Intelligent Information Systems

(MA-INF 3203)

Prof. Dr. Rainer Manthey

WS 2016/17
Vita Rainer Manthey

1953 Wilhelmshaven

1973 Kiel

University of Kiel
Informatics/Mathematics
Student (Diploma 1979)
Research assistant (PhD 1984)

1984 München

European Computer-Industry Research Centre (ECRC)
Researcher/Teamleader

1992 Bonn

University of Bonn
Professor
Modules Offered by the IDB Group

IDB (Intelligent Databases) Group:
Prof. Dr. Rainer Manthey
PD Dr. Andreas Behrend
Sahar Vahdati, MSc*

(* partially with Prof. Auer’s group)

WS
Intelligent Information Systems
(MA-INF 3203)
NEW: 4 hrs lecture/week

SS
Temporal Information Systems
(MA-INF 3302)

WS+SS
Seminar Selected Topics in Intelligent IS
(MA-INF 3210)
NEW: 2 hrs lecture/week (?)

Lab Intelligent Information Systems
(MA-INF 3313)

Scientific Writing
Prof. Dr. Rainer Manthey:

- **28.2.2019**: Day of Retirement
- => 5 semesters of teaching left (only in MSc program):
  WS 16/17, SS 17, WS 17/18, SS 18, WS 18/19
- => Supervision of master theses ends 30.9.2018

PD Dr. Andreas Behrend:

- PD/Habilitation: Full qualifications for any kind of academic teaching (incl. thesis supervision), independent teaching schedule
- Position at Uni Bonn ends at **30.6.2018** (at latest)
- At most 3 semesters of teaching left: **WS 16/17, SS 17, WS 17/18**
- Last two years: Guest professor at universities in Dresden, Marburg, Halle
- This semester: teaching in Bachelor and MSc LSI
- Next semester(s): teaching still open
### “Geography” of Our Institute

<table>
<thead>
<tr>
<th>Department I:</th>
<th>Department II:</th>
<th>Department III:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prof. Ro. Klein</td>
<td>Prof. Re. Klein</td>
<td>Prof. Wrobel</td>
</tr>
<tr>
<td>Prof. Kratsch</td>
<td>Prof. Weber</td>
<td><strong>Prof. Manthey</strong></td>
</tr>
<tr>
<td></td>
<td>Prof. Hullin</td>
<td>Prof. Gall</td>
</tr>
<tr>
<td></td>
<td>Jun.-Prof. Schultz</td>
<td>Prof. Auer</td>
</tr>
<tr>
<td></td>
<td>Jun.-Prof. Yao</td>
<td>Prof. Lehmann</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Department IV:</th>
<th>Department V:</th>
<th>Department VI:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prof. Martini</td>
<td>Prof. Röglin</td>
<td>Prof. Behnke</td>
</tr>
<tr>
<td>Prof. Meier</td>
<td>Prof. Blum</td>
<td>Prof. Anlauf</td>
</tr>
<tr>
<td>Prof. Smith</td>
<td></td>
<td>Prof. Bennewitz</td>
</tr>
<tr>
<td>apl. Prof. Kurth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun.-Prof. Reinhardt</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(**Department III:** in Römerstraße; all **other departments:** in Friedrich-Ebert-Allee)
Groups in Department III

- Main area of Research (and Teaching) in Department III: Intelligent Systems

- Until 2013: Information Systems and Software Engineering

- Since then: New department head Prof. Wrobel
  - Who is also Director of the Fraunhofer Institute IAIS and Director of B-IT
    (IAIS: Intelligent Analysis and Information Systems; B-IT: Bonn-Aachen International Center for Information Technology)

  - Recently: Two big new research groups founded (associated with IAIS as well)
    - EIS – Enterprise Information Systems (Prof. Auer) ~ 40 members
    - SDA – Smart Data Analytics (Prof. Lehmann) ~ 20 members

- Since 2013: Computer Vision group (Prof. Gall) 10 members
- Since 1992: Intelligent DB group (Prof. Manthey) 2 members

- Teaching (main lectures in MSc):
  - Wrobel: Machine Learning, Data Science & Big Data
  - Auer: Semantic DataWeb Technologies, Enterprise Information Systems
  - Lehmann: Knowledge Graph Analysis
  - Gall: Computer Vision I + II
  - Manthey: Intelligent IS, Temporal IS
Intelligent Information Systems

WS 2016/17

Organisation
• Every week: 4 hours lecture + 2 hours exercises (9 credits)

• **Lectures:** Monday afternoon (14:15 – 15:45) + **Wednesday** morning (10:30 – 12:00)

• **Exercises:** Wednesday afternoon (12:45 – 14:15)

• Lecture and exercises in the **same room** (A 207).

