabe, daß sich die bewiesene chronische Polyarthrit

is im Stadium I befindet, wird jedoch durch die Röntgenbefunde

- "keine Erosionen oder Usuren der Gelenke"
 - "keine Subluxation der Gelenke"
 - "keine Ankylose der Extremitätengelenke"
- abgesichert (gekennzeichnet durch -).

2.4.6.5.1 Diskussion

CADIAG-2 gibt dem Arzt ein Hilfsmittel in die Hand, das zu einer Verbesserung sowie Beschleunigung der klinisch-diagnostischen Entscheidung im konkreten Fall führen kann.

Einsatzmöglichkeiten:

- als elektronisches Nachschlagewerk für mögliche Diagnosen, wobei die Eingabe von einem oder wenigen Symptomen zur Ausgabe aller möglichen Diagnosen führt
- als komplexes on-line Konsultationssystem für spezielle Fälle zur detaillierten und vollständigen differentialdiagnostischen Abklärung der Erkrankung des Patienten.

ABER:

CADIAG-2 ist - obwohl es auf umfangreiches medizinisches Wissen zurückgreifen kann und anhand von hunderten klinischen Fällen getestet wurde - ein experimentelles System, daß weiter ausgebaut und verifiziert werden muß.

Weiterentwicklungen:

- CADIAG-3 verarbeitet rekursive Ausdrücke und verwendet unterschiedliche Fuzzy-Operationen.
- In CADIAG-4 erfolgt eine Erweiterung auf Basis von Client-Server-Technologien mit Anbindung an Systemen der Radiologie und des zentralen Labors sowie an die Literaturdatenbank MEDLINE.

2.4.7 Modellbasierte Diagnostik-Expertensysteme

Kausale Modelle sind eine eher deklarative Form der Wissensrepräsentation, mit denen Struktur, Verhalten und Funktion eines Systems relativ unabhängig von einer konkreten Problemstellung beschrieben werden.

Bei kausalen Modellen ist eine Diagnose dann konsistent, wenn alle beobachteten Symptome aus ihr abgeleitet werden können.

Die einfachste Wissensrepräsentation dafür ist die direkte Zuordnung von Ursachen (Diagnosen) zu ihren Wirkungen (Symptomen), z.B.

$$D_1 \rightarrow \{S_1, S_2, S_3\} \quad \text{und} \quad D_2 \rightarrow \{S_1, S_3, S_4, S_5\}$$

Die kausale Diagnostik ist dann eine Anwendung von Algorithmen zum Finden einer minimalen Mengenüberdeckung.

Die modellbasierten Diagnostiksysteme werden unterteilt in:

- funktionale Diagnostik (Verwendung: meist in technischen Bereichen)
- überdeckende Diagnostik (Verwendung: meist in medizinischen Bereichen)

Bei der überdeckenden Diagnostik wird eine Lösung gesucht, mit der alle Symptome erklärt werden können. Dabei versucht man mit einer einfachen Wissensrepräsentation, nämlich „Zustand 1 verursacht Zustand 2“, auszukommen. Dabei sind

- „Endzustände ohne Wirkungen“ Symptome (Zustand 3 in der Abbildung 1.9)
- „Anfangszustände ohne Ursachen“ Diagnosen (Zustand 1, Zustand 3a)

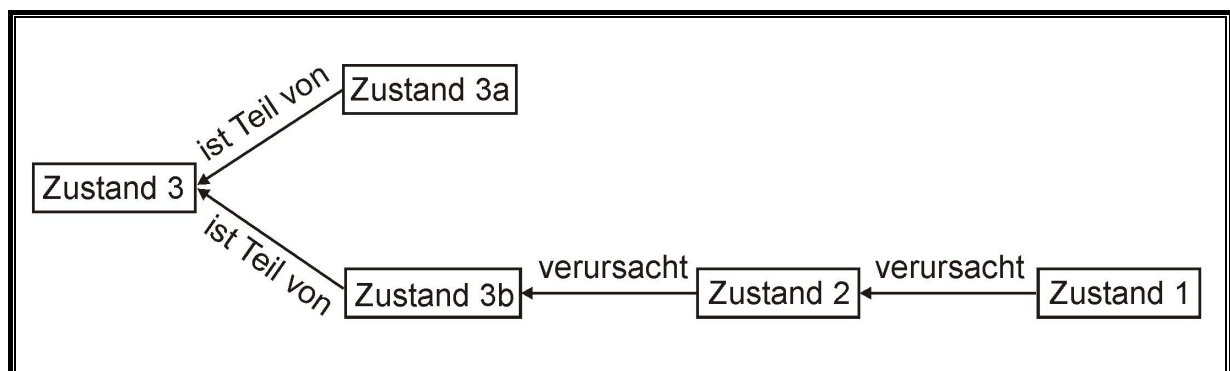


Abb.1.9. Wissensrepräsentation bei der überdeckenden Diagnostik [Puppe (1991) 89]

