

	Lehrveranstaltung	Grundlagen von Datenbanken		WS 2016/17
	Aufgabenzettel	3 (Lösungsvorschläge)		
	Gesamtpunktzahl	40		
	Ausgabe	Mi. 16.11.2016	Abgabe	Fr. 02.12.2016

1 Informationsmodellierung mit dem Entity-Relationship-Modell

[16 P.]

Erfassen Sie die im Folgenden beschriebenen Informationsstrukturen in einem ER-Diagramm. Beziehen Sie sich dabei genau auf die gegebene Beschreibung, ohne weiteres Wissen zu möglicherweise ähnlichen Anwendungsbereichen einfließen zu lassen. Markieren Sie in Ihrem Entwurf Primärschlüssel durch Unterstreichung und konkretisieren Sie die Abbildungstypen durch **Kardinalitätsrestriktionen** (Notation: [min;max]).

Verwenden Sie unbedingt die aus der Vorlesung bekannte Notation.

Benutzen Sie möglichst wenige Entitäten (Ausnahme: Vererbung). Existenzabhängigkeiten sollen NICHT modelliert werden.

Für jedes Vereinsmitglied werden neben einer eindeutigen Mitgliedsnummer auch Name und Eintrittsdatum gespeichert. Musiker sind besondere Vereinsmitglieder, die Instrumente spielen können. Ein Instrument besitzt eine eindeutige Bezeichnung und mehrere Zubehörteile. Ein Musiker kann beliebig viele Instrumente spielen und ein Instrument kann von beliebig vielen Musikern gespielt werden.

Außerdem kann ein Musiker beliebig viele Musikstücke komponieren; ein Musikstück besitzt einen eindeutigen Titel und wird von genau einem Musiker komponiert. Chorleiter sind besondere Musiker, die eine gewisse Anzahl an Chorjahren aufweisen.

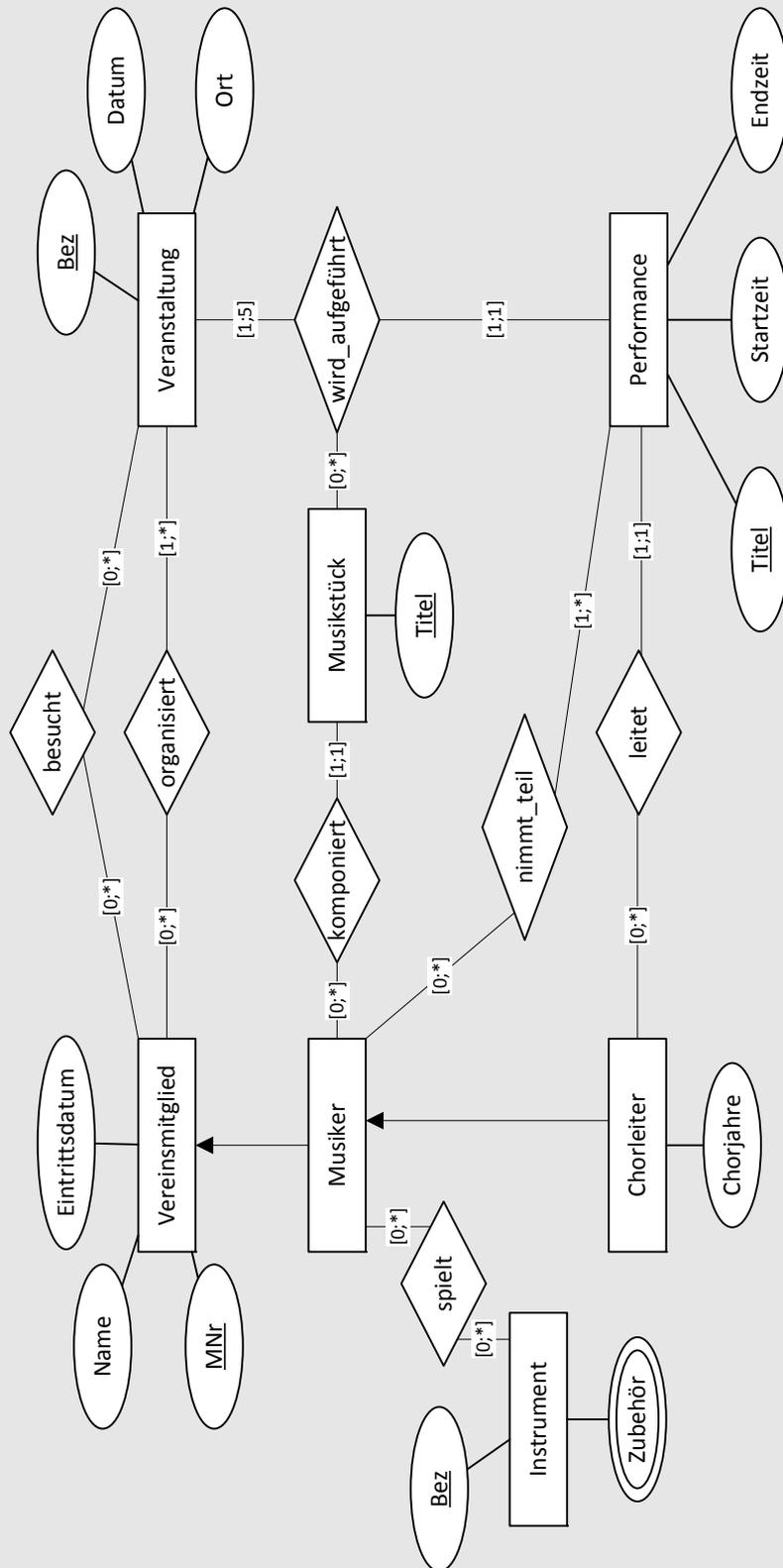
Weiterhin gibt es Performances, die jeweils eine Startzeit, eine Endzeit sowie einen eindeutigen Titel besitzen und von genau einem Chorleiter geleitet werden. Jeder Chorleiter kann allerdings beliebig viele Performances leiten. An einer Performance nimmt mindestens ein Musiker teil. Jeder Musiker kann an beliebig vielen Performances teilnehmen.

