

Modellrevision in vernetzten, unsicheren Wissensgebieten

Heinz-Jürgen Thole

OFFIS Institut, Forschungs-und Entwicklungsbereich: Informations- und
Kommunikationssysteme im Gesundheitswesen

E-Mail: Heinz-Juergen.Thole@informatik.uni-oldenburg.de

Modellbildung ist in Bereichen komplexen unsicheren Wissens ein schwieriges Geschäft, indem ein erster Modellierungsversuch durch einen iterativen, oft langwierigen Revisionsprozeß schrittweise Realitätsnähe gewinnt. Im besonderen bei kooperativer Modellentwicklung benötigen die Fachexperten die Hilfe einer Unterstützungskomponente. Vorteilhaft ist das Aufzeigen von Gemeinsamkeiten und Unterschieden ihrer Modelle. Bei revidierenden Eingriffen bewahrt die Kenntnis der Konsequenzen vor Folgefehlern. Eine *qualitative Modellrevisionskomponente* wurde bereits vorgestellt (Schröder, Thole, Willms, 1996). Das Thema dieses Beitrags ist eine *quantitative Modellrevisionskomponente*. Die entropiebasierte Revisionskomponente ist in die Bayes-Netz (BN) Modellierungsumgebung MEDIKUS integriert. Das Akronym MEDIKUS steht für: Modellierung, Erklärung und Diagnoseunterstützung bei komplexen, unsicheren Sachverhalten, es ist ein Forschungsprojekt vom OFFIS Institut (Oldenburger Forschungs- und Entwicklungsinstitut für Informatik-Werkzeuge und -Systeme). Durch die neue Revisionskomponente werden fehlerhafte bzw. ungenaue Modelle an Restriktionen angepaßt, die z. B. Aussagen neuer Daten bzw. wissenschaftlicher Erkenntnisse beinhalten, die gelten müssen damit das entwickelte Modell valide ist. Dabei werden die intransparenten Konsequenzen der Eingriffe aufgezeigt. Nach einer kurzen Beschreibung der MEDIKUS Komponenten, folgt in Abschnitt 3 die Darstellung der quantitativen Modellrevision. Zuerst werden an einem BN-Modell aus der Umweltmedizin die vielfältigen Formulierungsmöglichkeiten für Restriktionen vorgestellt. Danach geben wir in Abschnitt 4 als Ausblick ein Beispiel für den Einsatz der *quantitativen* Modellrevisionsunterstützung bei *kooperativer* Modellkonstruktion.

1. Einleitung

Viele Wissensgebiete (wie z.B. Marketing, Investitionsplanung, epidemiologische Umweltmedizin, aber auch die Spracherkennung) zeichnen sich dadurch aus, daß das relevante Wissen hochgradig vernetzt und unsicher ist. Prognose, Diagnose und Erklärungen von Einflüssen, Ursachen und Symptomen sind daher wegen ihrer schwer überschaubaren Fern-, Neben- und Wechselwirkungen komplex und fehleranfällig. Eine Unterstützung bei der Konstruktion von Erklärungsmodellen und bei der Entwicklung von Diagnosestrategien ist deshalb überaus wünschenswert und wichtig.

Das Ziel des OFFIS Projekts MEDIKUS besteht in der Entwicklung eines Softwarewerkzeugs zur Unterstützung der Modellkonstruktion und bei der Entwicklung geeigneter Planungs- und Diagnosestrategien. Den "Flaschenhals" bei der Modellkonstruktion bildet die Wissensakquisition. Zum einen neigen Experten eher dazu, vorsichtige qualitative Einschätzungen von

Einflüssen zu äußern, als sich auf präzise quantitative Angaben festzulegen. Bei ihren Aussagen verwenden sie linguistische Variablen wie zum Beispiel "... beeinflusst oft ..." oder "... wirkt häufig entgegen ...". Ein Forschungsziel in MEDIKUS besteht in der Entwicklung von Verfahren zur Bestimmung einer Semantik dieser linguistischen Variablen durch die Zuordnung von Wahrscheinlichkeiten. Zum anderen sind die Expertenangaben in der Regel unvollständig, da die nötigen, objektiven Daten nicht vorliegen. Die Modelle können oft nur unvollständig quantifiziert werden. Zum Teil gehen subjektive Schätzungen in die Modellkonstruktion ein. Diese können zu Inkonsistenzen im Modell bzw. unerwarteten Modellprognosen führen. Daher ist die Entwicklung einer automatischen, quantitativen Modellrevisionskomponente zu einem wichtigen Forschungsthema im MEDIKUS-Projekt geworden.

