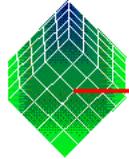


– Schwerpunkt: Grafische Modellierungsnotationen –



Multidimensionale

Datenbanken

Vortrag auf dem 4. Workshop des GI-Arbeitskreises
am 27.4.1998 in Darmstadt

- ❑ **Vorstellung**
- ❑ **Grundlagen**
- ❑ **1.1 Entity Relationship Model**
- ❑ **1.2 Objekttypenmethode**
- ❑ **1.3 Exkurs: Modellierungsansatz der G.S.E.**
- ❑ **2 Application Design for Analytical Processing Technologies**
- ❑ **3 Objektorientierte Modellierung mit der UML**
- ❑ **Ausblick**

Vorstellung

Referent: Dipl.-Inform. Andreas Totok

Technische Universität Carolo-Wilhelmina zu Braunschweig

Institut für Wirtschaftswissenschaften
Abteilung Controlling und Unternehmensrechnung
Pockelsstraße 14
D-38106 Braunschweig

Telefon: 05 31/ 391-36 06

Fax: 05 31/ 391-81 21

E-Mail: A.Totok@tu-bs.de

WWW: <http://www.tu-bs.de/institute/wirtschaftswi/controlling/staff/atotok/atotok.html>



Dissertationsthema: Modellierung von multidimensionalen Datenstrukturen für
Controllinganalysen (Arbeitstitel)

(Folien werden im PDF-Format auf meiner Homepage zum Download bereitgestellt)



Semantische betriebswirtschaftliche Modellierung

- ❑ „Ein Datenmodell ist ein strukturiertes Abbild der Daten eines fest abgegrenzten Teils der wahrgenommenen Realität, die für eine bestimmte Anwendung bzw. für bestimmte Anwender relevant sind, einschließlich der zwischen ihnen bestehenden Beziehungen.“
Maier 1998, S. 130
- ❑ Für die semantische Modellierung von *betriebswirtschaftlichen* Zusammenhängen reicht eine Sicht auf die reine Datenstruktur nicht aus, da
- ❑ eine semantische Modellierung die **Bedeutung** der handelnden *Personen*, ihrer *Verpflichtungen, Rollen, Handlungen* und *Mitteilungen* interpretieren muß. Betriebswirtschaftlich gesehen werden die Geschäftsprozesse abgebildet.
- ❑ Eine *semantische betriebswirtschaftliche* Modellierung beinhaltet daher unterschiedliche Sichten, wie für **Daten, Funktionen**, Organisation sowie für die Steuerung von dynamischen Abläufen.
- ❑ Kritischer Erfolgsfaktor für die Modellierung (und Implementierung) entscheidungsunterstützender Systeme ist der Deckungsgrad des Informationsbedarfs der Anwender.
- ❑ *Den Ausgangspunkt einer betriebswirtschaftlichen Modellierung muß eine grafische Notation bilden. Textuelle Beschreibungen sollten nur der Präzisierung dienen.*



Anforderungen an Datenmodelle

Kriterien für die Wissensrepräsentation in semantischen Datenmodellen:

Verständlichkeit

Der Anwender aus der Fachabteilung und der Systemdesigner müssen das enthaltene strukturelle Wissen verstehen und Informationen entnehmen können.

Ausdrucksstärke

Möglichst viele Sachverhalte (Daten, Regeln, Operationen) sollten mit größtmöglicher Präzision formalisiert dargestellt werden können.

Zerlegbarkeit

Sachverhalte sollten sich in der Darstellung einfach zerlegen lassen.

Validierbarkeit

Die Richtigkeit der Inhalte muß durch Verständlichkeit und Formalisierung sichergestellt werden können.

Veränderbarkeit

Modelle müssen veränder- und erweiterbar sein.

(Verarbeitungsgeschwindigkeit)

Hars 1994, S. 47 f



Einsatzbereiche von Datenmodellen

	Entwurf von DV-Systemen	Informationsmanagement (Nutzung des Produktionsfaktors Information)
Klärungs- funktion	<ul style="list-style-type: none"> • Anforderungsspezifikation • Integration heterogener Systeme • Reengineering 	<ul style="list-style-type: none"> • Begriffsklärung • Informationsbedarfsanalyse
Dokumentations- funktion	<ul style="list-style-type: none"> • Dokumentation der DV-Systeme • Kommunikation mit dem Anwender 	<ul style="list-style-type: none"> • Globales Informationsverzeichnis • Zugriffsmedium • Begriffsdokumentation (Organisationshandbuch) • Einarbeitung von Mitarbeitern
Gestaltungs- funktion	<ul style="list-style-type: none"> • Projektabgrenzung • Schnittstellendefinition • Trennung fachlicher und DV-technischer Aspekte (Wiederverwendbarkeit) • Generieren von Datendefinitionen • Auswahl von Software • Konfiguration 	<ul style="list-style-type: none"> • Schwachstellenanalyse • Informationsbewertung • Verantwortlichkeit für Information • Reintegration von Funktionen

Hars 1994, S. 29



Multidimensionale Modellierung

- ❑ Multidimensionale Datenmodelle werden zur Konzeption von Data-Warehouse-Datenbanken bzw. OLAP-Servern, also für **dispositive Informationssysteme** benötigt.
- ❑ Die **Denkweise** unterscheidet sich grundsätzlich von der Modellierung von operativen Systemen, da **Informationsobjekte für Managemententscheidungen** im Vordergrund stehen.
- ❑ Die primär für die Abbildung von operativen Geschäftsprozessen benutzten konventionellen Modellierungsnotationen trennen organisatorische Aspekte der Datenhaltung von der Interaktion der Datenbasis mit den Anwendungsprogrammen.
- ❑ Daher sollten diese Notationen nicht unmodifiziert für multidimensionale Systeme angewendet werden.
- ❑ Die grundlegenden Elemente eines multidimensionalen Modells sind:
 - **Variable bzw. Kennzahlen** (synonym Variables, Facts, Measures, measured Facts)
 - **Dimensionen** (betriebswirtschaftliche Entscheidungsobjekte)
 - **Formeln** (Konsolidierung, Kennzahlenberechnung, betriebswirtschaftliche Methoden)
- ❑ Die Darstellungsformen der im folgenden vorgestellten Notationen orientieren sich meist am Star-Schema mit Fakten- und Dimensionstabellen.
- ❑ Die Notation muß generell aber unabhängig von der physikalischen Implementierung sein.



Beispielhafte Problemstellung im Controlling

- Ein Unternehmen möchte im Rahmen seines Vertriebscontrolling umfangreiche Auswertungen vornehmen.
 - Betriebswirtschaftliche Variable
 - ◆ Bruttoerlös, Erlösschmälerungen, *Nettoerlös*, Variable Kosten, *Deckungsbeitrag I*
 - Dimensionen
 - ◆ Artikel: Einzel, Gruppe, Marke sowie Sondermodellgruppe
 - ◆ Szenarien: Plan, Ist, *Abweichung*
 - ◆ Vertriebsweg: Filiale, *Gebiet*, *Land*, *Region* jeweils mit Attributen
 - ◆ Zeit: *Jahr*, *Quartale* (I, II, III, IV), *Monate* (Januar bis Dezember), Tage (monatsabhängig)
 - Formeln: Alle kursiv angegebenen Elemente sowie Konsolidierungen
- Besonderheiten:
 - Tagesinformationen sollen nicht im dispositiven System abgelegt werden, sondern werden im Bedarfsfall mit einem SQL-Durchgriff aus operativen Systemen abgeholt.
 - Bei der Sondermodellgruppe handelt es sich nur um eine alternative Sichtweise.
 - Die Elemente der jeweiligen Hierarchieebene der Zeitdimension haben eine feste logische Reihenfolge (z.B. Januar bis Dezember).