Es gibt außerdem Veranstaltungen, die eindeutig über ihre jeweiligen Bezeichnungen identifizierbar sind und jeweils ein Datum und einen Ort besitzen. Ein Musikstück wird im Rahmen einer Performance auf einer Veranstaltung aufgeführt. Eine Performance wird dabei nur genau einmal aufgeführt. Jede Veranstaltung umfasst mindestens eine und maximal fünf Performances. Ein Musikstück kann beliebig oft aufgeführt werden.

Eine Veranstaltung wird von mindestens einem Vereinsmitglied organisiert und von beliebig vielen Vereinsmitgliedern besucht. Jedes Vereinsmitglied kann beliebig viele Veranstaltungen organisieren und auch beliebig viele Veranstaltungen besuchen.



Lösungsvorschlag:

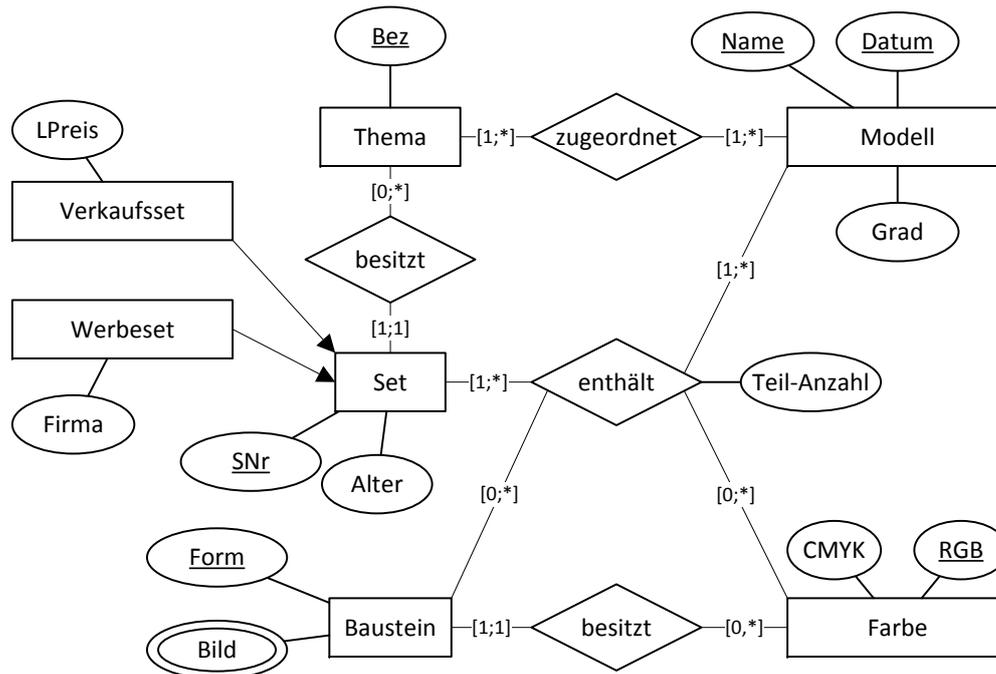


	Lehrveranstaltung	Grundlagen von Datenbanken		WS 2016/17
	Aufgabenzettel	3 (Lösungsvorschläge)		
	Gesamtpunktzahl	40		
	Ausgabe	Mi. 16.11.2016	Abgabe	Fr. 02.12.2016

2 Abbildung eines ER-Diagramms auf das relationale Datenmodell

[11 P.]

Gegeben sei folgendes ER-Diagramm:



- a) Entwickeln Sie aus dem dargestellten ER-Diagramm ein entsprechendes relationales Datenbankschema anhand der in der Vorlesung erläuterten Abbildungsregeln. Stellen Sie sicher, dass Ihr Datenbankschema die minimale Anzahl von Relationen aufweist. Verwenden Sie das **Partitionierungsmodell**, um die Vererbung abzubilden. Stellen Sie das resultierende DB-Schema dar, indem Sie die notwendigen Relationenschemata in der Form

$Relation(Attribut_1, Attribut_2, \dots, Attribut_n)$

