

Übung zur Vorlesung *Grundlagen: Datenbanken* im WS16/17

Harald Lang, Linnea Passing (gdb@in.tum.de)

<http://www-db.in.tum.de/teaching/ws1617/grundlagen/>

Blatt Nr. 02

Hausaufgabe 1

Erstellen Sie ein ER-Modell womit sich kausale Zusammenhänge darstellen lassen (Prinzip von Ursache und Wirkung). Nehmen Sie an, dass eine Ursache mehrere Wirkungen haben kann, und dass eine Wirkung auf maximal eine Ursache zurückzuführen ist. Geben Sie die Funktionalitäten an. Verwenden Sie die (min,max)-Notation.

Lösung:

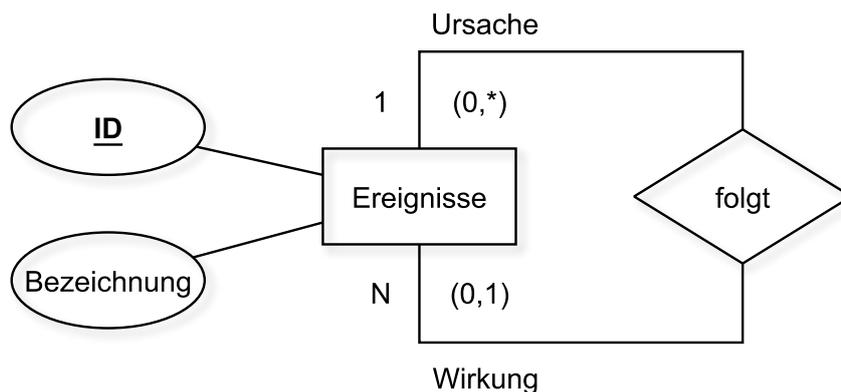


Abbildung 1: ER-Modellierung von kausalen Zusammenhängen

Man kann von Ursache und Wirkung abstrahieren und beides als Ereignis betrachten. Ereignisse können als Ursache sowie als Wirkung auftreten. Für die Relation folgt gilt dann:

$$\text{folgt} \subseteq \text{Ursache} : \text{Ereignisse} \times \text{Wirkung} : \text{Ereignisse}$$

Eine Ausprägung könnte wie folgt aussehen:

(min,max)-Notation

Ursache (0,*)

Sprich: Ein Ereignis kommt in der Rolle "Ursache" minimal 0 mal und maximal beliebig oft vor.

<i>Ursache</i>		<i>Wirkung</i>	
<i>ID</i>	<i>Bezeichnung</i>	<i>ID</i>	<i>Bezeichnung</i>
0	Big Bang	1	Beginn der Zeit
	⋮		⋮
10	Fahrzeugpanne	11	Stau auf der A99
11	Stau auf der A99	12	Student verpasst Vorlesung
11	Stau auf der A99	13	Telekom-Techniker verspätet sich mit DSL-Anschluss
	⋮		⋮

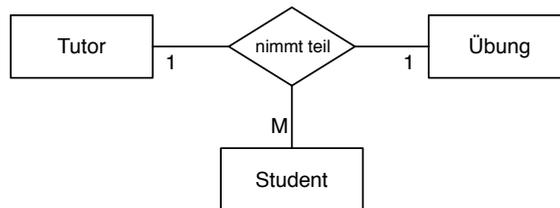
Zum aktuellen Zeitpunkt ist noch nicht bekannt, welche Auswirkungen das Ereignis “*Student verpasst Vorlesung*” haben wird, es taucht also 0 mal als Ursache auf. Das Ereignis “*Stau auf der A99*” löst in der oben gezeigten Ausprägung zwei Folgeereignisse aus: “*Student verpasst Vorlesung*” und “*Telekom-Techniker verspätet sich mit DSL-Anschluss*”. Denkbar sind beliebig viele (weitere) Auswirkungen.

