Context-aware databases: problems and solutions

Letizia Tanca
Politecnico di Milano (*)

(*) joint work with the Context-ADDICT team:
C. Bolchini, C. A. Curino, G. Orsi, E. Quintarelli, R. Rossato, F. A. Schreiber

PersDB 2008, Auckland

Motivations (1)

Keywords for pervasive systems:
mobility, context-dependence, adaptability, multimodality, multi-channel delivery, embedded systems, ubiquity, navigation, incremental discovery, sensors, power-awareness, computational power, storage space…

- **Many large applications require the decomposition of an information base into (possibly overlapping) subsets**
- **Content and services available at different sources and places, and user is an integral part of numerous applications, interacting with:**
  - **service providers, product sellers, governmental organisations, friends and colleagues, sensing devices**

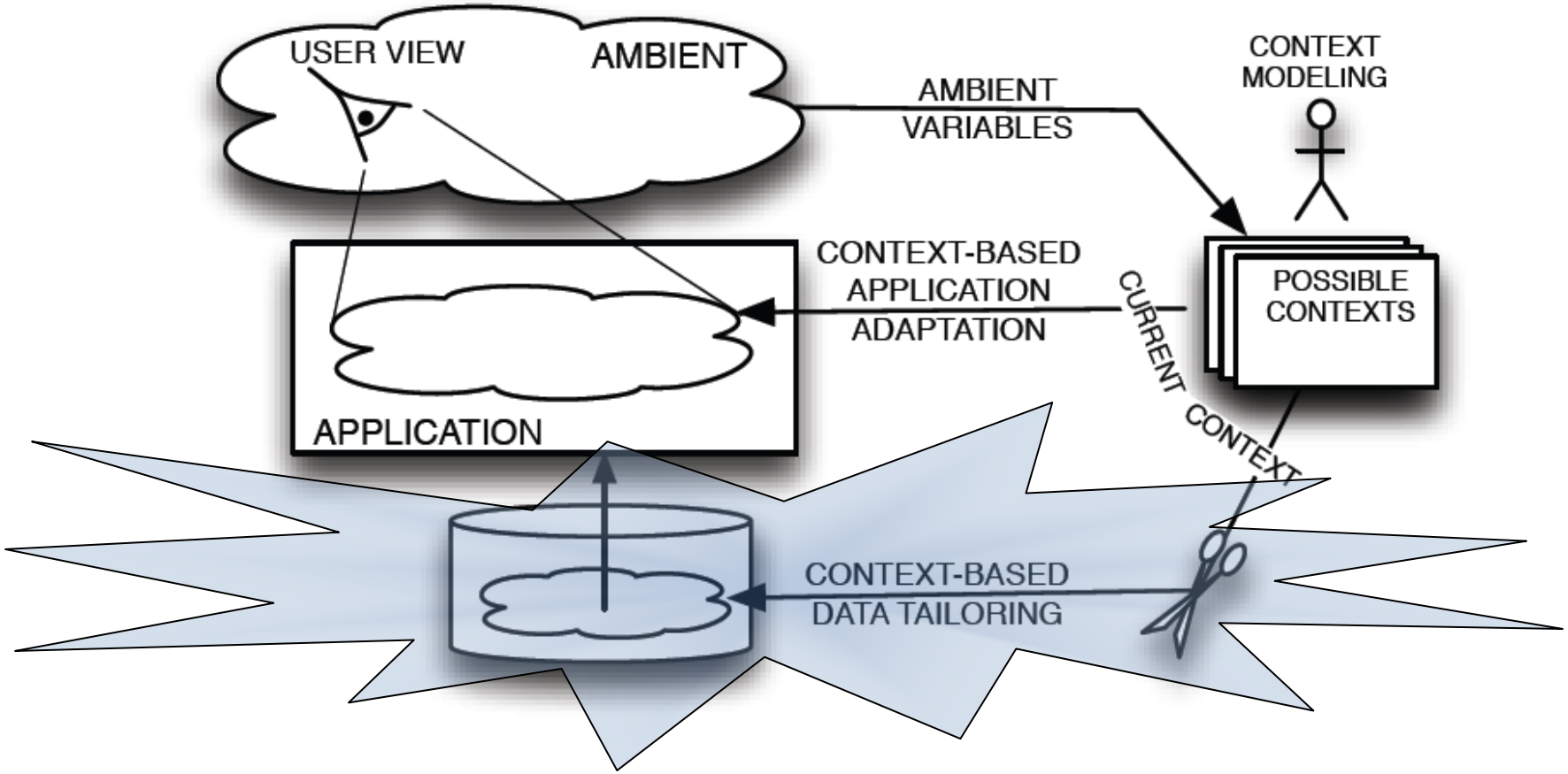
→ **users should not be confused by** *information noise*

→ **the information amount should be** *reduced* **in order to comply with the** *physical limitations* **of small systems**

Further motivations (2)

- *Embedded system design* **has a very important place in the the current SW/HW market -- an embedded system is an engineering artifact involving computation that is <u>subject to physical constraints</u>**
- **Parnas (ICSE 1978 !!!) : design SW for ease of extension** *and contraction*
- **Need for tailoring, not only systems and applications, but also DATA**

