DATA PROCESSING MODELING IN DECISION SUPPORT SYSTEMS

Escuela Técnica Superior de Ingeniería de Sistemas Informáticos
Madrid Politechnyc University

Concepción M. Gascueña, Rafael Guadalupe
cmgascue@eui.upm.es, r.guadalupe@topografia.upm.es

Iceis2012, Wroclaw. Poland

Contents

Introduction

- Introduction
- **Motivation**
- Objectives
- Proposal
- Example
- Conclusions
- Future Work

Due to the advancement of both, information technology in general, and databases in particular; data stored devices are becoming cheaper and data processing speed is increasing.

As result of this, organizations tend to store large volumes of data with holding great potential information.

Decision Support Systems (DSS) try to use the stored data to obtain valuable information for organizations.

They try to extract the information **processing of data** in a certain way; allowing managers to make decisions and predict future trends.

"Predicting the future by studying the past."

Introduction

- Introduction
- **Motivation**
- Objectives
- Proposal
- Example
- Conclusions
- Future Work

We propose a methodology to develop DSS in the Analysis phase, respective of **data processing modeling** and following **Software Engineering** processes.

Motivation

- Introduction
- **Motivation**
- Objectives
- Proposal
- Example
- Conclusions
- Future Work

One of the **challenges of Software Engineering** (SE), is to propose:

- Rules, process, guidelines and models that address Software development:
  - Quickly, efficiently, and in a specific and unambiguous manner and resulting in a quality product.

Motivation

- Introduction
- **Motivation**
- Objectives
- Proposal
- Example
- Conclusions
- Future Work

The techniques mostly used in Software Engineering are:

- *Data modeling*, proposes structures to store the data efficiently.
- *Use Cases* (UC), allows modeling the functionality of a software system at a high level of abstraction.

Motivation

- Introduction
- **Motivation**
- Objectives
- Proposal
- Example
- Conclusions
- Future Work

On the one hand

Decision Support Systems are based upon historical databases containing large amounts of data.

However, these systems are not always based on databases built for this purpose, which we don't consider efficient.

We believe the Decision Support Systems must be based on data warehouses (DW), or multidimensional databases; and following specific multidimensional (MM) data models.

Motivation

- Introduction
- **Motivation**
- Objectives
- Proposal
- Example
- Conclusions
- Future Work

On the other hand

The User Cases are almost always used in a particular way for each system.

They are "tailored" by the applications that they model.

We think it would be desirable **to have User Cases "*patterns*"** that could be **reused** by most systems that need the same functionalities.

Objectives

- Introduction
- Motivation
- **Objectives**
- Proposal
- Example
- Conclusions
- Future Work

We wish

To show a general **reusable User Case**, as a "***pattern***", which may be used as a ***guide*** in the development of Decision Support Systems to model the **data processing functionality**.

Also **using specific multidimensional** data models.

Proposal

- Introduction
- Motivation
- Objectives
- **Proposal**
- Example
- Conclusions
- Future Work

We will use **data models** and **User Cases** to propose a **guide** for development of DSS; proposing:

- **Using appropriate conceptual MM** data models that reflect the basic starting data required to develop a DW or Multidimensional Data Base.
- **Using a Use Case** "***pattern"*** to represent the functionality of any DSS, regarding **data processing**.

*This will allow us to reflect **how to obtain derived data, dynamically and automatically**, necessary for any analysis.*

Proposal, Multidimensional Concepts

- Introduction
- Motivation
- Objectives

Proposal

Multidimensional Concepts

- Example
- Conclusions
- Future Work

The MM models **should** represent the data **focused to analysis** at the earliest stages of the DSS development.

FactEntity Multidimensional Model

- Introduction
- Motivation
- Objectives

Proposal

- Example
- Conclusions
- Future Work

Next, let us present the FactEntity, FE, model, which is chosen to handle the multidimensional semantics in this work.

This model also supports spatial data.