1.1 Entity Relationship Model (ERM)

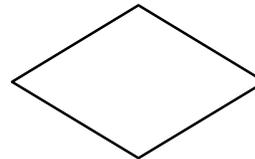
Grundelemente



Entity

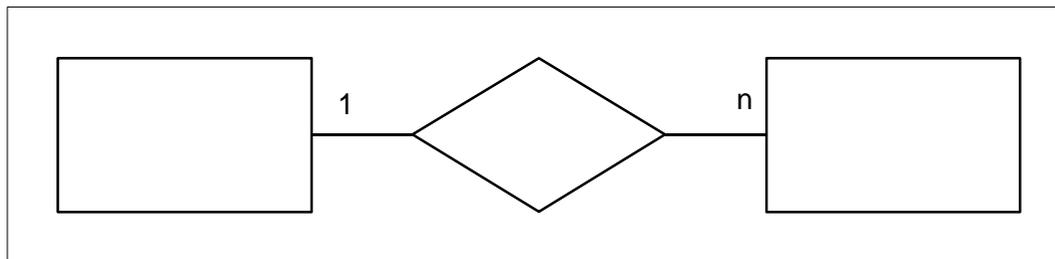


Attribut

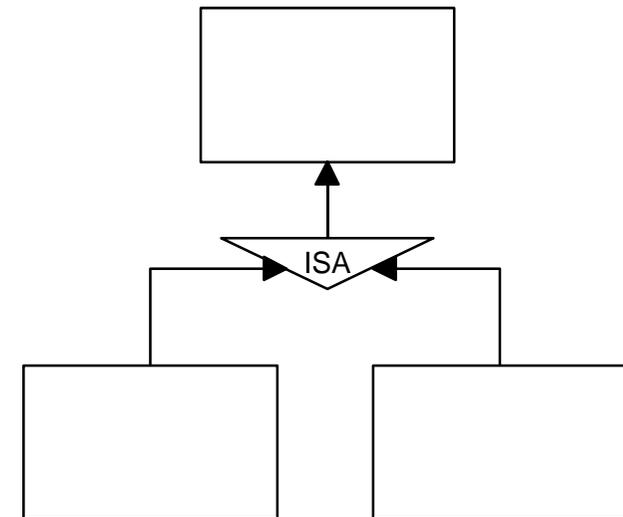


Beziehung

Erweiterungen



Aggregation durch Beziehung, aufgefaßt als Entity mit Kardinalitäten



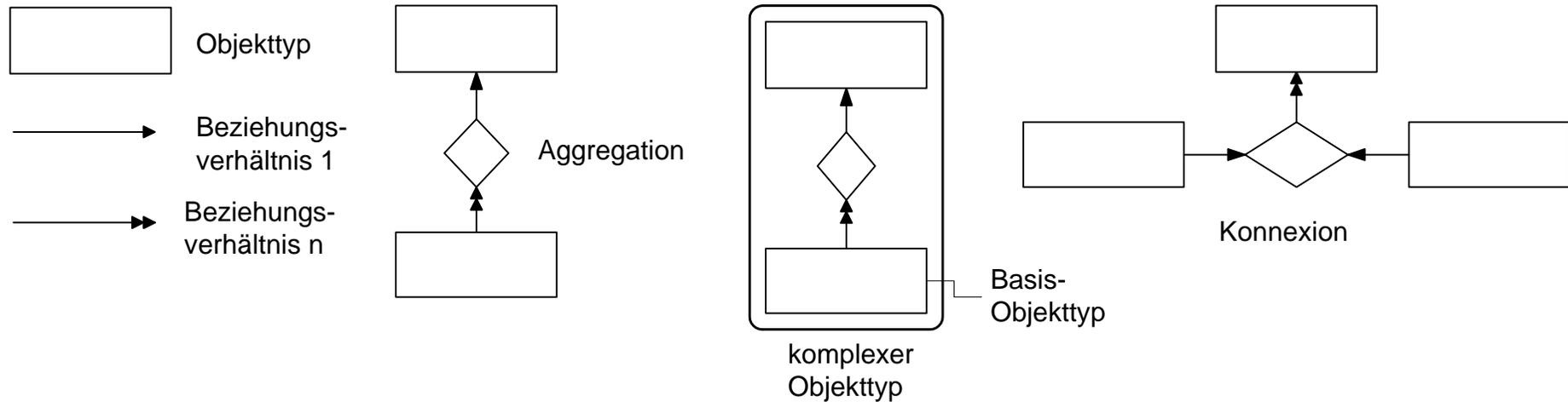
Spezialisierung

Chen 1976; zu eER siehe z.B. Biskup 1995, S. 51 ff



1.2 Objekttypenmethode (OTM)

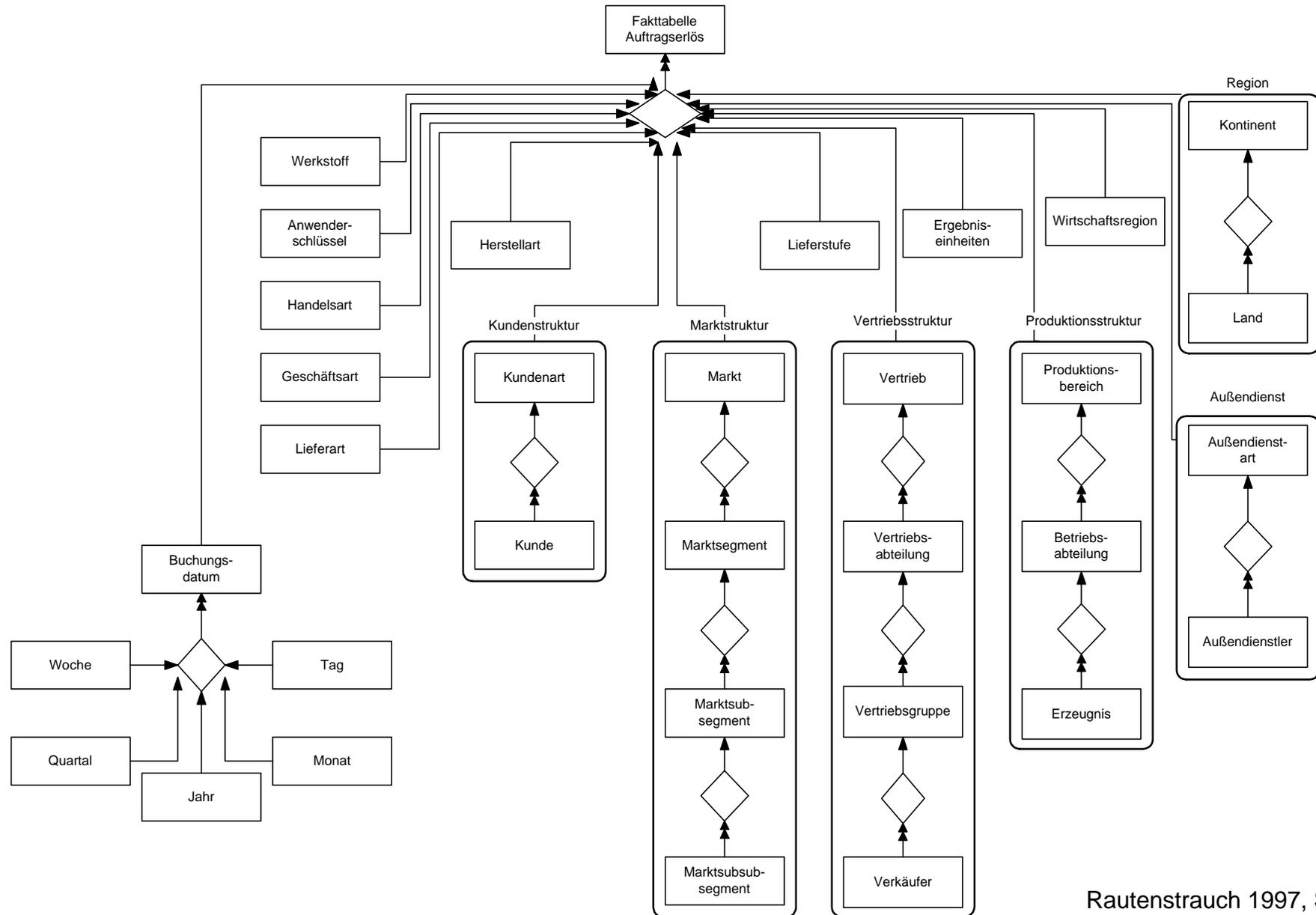
Ausgewählte Notationselemente



Ortner/Söllner 1989, S. 33 ff; Rautenstrauch 1997, S. 7



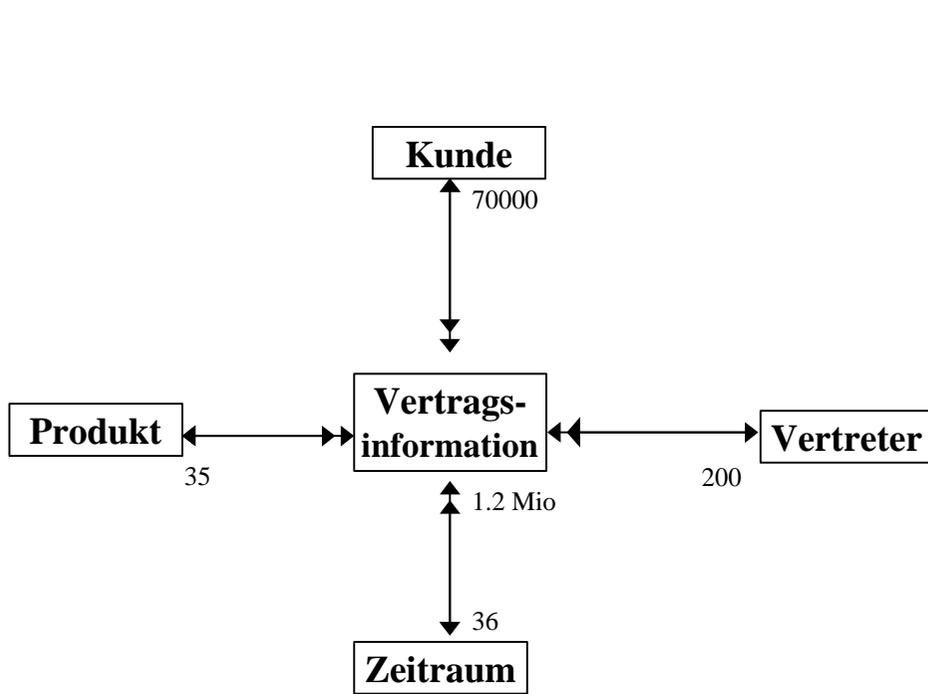
Ein Beispiel mit OTM



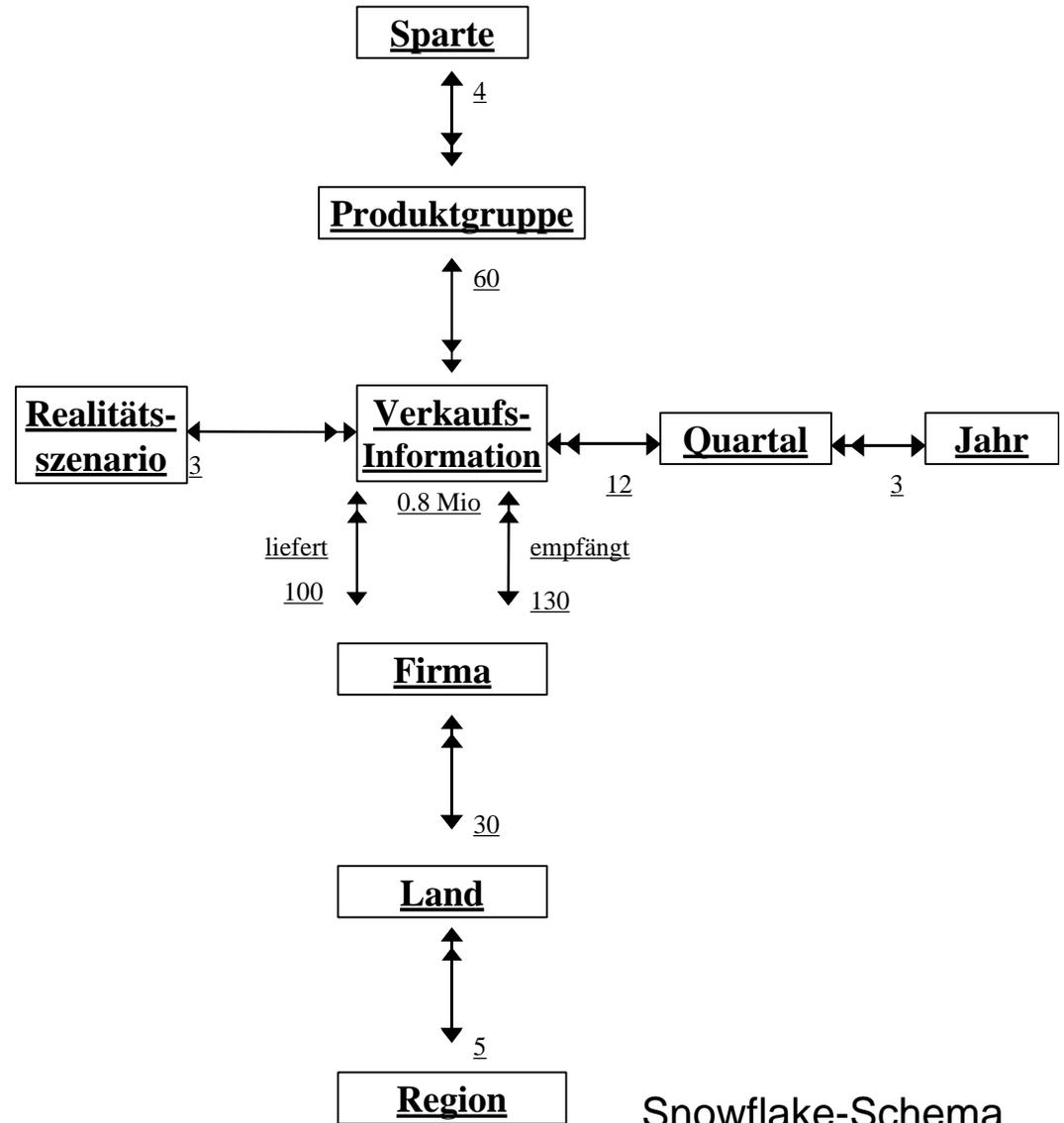
Rautenstrauch 1997, S. 9



1.3 Exkurs: Modellierungsansatz der G.S.E.



Star-Schema