Context-aware data tailoring



Context is the scissors

- **derived from the Latin** *cum* **(with or together) and** *texere* **(to weave),**
- **describes a context not just as a profile, but as** *an active process dealing with the way humans weave their experience within their whole environment, to give it meaning***.**

Accordingly, in general…

- **Context can contribute to the meaning that must be inferred from the adjacent world**
- **Such meaning ranges from the references intended for indefinite indications such as "take that" to the shared reference frame of ideas and objects that are suggested by a situation**
- **Context goes beyond immediate binding of variables (→** *semantics***) to the establishment of** *a framework for communication based on shared experience*
- **Such a shared framework provides a collection of roles and relations to organize meaning for a piece of information**

Accordingly, context may involve…

- *viewpoint* - **classifying items according to a viewpoint of some agent**
- *topic* - **any item in the information base is relevant to one or more topics, such as marketing, Canada, etc.,**
- **the setting of** *focus* **by ignoring issues not relevant to a specific situation,**
- **the modeling of limited availability of or restricted** *access* **to certain kinds of information,**
- **the description of situations with** *varying degree of evidence* **as in reasoning processes or in cooperative work,**
- *time slot* - **viewing an entity from just one instant of time as opposed to representing all the properties it gathered throughout its entire life-time.**
- *Location* – **viewing an entity from the different perspectives involved by the position of the user or of the entity itself**

Underlying, often implicit principles

- **Principle 1 (Locality): reasoning uses only part of what is potentially available (e.g., what is known, the available inference procedures). The part being used while reasoning is what we call *context* (of reasoning);**
- **Principle 2 (Compatibility): there is compatibility among the kinds of reasoning performed in different contexts.**

(*Chiara Ghidini, Fausto Giunchiglia,* **Local Models Semantics, or Contextual Reasoning = Locality + Compatibility,** *Artificial Intelligence, 2001*)

The 5 "W" questions of context

WHY

WHO

WHERE

WHEN

WHAT

hoW

(Proceedings of the CHI 2000 Workshop on "The What, Who, Where, When, Why and How of Context-Awareness", David R. Morse, Anind K. Dey, 2000, Georgia Institute of Technology)

What is context?

Context-based application example

Automated support for a natural history museums visitors, endowed with a portable device which reacts to a change of context by

- **adapting the user interface to the different** *abilities* **of the visitor – from low-sighted people to very young children –;**
- **providing different information contents based on the different** *interests/profiles* **of the visitor (geology, paleontology, . . . scholar, journalist, . . .), and on the room s/he is currently in;**
- *learning*, **from the current situation and the previous choices performed by the visitor, what information s/he is going to be interested in** *next*;
- **providing the visitor with** *appropriate services* **– to purchase the ticket for a temporary exhibition, or to reserve a seat for the next in-door show on the life of dinosaurs –;**
- **deriving** *location information* **from sensors which monitor the user environment;**
- **providing knowledge of the** *sorrounding people* **in terms of their roles and respective contexts,** *as related to the user*
- **providing** *active features* **within the various areas of the museum, which alert visitors with hints and stimuli on what is going on in each particular ambient.-**

Context

- Different meanings in different realms of Computer Science
- Is context a matter of….?….

Is context a matter of....?

(a) capability to *adapt content presentation* **to different channels or to different devices**

(b) modeling *location and environment* **aspects**

(c) modeling what *the user is doing*

(d) *agreement and sharing* **among groups of peers**

(e) *service/data reduction*

hoW

is context represented ??

Categories for context model interpretation – Modeled Aspects

- **Space:** whether the system manages location
  - Values: YES/NO
- **Time:** whether the system takes into account the time dimension
  - Values: YES/NO
- **Space/Time Coordinates:** whether the space and time are represented absolutely (e.g. Global Time and GPS coordinates) or relatively (e.g. "near something", "last month", "after that")
  - Values: RELATIVE / ABSOLUTE

Categories for context model interpretation – Modeled Aspects

- **Subject:** the point of view used to describe the context itself: context as if perceived by the user, or from the application point of view (the user itself is a portion of the context)
  - Values: USER / APPLICATION
- **User profile:** if and how the user is represented (profiled) within the context model
  - Values: ROLES / FEATURE-BASED
- **Context History:** whether the history of contexts is part of, or relevant for, the context itself, or instead the context is considered as a static picture of the current instant
  - Values: YES/NO

Categories for context model interpretation - Representation

- **Formalism Type:** the class of formalism used to capture the context
  - Values: Key-value based / Mark-up Scheme based / Logic based / Graph based / Ontology based
- **Flexibility:** is it possible to capture any context with this model?
  - Values: Application-domain bound / General
- **Level of formality:** how formal is the model
  - Values: High / Low

Categories for context model interpretation - Representation

- **Variable granularity management:** ability to model context aspects at different levels of detail
  - Values: YES/NO
- **Constraints on valid contexts:** whether the system provides ways to control the generation of invalid contexts
  - Values: YES/NO

Categories for context model interpretation – context usage

- **Context construction:** the possibility to build the context model (or instance) as the result of a joint effort
  - Values: Centralized / Distributed
- **Context dynamicity:** can the context be modified at run time
  - w.r.t. a single context:
    - Values: run-time vs. design time
  - w.r.t. the context model (dynamic reconfiguration of the context model itself):
    - Values: YES/NO
- **Multi-context management:** the possibility to represent and manage within a single instance of the model all the possible contexts, as opposite to a model where each instance represents a context.
  - Values: YES/NO

