

LERNER-INFERENZEN DURCH KONSTRUKTION BAYES'SCHER NETZE AUS DOMÄNENWISSEN

István-Tibor Nébel

Tagesklinik für kognitive Neurologie
Universität Leipzig
Liebigstr. 22a
D-04103 Leipzig
nebt@medizin.uni-leipzig.de

Abstract: Der Erfolg von e-Learning-Systemen hängt entscheidend vom Grad der Lerneradaptivität ab. Adaptive tutorielle Systeme, die sich in einem bestimmten Maße an den Lerner anpassen können, benötigen hierfür Verfahren des maschinellen Lernens. Dieser Beitrag stellt ein Verfahren vor, das mit Hilfe von Domänenwissen über einen Lerner Bayes'sche Netze konstruiert, um Inferenzen über den Lerner bilden zu können. Diese zusätzlichen Lernerinformationen werden dem Adaptionsprozess zur Verfügung gestellt, um eher und qualitativ bessere Inferenzen bilden zu können.

1 Motivation

Grundlage der Adaptivität in tutoriellen Systemen sind die Informationen über den Lerner und die Inferenzen, die aus diesen Informationen gebildet werden. Aus diesen Informationen resultieren individuelle Anpassungen des Systems an den einzelnen Lerner. Meist werden die Lerner-Informationen jedoch erst nach mehreren Interaktionsequenzen durch verschiedene Verfahren der Lerner-Modellierung erlangt. In der Arbeit des Autors [Ne05] wird ein Rahmensystem vorgeschlagen, das den Adaptionsprozess in intelligenten tutoriellen Systemen unterstützt, indem domänenspezifische Lernerinformationen einem adaptiven Anwendungssystem angeboten werden.

Die hierfür entwickelte Methode zur Lerner-Modellierung zieht Domänenwissen heran, um das Lernermodell induktiv anzureichern. Das Rahmensystem ist in der Lage, die funktionalen Abhängigkeiten von Konzepten und deren Attributen einer Domäne mit dem vorhandenen Wissen über den Lerner abzugleichen.

Die zu modellierenden Domänen beinhalten im hohen Maße unsicheres und unvollständiges Wissen. Es musste also eine Methode gefunden werden, um mit unsicherem Wissen über den Lerner aus Domänenwissen Schlussfolgerungen ziehen zu können. Hierfür eignen sich bekannterweise Bayes'sche Netze.

Dieser Beitrag behandelt die Konstruktion dieser Netze aus Domänenwissen, um neue Lernereigenschaften zu akquirieren und sie dem Adaptionsprozess zur Verfügung zu stellen.

Patienten sind hochmotivierte Lerner. Aus diesem Grund wurde eine medizinische Domäne – der Diabetes mellitus – gewählt und daran exemplarisch die Integration des Domänenwissen über den Diabetes mellitus in das Lerner- und Patientenprofil demonstriert.

2 Bayes'sche Netze

Ein geeigneter Ansatz Schlussfolgerungen unter Unsicherheit zu ziehen, sind Bayes'sche Netze. Dieser probabilistische Inferenzalgorithmus ermöglicht die systematische Berücksichtigung und differenzierte Darstellung von Unsicherheit [Wi02]. Bayes'sche Netze eignen sich im besonderem Maße für die individuelle Benutzer- bzw. Lernermodellierung, da individuelle Vorhersagen anhand benutzerspezifischer Eingangsparameter getroffen werden können [Sc98].

Ein Bayes'sches Netz $B = (G, \Theta)$ für eine Menge $X = X_1, \dots, X_n$ von Zufallsvariablen besteht aus zwei Teilen:

1. Einem gerichteten azyklischen Graphen $G = (X, E)$, dessen Knoten den Zufallsvariablen beziehungsweise den Domänenkonzepten entsprechen und mit dessen Kanten die bedingten Unabhängigkeiten zwischen den Variablen kodiert werden. Man spricht von G als der Struktur von B .
2. Einer Menge $\Theta = \Theta_1, \dots, \Theta_n$ von mit den Variablen assoziierten Tabellen bedingter Wahrscheinlichkeiten (CPT) $\Theta_i = P(X_i | Pa(X_i))$, $i = 1, \dots, n$. Sie beinhalten als Einträge die bedingten Wahrscheinlichkeiten $\Theta_{ijk} = P(x_{ij} | pa_k(X_i))$ der n_i Zustände x_{ij} , $j = 1, \dots, n_i$, der Variablen X_i konditioniert auf die möglichen Zustandskombinationen $pa(X_i)$ der Eltern $Pa(X_i)$. Mit $pa_k(X_i)$ wird die k -te der Zustandskombination $pa(X_i)$ der Eltern bezeichnet. Besitzt ein Knoten keine Eltern, dann beinhaltet seine CPT unbedingte A-priori-Wahrscheinlichkeiten $P(x_{ij})$, d.h., $\Theta_i = P(X_i)$.

In Bayes'sche Netzen können die Lernereigenschaften in Form von Wahrscheinlichkeitsverteilungen innerhalb der CPT dargestellt werden. Die Wahrscheinlichkeit gibt an, wie sicher eine Eigenschaft zutrifft. Das Netz verwaltet, gegenüber anderen Ansätzen, die Sicherheiten mit der aufeinanderfolgende Inferenzen gezogen werden.

In einem Bayes'schen Netz wird die gemeinsame Wahrscheinlichkeitsverteilung eines Knotens als Produkt lokaler Wahrscheinlichkeitsverteilungen der Elterknoten repräsentiert:

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | Pa(X_i)) \quad (1)$$

Mit Hilfe der Formel 1 lässt sich für jede beliebige Parameterausprägung (X_1, \dots, X_n) die Eintrittswahrscheinlichkeit von voneinander unabhängigen Zufallsvariablen berechnen.

3 Architektur und Arbeitsweise

Die Referenzdomäne soll im Folgenden aus einen Teilbereich der Medizin – dem Diabetes mellitus – modelliert werden. Für die Modellierung von Domänen und Domänenwissen kommen Ontologien ([Ka03], [Gr93]) zur Repräsentation des Domänenwissens und Berechnungsregeln zur Definition von funktionalen Abhängigkeiten zwischen den Domänenkonzepten zum Einsatz.

3.1 Berechnungsregeln

Funktionale Abhängigkeiten zwischen den Domänenkonzepten werden in Form von Berechnungsregeln definiert. Sprachen wie RuleML, TRIPLE, DLML sind gegenwärtig nicht ausdrucksstark genug, um beispielsweise mathematische Funktionen zu definieren. Infolge dessen wurde ein eigenes Format, das sich an DLML anlehnt entwickelt, das es ermöglicht, funktionale Abhängigkeiten zwischen den Domänenkonzepten zu definieren. Die Spezifikation dieses Formates enthält Funktionen zum Definieren von:

- mathematischen Zusammenhängen,
- Fuzzy-Zusammenhängen,
- probabilistischen Zusammenhängen.

Interessant sind im Kontext dieses Artikels die *probabilistischen Zusammenhänge*. Hier können Studienergebnisse einfließen, die besonders in der betrachteten Domäne in hoher Zahl anzutreffen sind. Da es sich meist um Aussagen mit einer bestimmten Wahrscheinlichkeit handelt, können im besonderem Maße die oben beschriebenen Bayes'sche Netze angewendet werden, die die Verarbeitung und Inferenzbildung unter Unsicherheiten erlauben. Beispielsweise besagt eine Studie von Matthaei [Ma00], dass 90% aller Typ 2-Diabetiker über 60 Jahre übergewichtig sind. Diese Aussage kann in geeigneter Form innerhalb der Berechnungsregeln notiert werden:

```
1: <dl:DEF>$gewichtsklassifikation.  
2:   <dl:CASE bool=AND ofValue=$diabetes-typ.,$alter.>  
3:   var(Typ2),range(>60,+INF) : uebergewichtig(0.9);  
4:   </dl:CASE>  
5: </dl:DEF>
```

3.2 Inferenzbildung

Dieser Abschnitt beschreibt die Konstruktion der Bayes'schen Netze mit Hilfe

- der Lernereigenschaften, die im Lernerprofil hinterlegt sind,
- des in Domänen-Ontologien repräsentierten Domänenwissens und
- der Berechnungsregeln, die Attribute bzw. Lernereigenschaften aus funktionalen Abhängigkeiten zwischen Domänenkonzepten definieren.