anführen und dabei jeweils den Primärschlüssel unterstreichen. Gegebenenfalls enthaltene Fremdschlüssel sind zu „unterstricheln“ und durch die aus den Übungen bekannte Pfeilnotation zu spezifizieren:

$\underline{Attr_i} \rightarrow Rel_b. \underline{Attr_j}$

Hinweis zur Semantik von 1:n-Beziehungen: Nach dem dargestellten ER-Diagramm besitzt jeder Baustein genau eine Farbe, während jede Farbe für beliebig viele Bausteine benutzt werden kann. Die Semantik aller anderen 1:n-Beziehungen ist entsprechend.

	Lehrveranstaltung	Grundlagen von Datenbanken		WS 2016/17
	Aufgabenzettel	3 (Lösungsvorschläge)		
	Gesamtpunktzahl	40		
	Ausgabe	Mi. 16.11.2016	Abgabe	Fr. 02.12.2016

Lösungsvorschlag:

Thema(Bez)

Modell(Name, Datum, Grad)

*zugeordnet(Thema → Thema.Bez,
(Modell, MDatum) → (Modell.Name, Modell.Datum))*

Farbe(RGB, CMYK)

Baustein(Form, Farbe → Farbe.RGB)

Bild(Baustein → Baustein.Form, Bild)

Set(SNR, Alter, Thema → Thema.Bez)

Verkaufsset(SNR → Set.SNR, LPreis)

Werbeset(SNR → Set.SNR, Firma)

*enthält(Set → Set.SNr, Baustein → Baustein.Form, Farbe → Farbe.RGB,
(Modell, MDatum) → (Modell.Name, Modell.Datum), Teil-Anzahl)*

b) Wieso ist die Verwendung des Hausklassenmodells problematisch?

[1 P.]

Lösungsvorschlag:

Wird die Vererbung mit dem Hausklassenmodell umgesetzt, so referenziert die enthält-Relation lediglich den Entitätentyp Set, aber nicht die erbenenden Entitätentypen Verkaufsset und Werbeset.

	Lehrveranstaltung	Grundlagen von Datenbanken		WS 2016/17
	Aufgabenzettel	3 (Lösungsvorschläge)		
	Gesamtpunktzahl	40		
	Ausgabe	Mi. 16.11.2016	Abgabe	Fr. 02.12.2016

3 Relationale Algebra und SQL

[9 P.]

