21st International Conference on Conceptual Modeling (ER 2002), p. 199-213: Lecture Notes in Computer Science 2503, Tampere (Finland), October 7-11 2002.

Multidimensional Modeling with UML Package Diagrams

Sergio Luján-Mora
Juan C. Trujillo

Il-Yeol Song

Universitat d'Alacant
Universidad de Alicante

Drexel

Contents

Contents

- Introduction
- Multidimensional Modeling by UML
- Package Design Guidelines
- A Case Study
- Package Stereotypes
- MD Modeling in Rational Rose
- Conclusions and Future Work

Introduction

Different conceptual approaches for DW

- Data warehouses, MD databases, OLAP applications
  - Multidimensional (MD) modeling
- Different approaches for conceptual modeling (graphical approaches):
  - Golfarelli *et al.*
  - Husemann *et al.* → No all properties considered
  - Sapia *et al.*
  - Tryfona *et al.* → Own graphical notations
  - Abello *et al.* ↓
  - … Learn a new notation

Introduction

Why UML ?

- UML → Standard OO modeling language
  - Minimize the efforts in learning new notations
- Extensible language
  - Stereotypes, tagged values, constraints → Profile
- OCL: well-formedness rules of the new defined elements
  - Avoids an arbitrary use of our extension
- Extensions can be programmed in Rational Rose
- Our conceptual MD modeling approach is based on the UML *(Trujillo et al., 2001)*

Introduction

A problem → Large and complex DW's

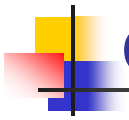
- Modeling complex and large data warehouses
  - MD model can become very complex to understand
- Other approaches use flat diagrams
  - Not suitable for designing complex and large DW's
  - Cluttered diagrams difficult to read

Introduction

A problem → Large and complex DW's

- UML grouping mechanism → Packages
  - Different levels of abstraction
  - Not restricted to flat diagrams
- However, UML does not provide guidelines
- We provide our own Guidelines for MD modeling
  - Correct and natural use for MD modeling

Contents



Contents

- Introduction
- Multidimensional Modeling by UML
- Package Design Guidelines
- A Case Study
- Package Stereotypes
- MD Modeling in Rational Rose
- Conclusions and Future Work

MD Modeling by UML

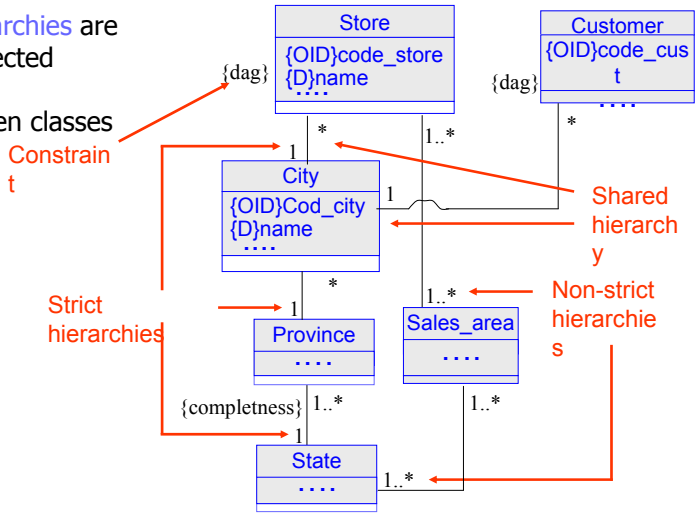
- Structural and dynamic properties of MD modeling by using the UML
  - Structural properties → UML class diagram
  - Dynamic properties → state and interaction diagrams
- Main Structural MD properties:
  - Many-to-many relationships fact – dimensions
  - Degenerate dimensions
  - Derived measures, Additivity
  - All kinds of classification hierarchies
  - Categorization of dimensions, etc.

