

## Prüfung über **Datenbanken**, **Künstliche Intelligenz** und **Knowledge Representation**.(praktische)

**Datenbanken:** nach Kemper/Eickler 1-15. **KI & KR** nach Vorlesung bzw. Büchern von Norvig  
AI: A modern approach, 1. (!) Auflage sowie Levesque/Brachmann: Knowledge Representation  
& Reasoning

Datum: 11.04.06

Note: 1,3

~ 50 min.

### **Datenbanken:**

*Sie hatten ja verschiedene Techniken zur Modellierung kennen gelernt, Entity-Relationship sowie UML. Wie funktionieren die denn und worin unterscheiden sie sich?*

Erstmal ER erklärt: Entities, Relationen, Rollen, Attribute der Entities und Keys. Sichten kurz erwähnt. Bei UML stand ich erstmal auf dem Schlauch (ich hatte XML verstanden und verwirrt, weil das ja erst sehr weit hinten im DB Buch ist ;-)) Dann hab ichs doch noch gemerkt und eben UML aus der objektorientierten Softwareentwicklung, Methoden und Attribute in eine Entität bzw. Klasse usw..

Er wollte dann noch auf ein spezielles Feature von UML raus, ich habe Use Cases erwähnt, er wollte aber etwas von UML-Interaktionsdiagrammen hören, da hab ich ziemlich rumgerätselt.

*Wie wandelt man denn so eine n:m Relation im ER-Modell in das relationale Schema um?*

Erklärt, alle Schlüssel der n Entity + alle Schlüssel der m Entity + Attribute der Relation, neuer Schlüssel setzt sich aus allen vorherigen Schlüsseln zusammen, bei 1:1 kann man den Schlüssel wählen, 1:N der von N Ent.

*Gegeben folgendes Relationsschema: "Verein Trainer Spiel Datum". Ist das ein gutes Relationsschema?*

Fkt. Abh.bestimmen, (Hat ein Verein mehrere Trainer? Jugendmannschaften, 1. Liga usw., hier nein!) Verein -> Trainer, Trainer -> Verein, Schlüssel ist aber Verein, Datum. Trainer aus NichtPrim und fkt. abh von Verein -> nicht in 2NF. Ich hatte erst den Superschlüssel mit Verein, Spiel, Datum.

*Wie würde man denn so eine Einschränkung konstruieren, dass an einem Datum nur genau ein Spiel stattfindet?*

k/A was da gemeint war, hab da rumgedrückt (in bessere NF überführen, Integrity Checks), war aber wohl falsch, haben das dann übersprungen

*Was sind Sichten?*

Für Sichten gibt es keine Modellierung im rel. Modell -> Views in SQL erklärt.

*Was für Einschränkungen gibt es da bei SQL?*

Views updatefähig gdw nur eine Relation, Schlüssel referenziert (Tupel eindeutig identifizierbar) und keine Aggregatfunktionen. SQL Updatefähige Sichten sind Teilmenge von möglichen Sichten.

*Wieso diese Einschränkung?*

??? Umh, Entscheidung ist kostenintensiv?

*Wie schwierig?*

Exponentiell?

*Nein, sogar unentscheidbar! ;-)*

### **Knowledge Representation:**

*Was ist denn so der Unterschied zwischen DB und KB?*

Im wesentlichen Domain Closure (CWA vergessen zu erwähnen), das ist dann auch die analoge Einschränkung im rel. Tupel/Domänenkalkül, dass man nur über Objekte der Domain spricht.

*Und was kann dann sonst schiefgehen?*

Bringt Unendlichkeit rein, bin dann noch auf das Bsp. aus dem KR-Buch eingegangen (endliche Anzahl z.B. von Flug(a,b), Flug(c,d) und Formel alle x die nicht Flug(x,d) -> unendlich. Hab dabei aber ziemlich rumgestottert.

*Diese Art von Schließen ist ja nicht*

monoton

*und welche anderen Arten von nichtmonotonem Schließen kennen sie so?*

Im Wesentlichen zu nennen Default Logics und Autoepistemic Logic. Dann noch Minimal Entailment...

*Dann das tweety -> Pinguin -> Vogel -> Fliegen Inheritance Netzwerk bekommen, wobei noch eine negative Kante von Pinguin nach Fliegen. Übersetzen sie das mal in Default Logic.*

Da hab ich mich ziemlich ungeschickt angestellt, deswegen schreib ich hier mal keine Lösung hin. Im Weiteren Verlauf wollte er aber dann noch wissen, wie man sicherstellt, dass tweety nicht fliegt, das ging in dem man eine zusätzlich nicht Regel konjunktiv in den Mittelteil (beta) einfügt. Hab ich nicht hinkommen, wir hatten da aber in den Übungen mal so eine Aufgabe dazu.

*Wie schwer sind solche Default Logiken?*

Da hab ich dann erwähnt das FO unentscheidbar ist, Default wußte ich nicht. Lakemeyer meinte, dass Default noch schwieriger als FO und damit unentscheidbar.