Wirkung (0,1)

Sprich: Ein Ereignis kommt in der Rolle “*Wirkung*” minimal 0 mal und maximal 1 mal vor.

Den Beginn einer Kausalkette bildet ein Ereignis, welches keine Wirkung einer Vorangegangenen Ursache ist. So ist der “*Big Bang*” ein Ereignis welches keine (bekannte) Ursache jedoch viele Auswirkungen hat. Er taucht also nur in der Rolle “*Ursache*” auf und somit 0 mal als Wirkung. Das Ereignis “*Stau auf der A99*” wurde durch die “*Fahrzeugpanne*” verursacht. Nach der Definition in der Aufgabenstellung gilt, dass ein Ereignis maximal eine Ursache hat. Es taucht somit auf der rechten Seite als Wirkung maximal 1 mal auf.

Hausaufgabe 2



Ignorieren Sie die Funktionalitätsangaben und beantworten Sie:

- Wie viele partielle Funktionen der Form $A \times B \rightarrow C$ können in einer ternären Beziehung auftreten (Ignorieren Sie beim Zählen die Reihenfolge auf der linken Seite der Beziehung).
- Nennen Sie alle möglichen partiellen Beziehungen in der hier gezeigten Beziehung „nimmt teil“.
- Nennen Sie für jede Funktion in Prosa, welche Einschränkung diese darstellt, falls sie gilt.

Unter Berücksichtigung der Funktionalitätsangaben:

- Welche partiellen Funktionen gelten hier?

Lösung:

Es gibt drei mögliche partielle Funktionen:

$$Tutor \times Uebung \rightarrow Student \quad (1)$$

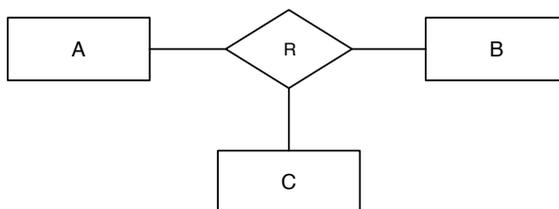
$$Tutor \times Student \rightarrow Uebung \quad (2)$$

$$Uebung \times Student \rightarrow Tutor \quad (3)$$

Würde Funktion 1 gelten, so darf ein Tutor pro Übung nur einen Studenten haben. Gilt Funktion 2, so darf ein Student bei einem Tutor nur eine Übung besuchen. Gilt Funktion 3, so darf es für eine konkrete Übung nur einen Tutor geben.

Zu den in der Abbildung gezeigten Kardinalitätsangaben „passen“ die partiellen Funktionen 2 und 3, weshalb diese für das Beispiel gelten. 1 gilt hingegen - wie auch bei uns im Übungsbetrieb - nicht.

Hausaufgabe 3



Angenommen, lediglich die partielle Funktion

$$A \times C \rightarrow B$$

gilt. Beschriften Sie die Abbildung mit Funktionalitätsangaben.

Beantworten Sie nun die Frage, wie Funktionalitätsangaben aus partiellen Funktionen und umgekehrt ermittelt werden können. Merken Sie sich die Antwort für die Klausur ;-)

Lösung:

An der Entität B wird eine 1 annotiert, an A und C jeweils ein N bzw. M.

