INFORMATIK BIOMETRIE und EPIDEMIOLOGIE

IN MEDIZIN UND BIOLOGIE

Bd. 26 (1995) Inhaltsverzeichnis

BRIGL, B.; RINGLEB, P.; STEINER, T.; MANN, G.; LEINER, F.; GRAU, A.; HACKE, W.; HAUX, R.: Multiple Verwendbarkeit Klinischer Dokumentationen am Beispiel eines wissensbasierten klinischen Arbeitsplatzsystems in der Neurologie 240

BRÜNING, J.; BECKER, R.; ENTEZAMI, M.; LOY, V.; H.; VONK, R.; WEITZEL, TOLX-DORFF, T.: Das wissensbasierte System ADNEXPERT zur Unterstützung der sonografischen Diagnose von Adnextumoren 166

BÜRSNER, S.; SPRECKEISEN, C.; NEBE, J.; SPITZER, K.: Epistemologische Adäquatheit bei der Akquisition und formalen Repräsentation neurologischen Wissens 229

DONG, M.; KRÄMER, U.; DOLGNER, R.; RANFT, U.: Population-Based Attributable Risks of Respiratory Diseases Resulting from Air Pollution: A re-analysis of data from the study on 55 year old women in the Ruhr district, 1990 289

ENGLER, C.; FÜHRER, A.; HEMPEL, P.; BUSCHER, H.-P.: HEPA-CADS : Entwicklung eines tutoriell nutzbaren Expertensystems zur Diagnostik von Lebererkrankungen 275

EYSENBACH, G.: Computereinsatz und Computerkenntnisse unter Medizinstudenten 56

FÖRSTER, M.; WASCHULZIK, T.; KIRCHNER, K.; MOSER, W.; BÖHM, V.; BRAUER, W.; ENGELBRECHT, R.; KOSCHINSKY, T.; ENTENMANN, G.; DIADOQ: Förderschwerpunkt Wissensbasen in der Medizin (MEDWIS) 198

GIERL, L.; ERTÜZÜN, S.; POLLWEIN, B.; SCHILLING, V.; HOERNSTEIN, M.; KRÄMLING, H.-J.; HORNUNG, H. M.: Handlungszyklen in einem heterogenen KKS: Dezentrale Systeme und Kommunikation am Beispiel der operativen Therapie 33

HEINRICH, J.; WICHMANN, H.-E.; BECK, E.; NOWAK, D.; JÖRRES, R.; MAGNUSSEN, H.; CLAUSSEN, M.; BERGER, J.: Die Verbreitung von Asthma und Atemwegssymptomen bei Erwachsenen in Erfurt und Hamburg. Erste Ergebnisse der deutschen Zentren des EC Respiratory Health Survey 297

HOLLÄNDER, N.; RZANY, B.; MOCKENHAUPT, M.; SCHÖPF, E.; SCHUMACHER, M.: Statistische Ansätze zur Beurteilung seltener Arzneimittelrisiken in einer Fallstudie 154

INGENERF, J.: Konzeptuelle Meta-Modellierung zur Steuerung zu einer kompositionellen semantischen Sprachanalyse – am Beispiel lexikalischer Ambiguität 18

JOSTES, C.; PACZKOWSKI, J.; SCHRÖDER, J. P.: Integration im Krankenhaus – Offene Kommunikation mit dem standardisierten Anwendungsprotokoll HL7 47

KNOLLE, H.: Die Aids-Epidemie unter homosexuellen Männern: Vergleich einer Modellrechnung mit Daten aus der Schweiz 7

PESCH, B.; JÖCKEL, K. H.; WICHMANN, H.-E.: Luftverunreinigungen und Lungenkrebs 134

PIEPHO, H.-P.: Methods for estimating missing genotype-location combinations in multilocation trials – an empirical comparison 335

PROKOSCH, H. U.; DUDECK, J.: Evaluation der Akzeptanz eines Krankenhaus-Informationssystems bei den ärztlichen Mitarbeitern eines Universitätskrankenhauses 107

PROKOSCH, H. U.; BÜRKLE, T.; STORCH, J.; STRUNZ, A.; MÜLLER, M.; DUDECK, J.; DIRKS, B.; KELLER, F.: MDD-GIPHARM: Design and Realization of a Medical Data Dictionary for Decision Support Systems in Drug Therapy 250

OHMANN, C.; PLATEN, C.; BELENKY, G.; FRANKE, C.; OTTERBECK, R.; LANG, K.; RÖHER, H.-D.: Expertensystem zur Unterstützung von Diagnosestellung und Therapiewahl bei akuten Bauchschmerzen 262

OPPEL, U. G.; FISCHER, U.; HEPP, K. D.; HERFURTH, S.; HIERLE, A.; LIESENFELD, B.; LÜDDEKE, H.-J.; SALZSIEDER, E.; SELL, C.: Kausal-probabilistische Expertensysteme im MEDWIS-Projekt DIADOQ 215

SAKIA, R. M.: An Empirical Comparison of three bias Estimating Procedures due to Retransformation 1

SCHOPEN, M.: Die logische Struktur der ICD-10 (Systematik) und ihre Beschreibung mit SGML 121

SWART, E.; RICHTER, I.; ROBRA, B.-P.; GRAD, A.: Kein erhöhtes Risiko für Kolonkarzinome nach Entfernung kolorektaler Adenome – Ergebnisse eines flächendeckenden registergestützten Follow-up 308

TEILLE, K.; JUNG, E.; ENGELBRECHT, R.: Medizinische Begriffe im DIABCARD-Repository 321

Tutorial

WELLEK, S.: Einführung in die statistische Methodik von Äquivalenzstudien 81

Editorial 185

Berichte

EIMEREN, W. VAN: MEDWIS 67

Empfehlungen

WINDELER, J.; TRAMPISCH, H. J.: Empfehlungen zur Durchführung von Studien zur therapeutischen Äquivalenz 350

Buchbesprechungen 78, 181, 286

Institut für Medizinische Biometrie
und Medizinische Informatik
der Universität Freiburg
Abt. Medizinische Biometrie u. Statistik
Stefan-Meier-Straße 26
79104 Freiburg

Editorial

Die Forschung im Bereich der wissensbasierten Systeme in der Medizin war in den letzten Jahren einem starken Wandel unterzogen. Auch in Fachkreisen der Medizinischen Informatik wurde dieses Forschungsgebiet noch vor fünf Jahren überwiegend mit dem Begriff ‚Grundlagenforschung' assoziiert. Das Forschungsinteresse galt Fragen wie beispielsweise der Verarbeitung unsicheren Wissens mittels ‚Neuronaler Netze' oder ‚Kausal Probabilistischer Netze'. Häufig diente das Anwendungsgebiet Medizin nur der Evaluation neuer methodischer Ansätze. Zwar besteht nach wie vor ein großer Bedarf an Grundlagenforschung – so ist zum Beispiel das Problem der Repräsentation zeitlichen Wissens noch unzureichend gelöst – inzwischen existiert jedoch ein für viele medizinische Problemstellungen ausreichendes Repertoire an Methoden. Damit gewinnen Probleme, die direkt mit dem Einsatz wissensbasierter Systeme im medizinischen Umfeld verbunden sind, zunehmend an Bedeutung für die Forschung. Ziele einer solchen anwendungsorientierten Forschung sind u. a. die Integration wissensbasierter Systeme in Krankenhausinformationssysteme, die Unterstützung der Pflege und Fortschreibung existierender Wissensbasen oder die Evaluation wissensbasierter Systeme.

Die Diskussion über die Bedeutung wissensbasierter Systeme für die einzelnen Einsatzbereiche in der Medizin wird nicht mehr durch die prinzipielle Realisierbarkeit von Systemen bestimmt, sondern durch die Bedingungen, unter denen ihre Entwicklung und ihr Einsatz stattfinden. Dazu gehören die Kosten, die mit der Konstruktion und dem Einsatz wissensbasierter Systeme verbunden sind, ebenso wie die Sicherheit der Systeme und die Gewährleistung einer kontinuierlichen Pflege des Wissens.

Diese Entwicklung spiegelt sich auch in dem Förderschwerpunkt ‚Medizinische Wissensbasen' (MEDWIS) des Bundesministeriums für Bildung, Wissenschaft, Forschung und Technologie wider. In der laufenden Phase II von MEDWIS werden, beginnend mit dem Jahr 1992, derzeit ca. 40 Forschungsvorhaben aus den Bereichen Neurologie, Operative Therapie, Stoffwechselerkrankungen, Arzneimittel, Intensivmedizin, Labormedizin, bildgebende Verfahren sowie Lehr- und Tutoringsysteme gefördert (eine ausführliche Darstellung der Ziele und des Standes des Förderschwerpunktes MEDWIS findet sich in [VAN EIMEREN 1995]). Die Vorhaben sind zu Verbunden zusammengefaßt, in denen medizinische, medizininformatische und teilweise auch industrielle Partner zusammenarbeiten. Die Phase II ist geprägt durch die Entwicklung oder Adaptierung geeigneter Methoden und Werkzeuge, die zur Lösung konkreter medizinischer Probleme in den oben genannten Bereichen verwendet werden können. Die Eignung dieser Methoden und Werkzeuge soll zum Ende der Phase II durch die Entwicklung von Demonstratoren belegt werden. Erfahrungen aus den einzelnen Vorhaben werden in informatischen Arbeitskreisen zusammengetragen und weitergegeben. Ziel der ab Herbst 1995 beginnenden Phase III ist die Weiterentwicklung von Demonstratoren zu im klinischen Umfeld einsetzbaren Prototypen. Eine verstärkte Einbeziehung industrieller Partner soll dabei den Wissenstransfer von Forschungseinrichtungen zur Industrie verbessern, die Integration in existierende EDV-Anwendungen fördern sowie die Multiplikation und Pflege der entwickelten Wissensbasen unterstützen.

Das hier vorliegende Heft ist das zweite Themenheft, das ausschließlich dem Forschungsgebiet der Medizinischen Wissensbasen gewidmet ist. Aus dem Förderschwerpunkt MEDWIS wurden Arbeiten ausgewählt, die ein möglichst breites Spektrum grundlagen- und anwendungsorientierter Forschungstätigkeit repräsentieren. Wenn das Themenheft auch kein geschlossenes Bild der Forschung im Bereich medizinischer Wissensbasen in Deutschland präsentieren kann, so werden doch wesentliche Facetten beispielhaft deutlich.

G. Mann
(MEDWIS-Projektkoordination)

VAN EIMEREN, W. (1995): MEDWIS. Informatik, Biometrie und Epidemiologie in Medizin und Biologie **26** (1), 67–77.

Informatik, Biometrie und Epidemiologie in Medizin und Biologie **26** (3),
187 – 197, ISSN 0943-5581
© Verlag Eugen Ulmer GmbH & Co., Stuttgart; Gustav Fischer Verlag KG, Stuttgart

Eine Entscheidungshilfe zum Umgang mit unsicherem und zeitlichem Wissen

Decision-Support-System for Handling Uncertain and Temporal Knowledge

Kerstin Schill, Elisabeth Umkehrer, Ernst Pöppel

Kurztitel:
Entscheidungshilfe

Schlagworte:
Wissensrepräsentation, zeitliches Schließen, unsicheres Schließen, Entscheidungshilfe

Keywords:
Knowledge representation, temporal reasoning, uncertain reasoning, decision-support system

Zusammenfassung

Das Wissen über einen Bereich (z. B. Intensivmedizin) liegt meist nur mit einer gewissen (Un-)Sicherheit vor, die sich sowohl auf das „statische" wie auch auf das temporale Wissen beziehen kann. Eine zentrale Frage bei der Entwicklung jedes wissensbasierten Systems ist daher die Auswahl einer geeigneten Theorie, mit der die Unsicherheit und Temporalität des Wissens erfaßt werden kann. Wir haben deshalb ein erstes Konzept einer automatisierten Entscheidungshilfe entwickelt, die den Benutzer bei der Frage unterstützen soll, welche Theorie geeignet ist, um das gegebene Wissen adäquat zu repräsentieren.
*Für das Modul zur Auswahl von temporalen Ansätzen wurde das „Intervall" als einheitliche Bezugsgröße gewählt. Dabei wurden bisher zwei temporal-logische Ansätze integriert, der Intervallansatz (*ALLEN *1983, 1984) und der Semi-Intervall-Ansatz (*FREKSA *1992). Abhängig von der zeitlichen Unsicherheit und Ungenauigkeit des Wissen kann automatisch die geeignete Methode zum Schließen gewählt werden. Desweiteren wurde der Umgang mit Referenzintervallen und ebenso das Konzept des Referenzzeitpunktes eingebunden, um eine Teilwelt auswählen und innerhalb dieser zeitlich schließen zu können.*

Für die Repräsentation von Unsicherheitstheorien wurde eine auf Carnap (CARNAP1950) basierende Metasprache entwickelt, die einen einheitlichen formalen Bezugsrahmen bietet, um die folgenden unterschiedlichen Unsicherheits-Theorien voneinander abzugrenzen und die Auswirkungen ihrer Schlußfolgerungstechniken darzustellen: die Bayes Theorie (PEARL 1986), (PEARL 1988), (CHARNIAK 1983), die Dempster-Shafer Theorie nach Shafer (SHAFER 1976), die Dempster-Shafer Theorie nach Smets (SMETS 1988), (SMETS und KENNES 1989), Fuzzy-Set basierte Theorien (ZADEH 1965), (ZADEH 1979), (DUBOIS und PRADE 1980), sowie die Upper/Lower Probability Theorie nach Fagin und Halpern (HALPERN und FAGIN 1992).

Summary

Knowledge concerning a particular domain (e.g., the medical domain) is generally uncertain with respect to both "static" and temporal knowledge. Therefore, a central question concerning the development of every knowledge-based system is the selection of an appropriate theory with which the uncertainty and temporality of knowledge can be captured. For this reason we have developed a preliminary concept of an automated decision-support system, which supports the user in the selection of an appropriate theory for representing the given knowledge.

In the module for the selection of temporal approaches, we have used the interval as the temporal primitive. We have integrated two temporal logic approaches, the interval-approach and the semi-interval-approach. Depending on the temporal uncertainty and inaccuracy of the knowledge, the appropriate method for reasoning is chosen. Furthermore, the system allows the use of points in time and intervals of time as a reference for the selection of parts of the temporal knowledge.

For the representation of uncertainty theories, we developed a meta-language based on the work of Carnap (CARNAP 1950), which provides a formal framework to delimit the following uncertainty theories, and to represent the effect of their reasoning techniques: the Bayes theory (PEARL 1986), (PEARL 1988), (CHARNIAK 1983), the Dempster-Shafer theory according to Shafer (SHAFER 1976), the Dempster-Shafer theory according to Smets (SMETS 1988), (SMETS and KENNES 1989), the Fuzzy Set based theories (ZADEH 1965), (ZADEH 1979), (DUBOIS and PRADE 1980), and the Upper/Lower-Probability theory according to Fagin and Halpern (HALPERN und FAGIN 1992).

1 Einführung

Die theoretischen Ansätze, die zum Umgang mit unsicherem und zeitlichem Wissen existieren, sind umfangreich und problemspezifisch, d. h. es gibt keinen besten Ansatz für alle Problemstellungen (SCHILL 1990), (SAFFIOTTI 1987). Dieser Umstand bedeutet, daß bei der Entwicklung medizinischer wissensbasierter Systeme in zeitaufwendiger Weise untersucht werden muß, welcher Ansatz zur Wissensmodellierung geeignet ist. Erschwerend kommt dabei hinzu, daß die Darstellung der existierenden Theorien nicht aus einer einheitlichen Sichtweise erfolgt, sondern nahezu jeder Ansatz gemäß einer eigenen formalen Sprachtradition formuliert wird. Dies erschwert die Theorien voneinander abzugrenzen und ihre Techniken zum Schließen zu vergleichen. Das hat häufig zur Folge, daß gewisse Formalisierungsmöglichkeiten gar nicht erst einbezogen werden (wie z. B. für lange Zeit der Ansatz von Dempster/Shafer) oder nur gewisse Aspekte aus den theoretischen Ansätzen extrahiert werden, mit der Konsequenz, daß diese Methoden im formalen Sinne nicht mehr evaluiert werden können (SCHILL 1990).

Im folgenden sollen zwei Module vorgestellt werden, die eine Basis für eine Entschei-

dungshilfe bilden sollen. Ziel ist es, jeweils mit einer einheitlichen „Sprache" die verschiedenen Ansätze zum Umgang mit zeitlichem und unsicherem Wissen darstellen und voneinander abgrenzen zu können, um damit dem Benutzer die Entscheidung einer geeigneten Auswahl zu vereinfachen.

2 Zeitliches Wissen

Die existierenden Theorien zum Umgang mit zeitlichem Wissen lassen sich in zustandsbasierte Methoden, wie das Situationenkalkül (MCCARTHY 1969) und das Ereigniskalkül (Event calculus) (KOWALSKI 1986), in Ansätze, die auf nichtlinearer Zeit basieren, sowie Linearzeitmodelle einteilen. Die Linearzeitmodelle lassen sich wiederum in punktorientierte, quantitative Ansätze und intervall-orientierte, qualitative Ansätze untergliedern. Eine wesentliche Frage bei der Entwicklung des Moduls zum Umgang mit zeitlichem Wissen war zunächst, mit welchem Sprachkonzept(en) möglichst viele dieser Ansätze darstellbar sind.

2.1 Intervalle als Basisgröße

Um einen möglichst einheitlichen Rahmen für die Darstellung der Ansätze zum Umgang mit zeitlichem Wissen zu gewährleisten, wurde das „Intervall" als Bezugsgröße gewählt. Es zeigt sich nämlich, daß zeitliches Wissen in sehr mächtiger Weise durch die zeitlichen Bezüge zwischen Ereignissen definiert werden kann, wobei die Ereignisse durch Intervalle repräsentiert werden können. Desweiteren kann eine Intervalldarstellung bei dem Umgang mit unterschiedlichen zeitlichen Granularitäten sehr hilfreich sein. Ein weiteres wichtiges Argument für die Verwendung von Intervallen ist, daß sie bei zwei sehr mächtigen Formalismen direkt als Basisgröße verwendet werden, nämlich dem Intervallansatz nach Allen (ALLEN 1983, 1984) und dem Semi-Intervall-Ansatz nach Freksa (FREKSA 1992). Hinzukommt, daß Ansätze bei denen andere Darstellungsformen für das zeitliche Wissen genutzt werden, wie z. B. Zeitpunkte, in Intervallansätze integrierbar sind (VILAIN 1982). Weiter spricht für die Wahl des „Intervalls" als Bezugsgröße, daß auch nicht-intervallbasierte Ansätze auf die Darstellung von Intervallen zurückgreifen müssen. So muß ein wichtiger Vertreter der punktorientierten Theorien, das Axiomensystem von McDermott (MCDERMOTT 1992), das wiederum eine erweiterte Version des Situationenkalküls ist, bei einer Darstellung von Abläufen und Ereignissen, aus Zeitpunkten Zeitintervalle definieren.
Neben den oben genannten Argumenten für die Wahl des Intervalls als Bezugsgröße spricht aber auch die gute „kognitive Zugänglichkeit", die gerade für eine Entscheidungshilfe wesentlich ist. Daß Intervalle eine wesentliche Basis für den „menschlichen" Umgang mit zeitlicher Information spielen, wurde bereits von Galton (GALTON 1990) hervorgehoben. Ebenso legen Modelle für die zeitliche Informationsverarbeitung im Gehirn (PÖPPEL und SCHILL 1992) nahe, daß Intervalle wichtige Bezugsgrößen für den Menschen sind.

2.2 Die Intervall-logischen Ansätze

Erster Schritt bei der Entwicklung der Entscheidungshilfe zum Umgang mit zeitlichem Wissen war die Integration der Ansätze von Allen und Freksa, die beide qualitative Ansätze zum Umgang mit temporalen Wissen sind, d. h. Ansätze, mit denen in nichtnumerischer Weise zeitliche Beziehungen zwischen Ereignissen dargestellt werden können und damit zeitlich geschlossen werden kann. Die Ansätze sollten so integriert

werden, daß abhängig vom zur Verfügung stehenden Wissen des Benutzers, das Modul den geeigneten Ansatz zum Schließen auswählt. Im folgenden sollen die Ansätze kurz dargestellt werden.

Die Entwicklung von intervall-logischen Ansätzen geht im wesentlichen auf Allen (ALLEN 1982), (ALLEN 1984) zurück. Der Allen'sche Ansatz ermöglicht qualitatives zeitliches Wissen über Ereignisse zu repräsentieren und konsistent mit Hilfe transitiver Regeln zeitlich zu schließen. Der Ansatz von Allen setzt das Wissen über Anfänge *und* Enden von Ereignissen voraus, wobei Ereignisse durch zeitliche Intervalle repräsentiert werden. Für den von Allen entwickelten „Intervall-Relation-Reasoner" werden 13 zeitliche Relationen definiert, die bis auf die „equal"-Relation nicht symmetrisch sind und erlauben, alle möglichen zeitlichen Beziehungen zwischen Ereignissen darzustellen.

Über eine Transitivitätstafel werden die Axiome zum Schlußfolgern mit den 13 Relationen angegeben, wobei jeweils die transitive Hülle erzeugt werden kann, wenn ein neuer Fakt hinzukommt. D. h. sämtliche zeitliche Konsequenzen werden erzeugt, wenn ein neuer Fakt hinzukommt und es liegt so die gewünschte vollständige zeitliche Information bzgl. des aktuellen Wissens vor. Die Technik kann als „Constraint-Propagation" angesehen werden, wobei eine maximale Einschränkung des Wertebereichs aller Variablen in Abhängigkeit von den aktuellen Fakten stattfindet.

Wenn das zeitliche Wissen ungenau ist, d. h. nicht sicher ist, in welcher Relation zwei Ereignisse zueinander stehen, kann dies innerhalb des Intervall-logischen Ansatzes durch die Disjunktion sämtlicher möglicher Relationen dargestellt werden.

Wenn das Wissen „unvollständiger" ist, d. h. zeitliches Wissen nur bezüglich des Anfangs oder nur bezüglich des Endes eines Ereignisses bekannt ist, liegt es nahe, eine kondensierte Form des Allen'schen Ansatzes zu benutzen, die von Freksa (FREKSA 1992) auf der Basis von Semi-Intervallen entwickelt wurde. Dieser Ansatz beruht auf der Definition der konzeptuellen Nachbarschaft. Primitive sind dabei die Anfänge oder Enden von Ereignissen (Halbintervalle). Der Ansatz ermöglicht mittels der definierten Relationen, die zeitlichen Beziehungen zwischen Ereignissen, die nur durch ihren Anfang oder ihr Ende charakterisiert sind, zu repräsentieren und damit zu schließen.

Der Ausgangspunkt bei der Entwicklung des Semi-Intervall-logischen Ansatzes ist der Begriff der konzeptuellen Nachbarschaft, mit dem „grobkörniges zeitliches" Wissen erfaßt werden kann, indem „feinere" zeitliche Beziehungen, die in einer konzeptuellen Nachbarschaft stehen, miteinander verschmolzen werden. Zwei zeitliche Relationen sind dabei konzeptuell benachbart, wenn eine minimale zeitliche Veränderung der involvierten Intervalle einen direkten Wechsel zwischen diesen Relationen hervorrufen kann. So werden z. B. die Allen'schen Relationen „before" und „meets" zu einer semi-intervall-logischen Relation „precedes" verschmolzen. Um mit diesem groben unvollständigen Wissen umgehen zu können, ordnete Freksa die dreizehn Relationen von Allen nach dem Kriterium der konzeptuellen Nachbarschaft. Die wichtigsten Nachbarschaften, die als Inferenzen in Allens Transitivitätstafel auftreten, werden im Semi-Intervall-Ansatz mit eigenem Namen belegt.

2.3 Konsultationsabhängige Auswahl der beiden Ansätze

Die beiden oben dargestellten Ansätze ermöglichen somit jeweils unterschiedlich (un)vollständiges und (un)genaues zeitliches Wissen zu repräsentieren und damit zu schließen. Beide Ansätze ergänzen sich auf diese Weise und stellen zusammen ein wichtiges Werkzeug zur Repräsentation von zeitlichem Wissen und zum zeitlichen Schließen dar. Bei der Erstellung des Moduls zur Auswahl des zeitlichen Wissens wurde genutzt, daß es eine Kompatibilität der beiden Ansätze gibt, die sich in der gemeinsamen Transitivitätsmatrix wiederspiegelt. Die Realisierung des Moduls erfolgt so, daß das zur Verfügung stehende zeitliche Wissen sowohl durch Intervalle wie auch durch Semi-

Intervalle eingebracht werden kann. Die gegebenen zeitlichen Relationen zwischen zwei Intervallen und Semi-Intervallen werden in einer Adjazenzmatrix repräsentiert. Die Belegung der Komponenten der Vektoren dieser Matrix definiert dabei jeweils die zeitlichen Relationen zwischen zwei Intervallen. Die inversen Relationen werden jeweils automatisch in der Adjazenzmatrix aktualisiert. Für das zeitliche Schließen mit den Relationen in der Adjazenzmatrix, also zum Erstellen der transitiven Hülle, wurde der Warshall-Algorithmus an das Problem angepaßt und implementiert.
Folgende Auswahlkriterien bestimmen dabei die Auswahl des zeitlichen Ansatzes:
- Werden ausschließlich intervall-logische Relationen eingegeben, bleibt das System auf dieser feinen zeitlichen Granularität.
- Bei ausschließlich Semi-Intervall-logischen Relationen wird mit der groben zeitlichen Granularität, den nachbarschaftsbasierten Zeitrelationen, geschlußfolgert.
- Im „Mischfall", bei dem aktuell das Wissen sowohl durch Intervall- wie auch durch Semi-Intervall-logische Relationen dargestellt ist, werden die Allen'schen Relationen in Nachbarschaftsrelationen transformiert und es wird mit dem Semi-Intervall-logischen Ansatz geschlußfolgert.

2.4 Die Auswahl von Teilwelten

Unabhängig davon welcher der beiden Ansätze genutzt wird, kann bei einer großen Anzahl von Ereignissen die Betrachtung der gesamten Hülle unübersichtlich werden. Hinzu kommt, daß in vielen Anwendungsfällen nur ein bestimmter Teilbereich des insgesamt zur Verfügung stehenden zeitlichen Wissens von Bedeutung für den Benutzer ist und eine Teilweltbetrachtung wünschenswert wäre. So ist z. B. bei gewissen medizinischen Problemstellungen nur dasjenige Wissen relevant, das sich auf den Zeitraum *vor* einer bestimmten Therapieform bezieht.
Eine Möglichkeit Teilwelten zu strukturieren und auszuwählen, besteht in der Definition von Referenzhierarchien (ALLEN 1983). Dabei müssen mehrere Ereignisse in einem sie umfassenden Zeitabschnitt hierarchisch gegliedert werden können. Dies konnte quasi direkt durch den Allen Ansatz für das Modul realisiert werden.
Da das zeitliche Wissen jedoch häufig nicht in einer solchen eindeutigen zeitlichen Baumstrukturierung vorliegt, wurde für das Modul, um einen möglichst flexiblen Umgang bei der Auswahl von Teilwelten zu erlauben, die Definition von beliebigen Referenzintervallen ermöglicht, die zur Auswahl von Teilbereichen der gesamten transitiven Hülle dienen können. Auf dieser Definition aufbauend kann innerhalb der Teilwelten geschlußfolgert werden und nach Bedarf die veränderte Teilwelt auch wieder in das gesamte Wissen integriert werden.

2.5 Zeitpunkte als zeitliche Anker

Die bisher berücksichtigten Ansätze zum Umgang mit zeitlichem Wissen schließen zunächst den Zeitpunkt als zeitliche Darstellungsgröße aus. Allerdings kann sich diese Anforderung abhängig vom gegebenen Wissen als notwendig erweisen (z. B. zum Zeitpunkt der ersten Medikation). Dabei wird der Zeitpunkt als minimales Ereignis aufgefaßt. Dieses minimale Ereignis ist jedoch kontextabhängig, und kann wiederum in kleinere Ereignisse unterteilt werden. So können verschiedene medizinische Experten unterschiedliche Interpretationen bezüglich der zeitlichen Auflösung des Ereignisses „Medikation" haben. Um die Ausdrucksmöglichkeiten in dem zeitlichen Schlußfolgerungsmodul weiter zu erhöhen, wurde das Konzept des Zeitpunktes hinzugenommen. Entgegen der mathematischen Sichtweise, in der der Punkt ohne Ausdehnung

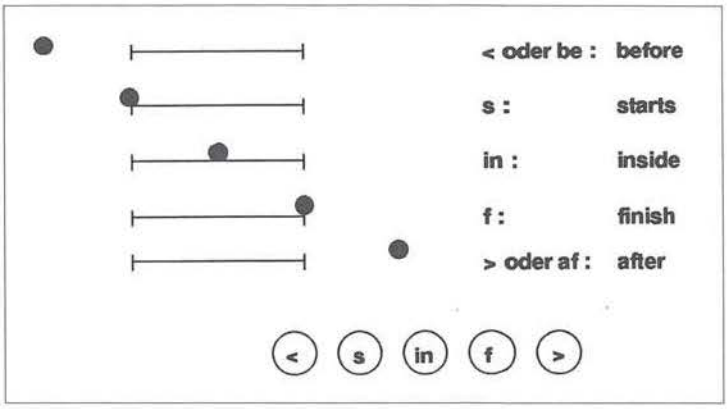


Abbildung 1: Die Beziehungen zwischen Punkt und Intervall mit ihrer logischen Anordnung

definiert wird, wurde von uns der Zeitpunkt als Intervall dargestellt, dessen Dauer kleiner ist als eine von der jeweiligen Betrachtungswelt abhängige Größe ε. Die Länge dieses Zeitpunktes wird nicht spezifiziert. Die Einführung des Zeitpunktes wird genutzt, um Teilwelten abhängig von Zeitpunkten zu definieren. Dabei stellte sich die Frage, wie Intervalle und Zeitpunkte zueinander in Relation stehen können und wie dies konsistent mit den verwendeten Logiken definiert werden kann. Es zeigte sich, daß für die Beschreibung aller möglichen zeitlichen Relationen zwischen Zeitpunkt und Intervall fünf Relationen des Intervallbasierten Ansatzes gelten können, nämlich: „before", „starts", „inside" (= „during"), „finish" und „after" (siehe Abbildung 1).
Auf diese Relationen wurde nun von uns der Ansatz der konzeptuellen Nachbarschaft angewendet, um damit „unvollständiges" Wissen auch im Bezug auf Zeitpunkte und Semi-Intervalle handhaben zu können. Dies führte zu der Definition der Relation „headsided", die durch die Verschmelzung der Disjunktion von „before", „starts" und „inside" entsteht und der Definition der Relation „tailsided", die durch die Disjunktion der Relationen, „inside", „finish" und „after" entsteht. Damit wurde bereits eine Basis für das Schlußfolgern mit Punkten und Intervallen sowie mit Punkten und Semi-Intervallen definiert.

3 Unsicheres Wissen

Um die Bedeutung von Unsicherheitstheorien miteinander vergleichen zu können, benötigt man ebenso einen Rahmen, innerhalb dessen ein solcher Vergleich überhaupt sinnvoll ist, d. h. man benötigt einen allgemeinen Formalismus für die Repräsentation von unsicherem Wissen und dem Schlußfolgern mit unsicherem Wissen.

Im folgenden werden zunächst die Gemeinsamkeiten der von uns untersuchten Unsicherheitstheorien (siehe Einleitung) beschrieben, dann ein Überblick über den von uns entwickelten Formalismus gegeben und schließlich wird gezeigt, wie die verschiedenen Unsicherheitstheorien in den Formalismus abgebildet werden können.
Das den Unsicherheitstheorien gemeinsame ist, daß sie eine Menge definieren und Elementen, bzw. Verknüpfungen von den Elementen dieser Menge einen Wert oder ein Wertepaar aus dem Intervall [0, 1] zuordnen. Ebenso ist eine Menge von Formeln gegeben, die die Menge der Element-Wert-Zuordnungen verändern. Die Menge der Element-Wert-Zuordnungen repräsentiert das unsichere Wissen und das Anwenden von Formeln wird als Schlußfolgern bezeichnet, d. h. also, daß das Schlußfolgern in einem gewissen Sinn eine Abbildung einer Wissensrepräsentation auf eine andere Wissensrepräsentation darstellt. Der Unterschied zwischen den Unsicherheitstheorien besteht

nun genau darin, in welcher Weise diese Element-Wert-Zuordnungen erfolgen und welcher Art die verwendeten Formeln sind. Sowohl die verschiedenen Repräsentationsformen, wie auch die verschiedenen Schlußfolgerungsarten sollen nun in den allgemeinen Formalismus abgebildet werden können.

3.1 Allgemeiner Formalismus für unsicheres Wissen

Die grundlegende Annahme des Formalismus ist: Ein Subjekt macht *Unterscheidungen* in der Umwelt. Eine Unterscheidung kann als Grenze gesehen werden, die die Umwelt in zwei Teile teilt. Etwas ist oder es ist nicht. Ähnlich einer Aussage. Sie ist entweder wahr oder falsch, ein drittes gibt es nicht. Aber eine Unterscheidung ist keine Aussage. Eine mögliche Unterscheidung muß nicht ausdrückbar sein. Daß eine Unterscheidung gemacht werden kann, impliziert nicht, daß sie benannt werden kann.
Aus elementaren Unterscheidungen lassen sich komplexe Unterscheidungen bilden. Eine komplexe Unterscheidung ist eine Grenze, die aus Grenzen der elementaren Unterscheidungen zusammengesetzt ist. Sie teilt die Umwelt somit wiederum in zwei Teile.
Wenn einem Subject n elementare Unterscheidungen möglich sind, dann wird seine Umwelt dadurch in 2^n Teile zerlegt. Jeder der durch die Unterscheidungen erzeugten Teile ist eine mögliche Beschreibung der Umwelt.
Manche Unterscheidungen können von dem Objekt benannt werden. Diese Unterscheidungen werden als *Konzepte* bezeichnet.
Eine weitere Annahme, die wir machen, ist, daß die Umwelt auf das Subjekt einwirkt und dadurch Teile, bzw. Beschreibungen der Umwelt unmöglich werden. Wir nennen dies *Beobachtung*.
Ein momentaner Zustand des Subjekts mit einer festen Anzahl von möglichen Unterscheidungen und Konzepten und einer Menge von unmöglichen Beschreibungen wird als *Sprache* bzw. *Wissen* bezeichnet. Wissen und Sprache sind hier also äquivalent.
Diese Sichtweise von Wissen erscheint auf den ersten Blick etwas ungewöhnlich. Dies resultiert daraus, daß üblicherweise unter Wissen nur das ausdrückbare Wissen verstanden wird. Hier umfaßt der Begriff Wissen beides – explizites und implizites Wissen.
Unsere These ist nun, daß alle oben genannten Unsicherheitstheorien jeweils eine Menge von Sprachen bzw. Wissen repräsentieren (nicht der Formalismus selbst, sondern die jeweilige Instantiierung). Die Unsicherheitstheorien beziehen sich also auf das Wissen eines Subjekts, d. h. es gibt unterschiedliche Arten, um das Wissen eines Subjekts zu repräsentieren. Aus dieser These folgt, daß eine Formalisierung dessen, was wir hier informell beschrieben haben, ermöglicht, die verschiedenen Repräsentationsformen, wie auch die verschiedenen Schlußfolgerungsarten in diesen Formalismus abzubilden.
Die Grundlage des hier vorgestellten Formalismus ist Carnaps *Logical Foundations of probability* (CARNAP 1950). Es kann im folgenden nur ein Überblick gegeben werden.

3.1.1 Die Sprachmenge S

Ein Element der Sprachmenge S wird mit $S_N^M/\mathscr{E}$ bezeichnet. (Erklärung der Symbole M, N und $\mathscr{E}$ weiter unten). In diesem Formalismus wird Wissen mit Sprache gleichgesetzt. Wenn von Wissen gesprochen wird, ist damit also eine Sprache gemeint.
Definition: Es gibt zwei Arten von **Zeichen in** $S_N^M/\mathscr{E}$:

1. *Konzeptkonstanten* (Θ). Es gibt M Konzeptkonstanten $c_1, c_2, \ldots, c_M$. Eine Menge von Konzeptkonstanten wird mit C oder C_i bezeichnet.
2. *Unterscheidungskonstanten* (Δ). Es gibt N Unterscheidungskonstanten $d_1, d_2, \ldots, d_N$, und $\Theta \cup \Delta = \Omega$.

Definition: Es gibt nur eine Form für einen **Ausdruck in** $S_N^M/\mathscr{E}$:
ein Element aus 2^Ω. Es wird *Beschreibung* genannt.
Definition: Es gibt eine **Abbildung in** $S_N^M/\mathscr{E}$:
Der Sprache $S_N^M/\mathscr{E}$ ist die Funktion $\mathscr{E}: 2^\Omega \to \{0, 1\}$ zugeordnet.
Definition: Basierend auf der Abbildung $\mathscr{E}$ lassen sich zwei **Klassen von Beschreibungen** definieren:

1. Die Menge $\{x \in 2^\Omega \mid \mathscr{E}(x) = 1\}$ wird Menge der *möglichen Beschreibungen* genannt (Λ). Eine Teilmenge aus Λ wird mit λ oder λ_i bezeichnet. Ein Element aus Λ wird mit x oder x_i bezeichnet.
2. Die Menge $\{x \in 2^\Omega \mid \mathscr{E}(x) = 0\}$ wird Menge der *unmöglichen Beschreibungen* genannt.

Definition: Wir vereinbaren **abkürzende Schreibweisen für Teilmengen von möglichen Beschreibungen**. Die allgemeine Form der Abkürzung : λ^K, wobei K *Bezeichner* genannt wird. Beispiele hierfür sind:

1. Statt $\{x \mid \forall c \in C: c \in x\}$: λ^C.
2. Statt $\{x \mid \exists c \in C: c \in x\}$: λ^C.
3. Statt $\Lambda - \lambda^K$: λ^K.

Definition: Basierend auf der Klasseneinteilung der Beschreibungen läßt sich ein **Grad des Glaubens in einen Bezeichner K bezüglich eines Bezeichners** K_w definieren:

$$G(K, K_w) = \frac{|\lambda^K \cap \lambda^{K_w}|}{|\lambda^{K_w}|}.$$

Definition: Basierend auf der Klasseneinteilung der Beschreibungen läßt sich eine Menge von **Zuordnungsfunktionen ZF** definieren:

$$ZF(x) = K \quad \text{mit} \quad x \in \lambda^K.$$

Während die Glaubensfunktion immer eindeutig ist, gibt es in der Regel zu einer Sprache eine Menge von Zuordnungsfunktionen.

3.1.2 Repräsentation von Sprachmengen

Die Funktionen Grad des Glaubens und Zuordnungsfunktion können dazu verwendet werden, um eine Menge von Sprachen festzulegen. Wird eine Konzeptmenge definiert und dazu eine Funktion basierend auf dem Grad des Glaubens oder einer Zuordnungsfunktion, dann gibt es eine Menge von Sprachen, die mit dieser Definition übereinstimmen. Wir können auf dieser Grundlage eine Menge von möglichen Wissensrepräsentationsformen erzeugen.
Der Unterschied zwischen den beiden Funktionen soll hier verdeutlich werden. Auf zwei gelben Tellern (ein kleiner und ein großer) und zwei blauen Tellern (ein kleiner und ein großer) sind 4 Pralinen verteilt. Wenn Ihnen jemand erzählt, daß sich auf den beiden gelben Tellern zusammen drei viertel der Pralinen befinden und auf den zwei kleinen Tellern zusammen ein viertel, und Sie Pralinen sehr gerne essen, dann werden Sie sicher den gelben, großen Teller auswählen. Wenn Ihnen allerdings jemand berichtet, daß er nun drei viertel Pralinen nehmen wird und sie auf die zwei gelben Teller und dann den Rest auf die kleinen Teller verteilen wird, dann scheint der kleine gelbe Teller die bessere Wahl zu sein. In beiden Situationen werden zur Beschreibung der Situation die selben Verhältnisangaben benutzt, jedoch mit verschiedener Bedeutung.
Diese Wissensrepräsentationen können nun mit jeweils einer Repräsentationsform der oben genannten Unsicherheitstheorien gleichgesetzt werden, d. h. die Repräsentationsformen der verschiedenen Unsicherheitstheorien können in den Formalismus abgebildet werden. Offen bleiben noch die Abbildungen der Schlußfolgerungsarten in den Formalismus. Die Schlußfolgerungsarten entsprechen nun Abbildungen zwischen den Sprachmengen. Diese werden wir im nächsten Abschnitt behandeln.

3.1.3 Abbildungen von Sprachen

Diese Abbildungen werden im Hinblick auf die verschiedenen Schlußfolgerungsarten der Unsicherheitstheorien definiert.

Definition: **Beobachtung**:

- $B: S_N^M/\mathscr{E}_i \rightarrow S_N^M/\mathscr{E}_j$, wobei nur der Übergang von $\mathscr{E}_i(x) = 1$ zu $\mathscr{E}_j(x) = 0$ möglich ist, nicht von $\mathscr{E}_i(x) = 0$ zu $\mathscr{E}_j(x) = 1$ (daraus folgt: $|\Lambda_{\mathscr{E}_j}| < |\Lambda_{\mathscr{E}_i}|$). B_{λ^K} sei die Beobachtung, die für alle x mit $x \notin \lambda^K$ $\mathscr{E}_j(x) = 0$ setzt.

Die Beobachtungsabbildung macht somit mögliche Beschreibungen zu unmöglichen Beschreibungen:

Definition: **Beobachtung** verallgemeinert auf Mengen von Sprachen:

1. $B_{\lambda^K}: \{S_N^M/\mathscr{E}_i\} \rightarrow \{S_N^M/\mathscr{E}_j\}$, wobei $\{S_N^M/\mathscr{E}_j\} = \{B_{\lambda^K}(S_N^M/\mathscr{E}_i)\}$.
2. $B_{\lambda^K}^w: \{S_N^M/\mathscr{E}_i\} \rightarrow \{S_N^M/\mathscr{E}_j\}$, wobei $\{S_N^M/\mathscr{E}_j\} = \{S_N^M/\mathscr{E}_i \mid \mathscr{E}_i(x) = 0$ für alle x mit $x \notin \lambda^K\}$.

Eine Verallgemeinerung der Beobachtungsabbildung auf Mengen von Sprachen kann auf zwei Arten geschehen. Einmal macht die Beobachtung in jeder einzelnen Sprache einzelne Beschreibungen unmöglich. Die zweite Art entfernt Sprachen aus der Sprachmenge, d. h. also „einzelne Wissen" werden unmöglich gemacht.

Definition: **Kombination**:

- $KOMB(\{S_N^M/\mathscr{E}_i\}, \{S_N^M/\mathscr{E}_j\}) = \{S_N^M/\mathscr{E}_i\} \cap \{S_N^M/\mathscr{E}_j\}$.

Bei der Kombinationsabbildung wird aus zwei Sprachmengen die Schnittmenge gebildet.
Definition: **Selektion**:

- $SEL(\{S_N^M/\mathscr{E}_i\}) = \{S_N^M/\mathscr{E}_j\}$, mit $\{S_N^M/\mathscr{E}_j\} \subseteq \{S_N^M/\mathscr{E}_i\}$.

Die Selektionsabbildung bildet eine Sprachmenge auf eine Teilmenge ab.
Definition: **Identität**:

- $ID(S_N^M/\mathscr{E}) = S_N^M/\mathscr{E}$.

Die Identitätsabbildung bildet eine Sprachmenge auf sich selbst ab.
Auf diese Abbildungen können alle Formeln der oben genannten Unsicherheitstheorien abgebildet werden. Im folgenden sollen einige Beispiele aufgezählt werden. Die Berechnung der bedingten Wahrscheinlichkeit aus der Wahrscheinlichkeit der Konjunktion und der Wahrscheinlichkeit der Bedingung erweist sich als eine Beobachtungsabbildung. Die „Dempster Rule of Combination" stellt eine Kombinationsabbildung mit einer anschließenden Selektion dar. Die „Dempster Rule of Conditioning" ist eine Beobachtung bezüglich Wissen, während das „Upper/Lower-Conditioning" eine Beobachtung bezüglich Beschreibungen darstellt. Die Und-Operatoren der Fuzzy-Set-Theorie erweisen sich als Selektionen.

