

# Zusammenfassung Datenbanken

Heosibel

July 11, 2021

Sommersemester 2021

Professor: Dr.-Ing. Maik Thiele

Dieses Dokument enthält eine Zusammenfassung der Vorlesung Datenbanken vom SS 2021. Diese Dokument ist kein offizielles vom Lehrstuhl und sollte nicht als Referenz verwendet werden. Die Erstellung ist in Vorbereitung auf die Prüfung im SS 2021 aus den Folien geschehen und kann Ungenauigkeit und/oder Fehler enthalten!

# **1 Einführung**

## **1.1 Definition Datenbank**

- logisch konsistent
- besitzt eine bestimmte Bedeutung
- repräsentiert einen Ausschnitt der realen Welt

## **1.2 Ziel einer Datenbank**

- effektives + effizientes Speichern
- Wiederfinden von Daten
- Analyse von Daten

## **1.3 Gründe für DBS**

- Effizienz und Skalierbarkeit
- Fehlerbehandlung und Toleranz
- Mehrbenutzersynchronisation
- Sicherstellung der Datenintegrität
- Deklarative Anfragesprachen ( Benutzer sagt was und nicht wie die Daten geholt werden)
- Datenunabhängigkeit

## **1.4 ANSI-SPARC-Architektur**

- Externe: Definition externer Schemata (Nutzer oder anwendungsspezifische Sichten)
- Logische: definiert die logischen Datenstrukturen und deren Beziehungen
- Interne: Festlegung der Art und Weise der Speicherung

## **1.5 konzeptuelle Modelle**

- Entity-Ralationship-Modell (ER-Modell)
- Unified Modeling Language (UML)

## **1.6 Logische Modelle**

- Hierarchische Modelle
- objektorientierte Modell
- Netzwerkmodelle
- Relationale Modelle
- ...

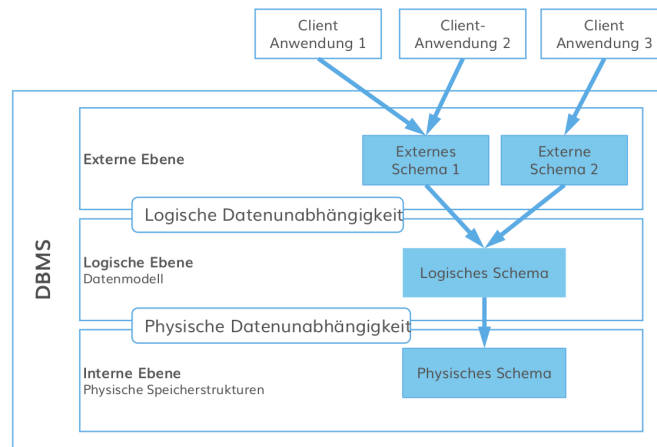


Figure 1: ANSI-SPARC-Architektur

### 1.7 Datenverwaltung ohne DBMS (Probleme)

- kein standardisiertes Format
- Duplikate
- hoher Aufwand bei Kombination von Daten
- Dateninkonsistenz
- maßgeschneidert aber ineffizient
- kein Mehrbenutzerzugriff

## 1.8 Information Retrieval

- Def: finden von Information unstrukturierter Art das eine Informationssnachfrage genügt
- Aufgabe:
  - Suche nach Informationen
  - Strukturierung von meist unstrukturierten Daten
  - Befriedigung des Informationsbedürfnisses eines Benutzers
  - Bewältigung von großen Datenmengen

## 1.9 Arten von Daten

- strukturierte Daten (z.B. Relationen)
- unstrukturiert Daten
- semistrukturierte Daten

## **2 Konzeptueller Entwurf**

### **2.1 Phasen des Entwurfes**

#### **2.1.1 Konzeptueller Entwurf**

- semantisches Modell der Objekttypen
- Beziehungen der Objekttypen

#### **2.1.2 Logischer Entwurf**

- Transformation semantischen Modells in DBS-spezifisches Datenmodell
- z.B. ER-Modell

#### **2.1.3 Physischer Entwurf**

- Einrichten der Datenbank + internes Schemata
- eventuell laden von Daten

### **2.2 Entity-Relationship-Modell**

- Entität: existiert in der Welt und unterscheidet sich von anderen Entitäten
- Attribut: relevantes Merkmal
- Beziehung: Zusammenhänge zwischen Entitäten

### **2.3 Funktionalitäten**

- One-To-One: (1:1) Beziehung von einer Entität zu höchstens einer anderen
- Many-To-One: (1:N)
- Many-To-Many: (N:M) Es liegt keine Beschränkung vor
- n-stellige Beziehungen (z.B. betreuen: Professoren x Studenten → Seminarthemen)

