	Lehrveranstaltung	Grundlagen von Datenbanken		WS 2017/18
	Aufgabenzettel	3 (Lösungsvorschläge)		
	Gesamtpunktzahl	41		
	Ausgabe	Mi. 15.11.2017	Abgabe	Fr. 01.12.2017

1 Informationsmodellierung mit dem Entity-Relationship-Modell

[15 P.]

McAcid's benötigt ein neues Burgastisches Kassensystem, bei dem eine relationale Datenbank verwendet werden soll.

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]). Sollte eine Restriktion nicht aus dem Text hervorgehen, kann ein unrestringierter Abbildungstyp ([0;*]) angenommen werden.

Verwenden Sie unbedingt die aus der Vorlesung bekannte Notation. Benutzen Sie möglichst wenige Entitäten (Ausnahme: Vererbung). Eine etwaige (Existenz-)Abhängigkeit soll nur dann modelliert werden, wenn dies eindeutig aus der Beschreibung hervorgeht (z.B. durch die Angabe eines schwachen Schlüssels).

Kunden besitzen eine eindeutige Kennung (KID), einen Namen und möglicherweise mehrere Adressen. Eine Bestellung wird eindeutig über ihre Quittungsnummer (QNr) identifiziert und hat einen Zeitstempel. Während Kunden beliebig viele Bestellungen tätigen können, wird jede Bestellung jedoch von genau einem Kunden getätigt.

Eine Bestellung enthält außerdem mindestens ein Produkt, wobei ein Produkt einen Preis und einen eindeutigen Namen besitzt und in beliebig vielen Bestellungen enthalten sein kann.

Burger, Getränke, Beilagen und Menüs sind Produkte, wobei für ein Getränk zusätzlich die Größe gespeichert wird. Ein Menü beinhaltet genau einen Burger, ein Getränk und eine Beilage. Ein Burger kann maximal zu einem Menü gehören. Ein Getränk und eine Beilage gehören zu beliebig vielen Menüs.

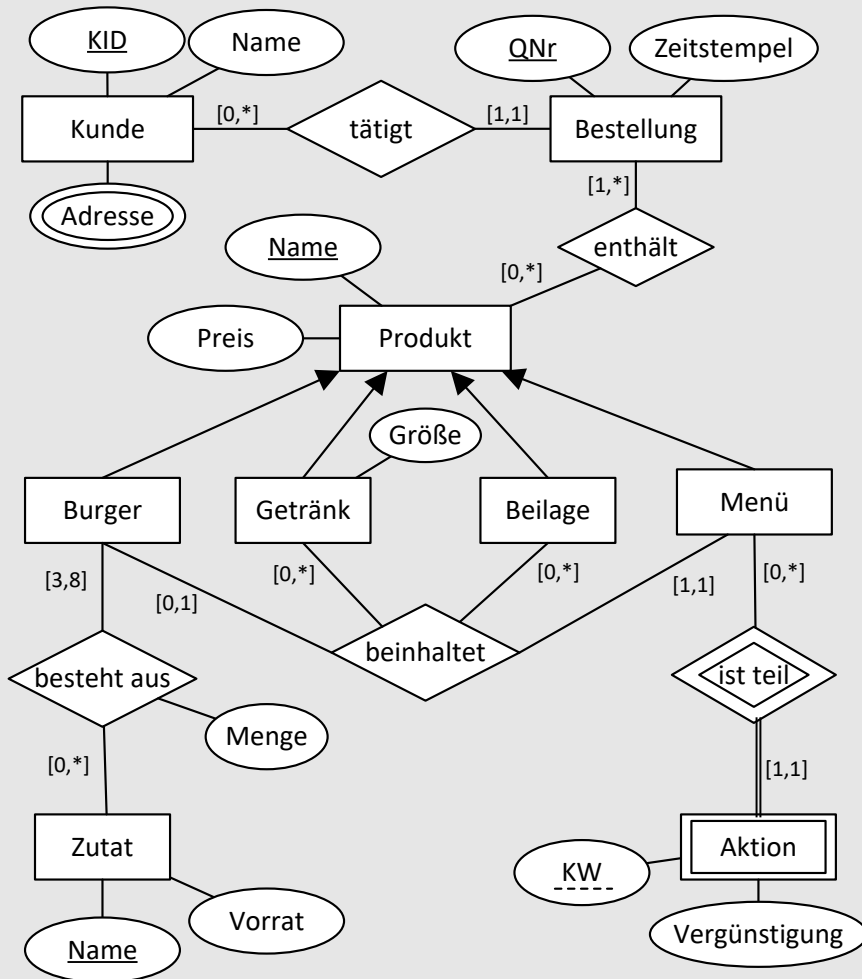
Für eine Zutat, welche eindeutig über ihren Namen identifiziert werden kann, wird der Vorrat gespeichert. Eine Zutat kann Bestandteil beliebig vieler Burger sein. Burger bestehen jedoch aus drei bis acht Zutaten, wobei dabei die Menge von Interesse ist.


