The OO-Binary Relationship Model: A Truly Object-Oriented Conceptual Model

Olga De Troyer

No. 27

This paper has been published in:

ISSN 0924-7807

©1991. Institute for Language Technology and Artificial Intelligence, Tilburg University, P.O.Box 90153, 5000 LE Tilburg, The Netherlands Phone: +3113 663113, Fax: +3113 663110.
The O0-Binary Relationship Model:
A Truly Object Oriented Conceptual Model

O. De Troyer
Tilburg University
Infolab
P.O. Box 90153
5000 LE Tilburg
The Netherlands.

Abstract.

Conventional conceptual models like the Binary Relationship Model (also known as NIAM) or the Entity-Relationship Model do not fit well with the promising object oriented database systems. In this paper we show that is possible to turn such a conventional conceptual model (in particular the Binary Relationship Model) into a truly object oriented conceptual model which combines the assets of the conventional model with the advantages of the object oriented approach.

1. Introduction.

In recent years there has been a substantial influence of object-oriented (00) programming languages to the Data Base Management System (DBMS) technology. This has resulted in the development of a number of object-oriented data models and systems [e.g. 1, 2, 3, 4, 9, 21, 29]. An object-oriented modelling method offers a number of important advantages, such as:
(1) All information is modelled through a single concept, namely objects. Each object has its own identity.
(2) The encapsulation mechanism that packs the data (instance variables) and behaviour (methods) of an object together, protects the data from corruption by other objects and hide low level implementation details from the rest of the system (information hiding).
(3) Similar objects are grouped together into classes. All objects belonging to the same class are described by the same instance variables and methods.
(4) The concept of class allows to define new abstract data types which enjoys the same rights and privileges of the built-in data types.
(5) Classes may be organized into subclass hierarchies. The instance variables and the methods specified for a class are inherited by all of its subclasses. In addition, the subclasses may have properties (instance variables as well as methods) of their own. This generalization/specialization mechanism allows to abstract away the detailed differences of several class descriptions and to factor out the commonalities as a common superclass. The technique of inheritance makes it unnecessary to restate the properties of the superclass for the

* This work was supported by the Basic Research Action IS-CORE of the Commission of the European Communities.
subclasses. In addition it will localize modifications due to changing requirements.

(6) Through the technique of overloading systems are simplified because the same method name may be used to activate an operation common to different types of objects but differently implemented for each type of object. It also makes the system easier to extend; if a new type of object, also bearing this operation is added, the existing objects are totally unaffected by this change.

A special case of overloading is overriding. This technique redefines a method of a class in a subclass. It allows special cases to be handled more efficiently and easily without having an impact on other objects.

(7) Objects can also contain other objects. Such objects are usually called composite objects. They are important because they can represent far more sophisticated structures than simple objects can.

On the other hand, conceptual data models (sometimes also called semantic models) are used during the information analysis stage in the development of an information system (IS) to translate the user's information requirements into a supposedly precise, complete and unambiguous description, the so-called conceptual schema (CS). One of the important functions of such a conceptual schema is to serve as a tool to communicate among all parties involved, the semantics of the Universe of Discourse (the application domain containing the given situation), abbreviated UoD. Therefore this description should ideally only deal with conceptual issues; i.e. the various object types of the UoD, their associations and the integrity constraints. No representation or implementation oriented details should be included. Looking too early into problems such as organization of the UoD objects into records, normalization, access strategies, representations, etc., distracts the specifiers' attention from mastering the real UoD problems. Well known examples of early conceptual models are Entity-Relationship models [e.g. 8], Binary Relationship Model [e.g. 26, 32], Object-Role Model [e.g. 14], Functional Model [e.g. 28]. Later on in the development of an information system, (at least for conventional DBMSs) the CS may become input to the design of a data base schema, which is (was) usually expressed in terms of record types, key fields and relationships between records. Until now, there is no clear connection of a conceptual schema to an OO-DBMS. It may even be questioned if a connection to an OO-system is possible and opportune. Some people will even argue that OO-data models makes conceptual models completely redundant.