Snowflake-Schema

Altenpohl/Huhn/Schwab/Zeh 1997, S. 18 und 20



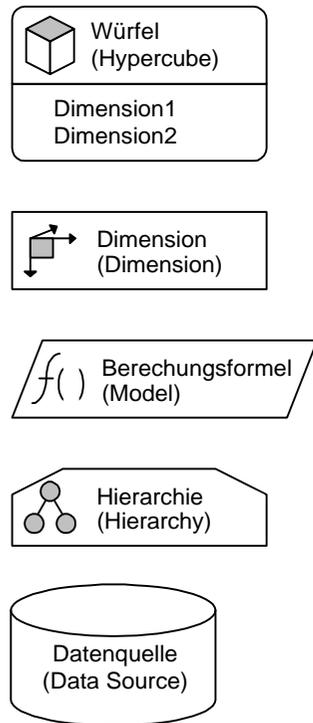
Fazit ERM

- ❑ **Reine erweiterte ER-Modelle sind für die Modellierung multidimensionaler Strukturen ungeeignet. Das eER-Modell muß um weitere Konstrukte ergänzt werden.**
- ❑ **Datensicht allein reicht nicht, daher sollten funktionale und organisatorische Sicht sowie implementierungsnähere Ebenen ergänzt werden.**
- ❑ **Weiterhin müssen sich auch dynamische Aspekte abbilden lassen.**
- ❑ **Vorschlag – Hinzunahme von weiteren Diagrammtypen, wie z.B. in der der Architektur integrierter Informationssysteme (ARIS) vorhanden:**
 - Funktionsbaum
 - Ereignisdiagramm/ereignisgesteuerte Prozeßkette
 - Relationendiagramm (z.B. für ROLAP-Implementierungen oder SQL-Durchgriff)