2. Überblick MEDIKUS

Zur Unterstützung des Modellierers, der ein ablauffähiges Erklärungs- und Diagnosemodell für sein Wissensgebiet erstellen will, verfügt MEDIKUS über 6 Modellierungskomponenten:

- einen linguistischen Modelleditor
- einen graphischen Modelleditor
- eine Komponente zur qualitativen Modellrevision
- eine Komponente zur quantitativen Modellspezifikation
- eine Erklärungskomponente
- eine Komponente zur quantitativen Modellrevision

Die ersten fünf Teilkomponenten werden hier nur kurz beschrieben, da sie in früheren Projektphasen entwickelt wurden. Daran schließt sich die Skizze unserer quantitativen Modellrevisionskomponente an. Sie ist eine komplette Neuentwicklung, die ähnliche Ansätze in ihrer Ausdrucksfähigkeit übersteigt. Das Anwendungsbeispiel stammt aus der Umweltmedizin und beschreibt die Risiken von BTX-Aromaten, Phenolen und Formaldehyd.

MEDIKUS stellt dem Anwender wahlweise einen linguistischen und einen graphischen Editor zur Verfügung (FOLCKERS et al., 1996; SCHRÖDER, MÖBUS, FOLCKERS, THOLE, 1996). In dem *linguistischen Editor* gibt der Anwender in vereinfacht-natürlichsprachlicher Form eine Beschreibung des zu modellierenden Sachverhalts ein. In dem *graphischen Editor* werden die Modelle mit der wahrscheinlichkeitstheoretischen Methode der Bayes-Netze, gerichtete azyklische Graphen, konstruiert (Abb. 1) (NEAPOLITAN, 1990; PEARL, 1988). Will der Modellierer den graphischen Editor nicht benutzen, können die vereinfacht-natürlichsprachlichen Aussagen in Graphen übersetzt werden.

Die *qualitative Modellrevision* basiert auf den in den Bayes-Netzen implizit enthaltenen Abhängigkeits- und Unabhängigkeitsaussagen (FOLCKERS et al., 1996; SCHRÖDER, MÖBUS, FOLCKERS, THOLE, 1996). Der Anwender kann überprüfen lassen, ob seine Annahmen über Abhängigkeiten und Unabhängigkeiten mit denen seines Modells übereinstimmen. Er gibt seine Annahmen ein, indem er bei gegebenen Daten/Beobachtungen Aussagen macht, welche zusätzlichen neuen Daten/Informationen zur Stärkung bzw. Schwächung einer zu untersuchenden Hypothese beitragen können und welche nicht relevant sind. Bei Abweichungen macht die Komponente Revisionsvorschläge.

Die Komponente zur *quantitativen Modellspezifikation* dient zur Bestimmung einer wahrscheinlichkeitstheoretischen Semantik linguistischer Variablen (SCHRÖDER, MÖBUS, THOLE, 1996; MÖBUS, SCHRÖDER, 1997a, 1997b). Zu univariaten Aussagen, die sich nur auf eine Modellvariable beziehen, gibt es bereits reichhaltige Literatur. Gänzlich unbekannt

ist aber die entsprechende Semantik multivariater Modellierungsaussagen (wie z.B.: "das Auftreten von X hat Y ziemlich oft zur Folge"). Im MEDIKUS-Projekt ist ein Verfahren entwickelt worden, das die probabilistische Semantik dieser Modellierungsaussagen bestimmt.