Categories for context model interpretation – context usage

- **Context info quality monitoring:** is it possible to control and correct automatically acquired context info (e.g. when sensors are present)?
  - Values: YES/NO
- **Reasoning:** whether the context model enables reasoning on context information to obtain more abstract or more complex context descriptions
  - Values: YES/NO

Categories for context model interpretation – context usage

- **Ambiguity and incompleteness management:** in case of ambiguous or incomplete context information, is the system able to ``interpolate'' somehow the context information and reconstruct a plausible ``current context''?
  - Values: YES/NO

- **Automatic Learning Features:** whether the system includes context learning features, e.g. by observing user behavior during browsing the system may autonomously learn user's preferences
  - Values: YES/NO

What is context?

Is context a matter of....?

(a) ...capability to adapt content presentation to different channels or to different devices:

- Different levels of granularity
- Often do not include a refined mechanism of location and time awareness
- User profiling is often present, feature-based
- Context specification quite informal
- Not very flexible (designed for specific applications)
- Automatic context learning (often available)
- No context reasoning

Is context a matter of....?

(b) ...modeling location and environment aspects

- Precise treatment of the time and space coordinates
- Ability to model the context in a highly flexible way
- Context reasoning provided, offering a powerful abstraction mechanism
- Information quality control, since in this case the system acquires location and time information from various kinds of sensors
- Ability to deal with information ambiguity

Is context a matter of....?

(c) ...modeling what the user is doing

- Context history is modeled
- Reasoning mechanisms are needed
- Time and space are taken into account as related to the user's current activity
- Different levels of formality
- Automatic learning (available in some systems) used to guess user activity from sensor readings

Is context a matter of....?

(d) ... agreement and sharing among groups of peers

- Focus on reaching an agreement about a context shared among peers
- Context definition reached in a distributed fashion
- Context reasoning is present
- Quality monitoring is present
- Ambiguity and incompleteness are managed
- Context model rather well formalized
- Primitive treatment of location, time and user profiling

Is context a matter of....?

(e) ...services/data reduction

- Selection of data, but also of relevant functionalities and services
- Time, space and user profile highly developed and well formalized
- Possibility to describe different kinds of contexts, thus flexibility is high
- Dynamic context construction; dynamic context model adaptation
- Variable level of granularity
- Constraints on valid contexts can be expressed
- Context history and reasoning often not provided

Context model features and systems exposing them

System	Space	Time	Space/Time coordinates (Relative or Absolute)	Context history	Subject (User or Application)	User profile (Role or Features based)	Variable context granularity	Valid context constraints	Type of formalism: Key-value-based	Type of formalism: Mark-up based	Type of formalism: Logic-based	Type of formalism: Graph-based	Type of formalism: Ontology-based	Formality level (High or Low)	Flexibility	Context construction (Distributed or Centralized)	Context reasoning	Context quality monitoring	Ambiguity/Incompleteness mgmt.	Automatic learning features	Multi-context model
ACTIVITY	+		A	+	U	F	+					+		L	+	C	+				+
CASS	+	+		+	U						+			L		D	+				
CoBrA	+	+	A		A	F							+	H	+	D	+		+	+	
CoDaMoS	+	+	R/A		A	F							+	H	+	D	+	+			
COMANTO	+	+	R/A		A	F							+	H	+	D	+				
Context-ADDICT	+	+	R/A		A	R	+	+		+		+	+	H	+	C					
Conceptual-CM	+	+	R	+	A	R						+		L	+	C	+		+	+	+
CSCP					A	F				+				L		C					+
EXPDOC		+	R	+	U	F					+	+		H		C				+	
FAWIS				+	U	F	+		+					H		C			+	+	+
Graphical-CM	+	+	R		A	F						+		H	+	C		+	+		+
HIPS/HyperAudio	+	+	A	+	U	F			+					L		C	+		+	+	
MAIS	+	+	A		A	F	+						+	H		C					
SCOPES					A						+			H		D	+		+		+
SOCAM	+	+	R/A		A	F							+	H	+	D	+	+			
U-Learn	+	+	A		U	F		+					+	H	+	D					+

Conclusion on context models in computer science

- **Due to the complexity of the problem as a whole and to the multitude of different applications, the best models** *have a well defined focus*, **and try to support only** *one of the above mentioned context-modeling categories*

  C. Bolchini, C. A. Curino, E. Quintarelli, F. A. Schreiber, L. Tanca:A Data-oriented Survey of Context Models, SIGMOD Record, Dec. 2007. http://poseidon.elet.polimi.it/ca/

- **Context in** *information systems*: **the** *viewpoint* **abstraction**

Context as abstraction

- **abstraction mechanisms are the indispensable means** *to deal with complexity in information system design* **(see abstractions for database design).**
- **the** *viewpoint abstraction* **has received little attention**
- **context as a** *viewpoint mechanism* **that takes into account implicit background knowledge**

Context in Information Systems

- **Context as a selector of *workspaces***

  → **information-system oriented**

  M. Theodorakis, A. Analyti, P. Constantopoulos, N. Spyratos: A theory of contexts in information bases. Inf. Syst. (IS) 27(3):151-191 (2002)

  R. Motschnig-Pitrik: A Generic Framework for the Modeling of Contexts and its Applications, Data & Knowledge Engineering, 2000.