### **2.4 Min/Max Notation**

- min: jede Entität dieses Typs steht mind. min-mal in Beziehung
- max: jede Entität dieses Typs steht höchstens max-mal in Beziehung
- Sonderfälle
  - min = 0: braucht keine Beziehungen
  - max = \*: kann beliebig oft in Beziehung stehen



Figure 2: Min-Max-Notation Beispiel Zug

## 2.5 erweiterte Attribut-Typen

- zusammengesetzte Datentypen
- mehrwertige Datentypen
- abgeleitet Attributwerte

## 2.6 Spezialisierung und Generalisierung

- jede Instanz der Subklasse ist auch Instanz der Klasse

Generalisierung (bottom-up): Unterdrückung der Unterschiede der Objekte

Spezialisierung (top-down): Betonung der Unterschiede

d (disjunkt): Instanzen der Unterklasse disjunkt

o (overlap): Instanzen können sich überlappen

totale Spezialisierung: Jede Instanz der Superklasse muss mind. eine Instanz der Subklasse sein

partielle Spezialisierung: eine Instanz kann zu einer Subklasse gehören

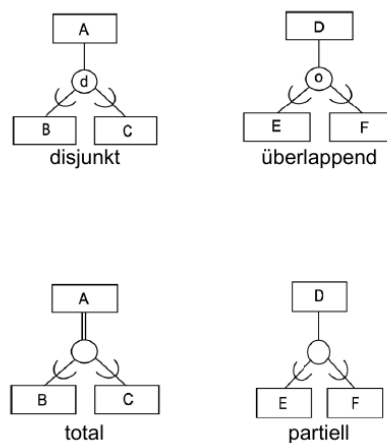


Figure 3: Spezialisierung und Generalisierung Symbole

## 3 Logischer Entwurf

### 3.1 Grundlagen

Tupel: Element einer Relation

Kardinalität: Kardinalität einer Relation: Anzahl der Tupel in einer Relation

Relation Menge von Tupeln

### 3.2 Primärschlüssel

Eindeutigkeit: jeder Schlüsselwert ist einzigartig und nicht doppelt vertreten

Definiertheit: Jedes Objekt hat einen Schlüsselwert

Minimalität es kann kein Teil des Schlüssels weggelassen werden sodass immer noch die Eigenschaften

### 3.3 Fremdschlüssel

Verwendung: Ausdrücken von Beziehungen (verweist auf eine andere Relation wo dieser ein Primärschlüssel ist)

Eigenschaften: Definiertheit und Minimalität

### 3.4 ER-Modell in Relationales Modell

1. Übersetzung von Entitäten
2. Übersetzung von Attributen
3. Übersetzung von 1:1 Beziehungen
4. Übersetzung von 1:N Beziehungen
5. Übersetzung von M:N Beziehungen
6. Übersetzung von Beziehungen zwischen mehr als zwei Relationen
7. Übersetzung rekursiver Beziehungen
8. Übersetzung von Attributen an Beziehungen
9. Übersetzung von Vererbungsbeziehungen

### 3.5 Abbildungsvarianten

#### 3.5.1 Horizontale Partitionierung

- jedes Objekt ist genau ein Tupel einer Relation  $\rightarrow$  gleiche ID heißt nicht das selbe Objekt
- Zusammenführung der Gesamtrelation entstehen viele NULL Objekte

### 3.5.2 Vertikale Zerlegung

- Gesamtheit aller Attribute einer Objektes nur durch den Verbund erhaltbar

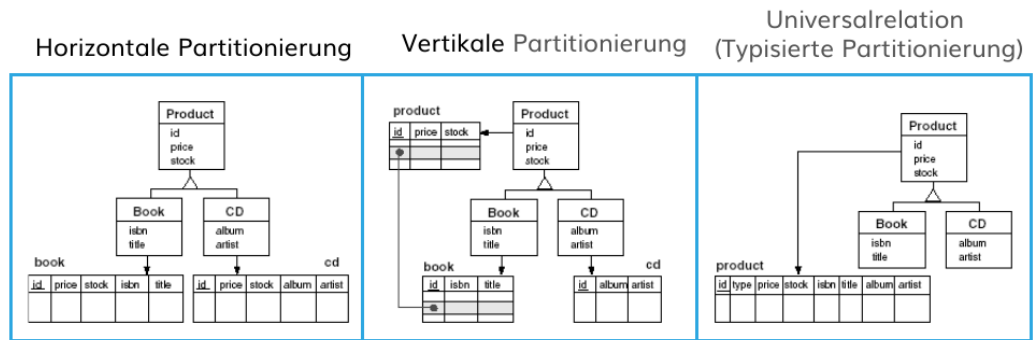


Figure 4: Vertikale und Horizontale Partitionierung



## 4 Relationale Algebra