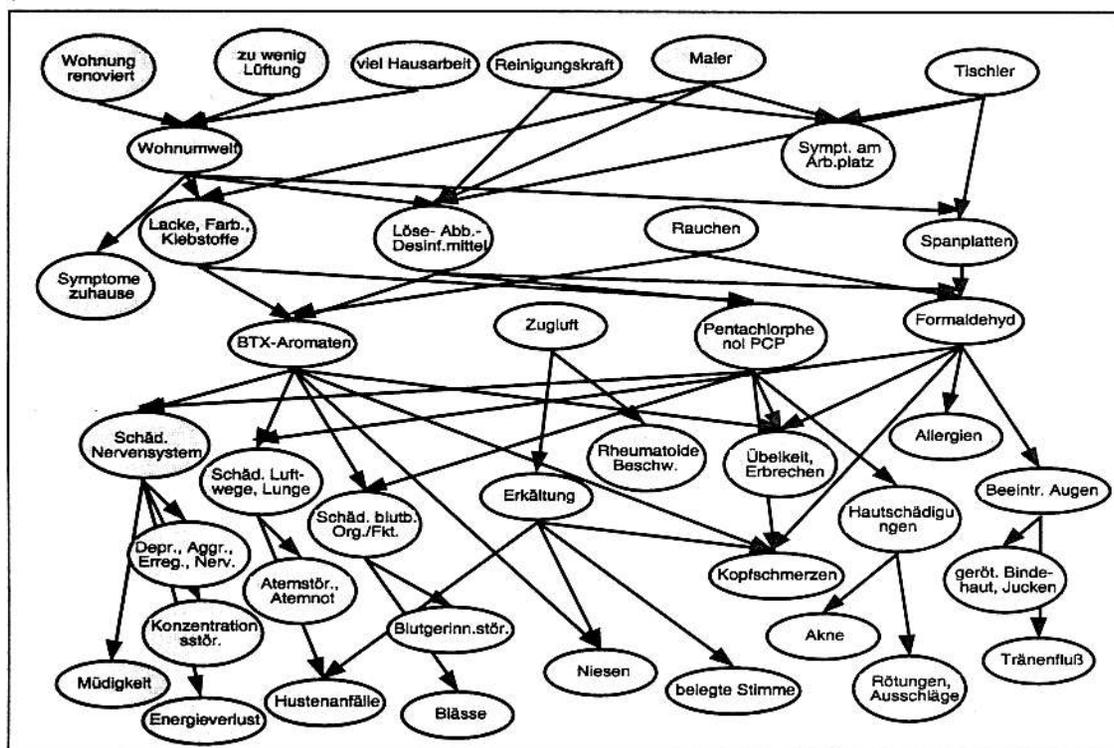


Abb.1: Der Graphische Editor: Ein Modell aus dem Bereich Umweltmonitoring

Falls dem Anwender die Diagnosen und Prognosen des Modells nicht transparent genug sind, kann er die *Erklärungskomponente* nutzen (LÜDTKE, 1997). Die Erklärungen erfolgen domänenunabhängig und basieren auf allgemeinen kognitionswissenschaftlichen Regeln des menschlichen Schließens und Erkennens. Es werden Fragen der Form: "Warum ist der Verdacht für ... bei den gegebenen Daten so stark angestiegen?" beantwortet.

3. Die quantitative Modellrevision

In MEDIKUS wird eine *quantitative Modellrevisionskomponente* integriert. Sie dient zur automatischen Anpassung der Modelle an reale, beobachtbare Daten bzw. an erwartete Modellaussagen. Bei fehlerhaften Prognosen oder Diagnosen seines Modells ist der Anwender nicht gezwungen, solange neue Spezifikationen einzugeben, bis sein Modell das reale Verhalten des modellierten Sachverhalts zeigt. Die quantitative Revisionskomponente paßt die vorhandenen Spezifikationen an die gegebenen Daten an.

Die quantitative Modellrevision basiert auf dem Konzept der minimalen relativen Entropie: bei der Erfüllung der Restriktionen bleiben die notwendigen Änderungen minimal. Inkonsistente, widersprüchliche Restriktionen werden dem Anwender zurückgemeldet. Die mathematischen Grundlagen bilden nichtlineare Optimierung und geometrische Programmierung (CSISZÁR, 1975; FLETCHER, 1987). Bei der relativen Entropie wird der Abstand einer gegebenen Verteilung P_0 von Zufallsvariablen und einer zweiten Verteilung P_1 , in der vorgegebene Restriktionen gelten, minimiert (siehe Appendix). Die Basis der Berechnung bildet die gemeinsame Verteilung der Zufallsvariablen. In BNs ist die gemeinsame Verteilung durch

apriori und bedingte Wahrscheinlichkeiten parametrisiert. Die Werte der Parameter werden bei der quantitativen Modellspezifikation eingegeben. In MEDIKUS erfolgt die Minimierung auf einem Cliquenbaum, einer Baumrepräsentation von BNs deren Knoten Cliquen genannt werden und eine (Teil-)Menge der BN-Knoten bzw. Zufallsvariablen enthalten (NEAPOLITAN, 1990). Die Basis der Berechnung bildet die gemeinsame Verteilung der Clique, in der die Randverteilung aller BN-Knoten einer Restriktion berechnet werden kann. Bei der Erfüllung der Restriktionen werden die notwendigen Änderungen der vom Modellierer spezifizierten apriori und bedingten Wahrscheinlichkeiten in MEDIKUS angezeigt. Bei unerwünschten bzw. fehlerhaften Abweichungen kann ein Modellierer die Wahrscheinlichkeit mit dem korrekten Wert der Restriktionsmenge hinzufügen. Durch Start der automatischen Modellrevision erhält der Modellierer ein Modell ohne die unerwünschten Abweichungen, das auch die ursprünglich vorgegebenen Restriktionen enthält. So unterstützt MEDIKUS Experten durch automatische quantitative Modellrevision mit Aufzeigen von Konsequenzen ihrer Revisionen bei der Entwicklung valider Modelle.

Abbildung 1 zeigt zur Illustration ein komplexes Modell aus dem Bereich des Umweltmonitoring. Dieses Modell umfaßt die von BTX-Aromaten, Phenolen und Formaldehyd ausgehenden Gefahren und Risiken. Am oberen Rand des Graphen befinden sich die Lebens- und Arbeitsumstände (z.B. "Wohnung renoviert" oder "Maler"), es folgen die dabei auftretenden Expositionen (z.B. "Lacke, Farb., Klebstoffe") und die daraus resultierende Krankheitsbilder, sowie die Syndrome, (z.B. "Schäd. Nervensystem"). Den unteren Rand des Graphen bilden die Symptome, unter denen die betroffenen Personen leiden (z.B. "Energieverlust" oder "Kopfschmerzen").