Die prinzipielle Arbeitsweise des entwickelten Rahmensystems berechnet in einem iterativen Prozess alle berechenbaren Attribute eines Lerner dieser Domäne und stellt die Inferenzen dem Lernerprofil zur Verfügung. Das Lernerprofil wird somit maximal mit neuen domänenspezifischen Attributen beliefert.

3.3 Konstruktion Bayes'scher Netze

Anhand von drei aufgeführten Studienergebnissen soll im Folgenden die Konstruktion Bayes'scher Netze beispielhaft erläutert werden¹:

Studie 1: *8% der über 60-jährigen sind Diabetiker.*

Studie 2: *80% der Diabetiker haben einen Typ-2-Diabetes.*

Studie 3: *90% aller Typ-2-Diabetiker über 60 Jahre sind übergewichtig.*

In folgender Tabelle 1 sind die explizit verfügbaren Lernerprofil-Daten und die Inferenzen des Rahmensystems, im Vorgriff auf den zu erläuternden Berechnungs- und Konstruktionsprozess, dargestellt.

Lernerprofil-Daten		Inferenzen des Rahmensystems				
Fakt	Wert	Attribut	Wert	Konfidenz	Regel	Studie
Geburtsdatum	23.06.1934	Alter	71	1,0000	1	-
		Diabetes	true	0,0800	4	1
		Diabetes-Typ	Typ 2	0,0640	2	2
		Gewichtsklassifikation	übergewichtig	0,0576	3	3

Tabelle 1: Lernerprofil mit vorhandenen und inferierten Lernereigenschaften

Den Inferenzen des Rahmensystems aus Tabelle 1 lagen folgende Regeln zugrunde:

Regel 1:
`<dl:DEF>$alter.
 <dl:CALC>SYSTEM$today(Y).- $geburtsdatum.
 </dl:CALC>
 </dl:DEF>`

Regel 2 / Studie 2:
`<dl:DEF>$diabetes-typ.
 <dl:CASE ofValue=$diabetes.>
 true:Typ2(0.8);
 true:Typ1(0.2);
 </dl:CASE>
 </dl:DEF>`

Regel 3 / Studie 3:
`<dl:DEF>$gewichtsklassifikation.
 <dl:CASE bool=AND ofValue=$diabetes-typ., $alter.>
 var(Typ1), range(0,+INF):normalgewichtig(0.75);
 var(Typ2), range(>60,+INF):übergewichtig(0.9);
 </dl:CASE>
 </dl:DEF>`

Regel 4 / Studie 1:
`<dl:DEF>$diabetes.
 <dl:CASE ofValue=$alter.>
 range(>60,+INF):true(0.08);
 </dl:CASE>
 </dl:DEF>`

¹Die realen Lernerprofil-Daten sind weitaus umfangreicher, als in diesem Beispiel angegeben.

Das Attribut *Alter* wird aufgrund der in Regel 1 definierten Systemfunktion berechnet, indem das aktuelle Datum vom Geburtsdatum subtrahiert wird. Mit Regel 4 wird das Studienergebnis 1 definiert: Da das Alter über 60 Jahre liegt, kann auf einen *Diabetes* geschlossen werden. Die Regel 2 repräsentiert das Studienergebnis 2, indem der *Diabetes-Typ* inferiert werden kann. Nunmehr kann Regel 3 angewendet werden, die, aufgrund von Studie 3, aus *Diabetes-Typ* und *Alter* mit einer bestimmten Sicherheit die *Gewichtsklassifikation* Übergewicht errechnet.

Die Konstruktion des Bayes'schen Netzes erfolgt demnach sukzessive aus den zu berechnenden Attributen. Welches Attribut dabei Eltern- und/oder Kindknoten repräsentiert, ist von der Regeldefinition innerhalb der Regelbeschreibungssprache abhängig. Das resultierende Netz für die Berechnung der *Gewichtsklassifikation* ist in Abbildung 1 visualisiert.

