Grundlagen des relationalen Modells

- Das relationale Modell
- Verfeinerung des relationalen Schemas
- Relationale Algebra
- Relationenkalkül

Kapitel 3

Grundlagen des relationalen Modells

Seien D_1, D_2, …, D_n Domänen (Wertebereiche, Mengen)

- Eine *Relation* ist eine Teilmenge $R \subseteq D_1 \times \ldots \times D_n$

  Bsp.: Telefonbuch $\subseteq$ *string x string x integer*

- Ein *Tupel* ist jedes Element $t \in R$ von R

  Bsp.: t = („Mickey Mouse", „Main Street", 4711)

- *Schema:* legt die Struktur der gespeicherten Daten fest

  Bsp.:

Telefonbuch: {[*Name: string, Adresse: string, <u>Telefon#:integer</u>*]}

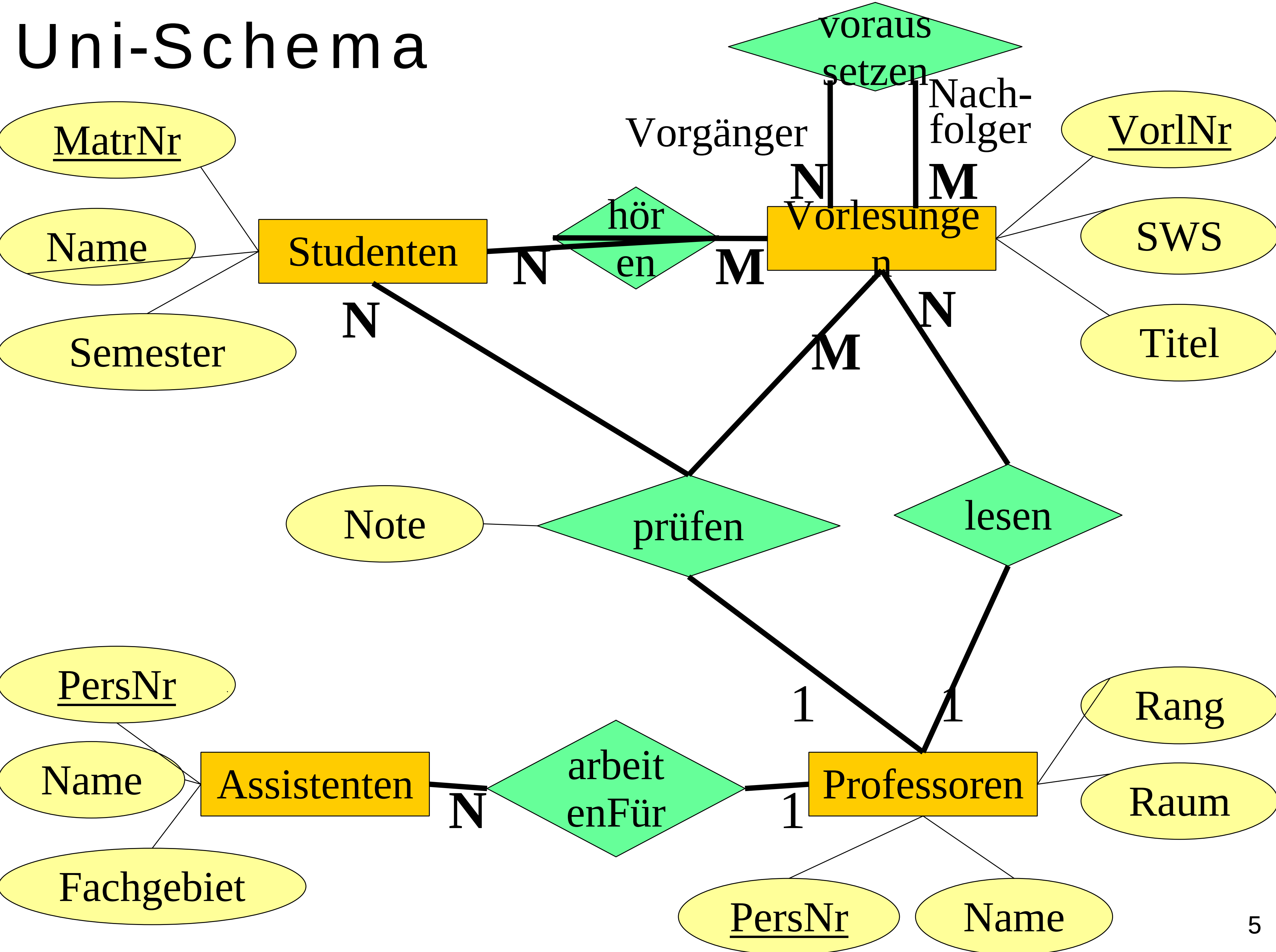
Telefonbuch		
Name	**Straße**	**<u>Telefon#</u>**
Mickey Mouse	Main Street	4711
Mini Mouse	Broadway	94725
Donald Duck	Broadway	95672
...	...	...

- **Ausprägung:** der aktuelle Zustand der Datenbasis
- **Schlüssel:** minimale Menge von Attributen, deren Werte ein Tupel eindeutig identifizieren
- **Primärschlüssel:** wird unterstrichen
  - Einer der Schlüsselkandidaten wird als Primärschlüssel ausgewählt
  - Hat eine besondere Bedeutung bei der Referenzierung von Tupeln

Telefonbuch		
Name	**Straße**	**<u>Telefon#</u>**
Mickey Mouse	Main Street	4711
Mini Mouse	Broadway	94725
Donald Duck	Broadway	95672
...	...	...

- Die Festlegung eines (Primär-)Schlüssels ist eine Designentscheidung.
- Bei einer gegebenen Datenbank wird dann bei einer Konsistenzprüfung überprüft, ob sie dieser Einschränkung gehorcht.