Bevor das Modell zur Bearbeitung echter Fälle herangezogen werden kann, muß gründlich getestet werden, ob die Diagnosen bzw. Prognosen des Modells korrekt sind. Bei auftretenden Abweichungen ist eine Berichtigung des Modells erforderlich. Die Ursache des Fehlverhaltens des Modells ist oft schwer nachzuvollziehen, falls der Sachverhalt noch nicht ausführlich erforscht worden ist. Alle quantitativen Spezifikationen des Modells können über Fern- und Nebenwirkungen die Modelldiagnosen und -prognosen beeinflussen. MEDIKUS stellt für diese Situationen eine automatische Modellrevisionskomponente zur Verfügung, die das unzulängliche Modell an (durch reale Daten) gegebene Restriktionen anpaßt.

In MEDIKUS gibt es drei Klassen von Restriktionen:

- Gleichungs-Restriktionen
- Ungleichungs-Restriktionen
- Vergleichende-Restriktionen

Die Restriktionen jeder Klasse können:

- einzelne Modellgrößen
- prognostische Einflüsse
- diagnostische Einflüsse

betreffen. Die prognostischen Einflüsse sind Ursache-Wirkungs-Beziehungen, die der Pfeilrichtung der Kanten des Graphen (Abb. 1) entsprechen. Die diagnostische Richtung führt von den Auswirkungen zu den Ursachen.

Abbildung 2 zeigt Gleichungs-Restriktionen in Form von Wahrscheinlichkeitsaussagen. Im oberen Teilbereich befinden sich je eine univariate, eine prognostische und eine diagnostische Restriktion. Alle drei sind lokal, d.h. sie sind entweder univariat oder sie betreffen nur

benachbarte, direkt verbundene Modellgrößen. In dem unteren Eingabebereich steht ein Beispiel für globale Restriktionen, da die Modellgrößen "Rauchen", "Müdigkeit" und "Hustenanfälle" nicht durch direkte Kanten verbunden sind.

Abbildung 3 zeigt Beispiele für Ungleichungs- und vergleichende Restriktionen. Mit den Ungleichungsrestriktionen können "vorsichtige" Restriktionen formuliert werden, z.B. die Aussage über den Zusammenhang zwischen auftretender Müdigkeit, Hustenanfällen und Rauchen (obere Hälfte). Durch vorsichtige Restriktionen werden keine punktgenauen Wertevorgaben für die Wahrscheinlichkeiten formuliert, sondern es werden Intervalle bestimmt, in denen die Wahrscheinlichkeitswerte liegen müssen. Vergleichende Restriktionen erlauben die Beschreibung von Beziehungen zwischen den Modellwahrscheinlichkeiten, z.B. bei vorliegenden Daten für Müdigkeit und Hustenanfälle steigt die Wahrscheinlichkeit des Rauchens um 0.3 an.

Eingabe von Restriktionen

Restriktionen für Netz #1

p(Formaldehyd = ja) = 0.3
 p(Formaldehyd = ja | Allergien = ja) = 0.4
 p(Spanplatten = ja | Tischler = nein) = 0.45

p(Rauchen = ja | Müdigkeit = ja, Hustenanfälle = ja) = 0.55

Buttons: Entfernen, Hinzufügen, OK

Abb.2: Univariate, prognostische, diagnostische und globale Gleichungs-Restriktionen

Eingabe von Restriktionen

Restriktionen für Netz #1

p(Rauchen = ja | Müdigkeit = ja, Hustenanfälle = ja) >= 0.5

p(Rauchen = ja | Müdigkeit = ja, Hustenanfälle = ja) >= p(Rauchen = ja) + 0.3

Buttons: Entfernen, Hinzufügen, OK

Abb.3: Ungleichungs- und vergleichende Restriktionen

Durch die vielfältigen Formulierungsmöglichkeiten für Restriktionen geht die automatische Revisionskomponente von MEDIKUS über bestehende vergleichbare Systeme wie zum Beispiel SPIRIT hinaus (MEYER, RÖDDER, 1996).

