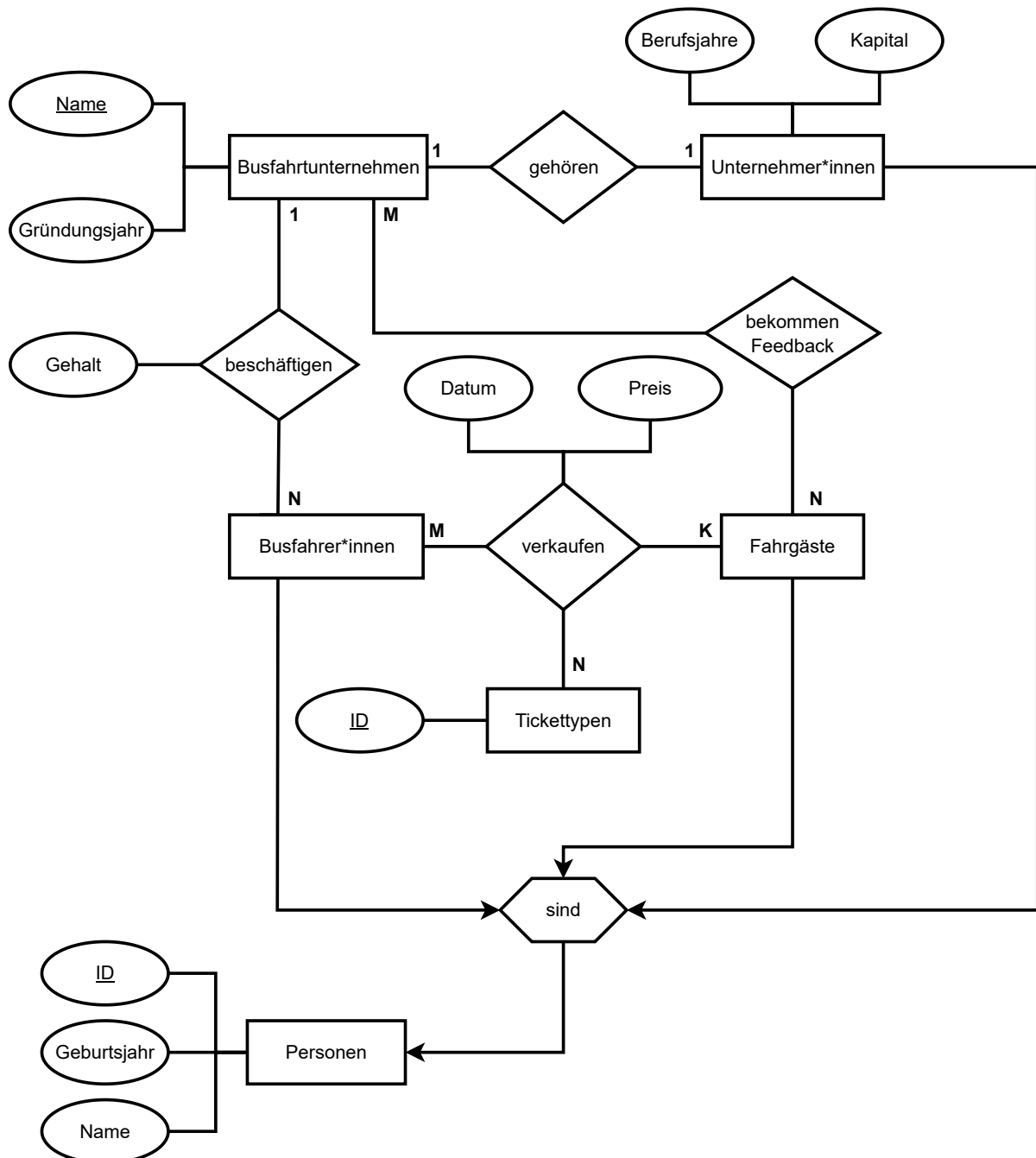


1 Überführung in das Relationale Modell (7 Punkte)

Überführen Sie das folgende ER-Diagramm in ein Relationales Modell. Vereinfachen Sie das Relationenschema so weit wie möglich und nutzen Sie sinnvolle Typen. Verwenden Sie die Notation aus der Vorlesung.



