



Übung zur Vorlesung *Grundlagen: Datenbanken* im WS22/23
Michael Jungmair, Stefan Lehner, Moritz Sichert, Lukas Vogel (gdb@in.tum.de)
<https://db.in.tum.de/teaching/ws2223/grundlagen/>

Blatt Nr. 02

Die Tutorübungen am Dienstag 01.11.2022 entfallen aufgrund des Feiertags.

Hausaufgabe 1

Lesen Sie zunächst das im Folgenden beschriebene Szenario eines Onlineshops:

In der Datenbank sollen die Kunden des Unternehmens mit ihrem Namen, ihrer Adresse und dem zur Adresse zugehörigen Land, sowie dem Geburtsdatum und dem Geschlecht erfasst werden.

Kunden können Bestellungen tätigen, wobei für die Bestellungen der Gesamtstatus (z.B. „Auftrag erteilt“, „In Bearbeitung“, „Abgeschlossen“), das Auftragsdatum sowie der Gesamtbetrag gespeichert werden müssen.

Bestellungen können mehrere Artikel umfassen, wobei für jeden Artikel der aktuelle Status (z.B. „In Auslieferung“, „Geliefert“, „Nicht verfügbar“), der Artikelpreis inkl. Steuern, die Steuern, die Artikelmenge (also wie oft der bestellte Artikel geliefert werden soll), sowie die Position innerhalb der Bestellung gespeichert werden müssen. Außerdem muss für jeden bestellten Artikel gespeichert werden, von welchem Lieferanten er stammt.

Für die Bereitstellung von Artikeln sind Lieferanten zuständig, wobei derselbe Artikel von unterschiedlichen Lieferanten zu unterschiedlichen Lieferpreisen angeboten werden kann. Lieferanten haben begrenzte Kapazitäten und einen Artikel nur begrenzt vorrätig.

Die Bereitstellungen der Artikel durch Lieferanten sind unabhängig davon zu erfassen, ob sie jemals bestellt worden sind. Es kann auch vorkommen, dass zum Zeitpunkt der Artikelbestellung noch kein Lieferant feststeht und erst später in der Datenbank nachgetragen werden soll.

Jeder Artikel besitzt eine unverbindliche Preisempfehlung, die unabhängig von den Lieferanten ist. Lieferanten sollen mit Namen und Adresse sowie dem der Adresse zugehörigen Land in der Datenbank gespeichert werden. Es soll auch erfasst werden, welche Länder von einem Lieferanten beliefert werden.

Der Onlineshop stellt außerdem die Anforderung an das neue Datenbankschema, dass wichtige Analysen getrennt nach Kontinent durchgeführt werden können, damit beispielsweise die Absatzzahlen eines Produktes in Europa mit denen in Amerika verglichen werden können.

Entwerfen Sie ein ER-Modell für dieses System. Sie müssen dabei nur die in der Aufgabe beschriebenen Sachverhalte berücksichtigen. Gehen Sie dazu wie folgt vor:

- a) Identifizieren Sie alle relevanten Entity-Typen.

- b) Identifizieren Sie alle notwendigen Relationships zwischen den Entity-Typen. Achten Sie darauf, dass die im Szenario beschriebenen Sachverhalte eingehalten werden.
- c) Ergänzen Sie die Entitäten (und wenn nötig auch die Relationships) um Attribute. Führen Sie, falls nötig, neue Schlüsselattribute ein, die zur eindeutigen Identifizierung der Entitäten erforderlich sind.
- d) Ergänzen Sie die Relationships mit Funktionalitätsangaben. Fügen Sie auch (min,max)-Angaben hinzu.
- e) Welche im Szenario beschriebene Sachverhalte können in der Modellierung nicht berücksichtigt werden?

Scannen Sie ihre Lösung und schicken Sie sie bereits vor der Übung an Ihren Tutor, damit während der Übung mehrere Lösungsansätze miteinander verglichen werden können und nicht jeder einzelne Lösungsvorschlag separat an das Whiteboard gezeichnet werden muss.