Uni-Schema



voraus
setzen
Vorgänger
Nach-
folger
N
M
MatrNr
Name
Semester
Studenten
hör
en
N
M
Vorlesungen
VorlNr
SWS
Titel
N
M
N
Note
prüfen
lesen
1
1
PersNr
Name
Fachgebiet
Assistenten
arbeit
enFür
N
1
Professoren
Rang
Raum
PersNr
Name

Relationale Darstellung von Entitytypen

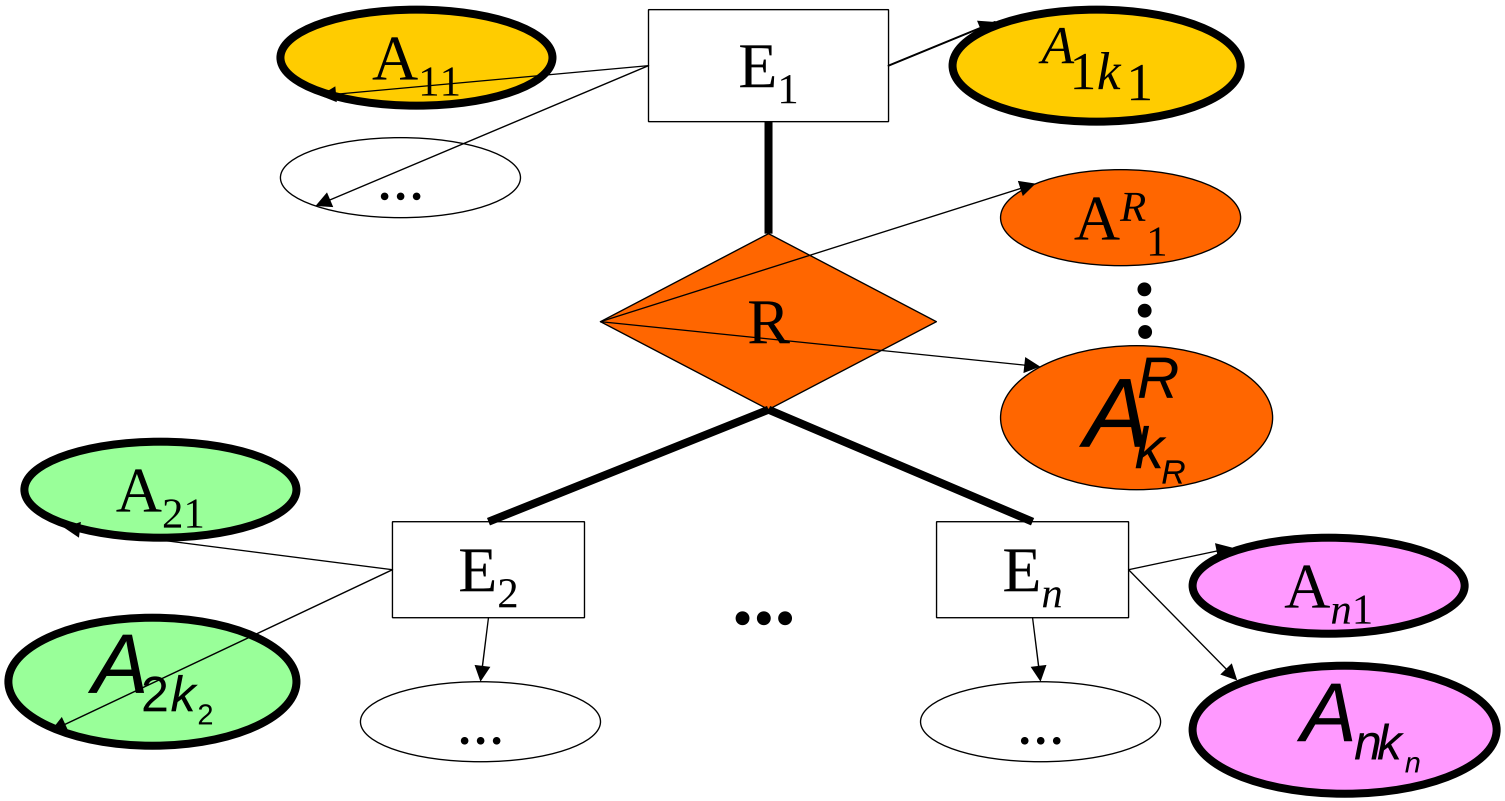
Studenten: {[<u>MatrNr:integer</u>, *Name*: *string, Semester*: *integer*]}

Vorlesungen: {[<u>VorlNr:integer</u>, *Titel*: *string, SWS*: *integer*]}

Professoren: {[<u>PersNr:integer</u>, *Name*: *string, Rang*: *string,*
Raum: *integer*]}

Assistenten: {[<u>PersNr:integer</u>, *Name*: *string, Fachgebiet*: *string*]}

Relationale Darstellung von Beziehungen



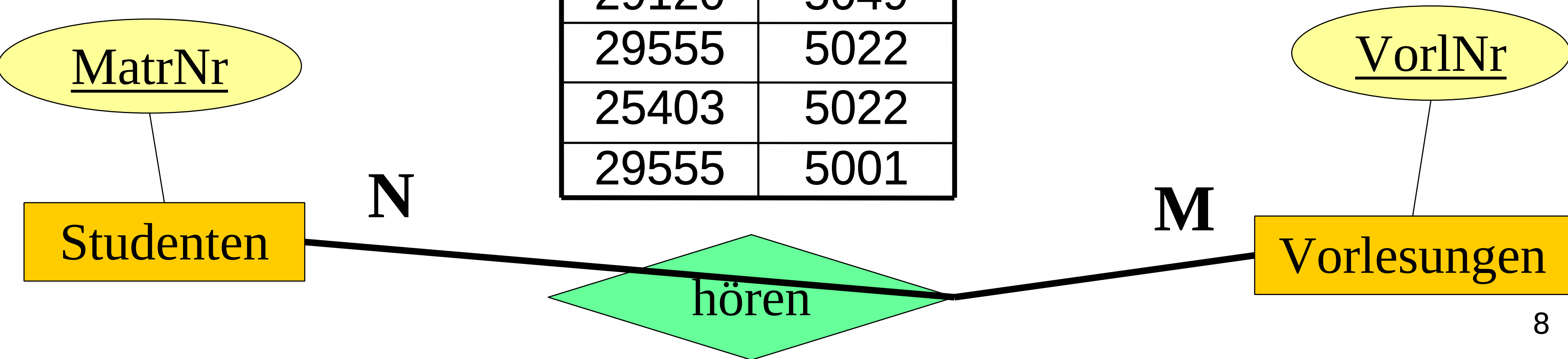
$$R\!:\!\{[\underbrace{A_{11},\ldots,A_{1k_1}}_{\text{Schlüssel von } E_1},\underbrace{A_{21},\ldots,A_{2k_2}}_{\text{Schlüssel von } E_2},\ldots,\underbrace{A_{n1},\ldots,A_{nk_n}}_{\text{Schlüssel von } E_n},\underbrace{A_1^R,\ldots,A_{k_R}^R}_{\text{Attribute von } R}]\}$$

Ausprägung der Beziehung *hören*

Studenten	
MatrNr	...
26120	...
27550	...
...	...

hören	
MatrNr	VorlNr
26120	5001
27550	5001
27550	4052
28106	5041
28106	5052
28106	5216
28106	5259
29120	5001
29120	5041
29120	5049
29555	5022
25403	5022
29555	5001

Vorlesungen	
VorlNr	...
5001	...
4052	...
...	...