MD Modeling by UML

Dimensions and classification hierarchies

Dimensions → dimension classes

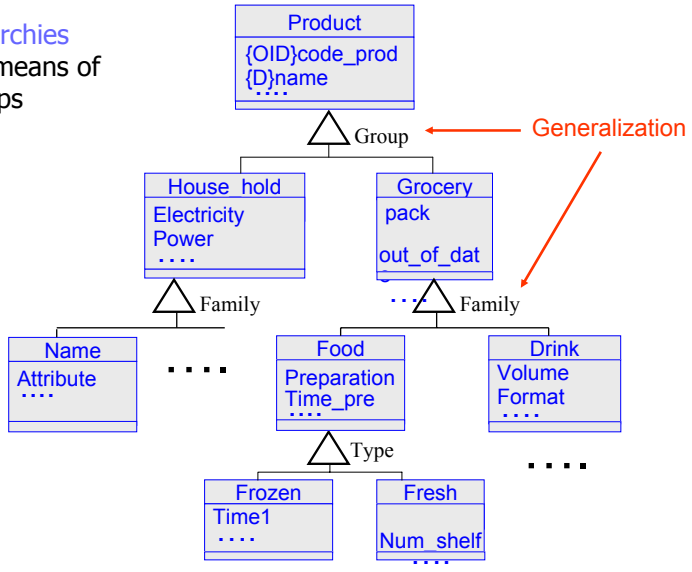
Classification hierarchies are considered by Directed Acyclic Graphs of association between classes



MD Modeling by UML

Dimensions and specialization hierarchies

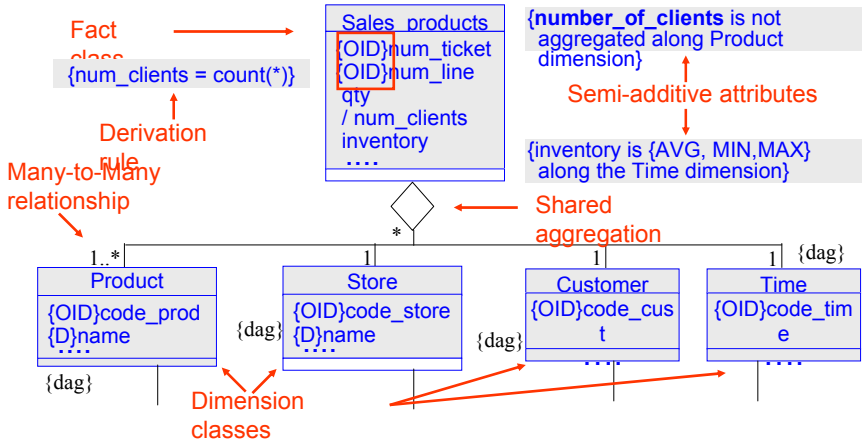
Specialization hierarchies are considered by means of subtype relationships



MD Modeling by UML

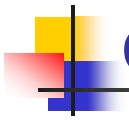
Facts

Fact classes are considered as a shared aggregation relationship between dimension classes



Contents

- Introduction
- Multidimensional Modeling by UML
- Package Design Guidelines
- A Case Study
- Package Stereotypes
- MD Modeling in Rational Rose
- Conclusions and Future Work



Contents

- Introduction
- Multidimensional Modeling by UML
- Package Design Guidelines
- A Case Study
- Package Stereotypes
- MD Modeling in Rational Rose
- Conclusions and Future Work

Package Design Guidelines

- Based on our experience in real-world cases
  - These guidelines express the natural way of understanding MD modeling
- 14 guidelines (G)
  - A very simple yet powerful design of MD models

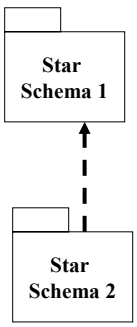
Package Design Guidelines

- (G0a) The design process is divided into three levels
  - Level 1: Model definition
  - Level 2: Star schema definition
  - Level 3: Dimension/fact definition
- (G0b) Define fact/dimensions and the shared dimensions