Lösung: Für die Teilaufgaben a) bis d) siehe Abbildung 1, Abbildung 2 und Abbildung 3.

- a) In der Beispiellösung werden auch *Länder* und *Kontinente* als eigenständige Entity-Typen erfasst. Das zugehörige Land könnte für *Kunden* und *Lieferanten* auch als Attribut gespeichert werden. Jedoch hätte diese Designentscheidung den Nachteil, dass die (mitunter längeren) Namen der Länder mehrmals persistiert werden müssten. Außerdem existieren *Länder* unabhängig davon, ob der Shop *Kunden* oder *Lieferanten* hat, die aus diesen entstammen. Falls sich der Shop später entschließen sollte, zusätzliche Informationen zu Ländern wie beispielsweise unterschiedliche Steuersätze oder Zeitzonen zu speichern, ist dieses Design ebenfalls schwieriger zu erweitern.

In der Beispiellösung werden die einzelnen *Artikel* in einer *Bestellung* als separater Entity-Typ *Bestellpositionen* erfasst. Dieser ist notwendig, da *Bestellungen* beliebig viele Positionen enthalten können. Bestellpositionen selbst haben hier *keinen* Primärschlüssel - nur die Position in Verbindung mit der Bestellnummer kann eine Entität eindeutig identifizieren.

- b) Die in der Beispiellösung angegebene ternäre Relationship *liefern* speichert zu jeder Bestellposition, welchen Artikel sie betrifft und von welchem Lieferanten sie geliefert wird. Da jedoch im Szenario verlangt wird, dass bereitgestellte Artikel unabhängig davon gespeichert werden sollen, ob sie überhaupt jemals bestellt worden sind, benötigen wir eine weitere Relationship *bereitstellen*. In dieser Relationship speichern wir zu jedem von Lieferanten bereitgestellten Artikel zusätzlich den Lieferpreis und die verfügbare Anzahl, da diese Attribute abhängig vom Lieferanten sind.
- c) Siehe Abbildung 2.
- d) Siehe Abbildung 3.
- e) Einige mögliche Antworten:
 - Die Tatsache, dass Lieferanten zum Zeitpunkt der Bestellung noch nicht feststehen müssen, lässt sich weder durch Funktionalitäten noch durch (min, max)-Angaben in der Beispiellösung ausdrücken.
 - Die Tatsache, dass eine Bestellposition nur genau einen Artikel und einen Lieferanten enthalten darf. In unserem ER-Diagramm könnte laut Funktionalitätsangaben eine Bestellposition mehrere Artikel und Lieferanten enthalten.