Beziehungen unseres Beispiel-Schemas

hören : {[MatrNr: integer, VorlNr: integer]}

lesen : {[PersNr: integer, VorlNr: integer]}

arbeitenFür : {[AssistentenPersNr: integer, *ProfPersNr: integer*]}

voraussetzen : {[Vorgänger: integer, Nachfolger: integer]}

prüfen : {[MatrNr: integer, VorlNr: integer, PersNr: integer,
Note: decimal]}

Schlüssel der Relationen

hören : {[MatrNr: integer, VorlNr: integer]}

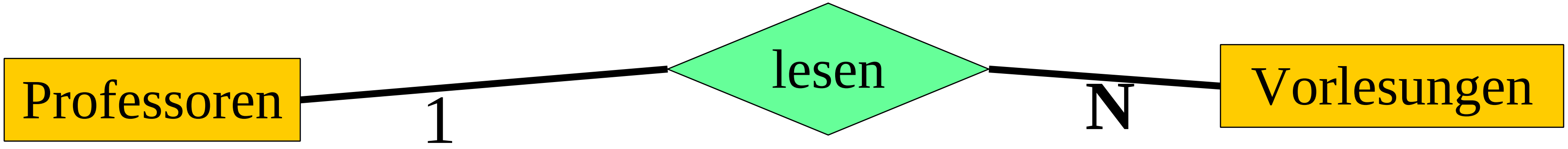
lesen : {[PersNr: integer, VorlNr: integer]}

arbeitenFür : {[AssistentenPersNr: integer, *ProfPersNr: integer*]}

voraussetzen : {[Vorgänger: integer, Nachfolger: integer]}

prüfen : {[MatrNr: integer, VorlNr: integer, PersNr: integer,
Note: decimal]}

Verfeinerung des relationalen Schemas



1:N-Beziehung

- Initial-Entwurf

  Vorlesungen : {[<u>*VorlNr*</u>*, Titel, SWS*]}

  Professoren : {[<u>*PersNr*</u>*, Name, Rang, Raum*]}

  lesen: {[<u>*VorlNr*</u>*, PersNr*]}

Verfeinerung des relationalen Schemas

1:N-Beziehung

- Initial-Entwurf

  Vorlesungen : {[*<u>VorlNr</u>, Titel, SWS*]}
  Professoren : {[*<u>PersNr</u>, Name, Rang, Raum*]}
  lesen: {[*<u>VorlNr</u>, PersNr*]}

- Verfeinerung durch Zusammenfassung

  Vorlesungen : {[*<u>VorlNr</u>, Titel, SWS,* ***gelesenVon***]}
  Professoren : {[*<u>PersNr</u>, Name, Rang, Raum*]}

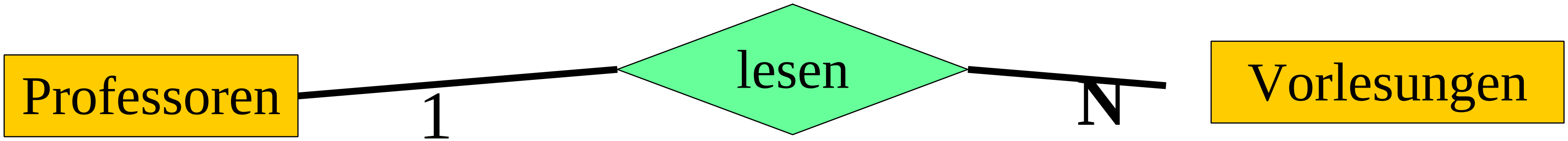
Regel

Relationen mit gleichem Schlüssel kann man zusammenfassen
aber nur diese und keine anderen!

Ausprägung von *Professoren* und *Vorlesung*

Professoren			
PersNr	Name	Rang	Raum
2125	Sokrates	C4	226
2126	Russel	C4	232
2127	Kopernikus	C3	310
2133	Popper	C3	52
2134	Augustinus	C3	309
2136	Curie	C4	36
2137	Kant	C4	7

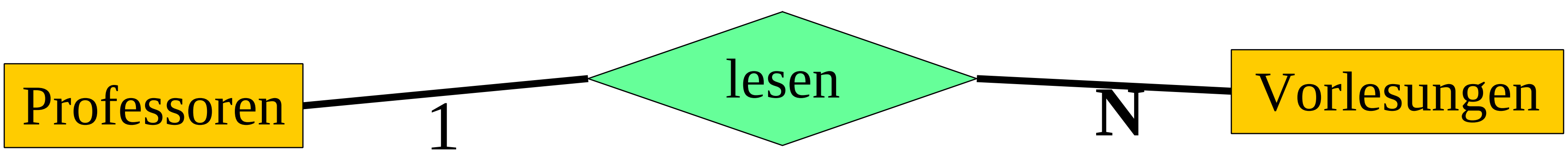
Vorlesungen			
VorlNr	Titel	SWS	Gelesen Von
5001	Grundzüge	4	2137
5041	Ethik	4	2125
5043	Erkenntnistheorie	3	2126
5049	Mäeutik	2	2125
4052	Logik	4	2125
5052	Wissenschaftstheorie	3	2126
5216	Bioethik	2	2126
5259	Der Wiener Kreis	2	2133
5022	Glaube und Wissen	2	2134
4630	Die 3 Kritiken	4	2137



Vorsicht: So geht es NICHT

Professoren				
PersNr	Name	Rang	Raum	liest
2125	Sokrates	C4	226	5041
2125	Sokrates	C4	226	5049
2125	Sokrates	C4	226	4052
...	...	...	...	...
2134	Augustinus	C3	309	5022
2136	Curie	C4	36	??