3.2 Entscheidungshilfe zum unsicheren Wissen

Die Repräsentationsform und das Schlußfolgern der einzelnen Unsicherheitstheorien haben nun eine konkrete Bedeutung innerhalb des allgemeinen Formalismus.
Auf dieser Basis können wir nun die Unsicherheitstheorien bezüglich ihrer Repräsentationsformen und Schlußfolgerungsmöglichkeiten klassifizieren und damit eine Entscheidungshilfe für die Auswahl einer Unsicherheitstheorie entwickeln. Da vom Benutzer

nicht erwartet werden kann, daß er allein auf Grund der Definition dieses Formalismus in der Lage ist, Entscheidungen zu treffen, wie z. B., ob eine Repräsentation auf der Basis von Glaubensgraden oder Zuordnungsfunktionen angemessener ist, ist auch für das Modul zum unsicheren Wissen ein wesentliches Prinzip, die im Formalismus verwendeten Begriffe mit Hilfe von Beispielen zu veranschaulichen.

4 Diskussion

Es wurden zwei Module entwickelt, die eine Basis für eine Entscheidungshilfe zum Umgang mit zeitlichem und unsicherem Wissen bilden. Diese Module ermöglichen bereits jetzt die Integration unterschiedlichster Ansätze in einer einheitlichen Sprache. Dies erleichtert die Auseinandersetzung mit den existierenden Theorien und fördert den Einsatz von vorhersagbaren Methoden in medizinischen wissensbasierten Systemen. Darüberhinaus bietet die Darstellung durch eine einheitliche Sprache die Möglichkeit die Ansätze besser voneinander abzugrenzen und ihre jeweilige Mächtigkeit beurteilen zu können. Genau diese Transparenz stellt aber auch eine wichtige Voraussetzung für die theoretischen Weiterentwicklungen der Ansätze selbst dar.
Die zukünftige Entwicklung der Entscheidungshilfe wird durch zwei Richtungen bestimmt. Um diese Entscheidungshilfe umfassender zu machen, sollen einerseits weitere wichtige Ansätze zum Umgang mit zeitlichem und unsicherem Wissen integriert werden. Dazu muß u. a. untersucht werden, inwieweit die bisher definierten Sprachen und Bezugsgrößen ausreichend sind.
Andererseits sollen die Benutzer noch stärker durch eine anschauliche und interaktive Darstellung der Ansätze bei ihrer Entscheidung unterstützt werden. Dafür sollen domänenspezifische medizinische Beispiele, wie z. B. aus der Intensivmedizin dargestellt werden. Desweiteren sollen Simulationsmodule entwickelt werden, die ermöglichen, daß der Benutzer mit Ausschnitten seines Wissens die Ansätze erproben kann. Diese Simulation soll sich sowohl auf den statischen Aspekt der Repräsentation des unsicheren und zeitlichen Wissens beziehen, wie auch auf die Auswirkung der jeweiligen Schlußfolgerungstechniken.

Dieses Projekt wurde im Rahmen des BMFT-Förderschwerpunktes „Wissensbasen in der Medizin" (MEDWIS) entwickelt.

5 Literatur

ALLEN, J. F. (1983): Maintaining kowledge about temporal intervals, *Commun. ACM* **26** (11).
ALLEN, J. F. (1984): Towards a general theory of action and time, *Artificial Intelligence* **23**, 123–154.
CARNAP, R. (1950): Logical Foundations of Probability, (Routledge and Kegan Paul, Ltd., London).
CHARNIAK, E. (1983): The Bayesian Basis of Common Sense Medical Diagnosis, in: *Proceedings of AAAI-83*, (Washington, D.C:).
DUBOIS, D. und PRADE, H. (1980): Fuzzy Sets and Systems: Theory and Applications, (Academic Press, New York).
FREKSA, C. (1992): Temporal reasoning based on semi-intervals, *Artificial Intelligence* **54**, 199–227.
GALTON, A. (1990): A critical Examination of Allen's Theory of Action and Time, *Artificial Intelligence* **42**, 159–188.
HALPERN, J. Y. and FAGIN, R. (1992): Two views of belief: belief as generalized probability und belief as evidence, *Artificial Intelligence* **54**, 275–317.
KOWALSKI, R. A. und SERGOT, M. J. (1986): A logic based calculus of events, New generation computing, (Ohmsha ltd. und Springer), 67–95.

MCCARTHY, J. und HAYES, P. J. (1969). Some philosophical problems from the standpoint of Artificial Intelligence, in: *Machine Intelligence 4*, Meltzer und Mitchie (Eds.), (University press, Edingburgh).
MCDERMOTT, D. (1992): A Temporal logic for reasoning about processes and plans, *Cognitive Science* **6**, 101–155.
PEARL, J. (1986): Fusion, Propagation, and Structuring in Belief Networks, *Artificial Intelligence* **29**, 241–282.
PEARL, J. (1988): Probabilistic Reasoning in Intelligent Systems, (Morgan Kaufmann).
PÖPPEL, E. und SCHILL, K. (1992): Zeitliche Koordinationsprobleme mentaler Prozesse, *KI* **2**, 7–12.
SAFFIOTTI, A. (1987): An AI view of the treatment of uncertainty, *The Knowledge Engineering Review* **2** (2) (Cambridge University Press, UK).
SCHILL, K. (1990): Medizinische Expertensysteme, (Oldenbourg, München).
SHAFER, G. (1976): A Mathematical Theory of Evidence, (Princeton University Press, Princeton).
SMETS, P. (1988): Belief Functions, in: *Non-Standard Logics for Automated Reasoning*, Smets, Mamdani, Dubois und Prade (Eds.), (Academic Press, London).
SMETS, P. und KENNES, R. (1989): The transferable belief model: Comparison with Bayesian Models, *TR/IRIDIA/89-1* (IRIDIA, Université Libre de Bruxelles, Bruxelles).
VILIAN, M. B. (1982): A system for reasoning about time, in: *Proc. of the AAAI-82*, 197–201.
ZADEH, L. A. (1965): Fuzzy Sets, *Information and Control* **8**, 338–353.
ZADEH, L. A. (1979): A theory of approximate reasoning, in: *Machine Intelligence*, Vol. 9, Hayes, Michie und Mikulich (Eds.), (Halstead Press, New York), 149–194.

Anschrift der Verfasser: Kerstin Schill, Elisabeth Umkehrer, Ernst Pöppel, Institut für Medizinische Psychologie, Goethestr. 31, 80336 München

Informatik, Biometrie und Epidemiologie in Medizin und Biologie **26** (3), 198–214, ISSN 0943-5581
© Verlag Eugen Ulmer GmbH & Co., Stuttgart; Gustav Fischer Verlag KG, Stuttgart

Förderschwerpunkt Wissensbasen in der Medizin (MEDWIS)

Implementing a Knowledge Base for the Diagnosis of Secondary Failure by Neural Networks using a structured Development Method for Feedforward Neural Networks (SENN)

Erstellung einer Wissensbasis für die Diagnose des Sekundärversagens mittels Neuronaler Netze unter Verwendung einer struktuierten Entwicklungsmethode für Feedforwardnetze (SENN)

M. Förster[1]), T. Waschulzik[2]), K. Kirchner[3]), W. Moser[2]), V. Böhm[2]), W. Brauer[4]), R. Engelbrecht[2]), T. Koschinsky[3]), G. Entenmann[3]), DIADOQ*)

Kurztitel:
DIADOQ: ANN-Knowledge Base for Secondary Failure

Key Words:
neural networks, knowledge base, secondary failure, structured development method, quality assurance, data preparation, data representation, data encoding, SENN

Schlüsselworte:
Neuronale Netze, Wissensbasen, Sekudärversagen, strukturierte Entwicklungsmethode, Qualitätssicherung, Datenvorbereitung, Repräsentation, Kodierung, SENN

[1]) Institut für Medizinische Informatik und Biometrie Dresden
[2]) GSF-MEDIS-Institut Oberschleißheim
[3]) Diabetes Forschungsinstitut Düsseldorf
[4]) Technische Universität München, Institut für Informatik München
*) DIADOQ – Optimized Care Through Knowledge-based Quality Assurance: Diabetes Mellitus. The organisations involved are: Institute for Diabetes Research, Düsseldorf; GSF – MEDIS Insitute for Medical Informatics, Neuherberg; Institute for Diabetes, Karlsburg; University Hospital of Erlangen-Nürnberg; Nürnberg; IMIB – Institute for Medical Informatics and Biometry, Dresden; Department of Diabetes, München-Bogenhausen; Boehringer Mannheim GmbH, Mannheim; Institute for Mathematics, Ludwig Maximilians University, München. DIADOQ is part of the German MEDWIS (Medical Knowledge-bases) research programme

Abstract

The growing number of projects using Artificial Neural Networks (ANNs) in medical care makes it necessary to analyse the productivity and the ability of quality assurance of different ANN approaches.
After a short presentation of the DIADOQ project and an introduction of ANNs, this paper addresses the problem specification, data preparation as well as the development of appropriate representations, the selection of suitable encodings and the combination of encodings in neural network projects. In addition, strategies for testing and evaluating the network are described. It is demonstrated how networks with fast learning properties and simple network structures can be successfully applied for medical tasks.
During evaluation these ANNs are analysed interpreted for quality assurance purpose. Finally, a standard approach and the hereafter presented SENN method are compared from the point of efficiency and the quality of the results.

Zusammenfassung

Die wachsende Anzahl von Projekten, die die Technologie der Künstlichen Neuronalen Netze (ANN) im Gesundheitswesen anwenden, erfordert eine Analyse der Produktivität und der Möglichkeiten zur Qualitätssicherung unterschiedlicher neuronaler Ansätze.
Nach einer kurzen Vorstellung des DIADOQ-Projektes und einer Einführung in die ANN behandelt dieser Beitrag die Problemspezifikation, die Datenvorverarbeitung die Entwicklung von geeigneten Repräsentationen der Daten, die Auswahl geeigneter Kodierungen und deren Kombination in Projekten mit ANN. Desweiteren werden Strategien zum Test und zu Evaluation des Netzes beschrieben. Es wird demonstriert, wie schnell lernende ANN mit einer einfachen Netzwerkstruktur erfolgreich für medizinische Problemstellungen Verwendung finden können.
In der Evaluationsphase können die ANN zur Qualitätssicherung analysiert und reinterpretiert werden.
Abschließend wird eine Standardmethode zur Entwicklung neuronaler Netze mit der im folgenden näher beschriebenen SENN-Methodik in Bezug auf die Effizienz der Erstellung und die Qualität der Ergebnisse verglichen.

1. Introduction

The medical knowledge for an optimized diabetes care is largely available and e.g. published in consensus guidelines of the international medical societies [NIDDM, IDDM 1993]. Nevertheless, problems in the management of diabetes patients occur mainly due to difficulties in the application of the available knowledge at the appropriate time during the process of care. In the St. Vincent Declaration [KRANS et al. 1992], an initiative of the International Diabetes Federation/Europe and the World Health Organization/ Europe, targets like the reduction of diabetes-related amputations for at least 50% should be reached within the following 5 years to improve the quality of life of Diabetes patients. Quality management is one of the key measures to implement these targets.
The members of the consortium regard DIADOQ as one part in a network of different activities for quality management of Diabetes. DIADOQ aims at assuring the quality of the process of health care delivery by means of knowledge bases for clinically relevant problems within diagnosing, therapy planning, and monitoring.
On the one hand, outcome indicators like blood glucose and HbA1c are already monitored e.g. in the DiabCare initiative [PIWERNETZ et al. 1993] and compared to those

of other institutions or accepted standards. From such feedback, improvement of the quality of treatment is expected. On the other hand, knowledge bases mainly for complex medical problems like insulin dosage adjustment [HEIJLESEN et al. 1993] have been developped but not integrated into the computer-based medical record (CPR). Complementary to these approaches DIADOQ aims at directly supporting the process of health care delivery through the integration of knowledge bases into a CPR system and therefore into clinical routine. ANNs represent one approach to implement knowledge bases particularly well suited for classification (such as diagnoses and prognoses) and approximation problems.
The DIADOQ CPR system adopts a problem-oriented medical record as a natural model of the CPR for managing a chronic disease. Patient problems organised in a problem list provide a concise medical summary of the patient and serve as justifications of performed medical acts. They provide the necessary contextual background to determine if a medical act has or has not to be performed. The recording of information within the CPR system concerning the medical act leading to each medical item lays the foundation for documenting the process of health care provision. This representation is the prerequisite for the implementation of guidelines and the identification of those steps within the process of care where knowledge-based support beneficially could be applied.

2. Neural Networks as knowledge bases

An ANN is a highly adaptive system that is able to learn based on examples. In contrast to other methods of knowledge representation, the designer of ANNs needs only few explicit model knowledge which is in general easy to obtain.
An ANN consists of a set of highly interconnected neurons. General knowledge is stored in a distributed manner in the weights of the connections of the ANN whereas the current neuron activities represent a specific case.
Multilayer feedforward networks are universal approximators. Thus, they can extract arbitrary functional dependencies between input and output variables from a statistically representative data set provided such dependencies exist. This allows a functional use of case-based knowledge.
The result of the learning process ("training") corresponds to the solution of a non-linear optimisation problem. The training aims at adjusting all free model parameters of the ANN (its weights) in a way such that the mean square difference between the desired and the current output of the ANNs is minimised for the examples of the training set. The more complex the mapping is and the more weights have to be adjusted the higher is the probability to obtain only a sub-optimal solution. Another important drawback of larger ANNs is that they require much more for training than smaller ANNs.
As shown in Figure 1, problem relevant input variables are mapped onto the outputs of the neurons of the input layer whereas output variables are mapped onto the output of those of the output layer. The type of mapping, the structuring of additional layers and connections as well as the selection of suitable learning algorithms and their parameter adjustment have to be done by the designer of the ANN. Although many heuristics for this are known, these tasks are commonly performed in a more or less empirical way since no general solution exist.
An increasing number of papers applying neural networks to medical problems [ANDREASSEN 1993, IFEACHOR 1994] have been published for the last years. In medical applications, we have often to deal with the following circumstances:
- sufficient formalised knowledge is unavailable
- onyl a small number of examples is available

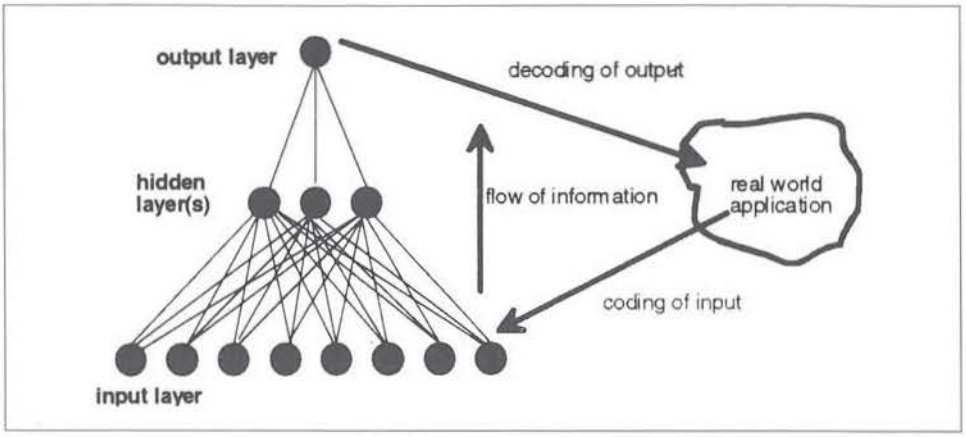


Figure 1: A multilayer feedforward neural network

- a lot of possible casual relations must be taken into consideration
- large normal parameter variations
- decisions have an enormous impact on the patient

On the other hand, neural networks should be used for medical tasks only if:

- a sufficiently large number of examples of high quality is availabe (evaluated examples without data collection artefacts, using standardised collection)
- sufficient formalised knowledge for a symbolic approach is unavailable and the problem has to deal with fuzzy[1]) circumstances
- only a number of variables than can be managed influence the problem solving process
- an expert of the domain accompanies the project.

Especially the mismatch between the available and required number of examples (which strongly depends on the number of variables influencing the output that has to be determined) often endangers the quality of the project. If the number of available examples is to small for an appropriate statistical analysis, neural networks should not be applied either. Ignoring this may lead to severe quality assurance problems. In order to improve the efficiency and in order to assure the quality of the projects a method for a structured development of neural networks is necessary. This method has been developed for several years [WASCHULZIK 1988, WASCHULZIK 1990 a] based on the experience of industrial [WASCHULZIK 1990B, WASCHULZIK 1991] and medical projects [QUANDT 1993, WASCHULZIK 1993, DROSTE 1994, DÄSCHLEIN 1994].

3. Medical problem

Type II diabetes is characterised by an abnormal insulin secretion and/or a decreased insulin effect on insulin sensitive tissues (insulin resistance) [YKI-JÄRVINEN 1994]. During the first stage of Type II Diabetes diet alone or in combination with oral antidiabetic drugs can compensate for diabetes related metabolic changes.

Many patients treated with oral hypoglycemic agents may develop a loss of therapeutic efficacy which finally results in an additional need for long-term insulin treatment to

[1]) Here the word fuzzy is used in the orginal way and not in the context of fuzzy-set-theory

achieve sufficient metabolic control. This stage of type II diabetes is referred to as secondary failure (SF) of the oral diabetic drug treatment. The pathophysiological factors leading to secondary failure are often unknown. Only a minority of these patients shows a decrease in insulin secretion. Such a persistent SF has to be distinguished from temporary metabolic derangements due to e.g. transient infections or the use of diabetogenic drugs for a limited period of time. Obviously there is an urgent need for a reliable and real-time differential diagnoses of SF, to select and start an appropriate treatment to assure the best quality of life with the least degree of diabetes related complications [EUROPEAN NIDDM POLICY GROUP 1990].
The problem "Secondary failure in type II diabetes" was approached in two different ways in order to compare the proposed method with a standard method. Data of 573 diabetic patients (Table 1) were available for the differential diagnosis of SF which can be considered a classification problem. Additonal data are being collected for purposes of validation.

Table 1: Number of cases in the data sets

Data Set	SF	No SF	Total
Training Set	219	113	332
Test Set	114	53	167
Assessment Set	55	19	74
Total	388	185	573

4. Structured development method for feedforward neural networks-SENN

The SENN method (Strukturierte Entwicklungsmethodik für vorwärtsgerichtete Neuronale Netze) is based on a system life cycle model and refers to techniques commonly used in literature. The new aspect is the integration and combination of these techniques for quality assurance and efficieny purposes and their support by appropriate tools. This paper describes how a knowledge base for the diagnose of SF is realised by neural components. The SENN method is described in more detail in [WASCHULZIK 1995].

4.1. System analysis

The set of input and output variables and a set of possible dependencies for each of the output variables from the input variables is one of the results of the system analysis in neural network projects. These dependencies as well as additional information on the appropriate representation and encoding of these variables are documented in a table of dependencies which is the basis for further system development. We distinguish between the representation of the variables (for example, whether a disease is classified in ICD9 or SNOMED) and the encoding of this representation for the neural network.
In some projects like the set-up of a system for the support of the differential diagnosis of the SF, the table of dependencies can degenerate since the only output variable depends on a list of input variables. Only unstructured variables are used in this approach for the estimation of SF. Thus, no special techniques for the representation and encoding of the variable had to be used. This is favourable for the conventional approach since SENN would have made good use of a more complicated internal structure of the variable to enhance the information processing of the network. However, the lack of structured variables in this application supports the comparison of the methods.
In the system analysis, it is also determined how many examples will be necessary and how they are available.

4.2. System design

During the design phase of the system it has to be ensured that all variables needed in the table of dependencies are available and that a sufficiently large number of examples can be collected in exactly the same way.
To assure the quality of the resulting system, we have to take care that the results produced by our prototype are of the same quality as in the future real-world application. Accordingly, we have to split up our data set into three parts:

- training set (to extract the knowledge according to the selected model)
- test set (to adjust the parameters and to compare the different combinations of the variables)
- assessment set (to determine which results we will obtain when we apply our system in the field)

The asseessment set is necessary since the system may not only overadapt to the training set but also to the test set during system development. Furthermore, while testing different parameter settings for the simulation and models many decisions are made based on the results of these systems on our test set. Thus, we can not expect to obtain the same results in the practical application as on our test set. We may only use the assessment set once in a project to determine the final result for the expected behaviour of the system. After this evaluation, the assessment data set is lost for further assessments. For the final test, the training set and test set can both be used as training data since the assessment set now serves as test set. Thus, all available examples are used efficiently.

Table 2: Features used by the neural net (C = categorical value, B = binary, O = ordinal, N = number)

Number	Variable	Type
1	sex	CB
2	age	N
3	duration of diabetes	N
4	duration of treatment with sulfonylureas	N
5	duration of poor metabolic control	N
6	body-mass-index (BMI)	N
7	weight change within the last three month	CO
8	fasting blood glucose	N
9	postprandial blood glucose	N
10	HbA1	N
11	acetone	CO
12	leukocytes	N
13	cholesterol	N
14	trigylcerides	N
15	sedimentation rate of blood (BSR)	N
16	diabetogenic drugs	CB
17	infections	C
18	heart disease	C
19	thyroid disease	C
20	urinary tract infection	C

One of the major advantages of using neural networks is that usually only few lines of code have to be written for a new application. In software projects, one general principle of quality assurance is the application of standard components wherever possible. Often only interface routines and routines for transformations of the raw data have to be written for the neural components. This is one reason for the contentment of users in neural systems.

4.3. Realisation of the system

The realisation of the system starts with the collection of the data sets and the selection of the examples. In our applications, the information about the SF and the input variables in Table 2 had to be registered. Afterwards the appropriate representation has to be chosen followed by the encoding of the variables and the configuration of the network. The neural network can be integrated into the system after the training and testing phase.

4.3.1. Representation of the variables

Sometimes it is advantageous to pre-process raw data before applying to the network. For example, in object recognition, it is not useful to present the raw pixel information to the neural network since cases with similar outputs might have very different representations at the input. It is necessary to enhance the similarities between the input and output of the neural network. This can be done by ignoring useless information because, generally, neural networks are continuous approximators that map similarities from the input to the output with some non-linear distortions. Thus, we have to switch from grey values to contrast pictures. As another step to enhance similarities, we have to switch to histograms of edges and combinations of edges. This simple representation is of good use for translation- and scale-invariant object recognition [POLZER 1992].
The quality of the representation can be scrutinised by using and ARE (Associative Recall of Examples). This is achieved by calculating the distance between the representation of the data set and the current example. The representation is suitable if similar patterns yield small distances in between, otherwise the representation has to be improved. The efforts for the development of representations and the development of networks must be well-balanced. The importance of the representation is also stressed in the literature: "A good representation does most of the work. Neural networks no matter what kind are weak computers, but a proper representation, coupled with the abstracting and co-operative effects of the network, can become extremely fast, powerful, and accurate." A. Rosenfeld in [ANDERSON 1988] according to [NELSON 1990, page 155]. In further investigations we will check whether the results of the ARE using the representations and later the encodings can be used as early quality indicators in the quality assurance process.

4.3.2. Encoding of the variables

The encoding has to present the representations to the network in an appropriate way. It can also be used to combine variables in different ways. In the following a small subset of possible encodings and combination techniques for variables is described.

Categorical variables:
An euclidian metrics over values of categorical variable is not defined not even if an order for possible values exists. This fact forbids to map values from categories to numbers and to use these numbers for the activation of neurons in a neural network since otherwise the next step of information processing would implicity assume an euclidian or similar metrics. Although this procedure of encoding different values of one category in one neuron is frequently applied we have to be aware that this mapping implicitly introduces an euclidian metrics over the possible values which is used by the neutral network. Doing this is like adding a wrong formula to a set of formulas that are used for the solution of a problem. Consequently, the "one of n"-encoding should be preferred. It uses one neuron for each possible value of the category. If a value has to be encoded only the according neuron is activated.
In case an order is defined over the possible values, the neighbourhood in the order can be presented to the network by activating those neurons (with lower activation) which encode the categories neighboured to the active category. In this case, a special connection exists for modelling the underlying dependencies for every possible value.
An example of an encoding for hierarchical categorical variables (e.g. variables encoded with ICD9 or SNOMED) can be found in [WASCHULZIK 1993].

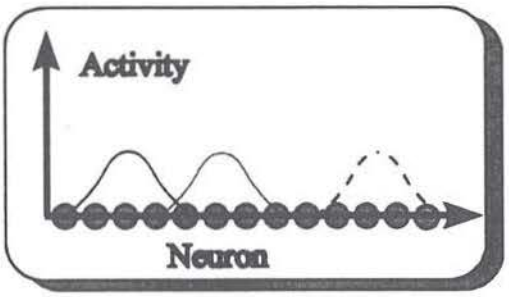


Figure 2: The topological encoding

Continuous variables:
Over continuous variables, an euclidian metrics can be assumed in most cases. Hence, we can initialise the activity of the encoding neuron with an activity calculated from the value of the variable. Based on very good experiences [QUANDT 1993, DROSTE, 1994, WASCHULZIK 1993] we prefer the topological encoding [GEIGER 1990].
This distributed encoding scheme uses up to n neurons to represent one contiuous variable and allows to model even complex functions with a simple perceptron [ELDRACHER 1993] and thus avoids the need of a hidden layer in the neural network.

4.3.3. Combination of variables by encodings

In most cases a set of input variables is used to determine the output variable(s). According to the desired properties of the combination, there are different possibilities for the combination of the variables.

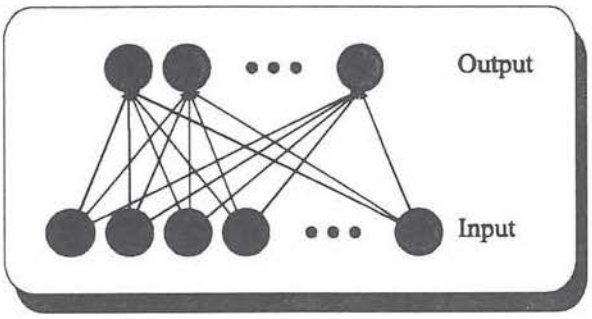
Parallel encoding:
Parallel encoding is the most common type of variable combination. The encoding of each variable is determined by calculating the activity of each neuron that represents the value of the variable. This encoding is referred to as the basic encoding of the variable.
Basic encodings of the variables are simply put side by side into the input layer of the network without carrying out specific operations. The total number of neurons for the representation of all input variables equals hence the sum of the numbers of the neurons used for the basic encodings.
This simple approach is advantageous when many variables of different types have to be combined.

Conjunctive encoding:
The conjunctive encoding is built upon the basic encodings of the variables that are to be combined. One neuron is used for every possible combination of the neurons from the different basic encodings. The activity of the resulting neuron is calculated by multiplying the activities of the neurons from the basic variable encodings that are combined. The conjunctive encoding allows in principal to model every possible mapping from the input to the output space as it was shown for RBF-Networks in [PARK 1991, BENAIM 1994][2]. The disadvantage of conjunctive encoding is the large number of needed neurons as discussed in [HINTON 1986, p 90].
One result of our investigation is that conjunctive encoding was less important in our pratical applications than we had expected. In most of our applications, the conjunctive encoding did not yield better results than the parallel encoding with appropriately represented variables. These results lead to the impression that the information about the conjuntion of variables in most applications is less important.

[2]) The topological encoding combined with the conjunctive encoding of all variables corresponds exactly to a RBF network if the input variables are equally distributed. The exact distribution of the RBF-neurons is not necessary for the proofs.

4.3.4. Network structure

The first applications we rendered by neural networks were completely neural solutions involving complex structured neural networks for pattern recognition tasks [WASCHULZIK 1988, WASCHULZIK 1990b]. The complexity of the networks was reduced stepwise to improve efficiency and quality assurance. For the same reason we directed our attention towards the development and improvement of the representations and encodings.
For the majority of the applications we prefer the simple perceptron with one input and output layer where each input neuron is connected with each output neuron ROSENBLATT 1958].



The directed connections are optimised by a gradient descent algorithm, the so called δ-rule [WIDROW 1960]. The advantages of the perceptron are its good generalisation properties, the fast convergence of the training algorithm and the possibility to interpret the network structure.

Figure 3: The simple perceptron

4.4. Evaluation

The evaluation consists of verification, validation and assessment [ENGELBRECHT 1994].

4.4.1. Verification

The verification is made by analysing the structure of the trained network which is important especially then the results on the assessment data set are better than expected. The competence of a network in a special context can also be evaluated by an ARE-retrieving similar cases to one special example (which is representing the context) from the training set using the same representations and encodings that are used for the initialisation for the input and output neurons. If the ARE finds many cases that are similar to such an example from the applicational view then the competence of the network is good in the actual context. Furthermore, the training data set can be split up into two equally sized parts and the resulting network structurs can be compared after training.

4.4.2. Validation

Beside the tests with the test set a data set with artificial examples, e.g., prototypical cases, can be set up and the results can be compared with textbook knowledge and discussed with an expert of the domain. It is also interesting to analyse the examples which yield bad results.
During this validation phase it is necessary that the medical expert "plays" with the prototype to decide whether the system meets his expectations or not. This could also be considered a "subjective" validation.

4.4.3. Assessment

The assessment is made on the whole by the tests with the assessment data set. Also statistical tests are recommended to determine whether the results yield by the neural network differ significantly from random choice.

Finally, the ANN must be tested with independent data from other sources. Unfortunately, this is expensive and difficult to achieve as long as no standard electronic medical record is used in practice. However, it should be stressed that without this it is not allowed to transfer the knowledge base to other places.

5. A conventional approach

In our first approach, we used conventional tools and mainly conventional techniques despite some influences from SENN introduced by former discussions about neural network development methods.
NeuralWare's[3] tool DataSculptor® V1.53 was used for data pre-processing. DataSculptor® imports different database formats and offers many different ways of transforming raw data.
NeuralWorks Professional® V5.0 was used for creating the networks and performing the simulations. The creation of each network has to be done manually by using dialogues in a graphical environment where architectures and learning paradigms can be chosen from a wide range. Each network has bo be connected with two files: one containing the training set, the other containing the test set. Both files are created by DataSculptor. All network and learning parameters including the number of input and output variables must be set by the user since they cannot be derived from the data files. During learning and testing, different protocols are written to disk (for later analysis) and to the screen for instant information.

5.1. Description of the simulations

Binary features were coded by one neuron; ordinal featurs were coded by 0 to n neurons using a thermometer code and all other categorical features were coded as 1-of-n. Different coding schemes were used for numeric features. Missing values were estimated in rare cases; the a-priori-probability was used for categorical and the mean for numeric features. In order to obtain a minimal net with best generalisation characteristics the following approaches were undertaken (see also [FÖRSTER 1994]):

A1) permanent decrease in the total number of free parameters of a multi-layer perceptron with one hidden layer trained by back-propagation [WERBOS 1974, RUMELHART 1986] by training over-dimensioned (large) nets and afterwards eliminating weights of minor importance (weight elimination) and/or pruning of neurons in the hidden layers that had only minor influence on the overall performance of the net

A2) permanent increase in the total number of free parameters of the net by adding new neurons until no further improvement of the classification quality can be reached (Cascade Correlation [FAHLMAN 1991]

B) numeric features were coded in three different ways:
   - B1) caled input, unchanged distribution
   - B2) scaled input after suitable transformation of the distribution (e.g. log()) in order to achieve a more homogeneous distribution
   - B3) like B2, in addition, each numeric input is coded as a fuzzy variable where every neuron is activated according to a membership function (5 neurons per feature, e.g., blood glucose: very low, low, medium, high, very high).

[3]) We used this tools because NeuralWare is according to our information world market leader in the area of tools for the development of neural networks and DataSculptur is the only well-known commercial tool for the preparation of data for neural networks.

This fuzzy coding was used in order to preserve similarities of the transformation to a maximal extent which would have been partly lost by applying the straight-forward splitting into categorical medical categories.

5.2. Results of the simulations

The criteria for stopping the learning procedure was the (at a certain point decreasing) classification rate on the test set instead of the mean square error on the training set. Therefore, a cyclic testing based on the test set of the net during learning was performed and the currently best net was saved.
The results of the simulation can be summarised as follows:

A1) A bad convergence during learning produced relatively poor initial configurations. Neither weight elimination nor pruning achieved better generalisation using these starting configurations

A2) Nets of different complexity evolved that did not achieve better generalisation than nets of smaller complexity.

After these disappointing results, different coding schemes for numeric input features were applied (B1 – B3) to networks with one or no hidden layer yielding much better results. During training, every class is presented with the same frequency, i.e., the information of the a-priori probability is discarded. In order to verify the results, a logistic regression was performed using the program BMDP®. Afterwards, another net was created using only features selected by the logistic regression.
Keeping in mind that a subset of features produces similar simulation results, three other different subsets of features were tested. The feature selection was based on sensitivity analysis of trained nets as well as on some early results obtained by HEAD. A fuzzy coding similar to B3 was used. It can be stated that networks using only a subset of the most important variables yield similar results as the networks using all input variables.
A more detailed description of the results and the approach can be found in [FÖRSTER 1994] and [WASCHULZIK 1995].

6. Approach using SENN

The same diagnosis of the SF was solved by using the SENN-approach supported by HEAD (Hybrid Expert system for the Analysis of large Data sets) [NÖHMEIER 1994]. HEAD is an integrated tool for the data preparation and the development of neural networks using SENN. It is based on routines developed for the analysis of an epidemiological study using neural networks [QUANDT 1993] and on an optimised simulator for perceptrons [GROTHE 1993]. The simulations were performed independently from the conventional approach.

6.1. Data preparation, creation and simulation of the neural networks

It was agreed upon that the available data set should be completely used for the simulations and that the assement should be performed with examples collected later and with examples from other medical institutions. The available data set was split up automatically by HEAD into a training and test set. Within HEAD, a network is designed by selecting the variables and their encodings for the simulations. The selection of the encodings is supported by proposals suggested by HEAD after the analysis of the available values for each variable. After the selections, the network is automatically

generated, trained, tested and the results are analysed. The results are presented on-line and are also available in a file that can be easily analysed by shell-scripts and programs. This approach dramatically reduces the amount of time required for the development of a network and, in addition, it avoids errors.
Categorial variables were encoded in the same way as in the conventional approach (1-of-N). Applying topological encoding, numeric variables were coded by different numbers of neurons.

6.2. Results of the simulations

After preparing the data set for the application of HEAD we made simulations in order to optimise the number of neurons for the representation of the numeric variables. In order to improve the mean value of sensitively and specificy both classes were trained with the same frequency. Then we tested the first combinations containing all input variables and obtained the results of S1. As a next step we combined only a subset of the variables. The best result (sensitivity + specificy)/2 = 72,4%) was obtained in simulation S2. For the next simulation, we changed to the routine level of HEAD and computed all combinations with three input-variables using both parallel and conjunctive encoding. Simulation S3 was the best combination and considered only three input variables. The simulations using conjunctive encoding achieved worse results than those using parallel encoding.
Examining the results of S3 the effect of overadaption on the test set can be recognised (the results on the test set are better than on the training set). The reason for this is the small number of free parameters which inhibits the overadaption effect on the training data set. The selection of the best results from a large number of simulations encourages the overadaption on the test set which is stronger than the overadaption effect on the training set.
We also tested some combinations of up to 6 variables where the simulation S4 achieved the best results.
In order to achieve better classification results we altered the training process in such a way that the classes were trained taking into account their a-priori probabilities. Now simulations using conjunctive encoding achieved better results. For the resulting networks, a big difference between the sensitivity and the specificy can be stated on the test set.

Table 3: Simulation results on the test set (approach with SENN)

Simulation			Test Set			Training Set
	Input variables (numbers from Table 2)	number of free parameters	classi-fication rate	sen-sitivity	specificy	classi-fication rate
S1	all input variables	418	71%	85%	43%	90%
S2	4,6,8,10,12	60	71%	69%	75%	73%
S3	4,6,12	40	74%	78%	66%	71%
S4	1,4,6,8,9,12	68	73%	84%	50%	77%
S5	2×4×6	128	74%	94%	32%	71%
S6	2×4×6, 16	132	76%	92%	43%	72%
S7	2×4×6, 19, 11	150	71%	71%	71%	73%

Table 4: Simulation results on the assessment set (approach with SENN)

	Assessment Set (network trained with training set)			Assessment Set (network trained with training and test set)			Trainig Set & Test Set (network trained with training and test set)		
	classi-fication rate	sensi-tivity	speci-ficy	classi-fication rate	sensi-tivity	speci-ficy	classi-fication rate	sensi-tivity	speci-ficy
S1	72%	89%	26%	74%	96%	10%	83%	93%	63%
S2	71%	72%	68%	79%	90%	47%	78%	87%	60%
S3	64%	67%	57%	64%	69%	52%	71%	78%	58%
S4	71%	80%	47%	81%	83%	73%	71%	70%	72%
S5	74%	98%	5%	78%	100%	15%	73%	93%	32%
S6	72%	92%	15%	60%	80%	5%	71%	82%	48%
S7	74%	94%	15%	75%	100%	5%	72%	93%	30%

6.3. Evaluation

Finally we ran simulations in order to evaluate the results on the assessment set. The tests with the assessment data set show that S1, S5, S6 and S7 reach a very good sensitivity but such a bad specificy that they will be hardly of any use in the practical application.

The overadaption effect in simulation S3 was confirmed. The results on the assessment set of S2 and S4 are astonishing good. We can not expect to obtain these results in a practical application. It has to be critically analysed whether the assessment set is statistically representative. In addition, further assessment cases have to be collected. However, much evaluation work, e.g. the analyses of the wrong classified examples, has to be carried out.

Interpretation of the weights

ANNs are often criticised because they are said to be "black boxes". However, this holds true only for some widespread types of ANNs such as the multi layer feedforward network commonly referred to as the "backpropagation network". One important aspect of using ANNs in medicine is therefore to use those types of ANNs whose structure can be easily analysed for verification purposes.

The decision to have one output neuron for each category (one for SF-yes and one for SF-no) is primarily based on considerations of interpretability of the trained network since for a bi-

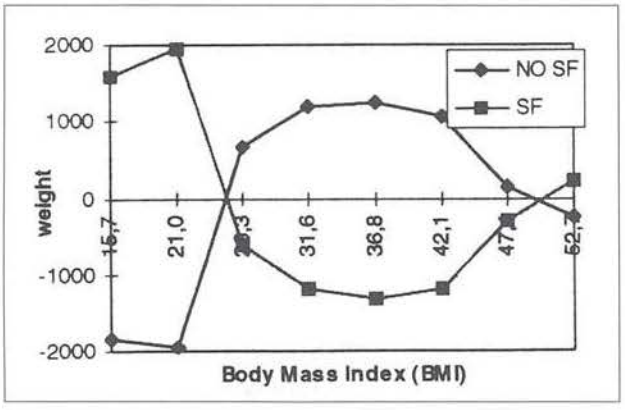


Figure 4: Resulting weights for the neurons topologically coding the body-mass index (BMI)

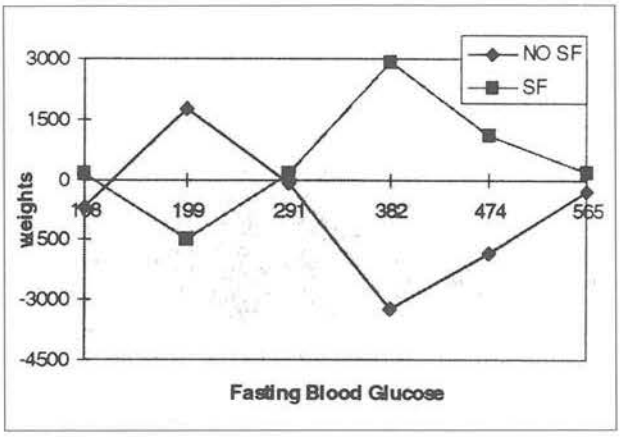


Figure 5: Resulting weights for the neurons topologically coding fasting blood glucose (FBG)

nary decision one neuron would be sufficient.
Figure 4 and Figure 5 show two examples of the interpretation of the weights the ANN has learned. The solid line indicates the influence of the neurons on the diagnosis "no secondary failure", the dotted line the influence on "secondary failure". Note that the leftmost neuron is coding a variable's mimimum value and the rightmost neuron its maximum value. Mainly the difference between the two lines is of interest since this shows the ability of the neurons to differentiate. Interpreting Figure 4, we can state that a very low BMI is a strong vote for SF, medium and quite high BMI are in favour of no SF, whereas very high BMI does not contribute very much to the decision. Note that the predictive value of the leftmost and rightmost neurons are sometimes of limited value since examples from these extreme categories are often under-represented provided the coded variable has a kind of normal distribution.
Note that the topological coding is also advantageous for a re-interpretation since a variable is represented by a set of neurons.
As a result, discussions with medical experts are facilitated and data artefacts or overtrained ANNs can be found.

How good is a classification rate of 70%, 75% or 80%?
In order to objectively compare a knowledge base with human experts decisions must be taken at the same point in time and based on the same information. In other words, a knowledge base should compete with human experts who are in exactly the same situation.
Hence, it is necessary that some external medical experts answer the same questions as the ANN, i.e. some 100 cases are randomly chosen and given to experts who have to answer them (yes, no, can not decide). We expect quite different results from different experts and generally classification rates below 100%. An evaluation study [ENTENMANN 1995] revealed that in an ambulance 34.1% out of 167 patients had a wrong positive diagnosis.

7 Comparison of the SENN and the conventional approach

The conventional and the SENN-approach led to similar results after additional information on a proper encoding of the variables and on good combinations of data fields both obtained by SENN was used in the conventional approach. Until now, the main difference is to be found in the total time spent in each approach (Table 5).
Mistakes during the analysis in the conventional approach due to the complexity and number of different protocol files created during simulations and tests are another important disadvantage of the conventional approach. Because of the lacking integration

Table 5: Total time spent for the development

activity	conventional approach	SENN approach
total spent time	3 month	2 weeks[4])
pre-processing	35%	40%
creation of network	15%	
simulations	5%	
analysis and selection of networks	35%	40%
evaluation	10%	20%

of tools in the conventional approach, this approach could not cover as many trials as SENN did.

8 Summary

Efficiency and quality assurance in medical projects can be enhanced by using the SENN method and neural networks without hidden layers. The higher efficiency mainly depends on the approach of developing proper representation and encodings for the combination of variables instead of testing different network types (like cascade correlation, back propagation ...) with a time consuming parameter adjustment. The SENN-based development of neural networks can be better supported by tools because network parameters for the training of the perceptron can be adjusted automatically. Thus, the training and the test of networks can be conducted without any further manual interaction of the user. This also improves the quality of the project. The use of representations and encodings for an associative recall of examples might also enable both an early evaluation of the first project phases as well as context dependent analysis of the competence of the network.

References

ANDERSON, J. A., ROSENFELD, E. (Eds.) (1988): Neurocomputing-Foundations of Research MIT Press.

ANDREASSEN, S., ENGELBRECHT, R., WYATT, J. (Eds.) (1993): Technology and Information 10, Artificial Intelligence in Medicinic, p. 466–476. IOS Press, Amsterdam.

BENAIM, M. (1994): On functional Approximation with Normalized Gaussian Units Neural Computation **6**, p. 319–333.

BOEHM, B. W. (1988): A spiral model of software development and enhancement, Computer, **May 1988**, pp. 61–72.

DÄSCHLEIN, R. WASCHULZIK, T., BRAUER, W., ARNOLDI, H.-M. R., ENGELMEIER, K.-H. (1994): Segmentation of Computertomographies with Neural Networks based on Local Features. In: Ifeachor, E., Rosen, K. (eds), Proceedings of the International Conference on Neural Networks and Expert Systems in Medicine and Healthcare, University of Plymouth, Plymouth.