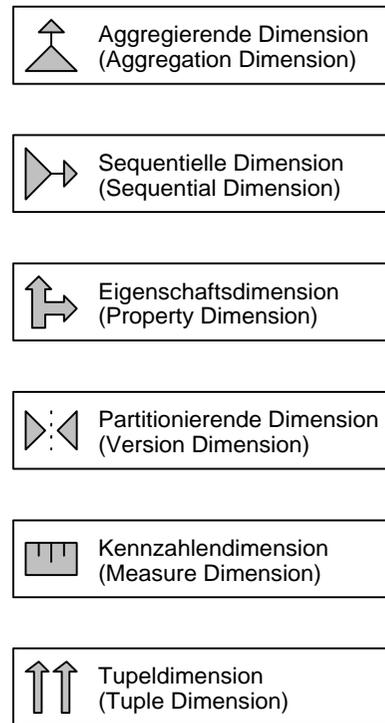


2 Application Design for Analytical Processing Technologies (ADAPT)

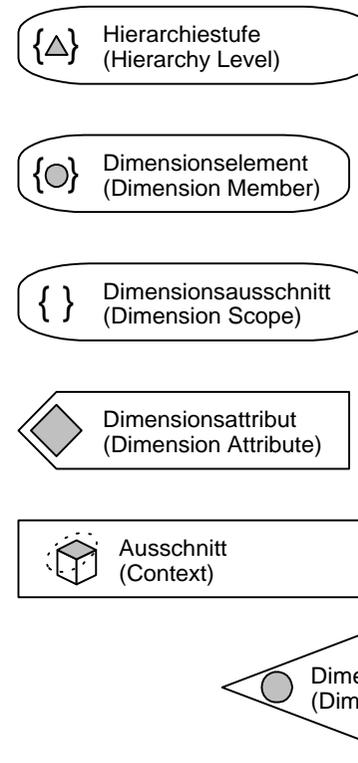
Kernelemente



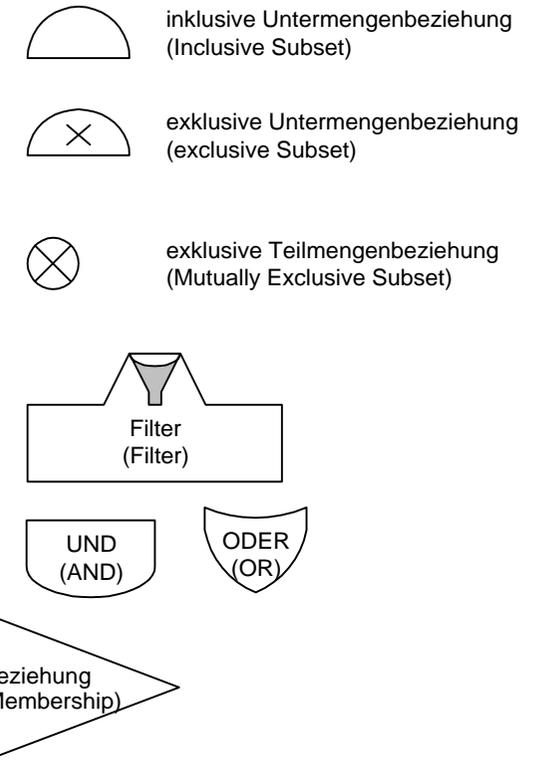
Dimensionstypen



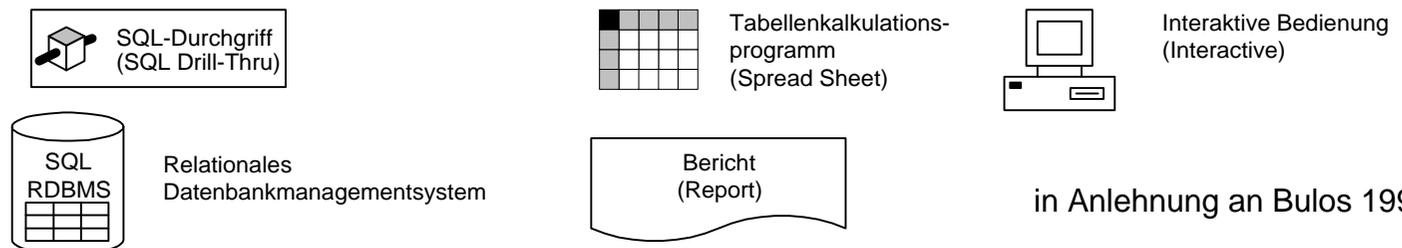
Dimensionselemente



Beziehungstypen



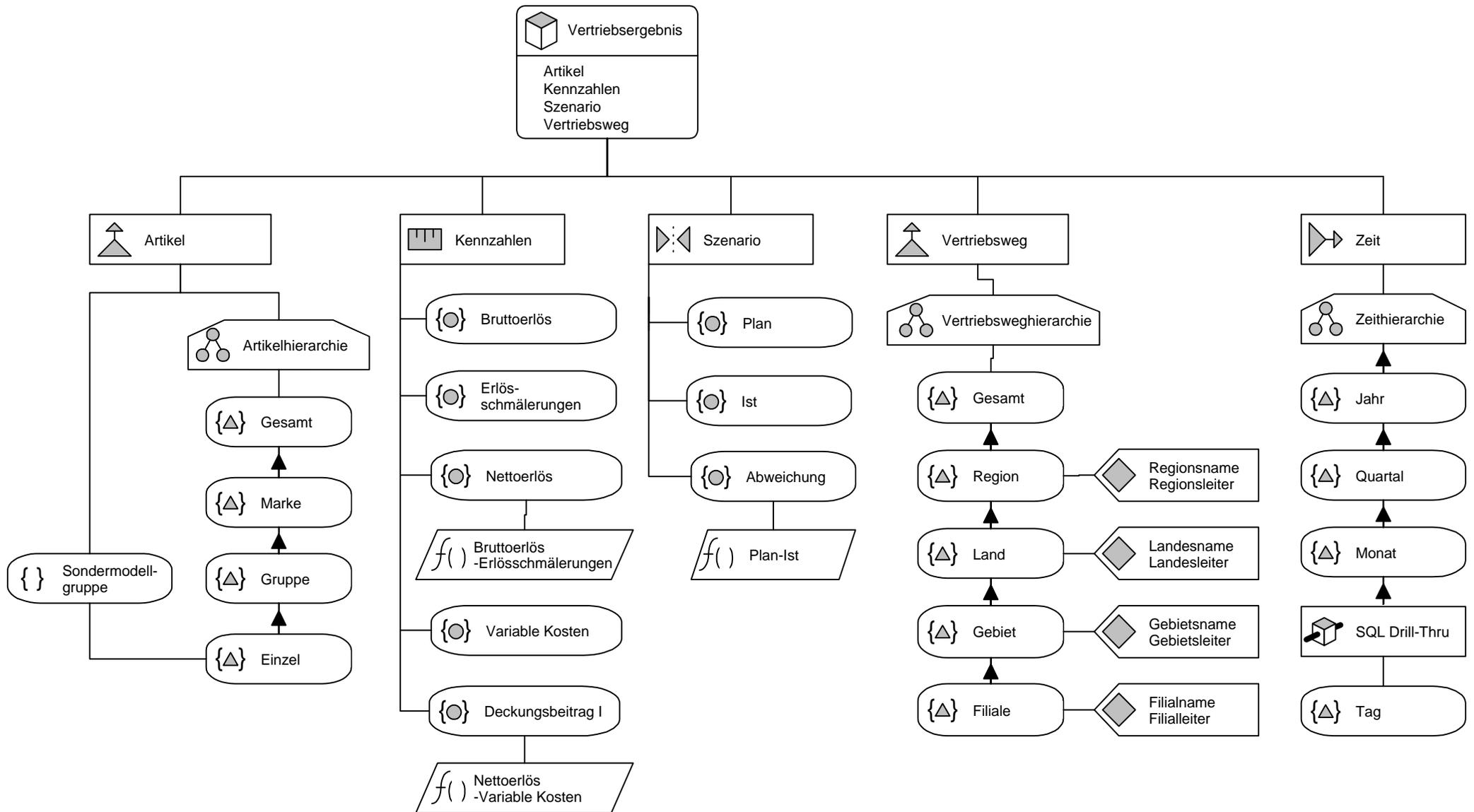
Sonstiges



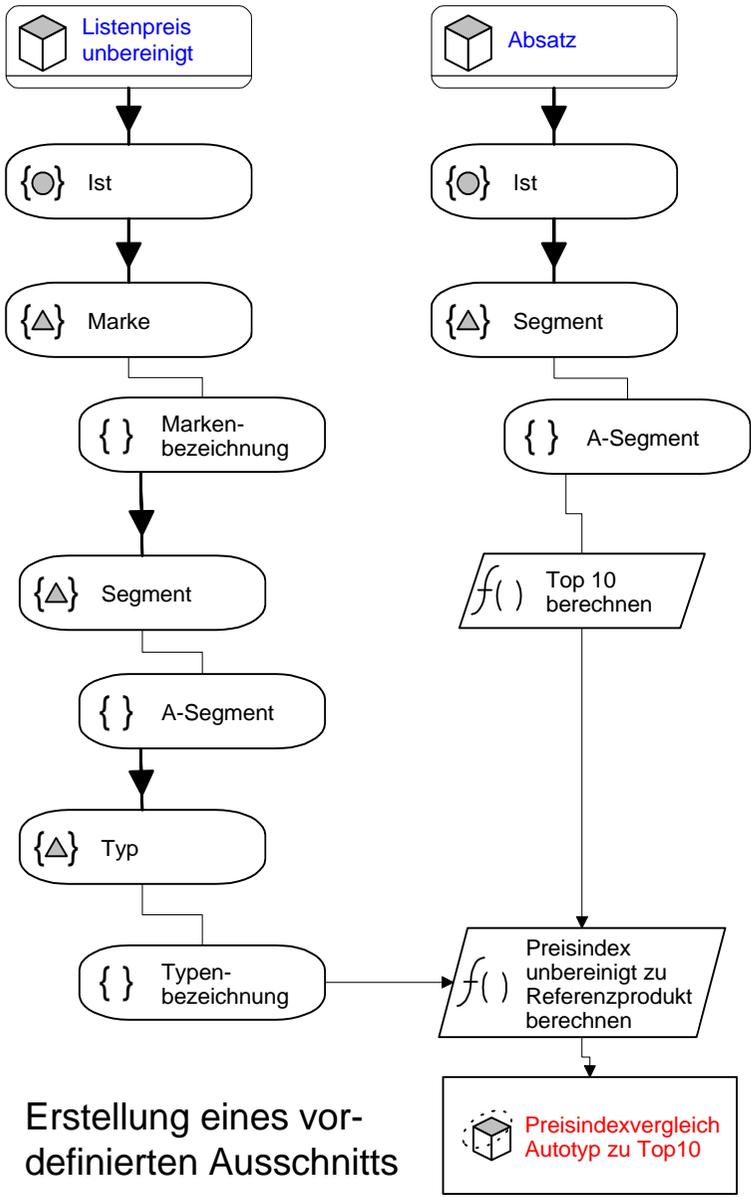
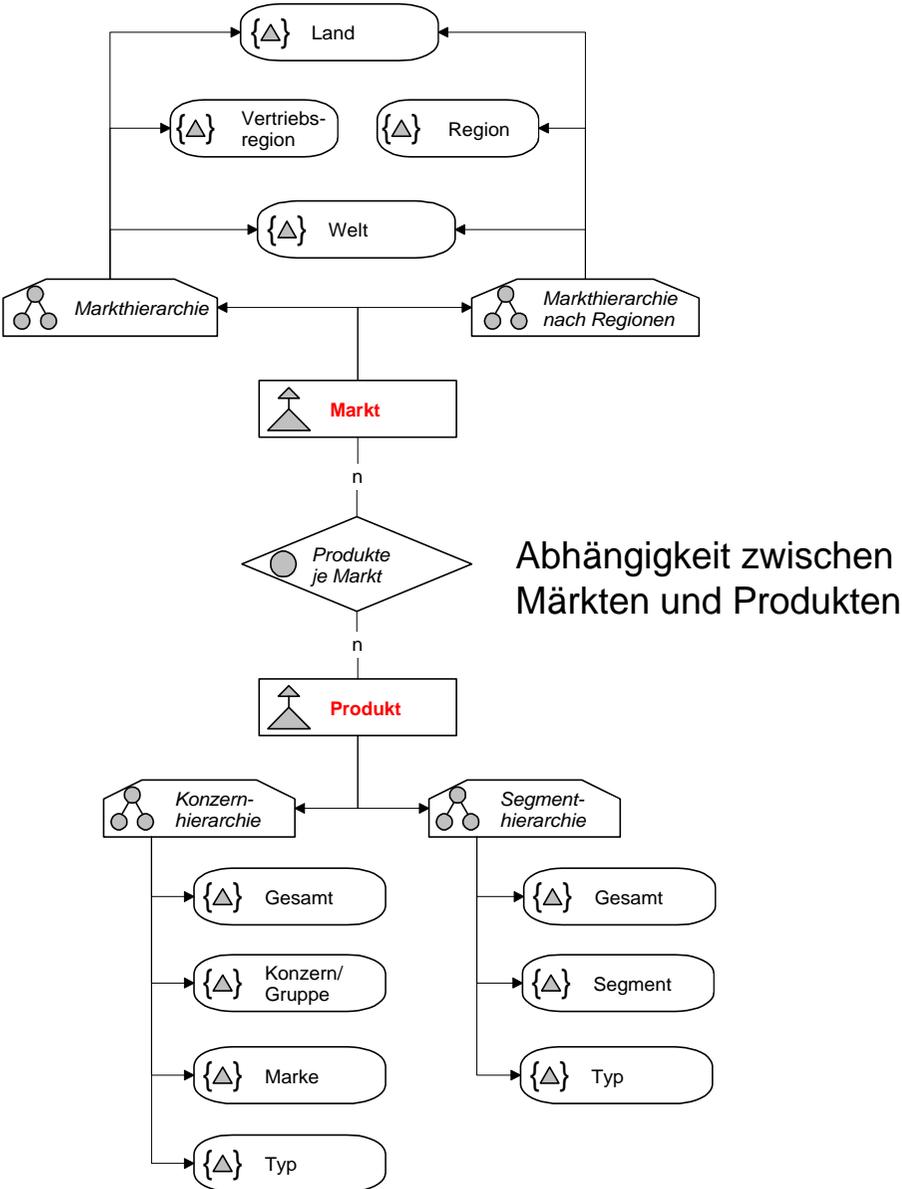
in Anlehnung an Bulos 1996, S. 34