Proposal, FactEntity Model

- Introduction
- Motivation
- Objectives

Proposal

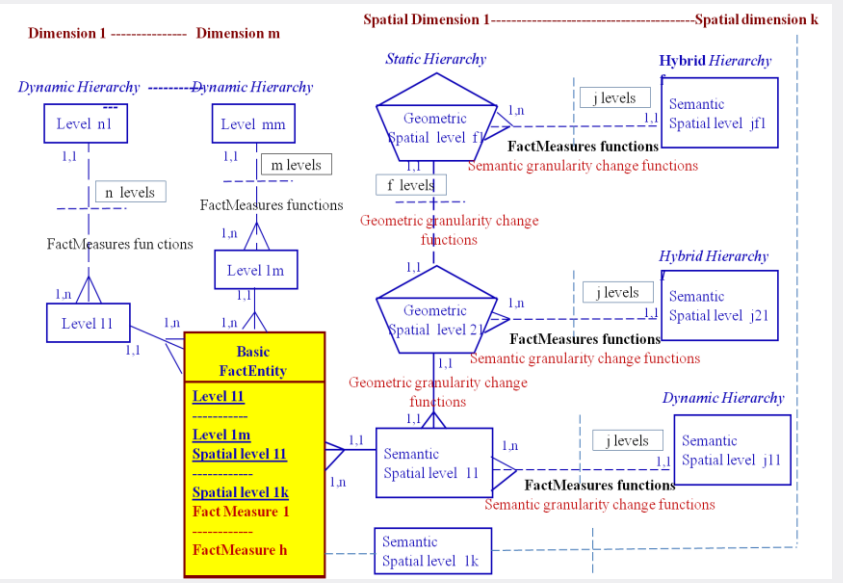
- Example
- Conclusions
- Future Work

The **FactEntity** model distinguishes between:

- **basic data** (existing data)
- **derived data (**data obtained **by processing the basic data** according to the analysis criteria)

Facts and dimensions are combined to obtain the named factEntities, these can be **basic** and **virtual**.

Constructors of FE model, which also represents the **functions** that will be used in **processing data** (and also spatial representations)



Proposal, FactEntity Model

- Introduction
- Motivation
- Objectives

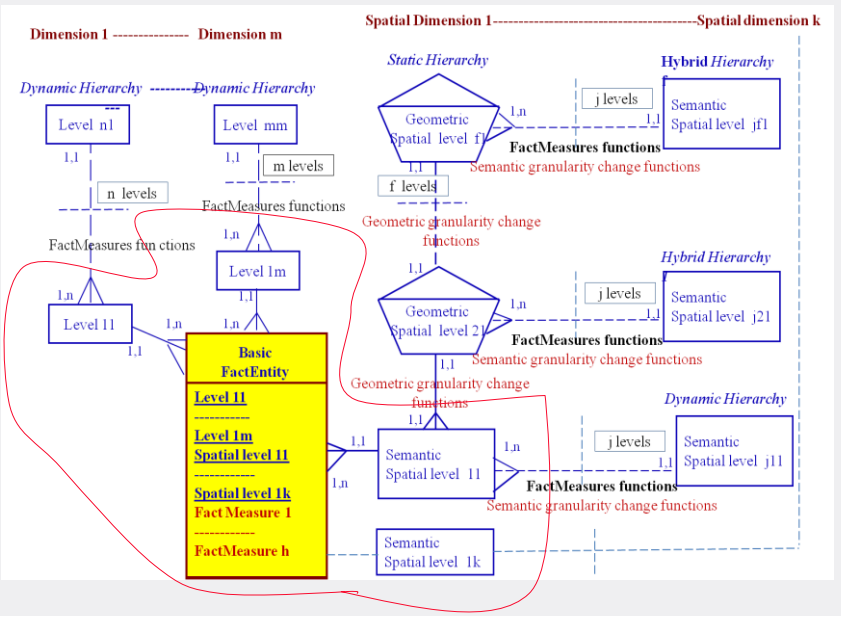
Proposal

- Example
- Conclusions
- Future Work

A **Basic factEntity** is composed of basic data such as:

- A *Basic fact*, object of study
- The *leaf levels* of their associated dimensions (levels of minimum granularity)
- This is represented explicitly in the scheme

The red line surrounds **the basic data** represented in the model



Proposal, FactEntity Model

- Introduction
- Motivation
- Objectives

Proposal

- Example
- Conclusions
- Future Work

The Rollup Operator:

Is used to navigate between the hierarchical levels

The processing data:

Produces **Virtual factEntities** which are generated when the **Rollup operator** is applied.