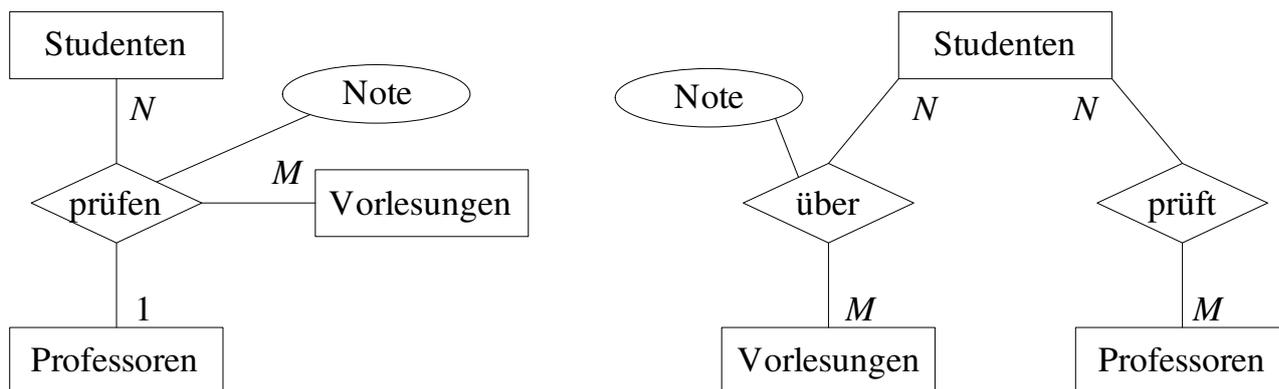
Eine einfache Daumenregel ist, dass an die Entität, die auf der rechten Seite des Pfeiles einer geltenden partiellen Funktion steht, eine 1 annotiert wird. Es bietet sich daher an, für die sichere Bestimmung der Kardinalitätsangaben grundsätzlich die möglichen partiellen Funktionen aufzustellen und zu überlegen, welche Einschränkungen gewünscht sind.

Hausaufgabe 4

Beim konzeptuellen Entwurf hat man gewisse Freiheitsgrade hinsichtlich der Modellierung der realen Welt. Unter anderem hat man folgende Alternativen, die Sie an unserem Universitätsschema beispielhaft illustrieren sollten:

- Man kann ternäre Beziehungen in binäre Beziehungen transformieren.

Betrachten Sie dazu die Beziehung *prüfen* und erläutern Sie die Vor- und Nachteile einer solchen Transformation.



Studenten \times Vorlesungen \rightarrow Professoren

Keine Einschränkungen

Abbildung 2: Auflösen der ternären Beziehung *prüfen* in binäre Beziehungen

- Man hat manchmal die Wahl, ein Konzept der realen Welt als Beziehung oder als Entitytyp zu modellieren. Erörtern Sie dies wiederum am Beispiel der Beziehung *prüfen* im Gegensatz zu einem eigenständigen Entitytyp *Prüfungen*.
- Ein Konzept der realen Welt kann manchmal als Entitytyp mit zugehörigem Beziehungstyp und manchmal als Attribut dargestellt werden. Ein Beispiel hierfür ist das Attribut *Raum* des Entitytyps *Professoren* im bekannten Uni Schema. Diskutieren Sie die Alternativen.

Lösung:

Ziel dieser Aufgabe ist es, alternative Entwürfe zu erstellen und bezüglich ihrer Anwendbarkeit zu analysieren. Unter Anwendbarkeit ist unter anderem zu verstehen, ob in der neuen Modellierung dieselben Informationseinheiten wie in der ursprünglichen abgebildet werden können, ob Konsistenzbedingungen eingehalten werden und ob die reale Welt in der modellierten Miniwelt sinnvoll wiedergegeben ist.

Erste Teilaufgabe: Transformation der ternären Beziehung in binäre Beziehungen

Abbildung 2 zeigt einen Ansatz, die ternäre Beziehung *prüfen* durch binäre Relationen auszudrücken. Durch die ursprüngliche Modellierung (links in der Abbildung) wird folgende Konsistenzbedingung ausgedrückt:

$$\text{prüfen} : \text{Studenten} \times \text{Vorlesungen} \rightarrow \text{Professoren} \quad (4)$$