Beispiel mit ADAPT



Beispiel Automobilbranche

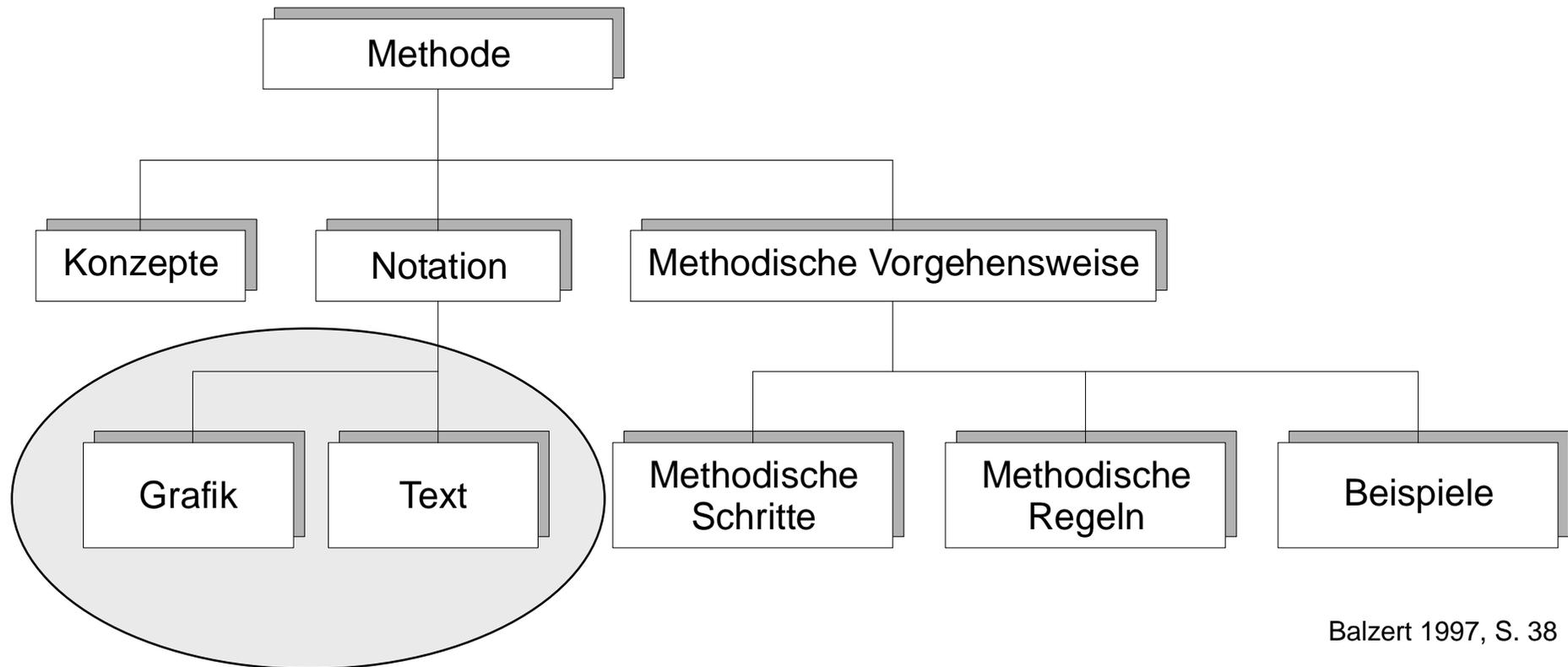


Fazit ADAPT

- ❑ **ADAPT ist eine speziell für den multidimensionalen Entwurf ausgelegte Modellierungsnotation.**
- ❑ **Der Umfang an Modellierungskonstrukten ist sehr groß:**
 - Vorteil: auch für komplexe Sachverhalte gibt es geeignete Darstellungsweisen.
 - Nachteil: Notationsvielfalt muß erlernt und auch korrekt angewendet werden (ähnliches Problem existiert bei ARIS).
- ❑ **Verschiedene Sichten werden in einem Diagrammtyp vermischt.**
 - Bei größeren Projekten sollten Modellierungssichten und -ebenen durch Konvention aufgeteilt werden.
- ❑ **Aus unserer Projekterfahrung bilden ADAPT-Diagramme eine gute Diskussionsgrundlage für Gespräche zwischen Anwendern aus Fachabteilungen und OLAP-Designern.**



3 Objektorientierte Modellierung



Balzert 1997, S. 38

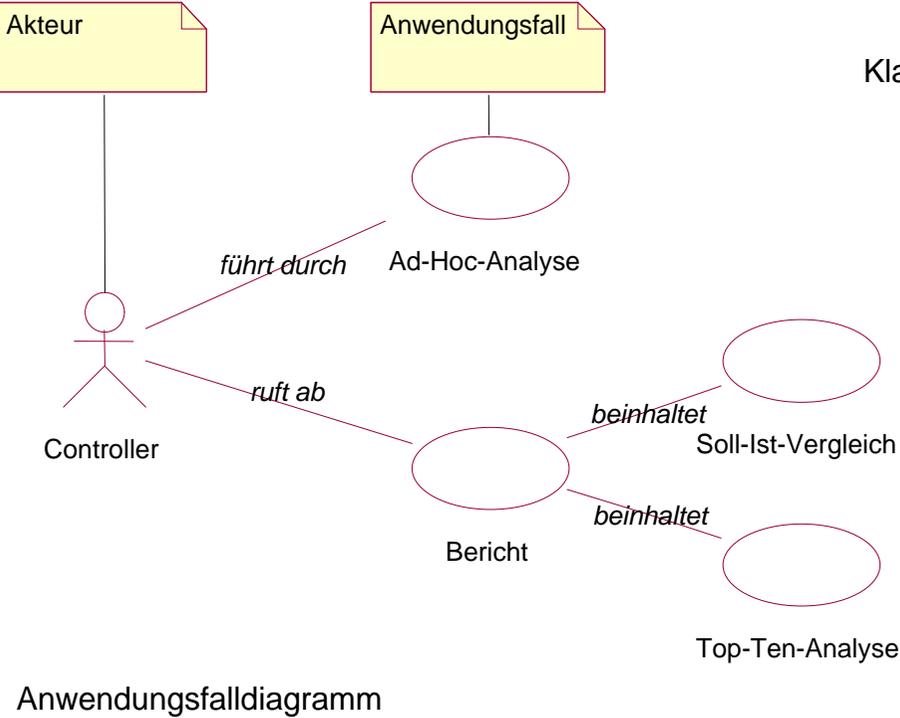
- ❑ **Mit *objektorientierter Modellierung* soll hier der Einsatz einer objektorientierten (grafischen) Notation vorgestellt werden.**

