

Verwaltung von Daten – Teil2

Entwurf eines relationalen Schemas

Redundanzen und Anomalien

Relationale Datenbanken speichern Daten in verschiedenen Tabellen. Im Idealfall werden in einer Tabelle die Daten zusammengefasst, die sich einer Entität oder Beziehung des entsprechenden Ausschnitts der realen Welt zuordnen lassen. In Teil 1 haben Sie gelernt, Daten mithilfe von SQL-Anfragen aus der Datenbank auszulesen. Dabei lassen sich Daten unterschiedlicher Tabellen mithilfe von Primär- und Fremdschlüssel nach Bedarf verknüpfen. Die Aufteilung in verschiedene Tabellen stellt also keinen Nachteil bezüglich der Informationsgewinnung dar. Vielmehr hat sie den Vorteil, dass Daten nicht redundant gespeichert werden müssen und flexibel miteinander in Zusammenhang gebracht werden können. In den Eingangsbetrachtungen von Teil 1 wurde aber auch bereits deutlich, dass die Einteilung der Entitäten und Beziehungen eine Modellierungsentscheidung und keineswegs eindeutig ist. Eine ungünstige Modellierung kann daher zu Tabellen führen, die redundante Daten enthalten oder in Bezug auf die Operationen Einfügen, Ändern und Löschen problematisch sind. Vielleicht werden auch nachträglich Attribute in den Tabellen ergänzt, die zu entsprechenden Schwierigkeiten führen, oder es steht eine Tabellenstruktur zur Verfügung, die noch nicht für eine relationale Datenbank optimiert wurde.

Im Folgenden sollen die Probleme, die beim Einfügen, Ändern und Löschen von Daten in Tabellen auftreten können, näher untersucht werden. Dadurch erhält man Hinweise, worauf beim Entwurf der Tabellen zu achten ist.

Im ER-Diagramm des Sportvereins TSV Fitstadt war vorgesehen, die Sportstätte als Attribut der Entität Kurse zu modellieren. Wenn wir nun noch weitere Daten zu der Sportstätte, wie die Hallengröße und die Kontaktdaten des Ansprechpartners¹ für die Reservierung von Hallenzeiten speichern möchten, ergibt sich das folgende Tabellenschema für die Tabelle Kurse:

KursID	Sportart	Sportstaette	Hallen-groesse	Ansprech-partner	Telefon	Wochen-tag	Uhrzeit	Dauer	Trainer
...
TWJ	Turnen	Sporthalle IGS	300	Herr Freund	0151-73451239	Dienstag	16:00	60	26
HHE	Handball	Sporthalle 2	200	Frau Holle	0162-192837465	Mittwoch	17:30	90	25
HHJ	Handball	Sporthalle 2	200	Frau Holle	0162-192837465	Donners-tag	17:00	60	25
...

Abbildung 1: Erweiterte Tabelle Kurse

Diese Tabelle enthält redundante Daten, da *Hallengroesse*, *Ansprechpartner* und zugehörige *Telefonnummer* für jeden Kurs, der in einer Sportstätte stattfindet, neu eingetragen werden, aber die Daten immer gleich sind. **Redundante Daten** erhöhen den Speicherbedarf und können zu

¹ Zur besseren Lesbarkeit und Umsetzbarkeit im Zusammenhang mit Datenbanken wird das generische Maskulinum und exemplarisch die männliche Form verwendet. Diese Formulierung umfasst hier gleichermaßen und gleichberechtigt alle Geschlechter.

Änderungsanomalien führen. Ändert sich die Telefonnummer eines Ansprechpartners, müssen mehrere Datensätze geändert werden und es kann passieren, dass die Aktualisierung in einem Datensatz vergessen wird. Das würde zu einem inkonsistenten Datenbestand mit widersprüchlichen Daten führen. Man spricht in diesem Fall von einer **Änderungsanomalie**. Hat der Verein eine Kooperation für eine neue Sportstätte geschlossen, in der aber aktuell noch keine Kurse stattfinden, ergibt sich beim Einfügen der neuen Sportstätte in die Datenbank das Problem, dass dieser noch kein Wert für den Primärschlüssel zugeordnet werden kann. Folglich lässt sich die Sportstätte ohne zugehörigen Kurs gar nicht in den Datenbestand aufnehmen. In diesem Fall spricht man von einer **Einfügeanomalie**. Werden hingegen alle Kurse, die bislang in Sporthalle 2 stattgefunden haben, eingestellt und aus der Tabelle Kurse gelöscht, so gehen auch die Informationen zu Sporthalle 2, die Hallengröße und der Ansprechpartner, verloren. Diese Daten werden aber für die Einrichtung neuer Kurse in dieser Sporthalle 2 noch benötigt. Werden auf diese Weise Daten ungewollt gelöscht, spricht man von einer **Löschanomalie**. Im Folgenden werden die verschiedenen Anomalien, die auftreten noch einmal allgemein definiert:

Änderungsanomalie (Update-Anomalie): Wenn das Ändern eines Datensatzes aufgrund der redundanten Speicherung der Daten zwangsläufig das Ändern weiterer Datensätze nach sich zieht, spricht man von einer Änderungsanomalie. Das ist nicht nur aufwändig, sondern kann auch zu inkonsistenten Daten führen, wenn die Änderung an einer Stelle vergessen wird.

Einfügeanomalie: Eine Einfügeanomalie tritt auf, wenn eine Information nicht eingefügt werden kann, ohne weitere Informationen, die vielleicht noch gar nicht bekannt sind, einzufügen, da sonst ein Schlüsselattribut fehlt oder andere Zellen, die nicht leer bleiben dürfen, fehlen. Von einer Einfügeanomalie spricht man also, wenn durch das Einfügen von Daten ungewollt weitere Daten eingefügt werden müssen.

Löschanomalie: Eine Löschanomalie liegt vor, wenn durch das Löschen von Daten weitere Informationen verloren gehen, die aber gar nicht gelöscht werden sollen.

Aufgabe 1:

- a) Eine Musik-Band bietet Ihre Merchandise-Artikel über Ihre Webseite an. Die Tabelle, in der die eingehenden Bestellungen gespeichert werden, hat die Struktur in Abbildung 2. Untersuchen Sie die Tabelle auf Redundanzen und Anomalien.
- b) Erläutern Sie ob und ggf. unter welchen Voraussetzungen die Musik-Band die Daten in der Tabelle in Abbildung 2 erheben und speichern darf.

Auftragsnr	Artikel	Lagerbestand	Anzahl	Datum	Kleidergröße	Kundenr	Name	E-Mail	Adresse	Kontonr
23431	T-Shirt A	34	1	12.03.2021	XL	123456	Bruno Bär	xy@web.de	Große Str. 4, 32453 Fanstadt	DE12 2312 3232 3233 2323 21
23432	Tasse schwarz	16	2	12.03.2021	M	654321	Lisa Lustig	ab@gmx.de	Musterstr. 7, 23412 Fitstadt	DE42 2345 4367 9876 3452 23
23456	Button	60	5	30.04.2021	XL	123456	Bruno Bär	xy@web.de	Große Str. 4, 32453 Fanstadt	DE12 2312 3232 3233 2323 21
23437	Tasse schwarz	14	2	24.06.2021	M	654321	Lisa Lustig	ab@gmx.de	Musterst. 7, 23412 Fitstadt	DE42 2345 4367 9876 3452 23
23463	Button	52	8	15.07.2021	S	91827	Susi Pink	blub@web.de	Breite Str. 5, 78564 Neudorf	DE34 3124 3421 4323 5637 43
...