Package Design Guidelines

Level 1: Model definition

- (G1) Package → Star schema (i.e. fact considered)
- (G2) Dependency → Share at least one dimension
  - (G2a) Decide star schemas hosting shared dimensions
  - (G2b) Group them together → Minimize dependencies

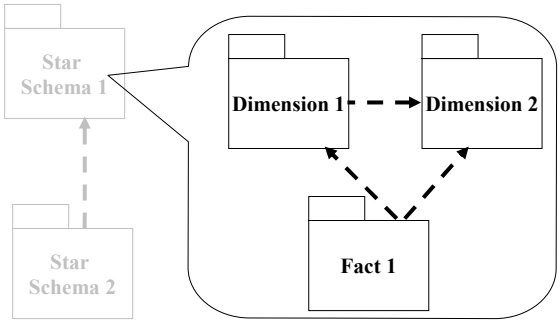


Level 1

Package Design Guidelines

Level 2: Star Schema definition

- (G3) Package → Fact and dimension
- (G4) Dependency
  - (G4a) From the fact package to each dim package
  - (G5) Do not define a dimension twice: import it
  - (G6) Btn dim packages → sharing hierarchy levels

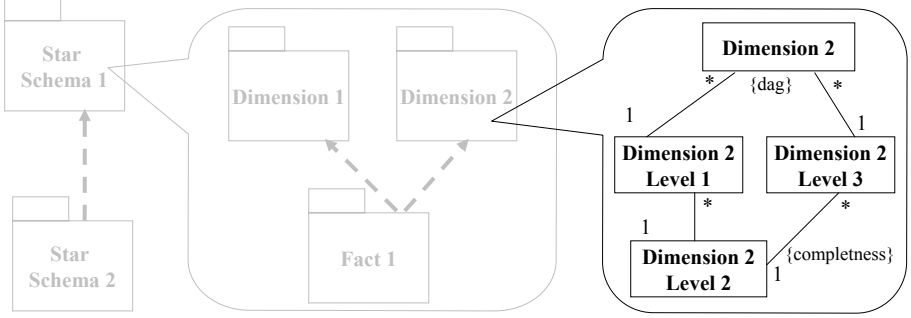


Level 1 **Level 2**

Package Design Guidelines

Level 3: Dimension/Fact definition

- (G7) Dim package → define dim and base classes
- (G8) Fact package → define the fact class and import the dimension classes
- (G9) In a dimension package, import shared levels
- (G10) Import only the required levels

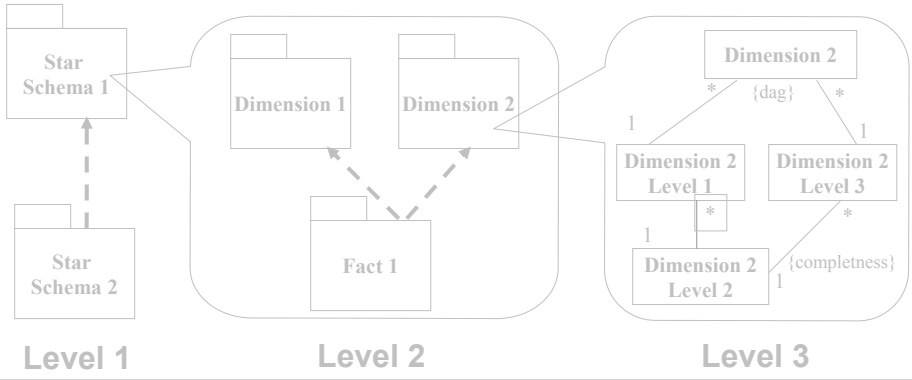


Level 1 Level 2 **Level 3**

Package Design Guidelines

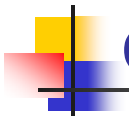
Level 3: Dimension/Fact definition

- (G7) Dim package → define dim and base classes
- (G8) Fact package → define the fact class and import the dimension classes
- (G9) In a dimension package, import shared levels
- (G10) Import only the required levels