Unified Modelling Language (UML)

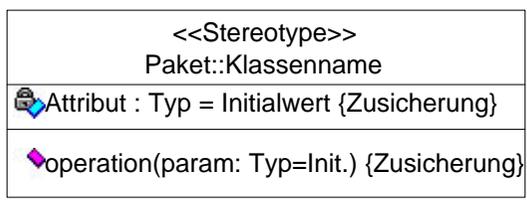
- ❑ **basiert schwerpunktmäßig auf Modellen von Booch, Jacobsen und Rumbaugh**
- ❑ **wurde als Bestandteil der Object Management Architecture (OMA) durch Object Management Group (OMG) aufgenommen**
- ❑ **UML ist Sprache und Notation, aber keine Methode (kann aber Basis für verschiedene Methoden sein)**
- ❑ **umfaßt verschiedene Sichten:**
 - **Anwendungsfall:** beschreibt die Interaktion zwischen Anwender und System
 - **Logisch:** zur Modellierung von statischen und dynamischen Zusammenhängen
 - **Komponenten:** gliedert die Organisation und Zusammenhänge eines komponentenweisen Aufbaus der Implementierung



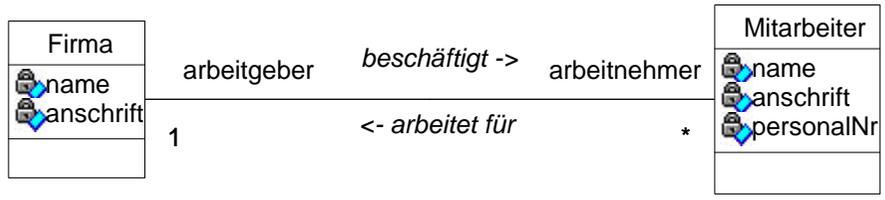
Ausgewählte statische Notationselemente



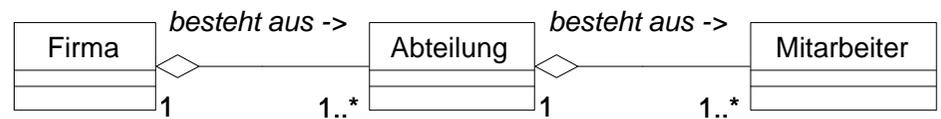
Klassendiagramm



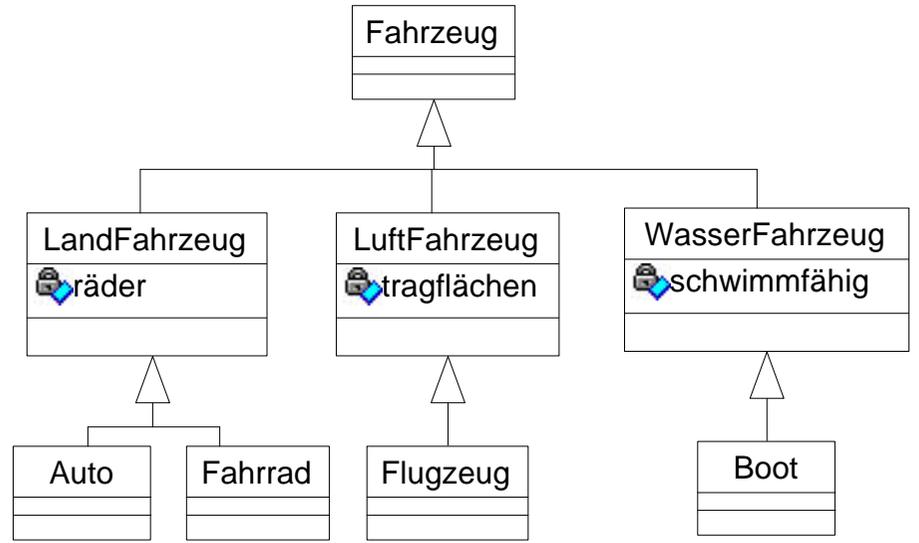
Assoziation



Aggregation

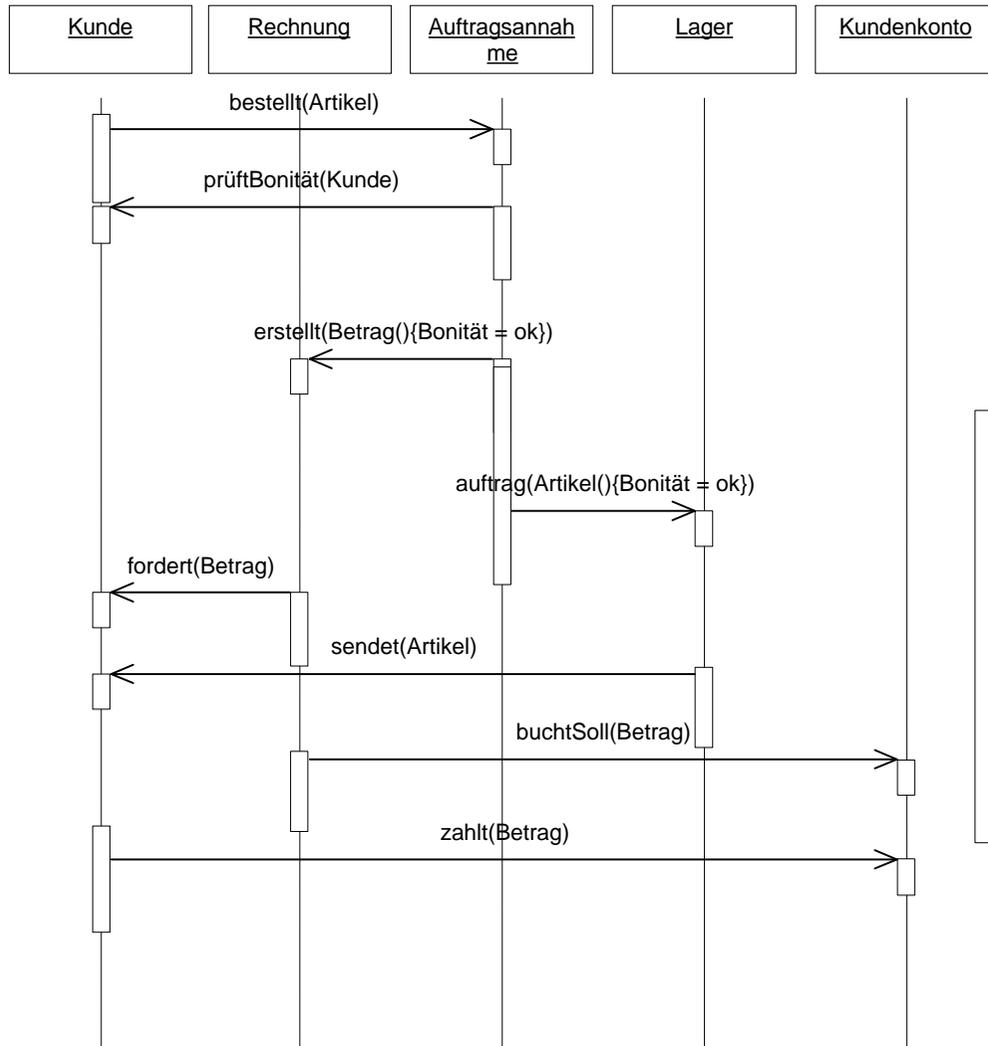


Vererbung

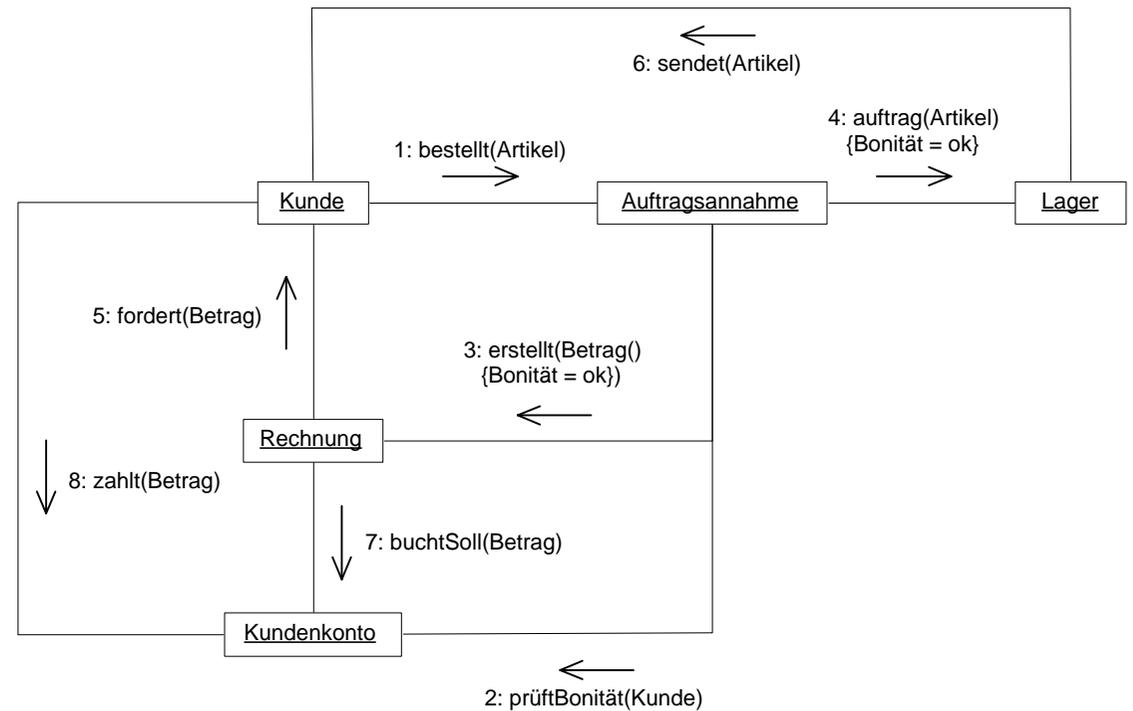


Anwendungsfalldiagramm

Ausgewählte dynamische Notationselemente



Sequenzdiagramm



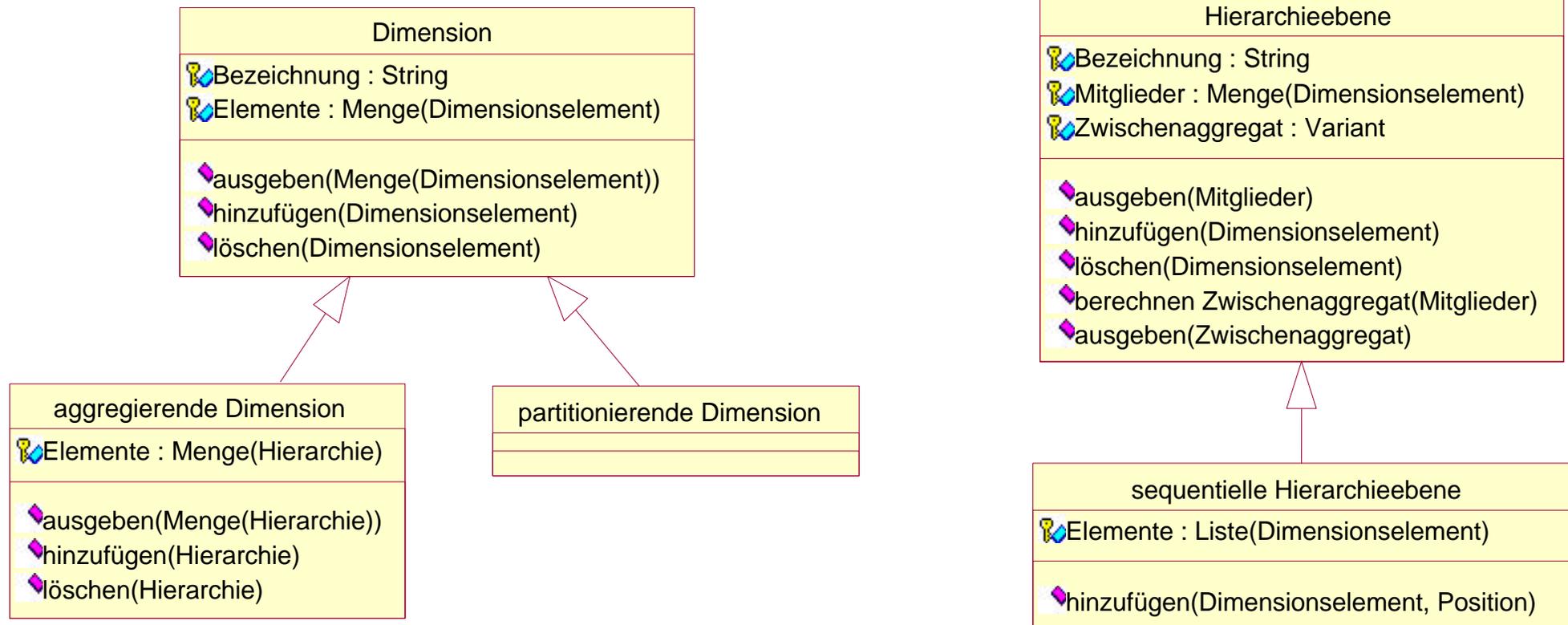
Kollaborationsdiagramm



Notation einer multidimensionalen Datenstruktur

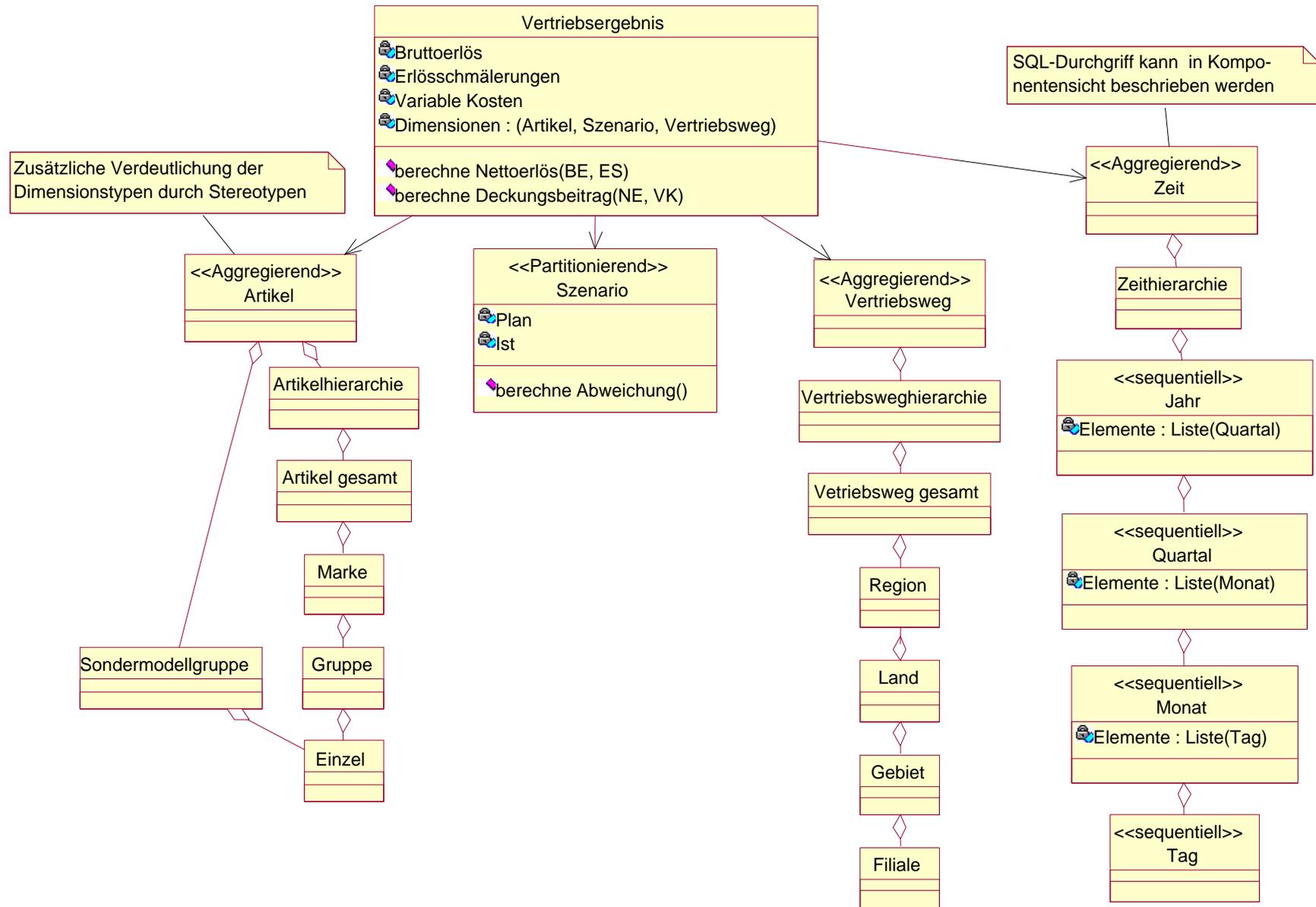


Notation von speziellen Klassen

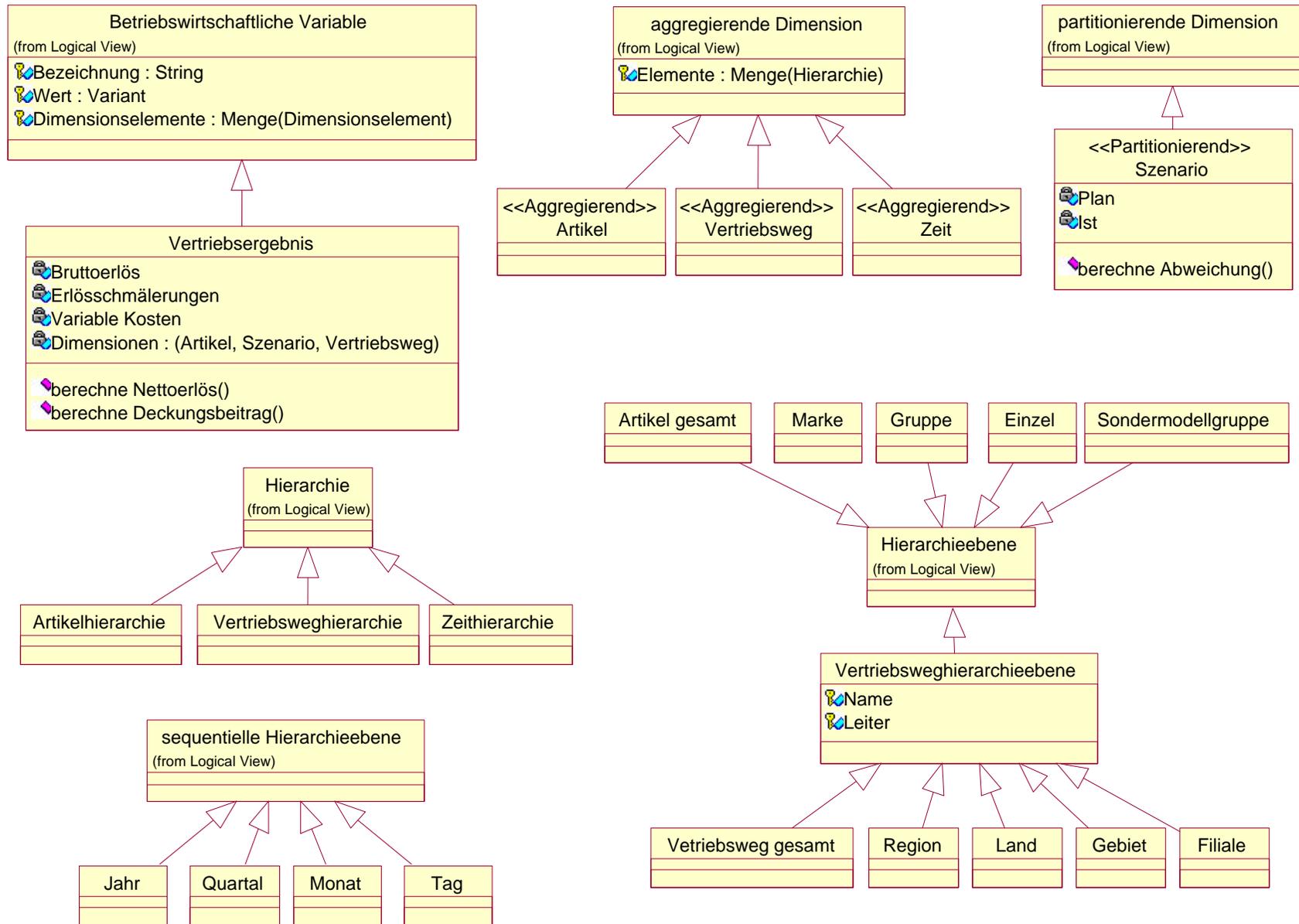


Die Hierarchieebenen der Zeitdimension besitzen *Liste* statt *Menge* als Typ für Elemente, um Reihenfolgebeziehung darzustellen.

Vertriebsergebnisrechnung als Anwendungsfall



Zuordnung der verwendeten Klassen



Fazit Objektorientierte Modellierung

- ❑ Reihenfolgebeziehungen von Elementen lassen sich als Liste darstellen (Benutzung von abstrakten Daten-/Objekttypen in Grafik möglich).
- ❑ Erweiterung der Notation durch Stereotypen möglich.
- ❑ **Es gelten die allgemeinbekanntesten Vorteile der objektorientierten Modellierung, insbesondere**
 - sind Daten und Funktionen eng miteinander verknüpft – betriebswirtschaftliche Variable und Dimensionen lassen sich direkt mit ihren Funktionen verbinden,
 - lassen sich objektorientierte Modelle sehr leicht in objektorientierten oder objektrelationen DBMS implementieren,