Abbildung 2 :Datenspeicherung eines Webshops

Aufgabe 2: Ein Zoo speichert eine Übersicht der Tiere mit ihrem Aufenthaltsort und den zuständigen Tierpfleger, die im Notfall zu kontaktieren sind, in einer Tabelle, die wie in Abbildung 3 aufgebaut ist. Untersuchen Sie die Tabelle auf Redundanzen und Anomalien.

Name	Tierart	Geburtsdatum	GehegeNr	Lage	Größe	Tierpfleger	Handynr	Festnetz
Paula	Giraffe	12.03.2005	8	Sambesi	2000	Piet Grimm	01253657894	0511123456
Strolch	Giraffe	16.04.2007	8	Sambesi	2000	Piet Grimm	01253657894	0511123456
Paula	Löwe	15.06.2015	10	Sambesi	1500	Gabi Pflug	01478956235	null
Robert	Löwe	12.08.2013	10	Sambesi	1500	Gabi Pflug	01478956235	null
Nala	Tiger	14.08.2009	16	Dschungel	1000	Gabi Pflug	01478956235	null
...

Abbildung 3: Datenspeicherung eines Zoos

Aufgabe 3: Ein Tierheim speichert in einer Tabelle mit der Struktur in Abbildung 4 Daten über seine aktuellen tierischen Bewohner. Untersuchen Sie die Tabelle auf Redundanzen und Anomalien.

Name	Geb-Datum	Geschlecht	Spezies	Tierart	Familie	Ordnung	Klasse
Luna	05.03.2020	w	Siamkatze	Hauskatze	Katzen	Raubtiere	Säugetiere
Pucki	06.05.2018	m	Perserkatze	Hauskatze	Katzen	Raubtiere	Säugetiere
Fipsi	05.02.2021	w	Hansi-Bubi	Wellensittich	Eigentliche Papageien	Papageien	Vögel
Frida	03.06.2020	w	Hansi-Bubi	Wellensittich	Eigentliche Papageien	Papageien	Vögel
Blacky	09.12.2019	m	Kongo-Graupapagei	Graupapagei	Eigentliche Papageien	Papageien	Vögel
...

Abbildung 4: Datenspeicherung eines Tierheims

Modellierung mithilfe von ER-Diagrammen

Redundanzen und Anomalien kommen häufig dadurch zustande, dass die Daten, die sich unterschiedlichen Entitäten zuordnen lassen, in einer gemeinsamen Tabelle gespeichert werden. Um dies zu vermeiden, ist eine gut durchdachte Modellierung in einem ER-Diagramm hilfreich, um im Vorfeld zu überlegen, welche Entitäten unterschieden werden müssen und in welchen Beziehungen diese zueinanderstehen. Treten dann noch Redundanzen und Anomalien in den Tabellen auf, können die Abhängigkeiten der Attribute genauer betrachtet werden, um die Tabellen ggf. weiter aufzuteilen (→ s. Optimierung einzelner Tabellen). Wir beginnen mit der Modellierung mithilfe eines ER-Diagramms.

Beim Entwurf einer Datenbank müssen die Daten, die einen Ausschnitt aus der realen Welt, z. B. eine Schule, einen Verein, einen Versandhandel oder eine Arztpraxis, beschreiben, strukturiert werden, indem der Ausschnitt der realen Welt in einem ER-Diagramm abgebildet wird. Dazu werden die Daten den verschiedenen Entitäten und ihren Beziehungen zugeordnet.

Manchmal ist es nicht ganz einfach, zwischen Attributen und Entitäten zu unterscheiden. Im ER-Diagramm des Sportvereins Fitstadt wird die Sportstätte beispielsweise als Attribut der Entität Kurs modelliert. Das ist in Ordnung, solange die Sportstätte mit einem einzelnen Wert, hier dem Namen, beschrieben werden kann. Möchte man hingegen weitere Informationen, die die Sportstätte näher beschreiben, wie die Adresse, die Größe oder die Anzahl der Sitzplätze auf der Tribüne, in der Datenbank speichern, ist es sinnvoller, die Sportstätte als eigene Entität zu modellieren. Andernfalls würden diese Daten für jeden Kurs, der in der Sportstätte stattfindet, redundant gespeichert. Eine entsprechende Modellierung zeigt Abbildung 5.

Durch Unterstreichen werden im ER-Diagramm bereits die Attribute, die den Primärschlüssel bilden, gekennzeichnet. Zur Erinnerung: Entitäten werden als Rechtecke dargestellt, Attribute in Ovalen und die Beziehungen in Rauten.