- **Context as a selector of *views* or *facets***

  → **database oriented**

  Y. Roussos Y. Stavrakas V. Pavlaki: Towards a Context-Aware Relational Model, In "Contextual Representation and Reasoning" Workshop (CRR'05), held in conjunction with CONTEXT'05, Paris, 2005

  K. Stefanidis, e. Pitoura and p. Vassiliadis: A Context-Aware Preference Database System, J. PERVASIVE COMPUT. & COMM. 1 (1), March 2005, Troubador pub. Ltd.

  Ours

Conclusions

- Even within Information Systems, context is used for different purposes:
  - to provide access to different facets of the same object or group of objects
  - to provide/equip different users with specific functions in various situations
  - to tailor data or services in different "shapes"
  - to associate data with different preference for values in different situations
- Fundamental underlying concepts:

  - VIEW
  - CONTEXT DIMENSION

Context-aware Data Tailoring
in the Context-ADDICT project

The space-reduction point of view

A contribution from the database research area

CHALLENGES:

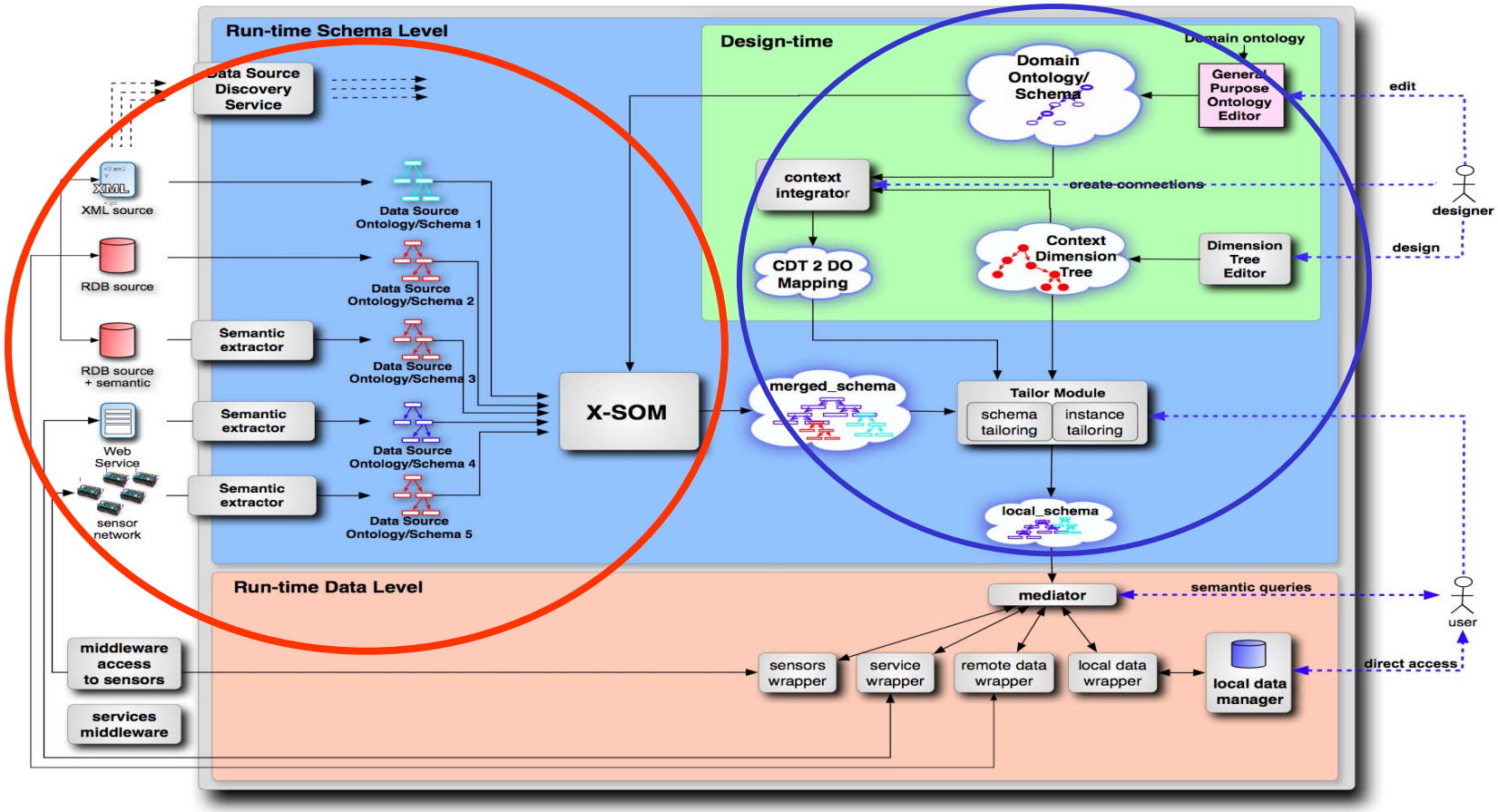
- Involved data volumes
- Data heterogeneity
- Data dynamicity
- Data distribution
- Scalability of the personalization solutions