• On **Wednesdays** special **timing** conventions:
  • Lecture starts 15 minutes **later** than usual: 10:30 rather than 10:15 a.m..
  • **45 minutes break** between lecture and exercises (rather than just 30 minutes)
  • Exercises start 30 minutes **later** than usual: 12:45 rather than 12:15.
  • Exercises end 30 minutes **later** than usual: 14:15 rather than 13:45.
  • **Next** lecture (Prof. Auer) starts 14:30 rather than 14:15!
<table>
<thead>
<tr>
<th></th>
<th>Monday</th>
<th>Wednesday</th>
</tr>
</thead>
<tbody>
<tr>
<td>October</td>
<td>17</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>26</td>
</tr>
<tr>
<td>November</td>
<td>31</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>30</td>
</tr>
<tr>
<td>December</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>21</td>
</tr>
<tr>
<td>January</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>27</td>
</tr>
<tr>
<td>February</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>10</td>
</tr>
</tbody>
</table>

- 18 lectures
- 8 exercises
- Begin of exercises
- 2016
- 2017
- Dies academicus
- Xmas break (2 weeks)
- 10 lectures
- 5 exercises
Exercises and Exams: „Rules of the Game“

- **Exercises:**
  - In the **same** room every Wednesday, following the lecture after 45 minutes break, for the entire auditorium, **no** small groups.
  - Exercises held by Prof. Manthey and/or Sahar Vahdati.
  - **Goals:**
    - **To make you fit for the exam!**
    - **Attention!** There are **too few** exercise weeks (for a long lecture)!
    - To provide some „hands on“ experience with theoretically introduced concepts.
  - **Participation** will **not** be checked, but is strongly **recommended**!!
  - **No** prerequisites for getting **admission** to exams!
  - **No** „homework“ to be delivered, but motivation/encouragement for individual activity provided in exercises.
  - **No** individual feedback possible.

- **Exams:**
  - **Registration** for exams: December 1 till 22
    (online via BASIS)
  - **Written exams** for both exam dates
    (120 minutes duration)
  - **Exam dates to be determined:**
    Most likely end of February + end of March
Master Programme in Computer Science
(MA-INF 3203)
Lecture "Intelligent Information Systems" (WS 2016/17)
Prof. Dr. Rainer Manthey

News

- The 1st IIS lecture (this year) will take place on Wednesday, October 19, 2016. On Wednesdays, the lecture will always begin at 10:30 (a.m.) and last till 12:00. After a 45 minutes break, the exercises will begin at 12:45 (and last till 14:15). Both, lecture and exercises, will take place in lecture hall A 207 (Romerstrasse, 2nd floor). Monday lectures take place in the same hall. They will start "regularly" at 14:15 (and last till 15:45). There is no lecture yet on October 17, and there are no exercises yet on October 19.

- From this semester (WS 2016/17) onwards, IIS will be offered in a new format, i.e., with 4 rather than 2 lecture hours per week. Additional topics will be covered by IIS (as compared to previous years), and more time will be available for some of the old topics (due to experiences from the past). Consequently, 9 rather than 6 credits will be granted for all participants passing the exam. In the future, IIS exams will last two full hours (rather than 90 minutes before).

- There will be no possibility for students who passed the old (shorter) lecture to attend the module again. IIS will be offered only 3 more times from now on: WS 2016/17, WS 2017/18, and finally WS 2018/19. In early 2019, Prof. Manthey will retire.

- Have a look at follow-up page "Background" (link in the column to the left) even before the lecture starts... and use the material provided there. A link to the DAS system (for practically working with concepts and languages relevant in IIS) can be found on this page, too.

- Please direct questions about the module by mail to Prof. Manthey (manthey@cs.uni-bonn.de).

last modified: 30.9.2016

Slides for download!
There is no textbook which could be recommended for this lecture . . .

. . . just the slides serve as a substitute instead (representing a compromise between a good background presentation and too much text)
A Word of Warning

But:

Only a small fraction of the attendees will have a chance to get a place in seminars or to do a master thesis in this area!

IIS 2015: Over 100 participants in the exam !!
Background

Intelligent Information Systems

WS 2016/17
This is the most commonly agreed view on the concept of an IS in informatics – provided people agree on the meaning of DBS!!
Users and application programmes

DBMS: Data Base Management System
(Many powerful application-independent services: schema mgt, query optimization, storage mgt, transaction mgt, etc.)
This lecture will be more accurately concerned with

**Intelligent Database Systems**

rather than with

**Intelligent Information Systems**

The naming of the module is more a matter of convention rather than precision!
"Real" DBMS support a separate kind of DB-specific "programming language" for accessing and manipulating data in the DB: query language.

In contrast to the external imperative programming languages, a query language is usually a declarative language, the performance of which is optimised by the DBMS.

"Programs" of the query language may be stored in the data dictionary within the DB.
Relational Data Model and SQL

- The most widely used data model nowadays is the relational model (introduced around 1970). Relations are the mathematical basis for data represented in tables (rows/columns).