Vorlesungen		
VorlNr	Titel	SWS
5001	Grundzüge	4
5041	Ethik	4
5043	Erkenntnistheorie	3
5049	Mäeutik	2
4052	Logik	4
5052	Wissenschaftstheorie	3
5216	Bioethik	2
5259	Der Wiener Kreis	2
5022	Glaube und Wissen	2
4630	Die 3 Kritiken	4



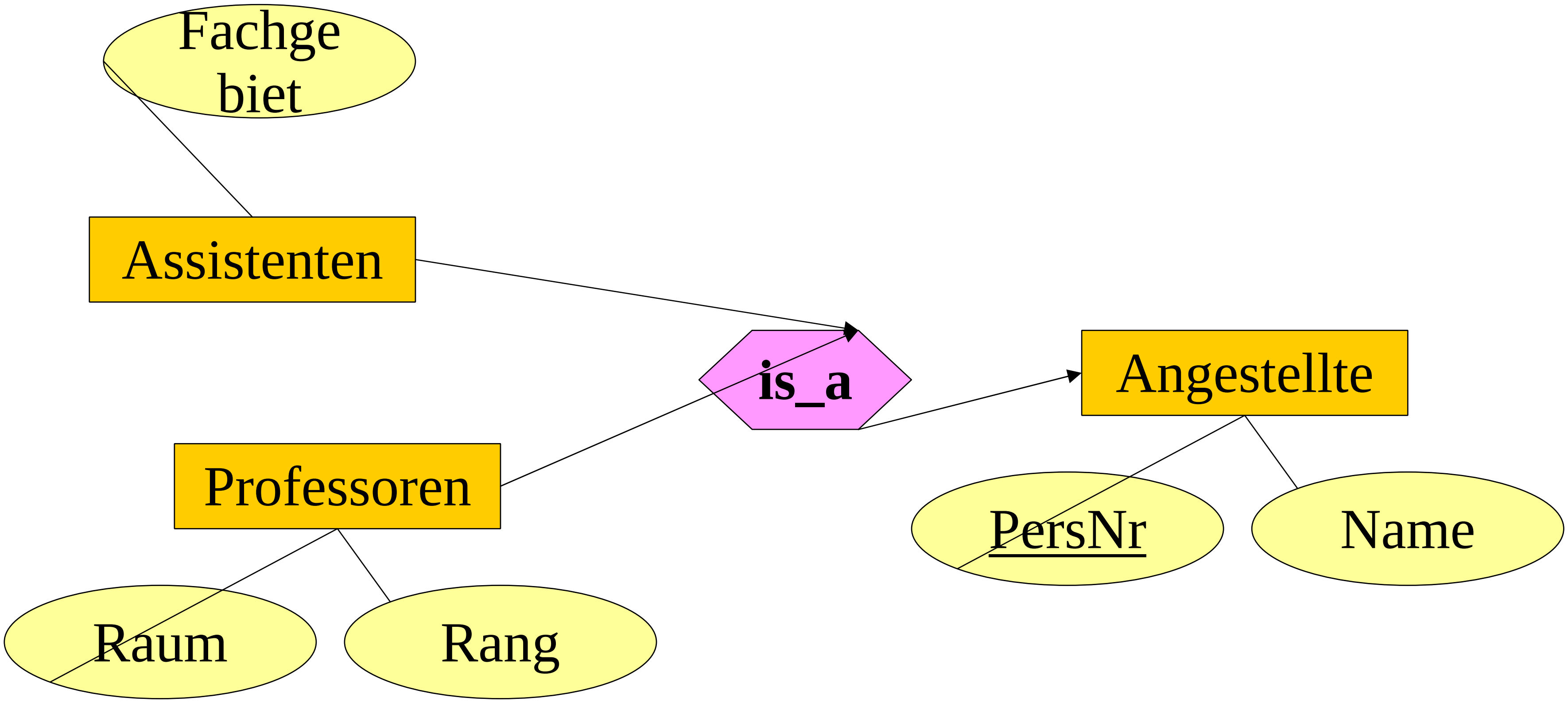
Vorsicht: So geht es NICHT: Folgen→Anomalien

Professoren				
PersNr	Name	Rang	Raum	liest
2125	Sokrates	C4	226	5041
2125	Sokrates	C4	226	5049
2125	Sokrates	C4	226	4052
...	...	...	...	...
2134	Augustinus	C3	309	5022
2136	Curie	C4	36	??

Vorlesungen		
VorlNr	Titel	SWS
5001	Grundzüge	4
5041	Ethik	4
5043	Erkenntnistheorie	3
5049	Mäeutik	2
4052	Logik	4
5052	Wissenschaftstheorie	3
5216	Bioethik	2
5259	Der Wiener Kreis	2
5022	Glaube und Wissen	2
4630	Die 3 Kritiken	4

- Update-Anomalie: Was passiert wenn Sokrates umzieht?
- Lösch-Anomalie: Was passiert wenn „Glaube und Wissen" wegfällt?
- Einfügeanomalie: Curie ist neu und liest noch keine Vorlesungen? (→ Funktionale Abhängigkeiten)