ANSWERS:

- Reduction of data volumes → context-aware data tailoring
- Data integration

ContextADDICT ARCHITECTURE

On-the-fly data integration + data reduction via tailoring



Context-ADDICT

System	Space	Time	Space/Time coordinates (Relative or Absolute)	Context history	Subject (User or Application)	User profile (Role or Features based)	Variable context granularity	Valid context constraints	Type of formalism: Key-value-based	Type of formalism: Mark-up based	Type of formalism: Logic-based	Type of formalism: Graph-based	Type of formalism: Ontology-based	Formality level (High or Low)	Flexibility	Context construction (Distributed or Centralized)	Context reasoning	Context quality monitoring	Ambiguity/Incompleteness mgmt.	Automatic learning features	Multi-context model
ACTIVITY	+		A	+	U	F	+					+		L	+	C	+				+
CASS	+	+		+	U						+			L		D	+				
CoBrA	+	+	A		A	F							+	H	+	D	+		+	+	
CoDaMoS	+	+	R/A		A	F							+	H	+	D	+	+			
COMANTO	+	+	R/A		A	F							+	H	+	D	+				
Context-ADDICT	+	+	R/A		A	R	+	+		+		+	+	H	+	C					
Conceptual-CM	+	+	R	+	A	R						+		L	+	C	+		+	+	+
CSCP					A	F				+				L		C					+
EXPDOC		+	R	+	U	F					+	+		H		C				+	
FAWIS				+	U	F	+		+					H		C			+	+	+
Graphical-CM	+	+	R		A	F						+		H	+	C		+	+		+
HIPS/HyperAudio	+	+	A	+	U	F			+					L		C	+		+	+	
MAIS	+	+	A		A	F	+						+	H		C					
SCOPES					A						+			H		D	+		+		+
SOCAM	+	+	R/A		A	F							+	H	+	D	+	+			
U-Learn	+	+	A		U	F		+					+	H	+	D					+

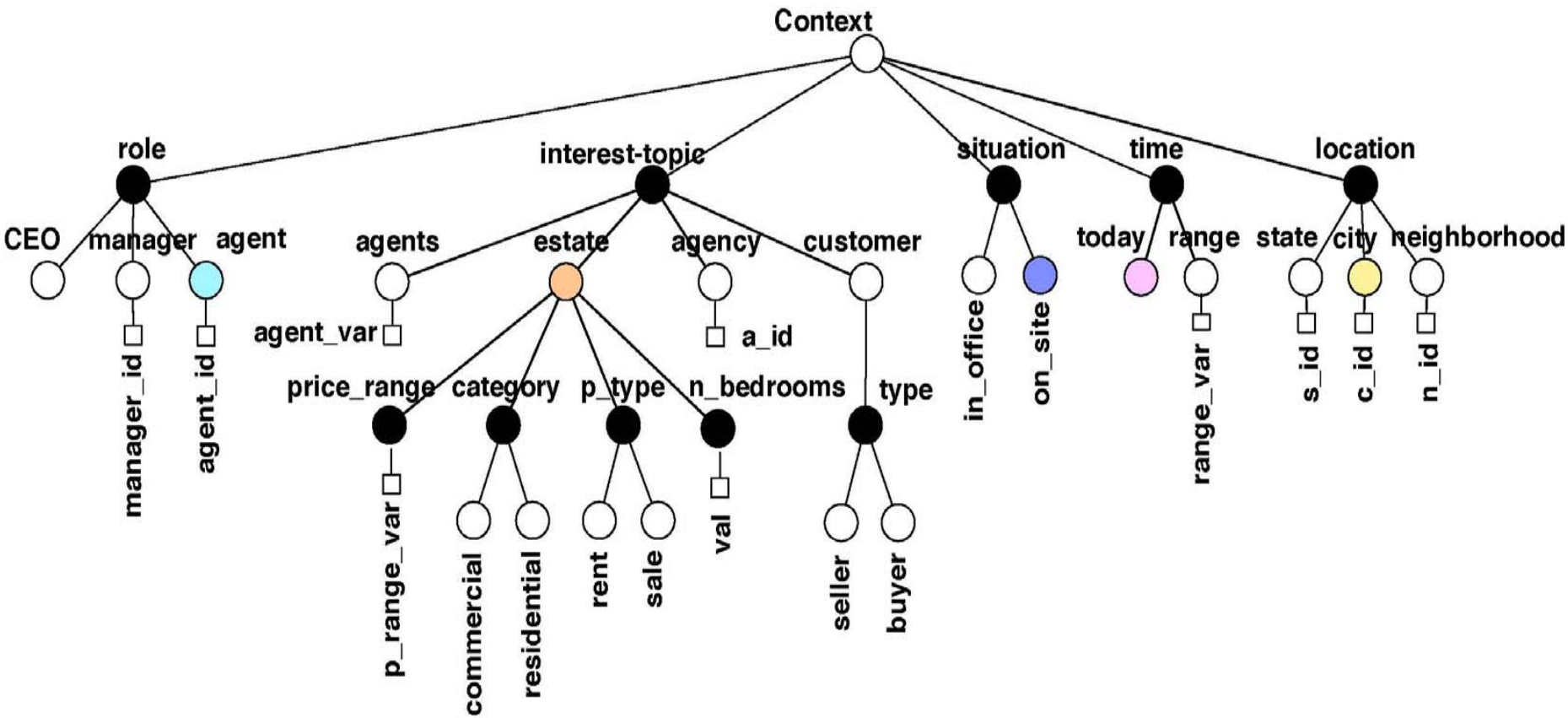
Ambient (or Context) Dimensions

- different points of view the device data are viewed from
- they drive the portion of data to be selected, for instance to be stored on a portable device
  - views over the global schema