Although most of the earlier mentioned conceptual models have already incorporated some of the attractive features of the OO-approach (e.g. subtype hierarchies together with inheritance of properties and aggregation hierarchies) [e.g. 18, 35] there is still a big gap between these conceptual models and OO-models. The main reasons for this are:

(1) Most of these conventional conceptual models do not support the specification of the behaviour of the IS, and if they support it there is usually a clear separation between the description of the information
structure (the data) and the functionality (behaviour) of the IS. In OO-approaches specification of data and behaviour is inextricably bound.

(2) Conventional models describe the UoD as a bunch of entities and their relationships and interactions are described from a global viewpoint. The application programs are described in a top down way. The approach is usually based on functional decomposition- and modular programming techniques. This is completely different in the OO-approach. The OO-approach does not start with the tasks or functions to be performed by the system, but rather with the objects that are needed in order to perform the tasks. Once these objects are correctly represented, they can be used to solve a wide variety of tasks including the original one. Objects are specified as general-purpose building blocks and their relationships and interactions with other objects are described locally. Systems are built by assemble objects. In this sense, the OO-approach is rather a bottom up approach.

(3) OO-systems usual offers a large range of new data types like e.g. sets, arrays, bitmaps,... together with the possibility to define others. These are needed because of the complexity of the data of non-conventional information systems. Conventional conceptual models are not able to support the specification (at an abstract level) of the use or definition of such data types.

Although in a recent paper [20] Maier argues that it is not possible to define a general object oriented data model because of the diversity of the features supported by the current available OO-DBMSs, it is our opinion that even in an OO-DBMS environment a conceptual schema is still desirable. Its function remains the same as in a conventional environment; to serve as a specification reference by giving an unambiguous description of the conceptual issues of the system which can also be understood by non-technical persons. The graphical representation technique used to represent the specifications plays an important role in the success of conceptual models as a communication tool between information analysts, database experts and non-technical persons.

Only, the earlier mentioned conceptual models as such cannot be combined with an OO-data model without sacrificing a number of the advantages of the OO-approach.

In this paper we show that it is possible to turn a conventional conceptual model (the Binary Relationship Model, also known as NIAM) into an OO-conceptual model without sacrificing the main assets of the model. The step from an OO-conceptual schema to an OO-implementation will then be much easier and the full power of the OO-approach can then be exploited. We have opted for the Binary Relationship Model (BRM) because it has a number of abstraction capabilities which are in particular suitable for conceptual information analysis.

The paper is organized as follows: In section 2 we briefly describe the BR model. Section 3 presents our object oriented version of the BR model and in section 4 we summarize differences and similarities between the conventional and the OO-BR model and discuss the difference with other approaches. The meta-schema of the OO-BR model is given in the appendix.
2. The Binary Relationship Model (BR model).

The descriptions of the BR model (under different names) appear in several forms in the literature ([12, 16, 19, 22, 24, 25, 26, 32, 34 ...]). The Binary Relationship Model has proven to be successful in several large scale industrial projects [e.g. 13, 31].

Its main abstraction capabilities are:

1. **Abstraction from lexical representation.**
   The objects of the UoD are classified into lexical and non-lexical objects. Examples of non-lexical objects are employees, departments. The lexical objects are objects which may be used to represent the non-lexical objects of the UoD; e.g. employee names, employee numbers, department names. Although there is an explicit distinction between lexical and non-lexical objects no decision whatsoever is made about which lexical object (or combination) will be used to represent a non-lexical object. Only possible lexical representations are identified.

2. **Abstraction from instance level.**
   This abstraction capability is well known and present in most conceptual models. It means that the building blocks are the object types (OT) instead of the individual objects.