Zur Verbesserung des Systems werden die Zustände und ihre Beziehungen mit Parametern wie

- Zeitpunkt
 - Dauer
 - Schweregrad
- versehen.

Wenn in einem kausalen Modell verschiedene Erklärungen für gegebene Symptome existieren, wird die einfachste Erklärung als beste Lösung gewählt („Occam’s Razor“-Entscheidungsmethode).

Nachteile (von kausalen Modelle):

- unvollständiges Wissen
- Ineffizienz

Nicht alle Wechselwirkungen können in ihm berücksichtigt werden, da die Beschreibung zu komplex würde. Vor allem kann über die Wahrscheinlichkeit des Auftretens möglicher Ereignisse in kausalen Modellen nicht viel ausgesagt werden, da diese auch von der Vorgeschichte des Systems abhängt. Kausale Modelle können also nur mögliche Erklärungen geben, die durch Erfahrungswissen überprüft werden müssen.

2.4.7.1 CADUCEUS

CADUCEUS wurde zu Beginn der 70-iger Jahre vom Informatiker Harry E. Pople gemeinsam mit dem Facharzt für innere Medizin Dr. Jack D. Myers entwickelt. Als Weiterentwicklung des rein assoziativen Vorgängersystems INTERNIST deckt es fast das gesamte Gebiet der inneren Medizin ab; derzeit sind in der Datenbank über 600 Krankheiten gespeichert. Es repräsentiert und verwendet eine große Diagnosehierarchie, in der 2 Typen von Beziehungen unterschieden werden:

- Nosologische (hierarchische) Beziehungen:
A ist eine Unterdiagnose von B, wenn A spezieller ist und mindestens ein wichtiges Symptom mit B gemeinsam hat.
- Pathophysiologische Beziehungen:
A ist durch B verursacht.

Die Inferenzstrategie von CADUCEUS:

Erste Phase:

Mit Hilfe der vorhandenen Symptome werden einige allgemeine Pathokonzepte aufgestellt.

Zweite Phase:

Das System versucht, die so etablierten Pathokonzepte möglichst einfach zu erklären. Hierzu werden 6 Syntheseoperatoren benutzt. Sie repräsentieren die verschiedenen primitiven Verknüpfungsmöglichkeiten zweier Knoten (Symptom oder Pathokonzept) im Netzwerk über nosologische und pathophysiologische Beziehungen:

- A ist Unterdiagnose von B.
- A ist Ursache von B.
- A und B haben eine gemeinsame Ursache.
- A und B haben eine gemeinsame Unterdiagnose.
- A hat als Ursache eine Unterdiagnose von B.
- A und B haben Ursachen, von denen die eine die Unterdiagnose der anderen ist.

CADUCEUS verfügt über Planungsbeziehungen, durch welche die Suche zwischen Knoten, die den Einsatz mehrerer Syntheseoperatoren nötig machen, effizienter gestaltet werden kann. Wenn ein Knoten zu wenigstens einer Unterdiagnose der Grobdiagnose eine Kausalbeziehung hat, wird der Knoten mit allen Grobdiagnosen verbunden. Bei der Etablierung einer Wissensbank werden die Planungsbeziehungen automatisch hergestellt. Da CADUCEUS pathophysiologische Beziehungen verwendet, ist es im Vergleich zu INTERNIST eher geeignet, zusammenhängende Erklärungen für Symptome aus verschiedenen Organsystemen zu finden. Die Verdachtsgenerierung ist bei CADUCEUS im Vergleich zu INTERNIST wesentlich verbessert.

2.4.7.2 ABEL

ABEL wurde 1981 von Ramesh, Patil, Schwarz und Szolovitz entwickelt.