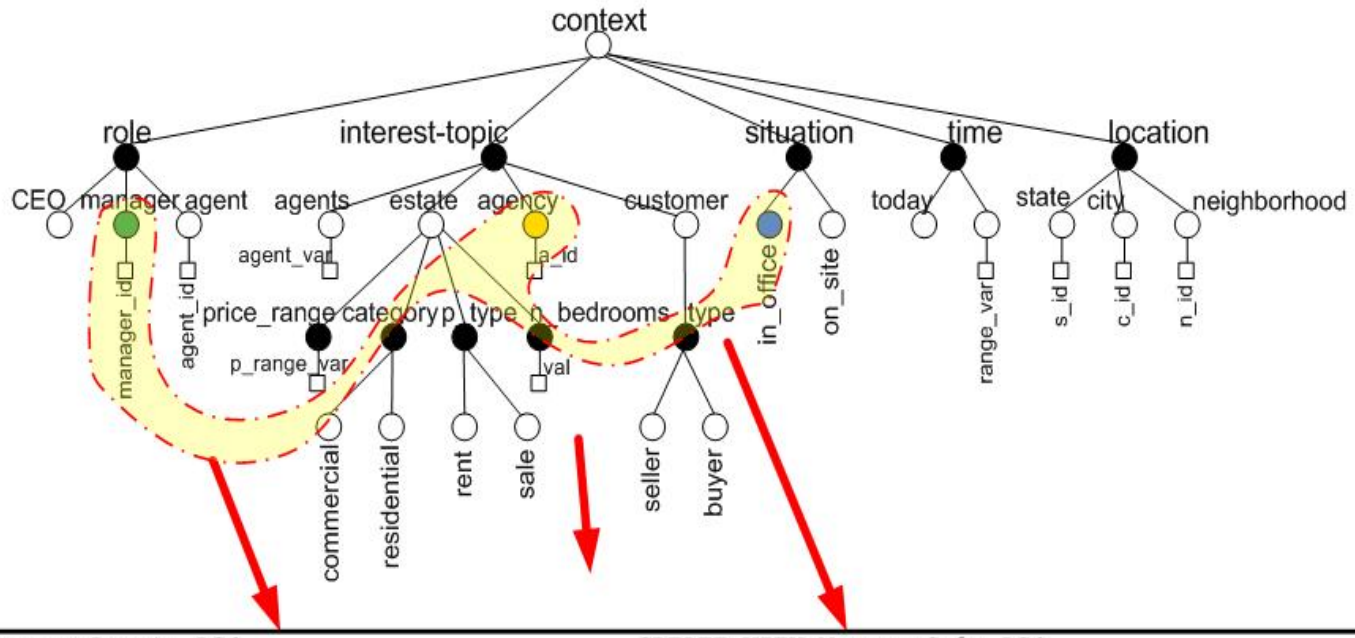
The real-estate example

OWNER(IdOwner, Name, Surname, Type, Address,City, PhoneNumber)
ESTATE(IdEstate, IdOwner, Category, Area, City, Province, RoomsNumber,
Bedrooms, Garage, SquareMeters, Sheet, CadastralMap)
CUSTOMER(IdCustomer, Name, Surname, Type, Budget, Address, City, PhoneNum)
AGENT(IdAgent, Name, Surname, Office, Address,City,Phone)
AGENDA(IdAgent, Data, Hour, IdEstate, ClientName)
VISIT(IdEstate, IdAgent, IdCustomer, Date, ViewDuration)
SALE(IdEstate, IdAgent, IdCustomer, Date, AgreePrice, Status)
RENT(IdEstate, IdAgent, IdCustomer, Date, RatePrice, Status, Duration)
PICTURE(IdPicture, IdEstate, Date, Description, FileName)

Context dimension tree: representation of the context dimensions



The agency manager, when in the office




```
CREATE VIEW ManagerAgents AS{
  SELECT P.*
  FROM Personnel AS P
  WHERE agentManager=$manager_id
}

CREATE VIEW ManagerEstates AS{
  SELECT E.*
  FROM Estate AS E, Agenda AS A
  WHERE E.estateID = A.estateID AND
        A.agentID IN (SELECT personellID
                      FROM ManagerAgents)
}
```

```
CREATE VIEW ManagerSale AS{
  SELECT S.*
  FROM Sale AS S
  WHERE S.agentID IN (SELECT personellID
                      FROM ManagerAgents)
}

CREATE VIEW ManagerRent AS{
  SELECT R.*
  FROM Rent AS R
  WHERE R.agentID IN (SELECT personellID
                      FROM ManagerAgents)
}
```