Ein Menü kann beliebig häufig teil einer Aktion sein. Einer Aktion ist genau ein Menü zugeordnet. Eine Aktion besitzt eine Kalenderwoche (KW) und eine Vergünstigung. Die Kalenderwoche identifiziert die Aktion nur im Zusammenhang mit dem Namen des zugehörigen Menüs eindeutig.



Lehrveranstaltung	Grundlagen von Datenbanken		WS 2017/18
Aufgabenzettel	3 (Lösungsvorschläge)		
Gesamtpunktzahl	41		
Ausgabe	Mi. 15.11.2017	Abgabe	Fr. 01.12.2017

Lösungsvorschlag:

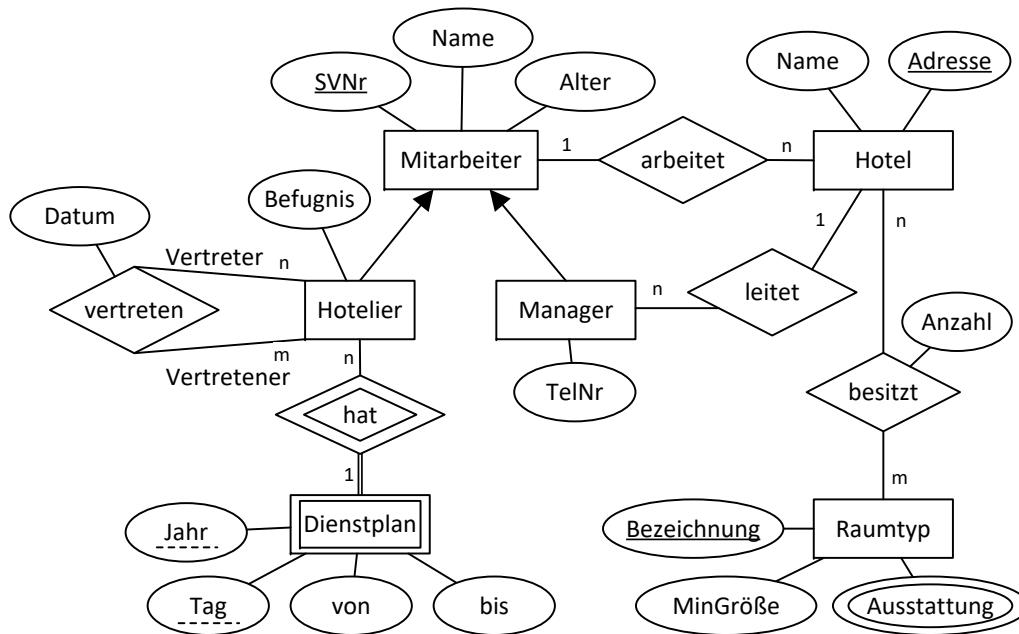


	Lehrveranstaltung	Grundlagen von Datenbanken		WS 2017/18
	Aufgabenzettel	3 (Lösungsvorschläge)		
	Gesamtpunktzahl	41		
	Ausgabe	Mi. 15.11.2017	Abgabe	Fr. 01.12.2017

2 Abbildung eines ER-Diagramms auf das relationale Datenmodell

[9 P.]