DROSTE, K., BOLLSCHWEILER, E., WASCHULZIK, T., ENGELBRECHT, R., MARUYAMA, K., SIEWERT, J. R. (1994): Preoperative Prediction of Lymph Node Metastasis in Gastric Cancer with Neural Networks, Journal of Cancer Research and Clinical Oncology, Supplement to Volume 120, 21st National Cancer Congress of the German Cancer Society, Abstracts of Lectures and Poster, Springer International.

EGMONT, PETERSON, M., TALMON, J. L., BRENDER, J., NCNAIR, P. 1994: On the quality of neural net classifiers. Artificial Intelligence in Medicine, Vol. **6**, No. 5, pp. 359–381.

[4]) The time needed for the simulations was rounded up, because the simulations did also run on days where no user interactions where required

EHRLICHER M. (1993): Untersuchungen eines strukturadaptiven Algorithmus zum Aufbau von RBF-Netzen mit topologisch codierter Ausgabeschicht. – Diploma theses at the Technische Universität München, Institut für Informatik, Lehrstuhl Prof. W. Brauer.

ELDRACHER, M., WASCHULZIK, T. (1993): A Topologically Distribued Encoding to Facilitate Learning. Journald of Systems Engineering **3**, p. 110–119. Springer Verlag, London.

ENGELBRECHT, R., RECTOR, A., MOSER, W. (1994): Verification and validation. In: Assessment and Evaluation of Information Technologies in Medicine (Eds.: van Gennip E. M. S. J., Talmon J. L.) Amsterdam, IOS Press, p. 51–66.

ENTENMANN, G., ROSENBAUER, J., KIRCHNER, K., KOSCHINSKY, T., HAUNER, H. GRIES, F. A. (1995): Wie zuverlässig ist die Diagnose des Sekundärversagens der Sulfonylharnstofftherapie bei Typ II Diabetikern. Submitted for the 30th Diabetes congress of the Deutsche Diabetes Gesellschaft, Nürnberg.

EUROPEAN IDDM POLICY GROUP (Eds) (1993): Consensus Guidlines for the Management of Insulin-dependent (Type 1) Diabetes, Medicom Europe BV, Bussum, Niederland.

EUROPEAN NIDDM POLICY GROUP (1990): A Desktop Guide for the Management of non-insulin-dependent diabetes melitus. IDF Bulletin 25, 1990 1.

FAHLMAN, S. E., LEBIERE, C. (1991): The Cascade-Correlation Learning Architecture. School of Computer Science, Carnegie Mellon University, Pittsburg, USA.

FÖRSTER, M., KIRCHNER, K., WASCHULZIK, T., et al. (1994): DIADOQ: „Untersuchungen zum Aufbau von Wissensbasen in der Diabetologie mittels Neuronaler Netze". 39. Jahrestagung der GMDS, Dresden.

FRITZKE, B. (1993): Growing Cell Structures – A Self-organizing Network for Unsupervised and Supervised Learning, ECSI TR-93-026.

GEIGER, H. (1990): Storing and Processing Information in Connectionist Systems. – In: R. Eckmiller, editor, Advanced Neural Computers, p. 271–277. Elsevier Science Publishers B. V. (North Holland).

GROTHE, F. (1993): Erstellen eines halbautomatischen Systems zur Analyse großer Datenmengen mit neuronalen Netzen. – Diploma thesis at the Technische Universität München, Institut für Informatik, Lehrstuhl Prof. W. Brauer.

HEJLESEN, O. K., ANDREASSEN, S., ANDERSEN, S. K. (1993): Implementation of a Learning Procedure for Multiple Observation in a Diabetes Advisory System Based on Causal Probabilistic Networks. In: Andreassen, S., Engelbrecht, R., Wyatt J. (Eds.) Technology and Informatics 10, Artifical Intelligence in Medicine, IOS Press, Amsterdam, 63–74.

HESSE, W., MERBETH, G., FRÖHLICH, R. (1992): Software-Entwicklung, In: Handbuch der Informatik Band 5.3. Oldenbourg Verlag München, Wien.

HINTON, G. E., MCCLELLAND, J. L., RUMMELHART, D. E. (1986): Distributed Representations p. 77–109. In: Rummelhart, D. E., McClelland, J. L. and the PDP Research Group: Parallel Distributed Processing, MIT Press.

IFEACHOR, E., ROSEN, K. (Eds.) (1994): Proceedings of the Inernational Conference on Neural Networks and Expert Systems in Medicine and Healthcare, University of Plymouth, Plymouth.

KRANS, H. M. J., PORTA, M., KEEN, H. (Eds.) (1992): Diabetes Care and Research in Europe: The St. Vincent Declaration Action Programme. World Health Organization, Regional Office for Europe, Copenhagen; International Diabetes Federation, European Region.

MINSKY, M., PAPERT, S. (1969). Perceptrons, Cambridge, MA: MIT-Press.

MOODY, J., DARKEN, C. (1989): Fast learning in networks of locally-tuned processing units, Neural Computation, Vol. 1, p. 281–294.

NELSON, M. M., INLLINGWORTH, W. T. (1990): A practical Guide to Neural Nets, Addison-Wesley Publishing Company.

NÖHMEIER, M., WASCHULZIK, T., BRAUER, W., GROTHE, F., ENGELBRECHT, R. (1994): A Rule-based System for Realisation of Medical Projects with Neural Networks (HEAD). In: Ifeachor, E., Rosen, K. (eds.), Proceedings of the International Conference on Neural Networks and Expert Systems in Medicine and Healthcare, University of Plymouth, Plymouth.

PARK, J., SANDBERG, I. W. (1991): Universal approximation using radial-basis-function networks. Neural Computation **3**, p. 246–257.

PIWERNETZ, K., MASSI-BENEDETTI, M., VERMEJ, D., KÖCK, C. M., DE LEIVA, A., LANDGRAF, R. (1993): DiabCare Thinkshop "Quality network Diabetes". Diabetes, Nutrition & Metabolism **6**, 107–122.

POLZER, A. (1992): Invariante Mustererkennung durch neuronale Netzwerke mit großen rezeptiven Feldern. – Diploma thesis at the Technische Universität München, Institut für Informatik, Lehrstuhl Prof. W. Brauer.

QUANDT, K., WASCHULZIK, T., LEWIS, M., HÖRMANN, A., ENGELBRECHT, R., BRAUER, W. (1993): Evaluation of epidemiologic data with neural networks. World Congress on Neural Networks 1993, Proceedings. Vol. 1, p. 257–260. Lawrence Erlbaum Associates Inc., Hillsdale.

ROSENBLATT, F. (1958): The perceptron: a probabilistic model for information storage and organisation in the brain. Physiological Revies **65,** p. 386–408.

RUMELHART, D. E., HINTON, G. E., WILLIAMS R. J. (1986): Learning representation by back-propagating errors. Nature **323**, p. 533–536.

WASCHULZIK, T. (1988): Optische Musterkennung in neuronalen Architekturen. – Diploma thesis at the Techische Universität München, Institut für Informatik, Lehrstuhl Prof. Brauer.

WASCHULZIK, T., BÖLLER, D., BUTZH, D., GEIGER, H., WALTER, H. (1991): Neuronale Netze in der Automatisierungstechnik. In: Brauer, W., Hernández, D. (eds.), Informatik Fachberichte 291, p. 486–497. Springer Verlag, Berlin.

WASCHULZIK, T., GEIGER, H. (1990a): Eine Entwicklungsmethodik für strukturierte konnektionistische Systeme. – In: G. Dorffner (eds.), Informatik Fachberichte 252, Konnektionismus in Artifical Intelligence and Kognitionsforschung, p. 202–206. Springer Verlag, Berlin.

WASCHULZIK, T., GEIGER, H. (1990b): Theorie und Anwendung strukturierter Konnektionistischer Systeme. – In: G. Dorffner, editor, Informatik Fachbericht 252, Konnektionismus in Artificial Intelligence und Kognitionforschung, p. 143–152, Springer Verlag, Berlin.

WASCHULZIK, T., QUANDT, K., LEWIS, M., HÖRMANN, A., ENGELBRECHT, R., BRAUER, W. (1993): Evaluation of an epidemiological data set as an example of the application of neural networks to the analysis of large medical data sets. In: Adreassen, S., Engelbrecht, R., Wyatt , J. (eds.), Technology and Informatics 10, Artificial Intelligence in Medicine, p. 466–476. IOS Press, Amsterdam.

WASCHULZIK, T., BRAUER, W., FÖRSTER, M. KIRCHNER, K., ENGELBRECHT, R., SCHÜTZ, T., KOSCHINSKY, T., ENTENMANN, G. (1995): Quality Assurance and increased Efficiency in Medical Projects with Neural Networks by using a structured Development Method for Feedforward Neural Networks d(SENN). Accepted for the AIME95 conference, Pavia.

WEED, L. L. (1991): Knowledge Coupling: Springer Verlag, New York, USA.

WERBOS, P. J. (1974): Beyond regression: new tools for prediction and analysis in the behavioral sciences. Ph. D. thesis. Harvard University, Cambridge, MA.

WIDROW, B., HOFF, M. E. (1960): Adaptive switching circuits. 1960 IRE WESCON Convention Record, New York: IRE, p. 96–104.

YKI-JÄRVINNEN, H. (1994): Pathogenesis of non-Insulin-dependent diabetes mellitus. The Lancet Vol. 343, p. 91–95.

Anschrift der Verfasser: M. Förster: Institut für Medizinische Informatik, Fetscherstr. 74, D-01307 Dresden, Tel.: ++49 (0)35 14 58 33 45, Fax ++49 (0)35 14 58 43 17, foerster@rcs.urz.tu-dresden.de

T. Waschulzik, W. Moser, V. Böhm, R. Engelbrecht: GSF-MEDIS-Institut, PF 1129, D-857558 Oberschleißheim, Tel.: ++49 (0)89 31 87 44 76, Fax: ++49 (0)89 31 87 30 08, waschu@gsf.de

K. Kirchner, T. Koschinsky, G. Entenmann: Diabetes Forschungsinstitut, Auf'm Hennekamp 65, D-40225 Düsseldorf, Tel.: ++49 (0)21 13 38 26 11, Fax.: ++49 (o)21 13 38 26 00, kirchner@astro.dfi.uni-duesseldorf.de

W. Brauer, Technische Universität München, Institut für Informatik, D-80290 München, Germany , Tel.: ++49 (0)89 21 05 24 01, brauer@informatik.tu-muenchen.de

Informatik, Biometrie und Epidemiologie in Medizin und Biologie **26** (3),
215 – 228, ISSN 0943-5581
© Verlag Eugen Ulmer GmbH & Co., Stuttgart; Gustav Fischer Verlag KG, Stuttgart

Kausal-probabilistische Expertensysteme im MEDWIS-Projekt DIADOQ

CPN Expert Systems in Diabetics

U. G. Oppel[1]), U. Fischer[3]), K. D. Hepp[2]), S. Herfurth[4]), A. Hierle[1]), B. Liesenfeld[2]), H.-J. Lüddeke[2]), E. Salzsieder[3]), C. Sell[3])

Stichworte:

Kausal-probabilistische Netze, Expertensysteme; Evaluation; Glukose-Insulin-Stoffwechsel, diabetische Nephropathie, diabetische Retinopathie

Zusammenfassung

Die von der DIADOQ-Gruppe des BMBF-Förderschwerpunktes MEDWIS angewandten Methoden der Konstruktion und der Evaluation von kausal-probabilistischen Expertensystemen werden beschrieben: Wir skizzieren anhand von vier Beispielen drei verschiedene Methoden der Konstruktion der die Wissensbasis dieser Expertensysteme darstellenden kausal-probabilistischen Netze und drei verschiedene Methoden der statistischen Evaluation solcher Expertensysteme. Der Graph und die Markov-Kerne dieser kausal-probabilistischen Netze werden aus Kompartmentmodellen durch zeitliche Diskretisierung und Stochastifizierung, aus deterministischen Expertensystemen durch Stochastifizierung oder aus überwiegend qualitativem medizinischen Wissen durch verallgemeinerte lineare Modelle gewonnen. Die statistische Evaluation erfolgt durch Monte-Carlo-Simulation der Netze, durch statistische Tests von Marginalverteilungen der zu den Netzen gehörigen multivariaten Wahrscheinlichkeitsverteilungen und durch Distanzmessungen solcher Marginalverteilungen und entsprechender empirischer Verteilungen. Als Beispiele dienen Expertensysteme zum Glukose-Insulin-Metabolismus, zur Progression und Differentialdiagnose der diabetischen Nephropathie und zur Risikoabschätzung der diabetischen Retinopathie.

[1]) Mathematisches Institut der Ludwig-Maximilians-Universität München
[2]) Diabeteszentrum München-Bogenhausen
[3]) Institut für Diabetes Karlsburg
[4]) Augenklinik der Universität Greifswald

Keywords:

Causal probabilistic network, Bayesian network, expert system; glucose-insulin metabolism, diabetic nephropathy, differential diagnosis, diabetic retinopathy

Abstract

We sketch the methods of construction and evaluation of several expert systems based on causal probabilistic networks which have been designed by the DIADOQ group for the estimation of prognosis of diabetic nephropathy and retinopathy and for the simulation of the glucose-insulin metabolism.

Kurztitel:

CPN-Expertensysteme in DIADOQ

1. Einleitung

Im Förderschwerpunkt MEDWIS des BMBF will das Projekt „DIADOQ: Diabetes mellitus, optimierte Betreuung durch wissensbasierte Qualitätssicherung" durch Bereitstellung medizinischen Wissens die Qualität der Langzeitbetreuung von Patienten mit Diabetes mellitus in Bezug auf die diagnostischen Prozesse, die Therapiefindung und -planung sowie deren Überwachung verbessern. Hierzu sollen modulare Wissensbasen aufgebaut werden, um die einzelnen Elemente der Betreuung wirkungsvoll im Rahmen eines Dokumentationssystems zu unterstützen. Die eingesetzten Methoden der routinisierten Qualitätssicherung beziehen sich auf die Versorgung jedes einzelnen Patienten und die Qualitätsstandards der gesamten Einrichtung. Maßstab für die verbesserte Versorgung ist neben der Ergebnisqualität aller Einzelschritte der effiziente Ressourceneinsatz.

Zu DIADOQ haben sich neun Forschergruppen unter Leitung von R. Engelbrecht (GSF, München), U. Fischer (IDK, Karlsburg-Greifswald), A. F. Gries (DFI, Düsseldorf; Koordinator), K. D. Hepp (KMB, München), N. Jersch (Boehringer-Mannheim), H. Kunath (IMIB, Dresden), Th. Koschinsky (DFI, Düsseldorf), U. G. Oppel (LMU, München) und R. B. Sterzel (KSN, Nürnberg) zu einem Forschungsverbund zusammengeschlossen.

Ziel des Projektes ist also die Verbesserung der Betreuung der Typ I und Typ II Diabetiker durch Bereitstellung wissensbasierter Systeme. Diese wissensbasierten Systeme sind eingebunden in ein Dokumentationssystem zur Erfassung und Verwaltung von (Patienten- und Theorie-) Daten. Von diesem Dokumentationssystem aus werden sogenannte Protokolle und von diesen wiederum spezielle, durch neuronale oder kausal-probabilistische Netze dargestellte Wissensbasen aufgerufen.

Insbesondere die in Form kausal-probabilistischer Netze dargestellten Wissensbasen dienen zur Bearbeitung sehr komplexer Situationen. Sie sind auf Grund ihrer Struktur und der für sie vorhandenen Algorithmen zur Diagnose, Prognose und Simulation multifaktorieller Zusammenhänge geeignet.

Viele Bereiche ärztlichen Handelns laufen auf der Basis eines deterministischen Regelsystems ab. Die oben erwähnten Protokolle stellen ein solches geregeltes Handlungs- und Entscheidungsschema dar. Bei der Abarbeitung eines solchen Protokolles kommt es aber zu Situationen, wo das dem Arzt zur Verfügung stehende Wissen nicht ausreicht,

um die im Rahmen des Protokolls notwendigen Entscheidungen zu treffen. Für solche Situationen werden wissensbasierte Systeme bereitgestellt, beispielsweise auf kausal-probabilistischen Netzen basierende Expertensysteme.
In einem kausal-probabilistischem Netz **(CPN)** werden stochastische Ursache-Wirkungsbeziehungen eines größeren Systems von Eingabe- und Zielgrößen (z. B. Symptome, Befunde, Krankheiten) qualitativ durch einen gerichteten Graphen und quantitativ durch bedingte Wahrscheinlichkeiten beschrieben. Es hat sich im bisherigen Verlauf des Projektes DIADOQ gezeigt, daß entgegen ursprünglicher Annahmen die für ein kausal-probabilistisches Netz notwendigen bedingten Wahrscheinlichkeiten nicht aus dem umfänglichen, aber in Anbetracht der Komplexität der Fragestellung nur spärlichen Datenmaterial ermittelt werden können. Darum wurden von der DIADOQ-Arbeitsgruppe der LMU Methoden entwickelt, welche die Ermittlung der für die Konstruktion eines CPN notwendigen Daten erlauben. Diese Methoden bestehen meist aus einer Kombination statistischer Verfahren mit Expertenbefragungen.
Ein CPN bietet auf Grund seiner Struktur eine Möglichkeit zur Modellierung äußerst komplexen medizinischen Wissens. Die im Rahmen der HUGIN-Shell gegebenen Möglichkeiten zur Eingabe neuen Wissens und zu dessen Propagierung erlauben (durch Algorithmen auf der Basis der iterierten Anwendung des Satzes der totalen Wahrscheinlichkeit und des Satzes von Bayes) Vorwärts- und Rückwärtsschlüsse, die ein noch so qualifizierter Wissenschaftler nicht leisten könnte. Diese enorme „Denkfähigkeit" eines kausal-probabilistischen Expertensystems erlaubt dann selbst in sehr komplexen medizinischen Situationen noch Diagnose, Prognose und Simulation.
Die Übersetzung des medizinischen Expertenwissen in die für ein CPN notwendige quantitative Form muß durch geeignete Evaluationsprozeduren gesichert werden. Hierzu hat die LMU-Arbeitsgruppe geeignete Methoden entwickelt und angewandt. Diese Methoden sind für eine objektive Evaluation eines CPN unerläßlich, sie beruhen im wesentlichen auf einer Methode zur Monte-Carlo-Simulation eines CPN und einer Methode zur statistischen Testung (der Marginalverteilungen der assoziierten multivariaten Wahrscheinlichkeitsverteilung) des CPN mit Hilfe von Signifikanztests und Distanzmaßen. Die Kombination dieser beiden Methoden erlaubt die statistisch objektive Überprüfung eines CPN anhand von Patientendaten.

2. Kausal-probabilistische Netze

Die Modellierung biologischer Phänomene und Prozesse mit unsicherem Wissen kann mit kausal-probabilistischen Netzen (CPN) geschehen, welche im Englischen auch „Bayesian belief network" oder „causal probabilistic network" genannt werden; vgl. z. B. PEARL (1988), ANDREASSEN et al. (1987), LAURITZEN et al. (1988) und LAURITZEN (1992).
Bei kausal-probabilistischen Expertensystemen wird die Wissensbasis als CPN modelliert. Die Wissensbasis wird dabei in Form eines Netzes dargestellt, dessen Knoten stochastische Größen (Zufallsvariable) und dessen gerichtete Kanten qualitativ stochastische Abhängigkeitsbeziehungen dieser Zufallsvariablen beschreiben. Diese stochastischen Abhängigkeitsbeziehungen beschreiben beispielsweise mit Unsicherheit behaftete Ursache-Wirkungsrelationen, wie sie bei komplexen medizinischen Situationen häufig auftreten. Quantitativ werden diese Beziehungen durch bedingte Wahrscheinlichkeiten beschrieben. Sie stellen zusammen mit der Netzstruktur das Expertenwissen dar. Aufgrund dieses Expertenwissens können die Wahrscheinlichkeitsverteilungen der Zufallsvariablen automatisch vom System bestimmt werden. Bei der Eingabe neuer Informationen berechnet HUGIN die Wahrscheinlichkeitsverteilungen neu, und zwar im Einklang mit dem Expertenwissen und unter Berücksichtigung der neuen In-

formationen. Dies geschieht mittels einer die Topologie des Netzes ausnutzenden Folge von Rückwärts- und Vorwärtsverkettungen. Die Rückwärtsverkettungen basieren auf dem Satz von Bayes und damit auf einer bewährten Methode des stochastisch korrekten Schließens von der Wirkung auf die Ursache. Die Vorwärtsverkettungen basieren auf dem Satz der totalen Wahrscheinlichkeit und damit auf einer stochastisch korrekten Methode des Schließens von der Ursache auf die Wirkung. Dieser komplexe Prozeß der Bewertung und Neubewertung wird durch die Inferenzmaschine der Shell HUGIN erledigt; erst sie ermöglicht den Einsatz dieser Methodik zur Modellierung realistischer Systeme und zur Diagnose („Rückwärtsschließen") und Prognose („Vorwärtsschließen").

Medizinisches Wissen besteht aus Kenntnissen über Einfluß- und Zielgrößen und deren Zusammenspiel. Dieses Wissen besteht aus einer Mischung aus reinem Datenwissen, aus strukturellem Wissen, welches oft nur durch sporadische Daten gestützt wird, seltener aus multifaktoriellem Wissen, zu einem großen Teil aus Analogien und mehr oder weniger gesicherten Übertragungen, aus deterministischen und stochastischen Modellen und aus qualitativen und quantitativen Theorien. Wenn man medizinisches Wissen in Gestalt eines CPN darstellen will, so muß man einen gerichteten Graphen finden, dessen Knoten die Einfluß- und Zielgrößen repräsentieren und der die (stochastische) Abhängigkeitsstruktur dieser (im allgemeinen stochastischen) Größen qualitativ charakterisiert. Oft läßt sich dieser Graph in Zusammenarbeit mit medizinischen Experten im Rahmen der angestrebten Genauigkeit relativ leicht finden.

Darüber hinaus braucht man zu jedem Knoten eines CPN einen Markov-Kern. Ein solcher Markov-Kern besteht aus einer (im allgemeinen großen) Anzahl von bedingten Wahrscheinlichkeiten. Jede dieser bedingten Wahrscheinlichkeiten ist eine Zahl, welche die Wahrscheinlichkeit des Eintretens eines bestimmten Zustandes der durch den Knoten repräsentierten Zufallsvariable angibt unter der Bedingung, daß ganz bestimmte Zustände der durch die Elternknoten des vorgegebenen Knotens repräsentierten Zufallsvariablen gegeben sind. In dem hier und im folgenden immer betrachteten Fall von Zufallsvariablen mit endlich vielen Zuständen besteht also ein Markov-Kern für einen Knoten aus einer Familie von Zahlen, welche mit den Zuständen der durch den Knoten und seine Elternknoten repräsentierten Zufallsvariablen indiziert sind. Betrachten wir ein Beispiel: Ein Knoten habe 5 Elternknoten. Hat jede der zugehörigen Zufallsvariablen 10 Zustände, so besteht der Markov-Kern aus 10^6 (= 1000000) bedingten Wahrscheinlichkeiten. Zur Schätzung derartig vieler Wahrscheinlichkeiten werden aber wohl kaum genügend Daten zur Verfügung stehen. Also müssen für die Ermittlung der für ein CPN benötigten Markov-Kerne geeignete Methoden entwickelt werden.

Zur Konstruktion der Markov-Kerne eines CPN muß man so weit wie möglich Strukturen und quantitative Theorien und Modelle ausnutzen und die wegen der großen Komplexität nur sporadischen Daten durch geeignete Methoden ergänzen, interpolieren, extrapolieren und kombinieren. Dies ist eine Vorgehensweise, die durchaus medizinischem Denken entspricht. Aber je mehr deterministische oder stochastische quantitative Theorie vorhanden ist, um so eher wird man eine befriedigende Darstellung eines medizinischen Wissensbereiches durch ein CPN erhalten.

Die LMU-Arbeitsgruppe hat beispielsweise Methoden zur Transformation von durch deterministische Differentialgleichungssysteme beschriebenen Kompartmentmodellen für den Glukose-Insulin-Stoffwechsel in ein CPN entwickelt und Werkzeuge zur Automatisierung dieser Transformation erstellt. Ein Ergebnis der Anwendung dieses Transformationsverfahrens ist das unten näher beschriebene CPN „Karlsburger Modell 8/94", welches aus dem in Karlsburg am IDK von U. Fischer und E. Salzsieder und ihren Mitarbeitern entwickelten Kompartmodell des Glukose-Insulin-Stoffwechsels abgeleitet wurde. Diese Transformation führt zu einem erstens realistischeren stochastischen Modell des Glukose-Insulin-Stoffwechsels und erlaubt zweitens eine sehr einfache Handhabung

des Modells durch den nicht notwendig mathematisch geschulten Mediziner zur Diagnose, Prognose und sonst sehr aufwendigen Bestimmung individueller Modellparameter.
Außerdem wurden von der LMU-Arbeitsgruppe zur Bestimmung der Markov-Kerne lineare Modelle, Momentenschätzer und Monte-Carlo-Simulationsverfahren entwickelt und bei zwei der nachfolgend aufgeführten Beispiele (CPN „Nephropathie 12/94" zum Progressionsrisiko der diabetischen Nephropathie und CPN „Retinopathie 2/95" zur Retinopathie diabetischer Schwangerer) angewandt. Schließlich hat die KMB-Arbeitsgruppe eine Methode zur Transformation regelbasierter Expertensysteme entwickelt und bei dem nachfolgend beschriebenen CPN „Differentialdiagnose" durchgeführt.
Wenn endlich eine medizinische Wissensbasis in Form eines CPN dargestellt ist, so muß überprüft werden, ob die angewandten Methoden und die verwendeten Daten auch zu einem richtigen Ergebnis führen. Wir wissen, daß ein CPN in eindeutiger Weise zu einem Wahrscheinlichkeitsmaß $\mathbb{P}$ führt, welches dieses Wissen im Sinne der Stochastik vollständig beschreibt. Wir könnten also in einem beispielsweise statistischen Testverfahren überprüfen, ob dieses $\mathbb{P}$ mit unserem Datenwissen übereinstimmt. Da $\mathbb{P}$ im allgemeinen eine hochdimensionale multivariate Wahrscheinlichkeitsverteilung ist, werden wir im allgemeinen wohl nicht genügend Daten haben, um einen Signifikanztest ausreichender Güte durchführen zu können. Wir können statt $\mathbb{P}$ aber sicherlich niederdimensionale Marginalverteilungen von $\mathbb{P}$ mit größeren Erfolg statistisch testen. An Hand des CPN „Nephropathie 12/94" wurden exemplarisch Marginalverteilungen durch Signifikanztests (Chi-Quadrat-Anpassungstest) und durch asymmetrische und symmetrische Distanzmaße (z. B. Kullback-Leibler-Divergenz; L_1- und L_2-Norm) überprüft. Ein wichtiges Mittel bei dieser Art von statistischer Überprüfung eines CPN sind Methoden zur Monte-Carlo-Simulation.
Neben dieser objektiven statistischen Evaluation eines CPN gibt es auch „weichere" Überprüfungsmethoden. So kann ein erfahrener Arzt beispielsweise das CPN mit der Shell benutzen, um damit Diagnosen oder Prognosen durch Eingabe und Propagierung von Evidenz zu erstellen und diese mit „eigenen" Diagnosen und Prognosen oder Beispielfällen zu vergleichen.

3. Beispiele von kausal-probabilistischen Netzen im Bereich Diabetes

3.1. Das CPN „Karlsburger Modell 8/94"

3.1.1. Aufgabe des CPN „Karlsburger Modell 8/94"

Training der Diagnose- und Prognosefähigkeit und der Therapieplanung

Der insulinbehandelte Diabetes mellitus stellt hohe Ansprüche an das Engagement, Wissen und Können aller Personen, die am Versorgungsprozeß beteiligt sind (Patienten, Ärzte, Diabetesberater). Um die Betreuungsqualität aufrechtzuerhalten und ihre Verbesserung langfristig sichern zu können, ergeben sich spezifische Aufgaben an das Diabetesmanagement. Hierzu dienen auch modellbasierte Simulationssysteme für die Vorhersage von Auswirkungen therapeutischer Maßnahmen auf den Glukosestoffwechsel und somit zur Unterstützung der Auswahl geeigneter Therapiestrategien, die es dem Nutzer gestatten sollen, seine Entscheidungen an Hand des im Modell integrierten Expertenwissens rechnergestützt zu überprüfen.
Nicht Diabetes-Spezialisten, sondern medizinisch qualifizierte Personen (Studenten, Teilnehmer an Qualitätszirkeln) sollen mit dem auf der Shell HUGIN implementierten CPN „Karlsburger Modell 8/94" und seinen Verfeinerungen darin unterstützt werden, das komplexe Zusammenspiel der verschiedenen Komponenten und Aspekte der Diabetesbehandlung leichter zu erfassen. Durch Eingabe individueller Daten (Glukose-

spiegel, Insulinspiegel, Glukosezufuhr, Insulingabe, körperliche Aktivität) werden Diagnose, Prognose und Therapieplanung bei der Diabetesbehandlung mit diesem CPN wirkungsvoll unterstützt.

3.1.2. Generierung des CPN „Karlsburger Modell 8/94"

Die Modellierung des Glukose-Insulin-Stoffwechsels als CPN „Karlsburger Modell 8/94" basiert auf dem am Institut für Diabetes in Karlsburg seit 1983 entwickelten und vielfach getesteten Kompartmentmodell, welches durch das folgende System von vier gekoppelten Differentialgleichungen charakterisiert wird:

$$dx/dt = u + G_{exg} \tag{1}$$

(x Glukosekonzentration)

$$du/dt = -(b_1 + b_2)\,u - b_3(y + e) + b_1(b_0 - G_{exg}) \tag{2}$$

(u Rate des gesamten endogenen Glukosemetabolismus)

$$dy/dt = -ky + I_{exg} \tag{3}$$

(y Insulinkonzentration)

$$de/dt = -ke + E_{exg} \tag{4}$$

(e Insulin-Äquivalent zu körperlicher Aktivität)

G_{exg} exogene Glukosezufuhr
I_{exg} exogene Insulinzufuhr
E_{exg} physische Aktivität
b_0, b_1, b_2, b_3 und k: individuelle Parameter des Systems

Bei der Transformation des Differentialgleichungssystems (1)–(4) in ein CPN werden die Zustandsvariablen x, u, y und e zeitlich diskretisiert und in Knoten X_n, U_n, Y_n und E_n überführt. Auch die individuellen Parameter des Systems b_0, b_1, b_2, b_3 und k und die zeitlich variablen „Kontrollfunktionen" G_{exg}, I_{exg} und E_{exg} werden als Knoten dargestellt. Die für das CPN notwendigen Markov-Kerne erhält man durch Lösen einer sehr großen Zahl von Anfangswertproblemen und durch Stochastifizierung der durch diese Lösungen der Anfangswertprobleme bestimmten deterministischen Übergänge. Die LMU-Arbeitsgruppe hat das Programmpaket SNULLIND erstellt, welches diese Aufgabe automatisch erledigt. Für Einzelheiten zum Transformationsverfahren und zum resultierenden CPN vergleiche man OPPEL et al. (1993) und OPPEL (1995).

3.1.3. Evaluation des CPN „Karlsburger Modell 8/94"

Geht die Stochastifizierung (z. B. im Sinne der schwachen Konvergenz von Verteilungen) gegen Null, so konvergiert der stochastifizierte Glukose-Insulin-Prozeß gegen den durch das Differentialgleichungssystem (und eventuell eine Startverteilung) gegebenen „deterministischen" Glukose-Insulin-Prozeß (ebenfalls im Sinne der schwachen Konvergenz von Verteilungen). Insbesondere erhält man für feste Parameter und feste Anfangs- oder Endwerte des Glukose- und des Insulinspiegels bei fest vorgegebenen zeitlich variablen Kontrollfunktionen G_{exg}, I_{exg} und E_{exg} als Grenzfall des stochastifizierten und in seiner Stochastifizierung gegen Null gehenden Prozesses den für diese Parameter und Werte durch das deterministische Differentialgleichungssystem bestimmten Glukose-Insulin-Verlauf. Da das deterministische Karlsburger Kompartmentmodell in der Vergangenheit in vielfacher Weise getestet und als vernünftig bestätigt wurde, erübrigt sich damit eine statistische Evaluation des aus dem Karlsburger Modell durch diese Transformation gewonnenen CPNs. Allenfalls sind die Tauglichkeit der Diskretisierung und der Grad

der Stochastifizierung zu überprüfen. Man vergleiche hierzu die in OPPEL (1995) und OPPEL et al. (1995A) gemachten Bemerkungen.

3.1.4. Literatur zum CPN „Karlsburger Modell 8/94"

OPPEL et al. (1993), OPPEL (1995), OPPEL (1995), OPPEL et al. (1995B), SALZSIEDER et al. (1994); FISCHER et al. (1980), FISCHER et al. (1986), SALZSIEDER et al. (1985), SALZSIEDER et al. (1993), SALZSIEDER et al. (1994).

3.2. Das CPN „Nephropathie 12/94"

3.2.1. Aufgabe des CPN „Nephropathie 12/94"

Progressionsbeurteilung der diabetischen Nephropathie bei Diabetes mellitus Typ 1

Durch das CPN „Nephropathie 12/94" soll die Risikoabschätzung der Progression erfolgen. Entscheidend hierbei ist das Vorliegen von Verlustvariablen, wie HbA1c, Mikroalbumin, alpha-1-Mikroglobulin, Therapie mit ACE-Hemmern und Blutdruck über einen Zeitraum von zwei Jahren. Mit Abnahme des Beobachtungszeitraums sinkt die Genauigkeit der Progressionsbeurteilung. Derzeit wird die Länge der notwendigen Beobachtungszeit und die Anzahl der entscheidenden Parameter geprüft, welche zur Beurteilung des Progressionsrisikos in einem Zeitraum von wenigen Jahren notwendig sind. Feststehende Größen wie Alter, Nikotinkonsum oder genetische Belastung wie z. B. diabetische Nephropathie der Eltern oder kardiovaskuläre Risikokonstellation der Eltern gehen in die Risikobeurteilung mit ein.

3.2.2. Generierung des CPN „Nephropathie 12/94"

Bei der Konstruktion eines kausal-probabilistischen Netzes muß eine Auswahl von für diese Fragestellung möglichst relevanten quantitativen oder qualitativen Einflußgrößen erfolgen. Danach ist der gerichtete Graph als die qualitative Darstellung der Interdependenzen dieser Einflußgrößen zu ermitteln. Anschließend müssen die diese Interdependenzen quantitativ beschreibenden bedingten Wahrscheinlichkeiten der Markov-Kerne des zu konstruierenden CPNs mit geeigneten Methoden ermittelt werden.
In enger Zusammenarbeit der DIADOQ-Projektpartner KMB und LMU wurde die Fragestellung präzisiert, die Variablen ausgewählt und der Graph und die Markov-Kerne ermittelt. Letzteres geschah mit Hilfe einer Methode der linearen Modelle. Eine detailliertere Beschreibung der Generierung des CPN „Nephropathie 12/94" findet man in OPPEL et al. (1995B).

3.2.3. Knoten des CPN „Nephropathie 12/94"

Alter des Patienten
Medikation
Bekommt der Patient ACE-Hemmer?
HbA1c-Wert
Blutdruckwerte
Nephropathie-Risiko
Anamnese der Eltern
Rauchgewohnheiten des Patienten (Nikotin)
Urinalbuminwerte
alpha-1-Mikroglobulin (Urinwerte)

Eine detailliertere Beschreibung der Bedeutung der Knoten und der gewählten Diskretisierung findet man in OPPEL et al. (1995B).
Mittels der Knoten soll das Progressions-Risiko abgeschätzt werden. Dabei haben die einzelnen Faktoren ein unterschiedlichen Einfluß auf das Risiko (also: Gewichtung). Den stärksten Einfluß haben: HbA1c, Albumin, RR; nächster Faktor in der Rangfolge ist: ACE; dann: Alter; den schwächsten Einfluß haben: Anamnese, alpha-1-Mikroglobulin, Nikotin.

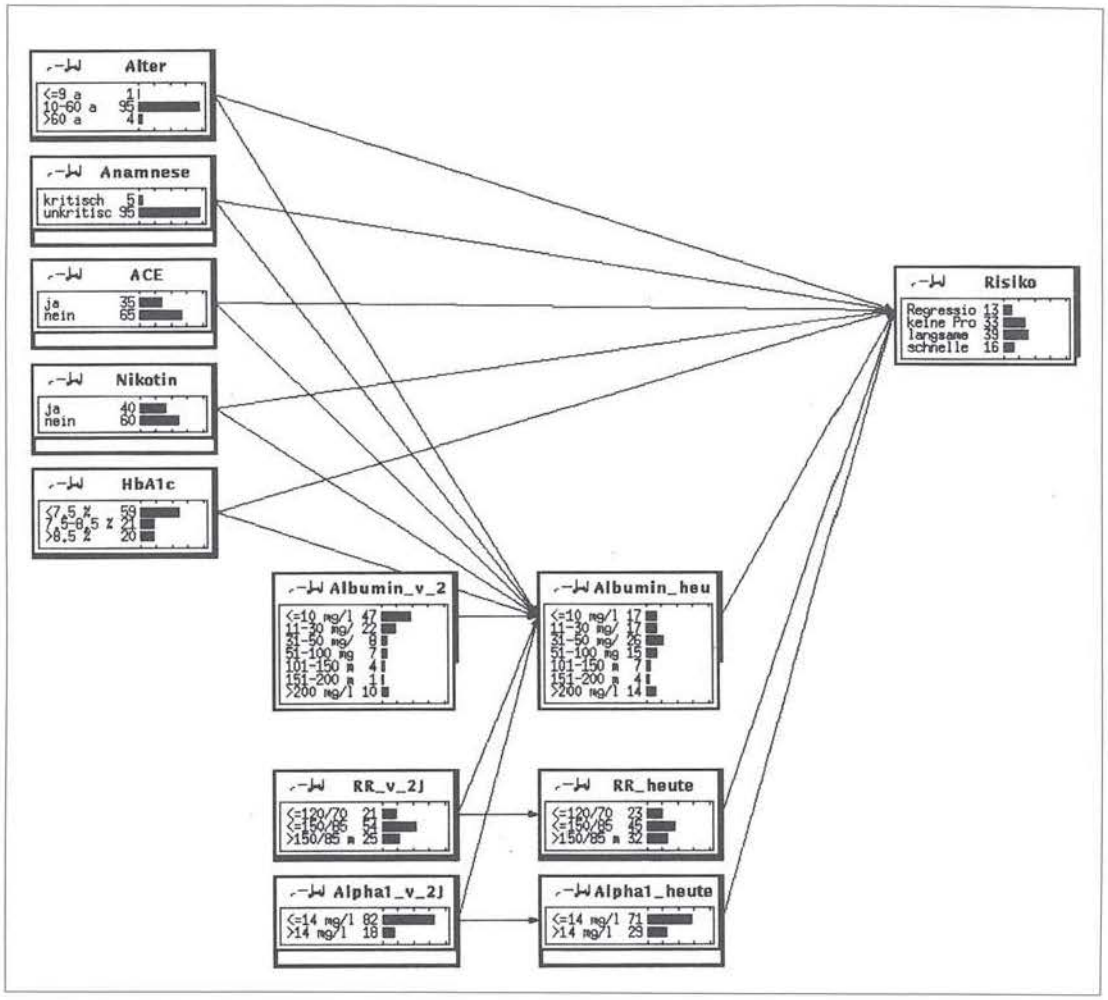


Abbildung 3.2.1: CPN „Nephropathie 12/94". Graph, alle Fenster geöffnet. Die Fenster zeigen die eindimensionalen Marginalverteilungen der durch die zugehörigen Knoten bestimmten Variablen im Initialzustand

3.2.4. Überprüfung des CPN „Nephropathie 12/94"

Um das CPN „Nephropathie 12/94" zu testen, stehen mehrere Möglichkeiten zur Auswahl, die alle benutzt wurden:

- Ein medizinischer Experte (möglichst ein anderer, als derjenige, der das Netz generierte) gibt ihm vertraute Situationen als Evidenzen in der CPN-Shell ein und vergleicht die Prognose, die das Netz liefert, mit seiner eigenen Einschätzung des Falls.
- Es werden reale Falldaten (mit bekannter Risiko-Prognose) von Hand als Evidenzen in das CPN eingegeben und die Prognosen des Netzes mit denen des Experten (bzw. im retrospektiven Fall: mit den Verlaufsdaten) verglichen. (Dies soll durch ein noch zu entwickelndes Werkzeug speziell bei größeren Datenmengen auch automatisch

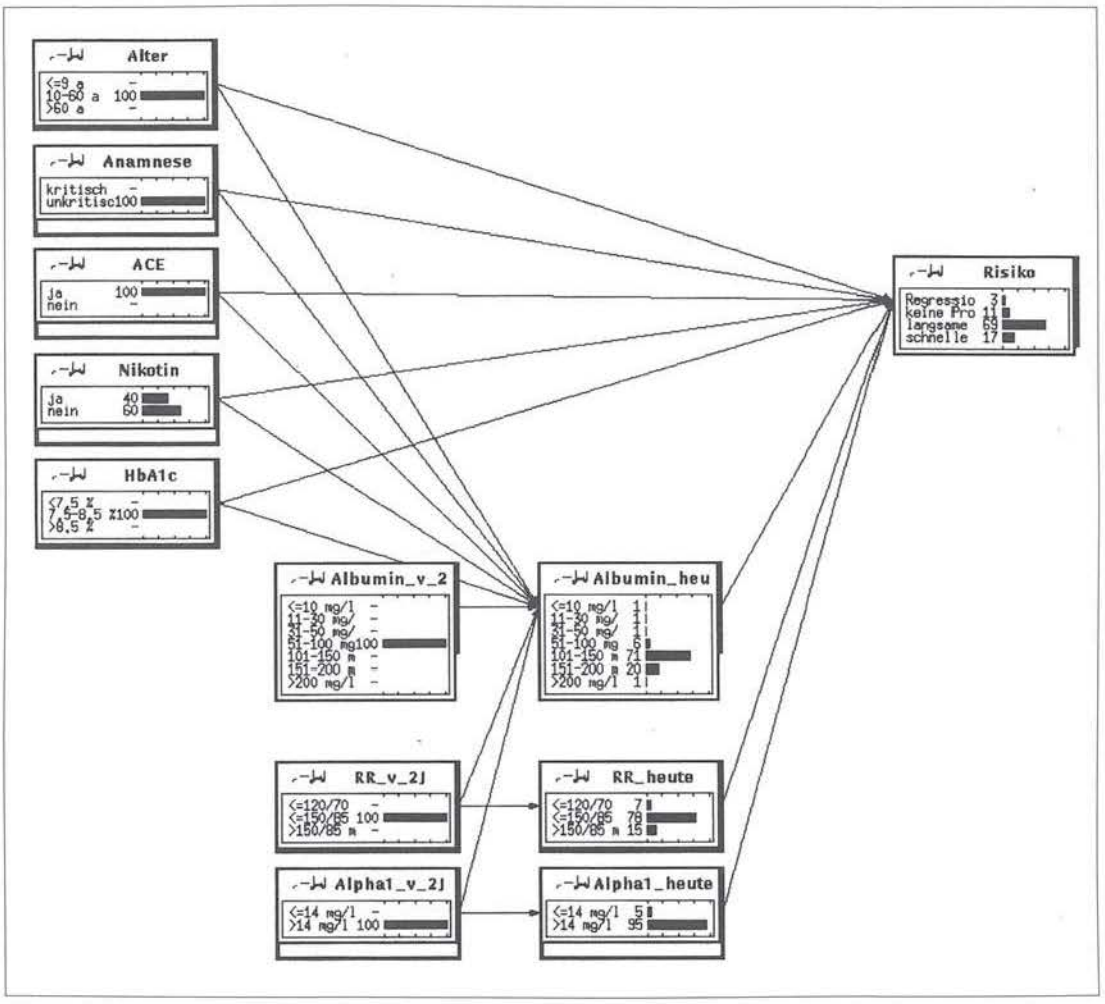


Abbildung 3.2.2: CPN „Nephropathie 12/94". Graph, alle Fenster geöffnet. Die Fenster zeigen die eindimensionalen Marginalverteilungen der durch die zugehörigen Knoten bestimmten Variablen nach Propagierung der Evidenz. Evidenz (100%-Balken) ist eingegeben in die Knoten: Alter, Anamnese, ACE, Nikotin, HbA1c; Albumin-v-2J

geschehen können, und zwar in dem Sinne, daß bei großen Datensätzen nicht mehr jeder Fall von Hand eingetragen werden muß.)