Example

- An agent (whose id. is 23564), ready to take prospective buyers to visit the residential estate properties located in the “Piola” area ($zone id=“Piola”)
- The current context C is the conjunctive propositional formula:

```
role = agent(“23564”)
^ category = residential
^ type = sale
^ zone = “Piola”
^ situation = on_site
^ time = today(getdate())
```

Context Modeling: constraints

- Not all configurations make sense: e.g., there is no point in combining the CEO role with the on-site situation
- The designer can express constraints on the possible combinations of the context elements, so that, when the complete set of contexts is combinatorially generated, many contexts be discarded
- Constraints are specified over the tree by means of standard logical formulae
- The most common constraints are the forbid (or "useless-context") constraints

Some interesting constraints

- *useless context (forbid)* **constraints allow the context designer to specify configurations that are not significant (either meaningless context situations or irrelevant for the application). In the real estate example, the CEO never goes on site:**

$$\neg\ (role = CEO\ \wedge$$

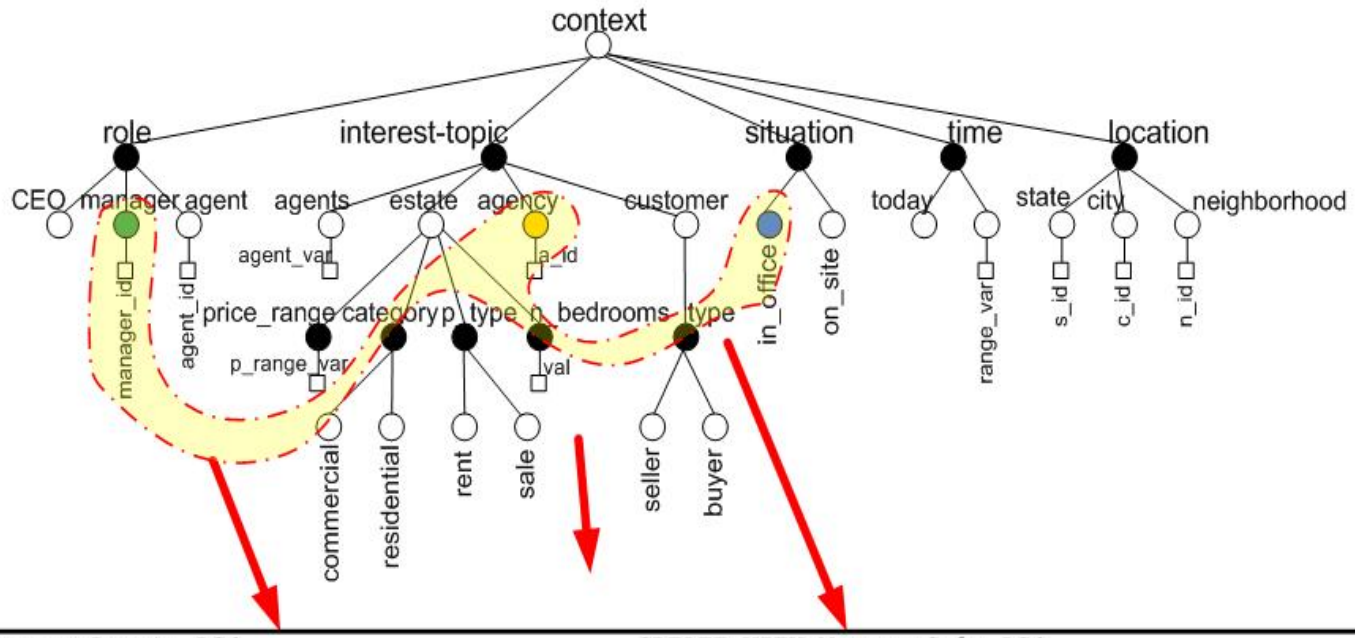
$$(Desc(situation = on\_site)\ \vee\ situation = on\_site\ )$$

- *preferred-detail* **constraints allow the designer to express the level of detail to be preferred for a dimension, w.r.t. the other dimensions' values. In the agency example, the CEO has access to all the data related to estates, other roles may be interested in lower level details :**

$$\neg\ (role = CEO$$

$$\wedge\ Desc(interest\text{-}topic = estates))$$

Recall what we want to do




```
CREATE VIEW ManagerAgents AS{
  SELECT P.*
  FROM Personnel AS P
  WHERE agentManager=$manager_id
}

CREATE VIEW ManagerEstates AS{
  SELECT E.*
  FROM Estate AS E, Agenda AS A
  WHERE E.estateID = A.estateID AND
        A.agentID IN (SELECT personellID
                      FROM ManagerAgents)
}
```