3. **Abstraction from aggregation.**
   All associations between object types are expressed by means of binary relationships (facts). No fore-ordered data organization is imposed. No decision has to be taken whether an object is entity or attribute.

4. **Abstraction by generalization/specialization.**
   As in the OO-approach, the BR model supports the notion of object type sub-hierarchy and inheritance of properties along this hierarchy. An object type sub-hierarchy captures the "is-a" relationship between an object type and its subtypes. Multiple supertypes for a subtype are allowed.

   In addition to these abstraction capabilities, the BR model explicitly addresses the issue of constraints. Constraints constitute the semantical component of the conceptual schema; they ensure that the schema meaningfully reflects the UoD. A large variety of constraint types are supported. Other constraints may be expressed in one of the constraint languages proposed in the literature [36, 32, 11].

   We adopt the well known "NIAM" graphical notation [12], [32] for the BR concepts:

   ![Diagram](image)

   a NOLOT -- (NOn-Lexical Object Type)
a LOT -- (Lexical Object Type). A LOT may be involved in one fact only, with a NOLOT.

a Fact. The "boxes" are called Roles. Each Fact involves exactly two Object Types (which may be the same).

a Sublink - the subtype occurrences implicitly inherit all properties of the supertype. Subtypes need not be disjoint and not all of a NOLOT's occurrences need to be in one of its subtypes.

Certain constraint types occur so frequently and are so fundamental that they have a graphical representation as well. We only introduce the graphical representation of some constraint types:

The Identifier constraint (simple functional dependency) is drawn as a line over the key-role.

A Total Role constraint stating that each instance of an OT should participate in a given Role is represented by a "V" sign.

A Total Union constraint is a generalization of a Total Role constraint. Each instance of the OT should participate in at least one of the indicated Roles or Sublinks.

The Exclusion constraint expresses the mutually exclusion of a number of Subtypes.
A Uniqueness constraint is a generalization of the Identifier constraint. An instance of an Object Type is identified by an instances-combination of the indicated Object Types.

The classical BR model as such cannot be called object-oriented, because:
(1) Objects do not have an object identity. They only exist through their properties; objects with the same properties are considered to be the same.
(2) Behaviour of objects is not specified.
(3) Data cannot be encapsulated into an object.
(4) There is no way to define new abstract data types.
(5) There is no concept such as composite object.

3. An object oriented BR model (OO-NIAM).

In this section we introduce an object oriented version of the BR model. This is not done by simply adding extra features. Instead we build up the OO-BR model from scratch, but we follow most of the principles of the classical BR model and combine them with OO-principles.

3.1. Objects, Object Types and Subtypes.

Following the OO-principles, an object is constituted by some private data (the state) and a set of operations (the behaviour) inextricably bound to it. In the classical BR model the concept of Object Type (OT) is only used to classify objects. In our OO version of the BR model, the description of state and behaviour of the objects are encapsulated into the OT. The OT defines the attribute relations, methods and constraints of the objects of this type. Attribute relations describe the state of an object, methods describe the behaviour of the object and the constraints restrict the possible states and the behaviour of an object. We use the term properties to denote attribute relations, methods and constraints.

In addition to the properties of the objects, an OT (which can also be considered as an object) may also described its own properties; they are called type-properties (type-attributes, type-methods and type-constraints).

As in regular NIAM, an OT will be represented by a circle.

An OT may be subtype of one or more supertypes. The properties of the supertypes are inherited by the subtypes. We keep the strict "is-a" meaning of the BR-model sublink concept; each object of the subtype is also an object of the supertype. We also keep the graphical NIAM representation for a sublink; an arrow pointing from the
subtype to the supertype. The BR-model exclusion and totality constraints expressible on sublinks keep their meaning.