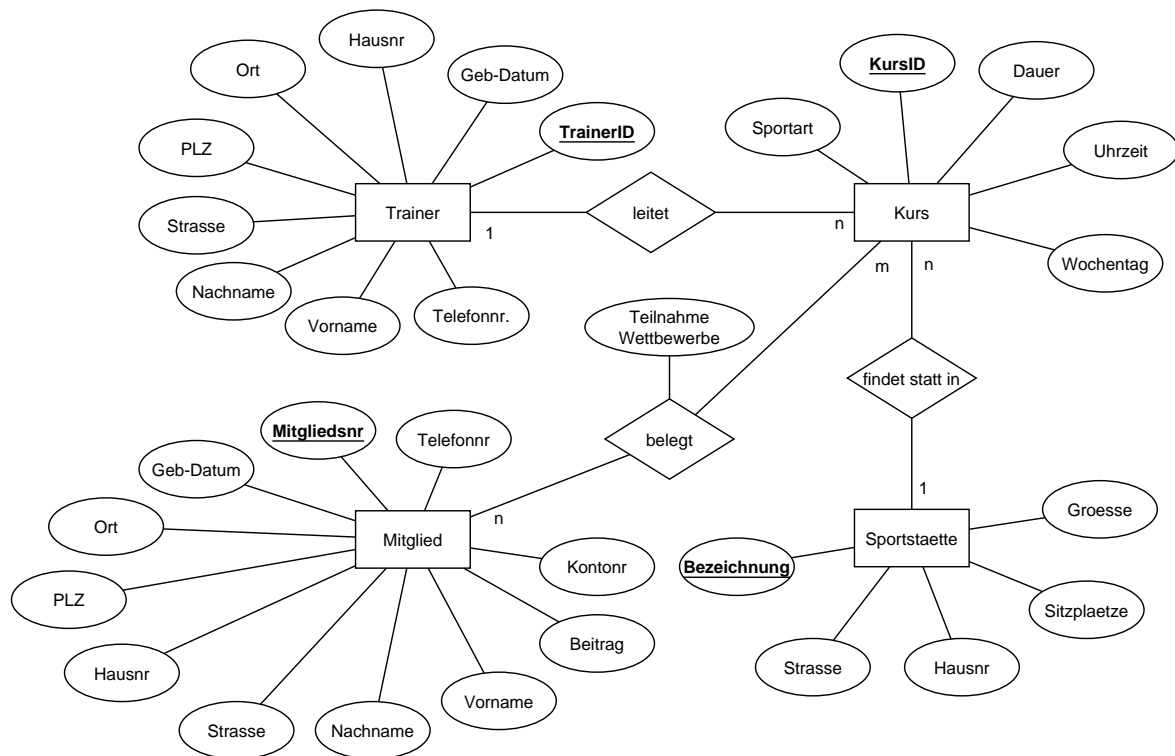


Abbildung 5: ER-Diagramm des TSV Fitstadt mit "Sportstaette" als eigener Entität

Kardinalitäten

Für die spätere Umsetzung des ER-Diagramms in Tabellen ist entscheidend, ob ein Element einer Entität eine Beziehung mit höchstens einem oder mit mehreren Elementen der anderen Entität eingehen kann. Man unterscheidet daher 1:1, 1:n und n:m-Beziehungen. Die Angabe 1 bzw. n wird als **Kardinalität** bezeichnet und an den Enden der Beziehung neben den Entitäten vermerkt. Im TSV Fitstadt gibt es pro Kurs nur einen Trainer. Deshalb steht an der Beziehung *leitet* auf der Seite der Entität *Trainer* eine 1. Ein Trainer kann aber beliebig viele Kurse leiten. Deshalb steht auf der Seite der Entität *Kurs* ein n. Ein Mitglied kann mehr als einen Kurs belegen und ein Kurs hat mehr als einen Teilnehmer, deshalb handelt es sich in diesem Fall um eine n:m-Beziehung. Sowohl bei 1:1, 1:n als auch bei n:m-Beziehungen ist es möglich, dass einige Elemente der Entität gar keine Beziehung eingehen. Das wäre z. B. der Fall, wenn ein Trainer aktuell keine Kurse leitet oder ein Mitglied keine Kurse belegt.

Aufgabe 4: Begründen Sie, dass es sich bei der Beziehung *findet statt in* zwischen Sportstätte und Kurs um eine 1:n Beziehung handelt.

Aufgabe 5:

- Ordnen Sie den schematischen Darstellungen in Abbildung 5 die Beziehungstypen 1:1, 1:n bzw. n:m zu.
- Geben Sie Beispiele für 1:1, 1:n bzw. n:m-Beziehungen an.

Elemente einer Entität können auch eine Beziehung mit den Elementen der gleichen Entität eingehen. Man spricht in diesem Fall von einer **rekursiven Beziehung**. Wenn im TSV Fitstadt für jeden Trainer ein anderer Trainer festgelegt wird, der im Krankheitsfall die Vertretung übernimmt, so kann dies im ER-Diagramm durch die rekursive Beziehung *vertritt* wie in Abbildung 7 dargestellt werden. Um die jeweilige Funktion der beiden an der Beziehung beteiligten Elemente der Entität zu verdeutlichen, wurden die beiden Enden der Beziehung hier genauer bezeichnet. Man spricht in diesem Fall von **Rollen**.

Aufgabe 6:

- Erläutern Sie, was die Kardinalitäten der Beziehung *vertritt* in Abbildung 7 über die Organisation der Vertretungen im TSV Fitstadt aussagt.
- Diskutieren Sie alternative Lösungen in der realen Welt und die entsprechende Modellierung im ER-Diagramm.

Aufgabe 7: Für die Schulbuchausleihe sollen folgende Informationen in einer Datenbank verwaltet werden: Es wird eine Übersicht benötigt, welche Schulbücher zur Ausleihe zur Verfügung stehen. Zu jedem **Schulbuch** werden **Titel**, **Verlag**, **ISBN**, **Fach** und **Jahrgang** vermerkt. Ein **Exemplar** eines Schulbuchs wird anhand der **Inventarnummer** identifiziert und zusätzlich der **Zustand** vermerkt. Zu jedem **Schüler** werden **Vorname**, **Nachname** und die **E-Mail-Adresse** gespeichert. Den Schülern werden außerdem ihre **Klassen zugeordnet**. Jede **Klasse** erhält neben der **Bezeichnung** und der **Raumnummer** eine **ID**. Wenn ein Schüler ein Exemplar eines Buches **ausleiht**, werden dazu das **Ausleihdatum** und das **Rückgabedatum** vermerkt.

Den Schülern werden außerdem ihre **Klassen zugeordnet**. Jede **Klasse** erhält neben der **Bezeichnung** und der **Raumnummer** eine **ID**. Wenn ein Schüler ein Exemplar eines Buches **ausleiht**, werden dazu das **Ausleihdatum** und das **Rückgabedatum** vermerkt.

- Modellieren Sie den Ausschnitt der Miniwelt „Schulbuchausleihe“ mithilfe eines ER-Diagramms.
- Erläutern Sie, welchen Vorteil es hat, der Klasse eine ID als Primärschlüssel zuzuordnen, anstatt die Bezeichnung als Schlüsselattribut zu verwenden.

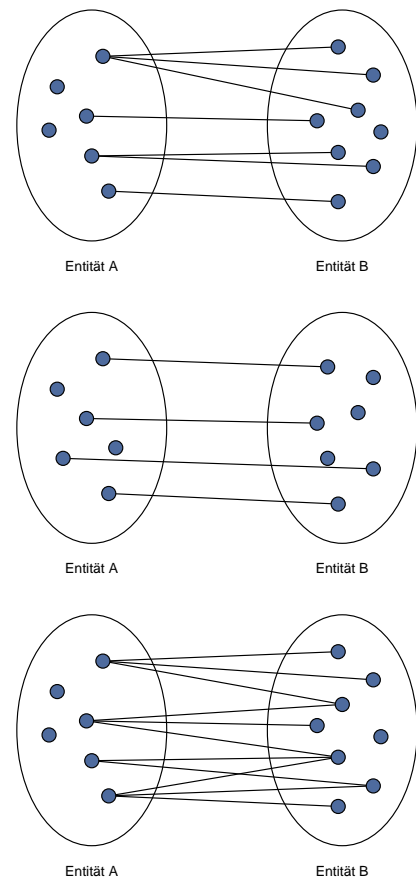


Abbildung 6: schematische Darstellung von Beziehungen mit unterschiedlichen Kardinalitäten

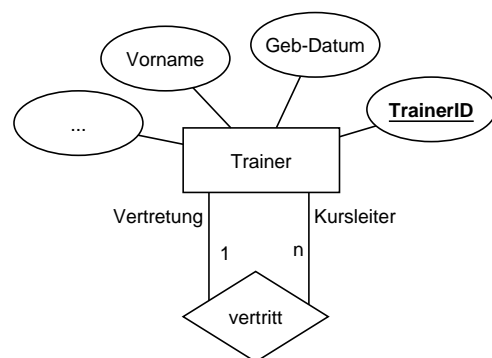


Abbildung 7: Beispiel für eine rekursive Beziehung

- c) Diskutieren Sie, unter welchen Voraussetzungen es sinnvoll ist, den Verlag als eigene Entität zu modellieren.
- d) Um die Schüler über die Klassenlehrer zu erreichen, soll diese Zuordnung ebenfalls in der Datenbank enthalten sein. Diskutieren Sie, wie die Klassenlehrer sinnvoll in dem ER-Diagramm ergänzt werden können.
- e) Die Schüler müssen pro Schuljahr eine Leihgebühr bezahlen. Die jeweilige Höhe der Leihgebühr und ob diese bereits bezahlt wurde, soll ebenfalls in der Datenbank gespeichert werden. Ergänzen Sie das ER-Diagramm entsprechend.
- f) Schüler, deren Geschwister ebenfalls an der Schulbuchausleihe teilnehmen, erhalten eine Ermäßigung der Leihgebühr. Deshalb soll in der Datenbank auch aufgenommen werden, welche Schüler Geschwister sind.
- g) Vergleichen Sie das ER-Diagramm mit der Schulbauchausleihe an Ihrer Schule. Müssen weitere Informationen ergänzt werden oder werden einige gar nicht benötigt?