 - können Projekte durch Wiederverwendung vereinfacht werden.
- ❑ **Nachteil: Objektorientiertes Denken muß erlernt und auch von Vertretern der Fachabteilung verstanden werden.**
- ❑ **Man sollte sich jetzt auf die UML als neuen Standard einstellen.**



Ausblick

- ❑ Die wirklichen Probleme treten erst bei der praktischen Modellierung und anschließenden Implementierung auf:
 - Schwierig zu modellieren sind sich im Zeitablauf verändernde Dimensionen (slowly changing Dimensions), wie die Neuorganisation von Artikelgruppen oder von Verantwortungsbereichen (Profit Center).
 - Teilweise lassen sich modellierte Eigenschaften mit OLAP-Produkten gar nicht implementieren:
 - ◆ umfangreiche zusätzliche Attribute können nicht geführt werden
 - ◆ bestimmte Variable lassen sich nicht gleichzeitig in einer Auswertung anzeigen, da sie unterschiedliche Dimensionen besitzen bzw. wenn sie sich anzeigen lassen, dann nur mit gleichen Dimensionen
- ❑ Bisher konnte sich noch keine Notation für die multidimensionale Modellierung dominierend durchsetzen. Die vorgestellten Notationen sind mehr oder weniger gut geeignet:
 - ER-Modelle lassen sich mit Hilfe von geeigneten Ergänzungen einsetzen.
 - ADAPT ist eine umfangreiche, einfach zu erlernende Notation aus der Praxis der Unternehmensberatung.
 - Die Eignung von objektorientierten Konstrukten für die multidimensionale Modellierung ist bisher noch unzureichend untersucht worden und einer weiteren Betrachtung.



Quellen und Literaturhinweise

- **Altenpohl, U.; Huhn, M.; Schwab, W.; Zeh, T. (1997):**
Datenmodellierung Data Warehouse – ein Lösungsvorschlag mittels ER-Modellierung, Interner Bericht der G.S.E. UAG Rhein-Main 1997
- **Balzert, H. (1997):**
Wie erstellt man ein objektorientiertes Analysemodell?, in: Informatik-Spektrum, 1/1997, S. 38-47
- **Bulos, D. (1996):**
A New Dimension, in: Database Programming & Design, 6/1996, S. 33-37
- **Biskup, J. (1995):**
Grundlagen von Informationssystemen, Braunschweig Wiesbaden (Vieweg) 1995
- **Chen, P. P. (1976):**