– Mit Hilfe von Monte-Carlo-Simulationen entsprechend den Markov-Kernen, die die Wissensbasis repräsentieren, werden Realisationen der durch das Netz gegebenen Wahrscheinlichkeitsverteilung simuliert und mit realen Daten verglichen (z. B. mit dem Chi-Quadrat-Anpassungstest, der Kullback-Leibler-Divergenz, der L_1- und L_2-Norm); vgl. OPPEL et al. (1995A) und OPPEL et al. (1995B). Dabei erhält man Aufschluß über die zu erwartenden Schwankungen der Daten (wie viele Datensätze müßten gesammelt werden, um zuverlässige Verteilungen zu erhalten?), sowie ein geeignetes Verwerfungskriterium.

3.2.5. Literatur zum CPN „Nephropathie 94/12"

LIESENFELD et al. (1994), OPPEL et al. (1995A), OPPEL et al. (1995B).

3.3. Das CPN „Differentialdiagnose"

3.3.1. Aufgabe des CPN „Differentialdiagnose"

Differentialdiagnose zur Abgrenzung der diabetischen Nephropathie von einer nichtdiabetischen

Die Differentialdiagnose gliedert sich in die Teilschritte „Erkennung" und „Differenzierung" einer Nephropathie.

Im *Erkennungsschritt* werden Serumkreatinin, Proteinurie, Leukozyturie oder Hämaturie im wesentlichen qualitativ beurteilt, um überhaupt das Vorliegen einer Nierenpathologie anzuzeigen. Das Vorliegen einer Albuminurie berechtigt allein bei weitem nicht zur Stellung einer bestimmten Diagnose, insbesondere nicht der der diabetischen Nephropathie.

Im *Differenzierungsschritt* werden Parameter wie Mikroalbumin, alpha-1-Mikroglobulin, alpha-2-Makroglobulin, Immunglobulin G und N-Acetyl-beta-D-Glukosaminidase und deren Relation zueinander zur Einordnung der Proteinurie in interstitiell, glomerulär, postrenal oder prärenal herangezogen. Die klinischen Angaben Diabetes mellitus oder Retinopathia diabetica erlauben dann mit einiger Sicherheit die Zuordnung zur Gruppe der diabetischen Nephropathie, sofern ein typisches Urineiweißmuster vorliegt und keine interkurrenten Erkrankungen bestehen.

3.3.2. Generierung des CPN „Differentialdiagnose"

Das CPN „Differentialdiagnose" wurde von der KMB-Arbeitsgruppe auf der Basis eines von W. Hofmann und W. G. Guder und ihren Mitarbeitern am KMB entwickelten deterministischen Regelsystems zur Differentialdiagnose der diabetischen Nephropathie konstruiert. Zunächst wurde ein diesem Regelsystem entsprechendes regelbasiertes Expertensystem in groben Zügen konzipiert. Mit einer nicht ganz einfach zu findenden Diskretisierung und einer anschließenden Stochastifizierung wurde das regelbasierte Expertensystem in ein CPN transformiert. Mit den damit möglichen probabilistischen Vorwärts- und insbesondere Rückwärtsschlüssen erhält das auf diesem CPN basierende kausal-probabilistische Expertensystem eine ganz andere Qualität als das ursprüngliche regelbasierte Expertensystem.

3.3.3. Knoten des CPN „Differentialdiagnose"

Stix für Hämoglobin	Stix für Leukozyten
N-Acetyl-Glukosaminidase (NAG)	Gesamteiweiß
Alpha-1-Mikroglobulin	Alpha-2-Makroglobulin
Albumin	Immunoglobulin G
Diabetes	Quotient Alpha-1-Mikro/Albumin
Quotient Alpha-2-Makro/Albumin	Quotient Immunglobulin G/Albumin
Selektivität	Prärenale Proteinurie
Differentialdiagnose	Hämaturie
Harnwegsinfekt (UTI)	Diabetische Nephropathie
Proteinurie	Urinstix

Eine detailliertere Beschreibung findet man in LIESENFELD et al. (1994) und OPPEL et al. (1995B); vgl. auch GUDER et al. (1993), HOFMANN et al. (1993).

3.4. Das CPN „Retinopathie 2/95"

3.4.1. Aufgabe des CPN „Retinopathie 2/95"

Progressionsbeurteilung der diabetischen Retinopathie bei Schwangeren mit Diabetes mellitus Typ 1

Obwohl die Behandlung des insulinpflichtigen Diabetes mellitus in den letzten Jahren einen bemerkenswerten Wandel erfahren hat, wird die diabetische Schwangerschaft auch heute als Hochrisikosituation für Mutter und Kind angesehen (IDDM-Konsensus 1993). Die Entwicklung oder Progression der diabetesspezifischen Retinopathie stellt für die Mutter das Hauptrisiko dar. Ausprägung und Schweregrad sind dabei von präkonzeptionell bereits vorhandenen Einflußfaktoren abhängig, deren prognostische Bewertung sowohl im Einzelfall als auch im komplexen Zusammenwirken problematisch ist.

3.4.2. Generierung des CPN „Retinopathie 2/95"

Die Konstruktion dieses CPNs erfolgte in enger Zusammenarbeit der Arbeitsgruppen des IDK und der LMU. Zunächst wurde von der IDK-Arbeitsgruppe das Problemformat und die Fragestellung zur Retinopathie der diabetischen Schwangeren präzisiert: Wie wird sich auf Grund der Konstellation der Einflußvariablen in der prägraviden Phase das Retinopathiestadium im Verlauf der Schwangerschaft voraussichtlich verändern, und welche Maßnahmen sind geeignet, dies günstig zu beeinflussen? Danach wählte die IDK-Arbeitsgruppe unter Berücksichtigung ophthalmologischer Gesichtspunkte die Variablen aus, die zur Beschreibung des Modells als

- notwendig (im Sinne von medizinisch relevant)
- zumindest in Spezialeinrichtungen (Diabetologen, Ophthalmologen) verfügbar

erachtet wurden. In einem ersten Testverfahren wurden diese 12 Variablen dann auf ihre Korrelation untereinander und auf die Zielvariable, das Retinopathiestadium am Ende der Schwangerschaft, untersucht. Von der LMU-Arbeitsgruppe wurde zunächst eine vorläufige Netzstruktur erstellt, die relativ unstrukturiert war und nur auf den unvollständigen statistischen Ergebnissen und groben ophthalmologischen Wertungen basierte.

Von dieser Grundlage ausgehend, wurde unter Berücksichtigung zusätzlichen ophthalmologischen Wissens und eines formalisierten Bewertungsverfahrens die vorliegende Netzstruktur von beiden involvierten Arbeitsgruppen gemeinsam entwickelt. Dabei wurden zunächst die Eingangsvariablen in vier Hauptgruppen aufgegliedert: Basis-, Einfluß-, Stoffwechsel- und Risikofaktoren. Diese Gruppierung entspricht weitgehend den Entscheidungskriterien der Ophthalmologen und dient als Grundlage sowohl für die Prognose des Retinopathiestadiums als auch der Beurteilung der Stabilität des Retinopathiestadiums. (Manchmal wird es für den Patienten als günstiger erachtet, wenn er ein zwar höheres, aber dafür stabileres Retinopathiestadium am Ende des Betrachtungszeitraumes erreicht.)

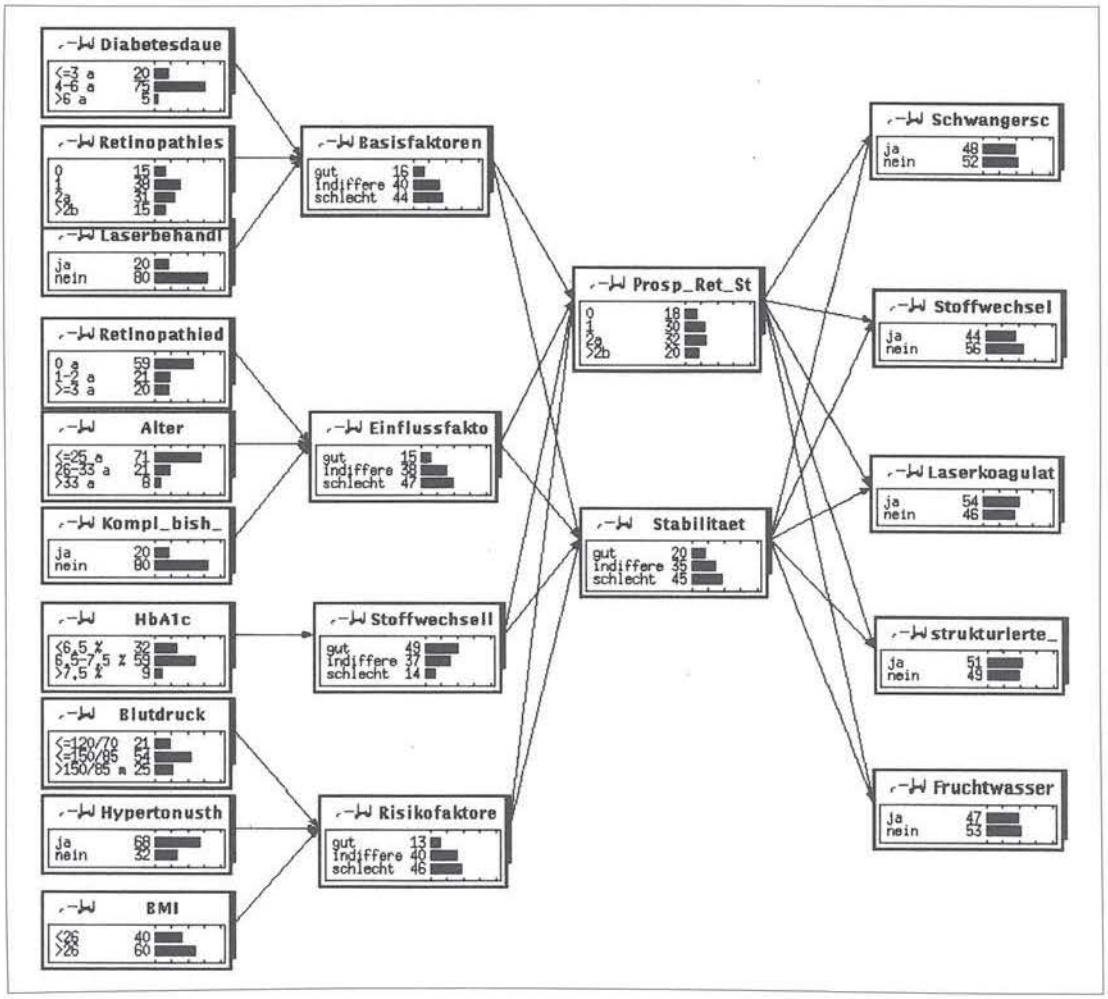
Anhand eines Bewertungsschemas wurden die Eingangsvariablen aus ophthalmologischer Sicht in Bezug auf die Gruppen gewichtet. So wurden beispielsweise die Zustände des Retinopathiestadiums vor der Schwangerschaft (0, 1, 2a, 2b) mit Gewichten versehen (Stadium 0 = sehr gut = Gewicht 2; Stadium 2b = sehr schlecht = Gewicht −2), diese mit dem Gewicht für den Knoten „Retinopathiestadium" multipliziert und für alle Eltern des Knotens „Basisfaktoren" aufsummiert. Dieser Score wurde zum Teil linear, zum Teil logarithmisch auf den Wertebereich (gut, indifferent, schlecht) abgebildet. Analog wurden die anderen Gruppen bewertet.

Aus der Prognose des Retinopathiestadiums und dessen Stabilität werden dann mittels eines auf einigen einfachen Regeln beruhenden Schemas die einzuleitenden Maßnahmen vorgeschlagen (z. B. bei wenig stabiler prognostizierter Retinopathie in mittlerem Stadium unbedingt vor Schwangerschaft eine Stoffwechselberuhigung erreichen). Für weitere Einzelheiten siehe OPPEL et al. (1995B).

3.4.3. Die Knoten des CPN „Retinopathie 2/95"

Diabetesdauer	Basisfaktoren
Retinopathie-Stadium	Einflußfaktoren
Laser-Behandlung	Stoffwechsellage
Retinopathie-Dauer	Risikofaktoren
Alter	prospektives Retinopathie-Stadium
Komplikationen bei bisherigen Geburten	prospektive Stabilität des R.-Stad.
HbA1c-Werte	Schwangerschaft
Blutdruck	Stoffwechselberuhigung
Hypertonus-Therapie	Laserkoagulation
Body-Mass-Index	strukturierte Schulung
	Fruchtwasser-Insulin

Abbildung 3.4.1: CPN „Retinopathie 2/95". Graph, alle Fenster geöffnet. Die Fenster zeigen die eindimensionalen Marginalverteilungen der zu den entsprechenden Knoten gehörigen Variablen im Initialzustand



3.4.4. Evaluation des CPN „Retinopathie 2/95"

Dieses CPN wird zur Zeit von den Arbeitsgruppen des DFI und des KMB durch Inspektion überprüft. Im Anschluß daran soll eine statistische Überprüfung mit den nur beim IDK ausreichend vorhandenen Daten erfolgen.

Literatur

ANDREASSEN, S., WOLDBY, M., FALCK, B., ANDERSEN, S. K. (1987): MUNIN — A Causal Probabilistic Network for Interpretation of Electromyographic Findings. Proceedings of the Tenth International Joint Conference on Artificial Intelligence, 366–372, Milano, Italy, August 1987.

FISCHER, U., JUTZI, E., BOMBOR, H., FREYSE, E.-J., SALZSIEDER, E., ALBRECHT, G., BESCH, W., BRUNS, W. (1980): Assessment of an algorithm for the artificial B-cell using the normal glucose-insulin relationship in diabetic dogs and men. Diabetologia **18** (1980), 97–107.

FISCHER, U., SALZSIEDER, E. (1986): New technologies in diabetes management. In: Serrano-Rios, M., Lefebre, P. J. (Eds.): Diabetes 1985. Amsterdam: Elsevier Publisher, 1986; 1036–1042.

GUDER, W. G., HOFMANN, W. (1993): Differentiation of Proteinuria and Haematuria by Single Protein Analysis in Urine. Clin. Biochem. **26** (1993), 277–288.

HOFMANN, W., SEDLMEIR-HOFMANN, C., IVANDIC, M., SCHMIDT, D., GUDER, W. G., EDEL, H. (1993): Befundung von Urin-Protein-Mustern auf der Basis klinisch gesicherter Patientenkollektive. Lab. Med. **17** (1993), 502–512.

LAURITZEN, S. L., SPIEGELHALTER, D. (1988): Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems. Journal of the Royal Statistical Society Series B **50** (2), 157–224.

LAURITZEN, S. L. (1992): Propagation of Probabilities, Means, and Variances in Mixed Graphical Models. JASA **87** (1992), 420, 1098–1108.

LIESENFELD, B., HIERLE, A. (1994): A New Clinical Decision Support Tool for Differential Diagnosis, Likelihood of Development and Progression of Diabetic Nephropathy in Insulin-Dependent Diabetics. Proceedings of the International Conference on Neural Networks and Expert Systems in Medicine and Health-care (NNESMED '94), Plymouth, August 24–26, 1994; 341–345.

OPPEL, U. G., HIERLE, A., JANKE, L., MOSER, W. (1993): Transformation of Compartmental Models into Causal Probabilistic Networks. In: Andreassen, S., Engelbrecht, R., Wyatt, J. (eds.): Artificial Intelligence in Medicine. Proceedings of the 4th Conference on Artificial Intelligence in Medicine Europe, 3–6 October 1993, Munich. IOS Press: Amsterdam–Oxford–Washington–Tokyo, 1993, p. 87–98.

OPPEL, U. G. (1994): Series of Causal Probabilistic Networks induced by Systems of Differential Equations for Diagnosis and Prognosis of Metabolic Processes. Proceedings of the Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU '94), Paris, July 4–8, 1994; 293–298.

OPPEL, U. G. (1995): Causal Probabilistic Networks and Their Application to Metabolic Processes. Proceedings of the Second Gauss Symposium, München, August 2–7, 1993. De Gruyter: Berlin, to appear 1995.

OPPEL, U. G., HIERLE, A., NOORMOHAMMADIAN, M. (1995A): Construction, Simulating and Testing of Causal Probabilistic Networks. To appear in: Proceedings of the "Invitational Workshop: Mathematical and Statistical Methods in Artificial Intelligence" (ISSEK '94), Udine (Italy), September 6–8, 1994.

OPPEL, U. G., HERFURTH, S., HIERLE, A., LIESENFELD, B., LÜDDEKE, H.-J., OPPEL, T., SALZSIEDER, E., SELL, C. (1995B): Bericht zum Stand der Entwicklung kausal-probabilistischer Expertensystem im MEDWIS-Projekt DIADOQ. Stand: Februar 1995. Mathematisches Institut der L-M-Universität München.

PEARL, J. (1988): Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann: San Mateo, California, USA.

SALZSIEDER, E., ALBRECHT, G., FISCHER, U., FREYSE, E.-J. (1985): Kinetic modeling of the glucoregulatory system to improve insulin therapy. IEEE Transactions Biomed. Eng. **32** (1985), 846–855.

SALZSIEDER, E., ALBRECHT, G., FISCHER, U., STÖWHAS, H., FRITZE, K., STEPHEN, R. L., CARSON, E. R. (1993): Estimation of individual action profiles of therapeutically administered insulin based on blood glucose levels in diabetic dogs and patients. Diab. Nutr. Metab. **6** (1993), 261–275.

SALZSIEDER, E., FISCHER, U., HIERLE, A., OPPEL, U. G. (1994): A Causal Probabilistic Network Associated to the Karlsburg Model of the Glucose-Insulin Metabolism for Approximate Assessment of Parameters, Diagnosis and Prognosis. Proceedings of the International Conference on Neural Networks and Expert Systems in Medicine and Healthcare (NNESMED '94), Plymouth, August 24–26, 1994; 112–120.

Korrespondenzadresse: Ulrich G. Oppel, Etterschlager Str. 10, 82237 Wörthsee, Telefon: 08153 – 7285 (eventuell auch: 089 – 2394 – 4489), Fax: 08153 – 89217

Informatik, Biometrie und Epidemiologie in Medizin und Biologie **26** (3),
229 – 239, ISSN 0943-5581
© Verlag Eugen Ulmer GmbH & Co., Stuttgart; Gustav Fischer Verlag KG, Stuttgart

Epistemologische Adäquatheit bei der Akquisition und formalen Repräsentation neurologischen Wissens

Epistemological Adequacy of the Acquisition and Formal Representation of Neurological Knowledge

Simone Bürsner[1]), Cord Spreckelsen[1]), Johannes Nebe[2]), Klaus Spitzer[1])

Kurztitel:
Wissensmodellierung in der Neurologie

Zusammenfassung

Bei der Wissensakquisition in der Neurologie unterscheiden wir zwei Formen von Wissensakquisition: Wissensakquisition auf konzeptueller Ebene und interaktive Wissensakquisition durch neurologische Fachexperten beim Füllen der Wissensbank mit konkreten Inhalten. Ziel des vorliegenden Papiers ist es, die Problematik der einzelnen Entwicklungsschritte im Rahmen der Wissensakquisition auf konzeptueller Ebene aufzuzeigen. Der dabei erarbeitete Ansatz zur Wissensrepräsentation ist gekennzeichnet durch das Bemühen um eine Repräsentation neurologischen Wissens, die zwar in der Form einheitlich ist, aber epistemologische Differenzierungen explizit zum Ausdruck bringt.
Ziel ist eine solche Modellierung von Domänenwissen, die den Aufsatz abstrakter Problemlösemodelle nicht ausschließt, deren leitende Intention es allerdings ist, die Wissensbank auf einen menschlichen ‚Problemlöser' auszurichten und für den menschlichen Umgang mit Wissen transparent zu gestalten.

Summary

With regard to the knowledge acquisition in neurology we distinguish two different forms of knowledge acquisition: knowledge acquisition on the conceptual level and interactive knowledge acquisition by neurological experts filling the knowledge base with concrete contents. It is the major aim of this paper to point out the difficulties with respect to the

[1]) Universität Heidelberg, Institut für Medizinische Biometrie und Informatik, Abt. Medizinische Informatik, Heidelberg
[2]) Universität Essen, Neurologische Klinik, Essen

different steps of knowledge acquisition on the conceptual level. Our approach to knowledge acquisition is characterized by the endeavour towards a representation of neurological knowledge, which is formally homogeneous, but which nevertheless makes epistemological differences explicit.
We are aiming at a domain modelling approach, which does not exclude the attachment of abstract problem solving models. However, the major intention is to adjust the knowledge base to a human problem solver and to make it transparent for the knowledge management by human experts.

Keywords:

spiralförmige Wissensakquisition; Domänenmodellierung; Epistemologie; neurologisches Wissen; spiral-shaped knowledge acquisition; modelling domain knowledge; epistemology; neurological knowledge

1. Einleitung

In der Neurologie fehlt es wie in anderen medizinischen Disziplinen bis heute an wissensbasierten Systemen, die verschiedenste Arten neurologischen Allgemein- und Spezialwissens umfassen können und dieses angepaßt an die jeweils aktuelle klinische Fragestellung möglichst schnell verfügbar machen. Einen Beitrag zur Behebung dieses Defizites können *Assistenzsysteme* leisten. Sie streben keinesfalls an, den medizinischen Fachexperten zu ersetzen. Dieser bleibt in der Rolle des ‚Problemlösers', dem aber durch das Assistenzsystem ein erweiterter Wissensumfang perspektivisch zur Verfügung gestellt wird. In Analogie zu einer Inferenzmaschine greift der Fachexperte dabei auf das während verschiedener Problemlösungsschritte benötigte, situationsspezifische Wissen zu. Diese umfassende Form der Entscheidungsunterstützung übersteigt den Leistungsumfang eng begrenzter diagnostischer und therapeutischer Problemlöser (Expertensysteme) [SPITZER & BÜRSNER 1994]. Die mögliche Realisierung eines solchen Assistenzsystems präsentiert sich dem Fachexperten als ‚intelligenter' Hypertext. Allerdings können konventionelle Hypertextsysteme das erforderliche kontextsensitive Retrieval neurologischen Wissens wegen ihrer seitenorientierten Form der Wissensrepräsentation [BÜRSNER 1995] nicht leisten.
Das Konzept eines neurologischen Assistenzsystems ist eine Verallgemeinerung und Erweiterung der Systemspezifikation von NeuroN, einem wissensbasierten, hypermedialen Nachschlagewerk für zerebrovaskuläre Erkrankungen [BÜRSNER et al. 1994]. Die folgenden Ausführungen basieren auf Erkenntnissen, die im Rahmen der Wissensakquisition und der formalen Wissensrepräsentation im NeuroN-Projekt gewonnen wurden.

1.1. Problematik

Die Anwendbarkeit des Assistenzsystems ist direkt von der Qualität der Wissensakquisition und der formalen Wissensrepräsentation abhängig: insbesondere gehört zum Umfang des zu repräsentierenden Wissens auch solches Expertenwissen, das nur implizit gegeben ist. Beispielsweise müssen variable, situationsspezifische Perspektiven akquiriert werden [BÜRSNER & SCHMIDT 1995]. Hierzu ist es erforderlich, die Wissensakquisition

so weit wie möglich in die Hände des Fachexperten selbst zu legen. Dazu bedarf es eines Prozesses der *interaktiven Wissensakquisition*, der adäquate Repräsentationsstrukturen für verschiedenste Arten neurologischen Wissens voraussetzt. Die Konstruktion dieser Strukturen erfordert die Erhebung, Strukturierung und Formalisierung eines epistemologisch vielfältig differenzierten Wissens.
Es wurde früh erkannt, daß die Wissensakquisition dieser Differenzierung des Domänenwissens Rechnung tragen muß. KL-ONE stellt dazu – ausgehend von sogenannten *„epistemological primitives"* – Strukturierungshilfsmittel zur Verfügung [BRACHMANN & SCHMOLZE 1985].
Es folgen hier zunächst eine Reihe von Begriffserklärungen (modellbasierte Wissensakquisition, Epistemologie und Domänenontologien). Anschließend werden in Abschnitt 2 die Anforderungen dargestellt, welche durch die epistemologische Vielfältigkeit an die Wissensakquisition insgesamt und im besonderen an die formale Wissensrepräsentation gestellt werden. In Abschnitt 3 sollen Lösungsansätze für diese Problematik dargestellt werden, anhand derer wir die Strukturierungsprobleme bei der Entwicklung eines epistemologisch adäquaten Repräsentationsmodells für neurologisches Wissen veranschaulichen.

1.2. Begriffserklärungen

Ausgangspunkt der Betrachtungen ist die *modellbasierte Wissensakquisition in der Neurologie*. In diesem Rahmen analysieren wir die Problematik *epistemologischer Differenzierungen* des Domänenwissens und betrachten deren Konsequenzen für den Aufbau von *Domänenontologien*.

1.2.1. Modellbasierte Wissensakquisition in der Neurologie

Bei der modellbasierten Wissensakquisition in der Neurologie unterscheiden wir zwei Formen von Wissensakquisition:

(1) *Wissensakquisition auf konzeptueller Ebene* mit dem Ziel, ein Repräsentationsmodell neurologischen Wissens und ein entsprechendes Wissensakquisitionswerkzeug zu schaffen. Dieses Werkzeug dient einerseits der Akquisition, Operationalisierung und Pflege der Repräsentationsstrukturen und andererseits der interaktiven Akquisition neurologischen Wissens zum Aufbau einer entsprechenden Wissensbank.
Das Arbeiten auf konzeptueller Ebene ist eine deklarative und typisierende Strukturierung von Wissensbankinhalten. Vergleichbar ist dies mit der Klassendefinition in objektorientierten Programmierumgebungen, der Schemadefinition im Datenbankbereich und den Konzeptbeschreibungen bei der terminologischen Wissensrepräsentation, wie z. B. in der T-Box von KL-ONE [BRACHMANN & SCHMOLZE 1985].
(2) *Interaktive Wissensakquisition* durch neurologische Fachexperten beim Füllen der Wissensbank mit konkreten Inhalten.

In dem vorliegenden Papier beschränken wir uns auf die Darstellung der Problematik der einzelnen Entwicklungsschritte im Rahmen der Wissensakquisition auf konzeptueller Ebene.

1.2.2. Epistemologie

Der Begriff Epistemologie wird in der Informatik diffus verwendet. Daher ist zunächst zu klären, in welchem Sinn im folgenden von Epistemologie oder epistemologischen Differenzierungen gesprochen wird.
In der angelsächsischen Philosophie wird Epistemologie oft nicht scharf von Erkenntnis- oder Wissenschaftstheorie unterschieden. Eine genauere Begriffsbestimmung bildete sich im französischen Sprachraum heraus. Epistemologie wird dort als kritisches Studium

wissenschaftlicher Erkenntnis, einer Betrachtung ihrer zugrundeliegenden Prinzipien und Strukturen sowie ihrer theoretischen und praktischen Methoden *im Nachhinein* verstanden. Entscheidend ist, daß Epistemologie – so verstanden – darauf verzichtet, wissenschaftliche Erkenntnisprozesse an vorher gegebenen Prinzipien zu messen, und stattdessen einen hohen deskriptiven Anteil hat. Sie wird so einer möglichen Vielfältigkeit von Erkenntnisprozessen gerecht. Die mögliche Vielfältigkeit von Erkenntnisprozessen spiegelt sich in einer entsprechenden Vielfältigkeit von Wissen wider (vgl. u. a. [CANGUILHEM 1979].

Aus pragmatischen Gründen können die am Wissen selbst beobachtbaren Differenzierungen in den Vordergrund gestellt werden. Ein solches Vorgehen sieht davon ab, die kognitiven Prozesse, die zu diesen Differenzierungen führen, selbst zu klären. Stattdessen wird lediglich die Existenz solcher Prozesse unterstellt. Die modellhafte Reproduktion oder auch nur die Klassifizierung der Differenzierungen wird zum Gegenstand einer epistemologischen Betrachtung. In diesem Sinne erklären BRACHMANN und SCHMOLZE ihre Verwendung des Ausdrucks *„epistemology"*, mit dem sie „[. . .] *not a theory of any particular domain* [. . .] – *but part of a generative theory of the structure and limits of thought for a rational agent*" bezeichnen [BRACHMANN & SCHMOLZE 1985].

In Fortsetzug dieser Tendenz zur Abstraktion von dahinterliegenden kognitiven Prozessen sprechen beispielsweise WETTER und WOODWARD zunächst von der Annahme, daß dem *„knowledgeable behaviour"* kognitive, informationsverarbeitende Prozesse des *„knowers"* zugrundeliegen [WETTER & WOODWARD 1990]. Sie trennen in der Folge jedoch die Identifizierung, Rekonstruktion und Simulation solcher Prozesse von den charakteristischen Strukturen des Wissens selbst und sprechen vom epistemologischen Umfang einer Wissensrepräsentationssprache als den *„concepts and terms to describe knowledge"*.

Diese Sprechweise soll hier übernommen und noch schärfer gefaßt werden: Die Fragen, wie, warum oder wozu etwas gewußt wird, zielen auf verschiedene Aspekte (Modalität, Kausalität, Finalität) der Beschreibung von Wissen, welche über eine rein inhaltliche Charakterisierung des Wissens hinausgehen. Charakteristika, die in entsprechender Weise das Wissen selbst betreffen, sollen hier mit dem Ausdruck *epistemologische Charakteristika* bezeichnet werden. Unterscheidet sich Wissen hinsichtlich dieser Charakteristika, so wird im folgenden von *epistemologischen Differenzierungen* gesprochen. Werden solche Differenzierungen bei der Modellierung von Domänenwissen explizit repräsentiert, so soll das Repräsentationsmodell *epistemologisch adäquat* heißen.

1.2.3. Domänenontologien

Die Forderung nach epistemologischer Adäquatheit hat unmittelbare Konsequenzen für den Aufbau von Domänenontologien.

Domänenontologien wurden als abstrakte Bausteine zur Präfigurierung eines konkreten Domänenmodells eingeführt: „*We have attempted to determine a reasonable set of underlying object and relation types for knowledge structuring.*" [BRACHMANN & SCHMOLZE 1985].

Schon diese Strukturierung von Wissen kann im obigen Sinne als epistemologisch motiviert gelten. Wir gehen jedoch davon aus, daß es nicht ausreicht, ein Vokabular zu schaffen, um generell über Domänenwissen reden zu können. Eine zusätzliche Schwierigkeit besteht in der Abbildung der epistemologischen Differenziertheit des Domänenwissens. Domänenontologien müssen daher, um der Forderung nach epistemologischer Adäquatheit Rechnung zu tragen, Ausdruckmittel auch für epistemologische Differenzierungen sein.

Eine speziellere Diskussion ausgewählter Aspekte der Entwicklung von Domänenontologien erfolgt in Abschnitt 4.

2. Probleme mit epistemologischer Vielfältigkeit bei der Wissensakquisition und formalen Wissensrepräsentation

Bei der Entwicklung eines wissensbasierten Systems tritt das Problem der epistemologischen Vielfältigkeit auf. Dieses Problem ist weder durch eine Beschränkung der Domäne noch durch die Verwendung einer hybriden Form der Wissensrepräsentation zu umgehen:

(1) Zwar wird die Granularität des Wissens einerseits durch die Beschränkung auf ein bestimmtes Interessensgebiet und andererseits auf die Modellierung einzelner relevanter Perspektiven auf das Domänenwissen vergröbert. Dennoch hat die erste Einschränkung keine Auswirkung auf die epistemologische Vielfältigkeit und lediglich aus der zweiten resultiert eine Beschränkung der Breite epistemologischer Vielfältigkeit.

(2) Aus der Sicht der Wissensrepräsentation scheint es möglich, durch hybride Wissensrepräsentationsformate beliebige Wissensinhalte zu operationalisieren. Dadurch wird jedoch das Problem epistemologischer Adäquatheit nicht bewältigt. Zum Beispiel umfaßt die Ätiologie einer Krankheitsentität kausale Aspekte, die sich im allgemeinen hinsichtlich epistemologischer Charakteristika nicht unterscheiden. Die Adäquatheit der Modellierung betreffend, ist darauf hinzuweisen, daß der Versuch, Kausalität in einem hybriden Repräsentationsmodell – d. h. unter Verwendung verschiedener Formate – zu repräsentieren, die genannte Homogenität zerstört.

Probleme mit hybriden Wissensrepräsentationsformalismen beispielweise hinsichtlich der Repräsentation kausalen Wissens werden bereits in [VAN HEIJST, POST & SCHREIBER 1994] beschrieben. Statt einer Integration verschiedener Repräsentationsformalismen auf der Grundlage der „Form" des Wissens, schlagen VAN HEIJST et al. eine wissensbasierte Integration vor. Eine solche basiert auf dem „Inhalt" des Wissens.
Festzustellen ist ein Mangel an Ansätzen, die der Notwendigkeit einer epistemologisch adäquaten Repräsentation von Domänenwissen Rechnung tragen oder Methoden zur Bewältigung epistemologischer Differenzierungen bereitstellen.
Das Problem epistemologischer Differenzierungen von Domänenwissen verlangt nach weitergehender Behandlung. Zu berücksichtigen sind neben den Konsequenzen für den Repräsentationsformalismus auch die Auswirkungen auf die Strukturierungen des Prozesses der Wissensakquisition. Es liegt nahe, daß hierbei sukzessive bereits repräsentiertes Wissen in epistemologisch unterschiedlichen Kontexten wiederverwendet oder dabei neu strukturiert und revidiert werden muß. Einen möglichen Lösungsweg stellt ein spiralförmiges Vorgehensmodell dar.

3. Wissensakquisition in spiralförmigem Vorgehensmodell

Generell erweist sich die wiederholte Revision gefundener Repräsentationsstrukturen im Wissensakquisitionsprozeß als notwendig. Allein dies macht es erforderlich, die zur Wissensakquisition nötigen Entwicklungsschritte wiederholt, d. h. zyklisch, durchzuführen. Hierzu kommt als weitere Anforderung die sukzessive Integration bereits gefundener Repräsentationsstrukturen in Metakonstrukten, welche auch epistemologische Differenzierungen erlauben. Aus der zyklischen Anwendung der Entwicklungsschritte resultiert eine spiralförmige Evolution des Repräsentationsmodells.
Im folgenden gehen wir auf die besonderen Anforderungen und Probleme ein, die nach unserer Erfahrung mit einer solchen Vorgehensweise bei der Wissensakquisition verbunden sind. Anhand von Beispielen werden die Probleme und Lösungsansätze erläutert.

3.1. Strukturierung – Probleme und Lösungen

Die geforderte Revidierbarkeit und Integrationsfähigkeit der Repräsentationsstrukturen und -inhalte setzt voraus, daß das Prinzip der *unilokulären Wissensrepräsentation* eingehalten wird. Dies bedeutet, daß die Wissensbasis Sachverhalte, welche ein einzelnes identifiziertes Objekt betreffen, so organisiert, daß ein Zugriff auf sie an einer einzigen, gemeinsamen Stelle erfolgen kann.

3.1.1. Identifikation von Grundstrukturen

Kernwissen: Mit der Forderung unilokulärer Wissensrepräsentation ist die Modellierung atomarer Wissensbausteine verbunden. Dazu müssen *Kernobjekte* und *Kernbeziehungen* identifiziert werden, indem in einem medizinischen Quellentext oder von einem Experten zentrale Begriffe gefunden und miteinander in Beziehung gesetzt werden. Eine offensichtliche Schwierigkeit dabei ist, daß die übliche medizinische Sprache das Vorhandensein eines Kernobjekts in Fällen suggeriert, in denen geringe sprachliche Differenzierungen das Erkennen erheblicher sachlicher Unterscheidungen verdecken.

Zum Beispiel ist der Begriff ‚Schwindel' in der geläufigen medizinischen Terminologie unscharf: Ein Schwindel mit dem *Attribut* ‚drehend' hat von der neurologischen Bedeutung her kaum Gemeinsamkeiten mit einem ‚diffusen' Schwindel. Denn während der Drehschwindel das diagnostische Denken stark bündelt und auf eine Läsion des Vestibularapparates oder anderer Strukturen der Gleichgewichtssteuerung hinweist, läßt ein ‚diffuser' Schwindel mehr Möglichkeiten offen; hinter ihm verbirgt sich evtl. ein leichteres Schwindelgefühl bei Hyperventilation oder ein Schwarzwerden vor Augen bei orthostatischer Dysregulation. Daher sollte für eine adäquate Modellierung des Sachverhaltes davon abgesehen werden, ein einziges Kernobjekt ‚Schwindel' zu definieren.

Kernrelationen: Selbst wenn Kernobjekte identifiziert werden konnten, bleibt es problematisch, unter den Beziehungen zwischen ihnen einen Kern von Relationen, der für die Objekte essentiell und definierend ist, von zusätzlichen Relationen, in die die Objekte abhängig vom jeweiligen Kontext eintreten, zu unterscheiden. Im folgenden sprechen wir in diesem Sinne von *internen* und *externen Relationen* (s. die ausführlichere Darstellung hierzu in [BÜRSNER & SPRECKELSEN 1995]).

Beispielsweise wird der Morbus Menière durch die Trias der Symptome rezidivierende Drehschwindelattacken, einseitige Hörminderung und Tinnitus beschrieben. Diese Beziehungen zwischen der Krankheit und den Symptomen stellen interne Relationen dar. Eine externe Relation wäre die differentialdiagnostische Beziehung zwischen den Krankheiten ‚benigner paroxysmaler Lagerungsschwindel' und Morbus Menière, die wegen des gemeinsamen Symptoms ‚rezidivierende Drehschwindelattacken' besteht. Die von einem Objekt ausgehenden internen Relationen stellen *objektzentriertes Wissen* dar, während externe Relationen zum *nicht-objektzentrierten Wissen* gehören, welches auch Beziehungen zwischen mehr als zwei Objekten umfassen kann.

Parametrierung: Viele Beziehungen gelten nur unter bestimmten Voraussetzungen, welche bei der Beschreibung der Beziehungen berücksichtigt werden müssen. Beispielsweise würde die Kombination aus den Symptomen Hörminderung und Tinnitus bei Frauen mit einer Wahrscheinlichkeit X auf einen Morbus Menière deuten, während für Männer eine niedrigere Wahrscheinlichkeit Y anzunehmen ist. Zur Modellierung solcher Sachverhalte führen wir die Möglichkeit ein, Beziehungen durch *Parameter* zu qualifizieren. Dabei kann, wie im Beispiel eine strukturell identische Beziehung durch zwei parallele, aber unterschiedlich parametrierte Verweise zu modellieren sein.

Eine weitere Schwierigkeit ergibt sich dabei durch die zum Teil vielfältigen Möglichkeiten, einen Sachverhalt darzustellen. So kann es sich z. B. bei einer Beziehung zwischen den Objekten A und B als ungünstig herausstellen, eine Bedingung dieser Beziehung durch eine auf den Fall der Beziehung begrenzte Attributierung des Objekts B darzustellen,

statt die Bedingung in die Parametrierung der Beziehung einfließen zu lassen. Grundsätzlich ist bei einer adäquaten Modellierung darauf zu achten, daß das angebundene Objekt nur so allgemeingültig attributiert wird, daß es ein Kernobjekt bleibt.

3.1.2. Strukturierung komplexer Zusammenhänge

Die hiermit erreichte Modellierungsmächtigkeit erwies sich als zu gering: Lediglich durch Kernobjekte und qualifizierte Beziehungen zwischen diesen lassen sich bestimmte Teilbereiche neurologischen Wissens nicht erfassen.

Metakonstrukte: Zur Repräsentation komplexerer Zusammenhänge ist es notwendig, aus dem Netz der Kernbeziehungen herauszutreten und dabei beliebige Teilbereiche des Netzes einzeln bezeichnen zu können. Repräsentationsobjekte, die auf solche Teilbereiche des Kernwissens verweisen, heißen *Metakonstrukte*. Demnach strukturieren Metakonstrukte nicht nur einzelne Konzepte des Kernwissens, sondern selektieren auch bestimmte Beziehungen zwischen diesen und unterdrücken andere.
In der Pathophysiologie der Subarachnoidalblutung ist die Blutung aus einem basalen Hirngefäß von zwei Faktoren abhängig: Gewöhnlich liegt eine schwache Stelle in der Wand einer Arterie, z. B. ein Aneurysma, vor; ausgelöst wird die Blutung durch eine Erhöhung des intravasalen Drucks. Zur Beschreibung eines solchen Zusammenhangs, bei dem zwei ursächliche Faktoren einen neuen pathophysiologischen Zustand erzeugen, haben wir das Werkzeug eines Metakonstruktes eingeführt. In unserem Beispiel liegen dem Metakonstrukt externe Relationen zwischen drei Objekten zugrunde.
Mit Hilfe von Metakonstrukten lassen sich aber auch umfangreichere Zusammenhänge beschreiben: Die Blutung aus einem basalen Hirngefäß steigert das Volumen im Subarachnoidalraum. Dadurch erhöht sich der intrakranielle Druck, und als Folge werden die cerebralen Gefäße komprimiert. Dies führt zu einer Verringerung der oben erwähnten Blutung. Hier ist demnach ein Zyklus beschrieben, in dem neben verstärkenden Einflüssen auch eine dämpfende Wirkung (Schwächung der Blutung durch Kompression der Hirngefäße) enthalten ist und bei dem sich dadurch möglicherweise ein stabiler Zustand einstellt.
Doch auch Zyklen, in denen die Objekte durchgehend verstärkend aufeinander wirken, lassen sich mit einem Metakonstrukt beschreiben. In unserem Beispiel sind die Objekte intrakranielle Drucksteigerung und Gefäßkompression gleichzeitig Bestandteil eines solchen Circulus vitiosus, der durch die Folge *intrakranielle Drucksteigerung → Gefäßkompression → hypoxischer Hirnschaden → Hirnödem → intrakranielle Drucksteigerung* gebildet wird.
Hier bestehen parametrierte kausale Kernbeziehungen zwischen Kernobjekten, z. B zwischen der Gefäßkompression und dem hypoxischen Hirnschaden. Natürlich gibt es außerhalb des hier betrachteten pathophysiologischen Themas weitere Ursachen für einen hypoxischen Hirnschaden, z. B. eine globale Abnahme der Sauerstoffkonzentrationen bei einer Asystolie. Um den pathophysiologischen Zusammenhang einer Subarachnoidalblutung zu beschreiben, ist es unabdingbar, aus dem Netz der Kernbeziehungen herauszutreten und durch die Verwendung von Metakonstrukten aus dieser Perspektive manche Beziehungen zu selektieren und andere zu unterdrücken.
Die Zusammenfassung einer solchen Auswahl von Kernbeziehungen mit Hilfe eines Metakonstruktes etabliert externe Relationen zwischen den beteiligten Objekten. Dazu muß im operationalisierbaren Repräsentationsmodell auf alle bisher eingeführten Modellkomponenten ein separater Zugriff möglich sein, sowohl auf Kernobjekte als auch auf Kernbeziehungen. Es hat sich hierbei als wichtig erwiesen, grundsätzlich zu unterscheiden, ob eine adäquate Modellierung von Sachverhalten darin bestehen kann, bereits vorhandene Kernbeziehungen anzusprechen, oder ob es nötig ist, weitere Kernbeziehungen zu etablieren.

In Analogie zur oben beschriebenen Schwierigkeit der Identifikation von Kernobjekten stellt sich hier das Problem, sinnvolle Zusammenfassungen zu finden, also im Domänenwissen geeignete Metakonstrukte zu identifizieren.