Aufgabe 8: Ein Lehrerkollegium möchte die Klausuren der Oberstufe künftig mithilfe eines Programms koordinieren, das die entsprechenden Daten in einer Datenbank speichert. Der Informatikkurs entwirft dafür das ER-Diagramm in Abbildung 8.

- a) Beurteilen Sie das ER-Diagramm hinsichtlich der Modellierung und optimieren Sie es ggf.
- b) Beurteilen Sie das ER-Diagramm unter datenschutzrechtlichen Gesichtspunkten.

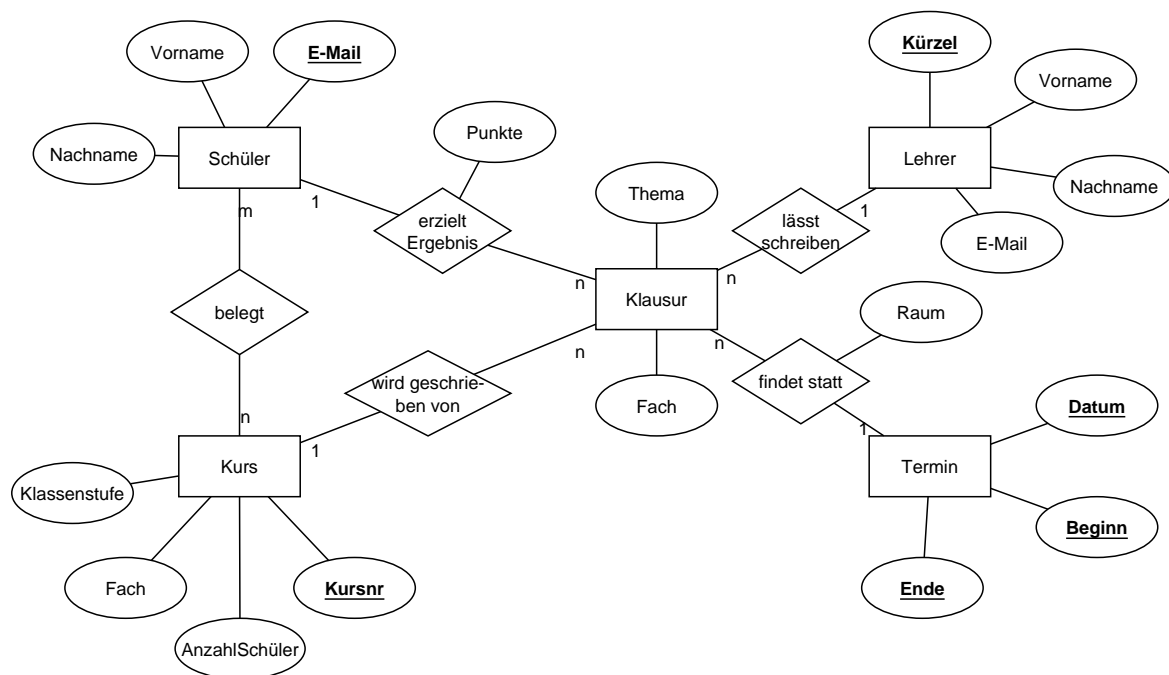


Abbildung 8: ER-Diagramm für eine Klausurplanung

Aufgabe 9: Beschreiben Sie die ER-Diagramme einer Fahrschule in Abbildung 9 und Abbildung 10 und gehen Sie auf die Unterschiede ein. Diskutieren Sie Vor- und Nachteile der beiden Modellierungen.