The entity-relationship-model – towards a unified view of data, in ACM Transactions on Database Systems, 1/1976, S. 9-36
- **Fowler, M. (1997):**
Analysis Patterns - Reusable Object Models, Massachusetts (Addison-Wesley) 1997
- **Gabriel, R.; Gluchowski, P. (1997):**
Semantische Modellierungstechniken für multidimensionale Datenstrukturen, in: HMD – Theorie und Praxis der Wirtschaftsinformatik, 195/1997, S. 18-37
- **Hars, A. (1994):**
Referenzdatenmodelle – Grundlagen effizienter Datenmodellierung, Wiesbaden (Gabler) 1994
- **Holthuis, J. (1997):**
Modellierung multidimensionaler Daten - Modellierungsaspekte und Strukturkomponenten, Arbeitsberichte des Lehrstuhls für Informationsmanagement und Datenbanken, European Business School, Oestrich-Winkel 1997
- **Kimball, R. (1997):**
A Dimensional Modeling Manifesto – Drawing The Line Between Dimensional Modeling and ER Modeling Techniques, in: DBMS Online, 8/1997, <http://www.dbmsmag.com/9708d15.html>, 17.10.1997
- **Maier, R. (1998):**
Nutzen und Qualität der Datenmodellierung – Ergebnisse einer empirischen Studie, in: Wirtschaftsinformatik, 2/1998, S. 130-149
- **Oestereich, B. (1997):**
Objektorientierte Softwareentwicklung: Analyse und Design - mit der Unified Modeling Language, 3. Aufl., München Wien (Oldenbourg) 1997
- **Ortner, E.; Söllner, B. (1989):**
Semantische Datenmodellierung nach der Objekttypenmethode, in: Informatik-Spektrum, 12/1989, S. 31-42
- **Rautenstrauch, C. (1997):**
Modellierung und Implementierung von Data-Warehouse-Systemen, Arbeitspapier, Otto-von-Guericke-Universität, Magdeburg, 1997
- **Scheer, A.-W. (1998):**
ARIS - Modellierungsmethoden, Metamodelle, Anwendungen, 3. Aufl., Berlin et al. (Springer) 1998
- **Schelp, J. (1998):**
Konzeptionelle Modellierung mehrdimensionaler Datenstrukturen, in: Chameni, P.; Gluchowski, P. (Hrsg.): Analytische Informationssysteme, Berlin et al. (Springer) 1998, S. 263-276
- **Totok, A. (1997):**
Data Warehouse und OLAP als Basis für betriebliche Informationssysteme, Arbeitsbericht Nr. 97/03, Institut für Wirtschaftswissenschaften, Technische Universität Braunschweig, 1997