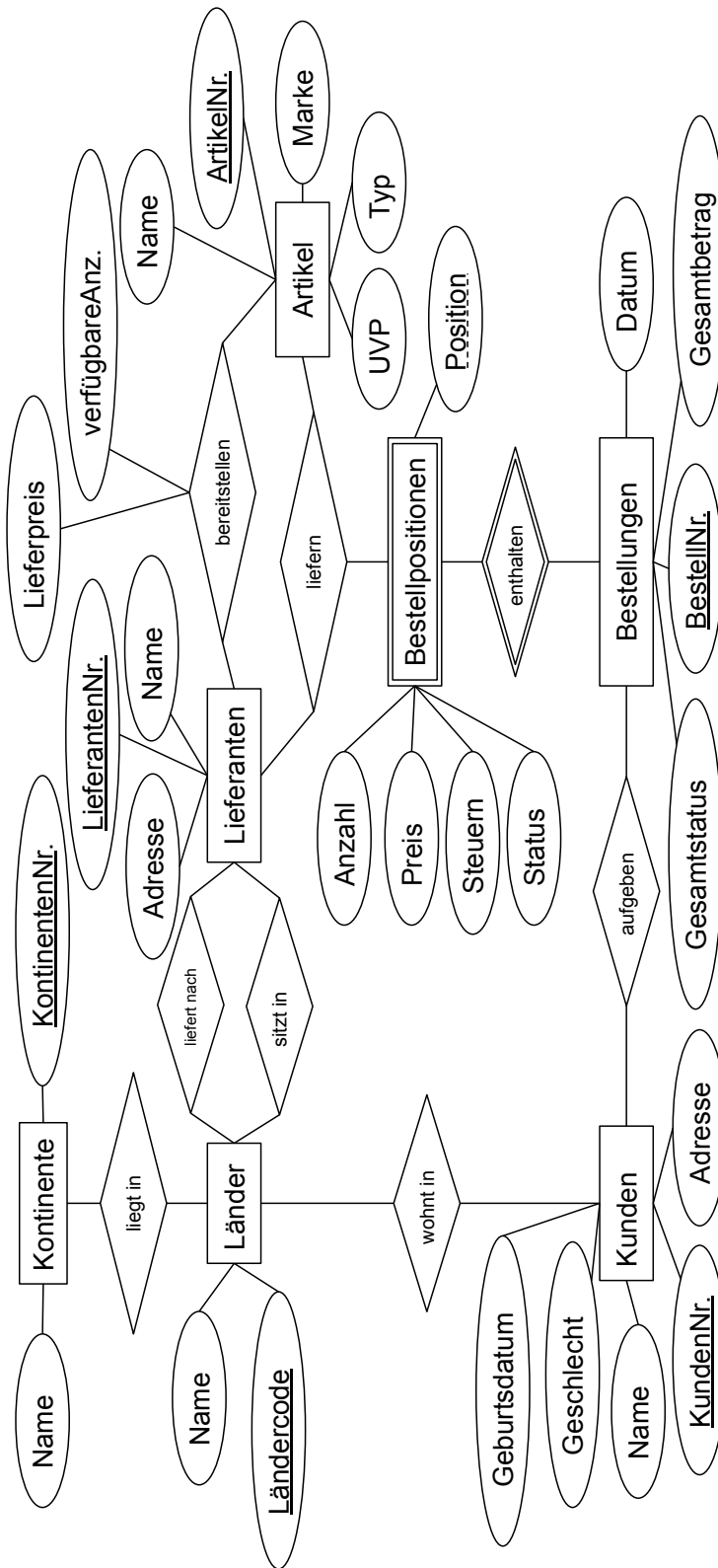


Abbildung 1: ER-Diagramm des Onlineshops

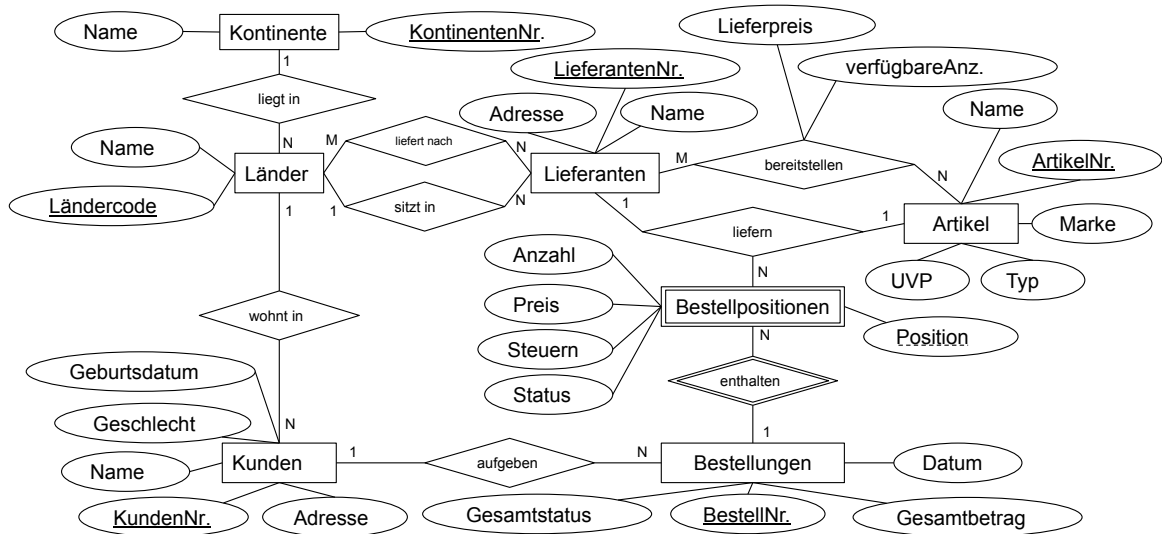


Abbildung 2: ER-Diagramm des Onlineshops mit Funktionalitätsangaben

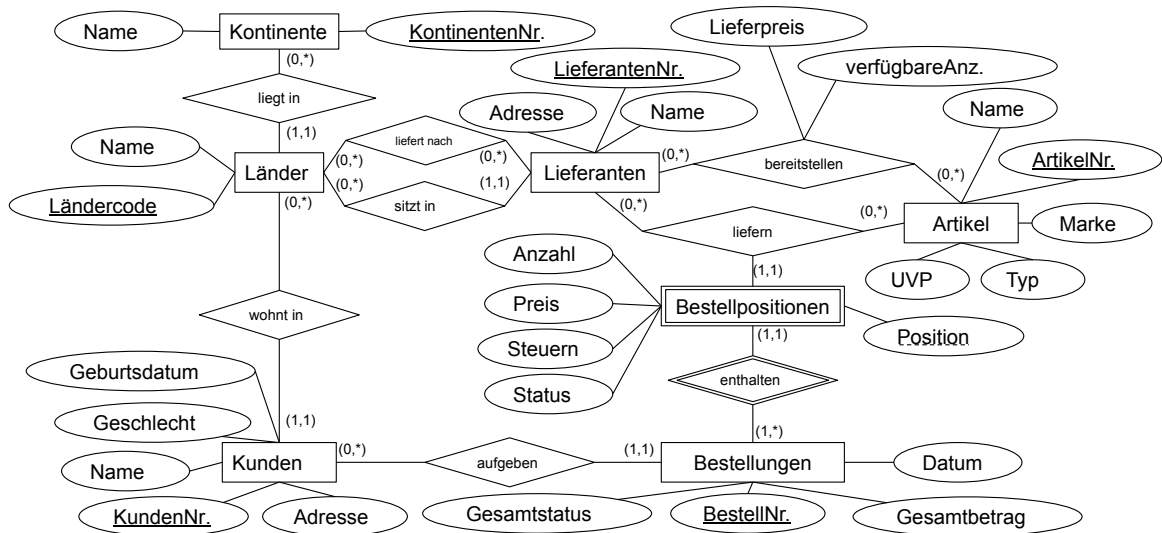
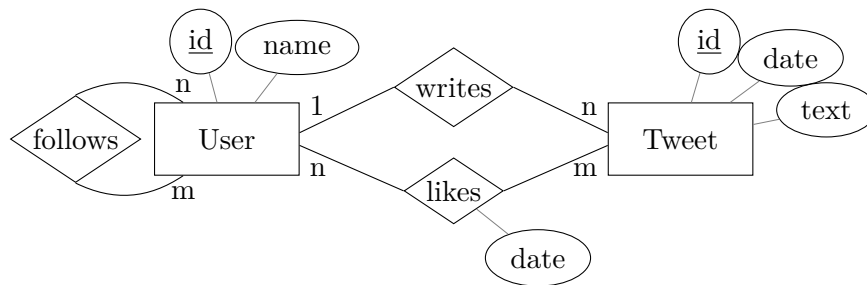


Abbildung 3: ER-Diagramm des Onlineshops mit (min, max)-Angaben

Hausaufgabe 2

Gegeben sei folgendes ER-Diagramm, das User, deren Tweets, Likes und Follows modelliert:



- Übertragen Sie das ER-Modell in ein relationales Schema.
- Verfeinern Sie das relationale Schema durch Elimination von Relationen.
- Kann das Attribut *date* des Entity-Typs *Tweet* stattdessen der Relationship *writes* zugeordnet werden? Kann das Attribut *date* der Relationship *likes* dem Entity-Typen *Tweet* zugeordnet werden?