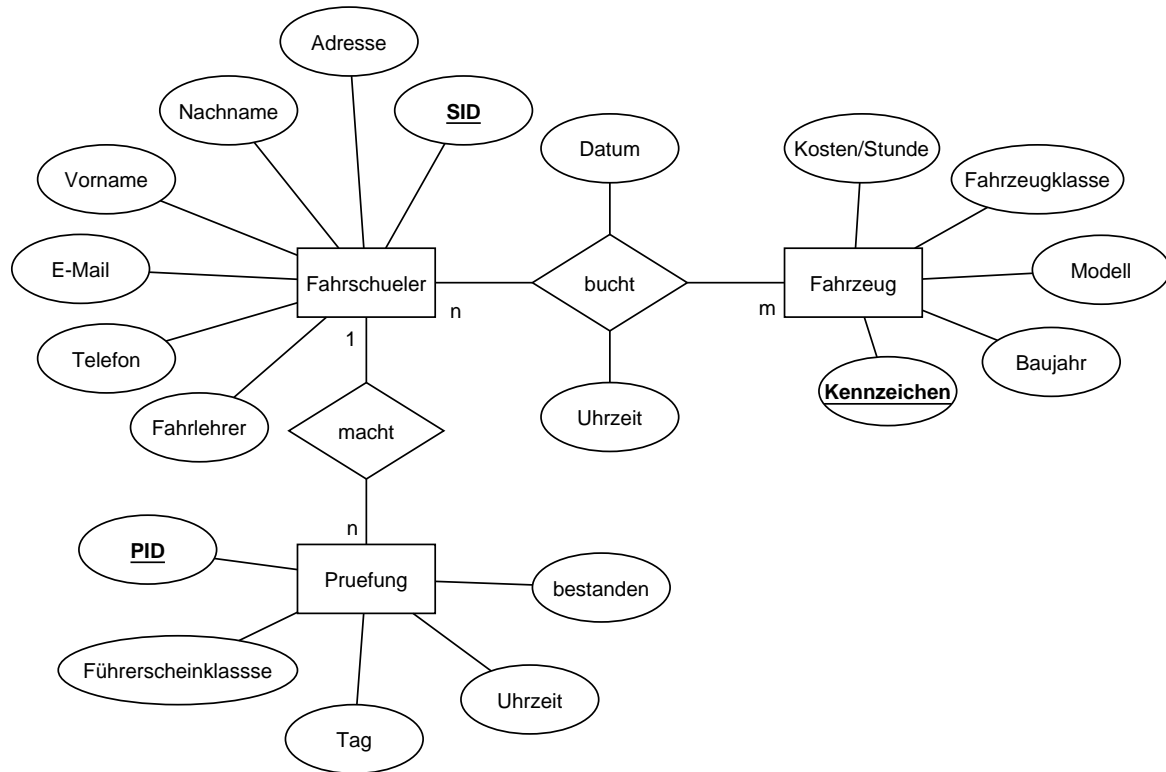


Abbildung 9: ER-Diagramm einer Fahrschule - Variante A

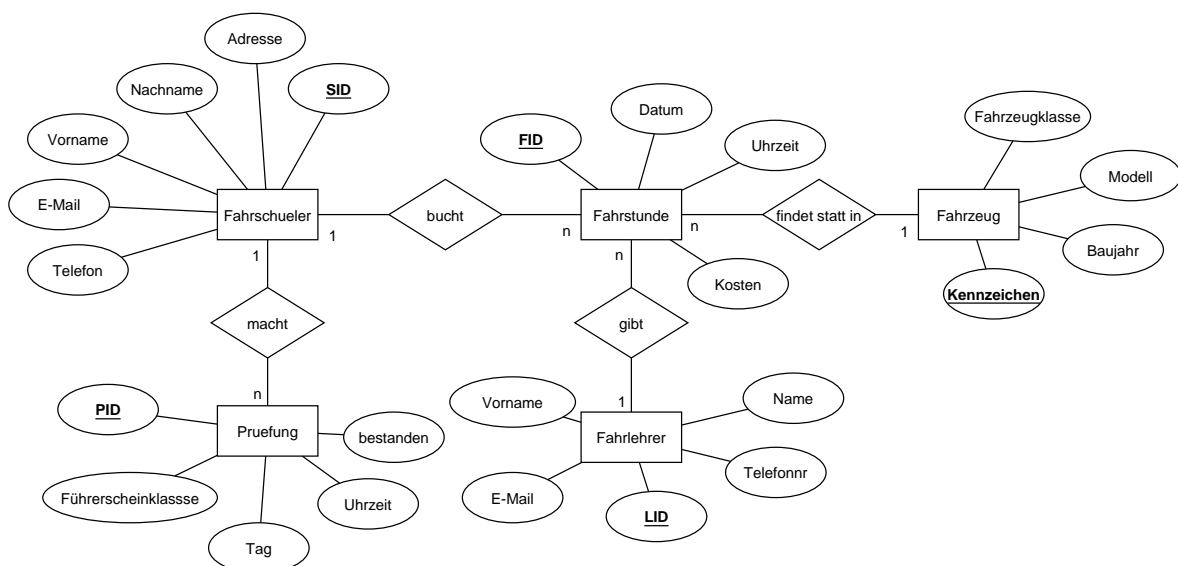


Abbildung 10: ER-Diagramm einer Fahrschule - Variante B

Aufgabe 10: Wählen Sie einen Weltausschnitt, für den Sie ein ER-Diagramm erstellen. Diskutieren Sie dazu in der Kleingruppe, welche Informationen jeweils benötigt werden. Vielleicht können Sie sogar ein Experteninterview durchführen. Beispiele für mögliche Szenarien sind:

- eine Autovermietung
- eine Vermietung von Ferienwohnungen
- ein Lieferservice für Pizza
- ein Fitnessstudio

Exkurs: Datenbankentwurf mithilfe unterschiedlicher Sichten

Komplexe Datenbanksysteme, auf denen später unterschiedliche Anwendungen arbeiten sollen, können oft nur überblickt und modelliert werden, indem man zunächst einzelne ER-Diagramme für die Anforderungen unterschiedlicher Bereiche entwirft. Man spricht hier von unterschiedlichen **Sichten**. In einem zweiten Schritt werden die einzelnen ER-Diagramme zu einem Gesamtmodell zusammengeführt, das schließlich alle Anforderungen berücksichtigt. Da es zwischen den Bereichen und den entsprechenden ER-Diagrammen Schnittmengen geben wird, muss auf folgende Aspekte geachtet werden:

- **Redundanzfreiheit:** Entitäten und Beziehungen, die in mehreren Teildiagrammen eine Rolle spielen, dürfen nur einmal im Gesamtmodell enthalten sein.
- **Widerspruchsfreiheit:** Die modellierten Sachverhalte dürfen sich weder inhaltlich noch strukturell widersprechen. Beispielsweise darf eine Entität *Klausur* nicht an anderer Stelle als Beziehung *schreibt Klausur* dargestellt werden. Auch ein Attribut darf nicht gleichzeitig als Entität dargestellt werden.
- **Synonyme bereinigen:** Der gleiche Sachverhalt darf nicht unterschiedliche Bezeichnungen haben. Zum Beispiel darf es nicht die Entität Putzfrau und Reinigungskraft geben, wenn beide dieselbe Gruppe von Personen bezeichnen.
- **Homonyme bereinigen:** Es darf nicht derselbe Begriff für unterschiedliche Sachverhalte verwendet werden. Es darf beispielsweise nicht zwei Entitäten Klasse geben, wobei die eine die Schulklasse als Gruppe und die andere die Jahrgangsstufe meint.

Aufgabe 11: Anhand der Daten, die in den unterschiedlichen Bereichen einer Schule anfallen, soll das Erstellen eines ER-Diagramms mithilfe unterschiedlicher Sichten einmal umgesetzt werden.

Phase 1: Erstellen der Sichten

Bilden Sie Kleingruppen. Jede Kleingruppe erstellt ein ER-Diagramm für einen der drei Bereiche **Sekretariat**, **Stunden- und Vertretungsplan** sowie **Bibliothek**. Im Folgenden erhalten Sie für Ihren Bereich einen Überblick über die Anforderungen. Nehmen Sie ggf. noch Ergänzungen vor. Zwecks besserer Lesbarkeit wird nur die maskuline Form verwendet, Schülerinnen, Lehrerinnen usw. sind hier genauso gemeint.

1. **Sekretariat:** Im Sekretariat werden die persönlichen Daten der Schüler und Lehrer verwaltet. Dazu gehören Name, Adresse, Geburtsdatum und Telefonnummer. Für jeden Schüler werden zusätzlich die Konfession und die Namen der Eltern gespeichert sowie die Klasse, in die die Schüler gehen.
Für jeden Lehrer werden zusätzlich seine Unterrichtsfächer vermerkt und von welcher Klasse er Klassenlehrer ist.

Die Schüler der fünften Klasse bekommen jeweils einen älteren Schüler als Paten zugeteilt.

2. **Stunden- und Vertretungsplan:** Der Stundenplankoordinator, der auch die Vertretungspläne erstellt, benötigt folgende Informationen: Die Stundenpläne der einzelnen Klassen, wobei zu jeder Stunde das Fach, der Raum und der Lehrer angegeben sein muss. Außerdem müssen die Größe der Klasse (Anzahl Schüler), der Klassenraum sowie der Klassenlehrer abrufbar sein. Zu den Räumen werden folgende Angaben benötigt: die Raumnummer, die Größe in Quadratmetern, die aktuelle Anzahl der Sitzplätze und um welche Art von Fachraum es sich ggf. handelt.

Um den Vertretungsunterricht sinnvoll zu planen, muss außerdem bekannt sein, welcher Lehrer, welche Fächer unterrichtet, und welche Lehrer sich in bestimmten Stunden für spontane Vertretungseinsätze bereithalten. Jeder Lehrer bekommt höchstens zwei Stunden zugeteilt, in denen er sich bereithalten muss. Damit kein Lehrer übermäßig belastet wird, wird vermerkt wie viele Stunden ein Lehrer im aktuellen Schuljahr schon vertreten hat.

3. **Bibliothek:** In der Bibliothek werden alle Bücher elektronisch verwaltet. Zu jedem Buch werden eine Signatur, der Titel, der Autor, das Erscheinungsjahr und der Verlag gespeichert. Alle Lehrer und Schüler sind nutzungsberechtigt. Sie werden mit Namen, Adresse, Telefonnummer und E-Mail gespeichert.