*Wir hatten ja noch eine andere Logik, Description Logic. Wie fkt. die und ebenfalls so schwierig?*

Kurz DL beschrieben, Interpretation mit Domain und Interpretationsfkt., Extension bestimmt durch logische Operatoren etc. polynomiell lösbar durch structure matching.

*Wie genau?*

$N^2$

*Immer?*

Nein, nur wenn man auf den RESTR Operator verzichtet. Sollte den hinschreiben, hab dann natürlich  $\{x \text{ aus } D \mid \dots\}$  hingeschrieben anstatt  $\{(x,y) \text{ aus } D \mid \dots\}$  (man spricht ja über Rollen). Auf Ls Hinweis dies sei ja der ALL Operator (als Hint gedacht) wies ich natürlich neunmalklug auf die Konjunktion statt den Pfeil hin  $\rightarrow$ , auch der Hinweis: Wie ist denn  $y$  quantifiziert? brachte mich nicht darauf.

### **Künstliche Intelligenz:**

*Was ist Lernen in der KI?*

Induktives Lernen erklärt, von Paaren von  $x, f(x)$  versuchen Funktion zu approximieren, DTs, Neural Networks im Zshg. mit PAC Learning dann auch Dls.

*Wie funktioniert so ein D-Tree?*

Boolsche Fkt, Attribute nach Informationsgehalt betrachten, Inf.gehalt mit Information Theory bestimmen, Attribut als Knoten in Baum, nach diskreten Werten des Attr. dann Bsp einordnen. Sukzessive fortsetzen. Wenn keine Attr. mehr  $\rightarrow$  Noise.

*Was entscheidet man dann?*

Mehrzahl Bsp betrachten und anhand daran entscheiden.

*Warum DLs?*

Motiviert durch PAC Learning: Wie nah bin ich an der tatsächlichen Fkt. dran? PAC: Nach gewisser Anzahl Bsp würde eine sehr falsche Fkt. "auffliegen". Aber allgemein Hypothesenraum  $2^{2^n} \rightarrow$  exponentiell viele Bsp. betrachten. Abhilfe  $k$ -DLs mit  $k$  Literalen,  $n$  Attributen gibt polynomiellen Aufwand.

*Was ist dann der Nachteil bei DLs?*

Schränke Hypothesenraum ein, nicht mehr jede beliebige boolsche Fkt. darstellbar sondern nur noch eben max  $k$ -Lit etc.

*Was sind denn Belief Networks?*

Motiviert durch Joint Probability Table (zu groß, aber jede W'keit berechenbar mit JPT), einfachere Darstellung. BN kurz beschrieben.

*Was ist dann da noch nötig?*

Conditional Independence, hab dann noch etwas ausgeholt (Bayesian Update, Bayes allein wenig nützlich weil bei  $P(X|E_1 \dots E_N)$  zwar von diagnostisch  $\rightarrow$  kausal, aber immer noch Conditional Probability Table von  $2^N$ , daher conditional independence  $\rightarrow$  Einzelw'keiten.

*Und was ist dann eine solche Conditional Independence im Belief Network?*

Wollte dann da loslegen mit d-Separation, aber L wollte einfach nur hören, dass die Pfeile Abhängigkeiten modellieren. Oder so ;-)

*Gegeben sei folgendes Netzwerk (Es war so ein Abhängigkeitsnetzwerk mit A,B,C). Ich sollte dann  $P(A, \text{nicht } C)$  berechnen. Ich meine mit Kanten  $A \rightarrow C$ ,  $B \rightarrow C$  und  $A \rightarrow B$ . Bin aber nicht mehr sicher.*

Dann mit Marginalisierung über die Fälle B und nicht B aufgeschrieben. Müsste dann analog zu oben  $P(\text{nicht } C | A, B) * P(B|A) * P(A) + P(\text{nicht } C | A, \text{nicht } B) * P(\text{nicht } B|A) * P(A)$ . Glaube L meinte dann noch man könne da was rauskürzen, ich habs aber nicht direkt gesehen und dann war auch schon Schluss.

### **Bemerkungen:**

Ich hab mich insbesondere bei den Bsp, die er so brachte relativ ungeschickt angestellt (z.B. bei Default Logic, auch die fkt. Abh. waren nicht so ganz sauber). Kann nur jedem empfehlen, außer der Theorie auch sicherzustellen, dass man solche Beispiele im Schlaf beherrscht, da durch die Nervosität noch ein zusätzlicher Schwierigkeitsfaktor hinzukommt.

Ansonsten ist Lakemeyer ein sehr fairer Prüfer und bleibt auch sehr ruhig und ist nicht "pampig", wenn man mal was nicht kann. (Hab das auch schon anders erlebt ;-)

Einzig unangenehm empfand ich, dass er - während ich etwas erklärte - schon parallel mit Skizzen anfing (z.B. Inheritance Network aufmalt oder Datenbankschema), wo man dann natürlich draufguckt, was dann ablenkt und zudem fast so wirkt als würde er nicht zuhören. Vielleicht bin ich da auch etwas überempfindlich ;-)