Contents

- Introduction
- Multidimensional Modeling by UML
- Package Design Guidelines
- A Case Study
- Package Stereotypes
- MD Modeling in Rational Rose
- Conclusions and Future Work



Contents

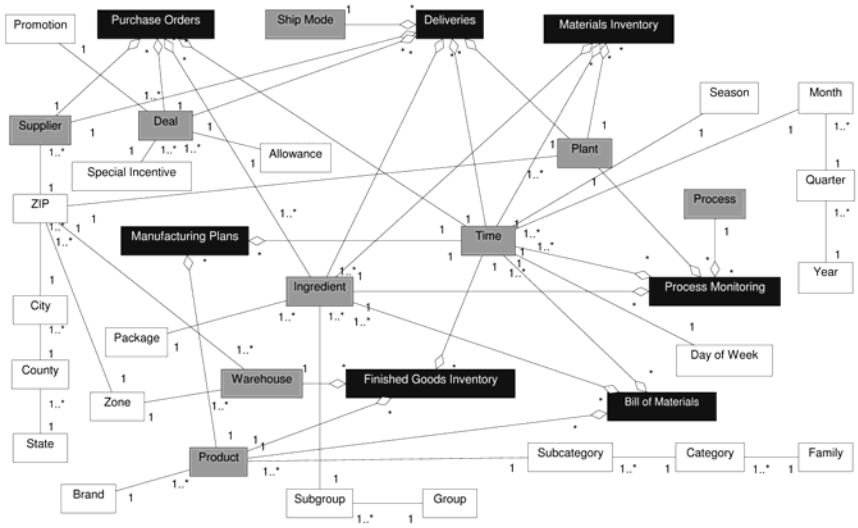
- Introduction
- Multidimensional Modeling by UML
- Package Design Guidelines
- A Case Study
- Package Stereotypes
- MD Modeling in Rational Rose
- Conclusions and Future Work

A Case Study

- The *supply value chain* (chapter 5 of *"The data warehouse toolkit"*, Ralph Kimball)
  - "The supply side of the business consists of the steps needed to manufacture the products from original ingredients or parts…"
- Includes:
  - 7 facts
  - 9 dimensions