As in Smalltalk [9, 17], we consider a single OT-sub-hierarchy for the entire schema. All OTs are ultimately subtype of the pre-defined OT "Object" which capture all objects of the UoD. According to the BR model principles, objects are classified into lexical and non-lexical objects. This means that the OT "Object" has two (disjoint) (pre-defined) subtypes "Lexical Object" and "Non-Lexical Object". The lexical objects may be further divided into e.g. "Boolean", "Number", "String" and "Text". See figure 1.

Each UoD-specific OT is a subtype of either "Lexical Object" or "Non-Lexical Object". For example, document numbers are lexical objects while documents are non-lexical objects (see figure 2).
We make the convention that UoD-defined OTs which are subtypes of "Lexical Object" will be represented by a dotted circle and we call them LOTs (similar as in regular NIAM). UoD-defined OTs which are subtype of "Non-Lexical Object" will be represented by a solid circle and are called NOLOTs (also similar as in regular NIAM). By this convention the sublink arrow to the respectively supertype "Lexical Object" or "Non-Lexical Object" can be omitted (see figure 3).

![Figure 3](image)

**figure 3**: graphical convention for LOTs and NOLOTs.

The (abstract) data type of a LOT may be specified by declaring the LOT a subtype of one of the pre-defined subtypes of "Lexical Object". For an example see figure 4; document numbers are defined to be numbers. Note that it is not necessary to specify the data type of LOT. It may be left unspecified till the implementation phase.

![Figure 4](image)

**figure 4**: data type for a LOT

By introducing new subtypes of "Lexical Object" new abstract data types may be defined by the user himself. See for instance figure 5 where two new lexical types "Name" and "Date" are introduced.
Note that in figure 5 we have NOT used the dotted circle convention for "Name" and "Date" this to indicate that "Name" and "Date" are UoD independent OTs. This allows the user to extend the pre-defined subtype-hierarchy without limitations in a UoD independent way. These new defined subtypes can later be used in different conceptual schemas.

3.2. Object Type Definition.

In the OO-version of the BRM, the OT concept is the mechanism to hide the description of the properties of the objects of this type. To describe these properties we will use a graphical representation technique very similar to the usual representation technique for the BR model.

The OT under description is represented as a circle embedded in a square. Attribute relations are considered to be special facts; they describe binary links between the objects of the OT under description and some other objects. The usual BRM constraints are used to restrict the possible instances of the attribute relations. For type-attributes the same graphical representation as for attribute relations is used but the role connected to the OT under description is labelled with a "T". Figure 6 represent the state part of the description of the OT "Document". The fact F1 is a type-attribute for the OT "Document". It specifies the "maximum size" of any document object. The graphical constraints specified for F1 state that there is only one maximum size and this value must be known even if there are no document objects. The objects of "Size", "Title" and "Doc_Number" are encapsulated in Document objects; they are called local OTs because no other object has access to a Document object's Size, Title or Doc-Number object. In addition to local object attributes, an object may also have attributes (or type-attributes) which refer to global OTs. Global OTs are defined independently of this OT definition. To distinguish a global OT from a local OT, we put a dotted square around the global OT. In the example of figure 6, the OT "Document" has an attribute relation with the global OT "Keyword".
The difference between an object of a local OT and a global OT is that a local object can only be accessed by the objects with which it is involved in an attribute relation while a global object can be accessed (through its methods) by other objects as well. For a global object only the attribute relation is hidden, the objects themselves are not hidden. For instance in the example of figure 6, it is not possible for some object to access directly the Keyword objects contained in a given Document object because this link is hidden in the Document object. However, individual Keyword objects are accessible by other objects.