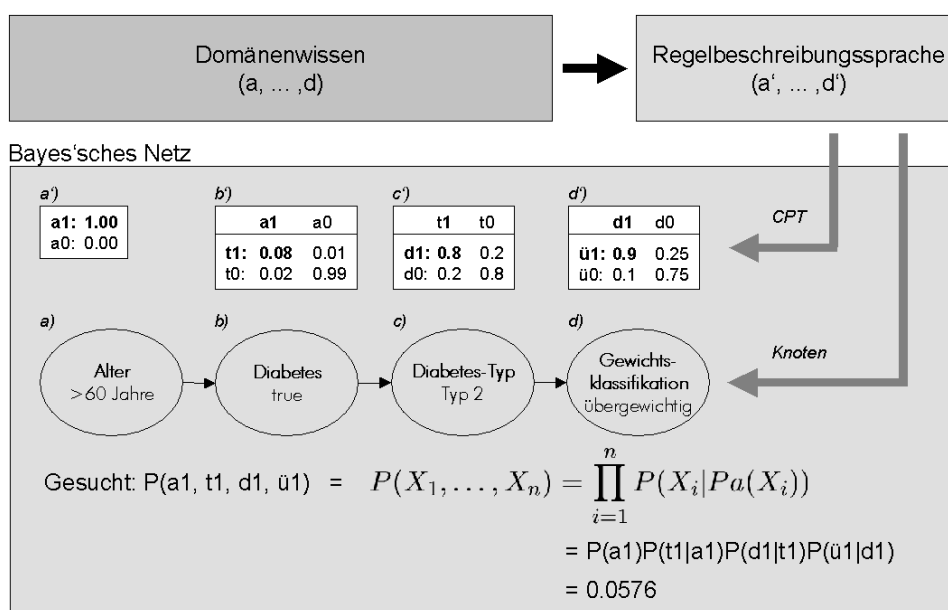


Abbildung 1: Konstruktion eines Bayes'schen Netzes aus Domänenwissen

Die konkrete Berechnung der *Gewichtsklassifikation* erfolgt nach (1) mit:

$$P(a1)P(d1|a1)P(t1|d1)P(u1|t1) = 1,0 * 0,008 * 0,8 * 0,9 = 0,0576 \quad (2)$$

Die berechnete Konfidenz von 0,0576 kann insofern interpretiert werden, als das der Lerner mit rund 5,8% als übergewichtig eingeschätzt wird. Die berechneten Inferenzen des Bayes'schen Netzes werden in das Lernerprofil (siehe Tabelle 1) eingetragen und stehen fortan für die weitere Verarbeitung zur Verfügung. Das nachfolgende XML-Skript spiegelt das Lernerprofil wider, indem in Abschnitt A die originalen Lernerprofil-Daten und in Abschnitt B das zusätzlich eingefügte Domänenwissen aufgeführt ist. Über einen Zugriff des adaptiven e-Learning-Systems auf die instanziierten XML-Attribute erlangt das System Kenntnis über domänenspezifische

Lernereigenschaften. Die Interpretation dieser Lernereigenschaften obliegt der didaktischen Konzeption des e-Learning-Systems und wird deshalb nicht an dieser Stelle eingehender diskutiert.

```

<?xml version="1.0"?>
<!--Abschnitt A: Lerner Profil Eintraege-->
<root>
  <user uid='0'>
    <name>uebergewichtige Testperson</name>
    <geburtsdatum>23.06.1934</geburtsdatum>
    ...
  </user>
  ...
</root>

<!--Abschnitt B: Domaenenwissen-->
<context>
  ...
  <Spezifik>
    ...
    <alter>71</alter>
    <diabetes>true(0.08)</diabetes>
    <diabetes-ty>Typ2(0.064)</diabetes-ty>
    <gewichtsklassifikation>uebergewichtig(0.0576)</gewichtsklassifikation>
    ...
  </Spezifik>
</context>

```

4 Summary

Wenn bekannt ist, dass sich ein Lerner eines tutoriellen Systems in einer abgeschlossenen Domäne bewegt, liegt es Nahe diese Domäneninformationen dem Adaptionprozess eines adaptiven Systems zuzuführen. Der Artikel beschreibt, wie diese Domäneninformationen mittels Bayes'schen Netzen genutzt werden können, um Inferenzen über einen Lerner zu bilden. Hierzu werden die funktionalen Abhängigkeiten zwischen Domänenkonzepten zur Konstruktion der Bayes'schen Netze verwendet.