Lösung:

Bei einer Lösung mit Vererbung sind 8 Relationen zu modellieren. Angedacht sind hierbei 0,5 Punkte für die Relationen 'Personen', 'Tickets' und 'Fahrgäste'. Bei den Relationen 'Busfahrtunternehmen' und 'Unternehmer*innen' wird die Relation, bei der man sich für den Fremdschlüssel entschieden hat mit einem Punkt bewertet, die andere mit 0,5. 'Busfahrer*innen', sowie 'bekommen Feedback' werden mit jeweils einem Punkt bewertet und 'verkaufen' mit 2 Punkten. Es gibt jeweils 0,5 Punkte Abzug für fehlende Attribute, falsche Schlüsselattribute, falsche Fremdschlüssel, unsinnige Typen, usw.). Ob beispielsweise string oder varchar als Typ ausgewählt wird, spielt bei der Punktevergabe allerdings keine Rolle.

```
[Personen] : {[ID:int, Name:varchar, Geburtsjahr:int]}
[Tickettypen] : {[ID:int]}
[Busfahrer*innen] : {[Person_ID:(Personen→ID),
                    Busfahrtunternehmen_Name:(Busfahrtunternehmen→Name), Gehalt:int]}
[Fahrgäste] : {[Person_ID:(Personen→ID)]}
[Busfahrtunternehmen] : {[Name:varchar, Gründungsjahr: int]}
[Unternehmer*innen] : {[Person_ID:(Personen→ID),
                    Busfahrtunternehmen_Name:(Busfahrtunternehmen→Name),
                    Berufsjahre: int, Kapital: int]}
[bekommen Feedback] : {[Fahrgast_ID:(Fahrgäste→Person_ID),
                    Busfahrtunternehmen_Name:(Busfahrtunternehmen→Name)]}
[verkaufen] : {[Tickettyp_ID:(Tickettypen→ID),
                    Busfahrer*in_ID:(Busfahrer*innen→Person_ID),
                    Fahrgast_ID:(Fahrgäste→Person_ID),Datum:date, Preis:int]}
```

oder alternativ:

```
[Personen] : {[ID:int, Name:varchar, Geburtsjahr:int]}
[Tickettypen] : {[ID:int]}
[Busfahrer*innen] : {[Person_ID:(Personen→ID),
                    Busfahrtunternehmen_Name:(Busfahrtunternehmen→Name), Gehalt:int]}
[Fahrgäste] : {[Person_ID:(Personen→ID)]}
[Busfahrtunternehmen] : {[Name:varchar,
                    Unternehmer*in_ID:(Unternehmer*innen→Person_ID), Gründungsjahr: int]}
[Unternehmer*innen] : {[Person_ID:(Personen→ID), Berufsjahre: int, Kapital: int]}
[bekommen Feedback] : {[Fahrgast_ID:(Fahrgäste→Person_ID),
                    Busfahrtunternehmen_Name:(Busfahrtunternehmen→Name)]}
[verkaufen] : {[Tickettyp_ID:(Tickettypen→ID),
                    Busfahrer*in_ID:(Busfahrer*innen→Person_ID),
                    Fahrgast_ID:(Fahrgäste→Person_ID),Datum:date, Preis:double]}
```

2 Übersetzung in relationale Algebra (8 Punkte)

Gegeben sei folgendes Relationenschema.