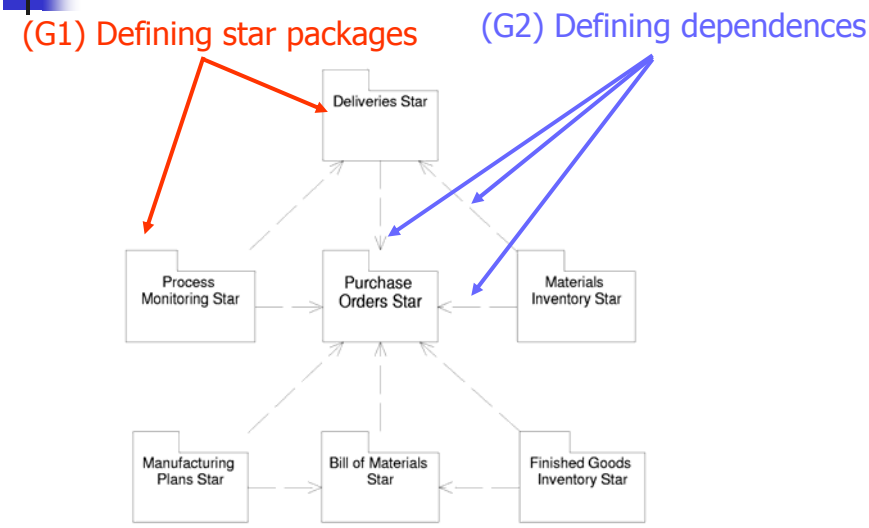
A Case Study

An example of a cluttered diagram for a complex DW



A Case Study

Level 1: Model definition

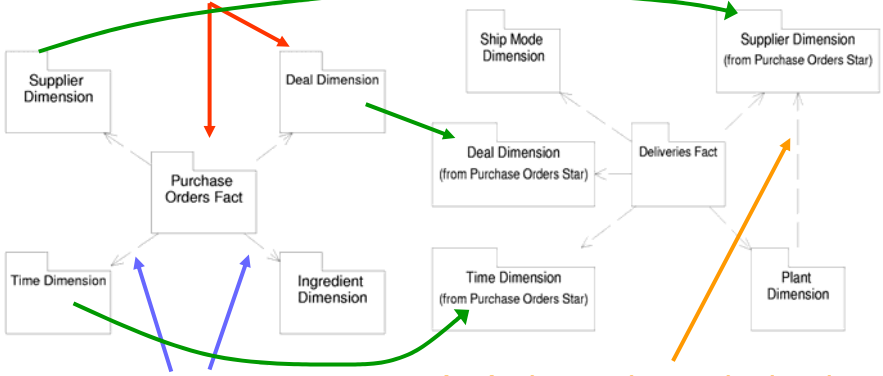


A Case Study

Level 2: Star schema definition: Purchase Orders

(G3) Defining fact and dimension packages

(G5) Importing dimensions



(G4) Defining dependences

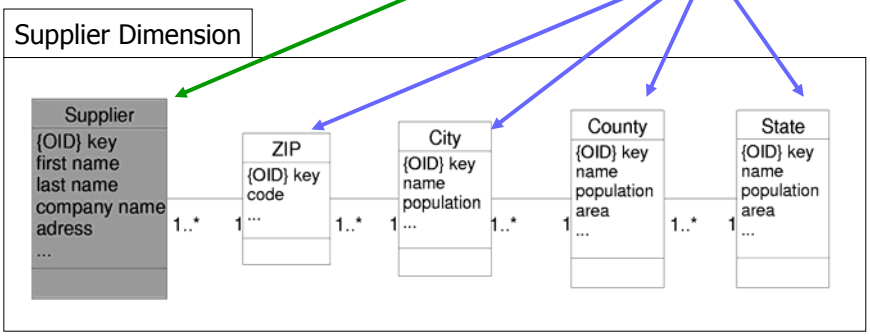
(G6) Sharing hierarchy levels

A Case Study

Level 3: Dimension/Fact definition: Supplier

- Defining dimensions

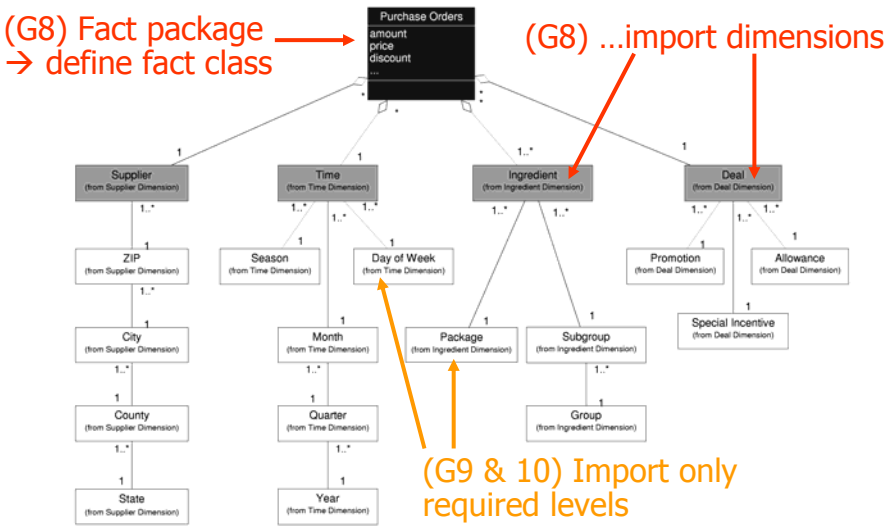
(G7) In a Dim package → define dimension and base classes