Proposal, FactEntity Model

- Introduction
- Motivation
- Objectives

Proposal

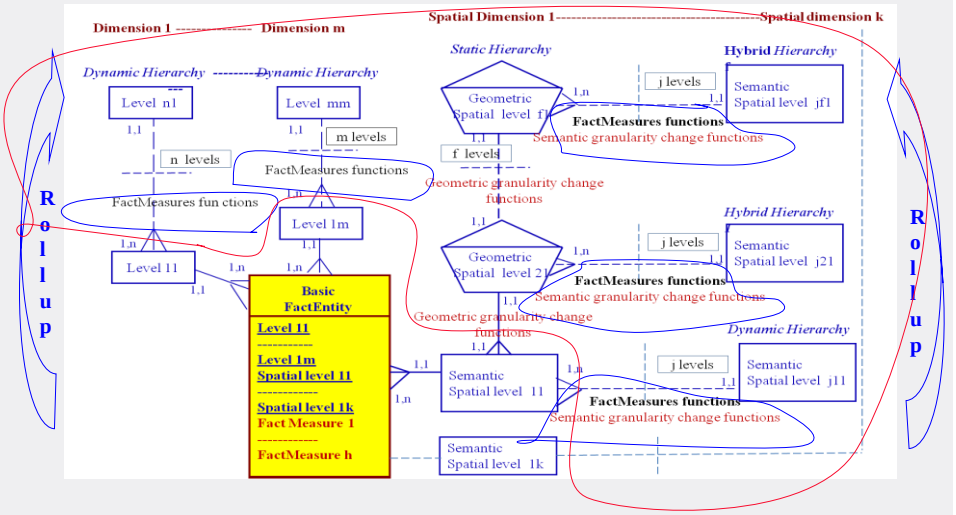
- Example
- Conclusions
- Future Work

Virtual factEntities are composed of derived data, such as:

- *Derived facts* **(The functions** represented in the diagram on **fact measure are applied)**
- *Cartesian subgroups* of their associated dimensions, where at least one dimension is involved with a level, greater than leaf level

These are not represented explicitly in the scheme, but...The necessary elements to generate them, in an automatic way, are represented.

The red line surrounds necessary elements for **Processing data**



When, we perform a **Rollup** over the dimensional levels and we apply the **functions** represented in the diagram on **fact measures**, we obtain the **Virtual factEntities** or **derived data.**

Proposal, USE CASE, UC

- Introduction
- Motivation
- Objectives

Proposal

- Example
- Conclusions
- Future Work

Next we will show the generic User Case model proposed to gather:

Decision Support System **functionalities**

in regard to **the processing of basic data**.

To develop this UC we have used:

- The data model **FactEntity**
- The **Gascueña algorithm**

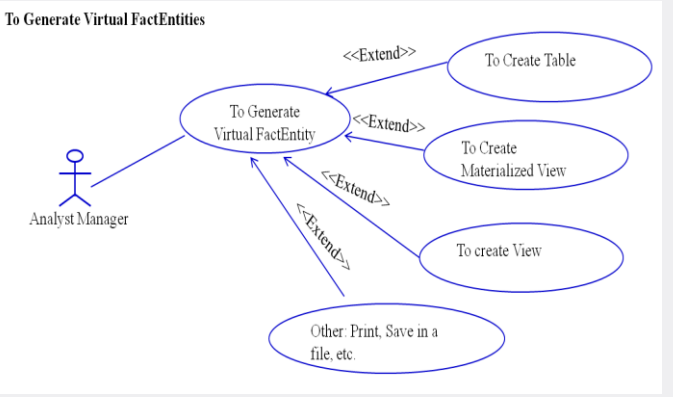
Proposal, USE CASE, UC, Diagram

- Introduction
- Motivation
- Objectives

Proposal

- Example
- Conclusions
- Future Work

Here we can see the *Generate Virtual factEntity UC*, and four associated Use Cases, which show different ways of stored the virtual FactEntities generated:



Proposal, USE CASE, UC, Diagram

- Introduction
- Motivation
- Objectives

Proposal

- Example
- Conclusions
- Future Work

Conclusions

- Introduction
- Motivation
- Objectives
- Proposal
- Example
- **Conclusions**
- Future Work

In this paper we have proposed a methodology, which attempts to serve as a **generalized guide** for the development of DSS **following the Software Engineering guidelines**

Our proposal is framed within the **Analysis phase** of the software development process life cycle.

DATA PROCESSING MODELING IN DECISION SUPPORT SYSTEMS

Polytechnic of Madrid University

Concepción M. Gascueña
cmgascue@eui.upm.es

Iceis2012, Wroclaw. Poland