Classifying the Semantics of Relationships in Conceptual Modeling by Categorization of Roles

Maria Bergholtz, Paul Johannesson

Department of Computer and System Sciences
Stockholm University and the Royal Institute of Technology
Electrum 230, S-164 40 Kista, Sweden
Email: {maria, pajo}@dsv.su.se

Abstract. In many information systems applications, it is important to identify similarities and differences between conceptual schemas. In order to support designers in this activity, ontologies can be utilised. Experience indicates that it is particularly difficult to compare relationships in conceptual models. Therefore, ontologies should be able to provide adequate support for this task. The contribution of this paper is an ontology for classifying and analysing relationships in conceptual modelling. The ontology is based on an analysis of actors and roles in conceptual modelling. Three requirements on an ontology for classifying relationships are introduced: high granularity, ease of application and explicit representation of similarities.

1 Introduction

In recent years, the use of ontologies has been proposed as a means for supporting different activities in conceptual modelling such as schema concept analysis and schema integration. An ontology is commonly defined as an explicit specification of a conceptualisation, where a conceptualisation consists of fobjects, concepts and other entities that are assumed to exist in some area of interest and the relationships that hold among themî, |Gr95|. A problem in representing and classifying relationships in ontologies is the fact that they are represented in conceptual models in many different ways, e.g. as binary or n-ary relationships or as entity types in their own right. The purpose of this paper is to propose an ontology for classifying and analysing relationships in conceptual modelling. We believe that a classification of relationships should satisfy the following three requirements: *High granularity*. The classification shall distinguish between different relationships and not put dissimilar relationships into the same category. *Ease of application*. It shall be easy to classify the relationships in a schema according to the ontology. *Explicit representation of similarities*. The ontology shall classify similar types of relationships in such a way that their similarity is explicitly represented.

2 Related research

A comprehensive work on the classification of relationships in conceptual modelling is that of Purao, Storey, and Ullrich, |UPS00|. Our approach is in many respects similar to

theirs, as we also propose an architecture partially built on the primitives in [UPS00]. The main distinguishing features of our approach is that we strive to achieve a higher granularity of the classification, and that we utilise speech act theory as a theoretical basis for analysing the roles in a relationship. Speech act theory [Se79] has been used for analysing processes as well as for clarifying the creation of obligations between actors, [JW99]. Some of this work has resulted in complete languages for communication in organisational settings [DW95]. In contrast, our use of speech act theory is light-weight as we only intend to provide a classification of relationships. A relationship between two or more entities can be looked upon from the perspective of each of these entities. We will adopt the UML notion and call these views the *roles* of the relationship. Furthermore, the roles often mirrors the actions carried out via the relationship. When this is the case, the involved entities are called *actors*. We advocate that the classification of a relationship is the aggregation of the classifications of all the (actor-) roles participating in a relationship.

3 An Ontology for classifying relationships

We define an ontology as a 6-tuple <COMMON, SA, ORG, ACTIVITY, RoleLabel, RelCat>. The first four components in the ontology refer to basic categories of relationships: *common relationships, speech acts, organizational relationships,* and *activities.*

i. COMMON RELATIONSHIPS
The first category consists of common relationships that correspond to the data-abstractions [GS99] common in conceptual modelling: *is-a* (generalization/specialization), *instance-of* (instantiation), *member-of, part-of* (aggregation) and *power type [Od94].* The relationships in this group are domain independent as well as static, i. e. they are not (or rarely) subject to change and accordingly not influenced by activities performed by actors.
ii. SPEECH ACTS
The relationships in the second category correspond to the speech acts of speech act theory as developed by John Searle [Se79]. Searle proposes a taxonomy of speech acts consisting of five classes, also called illocutionary points: assertives, comissives, directives, declaratives and expressives. From a business perspective, the three middle ones are the most frequent. Speech acts can also be viewed with respect to their so called *illocutionary force.* An illocutionary force can be thought of as an illocutionary point plus various qualifiers. [DW95] distinguishes between the qualifiers *power, authority and charity* as a grounding for a speech act. The example of Fig. 1 illustrates how these qualifiers are employed in the classification of relationships.
iii. ORGANIZATIONAL RELATIONSHIPS
The third group of relationships consists of organisational relationships. An organisational relationship is a relationship between an actor and another object that can exist only within an organisational context. A set of organizational relationships has to be defined for each specific domain. An example of such a set is the following, appropriate for a business context (these relationships are similar to a subset of those in group 1 of [UPS00]) : *{ëis-owner-ofí, ëis-owner-of-moneyí ëis-in-control-ofí, ëis-*

assigned-toí, ëis-subjected-toí,i is-creator-ofí, ëhas attitude towardsí, ëfollows-or-precedesí, ëis-destroyer-ofí, ëis-subjected-toí}.