A Case Study

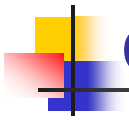
Level 3: Dimension/Fact definition: Supplier

- Defining facts



Contents

- Introduction
- Multidimensional Modeling by UML
- Package Design Guidelines
- A Case Study
- Package Stereotypes
- MD Modeling in Rational Rose
- Conclusions and Future Work



Contents

- Introduction
- Multidimensional Modeling by UML
- Package Design Guidelines
- A Case Study
- Package Stereotypes
- MD Modeling in Rational Rose
- Conclusions and Future Work

Package Stereotypes

- UML can be formally extended to be adapted to a specific method, organization, or user
  - Extension is still UML standard
- UML Extensibility Mechanisms
  - Stereotypes
  - Tagged values
  - Constraints

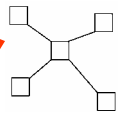
Package Stereotypes

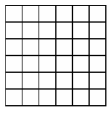
- **Stereotype**:
  - A new model element that specializes a UML element (Class, Attribute, Package, Association, etc.)
- **Tagged value**:
  - A new property of a model element
- **Constraint**:
  - Refines the semantics of a model element → Informal or formal (Object Constraint Language, OCL)

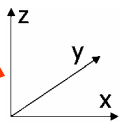
Package Stereotypes

Our package extension

- Extension summary:
  - 3 stereotypes:
    - **StarPackage**
    - **FactPackage**
    - **DimensionPackage**
  - 8 constraints



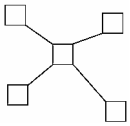




Package Stereotypes

An example of a package stereotype definition

- Name: **StarPackage**
- Base class: **Package**
- Description: **Packages of this stereotype represent MD star schemas**
- Icon:



- Tagged values: **None**

Package Stereotypes

- Constraints:
  - A StarPackage can only contain FactPackages or DimensionPackages:

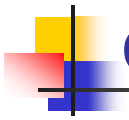
    self.contents->forAll(oclIsTypeOf(FactPackage) or oclIsTypeOf(DimensionPackage))

  - A StarPackage can only contain one FactPackage:

    self.contents->select(oclIsTypeOf(FactPackage))->size <= 1

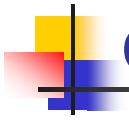
  - There are no cycles in the dependency structure:

    not self.allSuppliers->includes(self)



Contents

- Introduction
- Multidimensional Modeling by UML
- Package Design Guidelines
- A Case Study
- Package Stereotypes
- MD Modeling in Rational Rose
- Conclusions and Future Work