Im Folgenden sind vier Relationen mit den unten dargestellten Schemata und Ausprägungen gegeben:

Buch	Titel	Erscheinungsjahr	Seitenzahl	Verlag
	Schall und Wahn	1929	304	Diogenes
	Als ich im Sterben lag	1930	173	Diogenes
	Hundert Jahre Einsamkeit	1967	480	Fischer
	Der Fremde	1942	160	rororo
	Krieg und Frieden	1869	1536	Anaconda
	Anna Karenina	1878	991	Anaconda
	Schuld und Sühne	1866	752	Deutscher Taschenbuch Verlag
	Requiem für einen Traum	1978	316	Rowohlt
	Der Talisman	1984	714	Heyne

Person	PID	Vorname	Nachname	Lieblingsbuch
	1	Leo	Tolstoi	Schuld und Sühne
	2	Fjodor	Dostojewski	Krieg und Frieden
	3	Hubert	Selby	Der Fremde
	4	Albert	Camus	Schuld und Sühne
	5	William	Faulkner	Schuld und Sühne
	6	Stephen	King	Hundert Jahre Einsamkeit
	7	Peter	Straub	Schall und Wahn
	8	Gabriel	García Márquez	Requiem für einen Traum

Lieblingsbuch → Buch.Titel

Schreibt	Autor	Buch
	1	Krieg und Frieden
	1	Anna Karenina
	2	Schuld und Sühne
	3	Requiem für einen Traum
	4	Der Fremde
	5	Schall und Wahn
	5	Als ich im Sterben lag
	6	Der Talisman
	7	Der Talisman
	8	Hundert Jahre Einsamkeit

Autor → Person.PID, Buch → Buch.Titel

Begutachtet	Lektor	Buch
	2	Anna Karenina
	1	Schuld und Sühne
	8	Requiem für einen Traum
	6	Requiem für einen Traum
	5	Der Fremde
	4	Als ich im Sterben lag
	2	Krieg und Frieden
	7	Hundert Jahre Einsamkeit

Lektor → Person.PID, Buch → Buch.Titel

	Lehrveranstaltung	Grundlagen von Datenbanken		WS 2016/17
	Aufgabenzettel	3 (Lösungsvorschläge)		
	Gesamtpunktzahl	40		
	Ausgabe	Mi. 16.11.2016	Abgabe	Fr. 02.12.2016

a) Übersetzen Sie die folgenden umgangssprachlich formulierten Anfragen in einen zugehörigen Ausdruck der relationalen Algebra. Werten Sie die Ausdrücke aus und geben Sie jeweils die Ergebnisrelation an. [3 P.]

i) „Die Titel der Bücher, die mehr als 200 Seiten besitzen und nach 1950 erschienen sind.“

Lösungsvorschlag:

$\pi_{\text{Titel}}(\sigma_{\text{Seitenzahl} > 200 \wedge \text{Erscheinungsjahr} > 1950} \text{Buch})$

Ergebnis = {(Hundert Jahre Einsamkeit), (Requiem für einen Traum), (Der Talisman)}

ii) „Die Vor- und Nachnamen der Personen, die das Buch mit dem Titel 'Der Talisman' geschrieben haben.“

Lösungsvorschlag:

$\pi_{\text{Vorname}, \text{Nachname}}(\text{Person} \bowtie_{\text{PID}=\text{Autor}} (\sigma_{\text{Buch}='DerTalisman'} \text{Schreibt}))$

Ergebnis = {(Stephen, King), (Peter, Straub)}

iii) „Die Vor- und Nachnamen der Personen, die ihr Lieblingsbuch begutachtet haben.“

Lösungsvorschlag:

$\pi_{\text{Vorname}, \text{Nachname}}(\text{Person} \bowtie_{\text{PID}=\text{Lektor} \wedge \text{Lieblingsbuch}=\text{Buch}} \text{Begutachtet})$

Ergebnis = {(Leo, Tolstoi), (Gabriel, García Márquez), (Fjodor, Dostojewski)}

b) Interpretieren Sie die folgenden relationalen Ausdrücke, indem Sie eine umgangssprachliche Beschreibung sowie die Ergebnisrelation angeben (Hinweis: die Syntax $A \leftarrow B$ entspricht dem relationalen Umbenennungs-Operator, der ein Attribut mit dem Namen B in A umbenennt). [3 P.]

i) $(\pi_{\text{Titel}}(\text{Buch}) - \pi_{\text{Titel}}(\rho_{\text{Titel} \leftarrow \text{Buch}}(\text{Begutachtet}))) \bowtie \text{Buch}$

Lösungsvorschlag:

„Die Titel, die Seitenzahl, das Erscheinungsjahr und der Verlag der Bücher, die nie von einem Lektor begutachtet wurden.“

Ergebnis = {(Schall und Wahn, 1929, 304, Diogenes), (Der Talisman, 1984, 714, Heyne)}