Es ist ein System, das kausal Wissen repräsentiert und dabei ein wesentlich präziseres pathophysiologisches Modell als CADUCEUS benutzt. ABEL ist eines der wenigen Expertensysteme, bei dem die Reihenfolge der zu stellenden Fragen eine Rolle spielt.

Die Merkmale von ABEL sind

- Krankheitsbeschreibungen auf verschiedenen Abstraktionsebenen
- Entwicklung von Techniken, die verschiedene Effekte zu einem Schweregrad eines Zustandes quantitativ summieren.

In ABEL kann eine quantitative Beschreibungsebene verwendet werden, da das System in seinem Anwendungsbereich (Störungen im Säure-, Base- und Elektrolythaushalt) pathophysiologische Zustände mittels Laborwerten quantitativ errechnen kann.

Nach Eingabe der Labordaten (PH-Wert und Konzentration verschiedener Elektrolyte: Na, K, CL, HCO_3 , pCO_2) baut ABEL alle konsistenten patientenspezifischen Modelle (PSM's) auf.

Ein pathophysiologischer Zustand besitzt verschiedene Attribute wie

- Schweregrad
- Startzeit
- Dauer

Die kausalen Beziehungen zwischen verschiedenen Zuständen werden durch Constraints (mapping relations) dargestellt, die in beiden Richtungen (Ursache → Wirkung und Wirkung → Ursache) ausgewertet werden können.

Ein Constraint (Abb. 1.10.) ist als eine Menge von LISP-Funktionen implementiert, die die Attribute des einen Zustandes aus den Attributen des anderen Zustandes berechnen. Zusätzlich können Kontext- und Default-Bedingungen für einen Constraint angegeben werden.

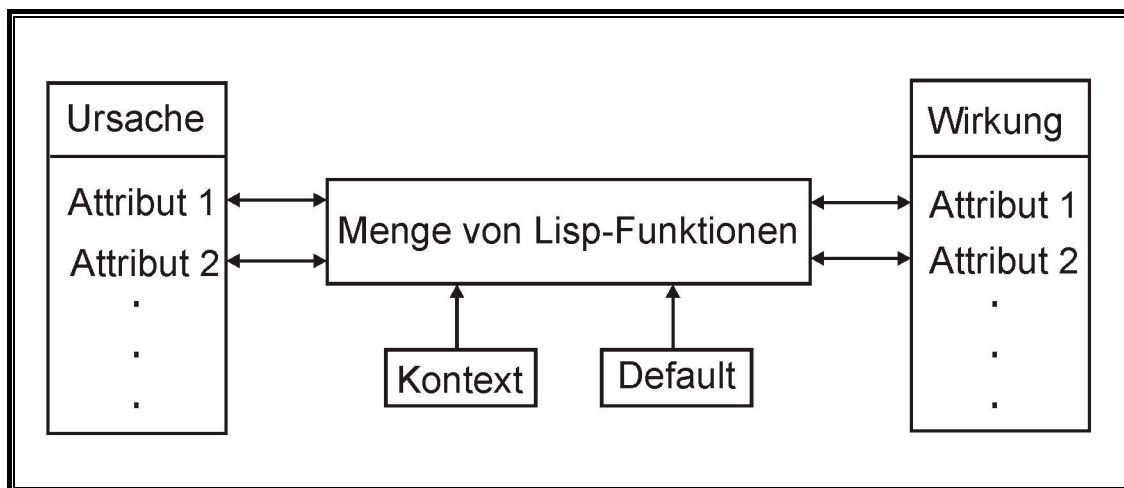


Abb. 1.10. Relation in ABEL zwischen 2 pathophysiologischen Zuständen (Ursache und Wirkung)
Beispiele für Attribute eines Zustandes sind Schweregrad, Startzeit und Dauer.

In ABEL werden Rückkopplungsschleifen repräsentiert, d.h. kausale Beziehungen, bei denen ein Zustand über verschiedene Zwischenzustände auf sich selbst verstärkend oder abschwächend zurückwirkt.