Contents

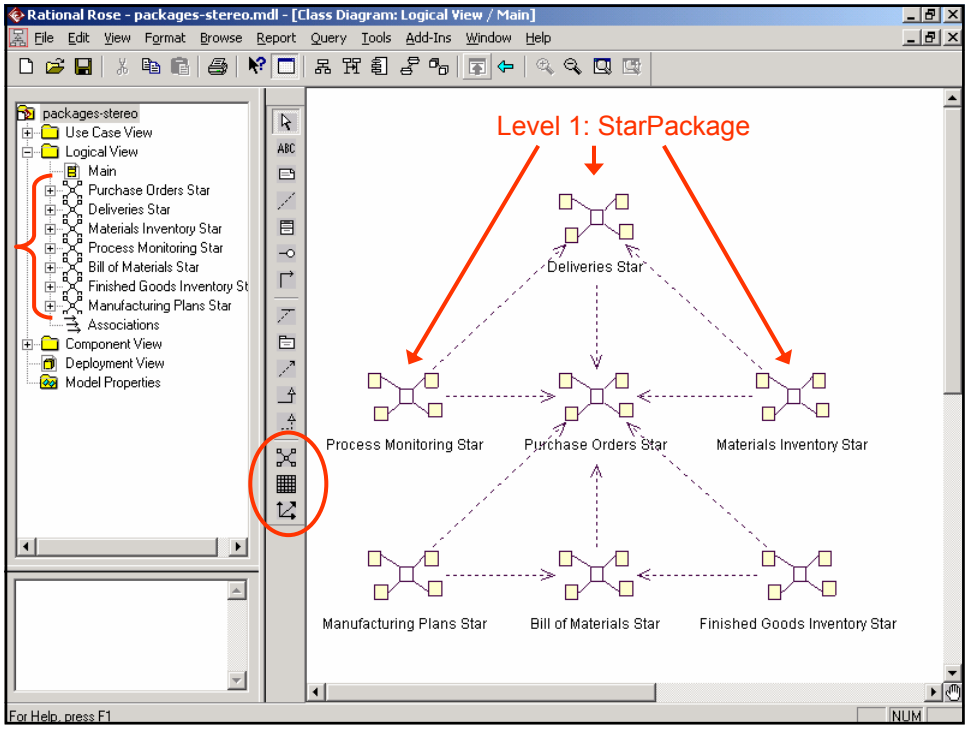
- Introduction
- Multidimensional Modeling by UML
- Package Design Guidelines
- A Case Study
- Package Stereotypes
- MD Modeling in Rational Rose
- Conclusions and Future Work

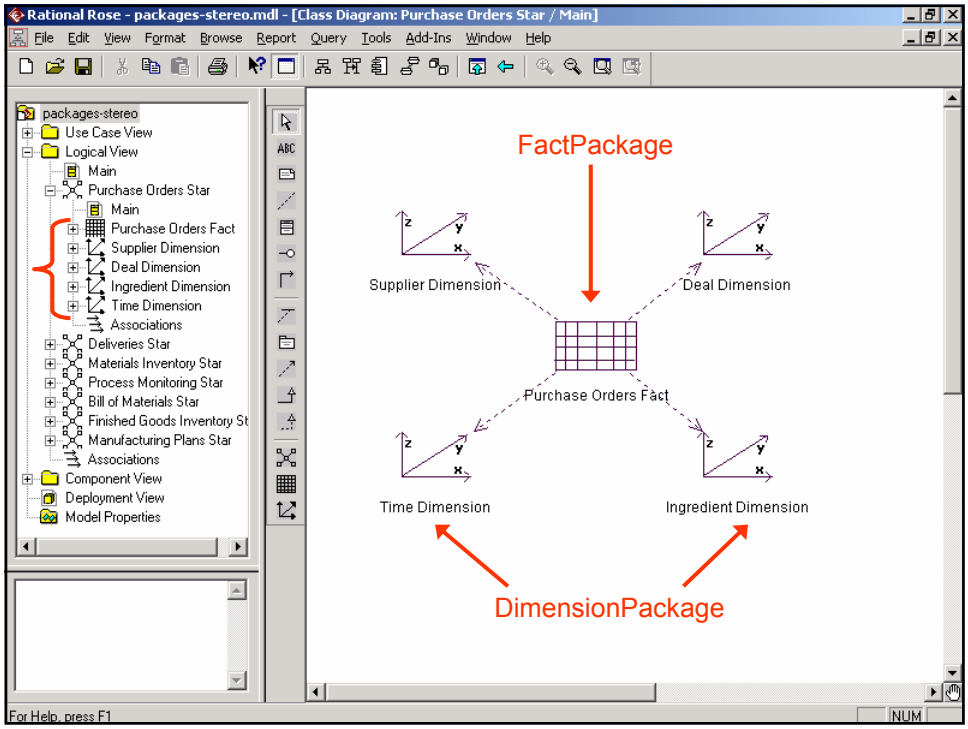
MD Modeling in Rational Rose

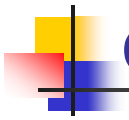
- Rational Rose (RR) is one of the most well-known visual modeling tools
- RR is extensible by means of add-ins through the Rose Extensibility Interface:
  - Main menu items
  - Stereotypes
  - Properties (*tagged values*)
  - Data types
  - Event handling
  - Scripts
  - …