	Lehrveranstaltung	Grundlagen von Datenbanken		WS 2016/17
	Aufgabenzettel	3 (Lösungsvorschläge)		
	Gesamtpunktzahl	40		
	Ausgabe	Mi. 16.11.2016	Abgabe	Fr. 02.12.2016

$$\text{ii) } \pi_{\text{Vorname, Nachname}}((\pi_{\text{Autor}}(\text{Schreibt}) \cap \pi_{\text{Autor}}(\rho_{\text{Autor} \leftarrow \text{Lektor}} \text{Begutachtet})) \bowtie_{\text{Autor=PID}} \text{Person})$$

Lösungsvorschlag:

„Die Vor- und Nachnamen der Personen, die sowohl als Autor als auch als Lektor tätig waren.“

Ergebnis = {(Leo, Tolstoi), (Fjodor, Dostojewski), (Albert, Camus), (William, Faulkner), (Stephen, King), (Peter, Straub), (Gabriel, García Márquez)}

$$\text{iii) } \pi_{\text{Vorname, Nachname}}(\text{Schreibt} \bowtie_{\text{Autor=Lektor}} (\text{Begutachtet} \bowtie_{\text{Lektor=PID}} \text{Person}))$$

Lösungsvorschlag:

„Die Vor- und Nachnamen der Personen, die sowohl als Autor als auch als Lektor tätig waren.“

Ergebnis = {(Leo, Tolstoi), (Fjodor, Dostojewski), (Albert, Camus), (William, Faulkner), (Stephen, King), (Peter, Straub), (Gabriel, García Márquez)}

c) Übersetzen Sie die folgenden Ausdrücke in entsprechende SQL-Ausdrücke und geben Sie die jeweiligen Ergebnisrelationen an. [3 P.]

i) Vor- und Nachnamen der Autoren ohne Duplikate von Büchern mit mehr als 500 Seiten.

Lösungsvorschlag:

```
SELECT DISTINCT Vorname, Nachname
FROM Person, Schreibt, Buch
WHERE Person.PID=Schreibt.Autor
AND Schreibt.Buch=Buch.Titel
AND Buch.Seitenzahl>500
```

= {(Leo, Tolstoi), (Fjodor, Dostojewski), (Stephen, King), (Peter, Straub)}

ii) Die Titel aller Bücher, die einen Autor haben, der schon mal ein Buch begutachtet hat.

Lösungsvorschlag:

```
SELECT DISTINCT s.Buch
FROM Schreibt s, Begutachtet b
WHERE s.Autor=b.Lektor
```

oder (weniger effizient):

```
SELECT DISTINCT Buch
FROM Schreibt
WHERE Autor IN (SELECT Lektor
FROM Begutachtet)
```

	Lehrveranstaltung	Grundlagen von Datenbanken		WS 2016/17
	Aufgabenzettel	3 (Lösungsvorschläge)		
	Gesamtpunktzahl	40		
	Ausgabe	Mi. 16.11.2016	Abgabe	Fr. 02.12.2016

oder (noch weniger effizient):

```
SELECT DISTINCT s.Buch
FROM Schreibt s
WHERE EXISTS (SELECT *
              FROM Begutachtet b
              WHERE s.Autor=b.Lektor)
```

= {(Als ich im Sterben lag), (Anna Karenina), (Der Fremde), (Der Talisman), (Hundert Jahre Einsamkeit), (Krieg und Frieden), (Schall und Wahn), (Schuld und Sühne)}

iii) $\pi_{\text{Vorname, Nachname}}(Person \bowtie (\pi_{PID}(Person) - \pi_{Lektor}(Begutachtet)))$

Lösungsvorschlag:

Natürlichsprachlich: Die Vor- und Nachnamen aller Personen, die noch kein Buch begutachtet haben, ohne Duplikate.

```
SELECT DISTINCT p.Vorname, p.Nachname
FROM Person p
WHERE p.PID NOT IN (SELECT b.Lektor
                  FROM Begutachtet b)
```

oder (weniger effizient):

```
SELECT DISTINCT p.Vorname, p.Nachname
FROM Person p
WHERE NOT EXISTS (SELECT b.Lektor
                 FROM Begutachtet b
                 WHERE p.PID=b.Lektor)
```

oder (ziemlich direkt übersetzt):

```
SELECT DISTINCT p.Vorname, p.Nachname
FROM Person p, (SELECT PID FROM Person
               WHERE PID NOT IN (SELECT Lektor
                               FROM Begutachtet)) AS p2
WHERE p.PID=p2.PID
```

= {(Hubert, Selby)}

	Lehrveranstaltung	Grundlagen von Datenbanken		WS 2016/17
	Aufgabenzettel	3 (Lösungsvorschläge)		
	Gesamtpunktzahl	40		
	Ausgabe	Mi. 16.11.2016	Abgabe	Fr. 02.12.2016

4 Algebraische Optimierung

[4 P.]