4. Kooperative Modellentwicklung

Bei kooperativer Modellentwicklung gibt die automatische Modellrevisionskomponente von MEDIKUS hilfreiche Unterstützung. Das Aufdecken und Anzeigen der intransparenten Konsequenzen von Eingriffen hilft Folgefehler und Mißverständnisse zu vermeiden. Im folgenden fiktiven Szenario entwickeln zwei Experten zur Illustration ein gemeinsames Modell. Als Anfangsmodell des ersten Experten (Abb.4) greifen wir ein Beispiel aus einem Standardlehrbuch über BNs auf (NEAPOLITAN, 1990). Das Modell beschreibt personenbezogen mögliche Zusammenhänge zwischen Tuberkulose und Lungenkrebs. Die Abkürzungen der Knoten des BN stehen für folgende Prädikate:

- a = "Hat die Person eine Asienreise unternommen?" (a1 = ja / a2 = nein)
- b = "Ist bei der Person Tuberkulose präsent?" (b1 = ja / b2 = nein)
- c = "Ist bei der Person Lungenkrebs oder Tuberkulose präsent?" (c1 = ja / c2 = nein)
- d = "Hat die Röntgenuntersuchung der Person ein positives Ergebnis?" (d1 = ja / d2 = nein)
- e = "Ist bei der Person Lungenkrebs präsent?" (e1 = ja / e2 = nein)
- f = "Raucht die Person Tabak?" (f1 = ja / f2 = nein)
- g = "Ist bei der Person Bronchitis präsent?" (g1 = ja / g2 = nein)
- h = "Ist (Sind) bei der Person Dyspnoe (Atembeschwerden) präsent?" (h1 = ja / h2 = nein)