[Personen] : {[PID: int, Name: string, Geburtsjahr: int, Wohnort: string]}
 [Schüler*innen] : {[SID:(Personen→PID), Klassenstufe: int, Klassenraum: int, Schulform: string]}
 [Klausuren] : {[KID: int, Fach: string, Thema: string, Dauer: int]}
 [Lehrer*innen] : {[LID:(Personen→PID), Hauptfach:string, Gehalt:int, Dienstjahre:int]}
 [unterrichten] : {[Schüler*in:(Schüler*innen→SID), Datum:date, Uhrzeit:time,
 Lehrer*in:(Lehrer*innen→LID), Fach: string]}
 [korrigieren] : {[Klausur:(Klausuren→KID), Schüler*in:(Schüler*innen→SID),
 Lehrer*in:(Lehrer*innen→LID), Note: int]}

Um die Lesbarkeit Ihrer Ausdrücke zu verbessern, dürfen Sie Teilergebnisse mit Namen versehen. Benennen Sie zum Beispiel einen Ausdruck wie folgt

$$R_3 := R_1 \bowtie_{A=B} R_2$$

so kann R_3 nun in darauffolgenden Ausdrücken verwendet werden. Zudem dürfen Sie folgenden syntaktischen Zucker benutzen, um mehrere Attribute gleichzeitig umzubenennen

$$\rho_{A' \leftarrow A, \dots, Z' \leftarrow Z} R = \rho_{A' \leftarrow A} (\dots (\rho_{Z' \leftarrow Z} R))$$

Übersetzen Sie die folgenden umgangssprachliche Anfragen in Ausdrücke der relationalen Algebra:

1. Alle Informationen über die Personen, die aus Neunkirchen kommen und vor 2009 geboren wurden.
2. Die Klassenstufe der Schüler*innen, die den Klassenraum 105 haben.
3. Die SID der Schüler*innen, die die 10. Klassenstufe besuchen und schon einmal im Fach 'Biologie' unterrichtet wurden.
4. Der Name und der Wohnort der Lehrer*innen mit einem Gehalt von über 3000 €.
5. Die Klassenstufe der Schüler*innen, die am 26.10.2021 von einer/einem Lehrer*in unterrichtet wurden, die/der bereits mehr als 10 Dienstjahre hat.
6. Der Wohnort der Lehrer*innen, die Mathematik als Hauptfach haben oder es schon einmal unterrichtet haben.
7. Die Dauer der Klausuren, in denen ein*e Schüler*in eines Gymnasiums die Note 2 bekommen hat, der/die schon einmal im Fach der Klausur unterrichtet wurde.
8. Das Thema der Klausuren, die nicht von einem/einer Lehrer*in korrigiert wurden, der/die das Fach der Klausur als Hauptfach hat.

Lösung:

Vorgesehen ist ein Punkt pro Anfrage (je 0,5 Punkte Abzug für inkorrekte (Join-)Prädikate oder Aggregate, unterschiedliche Schemata bei Mengenoperationen, nicht-disjunkte Schemata bei Joins, uneindeutige Bezeichner, inkorrekte Bezeichner durch Verwechslung von Umbenennung und Teilergebnisnotation, usw.). Wichtig ist hierbei nur, dass die Anfragen das korrekte Ergebnis produzieren, nicht, dass sie mit der Musterlösung übereinstimmen.