Gegeben seien die aus Aufgabe 3 bekannten Relationenschemata.

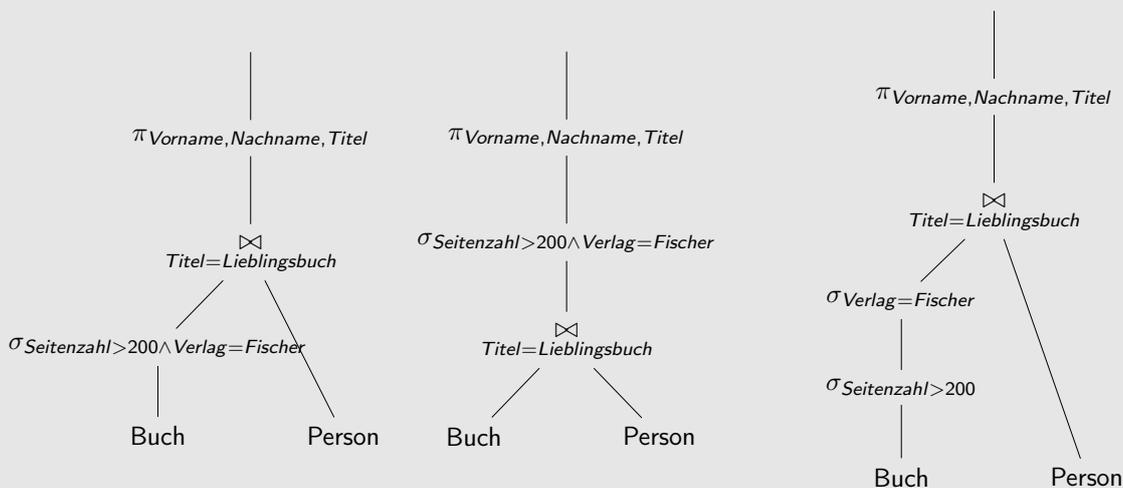
In jeder der folgenden drei Unteraufgaben sind jeweils drei relationale Ausdrücke angegeben. Alle drei Ausdrücke liefern dasselbe Ergebnis zurück und sind daher semantisch äquivalent. Die jeweiligen Ausdrücke unterscheiden sich jedoch in ihrem Optimierungsgrad. Zeichnen Sie zu jedem relationalen Ausdruck einen Operatorbaum und bestimmen Sie, welcher der drei Operatorbäume den höchsten Optimierungsgrad besitzt. Begründen Sie Ihre Entscheidung mit Hilfe der in der Vorlesung behandelten Optimierungsheuristiken I-VII.

$$A1. \pi_{\text{Vorname}, \text{Nachname}, \text{Titel}} \left(\left(\sigma_{\text{Seitenzahl} > 200 \wedge \text{Verlag} = \text{Fischer}}(\text{Buch}) \right) \bowtie_{\text{Titel} = \text{Lieblingsbuch}} \text{Person} \right)$$

$$A2. \pi_{\text{Vorname}, \text{Nachname}, \text{Titel}} \left(\sigma_{\text{Seitenzahl} > 200 \wedge \text{Verlag} = \text{Fischer}}(\text{Buch} \bowtie_{\text{Titel} = \text{Lieblingsbuch}} \text{Person}) \right)$$

$$A3. \pi_{\text{Vorname}, \text{Nachname}, \text{Titel}} \left(\left(\sigma_{\text{Verlag} = \text{Fischer}}(\sigma_{\text{Seitenzahl} > 200}(\text{Buch})) \right) \bowtie_{\text{Titel} = \text{Lieblingsbuch}} \text{Person} \right)$$

Lösungsvorschlag:



A1:

A2:

A3:

Gemäß Optimierungsheuristik I sind in den Ausdrücken A1 und A3 alle Selektionen soweit wie möglich zu den einzelnen Relationen verschoben wurden. In Ausdruck A1 sind zudem beide Selektionen zu einer einzelnen zusammengefasst (Optimierungsheuristik IV). Ausdruck A1 besitzt demnach den höchsten Optimierungsgrad.