Demgegenüber tritt bei der vorgeschlagenen Modellierung mittels binärer Relationen ein Semantikverlust auf. Durch die allgemeineren N:M-Beziehungen wird obige Konsistenzbedingung nicht mehr abgebildet. Somit ist das Modell der ternären Beziehung in diesem Fall ausdrucksstärker. Zwar lassen sich Prüfungsergebnisse in der alternativen Modellierung abbilden, allerdings geht Aussagekraft verloren. Abgebildet ist, dass Studenten über den Stoff von Vorlesungen geprüft werden, sowie dass Studenten von Professoren geprüft werden. Der Zusammenhang, welche Professoren welche Studenten in welchen Vorlesungen prüfen, ist aber nicht mehr ohne weiteres gegeben. Indirekt lösen lässt sich dies durch die Aufnahme des zusätzlichen Attributs *Prüfungstermin* in die Relation *über* und auch in *prüft*. Da der zusätzlich aufgeführte Prüfungstermin eine Prüfung eindeutig festlegt, lässt sich

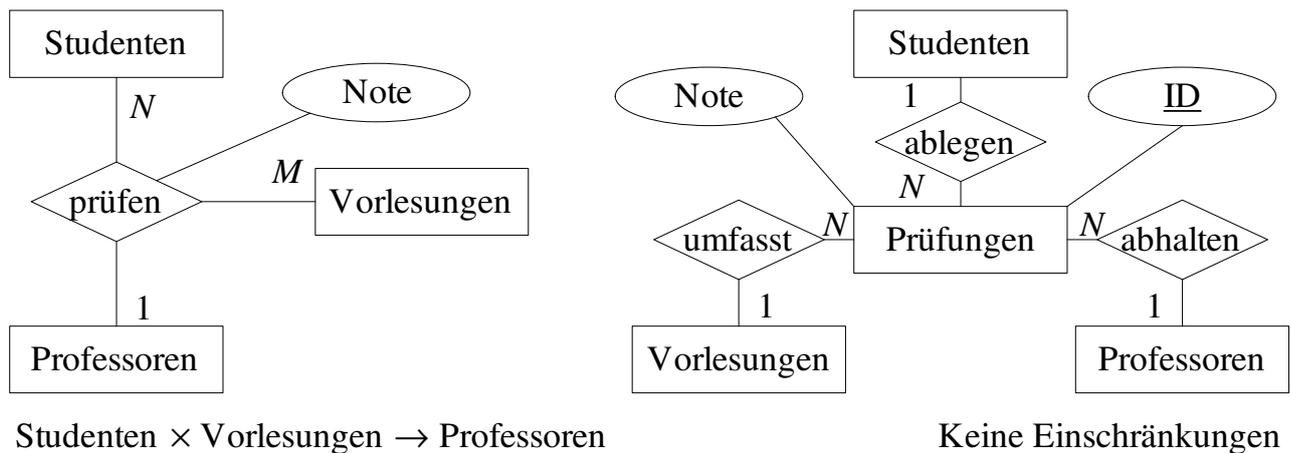


Abbildung 3: Modellierung von Prüfungen als Entitytyp

die Information über eine Prüfung aus beiden Relationen erhalten. Allerdings muss für eine konsistente Extension sichergestellt werden, dass zu einem Eintrag in *über* auch ein passender Eintrag in *prüft* enthalten ist. Die gezeigte alternative Modellierung weist also klare Nachteile gegenüber der ursprünglichen ternären Beziehung auf.

Die alternative Modellierung einer ternären Beziehung durch mehrere binäre kann (abhängig von den zu modellierenden Anforderungen) im Allgemeinen folgende Nachteile aufweisen:

- Es tritt ein Semantikverlust auf.
- Es besteht die Möglichkeit, inkonsistente Datenbankzustände zu generieren. Gegebenenfalls ist eine Konsistenzüberprüfung der Datenbank erforderlich.
- Die reale Welt wird in der Miniwelt unzureichend wiedergegeben.