Relationale Modellierung der Generalisierung

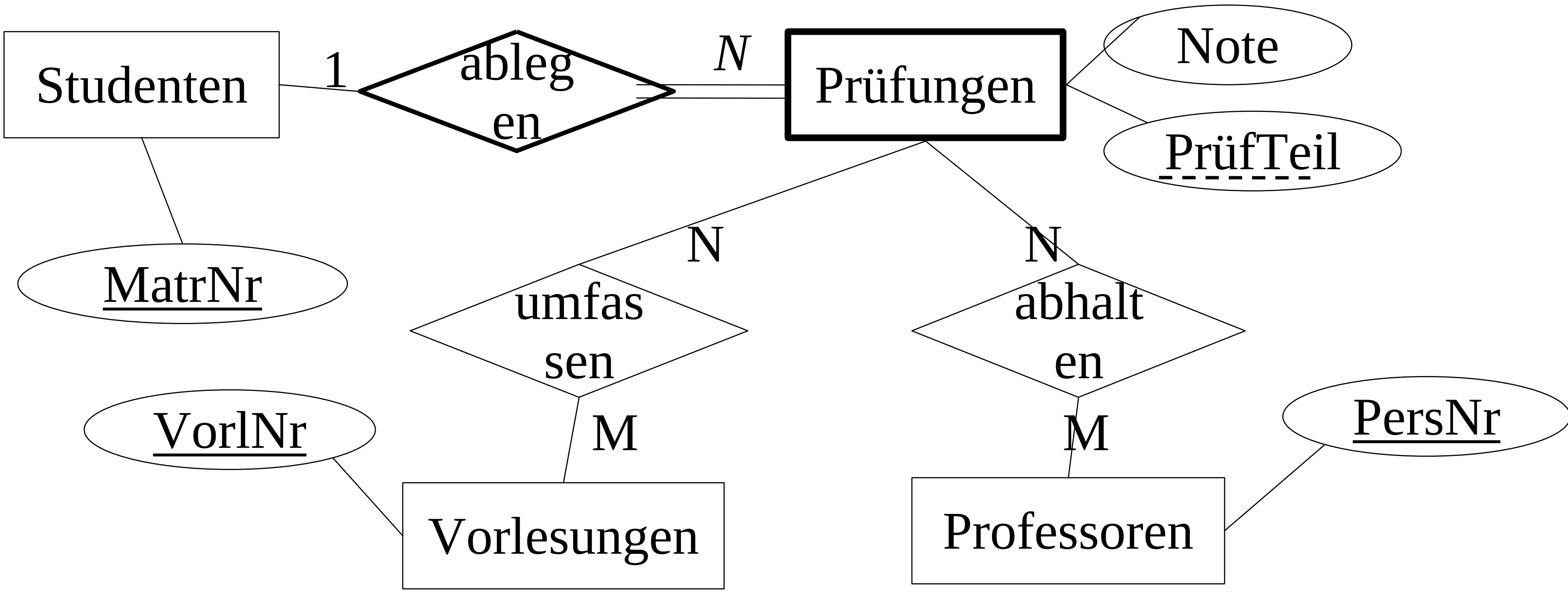


Angestellte: {[*<u>PersNr</u>*, *Name*]}

Professoren: {[*<u>PersNr</u>*, *Rang*, *Raum*]}

Assistenten: {[*<u>PersNr</u>*, *Fachgebiet*]}

Relationale Modellierung schwacher Entitytypen



Prüfungen: <u>{[MatrNr: integer, PrüfTeil: string</u>, Note: integer]}

umfassen: <u>{[MatrNr: integer, PrüfTeil: string, VorlNr: integer</u>]}

abhalten: <u>{[MatrNr: integer, PrüfTeil: string, PersNr: integer</u>]}

Man beachte, dass in diesem Fall der (global eindeutige) Schlüssel der Relation *Prüfung* nämlich *MatrNr* **und** *PrüfTeil* als Fremdschlüssel in die Relationen *umfassen* und *abhalten* übernommen werden muss.

Die relationale Uni-DB

Professoren

PersNr	Name	Rang	Raum
2125	Sokrates	C4	226
2126	Russel	C4	232
2127	Kopernikus	C3	310
2133	Popper	C3	52
2134	Augustinus	C3	309
2136	Curie	C4	36
2137	Kant	C4	7

Studenten

MatrNr	Name	Semester
24002	Xenokrates	18
25403	Jonas	12
26120	Fichte	10
26830	Aristoxenos	8
27550	Schopenhauer	6
28106	Carnap	3
29120	Theophrastos	2
29555	Feuerbach	2

Vorlesungen

VorlNr	Titel	SWS	gelesen von
5001	Grundzüge	4	2137
5041	Ethik	4	2125
5043	Erkenntnistheorie	3	2126
5049	Mäeutik	2	2125
4052	Logik	4	2125
5052	Wissenschaftstheorie	3	2126
5216	Bioethik	2	2126
5259	Der Wiener Kreis	2	2133
5022	Glaube und Wissen	2	2134
4630	Die 3 Kritiken	4	2137

voraussetzen

Vorgänger	Nachfolger
5001	5041
5001	5043
5001	5049
5041	5216
5043	5052
5041	5052
5052	5259

hören

MatrNr	VorlNr
26120	5001
27550	5001
27550	4052
28106	5041
28106	5052
28106	5216
28106	5259
29120	5001
29120	5041
29120	5049
29555	5022
25403	5022

Assistenten

PerslNr	Name	Fachgebiet	Boss
3002	Platon	Ideenlehre	2125
3003	Aristoteles	Syllogistik	2125
3004	Wittgenstein	Sprachtheorie	2126
3005	Rhetikus	Planetenbewegung	2127
3006	Newton	Keplersche Gesetze	2127
3007	Spinoza	Gott und Natur	2126

prüfen

MatrNr	VorlNr	PersNr	Note
28106	5001	2126	1
25403	5041	2125	2
27550	4630	2137	2

Die relationale Algebra

- σ Selektion
- π Projektion
- x Kreuzprodukt
- $\bowtie$ Join (Verbund)
- ρ Umbenennung
- $-$ Mengendifferenz
- $\div$ Division
- $\cup$ Vereinigung
- $\cap$ Mengendurchschnitt
- $\rtimes$ Semi-Join (linker)
- $\ltimes$ Semi-Join (rechter)
- ⟕ linker äußerer Join
- ⟖ rechter äußerer Join

Die relationalen Algebra-Operatoren

$\sigma_{\text{Semester} > 10}$ (Studenten)

Selektion

$\sigma_{\text{Semester} > 10}$ (Studenten)		
MatrNr	Name	Semester
24002	Xenokrates	18
25403	Jonas	12

In der Selektion $\sigma_F(R)$ ist das Selektionsprädikat F eine Formel, die aufgebaut ist aus

- Attributnamen von R und Konstanten
- $=, <, >, \leq, \geq, \neq$
- den logischen Operatoren $\wedge, \vee, \neg$

Das Ergebnis von $\sigma_F(R)$ besteht aus allen Tupeln $t \in R$, die F erfüllen, wenn jedes Auftreten eines Attributes A durch den Wert t.A ersetzt wird.