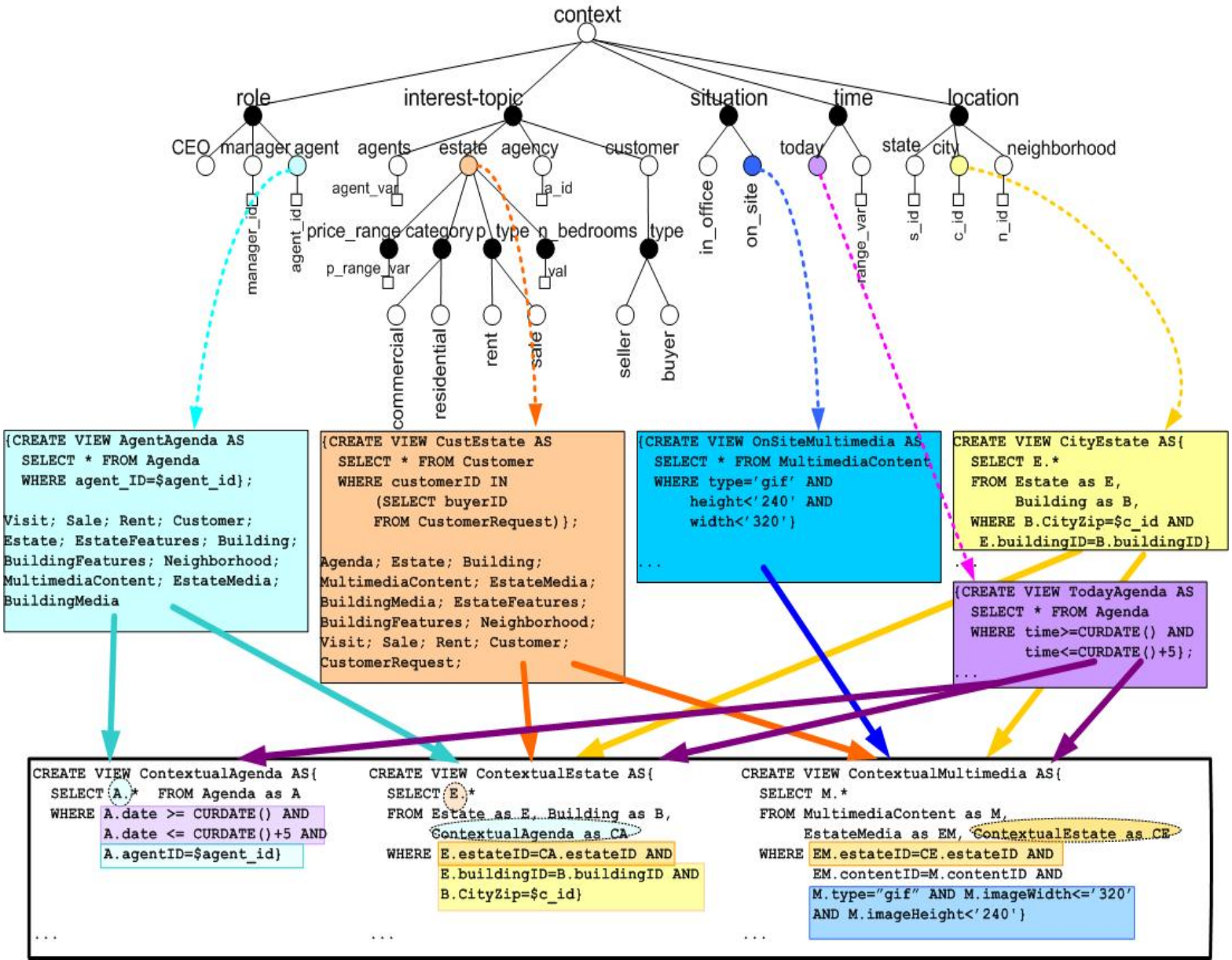
```
CREATE VIEW ManagerSale AS{
  SELECT S.*
  FROM Sale AS S
  WHERE S.agentID IN (SELECT personellID
                      FROM ManagerAgents)
}

CREATE VIEW ManagerRent AS{
  SELECT R.*
  FROM Rent AS R
  WHERE R.agentID IN (SELECT personellID
                      FROM ManagerAgents)
}
```

Node-based area assignment

- **The relevant area, or view, related to a context C, is denoted by** $\mathcal{R}el(C)$
- **Assigning relevant areas to all possible (valid) contexts of a tree is a very time-consuming task**
- **Less time-consuming, but more difficult from a conceptual viewpoint, is deriving the context view from the composition of relevant areas** *of the component nodes.* **Thus:**
  - *Relevant area assignment*: a *partial view*, expressed as a set of relational algebra expressions, is associated with each context element.
  - *View composition by algebraic operators*: the previously defined partial views, associated with the context elements in a configuration, are properly *combined* by means of an appropriate integration operator, to automatically obtain a *final view* defining all the information relevant for that configuration

Composing relevant areas



CONCLUSIONS

Much has still to be done on data tailoring:

- **Extend to other data models (e.g. XML, ontologies)**
- **Study more and different operators and their use (double union, double join)**
- **Study more operator properties**
- **http://poseidon.elet.polimi.it/ca/**

Related work in our group at PoliMI

- Context-ADDICT project
  - X-SOM ontology integration tool → <u>on-the fly</u> data integration
- Ontological support for schema evolution
- Data mining for:
  - (1) Approximate, intensional queries, (2) Constraint extraction for optimization, (3) Constraint relaxation for system adaptability, (4) Dataguide extraction for XML query support
- PoLiDBMS: a DBMS for very small DB's
- PerLa: PERvasive database LAnguage (sensors, RFiD, ..)

Q & A