In der zugrundeliegenden Arbeit des Autors [Ne05] für diesen Artikel wurde ein Rahmensystem entwickelt, das die inferierten Lernereigenschaften in das Lernerprofil einer adaptiven Anwendung zurücküberträgt. Im Folgenden werden einige Ergebnisse dieser Arbeit genannt, deren jeweilige Methodik nicht in diesem Artikel aufgeführt ist².

Das exemplarisch konzeptionierte Domänenwissen beinhaltet das Wissen über Patienten mit Diabetes mellitus. Dieser medizinische Kontext des Lerners ist in einer Domänen-Ontologie abgebildet, die darüber hinaus Berechnungsregeln zum Inferieren weiterer Lernereigenschaften enthält. Diese Berechnungsregeln definieren die funktionalen Abhängigkeiten zwischen den Domänen-Konzepten. Eine neu entwickelte Beschreibungssprache bietet Funktionen zum Definieren von mathematischen, Fuzzy- und probabilistischen Zusammenhängen, die zwischen den Konzepten existieren können. Ein Inferenz-Modul versucht, das Lernerprofil maximal mit

²Der Leser sei zur detaillierten Beschreibung des Rahmensystems, des medizinischen Anwendungssystems und zur Evaluierung auf die Originalarbeit des Autors [Ne05] verwiesen.

neuen Eigenschaften zu versorgen, indem mittels der Berechnungsregeln und den Attributen des Lernerprofils alle verfügbaren Attribute berechnet werden. Dieser Prozess terminiert, wenn keine weitere Regel wahrscheinlichere Attribute liefert, als die, die bereits instanziiert sind. Für die Weiterverarbeitung von probabilistischen Attributen, beispielsweise aus Studienergebnissen, kommt das automatisch konstruierte Bayes'sche Netz zum Einsatz. Das Netz berechnet für einen definierten Zusammenhang zwischen den Attributen die Konfidenz, die im bestem Fall die Wahrscheinlichkeit des Auftretens sein kann.

Im praktischen Einsatz hat sich die entwickelte Methode bewährt. Für ein konkretes, im Einsatz befindliches, adaptives e-Learning-System zum Thema der Unterzuckerungen bei Patienten mit Diabetes mellitus wurde das Rahmensystem eingesetzt. In einer vergleichenden Evaluierung mit 120 Diabetes-Patienten zeigte sich, dass die generierte Adaptivität in diesem medizinischen Anwendungssystem durch die Inferenzen des Rahmensystems unterstützt werden konnte, indem die implementierten Adaptionstechniken von den zusätzlichen domänenspezifischen Lerneigenschaften profitierten.

5 Literaturverzeichnis

- [Gr93] Gruber, T.R.: A Translation Approach to Portable Ontology Specifications, *Knowledge Acquisition* 5(2), 1993; S. 199-220
- [Ka03] Kalfoglou, Y.; Schorlemmer, M.: Ontology mapping: The state of the art, *Knowledge Engineering Review*, 18(2), 2003
- [Ma00] Matthaei, S. et. al.: Pathophysiology and pharmacological treatment of insulin resistance, *Endocrine Review*, 21(6), 2000; S. 585-618
- [Ne05] Nebel, I.T.: Domänenspezifische Benutzermodellierung für intelligente tutorielle Systeme im medizinischen Kontext, Dissertation, Christian Albrechts Universität zu Kiel, 2005
- [Sc98] Schäfer, R.: Benutzermodellierung mit dynamischen Bayes'schen Netzen als Grundlage adaptiver Dialogsysteme, Dissertation, Universität des Saarlandes, 1998
- [Wi02] Wittig, F.: Maschinelles Lernen Bayes'scher Netze für benutzeradaptive Systeme, Dissertation, Universität des Saarlandes, 2002