Die relationalen Algebra-Operatoren

Projektion

Π_{Rang}(Professoren)

Π_{Rang}(Professoren)
Rang
C4
C3

• Die Projektion wählt (eine oder mehrere) Spalten der Relation aus.

• Duplikate im Ergebnis werden nur einmal gelistet (aufgrund der Mengensemantik des Relationenkalküls)

Die relationalen Algebra-Operatoren

Vereinigung

Studenten		
MatrNr	Name	Semester
24002	Xenokrates	18
25403	Jonas	12
26120	Fichte	10

$\cup$

Studenten		
MatrNr	Name	Semester
26830	Aristoxenos	8
27550	Schopenhauer	6
28106	Carnap	3
29120	Theophrastos	2
29555	Feuerbach	2

=

Studenten		
MatrNr	Name	Semester
24002	Xenokrates	18
25403	Jonas	12
26120	Fichte	10
26830	Aristoxenos	8
27550	Schopenhauer	6
28106	Carnap	3
29120	Theophrastos	2
29555	Feuerbach	2

- Relationen mit gleichem Schema können vereinigt werden.

Die relationalen Algebra-Operatoren

Differenz

Studenten		
MatrNr	Name	Semester
24002	Xenokrates	18
25403	Jonas	12
26120	Fichte	10
26830	Aristoxenos	8
27550	Schopenhauer	6
28106	Carnap	3
29120	Theophrastos	2
29555	Feuerbach	2

\-

Studenten		
MatrNr	Name	Semester
26830	Aristoxenos	8
27550	Schopenhauer	6
28106	Carnap	3
29120	Theophrastos	2
29555	Feuerbach	2
12345	Waalkes	22

=

Studenten		
MatrNr	Name	Semester
24002	Xenokrates	18
25403	Jonas	12
26120	Fichte	10

- Von Relationen mit gleichem Schema kann die Mengendifferenz gebildet werden.

Die relationalen Algebra-Operatoren

Kartesisches Produkt

Professoren x hören

Professoren				hören	
PersNr	Name	Rang	Raum	MatrNr	VorlNr
2125	Sokrates	C4	226	26120	5001
...	...	...	...	...	...
2125	Sokrates	C4	226	29555	5001
...	...	...	...	...	...
2137	Kant	C4	7	29555	5001

- Das Schema enthält alle Attribute beider Relationen.
- Die Relation enthält alle $n \times m$ möglichen Kombinationen der jeweiligen Tupel der beiden Relationen.

Die relationalen Algebra-Operatoren

Kartesisches Produkt

Professoren x hören

Professoren				hören	
PersNr	Name	Rang	Raum	MatrNr	VorlNr
2125	Sokrates	C4	226	26120	5001
...	...	...	...	...	...
2125	Sokrates	C4	226	29555	5001
...	...	...	...	...	...
2137	Kant	C4	7	29555	5001

- Problem: riesige Zwischenergebnisse
- Beispiel: (Professoren x hören)
- "bessere" Operation: Join (siehe unten)

Die relationalen Algebra-Operatoren

Umbenennung

- Umbenennung von Relationen
- Beispiel: Ermittlung indirekter Vorgänger 2. Stufe der Vorlesung 5216

$$\Pi_{\text{V1. Vorgänger}}(\sigma_{\text{V2. Nachfolger=5216} \wedge \text{V1.Nachfolger = V2.Vorgänger}}(\rho_{\text{V1}}(\text{voraussetzen}) \times \rho_{\text{V2}}(\text{voraussetzen})))$$

- Umbenennung von Attributen

$$\rho_{\text{Voraussetzung} \leftarrow \text{Vorgänger}}(\text{voraussetzen})$$

Formale Definition der Algebra

- Basisausdrücke
- Relation der Datenbank oder
- konstante Relationen

Operationen

- Selektion: $\sigma_p (E_1)$
- Projektion: $\Pi_S (E_1)$
- Kartesisches Produkt: $E_1 \times E_2$
- Umbenennung: $\rho_V (E_1)$, $\rho_{A \leftarrow B} (E_1)$
- Vereinigung: $E_1 \cup E_2$
- Differenz: $E_1 - E_2$

Weitere Operationen können aus diesen zusammengesetzt werden →

Der natürliche Verbund (Join)

Gegeben seien:

- $R(A_1, \ldots, A_m, B_1, \ldots, B_k)$
- $S(B_1, \ldots, B_k, C_1, \ldots, C_n)$

$$R \bowtie S = \Pi_{A1,\ldots,\, Am,\, R.B1,\ldots,\, R.Bk,\, C1,\ldots,\, Cn}(\sigma_{R.B1=S.B1 \,\wedge \ldots \wedge\, R.Bk = S.Bk}(R x S))$$

$R \bowtie S$											
$R - S$				$R \cap S$				$S - R$			
A_1	A_2	…	A_m	B_1	B_2	…	B_k	C_1	C_2	…	C_n
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Professoren			
PersNr	Name	Rang	Raum
2125	Sokrates	C4	226
2126	Russel	C4	232
2127	Kopernikus	C3	310
2133	Popper	C3	52
2134	Augustinus	C3	309
2136	Curie	C4	36
2137	Kant	C4	7

Studenten		
MatrNr	Name	Semester
24002	Xenokrates	18
25403	Jonas	12
26120	Fichte	10
26830	Aristoxenos	8
27550	Schopenhauer	6
28106	Carnap	3
29120	Theophrastos	2
29555	Feuerbach	2

Vorlesungen			
VorlNr	Titel	SWS	gelesen Von
5001	Grundzüge	4	2137
5041	Ethik	4	2125
5043	Erkenntnistheorie	3	2126
5049	Mäeutik	2	2125
4052	Logik	4	2125
5052	Wissenschaftstheorie	3	2126
5216	Bioethik	2	2126
5259	Der Wiener Kreis	2	2133
5022	Glaube und Wissen	2	2134
4630	Die 3 Kritiken	4	2137

voraussetzen	
Vorgänger	Nachfolger
5001	5041
5001	5043
5001	5049
5041	5216
5043	5052
5041	5052
5052	5259

hören	
MatrNr	VorlNr
26120	5001
27550	5001
27550	4052
28106	5041
28106	5052
28106	5216
28106	5259
29120	5001
29120	5041
29120	5049
29555	5022
25403	5022

Assistenten			
PerslNr	Name	Fachgebiet	Boss
3002	Platon	Ideenlehre	2125
3003	Aristoteles	Syllogistik	2125
3004	Wittgenstein	Sprachtheorie	2126
3005	Rhetikus	Planetenbewegung	2127
3006	Newton	Keplersche Gesetze	2127
3007	Spinoza	Gott und Natur	2126

prüfen			
MatrNr	VorlNr	PersNr	Note
28106	5001	2126	1
25403	5041	2125	2
27550	4630	2137	2

Join-Beispiel

Studenten $\bowtie$ hören			
MatrNr	Name	Semester	VorlNr
26120	Fichte	10	5001
27550	Jonas	12	5022
28106	Carnap	3	4052
...	...	...	...