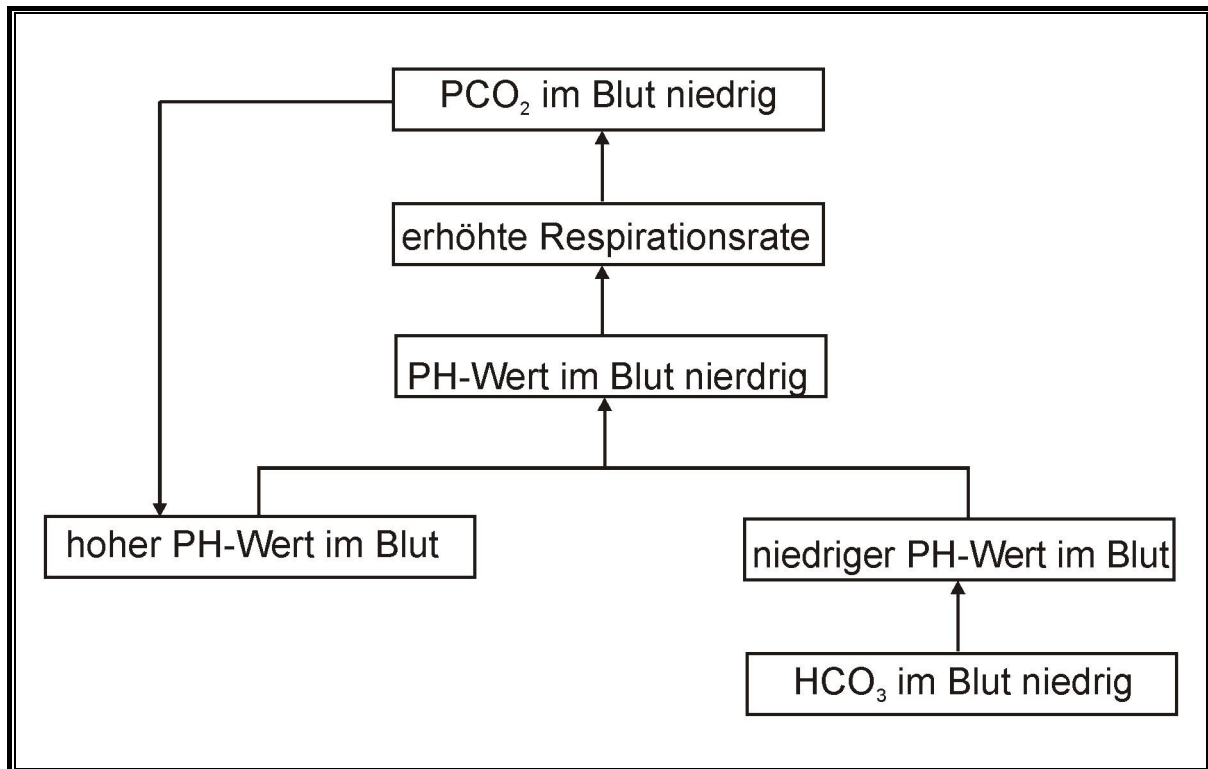


Abb. 1.11. Rückkopplungsschleife in ABEL [Puppe (1987) 47]

ABEL ordnet quantitativen Meßwerten im Zuge der Datenvorverarbeitung qualitative Zustände zu. Unter Berücksichtigung der Attribute wie Schweregrad, Startzeit und Dauer werden dann die Abhängigkeiten der Zustände untereinander ermittelt.

Die Attribute dienen

- zum Überprüfen, ob ein Zustand stark genug ist, einen anderen zu verursachen
- zur Berechnung der Wirkungen verschiedener Ursachen

Zuerst werden ein Modell bzw. bei nicht eindeutigen Fällen mehrere Modelle aufgebaut, um den Zustand des Patienten zu beschreiben.

Dann muß eine minimale Menge von Diagnosen gefunden werden, die die unerklärten Zustände im patientenspezifischen Modell verursachen können.

Verschiedene kausale Beziehungen können die Komplexität und Genauigkeit eines Modells beeinflussen. Einige dieser Kausalbeziehungen sind:

- verschiedene Typen von Beziehungen
z.B. linear, multiplikativ, reversibel und irreversibel
- Abhängigkeit einer Beziehung von
 - verstärkenden oder abschwächenden Faktoren (Katalysatoreffekt)
 - Randbedingungen
- zeitliche Beziehungen zwischen Ursache und Wirkung
z.B. gleichzeitig, verzögert oder überlappend
- multikausale Beziehungen

2.5 Statistische Expertensysteme

Die statistische Argumentation hat in den letzten Jahren für die medizinische Forschung zunehmend Bedeutung erlangt. Durch die steigende Verfügbarkeit von relativ leicht zu benutzenden Statistik-Softwaresystemen ist ihr Einsatz zur raschen und effizienten Datenanalyse und Hypothesentestung weit verbreitet. Daher sind nicht nur Statistiker Benutzer derartiger Programmsysteme, sondern vielfach auch Substanzwissenschaftler (z.B. Mediziner), die eine Unterstützung bei der korrekten Auswahl des einzusetzenden statistischen Verfahrens benötigen.