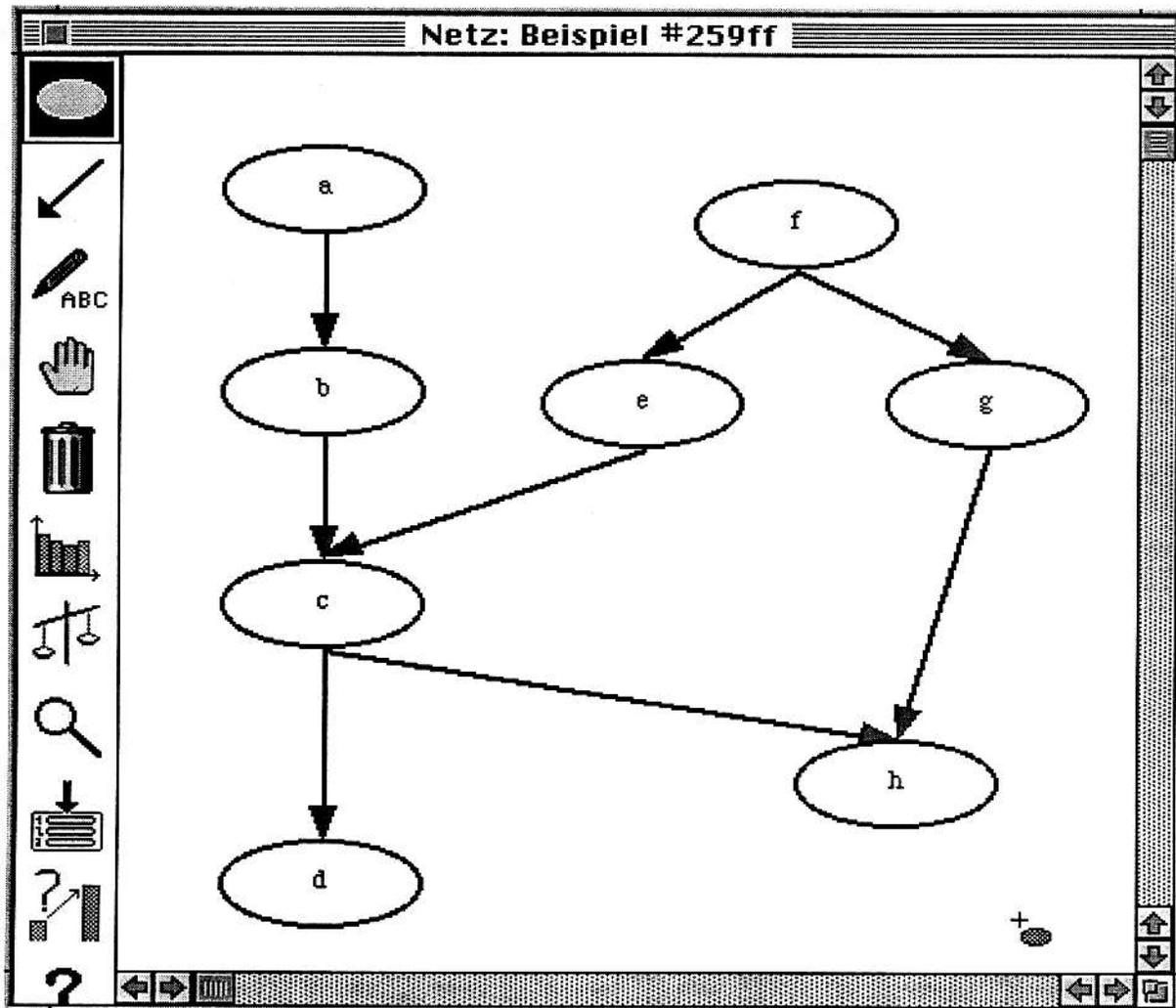


Abb.4: Das Anfangsmodell des ersten Experten

Das Modell liefert eine Auftretenswahrscheinlichkeit für Dyspnoe ($h=h_1$) von 0.436. Sein Kollege, ein zweiter Domänenexperte hat neuere Unterlagen nach denen dieser Wert (z.B. außerhalb von Ballungsräumen) zu hoch ist. Er sagt, daß die Auftretenswahrscheinlichkeit für Dyspnoe lediglich gleich 0.3 ist. Da dieser Wert dem ersten Experten jedoch zu niedrig erscheint, einigen die beiden sich nach einer kurzen Diskussion auf den Wert 0.35. Durch den Einsatz der automatischen Modellrevisionskomponente erzeugen die beiden ein revidiertes Folgemodell. Das Anfangsmodell wird iterativ an die Restriktion $p(h=h_1) = 0.35$ angepaßt, der Fortschritt der Restriktionsannäherung wird den Modellierern zurückgemeldet (Abb.5). Die Modellierer haben in jedem Iterationsschritt die Möglichkeit die Berechnung bei ausreichender Annäherung der Restriktionen zu stoppen. Die erreichten Näherungswerte bilden dann die gültigen Restriktionen. In dem Beispiel brechen die beiden Experten die Berechnung nach der dritten Iteration ab. Daraus ergibt sich die Restriktion: $p(h=h_1) = 0.3581$. Nach erfolgreicher Berechnung werden die Änderungen der Modellparameter angezeigt (Tab.1).

Revisionsbericht: Beispiel #259ff

1. Iteration
 Sind die erreichten Werte ausreichend? Ja Nein

Ziel: $p(h = h_1) = 0.35$

Aktueller Wert: $p(h = h_1) = 0.3910$

Revisionsbericht: Beispiel #259ff

3. Iteration
 Sind die erreichten Werte ausreichend? Ja Nein

Ziel: $p(h = h_1) = 0.35$

Aktueller Wert: $p(h = h_1) = 0.3581$

Abb.5: Iterativer Revisionsbericht

Modellparameter des Anfangsmodells entwickelt von dem ersten Experten.	Modellparameter nach der ersten Revision Erreichte Restriktion: $p(h=h_1) = 0.3581$
$p(h=h_1 c=c_1, g=g_1) = 0.9$ $p(h=h_1 c=c_1, g=g_2) = 0.7$ $p(h=h_1 c=c_2, g=g_1) = 0.8$ $p(h=h_1 c=c_2, g=g_2) = 0.1$ $p(h=h_2 c=c_1, g=g_1) = 0.1$ $p(h=h_2 c=c_1, g=g_2) = 0.3$ $p(h=h_2 c=c_2, g=g_1) = 0.2$ $p(h=h_2 c=c_2, g=g_2) = 0.9$	$p(h=h_1 c=c_1, g=g_1) = 0.829$ $p(h=h_1 c=c_1, g=g_2) = 0.5569$ $p(h=h_1 c=c_2, g=g_1) = 0.683$ $p(h=h_1 c=c_2, g=g_2) = 0.0564$ $p(h=h_2 c=c_1, g=g_1) = 0.1709$ $p(h=h_2 c=c_1, g=g_2) = 0.443$ $p(h=h_2 c=c_2, g=g_1) = 0.3169$ $p(h=h_2 c=c_2, g=g_2) = 0.9435$

Tab.1: Die Auswirkungen der automatischen Revision auf die Modellparameter

Der erste Experte ist mit der Änderung der bedingten Wahrscheinlichkeit $p(h=h1 | c=c1, g=g1)$ nicht einverstanden. Er sagt: Die Wahrscheinlichkeit für Dyspnoe gegeben Lungenkrebs und Tuberkulose und Bronchitis soll bei der Modellrevision nicht geändert werden.

Diese Aussage führt zu der neuen Restriktionsmenge:

$$\{p(h=h1) = 0.3581, p(h=h1 | c=c1, g=g1) = 0.9\}$$

Auch der zweite Experte findet bei der Durchsicht der geänderten Modellparameter eine Auswirkung der Modellrevision, die aus seiner Sicht nicht auftreten darf. Er ist mit der Änderung der bedingten Wahrscheinlichkeit $p(h=h2 | c=c2, g=g2)$ nicht einverstanden.