Leiht ein Benutzer ein Buch aus, werden dazu das Datum der Ausleihe und das späteste Rückgabedatum vermerkt. Benutzer können ein Buch auch vormerken lassen, so dass es für sie reserviert wird, sobald es zurückgegeben wurde und sie mit dem Ausleihen an der Reihe sind. Wird ein Buch nicht rechtzeitig zurückgegeben, muss der Benutzer eine Mahngebühr bezahlen.

Die Betreuung der Bibliothek wird von Lehrern und Schülern übernommen, die vorher eine entsprechende Schulung erhalten haben. Sie bekommen regelmäßige Zeiten zugeteilt, in denen sie in jeder Woche in der Bibliothek Aufsicht führen und die Ausleihe verwalten. Die Bibliothek ist in den großen Pausen und für eine Stunde nach Schulschluss geöffnet.

Phase 2: Zusammenführung der Sichten

Bilden Sie neue Kleingruppen, in denen für jede Sicht ein bis zwei Entwickler des entsprechenden ER-Diagramms vertreten sind. Führen Sie die drei ER-Diagramme zu einer Gesamtmodellierung zusammen. Überprüfen Sie das vollständige ER-Diagramm auf Redundanzen, Widerspruchsfreiheit und Namenskonflikte.

Überführung eines ER-Diagramms in ein relationales Schema

Abbildung von Entitäten

Um die Modellierung eines ER-Diagramms auf entsprechende Tabellen einer relationalen Datenbank abzubilden, wird im ersten Schritt für jede Entität eine Tabelle mit den entsprechenden Attributen angelegt. Wurden die Schlüsselattribute im ER-Diagramm bereits gekennzeichnet, können sie ebenfalls direkt übernommen werden. Ansonsten müssen Sie geeignet definiert werden.

Abbildung von n:m- und 1:n-Beziehungen

Im zweiten Schritt erfolgt die Abbildung der Beziehungen. Ob für eine Beziehung eine eigene Tabelle eingerichtet wird oder ob die Beziehung in eine der Tabellen der beteiligten Entitäten eingetragen

werden kann, hängt von den Kardinalitäten ab. Für das ER-Diagramm des TSV-Fitstadt in Abbildung 11 ergeben sich die folgenden Tabellen:

Trainer (TrainerID, Vorname, Nachname, Telefonnr, Strasse, Hausnr, PLZ, Ort, Geb-Datum)

Kurs (KursID, Dauer, Uhrzeit, Wochentag, Sportart, ↑TrainerID, ↑Sportstaette_Bezeichnung)

Sportstaette (Bezeichnung, Groesse, Sitzplaetze, Hausnr, Strasse)

Mitglied (Mitgliedsnr, Vorname, Nachname, Telefonnr, Strasse, Hausnr, PLZ, Ort, Geb-Datum, Beitrag, Kontnr)

Belegung (↑Mitgliedsnr, ↑KursID, TeilnahmeWettbewerbe)

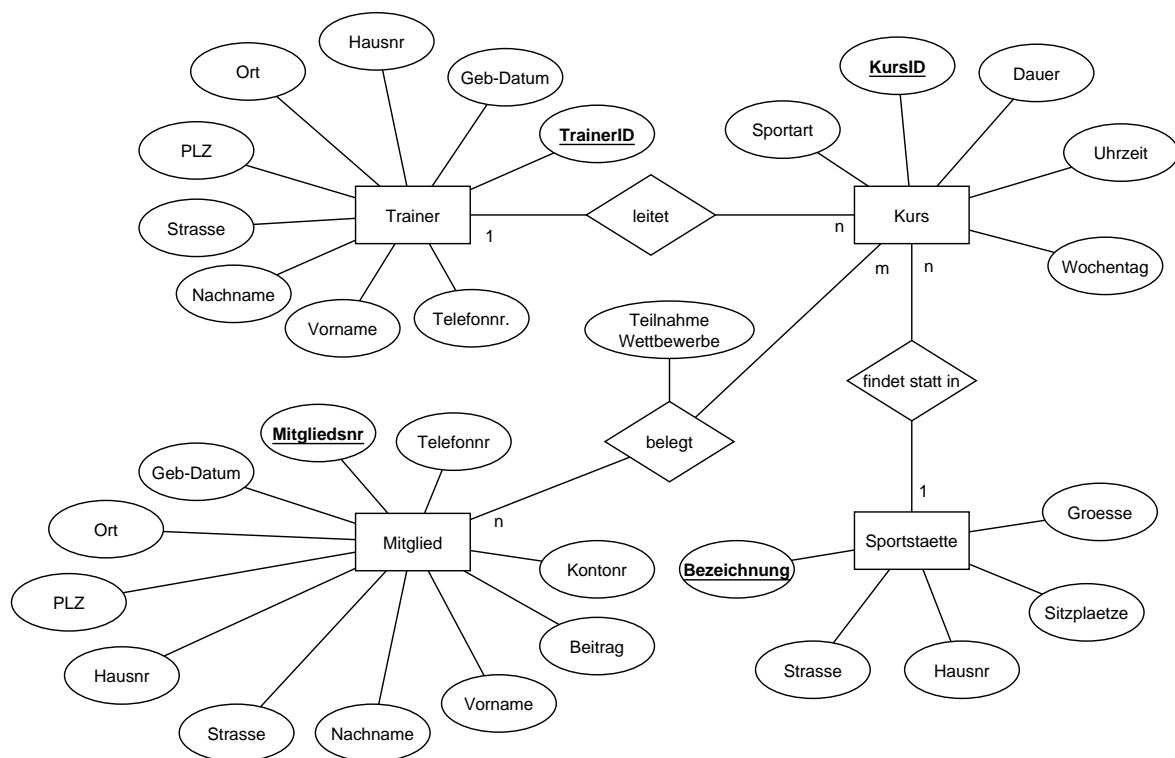


Abbildung 11: ER-Diagramm des TSV-Fitstadt

Aufgabe 12:

- Untersuchen Sie die Umsetzung der Beziehungen des ER-Diagramms des TSV-Fitstadt in dem relationalen Schema (Tabellen). Formulieren Sie Regeln für die Abbildung von n:m und 1:n-Beziehungen auf Tabellen.
- Erläutern Sie, warum ein unterschiedliches Vorgehen hier sinnvoll ist.

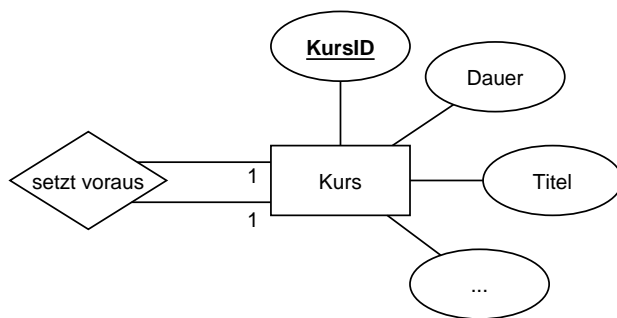
Abbildung von 1:1-Beziehungen

Bei einer 1:1 Beziehung sollte zunächst noch einmal geprüft werden, ob es sich tatsächlich um zwei eigenständige Entitäten handelt oder ob die Attribute der einen Entität auch der anderen zugeordnet werden können, so dass die zweite Entität und die Beziehung nicht benötigt werden.