Lösung:

a)

User : { [id,name] }
Tweet : { [id,date, text] }
follows : { [follower_id,follows_id] }
writes : { [tweet_id,user_id] }
likes : { [user_id,tweet_id,date] }

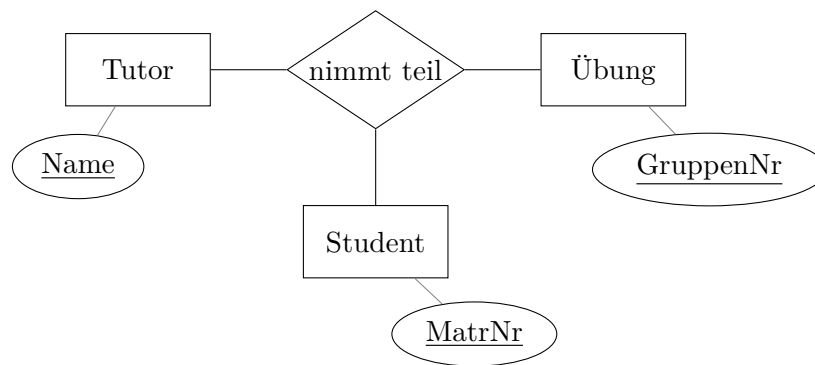
- b) Die Relationen *Tweet* und *writes* können zusammengefasst werden, da sie den gleichen Schlüssel haben und die Relationship eine 1-zu-n-Beziehung ist. Eine weitere Verfeinerung ist nicht möglich.

User : { [id,name] }
Tweet : { [id,user_id, date, text] }
follows : { [follower_id,follows_id] }
likes : { [user_id,tweet_id,date] }

- c) Das Attribut *date* kann auch *writes* zugeordnet werden. Insbesondere nach der Verfeinerung sind alle Attribute von *writes* und von *Tweet* in einer Relation zusammengefasst. Für das relationale Schema ist es also egal, ob das Attribut ursprünglich zu *writes* oder zu *Tweet* gehört hat.

Das Attribut *date* der Relation *likes* kann keinem anderen Entity-Typen zugeordnet werden, da weder *User* noch *Tweet* wegen ihrer n-zu-m-Beziehung ein eindeutiges „like_date“-Attribut haben können.

Hausaufgabe 3



Angenommen, das hier modellierte Übungssystem entspricht dem Übungssystem in Grundlagen: Datenbanken. Bestimmen Sie die MinMax Angaben so, dass folgende Einschränkungen modelliert werden:

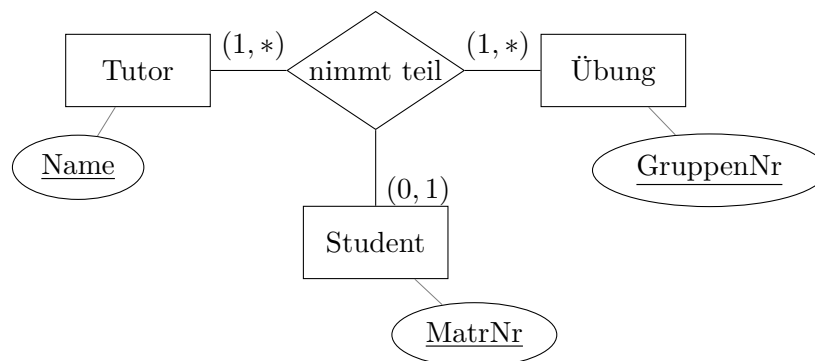
- Ein Tutor hält mindestens eine Übung.
- Eine Übung wird von mindestens einem Studenten besucht.
- Ein Student kann höchstens eine Übung besuchen.