3.1.3. Strukturierung von ‚Expertenwissen'

Zu den besonderen Herausforderungen der Wissensmodellierung zählt die Darstellung des oft nur implizit gegebenen ‚Expertenwissens'. Dazu zählt beispielsweise: Regelwissen, situationsabhängiges Wissen, didaktisches Wissen, strategisches Wissen zum wissenschaftlichen Erkenntnisgewinn, zum Diagnostizieren und zur Informationsbeschaffung, Expertenwissen umfaßt zwar komplexes Sachwissen, ist aber in weiten Teilen auch Wissen über den Umgang mit neurologischem Wissen.
Metakonstrukte sind ein geeignetes Werkzeug, um Wissen aufzubereiten, das einem Benutzer systematisch vermittelt werden soll. Dabei dienen sie sowohl der Gliederung und Themenbildung als auch der Darstellung zeitlicher, logischer oder kausaler Zusammenhänge, wie etwa im oben beschriebenen Beispiel der Subarachnoidalblutung. Diese Zusammenhänge sind notwendig, um Wissen erklärend darzustellen und didaktisch zu ordnen.

Regelwissen: Schwieriger ist es beispielsweise, einem Experten geeignete Beschreibungsmittel für das Wissen an die Hand zu geben, das bei der Sicherung einer Diagnose gebraucht wird. Hier bildet das Wissen, welche Symptome bei einer bestimmten Krankheit in welcher Häufigkeit vorkommen, nur die Beschreibungsgrundlage. Das klinische Diagnostizieren arbeitet dann vorwiegend mit dem Erwägen und dem Ausschließen von Differentialdiagnosen. In der Frage, welche Differentialdiagnosen erwogen werden müssen, geht das benötigte Wissen über die bloße Feststellung, welche Krankheiten häufig gemeinsame Symptome besitzen, deutlich hinaus. Beispielsweise fließen Zusatzinformationen wie das Alter, das Geschlecht, die Vorerkrankungen des Patienten, seine gegenwärtige Medikation und bereits vorliegende apparativ-diagnostische Ergebnisse in die Beurteilung der Situation ein. Alle diese Informationen verändern die Wertigkeit der in Betracht zu ziehenden Diagnosen. Ferner ist es möglich, daß bei manchen Krankheiten sofort Konsequenzen für die erste Versorgung oder die Therapie resultieren würden, auch diese bloße Möglichkeit beeinflußt das diagnostische Vorgehen. Zu beachten ist außerdem, daß nicht alle von einem Experten verwendeten Folgerungen die gleiche Qualität haben. Im Sinne der Strukturierung sind hier unbedingt kausal begründete von eher regelhaften, empirisch begründeten Folgerungen zu unterscheiden.

Kontextsensivitität: Aus struktureller Sicht ist ein Charakteristikum des Vorgehens zur diagnostischen Sicherung, daß geltende Zusammenhänge erst unter der Perspektive der jeweiligen diagnostischen Situation bewertet und nach jedem Schritt neu gewichtet werden. Dabei erscheinen die Objekte situationsabhängig in bestimmten Rollen. Solche Rollen können in Metakonstrukten abgebildet werden, die ein Objekt in einer bestimmten Rolle beschreiben. Damit ist das Wissen zum Agieren in einer Rolle unabhängig vom Kernwissen. Z. B. mag eine Krankheit für eine Zeitlang die Rolle einer Verdachtsdiagnose annehmen; dann würde durch das Metakonstrukt der Verdachtsdiagnosenrolle bestimmt, welche Differentialdiagnosen zu betrachten sind und welche Untersuchungen die Abgrenzung zwischen Verdachtsdiagnose und Differentialdiagnosen am besten ermöglichen.
Prinzipiell scheinen die Metakonstrukte eine wichtige Rolle bei der Darstellung auch des Expertenwissens zu spielen. Offensichtlich werden für diesen Fall aber erweiterte Metakonstrukte benötigt, wobei die Erweiterungen, die auf das jeweilige Problem zugeschnitten sind, regelartig die entsprechend gewichteten Ausschnitte des Wissens erzeugen müssen.

3.2. Epistemologische Aspekte

Schon die Unterschiede von objektzentriertem und nicht-objektzentriertem Wissen entspricht einer ersten epistemologischen Differenzierung. In diesem Sinne nennen schon BRACHMANN und SCHMOLZE die Basiskonzepte ihrer Wissensrepräsentationssprache *„epistemological primitives"* [BRACHMANN & SCHMOLZE]. Die oben eingeführten Metakonstrukte und ihre geforderten Erweiterungen erlauben auf der Grundlage der Kernstrukturen die Darstellung weiterer epistemologischer Differenzierungen. Dabei wird mittels der Metakonstrukte auf bereits repräsentiertes Wissen Bezug genommmen. Insofern lassen sich die Metakonstrukte wenigstens zum Teil semantisch qualifizieren als *„concepts and terms to describe knowledge"* [WETTER & WOODWARD 1990].
In Bezug auf die Wissensrepräsentation ist es nach dem Gesagten erforderlich, einen Repräsentationsformalismus zu finden, der selbst epistemologisch neutral ist. Das bedeutet, daß ein auf allen Stufen der Wissensrepräsentation einheitlicher Formalismus zur Verfügung gestellt wird, der es erlaubt, im Prozeß der Wissensakquisition sukzessiv komplexere Strukturen zu erzeugen und dann explizit mit einer epistemologisch differenzierten Semantik zu versehen. Hierzu wurden im einzelnen Kernobjekte, Kernrelationen, deren Parametrierungen, Metakonstrukte und deren Erweiterungen in einem einheitlichen, graphentheoretischen Formalismus beschrieben. Während das vorliegende Papier das Ziel hat, die Notwendigkeit einer entsprechenden epistemologisch differenzierten Wissensrepräsentation durch Hinweise auf das Domänenwissen selbst zu motivieren, soll in einer späteren Veröffentlichung der Formalismus detailliert beschrieben werden.

4. Diskussion

Der vorgestellte Ansatz zum Umgang mit Problemen epistemologischer Adäquatheit in der Wissensakquisition und formalen Wissensrepräsentation wird abschließend hinsichtlich seiner Konsequenz für die Entwicklung wissensbasierter Systeme betrachtet. Dabei werden Unterschiede und Gemeinsamkeiten zu wichtigen bisherigen Konzepten für die Repräsentation von Domänenwissen diskutiert. Der Beitrag des vorgestellten Ansatzes zu einer Verringerung der Kluft zwischen einer abstrakten Modellierung von Domänenstrukturwissen und der anwendungsbezogenen Domänenmodellierung wird skizziert.

4.1. Andere Ansätze

Die Bemühungen um methodische Unterstützung bei der Wissensakquisition im Rahmen der Entwicklung wissensbasierter Systeme konzentrierten sich lange Zeit auf die Modellierung von Problemlösewissen. Es entstanden Bibliotheken mit abstrakten, wiederverwendbaren Problemlösemodellen (z. B. [BREUKER et al. 1987]). Erst vor einigen Monaten erschien die CommonKADS Library [BREUKER & VAN DE VELDE 1994] mit einer Sammlung von Problemlösekomponenten für verschiedene Bereiche des Knowledge Engineering. Im Bereich des Medical Knowledge Engineering wurde das ST-Modell [RAMONI et al. 1992; STEFANELLI & RAMONI 1992] als ein abstraktes Modell zur Modellierung des Problemlösevorgehens in Diagnostik, Therapieplanung und Patientenmonitoring entwickelt.
Schwierigkeiten bei der Anwendung von solchen abstrakten Problemlösemodellen bei der Modellierung konkreter Anwendungsprobleme zeigten den Bedarf nach methodischer Unterstützung bei der Strukturierung des Domänenwissens, auf dem ein Problemlösemodell operiert. Es entstanden in der Folge Ansätze zum Aufbau von Ontologien (vgl. [PIRLEIN und STUDER 1994]). So werden z. B. in der CommonKADS-

Methodologie zwei Typen von Ontologien unterschieden: *„domain ontologies"* und *„model ontologies"* (z. B. [WIELINGA et al. 1993]). Erstere liefern Beschreibungsmittel für eine bestimmte Domäne, während letztere eine Sammlung generischer Beschreibungsmittel darstellen, die von einem Problemlösemodell dazu verwendet werden, Bezüge zum Domänenwissen herzustellen.

4.2. Verringerung der Kluft zwischen abstraktem Domänenstrukturwissen und anwendungsbezogenem Domänenwissen

Schon die CommonKADS-Methodologie reflektiert eine konzeptuelle Kluft zwischen abstrakten Problemlösemodellen und Domänenontologien. Zur Überbrückung dieser Kluft wurden die *„model ontologies"* als *„meta descriptions, described in terms of the domain ontologies"* [ABEN 1994] eingeführt. Sie sind spezifisch für die Problemlösemodelle.
Im Rahmen der Unterscheidung von *„model ontologies"* und *„domain ontologies"* führt ABEN das Konzept des *„ontology mapping"* [ABEN 1994] ein. Dieses ist der Versuch, Ontologien (gewissermaßen im Sinne einer Koordinatentransformation) aufeinander abzubilden und speziell taskspezifische Ontologien durch eine Basisontologie, die weitgehend domänen- und taskunabhängig ist, auszudrücken. Zwar finden sich Aspekte einer ‚vertikalen Abbildung' von taskspezifischen Ontologien auf eine Basisontologie auch in den Beziehungen zwischen Metakonstrukten und Kernstrukturen wieder, das hier dargestellte Problem der inneren epistemologischen Differenzierung eines Domänenmodells wird vom *„ontology mapping"* jedoch nicht berührt.
Der hier vorgestellte Ansatz zielt auf eine Modellierung von Domänenwissen, die den Aufsatz abstrakter Problemlösemodelle nicht ausschließt, deren leitende Intention es allerdings ist, die Wissensbank auf einen menschlichen ‚Problemlöser' auszurichten und für den menschlichen Umgang mit Wissen transparent zu gestalten. Vor diesem Hintergrund zeigt sich parallel zur bereits erwähnten konzeptuellen Kluft eine andersgeartete Schwierigkeit. Dem menschlichen ‚Problemlöser' ist ein Werkzeug zur interaktiven Strukturierung von Domänenwissen an die Hand zu geben. Damit dieses seinen wechselnden Anforderungen gerecht werden kann, muß es epistemologische Differenzierungen des Domänenwissens vermitteln. Indem die hier aufgeführten Metakonstrukte es erlauben, Domänenwissen epistemologisch differenziert zu organisieren, können sie als Kern eines solchen Werkzeuges fungieren. Außerdem lassen sie sich auch auf ein abstraktes Problemlösemodell ausrichten.

5. Literatur

ABEN, M. (1994). Canonical Functions: CommonKADS Inferences. In BREUKER, J. and VAN DE VELDE, W. (eds.) (1994): CommonKADS Library for Expertise Modelling. Reusable problem solving components, 89–119, IOS Press, Amsterdam.

BRACHMANN, R. J. and SCHMOLZE, J. G. (1985): An overview of the KL-ONE knowledge representation system, Cognitive Science **9**, 171–216.

BREUKER, J. A., WIELINGA, B. J., VAN SOMEREN, M., DE HOOG, R., SCHREIBER, A. T., DE GREEF, P., BREDEWEG, B., WIELEMAKER, J., BILLAULT, J. P., DAVOODI, M. and HAYWARD S. A. (1987): Model Driven Knowledge Acquisition: Interpretation Models. ESPRIT Project P1098 Deliverable D1 (task A1), University of Amsterdam and STL Ltd.

BREUKER, J. and VAN DE VELDE, W. (eds.) (1994): CommonKADS Library for Expertise Modelling. Reusable problem solving components. IOS Press, Amsterdam.

BÜRSNER, S., SPITZER, K., WÜNNEMANN, J., NEBE, J., HAUX, R., DIENER, H.-C. (1994): NeuroN – Rechnerunterstütztes Neurologisches Nachschlagewerk. In PÖPPL, S. J., LIPINSKI, H.-G.,

MANSKY, T. (Hrsg.): Medizinische Informatik: Ein integrierender Teil arztunterstützender Technologien. 38. Jahrestagung der GMDS 1993, S. 186–192, Medizin Verlag München.

BÜRSNER, S. (1995): Eine Kategorisierung von Hypertext-Werkzeugen. In: BÜRSNER, S., FOERTSCH, C. (Hrsg.): Beiträge zum 6. Workshop „Hypertext und Künstliche Intelligenz". FORWISS-Report FR-1995-002, Bayerisches Forschungszentrum für Wissensbasierte Systeme (FORWISS), Erlangen, März 1995, ISSN 0938-0035, S. 1–4.

BÜRSNER, S. and SCHMIDT, G. (1995): Building Views on Conceptual Models for Structuring Domain Knowledge. In Proceedings of KAW '95: Ninth Knowledge Acquisition for Knowledge-Based Systems Workshop, SRDG Publication, Dept. of Computer Science, Univ. of Calgary, Calgary, March 1995. pp. 24-1–24-20.

BÜRSNER, S. and SPRECKELSEN, C. (1995): Internal and external relations and their relevance for the construction of domain ontologies. In GAPPA, U., VOSS, H. (Eds.): Proc. Knowledge Engineering Forum 95. Concepts and Architectures for Reuse. Arbeitspapiere der GMD 903, D-53754 Sankt Augustin. S. 36–39.

CANGUILHEM, G.(1979): Wissenschaftsgeschichte und Epistemologie: gesammelte Aufsätze (übersetzt von BISCHOF, M. und SEILER, W., hrsg. v. LEPENIES, W.), Suhrkamp, Frankfurt a. M.

MOORE, E. G. (1992): Philosophical Studies. Kegan Paul Trench Trubner & Co. London, pp. 276–309.

PIRLEIN, T. and STUDER, R. (1994): KARO: An Integrated Environment for Reusing Ontologies. In STEELS, L., SCHREIBER, G., VAN DE VELDE, W. (Eds.): A Future for Knowledge Acquisition. 8th European Knowledge Acquisition Workshop, EKAW '94, Hoegaarden, Belgium, September 1994, pp. 200–225.

RAMONI, M., STEFANELLI, M., BAROSI, G., MAGNANI, L. (1992): An epistemological framework for medical knowledge based systems. IEEE Transactions on Systems, Man and Cybernetics **22**, 1361–1370.

SPITZER, K. und BÜRSNER, S. (1994): Wissensbasierte Systeme in der Medizin. it+ti-Informationstechnik und Technische Informatik **36** (6); R. Oldenburg Verlag, München, S. 54–59.

STEFANELLI, M. RAMONI, M. (1992): Modeling Medical Knowledge Based Systems. Informatik, Biometrie und Epidemiologie in Medizin und Biologie **23** (4), 215–227.

VAN HEIJST, G., POST, W. und SCHREIBER, A. T. (1994): Knowledge based integration of representation formalisms. In COHN, A. (ed.): 11th European Conference on Artifical Intelligence (ECAI' 94), p. 319–323, John Wiley & Sons, Ltd.

WETTER, T. and WOODWARD, B. (1990): Towards a Theoretical Framework for Knowledge Acquisition. In Proceedings of KAW '90; 5th Knowledge Acquisition for Knowledge-Based Systems Workshop, SRDG Publication, Dept. of Computer Science, Univ. of Calgary, Calgary, November 1990.

WIELINGA, B. J., VAN DE WELDE, W., SCHREIBER, G., AKKERMAN, H. (1993): Towards a Unification of Knowledge Modelling Approaches. In DAVID, J.-M., SIMMONS, R. (eds.): Second Generation Expert Systems, 299–335. Springer Berlin, Heidelberg.

Anschrift der Verfasser: Simone Bürsner, Cord Spreckelsen, Klaus Spitzer, Universität Heidelberg, Institut für Medizinische Biometrie und Informatik, Abt. Medizinische Informatik, Im Neuenheimer Feld 400, D-69120 Heidelberg, email: {sbuersner, cpreckelsen}@krzmail.krz.uni-heidelberg.de

Johannes Nebe, Universität Essen, Neurologische Klinik, Hufelandstr. 55, D-45147 Essen

Informatik, Biometrie und Epidemiologie in Medizin und Biologie **26** (3), 240 – 249, ISSN 0943-5581
© Verlag Eugen Ulmer GmbH & Co., Stuttgart; Gustav Fischer Verlag KG, Stuttgart

Multiple Verwendbarkeit Klinischer Dokumentationen am Beispiel eines wissensbasierten klinischen Arbeitsplatzsystems in der Neurologie

Multiple Usability of Clinical Documentations Demonstrated on a Knowledge-Based Clinical Workstation for Neurology

B. Brigl[1]), P. Ringleb[2]), T. Steiner[2]), G. Mann[3]), F. Leiner[4]), A. Grau[2]), W. Hacke[2]), R. Haux[1])

Kurztitel
Multiple Verwendbarkeit Klinischer Dokumentationen

Short title
Multiple Usability of Clinical Documentations

Schlüsselwörter
Multiple Verwendbarkeit, Wissensbasierte Systeme, Klinische Arbeitsplatzsysteme, Klinische Dokumentation

Keywords
multiple usability, knowledge-based systems, medical workstations, clinical documentation

Zusammenfassung

Wissensbasierte Systeme werden nur dann in der klinischen Routine erfolgreich eingesetzt, wenn sie in umfassendere Anwendungssysteme integriert sind. Der vorliegende Beitrag stellt daher das Modell eines wissensbasierten klinischen Anwendungssystems vor, das Funktionen zur Entscheidungsunterstützung, zur Unterstützung von Routinetätigkeiten und

[1]) Universität Heidelberg, Institut für Medizinische Biometrie und Informatik, Abteilung Medizinische Informatik
[2]) Neurologische Universitätsklinik Heidelberg
[3]) GSF, MEDIS-Institut Oberschleißheim
[4]) Universität-GH Essen, Institut für Medizinische Informatik, Biometrie und Epidemiologie

zur klinischen Forschung integriert. Kern des Modells ist ein medizinisches Merkmalsverzeichnis, das ein anwendungsunabhängiges, kontrolliertes Vokabular zur Verfügung stellt. Am Beispiel eines wissensbasierten klinischen Arbeitsplatzsystems für eine Neurologische Intensivstation wird die Umsetzung dieses Modells aufgezeigt.

Summary

The success of knowledge-based systems in clinical routine depends on their integration in broader application systems. Therefore, this paper describes a model for a knowledge-based clinical workstation which integrates functions for decision support, routine activities and clinical research. Core of that model is a medical data dictionary, which offers an application-independent controlled vocabulary. The realisation of that model will be demonstrated on a knowledge-based clinical workstation for a neurological intensive care unit.

1. Problematik und Zielsetzung

Wissensbasierte Systeme in der Medizin sind nur dann in der klinischen Routine sinnvoll einsetzbar, wenn sie in umfassendere Anwendungssysteme integriert sind (vgl. auch CHUTE, CESNIK und VAN BEMMEL 1994, CIMINO 1994). Nur dann werden sie einen Beitrag für die Verbesserung der Patentenversorgung leisten können. Ein erfolgreiches Beispiel dafür ist das Anwendungssystem HELP mit den darin integrierten wissensbasierten Funktionen, bei denen ein konkreter Nutzen für die Patientenversorgung nachgewiesen werden konnte (HAUG, GARDNER et al. 1994). Die Erfassung von Patientendaten zum Zwecke der Entscheidungsunterstützung darf nicht zu einem wesentlichen Mehraufwand und damit zu einer zusätzlichen Belastung für Personal und Patienten führen. Patientendaten, die im Rahmen der Klinischen Dokumentation (LEINER und HAUX a) aufgezeichnet werden, sollen daher für unterschiedliche Zwecke benutzt werden können (HAUX 1989, LEVY, LAWRENCE 1992). Wir sprechen hierbei von der *multiplen Verwendbarkeit* von Patientendaten. Vorstellbar sind neben der Patientenversorgung (dazu zählt auch die patientenspezifische Entscheidungsunterstützung) die Qualitätssicherung, die klinische Forschung, die Krankenhausberichterstattung und die medizinische Ausbildung. Jeder dieser genannten Zwecke stellt unterschiedliche Anforderungen an eine klinische Dokumentation.
Ein Ziel des MEDWIS-Projektes A4 ist es daher, ein Modell eines klinischen Arbeitsplatzsystems zu erarbeiten, das die multiple Verwendbarkeit von Patientendaten gewährleistet. Anhand eines darauf basierenden Anwendungssystems (DAWIN II) sollen die Möglichkeiten und Grenzen der multiplen Verwendbarkeit von Patientendaten geprüft werden. Beispielhaft werden die Bereiche Patientenversorgung und Klinische Forschung betrachtet. Die Nutzung der dokumentierten Patientendaten für die patientenspezifische Entscheidungsunterstützung ist dabei von besonderem Interesse.
Die vorliegende Arbeit stellt das Modell eines wissensbasierten klinischen Arbeitsplatzsystems vor und berichtet anhand von Beispielen aus einer neurologischen Intensivstation über die bisher vorliegenden Erfahrungen.

2. Multiple Verwendbarkeit von Patientendaten

Als Beispiele für die multiple Verwendbarkeit von Patientendaten betrachten wir im Rahmen der Patientenversorgung die Berichtschreibung und die patientenspezifische Entscheidungsunterstützung sowie im Rahmen der Klinischen Forschung den Aufbau

und die Auswertung klinischer Register. Im folgenden sollen einige wesentliche Aspekte der multiplen Verwendbarkeit von Patientendaten kurz erläutert werden (LEINER und HAUX b).

Datenqualität

Die Verwendung von Daten für unterschiedliche Zwecke stellt bestimmte Anforderungen an die Qualität einer Klinischen Dokumentation. Typische Qualitätsmaße sind die folgenden Kriterien:

Reliabilität: Maß für die Unabhängigkeit eines Beobachtungsergebnisses von externen Störgrößen wie beispielsweise von der Person des Beobachters.

Validität: Maß für den Grad der Übereinstimmung zwischen dem erhobenen Datum und dem realen Umstand, den es beschreiben soll.

Vollständigkeit: Anteil der erhebbaren Merkmale, die tatsächlich aufgezeichnet wurden. Der notwendige Grad an Vollständigkeit ist abhängig von den Fragestellungen, die anhand der Dokumentation beantwortet werden sollen.

Konsistenz: Widerspruchsfreiheit der einzelnen Daten untereinander. Inkonsistenzen weisen auf falsche oder unvollständige Daten hin, und ihre Aufdeckung fördert daher die Validität und Vollständigkeit der Dokumentation. Die Aufdeckung von Widersprüchen geschieht meist über die Formulierung heuristischer Konsistenzbedingungen. Die Konstruktion formaler semantischer Modelle des Wirklichkeitsausschnittes, an deren Strukturen die Konsistenz überprüft werden kann, bringt viele Schwierigkeiten mit sich.

Da bereits aufgezeichnete Daten automatisch in den Arztbrief integriert werden sollen, schlägt sich die Qualität der Dokumentation direkt in der Qualität des Arztbriefes nieder. Für Verfahren der Entscheidungsunterstützung müssen bestimmte Daten aufgezeichnet sein. Reliabilität und Validität erhöhen die Richtigkeit der Entscheidungsvorschläge. Im Sinne einer kritisierenden Funktion kann Entscheidungsunterstützung auch für die Überprüfung der Konsistenz einer Dokumentation eingesetzt werden. Die Auswertung Klinischer Register stellt besonders hohe Anforderungen an die Datenqualität. Die Vollständigkeit der Dokumentation für definierte Auswertungskollektive muß laufend überprüft werden, weil sonst die Gefahr statistischer Verzerrungen durch selektive Merkmalserhebung besteht (HAUX 1989).

Standardisierte Dokumentation

Die multiple Verwendbarkeit von Patientendaten setzt im allgemeinen eine standardisierte Dokumentation bis auf die Ebene der Merkmale und der Merkmalsausprägungen voraus. Der Dokumentation hat ein kontrolliertes Vokabular zugrundezuliegen. Eine standardisierte Dokumentation ist insbesondere dann notwendig, wenn eine Beobachtungsgleichheit verlangt wird (beispielsweise für statistische Auswertungen), wenn die Vollständigkeit der Dokumentation möglichst hoch sein soll und wenn Daten zwischen verschiedenen Anwendungssystemen ausgetauscht werden sollen.

Um im Rahmen der Auswertung Klinischer Register die Vergleichbarkeit von Beobachtungen zu gewährleisten, ist eine standardisierte Dokumentation nötig. Die gängigen statistischen Verfahren setzen metrische Merkmale mit definiertem Bezugssystem oder kategoriale Merkmale mit klassierten Ausprägungen voraus. Auch entscheidungsunterstützende Funktionen bauen auf einer standardisierten Dokumentation auf. Das zugrundeliegende Modell des Wissensgebietes wird meist über ein kontrolliertes Vokabular definiert. Allerdings ist zu beachten, daß durch eine vollständige Standardisierung unter Umständen wichtige Informationen verloren gehen können und daß manche Sachverhalte nur schwer standardisierbar sind. Eine freitextliche Dokumentation von Daten sollte daher immer vorgesehen sein, auch wenn diese Daten dann lediglich für die Berichtschreibung zur Verfügung stehen.

3. Konzeptuelles Modell eines wissensbasierten klinischen Arbeitsplatzsystems

Merkmalsverzeichnis
Den Kern des Modells eines wissensbasierten klinischen Arbeitsplatzsystems bildet ein medizinisches Merkmalsverzeichnis, mit dem das für den Anwendungsbereich relevante Vokabular definiert wird. In diesem werden die zu dokumentierenden Merkmale und deren Merkmalsausprägungen festgelegt. Dabei kann es sich um metrische, kategoriale oder freitextliche Merkmale handeln. Den Ausprägungen kategorialer Merkmale sind sogenannte Semantikcodes zugeordnet – wie beispielsweise ‚Normalbefund', ‚pathologisch', ‚nicht erhoben', ‚nicht prüfbar' –, die die strukturelle Integrität der Daten (vgl. MANN 1994) gewährleisten und später bei der Datenerfassung und Berichtschreibung verwendet werden. Jedes Merkmal besitzt als eine mögliche Ausprägung ‚nicht erhoben', womit die Vollständigkeit der Dokumentation kontrollierbar wird. Die Merkmale erhalten eine hierarchische Struktur.

Schema einer Patientendatenbank
Aus dem Merkmalsverzeichnis läßt sich automatisch das Schema der Patientendatenbank ableiten. Jedes Datenbankschema, das über die Metadatenbank erzeugt wird, ist das Schema einer Assoziierungsdatenbank im mRM/T (vgl. MANN 1989, WINTER und HAUX 1994) und erfüllt damit grundsätzlich einige wichtige Aspekte der semantischen Integrität.

Datenerfassung
Die Datenerfassung erfolgt über hierarchisch strukturierte Erhebungsbögen, die über das Merkmalsverzeichnis generiert werden können. Diese hierarchische Struktur und die Verwendung der Semantikcodes verringern den Aufwand für die Erfassung. Ist einer gewählten Merkmalsausprägung beispielsweise der Semantikcode ‚Normalbefund' zugeordnet, so erhalten alle hierarchisch darunterliegenden Merkmale ebenfalls die Ausprägung mit diesem Semantikcode zugeordnet. Damit wird die Vollständigkeit und die Konsistenz der Daten erhöht. Bei den kategorialen Merkmalen ist aus n Ausprägungsmöglichkeiten genau eine auszuwählen (exklusive Auswahl). Zu untergeordneten Erhebungsbögen gelangt man nur in Abhängigkeit der gewählten Merkmalsausprägung. Viele der Erhebungsbögen müssen bei der Datenerfassung daher nicht betreten werden. Die Datenerfassung muß Steuerungsmechanismen aufweisen, die die Vollständigkeit der Klinischen Dokumentation bezüglich definierter Auswertungskollektive gewährleisten.

Befundschreibung
Das Merkmalsverzeichnis wird auch für die Erstellung von Berichtsformularen (Untersuchungsbefunde, Arztbriefe) verwendet. Jedem Merkmal-Merkmalsausprägungs-Paar wird dabei eindeutig ein Textbaustein zugeordnet. Aufgrund der Semantikcodes kann festgelegt werden, ob bei Vorliegen einer bestimmten Merkmalsausprägung der zugehörige Textbaustein ausgegeben werden soll oder nicht (im allgemeinen werden nur pathologische Merkmale ausgegeben). In Abhängigkeit der erfaßten Merkmale werden Berichte dynamisch generiert.

Aufbau und Auswertung Klinischer Register
Das Merkmalsverzeichnis dient der Festlegung von Ein-/Ausschlußkriterien für die Definition von Auswertungskollektiven. Zudem können ‚Muß'-Merkmale angegeben werden, die für bestimmte Auswertungskollektive zu erheben sind. Diese werden deklarativ in einer Wissensbank abgelegt. Zusätzlich zu erhebende Merkmale müssen in

das Merkmalsverzeichnis aufgenommen werden. Sie sind in das Erhebungsbogenschema zu integrieren.
Für das Finden von Patienten mit bestimmten Eigenschaften und die Ausgabe bestimmter Merkmale zur Weiterverarbeitung mit einem Statistischen Auswertungssystem werden ebenfalls Bedingungen über das Merkmalsverzeichnis definiert.

Patientenspezifische Entscheidungsunterstützung
Das Merkmalsverzeichnis unterstützt die Definition von Objekten einer Wissensbank, indem es ein kontrolliertes medizinisches Vokabular zur Verfügung stellt. Als Schnittstelle zwischen Klinischer Dokumentation und patientenspezifischer Entscheidungsunterstützung sichert es die referentielle Integrität zwischen Datenbank und Wissensbank.
In der Wissensbank können beliebige medizinische Objekte abgelegt werden. Diese müssen über das Merkmalsverzeichnis definiert werden. Jedem Objekt ist daher ein (beliebig komplexer) logischer Ausdruck zugeordnet, in dem Merkmal-Merkmalsausprägungs-Paare über die Operatoren ‚und', ‚oder' miteinander verknüpft sind. Um die semantische Integrität der Wissensbank zu gewährleisten, sind Integritätsbedingungen festzulegen, die die Möglichkeiten der Verknüpfung von Merkmal-Merkmalsausprägungs-Paaren beschränkt. Beispielsweise kann ein Krankheitszeichen nur über Merkmale definiert sein, die selbst Krankheitszeichen oder Eigenschaften eines Krankheitszeichens sind. Über diese logischen Ausdrücke können Merkmale, die bei der Datenerfassung aufgezeichnet wurden und in der Patientendatenbank gespeichert sind, direkt abgebildet werden auf Objekte der Wissensbank. Es können Wissensmodelle für unterschiedliche Fragestellungen mit Hilfe des gleichen Vokabulares definiert werden. Dies eröffnet Möglichkeiten der Integration verschiedener Wissensbanken. Verfahren der patientenspezifischen Entscheidungsunterstützung können beispielsweise in die Berichtschreibung oder in die Datenerfassung integriert sein.

Werkzeuge
Für die genannten Verfahren stehen spezifische Werkzeuge sowohl auf der Systemverwaltungsseite als auch auf der Benutzungsseite zur Verfügung. Ein Erhebungsbogeneditor erleichtert die Strukturierung der Erhebungsbögen, ein Formulareditor dient der Definition von Berichtsformularen, ein Registereditor dient der Definition von Auswertungskollektiven. Diese sind Teil eines Metadatenverwaltungssystems. Wissensakquisitionswerkzeuge als Teil des Wissensverwaltungssystems dienen dem Aufbau und der Pflege von (unterschiedlichen) Wissensbanken. Innerhalb des Benutzungssystems gibt es Funktionen für die Patientendatenverwaltung, die Datenerhebung, die Berichtschreibung, die Auswertung Klinischer Register und die patientenspezifische Entscheidungsunterstützung.
Die Strukturen der Erhebungsbogenhierarchie und der Berichtformulare sind zusammen mit dem Merkmalsverzeichnis in einer Metadatenbank gespeichert.

Integration mit anderen informationsverarbeitenden Verfahren
Patientenspezifische Entscheidungsunterstützung erfordert eine umfassende Dokumentation der Behandlung. Durch die Einrichtung einer Kommunikationsschnittstelle zum Kommunikationssystem eines Krankenhausinformationssystems können Daten, die bereits in anderen Anwendungssystemen erfaßt wurden, in die Patientendatenbank des wissensbasierten klinischen Arbeitsplatzsystems integriert werden und stehen somit für entscheidungsunterstützende Funktionen zur Verfügung. Eine multiple Verwendbarkeit von Patientendaten ist somit auch anwendungssystemübergreifend gewährleistet.
Einen Gesamtüberblick über das Modell eines wissensbasierten klinischen Arbeitsplatzsystems gibt Abbildung 1.

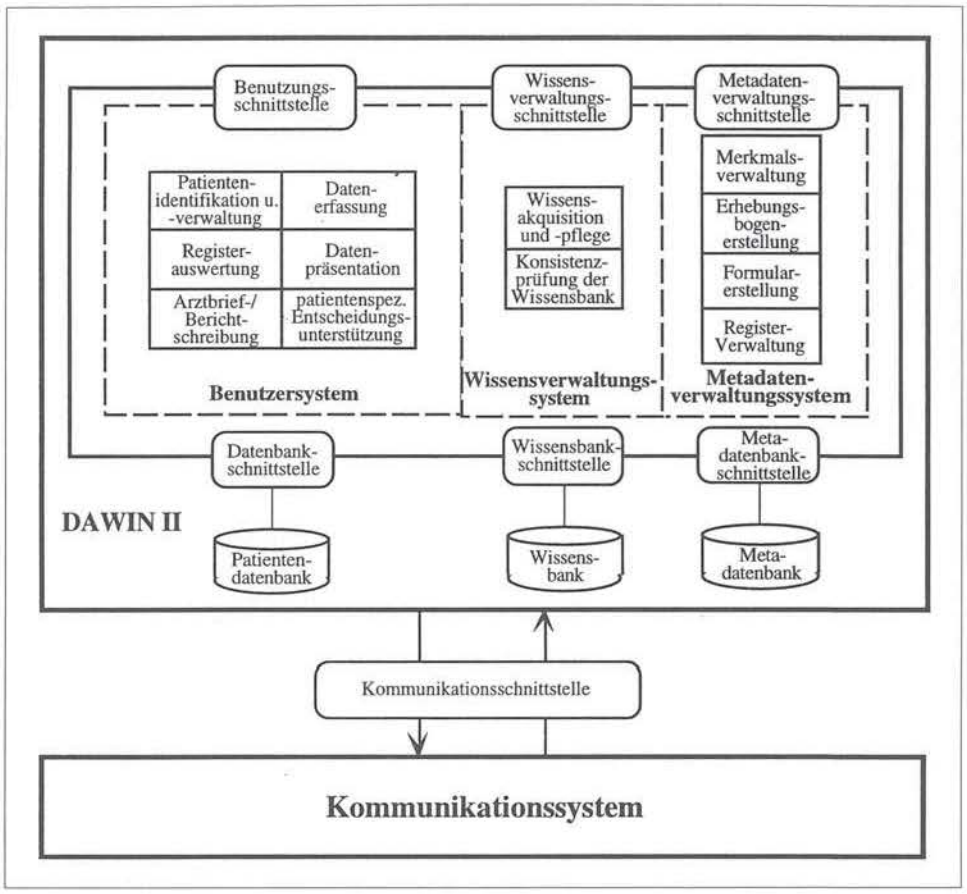


Abbildung 1: Modell eines wissensbasierten klinischen Arbeitsplatzsystems.

4. Beispiel: Neurologische Intensivstation

In der Neurologischen Intensivstation des Universitätsklinikums Heidelberg wird der klinische Befund standardisiert erhoben. Eine weitere Standardisierung, beispielsweise Anamnese, Medikation etc. ist geplant. Das Merkmalsverzeichnis dieser Klinischen Dokumentation umfaßt ca. 500 Merkmale. Ein Ausschnitt des Merkmalsverzeichnisses findet sich in Abbildung 2.
Die in der Patientendatenbank dokumentierten Daten sollen für die Befund- und Arztbriefschreibung und die patientenspezifische Entscheidungsunterstützung verwendet werden.

Datenerfassung und Arztbriefschreibung
Aus dem Merkmalsverzeichnis werden hierarchisch strukturierte Erhebungsbögen für die Datenerfassung abgeleitet (Abbildung 3). Aus den erhobenen Daten können Befundberichte und Arztbriefe generiert werden. Dafür sind den Merkmal-Merkmalsausprägungs-Paaren statische Textbausteine hinterlegt (Abbildung 4).

Merkmale	Ausprägungen
Bewußtsein und psychischer Befund	Normal, Pos. Befund, N.e., N.pr.
Blickmotorik	Normal, Pos. Befund, N.e., N.pr.
Skew deviation	Links tiefer als rechts, Rechts tiefer als links, Nein, N. e.
Pathologischer Nystagmus	Ja, Nein, N. e.
Blickwendung	Nach rechts, Unklar, Keine, Nach links, N. e.
Vertikale Blickparese	Nach unten, Nach oben, Nein, N. e., N. p.
Horizontale Blickparese	Nach links, Nach rechts, Nach beiden Seiten, Nein, N. e., N. p.
Konvergenzparese	Ja, Nein, N. e., N. p.
Lichtreaktion und Cornealreflex	Erhoben, N.e.
Nackensteifigkeit	Normal, Pos. Befund, N.e., N.pr.

Abbildung 2: Ausschnitt aus dem Merkmalsverzeichnis. (Bedeutung der Abkürzungen: N. e. — Nicht erhoben, N. pr. — Nicht prüfbar)

Patientenspezifische Entscheidungsunterstützung

Für die patientenspezifische Entscheidungsunterstützung sollen zwei Funktionen angeboten werden:

(1) Überprüfung der Qualität der Befundung

Es sollen Patientendaten und die aus ihnen abgeleiteten Befunde qualitativ und quantitativ beurteilt werden. Insbesondere soll analysiert werden, ob Widersprüche zwischen der Diagnose und den erhobenen Untersuchungsergebnissen bestehen, ob die vorliegenden Befunde für eine Diagnosefindung ausreichend sind, ob weitere Diagnosen denkbar sind und ob die gewählte Therapie korrekt war. Idealerweise wird die Qualität der Daten zum Zeitpunkt der Arztbriefschreibung überprüft, da dann alle notwendigen Informationen vorliegen. Die Verwendbarkeit dieser Funktion zu einem früheren Zeitpunkt und mit unvollständigen Daten ist im Hinblick auf eine Entscheidungsunterstützung für einen konkreten Patienten wünschenswert.

Für diese Funktion ist der Aufbau einer umfassenden neurologischen Wissensbank notwendig, die die Objekttypen ‚Anamnestische Befunde', ‚Klinische Befunde', ‚Apparative Befunde', ‚Ätiologie', ‚Diagnose' und ‚Therapie' und deren Beziehungen umfaßt. Das Wissensgebiet wird daher zunächst auf das Krankheitsbild ‚Schlaganfall' eingeschränkt.

Ein Ausschnitt eines Wissenserhebungsbogens, in dem die Definition eines Objektes der Wissensbank über Merkmale des Merkmalsverzeichnisses gezeigt wird, findet sich in Abbildung 4. Zudem sind geeignete Inferenzmodelle zu entwickeln.

(2) Berechnung neurologischer Scores

Neben diesen retrospektiven kritisierenden Funktionen sollen auch entscheidungsunterstützende Funktionen zur Verfügung stehen, die aktuell zum Zeitpunkt der Datenerhebung angewandt werden können. Hierzu zählt die Berechnung neurologischer Score-Werte, z. B. die NIH-Stroke-Scale oder die Glasgow-Coma-Scale. Die NIH-Stroke-Scale beurteilt die Schwere eines Schlaganfalls und bezieht beispielsweise Merkmale, die be-

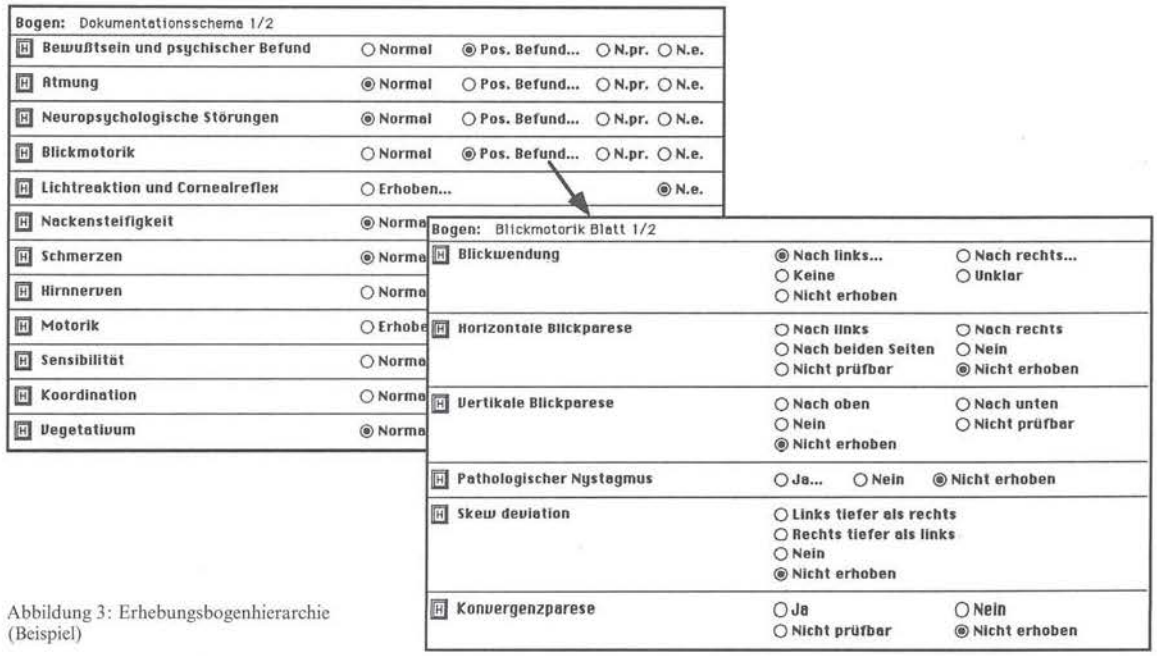


Abbildung 3: Erhebungsbogenhierarchie (Beispiel)

Abbildung 4: Definition von Objekten der Wissensbank (Beispiel)

Objektname: **Visueller Neglect links**

Objekttyp: Krankheitszeichen

Bezug zum Merkmalsverzeichnis:

Art des Neglects = visuell und **Seite des Neglects = links**

Assoziierte Diagnosen: ...

Lokalisation: ...

treffend des Bewußtseins (bsp. Vigilanzgrad) und der Motorik von Armen und Beinen mit ein.

5. Stand der Realisierung

Auf der Grundlage des vorgestellten Modells für ein wissensbasiertes klinisches Arbeitsplatzsystem wurde das Daten- und Wissensbanksystem DAWIN II realisiert. Hardware-Plattform sind Apple Macintosh Rechner (Macintosh II fx, Macintosh Quadra, Powerbook 540). Als Entwicklungswerkzeuge wurden das relationale Datenbankverwaltungssystem ORACLE und die objektorientierte Entwicklungsumgebung SMALLTALK/V verwendet. Realisiert ist das Metadatenverwaltungssystem mit den Funktionen Merkmalsverwaltung, Erhebungsbogenerstellung und Formularerstellung und das Benutzersystem für die Datenerfassung, die Registerauswertung und die Arztbrief-/Berichtschreibung. Die automatische Berechnung neurologischer Scores ist in die Datenerfassung integriert. Zudem existiert ein Wissensakquisitionswerkzeug zum Aufbau einer Wissensbank für die Beurteilung der Konsistenz einer Klinischen Dokumentation. Bis Ende des Jahres 1995 werden die genannten noch fehlenden Funktionen prototypisch realisiert sein. Eine Kommunikationsschnittstelle zum Heidelberger Kommunikationssystem (HeiKo) ist für Macintosh-Rechner realisiert und in DAWIN II eingebunden. Damit können derzeit Patientenstammdaten aus der Zentralen Patientendatenbank übernommen werden.
Seit Januar 1995 wird DAWIN II zur Dokumentation in der Intensivstation der Neurologischen Universitätsklinik Heidelberg in Routine eingesetzt. Es werden die Eingangs- und Entlassungsuntersuchungen aller Patienten aufgezeichnet. Die mit DAWIN II generierten Arztbriefe werden regelmäßig verschickt.