1. $\sigma_{\text{Wohnort} = \text{'Neunkirchen'} \wedge \text{Geburtsjahr} < 2009} \text{Personen}$
2. $\pi_{\text{Klassenstufe}}(\sigma_{\text{Klassenraum} = 105} \text{Schüler*innen})$
3. $\pi_{\text{SID}}(\sigma_{\text{Klassenstufe} = 10} \text{Schüler*innen}) \cap \pi_{\text{SID}}(\rho_{\text{SID} \leftarrow \text{Schüler*in}}(\sigma_{\text{Fach} = \text{'Biologie'}} \text{unterrichten}))$
4. $\pi_{\text{Name}, \text{Wohnort}}(\text{Personen} \bowtie_{\text{PID} = \text{LID}} (\sigma_{\text{Gehalt} > 3000} \text{Lehrer*innen}))$
5. $R := (\sigma_{\text{Dienstjahre} > 10}(\text{Lehrer*innen})) \bowtie_{\text{LID} = \text{Lehrer*in}} (\sigma_{\text{Datum} = 26.10.2021} \text{unterrichten})$
 $\pi_{\text{Klassenstufe}}(\text{Schüler*innen} \bowtie_{\text{SID} = \text{Schüler*in}} R)$
6. $R_1 := \pi_{\text{LID}}(\sigma_{\text{Hauptfach} = \text{'Mathematik'}} \text{Lehrer*innen})$
 $R_2 := \rho_{\text{LID} \leftarrow \text{Lehrer*in}}(\pi_{\text{Lehrer*in}}(\sigma_{\text{Fach} = \text{'Mathematik'}} \text{unterrichten}))$
 $\pi_{\text{Wohnort}}((R_1 \cup R_2) \bowtie_{\text{LID} = \text{PID}} \text{Personen})$
7. $R_1 := \rho_{\text{Schüler*in} \leftarrow \text{Schüler*in}, \text{Lehrer*in} \leftarrow \text{Lehrer*in}, \text{Fach} \leftarrow \text{Fach}} \text{unterrichten}$
 $R_2 := (\sigma_{\text{Note} = 2} \text{korrigieren}) \bowtie_{\text{Schüler*in} = \text{SID}} (\sigma_{\text{Schulform} = \text{'Gymnasium'}} \text{Schüler*innen})$
 $\pi_{\text{Dauer}}(R_1 \bowtie_{\text{Schüler*in} = \text{Schüler*in} \wedge \text{Fach} = \text{Fach}} (\text{Klausuren} \bowtie_{\text{KID} = \text{Klausur}} R_2))$
8. $R_1 := \text{Klausuren} \bowtie_{\text{KID} = \text{Klausur}} \text{korrigieren}$
 $R_2 := R_1 \bowtie_{\text{Lehrer*in} = \text{LID} \wedge \text{Fach} = \text{Hauptfach}} \text{Lehrer*innen}$
 $\pi_{\text{Thema}}((\pi_{\text{KID}, \text{Thema}} R_1 - \pi_{\text{KID}, \text{Thema}} R_2)$

3 Implementierung KeyRelation (5 Punkte)

Implementieren Sie die fehlenden Teile in der mit *Exercise* markierten Zelle im beigefügten Jupyter Notebook assignment02.ipynb.

Die Klasse `KeyRelation` soll dabei die Klasse `Relation` so erweitern, dass Schlüsselattribute spezifiziert werden können und beim Einfügen von Tupeln berücksichtigt werden. Zudem werden sogenannte UNIQUE-Attribute eingeführt, deren Werte unabhängig von den Schlüsselattributen ebenfalls nur einmal in einer Tabelle vorhanden sein dürfen. Verwenden Sie dazu geeignete Datenstrukturen, die Ihnen das Prüfen von Duplikaten in *konstanter Laufzeit* $\mathcal{O}(1)$ ermöglichen. Für einen einfachen Einstieg haben wir Ihnen bereits einen Unit Test mit mehreren Testfällen im Notebook zur Verfügung gestellt. Ihre finale Abgabe muss auf allen gültigen Eingaben das korrekte Ergebnis berechnen. Die Qualität Ihrer Implementierung wird bei der Punktevergabe berücksichtigt.

Abgabe

Lösungen sind in Teams von 2 bis 3 Studierenden bis zum 5. Mai 2022, 10:15 Uhr über Ihre persönlichen Statusseite im CMS einzureichen. Bitte notieren Sie Namen und Matrikelnummer aller Teammitglieder auf der Abgabe. Pro Team darf nur eine Abgabe hochgeladen werden.

Ihre Abgabe muss dem folgenden Format entsprechen:

```
abgabe.zip
├── abgabe.pdf
└── jupyter.txt
```

Hierbei enthält `abgabe.pdf` Ihre Lösungen zu Aufgaben 1 und 2 und `jupyter.txt` Ihre Lösung zu Aufgabe 3. Achten Sie darauf, dass Sie nur die von Ihnen zu ergänzenden Jupyter Zellen so kopieren, dass Einrückung und Formatierung korrekt sind.

Abgaben, die nicht den oben angegebenen Vorgaben entsprechen, führen zu Punktabzug.