Zweite Teilaufgabe: Modellierung als Entitytyp anstelle einer Beziehung

Abbildung 3 zeigt eine alternative Modellierung der *prüfen*-Beziehung über einen Entitytyp *Prüfungen*. Auch in diesem Fall tritt erneut ein Semantikverlust auf. Es ist möglich, dass in der Modellierung mittels Entitytyp eine Prüfung existiert, zu der z.B. noch kein Prüfer feststeht, bzw. der Prüfer nicht mehr existiert. Möchte man dies in der Modellierung ausdrücken, müsste man zur *(min,max)*-Notation übergehen, mittels derer man fordern kann, dass eine Prüfung genau je einmal in den Relationen *ablegen*, *abhalten* und *umfasst* auftritt. Außerdem kann auch obige Konsistenzbedingung (4) nicht zugesichert werden. Zwar legt eine *Prüfungen*-Instanz über die angesprochenen Relationen den Studenten / die Studentin, die geprüfte Vorlesung und den / die prüfende(n) Professor(in) fest. Allerdings bestimmt die Kombination *Studenten* \times *Vorlesungen* nun nicht mehr *Professoren*. Denn es ist durch den Entwurf nicht ausgeschlossen, dass es beispielsweise zwei unterschiedliche Prüfungen gibt, die der Student Fichte im Fach Ethik ablegt. Nur einmal lässt er sich bei Professor Sokrates und ein andermal bei Professorin Curie darüber prüfen. Andererseits lassen sich manche Aspekte der Modellierung mittels Entity genauer erfassen, als dies in der ursprünglichen Modellierung der Fall ist. So ist in obigem Beispiel etwa spezifiziert, dass pro Prüfung genau eine Vorlesung geprüft wird.

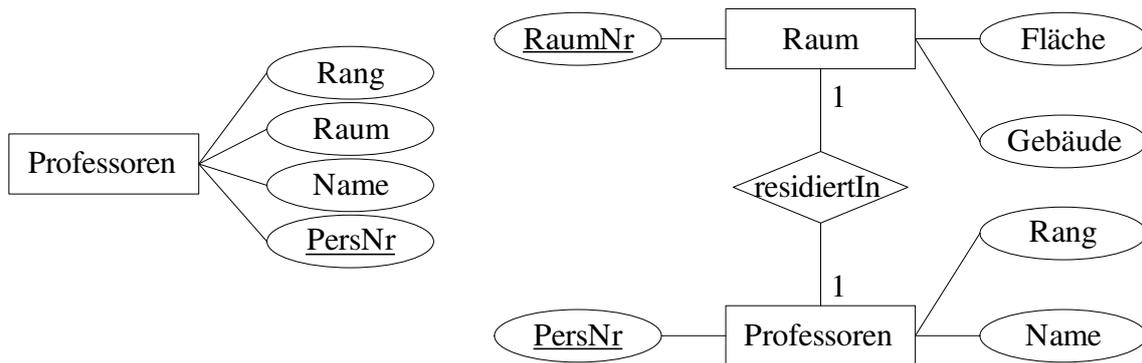


Abbildung 4: Zwei alternative Modellierungen um auszudrücken, dass Professoren ein Raum zugeordnet ist

Dritte Teilaufgabe: Alternative Modellierungen über Attribute oder Beziehungstypen

Abbildung 4 zeigt zwei Modellierungsansätze, die ausdrücken, dass jedem Professor ein Raum zugewiesen ist. Links die Darstellung mittels Attribut, rechts über die Beziehung *residiertIn*. Generell ist eine Modellierung über eine Beziehung mit einem eigenständigen Entity (*Raum*) dann angebracht, wenn entsprechend detaillierte Informationen zu einem Raum nötig sind. Dies kann z.B. dann der Fall sein, wenn die Anwendungssicht der Abteilung Gebäudetechnik in das Modell integriert werden muss. Möchte man die Raumdaten für jeden Professor abfragen, dann zieht diese Modellierung in der Regel eine weniger effiziente Anfrageauswertung nach sich.