- All relational DBMS support a predominant declarative query language based on logical and algebraic operators: SQL (Structured Query Language)
A good background in relational databases and in SQL is expected from everybody attending this lecture!! SQL will frequently be used during the semester, even though we are going to learn a different relational language!

Material for self-study (in case your background is weak, dated, or missing):

- Extra slides via IIS homepage
- Cheap and easy tutorials from the Schaum's series
Motivation

Intelligent Information Systems

WS 2016/17
Intelligent *Database* or Intelligent *Database System*?

Where is „intelligence“ located?
In the **DB** or in the **DBMS**?
Or even outside the **DBS**?
Certainly required: „Intelligent“ behaviour of the DBS, i.e., generic (application-independent) services inside the DBMS, able to „simulate intelligence“
Also certainly required: „Knowledge“ about the resp. application domain in the DD (Data Dictionary)

„Knowledge“: Rules from the application domain as a basis for drawing intelligent conclusions from stored data
IDBS: „Traditional“ Approach with External System Components

Inference System, Agent System, Expert System

Knowledge Base, Rule Base

DBS „loose coupling“

Preferred by many: Move „Intelligence“ and „Knowledge“ out of the DBS
Approach favoured by our research group (and thus in this lecture):

- Try to reach as much „intelligence“ as possible using existing DB technology!
- Identify weaknesses of this technology and think about reasonable extensions, without leaving the DB context!
This approach – which is a special one – explains the drawing on the title slide of this lecture.

Therefore:
Theory and Practice of the established research area of „Deductive Databases“ will be at the core of this lecture.

The essence of this area of research can be described as follows:

How to analyse data using stored queries (in SQL: views) that serve as declarative analytical programs?
Datalog and SQL

- Research in deductive databases has a nearly 40-years history (as old as SQL), but has been using a **different** declarative language (not SQL!) most of the time, strongly influenced by the logic programming language PROLOG: **Datalog**

- Nearly all publications in this area have been using **Datalog** – that‘s why we will use Datalog during this lecture, too (and you will have to learn it!).

- Many results of DDB research have been transferred to the **SQL** world recently! That‘s why SQL will also be appearing throughout the lecture in various places.

**SQL:**
- used in **industry** and commerce
- supported by many DBMS **products**
- **standardized**
- **user-friendly** („controlled English“)
- **rich** set of syntactic features

**Datalog:**
- used in **academia** only
- just few academic **protoypes**
- **no** standards
- **mathematical** style
- **minimalistic** syntax
Datalog vs. SQL: Comparison in a Nutshell

**Datalog rules**

\[
\begin{align*}
  s(X) & \leftarrow p(X,Y). \\
  s(X) & \leftarrow r(Y,X). \\
  t(X,Y,Z) & \leftarrow p(X,Y), r(Y,Z). \\
  w(X) & \leftarrow s(X), \text{not} \ q(X).
\end{align*}
\]

**SQL views**

```sql
CREATE VIEW s AS
  (SELECT a FROM p)
  UNION
  (SELECT b FROM r);

CREATE VIEW t AS
  SELECT a, b, c FROM p, r
  WHERE p.b = r.a,

CREATE VIEW w AS
  (TABLE s)
  MINUS
  (TABLE q);
```
Datalog Basics on a Single Slide

Constants

Facts

p(1,a).
p(2,b).
p(3,c).
q(2).
q(5).
r(a,1).
r(a,2).
r(b,3).

Rules

s(X) ← p(X,Y).
s(X) ← r(Y,X).
t(X,Y,Z) ← p(X,Y), r(Y,Z).
w(X) ← s(X), not q(X).

Variables

Relation Names

p, q, r: Base relations
s, t, w: Derived relations

Conjunction

Negation
Structure of the Course

This is how the lecture will be structured – the number of lectures might be slightly varying in „real life“

<table>
<thead>
<tr>
<th>1. Organisation and Motivation</th>
<th>1 lecture</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Deduction in Datalog and SQL</td>
<td>7 lectures</td>
</tr>
<tr>
<td>3. Semantics of Deductive Databases</td>
<td>6 lectures</td>
</tr>
<tr>
<td>4. Efficient Query Evaluation in DDBs</td>
<td>6 lectures</td>
</tr>
<tr>
<td>5. Efficient Update Propagation in DDBs</td>
<td>7 lectures</td>
</tr>
<tr>
<td>6. Perspectives</td>
<td>1 lecture</td>
</tr>
</tbody>
</table>

Timetable:
- Chapter 1: today
- Chapter 2: Oct/Nov
- Chapter 3: Nov/Dec
- Chapter 4: Dec/Jan
- Chapter 5: Jan/Febr
- Chapter 6: 8.2.2017

Chapter 5 is new this year!
In the next lecture, we will discuss a typical example of a relational database providing a lot of opportunities for static analysis of data using stored queries:

**Genealogical Databases**

Genealogy is the discipline of exploring family relationships between persons and their ancestors/descendants.
How to „put a family tree into a (relational) database“?