Gegeben sei folgendes ER-Diagramm:



Zusätzlich gelte folgende Integritätsbedingung: „Jede Vertretung ist eindeutig identifizierbar durch die Kombination aus Vertreter, Vertretener und Datum der Vertretung.“


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 **vertikale Partitionierung**, um die Vererbung abzubilden. Stellen Sie das resultierende DB-Schema dar, indem Sie die notwendigen Relationenschemata in der Form

$$Relation(\underline{Attribut_1}, \underline{Attribut_2}, \dots, \underline{Attribut_n})$$

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

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

Hinweis zur Semantik von 1:n-Beziehungen: Nach dem dargestellten ER-Diagramm ist jeder Mitarbeiter in max. einem Hotel angestellt. Ein Hotel hat beliebig viele angestellte Mitarbeiter. Die Semantik aller anderen 1:n-Beziehungen ist entsprechend.

	Lehrveranstaltung	Grundlagen von Datenbanken		WS 2017/18
	Aufgabenzettel	3 (Lösungsvorschläge)		
	Gesamtpunktzahl	41		
	Ausgabe	Mi. 15.11.2017	Abgabe	Fr. 01.12.2017

Lösungsvorschlag:

Mitarbeiter(SVNr, Name, Alter, Arbeitsplatz → Hotel.Adresse)

Hotelier(SVNr → Mitarbeiter.SVNr, Befugnis)

Manager(SVNr → Mitarbeiter.SVNr, TelNr)

vertreten(Vertreter → Hotelier.SVNr, Vertretener → Hotelier.SVNr, Datum)


Dienstplan(Hotelier → Hotelier.SVNr, Jahr, Tag, von, bis)

Hotel(Adresse, Name, Manager → Manager.SVNr)

Raumtyp(Bezeichnung, MinGröße)

Ausstattung(Raumtyp → Raumtyp.Bezeichnung, Ausstattung)

besitzt(Hotel → Hotel.Adresse, Raumtyp → Raumtyp.Bezeichnung, Anzahl)

	Lehrveranstaltung	Grundlagen von Datenbanken		WS 2017/18
	Aufgabenzettel	3 (Lösungsvorschläge)		
	Gesamtpunktzahl	41		
	Ausgabe	Mi. 15.11.2017	Abgabe	Fr. 01.12.2017

3 Relationale Algebra

[12 P.]

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

<u>Personal</u>	<u>PID</u>	Vorname	Nachname	Geburt	Wohnort	<u>Abteilung</u>
	4	Peter	Müller	1962-07-25	Hamburg	2
	8	Bianca	Lohse	1982-01-13	Kiel	4
	11	Murat	Sahir	1990-03-16	Hamburg	2
	21	Frank	Siebenstein	1975-12-02	Norderstedt	1
	22	Bernd	Schmidt	1973-11-26	Norderstedt	1
	24	Ulrike	Müller	1963-10-07	Hamburg	2
	31	Jochen	Fuhrmann	1958-05-09	Stade	2

Abteilung → Abteilungen.AID

<u>Abteilungen</u>	<u>AID</u>	Name	<u>Leiter</u>
	1	Controlling	21
	4	Marketing	8
	2	Einkauf	4

Leiter → Personal.PID

<u>ProjektArbeiter</u>	<u>PrID</u>	<u>PID</u>
	15	21
	15	22
	36	8
	36	11
	36	31

PrID → Projekte.PrID, PID → Personal.PID

<u>Projekte</u>	<u>PrID</u>	Name	<u>Leiter</u>	Budget
	15	Prozessoptimierung	22	10.000
	36	B.L.I.C.K.F.A.N.G	8	7.500

Leiter → Personal.PID

- a) Überprüfen Sie die Korrektheit der folgenden Ausdrücke der relationalen Algebra. Prüfen Sie dabei sowohl die Korrektheit der Syntax als auch der Semantik, d. h. ihre Gültigkeit in Bezug auf das vorliegende Datenbankschema. Begründen Sie jeweils Ihre Antwort.