6. Diskussion

Das hier vorgestellte Modell eines wissensbasierten klinischen Arbeitsplatzsystems gewährleistet, daß einmal aufgezeichnete Patientendaten für unterschiedliche Zwecke verwendet werden können. Der Mehraufwand für die Datenerfassung, der durch den Einsatz entscheidungsunterstützender Funktionen entsteht, reduziert sich daher auf die Dokumentation zusätzlicher Merkmale, die für die Lösung bestimmter Problembereiche notwendig sind. Damit ist eine wichtige Voraussetzung geschaffen, die Akzeptanz wissensbasierter Systeme in der klinischen Routine zu verbessern.
Bei der Einführung von Anwendungssystemen steht immer der konkrete Nutzen für die Patientenversorgung im Vordergrund. Anwender fragen nach einer spürbaren Erleichterung bei Routinearbeiten, die der Behandlung der Patienten zugute kommt. Auf

Aspekte der Benutzerfreundlichkeit und der Effizienz ist daher besonders zu achten. Erst wenn ein Anwendungssystem in dieser Hinsicht akzeptiert ist, wenn beispielsweise eine automatische Generierung von Arztbriefen die konventionelle Arztbriefschreibung abgelöst hat, kann die Einführung entscheidungsunterstützender Funktionen Erfolg haben. Die Integration wissensbasierter klinischer Arbeitsplatzsysteme in Krankenhausinformationssysteme ist insofern von besonderer Bedeutung, als der Austausch von Patientendaten mit anderen Anwendungssystemen eine beträchtliche Arbeitserleichterung mit sich bringen kann. Aspekte des strategischen Managements von Krankenhausinformationssystemen (vgl. WINTER 1994) sind daher bereits bei der Systemspezifikation zu berücksichtigen. Vorgaben in Form eines Rahmenplans für ein Krankenhausinformationssystem, die die Art der Hardware, Kommunikationsschnittstellen, Art der Datenhaltung etc. beinhalten können, sollten mit einbezogen werden, um die Interoperabilität zu verbessern und den Aufwand für die Systempflege möglichst gering zu halten.

Literatur

CHUTE, C. G., CESNIK, B., VAN BEMMEL, J. H. (1994): Medical data and knowledge management by integrated medical workstations: summary and recommendations. International Journal of Bio-Medical Computing **34**, 175–183.

CIMINO, J. J. (1994): Data storage and knowledge representation for clinical workstations. International Journal of Bio-Medical Computing **34**, 185–194.

HAUG, P. J., GARDNER, R. M., TATE, K. E., EVANS, R. S., EAST, T. D., KUPERMAN, G., PRYOR, T. A., HUFF, S. M., WARNER, H. R. (1994): Decision Support in Medicine: Examples from the HELP System. Comp. Biomed. Res. **27**, 396–418.

HAUX, R. (1989): Knowledge-Based Decision Support for Diagnosis and Therapy: On the Multiple Usability of Patient Data. Methods of Information in Medicine **28**, 69–77.

LEINER, F., HAUX, R. (a): Systematic Planning of Clinical Documentation – Part I: The Documentation Protocol. Zur Veröffentlichung eingereicht.

LEINER, F., HAUX, R. (b): Systematic Planning of Clinical Documentation – Part II: Design Aspects. Zur Veröffentlichung eingereicht.

LEVY, A. H., LAWRENCE, D. P. (1992): Data Acquisition and the Computer-Based Patient Record. In: BALL, M. H., COLLEN, M. F. (Hrsg.) (1992): Aspects of the Computer-Based Patient Record. Springer. New York.

MANN, G. (1994): Integration wissensbasierter Systeme in der Medizin am Beispiel eines Daten- und Wissensbanksystems in der Neurologie. Bericht Nr. 1/94 der Abteilung Medizinische Informatik, Universität Heidelberg, 1994.

MANN, G. (1989): Studienplanung und Datenbankentwurf für ein klinisches Register bei cerebralen Durchblutungsstörungen. Bericht Nr. 4/1989 des Institutes für Medizinische Informationsverarbeitung der Universität Tübingen.

WINTER, A. (1994): Beschreibung, Bewertung und Planung heterogener Krankenhausinformationssysteme. Habilitationsschrift, Fakultät für Theoretische Medizin der Universität Heidelberg.

WINTER, A., HAUX, R. (1994): A Modified RM/T Data Model for a Universal Relation View. Bericht Nr. 4/94 der Abteilung Medizinische Informatik, Universität Heidelberg, 1994 und zur Veröffentlichung eingereicht.

Anschrift der Verfasser: B. Brigl, R. Haux, Universität Heidelberg, Institut für Medizinische Biometrie und Informatik, Abteilung Medizinische Informatik, Im Neuenheimer Feld 400, 69120 Heidelberg

P. Ringleb, T. Steiner, A. Grau, W. Hacke, Neurologische Universitätsklinik Heidelberg, Im Neuenheimer Feld 400, 69120 Heidelberg

G. Mann, GSF, MEDIS-Institut, Postfach 1129, 85758 Oberschleißheim

F. Leiner, Universität-GH Essen, Institut für Medizinische Informatik, Biometrie und Epidemiologie, Hufelandstr. 55, 45122 Essen

Informatik, Biometrie und Epidemiologie in Medizin und Biologie **26** (3), 250 – 261, ISSN 0943-5581
© Verlag Eugen Ulmer GmbH & Co., Stuttgart; Gustav Fischer Verlag KG, Stuttgart

MDD-GIPHARM: Design and Realization of a Medical Data Dictionary for Decision Support Systems in Drug Therapy

MDD-GIPHARM: Design und Realisierung eines Medical Data Dictionaries für Wissensverarbeitende Systeme in der Arzneitherapie

H. U. Prokosch[1]), T. Bürkle[1]), J. Storch[1]), A. Strunz[1]), M. Müller[1]), J. Dudeck[1]), B. Dirks[2]), F. Keller[3])

Kurztitel
Medical Data Dictionary für den Arzneimittelbereich

Short title
Medical Data Dictionary for Drug Therapy Support

Stichworte
Medizinisches Data Dictionary, Wissensverarbeitung in der Arzneitherapie, Medical Data Dictionary, Drug Therapy Support

Summary

Medical data dictionaries (MDDs) play an important role in the development of medical information systems (MIS) and have been applied to support the integration of medical decision support capabilities into MIS environments by linking medical knowledge bases with clinical patient databases. MDD-GIPHARM has been designed for exactly this reason: to act as an integrator between different drug information sources, a drug therapy knowledge base and the WING patient database at Gießen University Hospital. In this context it shall efficiently support the quality assurance measures in drug therapy which shall be established within the MEDWIS project WAIN. This paper describes the design and implementation as well as the current status of MDD-GIPHARM.

[1]) Institut für Medizinische Informatik, Justus-Liebig-Universität Gießen
[2]) Univ. Klinik für Anästhesiologie, Klinikum der Universität Ulm
[3]) Medizinische Klinik II, Eselsberg, Universitätsklinikum Ulm

Zusammenfassung

Medizinische Data Dictionaries (MDDs) haben eine grundlegende Bedeutung bei der Entwicklung medizinischer Informationssysteme (MIS) und wurden vielfach eingesetzt, um die Einbindung wissensverarbeitender Funktionen in MIS zu unterstützen. Das MDD verknüpft in diesem Zusammenhang die Inhalte medizinischer Wissensbasen mit den Daten in klinischen Patientendatenbanken. MDD-GIPHARM wurde mit eben dieser Zielsetzung konzipiert und entwickelt: es agiert als Integrator zwischen verschiedenen Arzneimittelinformationssystemen, einer Wissensbasis zur Unterstützung der Arzneimitteltherapie und der Patientendatenbank des Krankenhaus-Informationssystems WING am Gießener Universitätsklinikum. In diesem Umfeld unterstützt es die Einführung qualitätssichernder Maßnahmen im Bereich der Arzneimitteltherapie, der Hauptzielsetzung des MEDWIS Projekts WAIN. Die vorliegende Publikation beschreibt Design und Realisierung sowie den aktuellen Status von MDD-GIPHARM.

1. Introduction

The importance of medical data dictionaries (MDDs) for the design and implementation of medical information systems (MIS) has been mentioned by several authors (e.g. LINNARSSON and WIGERTZ, 1989; MICHEL et al., 1989; CIMINO et al., 1989 and HUFF et al., 1987). ANDERSON (1986) has emphasized, that data dictionaries are a way forward to write meaning and terminology into MIS. Especially, when decision support capabilities shall be integrated into MIS and medical knowledge bases need to be linked with clinical databases, a robust and flexible MDD is the "heart and soul" for the system's success (HUFF et al., 1987; LINNARSSON and WIGERTZ, 1989). Several of the early computer-based medical record systems have already been designed around extensive, well structured data dictionaries. In a comprehensive review of five different medical record systems (STOR, HELP, RMRS, TMR and COSTAR), MCDONALD (1988) has concluded that a data dictionary approach, which all of these systems have in common, is the most adequate way for the development of electronic medical record systems. In these early MIS the role of the MDD was mainly restricted to a sequential list with all the terms and concepts applied in the respective MIS environment, linking a textual description and an internal code. For computational efficiency some of the MDDs have additionally supported the hierarchical organization of the included terms and the definition of synonyms. However, the ever increasing need to integrate information coming from different sources, and often available in incompatible formats, has forced a reevaluation of the MDD design requirements and assumptions. Consequently, HUFF et al. (1987), LINNARSSON and WIGERTZ (1989), CIMINO et al. (1989) and ROCHA et al. (1994) have presented more recent work on the design characteristics of medical data dictionaries.
In a definition given by PROKOSCH et al. (1992) an MDD is characterized by three typical features:

- it's capability to establish a controlled vocabulary as a central thesaurus for the relevant application area,
- it's capability to represent causal relationships existing between the different objects of the MDD in form of a semantic network, and
- it's capability to provide efficient mappings between it's own vocabulary and external vocabularies as well as standardized international nomenclatures and knowledge sources.

In a more recent work of HUFF and CIMINO (1995) an MDD has been decomposed into three complementary pieces: a vocabulary, an information model (to describe the structure of clinical data) and a knowledge base (to define hierarchies, semantic relationships and more complex rule-based relationships). An important issue, when designing an MDD is the scope of its application in the respective MIS which it shall support and the type of functionality which shall be supported. Thus the above features, as well as the decomposition, introduced by HUFF and CIMINO (1995) shall not be considered as mandatory, but as possible issues, which may be supported by an MDD in a particular setting.

At Gießen University Hospital research on the area of medical data dictionary design has been initiated during the test implementation of the HELP system (PROKOSCH et al., 1991) and has led to the development of the Gießen Medical Data Dictionary (GMDD) (MICHEL et al., 1989). Even though some of the restrictions detected within PTXT have been eliminated with the GMDD development (e.g. polyhierarchical relationships, domain management features and the linkage between the medical data dictionary and the structural data dictionary of the relational patient database have been supported within the GMDD), its capabilities to efficiently represent semantic relationships and mappings to external information sources were still limited. Thus, a more flexible design (already roughly described by PROKOSCH et al. (1992)) has been investigated. A first application area was the development of a rheumatology information system (AMIRI, 1993; KRAUSE, 1993 and PROKOSCH et al. 1995). In this new development the semantic network representation capability has been efficiently supported.

MDD-GIPHARM is the MDD developed to serve as one core component of the MEDWIS project WAIN (a pharmacological knowledge base focused on the areas of intensive care and nephrology). Together with a second MDD-component (realized at the department of informatics in psychotherapy of Ulm University), which serves as a glossary module to establish a consistent glossary for the WAIN project, it forms the medical data dictionary of the WAIN project. In the following however, we only focus on describing the conceptual design and realization of the medical data dictionary component realized at Gießen University Hospital.

The MDD-GIPHARM design and development is based on the experiences of all the previous MDD research performed at the Department of Medical Informatics, Gießen University Hospital. In addition to the features already supported in our earlier MDD developments, MDD-GIPHARM has been designed to act as a very flexible tool, which allows to link various heterogeneous information sources on drug/pharmaceutical substances and integrate them within one user-friendly application environment.

2. MDD-GIPHARM Database Design

The core database of MDD-GIPHARM comprises three files which are extensively referencing each other (figure 1). The OBJECT table is applied to establish a master index for the controlled vocabulary and shall contain one record for every object which shall be applied within the WAIN environment. The RELATION table contains one entry for every type of relationship which may be defined between any two terms defined in the OBJECT table. Finally the RELATIONSHIP table is used to actually describe the semantic relationships which may exist between two objects of MDD-GIPHARM. Semantic relationships are described in terms of quadrupels, where the first and the third item represent the two objects between which a particular relationship exists. The second item represents the respective type of relationship and the fourth item is applied to restrict the validity of the described relationship to a particular context if necessary.

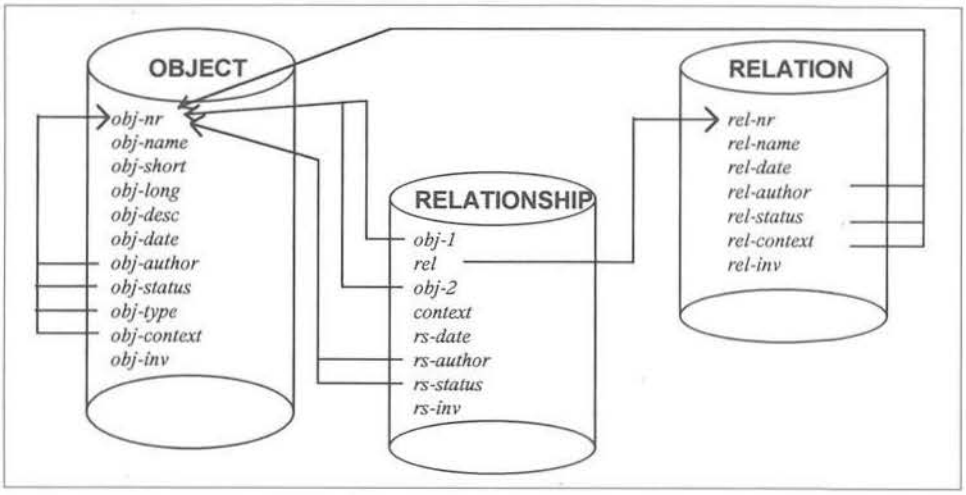


Figure 1: The core tables of MDD-GIPHARM and their linkages

2.1. Object Definition

One entry in the OBJECT table is uniquely identified by a numeric object identification number *(obj-nr)* or by its unique alphanumeric object name *(obj-name)*. The table elements *obj-short*, *obj-long* and *obj-descr* are applied to express the meaning of an object with different length of text, or to document the actual usage of an object within MDD-GIPHARM. The fields *obj-date*, *obj-author*, *obj-status* and *obj-inv* have the purpose to keep track of an object's history (when was it first asserted, who asserted it, what is its status?) and to support the MDD maintenance. The currently defined legal values for the object status comprise test, routine and invalid. Deleting an object from the MDD is explicitly not allowed, since old entries in a patient database or in MDD-GIPHARM (see the section on self-reference and domain management) may have already used this code and would consequently loose their reference when the respective object would be deleted. If usage of an object shall not be allowed any longer from a certain point of time, its status needs to be changed from routine to invalid and the field *obj-inv* shall then contain the date when the object became invalid. Should a more detailed need for history tracking arise in MDD-GIPHARM's future usage, this could be solved with an additional history table (compare for example the VOSER Object history table (Rocha et al., 1994)). Finally the fields *obj-type* and *obj-context* are applied to express some more semantic information about the respective MDD-GIPHARM object. The domain of the field *obj-type* currently comprises attributes, attribute-entries, concepts and semantic types, but may be dynamically expanded when the need arises. The field *obj-context* documents the original context from which a particular object has been extracted in order to include it in MDD-GIPHARM. Drug names for example have either been imported from the german drug information system "Rote Liste" or from our hospital's internal formulary. Other more general objects, e.g. the global terms "drug substance" or "mdd-author" are associated with the context MDD-GIPHARM, in order to depict their general application within MDD-GIPHARM which is not linked with any external source.

2.2. Relations and Relationships

Types of relationships are represented analogue to objects in the table RELATION. They are identified by either their unique relation identification number *(rel-nr)* or their unique name *(rel-name)*. *Rel-date*, *rel-status*, *rel-author* and *rel-inv* are again applied for maintenance purposes and history tracking. *Rel-context* may eventually document when a particular type of relationship has been extrated from an external source. Until now however, the only context applied is MDD-GIPHARM itself. Types of relationships which are currently included in MDD-GIPHARM are *is a*, *contains*, *belongs to* and *has attribute-set*. Additional types of relationships (e.g. *is indicated by*, *is contraindicated by*, *has adverse drug reaction*) may be added in a later project phase whenever the practical need arises.

The actual semantic network of MDD-GIPHARM is represented by the first three elements of the RELATIONSHIP table, which are *obj-1*, *rel* and *obj-2*. These are three numeric fields which link one entry in the RELATIONSHIP table with two entries in the OBJECT table (via the *obj-nr*) and one entry in the RELATION table (via the *rel-nr*). Here the field *context* describes the context in which the respective relationship between two objects is valid (in contrary to the OBJECT and RELATION tables where it refers to the original source of an item). The fields *rs-date*, *rs-status*, *rs-author* and *rs-inv* again serve maintenance purposes and history tracking.

The fields *obj-status*, *obj-author*, *obj-type*, *obj-context*, *rel-status*, *rel-author*, *rel-context*, *rs-status*, *rs-author* and *context* are numeric and may contain object identification numbers to refer to other objects defined within MDD-GIPHARM. Thus, MDD-GIPHARM is completely self-referencing (in a similar manner as it is VOSER, compare ROCHA et al. (1994)). More details about this self-reference and the management of the particular field domains is presented in section 4.

3. Describing Medical Facts within MDD-GIPHARM

The general concept of the MDD-GIPHARM database structure shall be illustrated by the following example. Let us assume we need to represent the facts that the drug "Lasix 500" contains the drug substance "Furosemid", that "Lasix 500" is a drug in the "Rote Liste" but also in the formulary of Gießen University Hospital (in the following abbreviated with "Form-GUH"), where it however is named "Lasix Tabl. 500 mg" and that "Furosemid" belongs to the group of drug substances, which in Form-GUH has been named "Saluretika". In order to include the above terms in the MDD-GIPHARM vocabulary the following entries are needed in the tables OBJECT and RELATION (the columns *obj-short*, *obj-long*, *obj-desc*, *obj-inv* and *rel-inv* are omitted because of the limited space availability; in the below examples however, they would not provide any additional information). The entries in the columns *obj-author*, *obj-status*, *obj-type*, *obj-context*, *rel-author*, *rel-status* and *rel-context* are coded. Decoding may be performed by refering to the respective value in column *obj-nr*. The objects 1 – 11, 45 and 5400 are not directly associated with the above example, but were necessary to consistently define the objects relevant for the example, since object-codes in the columns *obj-author*, *obj-status*, *obj-type*, *obj-context*, *rel-author*, *rel-status* and *rel-context* require their existence.

In order to represent the above mentioned causal relationships the following record entries have been made in the RELATIONSHIP table. All entries in the original table are coded and refer to the respective items in the OBJECT and RELATION table. For easier reading however, the corresponding object-name and relation-name have been used in the below illustration. The first 11 relationships have been entered to define

OBJECT (exemplary display of record entries)

obj-nr	obj-name	obj-date	obj-author	obj-status	obj-type	obj-context
1	Prokosch	1.7.94	1	2	3	4
2	test	1.7.94	1	2	3	4
3	attribute-entry	1.7.94	1	2	3	4
4	MDD-GIPHARM	1.7.94	1	2	3	4
5	attribute	1.7.94	1	2	3	4
6	mdd-author	1.7.94	1	2	5	4
7	object-type	1.7.94	1	2	5	4
8	context	1.7.94	1	2	5	4
9	status	1.7.94	1	2	5	4
10	concept	1.7.94	1	2	3	4
11	semantic type	1.7.94	1	2	3	4
45	import from RL	1.8.94	1	2	3	4
131	Rote Liste (RL)	1.8.94	1	2	10	4
132	drug substance (RL)	1.8.94	1	2	11	4
133	drug (RL)	1.8.94	1	2	11	4
140	Lasix 500	1.8.94	45	2	3	131
1120	Furosemid	1.8.94	45	2	3	131
5400	import from form-guh	1.9.94	1	2	3	4
5444	Form-GUH	1.9.94	1	2	10	4
5450	drug substance group (FGUH)	1.9.94	1	2	11	4
5451	drug substance (FGUH)	1.9.94	1	2	11	4
5452	drug (FGUH)	1.9.94	1	2	11	4
5460	Lasix Tabl. 500mg	1.9.94	5400	2	3	5444
5511	Saluretika	1.9.94	5400	2	3	5444

RELATION (exemplary display of record entries)

rel-nr	rel-name	rel-date	rel-author	rel-status	rel-context
1	is a	1.7.94	1	2	4
2	contains	1.7.94	1	2	4
3	belongs to	1.7.94	1	2	4

Figure 2: Exemplary entries in the OBJECT and RELATION tables (the columns *obj-author*, *obj-status*, *obj-type*, *obj-context*, *rel-author*, *rel-status* and *rel-context* are coded. They can be decoded based on the respective entries in the *obj-nr* column. Columns and rows are only partially displayed)

legal values for some of the fields in the OBJECT and RELATION table (compare domain definition in the next section).

4. Self-Reference and Domain Management

It has already been mentioned in the above section and is illustrated in figures 1 and 2 that MDD-GIPHARM is highly self-referencing. This means that all coding information, not only for clinical applications supported by MDD-GIPHARM, but also for the

obj-1	rel	obj-2	context
Prokosch	is a	mdd-author	MDD-GIPHARM
import from RL	is a	mdd-author	MDD-GIPHARM
import from form-guh	is a	mdd-author	MDD-GIPHARM
test	is a	status	MDD-GIPHARM
attribute-entry	is a	object-type	MDD-GIPHARM
attribute	is a	object-type	MDD-GIPHARM
concept	is a	object-type	MDD-GIPHARM
semantic type	is a	object-type	MDD-GIPHARM
MDD-GIPHARM	is a	context	MDD-GIPHARM
Rote Liste (RL)	is a	context	MDD-GIPHARM
Form-GUH	is a	context	MDD-GIPHARM
.....	.....	.....	.....
Lasix 500	is a	drug (RL)	Rote Liste (RL)
Furosemid	is a	drug substance (RL)	Rote Liste (RL)
Lasix 500	contains	Furosemid	Rote Liste (RL)
Lasix Tabl. 500mg	is a	drug (FGUH)	Form-GUH
Furosemid	is a	drug substance (FGUH)	Form-GUH
Saluretika	is a	drug substance group (FGUH)	Form-GUH
Lasix Tabl. 500mg	contains	Furosemid	Form-GUH
Furosemid	belongs to	Saluretika	Form-GUH

Figure 3: Exemplary entries in the RELATIONSHIP table (only partial display of columns and rows). All entries have been decoded (based on *obj-nr* and *rel-nr*) for better readability

coding of structured attribute-fields in MDD-GIPHARM itself, are defined within MDD-GIPHARM. The entries *obj-1*, *rel* and *obj-2* in RELATIONSHIP are natural links to OBJECT and RELATION representing the semantic network structure. All additional links illustrated in figure 1 (from *obj-author*, *obj-status*, *obj-type*, *obj-context*, *rel-author*, *rel-status* and *rel-context* to *obj-nr* in the OBJECT table) are self-referencing links where the respective entries have to be previously stored as objects in MDD-GIPHARM. Since for these fields not all objects defined in OBJECT represent legal values, the respective field domains have been defined using the RELATIONSHIP table. Legal values for the field *obj-author* for example may be derived from the OBJECT table by additionally applying the following selection criteria to the RELATIONSHIP table:

```
SELECT  o1.obj-name FROM rs relationship, o1 object, o2 object, r relation
WHERE   o1.obj-nr = rs.obj-1   AND
        rs.rel = r.rel-nr      AND
        r.rel-name = "is a"    AND
        o2.obj-nr = rs.obj-2   AND
        o2.obj-name = "mdd-author";
```

Thus, referential integrity criteria, defining the domains of every such coded field in OBJECT, RELATION and RELATIONSHIP may also be described based on the entries defined in RELATIONSHIP. An exemplary pseudocode description for the domain of *obj-author* is the following:

Domain (OBJEKTE.*obj-author*) = all x from OBJEKTE.*obj-nr*
for which an entry of the contents
x, is a, mdd-author
exists in RELATIONSHIP

These domain definitions have been applied in the realization of the management programs for MDD-GIPHARM. Selection-lists which pop up for the selection of the appropriate entry in the structured data entry fields are based on queries using the above logic. The advantage of this kind of self-reference and domain definition is that extensions in the scope of MDD-GIPHARM (for example adding new authors or new object-types) can be easily obtained by only entering the respective items in the OBJECT table and adding them to the relevant domains by adding new relationships. Thus, MDD-GIPHARM is highly flexible and can be dynamically adopted to new requirements.
It is however intuitiv (and obvious when looking at figure 2) that not all of the MDD-GIPHARM entries could have been stored based on this restrictive user-interface. An initial set of objects, relations and relationships (among others objects 1 – 11 from figure 2, the relation *is a* and some of the first 11 relationships in figure 3) needed to be instantiated manually.

5. Linking Heterogeneous Drug Information Sources

For the flexible integration of any drug information source the following assumptions have been made:
- the respective information source is either available in any relational database management system environment or at least in the form of unformatted ascii text files,
- the relevant drug information is associated with any of the major semantic types defined within MDD-GIPHARM (currently these are drugs, drug substances and drug substance groups),
- the relevant drug information can be uniquely identified within its original source
- a mappings (either automatically or manually) between the unique information identifiers in the original source and the respective objects of this semantic type in MDD-GIPHARM can be established.

If was further considered, if information from external sources should be directly integrated within the core tables of MDD-GIPHARM or if a loose coupling should be applied. It was decided to only establish loose coupling between MDD-GIPHARM (which is however absolutely transparent to the clinical end-user) and the contents of external sources in the current project status. Only if an external source does explicitly comprise a semantic network structure it will be investigated, if the information represented in the network can be integrated into the MDD-GIPHARM core tables without resulting into too much of redundancy or even inconsistency.
Loose coupling can be established by applying the relation "has attribute-set" in order to define relationships of the form ⟨semantic-type⟩ has attribute-set ⟨name of the external source⟩, where ⟨semantic type⟩'s are objects within MDD-GIPHARM with the object-type "semantic type" (e.g. the objects 132, 133, 5450, 5451 and 5452 in figure 2) and ⟨name of the external source⟩ is the filename of the respective datafile containing the external information which shall be linked with MDD-GIPHARM. Currently two sets of drug information sources, pharmacokinetic data and monographies, both linked with "drug substance (RL)" (object 132) as the respective semantic type have been integrated within MDD-GIPHARM. Due to space limitations the detailed design considerations which have been investigated and the final structures implemented for the integration of external sources can not be comprehensively covered in this paper.

6. Current Contents

The contents of MDD-GIPHARM may grow depending on the needs of the clinical applications within WAIN which MDD-GIPHARM is supposed to support. In the current preliminary version we have focused on representing the major pharmacological semantic types, such as drugs, drug substances and drug substance groups (the inclusion of new relationships to represent indications, contraindications and adverse drug reactions is already foreseen as a next step). From the original view at Gießen University Hospital it was important to incorporate the drug formulary of our hospital into MDD-GIPHARM. As a general reference for the integration of externally available pharmacological information it has been currently decided to include the drug, and drug substance items from the german drug information base „Rote Liste" (provided by the German Association of Pharmaceutical Companies). It is obvious however, that more scientifically approved drug information sources need to be additionally linked with the current MDD-GIPHARM objects. The american drug information system DRUGINDEX™ from Micromedix Inc. (as it is available as a stand-alone drug information system at every clinical workstation connected to the Gießen University Hospital network (compare PROKOSCH et al. 1994)) may serve as an excellent example for a much richer contents on actual drug information data and literature references. Nevertheless, integrating this information source into MDD-GIPHARM is currently not possible, since DRUGINDEX™ is only provided as a full information retrieval system and its database structure is not directly accessible from any other application. Other German drug information systems (e.g. the ABDA database or the SCHOLZ drug information system) may be easier to integrate, but are not freely available, so that integration efforts could not be started yet according to financial reasons.
Thus, in the current preliminary status, the only causal relationships represented within MDD-GIPHARM are the composition of drugs and drug substances, the grouping of drug substances into drug substance groups within the formulary of Gießen University Hospital and the associations between drugs/drug substances within our drug formulary and the Rote Liste contents. MDD-GIPHARM today comprise about 18000 objects, 4 relations and more than 50000 actual relationships. Among the 18000 objects there are currently 13423 drug names, 1134 drug substance names and 338 drug substance groups.
In order to illustrate MDD-GIPHARM's capabilities to act as an integrator between locally applied terminologies (e.g. brand names and drug substance names used within our hospital formulary) and drug information sources which are externally provided, either by other research institutions or by industrial companies, exemplary links have been established to two non-commercial drug information sources. These are firstly, a comprehensive set of drug monographies (restricted to a subset of drugs with particular relevance in the intensive care environment, established by one of the authors (B.D.)) and secondly, a pharmacokinetic drug information base, which is under development in the nephrology department of Ulm University (head F.K.). Since both sources are currently still under development and need final pharmacological validation (the final establishment and clinical validation of these drug information sources is one of the goals of the WAIN project), our current integration steps have not focused on really integrating the full contents of these sources, but rather on their exemplary integration based on some example drugs.

7. Results

MDD-GIPHARM is designed to serve as a distributed client/server-based medical data dictionary which may be applied in a cooperative project environment even over the

internet. The initial prototype version has been completely realized with MS-ACCESS on a DOS-PC. After having established this prototype version, the database has been exported to different Unix-based RDBMS (Sybase and Ingres). Typical information retrieval queries have been defined and used in a benchmark, to compare the performance of the ACCESS database server, Sybase and Ingres. When the DOS-PC is equipped with a 486 DX66 CPU and a minimum of 16 MB memory, navigation performance to perform the defined benchmark queries is satisfactory. Since on a DOS-PC MDD-GIPHARM can only be used in a single-user version it was decided to keep MS-ACCESS (and the DOS-PC) as a development platform, but to also realize a client/server version, in which a Unix-workstation with Sybase as RDBMS is applied as the database server and multiple ACCESS-clients can access this server using ODBC.
Finally the identical database structure (as illustrated in figure 1) has been established on the Tandem mainframe using the RDBMS Nonstop SQL. In this environment MDD-GIPHARM is actively applied to support the development of a structured drug therapy documentation module of the WING (PROKOSCH et al., 1991) hospital information system at Gießen University Hospital. Furthermore, this drug therapy documentation module will be enhanced with decision support capabilities in the upcoming months and MDD-GIPHARM will then play a major role in linking the knowledge base with the actual patient database and the applied medical terminology. In this context the PC-based version or the client-server version is applied for the MDD-GIPHARM management (and the cooperation betweeen the three WAIN partners) and database uploads to the Tandem MDD-GIPHARM version will be regularly performed.

8. Discussion

MDD-GIPHARM is a medical data dictionary, realized in order to support the computer-based drug therapy documentation within a university hospital and the various medical decision support functions which shall be established during the MEDWIS project WAIN. Its design is based on the long experience with medical data dictionaries at Gießen University Hospital (MICHEL et al., 1989; PROKOSCH et al., 1992; PROKOSCH et al., 1995), but also on the experiences and concepts which have been described by other groups who extensively work on developing medical data dictionaries (e.g. LINNARSSON and WIGERTZ, 1989; HUFF et al., 1987; CIMINO et al., 1989; LINDBERG et al., 1993; ROCHA et al., 1994). As mentioned by HUFF and CIMINO (1995) the scope of an MDD and the features it provides depend on the clinical application which shall be supported by it. Thus, the different characteristics of MDD-GIPHARM have to be compared with the objectives defined within the WAIN project. These were the following:

(1) to ensure a common understanding of the medical/medical informatics vocabulary used by the different WAIN partners, coming from different specialties, a glossary should be established were the major terms used within the project should be consistently defined and referenced with synonyms, associations, sub- and super-classes,

(2) to support the definition of a controlled vocabulary of all terms (e.g. brand names, substance names, drug substance groups and symptoms) which shall be applied in any form within a WAIN application,

(3) to allow the definition of basic factual knowledge on drugs and drug-related objects in order to provide an efficient basis for the implementation of medical decision support capabilities,

(4) to implement a medium which allows to link the pharmacological information sources established by the different project partners, but also drug information sources provided by industrial vendors to locally applied drug formularies in order to support a common presentation of such information within one clinical application,
(5) to provide a user-friendly interface for maintaining the MDD and for navigating through its contents,
(6) to be transferable to the Tandem mainframe environment, on which the WING hospital information system is implemented,
(7) to be capable of supporting a distributed cooperative MDD maintenance environment where partners located in different institutions may access the same MDD over the internet,
(8) to support the development of "intelligent" drug order entry programs (PROKOSCH et al., 1989; PROKOSCH and PRYOR, 1988) which shall be easy and fast to use by busy physicians,
(9) to be capable of linking the Arden-Syntax Editor (which shall be developed for the WAIN knowledge base management at Gießen University) with the locally used terminology and database structure.

The first objective, establishing a glossary of terms, is fulfilled by a glossary subcomponent developed at Ulm University (but not described in detail within this paper). The contents of the glossary was stepwise filled by the project partners and has also been provided as a WAIN contribution to the *MEDWIS Working Group Terminology*.
As described above, objectives two, three and four have been efficiently met by the MDD-GIPHARM realization. A controlled vocabulary for drug brand names, drug substance names and drug substance groups has been established and basic causal relationships between these objects are described within the RELATIONSHIP table. Furthermore, as illustrated in section 5, mechanisms have been defined which support the integration of different heterogeneous drug information sources and their linkage with a locally defined terminology.
By selecting Windows-PCs (and MS-ACCESS) as the MDD-GIPHARM development environment a user-friendly maintenance and navigation screen layout could be designed. Since a standard RDBMS was applied using SQL as query language, the database structure is easily transferable to any other RDBMS independent of its hardware platform. In order to support objectives six and seven, the MDD-GIPHARM database has already been transfered to the Tandem mainframe (within NonStop SQL) and to a UNIX platform (with SYBASE and INGRES as RDBMS). The NonStop-SQL version (on Tandem) currently supports the development of a structured drug entry program (within WING), which shall later be enhanced with decision support capabilities. Since MS-ACCESS provides an ODBC-interface to SYBASE, distributing MDD-GIPHARM in a client/server version with Windows clients and a UNIX (SYBASE) database server could be accomplished. Thus, distributed cooperative MDD maintenance can also be supported.
The MDD-GIPHARM feature to support the development of "intelligent" drug order entry programs has not been described in this paper, but was already efficiently applied to quickly realize a prototype version of the WING drug order entry component for demonstration purposes in order to illustrate the planned program functionality to future clinical users.
The last objective (concerning the integration of the Arden Syntax knowledge base editor) has not been tackled yet, but will be approached in a later project phase. The current database design however is flexible enough to support such a feature with only minor extensions.

References

AMIRI, F. (1993): DV-gestützte medizinische Dokumentation in der Rheumatologie mit integrierter Arztbriefschreibung. Diplomarbeit, Fachhochschule Gießen-Friedberg, Institut für Medizinische Informatik der Justus-Liebig Universität Gießen.

ANDERSON, J. (1986): Data Dictionaries – A Way Forward to Write Medical Meaning and Terminology into Medical Information Systems. Methods of Information in Medicine **25**, 137–138.

CIMINO, J. J., HRIPCSAK, G., JOHNSON, S. B. and CLAYTON, P. D. (1989): Designing an Introspective, Multipurpose, Controlled Medical Vocabulary. In: Kingsland III, L. C. (ed.) Proceedings SCAMC 1989, 513–518.

HUFF, S. M., CRAIG, R. B., GOULD, B. L., CASTAGNO, D. L. and SMILAN, R. E. (1987): Medical Data Dictionary for Decision Support Applications. In: STEAD, W. W. (ed.) Proceedings SCAMC 1987, 310–317.

HUFF, S. M. and CIMINO, J. (1995): Medical Data Dictionaries and their Use in Medical Information System Development. In: PROKOSCH, H. U., DUDECK, J. (eds.) Hospital Information Systems: Current Status and Future Trends, Elsevier, Amsterdam, in Press.

KRAUSE, D. (1993): Integration wissensverarbeitender Funktionen in ein Rheumatologisches Informationssystem. Diplomarbeit, Fachhochschule Gießen-Friedberg, Institut für Medizinische Informatik der Justus-Liebig Universität Gießen.

LINDBERG, A. A. B., HUMPHREYS, B. L. and MCCRAY, A. T. (1993): The Unified Medical Language System. IMIA Yearbook of Medical Informatics 93, Schattauer Verlag, Stuttgart, 41–51.

LINNARSSON, R. and WIGERTZ, O. (1989): The Data Dictionary – A Controlled Vocabulary for Integrating Clinical Databases and Medical Knowledge Bases. Methods of Information in Medicine **28**, 78–85.

MICHEL, A., PROKOSCH, H. U. and DUDECK, J. (1989): Concepts for a Medical Data Dictionary. In: BARBER, B. et al. (eds.) MEDINFO 89 Proceedings, North-Holland, Amsterdam, 805–808.

PROKOSCH, H. U., AMIRI, F., KRAUSE, D., NEEK, G. and DUDECK, J. (1995): A Semantic Network Model for the Medical Record of a Rheumatology Clinic. MEDINFO 95 Proceedings, in press.

PROKOSCH, H. U., PUHLE, B., MÜLLER, M., WAGNER, R., JUNGHANS, G., MARQUARDT, K. and DUDECK, J. (1994): From HIS to IAIMS: Expanding the Scope of Information Processing Applications in a German University Hospital. In: OZBOLT, J. G. (ed.) SCAMC Proceedings 1994, Hanley & Belfus, Philadelphia, 115–119.

PROKOSCH, H. U., DUDECK, J. and MICHEL, A. (1992): Standards for Data Dictionaries. In: BAKKER, A. R. et al. (eds.) Hospital Information Systems: Scope-Design-Architecture. Elsevier, Amsterdam, 189–195.

PROKOSCH, H. U., DUDECK, J., JUNGHANS, G., MARQUARDT, K., SEBALD, P. and MICHEL, A. (1991): WING – Entering a New Phase of Electronic Data Processing at the Gießen University Hospital. Methods of Information in Medicine **30**, 289–298.

PROKOSCH, H. U., WONG, T. W. and PRYOR, T. A. (1989): Medication Ordering Based on a Predictive Knowledge Base. Artificial Intelligence in Medicine **1**, 41–48.

PROKOSCH, H. U. and PRYOR, T. A. (1988): Intelligent Data Acquisition in Order Entry Programs. In: GREENES, R. A. (ed.) Proceedings SCAMC 1988, 454–458.

ROCHA, R. A., HUFF, S. M., HAUG, P. J. and WARNER, H. R. (1994): Designing a Controlled Medical Vocabulary Server: The VOSER Project. Computers and Biomedical Research **27**, 472–507.

Korrespondenzadresse: PD Dr. H. U. Prokosch, Institut für Medizinische Informatik, Justus-Liebig-Universität Gießen, Heinrich-Buff-Ring 44, 35392 Gießen

Informatik, Biometrie und Epidemiologie in Medizin und Biologie **26** (3),
262 – 274, ISSN 0943-5581
© Verlag Eugen Ulmer GmbH & Co., Stuttgart; Gustav Fischer Verlag KG, Stuttgart

Expertensystem zur Unterstützung von Diagnosestellung und Therapiewahl bei akuten Bauchschmerzen

C. Ohmann[1]), C. Platen[1]), G. Belenky[1]), C. Franke[2]), R. Otterbeck[1]), K. Lang[1]), H.-D. Röher[2]

Zusammenfassung

Trotz einiger Erfolge hat sich die computerunterstütze Diagnose bei akuten Bauchschmerzen nur in wenigen Kliniken durchsetzen können. Die vorhandenen Programme sind wenig benutzerfreundlich, die methodischen Ansätze zu einfach und die Datenbanken nicht mehr aktuell. Ziel des MEDWIS-Projektes ist es die Schwächen vorhandener Programme zu überwinden und ein verbessertes System zur Unterstützung von Diagnosestellung und Therapiewahl zu entwickeln, umfassend zu evaluieren und in die klinische Routine einzuführen. Die informatische Gesamtkonzeption umfaßt drei Komponenten: Data dictionary, klinische Datenbank mit Dokumentationsprogramm und Wissensbasis. Alle Komponenten werden auf der Basis der objektorientierten Analyse und des Designs entwickelt und als Schale in C++ unter WINDOWS realisiert. Das Data dictionary enthält Terminologieobjekte verbunden mit internen und externen Relationen und Dokumentobjekte. Es kontrolliert sowohl das Dokumentationsprogramm als auch die Wissensbasis. Das Dokumentationsprogramm umfaßt die klinischen Daten, wie z. B. Anamnese, klinische Untersuchung und Verlauf. Mehrere Eingabemodi wurden getestet, als Standardform wurde die formularbasierte Eingabe mit der Maus gewählt. Mit dem Programm wird zur Zeit an 14 deutschen Kliniken eine prospektive Datenbank aufgebaut. Die Wissensbasis befindet sich noch im Aufbau und soll Regeln, automatisch erzeugt aus prospektiven Datenbanken, Diagnosescores und probabilisistische Netzwerke umfassen. Mehrere Scores und verschiedene Algorithmen zur automatischen Regelerzeugung wurden bereits auf einer prospektiven Datenbank evaluiert, ein eigener deutscher Score zur Diagnose der Appendicitis entwickelt. Das Schwergewicht der laufenden Arbeiten liegt in der Integration der bisher entwickelten Komponenten und in der erweiterten Testung von Algorithmen zur automatischen Regelerzeugung.

[1]) Funktionsbereich Theoretische Chirugie
[2]) Klinik für Allgemein- und Unfallchirurgie, Heinrich-Heine-Universität, Düsseldorf

Schlüsselwörter:

akute Bauchschmerzen, Diagnose, Wissensbasis, Score, automatische Regelerzeugung, Evaluierung, Data dictionary

Summary

Despite sucessful studies, computer-aided diagnosis in acute abdominal pain has not been established in clinical routine. The existing programs are not user-friendly, the methodological approaches too simple and the clinical databases not up-to-date. It is the aim of the MEDWIS-project to overcome the weaknesses of these systems and to develop an improved system and to evaluate and introduce it into clinical routine. The informatical conception covers the components: Data dictionary, clinicial database with a documentation program and a knowledge base. All components will be developed by object-oriented analysis and design as a shell in C++ under WINDOWS. The data dictionary contains terminology objects with internal and external relations and document objects. It controls both the documentation program and the knowledge base. The documentation progam covers the clinical data, such as history, clinical examination and course of the disease. Several data entry modes were tested, the form-based data entry by mouse has been selected. Currently the program is used to build up a prospective database in 14 german hospitals. The knowledge base is under development and will consist of rules, automatically generated from prospective databases, diagnostic scores and probabilistic networks. Several scores and automatic rule induction techniques have been evaluated on a prospective database, a new german score for acute appendicitis has been developed. Work is focussing now on the integration of the developed components and increased testing of automatic rule induction techniques.