In addition to the graphical constraints, text constraints to restrict the states of an OT may be specified as well. Such a text constraint may access local objects directly (or through local methods), but properties hidden in the definition of global OTs should be accessed through proper methods. An example of a simple text constraint for the OT "Document" would be

\[
\text{ByteSize of Size of a Document is } \leq \text{ByteSize of Size max of Document}
\]

### 3.3 Methods.

In addition to attribute relations and constraints the definition of an OT also includes a description of the methods of the objects of this type. The definition of a method may be given using some OO-language. As for the constraint language, methods may directly access the attribute relations and the local objects of the OT under description, all global objects should be accessed through the proper methods. A description of this language will be given elsewhere. An example of such a language can be found in e.g. [33, 11].
A method will be graphically represented by a rectangle. A fat arrow points to the OT of which it is a method. The OTs of the in- and return-objects are specified by connecting the method box and those OTs by in- resp. out-going arrows. The arrows may be labelled with a name to distinguish two different in- or return-objects of the same OT. To make the difference between an object-method and a type-method, we put a dot inside the OT for an object-method. Figure 7 is an example of the graphical representation of the object-method "Info" for the OT "Document" Figure 8 is an example of a type-method for the same OT.

3.4. Fact Types.

In most OO-system, instance variables are the only means for relating objects. We have chosen to support also explicit relations. Explicit relations are not encapsulated in some object. We have found that during
information analysis the explicit representation of relationships between objects may be of great value. Very often there is no conceptual reason for considering one object attribute of the other object or vice versa. Introducing a new object to represent the relationship between the objects is not always a desirable conceptual solution. Representing the relationship in both objects is a kind of redundancy which we want to avoid as much as possible during conceptual modelling. In our opinion, explicit relations will also meet the so called "ravioli code" problem [30]. "Ravioli code" is the OO-version of "spaghetti code", it refers to lots of tiny well structured objects that are easy to understand in isolation, but whose interactions are nearly impossible to decipher.

Other OO-approaches which also support explicit relations are e.g. [5] and [27].

As in regular NIAM, we use fact types to model explicit relations.

Figure 9 shows an example of an explicit relation between the OTs "Document" and "Folder"; a Folder object may contain several Document objects and a Document object may be placed in different Folder objects.

3.5. Inheritance.

A subtype inherits all properties of its supertypes. Contrary to most OO-systems, we only allow a limited form of overriding of the inherited properties. Along the subtype hierarchy, inherited attribute relations and constraints can only become more specific. For attribute relations inherited by a subtype this means that additional constraints may be specified which do not apply to the supertype. Also the OT of the attribute may become more specific; i.e. only a subtype of the original OT may be used.

In this way, at least for the structural part, we avoid problems with multiple inheritance. Multiple inheritance raises problems if an OT inherits the same property of two (or more) supertypes and the property is differently defined for each supertype. Since we only allow specialization, the subtype inherits the specializations given for each supertype. This will not cause any problems; in the worst case no object will ever satisfy the specialization. However, such a situation can be detected by a CASE (Computer Aided Software Engineering) tool. For an example of such a tool see [10, 12].
An example of inheritance of attribute relations is given in figure 10. In the graphical representation, the inherited attribute relation are shadowed. The additional constraints and the more specific OT are drawn in full lines.

![Figure 10: Example of inheritance and specialization of attribute relations.](image)

Figure 10 shows an example; a Sprite_Document is defined as a subtype of a Document (see figure 6). They may only include Keywords which are registrated as Sprite_Keyword (OT-specialization of F5). Each Sprite_Document must refer to at least one Sprite_Keyword (extra constraint for F5).

Another example is given in figure 11. A Template_Document is a subtype of Document (see figure 6). A Template_Document includes a Style_Definition (extra attribute) but has no Keywords (extra constraint for F5). The latter is graphically represented by putting a vertical bar on the role connecting the OT "Keyword" to the OT "Template_Document".

![Figure 11: Example of subtype definition.](image)

The implementation of the methods can be overwritten completely. So for methods we have to deal with the multiple inheritance problem. For a possible solution we refer to the literature [6, 7].
3.6. Type constructors.

It is possible to see type constructors such as collection, bag, set, tree, list,... as objects (see figure 12).