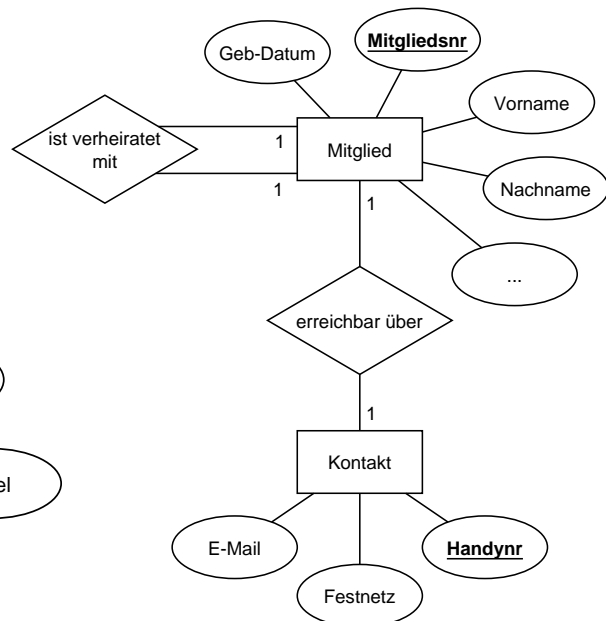
Handelt es sich tatsächlich um zwei Entitäten, ist zu prüfen, in welche Tabelle die Beziehung in Form des Fremdschlüsselattributs aufgenommen werden sollte. Ein Kriterium ist dabei die Vermeidung von leeren Datenfeldern in den Tabellen. Stellt die eine Entität eigentlich Attribute der anderen dar, wäre es möglich, die beiden beteiligten Entitäten vollständig in einer Tabelle zusammenzufassen. Auch

hier ist die Vermeidung leerer Datenfelder ein Kriterium. Insbesondere bei rekursiven 1:1-Beziehungen ist außerdem darauf zu achten, dass die Attribute in der Tabelle unterschiedliche Namen erhalten.

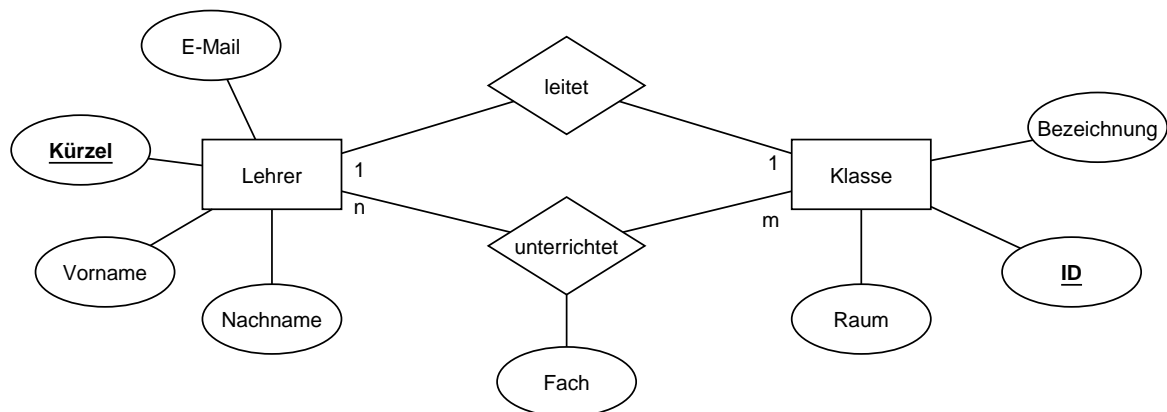
Aufgabe 13: Diskutieren Sie für die 1:1-Beziehungen in den Beispielen 1 bis 3 jeweils, welche Aufteilung der Tabellen Sie für sinnvoll erachten.



Beispiel 1 zu Aufgabe 13



Beispiel 2 zu Aufgabe 13



Beispiel 3 zu Aufgabe 13

Aufgabe 14: Überführen Sie die ER-Diagramme einer Fahrschule aus Aufgabe 9 jeweils in ein relationales Schema.

Optimierung einzelner Tabellen

Die Tabellen, die aus einem ER-Diagramm abgeleitet wurden, sollten noch einmal auf Redundanzen und Anomalien überprüft werden. Außerdem ist sicherzustellen, dass die Werte aller Attribute **atomar**² sind. Wenn Tabellen nachträglich um zusätzliche Attribute erweitert wurden oder Tabellen aus anderer Quelle übernommen werden, ist eine solche Überprüfung ebenfalls sinnvoll.

Hat man redundante Daten und Anomalien in einer Tabelle identifiziert, sollte die Tabellenstruktur überdacht und nach weiteren Aufteilungsmöglichkeiten in mehrere Tabellen gesucht werden. Dazu kann auf der Ebene der Entitäten und Beziehungen noch einmal geprüft werden, ob diese nicht sauber getrennt wurden. So kann man neben den Kursen in der Tabelle *Kurse* aus Abbildung 12 die eigenständigen Entitäten *Sportstätte* und *Ansprechpartner* ausfindig machen.

Kurse

KursID	Sportart	Sportstaette	Hallen-groesse	Ansprech-partner	Telefon	Wochentag	Uhrzeit	Dauer	Trainer
...
TWJ	Turnen	Sporthalle IGS	300	Herr Freund	0151-73451239	Dienstag	16:00	60	26
HHE	Handball	Sporthalle 2	200	Frau Holle	0162-192837465	Mittwoch	17:30	90	25
HHJ	Handball	Sporthalle 2	200	Frau Holle	0162-192837465	Donnerstag	17:00	60	25
...

Abbildung 12: Wiederholung der erweiterten Tabelle *Kurse* aus Abbildung 1

Alternativ kann die Abhängigkeit³ der einzelnen Attribute betrachtet werden. Im Beispiel der Tabelle *Kurse* hängen die Hallengröße, der Ansprechpartner und die Telefonnummer von der Sportstätte ab. Da das Attribut *Sportstaette* aber kein Schlüsselattribut ist, kann eine Sportstätte mehrfach in der Tabelle auftreten. Aufgrund der Abhängigkeit wiederholen sich die Werte für Hallengröße, Ansprechpartner und Telefonnummer für jede Wiederholung der Sportstätte. Daher ist es sinnvoll eine separate Tabelle für die Sportstätte einzurichten. In der Tabelle *Kurse* reicht es dann, den Primärschlüssel der Tabelle *Sportstaetten* als Fremdschlüssel aufzunehmen. Die abhängigen Attribute werden dann nur einmalig in der Tabelle *Sportstaetten* gespeichert.

Kurse

KursID	Sportart	Sportstaette	Wochentag	Uhrzeit	Dauer	Trainer
...
TWJ	Turnen	Sporthalle IGS	Dienstag	16:00	60	26
HHE	Handball	Sporthalle 2	Mittwoch	17:30	90	25
HHJ	Handball	Sporthalle 2	Donnerstag	17:00	60	25
...

Sportstaetten

Sportstaette	Hallengroesse	Ansprechpartner	Telefon
...
Sporthalle IGS	300	Herr Freund	0151-73451239
Sporthalle 2	200	Frau Holle	0162-192837465
...

Abbildung 13: Aufteilung der Tabelle *Kurse* in die Tabellen *Kurse* und *Sportstaetten*

² Zur Erinnerung: Man sagt, dass die Wertebereiche der Attribute atomar sind, wenn sie aus unteilbaren Informationseinheiten bestehen. Das heißt, ein Attribut darf keine Listen oder zusammengesetzte Werte enthalten. Der Wert „Peter Müller“ wäre beispielsweise nicht atomar, da er sich in Vor- und Nachname zerlegen lässt.

