

# Modellrevision in vernetzten, unsicheren Wissensgebieten

**Projektleiter** Prof. Dr. C. Möbus  
**Ansprechpartner** H.-J. Thole  
**Tel.** (04 41) 97 06-1 73  
**E-Mail** heinz-juergen.thole@offis.uni-oldenburg.de  
**Laufzeit** 07/1994 bis 06/1998

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. Prognosen, Diagnosen 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 Software-Werkzeugs zur Unterstützung der Modellkonstruktion und bei der Entwicklung geeigneter Planungs- und Diagnosestrategien. Das Akronym MEDIKUS steht für Modellierung, Erklärung und Diagnoseunterstützung bei komplexen, unsicheren Sachverhalten. 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.

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

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

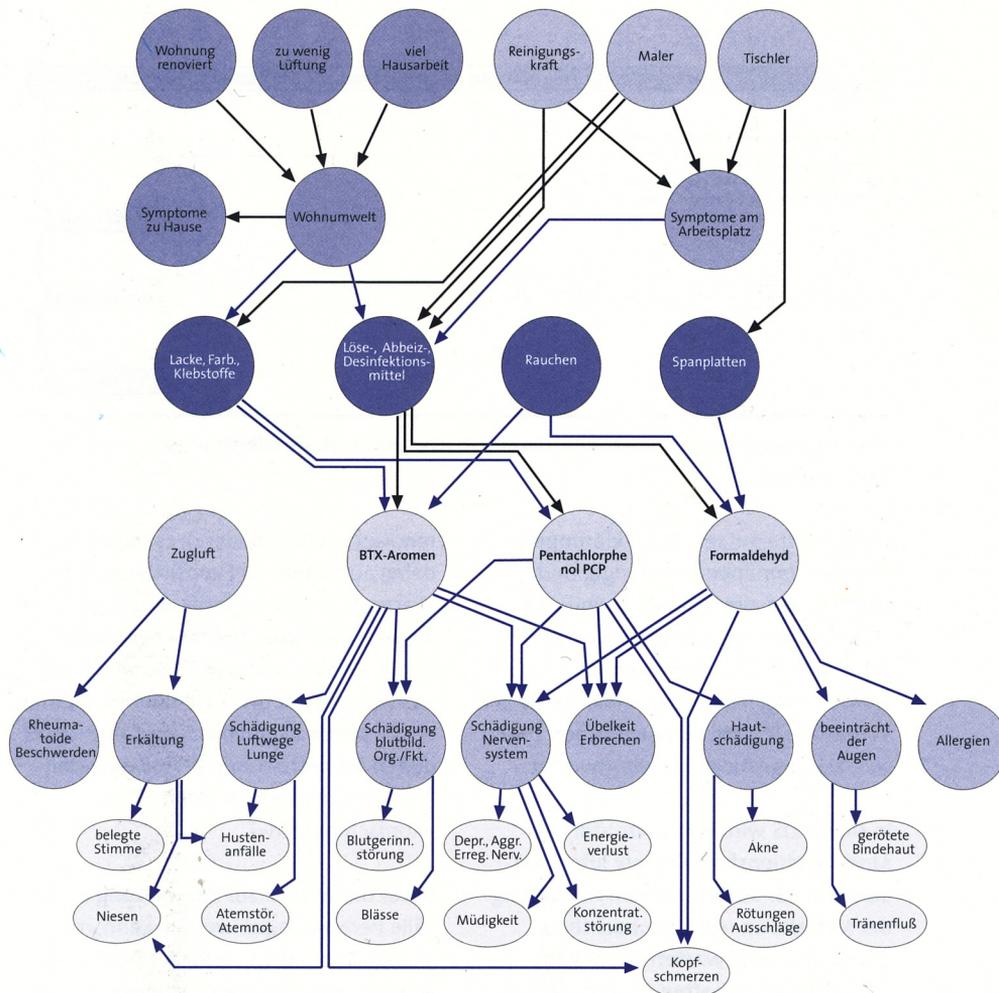


Abb. 1: Der graphische Editor: Ein komplexes Modell aus dem Bereich Umweltmonitoring

gerichtete azyklische Graphen, konstruiert (Abb. 1). 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. 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. 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.