Drei-Wege-Join

(Studenten $\bowtie$ hören) $\bowtie$ Vorlesungen

(Studenten $\bowtie$ hören) $\bowtie$ Vorlesungen						
MatrNr	Name	Semester	VorlNr	Titel	SWS	gelesenVon
26120	Fichte	10	5001	Grundzüge	4	2137
27550	Jonas	12	5022	Glaube und Wissen	2	2134
28106	Carnap	3	4052	Wissenschftstheorie	3	2126
...	...	...	...	...	...	...

Allgemeiner Join (Theta-Join)

- Gegeben seien folgende Relationen(-Schemata)
  - R(A1, ..., An) und
  - S(B1, ..., Bm)

$$R \bowtie_{\theta} S = \sigma_{\theta}(R \times S)$$

$$R \bowtie_{\theta} S$$

$R \bowtie_{\theta} S$							
R				S			
A_1	A_2	...	A_n	B_1	B_2	...	B_m
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Andere Join-Arten

- natürlicher Join

L		
A	B	C
a_1	b_1	c_1
a_2	b_2	c_2

$\bowtie$

R		
C	D	E
c_1	d_1	e_1
c_3	d_2	e_2

=

Resultat				
A	B	C	D	E
a_1	b_1	c_1	d_1	e_1

- linker äußerer Join

L		
A	B	C
a_1	b_1	c_1
a_2	b_2	c_2

$⟕$

R		
C	D	E
c_1	d_1	e_1
c_3	d_2	e_2

=

Resultat				
A	B	C	D	E
a_1	b_1	c_1	d_1	e_1
a_2	b_2	c_2	-	-

- rechter äußerer Join

L		
A	B	C
a_1	b_1	c_1
a_2	b_2	c_2

⟖

R		
C	D	E
c_1	d_1	e_1
c_3	d_2	e_2

=

Resultat				
A	B	C	D	E
a_1	b_1	c_1	d_1	e_1
-	-	c_3	d_2	e_2

Andere Join-Arten

- äußerer Join

L		
A	B	C
a_1	b_1	c_1
a_2	b_2	c_2

⟗

R		
C	D	E
c_1	d_1	e_1
c_3	d_2	e_2

=

Resultat				
A	B	C	D	E
a_1	b_1	c_1	d_1	e_1
a_2	b_2	c_2	-	-
-	-	c_3	d_2	e_2

- Semi-Join von L mit R

L		
A	B	C
a_1	b_1	c_1
a_2	b_2	c_2

⋉

R		
C	D	E
c_1	d_1	e_1
c_3	d_2	e_2

=

Resultat		
A	B	C
a_1	b_1	c_1

Andere Join-Arten (Forts.)

- Semi-Join von R mit L

L		
A	B	C
a_1	b_1	c_1
a_2	b_2	c_2

$\ltimes$

R		
C	D	E
c_1	d_1	e_1
c_3	d_2	e_2

=

Resultat		
C	D	E
c_1	d_1	e_1

Die relationale Division

Bsp.: Finde MatrNr der Studenten, die **alle** vierstündigen Vorlesungen hören

$$L := \Pi_{\text{VorlNr}}(\sigma_{\text{SWS=4}}(\text{Vorlesungen}))$$

$$\text{hören} \div \overbrace{\Pi_{\text{VorlNr}}(\sigma_{\text{SWS=4}}(\text{Vorlesungen}))}^{L}$$

Definition der Division

- $t \in R \div S$, falls für jedes $ts \in S$ ein $tr \in R$ existiert, so dass gilt:
  - $tr.S = ts.S$
  - $tr.(R-S) = t$

R	
M	V
m_1	v_1
m_1	v_2
m_1	v_3
m_2	v_2
m_2	v_3

$\div$

S
V
v_1
v_2

$=$

$R \div S$
M
m_1

- Die Division $R \div S$ kann auch durch Differenz, Kreuzprodukt und Projektion ausgedrückt werden.

$$R \div S = \Pi_{(R-S)}(R) - \Pi_{(R-S)}((\Pi_{(R-S)}(R) \times S) - R)$$

Mengendurchschnitt

Als Beispielanwendung für den Mengendurchschnitt (Operatorsymbol $\cap$) betrachten wir folgende Anfrage: Finde die *PersNr* aller C4-Professoren, die mindestens eine Vorlesung halten.

$$\Pi_{\text{PersNr}}(\rho_{\text{PersNr}\leftarrow\text{gelesenVon}}(\text{Vorlesungen})) \cap \Pi_{\text{PersNr}}(\sigma_{\text{Rang=C4}}(\text{Professoren}))$$

- Mengendurchschnitt nur auf zwei Argumentrelationen mit gleichem Schema anwendbar
- Deshalb ist die Umbenennung des Attribute *gelesenVon* in *PersNr* in der Relation *Vorlesungen* notwendig
- Der Mengendurchschnitt zweier Relationen $R \cap S$ kann durch die Mengendifferenz wie folgt ausgedrückt weden:

$$R \cap S = R - (R - S)$$

Der relationale Tupel-Kalkül

Eine Anfrage im relationalen Tupel-Kalkül hat die Form

$$\{t \mid P(t)\}$$

mit P(t) Formel.