³ Abhängigkeit bedeutet hier, dass der Wert eines Attributs die Werte der abhängigen Attribute eindeutig bestimmt.

Wenn wir davon ausgehen, dass ein Ansprechpartner für mehrere Sportstätten zuständig sein kann, enthält die Tabelle *Sportstaetten* ggf. immer noch redundante Daten. Die Telefonnummer hing in der ursprünglichen Tabelle *Kurse* nämlich gar nicht direkt von der Sportstätte ab. Die Telefonnummer lässt sich vielmehr dem Ansprechpartner zuordnen. Um die redundante Speicherung der Telefonnummer zu vermeiden, können wir eine weitere Tabelle *Kontakte* für die Ansprechpartner einrichten:

Sportstaetten

Sportstaette	Hallengroesse	Ansprechpartner
...
Sporthalle IGS	300	Herr Freund
Sporthalle 2	200	Frau Holle
...

Kontakt

Ansprechpartner	Telefon
...	...
Herr Freund	0151-73451239
Frau Holle	0162-192837465
...	...

Abbildung 14: Aufteilung der Tabelle *Sportstaetten* in die Tabellen *Sportstaetten* und *Kontakte*

Für die Tabelle *Kontakte* wäre noch zu überlegen, ob die Namen eindeutig sind oder ob hier ein künstliches Schlüsselattribut in Form einer ID benötigt wird. Weiterhin sind die Werte des Attributs *Ansprechpartner* nicht atomar, da sie in die Informationen *Anrede* und *Name* zerlegt werden können. Eine Verbesserung stellt daher die folgende Tabellenstruktur dar:

Sportstaetten

Sportstaette	Hallengroesse	Ansprechpartner
...
Sporthalle IGS	300	5
Sporthalle 2	200	6
...

Kontakte

AnsprechpartnerID	Anrede	Name	Telefon
...
5	Herr	Freund	0151-73451239
6	Frau	Holle	0162-192837465
...

Abbildung 15: Zerlegung der Attribute der Tabelle *Ansprechpartner* in atomare Wertebereiche

Zu einem ähnlichen Ergebnis kommen wir, wenn wir uns auf der Modellierungsebene überlegen, dass Kurse, Sportstätten und die Betreiber von Sportstätten jeweils eigene Entitäten mit zugeordneten Attributen darstellen, die miteinander in 1:n-Beziehungen stehen. Bei einer gut durchdachten Modellierung haben wir also am Ende weniger Arbeit beim Beseitigen von Redundanzen, Anomalien und der Aufteilung von Attributen in atomare Werte.

Spezialfall zusammengesetzte Primärschlüssel

Bei manchen Tabellen wird der Primärschlüssel aus zwei oder mehr Attributen zusammengesetzt. In diesem Fall sollte auf der Ebene der Abhängigkeiten geprüft werden, ob es Attribute gibt, die bereits durch eines der Schlüsselattribute eindeutig festgelegt sind. Auch dies kann einen Hinweis darauf liefern, dass hier Attribute in eine eigenständige Tabelle ausgelagert werden können, um Redundanzen zu vermeiden.

Abbildung 16 zeigt eine Tabelle in der Patientendaten verwaltet werden. Die Attribute *PatientID* und *PZN* (Pharmazentralnummer) bilden den Primärschlüssel. Die Attribute *Vorname*, *Nachname* und *ZimmerNr* lassen sich dem Patienten zuordnen und hängen daher nur von der *PatientID* ab. Die Attribute *Darreichungsform*, *Einheit* für die Dosierung und *Hersteller* hängen hingegen nur von der *PZN* ab. Nur für das Attribut *Dosierung* wird die Kombination aus beiden Schlüsselattributen *PatientID* und *PZN* benötigt. Es lassen sich daher die Entitäten *Patient* und *Medikament* ausfindig machen, die über eine Beziehung *nimmt ein* mit einer individuellen Dosierung miteinander verknüpft sind. Um Redundanzen zu vermeiden, erfolgt daher eine Aufteilung der Tabelle *Patienten* in drei Tabellen: *Patienten*, *Medikamente* und *Medikationsplan* (s. Abbildung 17).

Patienten

<u>Patient ID</u>	Vorname	Nachname	ZimmerNr	<u>PZN</u>	Dosierung	Einheit	Darreichungsform	Hersteller
1	Susi	Müller	101	2345689	1	Stück	Tablette	Pharmacon
1	Susi	Müller	101	7894561	10	ml	Saft	Geplus
2	Heiner	Bach	302	2345689	2	Stück	Tablette	Pharmacon
2	Heiner	Bach	302	3764825	1	Stück	Zäpfchen	Fixi
...

Abbildung 16: Tabelle Patienten

Patienten

<u>PatientID</u>	Vorname	Nachname	Zimmer-Nr
1	Susi	Müller	101
2	Heiner	Bach	302
...

Medikamente

<u>PZN</u>	Einheit	Darreichungsform	Hersteller
2345689	Stück	Tablette	Pharmacon
7894561	ml	Saft	Geplus
3764825	Stück	Zäpfchen	Fixi
...

Medikationsplan

<u>↑PatientID</u>	<u>↑PZN</u>	Dosierung
1	2345689	1
1	7894561	10
2	2345689	2
2	3764825	1
...

Abbildung 17: Aufteilung der Tabelle Patienten in drei Tabellen: Patienten, Medikamente und Medikationsplan

Auf der Ebene der Modellierung ergeben sich die gleichen Tabellen, wenn die Entitäten *Patient* und *Medikament* mit einer n:m-Beziehung *nimmt ein* verbunden werden, die das Attribut *Dosierung* erhält.

Aufgabe 15: Untersuchen Sie die Abhängigkeiten der Attribute in der Tabelle des Webshops (Abbildung 2 aus Aufgabe 1) und zerlegen Sie sie geeignet, so dass keine Redundanzen und Anomalien mehr auftreten. Stellen Sie außerdem sicher, dass alle Attribute atomare Wertebereiche haben.

Aufgabe 16: Untersuchen Sie die Abhängigkeiten der Attribute in der Tabelle des Zoos (Abbildung 3 aus Aufgabe 2) und zerlegen Sie sie geeignet, so dass keine Redundanzen und Anomalien mehr auftreten. Stellen Sie außerdem sicher, dass alle Attribute atomare Wertebereiche haben.

Aufgabe 17: Untersuchen Sie für die Tabelle des Tierheims aus Aufgabe 3 (Abbildung 4), inwieweit Ihnen eine Zerlegung in einzelne Tabellen zur Vermeidung von Redundanzen sinnvoll erscheint.

Hinweis

Die Materialien erheben keinen Anspruch auf Vollständigkeit hinsichtlich der für die Abiturprüfung erwarteten Kompetenzen. Verbindlich für das Abitur in Niedersachsen sind allein das niedersächsische Kerncurriculum für die gymnasiale Oberstufe sowie die ergänzenden Hinweise in der jeweils aktuellen Fassung.

Dieses Werk ist lizenziert unter einer [Creative Commons Namensnennung - Nicht-kommerziell - Weitergabe unter gleichen Bedingungen 4.0 International Lizenz](#). Ausgenommen von dieser Lizenz ist das INFSII-Logo.