

## Information Retrieval

Sascha Kriewel

sascha.kriewel@uni-duisburg.de

## Übungsblatt 11

Bearbeitung bis **11. Juli 2005**  
**keine Abgabe**

---

### Aufgabe 1: Datalog 1

Es sei eine Datenbasis mit Fakten für die folgenden Prädikate gegeben:

`maennlich(X)`: Person X ist maennlich

`weiblich(X)`: Person X ist weiblich

`verheiratet(X,Y)`: Person X und Y sind miteinander verheiratet

`kind(X,Y)`: Person X ist Kind von Person Y

`lebt(X)`: Person X ist noch nicht verstorben

Wie sehen dann Regeln für die folgenden Beziehungen aus: Tochter, Bruder, Nichte, Vetter, Einzelkind, Schwager, männlicher Vorfahr, Waisenkind, Witwe?

### Aufgabe 2: Datalog 2

Es sei eine Datenbasis mit Fakten über verschiedene Arten von multimedialen Dokumenten gegeben. Es soll ein dreistelliges Prädikat `spezieller/3` (`spezieller(DOKART1,DOKART2,MERKMAL)`) genutzt werden, das ausdrücken soll, das ein *Dokumenttyp*<sub>1</sub> spezieller als ein anderer *Dokumenttyp*<sub>2</sub> ist, weil er ein zusätzliches Merkmal besitzt.

- Überlege Dir einige beispielhafte Fakten.
- Wie könnten Regeln aussehen, die zu einem Begriff liefern, ob es sich dabei um ein Merkmal für Dokumentarten im obigen Sinne handelt?
- Wie könnten Regeln aussehen, die zu zwei Dokumentarten liefern, ob sie Geschwister sind (also eine **direkte** gemeinsame Oberart haben)?
- Wie könnten Regeln aussehen, die zu einer Dokumentart ihre Merkmale liefern (rekursiv)?
- Überlege, wie sich das Beispiel erweitern liesse, wenn man statt Datalog probabilistisches Datalog (pDatalog) einsetzt.

### Aufgabe 3: Markup-Sprachen

Betrachte das Vorlesungsverzeichnis der Universität Duisburg-Essen. Überlege Dir, wie Du Auszeichnungen (Markup) für dieses Dokument bezüglich der logischen (nicht der Layout-) Struktur in XML oder SGML festlegen würdest. Welche Elemente gibt es, wie sind diese strukturiert?

Wäge bei Modellierungsvarianten ab, welche für spätere Anwendungszwecke (also z.B. der Suche nach bestimmter Information im Vorlesungsverzeichnis) wie gut geeignet ist:

- Was sollte eher als Daten (Inhalt) und was als Metadaten (Elemente) modelliert werden?
- Wo sind hierarchische Strukturen sinnvoll?
- Wie granular muß die Modellierung sein?
- Was ist mit Nullwerten?

Versuche zumindest exemplarisch einige Elemente formal festzulegen, ohne eine komplette DTD zu schreiben. Notiere einen Auszug des Vorlesungsverzeichnis mit dem gewählten Markup.