iv. ACTIVITIES

Relationships of group four are activities, generally with a predetermined extension in time and often concluded by a speech act. This group of relationships are, compared to the other three groups, rarely present in a conceptual model. One case where this kind of relationship occurs is when it is meaningful to preserve information of the duration of the activity itself. Similar to the relationships of group III the members of this group may vary from domain to domain.

v. GRAMMAR

The last two components of the ontology, i.e. the classifications of roles (RoleLabel) and relationships (RelCat), are defined by the grammar defined below. Recall that COMMON, ORG and ACTIVITY denote the relationship categories ëCommon relationshipsí, ëOrganizational relationshipsí and ëActivitiesí, respectively, as defined above. Let ëDeclí, ëComí, ëDirí denote speech acts. An authorized directive is denoted Dir_a whereas Dir_p and Dir_c stands for directives based on power and charity, respectively.

I. Dir ← Dir_p Dir_a | Dir_c
II. Decl ← $Decl_p$ $Decl_a$ $Decl_c$
III. ZeroLabel ← ε
IV. SALabel ← Decl(ORG)
V. SALabel ← Dir(SALabel)
VI. SALabel ← Com(SALabel)
VII. RoleLabel ← ZeroLabel SALabel | RoleLabel¯
VIII. RelCat ← COMMON| ORG ACTIVITY| ACTIVITYDecl(ORG) | RoleLabel¯

We observe that the *declarative* speech act is the only one that can result in state changes of the relationships in ORG, see rule IV. Consequently a request to change status, for instance to request to *rent* something, should be thought of as a directive to change the status of the relationship verb-phrase ëis-in-control-ofí, i. e. a dir(decl · is-in-control-of).

4 Application of the ontology in classifying relationships

Figure 1 illustrates five relationship examples, each of which has been classified through the labeling of the two actor roles involved in each relationship. We will assume that the set of organisational relationships is confined to {ëis-in-control-ofí, ëis-owner-ofí, íis-owner-of-moneyí} as discussed in section 3, e.g. we extend the previously defined grammar grammar by the following rule: 0: ORG ← ëin-control-ofí ëis-owner-ofí | ëis-owner-of-moneyí.