Key words:

acute abdominal pain, diagnosis, knowledge base, score, automatic rule induction, evaluation, data dictionary

Trotz moderner Technologie stellt die Diagnose bei akuten Bauchschmerzen auch heute noch ein erhebliches Problem dar (13). In einer aktuellen prospektiven Studie wurden Fehlerraten bei der Diagnosestellung von ca. 40% beim jungen unerfahrenen Arzt und ca. 20% beim erfahrenen Chirurgen beobachtet (20): Die Folge sind zum Teil gravierende Behandlungsfehler mit unnötig oder zu spät durchgeführten Operationen. Mehrere Studien haben gezeigt, daß durch Einführung der computerunterstützen Diagnose die Ergebnise signifikant verbessert werden können (1, 19). Trotz der Erfolge hat sich diese Methode nur in wenigen Kliniken durchsetzen können (6, 11, 16). Die Gründe hierfür sind vielfältig. Die angebotenen Programme sind wenig benutzerfreundlich, die methodischen Ansätze zu einfach (vorwiegend Bayes Theorem mitUnabhängigkeitsbedingung) und die zugrunde liegenden Datenbanken nicht mehr aktuell. Eine Testung des Programmes der Weltgesellschaft für Gastroenterologie auf einer deutschen Datenbank ergab enttäuschende Ergebnisse (Richtigkeit 50%; (20)). Da es sich bei der computerunterstützten Diagnose jedoch um einen hoffnungsvollen Ansatz handelt, wurde im Rahmen eines MEDWIS-Projektes ein Antrag formuliert, der die Entwicklung eines Expertensystems zur Diagnosestellung und Therapiewahl bei akuten Bauschschmerzen zum Ziel hat (22). Durch eine umfassende klinische Standardisierung, die Erstellung

qualitätskontrollierter prospektiver Datenbanken mit großer Fallzahl, die Anwendung vielversprechender Methoden der automatischen Regelerzeugung und durch die Erstellung eines benutzerfreundlichen Programms sollten die Schwächen vorhandener Systeme überwunden werden.

Methoden

Ziel des MEDWIS-Projektes ist es, ein Expertensystem zur Unterstützung von Diagnosestellung und Therapiewahl bei akuten Bauchschmerzen zu entwickeln, dieses umfassend in mehreren Studien zu evaluieren und letztendlich durch den Einsatz dieses Systems eine Verbesserung der Krankenversorgung in der Routine zu erreichen. Das methodische Vorgehen sieht zunächst in einer parallelen Entwicklung die Erarbeitung einer informatischen Gesamtkonzeption sowie die Erstellung einer klinischen Standardisierung vor. Die informatische Gesamtkonzeption umfaßt drei wesentliche Komponenten.

a) Data dictionary

An übergeordneter Stelle steht ein kontrolliertes Vokabular (Data dictionary, (15)). Das Data dictionary enthält alle medizinischen Terme mit ihren Eigenschaften, ihren Relationen untereinander und Abbildungen in externe Klassifikationen (z. B. SNOMED; 2). Durch das Data dictionary sollen die klinische Datenbank und die Wissensbasis integriert werden. Für die Konzeption des Data dictionaries wurden die Methoden der objektorientierten Analyse und des objektorientierten Designs eingesetzt (7). In Weiterentwicklung zur Literatur enthält unser Data dictionary zwei Komponenten: Terminologieobjekte und Dokumentobjekte. Terminologieobjekte sind die medizinischen Terme, die mit Hilfe eines semantischen Netzes intern mit anderen Terminogieobjekten und extern mit Standardklassifikationen verbunden sind. Ein Terminologieobjekt wird u. a. durch eine Nummer, Name, Kurzname, Definition, charakterisiert. Bei den semantischen Relationen werden, „ist ein", „ist Teil von", „ist äquivalent", „ist assoziiert mit", „ist lokalisiert" verwendet. Dokumentobjekte umfassen die üblichen bei der klinischen Dokumentation benötigten Daten. Hierzu zählen klinische Parameter, Antwortmöglichkeiten und ihre Zusammenfassung in medizinischen Teildokumenten und Dokumenten. Als Beispiel kann das Dokument Anamnese dienen, welches als Parameter „Schmerzort" und als Antwortmöglichkeit „rechter unterer Quadrant" enthält. Dokumentobjekte verfügen wiederum über eine Nummer, Name, Kurzname, Definition, ein Parameter enthält darüber hinaus den Datentyp, Wertebereich, Maßeinheit, Ein- und Ausgabeprozeduren, etc. Jedes Dokumentobjekt läßt sich eindeutig mit einem Terminologieobjekt verknüpfen.

b) Klinische Datenbank und Dokumentationsprogramm

Die klinische Datenbank, die zweite Komponente des Systems, umfaßt die klinischen Daten, wie z. B. Anamnese und klinische Untersuchung. Die klinische Datenbank ist gekoppelt an ein Dokumentationsprogramm für die direkte Benutzung durch Ärzte in der klinischen Routine. Um eine möglichst große Akzeptanz zu erreichen, wurde ein solches Dokumentationsprogramm in Zusammenarbeit mit klinisch tätigen Chirurgen entwickelt. Bei der Konzeption wurden wiederum die Methoden der objektorientierten Analyse und des Designs verwendet (7). Um eine möglichst große Akzeptanz zu erreichen, wurden verschiedene Eingabemodalitäten entwickelt und in einer Laborstudie durch 10 Chirurgen umfassend evaluiert (24). Dabei wurden die formularbasierte Eingabe mit

der Maus, die formularbasierte Eingabe mit dem Digitalisierungstablett und die menügesteuerte Eingabe mit der Tastatur untersucht. Von den beteiligten Ärzten wurden die Daten von insgesamt 54 schriftlichen Befunden in zufallsmäßiger Reihenfolge mit den verschiedenen Eingabemodi (je 18mal) in den Computer eingegeben. Zielgrößen stellten der Zeitaufwand für die Dateneingabe und die Qualität der Dokumentation (Vollständigkeit, Fehler) dar. Aufgrund der Laborstudie sollte sich dann für eine Eingabemodalität entschieden werden. Das in dieser Form aufgebaute Dokumentationsprogramm soll in die klinische Routine von 14 Kliniken eingeführt und in einer prospektiven Studie angewendet werden. Ziel ist der Aufbau einer qualitätskontrollierten, prospektiven Datenbank mit großer Fallzahl, die als Grundlage für die Anwendung wissensbasierter Methoden dienen soll.

c) Wissensbasis

Den dritten Bestandteil des Expertensystems stellt die sogenannte Wissensbasis zur Diagnoseunterstützung dar. Die Wissensbasis soll regelbasierte Systeme mit automatisch aus prospektiven Datenbanken erzeugten Regeln, Diagnosescores und evtl. probabilistische Netzwerke umfassen. Um eine geeignete Vorauswahl treffen zu können, werden vielversprechende Algorithmen der automatischen Regelerzeugung auf einer vorhandenen prospektiven Datenbank mit 1254 Patienten getestet (21). Die Ergebnisse sollen mit dem Standardmodell, dem Bayes Theorem mit bedingter Unabhängigkeit, verglichen werden. Bei den in der Untersuchung stehenden Algorithmen handelt es sich unter anderem um ID3, NewID, PRISM, C4.5, CN2 und ITRULE (21). Bei ID3 handelt es sich um den klassischen Ansatz, bei dem Parameter und Werte anhand eines informationstheoretischen Ansatzes (Entropie) selektiert werden (28). C4.5 und NewID stellen Weiterentwicklungen von ID3 dar (29, 33). PRISM maximiert die A-posteriori Wahrscheinlichkeiten für eine Diagnose (4). ITRULE benutzt das informationstheoretische J-Maß, um die informativsten Regeln aus einer Datenbank zu finden (31). CN2 verbindet einen informationstheoretischen Ansatz (Entropie) mit einer Signifikanzmetrik um die Regelqualität zu verbessern (5). Für die Evaluierung wird die Datenbasis zufallsmäßig in ein Trainingsset (n = 839) und in ein separates Testset (n = 415) aufgeteilt. Als Leistungsmaße dienen die Gesamtrichtigkeit, die krankheitsspezifische Richtigkeit und der prädiktive Wert für jede Diagnose.

Besonders gut geeignet für eine Diagnoseunterstützung sind die Diagnosescores. Diese Instrumente sind für den Kliniker leicht verständlich und ohne Mühe anwendbar. Für das zu entwickelnde Expertensystem sollen zunächst existierende Scores auf der vorhandenen deutschen Datenbank (n = 1254) getestet werden. Sollte sich herausstellen, daß diese Scores nicht die notwendige Leistungsfähigkeit haben, so sollen neue Scores entwickelt werden (18, 23). Um die für das Diagnosestellung signifikanten Faktoren zu ermitteln, wird eine schrittweise, logistische Regression durchgeführt. Dieses Verfahren wird zuerst für die akute Appendizitis angewendet. Sollte sich das Vorgehen bewähren, so werden Scores auch für andere Problemstellungen entwickelt, z. B. Ileus. Die Testung vorhandener und eigener Scores erfolgt mit Hilfe von Operator-Annahme-Kennlinien (ROC-Kurven). Damit die Analyse nicht zu überoptimistischen Ergebnissen führt, sollen weitere prospektive Datenbanken zur unabhängigen Evaluierung herangezogen werden. Zur Verfügung steht dabei eine niederländische Datenbank mit 1346 Fällen (18). Die Methode der probabilistischen Netzwerke soll zu einem späteren Zeitpunkt evaluiert werden und ist nicht Gegenstand der vorliegenden Arbeit.

Die vielversprechendsten Ansätze regelbasierter Systeme und von Diagnosescores sollen in die Wissensbasis eingebaut und über das Data dictionary mit dem Dokumentationsprogramm verbunden werden. Da vorhandene Expertensystemshells nicht über ausrei-

chende benutzerfreundliche Schnittstellen für die Anwendung in der klinischen Routine verfügen, ist die Entwicklung einer eigenen Schale und Benutzeroberfläche notwendig. Beispielhaft wurde diese Konzeption für Diagnosescores mit Hilfe objektorientierter Analyse und Design entwickelt (7). Um eine möglichst große Akzeptanz bei Klinikern zu erreichen, soll die volle Kontrolle der Wissensbasis in den Händen des Arztes liegen. Zu diesem Zweck wird ein sogenannter Wissensmanager entwickelt, der nach Bedarf durch den Arzt konsultiert wird. Bei der Konzeption des Wissensmanagers wurde auf ein Standardmodell der Wissensinferenz zurückgegriffen (32).

Ergebnisse

a) Data dictionary

Das objektorientierte Design eines Data dictionaries wurde mit Hilfe der C++-Sprache gemäß der standardisierten Kriterien der objektorientierten Progammierung umgesetzt (7). Grundlage bildet eine Schale („Shell"), die mit Hilfe eines Data dictionary-Editors mit Daten gefüllt werden kann. Der Data dictionary-Editor erlaubt die Eingabe, das Editieren und das Löschen von Terminologieobjekten. Zur Orientierung und Navigierung innerhalb der Terminologieobjekte können graphische Darstellungen von Terminologieobjekten mit Nachbarrelationen verwendet werden. In Abbildung 1 sind beispielhaft verschiedene Terminologieobjekte und interne Relationen dargestellt. Der Editor für die Dokumentobjekte besitzt ein ähnliches Funktionsspektrum, d. h. die Eingabeeditierung und Löschung von Dokumentobjekten ist möglich. Für jedes Dokumentobjekt können Relationen zu Terminolgieobjekten definiert und editiert werden. In Abbildung 2 ist ein Bildschirmausdruck für das Dokumentobjekt „Schmerzort, jetzt", ein klinischer Parameter der Anamnese, dargestellt. Mit Hilfe der Schale wurde basierend

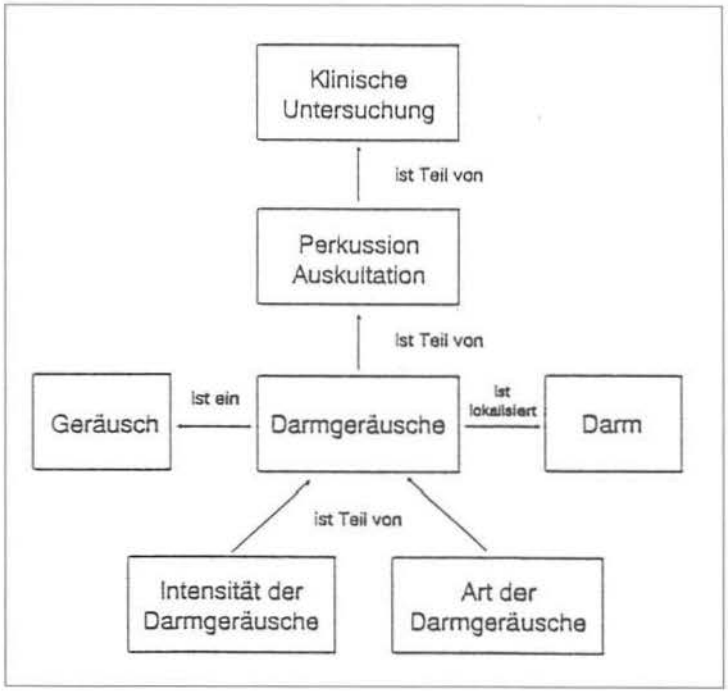


Abbildung 1: Semantisches Netzwerk für Terminologieobjekte (Beispiel)

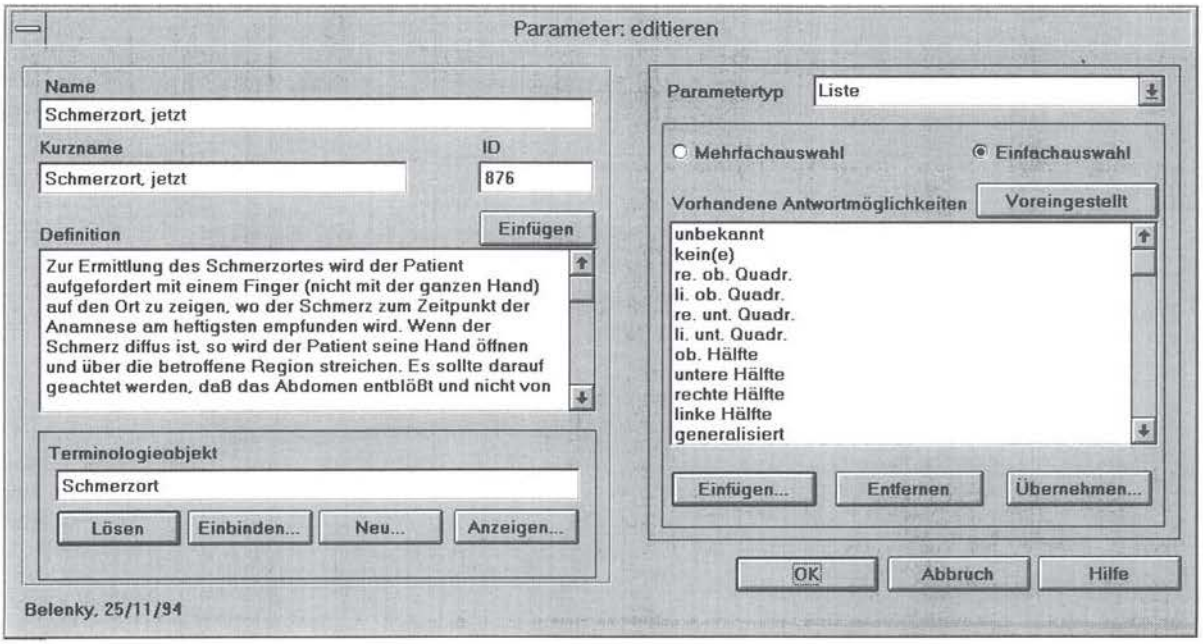


Abbildung 2: Dokumentobjekt: Parameter „Schmerzort, jetzt"

auf einer Standardisierung klinischer Experten ein Data dictionary für die Diagnoseunterstützung bei akuten Bauchschmerzen aufgebaut.

b) Klinische Datenbank und Dokumentationsprogramm

Die Evaluierung der verschiedenen Eingabemodalitäten in der Laborstudie ergab erhebliche Unterschiede speziell für die Eingabe der Anamnese und der klinischen Untersuchung. Eine komplette schriftliche Anamnese und klinische Untersuchung konnte mit der formularbasierten Eingabe durch das Digitalisierungstablett in ca. $2^1/_2$ Minuten, mit der formularbasierten Eingabe durch die Maus in $3^1/_2$ Minuten und durch die menügesteuerte Eingabe mit der Tastatur in 5 Minuten eingegeben werden (Tabelle 1 (24)). In Abhängigkeit von EDV-Kenntnissen und -fertigkeiten wurden zum Teil beträchtliche Unterschiede zwischen den Ärzten beobachtet. Gehäuft wurden fehlende Eingaben bei der formularbasierten Eingabe mit der Maus (50% der Befunde) und bei dem Digitalisierungstablett (72% der Befunde) festgestellt. Wesentliche Probleme ergaben sich bei der menügesteuerten Eingabe mit der Tastatur aus der sequentiellen Eingabe der Parameter, wodurch der Überblick über den Gesamtbefund verlorengeht. Schwierigkeiten bei der formularbasierten Eingabe mit der Maus waren die Orientierung im Bildschirmformular, der Wechsel des Eingabemediums bei Eingabe von Zahlen und Probleme mit der Windowsoberfläche. Bei der formularbasierten Eingabe mit dem Digitalisierungstablett ergaben sich zusätzliche Schwierigkeiten beim Einlegen des Dokumentationsbogens und der Kontrolle der Dateneingabe durch den Arzt auf dem Bildschirm. Keines der Verfahren war in allen Aspekten überlegen. Von den Ärzten selber wurde die formularbasierte Eingabe mit dem Digitalisierungstablett (6 Ärzte) und die formularbasierte Eingabe mit der Maus (4 Ärzte) bevorzugt. Für die Entwicklung des endgültigen Dokumentationsprogramm wurde in Absprache mit den klinischen Partnern die formularbasierte Eingabe mit der Maus gewählt (Abb. 3).

c) Wissensbasis

Die Testung der verschiedenen Algorithmen zur automatischen Regelerzeugung ergab keine Unterschiede bei der Gesamtrichtigkeit, ausgenommen für NewID mit einer Richtigkeit von nur 40% (Tabelle 2, (21)). Diese resultiert jedoch aus der Überspezialisierung mit 413 Regeln bei einer Richtigkeit von 92% in dem Trainingsset. Zwischen den verschiedenen Algorithmen wurden zum Teil erhebliche Unterschiede in Bezug auf bestimmte Diagnosen beobachtet. Keiner der Algorithmen erzielte eine Richtigkeit, die wesentlich über der des Standardmodells lag. ID3, PRISM und CN2 erzielten ähnliche

Tabelle 1: Zeitaufwand für die Dokumentation (Median (Bereich)) (24)

Zeitaufwand (Sek.)	Eingabemodus					
	formularbasiert mit Maus (n = 18)		formularbasiert mit Digitalisierungstablett (n = 18)		menügesteuert mit Tastatur (n = 18)	
Programmaufruf	32	(20–100)	24	(16–81)	19	(14–37)
Eingabe Stammdaten	89	(43–160)	104	(35–179)	74	(42–152)
Eingabe Anamnese	108	(75–200)	69	(36–126)	155	(98–282)
Eingabe Untersuchung	99	(57–148)	77	(40–140)	139	(50–255)
Programmabschluß	7	(4–10)	7	(4–15)	11	(5–45)

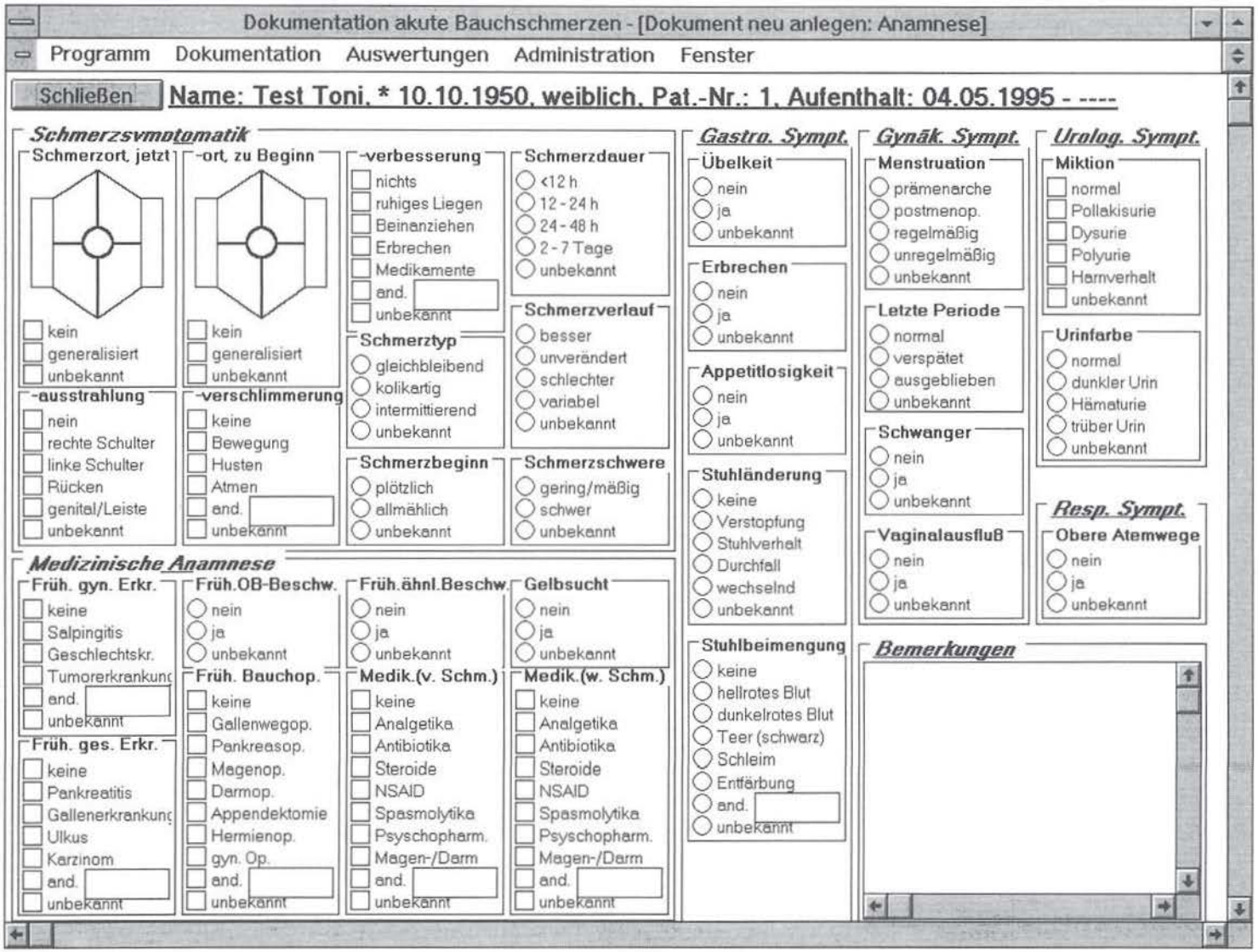


Abbildung 3: Dokumentation der Anamnese (22)

Ergebnisse mit sehr hohen Richtigkeiten für die diagnostische Kategorie „unspezifische Bauchschmerzen", mittlere Richtigkeiten für „Appendizitis" oder „Cholezystitis" und niedrige oder sehr niedrige Richtigkeiten für alle anderen Erkrankungen. Die Ergebnisse mit C4.5 unterscheiden sich von den drei oben genannten Algorithmen dadurch, daß die Richtigkeit für unspezifische Bauchschmerzen niedriger und für Appendizitis und Cholezystitis höher war. ITRULE war nur in der Lage drei Krankheiten vorherzusagen, in der Mehrzahl aller Fälle „unspezifische Bauchschmerzen". Das Standardmodell (Bayes Theorem) mit bedingter Unabhängigkeit zeigte eine hohe Richtigkeit für Appendizitis und Cholezystitis, eine mittlere Richtigkeit für unspezifische Bauchschmerzen, für Dyspepsie und perforierten Ulkus. Für vier der selteneren Erkrankungen wurden im Gegensatz zu den Algorithmen der automatischen Wissensgenerierung immerhin eine Richtigkeit von noch 30% und mehr erzielt.

Tabelle 2: Diagnostische Richtigkeit der Algorithmen zur automatischen Regelerzeugung (21)

Methode	Diagnostische Richtigkeit (%)	
	Training (n = 839)	Testung (n = 415)
Bayes	62	45
ID3	53	48
NewID	92	40
PRISM	54	45
CN2	61	47
C4.5	75	46
ITRULE	45	43

Die Testung von 10 Scores für Appendizitis auf der deutschen Datenbank ergab im Gegensatz zur Literatur enttäuschende Ergebnisse (23). Zwischen den Scores wurden erhebliche Unterschiede beobachtet. Aus den Untersuchungen wurde der Schluß gezogen, daß die publizierten Daten möglicherweise optimistisch verfälscht sind. Es wurde die Entscheidung zur Neuentwicklung eines Diagnosescores für Appendizitis getroffen. Bei der univariaten Analyse mit der deutschen Datenbank (n = 1254, (20)) erwiesen sich 15 Einzelfaktoren als statistisch signifikant für die Diagnose einer akuten Appendizitis. Die multivariate Analyse mit Hilfe der schrittweisen logistischen Regression ermittelte als statistisch signifikante Faktoren die in Tabelle 3 dargestellten Parameter. Die Gewichtung wurde durch Multiplikation der Koeffizienten mit dem Faktor 2 und einer entsprechenden Rundung berechnet. In Abbildung 4 sind die ROC-Kurven für diesen Score angewendet auf die beiden Datenbanken (deutsche Datenbank: n = 1254, niederländische Datenbank: n = 1346 (18)) dargestellt. Im Gegensatz zu den internationalen Scores sind die Ergebnisse wesentlich besser und erreichen zumindest näherungsweise die vorformulierten Standards. Für die Anwendung des Scores werden die folgenden Empfehlungen gegeben: Ausschluß von Appendizitis bei einem Punktwert unter 6, Monitoring und Rescoring nach 6 Stunden bei einem Score von 6 bis 11,5 und

Tabelle 3: Diagnosescore für akute Bauchschmerzen (18)

Parameter	Gewicht
Druckschmerz im rechten unteren Quadranten	4,5
Loslaßschmerz	2,5
keine Miktionsbeschwerden	2
kontinuierlicher (andauernder) Schmerz	2
Leukocyten $\geq 10000/\mu l$	1,5
Alter < 50 Jahre	1,5
Schmerzwanderung zum rechten unteren Quadranten	1
Abwehrspannung	1

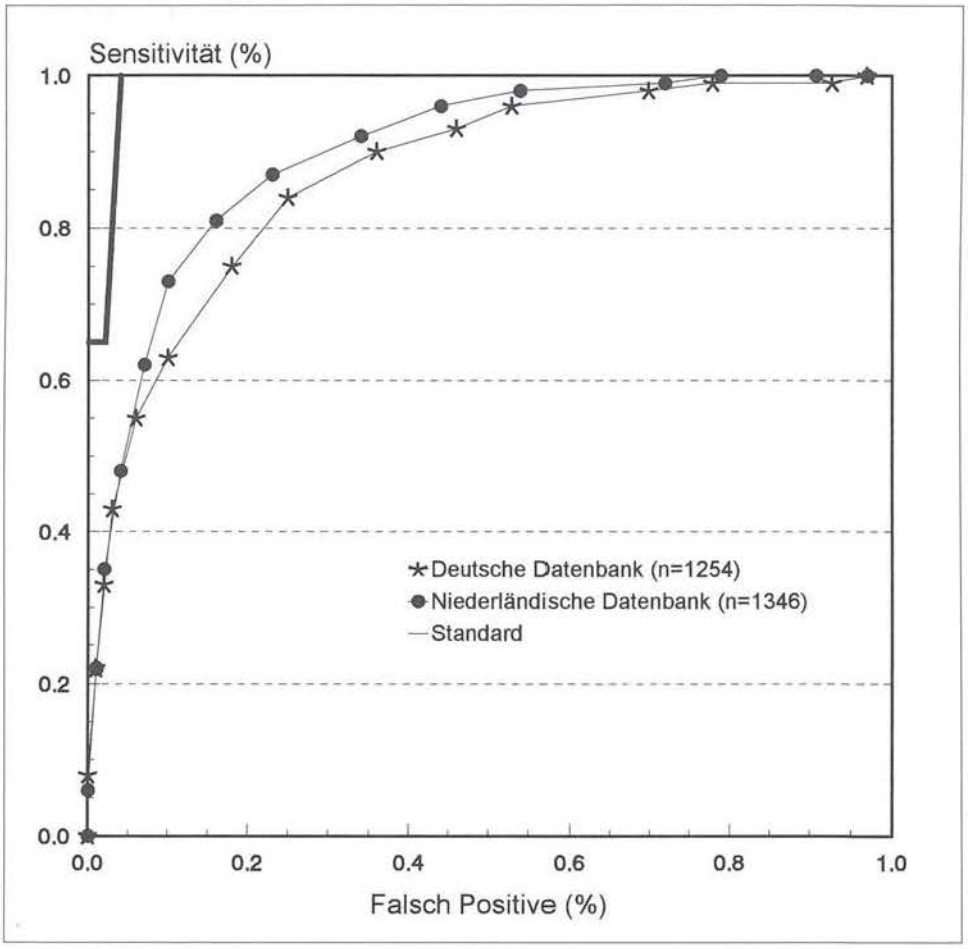


Abbildung 4: Operator-Annahme-Kennlinie (ROC-Kurve) für den Diagnosescore (18)

Evaluierung einer Appendektomie bei einem Punktwert von ≥ 12. Die Leistungsfähigkeit des Scores wird zur Zeit in einer prospektiven Interventionsstudie mit einer Prä-, Test- und Postphase in 14 deutschen Kliniken getestet.
Für die Darstellung des Wissens wurden wiederum die Methoden der objektorientierten Analyse und des Designs herangezogen. Eine Konzeption wurde beispielhaft für Scores erarbeitet. Dabei lassen sich alle gängigen Scores im System abbilden. Die Konzeption erlaubt neben der Anwendung der Scores auf klinische Daten auch die standardisierte Darstellung am Bildschirm, die Identifikation von Scores mit Hilfe eines Suchmusters und die Präsentation zusätzlicher Informationen, wie z. B. Evaluierungsstudien.

Diskussion

Die Verwendung eines Data dictionaries bei wissensbasierten Systemen stellt heute einen Standard dar (15, 26). Ein Data dictionary ist notwendig, um konsistente und richtige

Daten über die Zeit und zwischen verschiedenen Nutzern zu erreichen. Das Data dictionary-Konzept wurde in verschiedenen Krankenhausinformationssystemen umgesetzt, so z. B. bei HELP und WING (14, 25). Das Konzept des Data dictionaries wurde von uns konsequent weiterentwickelt. Eine Besonderheit stellt dabei die Trennung von medizinischen Termen und Dokumentobjekten dar, dies ist in der Regel in der Literatur nicht vorgesehen. Diese Vorgehensweise ist sinnvoll, da sich Definitionen erfahrungsgemäß schneller ändern als die medizinischen Terme selbst und Kliniker sich zwar häufiger über Begriffe einigen können, bei der klinischen Anwendung jedoch erhebliche Differenzen bestehen (30). Zu den wesentlichen Elementen eines Data dictionaries zählen die internen Relationen zwischen den medizinischen Termen. Diese können mit Hilfe einer Hierarchie oder auch mit Hilfe eines semantischen Netzwerkes (26) modelliert werden. Die Darstellung interner Relationen mit Hilfe eines semantischen Netzwerkes ist von besonderem Vorteil. Sie erlaubt den Aufbau von Dokumentobjekten für die Dokumentation, die Unterstützung von Userviews, die Unterstützung bei der Erzeugung von schriftlichen Befunden und Arztbriefen und die Unterstützung neuerer Dokumentationskonzepte (3, 17).

Voraussetzung für die Anwendung wissensbasierter Systeme ist die Dokumentation klinischer Daten. Durch mehrere Studien konnte gezeigt werden, daß der Zeitaufwand für eine computerunterstützte, strukturierte Dokumentation in der Regel höher ist als für die herkömmliche Dokumentation (10). Um einen hohen Grad der Akzeptanz zu erreichen, muß dieser Mehraufwand so gering wie möglich gehalten werden. In unserer Studie erwies sich die Dateneingabe mit den Digitalisierungstablett allen alternativen Modalitäten überlegen. Wegen technischer und Handhabungsprobleme (z. B. falsches Einlegen des Bogens) konnten wir uns für eine breite Einführung dieser Eingabemöglichkeit in die Routine nicht entschließen. Stattdessen führten wir die formularbasierte Eingabe mit der Maus ein, die im Vergleich zum Digitalisierungstablett zu einer geringeren Fehlerrate führt, aber dennoch den Zeitaufwand im Vergleich zu herkömmlichen Eingabe mit Tastatur und DOS-Programmen reduziert. Im Rahmen des MEDWIS-Projektes wird eine prospektive qualitätskontrollierte Datenbank aufgebaut, die als Basis für das wissensbasierte System dienen soll. Der Aufbau einer solchen Datenbank ist nicht nur wegen mangelnder Datenqualität vorhandener Datenbanken notwendig, sondern auch wegen der erheblichen Problematik der goegraphischen Variation (8). In der Literatur wurde diese gerade für die Bauchschmerzen mehrfach beschrieben und ist sicherlich zum Teil verantwortlich für die schlechten Ergebnisse bei der Anwendung englischer Datenbanken auf nicht-englische Fälle (z. B. (20)).

Die am häufigsten angewendete Methode bei der Diagnoseunterstützung der akuten Bauchscherzen ist das Bayes Theorem mit der Voraussetzung der bedingten Unabhängigkeit (19). Diese Methodik wurde in 75% aller Systeme angewandt. Techniken der künstlichen Intelligenz, wie z. B. regelbasierte Systeme, wurden nur selten eingesetzt. Dies resultiert aus der Tatsache, daß sich in der Regel die Zusammenhänge bei akuten Bauchschmerzen nicht über kausale pathophysiologische Zusammenhänge modellieren lassen. Techniken der automatischen Regelerzeugung wurden mit großem Erfolg in anderen klinischen Gebieten angewendet (12, 34). Erste Applikationen bei der Diagnose akuter kindlicher Bauchschmerzen wurden publiziert (9). Die Ergebnisse einer ersten eigenen Testung an einer vorhanden deutschen prospektiven Datenbank waren jedoch enttäuschend (21). Es konnten keine Verbesserungen gegenüber dem Standardmodell erzielt werden. Dabei muß jedoch berücksichtigt werden, daß die vorhandene Datenbank auf keiner standardisierten Terminologie basierte und außerdem die Datensammlung nicht ausreichend qualitätskontrolliert durchgeführt wurde. Die Ergebnisse sind deshalb nur als vorläufig zu bezeichnen. Verbesserungen sind zu erwarten bei Erhöhung der Fallzahl und bei einer Verbesserung der Datenqualität. Darüber hinaus könnte es interessant sein die Stärken der verschiedenen Algorithmen in einem hybriden Ansatz

zu verbinden. Scores werden aus verschiedenen Gründen als Entscheidungsunterstützungsinstrument propagiert. Sie sind nicht invasiv, zusätzliche Geräte sind nicht erforderlich und sie basieren auf Daten, die im Rahmen der routinemäßigen Anamneseerhebung und klinischen Untersuchung sowieso erhoben werden. Leider führte die geographische Variation zu denkbar schlechten Ergebnissen bei der Anwendung ausländischer Scores auf deutsche Daten (23). Durch die Entwicklung eines eigenen Scores konnten die Ergebnisse erheblich verbessert werden (18). Die Leistungsfähigkeit der Scores in der klinischen Routine wird zur Zeit in einer laufenden Studie untersucht. Darüber hinaus wird evaluiert, für welche Krankheiten eine Entwicklung weiterer Scores sinnvoll ist. Die objektorientierte Implementation der Scores ist so gestaltet, daß MLM-Module automatisch erzeugt werden können (27). Insgesamt läßt sich sagen, daß einzelne Elemente des wissensbasierten Systems bereits erfolgreich entwickelt und evaluiert wurden. Das Schwergewicht der laufenden Arbeiten liegt zum einem in der Integration der entwickelten Komponenten, zum anderen in der erweiterten Testung von Algorithmen zur automatischen Regelerzeugung.

Danksagung

Dem Bundesministerium für Forschung und Technologie wird für seine Unterstützung im Rahmen des MEDWIS-Programmes gedankt (Projekt A 70).

Literatur

[1] ADAMS, I. D., CHAN, M., CLIFFORD, P. C., COOKE, W. M., DALLOS, V., DE DOMBAL, F. T., EDWARDS, M. H., HANCOCK, D. M. HEWETT, D. J., MCINTYRE, N., SOMERVILLE, P. G., SPIEGELHALTER, D. J., WELLWOOD, J., WILSON, D. H. (1986): Computer aided diagnosis of acute abdominal pain: a multicentre study. British Medical Journal **293**, 800–804.

[2] BELENKY, G., PLATEN, C., LANG, K., OHMANN, C. (1994): Informatische Gesamtkonzeption für ein Expertensystem zur Unterstützung von Diagnosestellung und Therapiewahl bei akuten Bauchschmerzen. In: Pöppl, S. J., Lipinski, H. G., Mansky, T. (Hrsgb.): Medizinische Informatik. Ein integrierender Teil arztunterstützender Technologien. 38. Jahrestagung der GMDS, Lübeck, September 1993, MMV Medizin Verlag München, 182–185.

[3] BERNAUER, J. (1991): Designing a terminology controlled interface for report generation. In: Timmers, T., Blum, B. I. (Eds), Software engineering in medical informatics, Elsevier Science Publishers, 441–462.

[4] CENDROWSKA, J. (1988): PRISM: An algorithm for inducing modular rules. Knowledge-Based-Systems **1**, 255–276.

[5] CLARK, P., NIBLETT, T. (1989): The CN2 induction algorithm. Machine Learning **3**, 261–283.

[6] CLYMA, J. A., LANCASTER, G. (1993): Computer aided diagnosis of abdominal pain: implementation difficulties in the North Western Region. Archives of Emergence Medicine **10**, 314–320.

[7] COAD, P., YOURDON, E. (1991): Object-oriented analysis and object-oriente design. Prentice Hall, Englewood Cliffs.

[8] DE DOMBAL, F. T., STANILAND, J. R., CLAMP SE (1981): Geographical variation in disease-presentation. Does it constitute a problem and an information science help? Medical Decision Making **1**, 59–69.

[9] GAGA, L., MOUSTAKIS, V., VLACHAKIS, Y., CHARISSIG, G. ID$^+$: Enhancing medical knowledge acquisition with machine learning. Applied Artifical Intelligence (in press).

[10] GARRET, L. E., HAMMOND, W. E., STEAD, W. W. (1986): The effects of computerized medical records on provider efficiency and quality of care. Methods of Information Medicine **25**, 151–157.

[11] GUNN, A. A. (1991): The acute abdomen: the role of computer-assisted diagnosis. Bailliere's Clinical Gastroenterology **5**, 639–665.

[12] HADZIKADIC, M. (1992): Automatic design of diagnostic systems. Artifical Intelligence in Medicine **4**, 329–342.
[13] HOFFMANN, J., RASMUSSEN, O. (1989): Aids in the diagnosis of acute appendicitis. British Journal of Surgery **76**, 774–779.
[14] KUPERMAN, G. J., GARDNER, R. M., PRYOR, T. A. (1991): HELP: A dynamic hospital information system. Springer-Verlag, New York.
[15] LINNARSSON, R., WIGERTZ, O. (1989): The data dictionary – a controlled vocabulary for integrating clinical databases and medical knowledge bases, Methods of Information in Medicine **28**, 78–85.
[16] MCADAM, W. A. F., DAVENPORT, P., BROCK, B. M., CHAN, M., ARMITAGE, T., DE DOMBAL, F. T. (1990): Twelve years experience of computer-aided diagnosis in a distict general hospital. Annals of the Royal College of Surgeons of England **72**, 140–146.
[17] MOORMAN, P. W., VAN GINNEKEN, A. M., VAN DER LEI, J., VAN BEMMEL, J. H. (1994): A model for structured data entry based on explicit descriptional knowledge. Methods of Information in Medicine **33**, 454–463.
[18] OHMANN, C., FRANKE, C., YANG, Q., MARGULIES, M., CHAN, M., VAN ELK, P. J., DE DOMBAL, F . T., RÖHRER, H.-D. (1995): Diagnosescore für akute Appendicitis. Chirurg **66**, 135–141.
[19] OHMANN, C., KRAEMER, M. (1992): Evaluierung von Entscheidungsuntersützungssystemen bei der Diagnose von akuten Bauchschmerzen. Biometrie und Informatik in Medizin und Biologie **23**, 107–111.
[20] OHMANN, C., KRAEMER, M., JÄGER, S., SITTER, H., POHL, C., STADELMAYER, B., VIETMEIER, P., WICKERS, J., LATZTKE, L., KOCH, B., THON, K. (1992): Akuter Bauchschmerz – standardisierte Befundung als Diagnoseunterstützung. Chirurg **63**, 113–123.
[21] OHMANN, C., MOUSTAKIS, V., YANG, Q., LANG, K. Evaluation of automatic knowledge acquisition techniques in the diagnosis of acute abdominal pain. Artificl Intelligence in Medicine (in press).
[22] OHMANN, C., PLATEN, C., BELENKY, G. (1994): Computerunterstützte Diagnose bei akuten Bauchschmerzen. Biomedical Journal **39**, 15–18.
[23] OHMANN, C., YANG, Q., FRANKE, C. and the abdominal pain study group (1995): Diagnostic scores for acute appendicitis. European Journal of Surgery **161**, 273–281.
[24] PLATEN, C., OHMANN, C., BELENKY, G., LANG, K., FRANKE, C. Evaluierung von verschiedenen Dateneingabemöglichkeiten bei der Computerdokumentation von akuten Bauchschmerzen (MEDWIS-Projekt A 70). Proceedings GMDS-Tagung 1994 (im Druck).
[25] PROKOSCH, H. U., DUDECK, J., JUNGHANS, G., MARQUARDT, K., SEBALD, P., MICHEL, A. (1991): WING – entering a new phase of electronic data processing at the Giessen University Hospital. Methods of Information in Medicine **30**, 289–298.
[26] PROKOSCH, H. U., DUDECK, J., MICHEL, A. (1992): Standards for data dictionaries. In: Bakker, A. R. et al. (Eds.) Hospital Information Systems. Elsevier Science Publishers, 189–195.
[27] PRYOR, T. A., HRIPCSAK, G. (1993): The arden syntax for medical logic modules. International Journal of Clinical Monitoring and Computing **10**, 215–224.
[28] QUINLAN, J. R. (1986): Induction of decision trees. Machine Learning, **1**, 81–106.
[29] QUINLAN, J. R. (1992): Programs in Machine Learning. Morgan Kaufmann, California.
[30] RECTOR, A. L., NOWLAN, W. A., KAY, S. (1992): Conceptual knowledge: the core of medical information systems. In: Lun, K. C. et al. (Eds.): MEDINFO 92. Elsevier Science Publishers, 1420–1625.
[31] SMYTH, P., GOODMAN, R. M. (1992): An information theoretic approach to rule induction from databases. IEEE Transactions on knowledge and data engineering **4**, 301–316.
[32] STEFANELLI, M., RAMONI, M. (1992): Modelling medical knowledge based systems, Informatik, Biometrie und Epidemiologie **23**, 215–227.
[33] The Turing Institute (1993): NewId Software System. Georg House, 36 North Hannover, St. Glasgow GI2AD, Scotland.
[34] WIDMER, G., HORN, W., NAGELA, B. (1993): Automatic knowledge base refinement: learning from examples and deep knowledge in rheumatology. Artifical Intelligence in Medicine **5**, 225–244.