Falls dem Anwender die Diagnosen und Prognosen des Modells nicht transparent genug sind, kann er die *Erklärungs-*



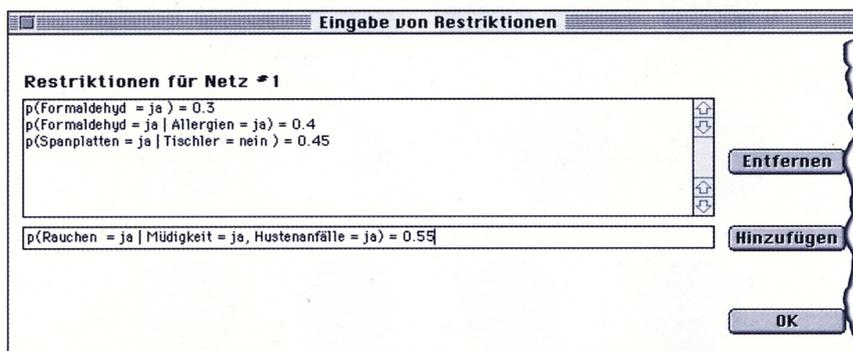


Abb. 2: Univariate, prognostische, diagnostische und globale Gleichungs-Restriktionen

komponente nutzen. Die Erklärungen erfolgen domänenunabhängig und basieren auf allgemeinen kognitions-wissenschaftlichen Regeln des menschlichen Schließens und Erkennens. Es werden Fragen der Form: »Warum ist der Verdacht für ... bei den gegebenen Daten so stark angestiegen?« beantwortet.

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.

Diese Revision basiert auf dem Prinzip der minimalen relativen Entropie: bei der Erfüllung der Restriktionen bleiben die notwendigen Änderungen minimal. Inkonsistente, widersprüchliche Restriktionen werden dem Anwender zurückgemeldet.

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, Farben, Klebstoffe«) und die daraus resultierende Krankheitsbilder, sowie die Syndrome, (z. B. »Schädigung des Nervensystems«). 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 sowie
  - vergleichende Restriktionen.
- Die Restriktionen jeder Klasse können
- einzelne Modellgrößen,
  - prognostische Einflüsse und
  - diagnostische Einflüsse

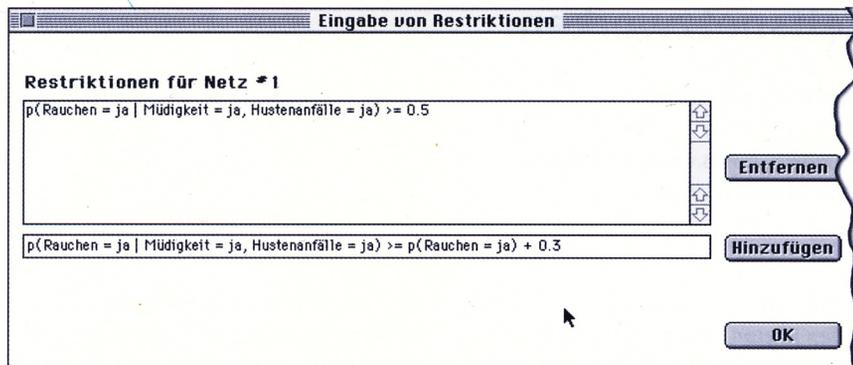


Abb. 3: Ungleichungs- und vergleichende Restriktionen

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 Constraints formuliert werden, z. B. die Aussage über den Zusammenhang zwischen auftretender Müdigkeit, Hustenanfällen und Rauchen (obere Hälfte). 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.

Aktuelle Anwendungsfelder von MEDIKUS sind außer den oben genannten z. B. die Wissensstanddiagnose in CBT-Systemen, Optimierungen im Angebotswesen, Konfiguration und Datenverwaltung in Betrieben des verfahrenstechnischen Aggregatebaues (z. B. in der Lebensmitteltechnik). Geplant ist eine Zusammenarbeit mit dem DIL (Deutsches Institut für Lebensmitteltechnik) im Rahmen eines EU-Projekts, in das Aggregatehersteller aus dem Bereich der Misch-, Dispergier- und Pumptechnik als Industriepartner eingebunden werden.

In dem aus EU-Mitteln (ESF) beantragten ADAPT-Projekt ISO\ESPC wird die um Einflußdiagramme erweiterte Modellierungsumgebung von MEDIKUS zur Planung und zur Schulung der statistischen Prozeßkontrolle zur Zertifizierung nach ISO 9000 ff. eingesetzt werden. Einflußdiagramme bilden eine Fortentwicklung von Bayes-Netzen, die die explizite Modellierung von Entscheidungsstrategien und ihren Bewertungen unterstützt. Das Projekt hat einen nationalen Partner (BBZ Fulda) und ist in ein transnationales Konsortium von fünf europäischen Partnern eingebunden. ←