MD Modeling in Rational Rose

- Our add-in customizes:
  - Stereotypes
    - Stereotype configuration file
  - Constraints
    - The Rose Extensibility Interface does not allow us to directly define new OCL constraints
    - Menu item *MD Validate* in the Menu configuration file
      - Runs a Rose script that validates a MD model checking all defined constraints

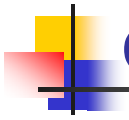

Rational Rose - packages-stereo.mdl - [Class Diagram: Logical View / Main]
File Edit View Format Browse Report Query Tools Add-Ins Window Help
packages-stereo
Use Case View
Logical View
Main
Purchase Orders Star
Deliveries Star
Materials Inventory Star
Process Monitoring Star
Bill of Materials Star
Finished Goods Inventory St
Manufacturing Plans Star
Associations
Component View
Deployment View
Model Properties
Level 1: StarPackage
Deliveries Star
Process Monitoring Star
Purchase Orders Star
Materials Inventory Star
Manufacturing Plans Star
Bill of Materials Star
Finished Goods Inventory Star
For Help, press F1
NUM


Rational Rose - packages-stereo.mdl - [Class Diagram: Purchase Orders Star / Main]
File Edit View Format Browse Report Query Tools Add-Ins Window Help
packages-stereo
Use Case View
Logical View
Main
Purchase Orders Star
Main
Purchase Orders Fact
Supplier Dimension
Deal Dimension
Ingredient Dimension
Time Dimension
Associations
Deliveries Star
Materials Inventory Star
Process Monitoring Star
Bill of Materials Star
Finished Goods Inventory St
Manufacturing Plans Star
Associations
Component View
Deployment View
Model Properties
FactPackage
Supplier Dimension
Deal Dimension
Purchase Orders Fact
Time Dimension
Ingredient Dimension
DimensionPackage
For Help, press F1
NUM



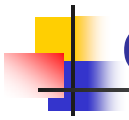
Contents

- Introduction
- Multidimensional Modeling by UML
- Package Design Guidelines
- A Case Study
- Package Stereotypes
- MD Modeling in Rational Rose
- Conclusions and Future Work



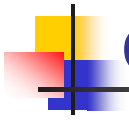
Contents

- Introduction
- Multidimensional Modeling by UML
- Package Design Guidelines
- A Case Study
- Package Stereotypes
- MD Modeling in Rational Rose
- Conclusions and Future Work



Conclusions and Future Work

- MD modeling approaches use "flat design"
  - Not suitable for huge and complex MD models
- In this paper…
  - UML packages → MD modeling at three levels
    - Level 1: Model definition
    - Level 2: Star schema definition
    - Level 3: Dimension/Fact definition
- 14 design guidelines



Conclusions and Future Work

- UML extension for MD modeling
  - StarPackage
  - FactPackage
  - DimensionPackage
- OCL to specify the constraints, avoiding an arbitrary use of the extension
- Rational Rose add-in
  - This extension will be available on the web http://gplsi.dlsi.ua.es/gplsi/areasf.htm

Conclusions and Future Work

- Future work:
  - Providing a UML profile for our approach (including packages). *UML'02*
  - Automatic generation of database schema into object-oriented and object-relational databases
  - Extension allowing OLAP operations on the Web
  - Methodology

Multidimensional Modeling with UML Package Diagrams

Sergio Luján-Mora
Juan C. Trujillo

Il-Yeol Song

Universitat d'Alacant
Universidad de Alicante

Drexel