Durch die Entwicklung von "intelligenter" Software wird versucht - zumindest im Fall von Routineproblemen - die benötigte statistische Beratung bereitzustellen. Solche Programmsysteme simulieren in einem Mensch-Maschine-Dialog, wobei das System die Funktion des Statistikexperten übernimmt, das Wechselspiel einer statistischen Beratung. Auf Grund der jeweiligen Problemspezifikation durch den Benutzer wird versucht, eine zielführende statistische Vorgangsweise abzuleiten. Auf diese Weise sollen

- methodische Fehler vermieden
- qualifizierte Fachstatistiker von Routinekonsultationen entlastet und dadurch freie Kapazität für anspruchsvollere Beratungstätigkeiten geschaffen werden.

Statistische Expertensysteme werden seit etwa 1980 diskutiert. Trotz einer Fülle von Entwicklungsansätzen und Vorschlägen zur Gestaltung von statistischen Expertensystemen sind praktisch brauchbare Systeme nur in geringem Umfang verfügbar. Soweit solche Programme überhaupt das Stadium akademischer Prototypen überschritten haben, decken sie nur kleine Methodenbereiche der Statistik ab.

Existierende statistische Expertensysteme		
Name	Bereich	Entwickler
GLIMPSE	Lineare Modelle	Imperial College of Science and Technology, London
CADEMO	Statistische Versuchsplanung	Forschungszentrum für Tierproduktion, Dummerstorf-Rostock
KENS	Nichtparametrische Statistik	Institute of Psychology, London

Die Beschränkung statistischer Expertensysteme auf kleine Methodenbereiche ist problematisch, da zur Beurteilung der sinnvollen Verwendbarkeit eines derartigen Systems weitreichende statistische Kenntnisse beim Benutzer erforderlich sind. Reine statistische Beratungssysteme operieren überdies oftmals unabhängig von konkreten Beobachtungsdaten und reagieren daher nicht auf offenkundige Anomalien in den Daten oder auf Inkonsistenzen zwischen Daten und verwendetem Modell. Ebenso wenig können sie den Statistikanwender von den z.T. umfangreichen technischen Details der Durchführung einer statistischen Auswertung entlasten.

Die Realisierung von statistischen Expertensystemen erweist sich aus folgenden Gründen als schwierig:

- die Breite des abzudeckenden statistischen Verfahrensspektrums
- die Diversität statistischer Strategien
- die häufig unpräzisen Ausgangsfragestellungen
- die unumgängliche Formalisierung der statistischen Entscheidungslogik

2.5.1 Diskussion

Die statistische Beratungstätigkeit umfaßt

- die Versuchsplanung
- die statistische Modellformulierung
- die Abwicklung konkreter Auswertungen
- die Darstellung und Interpretation der Ergebnisse.

Angesichts dieses breiten Spektrums scheint gegenwärtig die praktische Realisierung eines statistischen Konsultationssystems nur bei Beschränkung auf ein sehr kleines Anwendungsgebiet erfolgversprechend zu sein; selbst dann sind tiefliegende methodische Probleme (z.B. die Frage des multiplen Testens) zu klären. Umgekehrt wäre eine Einbeziehung medizinischen Wissens zu überlegen, um die Systembenutzer hinsichtlich der statistischen Modellformulierung besser unterstützen zu können.

Die Entwicklung statistischer Konsultationssysteme für Nichtstatistiker wirft das Problem der Verantwortlichkeit verstärkt auf:

Expertensysteme, die - wie in der Medizin - einen Experten unterstützen sollen, lasten legitimerweise die Letztverantwortlichkeit dem Anwender auf.

Bei einem System für statistische Laien ist dies schwerlich vertretbar und daher liegt ein wesentliches Moment der Systemgestaltung in der Erreichung einer gewissen Sensitivität gegen eine inadäquate, jedoch unbeabsichtigte Verwendung des im System abgebildeten Methodenbereiches.