[4 P.]

$$i) \sigma_{Geburt="1962-07-25"}(Abteilungen \bowtie_{Leiter=PID} (\pi_{PID, Vorname, Nachname}(Personal)))$$

Lösungsvorschlag:

Die Selektion anhand der Spalte *Geburt* ist nicht möglich. Diese ist zwar in der Relation *Personal* vorhanden, wurde aber durch die Projektion auf die Spalten *Vorname* und *Nachname* ausgeblendet.

$$ii) \pi_{Vorname}(Personal) \cup \pi_{Geburt}(Personal)$$

	Lehrveranstaltung	Grundlagen von Datenbanken		WS 2017/18
	Aufgabenzettel	3 (Lösungsvorschläge)		
	Gesamtpunktzahl	41		
	Ausgabe	Mi. 15.11.2017	Abgabe	Fr. 01.12.2017

Lösungsvorschlag:

Die Vereinigung ist nicht möglich, da die beiden zu vereinigenden Projektionen nicht vereinbar sind (unterschiedliche Attribute).

iii) $\sigma_{Budget < 1000 \wedge Budget > 5000}(Projekte)$

Lösungsvorschlag:

Dieser Ausdruck liefert zwar in jedem Fall ein leeres Resultat, ist aber dennoch korrekt.

iv) $\pi_{Name}(Abteilungen) \bowtie_{PID=Leiter} \pi_{Vorname, Nachname}(Personal)$

Lösungsvorschlag:

Der Verbund ist nicht möglich, da die Verbundsattribute durch die vorhergehenden Projektionen ausgeblendet wurden.


b) 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. [4 P.]

i) Namen von Projekten, in denen junge Mitarbeiter (Geboren in oder nach 1990) mitarbeiten

Lösungsvorschlag:

$\pi_{Name}(\sigma_{Geburt \geq "1990-01-01"}(Personal) \bowtie ProjektArbeiter \bowtie Projekte)$

Ergebnis = {(B.L.I.C.K.F.A.N.G.)}

	Lehrveranstaltung	Grundlagen von Datenbanken		WS 2017/18
	Aufgabenzettel	3 (Lösungsvorschläge)		
	Gesamtpunktzahl	41		
	Ausgabe	Mi. 15.11.2017	Abgabe	Fr. 01.12.2017

- ii) Persönliche Daten (Vor- und Nachname, Geburtsdatum) aller an dem Projekt *Prozessoptimierung* beteiligten Mitarbeiter.

Lösungsvorschlag:

$\pi_{\text{Vorname, Nachname, Geburt}}(\sigma_{\text{Name}=\text{"Prozessoptimierung"}}(\text{Projekte}) \bowtie \text{ProjektArbeiter} \bowtie \text{Personal})$

Ergebnis = {(Frank, Siebenstein, 1975-12-02), (Bernd, Schmidt, 1973-11-26)}

- iii) Mitarbeiter, die an keinem Projekt mitarbeiten.

Lösungsvorschlag:

$\text{Personal} - \pi_{\text{PID, Vorname, Nachname, Geburt, Wohnort, Abteilung}}(\text{Personal} \bowtie \text{ProjektArbeiter})$

Ergebnis = {(4, Peter, Müller, 1962-07-25, 2), (24, Ulrike, Müller, 1963-10-07, 2)}

- iv) Vor- und Nachnamen der Abteilungskollegen von Jochen Fuhrmann (PID=31).

Lösungsvorschlag:

$\pi_{\text{Vorname, Nachname}}(\sigma_{\text{PID} \neq 31}(\text{Personal} \bowtie \pi_{\text{Abteilung}}(\sigma_{\text{PID}=31}(\text{Personal}))))$

Ergebnis = {(Peter, Müller), (Murat, Sahir), (Ulrike, Müller)}

- c) Interpretieren Sie die folgenden relationalen Ausdrücke, indem Sie eine umgangssprachliche Beschreibung sowie die Ergebnisrelation angeben. [4 P.]