Betrachten Sie nun die folgende Ausprägung, die die Beziehung modellieren soll:

| Name | GruppenNr | MatrNr |
|---------|-----------|--------|
| ⋮ | ⋮ | ⋮ |
| Lang | G12 | 23 |
| Passing | G27 | 42 |
| Passing | G27 | 43 |
| ⋮ | ⋮ | ⋮ |
| Passing | G28 | 97 |
| Passing | G28 | 98 |
| Passing | G28 | 99 |
| ⋮ | ⋮ | ⋮ |

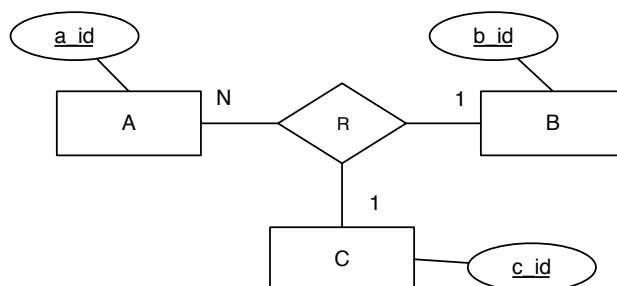
Welche Beziehung besteht zwischen der MinMax Notation und einer solchen Ausprägung?

Lösung:



Im Bezug auf die tabellarische Repräsentation der Beziehung sagen die MinMax Angaben gerade aus, wie oft ein konkreter Wert, etwa der Name des Tutors *Lang*, minimal und maximal vorkommen darf. Die MinMax-Angaben lassen sich also leicht herleiten, indem man sich die tabellarische Repräsentation einer Beziehung vorstellt und überlegt, welche Einschränkungen für die Wiederholung eines konkreten Wertes in dieser Repräsentation gelten sollen. Dadurch, dass im Beispiel keine Matrikelnummer mehrfach vorkommt wird etwa die Einschränkung modelliert, dass ein Student nicht mehrere Übungen besuchen darf.

Hausaufgabe 4



- Welche partiellen Funktionen gelten?
- Setzen Sie das ER Modell in Relationen um.
- Bestimmen Sie einen Schlüssel für die Beziehung R, so dass möglichst viele Einschränkungen aus dem ER Modell auch in der Relation für die Beziehung modelliert werden.
- Wieso ist ein Semantikverlust zunächst unvermeidbar? Welche Einschränkung müsste der Relation hinzugefügt werden, um die volle Semantik des ER Modells zu modellieren? ¹

Lösung:

Es gelten die partiellen Funktionen

$$A \times C \rightarrow B \quad (1)$$

$$A \times B \rightarrow C. \quad (2)$$

Aus dem Modell entstehen die folgenden Relationen:

$$A \quad : \quad \{\{a_id\}\} \quad (3)$$

$$B \quad : \quad \{\{b_id\}\} \quad (4)$$

$$C \quad : \quad \{\{c_id\}\} \quad (5)$$

sowie entweder

$$R \quad : \quad \{\{a_id, b_id, c_id\}\} \quad (6)$$

oder

$$R \quad : \quad \{\{a_id, b_id, c_id\}\} \quad (7)$$

¹Diese Teilaufgabe ist keine Hausaufgabe. Diskutieren Sie dies in der Übung!

Egal ob zur Modellierung der Beziehung die Variante in Gleichung 6 oder 7 verwendet wird, kann hier nur die Einschränkung einer der zwei im ER Modell geltenden partiellen Funktionen in die Modellierung mittels Relationen übertragen werden. Eine andere Wahl des Schlüssel außer den hier gezeigten ist entweder falsch, da sie zu starke Anforderungen stellt (etwa wenn lediglich ein Attribut als Schlüssel markiert wurde) oder semantisch schwächer (etwa die Markierung aller Attribute als Schlüssel).

Der Semantikverlust ist zunächst unvermeidbar, da lediglich eine Kombination von Attributen als Primärschlüssel der Beziehung markiert werden kann. Jede Kombination dieser Attribute darf daher nur einmal in der Ausprägung der Beziehungsrelation auftreten, was es erlaubt, die Einschränkung einer der zwei geltenden partiellen Funktionen in die Modellierung als Relation zu übertragen. Die andere geltende partielle Funktion kann nicht in die Relationenmodellierung übernommen werden. Hierfür wäre es notwendig, für eine weitere Menge von Attributen die Einschränkung modellieren zu können, dass jede Kombination dieser Attribute nur einmal in der Ausprägung der Relation auftreten kann. Dies ist beispielsweise in SQL mittels dem `unique`-Schlüsselwort möglich.