Die Aussagen der beiden Experten führen zu der Restriktionsmenge:

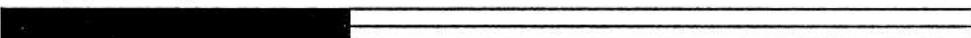
$$\{p(h=h1) = 0.35, p(h=h1 | c=c1, g=g1) = 0.9, p(h=h2 | c=c2, g=g2) = 0.9\}$$

Durch den wiederholten Einsatz der automatischen Modellrevisionskomponente erhalten die beiden Experten bereits nach der zweiten Iteration ein gemeinsames Modell, in dem die neuen Restriktionen zufrieden stellend angenähert sind (Abb.6). Das fiktive Szenario endet hiermit, bei verteilter Modellentwicklung in komplexen Bereichen unsicheren Wissens muß wiederholt geprüft werden, ob die notwendigen Auswirkungen auf die Modellparameter erwünscht sind.

Revisionsbericht: Beispiel #259ff		
2. Iteration		
Sind die erreichten Werte ausreichend?	Ja	Nein

Ziel: $p(h=h1 c=c1, g=g1) = 0.9$	<input checked="" type="radio"/>	<input type="radio"/>
		
Aktueller Wert: $p(h=h1 c=c1, g=g1) = 0.8976$		

Ziel: $p(h=h2 c=c2, g=g2) = 0.9$	<input checked="" type="radio"/>	<input type="radio"/>
		
Aktueller Wert: $p(h=h2 c=c2, g=g2) = 0.9022$		

Ziel: $p(h = h1) = 0.3581$	<input checked="" type="radio"/>	<input type="radio"/>
		
Aktueller Wert: $p(h = h1) = 0.3581$		

	<input type="button" value="OK"/>	

Abb.6: Die gemeinsamen Restriktionen der beiden Experten sind erfüllt

Durch das Aufzeigen der Konsequenzen der automatischen quantitativen Modellrevision begünstigt MEDIKUS im besonderen die kooperative Modellentwicklung. Durch die Integration der MEDIKUS-Komponente zur quantitativen Modellspezifikation (siehe oben) können linguistische Variablen anstelle von Wahrscheinlichkeitswerten bei der Formulierung der Restriktionen von den Domänenexperten verwendet werden. Dadurch wird die Expertendiskussion bei kooperativer Modellkonstruktion erleichtert.

5. Literatur

CSISZÁR, I., I-Divergence Geometry of Probability Distributions and Minimization Problems, in *The Annals of Probability*, 1975

FLETCHER, R., *Practical Methods of Optimization*, New York: John Wiley & Sons, 1987

FOLCKERS, J., MÖBUS, C., SCHRÖDER, O., THOLE, H.-J., An Intelligent Problem Solving Environment for Designing Explanation Models and for Diagnostic Reasoning in Probabilistic Domains, in C. FRASSON, G. GAUTHIER, A. LESGOLD (eds), *Intelligent Tutoring Systems, Proceedings of the Third International Conference ITS 96*, Montreal, June 1996, Berlin: Springer (LNCS 1086), 1996

LÜDTKE, A. Abstraktionsbasierte Erklärungen für Prognose und Diagnose mit Bayes-Netzen, in: C. Herzog (ed.), *Beiträge zum 8. Arbeitstreffen der GI-Fachgruppe 1.1.5/7.0.1 "Intelligente Lehr- und Lehrsysteme"*, Duisburg, 18.-19. September 1997, "Blaue Berichte" der TU München, 1997

MEYER, C.-H., & RÖDDER, W., Probabilistic Knowledge Representation and Reasoning at Maximum Entropy by SPIRIT, in *KI-96: Advances in Artificial Intelligence Proceedings of the 20th Annual German Conference on Artificial Intelligence* (Ed.: S. Hölldobler), Berlin: Springer, 1996

MÖBUS, C., SCHRÖDER, O., Building Domain Models by Novices in Stochastics: Towards the Probabilistic Semantics of Verbalized Stochastic Relations, in: B. Boulay, R. Mizoguchi (eds.): *Artificial Intelligence in Education*, Amsterdam: IOS Press, 1997a

MÖBUS, C., SCHRÖDER, O., Unterstützung der Bildung stochastischer Modelle: Von qualitativen verbalen Relationsbeschreibungen zu quantitativen Beziehungen, in: W. Krause, U. Kotkamp, R. Goertz (eds.): *KogWis97, Proceedings der 3. Fachtagung der Gesellschaft für Kognitionswissenschaft*, Jena: Friedrich-Schiller-Universität, 1997b

NEAPOLITAN, R. E.: *Probabilistic Reasoning in Experts Systems*, New York: John Wiley & Sons, 1990

PEARL J., *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, San Mateo: Morgan Kaufman, 1988

SCHRÖDER, O., MÖBUS, C., FOLCKERS, J., THOLE, H.-J., Supporting the Construction of Explanation Models and Diagnostic Reasoning in Probabilistic Domains, In D.C. Edelson, E.A. Domeshek (eds): *Proceedings of the International Conference on the Learning Sciences ICLS 96*, Charlottesville, VA: AACE, 1996