- i) $\pi_{\text{Vorname, Nachname}}(\sigma_{\text{Budget} > 8000}(\text{Projekte}) \underset{\text{Leiter}=\text{PID}}{\bowtie} \text{Personal})$

Lösungsvorschlag:

Vor- und Nachnamen der Projektleiter, die über ein Budget von mehr als 8.000 EUR verfügen.


Ergebnis = {(Bernd, Schmidt)}

- ii) $\pi_{\text{PID, Vorname, Nachname, Geburt, Wohnort, Abteilung}}(\text{Personal} \underset{\text{PID}=\text{Leiter}}{\bowtie} \text{Projekte}) \cap \pi_{\text{PID, Vorname, Nachname, Geburt, Wohnort, Abteilung}}(\text{Personal} \underset{\text{PID}=\text{Leiter}}{\bowtie} \text{Abteilungen})$

Lösungsvorschlag:

Mitarbeiter, die gleichzeitig Projektleiter und Abteilungsleiter sind.

Ergebnis = {(8, Bianca, Lohse, 1982-01-13, Kiel, 4)}

	Lehrveranstaltung	Grundlagen von Datenbanken		WS 2017/18
	Aufgabenzettel	3 (Lösungsvorschläge)		
	Gesamtpunktzahl	41		
	Ausgabe	Mi. 15.11.2017	Abgabe	Fr. 01.12.2017

iii) $\pi_{Name}(\pi_{Abteilung}(\sigma_{Name="B.L.I.C.K.F.A.N.G."}(\text{Projekte}) \bowtie \text{ProjektArbeiter} \bowtie \text{Personal}) \bowtie_{Abteilung=AID} \text{Abteilungen})$

Lösungsvorschlag:

Die Namen aller an dem Projekt *B.L.I.C.K.F.A.N.G.* beteiligten Abteilungen.


Ergebnis = {(Marketing), (Einkauf)}

iv) $\pi_{PrID}(\sigma_{Name="Controlling"}(\text{Abteilungen}) \bowtie_{AID=Abteilung} \text{Personal} \bowtie \text{ProjektArbeiter}) \bowtie \text{Projekte}$

Lösungsvorschlag:

Alle Projekte an denen Mitarbeiter der Abteilung *Controlling* mitwirken.

Ergebnis = {(15, Prozessoptimierung, 22, 10.000)}

	Lehrveranstaltung	Grundlagen von Datenbanken		WS 2017/18
	Aufgabenzettel	3 (Lösungsvorschläge)		
	Gesamtpunktzahl	41		
	Ausgabe	Mi. 15.11.2017	Abgabe	Fr. 01.12.2017

4 Algebraische Optimierung


[5 P.]

Betrachten Sie erneut das Datenbankschema aus Aufgabe 3. In der folgenden Aufgabe sind zwei relationale Ausdrücke angegeben. Beide Ausdrücke liefern dasselbe Ergebnis zurück und sind daher semantisch äquivalent, unterscheiden sich jedoch in ihrem Optimierungsgrad. Zeichnen Sie zu jedem relationalen Ausdruck einen Operatorbaum und bestimmen Sie, welcher der zwei Operatorbäume den höchsten Optimierungsgrad besitzt. Begründen Sie Ihre Entscheidung mit Hilfe der in der Vorlesung behandelten Optimierungsheuristiken I-VII (Folien 62-71, Kapitel 4.).

Für die zugehörige Datenbank werden in dieser Aufgabe jedoch folgende Kardinalitäten angenommen:

$Card(Personal) = 2000$, $Card(Abteilungen) = 40$, $Card(Projekte) = 50$, $Card(ProjekteArbeiter) = 1000$