Korrespondenzadresse: Prof. Dr. C. Ohmann, Funktionsbereich Theoretische Chirurgie, Klinik für Allgemein- und Unfallchirurgie, Heinrich-Heine-Universität, Moorenstr. 5, 40225 Düsseldorf, Tel.: 02 11/3 11-61 42, Fax: 02 11/31 69 78

Informatik, Biometrie und Epidemiologie in Medizin und Biologie **26** (3),
275 – 285,
ISSN 0943-5581
© Verlag Eugen Ulmer GmbH & Co., Stuttgart; Gustav Fischer Verlag KG, Stuttgart

HEPA-CADS: Entwicklung eines tutoriell nutzbaren Expertensystems zur Diagnostik von Lebererkrankungen

HEPA-CADS: Development of a diagnostic system for tutoring in Hepatology

C. Engler, A. Führer, P. Hempel, H.-P. Buscher

Kurztitel:

Tutorielles Expertensystem für die Hepatologie

Stichwörter:
Tutorsystem, Expertensystem, Leber, Diagnostik

Key words:
tutor system, expert system, liver, diagnosis

Abstract

The development of HEPA-CADS is described. It is a computer-aided tool for the support of diagnostic processes in hepatology which is designed to fulfill tutorial functions. It consists of three components: a hierarchy of question sets and questions for the inquiry of symptoms, a hierarchy of hepatologic diagnoses, and rules which control the dialogue and connect symptoms with diagnoses. The sequence of questions being conditioned by rules regards pathophysiologic considerations. The tutorial advantage of the program is supported by explanations and illustrations and by the use of exemplary cases of hepatologic diseases.

Zusammenfassung

HEPA-CADS ist ein computergestütztes System, das die Diagnostik von Lebererkrankungen begleiten und tutoriell nutzbar sein soll. Es besteht aus einer Symptomerfassungs-, einer Diagnose- und einer Erklärungskomponente, einem verknüpfenden Regelwerk und

Trainingsfällen. Die Symptomerfassungskomponente enthält die zu den Symptomen führenden Fragen, die nach systematischen, kontextuellen und pathophysiologischen Gesichtspunkten geordnet wurden. Eine Ordnung der Fragen wurde hergestellt 1. durch Nachfragen, die immer nach einer bestimmten Antwort kommen (sie dienen zu ihrer Spezifizierung), 2. durch Anbindung von Fragen an eine oder mehrere Frageklassen, deren Aufruf alle unter ihnen stehenden Fragen aktiviert (sie dienen zur Aufrufbarkeit von der Diagnoseseite her) und 3. durch Regeln, die Fragen aus verschiedenen Frageklassen indizieren (sie dienen zur Überleitung zu anderen diagnostischen Gedankengängen). Die Diagnosekomponente enthält alle hepatologischen Diagnosen nach nosologischen Gesichtspunkten geordnet. Das Regelwerk verknüpft die Frageklassen und Fragen der Symptomerfassungsseite untereinander und mit der Diagnoseseite. Es enthält Komponenten, die zur Interpretation von Symptomen hinsichtlich pathophysiologischer Zustände, zur Wichtung von Diagnosen und zur von Diagnosen ausgehenden Indizierung weiterer Datenerhebung dienen. Durch wahlweisen Dialog entsprechend den unterschiedlichen Anforderungen (ausführliche Symptomerhebung oder rasche und straffe Diagnostik) können verschiedene tutorielle Ziele verfolgt werden. Die tutorielle Nutzbarkeit wird durch eine ausführliche Erklärungskomponte, durch Bilddokumentationen und durch zum Training bereitgestellte exemplarische Krankheitsfälle unterstützt.

Entwicklungsziel

Seltene Erkrankungen und schwierige Differentialdiagnosen erforden das Wissen medizinischer Experten. Um solches Wissen für Konsile und zu Ausbildungszwecken jederzeit verfügbar zu machen, werden Expertensysteme entwickelt, die unter Berücksichtigung des aktuellen Wissenstandes das diagnostische Denken erfahrener Ärzte nachahmen und fallbezogen Entscheidungsvorschläge unterbreiten sollen. Wir entwickeln ein computergestütztes Entscheidungssystem, das die Diagnostik in der Hepatologie unterstützen und gleichzeitig zur Ausbildung von Studenten und Ärzten nutzbar sein soll (HEPA-CADS: computer-aided decision support für Lebererkrankungen). Es ergänzt die parallel stattfindende Entwicklung von Programmen zur Diagnostik in der Rheumatologie (SCHEWE et al. 1990) und Hämatologie (PUPPE, B. 1995) und soll dazu beitragen, die Innere Medizin abzudecken. Damit HEPA-CADS für die gesamte Gastroenterologie ausbaubar ist und die Verdächtigung und Sicherung möglichst differenzierter Diagnosen unterstützen kann, soll die Symptomerfassungskomponente einen ausführlichen und erweiterungsfähigen Fragenkatalog enthalten. Tutorielle Gesichtspunkte sollen bereits im Konzept der Symptomerfassungskomponente berücksichtigt werden. Die Reihenfolge der Fragen soll pathophysiologische Überlegungen reflektieren. Symptome, technische Untersuchungen und Diagnosen sowie die vom System generierten Entscheidungsvorschläge sollen mit Erklärungen, Bildmaterial und Videoclips unterstützt werden. Die Organisation der Wissensbasis von HEPA-CADS, die diesen verschiedenen Anforderungen genügen soll, wird hier beschrieben und diskutiert.

Expertensystem-Werkzeug

Als Expertensystem-Werkzeug wird D3 der Entwicklergruppe von Prof. Puppe, Würzburg (FRG), in der aktuellen Version verwendet. Es ermöglicht unter anderem die Indikation von Frageklassen, Fragen und Nachfragen durch Symptome und Diagnosen im vorwärts- und rückwärtsschließenden Verfahren nach Hypothese-Test-Strategien und die differenzierte Wichtung von Diagnosen durch Symptome (PUPPE, F. 1990; PUPPE, F. 1991).

Definitionen

Im Rahmen der vorliegenden Arbeit werden folgende Definitionen verwendet: Unter *Symptomerfassung* wird die Erfassung medizinischer Symptome verstanden. Die Symptomerfassungskomponente von HEPA-CADS besteht aus Fragen zur Erhebung von für die Diagnostik relevantem Wissen.

Unter *Symptomen* werden solche anamnestischen Angaben und Befunde aus körperlicher Untersuchung oder technischen Untersuchungen (medizinische Daten) verstanden, die Informationen hinsichtlich möglicher Diagnosen ergeben. Symptome können aus Daten technischer Untersuchungen bestehen (Beispiel: der Befund ‚Quickwert erniedrigt' ist ein Symptom der Zwischenebenendiagnose ‚Gerinnungsstörung') oder sich aus ihnen ergeben (Beispiel: die Befunde ‚Bilirubin erhöht' und ‚Bilirubin im Serum überwiegend direkt reagierend' sind zusammen verwertbar als Symptom für die Zwischenebenendiagnose ‚Cholestase').

Unter *medizinischen Daten* werden Daten verstanden, die selbst Symptome sein können (Beispiel: Daten zu ‚Quickwert') oder mit anderen Daten zusammen als Symptome verwertbar sind (Beispiel: ‚Bilirubin im Serum' zusammen mit ‚Bilirubin, gesamt' und ‚Bilirubin, direkt' ergibt eine Information als Symptom).

Unter *Frageklassen* werden Ordnungsgruppen verstanden, die einen gemeinsamen Aufruf der unter ihnen zusammengefaßten Fragen sowohl von einem Symptom, als auch von einer Diagnose her ermöglichen. Fragen können je nach dem erforderlichen Ordnungsgefüge verschiedenen Frageklassen zugeordnet sein. Beispiel: Die Frage nach ‚Palmarerythem' steht unter den Frageklassen ‚Leberhautzeichen' und ‚Hand'.

Unter *technischen Untersuchungen* werden alle diejenigen Untersuchungen verstanden, die einen apparativen Aufwand erfordern und medizinische Daten der Symptomerfassung zugänglich machen.

Unter *Aussagekraft* (evoking strength) von Symptomen wird ihr empirischer Hinweiswert bezüglich in Frage kommender Diagnosen verstanden; sie wird für ihre Wichtung benötigt.

Unter *Diagnosen* werden Symptominterpretationen, Krankheitsentitäten (Beispiel: ‚primär biliäre Zirrhose') und die Oberbegriffe von Krankheitsentitäten (Beispiel: ‚Leberzirrhose') verstanden.

Unter *‚Diagnosekomplexe'* werden Komplexe von solchen Diagnosen verstanden, die in systematischem oder pathophysiologischem Zusammenhang miteinander stehen (Beispiel: ‚Leberzirrhose' steht in einem pathophysiologischen Zusammenhang mit ‚Koagulopathie' und ‚portale Hypertension').

Unter *‚Zwischenebenendiagnosen'* werden Diagnosen verstanden, die keine Enddiagnosen im Sinne von Krankheitsentitäten darstellen und einer weiteren Spezifizierung bedürfen. Sie sind Diagnoseoberbegriffe (Beispiel: ‚Leberzirrhose') oder pathophysiologische Zustände mit Diagnosecharakter (Beispiel: ‚portale Hypertension').

Unter *Differentialdiagnosen* werden Diagnosen der gleichen Differenzierungsebene verstanden, die um die höchste Priorität konkurrieren (Beispiel: ‚primär biliäre Zirrhose' und ‚Leberzirrhose bei primär sklerosierender Cholangitis'; beide Diagnosen stehen auf gleicher Ebene unter den Zwischenebenendiagnosen ‚Leberzirrhose' und ‚Cholangitis, chronisch').

Unter *Symptominterpretationen* werden die Interpretationen von Symptomen hinsichtlich pathophysiologischer Zustände verstanden (Beispiel: ‚Bilirubinkonzentration im Serum erhöht' und ‚Aktivität der alkalischen Phosphatase im Serum erhöht' werden als ‚laborchemische Cholestase' interpretiert). Symptominterpretationen werden verwendet, wenn der durch sie repräsentierte pathophysiologische Zustand nicht als Zwischenebenendiagnose erscheinen soll.

Unter *Wichtung* von Symptomen hinsichtlich verschiedener Diagnosen wird die Vergabe von positiven oder negativen Punkten auf die Punktekonten der durch diese Symptome ansteuerbaren Diagnosen verstanden.

Aufbau der Wissensbasis

Die Wissensbasis enthält (a) eine Diagnoseseite, in der die Diagnosen nach nosologischen Gesichtspunkten hierarchisch geordnet vorliegen (Abb. 1), (b) eine Symptomerfassungskomponente, in der Fragen nach den für die Diagnostik relevanten Symptomen repräsentiert sind und (c) ein Regelwerk, welches den Dialog steuert und die Beziehungen zwischen Symptomen und Diagnosen herstellt. Die Symptomerhebung folgt dem Regelschema: ‚wenn Symptom a oder Diagnose b vorliegt, dann frage die Fragen der Frageklasse c oder die Frage d (oder eine beliebige Zusammenstellung von Fragen oder Frageklassen)'. Die Regeln greifen auf eine Ordnung von Fragen und Frageklassen zurück, die nach systematischen, kontextuellen und pathophysiologischen Gesichtspunkten erstellt worden ist. Die Ordnung nach systematischen Gesichtspunkten ermöglicht eine Übersichtlichkeit für den Entwickler, diejenige nach kontextuellen und pathophysiologischen Gesichtspunkten eine Symptomerhebung in medizinisch sinnvoller und tutoriell nutzbarer Reihenfolge.

Eigenschaften der Symptomerfassungskomponente

Vollständigkeit der Symptomerfassungskomponente. Die Symptome, die bei den in HEPA-CADS berücksichtigten Lebererkrankungen vorkommen, wurden vollständig in der Symptomerfassungskomponente repräsentiert. Dazu wurden die Diagnosen und die bei ihnen vorkommenden Symptome aus Lehr- und Handbüchern der Hepatologie und aus veröffentlichten Statistiken gesammelt und nach systematischen Gesichtspunkten geordnet. Die Vollständigkeit der Symptome wurde entwicklungsbegleitend an klinischen Fällen und veröffentlichten Fallbeschreibungen überprüft.
Erweiterungsmöglichkeit der Symptomerfassungskomponente. Die Symptomerfassungskomponente von HEPA-CADS wurde hierarchisch aufgebaut. Es wurde eine Frageklassenordnung entwickelt, die vom Allgemeinen ins Spezielle führt. Jeder Frageklassenoberbegriff wurde so formuliert, daß weitere Frageklassen angefügt werden können. Ebenso können an bestehende Frageklassen weitere Fragen und an bestehende Fragen weitere Antwortalternativen angebunden werden, ohne daß das bestehende Gefüge aus Regeln gefährdet wird. Damit ist eine Erweiterbarkeit von HEPA-CADS gewährleistet.
Voraussetzungen für die vollständige Erfassung aller notwendigen Symptome im Dialog. Um die Symptome eines Krankheitsbildes einzeln oder gebündelt rasch und vollständig erhebbar zu machen, wurden die zu ihnen führenden Fragen zusätzlich zu ihrer Ordnung nach systematischen Gesichtspunkten auch nach ihrer kontextuellen Zusammengehörigkeit und nach pathophysiologischen Gesichtspunkten gebündelt. Dies gilt auch für Fragen nach Symptomen, die *gegen* eine Diagnose sprechen und vor ihrer Etablierung notwendigerweise ausgeschlossen werden müssen. Beispiel: Die Diagnose ‚Gilbert-Meulengracht Syndrom' wird gestellt, wenn ‚Bilirubin-gesamt' leicht erhöht ist, die ‚Cholestaseenzyme' und ‚Transaminasen' nicht erhöht sind und die Symptominterpretation ‚Hämolyse' nicht vorliegt. Die Symptome, die zur Symptominterpretation ‚Hämolyse' führen, wurden daher so angeordnet, daß sie bei Verdächtigung der Diagnose ‚Gilbert-Meulengracht Syndrom' *en bloc* erhebbar sind.

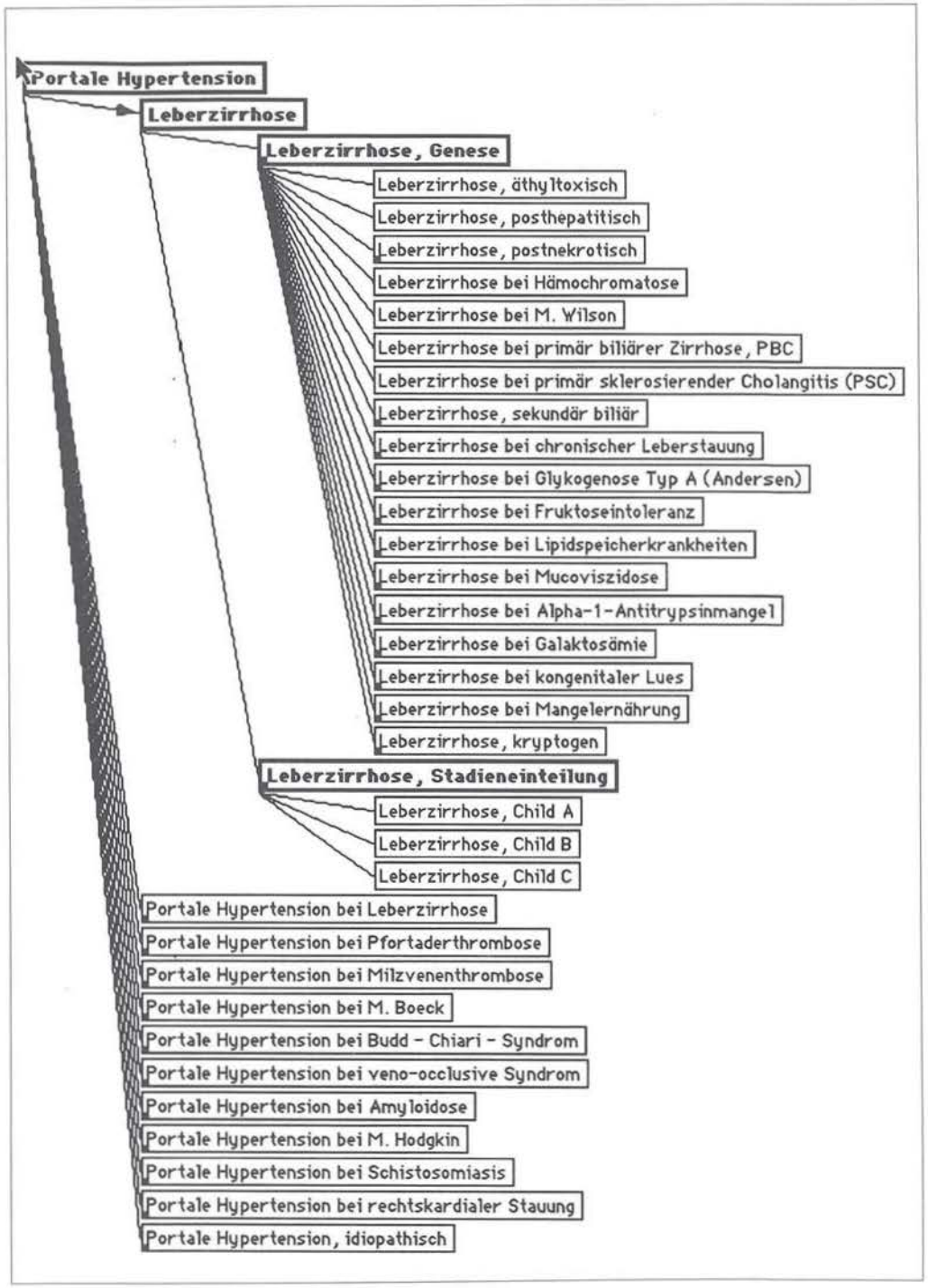
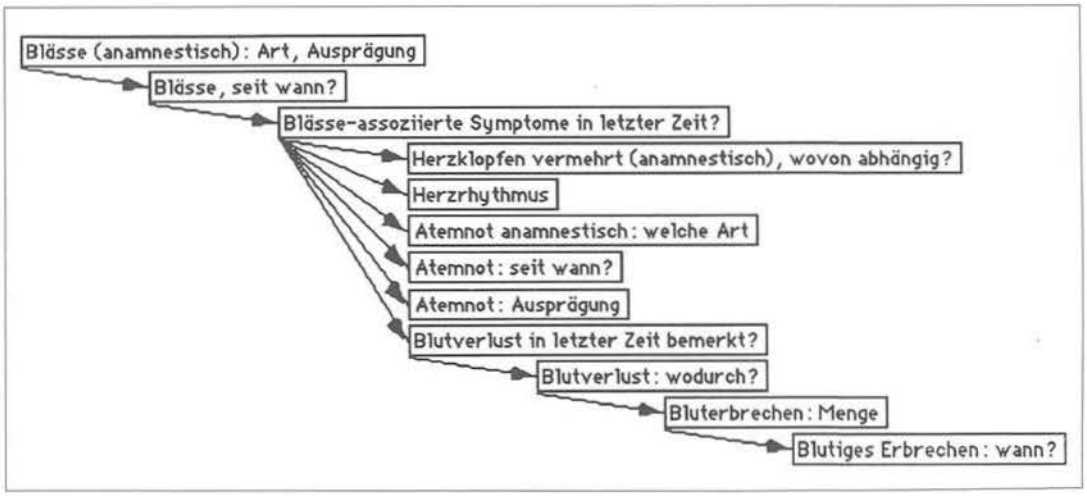


Abbildung 1: Diagnosehierarchie von ‚portale Hypertension' und ‚Leberzirrhose' (Ausschnitt)

Mechanismus zur Bündelung von Fragen

Um Symptome *en bloc* erheben zu können, wurden drei Wege beschritten:
1. *Bündelung durch Nachfragen.* Fragen, deren Antwortalternativen selbst noch spezifiziert werden müssen, um eine sinnvolle Wichtung von Diagnosen durchführen zu können oder um andere Fragen oder Frageklasen zu indizieren, führen direkt zu Nachfragen. Beispiele: Die Antwortalternative ‚Umfangszunahme des Bauchs' der Frage nach Symptomen im Abdomen hat immer die Nachfragen ‚Umfangszunahme, um wieviel' und ‚- in welcher Zeit' zur Folge. Auf die Bejahung der Frage nach Bauchschmerzen folgen immer die Nachfragen nach ihrer Lokalisation, Qualität, Stärke, Dauer und Beeinflußbarkeit. Eine Wichtung der Diagnosen erfolgt erst, wenn eine Antwort aus der letzten Nachfrageebene vorliegt.
2. *Bündelung durch Frageklassen.* Eine Bündelung durch Anbindung mehrerer Fragen an Frageklassen erlaubt eine gemeinschaftliche Indizierung von einer Diagnose aus. Eine Bündelung in Frageklassen ersetzt die getrennte Indikation der einzelnen Fragen durch multiple Regeln. Beispiel: Die Fragen nach Leber-Haut-Zeichen wurden nach systematischen Gesichtspunkten dort eingeordnet, wo sie am Körper auftreten (Palmarerythem unter ‚Symptome an den Händen' und Lacklippen unter ‚Symptome im Gesicht'); sie wurden zusätzlich unter der Frageklasse ‚Leber-Haut-Zeichen' zusammengefaßt. Symptome, die über die Aktivierung verschiedener Frageklasssen mehrfach erhebbar sind, werden im Dialog jedoch nur einmal erfragt.
3. *Bündelung durch antwortgetriggerte Regeln.* Einzelne Fragen verschiedener Frageklassen können innerhalb der Symptomerfassungsebene direkt angesteuert werden. Eine Indizierung ganzer Frageklassen, wie dies von der Diagnosenebene aus erforderlich ist, kann innerhalb der Symptomerfassungsebene vermieden werden. Dieser Mechanismus erlaubt einen Diagolg nach pathophysiologischen oder speziellen systematischen Gesichtspunkten (Abb. 2).

Abbildung 2: Fragesequenz, wenn Blässe als Eingangssymptom ausgewählt wird. Über Fragen nach den Ursachen gelangt man zur oberen GI-Blutung, die den Weg zur ‚portalen Hypertension' und weiter zur ‚Leberzirrhose' bahnt. Die Anschlußfragen gehören zu unterschiedlichen Frageklassen. Darstellung der Regelbeziehungen



Voraussetzungen für den Dialogbeginn aus verschiedenen Richtungen

Der Dialogbeginn muß entsprechend den Anforderungen in der Praxis variabel sein. Der erste Verdacht auf das Vorliegen einer Erkrankung entsteht durch ganz unterschiedliche Erstsymptome. Sie müssen aus den zu Beginn des Dialogs gestellten Übersichtsfragen hervorgehen. Die Eingangsfragen wurden daher so gestellt, das Basiswissen über den Patienten erhoben und die Richtung der Symptomerfassung festgelegt wird. Erstsymptome können amamnestische Informationen, Statusbefunde oder technische Untersuchungen sein. Es folgen Übersichtsfragen nach den spezielleren Bereichen, aus denen die vorliegenden Informationen kommen. Je nach Einstieg und Akuität des Krankheitsbildes werden unterschiedliche Fragefolgen indiziert. Beispiel: Als Einstiegsmöglichkeiten für die Diagnostik der ‚Leberzirrhose' mit ihren differenzierten Unterdiagnosen (wie ‚Leberzirrhose posthepatitisch' oder ‚Leberzirrhose bei Hämochromatose') wurden neben den bei Lebererkrankungen vorkommenden unspezifischen anamnestischen Symptomen auch ‚pathologische Laborwerte der Leber', ein auffälliger Befund bei der ‚Sonographie der Leber', ‚Bluterbrechen' und ‚Aszites' berücksichtigt. Das Einstiegssymptom ‚Bluterbrechen' führt über eine ‚ÖGD' zu den Diagnosen ‚obere gastrointestinale Blutung' und ‚Ösophagusvarizenblutung' und damit zur Verdächtigung der Zwischenebenendiagnose ‚portale Hypertension'. Von dort werden weitere, darunter auch sonographische Symptome erfragt, die in Richtung ‚Leberzirrhose' weisen. Der Einstieg über das primäre Symptom ‚Transaminasen mäßig erhöht' führt dagegen über die Zwischenebenendiagnose ‚Hepatopathie' und den von dort ausgehenden längeren Weg der verschiedenen anamnestischen Fragen schließlich auch zur Frage ‚Sonographie Leber' und damit zum Verdacht auf ‚Leberzirrhose'.

Dialogführung auf der Symptomerfassungsebene

Um rasch eine sinnvolle Auswahl der zur Verdachtsgenerierung oder Sicherung einer Diagnose notwendigen Fragen nach Symptomen zu erhalten, wurden zu jedem Symptomerhebungskomplex Eingangsfragen konstruiert. Sie stellen eine Auswahl grobgerasterter Antwortalternativen zur Verfügung, mit der diejenigen Frageklassen angesteuert werden, aus denen Symptome schon bekannt sind, und die daher für die weitere Symptomerhebung benötigt werden. Eingangsfragen dienen damit der Kürze des Dialogs (Abb. 3). Die in der Symtomerfassungskomponente von HEPA-CADS wirksamen Regeln legen eine Reihenfolge von Fragen fest, die allein durch die jeweiligen Antwortalternativen und nicht durch die verdächtigten Diagnosen bestimmt wird. Dieser Weg der Dialogsteuerung wurde dann gewählt, wenn weitere Fragen immer in der Folge einer bestimmten Antwort gefragt werden sollen und nicht den aktuellen Verdacht auf eine Diagnose oder ihre Sicherung zur Voraussetzung haben. Beispiel: Die Antwortalternative ‚Umfangszunahme des Bauchs' der Frage nach Symptomen im Abdomen hat immer die Frage ‚Gewichtszunahme?' zur Folge, weil dadurch der Verdacht auf Wassereinlagerung gelenkt werden kann, so daß der Weg in Richtung ‚Aszites' gebahnt wird. Hier wird ein pathophysiologischer Zusammenhang hergestellt, der tutoriell nutzbar ist. Ohne eine solche Regel würde die Frage nach einer Gewichstszunahme zwar auch gestellt, aber an einer Stelle im Dialog, der von einer ‚gedankenlosen' Systematik bestimmt wird.

Dialogsteuerung von der Diagnosenebene aus

Die Verdächtigung oder Etablierung einer Diagnose führt in der Praxis oft zu Nachfragen nach Symptomen, die zu ihrer Sicherung nötig sind oder die von ihr auf weitere Diagnosen schließen lassen. Beispiel: Die Etablierung der Diagnose ‚Leberzirrhose' führt zur Frage nach Symptomen einer ‚portalen Hypertension' (Sonographie des Abdomens und

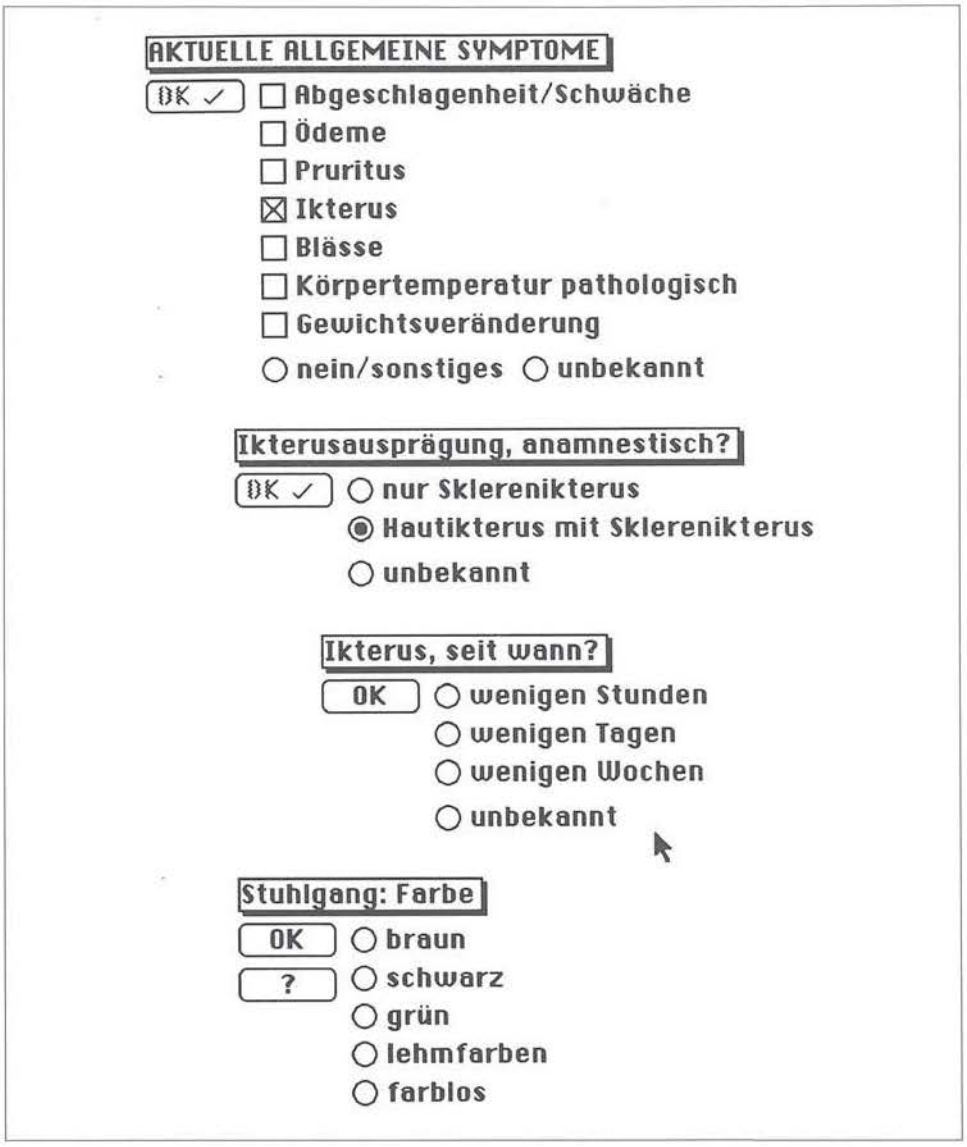


Abbildung 3: Eingangsfrage und Anschlußfragen, wenn als Leitsymptom ‚Ikterus' gewählt wird. Darstellung der Fragen im Dialog

Duplexsonographie) und einer ‚Gerinnungsstörung' (Labor – Leberzirrhose). Das Regelwerk von HEPA-CADS wurde so erstellt, daß diese Anforderungen berücksichtigt werden. Die Erhebung von Symptomen wird durch die in Verdacht geratenen Diagnosen komplexbezogen soweit vervollständigt, bis sie gesichert sind. Die Symptomerhebung nach Diagnosen und Diagnosekomplexen dient zur Straffung des Dialogs und zur inhaltlichen Nachvollziehbarkeit der Fragereihenfolge für den Benutzer. Sie ist eine Voraussetzung für die tutorielle Anwendbarkeit des Programms.

Tutorielle Gesichtspunkte beim Aufbau von HEPA-CADS

Voraussetzung für eine tutorielle Nutzung von HEPA-CADS ist die Möglichkeit einerseits zur vollständigen, andererseits aber auch zur straffen, zielgerichteten und nach pathophysiologischen Gesichtspunkten geordneten Symptomerhebung. Die dazu notwendigen Fragen wurden möglichst knapp und eindeutig formuliert. Die Zahl der Antwortalternativen pro Frage wurde aus Gründen guter Übersichtlichkeit meist unter 6 gehalten. Zu erklärungsbedürftigen Antwortalternativen, Symptomen und technischen Untersuchungen wurden Definitionen und Erläuterungen bereitgestellt, die bei Bedarf im Dialog aufgerufen werden können. Zudem sollen bildliche und graphische Darstellungen eingebunden werden, die zur Veranschaulichung von Sachverhalten und Symptomen dienen. Diagnosen werden mit lehrbuchartigen Kurzinformationen versehen, die dem Tutanden jederzeit das notwendige Wissen über ein Krankheitsbild hypertextartig zur Verfügung stellen. Es wurden zudem exemplarische Krankheitsbilder der verschiedenen Lebererkrankungen – beginnend mit den Formen der Leberzirrhose – für eine Falldatenbank zusammengestellt. Sie sollten zur entwicklungsbegleitenden Evaluation auf Vollständigkeit und Richtigkeit und als tutorielle ‚Trainingsfälle' dienen.

Diskussion

HEPA-CADS ist ein computergestütztes System, das diagnostische Entscheidungsfindungen im Bereich der Hepatologie unterstützen und tutoriell nutzbar sein soll. Die Brauchbarkeit eines solchen Systems im praktischen klinischen Alltag ist noch immer nicht definitiv nachgewiesen. Obwohl einige Programme käuflich zu erwerben und punktuell im klinischen Einsatz sind (ADLASSNIG, K. P. et al. 1989; BARNET, G. O. et al. 1987; BERNER, E. S. et al. 1994; DARMONI, S. J. et al. 1992; EYSENBACH, G. 1994; MILLER, R. et al. 1986; SCHEWE, S. et al. 1990; WARNER jr. H. R. 1989), hat sich doch keines bisher in der Praxis breit durchsetzen können. Auch wenn einige von ihnen speziell für die Ausbildung konzipiert sind (CLANCEY, W. 1987; CONSOLE, L. et al. 1992; SCHEWE, S. et al. 1990; THEODOROPOULOS, G. et al. 1994), ist für keines der tutorielle Nutzen nachgewiesen. Die Gründe liegen wahrscheinlich in der Komplexität der Programme und der Langatmigkeit des Dialogs.

Voraussetzung für eine Akzeptanz von HEPA-CADS scheint uns daher eine straffe Dialogführung zu sein. Sie ist kaum durch eine Beschränkung des diagnostischen Umfangs (beispielsweise auf seltene Lebererkrankungen) erreichbar, da die häufigen und ‚trivialen' Erkrankungen als Differentialdiagnosen der seltenen vorkommen und damit ebenfalls repräsentiert sein müssen. Die entwicklungsbegleitende Evaluation hat ergeben, daß die Zahl der für eine Diagnose einzugebenden Symptome meist zwischen 10 und 15 liegt, und daß zu ihrer Erhebung etwa 60–70 Fragen zu beantworten sind. Zur großen Zahl der Fragen trägt bei, daß meistens Ausprägung und Zeitverlauf der Symptome erfragt werden müssen. Bei zügiger Konsultation von HEPA-CADS wird damit ein Dialog etwa 15–20 Minuten dauern. Eine Beschränkung auf wenige Symptome – beispielweise nur auf die anamnestischer Angaben – wird mit einer deutlich schlechteren Differenzierbarkeit der Diagnosen erkauft. Da die Diagnosestellung seltener Lebererkrankungen in der Regel auch die Befunde aus körperlicher Untersuchung und speziellen technischen Untersuchungen erfordert, wurde eine Beschränkung auf anamnestische Angaben im Konzept von HEPA-CADS nicht verwirklicht und eine breite Symptomerfassung konzipiert. Programme mit einem zu engen Diagnosespektrum werden wegen zu geringer Entscheidungsunterstützung in der Praxis auch kaum attraktiv sein.

Eine straffe Dialogführung ist Aufgabe des Regelwerks. In HEPA-CADS wurde die Option einer raschen zielgerichteten Diagnostik zur Verfügung gestellt. Wenn sie gewählt wird, werden alle für die Verdächtigung und Sicherung einer Diagnose notwendigen Symptome erfragt, nicht aber alle bei ihr vorkommenden. Dieser Modus lehnt sich an das Denken eines hepatologischen Experten an und führt den Nutzer in der Weise durch die Diagnostik, daß eine zum jeweiligen Zeitpunkt gemäß systematischer oder pathophysiologischer Überlegungen sinnvolle Informationserhebung indiziert wird. Beim Aufbau des Regelwerks wurde darauf geachtet, daß diejenigen Untersuchungen, die für die gerade am höchsten gewichtete Diagnose die meisten Punkte ergeben und am wenigsten aufwendig sind, als erste indiziert werden. Die Erhebung von Informationen bezüglich einer Diagnose wird nach ihrer Sicherung nicht durchgeführt. Auf diese Weise gehorcht die Vorgehensweise derjenigen, die für die ‚rationale Diagnostik' gefordert wird.

HEPA-CADS verfolgt in einer weiteren Option die Anforderung einer vollständigen Erhebung von gastroenterologisch-hepatologischer Anamnesie und körperlichem Untersuchungsbefund. In diesem Modus werden alle bei Lebererkrankungen vorkommenden Symptome erfragt, auch wenn sie für einen aktuellen Fall nicht erforderlich sind. Alle Symptome werden mit Informationen zu ihren differentialdiagnostischen Bedeutungen hinterlegt. Damit werden die Vorbedingungen für eine tutorielle Nutzung geschaffen, die die Erlernung einer vollständigen klinischen Informationserhebung zum Ziel hat.

Während seiner Entwicklung wurde HEPA-CADS anhand von klinischen Fällen ständig auf Vollständigkeit und Richtigkeit evaluiert. Das Regelwerk hat sich dabei aus zwei Gründen als besonders kritisch herausgestellt. Erstens kann die Verwirklichung des Prinzips multipler Einstiege im Dialog zu Fragereihenfolgen führen, die tutoriell nicht erwünscht sind. Zweitens kann bei Verwirklichung der Option einer ‚rationalen Diagnostik' im Dialog ein rasches Springen zwischen Fragen aus Anamnese, Status und den einzelnen technischen Untersuchungen resultieren, das den Nutzer, insbesondere einen Tutanden, verwirren kann. Diese unerwünschten Dialogführungen wurden durch geeignete Einrichtung und Wahl von Frageklassen bei der Indizierung einer Symptomerhebung vermieden.

Mit den zur Verfügung stehenden Optionen eines straffen oder eines ausführlichen Dialogs zur Symptomerhebung werden nur die Vorbedingungen für eine tutorielle Anwendung erfüllt. Später sollen im Tutormodus Krankheitsfälle vorgegeben und gelöst werden, wobei dem Tutanden Vorschläge zum weiteren diagnostischen Vorgehen unterbreitet und seine Entscheidungen vom System kritisiert werden. Der tatsächliche tutorielle Wert eines solchen Lern- und Lehrmittels muß sich durch ausführliche Evaluation noch erweisen.

Literatur

ADLASSING, K.-P., AKHAVAN-HEIDARI, M. (1989): CADIAG-2/GALL: An experimental diagnostic expert system for gallbladder and biliary tract diseases. Artificial Intelligence in Medicine **1**, 71–77.

BARNETT, G. O., FELDMANN, M. J., KAHN, J. A. (1987): DXplain – An evolving diagnostic decision support system. JAMA **258**, 67–74.

BERNER, E. S., WEBSTER, G. D., SHUGERMAN, A. A. et al. (1994): Performance of four computer-based diagnostic systems. New Engl. J. Med. **330**, 1792–1796.

CLANCEY, W. Knowledge-based Tutoring (1987): The GUIDON-Program. MIT Press.

CONSOLE, L., MOLINO, G., RIPA DI MEANA, V., TORASSO, P. (1992): LIED – Liver: Information, Education and Diagnosis. Meth. Inform. Med. **31**, 284–297.

DARMONI, S. J., POYNARD, T. (1992): Computer-aided decision support in hepatology (Review). Scand, J. Gastroenterol. **27**, 889–896.
EYSENBACH, G. (1994): Computer-Manual, Urban u. Schwarzenberg.
MILLER, R., MASARI, F. E., MYERS, J. (1986): Quick medical reference (QMR) for diagnostic assistance. MD Computing **3**, 34–48.
PUPPE, B. (1995): Reflections on building medical decision support systems and corresponding implementation in diagnostic shell D3. AIME (akzeptiert).
PUPPE, F. (1990): Problemlösungsmethoden in Expertensystemen. Berlin: Springer.
PUPPE, F. (1991): Einführung in Expertensysteme. Berlin: Springer.
SCHEWE, S., HERZER, P., KRÜGER, K. (1990): Prospective application of an expert system for the medical history of joint pain. Klinische Wochenschrift **68**, 466–471.
THEODOROPOULOS, G., LOUMOS, V. (1994): Parasitology tutoring system: a hypermedia computer-based application. Comput. Methods Programs Biomed. **42**,113–118.
WARNER, Jr. H. R. (1989): Iliad: Moving medical decision making into new frontiers. Meth. Inform. Med. **28**, 370–372.

Korrespondenzadresse: Prof. Dr. H.-P. Buscher, DRK-Krankenhaus Köpenick, Salvador-Allende-Str. 2–8, 12559 Berlin, Tel.: (030) 6587-3319, Fax: (030) 6587-3321

BUCHBESPRECHUNGEN/BOOKREVIEWS

PAUL, N.
Medizinische Wissensbasen: Vom Wissensmodell zur Repräsentation
1995, Frankfurt am Main, Berlin, Bern, New York, Paris, Wien: Peter Lang
ISBN 3-631-48349-X

Zunehmend wird bei der Konstruktion wissensbasierter Systeme auf die lange Zeit gebräuchliche 'ad hoc-Modellierung', d. h. die direkte Transformation medizinischen Wissens in konkrete Repräsentationsformalismen eines zu realisierenden Anwendungssystems, verzichtet. Mittels konzeptueller Wissensmodelle wird versucht, relevantes medizinisches Wissen erst explizit zu machen und dann schrittweise in eine formale Repräsentation zu überführen. Die Vorteile liegen unter anderem in einer klareren Trennung zwischen Domänenwissen und argumentationsstrategischem Wissen.
In dem aus einer Dissertationsschrift entstandenen Buch begründet Norbert Paul ausführlich die Notwendigkeit der Trennung zwischen Wissensmodellierung und Repräsentation. Am Beispiel differentialdiagnostischer Fragestellungen aus der Nephrologie wird, ausgehend von Flußdiagrammen, schrittweise ein Wissensmodell erstellt und unter Verwendung objektorientierter Modelle in eine Wissensbasis überführt. Interessant macht das Buch nicht die Bearbeitung dieses, doch sehr einfach gehaltenen Beispiels, sondern die ständige Reflexion der Behandlung unterschiedlicher Wissensarten und deren Bedeutung.
Das Buch ist angesiedelt im Schnittbereich der Wissenschafts- und der Medizintheorie einerseits und der Konstruktion wissensbasierter Systeme als anwendungsbezogenem Zweig der Medizinischen Informatik andererseits. Daher richtet sich das Buch, auch wenn der Titel dieses nahezulegen scheint, nicht an Leser, deren primäres Interesse der Anwendung von Methoden und Modellen zur Konstruktion wissensbasierter Systeme in der Medizin gilt; Zielgruppe sind vielmehr Leser, deren Augenmerk sich auf Fragen richtet, die sich mit den zur Wissensmodellierung und Repräsentation verwendeten Methoden selbst beschäftigen.

G. Mann Neuherberg

BEHRENDS, E.
Überall Zufall. Eine Einführung in die Wahrscheinlichkeitsrechnung
1994, 214 S., DM 29,80, ISBN 3-411-17091-3
BI-Wissenschaftsverlag, Mannheim-Leipzig-Wien-Zürich

Eine Einführung in die Wahrscheinlichkeitsrechnung kann auch einfach und leicht lesbar sein, wie das vorliegende Buch zeigt. Dabei muß aber festgehalten werden, daß unter dieser leichten Lesbarkeit die Exaktheit der Darstellung nicht zu leiden braucht. Dies bedingt auch, daß vom Leser eine Bereitschaft zum Mitdenken erwartet wird, so wie Erfahrungen im Umgang mit Computern nützlich sind.
Behandelt werden nach einer ausführlichen Einleitung die wichtigsten Wahrscheinlichkeitsräume sowie die Grundbegriffe der Theorie (Zufallsvariable, Unabhängigkeit, Binomial-, Exponential- und Normalverteilung) und im letzten Teil die Gesetze der großen Zahlen.
Geeignet erscheint das Buch für Studenten, Schüler (Leistungskurs) und Mathematiklehrer.

H. Geidel, Stuttgart

RETHERFORD, R. D.; CHOE, M. K.
Statistical models for causal analysis
1993, 258 S., £ 32,95, ISBN 0-471-55802-8
J. Wiley & Sons, Inc., New York

„Causal analysis" ist vielleicht etwas hochgegriffen. Explorative Datenanalyse trifft den Inhalt besser. Behandelt werden Methoden der linearen Regression, der multiplen Klassifikation, der Pfadanalyse, der logistischen Regression sowie Überlebensmodelle. Ergänzt werden die Darstellungen durch Progammlistings für SAS, BMDP und LIMDEP. Die Beipiele wurden aus einem Datensatz des „Fiji Fertily Survey 1974" genommen.
Hervorzuheben ist die klare, leicht verständliche Darstellung, die anhand der Beispiele die verschiedenen Methoden, ihre Grenzen und mögliche Randbedingungen erläutert. Von daher möchte man sich eine deutsche Bearbeitung dieses Textes als vorlesungsbegleitende Lektüre für Medizinstudenten wünschen.

H. Geidel, Stuttgart