SCHRÖDER, O., MÖBUS, C., THOLE, H.-J., Acquiring Knowledge from Linguistic Models in Complex, Probabilistic Domains, in P. BRNA, A. PAIVA, J. SELF (eds), *Artificial Intelligence in Education, Proceedings of the European Conference on AI in Education 1996*, Lisbon, Portugal, Sept. 30 - Oct. 2, 1996

SCHRÖDER, O., THOLE, H.-J., WILLMS, J., Zur Modellierung und Unterstützung des kooperativen und kompetitiven Wissenserwerbs mit Bayes-Agenten, in: M. Thielscher, S.-E. Bornscheuer (Hrsgb), *Fortschritte der künstlichen Intelligenz, Aus den Workshops der 20. Deutschen Jahrestagung für Künstliche Intelligenz*, Dresden, 17.-19. 9. 1996, Dresden University Press, 1996

6. Appendix

Definition der relativen Entropie:

Für zwei gemeinsame Verteilungen P_0 und P_1 mit n Komponenten ist die relative Entropie von P_1 in Bezug auf P_0 definiert durch:

$$R(P_0, P_1) = \sum_{i=1}^n p_1(v_i) \cdot \log \left(\frac{p_1(v_i)}{p_0(v_i)} \right)$$

Berechnung der Verteilung P_1 :

Gegeben die gemeinsame Verteilung P_0 (eines BNs bzw. einer Clique eines BN)

Gegeben die Restriktion: $p(a=a_1 | b=b_1) = x$ (wobei a und b Knoten des BN sind)

$$p_1(v_i) = \begin{cases} p_0(v_i) \cdot \sigma & , v_i = (\dots, b \neq b_1, \dots) \\ p_0(v_i) \cdot \sigma \cdot \alpha^x & , v_i = (\dots, a \neq a_1, \dots, b = b_1, \dots) \\ p_0(v_i) \cdot \sigma \cdot \alpha^{1-x} & , v_i = (\dots, a = a_1, \dots, b = b_1, \dots) \end{cases}$$

mit:

$$\alpha = \frac{x}{1-x} \cdot \frac{p_0(a \neq a_1, b = b_1)}{p_0(a = a_1, b = b_1)}$$

σ ist der Normalisierungsfaktor

**Intelligente Unterstützung in offenen
Lehr- / Lernumgebungen ?**

Olaf Schröder (Hrsg.)

TZI-Bericht Nr. 8, 1998

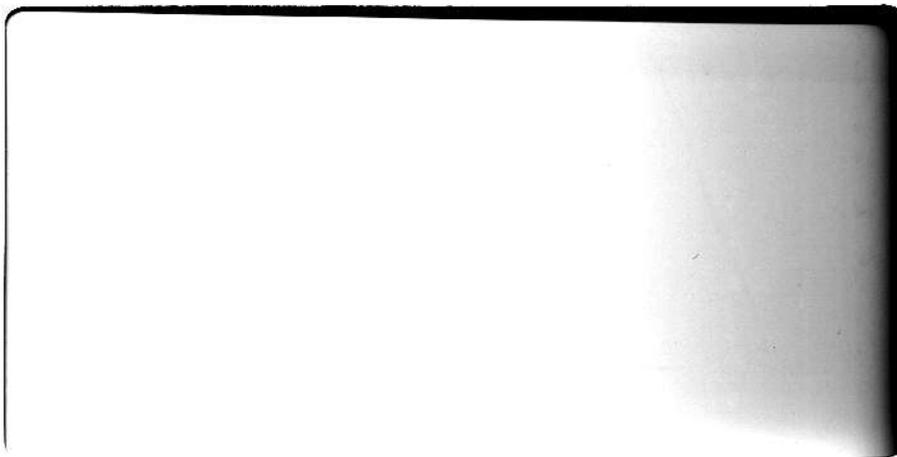
*Von der Theorie zur Praxis:
Entwicklung und Transfer
von Informatik-Technologien*

**Intelligente Unterstützung in offenen
Lehr- / Lernumgebungen ?**

Olaf Schröder (Hrsg.)

TZI-Bericht Nr. 8, 1998

Beiträge zum Workshop auf der 22. Jahrestagung Künstliche Intelligenz (KI-98)
vom 15.-17. September in der Universität Bremen



Ansprechpartner

Sprecher des TZI

Prof. Otthein Herzog

Geschäftsführer des TZI

Dr. Andreas Günter

Technologie-Zentrum Informatik
FB Mathematik und Informatik
Universität Bremen
Postfach 330440
D - 28334 Bremen

Telefon: +49 421 218-7272/-7090/-2894

Fax: +49 421 218-7196

E-Mail: tzi@informatik.uni-bremen.de

www: <http://www.informatik.uni-bremen.de/grp/tzi>



Technologie-Zentrum Informatik