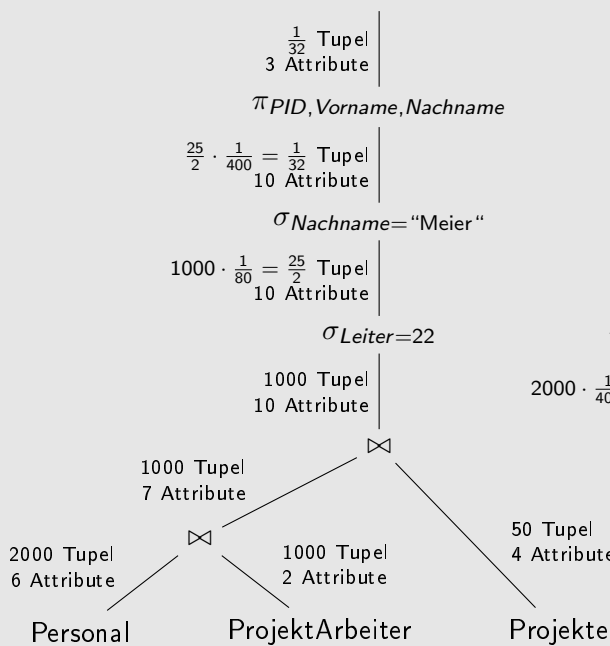
Es gibt 50 verschiedene Werte für das Attribut Wohnort und je 400 verschiedene Vor- und Nachnamen. Jeder 25. Personalangestellte ist ein Projektleiter. Die Namen von Abteilungen sind eindeutig.

	Lehrveranstaltung	Grundlagen von Datenbanken		WS 2017/18
	Aufgabenzettel	3 (Lösungsvorschläge)		
	Gesamtpunktzahl	41		
	Ausgabe	Mi. 15.11.2017	Abgabe	Fr. 01.12.2017

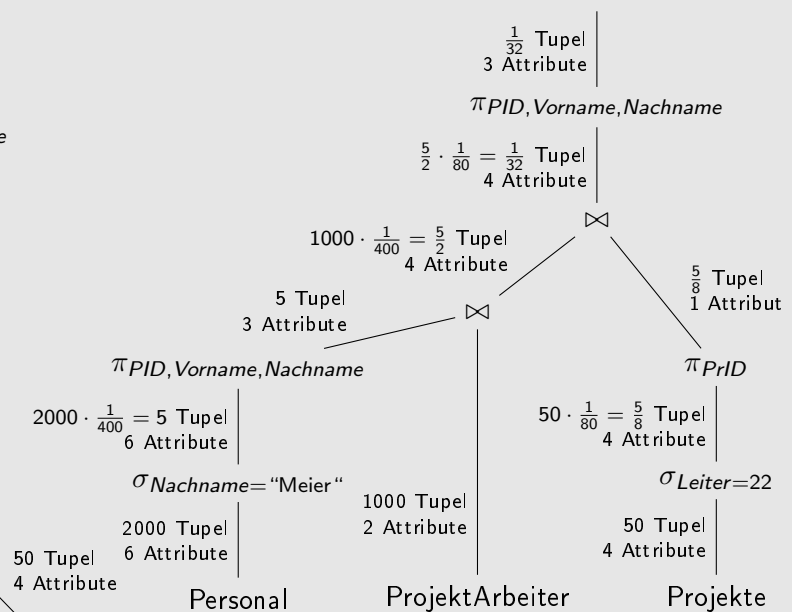
- a) $\pi_{PID, Vorname, Nachname}(\sigma_{Nachname=Meier}(\sigma_{Leiter=22}((Personal \bowtie ProjektArbeiter) \bowtie Projekte)))$
- b) $\pi_{PID, Vorname, Nachname}(((\pi_{PID, Vorname, Nachname}(\sigma_{Nachname="Meier"} Personal)) \bowtie ProjektArbeiter) \bowtie (\pi_{PrID}(\sigma_{Leiter=22} Projekte)))$

Lösungsvorschlag:

a)



b)



Der zweite Operatorbaum besitzt den höchsten Optimierungsgrad, da auch hier Projektionen und Selektionen so früh wie möglich ausgeführt werden. Der Verbund zwischen den Relationen Personal und ProjektArbeiter führt zu einem Zwischenergebnis von etwa $\frac{5}{2}$ Tupeln mit 4 Attributen (vgl. 1000 Tupel mit 7 Attributen) und der nachfolgende Verbund mit der Projekte-Relation liefert $\frac{1}{32}$ Tupel mit 4 Attributen (vgl. 1000 Tupel mit 10 Attributen).