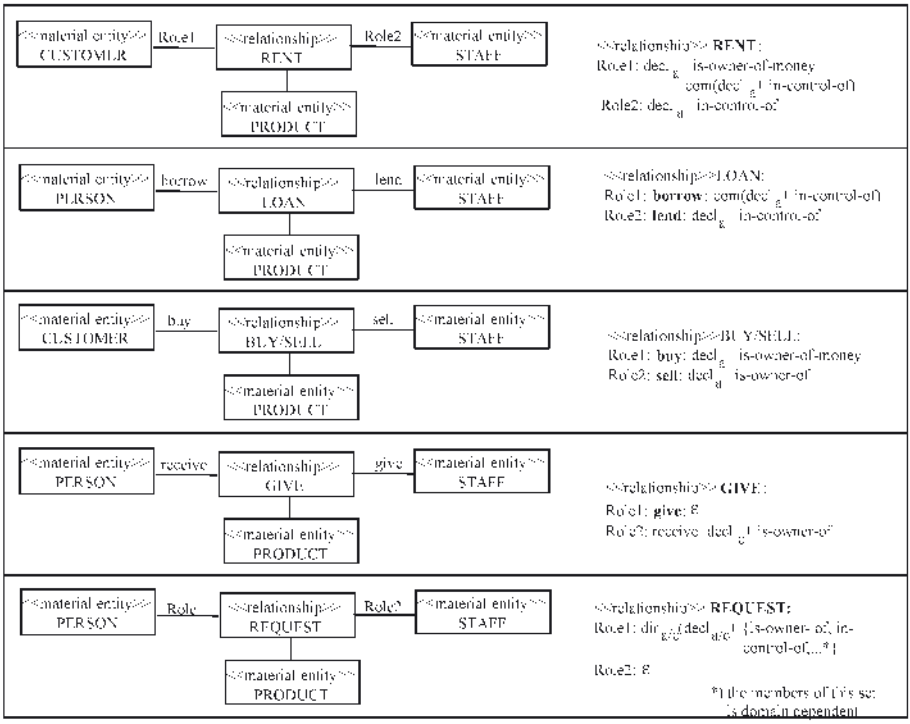

Figure 1. Five relationships with two actors

RENT and LOAN both deal with the transfer of control between actors. The relevant member of the set of organisational relationships is thus čis-in-control-ofî in both cases. The illocutionary force is also the same, *authority*, for the speech acts involved. The significant difference between a LOAN relationship and RENT relationship is that the value transfer in the LOAN, in addition to a transfer of a product, also include transfer of money, i. e. a declarative of ownership of money on behalf of the customer to the staff authorizing the rental process.

Inspecting the BUY/SELL- and GIVE-relationships reveals some differences, the most significant being the difference in illocutionary point on the classification of Role2, (ìsellî versus ìgiveî). The roles of ìsellî and ìgiveî are both classified as declaratives of ownership, i. e. the actor performing Role2 is declaring a transfer of ownership to the actor of Role1. The illocutionary force of the declarative of the ownership of the ìsellî -role is *authority* whereas it is *charity* with regard to the ìgiveî -role.

REQUEST, finally is probably the most generic relationship of the five. A request can refer to all of the four relationships (LOAN, RENT, BUY/SELL, GIVE). The choice of members of the set of organisational relationships must be established reflecting the domain at hand.

5 Conclusions

In Section 1, three requirements on ontologies for classifying relationships were introduced. We will now consider how well the proposed ontology satisfies these requirements.

High granularity: Although the ontology is based on a small number of primitives, it is our experience from a number of preliminary case studies that it is able to distinguish even among fine nuances of relationships. One reason for this capability is the possibility to form nested speech acts (as manifested in the recursive definition in the grammar of Section 3). This enables the build-up of a large and varied set of roles that actors can play in a relationship. *Ease of application*: It might be argued that our ontology builds on a more complex foundation than other approaches, such as [UPS00], and that it, therefore, will be more difficult to apply. However, we believe that an adequate tool support, in the form of a question and answering system, can substantially alleviate the task of applying the ontology to specific conceptual schemas. *Explicit representation of similarities*: A strength of the proposed ontology is that it clearly represents the similarities and differences between relationships. This is achieved by the possibility to combine the primitives of the ontology (as manifested in the composition rules in the grammar of Section 3).

References

[DW95] T. Dignum and H. Weigand, "Modelling Communication between Cooperative Systems", in proceedings of Conference of *Computer Advanced Information System Engineering (CAiSE)* , pp. 1995.

[GS99] Goldstein, Robert, C., Storey, Veda, C.: "Data abstractions: Why and How?", *Data and Knowledge Engineering 29 (1999), pp. 293-311.*

[Gr95] T. Gruber, "Toward principles for the design of ontologies used for knowledge sharing", *International Journal of Human Computer Studies*, No 43, pp. 907 - 928. 1995.

[JW99] P. Johannesson and P. Wohed, "The Deontic Pattern - a Framework for Domain Analysis in Information Systems Design", *Data & Knowledge Engineering 31*, pp. 135-153, 1999.

[Od94] Martin, J., Odell, J.: *Object-Oriented Methods. A Foundation,* Prentice Hall 1994

[Se79] Searle, J.R., "A taxonomy of Illocutionary Speech Acts", in Expressions and Meaning: Studies in the Theory of Speech Acts, pp. 1-29, Cambridge Press, Binghampton, New York 1979

[UPS00] Ullrich, H., Purao, S., and Storey, V.C., "An Ontology for Classifying the Semantics of Relationships in Database Design". *Proceedings of NLDB2000*. Versailles 28-30 June 2000