Beispiele:

- C4-Professoren
  - $\{p \mid p \in \text{Professoren} \wedge p.\text{Rang} = \text{'C4'}\}$

- Studenten mit mindestens einer Vorlesung von Curie

$$\begin{aligned}\{s \mid s \in \text{Studenten} \\ &\wedge \exists h \in \text{hören}(s.\text{MatrNr}=h.\text{MatrNr} \\ &\wedge \exists v \in \text{Vorlesungen}(h.\text{VorlNr}=v.\text{VorlNr} \\ &\wedge \exists p \in \text{Professoren}(p.\ \text{PersNr}=v.\text{gelesenVon} \\ &\wedge p.\text{Name} = \text{'Curie'})))\}\end{aligned}$$

- Wer hat **alle** vierstündigen Vorlesungen gehört

$\{s \mid s \in \text{Studenten} \land \forall v \in \text{Vorlesungen } (v.\text{SWS}=4 \Rightarrow$
$\exists h \in \text{hören}(h.\text{VorlNr}=v.\text{VorlNr} \land h.\text{MatrNr}= s.\text{MatrNr}))\}$

Definition des Tupelkalküls

Atome

- s $\mid$ R, mit s Tupelvariable und R Relationenname
- s.A ϕt.B, mit s und t Tupelvariablen, A und B Attributnamen und ϕ Vergleichsperator ($=$, $\neq$, $\leq$, ...)
- s. A ϕ c mit c Konstante

Formeln

- Alle Atome sind Formeln
- Ist P Formel, so auch $\neg$P und (P)
- Sind P_1 und P_2 Formeln, so auch $P_1 \wedge P_2$, $P_1 \vee P_2$ und $P_1 \Rightarrow P_2$
- Ist P(t) Formel mit freier Variable t, so auch

  $\forall t \in R(P(t))$ und $\exists\, t \in R(P(t))$

Sicherheit

- Einschränkung auf Anfragen mit endlichem Ergebnis.
- Die folgende Beispielanfrage

$$\{n \mid \neg (n \in \text{Professoren})\}$$

  ist nicht sicher, denn das Ergebnis ist unendlich.
- Lösung durch Zusatzbedingung: Das Ergebnis des Ausdrucks muss Teilmenge der Domäne der Formel sein.
- Die Domäne einer Formel enthält
  - alle in der Formel vorkommenden Konstanten
  - alle Attributwerte von Relationen, die in der Formel referenziert werden

Der relationale Domänenkalkül

Ein Ausdruck des Domänenkalküls hat die Form

$$\{[v_1, v_2, \ldots, v_n] \mid P(v_1, \ldots, v_n)\}$$

mit $v_1, \ldots, v_n$ Domänenvariablen und P Formel.

Beispiel: MatrNr und Namen der Prüflinge von Curie

$\{[m, n] \mid \exists s ([m, n, s] \in \text{Studenten} \wedge \exists v, p, g ([m, v, p, g] \in \text{prüfen}$
$\wedge \exists a, r, b([p, a, r, b] \in \text{Professoren} \wedge a = \text{'Curie'})))\}$

Während im Tupelkalkül Variablen für ganze Tupel stehen, stehen sie hier für einzelne Werte.

Sicherheit des Domänenkalküls

- Sicherheit ist analog zum Tupelkakkül
- zum Beispiel ist

  $$\{[p,n,r,o] \mid \neg ([p,n,r,o] \in \text{Professoren})\}$$

  nicht sicher.
- Zur Definition der Sicherheit benötigen wir:

  Def.: Die *Domäne* eines Prädikats *P* besteht aus der Menge aller in ihm enthaltenen Konstanten sowie aus der Vereinigung der Domänen der in ihr auftretenden Prädikate.

Ein Ausdruck

$$\{[x1, x2, ..., xn] \mid P(x1, x2, ..., xn)\}$$

ist sicher, falls folgende drei Bedingungen gelten:

1. Falls Tupel $[c_1, c_2, ..., c_n]$ mit Konstante c_i im Ergebnis enthalten ist, so muss jedes c_i $(1 \leq i \leq n)$ in der Domäne von P enthalten sein.

2. Für jede existenz-quantifizierte Teilformel $\exists x(P_1(x))$ muss gelten, dass P_1 nur für Elemente aus der Domäne von P_1 erfüllbar sein kann - oder evtl. für gar keine. Mit anderen Worten, wenn für eine Konstante c das Prädikat $P_1(c)$ erfüllt ist, so muss c in der Domäne von P_1 enthalten sein.

3. Für jede universal-quantifizierte Teilformel $\forall x(P_1(x))$ muss gelten, dass sie dann und nur dann erfüllt ist, wenn $P_1(x)$ für alle Werte der Domäne von P_1 erfüllt ist. Mit anderen Worten, $P_1(d)$ muss für alle d, die nicht in der Domäne von P_1 enthalten sind, auf jeden Fall erfüllt sein.

Ausdruckskraft

Die drei Sprachen

1. relationale Algebra,
2. relationaler Tupelkalkül, eingeschränkt auf sichere Ausdrücke und
3. relationaler Domänenkalkül, eingeschränkt auf sichere Ausdrücke

sind **gleich mächtig**. Dies deutet darauf hin, dass man eine kanonische Stufe der Ausdrucksstärke erreicht hat.

(Ähnliches gilt z.B. für die (Turing-)Berechenbarkeit, die unabhängig in verschiedenen Theorien äquivalent definiert wurde.)

Ausdruckskraft

Warum erlaubt man nicht gleich die ganze Prädikatenlogik als Anfragesprache?

- Anfragen würden nicht unbedingt terminieren, denn es kann
  - unendliche Ergebnise geben (unsichere Ausdrücke) und
  - noch schlimmer: die Prädikatenlogik ist nicht entscheidbar.

Ausblick: Manchmal will man eigentlich etwas mehr Ausdrucksstärke als relationale Algebra:

- Datalog bietet Rekursion (z.B.: Welche Metrostationen sind an einem Streiktag erreichbar?)
- F-Logic bietet Objektorientierung, Operatorüberladung, etc..
- OWL DL bietet gerade noch entscheidbare Logik (siehe Vorlesung „Künstliche Intelligenz“)
- ...

Relationale Algebra und SQL

SQL ähnelt vom Aussehen her dem relationalen Tupelkalkül.

Die Semantik von SQL wird aber unter Verwendung der relationalen Algebra spezifiziert, wie wir weiter hinten sehen werden.