![figure 12 : pre-defined type constructors.](image)

The definition of these OTs as pre-defined OTs make it possible for the user to define UoD-specific subtypes of these OTs which inherit all built-in methods such as e.g. first/next for "List". Usually, UoD-specific subtypes of the OTs of the "Collection" family-tree (such as e.g. "Keyword_List", "Document_Set") only restrict the elements of these collection to a certain UoD-specific OT. If this is the case, we make the convention that the OT is not represented as a subtype of "Collection" or one of its subtypes, but the name of the OT is composed of the name of the chosen type constructor and the name of the OT of the elements, e.g. "List_of_Keyword" is a List Collection of Keyword objects.

Of course the user is free to define new UoD independent subtypes of "Collection" or other type constructors (like e.g. bitmaps) in order to accommodate his needs.

3.7. Meta Object Types.

Ots, fact types, constraints, methods, etc. may also be considered as objects of some (meta-)OTs. These meta-OTs together form the meta-schema which describe the OO-BR model.

We can consider the meta-schema as an integral part of the model, deriving a self-describing system. This has several advantages, for a detailed discussion see e.g. [23]. One of the advantages is that OTs may then be used
as in- and return-objects for methods. In this way objects may be asked for their type(s) (OTs), OTs may be asked for their subtypes, methods, attribute relations and so on.

6. Conclusions.

In the previous section we have presented an object oriented version of the well known Binary Relationship Model (also known as NIAM). The new model has been built from scratch following most of the principles of the conventional BRM in combination with OO-principles. The differences and similarities with the conventional BR model can be summarized as follows.

In contrast with the conventional BRM, the OO-BRM supports:
(1) encapsulation of properties,
(2) specialization of inherited properties,
(3) specification of (abstract) data types for lexical object types,
(4) definition of new abstract data types,
(5) type constructors as object types,
(6) definition of new type constructors,
(7) specification of the behaviour of objects as an integrated part.

The following BRM characteristics still apply for the OO-BRM:
(1) same graphical representation,
(2) same support for constraint specification (see also below),
(3) distinction between lexical and non-lexical object types,
(4) explicit relationship between object types,
(5) same subtype mechanism.

In principle, for supporting the specification of constraints the OO-BRM has the same capabilities as the conventional BRM. However because of the encapsulation principles of the OO-BRM, constraints which involve properties of multiple object types have to be expressed differently. A constraint defined in the context of one object type can only access the properties of another object type through the methods of this object type. As a consequence, it will be less obvious which are all the properties involved in a certain constraint. On the other hand, the definition of OTs may be changed without affecting the constraints attached to other OTs. This is not always the case for the conventional BRM.

Related work can be divided into two groups. To the first group belong the newly defined object models e.g. [15, 33, 37, 38, 39]. These models are fully object oriented but have the disadvantage to be new and not be related with a well-tried and widely accepted model. Therefore it will be hard to introduce them in the sort term into big companies.

The second group contains work that extends the conventional models (mainly the E-R model) with OO-features such as complex objects, subtypes, generalizations and new data types e.g. [18, 35]. Although these extensions
are valuable contributions to the models, they do not turn the model into a real object oriented model. We would rather call them object directed.

We argue that our model is truly object-oriented. Very important in achieving this goal was the introduction of a single subtype hierarchy, including type constructors and abstract data types, of which each user-defined object type is implicitly a subtype.

We believe that the proposed OO-BR model is a valuable contribution in order to reduce the gap between conventional conceptual models and the promising OO-DBMSs. In conventional environments the BR model has proven to be successful. By turning it into a truly OO model without actually affecting its main characteristics, it will also be possible to benefit from the advantages offered by the OO-approach when it is used in a non-conventional environment which demands for an OO solution.
Bibliography and References.


[12] De Troyer O., Meersman R., Verlinden P., "RIDL* on the CRIS Case: A Workbench for NIAM". In "Computerized Assistance During the


[37] Su S.Y.W. "An Object-Oriented Semantic Association Model (OSAM*)". In AI in Industrial Engineering and Manufacturing:

